Biophysical Studies of Protein Assemblies

Basile Isidore Martin Wicky

St Catharine's College

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

September 2018

À ma famille,

Maya, Didier, Marie et Ulysse

À Viviane

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the Preface or specified in the text. It is not substantially the same as any that I have submitted for a degree or diploma or other qualification at the University of Cambridge or any other University, and no part has already been or is concurrently being submitted for any degree, diploma or other qualification. It does not exceed the 60,000-word limit, as imposed by the Physics and Chemistry Degree Committee.

Basile I. M. Wicky

Cambridge, September 2018

Summary

Proteins are synthesised as linear polymeric chains. The subtle energetic interplay of interatomic interactions results in chain folding, through which proteins may acquire defined structures. This spatial organisation is encoded by the protein sequence itself; the so-called thermodynamic hypothesis formulated by Anfinsen in 1961. A defined structure is often considered a pre-requisite to protein function, but widespread existence of intrinsically disordered proteins (IDPs) has prompted a re-evaluation of the ways biological function may be encoded into polypeptide chains. Furthermore, proteins often exist as part of multi-component entities, where regulation of assembly is integral to their properties. The interplay between disorder, oligomerisation and function is the focus of this thesis.

Some IDPs fold conditionally upon interacting with a partner protein; a process known as coupled folding and binding. What are the biophysical advantages and consequences of disorder in the context of these interactions? A common feature of IDPs is their sequence composition bias, with charged residues being often over-represented. It is therefore tempting to speculate that electrostatic interactions may play a major role in coupled folding and binding reactions. Surprisingly, the opposite was found to be true. Charge-charge interactions only contributed about an order of magnitude to the association rate constants of two contrasting model systems. The lack of pre-formed binding interfaces—a consequence of disorder—might preclude electrostatic acceleration from complementary patches.

By looking at the role of the sequence, many studies have taken a protein-centric approach to understanding disorder. Yet there is paucity of data about the effect of extrinsic factors on interactions involving disordered partners. Investigating the role of co-solutes, it was discovered that the kinetic and thermodynamic profiles of coupled folding and binding reactions were sensitive to ion-types. This effect followed the Hofmeister series, and occurred at physiological concentrations of salt. The sensitivity of coupled folding and binding reactions—a consequence of the lack of stability of IDPs—might be advantageous.

Given the role of ions in biology, this 'biophysical sensing' could be a mechanism of physiological relevance, allowing modulation of protein-protein interactions involving disordered partners in response to changes in their environments.

In cells, signalling networks are often multi-layered, and involve competing protein-protein interactions. The interplay between the biophysical characteristics of the components, and the behaviour of the network were investigated in a model tripartite system composed of folded and disordered proteins. The BCL-2 family regulates the intrinsic pathway of apoptosis through control of mitochondrial outer-membrane permeabilisation; a result of BAK and BAX oligomerisation. Through a shared homology motif (termed BH3), the subtle balance of their interactions determines cellular fate at the molecular level. Characterisation of the model under simple biochemical conditions revealed large differences in affinities among binary interactions; the consequence of the lifetime of the complexes, not their speed of association. A membrane-like environment, re-created using detergents, allows the oligomerisation of BAK and BAX *in vitro*. Furthermore, investigation of the tripartite system under detergent conditions showed that regulation of the network was the result of competing hetero- and homo-oligomerisation events. Relationships to their biophysical properties were gained by probing their energy landscapes using protein folding techniques. The connection between the biophysical properties of the components of the network and their interactions provides a molecular explanation for the regulation of apoptosis.

This thesis offers insights into the ways structured assemblies and environmentally responsive disorder elements may encode functions into proteins.

Acknowledgments

First of all, I would like to thank Jane. A few years back, a naive chemist emailed her. He was hoping to undertake a PhD in molecular biophysics because he thought that large wiggly molecules that attained three-dimensional structures autonomously were fascinating. A few years later, the metamorphosis is (hopefully!) complete. I did get the chance to learn about protein folding, and along the way I also discovered that some proteins lack a structure, that some do when they bind other molecules, that others oligomerise, and that the whole conundrum depends on solution conditions. I have learnt so much more than I could have imagined, and this would not have been possible without coming to Cambridge. So thank you Jane for accepting me as part of your group, and for giving me the opportunity to explore some of my crazier ideas.

My time in lab 290 would not have been the same without all the people that were there. A special mention goes to my deskmate for three years, Michael. You were always up for a bit of scientific jousting, cheering me up when experiments failed, or putting me back on track when I became trapped in the literature limbo. The many debates we had over the years have profoundly helped my scientific learning, and you were thoroughly missed during the last few months of my PhD. Another special mention goes to Liza (my own personal genie) and Carol ('prata ou proteinas?'). Your friendships and *joie de vivre* were so important to the atmosphere of the IDP bay! Tristan, Quenton and Jeff, you were brilliant! I wished we could have spent more time together! Special thanks to Tristan, for all the help with BAX, and our Bee Gees singing times by the AKTAs. To Andrew, a labmate-turned-landlord, thank you for offering me a place in your house during the last year of my PhD. Living with people you know is great, living with a friend is even better! To all the other Clarke group members, Dominika (number 1!), Adrian, Lee, Pavel, Charlie, Sophie, and Katie, thank you for your help and friendships over the years. My gratitude also goes to Sarah, who dedicated so much of her time to train me when I first joined the lab, and whose rigorous

scientific approach is an inspiration. Finally, the list of Clarke group members would not be complete with mentioning Annette; you were there until the very end, and provided much needed company during the arduous months of writing. Thank you for bearing with me!

I could not have done my PhD without the many people that have helped along the way; thank you to Jo and Helen, Gary and Kevin, Ryan and Paul, Anita, Asha, and Peter Sharratt. Thank you as well to Katherine Stott, who has been so kind to teach me about AUC. To my scientific collaborators, Prof. Andrew Wilson, Prof. Thomas Edwards, and Phil Rowell, who did the EM work. To Prof. Carol Robinson, for accepting me as a fleeting member of her lab, and giving me the opportunity to do some fancy native MS. And of course to Dr Kallol Gupta, who did not shy away from recording spectra until the wee hours of the morning when exciting results were coming our way.

A good life is one shared with friends, and I have been fortunate enough to meet many at Catz. Thank you for the amazing moments we spent over the past four year! In particular, to all the Slodgers: JMG, Josh, Donna, Phil, Alex, Sarah, Rachel, Geoff, Peter, Ayhan, who were always up for a late night debate on pretty much anything. To all the friends I made at SCCBC: Chris, Will, Casper, James, Jack, Rory, Michael, Stuart, Adam, Toby, Nick, Michael, and so many more! We did not get the results we wanted, but I would do it all again in a heartbeat. The many hours we spent together on the river have kept me (somewhat) sane through this PhD journey.

Ces remerciement ne serait pas complet sans mentionner ma famille; Papa et Maman, Ulysse et Marie, vous avez toujours été là pour moi, vous avez cru en moi, et vous avez su me montrer que ce qui est important dans la vie, c'est de faire ce qu'on aime. Cette thèse est un symbole de cette philosophie de vie, et c'est pour ça que je vous la dédie.

Und endlich, zu dir Viviane; ohne dir hättes ich nicht gschaft. Du bisch immer da xi für mir, und du hash mir immer unterstützt. Merci und viel meh, you deserve all the Shampoo!

Abbreviations

A1	Bcl-2-related protein A1 (UniProt:Q07440)
AEW	average emission wavelength
a.u.	arbitrary units
α0α1	first two domains of erythrocytic α-spectrin (UniProt:P02549)
β16β17	last two domains of erythrocytic β-spectrin (UniProt:P11277)
BAK	Bcl-2 homologous antagonist/killer [protein] (UniProt:Q16611)
BAX	apoptosis regulator BAX [protein] (UniProt:Q07812)
bp	base pair
BCL-X_L	Bcl-2-like protein 1, isoform 1 (UniProt:Q64373-1)
BCL-2	B-cell lymphoma 2
BH3	Bcl-2 homology domain 3
BID	BH3-interacting domain death agonist [protein] (mouse, UniProt:P70444; human, UniProt:P55957)
BIM	Bcl-2-like protein 11 [protein] (UniProt:O43521)
BS3	bis(sulfosuccinimidyl)suberate
CD	circular dichroism [spectroscopy]
CMC	critical micelle concentration
CuPhe	Cu^{II}(1,10-phenanthroline)$_3$
C8E4	tetraoxyethylene monooctyl ether
C12E8	octaoxyethylene monododecyl ether
D	denatured state
[den]	concentration of denaturant
DNA	desoxyribonucleic acid
dsDNA	double-stranded DNA
DTT	dithiothreitol

ΔC_{p}	heat capacity change at constant pressure
ΔG	Gibbs free energy change
ΔH	enthalpy change
ΔS	entropy change
E_a	activation energy
E.coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
FITC	fluorescein isothiocyanate
FPLC	fast protein liquid chromatography
GB1	B1 domain of protein G
GdmCl	guanidinium chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
I	ionic strength
IDP	intrinsically disordered protein
IDR	intrinscially disordered region
IPTG	isopropyl β-D-1-thiogalactopyranoside
K_{d}	equilibrium dissociation constant
k_{obs}	observed rate constant (expressed in s^{-1})
k_{on}	bimolecular association rate constant (expressed in M^{-1} s^{-1})
k_{off}	unimolecular dissociation rate constant (expressed in s^{-1})
LB	lysogeny broth
MCL-1	induced myeloid leukemia cell differentiation [protein] (mouse, UniProt:P97287; human, UniProt:Q07820)
MOM	mitochondrial outer membrane
MOMP	mitochondrial outer membrane permeabilisation
MOPS	3-(*N*-morpholino)propanesulfonic acid
MRE	mean residue ellipticity (expressed in deg cm^2 s^{-1})
MS	mass spectrometry
MWCO	molecular weight cut-off
N	native state

Ni-NTA	nickel nitrilotriacetic acid [resin]
NMR	nuclear magnetic resonance [spectroscopy]
OD_{600}	optical density (absorbance) at 600 nm
OGP	octyl β-D-glucopyranoside
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PES	polyethersulfone
pI	isoelectric point
PPI	protein-protein interaction
PS20	polyoxyethylene (20) sorbitan monolaurate (TWEEN®)
PTM	post-translational modification
PUMA	p53 upregulated modulator of apoptosis [protein] (mouse, UniProt:Q99ML1; human, UniProt:Q9BXH1)
PVDF	polyvinylidene fluoride
RMSD	root-mean-square deviation (expressed in Å)
rpm	revolutions per minutes
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
$t_{1/2}$	half-life
τ	mean lifetime
TAMRA	5-carboxytetramethylrhodamine
T_{m}	melting temperature midpoint
Tris	tris(hydroxymethyl)aminomethane
TS, or ‡	transition state
UV-vis	ultraviolet-visible [spectroscopy]
2×TY	double yeast extract and tryptone

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the public eye, DNA symbolises the abstract molecular concepts of biological sciences. Since the seminal work of Crick, Franklin, Watson and Wilkins on the structure of the 'molecule of heredity' (Watson & Crick, 1953), the double-helix has fascinated by its elegant simplicity, and soothing symmetry. But if DNA contains the blueprints of life, it really is the products of its molecular deciphering—proteins—that constitute the fabric of biological processes. What Charles Tanford and Jacqueline Reynolds called 'nature's robots' (Tanford & Reynolds, 2001), constitute the molecular cogs at the heart of the machinery of life. The palette of functions attributed to proteins encompasses virtually every aspects of an organism's inner workings; enzymes capable of accelerating chemical reactions by orders of magnitude (Wolfenden & Snider, 2001), structural architectures conferring mechanical properties to cells—*e.g.* the giant muscle protein titin (Wang *et al.*, 1979, Trinick *et al.*, 1984)—nanomachines at the centre of metabolic processes—*e.g.* ATPases (Abrahams *et al.*, 1994), antibodies providing the basis for acquired immunity (Wu & Kabat, 1970, Al-Lazikani *et al.*, 1997), phase-separating proteins capable of generating membrane-less cellular compartments (Brangwynne *et al.*, 2009, Boeynaems *et al.*, 2018), and the nodes at the heart of signalling networks (Pawson & Nash, 2000). These are just a few of the functional attributes of proteins, and many more could be listed. But what is the molecular basis for protein function? And how do species composed of a common set of twenty building blocks give rise to so many different structural, functional, and biological outcomes?

1.1 The origin of function in proteins

Proteins are produced as linear polypeptide chains, formed by covalent linkage of amino acid units. The precise sequential arrangement of these chemical building blocks is what defines proteins and their roles. However, the link between sequence and function, is not a trivial one. Even the same sequence expressed in reverse does not have the same properties (Lacroix *et al.*, 1998b, English *et al.*, 2018). It is often assumed that proteins owe their functions to their structures, which is examplified by the requirement for enzymes to have well-defined active sites that are complementary to the transition state of the reactions they catalyse (Pauling, 1946). Proteins acquire their structures through folding, which is briefly reviewed below. However, not every sequence does fold, and the case of intrinsically disorder proteins is described later (*vide infra*).

1.1.1 Protein folding

The physical origin of protein folding is rooted in the subtle energetic interplay between *intra*- and *inter*-molecular interactions. The former are between atoms and functional groups present on the polypetide chain, while the latter are between the protein and the molecules present in the surroundings; the most abundant of which is water. The nature of these interactions varies, and so do their energetic contributions to the stability of proteins (Dill, 1990). These forces include; van der Waals interactions between non-charged atoms, hydrogen-bonding between (mostly) amide and carbonyl groups of the protein main chain, and electrostatic interactions between charged residues. However, one of the most important contributions to the folding energy of proteins is the 'hydrophobic effect' (for a recent review, see Southall *et al.* (2002)). This phenomenon is the result of solvent-solvent interactions prevailing over solvent-protein interactions; folding of the protein chain releases water molecules, which contributes favourably to the entropy of the reaction (Privalov & Makhatadze, 1993, Makhatadze & Privalov, 1996). In contrast, some molecular events contribute unfavourably, such as the loss of conformational entropy of the chain upon it gaining structure (for a recent quantification, see Baxa *et al.* (2014)).

This subtle balance between favourable and unfavourable energetic contributions results in proteins being only marginally stable; in the order of 5–15 kcal/mol (Pace, 1975, Privalov, 1979, Fersht, 1999). This value barely corresponds to the strength of a few of hydrogen bonds in water (van der Spoel *et al.*, 2006), a very small number considering the thousands

of atoms composing proteins, and the even larger number of possible interactions. It is important to point out that the marginal stability of natural proteins is *not* an inherent property of polypeptidic chains *per se*. Indeed, proteins with folding free energies in the order of −60 kcal/mol have been designed (Huang *et al.*, 2014). Instead, it appears that typical protein stabilities are 'good enough' for biological systems, and have not been under any evolutionary pressure to become more stable (Williams *et al.*, 2006).

The entirety of the information necessary for proteins to fold—at least for monomeric systems—is encoded in the sequence itself. This was demonstrated by Anfinsen and colleagues, who showed that ribonuclease A could be unfolded reversibly *in vitro*, excluding the need for co-factors for proteins to attain their structures (Anfinsen *et al.*, 1961). This discovery lead to the formulation of the thermodynamic hypothesis, which stipulates that the native state of a protein is its thermodynamic minimum (Anfinsen, 1973). This is not always true. Indeed, some protein structures are kinetically trapped into metastable states (*e.g.* serpins (Wang *et al.*, 1996)). Moreover, it has been suggested that the amyloid state might be the true thermodynamic minimum of many proteins (Gazit, 2002, Dobson, 2003, Baldwin *et al.*, 2011). Nevertheless, Anfinsen's hypothesis remains valid for many of the systems studied to date.

Protein are very large molecules, and the number of conformations accessible to a polypetide chain raises questions about how the native state may be reached; the so-called 'folding problem' (Dill *et al.*, 2008). It was realised early on that if proteins chains were to find their energetic minima by exhaustive sampling of all their possible conformations, proteins would take longer than the age of the universe to fold. However, they do so in much shorter timescales. This discrepancy is often referred to as the 'Levinthal paradox' (Levinthal, 1969). This apparent paradox is resolved by the fact that folding reactions occur along defined pathways. Decades of work have revealed that most proteins fold through variations of the 'nucleation-condensation' mechanism (Itzhaki *et al.*, 1995, Fersht, 1997, Daggett & Fersht, 2003). For such processes, a few critical interactions are generated at the transition state (the nucleation), followed by the formation of the remainder of the structure to reach the native state (the condensation). Thus, folding nuclei reduce the problem of sampling a large conformational space by directing the search along defined pathways. This concept has also been described in terms of funnelled energy landscapes (Onuchic & Wolynes, 2004).

In summary, protein chains fold because they have an energetic advantage to do so. The

native state acquired after the folding process is encoded on the energy landscape by the sequence itself, and the resulting structure defines the properties of the protein. Thus, a link between sequence and function can be drawn. But not all proteins fold.

1.1.2 Intrinsically disordered proteins

When Kendrew and his co-workers published the first structure of a protein in 1958 (Kendrew *et al.*, 1958), they opened the avenue to an entirely new field of science; structural biology. The past half century has seen an enormous growth in the number of systems that have had their structures solved (144,464 at the time of finishing this thesis). However, since the prevalent technique for obtaining these atomistic models remains X-ray crystallography—which inherently requires the formation of crystals—our view of proteins has been biased towards folded, globular states (Cossio *et al.*, 2010). In fact, this probably played a role in shaping the structure-function paradigm, and this notion has largely dominated the field of protein science since the advent of the structural biology era.

The more recent discovery that a large portion of the proteome is composed of (partially) unstructured proteins has challenged the view that an ordered state is a prerequisite to function (Wright & Dyson, 1999, Uversky *et al.*, 2000, Dunker *et al.*, 2001, Tompa, 2011). Despite their lack of structure, intrinsically disordered proteins (IDPs) are involved in numerous cellular tasks (Dyson & Wright, 2005, Wright & Dyson, 2015). Moreover, they are often implicated in pathological conditions (Uversky *et al.*, 2008), pointing at their importance in cellular homoeostasis. The fact that proteins lacking a structure may be functional has forced a re-evaluation of how nature may encode function into protein molecules (Wright & Dyson, 1999, Uversky & Dunker, 2010).

IDPs are characterised by a lack of uniquely-defined state; instead they populate large conformational ensembles resembling unfolded proteins (Eliezer, 2009, van der Lee *et al.*, 2014a, Chebaro *et al.*, 2015), and are best described by the concepts of polymer physics (Hofmann *et al.*, 2012). Compared to folded proteins, IDPs typically show lower sequence complexity, and amino acid composition biases; large hydrophobic residues tend to be under-represented, while prolines, glycines and charged residues are over-represented (Uversky *et al.*, 2000, Romero *et al.*, 2001, Lise & Jones, 2005, Theillet *et al.*, 2013). These sequence-based characteristics have provided the basis for bioinformatic mining of proteomes (Weathers *et al.*, 2004), which has revealed that over 30% of eukaryotic proteins

are expected to contain long disordered stretches (>30 amino acid long) (Ward *et al.*, 2004, Peng *et al.*, 2014).

But why are IDPs unfolded under physiological conditions? The sequence biases described above put intrinsically disordered proteins in a particular region of the charge-hydrophobicity phase space (Uversky *et al.*, 2000). The hydrophobic effect—so crucial to protein folding—is expected to be weak in this region, thus providing an explanation for the lack of autonomous folding of IDPs. Moreover, it has also been postulated that a very small amount of excess chain entropy compared to folded proteins ($\Delta\Delta S = 1.5$ cal mol^{-1} K^{-1} per residue) would be sufficient to promote the unfolded state under physiological conditions (Rajasekaran *et al.*, 2016). This value corresponds to only a 2-fold increase in the number of conformations available per residue. Hence, it is possible that the nature of some over-represented amino acids (*e.g.* glycines) provides enough extra degrees of freedom to offset the folding free energy of the chain; explaining the disordered nature of IDPs under physiological conditions.

The fact that intrinsic disorder appears to be more prevalent in higher organisms (Ward *et al.*, 2004, Dunker *et al.*, 2015)—usually in signalling networks and transcriptions; essential features of multicellularity—clearly points at protein disorder as having evolutionary significance. In fact, its evolutionary conservation is non-trivial (Schlessinger *et al.*, 2011), which further supports the notion of its biological importance. The advantages of IDPs have been argued, and many postulates put forward over the years (Liu & Huang, 2014). Some of the features of protein disorder in relationship to their potential biological advantages are discussed below.

It has been hypothesised that functions encoded in IDPs might be more tolerant to mutations. For folded proteins, core mutations usually have significant consequences on stability (Tokuriki *et al.*, 2007), which can result in disrupted structures. Protein unfolding is likely to prevent its function, and many diseases are indeed associated with loss-of-stability mutations (Yue *et al.*, 2005). This stability-function relationship puts pressure on important positions within proteins (Mirny & Shakhnovich, 1999), making the scaffold less amenable to changes. In contrast, the sequence space tolerant to mutations is expected to be higher for IDPs; since there is no structure to disrupt, positions outside of binding regions should not have much of an impact. This hypothesis appears to be supported by the fact that disordered regions tend to be less conserved than ordered ones (Chen *et al.*, 2006). This larger mutational tolerance could also allow higher evolutionary rates (Chen *et al.*, 2006, Brown

et al., 2011), although it has been shown that disordered regions do not always evolve faster. Furthermore, it has also been proposed that a lack of a defined structure might allow alternative splicing whilst avoiding structural complications (Romero *et al.*, 2006). This is expected to enable functional diversification while maintaining genome economy.

Because IDPs are disordered and expanded, more amino acids are accessible for post-translational modifications (PTMs), and disordered motifs have indeed been found to more heavily modified. Phosphorylation, for example, has been shown to affect the function of disordered proteins (Tompa *et al.*, 2014, Bah *et al.*, 2015, Bah & Forman-Kay, 2016, Dahal *et al.*, 2017b). Ubiquitination also appears to be more prevalent (Edwards *et al.*, 2009), which is consistent with the observed shortened cellular half-lives of disordered proteins (van der Lee *et al.*, 2014b, Fishbain *et al.*, 2015). This higher propensity to PTM might allow the activity of IDPs to be fine-tuned, which could be an advantage in signalling networks. This notion is supported by the finding that the abundance of disordered proteins inside cells is under tight control (Gsponer *et al.*, 2008, Babu *et al.*, 2011). Moreover, IDPs are often hub proteins, with a higher average number of partners than folded proteins (Haynes *et al.*, 2006). Thus, intrinsic disorder appears to confer a functional advantage for biological processes such as signalling and transcription (Shammas, 2017).

From a biophysical point of view, protein disorder has been postulated to offer a canvas for producing interactions that have high specificity but low affinity (Zhou, 2012). It also been claimed that the higher capture radius of disordered chains might increase their association speed by the so-called 'fly-casting' mechanism (Shoemaker *et al.*, 2000). However, this notion has been criticised, as the gain in capture radius is offset by the slower diffusional coefficient, which results in almost no net advantage (Huang & Liu, 2009). The disordered nature of IDPs has also been postulated to allow the affinity to be tuned through modulation of the entropic cost of binding (Flock *et al.*, 2014). Surveys of protein-protein interactions involving disordered partners revealed that, on average, such interactions are slightly weaker than PPIs involving exclusively folded proteins (Teilum *et al.*, 2015), and that the difference in affinity could be explained by shorter lifetimes of the bound complexes (Shammas *et al.*, 2012).

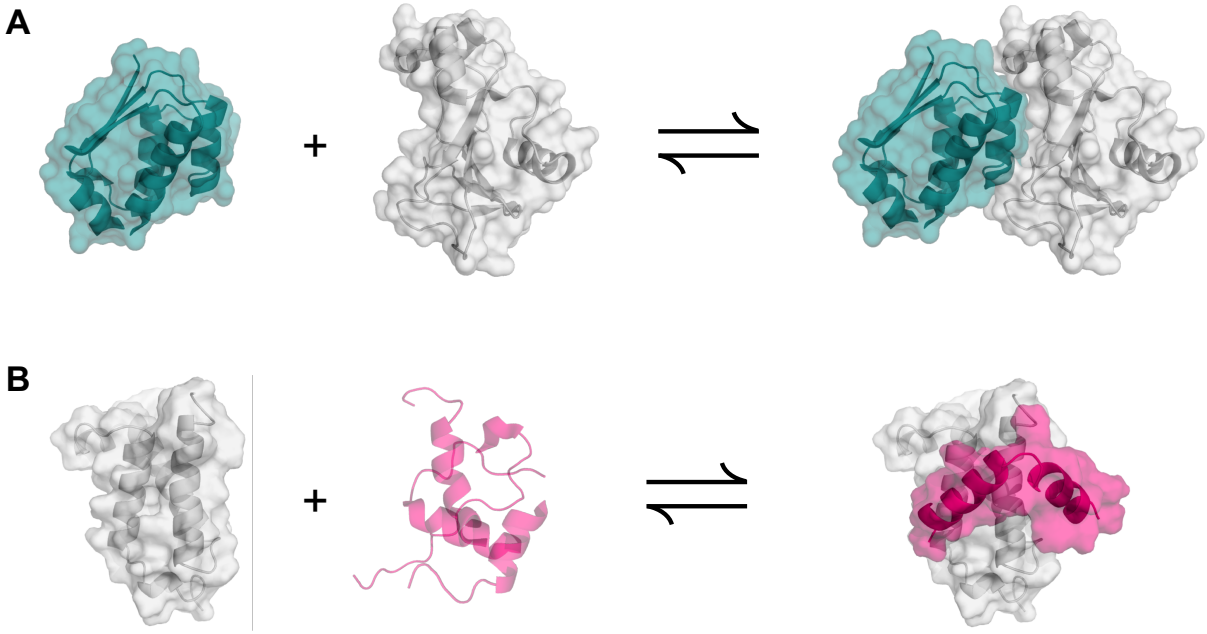

Figure 1.1 – Comparaison between 'classical' protein-protein interactions, and coupled folding and binding reactions. (**A**) Interaction between barnase and barstar, where both proteins are folded in isolation (PDB:2ZA4). (**B**) Interaction between KIX and phosphorylated KID. Here only KIX is folded in isolation, while the IDP pKID samples an ensemble of conformations prior to forming its complex (PDB:1KDX).

1.1.3 Coupled folding and binding

Intrinsically disordered proteins lack the capacity to fold autonomously. However, as alluded in the previous section, they are functional, and often interact with partner macromolecules. Many of these interactions result in the IDP gaining a defined structure in the bound state (Dyson & Wright, 2002, Wright & Dyson, 2009). These 'coupled folding and binding' reactions are in effect cases of conditional folding, where the binding of a partner protein provides the free energy necessary for the IDP to fold (Fig. 1.1). This particular class of protein-protein interactions is particularly important in signalling systems (Wright & Dyson, 2015).

The presence of a folding dimension on-pathway to binding raises an important question; what comes first—folding or binding? Indeed, there might be functional consequences depending on the mechanism sampled by these interactions. For coupled folding and binding reactions, two extreme mechanisms are possible: on one end of the spectrum the induced-fit process—every conformation of the IDP is binding-competent, and folding occurs after binding—and on the other end, conformational selection—only a subset of unbound states are binding-competent, and folding *de facto* precedes binding (Fig. 1.2). It is emphasised

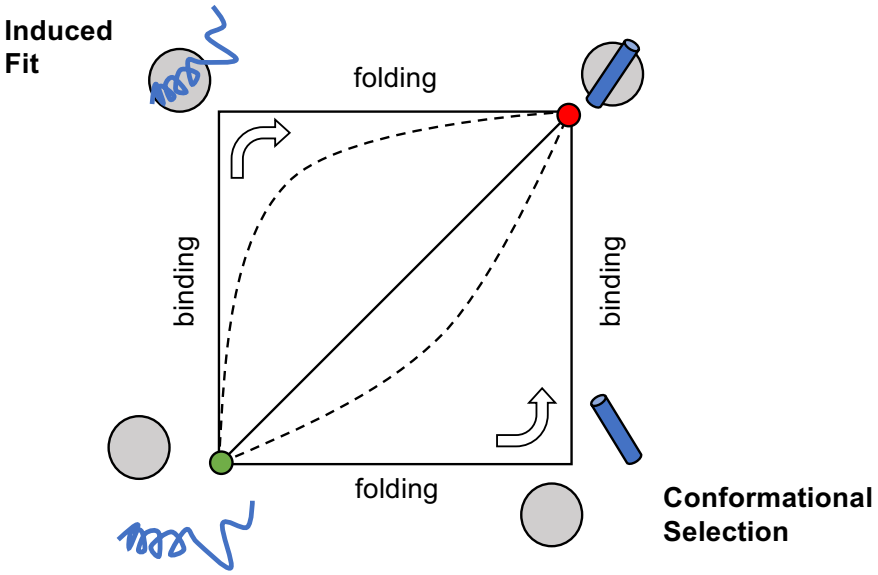

Figure 1.2 – Schematic representation of coupled folding and binding reactions. The folded partner is represented as a grey sphere, and the IDP in blue. The entire reaction from separate components (green dot) to complex (red dot) is divided into two dimensions: folding (horizontal), and binding (vertical). A pure induced fit process would proceed *via* the top left corner, while a pure conformational selection mechanism would go through the bottom right corner. The fully concerted mechanism is represented by the diagonal line, and trajectories with mechanistic elements of both are represented by dashed lines.

that these mechanisms only represent the extrema of the spectrum, and coupled folding and binding reactions with elements of both are possible (Hammes *et al.*, 2009, Greives & Zhou, 2014). Moreover, it is noted that some complexes remain unstructured in the bound state (Borgia *et al.*, 2018). Such assemblies are sometimes referred to as 'fuzzy complexes' (Tompa & Fuxreiter, 2008, Fuxreiter, 2018). Although mechanistic conclusions drawn from thermodynamic arguments can often be found in the literature—such as the inference of conformational selection mechanisms based on the presence of bound-like structures in the free state of the IDP—only kinetic analyses can properly assess mechanisms (Kiefhaber *et al.*, 2012, Gianni *et al.*, 2014, Dogan & Jemth, 2014, Gibbs & Showalter, 2015, Shammas *et al.*, 2016, Gianni *et al.*, 2016).

The nature of the transition state of coupled folding and binding reactions has been probed using Φ-value analysis—a technique initially developed for studying protein folding (Matouschek *et al.*, 1989). Over the past few years, multiple systems have been investigated, and the amount of native structure present at the rate-determining step of these reactions reported (Bachmann *et al.*, 2011, Dogan *et al.*, 2013, Giri *et al.*, 2013, Hill *et al.*, 2014, Rogers *et al.*, 2014a, Dahal *et al.*, 2017a). Most of these studies revealed transition states

with very little structure, suggesting that conformational selection is unlikely to be the dominant mechanism of coupled folding and binding reactions. This notion is supported by the finding that disrupting the residual helicity of the free IDP only marginally affected the association rate constant of the PUMA:MCL-1 model system (Rogers *et al.*, 2014b).

Given the apparent prevalence of the induced-fit process, it is tempting to speculate that the partner protein might play in important role in determining the folding mechanism of the IDP. Interestingly, the opposite was found to be true (Crabtree *et al.*, 2018). In this study, the authors demonstrated that it is the IDP that dictates its own folding pathway, and that the transition state is *not* templated by the partner protein. The interplay between partner and disordered protein was also investigated in terms of the energetics of the binding interface (Jemth *et al.*, 2014). Interestingly, a significant amount of energetic frustration was discovered, which the authors attributed to the promiscuity of the partners—making the optimisation of multiple interfaces at once difficult. It will be interesting to see if this conclusion holds for systems with fewer binding partners, and whether the level of energetic frustration is correlated with the level of promiscuity.

So far, most studies have focused on the protein components of coupled folding and binding reactions. For example, the influence of sequence on residual helicity—and its impact on coupled folding and binding reactions—has been an important focus of recent studies (Rogers *et al.*, 2014b, Crabtree *et al.*, 2017, Dahal *et al.*, 2018). These reports have revealed that changes in IDP:partner affinities are mostly the results of modulations of the lifetime of the bound state, and *not* a consequence of altered association rates. These protein-centric approaches have provided valuable insights into our understanding of the factors affecting coupled folding and binding reactions. Howevever, IDPs are also extremely sensitive to solution conditions. Indeed, it has be shown that temperature (Wuttke *et al.*, 2014), molecular crowding, (Soranno *et al.*, 2014), pH (Hofmann *et al.*, 2013), and ionic-strength (Müller-Späth *et al.*, 2010) all crucially affect the collapse, and geometrical dimensions of IDP ensembles. Yet, little is known about the effect of solution conditions on coupled folding and binding reactions. This question is the focus of the work described in Part I of this thesis.

1.1.4 Oligomerisation

Protein-protein interactions (PPIs), and the formation of protein assemblies, are central to biology. Ranging from functional oligomers (Perham, 1975, Nooren & Thornton, 2003, Marianayagam *et al.*, 2004), to debilitating amyloid fibril formation (Knowles *et al.*, 2014), the combination of multiple polypetidic chains into larger non-covalent assemblies occurs throughout the protein world. Coupled folding and binding only represents one such form of assembly. But the range, size, stoichiometries, symmetries, and function of protein oligomers encompasses much more (Fig. 1.3). In fact, a large fraction of the proteome of most organisms exists in the oligomeric state (Goodsell & Olson, 2000), and these assemblies are often integral to the function of these proteins (Griffin & Gerrard, 2012). It has actually been estimated that the average assembly state of *E. coli* proteins is four (Goodsell, 1991).

The role of oligomerisation is also fundamental to molecular evolution, and it has been speculated that quaternisation of structure is central to the appearance of new proteins and functions (Lupas *et al.*, 2001, Smock *et al.*, 2016); by harnessing the free energy gain from an interaction, evolutionary nascent structures may be stabilised enough to be functional. Interestingly, the relationship between assembly and evolution might be so close that many proteins are apparently on the cusp of multimerisation; only a few mutations away from forming large assemblies (Garcia-Seisdedos *et al.*, 2017). For other systems, the oligomeric state is actually already favoured, but its access prevented for kinetic reasons (Baldwin *et al.*, 2011). However, the case of misfolding and aggregation—where the assembly is off-pathway to the native, functional, state of the protein—represents another set of questions and challenges somehow distinct from functional oligomerisation. In comparison to our understanding of folded proteins, the topological classification (Levy *et al.*, 2006, Ahnert *et al.*, 2015), assembly mechanisms (Levy *et al.*, 2008), and evolutionary origin of protein oligomers (Marsh & Teichmann, 2015) is only starting to emerge (*N.B.* these comments mostly apply to *functional* oligomers).

For any set of polypeptide chains, whether they fold individually, remain disordered, or produce multimeric assemblies, is a result of their composite energy landscape, *i.e.* the thermodynamic distribution, and kinetic accessibility of these states. Because of the multi-molecular nature of oligomers—and the inherent difficulties associated with studying concentration-dependent processes—much less is known about the thermodynamics

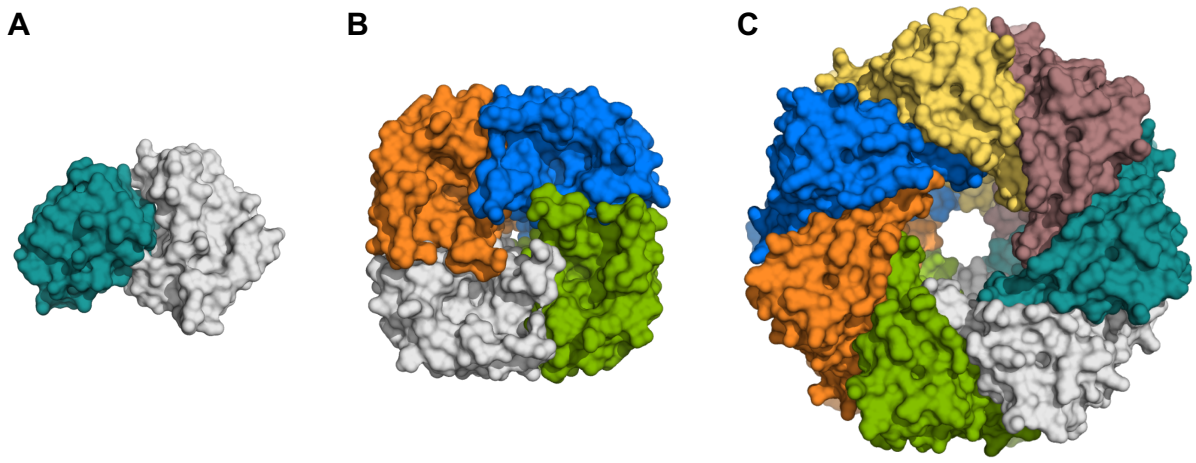

Figure 1.3 – Examples of protein assemblies. (**A**) Hetero-dimer between barnase and barstar (PDB:2ZA4). (**B**) Homo-tetramer of aquaporin (PDB:2ABM). (**C**) Homo-heptamer of α-hemolysin (PDB:3M2L). Each protein chain is coloured differently.

and kinetics of functional oligomerisation than protein folding. The BCL-2 family is an important system where oligomerisation is a key determinant of the biological function (Czabotar *et al.*, 2014, Kale *et al.*, 2017). These assemblies are regulated, *i.e.* not 'constitutive', and as such, represent an interesting starting point to study the biophysics of protein oligomerisation in the context of an interaction network. In part II of this thesis, some fundamental aspects of BCL-2 regulation are explored. By drawing a link between the biophysical characteristics of the components and their interactions, insights into the emerging properties of the network are gained. These provide a mechanistic understanding of the regulation of this biological process.

1.2 Aims of this thesis

When starting the work described in this thesis, intrinsic protein disorder had already been recognised as prevalent in biology, and not just an occasional oddity (Ward *et al.*, 2004, Wright & Dyson, 2015). The realisation that proteins can function in the absence of a pre-defined fold challenges many of the canonical tenets of molecular biology, and raises questions about the functional role of disorder. Understanding why it might have arisen during evolution leads to the question: what, if any, are the advantages and/or consequences of being disordered? Part I aims at understanding the biophysical response of coupled folding and binding reactions to their environment, *i.e.* the solution conditions. While much of the early work done in the field has focused on understanding the role of

intrinsic elements (the sequence), much less is known about the impact of extrinsic factors (the solution). Work described in Chapter 4 suggests that electrostatics might only play a minor role in the context of PPIs involving IDPs; a potential consequence of protein disorder. Chatper 5 demonstrate that coupled folding and binding reactions are sensitive to subtle changes in their environment, specifically ion-types. This 'biophysical sensing'—which can be attributed to the features of the energy landscape of IDPs—could be an advantage of protein disorder; allowing protein-protein interactions to be modulated in response to changes in their surroundings.

Coupled folding and binding reactions are often involved in signalling pathways. At the systems level, these networks are usually composed of competing interactions, and multi-layered architectures. However, an understanding of how the kinetic and thermodynamic signatures of the components might affect the properties of the network is not well-understood. These questions were investigated in Part II of this thesis by extending binary bindings to a tripartite system composed of competing interactions. The BCL-2 family was used as a model system, and the role of order-disorder transitions in the regulation of oligomerisation was investigated. It was discovered that folded, embedded motifs could interact upon partial unfolding of the proteins. Moreover, the interplay between homo- and hetero-oligomerisation was key to the regulation of the assembly state. This 'hidden disorder' highlights that conformational plasticity—in the form of conditional folding (binding-induced folding), or conditional binding (unfolding-dependent binding)—might provide additional regulatory layers for modulating protein-protein interactions, oligomerisation, and biological function.

Chapter 2

Thermodynamics and kinetics of protein folding and interactions

This chapter describes some general principles of thermodynamics and kinetics pertinent to the study of protein folding and protein-protein interactions. More specific models are described in the text where appropriate.

Proteins are not 'special' molecules. They obey the same physical laws underlying the rest of chemistry. Thus, their properties and functions can be described using theories and concepts borrowed from physical chemistry. Thermodynamics provides the mathematical framework to describe systems at equilibrium, while chemical kinetics gives insights into the ways perturbed systems return to a new state of equilibrium.

2.1 Systems at equilibrium and thermodynamics

2.1.1 The definition of spontaneity

It is often useful to know if a reaction is spontaneous, *i.e.* whether it can proceed without an input of energy. Whilst this knowledge does not provide any information about kinetics, it provides the basis to assess feasibility. Moreover, an understanding of the energetic of a system allows its equilibrium properties to be described.

Given enough time, all systems reach a state of equilibrium. This energy minima—where macroscopic aspects of the ensemble no longer change—is described by the laws and theories of thermodynamics. The application of its principles allows molecular systems to be quantified, and provides information on the equilibrium energetics of protein folding,

protein-protein interactions, and other aspects of protein biophysics.

The second law of thermodynamics states that for any process to be spontaneous, the net change in entropy of the universe must be positive:

$$\Delta S_{\text{uni}} = \Delta S_{\text{sys}} + \Delta S_{\text{surr}} \geq 0 \tag{2.1}$$

where ΔS_{uni} is the change of entropy of the universe for that process, and ΔS_{sys} and ΔS_{surr} represent the entropy changes of the system and the surroundings respectively. This distinction is purely arbitrary, and the boundaries of what represents the 'system' are often simply chosen to provide a convenient description of the process. In the context of typical biophysical experiments, the system may be the the tube containing the protein under investigation, leaving the surroundings to be everything else.

Entropy is defined statistically by the number of microstates available to the system:

$$S = k_{\text{B}} \ln W \tag{2.2}$$

where k_{B} is the Boltzmann constant, and W the number of microstates. For proteins, this concept might be understood in terms of the number of available conformations the backbone and side-chains may sample. It is also important to consider the contributions from the solvent and other co-solutes to the entropy of the system. For example, an unfolded protein might force water molecules to gain order and 'rigidify' in order to solvate the chain. Upon folding, these water molecules are released, which contributes favourably to the entropy of the system (Privalov & Makhatadze, 1993, Makhatadze & Privalov, 1996). On the other hand, the loss of conformational states available to the protein would contribute unfavourably (Baxa *et al.*, 2014). This example highlights the importance of considering *all* the components of the system when assessing its energetics.

The entropy of the surroundings, however, is not as simple to rationalise using this approach. In fact, assessing the spontaneity of a reaction directly from the second law is cumbersome; for any process, it would require the change of entropy of the entire surroundings to be quantified, which would be difficult to compute.

This problem may be solved by describing the entropy of the universe purely in terms of descriptors relating to the system. This can be achieved by using the relationships between thermodynamic quantities. Starting from the classical definition of entropy:

$$\mathrm{d}S = \frac{\delta q_{\mathrm{rev}}}{T} \tag{2.3}$$

where δq_{rev} represents the heat exchanged for a reversible process. Typical biochemical and biophysical experiments are performed at constant pressure. Under these conditions, the heat exchanged between the system and the surroundings is equal to the enthalpy ($H = U + pV$), and thus:

$$\mathrm{d}H = \delta q \tag{2.4}$$

Since H is a state function, $\mathrm{d}H_{\mathrm{surr}} = -\mathrm{d}H_{\mathrm{sys}}$, and the entropy change of the surroundings may be written as a function of the enthalpy change of the system:

$$\mathrm{d}S_{\mathrm{surr}} = \frac{\mathrm{d}H_{\mathrm{surr}}}{T} = \frac{-\mathrm{d}H_{\mathrm{sys}}}{T} \tag{2.5}$$

By combining this definition with Equation 2.1, the second law can be written as:

$$\Delta S_{\mathrm{uni}} = \frac{-\Delta H_{\mathrm{sys}}}{T} + \Delta S_{\mathrm{sys}} \geq 0 \tag{2.6}$$

which can be re-arranged to:

$$-T\Delta S_{\mathrm{uni}} = \Delta H_{\mathrm{sys}} - T\Delta S_{\mathrm{sys}} \leq 0 \tag{2.7}$$

Here, the second law is written purely in terms of quantities relating to the system, avoiding the need to quantify the entropy change of the surroundings. By defining $\Delta G_{\mathrm{sys}} = -T\Delta S_{\mathrm{uni}}$, and dropping the subscript notation since all thermodynamic quantities are properties of the system, the familiar equation for the Gibbs free energy used in chemistry can be recovered:

$$\Delta G = \Delta H - T\Delta S \leq 0 \tag{2.8}$$

which corresponds to the change of Gibbs free energy ($G = H - TS$) at constant temperature. Therefore, for a process to be spontaneous, the free energy of the system needs to *decrease*, which relates to an *increase* of the entropy of the universe.

2.1.2 Relationship between free energy and equilibrium distribution

The sign of ΔG is useful, as it contains the information about the direction in which a reaction is spontaneous. But the value of the Gibbs free energy also describe the equilibrium position of the system; an important value for quantifying the distribution between states in an ensemble. The relationship between energy and equilibria is described below. The case is made for ideal gases, and the result translated into the more biophysically relevant setting of solutions.

Starting from the differential definition of the Gibbs free energy:

$$\mathrm{d}G = V\mathrm{d}p - S\mathrm{d}T \tag{2.9}$$

and considering a process involving ideal gases ($pV = nRT$) at constant temperature ($\mathrm{d}T = 0$), an infinitesimal change in free energy can be written as:

$$\mathrm{d}G = \frac{nRT}{p}\mathrm{d}p \tag{2.10}$$

The change in free energy as a function of pressure can therefore be obtained by integration, which gives the following result:

$$G(p_2) - G(p_1) = nRT \ln\left(\frac{p_2}{p_1}\right) \tag{2.11}$$

This is often expressed with reference to the standard pressure of 1 bar (denoted $p^{\ominus}$). Therefore, the Gibbs free energy of a gas at pressure p can be described by reference to its standard state by:

$$G(p) = G^{\ominus} + nRT \ln\left(p/p^{\ominus}\right) \tag{2.12}$$

where $G^{\ominus}$ is the free energy of the gas at the standard pressure ($p^{\ominus}$). For a mixture of gases, the Gibbs free energy is expressed by the sum of the free energies of its constituents:

$$\begin{aligned} G_{\text{mix}} &= \sum_{i=1}^{N} \left(G_i^{\ominus} + n_i RT \ln\left(p_i/p^{\ominus}\right)\right) \\ &= \sum_{i=1}^{N} G_i^{\ominus} + RT \sum_{i=1}^{N} n_i \ln\left(p_i/p^{\ominus}\right) \end{aligned} \tag{2.13}$$

Now assuming that this mixture goes from an arbitrary state 1 to another arbitrary state 2, the change in Gibbs free energy for this process ($\Delta_{\text{r}}G$) is described by:

$$\Delta_{\text{r}}G = G_{\text{mix}}(2) - G_{\text{mix}}(1) \tag{2.14}$$

and combining Equations 2.13 and 2.14, the equation becomes:

$$\begin{aligned} \Delta_{\text{r}}G = &\left[\sum_{i=1}^{N} G_i^{\ominus}(2) + RT \sum_{i=1}^{N} n_i \ln\left(p_i(2)/p^{\ominus}\right) \right] \\ &- \left[\sum_{i=1}^{N} G_i^{\ominus}(1) + RT \sum_{i=1}^{N} n_i \ln\left(p_i(1)/p^{\ominus}\right) \right] \end{aligned} \tag{2.15}$$

which, using the properties of logarithms, can be re-arranged to

$$\Delta_{\text{r}}G = \Delta_{\text{r}}G^{\ominus} + RT \ln\left(\frac{\prod_{i=1}^{N} (p_i(2)/p^{\ominus})^{n_i}}{\prod_{i=1}^{N} (p_i(1)/p^{\ominus})^{n_i}} \right) \tag{2.16}$$

where $\Delta_{\text{r}}G^{\ominus}$ represent the Gibbs free energy change for the reaction going from pure reactants at standard pressure, to pure products at standard pressure.

In ideal solutions, the equivalent of partial pressures are concentrations ($p_i \propto c_i = n_i/V$), and the standard concentration ($c^{\ominus}$) is usually taken to be 1 M. Thus, Equation 2.16 can be written for the case of solution equilibria as:

$$\Delta_{\text{r}}G = \Delta_{\text{r}}G^{\ominus} + RT \ln\left(\frac{\prod_{i=1}^{N} (c_i(2)/c^{\ominus})^{n_i}}{\prod_{i=1}^{N} (c_i(1)/c^{\ominus})^{n_i}} \right) \tag{2.17}$$

It is noted that for real solutions, *activities* should be used instead of concentrations ($a_i = \gamma_i \cdot c_i$). These represent the 'effective' concentration of dissolved species when the precepts of ideality are no longer satisfied. However, for sufficiently dilute solutions—which is normally the case in typical biophysical experiments—interactions between solutes may be neglected, in which case the approximation of concentration is acceptable.

For a reaction at equilibrium, $\Delta_r G = 0$. Taking the simple case of an equilibrium between 2 states ($A \rightleftharpoons B$), Equation 2.17 becomes the familiar relationship:

$$\Delta G^{\ominus} = -RT \ln \frac{(c_B/c^{\ominus})}{(c_A/c^{\ominus})} = -RT \ln K_{\text{eq}} \tag{2.18}$$

where the subscript for reaction was omitted for simplicity. Therefore, the standard Gibbs free energy of a process is related to its composition at equilibrium, which is described by the constant K_{eq}.

2.1.3 Free energy of protein folding

Most small proteins fold cooperatively and reversibly, with the native state being attained from the denatured state in an 'all-or-nothing' manner; intermediates—if they exist—are too high in energy to be populated at equilibrium. Such folding processes are best described by a 2-state model:

$$D \rightleftharpoons N$$

where D and N represent the denatured and native state of the protein respectively. The equilibrium constant of the system is thus simply:

$$K_{\text{N-D}} = \frac{[N]}{[D]} \tag{2.19}$$

where $[N]$ and $[D]$ represent the concentrations of native and denatured states at equilibrium, and $K_{\text{N-D}}$ is the folding equilibrium constant. It is customary in protein biophysics to talk about *unfolding*, which is simply the inverse reaction ($K_{\text{D-N}} = K_{\text{N-D}}^{-1}$). Using Equation 2.17, the free energy of unfolding is expressed by:

$$\Delta G^{\ominus}_{\text{D-N}} = -RT \ln K_{\text{D-N}} = -RT \ln \frac{[D]}{[N]} \tag{2.20}$$

The typical stabilities of small globular proteins (5–15 kcal/mol, (Fersht, 1999)) implies that at equilibrium, $[D]$ is extremely small. So small, in fact, that quantifying stability by direct measurements of the distribution of native and denatured states in buffer is almost impossible. Instead, protein folding stabilities are typically measured by applying a perturbation to the system. This may be achieved by changing the temperature, which is the approach used in differential calorimetry experiments. Unfortunately, many proteins

do not unfold reversibly when heated, and usually start to aggregate. This precludes an accurate determination of thermodynamic parameters, as reversibility is a pre-requisite to such analyses. The use of chemical denaturants provides an alternative to this problem, and many proteins do unfold reversibly in the presence of *e.g.* urea or GdmCl. It has been empirically demonstrated that the folding free energy of a protein is linearly dependent on the concentration of denaturant (Tanford, 1968, Pace, 1986).

$$\Delta G_{\text{D-N}}^{[\text{den}]} = \Delta G_{\text{D-N}}^{\text{buffer}} - m \cdot [\text{den}] \tag{2.21}$$

where $\Delta G_{\text{D-N}}^{[\text{den}]}$ represents the free energy difference between the folded and denatured states at a particular concentration of denaturant, $\Delta G_{\text{D-N}}^{\text{buffer}}$ is the free energy in buffer (the value of interest), and m is a constant of proportionality. Thus, it is possible to estimate the unfolding free energy of a protein by performing a denaturation titration, and extrapolating the result back to buffer.

Measuring ΔG for protein folding

In this thesis, ensemble spectroscopic techniques were used to determine $\Delta G_{\text{D-N}}$. Under any given condition, the observed signal is the sum of its spectral components. In the case of protein folding, these components are the spectroscopic signals associated with the native and denatured states, weighed by their relative abundances:

$$S_{\text{obs}} = f_N \cdot S_N + f_D \cdot S_D = S_D + (S_N - S_D) \cdot f_N \tag{2.22}$$

The fraction of folded protein (f_N) can be expressed as a function of the equilibrium constant of the unfolding reaction ($K_{\text{D-N}}$):

$$f_N = \frac{[N]}{[D] + [N]} = \frac{1}{\frac{[D]}{[N[} + 1} = \frac{1}{K_{\text{D-N}} + 1} \tag{2.23}$$

By combining Equations 2.20 and 2.21, the equilibrium constant can be expressed as the function of the free energy difference in buffer, the m-value, and the concentration of denaturant:

$$K_{\text{D-N}} = \exp\left(-\frac{(\Delta G_{\text{D-N}}^{\text{buffer}} - m \cdot [\text{den}])}{RT}\right) \tag{2.24}$$

Finally, combining Equations 2.22, 2.23, and 2.24 leads to:

$$S_{\text{obs}} = S_D + (S_N - S_D) \cdot \left[1 + \exp\left(-\frac{(\Delta G^{\text{buffer}}_{\text{D-N}} - m \cdot [\text{den}])}{RT}\right)\right]^{-1} \tag{2.25}$$

which relates the change in spectroscopic signal (the dependent variable) to the change in denaturant concentration (the independent variable). In this equation, S_D, S_N, m, and $\Delta G^{\text{buffer}}_{\text{D-N}}$ are fitting parameters. If the signals of D and N change with the concentration of denaturant, linear correction terms may be added. This equation allows the folding free energy from chemical denaturation curves of proteins that are 2-state at equilibrium to be estimated.

For proteins that do not participate in multimeric assemblies, their stabilities are *independent* of their concentrations. This can be demonstrated from the thermodynamic treatment shown above, and replacing concentrations by molar fractions of the total protein concentration ($x_i = c_i/[P]$). Taking Equation 2.17, and expressing the concentrations in terms of molar fractions for the protein unfolding case:

$$\Delta_{\text{r}} G_{\text{D-N}} = \Delta_{\text{r}} G^{\ominus}_{\text{D-N}} + RT \ln \frac{(x_D[P]/c^{\ominus})}{(x_N[P]/c^{\ominus})} = \Delta_{\text{r}} G^{\ominus}_{\text{D-N}} + RT \ln \frac{x_D}{x_N} \tag{2.26}$$

where x_D and x_N are the molar fractions of unfolded and folded species respectively, and $[P] = [D]+[N]$ represent the total protein concentration. Because this term cancels out, the reaction is only dependent on the *relative* concentration of the species. This concentration independence is an important properties of unimolecular processes.

2.1.4 Protein-protein interactions

Unlike protein folding reactions, protein-protein interactions are multimeric processes. Taking a simple case, for example a 2-state hetero-dimerisation reaction—where the species exist either separate or in complex—the reaction schemes becomes:

$$A + B \rightleftharpoons AB$$

and the associated equation to describe the free energy of the reaction (in terms of molar fractions):

$$\Delta_{\text{r}} G = \Delta_{\text{r}} G^{\ominus} + RT \ln \frac{(x_{AB} \cdot [P]/c^{\ominus})}{(x_A \cdot [P]/c^{\ominus})(x_B \cdot [P]/c^{\ominus})} = \Delta_{\text{r}} G^{\ominus} + RT \ln \frac{x_{AB}}{x_A \cdot x_B} \frac{c^{\ominus}}{[P]} \tag{2.27}$$

where $[P] = [A] + [B] + 2\dot{[}AB]$ is the total concentration of protein defined in terms of monomers, and the molar fractions x_i have the same meaning as above. It can be seen that, since the term $[P]$ no longer cancels out, the free energy of the system now depends on both the *relative* and *absolute* concentrations. This is the reason why the equilibrium position of multimolecular processes—and thus their apparent stabilities—shifts with changes in concentrations. The same applies for homo-oligomerisation reactions.

Unlike for protein folding, where an external perturbation is necessary in order to measure the ΔG of the system, the concentration-dependence of binding reactions implies that their free energies can be directly determined from titration experiments. Protein-protein interactions are often defined by their equilibrium dissociation constants:

$$K_{\mathrm{d}} = \frac{[A][B]}{[AB]} = \frac{([A]_{\mathrm{tot}} - [AB])([B]_{\mathrm{tot}} - [AB])}{[AB]} \tag{2.28}$$

where $[A]$, $[B]$, and $[AB]$ represent the concentrations of both partners and the complex at equilibrium, while $[A]_{\mathrm{tot}}$ and $[B]_{\mathrm{tot}}$ are the total concentration of each protein. This equation can be re-arranged to a quadratic, which has the following solutions:

$$[AB] = \frac{[A]_{\mathrm{tot}} + [B]_{\mathrm{tot}} + K_{\mathrm{d}} \pm \sqrt{([A]_{\mathrm{tot}} + [B]_{\mathrm{tot}} + K_{\mathrm{d}})^2 - 4[A]_{\mathrm{tot}}[B]_{\mathrm{tot}}}}{2} \tag{2.29}$$

Here the concentration of complex at equilibrium is expressed as a function of the total concentration of each species, and the equilibrium dissociation constant of the reaction.

For an equilibrium binding experiments, one partner is typically labelled and kept at a constant concentration, while the other partner is titrated. The signal change associated with the labelled component as a function of the total concentration of titrant is recorded, and the binding isotherm fitted to the following equation:

$$S_{\mathrm{obs}} = S_A + (S_{AB} - S_A) \cdot \frac{[A]_{\mathrm{cst}} + [B] + K_{\mathrm{d}} - \sqrt{([A]_{\mathrm{cst}} + [B] + K_{\mathrm{d}})^2 - 4[A]_{\mathrm{cst}}[B]}}{2 \cdot [A]_{\mathrm{cst}}} \tag{2.30}$$

where S_{obs} is the observed signal (the dependent variable, *e.g.* anistropy), S_A is the signal of the free labelled species, S_{AB} its signal in complex, K_{d} is the equilibrium dissociation constant, $[A]_{\mathrm{cst}}$ the concentration of labelled species that is kept constant, and $[B]$ the concentration of titrant (the independent variable). Thus, the free energy of binding can be calculated from the dissociation constant by:

$$\Delta G^{\ominus} = -RT \ln\left(K_{\mathrm{d}}/c^{\ominus}\right) \tag{2.31}$$

Throughout this thesis, values for equilibrium constants were reported with units of M or its derivatives. This is strictly speaking incorrect, as K's should be dimensionless. However, having units is convenient for titration experiments, as it relates the value of K_{d} to the concentration at which half the titrated protein is in complex (provided that its concentration is much lower than K_{d}).

2.2 Systems out of equilibrium and kinetics

When systems are perturbed, they spontaneously return to a new state of equilibrium over a given amount of time. The features of these processes hold important mechanistic information about the transformation, and kinetic analyses provide insights into the molecular events that constitute reactions. Kinetics was employed in this thesis to investigate mechanistic aspects of protein (un)folding and interactions.

2.2.1 The rate of reactions

For an arbitrary reaction of the form:

$$aA + bB + ... \underset{k_2}{\overset{k_1}{\rightleftharpoons}} cC + dD + ...$$

where a, b, c, and d are the stoichiometric coefficient of the reaction, the rate of reaction is defined by:

$$\nu = -\frac{1}{a}\frac{\mathrm{d}[A]}{\mathrm{d}t} = -\frac{1}{b}\frac{\mathrm{d}[B]}{\mathrm{d}t} = \frac{1}{c}\frac{\mathrm{d}[C]}{\mathrm{d}t} = \frac{1}{d}\frac{\mathrm{d}[D]}{\mathrm{d}t} \tag{2.32}$$

and the the rate *law* of the reaction is described by:

$$\nu = k_1[A]^{\alpha}[B]^{\beta} - k_2[C]^{\gamma}[D]^{\delta} \tag{2.33}$$

where k_1 and k_2 are the rate constants for the forward and reverse reaction respectively, and α, β, γ, and δ represent the partial orders of the different species. It is noted that, except for elementary steps, these partial orders do not necessary relate to the stoichiometry of the reaction, and need to be determined experimentally (usually by the isolation method).

It is often useful to know the integrated form of the rate law, which allows the concentration of species as function of time to be determined. For simple systems, analytical solutions usually exist, while for more complicated functions, numerical integration might be necessary.

2.2.2 Protein folding

Many protein folding reactions have been shown to be consistent with a simple 2-state process that only involves a single transition state. For such scenarios, the reaction scheme is represented by:

$$D \underset{k_\mathrm{u}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} N$$

and the rate of the reaction is defined by:

$$\frac{\mathrm{d}[N]}{\mathrm{d}t} = -\frac{\mathrm{d}[D]}{\mathrm{d}t} = k_\mathrm{f}[D] - k_\mathrm{u}[N] \tag{2.34}$$

where $[D]$ and $[N]$ represent the concentration of denatured and native states respectively, and k_f and k_u are the rate constants of folding and unfolding. By recognising that $[P]_\mathrm{cst} = [N] + [D]$ (where $[P]$ is the total protein concentration), this relationship can be expressed as a first-order ordinary differential equation, which can be integrated to yield:

$$[N]_t = C \cdot \exp(-k_\mathrm{obs} t) + \frac{k_\mathrm{f}[P]}{k_\mathrm{obs}} \tag{2.35}$$

where $k_\mathrm{obs} = k_\mathrm{f} + k_\mathrm{u}$, and C is the integration constant, which represents the amplitude of the reaction.

At equilibrium, the macroscopic features of the system are constant, and the concentrations of each species are at steady-state:

$$\frac{\mathrm{d}[N]}{\mathrm{d}t} = -\frac{\mathrm{d}[D]}{\mathrm{d}t} = k_\mathrm{f}[D] - k_\mathrm{u}[N] = 0 \tag{2.36}$$

Therefore, the forward and reverse reactions equate each other:

$$k_\mathrm{f}[D] = k_\mathrm{u}[N] \iff \frac{k_\mathrm{u}}{k_\mathrm{f}} = \frac{[D]}{[N]} = K_\mathrm{D\text{-}N} \tag{2.37}$$

Thus, the equilibrium constant of the reaction is proportional to the ratio of its reverse and forward rate constants, which allows thermodynamic parameters to be determined from kinetic experiments, and *vice versa*.

As for equilibrium folding studies, the kinetics of protein folding is usually studied using chemical denaturants (Tanford, 1968, 1970). Using the results from transition-state theory, the rate of a reaction is related its free energy of activation ($\Delta G^{\ddagger}$) by:

$$k = A \cdot \exp\left(\frac{-\Delta G^{\ddagger}}{RT}\right) \tag{2.38}$$

where A is a pre-exponential factor. Assuming the same linear dependence for the free energy of activation as that described in Equation 2.21, the folding rate constant of a protein as a function of the concentration of denaturant can be modelled by:

$$k_{\mathrm{f}}^{[\mathrm{den}]} = A \cdot \exp\left(\frac{-\left(\Delta G_{\ddagger\text{-N}}^{\mathrm{buffer}} - m_{\ddagger\text{-N}}[\mathrm{den}]\right)}{RT}\right) \tag{2.39}$$

where $k_{\mathrm{f}}^{[\mathrm{den}]}$ is the folding rate constant at [den], $\Delta G_{\ddagger\text{-N}}^{\mathrm{buffer}}$ represent the free energy of activation in buffer, and $m_{\ddagger\text{-N}}$ is a proportionality constant. By taking the logarithm on both sides:

$$\ln k_{\mathrm{f}}^{[\mathrm{den}]} = \ln A - \frac{\left(\Delta G_{\ddagger\text{-N}}^{\mathrm{buffer}} - m_{\ddagger\text{-N}}[\mathrm{den}]\right)}{RT} \tag{2.40}$$

and recognising that:

$$\frac{-\Delta G_{\ddagger\text{-N}}^{\mathrm{buffer}}}{RT} = \ln k_{\mathrm{f}}^{\mathrm{buffer}} - \ln A \tag{2.41}$$

the equation can be re-written as:

$$\ln k_{\mathrm{f}}^{[\mathrm{den}]} = \ln k_{\mathrm{f}}^{\mathrm{buffer}} + \frac{m_{\ddagger\text{-N}}}{RT}[\mathrm{den}] \tag{2.42}$$

Thus, the folding rate of a protein in buffer may be estimated by performing experiments at different concentrations of denaturant, and extrapolating the result back to buffer. It is noted that the pre-exponential factor (A) is assumed to be independent of [den]. A similar treatment allows an expression for the unfolding rate constant to be obtained, and k_{f} and k_{u} are usually estimated by fitting chevron plots, which capture the results from both folding and unfolding experiments.

2.2.3 Protein-protein interactions

Despite the apparent complexity of protein-protein interactions—especially when folding is involved—many systems can actually be described by simple 2-state kinetic models, at least when monitored by ensemble spectroscopic techniques. A bimolecular, 2-state (single TS, no intermediate), reversible reaction can be modelled according to the following scheme:

$$A + B \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} AB$$

The rate law for such a process is defined by:

$$\frac{\mathrm{d}[AB]}{\mathrm{d}t} = k_{\text{on}}[A][B] - k_{\text{off}}[AB] \tag{2.43}$$

Pseudo-first order conditions

If either A or B is in excess, *i.e.* its concentration is assumed to remain constant over the course of the reaction, then this second order reaction simplifies to a pseudo-first order process:

$$\frac{\mathrm{d}[AB]}{\mathrm{d}t} = k'[A] - k_{\text{off}}[AB] \tag{2.44}$$

where $k' = k_{\text{on}}[B]_{\text{cst}}$. By recognising that $[A] = [A]_{\text{tot}} - [AB]$ (where $[A]_{\text{tot}}$ is the total concentration of A), it is possible to re-arrange Equation 2.44 to a first-order differential equation, which can be analytically integrated, giving Equation 2.45.

$$[AB]_t = C \cdot \exp(-k_{\text{obs}}t) - \frac{k_{\text{off}}[A]_{\text{tot}}}{k_{\text{obs}}} \tag{2.45}$$

where $k_{\text{obs}} = k_{\text{on}}[B]_{\text{cst}} + k_{\text{off}}$, and C is the integration constant. If the reaction is followed spectroscopically (*e.g.* fluorescence or CD), it is customary to express the signal at any given time by:

$$S(t) = S_0 + \Delta S \cdot \exp(-k_{\text{obs}}t) \tag{2.46}$$

where the integration constant C and the term $-\frac{k_{\text{off}}[A]_{\text{tot}}}{k_{\text{obs}}}$ from Equation 2.45 are represented by ΔS and S_0 respectively. The latter corresponds to the signal at the start of the reaction, while the former indicates the change in signal upon complex formation.

Fitting Equation 2.46 to data collected under pseudo-first order conditions allows an apparent rate constant (k_{obs}) to be obtained. Repeating the experiment at different concentrations of the species in excess, and fitting the data (k_{obs} *vs* [excess]) to a straight line, allows the association rate constant (k_{on}) to be determined by taking the gradient of the linear fit. While in theory this approach also yields the dissociation rate constant (which corresponds to the intercept), the extrapolation of the fit is usually associated with large errors. Therefore, it is customary to obtain k_{off} through out-competition dissociation experiments. Alternatively, k_{off} can be obtained indirectly by measuring the equilibrium dissociation constant, and using the relationship $K_{\text{d}} = k_{\text{off}}/\ k_{\text{on}}$, which is appropriate for 2-state systems since:

$$k_{\text{on}}[A]_{\text{eq}}[B]_{\text{eq}} = k_{\text{off}}[AB]_{\text{eq}} \Longleftrightarrow \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[A]_{\text{eq}}[B]_{\text{eq}}}{[AB]_{\text{eq}}} = K_{\text{d}} \tag{2.47}$$

Second-order conditions

Pseudo-first order conditions typically require multiple measurements at different concentrations. For systems with 2-state kinetics, binding rate constants can be directly obtained from single measurements under bimolecular conditions, and fitting the data to appropriate models. This approach was used in Part I of this thesis.

A second-order approach for fitting measurements obtained under reversible, near-equimolar, conditions has been described (Shammas *et al.*, 2013). This model is valid for reactions performed at concentrations close to K_{d}, *i.e.* when both forward and reverse reactions are significant.

$$F = F_0 + \Delta F \frac{(b-z)(1-\exp(zk_{\text{on}}t))}{2\left(\left(\frac{b-z}{b+z}\right)\exp(zk_{\text{on}}t) - 1\right)} + a \cdot t \tag{2.48}$$

where F is the fluorescence signal, F_0 the initial fluorescence, ΔF the fluorescence amplitude of the reaction, a a linear drift term, $b = -(K_{\text{d}} + (1+x)[A])$, $z = (K_{\text{d}}^2 + 2(1+x)K_{\text{d}}[A]_0 + (x^2 - 2x + 1)[A]_0^2)^{1/2}$, and $x = [B]_0/[A]_0$ with A and B the two proteins involved in the binding reaction. All but x are fitting parameters. This equation allows both k_{on} and K_{d} (and therefore k_{off}) to be determined, provided that the contribution from the dissociation reaction is significant. This fitting strategy was applied to study the kinetics of spectrin association, as well as the association of MCL-1 and PUMA at low concentrations. For reactions performed under bimolecular conditions, but far away from the K_{d} (where

the dissociation reaction is negligible), the following near-equimolar, irreversible, binding model has been described (Malatesta, 2005):

$$F = F_0 + \Delta F[A]_0 \left(\frac{x - x\exp(-k_{\text{on}}t(1-x)[A]_0)}{1 - x\exp(-k_{\text{on}}t(1-x)[A]_0)} \right) + a \cdot t \tag{2.49}$$

where all the terms have the same meaning as for the reversible model. This equation was used to obtain k_{on} for most of the binding kinetics between PUMA and MCL-1 reported in Part I of this thesis.

Chapter 3

Materials and methods

All chemicals were purchased from commercial suppliers and used without further purification unless stated otherwise.

Chemicals and reagents were weighed on either a PB1502-S balance (Mettler Toledo, for masses >10 g), or an AS60 precision balance (Ohaus).

Reagents and proteins were filter-sterilised using syringe-driven membrane filter units with 0.22 μm pore sizes (PES or PVDF membranes, Millex).

Water used for preparing buffers and reagent was obtained from purifier units at a resistivity of 18.2 MΩ · cm (at 25 °C).

3.1 Buffers and common reagents

3.1.1 Common reagents

Ampicillin (ampicillin sodium salt, Fischer Scientific) stock solutions were prepared at a concentration of 100 mg/mL in water, filter-sterilized, aliquoted and stored at −20 °C. Stock solutions were used as 1/1000 dilutions (100 μg/mL final concentration).

IPTG (isopropyl β-D-1-thiogalactopyranoside, Thermo Scientific) stock solutions were prepared at a concentration of 1 M in water, filter-sterilized, aliquoted and stored at −20 °C. IPTG was used to induce the *lac* operon at final concentrations of 0.1–1 mM.

Thrombin (thrombin from bovine serum, Sigma) stock solutions were prepared at a concentration of 1 U/μL in water, filter-sterilized, aliquoted and stored at −20 °C. Stock solutions were used to cleave fusion proteins containing a thrombin cleavage site (LVPR-/-GS) at a final concentration of 10 U/mL.

DTT (dithiothreitol, Fischer Scientific) stock solutions were prepared at a concentration of 1 M in water, filter-sterilized, aliquoted and stored at −20 °C. Stock solutions were used to reduce disulphide bonds at final concentrations ranging from 1–5 mM.

LB (lysogeny broth) medium was prepared by dissolving 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl (added as capsules, MP Biomedicals) into 1 L of water, and autoclaving for 20 min at 120 °C.

2×TY (double yeast-tryptone) medium was prepared by dissolving 16 g of tryptone, 10 g of yeast extract and 5 g of NaCl (added as capsules, MP Biomedicals) into 1 L of water, and autoclaving for 20 min at 120 °C.

EDTA (ethylenediaminetetraacetic acid, VWR Chemicals) stock solutions were prepared at a concentration of 0.25 M by dissolving 93.06 g of the disodium dihydrate salt into 1 L of water, and correcting the pH to 8.0. EDTA was added to buffers at final concentrations of 1–5 mM.

Imidazole (Acros Organics) stock solutions were prepared at a concentration of 1 M by dissolving 34.04 g in 500 mL of water, and filtered. Imidazole was added to buffers at concentrations up to 500 mM for eluting His-tagged proteins from nickel resins.

PS20 (polyoxyethylene (20) sorbitan monolaurate, Fisher Scientific, commercial name TWEEN® 20) stock solutions were prepared at a concentration of 80 mM (9% v/v, 1000×CMC) by adding 90 μL of neat liquid to 910 μL of water. This detergent was used to induce the oligomerisation of BAK and BAX.

C12E8 (octaoxyethylene monododecyl ether, Anatrace) was purchased as a 25% w/w stock in water (∼470 mM). Working stock solutions were prepared at 90 mM (1000×CMC)

by adding 192 μL of stock to 808 uL of water. This detergent was used to induce the oligomersation of BAK and BAX.

C8E4 (tetraoxyethylene monooctyl ether, Anatrace) was purchased as a 50% w/w stock in water (∼1.6 M). Working stock solutions were prepared at 80 mM (10×CMC) in ammonium acetate buffer, or used neat. This detergent was used for screening oligomerisation conditions in Chapter 7.

OGP (octyl β-D-glucopyranoside, Sigma-Aldrich) stock solutions were prepared at a concentration of 100 mM (4×CMC) by dissolving 58.5 mg in 2 mL of ammonium acetate buffer. This detergent was used for screening oligomerisation conditions in Chapter 7.

CuPhe (Cu^{II}(1,10-phenanthroline)$_3$) stock solutions were prepared at a concentration of 25 mM (complex) by dissolving 0.016 g of $CuSO_4$ (Acros Organics) and 0.054 g of 1-10-phenantroline (Sigma Aldrich) into a 4:1 water:ethanol mixture (3.2 mL water, 0.8 mL ethanol). This reagent was used to induce disulfide bond formation.

3.1.2 Buffers

All buffers were prepared with ultrapure water, and filter-sterilised by passing through a 0.2 μm cellulose acetate membrane (Sartorius Stedium Biotech). All buffers used in biophysical experiments were prepared using volumetric glassware. Unless stated otherwise, buffers were prepared gravimetrically without pH adjustment. The pH of each buffer was checked by using a PHM210 pH-meter (Radiometer Analytical) following a two-point calibration with standard solutions (pH 4.00 and 10.00, Hannah Instruments). If necessary, the pH was adjusted using the appropriate set of acids and bases.

Preparative buffers for spectrin:

sPBS buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.88) was prepared as a 10× stock by dissolving 54.285 g of $Na_2HPO_4{\cdot}2H_2O$, 30.42 g of $NaH_2PO_4{\cdot}2H_2O$ and 87.54 g NaCl into 1 L of water. This stock was used to make 1× dilutions that included variable concentrations of imidazole.

Tris buffer for ion-exchange chromatography (20 mM Tris, pH 8.0) was prepared as a 20× stock by dissolving 35.30 g of tris(hydroxymethyl)aminomethane hydrochloride and 21.08 g of tris(hydroxymethyl)aminomethane into 1 L of water. This stock was used to make 1×, and 1× + 1 M NaCl (58.44 g/L) buffers used in ion-exchange chromatography, and for the preparation of Factor Xa cleavage buffer.

Preparative buffers for MCL-1:

PBS buffer (10 mM sodium phosphate, 137 mM NaCl, 3 mM KCl, pH 7.4) was prepared as a 20× stock by dissolving 28.83 g of $Na_2HPO_4{\cdot}2H_2O$, 5.93 g of $NaH_2PO_4{\cdot}2H_2O$, 159.54 g of NaCl, and 4.03 g of KCl into 1 L of water. This stock was used to make 1× dilutions that included variable concentrations of imidazole.

HEPES buffer for ion-exchange chromatography (10 mM HEPES, pH 7.5) was prepared as a 20× stock by dissolving 27.17 g of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, and 21.86 g of sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonate into 1 L of water. This stock was used to make 1×, and 1× + 1 M NaCl (58.44 g/L) buffers used for ion-exchange chromatography.

Preparative buffers for GB1-X_{BH3} constructs:

Factor Xa cleavage buffer (20 mM Tris pH 8.0, 50 mM NaCl, 5mM $CaCl_2$) was prepared by mixing 50 mL of Tris ion-exchange stock buffer, 2.92 g of NaCl, and 5 mL commercial $CaCl_2$ solution (1.0 M, Fluka Analytical) into 1 L of water. This buffer was used to cleave the GB1-PUMA (mouse) construct.

Biophysical buffers for Part I:

All ionic strength studies were performed in 10 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) with variable concentrations of the salts investigated. To minimise the necessity for pH corrections, approximate ratios of MOPS acid and MOPS base were calculated using an on-line server (www.liverpool.ac.uk/pfg/Research/Tools/BuffferCalc/Buffer.html), which accounts for the effect of ionic strength and temperature on the pK_a values of buffer components. Variable concentrations of either NaCl (Fisher Chemcial), KCl (VWR Chemicals), LiCl (Atom Scientific), $MgCl_2$ (Fluka Analytical), $CaCl_2$ (Fluka Analytical), NaBr

(Breckland Scientific), or NaI (Breckland Scientific) were added to modulate the ionic strengths of the buffers. $MgCl_2$ and $CaCl_2$ were added from commercial 1.0 M solutions (Fluka Analytical) in order to minimize errors associated with the hygroscopicity of these salts. After preparation, pHs were checked, and corrected if necessary. The exact ionic strength of each buffer was precisely back-calculated using:

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2 \tag{3.1}$$

where c_i is the concentration of a specific ion and z_i its net charge. All ions were taken into account, including the contribution from NaOH or HCl if the buffers were pH-corrected. For MOPS, the concentration of ionised species was calculated using the Henderson-Hasselbach equation, assuming a pK_a of 7.20 at 25 °C:

$$\mathrm{pH} = \mathrm{p}K_\mathrm{a} + \log_{10}\left(\frac{[\mathrm{MOPS}^-]}{[\mathrm{MOPS}^\pm]}\right) \tag{3.2}$$

Zwitterionic species do not influence the ionic strength (Stellwagen *et al.*, 2008), therefore only [MOPS^-] was included in the calculation of I, resulting in a contribution of about 4 mM from the buffer.

Buffers for spectrins were prepared as 1× stocks, and the proteins were buffer-exchanged using HiTrap Desalting columns (GE Healthcare). Buffers for PUMA:MCL-1 were prepared as 2× stocks. The lyophilised protein/peptide were reconstituted in water containing 0.1% (v/v) PS20, followed by mixing with buffers in a 1:1 volume ratio.

Preparative buffers for BAK and BAX:

HEPES buffer (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, pH 7 at 25 °C) was prepared by dissolving 7.41 g of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 2.29 g of sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonate, 11.69 g of NaCl, and 8 mL of EDTA stock solution into 2 L of water.

Biophysical buffers for Part II:

Sodium phosphate buffer (50 mM sodium phosphate, pH 7.0, $I \approx 100$ mM) was prepared by dissolving 10.57 g of $Na_2HPO_4{\cdot}2H_2O$ and 6.33 g of $NaH_2PO_4{\cdot}2H_2O$ into 2 L of

water. This buffer was used for all the experiments presented in Part II, except the native MS, and some of the initial detergent screening.

Ammonium acetate buffer (100 mM ammonium acetate, pH 7.0, $I \approx 100$ mM) was prepared by dissolving 7.71 g of ammonium acetate into 1 L of water. The pH was adjusted to 7.0 using acetic acid and ammonia. This buffer was used for all the native MS experiments, and some of the initial detergent screening.

Buffer for molecular biology:

TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.3) was prepared as a 50× stock by adding 242 g of tris(hydroxymethyl)aminomethane, 57 mL of glacial acetic acid, and 200 mL of EDTA stock to 1L of water. This buffer was used to prepare agarose gels for DNA electrophoresis.

3.2 Molecular biology

3.2.1 Bacterial strains and vectors

Competent cells:

All biological work was performed in chemically competent *Escherichia coli* cells. Cloning and DNA preparations were performed in either XL-1 Blue, or DH5α, which are *endA* and *recA* deficient strains. Deficiency in the *endA* endonuclease ensures greater yields in DNA minipreps, while *recA* deficiency prevents unwanted recombinations, increasing the stability of the insert. Recombinant protein expressions were performed in either BL21(DE3), C41(DE3), or C41(DE3)pLysS. These strains carry a DE3 recombinant phage harboring the T7 RNA polymerase gene under control of the *lac* operon. Thus, they are suited for T7 promoter-based expression systems. C41(DE3) (Miroux & Walker, 1996) is a derivative of BL21(DE3) that contains additional mutations advantageous for the production of toxic proteins. C41(DE3)pLysS is a derivative of C41(DE3) that harbours the pLysS plasmid, which constitutively expresses T7 lysosyme. This natural inhibitor of T7 RNA polymerase reduces basal expression levels of recombinant proteins prior to induction, allowing better cell growth when the protein is toxic.

Chemically competent cells were prepared from a starter culture (5 mL LB, no antibiotic) of the desired strain. Following overnight incubation at at 37 °C, 1 mL was used to inoculate 100 mL of LB, and the cells were grown at 37 °C until an OD_{600} of 0.3–0.5 was reached. At this point, the culture was put on ice for 20 min. Following centrifugation, the growth media was discarded, and the cells were re-suspended in 20 mL of pre-chilled 100 mM $CaCl_2$ containing 15% (v/v) glycerol. After another 20 min of incubation on ice, the sample was centrifuged, the supernatant discarded, and the cells re-suspended in 4 mL of the calcium/glycerol solution. Aliquots (100 µL for XL-1 Blue and DH5α, 50 µL for BL21(DE3), C41(DE3), and C41(DE3)pLysS) were flash frozen in liquid nitrogen, and stored at −80 °C.

Plasmids:

Constructs of spectrins (α0α1 and β16β17), MCL-1 (mouse and human), and GB1-X_{BH3} were sub-cloned into a modified version of pRSET A. This vector contains a N-terminal hexahistidine-tag followed by a thrombin cleavage site, the whole under the control of a T7 promoter. The plasmid also contains an ampicillin resistance gene. Due to cloning artefacts, an extra two residues (GS) were appended at the N-terminus of thrombin-cleaved proteins expressed from this vector. For GB1-X_{BH3} constructs, the thrombin cleavage site after the His-tag was mutated out.

Constructs of BAK and BAX were sub-cloned into the pTXB1 vector (New England Biolabs), which contains a C-terminal intein, followed by a chitin binding domain. The plasmid also contains an ampicillin resistance gene. These constructs were purified by binding to chitin resin, followed by self-cleavage of the intein induced by the addition of DTT.

Transformations:

The plasmid of interest (1 µL) was mixed with one aliquot of competent cells of the appropriate type, followed by incubation on ice for 20 min. Cells were heat-shocked (45 s at 42 °C), returned to ice, and incubated for another 5 min. The aliquot was completed to 1 mL with either LB or 2×TY, followed by 45 min of incubation at 37 °C. Cells were harvested by centrifugation, re-suspended, plated on 2×TY-agar plates containing 100 µg/mL of ampicillin, and incubated overnight at 37 °C.

3.2.2 Cloning

Primer design:

Primers for amplification were designed to have ∼20 nucleotides of overlap with the insert. Primers for site-directed mutagenesis were designed to have ∼15 nucleotides on both sides of the mismatched site. Nucleotide mutations were performed to match the codon usage bias of *E. coli*. Primers for In-Fusion® cloning were designed to have ∼20 nucleotides matching the sequence of interest, and homologous overlaps were made to be 15 bp long. Where possible, all primers were designed to contain two consecutive G/C nucleotides at their 3' end. Primers were ordered from Invitrogen (desalted), and reconstituted in water to a final concentration of a 100 µM. Forward/reverse primer mix used in PCR were made at final concentrations of 10 µM (each).

Polymerase chain reaction:

DNA amplification was performed *in vitro* using the polymerase chain reaction (PCR). Template and primers were assembled with a thermostable DNA polymerase (Phusion high-fidelity DNA polymerase, Thermo Scientific), and the amplification was performed by repeated cycles of denaturation-annealing-extension. Typical PCR mixtures were assembled as described in Table 3.1

Table 3.1 – Composition of PCR mixtures.

	Volume / µL	Final concentration
Water	30.5	–
5× Phusion HF buffer	10	1×
dNTP mix	5	200 µM (each)
Primer mix	2.5	0.5 µM (each)
Template	1	1–5 ng/µL
Phusion HF polymerase	1	1 U

Routine thermocycling protocols are described in Table 3.2. Annealing temperatures were adapted to the lowest primer:template melting temperature (estimated from the GC content) or experimentally optimised in the case of difficult reactions. Extension times were adapted to the length of the amplicon, and the number of cycles was reduced for long

PCR products in order to minimise the accumulation of mutations (*e.g.* for full vector amplification during In-Fusion® reactions).

Table 3.2 – Thermocycling protocols. The annealing temperature (x) was varied depending on the primers. The extension step was adapted for the length of the amplicon. The number of cycles (denaturation-annealing-extension) was varied between 25–35.

Step	Temperature / °C	Time / s
Initial denaturation	98	30
Denaturation	98	10
Annealing	x	10
Extension	72	30/kb
Final extension	72	300

Site-directed mutagenesis:

Point mutations were obtained by performing PCR with the appropriate set of mutagenic primers. Following amplification, 1 µL of *Dpn*I (Thermo Scientific) was added to the PCR product, and the sample was incubated at 37 °C for 3 h. More *Dpn*I (up to 3 µL) and longer incubation times (up to overnight) were employed for reactions that resulted in large numbers of false positive transformants. After template digestion, 10 µL were transformed into either XL-1 Blue or DH5α for *in vivo* plasmid amplification. Results were confirmed by sequencing.

Traditional cloning:

Traditional cloning using restriction enzymes was used to produce GB1-X_{BH3} constructs and human MCL-1. Synthetic genes were ordered from GenScript (codon-optimised for *E. coli*) and received in pUC57 vectors. Inserts were amplified by PCR using the appropriate set of primers, followed by digestion with pairs of restriction enzymes to generate overhanging ends (GB1-X_{BH3}: *Nde*I/*Eco*RI; MCL-1: *Bam*HI/*Eco*RI, Thermo Scientific). The receiving vector (pRSET A) was digested with matching sets of enzymes in the presence of phosphatase (Thermo Scientific). Reactions were performed at 37 °C for 1–3 h in the buffer recommended by the manufacturer. Digested products were purified by gel electrophoresis (1% agarose in TAE buffer, 100 V for 30 min) and stained with GelRed® (Biotium).

Bands were excised, and the DNA was extracted using a commercial kit according to the manufacturer's protocol (QIAquick® Gel Extraction Kit, Qiagen). Vectors and inserts were mixed in a 1:1–5 molar ratio, and ligated with T4 DNA ligase (Thermo Scientific) for 1 h at room temperature, or overnight at 4 °C. Ligation products were transformed into either XL-1 Blue or DH5α, minipreped, and sent for sequencing.

Seamless cloning:

Seamless cloning using In-Fusion® (Clontech) was used to produce the constructs of BAK and BAX. Synthetic genes were ordered from GenScript (codon-optimised for *E. coli*) and received in pUC57 vectors. Inserts (in pUC57) and receiving vectors (pTXB1) were amplified by PCR using appropriate sets of primers. These were designed to have insert/vector pairs with 15 bp of homologous overlap at each cloning site. PCR reactions were digested with *Dpn*I (Thermo Scientific) to remove templates. In-Fusion® reactions were performed by mixing 3 μL of water, 0.5 μL of insert and vector (each), and 1 μL of In-Fusion® buffer/enzyme mix, followed by incubation at 50 °C for 15 min. This cloning method utilises the combination of effects from a polymerase and an exonuclease. The latter creates 5' overhangs on the insert and the vector—which can then anneal— while the polymerase seals the gaps, and phosphodiester bonds are ultimately re-formed in *E. coli*. The products were transformed into either XL-1 Blue or DH5α, and the results confirmed by sequencing.

Sequencing:

All cloning results were confirmed by sequencing. Single colonies from overnight cultures on Petri dishes were used to inoculate 5 mL of media containing ampicillin. Cells were grown overnight at 37 °C, and plasmid DNA purified the next day using commercial miniprep kits (QIAprep® Spin Miniprep Kit, Qiagen). In brief, cells were pelleted, re-suspended in buffer, lysed under alkaline conditions in the presence of a detergent, and then neutralised; leading to the precipitation of chromosomal DNA, proteins, and lipids. The supernatant containing plasmid DNA was applied to a spin column composed of a silica matrix capable of binding charged DNA, washed with ethanol-containing buffer, and eluted with water. DNA concentrations were calculated by measuring the absorption at 260 nm ($\varepsilon_{260}^{\mathrm{dsDNA}} = 50$ ng^{-1} μL cm^{-1}).

Samples (10 µL at ∼100 ng/µL of plasmid DNA) were sent to Genewiz for Sanger sequencing with T7 primers. For constructs >1000 bp, or for reactions giving poor read-throughs, additional sequencing reactions from the 3' end were performed using T7 terminator primers. Results were analysed with CLC Sequence Viewer (CLC Bio).

3.3 Protein expression and purification

3.3.1 Protein purity and identity

Protein purity was assessed by SDS-PAGE on NuPAGE™ 4–12% Bis-Tris pre-cast gels (Invitrogen). Electrophoresis was performed in NuPAGE™ MES SDS running buffer (Invitrogen) for 35 min at a constant current of 120 mA. Protein samples were prepared in NuPAGE™ LDS sample buffer (Invitrogen), followed by denaturation at 95 °C for 5 min. Typical loading quantities were ∼0.1 nmol. If reducing conditions were necessary, 5 mM DTT was added prior to denaturation. Mass estimates were obtained by comparing the electrophoretic mobility of the samples with that of molecular weight markers (PageRuler™ Unstained Protein Ladder, Thermo Scientific). Gels were stained overnight with InstantBlue™ (Expedeon), and imaged on a BioDoc-It Imaging System (UVP).

The identity of the purified proteins were confirmed by positive ion LC-MS electro-spray on a Xevo G2-S QTof instrument (Waters). All samples were processed by the mass spectrometry facility of the Department of Chemistry of the University of Cambridge.

3.3.2 Spectrins

Both erythrocyte spectrin proteins from *Homo sapiens* were produced as truncated versions of their respective gene products. These constructs were designed, cloned, and tested by Dr Lee Gyan Kwa (Hill *et al.*, 2014). Construct α0α1 is composed of the first partial (α0) and full (α1) domains of α-spectrin (UniProt:P02549, residues 2–163, Fig. 3.1). Construct β16β17 is composed of the last full (β16) and partial (β17) domains of β-spectrin (UniProt:P11277, residues 1898–2083, Fig. 3.2).

Both spectrin proteins were expressed and purified according to the same protocol. Plasmids were transformed into C41(DE3), and cells were grown overnight at 37 °C on 2×TY-agar plates containing ampicillin. Pre-cultures (5 mL) were made from scrapes, and used

```
P02549      1  -MEQFPKETVVESSGPKVLETAEEIQERRQEVLTRYQSFKERVAERGQKLEDSYHLQVFK   59
a0a1        1  GSEQFPKETVVESSGPKVLETAEEIQERRQEVLTRYQSFKERVAERGQKLEDSYHLQVFK   60
               ************************************************************

P02549     60  RDADDLGKWIMEKVNILTDKSYEDPTNIQGKYQKHQSLEAEVQTKSRLMSELEKTREERF  119
a0a1       61  RDADDLGKWIMEKVNILTDKSYEDPTNIQGKYQKHQSLEAEVQTKSRLMSELEKTREERF  120
               ************************************************************

P02549    120  TMGHSAHEETKAHIEELRHLWDLLLELTLEKGDQLLRALKFQQYVQECADILEWIGDKEA  179
a0a1      121  TMGHSAHEETKAHIEELRHLWDLLLELTLEKGDQLLRALKFQQY----------------  164
               ****************************************
```

Figure 3.1 – Protein sequence of the α0α1 construct compared to the gene product of α-spectrin (residues 180–2419 are omitted). The cloning artefact (extra GS residues at the N-terminus) is highlighted in bold.

```
P11277   1861  RLQTAYAGEKAEAIQNKEQEVSAAWQALLDACAGRRTQLVDTADKFRFFSMARDLLSWME  1920
b16b17      1  ----------------------------------GSQLVDTADKFRFFSMARDLLSWME    25
                                                 :**********************

P11277   1921  SIIRQIETQERPRDVSSVELLMKYHQGINAEIETRSKNFSACLELGESLLQRQHQASEEI  1980
b16b17     26  SIIRQIETQERPRDVSSVELLMKYHQGINAEIETRSKNFSACLELGESLLQRQHQASEEI    85
               ************************************************************

P11277   1981  REKLQQVMSRRKEMNEKWEARWERLRMLLEVCQFSRDASVAEAWLIAQEPYLASGDFGHT  2040
b16b17     86  REKLQQVMSRRKEMNEKWEARWERLRMLLEVCQFSRDASVAEAWLIAQEPYLASGDFGHT   145
               ************************************************************

P11277   2041  VDSVEKLIKRHEAFEKSTASWAERFAALEKPTTLELKERQIAERPAEETGPQEEEGETAG  2100
b16b17    146  VDSVEKLIKRHEAFEKSTASWAERFAALEKPTTLELKERQIAE-----------------   188
               ****************************************

P11277   2101  EAPVSHHAATERTSPVSLWSRLSSSWESLQPEPSHPY                         2137
b16b17    189  ------------------------------------                          188
```

Figure 3.2 – Protein sequence of the β16β17 construct compared to the gene product of β-spectrin (residues 1–1860 are are omitted). The cloning artefact (extra GS residues at the N-terminus) is highlighted in bold.

to inoculate 1 L cultures (LB). Cells were incubated at 37 °C until an OD_{600} of 0.4–0.6 was reached. At this point, protein production was initiated by the addition of IPTG to a final concentration of 0.1 mM, and the temperature was dropped to 25 °C for overnight expression. Cells were harvested by centrifugation (5,000 rpm, 15 min, 4 °C), the pellets re-suspended in sPBS containing 25 mM imidazole, and sonicated on ice (3–5 min, 15 s pulses at 45 s intervals). Debris were cleared by centrifugation (18,000 rpm, 45 min, 4 °C), and proteins were purified from the soluble fraction by binding to Ni-NTA resin (Agarose Bead Technologies, 5 mL of slurry per litre of culture equivalent). After 1 h at 4 °C, the resins were washed with sPBS containing 25 mM imidazole, and bound proteins were released by overnight thrombin cleavage at room temperature (∼100 U per litre of

culture equivalent). The proteins were further purified by size-exclusion chromatography (SEC) on a ÄKTA FPLC system (GE Healthcare), using a Superdex 75 26/600 columm (GE Healthcare) equilibrated in sPBS. The proteins were eluted at an isocratic flow-rate of 2 mL/min, and the peaks were analysed by SDS-PAGE. Pure fractions were pooled, filter-sterilised, and stored at 4 °C.

Although not necessary, an additional ion-exchange step increased the yield of β16β17. The later part of the SEC peak co-eluted with a smaller contaminant (most likely a C-terminal truncation of the protein), and these fractions always had to be discarded. This impurity was removed by performing an ion-exchange step prior to SEC. Following overnight thrombin cleavage, the resin supernatant containing the protein was diluted in buffer A (20 mM Tris, pH 8.0) to reduce the ionic strength of the solution. This sample was loaded onto a HiTrap Q HP anion-exchange column (GE Healtcare) at a flow-rate of 2 mL/min, followed by washing with buffer A until conductivity and UV readings returned to baseline. The protein was eluted by applying a linear gradient of buffer B (20 mM Tris, pH 8.0, 1 M NaCl, 0–25% over 90 mL at 2 mL/min), which separated the contaminant (elution at ∼10% B) from the protein (elution at ∼15% B). Clean fractions were pooled, and further purified by SEC (*vide supra*).

Identities of the proteins were confirmed by MS (α0α1: 19420.91(±0.17) Da, theoretical 19420.79 Da; β16β17: 22037.59(±4.42) Da, theoretical 22036.99 Da).

3.3.3 MCL-1

The construct of mouse MCL-1 (induced myeloid leukemia cell differentiation protein) was designed, cloned, and tested by Dr Joseph M. Rogers (Rogers *et al.*, 2013). It only contains the globular part of the protein (UniProt:P97287, residues 152–308, Fig. 3.3); its N-terminal disordered tail and C-terminal membrane-anchoring sequence were not included in the construct, in line with previous structural work (Day *et al.*, 2005, 2008).

The construct of human MCL-1 was designed to match that of the mouse version. The synthetic gene was ordered from GenScript, and cloned as described (*vide supra*). The protein was three residues longer at the N-terminus; including additional glutamates from the sequence, which were incorporated to shift the pI away from neutral, and circumvent

```
P97287          1  MFGLRRNAVIGLNLYCGGASLGAGGGSPAGARLVAEEAKARREGGGEAALLPGARVVARP   60
MCL-1_mouse     1  ------------------------------------------------------------     0

P97287         61  PPVGAEDPDVTASAERRLHKSPGLLAVPPEEMAASAAAAIVSPEEELDGCEPEAIGKRPA  120
MCL-1_mouse     1  ------------------------------------------------------------     0

P97287        121  VLPLLERVSEAAKSSGADGSLPSTPPPPEEEEDDLYRQSLEIISRYLREQATGSKDSKPL  180
MCL-1_mouse     1  ------------------------------GSEDDLYRQSLEIISRYLREQATGSKDSKPL   31
                                                .****************************

P97287        181  GEAGAAGRRALETLRRVGDGVQRNHETAFQGMLRKLDIKNEGDVKSFSRVMVHVFKDGVT  240
MCL-1_mouse    32  GEAGAAGRRALETLRRVGDGVQRNHETAFQGMLRKLDIKNEGDVKSFSRVMVHVFKDGVT   91
                   ************************************************************

P97287        241  NWGRIVTLISFGAFVAKHLKSVNQESFIEPLAETITDVLVRTKRDWLVKQRGWDGFVEFF  300
MCL-1_mouse    92  NWGRIVTLISFGAFVAKHLKSVNQESFIEPLAETITDVLVRTKRDWLVKQRGWDGFVEFF  151
                   ************************************************************

P97287        301  HVQDLEGGIRNVLLAFAGVAGVGAGLAYLIR                               331
MCL-1_mouse   152  HVQDLEGG------------------------                               159
                   ********
```

Figure 3.3 – Protein sequence of mouse MCL-1 compared to the construct used in this thesis. The cloning artefact (extra GS residues at the N-terminus) is highlighted in bold.

potential solubility issues at pH 7. The construct also included the mutation C286S, so that the use of reducing agents could be avoided. As for the mouse version, the disordered N-terminus, and the C-terminal transmembrane regions were excluded (UniProt:Q07820, residues 168–327, Fig. 3.4).

Protocols for producing mouse and human versions were identical. Plasmids were transformed into C41(DE3), and the cells were grown overnight at 37 °C on 2×TY-agar plates containing ampicillin. Pre-cultures (5–10 mL) were prepared from scrapes, and used to inoculate 1 L cultures (LB) that were incubated at 37 °C until an OD_{600} of ∼0.6 was reached. Protein expression was induced by addition of IPTG (1 mM final concentration), and the temperature was reduced to 18 °C. After overnight expression, cells were harvested (5,000 rpm, 15 min, 4 °C), the pellet re-suspended in PBS buffer containing 25 mM imidazole, and sonicated on ice (3–5 min, 15 s pulses with 45 s intervals). Debris were cleared by centrifugation (18,000 rpm, 45 min, 4 °C), and the proteins purified from the supernatant by binding to Ni-NTA resin (5 mL of slurry per litre of culture equivalent) for ∼1 h at 4 °C, followed by three washes with PBS buffer containing 25 mM imidazole. Unlike spectrins, MCL-1 could not be cleaved off the resin. Thus, proteins were eluted by addition of PBS containing 500 mM imidazole, followed by buffer exchanging to thrombin cleavage buffer (20 mM Tris, 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5). This step was carried out by using

```
Q07820          1  MFGLKRNAVIGLNLYCGGAGLGAGSGGATRPGGRLLATEKEASARREIGGGEAGAVIGGS   60
MCL-1_human     1  ------------------------------------------------------------     0

Q07820         61  AGASPPSTLTPDSRRVARPPPIGAEVPDVTATPARLLFFAPTRRAAPLEEMEAPAADAIM  120
MCL-1_human     1  ------------------------------------------------------------     0

Q07820        121  SPEEELDGYEPEPLGKRPAVLPLLELVGESGNNTSTDGSLPSTPPPAEEEEDELYRQSLE  180
MCL-1_human     1  ----------------------------------------------GSEEEEDELYRQSLE   15
                                                                  :*************

Q07820        181  IISRYLREQATGAKDTKPMGRSGATSRKALETLRRVGDGVQRNHETAFQGMLRKLDIKNE  240
MCL-1_human    16  IISRYLREQATGAKDTKPMGRSGATSRKALETLRRVGDGVQRNHETAFQGMLRKLDIKNE   75
                   ************************************************************

Q07820        241  DDVKSLSRVMIHVFSDGVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESITDVLVR  300
MCL-1_human    76  DDVKSLSRVMIHVFSDGVTNWGRIVTLISFGAFVAKHLKTINQESSIEPLAESITDVLVR  135
                   *******************************************.**************

Q07820        301  TKRDWLVKQRGWDGFVEFFHVEDLEGGIRNVLLAFAGVAGVGAGLAYLIR            350
MCL-1_human   136  TKRDWLVKQRGWDGFVEFFHVEDLEGG-----------------------            162
                   *************************
```

Figure 3.4 – Protein sequence of human MCL-1 compared to the construct used in this thesis. The cloning artefact (extra GS residues at the N-terminus) is highlighted in bold. The construct also contained the mutation C286S.

centrifugal concentrators (Sartorius Stedium Biotech, 5,000 Da MWCO) and performing repeated cycles of concentration/dilution. EDTA (5 mM) was included at this stage to reduce precipitation, thus calcium chloride (10 mM $CaCl_2$) had to be re-introduced at the end of the procedure. Thrombin (200 U per litre of culture equivalent) was added, and proteolytic cleavage was carried out overnight at room temperature. The proteins were further purified by SEC on a Superdex 75 26/600 (GE Healthcare). The column was equilibrated in biophysical phosphate buffer, and the proteins were eluted at an isocratic flow-rate of 2 mL/min. Purity of the fractions were analysed by SDS-PAGE, and pooled accordingly. For short-term use, samples were filter-sterilised and stored at 4 °C. For long-term storage, the proteins were buffer-exchanged into water (HiTrap Desalting columns, GE Healthcare), freeze-dried and stored at −20 °C. For native MS (Chapter 8), His-tagged MCL-1 was used instead of the tag-less version. This construct was simply obtained by skipping the thrombin cleavage step, and performing SEC straight after eluting the protein from the nickel resin.

Because of the presence of an impurity that co-eluted as a shoulder peak during SEC, some of the fractions always had to be discarded. The overall yields were improved by performing an ion-exchange step prior to SEC, which separated this contaminant. After overnight

thrombin cleavage, the proteins were loaded onto a HiTrap SP HP cation-exchange column (GE Healthcare) at flow-rate of 1 mL/min. The column was washed with buffer A (HEPES ion exchange buffer) until conductivity and UV readings returned to baseline. The proteins were eluted with a stepped gradient of buffer B (HEPES ion-exchange buffer + 1 M NaCl; 0–9% B over 20 mL, 9–13% B over 25 mL, 13–20% B over 20 mL). Clean fractions were pooled, and further purified by SEC (*vide supra*).

Identities of the proteins were confirmed by MS (mouse MCL-1: 17966.57(±4.27), theoretical 17965.33 Da; human MCL-1: 18409.52(±4.33), theoretical 18409.74 Da).

3.3.4 BAK and BAX

The constructs of BAK (UniProt:Q16611, residues 16–185) and BAX (UniProt:Q07812, residues 1–171) were designed to match the structural data available for these proteins (Wang *et al.*, 2009, Suzuki *et al.*, 2000). In both cases, the C-terminal transmembrane helix was removed to aid solubility. In the case of BAK, a short disordered segment at the N-terminus was also excluded. None of the cysteine residues being involved in intramolecular disulfide bonds, they were mutated to serines to avoid the use of reducing agents. Synthetic genes were ordered from GenScript and sub-cloned into pTXB1 (New England Biolabs) as described (*vide supra*).

```
Q16611     1  MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM   60
BAK        1  --------------MEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM   46
                            ****************************************

Q16611    61  VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE  120
BAK       47  VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE  106
              ************************************************************

Q16611   121  SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA  180
BAK      107  SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHSIARWIAQRGGWVAA  166
              ********************************************.***************

Q16611   181  LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS                               211
BAK      167  LNLGN--------------------------                               171
              *****
```

Figure 3.5 – Protein sequence of BAK compared to the construct used in this thesis. The cloning artefact (M residue at the N-terminus) is highlighted in bold. The construct also contained the mutation C166S.

Expression and purification protocols for each protein were broadly similar. Differences are indicated in the text where relevant. Disulfide mutants of BAK were prepared according

```
Q07812     1  MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPVPQDASTKKLS   60
BAX        1  MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPVPQDASTKKLS   60
              ************************************************************

Q07812    61  ECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMFSDGNFNWGRVVALFYFASKL  120
BAX       61  ESLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMFSDGNFNWGRVVALFYFASKL  120
              *.************************************************************

Q07812   121  VLKALCTKVPELIRTIMGWTLDFLRERLLGWIQDQGGWDGLLSYFGTPTWQTVTIFVAGV  180
BAX      121  VLKALSTKVPELIRTIMGWTLDFLRERLLGWIQDQGGWDGLLSYFGTPTWQ---------  171
              *****.*******************************************

Q07812   181  LTASLTIWKKMG                                                  192
BAX      172  ------------                                                  171
```

Figure 3.6 – Protein sequence of BAX compared to the construct used in this thesis. The construct had two mutations (C62S and C126S).

to the same protocol as wild type BAK. Plasmids were transformed in BL21(DE3) (BAK) or C41(DE3) (BAX), and the cells were grown overnight at 37 °C on 2×TY-agar plates containing ampicillin. Pre-cultures were made from these plates, and used to inoculate 1 L cultures (LB), which were incubated at 37 °C until an OD_{600} of ∼0.6 was reached. BAK expression was carried out at 37 °C for 4 h following induction with 1 mM IPTG. BAX was induced with 0.1 mM IPTG, and the expression performed overnight at 28 °C. Cells were harvested by centrifugation (5,000 rpm, 15 min, 4 °C), the pellet re-suspended in HEPES buffer, and sonicated on ice (3–5 min, 15 s pulses at 45 s intervals). Debris were cleared by centrifugation (18,000 rpm, 45 min, 4 °C), and the proteins were purified from the soluble fraction by binding to chitin resin (New England Biolabs). Supernatants were passed through the resin at 1 mL/min using a peristaltic pump, followed by washing with HEPES buffer at 2 mL/min (15 CVs). These procedures were performed at 4 °C. Self-cleavage of the inteins were induced by adding HEPES buffer containing 50 mM DTT, isolating the column, and incubating the resin overnight at room temperature. The next day, cleaved proteins were eluted from the column, concentrated, and purified by SEC (Superdex 75 26/600, equilibrated in biophysical phosphate buffer) at an isocratic flow-rate of 2 mL/min. Small amounts of affinity tag always got stripped from the column, and co-eluted with the protein peaks. Therefore, a chitin pull-down was necessary. Fresh resin was prepared in biophysical phosphate buffer, and added to the pooled SEC fractions. The purity of the supernatant was confirmed by SDS-PAGE, filter-sterilised and kept at 4°C.

For long-term storage of BAK and BAX, the proteins were flash-frozen in liquid nitrogen

and kept at −80 °C. Before using these samples, size-exclusion chromatography was used to remove small-scale aggregates that resulted from the freeze-thaw process (especially in the case of BAX). Samples were never frozen more than once.

Identities of the proteins were confirmed by MS (BAK: 19213.01(±8.55), theoretical 19217.64 Da). The mass of BAX was confirmed by native MS, and was found to be within 1 Da of its theoretical value (18933.65 Da).

3.3.5 GB1-X_{BH3}

Some BH3 peptides were produced recombinantly as GB1 fusions (Immunoglobulin G-binding protein G) to aid expression and solubility (Cheng & Patel, 2004). Sequences of ∼35 amino acids encompassing the BH3 motif, and ∼10 flanking residues were N-terminally fused to a His-tagged GB1 *via* a proteolytic cleavage site. The constructs of mouse $PUMA_{BH3}$ and its W133F/N149A mutant (p53 upregulated modulator of apoptosis, UniProt:Q99ML1, residues 127–161) contained a Factor Xa cleavage site, and the M144A mutation that helps reduce the oligomerisation propensity of the peptides (Rogers *et al.*, 2014b). The construct of human $PUMA_{BH3}$ (UniProt:Q9BXH1, residues 125–161) contained a thrombin cleavage site, and the same M144A mutation. The construct of human BAX_{BH3} (UniProt:Q07812, residues 49–83) contained a thrombin cleavage site, the C62S mutation, and two extra residues (GW) at the N-terminus. The tryptophan was added to allow spectroscopic quantifications of the peptide concentration. Sequences of the full constructs are shown in Fig. 3.7.

Expression and purification protocols were broadly similar to those of MCL-1 and spectrin. In brief, expressions were carried out at 37 °C for 4 h in C41(DE3) or C41(DE3)pLysS cells. After sonication, the constructs were bound to nickel resin, and the peptides cleaved off with protease overnight at room temperature. Factor Xa (New England Biolabs) was used to cleave mouse $PUMA_{BH3}$ (and its W133F/N149A mutant) in Factor Xa cleavage buffer. Thrombin was used to cleave human $PUMA_{BH3}$ and BAX_{BH3} in PBS containing 25 mM imidazole (leftover from the nickel resin washes). Peptides were first purified by anion-exchange chromatography (HiTrap Q HP, GE Healthcare) with a 20 mM Tris pH 8.0/1M NaCl buffer system (0–20% over 50 mL, elution at ∼15%). The final purification step was size-exclusion chromatography (Superdex 30, 26/600, GE Healthcare) into biophysical

Mouse $PUMA_{BH3}$
MRGSHHHHHHGLVPRGSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYD
DATKTFTVTENYDIPTSHGIEGR-RVEEEEWAREIGAQLRRAADDLNAQYERRRQEEQH

Human $PUMA_{BH3}$
MRGSHHHHHHGLVGSGSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYD
DATKTFTVTENYDIPTSHGGLVPR-GVRGEEEQWAREIGAQLRRAADDLNAQYERRRQEEQQ

Human BAX_{BH3}
MRGSHHHHHHGLVGSGSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYD
DATKTFTVTENYDIPTSHGGLVPR-GWPVPQDASTKKLSESLKRIGDELDSNMELQRMIAAV

Figure 3.7 – Protein sequences of GB1-X_{BH3} peptide constructs. Products resulting from proteolitic cleavge are underlined. Residues that differ from the gene products (mutations or cloning artefacts) are highlighted in blue. The extra mutations present in the 'spectroscopically silent' version of mouse PUMA are highlighted in yellow.

phosphate buffer, and the peptides were stored frozen (−80 °C) until used.

Identities of the proteins were confirmed by MS (human $PUMA_{BH3}$: 4441.93(±0.50) Da, theoretical 4442.75 Da; BAX_{BH3}: 4113.82(±0.92), theoretical 4114.70 Da). Although nothing could be detected by SDS-PAGE, MS analysis of BAX_{BH3} showed the presence of impurities. Based on their masses, these appeared to be C-terminal truncations of the peptide (−2, and −5, residues respectively). This sample was only used in a couple of experiments.

3.4 Synthetic peptides

Most of the work involving BH3 motifs was actually performed with chemically synthesised versions of the peptides. These were ordered from commercial suppliers (HPLC purity >98%), re-constituted in biophysical buffers from their lyophilised stocks, and used without further purification.

3.4.1 Peptides used in Part I

All experiments involving PUMA that are presented in Part I (except the out-competition dissociation experiments, *vide infra*) were performed with a 34 residue long sequence containing the 15 residues from the BH3 motif of *Mus musculus* (UniProt:Q99ML1, residues 128–161). This peptide was synthesised by Selleck Chemicals, and included the M144I mu-

tation used in the NMR structure (PDB:2ROC) (Day *et al.*, 2008). Termini were protected by N-terminal acetylation and C-terminal amidation.

Ac-VEEEEWAREIGAQLRRIADDLNAQYERRRQEEQH-NH_2

tamra-RVEEEEWAREIGAQLRRAADDLNAQYERRRQEEQH

Figure 3.8 – Sequences of the mouse PUMA peptides used in Part I. The M144 mutations are highlighted, and the BH3 motifs underlined.

Out-competition dissociation experiments were performed with a peptide containing an extra residue at the N-terminus (UniProt:Q99ML1 residues 127–161), which was ordered from Biomatik. It contained the M144A mutation (instead of M144I) that had been shown to reduce oligomerisation of the peptide (Rogers *et al.*, 2014b). It also contained a TAMRA dye at the N-terminus to enable extrinsic fluorescence measurements (*vide infra*).

3.4.2 Peptides used in Part II

BH3 motifs used in Part II of this thesis were ordered from Biomatik as 35 amino acid long peptides composed of the BH3 sequence (15 residues), plus 10 flanking residues on both sides. Each construct also contained a N-terminal 5-carboxytetramethylrhodamine (TAMRA) dye for extrinsic fluorescence experiments (Fig. 3.9). This rhodamine derivative was chosen for its price, and photo-chemical stability (Dempsey *et al.*, 2011). Some kinetic processes required long acquisition times (multiple hours), and more 'standard' fluorophores (*e.g.* FITC) were prone to severe photobleaching over these timescales. It is noted that the presence of dyes can affect binding properties. However, it was shown not to be the case in the context of TAMRA-labelled BH3 peptides binding BCL-2 proteins (Crabtree *et al.*, 2018).

All sequences were based on gene products from *Homo sapiens*: t-$PUMA_{BH3}$ (UniProt: Q9BXH1, residues 127–161, M144A); t-BID_{BH3} (UniProt:P55957, residues 76–110); t-BAK_{BH3} (UniProt:Q16611, residues 64–98, P64A); t-BAX_{BH3} (UniProt:Q07812, residues 49–83, P49G and C62A); t-BIM_{BH3} (UniProt:O43521, residues 138–172); t-MCL-1_{BH3} (UniProt:Q07820, residues 199–233). Sequences of the peptides are shown in Fig. 3.10.

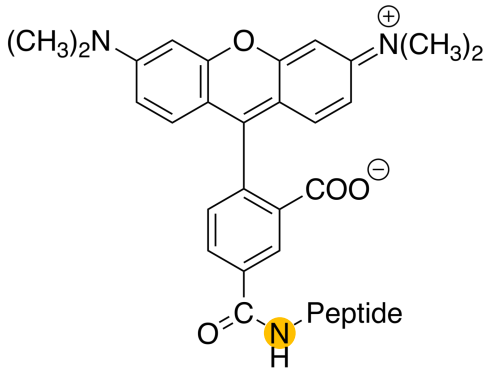

Figure 3.9 – Chemical structure of 5-carboxytetramethylrhodamine (TAMRA). The dye is zwitterionic at neutral pH, and thus does not contribute to the net charge of the peptide. The nitrogen atom of the peptide's N-terminus is highlighted.

```
                                        BH3
t-PUMA_BH3    tamra-RGEEEQWAREIGAQLRRAADDLNAQYERRRQEEQQ
t-BID_BH3     tamra-SESQEDIIRNIARHLAQVGDSMDRSIPPGLVNGLA
t-BAK_BH3     tamra-GLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQ
t-BAX_BH3     tamra-GVPQDASTKKLSEALKRIGDELDSNMELQRMIAAV
t-BIM_BH3     tamra-EPADMRPEIWIAQELRRIGDEFNAYYARRVFLNNY
t-MCL-1_BH3   tamra-MGRSGATSRKALETLRRVGDGVQRNHETAFQGMLR
```

Figure 3.10 – Protein sequences of BH3 peptides used in Part II. Each construct contained a N-terminal TAMRA dye. Residues that are different from the gene products are highlighted in blue.

3.5 Biophysical techniques

3.5.1 Protein concentration

Protein and peptide concentrations were determined spectroscopically by absorbance measurements on a Cary 60 UV-Vis spectrophotometer (Agilent Technologies) using the Beer-Lambert law:

$$\log_{10}\left(\frac{I_0}{I}\right) = A_\lambda = \varepsilon_\lambda \cdot c \cdot l \tag{3.3}$$

where I and I_0 represent the light intensities reaching the detector in the presence and absence of the sample respectively, A_λ the absorbance at wavelength λ, ε_λ the extinction coefficient at the same wavelength (in M^{-1} cm^{-1}), c the protein concentration (in M) and l the optical path length (in cm).

Measurements were performed by scanning the samples between 200–400 nm (or 200-800

nm when TAMRA was present), and buffer-subtracting the results. In case of observable deviations between buffer and sample spectra outside of the chromophore(s) spectral range—indicative of light scattering due to aggregation or large impurities—the samples were filtered, and the measurements repeated. Absorbances were kept within the dynamic range of the instrument, and the sample diluted if necessary. For experiments where the results were directly dependent on the accuracy of the protein concentration measurement (*e.g.* association kinetics), dilution factors were obtained by weighing components on a precision balance, and averaging the results from three independent measurements.

Table 3.3 – Experimentally determined extinction coefficients for the proteins used in this thesis.

	ε_λ / M^{-1} cm^{-1}
ε_{280} ($\alpha0\alpha1$)	18,320
ε_{280} ($\beta16\beta17$)	32,600
ε_{280} (mouse MCL-1)	22,158
ε_{280} (acetylamidated $PUMA_{BH3}$)	7,113
ε_{280} (mouse $PUMA_{BH3}$ W133F/N149A)	1,277
ε_{280} (human MCL-1)	23,782
ε_{280} (BAK)	28,190
ε_{280} (BAX)	32,498
ε_{280} (human $PUMA_{BH3}$)	6,925
ε_{280} (BAX_{BH3})	6,000
ε_{555} (TAMRA)	83,000

All extinction coefficients were experimentally determined, and the values reported in Table 3.3. Values for spectrin proteins were previously determined by Dr Stephanie A. Hill by the method of Gill & von Hippel (1989). Extinction coefficients for mouse MCL-1, acetyamidated PUMA, and mouse PUMA W133F/N149A were determined by Dr Joseph M. Rogers using amino acid analysis (Rutherfurd & Gilani, 2009). Those of BAK, BAX, human MCL-1, human $PUMA_{BH3}$, and BAX_{BH3} were determined using the same method (analyses performed by Peter Sharratt, Department of Biochemistry, University of Cambridge). These experimental values were typically within 10% of the coefficients calculated using ProtParam (http://web.expasy.org/protparam/), which computes ε_{280} as a linear

combination of extinction coefficient for W, Y and disulphide bonds, each being averaged contextual values calculated from a representative data set (Pace *et al.*, 1995).

Concentrations of dye-labelled peptides were obtained by measuring the abosrbance of TAMRA at 555 nm. Its experimental extinction coefficient was determined from a large data set of TAMRA-labelled peptides that were quantified by amino acid analysis, and found to be independent of the protein sequence (Crabtree *et al.*, 2018). The value of the extinction coefficient was sensitive to the self-assembly of the peptide, resulting in a change of the A555/A523 ratio for BH3 motifs with strong oligomerisation propensities. If this ratio became significantly smaller than ∼2.5—indicative of the presence of oligomers—the peptide was diluted, and the concentration of the stock was obtained by back-calculation from the diluted sample measurement.

3.5.2 Chemical denaturation

Equilibrium chemical denaturation experiments were perfomed in the presence of either guanidinium chloride (GdmCl, MP Biomedical) or urea (VWR Chemicals). Denaturant solutions were prepared gravimetrically and volumetrically using stock solutions of buffers, or solid constituants (to avoid dilution of the buffer components). Once made, GdmCl solutions were kept in the dark at 4 °C, and urea solutions were kept frozen at −20 °C. Concentrations of denaturant were checked by measuring the refractive indices of the solutions compared to buffer (at 25 °C), and using Equation 3.4 for GdmCl (Nozaki, 1972), or Equation 3.5 for urea (Warren & Gordon, 1966).

$$[\text{GdmCl}] = 57.147(\Delta h) + 38.68(\Delta h)^2 - 91.60(\Delta h)^3 \tag{3.4}$$

$$[\text{Urea}] = 117.66(\Delta h) + 29.753(\Delta h)^2 + 185.56(\Delta h)^3 \tag{3.5}$$

where Δh is the difference of refractive index between denaturant and buffer solutions at the sodium D line.

Equilibrium denaturation curves were prepared by making 68 samples of increasing denaturant concentration, ranging from 0 to ∼7 M. Buffer/denaturant solutions (800 µL total volume) were precisely dispensed using a Microlab 500 liquid handling unit (Hamilton), followed by the addition of 100 µL of protein solution. Samples were incubated at 25 °C

(2 h to 2 days) before performing measurements. Changes in intrinsic fluorescence and CD signal were used to monitor denaturation, and data were acquired on a Cary Eclipse Fluorescence Spectrophotometer (Varian) and a Chirascan Circular Dichroism Spectrometer (Applied Photophysics) respectively. Both instruments were thermostated at 25 °C, and the samples were measured in parallel. CD signals were recorded at 222 nm, followed by conversion to MRE units. Fluorescence emission spectra were recorded from 300 to 400 nm following excitation at 280 nm, and the data were reported as average emission wavelengths (AEW):

$$\mathrm{AEW} = \frac{\sum \lambda_i F_i}{\sum F_i} \tag{3.6}$$

where F_i is the fluorescence intensity at wavelength λ_i. It is noted that if the quantum yields of the different states are significantly different, the use of AEW is inappropriate, and linear decompositions of the spectra should be used instead. The data for human MCL-1 denaturation curves (Fig. 9.1) were fitted to a 2-state denaturation model (Equation 2.25) in order to extract thermodynamic parameters. Curves for BAK and BAX were not fitted.

3.5.3 Circular dichroism spectroscopy

Circular dichroism spectroscopy was performed on Chirascan instruments (Applied Photophysics) thermostated at 25 °C, unless stated otherwise. Because buffer components (*e.g.* chloride ions) absorb strongly in the far-UV, a delicate balance of protein concentration, path length, and data averaging had to be optimised for most experiments. When possible, conditions were chosen that gave ellipticities (at 222 nm) around −10 mdeg or more. Raw signals were buffer-subtracted, and converted to MRE values using the following relationship:

$$(\mathrm{MRE\ /\ deg\ cm^2\ dmol^{-1}}) = \frac{(\Theta/\mathrm{mdeg})}{(l/\mathrm{mm}) \cdot (c/\mathrm{M}) \cdot n} \tag{3.7}$$

where Θ represents the buffer-subtracted signal in mdeg, l is the optical path length (in mm), c the concentration (in M), and n the number of amino acids in the protein. MRE values represent CD 'extinction coefficients' that are further corrected for the number of CD 'chromophores', *i.e.* amide bonds. This allows for a direct comparison between different proteins, as it reports the spectrum *per* residue in a given protein context.

Estimation of helicities form MRE values at 222 nm were obtained by using the method of Muñoz & Serrano (1995), which consists of a linear interpolation between parameters representing a pure α-helix, and a pure coil:

$$\%\text{Helicity} = 100 \cdot \left(1 + \frac{(\text{MRE}_{222} - \text{MRE}_{\text{helix}})}{(\text{MRE}_{\text{coil}} - \text{MRE}_{222})}\right)^{-1} \tag{3.8}$$

where $\text{MRE}_{\text{helix}}$ is the value of a fully α-helical structure at 25 °C ($-35{,}791$ deg cm^2 dmol^{-1}), and MRE_{coil} represents the value of a pure coil at the same temperature (-725 deg cm^2 dmol^{-1}). These parameters are temperature-corrected values of a pure helix and a pure coil at 0 °C (described in the paper).

Structural spectra:

Spectra were typically recorded from 200 to 260 nm every 1 nm, with an averaging time per point of 5–20 s (adaptive sampling), and a bandwidth of 1–2 nm. Optical path lengths were usually between 2–4 mm, and protein concentrations between 1–20 μM.
In order to assess the change of structural content upon binding, the arithmetic average of the scans of each protein in isolation was compared to the spectrum of the complex, highlighting any gain or loss of structure. The average was calculated according to the following equation:

$$\text{MRE}_{\text{average}} = \frac{\sum c_i n_i \text{MRE}_i}{\sum c_i n_i} \tag{3.9}$$

where $\text{MRE}_{\text{average}}$ represents the arithmetic average, MRE_i the MRE value of protein i on its own, n_i its length, and c_i its concentration (in the mixture). It is noted that because MRE values are used, averages for any hypothetical set of concentrations may be calculated. Thus, structural spectra of the mixtures did not have to be obtained with protein concentrations matching that of their individual scans.

Oligomerisation kinetics:

The oligomerisation kinetics of BAK and BAX, induced by either GdmCl or detergent, were monitored by following the change of CD signal at 222 nm (2 nm bandwidth). The sampling interval was adapted for the length of the reaction in order to optimise the signal-to-noise ratio. The reactions were initiated by adding detergent or denaturant form concentrated stocks directly into the cuvette containing proteins, mixing by inverting the sample a few

times, and starting data acquisition. The dead time incurred from this manual mixing was taken into account by correcting the timebase during data analysis.

Temperature denaturation:

For temperature denaturation, cuvettes were sealed with Teflon caps to avoid evaporation of the solutions. Typical protein concentrations were 1–5 μM for a path length of 2 mm. The temperature was raised in steps, and scans (1 nm data intervals, 2 s adaptive sampling) were acquired every 2(±0.2) degrees. Data were usually acquired between 10–90 °C, and spectra were collected during both heating and cooling stages.

3.5.4 Fluorescence spectroscopy

Intrinsic and extrinsic fluorescence measurements were acquired on Cary Eclipse Fluorescence Spectrophotometers (Agilent). Samples were held in either a 1 mL cuvette (for low concentrations) or a 60 μL cuvette. The excitation and emission slit widths, the voltage applied across the photomultiplier tube, and the time-averaging per point, were modified in parallel depending on the protein concentration, and the requirement of the experiment. This was done in order to stay within the dynamic range of the instrument, optimise the signal-to-noise ratio, and reduce photobleaching.

Fluorescence emission spectra:

Intrinsic fluorescence emission spectra (for chemical denaturation curves) were obtained by exciting at 280 nm and recording between 300–400 nm at 1 nm intervals. Scans were typically acquired at 200 nm/min. For these experiments, the emission slit width was kept as small as possible to avoid spectral overlap between adjacent wavelengths.

Equilibrium binding:

Equilibrium binding dissociation constants between BAK and BH3 motifs were obtained by measuring the fluorescence anisotropy of TAMRA-labelled peptides as a function of BAK concentration. Experiments were performed by using a manual polarisation accessory, exciting the dye 555 nm and recording the fluorescence at 575 nm. Fluorescence anisotropy relates to the tumbling of the dye in solution, which provides a proxy for binding; a small peptide on its own tumbles faster than when bound to a larger protein. The fluorescence

anisotropy of TAMRA-labelled peptides was determined experimentally by quantifying the ratio of the polarised component to the total fluorescence intensity of the dye (Lakowicz, 2006):

$$R = \frac{I_{\mathrm{VV}} - GI_{\mathrm{VH}}}{I_{\mathrm{VV}} + 2GI_{\mathrm{VH}}} \tag{3.10}$$

where R is the fluorescence anisotropy, I_{VV} and I_{VH} are the vertically- and horizontally-polarised components of the fluorescence intensity, and $G = I_{\mathrm{HV}}/I_{\mathrm{HH}}$ is a correction factor that accounts for the instrument's differential detection sensitivities in the vertical and horizontal polarisation planes (where I_{HV} and I_{HH} have analogues meanings to I_{VV} and I_{VH}).

Binding induces changes in the intensity of the polarised component of the fluorescence, but it might also have an impact on the total fluorescence of the dye. This would affect the calculation of R, which would impact the shape of the binding curve, and thus the value of K_{d} obtained from fitting. In order to circumvent this problem, the change in total fluorescence intensity upon binding needs to be taken into account, and the anisotropy corrected accordingly (Dandliker *et al.*, 1981):

$$R_{\mathrm{corr}} = \frac{x \cdot R_{\mathrm{bound}} \cdot (I_{\mathrm{free}}/I_{\mathrm{bound}}) + R_{\mathrm{free}}}{1 + x \cdot (I_{\mathrm{free}}/I_{\mathrm{bound}})} \tag{3.11}$$

where R_{corr} represents the corrected anisotropy, R_{free} and R_{bound} the anisotropy of free and bound peptide respectively, I_{free} and I_{bound} the total fluorescence intensities of the free and bound states respectively ($I = I_{\mathrm{VV}} + 2I_{\mathrm{VH}}$), and $x = (R - R_{\mathrm{free}})/(R_{\mathrm{bound}} - R)$.

Binding curves were prepared by serial dilutions of BAK in the presence of a constant concentration of dye-labelled peptide (typically 1 μM). BAK was concentrated to ∼1 mM (3,000 MWCO centrifugal filter units, Millipore), and its concentrations precisely determined. Serial dilutions (1:1 volumes) were performed by hand without changing the dials on the pipette in order to reduce errors associated with inaccurate volume dispensing. Samples were transferred to a 60 μL cuvette, and the fluorescence intensity resulting for each four polarisation setting measured in triplicate, and averaged. These readings were converted to anisotropy values as described above. Equilibrium dissociation constants (K_{d}) were obtained by fitting the corrected anisotropy signal as a function of BAK concentration to a 2-state hetero-dimerisation model (Equation 2.30).

Oligomerisation kinetics:

BAK and BAX detergent-induced oligomerisation kinetics were monitored by excitation at 280 nm, and measuring changes in fluorescence intensity at 330 nm. Proteins in buffer were placed in the cuvette, and the reactions were initiated by addition of detergent from 10× stocks using a pipette. The solutions were mixed by pipetting up and down, and the measurements started. The dead time between the addition of detergent and the start of the data acquisition were accounted for by modifying the timebase during analysis. Data were fitted to the sum of two exponential decay functions (Chapter 7 and 8).

Out-competition dissociation kinetics:

Dissociation rate constants for complexes between BH3 motifs and MCL-1 (mouse and human) were obtained by performing out-competition experiments. Complexes of MCL-1 and TAMRA-labelled BH3 peptides were pre-formed by mixing equimolar amounts of the components at typical concentrations of 5 μM. A solution of the out-competitor (mouse or human PUMA_{BH3}, un-labelled) was placed in a fluorescence cuvette, and irreversible dissociation was initiated by making a 10-fold dilution of the complex into the un-labelled BH3 peptide solution using a manual pipette. The sample was mixed by inverting the cuvette a few times, and the measurement was started. This method incurred dead times of ∼20 s, which were taken into account during analysis. Data were collected by exciting TAMRA at 555 nm and measuring the change in fluorescence intensity at 575 nm. In certain cases, loss of fluorescence polarisation (obtained using a manual fluorescence polarisation accessory set to V/V) gave a better signal-to-noise ratio, and was used instead of fluorescence intensity. Data were fitted to a single exponential decay function. In certain cases, a linear drift term was added to account for photo-bleaching (indicated in the text where appropriate). Experiments were repeated in the presence of different excess quantities of out-competitor (50–400 molar excess over complex). Irreversible dissociation was confirmed by the absence of dependence of k_{obs} on the fold-excess of competitor. Concentration-independent values were averaged, giving k_{off}.

In Chapter 6, dissociation rate constants for un-labelled BH3 peptides were also obtained. Changes in intrinsic, instead of extrinsic, fluorescence intensities were monitored by exciting at 280 nm, and recording at 330 nm. For these experiments, the 'spectroscopically silent' version of mouse PUMA_{BH3} (W133F/N149A) was used instead of the wild type (Rogers

et al., 2014b). This peptide does not give an appreciable signal change upon binding, thus only the dissociation event is detected. Moreover, the W133F mutation reduces its fluorescence, which decreases background signal, and improves the signal-to-noise ratio of the experiment.

3.5.5 Stopped-flow kinetics

For fast reactions, stopped-flow kinetic measurements were performed. These were recorded on either a SX18 or SX20 stopped-flow spectrophotometer (Applied Photophysics). Measurements temperature were 25 °C, unless stated otherwise. Reactions were monitored by either following the change in intrinsic fluorescence (exciting at 280 nm, and using a 320 nm longpass filter), or extrinsic fluorescence (TAMRA, exciting at 555 nm, and using a 570 nm longpass filter). In some cases, the signal change observed from fluorescence intensity was poor, and fluorescence anisotropy was recorded instead (using a FP1 fluorescence polarisation accessory, Applied Photophysics). Drive ramp pressures were typically set to 3 bar, and the first 5 ms of the measurements were discarded due to mixing artefacts. For fast reactions (<2 s), a pressure hold was employed, which reduced the dead time to 1 ms. Data were acquired on a linear timebase, adjusted for the length of the reaction ($\sim 10 \cdot t_{1/2}$ for first order processes), and 2,000 points were typically collected for each trace. Association reactions were performed using 1:1 volume mixing, and out-competition dissociation experiments were done using 1:10 mixing.

Association reactions:

Spectrin association was monitored by change in intrinsic fluorescence over 1000 s, and 10–15 traces were recorded for each condition. The traces were individually fit to a near-equimolar, reversible, bimolecular, 2-state model including a linear drift term (Equation 2.48). Most experiments were performed with final concentrations in the range of 5–10 μM.

PUMA:MCL-1 association was monitored over 2–20 s (adjusted for the slowing-down in association rate with increasing ionic strength). For each condition, 20–30 traces were collected, averaged, and fitted using a near-equimolar, irreversible, bimolecular, 2-state model including a linear drift term (Equation 2.49). Final concentrations were typically in

the range of 0.5–1 μM.

PUMA:MCL-1 dissociation rate constants were also indirectly obtained from association reactions under reversible conditions (Appendix C). For these experiments, binding was performed under low-nM concentrations (∼20 nM) of MCL-1 and acetylamidated PUMA. The data was fitted to the near-equimolar, reversible binding model Equation 2.48) with k_{on} fixed to the value obtained under irreversible conditions (to increase the robustness of the fitted parameter K_{d}). The dissociation rate constants were obtained using $k_{\mathrm{off}} = k_{\mathrm{on}} \cdot K_{\mathrm{d}}$, which is a valid assumption for a 2-state reaction.

Association of MCL-1 and BAK with BH3 motifs (Chapter 6) were obtained under pseudo-first order conditions. For these reactions, TAMRA-labelled peptides were typically at concentrations of 100–500 nM (final), and the partner protein was present in at least a 10-fold excess. Under these conditions, reactions fitted to a single exponential decay function. The gradient of the line between k_{obs} and the concentration of excess partner gave k_{on}. Some association rate constants between MCL-1 and BH3 motifs were obtained using unlabelled peptides (and monitoring intrinsic fluorescence). These reaction were performed under reverse-pseudo-first order conditions (peptide in excess), and also fitted to a single exponential decay function.

Dissociation reactions:

For some out-competition dissociation reactions (*vide supra*), the rates were too fast to be measured on a fluorimeter. For these instances, stopped-flow kinetics using 1:10 mixing was performed. General principles were identical to the experiments performed by manual mixing and using the fluorimeters. The only difference was the reduction in dead time, thus allowing fast dissociation events to be observed. Reactions were monitored by change in intrinsic fluorescence intensity.

3.5.6 Cross-linking

Chemical cross-linking:

Detergent-treated, and GdmCl-treated BAK/BAX oligomers were cross-linked with mixtures of EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, Thermo Scientific) and BS3 (bis(sulfosuccinimidyl)suberate, Thermo Scientific). Different protein : EDC : BS3 ratios were employed, and these are indicated in the text where appropriate.

Cross-linker stock solutions were prepared in biophysical phosphate buffer, and used directly. Proteins were pre-incubated in detergent or GdmCl, and cross-linkers added from 10× stocks, followed by incubation at room temperature for 2 h. Samples were mixed with NuPAGE loading buffer, heat-denatured, and analysed by SDS-PAGE. In the case of denaturant-induced oligomerisation, overnight dialysis (3,500 MWCO Slide-A-Lyzer® Mini Dialysis Unit, Thermo Scientific) was performed to remove GdmCl, before adding loading buffer, and running the gel.

Disulfide cross-linking:

Oxidation of BAK cysteine double mutants (Chapter 7) was performed using the redox catalyst CuPhe (Kobashi, 1968). Stock solutions (25 mM) were added to protein solutions at a final concentration of 0.5 mM, followed by incubation on ice for 30–45 min. The reactions were quenched by addition of EDTA (2 mM final concentration), followed by overnight dialysis at room temperature (3,500 Da MWCO Slide-A-Lyzer® Mini Dialysis Unit, Thermo Scientific). Results were analysed by SEC, and control experiments were obtained by reducing the respective disulfide mutant with 50 mM DTT for 10 min at room temperature. The oxidation step was performed either before, or after, the addition of detergent (details in the text).

3.5.7 Analytical size-exclusion chromatography

Analytical size-exclusion chromatography experiments were performed on an ÄKTA FPLC system (GE Healthcare), using a Superdex™ 200 Increase 10/300 GL column (GE Healthcare). The column was equilibrated in biophysical phosphate buffer (2 CVs), except for some initial detergent screenings (ammonium acetate buffer, Chapter 7), and the denaturant SEC experiments (biohysical phosphate buffer + x M GdmCl, Chapter 9). Elutions were performed at 0.75 mL/min, except for the 'GdmCl capture' experiments (1.4 mL/min, Chapter 9). Injection volumes were between 100–500 µM, with typical protein concentrations being in the range of 10–20 µM. Elutions chromatograms were recorded by measuring the absorbance at 280 nm through a 0.5 cm flow-cell.

The calibration of the column was performed using chromatographic standards (Kostanski *et al.*, 2004). The void volume was determined by injecting Blue Dextran (Sigma) at 1

mg/mL, and was found to be 8.15 mL. The relationship between elution volume and protein size was obtained by injecting a mixture of globular protein standards (GE Healthcare): Thyroglobulin (669,000 Da); Ferritin (440,000 Da); Aldolase (158,000 Da); Conalbumin (75,000 Da); Ovalbumin (44,000 Da); Carbonic anhydrase (29,000 Da); Ribonuclease A (13,700 Da); 3 mg/ mL each. The result of the calibration is shown in Fig. 3.11, and the correlation is described by Equation 3.12, where $V_{elution}$ is the elution volume of the peak maximum (in mL).

$$\log_{10}(\mathrm{MW/Da}) = -0.1942 \cdot V_{elution} + 7.5685 \tag{3.12}$$

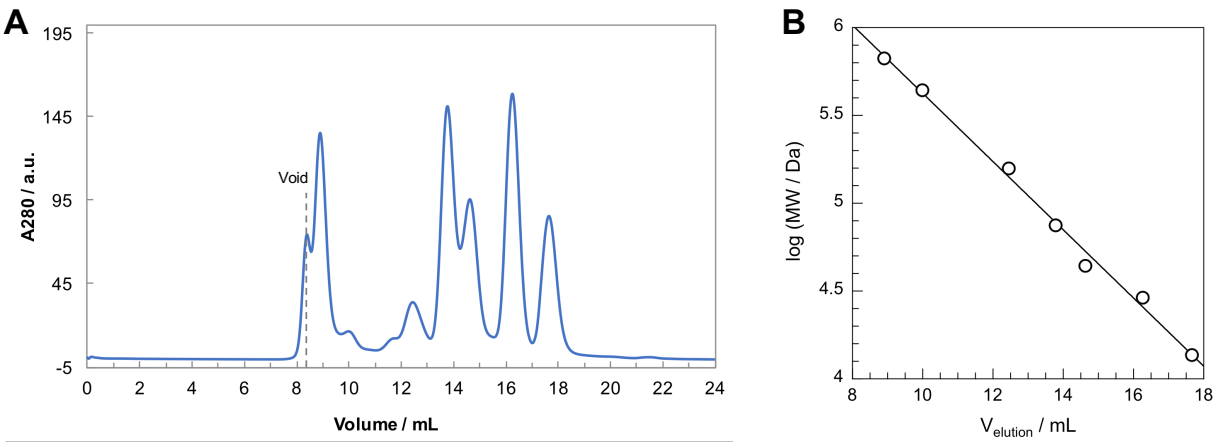

Figure 3.11 – Calibration of Superdex 200 10/300 column. (**A**) Elution of the mixture of protein standards. (**B**) Correlation plot. Details are described in the text.

3.5.8 Native mass spectrometry

Samples for native mass spectrometry were prepared in Cambridge. Proteins were expressed and purified as described (*vide supra*). Monomers (in biophysical phosphate buffer) and lyophilised peptides were transported (refrigerated) to the laboratory of Prof. Carol V. Robinson (University of Oxford). Protein stocks were buffer-exchanged to ammonium acetate buffer on-site using size-exclusion chromatography (Superdex 75 10/300).

Mixtures of proteins were assembled in buffer to final concentrations of 5 µM (total, equimolar distributions), before adding PS20 (5×CMC). Oligomerisation was left to proceed at room temperature for at least 3 h before measuring the samples. All spectra were acquired on a Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific) at a HCD cell pressure of 7 mL/min of argon, and a HCD energy of 40 eV. The data were anal-

ysed with Xcalibur (Thermo Scientific). All experiments were performed with Dr Kallol Gupta, who also did the final data analysis.

3.6 Data analysis

3.6.1 Fitting

Experimental data were fitted to appropriate physical models by non-linear least squares minimisation routines. These were performed with either KaleidaGraph (Synergy Software), pro Fit (QuantumSoft), or Wolfram Mathematica (Wolfram Research). For non-trivial functions, values of the parameters used during initialisation were varied to assess the robustness of the fit, and confirm the convergence of the model.

Where possible, data from repeats were individually fitted, and the parameters averaged, allowing to obtain standard deviations and standard error of the means. Alternatively, data sets were averaged prior to fitting, and uncertainties represent fitting errors.

3.6.2 Uncertainties

Errors associated with the values reported in this thesis were either fitting errors, standard deviations, or standard error of the means. These are indicated in the text as appropriate. Statistical parameters were estimated assuming that the measurements were normally distributed around the mean. Standard deviations (σ) were calculated using the Bessel correction:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{3.13}$$

where N represents the size of the sample, and $\overline{x}$ is the mean of the data set. From this, standard error of the means were obtained by:

$$\sigma_{\text{mean}} = \frac{\sigma}{\sqrt{N}} \tag{3.14}$$

When parameters needed to be calculated from algebraic manipulations of values having associated errors, uncertainties were propagated using the relationships described in Table 3.4.

Table 3.4 – Propagation of uncertainties

Relationship	Error in the result
$X = A \pm B$	$\delta X = \sqrt{(\delta A)^2 + (\delta B)^2}$
$X = cA$	$\delta X = c\delta A$
$X = c(A \cdot B)$ or $X = c(A/B)$	$\delta X = X\sqrt{(\delta A/A)^2 + (\delta B/B)^2}$
$X = cA^n$	$\delta X = X \cdot \lvert n \rvert \cdot (\delta A/A)$
$X = \ln(cA)$	$\delta X = (\delta A/A)$
$X = \exp(cA)$	$\delta X = X \cdot c \cdot \delta A$

3.6.3 Protein structure visualisation

Protein structures were visualised using PyMOL (Schrödinger) using atomistic coordinates available from the Protein Data Bank (https://www.rcsb.org). PDB codes of the relevant structures are cited in the text where appropriate.

Part I

Effect of solution conditions on protein-protein interactions involving disordered partners

Chapter 4

Electrostatics in coupled folding and binding

Some of the results presented in this Chapter were published in: Wicky, B. I. M., Shammas, S. L. & Clarke, J. (2017). Affinity of IDPs to their targets is modulated by ion-specific changes in kinetics and residual structure. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 9882–9887.

4.1 Introduction

Intrinsically disordered proteins (IDPs) and proteins with intrinsically disordered regions (IDRs) constitute a large proportion of the proteome, especially of eukaryotic organisms (Wright & Dyson, 1999, Dunker *et al.*, 2001, Ward *et al.*, 2004, Tompa, 2011, van der Lee *et al.*, 2014a, Wright & Dyson, 2015). These disordered regions are characterised by a lack of a uniquely-defined structure, instead populating many near-isoenergetic conformations (Eliezer, 2009, Chebaro *et al.*, 2015). Despite their structural heterogeneity, IDPs are functional and involved in numerous cellular tasks (Uversky *et al.*, 2008, Wright & Dyson, 2015). Coupled folding and binding reactions—where an IDP folds upon binding to its target protein—constitute an important class of protein-protein interactions (PPIs). Compared to PPIs involving folded partners, these reactions possess an additional folding dimension. Moreover, their disordered nature and marginal stabilities make their structural ensembles particularly susceptible to changes in solution conditions. For example, it has been demonstrated that changes in solvent excluded volume and ionic strength can significantly affect the radius of gyration of disordered proteins (Müller-Späth *et al.*,

2010, Soranno *et al.*, 2014). Considering these extra aspects specific to PPIs involving disordered partners, factors affecting affinities and lifetimes are yet to be completely understood (Gibbs & Showalter, 2015). Much of the early work in the field has focused on the protein (sequence) determinant of these reactions (Iešmantavičius *et al.*, 2014, Rogers *et al.*, 2014b,a, Toto *et al.*, 2014, Toto & Gianni, 2016). However, the role of environment (solution) conditions on coupled folding and binding has largely been ignored in biophysical studies, despite the established effect on IDP structural ensembles (Müller-Späth *et al.*, 2010, Soranno *et al.*, 2012, Rogers *et al.*, 2013, Dogan *et al.*, 2015).

4.2 Aims

The role of electrostatics is of particular interest in the context of PPIs involving disordered partners. At the sequence level, IDPs often show compositional biases; prolines, glycines and charged residues tend to be over-represented compared to their distributions in folded proteins (Uversky *et al.*, 2000, Romero *et al.*, 2001). Moreover, the sequence complexity of disordered proteins is usually lower. These features—in particular the over-representation of charged residues—raises questions about the role of electrostatics in coupled folding and binding reactions. For example, in the context of IDPs, changing ionic strength of the solution may affect long-range electrostatics and chain collapse; both of which could have consequences on binding kinetics. These questions are investigated in Part I of this thesis. In this Chapter, the dependence of k_{on} on the ionic strength of the solution was measured for two model systems. Estimations of the basal rate constants in the absence of charge-charge interactions were obtained by applying a Debye-Hückel-like formalism. By comparing these values with kinetic results obtained at low, or physiological ionic strengths, the role of electrostatics in coupled folding and binding reactions could be evaluated. Ionic-strength dependencies were obtained by NaCl titrations; the 'canonical salt' for this type of studies. In the next Chapter, this work was extended to other types of ions.

4.3 Deconvoluting electrostatics experimentally

Despite the apparent complexity of coupled folding and binding reactions, the 2-state approximation—where the partners exist either as separate species, or in complex—has been demonstrated for range of systems. For such cases, the affinity of the bound complex

is determined simply by the ratio of its dissociation and association rate constants; $K_{\mathrm{d}} = k_{\mathrm{off}}/k_{\mathrm{on}}$. For reactions having early transition states, dissociation rate constants are predominantly determined by the energetics of the contacts formed in the bound structure. These are determined by the protein sequences, which also influences association rate constants. However, k_{on} may also be affected by solution conditions.

4.3.1 The association rate constant

For two proteins modelled as hard spheres that freely diffuse in solution, the bimolecular association rate constant may be expressed by the Smoluchowski-Stokes equation:

$$k_{\mathrm{on}} = \frac{2RT}{3\eta}\frac{(r_A + r_B)^2}{r_A r_B} \tag{4.1}$$

where r_A and r_B are the radii of the two spheres, T the thermodynamic temperature, R the gas constant, and η the viscosity of the solution. For particles of the size of typical globular proteins in water at 25 °C, $k_{\mathrm{on}} = 10^9$–10^{10} $\mathrm{M^{-1}\ s^{-1}}$ (Berg & von Hippel, 1985). This range constitutes the 'diffusion-limit' for species that are reactive over their whole surfaces, *i.e.* every collision leads to binding. However, not all encounters are productive. Indeed, protein molecules might only be able to form complexes if they collide with the right orientations. This anisotropic reactivity can be taking into account by adding a scaling factor $k_{\mathrm{on}}^{\mathrm{corr}} = A \cdot k_{\mathrm{on}}$. From theoretical studies, a revised diffusion-limit of 10^5–10^6 $\mathrm{M^{-1}\ s^{-1}}$ is often quoted for folded proteins (Berg, 1985, Zhou, 1993)

These estimations put an upper-bound limit for diffusion-limited processes that are energetically unbiased. However, for proteins in solution, inter-molecular forces will influence their encounters (Schreiber *et al.*, 2009). These 'steering forces' depend on the interaction potential between the partners. Using approximations from molecular mechanics, the inter-molecular potential energy can be described as the sum of a Coulombic term (V_{el})—to model charged interactions—and a Lennard-Jones potential (V_{vdW})—to describe non-charged interactions:

$$V_{\text{inter-molecular}} = V_{\mathrm{el}} + V_{\mathrm{vdW}} = \frac{Q_A Q_B}{4\pi\epsilon r} + \sigma\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right] \tag{4.2}$$

where Q_A and Q_B are the charges of the interacting species, ϵ is the permittivity of the medium, r the separation distance, σ the depth of the energy well, and r_0 the van der

Waals contact distance. Because of the presence of these inter-molecular forces, proteins in solution do not freely diffuse. The extent to which the diffusion is biased depends on the separation distance between the interacting partners, and the solution conditions. Non-charged interactions (van der Waals interactions) are 'short-range' because $V_{\text{vdW}} \propto 1/r^6$ (*N.B.* the other term represents repulsions from orbital overlap, and is even shorter-range). Therefore, their contributions to the formation of the initial encounter complex will be negligible. They are only going to modulate k_{on} by affecting interactions when the proteins are already in contact. On the other hand, Coulombic interactions are 'long-range' because $V_{\text{el}} \propto 1/r$. Thus, electrostatic interactions may play a significant role in the modulation of the association rate constant; by steering the partners towards each other (Schreiber *et al.*, 2009).

4.3.2 The Debye length

The magnitude of electrostatic interactions is dependent on both the separation distance between the partners, and the composition of the interstice. For a continuous medium, this is described by the dielectric constant in the calculation of the Coulombic term (*cf.* Equation 4.2). However, proteins in solution are screened by solvent molecules, and ions from buffer components. This affects the effective distance over which electrostatic interactions may occur. This effective distance is described by the Debye length (κ^{-1}), which represents the mean radius of the ionic atmosphere around a charged species (Girault, 2007). This quantity is important to understand electrostatic interactions in solution, as it represents the screening distance—the distance beyond which charge-charge interactions become negligible. The Debye length is expressed by the following relationship:

$$\kappa^{-1} = \sqrt{\frac{\epsilon_r \epsilon_0 k_B T}{2000 N_A e^2 I}} \tag{4.3}$$

where ϵ_r is the relativity permittivity of the medium, ϵ_0 the permittivity of free space, k_B the Botzmann constant, T the thermodynamic temperature, N_A the Avogadro number, e the elementary charge, and I the ionic strength expressed in mol/L:

$$I = \frac{1}{2} \sum_i c_i z_i^2 \tag{4.4}$$

where c_i is the molar concentration of ion i, and z_i its net charge. The Debye length

as a function of ionic strength for a range typical of biochemical conditions is shown in Fig. 4.1. It is evident that changing the ionic strength of a solution (*e.g.* by adding salt to the buffer) will significantly impact the distance over which proteins may experience electrostatic forces (Fig. 4.1B). This might have consequences for their association kinetics, and thus the affinity of their interactions.

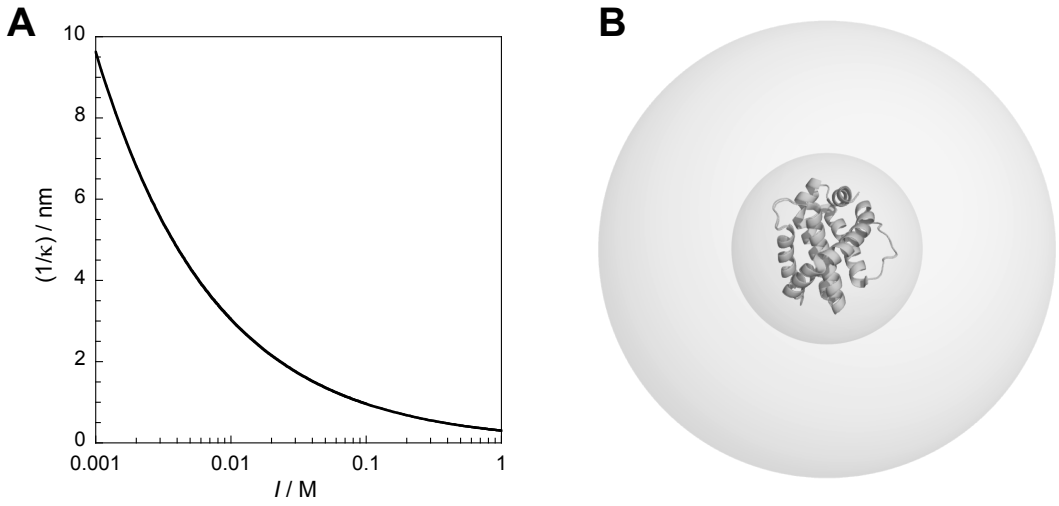

Figure 4.1 – Debye length as a function of ionic strength (**A**) Plot of the Debye length between 1–1000 mM ionic strength. (**B**) Debye length for 4 mM (outer sphere) and 100 mM (inner sphere) plotted onto MCL-1 (PDB:1WSX).

4.3.3 Association rate constants in the absence of electrostatics

For the interaction between folded proteins, it has been shown that the effect of the ionic strength of the solution on the association rate constant could be captured by a Debye-Hückel-like formalism (Schreiber & Fersht, 1996, Vijayakumar *et al.*, 1998). This finding suggests that the rate-limiting step is correlated with the electrostatic potential between the proteins, and that this potential is altered by the ionic strength. Using this theoretical framework, a basal rate constant (the association in the absence of long-range electrostatic interactions, $k_{\text{on}}^{I=\infty}$) can be estimated. Here, a re-arranged version of the equation proposed by Vijayakumar *et al.* (1998) was used:

$$\ln k_{\text{on}} = \ln k_{\text{on}}^{I=\infty} - \frac{AB\sqrt{I}}{Bd \cdot (Bd + \sqrt{I})} \tag{4.5}$$

with:

$$\begin{cases} AB = e\frac{Q_A Q_B}{k_B T \epsilon}\sqrt{\frac{8\pi N_A}{k_B T \epsilon}} \\ Bd = e\sqrt{\frac{8\pi N_A}{k_B T \epsilon}}d \end{cases} \tag{4.6}$$

where AB, Bd, and $\ln k_{\text{on}}^{I=\infty}$ are the free-fitting parameters. In these equations, Q_A and Q_B represent the charges of the proteins, d is the separation distance of the encounter complex, k_{B} is the Boltzmann constant, N_{A} the Avogadros number, T the temperature, e the elementary charge, and ϵ the permittivity of water ($\epsilon = \epsilon_0 \cdot \epsilon_r$). The AB term contains information about the product charge of the complex, and Bd relates to the approach distance of the encounter complex.

It is noted that the Debye-Hückel treatment of electrolytes is only valid for cases where the interactions between ions are weak when compared to the thermal energy available (Girault, 2007). In practice, this usually means very dilute solutions (a few millimolar at most). However, the work presented here investigated salt concentrations up to 1 M (and in the work by Schreiber & Fersht (1996), up to 2 M NaCl were employed). Since most conditions were outside of the regime for which the assumptions underlying the Debye-Hückel theory are valid, fitted parameters have to be interpreted with care. Although the separation distance and the product charge obtained from fitting appeared physically reasonable—*e.g.* values for d were in the single-digit Å range—these parameters were not analysed. Instead, this model was only used for extrapolating association rate constants to infinite ionic strength. By comparison with experimental values, the effect of charged interactions under specific conditions was estimated.

4.4 Model systems

For this study, two model coupled folding and binding reactions were investigated. These systems were chosen for their contrasting biophysical properties, and because they had already been extensively characterised in the group. Affinities of the complexes, association rate constants, topology of the bound state, and structure of the transition state were all different. Thus, it was rationalised that observations made for both systems would be a consequence of their disorder; providing generalisable insights into the consequences of disorder on protein-protein interactions.

4.4.1 The spectrin tetramerisation domain

Spectrin proteins are long, elongated, multi-domain proteins having important structural roles. In particular, the erythrocytic isoform is an important component of the structural network giving red blood cells their elastic properties (Baines, 2009). The fundamental unit of structure of any spectrin protein—the 'spectrin domain'—is a 106-residue long domain that folds into a three-helix bundle (named A, B and C, with C forming a continuous, single helix with A of the next domain). Those helical bundles are then arranged in tandem to form multi-domain proteins (Speicher & Marchesi, 1984). Erythrocyte spectrin exists in two forms: the α chain, which is composed of 20 full domains (α1–α20) preceded by one partial, disordered N-terminal domain (α0), and the β chain, which possesses 16 full spectrin domains (β1–β16) followed by one partial, unstructured C-terminal domain (β17). The unstructured parts of each chain can dimerise to form a new, non-covalent, spectrin topology termed the 'tetramerisation domain' (*N.B.* in spite of its name, this assembly is actually dimeric, Fig. 4.2) (Speicher *et al.*, 1993). The spectrin dimers associate laterally in an anti-parallel fashion to from a tetramer composed of two α chains and two β chains, which supports red blood cell membranes.

The formation of the tetramerisation domain involves IDRs from both interacting partners that fold upon binding. Thus, this association can be considered a coupled folding and binding reaction (Shammas *et al.*, 2012). Mutations in this region have been linked to important hematological phenotype changes (Gaetani *et al.*, 2008). Studies of this process are typically done on shortened version of the spectrin proteins that abolish lateral association. However, α0 and β17 (the two partial domains of each chain) do not associate by themselves and need to be flanked by one adjacent (fully folded) domain each (Dr Stephanie A. Hill, unpublished results). In the context of kinetic studies of the tetramerisation reaction, constructs α0α1 (UniProt:P02549, residues 2–163) and β16β17 (UniProt:P11277, residues 1898-2083) were designed, which still associate and were used as minimal surrogates for the study of full-length spectrin association (Ipsaro *et al.*, 2010).

Despite the topological complexity of the bound state, the tetramerisation reaction is consistent with a 2-state mechanism when monitored by change in intrinsic fluorescence (Shammas *et al.*, 2012). This was confirmed by performing the reaction under pseudo-first-order conditions with α0α1 in excess. The data fitted a single exponential decay function, consistent with a reaction possessing a single transition state (data not shown).

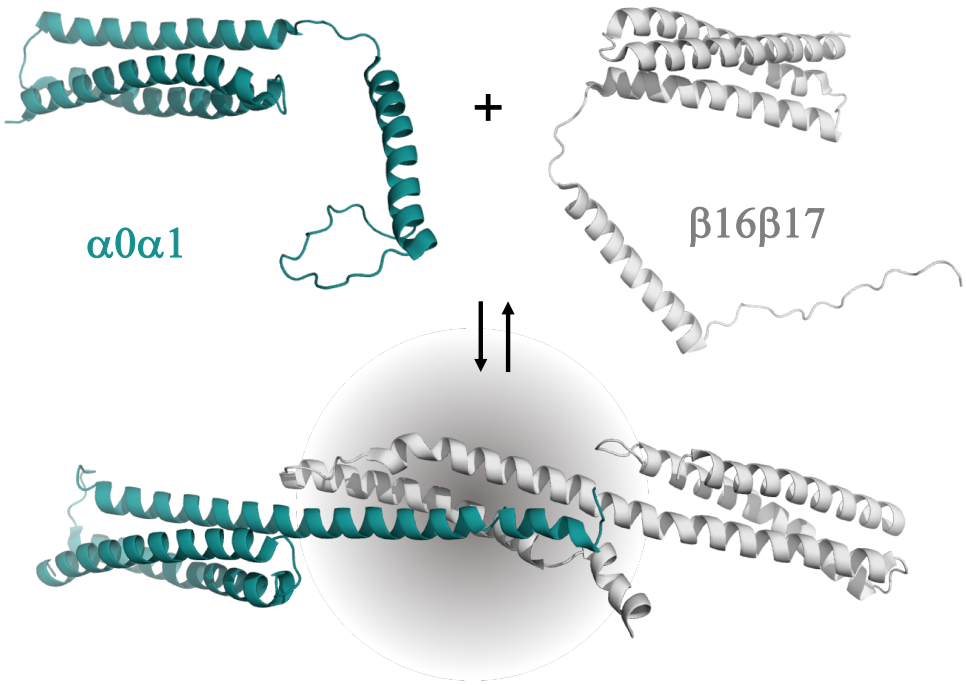

Figure 4.2 – Coupled folding and binding of spectrin. In isolation, both spectrin chains have a folded domain (three-helix bundle) adjacent to a disordered segment. Upon binding, these segments fold to a spectrin domain topology (the 'tetramerisation domain', highlighted by the shaded area). The structure of the complex is based on PDB:3LBX. The representation of α0α1 is based on one of the ten structure of its NMR ensemble (PDB:1OWA). No structure of β16β17 exists. Its depiction is purely for visualisation purposes (based on PDB:1OWA).

All subsequent work was performed by near-equimolar mixing and fitting the data to a reversible bimolecular binding model (*cf.* Chapter 2). This strategy allowed k_{on} and K_{d} (and hence k_{off}) to be extracted from measurements performed at one set of concentrations.

The association of spectrin domains is surprisingly slow (6.3 10^2 M^{-1} s^{-1}, (Shammas *et al.*, 2012)). However, since the dissociation rate constant is slow as well (2.6 10^{-4} s^{-1}), the affinity of the complex remains sub-μM (0.4 μM, (Mehboob *et al.*, 2003, Gaetani *et al.*, 2008, Shammas *et al.*, 2012)). Interestingly, the activation energy for the binding reaction (E_a = 11.3(±0.1) kcal mol^{-1}) was found *not* to be significantly different from a system with a very fast association rate constant (*cf.* Appendix A). Therefore, the slow association kinetics observed for spectrin does not seem to be due to a large activation energy barrier. The nature of the transition state has been probed by Φ-value analysis of the domains undergoing coupled folding and binding (Hill *et al.*, 2014). This work revealed a significantly structured transition state; partially helical segments of β17 dock onto a mostly pre-formed α0. Together with the lack of a large activation barrier, these results suggest

that the slow association of spectrin domains is probably a consequence of a conformational selection mechanism. It is noted that while these observations would be consistent with such a scenario, only a formal kinetic analysis could distinguish between induced fit and conformational selection (Gianni *et al.*, 2014, Shammas *et al.*, 2016).

4.4.2 The PUMA:MCL-1 system

Induced myeloid leukemia cell differentiation protein (MCL-1) and p53 up-regulated modulator of apoptosis (PUMA) are two members of the B-cell lymphoma 2 (BCL-2) family (for recent reviews, see Czabotar *et al.* (2014), Kale *et al.* (2017)). These proteins are key regulators of the intrinsic pathway of apoptosis (*cf.* Part II). MCL-1 belongs to the anti-apoptotic class. These proteins have a distinct globular fold comprised of seven amphiphatic α-helices packing around a central hydrophobic ones (Petros *et al.*, 2004). The resulting structures have a hydrophobic groove on their surfaces, which constitute the site of interaction for pro-apoptotic BCL-2 proteins (Sattler *et al.*, 1997). Murine MCL-1 is 331 amino acid long, and contains a 151 amino acid long disordered N-terminus not present in other BCL-2 proteins. It also contains a 23 amino acid C-terminal membrane anchoring sequence. These segments were not included in the construct used in the work presented in this thesis; only the globular BCL-2 fold (UniProt:P97287, residues 152–308) was studied, in line with previous work (Rogers *et al.*, 2013, 2014b,a).

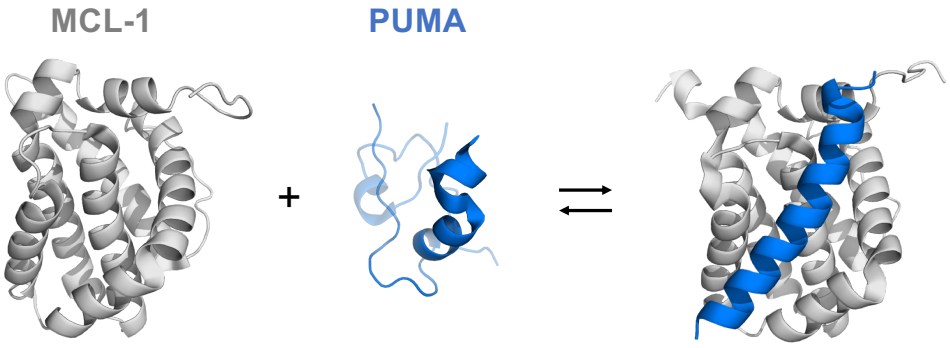

Figure 4.3 – Coupled folding and binding of PUMA with MCL-1. PUMA in isolation is disordered, and folds to a contigous α-helix upon binding to the surface of MCL-1. In contrast, MCL-1 is folded on its own (PDB:1WSX) and undergoes minimal structural change upon forming the complex (PDB:2ROC, RMSD = 1.3 Å over 142 Cα).

PUMA belongs to the pro-apoptotic class of BCL-2 proteins. These members are mostly composed of disordered proteins. PUMA is 193 amino acid long, of which approximately 30 can fold upon binding BCL-2 partners. For the work presented in this chapter, a 34 residue-

long sequence containing the 15 residues from the BH3 motif of PUMA from *Mus musculus* (UniProt:Q99ML1, residues 128–161) was used instead of the full-length IDP. This sequence included the M144I mutation used in the NMR structure of the PUMA:MCL-1 complex (PDB:2ROC, (Day *et al.*, 2008)).

Earlier work in the group has shown that binding of PUMA to MCL-1 is a fast process ($1.6\ 10^{7}\ M^{-1}\ s^{-1}$, (Rogers *et al.*, 2013)). Despite its fast association, it was demonstrated that this binding is *not* diffusion-limited, and does possesses an activation energy. With a dissociation rate constant of $1.6\ 10^{-3}\ s^{-1}$, the affinity of the resulting complex is very tight ($K_{\mathrm{d}} = 0.1$ nM). This coupled folding and binding reaction also displayed 2-state kinetics. Therefore, all data was collected under second-order conditions, and the data fitted to an irreversible binding model (*cf.* Chapter 2, valid for reactions performed at concentrations largely exceeding K_{d} , where the contribution from k_{off} can be ignored). By performing a 'proline-scan' of PUMA to disrupt the helicity of the free IDP, it was shown that residual structure was not a key determinant of the association rate constant (Rogers *et al.*, 2014b). Moreover, Φ-value analysis of the binding reaction revealed a mostly unstructured transition state, with the exception of some slight helicity at the N-terminus of PUMA (Rogers *et al.*, 2014a). Together, these results strongly support the notion that the binding of PUMA to MCL-1 is (mostly) an induced-fit process.

The biophysical characteristics of these two model systems under physiological-like conditions are summarised in Table 4.1.

Table 4.1 – Thermodynamic, kinetic and mechanistic signatures of the two model systems under physiological-like conditions. These data are taken from previously published work (references in the text).

	k_{on} / M^{-1} s^{-1}	k_{off} / s^{-1}	K_d	TS (from Φ-value analysis)	Conditions
α0α1:β16β17	6.3×10^2	2.6×10^{-4}	0.4 μM	Transiently helical β17 docks onto a more structured α0	25 °C 50 mM Na-PO_4, 150 mM NaCl pH 7 $I \sim 250$ mM
PUMA:MCL-1	1.6×10^7	1.6×10^{-3}	0.1 nM	Mostly unstructured with few native interactions	25 °C 50 mM Na-PO_4 pH 7 $I \sim 100$ mM

In addition to their different biophysical signatures, these two systems also have different number of charged residues. This might be expected to affect the electrostatic component of their interactions. At neutral pH, both spectrin proteins are negatively charged (-7 and -5 for α- and β-spectrin respectively). In contrast, MCL-1 is predicted to be positively charged ($+1$), and PUMA negatively (-4). Statistics on the number and type of charged residues can be found in Appendix B. The distribution of Asp/Glu and Arg/Lys were plotted onto the structures of the bound structures of spectrin and PUMA:MCL-1 for visual representation (Fig. 4.4).

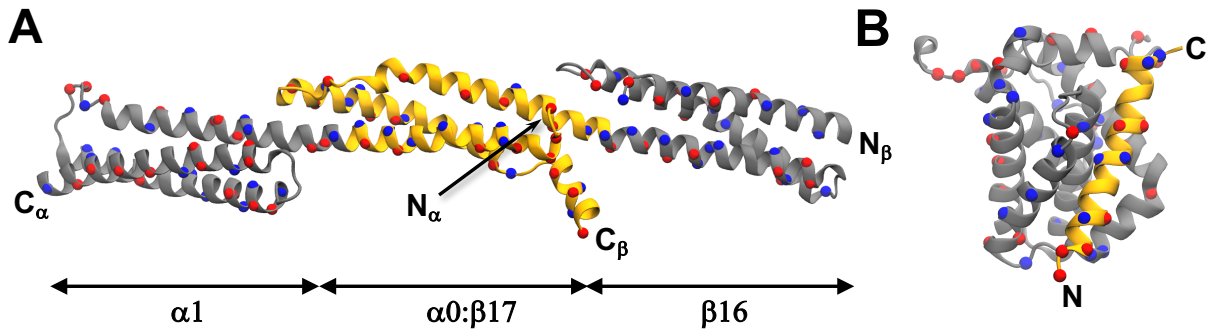

Figure 4.4 – Charged residues present in the model systems. The domains/proteins that are folded in isolation are depicted in grey. The parts that are disordered in isolation are depicted in gold. Asp/Glu and Lys/Arg residues are coloured onto their Cα (represented as spheres) in red and blue respectively. (**A**) Structure of the spectrin tetramerisation domain (α0:β17) flanked by its respective folded domains α1 and β16 (PDB:3LBX). (**B**) Structure of the PUMA:MCL-1 complex (PDB:2ROC).

4.5 Electrostatics only marginally affects association

For this study, NaCl was used to screen charge-charge interactions, as it is the 'canonical salt' for this type of studies. The association rate constants of spectrin and PUMA:MCL-1 were measured between 4 mM—no added salt, ionic strength contributed exclusively by the buffer—and 1 M ionic strength. Over that range, the effect of salt concentration on viscosity was negligible, and could therefore be neglected. The data were fitted to the Debye-Hückel-like model (Equation 4.5), and the results are presented in Fig. 4.5. The natural logarithm of the association rate constants were plotted against the inverse square-root of the ionic strength, which directly relates to the Debye length under these conditions. The more intuitive scale of ionic strength increases from right to left, and the intercept with the y-axis represent infinite ionic strength. Thus, the intercept of the fit represents

(the natural logartihm of) the basal rate constant.

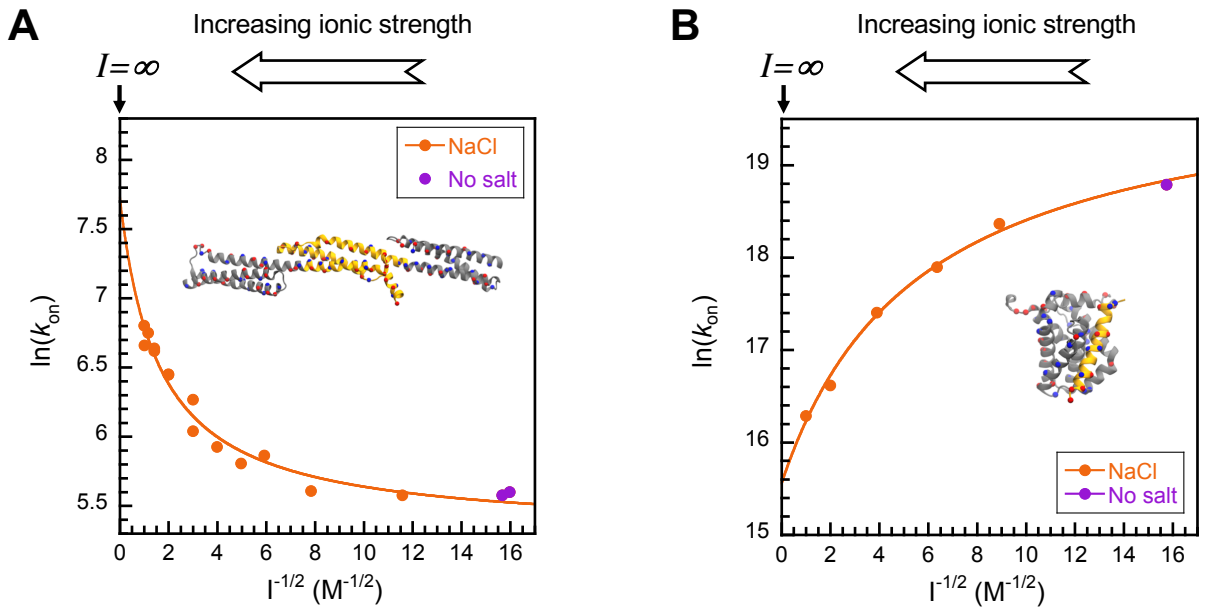

Figure 4.5 – Ionic strength dependence of the association rate constants (k_{on}) modulated by NaCl for spectrin (**A**) and PUMA:MCL-1 (**B**). Increasing salt concentration accelerated spectrin association but reduced the speed of PUMA binding MCL-1. Solid lines are fits to the Debye-Hückel-like model (Equation 4.5), where the intercept represents the association at infinite ionic strength. The ionic strength of the buffer without salt was 4 mM

Table 4.2 – Association rate constants at selected ionic strengths. Errors for $I = 4$ mM represent standard error of the mean. Errors for $I = \infty$ were propagated from the fitting error of $\ln k_{on}^{I=\infty}$.

	k_{on} / M^{-1} s^{-1}	
	$I = 4$ mM	$I = \infty$
Spectrin	$2.13(\pm0.03)\ 10^2$	$2.3(\pm0.6)\ 10^3$
PUMA:MCL-1	$1.45(\pm0.04)\ 10^8$	$6(\pm1)\ 10^6$

These results showed that the association of spectrin is slowed down by electrostatic repulsion (Fig. 4.5A). Indeed, binding became faster as the ionic strength of the solution was increased, indicating that NaCl screened repulsive long-range charge interactions. However, the contribution of these repulsive charge-charge interactions to the association rate constant at 4 mM ionic strength was only modest (~10-fold, *cf.* Table 4.2).

This unfavourable electrostatic effect might be expected from the knowledge of the overall net charges of the proteins; both spectrin are negatively charged at the pH of these exper-

iments. Interestingly, these repulsive charge-charge interactions were observed despite the presence of electrostatically complementary binding interfaces in the bound structure (Fig. 4.6). This observation highlights the importance of long-range electrostatics (considering overall net charges) over local ones. Indeed, if the binding interfaces were the determinant of the electrostatic component of k_{on}, the opposite trend with respect to the ionic strength dependence would have been expected. However, it noted that this interpretation is not that straightforward in the context of IDPs. In isolation, disordered proteins sample many conformations, and binding interfaces do not exist until the bound structure is formed. Only once coupled folding and binding has occurred to these patches form.

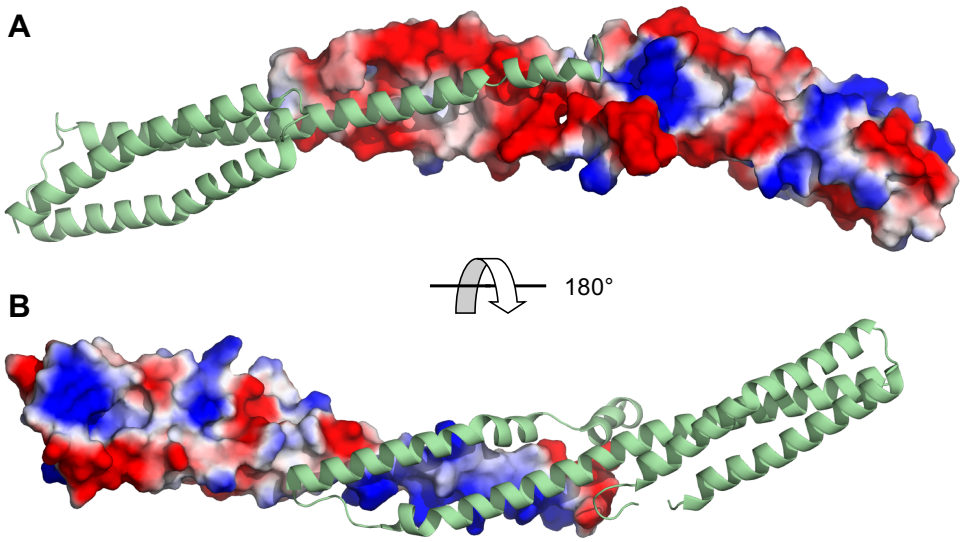

Figure 4.6 – Relative electrostatic potential contact maps for spectrin proteins. (**A**) Map of β16β17 with α0α1 displayed in cartoon representation. (**B**) Map of α0α1 with β16β17 in cartoon representation. Structures based on PDB:3LBX, and electrostatic maps generated with PyMOL using a continuum Poisson-Boltzmann model. The colour scale ranges from −54 to +54 (red and blue respectively, in units of k_BT/e).

The magnitude of the association rate constant between PUMA and MCL-1 under physiological ionic strength might suggest binding with a strong electrostatic component. Perhaps surprisingly, the opposite was found to be true. Indeed, binding of the two proteins was only accelerated ∼25-fold in the absence of salt (*cf.* Table 4.2). Thus, the fast association kinetics observed for this system was not a consequence of k_{on} having a large electrostatic steering component. This appears to be in spite of PUMA and MCL-1 having opposite charges, and their binding interfaces showing electrostatic complementarity (Fig. 4.7). By analogy to the argument made for spectrin, it appears that interface complementarity does not necessarily result in large electrostatically-accelerated association rate constants. Once

again, the lack of pre-formed interfaces in isolation might be the reason behind the lack of enhancement.

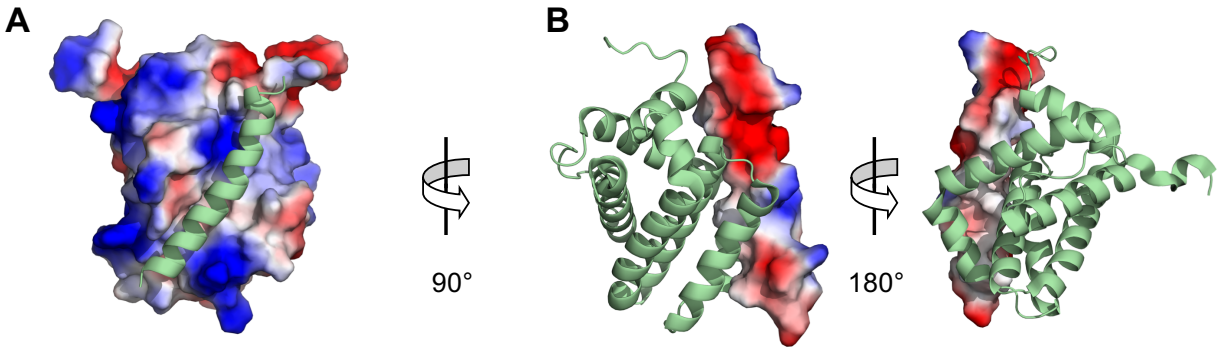

Figure 4.7 – Relative electrostatic potential contact maps for MCL-1 and PUMA proteins. (**A**) Map of MCL-1 with PUMA displayed in cartoon representation. (**B**) Map of PUMA with MCL-1 in cartoon representation. Structures based on PDB:2ROC, and electrostatic maps generated with PyMOL using a continuum Poisson-Boltzmann model. The colour scale ranges from -54 to $+54$ (red and blue respectively, in units of k_BT/e).

4.6 Discussion

In this chapter, the effect of electrostatics on the association rate constant of two IDP systems having very different kinetic, thermodynamic and mechanistic signatures was investigated. Under physiological-like conditions, spectrin associates slowly, with extensive structure present at the transition state (Shammas *et al.*, 2012, Hill *et al.*, 2014), while the PUMA:MCL-1 complex is formed rapidly with a mostly unstructured transition-state (Rogers *et al.*, 2013, 2014a). They also proceed through different mechanisms; the binding between PUMA and MCL-1 is consistent with a (mostly) induced fit process, while spectrin association contains significant elements of conformational selection.

IDPs generally contain a higher proportion of charged residues than folded proteins (Uversky *et al.*, 2000, Romero *et al.*, 2001, Weathers *et al.*, 2004, Lise & Jones, 2005). This sequence-level bias, as well as the patterning of charges, has been shown to be important in dictating the overall geometrical features of IDPs (Müller-Späth *et al.*, 2010, Mao *et al.*, 2010, Das & Pappu, 2013). However, less is known about its impact on the kinetics of coupled folding and binding reactions. Surprisingly—and despite their marked differences in binding affinities, net charges, and number of charged residues (*cf.* Appendix B)—long-range electrostatics only marginally contributed to the association rate constants of

these reactions. In both cases, the effect was about an order of magnitude (*cf.* Table 4.2). This observation contrasts with the enhancement reported for some PPIs involving folded proteins, where charge-charge interactions have been shown to provide important contributions to their association rate constants (2-5 orders of magnitude) (Wallis *et al.*, 1995, Mei *et al.*, 1996, Schreiber & Fersht, 1996, Darling *et al.*, 2002). Given that convergent results were obtained for two contrasting systems, it is tempting to speculate that weak electrostatic enhancement might be a common feature of PPIs involving IDPs. Literature reports revealed similarly small effects (≤ 1 order of magnitude) for other coupled folding and binding systems, thus providing additional support for this hypothesis (Goldberg & Baldwin, 1998, Shammas *et al.*, 2013, Dogan *et al.*, 2015).

The notion that there might be a fundamental difference in 'electrostatic sensitivity' between disordered and folded systems is elegantly supported by a recent study from Papadakos *et al.* (2015). By introducing a single, non-charged, point mutation in the core of E3 rRNase, the authors were able to transform this folded protein into an IDP that retained binding to its partner (Im3). The consequence of introducing disorder was a drastic reduction in affinity, which was almost entirely accounted for by a reduction in k_{on}. Importantly, the association rate constant became almost insensitive to ionic strength. The affinity of the doubly folded system changed by 3 orders of magnitude when the NaCl concentration was varied between 20–500 mM. In stark contrast, the affinity of the disordered variant changed by a mere 20-fold over the same ionic strength range. These results provide strong support for the hypothesis that protein-protein interactions involving disordered partners cannot benefit from the same electrostatic rate-enhancement available to folded systems.

The reason behind this effect is unclear, but some speculations may be formulated. Disordered proteins lack a structure in isolation. Thus, binding interfaces found in the bound state do not exist prior to the formation of the complex. This lack of pre-formed (complementary) binding interfaces—and the existence instead of rapidly re-configuring ensembles—might preclude directional steering of the partners. This would be expected to prevent electrostatically-enhanced association. Indeed, differences in charge complementarity have been reported to account for as much as 3 orders of magnitude to the association rate constants of various cytokine proteins (Pang *et al.*, 2011). It is noted that the recent study by Borgia *et al.* (2018) on proTα and its partner H1 presented a system with extreme ionic strength dependence despite involving disordered partners. While it is to be seen whether this effect is accounted for by modulation of the association rate

constant, the structural differences of this system should be highlighted. Indeed, unlike the coupled folding and binding reactions investigated here, the complex between H1 and proTα remains completely disordered in the bound state. Thus, this system should be under no constraints of complementary 'patches' to determine its affinity.

Despite apparently not being able to benefit from electrostatic acceleration, many coupled folding and binding reactions are fast events (Shammas *et al.*, 2012). It is possible that the reduced orientational constraints required for productive encounters of disordered systems might counter-balance this 'deficiency', at least for induced-fit processes. Thus, it is interesting to speculate that, while PPIs involving folded or disordered partners may both achieve fast association, the physical origin of these rates might be different. This could provide alternative ways for modulating complex formation, which—given the abundance of IDPs in signalling systems—may have regulatory consequences.

Chapter 5

The role of ion-types

Some of the results presented in this Chapter were published in: Wicky, B. I. M., Shammas, S. L. & Clarke, J. (2017). Affinity of IDPs to their targets is modulated by ion-specific changes in kinetics and residual structure. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 9882–9887.

5.1 Introduction

Intrinsically disordered proteins (IDPs) are characterised by a lack of defined structure. Upon interacting with certain partner proteins, some IDPs undergo coupled folding and binding. These reactions are *de facto* protein-protein interactions (PPIs), but the presence of an additional folding dimension may impart new consequences and/or advantages to this type of biomolecular interactions. Productive binding events might only occur with specific conformational state of the IDP—the conformational selection mechanism—or instead, any collision might lead to folding on the surface of the partner protein—the induced-fit mechanism. These two descriptions only represent extrema, and a whole spectrum of reactions with elements of both is in theory possible. Modulations of either the folding or the binding dimensions are expected to impact the overall kinetic profile, and thus the affinity, of PPIs that involve IDPs. The ramifications of this mechanistic plasticity are yet to be fully deciphered.

Unlike folded proteins, IDPs populate many ensembles of rapidly inter-converting conformations with marginal structural stabilities (Jensen *et al.*, 2014, Chebaro *et al.*, 2015). These transient structures are much more sensitive to solution conditions than ordered proteins. It has been demonstrated that the geometric properties of disordered proteins

are affected by changes in solution conditions; ionic strength (Müller-Späth *et al.*, 2010) and solvent excluded volume (Soranno *et al.*, 2014) can influence the extent of chain collapse. In fact, IDPs are best described by concepts from polymer physics (Hofmann *et al.*, 2012), and the knowledge acquired about PPIs involving folded proteins might only be partially transferable.

The sensitivity of disordered proteins to solution conditions raises questions about its implications for coupled folding and binding reactions. In the previous chapter, the role of ionic strength was investigated. Despite the charged nature of IDPs, it was discovered that electrostatics did not significantly contribute to the association rate constant; a possible consequence of disorder. In order to perform this analysis, sodium chloride was used to screen charge-charge interactions in solution, which is the standard approach for this type of studies. However, since ionic strength may affect the ensemble properties of IDPs, is the observed effect purely electrostatics, or does it have structural elements as well? The predictor AGADIR (Lacroix *et al.*, 1998a) for example, shows a reduction of helicity for PUMA with increasing ionic strength (Fig. 5.1). These differences might directly impact how it interacts with MCL-1. This re-modelling of the energy landscape of the IDP by subtle changes in its environment calls for a more careful examination of the influence of the context on coupled folding and binding events.

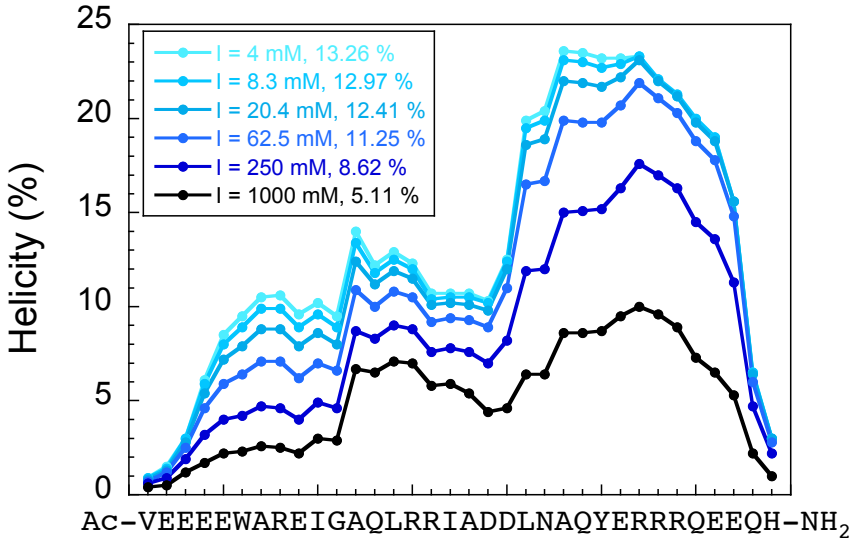

Figure 5.1 – Prediction of PUMA helicity as a function of ionic strength. AGADIR plots of PUMA showing residue-level and overall helicities (indicated in the legend). The peptide shows a reduction in helicity with increasing ionic strength (values of I matched that of the experiments from Fig. 4.5).

5.2 Aims

Compared to their folded counterparts, intrinsically disordered proteins are clearly more sensitive to changes in solution conditions. The consequences of this 'environmental sensitivity' on the interaction between IDPs and their partners were investigated in this chapter. Kinetic and structural techniques were employed to study the effect of charged co-solutes on the model coupled folding and binding systems presented in Chapter 4. Surprsingly, it was found that association and dissociation rates—and thus the affinity of the complex—were ion-type dependent, and not a simple consequence of ionic strength. It was discovered that these effects were rooted in the structural sensitivity of the IDP, and that it followed the Hofmeister series. Incidentally, this behaviour allowed a deconvolution of electrostatics from structural effects in the PUMA:MCL-1 system, providing new insights into the elements contributing to the binding of this system.

5.3 Beyond electrostatics: ion-specific effects

Ionic strength is, by definition, assumed to be independent of the nature of the ion beyond its charge. As part of the Debye-Hückel formalism described in the previous chapter, it is also implicitly assumed that ions in solution only affect reaction kinetics by screening charge-charge interactions. However, it is also possible that they influence the structural ensemble of the IDP. Sodium chloride was used for screening, as it is normally considered the 'canonical' salt for that kind of studies. If the electrostatic modulation obtained with NaCl does indeed represent 'pure' electrostatic effect, then changing the screening agent to another salt should have no impact on the result. To test this hypothesis in the context of coupled folding and binding reactions, one ion-type was systematically varied, while keeping the counter-ion constant. Chloride salts of monoatomic cations were chosen to avoid possible consequences arising from the specific geometries of polyatomic ions. Experiments focused on the biologically-relevant cations K^+, Na^+, (Li^+), Mg^{2+}, Ca^{2+} to study both 1:1 and 1:2 electrolytes. All experiments were performed between 4 mM (no salt added, contribution from the buffer only) and 1 M ionic strength. It is noted that due to the squared contribution of the charge of the species to the ionic strength (*cf.* Equation 4.4), only a third of the salt concentration of 1:2 electrolytes (*e.g.* $CaCl_2$) is necessary to achieve to same ionic strength as 1:1 electrolytes (*e.g.* NaCl).

5.3.1 Different salts affect rates of complex formation beyond ionic strength effects alone

The association rate constants for spectrins and PUMA:MCL-1 were measured under the same range of ionic strengths used in Chapter 4, but in the presence of different salts. The results are shown in Fig. 5.2, and the data for NaCl reproduced for comparison. Similarly to the effect of sodium chloride, the association rate constants between spectrins became faster with increased ionic strength, and binding between PUMA and MCL-1 became slower. However, the results clearly did not overlay, and the association rate constants appeared ion-specific. This suggests that the presence of salts gives rise to more than electrostatic screening. Consequently, using the Debye-Hückel-like model to fit the ionic strength series for each salt yielded different basal rate constants. This confirms that the observed effect is more than the result of screening charge-charge interactions, as otherwise the basal rate constants would have been expected to converge. The discrepancy between salts was largest for the highest ionic strengths studied, indicating concentration-dependent effects. Importantly, however, systematic deviations were observed for values as low as I = 10 mM, highlighting the extreme sensitivity of these systems to changes in solution conditions.

The effect of the different salts on each system showed comparable trends. The 2+ ions led to the largest change in rate constants between 4 mM and 1 M ionic strength, despite being present at lower concentrations (∼333 mM, since 1 M ionic strength is achieved with a third of the salt concentration). Broadly speaking, sodium and potassium gave rise to the smallest modulation in association rates, and lithium's effect was intermediate. This is most clearly seen for the PUMA:MCL-1 system (Fig. 5.2B). The effect was substantial, with the largest difference (KCl *vs* $CaCl_2$) being about 3-fold at the same ionic strength (1 M), and more substantial still if simply considering concentration, since the concentration of calcium and magnesium ions are 1/3 of the 1+ ions before normalising for ionic strength.

For PUMA:MCL-1 the nature of the anion was also systematically varied. There was a clear difference between each salt at 1 M ionic strength. The order $Cl^- < Br^- < I^-$ is illustrated in Fig. 5.2B. NaI led to a larger change in association rate constant than any of the divalent cations, clearly highlighting that the charge state is not an accurate predictor for rationalising the effect of the different salts.

These results demonstrate that the concept of ionic strength alone cannot explain the ob-

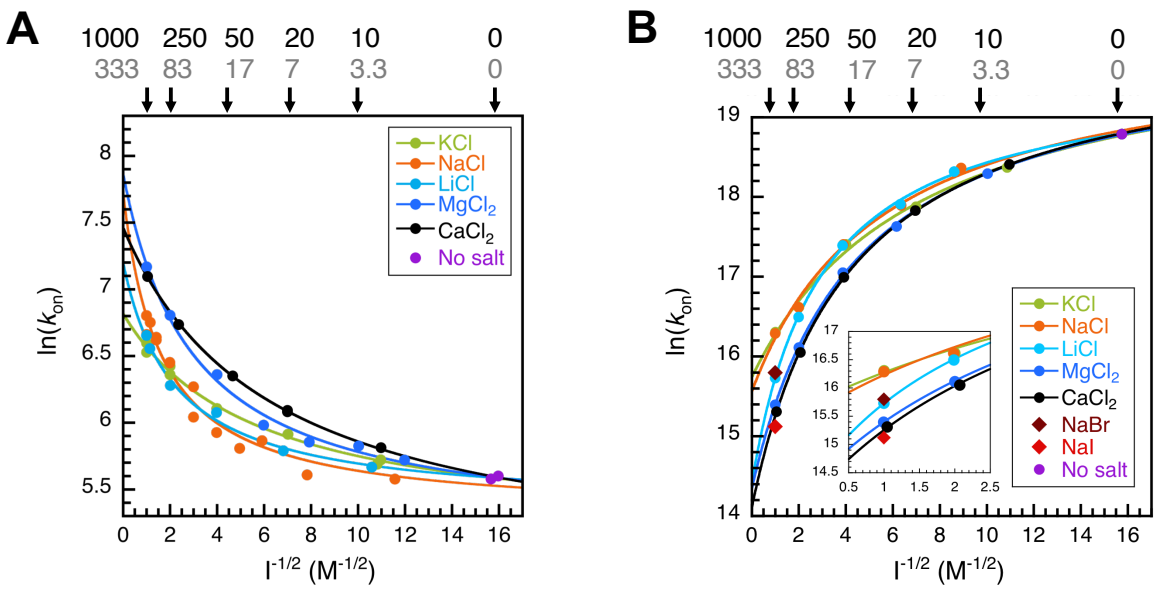

Figure 5.2 – Association kinetics under a range of ionic strengths and salt types. Solid lines represent fits to the Debye-Huckel-like model (Equation 4.5). (**A**) Association kinetics for spectrin. (**B**) Association kinetics for PUMA:MCL-1. The inset highlights the ion-specificity of the association rate constant at 1 M ionic strength. The value measured in the absence of added salt (magenta, $I = 4$ mM) is common to all fitted lines. Some values of salt concentrations (in mM) corresponding to the x-scale are indicated above each plot for reference. Values for 1:1 and 1:2 electrolytes are shown in black and grey respectively.

served outcomes, as the plots were all normalised for it. Moreover, the basal rate constants were different. If the effect of each salt was through screening electrostatic interactions, then the rates constants at infinite ionic strength should have converged.

5.3.2 Different salts also modulate the rate of complex dissociation

The fact that association rate constants were modulated by the addition of salts beyond their impacts on ionic strength raises questions about their potential effect on the lifetimes of the bound state. The unimolecular nature of complex dissociation implies no long-range electrostatic steering. Thus, modulations of k_{off} would confirm ion-specific effects of a different nature to ionic strength. This hypothesis was tested by performing out-competition dissociation experiments of PUMA:MCL-1 in the presence of each salt ($I = 1$ M), and in buffer only ($I = 4$ mM, Table 5.1). It is noted that k_{off} was obtained with a slightly different PUMA peptide than the one used for the association reactions. Out-competition experiments required the presence of a dye (attached to the N-terminus) for

monitoring the dissociation. In order to exclude that the results reported in Table 5.1 were due to this modification, an alternative set of experiments using the dye-free version of PUMA were performed (*cf.* Appendix C). Although the absolute rates were slightly different, both the ion-specificity and the trend were conserved, hence confirming that the results were not a consequence of the modification of the peptide.

Table 5.1 – Kinetic and thermodynamic parameters for PUMA binding MCL-1 in the presence of different salts (at 1 M ionic strength) and no salt conditions. k_{on} were obtained from irreversible association experiments between MCL-1 and acetylamidated PUMA. k_{off} were obtained from out-competition dissociation experiments of t-PUMA:MCL-1 complex. Affinity constants were calculated by $K_d = k_{off}/k_{on}$. Errors represent standard error fo the mean, which were propageted in the case of K_d.

	$k_{on} \times 10^6$ / M^{-1} s^{-1}	$k_{off} \times 10^{-3}$/ s^{-1}	K_d / nM
No salt	145(±4)	1.5(±0.1)	0.011(±0.001)
KCl	12.1(±0.4)	2.42(±0.06)	0.20(±0.01)
NaCl	11.9(±0.2)	2.4(±0.2)	0.21(±0.02)
LiCl	6.8(±0.4)	2.9(±0.2)	0.43(±0.04)
$MgCl_2$	4.9(±0.3)	4.0(±0.4)	0.82(±0.09)
$CaCl_2$	4.5(±0.3)	5.29(±0.03)	1.19(±0.08)
NaBr	7.3(±0.3)	5.7(±0.2)	0.79(±0.04)
NaI	3.7(±0.1)	–	–

As with the association experiments, ion-specific changes in the rate of complex dissociation were observed. The trend for the different salts was identical to that for association, and there was an inverse correlation between the association and dissociation rates, *i.e.* the slower the complex forms, the faster it dissociates. As for the association, the largest change in complex dissociation (KCl *vs* $CaCl_2$) was significant, and amounted to ∼2-fold. Ionic strength also had an impact. The lifetime of the complex was longer in buffer-only than in any of the 1 M conditions. However, the effect was much smaller than on association. This is expected given that dissociation reactions—unlike bimolecular reactions—do not involve long-range electrostatic interactions that could be screened.

Interestingly, the inverse correlation between k_{on} and k_{off} means that the ion-specific effects compound in terms of binding affinity. Indeed, K_d values were shifted by more than if only an effect on the association or dissociation rate constants were observed (Table 5.1 and

Fig. 5.3). Taking potassium and calcium at 1 M ionic strength as an example, there is a ∼3-fold difference in k_{on} and a ∼2-fold difference in k_{off}, which results in a 6-fold shift in affinity. It is emphasised that this effect is purely due to the nature of the ion since the results are within the same ionic strength, therefore excluding long-range electrostatic effects as the origin of these differences.

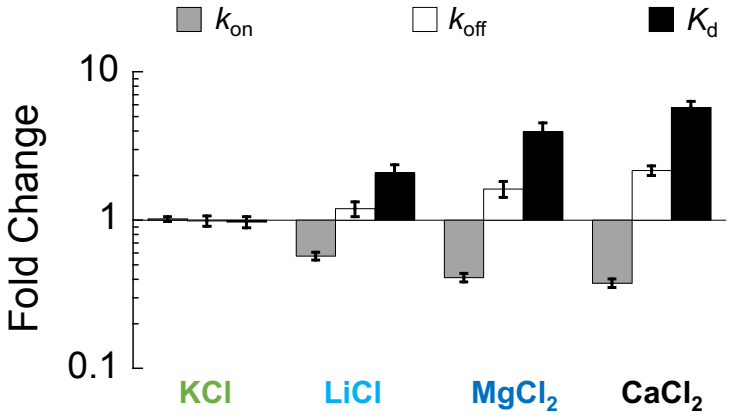

Figure 5.3 – Fold changes compared to NaCl for the interaction between PUMA and MCL-1. The data refers to the 1 M ionic strength conditions. Error bars were obtained from error propagation of the standard error of the mean reported in Table 5.1.

The dissociation rate constants for spectrin were indirectly obtained from performing the association reactions under reversible conditions. However, the fitting of K_d—used to calculate k_{off}—was less robust than the k_{on} parameter, and the analysis was not pursued in great details. The values for strongly contrasting salts (NaCl and $CaCl_2$) and buffer-only are reported in Table 5.2. Interestingly, the effect of ion-type on the dissociation rate constant appeared more pronounced in the case of spectrin than PUMA:MCL-1. Indeed, a ∼6-fold reduction in the lifetime of the complex was observed when switching between sodium and calcium (at the same ionic strength). The transition state of spectrin is relatively structured, and thus it is reasonable to assume that it might occur late on the reaction coordinate. In contrast, the transition state for the binding of PUMA to MCL-1 is almost completely unfolded, and by analogy would occur early. By applying a Hammond-like postulate, the bound state of spectrin is expected to be closer in energy to its TS (relatively). Therefore, small energetic changes would be proportionally more important for the dissociation reaction of spectrins than PUMA:MCL-1. This would result in a more pronounced effect on the k_{off} of spectrin.

Despite a greater impact on k_{off}, the overall effect on the affinity of spectrin was not as pronounced (∼3-fold). This might be because electrostatic screening competes with salt-

specific effects. For PUMA:MCL-1 *both* ionic strength and calcium reduced the speed of association. Provided that the effect of switching from sodium to calcium is similar for spectrin, the ion-specfic outcome would be a reduced association rate. However, greater ionic strength accelerates the spectrin reaction. Thus, the two effects would cancel each other out, which would explain the K_{d} results.

Table 5.2 – Kinetic and thermodynamic parameters for the interaction between spectrin proteins in the presence of different salts (at 1 M ionic strength) and no salt condition. k_{on} and K_{d} were obtained from reversible association experiments. Dissociation rate constants were calculated by $k_{\mathrm{off}} = k_{\mathrm{on}} \cdot K_{\mathrm{d}}$. Errors represent standard error fo the mean, which were propageted in the case of k_{off}.

	$k_{\mathrm{on}} \times 10^2$ / $\mathrm{M^{-1}\ s^{-1}}$	$k_{\mathrm{off}} \times 10^{-4}$/ $\mathrm{s^{-1}}$	K_{d} / μM
No salt	2.1(±0.1)	4(±1)	1.9(±0.5)
NaCl	7.8(±0.1)	3.3(±0.5)	0.42(±0.06)
$CaCl_2$	12.06(±0.08)	19.0(±0.6)	1.58(±0.05)

5.4 Hofmeister effect on coupled folding and binding

5.4.1 The amount of residual structure of the IDP is ion-specific

The absence of convergence for the basal association rate constants in the presence of different ions was not consistent with a simple screening effect. Moreover, the dissociation rate constants also revealed a dependence to the nature of the salt. These reactions being unimolecular, consequences resulting from altered long-range electrostatics can be excluded. Prediction of the helicity of PUMA as a function of ionic strength (Fig. 5.1) suggests that the presence of salt might lead to possible structural effects. This question was investigated using circular dichroism (CD) spectroscopy, which allows bulk secondary structure properties of proteins, and changes in residual helicity to be determined (Kelly *et al.*, 2005, Greenfield, 2006).

Consistent with the prediction from AGADIR, PUMA showed a reduction in helicity with increasing ionic strength (Fig. 5.4A). However, the residual structure was also ion-dependent. This effect was far from negligible, with ion-specific changes accounting for about *half* of the overall change in helicity (the rest being due to ionic strength). Similar to the kinetics findings, structural changes did not appear to be a consequence of the

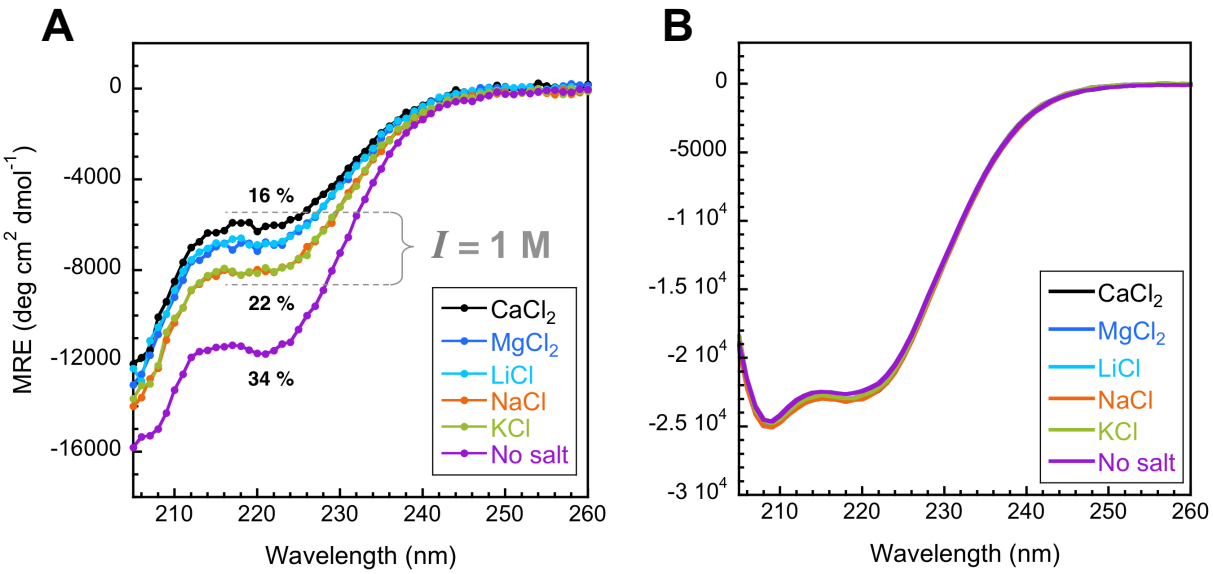

Figure 5.4 – The structural effects of ionic strength and salt-types were investigated by circular dichroism spectroscopy. Buffer-only (I = 4 M) was compared with the effect of different salts at 1 M ionic strength. (**A**) The residual helicity of PUMA was affected by ionic strength, indicated by the lower MRE values in the presence of any of the salts. Importantly, the residual structure was also sensitive to the nature of the ion. (**B**) The structure of MCL-1 was not affected by ionic strength and salt-type at the concentrations used for these experiments. These results highlight the sensitivity of IDPs to both ionic strength and the nature of the co-solute.

charge of the ion. Indeed lithium and magnesium had comparable effects. No changes due to either ionic strength or ion-type were observed for the folded protein MCL-1 (Fig. 5.4B). This result is expected given the range of salt concentrations used in this study ($\leq$ 1 M), which is typical of biochemical buffers. However, the stark contrast with PUMA clearly highlights the higher sensitivity of IDPs towards changes in solution conditions.

Unfortunately, no data could be obtained for bromide and iodide, as these ions strongly absorb in the far-UV range used for CD spectroscopy. Similarly, no conclusion could be drawn from the data collected on spectrins. Because each IDR is flanked by a folded domain, the signal from the disordered regions were completely masked by the strong helical signal of these domains. Therefore, it was not possible to assess the subtle effects of salts and ionic strength on the IDRs of spectrin proteins (data not shown).

Interestingly, the binding rate constants of PUMA to MCL-1 correlated with the structural content of free PUMA observed by CD spectroscopy (Fig. 5.5). These relationships strongly suggest that the ion-specific kinetics results are actually of structural origin. For the association, destabilising the nascent structure of PUMA reduced the binding speed (Fig. 5.5A). It is noted that this correlation does not imply conformational selection and

is equally consistent with an induced-fit mechanism (Shammas *et al.*, 2016). Indeed, the reaction being kinetically 2-state, the association rate constant contains both the binding and the folding steps. Thus, modulation of the folding rate *after* binding would result in the same observable outcome. For the dissociation, the less helical the free IDP, the shorter the lifetime of the complex (Fig. 5.5B). This result might at first seem surprising. Indeed, the unbound state of PUMA does not exist on the reaction coordinate between the bound structure and the transition state. So how does the helicity of free PUMA affect the dissociation reaction? MCL-1 was structurally unaffected by the presence of salts (Fig. 5.4B), hence it is reasonable to assume that the folded complex would be the same. Therefore, changes in the lifetime of the bound state would have to happen from modulation of the energetics of the transition state. Under this assumption, the correlation with the residual structure of free PUMA would simply represent a proxy for the helicity of the TS.

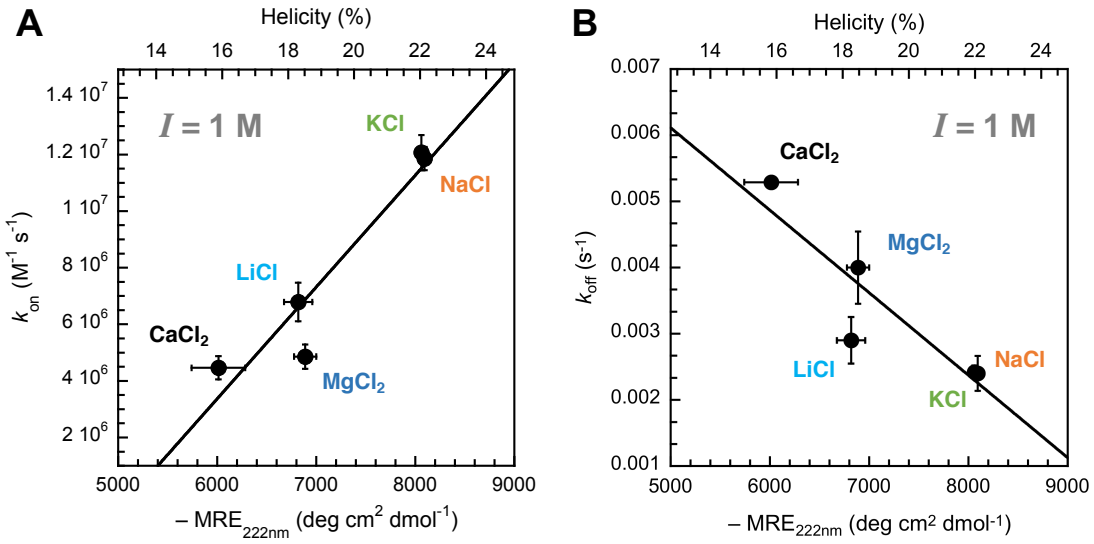

Figure 5.5 – Binding rates of PUMA to MCL-1 (Table 5.1) correlate with its ion-dependent helicity (Fig. 5.4). Using the MRE value at 222 nm as a proxy for helicity, positive and negative correlations were found with the association rate constant (**A**) and dissociation rate constant (**B**) respectively. Interestingly, the fold changes (over the range of MRE values available) were similar in both cases (∼2.5-fold). The helical content in % is indicated above the plot for reference (estimated using the method of Muñoz & Serrano (1995)). Error bars represent standard deviations.

The results presented in this section highlight that: *i*) ionic strength affects the stability of the transient helix of PUMA, *ii*) the amount of residual structure is ion-dependent, and *iii*) binding rates correlate with these structural changes. It is important to stress that these effects are not simply a consequence of the charge-state of the ion. This is evident from the similar CD spectra in the presence of Li^+ or Mg^{2+}. Furthermore, the charge density of

the ions—which has been reported to affect RNA folding (Koculi *et al.*, 2007)—does not explain the kinetic results observed for the anion series (Fig. 5.2B); the trend would be expected to be inverse if that was the case. Rather, these trends in kinetics and structural changes follow the Hofmeister series of these ions (Fig. 5.6).

5.4.2 The Hofmeister series

The Hofmeister series provides a qualitative ranking of ions with respect to their abilities to solubilise (salting in) or precipitate (salting out) proteins. Typical examples include the use of guanidinium chloride (or thiocyanate) for denaturing proteins in chemical unfolding experiments, or the use of ammonium sulfate precipitation in fractionated protein purification protocols. Anions tend to have a stronger effect than cations. This classification of ions and their associated effect on protein stability has long been established (Hofmeister, 1888), and has been the focus of extensive research over the years (Baldwin, 1996, Gurau *et al.*, 2004, Zhang & Cremer, 2006, 2010, Lo Nostro & Ninham, 2012).

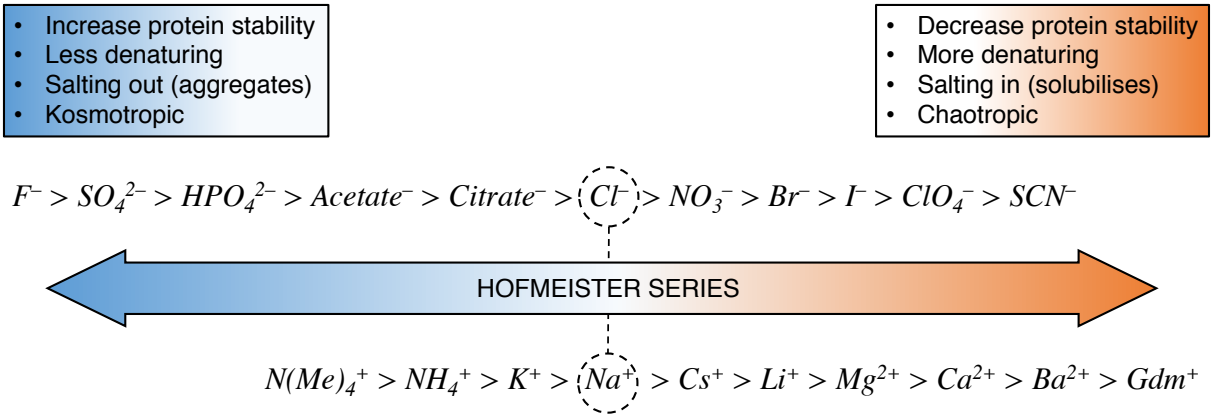

Figure 5.6 – The Hofmeister series provides a qualitative ranking of ions with respect to their ability to solubilise (salting in) or precipitate (salting out) proteins. It is shown centered around sodium and chloride, which are traditionally taken as references of 'neutral' effects within each series.

The exact physical principle behind the Hofmeister effect remains controversial, but binding to peptide backbones and charged residues seem to be at the origin of altered stabilities (Okur *et al.*, 2017). Folded proteins are stable because the sum of their intra-molecular interactions out-weighs the sum of the inter-molecular interactions they would form if they were unfolded. These putative inter-molecular interactions may be to solvent molecules, or co-solutes. However, since these events are multi-molecular, changes in concentration will affect the positions of the equilibria. Thus, even if typical ions have weak affinities

for polypeptides, once sufficiently high concentrations are present, inter-molecular binding starts to compete with intra-molecular interactions, resulting in chain unfolding. It is important to emphasise that the Hofmeister effect is not contingent on particular binding sites, thus making it relatively sequence-insensitive.

While the experiments presented in this chapter do not answer the atomistic details of ion-specificity, it demonstrated sizeable structural effects at low concentrations of 'common' salts, and kinetic effects became apparent for concentrations as low as 10 mM. The correlations between helicity change and the binding affinities suggest structural origins to these ion-specific modulations. But why are coupled folding and binding reactions affected under conditions where folded proteins are insensitive? Indeed, folded proteins usually require multi-molar concentrations of much more 'denaturing' salts (*e.g.* GdmCl) before structural effects become apparent. It is tempting to speculate that the marginal folding stability of IDPs, as well as their larger solvent-accessible surface area, are the reasons for their greater sensitivity.

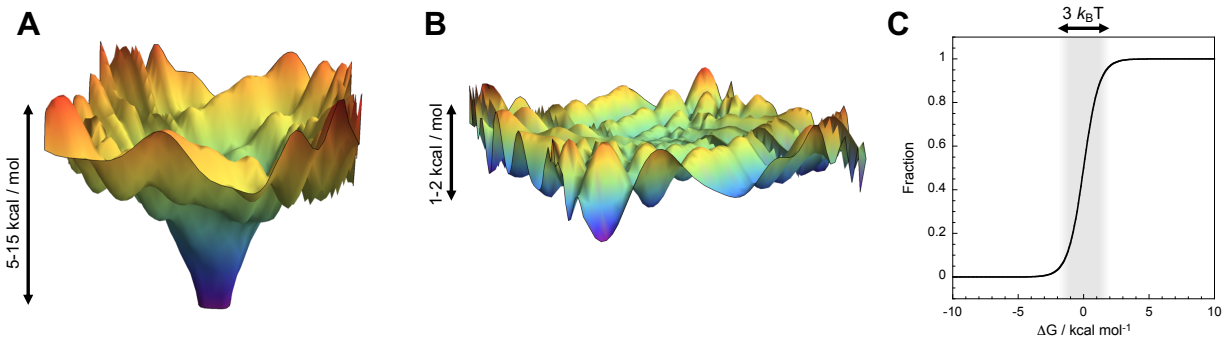

Figure 5.7 – Marginal stabilties of IDPs might explain their sensitivies to solution conditions. (**A**) Schematic representation of the energy landscape of a typical folded protein. The native state is separated from higher-energy states (*e.g.* the denatured ensemble) by 5–15 kcal/mol (Fersht, 1999). (**B**) In contrast, the energy landscape of an IDP is expected to be much shallowerer, with many near-isoenergetic minima that the protein can sample. For these different states to be populated at room temperature, energy differences between them need to be within a few k_BT (1–2 kcal/mol). It is noted that these landscapes were not calculated from physical models, and are shown here purely for visualisation purposes. (**C**) Fractional distribution as a function of the free energy difference between two states. For systems with marginal stabilities (shaded area), small changes in relative free energies strongly affect the distribution.

This hypothesis is illustrated in Fig. 5.7. Typical folded proteins have funnelled energy landscape (Onuchic *et al.*, 1997), with at least one deep minima (Fig. 5.7A). The energy difference to alternative states (*e.g.* denatured ensembles) are typically in the order of 5–15 kcal/mol. In contrast, IDPs do not have a singly-defined structure; they populate

ensembles of conformations (Fig. 5.7B). Since these different states are all accessible at room temperature, differences in stabilities have to be within a few $k_{\mathrm{B}}T$ (1–2 kcal/mol). Hofmeister salts such as GdmCl destabilise proteins by affecting the free energy difference between native and denatured states, and this effect is linearly dependent on denaturant concentration (Tanford, 1968, Pace, 1986). Assuming that the salt effects reported in this chapter are akin to that of GdmCl—a reasonable assumption given that they are all part of the Hofmeister series—they would be expected to affect the stability of proteins as well. The magnitude of the effect ($\Delta\Delta G$) should be the same for a given salt, regardless of whether the protein is MCL-1 or PUMA. However, since PUMA is only marginally stable, small free energy differences would result in large re-distributions between states (Fig. 5.7C, shaded area). This is a direct consequence of the flatter energy landscape expected of IDPs. In contrast, the same $\Delta\Delta G$ would have no *observable* effect on the folded distribution of MCL-1. Thus, it appears that the properties of the energy landscape of IDPs make them particularly susceptible to small changes in solution conditions, which in turn might affect coupled folding and binding reactions. This might be a functional consequence of protein disorder.

5.5 Deconvoluting structural from electrostatic effects

Ionic strength, regardless of ion-type, destabilised the nascent helical structure in PUMA. This is evident from Fig. 5.4A, where PUMA in the presence of all salts at 1 M ionic strength is less helical than under buffer-only condition. Although the absolute values differ, the same trend was predicted using AGADIR (Fig. 5.1). Thus, changes in ionic strength of the solution may affect association rate constants of IDPs by two different mechanisms; through shielding of long-range electrostatic interactions, and through changes in residual structure. But what is the relative weight of each effect on the rate of formation of the IDP:partner complex? Incidentally, the ion-specific structural changes observed under conditions of identical ionic strength provide a means to deconvolute 'pure' electrostatics from structural effects.

The correlation obtained from Fig. 5.5A (helicity *vs* association rate constant for the different salts at 1 M ionic strength) can be used to estimate the association rate constant for any arbitrary values of MRE at the same ionic strength (Fig. 5.8). Therefore, electrostatic and structural effects can be deconvoluted by comparing observed rate constants

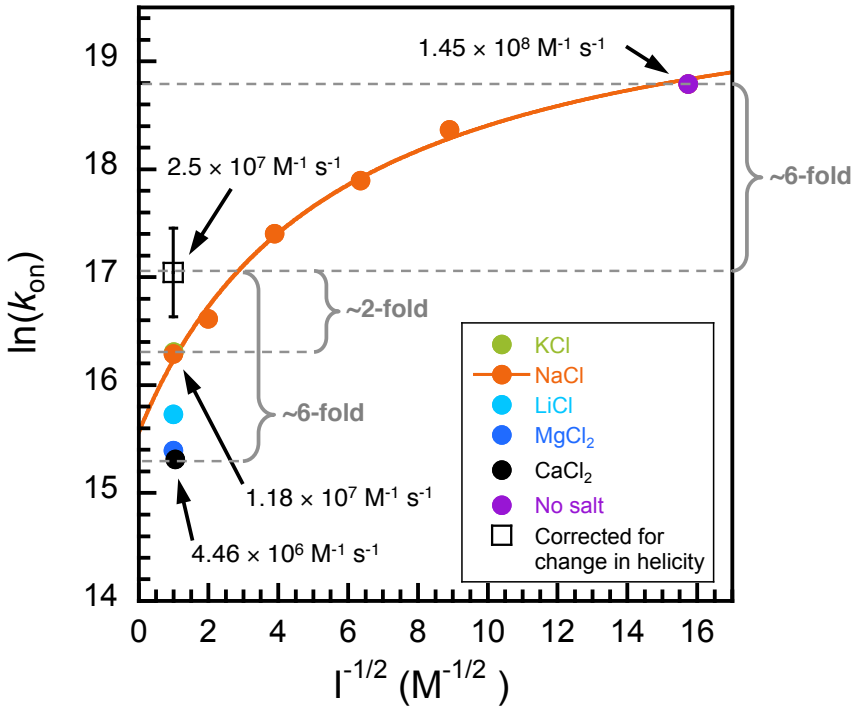

Figure 5.8 – Contribution from the helicity of PUMA to its association rate constant. Shown as an open square is the predicted rate constant (at 1 M ionic strength) assuming no change in helicity with respect to the buffer-only condition. Comparison with observed values indicates a 2- to 6-fold discrepancy, which can be attributed to structural effects. The extrapolated point (open square) was obtained by using the correlation from Fig. 5.5A (k_{on}= –2.03 10^7 + 3948.6 (–MRE_{222nm})), and the MRE value of PUMA under no salt condition (–11536 deg cm^2 $dmol^{-1}$). The data, and Debye-Hückel fit, for the association in the presence of NaCl is reproduced for comparaison. Error bars represent relative errors (obtained by uncertainty propagation).

with estimates corrected for changes in helicity. Using the MRE_{222nm} value at 4 mM ionic strength (buffer only), the extrapolated k_{on} at 1 M ionic strength becomes 2.5(±1.0) 10^7 M^{-1} s^{-1} (open square). This value corresponds to the association rate constant at 1 M ionic strength assuming that PUMA has the same helicity as the peptide in buffer (I = 4 mM). Therefore, the change in k_{on} over the range 4 mM to 1 M ionic strength can be attributed purely to screening of charge-charge interactions, and excludes structural effects. Taking NaCl as an example, this suggests that the ∼12-fold decrease in k_{on} observed over that range is actually ∼6-fold electrostatic, and 2-fold due to the reduction in helicity. The effect is even more pronounced for *e.g.* $CaCl_2$, where the ∼35-fold change is 6-fold electrostatic, and ∼6-fold structural; *half* of the observed change in k_{on} results from a loss of the intrinsic helical structure of the IDP. Intriguingly, these results imply that the association of PUMA with MCL-1 is even less electrostatically enhanced than previously thought. This finding further supports the notion that one of the consequence of disorder on PPIs might be a

marginal electrostatic component to their association rate constants (*cf.* Chapter 4).

In the PUMA:MCL-1 case, adding salt had a compounding effect on the association; it reduced long-range electrostatic attractions, and disrupted the residual structure of the IDP, both of which contributing to a slower binding. It is possible, even probable, that in some systems the interplay between charge-charge interactions and structural effects might be opposing. In the spectrin system, for example, increased ionic strength accelerated association, whilst salts are likely to decrease residual structure (assuming the same effect observed on PUMA), and thereby decrease the on-rate. Thus, the apparent outcome of adding salt on k_{on} might be less, as instead of compounding, the effects would cancel each other out. The findings from this section suggest that use of different salts while keeping the ionic strength constant could be used for mechanistic investigations, allowing the deconvolution of structural and electrostatic effects in coupled folding and binding reactions.

5.6 Discussion

Despite their importance and prevalence, far less is understood about the fundamental biophysics of IDP:partner interactions than about PPIs involving structured proteins (Gibbs & Showalter, 2015, Shammas *et al.*, 2016). In particular, the role of solvent conditions and co-solutes are usually neglected, despite their known effects on IDP structural ensembles (Müller-Späth *et al.*, 2010, Soranno *et al.*, 2014). In Chapter 4, the role of electrostatics in coupled folding and binding was investigated. It was discovered that, despite their sequence bias towards charged residues, association reactions involving IDPs were not significantly accelerated by charge-charge interactions. It is postulated that the lack of pre-formed binding interfaces might preclude directional steering; a possible consequence of protein disorder. However, in order to perform this analysis, NaCl titrations had to be conducted. Changes in ionic strength have previously been reported to affect the structural ensemble of IDPs, raising questions about its effect on binding reactions.

In this chapter, a systematic analysis of the role of charged co-solutes on coupled folding and binding reactions was performed. The two contrasting intrinsically disordered systems investigated in Chapter 4 were used for this study. The results revealed that binding affinities were ion-specific, even when normalised for ionic strength. This highlighted that adding salts lead to more than electrostatic screening. By deconvoluting the stability of the complex into its kinetic components, it was possible to show that affinity changes stemmed from variation in both association *and* dissociation rate constants. The modulation of the lifetime of the complex was a clear indication that effects beyond long-range electrostatics were involved. These ion-specific differences were linked to structural changes in the free IDP, and probably the transition-state as well. Interestingly, the trends in kinetics and affinity were found to relate to the Hofmeister series.

Importantly, these ion-specific results were observed at low concentrations of salts, which appears to have been unappreciated so far. While the Hofmeister effect has been known in the context of folded proteins for a long time (Hofmeister, 1888), structural repercussions are usually not observed until multi-molar concentrations of ions are present. In stark contrast, the work presented here revealed systematic deviations in kinetic profiles for concentrations as low as 10 mM. It is tempting to suggest that the higher structural sensitivity of IDPs—even to modest changes in solution conditions—might be a consequence of their marginal folding stabilities. Because of the presence of a folding dimension

that lies on-pathway to binding, this sensitivity translates into modulation of both binding kinetics and affinity. Since the effect appears to be related to the Hofmeister series—a sequence-independent phenomenon—and because converging results were obtained for two contrasting model systems, these findings are anticipated to be of general applicability in the context of PPIs involving disordered partners.

Although the effects reported here were relatively modest (within an order of magnitude), they occurred at physiological concentrations of salts. It has been shown that relatively small changes in affinity—stemming from altered residual structure of the IDPs—can have significant physiological consequences. In the context of p53 binding MDM2, for instance, changes in residual structure upon mutation resulted in a 10-fold shift in K_d that strongly impaired cellular function (Borcherds *et al.*, 2014). Thus, it is interesting to speculate that the sensitivity to environmental conditions reported here may have physiological implications, especially given the asymmetric ion-profiles of different cellular compartments (Alberts *et al.*, 1989), the role of ion fluxes in signalling pathways (Clapham, 2007), and the importance of charged osmolytes in maintaining cellular function (Yancey *et al.*, 1982).

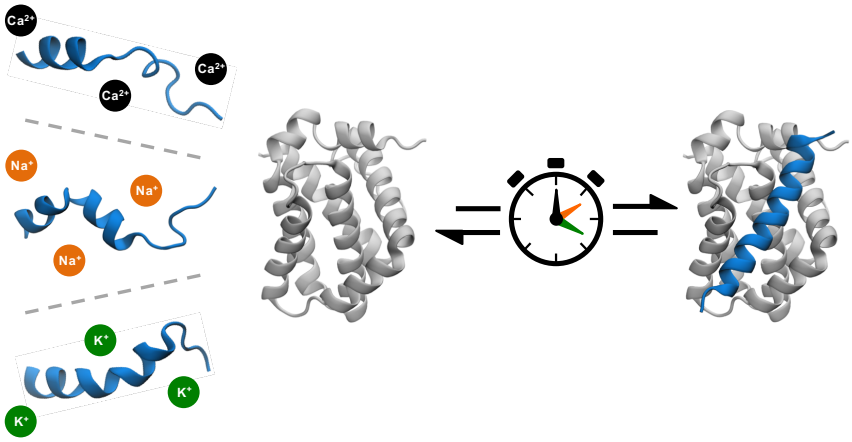

Figure 5.9 – Summary of the ion-specific effect observed in coupled folding and binding. Different salts modify the kinetic and thermodynamic profiles of binding reactions through structural alteration of the disordered partner.

Part II

Oligomerisation and regulation in a tripartite protein network

Chapter 6

Interactions within the BCL-2 family

6.1 Introduction

The theory of evolution states that traits that provide a fitness advantage to an organism will amplify within a population under natural selection. In other terms, a mutation or alteration that provides a survival advantage will be propagated to a greater extent. The concept of survival and fitness is easily understood for unicellular organisms; better survival of the cell means better fitness of the organism. However, multicellular organisms represent a more complex challenge. While the overall organism's fitness is under selective pressure, survival of each individual cell might not prove advantageous. Therefore, some cells might need to be sacrificed to benefit the organism as a whole, a process known as tissue homeostasis. This altruistic death phenomenon also occurs during development; when the cells of an embryo die to delineate the fingers, or the tail of a tadpole gives way to allow the metamorphosis into the frog stage. These examples highlight that in the pursuit of life, death is sometimes necessary. But this simple concept requires extremely careful planning at the cellular and molecular level. Cellular death comes in different forms. Necrosis—the uncontrolled burst of a cell following extensive cellular injury—can be dangerous to the organism (Proskuryakov *et al.*, 2003). The spill of cellular components following membrane rupture triggers immune responses and lead to inflammation. An alternative form of death does exist, called apoptosis (Kerr *et al.*, 1972, Elmore, 2007). In contrast to necrosis, the events leading to the demise of the cell are carefully controlled and programmed at the biochemical level. This molecular orchestration ensures immunogenic factors are carefully handled.

The phenomenon of apoptosis is regulated by two different pathways (Elmore, 2007). Both start with specific stimuli, which trigger sets of biochemical events. Ultimately, both pathways converge to the activation of caspases, a set of intracellular proteases that exist as inactive precursors prior to apoptosis. These go on to degrade most of the cellular interior, resulting in the death of the cell. Responsible for the control of apoptosis are the *intrinsic* and *extrinsic* pathways, named after the point of origin of their triggers. The extrinsic pathway is gated by death receptors, extracellular receptors that bind specific ligands, which triggers the formation of the death-inducing-signalling-complex, a multi-protein complex that activates caspases. In contrast, the intrinsic pathway is activated following intracellular apoptotic stimuli such as oxygen depravation, genotoxic damage and other forms of cellular stresses. These cues are integrated and processed downstream by the BCL-2 family of proteins, which controls mitochondrial outer-membrane permeabilisation (MOMP) (Cory & Adams, 2002). Loss of membrane integrity allows apoptogenic factors such as cytochrome c to leak from the mitochondrial inter-membrane space. Presence of these elements in the cytosol leads to caspase activation and proteolytic degradation of the cellular interior (Kluck *et al.*, 1997). Once started, the caspase cascade becomes an amplifying chain reaction, therefore MOMP is regarded as the point of no return of the entire process. The BCL-2 proteins regulate this pivotal step through a network of pro- and anti-apoptotic members. Thus, the balance of interactions within this network determines cellular fate at the biophysical level (Fig. 6.1B)(Czabotar *et al.*, 2014, Kale *et al.*, 2017). Beyond its physiological role in healthy cells, understanding the network has important biomedical implications. De-regulation of the balance of interactions is often implicated in cancer (Hanahn & Weinberg, 2011), where cells evade apoptosis by resisting MOMP. It is also involved in auto-immune diseases (Eguchi, 2001). Some viruses express BCL-2 homologues acting as molecular decoys (*e.g.* BHRF1 form the Epstein-Barr virus), preventing premature death of the cells they infect by hijacking their apoptosis machinery (Kvansakul & Hinds, 2013). Studying the BCL-2 family and the biophysics of their interactions should aid our molecular understanding of the mechanisms of apoptosis. This work would also expand the study of PPIs involving competing interactions.

BCL-2 members are related by the presence of at least one BCL-2 homology (BH) motif (Aouacheria *et al.*, 2013). All share the BH3 motif, while subsets share more. This BH3 motif is thought to be the site of interaction between the different BCL-2 proteins. At the sequence and functional level, the network can be divided into three (Fig. 6.1A).

At the structural level, most BH3-only proteins are found or predicted to be intrinsically disordered (Rautureau *et al.*, 2010). In contrast, anti-apoptotic proteins fold into globular structures composed of seven amphiphatic helices packing around a central hydrophobic one (Fig. 6.1C) (Petros *et al.*, 2004). BCL-2 proteins display a groove on their surfaces, where BH3-only protein can bind as a single, contiguous α-helix (Fig. 6.1D). Surprisingly, the effectors BAK and BAX exist in soluble form, with a structure virtually identical to their anti-apoptotic counterparts despite low sequence identity (<30%, Fig. 6.1E). This raises question about the biophysical signatures that differentiate these two functionally distinct proteins; how does a soluble protein become a pore-forming one?

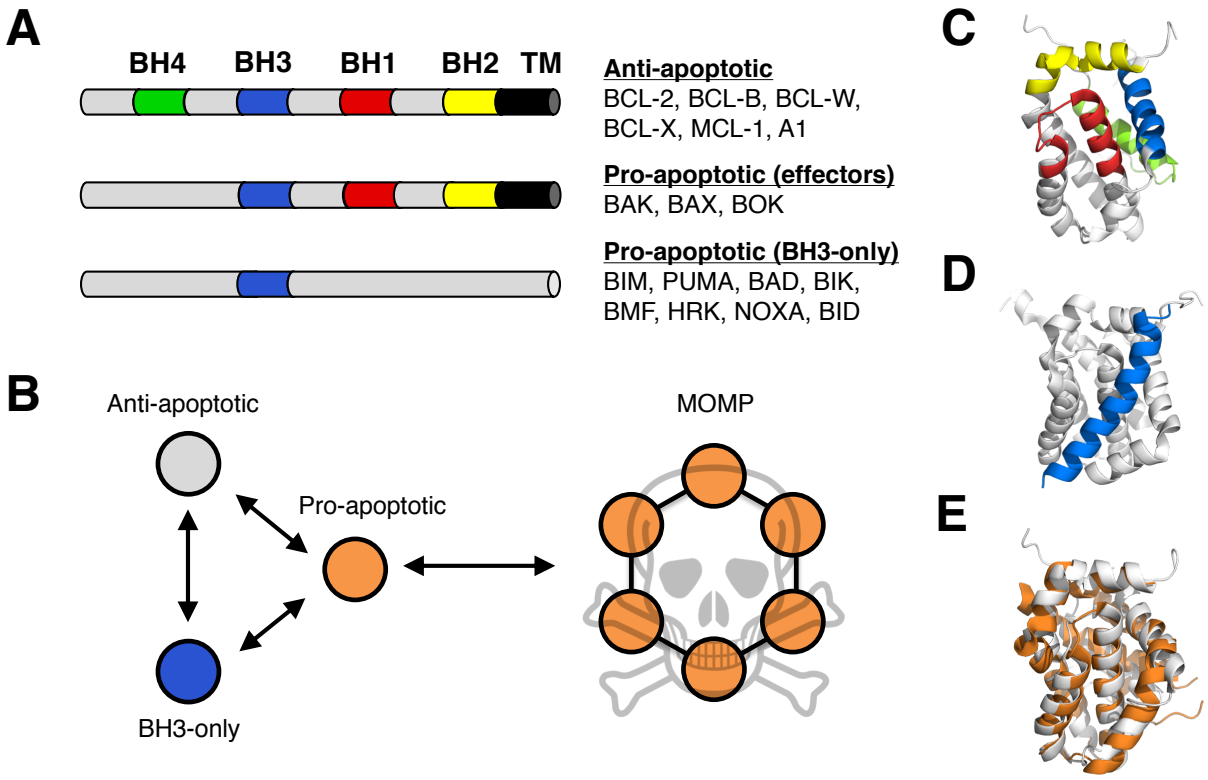

Figure 6.1 – The BCL-2 family. (**A**) Sequence and functional classification of the different BCL-2 proteins. Most multi-motifs BCL-2 have a putative transmembrane (TM) domain. (**B**) Interaction network controlling the oligomerisation of BAK/BAX leading to MOMP. (**C**) Structure of BCL-2 (PDB: 1G5M) with its BH motifs highlighted (colour scheme identical to (**A**)). (**D**) Structure of MCL-1 bound to $PUMA_{BH3}$ at the canonical groove (PDB:2ROC). (**E**) Structural alignment of BCL-2 (PDB:1G5M) and BAK (PDB:2YV6) showing the structural similarities of the pro- and anti-apoptotic BCL-2 proteins. The orientations of the structures in **C**-**E** are the same.

Mechanistically, the regulation of BAK and BAX oligomerisation is debated (Czabotar *et al.*, 2014). Different models have been proposed, each being a variation of the scheme presented in Fig. 6.1B. In the direct activation model, BH3-only proteins directly engage

BAK and BAX, triggering their oligomerisation. This is prevented by anti-apoptotic BCL-2 proteins, which sequester the BH3-only. In contrast, the indirect activation model postulates that BAK and BAX are constitutively 'active', and kept in check by anti-apoptotic members by binding to them. Presence of BH3-only proteins displaces this equilibrium by binding anti-apoptotic members, thus freeing BAK and BAX for oligomerisation. A unified model has also been proposed, which contains mechanistic elements of both to accommodate the observed effect of different BH3-only proteins in cellular studies (Llambi *et al.*, 2011).

6.2 Aims

To date, most studies have taken a 'top-down' approach, making gene knockouts and trying to delineate the role of each BCL-2 and their interactions. While these studies have proven invaluable in describing the network, promiscuous binding and cross-talk between the different members have led to contradictory conclusions. Structurally, most investigations have focused on anti-apoptotic proteins binding BH3 peptides, but little structural or mechanistic investigations of the pore and its formation has been undertaken to date.

Part II of this thesis aims to understand the biophysical characteristics that underlays the BCL-2 network. Using a reduced set of proteins with representatives of each subfamily member, together with kinetic and thermodynamic analyses, mechanistic insights into the regulation of a tripartite protein network were gained. This Chapter describes the tripartite interactions in buffer to understand the behaviour and distribution of states under standard biochemical conditions. Chapter 7 explores the oligomerisation of the effector BCL-2 proteins BAK and BAX. Detergents were employed to provide membrane-like environments, and the mechanism and structural aspects of the oligomerisation were investigated. Given the important role of detergents in regulating the multimerisation of BAK and BAX, the tripartite interaction network was further investigated in this context in Chapter 8. Finally, Chapter 9 explores the link between the conformational changes induced by detergents, and the energy landscape of BAK and BAX.

6.3 Model tripartite system

The BCL-2 family is composed of about twenty proteins in humans. A reduced system needed to be selected to allow a practical number of interactions to be studied, while retaining the features of the network. At least one member of each sub-family (pro-apoptotic, anti-apoptotic, effector) was required to recapitulate all the features. Some simple selection rules were followed:

- Availability of structural data
- Existence of purification protocol in the literature
- Some biophysical knowledge

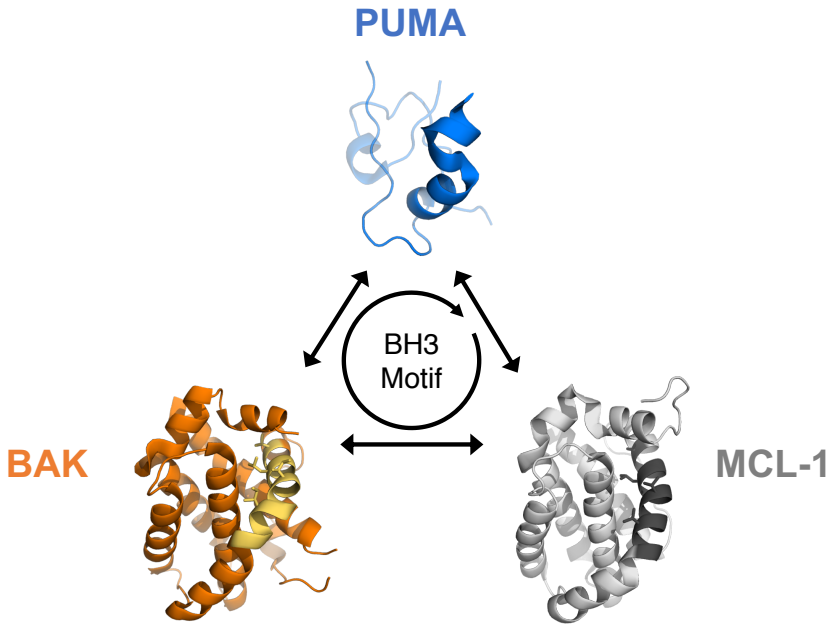

Figure 6.2 – Model tripartite system. Both BAK (PDB:2YV6) and MCL-1 (PDB:2MHS) are monomeric and folded, while BH3-only peptides are intrinsically disordered (here a schematic represenation of disordered PUMA). The BH3 region of BAK and MCL-1 is highligthed in a different shade within their structures. Note that the residues that compose the hydrophobic part of the amphiphatic BH3 motif (in sticks representation) face inwards, and are inaccessible for interactions in these folded states.

Only human (*Homo sapiens*) and mouse (*Mus musculus*) orthologues were investigated as candidates. Being the most studied, a wealth of data is already available (structural, biochemical, biophysical). Owing to their biological relevance, human proteins represent the overwhelming majority of the literature. The murine versions of MCL-1 and PUMA have

already been extensively studied in the group, including detailed kinetic, thermodynamic and mechanistic analyses (Rogers *et al.*, 2013, 2014b,a). Furthermore, most interactions between anti-apoptotic BCL-2 and BH3-only proteins have been reported (Ku *et al.*, 2011). These studies include thermodynamic (and to an extent kinetic) data, as well as the structures of many BCL-2:BH3 complexes. Because only structures of human effector BCL-2 (BAK, BAX) were available, it was decided to study exclusively human versions of the proteins. MCL-1 was selected as the anti-apoptotic member. Both BAK and BAX were studied, with more emphasis put on BAK due to its ease of expression and purification. BH3-only proteins were all investigated as 35-mers—encompassing the 15 residues of the BH3 motif, and 10 flanking residues on both sides—in line with previous work. These were initially expressed as GB1 fusions, and cleaved from the expression and solubilisation tag after purification. However, due to time and spectroscopic constrains, TAMRA-labelled peptides were ultimately ordered and used for most of the work presented in this thesis (thereafter labelled as t-X_{BH3}, where X is any BCL-2 protein). Proteins were expressed, purified, and biophysically characterised in the strict absence of detergents. All biophysical experiments were performed in 50 mM sodium phosphate, pH 7.0 at 25 °C unless stated otherwise.

6.4 Binary interactions among BCL-2 members

6.4.1 MCL-1 and BH3 motifs

Mouse MCL-1 has previously been reported to associate tightly with BH3-only proteins, with most interactions being either low or sub-nM (Ku *et al.*, 2011). Kinetic profiles of these interaction reveals fast binding—albeit not diffusion-limited—and slow dissociation rate constants (Rogers *et al.*, 2013). Human MCL-1 was expressed in a manner analogous to its murine counterpart (*cf.* Materials and Methods), producing monomeric protein. Investigations of the interaction between MCL-1 and different BH3 motifs were performed to gain insights into the characteristics of these interactions as the model for BH3 binding anti-apoptotic proteins part of the tripartite network.

As a first step to demonstrate the presence of coupled folding and binding events, equilibrium structural studies by circular dichroism spectroscopy were performed. The CD spectrum of MCL-1 shows the typical profile for an α-helix with dips at 208 and 222 nm.

In contrast, BH3 peptides have characteristic 'random-coil' profiles. Comparing the spectra of the complex between MCL-1 and BH3 peptides with that expected if the two proteins were not interacting shows an increase of structure. This is indicative of a coupled-folding and binding reaction, which can be attributed to the folding of the BH3 motif to an α-helix onto the groove of MCL-1 as depicted in Fig. 6.1D.

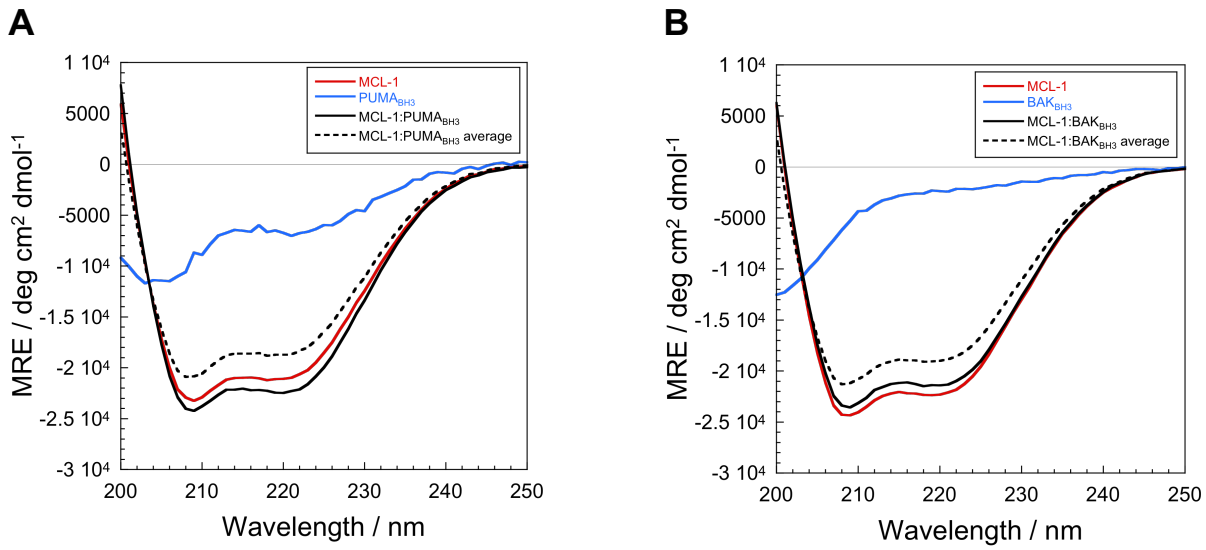

Figure 6.3 – Coupled folding of BH3 motifs with MCL-1 probed by circular dichroism spectroscopy. Each protein concentration is 5 μM. BH3 peptides have a TAMRA dye at the N-terminus. (**A**) Interaction between MCL-1 and the BH3 motif of PUMA. (**B**) Interaction between MCL-1 and the BH3 motif of BAK. In both cases the amount of structure of the complex (solid black line) is greater than the average of its parts (dashed black line), indicating coupled folding and binding.

Next, kinetics was measured for MCL-1 interacting with different BH3 motifs. Association kinetics was performed under pseudo-first-order conditions with MCL-1 in excess using stopped-flow experiments. Conditions were: hundreds of nM for the BH3 peptides, and at least 10-fold excess of the partner protein. Experiments were performed with dye-free BH3 peptides, following intrinsic tryptophan fluorescence, or with TAMRA-labelled equivalents and monitoring either the fluorescence or anisotropy of the dye. Results were broadly similar, and indicated few differences due to either the presence of the dye or the kinetic probe employed. These results confirmed the validity of using dye-labelled versions of BH3 peptides for probing the kinetics of these reactions. Overall, all reactions appeared 2-state (single exponential kinetics) and were fast under the conditions of the experiments (k_{on} = 10^6 – 10^7 M^{-1} s^{-1}, *cf.* Table 6.1).

Dissociation rate constants for MCL-1:BH3 complexes were measured by performing out-

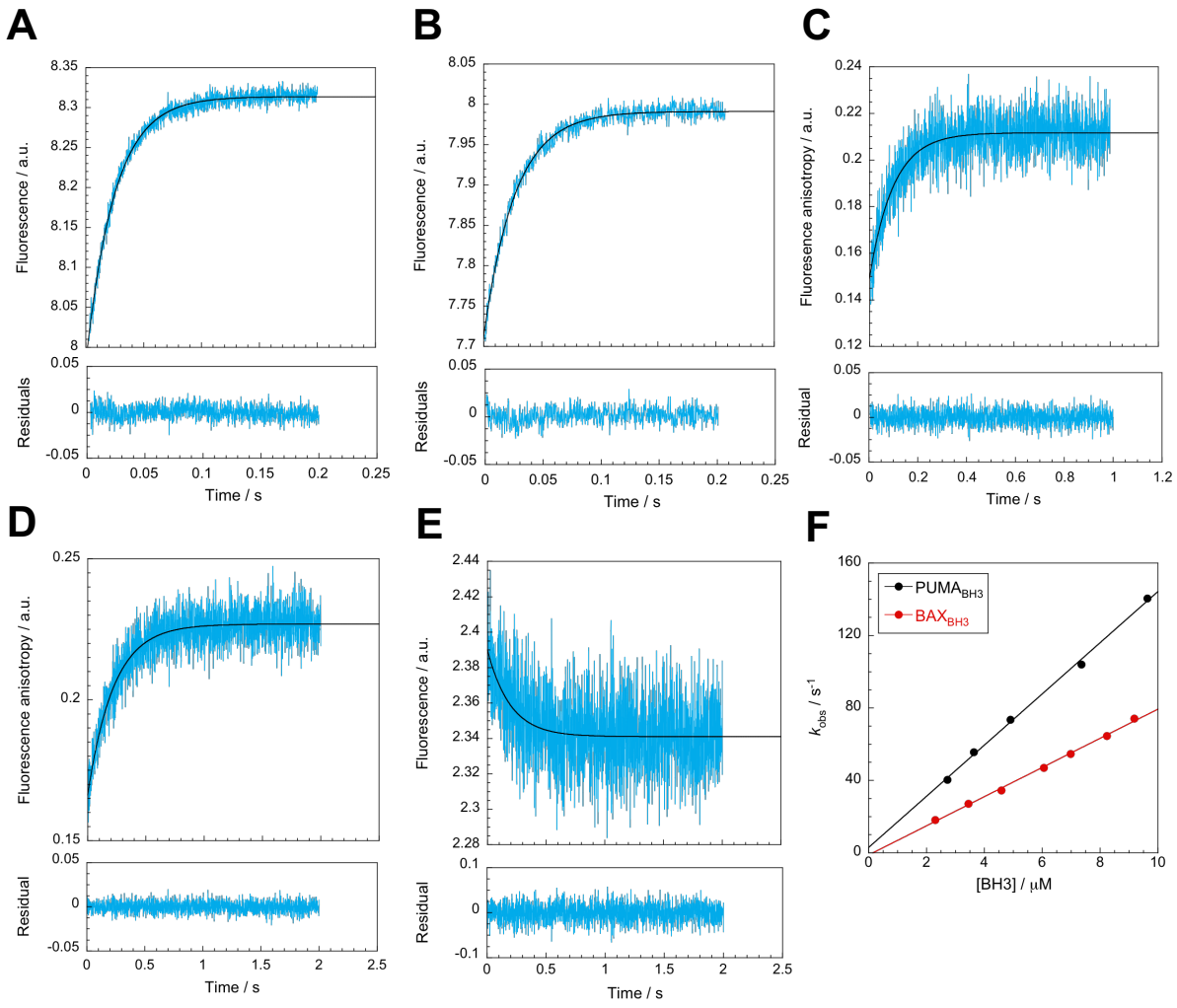

Figure 6.4 – Association kinetics between MCL-1 and BH3 motifs. (**A**) Association between 0.5 μM $PUMA_{BH3}$ and 5 μM MCL-1 followed by intrinsic fluorescence intensity change. (**B**) Association between 0.5 μM BAX_{BH3} and 5 μM MCL-1 followed by intrinsic fluorescence intensity change. (**C**) Association between 0.14 μM t-BAK_{BH3} and 1.4 μM MCL-1 followed by extrinsic fluorescence anisotropy change. (**D**) Association between 0.14 μM t-MCL-1_{BH3} and 1.4 μM MCL-1 followed by extrinsic fluorescence anisotropy change. (**E**) Same, but followed by extrinsic fluoresence intensity change. (**F**) Pseudo-first-order plots for $PUMA_{BH3}$ and BAX_{BH3} binding MCL-1. The gradient to the line represents the association rate constant of the reaction (k_{on}). For **A**–**E**, the residuals to single exponential fits are indicated in the corresponding lower panels.

competition experiments. MCL-1 and the BH3 peptide of interest were pre-formed as 1:1 complexes at concentrations of 1–3 μM. Rapid dilution into a large excess of out-competitor was performed, and dissociation followed spectroscopically. If the affinity of the out-competitor for MCL-1, and its excess, are large enough, then the observed dissociation rate constant corresponds to the dissociation of the BH3 peptide from MCL-1. This was confirmed by performing the dissociation experiments at different excess concentrations of competitor, and confirming that the observed rate constants were indeed

independent of out-competitor concentrations. Fast reactions (<1000 s) were measured by stopped-flow techniques, while longer traces were recorded by manual mixing on a fluorimeter. Complexes between MCL-1 and dye-free BH3 peptides were out-competed by mouse $PUMA_{BH3}$ W133F/N149A, a spectroscopically silent and less fluorescent version of PUMA. This allowed large excesses of it to be used, while still being able to measure tryptophan fluorescence changes upon complex dissociation. TAMRA-labelled BH3 peptides were out-competed by $PUMA_{BH3}$, and the reaction followed by observing the change of dye fluorescence upon dissociation. In certain cases loss of fluorescence polarisation (V/V, using a manual polarisation device) gave better signal-to-noise ratios, and was employed instead of fluorescence intensity for monitoring the dissociation reaction.

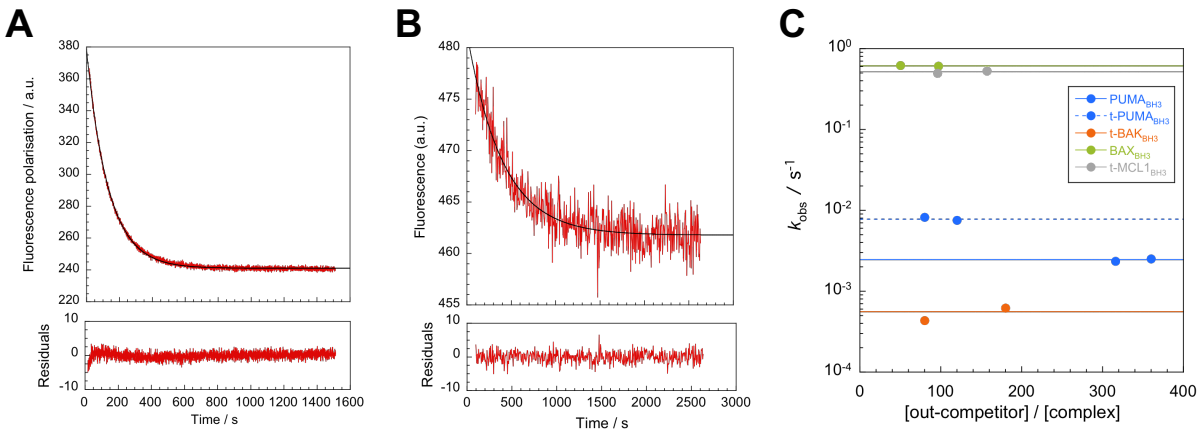

Figure 6.5 – Dissociation kinetics between MCL-1 and BH3 motifs. (**A**) Dissociation of pre-formed t-$PUMA_{BH3}$:MCL-1 (0.2 μM complex) by out-competition with 24 μM $PUMA_{BH3}$ (120-fold excess). The reaction was monitored by following the loss of fluorescence polarisation intensity of the TAMRA dye. (**B**) Dissociation of pre-formed $PUMA_{BH3}$:MCL-1 (0.3 μM complex) by out-competition with 108 μM mouse $PUMA_{BH3}$ W133F/N149A (360-fold excess). The reaction was followed by monitoring the change of intrinsic fluorescence. The residuals to single exponential fits are indicated in the corresponding lower panels. (**C**) Dissociation rate constants of MCL-1:BH3 complexes as a function of out-competitor excess.

All reactions fit to single exponential functions. A linear drift term was sometimes employed to account for photobleaching and/or instrumental drift. This further confirms the 2-state nature of the interaction between MCL-1 and BH3 motifs, as observed in association experiments. All reactions investigated were slow (k_{off} < 1 s^{-1}), but span multiple orders of magnitude depending on the specific BH3 motif. Therefore, the lifetime of the different complexes ranged from about 2 s to more than 30 min.

Both association and dissociation kinetic experiments were single exponential for the reac-

tion between MCL-1 and the BH3 peptide investigated. This kinetic profile is consistent with 2-state reactions without significant population of an intermediate. For such systems, the affinity of the complex can be expressed as the ratio of the dissociation rate constant over the association rate constant ($K_d = k_{off} / k_{on}$). Using this relationship and the rate constants described above, the affinity of BH3 peptides for MCL-1 were calculated. The results are summarised in Table 6.1.

Table 6.1 – Kinetic and thermodynamic profiles of MCL-1 binding BH3 peptides. Error on the rates are fitting errors. Equilibrium affinity obtained from kinetic data, and error obtained by error propagation. Assocation constants for dye-labelled peptides were obtained from a single assocation reaction using $k_{on} = (k_{obs} - k_{off})/[\text{MCL-1}]$.

BH3 peptide	k_{on} / M^{-1} s^{-1}	k_{off} / s^{-1}	K_d / nM
$PUMA_{BH3}$	1.41(±0.04) 10^7	2.46(±0.05) 10^{-3}	0.174(±0.06)
BAX_{BH3}	8.6(±0.3) 10^6	0.613(±0.05)	71(±3)
t-BAK_{BH3}	7.2(±0.3) 10^6	5.6(±0.6) 10^{-4}	0.077(±0.009)
t-$MCL\text{-}1_{BH3}$	2.9(±0.1) 10^6	0.52(±0.01)	182(±7)

In line with other studies performed on BH3 peptides binding MCL-1, all interactions were tight. It is interesting to notice that the changes in affinities can be almost exclusively attributed to the variation in dissociation rate constants, while all association rate constants were found to be within an order of magnitude. This observation corroborates recent reports in the literature of both BCL-2 systems (Dahal *et al.*, 2018) as well as others (Crabtree *et al.*, 2017), which points at the lifetime of the complex as the main determinant of affinity, not its speed of formation. Thus, modulation of the sequence of the IDP can tune the affinity to its partner protein without significantly affecting how fast it forms.
It is interesting to notice that MCL-1 has significant affinity for its own BH3 motif. This result is unexpected given that there is not biological report of MCL-1 interacting with itself (or to other BCL-2 proteins *via* its BH3).

6.4.2 BAK and BH3 motifs

Despite their importance within the BCL-2 network, there is a striking paucity of biophysical data regarding the interaction of BAK and BAX interacting with BH3-only proteins. This is surprising given the postulated role of these motifs in the regulation of BAK and

BAX, specifically in the direct activation model. In contrast, the interaction of anti-apoptotic BCL-2 and BH3-only proteins has been well-documented (Ku *et al.*, 2011), with many complexes available in the Protein Data Bank, as well as affinity measurements. One reason might be that BAK and BAX are more challenging to express.

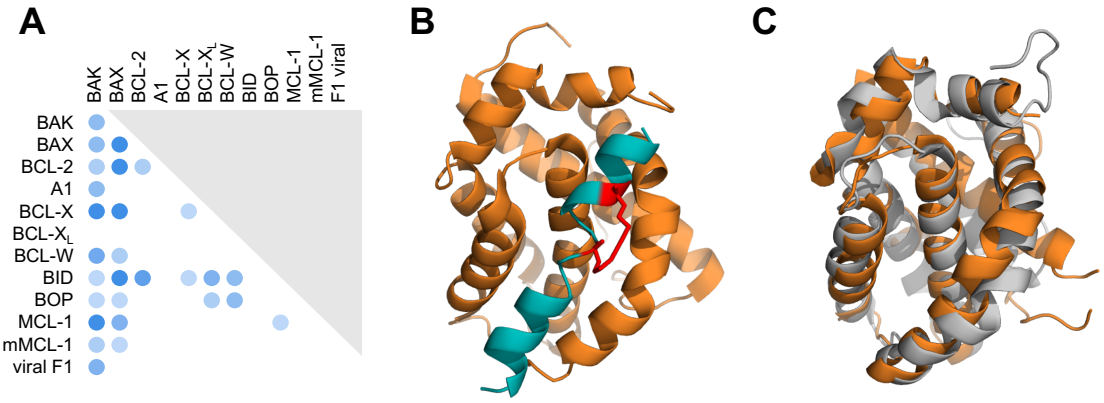

Figure 6.6 – Binding partners of BAK. (**A**) Interactions between BAK and other proteins as reported in UniProt. (**B**) Structure of BAK in complex with a stappled BID as determined by NMR spectroscopy (PDB:2M5B). The staple is represented as red sticks. (**C**) Structural alignment of BAK (orange, PDB:2YV6) and MCL-1 (grey, PDB:2MHS). The RMSD over 122 C_α is 2.64 Å and the sequence identity 20.93%.

The UniProt database lists multiple binding partners for BAK. These are based on a variety of experimental methods (pull-down assays, co-localisation experiments, etc...) and might not necessarily reflect strict physical interactions. A single structure of a complex has been reported between monomeric BAK and a stapled version of BID (Moldoveanu *et al.*, 2013). The study also reports a much lower affinity ($K_d \sim 10$ µM) than for typical anti-apoptotic:BH3 interactions, despite binding at the canonical groove.

BAK could be expressed and purified as a monomeric protein, and remained stable without detectable oligomerisation over months when stored in buffer in the strict absence of detergent. Expression and purification protocol are reported in Materials and Methods. Monomeric BAK and dye-free BH3 peptides or TAMRA-labelled versions were employed to gain insights into the interactions between effector and BH3-only proteins.

Interactions between BAK and different BH3 peptides were investigated by circular dichroism spectroscopy to test for the presence of coupled folding and binding events. Given the structural homology between BAK and MCL-1—thus the presence of the BH3 binding groove—and the existence of a structure of the BAK:BID complex, similar results to MCL-1 were expected (*cf.* Fig. 6.3). Indeed, binding of BH3 motifs to MCL-1 being in the low to sub-nM, it was rationalised that binding could be observed in the µM concentration

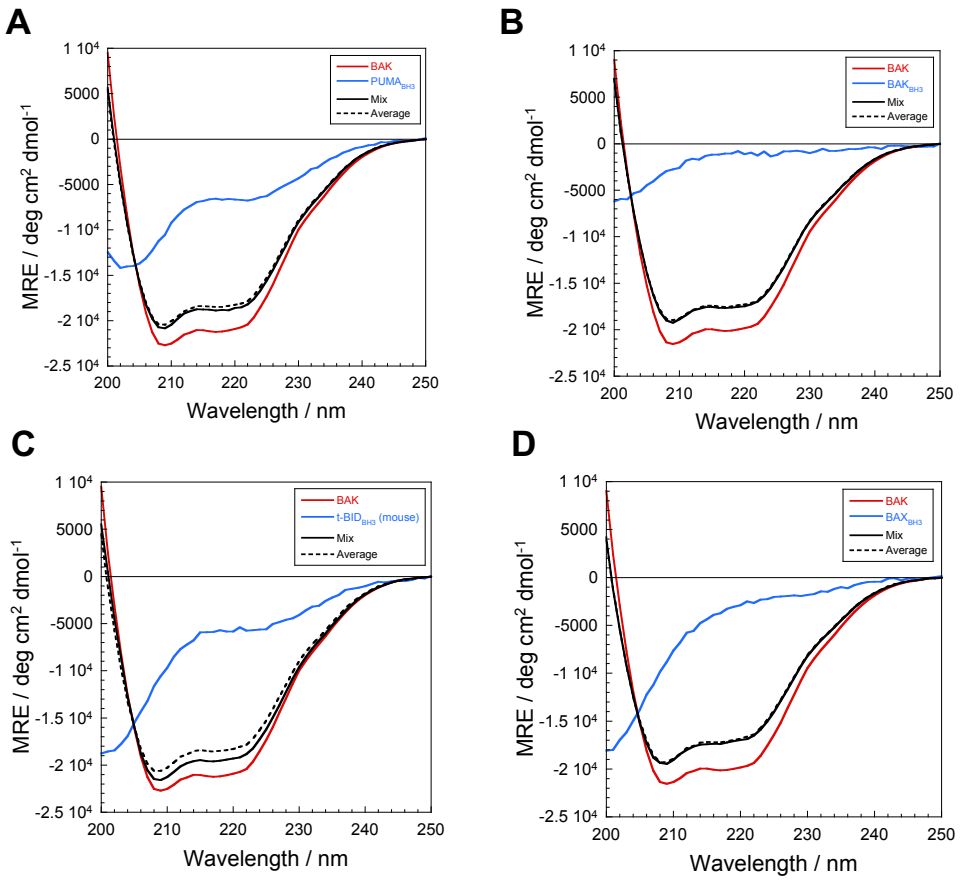

Figure 6.7 – Coupled folding of BH3 motifs with BAK probed by circular dichroism spectroscopy. (**A**) Mixing BAK and $PUMA_{BH3}$ shows a small increase in structure. (**B**) Mixing BAK and BAK_{BH3} shows no increase in structure. (**C**) Mixing BAK and mouse t-BID_{BH3} displays evidence of binding. (**D**) No interaction observed between BAK and BAX_{BH3}. Each protein concentration is 10 μM except BAK_{BH3} (7.3 μM). MRE values of the mix (solid black line) greater than the average of its parts (dashed black line) is indicative of coupled folding and binding.

range even if the affinity was shifted by multiple orders of magnitude. Surprisingly, interactions between BAK and BH3 peptides could hardly be observed at concentrations of 10 μM. Mouse BID_{BH3} (71.4% identiy with human for the 35-mer) shows the largest gain in structure, but still remains marginal when compared to MCL-1:BH3 complexes. Thus, the affinity of these interactions are likely to be greater than the tens of micromolar. This suggests that BAK does not bind BH3 motifs tightly, unlike its pro-apoptotic structural homologues. Binding of BAK and t-BIM_{BH3} was also investigated. Unfortunately, the strong self-oligomerisation of BIM_{BH3} meant that the peptide was already fully helical at 5 μM, precluding the observation of binding through gain of structure.

To confirm the results from CD spectroscopy, the affinities between BAK and BH3 peptides were quantified by measuring equilibrium dissociation constants (K_d). TAMRA-labelled

versions of the peptides were employed, and BAK was titrated to concentrations of about 1 mM with the intent of obtaining saturated binding even for strongly shifted K_d's. Formation of the bound complex was followed by measuring the change in fluorescence anisotropy of the dye on a fluorimeter equipped with a manual polarisation accessory. Because fluorescence anisotropy reports on the tumbling of the fluorophore in solution, which relates to the size of the complex, it provides a relatively unbiased reading of binding. The binding isotherms of the six BH3 motifs investigated are reported in Fig. 6.8. The data were fitted to a 2-state binding model to obtain K_d, and the values are summarised in Table 6.2. Due to the lack of of binding saturation, the uncertainties on the affinity reported are large, and the values should be interpreted with care. However, these experiments clearly demonstrate that all BH3 motif investigated (except BIM), have very weak affinities to BAK, with a lower-bound estimate of ~100 μM. Indeed, for a 2-state hetero-dimerisation reaction, 90% complex is expected at a concentration 10-fold higher than K_d. Since no clear baseline can be seen when approaching 1 mM BAK, K_d's cannot be tighter than 100 μM. Note that the baselines are different due to the local flexibility of the dye in its peptide context.

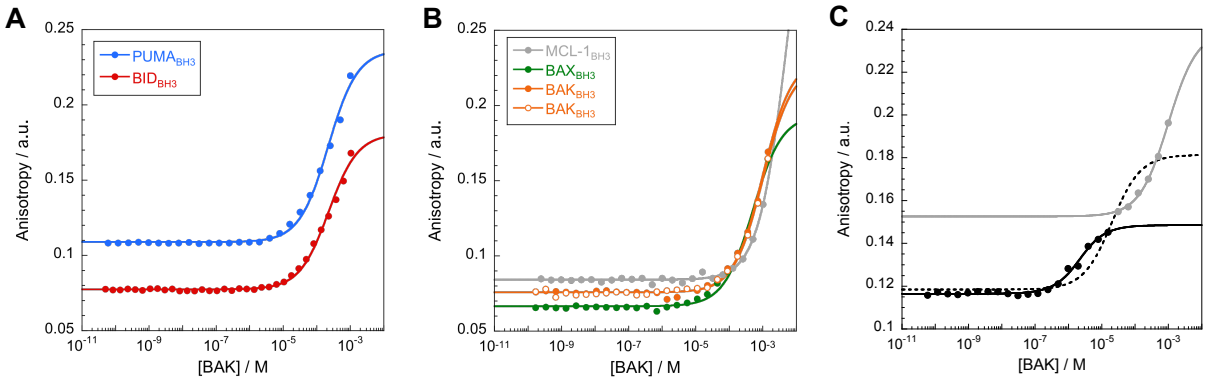

Figure 6.8 – Binding isotherms between BAK and TAMRA-labelled BH3 peptides followed by fluorescence anisotropy. (**A**) Binding between BAK and BH3 motifs from BH3-only proteins. (**B**) Binding between BAK and BH3 motifs from folded BCL-2 proteins. Repeats are indicated for BAK_{BH3}. (**C**) Binding between BAK and t-BIM_{BH3}. The whole dataset did not fit well to the 2-state binding model (dashed line), possibly due to the presence of an intermediate. Therefore, each transition was fitted individually. The concentration of dye-labelled peptide was 1 μM in all cases. Solid lines represent fits to a 2-state binding model.

The affinity of BID for BAK reported here is about an order of magnitude higher than previously measured by AUC (Moldoveanu *et al.*, 2013). While the temperature of their experiment was 5 °C lower, possibly accounting for some of the discrepancy, the likely

source of the difference is the peptide sequence itself. Their experiments was performed with a shorter version of BID that lacked two consecutive proline residues just outside of the BH3 motif. These were included in our construct. It has been shown that helix-flanking prolines in IDPs modulate the affinity of their interactions, primarily through modulation of their dissociation rate constants (Crabtree *et al.*, 2017). Therefore, it is not surprising that absence of these residues just outside the BH3 region would affect binding affinity.

The binding between BIM_{BH3} and BAK appeared biphasic and could not be fitted well as a whole. Instead, each transition was fitted individually, yielding K_d's of 2 μM and 1 mM respectively. The presence of two transition can be attributed to the strong propensity of the BIM peptide to self-oligomerise as observed by CD spectroscopy. This was not investigated further. Interestingly, binding of BH3 motifs from folded BCL-2 proteins (MCL-1, BAK and BAX) was even weaker. This is especially surprising given the postulated role of BAK's own BH3 in its self-oligomeristion (see Chapter 7).

BAK and BH3 motifs do not associate tightly, and the origin of this phenomenon was investigated. We asked the question; does slow association or fast dissociation explain these weak K_d's? Stopped-flow techniques were employed, and association reactions performed under pseudo-first-order conditions with BAK in excess. Dye-labelled versions of the peptides were used, and the reactions monitored by change in extrinsic fluorescence or anisotropy. Pseudo-first-order plots of the association between BAK and BH3 motifs can be found in Fig. 6.9A. The gradient of the line represents k_{on}.

Reactions with mouse BIM were followed by both fluorescence and anisotropy. Kinetic traces collected using either method fitted to a single exponential decay function, and resulted in very similar association rate constants (k_{on}^{fluo} = 2.7(±0.2) 10^6 M^{-1} s^{-1} and k_{on}^{aniso} = 2.4(±0.1) 10^6 M^{-1} s^{-1}). However the intercepts of the lines—which represent the dissociation rate constant of the reaction—do not match. These differences are likely the result of the differential sensitivity of both methods towards additional kinetic events. Anisotropy changes reflect slower tumbling of the dye, which is purely dependent on the formation of the complex. Fluorescence intensity changes are instead related to changes in the local environment of the dye, and these can be the consequence of non-binding events. For example, when shooting t-BIM_{BH3} into buffer, a rate with a very small amplitude can be observed by fluorescence (k_{obs} = 0.9(±0.3) s^{-1}). This is attributed to the dissociation of BIM_{BH3} oligomers, which is a common feature of BH3 peptides with high helical propensities. The same experiment monitored by anisotropy did not show any signal change,

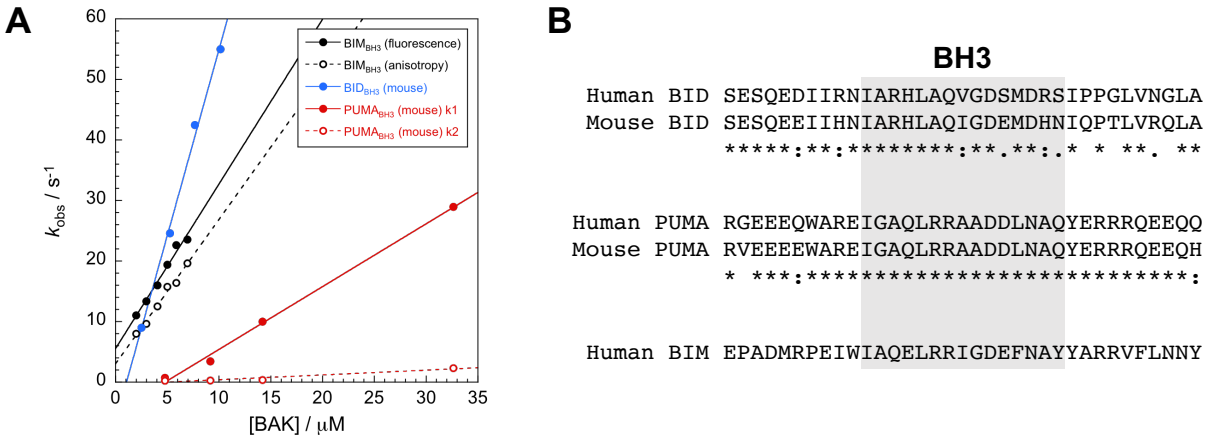

Figure 6.9 – Association kinetics between BAK and TAMRA-labelled BH3 peptides. (**A**) Pseudo-first-order plot with BAK in excess. Reactions were followed by fluorscence intensity change, and BIM was also followed by anisotropy (black dashed line). Association traces were fitted to a single exponential function, except PUMA, which was fitted with a double exponential function. Final concentrations of peptides were 100 nM, 150 nM and 250 nM for BID, BIM and PUMA respectively. (**B**) Sequences of the peptides used for these experiments.

probably due to the lower sensitivity of the technique itself, and the very small size change associated with peptide dimerisation. Because the signal-to-noise ratio of the association between t-BIM$_{BH3}$ and BAK is poor, separate kinetic phases could not be deconvoluted, and the traces appeared single exponential. Therefore k_{obs} from fluorescence will be offset by the presence of the BIM oligomer dissociation rate, explaining the discrepancies in intercepts. However, this rate being independent of the concentration of BAK, the slopes are unaffected, giving confidence in the value obtained for k_{on}. Moreover, estimation of the equilibrium dissociation constant by using these two kinetic dataset yields K_d's of 2.0(±0.4) and 1.3(±0.3) μM from fluorescence and anisotropy respectively, which are close to 1.7(±0.3) μM obtained from the equilibrium binding experiment (Fig. 6.8C).

Reactions between t-BID$_{BH3}$ and t-PUMA$_{BH3}$ and BAK also revealed fast association rate constants (k_{on} = 6.2(±0.3) 10^6 M^{-1} s^{-1} and k_{on} = 1.04(±0.04) 10^6 M^{-1} s^{-1} respectively). Both had negative intercepts, which is physically impossible as it would imply negative dissociation rate constants. These results were attributed to the the poor signal-to-noise ratios of these reactions, and the inherent error associated with extrapolation. This problem was accentuated by the weak affinities of these interactions (*cf.* Table 6.2), which resulted in low amounts of complex formation at the concentrations used for these experiments. Complex formation was estimated at 1–4% for BID, and 2–13% for PUMA. For the association with PUMA, the data fitted to a double exponential function. The

second phase was slow, and appeared almost concentration-independent. Averaging these values yields $k_{\text{obs}} = 0.8(\pm 0.3)$ s^{-1}, which matches the rate observed for PUMA oligomer dissociating when shot in buffer, as well as the rate of BIM oligomer dissociating. This result suggest that the association of PUMA with BAK is not different from the other BH3 motifs, and displays 2-state kinetics as well.

Because of the difficulties of studying weak complexes that associate fast—and the poor signal quality—the absolute values need to be interpreted with care. However, it is clear from comparing the results obtained here that all reactions are fast, and within an order of magnitude of each other. This is similar to the associations between MCL-1 and BH3 motifs (*cf.* Table 6.1). It is striking that the binding of BH3 motifs to BAK are orders of magnitude weaker than to MCL-1 (*e.g.* 6 orders of magnitude for PUMA), yet the speed at which they associate is comparable. The implication is that the entirety of the shifts are accounted for by accelerated dissociation rates, meaning that complexes have extremely short lifetimes. A back-of-the-envelope calculation—using the 2-state approximation and the binding parameters reported—indicates lifetimes for bound structures of BAK and BH3 peptides in the range of 0.1–50 ms. It is interesting to consider that from the motifs investigated, BIM appears the tightest at 2 μM, yet is still an order of magnitude weaker than MCL-1:MCL-1$_{\text{BH3}}$; an interaction supposedly irrelevant under physiological conditions since MCL-1 remains folded with its BH3 motif embedded. Could that suggest that the interaction between BAK and BH3 motifs might have been negatively selected during evolution?

Table 6.2 – Kinetic and thermodynamic profiles of BAK binding BH3 peptides. Dissociation rate constants were obtained assuming 2-state reactions and using $k_{\text{off}} = k_{\text{on}} \cdot K_{\text{d}}$ (for BID and PUMA, mouse and human peptides were assumed to share the same parameters). Errors represent fitting errors.

BH3 peptide	k_{on} / M^{-1} s^{-1}	k_{off} / s^{-1}	K_{d} / μM
t-PUMA$_{\text{BH3}}$	$1.04(\pm 0.04)\ 10^6$	$2.3(\pm 0.2)\ 10^2$	$220(\pm 20)$
t-BID$_{\text{BH3}}$	$6.2(\pm 0.3)\ 10^6$	$1.4(\pm 0.1)\ 10^3$	$230(\pm 20)$
t-BIM$_{\text{BH3}}$	$2.6(\pm 0.2)\ 10^6$	$4.4(\pm 0.9)$	$1.7(\pm 0.3)$
t-BAK$_{\text{BH3}}$	–	–	$1000(\pm 200)$
t-BAX$_{\text{BH3}}$	–	–	$500(\pm 60)$
t-MCL-1$_{\text{BH3}}$	–	–	$5000(\pm 3000)$

6.4.3 BAK and MCL-1

Having investigated the binding between BH3 motifs with both pro- and anti-apoptotic BCL-2 proteins, the interaction between BAK and MCL-1 (if any) remained to be studied. Unlike BH3 motifs—which are disordered peptides—both interacting partners are folded, globular proteins. Moreover, their respective BH3's form part of their cores, and appear inaccessible for interactions (Fig. 6.2). From the K_d measured using BAK_{BH3} and MCL-1_{BH3}, it is clear that if these two proteins are to interact *via* their BH3 motifs, exposure and binding of BAK_{BH3} to MCL-1 is the more likely scenario. This hypothesis is based on the premise that BH3 motifs are the point of interaction between the different BCL-2 protein, and does does not consider additional interactions, or a different binding interface altogether.

As a simple test to see whether these two proteins were capable of hetero-dimerising under standard buffer conditions, analysis by size-exclusion chromatography (SEC) was performed. Since both proteins are about the same size, any stable interaction would lead to a significant shift in both mass and hydrodynamic radius. This would manifest itself by a peak with a smaller elution volume than that of the monomers. The result can be found in Fig. 6.10A.

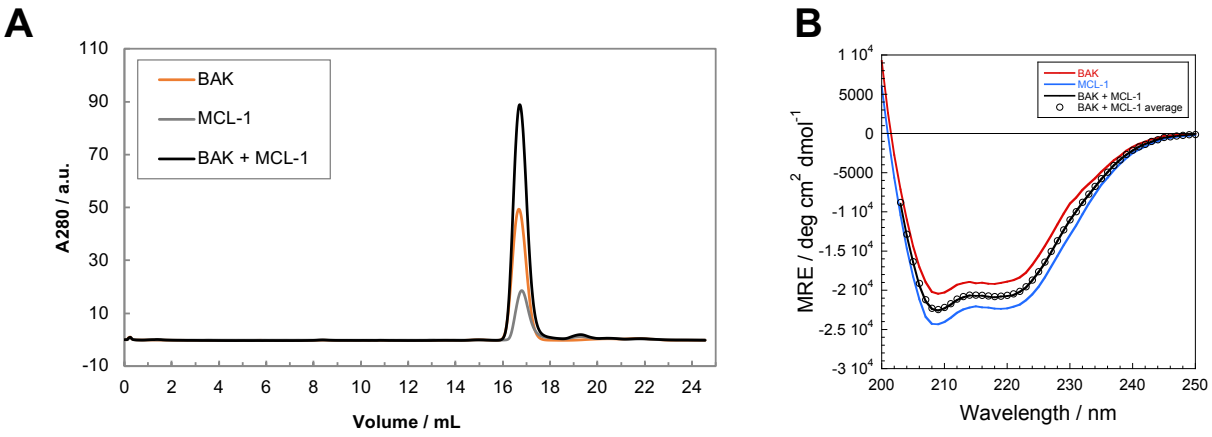

Figure 6.10 – BAK and MCL-1 do not interact in buffer (**A**) SEC analysis of BAK, MCL-1 and their mixture shows no multimerisation after overnight incubation at 25 °C. SEC was performed at 10 μM of each protein on a Superdex 200 10/300 column. (**B**) CD spectroscopy reveals no structural change after overnight incubation at 25 °C. The concentration of each protein was 5 μM.

It is evident that the elution chromatogram of the mix is the sum of its parts, *i.e.* the control experiment when only a single protein was injected. Furthermore, the elution volume of the mixture is 16.7 mL, which corresponds to a globular protein of 21 kDa (*cf.*

column calibration in Materials and Methods). Together, these results demonstrate that both proteins remain monomeric, and that dimerisation or higher-order oligomerisation can be excluded.

CD spectroscopy was also employed to see whether any structural change—which might suggest some form of interaction—could be observed. More specifically, if BAK had to partially unfold in order to bind MCL-1, the mixture would show a reduced helical profile. Fig. 6.10B shows that the secondary structure of the mixture was identical to the average of its parts, clearly indicating an absence of structural change. This result does not eliminate the possibility of an interaction without structural change. However, in combination with the SEC analysis it appears unambiguously that BAK and MCL-1 are incapable of interacting at these concentrations under buffer conditions alone, despite the high affinity of MCL-1 for the BH3 motif of BAK. Evidently, if these two proteins are to interact as part of their function, standard biochemical conditions do not recapitulate this behaviour (*cf.* Chapter 8 for more details).

It should be pointed out that while this lack of interaction might be surprising from a biological standpoint, it is fully comprehensible from a thermodynamic point of view. Assuming the simple case scenario where BAK needs to unfold to expose its BH3, allowing it to interact with MCL-1, the following thermodynamic relationship can be expressed:

$$\Delta G_{\text{bind}} = \Delta G_{\text{BH3}} + \Delta G_{\text{D-N}} \tag{6.1}$$

where ΔG_{bind} represents the binding between BAK and MCL-1, ΔG_{BH3} the binding of BAK_{BH3} to MCL-1, and $\Delta G_{\text{D-N}}$ is the free energy of unfolding of BAK. The lack of hetero-dimerisation implies that the binding free energy of MCL-1:BAK_{BH3} is insufficient to offset the folding free energy of BAK in buffer. ΔG_{BH3} can be calculated from the corresponding binding equilibrium constant, and is worth 13.8 kcal mol^{-1}. A upper bound estimate of ΔG_{bind} can be obtained from the SEC analysis done at 10 μM protein (Fig. 6.10A). Assuming a detection limit of 100 nM (corresponding to 1% conversion), the resulting equilibrium binding constant would need to be 1 mM or higher, corresponding to $\Delta G_{\text{bind}} \leq$ 4.1 kcal mol^{-1}. Therefore, an unfolding free energy of 9.7 kcal mol^{-1} would be sufficient to prevent hetero-dimerisation at 10 μM, which is well within the realm of possible stabilities for a protein of that size. This back-of-the-envelope calculation does not account for additional interactions beyond the BH3 motif, and assumes a complete unfolding of BAK,

which might be incorrect. Nevertheless, it suggests that 'standard' folding stabilities are sufficient to prevent binding under μM concentrations, even for sequence motifs with sub-nM affinities. It is noted that hetero-dimerisation is still possible—complex formation is concentration-*dependent*, while folding is concentration-*independent*—but would require significantly higher concentrations of proteins.

6.5 Interactions within the ternary system

Investigating the binary interactions between the different components of the model tripartite system revealed that BH3 motifs associate fast with both BAK and MCL-1. However, orders of magnitude differences in affinities were observed, which can be attributed to large differences in dissociation rate constants. Consequently, under equilibrium conditions the distribution of states for the ternary mixture should be heavily skewed towards MCL-1:BH3 complexes, with BAK present as a spectator protein. This would be true if the properties of the binary interactions are additive, *i.e.* absence of cooperativity when all components are in the presence of each other. This hypothesis was tested by both SEC and CD spectroscopy using TAMRA-labelled PUMA_{BH3} (Fig. 6.11).

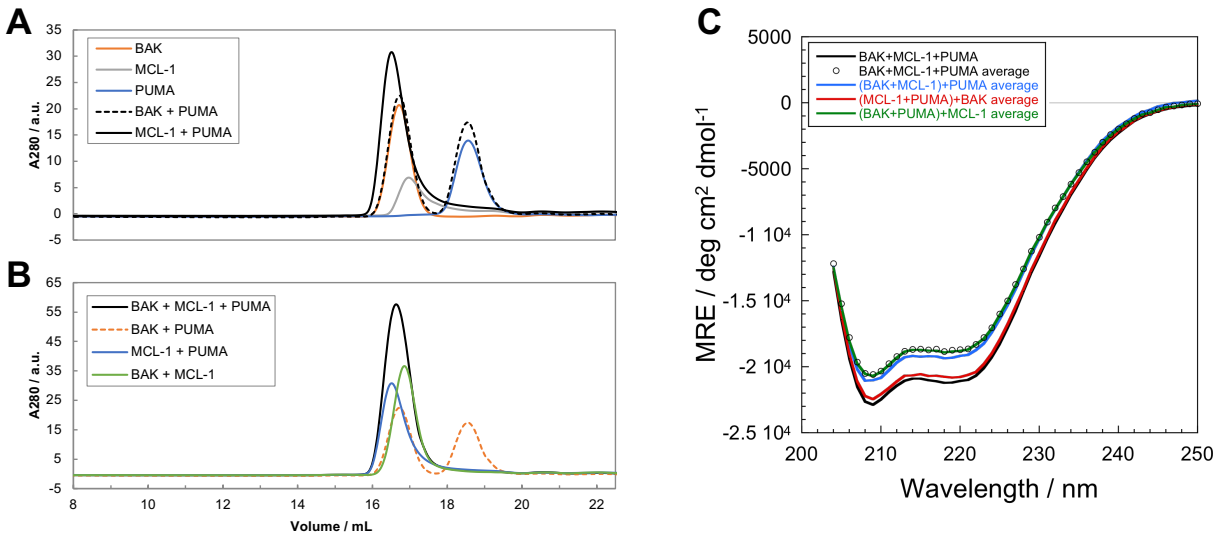

Figure 6.11 – Analysis of interactions within the ternary system composed of BAK, MCL-1 and PUMA. (**A**) Consistent with K_{d} values, the interaction between PUMA and MCL-1, but not with BAK, can be detected by SEC. (**B**) The profile of the ternary mixutre is explained by the interaction of PUMA with MCL-1. SEC was performed on a Superdex 200 10/300 column. (**C**) CD spectroscopy reveals that the entirety of the structural change observed for the ternary mixture can be accounted for by the interaction of PUMA with MCL-1. Protein concentrations were 5 μM, and samples were incubated overnight at 25 °C.

As expected from the knowledge of the equilibrium dissociation constant of BAK:PUMA, no interaction was observed by SEC, and the dye-labelled peptide eluted after the protein peak (18.5 mL). The total lack of binding was confirmed by: *i*) the absence of absorbance at 555 nm (TAMRA) in the protein peak (16.7 mL), and *ii*) SDS-PAGE analysis. In contrast, PUMA forms a complex with MCL-1, which can be isolated by SEC. The free peptide peak at 18.5 mL disappears, and the dye can be observed in the protein peak at 16.5 mL (by absorbance at 555 nm and SDS-PAGE). Furthermore, while free MCL-1 elutes at 17 mL, addition of PUMA shifts the elution to 16.5 mL, consistent with the formation of a larger complex. This again is consistent with the reported K_{d} value for this interaction. When mixing all three proteins together, the SEC profile contains a single peak (Fig. 6.11B). The elution volume is consistent with the absence of multimers of BAK and MCL-1. Furthermore, no free peptide remains—confirmed by the absence of an elution peak at 18.5 mL—which can be explained by its interaction with MCL-1. Analysis by CD spectroscopy (Fig. 6.11C) shows that the ternary system (black line) is more structured than the average of its individual part (black circles). The totality of this structural change can be accounted for by the coupled folding and binding of PUMA with MCL-1 (red line). These results demonstrate that the properties of the binary interactions are additive, and that the behaviour of the ternary mixture can be fully explained by the interaction between MCL-1 and PUMA, with BAK as spectator protein.

6.6 Discussion

This chapter describes the selection and characterisation of a tripartite interaction network composed of members from the BCL-2 family of proteins. MCL-1 was selected as the anti-apoptotic representative due to the extensive biophysical characterisation of its murine counterpart. BAK was chosen for this initially study due to its ease of expression and purification given the large amounts required for some of the experiments described in this chapter. Finally, different BH3 motifs (both from BH3-only proteins, and from folded BCL-2's) were tested for interaction with BAK and MCL-1. While some initial investigations involved dye-free peptides, most of the subsequent work was performed with TAMRA-labelled variants. However, differences between dye-labelled and dye-free peptides were limited, and do not affect the conclusions.

Binary interactions between each member of the model system were characterised, both by equilibrium and kinetic methods where possible. All BH3 motifs investigated interacted fast and tightly with MCL-1, in line with previous work (Ku *et al.*, 2011, Dahal *et al.*, 2018). Association rate constants were within an order of magnitude of each other, and affinities controlled by dissociation rate constants. In stark contrast, none of the BH3 motif investigated 'interacted' with BAK (BIM being the only exception). Affinities were orders of magnitude less tight than with MCL-1, and in the high μM range. This result suggests that in the cell, physiological concentrations of these proteins are unlikely to be sufficient to form any BAK complexes, and instead all BH3's would interact with MCL-1. Interestingly, despite weak binding, associations with BAK (when measured) were as fast as with MCL-1, thus indicating once more than the lifetime of the complex determines its affinity. This also suggest that there is no particular energetic barrier to binding BAK, *e.g.* complicated folding on-pathway to the bound complex. While this result would appear to be in favour of the 'hit-and-run' mechanism (Chipuk & Green, 2008), the lack of resulting oligomerisation determined by SEC does not support this model (*cf.* Chapter 8). Moreover, it is noted that physico-chemical considerations would also disfavour this mechanism. Indeed, the net energetic sum of a binding event followed by dissociation is null, thus the only possibility for a BH3 to trigger oligomerisation would be through a catalytic effect. However this would suggest that BAK oligomers are the thermodynamically favoured state of the protein, and that its conversion is only kinetically limited. Since BAK remained monomeric over months when kept in buffer at 4 °C, this scenario is unlikely (*cf.* Chapter 7).

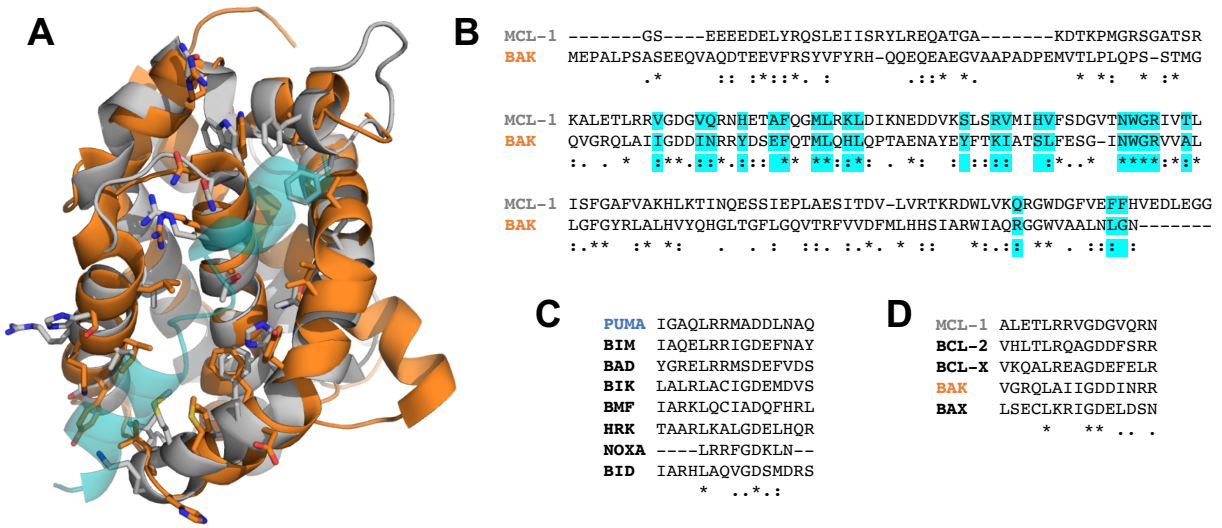

Figure 6.12 – Sequence and structural homolgoy between BAK and MCL-1. (**A**) Structural comparaison between BAK (orange, PDB:2M5B) and MCL-1 (grey, PDB:2MHS). Interface residues of the BH3 groove are highlighted in sticks representation. BID bound to BAK is shown in transparency to aid visualisation of the groove. The RMSD over 131 C_α is 2.76 Å. (**B**) Sequence alignnment of BAK and MCL-1. Interface residues of the BH3 groove are highlighted in cyan. The identites over the full sequence and the interface are 19.9% and 34.8% respectively. (**C**) BH3 motifs from BH3-only proteins. These are part of larger, intrinsically disordered proteins. (**D**) BH3 motifs from anti-apoptotic (MCL-1, BCL-2, BCL-X) and effector (BAK and BAX) BCL-2 proteins. These sequences are part of folded regions within their respective proteins.

The lack of affinity of BAK for BH3 motifs is intriguing. While it is entirely possible that these interactions have not been under selective pressure, the structural homology between BAK and MCL-1 might suggest some 'residual affinity'. While the overall sequence identity of the two proteins is low (19.9%), it is not unlike sequence identities between MCL-1 and other anti-apoptotic BCL-2 proteins, which all bind BH3 motifs tightly (Ku *et al.*, 2011). Also, the residues that compose the groove are more conserved than the sequence as a whole (Fig. 6.12B). Furthermore, BH3 motifs are poorly conserved (Fig. 6.12C, D), yet MCL-1 binds many tightly, suggesting a level of promiscuity. This argument is further supported by the sub-μM affinity of MCL-1 for its own BH3 motif; an interaction that should not have been under selective pressure given its lack of biological relevance. Thus, it may indicate that a simple amphiphatic helix might suffice to achieve reasonable binding affinity. Given the structural and sequence homology of the grooves, higher affinities to BAK would be expected. The fact that it is not the case might suggest a form of negative selection, whereby BH3 affinities for BAK would have been selected *against* during evolution. This might have happened to prevent stabilisation of BAK through protein-protein interactions,

which would have compromised its oligomerisation and function (*cf.* Chapter 7). While it is a plausible hypothesis, more experiments would be warranted to confirm its veracity.

To complete the characterisation of the tripartite system in buffer, the potential interaction between BAK and MCL-1 was investigated. Given the affinity of their respective BH3 motifs to each other, it was anticipated that BAK might bind MCL-1 through its BH3. However, this region being folded, and the hydrophobic interface being buried, unfolding or significant structural rearrangement would be necessary. No interaction between the two proteins could be observed under experimental concentrations; probably a consequence of the unfolding/rearrangement energy that would be required to expose the BH3 motif of BAK. Furthermore, analysis of the full tripartite mixture did not reveal any behaviour that was not the sum of its binary interactions, thus excluding any cooperativity when all three sub-classes of BCL-2 proteins are present.

In summary, studies performed on a simplified network of BCL-2 proteins under standard biochemical conditions (dilute solutions, aqueous buffer) revealed that MCL-1 and BH3 motifs interact tightly, forming long-lived complexes. Under these conditions BAK remains monomeric, and showed an absence of interaction with either anti-apoptotic or BH3-only proteins. This spectator role is not in-line with the known biological function of BAK (and BAX). While the biophysical investigations performed here were informative in providing the initial characterisation of the network, it is clear that buffer conditions do not fully recapitulate the biological behaviour of this system. Thus conditions mimicking this behaviour more closely were investigated, and the results presented in the next two chapters.

Chapter 7

Oligomerisation of BAK and BAX

For this chapter, BAX was studied alongside BAK. Its expression and purification required significant optimisation in order to obtain sufficient amounts for the study. This was the work of Tristan O. Kwan. Native mass spectrometry was carried out in the group of Prof. Carol V. Robinson (University of Oxford) with Dr Kallol Gupta. Negative-stain electron microscopy was performed by Philip Rowell at the University of Leeds under the supervision of Prof. Andrew Wilson and Prof. Thomas Edwards.

7.1 Introduction

BAK and BAX are the biochemical executioners of apoptosis (Dewson & Kluck, 2009, Peña-Blanco & García-Sáez, 2018). When programmed cell death is activated, they form pores in the mitochondrial outer-membrane, which disrupt its integrity. The resulting leakage of mitochondrial inter-membrane components (*e.g.* cytochrome c) activates caspases, leading to cellular demise. Thus, the control of BAK and BAX oligomerisation represents the pivotal point of the regulation of apoptosis by the BCL-2 network (Czabotar *et al.*, 2014, Kale *et al.*, 2017). Despite its importance, the structural and mechanistic aspects of this process are still poorly understood.

More than two decades of biochemical work have allowed a better understanding of the different aspects of this regulation by BCL-2 members. In particular, studies using gene knock-outs have aimed to understand the relationship and selectivity of the different BCL-2 components. However, since BAK, BAX and most anti-apoptotic BCL-2 proteins are membrane-localised prior to the onset of apoptosis, a clear distinction between direct and indirect effects from the presence or absence of certain members is difficult. This has led to

contradictory results in the literature, and the emergence of different models to explain the regulation of BAK/BAX oligomerisation (Czabotar *et al.*, 2014). In the *direct* activation model, BAK/X exist in an inactive state, which upon 'transient' interaction with specific 'activator' BH3-only proteins, undergo oligomerisation. In this model, anti-apoptotic BCL-2 proteins sequester BH3-only proteins, preventing their interactions with BAK and BAX, thus stopping apoptosis. In contrast, the *indirect* activation model stipulates that BAK and BAX are constitutively 'active', and require anti-apoptotic proteins to impede their natural propensity to oligomerise. Here, BH3-only proteins displace sequestered BAK and BAX, freeing them from their 'chaperones', promoting apoptosis by allowing the spontenous pore-formation of BAK and BAX. Furthermore, 'unified' (Llambi *et al.*, 2011), 'embedded together' (Leber *et al.*, 2007), and 'interconnected hierarchical' (Chen *et al.*, 2015) models have also been proposed. Differing in their specifics and complexities, they are fundamentally combinations of the direct and indirect models; the dominant pathway is circumstantial, and depends on the BCL-2 proteins under consideration.

While these models all explain the macroscopic features of cellular apoptosis, they are fundamentally different at the molecular level. Their descriptions of the interactions between anti-apoptotic and BH3-only proteins overlap, but not that of the behaviour of BAK and BAX. Indeed, in the direct activation model, BAK and BAX are stable, monomeric proteins which require 'activators' to start oligomerising, while in the indirect activation model their oligomerisation is spontaneous, and needs to be, prevented. These models are in opposition from a thermodynamic standpoint. It is noted that the process of BAK/BAX oligomerisation is not known to require any form of cellular energy.

Structurally, BAK/BAX and their oligomers are less well understood than their anti-apoptotic counterparts. Compared to the number of structures of anti-apoptotic BCL-2 proteins—in complex or not with BH3 peptides—structural information about BAK and BAX oligomers are crucially lacking (Cosentino & García-Sáez, 2017). It is perhaps surprising to consider that both proteins have had their monomeric state characterised, showing the typical soluble, monomeric and globular BCL-2 fold (Suzuki *et al.*, 2000, Moldoveanu *et al.*, 2006), yet their ultimate structural state is a membrane-embedded oligomer. How this change occurs is not fully understood (Peña-Blanco & García-Sáez, 2018). Structures of helix-swapped dimers of both BAK and BAX have been reported (Czabotar *et al.*, 2013, Brouwer *et al.*, 2014). However, the authors claim them to be off-pathway to the apoptotic oligomers. Instead, the consensus in the field is that both

proteins homo-oligomerise *via* their BH3 motifs, whereby the BH3 motif of one BAK (or BAX) molecule is exposed and binds to the groove of another one (Dewson *et al.*, 2008, 2012). These dimers then assemble into larger oligomers. Tentative structures of these dimers have also been reported for both BAK and BAX (Czabotar *et al.*, 2013, Brouwer *et al.*, 2014). However, severe truncations—and fusion to GFP—were required to crystallise theses homo-dimers. Therefore, it appears unclear whether these structures retain validity in representing the postulated first oligomeric step of pore-formation.

Alternative techniques have also reported that BAK and BAX oligomers are based on units of dimers (Subburaj *et al.*, 2015), thus adding weight to this hierarchical architecture description, but the structure of both the dimeric unit and the pore as the whole appears unclear. It is noted that time-lapsed microscopy studies have revealed that in their cellular context, BAK and BAX pores grow to very large structures, even allowing herniation of the mitochondrial matrix (McArthur *et al.*, 2018). Clearly, these pore structures appear heterogeneous, and suggest a level of plasticity to the oligomerisation of BAK and BAX. This contrasts with classic structural biology expectations of symmetry and well-resolved assembly states.

7.2 Aims

In Chapter 6, investigation of BAK under simple buffer conditions revealed the absence of oligomer formation, neither spontaneously, nor in the presence of peptides of BH3 motifs. In appeared that BAK was a acting like a 'spectator' molecule, in contradiction with its known biological role. An understanding of the network and its regulation requires a grasp of oligomerisation itself, which is the focus of this chapter. Using detergents to mimick a membrane-like environment, oligomerisation of BAK and BAX was achieved *in vitro*. This process was spontaneous, and did not require the presence of any other BCL-2 protein. The specificity of this oligomerisation to pro-apoptotic BAK and BAX is demonstrated, as well as the physiological relevance of resulting detergent-treated oligomers. Disulfide stapling was also employed to gain insight into the topology of the complex. Finally, kinetic investigation of the oligomerisation process suggests that a detergent-induced structural change is the rate-limiting step to the assembly, thus providing a mechanistic insight into the conversion of soluble monomeric proteins to membrane-embedded oligomeric structures.

7.3 Oligomerising BAK and BAX in detergents

7.3.1 The difficulty of making monomeric BAK and BAX

Constructs of human BAK and BAX were made to match available structures. Human BAK (UniProt:Q16611 16–185, C166S) and human BAX (UniProt Q07812 1–171, C62S/C126S) both had their C-terminal trans-membrane helix removed, matching the structures, and most of the literature on these proteins. Indeed, presence of these helices can dramatically reduce solubility, and often leads to aggregation. Since the cysteines do not form disulfide bridges, they were mutated to serines to avoid the use of reducing agents. Expression and purification of BAK and BAX were initially attempted as N-terminal His-tagged versions of the proteins; an adaptation of the protocol for making MCL-1. However, these proved unsuccessful with an almost complete lack of expression. Both proteins were sub-cloned into the pTXB1 vector, which possesses a C-terminal intein, followed by a chitin binding domain (CBD). It was hopped that the large tag (27 kDa) would aid expression and solubility, and that the protease-free tag removal strategy would reduce potential proteolytic degradation. After optimisation of the expression conditions, it was possible to obtain both constructs in sufficient amounts for the intended study. Typical yields after purification were 10 mg and 1.5 mg (per litre of culture) for BAK and BAX respectively.
Despite their structural similarities with anti-apoptotic BCL-2 proteins, BAK and BAX proved relatively difficult to express and purify. Moreover, obtention of monomers was more challenging. During the purification process, significant amounts formed higher-order oligomers species which eluted at the void volume on a Superdex 75 column (Fig. 7.1). Given the fractionation range of this resin, these species must have apparent masses of 70 kDa or more; corresponding to tetramers, or higher oligomeric species. Furthermore, the entire purification process had to be performed in the strict absence of detergent, or almost no monomer could be recovered.
In contrast, MCL-1 (as well as BCL-X_L and A1, which have been studied by other members of the group) was readily obtained as monomer (Fig. 7.1C), and in good yield. Furthermore, presence or absence of detergents in purification buffers did not affect the outcome. Evidently, it is impossible to make any firm conclusion based on this observation alone; vectors, tags, expressions and purification conditions being different, the number of variable and their effects cannot be systematically assessed. However, it is interesting to notice the differences in behaviours between pro- and anti-apoptotic proteins during these

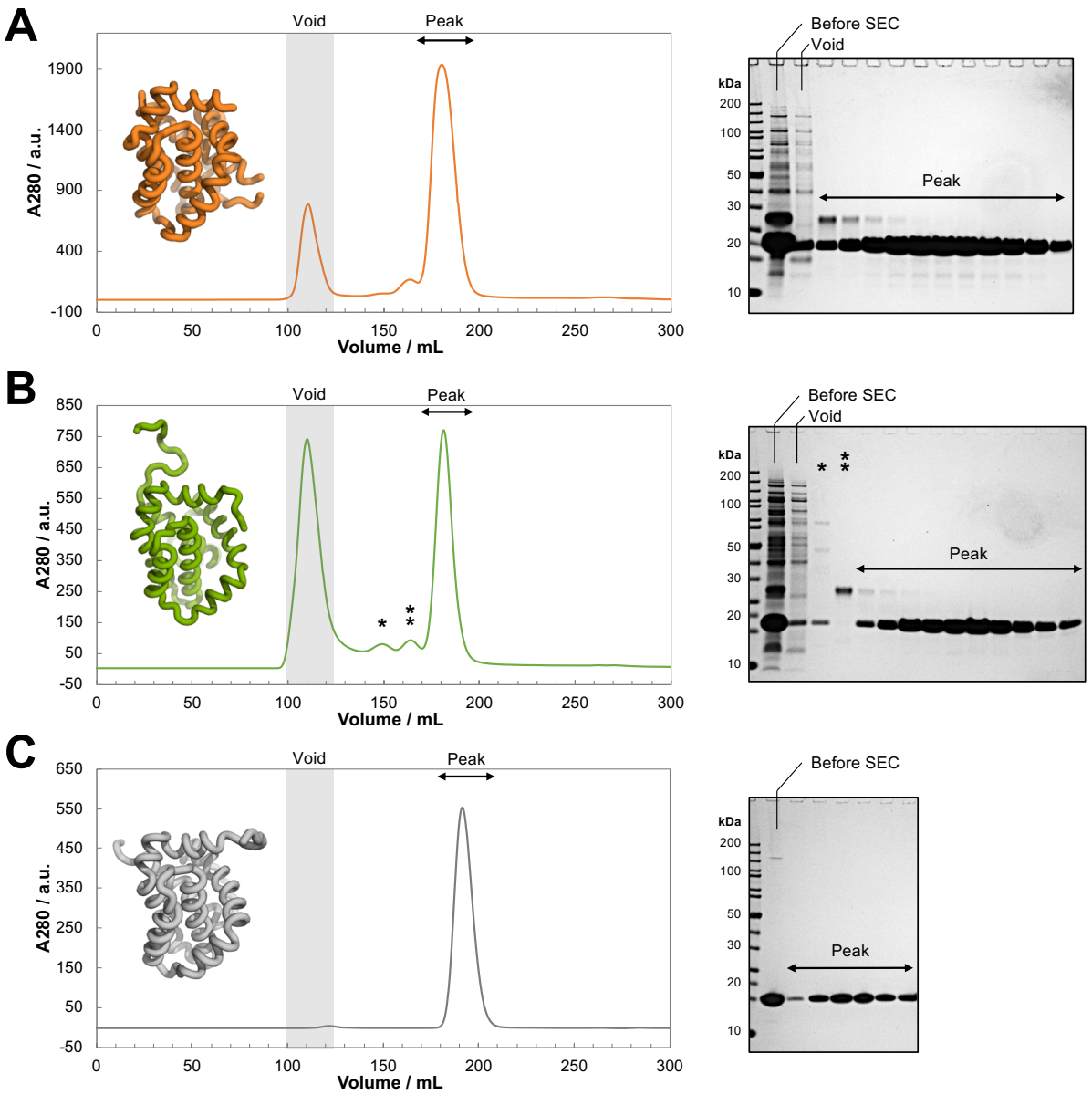

Figure 7.1 – Preparative size-exclusion chromatography (SEC) of BAK, BAX, and MCL-1. (**A**) SEC of BAK from a 4 L preparation. (**B**) SEC of BAX from a 8 L preparation. (**C**) SEC of mouse MCL-1 from a 2 L preparation. Note that for BAK and BAX, a chitin pull-down step was added after SEC to remove the intein-CBD contaminant (27 kDa). SEC was perfomed on a Superdex 75 26/600 column equilibrated in 50 mM sodium phosphate pH 7.0 at room temperature. Shaded area represent the void volume of the column. SDS-PAGE analyses of the fractions are indicated to the right of each chromatogram. Structures of the proteins are shown for reference (BAK, PDB:2YV6; BAX, PDB:1F16; MCL-1, PDB:1WSX).

preparative steps. Could they be a reflection of their biological function? Could BAK and BAX be more sensitive to bacterial lipids during *e.g.* sonication because of their roles as pore-forming proteins? A more systematic study using detergents and monomeric proteins was therefore undertaken.

7.3.2 Detergent screening

The known biological function of BAK and BAX is their ultimate assembly into pores at the mitochondrial outer-membrane (MOM). Biological membranes are complex. Constituted of different lipids and proteins, there composition can vary as a function of cell cycle, or in response to stimuli. Moreover, the presence or absence of specific lipids can affect and/or regulate membrane proteins directly (Singer & Nicolson, 1972, Laganowsky *et al.*, 2014, Gupta *et al.*, 2017). Understanding the effect of this complexity on the oligomerisation of BAK and BAX was beyond the scope of this work. However, in order to gain insights into the oligomerisation process, it was necessary to obtain a simplified system capable of mimicking a membrane environment. Detergents have often been used in the investigation of membrane proteins, aiding solubilisation and structural studies (Garavito & Ferguson-Miller, 2001, Seddon *et al.*, 2004). It has been noticed early in the study of BAK and BAX that some detergents are capable of inducing dimerisation and expose epitopes consistent with apoptotic pores, thus validating this approach (Hsu & Youle, 1997).

Different detergents were tested for their abilities to oligomerise monomeric BAK or BAX. Detergents are classified as either ionic, non-ionic, or zwitterionic, based on the nature of their head-group. To ensure downstream compatibility with mass spectrometry analysis, only non-ionic detergents were employed. The molecular structures of the detergents investigated is illustrated in Fig. 7.2.

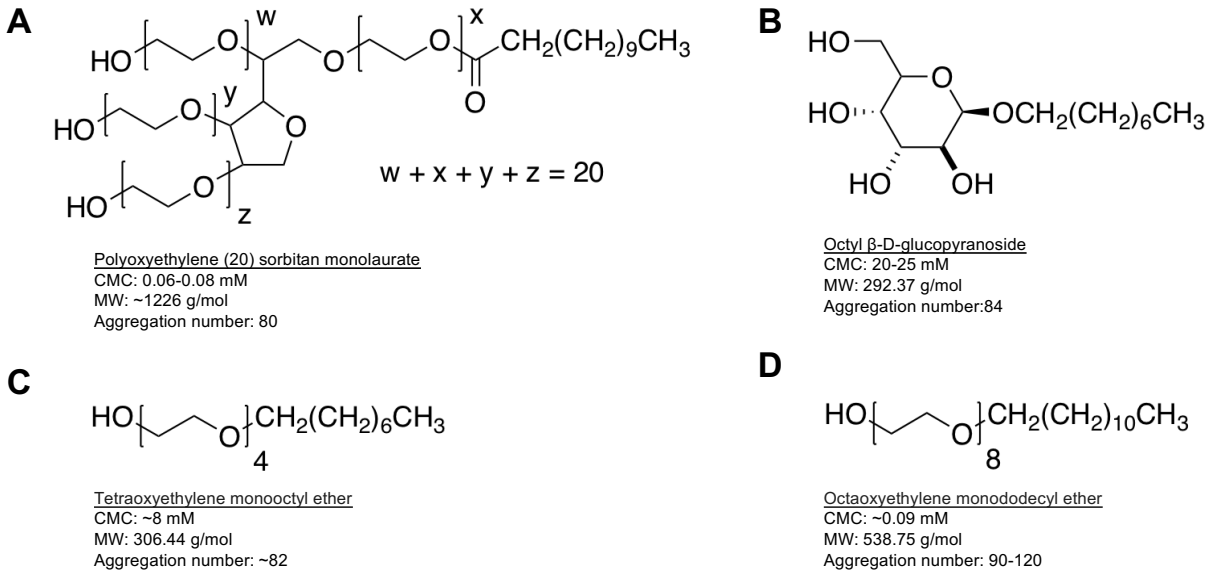

Figure 7.2 – Structures and properites of the detergents used for the oligomerisation of BAK and BAX. (**A**) PS20, (**B**) OGP, (**C**) C8E4, (**D**) C12E8.

Owing to their amphiphatic character, detergent molecules can self-assemble into micelles. This process is concentration-dependent, and depends on the physico-chemical properties of the solution and the detergent itself. Due to the highly cooperative nature of the assembly, the transition is sharp, and happens over a very narrow concentration range. This threshold value is called the critical micelle concentration (CMC). Addition of detergent above this value affects mainly the micelle concentration, while the monomer concentration remains mostly constant. Only micelles are capable of solubilising membrane proteins, thus it is important to work at concentrations above the CMC (Garavito & Ferguson-Miller, 2001). For the detergent investigated here, CMC's span multiple orders of magnitude (values indicated in Fig. 7.2), hence the concentrations used ranged from µM to mM.

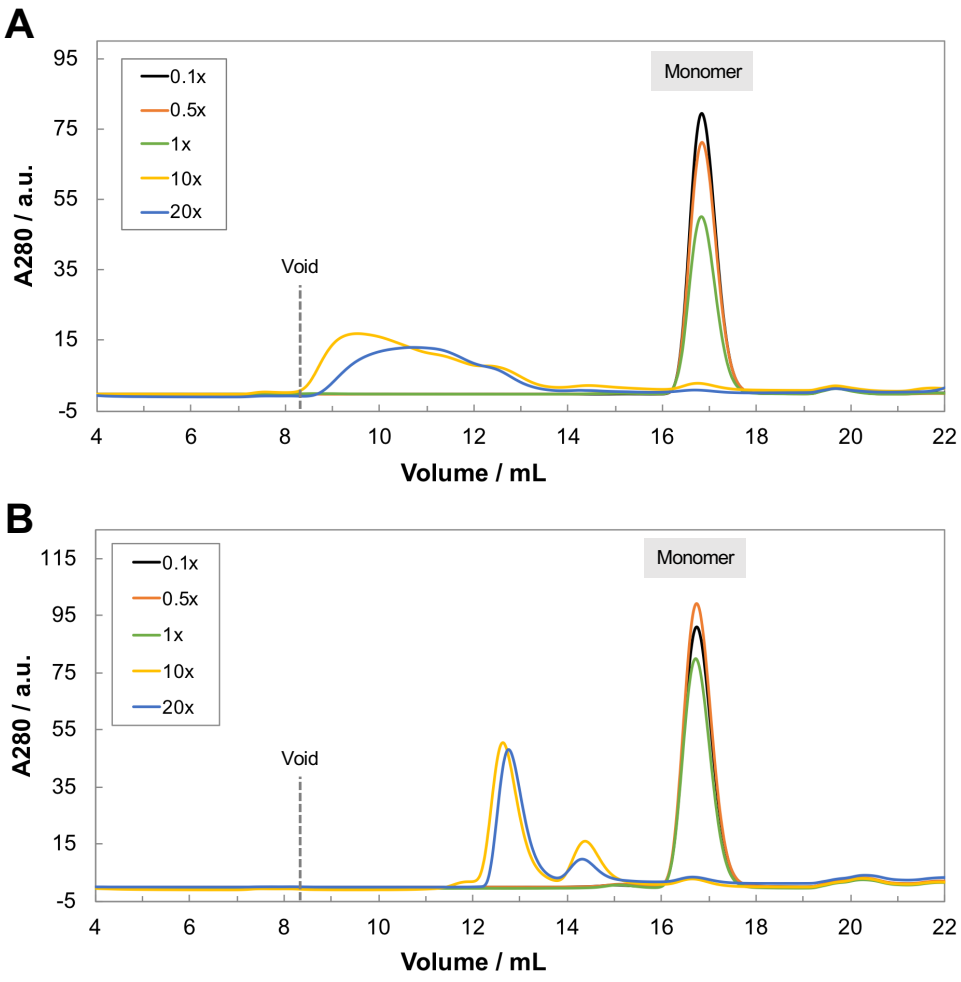

Figure 7.3 – Example of detergent titration for C12E8. Monomer depletion started ~1×CMC. For this detergent, the nature of the oligomer(s) formed depends on the protein. (**A**) BAK at a concentration of 20 µM. (**B**) BAX at a concentration of 24 µM. The concentration of C12E8 is indicated as a multiple of its CMC in the legend. SEC was performed on a Superdex 200 10/300 equilibrated in 100 mM ammonium acetate pH 7.0. Protein concentrations refer to the concentration of application.

Monomeric BAK or BAX were incubated with detergents at different values of their CMC's, and the results analysed by size-exclusion chromatography (SEC). Incubations and chromatography were performed at room temperature or 25 °C. All analyses were carried out on a Superdex 200 10/300 column, and samples were eluted at a flow rate of 0.75 mL/min. Details about the calibration of the column can be found in Materials and Methods. For both BAK and BAX in combination with any detergent, the proteins remained monomeric in the presence of sub-CMC concentrations. Oligomerisation to larger species started to occur from ∼1×CMC onwards (Fig. 7.3), consistent with the formation of 'membrane-like' species. Indeed, it demonstrates that the presence of micelles is necessary, and excludes an event triggered by the detergent molecules themselves. When subjecting mouse MCL-1 to the same treatment, the protein remained monomeric (Fig. 7.5C). These results support the notion that the oligomerisation induced by detergents is specific, and relates to a membrane-associated state only accessible to pro-apoptotic BCL-2 proteins. This detergent-sensitivity can therefore be regarded as reflecting the biological function of BAK and BAX.

While all detergents depleted the monomeric states of BAK and BAX, the nature of the resulting oligomer(s) was highly dependent on both the protein and the detergent. For BAK, all but PS20 lead to the formation of large oligomeric states, which eluted at, or close to the void volume of the column (Fig. 7.5A). On a Superdex 200 matrix, this corresponds to an apparent mass of ≥600 kDa, *i.e.* ≥30-mer. The lack of defined Gaussian distributions—and the presence instead of trailing peaks—suggests a absence of homogeneity, most likely aggregate-like assemblies. In contrast, PS20 allowed the formation of a well-defined peak at ∼12.2 mL, corresponding to an apparent mass of 160 kDa (an octamer). This value is only indicative. Indeed, the relationship between elution volume and mass is based on a calibration using globular proteins (see Materials and Methods). Since separation by SEC is based on hydronamic radii, non-globular proteins/oligomers—which have larger hydronamic radii than globular species of the same mass—will elute earlier, resulting in an over-estimation of their masses. Moreover, the presence of detergent might affect the elution profile. Because BAK and BAX are expected to form pore-like oligomers, theses shapes would significantly affect their hydronamic radii and resulting elution volumes. Therefore, it is not appropriate to use SEC to assess the oligomeric state of the protein. However, it reasonable to conclude that this species is larger than a dimer, which (if globular) would elute at ∼15 mL.

An estimate of the detergent:protein stoichiometry was obtained for BAK and PS20. Since most detergents have CMC values in the mM range, their concentrations is usually not limiting under normal working concentrations ($\geq$ 1×CMC). In contrast, PS20 having a very low CMC value (80 µM), it concentration might potentially become limiting under µM concentrations of protein. This would be the case if a fixed stoichiometry of detergent:protein was required to stabilise the oligomer. This hypothesis was tested by titrating PS20 over a 10-fold range (5–50×CMC, 0.4–4 mM) onto a constant concentration of BAK (50 µM). The results can be found in Fig. 7.4. Under these conditions, the threshold for complete conversion to oligomer was not reached until 25–30×CMC (1.25–1.5 mM). Therefore, a minimum of 40–50 molecules of PS20 per BAK monomer appears necessary to form the oligomer.

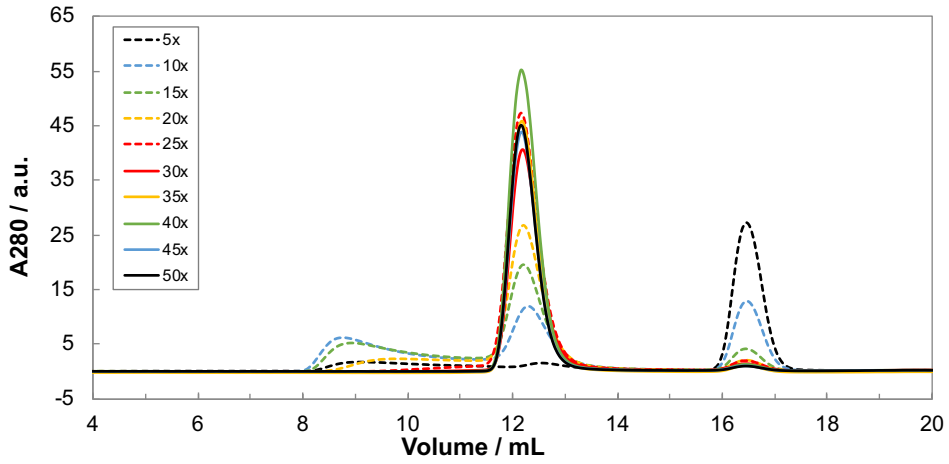

Figure 7.4 – Estimation of the ratio of PS20 to BAK for complete conversation. BAK concentration was 50 µM and the detergent added as multiple of its CMC (indicated in the legend). Pre- and post-threshold concentrations are indicated as dashed and solid lines respectively. SEC was performed on a Superdex 200 10/300 equilibrated in 50 mM sodium phosphate pH 7.0.

In comparison to BAK, the behaviour of BAX in the presence of C12E8, C8E4, and OGP was markedly different. PS20 lead to the formation of a single peak eluting at ~11.9 mL (apparent mass 180 kDa, 9-mer), comparable to the oligomer of BAK in the same detergent. The other detergents did not result in singly-defined oligomeric states, and instead indicated the presence of multiple individual peaks (Fig. 7.5). These profiles were different from the trailing peaks observed for BAK, suggesting a range of well-defined structural states, not an unspecific aggregate-like behaviour. Some identical states were obtained using different detergent— indicated by peaks at similar elution volumes—supporting the notion that defined protein-specific oligomeric states were generated. Interestingly, the larger

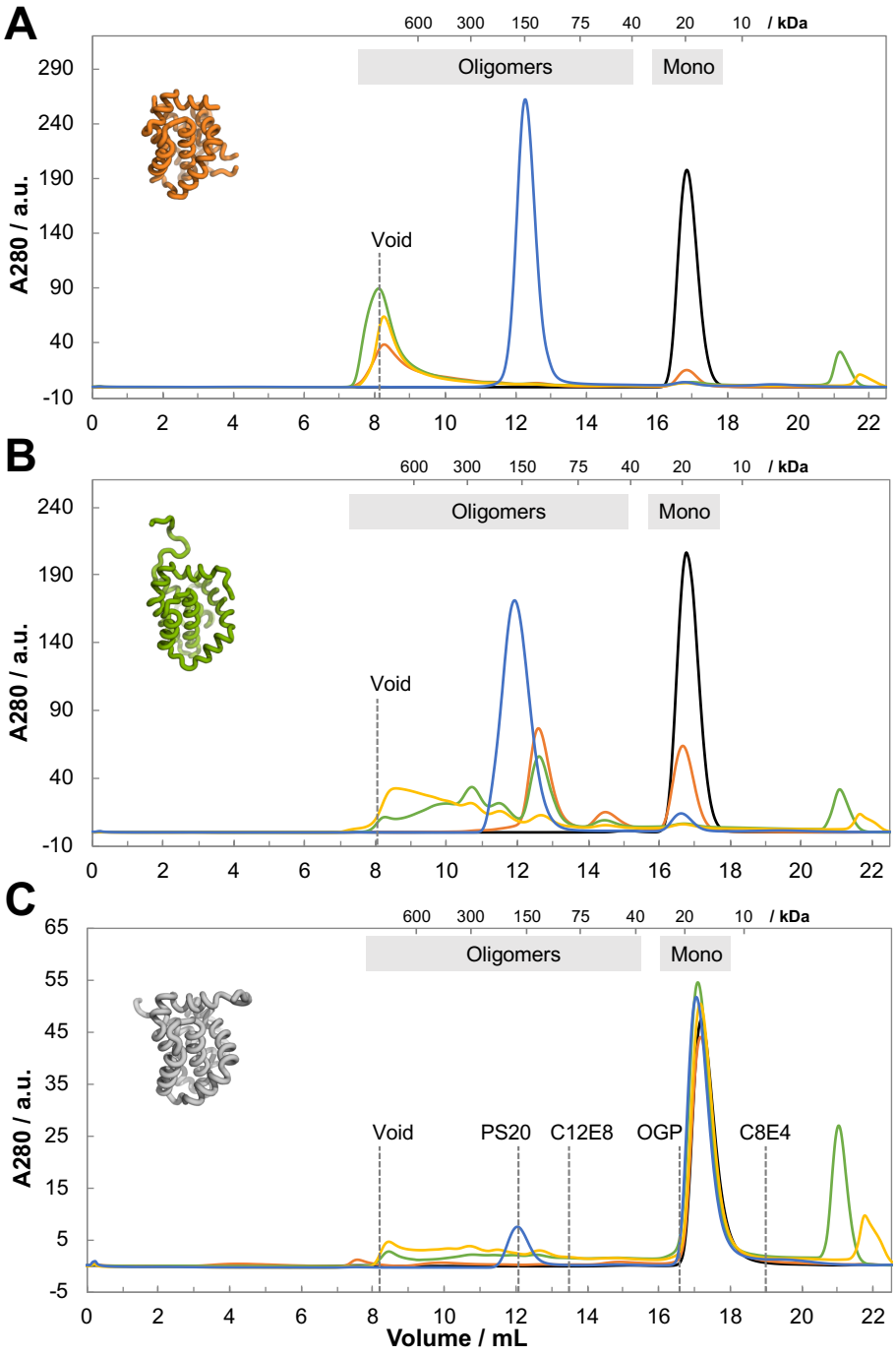

Figure 7.5 – Effect of detergents on pro-apoptotic BAK and BAX, and anti-apoptotic MCL-1. The different conditions were: no detergent (black); C12E8 at 5×CMC (orange); C8E4 at 2×CMC (green); OGP at 2×CMC (yellow); PS20 at 75×CMC (blue). (**A**) BAK at 50 µM. (**B**) BAX at 50 µM. (**C**) mouse MCL-1 at 20 µM. The elution volumes of the free micelles for each detergent is indicated for reference. Note that the peak for PS20 in **C** is solely due to the absorbance of the detergent, and does not contain any protein (confirmed by SDS-PAGE). Instead of oligomerisating it, the smaller detergents OGP and C8E4 appeared to partially unfold MCL-1 (indicated by anomaleously late elution peaks after the monomer). Experiments were perfomed on a Superdex 200 10/300 equilibrated in 100 mM ammonium acetate pH 7.0. Some masses (determined by calibration) are reported above each plot for reference.

detergents PS20 and C12E8 appeared to produce fewer oligomers, while the smaller C8E4 and OGP gave rise to more heterogeneous distributions.

It is worth mentioning that the elution profiles observed for some detergents might be to some extent a consequence of detergent removal. All SEC experiments were performed with the column equilibrated in buffer containing no detergent. For all but PS20, elution volumes of the free micelle were different from that of the oligomers (Fig. 7.5). Therefore, if interactions with the detergents are weak, separation from the protein might occur during the elution. This might have the consequence of destabilising the oligomer and breaking it into smaller components, hence explaining the presence of multiple oligomeric species. Indeed, when re-analysing BAX peaks isolated after C12E8 treatment, break-down to smaller units was observed.

The effect of detergent removal was tested by pre-forming the oligomers in detergent, followed by addition of BioBeads (Bio-Rad). These porous polystyrene beads have a high adsorbing capacity for organic and hydrophobic compounds, thus being useful in removing detergents from aqueous solutions. Fig. 7.6 shows the results for BAK and BAX with either PS20 or C12E8. By comparing the results between detergent-treated, and detergent-treated followed by addition of BioBeads, the stability of the oligomers in the absence of detergents can be assessed. Moreover, it provides a test for the reversibility of the oligomerisation reaction. For BAK, partial reversibility after removal of PS20 was observed, indicated by a larger monomer peak after treatment with BioBeads. The same procedure with C12E8 appeared completely irreversible, denoted by the absence of any monomer after removal of detergent. Instead, the protein eluted entirely with the void volume, most likely due to complete aggregation. For BAX, removal of PS20 appeared to have little effect. However, little conversion of the monomer occurred in the first place, potentially explaining this result. In contrast to BAK, the effect of C12E8 on BAX appeared partially reversible, noticeable by the reduction of oligomer peaks—and an increase in the monomer peak—following BioBeads treatement. These results indicate that: *i*) the oligomers of BAK and BAX are not stable in the absence of detergent, and *ii*) the reaction is partially reversible in some cases. For the irreversible cases, exposure of hydrophobic parts of the proteins following detergent removal most likely leads to non-specific interactions. If these occur faster than refolding to the monomeric state, aggregation is the most likely outcome. Together, these screening results demonstrate the feasibility of oligomerising BAK and BAX in detergent micelles. This sensitivity to a membrane-like environment appears spe-

cific to these anti-apoptotic proteins, in line with their biological roles. Surprisingly, a level of heterogeneity was observed (at least for BAX). Whether this reflects the intrinsic property of the protein, or the result of different detergent:protein interactions is unclear. PS20 giving the most homogeneous data for both BAK and BAX, subsequent work mostly focused on this detergent.

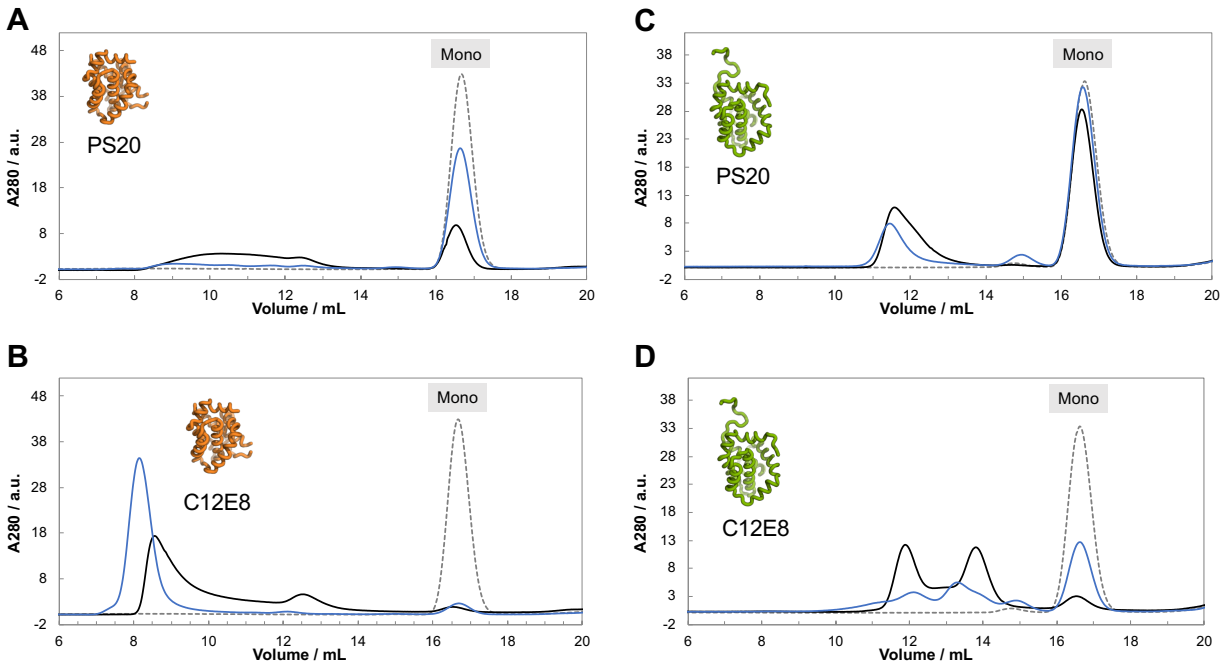

Figure 7.6 – Effect of detergent removal. Proteins (10 μM) were incubated in the presence of either PS20 or C12E8 (5×CMC). Detergent removal was achieved by incubating the oligomers with BioBeads (blue) and compared to the samples where the detergent was not removed (black). The hydrophobic surface of the BioBeads did not have any effect on the monomeric proteins (dashed grey line). (**A**) and (**B**) effect of PS20 removal from BAK and BAX respectively. (**C**) and (**D**) effect of C12E8 removal from BAK and BAX respectively. SEC was performed on a Superdex 200 10/300 column equilibrated in 50 mM sodium phosphate pH 7.0.

7.3.3 Cross-linking studies

In the case of PS20, the elution volumes of BAK and BAX oligomers were close to that of the free micelles. This raises questions regarding the nature of the species observed by SEC; could it be an unfolded monomer in a large micelle? CD spectroscopy revealed that the oligomer remained folded. Besides a slight structural change, this indicated that the detergent did not unfold the protein. These results are described in more details in Section 7.5. The presence of oligomers in detergent-treated samples was assessed by chemical cross-linking.

A dual system composed of BS3 (bis(sulfosuccinimidyl)suberate) and EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) was employed to increase the chances of capturing larger species, which are inherently harder to cross-link. EDC reacts with carboxyl moities (C-terminus, D/E), forming an *O*-acylisourea intermediate that can then undergo nucleophilic attack by an deprotonated amine (N-terminus, K). This reaction forms new amide bonds between carboxly groups and amines, incorporating none of the atoms of EDC. Hence, it is often termed a 'zero-length' cross-linker. However, the absence of a spacer group requires the two residues to be in very close proximity, thus reducing the efficacy of EDC on its own. This shortcoming was mitigated by using BS3, a *N*-hydroxysulfosuccinimide (NHS) esters with a spacer arm of 11.4 Å. This homo-bifunctional reagent is specific towards nucleophilic lysines, as well as *O*-acylisoureas (EDC-activated carboxyl groups). Therefore, using it in conjunction with EDC allows cross-linking of aspartate and glutamate residues with lysines that are otherwise too far apart to be reacted together. Note that BS3 is the sulfonyl derivative of the more commonly used disuccinimidyl suberate (DSS) cross-linker. This modification makes it water-soluble, thus avoiding the use of organic solvents, which could have interfered with the detergent micelles.

BAK and BAX were oligomerised in the presence of either PS20 or C12E8, followed by the addition of cross-linker from fresh stocks prepared in buffer. The reactions were optimised by varying the protein:EDC:BS3 ratios, as well as the amount of detergent present. This was necessary to find the right balance between intra- and inter-molecular reactivity, as well as ensuring proper oligomerisation without interfering too much with the chemistry. Overall, cross-linking in the presence of either detergent revealed higher order species, thus confirming that the SEC peaks do contain oligomers, and not simply an 'micelle-encapsulated' folded monomer. Up to hexamers could be detected for BAK in either PS20 or C12E8. For BAX, tetramers were the largest species that could be observed.

Interestingly, despite containing only one lysine residue (and the N-terminus), BAK appeared to cross-link better than BAX, which contains far more lysines. Cross-linking of BAK in either PS20 or C12E8 revealed similar outcomes, while better results were obtained for BAX in C12E8. These results might reflect differences in accessibilities of reactive residues for the oligomeric state of each protein, as well as a detergent-specific interference with the reactions.

Surprisingly, the control experiment in the absence of detergent revealed a small amount of dimer for BAK, despite this species not being observed by SEC. However, this result

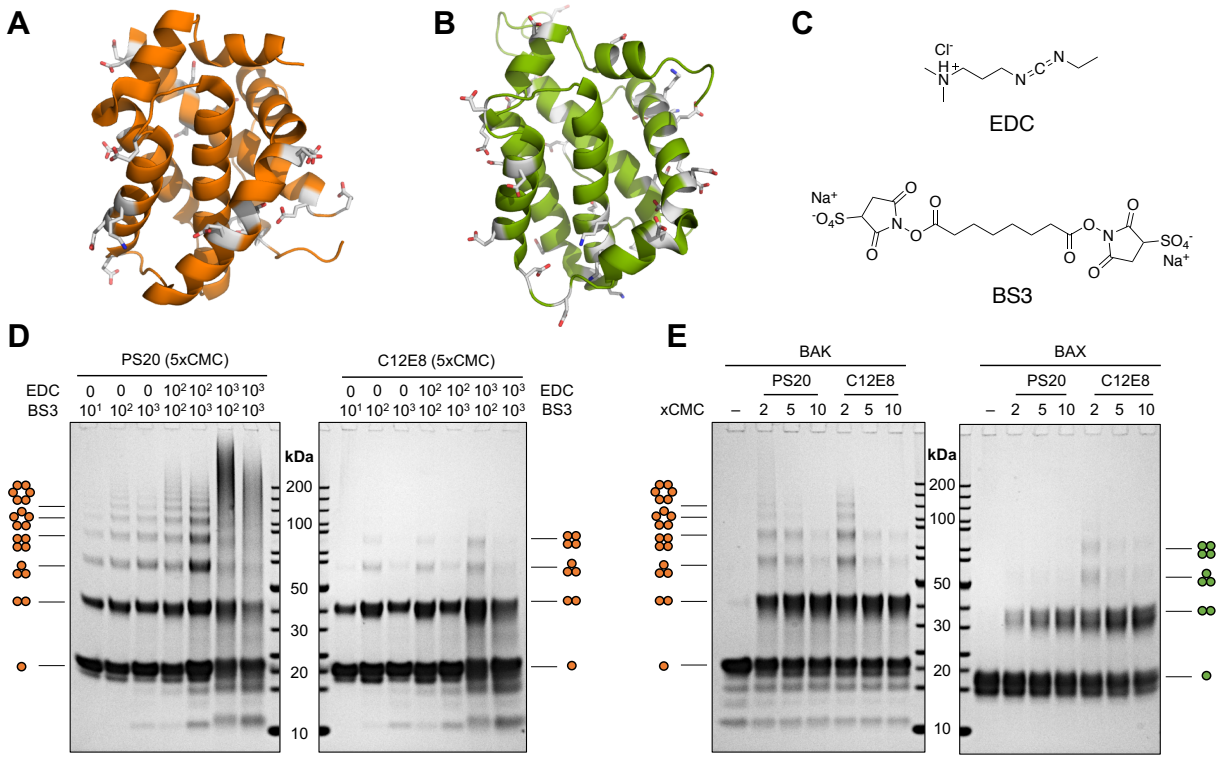

Figure 7.7 – Chemical cross-linking of BAK and BAX oligomers. (**A**) Structure of BAK (PDB:2YV6) highlithing D/E/K residues (6/13/1). (**B**) Structure of BAX (PDB:1F16) highlithing D/E/K residues (13/10/7). Note that the termini are also reactive. (**C**) Chemical structures of the cross-linker used for this study. (**D**) Effect of varying the protein:EDC:BS3 ratio at constant detergent concentration. BAK concentration was 25 µM and the molar excess of each cross-linker indicated above the lane. (**E**) Effect of varying detergent concentration at constant protein:cross-linker ratio (1:300:300). BAK and BAX concentrations were 20 µM, and detergent concentrations are indicated as multiples of their CMC's above each lane.

can be explained by the fact that chemical cross-linking is an out-of-equilibrium technique. Indeed, the dimer might exist in very low abundance at equilibrium, at a concentration too low to be detected. Cross-linking the dimer shifts the monomer-dimer equilibrium, and over the course of the reaction (2 hours) more and more dimers are 'pulled out' of the equation. Therefore, this species appears more populated than it would under equilibrium conditions, thus explaining the result. Another artefact of chemical cross-linking is the presence of bands lower than the monomer. While these could be the result of chemical degradation of the protein, a more likely explanation is intra-molecular cross-linking. This results in anomalous band migration, due to the protein chain not linearising fully under denaturing conditions.

Analyses of PS20-induced oligomers showed single peaks by SEC, but cross-linking of these species reveals multiple bands. This heterogeneity appears in contradiction with the

results obtained by chromatography, however it is most likely another artefact of chemical cross-linking. Taking the hexamer as an example, at least five independent reactions are required to covalently link all subunits together. The likelihood of reacting a new cross-linker molecule on an oligomer that has already been cross-linked four times is lower than this cross-linker reacting with another (less reacted) oligomer. Assuming the simplest case of independent probabilities, fully cross-linking an oligomer containing n subunits scales as $(n-1)^P$, where P represents the probability of cross-linking. Therefore, the larger the oligomer, the less likely it is to be fully cross-linked. Subunits that are not covalently modified would then be lost upon running denaturing SDS-PAGE, and the oligomer would appear smaller. This is also consistent with the observation that the larger the oligomer, the fainter the band.

Since chemical cross-linking is not stoichiometric, the distribution of bands does not necessarily reflect the solution-state of the oligomer. So while these results do not confirm the exact nature of the species observed by SEC, it does prove that assemblies up to hexamers do exist for BAK treated with PS20. Tetramers were confirmed for BAX. However, this might be due to the lower reactivity of this protein, the interference from detergents, or the lower conversion to oligomers at the detergent concentrations used in cross-linking experiments. These questions were further investigated in Section 7.4.

7.3.4 Physiological relevance of detergent-treated oligomers

The results presented in the previous sections demonstrate the feasibility of oligomerising BAK and BAX with detergents. This behaviour appears specific to these anti-apoptotic proteins, thus suggesting a link to their biological roles. However, detergents are not membranes. They are not laminar, but form spherical micelles instead. This discrepancy can affect detergent:protein interactions. Moreover, the composition of biological membranes—which includes various types of lipids, with differences in both head groups and hydrophobic chain—is not recapitulated. For these reasons, it is important to demonstrate that the detergent-treated state does capture the known biological function. This is typically achieved using some form of functional assay.

For an enzyme, that might be showing retained catalytic function in the presence of the detergent. For a 'structural' protein such as the case at hand, it is less straightforward. One possibility would have been to use dye-release assays from liposomes. In these experiments,

dye-labelled dextran molecules are encapsulated in liposomes formed of MOM lipids. BAK and BAX are then added to the liposomes, and the release of the large dextran molecules monitored by fluorescence, indicating permeabilisation. However, this would have been difficult to perform in the context of the work of this thesis. Indeed, formation of the oligomers is contingent on the presence of detergents. These would probably interfere with the assay by breaking the liposomes. Indeed, addition of detergent (*e.g.* Triton X-100) is often the method of choice for fully releasing the dye and obtain the reading for maximum permeabilisation, which is used for data normalisation. Therefore, an alternative approach was required.

An elegant biochemical study used engineered disulfide bonds in BAK to probe its conformational re-arrangements upon pore formation (Iyer *et al.*, 2016). In this work, the authors introduced pairs of cysteine residues that were in close proximity in the monomeric state. These mutants were stably expressed in $Bak^{-/-}$ $Bax^{-/-}$ double knock-outs mouse embryonic fibroblasts (MEFs). Apoptosis was then triggered by addition of truncated BID (tBID), and the release of cytochrome c from mitochondria monitored by western blotting—indicating pore formation. By showing that the formation of certain disulfide bonds prevented release of cytochrome c, the authors were able to demonstrate that parts of the protein had to undergo conformational rearrangements upon pore formation. These were prevented by disulfde stapling, thus preventing oligomerisation and MOMP.

The same disulfide stapling strategy was used as a proxy for testing the physiological relevance of the oligomers generated using detergents. Three of the disulfide mutants reported to suppress MOMP were selected (A28C/L163C, Y41C/A79C, V142C/F150C). In addition, two other mutants were designed to study the effect of linking other parts parts of the protein (Q77C/Q184C, T116C/H165C). Cysteine double mutants of BAK were generated by site-directed mutagenesis, expressed, and purified in the same way as the wild-type construct. Monomers were buffer-exchanged into 50 mM sodium phosphate pH 7.0 containing no reducing agent. Intra-molecular disulfide bonds were generated by incubating the proteins with the redox catalyst Cu^{II}(1,10-phenanthroline)$_3$ (CuPhe, Kobashi (1968)). After quenching with EDTA, the samples were dialysed. This step was necessary to remove the Cu:EDTA complex, which interfered with downstream spectroscopic detection due to its strong absorbance at 280 nm (ε_{280} = 42211 M^{-1} cm^{-1}). These BAK disulfide mutants were then incubated in the presence of PS20, and the extent of oligomerisation assessed by SEC. Addition of DTT prior to adding detergent served as controls.

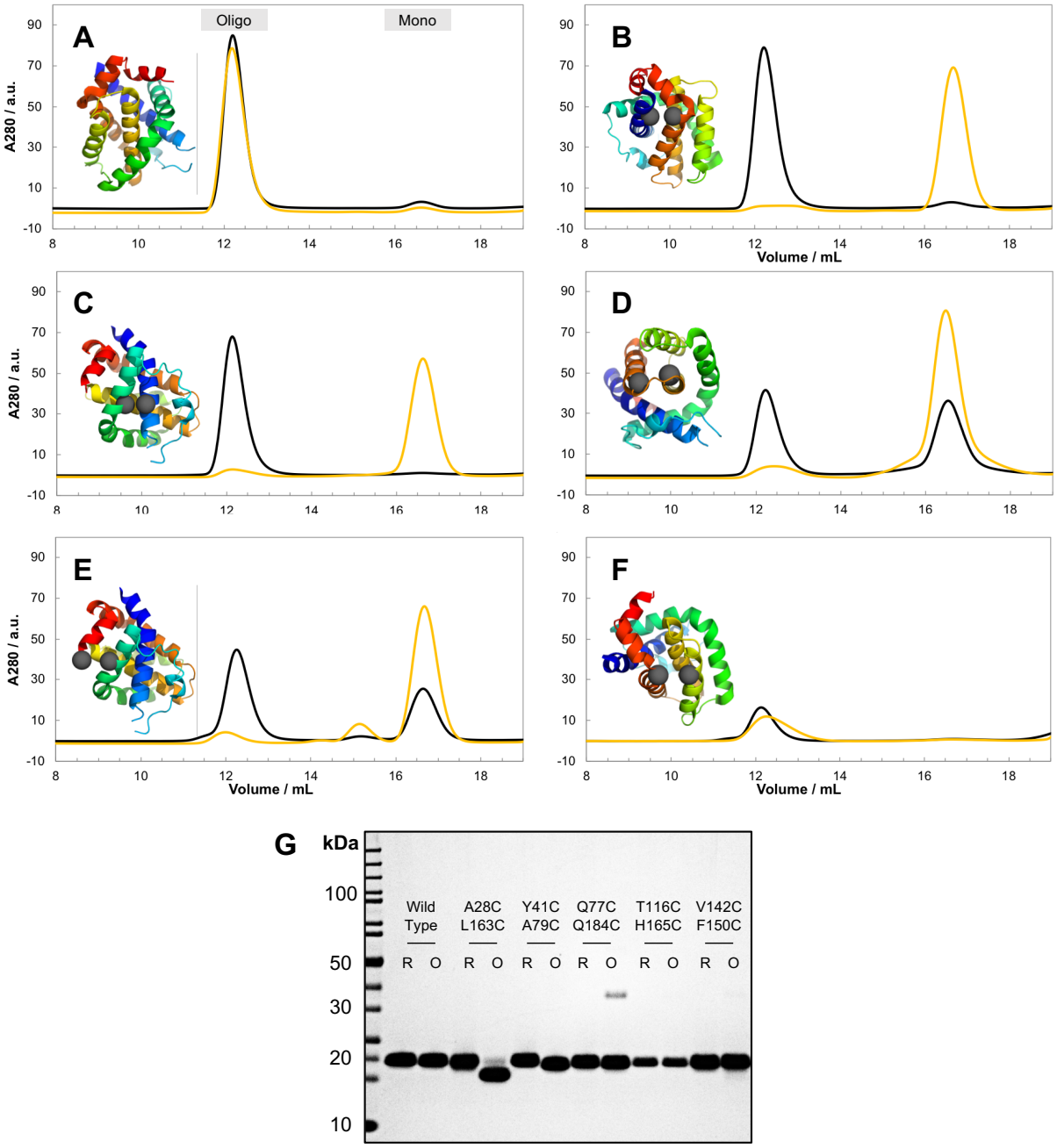

Figure 7.8 – Disulfide mutants of BAK incapable of MOMP were used as proxies for determining the physiolgical relevance of oligomerisation in detergent. For each double mutant, 20 μM of protein was treated with PS20 (20×CMC), and the oligomerisation assessed by SEC (Superdex 200 10/300). The disulfide form (yellow) was compared to the dithiol form (black) to devonvolute the effect of mutations from stapling. Structures of BAK (PDB:2YV6) with cysteine mutations shown as grey spheres are indicated for each double mutant. (**A**) BAK wild-type (no cysteines). (**B**) A28C/L163C. (**C**) Y41C/A79C. (**D**) V142C/F150C. (**E**) Q77C/Q184C. (**F**) T116C/H165C at 6 μM. **E** and **F** have not been assessed in mitochondrial assays. (**G**) SDS-PAGE analysis of oxidised (O) and reduced (R) forms of the double mutants.

All three cysteine double mutants that were incapable of MOMP also failed to oligomerise in the presence of PS20. This effect was specific to the presence of disulfide bonds, as the

dithiol forms of the proteins were still capable of oligomerising under the same conditions (Fig. 7.8A–D). Thus, the cysteine mutations themselves were not responsible for the lack of oligomerisation; the formation of the disulfide bond was. V142C/F150C was an exception, showing incomplete oligomerisation in its dithiol form, therefore indicating that mutations in this part of the protein destablises the oligomer. Interestingly, Q77C/Q184C and T116C/H165—the two designed constructs—behaved very differently. The former showed suppressed oligomerisation, although as for V142C/F150C, this could be attributed in part to the mutations themselves. In contrast, the latter appeared to have no effect; both the disulfide and dithiol forms underwent oligomerisation in the presence of detergent, indicating that these parts of the protein remain in proximity in the oligomer. Further results on the use of disulfide stapling for topological mapping can be found in Section 7.4.2.

It is noted that since the DTT-reduced controls were first treated with CuPhe as well, oxidative damage, or effects different from the presence of the disulfide bond can be excluded. This is further supported by the fact that BAK wild-type—which does not contain cysteines—behaved the same regardless of the treatment to which it was subjected. Complete intramolecular disulfide bond formation was confirmed for BAK A28C/L163C. SDS-PAGE analysis revealed a downward shift of the monomer band following oxidation, consistent with the formation of an intramocular bond (Fig. 7.8G). By extension, disulfide bond formation was assumed to occur for all other mutants, despite the lack of anomalous gel migration. SDS-PAGE also revealed the absence of intermolecular bond formation, thus excluding covalent multimerisation as the reason behind the observed behaviour. Q77C/Q184C showed a small amount of dimer, potentially due to slight helix fraying at the C-terminus.

Together, these results prove that the suppression of oligomerisation in PS20 can be attributed exclusively to the formation of an intramolecular disulfide bond. By extrapolation to the biochemical data (Iyer *et al.*, 2016), it can therefore be assumed that the oligomeric state(s) obtained using detergents is equivalent to the structural state responsible for pore formation under physiological conditions.

Folding studies have shown that stabilisation of proteins from intramolecular disulfide bonds occurs primarily through destabilisation of the denatured state, not through stabilisation of the native fold (Pace *et al.*, 1988). While this work does not involve unfolding of the protein, nor does it make use of denaturant, it is interesting to draw parallels. Here

the oligomer is the state that is too high in energy to be significantly populated in buffer (the 'unfolded state'), and the thermodynamic perturbation comes from the presence of detergent (the 'denaturant'). If the detergent acts by shifting the distribution of states, the implication is that the formation of an intramolecular disulfide bond destabilises the oligomer, thus favouring the monomer.

7.4 The nature of the oligomer

The structural nature of BAK and BAX oligomers has been extensively investigated. However, despite decades of work, their topologies and structures remain elusive. Since pore formation relates to cell death, the oligomers can not be accumulated, thus precluding isolation from biological samples. Therefore, not only is re-constitution necessary, but the membrane-embedded nature of the oligomer represents a challenge for most structural techniques. Insights have been gained from biochemical studies using disulfide stapling to map interactions in the MOM (Dewson *et al.*, 2008, 2012) , cryo-EM on nanodisc-embedded BAX (Xu *et al.*, 2013), AFM studies in supported lipid bilayers (Salvador-Gallego *et al.*, 2016), modelling using DEER constraints (Bleicken *et al.*, 2014), and X-ray crystallography of dimers (Czabotar *et al.*, 2013, Brouwer *et al.*, 2014), yet the structure of the oligomer remains debated.

The work presented here demonstrated that BAK and BAX can be oligomerised to physiologically relevant states in the presence of the detergent PS20. This discovery opens new avenues for studying structural aspects of these assemblies, which are the subject of the following sections.

7.4.1 Oligomers are constructed of dimer units

Recent advances in native mass spectrometry (MS) have made this technique particularly suitable for investigating membrane proteins (Laganowsky *et al.*, 2013). Using a delicate balance of ionisation and collision energies, proteins can be de-solvated and transferred into the gas-phase without unfolding them, thus allowing to record charge states of folded ensembles. The same principle can be applied to oligomers, provided that the quaternary structure survives the ionisation (Benesch & Robinson, 2006). By extension, this technique can be employed to study membrane proteins solubilised in detergents. The added difficulty

resides in the necessity to break the protein:detergent complex, which needs to be performed mildly to retain the membrane protein intact. This can be particularly challenging for membrane-embedded quaternary assemblies.

The ability to oligomerise BAK and BAX in non-ionic detergents opened the possibility to study these complexes by native MS. This technique was employed to confirm oligomerisation, and gain insights into the stoichiometries of these assemblies. Further experiments using native MS to study the regulatory mechanism of BCL-2 proteins are reported in the next chapter.

Monomeric BAK and BAX were purified in sodium phosphate buffer, and transferred to the laboratory of Prof. Carol V. Robinson (University of Oxford). The proteins were buffer-exchanged into ammonium acetate buffer matching the pH and ionic strength of the phosphate buffer on-site. Samples were prepared at 5 μM concentration, and oligomerisation initiated by addition of detergent. Both PS20 and C12E8 were tested. PS20 gave better spectra, and further efforts were focused on this detergent. Mass spectra were recorded on a Q Exactive Quadrupole-Orbitrap mass spectrometer (Thermofisher Scientific). All data were collected with Dr Kallol Gupta, who also performed the final analysis.

Oligomers of both proteins were detected in PS20. These results were contingent on the presence of detergent—spectra of proteins in buffer did not reveal higher-order species—confirming that a hydrophobic environment is a pre-requisite to oligomerisation. Up to hexamers could be recorded for BAK, while under the same conditions, pentamers of BAX could be observed (Fig. 7.9). These results are consistent with cross-linking experiments, which captured higher-order species for BAK than BAX. Qualitatively, both proteins showed similar spectra. However, despite equivalent SEC profiles, the distributions of BAK and BAX oligomers in native MS were different. The distribution of BAK oligomers was shifted towards larger species when compared to BAX. The spectra were acquired under identical conditions, making conclusions about the relative energies of each assembly possible. Thus, it can be inferred that BAK oligomers are more stable than BAX in the gas-phase. For example, BAK dimers were the smallest units observed under these conditions, while monomers were present for BAX.

The distribution of species was dependent on the collision energy employed. Increasing it led to a gradual disappearance of higher-order oligomers, and an increase in monomer/dimer population. Further increases started to unfold the proteins, noticeable by the formation of narrow distributions at small m/z values—corresponding to the high charge states typical

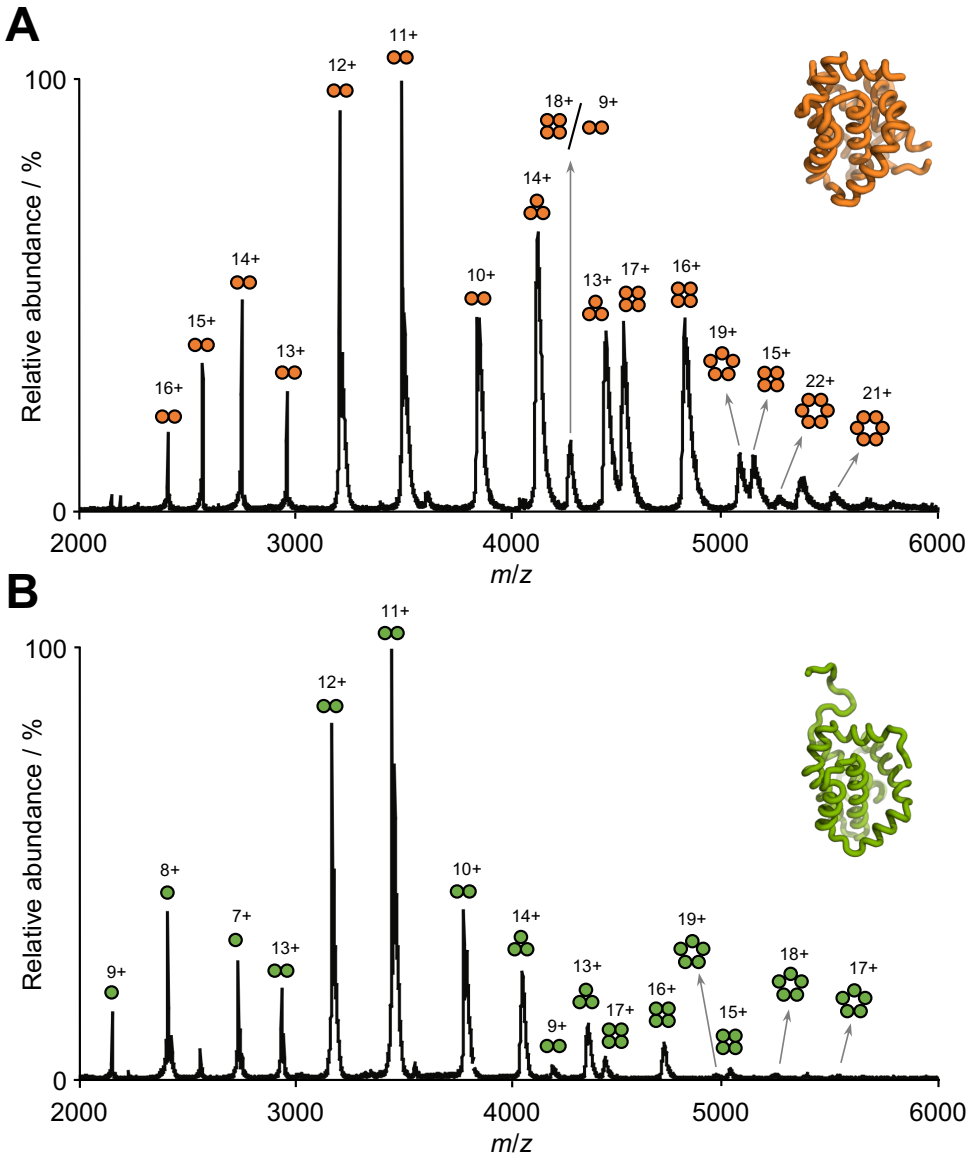

Figure 7.9 – Native MS of BAK and BAX oligomers in PS20 (5×CMC). (**A**) Ensemble of BAK oligomers. Up to hexamers could be detected. (**B**) Ensemble of BAX oligomers. Up to pentamers could be detected.

of unfolded chains. Therefore, while these data are informative in confirming the presence or absence of given stoichiometries, their relative abundance in the gas-phase does not reflect the solution-state. However, the absence of highly asymmetric charge states suggests that gas-phase fragmentation was not the source of the smaller oligomers. In other words, all oligomers already exist in solution, but their relative amounts cannot be deduced from these native MS experiments.

This heterogeneity appears surprising when considering the results obtained from SEC. While most detergent showed a distribution of species, PS20 led to an apparently homogenous peak for both BAK and BAX. It was rationalised that the co-elution of free micelles

with the oligomers in the case of PS20 resulted in greater stability. However, could different protein:detergent ratio for the various oligomers lead to an ensemble of species with similar elution volumes, thus masking an underlying heterogeneity? The low μM concentrations used in detergent screening did not offer sufficient resolution, so oligomerisation at higher concentration was performed, and the result analysed by SEC (Fig. 7.10).

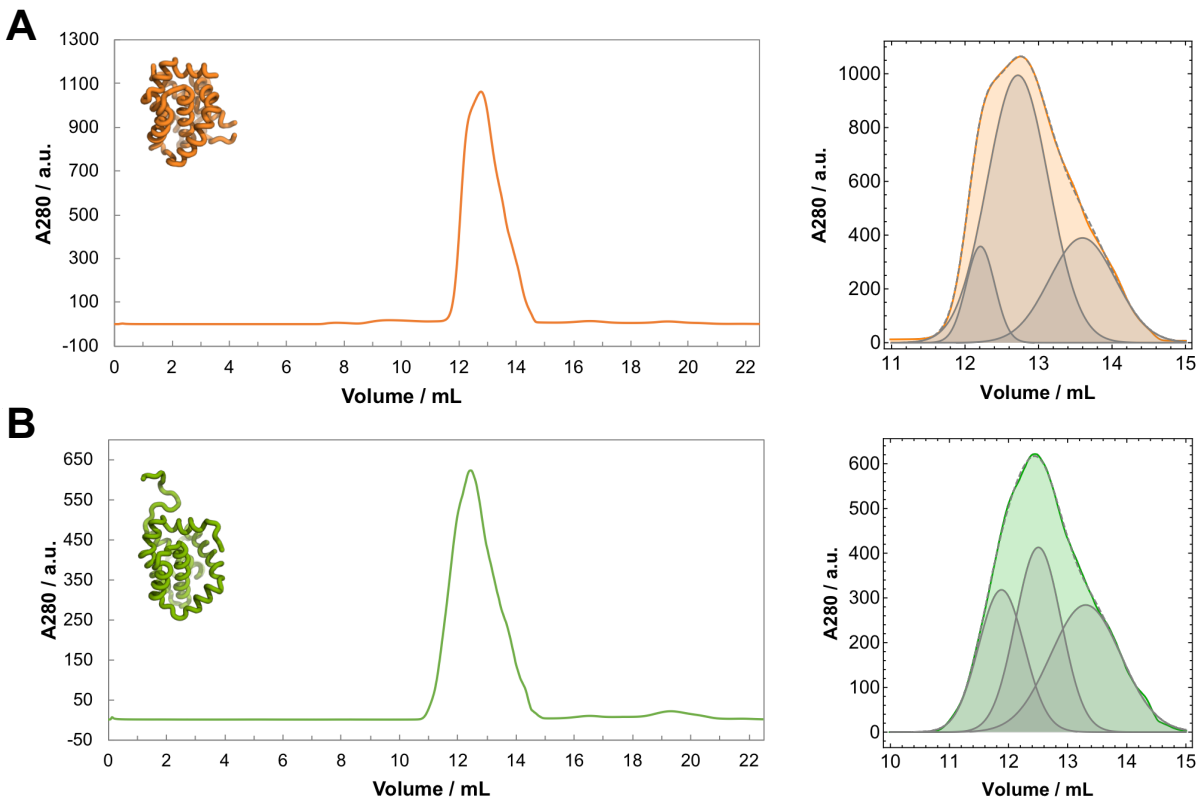

Figure 7.10 – BAK and BAX oligomers are heterogneous at high concentration. (**A**) SEC of BAK at ∼450 μM in the presence of ∼90 mM PS20 (1100×CMC). (**B**) SEC of BAX at ∼250 μM in the presence of ∼90 mM PS20 (1100×CMC). Right panels show the oligomer peak fitted to a triple Gaussian function (dashed grey line). The individual Gaussians are reproduced for visualisation. SEC was performed on a Superdex 200 10/300 equilibrated in 50 mM sodium phosphate pH 7.0.

These results highlight that even in PS20, the oligomeric distributions of BAK and BAX are also heterogeneous when carefully analysed by SEC. Although the distribution of masses observed by native MS would suggest a wider ranger of elution volumes, the distribution recorded by SEC is relatively narrow. While native-MS looks at 'naked' oligomers stripped of their detergents, these assemblies are part of micelles in solution. These protein:detergent complexes will impact the hydrodynamic radii of the particles, and thus affect their elution volumes. Therefore, it is possible that the narrow distribution of elution volumes observed by SEC could be the consequence of different protein:detergent ratios.

Interestingly, even-numbered oligomers seemed slightly more abundant. Moreover, upon ramping the collision energy, higher-order oligomers disappeared in favour of dimers. These appeared to be more stable than the larger species. These results might suggest different oligomerisation interfaces—monomers assemble into sturdy dimer units, which further assemble through weaker dimer:dimer interactions to form larger oligomers. Therefore, it is tempting to speculate that the oligomers possess at least two sets of energetically distinct interfaces. Addition of monomers to these units also appears possible, resulting in odd-numbered assemblies.

In summary, these native MS investigations of PS20-treated BAK and BAX confirmed the oligomerisation of these proteins into larger assemblies. It also revealed heterogeneous ensembles. These results are consistent with a recent study in supported lipid bilayers (Subburaj *et al.*, 2015), which also showed a distribution of oligomeric states by single-particle fluorescence microscopy; further validating the physiological relevance of PS20-treated oligomers.

7.4.2 Investigations of the dimer topology

Results from native MS seemed to reveal a hierarchy to the oligomerisation of BAK and BAX; sturdy dimers and weaker larger assemblies. It is interesting to consider the existence of this dimer 'denominator' in light of the structural studies undertaken to date. Three different types have been crystallised for both BAK and BAX, thus indicating some form of energetic promiscuity to the assembly of these proteins. Could they be related to the dimers constituting the oligomers formed in PS20? Using biophysical data, and an extension of the disulfide stapling presented in Section 7.3.4, the oligomers presented here were assessed for compatibility with these structures.

Despite a sequence identity of 21.1%, the structures reported for BAK and BAX show great similarities. The dimers can be classified according to their topologies as: *i*) side-on, *ii*) BH3-in-groove, and *iii*) helix-swapped (Fig. 7.11). Note that the side-on dimers are topologically similar (absence of swapping), but unlike BH3-in-groove and helix-swapped dimers, have different interfaces. Each structure is discussed in turn, and compared to the data obtained for BAK in PS20. While no data was collected for BAX, its identical biological function, as well as the structural homology of both its monomer and dimers, suggests that conclusions are probably transferable.

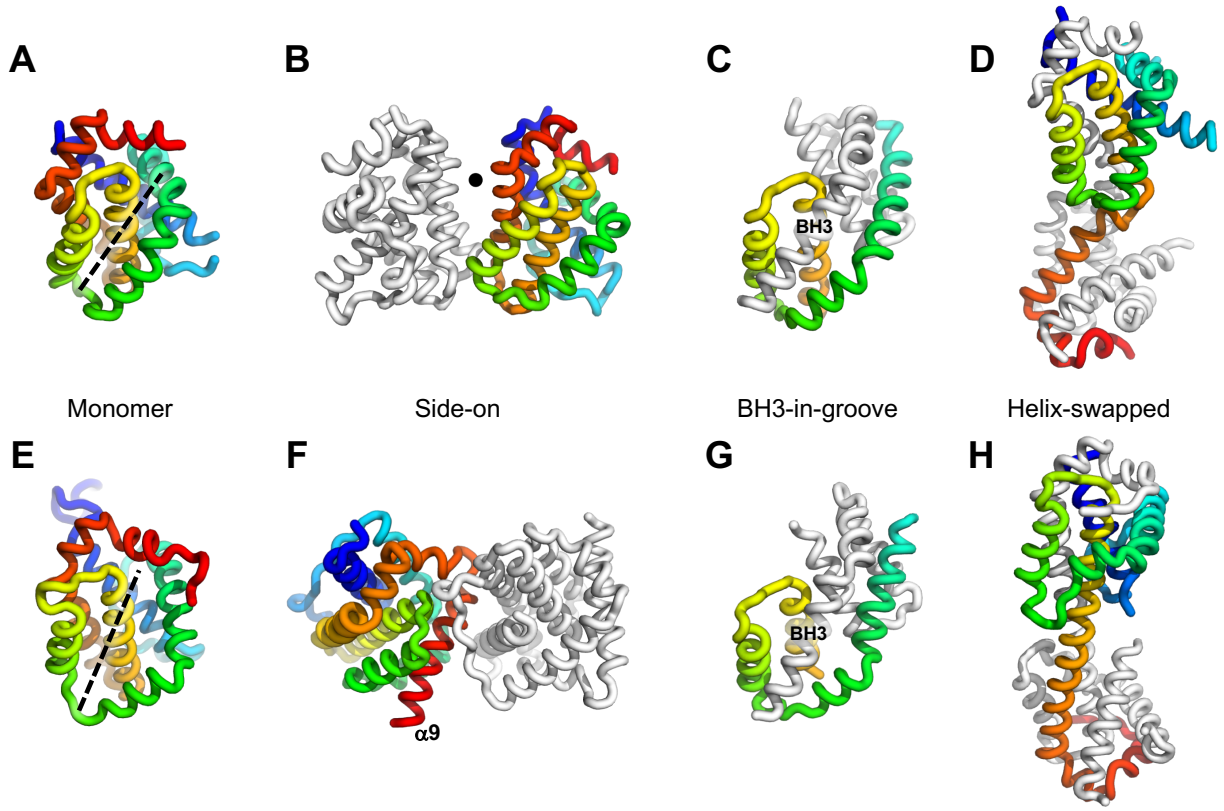

Figure 7.11 – Structures of BAK (top) and BAX (bottom) reported in the literature. Each monomer is coloured from N (blue) to C (red) terminus. For dimers, only one protomer is coloured, and the other one is shown in grey. BH3-in-groove structures only contain helices α2–α5, which is highlighted by matching the colour gradient to aid comparison. The canonical BH3 groove is indicated by a dashed line on the monomer structures. (**A**) PDB:2YV6; (**B**) PDB:2IMT, the zinc atom is depicted as a black sphere; (**C**) PDB:4U2V; (**D**) PDB:4U2U; (**E**) PDB:1F16, the 9$^{\text{th}}$ transmembrane helix is not shown; (**F**) PDB: 4S0O, the 9$^{\text{th}}$ transmembrane helix is shown; (**G**) PDB:4BDU; (**H**) PDB:4BD7.

The side-on dimer of BAK (Fig. 7.11B) was obtained in the presence of μM concentrations of zinc through a coordination site at its interface (Moldoveanu *et al.*, 2006). The author also demonstrated dimerisation using NMR spectroscopy, thus excluding a crystallisation artefact. However, it was only observed in the presence of zinc, and at high concentrations of protein. In contrast, BAK readily oligomerises in the low-μM range—and in the absence of zinc—when PS20 is present, suggesting a different dimer. A similar argument also applies for the side-on dimer of BAX (Fig. 7.11F) (Garner *et al.*, 2016). This structure has an apparent K_{d} of ∼250 μM, again precluding its formation under the low-μM concentrations used here. Most importantly, both side-on dimer show no topological complexity, *i.e.* no structural rearrangements. This is inconsistent with disulfide stapling results (*cf.* Section 7.3.4). Therefore it was concluded that these side-on dimers are not the same as the ones

formed in the presence of PS20.

The helix-swapped dimer of BAX (Fig. 7.11H) was obtained by incubating the protein with *n*-octyl β-D-maltoside—a non ionic-detergent (Czabotar *et al.*, 2013). The same topology was obtained for BAK by addition of CHAPS (a zwitterionic detergent) and BID_{BH3}, although the latter did not co-crystallise (Fig. 7.11D) (Brouwer *et al.*, 2014). In both structures, helices α6–α8 from each protomer swap over, resulting in two globules with folds equivalent to the monomers. These 'domains' are composed of helices α1–α5 from one chain and α6–α8 from the other. Except for the presence/absence of a kink between helices α5 and α6—the interface between the two globules—both helix-swapped dimers of BAK and BAX are similar. It is noted that the authors argue against their physiological relevance, mainly on the basis of inconsistency with biochemical data regarding the oligomer interfaces formed during apoptosis.

PS20 generated oligomers bigger than dimers. However, since these structures were also obtained in the present of detergents, could the dimers observed by native MS be composed of the same helix-swapped unit? This question was addressed by disulfide cross-linking using the cysteine double mutants presented in Section 7.3.4. It was shown that for all but T116C/H165C, generation of the disulfide link *before* addition of detergent prevented oligomer formation. Hence, these parts of the protein need to separate—at least transiently—in order for the oligomer to form.

In the helix-swapped dimer, each globule has the same topology as the monomer. Therefore, any contacting pair existing in the monomer will also exist in the dimer, although in the context of the helix-swapped structure, some contacts will be formed *between* chains. Thus, if disulfide formation was used to capture these contacts, both *inter*- and *intra*-molecular bonds may be formed; the outcome depending on the position of the pair. In other words, formation of the disulfide bond *after* the addition of detergent—*i.e.* after oligomerisation took place—should result in covalent dimer formation only for some cysteine double mutants. Therefore, by looking at the presence or absence of monomers and dimers for the different disulfide constructs, the consistency of the helix-swapped dimer with respect to the structure formed in PS20 may be assessed. Covalent dimerisation from disulfide formation was evaluated by non-reducing denaturing SDS-PAGE. These results can be found in Fig. 7.12, alongside the positions of the cysteine mutations mapped onto the helix-swapped structure of BAK.

Analysis of dimer formation with respect to the positions of the mutated residues reveals

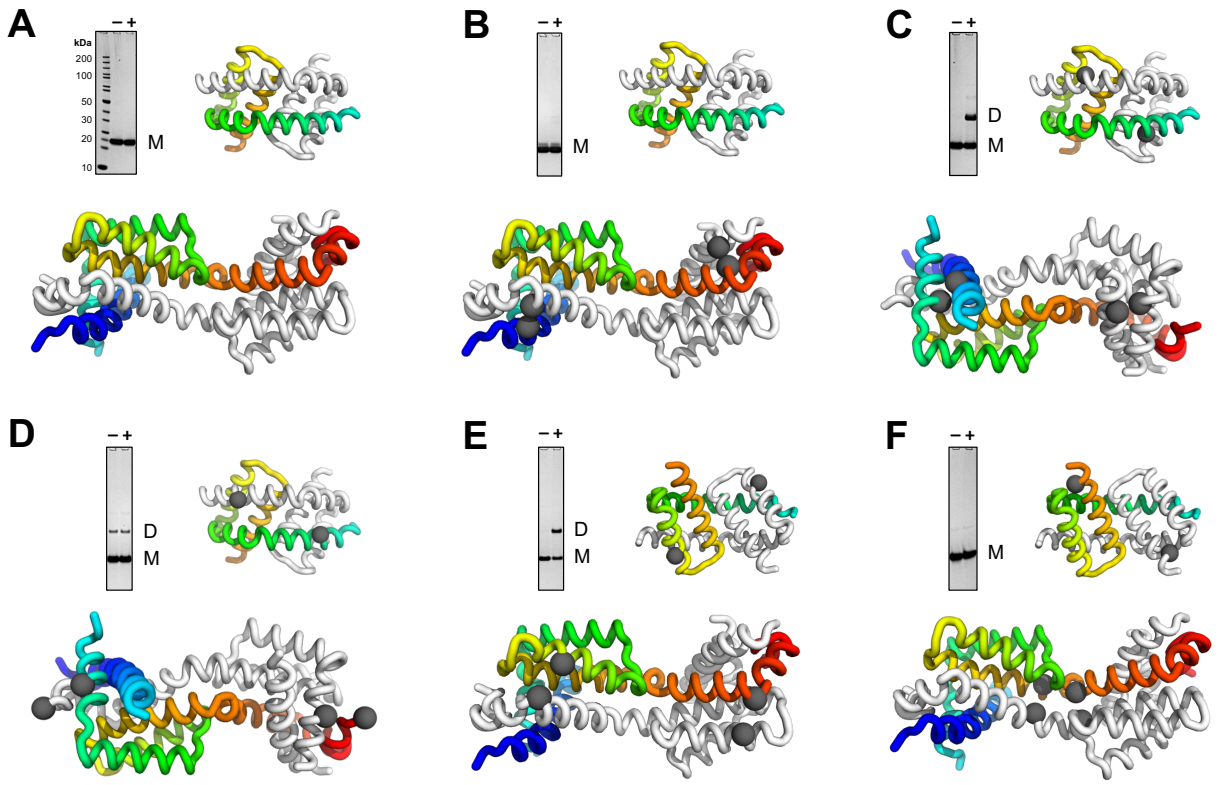

Figure 7.12 – Topolgical mapping of dimer structures with disulfide stapling. Positions of the cysteine mutations are indicated as grey spheres on the structures of BAK helix-swapped dimer (PDB:4U2U, bottom) and BH3-in-groove dimer (PDB:4U2V, top right). Inter-molecular disulfide bond formation was assessed by SDS-PAGE (top left) in the presence (+) or absence (–) of PS20 (20×CMC). (**A**) BAK wild-type (no cysteines); (**B**) A28C/L163C; (**C**) Y41C/A79C; (**D**) Q77C/Q184C; (**E**) T116C/H165C; (**F**) V142C/F150C.

that four out of the five constructs show data inconsistent with the helix-swapped structure. Indeed, A28C/L163C (Fig. 7.12B) should produce dimers, but only monomers are observed. Y41C/A79C (7.12C) should solely yield monomers, but dimer formation is evident. Based on the structure, Q77C/Q184C (7.12D) is expected to dimerise. Some appear to form, but the relative amount is identical to the control (absence of detergent). Slight oligomerisation in buffer had already been observed for this mutant (Fig. 7.8G), and was ascribed to helix fraying given the very terminal position of Q184C. The inter- and intra-cysteine distances in V142C/F150C (Fig. 7.12F) are very similar, thus inter-molecular cross-linking might compete with intra-molecular reactivity. However, some dimer should still form, which is clearly not the case. Only T116C/H165C (Fig. 7.12E) appears consistent. Both positive results (formation of dimers when only monomers would be expected) and negative results (absence of dimers when formation would be expected) were obtained. Moreover, the positions of these cysteine mutations represent a comprehensive coverage of

the structure (helices 1, 2, 4, 5, 6, 8). Altogether, these data provide compelling evidence that the helix-swapped topology is not the dimer at the heart of the BAK oligomers formed in PS20.

The BH3-in-groove hypothesis for the nucleation of BAK and BAX oligomers appears to be the prevailing model in the literature. Interestingly, its inception precedes any form of structural data. Its popularity seems to stem from the fact that it fits the general mantra that BH3 motifs are the molecular handles for the interactions of BCL-2 family members. The model has been supported by screening for loss-of-function mutations in cellular assays, revealing a concentration in the BH1 and BH3 motifs (Dewson *et al.*, 2008). While these data are consistent with a BH3-in-groove dimer, their lack of molecular resolution implies that alternative structures cannot be fully excluded—only the importance of these residue for function is certain, not their structural roles.

Recently, structures of these putative BH3-in-groove dimer units have been reported for both BAK and BAX (Czabotar *et al.*, 2013, Brouwer *et al.*, 2014). However, severe modifications of the proteins were required to obtain them. Out of eight helices (nine if considering the trans-membrane one), four were removed from the crystallised constructs. Moreover, fusions to green fluorescent proteins (GFP) were required to increase solubility and aid crystallisation. The results do show a BH3-in-groove topology (Fig. 7.11C, G), however the artificiality of these constructs questions the relevance of these structures. The positioning of the remainder of the protein with respect to this dimerisation unit is unclear. The authors claim their relevancies based on consistency with the aforementioned loss-of-function study (Dewson *et al.*, 2008), EPR measurements (Bleicken *et al.*, 2010), and photocross-linking (Zhang *et al.*, 2010). The lack of spatial resolution of these studies implies a degenerate set of structures consistent with these results. So while they do not disprove the BH3-in-groove structures, they do not prove its existence either.

Unfortunately, all the cysteine double mutants had at least one position outside of the BH3-in-groove structure (Fig. 7.12). Thus, the structure could neither be confirmed, nor disproved by simply using this dataset. However, comparing the disulfide dimer formation with the mapped positions does hold some information. Absence of dimer formation is a trivial result, but dimer formation is not; especially if a pair of residues can be excluded. This is the case for the double mutants Y41C/A79C and T116C/H165C. Indeed, each have one position that is too far in space in the BH3-in-groove structure to interact. Since disulfide bonding does occur, the implication is that the other position forms a spatial

pair in the dimer (Y41C and H165C respectively). Of course this result would only be true if the BH3-in-groove structure is indeed the dimer unit formed in PS20. Based on the BH3-in-groove structure, a model of BAX oligomers in liposomes was generated using double electron-electron resonnance (DEER) spectroscopy constraints (Bleicken *et al.*, 2014). This model allows the missing cysteine positions to be mapped. The result however, was inconsistent with the data presented here. Whether this discrepancy stems from the model, or the dimers formed in PS20 are different, is unclear.

The hypothesis that the BH3 motif is the dimerisation key was further analysed in the context of biophysical data. In Chapter 6, the affinity of BAK for its own BH3 motif was demonstrated to be vanishingly small (K_d = 1 mM). Accordingly, it seems unlikely that such an interaction would provide the driving force behind dimer nucleation. Although the BH3-in-groove dimer involves more interaction than just its BH3, the energetic of the other interactions would need to compensate for this weak affinity. However, since the K_d was obtained in buffer—where BAK is monomeric—some conformational differences in the dimer could shift the affinity. In order to test this hypothesis, BAK was incubated in the presence of a large excess of t-$\mathrm{BAK_{BH3}}$, and oligomerisation induced by addition of PS20. It was rationalised that if the oligomer had higher affinity for its motif than the monomer, binding might be detected under the condition of the experiment. More importantly, since the BH3 peptide would compete for the binding groove, inhibition of the oligomerisation would be expected to occur. Thus, reduction of the amount of oligomer formed would be supporting evidence for the BH3-in-groove model of dimerisation. The outcome was analysed by SEC, and the partitioning of the t-$\mathrm{BAK_{BH3}}$ peptide analysed by absorbance at 555 nm to detect the presence of TAMRA dye in SEC fractions.

No binding of the peptide to the oligomer was observed with a 30-fold excess of t-$\mathrm{BAK_{BH3}}$. This was confirmed by a complete lack of absorbance at 555 nm—the absorbance maximum of TAMRA—for the peak at ∼12.2mL. At the concentrations used for this experiment, ∼23% complex would be expected if the $\mathrm{BAK_{BH3}}$ peptide had the same affinity for the oligomer as the monomer. Therefore, absence of binding either suggests an occluded binding groove—which would be case in the BH3-in-groove scenario—or some other conformational rearrangement that would disrupt the binding interface. Integration of the oligomer peaks reveals a 20% reduction in surface area when the oligomerisation is conducted in the presence of the peptide. This value is close to the 23% mentioned earlier, thus suggesting that the reduction in oligomerisation is a consequence of the interaction

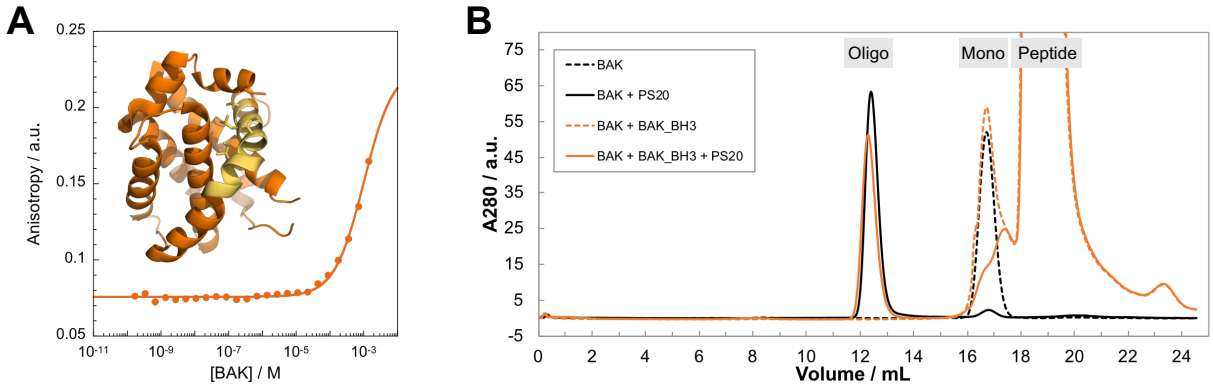

Figure 7.13 – BAK_{BH3} does not significantly prevent BAK oligomerisation. (**A**) Binding isotherm of the interaction between BAK and t-BAK_{BH3} shows a very weak affinity (K_d = 1 mM). The structure of BAK (PDB:2YV6) with its BH3 motif highlighted is shown for reference. (**B**) SEC analysis of the oligomerisation of BAK (10 µM) in the presence (solid orange line) or absence (solid black line) of 300 µM t-BAK_{BH3}. The concentration of PS20 was 20×CMC. The respective experiments in the absence of detergent are shown as dashed lines.

between BAK monomer and BAK_{BH3}. In other words, binding of the motif to the canonical groove of BAK inhibits its oligomerisation (*cf.* Section 7.5.2 for more details on this phenomenon). Together with the absence of binding to the oligomer—arguing for competing interfaces—these results are consistent the BH3-in-groove dimer model, although more experiments are warranted to confirm this scenario in the context of PS20.

Dimers of other BCL-2 proteins

Interestingly, two additional types of dimers have been described for BCL-2 proteins: *i*) α1 swapped, and *ii*) α5–6 swapped. These structures have not been reported for BAK or BAX, only for anti-apoptotic proteins. However, given their structural plasticity, could it be that these dimers do form, but have not been structurally characterised yet? In order to test this hypothesis, homology models of BAK for both topologies were generated using the SWISS-MODEL server (https://swissmodel.expasy.org) (Fig. 7.14). PDB:3INQ (BCL-X_L, Lee *et al.* (2009)) was used as template for the α1 swapped dimer, and PDB:2Y6W (BCL-W, Lee *et al.* (2011)) was used for the α5–6 swapped dimer. These models were assessed for their consistency with the disulfide cross-linking data (Fig. 7.12). Analysis of pair distances and expected monomer/dimer outcomes revealed that neither model could fully recapitulate the dataset. The α1 swapped dimer was inconsistent with the results from both A28C/L163C and T116C/H165C, while the α5–6 swapped dimer contradicted the Y41C/A79C data. Therefore, it can be concluded that these topologies do not represent

the structures formed when BAK oligomerises in PS20.

The combination of biophysics and disulfide cross-linking approaches presented in this section have allowed the oligomerisation of BAK in the context of known structures of its dimer to be analysed. It appeared that side-on and helix-swapped topologies were not consistent with the oligomers formed in PS20. Neither was it consistent with homology models based on other dimer topologies reported solely for anti-apoptotic BCL-2 proteins. Albeit not entirely conclusive, oligomerisation of BAK in PS20 appears to be compatible with the BH3-in-groove topology. This is the prevailing model in the literature, which is supported by cross-linking studies in apoptotic cells (Dewson *et al.*, 2008). Therefore, in addition of the data presented in Section 7.3.4, this further supports the physiological relevance of BAK oligomers in PS20. While no data was collected for BAX, structural and functional homology suggest that the conclusions reached for BAK are probably transferable.

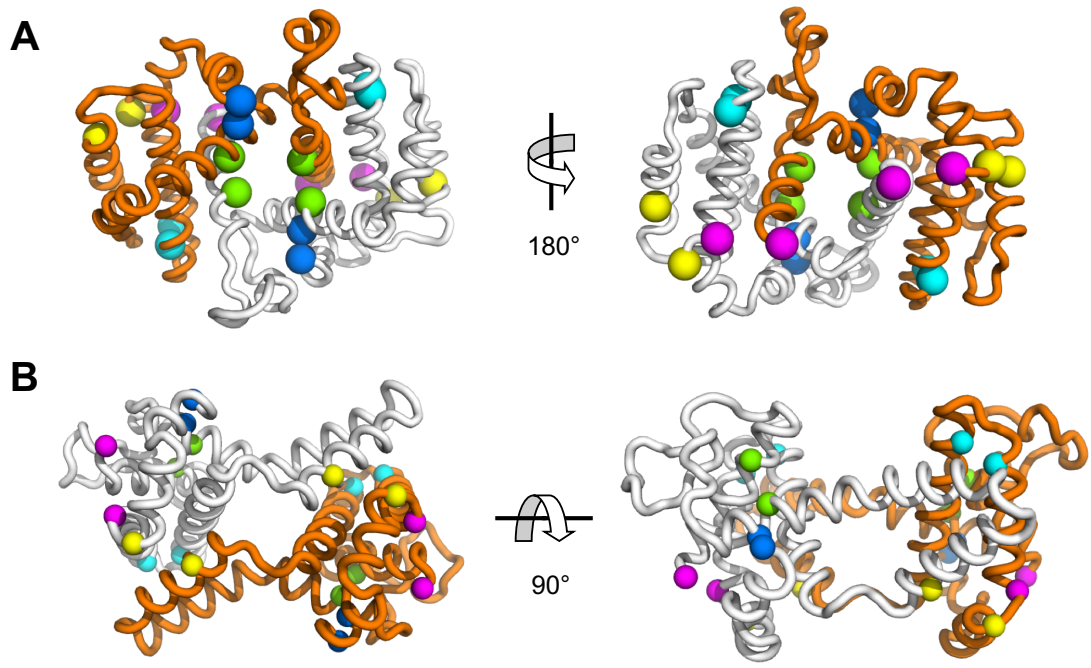

Figure 7.14 – Homology models of BAK based on dimer structures only reported for anti-apoptotic BCL-2 proteins. One protomer is coloured orange, and the other grey. (**A**) α1 swapped dimer based on template PDB:3INQ. (**B**) α5–6 swapped dimer based on template PDB:2Y6W. The cysteine double mutants are colour-coded: A28C/L163C (magenta); Y41C/A79C (green); Q77C/Q184C (blue); T116C/H165C (yellow); V142C/F150C (cyan).

7.4.3 Visualising the oligomer

The oligomerisation of BAK and BAX was demonstrated in detergents, and its physological relevance assessed. Native MS suggested a hierarchical organisation, with dimer units assembling into larger oligomers. The topology of this building block was investigated using disulfide stapling in conjunction with data from the literature. These experiments have provided valuable insights into these oligomers. However, an atomistic model would be desirable. Crystallisations of PS20-treated BAK and BAX were attempted, however these were not successful. Detergents being known to be problematic for crystallisation, it was the most likely cause of these failed attempts and this route was not pursued any further. Due to their large sizes, they are unfortunately unsuitable for solution-NMR characterisation, as their slow tumbling would likely lead to severe line broadening.

Cryo-EM is an attractive structural technique for solving complexes that have resisted crystallisation. However, even with recent advances in the field, the small size of the oligomer (∼120 kDa for the hexamer) still represents a challenge. As a first step towards potentially studying these assemblies using cryo-EM, negative-stain EM was attempted. Using class-averaging—a technique that enhances signal-to-noise by averaging many particles with identical orientations—it is in theory possible to obtain a relatively accurate description of the topology of an oligomer (Ohi *et al.*, 2004).

BAK and BAX monomers were prepared in 50 mM sodium phosphate pH 7.0 buffer. The proteins were shipped to our collaborators, together with detergents to induce oligomerisation on-site. Grid preparation, data collection and analysis were performed by Philip Rowell at the University of Leeds on a Tecnai T12.

Class averages of PS20-treated BAK revealed that some particles adopt pore-like shapes. These results are in line with the expected topology for these proteins given that MOMP is a result of pore-formation by BAK and BAX. This provides additional support to the notion that oligomerisation in detergent does relate to the apoptotic pore. However, these class-averaged results also reveal structural heterogeneity. While this does not contradict the results from native MS and chemical cross-linking, it appears at odds with the outcome from SEC. Heterogeneity was observed with some detergents, but PS20 showed an apparently homogeneous distribution. One possible explanation for these EM results lies in grid preparation. Very low concentrations are required to ensure that the particles are sufficiently dispersed. So after performing the oligomerisation reaction at µM concentra-

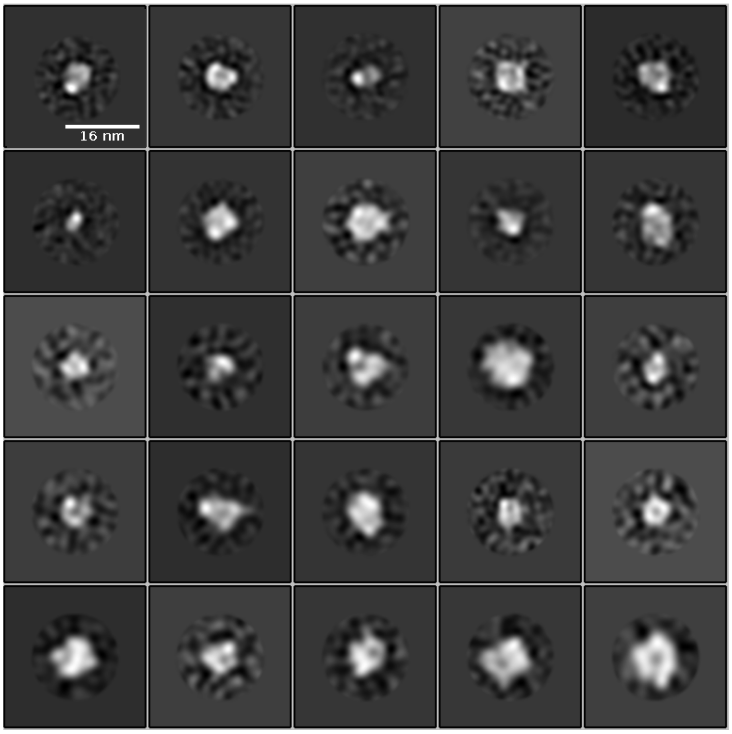

Figure 7.15 – Class-averages from negative-stain EM of BAK treated with PS20. Data collected and analysed by Philip Rowell (University of Leeds).

tions, the sample was diluted into buffer containing no detergent before being applied to the grid. This step also diluted PS20 below its CMC. It is therefore possible that the oligomer became unstable, and fell apart. Protein dilution into buffer containing detergent was attempted to circumvent this problem. Unfortunately, these samples did not adhere to the grid and could not be imaged. Optimisation will be required to find conditions where the oligomers remain stable, while also making grid preparation possible. This is currently work in progress.

7.5 Mechanism of oligomerisation

Oligomersation of BAK and BAX being responsible for MOMP, an understanding of its mechanism would provide valuable insights into the control of apoptosis. Having demonstrated the oligomerisation in detergent, its physiological relevance, and some of its structural aspects, biophysical methods were employed to study the mechanism of the process.

7.5.1 Oligomerisation is not spontaneous in buffer

The direct activation mechanism stipulates that certain BH3-only proteins (termed 'activators') directly engage BAK and BAX, triggering their oligomerisation . The model proposes a 'hit-and-run' transient interaction, which induces the conformational changes necessary to transform monomeric proteins into pore-forming ones (Eskes *et al.*, 2000, Wei *et al.*, 2000, Dai *et al.*, 2011, Li *et al.*, 2017). Molecular support for this mechanism is mostly based on liposome permeabilisation assays, where addition of certain BH3's accelerates dye-release (Leshchiner *et al.*, 2013). Some structural studies also suggest transient interactions between BAK/BAX and BH3 motifs based on NMR chemical shift experiments (Gavathiotis *et al.*, 2008, 2010, Moldoveanu *et al.*, 2013). The alternative indirect activation model specifies that the proteins do not require activation, but instead need to be kept in check because of their spontaneous propensity to oligomerise at the membrane. It is interesting to consider the physico-chemical implications of this 'hit-and-run' mechanism. Since binding is transient, any energy gained from the binding is lost upon dissociation. Thus, the net energetic gain is null, and the free energy necessary to the oligomerisation has to come from the protein itself. In other words, BAK and BAX monomers have to be thermodynamically unstable, and activators only act as catalysts of the oligomerisation. Since catalysts only accelerate already favourable reactions, the monomeric forms of BAK and BAX would represent kinetically-trapped, metastable, states of the proteins. Therefore, while these activators would increase the rate of reaction, oligomerisation should still proceed in their absence, just slower.

At odds with this model, an absence of oligomerisation was observed over months when stocks (tens of µM) of these proteins were kept at 4 °C. The same as been reported by others (Czabotar *et al.*, 2013), suggesting that the reaction is not spontaneous in buffer. These results contradict the underlying physico-chemical requirements of the 'hit-and-run' mechanism. However, biological reactions (*e.g.* folding and binding) can have significant

heat capacities, thus the effect of temperature can be complex, since both the enthalpy and entropy of the reaction are affected. In order to test whether the absence of oligomerisation at 4 °C was a reflection of the lack of spontaneity at more physiological temperatures—and not because of altered thermodynamic parameters—the effect of temperature on BAK and BAX was investigated (Fig. 7.16).

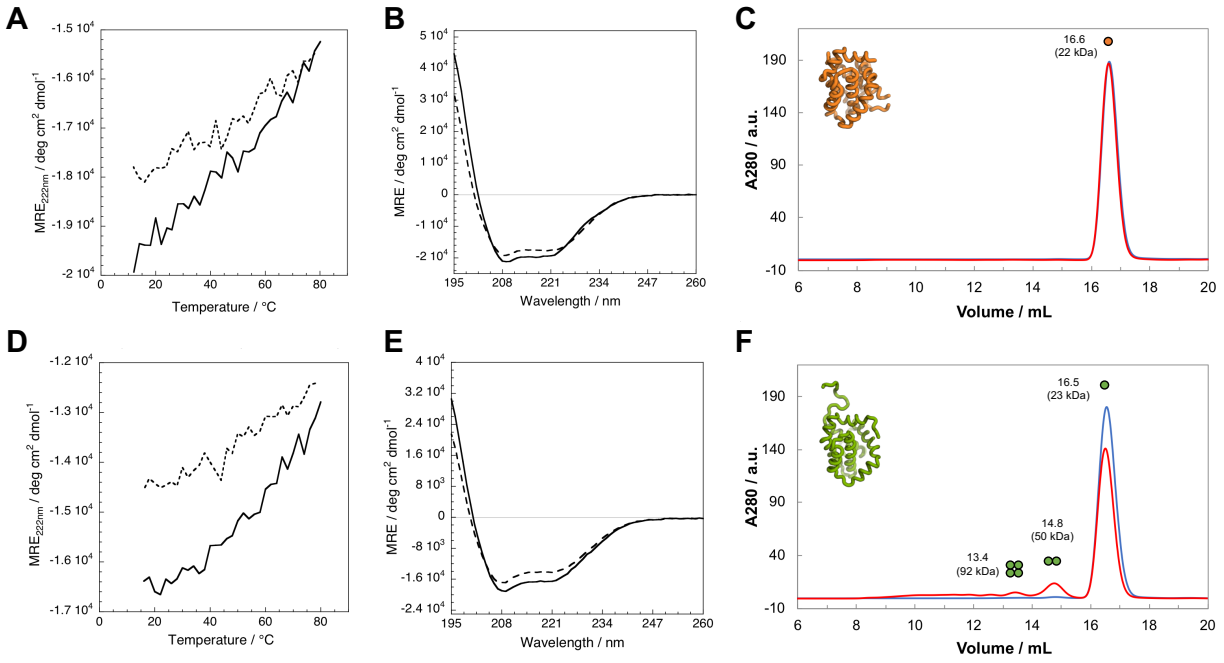

Figure 7.16 – Effect of temperature on BAK (**A**–**C**) and BAX (**D**–**F**). (**A** and **D**) temperature ramps (heating, solid line; cooling, dashed line) up to 80 °C followed by CD at 222 nm. Protein concentrations were 3 μM. (**B** and **E**) CD scans before (solid lines) and after (dashed lines) temperature ramping. (**C** and **F**) SEC analysis of 50 μM samples incubated overnight at room temperature (blue) or 50 °C (red). Experiments were performed on a Superdex 200 10/300 column. Approximate masses were calculated from the elution volumes and the calibration curve.

Heating either proteins up to 80 °C did not result in unfolding; MRE values at 222 nm were still reporting on the presence of significant helicity. Moreover, no cooperative transition could be observed. Thus, these small losses in structure are not consistent with oligomerisation or unfolding of the proteins, and are most likely due to helix fraying or loosening of the monomeric structure. The changes in structure appear irreversible, with the start and end points not overlaying. However, these results might be due to protein loss following aggregation at higher temperatures, although none could be observed by visual inspection. To confirm that these spectral changes were not related to changes in assembly state, SEC analysis of BAK and BAX following overnight incubations at either room temperature or 50 °C were performed. Strikingly, BAK remained entirely monomeric following these treat-

ments, while BAX appeared to undergo some slight oligomerisation—mostly to dimer—at 50 °C, but not at room temperature.

Higher temperatures should increase the rate of oligomerisation if the process was spontaneous, making it detectable. However, no oligomerisation was observed for BAK, and only minimal transformation was revealed for BAX. Thus, it can be concluded that oligomersiation in buffer is *not* a spontaneous process. It is noted that since the reaction is multimeric, concentration would influence the position of the equilibrium. However, even at 50 μM, oligomerisation did not appear to proceed spontaneously. Since this is much higher than the expected concentration of these proteins in cells (low nM concentrations have been reported for BAX (Polster *et al.*, 2003)), the process would be even less favourable at physiological concentrations.

7.5.2 Oligomerisation does not require 'activators'

These thermodynamic considerations appear to invalidate the 'hit-and-run' model, thus implying an absence of role for these 'activators'. Analysis of the interaction between BAK and BH3 motifs presented in the previous chapter revealed extremely weak affinities, despite fast associations. This resulted in complexes with very short lifetimes. While these findings appear in line with a 'hit-and-run' mechanism, only formation of oligomers would prove the model. This hypothesis was tested by incubating BAK for two days in the prescence of a 10-fold excess of either BID_{BH3} or $PUMA_{BH3}$. The experiment was performed in the presence and absence of detergent, and oligomer formation was monitored by SEC (Fig. 7.17).

When the experiments were performed in buffer (solid lines), BAK remained monomeric in the presence of either BID or PUMA. Evidently, presence of a large excess of these 'activator' BH3 did not trigger any form of oligomerisation. Thus, they are not sufficient for promoting oligomerisation, as expected from the thermodynamic analysis described in the previous section. The same experiments performed in the presence of PS20 (dashed lines) did show conversion of BAK monomers to oligomers. However, the results in the presence of peptides were not different from the control done in their absence. In fact, a slight reduction of the extent oligomerisation appears to occur; −7% and −12% from the integration of the PUMA and BID chromatograms respectively. This is similar to the observation made for BAK_{BH3} (*cf.* Section 7.4.2). These reductions were echoed by

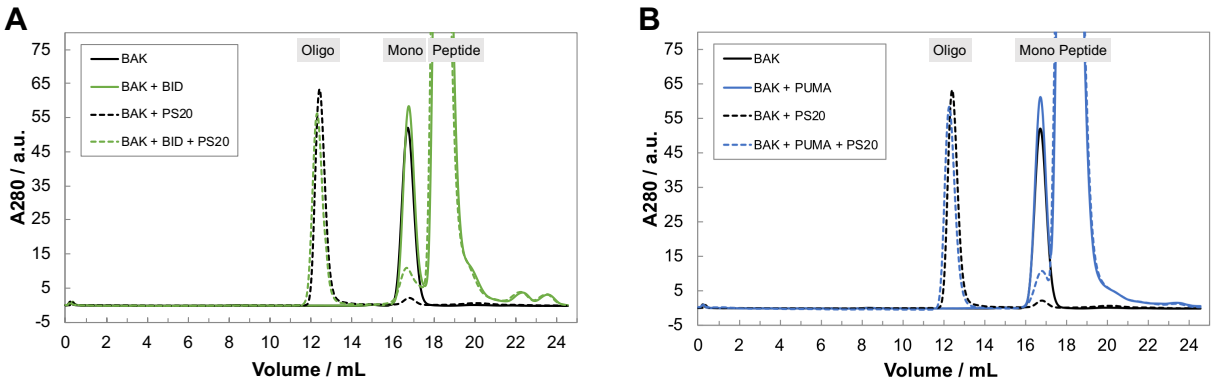

Figure 7.17 – BH3 peptides do not promote BAK oligomerisation. (**A**) Effect of t-BID_{BH3} on BAK. (**B**) Effect of t-$PUMA_{BH3}$ on BAK. Concentrations of protein and peptide were 10 μM and 100 μM respectively. Experiments were performed in the presence and absence of PS20 (20×CMC). Results were obtained on a Superdex 200 10/300 (equilibrated in 50 mM sodium phosphate pH 7.0) after two days of incubation at 25 °C.

increases in monomer peak sizes; although this might be due to peptide binding, since TAMRA absorbs at 280 nm. At these concentrations, 30% of BAK would be expected to be in complex in the absence of competing reactions. These results clearly demonstrate that BID and PUMA are not 'activators'.

Interestingly, it also suggest that binding of BH3 motifs to BAK or BAX, far from being activating, might instead be inhibitory. This hypothesis is supported by a recent report in the literature of the generation of an inhibitor of BAK oligomerisation (Brouwer *et al.*, 2017). The authors synthesised BH3 peptide mimicks containing non-natural amino acids that were capable of binding monomeric BAK with sub-μM affinities. These peptides bound at the canonical groove, and prevented oligomerisation. While the mechanism of this inhibition might be the prevention of BH3-in-groove dimer formation, as claimed by the authors, it is interesting to consider the thermodynamic hypothesis. If indeed the detergent acts by destabilising the monomer, then the free energy gained from binding could stabilise the fold. This would also prevent oligomerisation, thus equally explaining the outcome. Interestingly, this binding-induced stabilisation would further support the evolutionary hypothesis postulated in Chapter 6; binding of BH3 motifs to BAK and BAX would have been selected *against* to ensure that BAK and BAX remain on the cusp of stability.

These results demonstrate that binding of activator BH3 peptide does not trigger oligomerisation, disproving the 'hit-and-run' mechanism. Instead, binding at the canonical groove appears to inhibit the process, although more experiments would be warranted to confirm

these results. Together with the aforementionned thermodynamic considerations, these results demonstrate that oligomerisation of BAK and BAX cannot be triggered in buffer because it is not spontaneous. Instead, it appears to require the presence of a hydrophobic environment in order to occur, in line with the indirect activation model. The mechanism of this detergent-induced oligomerisation was investigated in the next section.

7.5.3 Thermodynamic effect of detergent

The resulted presented in the previous section, together with those of Chapter 6, clearly established that BAK and BAX are innocuous in buffer; they do not oligomerise, do not interact, and cannot be triggered. Yet, the presence of detergent completely transforms their assembly states. This leads to the obvious thermodynamic conclusion that the monomer is energetically favoured in buffer, and the presence of a hydrophobic environment—resulting from the addition of detergent—shifts that equilibrium in favour of the oligomer. An attempt at quantifying the thermodynamics of this phenomenon was made by looking at the partitioning of monomer and oligomer as a function of detergent concentration. BAX was chosen for the study because its oligomerisation was more gradual than BAK; allowing a better quantification of the transition.

Assuming an arbitrary 2-state homo-oligomerisation reaction involving n-subunits:

$$nA \rightleftharpoons A_n$$

where $K = [A_n]/[A]^n$, the following relationship can be written:

$$K = \frac{P(1-f)}{n(fP)^n} = \exp\left(\frac{-\Delta G}{RT}\right) \tag{7.1}$$

where K is the equilibrium constant, P the total protein concentration expressed in terms of monomer, f the fraction monomer, ΔG the free energy of oligomerisation, R the gas constant, and T the thermodynamic temperature. After five days of incubation, the extent of oligomerisation under each condition was assessed by SEC (Fig. 7.18A). Fraction monomer was calculated by curve integration. Free energies were calculated from the extent of conversation using Equation (7.1), and assuming oligomerisation to different assembly states ($n = 2, 3, 4, 5, 6$, Fig. 7.18B).

Regardless of the presumed oligomeric state, the free energy of the process appeared linearly dependent on detergent concentration in the transition region. Exceptions to that

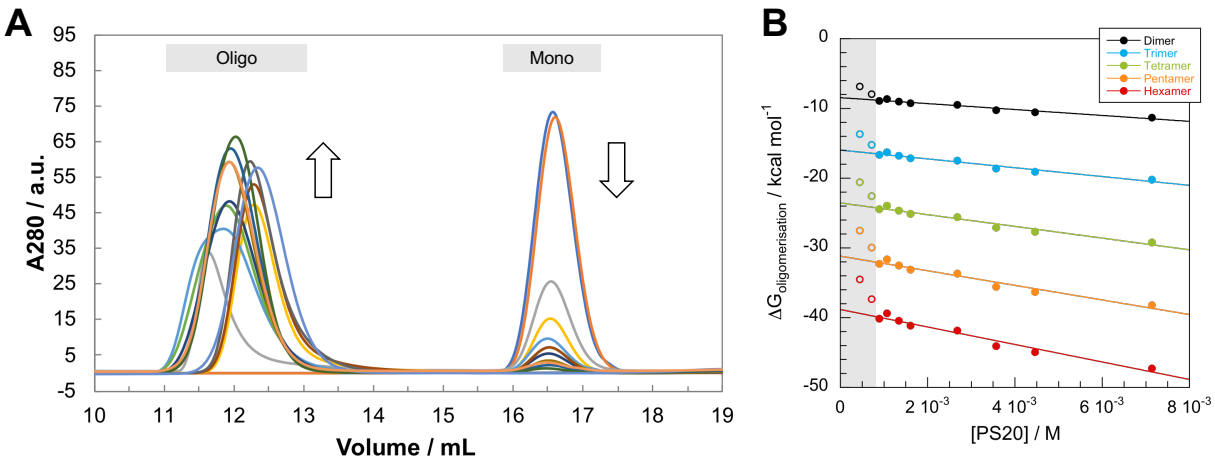

Figure 7.18 – Thermodynamic effect of PS20 on BAX. (**A**) BAX (15 µM) was incubated in the presence of PS20 (1–111×CMC) for 5 days at 25 °C, before analysis by SEC (Superdex 200 10/300). Increasing amounts of detergent gradually depleted the monomer in favour of the oligomer(s). (**B**) Monomer:oligomer ratios were obtained by integrating the chromatograms in **A**. Fractions were converted to free energies using Equation (7.1), assuming exclusive 2-state equilibria between monomer and the oligomeric state indicated in the legend. The free energies as a function of PS20 showed linear dependence for [PS20]/[BAX] >50. Values below this threshold (shaded area) were not included in the fit.

observation were found for concentrations where the detergent:protein ratio was less than fifty (shaded area). However, this is not surprising considering the observation made for BAK that a minimum ratio was necessary to allow oligomerisation to proceed properly (Fig. 7.4). Since fraction conversion could only be determined in the transition region, it was not possible to assess whether this linearity held over a wider range of detergent concentrations.

While it is interesting to draw parallels with the effect of chemical denaturant on protein folding stabilities, the analogy is limited. Indeed, the linear relationship between denaturant concentration and folding free energy is routinely used determine the stability of protein in buffer using the linear extrapolation method. Clearly, this linearity does not hold for all concentrations of detergent. If it was the case—and looking at the intercepts of Fig. 7.18B—the implication would be that the oligomers are the thermodynamically favoured states in buffer. Results shown in the previous section clearly demonstrate that this is not the case. Therefore, the effect of detergent is not as simple as that of chemical denaturation, with linearity only holding for certain concentration windows.

It is noted that this analysis is an over-simplification of the reality. Indeed, free energies were calculated assuming conversation to ‘pure’ oligomeric states. This is clearly not the

case given the heterogeneity of oligomeric states reported in this Chapter. Moreover, the distribution of states will likely be influenced by the concentration of detergent (*N.B.* the shift in elution volumes does not follow a trend, thus excluding an obvious relationship between detergent concentration and size of the oligomer). However, it does provide some insights into the role of PS20 in the oligomerisation of BAX.

In summary, oligomerisation in buffer appears to be thermodynamically unfavourable. Presence of detergent inverses the position of the equilibrium, favouring the oligomeric state over the monomer; possibly through stabilisation of the oligomer by providing a hydrophobic environment. The apparent linear dependence of the free energy to detergent concentration appears to support this hypothesis, although more experiments and better models are warranted to confirm it. Further thermodynamic analyses are reported in Chapter 9.

7.5.4 Oligomerisation is rate-limited by detergent-induced conformational changes

Oligomerisation is a slow process

In order to gain further insights into the mechanism of oligomerisation, kinetic analyses of the process were undertaken. As a first step towards characterisation, a time-dependent SEC analysis was performed. Protein and detergent concentrations were kept constant, and the extent of oligomerisation followed by doing SEC time-points.

Perhaps surprisingly, both reactions were slow and in the order of multiple hours under the conditions of these experiments. While this SEC method lacks temporal resolution—making a detailed kinetic analysis impossible—the time-evolution of the oligomerisation appears to be consistent with a first-order process. Fitting the data to a single exponential decay function yielded rate constants of 5.1 10^{-4} s^{-1} and 1.2 10^{-4} s^{-1} for BAK and BAX respectively. These correspond to half-lives of 22.7 min and 96.3 min, highlighting the slowness of the processes. Both k_{obs} were within 5-fold of each other, and since protein and detergent concentrations were identical, the intrinsic biophysical characteristics of the two proteins must be the origin of the observed differences. These biophysical disparities were further highlighted by their thermodynamics; while the conversation of BAK was close to stoichiometric at equilibrium, only three quarters of BAX monomers underwent oligomerisation under the same conditions.

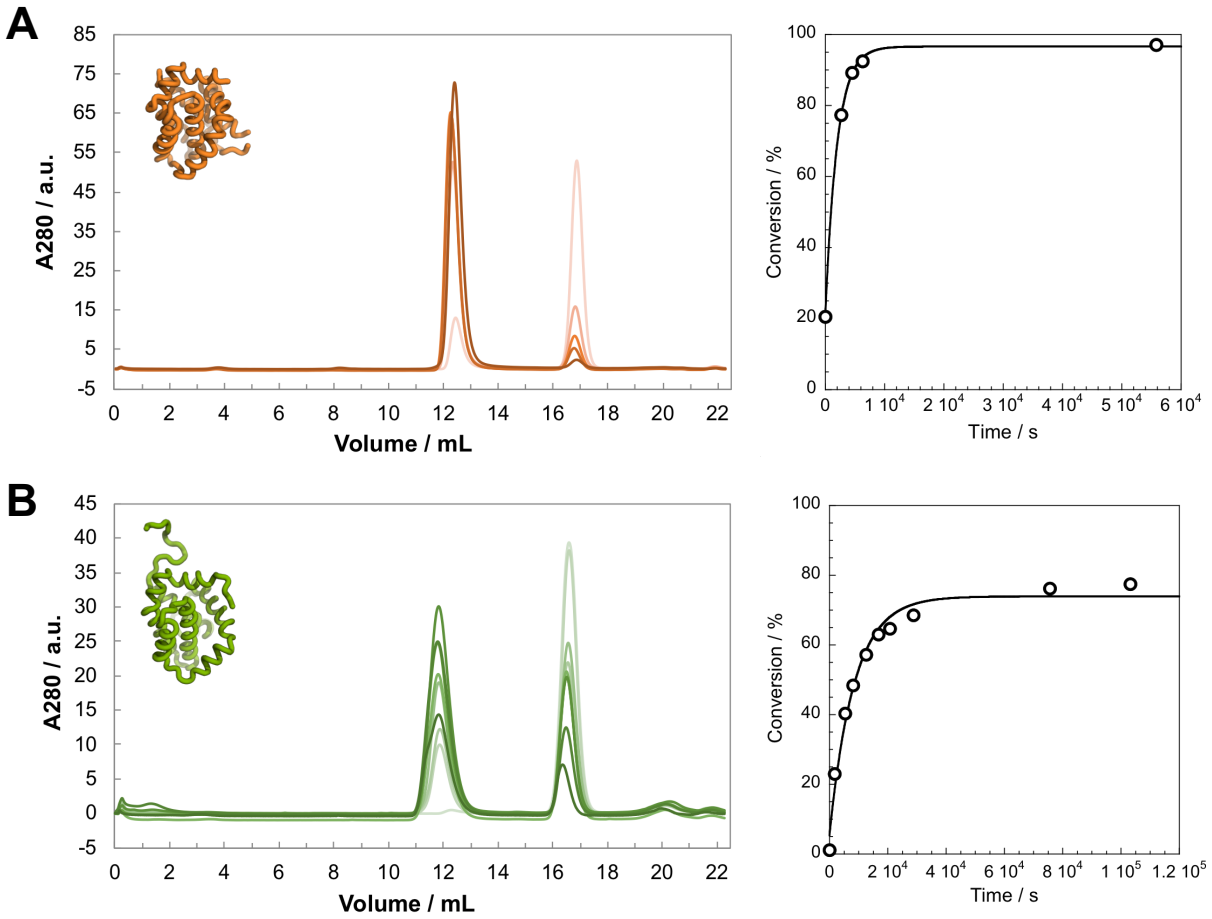

Figure 7.19 – Oligomerisation of BAK (**A**) and BAX (**B**) in detergent is slow. Protein concentrations were 18 µM, and oligomerisation was induced by addition of PS20 (22×CMC). The extent of conversion at each time-point was analysed by SEC on a Superdex 200 10/300 (run time ∼30 min, time refers to the moment of injection). Time-evolution is shown by colour gradients from light to dark. Right panels depict the percentage conversation (obtained by integration) as a function of time. Solid lines represent fits to a single exponential function (BAK: k_{obs} = 5.1 10^{-4} s^{-1}, BAX: k_{obs} = 1.2 10^{-4} s^{-1}).

Oligomerisation being a multi-molecular process—potentially involving many intermediates—a complex kinetic profile might have be expected. Interestingly, these results show a striking simplicity, suggesting a simple 2-state mechanism. The fact that the reaction appears first-order in spite of its multi-molecularity might suggest an unimolecular rate-limiting step, *i.e.* a conformational change or partial unfolding. However, given the method of detection, the resolution of subtle steps, or the presence of intermediates might not be achievable. Thus, a more sensitive method was required.

Spectroscopic signatures of oligomerisation

Spectroscopic techniques are well-suited for experiments that require accurate temporal resolutions. Provided that a signal change can be observed for the process of interest, spectroscopic probes allow kinetic events over a wide range of time-scales to be measured. For proteins, two intrinsic probes are readily at hand. Tryptophan (and to a lesser extent tyrosine) fluorescence is sensitive to the environment of the residue—informing on its tertiary context—while electronic spectroscopy of the amide bond using circularly polarised light reports on the secondary structure content of the protein. Both methods were tested for BAK and BAX, and the spectra of monomers and detergent-treated oligomers compared (Fig. 7.20).

A large fluorescence change was observed for both proteins (Fig. 7.20C, D). Interestingly, most of the change was in intensity, not in average emission wavelength. BAK shifted from 338 to 344 nm and BAX from 344 to 345 nm. These values are all consistent with 'buried' tryptophans, thus supporting the notion that the proteins remain folded upon addition of detergents. Unfolding would lead to a much more pronounced red-shift of the average emission wavelength (∼360 nm). The fact that BAX has a more red-shifted average emission wavelength than BAK in the monomeric state is most likely due to the presence of two extra solvent-exposed tryptophans (BAK has three tryptophan residues, while BAX has five). Solvent-exposure and unfolding being excluded, it can be concluded that the change in fluorescence intensity is a consequence of oligomerisation. The emission spectra of BAK were independent of PS20 concentration, confirming that these spectral changes were related to conformational changes of the protein. If local interactions between detergent and tryptophans were the cause of the change, then a concentration-dependent shift would be expected. For BAX, differences were observed between C12E8 and PS20. However, at theses detergent concentrations the proteins were not fully oligomeric, hence the results represent averages of monomer and oligomer emission spectra.

Circular dichroism spectra of BAK and BAX monomers show typical α-helical profiles, with dips at 208 and 222 nm (Fig. 7.20E, F, black lines). Importantly, the oligomers remain folded in either PS20 or C12E8 at a range of their concentrations. This is evident from the retention of strong α-helical profiles, although structural changes are clearly discernible. These are illustrated by plotting the differential between monomer and oligomer spectra (Fig. 7.20G, H). These changes are independent of the nature or concentration of the

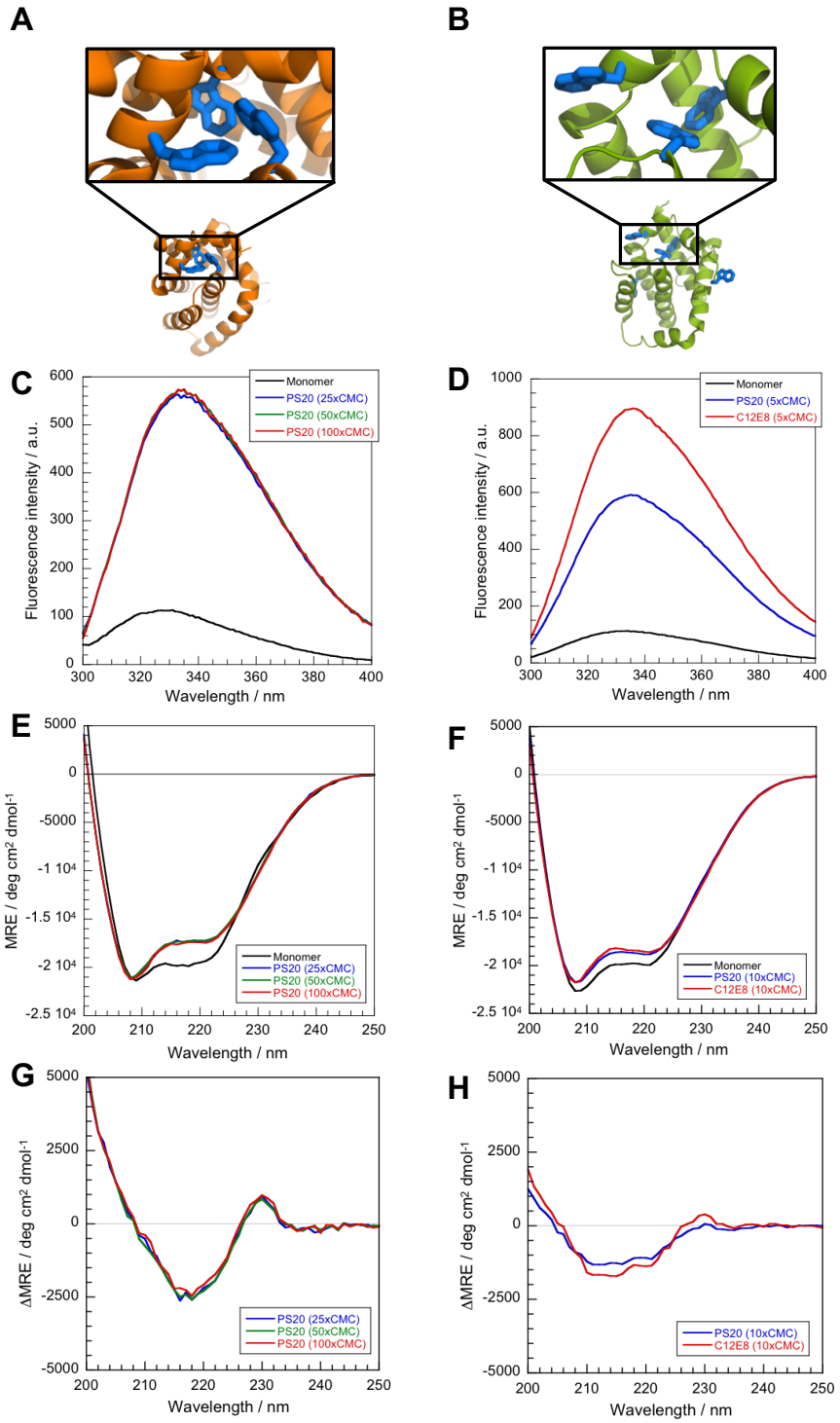

Figure 7.20 – Spectroscopic signatures of BAK (left) and BAX (right) monomers and oligomers. (**A**) Structure of BAK (PDB:2YV6). (**B**) Structure of BAX (PDB:1F16). Tryptophan residues are shown in blue, and the clusters highlighted. (**C**, **D**) Fluoresence emission spectra of monomers (black lines) and oligomers (coloured lines). Excitation wavelengths were 295 and 280 nm for BAK and BAX respectively. (**E**, **F**) Circular dichroism spectra of monomers (black lines) and oligomers (coloured lines). (**G**, **H**) Spectral differences between monomers and oligomers. Protein concentrations were 5 and 10 µM for BAK and BAX respectively. Detergents concentrations are indicated in the legends as multiples of their CMC's. All spectra were buffer-subtracted.

detergent—indicative of a switch to a distinct state, and are not consistent with gradual unfolding.

While at first glance these changes suggest a slight loss of helicity—due to a reduction of the MRE value at 222 mm—closer inspection indicates the presence of another phenomenon as well. Indeed, loss of helicity should result in reduction of both the 222 and 208 values. This is clearly not the case for BAK, although less obvious for BAX. Fig. 7.20G shows evidence of a negative-positive Trp π–π* exciton coupled band at 216 and 228 nm (Wu *et al.*, 2009, Dagil *et al.*, 2012). This can be explained by the presence of a tryptophan cluster (Highlighted in Fig. 7.20A). These Trp edge-to-face arrangements give rise to specific spectral signatures in the far-UV, which is otherwise dominated by the amide signal. Thus, it can be concluded that these arrangements are lost/modified upon oligomerisation. This signal change is weaker for BAX (Fig. 7.20H), probably due to the fact that only two of its tryptophans participate in edge-to-face interactions (Fig. 7.20B). It is noted that the spectral changes resulting from the oligomerisation are not fully accounted for by these excitons, which otherwise would have a negative-positive symmetry. Hence, additional structural changes do occur as well.

Together, these results demonstrate that BAK and BAX remain folded in the oligomeric state, and that oligomerisation is associated with specific spectroscopic signatures. Therefore, both fluorescence and circular dichroism spectroscopy can be employed to follow reaction kinetics.

Kinetics of oligomerisation

The kinetics of oligomerisation of BAK and BAX was investigated in the presence of the detergents PS20 and C12E8. Monomeric proteins and detergents were manually assembled, and the reaction followed either by changes in intrinsic fluorescence, or by monitoring signal changes at 222 nm by circular dichroism spectroscopy. Manual mixing caused a dead-time of ∼20–30 s. Therefore, any event occurring during this time period would have been missed using these techniques. However, starting MRE values were identical to monomers in equilibrium experiments, suggesting the absence of fast events.

Fig. 7.21 shows examples of kinetic traces recorded on different protein-detergent systems, and monitored either by CD or fluorescence. Although the conditions were different to the experiments performed by SEC, the observed rates are broadly similar. All reactions were slow, and oligomerisations took hours to complete ($t_{1/2} \sim 30$ min). Thus, the signals

observed spectroscopically do indeed report on the oligomerisation event itself. However, the greater resolution of these methods revealed the presence of an additional kinetic phase that was not obvious from the SEC time-course experiments. Indeed, traces did not fit to a single exponential, even if a linear drift term was included. This observation was independent of the detection method. Instead, a double exponential function was required to capture the signal change properly. This additional rate was within an order of magnitude slower than the phase previously observed by SEC (*N.B.* the assignment of the rate constants as 1 and 2 is purely arbitrary, and does not reflect a specific sequence of event). Clearly, the kinetics of oligomerisation is more complex than a simple first-order process. Differences between proteins, detergents, concentrations and detection methods all potentially influencing the rates, a more systematic approach was required. Because of its better characterisation, and ease of purification, efforts were focused on BAK. However, given these preliminary kinetic results, and the general similitude between the two proteins, BAX would probably show comparable observations.

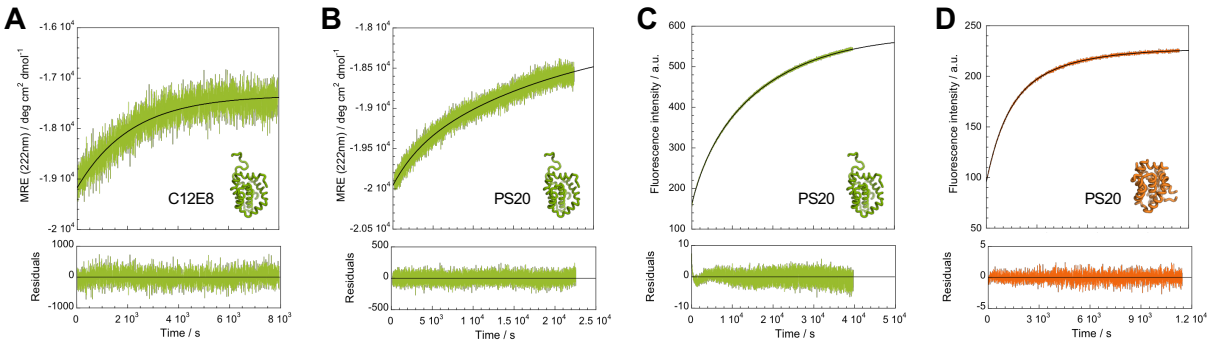

Figure 7.21 – Examples of oligomerisation kinetics. Different spectroscopic methods, proteins, detergents, and concentrations are all illustrated. (**A**) BAX (1 μM) in the presence of C12E8 (10×CMC) followed by CD. (**B**) BAX (10 μM) with PS20 (10×CMC) followed by CD. (**C**) Same experiment as **B** but followed by fluorescence. (**D**) Oligomerisation of BAK (1 μM) in PS20 (10×CMC) detected by fluorescence. Black lines represent fits to a double exponential function (except **A**, which is a single exponential). Lower panels show the residuals to the fits. Fitted values are reported in Table 7.1.

First, the potential differences between the two spectroscopic probes was assessed. While fluorescence reports on the general environment of the tryptophans (and tyrosines), CD describes the secondary structure content of the protein; although the presence of excitons implies that some of the observed signal change could be due to the trypthophans as well. Fig. 7.22 shows the results of a series of experiments where the concentration of BAK was kept constant, and and increasing amounts of PS20 was employed to induce

oligomerisation. Each reaction condition was followed by either monitoring the change in intrinsic fluorescence, or the variation of the 222 nm value using circular dichroism spectroscopy. The resulting kinetic traces were fitted to a double exponential function, and the values reported in Table 7.2.

Table 7.1 – Values of the fits from Fig. 7.21. Errors are fitting errors.

	BAX (1 μM) C12E8	BAX (10 μM) PS20	BAX (10 μM) PS20	BAK (1 μM) PS20
$k_1 \times 10^{-4}$ / s^{-1}	4.66(±0.09)	3.3(±0.3)	3.04(±0.06)	9.2(±0.1)
$k_2 \times 10^{-5}$ / s^{-1}	–	4.6(±0.5)	5.64(±0.03)	30.0(±0.4)
f_{Amp1}	–	0.17	0.15	0.52

Both spectroscopic methods appeared to report on the same events, and rate constants obtained from fitting either signals were broadly similar. Since fluorescence and CD relate to respectively tertiary and secondary structural changes, it can be concluded that these events occur simultaneously during oligomerisation. Interestingly, the fractional distribution of amplitude associated with each phase did not match between the two techniques. Fluorescence changes are arbitrary, and their absolute values cannot be interpreted. In contrast, absolute amplitude values obtained from CD hold structural details, *i.e.* the amount of structural change associated with each phase. Thus, it appears that most of the structural changes occur during the kinetic phase k_1. It is noted that due to the presence of the excitons, the interpretation of this structural change is not straightforward, and it cannot be ascribed simply to a loss of helicity.

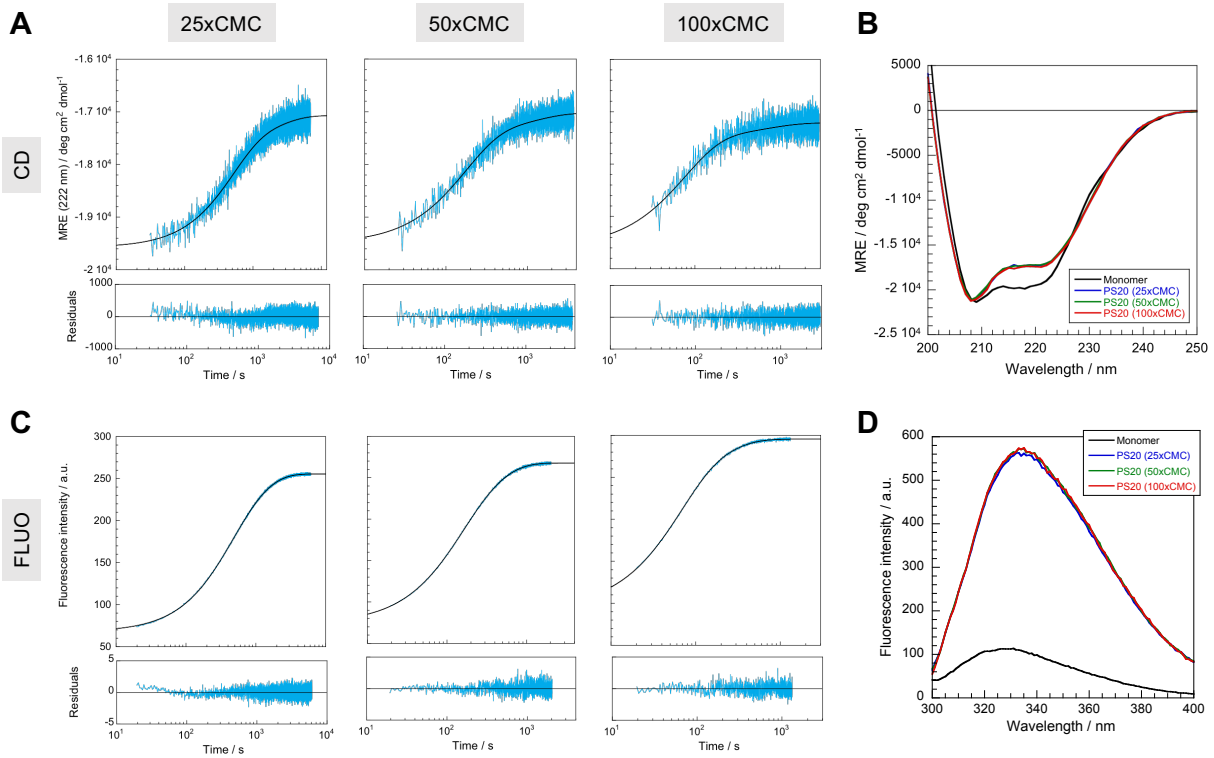

Figure 7.22 – Influence of the spectroscopic technique on kinetics. BAK (5 μM) was oligomerised in the presence of PS20 (25–100×CMC). (**A**) The reactions were monitored by circular dichroism spectroscopy at 222 nm. (**C**) The reactions were monitored by change in intrinsic fluorescence intensity. Lower panels show the residuals to a double exponential fit (black lines, values are reported in Table 7.2). (**B** and **D**) CD and fluorescence spectra of the different end points are identical, thus the final states are the same. All spectra were buffer-substracted. Note that the intial fluorescence intensity of the kinetic experiments shifts with increasing detergent concentrations. This is caused by light scattering due to the micelles.

Table 7.2 – Values of the fits from Fig. 7.22. Errors are fitting errors.

		25×CMC	50×CMC	100×CMC
CD	$k_1 \times 10^{-3}$ / s^{-1}	2.11(±0.08)	6.0(±0.2)	13.4(±0.9)
	$k_2 \times 10^{-3}$ / s^{-1}	0.47(±0.06)	0.9(±0.1)	1.53(±0.03)
	f_{Ampl}	0.76	0.82	0.86
FLUO	$k_1 \times 10^{-3}$ / s^{-1}	1.184(±0.006)	8.2(±0.1)	16.8(±0.2)
	$k_2 \times 10^{-3}$ / s^{-1}	3.01(±0.02)	3.16(±0.03)	5.41(±0.09)
	f_{Ampl}	0.52	0.54	0.73

A more detailed analysis of the effect of the spectroscopic probe on the fitted rate constants shows that while k_1's are virtually identical (less than 2-fold apart), k_2's are consistently lower when measured by CD. However, this is probably a consequence of the second phase having a much lower amplitude, and the signal-to-noise ratio being poorer for CD than fluorescence, thus making fitting less robust. Looking at the effect of detergent concentration on each rate constant reveals an interesting trend. Indeed, k_1 undergoes a ∼10-fold increase over a 4-fold change in PS20 concentration. In contrast, k_2 appears less sensitive, with only a ∼2-fold increase over the same range. The fractional amplitude was also dependent on the amount of detergent present; phase 1 became more dominant at higher PS20 concentration. This observation was independent of the spectroscopic method employed. The end-states of the different reactions were spectroscopically identical (Fig. 7.22B, D), suggesting that the increased amount of detergent did not affect the nature of the assembly. Thus, changes in kinetics must reflect a direct effect of the detergent on the rate of the oligomerisation reaction.

The effect of the detergent on the oligomerisation kinetics was investigated (Fig. 7.23). Fluorescence had a better signal-to-noise ratio, and required less protein than CD. Since the results obtained using either technique were similar, all subsequence experiments were solely monitored using fluorescence. PS20 and C12E8 were tested over a 10-fold range (10× and 100×CMC). Because oligomerisation is a multimeric process, the effect of protein concentration was evaluated by performing the reactions over a a 10-fold range as well (1 and 10 μM).

Results from these experiments lead to four main observations: *i*) Changing protein concentration had little effect; *ii*) The kinetics of oligomerisation is highly detergent-dependent; *iii*) Changing the concentration of detergent strongly affects the rate constants for PS20, but not those for C12E8; *iv*) Both rate constants were similarly affected when changing conditions, regardless of the dimension considered (nature of the detergent, protein concentration, excess CMC). The fold-changes along each dimension for each detergent are illustrated in Fig. 7.23 (right-hand side diagrams).

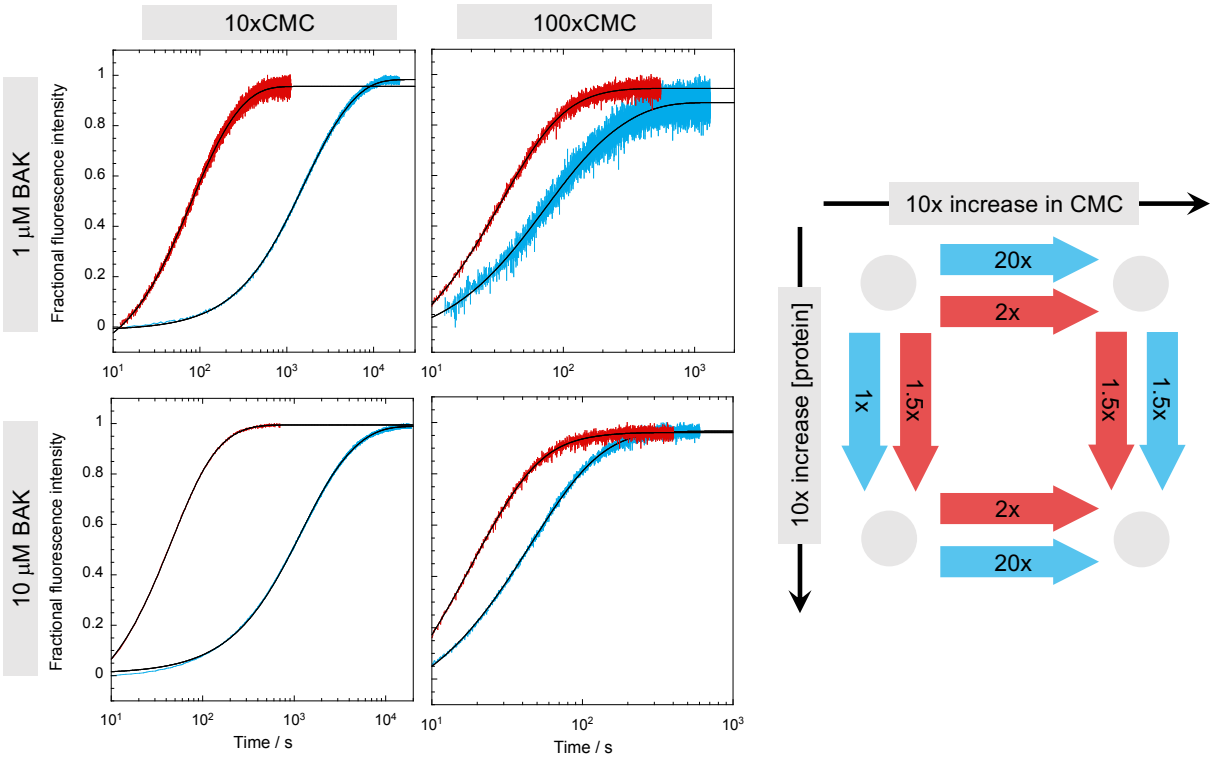

Figure 7.23 – Effect of detergent on the oligomerisation kinetics of BAK. Both C12E8 (red) and PS20 (blue) were tested over a 10-fold range against BAK. The protein concentration was also varied over a 10-fold range. Black lines represent fits to a double exponential function (values of the fits are reported in Table 7.3). Note the different time-scales. Diagrams on the right-hand side illustrate the approximate fold-changes when switching between the protein-detergent conditions indicated by the arrow. Fold-changes were similar for k_1 and k_2.

Table 7.3 – Values of the fits from Fig. 7.23. Errors are fitting errors.

		10×CMC		100×CMC	
		C12E8	PS20	C12E8	PS20
1 µM	$k_1 \times 10^{-2}$ / s^{-1}	1.88(±0.04)	0.1004(±0.0009)	3.31(±0.07)	1.91(±0.07)
	$k_2 \times 10^{-3}$ / s^{-1}	6.56(±0.08)	0.326(±0.002)	13.1(±0.7)	5.8(±0.1)
	f_{Ampl}	0.48	0.45	0.78	0.55
10 µM	$k_1 \times 10^{-2}$ / s^{-1}	2.55(±0.04)	0.1123(±0.0004)	5.70(±0.07)	2.46(±0.05)
	$k_2 \times 10^{-3}$ / s^{-1}	10.0(±0.3)	0.362(±0.001)	18(±1)	11.9(±0.6)
	f_{Ampl}	0.71	0.57	0.90	0.76

Perhaps surprisingly—despite spanning an order of magnitude in protein concentrations—the observed rate constants for 1 and 10 μM BAK were within 2-fold of each other. This was true regardless of nature or the amount of detergent present. The rates do show a slight concentration-dependence, excluding 'pure' unimolecular processes for these kinetic phases. However, it also exclude transitions states with high molecularities of monomers; these would be expected to have high concentration-dependences. In stark contrast to this weak dependence on protein concentration, changing the concentration of PS20 had a dramatic effect on both observed rate constants. It is noted that complete oligomerisation occurs for all protein:detergent ratios tested here. Therefore, partial conversions can be excluded as the source of these differences. Instead, it is evident that the concentration of detergent directly affects the free energy profile of the reaction, in line with its thermodynamic effect (Fig. 7.18). Interestingly, the magnitude of this effect was about 10-fold greater for PS20 than C12E8. The origin of this difference is unclear; the fact that BAK forms higher-order species in C12E8, or distinctive protein:detergent energetics could both potentially explain theses results.

Clearly, the mechanism of oligomerisation is a complex process. The presence of two rate constants, each dependent on multiple parameters highlight this intricacy. Interestingly, not all variables have the same effect; shown by the differential sensitivity of the rate constants to changes in conditions. A systematic approach was undertaken to deconvolute the oligomerisation reaction with respect to these different dimensions. A matrix of detergent-protein conditions was tested. Because of the less well-characterised end-state in C12E8, only PS20 was used in these experiments. Protein concentrations span 100-fold, while the concentration of detergent was varied 4-fold. Enough was required to ensure complete oligomerisation at every protein concentration—imposing the ∼40:1 detergent:protein ratio. However, too much could not be used either otherwise the reactions became to fast to be accurately measured. Therefore, a relatively limited window was amenable for experimental characterisation by manual mixing on a fluorescence spectrophotometer. Each condition was measured in duplicate, and the traces individually fitted to a double exponential function. The values were averaged and plotted against the two dimensions probed in this experiment: protein concentration and excess CMC of detergent.

Looking at the dependence of both rate constants with respect to each dimension revealed interesting trends, and provided insights into the mechanism of oligomerisation. Both rate constants appeared to have power-law dependences to the protein concentra-

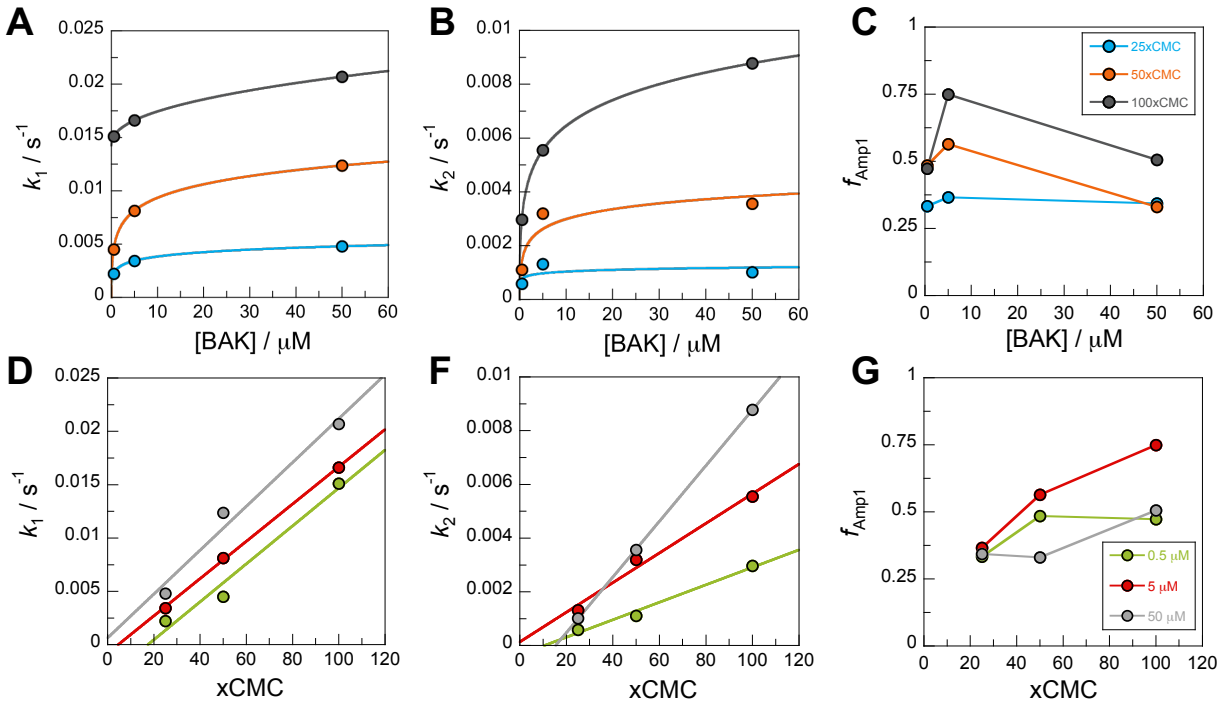

Figure 7.24 – Reaction matrix for deconvoluting the effect of protein and detergent concentration on the kinetics of oligomerisation. BAK (0.5, 5, 50 µM) was assembled with PS20 (25, 50, 100×CMC), and the reactions followed by intrinsic fluorescence. Each protein-detergent condition was individually fitted to a double exponential function. Values from two independent experiments were averaged, and the results plotted against one of the variables. (**A**–**C**) Plots of k_1, k_2 and the fractional amplitude of k_1 against the concentration of protein. Each data series represents a concentration of detergent. (**D**–**G**) k_1, k_2 and the fractional amplitude of k_1 against the concentration of detergent (indicated as a multiple of its CMC). Each data series represents a concentration of protein. (**A** and **B**) Solid lines represent fits to a power function ($a + b \cdot [\mathrm{BAK}]^c$). (**D** and **E**) Solid lines represent fits to a linear function. (**C** and **G**) Points were joined by straight lines to aid visualisation.

tion, irrespective of the amount of detergent present (Fig. 7.24A, B). The exponents being smaller than one under all conditions tested, transitions states with high molecularities of monomer can be excluded. Instead, it shows that the oligomerisation reaction becomes almost concentration-independent beyond ∼5 µM. This dependence was less pronounced at lower excesses of CMC (blue lines), suggesting that under these conditions the detergent might become more rate-limiting. The dependencies of k_1 and k_2 on protein concentration were broadly similar, therefore suggesting a 'uniform' effect on the oligomerisation reaction itself.

In stark contrast, the dependence of the rate constants on the concentration of detergent showed linear dependencies. These results demonstrate that PS20 has a much greater impact on the oligomerisation reaction than the protein concentration itself—despite the

reaction being multi-molecular in nature. Moreover, the dependence of k_1 was independent of protein concentration; indicated by the fact that the three lines in Fig. 7.24D are parallel (although this might no longer be the case if the intercepts were constrained to zero). This suggests that phase 1 might represents a unimolecular process that is directly affected by the detergent,*e.g.* a PS20-induced conformational change or unfolding event. However, the different intercepts also suggest that changing protein concentration affects a 'basal' component of the reaction. The dependence of k_2 was also linearly dependent on the amount of detergent, but each protein concentration series had a different gradient. Therefore, this kinetic phase is not concentration-independent. These results suggest that it is a multi-molecular possess, which free energy is logarithmically-dependent on detergent concentration, *e.g.* dimers assembling into larger oligomers.

By taking these results and making some simplifying assumptions, a model of the oligomerisation process can be drafted. From the plots of the rates as a function of protein concentration (Fig. 7.24A, B), the following relationships can be expressed:

$$\begin{cases} k_1 = k_{p1}[M]^x + C_{p1}([d]), & \text{where } 0 < x < 1 \\ k_2 = k_{p2}[M]^y + C_{p2}([d]), & \text{where } 0 < y < 1 \end{cases}$$

where $[M]$ represents the concentration of monomeric protein and $[d]$ the concentration of detergent. Here k_{p1} and k_{p2} are independent of detergent concentration. While not entirely true, their dependencies are relatively modest, thus this assumption might be acceptable.

From the plots of the rates as a function of protein concentration (Fig. 7.24D, F), the following relationships can be expressed:

$$\begin{cases} k_1 = k_{d1}[d]^1 + C_{d1}([M]) \\ k_2 = k_{d2}[d]^1 + C_{d2}([M]), & \text{where } k_{d2} = k'_{d2}[M]^z \end{cases}$$

Here k_{d1} is truly a constant, and the protein concentration dependence of k_{d2} is taken into account.

By combining these two sets of equations, and cancelling out the correction terms (the C_i's)—which assumes that these implicit terms are explicitly described in the other equation—the following equations can be written:

$$\begin{cases} k_1 = k_{p1}[M]^x + k_{d1}[d]^1, & \text{where } 0 < x < 1 \\ k_2 = k_{p2}[M]^y + (k'_{d2}[M]^z)[d]^1, & \text{where } 0 < y < 1 \end{cases}$$

Therefore, the observed rate constants k_1 and k_2 can be expressed as a function of protein and detergent concentrations. While simplistic and not entirely accurate given the assumptions that were made, it provides a framework for the characterisation of the oligomerisation. Looking forward, matrices with higher resolutions along each dimension (using for example a plate-reader assay), together with global fitting of the results, would allow the validity of this model to be tested.

It is noted that the presence of two rate constants could either imply sequential transition states, or parallel pathways. Moreover, the observed rate constants might not represent individual molecular events, and instead be a combinations step, *e.g.* forward and reverse reactions.

Clearly, the interplay between the different parameters makes a detailed interpretation difficult. However, fundamental observations can be made from the experiments presented in this section. The oligomerisation of BAK and BAX is clearly a slow process, which is mostly independent of protein concentration. Instead, the nature and amount of detergent is the main determinant of the reaction kinetics. Thus, a detergent-induced conformational change that is rate-limiting appears to be an appropriate model to explain the oligomerisation. Of elegant simplicity is the linear relationship between PS20 concentration and the observed rate constants. This result has interesting implications for the way membranes might affect the energy landscape of membrane proteins.

7.6 Discussion

In Chapter 6, investigations of the interactions between BAK, MCL-1 and BH3 proteins revealed a spectator role for BAK in buffer. Since it oligomerises and induces MOMP in a biological context, its function was clearly not fully re-capitulated under these conditions. This was attempted in this Chapter. The oligomerisation of both BAK and BAX—the two pro-apoptotic effector proteins of the BCL-2 family—were investigated to gain biophysical insights into this pivotal step of apoptosis.

Using detergents, both proteins could be oligomerised *in vitro*. Crucially, this detergent-

induced assembly was specific to BAK and BAX, and when MCL-1 was subjected to the same treatment it remained monomeric. Hence, these findings appear to reflect the biological role of these proteins, and validate the physiological relevance of these oligomers. This was further confirmed by introducing disulfide staples into BAK, and demonstrating that restraining the monomeric structure prevented oligomerisation. These results echo biochemical studies, which showed that the same approach prevented MOMP in mitochondria.

Different detergents lead to a range of outcomes in terms of the oligomers formed. Thus, a level of plasticity appears to underlie the oligomerisation of BAK and BAX. PS20 giving better-resolved results, subsequent work mostly focused on this detergent. Chemical cross-linking studies revealed structures as large as hexamers for BAK, and tetramers for BAX. These results were confirmed by native mass spectrometry. Interestingly, it also revealed that a range of oligomeric states were populated in the solution state. SEC had previously shown an apparently homogeneous distribution, an observation at odds with the results from cross-linking and native MS. However, chromatography performed at higher concentrations revealed an underlying heterogeneity. Thus, the apparent discrepancy at lower concentrations might be explained by the interplay between assembly states, detergents, and micelles on the elution profile—masking the presence of different oligomers.

Could this heterogeneity be an artefact of the use of detergents? Indeed, membrane-embedded proteins often show well-defined structures with elegant internal symmetries. The presence of multiple states for detergent-treated BAK and BAX contradicts this notion of molecular order. However, a recent study using time-lapse microscopy on live cells revealed that BAK and BAX proceed to form very large pores during apoptosis (McArthur *et al.*, 2018). These assemblies appear to be ill-defined in both size and shape. Thus, it seems that these proteins do not form well-defined structures in their biological contexts either; clashing with the 'crystallographic' picture of ordered oligomers. Accordingly, the results obtained here in detergent—showing an heterogeneity of oligomeric states—might in fact be a reflection of the natural plasticity of these proteins (Uren *et al.*, 2017).

Interestingly, native MS also suggested that oligomers might be formed of two energetically distinct interfaces; sturdy dimer units underscore the composition of larger assemblies. These results are consistent with another report from the literature, where the authors demonstrated that BAX oligomers in supported bilayers existed as multiples of dimers (Subburaj *et al.*, 2015). Therefore, the hierarchical structural organisation observed when

the oligomerisation was performed in PS20 appears consistent with experiments performed in membranes, further validating the relevance of these detergent-treated oligomers.

The nature of this possible dimer 'denominator' was investigated in light of structural reports from the literature. Three topologically distinct dimers have been solved for BAK and BAX, but only one is deemed physiologically relevant. Detergent-induced dimerisation of BAK and BAX had been previously reported, however these assemblies were shown to be inconsistent with the topologies found in apoptotic pores (Czabotar *et al.*, 2013). Using a combination of biophysics and disulfide cross-linking, two out of the three possibilities could be excluded as the structure formed in PS20; including the supposedly physiologically irrelevant helix-swapped dimer. This highlights the unique properties of this detergent with regard to BAK and BAX oligomerisation. The consensus in the literature—the BH3-in-groove dimer—appeared consistent with the oligomers formed in PS20, although more experiments are warranted to confirm it. Interestingly, the affinity of BAK for its own BH3 motif is extremely weak (*cf.* Chapter 6). Therefore, if this topology is indeed the nucleus of the oligomerisation, the actual BH3 is unlikely to be the driving force behind the formation of dimers. However, this might ensure the lack spontaneous homo-oligomerisation in aqueous conditions, allowing the control of the assembly to be made by membrane environments (see below, as well as the argument made in the previous Chapter about folding free energies and binding).

The ability to oligomerise these proteins *in vitro* opens new and exciting opportunities for structural characterisations. Therefore, in addition to these 'indirect' ways of looking at the oligomers, a more 'visual' approach was also attempted. Class-averages of negative-stain electron micrographs were obtained for BAK treated with PS20. These images revealed pore-like arrangements, validating the assemblies obtained in detergent, and demonstrating the exciting opportunities these findings provide. This is currently work in progress.

The mechanism, and regulation of BAK and BAX oligomerisation remains debated in the literature. Opposing models have been proposed, each with different physico-chemical implications pertaining to their validities. In the direct activation model, 'transient' interactions with specific BH3-only proteins trigger the oligomerisation, while in the indirect activation model pore-formation is spontaneous, but prevented by anti-apoptotic BCL-2 proteins. The validity, and likelihood, of these models were investigated using detergent systems in conjunction with biophysical techniques. Experiments with BH3 peptide revealed that they did not trigger oligomerisation in buffer. Because of their presumed cat-

alytic roles, this mechanism could only be valid if BAK and BAX were metastable, and the conversion to the oligomeric state kinetically hindered. This model was further disproved by demonstrating that the monomers were the most stable states in buffer. While some of the biological complexity is obviously lost in this minimal context, the thermodynamic constrains of the system appear to invalidate the indirect activation model.

While BH3 peptides were incapable of triggering oligomerisation on their own, addition of detergents readily promoted the event. A thermodynamic analysis of the effect of PS20 on the free energy of BAK oligomerisation suggested that it may act in a manner analogous to chemical denaturants; increasing amounts of detergent gradually stabilised the oligomer/destabilised the monomer. These results have important implications for the mechanism of BAK and BAX oligomerisation *in vivo*. Indeed, it suggests that monomers are only the preferred state in aqueous conditions, and that oligomers become favoured in hydrophobic environments. Therefore, these results support the indirect activation model—BAK and BAX monomers are inherently unstable, and spontaneously oligomerise in a membrane environment. This conclusion reached from studies performed with detergents is also supported by work done on a cellular system where all BCL-2 proteins were knocked-out (O'Neill *et al.*, 2016). Upon re-introduction of either BAK or BAX (without anything else), the cells went into apoptosis. Thus, it appears that these proteins do *not* require a trigger, and instead can induce MOMP through spontaneous oligomerisation at the membrane.

Finally, a mechanistic investigation of the oligomerisation in detergent was undertaken. Interestingly, this process was slow, with half-lives in the order 1–10 minutes depending on the conditions. This is compellingly close to the 5 min reported for the release of cytochrome c from mitochondria (Goldstein *et al.*, 2000, Rehm *et al.*, 2002, Albeck *et al.*, 2008), suggesting an enticing link to the mechanism of oligomerisation in a biological context. It was also discovered that the rate of oligomer formation was mostly independent of protein concentration—a surprising result given the multi-molecular nature of the reaction. A more detailed kinetic analysis revealed the presence of two phases, each being dependent on both protein and detergent concentrations. This interplay made the evaluation difficult. However, a simplified model was established, opening avenues for future work. Importantly, it was found that the detergent had a much bigger impact on the rate of oligomer formation. An apparently linear relationship between PS20 concentration and observed rate constants was noticed, revealing a surprising simplicity to the role of the detergent. These results sug-

gest that oligomerisation is mostly limited by a detergent-induced conformational change on-pathway to the assembly into larger species (illustrated schematically in Fig. 7.25). A mechanism of this nature might have interesting biological implications; putting a multimerisation reaction under (mostly) the control of an unimolecular process—and having this rate modulated by the nature of the hydrophobic environment—could provide a way of controlling apoptosis through alterations of membrane biophysics. Looking forward, the use of low-concentration single-molecule experiments might help to deconvolute the different steps of the oligomerisation. In particular, it should allow conformational changes to be separated from association reactions.

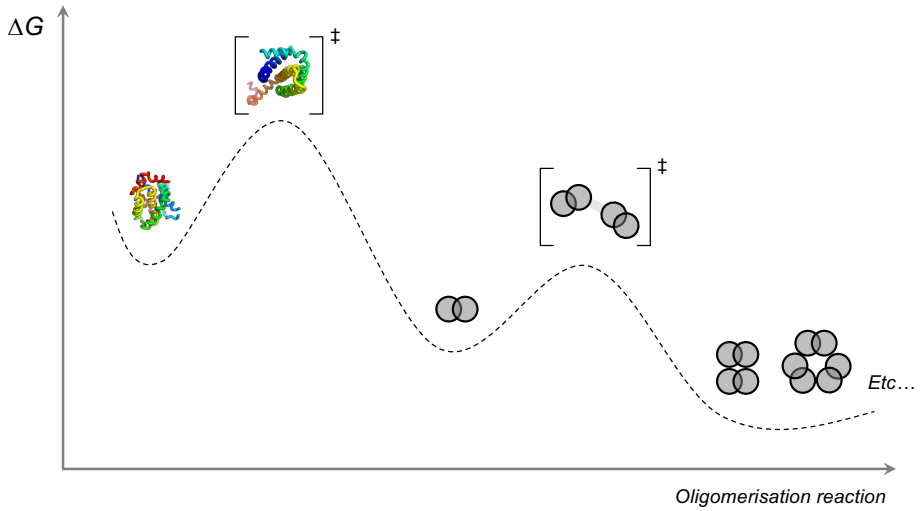

Figure 7.25 – Putative reaction diagram for the oligomersiation of BAK and BAX. The energy profile is depicted for the presence of detergent (the membrane). In buffer, the energetics of the different states would be shifted, and the monomer would the most stable species.

In summary, the results presented in this chapter demonstrate the possibility to oligomerise BAK and BAX to physiologically-relevant states in detergents. This ability offers new and exciting opportunities to study these elusive assemblies. Structural and mechanistic approaches have provided valuable insights into the regulation of this process. Moreover, it established a simplified framework to study the oligomerisation in the context of the BCL-2 network, which is the subject of the next Chapter.

Chapter 8

Mechanism of BCL-2 regulation

All native mass spectrometry results presented in this chapter were collected in the group of Prof. Carol V. Robinson (University of Oxford) with Dr Kallol Gupta, who also performed the final analysis of the data.

8.1 Introduction

In Chapter 6, the interactions of a model tripartite BCL-2 system was investigated under standard biochemical conditions. It was shown that BH3 motifs interact fast and tightly with the anti-apoptotic member MCL-1. In stark contrast—and despite high structural homology—these same motifs did not form complexes of appreciable lifetime with BAK. Thus, a ternary mixtures of these proteins at equilibrium conveys the role of spectator for BAK. Importantly, no oligomerisation was observed. Since BAK and BAX are known to assemble into pores at the mitochondrial outer-membrane, it was clear that these conditions did not recapitulate the full behaviour of these proteins.

The oligomerisation of BAK and BAX *in vitro* was achieved using detergents (Chapter 7). Characterisation of theses states revealed the formation of physiologically-relevant higher-order oligomers, based (mostly) on dimer units. Importantly, BAK and BAX monomers were shown to be stable in buffer, but the oligomers became the thermodynamically favoured states in the presence of a hydrophobic environment. These results suggest a spontaneous oligomerisation process at the membrane. Moreover, BH3 motifs did not appear to trigger the event, in contradiction with the direct activation model.

These findings raise interesting questions regarding the interplay between the underlying biophysical properties of BAK and BAX, and the regulation of the BCL-2 network.

The detergent-sensitivity of these pro-apoptotic proteins suggests that they would spontaneously induce MOMP. If so, how is apoptosis prevented in healthy cells? The indirect activation model stipulates that BAK and BAX are kept in check by anti-apoptotic BCL-2 proteins. However, no interaction between MCL-1 and BAK was observed in buffer.

8.2 Aims

Using the framework established in the previous two Chapters, the mechanism of BCL-2 regulation was investigated in the context of detergent. Interactions within the model tripartite system were re-visited to see how a membrane environment might affect protein-protein interactions. A combination of SEC, native MS, and kinetic experiments were employed to deconvolute the interactions. These results established a model whereby competition between homo- and hetero-oligomerisation within the BCL-2 family regulate the assembly of BAK/BAX.

8.3 BCL-2 interactions in the presence of detergent

The minimal tripartite BCL-2 system composed of BAK, MCL-1, and PUMA (described in Chapter 6) was used for most of the work presented in this chapter. For some experiments, BAX was studied alongside BAK, and additional BH3 peptides were also investigated. Only PS20 was used as detergent. All native MS experiments were performed with the the His-tagged version of MCL-1; this ensured a larger mass difference (compared to BAK and BAX), allowing better resolution between homo- and hetero-oligomers. Accordingly, some SEC and kinetic analyses were also performed with this construct, showing similar results to the tag-free construct.

In order to understand the tripartite interactions, the system was de-constructed into its components. The effect of detergent on individual proteins was characterised, followed by the investigation of all binary combinations, and finally the tripartite system was re-constructed. Using this approach, it was possible to deconvolute additive, from emerging properties.

8.3.1 Only BAK and BAX are affected by detergent

First, the effect of detergent on the assembly state of the different components was tested (Fig. 8.1). Monomeric proteins were incubated in the presence of PS20, and the results analysed by SEC. As described in the previous Chapter, both BAK and BAX underwent oligomerisation to higher-order species; indicated by the disappearance of the monomer peak (∼16.5 mL), and the appearance of an oligomer peak (∼12.2 mL). Note that the higher stability of BAX—about half of it remained monomeric, while under identical conditions BAK underwent stoichiometric conversion to the oligomeric state.

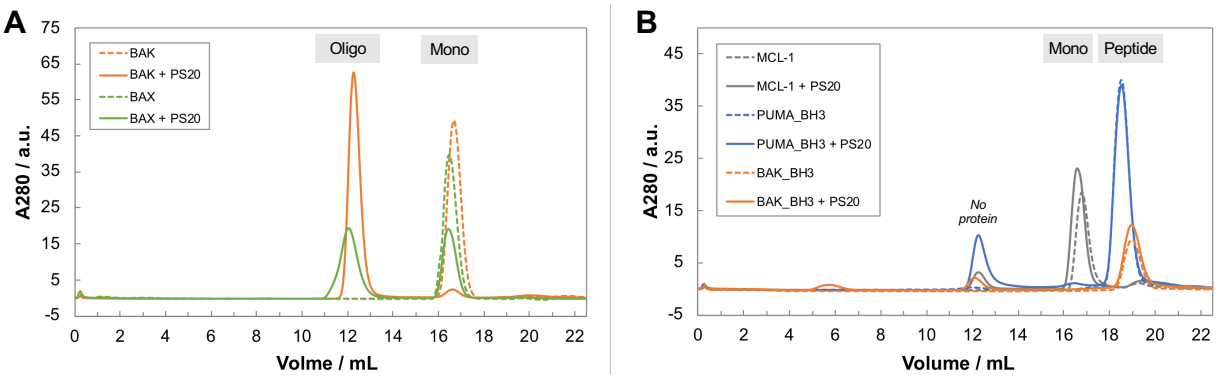

Figure 8.1 – Individually, only BAK and BAX are affected by detergent. SEC analysis of BAK, BAX, MCL-1, t-PUMA$_{BH3}$, and t-BAK$_{BH3}$ in the presence (solid lines) or absence (dashed lines) of PS20 (20×CMC). (**A**) Both BAK and BAX monomers are converted to oligomers in the presence of detergent. This is indicated by the disappearance of the monomer peak at ∼16.5 mL, and the concomitant apparition of an oligomer peak at ∼12.2 mL. Note that greater stability of BAX under these conditions. (**B**) MCL-1 and BH3 peptides are not affected by detergent. Elution profiles are consistent with monomeric states, and identical in the presence or absence of PS20. Note that the peaks at ∼12.2 mL do not contain protein (confirmed by SDS-PAGE), and are solely due to the absorbance of PS20 free micelles ($\varepsilon_{280} \approx 10$ M^{-1} cm^{-1}). Protein concentrations were 10 μM (except t-BAK$_{BH3}$, which was 5 μM). Analyses were performed on a Superdex 200 10/300 column equilibrated in 50 mM sodium phosphate pH 7.0 buffer.

In contrast, the assembly states of MCL-1 and BH3 peptides remained unperturbed by the presence of detergent (Fig. 8.1B). Elution profiles of the proteins in either buffer or PS20 overlaid, indicating the absence of oligomerisation. For MCL-1 and t-BAK$_{BH3}$, a slight increase in peak height could be observed. However, this was attributed to the detergent reducing non-specific interaction between the column matrix and the proteins. Importantly, the elution volumes remained unchanged, confirming the absence of oligomerisation of these proteins. A peak at ∼12.2 mL did appear for all injections that contained PS20. However, SDS-PAGE analysis revealed the absence of protein. Instead, the peak was attributed to

the elution of free micelles; PS20 absorbs slightly at 280 nm, and micelles also scatter light at this wavelength.

Together, these results confirm that detergent-induced oligomerisation is specific to effector BCL-2 proteins. It highlights a distinct biophysical signature that is not shared with their anti-apoptotic counterparts. Whilst it might not be surprising from the knowledge of their opposing biological functions, the similarities of their monomeric structures might suggest otherwise. Thus, it appears that the sequence—*not* the structure—determines function in the context of BCL-2 proteins.

8.3.2 Binary interactions with BH3 motifs are not affected by detergent

Having demonstrated that only BAK and BAX were individually affected by the presence of detergent, the effect of PS20 on binary protein-protein interactions was investigated. In buffer, peptides of BH3 motifs were all capable of binding tightly to MCL-1, while their interactions with BAK were too short-lived to form complexes with 'physiological' affinities. Their interactions in the presence of detergent were studied using SEC and CD, and the results compared to the outcomes obtained in buffer.

The binary interactions between MCL-1 and the two BH3 peptides tested (PUMA and BAK) appeared unaffected by the presence of PS20. SEC revealed complete binding under the conditions of the experiments—indicated by the disappearance of the peptide peak at ∼18.5 mL (Fig. 8.2A, D). Furthermore, CD showed the presence of coupled and binding in both cases, confirming the absence of structural effects from the presence of PS20 on the individual proteins or their complexes (Fig. 8.2B, C, E, F).

These results do not offer a quantification of the affinities in detergent. Thus, changes in binding rate constants and affinities cannot be excluded. However, the complete absence of free peptides at 5 μM protein concentrations observed by SEC confirms that K_{d}'s must be at least in the low-nM range. Clearly PS20 does not significantly impact the interaction between BH3 motifs and MCL-1.

In Chapter 7, BH3 peptides of BID, PUMA, and BAK were shown not to interact with the oligomeric state of BAK. However, these experiments were performed with large excesses of peptides, and it was impossible to assess whether the presence of detergent lead to the formation of monomer:BH3 interactions due to peak overlaps. SEC analyses were

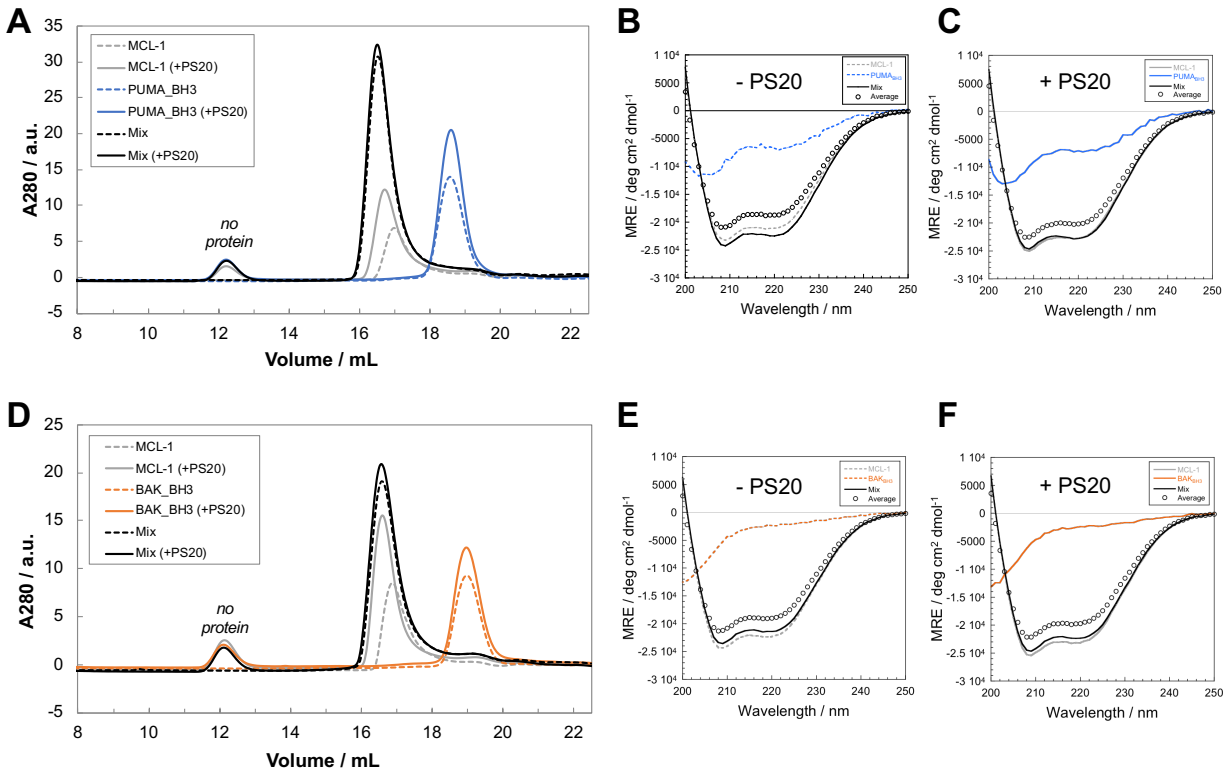

Figure 8.2 – The interaction between MCL-1 and BH3 peptides is not affected by the presence of detergent. (**A**, **D**) Formation of MCL-1:BH3 complexes monitored by SEC shows no consequence from the presence of detergent. Interaction is demonstrated by the disappearance of the free peptide peak (∼18.5 mL) in the mixture. (**B**, **C**, **E**, **F**) CD demonstrates equivalent structural interactions between and MCL-1 and BH3 peptides in the presence or absence of detergent. The spectra of the mixtures show more structure than the averages of the components, indicating coupled folding and binding. All experiments were performed at 5 μM protein concentration, and PS20 was present at 20×CMC. His-tagged MCL-1 and dye-lablled peptides were used. SEC was performed on Superdex 200 10/300.

performed under equimolar conditions (5 μM) to test for the binding of BH3 peptides to the monomeric state of BAK (Fig. 8.3A, D). The results clearly demonstrate absence of interactions, noticeable by the persistence of the free peptide peak at ∼18.5 mL. SDS-PAGE analyses and absorbance measurements at 555 nm (to test for the presence of TAMRA) were negative for both the oligomeric and monomeric peaks, thus confirming the absence of complex formation with either state of BAK.

Analysis by CD confirmed these SEC results, and revealed the absence of coupled folding and binding in detergent, consistent with the results obtained in buffer. Together, these experiment show that BH3 motifs are spectator of the oligomerisation of BAK in detergent. Moreover, it confirms that PS20 does not induce any conformational changes that make BAK able to bind BH3 peptides tightly.

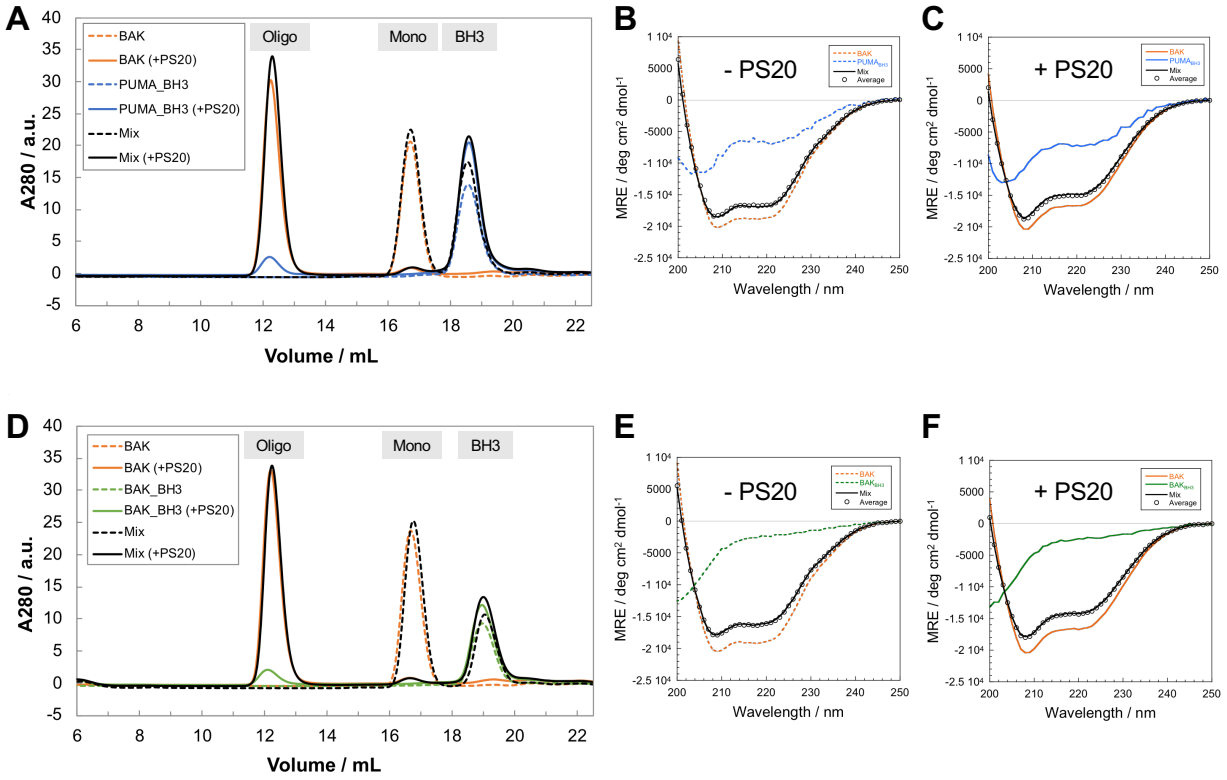

Figure 8.3 – Oligomerisation of BAK in detergent is unaffected by the presence of BH3 peptides, and they do not bind the monomeric, or the oligomeric state of BAK. (**A**, **D**) SEC analysis of BAK oligomerisation in PS20 shows no consequence from the presence of BH3 peptides. Oligomerisation is indicated by the disappearance of the monomer peak at ∼16.5 mL, and the concomitant apparition of an oligomer peak at ∼12.2 mL. Lack of interaction with BH3 peptides is shown by the persistance of the free peptide peak (∼18.5 mL) in the mixture. SDS-PAGE and A555 anaylses of the monomer and oligomer peaks confirmed the absence of dye-labelled peptides in these fractions. (**B**, **C**, **E**, **F**) CD demonstrates a lack of structural interactions between BAK and BH3 peptides in the presence or absence of detergent. The spectra of the mixtures show no gain in structure compared to the averages of the components, indicating an absence of coupled folding and binding. Note the differences in BAK spectra in the presence and absence of detergent due to the oligomerisation. All experiments were performed at 5 μM protein concentration, and PS20 was present at 20×CMC. TAMRA-lablled versions of the peptides were used. SEC was performed on Superdex 200 10/300.

The interaction of PUMA with MCL-1 in detergent—and the lack thereof with BAK—was confirmed by native MS (Fig. 8.4). The hetero-dimer of MCL-1 and PUMA was observed in the gas-phase, demonstrating the stability of this complex. In contrast, no hetero-dimer between BAK and PUMA was observed, consistent with the results obtained by SEC and CD. The spectrum showed a range of oligomeric species that was qualitatively comparable to the spectrum recorded in the absence of PUMA. This is consistent with the notion that BH3 peptides are spectator of the spontaneous oligomerisaiton of BAK in buffer.

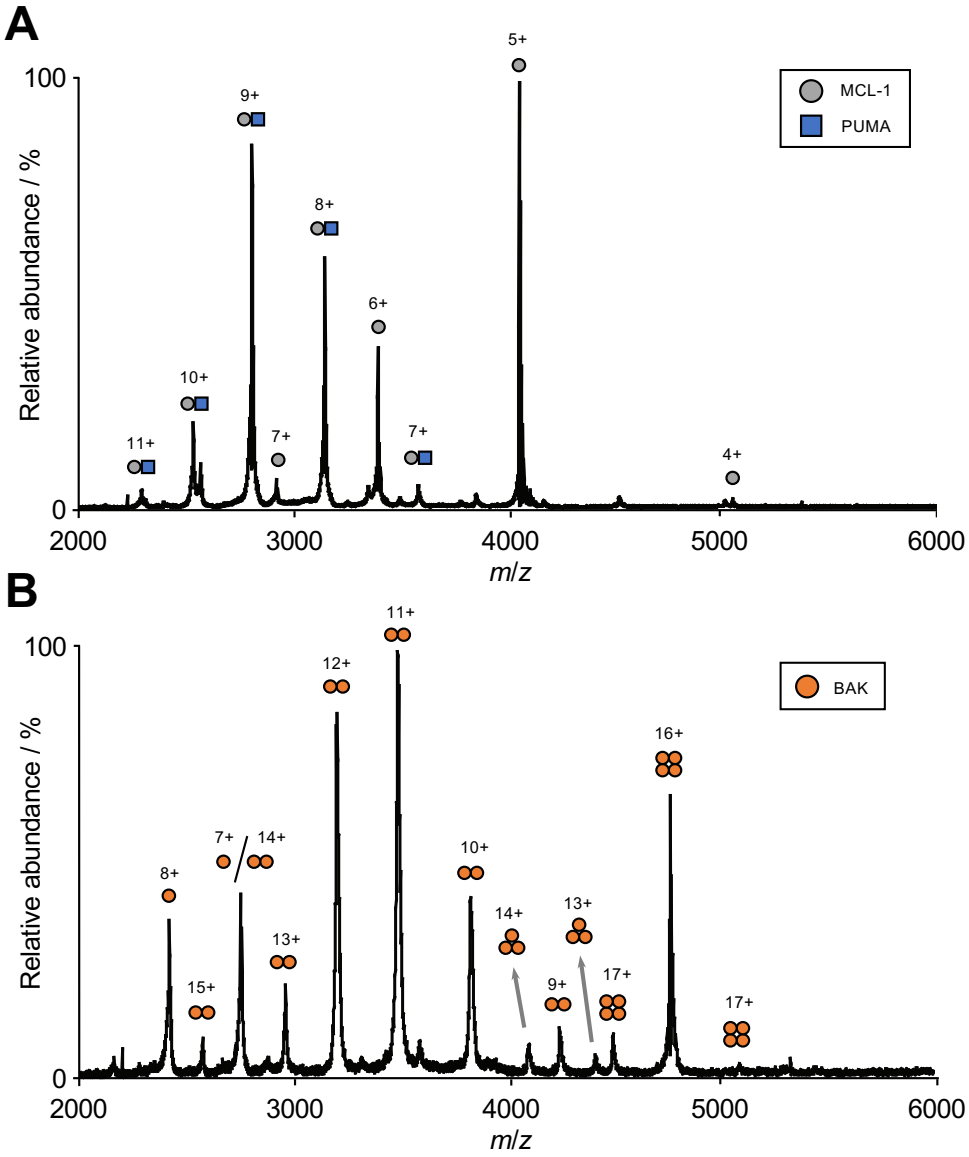

Figure 8.4 – Native MS of PUMA with either BAK or MCL-1 in the presence of detergent. (**A**) Formation of hetero-dimers of PUMA and MCL-1 in the presence of PS20 is consistent with the expected assembly state of these proteins, and shows that the detergent does not affect their interaction. (**B**) Incubation of BAK and PUMA in the presence of detergent only reveals homo-mers of BAK, thus demonstrating the lack of interaction between PUMA and either state of BAK. This spectrum is qualitatively comparable to Fig. 7.9B, hence confirming that PUMA does not affect the oligomerisation of BAK. Free PUMA could not be detected under the conditions of these experiments. Each protein concentration was 2.5 μM, and PS20 was present at 5×CMC.

8.3.3 MCL-1 hetero-dimerises with BAK and BAX, suppressing homo-oligomerisation

In stark contrast to the innocuous effect of PS20 on the binary interactions involving BH3 peptides, the presence of detergent completely altered the interaction profile of BAK/BAX with MCL-1. Indeed, these proteins did not interact at all in buffer, but formed hetero-dimers when PS20 was present

SEC analysis of the mixture of BAK and MCL-1 in the presence of PS20 showed the formation of a higher-order oligomeric species eluting at ∼11–12 mL (Fig. 8.5A). The peak was not symmetrical, and showed a shoulder with an elution volume matching that of BAK homo-oligomers (solid orange line). SDS-PAGE analysis of the fractions confirmed that the peak was composed of at least two species; eluting earlier was the complex between BAK and MCL-1, while the shoulder peak corresponded to homo-oligomers of BAK. Interestingly, the band intensities suggest a 1:1 complex, although this result has to be interpreted with care.

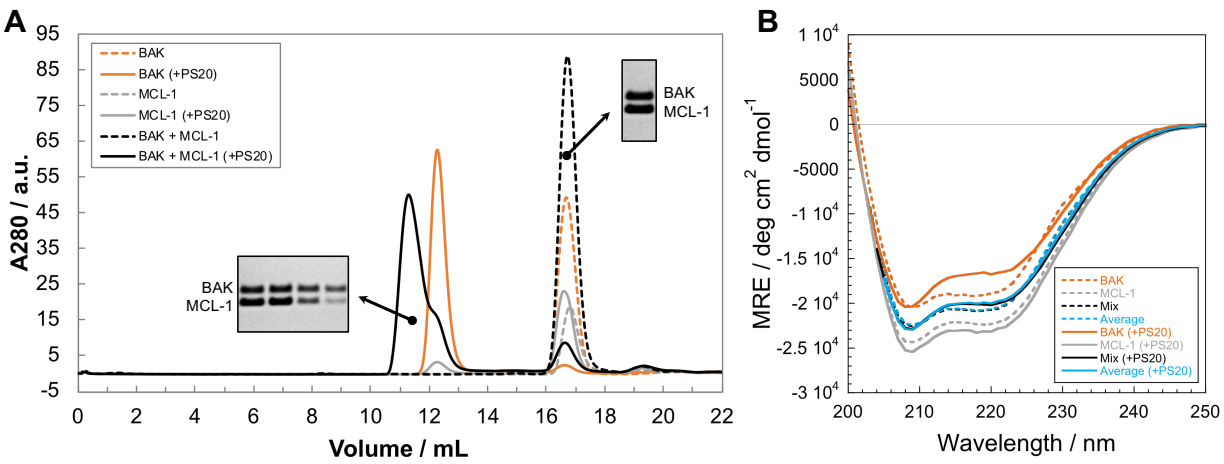

Figure 8.5 – BAK interacts with MCL-1 in detegent without loss of structure. (**A**) SEC analysis of BAK and MCL-1 (10 μM each) incubated in the presence of PS20 (20×CMC) reveals the formation of a larger assembly containing both proteins (solid black line). This species competes with the formation of the homo-oligomer of BAK (*cf.* shoulder peak). The two proteins do not interact in the absence of detergent (dashed black line). (**B**) CD analysis shows no structural change upon formation of the BAK:MCL-1 complex in PS20. The spectrum of the mixture (solid black line) is identical to the average of its parts (solid blue line). Spectra of BAK, MCL-1 and their mixture in the absence of detergent (dashed lines) are added for reference. Note the structural change of BAK upon oligomerisation. MCL-1 also shows a slight loss of structure in the prescence of detergent. Protein concentrations were 5 μM, and PS20 present at 20×CMC. His-tagged MCL-1 was used for the CD experiments.

Importantly, the formation of this hetero-meric complex was contingent on the presence

of detergent. SEC analysis of the mixture in the absence of detergent (dashed black line) showed an elution profile consistent with monomeric species. The peak contained both proteins, confirming their lack of interaction in buffer. It is noted that this result does not prove the interaction in itself, and is equally consistent with for example MCL-1 eluting earlier due to homo-oligomerisation, or interaction with the micelles. However, since MCL-1 on its own remained monomeric in the presence of detergent, its partitioning into the oligomer peak confirms its interaction with BAK.

The lack of hetero-merisation between BAK and MCL-1 in buffer was attributed to the buried nature of the BH3 motif—the probable site of interaction between these proteins. Therefore, the formation of a complex in the presence of detergent raises questions about the structural aspects of this hetero-merisation. For example, does the assembly of BAK and MCL-1 involve a 'coupled *unfolding* and binding' event? These questions were investigated by CD (Fig. 8.5B). The results clearly demonstrate that the complex remains folded. Moreover, the interaction between these two proteins in detergent is not associated with any structural changes. Indeed, the spectrum of the mixture (solid black line) overlays with the spectral average of its constituents. This observation was identical when the proteins were in buffer (dashed lines). However, it is noted that the structural effects of PS20 on BAK (compare the solid and dashed lines) might 'prime' it for interaction. In this scenario, the interaction itself does not induce any further structural changes, and the binding interface is already available prior to complex formation. This would be consistent with the BH3-in-groove model—dimerisation necessitates the exposure of the BH3 motif, hence it would become available for either homo- or hetero-oligomerisation.

The SEC profile of the hetero-mer suggests a large species. However, as previously noted, elution volumes in the presence of detergent are not representative. In order to characterise the stoichiometry of this detergent-induced hetero-meric complex, native MS was employed (Fig. 8.6). The mixture of BAX and MCL-1 was also analysed (Fig. 8.7). In both figures, the spectra in the absence of MCL-1 were reproduced for comparison.

Native MS spectra of detergent-treated BAK/BAX in the presence or absence of MCL-1 showed striking differences. While homo-oligomers as large as hexamers were observed for BAK when incubated on its own (Fig. 8.6A), addition of MCL-1 almost completely suppressed homo-oligomerisation (Fig. 8.6B). Only traces of homo-dimers could be observed, whilst most of BAK formed hetero-dimers with MCL-1. These results clearly demonstrate that: *i*) MCL-1 can interact with BAK in the presence of detergent to form hetero-dimers,

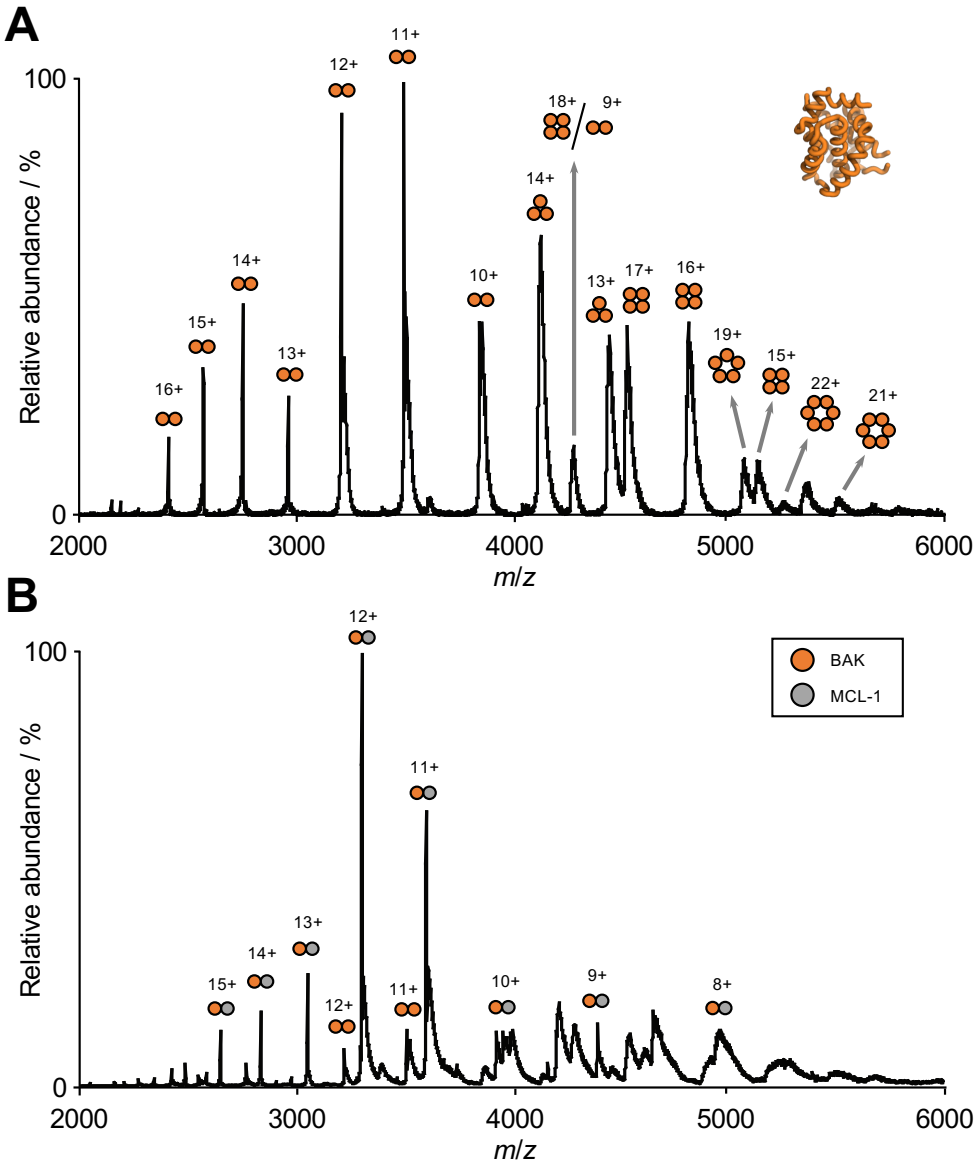

Figure 8.6 – MCL-1 suppresses the oligomerisation of BAK by forming hetero-dimers in detergent. (**A**) Reproduction from Fig. 7.9A, showing that BAK forms higher order oligomeric species (up to hexamers) in the presence of PS20. (**B**) When BAK is co-incubated with MCL-1 in the presence of detergent, homo-oligomerisation is suppressed in favour of BAK:MCL-1 hetero-dimers. Total protein concentrations were 5 μM (2.5 μM each for the mixture), and PS20 was present at 5×CMC.

and *ii*) the resulting complex prevents the homo-oligomerisation of BAK. Results for BAX were qualitatively similar. Larger oligomers disappeared in favour of BAX:MCL-1 hetero-dimeric species, although a significant proportion of dimers was still observed. However, the overall conclusion was similar to the effect of MCL-1 on BAK; hetero-dimerisation reduced the extent of homo-oligomerisation.

While the site of interaction between BAK/BAX and MCL-1 in these hetero-dimers is unknown, it is tempting to speculate that they occur *via* their BH3 motifs. Interestingly,

the affinities of MCL-1 for BAX_{BH3} and BAK_{BH3} are 71 and 0.077 nM respectively. Thus, the reduced effect of MCL-1 on BAX homo-oligomerisation (compared to BAK) might be a reflection of the 100-fold difference in binding affinites of MCL-1 for their respective BH3 motifs.

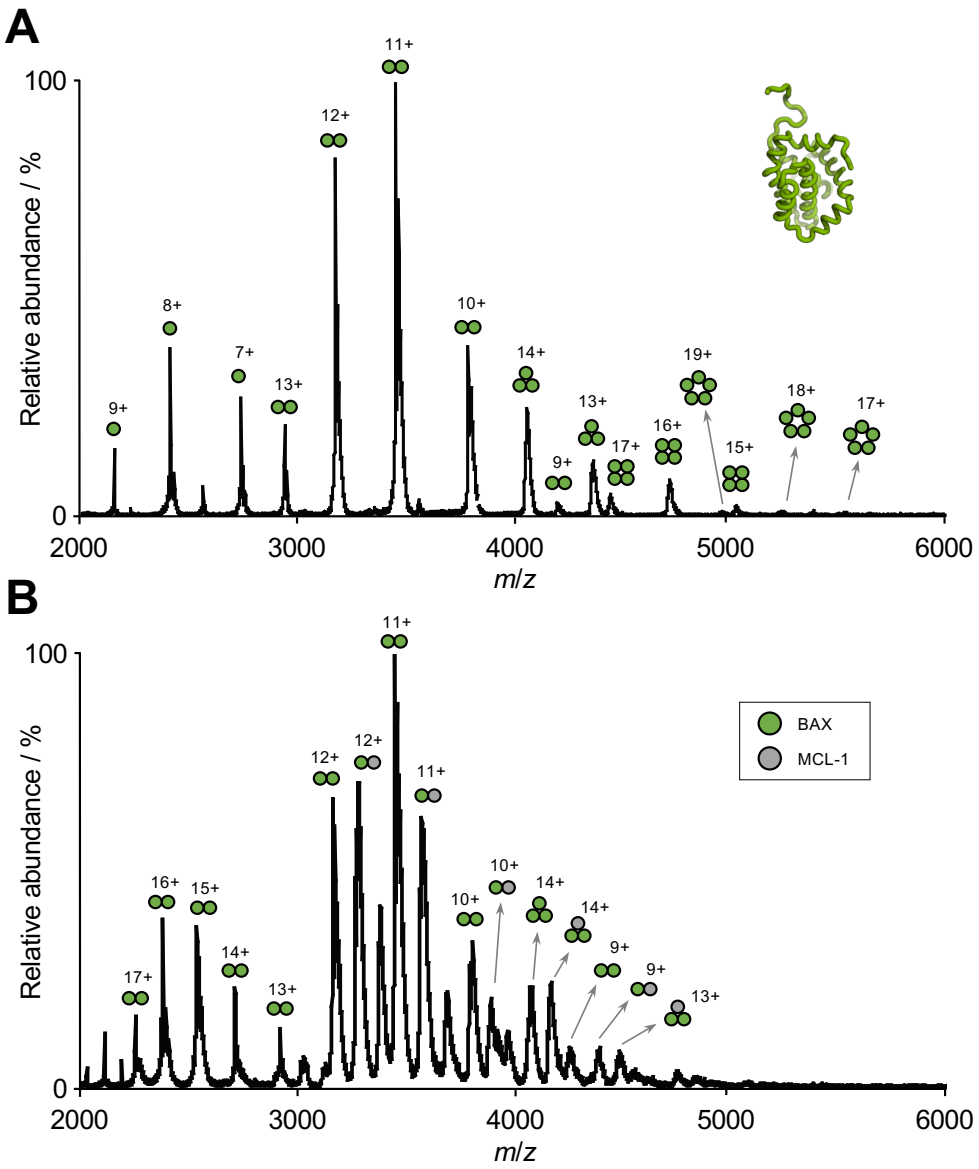

Figure 8.7 – MCL-1 reduces the oligomerisation of BAX by forming hetero-dimers in detergent. (**A**) Reproduction from Fig. 7.9B, showing that BAX forms higher order oligomeric species (up to pentamers) in the presence of PS20. (**B**) When BAX is co-incubated with MCL-1 in the presence of detergent, homo-oligomerisation is reduced in favour of BAX:MCL-1 hetero-dimers. These results are qualitatively comparable to the ones of BAK (Fig. 8.6). Total protein concentrations were 5 µM (2.5 µM each for the mixture), and PS20 was present at 5×CMC.

Together, these results conclusively demonstrate that MCL-1 is capable of preventing the homo-oligomerisation of BAK/BAX by binding to them. Moreover, the presence of hetero-

dimers—intstead of larger hetero-meric species—suggests that MCL-1 binding competes with the formation of the dimer 'denominator' of BAK/BAX. Thus, by preventing the nucleation step, it acts at the core of the homo-oligomerisation process. This result supports the notion of distinct oligomerisation interfaces (*cf.* Chapter 7).

Interestingly, the presence of the shoulder peak in SEC (Fig. 8.5A) indicates that, while the formation of BAK:MCL-1 species competes with the homo-oligomerisaiton of BAK, it does not completely abrogate it. The affinity of MCL-1 for the BH3 motif of BAK ($K_{\mathrm{d}} = 0.077$ nM) might suggest a more complete hetero-dimerisation at the concentrations of this experiment (10 μM of each protein). However, the presence of a competing homo-oligomerisation reaction precludes this simple interpretation. Thus, a more complete description was necessary in order to gain insights into the affinity of the hetero-dimer. Luckily, SEC showed the presence of suitably resolved peaks for each species, which allowed an estimation of the binding energetics between full-length BAK and MCL-1.

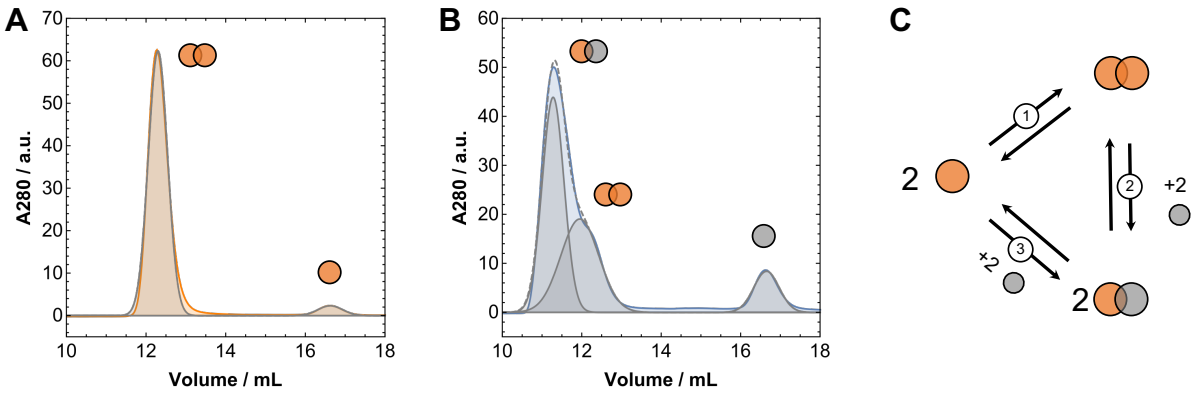

Figure 8.8 – Estimation of the affinity of the BAK:MCL-1 hetero-dimer. SEC chromatograms from Fig. 8.5A were fitted to linear combinations of Gaussian functions (dashed grey lines) to estimate the fractional distribution of each species. Individual Gaussian functions are reproduced for visualisation. (**A**) Elution profile of detergent-treated BAK fitted to a double Gaussian function. (**B**) Elution profile of the detergent-treated mixture of BAK and MCL-1 fitted to a triple Gaussian function. (**C**) Reaction scheme representing the competition between homo- and hetero-oligomeristion. The case of homo-dimerisation is illustrated, although the same relationships apply for any oligomeric state of BAK. Orange and grey circles represent BAK and MCL-1 monomers respectively.

The competing homo- and hetero-oligomerisation reactions can be expressed in the form of a thermodynamic triangle (Fig. 8.8C). Here the diagram represents the competition between homo- and hetero-dimerisation, but the same scheme applies for any homo-oligomerisation of BAK. From this reaction scheme, the following relationships can be

expressed (numbering as in Fig. 8.8C):

$$\begin{cases} \Delta G_1 = -RT \ln\left(\frac{[(\text{BAK})_n]}{[\text{BAK}]^n}\right) \\ \Delta G_2 = -RT \ln\left(\frac{[(\text{BAK:MCL-1})]^n}{[(\text{BAK})_n][\text{MCL-1}]^n}\right) \\ \Delta G_3 = -RT \ln\left(\frac{[(\text{BAK:MCL-1})]^n}{[\text{BAK}]^n[\text{MCL-1}]^n}\right) \end{cases} \tag{8.1}$$

where n represents the stoichiometry of the homo-oligomer of BAK. Furthermore, because of the triangular relationship between these three states:

$$\Delta G_3 = \Delta G_1 + \Delta G_2 \tag{8.2}$$

Estimations of the concentrations of each species were obtained from the SEC chromatograms of the reaction performed in PS20 (8.5A). Fitting to Gaussian functions allowed to obtain fractional distribution between states. This strategy assumed the presence of only three distinct species: monomers, hetero-dimers, and homo-oligomers (shown in Fig. 8.8 for the dimeric case). By assuming that the elution profiles are representative of the solution-state distributions—*i.e.* absence of significant exchange between states on the time-scale of the elution, which appears realistic given the absence of severely broaden peaks—the concentration of each species could be estimated from the fractional distributions. For the hetero-dimer, weighing by extinction coefficients was applied. These concentration values were used to calculate the thermodynamic parameters expressed in Equation 8.1.

Analysing the results obtained for the oligomerisation of BAK on its own (Fig. 8.8A) allowed ΔG_1 to be estimated. Similarly, analysis of the mixture (Fig. 8.8B) allowed ΔG_2 to be obtained. Direct estimation of ΔG_3 was not possible. Indeed, this reaction cannot be performed on its own, as the presence of detergent also allows the competing homo-oligomerisation reaction to proceed. Thus, the energetics of the hetero-dimerisation reaction was obtained indirectly from ΔG_1 and ΔG_2 (Equation 8.2). Note that this value is for the conversion of n moles of BAK with n moles of MCL-1 to generate n moles of hetero-dimers. This value is not very intuitive. Instead, the energetic per hetero-dimerisation event can by obtained by $\Delta G_3/n$. This quantity is more useful as it directly relate to the K_d of the hetero-dimer.

This system of equations assumes that BAK only populate a single oligomeric state. This assumption is clearly incorrect given the known heterogeneity observed by native MS.

Therefore, it was necessary to assess the dependence of the energetics on the assumed assembly state. This was done by calculating ΔG's for different scenarios of 'pure' homo-oligomerisation (Table 8.1).

Table 8.1 – Thermodynamic parameters estimated from Fig. 8.5 using Equations 8.1 and 8.2. Calculations were performed for the homo-dimeric ($n = 2$), tetrameric ($n = 4$), and hexameric ($n = 6$) cases. The hetero-dimeric dissociation constants associated with each presumed homo-oligomeric state were calculated using $K_\mathrm{d} = \exp\left(\frac{\Delta G_3}{nRT}\right)$.

	$n = 2$	$n = 4$	$n = 6$
ΔG_1 / kcal mol^{-1}	-10.3	-27.5	-44.9
ΔG_2 / kcal mol^{-1}	-9.0	-10.5	-11.8
ΔG_3 / kcal mol^{-1}	-19.4	-38.0	-56.7
$\frac{1}{n}\Delta G_3$ / kcal mol^{-1}	-9.7	-9.5	-9.4
K_d (BAK:MCL-1) / nM	79	110	131

Since the oligomeric coefficient is present on both side of reaction 2, the estimation of its energetics is not very dependent on the choice of n. In contrast, the estimation of ΔG_1 is strongly affected. This is a consequence of the stoichiometric asymmetry of the reaction, which becomes more pronounced as n increases. Because ΔG_3 is a linear combination of the energetics of reactions 1 and 2, its estimation varies significantly as well. This is consistent with the asymmetric stoichiometry of reaction 3. However, since the energetics of hetero-dimerisation is obtained after normalisation by n, its value is relatively independent of the presumed oligomeric state. Thus, while simplistic, this thermodynamic analysis provides a relatively robust estimate of the energetics of hetero-dimerisation between BAK and MCL-1. It is noted that this set of values are only valid for the specific conditions of the experiments. In particular, the concentration of detergent is expected to have an important impact on the position of these equilibria.

It is interesting to compare the estimated K_d values for the hetero-dimerisation of full-length BAK with MCL-1 ($K_\mathrm{d} \approx 100$ nM) with that of $\mathrm{BAK_{BH3}}$ ($K_\mathrm{d} = 0.077$ nM). This difference in affinity suggests that if the interaction with MCL-1 does occur *via* the BH3 motif of BAK, a significant amount of the binding energy is lost ($\Delta\Delta G \sim 4$ kcal mol^{-1}). This could be due to an unfavourable conformational re-arrangement in BAK, or an incomplete interaction of the BH3 motif due to structural constraints.

This lowered affinity might have important regulatory consequences. Indeed, the average

interface energetics for the homo-oligomer of BAK is about 5–8 kcal mol^{-1} (depending on the presumed oligomeric state used for the calculation). This value is not far from the ΔG for hetero-dimerisation (∼9.5 kcal mol^{-1}). Hence, a single mutation may shift the distribution of states sufficiently to abrogate the effect of MCL-1 on the homo-oligomerisation of BAK. This effect is more dramatic due to the presence of competing equilibria, and the compounding effect resulting from the oligomeric nature of these assemblies.

8.3.4 BH3 motifs displace BAK/BAX MCL-1 hetero-dimers

The results presented in the previous section provide the link between the biophysical properties of BAK/BAX, and the anti-apoptotic role of MCL-1; large-scale homo-oligomerisation of effector BCL-2 proteins may be 'capped' through hetero-dimerisation with their anti-apoptotic counterparts, thus preventing pore-formation. This line of evidence strongly support the indirect activation model. In order to confirm this behaviour, the role of BH3-only proteins was investigated. This was achieved by studying the minimal tripartite system in detergent with all its components. As for the analysis of the binary interactions, a combination of SEC, CD and native MS was employed.

As a first step to understand the effect of PUMA on the reaction between BAK and MCL-1, SEC analysis of the ternary mixture in detergent was performed (Fig. 8.9A). All three components were pre-mixed in buffer, followed by the addition of PS20. The results unequivocally showed a reversal of the outcome observed in the absence of PUMA. The oligomer peak shifted back to match that of the homo-oligomer of BAK in the absence of any other protein (orange line). SDS-PAGE confirmed that this peak only contained BAK. In stark contrast to the elution profile of BAK with MCL-1—where MCL-1 completely partitioned into the 'oligomer' peak at ∼11–12 mL (*cf.* Fig. 8.5A)—addition of PUMA resulted in MCL-1 eluting as a 'monomer' (∼16.5 mL). This was in fact not the monomer, but the complex between MCL-1 and PUMA. The interaction between PUMA and MCL-1 was confirmed by: *i*) the absence of a free peptide peak (∼18.5 mL), *ii*) SDS-PAGE analysis of the fractions, *iii*) absorbtion profiles consistent with the presence of the TAMRA-labelled peptide. UV-vis analysis also confirmed the absence of PUMA in the 'oligomer' fractions, consistent with previous experiments (*cf.* Section 8.3.2). These results conclusively demonstrate that PUMA binds MCL-1, preventing its hetero-oligomerisation with BAK. As a consequence of this 'inactivation', BAK is left to homo-oligomerise. From

the SEC profiles, the outcome appears identical to the result obtained with just BAK in detergent (orange line).

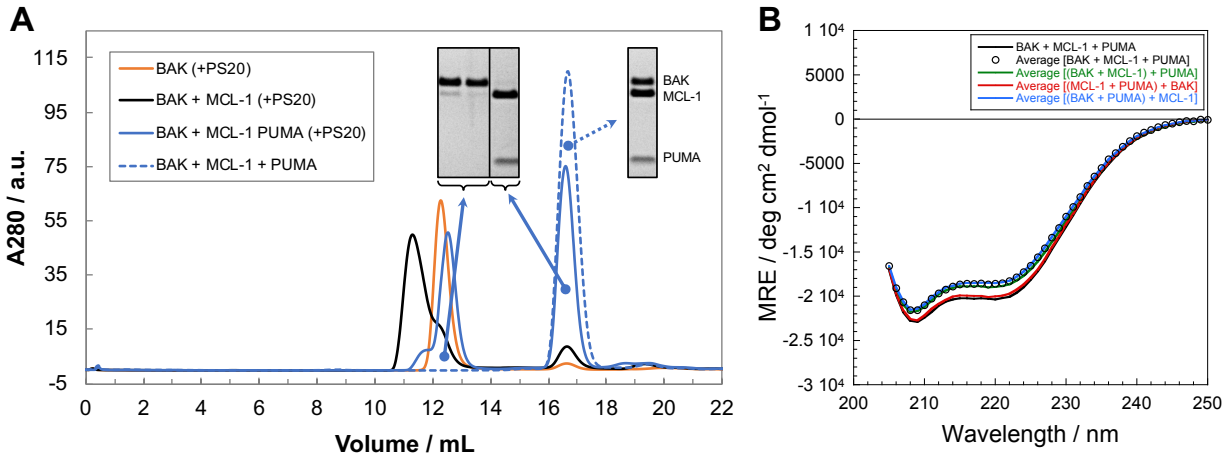

Figure 8.9 – PUMA binding to MCL-1 prevents its interaction with BAK. (**A**) SEC analysis of the ternary mixture (BAK, MCL-1, PUMA) in PS20 shows the existence of two species: BAK homo-oligomers, and PUMA:MCL-1 hetero-dimers. Binding of PUMA to MCL-1 prevents its hetero-oligomerisation with BAK, resulting in BAK homo-oligomerisation as when it is on its own (orange line). The hetero-oligomer of BAK and MCL-1 (black line) is reproduced here for comparaison. This partioning is specific to the presence of detergent, indicated by the absence of larger species in buffer (dashed blue line). Experiments were performed at 10 μM of each protein in the presence of PS20 (20×CMC) using a Superdex 200 10/300 colum. (**B**) CD analysis demonstrates that PUMA binds MCL-1 in the presence of both BAK, and detergent. The ternary mixture (black line) is more structured that the average of its individual parts (empty circles). Only the spectral average of PUMA:MCL-1 and BAK (red line) recapitulates this gain in structure, hence confirming the coupled folding and binding of PUMA to MCL-1. Experiments were performed at 5 μM of each protein in the presence of PS20 (20×CMC). His-tagged MCL-1 was used for the CD experiments.

When BAK and MCL-1 were incubated in the presence of detergent, the outcome was not the sum of its parts; a consequence of their interaction. The SEC results of the ternary mixture also showed an outcome different to the sum of its individual components. The result was consistent with the sum of the outcomes of BAK on its own and the mixture of PUMA and MCL-1. This specific additivity was confirmed by CD (Fig. 8.9B). The spectrum of the ternary mixture did not overlay with the average of its individual components, indicating a coupled folding and binding event. The fact that this interaction was between PUMA and MCL-1 was confirmed by the fact that only the spectral averages of the PUMA:MCL-1 complex with BAK oligomers recapitulated the change in structure. None of the other binary combination recapitulated the observable result. These results confirm the observations made by SEC

To confirm these SEC and CD results, native MS was employed to study the ternary mixtures of BAK/BAX, MCL-1 and PUMA in the presence of PS20. This also offered a more detailed picture of the outcome of the reactions. In particular, the stoichiometry of the different oligomeric states could be obtained. Fig. 8.10 shows the results for the mixture of MCL-1, PUMA and either BAK or BAX. Both reactions had qualitatively similar outcomes. Higher-order homo-oligomers were observed; in stark contrast to the results obtained in the absence of PUMA (Fig. 8.6 and 8.7). The presence of the BH3 peptide completely suppressed the formation of BAK/X:MCL-1 hetero-dimers. Consistent with the results from SEC and CD, MCL-1 formed hetero-dimers with PUMA. These native MS results are in effect the combination of the homo-oligomer spectra (Fig. 8.6A and 8.7A) and the PUMA:MCL-1 spectrum (Fig. 8.4A)—in line with the additivity argument formulated for SEC and CD. Therefore, native MS confirmed that PUMA scavenges MCL-1, preventing its hetero-dimerisation with either BAK or BAX. As a consequence, these proteins proceed to form homo-oligomers, and the outcomes become identical to the results obtained when neither PUMA, nor MCL-1 are present.

Experiments with BID instead of PUMA were also performed (data not shown). These spectra were qualitatively identical to the results obtained with PUMA, although the BID:MCL-1 complex could not be detected in the gas-phase. These experiments confirmed the common function of BH3-only proteins, and demonstrated that these outcomes were not dependent on the choice of the system. Importantly, these results obtained using BH3 peptides prove that the BH3 groove of MCL-1 forms part of the hetero-dimerisation interface between BAK/BAX and MCL-1. Although it does not conclusively demonstrate that the BH3 motifs of BAK/BAX constitute the reciprocal interfaces, it is tempting to speculate that they do.

It is noted that all the experiments described in this section were performed by pre-incubating the components in buffer, followed by addition of detergent. Since PUMA/BID and MCL-1 already interact in buffer, it raises the question whether the observed outcomes are a consequence of MCL-1 being unavailable to hetero-dimerise in the first place. This hypothesis was tested by performing the reactions in a sequential fashion; BAK was first oligomerised in detergent, followed by addition of MCL-1, and then PUMA was added. Between each step, sufficient time was allowed for the systems to reach equilibrium (multiple hours). It was found that the outcomes were independent of the sequence of events. Hence, the processes are fully reversible, and the results described here are representative

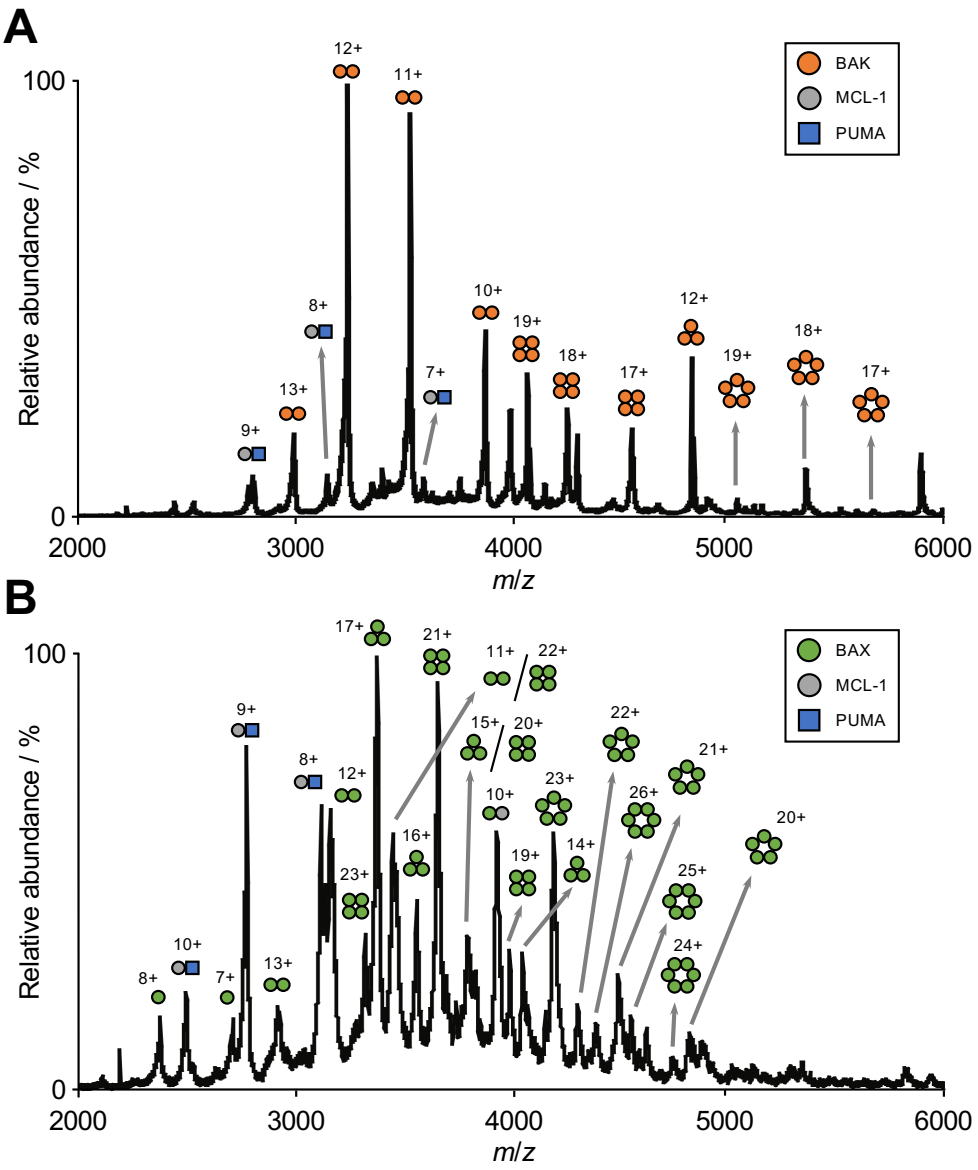

Figure 8.10 – PUMA displaces hetero-dimers between MCL-1 and BAK/BAX, leading to their homo-oligomerisation. The ternary mixtures (BAK or BAX, MCL-1, and PUMA) in PS20 show the formation of PUMA:MCL-1 hetero-dimers, and larger homomers of BAK/BAX. Thus, compared to the results in the absence of PUMA (Fig. 8.6 and 8.7), it demonstrates that PUMA leads to the homo-oligomerisation of BAK/BAX by binding MCL-1. This interaction prevents MCL-1 from hetero-dimerising with BAK/BAX, which normally caps the homo-oligomerisation. (**A**) Ternary mixture of BAK, MCL-1 and PUMA. (**B**) Ternary mixture of BAX, MCL-1 and PUMA. Protein concentrations were 2.5 µM for BAK/BAX and MCL1, and 5 µM for PUMA. PS20 was present at 5×CMC.

of the outcome that would be obtained from any starting point.

8.3.5 Kinetic investigations of BCL-2 regulation

Results presented in the previous sections showed that in detergent, BAK and BAX form hetero-dimers with MCL-1, and that these complexes are displaced by PUMA or BID. Furthermore, experiments performed by changing the sequence of events demonstrated that all the reactions were reversible. Therefore, the distribution of states obtained for the different systems (binary, ternary) represent thermodynamic equilibria of these mixtures. In this section, kinetics was employed to gain insights into the mechanism of some of these steps.

The effect of PUMA (or BID) on the binary mixtures of BAK/X and MCL-1 was striking, leading to an almost complete suppression of BAK/BAX:MCL-1 hetero-dimerisations. This process was shown to be reversible, and was not contingent on the pre-formation of the PUMA:MCL-1 complex. Thus, it is clear that BH3 peptides are capable of displacing these hetero-dimers. The mechanism of this coupled 'BH3-binding and BAK-dissociating' event was investigated. The interactions between peptides and MCL-1 being fast, similarly rapid reactions would be expected if the BH3 binding groove in BAK:MCL-1 was free. In contrast, if the interaction site was occupied, the rates would be expected to be much slower. These different scenarios were tested by pre-forming BAK:MCL-1 hetero-dimers in detergent, followed by addition of dye-labelled BH3 peptides. The fluorescence polarisation of the dye was used as probe for the binding reaction, thus excluding signal changes associated with other events. The reactions with t-PUMA$_{BH3}$ and t-BAK$_{BH3}$ were investigated (Fig. 8.11). Both reactions fitted a single exponential function. Moreover, the resulting rate constants were within 2-fold of the dissociation rate constant between BAK$_{BH3}$ and MCL-1 (*cf.* figure legend for details). This similarity—together with the apparent first-order kinetics of these reactions—strongly support the notion that the binding of BH3 peptides to the hetero-dimer is rate-limited by the dissociation of BAK from MCL-1. It also precludes the existence of an 'active' displacement mechanism. Moreover, it provides further evidence for the theory that the BH3 groove of MCL-1 forms part of the BAK:MCL-1 hetero-dimeric interface. The fact that BAK can be displaced by a peptide of its own BH3 motif strongly suggests that it represents the reciprocal interface of the hetero-dimer. However, this result alone does not constitute a formal proof of this hypothesis.

The study of BAK and BAX oligomerisation in detergent (*cf.* Chapter 7) revealed the

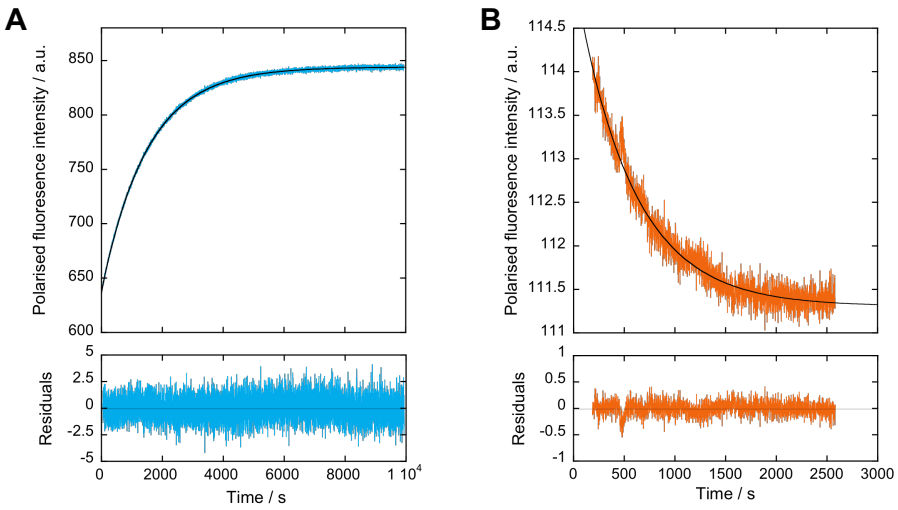

Figure 8.11 – Binding of BH3 peptides to BAK:MCL-1 hetero-dimers is rate-limited by the dissociation of the complex. BAK and MCL-1 (His-tagged) were pre-incubated in the presence of PS20 (20×CMC), followed by addition of dye-labelled BH3 peptides (containing PS20). The reactions were followed by monitoring the change of TAMRA fluoresence polarisation (V/V). Lower panels show the residuals to a single exponential fit (black lines). (**A**) Binding of t-PUMA$_{BH3}$ (reaction performed with 2.5 μM final protein concentrations, k_{obs} = 6.7 10^{-4} s^{-1}). (**B**) Binding of t-BAK$_{BH3}$ (reaction performed with 5 μM final protein concentrations, k_{obs} = 1.8 10^{-3} s^{-1}). These observed rate constants are similar to the dissociation rate constant of t-BAK$_{BH3}$ from MCL-1 in buffer (k_{off} = 5.6 10^{-4} s^{-1}), suggesting a dissociation-limited binding event.

presence of two slow kinetic phases. Both were dependent on the concentration of proteins and detergent, although the latter had the strongest impact on the rate of transformation. In this chapter, it was demonstrated that BAK forms hetero-dimers with MCL-1 in the presence of detergent. The mechanism of this interaction was investigated by performing kinetic measurements (Fig. 8.12). BAK and MCL-1 were pre-incubated in buffer, and the reaction was initiated by the addition of PS20. The hetero-dimerisation was followed by measuring the change in intrinsic fluorescence. The reaction was slow, and remained consistent with a double exponential decay. Interestingly, when compared to the identical reaction performed in the absence of MCL-1, both rate constants remained within 2-fold. Therefore, the formation of BAK:MCL-1 appears to be rate-limited by similar events to the homo-oligomerisation reaction. Together with the kinetic evidences gathered in Chapter 7, this result suggests that the rate of hetero-dimerisation between BAK and MCL-1 is mostly limited by conformational re-arrangements within BAK, *e.g.* exposure of its BH3 motif.

Addition of detergent to the ternary mixture was shown to result in the formation of homo-oligomers of BAK, and PUMA:MCL-1 complexes (Fig. 8.10A). Interestingly, the kinetic

analysis of this process revealed rate constants virtually identical to the reaction performed with BAK alone. In terms of BAK oligomerisation, the interaction between PUMA and MCL-1 in the ternary mixture lead to the same equilibrium outcome as when these proteins were absent. This kinetic result proves that the mechanism is also unaffected. Thus, the neutralisation of MCL-1 by BH3 peptides truly makes this complex a spectator of the homo-oligomerisation of BAK.

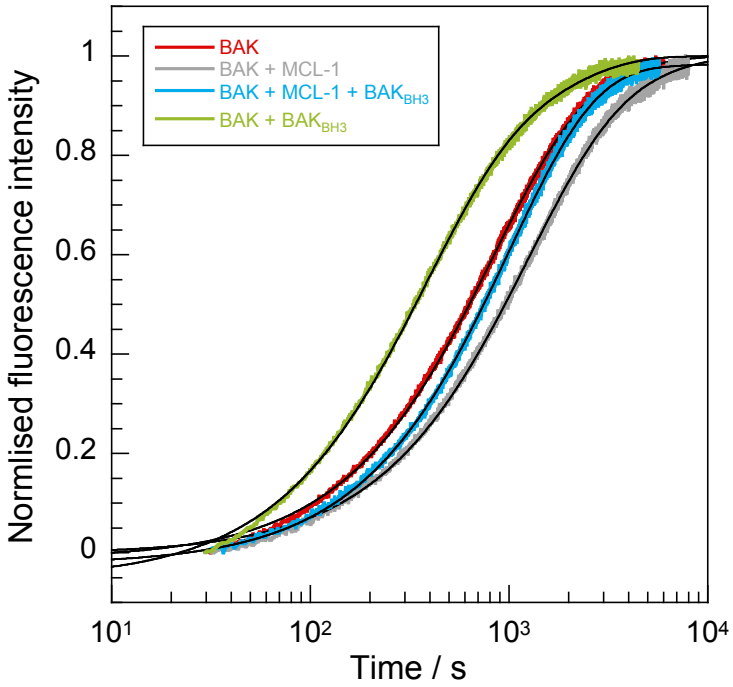

Figure 8.12 – Oligomerisation of BAK in the presence of other proteins. Components were pre-assembled in buffer (5 μM final concentrations), and the reaction initated by the addition of PS20 (20×CMC). Reactions were followed by monitoring the change in intrinsic fluorescence, and the data normalised to aid visualisation. Black lines represent fits to a double exponential function (with F_{final} set to 1). Results from the fits are reported in Table 8.2. His-tagged MCL-1 was used for these experiments.

Table 8.2 – Values of the fits from Fig. 8.12. All experiments were initated by the addition of PS20 (20×CMC). Errors are fitting errors.

BAK	+	+	+	+
MCL-1	–	+	+	–
BH3	–	–	+	+
$k_1 \times 10^{-3}$ / s^{-1}	1.300(±)0.005	0.876(±)0.004	0.992(±)0.002	2.92(±)0.01
$k_2 \times 10^{-4}$ / s^{-1}	4.25(±)0.06	2.88(±)0.04	0.3(±)0.1	7.26(±)0.05
f_{Amp1}	0.83	0.80	0.98	0.75

Surprisingly, adding BAK_{BH3} to BAK appeared to slightly accelerate the oligomerisation reaction. Due to the extremely low affinities of BAK for BH3 peptides (*cf.* Chapter 6), formation of complexes is not expected to occur to any significant extent under the concentrations of these kinetic experiments. BH3 peptides do form short-lived interactions with BAK, however binding would be expected to slow-down the oligomerisation by inhibiting it (*cf.* Chapter 7). Whilst the kinetic effect was very small, it is nevertheless puzzling. More experiments are warranted to clarify this result.

These preliminary experiments demonstrate that—similarly to the homo-oligomerisation—the hetero-dimerisation of BAK with MCL-1 is probably limited by events pertaining to structural changes in BAK. Moreover, the disruption of this complex by BH3 peptides appears to be entirely limited by the dissociation of BAK from MCL-1. Together, these findings paint a picture whereby most of the kinetic events within the BCL-2 network are determined by BAK (and BAX). Interestingly, the rate of BAK:MCL-1 dissociation initiated by the presence of BH3 peptides ($t_{1/2} \approx 20$ min) is not far from the time it takes for cytochrome c to be released from mitochondria in apoptotic cells (5 min, (Goldstein *et al.*, 2000)). Given the differences in conditions, especially temperature, it is tempting to speculate that the rate-limiting step to pore-formation is the dissociation of BAK/BAX from their anti-apoptotic counterparts.

8.4 Discussion

The results presented in this Chapter, together with those from Chapter 6 and 7, have lead to an understanding of the biophysical regulation of the BCL-2 network. Using a simplified tripartite system, and detergents as membrane-mimics, the properties of the this network were re-capitulated *in vitro*. Importantly, it allowed the mechanism of the regulation to be investigated on a molecular level, thus clarifying contradicting biochemical reports from the literature.

In Chapter 6, investigations of the tripartite system in buffer showed that BH3 peptides only interacted appreciably with MCL-1. BAK remained spectator of these interactions, and did not show any signs of oligomerisation; neither spontaneous, nor triggered. These results clearly demonstrated that, in the absence of a hydrophobic environment, the properties of this network could not explain the biological function of these proteins. In Chapter 7, membranes were mimicked by the use of detergents, and the oligomerisation of BAK and BAX re-created *in vitro*. This capability opened new avenues to study this process, and both mechanistic and structural investigations were undertaken. The oligomerisation was shown to be spontaneous in the presence of detergent, suggesting that BAK/BAX oligomers might the thermodynamically favoured states at the membrane. These findings have important biological and mechanistic implications, and point towards the indirect activation model for the regulation of MOMP.

The possibility to generate oligomers with detergent has allowed the tripartite system to be re-visited in a more relevant context. The results from these experiments were the focus of this Chapter, and the results are summarised in Fig. 8.13 (together with the results obtained in Chapters 6 and 7).

The presence of detergent completely altered the interaction profile between members of the tripartite system. While BAK/BAX and MCL-1 were incapable of interacting in buffer, they formed complexes in the presence of detergent. This is in line with a previous report (Liu & Gehring, 2010). More importantly, this hetero-dimerisation abrogated the homo-oligomerisation of BAK/BAX, thus providing a biophysical explanation for the anti-apoptotic character of MCL-1. Hetero-dimerisation appeared to act at the core of the pore-formation process by preventing the nucleation step—homo-dimer formation. Addition of PUMA (or BID) lead to the dissociation BAK/BAX:MCL-1, which resuscitated the homo-oligomerisation of BAK and BAX. Thus, BH3-only proteins initiate apoptosis

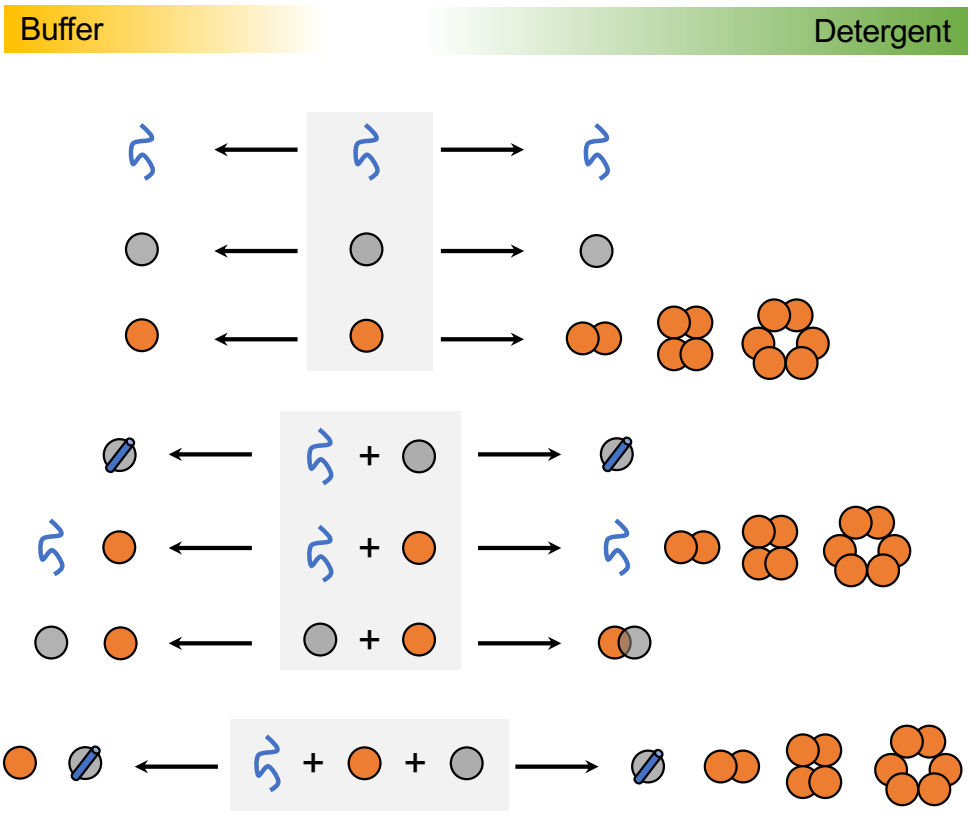

Figure 8.13 – Summary of tripartite states and interactions depending on the presence or absence of detergent. The diagram is divided into single-, dual-, and tri-component systems. The outcome in the absence of detergent (left-hand side) is compared to the result when detergent is present (right-hand side). These differences highlight the importance of the hydrophobic environment in mediating key interactions. These schemes summarises the results from Chapters 6, 7, and 8. Orange spheres denote BAK (and BAX), grey ones indicate MCL-1, and BH3 peptides are shown in blue. Detergent makes reference to PS20.

indirectly—by scavenging anti-apoptotic BCL-2 protein—not through direct triggering of the oligomerisation process itself. These results suggest that pore-formation is a collateral effect of the neutralisation of anti-apoptotic BCL-2 proteins; a consequence of the inherent thermodynamic instability of monomeric BAK and BAX in a membrane environment. A summary of this mechanism is schematically described in Fig. 8.14.

These findings—obtained using a minimal tripartite system and detergent—drew a clearer picture of the regulation of apoptosis by the BCL-2 network. The results presented in this thesis are fully consistent with the indirect activation model (Chen *et al.*, 2005, Willis *et al.*, 2005, 2007); pore-formation is spontaneous, and requires capping by anti-apoptotic members. BH3-only protein release BAK and BAX from their molecular chaperones, which leaves them free to oligomerise. This model is supported by a study which demonstrated that cells spontaneously undergo apoptosis if all BCL-2 proteins but BAK or BAX are

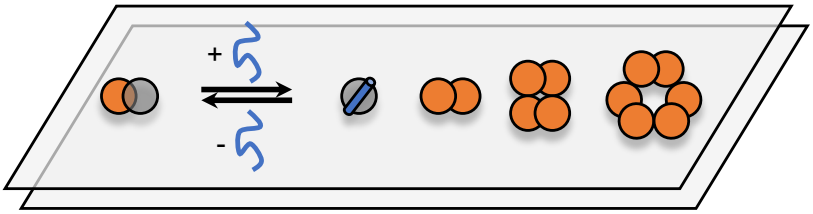

Figure 8.14 – Putative mechanism for the regulation of the BCL-2 interactome at the membrane. In non-apoptotic cells, BAK and BAX (orange spheres) form hetero-dimers with anti-apoptotic BCL-2 (grey spheres). This prevents their homo-oligomerisation. Competitive binding of BH3-only proteins (blue) to anti-apoptotic BCL-2 proteins displaces BAK and BAX, leaving them free to homo-oligomerise and form pores. Thus, the competition between homo-, and hetero-oligomerisation determines cellular fate at the the molecular level.

knocked-out (O'Neill *et al.*, 2016). It also echoes the results form a recent fluorescence cross-correlation spectroscopy study on the interactome of BAX, BCL-X_L and BID in unilamellar vesicles (Bleicken *et al.*, 2017). In conclusion, it appears that the process does not require to be triggered, but needs to be prevented instead. These results paint a picture where cellular survival is on the cusp of protein stability, and the underlying biophysical features of BCL-2 proteins are at the heart of the emerging properties of the network.

Chapter 9

Energy landscape of BCL-2 proteins

Chemical denaturation curves of BAX were performed by Tristan O. Kwan.

9.1 Introduction

The energy landscape theory of proteins underlies much of our modern understanding of folding (Bryngelson & Wolynes, 1987, Onuchic & Wolynes, 2004). An extension of the Anfinsen hypothesis, it describes proteins in terms of statistical mechanics. The sequence—and external conditions—determine a multi-dimensional energy surface which may be sampled by the protein. Hence, the energy landscape describes both the native state(s), as well as higher-energy intermediates. A full knowledge of the landscape would allow any protein to be characterised; not only thermodynamically, but also kinetically. Indeed, activation energies for transitions between states are also described by the landscape.

It is important to point out that the energy landscape of a given protein chain is not only encoded by the sequence itself, but by external factors as well. Changes in solution conditions (*e.g.* pH, or the presence of chemical denaturant) will re-model this energy surface, which may affect the distribution between states. The effect of urea and guanidinium chloride on the stability of the physiological states of proteins are well-known examples of this phenomenon. Moreover, if the chain can participate in multi-molecular reactions, changes in protein concentration will also impact the observed stabilities for the associated states, and thus their distributions. It is noted that under biological conditions the thermodynamically favoured distribution might not be reached (or only slowly) due to kinetic constraints (also described by the energy landscape) (Wang *et al.*, 1996). This is exemplified by the case of amyloid-forming proteins; their multimeric states are favoured over

their monomeric state, but the conversation is extremely slow. Therefore, the monomeric state of these proteins is metastable with respect to the amyloid under the concentrations present in the cell (Baldwin *et al.*, 2011).

BAK and BAX are known to oligomerise as part of their biological functions. The work presented in the previous three Chapters revealed that BAK and BAX formed oligomers in the presence of detergent, a result consistent with the membrane-like nature of these assemblies. However, the same proteins in buffer did not show appreciable oligomerisation over a large time-scale and temperature-range. Therefore, it is likely that the monomers are the thermodynamically favoured states of these proteins in the absence of a hydrophobic environment. Preliminary experiments aimed at characterising the thermodynamic effect of detergent point at the same conclusion.

9.2 Aims

The spontaneous formation of these oligomers in the presence of detergent raises interesting questions about the energy landscapes of BAK and BAX. Do these structures exists as high-energy states in buffer? This hypothesis was investigated in this Chapter using protein folding techniques. Chemical denaturation revealed the presence of a multimeric intermediate, and preliminary results suggest that this state is off-pathway to the unfolding of BAK. In Chapter 6, BAK and MCL-1 were shown not to interact in buffer. This was attributed to the buried nature of the BH3 motif of BAK, making it unavailable for interaction at the groove of MCL-1. However, in the presence of detergent hetero-dimerisation did occur (*cf.* Chapter 7). This interaction was re-visited using denaturant jumps in the absence of detergent. The formation of a complex between BAK and MCL-1 was generated in buffer by pre-unfolding BAK.

9.3 Chemical unfolding studies

All experiments presented in this Chapter were performed in the absence of detergent, except for Section 9.3.4.

9.3.1 BAK and BAX populate an intermediate

As a first step to characterise the energy landscape of BCL-2 proteins, equilibrium chemical unfolding curves were performed. Monomeric BAK, BAX, and MCL-1 were incubated in increasing amounts of denaturant (GdmCl for BAK and BAX, urea for MCL-1) and the results monitored by spectroscopic techniques. Intrinsic fluorescence and circular dichroism spectroscopy were employed in order to probe changes in both tertiary and secondary structure. (Fig. 9.1)

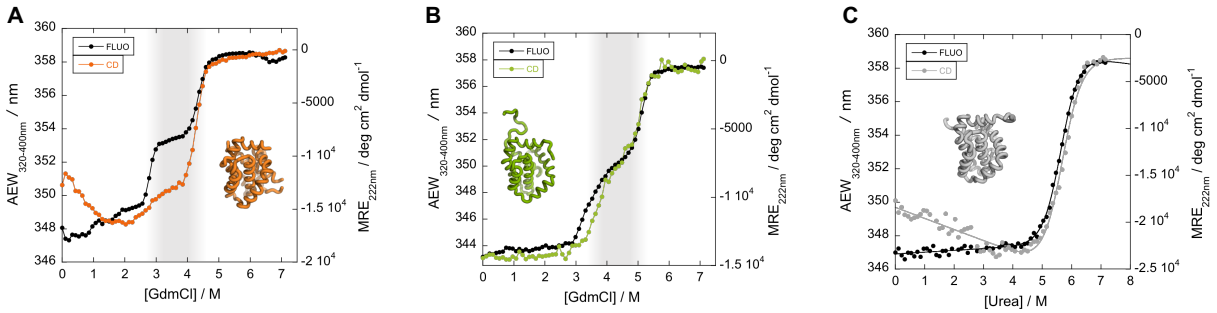

Figure 9.1 – Equilibration denaturation reveals that BAK and BAX populate an intermediate, but MCL-1 does not. Both intrinsic fluoresence (average emission wavelength (AEW) between 320 and 400 nm) and CD (mean residual ellipticity (MRE) at 222 nm) were used to monitor the process in parallel. (**A**) BAK at 0.9 μM. (**B**) BAX at 1 μM. Both curves show the presence of an intermediate (shaded area). (**C**) In contraste, MCL-1 (0.5 μM) shows no intermediate. Solid lines represent fits to a 2-state equilibrium unfolding model ($\Delta G^{\mathrm{FLUO}}_{\mathrm{D\text{-}N}} = 10.9 \pm 0.5$ kcal mol^{-1}, $\Delta G^{\mathrm{CD}}_{\mathrm{D\text{-}N}} = 10.1 \pm 0.4$ kcal mol^{-1}). For BAK and BAX, data points were joined by a solid line to aid visualisation

Multi-motif BCL-2 proteins (BAK, BAX and anti-apoptotic members) are all in the order of two hundred amino acids. Moreover, all monomeric structures characterised to date show a compact globular fold. Thus, from general protein folding considerations they might be expected to show 2-state equilibrium unfolding profiles. Analysis of MCL-1 did reveal that this was the case (Fig. 9.1C). The unfolding free energy obtained by fitting the data was in the order of 10–11 kcal mol^{-1}, again in line with the expected thermodynamic stability of a protein of that size.

In stark contrast, the unfolding curves of BAK and BAX both showed the presence of an intermediate (Fig. 9.1A,B shaded area). This result is interesting given the structural

similarities of BAK, BAX and MCL-1 (*cf.* Chapter 6). The resistance to denaturant appeared to be less for the anti-apoptotic protein. Indeed, it was completely unfolded in ∼7 M urea. Assuming that GdmCl is about twice a strong a denaturant, this would correspond to ∼3.5 M GdmCl. At these concentrations BAK and BAX are not fully unfolded yet (*cf.* MRE values), and instead populate an intermediate state.

It should be stressed that the presence of a single transition in the unfolding curve of MCL-1 is *not* an artefact of using the weaker denaturant urea instead of GdmCl—suggesting that an second transition might have been missed. The CD and fluorescence spectra of the end-point of the curve are consistent with a fully unfolded species (Fig. 9.2C, D). This confirms that the single transition observed for MCL-1 corresponds to unfolding of the protein, and not to the formation of an intermediate as seen for BAK and BAX.

Spectroscopic analysis of the intermediate of BAK indicated that this species is mostly folded (Fig. 9.2). The CD spectrum at 3.8 M showed that the protein retained a strong helical profile, although some structure was clearly lost—the intermediate has a MRE value at 222 nm about 25% lower than than the protein in buffer. This change would normally be interpreted as a loss of helicity. However, the presence of a Trp π–π^* exciton due to the edge-to-face arrangement of the tryptophans in BAK (*cf.* Chapter 7) make this interpretation more difficult. Therefore, the spectroscopic signature observed for the intermediate might be due in part to structural re-arrangements. This might explain the peculiar progression of the MRE signal in the unfolding curve of BAK; the signal first increased (up to ∼2 M) before starting to decrease. Interpreting this signal change as an initial gain of helicity in increasing amount of denaturant would be difficult to rationalise. However, it is possible that some subtle conformational changes affected the coupling efficiency of the Trp exciton.

These CD results are supported by the intrinsic fluorescence emission spectra of the same states (Fig. 9.2B). BAK in buffer (black line) shows a typical blue-shifted emission, consistent with buried tryptophans. At 5.1 M, the average emission wavelength was ∼360 nm, indicative of an unfolded state. The emission spectrum of the intermediate showed a profile half-way between these two extremes. Clearly the environment of the tryptophan is changing—indicated by both the increase in fluorescence intensity, and the shift in average emission wavelength. These results are consistent with the CD spectra for folded, intermediate, and unfolded BAK.

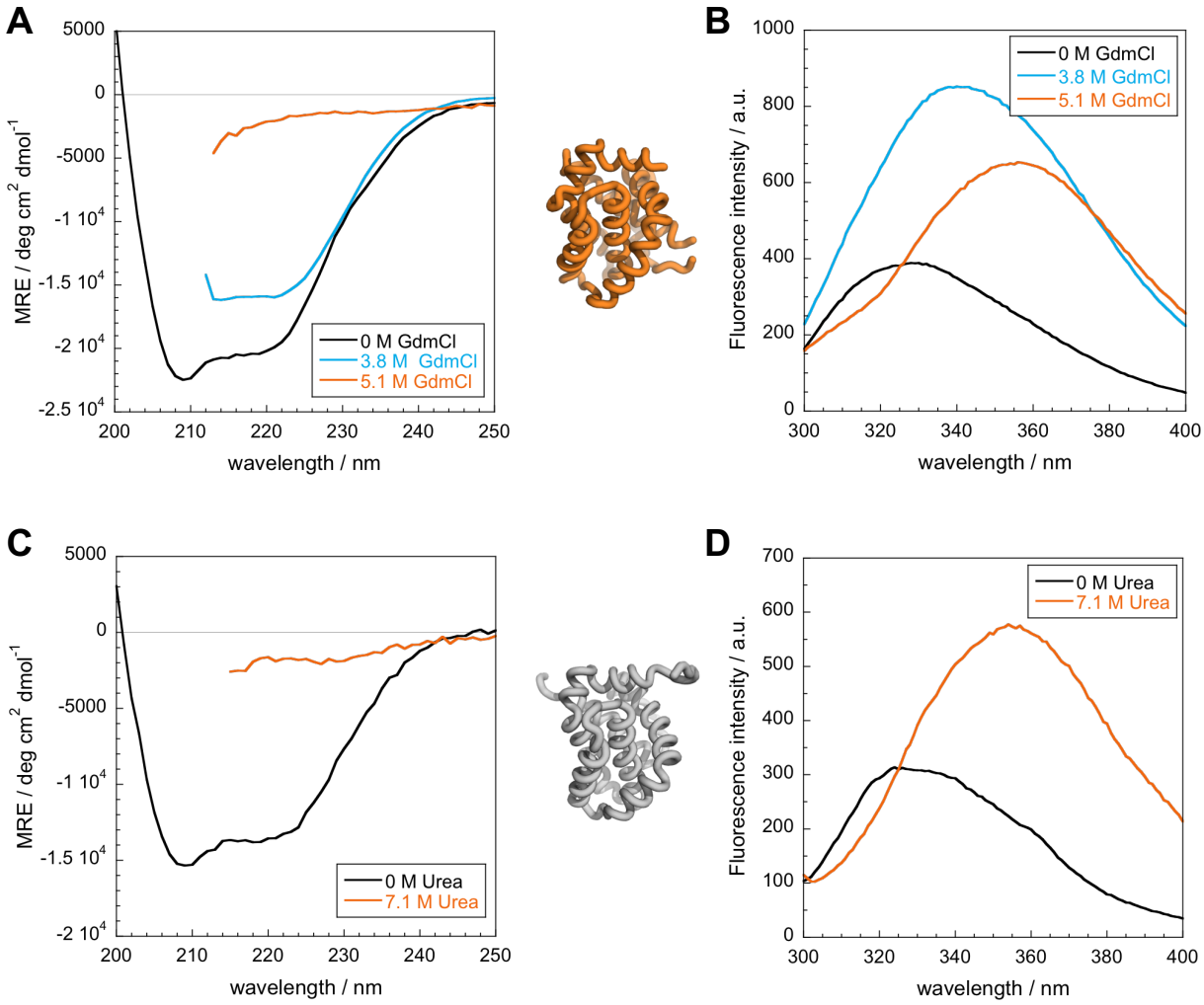

Figure 9.2 – Spectral properties of the intermediate of BAK suggest a partially folded species. (**A**) CD spectra of folded (0 M), intermediate (3.8 M), and unfolded (5.1 M) BAK. The spectra of the intermediate is consistent with a folded state, although some structure is lost. (**B**) Fluoresence emission spectra of the same states (following excitation at 280 nm). The average emission wavelength of the intermediate is more red-shifted than the folded state, but not as much as the unfolded state. Experiments were performed in GdmCl at a protein concentration of 1 μM. The CD spectra (**C**) and fluoresence emission spectra (**D**) of MCL-1 (0.5 μM) in urea are shown for comparaison. MCL-1 is fully unfolded in 7.1 M urea, which would correspond to ∼3.5 M GdmCl; the concentration of denaturant where BAK and BAX populate an intermediate.

9.3.2 The intermediate is multimeric

The presence of a high-energy intermediate on the energy landscape of BAK and BAX is interesting, and the link to the detergent-treated oligomers enticing. As a first step towards bridging these results, the multimeric nature of the intermediate needed to be demonstrated. Indeed, a partially folded monomeric state would be equally consistent with the results presented in the previous section.

In order to test whether the intermediate state observed in GdmCl was oligomeric, three

separate approaches were taken. Chemical cross-linking in denaturant was performed to investigate the presence of oligomers, and SEC was used to analyse their stoichiometries. Moreover, equilibrium chemical denaturation curves at different concentrations of protein were measured. The apparent stability of multimeric species is concentration-dependent. Thus, if the intermediates of BAK and BAX are indeed oligomers, then the positions of their transitions should shift with changes in protein concentration.

The concentration of BAK was varied 10-fold, and that of BAX 2-fold (Fig. 9.3A, B). In order to spare protein, the curves of BAK were performed at 1/10 of the volume normally used for chemical denaturation curves (90 µL instead of 900 µL). Due to the precision limit of the liquid dispensing robot, larger concentration steps were required, resulting in fewer total points.

For both BAK and BAX, changing the protein concentration shifted the positions of the transitions. Lower protein concentrations reduced the denaturant span over which the intermediate was populated. In the case of BAX, the two transitions even started to merge. These results are consistent with the following set of coupled equilibria:

$$nA \rightleftharpoons A_n \rightleftharpoons nU$$

where A represents the monomer, A_n the oligomer, and U the unfolded monomer. Increasing the total concentration of monomer always favours multimeric states, which results in larger apparent stabilities. In terms of chemical denaturation, this means a larger span of denaturant over which these states are populated. Not only does it imply that the oligomer will withstand larger concentrations of denaturant before fragmenting and unfolding, it also means that it will be formed at lower concentrations of denaturant (*i.e.* the monomer becomes comparatively less stable). Therefore, these equilibrium chemical denaturation results are consistent with oligomeric intermediates.

The multimeric nature of these species was further confirmed by chemical cross-linking (Fig. 9.3C). At 0 and 2 M GdmCl, results were consistent with the proteins populating monomeric states. For BAK, trace amounts of dimers could be observed. However these results might be attributed to the non-equilibrium nature of cross-linking experiments (*cf.* Chapter 7 for a more detailed discussion). The first transition occurs at lower concentrations of denaturant for BAK than BAX, potentially explaining this discrepancy. In the presence of 4 M GdmCl (the intermediate state), both proteins showed cross-linked species consistent with large-scale oligomers. Up to tetramers were clearly distinguishable

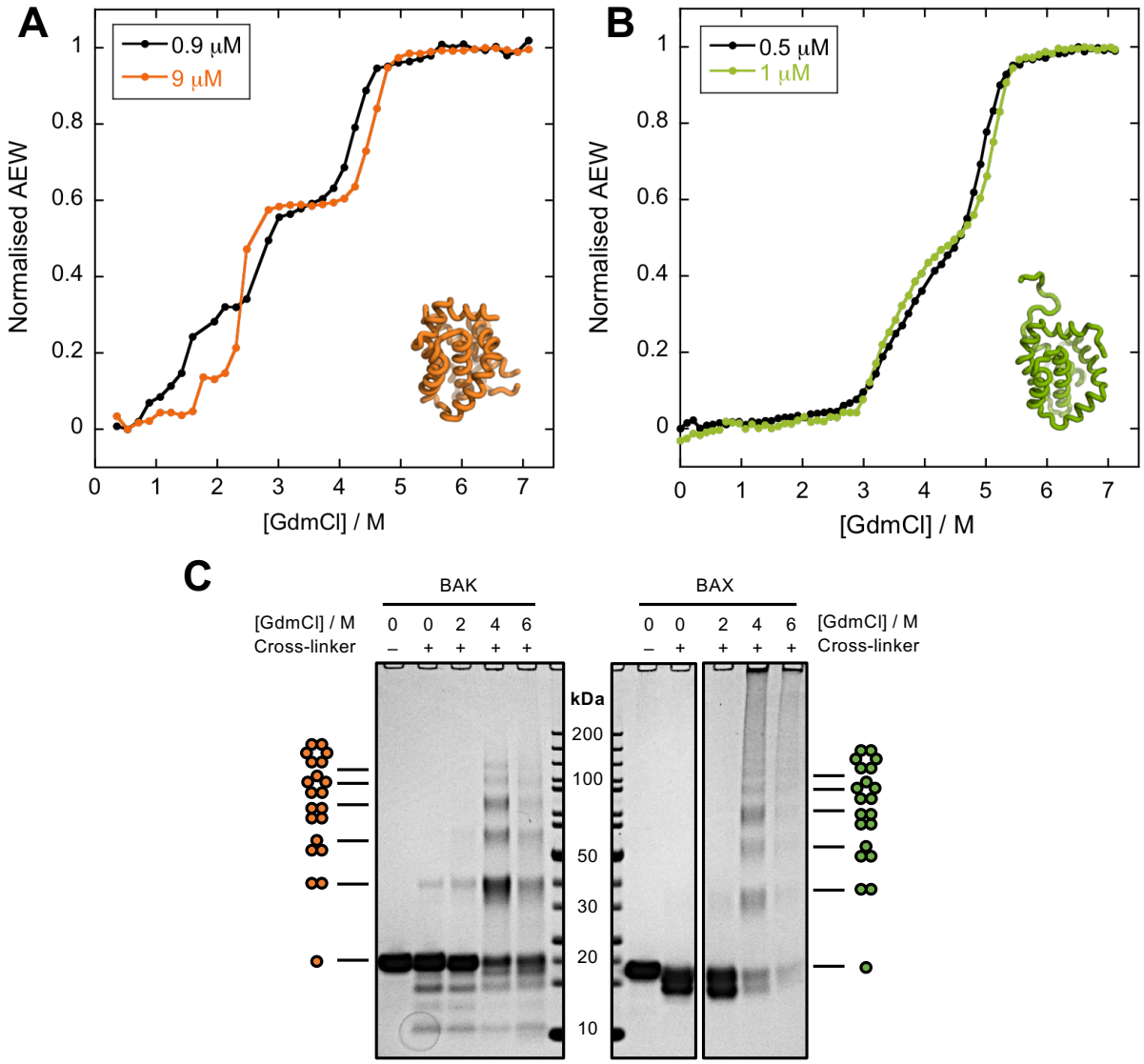

Figure 9.3 – The intermediates of BAK/BAX observed by equilibrium chemical denaturation are multimeric. (**A**, **B**) Equilibrium chemical denaturation curves of BAK/BAX performed at different protein concentrations. The positions of *both* transitions are concentration-dependent. At lower concentrations of protein, the transitions move towards each other and start to merge. This is consistent with a multimeric intermediate. Curves were measured by monitoring average emission wavelengths, and the data normalised. Points are joined by solid lines to aid visualisation. (**C**) Cross-linking of BAK and BAX in the prescence of increasing amounts of GdmCl. The intermediates (4 M GdmCl) show the presence of oligomers. Protein concentrations were 22.5 μM, and the protein:EDC:BS3 molar ratios 1:250:125. Samples were dialysed to remove GdmCl prior to SDS-PAGE.

for both BAK and BAX, and potentially some larger species. Therefore, these results confirm the oligomeric nature of the intermediate observed in equilibrium chemical unfolding experiments.

Surprisingly, multimeric species were also observed at 6 M GdmCl—where the proteins

are expected to be unfolded and monomeric. It is possible that the higher concentration of protein used in cross-linking shifted the second transition past this concentration of denaturant. Another explanation might be the dialysis step. Indeed, GdmCl had to be removed prior to SDS-PAGE. It is possible that, as the concentration of denaturant gradually decreased, oligomer started to form. Those would then have been captured by residual amounts of cross-linker.

While chemical cross-linking can reveal the presence of higher-order species, it does not inform on the solution-state distribution of these species. Since complete cross-linking is unlikely—especially for larger oligomers—the distribution of bands observed by denaturating SDS-PAGE will not necessarily be a true representation of the assembly states present in solution (*cf.* Chapter 7 for a more detailed discussion about this point). Therefore, an alternative method was necessary to quantify the heterogeneity of the oligomeric intermediate.

Size-exclusion chromatography was used to investigate the oligomeric nature and heterogeneity of the intermediate. This approach had previously been employed for studying the detergent-treated oligomers (*cf.* Chapter 7). The presence of micelles—and resulting effects on the hydrodynamic radii of the oligomers—complicated the analysis. However, the absence of detergent in the formation of these GdmCl-induced assemblies should allow a more straightforward interpretation.

The column was equilibrated in buffer containing GdmCl, and the samples were pre-incubated at the same concentration of denaturant for at least an hour prior to injection. Matching the conditions before and during analysis ensured that no re-equilibration of the system would occur during the elution. It is noted that while the system may be at equilibrium, microscopic reversibility might still result in exchanges between states on the time-scale of the elution ($\sim$30 min). This would affect the confidence with which the solution-state may be assessed form chromatographic analysis. However, the lack of severely broaden peaks suggested that the different species did not significantly exchange on the time-scale required to perform the analysis. Hence, the resulting profiles were assumed to be representative of the solution ensemble.

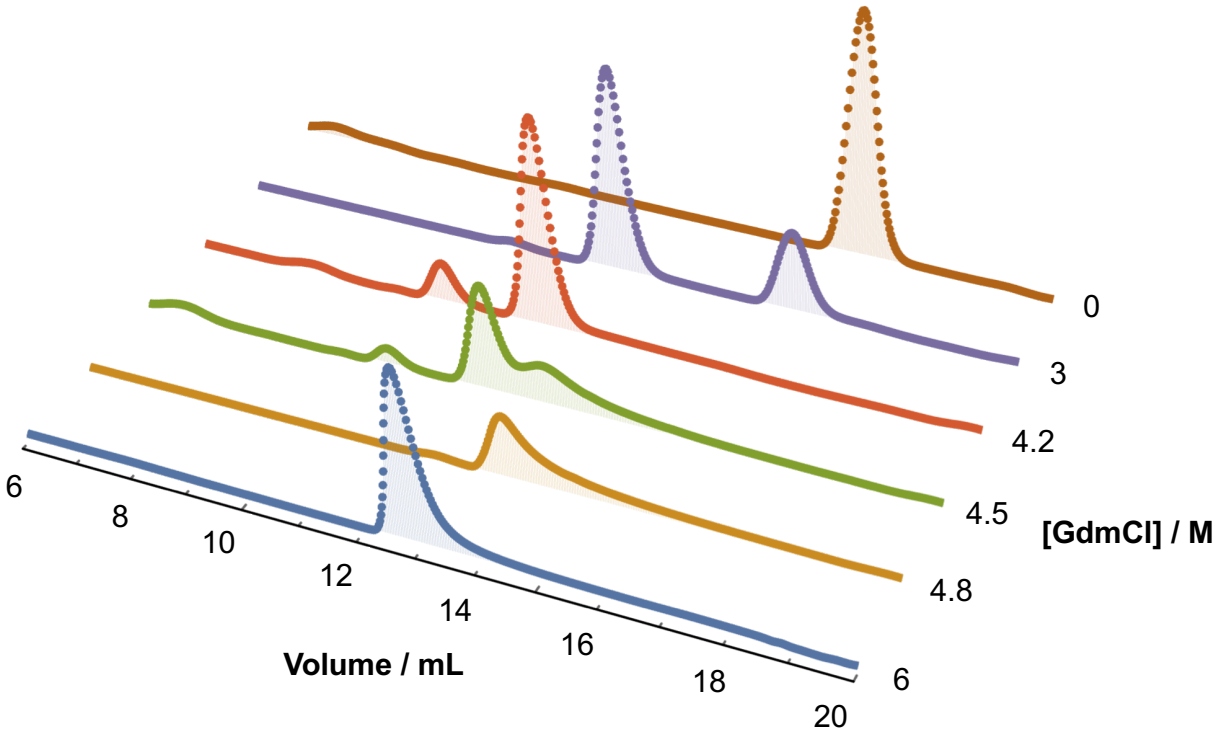

Figure 9.4 – SEC analysis of BAK in denaturant. The column (Superdex 200 10/300) was equilibrated in 50 mM sodium phosphate pH 7.0 buffer containing the indicated amount of GdmCl. BAK samples were pre-incubated ($\geq$ 1 h) at the same concentration of denaturant before injection onto the column. Protein concentrations were ~10 μM (except the 4.8 M condition, which was 5 μM). Elution volumes for each peak are reported in Table 9.1.

Table 9.1 – Elution volumes of BAK species from SEC performed in denaturant (Fig. 9.4).

	Elution volume / mL			
[GdmCl] / M	Folded	Oligomer 1	Oligomer 2	Unfolded
0.0	16.6	–	–	–
3.0	16.1	12.9	11.0	–
4.2	–	12.4	10.6	–
4.5	–	12.3	10.5	13.3
4.8	–	–	–	13.5
6.0	–	–	–	12.7

Consistent with the results presented in Fig. 9.1 and 9.3, the monomer (∼16.5 mL) underwent oligomerisation around 3 M GdmCl. Two peaks corresponding to multimeric species were observed, with a major species at ∼12.5 mL, and a minor species at ∼10.5 mL. These oligomers started to disappear at 4.5 M GdmCl, giving way to a single trailing peak at ∼13 mL. This transition was consistent with the unfolding of the protein observed using spectroscopic techniques.

Interestingly, the intermediate appeared to be relatively homogeneous; consisting of only two species. This result contrasts with the heterogeneity observed for the detergent-treated oligomers (*cf.* Chapter 7). It is noted that in the detergent case, SEC performed at low concentrations of protein also suggested homogeneity. Thus, it is possible that these distributions hide underlying heterogeneities as well. However, the absence of protein:detergent complexes should give greater confidence in the fact that these peaks represent unique oligomeric states.

At the protein concentrations used for these experiments, the distribution appeared to be dominated by the smaller oligomer (∼12.5 mL). Due to the presence of denaturant, analysis of oligomeric states from elution volumes is not straightforward; a smaller, more unfolded, oligomer may elute earlier than a larger, more compact, one. However, when the experiment at 4.2 M was performed with twice the protein concentration, the peak at ∼10.5 became proportionally more populated. This concentration-dependence is consistent with the notion that the species eluting at ∼12.5 mL is composed of fewer subunits.

The presence of denaturant affected the elution properties of all species—indicated by the shifts in elution volumes as the concentration of GdmCl was increased (*cf.* Table 9.1). For the monomer of BAK, going from 0 to 3 M GdmCl affected its elution volume by about 0.5 mL. For the most abundant oligomer, going from 3 to 4.5 M affected its elution volume by about 0.6 mL. Thus, an elution volume of ∼13.5–14 mL in buffer appears reasonable for this species. The back-calculated mass range (70–89 kDa, *cf.* Materials and Methods) would suggest a tetramer (77 kDa). However, this is without taking into account the probable non-globular structure of the oligomer, which should result in a larger elution volume than its mass would indicate. Thus, this species is unlikely to be the tetramer. Instead, it is probably the dimer. This would be consistent with the fact that it was also the most abundantly cross-linked oligomeric state. Following a similar line of reasoning, the less abundant species might be the tetramer. The trimer is unlikely given the significant difference in elution volumes. Larger species cannot be excluded, but would be improbable

given their low abundances in cross-linking experiments.

Further insights into the assembly of these oligomers were gained from performing out-of-equilibrium denaturing chromatography. Here, instead of matching the pre-injection conditions with that of the column, concentrations of denaturant were deliberately made different. Therefore, re-equilibration of the system occurred during elution, which gave insights into the assembly and kinetics of the oligomeristion process in chemical denaturant. In Fig. 9.5A and B, BAK and BAX were oligomerised in the presence of 4 M GdmCl, followed by SEC on a column equilibrated in buffer. Both proteins showed similar elution profiles, which included at least three species. The most abundant states eluted at ∼14 mL, which are likely to be the dimers (*vide supra*), and the second most abundant species might be tetramers. Compellingly, the extent of 're-folding' to the monomeric states were not very pronounced over the time-scale of these elutions. These results suggest that fragmentations of the oligomers are slow processes.

Not pre-incubating BAK before performing SEC at 4.5 M GdmCl resulted in a significantly different elution profile than when the sample was pre-incubated (Fig. 9.5C). This suggests that the oligomerisation in denaturant is a slow process, which is not completed on the time-scale of the elution (∼10 min). Interestingly, this result echoes the slow oligomerisation kinetics observed in detergent (*cf.* Chapter 7).

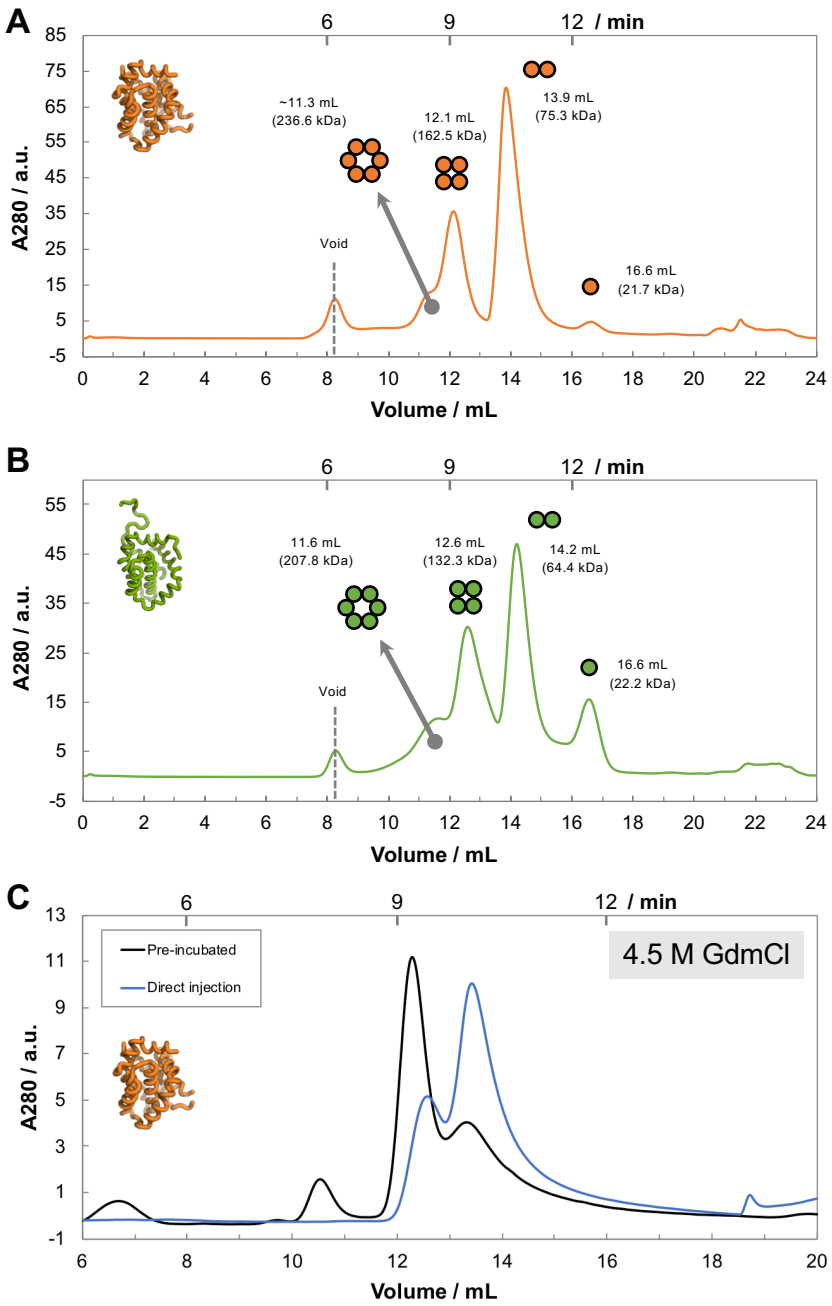

Figure 9.5 – Oligomers of BAK and BAX form and fragment slowly. Buffer-exchanging GdmCl from BAK/BAX oligomers reveals a slow breakdown. (**A**) BAK (25 μM) pre-incuabated in 4 M GdmCl for 2h30 before injection. (**B**) BAX (25 μM) pre-incubated in 4 M GdmCl for 3h before injection. SEC was performed on a Superdex 200 10/300 column equilibrated in 50 mM sodium phosphate pH 7.0 containing no denaturant. Masses were calculated from the corresponding elution volumes using the column calibration (*cf.* Materials and Methods). Presumed oligomeric states are indicated as spheres. (**C**) Oligomer formation is slow as well. The distribution changes if the sample is not pre-incubated, indicating that equilibrium is not reached on the time-scale of a SEC run. The column was equilibrated in buffer containing 4.5 M GdmCl. BAK was either pre-incubated at the same concentration of denaturant for 1h30, or directly injected. Protein concentrations were 11 and 14 μM for the pre-incubated and direct injection samples respectively. Times of elution after injection are indicated above each plot for reference.

About a third of the amount of putative dimer (∼12 mL) present at equilibrium was already formed after 9 min. Surprisingly, most of the protein appeared to be in the unfolded state, not the monomeric state. This interpretation is based on the trailing profile of the peak at ∼13.5 mL, which matches the elution profile of unfolded BAK. This result might suggest that unfolding is on-pathway to the formation of the oligomers.

Importantly, these experiments revealed slow processes for both the formation, and the fragmentation, of the oligomeric species. Further kinetic investigations were undertaken, and the results presented in the next section.

9.3.3 The intermediate is off-pathway to the unfolding of BAK

Preliminary chromatographic results indicated that formation of the oligomeric species in denaturant is slow. This is not dissimilar to the kinetic results obtained with detergents (*cf.* Chapter 7). In order to gain insights with higher temporal resolution, spectroscopic techniques were employed to study the oligomerisation and unfolding of BAK in GdmCl. Circular dichroism spectroscopy was used for these experiments, as it allowed an unambiguous assessment of unfolding.

Unlike for a 2-state folder, the presence of a populated intermediate makes kinetic analysis more challenging. As a first step, single- and double-jump experiments were performed (Fig. 9.6A). BAK was manually mixed into buffer containing GdmCl, and the reaction was followed on a CD spectrophotometer by monitoring the change in MRE value at 222 nm. For the single-jump experiment (x), monomeric BAK in buffer was mixed to a final concentration of GdmCl where the protein was fully unfolded (5.1 M). For the double-jump experiments (y and z), monomeric BAK in buffer was mixed to a final concentration of GdmCl where the protein was multimeric (3.8 M; reaction y). Once this reaction reached equilibrium, more GdmCl was added to the sample to fully unfold the protein (5.1 M; reaction z).

The end states of the single- and double-jump experiments were identical. This was confirmed by recording CD spectra of the samples after each kinetic experiment (Fig. 9.6B). BAK in 5.1 M GdmCl showed the typical profile of an unfolded protein, regardless of the path taken. The intermediate of the double-jump experiment (after reaction y) showed the profile of the intermediate observed in equilibrium chemical unfolding experiments.

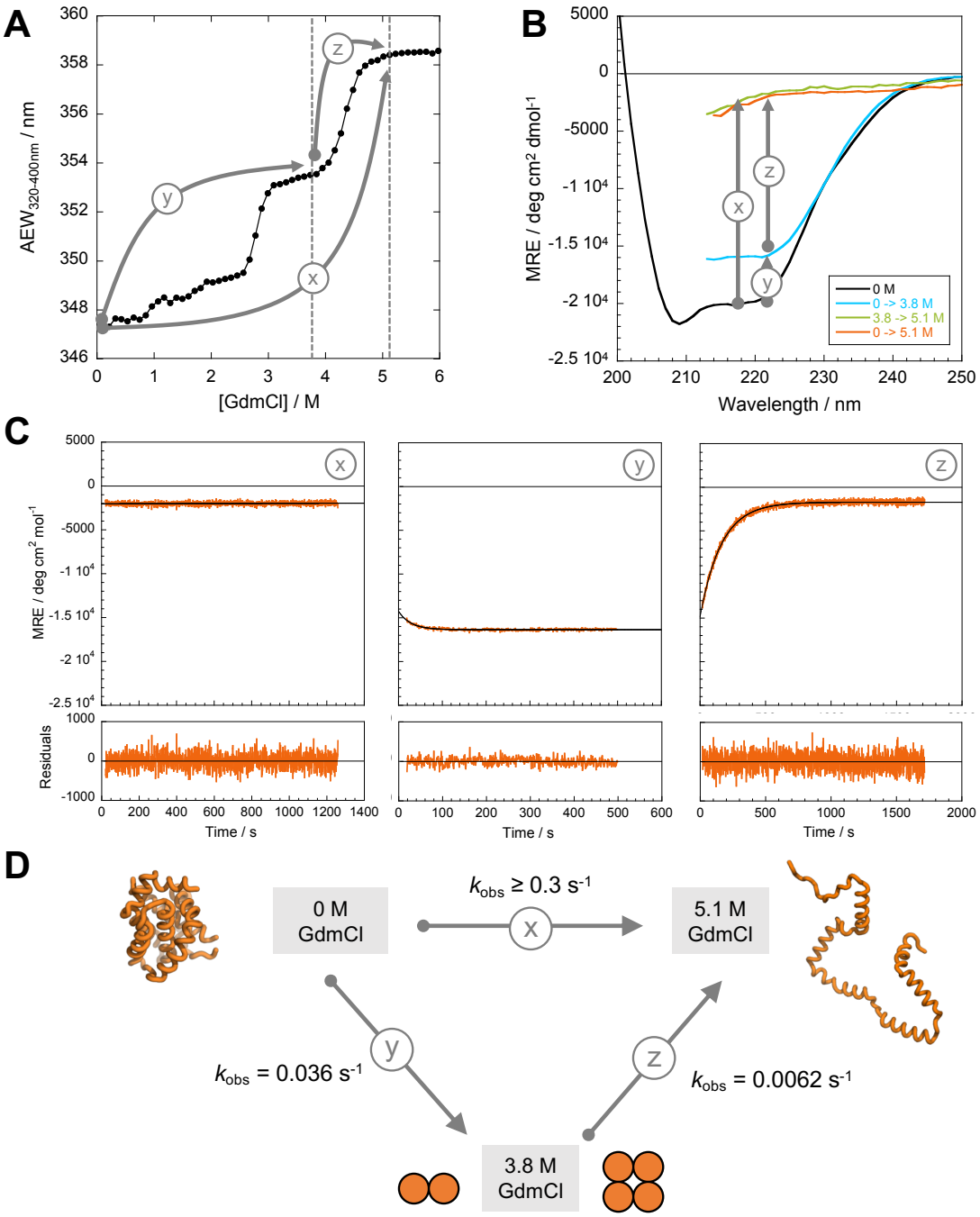

Figure 9.6 – Kinetic studies suggest that the oligomeric intermediate is off-pathway to the unfolding of BAK. Reactions were manually assembled, and the reaction monitored by CD. Both single-jump (x) and double-jump (y and z) experiments were performed. (**A**) The jumps are indicated on the equilibrium unfolding curve of BAK for visualisation. (**B**) CD spectra of monomeric BAK and the end-states of the GdmCl jumps. (**C**) Kinetic traces for the GdmCl jump experiments: X – jump from 0→5.1 M GdmCl, 24.5→1 μM protein; Y – jump from 0→3.8 M GdmCl, 24.5→2.45 μM protein; Z – jump from 3.8→5.1 M GdmCl, 2.45→1 μM protein. Double-jump reactions were fitted to a single exponential function, and x was fitted to a straight line. Lower panels indicate residuals to the fits. The same MRE scale was used to aid comparaison between the reactions. (**D**) Schematic of the results. Reaction x was over before the measurement started (∼20 s), therefore $k_{obs} \geq 0.3$ s^{-1}.

The kinetic traces for each reaction are shown in Fig. 9.6C. The single-jump experiment (reaction x) revealed that the protein had fully unfolded within the dead-time of the measurement (∼20 s). No signal change was detected past this time, and the MRE value at 222 nm was consistent with an unfolded state.

As expected from the SEC results, reaction y showed that the oligomeric state was reached slowly ($k_{\text{obs}} = 0.036$ s^{-1}). Moreover, the data fitted a single-exponential function, suggesting a first-order process. As with detergent, it appears that formation of the oligomer is not rate-limited by a multi-meric event. Instead, it is likely to be determined by a conformational change or partial unfolding of BAK. Surprisingly, the amplitude of the reaction was positive. From the CD spectra at equilibrium, the net signal change going from monomer to oligomer is negative. The MRE value at the end of the reaction matched that of the equilibrium experiment. Thus, the higher value at the start of the reaction suggest that the protein unfolded to some extent during the dead-time of the measurement. Following this undetected fast phase, a slower re-folding event lead to the oligomeric state. This is consistent with the chromatographic result obtained in Fig. 9.5C, where an apparent accumulation of unfolded BAK was observed during the formation of the oligomer.

Fragmentation of the oligomeric species to unfolded monomers (reaction z) was an even slower process ($k_{\text{obs}} = 0.0062$ s^{-1}). This reaction also fitted a single exponential function, again suggesting an unimolecular process as the rate-limiting step of the process. In contrast to the formation of the oligomers, there appeared to be no fast event during the dead-time of the measurement. Indeed, the start of reaction z matched the end point of reaction y.

These preliminary kinetic results contain surprising mechanistic information about the oligomerisation and unfolding of BAK in denaturant. Comparison between the kinetics from reaction x with the double-jump experiments (y and z) strongly suggest that the oligomerisation is off-pathway to the unfolding of the protein. If it was on-pathway, the unfolding would be limited by the fragmentation reaction (z), which should result in a much slower process. This is inconsistent with the fast unfolding observed in reaction x. While it appears that the oligomer is off-pathway to the unfolding, the reverse might not be true. Indeed, the starting MRE value of reaction y—together with the result presented in Fig. 9.5C—suggest that the unfolded state might be on-pathway to the oligomerisation. This intriguing result might explain the larger apparent stability of BAK/BAX; more denaturant is required to unfold them than MCL-1, although the proteins have similar

folds and sizes. The energetic accessibility of an oligomeric intermediate from the unfolded state would 'delay' the population of the unfolded state. This would result in a greater apparent stability of the monomer.

9.3.4 Relationship to detergent-treated oliogomers

The previous sections demonstrated the presence of oligomeric intermediates on the energy landscapes of BAK and BAX. This ensemble appeared to be mostly formed of dimers with some tetramers (at least for BAK). Thus, it is tempting to speculate a link to the oligomerisation observed in detergent (*cf.* Chapter 7). However, it is difficult to assess the similarity between these states in the absence of structural information.

Results presented in Chapter 7 suggested that detergents promote the assembly of BAK and BAX through a stabilisation of their oligomeric states (or a destabilisation of their monomers). Therefore, equilibrium chemical denaturation of BAK oligomers was performed as a way to bridge the two sets of results. The following was rationalised; if the intermediate observed on the landscape in buffer was related to the oligomers formed in detergent, then presence of detergent should lead to an apparent stabilisation of this intermediate. This should be reflected on the unfolding curve in a manner analogous to increasing the total concentration of monomer—the intermediate would be populated over a wider range of denaturant concentrations.

The curve of BAK in the presence of PS20 (10×CMC) showed that the intermediate did indeed become more stable (compare Fig. 9.7A, B). The unfolding of the protein was barely complete at 7 M GdmCl, while 5 M were sufficient in the absence of detergent. Moreover, the transition to the oligomer was already complete at 2 M when PS20 was present, while 3 M GdmCl were necessary to generate the intermediate in buffer. These results clearly highlight that the presence of detergent stabilises an intermediate on the energy landscape of BAK.

While it is therefore enticing to speculate a link between the two set of oligomers, some outstanding questions remain. In particular, the spectral properties of the different states obtained in either buffer or detergent did not completely match (Fig. 9.7D, E). Some possible reasons for these discrepancies are discussed below. The unfolded states appeared to have the same red-shifted average emission wavelengths. Moreover, their CD spectra showed a complete loss of α-helical structures. These results are consistent with unfolded

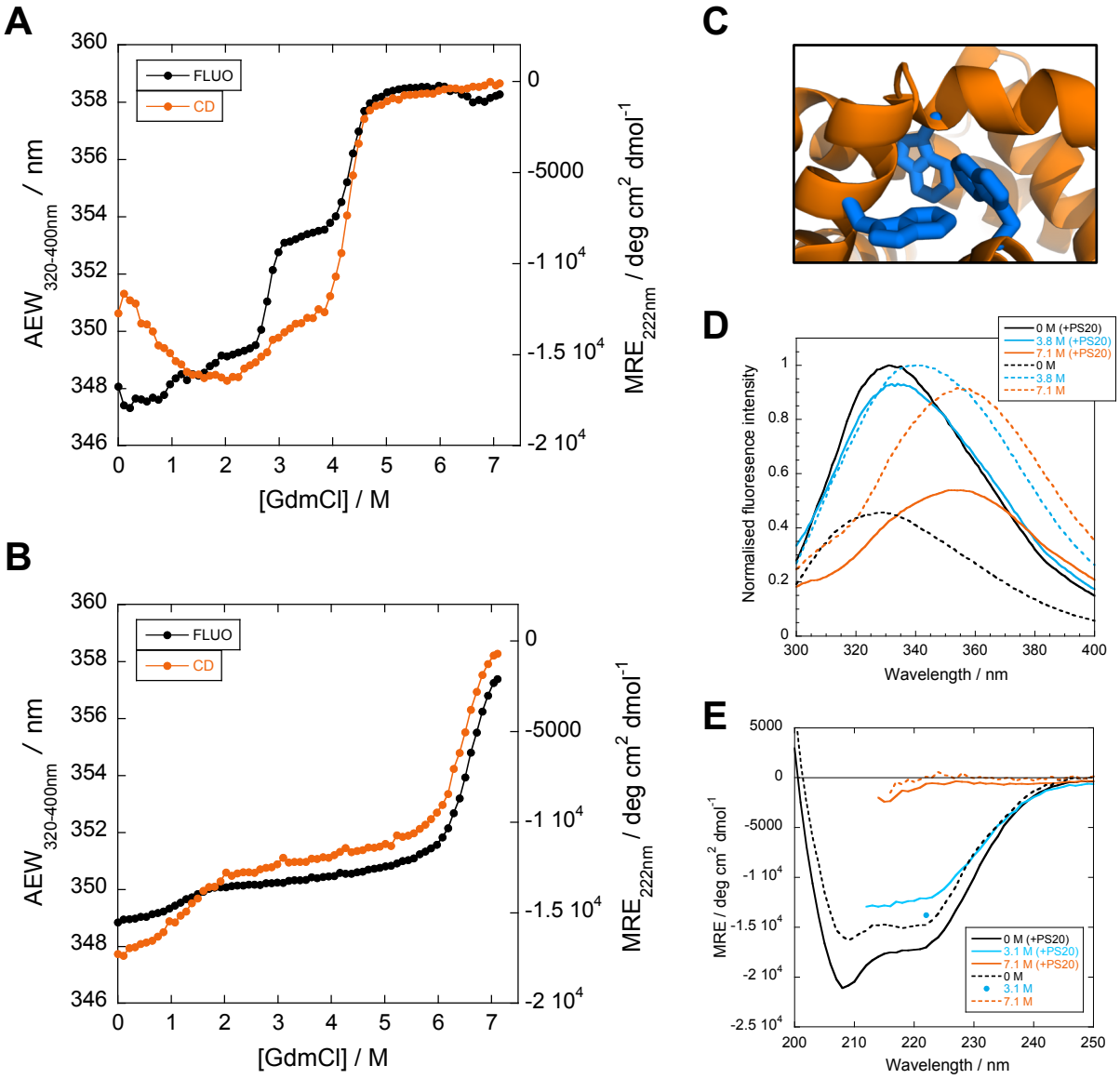

Figure 9.7 – The presence of detergent stabilises the intermediate. (**A**) Equilibrium chemical denaturation of BAK monomer (0.9 μM) in buffer (reproduced from Fig. 9.1A). (**B**) Equilibrium chemical denaturation of detergent-treated BAK oligomers (1 μM total monomer concentration, 10×CMC PS20). In comparaison to the result obtained in the absence of detergent, the region between both transitions spans a much wider range of denaturant concentration. This result suggests that the detergent stabilises this particular intermediate. (**C**) Tryptophan cluster in the structure of BAK monomoer (PDB:2YV6). These are the only tryptophans present in the construct of BAK used for this work. (**D**) CD spectra of BAK in the presence (solid lines) or absence (dashed lines) of detergent for selected points along the chemical denaturation curves. The intermediate appears mostly folded in both cases. (**E**) Fluorescence emission spectra of BAK in the presence (solid lines) or absence (dashed lines) of detergent for selected points along the chemical denaturation curves. It is noted that in the presence of detergent, the starting point and the intermediate of BAK have similar spectral properties. Data were normalised for ease of comparaison.

species in both cases. The greater fluorescence intensity observed in buffer might be a result of quenching from the presence of detergent. In contrast, the spectra of the starting points did not match. This is not surprising given that BAK is already oligomeric at the start of the curve. Interestingly, the spectra of the intermediate in buffer, and the starting and intermediate states observed in detergent show striking similarities. These results are consistent with the notion that the intermediate formed in denaturant relates to the oligomers formed in detergent. Slight discrepancies might again be attributed to the presence of PS20.

The main point of contention to link these oligomers, is the presence of an initial transition in the curve obtained in detergent. In buffer, this transition relates to the oligomerisation step. But since the protein is already multimeric in detergent, this first transition might be expected *not* to occur. One possible explanation for this discrepancy is the more heterogeneous ensemble of oligomers formed in detergent (*cf.* Chapter 7). Thus, it is possible that addition of denaturant breaks the large-scale oligomers, and 'normalises' the ensemble to (mostly) dimers. The much smaller amplitude observed for the first transition in Fig. 9.7B is consistent with this hypothesis.

These results comparing chemical denaturation curves in the presence and absence of PS20 revealed that an intermediate state is indeed stabilised by the presence of detergent. Hence, it is interesting to speculate that the high-energy intermediate present in buffer relates to the oligomers formed in detergent. Moreover, the stabilities, spectra, and chromatographic elution profiles of the oligomers obtained in both denaturant and detergent, indicate defined structural states. Thus, it is not unreasonable to assume that they are the same structure(s). This would relate the biological function of BAK (and BAX) to the features encoded on their energy landscapes. However, more experiments are warranted to confirm the link between these states.

9.4 Binding BAK to MCL-1 by unfolding

The interaction between BAK and MCL-1 is expected to occur *via* the BH3 motif of BAK. In the form of a peptide, this sequence stretch binds MCL-1 avidly (K_d = 0.077 nM). However, no interaction between full-length BAK and MCL-1 was observed in buffer (*cf.* Chapter 6). This lack of binding was attributed to the structural positioning of the BH3 motif within BAK; the contact interface points inwards, and is integral to the globular fold of the protein. However, these proteins form hetero-dimers in the presence of detergent (*cf.* Chapter 8); highlighting the possibility of forming this MCL-1:BH3 interaction in the context of full-length proteins.

The work presented in this Chapter showed a complex energy landscape for BAK and BAX. Both proteins populate an oligomeric intermediate, which appears to be off-pathway to the unfolding. These results raise question about the relationship of these oligomers to the ones formed in detergent. May their structural nature allow them to interact with MCL-1? Therefore, hetero-dimerisation between BAK and MCL-1 was revisited in the absence of detergent, using chemical denaturation instead.

It was rationalised that BAK should be able to interact with MCL-1 if it was unfolded—denaturing BAK should expose its BH3 motif, making it available for interaction with MCL-1. However, the amount of GdmCl necessary to unfold BAK would also result in MCL-1 unfolding. So instead of incubating both proteins in the presence of denaturant, an alternative strategy was necessary. Pre-unfolding BAK in GdmCl, followed by mixing with MCL-1 in buffer might allow the BH3-exposed unfolded state to be 'captured'. In this scenario, the binding reaction competes with the refolding of BAK (Fig. 9.8A). Assuming irreversible reactions—which is reasonable given the slow dissociation rate constant of BAK_{BH3} from MCL-1 ($k_{off} = 5.6\ 10^{-4}\ s^{-1}$; $t_{1/2} \approx 20$ min)—the following set of rate equations can be expressed:

$$\begin{cases} \frac{d[U]}{dt} = -k_f[U] - k_{on}[U][MCL\text{-}1] \\ \frac{d[F]}{dt} = k_f[U] \\ \frac{d[BAK{:}MCL\text{-}1]}{dt} = k_{on}[U][MCL\text{-}1] \end{cases} \tag{9.1}$$

where [U] represent the concentration of unfolded BAK, [F] the concentration of folded BAK, [MCL-1] the concentration of free MCL-1, k_f the folding rate constant, and k_{on} the bimolecular association rate constant. Moreover, [MCL-1] = $[MCL\text{-}1]_0$ – [BAK:MCL-1],

where $[\text{MCL-1}]_0$ is the concentration of MCL-1 at the start of the reaction. Formation of a hetero-dimer was assumed based on the results obtained in Chapter 8.

This system of differential equations was numerically integrated with different set of parameters to evaluate the feasibility of the 'capture' experiment. The concentration of BAK was fixed to 10 μM, a value deemed appropriate for experimental validations. The biomolecular association rate constant was set to 7.2 10^6 s^{-1}, which corresponds to the value of k_{on} for the binding of BAK_{BH3} to MCL-1. The folding rate constant of BAK being unknown, its value was varied over two orders of magnitude around typical refolding values to examine its influence (Fig. 9.8A). The concentration of MCL-1 was also varied by a 100-fold to assess its effect on the amount of complex formed (Fig. 9.8C). It is noted that only the fold-excess of MCL-1 over BAK affects the outcome, not the absolute concentrations of protein.

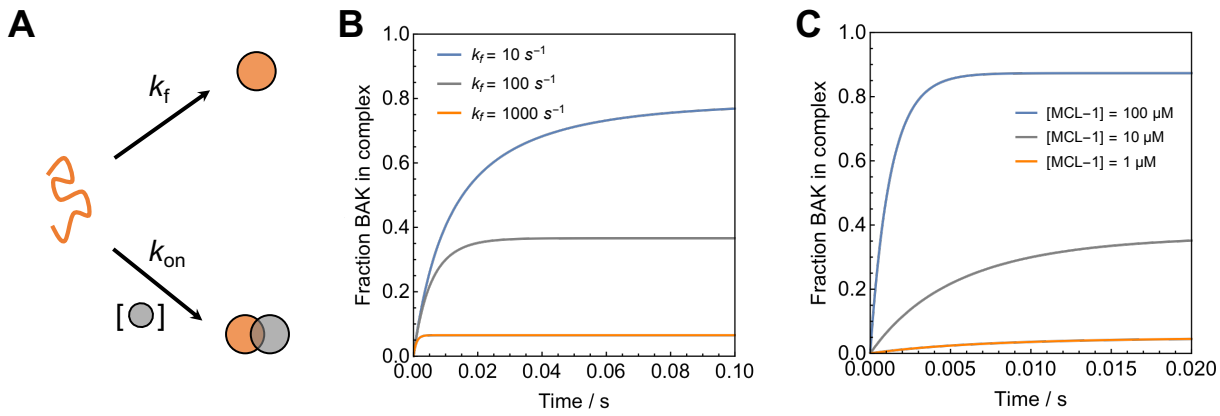

Figure 9.8 – Simulations for the capture of unfolded BAK by folded MCL-1. (**A**) Reaction scheme for the competition between refolding and binding of unfolded BAK. Irrevesible reactions are assumed. BAK and MCL-1 are represented in orange and grey respectively. (**B**, **C**) Kinetic simulations for formation of BAK:MCL-1 (indicated as a fraction of the total BAK concentration). Fixed parameters were: [BAK] = 10 μM, k_{on} = 7.2 10^6 s^{-1}. (**B**) Simulations varying k_f ([MCL-1] = 10 μM). (**C**) Simulations varying [MCL-1] (k_f = 100 s^{-1}). The fractional values at infinite time represent the equilibrium distribution between folded and captured BAK (under irreversible conditions).

These simulations indicated that hetero-dimerisation between BAK and MCL-1 should be possible. Moreover—and unless the refolding of BAK is really fast—an appreciable amount of complex should be formed with μM concentrations of protein.

Capture experiments were performed by doing different GdmCl jumps, and the outcome monitored by SEC. Experiments starting from the unfolded, the intermediate, and from the unfolded to the intermediate states were performed. BAK was pre-incubated in either 6 M (unfolded) or 3.5 M (oligomeric) GdmCl, before being mixed with MCL-1 in buffer. After

mixing, the remaining concentration of denaturant was either 1 M—where both proteins were expected to be folded—or 3.5 M—where BAK was expected to be oligomeric. It is important to highlight that the jumps discussed in this section only refer to the changes experienced by BAK. MCL-1 was always pre-incubated in buffer (except for Fig. 9.9D). Therefore, its jumps were 0→X M GdmCl. Following mixing, samples were directly injected onto the column, and the elution performed at 1.4 mL/min to reduce the time before detection. Under these conditions, the last peak eluted ∼12 min after mixing the samples. An experiment where both proteins were pre-incubated in the presence of 3.5 M GdmCl was also performed. These different conditions were tested to see whether the starting and final conditions made a difference to the formation of BAK:MCL-1 hetero-oligomers.

Unfolded BAK was captured by MCL-1 when jumped into conditions promoting the folded state of both proteins (6→1 M, Fig. 9.9A). A large-scale hetero-oligomer eluted at ∼10 mL, and SDS-PAGE analysis revealed the presence of both proteins. Relative band intensities also suggested a complex with equimolar stoichiometry. It is noted that the presence of both proteins in that peak does not prove interaction. However, the MCL-1 control experiment showed only a monomer. Since the presence of MCL-1 in the oligomer peak was conditional on the presence of BAK, it suggests that the two protein interact.

The oligomer peak was broad, not symmetrical, and close to the void volume of the column. Moreover, its elution volume indicated a very large species (∼420 kDa). Thus, it appeared that the hetero-oligomerisation of BAK and MCL-1 in buffer lead to the formation of large-scale aggregate-like species. This contrasts with the clean hetero-dimerisation observed in detergent (*cf.* Chapter 8). It is possible that hydrophobic parts of BAK become exposed in the complex. In the absence of a micellar environment, these would likely lead to aggregation.

Unfolded BAK was not captured by MCL-1 when the final condition was 3.5 M GdmCl (6→3.5 M, Fig. 9.9B). SDS-PAGE analysis revealed that the oligomer peak at ∼14 mL only contained BAK. Moreover, the elution profile of the mixture was clearly the sum of its parts (blue and red dashed lines), which is indicative of a lack of interaction. This result might be attributed to the fact that at this concentration of denaturant, the unfolded state of MCL-1 is favoured. Formation of a complex might be expected to stabilise the protein, but if unfolding occurs before binding, no complex would be formed. The folding kinetics of the murine homologue of MCL-1 (85.8 % identity) has been determined (Dr Joseph Rogers, unpublished results). From this data, the unfolding rate of MCL-1 under

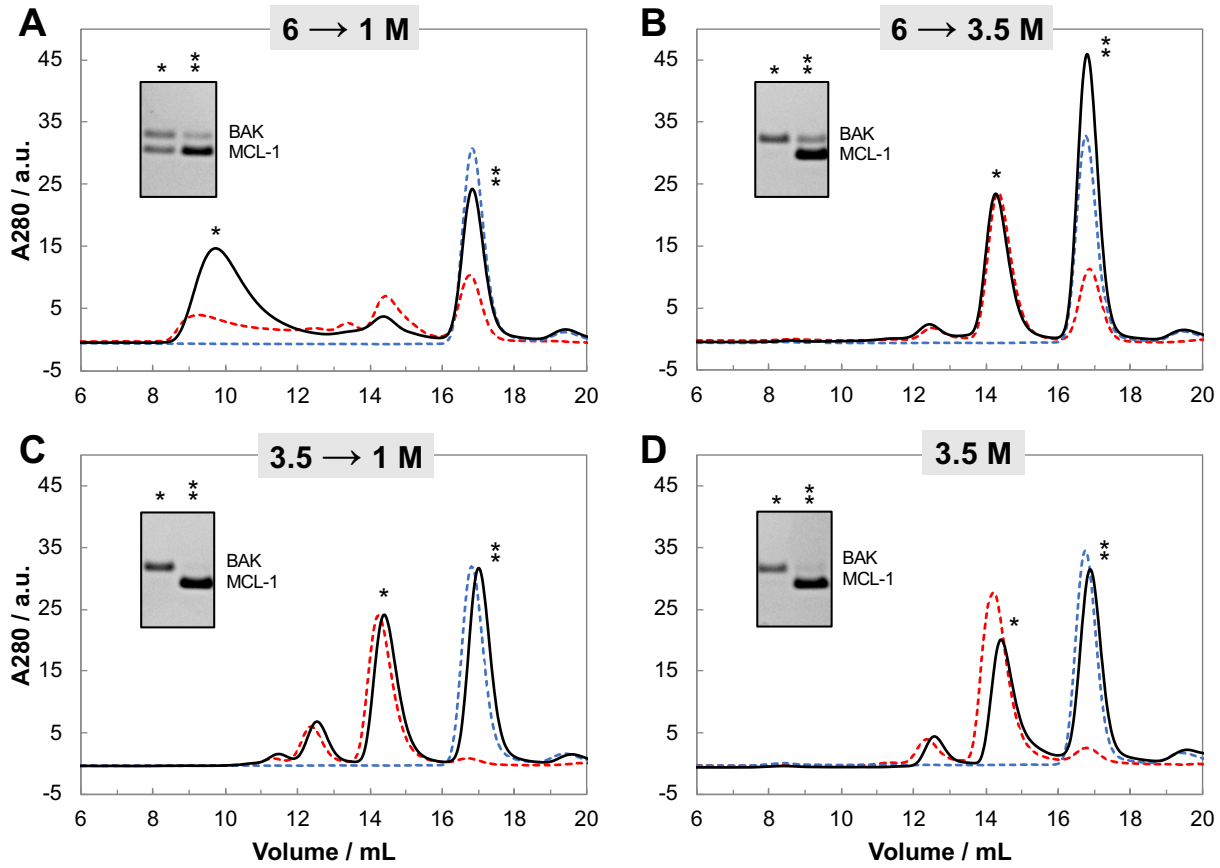

Figure 9.9 – BAK:MCL-1 hetero-oligomers can be formed by capture of fully unfolded BAK. Monomeric BAK was pre-incubated in the presence of GdmCl (2–6 h) before rapid mixing with MCL-1 (in buffer). The results were immediatlely analysed by SEC (Superdex 200 10/300; flow-rate = 1.4 mL/min; elution time for the last peak ∼12 min). The denaturant jumps undergone by BAK are indicated above each plot. Protein mixing experiments are represented by solid black lines, and control experiments (identical conditions but with only one protein) are indicated as dashed lines (red and blue for BAK and MCL-1 respectively). Peaks of the protein mixing experiments (solid black lines, indicated by astericks) were analysed by SDS-PAGE. The results are shown as insets on each chromatogram. Final concentrations of each protein were 10 μM. (**A**) Mixing unfolded BAK with MCL-1 to conditions where both proteins are expected to be folded and monomeric resulted in the formation of a hetero-oligomer peak. (**B**) Mixing unfolded BAK with MCL-1 to conditions where BAK is expected to form homo-oligomers did not result in the formation of mixed oligomers. (**C**) Mixing of oligomeric BAK with MCL-1 to conditions where both proteins are expected to be folded and monomeric did not result in the formation of hetero-oligomeric species. (**D**) Incubating both proteins together in the presence of 3.5 M GdmCl (3 h) did not result in hetero-oligomer formation.

the conditions of the experiment may be estimated ($k_{\text{obs}}^{3.5M} \approx 1$ s^{-1}). In contrast, the rate of association at the start of the reaction is only $k_{\text{obs}}^{t=0} = 7.2\ 10^{-4}$ M s^{-1}, and expected to drop quickly as MCL-1 unfolds. Thus, the lack of hetero-oligomerisation for the 6→3.5 M jump is not surprising; at the protein concentrations used in this experiment, unfolding of

MCL-1 competes kinetically with the association reaction.

When starting from the oligomeric state of BAK, no hetero-oligomerisation was detected (3.5→1 M, Fig. 9.9C). It is possible that the homo-oligomers of BAK have their BH3 motifs buried, which would prevent interaction with MCL-1. This would be the case in the BH3-in-groove scenario (*cf.* Chapter 7). While the oligomers are expected to dissociate in 1 M GdmCl, this reaction is slow. Indeed, no appreciable amount of monomer was observed for the BAK control experiment (Fig. 9.9C, red dashed line). Thus, binding of MCL-1 might be limited by the rate of fragmentation of the oligomers, which would explain the lack of hetero-oligomerisation.

Incubating both proteins at 3.5 M GdmCl also resulted in no interaction. Here, the combination of MCL-1 being unfolded, and the competition with the homo-oligomerisation reaction are probably the reasons behind this outcome.

Together, these results demonstrate the feasibility of forming BAK:MCL-1 hetero-oligomers by capture of unfolded BAK, and confirm the results obtained in detergent. It supports the notion that BAK does not interact with MCL-1 in buffer because the binding energy is not sufficient to offset the folding energy of BAK (*cf.* Chapter 6 for a more detailed discussion on that point). By capturing the unfolded state, the hetero-oligomer can be kinetically trapped. However, for this hypothesis to be true, BAK:MCL-1 formed by 'GdmCl capture' should revert back to monomeric proteins over time.

The reversibility of the hetero-oligomer was tested (Fig. 9.10). BAK was unfolded in 5.5 M GdmCl, and hetero-oligomerisation initated by mixing with MCL-1 to a final concentration of 1.5 M GdmCl. The sample was directly injected, and the result compared with injections performed 3 h and 18 h after mixing (Fig. 9.10A). The monomer peak increased over time, indicating that some of the complex dissociated to monomers. However, the reaction appeared extremely slow, and far from complete after 18 h.

As a more direct test of reversibility, the hetero-oligomer peak from a 6→1 M GdmCl jump was collected (orange area, Fig. 9.10B), and re-injected after 3 h (chromatogram shown as inset). This avoided an incubation step in low denaturant concentration before injection, which could potentially interfere with the reaction. The process was clearly reversible, and a monomer peak started to re-appear. Interestingly, the oligomer peak also appeared to split into at least two set of species.

These results indicate that the hetero-oligomerisation of BAK and MCL-1 in buffer is reversible, at least partially. Therefore, it suggests that the monomeric proteins are the

thermodynamically favoured states in buffer, in line with the observation and hypothesis made in Chapter 6. From the knowledge of the lifetime of MCL-1:BAK_{BH3} ($\tau \approx 30$ min), the slow dissociation kinetics observed here is not surprising. The fact that is appeared to be even slower might be due to additional interactions made in the context of full-length proteins.

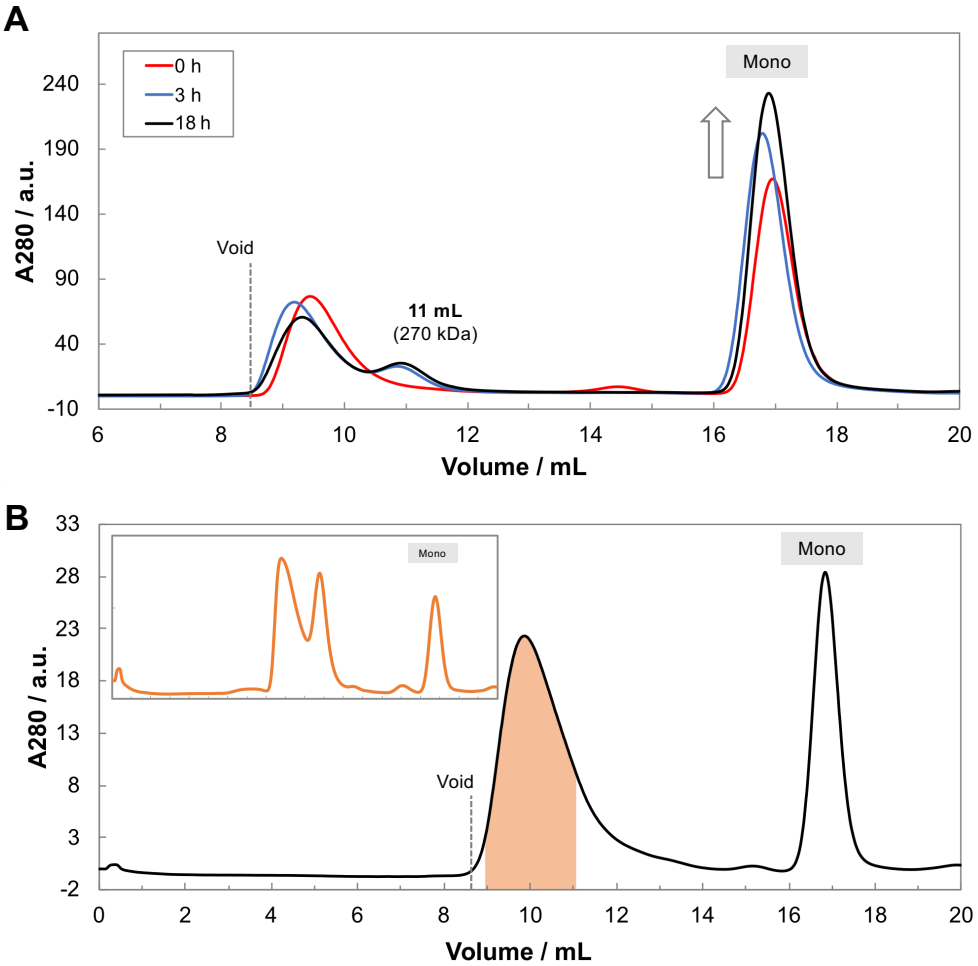

Figure 9.10 – Reversibility of GdmCl-induced hetero-oligomerisation of BAK and MCL-1. (**A**) BAK was pre-incubated in 5.5 M GdmCl for 30 min before mixing with His-tagged MCL-1 (final concentrations: 1.5 M GdmCl, 26 µM BAK, 67.4 µM his-MCL-1). The mixture was analysed by SEC directly, after 3h, and after 18 h. (**B**) BAK was pre-incubated in 6 M GdmCl for 1h30 before mixing with MCL-1 (final concentrations: 1 M GdmCl, 12.7 µM BAK, 13.2 µM MCL-1). This sample was directly injected onto the column, and the oligomer peak collected (orange area). This fraction was concentrated ($\sim$ 5×, to compensate for the dilution from the elution) and re-injected after 3 h of incubation at 25 °C (inset). SEC experiments were performed on a Superdex 200 10/300 equilbrated in 50 mM sodium phosphate pH 7.0 buffer.

9.5 Discussion

In this Chapter, the link between the energy landscape of BCL-2 proteins and the results obtained with detergents (Chapters 7 and 8) was investigated. Equilibrium denaturation experiments revealed the presence of an intermediate for BAK and BAX, but not for MCL-1. This intermediate was shown to be multimeric, and chromatographic analyses in denaturant suggested a dimer, with some tetramers as well. The link with the oligomers formed in detergent is enticing, and further experiments were performed to investigate this hypothesis. Chemical denaturation in the presence of PS20 showed an apparent stabilisation of the intermediate, suggesting that the detergent- and denaturant-induced oligomers are related. However, more experiments are warranted to asses the kinship of these species.

The use of denaturant allowed additional kinetic experiments aimed at investigating the oligomerisation reaction to be performed. Interestingly, formation of the oligomer in GdmCl was slow, echoing the results obtained with detergents. Its fragmentation to unfolded monomers was slow as well, indicating a state with significant kinetic stability. In contrast, direct unfolding of the protein was fast, suggesting that the oligomeric intermediate is off-pathway. However, preliminary experiments indicated that the reverse might not be true; chromatographic and spectroscopic results suggested that an unfolded intermediate might accumulate during the formation of the oligomer. This would imply that unfolding is on-pathway to the oligomerisation. A more complete kinetic characterisation of the system should allow the validation of this hypothesis.

These equilibrium and kinetic experiments provided new insights into the oligomerisation of BAK and BAX. Importantly, the oligomers seem to be already present on the buffer energy landscapes, but as high-energy states. This would be consistent with the notion that detergents promote oligomerisation through thermodynamic means; by favouring the oligomeric state over the monomeric one.

It is important to emphasise that the 2-state behaviour observed for MCL-1 is *not* specific to anti-apoptotic BCL-2 proteins. Equilibrium chemical denaturation experiments with A1 and BCL-X_L also revealed intermediates (data not shown). However these proteins remained monomeric when treated with detergents, as did MCL-1. Thus, the link between the presence of an intermediate on the energy landscape, and oligomerisation in membrane-like environments is not straightforward. Instead, it might suggest that BCL-2 proteins (pro- and anti-apoptotic) lie on a scale of 'oligomerisation propensity'. This hypothesis

is interesting in light of the evolutionary relationship between these proteins. Indeed, all members are paralogues of each other. Their last common ancestor might have been able to oligomerise, and anti-apoptotic members would have arisen later in evolution. These would have evolved not to form pores any more, but vestiges of their evolutionary past might still be encoded in their energy landscapes.

The hetero-oligomerisation between BAK and MCL-1 was studied in the context of chemical denaturation. These proteins were shown not to interact in buffer, despite a very high affinity of MCL-1 for the BH3 motif of BAK. This absence of binding was attributed to the buried nature of the motif. Using 'denaturant jumps', pre-unfolded BAK could be bound to MCL-1, demonstrating that full-length proteins are indeed capable of interacting in buffer. However, this complex did appear to slowly revert to monomers, suggesting a kinetically-trapped metastable state. This result further supports that the lack of interaction between BAK and MCL-1 in buffer is rooted in thermodynamics, as hypothesised in Chapter 6.

The work presented in this Chapter aimed at bridging the results for the oligomerisation and interaction of BCL-2's in detergent, with the energy landscapes of these proteins. More work is required to confirm some of these initial findings. However, a link between energy landscape, biophysical properties, membrane environments, and biological function seem to emerge.

Chapter 10

Conclusions and outlooks

10.1 Evolution of the field since 2014

It has now been almost two decades since intrinsically disordered proteins have been recognised to be functional and widespread in biology (Wright & Dyson, 1999, Uversky *et al.*, 2000, Dunker *et al.*, 2001). Because IDPs challenge many of the canonical tenets of molecular biology—in particular the structure-function paradigm—the role of protein disorder has been argued and debated (Zhou, 2012). The observation that IDPs are more common in higher organisms (Ward *et al.*, 2004), specifically in processes such as signalling and transcription, suggests that there might be an evolutionary link between multi-cellularity and protein disorder. Many arguments have been put forward, often without experimental validation. However, over the past half-decade, an increasing number of kinetic and mechanistic studies have been undertaken. These have shed light onto some of the biophysical consequences of protein disorder, paving the way to a functional understanding of their biological significance.

10.1.1 Coupled folding and binding reactions have unstructured transition states

Coupled folding and binding reactions start as disordered ensembles that ultimately collapse to singly-defined structural states; prompting much speculation about the presence of 'bound-like' structures in the disordered ensembles. The detection of such species in the free state of the IDP has led to the conclusion that conformational selection might define coupled folding and binding reactions (Fuxreiter *et al.*, 2004, Iešmantavičius *et al.*, 2014,

Krieger *et al.*, 2014). A few model reactions have had their transition states characterised by Φ-value analysis or analogous techniques (Bachmann *et al.*, 2011, Dogan *et al.*, 2013, Giri *et al.*, 2013, Hill *et al.*, 2014, Rogers *et al.*, 2014a, Dahal *et al.*, 2017a). Interestingly, most studies reported very little bound-like structure at the transition state; suggesting that the rate-determining step of these reactions is not the folding of the IDP, but the encounter between the partners. Although not a formal kinetic demonstration of mechanism (Gianni *et al.*, 2014, Shammas *et al.*, 2016), these reports suggest that induced-fit (sometimes termed dock-and-coalesce (Ou *et al.*, 2017)) might actually be the dominant mechanism for coupled folding and binding reactions. As more reports appear, it will be interesting to see whether this conclusion holds. Most systems reported to date have relatively simple bound topologies, which might explain the pre-dominance of unstructured transition states. However, for reactions necessitating more complex folding, the nature of the TS has been reported to be more structured (Hill *et al.*, 2014). Thus, it is possible that the topological complexity of the bound state may influence the binding mechanism.

10.1.2 Affinities are mostly determined by the lifetime of the bound state

The amount of bound-like structure present in the free state of the IDP might be expected to influence the speed of association, regardless of the mechanism—by increasing the observed binding rate (for a conformational selection mechanism), or through faster folding (for an induced-fit process). Early reports on the modulation of the residual helicity of the IDP p53 showed that it affected its affinity to Mdm2 (Borcherds *et al.*, 2014). Interestingly, it was shown that most of the effect was the result of altered dissociation rate constants (Crabtree *et al.*, 2017). In fact, for the systems reported so far, the lifetime of the bound state always appears to be the main determinant of affinity (Rogers *et al.*, 2014b, Crabtree *et al.*, 2017, Dahal *et al.*, 2017b, 2018). For these reactions, the complex is assumed to be unaffected by the changes in residual helicity observed in the free state of the IDP. Therefore, altered dissociation rate constants are expected to be the consequences of altered entropic costs of unbinding.

The fact that changes in affinity are mostly a result of altered k_{off} (not k_{on}) might be primarily a consequence of these reactions having unstructured transition states, *i.e.* occurring early on the reaction coordinate. In accordance with Hammond's postulate, these

TS will be energetically more closely related to the free state than the bound state. Thus, changes in ΔG should be mostly due to changes in $\Delta G^{\ddagger}_{\text{off}}$, which would explain these results.

10.1.3 Non-contacting residues can influence the affinity

For the interaction between two folded proteins, the residues present at the binding interface are expected to determine most—if not all—of the thermodynamics and kinetics of the reaction. In contrast, coupled folding and binding reactions have been shown to be sensitive to changes outside of the binding interface. For examples, it has been demonstrated that mutations of helix-flanking prolines can affect the affinity of IDP:partner interactions (Borcherds *et al.*, 2014, Crabtree *et al.*, 2017). Moreover, 'shuffling' of non-contacting residues—hence conserving the overall sequence composition—can affect the residual structure of the free IDP (Harmon *et al.*, 2016); which has been shown to affect binding (*vide supra*). The possibility to alter the affinity (and kinetics) of PPIs by making changes outside of the binding region might be a functional advantage of protein disorder. Indeed, the interface is a crucial determinant of specificity, and modulating the affinity without affecting it could be of importance in signalling networks.

10.1.4 Binding pathways are encoded into the IDP

In coupled folding and binding reactions, structuring of the IDP is conditional; and depends on the presence of a partner macromolecule. This might suggest that the binding partner plays a role in dictating the mechanism/pathway of the folding step. To an extent it does, and it has been shown that mutations of the folded protein can affect binding (Rogers *et al.*, 2014a). However, a comparative Φ-value analysis of coupled folding and binding reactions has revealed that the transition state is actually encoded in the IDP, *not* templated by the partner molecule (Crabtree *et al.*, 2018). This unilateral encoding of folding information might explain the energetic frustration observed for the interface of PPIs involving disordered partners (Jemth *et al.*, 2014). Whether this frustrations is a consequence of the promiscuity of IDPs, or their 'evolutionary youth', is not clear.

10.1.5 Phase-separation

Following the publication of series of accounts describing phase-separating proteins, the field of intrinsically disordered proteins has taken a new direction. Since the first observa-

tion of such a phenomenon *in cellulo* a decade ago (Brangwynne *et al.*, 2009), the number of proteins reported to undergo liquid de-mixing has exploded (Boeynaems *et al.*, 2018). Because the proteinaceous constitutents of these droplets are disordered both before and after separation, the IDP field has taken an interest in understanding the grammar, and properties of these processes. It has been demonstrated that high flexibility (due to the presence of glycines), and the possibility for many cation-π interactions (between K/R and F/Y residues), were important determinants of phase separation (Nott *et al.*, 2015, Boeynaems *et al.*, 2018). Here, the disordered nature of the protein is crucial, as it allows a large number of weak interactions to be formed, which results in liquid-to-gel properties. Polyvalent low-affinity binding motifs have also been shown to be important for enabling fast transport across nuclear pores (Milles *et al.*, 2015). Interestingly, the formation of these protein droplets appears to be very sensitive to solution conditions. The biological advantages of phase-separation is already becoming apparent; offering the possibility to generate cellular compartments *à la carte* in response to various stimuli (including PTMs). These membrane-less organelles provide delimited regions of altered physico-chemical properties *inside* the cell. These have been shown to concentrate certain macromolecules, while exuding others; and also affect the stability of nucleic acids (Nott *et al.*, 2016). Therefore, disorder allows phase-separated protein droplets to be formed, which potentially offer the canvas for new biochemistry.

10.2 Conclusions from this thesis

10.2.1 Electrostatics only contribute marginally to the association rate constant

IDPs are characterised by an amino acid content bias. Large hydrophobic residues tend to be under-represented, while prolines, glycines, and charged residues are over-represented (Uversky *et al.*, 2000, Romero *et al.*, 2001). The presence of a larger numbers of charges is particularly intriguing from the point of view of electrostatically-enhanced association. In Chapter 4, the contribution from charge-charge interactions to the association rate constant of two model coupled folding and binding systems was investigated. Surprisingly, the effect was within an order of magnitude, suggesting that electrostatics does not contribute much to k_{on}. This contrasts with reports of large enhancements observed for PPIs involving

folded partners (Schreiber & Fersht, 1996). The model systems being different in terms of mechanism, thermodynamic and kinetic signatures, it may be anticipated that these findings are representative of PPIs involving disordered partners. The few additional studies on the contribution of electrostatics to kinetic rates reported to date seem to agree with the results from this thesis. Furthermore, transforming a folded protein into an IDP was shown to affect the affinity for its partner through a drop of its electrostatically-enhanced association rate constant (Papadakos *et al.*, 2015), adding weight to this hypothesis.

For electrostatic forces to influence binding rates, they have to bias the interaction potential between the interacting partners. It is possible that for the magnitude of these force vectors to be sufficient, the presence of electrostatically complementary surfaces on the proteins is required. Disordered proteins lack a binding interface in isolation, and instead quickly re-configure. A consequence of this dynamics is a constant re-shuffling of force vectors, which might blunt steering effects. This might explain why PPIs involving disordered partners might not show electrostatically-enhanced association. It is noted that the recent report of a strongly ionic-strength-dependent affinity between disordered proteins appears to challenge this hypothesis (Borgia *et al.*, 2018). It will be interesting to see whether this effect is in the association or the lifetime of the complex. However, it is noted that since both partners remain entirely disordered in this 1:1 complex, no specific binding interfaces exist in either the unbound or the bound state. This difference might explain the discrepancy.

10.2.2 Solution conditions affect coupled folding and binding reactions

Folded proteins are often referred to as 'marginally stable'. However, they remain relatively tolerant to changes in solution conditions. In contrast, IDP ensembles are particularly sensitive to modifications in their environments (Müller-Späth *et al.*, 2010, Soranno *et al.*, 2014). Despite this, the role of solution conditions on coupled folding and binding reactions has not been thoroughly assessed. In Chapter 5, the effect of changing the nature of charged co-solutes was investigated. It was discovered that binding reactions were affected both thermodynamically and kinetically by the nature of the ions, and that the effect was beyond simple electrostatic considerations. The outcome was correlated with structural changes of the IDP, which followed the Hofmeister series. These results seem to indicate

that PPIs involving disordered partners might be more sensitive to their environments, which could be consequence of their shallower energy landscapes. Given their effects on IDP ensembles, it is anticipated that pH and temperature would have similar consequences (Wuttke *et al.*, 2014, Hofmann *et al.*, 2013). In fact, on-going work in collaboration with the group of Prof. John E. Walker (MRC-MBU) has revealed that the assembly state of the IDP IF1 (Walker *et al.*, 1987) is highly salt and pH-dependent. This sensitivity to solution conditions might represent a mechanism of physiological significance. Indeed, biological systems often experience fluctuations in salt gradients, pH, and presence/absence of specific co-solutes. Therefore, this 'biophysical sensing' might constitute a functional advantage of protein disorder.

10.2.3 Regulation of the BCL-2 network

The BCL-2 family of proteins are the gatekeepers of mitochondrial integrity (Czabotar *et al.*, 2014, Kale *et al.*, 2017). Upon apoptotic stimulation, a range of cellular cues are integrated and processed by this network, and the binary output life/death rendered. Despite decades of work, a detailed molecular understanding of the events underlying the function of the network is still lacking. In particular, the regulation of BAK/BAX oligomerisation—the pivotal step of apoptosis—is not fully understood (Cosentino & García-Sáez, 2017). Using a simplified tripartite system, the molecular origins of this biological process were investigated.

Interactions of BH3 motifs

One feature of the BCL-2 family is their shared BH3 motif (Aouacheria *et al.*, 2013). It is often claimed that this segment provides the point of interaction between BCL-2 members. Affinities between BH3-only proteins and anti-apoptotic BCL-2 proteins have been measured, and found to be very tight (Ku *et al.*, 2011). Surprisingly, peptides of BH3 motifs did not interact appreciably with BAK and BAX, despite the very high structural homology between these pro-apoptotic proteins and anti-apoptotic ones. The reason behind this lack of affinity is unclear. However, it is interesting to speculate that it might be the result of negative evolutionary selection. Indeed, even a physiologically irrelevant interaction (MCL-1:MCL-1_{BH3}) was orders of magnitude tighter, suggesting that the binding of amphiphatic helices should be relatively promiscuous. Therefore, one possible expla-

nation for this absence of binding could be to avoid stabilisation of BAK and BAX as a result of complex formation, which otherwise might negatively impact their propensity to oligomerise, and their ability to perform their biological function.

In some BCL-2 proteins, the BH3 motif forms part of the core of the protein. Studies on the network performed in buffer revealed a lack of interaction when the binding partners had their motifs embedded. The motifs themselves (in the form of peptides) had high affinities. Thus, it appeared that the free energy gain of binding was not sufficient to offset the folding free energy. These results indicated that simple biochemical conditions were not sufficient to re-capitulate the biological function of BCL-2 proteins.

Oligomerisation of BAK and BAX

The oligomerisation of BAK and BAX could be recapitulated *in vitro* using detergents as membrane mimics. These detergent-treated oligomers appeared to be physiologically relevant; supported by disulfide cross-linking studies (Iyer *et al.*, 2016), and the fact that effect of detergent was specific to pro-apoptotic BAK and BAX. Preliminary negative-stain EM revealed pore-like arrangements—in line with the expected biological function of these proteins—and analysis by native MS confirmed the formation of higher-order assemblies.

The 'direct' activation mechanism appears to be a popular explanation for describing the events leading to apoptosis (Shamas-Din *et al.*, 2013). Under the assumptions of this process, BH3-only proteins transiently interact with BAK/BAX; triggering their oligomerisation. However, BH3 peptides did not induce any form of assembly under the conditions of the experiments. Moreover, a thermodynamic analysis revealed that this mechanism is unlikely. Instead, the proteins appeared to spontaneously form oligomers in the presence of detergents—suggesting that a membrane environment might do the same.

Mechanism of BCL-2 regulation

The spontaneous oligomerisation observed in the presence of detergent is consistent with the 'indirect' activation mechanism (Chen *et al.*, 2005, Willis *et al.*, 2005). In this model, BAK and BAX are kept in check by anti-apoptotic members. Therefore, the interaction between BAK/BAX and MCL-1 was re-visited in the presence of detergent. In stark contrast to the lack of binding observed in buffer, the proteins were capable of hetero-dimerisation. Moreover, the consequence of this interaction was a complete sup-

pression of the homo-oligomerisation of BAK/BAX. BH3 peptides were capable of displacing these hetero-dimeric complexes; scavenging MCL-1, and leaving BAK and BAX free to oligomerise. By extrapolation to their biological context, this would imply MOMP and cell death. In summary, cellular fate is in the hands of thermodynamically unstable proteins that are prone to oligomerisation. The subtle balance between hetero- and homo-oligomerisation ensures the suppression of pore-formation under steady-state conditions, but this equilibrium can be broken by the presence of BH3-only proteins. These results provide a molecular understanding for the arbitration of apoptosis by the BCL-2 network that is consistent with the 'indirect' activation mechanism of apoptosis.

The oligomerisation is encoded on the energy landscape

Chemical unfolding studies revealed that both BAK and BAX could populate high-energy multimeric intermediates, and analysis by SEC suggested a dimeric species. Unfolding in the presence of detergent stabilised the intermediate. Thus, it is interesting to draw a parallel between detergent- and denaturant-induced oligomerisation. These results suggest a link between the behaviour of the proteins observed in detergent, and the energy landscape encoded by these proteins.

Preliminary kinetic experiments indicated that oligomerisation might be off-pathway to unfolding, but that unfolding could be on-pathway to oligomerisation. This would imply that the oligomerisation represents an energetic 'trap' for the unfolded protein chain. These mechanistic and equilibrium studies provide insights into how the energetic landscape of apparently simple globular proteins might encode complex assemblies and behaviours.

10.3 Future directions

In terms of coupled folding and binding reactions, it will be interesting to see whether the lack of electrostatic-enhancement, and the sensitivity to solution conditions, are indeed general features of PPIs involving disordered partners. Investigations of additional systems and other solvent perturbations should provide valuable insights into the physical components underlying the binding of disordered proteins. This should help characterise the functional relevance of protein disorder.

The detergent framework established for the oligomerisation studies of BCL-2 proteins offers new and exciting opportunities to further characterise this important biological sys-

tem. In particular, it will be interesting to complete the structural characterisation of the homo- and hetero-assemblies. A more in-depth kinetic analysis of the oligomerisation process should also allow mechanistic aspects of the process to be deconvoluted.

Because BCL-2 members have embedded binding motifs, their interactions can be rendered conditional on the folding stability of the proteins. How widespread this biophysical feature might be remains unknown. However, the presence of non-canonical binding interfaces and conditional binding might provide functional diversification while maintaining genome economy, which could be an evolutionary advantage (Cavalier-Smith, 2005).

One surprising aspect of the BCL-2 family was the presence of energetic intermediates on the unfolding landscape of some anti-apoptotic proteins (data not shown). Given the evolutionary relationship between pro- and anti-apoptotic members, it would be interesting to investigate the biophysical properties of an ancestral BCL-2 (Harms & Thornton, 2010). This could provide insights into the origin of oligomerisation within this protein family. Intriguingly, BCL-2 proteins are nuclear-encoded, but act upon the mitochondria. Therefore, such results could provide fascinating answers about the synergistic evolution between eukaryotes and mitochondria.

10.4 Concluding remarks

The thermodynamic hypothesis initially formulated by Anfinsen and his colleagues states that the native state of a protein is encoded by its sequence, and represents the energy minima of its landscape. The resulting structure is then assumed to provide the function. The work presented in this thesis offers some explanations for the ways energy landscapes may encode function into disordered proteins and oligomers—through the sensitivity of marginally stable states to solution conditions, and through conditional assemblage.

Appendices

Appendix A

Activation energy of spectrin association

The association of spectrin domains is slow under physiological-like conditions (k_{on} = 6.3 10^2 M^{-1} s^{-1}, (Shammas *et al.*, 2012)). This remains true in the absence unfavourable electrostatic interactions ($k_{\text{on}}^{I=\infty}$ = 2.27 10^3 M^{-1} s^{-1}, this thesis), thus excluding strong repulsive forces as the origin of this phenomenon. Could it therefore be due to a a high activation energy barrier? This hypothesis was investigated by performing the association kinetics at a range of temperatures. Final concentrations of proteins were twice those used for the ionic strength dependence study in order to obtain measurable rates at low temperatures. The Eyring-like plot of the natural logarithm of the rate against the inverse of the temperature was used to extract the activation energy of the reaction (the slope of the linear fit). Since the reaction is bimolecular, changes in buffer viscosities needed to be considered, as it would impact the collision frequency of the molecules. This is described by the following relationship:

$$k_{\text{on}} \propto A\frac{T}{\eta}\exp\left(-\frac{E_a}{RT}\right) \tag{A.1}$$

where A is the pre-exponential factor—or rate at infinite temperature—η is the viscosity coefficient, E_a the activation energy, R the gas constant and T the thermodynamic temperature. Solvent viscosities were determined using the empirical equation described below, where T is expressed in degree Celsius, and $\eta_{293} = 1.002 \cdot 10^{-3}$ Pa·s (Kestin *et al.*, 1978).

$$\log\left(\frac{\eta}{\eta_{293}}\right) = \left(\frac{20-T}{T+96}\right)\left(1.2364 - 1.37\cdot 10^{-3}(20-T) + 5.7\cdot 10^{-6}(20-T)^2\right) \qquad \text{(A.2)}$$

Performing the experiment between 5–30 °C demonstrated that the logarithm of the association rate constant was linearly dependent on the inverse of the temperature (Fig. A.1). Thus, by taking the gradient of the line, the activation energy could be estimated ($E_a = 11.3(\pm 0.1)$ kcal mol^{-1}). Unfortunately, an analogous analysis could not be performed for k_{off}, as the values obtained from fitting were not accurate enough.

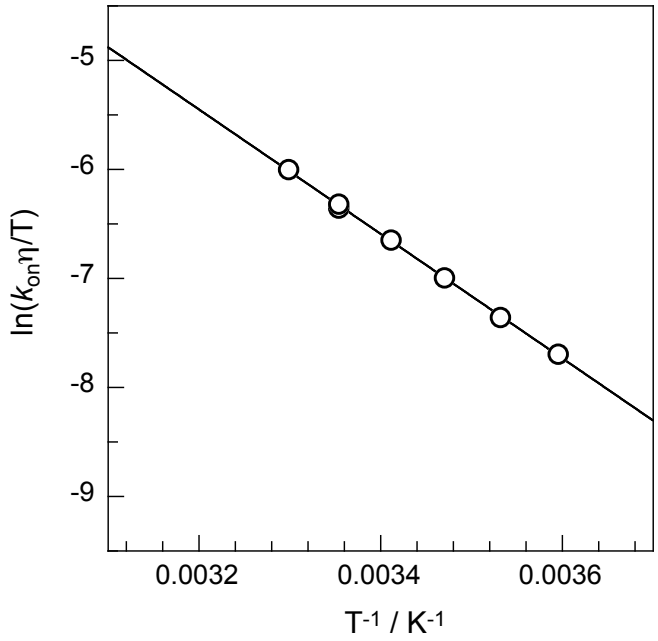

Figure A.1 – Temperature dependence of spectrin association. Eyring-like plot of k_{on} over the temperature range 5–30 °C. Experiments were performed in 50 mM sodium phosphate, 150 mM NaCl, pH 7.0. The two overlayed points correspond to independent measurements using different protein batches. The black line represents the linear fit to the data.

Interestingly, this activation energy was not significantly different from the value corresponding to a very fast associating system (KIX:cMyb, $k_{\text{on}} = 2.1\ 10^7$ M^{-1}s^{-1}, $E_a = 10.9(\pm 0.7)$ kcal mol^{-1}, (Shammas *et al.*, 2013)). Therefore, it appears that the origin of the slow kinetics of spectrin is *not* a high activation energy barrier. It is noted that using results from transition-state theory, the activation energy is closely related to the activation enthalpy ($E_a = \Delta H^{\ddagger} + RT$, (Atkins & Paula, 2010)). On the other hand, the activation entropy ($\Delta S^{\ddagger}$) is described by the intercept of the line at infinite temperature. However, this value also contains other contributions from the pre-exponential factor (*e.g.*

the transmission coefficient, κ), which means that $\Delta S^{\ddagger}$ cannot be directly estimated from an analysis of the temperature-dependence of the rate constant. Therefore, while the activation enthalpy of the reaction is consistent with a fast process, a large activation entropy might be the reason for the slow kinetics of spectrin association. This would be consistent with the large degree of structure formed at the transition state (Hill *et al.*, 2014).

It is noted that since the reaction involves a significant amount of folding, it might be expected to have a non-negligible $\Delta C_{\mathrm{p}}^{\ddagger}$ (Oliveberg *et al.*, 1995). This would affect the activation enthalpy according to the following relationship:

$$\Delta H_{T_2}^{\ddagger} = \Delta H_{T_1}^{\ddagger} + \Delta C_{\mathrm{p}}^{\ddagger}(T_2 - T_1) \tag{A.3}$$

Assuming a negative value for $\Delta C_{\mathrm{p}}^{\ddagger}$—an reasonable approximation considering that the transition state is more structured than the proteins in isolation—the activation energy obtained from Fig. A.1 would be under-estimated.

In order to gain further insights into the effect of the heat capacity of activation, a wider range of temperatures was investigated. Although association reactions were performed up to 60 °C, fitting became impossible above 35 °C; the fluorescence signal changed direction, and did not appear to probe association. Temperature denaturations experiments revealed that the complex dissociated at temperatures higher than ∼35 °C, thus explaining why no association could be measured above this point (Fig. A.2).

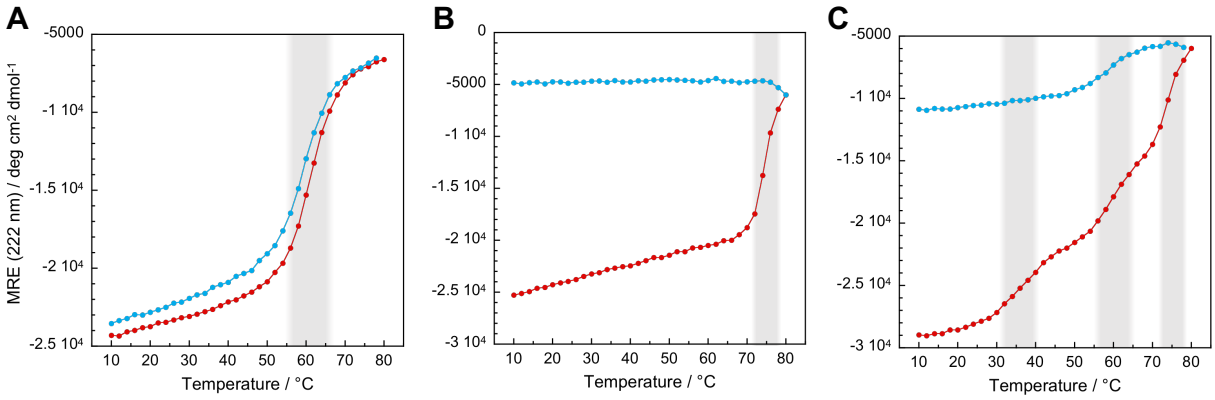

Figure A.2 – Temperature denaturation of spectrin proteins monitored by CD at 222 nm. Heating curves are shown in red, and cooling ones in blue. (**A**) α0α1 alone shows partial denaturation reversibility (concentration = 4.8 μM). (**B**) β16β17 alone shows higher temperature stability, but not denaturation reversibility (concentration = 8.5 μM). (**C**) Temperature denaturation of pre-formed complex shows three cooperative unfolding transitions but no reversibilty (α0α1 = 3.6 μM and β16β17 = 5 μM). All samples contained DTT. Points were joined by lines to aid visualisation. Unfolding transitions are highlighted as shaded area.

Temperature denaturation experiments were monitored by CD at 222 nm for both individual proteins, and the complex. Only α-spectrin showed partial refolding after temperature denaturation ($T_{\mathrm{m}} \sim 60$ °C). Although more stable with respect to temperature ($T_{\mathrm{m}} \sim 75$ °C), heat-induced unfolding of β-spectrin was not reversible. The denaturation curve for the complex appeared to show three cooperative unfolding transitions. By comparing the apparent mid-points for single proteins and the mixture, the transition corresponding to the dissociation of the complex could be assigned ($T_{\mathrm{m}} \sim 35$ °C).

Appendix B

Charged residues in spectrins, PUMA, and MCL-1

Number of charged residues and estimated isoelectric points for the proteins used for the work presented in Chapter 4 and 5.

- Net charges were calculated by summing the numbers of Asp, Glu, Arg and Lys, assuming full ionization at pH 7 (the pH at which the experiments were performed).
- Histidines were not included in net charge calculations, since their protonation states were not known. However, it is noted that they were only a few per protein. Moreover, assuming unperturbed pK_a's ($\sim$6), contributions to the net charges of the proteins would only be $\sim$10% of their numbers.
- pIs were estimated at the sequence level using ProtParam (http://web.expasy.org/protparam/).
- The sequence of PUMA_{BH3} was acetylamidated (both termini were protected and no longer contained any charges). This was considered for estimating the net charge of the peptide, but not the pI.
- The peptide t-PUMA_{BH3} had a N-terminal TAMRA dye (zwitterion with a neutral net charge) and a free C-terminus. This was considered for estimating the net charge, but not the pI.

	# Asp	# Glu	# Arg	# Lys	# His	# AA	% Asp+Glu	% Arg+Lys	% charged	Net charge	pI
α0α1	8	25	11	15	6	164	20.1	15.8	36.0	−7	5.48
α1 (folded)	8	15	6	12	6	116	19.8	15.5	35.3	−5	5.68
α0 (IDR)	0	10	5	3	0	48	20.8	16.7	37.5	−2	5.06
β16β17	7	26	17	11	4	188	17.5	14.9	32.4	−5	5.56
β16 (folded)	4	15	12	6	2	109	17.4	16.5	33.9	−1	6.20
β17 (IDR)	3	11	5	5	2	79	17.7	12.7	30.4	−4	5.11
MCL-1	11	12	14	10	4	159	14.5	15.1	29.6	+1	8.20
PUMA_{BH3}	2	8	6	0	1	34	29.4	17.6	47.0	−4	4.75
t-PUMA_{BH3}	2	8	7	0	1	35	28.6	20.0	48.6	−4	4.97

Appendix C

PUMA:MCL-1 binding under reversible conditions

Dissociation rate constants between PUMA and MCL-1 presented in Chapter 5 were obtained with a dye-labelled version of the peptide that also contained an extra residue at the N-terminus. In order to exclude that the observed results were due to the presence of the dye, experiments with acetyamidated PUMA (used for measuring k_{on}) were performed to obtain K_d. Although the absolute values were slightly different (∼5-fold), the trend with respect to the Hofmeister series was conserved (*cf.* Table 5.1).

Table C.1 – Kinetic and thermodynamic parameters for PUMA (acetylamidated) binding MCL-1 in the presence of different salts (at 1 M ionic strength) and no salt conditions. k_{on} were obtained from irreversible association experiments (reproduced from Table 5.1). K_d were obtained from reversible, low-nM, association experiments and fixing k_{on} to the value obtained under irreversible conditions. Dissociation rate constants were obtained from $k_{off} = K_d \cdot k_{on}$. Errors represent standard error fo the mean, which were propageted in the case of k_{off}.

	$k_{on} \times 10^6$ / M^{-1} s^{-1}	$k_{off} \times 10^{-2}$/ s^{-1}	K_d / nM
No salt	145(±4)	0.7(±0.2)	0.05(±0.01)
KCl	12.1(±0.4)	0.82(±0.07)	0.68(±0.06)
NaCl	11.9(±0.2)	1.29(±0.07)	1.09(±0.06)
LiCl	6.8(±0.4)	1.4(±0.1)	2.0(±0.2)
$MgCl_2$	4.9(±0.3)	0.61(±0.06)	1.25(±0.09)
$CaCl_2$	4.5(±0.3)	1.6(±0.2)	3.5(±0.3)

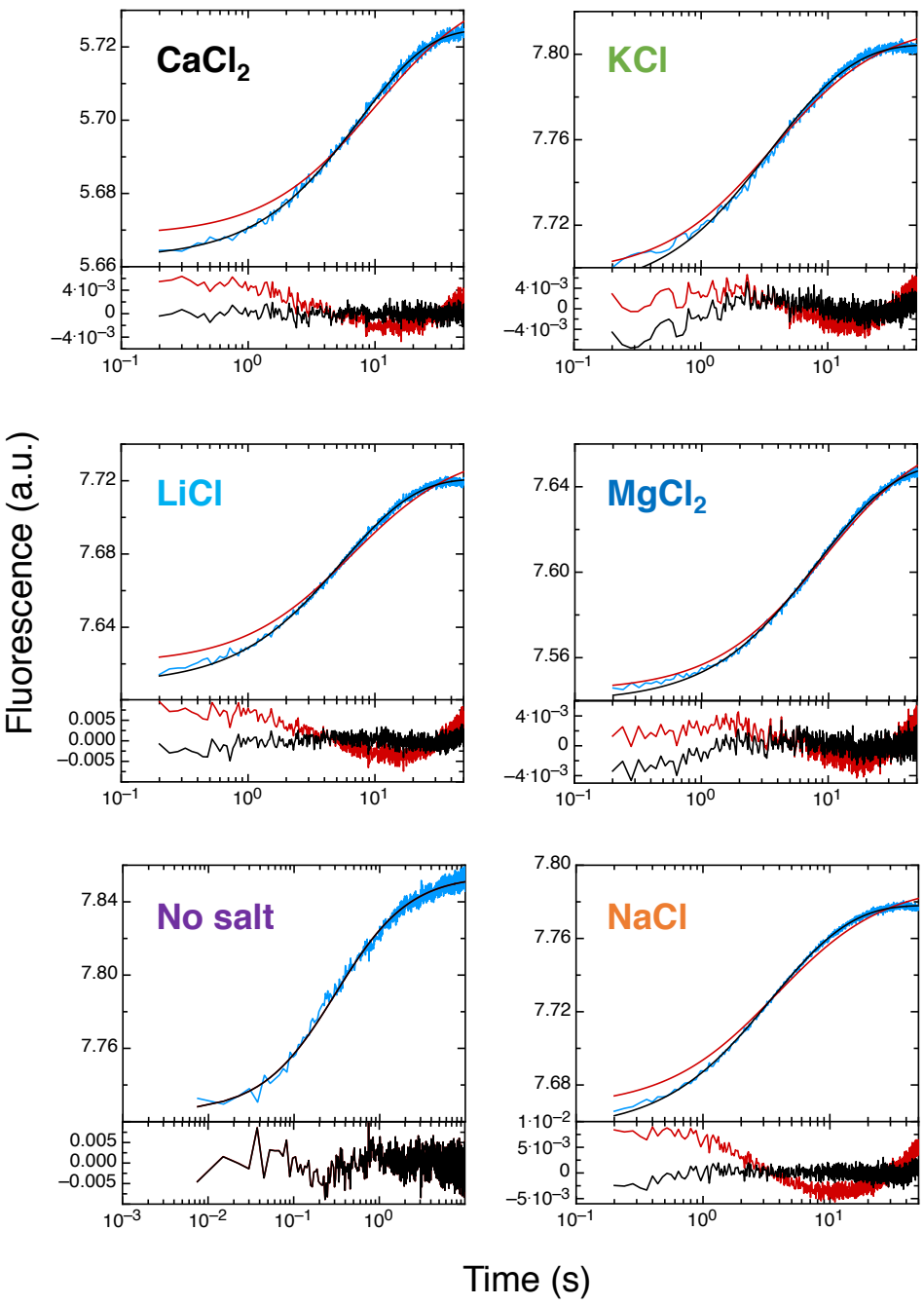

Figure C.1 – Binding of MCL-1 to PUMA under reversible conditions. Binding affinities of acetylamidated PUMA to MCL-1 were measured by performing low-nM (∼20 nM), near-equimolar, association experiments. At low concentrations, the dissociation reaction becomes significant and the data no longer fitted to an irreversible bimolecular binding model (Equation 2.49, red fits and residuals). Instead, the data was properly captured by a reversible bimolecular binding model (Equation 2.48, black fits and residuals). Fitting of the data to the reversible model (with k_{on} fixed to the value obtained under irrversible conditions) allowed K_d, and therefore k_{off} to be extracted. It is noted that the no salt condition fitted equally well to both models. This is consistent with a tighter affinity where the dissociation reaction is negligible. The experiments were performed for each salt at 1 M ionic strength and no salt conditions. Lower panels represent residuals to the fits, and the results are reported in Table C.1.

References

Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. (1994). Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature*, **370**, 621–628.

Ahnert, S. E., Marsh, J. A., Hernández, H., Robinson, C. V. & Teichmann, S. A. (2015). Principles of assembly reveal a periodic table of protein complexes. *Science*, **350**, aaa2245.

Al-Lazikani, B., Lesk, A. M. & Chothia, C. (1997). Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.*, **273**, 927–948.

Albeck, J. G., Burke, J. M., Spencer, S. L., Lauffenburger, D. A. & Sorger, P. K. (2008). Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLoS Biol.*, **6**, e299.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & And Walter, P. (1989). *Molecular biology of the cell*. Garland Publishing, New York, second edition.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.

Anfinsen, C. B., Haber, E., Sela, M. & White, F. H. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.*, **47**, 1309–1314.

Aouacheria, A., Rech de Laval, V., Combet, C. & Hardwick, J. M. (2013). Evolution of Bcl-2 homology motifs: Homology versus homoplasy. *Trends Cell Biol.*, **23**, 103–111.

Atkins, P. & Paula, J. D. (2010). *Atkins' physical chemistry*. Oxford University Press, ninth edition.

Babu, M. M., van der Lee, R., de Groot, N. S. & Gsponer, J. (2011). Intrinsically disordered proteins: Regulation and disease. *Curr. Opin. Struct. Biol.*, **21**, 432–440.

Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G. & Kiefhaber, T. (2011). Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 3952–3957.

Bah, A. & Forman-Kay, J. D. (2016). Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.*, **291**, 6696–6705.

Bah, A., Vernon, R. M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L. E. & Forman-Kay, J. D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature*, **519**, 106–109.

Baines, A. (2009). Evolution of spectrin function in cytoskeletal and membrane networks. *Biochem. Soc. Trans.*, **37**, 796–803.

Baldwin, A. J., Knowles, T. P., Tartaglia, G. G., Fitzpatrick, A. W., Devlin, G. L., Shammas, S. L., Waudby, C. A., Mossuto, M. F., Meehan, S., Gras, S. L., Christodoulou, J., Anthony-Cahill, S. J., Barker, P. D., Vendruscolo, M. & Dobson, C. M. (2011). Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.*, **133**, 14160–14163.

Baldwin, R. L. (1996). How Hofmeister ion interactions affect protein stability. *Biophys. J.*, **71**, 2056–2063.

Baxa, M. C., Haddadian, E. J., Jumper, J. M., Freed, K. F. & Sosnick, T. R. (2014). Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 15396–15401.

Benesch, J. L. & Robinson, C. V. (2006). Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Curr. Opin. Struct. Biol.*, **16**, 245–251.

Berg, O. G. (1985). Orientation constraints in diffusion-limited macromolecular association. The role of surface diffusion as a rate-enhancing mechanism. *Biophys. J.*, **47**, 1–14.

Berg, O. G. & von Hippel, P. H. (1985). Diffusion-controlled macromolecular interactions. *Annu. Rev. Biophys. Biophys. Chem.*, **14**, 131–160.

Bleicken, S., Classen, M., Padmavathi, P. V. L., Ishikawa, T., Zeth, K., Steinhoff, H. J. & Bordignon, E. (2010). Molecular details of Bax activation, oligomerization, and membrane insertion. *J. Biol. Chem.*, **285**, 6636–6647.

Bleicken, S., Hantusch, A., Das, K. K., Frickey, T. & Garcia-Saez, A. J. (2017). Quantitative interactome of a membrane Bcl-2 network identifies a hierarchy of complexes for apoptosis regulation. *Nat. Commun.*, **8**, 73.

Bleicken, S., Jeschke, G., Stegmueller, C., Salvador-Gallego, R., García-Sáez, A. J. & Bordignon, E. (2014). Structural model of active Bax at the membrane. *Mol. Cell*, **56**, 496–505.

Boeynaems, S., Alberti, S., Fawzi, N. L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., Tompa, P. & Fuxreiter, M. (2018). Protein phase separation: A new phase in cell biology. *Trends Cell Biol.*, **28**, 420–435.

Borcherds, W., Theillet, F. X., Katzer, A., Finzel, A., Mishall, K. M., Powell, A. T., Wu, H., Manieri, W., Dieterich, C., Selenko, P., Loewer, A. & Daughdrill, G. W. (2014). Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. *Nat. Chem. Biol.*, **10**, 1000–1002.

Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., Sottini, A., Soranno, A., Buholzer, K. J., Nettels, D., Kragelund, B. B., Best, R. B. & Schuler, B. (2018). Extreme disorder in an ultrahigh-affinity protein complex. *Nature*, **555**, 61–66.

Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F. & Hyman, A. A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, **5**, 1729–1732.

Brouwer, J. M., Lan, P., Cowan, A. D., Bernardini, J. P., Birkinshaw, R. W., van Delft, M. F., Sleebs, B. E., Robin, A. Y., Wardak, A., Tan, Iris, K., Reljic, B., Lee, E. F., Fairlie, D. W., Call, Melissa, J., Smith, B. J., Dewson, G., Lessene, G., Colman, P. M. & Czabotar, P. E. (2017). Conversion of Bim-BH3 from activator to inhibitor of Bak through structure-based design. *Mol. Cell*, **68**, 659–672.

Brouwer, J. M., Westphal, D., Dewson, G., Robin, A. Y., Uren, R. T., Bartolo, R., Thompson, G. V., Colman, P. M., Kluck, R. M. & Czabotar, P. E. (2014). Bak core and latch domains separate during activation, and freed core domains form symmetric homodimers. *Mol. Cell*, **55**, 938–946.

Brown, C. J., Johnson, A. K., Dunker, A. K. & Daughdrill, G. W. (2011). Evolution and disorder. *Curr. Opin. Struct. Biol.*, **21**, 441–446.

Bryngelson, J. D. D. & Wolynes, P. G. G. (1987). Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 7524–7528.

Cavalier-Smith, T. (2005). Economy, speed and size matter: Evolutionary forces driving nuclear genome miniaturization and expansion. *Ann. Bot.*, **95**, 147–175.

Chebaro, Y., Ballard, A. J., Chakraborty, D. & Wales, D. J. (2015). Intrinsically disordered energy landscapes. *Sci. Rep.*, **5**, 10386.

Chen, H.-C., Kanai, M., Inoue-Yamauchi, A., Tu, H.-C., Huang, Y., Ren, D., Kim, H., Takeda, S., Reyna, D. E., Chan, P. M., Ganesan, Y. T., Liao, C.-P., Gavathiotis, E., Hsieh, J. J. & Cheng, E. H. (2015). An interconnected hierarchical model of cell death regulation by the BCL-2 family. *Nat. Cell Biol.*, **17**, 1270–1281.

Chen, J. W., Romero, P., Uversky, V. N. & Dunker, A. K. (2006). Conservation of intrinsic disorder in protein domains and families: II. Functions of conserved disorder. *J. Proteome Res.*, **5**, 888–898.

Chen, L., Willis, S. N., Wei, A., Smith, B. J., Fletcher, J. I., Hinds, M. G., Colman, P. M., Day, C. L., Adams, J. M. & Huang, D. C. (2005). Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Mol. Cell*, **17**, 393–403.

Cheng, Y. & Patel, D. J. (2004). An efficient system for small protein expression and refolding. *Biochem. Biophys. Res. Commun.*, **317**, 401–405.

Chipuk, J. E. & Green, D. R. (2008). How do BCL-2 proteins induce mitochondrial outer membrane permeabilization? *Trends Cell Biol.*, **18**, 157–164.

Clapham, D. E. (2007). Calcium Signaling. *Cell*, **131**, 1047–1058.

Cory, S. & Adams, J. M. (2002). The BCL2 family: Regulators of the cellular life-or-death switch. *Nat. Rev. Cancer*, **2**, 647–656.

Cosentino, K. & García-Sáez, A. J. (2017). Bax and Bak pores: Are we closing the circle? *Trends Cell Biol.*, **27**, 266–275.

Cossio, P., Trovato, A., Pietrucci, F., Seno, F., Maritan, A. & Laio, A. (2010). Exploring the universe of protein structures beyond the protein data bank. *PLoS Comput. Biol.*, **6**, e1000957.

Crabtree, M. D., Borcherds, W., Poosapati, A., Shammas, S. L., Daughdrill, G. W. & Clarke, J. (2017). Conserved helix-flanking prolines modulate intrinsically disordered protein:target affinity by altering the lifetime of the bound complex. *Biochemistry*, **56**, 2379–2384.

Crabtree, M. D., Mendonça, C. A., Bubb, Q. R. & Clarke, J. (2018). Folding and binding pathways of BH3-only proteins are encoded within their intrinsically disordered sequence, not templated by partner proteins. *J. Biol. Chem.*, **293**, 9718–9723.

Czabotar, P. E., Lessene, G., Strasser, A. & Adams, J. M. (2014). Control of apoptosis by the BCL-2 protein family: Implications for physiology and therapy. *Nat. Rev. Mol. Cell Biol.*, **15**, 49–63.

Czabotar, P. E., Westphal, D., Dewson, G., Ma, S., Hockings, C., Fairlie, W. D., Lee, E. F., Yao, S., Robin, A. Y., Smith, B. J., Huang, D. C., Kluck, R. M., Adams, J. M. & Colman, P. M. (2013). Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. *Cell*, **152**, 519–531.

Daggett, V. & Fersht, A. R. (2003). Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.*, **28**, 18–25.

Dagil, R., Knudsen, M. J., Olsen, J. G., O'Shea, C., Franzmann, M., Goffin, V., Teilum, K., Breinholt, J. & Kragelund, B. B. (2012). The WSXWS motif in cytokine receptors is a molecular switch involved in receptor activation: Insight from structures of the prolactin receptor. *Structure*, **20**, 270–282.

Dahal, L., Kwan, T. O., Hollins, J. J. & Clarke, J. (2018). Promiscuous and selective: How intrinsically disordered BH3 proteins interact with their pro-survival partner MCL-1. *J. Mol. Biol.*, **430**, 2468–2477.

Dahal, L., Kwan, T. O., Shammas, S. L. & Clarke, J. (2017a). pKID Binds to KIX via an unstructured transition state with nonnative interactions. *Biophys. J.*, **113**, 2713–2722.

Dahal, L., Shammas, S. L. & Clarke, J. (2017b). Phosphorylation of the IDP KID modulates affinity for KIX by increasing the lifetime of the complex. *Biophys. J.*, **113**, 2706–2712.

Dai, H., Smith, A., Meng, X. W., Schneider, P. A., Pang, Y. P. & Kaufmann, S. H. (2011). Transient binding of an activator BH3 domain to the Bak BH3-binding groove initiates Bak oligomerization. *J. Cell Biol.*, **194**, 39–48.

Dandliker, W. B., Hsu, M. L., Levin, J. & Rao, B. R. (1981). Equilibrium and kinetic inhibition assays based upon fluorescence polarization. *Methods Enzymol.*, **74**, 3–28.

Darling, R. J., Kuchibhotla, U., Glaesner, W., Micanovic, R., Witcher, D. R. & Beals, J. M. (2002). Glycosylation of erythropoietin affects receptor binding kinetics: Role of electrostatic interactions. *Biochemistry*, **41**, 14524–14531.

Das, R. K. & Pappu, R. V. (2013). Conformations of intrinsically disordered proteins are infl uenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 13392–13397.

Day, C. L., Chen, L., Richardson, S. J., Harrison, P. J., Huang, D. C. S. & Hinds, M. G. (2005). Solution structure of prosurvival Mcl-1 and characterization of its binding by proapoptotic BH3-only ligands. *J. Biol. Chem.*, **280**, 4738–4744.

Day, C. L., Smits, C., Fan, F. C., Lee, E. F., Fairlie, W. D. & Hinds, M. G. (2008). Structure of the BH3 Domains from the p53-Inducible BH3-Only Proteins Noxa and Puma in Complex with Mcl-1. *J. Mol. Biol.*, **380**, 958–971.

Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods*, **8**, 1027–1040.

Dewson, G. & Kluck, R. M. (2009). Mechanisms by which Bak and Bax permeabilise mitochondria during apoptosis. *J. Cell Sci.*, **122**, 2801–2808.

Dewson, G., Kratina, T., Sim, H. W., Puthalakath, H., Adams, J. M., Colman, P. M. & Kluck, R. M. (2008). To trigger apoptosis, Bak exposes its BH3 domain and homodimerizes via BH3:groove interactions. *Mol. Cell*, **30**, 369–380.

Dewson, G., Ma, S., Frederick, P., Hockings, C., Tan, I., Kratina, T. & Kluck, R. M. (2012). Bax dimerizes via a symmetric BH3:groove interface during apoptosis. *Cell Death Differ.*, **19**, 661–670.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, **29**, 7133–7155.

Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. (2008). The Protein Folding Problem. *Annu. Rev. Biophys.*, **37**, 289–316.

Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, **426**, 884–890.

Dogan, J. & Jemth, P. (2014). Only kinetics can prove conformational selection. *Biophys. J.*, **107**, 1997–1998.

Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. (2015). Binding rate constants reveal distinct features of disordered protein domains. *Biochemistry*, **54**, 4741–4750.

Dogan, J., Mu, X., Engstrom, A. & Jemth, P. (2013). The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.*, **3**, 2076.

Dunker, A. K., Bondos, S. E., Huang, F. & Oldfield, C. J. (2015). Intrinsically disordered proteins and multicellular organisms. *Semin. Cell Dev. Biol.*, **37**, 44–55.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C. H., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C. & Obradovic, Z. (2001). Intrinsically disordered protein. *J. Mol. Graph. Model.*, **19**, 26–59.

Dyson, H. J. & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.*, **12**, 54–60.

Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, **6**, 197–208.

Edwards, Y. J., Lobley, A. E., Pentony, M. M. & Jones, D. T. (2009). Insights into the regulation of intrinsically disordered proteins in the human proteome by analyzing sequence and gene expression data. *Genome Biol.*, **10**, R50.

Eguchi, K. (2001). Apoptosis in autoimmune diseases. *Intern. Med.*, **40**, 275–284.

Eliezer, D. (2009). Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.*, **19**, 23–30.

Elmore, S. (2007). Apoptosis: A review of programmed cell death. *Toxicol. Pathol.*, **35**, 495–516.

English, L. R., Tischer, A., Demeler, A. K., Demeler, B. & Whitten, S. T. (2018). Sequence Reversal Prevents Chain Collapse and Yields Heat-Sensitive Intrinsic Disorder. *Biophys. J.*, **115**, 328–340.

Eskes, R., Desagher, S., Antonsson, B. & Martinou, J.-C. (2000). Bid induces the oligomerization and insertion of Bax into the outer mitochondrial membrane. *Mol. Cell. Biol.*, **20**, 929–935.

Fersht, A. (1999). *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*. Freeman, W. H.

Fersht, a. R. (1997). Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.*, **7**, 3–9.

Fishbain, S., Inobe, T., Israeli, E., Chavali, S., Yu, H., Kago, G., Babu, M. M. & Matouschek, A. (2015). Sequence composition of disordered regions fine-tunes protein half-life. *Nat. Struct. Mol. Biol.*, **22**, 214–221.

Flock, T., Weatheritt, R. J., Latysheva, N. S. & Babu, M. M. (2014). Controlling entropy to tune the functions of intrinsically disordered regions. *Curr. Opin. Struct. Biol.*, **26**, 62–72.

Fuxreiter, M. (2018). Fuzziness in protein interactions—A historical perspective. *J. Mol. Biol.*, **430**, 2278–2287.

Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.*, **338**, 1015–1026.

Gaetani, M., Mootien, S., Harper, S., Gallagher, P. G. & Speicher, D. W. (2008). Structural and functional effects of hereditary hemolytic anemia-associated point mutations in the alpha spectrin tetramer site. *Blood*, **111**, 5712–5720.

Garavito, R. M. & Ferguson-Miller, S. (2001). Detergents as tools in membrane biochemistry. *J. Biol. Chem.*, **276**, 32403–32406.

Garcia-Seisdedos, H., Empereur-Mot, C., Elad, N. & Levy, E. D. (2017). Proteins evolve on the edge of supramolecular self-assembly. *Nature*, **548**, 244–247.

Garner, T. P., Reyna, D. E., Priyadarshi, A., Chen, H.-C., Li, S., Wu, Y., Ganesan, Y. T., Malashkevich, V. N., Almo, S. S., Cheng, E. H. & Gavathiotis, E. (2016). An autoinhibited dimeric form of BAX regulates the BAX activation pathway. *Mol. Cell*, **63**, 485–497.

Gavathiotis, E., Reyna, D. E., Davis, M. L., Bird, G. H. & Walensky, L. D. (2010). BH3-triggered structural reorganization drives the activation of proapoptotic BAX. *Mol. Cell*, **40**, 481–492.

Gavathiotis, E., Suzuki, M., Davis, M. L., Pitter, K., Bird, G. H., Katz, S. G., Tu, H. C., Kim, H., Cheng, E. H., Tjandra, N. & Walensky, L. D. (2008). BAX activation is initiated at a novel interaction site. *Nature*, **455**, 1076–1081.

Gazit, E. (2002). The 'correctly folded' state of proteins: Is it a metastable state? *Angew. Chemie – Int. Ed.*, **41**, 257–259.

Gianni, S., Dogan, J. & Jemth, P. (2014). Distinguishing induced fit from conformational selection. *Biophys. Chem.*, **189**, 33–39.

Gianni, S., Dogan, J. & Jemth, P. (2016). Coupled binding and folding of intrinsically disordered proteins: What can we learn from kinetics? *Curr. Opin. Struct. Biol.*, **36**, 18–24.

Gibbs, E. B. & Showalter, S. A. (2015). Quantitative biophysical characterization of intrinsically disordered proteins. *Biochemistry*, **54**, 1314–1326.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.

Girault, H. H. (2007). *Electrochimie physique et analytique*. Presses Polytechniques et Universitaires Romandes.

Giri, R., Morrone, A., Toto, A., Brunori, M. & Gianni, S. (2013). Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 14942–14947.

Goldberg, J. M. & Baldwin, R. L. (1998). Kinetic mechanism of a partial folding reaction. 1. Properties of the reaction and effects of denaturants. *Biochemistry*, **37**, 2546–2555.

Goldstein, J. C., Waterhouse, N. J., Juin, P., Evan, G. I. & Green, D. R. (2000). The coordinate release of cytochrome c during apoptosis is rapid, complete and kinetically invariant. *Nat. Cell Biol.*, **2**, 156–162.

Goodsell, D. S. (1991). Inside a living cell. *Trends Biochem. Sci.*, **16**, 203–206.

Goodsell, D. S. & Olson, A. J. (2000). Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 105–153.

Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.*, **1**, 2876–2890.

Greives, N. & Zhou, H.-X. (2014). Both protein dynamics and ligand concentration can shift the binding mechanism between conformational selection and induced fit. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 10197–10202.

Griffin, M. D. & Gerrard, J. A. (2012). *The relationship between oligomeric state and protein function*. Springer, New York, NY.

Gsponer, J., Futschik, M. E., Teichmann, S. A. & Babu, M. M. (2008). Tight regulation of unstructured proteins: From transcript synthesis to protein degradation. *Science*, **322**, 1365–1368.

Gupta, K., Donlan, J. A., Hopper, J. T., Uzdavinys, P., Landreh, M., Struwe, W. B., Drew, D., Baldwin, A. J., Stansfeld, P. J. & Robinson, C. V. (2017). The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature*, **541**, 421–424.

Gurau, M. C., Lim, S. M., Castellana, E. T., Albertorio, F., Kataoka, S. & Cremer, P. S. (2004). On the mechanism of the Hofmeister effect. *J. Am. Chem. Soc.*, **126**, 10522–10523.

Hammes, G. G., Chang, Y.-C. & Oas, T. G. (2009). Conformational selection or induced fit: A flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 13737–13741.

Hanahn, D. & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, **144**, 646–674.

Harmon, T. S., Crabtree, M. D., Shammas, S. L., Posey, A. E., Clarke, J. & Pappu, R. V. (2016). GADIS: Algorithm for designing sequences to achieve target secondary structure profiles of intrinsically disordered proteins. *Protein Eng. Des. Sel.*, **29**, 339–346.

Harms, M. J. & Thornton, J. W. (2010). Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.*, **20**, 360–366.

Haynes, C., Oldfield, C. J., Ji, F., Klitgord, N., Cusick, M. E., Radivojac, P., Uversky, V. N., Vidal, M. & Iakoucheva, L. M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.*, **2**, e100.

Hill, S. A., Kwa, L. G., Shammas, S. L., Lee, J. C. & Clarke, J. (2014). Mechanism of assembly of the non-covalent spectrin tetramerization domain from intrinsically disordered partners. *J. Mol. Biol.*, **426**, 21–35.

Hofmann, H., Nettels, D., Schuler, B., Hofmann, H., Nettels, D. & Schuler, B. (2013). Single-molecule spectroscopy of the unexpected collapse of an unfolded protein at low pH Single-molecule spectroscopy of the unexpected collapse of an unfolded protein at low pH. *J. Chem. Phys.*, **139**, 121930.

Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D. & Schuler, B. (2012). Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 16155–16160.

Hofmeister, F. (1888). Zur Lehre von der Wirkung der Salze - Zweite Mittheilung. *Arch. für Exp. Pathol. und Pharmakologie*, **24**, 247–260.

Hsu, Y.-T. & Youle, R. J. (1997). Nonionic detergents induce dimerization among members of the Bcl-2 family. *J. Biol. Chem.*, **272**, 13829–13834.

Huang, P.-S., Oberdorfer, G., Xu, C., Pei, X. Y., Nannenga, B. L., Rogers, J. M., DiMaio, F. P., Gonen, T., Luisi, B. & Baker, D. (2014). High thermodynamic stability of parametrically designed helical bundles. *Science*, **346**, 481–485.

Huang, Y. & Liu, Z. (2009). Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: A critical assessment of the 'fly-casting' mechanism. *J. Mol. Biol.*, **393**, 1143–1159.

Iešmantavičius, V., Dogan, J., Jemth, P., Teilum, K. & Kjaergaard, M. (2014). Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chemie – Int. Ed.*, **53**, 1548–1551.

Ipsaro, J. J., Harper, S. L., Messick, T. E., Marmorstein, R., Mondragón, A. & Speicher, D. W. (2010). Crystal structure and functional interpretation of the erythocyte spectrin tetramerization domain complex. *Blood*, **115**, 4843–4852.

Itzhaki, L. S., Otzen, D. E. & Fersht, a. R. (1995). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.*, **254**, 260–288.

Iyer, S., Anwari, K., Alsop, A. E., Yuen, W. S., Huang, D. C., Carroll, J., Smith, N. A., Smith, B. J., Dewson, G. & Kluck, R. M. (2016). Identification of an activation site in Bak and mitochondrial Bax triggered by antibodies. *Nat. Commun.*, **7**, 11734.

Jemth, P., Mu, X., Engström, Å. & Dogan, J. (2014). A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.*, **289**, 5528–5533.

Jensen, M. R., Zweckstetter, M., Huang, J.-r. & Blackledge, M. (2014). Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.*, **114**, 6632–6660.

Kale, J., Osterlund, E. J. & Andrews, D. W. (2017). BCL-2 family proteins: changing partners in the dance towards death. *Cell Death Differ.*, **25**, 65–80.

Kelly, S. M., Jess, T. J. & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics*, **1751**, 119–139.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*, **181**, 662–666.

Kerr, J. F., Wyllie, A. H. & Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br. J. Cancer*, **26**, 239–257.

Kestin, J., Sokolov, M. & Wakeham, W. A. (1978). Viscosity of liquid water in the range −8 °C to 150 °C.

Kiefhaber, T., Bachmann, A. & Jensen, K. S. (2012). Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr. Opin. Struct. Biol.*, **22**, 21–29.

Kluck, R. M., Bossy-Wetzle, E., Green, D. R. & Newmeyer, D. D. (1997). The release of cytochrome c from mitochondria: A primary site for Bcl-2 regulation of apoptosis. *Science*, **275**, 1132–1136.

Knowles, T. P., Vendruscolo, M. & Dobson, C. M. (2014). The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.*, **15**, 384–396.

Kobashi, K. (1968). Catalytic oxidation of sulfhydryl groups by *o*-phenanthroline copper complex. *Biochim. Biophys. Acta*, **158**, 239–245.

Koculi, E., Hyeon, C., Thirumalai, D. & Woodson, S. A. (2007). Charge density of divalent metal cations determines RNA stability. *J. Am. Chem. Soc.*, **129**, 2676–2682.

Kostanski, L. K., Keller, D. M. & Hamielec, A. E. (2004). Size-exclusion chromatography – A review of calibration methodologies. *J. Biochem. Biophys. Methods*, **58**, 159–186.

Krieger, J. M., Fusco, G., Lewitzky, M., Simister, P. C., Marchant, J., Camilloni, C., Feller, S. M. & De Simone, A. (2014). Conformational recognition of an intrinsically disordered protein. *Biophys. J.*, **106**, 1771–1779.

Ku, B., Liang, C., Jung, J. U. & Oh, B. H. (2011). Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. *Cell Res.*, **21**, 627–641.

Kvansakul, M. & Hinds, M. G. (2013). Structural biology of the Bcl-2 family and its mimicry by viral proteins. *Cell Death Dis.*, **4**, e909.

Lacroix, E., Viguera, A. R. & Serrano, L. (1998a). Elucidating the folding problem of α-helices: Local motifs, long-range electrostatics, lonic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.*, **284**, 173–191.

Lacroix, E., Viguera, a. R. & Serrano, L. (1998b). Reading protein sequences backwards. *Fold. Des.*, **3**, 79–85.

Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J. & Robinson, C. V. (2014). Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*, **510**, 172–175.

Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. (2013). Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.*, **8**, 639–651.

Lakowicz, J. R. (2006). *Principles of fluorescence spectroscopy*. Springer, New York, NY, third edition.

Leber, B., Lin, J. & Andrews, D. W. (2007). Embedded together: The life and death consequences of interaction of the Bcl-2 family with membranes. *Apoptosis*, **12**, 897–911.

Lee, E. F., Czabotar, P. E., Yang, H., Sleebs, B. E., Lessene, G., Colman, P. M., Smith, B. J. & Fairlie, W. D. (2009). Conformational changes in Bcl-2 pro-survival proteins determine their capacity to bind ligands. *J. Biol. Chem.*, **284**, 30508–30517.

Lee, E. F., Dewson, G., Smith, B. J., Evangelista, M., Pettikiriarachchi, A., Dogovski, C., Perugini, M. A., Colman, P. M. & Fairlie, W. D. (2011). Crystal structure of a BCL-W domain-swapped dimer: Implications for the function of BCL-2 family proteins. *Structure*, **19**, 1467–1476.

Leshchiner, E. S., Braun, C. R., Bird, G. H. & Walensky, L. D. (2013). Direct activation of full-length proapoptotic BAK. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, E986–E995.

Levinthal, C. (1969). How to fold graciously. *Mössbauer Spectrosc. Biol. Syst. Proc.*, **67**, 22–24.

Levy, E. D., Erba, E. B., Robinson, C. V. & Teichmann, S. A. (2008). Assembly reflects evolution of protein complexes. *Nature*, **453**, 1262–1265.

Levy, E. D., Pereira-Leal, J. B., Chothia, C. & Teichmann, S. A. (2006). 3D complex: A structural classification of protein complexes. *PLoS Comput. Biol.*, **2**, 1395–1406.

Li, M. X., Tan, I. K. L., Ma, S. B., Hockings, C., Kratina, T., Dengler, M. A., Alsop, A. E., Kluck, R. M. & Dewson, G. (2017). BAK α6 permits activation by BH3-only proteins and homooligomerization via the canonical hydrophobic groove. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 7629–7634.

Lise, S. & Jones, D. T. (2005). Sequence patterns associated with disordered regions in proteins. *Proteins Struct. Funct. Genet.*, **58**, 144–150.

Liu, Q. & Gehring, K. (2010). Heterodimerization of BAK and MCL-1 activated by detergent micelles. *J. Biol. Chem.*, **285**, 41202–41210.

Liu, Z. & Huang, Y. (2014). Advantages of proteins being disordered. *Protein Sci.*, **23**, 539–550.

Llambi, F., Moldoveanu, T., Tait, S. W., Bouchier-Hayes, L., Temirov, J., McCormick, L. L., Dillon, C. P. & Green, D. R. (2011). A unified model of mammalian BCL-2 protein family interactions at the mitochondria. *Mol. Cell*, **44**, 517–531.

Lo Nostro, P. & Ninham, B. W. (2012). Hofmeister phenomena: an update on ion specificity in biology. *Chem. Rev.*, **112**, 2286–322.

Lupas, A. N., Ponting, C. P. & Russell, R. B. (2001). On the evolution of protein folds: Are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world? *J. Struct. Biol.*, **134**, 191–203.

Makhatadze, G. I. & Privalov, P. L. (1996). On the entropy of protein folding. *Protein Sci.*, **5**, 507–510.

Malatesta, F. (2005). The study of bimolecular reactions under non-pseudo-first order conditions. *Biophys. Chem.*, **116**, 251–256.

Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L. & Pappu, R. V. (2010). Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 8183–8188.

Marianayagam, N. J., Sunde, M. & Matthews, J. M. (2004). The power of two: Protein dimerization in biology. *Trends Biochem. Sci.*, **29**, 618–625.

Marsh, J. A. & Teichmann, S. A. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. *Annu. Rev. Biochem.*, **84**, 551–575.

Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature*, **340**, 122–126.

McArthur, K., Whitehead, L. W., Heddleston, J. M., Li, L., Padman, B. S., Oorschot, V., Geoghegan, N. D., Chappaz, S., Davidson, S., Chin, H. S., Lane, R. M., Dramicanin, M., Saunders, T. L., Sugiana, C., Lessene, R., Osellame, L. D., Chew, T. L., Dewson, G., Lazarou, M., Ramm, G., Lessene, G., Ryan, M. T., Rogers, K. L., Van Delft, M. F. & Kile, B. T. (2018). BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science*, **359**, eaao6047.

Mehboob, S., Jacob, J., May, M., Kotula, L., Thiyagarajan, P., Johnson, M. E. & Fung, L. W. M. (2003). Structural analysis of the αN-terminal region of erythroid and nonerythroid spectrins by small-angle X-ray scattering. *Biochemistry*, **42**, 14702–14710.

Mei, H., Wang, K., McKee, S., Wang, X., Waldner, J. L., Pielak, G. J., Durham, B. & Millett, F. (1996). Control of formation and dissociation of the high-affinity complex between cytochrome c and cytochrome c peroxidase by ionic strength and the low-affinity binding site. *Biochemistry*, **35**, 15800–15806.

Milles, S., Mercadante, D., Aramburu, I. V., Jensen, M. R., Banterle, N., Koehler, C., Tyagi, S., Clarke, J., Shammas, S. L., Blackledge, M., Gräter, F. & Lemke, E. A. (2015). Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors. *Cell*, **163**, 734–745.

Mirny, L. A. & Shakhnovich, E. I. (1999). Universally conserved positions in protein folds: Reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.*, **291**, 177–196.

Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.*, **260**, 289–298.

Moldoveanu, T., Grace, C. R., Llambi, F., Nourse, A., Fitzgerald, P., Gehring, K., Kriwacki, R. W. & Green, D. R. (2013). BID-induced structural changes in BAK promote apoptosis. *Nat. Struct. Mol. Biol.*, **20**, 589–597.

Moldoveanu, T., Liu, Q., Tocilj, A., Watson, M., Shore, G. & Gehring, K. (2006). The X-Ray Structure of a BAK Homodimer Reveals an Inhibitory Zinc Binding Site. *Mol. Cell*, **24**, 677–688.

Müller-Späth, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L., Nettels, D. & Schuler, B. (2010). Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 14609–14614.

Muñoz, V. & Serrano, L. (1995). Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J. Mol. Biol.*, **245**, 297–308.

Nooren, I. M. A. & Thornton, J. M. (2003). Diversity of protein-protein interactions. *EMBO J.*, **22**, 3486–3492.

Nott, T. J., Craggs, T. D. & Baldwin, A. J. (2016). Membraneless organelles can melt nucleic acid duplexes and act as biomolecular filters. *Nat. Chem.*, **8**, 569–575.

Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D. & Baldwin, A. J. (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell*, **57**, 936–947.

Nozaki, Y. (1972). The preparation of guanidine hydrochloride. *Methods Enzymol.*, **26**, 43–50.

Ohi, M., Li, Y., Cheng, Y. & Walz, T. (2004). Negative staining and image classification—powerful tools in modern electron microscopy. *Biol. Proced. Online*, **6**, 23–34.

Okur, H. I., Hladílková, J., Rembert, K. B., Cho, Y., Heyda, J., Dzubiella, J., Cremer, P. S. & Jungwirth, P. (2017). Beyond the Hofmeister series: Ion-specific effects on proteins and their biological functions. *J. Phys. Chem. B*, **121**, 1997–2014.

Oliveberg, M., Tan, Y. J. & Fersht, A. R. (1995). Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 8926–8929.

O'Neill, K. L., Huang, K., Zhang, J., Chen, Y. & Luo, X. (2016). Inactivation of prosurvival Bcl-2 proteins activates Bax / Bak through the outer mitochondrial membrane. *Genes Dev.*, **30**, 973–988.

Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. (1997). Theory of protein folding: The energy landscape perspective. *Annu. Rev. Phys. Chem.*, pp. 545–600.

Onuchic, J. N. & Wolynes, P. G. (2004). Theory of protein folding. *Curr. Opin. Struct. Biol.*, **14**, 70–75.

Ou, L., Matthews, M., Pang, X. & Zhou, H. X. (2017). The dock-and-coalesce mechanism for the association of a WASP disordered region with the Cdc42 GTPase. *FEBS J.*, **284**, 3381–3391.

Pace, C. N. (1975). The stability of globular proteins. *CRC Crit. Rev. Biochem.*, **3**, 1–43.

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.*, **131**, 266–280.

Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.*, **263**, 11820–11825.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.*, **4**, 2411–2423.

Pang, X., Qin, S. & Zhou, H. X. (2011). Rationalizing 5000-fold differences in receptor-binding rate constants of four cytokines. *Biophys. J.*, **101**, 1175–1183.

Papadakos, G., Sharma, A., Lancaster, L. E., Bowen, R., Kaminska, R., Leech, A. P., Walker, D., Redfield, C. & Kleanthous, C. (2015). Consequences of inducing intrinsic disorder in a high-affinity protein-protein interaction. *J. Am. Chem. Soc.*, **137**, 5252–5255.

Pauling, L. (1946). Molecular architecture and biological reactions. *Chem. Eng. News*, **24**, 1375–1377.

Pawson, T. & Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes Dev.*, pp. 1027–1047.

Peña-Blanco, A. & García-Sáez, A. J. (2018). Bax, Bak and beyond – mitochondrial performance in apoptosis. *FEBS J.*, **285**, 416–431.

Peng, Z., Mizianty, M. J. & Kurgan, L. (2014). Genome-scale prediction of proteins with long intrinsically disordered regions. *Proteins Struct. Funct. Bioinforma.*, **82**, 145–158.

Perham, R. N. (1975). Self-assembly of biological macromolecules. *Philos. Trans. R. Soc. B*, **272**, 123–136.

Petros, A. M., Olejniczak, E. T. & Fesik, S. W. (2004). Structural biology of the Bcl-2 family of proteins. *Biochim. Biophys. Acta - Mol. Cell Res.*, **1644**, 83–94.

Polster, B. M., Basañez, G., Young, M., Suzuki, M. & Fiskum, G. (2003). Inhibition of Bax-induced cytochrome c release from neural cell and brain mitochondria by dibucaine and propranolol. *J. Neuroscience*, **23**, 2735–2743.

Privalov, P. L. (1979). Stability of proteins: small globular proteins. *Adv. Protein Chem.*, **33**, 167–241.

Privalov, P. L. & Makhatadze, G. I. (1993). Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration.

Proskuryakov, S. Y., Konoplyannikov, A. G. & Gabai, V. L. (2003). Necrosis: A specific form of programmed cell death? *Exp. Cell Res.*, **283**, 1–16.

Rajasekaran, N., Gopi, S., Narayan, A. & Naganathan, A. N. (2016). Quantifying protein disorder through measures of excess conformational entropy. *J. Phys. Chem. B*, **120**, 4341–4350.

Rautureau, G. J., Day, C. L. & Hinds, M. G. (2010). Intrinsically disordered proteins in Bcl-2 regulated apoptosis. *Int. J. Mol. Sci.*, **11**, 1808–1824.

Rehm, M., Düßmann, H., Jänicke, R. U., Tavaré, J. M., Kögel, D. & Prehn, J. H. M. (2002). Single-cell fluorescence resonance energy transfer analysis demonstrates that caspase activation during apoptosis is a rapid process. *J. Biol. Chem.*, **277**, 24506–24514.

Rogers, J. M., Oleinikovas, V., Shammas, S. L., Wong, C. T., De Sancho, D., Baker, C. M. & Clarke, J. (2014a). Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 15420–15425.

Rogers, J. M., Steward, A. & Clarke, J. (2013). Folding and binding of an intrinsically disordered protein: Fast, but not 'diffusion-limited'. *J. Am. Chem. Soc.*, **135**, 1415–1422.

Rogers, J. M., Wong, C. T. & Clarke, J. (2014b). Coupled folding and binding of the disordered protein PUMA does not require particular residual structure. *J. Am. Chem. Soc.*, **136**, 5197–5200.

Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J. & Dunker, A. K. (2001). Sequence complexity of disordered protein. *Proteins Struct. Funct. Genet.*, **42**, 38–48.

Romero, P. R., Zaidi, S., Fang, Y. Y., Uversky, V. N., Radivojac, P., Oldfield, C. J., Cortese, M. S., Sickmeier, M., LeGall, T., Obradovic, Z. & Dunker, A. K. (2006). Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 8390–8395.

Rutherfurd, S. M. & Gilani, G. S. (2009). Amino acid analysis. *Curr. Protoc. Protein Sci.*, **58**, 11.9.1–11.9.37.

Salvador-Gallego, R., Mund, M., Cosentino, K., Schneider, J., Unsay, J., Schraermeyer, U., Engelhardt, J., Ries, J. & Garcia-Saez, A. J. (2016). Bax assembly into rings and arcs in apoptotic mitochondria is linked to membrane pores. *EMBO J.*, **35**, 389–401.

Sattler, M., Liang, H., Nettesheim, D., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H. S., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B. & Fesik, S. W. (1997). Structure of Bcl-xL–Bak peptide complex: Recognition between regulators of apoptosis. *Science*, **275**, 983–986.

Schlessinger, A., Schaefer, C., Vicedo, E., Schmidberger, M., Punta, M. & Rost, B. (2011). Protein disorder – a breakthrough invention of evolution? *Curr. Opin. Struct. Biol.*, **21**, 412–418.

Schreiber, G. & Fersht, A. R. (1996). Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.*, **3**, 427–431.

Schreiber, G., Haran, G. & Zhou, H.-X. (2009). Fundamental aspects of protein-protein association kinetics. *Chem. Rev.*, **109**, 839–860.

Seddon, A. M., Curnow, P. & Booth, P. J. (2004). Membrane proteins, lipids and detergents: Not just a soap opera. *Biochim. Biophys. Acta - Biomembr.*, **1666**, 105–117.

Shamas-Din, A., Kale, J., Leber, B. & Andrews, D. W. (2013). Mechanisms of action of Bcl-2 family proteins. *Cold Spring Harb. Perspect. Biol.*, **5**, a008714.

Shammas, S., Crabtree, M. D., Dahal, L., Wicky, B. I. M. & Clarke, J. (2016). Insights into coupled folding and binding mechanisms from kinetic studies. *J. Biol. Chem.*, **291**, 6689–6695.

Shammas, S. L. (2017). Mechanistic roles of protein disorder within transcription. *Curr. Opin. Struct. Biol.*, **42**, 155–161.

Shammas, S. L., Rogers, J. M., Hill, S. A. & Clarke, J. (2012). Slow, reversible, coupled folding and binding of the spectrin tetramerization domain. *Biophys. J.*, **103**, 2203–2214.

Shammas, S. L., Travis, A. J. & Clarke, J. (2013). Remarkably fast coupled folding and binding of the intrinsically disordered transactivation domain of cMyb to CBP KIX. *J. Phys. Chem. B*, **117**, 13346–13356.

Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. (2000). Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 8868–8873.

Singer, S. J. & Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science*, **175**, 720–731.

Smock, R. G., Yadid, I., Dym, O., Clarke, J. & Tawfik, D. S. (2016). De novo evolutionary emergence of a symmetrical protein is shaped by folding constraints. *Cell*, **164**, 476–486.

Soranno, A., Buchli, B., Nettels, D., Cheng, R. R., Muller-Spath, S., Pfeil, S. H., Hoffmann, A., Lipman, E. A., Makarov, D. E. & Schuler, B. (2012). Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 17800–17806.

Soranno, A., Koenig, I., Borgia, M. B., Hofmann, H., Zosel, F., Nettels, D. & Schuler, B. (2014). Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 4874–4879.

Southall, N. T., Dill, K. A. & Haymet, A. D. (2002). A view of the hydrophobic effect. *J. Phys. Chem. B*, **106**, 521–533.

Speicher, D. W., DeSilva, T. M., Speicher, K. D., Ursitti, J. A., Hembach, P. & Weglarz, L. (1993). Location of the human red cell spectrin tetramer binding site and detection of a related 'closed' hairpin loop dimer using proteolytic footprinting. *J. Biol. Chem.*, **268**, 4227–4235.

Speicher, D. W. & Marchesi, V. T. (1984). Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature*, **311**, 177–180.

Stellwagen, E., Prantner, J. D. & Stellwagen, N. C. (2008). Do zwitterions contribute to the ionic strength of a solution? *Anal. Biochem.*, **373**, 407–409.

Subburaj, Y., Cosentino, K., Axmann, M., Pedrueza-Villalmanzo, E., Hermann, E., Bleicken, S., Spatz, J. & García-Sáez, A. J. (2015). Bax monomers form dimer units in the membrane that further self-assemble into multiple oligomeric species. *Nat. Commun.*, **6**, 8042.

Suzuki, M., Youle, R. J. & Tjandra, N. (2000). Structure of Bax: Coregulation of dimer formation and intracellular localization. *Cell*, **103**, 645–654.

Tanford, C. (1968). Protein Denaturation. *Adv. Protein Chem.*, **23**, 121–282.

Tanford, C. (1970). Protein denaturation: Part C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.*, **24**, 1–95.

Tanford, C. & Reynolds, J. (2001). *Nature's robots: A history of proteins*. Oxford University Press.

Teilum, K., Olsen, J. G. & Kragelund, B. B. (2015). Globular and disordered—the non-identical twins in protein-protein interactions. *Front. Mol. Biosci.*, **2**, 40.

Theillet, F.-X., Kalmar, L., Tompa, P., Han, K.-H., Selenko, P., Dunker, A. K., Daughdrill, G. W. & Uversky, V. N. (2013). The alphabet of intrinsic disorder. *Intrinsically Disord. Proteins*, **1**, e24684.

Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L. & Tawfik, D. S. (2007). The stability effects of protein mutations appear to be universally distributed. *J. Mol. Biol.*, **369**, 1318–1332.

Tompa, P. (2011). Unstructural biology coming of age. *Curr. Opin. Struct. Biol.*, **21**, 419–425.

Tompa, P., Davey, N. E., Gibson, T. J. & Babu, M. M. (2014). A million peptide motifs for the molecular biologist. *Mol. Cell*, **55**, 161–169.

Tompa, P. & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.*, **33**, 2–8.

Toto, A. & Gianni, S. (2016). Mutational analysis of the binding-induced folding reaction of the mixed-lineage leukemia protein to the KIX domain. *Biochemistry*, **55**, 3957–3962.

Toto, A., Giri, R., Brunori, M. & Gianni, S. (2014). The mechanism of binding of the KIX domain to the mixed lineage leukemia protein and its allosteric role in the recognition of c-Myb. *Protein Sci.*, **23**, 962–969.

Trinick, J., Knight, P. & Whiting, A. (1984). Purification and properties of native titin. *J. Mol. Biol.*, **180**, 331–356.

Uren, R. T., O'Hely, M., Iyer, S., Bartolo, R., Shi, M. X., Brouwer, J. M., Alsop, A. E., Dewson, G. & Kluck, R. M. (2017). Disordered clusters of bak dimers rupture mitochondria during apoptosis. *Elife*, **6**, e19944.

Uversky, V. N. & Dunker, A. K. (2010). Understanding protein non-folding. *Biochim. Biophys. Acta – Proteins Proteomics*, **1804**, 1231–1264.

Uversky, V. N., Gillespie, J. R. & Fink, A. L. (2000). Why are natively unfolded' proteins unstructured under physiologic conditions? *Proteins Struct. Funct. Genet.*, **41**, 415–427.

Uversky, V. N., Oldfield, C. J. & Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: Introducing the D^2 concept. *Annu. Rev. Biophys.*, **37**, 215–246.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Old, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E. & Babu, M. M. (2014a). Classification of intrinsically disordered regions and proteins. *Chem. Rev.*, **114**, 6589–6631.

van der Lee, R., Lang, B., Kruse, K., Gsponer, J., de Groot, N. S., Huynen, M. A., Matouschek, A., Fuxreiter, M. & Babu, M. M. (2014b). Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep.*, **8**, 1832–1844.

van der Spoel, D., van Maaren, P. J., Larsson, P. & Tîmneanu, N. (2006). Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B*, **110**, 4393–4398.

Vijayakumar, M., Wong, K. Y., Schreiber, G., Fersht, A. R., Szabo, A. & Zhou, H. X. (1998). Electrostatic enhancement of diffusion-controlled protein-protein association: Comparison of theory and experiment on barnase and barstar. *J. Mol. Biol.*, **278**, 1015–1024.

Walker, J. E., Gay, N. J., Powell, S. J., Kostina, M. & Dyer, M. R. (1987). ATP synthase from bovine mitochondria: Sequences of imported precursors of oligomycin sensitivity conferral protein, factor 6, and adenosinetriphosphatase inhibitor protein. *Biochemistry*, **26**, 8613–8619.

Wallis, R., Moore, G. R., James, R. & Kleanthous, C. (1995). Protein-protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. *Biochemistry*, **34**, 13743–13750.

Wang, H., Takemoto, C., Akasaka, R., Uchikubo-Kamo, T., Kishishita, S., Murayama, K., Terada, T., Chen, L., Liu, Z. J., Wang, B. C., Sugano, S., Tanaka, A., Inoue, M., Kigawa, T., Shirouzu, M. & Yokoyama, S. (2009). Novel dimerization mode of the human Bcl-2 family protein Bak, a mitochondrial apoptosis regulator. *J. Struct. Biol.*, **166**, 32–37.

Wang, K., McClure, J. & Tu, A. (1979). Titin: major myofibrillar components of striated muscle. *Proc. Natl. Acad. Sci. U. S. A.*, **76**, 3698–3702.

Wang, Z., Mottonen, J. & Goldsmith, E. J. (1996). Kinetically controlled folding of the serpin plasminogen activator inhibitor 1. *Biochemistry*, **35**, 16443–16448.

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three Kingdoms of life. *J. Mol. Biol.*, **337**, 635–645.

Warren, J. R. & Gordon, J. A. (1966). On the refractive indices of aqueous solutions of urea. *J. Phys. Chem.*, **70**, 297–300.

Watson, J. D. & Crick, F. H. C. (1953). Molecular structure of nucleic acids. *Nature*, **171**, 737–738.

Weathers, E. A., Paulaitis, M. E., Woolf, T. B. & Hoh, J. H. (2004). Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett.*, **576**, 348–352.

Wei, M. C., Lindsten, T., Mootha, V. K., Weiler, S., Gross, A., Ashiya, M., Thompson, C. B. & Korsmeyer, S. J. (2000). tBID, a membrane-targeted death ligand, oligomerizes BAK to release cytochrome c. *Genes Dev.*, **14**, 2060–2071.

Williams, P. D., Pollock, D. D. & Goldstein, R. a. (2006). Functionality and the evolution of marginal stability in proteins: inferences from lattice simulations. *Evol. Bioinform. Online*, **2**, 91–101.

Willis, S. N., Chen, L., Dewson, G., Wei, A., Naik, E., Fletcher, J. I., Adams, J. M. & Huang, D. C. S. (2005). Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. *Genes Dev.*, **19**, 1294–1305.

Willis, S. N., Fletcher, J. I., Kaufmann, T., Van Delft, M. F., Chen, L., Czabotar, P. E., Ierino, H., Lee, E. F., Fairlie, W. D., Bouillet, P., Strasser, A., Kluck, R. M., Adams, J. M. & Huang, D. C. (2007). Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science*, **315**, 856–859.

Wolfenden, R. & Snider, M. J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.*, **34**, 938–945.

Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, **293**, 321–331.

Wright, P. E. & Dyson, H. J. (2009). Linking folding and binding. *Curr. Opin. Struct. Biol.*, **19**, 31–38.

Wright, P. E. & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.*, **16**, 18–29.

Wu, L., McElheny, D., Huang, R. & Keiderling, T. A. (2009). Role of tryptophan-tryptophan interactions in Trpzip β-hairpin formation, structure, and stability. *Biochemistry*, **48**, 10362–10371.

Wu, T. T. & Kabat, E. A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for anti-body complementarity. *J. Exp. Med.*, **132**, 211–250.

Wuttke, R., Hofmann, H., Nettels, D., Borgia, M. B., Mittal, J., Best, R. B. & Schuler, B. (2014). Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 5213–5218.

Xu, X. P., Zhai, D., Kim, E., Swift, M., Reed, J. C., Volkmann, N. & Hanein, D. (2013). Three-dimensional structure of Bax-mediated pores in membrane bilayers. *Cell Death Dis.*, **4**, e683.

Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. (1982). Living with water stress: Evolution of osmolyte systems. *Science*, **217**, 1214–1222.

Yue, P., Li, Z. & Moult, J. (2005). Loss of protein structure stability as a major causative factor in monogenic disease. *J. Mol. Biol.*, **353**, 459–473.

Zhang, Y. & Cremer, P. S. (2006). Interactions between macromolecules and ions: the Hofmeister series. *Curr. Opin. Chem. Biol.*, **10**, 658–663.

Zhang, Y. & Cremer, P. S. (2010). Chemistry of Hofmeister anions and osmolytes. *Annu. Rev. Phys. Chem.*, **61**, 63–83.

Zhang, Z., Zhu, W., Lapolla, S. M., Miao, Y., Shao, Y., Falcone, M., Boreham, D., McFarlane, N., Ding, J., Johnson, A. E., Zhang, X. C., Andrews, D. W. & Lin, J. (2010). Bax forms an oligomer via separate, yet interdependent, surfaces. *J. Biol. Chem.*, **285**, 17614–17627.

Zhou, H. X. (1993). Brownian dynamics study of the influences of electrostatic interaction and diffusion on protein-protein association kinetics. *Biophys. J.*, **64**, 1711–1726.

Zhou, H. X. (2012). Intrinsic disorder: Signaling via highly specific but short-lived association. *Trends Biochem. Sci.*, **37**, 43–48.