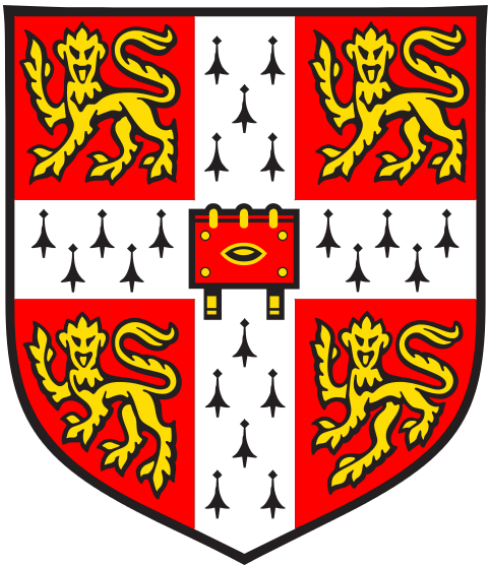
Studies of PARP-1 activation and inhibition using NMR spectroscopy



Tom Ogden

MRC Laboratory of Molecular Biology

Hughes Hall

University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

April 2019

Preface

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Thesis Summary

Studies of PARP-1 activation and inhibition using NMR spectroscopy

Tom Ogden

Poly(ADP-ribose)polymerase 1 (PARP-1) is a highly abundant multi-domain chromatin-associated enzyme found in all higher eukaryotic cell nuclei. It is the founding member of the PARP family of enzymes that modify themselves and other proteins by adding negatively charged poly(ADP-ribose) chains derived from NAD^+. PARP-1 is a key sensor of DNA strand breaks and becomes automodified in response to DNA damage. PARP-1 activation occurs following recognition of single or double strand breaks by PARP-1's N-terminal zinc finger domains. A multi-domain allosteric activation pathway leads to conformational changes in the HD subdomain of the catalytic domain that promote increased catalytic activity and lead to PARP-1 automodification and recruitment of DNA damage response proteins.

Small molecule inhibitors of PARP-1 catalytic domain activity are also the first compounds to be discovered which target tumour cells with deficiencies in homologous recombination via a synthetically lethal approach. PARP-1 inhibitors cause cytotoxicity by preventing PARP-1 from being released from sites of DNA damage, but the mechanistic basis of the wide differences in effectiveness between different compounds has yet to be definitively established. Therefore, the aims of the work described in this thesis are to further elucidate the changes that occur in the HD subdomain during catalytic domain activation, and to help establish how binding of different PARP inhibitors to the catalytic domain leads to the formation of trapped PARP-1 lesions.

The differences in chemical shift perturbations and relaxation parameters between the 360-residue wild-type catalytic domain and three constitutively partially active HD subdomain mutants were studied using NMR spectroscopy, and the wild-type PARP-1 catalytic domain was titrated with four different PARP inhibitors with varying levels of potency in order to compare the chemical shift perturbations, relaxation parameters and solvent exchange properties of the complexes.

Mutation of residues from key α-helices of the HD subdomain appeared to show that the arrangement of α-helices in the HD subdomain is important for maintenance of PARP-1 autoregulatory function. Additionally, the strength of HD subdomain contacts appears to be correlated with inhibitor binding affinity. Furthermore, binding of the PARP inhibitor EB-47 causes destabilisation of the same HD subdomain α-helices as does PARP-1 activation. Although further work is required to characterise the binding of EB-47 to full-length PARP-1, this finding suggests that EB-47 may promote formation of trapped PARP-1 lesions through a reverse allosteric mechanism.

Acknowledgments

The work presented here would not have been possible without the help that I received from many people. First and foremost, I would like to thank my supervisor David Neuhaus, for his patience, encouragement, sense of humour and support throughout the duration of this project, and for the large amount of time that he has spent patiently sharing his vast knowledge of NMR spectroscopy, listening to practice talks, writing scripts and reading the many drafts of this thesis. David wrote the scripts that were used in Sections 4.3 and 5.2 to globally scale chemical shift perturbations from multiple data sets against one another, and the script that mapped data from the CLEANEX and $^{2}H_2O$ solvent exchange experiments onto the structure of PARP-1 catalytic domain in Section 5.4.

I am also deeply grateful to Ji-Chun Yang, who performed NMR experiments, processed NMR data and patiently and expertly showed me how to perform sequential resonance assignment. Due to the enormous and daunting scale of the task, resonance assignment and analysis of relaxation data of the three PARP-1 catalytic domain mutants and four PARP inhibitor complexes was performed in collaboration with Ji-Chun.

I would also like to thank Laura Easton, who helped me get started in the lab and provided useful help and advice on all aspects of molecular cloning and protein expression and purification throughout the project. Prior to my arrival, Laura cloned the first ever PARP-1 catalytic domain fragment expressed in the Neuhaus group. Thank you also to William Hawthorne for his encouragement and helpful lab suggestions, and to Leo Kiss, Katy Hedgethorne and Harriet Crawley-Snowdon, the other current and past members of the Neuhaus group during my time here. It has been a great pleasure to work alongside all of you over the last three and a half years.

Much of the work described in Chapter 5 would not have been possible without the help of Kevin Embrey at AstraZeneca, who not only provided useful advice and many helpful observations but also facilitated productive collaboration with AstraZeneca throughout the duration of the work. To that end, I would also like to gratefully acknowledge the contribution of Marianne Schimpl at AstraZeneca, who very kindly supplied us with the co-ordinates of two unpublished crystal structures of PARP-1 catalytic domain in complex with the PARP inhibitors veliparib and olaparib that she had solved. These were used to display the mapping results for CSPs and solvent exchange experiments from wild-type PARP-1 catalytic domain in complex with these compounds in Sections 5.2 and 5.4.

I am also very grateful to our collaborator John Pascal of the Université de Montreal for the very helpful discussions shared with our group over the course of this project, most notably on the subjects of the catalytic domain mutants investigated in Chapter 4 and the PARP inhibitor EB-47 investigated in Chapter 5.

Across the wider LMB, I would like to thank Stephen McLaughlin of the LMB biophysics facility for showing me how to use the Nanotemper Prometheus NT.48 machine to perform thermal stability measurements on PARP-1 catalytic domain (Section 3.5), and Farida Begum who performed mass spectrometry on samples of the PARP-1 catalytic domain (Section 3.6). I would also like to thank the MRC for its financial support.

Finally, I would like to thank my friends from the LMB and beyond for many amazing experiences in Cambridge and further afield over the years, and my family for their support and encouragement along the way, without which completion of this thesis would not have been possible.

Contents

Binding of 58-bp dsDNA to the *Dictyostelium discoideum* mating-type protein MatA

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

List of Figures

List of Tables

Abbreviations and symbols

ADP-ribose	Adenosine diphosphate ribose
ADPr	ADP-ribosylation
ART	ADP-ribosyl transferase
BAD	Benzamide adenine dinucleotide
BER	Base excision repair
bp	Base pair
BRCT	Breast cancer susceptibility protein C-terminal
CAT	PARP-1 catalytic domain
CSP	Chemical shift perturbation
CV	Column volumes
D_2O	Deuterium oxide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSB	Double-strand break
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
HD	Helical domain
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
hNOE	heteronuclear NOE (see NOE below)
HR	Homologous recombination
HSQC	Heteronuclear single quantum coherence spectroscopy
HXMS	Hydrogen/deuterium exchange mass spectroscopy
IC_{50}	Half maximal inhibitory concentration
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_D	Dissociation constant
k_d	Off-rate
K_i	Inhibition constant
MART	Mono(ADP-ribosyl)transferase
MWCO	Molecular weight cut off
NAD^+	Nicotinamide adenine dinucleotide
nanoDSF	nano differential scanning fluorimetry

NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NiNTA	Nickel-NTA (nitriloacetic acid)
NMN	Nicotinamide mononucleotide
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NUS	Non-uniform sampling
PAR	Poly ADP-ribose
PARP	Poly(ADP-ribose)polymerase
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PMSF	Phenylmethylsulfonyl fluoride
ppm	Parts per million
RDC	Residual dipolar coupling
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SSB	Single-strand break
T_1	Longitudinal relaxation time
$T_{1\rho}$	Spin-locked transverse relaxation time
T_2	Transverse relaxation time
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEV^{pro}	Tobacco etch virus protease
TR-FRET	Time resolved fluorescent energy transfer
T_{onset}	Temperature at onset of melting
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimised spectroscopy

Chapter One

Introduction

Poly(ADP-ribose) polymerase 1 (PARP-1) is a highly abundant chromatin-associated enzyme found in all higher eukaryotic cell nuclei. It is the founding member of the PARP family of enzymes; these modify a range of target proteins via the posttranslational modification of poly ADP-ribosylation, in which negatively charged poly ADP-ribose (PAR) chains derived from NAD^+ are added (D'Amours et al., 1999).

PARP-1 plays a key role in the DNA damage response in higher eukaryotic cells. It is not involved directly in the chemical reactions of DNA strand break repair, but instead acts as a key initial sensor, particularly for DNA single-strand breaks that are the most common form of DNA damage. Upon becoming activated by such damage, PARP-1 responds by adding long, negatively-charged, branched chains of PAR to itself and other proteins. These long, branched chains effectively act as a signal to the repair machinery by providing a scaffold for recruitment of factors essential for multiple DNA repair pathways (Caldecott, 2007). PARP-1 is also the target of the first commercially available drugs to exploit the genetic principle of synthetic lethality, in which the combined loss of function of two genes or proteins promotes cell death, but the loss of a single one does not. Thus, tumours resulting from germline mutations in *BRCA1* and *BRCA2* genes that lack a functional homologous recombination pathway are sensitive to inhibition of PARP-1 catalytic activity using small molecules, whereas healthy cells are not (Bryant et al., 2005; Farmer et al., 2005).

In this introduction, the structure and metabolism of poly ADP-ribose will be discussed, as well as its predominant cellular functions including its role in the DNA damage response. The structural and biochemical information published to date relating to the function of the PARP family of proteins which carry out poly and mono-ADP-ribosylation will be outlined, with a focus on the structure and function of the catalytic domain and other domains of the protein PARP-1, the predominant generator of poly(ADP-ribose) in higher eukaryotic cells. Finally, the clinical utility, importance and putative mechanism of the inhibition of PARP-1 and ADP-ribosylation by PARP inhibitors will also be summarised.

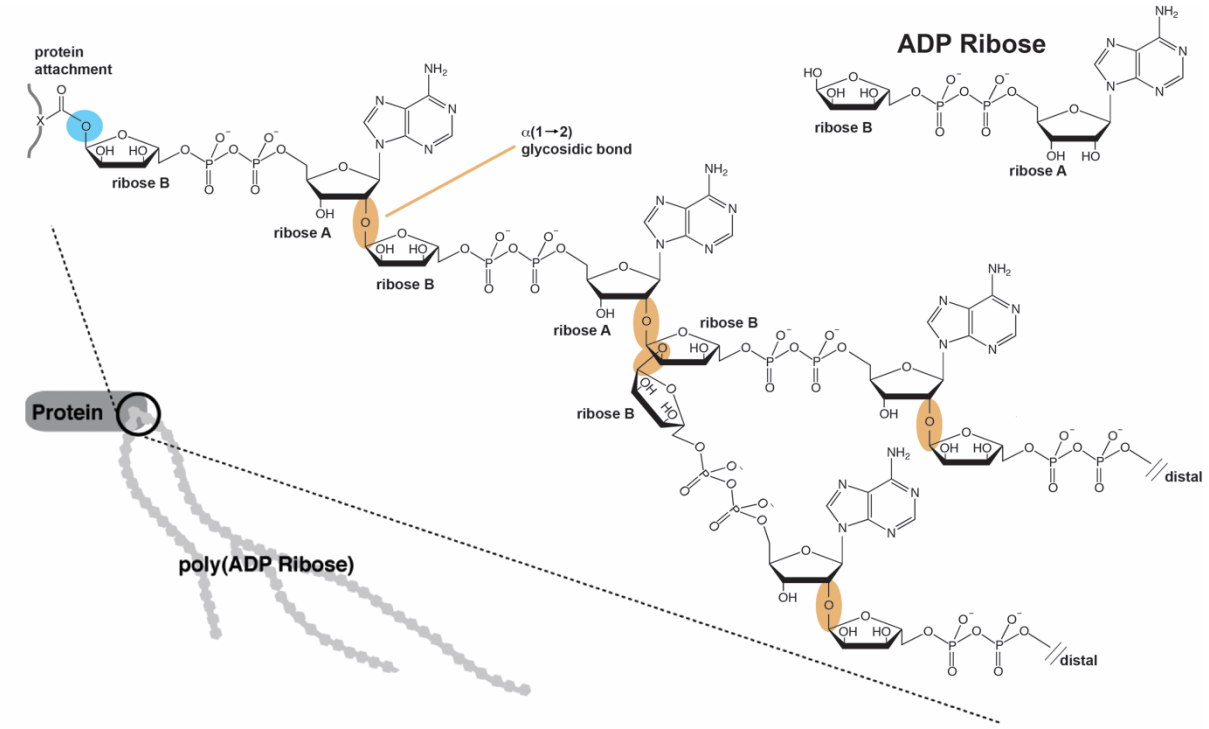


Figure 1.1 Chemical structure of poly ADP-ribose (PAR). PAR is a post translational modification that is covalently attached to side chains of certain residue types in acceptor proteins via formation of O-glycosidic linkages (shown here in blue). Polyenzyme members of the PARP family (see section 1.1.3) elongate PAR chains by catalysing the formation of α(1→2) glycosidic bonds (highlighted in orange) between ribose A of one unit of ADP ribose and ribose B of another. Branching occurs at a less frequent rate via formation of α(1→2) glycosidic bonds between ribose B moieties of two units of ADP ribose. PAR synthesis proceeds from the site of protein attachment to the distal ends of the PAR chains.

1.1 Poly ADP-Ribose and the PARP family

1.1.1 Poly ADP-ribose

Poly ADP-ribose (PAR) was first discovered by the group of Paul Mandel, who observed that addition of the NAD^+ precursor nicotinamide mononucleotide (NMN) to a nuclear fraction isolated from hen liver resulted in the enzymatic production of a polymer containing adenine that was resistant to digestion by both DNAse and RNAse (Chambon et al., 1963). While initially misidentified as polyA, the polymer was later identified as PAR and NAD^+ was later identified as the substrate of PARP-1 (Chambon et al., 1966).

Poly ADP-ribose is formed by the polymerisation of the ADP ribose moiety of NAD+ (Chambon et al, 1966). Its structure is well characterised (Figure 1.1), with ADP-ribose monomers linked by α(1→2) ribose-ribose O-glycosidic bonds (Miwa et al., 1981). Chain lengths are heterogenous and polymers also branch every 40 to 50 residues both *in vitro* and *in vivo* via the formation of additional α(1→2) ribose-ribose linkages (Hayashi et al., 1983; Alvarez-Gonzalez and Jacobson, 1987).

1.1.2 Mono ADP-ribose

Post-translational modification of proteins via the addition of a single ADP-ribose moiety to side chains of amino acid residues also occurs, and proteins that catalyse only this reaction are considered to be mono (ADP-ribosyl) transferases (MARTs), also sometimes called monoenzymes (Gupte et al., 2017). The first MARTs to be discovered were bacterial toxins, including from diphtheria, cholera and pertussis (Krueger and Barbieri, 1995). However, it has been established that MARTs perform widespread functions across many organisms, including in higher eukaryotes where the majority of members of the PARP family are considered to be MARTs (see further discussion in Section 1.1.4).

1.1.3 Production of ADP ribose by PARP-1

Poly ADP-ribosylation of acceptor proteins, using NAD as a substrate, is catalysed by poly ADP-ribose polymerases (PARPs), also sometimes referred to as polyenzymes (Gupte et al., 2017). Despite the polyenzyme PARP proteins' ability to catalyse formation of long PAR chains, the catalytic mechanism of ADP ribose formation from NAD is identical in PARPs and mono ADP ribosyl transferases (MARTs), with the three-dimensional structure of chicken PARP-1 catalytic domain revealing a high degree of structural homology and shared evolutionary history with bacterial toxins that have mono ADP-ribosyl transferase activity (Ruf et al., 1996). In all cases NAD^+ is recognised by a characteristic motif β-α-loop-β-α, a fold now designated as a "PARP signature" and used to identify and classify members of the PARP family and PARP paralogues (Amé, 2004). This motif includes the NAD^+ binding site, a 50-residue block of amino acids (residues 859-908 in PARP-1), which is the location for binding of a nicotinamide analogue in the published crystal structure of PARP-1 catalytic domain from chicken (Ruf et al., 1996), and which is highly conserved across PARP sequences in vertebrates (de Murcia and Ménissier de Murcia, 1994). Formation of PAR chains requires three distinct steps, and PARP-1 catalytic domain is able to perform all three:

1. Initiation: Formation of an ADP-ribose unit from NAD^+ and attachment of this unit to the sidechain of an amino acid via an O or N glycosidic linkage (a step also carried out by MARTs)
2. Elongation: Addition of one ADP-ribose unit to a growing chain by formation of a $\alpha(1\rightarrow2)$ O glycosidic bond from ribose A of one unit to ribose B of another unit, using the 2'-hydroxyl group from the ribose A at the chain terminus as an acceptor.

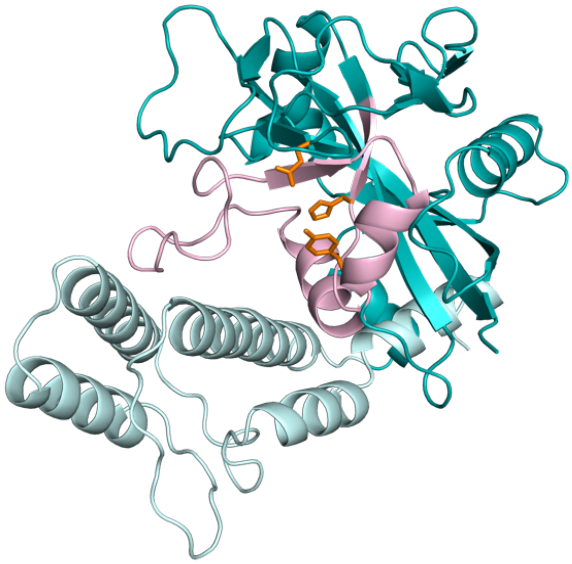


Figure 1.2 Structure of PARP-1 catalytic domain (pdb 1a26). The autoregulatory HD subdomain domain is shown in cyan. The C-terminal ART domain (teal) contains the "PARP signature" fold (pink). Side chains of residues H862, Y907 and E988 comprising the H-Y-E motif required for poly-ADP-ribosylation are shown in orange.

3. Branching: Formation of a α(1→2) O glycosidic bond between ribose B on two separate units of ADP-ribose. Each of these ADP ribose units can be elongated from the 2'-hydroxyl group from the ribose A moiety as described in step 2.

It has been proposed that both elongation and branching are permitted by PARP-1 because the active site is accessible to PAR chains in two different orientations, which facilitate formation of linear and branched linkages respectively (Ruf et al., 1998). The conserved PARP fold contains histidine and tyrosine residues that bind and orient NAD^+ in the required arrangement for formation of PAR, while a key glutamate residue is responsible for mediating transfer of ADP-ribose to the acceptor residue and is also considered to be essential for elongation, as its mutation abolishes poly ADP-ribosylation in favour of mono ADP-ribosylation (Ruf et al., 1998).

Only the side chains of certain amino acid types in PARP-1 and other targets of ADP-ribosylation are ADP-ribosylated and evidence strongly suggests that there is no sequence-specific targeting of residues in acceptor proteins of ADP-ribosylation, with ADP ribosylation patterns in PARP-1 itself and target proteins simply based on amino acid type and proximity to the catalytic centre of the ADP-ribosylating protein. The γ-COOH group of glutamic acid was the first acceptor of ADP-ribosylation to be identified (Riquelme et al., 1979; Burzio et al., 1979), with the -COOH group of aspartic acid also proposed later (Suzuki et al., 1985). Subsequent studies have reported ADP ribosylation of side chains from arginine, asparagine and lysine residues (Laing et al., 2011; Manning et al., 1984; Altmeyer et al., 2009). More recently, evidence has emerged to show that serine residues of histones

are ADP-ribosylated in response to DNA damage (Leidecker et al., 2016; Palazzo et al., 2018). It has also been reported that ADP-ribosylation of histone serine residues is dependent on interaction of histone PARylation factor 1 (HPF1) with PARP-1 or PARP-2 and that PARP-1 also automodifies its own serine residues when it becomes activated in the presence of HPF1 (Bonfiglio et al., 2017).

1.1.4 The PARP family

The entire PARP family contains 17 members, all of which share a conserved catalytic domain containing the PARP signature that was first discovered in PARP-1 (Gibson and Kraus, 2012). However, the members of the PARP family vary widely in their enzymatic activity, the presence or absence of N-terminal domains, subcellular localisations and cellular function. In addition, only four members of the PARP family (PARP-1, PARP-2, and tankyrases PARP-5a and PARP-5b) have been shown experimentally to exhibit poly ADP- ribosylation activity. All other members of the PARP family have been identified as MARTs, with the exception of PARP-9 and PARP-13 that appear to be catalytically inactive (Vyas et al., 2014). Differences in the ability of members of the PARP family to perform mono- or poly-ADP ribosylation can be explained by variations in the three residue H-Y-E motif found in the PARP-1 catalytic domain. In all MART members of the PARP family, the key glutamate residue has been replaced with isoleucine, leucine or tyrosine, leading to loss of polymerase activity (Gupte et al., 2017). PARP-3 and PARP-4 share the classical H-Y-E motif with the polyenzyme members of the family; however, they only display monoenzymatic activity due to their structurally distinct donor ("D") loops (position 880 to 900 in PARP-1 catalytic domain), indicating that the H-Y-E motif is necessary but not in itself sufficient to confer polyenzyme activity (Vyas et al., 2014). PARP-9 and PARP-13 lack the conserved histidine residue of this motif, which could disrupt NAD^+ binding and account for their catalytic inactivity (Vyas et al., 2014).

PARP-1 is distinguished from other members of the PARP family (including PARP-2 and PARP-3, the other DNA damage response PARPs) by its N-terminal domain structure (Barkauskaite et al., 2015) (see Figure 1.2). Human PARP-1 is a 113 kDa protein of 1014 residues and contains six independently folded domains connected by flexible linkers. N terminal F1 and F2 zinc-finger domains are responsible for initial binding to sites of DNA damage, leading to activation of the C terminal CAT domain, which itself contains two subdomains; the ADP-ribosyl transferase (ART) subdomain where the signature PARP fold is found and a regulatory helical domain (HD). In between these regions, a third zinc finger

domain (F3) structurally unrelated to the F1 and F2 domains, a BRCT domain containing acceptor sites for autoPARylation and a WGR domain (named after its Trp-Gly-Arg motif conserved within the PARP family) are found. PARP-2 and PARP-3 share a WGR domain and catalytic domain with PARP-1, complete with the HD regulatory subdomain, but have only short, disordered N-terminal regions in contrast to PARP-1's four N-terminal subdomains. In contrast to PARP-1, which can sense many forms of damaged DNA, DNA single-strand breaks must be 5' phosphorylated to activate PARP-2 or PARP-3 (Langelier et al., 2014).

1.1.5 Cellular functions of PARP-1 and PAR metabolism

As mentioned earlier, in higher eukaryotes PARP-1 is the key sensor of single-strand breaks, the most common form of DNA damage that occurs *in vivo,* and acute activation of PARP-1 is one of the earliest responses to this form of stress in eukaryotic cells. The main target of PAR chain addition by PARP-1 is PARP-1 itself, and, following DNA-damage dependent activation, addition of highly negatively-charged branched PAR chains to PARP-1 acts as a signal for recruitment of DNA repair complexes such as XRCC1 (Caldecott, 2007). Although PARP-1 itself does not participate in the covalent chemistry of DNA repair and must be released from the site of DNA damage in order for repair to proceed, it has been shown that activation of PARP-1 following binding to DNA single-strand breaks (SSBs) plays an important role in DNA repair processes including base excision repair (BER), homologous recombination (HR), non-homologous end-joining (NHEJ), nucleotide excision repair (NER) and single-strand break repair (SSBR) (Caldecott, 2007; De Vos et al., 2012). In addition to recognising single-strand breaks (SSBs), PARP-1 has been shown to recognise other less common forms of DNA damage, including double-strand breaks (DSBs), cruciforms and hairpins, and has also been implicated in detection and re-initiation of stalled replication forks (Bryant et al., 2009; Lonskaya et al., 2005; Potaman et al., 2005). The release of activated PARP-1 from sites of DNA damage is likely to occur at least in part as a result of electrostatic repulsion between DNA and the negatively-charged PAR chains resulting from automodification (Satoh and Lindahl, 1992) (Mortusewicz et al., 2007).

As well as its role in DNA damage sensing and repair, post-translational modification via poly ADP-ribosylation (both of PARP-1 itself and other proteins such as histones) has been shown to play a role in regulation of cellular processes such as chromatin remodelling, gene expression, histone chaperoning and even cell fate decisions, as hyperactivation of PARP-1 in response to severe genotoxic stress results in a form of regulated cell death

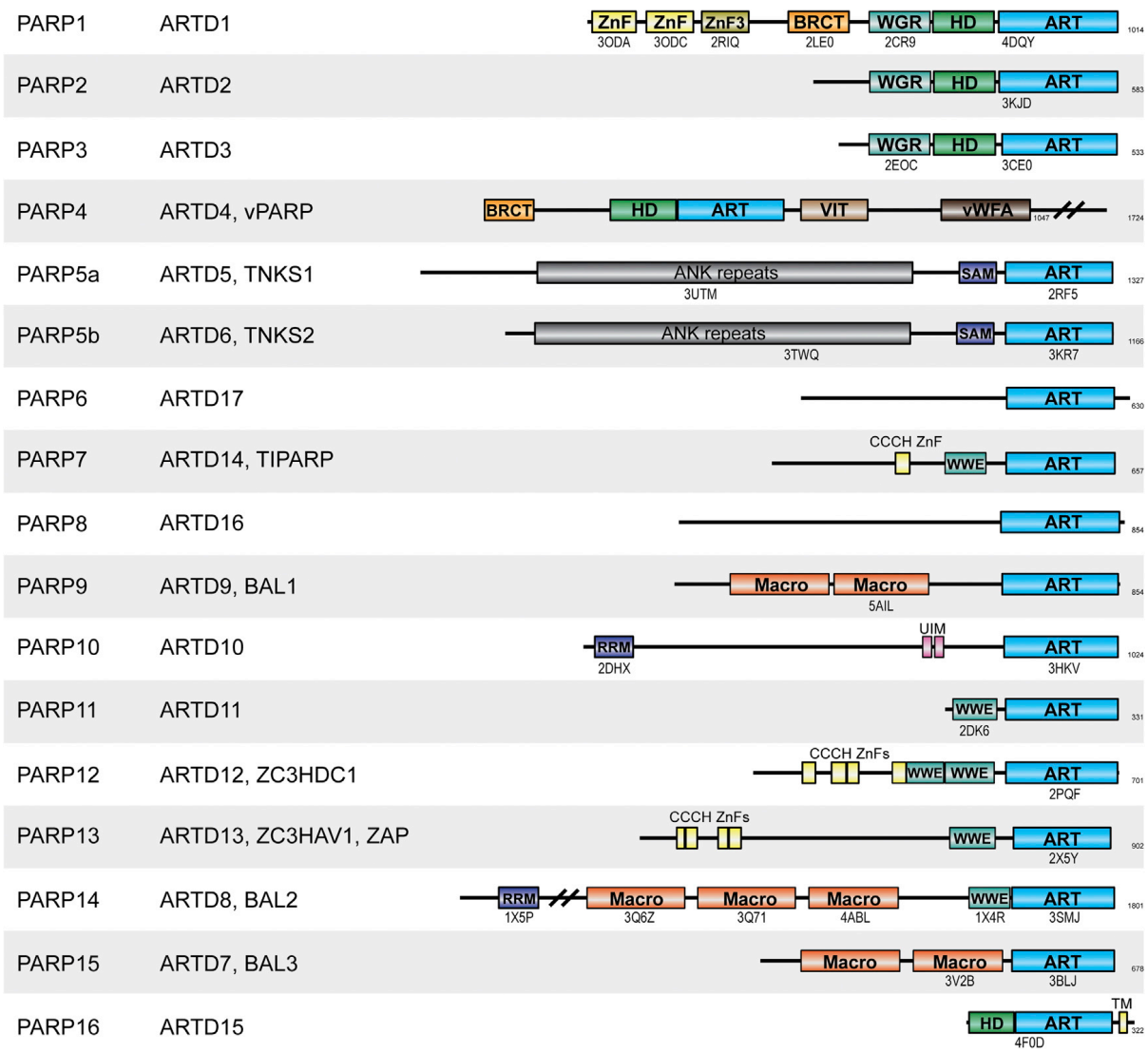


Figure 1.3 Domain structures of proteins of the PARP family. The names and domain structures of the 17 members of the PARP family, including alternative names. All members of the PARP family have an ART domain containing the PARP signature fold. Despite this, only PARP-1, PARP-2, and tankyrases PARP-5a and PARP-5b have been shown to exhibit poly-ADP ribosylation activity, with all other members now considered to be mono ADP-ribosyl transferases (MARTs) (Vyas et al., 2014). In PARP-1, PARP-2 and PARP-3, the ART is a subdomain of the catalytic domain and is preceded by an HD regulatory subdomain. Where structures of domains from the PARP family have been previously solved, a PDB code is listed underneath the domain. The number of amino acids in each protein is shown on the right-hand side of the figure. Abbreviations: ARTD, ADP-ribosyl transferase diphtheria type; vPARP, vault PARP; TNKS, Tankyrase; TIPARP, TCDD inducible PARP; BAL, B cell aggressive lymphoma; ZC3HDC1, zinc finger CCCH-type domain containing 1; ZC3HAV1, zinc finger CCCH-type antiviral 1; ZAP, zinc finger antiviral protein. Abbreviations of domain names: ZnF, zinc finger (PARP-1 zinc finger domains are referred to using simply F [for finger] elsewhere in this thesis); BRCT, BRCA1 C-terminal; WGR, conserved Trp-Gly-Arg motif domain; HD, helical domain; ART, ADP-ribosyl transferase domain; VIT, vault protein inter-alpha-trypsin; vWFA, von Willebrand type A; ANK, ankyrin; SAM, sterile alpha motif; CCCH ZnF, CCCH type zinc finger; WWE, three conserved residues W-W-E motif domain; Macro, macrodomain; RRM, RNA recognition motif; UIM, ubiquitin interaction motif; TM, transmembrane motif. Reproduced from Barkauskaite et al., 2015.

distinct from both apoptosis and necrosis referred to as “parthanatos”, a portmanteau of PAR and Thanatos, the ancient Greek personification of death (Galluzzi et al., 2012). In parthanatos, high quantities of PAR produced by PARP-1 in response to high levels of DNA

damage migrate from the nucleus to the cytosol, triggering release and transport to the nucleus of the mitochondrial protein Apoptosis Inducing Factor, and eventually chromatin condensation and cleavage of genomic DNA into large fragments by endonucleases (Fouquerel and Sobol, 2014). Although parthanatos is a functionally distinct process from apoptosis and PARP-1 does not participate in the apoptosis pathway, PARP-1 interdomain linkers are cleaved by caspases, abrogating PARP-1 activity and producing signature fragments (notably PARP-1 1-214) which are considered to be biomarkers for induction of apoptosis (Chaitanya et al., 2010). As a result of its role in parthanatos and ability to rapidly deplete cellular reserves of NAD^+ upon hyperactivation, PARP-1 activity must be tightly regulated, and turnover of PAR in the cell is rapid (Gupte et al., 2017).

In addition to poly and mono-PARP enzymes which act as "writers" of ADPr, "erasers" of ADPr which catalyse the removal of PAR chains or mono-ADP ribose from proteins also exist. These include ADP-ribosyl hydrolase 3 (ARH3) and PAR glycohydrolase (PARG), which catalyse cleavage of the ribose-ribose bonds in PAR but leave a single ADP-ribose moiety bound to the acceptor amino acid (Oka et al., 2006; Slade et al., 2011), and TARG, MacroD1 and MacroD2 which can catalyse the complete removal of ADPr by hydrolysing the ester bond between ADP-ribose and glutamic or aspartic acid side chains (Jankevicius et al., 2013; Rosenthal et al., 2013; Sharifi et al., 2013). Removal of ADPr by the NUDIX family of hydrolases results in the formation of a phosphoribose moiety attached to the acceptor side chain (Daniels et al., 2015).

1.1.6 PARP-1 DNA damage sensing and activation

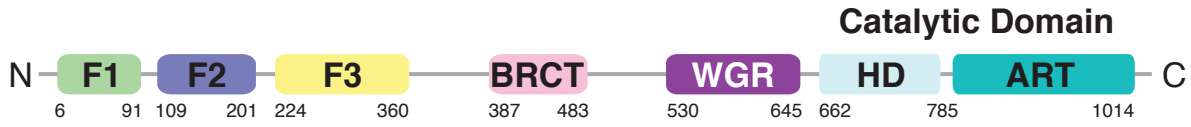


Figure 1.4 PARP-1 domain structure, domain nomenclature and colour scheme as used throughout this thesis.

NMR spectroscopy and X-ray crystallography have revealed the structural and mechanistic basis of the function of all PARP-1 domains in response to detection of DNA single strand and double strand breaks, with the exception of BRCT which does not appear to directly participate in the activation of PARP-1 following binding to the site of DNA damage (Altmeyer et al., 2009).

While SSBs are the most common form of DNA damage found in the cell, to date no crystal structures of PARP-1 in complex with an SSB have been obtained, presumably due

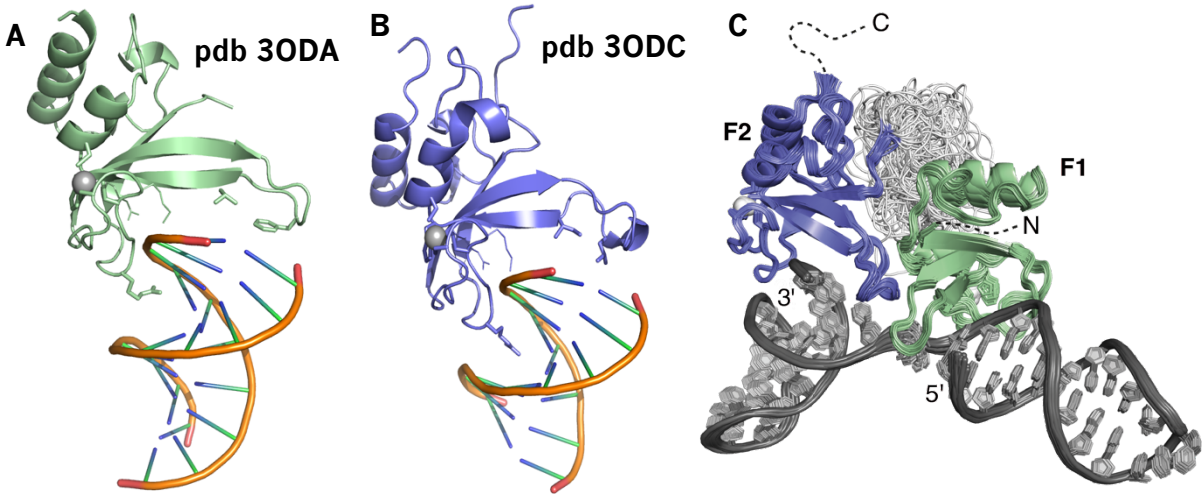


Figure 1.5 **The structural basis of F1 and F2 binding to DSBs and SSBs.** (A), (B). Crystal structures of F1 (A, green) and F2 (B, purple) domains bound to DNA blunt ends. (C) Hybrid X-ray and NMR structure of F1F2 bound to a model of an SSB (gapped dumbbell DNA ligand with a 5'-phosphorylated single nucleotide gap). The ensemble of all 78 acceptable structures derived from the structure determination process are shown.

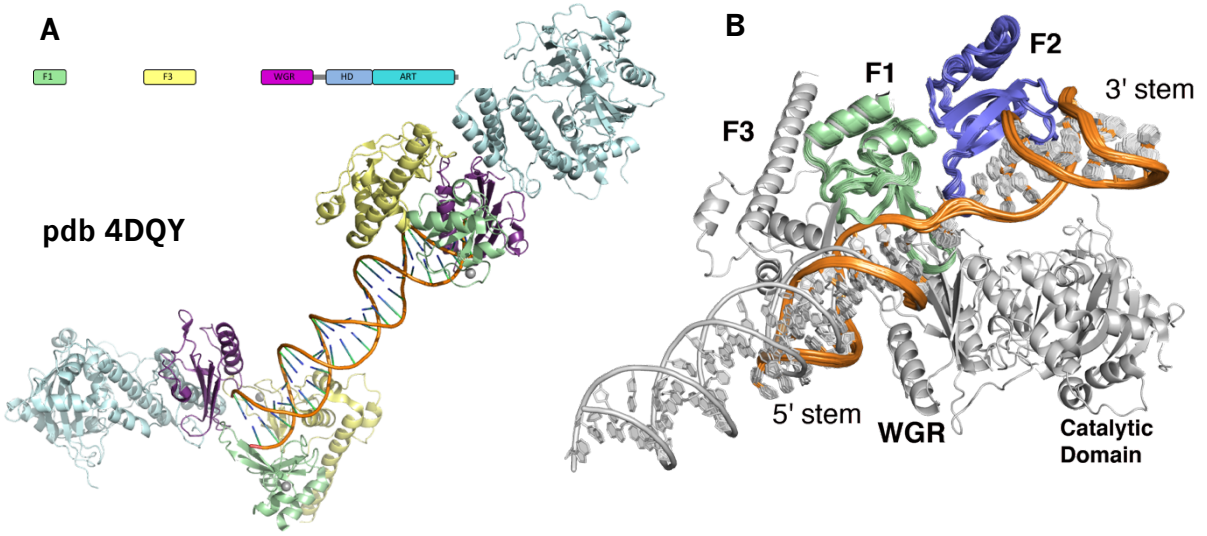


Figure 1.6. **Formation of interdomain contacts upon binding to a DNA duplex or SSB.** (A) Cartoon representation of PARP-1 F1, F3, WGR and CAT domains assembled at each end of a DNA duplex (pdb 4DQY). Each domain can be identified via the colour scheme in the domain structure. (B) Superposition of the common features (F1 and DNA stems) of F1F2 SSB complex from Eustermann et al, 2015 with pdb 4DQY.

to the inherent flexibility of the DNA around the intact strand that likely interferes with the formation of crystals. Previous work in the Neuhaus group (by Sebastian Eustermann), used a gapped DNA dumbbell as a model of an SSB and studied its interaction with PARP-1 F1F2 by NMR in combination with other biophysical techniques (EMSA, centrifugation and fluorescence). This revealed that recognition of SSBs is achieved by F1F2 acting as a monomer and in a single orientation, while parallel studies with the separated F1 and F2 domains showed that it is F2 that provides most of the affinity for DNA binding. Addition

of a nicked DNA dumbbell ligand instead of a gapped DNA dumbbell resulted in only minor differences in chemical shift perturbations (CSPs) in spectra of F1F2, indicating that PARP-1 recognises both forms of SSB in a highly similar manner (Eustermann et al., 2011).

Work in the Neuhaus group to obtain a solution NMR structure of F1F2 bound to an SSB dumbbell ligand continued, but in the event crystal structures of separate F1 and F2 domains bound to blunt ends of DNA duplexes (pdb 3ODA and pdb 3ODC; see Figures 1.4A and 1.4B) were published soon afterwards by the group of John Pascal (then at Thomas Jefferson University Philadelphia, now moved to Université de Montréal, Canada). These showed strong similarities with unpublished structural features of SSB binding by F1F2 observed by NMR in the Neuhaus group. Notably, in the crystal structures of both F1 and F2 on DNA blunt ends, the domains engaged the exposed nucleotide bases of DNA through a "base stacking loop" (to use the nomenclature introduced by Pascal), used a "phosphate backbone grip" comprising a set of suitably positioned basic sidechains to recognise the phosphate backbone of the DNA, and inserted an arginine side chain into the minor groove of the DNA in a non-base specific manner (Langelier et al., 2011). In combination, these interactions resulted in the positioning of each finger domain over the minor groove and 3' terminus of the DNA stem to which they bound.

Given that the existing NMR data (intermolecular NOEs and chemical shift perturbations upon DNA binding) were fully consistent with this same binding mode being adopted by both F1 and F2 in the complex of F1F2 with the gapped dumbbell DNA, the Neuhaus group devised a protocol for calculating a structure for this complex based on a combination of constraints derived from NMR experiments (F1-F2 and intermolecular NOEs as well as amide group RDCs) and similarity to the crystal structures (non-crystallographic symmetry constraints to adapted template copies of pdb 3ODA, for F1, and pdb 3ODC, for F2). The resulting hybrid NMR/X-ray structure is shown in Figure 1.4C). Very shortly after this structure for the complex of F1F2 with the gapped dumbbell was derived in the Neuhaus group, the Pascal group published a crystal structure (pdb 4DQY, Fig 1.5A) of the F1, F3, and WGR-CAT domains (as three separate protein chains) in complex with a DNA duplex, which showed that these domains assemble onto a DNA blunt end and thereby form a network of interdomain contacts (Langelier et al., 2012). They suggested that these contacts were involved in transmission of the DNA-damage-dependent activation signal to the catalytic domain, which as shown in this structure, does not itself make any contact with the DNA.

However, this crystal structure lacks the F2 domain of PARP-1, as well as the second stem of DNA that would be present for a complex with an SSB. It was therefore of great interest to superpose the F1F2 SSB complex structure with pdb 4DQY, using those parts of the structure that were common to both, i.e. the F1 domain and its associated DNA stem. This superposition (Figure 1.5B) strongly suggested that a similar domain arrangement might occur for a complex with an SSB, as well as showing how the F2 domain might participate in such a case. Given the complementary nature of the work in the two groups, the Neuhaus and Pascal groups initiated a collaboration at this point, and chemical shift perturbation and ^{15}N relaxation experiments were carried out by the Neuhaus group on complexes containing F1, F2, F3 and WGR domains (Eustermann et al., 2015). Overall, this work provided a number of insights into DNA damage binding and subsequent activation of PARP-1 (see Figure 1.6), summarised below:

- F1F2 binds to an SSB as a monomer and in a single orientation. F2 binds first due to its greater affinity. F1 then also binds, thereby forming a small, DNA-damage dependent interface between F1 and F2. This model of binding explains non-sequence specific PARP-1 DNA damage sensing activity by showing how F1F2 binding forces damaged DNA into a conformation that undamaged DNA cannot reach, with the DNA stems on each side of the damage site bent and twisted apart to expose both 5' and 3' ends to F1 and F2 respectively.
- F3 interacts with F1 and the DNA (this also occurs following binding of F1 to a DSB site) and thereby forms a composite interface to receive the WGR domain. Upon titration of WGR to an F1F2F3 complex, existing CSPs between other interfaces were reinforced, supporting the idea that formation of these interfaces is co-operative and stepwise.
- WGR binds to F1, F3 and the DNA, which creates a further composite interface on WGR and F3 for the catalytic domain to bind, which leads to activation.

Importantly, no direct contact between F1F2 or DNA and the catalytic domain occurs and only binding of F1F2 as a monomer in the orientation described above allows subsequent co-operative stepwise assembly of PARP-1 domains in this manner, as binding in the opposite orientation would not allow subsequent interaction of F1 with F3.

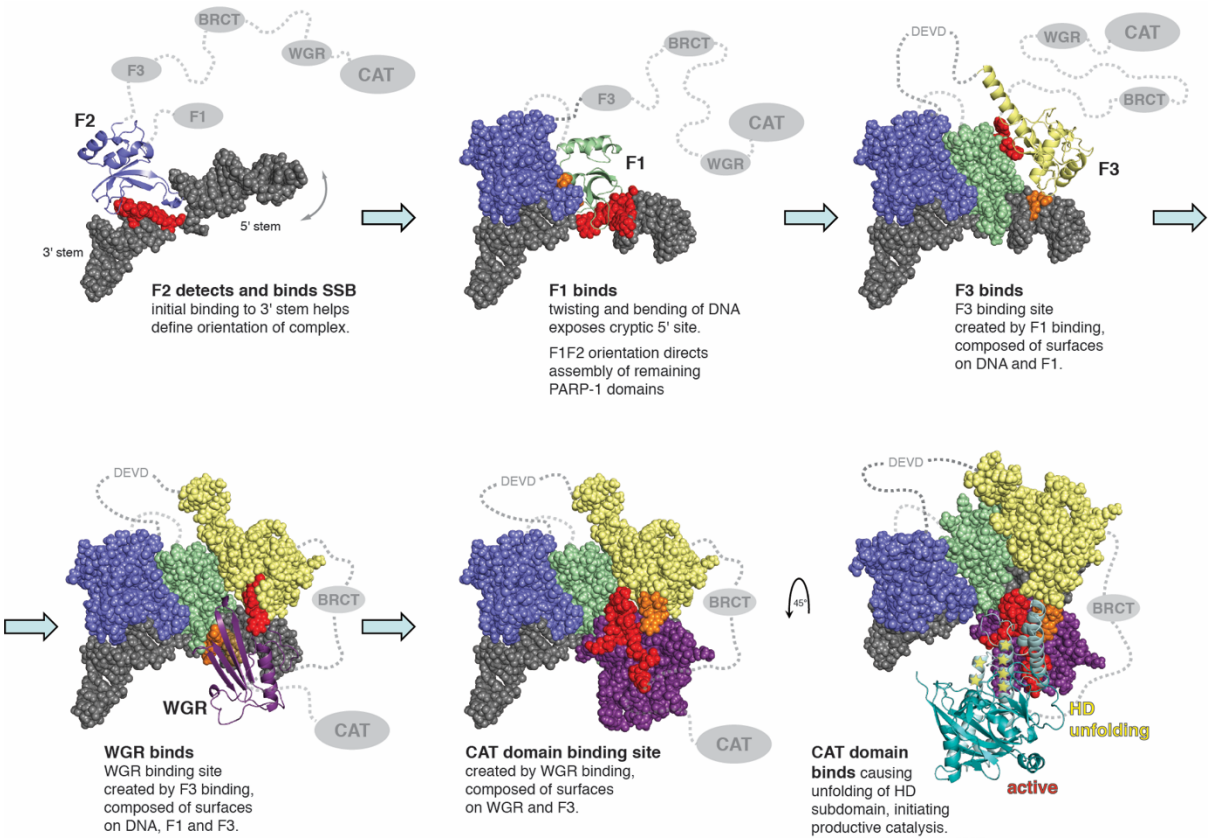


Figure 1.7 Allosteric activation mechanism of PARP-1 by DNA single-strand breaks. In the absence of damaged DNA, PARP-1 domains behave independently and are connected by disordered linkers. Recognition of an SSB by F1 and F2 initiates stepwise multidomain folding, leading to formation of an interface composed of surfaces on F3 and WGR for the catalytic domain to bind. Residues forming part of interfaces between domains are coloured red or orange. DEVD indicates the location of a caspase cleavage site between F2 and F3. Reproduced from Eustermann et al., 2015.

In contrast with the steps involved in allosteric transmission of the activation signal to the catalytic domain, less high-resolution structural or dynamic information is available concerning the events taking place in the catalytic domain itself during activation. Although the Pascal group was able to obtain a crystal structure which included PARP-1 catalytic domain and showed that the catalytic domain was somewhat destabilised and structurally distorted upon DNA binding (Langelier et al., 2012), their subsequent study of activated full-length PARP-1 using hydrogen/deuterium exchange-mass spectrometry (HXMS) unexpectedly revealed more profound effects, in that some parts of the catalytic domain's HD subdomain undergo significantly faster exchange of amide protons with solvent once DNA damage recognition has occurred. This suggests that at least some degree of secondary structure perturbation occurs, possibly in the form of local partial unfolding (Dawicki-McKenna et al., 2015). In conjunction with mutational and crystallographic studies carried out in in the same paper, this suggests that the HD subdomain is autoinhibitory in nature and

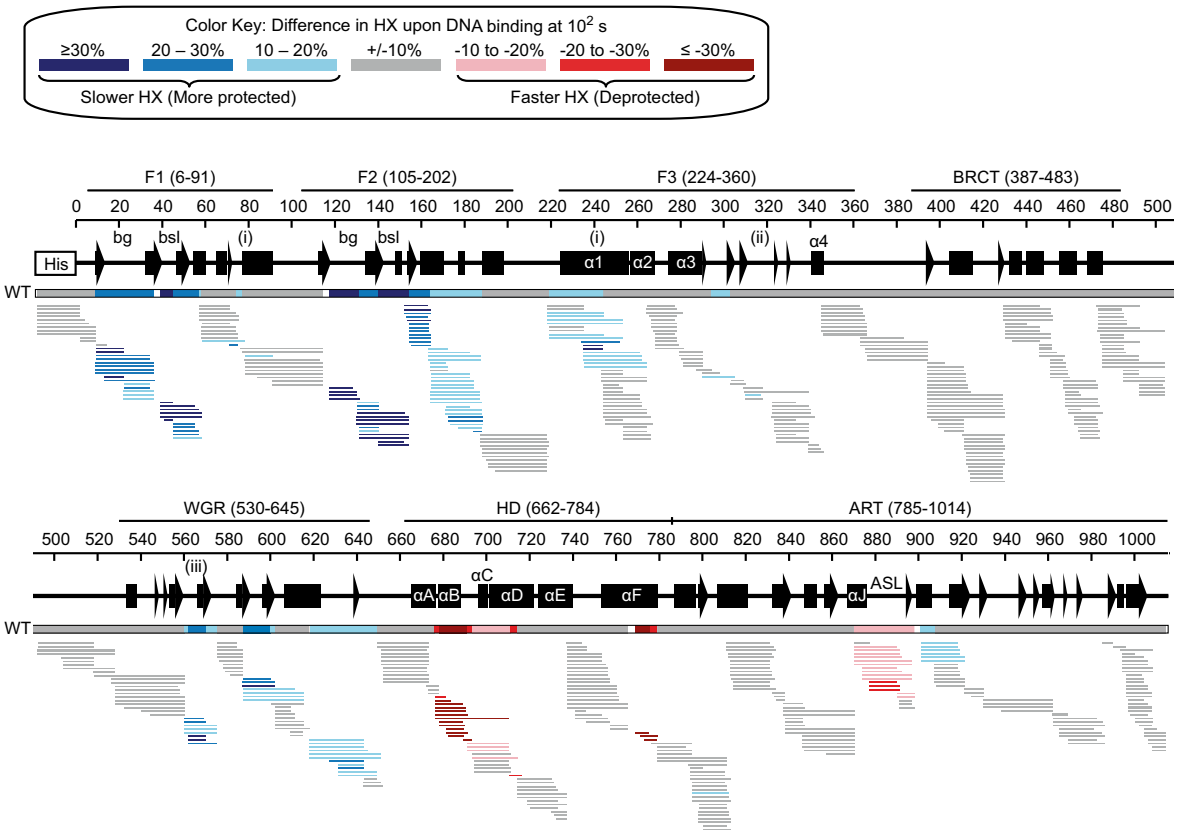


Figure 1.8 Changes in rates of solvent exchange in full-length PARP-1 following detection of DNA damage after 100 seconds. Slower rates of hydrogen exchange (HX) shown in blue are observed in fragment peptides at sites of interdomain contacts following activation. Faster rates of HX shown in red are observed in fragments from the active site location (ASL) and the HD subdomain of the catalytic domain. The extent of hydrogen exchange in the affected regions of the HD subdomain after 100 seconds is not consistent with maintenance of stable helices during catalytic domain activation and is more consistent with significant partial melting of the helical regions. Consensus behaviour of each region of the sequence was determined according to rates of exchange in peptide fragments. Individual fragments from the experiment and their differences in HX are shown below the domain structure, sequence numbering, secondary structure and consensus behaviour. Abbreviations of structural elements: bg, backbone grip; bsl, base stacking loop; (i), F1–F3 interface; (ii), F3–WGR–HD interface; (iii), WGR–HD interface. Reproduced from Dawicki-McKenna et al., 2015.

performs its regulatory function by blocking productive binding of NAD^+ to the ART domain in the absence of damaged DNA, and that HD unfolding upon PARP-1 activation relieves this autoinhibition. However, the intrinsic resolution of HXMS is limited to the level of fragment peptides, and even if a fully activated state was crystallised, the full extent and detailed nature of HD unfolding is unlikely to be captured at atomic resolution in a crystal structure.

1.2 Inhibition of PARP-1 using small molecules

1.2.1 Synthetically lethality of PARP-1 inhibition

PARP-1's role in DNA repair has made it a target for the first cancer therapies based on the principle of "synthetic lethality", in which the combined loss of function of two genes,

proteins or pathways results in cell death, but the loss of one of them does not (Ashworth et al., 2011). PARP-1$^{-/-}$ mouse knockouts are embryonically viable and fertile, but are demonstrably deficient in repair of SSBs (Conde, 2001; de Murcia et al., 1997). However, PARP-1 knockout in these mice does not lead to an associated increase in accumulations of mutations or tumours, due to a resulting increase in error-free repair of DNA double strand breaks by the homologous recombination pathway (de Murcia et al., 1997). However, in tumour cells containing mutations in the tumour-suppressing BRCA1 or BRCA2 genes associated with familial breast cancer (in which repair via homologous recombination is defective) it was subsequently discovered that inhibition of PARP-1 enzymatic activity by small-molecule PARP inhibitors, in conjunction with addition of alkylating agents or radiation, leads to a highly selective collapse in genome integrity in comparison with healthy cells (Bryant et al., 2005; Farmer et al., 2005). It was suspected that in the absence of PARP-1, unrepaired single-strand breaks collapse replication forks into double strand breaks, triggering repair via homologous recombination, and that selective toxicity in the mutant BRCA cells resulted from the accumulation of unrepaired double-strand breaks that would normally be repaired by homologous recombination. These findings represented the first demonstration of the synthetic lethality of PARP-1 inhibition *in vivo*.

Intriguingly, it was later shown by the group of Yves Pommier that the cytotoxicity of alkylating agents towards the tumour cells was more strongly enhanced by addition of PARP inhibitors than it was by siRNA-mediated knockdown of PARP-1 (Murai et al., 2012). This demonstrated that it is not simply the absence of signalling via ADP-ribosylation and lack of repair of single strand breaks by PARP-1 that causes cytotoxicity, but that the behaviour of PARP-1 itself is modified upon binding to small-molecule inhibitors in a manner which results in an increase in genomic instability in the absence of functional HR machinery. They showed that the cytotoxicity of PARP inhibitors *in vivo* arises from trapping of PARP-1 or PARP-2 enzymes at the sites of DNA damage; since inhibited PARP fails to be released from the DNA, it remains in place forming a bulky, toxic lesion that eventually collapses to form a double strand break. They also found that PARP inhibitors varied in their ability to form these trapped PARP lesions. They proposed that this could be due to prevention of automodification (thought to be required for release) upon binding of PARP inhibitors to the catalytic domain, but they also hypothesised that a component of the effect might arise because binding of inhibitors to the NAD^+ acceptor site of the PARP signature fold could enhance the binding affinity of PARP-1 to DNA via a reverse allosteric mechanism, in which the inter-domain and DNA-binding interfaces formed following

detection of DNA damage could be strengthened. Inhibitors could vary in their ability to promote this reverse-allosteric mechanism depending on factors such as their steric bulk and the precise mode of catalytic domain binding, although the detailed nature of such proposed differences remained unclear. The compounds veliparib, olaparib (the first PARP inhibitor to be approved by the FDA as a targeted therapy for breast and ovarian cancer in people with hereditary *BRCA1* or *BRCA2* mutations under the name Lynparza) and niraparib were initially tested by the Pommier group, with veliparib exhibiting negligible trapping activity and olaparib and niraparib exhibiting greater and equivalent trapping and cytotoxicity in *vivo* (Murai et al., 2012). Another PARP inhibitor with significantly higher potency for trapping, talazoparib, was later discovered (Shen et al., 2013). Talazoparib remains the most active PARP inhibitor (as defined by cytotoxicity and trapping potency) discovered to date. It has been estimated that talazoparib is approximately 100-fold more effective at forming trapped PARP-1 DNA complexes than olaparib (Murai et al., 2014). The chemical structures, *in vivo* cytotoxicity, and results of published biochemical and kinetic assays for these PARP inhibitors are summarised later in this thesis in Section 5.1.

1.2.2 Structures of PARP inhibitors in complex with PARP-1

The majority of PARP inhibitors discovered so far, including those described above, mimic the binding mode of the nicotinamide moiety of NAD^+ to the ART subdomain of the PARP-1 catalytic domain, containing groups that form hydrogen bonds with key residues in the conserved NAD^+ binding pocket (Steffen et al., 2013). Many crystal structures of the PARP inhibitors veliparib, olaparib and talazoparib have been published in complex with PARP catalytic domains from across the polyenzyme PARP family, but to date the only inhibitor to have been solved in complex with the entire PARP-1 catalytic domain is talazoparib (Aoyagi-Scharber et al., 2014). The structure of olaparib in complex with PARP-1 ART subdomain (i.e. catalytic domain ΔHD) has been solved by the Pascal group (Dawicki-McKenna et al., 2015) and structures of veliparib in complex with PARP-2 catalytic domain (Karlberg et al., 2010) and PARP-5b (Tankyrase 2) ART subdomain (Qiu et al., 2014), but not with PARP-1 CAT domain, have been deposited in the Protein Data Bank (PDB). Consequently, meaningful comparison of structural differences in binding of these compounds to PARP-1 and their relation to trapping ability or other mechanisms of potency up to this point has been challenging. The structure of talazoparib in complex with PARP-1 catalytic domain (Aoyagi-Scharber et al., 2014) revealed that as well as forming contacts with the nicotinamide binding pocket via hydrogen bonds and π-stacking interactions,

binding occurs at the D-loop in the outer edges of this pocket in the ART subdomain, and to the αF helix of the HD subdomain, a portion of which was identified by the Pascal group as becoming destabilised upon catalytic domain activation (Dawicki-McKenna et al., 2015) (see Figure 1.7). Comparison of binding at these sites with olaparib and veliparib is hindered by the lack of any published structure of PARP-1 in complex with olaparib including the HD subdomain, and the fact that key residues in both the D-loop and αF helix implicated in talazoparib binding differ between PARP-1 and PARP-2 (Aoyagi-Scharber et al., 2014). Additionally, all structural information relating to the binding and mechanism of these compounds to date has been derived from crystal structures, and whilst no major structural rearrangements have been observed, no exploration of potential changes in solvent exposure or dynamics distant from the binding site upon drug binding have been carried out via methods such as NMR spectroscopy.

1.2.3 Proposed mechanisms of PARP inhibition by small molecules

The reasons that trapping occurs *in vivo*, and the question of whether and how it can be reconstituted and studied *in vitro* remain to be fully resolved. It has been shown that veliparib, olaparib, niraparib and talazoparib all potentiate the cytotoxicity of alkylating agents to varying extents, indicating that all four compounds are capable of PARP-1 trapping (Hopkins et al., 2015). Differences in trapping ability were observed *in vivo* and were correlated with strength of catalytic inhibition, but in a non-linear fashion. Although talazoparib was 10-fold more potent than veliparib with respect to catalytic inhibition, it was 10,000 times more potent with respect to trapping activity. However, as part of the same study, no strengthening of PARP-1 binding to DNA upon addition of inhibitors, or vice versa, was observed *in vitro* in biochemical assays or kinetic studies in the absence of NAD^+. The presence of NAD^+ *in vitro* prevented PARP-1 incubated with inhibitors from being released, but in a linear IC_{50} dependent manner, suggesting that there was no reverse allosteric component to trapping for these compounds in their assay, and that variations in trapping ability are due purely to efficacy of catalytic inhibition. It remains unclear why the difference in talazoparib's trapping potency in cells and mouse models is so great compared to the difference in IC_{50} between the compounds, or why talazoparib was the only compound tested which could potentiate the effect of alkylating agents in vivo at concentrations below its IC_{50}.

Nonetheless there is evidence to suggest that some compounds do promote trapping via reverse allostery. Notably, the Pascal group showed using HXMS that binding of the non-hydrolyzable NAD^+ analogue benzamide adenine dinucleotide (BAD) to the catalytic domain in the presence of damaged DNA caused a decrease in HX at interdomain interfaces and sites of DNA binding, and that this was consistent with BAD/NAD^+ binding to already activated PARP-1 shifting the HD further towards an unfolded conformation and then stabilising it (Langelier et al., 2018). Hypothetically, persistence of activated ADP-ribosylated PARP-1 on DNA strand breaks following NAD^+ binding *in vivo* may contribute to more efficient downstream SSB repair. In contrast to the clinical PARP inhibitors above, the larger molecule BAD (molecular weight 662.442 g/mol) does not bind to the catalytic domain of inactivated full-length PARP-1, but rather binds only to PARP-1 ΔHD or to PARP-1 in the presence of DNA damage, indicating that the HD domain in its fully folded state imposes a complete steric block on NAD^+ binding and does not simply disfavour productive catalysis as was previously thought to be possible. However, PARP inhibitors bind to the catalytic domain with extremely high affinity that is much higher than for BAD, typically with nanomolar dissociation constants in comparison to the estimated K_D of 50 μM for BAD. Nonetheless, one PARP inhibitor with a chemical structure that has a high similarity to NAD^+ is EB-47 (Jagtap et al., 2004). John Pascal informed us that his group was unable to obtain a crystal structure of EB-47 in complex with PARP-1 catalytic domain, and we jointly speculated that this could be because binding of EB-47 may destabilise the HD subdomain (unpublished personal communication). The chemical structures of NAD^+, BAD and EB-47 are shown later in this thesis in Section 5.1, Figure 5.1. To date, only crystal structures of EB-47 in complex with PARP-2 ΔHD catalytic domain and tankyrase 2 (PARP-5b), which has no HD subdomain, have been published (Dawicki-McKenna et al., 2015; Haikarainen et al., 2013), demonstrating that EB-47 binds to the NAD^+ binding pocket of the PARP ART fold but providing no insight into the effect of EB-47 binding on the HD subdomain of PARP-1 catalytic domain. Therefore, the potential structural basis of any mechanistic differences between clinically approved PARP inhibitors and compounds such as EB-47 remains to be fully explored.

1.3 Project aims

Therefore, at the start of the work described in this thesis, we wished to address the following key questions. Firstly, we aimed to further characterise the changes described by the group of John Pascal (Dawicki-McKenna et al., 2015) that take place in the HD subdomain of the

catalytic domain following DNA damage detection by PARP-1 using NMR spectroscopy. How exactly does the secondary structure of these helical regions change upon activation and is there further evidence that this deprotection occurs as a result of partial loss of structure or even partial melting, as hypothesised by the Pascal group? Can the underlying causes of these dynamic changes be probed in solution at the level of individual residues (as opposed to peptide fragments) or at timescales faster than those investigated using HXMS (i.e. <10 seconds)?

Secondly, how might binding of the PARP inhibitors described above to PARP-1 catalytic domain lead to "trapping" of PARP-1 at sites of DNA damage, and what is the structural basis of the differences in potency between PARP inhibitors? As binding of these compounds to PARP-1 active sites has provided relatively few insights into these questions, determining whether any dynamic or conformational changes take place in the HD subdomain in these complexes will be key to answering this question. With the exception of talazoparib, to date no structures of complexes with these compounds have been solved by X-ray crystallography that include the PARP-1 HD subdomain, and even in the case of talazoparib there may be important dynamic changes in the HD subdomain that are not observed in the crystal structure (in a manner similar to the difference between the structure of the activated catalytic domain in pdb 4DQY and the HXMS data described above). Furthermore, if binding by any of the compounds described above to the catalytic domain leads to tightening of interdomain and DNA-binding interfaces in the rest of the full-length PARP-1 protein via a reverse allosteric mechanism, this is highly likely to be mediated through the HD subdomain, as formation of these interfaces following DNA binding leads to conformational changes in the HD subdomain.

With these questions in mind, Chapter 3 describes the work that was necessary to complete the sequential resonance assignment of backbone amides in the 360 residue, 40 kDa wild-type catalytic domain fragment, as well as three mutants of the catalytic domain and the wild-type catalytic domain in complex with four different PARP inhibitors. Chapter 4 describes work characterising the differences between the wild-type catalytic domain and these three catalytic domain mutants using NMR spectroscopy, all of which exhibit a constitutive partial DNA-independent increase in activity and are used as a model of PARP-1 activation in the isolated catalytic domain fragment. Chapter 5 describes experiments to characterise, compare and contrast the binding of the PARP inhibitors veliparib, olaparib, talazoparib and EB-47 to the wild-type PARP-1 catalytic domain in solution using NMR spectroscopy.

Chapter Two

Materials and methods

2.1 Molecular cloning

All sequences from human PARP-1 used in this work originated from a plasmid (pQE-T7) containing the DNA sequence coding for the human PARP-1 protein, codon-optimised for bacterial expression (Qiagen).

Initially, the codon-optimised PARP-1 catalytic domain fragment 660-1014 was cloned by Laura Easton into a modified pET28a vector which contained a DNA sequence encoding the *Bacillus stearothermophilus* dihydrolipoamide acetyltransferase lipoyl domain (referred to as lipoyl domain) as an N-terminal fusion tag. The lipoyl domain itself also included an N-terminal His_6 tag. To facilitate cleavage of this fusion tag, when the PARP-1 catalytic domain fragment was subcloned into this vector, a DNA sequence encoding a recognition sequence (ENLYFQG) for Tobacco Etch Virus protease (TEV^{pro}) was inserted each time a new fragment was cloned. However, in the fragment 660-1014 this left only two residues between the protease cleavage site and the beginning of the structured catalytic domain at residue 662, which sterically inhibited cleavage of the N-terminal fusion tag by TEV protease. To resolve this problem, three further versions of this vector were cloned to include additional residues from the CAT-WGR linker to increase the number of residues from the WGR-CAT linker and facilitate improved cleavage by TEV protease (see **Table 2.1**).

In each case, the DNA sequence of the desired length of PARP-1 catalytic domain fragment was amplified by PCR using Pfu Turbo DNA Polymerase (Agilent) according to the manufacturer's instructions and the sequence for the TEV protease recognition motif was added using the forward PCR primer. Following the PCR reaction, the amplified product was purified using a QIAquick PCR Purification Kit (Qiagen). The purified product and pET28a-lip vector were then digested with the restriction enzymes Nde1 and Xho1 (New England Biolabs), and digestion products were purified using agarose gel electrophoresis and a gel extraction kit from Qiagen. The digested vector was then dephosphorylated using alkaline phosphatase (Roche). The digested insert and vector were ligated together during an overnight incubation at 16°C with T4 DNA ligase (New England Biolabs) and the ligated product was transformed into chemically competent DH5α cells.

To achieve this, 10 μL of ligated plasmid vector was added to 50 μL DH5α cells in a round-bottomed 14 mL Polypropylene Falcon tube, gently mixed and left on ice for 30 minutes; the cells and vector were then heat shocked at 42°C for 45 seconds and allowed to recover on ice for 2 minutes before 450 μL SOB medium was added and cells were grown at 37°C in an incubator with shaking. Cells were plated on TYE agar plates containing kanamycin (Formedium, working concentration of 50 μg/mL) and incubated overnight at 37°C. Colonies were sequenced to confirm the presence of the chosen insert in the vector.

Although the TEV protease recognition motif was initially formed using the canonical residue G660 from the PARP-1 WGR-CAT linker, fragments containing a longer section of the CAT-WGR linker that did not begin with a canonical glycine residue from the PARP-1 sequence resulted in a non-canonical glycine at the N-terminus of the catalytic domain fragment following cleavage by TEV protease, as TEV protease cleaves its substrate between the Q and G residues of the recognition sequence. Proteins CAT2 V762 (G-PARP-$1_{656-1014}$), CAT3 V762 (G-PARP-$1_{652-1014}$) and CAT4 V762 (PARP-$1_{646-1014}$) were expressed in rich media and cleavage with TEV protease was equally efficient for all three extended linker lengths.

Additionally, since the codon-optimised gene that had originally been obtained from Qiagen corresponded to protein that differed by a single amino-acid from that used in the Pascal group (both are naturally occurring isoforms of PARP-1), in order to be consistent with their work we replaced valine 762 with alanine in the HD subdomain of all catalytic domain fragments. This change was introduced by cloning via PCR with Pfu Turbo DNA Polymerase using a site-directed mutagenesis approach based on Strategene's Quikchange II methodology. Following the mutagenesis reactions, parent DNA templates were digested using Dpn1 endonuclease (New England Biolabs). Reactions were subsequently transformed into DH5α cells and clones from TYE agar plates containing kanamycin were sequenced to confirm that the desired mutagenesis had taken place. All vectors cloned with this mutation are summarised in Table 2.1. CAT2, which contained the V762A mutation and the shortest of the extended linkers (G-PARP-$1_{656-1014}$) was considered to be the optimal fragment to use for studies of wild-type PARP-1 catalytic domain. Attempts to carry out mutagenesis of CAT4 V762 to alanine were repeatedly unsuccessful and were abandoned once CAT2 was found to be optimal. As used from this point onwards in this thesis, the abbreviation "CAT" implies that position 762 is alanine; on those occasions that it is

necessary to refer explicitly to fragments or constructs carrying valine at position 762, these will be referred to as CAT V762.

Name	Number	Protein	Vector	Fusion proteins, purification tags, protease cleavage sites
PARP-1 FL	216	PARP-1_{1-1014}	pQE-T7	His_6
CAT V762	217	PARP-$1_{660-1014}$ V762	pET28a-lip	His_6, lipoyl, TEV protease
CAT	218	PARP-$1_{660-1014}$ V762A	pET28a-lip	His_6, lipoyl, TEV protease
CAT2 V762	224	G-PARP-$1_{656-1014}$	pET28a-lip	His_6, lipoyl, TEV protease
CAT3 V762	225	G-PARP-$1_{652-1014}$	pET28a-lip	His_6, lipoyl, TEV protease
CAT4 V762	226	PARP-$1_{646-1014}$	pET28a-lip	His_6, lipoyl, TEV protease
CAT2	227	G-PARP-$1_{656-1014}$ V762A	pET28a-lip	His_6, lipoyl, TEV protease
CAT3	228	G-PARP-$1_{652-1014}$ V762A	pET28a-lip	His_6, lipoyl, TEV protease
CAT mutant L713F	230	G-PARP-$1_{656-1014}$ V762A L713F	pET28a-lip	His_6, lipoyl, TEV protease
CAT2 mutant L765F	231	G-PARP-$1_{656-1014}$ V762A L765F	pET28a-lip	His_6, lipoyl, TEV protease
CAT2 mutant L765A	232	G-PARP-$1_{656-1014}$ V762A L765F	pET28a-lip	His_6, lipoyl, TEV protease
TEV^{pro}	N/A	(TEV protease "SuperTEV" from Mark Allen with N-terminal lipoyl domain)	pRET3a	His_6, lipoyl
CAT2 E988A	N/A	G-PARP-$1_{656-1014}$ V762A E988A	pET28a-lip	His_6, lipoyl, TEV protease

Table 2.1 Summary of protein constructs used in this thesis. All fusion proteins, purification tags and protease cleavage sites were located at the N-terminus.

Constitutively partially active CAT mutants L713F, L765F and L765A were also cloned using the Quikchange II methodology. Reactions were performed with KOD Hot Start DNA Polymerase (Novagen) and overhanging primers for site-directed mutagenesis were designed by Chris Tate using the program Optimus Primer (Brown et al., 2010). All corresponding fragments of PARP-1 catalytic domain (including any artefacts from molecular cloning) are summarised in Table 2.1. From this point onwards, reference to wild-type catalytic domain refers to the protein fragment CAT2, post cleavage of N-terminal His_6-lipoyl fusion protein.

The plasmid used for bacterial expression of His_6-Lipoyl-TEV protease was a gift from Mark Allen and based on a pRET3a plasmid which confers ampicillin resistance.

2.2 Protein expression and purification using *E. coli*

All PARP-1 proteins were expressed in the *Escherichia coli* strain BL21 (DE3).

2.2.1 Expression of unlabelled PARP-1 catalytic domain

The plasmid CAT2 encoding wild-type PARP-1 catalytic domain was transformed into BL21 (DE3) cells using the protocol described above and plated on TYE agar plates containing kanamycin (Formedium, working concentration of 50 µg/mL). Plates were incubated overnight at 37°C. A single colony was picked from the plate and used to inoculate 50 mL LB medium supplemented with kanamycin, which was subsequently grown overnight at 37°C with shaking at 200 rpm. Starter cultures were used to inoculate LB medium supplemented with kanamycin, at a ratio of 1:40. For expression of unlabelled protein, 500 mL of expression medium was used per 2L flask to ensure adequate aeration of cells during growth and expression of recombinant protein. Following inoculation, expression cultures were grown at 37°C with shaking at 200 rpm until OD_{600} was between 0.8 and 1.0. Protein expression was then induced by addition of 1 M isopropyl ß-D-1-thiogalactopyranoside stock (IPTG, Melford) to give a final concentration of 2 mM and cells were incubated overnight (minimum of 16 hours) at 30°C with shaking at 200 rpm. Cells were pelleted by centrifugation at 4000 rcf for 20 minutes at 4°C in 1L polycarbonate Beckman bottles in a JLA 8.1000 rotor and pellets were resuspended in ice-cold PBS before a second centrifugation step was carried out in 50 mL Falcon tubes as described above. The supernatant was discarded and pellets were frozen at -20°C.

2.2.2 Expression of unlabelled PARP-1 catalytic domain mutants

Catalytic domain mutants were expressed and purified based on an approached outlined for production and purification of full-length human PARP-1 from *E. coli* (Langelier et al., 2017). Following transformation with plasmids encoding catalytic domain mutants or full-length PARP-1, cells were plated on TYE agar plates containing kanamycin and 10 mM benzamide (Sigma-Aldrich). To prepare TYE agar with kanamycin and benzamide, autoclaved TYE agar was kept at 50°C in a water bath, kanamycin was added and benzamide powder was dissolved to give a final concentration of 10 mM. Expression of unlabelled catalytic domain mutant proteins then followed a similar procedure to expression of wild-

type catalytic domain, but with all media supplemented with 10 mM benzamide. Expression was induced by addition of IPTG to give a final concentration of 0.5 mM and expression cultures were grown overnight at 20°C with shaking at 200 rpm. Cells were harvested as described for wild-type PARP-1 catalytic domain.

2.2.3 Expression of isotope-labelled catalytic domain and mutants

Expression of uniformly-labelled ^{15}N and/or ^{13}C isotopically labelled proteins (without deuteration) was carried out in M9 minimal media (composition in Table 2.2) supplemented with $^{15}NH_4Cl$ and/or $^{13}C_6$-glucose instead of their natural abundance equivalents. BL21 (DE3) cells transformed with plasmid encoding wild-type or mutant catalytic domain were plated and starter cultures inoculated as described for unlabelled expression. Starter cultures were then pelleted by centrifugation (4000 rcf, 20 mins at 4°C), supernatant discarded and pellets washed gently with ice-cold phosphate-buffered saline (PBS) and re-suspended into minimal media which was used to inoculate 500 mL minimal media per 2L flask. As with expression of unlabelled protein, starter culture media and minimal media for expression of isotope-labelled catalytic domain mutants contained 10 mM benzamide. Cells were grown at 37°C with shaking at 200 rpm until OD_{600} was between 0.8 and 1.0 and expression and harvesting was carried out for each protein as described for its unlabelled counterpart.

Expression of perdeuterated wild-type and mutant catalytic domains in conjunction with uniform ^{15}N and ^{13}C labelling was based on the approach described in (Marley et al., 2001) in which a high cell mass is cultivated in unlabelled rich media before *E. coli* cells are re-suspended into a smaller volume of isotopically enriched minimal media and recombinant protein expression is induced. Starter cultures were inoculated as described for unlabelled protein expression and used to inoculate 4x desired expression culture volume of LB medium containing kanamycin. Cells were grown at 37°C with shaking at 200 rpm until OD_{600} reached 0.6. Cells were then pelleted via centrifugation at 4000 rcf for 20 minutes at 4°C and pellets were gently washed with ice-cold PBS and resuspended in M9 minimal media containing ^2H_2O (Sigma-Aldrich), $^{15}NH_4Cl$ and $^{13}C_6$,$^2H_7$-glucose. Cells were incubated for 45 minutes at 37°C with shaking at 200 rpm and a further 45 minutes at the desired expression temperature (37°C for wild type catalytic domain and 20°C for catalytic domain mutants). Expression of recombinant protein was induced by adding IPTG to a final concentration of 2 mM.

Minimal Media	
Reagent	Concentration
Na_2HPO_4	50 mM
KH_2PO_4	25 mM
NaCl	10 mM
Thiamine	40 mg/L
$MgSO_4$	4 mM
$CaCl_2$	0.2 mM
Trace Elements	1x
Vitamin Mix	1x
NH_4Cl	1g/L
Glucose	4g/L

100x Trace Elements	
Reagent	Concentration (g/L)
Na_4EDTA	4
$FeCl_3 \cdot 6H_2O$	0.83
$ZnCl_4$	0.05
$CoCl_2 \cdot 6H_2O$	0.01
$CuCl_2$	0.01
H_3BO_3	0.01
$MnCl_2 \cdot 4H_2O$	1.6

1000x Vitamin Mix	
Reagent	
Biotin	All 1g/L
Choline	
Folic Acid	
Niacinamide	
D-pantothenate	
Pyridoxal	
Riboflavin	
B_{12}	

Table 2.2 Recipes used for the production of M9 minimal media, required for the expression of isotope-labelled proteins in *E. coli*.

Selective amino acid labelling of wild-type and mutant PARP-1 catalytic domains was also carried out using this approach in H_2O-based M9 minimal media containing $^{14}NH_4Cl$ and $^{12}C_6$-glucose. For selective ^{15}N labelling of lysine and arginine residues, M9 minimal media was supplemented with 0.1 g/L L-Lysine-α-^{15}N dihydrochloride, (Cambridge Isotope Laboratories), or L-Arginine-$^{15}N_4$ hydrochloride (Sigma-Aldrich) respectively. For selective ^{15}N labelling of leucine and isoleucine residues as well as joint selective ^{15}N labelling of glycine and serine residues, expression media was supplemented with 0.1g/L of all non-target proteinogenic amino acids respectively (all from Sigma-Aldrich) and 15 minutes prior to induction 0.1g/L L-Leucine-^{15}N, L-Isoleucine-^{15}N or Glycine-^{15}N respectively (all from Cambridge Isotope Laboratories) and 0.5 g/L of non-target unlabelled amino acids (including serine for joint glycine-serine labelling) were added to the media.

2.2.4 Purification of human PARP-1 catalytic domain and mutants

Wild-type and mutant PARP-1 catalytic domains were purified via the following steps: Ni-affinity purification, cleavage of His_6-tagged lipoyl domain using TEV protease, removal of TEV protease, cleavage product and uncleaved protein via a second Ni-affinity step and a final gel filtration step.

Cell pellets were resuspended in 35 mL Buffer A+. Cell pellets harvested from culture volumes of 4L or below were lysed by sonication (microprobe, 70% amplitude, 3s

on and 7s off at 4°C for 6 minutes total on-time). For pellets from culture volumes of greater than 4L, cells were re-suspended in 200 mL lysis buffer and cells were lysed by three passes through a cell disruptor (Constant) at 25 kpsi. Total cell lysate was cleared via centrifugation (27,000 rcf, 45 min, 4°C) and supernatant was incubated with 5 mL 50% Ni-NTA agarose resin (Qiagen) re-suspended in the chosen lysis buffer (following 2x wash steps in Millipore water) for 1.5 hours on a roller at 4°C. Resin was loaded onto a glass chromatography column (Bio-Rad) and flow-through was discharged by gravity and collected. The resin was washed with 2x 50 mL Wash Buffer and 50 mL High Salt Wash Buffer. Three elution steps in Elution Buffer were performed: a 10 mL step, a 5 mL step and a 5 mL step with a five-minute incubation period to make a total elution volume of 20 mL. Elution fractions were pooled and 2 mg His-tagged TEV protease was added to the sample. The mixture was then dialysed overnight in 1L HepA+ buffer in a dialysis membrane with a 3500 molecular weight cut-off (Snakeskin, Thermo Scientific).

Following dialysis, SDS-PAGE was carried out to verify completed cleavage of N-terminal His_6-lipoyl domain and the mixture was incubated with 2 mL 50% NiNTA re-suspended in HepA+ (following 2x wash steps in Millipore water) on a roller at 4°C for one hour to bind uncleaved protein, TEV protease and His_6-lipoyl domain (the last of these being the unwanted product of the cleavage by TEV protease). The mixture was passed through a Bio-Rad Poly-Prep column via centrifugation in a 50 mL Falcon tube and the flow-through containing cleaved PARP-1 catalytic domain was collected. The resin was then washed with 3 mL HepA+. The flow-through and wash were checked using SDS-PAGE to ensure that uncleaved protein, TEV protease and His_6-lipoyl domain had been removed, then pooled and concentrated in a Vivaspin 20 with a 10000 molecular weight cut-off (Sartorius Stedium) before purification via gel filtration. The sample was loaded onto a HiLoad 26/60 Superdex 75 or 200 gel filtration column (GE Healthcare) equilibrated with GF150+ gel filtration buffer and run at 2 ml/min with 4 mL fractions collected immediately following the void volume. Collected fractions were analysed using SDS-PAGE and fractions containing pure CAT domain were pooled. The concentration of pooled fractions was measured via A280 on a Nanodrop 2000 (Thermo Scientific). Pooled fractions were concentrated in a Vivaspin 20 as described above. Samples prepared for NMR spectroscopy were exchanged into CAT NMR Buffer by 5 repeated 10 x concentrations and dilutions in a Vivaspin 2 with a 10,000 molecular weight cut-off (Sartorius Stedium). For samples prepared for applications other than NMR, 20% glycerol was added and the samples were snap frozen in 2 mL Eppendorf tubes using dry ice and stored at -80°C.

Buffer	Composition
Buffer A+	50 mM Na-HEPES pH 8.0, 0.5 M NaCl, 0.5 mM TCEP, 1 mM PMSF, Roche EDTA-free Complete protease inhibitor tablets (1 tablet per 2L culture)
Wash Buffer	50 mM Na-HEPES pH 8.0, 0.5 M NaCl, 0.5 mM TCEP, 1 mM PMSF
High Salt Wash Buffer	50 mM Na-HEPES pH 8.0, 1 M NaCl, 0.5 mM TCEP, 1 mM PMSF
Elution Buffer	50 mM Na-HEPES pH 8.0, 0.5 M NaCl, 0.3 M imidazole, 0.5 mM TCEP, 1 mM PMSF
Hep A+ Buffer	50 mM Na-HEPES pH 8.0, 0.2 M NaCl, 0.5 mM TCEP, 5% glycerol, Roche EDTA-free Complete protease inhibitor tablets (1 tablet per litre dialysis buffer)
GF150+	50 mM Na-HEPES pH 8.0, 0.15 M NaCl, 0.5 mM TCEP
Low Salt Wash Buffer	50 mM Na-HEPES pH 8.0, 0.5 M NaCl, 10 mM imidazole, 0.5 mM TCEP, 1 mM PMSF
TEV Buffer A	50 mM Tris-HCl pH 8.0, 0.3 M NaCl, 2 mM β-mercaptoethanol, 1 mM PMSF
TEV Elution Buffer	50 mM Tris-HCl pH 8.0, 0.3 M NaCl, 0.1 M imidazole, 2 mM β-mercaptoethanol, 1 mM PMSF
2x TEV Dialysis Buffer	50 mM Tris-HCl pH 7.5, 5 mM DTT

Table 2.3 Summary of all buffers used in protein purifications. Abbreviations: HEPES: (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) Tris: Tris(hydroxymethyl)aminomethane TCEP: Tris(2-carboxyethyl)phosphine hydrochloride PMSF: Phenylmethylsulfonyl fluoride. DTT: Dithiothreitol. EDTA: Ethylenediaminetetraacetic acid.

2.2.5 Expression and purification of TEV protease

Cells from *E. coli* strain C41 (DE3) were transformed with the plasmid containing the DNA sequence of His_6-lipoyl TEV protease, plated on TYE agar plates containing ampicillin (Formedium, 100 μg/mL) and grown overnight in an incubator set to 37°C. Plates were scraped directly into 3L 2xTY medium supplemented with ampicillin (1 plate per 1L medium in a 2L flask) and cultures were grown at 37°C until OD_{600} reached 0.5. The temperature was lowered to 22°C and expression was induced via addition of 0.5 mM IPTG. Cells were harvested 4-6 hours after expression, as described for PARP-1 catalytic domain fragments.

Cell pellets were re-suspended in TEV Buffer A, then cells were lysed and lysate clarified as described for PARP-1 catalytic domain. The supernatant was incubated for 1 hour on a roller at 4°C with 10 mL 50% Ni-NTA pre-washed with Millipore water and equilibrated in TEV Buffer A. Resin was loaded onto a glass chromatography column (Bio-Rad) and flow-through was discharged by gravity and collected. TEV protease was eluted via 3 x 15 mL washes with TEV Elution Buffer with a five-minute incubation before

collection of each fraction. Eluates were pooled and dialysed into 2x TEV Dialysis buffer overnight and concentrated to 4 mg/mL before addition of 50% glycerol, snap freezing and storage at -20°C.

2.3 NMR spectroscopy

NMR data was acquired at the MRC-LMB MRS facility on Bruker Avance III HD 800, Avance III 600 or Avance I 600 spectrometers. Some data was also acquired at the MRC Biomolecular NMR centre at the Francis Crick Institute, London on a Bruker Avance III HD 950. All spectrometers used were equipped with triple resonance ($^{15}N/^{13}C/^{1}H$) cryoprobes. Unless otherwise specified, all NMR experiments performed using wild type PARP-1 catalytic domain fragment or mutants were performed at 25°C in NMR Buffer containing 50 mM Tris-d11-HCl pH 7.0 (at 25°C), 50 mM NaCl, 2 mM DTT d-10, 0.02% NaN_3 and 95:5 H_2O:D_2O in 5 mm NMR tubes (Norell).

2.3.1 Triple-resonance backbone experiments

TROSY-based HNCA, HN(CO)CA, HNCACB and HN(CO)CACB triple-resonance backbone experiments (Salzmann et al., 1998; Eletsky et al., 2001) were performed by Ji-Chun Yang on wild-type catalytic domain and catalytic domain mutant L713F using non-uniform sampling (NUS). For TROSY-based HNCA and HNCACB experiments, between 25% and 30% of data was collected under the NUS scheme, and for TROSY-based HNCOCA and HNCOCACB experiments, the data collection percentage was 20%. For wild-type catalytic domain, experiments were carried out on perdeuterated ^{15}N ^{13}C labelled wild-type catalytic domain (200 μM) in NMR Buffer at 30°C. For L713F mutant, experiments were carried out on both perdeuterated ^{15}N ^{13}C labelled (200 μM) and non-deuterated ^{15}N ^{13}C domain in NMR Buffer at 25°C. Data was processed using the programmes MddNMR (Orekhov et al., 2011) and NMRPIPE (Delaglio et al., 1995). Analysis of data, including sequential assignment of wild-type catalytic domain and catalytic domain L713F, was performed using the programme CcpNMR Analysis 2.4 (Vranken et al., 2005).

2.3.2 ^{15}N NOESY-HSQC experiments

^{15}N NOESY-HSQC experiments (Sklenář et al., 1993) with a mixing time of 70 ms were recorded from 400 μM samples of ^{15}N-labelled samples of wild-type PARP-1 catalytic domain, L713F, L765F and L765A catalytic domain mutants, and samples of wild-type

PARP-1 catalytic domain in complex with a 1:1.5 ratio of veliparib (300 μM catalytic domain), olaparib (400 μM), talazoparib (500 μM) and EB-47 (400 μM). All samples were made in NMR Buffer and catalytic domain samples were titrated with 50 mM stock solutions of PARP inhibitors in 100% DMSO-d_6 to a final ratio of 1:1.5 to ensure that the inhibitor was present in excess. Data for all samples was collected at 25°C on a Bruker Avance III HD 800 spectrometer, with the exception of the catalytic domain mutant L765F, for which data was collected on a Bruker Avance III HD 950 at the MRC Biomolecular NMR centre at the Francis Crick Institute, London. Both spectrometers were equipped with triple-resonance cryoprobes ($^{15}N/^{13}C/^{1}H$). Data analysis, including sequential assignment of catalytic domain mutants L713F, L765F, L765A and wild-type catalytic domain in complex with the PARP inhibitors veliparib, olaparib, talazoparib and EB-47, was performed using the programme CcpNMR Analysis 2.4 (Vranken et al., 2005).

2.3.3 Chemical shift perturbation

Analyses of backbone amide chemical shift perturbations (CSPs) relative to wild-type catalytic domain were performed for ^{15}N-labelled catalytic domain mutants and ^{15}N-labelled wild-type catalytic domain following titration with PARP inhibitors. In all cases, ^{15}N and ^{1}H chemical shifts of backbone amides from [^{15}N-^{1}H]-TROSY spectra were recorded and the chemical shift perturbation Δδ (measured in ppm) was calculated using the formula $\Delta\delta = ((\Delta\delta(^{1}H))^2+(\Delta\delta(^{15}N)/5)^2)^{1/2}$ (Williamson, 2013). The scaling of the ^{15}N shifts by a factor of 1/5 was used to compensate approximately for the different chemical shift ranges for the ^{1}H and ^{15}N chemical shifts; the same scaling factor was used for all amino acids and analyses of CSPs described in this dissertation. All CSP analyses described in this dissertation used amide chemical shifts from [^{15}N-^{1}H]-TROSY spectra acquired on a Bruker Avance III HD 800 spectrometer under the conditions described at the beginning of this section.

Chemical shift perturbation analyses of ^{15}N wild-type PARP-1 catalytic domain and ^{15}N PARP-1 catalytic domain mutants were achieved by comparing the amide chemical shifts assigned to each residue measured from [^{15}N-^{1}H]-TROSY spectra of wild-type and mutant domains.

Chemical shift perturbation analyses of ^{15}N-labelled wild-type PARP-1 catalytic domain upon binding to selected PARP inhibitors was achieved by titrating the catalytic domain to a 1:1.5 protein:inhibitor ratio (to ensure complete saturation) with veliparib, olaparib, talazoparib (all from Selleck) or EB-47 (Sigma-Aldrich) solubilised in deuterated

dimethyl sulfoxide (DMSO-d_6, Aldrich), and comparing the amide chemical shifts from [^{15}N-^{1}H]-TROSY spectra. Titrations were performed with an appropriate concentration of inhibitor to ensure that the final concentration of DMSO-d_6 did not exceed 5%, and a control titration was carried out to confirm that addition of DMSO-d_6 alone did not cause any CSPs.

Chemical shift perturbations from catalytic domain mutants and inhibitor titrations were mapped onto the structure of wild-type catalytic domain and complexes of the catalytic domain and inhibitors for display in Pymol using a combination of in-house scripts written by members of the Neuhaus group. First, CSP values exported directly from CCPN analysis were globally normalised to values between zero and one (using a shell script written by David Neuhaus), with the largest chemical shift perturbation found across all mutant and inhibitor complexes being set to one. This allowed for later simultaneous visual comparison between the mapping results across all datasets. For each dataset in turn, these scaled CSP values were then used as input to a python script (written by Christoph Brockmann) that converted them into RBG colour settings corresponding to a colour ramp running between two pre-defined colours. In this work, a CSP of zero corresponded to grey (RGB 0.6,0.6, 0.6) and a CSP of one corresponded to yellow (RGB 1.0, 1.0, 0.0). The output of this python script is a Pymol-command script (.pml file) that, when run in a Pymol session where the relevant structure has been pre-loaded, sets the colour for each residue to the RGB values calculated according to the defined colour ramp. For CSPs from catalytic domain mutants and wild-type catalytic domain in complex with EB-47, for which no solved crystal structures exist, globally scaled CSP values were mapped onto the crystal structure of wild-type PARP-1 catalytic domain (pdb 4DQY). For CSPs from wild-type PARP-1 catalytic domain in complex with talazoparib, globally scaled CSP values were mapped onto pdb 4PJT, a published crystal structure of talazoparib bound to the catalytic domain Aoyagi-Scharber et al., 2014). We gratefully acknowledge Marianne Schimpl at AstraZeneca who very kindly supplied us with the co-ordinates of unpublished crystal structures of PARP-1 catalytic domain in complex with veliparib or olaparib, which were used to display the mapping results for CSPs from wild-type PARP-1 catalytic domain in complex with these compounds.

2.3.4 ^{15}N relaxation analysis

Backbone ^{15}N longitudinal relaxation time (T_1), spin-locked transverse relaxation time ($T_{1\rho}$) and ^{15}N{^{1}H} steady state NOEs were determined for ^{15}N-labelled samples of wild-type

PARP-1 catalytic domain, L713F, L765F and L765A catalytic domain mutants, and samples of wild-type PARP-1 catalytic domain in complex with veliparib, olaparib, talazoparib and EB-47, using TROSY-based pulse sequences described by Lakomek et al. (2012). The spin lock field strength for all $T_{1\rho}$ experiments was 1.56kHz and all $T_{1\rho}$ measurements were corrected for the off-resonance tilted field using the equations described in Lakomek et al. (2012). All samples were made in NMR Buffer at a concentration of 400 µM and catalytic domain samples were titrated with 50 mM stock solutions of PARP inhibitors in 100% DMSO-d_6 to a final ratio of 1:1.5 to ensure that the inhibitor was present in excess. Data was collected at 25°C on a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{15}N/^{13}C/^{1}H$). T1 data was recorded for all samples using delays of 0, 120, 120, 280, 480, 800, 1200, 1200, 2000, 3200, 4800 and 6800 ms and T1ρ data was recorded for all samples using delays of 0.15, 2, 4, 4, 7, 12, 20, 34, 50 and 100 ms. Relaxation times were derived in Sparky using non-linear least square fitting of peak heights at different delays. In order to determine $^{15}N\{^{1}H\}$ steady state NOEs, CcpNMR Analysis 2.4 (Vranken et al., 2005) was used to calculate the peak intensity ratios in pairs of spectra recorded with and without ^{1}H saturation. A saturation time of seven seconds was used for steady state NOE experiments. For $T_{1\rho}$ values, error bars were plotted according to the data fitting error in Sparky. For heteronuclear NOE, error bars were plotted according to standard deviation determined on the basis of background noise levels as described in Farrow et al. (1994). τ_c values for backbone amides in all samples described above were calculated from T_1 and $T_{1\rho}$ values using the formula $\tau_c \approx \frac{1}{4\pi\upsilon_N}\sqrt{6\frac{T_1}{T_{1\rho}} - 7}$, where υ_N is the ^{15}N resonance frequency in Hz (Rossi et al., 2010).

2.3.5 CLEANEX experiments

Experiments were recorded for 400 µM ^{15}N-labelled samples of free wild-type PARP-1 catalytic domain and PARP-1 catalytic domain in complex with veliparib, olaparib, talazoparib and EB-47 using a phase-modulated CLEANEX pulse sequence (Hwang et al., 1998) with a mixing time of 150 ms. Data was collected at 25°C on a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{15}N/^{13}C/^{1}H$). Peak heights were normalised according to the difference between the average peak heights of N-terminal residues 657,658 and 660 relative to a 400 µM ^{15}N TROSY reference sample to give a measure of CLEANEX Relative Intensity.

2.3.6 $^{2}H_2O$ exchange experiments

Solvent exposure of backbone amides over a timeframe of 48 hours was compared in ^{15}N-labelled wild-type catalytic domain and ^{15}N-labelled wild-type catalytic domain in complex with PARP inhibitors. 500 µL, 100 µM samples of ^{15}N-labelled wild-type catalytic domain in H_2O-based NMR Buffer were snap frozen in dry ice and lyophilised overnight (>16 hours) in a 2 mL Eppendorf tube. Pellets were re-suspended in $^{2}H_2O$, samples were transferred to 5 mm NMR tubes and a repeating series of 16 [^{15}N-^{1}H]-TROSY experiments was immediately started with the aim of measuring the height of backbone amide peaks at repeating intervals over a period of 48 hours. For complexes of catalytic domain with PARP inhibitors, PARP inhibitor solubilised in 100% d_6-DMSO was added immediately after re-suspension of the pellet in $^{2}H_2O$ to give a final concentration of 150 µM (a 1:1.5 ratio to ensure complete saturation, as described above). Data was collected at 25°C on a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{15}N/^{13}C/^{1}H$). The heights of peaks from TROSY spectra were measured in CCPN analysis 2.4 (Vranken et al., 2005. A measure of relative residual intensity for every sample at 6-hour, 15-hour and 42-hour time points was calculated by scaling peak heights between TROSY spectra in order to normalise variations in concentration (according to the 1D proton spectra of the samples) and number of scans. Data from $^{2}H_2O$ and CLEANEX experiments was mapped onto the structure of wild-type catalytic domain and structures of the catalytic domain in complex with the relevant inhibitors for display in Pymol using a script written by David Neuhaus that coloured residues light blue (6 hours), blue (15 hours) or dark blue (42 hours) according to the latest time point for which signals were observed or red where signals were recorded from CLEANEX experiments. For residues where no signals were observed from $^{2}H_2O$ and CLEANEX experiments, residues were coloured white.

2.4 Thermal stability measurements

The thermal stability of wild-type PARP-1 catalytic domain and the catalytic domain mutants L713F, L765F and L765A was assessed by nano differential scanning fluorimetry (nanoDSF) using a Prometheus NT.48 instrument manufactured by NanoTemper. 500 µM samples of wild-type and mutant catalytic domain samples in a Tris-based buffer (50 mM Tris pH 7.0, 50 mM NaCl, 2 mM DTT) were loaded into capillaries, heated by the Prometheus NT.48 (30% excitation, 15-75°C, increasing at a rate of 2°C per minute) and scattering was recorded as a measure of protein denaturation. Scattering and T_{onset} values were exported from the Nanotemper software PR.ThermControl.

Chapter Three

Resonance assignment of PARP-1 catalytic domain

3.1 Sequential assignment and difficulties of assigning large proteins

Interpretation of data acquired from isotopically labelled proteins using NMR, whether in the form of chemical shift perturbations or dynamic experiments, is contingent upon resonance assignment. Following acquisition of 2D HSQC or TROSY spectra for a target protein, or relaxation or solvent-exchange data-based experiments, it is necessary to identify which two-dimensional cross peak corresponds to each backbone amide in the sequence of the protein. Only once this has been achieved can changes in chemical shifts, relaxation times, rates of solvent exchange or results from any other two-dimensional experiments be linked to a particular residue in the protein, allowing biological interpretation of the results at the level of individual residues.

Assignment of backbone amides using a sequential assignment approach requires acquisition of a set of three-dimensional triple-resonance spectra, each of which contain a ^{1}H, a ^{15}N and a ^{13}C dimension. These experiments transfer magnetisation between directly bonded nuclei, which have large scalar (through-bond) couplings, and can therefore correlate the chemical shifts of each NH group with the chemical shifts of the Cα or Cβ atoms in either the same residue (referred to as i) and the preceding residue (i-1), or the preceding residue only. This allows the Cα, Cβ and NH of the preceding residue to be identified and a chain of NH chemical shifts that are progressively linked to be built up. In principle, assignment of the residues in the entire sequence can be inferred from the Cα and Cβ chemical shifts of every residue in the chain, as these chemical shifts are highly characteristic according to the amino acid type. Ambiguity between amino acid types with similar Cα and Cβ chemical shifts can be resolved using less ambiguous chemical shifts from other residues in the sequence such as alanine, glycine, serine or threonine, with the aim being to complete the assignment of the chain as a whole.

As the molecular weight of a protein increases, resonance assignment becomes progressively more difficult. For proteins of increasing size, tumbling rates in solution are progressively slower, shortening the relaxation time of the NMR signal and reducing sensitivity of the experiments. Interpretation of the spectra also becomes more challenging as an increase in the number of NMR-visible nuclei in the protein results in greater

complexity of the spectra and leads to more overlap of peaks, in both 2D and in 3D experiments.

With this in mind, a range of strategies have been developed to aid the assignment of challenging proteins. These can take the form of isotopic labelling strategies, new NMR experiments or a combination of the two, and may aim to either improve the quality of the data collected, reduce spectral congestion or overlap in the recorded spectra, or both.

Most significantly, for NMR spectroscopy of proteins greater than 25 kDa in size, the data quality of triple-resonance spectra for assignment of backbone amides is greatly enhanced by deuteration of the target protein, in which spin-½ ^{1}H nuclei (protons) are substituted for nuclei of the heavier stable hydrogen isotope deuterium (^{2}H) that contain a proton and a neutron and have a spin quantum number of 1. Deuterium nuclei are not visible in proton NMR, but due to their smaller dipole moment, their presence in place of hydrogen in high molecular weight proteins very greatly reduces the contributions of proton-proton dipolar interactions and scalar couplings to the linewidths of remaining protons of interest in the sample (Grzesiek et al., 1993; Kay and Gardner, 1997). The results of this improvement in sensitivity are even more pronounced when deuteration is combined with the use of NMR experiments based on the TROSY (transverse relaxation optimised spectroscopy) effect, which selects for long-lived signal components (Tugarinov et al., 2006).

Deuteration of non-labile sites in a protein sample can be achieved by expression of a protein by *E. coli* in partially deuterated M9 minimal media that includes deuterium oxide (D_2O) instead of H_2O. In minimal media, deuteration can be easily combined with ^{15}N and ^{13}C labelling to produce triple-labelled samples for optimal acquisition of data from triple-resonance backbone experiments. Subsequent incubation of the protein in H_2O (e.g. during purification using H_2O-based buffers) exchanges deuterons with protons at labile sites, the positions (including backbone amides) where protons or deuterons can be easily exchanged with those from the solvent, ensuring that signals from these nuclei in these chemical groups can be detected during NMR experiments.

Labelling of individual amino acid types with ^{15}N and/or ^{13}C nuclei can also be carried out reliably for many residue types in *E. coli in vivo*. Acquisition of HSQC spectra from samples containing different selectively labelled ^{15}N residue types can reveal the residue types of individual crosspeaks when overlaid with a uniformly labelled HSQC or TROSY spectrum of the uniformly labelled protein (see Section 3.3). However, combinatorial amino acid type-selective labelling, in which samples are prepared containing

both ^{15}N and ^{13}C nuclei in selected pairs of residue types, can provide information on sequential connections as well as information about residue type alone. These strategies typically rely on sample preparation via cell free protein production, which permits straightforward incorporation of isotopically labelled amino acids into the protein of interest due to the absence of enzymes that can promote metabolic scrambling (Hoffman et al., 2018). However, as demonstrated by Butterwick and MacKinnon (2010), assignments for residues that could not be completed by analysing conventional triple-resonance spectra were obtained from only three samples produced *in vivo*, containing a combination of six labelled amino acid types containing ^{15}N, ^{13}C or both.

Signals from specific groups of residue types can also be selected from uniformly ^{15}N labelled samples using carefully selected triple resonance experiments. These experiments initially select XH_2 or XH_3 side chain groups (where X can be either ^{15}N or ^{13}C) via a step based on the MUSIC (multiplicity selective in phase coherence transfer) coherence transfer scheme, and can provide further specificity in combination with selective pulses and tuned delays, enabling visualisation of two dimensional spectra containing only ^{1}H-^{15}N correlations from groups of residue types such as valine, isoleucine and alanine, leucine and alanine, aspartic acid, asparagine and glycine and even serine residues alone (Schubert et al., 2001).

Standard triple resonance backbone experiments such as HNCA provide both intra-residue and sequential correlation peaks. In order to reduce complexity and spectral overlap intra-residue experiments in which only H, N and Cα nuclei of the same residue are correlated can be collected instead, with the aim of reducing the number of peaks in the spectrum, enhancing resolution of the spectra and reducing ambiguity by avoiding overlap with peaks from other experiments such as HN(CO)CA. An HN(CO)CA experiment can be performed separately in order to obtain sequential correlation peaks (Brutscher, 2002). In contrast, a strategy to obtain sequential assignments from a single HNCA spectrum has recently been developed (Robson et al., 2018). In this approach, a sample is prepared in *E. coli* with a combination of 2-^{13}C and 3-^{13}C pyruvate as the sole carbon source, with the aim of producing protein samples in which Cα, but not Cβ, nuclei are ^{13}C labelled, suppressing the coupling between these nuclei and permitting significantly enhanced resolution for i and i-1 Cα resonances.

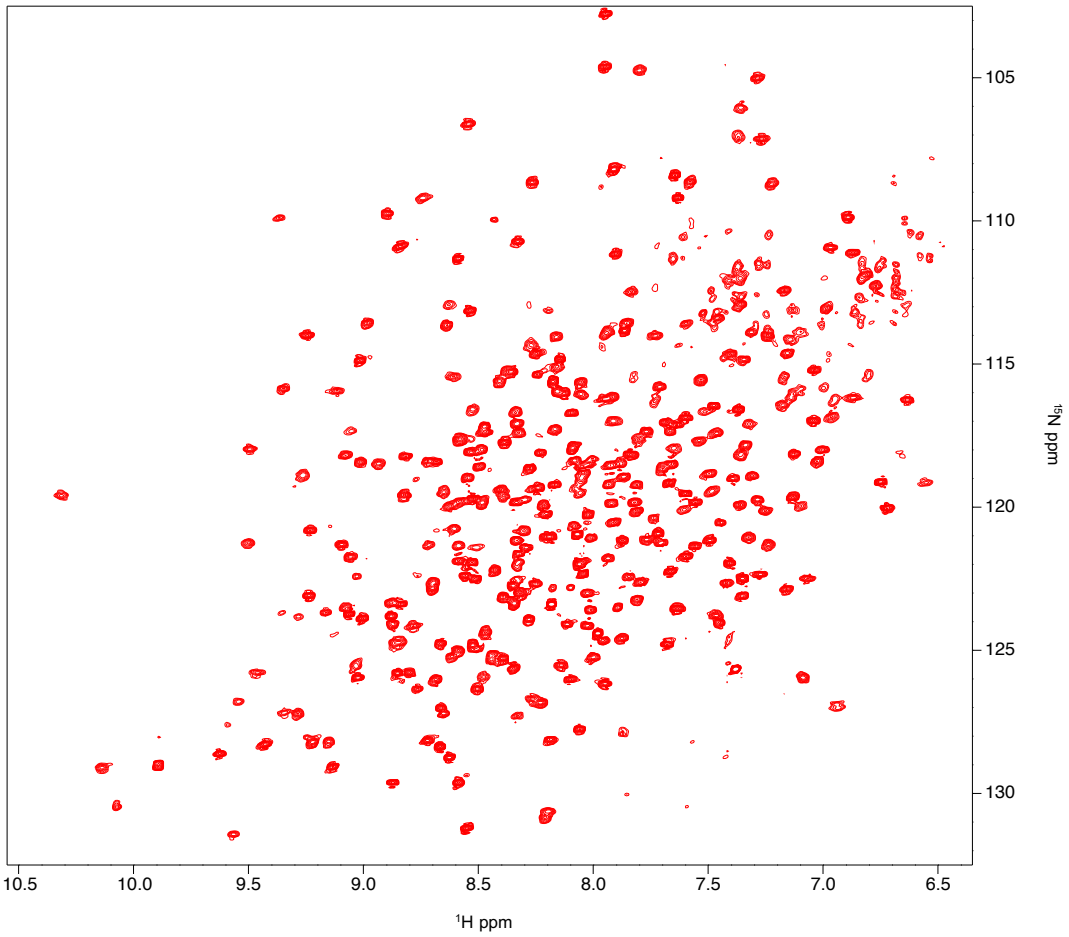


Figure 3.1. NMR spectrum of perdeuterated ^{13}C ^{15}N labelled wild-type catalytic domain. [^{15}N-^{1}H] TROSY spectrum of PARP-1 catalytic domain (656-1014). 200 µM in H_2O, 50 mM Tris d-11 pH 7.0, 50 mM NaCl, 2 mM DTT d-10, 5% D_2O. The spectrum was acquired at 30°C using a Bruker Avance III HD 800 spectrometer equipped with a triple resonance ($^{15}N/^{13}C/^{1}H$) cryoprobe. Assignments were completed and copied to the spectrum of ^{15}N labelled (non-deuterated) catalytic domain (see Figure 3.4), which was used as a reference for all subsequent downstream experiments.

3.2 Deuteration of PARP-1 wild-type catalytic domain

At 40 kDa in size and 360 residues in length, we considered that deuteration in conjunction with use of these experiments would be required for comprehensive sequential assignment of the PARP-1 catalytic domain using triple-resonance backbone experiments. Early in the project, expression trials and NMR experiments established that PARP-1 catalytic domain could be expressed and purified from *E. coli strain* BL21 (DE3) grown in rich media (LB medium) and that [^{15}N-^{1}H] HSQC and TROSY spectra could be acquired at 30°C from pure highly concentrated ^{15}N-labelled catalytic domain expressed in *E. coli* grown in H_2O-based M9 minimal medium (data not shown). Expression of recombinant protein from *E. coli* in deuterated media often results in lower yields than from H_2O-based media and therefore requires careful optimisation of expression parameters. Following these experiments, a small-scale trial of catalytic domain expression at different temperatures (18°C, 22°C, 25°C, 30°C and 37°C) and over different time periods (harvest of cells either

4 hours post-induction or overnight) was carried out in BL21 (DE3) in D_2O-based M9 minimal media made according to the recipe described in Chapter 2, and containing $^{14}NH_4Cl$ and $^{12}C_6$-glucose. Based on the approach described in Marley et al (2001), cells were initially grown in rich media, pelleted and re-suspended in a four-fold smaller volume of deuterated minimal media as described in Chapter 2. Due to the resulting density of the cultures, expression was induced in all cultures with 2 mM final concentration of IPTG. Yield of recombinant deuterated catalytic domain was negligible at temperatures below 30°C and highest overnight at 37°C (data not shown). On that basis, PARP-1 catalytic domain was then expressed in BL21 (DE3) overnight at 37°C from 250 mL D_2O-based M9 minimal media enriched with $^{15}NH_4Cl$ and $^{13}C_6$-glucose. A [^{15}N-^{1}H] TROSY spectrum (see Figure 3.1) was then recorded from the resulting sample (concentration ~200 µM). As with TROSY spectra recorded from non-deuterated catalytic domain, the partially deuterated catalytic domain was highly stable at 30°C and contained a number of well dispersed peaks approximately corresponding to the number of non-proline residues in the protein (343). Perdeuterated protein samples can require denaturation and refolding in H_2O-based buffers in order to ensure that solvent exchange occurs for all backbone amides, including those buried inside the protein fold with limited solvent exposure. However, this step was not required in the case of the catalytic domain. Work then proceeded with the acquisition of

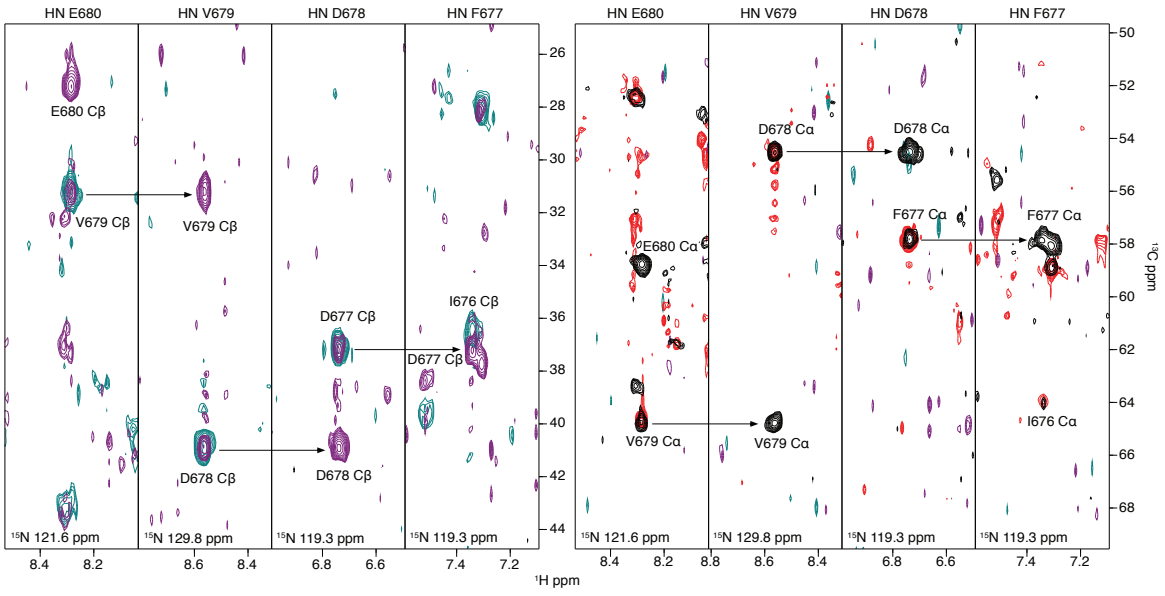


Figure 3.2 Backbone assignment of wild-type PARP-1 catalytic domain. TROSY versions of the triple resonance backbone experiments HNCA, HN(CO)CA, HNCACB and HN(CO)CACB were acquired at 30°C from a perdeuterated ^{13}C ^{15}N labelled wild-type catalytic domain sample (200 µM, 50mM Tris-d11-HCl pH 7.0, 50mM NaCl, 2mM DTT d-10, 5% D_2O) using Bruker Avance III HD 800 or Avance III 600 spectrometers equipped with triple resonance ($^{15}N/^{13}C/^{1}H$) cryoprobes. The selected strips show the assignment for the short section E680-F677. Superimposed HNCACB (blue) and HNCACOCB (teal) spectra, and superimposed HNCA (black) and HNCOCA spectra (red) reveal the connections between adjacent residues in the sequence. The assigned reference [^{15}N- ^{1}H] TROSY spectrum is shown in Figure 3.4.

suites of 3D triple-resonance spectra (TROSY-HNCA, TROSY-HN(CO)CA, TROSY-HNCACB and HN(CO)CACB) at 30°C as described in Section 2.3.1 in order to carry out resonance assignment of backbone amides. Using the parameters described in that section, non-uniform sampling (NUS) was employed to improve the quality of the spectra. Using these spectra, assignment of the backbone amide signals from the TROSY spectrum was performed in the programme CCPN Analysis 2.4 (Vranken et al, 2005) using the sequential approach described above in Figure 3.2.

3.3 Selective ^{15}N labelling of backbone amides for individual residue types

Although the typical Cα and Cβ chemical shifts recorded from triple-resonance backbone experiments are well characterised for each residue type, the possible chemical shifts of many residue types overlap with only some residue types exhibiting unique Cα and/or Cβ chemical shifts. This ambiguity can be partially resolved by selectively labelling the backbone amide moieties of individual residue types, revealing which of the amide signals from the uniformly ^{15}N-labelled domain are from that residue. For sequential assignment of larger proteins, 2D spectra from selectively labelled backbone amides can aid significantly in resolving ambiguities encountered during the assignment process by identifying the residue type of otherwise ambiguous or overlapping signals, as well as providing a number of starting points for chains of sequential assignments. Therefore, we attempted to selectively ^{15}N label backbone amides of a number of individual residue types during the assignment process.

3.3.1 Selective ^{15}N labelling of lysine, arginine, leucine and isoleucine residues

The feasibility of successfully selectively labelling backbone amide moieties varies by residue type, as selective ^{15}N labelling is achieved by supplementation of otherwise unlabelled M9 minimal media with the desired ^{15}N-labelled L-amino acid and atoms from some amino acid molecules may be metabolised into others by *E. coli*. For example, glutamate is a precursor for production of arginine, glutamine and proline, and glutamate and glutamine are inter-convertible by transaminases (Takeuchi et al., 2007) so supplementation of media with isotopically-labelled glutamate is likely to result in scrambling to these residue types, though this can be partially mitigated by the use of measures such as the use of auxotrophic strains or the addition of enzyme inhibitors. Additionally, selective labelling of backbone amides of residue types such as alanine, which

can be unambiguously distinguished from all or most others from triple-resonance spectra alone, is of limited utility for the purpose of resonance assignment. Therefore, the first amino acid types to be selectively labelled in the wild-type catalytic domain were lysine and arginine, which were labelled by supplementing unlabelled M9 minimal media with L-Lysine-α-^{15}N dihydrochloride, (Cambridge Isotope Laboratories), or L-Arginine-$^{15}N_4$ hydrochloride (Sigma-Aldrich) respectively as described in Chapter 2. Lysine was chosen due to the high proportion of lysine residues in the sequence (33 out of 360, representing 9.2% of all residues) and because its average Cα and Cβ shifts of 56.98 ppm and 32.76 ppm respectively (all values taken from Biological Magnetic Resonance Data Bank, Ulrich et al., 2008) both overlap considerably with Cα and Cβ shifts of other residues such as methionine (56.15 ppm and 32.93 ppm), phenylalanine (58.13 ppm and 39.92 ppm) and tyrosine (58.18 ppm and 39.26 ppm). Additionally, according to Chae (2000), backbone amides of lysine residues could be selectively labelled by addition of ^{15}N L-lysine alone and no additional steps, with only minor metabolic scrambling of isotope-labelled material. The 11 arginine residues make up a smaller proportion of the catalytic domain (3.1%), but its average Cα and Cβ shifts (56.81 ppm and 30.64 ppm) both overlap with residues including histidine (56.51 ppm and 30.22 ppm), glutamic acid (57.35 ppm and 29.95 ppm) and glutamine (56.61 ppm and 29.15 ppm). As with lysine, arginine is an end product in *E. coli*'s amino acid biosynthetic pathway and we hypothesised that mis-incorporation of ^{15}N-labelled arginine would therefore also be minimal.

In both cases, high quality [^{15}N-^{1}H] HSQC spectra were acquired from wild-type catalytic domain in which backbone amides of either lysine or arginine were selectively ^{15}N labelled (Figure 3.3). The number of peaks in each case corresponded to the expected number of residues of the labelled amino acid type from the wild-type catalytic domain sequence and overlays of each spectrum with spectra acquired from uniformly ^{15}N labelled catalytic domain revealed that chemical shifts of all peaks corresponded to a subset of peaks from the uniformly labelled spectra. Minimal isotopic scrambling occurred, as inferred from the very low number of smaller peaks observed and the fact that the numbers of strong peaks corresponded to the exact number of lysine and arginine residues expected from the sequence. In the case of arginine, minimal media was supplemented with uniformly ^{15}N labelled L-arginine which contained ^{15}N at side chain NH moieties in addition to the backbone NH. Strong peaks corresponding to side chain NH epsilons were visible in the selectively labelled ^{15}N arginine HSQC spectrum due to peak folding; this occurs when

peaks whose true frequencies fall outside the selected boundaries of the spectrum in the indirect dimension appear folded, or reflected, about one of the edges of the spectrum so that they appear at frequencies (and subsequently display chemical shifts) inside the spectral width after Fourier transformation. Folded peaks can be distinguished from unfolded (backbone) amide signals, as for the latter chemical shifts do not change when the spectral width is changed; in the case of spectra of uniformly ^{15}N labelled catalytic domain reported here, the greater spectral width of this experiment causes side chain NH epsilon signals to fold at chemical shifts outside the boundaries of the spectrum displayed in Figure 3.3.

Following the success of ^{15}N selective labelling for lysine and arginine, we sought to label more residue types in a similar way. PARP-1 catalytic domain contains 39 leucine residues (10.2% of all catalytic domain residues), the most of any residue type. When chemical shifts of both Cα (55.69 ppm) and Cβ (42.25 ppm) are considered, signals from leucine also overlap closely with average chemical shifts of aspartic acid (54.69 ppm and 40.86 ppm), which itself appears 25 times in the CAT domain sequence, and phenylalanine (58.13 ppm and 39.92 ppm), which appears 10 times. In order to selectively ^{15}N label leucine residues, a protocol was adapted from the approach described in Tong et al., (2008), which is described in detail in Chapter 2. This strategy involves supplementation of unlabelled M9 minimal media with both L-Leucine-^{15}N and unlabelled L-amino acids of all residue types except leucine, with the aim of saturating all biosynthetic pathways of non-leucine amino acid synthesis and disfavouring incorporation of ^{15}N atoms from labelled leucine into other amino acid types. This approach led to successful ^{15}N selective labelling of all 39 leucine residues with minimal isotopic scrambling (Figure 3.3). As this approach is simple, not specific to any single residue type, and requires no auxotrophic strains or specific inhibitors of *E. coli* enzymes, we later attempted to selectively ^{15}N label isoleucine residues in this manner. Although chemical shifts of isoleucine residues are typically considered to be distinctive, with Cα shifts commonly overlapping between isoleucine, valine and proline, but Cβ shifts of isoleucine easy to distinguish from the other two residue types, 25 isoleucine residues (6.9% of the catalytic domain sequence) were present in the protein sequence. Once again, all 25 residues of isoleucine were selectively labelled with this approach, with minimal isotopic scrambling observed. At this point, backbone amide assignments were completed for all lysine, arginine, leucine and isoleucine residues with the help of selective ^{15}N labelling, representing a full 30% of all 360 residues in the sequence.

3.3.2 Selective ^{15}N labelling of glycine and serine residues

According to results from Chae (2000), serine residues can be inexpensively labelled by supplementing otherwise unlabelled M9 minimal media with ^{15}N-Glycine, which leads to the expression of jointly ^{15}N glycine and serine labelled protein. ^{15}N labelled glycine is considerably cheaper then ^{15}N labelled serine, and Chae also found that in the strain BL21 (DE3), isotopically labelled serine scrambled to NH groups of glutamine or asparagine side chains. We attempted to selectively label serine residues (average Cα chemical shift of 58.74 ppm, Cβ 63.78 ppm) in order to distinguish them from threonine (average Cα chemical shift of 62.26 ppm, Cβ 69.69 ppm) in severely overlapped regions of the HNCA spectra, which also contained some particularly broad signals. Glycine residues can be easily distinguished from all other amino acid types, including serine, because glycine lacks a Cβ atom, and its Cα has a highly distinctive average chemical shift of 45.36. Two samples were prepared using slightly different methods, both using the adapted approach from Tong et al (2008) as described in Chapter 2. Both samples were supplemented with ^{15}N-Glycine, one with non-isotopically labelled serine and one without, and with both samples containing all 18 other amino acids in unlabelled form. We expected that the first sample would be glycine labelled only and that the second sample would contain labelled serine and glycine residues. However, HSQC spectra revealed that the first sample exhibited significant isotopic scrambling (data not shown), whereas in the second sample (Figure 3.3) only strong peaks corresponding to signals from glycine and serine residues were observed. Signals were observed for all 30 serine residues in the sequence. In the case of glycine, excluding the N-terminal residue which has no backbone amide moiety, backbone amide signals were observed for all glycines excluding G699, which remained unassigned across all spectra acquired from the catalytic domain, or G863, which was assigned in uniformly-labelled spectra using data from triple-resonance backbone experiments but exhibited low signal strength in a region highly overlapped with signals from other glycine and serine residues.

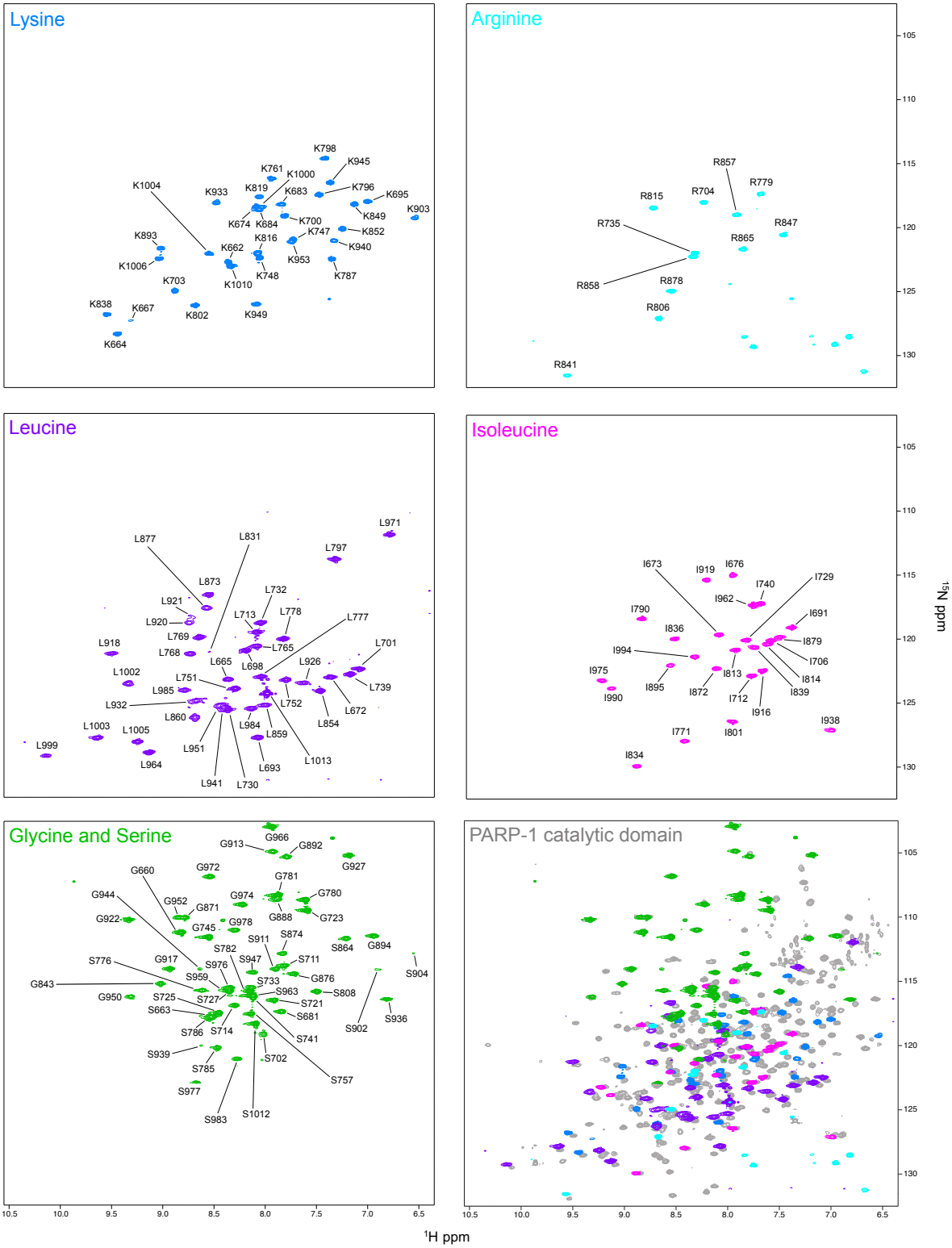


Figure 3.3 Selective ^{15}N labelling of backbone amides in individual residue types. Assigned [^{15}N-^{1}H] HSQC spectra of selectively ^{15}N labelled lysine (150 µM), arginine (250 µM), leucine (400 µM), isoleucine (400 µM), glycine and serine (400 µM) residues in wild-type PARP-1 catalytic domain, including superposition of [^{15}N-^{1}H] HSQC spectra of all selectively ^{15}N labelled residue types with [^{15}N-^{1}H] TROSY of PARP-1 catalytic domain (all experiments performed in NMR Buffer as described in Figure 3.1). Selectively labelled spectra of lysine, arginine, leucine and isoleucine residues were acquired on a Bruker Avance I 600 spectrometer at 30°C. HSQC spectra of ^{15}N labelled glycine and serine residues and [^{15}N-^{1}H] TROSY of uniformly ^{15}N labelled PARP-1 catalytic domain were acquired at 25°C on a Bruker Avance III HD 800 spectrometer. All spectrometers were equipped with triple-resonance (^{15}N/^{13}C/^{1}H) cryoprobes.

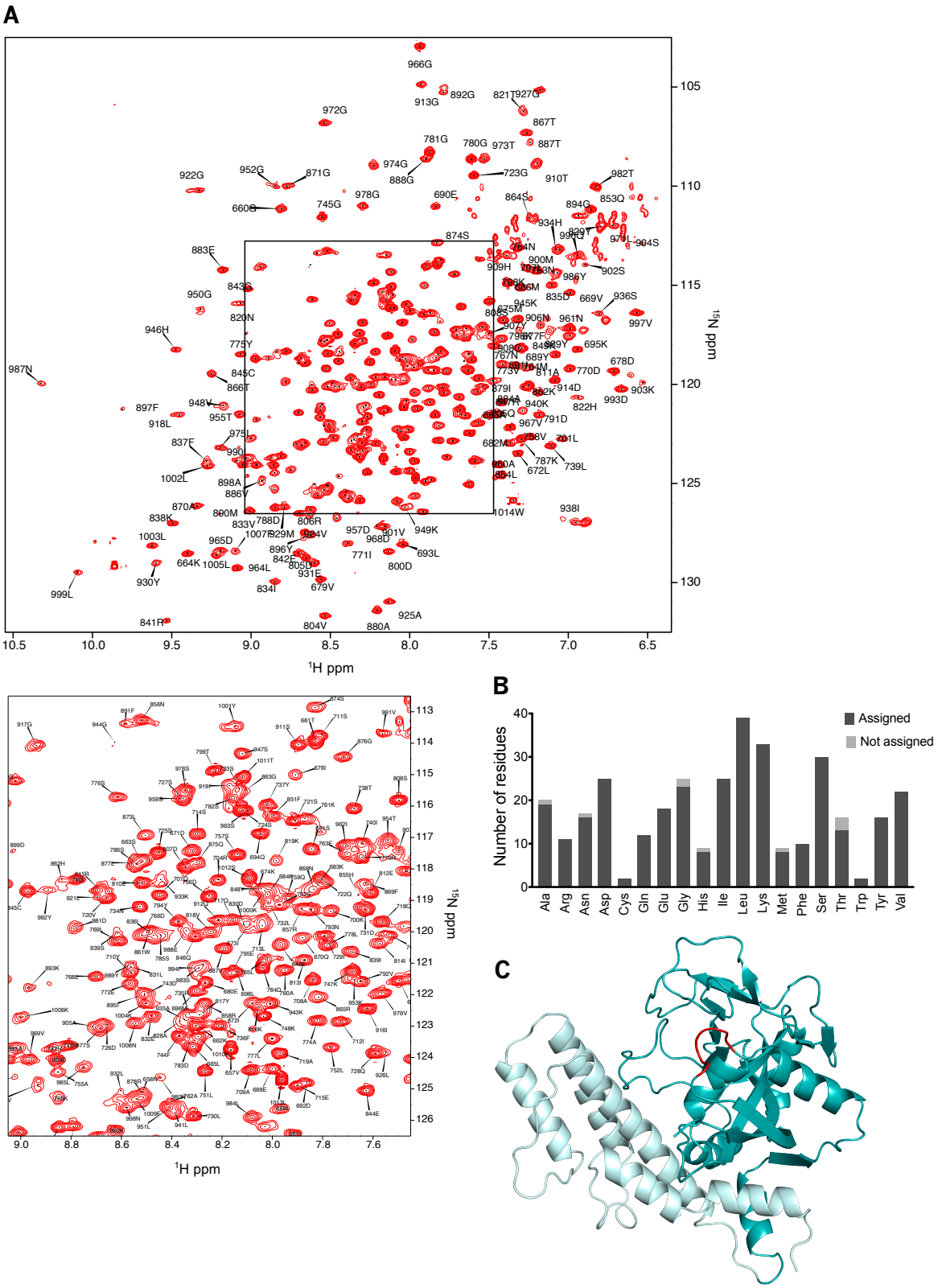


Figure 3.4 NMR resonance assignment of PARP-1 wild-type catalytic domain. (A) Assigned [^{15}N-^{1}H] TROSY spectrum of ^{15}N labelled (non-deuterated) PARP-1 (656-1014), 400 µM in NMR buffer at 25°C. Sequential assignments were performed based on triple-resonance backbone data acquired from ^{13}C ^{15}N ^{2}H labelled PARP-1 catalytic domain at 30°C and transferred to the TROSY spectrum of ^{15}N labelled PARP catalytic domain acquired at 25°C. This latter spectrum, shown here, was used as a reference for all subsequent experiments with catalytic domain mutants and PARP inhibitors. (B) Completion of backbone amide resonance assignments by residue type in the reference TROSY spectrum of wild-type PARP-1 catalytic domain. (C) Location of unassigned residues 823-827 (red) in the ART subdomain (teal) of PARP-1 catalytic domain.

3.4 Completed assignments of wild-type PARP-1 catalytic domain

Using a sequential assignment approach combined with acquisition of spectra from samples with selective ^{15}N labelling of individual residue types, resonances were assigned for over 96% of backbone amides from the 343 non-proline residues in the wild-type PARP-1 catalytic domain. Assignments from experiments run on the triple-labelled wild-type catalytic domain at 30°C were copied to the TROSY spectrum of ^{15}N labelled PARP-1 catalytic domain acquired at 25°C, which was used as a control and reference for subsequent experiments. Of the 11 non-proline residues that could not be assigned, 2 were residues from the N-terminus of the sequence (a non-canonical glycine residue and T656) and 5 were the residues 823-827 (sequence: ATTHN), a consecutive sequence of residues found in a single loop in the ART subdomain. Additionally, weak signals from NH moieties of residues K667 and L920 were assigned in spectra from the triple-labelled and selectively labelled samples but were not visible in the reference TROSY spectrum. The assigned reference TROSY spectrum, completion of assignment by residue type and the location of residues 823-827 are all shown in Figure 3.4.

3.5 Assignment of PARP-1 catalytic domain mutants and wild-type catalytic domain in complex with PARP inhibitors

During this project, we aimed to investigate both the structural basis of activation in the PARP-1 catalytic domain using constitutively partially active mutants of the PARP-1 catalytic domain fragment (Chapter 4) and the structural and mechanistic basis of PARP-1 inhibition by 4 different PARP inhibitors (Chapter 5). Therefore, following the resonance assignment of PARP-1 wild-type catalytic domain, it was also necessary to carry out assignment of as many backbone amides as possible in the chosen constitutively partially active mutants of the catalytic domain, as well the wild-type catalytic domain in complex with our chosen PARP inhibitors. Chemical shifts are exquisitely sensitive to the environment of individual nuclei, and are affected not only by changes in covalent bonding in a protein structure, but also by interactions with binding partners. Therefore, changes such as the introduction of a mutation in the HD subdomain or binding of a PARP inhibitor may introduce changes in the environment of many backbone amides, which means that the transfer of assignments to these spectra from those of the wild-type catalytic domain cannot be considered a trivial task, especially in light of the catalytic domain's number of residues and high molecular weight, obstacles which applied equally to the 3 mutant domain and 4 inhibitor complexes investigated.

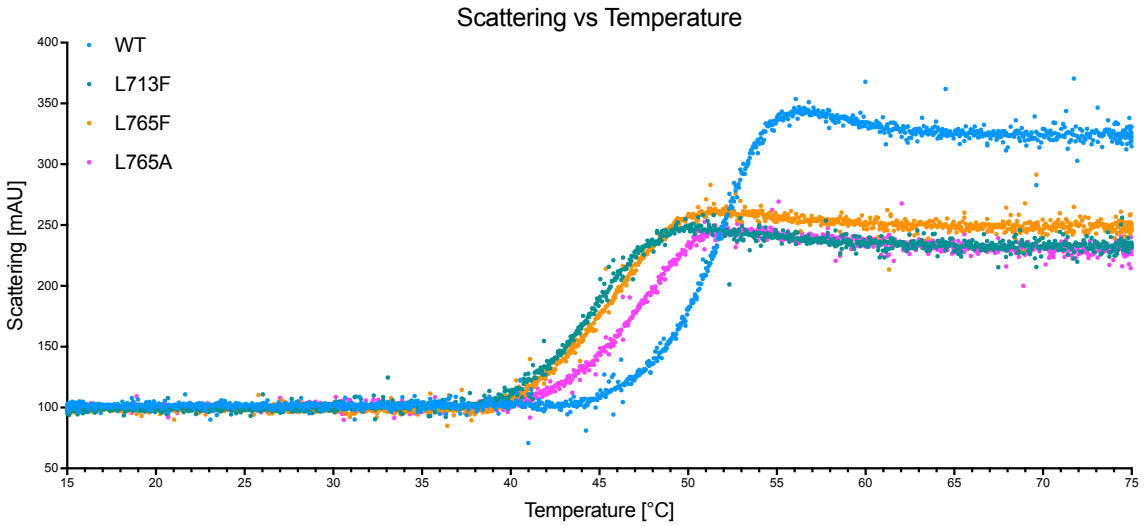


Figure 3.5 Thermal unfolding of wild type catalytic domain and catalytic domain mutants. Scattering from wild-type catalytic domain and catalytic domain mutants (all 500 µM in 50 mM Tris pH 7.0, 50 mM NaCl, 2 mM DTT) plotted against temperature. Samples were loaded into capillaries and heated by the Prometheus NT.48 (30% excitation, 15-75°C, increasing at a rate of 2°C per minute).

In the case of the constitutively partially active HD subdomain mutants L713F, L765F and L765A, we initially aimed to complete the resonance assignment by performing and analysing the same TROSY-based triple-resonance backbone experiments as carried out on the wild-type catalytic domain, in conjunction with perdeuteration, ^{13}C and ^{15}N labelling. We anticipated that for the majority of residues, there would be no large differences in Cα or Cβ chemical shifts, so that for these, as well as for backbone amide signals that only had small CSPs, assignments could be made largely by comparison against wild-type signals, with the possibility of making sequential assignments using explicit connectivity tracing through the triple resonance data in the most difficult cases. However, this strategy was hampered by the lower thermal stability of the mutant catalytic domains in comparison to the wild-type domain. Initial attempts to acquire TROSY spectra from ^{15}N-labelled samples of L713F at 30°C failed when the sample rapidly precipitated. In order to determine the highest temperature that spectra could be safely recorded without destroying the samples, thermal unfolding profiles of wild-type PARP-1 catalytic domain and all three mutants were generated via nanoDSF (differential scanning fluorimetry) using a Prometheus NT.48 machine from Nanotemper (Figure 3.5). This revealed that for the catalytic domain mutants, onset of melting (T_{onset}) occurred at lower temperature than for wild-type catalytic domain. In wild-type catalytic domain, T_{onset} occurred at 42.5°C, whereas L713F, L765F and L765A mutants, T_{onset} occurred at 34.8°C, 35.5°C, and 36.7 °C respectively. Although the T_{onset} temperatures obtained from the unfolding profiles were greater than 30°C, which we had already tested during acquisition of mutant domain spectra, it was clear that the thermal

stability of the catalytic domain mutants was lower than for the wild-type catalytic domain. Therefore, we attempted to acquire spectra from fresh ^{15}N labelled catalytic domain mutant samples at 25°C and found that high quality [^{15}N-^{1}H] TROSY spectra could be acquired from all mutants at this temperature (see Section 4.3), and that labelled mutant samples remained stable when stored at room temperature or 4°C. However, the quality of spectra from perdeuterated ^{13}C ^{15}N labelled L713F mutant at 25°C was not sufficient to make unambiguous assignments for all of the L713F mutant based on similarity of the signals to the wild-type domain. Although data recorded from HNCA and HNCOCA experiments was of high quality and signals from NH groups of some residues could be assigned based on proximity to wild-type signals in TROSY spectra and similarity of Cα chemical shift perturbations alone, the quality of HNCACB and HNCOCACB spectra was low and without high quality data from these experiments it was not possible to make unambiguous assignments on a sequential basis for difficult cases, such as for signals from residues in close proximity to the mutation site, which were highly perturbed by ring currents from the substitution of a leucine residue by a phenylalanine.

In the case of wild-type catalytic domain in complex with PARP inhibitors, the assignment task was complicated by the high binding affinities and slow dissociation rates of the inhibitors (summarise in Section 5.1), which bind tightly to the catalytic domain and form complexes that are in slow exchange on the NMR timescale. For complexes in fast exchange, with weak binding where the off rate of the binding partner is greater than the chemical shift difference between free and bound states, a single peak averaging the chemical shift of both states will be observed in the spectrum, and multi-point titration of binding partners leads to progressively larger chemical shift perturbations which can in principle be followed and assigned with no follow-up experiments. However, if the complexes exchange between states over a slow exchange timescale, two separate signals will be observed in the spectra with different chemical shifts, with their intensities corresponding to the size of each population. Assignment of complexes in slow exchange is consequently far more challenging, because even with multiple titration points, chemical shift perturbations cannot be progressively traced to the end point of the titration by eye, so the bound-state signals must be carefully assigned using further experiments. Assignment of four PARP inhibitor complexes using a sequential backbone assignment approach would have necessitated production of four further perdeuterated ^{15}N ^{13}C labelled samples, with

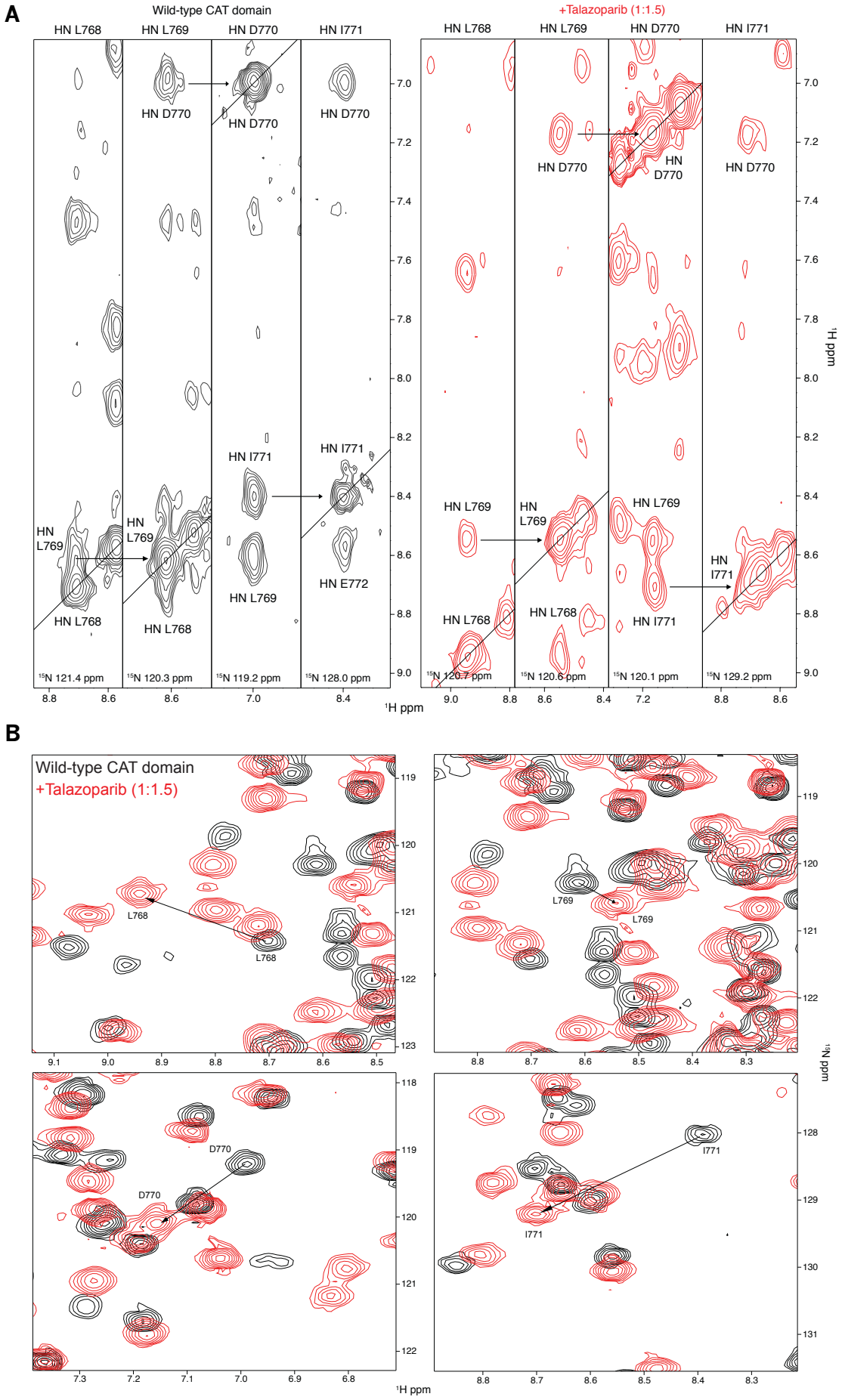

A
Wild-type CAT domain
+Talazoparib (1:1.5)
HN L768
HN L769
HN D770
HN I771
HN E772
^{15}N 121.4 ppm
^{15}N 120.3 ppm
^{15}N 119.2 ppm
^{15}N 128.0 ppm
^{15}N 120.7 ppm
^{15}N 120.6 ppm
^{15}N 120.1 ppm
^{15}N 129.2 ppm
^{1}H ppm
B
Wild-type CAT domain
+Talazoparib (1:1.5)
L768
L769
D770
I771
^{1}H ppm
^{15}N ppm

Figure 3.6 (previous page) Backbone assignment of catalytic domain mutants and wild-type catalytic domain in complex with PARP inhibitors using ^{15}N NOESY-HSQC. (A) Selected strips from ^{15}N NOESY-HSQC spectra of wild-type catalytic domain (black) and wild-type catalytic domain titrated with a 1:1.5 ratio of talazoparib (red) for the short section of L768-I771. Comparison of strips from free wild-type catalytic domain reveals NOE signals from backbone amides of neighbouring residues, enabling identification of the corresponding connections and residues in ^{15}N NOESY-HSQC spectra from PARP inhibitor complexes and catalytic domain mutants, even where chemical shift perturbations have occurred (see B). (B) Assigned backbone amides of the short sequence L768-I771 in [^{15}N ^{1}H] TROSY spectra of wild-type catalytic domain (black) and new assignments of backbone amides from the same sequence in in complex with talazoparib (red) made using ^{15}N NOESY-HSQC spectra. ^{15}N NOESY-HSQC and [^{15}N ^{1}H] TROSY spectra were acquired at 25°C from ^{15}N labelled catalytic domain samples (500 µL volume, 400 µM concentration) using a Bruker Avance III HD 800 spectrometer. 6 µL of a 50 mM talazoparib stock solution in 100% DMSO-d_6 was added to make a final concentration of 600 µM at a ratio of 1:1.5 to ensure that the catalytic domain was titrated to saturation with talazoparib.

accompanying spectrometer time required to carry out separate sets of HNCA, HNCOCA, HNCB and HNCOCB triple-resonance experiments on each complex.

Consequently, we adopted a different strategy in order to complete the assignment of the catalytic domain mutants, which was also later applied to assignment of TROSY spectra from complexes of the catalytic domain with PARP inhibitors. Instead of triple-resonance backbone data, we recorded a three-dimensional ^{15}N NOESY-HSQC experiments at 25°C from ^{15}N labelled wild-type catalytic domain, as well as from all of the catalytic domain mutants and wild-type catalytic domain samples titrated to saturation with the PARP inhibitors veliparib, olaparib, talazoparib and EB-47, as described in Section 2.3.2. The ^{15}N NOESY-HSQC experiment exchanges magnetisation between hydrogens using the nuclear Overhauser effect (NOE) before transferring the magnetisation to neighbouring ^{15}N and back to nearby ^{1}H nuclei for detection. This experiment exploits the NOE to reveal which nuclei are in close spatial proximity to one another and is typically used to generate restraints for structure calculations using NMR. However, it can also be used to make assignments, as when the experiment is visualised for a given amide signal, NOEs from neighbouring residues, including for the backbone amide, can be observed (see Figure 3.6). As with data from triple-resonance backbone experiments, assignments can be made by analogy to data from wild-type protein, and neighbouring residues can be identified, in this case via NOEs. This approach requires less time to collect high quality data (4 days for the PARP-1 catalytic domain as opposed to approximately 2 weeks to collect all triple-resonance backbone experiments) and sample preparation was significantly cheaper as no ^{13}C labelling or deuteration was required. Using this approach, assignments were largely completed for all mutants and inhibitor complexes (see Table 3.1). The backbone amide assignment of all mutants and complexes was completed to a similar level as for the wild-type catalytic domain, with the exception of the catalytic domain in complex with EB-47, for which only

88% of non-proline residues were assigned. The number of assignments was lower in comparison with complexes containing the other inhibitors because no signals were observed for some residues that were part of or in close proximity to the loop in the ART subdomain (residues 880-900) where the catalytic domain's active site is located.

Sample	Assigned	Not Assigned	Percentage
Wild-type (WT)	331	12	96.5%
L713F	323	20	94.2 %
L765F	326	17	95.0 %
L765A	328	15	95.6 %
WT + Veliparib (1:1.5)	328	15	95.6%
WT + Olaparib (1:1.5)	327	16	95.3%
WT + Talazoparib (1:1.5)	324	19	94.4%
WT + EB-47 (1:1.5)	302	41	88.0%

Table 3.1 Assignments completed for wild-type PARP-1 catalytic domain, catalytic domain mutants and complexes of wild-type catalytic domain with 4 PARP inhibitors. All catalytic domains contained 343 non-proline residues (out of 360 residues in total). All percentages are calculated as percentages of non-proline residues.

3.6 Spectral changes occurring following sample preparation

During initial acquisition of 2D TROSY spectra from both deuterated and non-deuterated ^{15}N labelled wild-type and mutant PARP-1 catalytic domains, it was observed that consistent and progressive small changes in chemical shifts occurred for a subset of peaks over the first 1-2 days after the sample had been prepared. After this time period, no further changes occurred and chemical shifts of all spectra acquired subsequently were identical. As the end point of the changes was consistently reproducible, all assignments and subsequent experiments were carried out once these changes had been allowed to progress effectively to completion. Therefore, as wild-type and mutant PARP-1 domains are stable at room temperature (22°C), samples were incubated at room temperature for 48 hours following purification before NMR experiments were carried out, or before addition of PARP inhibitors during PARP inhibitor titrations, in order to ensure that no changes would occur during data collection.

Because the peaks changed their chemical shifts smoothly over time and only a small proportion were affected, mostly in vicinity of the NAD^+ binding site, we concluded that the most plausible original of the effect was weak binding of some small molecule binding

partner that slowly changed its bound concentration over time. Based also on discussion with John Pascal, who had observed density corresponding to an unidentified molecule in the active site of catalytic domain crystals (unpublished personal communication), we hypothesised that this binding partner was located in the NAD^+ binding pocket.

However, we were not able to identify this compound using either ^1H NMR spectra of the wild-type catalytic domain, or via time-of-flight mass spectrometry of an inactive catalytic domain mutant, E998A, in both cases looking unsuccessfully for changes over time (data not shown), nor could we establish what process caused its binding equilibrium to change over time.

Chapter Four

Studies of PARP-1 activation in catalytic domain mutants

The mechanism of DNA damage recognition by the N-terminal zinc finger domains of PARP-1 and the subsequent formation of interdomain contacts leading to activation in the catalytic domain have been well characterised (Eustermann et al., 2015). Data from HXMS experiments in full-length PARP-1 has also provided insight into the activation process in the catalytic domain itself, revealing that the HD (helical domain) subdomain of the catalytic domain is autoinhibitory and blocks productive NAD^+ binding in the absence of DNA damage, and that upon DNA binding at the N-terminus of the protein some regions of the catalytic domain's HD subdomain undergo local partial unfolding, promoting substrate access to the active site (Dawicki-McKenna et al., 2015). However, the resolution of HXMS is limited to the level of fragment peptides and these events could not be studied at the level of individual residues or atomic resolution. Using three constitutively partially active mutants of the isolated catalytic domain fragment that were hypothesised to mimic the effects of DNA binding, the role of different helical regions in the HD was investigated using NMR spectroscopy.

4.1 Constitutively partially active catalytic domain mutants

Multiple catalytic domain mutants with elevated DNA-independent activity have been identified, all with mutations in the hydrophobic core of the HD subdomain. These mutants have a constitutive level of activity that is greater than wild-type PARP-1's basal level of activity and is independent of DNA-binding and the resulting allosteric activation mechanism involving other domains of PARP-1. The HD mutant L713F was the earliest mutant of this type to be identified (Miranda et al., 1995) via a random mutagenesis approach. L713F is located in the αD helix of the of the HD subdomain (Figure 4.1) and initial kinetic studies upon its discovery estimated that the L713F mutation increased catalytic domain activity by at least 9-fold. Later biochemical and kinetic analysis by the Pascal Group (Langelier et al., 2012) showed that L713F has DNA-independent catalytic activity 20-fold greater than the basal level, with no change in affinity for NAD^+. They also tested a novel mutant L765A, with a mutation site located in the αF helix, which exhibited near-identical properties to L713F. However, the catalytic activity of these mutants was not

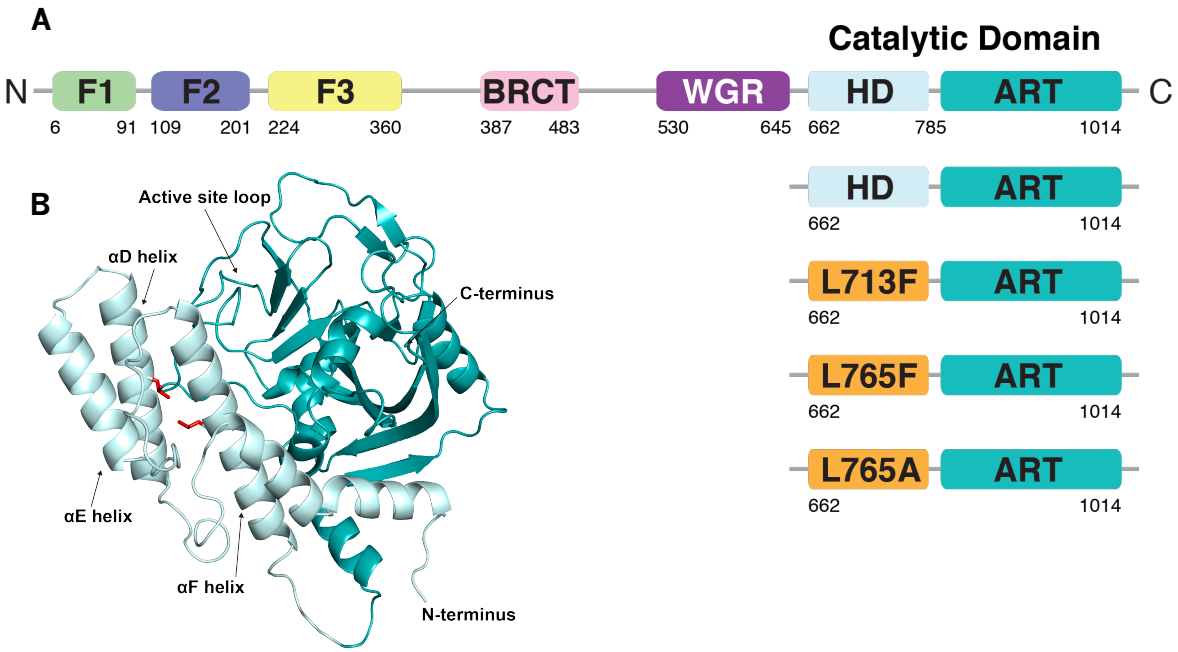


Figure 4.1 Constitutively partially active mutants of PARP-1 catalytic domain. (A) PARP-1 domain structure and colour scheme used throughout this thesis (top) and constructs of the catalytic domain used in this chapter (bottom). (B) Structure of PARP-1 catalytic domain (pdb 4DQY). Side chains of L713 (situated in the αD helix) and L765 (situated in the αF helix) are shown in red.

as great as the estimated 50-fold increase in the catalytic activity of full-length PARP-1 following binding to sites of DNA damage (Langelier et al., 2014). Nonetheless, subsequent study of the effect of the L713F mutation on HX behaviour by the Pascal group showed that at the time points 10^4 and 10^5 seconds, increases in HX occurred at the same locations of the HD subdomain as for full length wild-type PARP-1 following DNA damage binding, independent of changes in HX near the mutation site, which occur far earlier (Dawicki-McKenna et al., 2015). Therefore, the behaviour of this mutant, whilst only displaying a partial increase in activation with respect to fully activated wild-type PARP-1 or the fully activated DNA-independent mutant PARP-1 ΔHD, appeared to mimic the effect of the full-length wild-type PARP-1 following binding to DNA damage. Importantly, given the independence of the mutants' catalytic activity from the activity of the N-terminal domains of PARP-1, there is no difference in the effect on catalytic activity between the mutations in the context of full-length PARP-1 and the isolated catalytic domain fragment. Consequently, as we considered that the size (116 kDa) of full-length PARP-1 would be prohibitive for the initial study of catalytic domain activation using NMR spectroscopy, we hypothesised that introducing these mutations into the isolated catalytic domain fragment instead would act as a model system enabling us to study the mechanistic basis of activation in the PARP-1 catalytic domain fragment *in vitro* using NMR spectroscopy. Therefore, we sought to clone,

express and acquire spectra from the catalytic domain mutants L713F, L765A and L765F, the last of these being another mutation tested by the Pascal group with similar kinetic properties to L713F and L765A (unpublished personal communication with John Pascal). These particular mutants were selected because during expression trials carried out by the Pascal group, their effect on catalytic domain activity was greater than other single amino-acid substitutions in the HD subdomain mutations that were tested and expression of catalytic domain containing double mutations was significantly worse in *E. coli* (John Pascal, unpublished personal communication).

4.2 Optimisation of catalytic domain mutant expression in *E. coli*

In order to study the catalytic domain mutants, it was necessary to express and purify isotope-labelled proteins from *E. coli*. As wild-type PARP-1 catalytic domain had been successfully cloned and expressed from the modified pETt28a vector described in Section 2.1, primers were designed to clone the desired mutants from this same vector using a site-directed mutagenesis approach, also described in Section 2.1. However, in initial expression trials carried out in *E. coli* under the expression conditions previously used for wild-type catalytic domain, it became clear that the expression of all three mutants in *E. coli* was poor at all temperatures, IPTG concentrations and harvesting time points tested. Furthermore, it was not possible to purify the expressed mutant proteins from *E. coli* because all three mutant catalytic domains were persistently found in the insoluble fraction during the centrifugation step following cell lysis. As the transformed cells grew well prior to induction of protein expression with IPTG, we hypothesised that the greater catalytic activity of the constitutively partially active mutants was depleting levels of NAD^+ inside the *E. coli* cells following expression, causing the cells either to reject the plasmid or sequester the over-expressed recombinant protein in an unfolded state inside inclusion bodies. As a result, we adopted a published protocol from the group of John Pascal, which optimises the expression and purification of full-length wild-type and mutant constructs of PARP-1, PARP-2 and PARP-3 from *E. coli* via the supplementation of expression media with the low affinity PARP inhibitor benzamide (Langelier et al., 2017). Benzamide mimics NAD^+ (see Figure 4.2A) and binds to the active site of PARP family catalytic domains, inhibiting NAD^+ binding and PARP catalytic activity and consequently automodification of PARP proteins. With an IC_{50} (the concentration required for 50% inhibition *in vitro*) of 22 μM, benzamide is potent enough to inhibit 92% of PARP catalytic activity at a concentration of 1 mM (Banasik et al., 1992), but its binding affinity is also modest enough to allow it to be removed from PARP-

1 catalytic domains during purification of PARP proteins from *E. coli* (Langelier et al., 2017).

In order to test the effectiveness of benzamide, TY agar plates and 2x TY liquid media containing 10 mM benzamide were prepared as described in Chapter 2. For expression of non-codon optimised full-length PARP-1, PARP-2 and PARP-3 in *E coli*, the Pascal group carried out expression overnight at 16°C in Rosetta 2 (DE3) cells in LB media containing 10 mM benzamide, inducing expression with addition of IPTG to a final concentration of 0.2 mM. All catalytic domain constructs in our study were codon optimised and we had previously expressed wild-type PARP-1 catalytic domain (with N-terminal His_6-lipoyl tag) at both 37°C and 30°C overnight with a final IPTG concentration of 2 mM. Therefore, in order to compare the effect of supplementation with benzamide for wild-type and mutant catalytic domains, wild-type catalytic domain and catalytic domain mutants L713F, L765F and L765A were expressed overnight in the *E. coli* strain BL21 (DE3) at either 16°C or 37°C, in small-scale 10 mL culture volumes of LB media containing kanamycin, with either no benzamide or 10 mM benzamide. Expression in all cultures was induced with a final concentration of 0.2 mM IPTG and cells were harvested as described in Chapter 2, 16 hours after induction. Cells were then lysed using BugBuster (Novagen) and the soluble and insoluble fractions were separated via centrifugation according to the manufacturer's instructions. Aliquots from the total lysate (soluble and insoluble fractions) and supernatant (soluble fraction only) were analysed using SDS-PAGE (Figure 4.2B). In this expression trial, wild-type catalytic domain was observed in the soluble fraction under all conditions (16°C and 37°C, with and without the addition of benzamide). However, catalytic domain mutants were observed in the soluble fraction only following expression at 16°C in media supplemented with 10 mM benzamide. No catalytic domain mutants were found in the soluble fraction following expression at 16°C without benzamide or at 37°C under any conditions, although the presence of large bands corresponding to 51 kDa observed in the total lysate fractions indicate that the mutants were expressed in *E. coli* cells at 37°C following induction. At 16°C with 10 mM benzamide, large 51 kDa bands were observed in both fractions of WT, L713F and L765A lysates, in contrast to wells from the same temperature without benzamide, in which only faint bands no larger than background proteins were observed. Due to the significant improvement in expression and solubility of catalytic domain mutants L713F and L765A at 16°C with addition of 10 mM benzamide, all

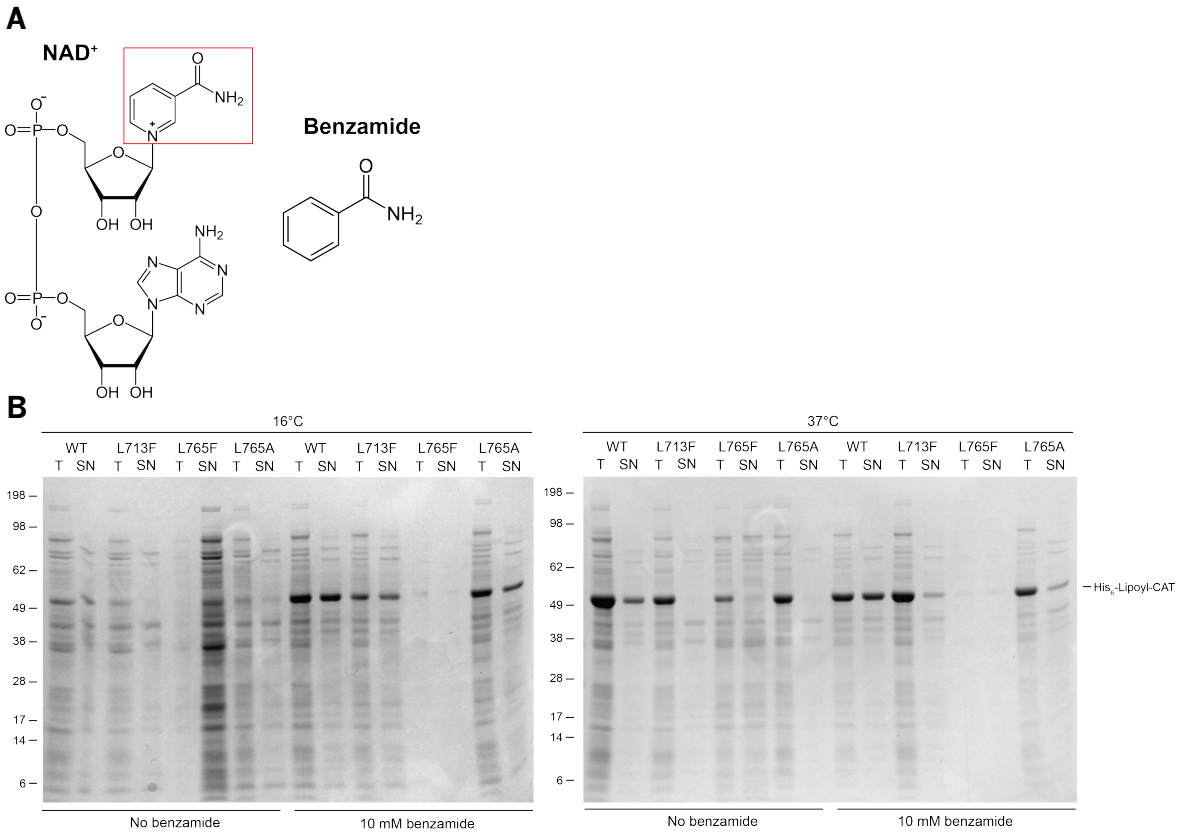


Figure 4.2 Addition of the low-affinity PARP inhibitor benzamide improves expression and solubility of constitutively partially active catalytic domain mutants in *E. coli* at low temperature. (A) Chemical structures of NAD^+ and benzamide. Benzamide closely resembles the nicotinamide moiety of NAD^+ (inside red outline). (B) Aliquots from total lysate (T) and supernatant from total lysate clarified by centrifugation (SN) analysed using SDS-PAGE. The presence of a 51 kDa band in T indicates that the domain was expressed in *E. coli*. The presence of a 51 kDa band in SN indicates that the domain was found in the soluble fraction.

catalytic domain mutants were subsequently expressed overnight in media containing 10 mM benzamide, including in M9 minimal media for later production of isotope-labelled samples. The fact that wild-type catalytic domain was observed in the soluble fraction at 37°C with and without benzamide, in contrast to the catalytic domain mutants (even with addition of benzamide) also suggested that the catalytic domain mutants exhibited lower thermal stability than the wild-type catalytic domain. This was subsequently confirmed by thermophoresis experiments and melting curves described in Chapter 3 (Section 3.5). However, after further optimisation of the expression conditions for the catalytic domain mutants (data not shown), it was established that in media including 10 mM benzamide, higher yields of soluble target proteins could be acquired with overnight expression at 20°C and induction of recombinant protein expression with a final IPTG concentration of 0.5 mM. Additionally, in Figure 4.2B, no protein bands are observed in any lysate fractions from benzamide cultures for the mutant L765F. However, follow-up tests of L765F expression in

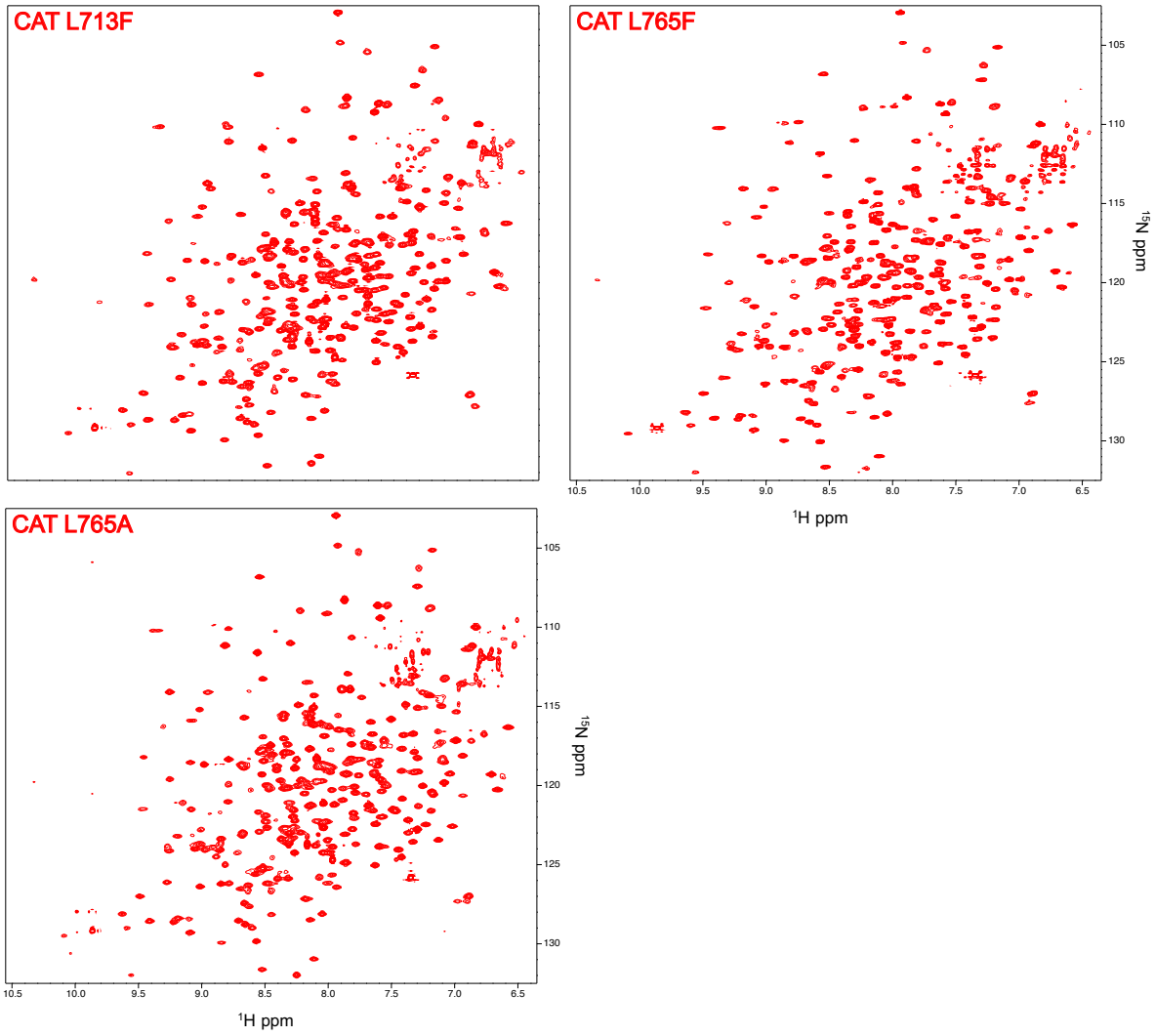


Figure 4.3 [^{15}N-^{1}H] TROSY spectra of PARP-1 catalytic domain mutants L713F, L765F and L765A. All spectra were acquired at 25°C from 400 µM samples of ^{15}N labelled catalytic domain mutants in NMR Buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C).

both rich and minimal media confirmed that L765F can be expressed and purified from *E. coli* under identical conditions to L713F and L765A (data not shown). The lack of bands in these wells, including those of endogenous proteins from *E. coli*, most likely indicates that cell lysis of these pellets by the BugBuster reagent was ineffective for these samples.

4.3 Mutation of residues in α-helices of the HD subdomain causes wide-ranging chemical shift perturbations in the HD and ART subdomains

Following optimisation of catalytic domain mutant expression and purification using benzamide, work proceeded with acquisition of high quality ^{15}N-^{1}H TROSY spectra from PARP-1 catalytic domain containing the L713F, L765F or L765A mutations. Optimisation

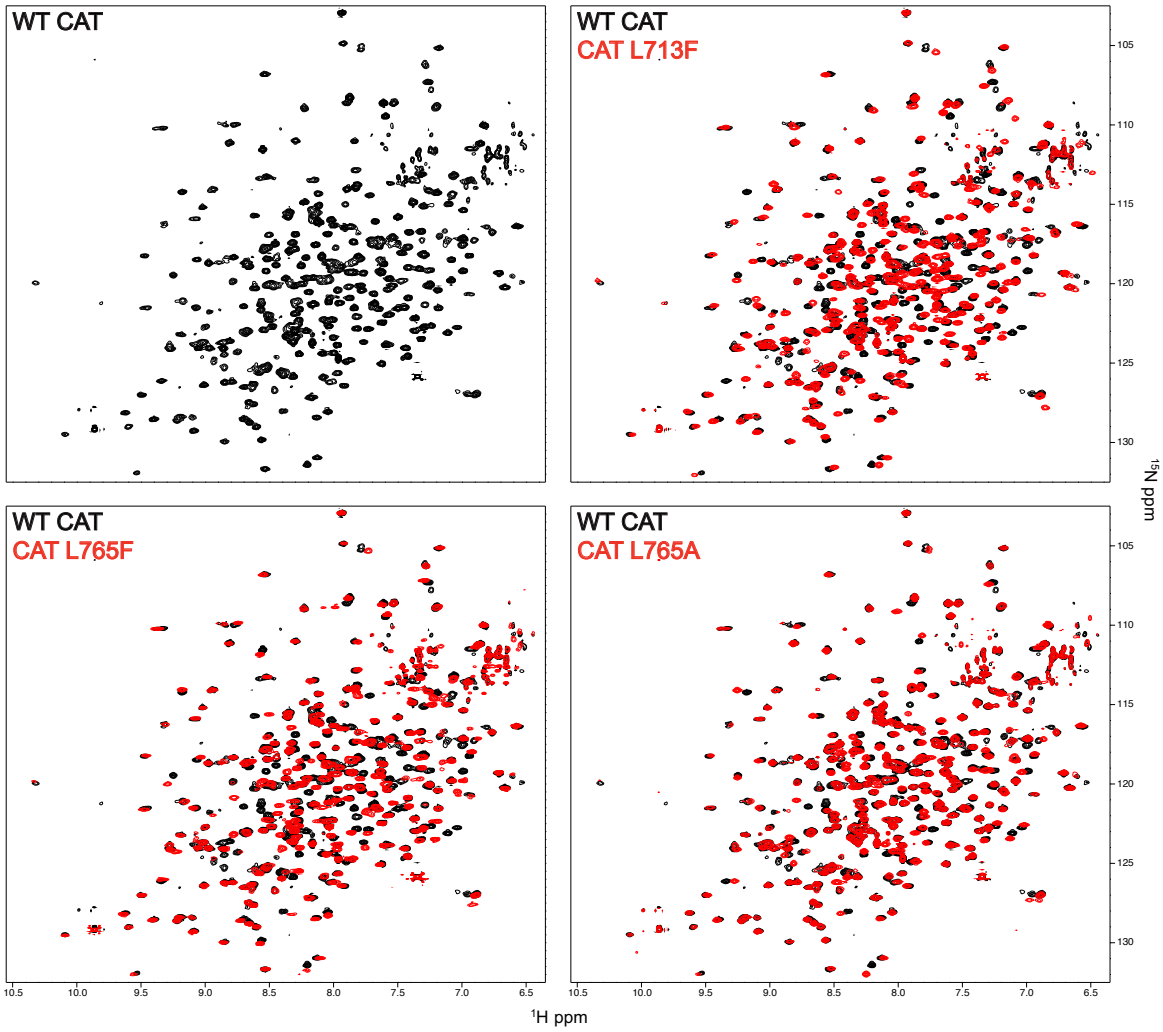


Figure 4.4 Comparison of [^{15}N-^{1}H] TROSY spectra of wild type PARP-1 catalytic domain and catalytic domain mutants L713F, L765F and L765A. All spectra were acquired at 25°C from 400 µM samples of ^{15}N labelled wild-type catalytic domain or catalytic domain mutants in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C).

of conditions for spectral acquisition is described in Section 3.5, and ^{15}N-^{1}H TROSY spectra were recorded for wild-type catalytic domain and all three HD mutants at 25°C in NMR buffer (Figure 4.3). As with the equivalent spectra of wild-type catalytic domain, the spectra of all three mutants were of high quality and well-dispersed with no random-coil shifts, indicating that the proteins were well-folded and homogenous. This suggested that no partial melting or significant, large-scale change in the secondary structure of any region of the HD subdomain had occurred. However, from comparison of spectra from the mutants with spectra from wild-type (Figure 4.4) it was also clear that the chemical shifts of many amide cross peaks had changed. Chemical shifts are extremely sensitive to changes in the nuclear environment and in the ^{15}N-^{1}H TROSY experiment, chemical shifts of all non-proline backbone amides can be used as a probe, allowing data to be interpreted at the level of

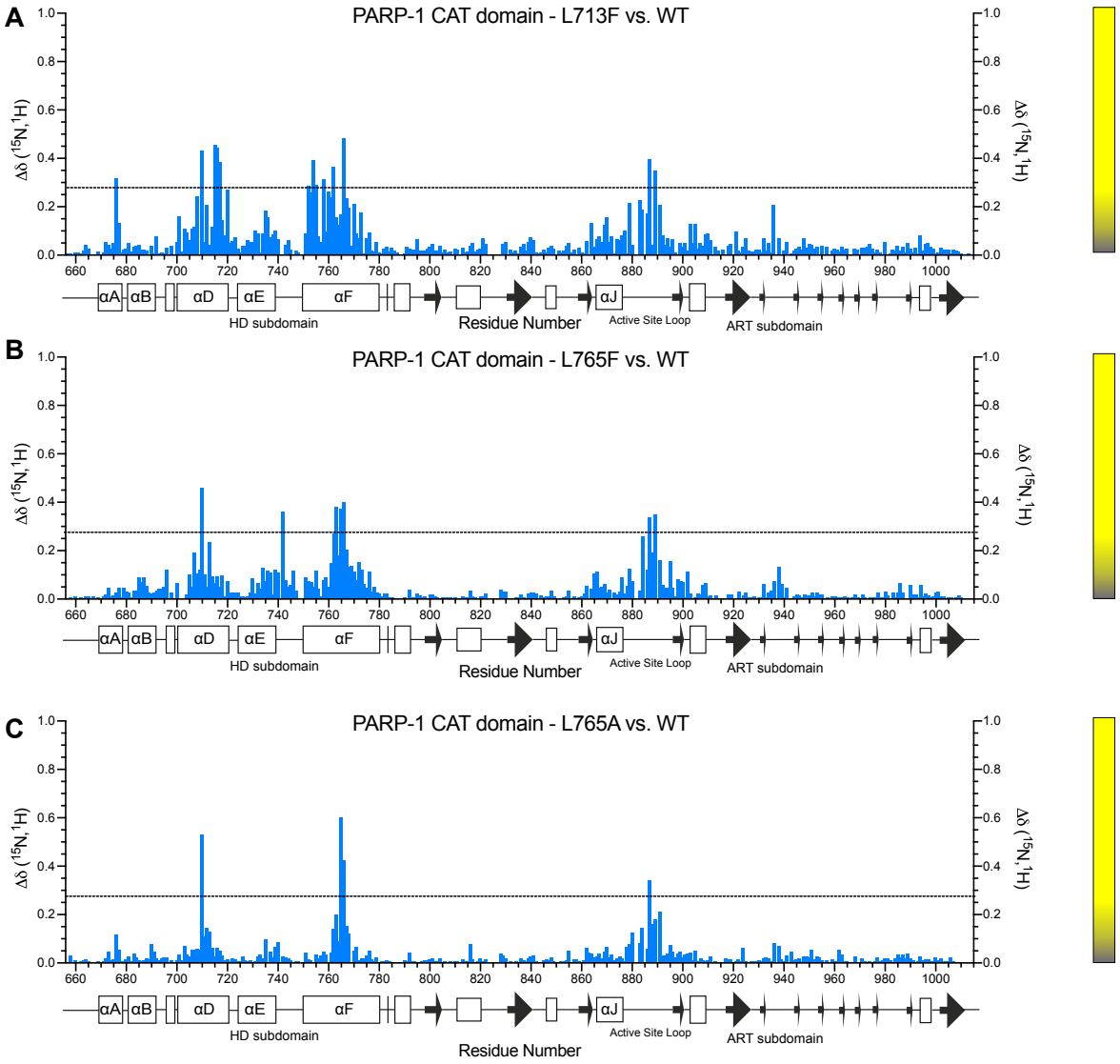


Figure 4.5 Chemical shift perturbations observed in PARP-1 catalytic domain as a result of mutations L713F, L765F and L765A. (A), (B) and (C). All chemical shift perturbations observed for mutations L713F, L765F and L765A respectively. Scaling of CSP mapping by colour to the catalytic domain structure in Figure 4.6 is shown next to each plot. Subdomain and secondary structure boundaries are shown below each plot. Chemical shift perturbations below 0.249, the upper limit of CSP scaling by colour, are enlarged in (D), (E) and (F).

individual residues. Although chemical shift perturbations near to the mutation site would be expected in the case of amino acid substitutions such as those of the catalytic domain mutants, given the expected properties of the constitutively partially active catalytic domain mutants, we hypothesised that any change in the regulatory properties of the HD subdomain resulting from the amino acid substitutions could have longer-range effects which might be detected using CSPs. The spectra acquired from the mutants were primarily assigned using data collected from ^{15}N NOESY-HSQC experiments (Section 3.5), with partial assignment of L713F also completed using data from HNCA and HNCOCA spectra. In order to assess the effect of each HD mutation on the catalytic domain, we then compared the chemical shift perturbations of ^{15}N-^{1}H TROSY spectra from each mutant with respect to spectra from the

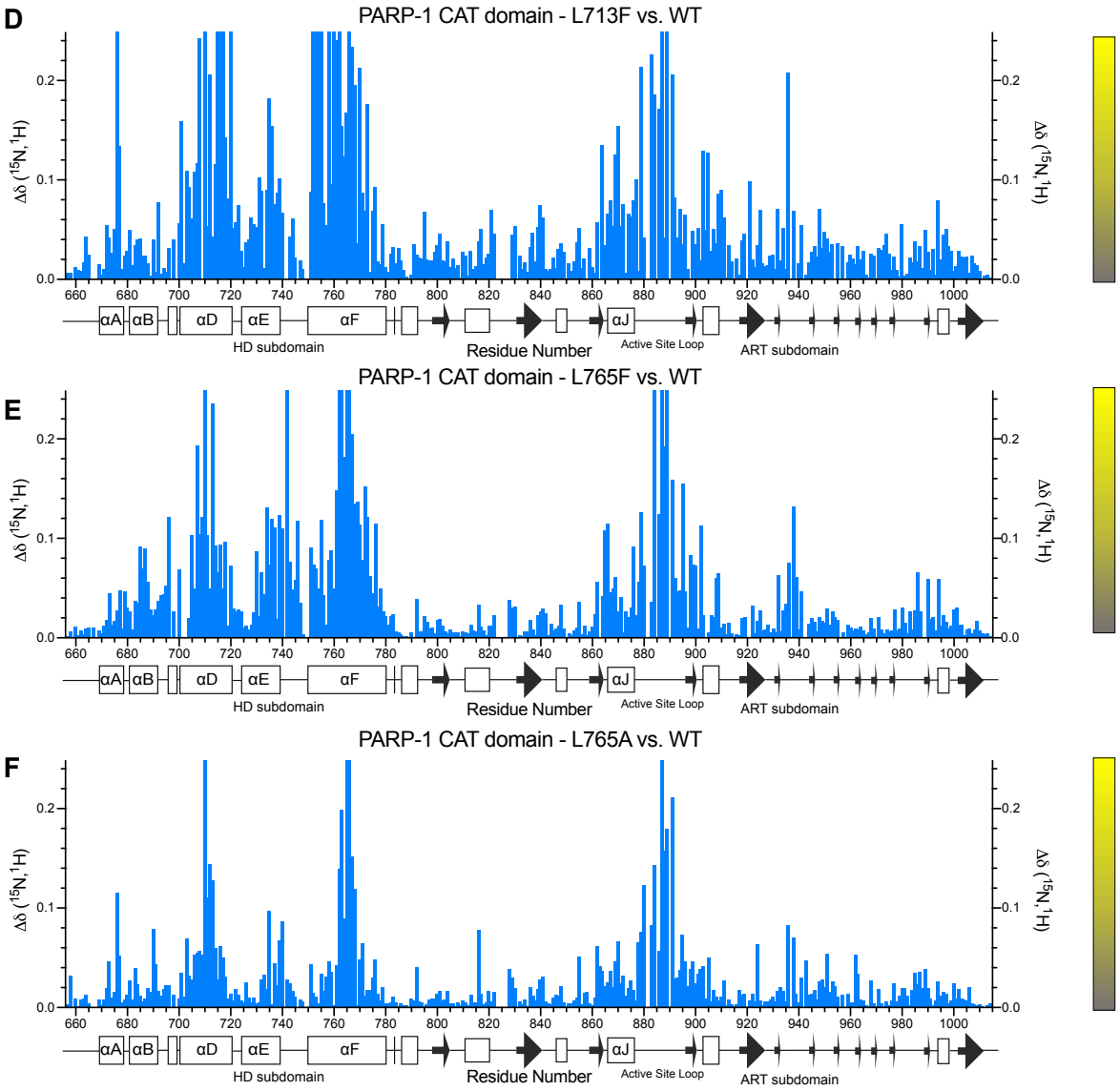


Figure 4.5 Chemical shift perturbations observed in PARP-1 catalytic domain as a result of mutations L713F, L765F and L765A. (D), (E) and (F). Enlarged plots of chemical shift perturbations below 0.249 from (A), (B) and (C) respectively. Subdomain and secondary structure boundaries are shown below each plot. Scaling of CSP mapping by colour to the catalytic domain structure in Figure 4.6 is shown next to each plot.

wild-type catalytic domain. Chemical shift perturbations of the backbone amides were plotted by residue (Figure 4.5) and subsequently mapped onto the previously solved structure of the wild-type catalytic domain as described in Chapter 2 number (Figure 4.6). Because such mapping can easily be dominated by the very largest CSP values, which can make the many informative but smaller variations that occur for other residues difficult to see, it is useful to define an upper threshold for the mapping process, whereby it reaches its maximum colour saturation at a value that is scaled down relative to the largest CSP value in the dataset. In this case, the CSPs from the mutant data sets were globally scaled against each other and CSP data sets from titration of the catalytic domain with PARP inhibitors (as described in Section 5.2), which contained larger CSPs than those found in the mutant data sets in this

chapter. In this case the threshold was set to 0.2, meaning that the colour ramp reaches its maximum of bright yellow at a CSP value that is 0.2 times the largest individual CSP value among all datasets. Among all CSP data collected for PARP-1 catalytic domain, the largest CSP value is 1.247 (residue G888 in the talazoparib data set from Section 5.2), meaning that a threshold of 0.2 scales all CSPs between 0.0 and 0.249 in size, and all CSPs greater in size than this are mapped with maximum colour saturation.

All three mutants displayed strong CSPs in the same regions of the HD and ART subdomains. In all cases, the largest chemical shift perturbations were observed in the αD, αE and αF helices in the HD subdomain (corresponding approximately to residues 700-720, 726-740 and 755-780 respectively) and region 880-900 of the ART subdomain (Figure 4.5), a loop region which is part of the PARP signature fold and the active site location of PARP-1. Although there was no significant difference in the size of the largest CSPs in these regions between each mutant, a greater number of CSPs were observed near the mutation sites for L713F and L765F compared to L765A. Although we knew that all three mutants possessed similar catalytic activity (John Pascal, personal communication) and had established that all three possessed similar thermal stability, in L765A one residue with an aliphatic side chain (leucine) is replaced with another (alanine), whereas in L713F and L765F leucine is replaced with phenylalanine, containing an aromatic benzyl ring. The introduction of an aromatic moiety in this manner may lead to large changes in chemical shifts of nearby chemical groups as a result of aromatic ring currents, as in the case of L713F and L765F. However, the fact that the locations of the largest CSPs observed were consistent between all three mutants, including L765A, despite the differences in location and nature of the mutations shows that the effects seen in the HD and ART subdomains are the result in a change in the behaviour of the HD subdomain. The observation of CSPs in the ART subdomain following the introduction of an activating mutation in the HD subdomain is consistent with previous studies showing that in wild-type inactive PARP-1, the HD is autoinhibitory in nature and blocks productive substrate access to the active site (Dawicki-McKenna et al., 2015). In this instance, disruption of the HD's autoregulatory function would increase the accessibility of the active site to the substrate. Even in the absence of the substrate, CSPs at this location could occur due to an increase in solvent exposure or due to disruption of contacts between the active site loop and the αF helix, the region of the HD subdomain in closest proximity to the active site loop. Mapping of the CSPs from each mutation onto the structure of the wild-type catalytic domain (Figure 4.6) reveals that introduction of a mutation into any one of the

α-helices αD, αE and αF in the HD subdomain causes wide-ranging CSPs in the other helices of this group, even if they are not immediate neighbours. For example, even in the L765F and L765A mutants (mutated in the mid-point of the αF helix), CSPs are observed in both the neighbouring αD helix, which immediately borders αF in the structure, but also the more distant αE helix, particularly in the case of L765F, where CSPs are observed among residues L735-745, the C-terminal section of the αE helix and beginning of the long αE-αF linker. Similar CSPs are also observed in the L713F mutant, despite the fact that the side chain of residue 713 faces towards the αF helix and not αE. However, the fact that all peaks in ^{15}N-^{1}H TROSY spectra of the mutants are well-dispersed and no random coil shifts are observed suggests that CSPs are not caused by partial melting of the helices. Instead, the largest CSPs observed in the HD subdomain suggest that mutations of residues in these α-helical regions disrupt hydrophobic interactions between side chains which face towards the neighbouring α-helices. For example, a CSP of 0.4, among the largest in the HD subdomain and catalytic domain as a whole, is observed for the residue Y710 in all three mutants (Figure 4.5 and Figure 4.7). The bulky hydrophobic, aromatic side chain of Y710 protrudes from the αE helix in the direction of the αF helix, in close proximity to the entire hydrophobic side chain of L765 which extends from the αF helix in the opposite direction. Substitution of L765 in the L765F and L765A mutants and consequent disruption of this contact presumably explains why Y710 exhibits large CSPs. In the case of L765F, the leucine side chain has been substituted for the bulky benzyl ring of the phenylalanine side chain, whereas in L765A the longer leucine side chain has been substituted for a single methyl group that comprises the side chain of alanine, removing two methyl groups from close proximity to the phenyl group in the Y710 side chain. The CSPs observed for Y710 of the same size are the same in L765A and L765F, despite the greater number of smaller chemical shift perturbations observed in the αD helix in L765F overall, showing that this effect is likely to be independent of the CSPs caused by ring currents from insertion of an aromatic side chain group. Y710 also lies in close proximity to L713 on the same helix and a CSP of the same magnitude is observed in the L713F mutant, in which a leucine side chain is also substituted for a phenylalanine. Overall, the presence of these CSPs in response to the introduction of mutations in regions of the HD subdomain proximal to the active site loop is consistent with the HD subdomain's regulatory function and its autoinhibitory nature in wild-type PARP-1 in the absence of DNA damage, and suggests that maintenance of this autoinhibitory function depends on the stability of interactions between key α-helices.

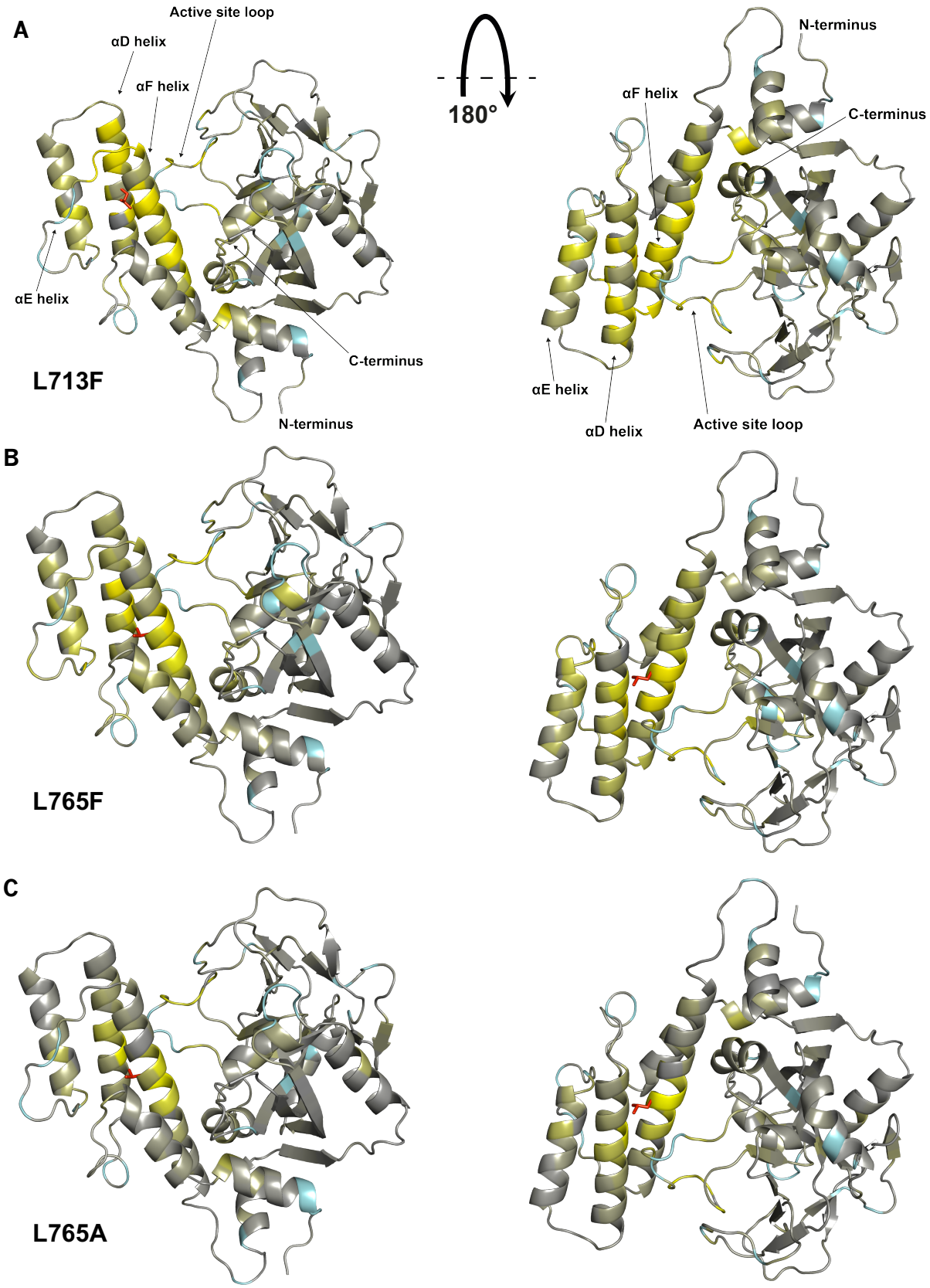


Figure 4.6 Mapping of chemical shift perturbations from PARP-1 catalytic domain mutants L713F, L765F and L765A. (A), (B) and (C). Chemical shift perturbations from Figure 4.5 globally scaled and mapped onto the structure of residues 662-1011 of wild-type PARP-1 catalytic domain (pdb 4DQY), as described in Chapter 2. CSPs are shown on a gradient from grey (no CSP) to yellow (all CSPs > 0.249). Residues that are incomparable due to missing assignments (including proline residues) are shown in cyan. The position of L713 or L765 side chains from wild-type catalytic domain are shown in red.

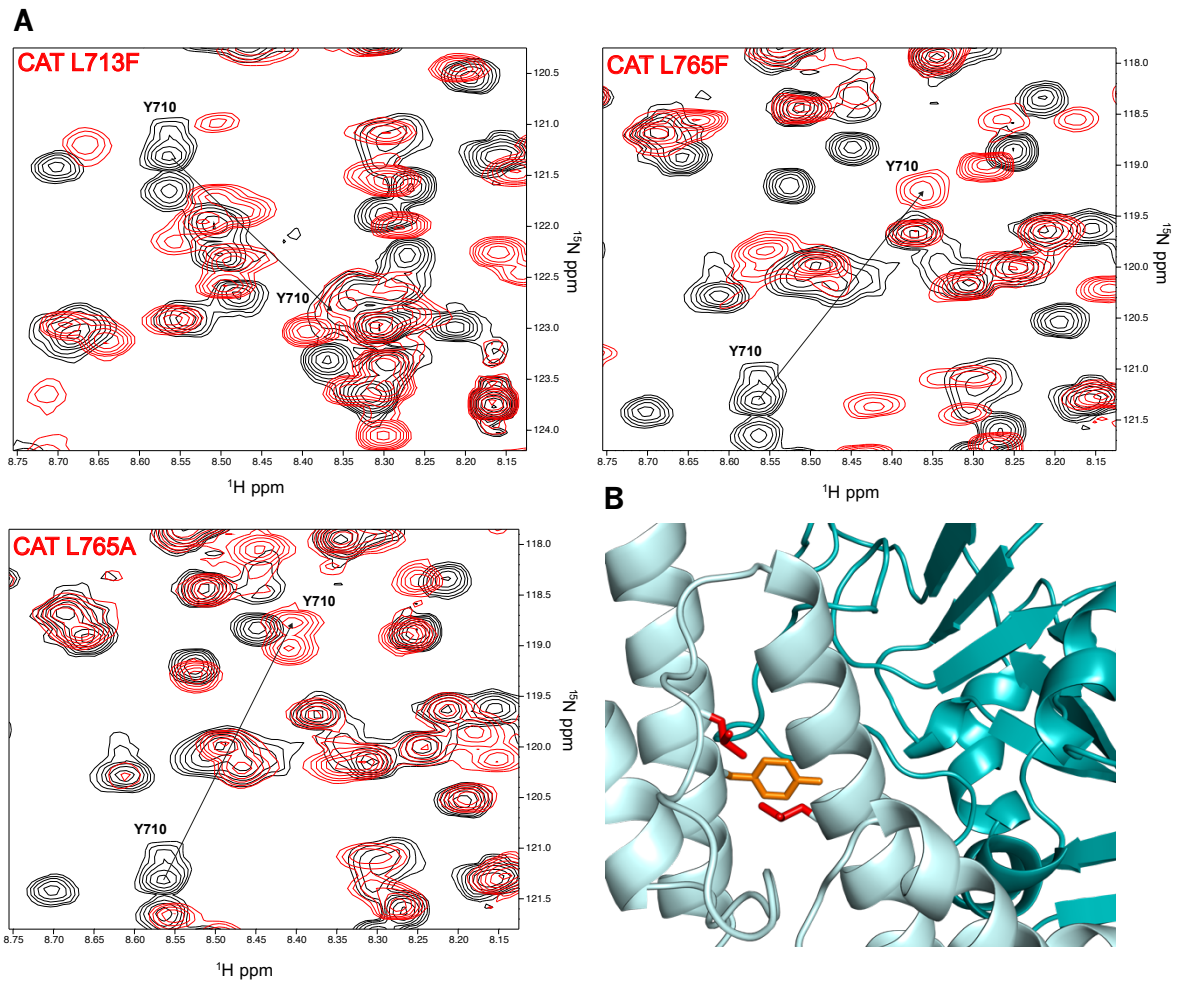


Figure 4.7 CSPs of Y710 in catalytic domain mutants. A) Chemical shifts of Y710 in catalytic domain mutants in comparison to wild-type. B) Location of Y710 side chain (orange) in wild-type PARP-1 catalytic domain. Side chains of residues L713 (αD helix, left) and L765 (αF, right) are shown in red. The HD subdomain is shown in cyan and the ART subdomain is shown in teal.

4.4 ^{15}N relaxation analysis of backbone amides in PARP-1 catalytic domain mutants

Analysis of chemical shift perturbations in the catalytic domain mutants showed that substitution of key leucine side chains in the α-helices of the HD subdomain disrupts interactions between the helices which are important for maintenance of autoinhibitory function. However, previous studies of activation following DNA binding in full-length PARP-1 reported that activation of the PARP-1 catalytic domain was mediated by changes in the secondary structure of HD helical regions and hypothesised that changes in the rate of solvent exchange in these regions were consistent with partial or even near-complete unfolding, leading to the loss of α-helical secondary structure (Dawicki-McKenna et al., 2015). From examination of the two-dimensional TROSY spectra recorded from the catalytic domain mutants, in which the distribution of the peaks was consistent with well-

folded protein samples, we concluded that complete unfolding of these regions did not occur in the mutants. Nonetheless, we sought to investigate whether any less substantial perturbation of α-helical secondary structure had occurred in these regions in the catalytic domain mutants by measuring the effect of the HD subdomain mutations on the dynamics of backbone amides in the catalytic domain using NMR spectroscopy. In NMR, the term relaxation describes how long it takes for a population of nuclei to return to equilibrium from an excited state. Relaxation times are generally longer for nuclei in disordered or linker regions in comparison to nuclei from regions with ordered secondary structure, in particular for α-helices, as the time taken for the NMR signal to decay in nuclei of interest is shortened by dipolar interactions and scalar coupling with neighbouring magnetically active nuclei, particularly protons, and this effect is greater for slowly moving regions in a molecule than for regions where internal flexibility causes motions to be faster. Chemical shift anisotropy (CSA), which arises when the electronic environment that surrounds a nucleus and shields it from the external magnetic field is non-spherical, also contributes to relaxation as it causes variation in the strength of local fields as the molecule tumbles. This effect is more pronounced at the high-strength magnetic fields used for biomolecular NMR as the rate of relaxation from this mechanism depends on the square of the externally applied field. Therefore, we hypothesised that if mutation of L713 or L765 disrupted the formation of α-helices at those sites, or any sites in close proximity, we would observe corresponding changes in the relaxation behaviour of the backbone amides of those nuclei.

For PARP-1 catalytic domain and the catalytic domain mutants, we measured backbone ^{15}N spin-locked relaxation times ($T_{1\rho}$, Figure 4.7), $^{15}N\{^{1}H\}$ steady state NOEs (Figure 4.8) and T_1 (longitudinal) relaxation times (data not shown). Each of these types of data have their own characteristic dependence on the rotational correlation time of the molecule (τ_c, defined as the average time that it takes for a particle to rotate by one radian), as well as on any faster internal motions or slow exchange processes that may be present for particular residues. $T_{1\rho}$ (which is quite similar in its behaviour to T_2, but measurement of which was preferred for technical reasons) includes a dependence on slow motions and exchange processes (millisecond to microsecond timescale) while the heteronuclear NOE and T_1 depend only on faster rates (nanosecond timescale) associated with longitudinal transitions. In the $T_{1\rho}$ experiment, the rate constant for decay of spin-locked ^{15}N magnetisation is measured, while the T_1 experiment measures the decay constant for longitudinal ^{15}N magnetisation. In the $^{15}N\{^{1}H\}$ heteronuclear experiment, two spectra are

recorded from the sample with and without ^{1}H saturation, and values for the $^{15}N\{^{1}H\}$ heteronuclear NOE are determined by the ratio of peak intensity in the two spectra for each amide group. Amide groups from relatively flexible or unstructured regions undergo faster motion and have lower $^{15}N\{^{1}H\}$ NOE values than amide groups of residues in more rigid or highly structured regions, where motion of amide groups is slower. By comparing the $T_{1\rho}$ and $^{15}N\{^{1}H\}$ NOE data, we aimed to assess the contributions of differing processes to the relaxation of backbone amides in the HD subdomain mutants. In addition, by also using the T_1 data, we could derive τ_c values for the samples, using the method described in Section 2.3.4. Values of τ_c are generally considered to be dependent on the molecular weight of the protein, assuming that it is approximately globular in shape. Although τ_c is theoretically defined as a single value for the molecule as a whole, one can calculate individual values for each residue of the protein and use these to estimate an average τ_c across ordered regions. Furthermore, if one estimates such an average across individual domains and then compares the results between domains, this can sometimes give an indication of whether the domains tumble as a single entity (if the domains have the same τ_c consistent with the overall size of the protein), or as flexibly linked semi-independent entities (if the domains have the appreciably different τ_c values roughly consistent with the sizes of the individual domains).

Both $T_{1\rho}$ (Figure 4.8) and heteronuclear NOE (Figure 4.9) data displayed clear variations in the dynamics of different residues across the HD subdomain in wild-type PARP-1 catalytic domain, which corresponded to regions of the catalytic domain structure with slower relaxation times and local motions in solution. For example, residues 780-785 could be clearly identified as the linker region between the HD and ART subdomains from the sharp increase in $T_{1\rho}$ and decrease in intensity ratios of the heteronuclear NOE. $T_{1\rho}$ values for the three residues G723, S724 and S725, which form part of a short linker region between the αD and αE helices, were 39 ms, 48 ms and 49 ms respectively, compared to an average $T_{1\rho}$ value of 25 ms for the majority of residues in the HD and ART subdomains. In data from the heteronuclear NOE experiment, the intensity ratios of these three residues were 0.47, 0.37 and 0.36, whereas intensity ratios for residues in nearby α-helices were typically between 0.7 and 0.8. This linker demonstrates significantly faster motion than any other inter-helical linkers in the HD subdomain, such as region 690-702, which in pdb 4DQY links the αB helix and the region which contains the short αC helix and the αD helix, or residues from 740-750, which link the αE and αF helices (although data was missing for some residues in this sequence). In the ART subdomain, which contains a far lower proportion of

the sequence with α-helical properties and more beta sheet and loop regions, no regions (with the exception of the disordered and highly mobile C-terminus) exhibited noticeable differences from the average $T_{1\rho}$ value (Figure 4.10) but hNOE intensity ratios (Figure 4.11) were significantly lower for residues S939, K940, L941 and G944 (values of 0.46, 0.35, 0.44 and 0.27 respectively) which appear to have no secondary structure in pdb 4DQY. However,

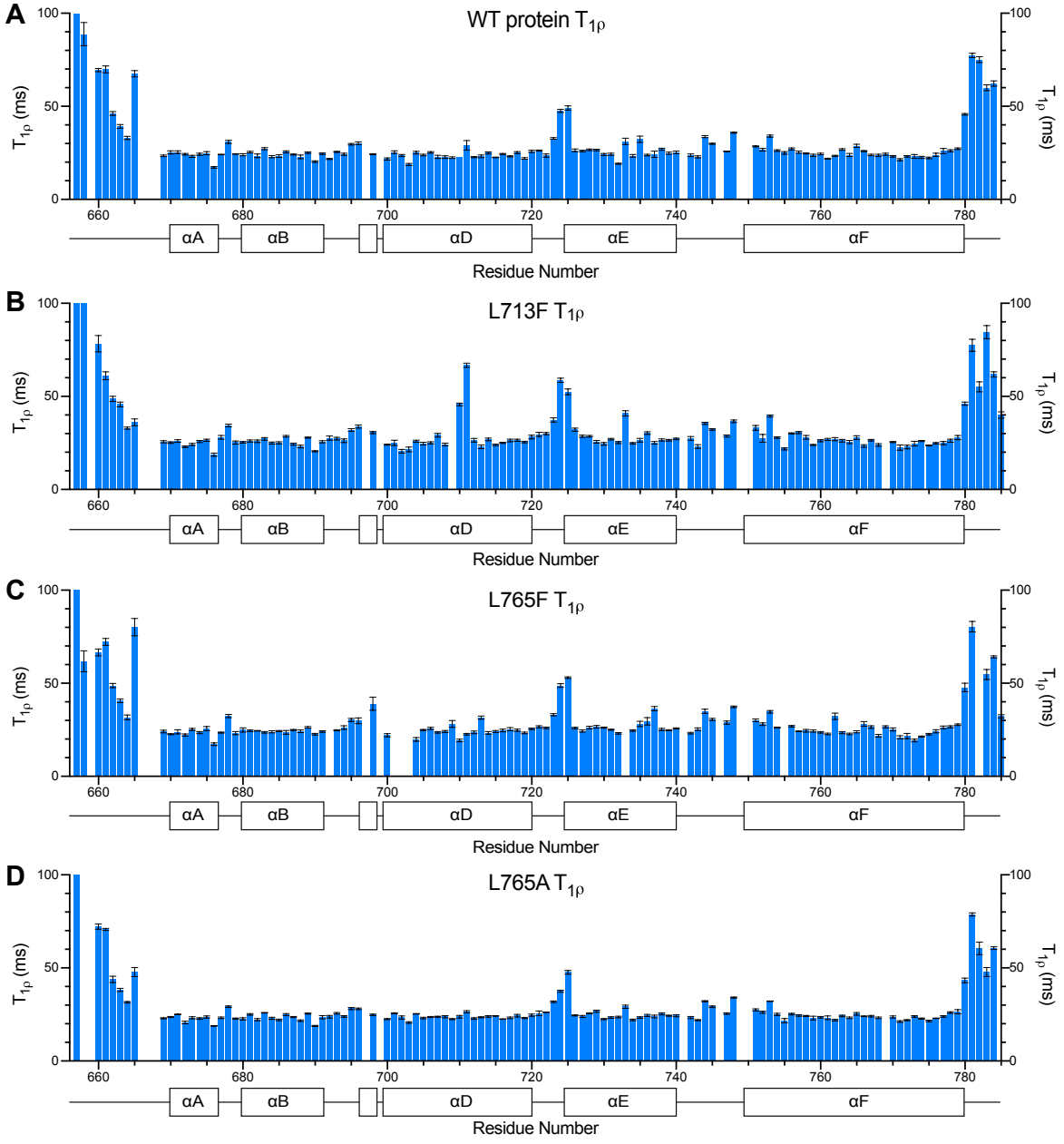


Figure 4.8 ^{15}N $T_{1\rho}$ measurements of HD subdomain residues from wild-type and mutant PARP-1 catalytic domain. (A), (B), (C) and (D). ^{15}N $T_{1\rho}$ measurements of HD subdomain residues (657-785) from wild-type catalytic domain and L713F, L765F and L765A mutants respectively. All data was recorded at 25°C from 400 µM samples of ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C). Error bars were plotted according to the data fitting error in Sparky as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

in general the residues of the active site loop exhibit no difference from the average intensity ratio of the hNOE. Furthermore, the results of both the $T_{1\rho}$ and hNOE measurements showed that there were no significant large differences in ^{15}N dynamics between wild-type PARP-1 catalytic domain and any of the constitutively partially active catalytic domain mutants. There was also no significant difference in τ_c between wild-type and mutant catalytic domains, indicating that there was no difference in tumbling speed of the domain in solution as a result of the HD subdomain mutations, or any difference in τ_c between the HD and ART

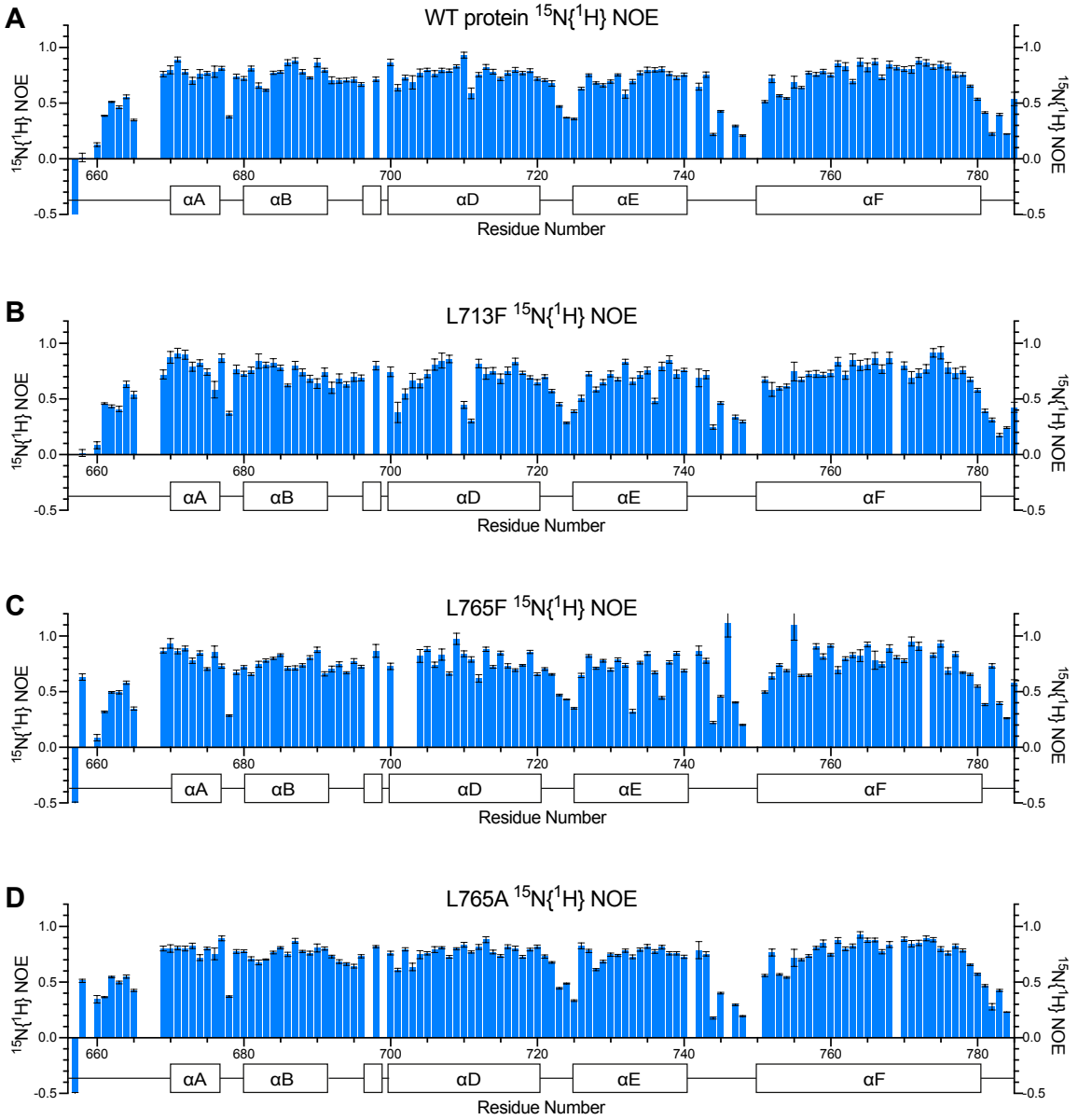


Figure 4.9 $^{15}N\{^1H\}$ NOE data from HD subdomain residues in wild-type and mutant PARP-1 catalytic domain. (A), (B), (C) and (D). $^{15}N\{^1H\}$ steady-state NOE data measurements of HD subdomain residues (656-785) in wild-type catalytic domain and L713F, L765F and L765A mutants respectively. All data was recorded at 25°C from 400 µM samples of ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^1H/^{15}N/^{13}C$). Error bars were plotted according to standard deviation determined on the basis of background noise levels as described in Section 2.3.4. Secondary structure boundaries are shown below each plot.

subdomains (Figure 4.12). This reveals that despite their separation by a highly flexible linker, both subdomains tumble as a single unit in solution and do not tumble independently. Overall, τ_c for the majority of backbone amide groups across the wild-type and mutant catalytic domains was approximately 25 ns, as would be expected for a single folded globular domain approximately 40 kDa in size (Rossi et al., 2010). In conjunction with the chemical

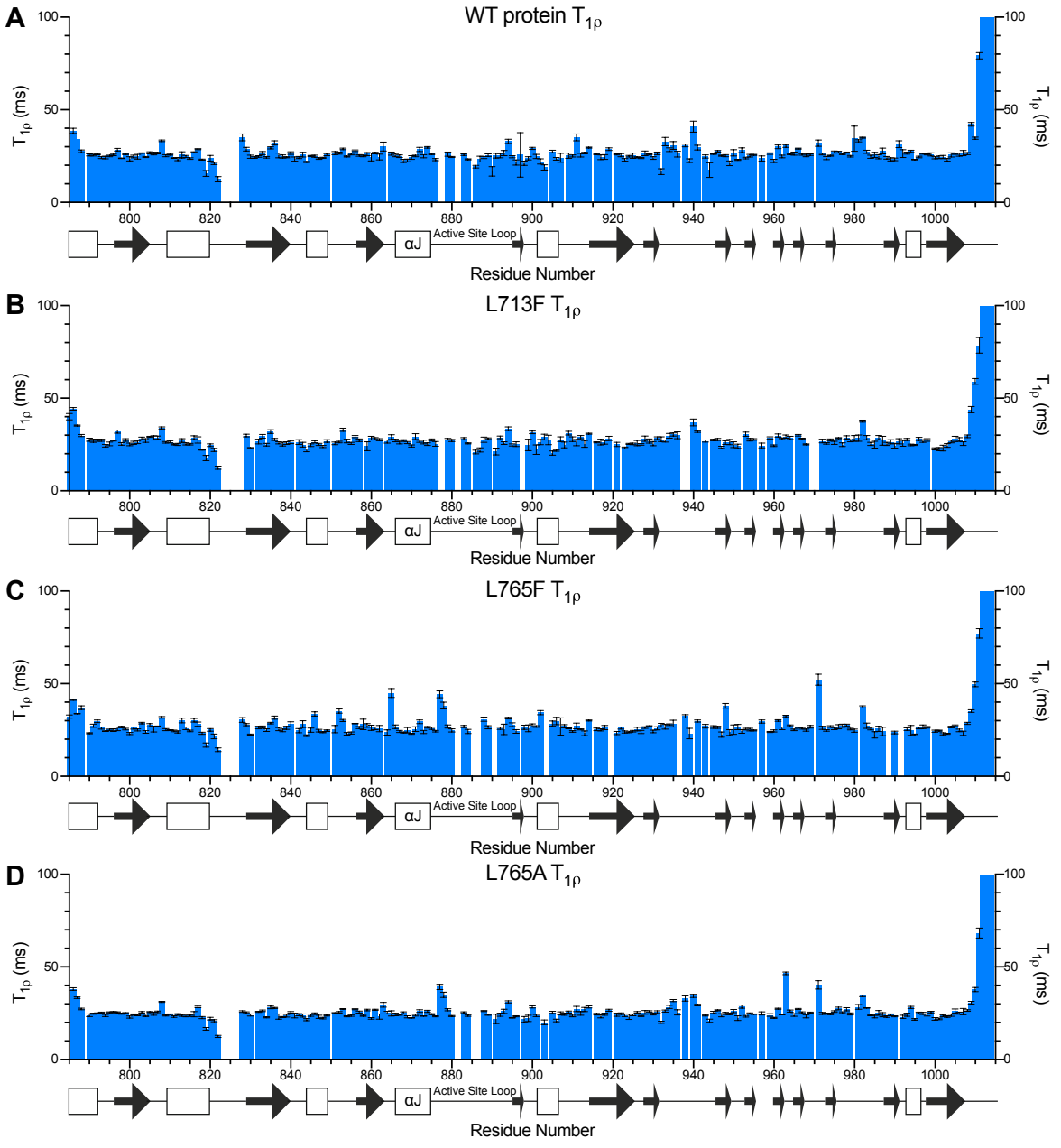


Figure 4.10 ^{15}N $T_{1\rho}$ measurements of ART subdomain residues from wild-type and mutant PARP-1 catalytic domain. (A), (B), (C) and (D). ^{15}N $T_{1\rho}$ measurements of ART subdomain residues (786-1014) from wild-type catalytic domain and L713F, L765F and L765A mutants respectively. All data was recorded at 25°C from 400 µM samples of ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C). Error bars were plotted according to the data fitting error in Sparky as described in Section 2.3.4. Secondary structure boundaries are shown below each plot.

shift perturbations observed for the HD subdomain mutations, this comparison of ^{15}N relaxation parameters between wild-type catalytic domain suggests that the difference in catalytic activity between the wild-type and mutant catalytic domains is the result of disruption of the packing between the α-helices and the active site, as well as between the α-helices themselves.

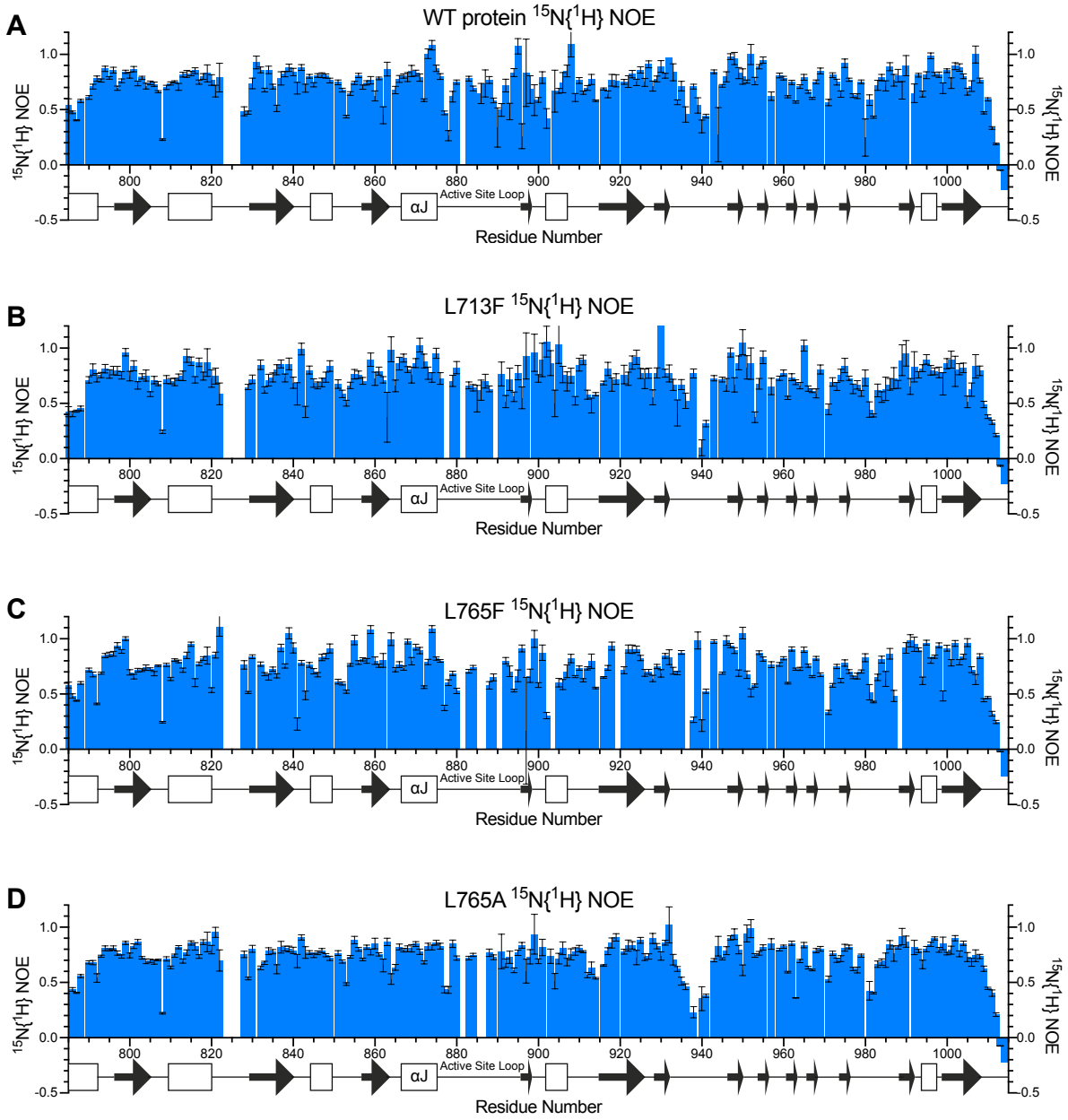


Figure 4.11 $^{15}N\{^{1}H\}$ NOE data from ART subdomain residues in wild-type and mutant PARP-1 catalytic domain. (A), (B), (C) and (D). $^{15}N\{^{1}H\}$ steady-state NOE data measurements of ART subdomain residues (786-1014) from wild-type catalytic domain and L713F, L765F and L765A mutants respectively. All data was recorded at 25°C from 400 μM samples of ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{1}H/^{15}N/^{13}C$). Error bars were plotted according to standard deviation determined on the basis of background noise levels as described in Section 2.3.4. Secondary structure boundaries are shown below each plot.

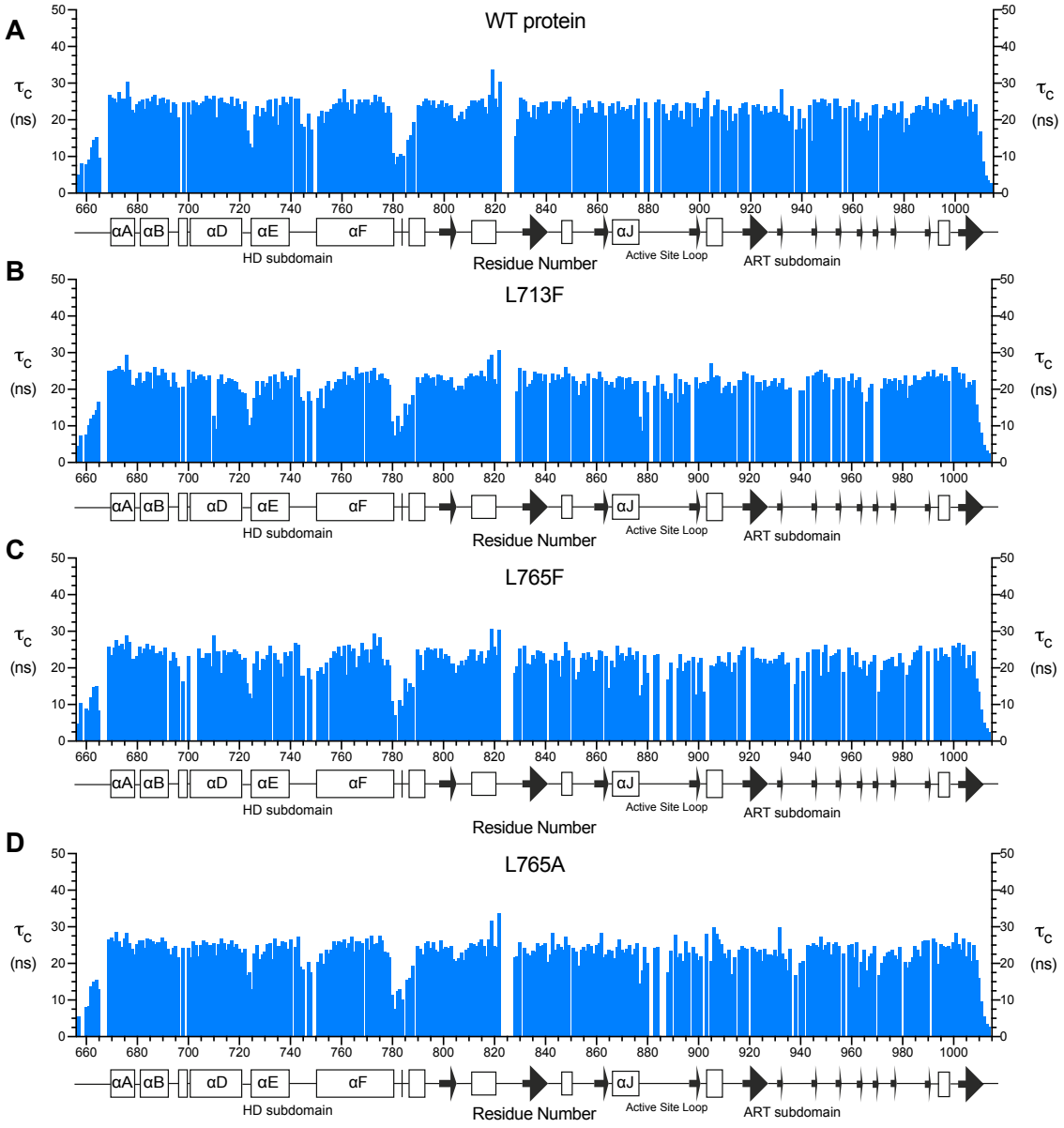


Figure 4.12 **τ_c values of wild-type and mutant PARP-1 catalytic domains.** (A), (B), (C) and (D). τ_c values calculated from T_1 and $T_{1\rho}$ measurements for wild-type catalytic domain and L713F, L765F and L765A mutants respectively. All T_1 and $T_{1\rho}$ measurements were recorded at 25°C from 400 µM samples of ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^1H/^{15}N/^{13}C$). τ_c values were calculated as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

4.5 Discussion

In PARP-1 catalytic domain, an increased rate of DNA-independent catalytic activity caused by disruption of α-helical packing by the HD subdomain mutations would be consistent with the autoinhibitory role of the HD subdomain identified by the Pascal group (Dawicki-McKenna et al., 2015), and explains why the basal catalytic activity of the mutants is greater than unactivated wild-type catalytic domain, but lower than full-length PARP-1 activated by DNA damage binding. The most active form of PARP-1 identified by the Pascal group in their previous studies is PARP-1 ΔHD, which exhibits constitutive, DNA-independent hyperactivation of PARP-1 in the complete absence of the HD subdomain (Dawicki-McKenna et al., 2015), and their subsequent kinetic studies of PARP-1 binding to the non-hydrolysable NAD^+ analogue benzamide adenine dinucleotide (BAD) confirmed that in the absence of activation by DNA damage, the PARP-1 HD subdomain sterically blocks NAD^+ from binding to the catalytic domain, rather than allowing binding but in a non-productive manner, as had previously been hypothesised might be possible (Langelier et al., 2018).

The location of the chemical shift changes observed in the catalytic domain mutants partially correspond to the areas of greatest deprotection following DNA binding by full-length wild-type PARP-1 in the Pascal group's previously described HXMS experiments. In the ART subdomain, CSPs were observed for the active site loop between residues 880 and 900 for all three mutants. In the HD subdomain, the areas of greatest deprotection after 100 seconds in the HXMS study were located in the αB helix between residues 680 and 700 (in close proximity to the αF helix in the catalytic domain structure), for which no large CSPs were observed in the mutants, and the C-terminal half of the αF helix (residues 765 to 780), for which CSPs were observed for all three mutants. In contrast, the location of the CSPs in all three mutations had greater similarities with the location of the differences in solvent exchange between full length wild-type and PARP-1 L713F following activation (Figure 4.13), with deprotection observed in the αD, αE and αF helices, and differences in deprotection in the αB helix only observed at later time points (10^4 and 10^5 seconds). In Section 5.4 of this thesis, slow solvent exchange of backbone amides in free PARP-1 catalytic domains and PARP inhibitor complexes was investigated over long time scales (3 hours to 48 hours) by re-suspending lyophilised PARP-1 catalytic domain in D_2O-based NMR buffer and measuring peak heights from TROSY spectra over time. This data revealed changes in secondary structure in key α-helices that were not detected by relaxation

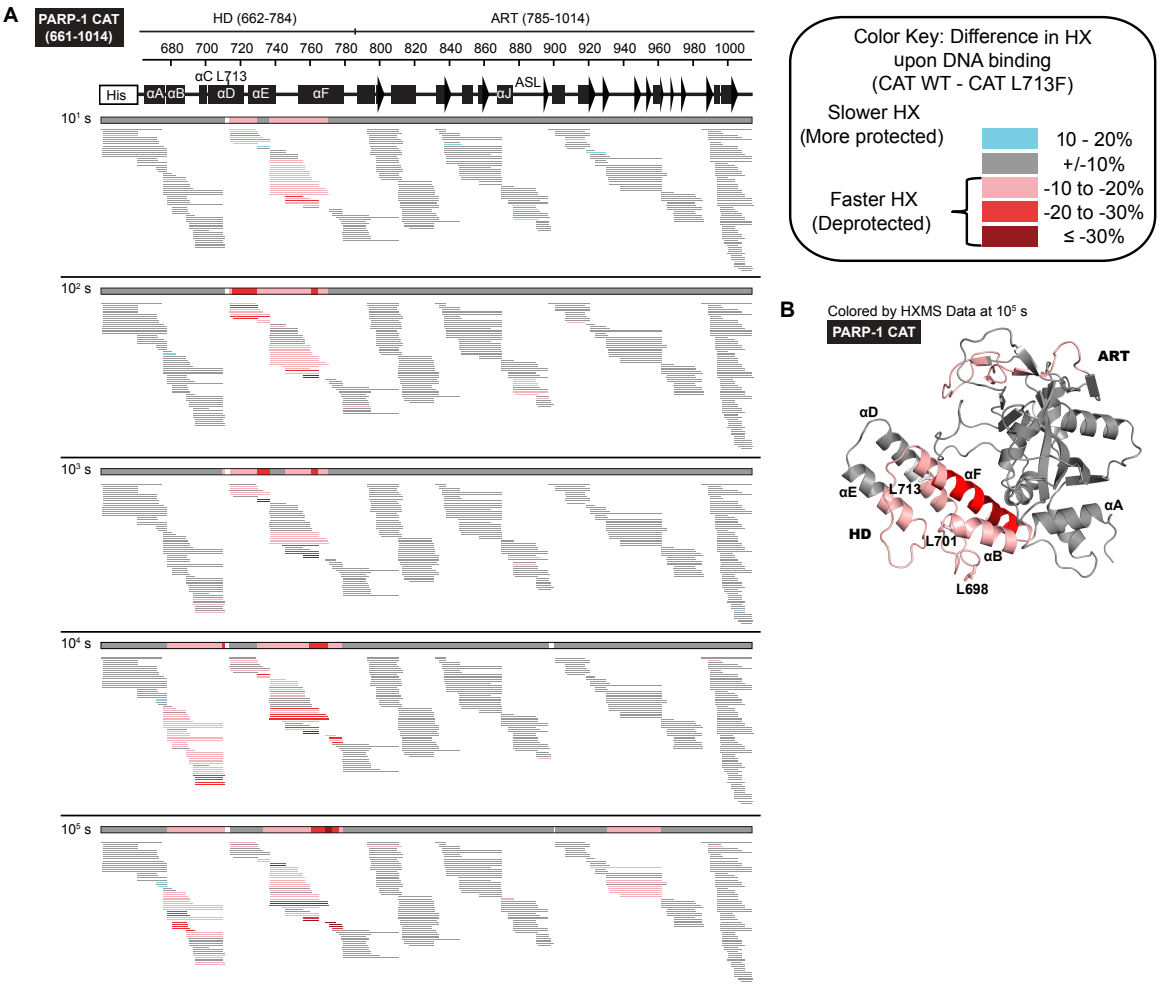


Figure 4.13 Differences in HX following DNA binding between full-length PARP-1 and PARP-1 L713F. (A) Percentage difference in catalytic domain deuteration following subtraction of L713F deuteration from wild-type full length PARP-1 at 10^1, 10^2, 10^3, 10^4 and 10^5 seconds. (B) HXMS data from the 10^5 s time point in (A) mapped onto the structure of PARP-1 catalytic domain (pdb 4DQY). Reproduced from Dawicki-McKenna *et al*, 2015.

experiments and could be compared with the Pascal group's HXMS data from later time points such as 10^5 seconds (27.78 hours) at the level of individual residues. Consequently, further characterisation of key α-helical regions in the HD subdomains of the catalytic domain mutants and comparison to wild type catalytic domain could potentially be performed using an experiment of this nature, and could possibly either provide new evidence of destabilisation or partial unfolding, or confirm its absence. If spectra could be acquired from full-length PARP-1 containing segmentally labelled wild-type PARP-1 catalytic domain in the presence of DNA damage, a speculative and highly challenging experiment discussed further in Chapter 6, NMR data from the HD subdomain mutants could be directly compared with data from fully-activated PARP-1 catalytic domain in order to further characterise both the wild-type activation process and the impact of the HD subdomain mutations.

In energetic terms, it could be suggested that mutation of key residues of the HD subdomain may lower the Gibbs free energy of activation required for the formation of ADP-ribose from NAD^+ in comparison to that required by the wild-type catalytic domain. PARP-1 activity *in vivo* is tightly regulated, as PARP-1 activation depletes vital cellular reserves of NAD^+. In the context of wild-type full-length PARP-1, the high energetic barrier to activation resulting from the presence of the wild-type HD subdomain could act to maintain the low basal rate of PARP-1 catalytic activity in the absence of DNA damage. The activation process of PARP-1 and propagation of the activation signal to the catalytic domain at the active site following detection of DNA damage by the N-terminal zinc-finger domains via the stepwise formation of interdomain contacts could then provide a sufficiently high quantity of energy to overcome this barrier, by altering the conformational state of the HD subdomain. However, activation of wild-type full-length PARP-1 in the presence of DNA damage does not represent the greatest possible modification of the rate of catalysis by the HD subdomain for the reaction, given that not only does the complete removal of the HD subdomain lead to constitutive hyperactivation in PARP1, but HXMS studies of the L713F mutation in the context of full-length PARP-1 showed that it displayed still greater deprotection in helical regions of the HD than that of wild-type even in the presence of DNA damage (Dawicki-McKenna et al., 2015), suggesting that even in the presence of DNA damage, key helical regions of the HD subdomain may restrain production of ADP-ribose to some extent in a full-length, post-activation context. In terms of Gibbs free energy required for activation of ADP-ribose synthesis from the NAD^+ substrate, mutation of residues in α-helices of the HD subdomain could therefore represent an intermediate level between unactivated and fully activated catalytic domain in full-length PARP-1. In that case, the state of the HD subdomain in the catalytic domain mutants, in which changes in helical packing or another process lower the energetic requirement of the reaction, may represent a partial step towards the more extensive changes observed in full-length wild-type PARP-1 following detection of DNA damage.

Chapter Five

Studies of inhibition in PARP-1 catalytic domain

In addition to its role in the DNA damage response, PARP-1 is also of biomedical interest due to the role of PARP inhibitors in anti-cancer therapies. Inhibition of PARP-1 catalytic activity by small molecules is selectively cytotoxic to tumour cells with deficient homologous repair machinery arising from mutations in the tumour-suppressing *BRCA1* or *BRCA2* genes, as these cells are dependent on repair mediated by PARP-1 and ADP-ribosylation. Numerous compounds that inhibit catalytic domain activity have been discovered, the majority of which bind with high affinity to the PARP signature region of the ART subdomain containing the active site loop. However, the cytotoxicity of these compounds *in vivo* has been shown to vary considerably, and even through orders of magnitude according to some estimates. The basis of this variation, and the mechanism of clinically availability PARP inhibitors *in vivo*, remains to be definitively established. Furthermore, no published structural data is available for many PARP inhibitors in complex with PARP-1 catalytic domain. In this chapter, complexes of wild-type PARP-1 catalytic domain with three clinically available PARP inhibitors with differential inhibitory efficacy *in vivo* were investigated using NMR spectroscopy and interpreted with the help of crystal structures. Additional comparisons of data from NMR experiments were carried out with a fourth PARP inhibitor, EB-47, for which no crystallographic data was available.

5.1 PARP inhibitors

The majority of all PARP inhibitors discovered so far, including those investigated during this study, mimic the binding of NAD^+'s nicotinamide moiety to the conserved PARP signature in the ART subdomain of PARP-1 catalytic domain, forming hydrogen bonds with residues from the conserved NAD^+ binding site (Steffen et al., 2013). Of the four PARP inhibitors used in this work, the compounds veliparib, olaparib and talazoparib were chosen because they are either clinically available or undergoing clinical trials for treatment of tumours with BRCA mutations, the kinetics of their binding to PARP-1 have been studied, and they exhibit well-characterised differences in cytotoxicity *in vivo*, whether as a result of differential ability to form trapped PARP-1 lesions or via inhibition of the ADP-ribosylation signal in response to DNA damage. Of these three compounds, veliparib (developed by

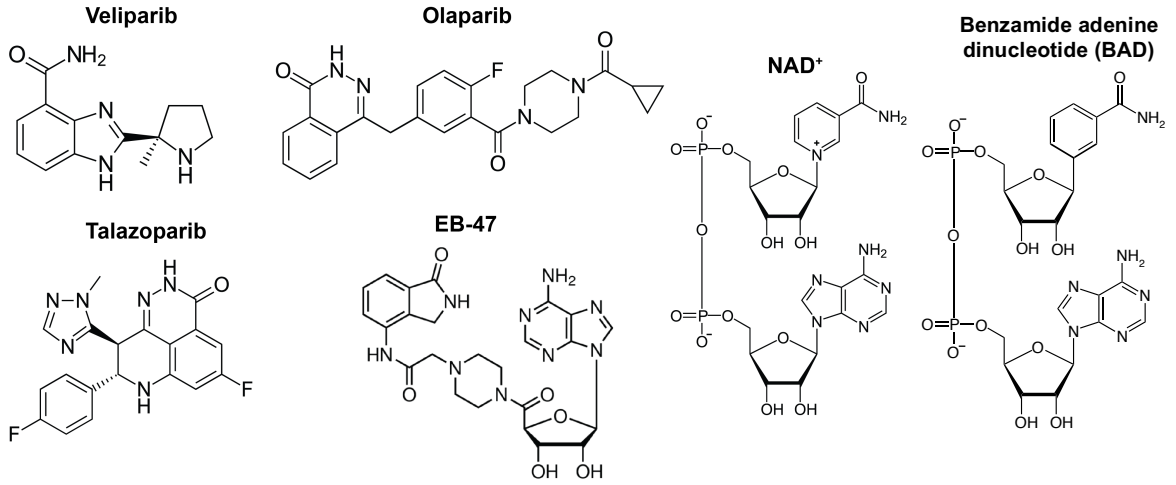


Figure 5.1 Chemical structures of PARP inhibitors used in this chapter, NAD$^+$ and the non-hydrolysable NAD$^+$ analogue benzamide adenine dinucleotide (BAD). In BAD, carbon replaces the nitrogen atom in the nicotinamide ring of NAD$^+$, forming a benzamide moiety.

AbbVie) has the lowest binding affinity and cytotoxicity, with olaparib (developed by AstraZeneca, and the first PARP inhibitor to be clinically approved and reach the market) the next highest in both properties and talazoparib (Pfizer) having the highest binding affinity and cytotoxicity of all clinically investigated compounds to date (Table 5.1).

	Veliparib	**Olaparib**	**Talazoparib**	**EB-47**
Molecular weight	244.29	435.08	380.35	610.45
Relative cytotoxicity *in vivo* (trapping)	Lowest (10-fold less potent than olaparib[*])	Middle (10-fold more potent than veliparib[*])	Highest (100-fold more potent than olaparib[**])	Unknown
K_D (no DNA, TR-FRET, nM)	2.0 ± 0.2[†]	0.86 ± 0.26[†]	0.47 ± 0.11[†]	Unknown
K_D (+SSB DNA, TR-FRET, nM)	3.3 ± 2.3[†]	1.3 ± 0.7[†]	0.5 ± 0.26[†]	Unknown
K_D (Biacore, nM)	4.4[†]	1.3[†]	0.17[†]	Unknown
k_d (Biacore, s^{-1})	7.0 x 10^{-3}[†]	3.2 x 10^{-4}[†]	6.3 x 10^{-5}[†]	Unknown
K_i (Biacore, nM)	3.7 ± 0.2[†]	1.3 ± 0.5[†]	0.5 ± 0.1[†]	Unknown
IC_{50} (nM)	39 ± 12[†]	7.9 ± 2.2[†]	4.1 ± 0.9[†]	45[‡]

Table 5.1 Clinical efficacy, biochemical data and kinetic data for PARP inhibitors studied in this chapter. Molecular weight and reported relative cytotoxicity ("trapping" as defined by potentiation of DNA damage from alkylating agents *in vivo*), equilibrium dissociation constants (K_D), off-rates (k_d) inhibition constants (K_i) and IC_{50} (half maximal inhibitory concentration) values of PARP inhibitors used in this chapter. [*]Murai *et al*, 2012. [**]Murai *et al*, 2014. [†]Hopkins *et al*, 2015. [‡]Jagtap *et al*, 2004.TR-FRET: Time-resolved fluorescence energy transfer.

However, with the exception of talazoparib (pdb 4PJT, Aoyagi-Scharber et al., 2014), no crystal structures of these compounds in complex with the whole wild-type PARP-1 catalytic domain have been published, with crystal structures including veliparib and olaparib either limited to complexes with PARP-2 catalytic domain or other members of the PARP family (Karlberg et al., 2010; Qiu et al., 2014), or solved in complex with PARP-1 catalytic domain ΔHD (Dawicki-McKenna et al., 2015), revealing the structural basis for binding of these compounds to the catalytic domain's active site but providing no insight into the behaviour of the HD subdomain following binding of the compounds. One mechanism that has been proposed by which binding of inhibitors at the active site location can cause *in vivo* toxicity involves "trapping" of PARP-1 proteins at sites of DNA damage, but competing hypotheses exist as to how this may occur *in vivo* for these three compounds. It has been suggested that trapping of these compounds is mediated through an allosteric pathway that in effect resembles the DNA damage activation pathway in reverse (Murai et al., 2012) and in which the HD subdomain would consequently play a pivotal role; if the structural changes in the HD subdomain that are responsible for the relief of autoinhibition following DNA damage recognition were partially or completely prevented by the presence of an inhibitor, then this would presumably increase the overall free energy change associated with DNA binding.

However, *in vitro* biochemical assays have suggested that these compounds are not able to increase the affinity of PARP-1 to DNA damage models in the absence of NAD^+ and that even in the presence of inhibitors and NAD^+, the slower off-rates of PARP-1 are directly correlated to binding affinity and the ability of compounds to inhibit ADP-ribosylation in a linear fashion, suggesting that trapping is caused by disrupted formation of the long negatively charged ADP-ribose chains required to discharge PARP-1 from sites of DNA damage (Hopkins et al., 2015). The importance of contact formation with the HD subdomain in this process in these three compounds has yet to be determined, as the structure of the whole catalytic domain has only been solved in complex with talazoparib, and not for other the other compounds used in this study that have lower binding affinities (Table 5.1). Therefore, investigation of the HD subdomain's role in PARP inhibition is a critical part of the evaluation of both the structural basis of differences in effectiveness between compounds and clarification of the mechanistic basis of trapping. In addition to the absence of published structural data for the HD subdomain in complexes of PARP-1 catalytic domain with PARP inhibitors, to our knowledge no attempts to compare dynamic differences between

complexes across different residues of the catalytic domain or the HD subdomain in solution have previously been made.

A fourth PARP inhibitor, EB-47, was also included following discussion with our collaborator John Pascal, who had found he was unable to obtain a crystal structure of a complex of EB-47 with the PARP-1 catalytic domain. We jointly speculated that this could be because binding of EB-47 might de-stabilise the HD subdomain (unpublished personal communication). The chemical structure of EB-47 (Figure 5.1) is significantly more similar to the structure of NAD^+ than other PARP inhibitors, with a far greater molecular weight (610.45 g/mol) that is closer to the molecular weight of NAD^+ (663.43 g/mol). Consequently, it could be the case that when bound, parts of the EB-47 structure occupy spatial regions that clash with the HD subdomain in a way that other, smaller, PARP inhibitors do not. To date, only the structure of PARP-2 CAT ΔHD has been solved by the Pascal group in complex with EB-47 (Dawicki-McKenna et al., 2015). The effectiveness *in vivo* of EB-47 in tumour cells, whether on the basis of cytotoxicity or hypothetical trapping ability, has not been evaluated, but with an IC_{50} of 45 nM for PARP-1 (Jagtap et al., 2004; Table 5.1), EB-47 exhibits high affinity for the PARP signature region and displays potent PARP inhibition *in vitro* in a similar manner to the other compounds investigated in this chapter. The chemical structures of all four inhibitors and NAD^+ are shown in Figure 5.1 and published biological, biochemical and kinetic data is summarised in Table 5.1.

5.2 Mapping of chemical shift perturbations in PARP inhibitor complexes

The chemical shifts of nuclei in a protein, particularly amide group protons, are highly sensitive to non-covalent interactions with binding partners. Furthermore, changes in chemical shifts can result not only from the primary effects of binding at the interaction site, but also from the secondary effects of conformational change at locations far away from the binding site. As with the HD subdomain mutants described in Chapter 4, measurements of perturbations in chemical shifts of backbone amides following binding of PARP inhibitors would allow us to probe changes in the catalytic domain at the level of individual residues. Furthermore, NMR experiments could be carried out in solution in the intact wild-type catalytic domain fragment, including the HD subdomain. In order to characterise, compare and contrast the binding of veliparib, olaparib, talazoparib and EB-47 to the catalytic domain, ^{15}N-labelled wild-type catalytic domain was titrated with stocks of these compounds re-suspended in DMSO-d_6 as described in Section 2.3.3. $^{15}N^1H$ TROSY spectra were recorded

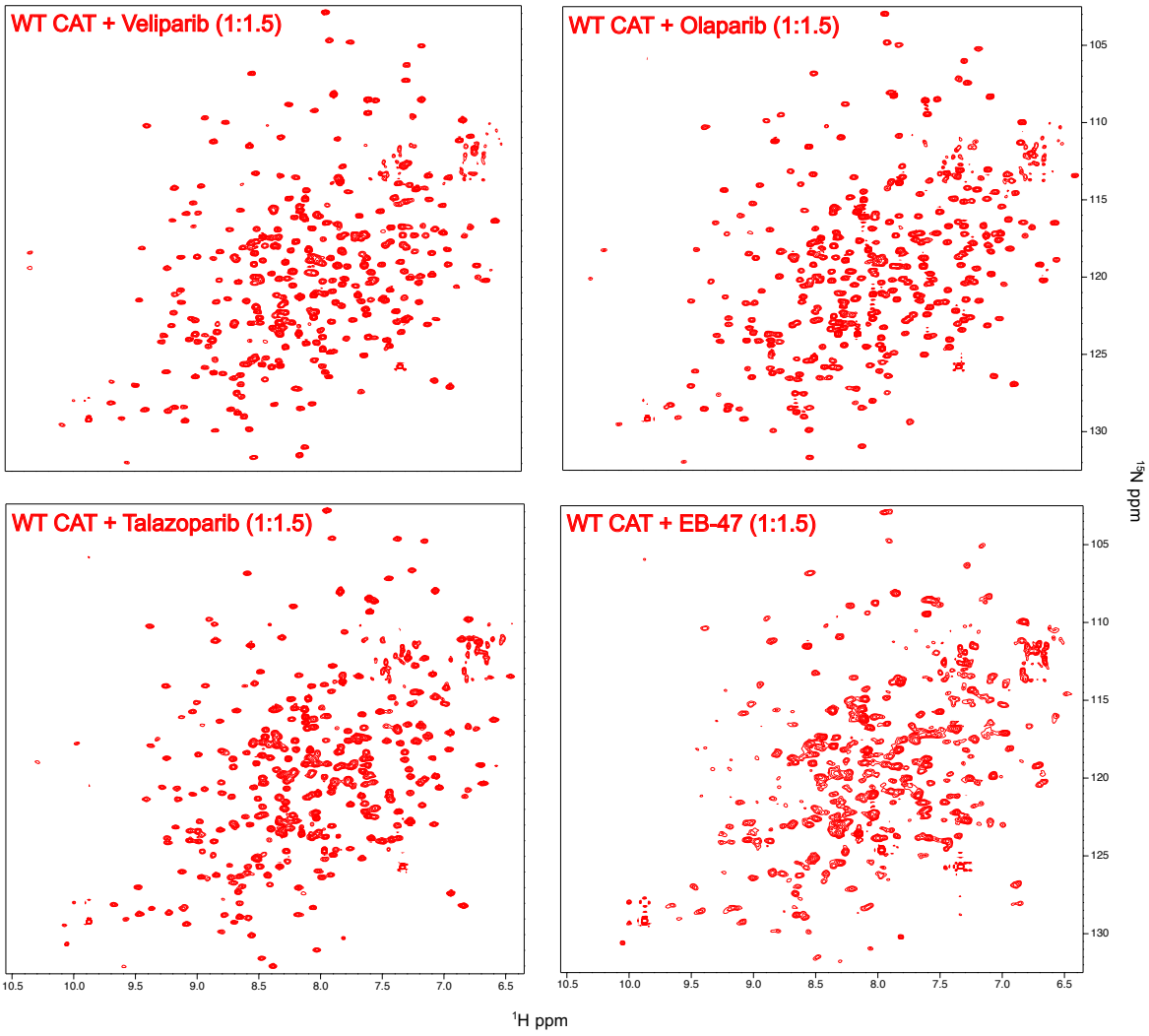


Figure 5.2 [^{15}N ^{1}H] TROSY spectra of PARP-1 catalytic domain in complex with veliparib, olaparib, talazoparib and EB-47. All spectra were acquired at 25°C from samples of ^{15}N labelled catalytic domain mutants in NMR buffer at concentrations of 300 µM (veliparib complex), 400 µM (olaparib and EB-47 complexes) or 500 µM (talazoparib complex). Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C)

(Figure 5.2), peaks were assigned using ^{15}N NOESY-HSQC experiments described in Section 2.3.2 and compared to the reference TROSY spectrum of free wild-type catalytic domain (Figure 5.3); chemical shift perturbations were then calculated as described in Section 2.3.3 and plotted by residue (Figure 5.4). During the titrations, the catalytic domain became saturated at a 1:1 ratio for all four compounds and no further chemical shift perturbations were observed beyond this point. Nonetheless, all CSPs were calculated following titration of compounds at a 1:1.5 ratio to ensure that all titrations had reached their endpoints. Chemical shift perturbations from each complex were then mapped onto crystal structures of that complex as described in Section 2.3.3, with the exception of the EB-47 complex, which was mapped onto the structure of free wild-type catalytic domain because no solved structure of that complex was available. CSPs from talazoparib binding were

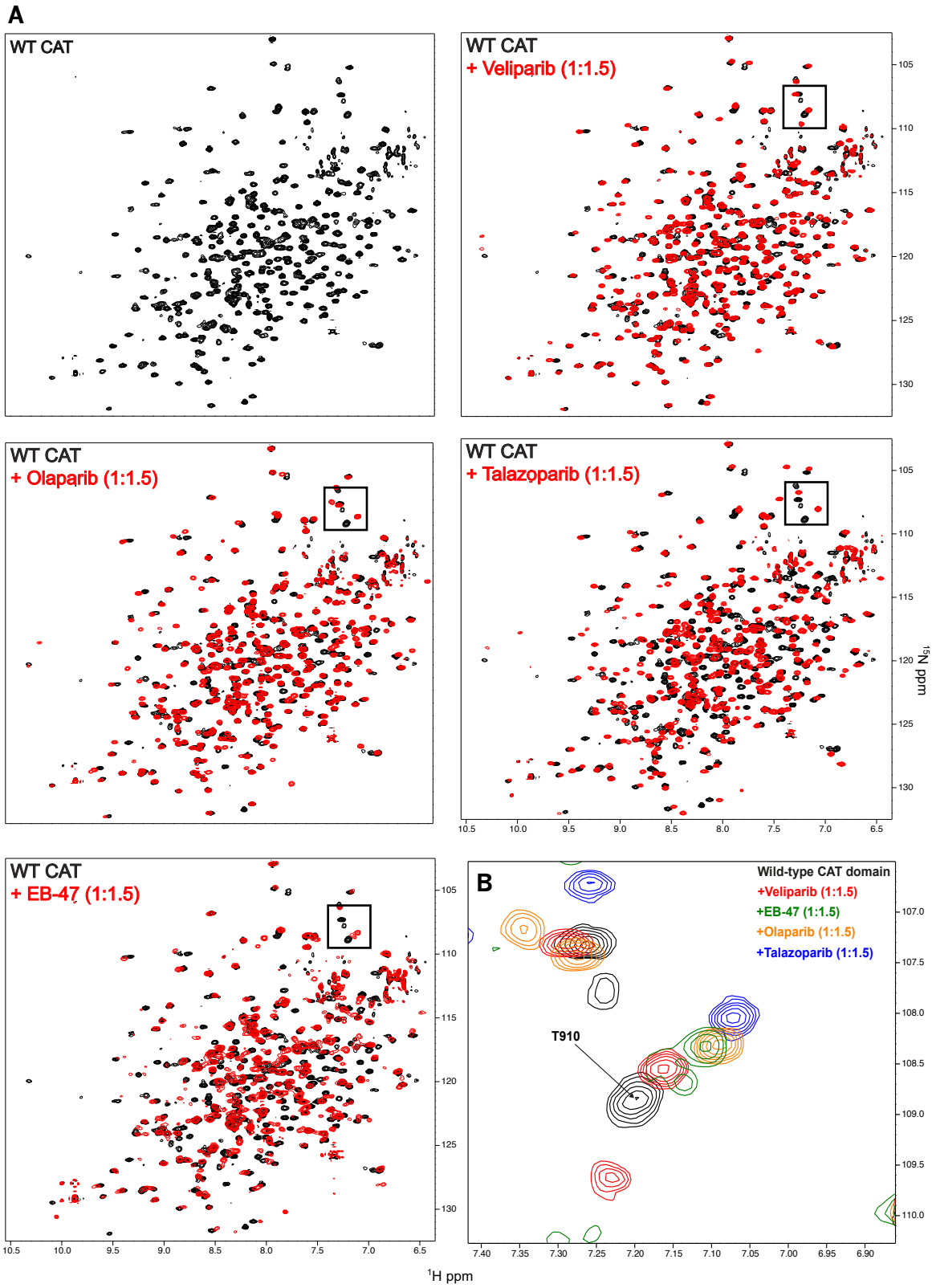


Figure 5.3 Comparison of [^{15}N ^{1}H] TROSY spectra of free PARP-1 catalytic domain and catalytic domain in complex with PARP inhibitors. (A) Spectra of inhibitor complexes shown in Figure 5.2 (red) are superimposed on the spectrum of free wild-type catalytic domain (black). The free catalytic domain spectrum had a concentration of 400 µM and data was collected under the same conditions described for the inhibitor complexes in Figure 5.2. (B) Superposition of expanded spectra from free catalytic domain and all PARP inhibitor complexes to show chemical shift perturbations for the active site residue T910.

mapped onto the co-ordinates of pdb 4PJT (Aoyagi-Scharber et al., 2014) and CSPs from complexes of the catalytic domain with veliparib and olaparib were mapped onto previously unpublished crystal structures of these complexes solved by Marianne Schimpl of AstraZeneca; we gratefully acknowledge her making these co-ordinates available to us. As with mapping of CSPs from catalytic domain mutants described in Chapter 4, the upper threshold was set to 0.2, and among all CSP data collected for PARP-1 catalytic domain, the largest CSP value is 1.247 (residue G888 in the talazoparib data set), meaning that a threshold of 0.2 scales all CSPs between 0.0 to 0.249 in size, and all CSPs greater in size than this are mapped with maximum colour saturation.

Visual comparison of the spectra from the four complexes (Figure 5.3), plotting of CSPs (Figure 5.4) and mapping of CSPs onto structures of the complexes (Figure 5.5) revealed both similarities and differences between complexes in the size and location of CSPs within the catalytic domain. All complexes displayed pronounced CSPs in the active site loop (corresponding approximately to residues 880-900) and the 10 residues immediately following it. The largest CSP from this region (and the complex as a whole) from the veliparib complex is found for residue F897 (0.772) and the largest CSPs from the complexes with olaparib and talazoparib are 1.21 (I879) and 1.24 (G888) respectively. In the 20 residues preceding the active site loop (residues 860-880, containing the ART subdomain helix αJ), the majority of which were under 0.2 in size, CSPs are also observed in all inhibitor complexes (Figures 5.4 and 5.5). The differences in CSP sizes between complexes is illustrated by the size of the CSP for the binding site residue T910 in each complex (Figure 5.3B). Of the three compounds, talazoparib has the largest chemical shift perturbation, followed by olaparib and then veliparib. In the case of EB-47, many assignments from the active site loop and neighbouring regions were missing in comparison to the other compounds, but the size of the largest CSP in the dataset (0.924, residue F897) and the size of the CSPs in the region 860-880 were intermediate between those of veliparib

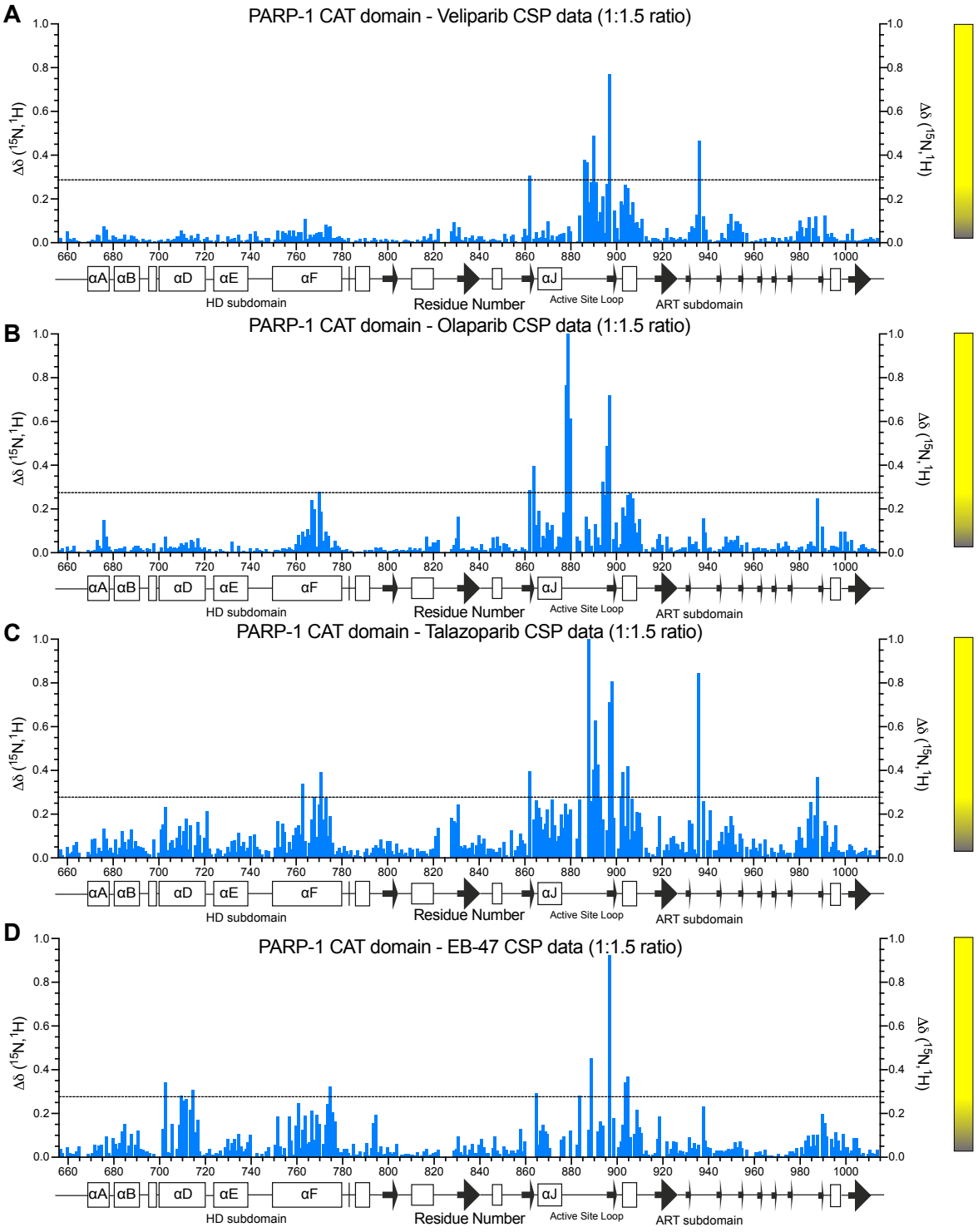


Figure 5.4 Chemical shift perturbations observed in complexes of PARP inhibitors with PARP-1 catalytic domain. (A), (B), (C) and (D). All chemical shift perturbations observed in spectra of 1:1.5 complexes of the catalytic domain with veliparib, olaparib, talazoparib and EB-47 (from Figure 5.3) respectively. Scaling of CSP mapping by colour to the catalytic domain structure in Figure 5.6 is shown next to each plot. Subdomain and secondary structure boundaries are shown below each plot. Chemical shift perturbations below 0.249, the upper limit of CSP scaling by colour, are enlarged in (E), (F), (G) and (H).

and olaparib, although it is possible that residues that remain unassigned could have CSPs that are greater in size than this, as these are typically the most difficult assignments to make in protein-drug complexes. Unlike for other compounds investigated, two peaks were

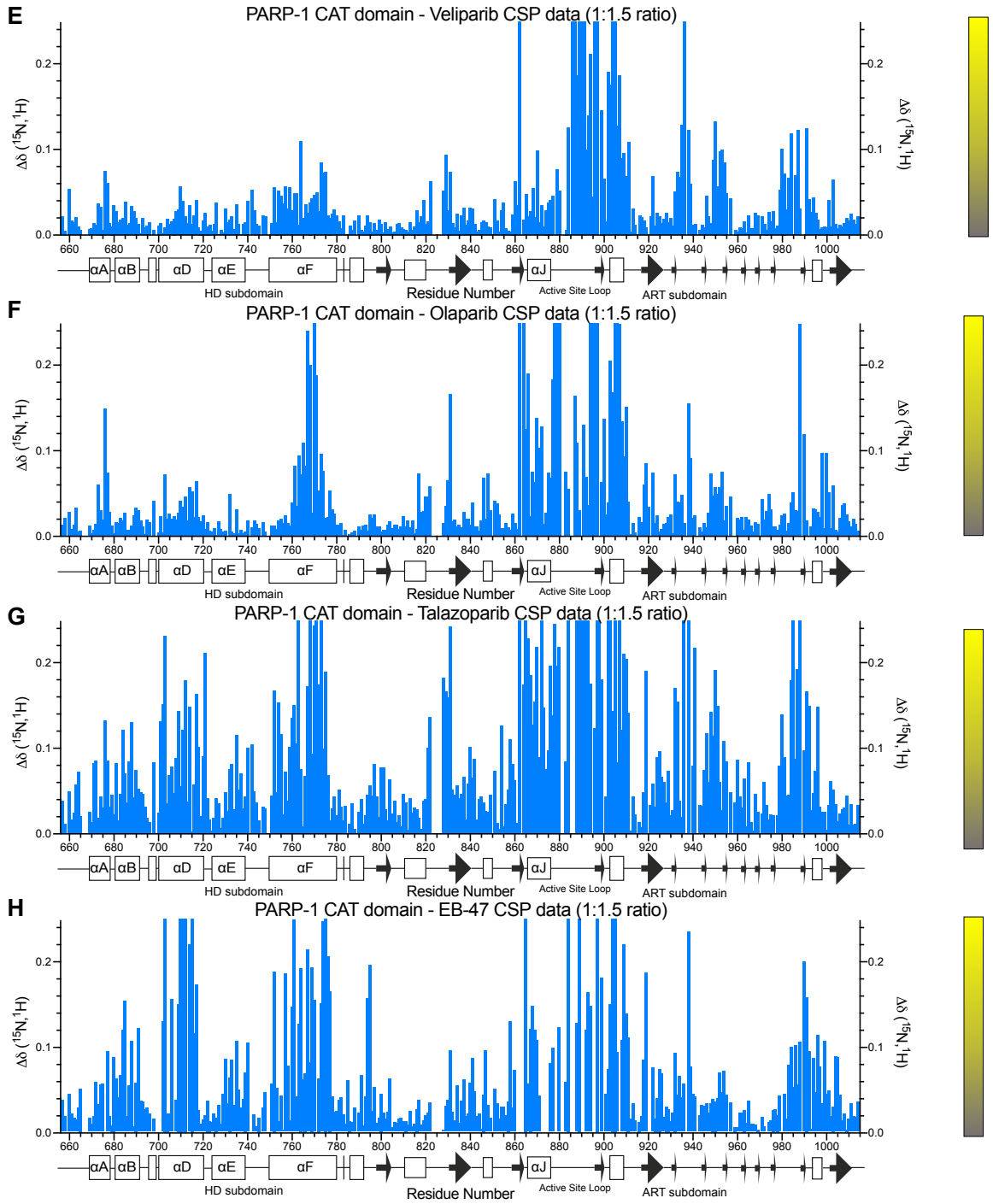


Figure 5.4 Chemical shift perturbations observed in complexes of PARP inhibitors with PARP-1 catalytic domain. (E), (F), (G) and (H). Enlarged plots of chemical shift perturbations below 0.249 from (A), (B), (C) and (D) respectively. Subdomain and secondary structure boundaries are shown below each plot. Scaling of CSP mapping by colour to the catalytic domain structure in Figure 4.6 is shown next to each plot.

observed at position T910 following titration of a 1:1.5 ratio of EB-47, with a single sharp peak observed that overlapped with the signal from the olaparib titration, and a broad weaker peak that overlapped with the single peak observed following titration with veliparib. This was the only peak in the spectrum of the EB-47 complex to exhibit this behaviour, although

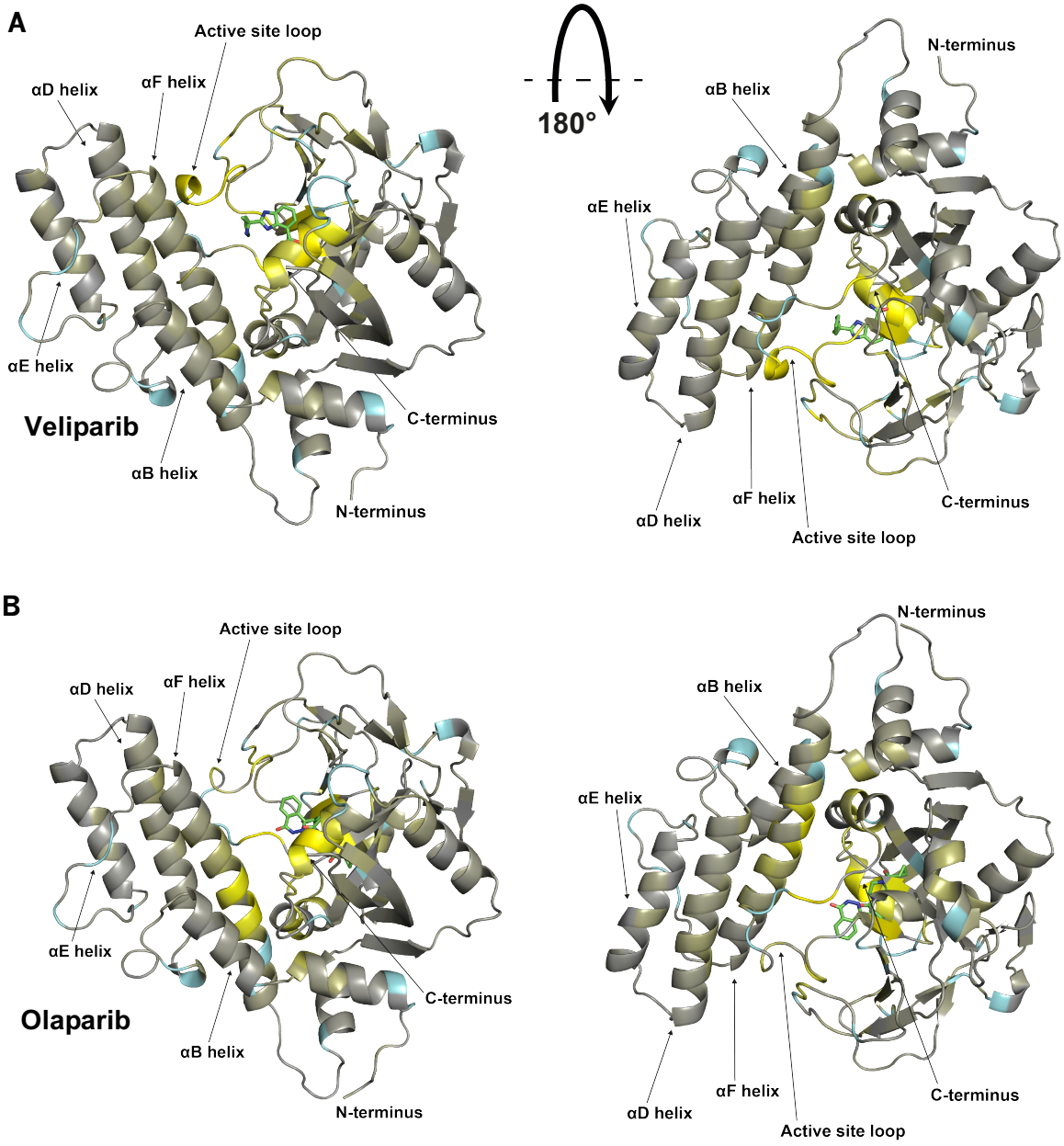


Figure 5.5 Mapping of chemical shift perturbations from PARP-1 catalytic domain in complex with veliparib, olaparib, talazoparib and EB-47. (A) and (B). Chemical shift perturbations from Figure 5.4 globally scaled and mapped onto the structure of PARP-1 catalytic domain in complex with veliparib (A) or olaparib (B) inhibitor complexes as described in Chapter 2. CSPs are shown on a gradient from grey (no CSP) to yellow (all CSPs > 0.249). Residues that are incomparable due to missing assignments (including proline residues) are shown in cyan Carbon atoms of PARP inhibitor molecules are shown in green, nitrogen atoms in blue and oxygen atoms in red.

it was clear that many signals in the TROSY spectrum of the EB-47 complex were either missing entirely or noticeably broader than their counterparts in the spectra of free catalytic domain and the other inhibitor complexes (Figure 5.2). The interpretation of these observations remains unclear. In the ART subdomain, chemical shift perturbations of differing sizes between complexes were also observed between residues 980 and 1000, a

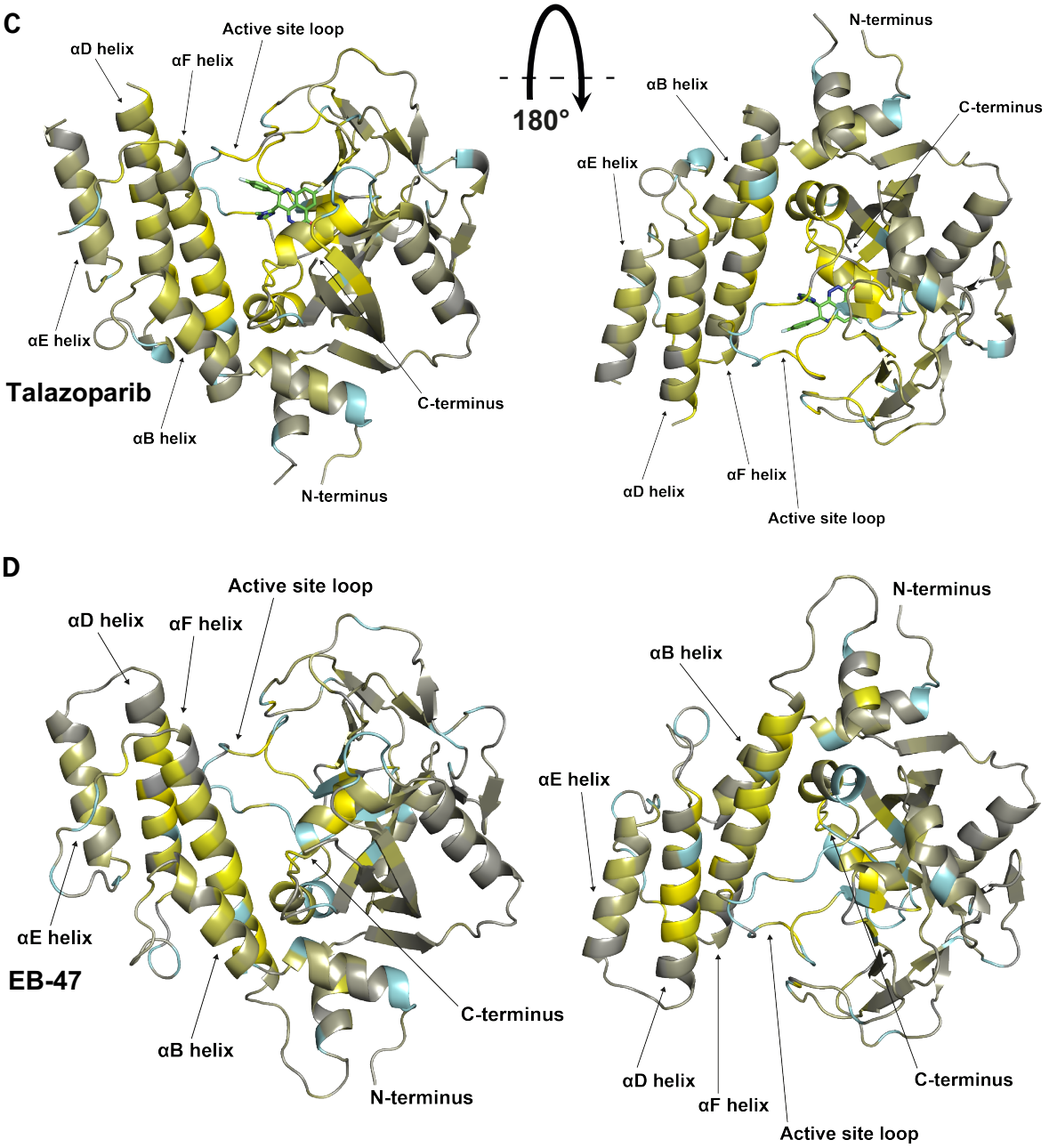


Figure 5.5 Mapping of chemical shift perturbations from PARP-1 catalytic domain in complex with veliparib, olaparib, talazoparib and EB-47. (C) and (D). Chemical shift perturbations from Figure 5.4 globally scaled and mapped onto the structure of PARP-1 catalytic domain in complex with talazoparib (C) or in the case of EB-47, free catalytic domain (D), as described in Chapter 2. CSPs are shown on a gradient from grey (no CSP) to yellow (all CSPs > 0.249). Residues that are incomparable due to missing assignments (including proline residues) are shown in cyan. Carbon atoms of PARP inhibitor molecules are shown in green, nitrogen atoms in blue and oxygen atoms in red.

region which contains the key catalytic residue E988. These CSPs were greatest in size in the talazoparib complex, followed by EB-47 and then roughly equivalent in size for veliparib and olaparib.

Although the largest CSPs observed in the HD subdomains of the complexes were smaller than those found in the ART subdomain, the differences in size and location of CSPs

in the HD subdomains between complexes were significantly larger. In the case of veliparib, no CSPs in these helices and across the HD subdomain as a whole were greater than 0.11 in size. However, in the case of olaparib, larger CSPs were observed between residues 760 and 780, corresponding to the C-terminal half of the αF helix, which borders the ART subdomain and the active site loop. In the case of talazoparib, even though (as for the other inhibitors) binding of the compound is to the PARP signature region of the ART subdomain, this causes wide-ranging CSPs across the α-helices of the HD subdomain. Talazoparib binding results in the largest and most widely-distributed CSPs in the αF helix, but also leads to CSPs in the neighbouring αD and αE helices, which do not occur for the complexes with olaparib or veliparib. Plotting of CSPs from the EB-47 complex and mapping onto the structure of free catalytic domain reveals that EB-47 binding also leads to large CSPs in the αF helix, comparable to those associated with talazoparib binding, and CSPs in residues 700-720 of the αD helix that are slightly larger than those observed for talazoparib. In the case of talazoparib and EB-47, small CSPs were also found in the αB helix (residues 680-690), which lies in close proximity to the C-terminus of the αF helix in the catalytic domain structure.

In the veliparib, olaparib and talazoparib complexes, larger chemical shift perturbations in the αF helix appear to be correlated to the size of CSPs in the ART subdomain's αJ helix, which lies in very close proximity to the C-terminus of the αF helix where the largest CSPs in the HD subdomain are observed. In the case of EB-47, many assignments between residues 860 and 880 could not be assigned due to line broadening in the spectra. The presence of these wide-ranging CSPs in talazoparib and EB-47 complexes, and the contrasting absence of these changes in veliparib and olaparib complexes, demonstrates that binding of small molecules to the PARP signature region in the ART subdomain can, in some cases but not others, lead to changes in areas of the HD subdomain that are distant from the active site, and suggests that as in the case of the HD subdomain mutants investigated in Chapter 4, effects can be propagated to more distant α-helices through the αF helix, the helix closest to the active site.

5.3 ^{15}N relaxation analysis of backbone amides in PARP-1-inhibitor complexes

In the case of veliparib, olaparib and talazoparib, CSPs in the αF and other helices of the HD subdomain appeared to be linked to the formation of contacts from the inhibitor to regions on the periphery of the NAD^+ binding pocket such as the αF helix, which could also be consistent with the hypothesised role of the HD subdomain in allosteric trapping of PARP-1 during PARP-1 inhibition. However, no large-scale structural rearrangements of the HD subdomain were observed in the crystal structures of these complexes, and in the case of the EB-47 complex, large CSPs in the HD subdomain were also observed and no crystallographic data was available to aid their interpretation.

Therefore, in order to help establish how the behaviour of key α-helices in the HD subdomain were affected by the binding of these inhibitors, we sought to further characterise the changes observed across the HD and ART subdomains following the binding of these four compounds by measuring relaxation parameters and rates of solvent exchange in these complexes using NMR spectroscopy. We hypothesised that if binding of the compounds affected the secondary structure of any regions of the catalytic domain, led to changes in the motion of linkers, or any significant change in the interdependence of the ART and HD subdomains, corresponding changes in relaxation parameters would be detected for backbone amides. We also hypothesised that binding of the compounds, perturbation of secondary structure or changes in motion of dynamic regions in the HD or ART subdomains resulting from inhibitor binding, including those over very long timescales of hours to days, should be reflected in differences in rates of solvent exchange. Therefore, as described in Section 4.4, we measured backbone ^{15}N spin-locked transverse relaxation times ($T_{1\rho}$) and $^{15}N\{^1H\}$ steady state NOEs for the HD and ART subdomains for the free catalytic domain and all four complexes (Figure 5.6-5.9), and also measured longitudinal relaxation times (T_1) in order to calculate the overall rotational correlation time (τ_c) of each complex (Figure 5.10).

In order to measure the rates of solvent exchange over very long timescales, ^{15}N-labelled, fully protonated wild-type PARP-1 catalytic domain was lyophilised overnight, re-suspended in a D_2O-based NMR buffer, and the heights of backbone amide peaks were measured by recording successive $^{15}N^1H$ TROSY spectra every three hours over a period of 48 hours as described in Section 2.3.6. We also measured the rate of proton exchange between the solvent and backbone amides over much faster timescales for free wild-type catalytic domain and all four complexes using a CLEANEX-based experiment as described

in Section 2.3.5. All solvent exchange data over these timescales is plotted by colour in Figure 5.11. Where signals were observed in the CLEANEX experiment or at varying time points in the D_2O exchange experiment, they were mapped onto the structure of the catalytic domain in Figures 5.12 and 5.13 using the same colour from Figure 5.11. Signals measured for the free catalytic domain were mapped onto pdb 4DQY, signals measured for veliparib, olaparib and talazoparib complexes were mapped onto the crystal structures of those complexes (described above), and in the absence of crystallographic data for the EB-47 complex, signals measured for that case were mapped onto pdb 4DQY. Unfortunately, we were unable use this approach to measure solvent exchange rates for residues in what might be called the "intermediate" timescale of solvent exchange, that was too slow to be detected by CLEANEX and too fast to be detected in the first TROSY timepoint, because there is no NMR experiment conveniently available to measure such rates. However, although this meant that coverage was incomplete, these experiments did allow us to evaluate solvent exchange both in particularly fast-exchanging flexible regions and also in relatively slow-exchanging regions, obtaining data at the level of individual residues.

In both $T_{1\rho}$ and heteronuclear NOE data, the particularly mobile regions of the catalytic domain that were identifiable from their relaxation times and motions in solution for the free catalytic domain, such as residues 723-725 and 780-785, were also clearly identifiable in the inhibitor complexes. In the ART subdomain, $T_{1\rho}$ data (Figure 5.8) and hNOE data (Figure 5.9) show no significant difference in relaxation between the free catalytic domain and the inhibitor complexes, although many assignments between residues 860 and 900 are missing in the EB-47 complex.

However, both $T_{1\rho}$ and heteronuclear NOE data revealed differences in ^{15}N relaxation in different regions of the HD subdomain. $T_{1\rho}$ data (Figure 5.6) shows that the relaxation times of residues 780-785, which form the highly flexible linker between the HD and ART subdomains, are noticeably lower for all drug complexes than they are for the free catalytic domain, and they are substantially lower in the talazoparib and EB-47 complexes. In the free protein, $T_{1\rho}$ values for these six residues were 56 ms, 96 ms, 76 ms, 65 ms, 64 ms and 31 ms respectively (all values rounded to the nearest millisecond). In the talazoparib complex, the values for these residues were 42 ms, 63 ms, 48 ms, 46 ms, 45 ms and 27 ms respectively, and for the EB-47 complex the respective values were 39 ms, 72 ms, 52 ms, 57 ms, 47 ms, and 29 ms. This suggests that the binding of these compounds to the active site reduces the motion of this highly flexible linker, perhaps through formation of contacts between the inhibitor and the αF helix. Interestingly, no corresponding effect on this linker is observed

in the heteronuclear NOE data in Figure 5.7, suggesting that this effect may occur due to slower chemical exchange between bound and unbound states in these complexes. However, it is important to note that in the crystal structure of the talazoparib complex no substantial structural rearrangement of either subdomains in relation to one another is observed (no data is available for EB-47) and that no overall difference in τ_c values was observed between the free catalytic domain and any complex in Figure 5.10, suggesting that binding of the compounds probably does not greatly affect the overall tumbling motion of the catalytic domain or the inter-dependence of the two subdomains.

In the heteronuclear NOE data (Figure 5.7), the overall intensity ratios of residues in the αE and αF helices (residues 725-740 and 760-775) were reduced with respect to free catalytic domain to varying extents across the four complexes, indicating that the rate of motion in these residues had probably increased. In the veliparib, olaparib and talazoparib complexes, these changes correlate with binding affinity and *in vivo* cytotoxicity, as of the three complexes the increases in motion are most pronounced in the talazoparib complex, followed by olaparib, with only very minor decreases in intensity ratio relative to free catalytic domain in the veliparib complex. In EB-47, the increase in flexibility of residues 750-760 (the C-terminus of the αE-αF linker and N-terminus of αF) 772-775 at the C-terminus of the αF helix was greater relative to all other complexes and free catalytic domain, but motion in the αE helix, while faster than in the free catalytic domain, did not increase to the same extent as in the talazoparib complex. Nonetheless, these heteronuclear NOE results suggest that binding of each of these four compounds to the active site in the ART subdomain may affect the conformation and dynamics of α-helices in the HD subdomain.

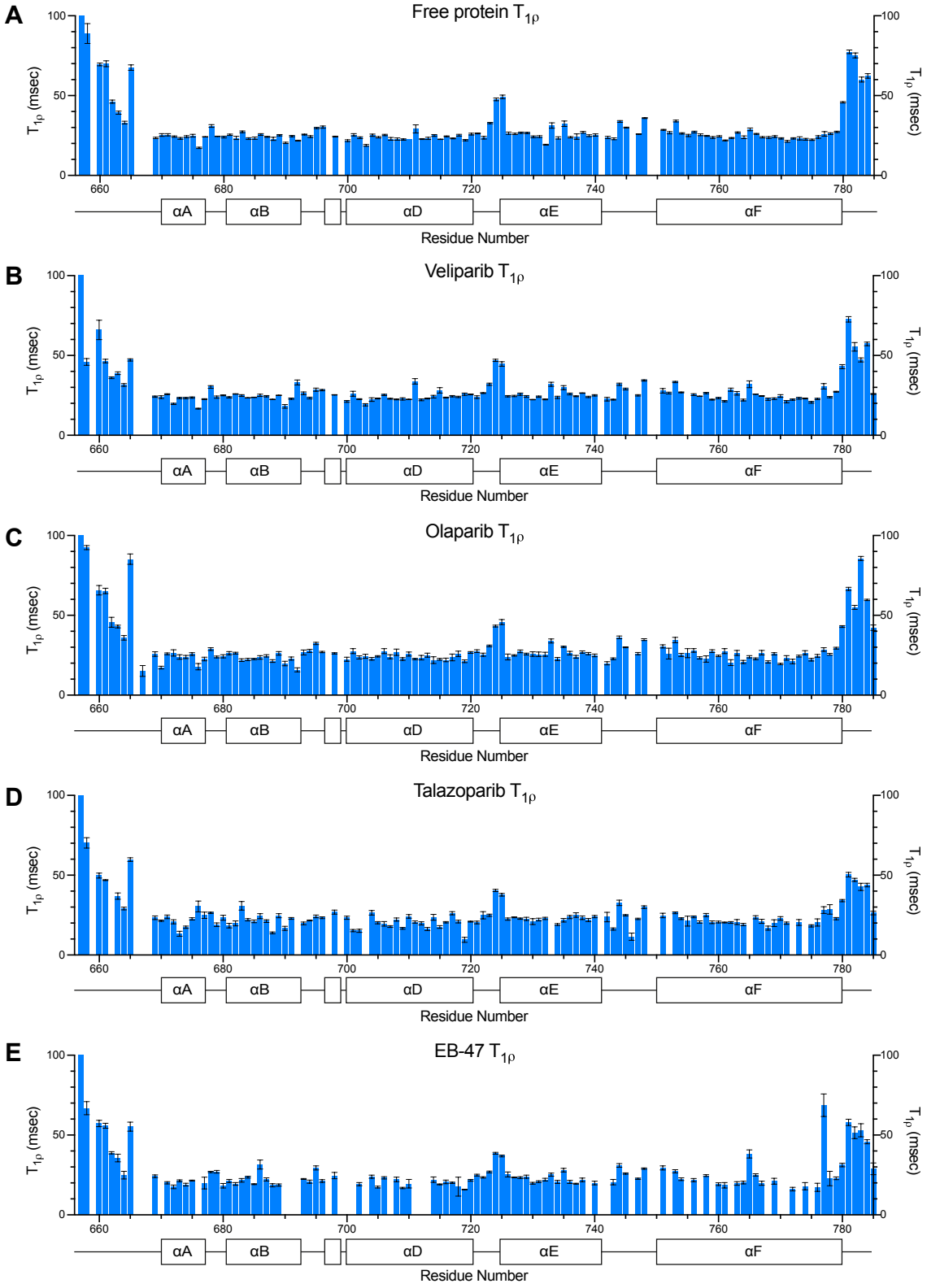


Figure 5.6 **^{15}N $T_{1\rho}$ measurements of HD subdomain residues in free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors.** (A), (B), (C), (D) and (E). ^{15}N $T_{1\rho}$ measurements of HD subdomain residues (656-785) from free catalytic domain (400 µM) and catalytic domain in complex with veliparib (300 µM), olaparib (400 µM), talazoparib (500 µM) and EB-47 (400 µM) respectively. All data was recorded at 25°C from ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C). Error bars were plotted according to the data fitting error in Sparky as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

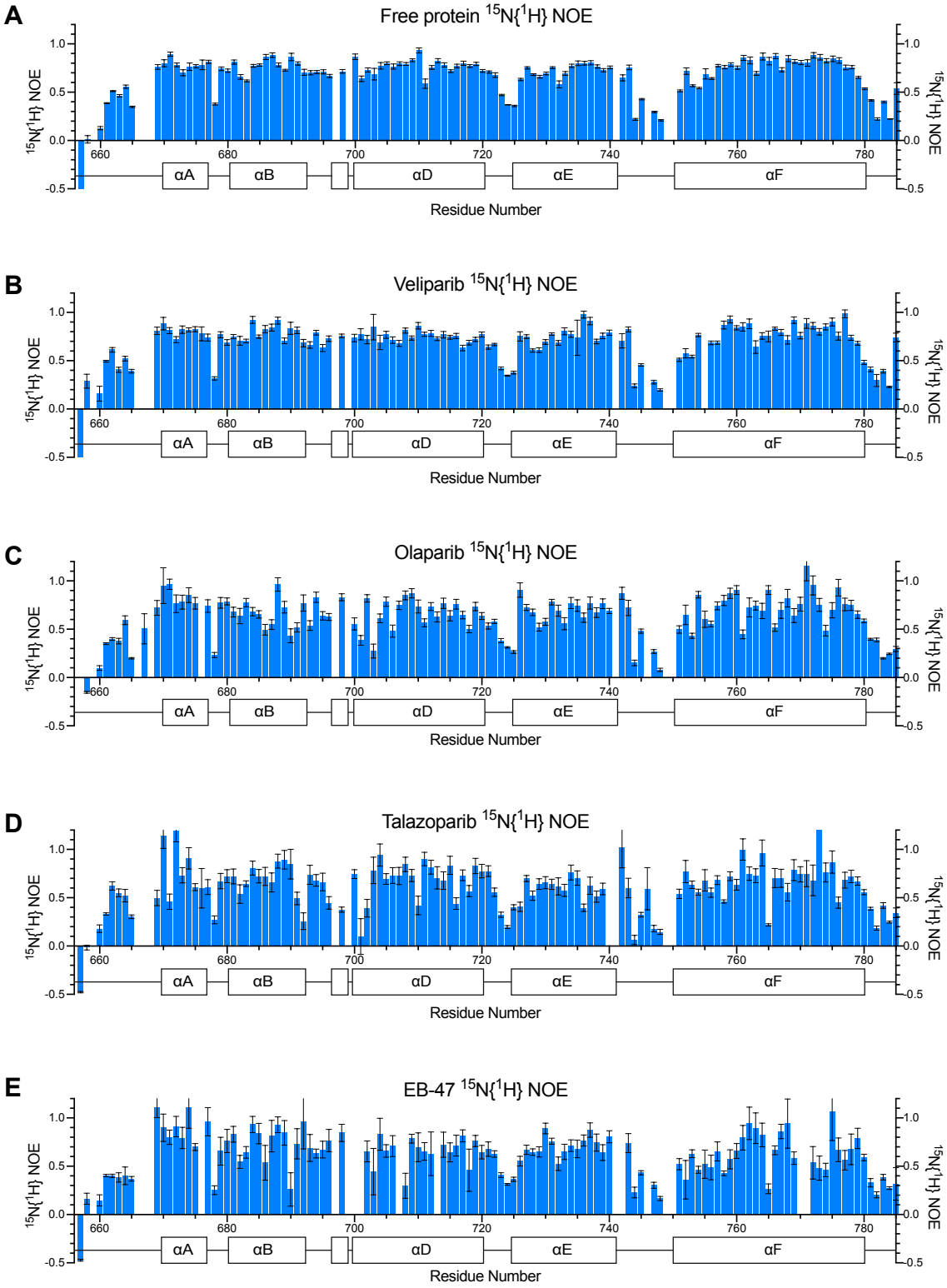


Figure 5.7 $^{15}N\{^{1}H\}$ NOE data from HD subdomain residues in free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. (A), (B), (C), (D) and (E). $^{15}N\{^{1}H\}$ steady-state NOE data measurements of HD subdomain residues (656-785) from free catalytic domain (400 µM) and catalytic domain in complex with veliparib (300 µM), olaparib (400 µM), talazoparib (500 µM) and EB-47 (400 µM) respectively. All data was recorded at 25°C from ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{1}H/^{15}N/^{13}C$). Error bars were plotted according to standard deviation determined on the basis of background noise levels as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

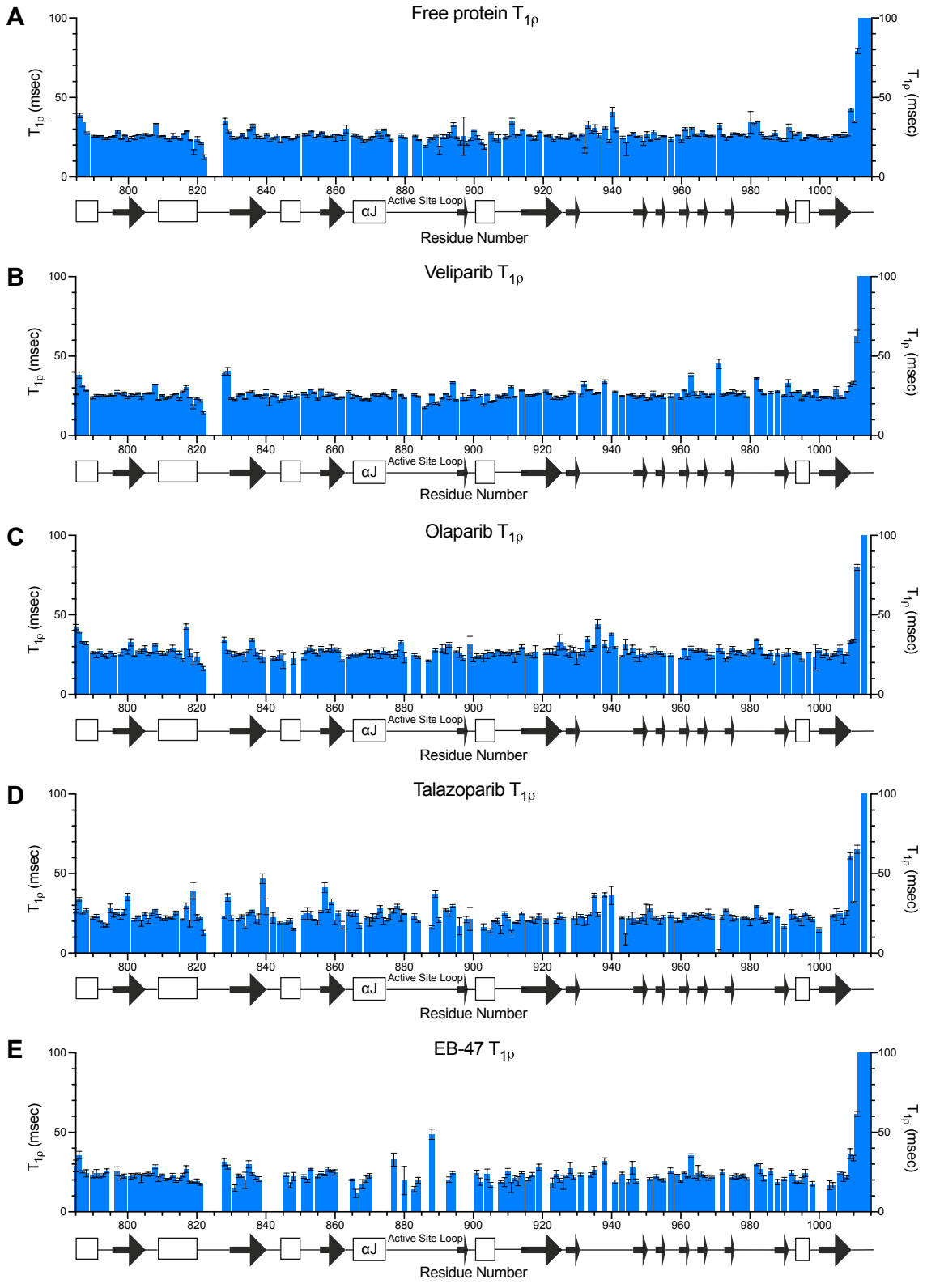


Figure 5.8 ^{15}N $T_{1\rho}$ measurements of ART subdomain residues in free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. (A), (B), (C), (D) and (E). ^{15}N $T_{1\rho}$ measurements of ART subdomain residues (786-1014) from free catalytic domain (400 µM) and catalytic domain in complex with veliparib (300 µM), olaparib (400 µM), talazoparib (500 µM) and EB-47 (400 µM) respectively. All data was recorded at 25°C from ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C). Error bars were plotted according to the data fitting error in Sparky as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

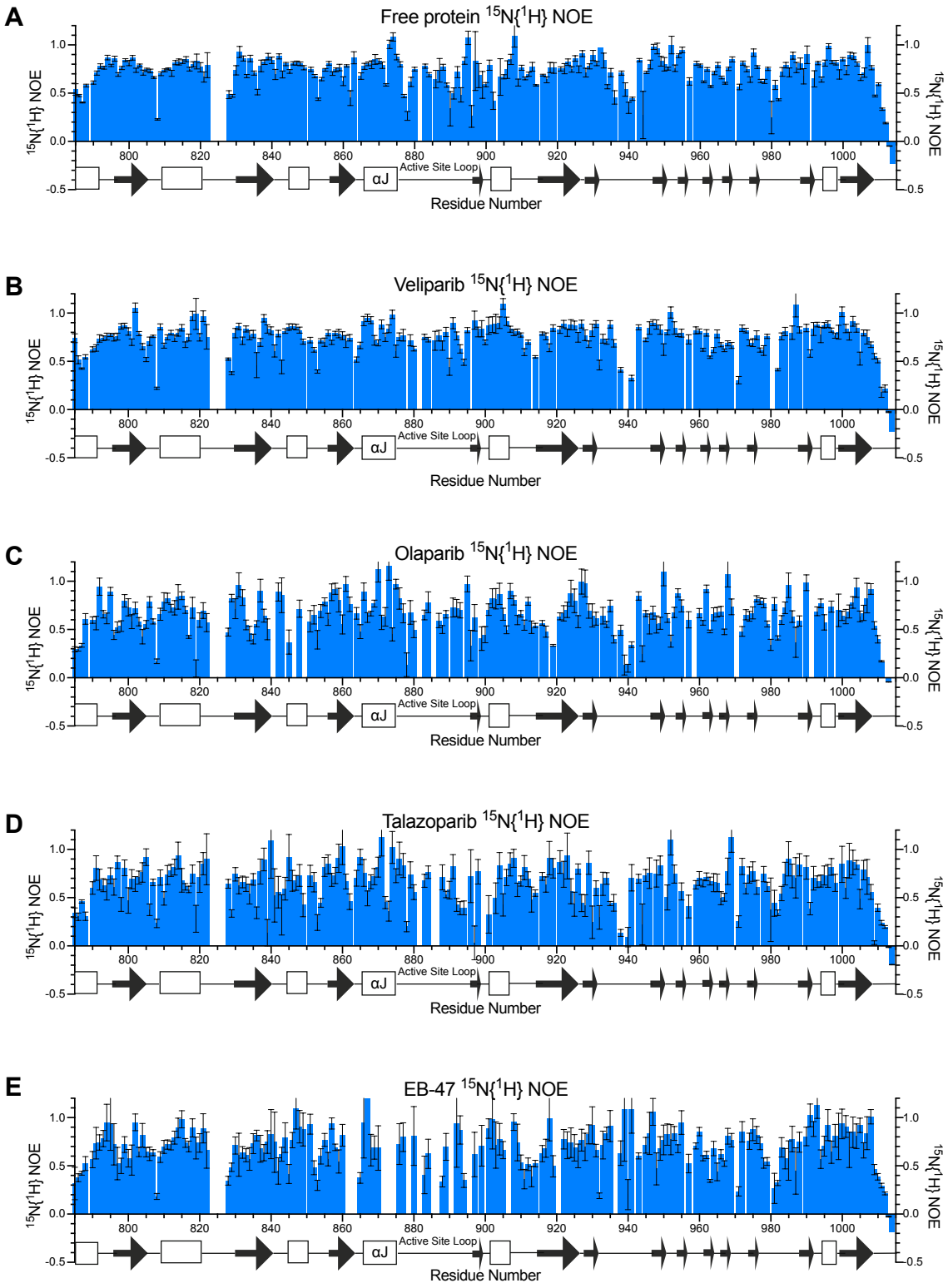


Figure 5.9 $^{15}N\{^{1}H\}$ NOE data from ART subdomain residues in free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. (A), (B), (C), (D) and (E). $^{15}N\{^{1}H\}$ steady-state NOE data measurements of ART subdomain residues (786-1014) from free catalytic domain (400 µM) and catalytic domain in complex with veliparib (300 µM), olaparib (400 µM), talazoparib (500 µM) and EB-47 (400 µM) respectively. All data was recorded at 25°C from ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{1}H/^{15}N/^{13}C$). Error bars were plotted according to standard deviation determined on the basis of background noise levels as described in Section 2.3.4. Secondary structure boundaries are shown below each plot.

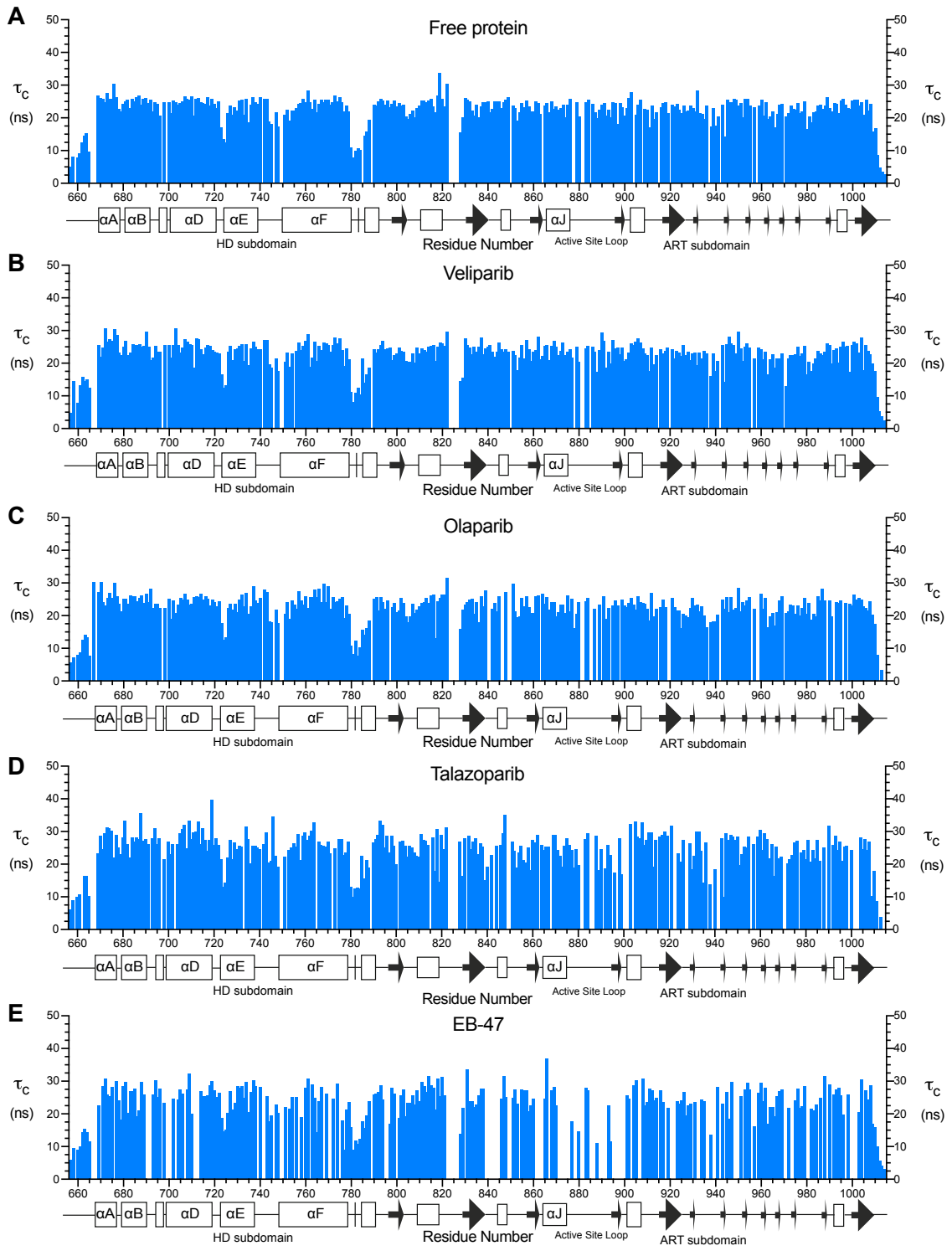


Figure 5.10 τ_c values of free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. (A), (B), (C), (D) and (E). τ_c values calculated from T1 and $T_{1\rho}$ measurements for free catalytic domain (400 µM) and catalytic domain in complex with veliparib (300 µM), olaparib (400 µM), talazoparib (500 µM) and EB-47 (400 µM) respectively. All T1 and $T_{1\rho}$ measurements were recorded at 25°C from ^{15}N labelled samples in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe ($^{1}H/^{15}N/^{13}C$). τ_c values were calculated as described in Section 2.3.4. Subdomain and secondary structure boundaries are shown below each plot.

5.4 EB-47 destabilises α-helical regions of the HD subdomain

Analysis of the solvent exchange data (Figure 5.11) and mapping of the data onto structures of the catalytic domain and inhibitor complexes (Figures 5.12 and 5.13) revealed pronounced differences between the EB-47 complex and all other samples. In the free catalytic domain and the veliparib, olaparib and talazoparib complexes, peaks were observed at later time points in the same regions of the ART subdomain (residues 810-820, 830-840, 860-880 including the αJ helix, 915-925, 945-955, and 985-1005 including the catalytic residue E988) and in αA (670-680), αB (680-690), αD (700-720), αE (725-740) and αF (760-775) helices. This indicated that these residues were protected from D_2O exchange, even over the 48-hour duration of the experiment, by the folds of the ART or HD subdomains. The αF helix was particularly well protected, with the exception of residues 755-760 at the N-terminus of the helix, for which no signals were observed in free catalytic domain or any inhibitor complex. Intense CLEANEX signals were observed at the N and C termini, and at previously identified loop regions such as the inter-subdomain linker at residues 780-785 and 935-945. CLEANEX signals of lower intensity were also observed for the αD-αE linker at residues 723-725 and although the intensity of these signals was lowest in the talazoparib complex, given the small number of residues in the linker and overall similarity of these signals in all five samples it is unclear whether this represents a significant change in the rate of exchange at faster timescales in this linker for the talazoparib complex.

However, despite the similarities observed in the CSP and relaxation data between EB-47 and other compounds (in particular talazoparib), the long-timescale solvent exchange properties of the HD subdomain in the EB-47 complex were completely different from those observed for the free catalytic domain and talazoparib, olaparib and veliparib complexes. As shown in Figure 5.13, many of the signals from backbone amides in the previously highly protected αB, αD and αF helices of the HD subdomain were not observed, indicating that a significant degree of deprotection consistent with a conformational change in these helices had taken place in this complex. This deprotection effect was also observed for the active site residues 860-870 that include the αJ helix, but was less apparent in the αE helix which is the furthest of the HD α-helices from the active site. Changes in solvent exchange of this magnitude suggested that EB-47 binding had significantly destabilised these α-helical regions, and that a corresponding destabilisation did not occur upon binding of the other compounds.

In the absence of a solved crystal structure of EB-47 in complex with the entire wild-type PARP-1 catalytic domain to confirm this hypothesis, we re-examined data from the ^{15}N NOESY-HSQC experiments that were used to complete the resonance assignment of backbone amides in these complexes described in Chapter 3. The backbone amides of residues in ordered α-helical regions would typically be expected to display strong NOE cross-peaks with the amides of sequentially neighbouring residues as a result of their strong dipolar interactions, and we had observed during the assignment process that many of these cross-peaks in the EB-47 complex appeared to be missing in comparison to their counterparts in the other inhibitor complexes. Therefore, we hypothesised that if EB-47 binding caused destabilisation of these α-helical regions, there would be a corresponding loss of cross peaks observed in the HD subdomain in comparison to free catalytic domain and complexes with other inhibitors. Plotting of intensity of all backbone amide cross-peaks in the HD subdomain in free catalytic domain and all inhibitor complexes (Figure 5.14) reveals that almost all backbone amide cross peaks in the main α-helical regions of interest (residues 700-720, including αD, and 755-780, including αF) identified in Figures 5.11, 5.12 and 5.13 are absent in the case of the EB-47 complex. In contrast, there is no significant difference in the number or intensity of cross-peaks (relative to neighbouring peaks) in the regions more distant from the active site containing αA (residues 660-680), αB (680-700) and αE helices (720-740), or linker regions such as residues 740-750, which did not exhibit significant differences in solvent exchange from the other samples. This suggests that, in contrast to the other PARP inhibitors studied as part of this investigation, EB-47 binding significantly destabilises the ordered helical regions of the HD subdomain closest to the active site location.

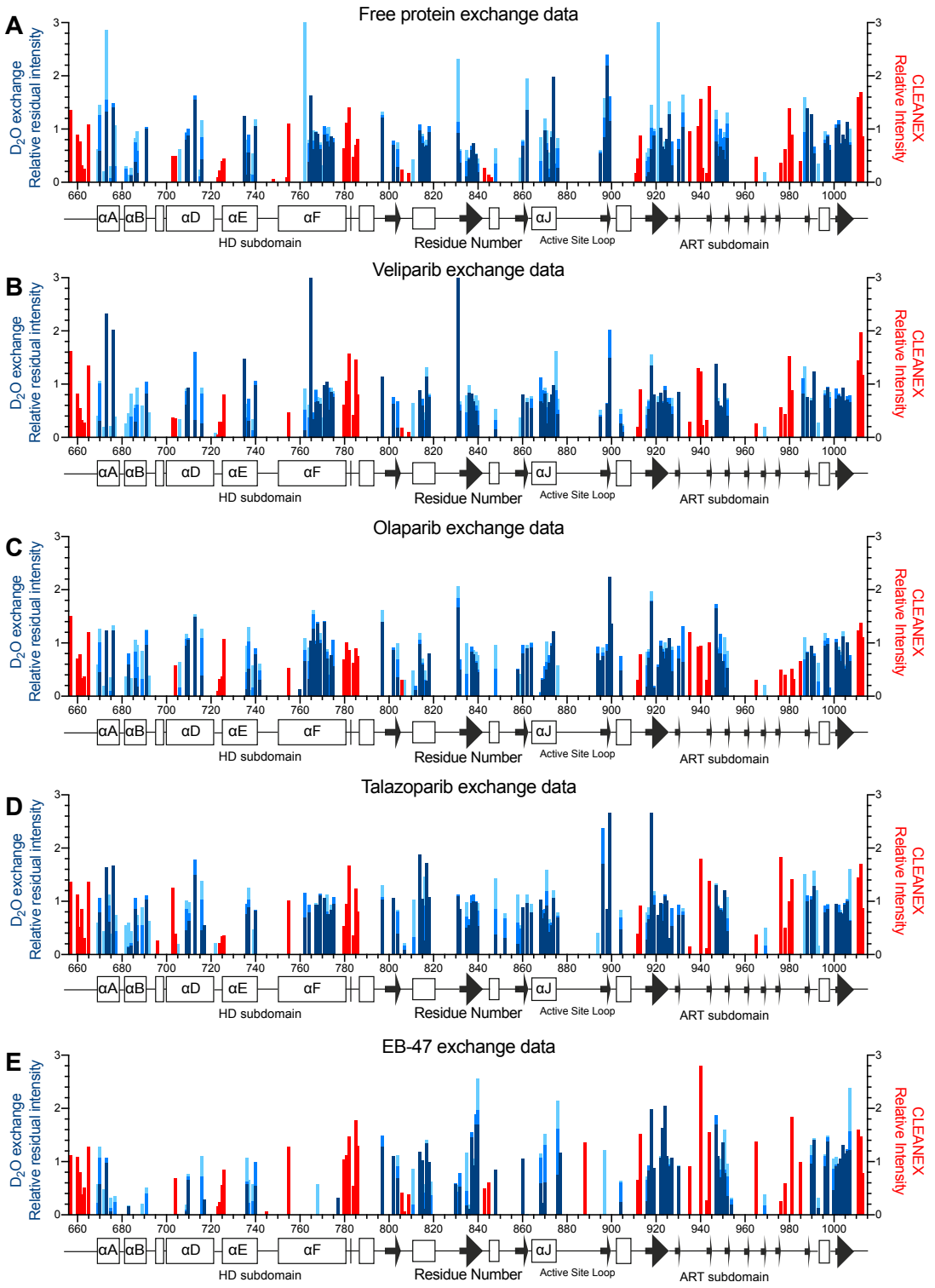


Figure 5.11 D_2O exchange and CLEANEX data for PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. $^{15}N^{1}H$ TROSY experiments were performed on ^{15}N labelled samples (500 µL volume, 100 µM concentration), every 3 hours for 48 hours following lyophilisation and re-suspension in D_2O-based buffer and peak heights were measured and normalised between samples as described in Chapter 2 to obtain relative residual intensity. Relative residual intensity of peaks after 6 hours (light blue), 15 hours (blue), and 42 hours (dark blue) was calculated as described in Section 2.3.6 and superimposed for free catalytic domain (A) and catalytic domain in complex with veliparib (B), olaparib (C), talazoparib (D) and EB-47 (E). CLEANEX data was collected from ^{15}N labelled samples at concentrations described in Figures 5.2 and 5.3 and CLEANEX relative intensity (red) calculated for each sample as described in Section 2.3.5 was plotted alongside relative residual intensity from the D_2O exchange experiment. Subdomain and secondary structure boundaries are shown below each plot.

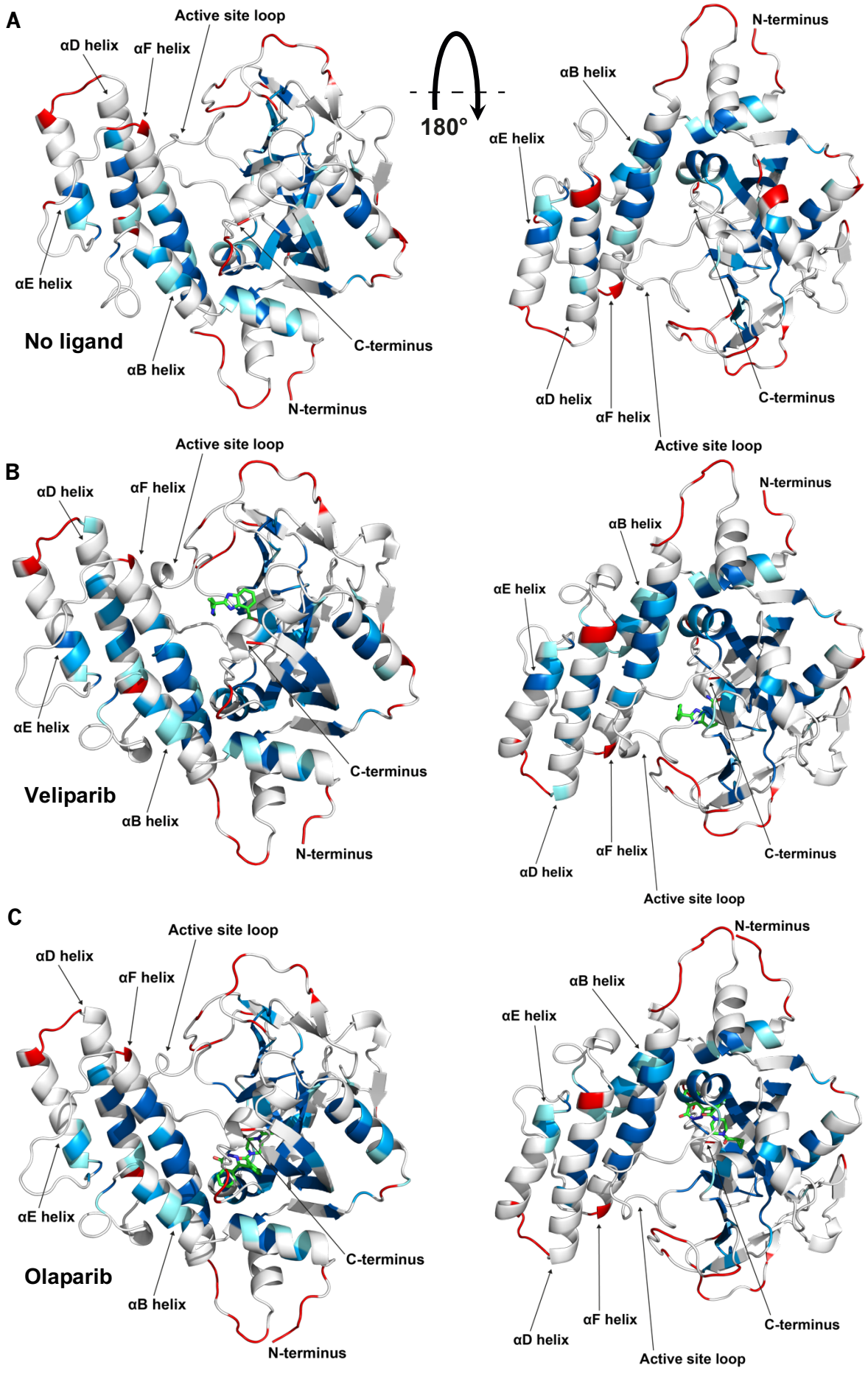

A
Active site loop
αD helix
αF helix
αE helix
No ligand
αB helix
C-terminus
N-terminus
180°
N-terminus
αB helix
αE helix
αD helix
αF helix
C-terminus
Active site loop
B
Active site loop
αD helix
αF helix
αE helix
Veliparib
αB helix
C-terminus
N-terminus
N-terminus
αB helix
αE helix
αD helix
αF helix
C-terminus
Active site loop
C
Active site loop
αD helix
αF helix
αE helix
Olaparib
αB helix
C-terminus
N-terminus
N-terminus
αB helix
αE helix
αD helix
αF helix
C-terminus
Active site loop

Figure 5.12 (previous page) Mapping of solvent exchange rates to the structures of free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with veliparib and olaparib. Residues in the structures of free catalytic domain (A) or catalytic domain in complex with veliparib (B) or olaparib (C) were coloured according to the relative residual exchange data from D_2O exchange experiments and CLEANEX data shown in Figure 5.11. Where data for the fastest-exchanging residues was acquired in CLEANEX experiments, residues were coloured red. Where data was acquired for slowly-exchanging residues in the D_2O exchange experiment, residues were coloured according to the latest time point for which signals were observed. Signals from the 6-hour (light blue), 15-hour (blue) and 42-hour (dark blue) time points shown in Figure 5.11 were mapped. Residues with solvent exchange rates intermediate between CLEANEX and D_2O exchanged time points, for which no signals were observed, were coloured white. Carbon atoms of PARP inhibitor molecules are shown in green, nitrogen atoms in blue and oxygen atoms in red. Mapping was performed as described in Section 2.3.6 using a script written by David Neuhaus.

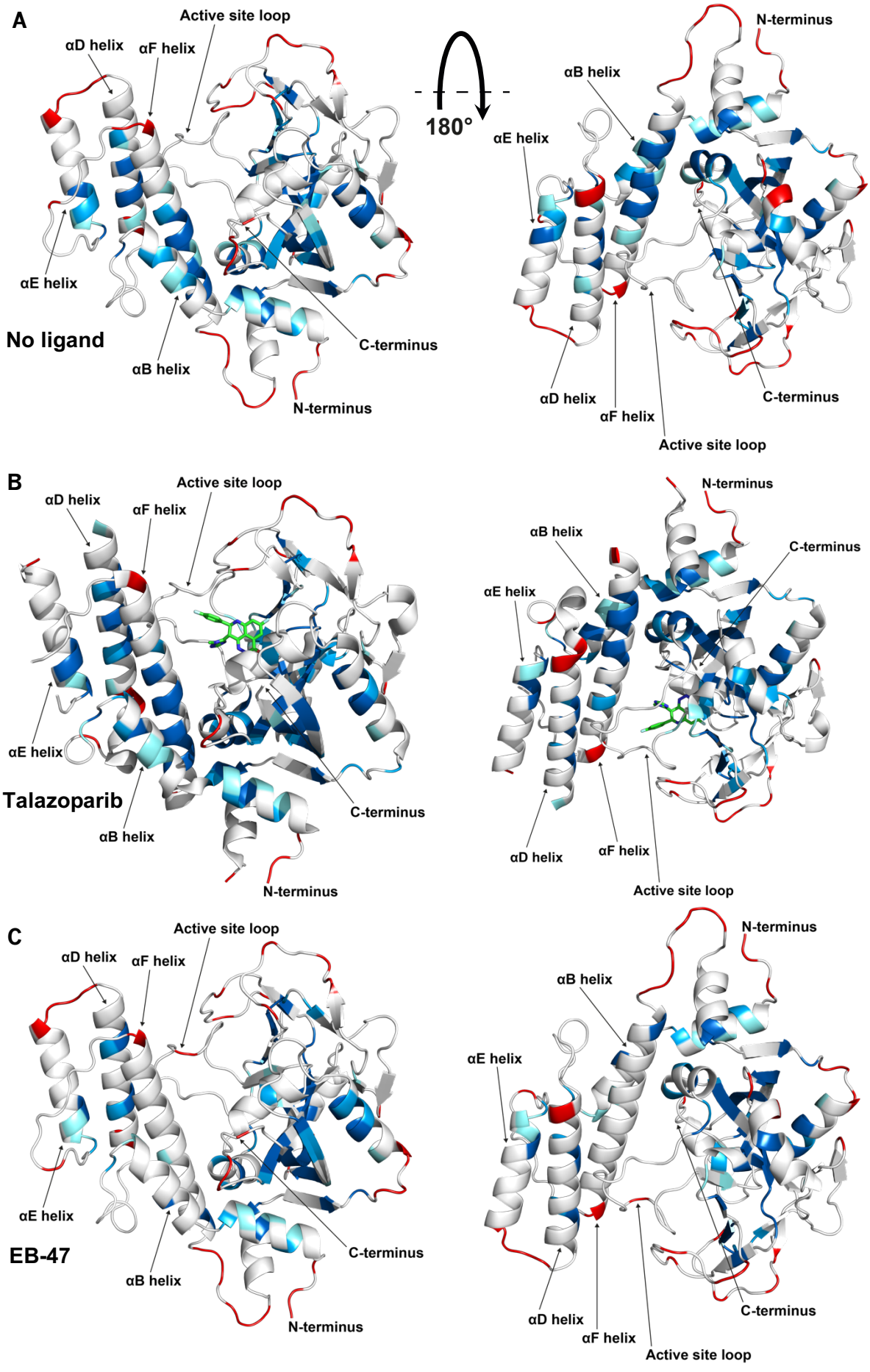

A
Active site loop
αD helix
αF helix
αE helix
No ligand
αB helix
C-terminus
N-terminus
180°
N-terminus
αB helix
αE helix
αD helix
αF helix
C-terminus
Active site loop
B
Active site loop
αD helix
αF helix
αE helix
Talazoparib
αB helix
C-terminus
N-terminus
N-terminus
αB helix
C-terminus
αE helix
αD helix
αF helix
Active site loop
C
Active site loop
αD helix
αF helix
αE helix
EB-47
αB helix
C-terminus
N-terminus
N-terminus
αB helix
αE helix
αD helix
αF helix
C-terminus
Active site loop

Figure 5.13 (previous page) Mapping of solvent exchange rates to the structures of free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with talazoparib and EB-47. Residues in the structures of free catalytic domain (A) or catalytic domain in complex with talazoparib (B) or EB-47 (C) were coloured according to the relative residual exchange data from D_2O exchange experiments and CLEANEX data shown in Figure 5.11. Where data for the fastest-exchanging residues was acquired in CLEANEX experiments, residues were coloured red. Where data was acquired for slowly-exchanging residues in the D_2O exchange experiment, residues were coloured according to the latest time point for which signals were observed. Signals from the 6-hour (light blue), 15-hour (blue) and 42-hour (dark blue) time points shown in Figure 5.11 were mapped. Residues with solvent exchange rates intermediate between CLEANEX and D_2O exchanged time points, for which no signals were observed, were coloured white. Carbon atoms of the talazoparib molecule are shown in green, nitrogen atoms in blue and oxygen atoms in red. Mapping was performed as described in Section 2.3.6 using a script written by David Neuhaus.

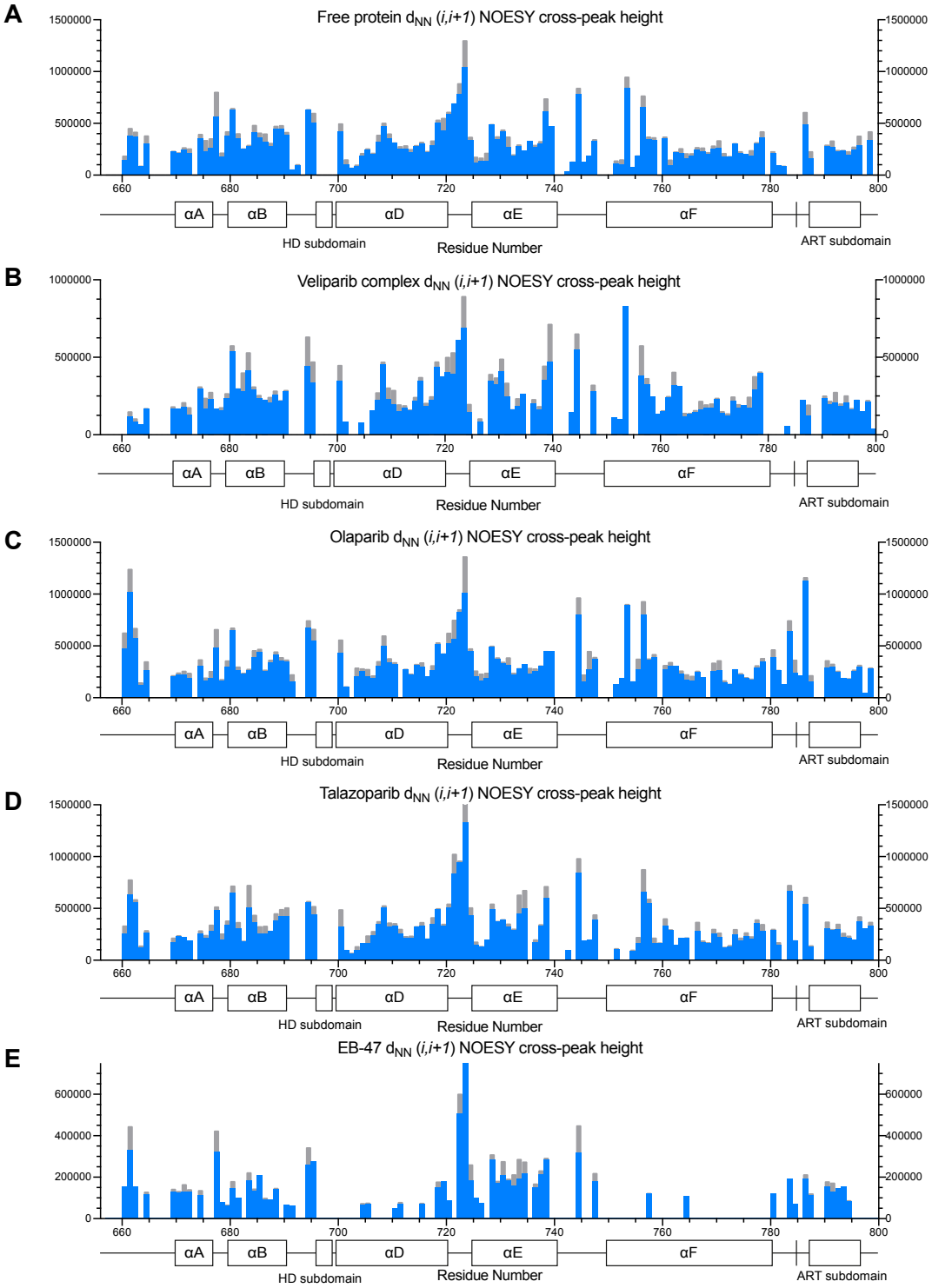


Figure 5.14 Backbone amide-amide cross-peaks from ^{15}N NOESY-HSQC experiments in the HD subdomain of free PARP-1 catalytic domain and PARP-1 catalytic domain in complex with PARP inhibitors. (A), (B), (C), (D), (E). Intensity of backbone amide-amide cross-peaks plotted by residue number for HD subdomain residues (656-800) for free catalytic domain (400 μM), veliparib (300 μM), olaparib (400 μM), talazoparib (500 μM) and EB-47 (400 μM). Bars representing sequential cross-peaks are positioned to span the pair of residues connected by the cross-peak and the average intensity of each pair of sequential cross-peaks is shown in blue, with the intensity of the largest cross-peak in each pair shown in grey. ^{15}N NOESY-HSQC data was collected at 25°C using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C) as described in Chapters 2 and 3. Subdomain and secondary structure boundaries are shown below each plot.

5.5 Discussion

In this chapter, differences were found in the size of chemical shift perturbations in key regions of both the HD and ART subdomains in complexes of PARP-1 catalytic domain fragment in complex with the PARP inhibitors veliparib, olaparib and talazoparib. Given the very high binding affinities of all these inhibitors (see Table 5.1), these differences are clearly not related to the degree of saturation, which was complete in all cases. Despite the similarity of the size and locations of HD subdomain CSPs between the talazoparib and EB-47 catalytic domain complexes, measurement of solvent exchange at long timescales revealed that binding of EB-47 is unique among the four compounds investigated in this chapter, as only EB-47 led to significant and clearly detectable destabilisation of the αF and αD helices in closest proximity to the active site.

For veliparib, olaparib and talazoparib, the similarity of solvent exchange data for these complexes to free wild-type catalytic domain, at least for all regions of the HD subdomain for which signals were observed at longer timescales (which includes the majority of residues from the αA, αB, αD and αF helices, and the five C-terminal residues of the αE helix), indicates that no significant destabilisation of the secondary structure of these α-helices occurred following binding of these compounds. However, differences in the size of CSPs in the αF and other helices of the HD subdomain appear to be linked to the presence or absence of contacts between the compounds and regions on the periphery of the NAD^+ binding pocket. For example, in the case of talazoparib, contacts with the side chains of residues Q759 and E763 from the αF helix have been described, as well as the disruption of hydrogen bonding between Y889 in the active site loop and D766 by talazoparib binding (Aoyagi-Scharber et al., 2014). In the previously unpublished crystal structure of the catalytic domain in complex with olaparib, an oxygen atom also protrudes towards the vicinity of the D766 side chain, but no such potential contacts are observed in the structure of the veliparib complex, which appears to form contacts only with residues from the active site of the ART subdomain. From the results of similar experiments performed on HD subdomain mutants in Chapter 4, it appears that packing of HD subdomain helices in relation to both one another and the active site location is important for autoinhibitory function (and the relief of this function upon binding to DNA damage), and that any perturbation of this packing between helices in the HD subdomain leads to corresponding changes in chemical shifts not only within the helices but also for residues of the active site. As talazoparib forms contacts with two residues from the αF helix and also disrupts a contact between the αF

helix and the loop at the periphery of the active site, these more disruptive changes could explain why talazoparib binding causes CSPs in the other HD subdomain helices bordering αF, whereas olaparib binding causes CSPs in αF only and veliparib binding leads to minimal changes in any chemical shifts of HD backbone amides.

Most importantly, the results of the CSP and solvent exchange data are in agreement with the results of previous biochemical and kinetic studies performed on these compounds *in vitro* (Hopkins et al., 2015; Rudolph et al., 2018), which have suggested that trapping of PARP-1 (or rather, slowed dissociation of inhibitor-bound PARP-1 from DNA) by these compounds *in vitro* occurs only in the presence of NAD^+ and appears to be correlated to binding affinity of the compounds in a linear fashion, as opposed to the non-linear differences (particularly for talazoparib) reported *in vivo* (Murai et al., 2012; Murai et al., 2014). As formation of long, branched, negatively charged poly ADP-ribose chains appears to be important for discharge of PARP-1 from DNA strand breaks, trapping of PARP-1 in the presence of NAD^+ in a manner that correlates closely with binding affinity indicates that inhibition of formation of these chains by PARP inhibitor binding is the mechanism that drives any decrease in off-rate of the DNA. If trapping of these compounds occurred via promotion of an allosteric pathway that mirrors the activation pathway in response to DNA damage, not only would decreased off-rates for PARP inhibitors be observed in *in vitro* assays in the presence of DNA, but pronounced conformational changes in key HD α-helices, with parallels to the partial unfolding observed following activation (Dawicki-McKenna et al., 2015) would also be expected.

However, this means that the results of our *in vitro* studies do not provide a direct mechanistic explanation of the reported highly elevated cytotoxicity of talazoparib *in vivo* with respect to the other two compounds (Murai et al., 2014), which is one order of magnitude higher than would be expected from comparison of PARP-1 binding affinity alone (Table 5.1). These larger than expected differences in potentiation of DNA damage by alkylating agents reported *in vivo* could perhaps occur as a result of off target effects in these compounds. The PARP inhibitors veliparib, olaparib, niraparib, rucaparib and talazoparib have been shown to exhibit varying off-target effects on protein kinases (Antolin and Mestres, 2014; Antolin et al., 2019), with some impacts observed at clinically relevant submicromolar concentrations. Although talazoparib (the compound with the biggest difference between the relative cytotoxicity predicted from the K_i and the effects observed *in vivo*) did not affect the kinases which were inhibited by rucaparib or niraparib, it is not

improbable that off-target effects occur for other classes of proteins and that these may also vary in *in vivo* assays in a cell-line dependent manner.

In the case of EB-47, deprotection in our D_2O exchange experiment was observed at the same locations as in the previously published HXMS data for WT full-length PARP-1 upon DNA binding. 100 seconds after DNA binding, deprotection was observed in the αB helix (between residues 680 and 700) and the C-terminal half of the αF helix (residues 760-775), where the fastest HX was observed after 100 seconds, as well as in the αD helix (residues 700-720) where moderately fast HX was also observed (Dawicki-McKenna et al., 2015). Therefore, there appear to be parallels between EB-47 binding and the changes that occur in the HD subdomain in full length PARP-1 following DNA binding at the N-terminus, which could suggest that EB-47 binding can promote allosteric trapping in full-length PARP-1 by mimicking the changes that occur in the HD subdomain during activation following DNA damage.

The chemical structure of EB-47 also resembles the structure of NAD^+ far more closely than it does the structures of veliparib, olaparib and talazoparib (Figure 5.1). Analysis by the Pascal group of binding by the non-hydrolysable NAD^+ analogue benzamide adenine dinucleotide (BAD) to PARP-1 in the presence of DNA damage reveals that BAD binding to the catalytic domain promoted stronger binding of N-terminal zinc finger domains to DNA (Langelier et al., 2018). In BAD (Figure 5.1), carbon replaces the nitrogen atom of NAD^+'s nicotinamide ring, instead forming a benzamide moiety. Tightening of inter-domain and DNA binding interfaces in response to NAD^+ binding (and a corresponding increase in the time spent bound to the site of DNA damage by activated PARP-1) may promote more effective signalling following detection of single-strand breaks *in vivo*. However, the structural basis of this effect has yet to be fully elucidated, as although the structure of BAD in complex with catalytic domain ΔHD was also solved by the Pascal group, revealing for the first time how the PARP- catalytic domain substrate interacts with the active site, no structure of any NAD^+ analogue has been solved in complex with a PARP-1 construct containing the HD subdomain. Langelier and co-workers also showed that BAD (and presumably NAD^+) does not bind to the unactivated catalytic domain due to the autoinhibitory presence of the HD subdomain. It was also shown recently that $NADP^+$ can function as an endogenous PARP inhibitor that suppresses ADP-ribosylation both *in vitro* and *in vivo* (Bian et al., 2019), confirming that molecules with strong structural similarity to NAD^+ can function as PARP-1 inhibitors, although the ability of $NADP^+$ binding to promote

an allosteric response in the same manner as BAD was not evaluated. Consequently, it would be of interest to evaluate the kinetics of EB-47 binding to PARP-1 bound to a model of DNA damage via a biophysical method such as surface plasmon resonance (SPR). If a decrease in the off-rate of DNA and inhibitor from PARP-1 was observed in the presence of EB-47 but not for veliparib, olaparib and talazoparib, this could provide insights into the relationship between PARP inhibitor structure and the different mechanisms of PARP inhibition.

Given the similarity of EB-47 in both chemical structure and size to BAD and NAD^+ (Figure 5.15) and the fact that BAD binding to unactivated PARP-1 catalytic domain is sterically blocked, there may also be differences in binding of EB-47 to activated catalytic domain in full-length PARP-1 bound to DNA. These differences could potentially be further investigated by NMR spectroscopy using a model of segmentally labelled catalytic domain in full-length PARP-1 discussed further in Chapter 6, although both sample preparation and acquisition of spectra from segmentally labelled domains in full-length PARP-1 is considered to be highly challenging.

Chapter Six

Conclusions and outlook

Detection of single-strand breaks by PARP-1 and its subsequent activation and automodification by ADP-ribosylation are critical processes in the DNA damage response in higher eukaryotic cells. Recently, the underlying structural basis of SSB and DSB recognition by the N-terminal zinc finger domains of PARP-1 has been discovered (Eustermann et al., 2011; Eustermann et al., 2015; Langelier et al, 2012), and much of the mechanistic basis of the subsequent interdomain interactions which lead to activation in the C-terminal catalytic domain has been characterised using NMR spectroscopy (Eustermann et al., 2015). It has also been shown by HXMS that the increase in catalytic domain activity following activation is unexpectedly due to partial loss of structure of α-helical regions in the HD subdomain of the catalytic domain (Dawicki-McKenna et al., 2015) and that in the absence of DNA damage, the HD subdomain sterically blocks NAD^+ substrate binding to the active site in the ART subdomain (Langelier et al., 2018), therefore indicating that conformational change in the HD subdomain improves substrate access to the active site. However, the resolution of HXMS means that so far this process has only been characterised at the level of individual peptides and HXMS timescales (between 10 seconds and 10^5 seconds) and the changes in the HD subdomain that lead to unfolding after binding of the catalytic domain to the binding site formed by the F3 and WGR domains of activated PARP-1 (Figure 1.6) remain to be characterised at higher resolution.

Additionally, PARP-1 activity is an important drug target for treatment of tumours with mutations in *BRCA1* and *BRCA2* genes, and potentially for other genes such as *RAD51* and *FANCA* that are also involved in homologous recombination (McCabe et al., 2006). Cells that are deficient in homologous recombination are reliant on PARP-dependent DNA repair pathways, and inhibition of PARP-1 catalytic domain activity by small molecules is the first anti-tumour treatment to be based on a synthetically lethal approach, in which the combined loss of function of two genes, proteins or pathways results in cell death, but the loss of one of them does not. This effect is not simply the result of abolition of ADP-ribosylation signalling, but rather arises because binding of PARP-1 inhibitors to the C-terminal active site prevents the release of PARP-1 from sites of DNA damage, thereby forming bulky, toxic lesions (Murai et al., 2012). However, it remains unclear whether this

trapping of PARP-1, and the differences in potency between known PARP inhibitors, is the direct result of differences in their binding affinities (and thus their abilities to inhibit formation of the PAR chains needed to discharge PARP-1 from sites of DNA damage) or is due to enhancement of DNA binding via a reverse allosteric mechanism that mirrors the allosteric pathway involved in PARP-1 activation.

By studying the differences between the wild-type catalytic domain fragment and constitutively partially active catalytic domain mutants that are hypothesised to partially mimic the effect of DNA binding, and by comparing the binding of different PARP inhibitors to the catalytic domain fragment using NMR spectroscopy, we aimed both to further characterise the changes that occur in the HD subdomain following PARP-1 activation and determine whether there were any differences between the complexes which could explain the mechanism of PARP inhibition for these compounds.

The discovery that partially activating mutations in key α-helices of the HD subdomain lead to chemical shift perturbations in neighbouring α-helical regions, as well as the active site in the ART subdomain, suggests that the packing of the α-helices is critical to the autoinhibitory function of the HD subdomain and that these HD subdomain mutants may represent an intermediate level between the unactivated and fully activated states of the catalytic domain in wild-type PARP-1.

In the case of all four PARP inhibitors investigated during the course of this work, chemical shift perturbations directly caused by binding were observed at the active site location in the ART subdomain. In the HD subdomain, comparison of the size and location of the observed chemical shift perturbations with crystal structures of the catalytic domain in complex with veliparib, olaparib and talazoparib showed that CSPs can occur as a result of formation of contacts on the boundary of the active site. In our experiments no systematic differences in solvent exchange rates from free catalytic domain were detected for these three complexes, but in contrast the EB-47 complex showed pronounced deprotection from solvent exchange in the key α-helical regions closest to the active site, indicating that EB-47 binding significantly destabilises these regions. This was subsequently confirmed by analysis of cross-peaks from ^{15}N NOESY HSQC experiments, which revealed that in contrast to the other complexes, cross-peaks between backbone amides of sequential residues in these regions were missing for the EB-47 complex alone. On the basis of these results and other in vitro studies that have been published (Hopkins et al., 2015; Rudolph et al., 2018), it appears likely that in the case of veliparib, olaparib and talazoparib, trapping is

due to the inhibition of ADP-ribosylation by active site binding alone. However, the case of EB-47 is more intriguing and the similarities between destabilisation of the HD subdomain and unfolding of the HD subdomain during activation suggest the unconfirmed possibility that EB-47 may promote trapping via a reverse allosteric pathway.

Although by employing CLEANEX and real-time D_2O exchange experiments, it was possible to examine the effects of inhibitor binding on solvent exchange at both fast and very slow timescales, these experiments could not provide data on exchange at intermediate rates, i.e. those residues whose amide NHs exchanged too slowly to be detected by CLEANEX but which exchange before the first TROSY timepoint (6 hours). Therefore, transient conformations of these complexes that exchange more quickly than the very slow timescale of the real-time experiment, or existed as a relatively low proportion of the dominant ground state, would be undetectable using this approach. NMR-based approaches to detect and study transient protein or protein-ligand conformations that exchange with highly populated ground state conformations at varying timescales of chemical exchange have been developed. Techniques such as CEST (chemical exchange saturation transfer) can be applied to spectra of biomolecules by applying a weak radiofrequency field at different frequencies within the chemical shift range of the chosen nucleus (e.g. ^{15}N) with the aim of detecting a loss in intensity of ground state peaks when the irradiation frequency is in close proximity to corresponding peaks of the excited state (Sekhar and Kay, 2013). CEST experiments have been successfully applied to systems with excited state lifetimes of 5 to 50 ms, where the excited conformation represented at least 1% of the population in solution (Vallurupalli et al., 2012). However, in the case of the protein-inhibitor complexes investigated during the course of this project, $T_{1\rho}$ relaxation data covering a similar timescale revealed no significant differences in $T_{1\rho}$ values between samples.

Where the conformational change in a system takes place in a slow exchange regime such that separate peaks are seen for the individual exchanging states (as can occur for slow ligand binding events or changes in secondary structure), a technique such as EXSY (Exchange SpectroscopY, also known as the ZZ-exchange experiment) may be used to quantify the precise timescales and kinetics of these changes. The peak of interest is defined or assigned by its chemical shift at each state and during the evolution period of this experiment, magnetisation is stored longitudinally for a variable length of time (T) while interconversion between the differing states A and B continues to occur, yielding a combination of self-peaks and cross-peaks that vary in strength in a time-dependent manner. By repeating experiments with differing values of T, curves can be generated according to

the intensities of the different species and rates of interconversion between the two states can be calculated. However, the length of the z magnetisation step is limited by the duration of T_1 for a given nucleus and so the exchange between the two states must occur at a rate comparable to or faster than the rate of longitudinal relaxation to be detectable, limiting the useful timescale of this technique (Kleckner and Foster, 2011).

If it could be shown that EB-47 binding promotes delayed release of PARP-1 from DNA damage sites *in vitro* using a biophysical technique such as SPR (discussed in Section 5.5), the changes promoted by EB-47 binding, as well as the changes in HD secondary structure that occur during activation, could be further investigated by NMR spectroscopy by acquiring spectra from segmentally labelled catalytic domain in the context of full-length PARP-1 bound to a DNA damage model. Segmental labelling of individual domains or continuous multi-domain fragments can be carried out using the transpeptidase Sortase A, which can be used to ligate two differentially labelled recombinant proteins to make an NMR sample. This approach has been successfully used in the Neuhaus lab in the past in order to reduce overlap and complexity of spectra (for instance in the PARP-1 F1-F2-F3 fragment; Eustermann et al., 2015). Using this approach, we have so far ligated isotopically labelled catalytic domain to an unlabelled PARP-1 ΔCAT construct to make full-length PARP-1, which presumably has the same DNA binding and catalytic domain activity as non-recombinant PARP-1. During the course of previous work to characterise the structural and mechanistic basis of SSB binding to PARP-1 F1 and F2 domains, DNA dumbbells which act as models of different forms of SSBs were already designed by Sebastian Eustermann of the Neuhaus group, which means that full-length sortase-ligated PARP-1 could be titrated with the DNA dumbbell inside the NMR tube and spectra acquired from fully activated catalytic domain in PARP-1.

This experiment is highly challenging in comparison to the study of the isolated catalytic domain fragment in isolation, as the sample would have a molecular weight of approx. 113 kDa in isolation and approx. 128 kDa in complex with the DNA dumbbell ligand. Additionally, preparation of the ligated sample at the molar concentration required for acquisition of high-quality NMR spectra has been severely limited by the poor yield of the multi-domain N-terminal fragment in *E. coli.*, as well as quite low yield of ligated product from the sortase ligation reaction. Nonetheless, to date we have acquired spectra from a single 40 μM sample of sortase-ligated full-length PARP-1 containing ^{15}N-labelled catalytic domain ligated to unlabelled N-terminal domains (Figure 6.1; sample prepared by Laura Easton), in the absence of a DNA dumbbell ligand. Although some backbone amide

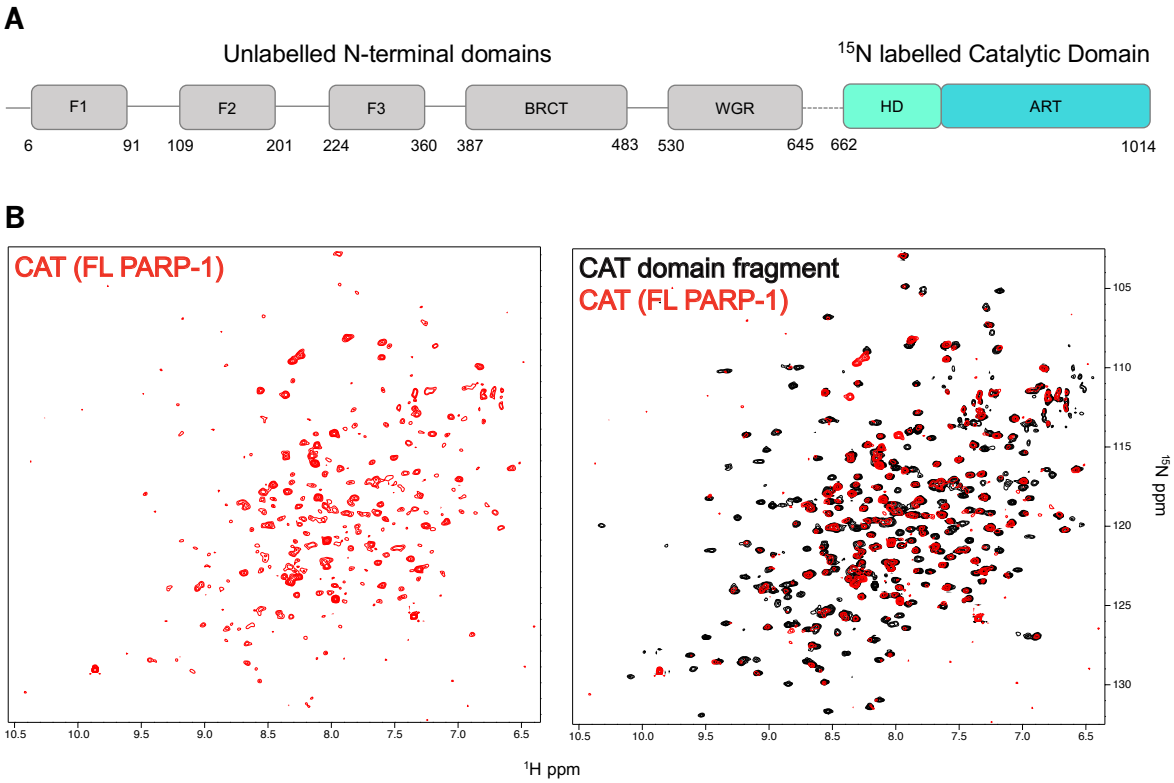


Figure 6.1 Segmental labelling of PARP-1 catalytic domain in full-length PARP-1. (A) Schematic of full-length PARP-1 containing segmentally labelled catalytic domain. Unlabelled domains are shown in grey, the WGR-CAT linker containing the sortase ligation site is represented by a dotted line. (B) [^{15}N ^{1}H] TROSY spectra of ^{15}N segmentally labelled PARP-1 catalytic domain (40 µM, red), alone and superimposed on the spectrum of the isolated catalytic domain fragment (400 µM, black). All spectra were acquired at 25°C in NMR buffer. Data was collected using a Bruker Avance III HD 800 spectrometer equipped with a triple-resonance cryoprobe (^{1}H/^{15}N/^{13}C).

signals are missing in full-length PARP-1 catalytic domain in comparison with the isolated catalytic domain fragment, a fairly high proportion overall are still observed and in almost all cases overlap precisely with the isolated catalytic domain fragment, with no chemical shift perturbations. Because the assignment of the full-length spectrum has not been completed, at present it remains unknown which signals are missing, and it also remains unclear whether their disappearance is due to signal loss from slower tumbling, and consequently faster relaxation caused by more efficient dipole-dipole interactions that would be expected from high molecular weight proteins, or by transient interactions with other PARP-1 domains under an intermediate exchange regime, or by a combination of these factors, or by another explanation. Other sharp peaks have appeared, which are likely (but yet to be confirmed) to be from additional residues in the sortase ligation sequence; these are in the linker that is expected to be relatively mobile. Finding solution conditions under which the complex of full-length PARP-1 with our DNA dumbbell ligand is soluble in sufficiently high concentration to allow NMR experiments is significantly more challenging

than working with the protein fragments studied in the group until now, but work by Laura Easton has found that a sodium phosphate buffer previously used in DNA binding studies of full-length PARP-1 (Langelier et al., 2018) appears to avoid precipitation at concentrations where NMR becomes feasible, paving the way for attempts to acquire spectra of the separately labelled catalytic domain in the context of activated, DNA-bound full-length PARP-1.

Any change in NMR buffer is likely to lead to changes in chemical shifts in the spectra and necessitate transfer of assignments to the new spectra, and it also remains unclear whether signals from backbone amides can be observed in the catalytic domain following activation, as both the extent of the independence of the catalytic domain from the other domains involved in activation and the strength of the interaction between the catalytic domain and its binding site interface (formed from surfaces on F3 and WGR) are unknown. However, acquisition of spectra from activated catalytic domain could shed light on these questions, as well as potentially allowing us to probe the changes observed in the HD subdomain in previously published HXMS experiments, and from the EB-47 complex in our solvent exchange experiments, with a view to establishing whether EB-47 binding promotes allosteric strengthening of CAT-WGR binding in full-length PARP-1. As assignments for PARP-1's 11 kDa WGR domain (which precedes the catalytic domain in the PARP-1 sequence and forms both contacts with DNA and part of the catalytic domain binding site during activation) have previously been completed by other members of the Neuhaus group, samples of full-length PARP-1 could also be prepared containing an isotope-labelled WGR-CAT fragment ligated to an unlabelled N-terminal F1-F2-F3-BRCT fragment. In the absence of DNA, we could expect to detect a high proportion of the backbone amide signals from the WGR domain in the absence of DNA (as already observed for the catalytic domain), as prior to activation all PARP-1 domains are known to tumble independently in a manner consistent with "beads on a string". However, as the WGR domain is known to form interdomain contacts with DNA, F1, F3 and the catalytic domain (Eustermann et al., 2015; see Section 1.1.6) in the DNA complex, we hypothesise that it does not tumble independently from DNA and other PARP-1 domains upon activation and that regardless of the results observed for the catalytic domain, this would be reflected in lower intensity, or complete absence, of signals from the WGR domain in the spectra following titration with a DNA dumbbell ligand. Therefore, repeating any experiments carried out on segmentally labelled catalytic domain in segmentally labelled WGR-CAT may function as an important control experiment, as if catalytic domain signals are observed

post-titration and WGR signals are not, this will confirm that PARP-1 activation is associated with loosening of the interdomain contacts between CAT and the rest of the complex.

Additionally, for high molecular weight systems, isotope labelling of methyl groups from amino acid side chains has often provided insights into the dynamics and function of molecular assemblies up to 1 MDa in size (Sprangers and Kay, 2007), as methyl groups very often have relatively sharp signals due to their rapid internal motions. In the case of the sortase-ligated catalytic domain, methyl groups of methionine, isoleucine, leucine and valine side chains could be labelled in a perdeuterated background (MILV labelling) and 2D [^{13}C-^{1}H] spectra acquired. Given that in this thesis we have identified the importance of inter-helical side chain contacts in leucine residues such as L713F and L765F in the autoinhibitory function of the HD subdomain, signals from methyl groups of side chains could be informative in characterising the changes that occur in these residues upon activation of full-length PARP-1.

The differences in catalytic domain binding observed between the PARP inhibitors investigated in this study have implications for our understanding of PARP-1 inhibition by small molecules. The fact that chemical shift perturbations are observed across multiple helical regions of the HD subdomain upon binding of talazoparib, a compound that makes relatively extensive contacts with side chain residues of the αF helix, and the active site loop that bridges the HD and ART subdomain (Aoyagi-Scharber et al., 2014), but not for olaparib and veliparib that make lesser contacts with the HD, confirms that HD subdomain contacts can also play a role in binding of compounds to PARP-1 catalytic domain.

Exploiting HD subdomain contacts may also present opportunities to improve the selectivity of PARP inhibitors. Given the highly conserved nature of the PARP signature (and therefore the active sites) across the 17 members of the PARP family, compounds that are designed to bind to the PARP-1 active site or ART subdomain alone are highly likely to bind to other members of the PARP family. Due to the diversity of functions across the PARP family, there is significant interest in generating highly selective PARP inhibitors for improved functional *in vivo* studies of individual PARPs, or multiple PARPs that are grouped by function (Steffen et al., 2013). One such study was recently reported for PARP-11, revealing that PARP-11 catalytic domain activity is required for localisation of PARP-11 to the nuclear envelope (Kirby et al., 2018). Of the 17 PARP family members, only PARP-1, PARP-2 and PARP-3 appear to have functions relevant to detection of DNA

damage and only PARP-1, PARP-2, PARP-3 and PARP-16 appear to contain autoregulatory HD subdomains (see Chapter 1, Figure 1.2). Differences also exist between the sequences of these HD subdomains that can have relevance to the binding mechanism of small molecules, as already shown in the case of talazoparib binding to PARP-1 and PARP-2 (Aoyagi-Scharber et al., 2014).

In the case of EB-47, it still remains to be confirmed in follow-up experiments whether the mechanism of inhibition differs in full-length PARP-1 from the other compounds investigated during this work. However, the relative similarity of EB-47 to NAD^+ and recent findings showing that binding of the NAD^+ analogue BAD to PARP-1 catalytic domain promotes reverse allosteric tightening of PARP-DNA binding, and that the NAD^+ derivative $NADP^+$ functions as an endogenous PARP inhibitor *in vivo*, suggest intriguing possibilities for the design of future PARP inhibitors.

In addition to PARP-1, PARP-2 and PARP-3 also show DNA-dependent activation, although to a lesser extent than does PARP-1, and one aim of future work in the Neuhaus group is to attempt to characterise the mechanism of activation in these cases also. In contrast to PARP-1, PARP-2 and PARP-3 each have only a short, disordered DNA binding domain and a WGR domain N-terminal to the catalytic domain, but the catalytic domain of both proteins contains an autoregulatory HD subdomain and an ART subdomain that contains the PARP signature fold (Chapter 1, Figure 1.2) and despite the differences in their domain structure and that of PARP-1, PARP-2 and PARP-3 share PARP-1's allosteric activation mechanism. Unlike for PARP-1, activation of PARP-2 or PARP-3 appears to require that there is a 5' phosphoryl group at a DNA strand break (Langelier et al., 2014), although PARP-3 possesses only mono-ADP-ribosyl transferase (MART) activity (Vyas et al., 2014), whereas PARP-2 (like PARP-1) has poly-ADP-ribosylation activity. However, the ADP-ribosylation activity of wild-type PARP-2 following activation is not as high as for PARP-1, with estimates varying between a two and four-fold greater level of stimulation above basal activity for PARP-1 than for PARP-2 (Langelier et al., 2014). Furthermore, it has been reported that in conjunction with WGR, the catalytic domain of PARP-2 plays a role in localisation of PARP-2 to DNA damage sites (Riccio et al., 2016) Consequently, it would also be of interest to replicate any studies of catalytic domain activation in full-length PARP-1 in PARP-2, in order to compare and contrast the changes that occur during activation of the domains and determine whether this difference in activated catalytic domain activity is the result of differences in HD subdomain behaviour. As PARP-2 has fewer N-terminal domains, it can form fewer interdomain and DNA contacts upon activation, and this may

reduce the amount of Gibbs free energy available for activation upon DNA binding, for example leading to less extensive changes in HD helical regions. Equally, this deficit could be compensated if PARP-2's catalytic domain interacts with DNA, in which case comparison with the events leading to PARP-1 activation would be of interest. Due to its smaller number of domains, the molecular weight of full-length PARP-2 is 62 kDa in size, substantially lower than PARP-1, although this is still considered to be a challenging, high molecular weight protein in NMR terms. If high-quality spectra can be acquired from full-length PARP-2, segmental labelling of WGR and catalytic domains using Sortase A could be employed to reduce spectral congestion in order to aid interpretation of PARP-2 spectra.

In conclusion, the work described in this thesis represents the first study of activation and inhibition in the PARP-1 catalytic domain at the level of individual residues in solution. The ability of solution NMR spectroscopy to both detect and characterise highly flexible regions and its sensitivity to environmental changes of backbone amides has provided new insights into the dynamic changes occurring in regions of the HD subdomain in these processes that could not have been determined using other structural methods. In particular, our findings demonstrate the importance of the relationship between the helical regions of the catalytic domain in PARP-1 autoinhibition, and our detection of differences in HD subdomain behaviour following binding of different PARP inhibitors, including destabilisation of HD helical regions by EB-47, may reflect the different mechanisms by which these compounds promote the formation of trapped PARP-1 lesions.

Bibliography

Altmeyer, M., Messner, S., Hassa, P.O., Fey, M., and Hottiger, M.O. (2009). Molecular mechanism of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. Nucleic Acids Res. *37*, 3723–3738.

Alvarez-Gonzalez, R., and Jacobson, M.K. (1987). Characterization of Polymers of Adenosine-Diphosphate Ribose Generated In vitro and In vivo. Biochemistry *26*, 3218–3224.

Amé, J.-C., Spenlehauer, C., and de Murcia, G. (2004). The PARP superfamily. Bioessays *26*, 882–893.

Antolín, A.A., Ameratunga, M., Banerji, U., Clarke, P., Workman, P., and Al-Lazikani, B. (2019). The off-target kinase landscape of clinical PARP inhibitors. bioRxiv 520023.

Antolín, A.A., and Mestres, J. (2014). Linking off-target kinase pharmacology to the differential cellular effects observed among PARP inhibitors. Oncotarget *5*, 3023–3028.

Aoyagi-Scharber, M., Gardberg, A.S., Yip, B.K., Wang, B., Shen, Y., and Fitzpatrick, P.A. (2014). Structural basis for the inhibition of poly(ADP-ribose) polymerases 1 and 2 by BMN 673, a potent inhibitor derived from dihydropyridophthalazinone. Acta Crystallogr F Struct Biol Commun *70*, 1143–1149.

Ashworth, A., Lord, C.J., and Reis-Filho, J.S. (2011). Genetic Interactions in Cancer Progression and Treatment. Cell *145*, 30–38.

Banasik, M., Komura, H., Shimoyama, M., and Ueda, K. (1992). Specific inhibitors of poly(ADP-ribose) synthetase and mono(ADP-ribosyl)transferase. J. Biol. Chem. *267*, 1569–1575.

Barkauskaite, E., Jankevicius, G., and Ahel, I. (2015). Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation. Mol. Cell *58*, 935–946.

Bian, C., Zhang, C., Luo, T., Vyas, A., Chen, S.H., Liu, C., Kassab, M.A., Yang, Y., Kong, M., and Yu, X. (2019). NADP $^{+}$ is an endogenous PARP inhibitor in DNA damage response and tumor suppression. Nat Commun *10*, 693.

Bonfiglio, J.J., Fontana, P., Zhang, Q., Colby, T., Gibbs-Seymour, I., Atanassov, I., Bartlett, E., Zaja, R., Ahel, I., and Matic, I. (2017). Serine ADP-Ribosylation Depends on HPF1. Mol. Cell *65*, 932–940.

Brown, A.M., Lo, K.S., Guelpa, P., Beaudoin, M., Rioux, J.D., Tardif, J.-C., Phillips, M.S., and Lettre, G. (2010). Optimus Primer: A PCR enrichment primer design program for next-generation sequencing of human exonic regions. BMC Research Notes *3*, 185.

Brutscher, B. (2002). Intraresidue HNCA and COHNCA Experiments for Protein Backbone Resonance Assignment. J. Magn. Reson. *156*, 155–159.

Bryant, H.E., Schultz, N., Thomas, H.D., Parker, K.M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N.J., and Helleday, T. (2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature *434*, 913–917.

Bryant, H.E., Petermann, E., Schultz, N., Jemth, A.-S., Loseva, O., Issaeva, N., Johansson, F., Fernandez, S., McGlynn, P., and Helleday, T. (2009). PARP is activated at stalled forks to mediate Mre11-dependent replication restart and recombination. Embo J. *28*, 2601–2615.

Burzio, L.O., Riquelme, P.T., and Koide, S.S. (1979). ADP-ribosylation of Rat-Liver Nucleosomal Core Histones. J. Biol. Chem. *254*, 3029–3037.

Butterwick, J.A., and MacKinnon, R. (2010). Solution Structure and Phospholipid Interactions of the Isolated Voltage-Sensor Domain from KvAP. J. Mol. Biol. *403*, 591–606.

Caldecott, K.W. (2007). Mammalian single-strand break repair: mechanisms and links with chromatin. DNA Repair (Amst.) *6*, 443–453.

Chae., Y.K. (2000). Recommendations for the Selective Labelling of [^{15}N]-labelled Amino Acids without Using Auxotrophic Strains. J Kor Magn Reson, *4*, 133-139.

Chaitanya, G.V., Steven, A.J., and Babu, P.P. (2010). PARP-1 cleavage fragments: signatures of cell-death proteases in neurodegeneration. Cell Communication and Signaling *8*.

Chambon, P., Weill, J.D., and Mandel, P. (1963). Nicotinamide mononucleotide activation of new DNA-dependent polyadenylic acid synthesizing nuclear enzyme. Biochem. Biophys. Res. Commun. *11*, 39–43.

Chambon, P., Weill, J.D., Doly, J., Strosser, M.T., and Mandel, P. (1966). On the formation of a novel adenylic compound by enzymatic extracts of liver nuclei. Biochem. Biophys. Res. Commun. *25*, 638–643.

Conde, C., Mark, M., Oliver, F.J., de Murcia, G., and Ménissier de Murcia, J. (2001). Loss of poly(ADP-ribose) polymerase-1 causes increased tumour latency in p53-deficient mice. Embo J. *20*, 3535–3543.

D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G.G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. Biochem. J. *342*, 249-268.

Daniels, C.M., Thirawatananond, P., Ong, S.-E., Gabelli, S.B., and Leung, A.K.L. (2015). Nudix hydrolases degrade protein-conjugated ADP-ribose. Sci Rep *5*, 18271.

Dawicki-McKenna, J.M., Langelier, M.-F., DeNizio, J.E., Riccio, A.A., Cao, C.D., Karch, K.R., McCauley, M., Steffen, J.D., Black, B.E., and Pascal, J.M. (2015). PARP-1 Activation Requires Local Unfolding of an Autoinhibitory Domain. Mol. Cell *60*, 755–768.

de Murcia, G., and Ménissier de Murcia, J. (1994). Poly(ADP-ribose) polymerase: a molecular nick-sensor. Trends Biochem. Sci. *19*, 172–176.

de Murcia, J.M., Niedergang, C., Trucco, C., Ricoul, M., Dutrillaux, B., Mark, M., Oliver, F.J., Masson, M., Dierich, A., LeMeur, M., et al. (1997). Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells. Pnas *94*, 7303–7307.

De Vos, M., Schreiber, V., and Dantzer, F. (2012). The diverse roles and clinical relevance of PARPs in DNA damage repair: Current state of the art. Biochem. Pharmacol. *84*, 137–146.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

Eletsky, A., Kienhofer, A., and Pervushin, K. (2001). TROSY NMR with partially deuterated proteins. J. Biomol. NMR *20*, 177–180.

Eustermann, S., Videler, H., Yang, J.-C., Cole, P.T., Gruszka, D., Veprintsev, D., and Neuhaus, D. (2011). The DNA-binding domain of human PARP-1 interacts with DNA single-strand breaks as a monomer through its second zinc finger. J. Mol. Biol. *407*, 149–170.

Eustermann, S., Wu, W.-F., Langelier, M.-F., Yang, J.-C., Easton, L.E., Riccio, A.A., Pascal, J.M., and Neuhaus, D. (2015). Structural Basis of Detection and Signaling of DNA Single-Strand Breaks by Human PARP-1. Mol. Cell *60*, 742–754.

Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N.J., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917–921.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (2002). Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. Biochemistry *33*, 5984–6003.

Fouquerel, E., and Sobol, R.W. (2014). ARTD1 (PARP1) activation and NAD(+) in DNA repair and cell death. DNA Repair (Amst.) *23*, 27–32.

Galluzzi, L., Vitale, I., Abrams, J.M., Alnemri, E.S., Baehrecke, E.H., Blagosklonny, M.V., Dawson, T.M., Dawson, V.L., El-Deiry, W.S., Fulda, S., et al. (2012). Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. Cell Death and Differentiation *19*, 107–120.

Gibson, B.A., and Kraus, W.L. (2012). New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. Nat. Rev. Mol. Cell Biol. *13*, 411–424.

Grzesiek, S., Anglister, J., Ren, H., and Bax, A. (1993). Carbon-13 line narrowing by deuterium decoupling in deuterium/carbon-13/nitrogen-15 enriched proteins. Application to triple resonance 4D J connectivity of sequential amides. J. Am. Chem. Soc. *115*, 4369–4370.

Gupte, R., Liu, Z., and Kraus, W.L. (2017). PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. Genes Dev. *31*, 101–126.

Haikarainen, T., Narwal, M., Joensuu, P., and Lehtiö, L. (2013). Evaluation and Structural Basis for the Inhibition of Tankyrases by PARP Inhibitors. ACS Med. Chem. Lett. *5*, 18–22.

Hayashi, K., Tanaka, M., Shimada, T., Miwa, M., and Sugimura, T. (1983). Size and Shape of Poly(Adp-Ribose) - Examination by Gel-Filtration, Gel-Electrophoresis and Electron-Microscopy. Biochem. Biophys. Res. Commun. *112*, 102–107.

Hoffmann, B., Löhr, F., Laguerre, A., Bernhard, F., and Dötsch, V. (2018). Protein labeling strategies for liquid-state NMR spectroscopy using cell-free synthesis. Prog Nucl Magn Reson Spectrosc *105*, 1–22.

Hopkins, T.A., Shi, Y., Rodriguez, L.E., Solomon, L.R., Donawho, C.K., DiGiammarino, E.L., Panchal, S.C., Wilsbacher, J.L., Gao, W., Olson, A.M., et al. (2015). Mechanistic Dissection of PARP1 Trapping and the Impact on In Vivo Tolerability and Efficacy of PARP Inhibitors. Mol. Cancer Res. *13*, 1465–1477.

Hwang, T.-L., van Zijl, P.C.M., and Mori, S. (1998). Accurate Quantitation of Water-amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. J. Biomol. NMR *11*, 221–226.

Jagtap, P.G., Southan, G.J., Baloglu, E., Ram, S., Mabley, J.G., Marton, A., Salzman, A., and Szabó, C. (2004). The discovery and synthesis of novel adenosine substituted 2,3-dihydro-1H-isoindol-1-ones: potent inhibitors of poly(ADP-ribose) polymerase-1 (PARP-1). Bioorg. Med. Chem. Lett. *14*, 81–85.

Jankevicius, G., Hassler, M., Golia, B., Rybin, V., Zacharias, M., Timinszky, G., and Ladurner, A.G. (2013). A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. Nat. Struct. Mol. Biol. *20*, 508–514.

Karlberg, T., Hammarström, M., Schütz, P., Svensson, L., and Schüler, H. (2010). Crystal structure of the catalytic domain of human PARP2 in complex with PARP inhibitor ABT-888. Biochemistry *49*, 1056–1058.

Kay, L.E., and Gardner, K.H. (1997). Solution NMR spectroscopy beyond 25 kDa. Curr Opin Struct Biol *7*, 722–731.

Kleckner, I.R., and Foster, M.P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics *1814*, 942–968.

Krishnarjuna, B., Jaipuria, G., Thakur, A., D'Silva, P., and Atreya, H.S. (2011). Amino acid selective unlabeling for sequence specific resonance assignments in proteins. J. Biomol. NMR *49*, 39–51.

Krueger, K.M., and Barbieri, J.T. (1995). The Family of Bacterial Adp-Ribosylating Exotoxins. Clin. Microbiol. Rev. *8*, 34–47.

Laing, S., Unger, M., Koch-Nolte, F., and Haag, F. (2011). ADP-ribosylation of arginine. Amino Acids *41*, 257–269.

Lakomek, N.-A., Ying, J., and Bax, A. (2012). Measurement of N-15 relaxation rates in perdeuterated proteins by TROSY-based methods. J. Biomol. NMR *53*, 209–221.

Langelier, M.-F., Planck, J.L., Roy, S., and Pascal, J.M. (2011). Crystal structures of poly(ADP-ribose) polymerase-1 (PARP-1) zinc fingers bound to DNA: structural and functional insights into DNA-dependent PARP-1 activity. J. Biol. Chem. *286*, 10690–10701.

Langelier, M.-F., Planck, J.L., Roy, S., and Pascal, J.M. (2012). Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. Science *336*, 728–732.

Langelier, M.-F., Riccio, A.A., and Pascal, J.M. (2014). PARP-2 and PARP-3 are selectively activated by 5' phosphorylated DNA breaks through an allosteric regulatory mechanism shared with PARP-1. Nucleic Acids Res. *42*, 7762–7775.

Langelier, M.-F., Steffen, J.D., Riccio, A.A., McCauley, M., and Pascal, J.M. (2017). Purification of DNA Damage-Dependent PARPs from E. coli for Structural and Biochemical Analysis. In Poly(ADP-Ribose) Polymerase: Methods and Protocols, A.V. Tulin, ed. (New York, NY: Springer New York), pp. 431–444.

Langelier, M.-F., Zandarashvili, L., Aguiar, P.M., Black, B.E., and Pascal, J.M. (2018). NAD+ analog reveals PARP-1 substrate-blocking mechanism and allosteric communication from catalytic center to DNA-binding domains. Nat Commun *9*, 844.

Leidecker, O., Bonfiglio, J.J., Colby, T., Zhang, Q., Atanassov, I., Zaja, R., Palazzo, L., Stockum, A., Ahel, I., and Matic, I. (2016). Serine is a new target residue for endogenous ADP-ribosylation on histones. Nat. Chem. Biol. *12*, 998–1000.

Lonskaya, I., Potaman, V.N., Shlyakhtenko, L.S., Oussatcheva, E.A., Lyubchenko, Y.L., and Soldatenkov, V.A. (2005). Regulation of poly(ADP-ribose) polymerase-1 by DNA structure-specific binding. J. Biol. Chem. *280*, 17076–17083.

Manning, D.R., Fraser, B.A., Kahn, R.A., and Gilman, A.G. (1984). ADP-ribosylation of transducin by islet-activation protein. Identification of asparagine as the site of ADP-ribosylation. Journal of Biological Chemistry *259*, 749–756.

Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. J. Biomol. NMR *20*, 71–75.

McCabe, N., Turner, N.C., Lord, C.J., Kluzek, K., Białkowska, A., Swift, S., Giavara, S., O'Connor, M.J., Tutt, A.N., Zdzienicka, M.Z., Smith, G.C.M., and Ashworth, A. (2006). Deficiency in the Repair of DNA Damage by Homologous Recombination and Sensitivity to Poly(ADP-Ribose) Polymerase Inhibition. Cancer Res. *66*, 8109–8115.

Miranda, E.A., Dantzer, F., O'Farrell, M., de Murcia, G., and De Murcia, J.M. (1995). Characterisation of a gain-of-function mutant of poly(ADP-ribose) polymerase. Biochem. Biophys. Res. Commun. *212*, 317–325.

Miwa, M., Ishihara, M., Takishima, S., Takasura, N., Maeda, M., Yamaizumi, Z., Sugimura, T., Yokoyama, S., and Miyazawa, T. (1981). The Branching and Linear

Portions of Poly(Adenosine Diphosphate Ribose) Have the Same Alpha(1-]2) Ribose-Ribose Linkage. J. Biol. Chem. *256*, 2916–2921.

Mortusewicz, O., Ame, J.-C., , Schreiber, V., and Leonhardt, H. (2007). Feedback-regulated poly(ADP-ribosyl)ation by PARP-1 is required for rapid response to DNA damage in living cells. Nucleic Acids Res. *35*, 7665–7675.

Murai, J., Huang, S.-Y.N., Das, B.B., Renaud, A., Zhang, Y., Doroshow, J.H., Ji, J., Takeda, S., and Pommier, Y. (2012). Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. Cancer Res. *72*, 5588–5599.

Murai, J., Huang, S.-Y.N., Renaud, A., Zhang, Y., Ji, J., Takeda, S., Morris, J., Teicher, B., Doroshow, J.H., and Pommier, Y. (2014). Stereospecific PARP trapping by BMN 673 and comparison with olaparib and rucaparib. Mol. Cancer Ther. *13*, 433–443.

Oka, S., Kato, J., and Moss, J. (2006). Identification and characterization of a mammalian 39-kDa poly(ADP-ribose) glycohydrolase. J. Biol. Chem. *281*, 705–713.

Orekhov, V.Y., and Jaravine, V.A. (2011). Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. Prog Nucl Magn Reson Spectrosc *59*, 271–292.

Palazzo, L., Leidecker, O., Prokhorova, E., Dauben, H., Matic, I., and Ahel, I. (2018). Serine is the major residue for ADP-ribosylation upon DNA damage. Elife *7*, e34334.

Potaman, V.N., Shlyakhtenko, L.S., Oussatcheva, E.A., Lyubchenko, Y.L., and Soldatenkov, V.A. (2005). Specific binding of poly(ADP-ribose) polymerase-1 to cruciform hairpins. J. Mol. Biol. *348*, 609–615.

Qiu, W., Lam, R., Voytyuk, O., Romanov, V., Gordon, R., Gebremeskel, S., Vodsedalek, J., Thompson, C., Beletskaya, I., Battaile, K.P., et al. (2014). Insights into the binding of PARP inhibitors to the catalytic domain of human tankyrase-2. Acta Crystallogr. D Biol. Crystallogr. *70*, 2740–2753.

Riccio, A.A., Cingolani, G., and Pascal, J.M. (2016). PARP-2 domain requirements for DNA damage-dependent activation and localization to sites of DNA damage. Nucleic Acids Res. *44*, 1691–1702.

Riquelme, P.T., Burzio, L.O., and Koide, S.S. (1979). ADP ribosylation of rat liver lysine-rich histone in vitro. Journal of Biological Chemistry *254*, 3018–3028.

Robson, S.A., Takeuchi, K., Boeszoermenyi, A., Coote, P.W., Dubey, A., Hyberts, S., Wagner, G., and Arthanari, H. (2018). Mixed pyruvate labeling enables backbone resonance assignment of large proteins using a single experiment. Nat Commun *9*, –11.

Rosenthal, F., Feijs, K.L.H., Frugier, E., Bonalli, M., Forst, A.H., Imhof, R., Winkler, H.C., Fischer, D., Caflisch, A., Hassa, P.O., et al. (2013). Macrodomain-containing proteins are new mono-ADP-ribosylhydrolases. Nat. Struct. Mol. Biol. *20*, 502–507.

Rossi, P., Swapna, G.V.T., Huang, Y.J., Aramini, J.M., Anklin, C., Conover, K., Hamilton, K., Xiao, R., Acton, T.B., Ertekin, A., Everett, J.K, Montelione, G.T.. (2010). A microscale protein NMR sample screening pipeline. J. Biomol. NMR *46*, 11–22.

Rudolph, J., Mahadevan, J., Dyer, P., and Luger, K. (2018). Poly(ADP-ribose) polymerase 1 searches DNA via a "monkey bar" mechanism. Elife *7*, 1046.

Ruf, A., de Murcia, J.M., de Murcia, G., and Schulz, G.E. (1996). Structure of the catalytic fragment of poly(AD-ribose) polymerase from chicken. Pnas *93*, 7481–7485.

Ruf, A., Rolli, V., de Murcia, G., and Schulz, G.E. (1998). The mechanism of the elongation and branching reaction of poly(ADP-ribose) polymerase as derived from crystal structures and mutagenesis. J. Mol. Biol. *278*, 57–65.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998). TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. Pnas *95*, 13585–13590.

Satoh, M.S., and Lindahl, T. (1992). Role of Poly(Adp-Ribose) Formation in Dna-Repair. Nature *356*, 356–358.

Schubert, M., Oschkinat, H., and Schmieder, P. (2001). MUSIC, Selective Pulses, and Tuned Delays: Amino Acid Type-Selective 1H–15N Correlations, II. Journal of Magnetic Resonance *148*, 61–72.

Sekhar, A., and Kay, L.E. (2013). NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. Proc. Natl. Acad. Sci. U.S.a. *110*, 12867–12874.

Sharifi, R., Morra, R., Appel, C.D., Tallis, M., Chioza, B., Jankevicius, G., Simpson, M.A., Matic, I., Ozkan, E., Golia, B., et al. (2013). Deficiency of terminal ADP-ribose protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. Embo J. *32*, 1225–1237.

Shen, Y., Rehman, F.L., Feng, Y., Boshuizen, J., Bajrami, I., Elliott, R., Wang, B., Lord, C.J., Post, L.E., and Ashworth, A. (2013). BMN 673, a novel and highly potent PARP1/2 inhibitor for the treatment of human cancers with DNA repair deficiency. Clin. Cancer Res. *19*, 5003–5015.

Sklenář, V., Piotto, M., Leppik, R., and Saudek, V. (1993). Gradient-Tailored Water Suppression for ^{1}H-^{15}N HSQC Experiments Optimized to Retain Full Sensitivity. J Magn Reson, Series A *102*, 241–245.

Slade, D., Dunstan, M.S., Barkauskaite, E., Weston, R., Lafite, P., Dixon, N., Ahel, M., Leys, D., and Ahel, I. (2011). The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. Nature *477*, 616–620.

Sprangers, R., and Kay, L.E. (2007). Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature *445*, 618–622.

Steffen, J.D., Brody, J.R., Armen, R.S., and Pascal, J.M. (2013). Structural Implications for Selective Targeting of PARPs. Front Oncol *3*, 301.

Suzuki, H., Buonamassa, D.T., Farina, B., and Leone, E. (1985). Effects of 2',5'-Oligoadenylates on Poly(Adp-Ribose)Transferase Activity. Italian Journal of Biochemistry *34*, 428–430.

Takeuchi, K., Ng, E., Malia, T.J., and Wagner, G. (2007). 1-^{13}C amino acid selective labeling in a ^{2} H^{15}N background for NMR studies of large proteins. J. Biomol. NMR *38*, 89–98.

Tong, K.I., Yamamoto, M., and Tanaka, T. (2008). A simple method for amino acid selective isotope labeling of recombinant proteins in E. coli. J. Biomol. NMR *42*, 59–67.

Tugarinov, V., Kanelis, V., and Kay, L.E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc *1*, 749–754.

Ulrich., E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D.E., Kent Wenger, R., Yao, Hongyang., H., and Markley, J.L. (2008) BioMagResBank. Nucleic Acids Res, *36*, D402-D408.

Vallurupalli, P., Bouvignies, G., and Kay, L.E. (2012). Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. J. Am. Chem. Soc. *134*, 8148–8161.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696.

Vyas, S., Matic, I., Uchima, L., Rood, J., Zaja, R., Hay, R.T., Ahel, I., and Chang, P. (2014). Family-wide analysis of poly(ADP-ribose) polymerase activity. Nat Commun *5*, 4426.

Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc *73*, 1–16.

Appendix: Binding of 58-bp dsDNA to the *Dictyostelium discoideum* mating-type protein MatA

The work described in this appendix concerns a publication arising from the project of a previous student in the group (Dr. Katy Hedgethorne), for which further experiments were required to respond to a referee's comments on a previously submitted paper after the original student had left. This work resulted in my becoming an author of the following publication:

Hedgethorne, K., Eustermann, S., Yang, J.-C., Ogden, T.E.H., Neuhaus, D., and Bloomfield, G. (2017). Homeodomain-like DNA binding proteins control the haploid-to-diploid transition in Dictyostelium. Science Advances *3*, e1602937.

The slime mould *Dictyostelium discoideum* is a species of social amoeba found in the soil of temperate climates in regions such as North America, Japan and eastern China (Swanson et al., 2019). *D. discoideum* is eukaryotic and unicellular, but as part of its lifecycle, cells aggregate to form a multi-cellular "slug" in which cells differentiate into distinct pre-stalk and pre-spore cells (Raper et al., 1940). The slug is capable of migration to the surface of the soil based on environmental cues and further differentiates upon reaching this point to form a fruiting body, in which the 20% of cells that form the stalk will die in an altruistic bid to ensure the survival of the remaining 80% (Flowers et al., 2010). As a unicellular eukaryotic organism that nonetheless undergoes differentiation, is easy to manipulate genetically and easily grown in the laboratory (Howard et al., 1988), *D. discoideum* has been used as a model organism for many decades and has contributed to our understanding of cellular processes such as cell differentiation and chemotaxis (Willard and Devreotes, 2006). Its susceptibility to infection by human pathogens including *Pseudomonas aeruginosa* and *Legionella pneumophilia* has also been useful for studies of host-pathogen interactions (Pukatzki, et al., 2002; Solomon et al., 2002).

However, *D. discoideum* is also of interest in its own right, due to its ability to undergo differential developmental pathways and multiple asexual and sexual lifecycles with different ploidy levels in response to differing environmental conditions. The social cycle of development described above culminates in the formation of a fruiting body to spread spores necessary for asexual reproduction, and all D. discoideum cells remain haploid throughout. However, if cells respond to starvation cues in dark humid conditions that are not favourable for fruiting body formation, *D. discoideum* cells can enter another

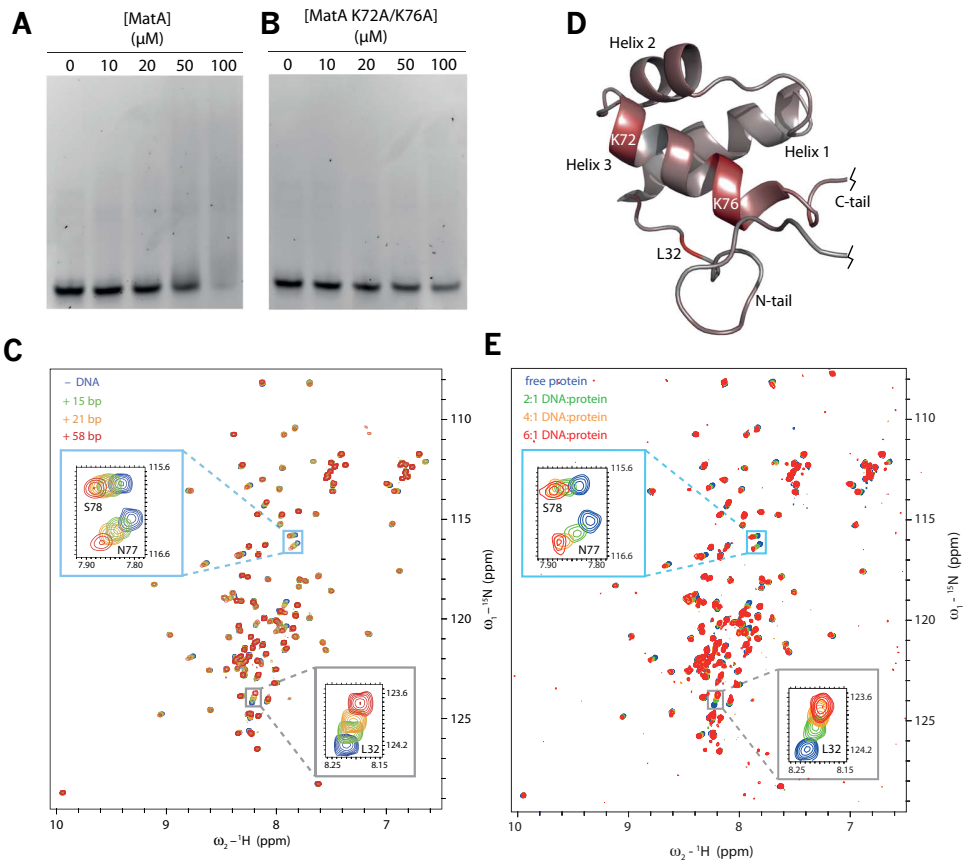
developmental pathway, in which zygotic diploid cells form by mating and (without cell division) cannibalise haploid cells to grow into large dormant cells called macrocysts (Filosa and Dengler, 1972). During dormancy, macrocysts have been reported to undergo meiosis before subsequently releasing haploid progeny, thereby completing the lifecycle (Macinnes and Francis, 1974). Furthermore, *D. discoideum* has three different mating types (type I, type II, and type III) (Bloomfield et al., 2010), with each mating type able to fuse with either of the other mating types in a pairwise manner to form a diploid zygote, but not its own type (Okamoto et al., 2016). Mating type for individual *D. discoideum* cells is determined by a single genetic locus, containing different genes that specify sexual compatibility of the different mating types. Type I cells encode only the gene *matA* at this locus, type III cells encode only the gene *matS* and type II cells encode the genes *matB* and *matC* that are homologues of *matA* and *matC* respectively (Bloomfield et al., 2010). matA and *matB* encode the 107-amino acid proteins MatA and MatB respectively, and *matS* encodes the 209 amino acid protein MatS. None of the genes found at this locus displayed any recognisable sequence homology to other known protein families or domains, and homologues of these genes have only been detected in the genomes of other closely related dictyostelid species (Bloomfield, 2011).

Therefore, the work described by Hedgethorne et al. (2017) prior to my involvement initially focused on determining the solution structures of the MatA and MatB proteins using NMR spectroscopy. In both cases, the small size and hydrophilic natures of these proteins made them good candidates for structural study via this method. In both cases, ^{15}N-^{1}H HSQC, ^{13}C-^{1}H HSQC, 3D data sets for the purposes of resonance assignment and ^{15}N NOESY-HSQC and ^{13}C NOESY HSQC experiments (both with mixing time of 150 ms) were collected and following resonance assignment, structure calculations were performed using restraints from NOESY spectra as described in the published article. In both cases, solution structures revealed that the proteins consisted of a folded core domain composed of approximately 50 amino acids flanked by long flexible tail regions, and that this folded core was arranged in the form of three α-helices to form a homeodomain-like fold. This suggested that MatA controls mating-type function by acting as a transcription factor, binding to DNA in order to regulate gene expression at loci that determine type I mating type identity.

Therefore, the authors next attempted to establish whether MatA had DNA-binding activity *in vitro*, as would be expected for a transcription factor. In order to do this, they performed an electrophoretic mobility shift assay (EMSA), in which increasing amounts of

MatA protein were added to double-stranded DNA (dsDNA) (Appendix Figure 1A). As the specific sequences that MatA binds *in vivo* are yet to be determined, the authors used a 57bp dsDNA oligonucleotide that was previously used in a crystal structure of the *S. cerevisiae* protein Matα2 bound to DNA (Wolberger et al., 1991; pdb 1APL). As the concentration of MatA increased, the DNA band observed in the EMSA became more smeared (Appendix Figure 1A). This indicated that MatA bound to this oligonucleotide, but most likely in a weak, non-specific manner, which could likely be because the oligonucleotide did not correspond to the unknown specific MatA recognition sequence found *in vivo*. After overlaying the MatA structural ensemble with the DNA-bound crystal structure of *S. cerevisiae* Matα2, the authors also hypothesised that MatA lysine residues K72 and K76 could form contacts with DNA. Mutation of these residues was shown to abolish MatA's DNA binding activity (Appendix Figure 1B). This also indicated that, as for other known homeodomain proteins, DNA binding activity appeared to be mediated via residues in the third α-helix of the homeodomain fold.

Work then proceeded to further characterise MatA DNA binding using NMR spectroscopy, via mapping of chemical shift perturbations in backbone amides of ^{15}N-labelled MatA following addition of a 4:1 excess of dsDNA oligonucleotides of varying lengths (Appendix Figure 1C). In this case, sequences of DNA from the *E. coli* expression vector pET30a were used as the specific MatA recognition sequence remained unknown. In this experiment, several peaks exhibited noticeable chemical shift perturbations upon addition of the DNA oligonucleotides, and the CSPs became progressively larger as the length of the dsDNA was increased, suggesting that MatA exhibits higher (non-specific) affinity for DNA as the sequence length increases. Many of the largest CSP values, such as those for residues E69, I75, K76 and N77, were observed for residues that formed part of the C-terminal helix of the MatA structure, further supporting the hypothesis of the authors that it is this helix that mainly interacts with the DNA. However, the observation of large CSPs for residues L32 and L35, located between the N-terminal tail and folded core region, also suggested the unconfirmed possibility that DNA-induced conformational changes may take place in MatA (Appendix Figure 1D). Later in this project, following initial submission of the manuscript, a referee requested follow-up experiments to confirm using NMR spectroscopy that MatA weakly interacted with the dsDNA in a manner consistent with fast exchange on the NMR timescale. Therefore, I prepared fresh ^{15}N-labelled MatA protein and performed and optimised titration experiments with Ji-Chun Yang, in which the protein was



Appendix Figure 1. Studies of MatA binding to DNA using NMR spectroscopy. (A) EMSA experiment showing that wild-type MatA binds weakly to a 57bp dsDNA oligonucleotide. (B) Mutation of residues K72 and K76 to alanine abolishes DNA binding in MatA. (C) Addition of dsDNA (80 μM) to ^{15}N labelled MatA (20 μM) causes chemical shift perturbations. (D) Chemical shift perturbations from MatA complex with 29bp dsDNA mapped onto the lowest energy structure of the core MatA folded region, showing that the largest chemical shift perturbations occur in helix 3 of the homeodomain fold. (E) Titration of MatA (20 μM) with increasing amounts of a 58bp dsDNA oligo reveals that MatA DNA binding is in fast exchange on the NMR timescale. NMR experiments were performed in 25 mM sodium phosphate (pH 6.0), 50 mM NaCl and 50 μM EDTA. Reprinted from Hedgethorne et al., 2017 with permission from AAAS.

titrated with an increasing ratio of the longest 58bp nucleotide and ^{15}N^{1}H HSQC spectra were recorded (Appendix Figure 1E). This experiment showed that the system was indeed in fast exchange, as increasing the concentration of the DNA increased the size of the CSPs smoothly and continuously as more DNA ligand was added; this increase of the size of the CSPs with increasing DNA concentration paralleled that seen as a function of increasing DNA oligo length in the earlier experiments (Appendix Figure 1C). Under conditions of fast exchange, the location of each peak represents an average chemical shift weighted by population. Therefore, as the concentration of the ligand increases, the weighting of the populations also increases in favour of the complex and the chemical shift perturbation increases in size. As the interaction between MatA and the DNA was so weak in this experiment, the protein remained unsaturated even at the highest ratio of DNA to protein

(6:1) tested in this experiment. Following re-submission of the manuscript including these observations, the manuscript was subsequently accepted for publication.

Bibliography

Bloomfield, G., Skelton, J., Ivens, A., Tanaka, Y., and Kay, R.R. (2010). Sex determination in the social amoeba Dictyostelium discoideum. Science *330*, 1533–1536.

Bloomfield, G. (2011). Genetics of sex determination in the social amoebae. Dev. Growth Differ. *53*, 608–616.

Filosa, M.F. and Dengler, R.E. (1972). Ultrastructure of macrocyst formation in the cellular slime mold, Dictyostelium mucoroides: extensive phagocytosis of amoebae by a specialized cell. Dev Biol *29*, 1–16.

Howard, P.K., Ahern, K.G. and Firtel, R.A. (1988). Establishment of a transient expression system for Dictyostelium discoideum. Nucleic Acids Res *16*, 2613–2623.

M. Okamoto, L. Yamada, Y. Fujisaki, G. Bloomfield, K. Yoshida, H. Kuwayama, H. Sawada, T. Mori, H. Urushihara (2016). Two HAP2-GCS1 homologs responsible for gamete interactions in the cellular slime mold with multiple mating types: Implication for common mechanisms of sexual reproduction shared by plants and protozoa and for male-female differentiation. Dev. Biol. *415*, 6–13.

Macinnes, M. A., & Francis, D. (1974). Meiosis in Dictyostelium mucoroides. Nature *251*, 321–324.

Pukatzki, S., Kessin, R.H. and Mekalanos, J.J. (2002). The human pathogen Pseudomonas aeruginosa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum. Proc Natl Acad Sci USA *99*, 3159–3164.

Raper, K.B. (1940). Pseudoplasmodium formation and organization in Dictyostelium discoideum. J Elisha Mitchell Sci Soc *56*, 241–282.

Solomon, J.M., Rupper, A., Cardelli, J.A. and Isberg, R.R. (2000). Intracellular growth of Legionella pneumophila in Dictyostelium discoideum, a system for genetic analysis of host-pathogen interactions. Infect Immun *68*, 2939–2947.

Swanson, A.R., Vadell, E.M. and Cavender, J.C. (1999). Global distribution of forest soil dictyostelids. J Biogeogr *26*, 133–148.

Willard, S.S. and Devreotes, P.N. (2006). Signaling pathways mediating chemotaxis in the social amoeba, Dictyostelium discoideum. Eur J Cell Biol *85*, 897–904.

Wolberger, C., Vershon A.K., Liu, B., Johnson, A.D., and Pabo C.O. (1991) 'Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions', Cell *67*, 517–528.

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C^{2}H$				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
G655	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T656	-	-	61.68	69.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
V657	8.16	123.57	61.58	31.93	8.16	123.61	8.16	123.74	8.17	123.76	8.17	123.75	8.16	123.66	8.16	123.69	8.15	123.54	8.15	123.59
N658	8.51	125.01	54.73	-	8.51	124.95	8.52	125.22	8.52	125.13	8.53	125.27	8.51	125.07	8.52	125.20	8.51	125.1	8.5	125.06
P659	-	-	-	32.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G660	8.83	111.11	45.02	-	8.79	110.97	8.8	111.09	8.82	111.17	8.81	111.13	8.86	111.26	8.83	111.21	8.86	111.25	8.85	111.27
T661	7.85	114.01	61.7	69.66	7.82	113.89	7.83	113.97	7.83	113.92	7.83	113.95	7.82	113.85	7.83	113.89	7.82	113.82	7.81	113.84
K662	8.34	122.71	55.11	33.09	8.30	122.83	8.3	122.96	8.31	122.98	8.31	122.94	8.29	122.92	8.31	123.04	8.28	122.89	8.31	122.94
S663	8.43	117.59	57.82	62.96	8.46	117.35	8.5	117.76	8.52	117.77	8.51	117.76	8.50	117.71	8.49	117.75	8.45	117.7	8.53	117.77
K664	9.43	128.53	55.22	31.37	9.40	128.39	9.43	128.65	9.41	128.57	9.42	128.59	9.41	128.53	9.40	128.64	9.47	128.79	9.43	128.57
L665	8.33	123.65	54.15	40.72	8.32	123.58	8.32	123.57	8.31	123.65	8.32	123.59	8.31	123.64	8.32	123.68	8.3	123.57	8.3	123.45
P666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
K667	9.29	127.56		-	8.03	118.40	8.05	118.71	9.28	127.5	-	-	9.29	127.58	9.29	127.60	-	-	-	-
P668	-	-	65.82	30.49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
V669	7.02	115.35	65.23	30.41	7.77	122.42	6.98	115.3	6.99	115.35	6.98	115.33	7.78	122.59	7.78	122.67	7.88	123.66	6.98	115.3
Q670	7.91	120.07	59.94	27.19	7.87	119.94	7.89	120.18	7.88	120.2	7.88	120.19	7.88	120.22	7.88	120.19	7.88	120.29	7.87	120.2
D671	8.31,7	117.25	-	39.23	8.31	117.17	8.32	117.4	8.33	117.42	8.32	117.42	8.33	117.40	8.32	117.42	8.36	117.12	8.33	117.4
L672	7.34	123.32	57.61	40.43	7.31	123.33	7.34	123.7	7.33	123.56	7.33	123.6	7.32	123.53	7.32	123.42	7.37	123.82	7.35	123.71
I673	8.08	119.7	62.43	-	8.04	119.39	8.05	119.86	8.05	119.69	8.07	119.71	8.06	119.67	8.04	119.58	8.09	119.7	8.01	119.79
K674	8.1	118.77	59.3	31.38	8.07	118.77	8.05	118.78	8.08	118.83	8.08	118.8	8.06	118.91	8.04	118.71	8.09	118.65	8.03	118.49
M675	7.45	116.68	58.35	32.97	7.41	116.75	7.37	116.95	7.4	116.82	7.4	116.83	7.41	116.79	7.43	116.70	7.38	116.83	7.37	116.84
I676	7.94	114.93	63.8	36.44	7.90	114.98	7.95	116.53	7.93	114.92	7.92	115.64	7.93	115.37	7.92	114.26	7.98	115.53	-	-
F677	7.35	118.13	57.85	37.22	7.30	118.10	7.35	117.47	7.35	118.32	7.31	117.82	7.32	117.85	7.31	118.53	7.35	117.76	7.36	117.76
D678	6.74	119.34	54.46	40.88	6.70	119.19	6.69	119.33	6.71	119.27	6.71	119.31	6.71	119.25	6.69	119.21	6.69	119.17	6.71	119.2
V679	8.56	129.76	64.79	31.19	8.56	129.84	8.57	129.64	8.57	130.06	8.56	129.79	-	-	-	-	-	-	8.55	129.91
E680	8.28	121.59	58.76	27.2	8.26	121.42	8.29	121.56	8.27	121.77	8.28	121.54	8.28	121.46	8.27	121.63	8.28	121.56	8.35	121.5
S681	7.88	117.22	61.34	62.02	7.85	117.21	7.79	117.47	7.82	117.31	7.81	117.44	7.82	117.37	7.85	117.30	7.77	117.61	7.79	117.55
M682	7.41	122.81	59.88	32.14	7.36	122.75	7.36	122.91	7.36	122.98	7.37	122.99	7.36	122.90	7.35	123.03	7.37	123.03	7.35	122.82
K683	7.83	118.46	55.79	28.35	7.84	118.56	7.82	118.47	7.85	118.44	7.82	118.35	7.81	118.44	7.79	118.53	7.85	118.36	7.82	118.17
K684	8.05	118.89	59.3	31.44	8.01	118.69	8.05	118.78	8.05	118.84	8.04	118.84	8.04	118.79	8.00	118.97	8.13	119.05	8.1	119.02
A685	7.53	121.42	54.43	16.85	7.49	121.33	7.45	121.4	7.41	121.41	7.47	121.47	7.48	121.37	7.51	121.66	7.43	121.47	7.36	121.15
M686	7.33	115.01	59.64	32.28	7.29	114.94	7.28	115.24	7.23	114.99	7.29	115.1	7.29	115.10	7.30	115.13	7.26	115.25	7.28	115.05
V687	8.21	120.47	65.71	30.67	8.19	120.38	8.2	120.45	8.13	120.21	8.2	120.53	8.21	120.53	8.19	120.51	8.28	120.66	8.24	120.72
E688	8.02	124.22	58.88	28.09	7.99	124.12	8	124.28	8.06	124.43	8	124.26	8.02	124.21	7.99	124.41	8.06	123.79	8.02	123.85
Y689	7.31	118.04	58.06	37.74	7.30	118.08	7.29	118.19	7.3	118.32	7.29	118.29	7.29	118.05	7.27	118.15	7.31	118.36	7.28	118.2
E690	7.87	111.12	57.14	25.43	7.83	110.95	7.82	110.84	7.81	111.05	7.81	110.63	7.85	111.10	7.83	110.87	7.8	110.67	7.87	111.25
I691	7.35	119.09	60.3	36.09	7.32	118.93	7.32	119.08	7.31	118.99	7.3	119.13	7.33	119.01	7.32	119.22	7.31	118.82	7.25	118.64
D692	7.89	124.74	51.88	39.77	7.86	124.60	7.9	124.86	7.9	124.72	7.86	124.85	7.86	124.80	7.87	124.89	7.88	124.6	7.87	124.72
L693	8.04	127.97	55.61	40.94	8.03	127.84	8.04	128.08	8.04	128.28	8.05	128.14	8.05	128.16	8.04	128.01	8.05	128.3	8.06	128.24
Q694	8.08	116.87	57.37	27.36	8.04	116.75	8.05	116.82	8.05	116.66	8.03	116.85	8.05	116.86	8.05	116.88	8.05	116.73	8.06	116.76

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C^{2}H$				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
K695	6.99	118.22	56.75	33.59	6.94	118.10	6.94	118.2	6.92	117.99	6.94	118.15	6.95	118.20	6.92	118.26	6.95	118.19	6.93	118.1
M696	8.31	121.91	51.98	32.36	8.30	121.79	8.27	121.99	8.33	122.46	8.29	121.97	8.30	121.80	8.28	121.87	8.29	121.98	8.31	121.76
P697	-	-	61.88	31.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L698	8.15	121.2	57.62	40.24	8.14	121.10	8.13	121.45	8.17	121.28	8.16	121.31	8.14	121.26	8.17	121.39	8.07	121.1	8.13	121.26
G699	-	-	45.44	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
K700	7.8	119.41	54.72	32.26	7.76	119.26	7.72	119.22	7.79	119.67	7.76	119.36	7.76	119.29	7.76	119.41	7.75	119.25	7.77	119.35
L701	7.07	122.73	55.49	41.7	7.03	122.62	7.18	123.21	-	-	7.03	122.59	7.04	122.70	7.02	122.69	7.16	122.56	7.19	122.97
S702	8.02	119.14	55.52	65.85	7.97	119.11	7.97	119.32	-	-	7.99	119.07	7.97	119.14	7.98	119.13	8.02	118.4	7.9	118.64
K703	8.86	125.22	59.99	31.19	8.83	125.12	8.79	125.79	8.67	125.02	8.82	125.13	8.84	125.30	8.85	125.06	8.91	126.45	8.95	126.94
R704	8.21	118.27	58.58	28.65	8.20	118.14	8.19	117.89	8.2	118.38	8.25	118.29	8.23	118.27	8.23	118.42	8.19	118.16	8.24	118.34
Q705	7.47	121.34	58.45	27.84	7.44	121.31	7.48	121.3	7.54	121.21	7.47	121.5	7.46	121.52	7.46	121.62	7.44	121.14	7.46	121.56
I706	7.55	120.04	65.17	37.35	7.53	119.97	7.6	119.59	7.58	119.96	7.55	119.92	7.55	120.15	7.56	120.15	7.59	119.83	7.65	119.58
Q707	8.48	118.71	59.79	28.16	8.43	118.59	8.48	118.32	8.26	118.55	8.41	118.96	8.45	118.92	8.44	118.72	8.41	118.72	-	-
A708	7.86	121.43	54.3	17.21	7.83	121.34	7.87	122.72	7.74	121.72	7.8	121.81	7.86	121.63	7.85	121.45	7.75	121.7	7.87	121.57
A709	8.09	124.23	55.01	17.34	8.08	124.23	7.93	123.27	7.98	124.12	8.03	124.19	8.10	124.27	8.09	124.53	8.04	123.69	7.98	123.86
Y710	8.58	121.07	61.8	37.33	8.55	121.13	8.32	123.03	8.36	119.24	8.41	118.92	8.60	121.53	8.60	121.26	8.5	121.33	8.59	122.71
S711	7.83	113.78	62.1	-	7.80	113.62	7.83	113.98	7.83	114.33	7.89	113.93	7.84	113.78	7.80	113.54	7.68	113.5	7.92	114.88
I712	7.76	122.84	63.85	35.94	7.72	122.73	7.53	122.45	7.73	123.11	7.87	123.08	7.72	122.76	7.77	122.83	7.57	122.47	7.46	122.72
L713/F713	8.06	119.73	57.36	40.35	8.03	119.60	8.01	119.95	8.27	119.97	8.16	120.09	8.04	119.74	8.02	119.83	8.07	119.72	8.12	120.16
S714	8.32	116.73	62.78	62	8.29	116.69	8.28	116.29	8.32	116.44	8.35	116.97	8.33	116.93	8.34	117.10	8.44	116.94	8.2	115.88
E715	7.93	124.78	59.82	29.18	7.89	124.65	7.59	123.19	7.84	124.7	7.86	124.85	7.90	124.82	7.93	125.11	7.89	124.91	7.67	123.81
V716	8.82	123.49	66.28	30.29	8.79	123.45	8.52	121.93	8.78	123.18	8.85	123.31	8.82	123.52	8.81	123.73	8.81	123.32	8.76	123.09
Q717	8.16	119.57	58.38	27.59	8.14	119.42	7.83	118.47	8.21	119.72	8.18	119.93	8.20	119.63	8.22	119.63	8.31	119.8	8.33	119.73
Q718	7.59	119.63	58.28	27.63	7.56	119.51	7.7	119.68	8.02	119.49	7.59	119.57	7.57	119.70	7.57	119.80	7.65	119.53	7.56	119.57
A719	8	123.76	54.54	17.21	7.96	123.65	7.91	123.52	7.96	123.8	7.97	123.88	7.96	123.87	7.97	123.94	7.96	123.74	7.96	123.85
V720	8.67	118.76	64.97	30.81	8.64	118.73	8.38	118.95	8.64	118.57	8.67	118.88	8.67	118.86	8.68	118.84	8.7	119.29	8.66	118.98
S721	7.92	116.31	60.89	62.65	7.89	116.25	7.87	116.64	7.93	116.52	7.92	116.5	7.93	116.50	7.91	116.44	8.11	116.77	7.93	116.51
Q722	7.82	118.34	55.66	28.33	7.80	118.16	7.78	118.56	7.82	118.41	7.82	118.35	7.81	118.27	7.81	118.23	7.85	118.31	7.83	118.18
G723	7.62	109.38	45.69	-	7.58	109.34	7.64	109.23	7.59	109.34	7.59	109.43	7.60	109.42	7.60	109.44	7.6	109.37	7.6	109.5
S724	8.14	116.1	58.42	63.15	8.12	116.04	8.11	116	8.14	116.15	8.13	116.13	8.12	116.11	8.12	116.11	8.12	116.08	8.11	116.21
S725	8.45	117.34	57.23	64.1	8.45	117.42	8.44	117.47	8.46	117.42	8.46	117.45	8.45	117.40	8.46	117.48	8.45	117.36	8.44	117.37
D726	8.68	123.02	57.17	39.4	8.67	119.97	8.63	123.1	8.68	122.99	8.68	123.04	8.69	122.87	8.66	122.95	8.69	122.99	8.69	123
S727	8.37	115.73	61.04	61.74	8.35	115.76	8.34	115.63	8.35	115.75	8.35	115.7	8.35	115.77	8.36	115.85	8.37	115.62	8.38	115.53
Q728	7.62	123.73	58.26	27.78	7.59	123.70	7.63	123.67	7.6	123.87	7.59	123.88	7.60	123.86	7.59	123.84	7.58	123.83	7.59	123.84
I729	7.8	120.03	62.99	35.74	7.78	119.99	7.81	119.89	7.79	120.26	7.79	120.08	7.79	120.12	7.80	120.13	7.77	120.14	7.76	119.99
L730	8.34	125.86	58.33	40.45	8.30	125.68	8.29	126.11	8.27	125.5	8.32	125.86	8.31	125.72	8.32	125.88	8.33	125.66	8.3	125.46
D731	7.65	119.35	57.44	40.72	7.58	119.60	7.69	119.69	7.64	119.19	7.63	119.29	7.63	119.22	7.63	119.30	7.62	119.18	7.61	119.15
L732	8.03	119.1	57.93	41.75	8.00	118.98	7.98	119.46	8.01	119.38	8	119.01	8.01	118.97	8.00	118.82	7.94	119.08	7.92	119.09
S733	8.15	115.26	62.46	-	8.12	115.13	8.13	115.55	8.12	115.64	8.15	115.49	8.15	115.47	8.14	115.45	8.21	115.54	8.21	115.37
N734	8.53	119.14	55.7	36.98	8.51	118.99	8.58	119.52	8.56	119.84	8.52	119.28	8.53	119.17	8.52	119.20	8.52	119.05	8.52	119.06

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
R735	8.3	122.14	59.69	29.29	8.27	122.09	8.15	122.28	8.3	122.29	8.27	122.14	8.35	122.47	8.35	122.52	8.22	122.37	8.33	122.47
F736	8.24	122.87	63.04	38.13	8.21	122.79	8.05	122.77	8.31	122.61	8.22	122.98	8.22	122.93	8.21	123.01	8.24	123.13	8.24	122.79
Y737	8.05	115.86	58.46	36.5	8.02	115.79	7.94	115.91	8.13	116.13	8.06	116.13	8.02	115.94	8.02	115.91	8.06	116.26	8.02	116.08
T738	7.7	115.91	65.23	68.46	7.66	115.91	7.6	115.81	7.68	116	7.66	115.95	7.67	116.04	7.67	116.07	7.66	115.94	7.65	115.94
L739	7.14	123.1	56.72	42.06	7.10	122.94	7.19	123.38	7.21	123.47	7.12	123.44	7.11	123.03	7.11	123.13	7.07	122.96	7.09	122.75
I740	7.67	117.3	54.99	34.62	7.64	117.13	7.57	117.09	7.74	117.44	7.64	116.73	7.62	117.33	7.66	117.15	7.65	117.07	7.53	117.21
P741	-	-	63.15	31.44	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H742	8.88	123.59	54.04	32.3	8.86	123.55	8.88	123.5	9.22	124.27	8.85	123.66	8.93	123.73	8.88	123.68	8.98	123.72	-	-
D743	8.53	122.01	51.93	40.13	8.51	121.97	8.48	121.85	8.44	121.86	8.5	121.89	8.52	121.86	8.49	121.97	8.47	121.85	8.51	121.94
F744	8.35	123.2	57.16	38.49	8.36	123.13	8.39	123.02	8.41	123.21	8.37	123.37	8.37	123.25	8.38	123.36	8.4	123.23	8.37	123.29
G745	8.57	111.48	46.9	-	8.54	111.44	8.53	111.49	8.57	111.83	8.56	111.6	8.56	111.55	8.56	111.56	8.55	111.48	8.56	111.57
M746	-	-	55.04	31.37	8.85	125.50	-	-	8.93	124.68	-	-	-	-	-	-	8.9	125.22	-	-
K747	7.71	121.27	55.52	32.29	7.68	121.17	7.69	121.33	7.67	121.44	7.68	121.29	7.68	121.17	7.68	121.29	7.66	121.15	7.67	121.14
K748	8.02	122.52	53.57	30.76	8.02	122.46	8.04	122.68	8.03	122.7	8.04	122.66	8.04	122.58	8.03	122.80	8.03	122.54	8.03	122.47
P749	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
P750	-	-	61.88	31.06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L751	8.26	124.2	54.26	42.05	8.26	124.25	8.3	124.03	8.26	124.34	8.26	124.28	8.27	124.52	8.27	124.40	8.28	124.67	8.26	124.76
L752	7.78	123.49	52.41	38.3	7.77	123.49	7.69	122.28	7.75	124.03	7.78	123.76	7.78	123.95	7.78	123.65	7.75	124.49	7.7	124.53
N753	7.23	114.17	51.15	38.54	7.19	114.09	7.44	114.03	7.22	114.6	7.2	114.34	7.20	114.33	7.19	114.33	7.21	114.51	7.24	114.32
N754	7.35	113.07	51.56	39.92	7.31	112.98	7.28	115.13	7.32	113.34	7.32	113.08	7.31	112.84	7.33	113.08	7.29	112.34	7.3	112.77
A755	8.87	124.24	54.74	17.58	8.85	124.25	8.81	125.94	8.95	124.78	8.89	124.5	8.90	124.42	8.87	124.40	8.95	124.7	8.89	124.37
D756	8.26	118.76	57.03	39.38	8.24	118.67	8.21	118.51	8.28	118.99	8.27	118.91	8.27	119.02	8.26	118.92	8.25	118.84	8.25	118.85
S757	8.15	117.4	61.52	62.33	8.13	117.35	8.1	117.34	8.15	117.48	8.15	117.71	8.18	117.71	8.15	117.65	8.2	117.66	8.27	118.17
V758	7.26	122.55	66.68	30.37	7.23	122.47	7.51	121.88	7.21	122.35	7.23	122.45	7.27	122.67	7.25	122.60	7.27	122.78	7.32	122.68
Q759	7.94	118.81	58.59	27.06	7.92	118.73	7.98	119.45	7.97	119.36	7.99	119.01	7.90	118.76	7.96	118.94	8.04	119.06	7.96	118.84
A760	7.98	121.21	54.6	17.48	7.94	121.01	7.7	120.87	7.98	121.01	7.94	121.13	7.95	121.08	7.97	121.05	7.9	120.58	7.82	120.9
K761	7.9	116.29	56	30.39	7.87	116.23	7.81	117.53	7.75	116.77	7.84	116.54	7.84	116.61	7.82	116.13	7.77	116.99	7.84	117.62
A762	8.43	125.33	55.54	15.73	8.39	125.31	8.11	124.31	8.39	124.05	8.48	126.19	8.43	125.21	8.37	125.42	8.49	125.66	8.36	125.35
E763	7.64	117.22	58.93	28.44	7.67	117.34	7.63	118.23	7.95	118.98	7.77	118.35	7.70	117.47	7.65	117.14	7.65	119.18	7.65	116.98
M764	7.28	119.15	58.79	31.27	7.24	119.01	7.35	119.36	7.39	119.71	7.33	119.14	7.30	119.63	7.20	119.45	7.28	119.46	7.29	120.11
L765/F765/A765	8.05	120.94	57.93	39.84	8.01	120.70	7.91	120.16	7.96	118.98	8.09	123.78	8.01	120.92	7.94	121.18	8.03	120.78	8.07	121
D766	8.53	119.94	56.97	39.6	8.48	119.83	8.06	119.52	8.22	121.46	8.44	118.03	8.50	119.98	8.56	120.48	8.46	120.01	8.45	120.06
N767	7.51	117.82	55.5	38.27	7.46	117.84	7.69	118.3	7.33	117.29	7.59	118.43	7.45	117.96	7.30	117.21	7.63	118.37	7.62	117.32
L768	8.71	121.45	57.66	40.06	8.70	121.19	8.87	121.84	8.78	120.85	8.8	120.95	8.74	121.39	8.60	120.55	8.93	120.7	8.75	121.74
L769	8.61	120.21	58.22	41.03	8.60	120.12	8.61	120.44	8.49	119.97	8.61	120.29	8.65	120.33	8.63	120.18	8.54	120.63	8.44	120.07
D770	7.03	118.96	57.69	39.97	6.99	119.14	7.14	119.91	6.99	119.78	7.05	119.4	6.94	119.17	7.11	117.94	7.15	120.11	6.85	118.93
I771	8.37	127.83	65.92	37.1	8.39	127.94	8.46	128.2	8.33	128.24	8.45	128.16	8.36	128.21	8.56	128.46	8.7	129.21	-	-
E772	8.57	121.66	60.64	26.96	8.56	121.60	8.5	121.83	8.42	121.37	8.55	121.72	8.57	121.53	8.62	121.62	8.61	122.48	8.58	121.94
V773	7.46	119.02	65.9	30.53	7.41	118.87	7.53	119.6	7.42	119.59	7.43	119.14	7.49	119.20	7.34	118.67	7.61	119.88	7.42	119.72
A774	7.86	122.68	55.34	18.12	7.82	122.82	7.81	122.84	7.86	123.05	7.82	122.82	-	-	-	-	-	-	7.65	123.7

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
Y775	9.09	118.42	62.33	37.35	9.06	118.34	9.13	118.6	9.03	118.32	9.1	118.57	9.12	118.72	9.04	118.48	9.25	118.74	9.32	119.4
S776	8.63	115.68	61.5	62.22	8.61	115.53	8.7	115.68	8.5	115.61	8.67	115.72	8.64	115.67	8.59	115.69	8.68	115.72	8.81	115.64
L777	8.02	123.29	56.91	41.31	7.91	122.90	7.99	123.49	8.02	123.27	8	123.49	7.94	123.16	7.85	122.98	7.98	123.1	7.85	123.7
L778	7.8	120.36	56.72	41.33	7.78	120.27	7.77	120.48	7.8	120.19	7.79	120.47	7.81	120.44	7.76	120.33	7.78	120.24	7.74	120.49
R779	7.66	117.56	56.32	28.98	7.64	117.42	7.69	117.47	7.62	117.56	7.66	117.58	7.66	117.49	7.65	117.76	7.62	117.44	7.69	117.64
G780	7.63	108.56	45.1	-	7.60	108.48	7.61	108.67	7.62	108.72	7.61	108.63	7.60	108.54	7.62	108.58	7.6	108.51	7.58	108.61
G781	7.88	108.35	44.74	-	7.86	108.21	7.86	108.32	7.89	108.3	7.87	108.28	7.88	108.22	7.87	108.23	7.83	108.09	7.85	108.2
S782	8.15	115.79	57.3	63.45	8.14	115.66	8.13	115.64	8.14	115.66	8.15	115.72	8.15	115.71	8.17	115.76	8.13	115.6	8.13	115.61
D783	8.33	123.45	54.14	40.61	8.31	123.46	8.29	123.52	8.32	123.65	8.32	123.57	8.32	123.64	8.33	123.54	8.31	123.56	8.31	123.51
D784	8.05	121.15	53.21	40.58	8.03	121.03	8.03	121.08	8.04	121.18	8.04	121.13	8.04	121.16	8.04	121.12	8.02	121.01	8.03	121.11
S785	8.47	119.99	59.16	63.07	8.46	120.01	8.47	120.16	8.45	120.11	-	-	8.52	120.37	8.45	120.09	8.46	120.33	8.45	120.35
S786	8.56	117.87	58.91	63.3	8.53	117.70	8.54	117.85	8.54	117.86	8.53	117.82	8.55	117.84	8.54	117.84	8.55	117.82	8.55	117.82
K787	7.33	122.7	54.96	33.83	7.29	122.65	7.28	122.76	7.3	122.79	7.29	122.79	7.29	122.76	7.29	122.89	7.27	122.64	7.29	122.72
N788	8.84	126.05	52.19	42.68	8.80	125.96	8.84	126.23	8.85	126.22	8.85	126.24	8.86	126.16	8.85	126.21	8.84	126.2	8.84	126.08
P789	-	-	64.89	31.72	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I790	8.81	118.43	63.65	36.22	8.76	118.22	8.78	118.33	8.78	118.31	8.78	118.33	8.80	118.34	8.79	118.36	8.8	118.26	8.8	118.31
D791	7.23	121.55	57.11	39.15	7.18	121.35	7.19	121.71	7.18	121.56	7.18	121.62	7.18	121.58	7.18	121.53	7.17	121.75	7.23	121.8
V792	7.69	121.45	65.63	31.06	7.65	121.35	7.63	121.67	7.68	121.64	7.68	121.54	7.65	121.55	7.66	121.54	7.64	121.61	7.63	121.58
N793	7.85	119.15	55.3	35.81	7.83	119.01	7.86	119.26	7.85	119.27	7.86	119.26	7.85	119.25	7.83	119.29	7.86	119.37	7.89	119.26
Y794	8.39	119.59	61.19	38.18	8.36	119.51	8.37	119.59	8.37	119.65	8.38	119.69	8.38	119.78	8.37	119.64	8.35	119.46	8.43	120.4
E795	8.18	121.26	59.03	28.45	8.15	121.11	8.1	121.22	8.14	121.28	8.16	121.32	8.15	121.27	8.18	121.40	8.1	121.25	8.04	120.58
K796	7.45	117.59	57.95	32.22	7.39	117.47	7.42	117.59	7.44	117.71	7.42	117.67	7.42	117.71	7.42	117.83	7.46	117.65	7.45	117.69
L797	7.3	114.09	55.19	40.76	7.25	113.92	7.24	114.14	7.26	114.18	7.25	114.12	7.26	114.07	7.28	114.23	7.23	114.51	7.31	114.18
K798	7.39	114.79	56.54	29.08	7.36	114.71	7.39	114.86	7.39	114.88	7.39	114.89	7.38	114.87	7.37	114.99	7.41	114.86	7.38	114.91
T799	8.23	114.82	61.79	72.36	8.20	114.69	8.23	114.98	8.23	114.88	8.23	114.92	8.24	114.96	8.22	114.84	8.23	114.95	8.23	115.02
D800	8.17	128.37	53.91	42.22	8.12	128.25	8.15	128.58	8.14	128.52	8.14	128.52	8.15	128.45	8.13	128.46	8.18	128.73	8.1	128.38
I801	7.95	126.4	61.2	38.31	7.92	126.31	7.96	126.37	7.93	126.41	7.94	126.44	7.91	126.41	7.92	126.39	7.98	126.35	7.91	126.55
K802	8.67	126.29	53.45	35.48	8.59	125.92	8.64	126.35	8.64	126.32	8.63	126.33	8.64	126.29	8.62	126.41	8.65	126.29	8.63	126.25
V803	9.04	123.93	62.82	31.5	9.04	123.80	9.05	123.95	9.06	124.02	9.05	123.99	9.05	124.02	9.06	124.10	9.05	123.99	9.04	123.95
V804	8.53	131.43	61.7	31.05	8.53	131.51	8.49	131.57	8.53	131.66	8.52	131.64	8.53	131.62	8.55	131.66	8.47	131.51	8.48	131.52
D805	8.65	128.58	54.6	41.38	8.65	128.50	8.65	128.83	8.66	128.8	8.66	128.81	8.66	128.76	8.66	128.72	8.64	128.9	8.65	128.79
R806	8.65	127.24	58.66	29.14	8.65	127.26	8.66	127.38	8.67	127.45	8.66	127.48	8.67	127.47	8.67	127.53	8.66	127.28	8.66	127.42
D807	8.37	117.91	53.32	39.98	8.34	117.80	8.34	117.89	8.34	117.93	8.34	117.95	8.35	118.00	8.36	117.99	8.35	117.84	8.35	117.96
S808	7.52	115.71	-	65.72	7.49	115.67	7.5	115.82	7.5	115.82	7.5	115.82	7.50	115.75	7.50	115.77	7.52	115.74	7.5	115.75
E809	8.86	124.11	57.75	28.58	8.84	123.85	8.85	124.05	8.86	124.08	8.86	124.06	8.87	124.06	8.87	124.09	8.86	124.03	8.86	124.05
E810	8.51	118.28	59.23	28.87	8.51	118.23	8.48	118.4	8.51	118.44	8.51	118.44	8.52	118.45	8.54	118.42	8.47	118.56	8.5	118.49
A811	7.28	119.97	54.6	17.95	7.25	119.93	7.25	119.96	7.25	120.03	7.25	120.03	7.25	120.05	7.26	120.09	7.26	119.89	7.25	120.02
E812	7.64	118.71	58.97	28.19	7.59	118.58	7.6	118.78	7.61	118.77	7.61	118.75	7.61	118.78	7.63	118.76	7.61	118.85	7.6	118.76
I813	7.89	120.73	64.69	37.13	7.85	120.72	7.89	120.94	7.88	120.92	7.89	120.95	7.89	120.90	7.88	120.86	7.91	120.9	7.89	120.86
I814	7.59	120.27	65.36	36.68	7.56	120.16	7.57	120.33	7.58	120.36	7.57	120.31	7.58	120.38	7.59	120.43	7.57	120.27	7.57	120.34

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
R815	8.7	118.59	60.94	29.92	8.67	118.41	8.69	118.71	8.69	118.69	8.7	118.82	8.68	118.72	8.69	118.66	8.7	118.77	8.69	118.72
K816	8.05	122.14	59.57	31.41	8.07	122.17	8.04	122.37	8.11	122.33	8.07	122.29	8.05	122.24	8.05	122.36	8.07	122.25	8.03	122.22
Y817	8.31	122.76	56.92	37.31	8.27	122.59	8.29	122.91	8.28	122.64	8.28	122.59	8.29	122.71	8.32	122.97	8.29	122.73	8.27	122.8
V818	8.28	119.92	65.79	31.03	8.24	119.86	8.22	119.92	8.25	120.01	8.24	120	8.26	120.08	8.27	120.07	8.25	120.04	8.23	119.97
K819	8.02	117.72	58.48	31.91	7.98	117.59	7.96	117.7	7.98	117.73	7.98	117.74	7.96	117.57	8.00	117.83	7.95	117.68	7.99	117.71
N820	9.1	115.92	54.36	37.41	9.06	115.70	9.05	115.81	9.07	115.87	9.07	115.87	9.08	115.90	9.11	116.02	9.03	115.83	9.08	115.84
T821	7.33	106.22	61.47	69.44	7.28	106.15	7.27	106.56	7.28	106.29	7.28	106.32	7.28	106.31	7.30	106.02	7.26	106.73	7.28	106.33
H822	7.01	120.65	55.26	30.24	6.96	120.57	6.88	120.69	6.92	120.75	6.92	120.65	6.88	120.58	6.98	120.46	6.81	120.82	6.9	120.65
A823	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T824	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T825	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T826	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T827	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A828	8.35	122.89	52.4	18.42	8.34	122.85	-	-	8.34	122.7	8.34	122.7	8.32	122.93	8.36	122.92	8.31	121.97	8.37	122.85
Y829	6.82	112.07	54.55	38.66	6.79	111.89	6.77	111.75	6.78	111.89	6.79	111.88	6.73	111.70	6.76	112.01	6.72	111.29	6.77	111.84
D830	8.17	119.88	52.21	43.32	8.13	119.76	8.07	119.51	8.09	119.66	8.09	119.73	8.09	119.68	8.14	120.08	8.03	119.07	8.07	119.72
L831	8.53	121.26	53.09	44.55	8.56	121.08	-	-	-	-	8.59	121.08	8.56	121.47	8.55	120.28	8.8	120.88	8.62	121.5
E832	8.5	122.56	53.7	32.37	8.47	122.49	8.49	122.54	8.48	122.67	8.49	122.64	8.47	122.62	8.49	122.76	8.5	122.41	8.5	122.61
V833	9	126.16	63.03	30.44	9.00	126.13	9	126.34	9.01	126.39	9.01	126.39	9.00	126.41	9.02	126.47	8.97	126.34	9	126.44
I834	8.86	129.85	62.17	36.87	8.84	129.79	8.83	129.93	8.86	129.97	8.85	129.95	8.84	129.86	8.85	129.92	8.8	129.81	8.81	129.82
D835	7.14	114.84	53.52	47.63	7.10	114.72	7.09	114.96	7.11	114.98	7.11	115	7.10	114.99	7.12	114.94	7.07	114.83	7.09	114.97
I836	8.49	119.87	60.49	39.82	8.48	120.11	8.48	120	8.47	120.2	8.48	120.1	8.47	120.22	8.47	120.16	8.45	120.33	8.46	120
F837	9.29	124.18	54.97	39.68	9.26	123.84	9.24	124.07	9.28	123.84	9.27	123.97	9.26	123.76	9.30	123.79	9.23	124.13	9.25	123.6
K838	9.53	126.91	55.09	33.31	9.49	126.80	9.47	126.97	9.49	127.01	9.49	126.97	9.49	126.97	9.51	127.11	9.48	126.98	9.48	126.96
I839	7.72	120.59	57.77	40.57	7.67	120.60	7.73	120.46	7.71	120.59	7.7	120.6	7.70	120.56	7.71	120.66	7.74	120.49	7.71	120.63
E840	8.01	120.46	54.67	30.91	7.97	120.22	7.96	120.06	7.97	120.29	7.97	120.29	7.97	120.26	7.97	120.35	7.96	119.94	7.97	120.13
R841	9.55	131.61	56.84	29.8	9.53	131.83	9.58	132.03	9.56	131.97	9.56	131.98	9.56	131.94	9.57	131.95	9.58	132.07	9.62	131.89
E842	8.72	128.43	58.37	28.64	8.68	128.27	8.71	128.55	8.73	128.6	8.71	128.55	8.72	128.52	8.70	128.48	8.78	128.73	8.69	128.36
G843	9	115.1	45.66	-	8.99	115.04	9.01	115.22	9.02	115.21	9.02	115.21	9.02	115.22	9.01	115.24	9	115.18	9.02	115.27
E844	7.67	125.06	60.08	31.41	7.62	124.85	7.63	125.02	7.64	125.08	7.63	125.04	7.63	125.03	7.63	125.03	7.66	125.15	7.61	124.96
C845	8.99	118.58	62.54	25.58	8.94	118.64	8.99	118.65	8.98	118.68	8.98	118.72	8.99	118.68	8.99	118.77	8.99	118.51	8.97	118.68
Q846	8.29	120.06	58.55	27.01	8.31	120.16	8.31	120.16	8.31	120.19	8.32	120.16	8.32	120.10	8.37	120.24	8.28	120.25	8.31	119.92
R847	7.44	120.79	58.02	29.71	7.42	120.89	7.4	120.76	7.42	120.85	7.41	120.86	7.42	120.83	7.43	120.84	7.44	120.89	7.41	120.44
Y848	8.03	118.61	57.7	39.16	7.98	118.41	8.04	118.79	8.03	118.78	8.04	118.79	8.05	118.63	7.97	118.70	8.05	118.75	8.04	118.95
K849	7.11	118.42	61.04	29.81	7.09	118.23	7.08	118.63	7.08	118.55	7.09	118.56	7.08	118.46	7.08	118.35	7.08	118.72	7.08	118.47
P850	-	-	-	29.86	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F851	8.03	116.24	59.62	37.74	7.98	116.23	7.98	116.34	8	116.31	8	116.32	8.00	116.08	8.03	116.14	7.95	116.05	7.98	116.06
K852	7.23	120.4	57.62	31.27	7.18	120.29	7.17	120.38	7.18	120.42	7.18	120.43	7.18	120.32	7.21	120.40	7.18	120.34	7.2	120.3
Q853	6.86	111.22	54.32	28.25	6.84	111.00	6.87	111.23	6.86	111.18	6.86	111.19	6.85	111.19	6.84	111.30	6.85	111.16	6.84	111.17
L854	7.45	124.45	54.9	41.61	7.41	124.47	7.42	124.36	7.42	124.53	7.43	124.51	7.42	124.43	7.42	124.53	7.42	123.99	7.45	124.55

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
H855	7.66	119.07	54.88	30.23	7.63	119.04	7.64	119.04	7.65	118.65	7.67	118.63	7.63	118.97	7.66	118.78	7.65	118.79	7.63	118.83
N856	8.52	113.22	53.16	36.24	8.51	113.12	8.51	113.23	8.52	113.26	8.52	113.25	8.51	113.29	8.52	113.35	8.49	113.2	8.5	113.25
R857	7.9	119.24	54.85	29.62	7.88	119.06	7.86	119.26	7.87	119.27	7.86	119.26	7.88	119.27	7.87	119.30	7.83	119.21	7.84	119.27
R858	8.32	122.35	53.62	34.47	8.30	122.27	-	-	8.27	122.28	8.27	122.21	8.27	122.29	8.28	122.31	8.2	122.73	8.27	122.93
L859	7.99	125.5	53	41.86	7.96	125.41	7.98	125.76	7.97	125.66	7.97	125.67	7.96	125.74	7.96	125.50	8.01	125.96	7.95	125.75
L860	8.67	126.45	53.01	45.71	8.63	126.46	8.64	126.53	8.64	126.56	8.63	126.63	8.63	126.91	8.65	126.63	8.63	126.99	8.64	126.97
W861	8.51	119.89	55.85	32.2	8.49	119.95	8.48	120.01	8.51	119.99	8.49	119.97	8.52	119.96	8.5	119.94	8.47	120.01	-	-
H862	8.9	118.32	56.75	31.02	8.84	118.35	8.79	118.45	8.83	118.63	8.82	118.65	8.72	116.93	8.87	116.92	8.81	116.43	-	-
G863	8.12	115.47	43.52	-	8.31	111.88	8.41	110.26	-	-	8.15	115.48	8.42	110.23	8.41	110.23	8.41	110.2	-	-
S864	7.25	111.72	57.68	64.01	7.23	111.63	7.24	111.62	7.2	111.59	7.21	111.56	7.25	111.51	7.34	113.52	7.15	112.36	-	-
R865	7.86	121.88	56.69	30.21	7.80	121.87	7.79	121.94	7.72	121.9	7.8	122.03	7.81	121.85	7.71	122.41	7.54	121.96	7.65	120.87
T866	9.24	119.18	66.75	67.94	9.23	119.26	9.25	119.78	9.29	120.01	9.25	119.66	9.23	119.41	9.35	120.29	9.28	120.59	9.2	119.39
T867	7.28	107.36	62.72	67.98	7.25	107.26	7.33	107.54	7.29	107.17	7.31	107.41	7.29	107.32	7.27	107.43	7.44	107.25	7.38	107.38
N868	7.88	118.6	53.88	38.29	7.83	118.42	7.82	118.43	7.88	118.53	7.74	118.5	7.89	118.74	7.82	118.28	7.83	118.01	7.85	117.93
F869	7.75	118.29	63.74	38.94	7.62	118.25	7.72	118.91	7.64	118.75	7.58	118.43	7.65	118.51	7.64	118.77	7.73	119.04	7.69	118.96
A870	9.36	125.93	55.19	15.63	9.31	125.90	9.18	126	9.34	126	9.27	126.12	9.24	126.10	9.48	126.09	9.11	125.96	9.23	126.1
G871	8.78	109.8	46.25	-	8.77	109.93	8.81	110.14	8.74	109.87	8.8	110.19	8.76	110.01	8.78	109.49	8.86	110.19	8.81	110.02
I872	8.08	122.28	64.68	37.15	8.04	122.32	8.14	122.39	8.06	122.4	8.05	122.38	8.06	122.13	8.2	122.44	8.33	122.64	-	-
L873	8.5	116.69	57.62	39.36	8.48	116.93	8.5	116.91	8.52	117.05	8.52	116.91	8.50	116.89	8.51	116.86	8.54	117.67	-	-
S874	7.84	112.7	61.24	63.24	7.81	112.67	7.88	113.06	7.81	112.83	7.84	112.93	7.80	112.84	7.8	112.82	7.89	112.8	-	-
Q875	8.33	117.82	55.29	28.98	8.29	117.70	8.29	118.06	8.33	117.82	8.3	117.92	8.32	117.96	8.31	117.42	8.34	117.37	-	-
G876	7.73	114.27	44.03	-	7.69	114.28	7.79	114.39	7.81	114.46	7.75	114.44	7.76	114.50	7.66	114.63	7.91	114.44	7.73	114.07
L877	8.55	117.86	54.88	39.63	8.54	117.68	8.45	117.53	8.55	117.71	8.54	117.85	8.55	118.03	8.47	118.62	8.42	117.88	8.56	118.21
R878	8.54	124.98	54.55	32.92	8.53	125.16	8.52	125.25	8.49	125.17	8.5	125.23	-	-	7.92	127.46	8.43	126.23	-	-
I879	7.47	119.65	59.86	38.09	7.46	119.62	7.64	120.47	7.59	120.08	7.55	120.11	7.53	120.11	6.89	114.56	7.63	120.42	-	-
A880	8.17	131.05	50.73	16.1	8.15	130.66	8.08	130.95	8.22	131.76	8.24	131.99	8.15	131.45	7.74	129.39	8.38	132	8.3	131.75
P881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
P882	-	-	65.17	31.08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E883	9.22	114.18	58.45	28.81	9.17	114.04	8.96	113.76	9.17	114.05	9.25	114.1	9.18	114.26	9.24	114.37	9.25	114.13	9.22	114.37
A884	7.3	120.1	49.61	17.8	7.27	120.04	7.23	121.03	7.03	120.48	7.18	120.44	7.17	119.94	7.3	120.40	7.03	120.61	7.01	119.93
P885	-	-	61.75	30.39	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
V886	8.93	124.72	64.64	30.58	8.93	124.60	9.09	125.21	8.9	125.46	8.88	124.78	8.82	123.05	9.05	125.90	-	-	8.86	123.32
T887	7.27	107.67	62.31	67.91	7.25	107.35	7.21	108.49	7.26	109.48	7.29	109.47	7.23	109.63	7.35	107.16	-	-	-	-
G888	7.9	108.5	44.79	-	7.89	108.50	7.88	108.81	8.08	108.9	8.02	109.2	8.03	109.26	7.9	108.08	8.53	113.97	8.01	108.78
Y889	7.03	117.49	56.55	39.98	7.00	117.49	7.2	118.93	7.16	119.07	7.1	118.23	7.22	118.34	7.02	117.58	7.19	116.65	7.38	116.47
M890	9.19	126.51	59.55	32.77	9.19	126.48	-	-	-	-	9.23	126.1	9.69	126.73	9.29	127.60	9.36	128.31	-	-
F891	8.59	113.27	56.39	38.61	8.57	113.25	8.46	114.22	8.47	113.91	8.41	113.97	8.31	113.44	8.71	113.15	7.97	114.03	-	-
G892	7.82	105.25	43.64	-	7.78	105.19	7.7	105.42	7.73	105.36	7.76	105.39	7.74	104.83	7.83	104.98	7.38	104.77	7.68	105.05
K893	8.99	121.88	56.98	28.8	8.96	120.74	8.98	121.97	8.94	121.97	8.93	121.79	8.83	121.60	8.98	121.76	8.7	121.15	8.81	121.65
G894	6.97	111.17	43.5	-	6.93	111.27	6.87	111.32	6.89	111.38	6.9	111.45	6.76	110.90	6.72	112.68	6.78	111.05	6.79	111.55

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
I895	8.57	122.01	60.65	37.62	8.51	121.84	8.54	122.05	8.39	122.25	8.48	121.88	8.51	121.92	8.3	121.56	-	-	-	-
Y896	8.63	127.63	57.77	38.13	8.64	127.64	8.72	127.86	8.62	127.66	8.64	127.59	8.43	127.08	8.18	128.14	8.79	127.74	-	-
F897	9.85	121.37	56.69	45.03	9.81	121.23	9.82	121.24	9.81	121.2	9.79	121.22	10.34	118.44	10.22	118.26	9.98	117.79	10.47	118.02
A898	9.1	124.65	51.13	22.31	9.02	124.61	9.07	117.24	8.96	124.35	8.99	124.49	-	-	9.05	125.93	9.16	120.66	-	-
N899	9.07	117.46	52.8	42.75	9.02	117.31	9.08	117.51	9.11	117.43	9.03	117.2	9.02	116.71	9.03	117.12	8.96	116.63	8.95	116.65
M900	7.07	113.34	53.86	33.72	7.05	113.20	7.06	113.33	7.08	113.54	7.09	113.25	7.10	113.45	7.17	113.59	-	-	-	-
V901	8.19	126.87	63.73	29.54	8.18	126.92	8.1	127.18	8.15	127.17	8.12	127.17	8.17	127.17	8.18	127.23	8.14	127.35	8.19	127.12
S902	6.99	113.91	60.32	62.03	7.03	114.10	6.89	113.83	6.8	113.61	6.86	113.88	6.75	113.34	6.92	113.80	6.63	113.55	6.95	113.73
K903	6.56	119.56	56.92	30.76	6.53	119.87	6.63	120.33	6.53	120.02	6.57	120	6.70	120.01	6.55	118.97	6.83	121.2	6.67	119.73
S904	6.6	112.72	61.14	63.77	6.54	112.86	6.49	113.03	-	-	6.52	112.91	6.32	113.67	6.41	113.44	6.45	113.52	6.48	114.64
A905	8.74	122.7	53.67	16.53	8.71	122.91	8.71	123.65	8.71	122.98	8.72	123.25	8.85	124.05	8.86	124.09	8.87	124.97	8.96	124.35
N906	7.19	116.67	54.88	37.45	7.16	116.74	7.2	117.08	7.17	116.9	7.18	116.99	7.05	116.87	6.94	116.30	7.11	117.31	7.08	117.1
Y907	7.54	117.3	58.65	36.49	7.43	116.67	7.53	117.61	7.48	117.44	7.51	117.55	7.32	117.28	7.35	116.44	7.28	116.54	-	-
C908	7.32	117.12	60.63	27.6	7.29	117.06	7.34	117.32	7.25	117.33	7.31	117.27	7.22	117.26	7.23	117.79	7.23	117.7	7.22	117.5
H909	7.43	113.56	56.42	26.3	7.37	113.39	7.42	113.94	7.34	113.25	7.38	113.55	7.35	113.93	7.47	113.29	7.22	112.81	7.43	114.57
T910	7.21	108.91	60.43	70.18	7.20	108.92	7.14	108.49	7.19	108.83	7.19	108.82	7.16	108.55	7.09	108.31	7.06	107.96	7.1	108.32
S911	7.92	113.96	57.03	65.12	7.89	113.96	7.85	113.86	7.89	113.98	7.89	113.92	7.84	113.58	7.89	113.84	7.82	113.38	7.83	113.57
Q912	8.21	119.43	58.97	26.99	8.19	119.39	8.21	119.62	8.21	119.65	8.21	119.64	8.22	119.62	8.21	119.63	8.21	119.62	8.23	119.64
G913	7.93	104.83	45.05	-	7.92	104.73	7.92	104.82	7.92	104.85	7.92	104.84	7.91	104.76	7.92	104.81	7.9	104.68	7.91	104.8
N914	7.12	119.9	51.72	40.31	7.07	119.70	7.08	119.91	7.08	119.81	7.08	119.84	7.07	119.86	7.08	119.80	7.06	119.89	7.07	119.86
P915	-	-	64.73	30.91	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I916	7.65	122.42	58.19	36.74	7.62	122.35	7.63	122.36	7.62	122.47	7.63	122.44	7.60	122.39	7.6	122.42	7.6	122.45	7.61	122.44
G917	8.96	113.8	42.86	-	8.93	113.83	8.94	114.03	8.94	114.08	8.95	114.12	8.96	114.13	8.95	114.04	8.93	114.11	8.94	114
L918	9.46	121.52	52.39	43.61	9.44	121.40	9.42	121.38	9.47	121.61	9.46	121.51	9.46	121.47	9.51	121.54	9.42	121.35	9.42	121.21
I919	8.2	115.38	60.16	38.03	8.17	115.22	8.2	115.52	8.17	115.25	8.17	115.43	8.17	115.39	8.18	115.78	8.23	116.26	8.19	116.28
L920	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L921	8.73	118.96	53.14	45.02	8.68	118.78	8.62	118.55	8.7	118.9	8.7	118.84	8.71	118.82	8.68	118.74	8.72	118.76	-	-
G922	9.36	110.05	43.2	-	9.32	110.01	9.34	110.15	9.36	110.23	9.34	110.2	9.40	110.26	9.4	110.32	9.38	110.3	9.39	110.42
E923	8.75	126.54	55.22	29.33	8.72	126.48	8.69	126.56	8.7	126.52	8.69	126.52	8.69	126.48	8.71	126.54	8.69	126.63	8.7	126.56
V924	8.65	127.58	59.81	32.1	8.65	127.50	8.64	127.7	8.61	127.64	8.67	127.73	8.63	127.71	8.6	127.55	8.64	128.03	8.61	127.79
A925	8.17	130.81	50.69	17.57	8.19	130.99	8.15	131.24	8.1	130.97	8.12	130.96	8.11	130.95	8.13	130.94	8.19	131.26	8.05	130.95
L926	7.6	123.73	55.53	40.84	7.56	123.79	7.56	123.89	7.55	123.86	7.55	123.84	7.54	123.89	7.55	123.79	7.49	124.04	7.53	123.91
G927	7.23	105.13	45.28	-	7.17	105.00	7.17	105.08	7.17	105.13	7.17	105.11	7.16	105.10	7.19	105.21	7.15	104.86	7.17	105.13
N928	11.55	128.66	52.23	37.53	11.53	128.65	11.54	128.79	11.54	128.79	11.53	128.8	11.50	128.76	11.51	128.82	11.52	128.65	11.51	128.73
M929	8.8	126.01	55.16	32.58	8.79	126.12	8.77	126.24	8.79	126.2	8.78	126.22	8.78	126.19	8.79	126.08	8.72	126.37	8.76	126.17
Y930	9.62	128.87	53.61	36.44	9.60	128.83	9.57	128.97	9.59	129.02	9.59	129.03	9.61	129.09	9.62	129.08	9.58	129	9.59	129.07
E931	8.59	128.94	55.77	29.57	8.59	128.78	8.59	128.95	8.6	129	8.6	128.98	8.61	128.98	8.6	129.02	8.58	128.91	8.59	128.95
L932	8.65	125.08	53.06	46.66	8.61	125.06	8.59	125.38	8.61	125.36	8.61	124.98	8.58	125.29	8.66	125.30	8.74	125.58	8.52	125.22
K933	8.45	118.31	55.65	33.44	8.43	118.24	8.46	118.31	8.45	118.32	8.44	118.44	8.44	117.97	8.44	118.24	8.61	118.38	8.44	118.09
H934	7.1	113.14	53.4	32.32	7.07	113.17	7.07	113.43	7.06	113.3	7.09	113.25	7.12	113.48	7.11	113.07	7.08	113.4	7.08	113.5

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C^{2}H$				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
A935	8.34	122.4	52.73	18.07	8.41	122.28	8.33	122.76	8.32	122.68	8.34	122.71	8.47	122.94	8.37	122.58	8.32	122.84	8.37	122.52
S936	6.84	116.31	57.37	65.09	6.79	116.26	7.01	116.58	6.76	116.75	6.77	116.8	7.27	116.75	6.82	116.45	7.66	116.73	6.78	116.69
H937	-	-	54.9	27.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I938	6.99	126.92	59.94	37.09	6.99	126.86	6.89	127.09	6.92	127.61	6.98	127.3	7.06	126.67	7.06	126.41	6.84	128.17	6.86	128.06
S939	8.62	120.29	-	-	8.41	120.04	-	-	8.59	120.31	8.62	120.11	8.59	119.77	8.62	119.58	8.36	-	8.69	120.17
K940	7.31	121.29	54.65	33.62	7.28	121.19	7.29	121.27	7.28	121.3	7.29	121.32	7.30	121.33	7.28	121.55	7.27	120.94	7.29	121.12
L941	8.39	125.66	52.47	40.45	8.38	125.59	8.41	125.99	8.41	125.96	8.4	125.9	8.38	125.81	8.37	125.81	8.57	126.08	8.35	125.63
P942	-	-	61.75	30.98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
K943	8.04	122.14	57.5	31.28	8.02	122.27	8.02	122.32	8.01	122.29	8.02	122.3	8.01	122.27	8.02	122.41	8.03	122.25	8.03	122.38
G944	8.61	113.81	44.66	-	8.62	113.84	8.63	114.05	8.63	114.04	8.63	114.04	8.62	114.03	8.63	114.00	8.61	114.12	8.65	114.09
K945	7.37	116.79	52.67	32.81	7.32	116.62	7.34	116.76	7.33	116.75	7.33	116.74	7.32	116.74	7.29	116.79	7.37	116.9	7.35	116.77
H946	9.49	118.2	57.49	-	9.47	118.09	9.44	118.16	9.46	118.21	9.46	118.21	9.44	118.12	9.47	118.21	9.37	117.99	9.44	118.14
S947	8.15	114.16	57.69	64.53	8.11	114.17	8.1	114.24	8.11	114.31	8.11	114.31	8.10	114.39	8.12	114.47	8.07	114.13	8.09	114.32
V948	9.2	121.03	60.38	34.47	9.17	120.92	9.1	120.97	9.15	121.06	9.14	121.07	9.17	121.11	9.23	121.28	9.03	121.05	9.18	121.22
K949	8.09	126.27	53.4	36.01	8.02	126.04	7.99	126.25	8.01	126.2	8	126.17	7.95	126.07	7.98	126.18	7.91	126.23	7.99	126.15
G950	9.35	116.05	44.03	-	9.15	117.28	9.27	116.1	9.31	116.25	9.31	116.25	8.83	120.98	8.87	121.15	8.7	120.69	9.29	116.36
L951	8.43	125.57	54.05	41.57	8.41	125.59	8.36	125.56	8.38	125.55	8.39	125.54	-	-	-	-	-	-	8.37	125.49
G952	8.86	109.92	43.19	-	8.83	109.79	8.82	110.06	8.85	109.92	8.85	109.92	8.93	109.71	8.9	109.89	8.9	109.85	8.9	109.82
K953	7.75	121.44	58.65	33.11	7.69	121.39	7.71	121.59	7.65	121.57	7.65	121.59	7.58	121.44	7.65	121.30	7.66	121.48	7.63	121.59
T954	7.59	117.06	61.2	69.74	7.54	116.95	7.55	117.1	7.56	117.12	7.56	117.13	7.63	117.16	7.53	117.15	7.47	116.71	7.59	117.17
T955	9.09	121.43	56.83	72.46	9.06	121.31	9.09	121.39	9.09	121.51	9.09	121.48	9.12	121.64	9.07	121.70	9.15	121.34	9.11	121.56
P956	-	-	63.56	31.61	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D957	8.49	126.63	52.06	40.7	8.45	126.53	8.44	126.86	8.45	126.76	8.45	126.74	8.43	126.58	8.47	126.53	8.41	126.81	8.43	126.65
P958	-	-	63.97	31.05	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S959	8.4	115.69	60.05	62.64	8.35	115.59	8.35	115.65	8.36	115.73	8.35	115.7	8.37	115.78	8.35	115.80	8.35	115.6	8.37	115.78
A960	7.47	123.99	50.56	19.7	7.44	123.86	7.46	124.05	7.45	124.06	7.45	124.05	7.44	124.07	7.42	124.04	7.52	123.98	7.43	124.04
N961	7.02	117.15	54.27	37.92	6.98	117.02	6.97	117.13	6.98	117.15	6.98	117.15	6.98	117.15	7.01	117.19	6.94	117.11	6.97	117.12
I962	7.76	117.56	58.64	41.01	7.72	117.30	7.7	117.4	7.71	117.3	7.7	117.52	7.70	117.22	7.69	117.20	7.68	117.39	7.68	117.3
S963	8.11	116.15	55.97	63.83	8.05	116.16	8.11	116.14	8.1	116.19	8.13	116.13	8.11	116.18	8.1	116.12	8.12	116.46	8.11	116.24
L964	9.12	129.33	53.02	42.64	9.10	129.23	9.1	129.36	9.09	129.32	9.1	129.32	9.09	129.26	9.08	129.17	9.08	129.33	9.08	129.24
D965	9.2	128.38	54.78	39.43	9.17	128.18	9.16	128.4	9.22	128.64	9.18	128.41	9.17	128.39	9.21	128.33	9.11	128.33	9.21	128.53
G966	7.93	102.91	44.94	-	7.92	102.84	7.93	102.93	7.94	102.94	7.94	102.94	7.94	102.94	7.95	102.98	7.94	102.91	7.94	102.98
V967	7.39	122.18	60.5	31.93	7.36	122.10	7.36	122.18	7.37	122.18	7.37	122.2	7.37	122.19	7.37	122.18	7.36	122.16	7.36	122.21
D968	8.22	127.04	55.01	41.95	8.19	126.94	8.19	127.09	8.19	127.17	8.19	127.16	8.18	127.17	8.22	127.18	8.18	127	8.21	127.12
V969	8.99	124.04	58.67	31.36	8.96	123.91	8.97	123.98	8.97	124.03	8.97	124.05	8.98	124.15	8.97	124.13	8.99	124.03	8.96	124.09
P970	-	-	60.85	28.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L971	6.74	112.34	53.11	39.97	6.68	111.93	6.7	111.82	6.72	111.88	6.72	111.85	6.73	111.92	6.74	112.19	6.7	111.93	6.72	111.9
G972	8.52	106.78	45.54	-	8.52	106.68	8.56	106.84	8.55	106.82	8.55	106.83	8.55	106.87	8.52	106.82	8.59	106.93	8.54	106.88
T973	7.55	108.73	61.16	69.5	7.53	108.56	7.54	108.75	7.53	108.65	7.53	108.65	7.53	108.61	7.52	108.50	7.56	108.73	-	-
G974	8.24	108.83	46.27	-	8.21	108.77	8.19	109.09	8.23	109	8.21	108.98	8.24	108.85	8.26	108.79	8.22	109.04	8.22	108.97

Chemical shift assignments for wild-type PARP-1 catalytic domain, catalytic domain mutants and catalytic domain-inhibitor complexes

	WT 30°C $^{15}N^{13}C$ ^{2}H				WT 25°C ^{15}N		L713F 25°C ^{15}N		L765F 25°C ^{15}N		L765A 25°C ^{15}N		Veliparib 1:1.5 25°C ^{15}N		Olaparib 1:1.5 25°C ^{15}N		Talazoparib 1:1.5 25°C ^{15}N		EB-47 1:1.5 25°C ^{15}N	
	H	N	CA	CB	H	N	H	N	H	N	H	N	H	N	H	N	H	N	H	N
I975	9.22	123.34	59.1	42.48	9.19	123.18	9.17	123.28	9.19	123.23	9.19	123.23	9.20	123.12	9.21	123.06	9.18	123.14	9.19	123.13
S976	8.33	115.42	58.11	62.44	8.33	115.31	8.35	115.58	8.35	115.5	8.35	115.64	8.37	115.52	8.35	115.55	8.35	115.39	8.36	115.66
S977	8.68	122.72	58.19	64.73	8.67	122.81	8.69	122.97	8.67	123.14	8.68	123.04	8.69	123.07	8.69	123.11	-	-	8.68	123
G978	8.31	110.89	45.09	-	8.28	110.86	8.3	111.02	8.3	111.01	8.3	111.01	8.31	111.00	8.3	110.96	8.31	111.01	8.3	110.96
V979	7.59	121.99	61.58	30.3	7.56	121.86	7.56	121.98	7.57	122	7.57	121.99	7.61	122.07	7.58	121.95	7.63	122.14	7.6	122.05
N980	8.44	124.49	-	-	8.43	124.42	8.42	124.39	8.43	124.51	8.44	124.55	8.43	124.14	8.45	124.52	8.41	123.94	-	-
D981	8.8	119.74	53	38.26	8.77	119.68	8.78	119.86	8.79	119.93	8.79	119.94	8.81	120.19	8.8	119.96	8.8	120.28	8.8	120.04
T982	6.86	110.12	57.59	69.43	6.83	109.93	6.82	110	6.83	110.01	6.82	110.02	6.83	109.88	6.83	110.00	6.79	109.86	6.82	109.96
S983	8.3	120.96	58.04	63.45	8.28	120.90	8.29	121.07	8.3	121.08	8.29	121.08	8.34	121.15	8.29	121.04	8.32	121.19	8.35	121.04
L984	8.12	125.84	54.18	40.42	8.08	124.93	8.06	125.82	8.07	125.87	8.07	125.86	8.03	125.35	8.06	125.72	7.96	124.41	8.05	125.42
L985	8.76	124.31	57.5	41.65	8.74	124.32	8.71	124.24	8.72	124.28	8.72	124.25	-	-	8.79	124.19	8.97	123.63	8.72	124.1
Y986	7.1	114.25	55.47	41.39	7.06	114.19	7.08	114.33	7.13	114.44	7.08	114.37	7.13	114.25	7.05	114.49	6.99	114.18	7.13	114.75
N987	10.32	119.75	52.61	38.5	10.31	119.70	10.34	119.8	10.34	119.83	10.33	119.8	10.34	119.35	10.33	120.10	10.29	119.01	10.32	119.69
E988	8.36	119.91	55.16	33.5	8.33	119.78	8.34	120.01	-	-	8.35	120.1	8.35	120.12	8.34	118.72	8.3	119.24	8.38	119.47
Y989	8.64	121.08	55.8	39.78	8.57	121.03	8.5	120.99	-	-	8.59	121.08	-	-	-	-	-	-	-	-
I990	9.15	123.92	60.19	39.73	9.10	123.67	9.06	123.84	9.07	123.56	9.07	123.86	9.04	122.65	9.04	123.30	9.05	123.36	9.05	122.87
V991	7.59	113.82	58.29	33.32	7.57	113.67	7.57	113.66	7.58	113.64	7.56	113.61	7.46	113.34	7.58	113.74	7.47	113.01	7.72	113.74
Y992	8.91	118.65	57.75	37.11	8.88	118.49	8.89	118.79	8.87	118.64	8.89	118.68	8.86	118.79	-	-	8.75	118.43	8.81	118.45
D993	6.71	120.26	52.43	43.35	6.66	120.15	6.65	120.25	6.66	120.28	6.66	120.3	6.64	120.21	6.66	120.20	6.67	120.35	6.66	120.27
I994	8.3	121.29	63.33	36.95	8.28	121.13	8.3	121.54	8.31	121.43	8.3	121.09	8.26	121.30	8.31	121.03	8.33	121.24	8.29	121.58
A995	9.06	123.65	52.64	18.25	9.02	123.52	9	123.72	9.02	123.71	9.02	123.71	9.01	123.68	9.01	123.61	9	123.39	9.01	123.51
Q996	6.98	113.32	56.53	29.92	6.94	113.20	6.95	113.67	6.95	113.59	6.94	113.56	6.94	113.58	6.96	113.37	6.98	114.18	6.97	114.02
V997	6.62	116.39	59.49	34.28	6.56	116.32	6.6	116.26	6.58	116.33	6.58	116.33	6.56	116.39	6.57	116.51	6.6	116.31	6.56	116.02
N998	8.6	125.45	49.08	39.54	8.57	125.46	8.57	125.69	8.58	125.65	8.57	125.64	8.59	125.63	8.52	125.23	8.59	125.69	8.57	125.8
L999	10.11	129.34	56.45	40.84	10.08	129.20	10.07	129.48	10.09	129.53	10.09	129.5	10.09	129.51	10.1	129.53	10.07	129.4	10	129.23
K1000	8	118.51	55.53	35.35	8.03	118.66	8.04	118.75	8.06	118.64	8.05	118.78	8.02	118.68	7.93	118.75	8.03	118.71	8	118.53
Y1001	8.19	113.39	55.46	-	8.14	113.44	8.13	113.41	8.18	113.5	8.15	113.48	8.18	113.57	8.2	113.54	8.17	113.47	8.19	113.58
L1002	9.29	124.08	54.95	39.67	9.28	124.05	9.26	123.99	9.27	124.16	9.26	124.11	9.28	124.16	9.28	124.15	9.25	123.84	9.26	124.15
L1003	9.64	127.85	54.15	-	9.62	128.12	9.64	128.05	9.64	128.18	9.64	128.13	9.69	128.09	9.68	128.28	9.68	128.06	9.65	128.16
K1004	8.54	122.12	54.8	-	8.49	122.12	8.48	122.41	8.51	122.28	8.49	122.28	8.50	122.22	8.5	122.34	8.47	122.35	8.42	122.19
L1005	9.22	128.5	52.98	44.55	9.21	128.40	9.22	128.66	9.21	128.41	9.21	128.57	9.22	128.62	9.22	128.58	9.22	128.59	9.15	128.34
K1006	9.01	122.62	54.34	34.06	8.99	122.56	8.99	122.76	9	122.67	9	122.77	9.01	122.76	9.01	122.89	8.97	122.8	9.03	122.89
F1007	9.13	128.3	55.41	38.36	9.09	128.16	9.11	128.3	9.1	128.4	9.1	128.39	9.09	128.37	9.12	128.54	9.1	128.31	9.11	128.35
N1008	8.57	122.71	51.31	38.05	8.54	122.65	8.55	122.91	8.55	122.94	8.55	122.95	8.55	122.97	8.51	122.82	8.57	122.97	8.55	122.99
F1009	8.55	125.13	59.66	38.84	8.54	125.13	8.53	125.25	8.53	125.28	8.53	125.27	8.54	125.22	8.56	125.40	8.51	125.14	8.54	125.31
K1010	8.32	123.33	55.85	32.01	8.28	123.19	8.3	123.43	8.3	123.37	8.3	123.41	8.30	123.26	8.27	123.38	8.29	123.26	8.29	123.3
T1011	8.12	115.01	61.33	69.07	8.11	114.99	8.11	115.07	8.11	115.07	8.12	115.05	8.11	114.97	8.11	115.02	8.09	114.95	8.1	114.91
S1012	8.09	118.15	57.52	63.19	8.09	118.15	8.09	118.29	8.09	118.31	8.1	118.31	8.08	118.24	8.08	118.22	8.08	118.22	8.08	118.22
L1013	7.96	124.69	54.98	41.24	7.96	124.59	7.96	124.72	7.96	124.74	7.96	124.73	7.97	124.65	7.96	124.64	7.96	124.67	7.97	124.64
W1014	7.36	125.83	57.9	29.37	7.34	125.69	7.34	125.85	7.34	125.85	7.35	125.84	7.34	125.75	7.34	125.82	7.33	125.69	7.33	125.69