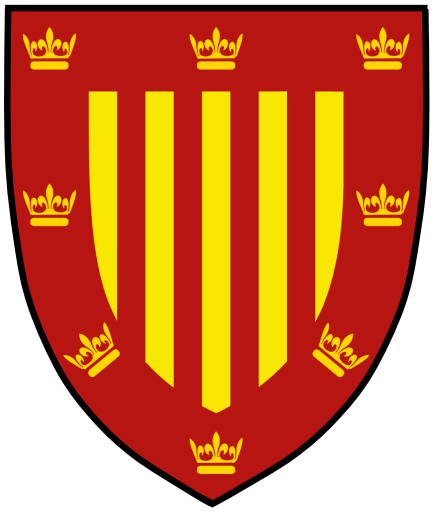
FROM ANTIBODY-DRUG CONJUGATES TO MASKED ANTIBODIES: BIOPHYSICAL INSIGHT FOR THE RATIONAL DESIGN OF FUTURE THERAPIES



Carolina Orozco

Peterhouse

Yusuf Hamied Department of Chemistry

University of Cambridge

This dissertation is submitted for the degree of *Doctor of Philosophy*

September 2021

To my family

Miriam, Manuel and Natalia Orozco

To Antoine

DECLARATION

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

Carolina Orozco

28th September 2021

ABSTRACT

FROM ANTIBODY-DRUG CONJUGATES TO MASKED ANTIBODIES: BIOPHYSICAL INSIGHT FOR THE RATIONAL DESIGN OF FUTURE THERAPIES

Carolina OROZCO

Monoclonal antibodies have become a major class of therapeutics over the past twenty-five years and, during this time, many successful antibody scaffolds have been based on wild-type humanized and human sequences of the immunoglobulin G isotype 1 (IgG1) subclass. In the past, several mutations have been engineered into these scaffolds, to optimise biological/physical properties or to allow site-specific conjugation. In the first project, the thermodynamic, thermal and kinetic stabilities of two sets of mutations were measured: one was a triple mutation in the C_H2 domain (L234F/L235E/P331S), and the other had an additional substitution in the C_H3 domain, S442C, to enable site-specific conjugation of a cytotoxic payload. Overall, results showed that the order of the domains in increasing stability were C_H2, C_H3 and Fab. The triple mutation was found to destabilize the C_H2 domain, while the substitution in the C_H3 domain did not affect stability significantly.

Antibody-drug conjugates have been one of the most actively developed classes of drugs in the past fifteen years combining the strengths of large and small molecule therapeutics. Recently, strategies involving the insertion of a cysteine and maleimide linkers to achieve site-specific conjugation have been developed in order to attain highly controlled drug to antibody ratios. One such antibody scaffold, Fc-C239i, that formed an unexpected disulfide bridge during manufacture, was characterized. A combination of mass spectrometry and biophysical techniques were used to understand how the additional disulfide bridge forms, interconverts, and changes the stability and structural dynamics of the antibody.

Alternative strategies to the use of maleimide linkers have recently been developed, e.g. DVP and tetraDVP linkers, by the Spring Group in the Department of Chemistry. The impact of the conjugation of DVP and tetraDVP-linkers on the stability and dynamics of the antibody trastuzumab was investigated. Results showed that the linkers destabilised the C_H2 and constant domains of the Fab, to a degree which is very similar to those observed for other antibody-drug conjugate scaffolds.

Although monoclonal antibodies have greatly improved cancer therapies, they can trigger

side effects due to on-target off-tumour toxicity. Recently, strategies have emerged to mask the antigen-binding site of antibodies, such that they are only activated at the tumour site. Here, the underlying mechanisms that determine what makes an effective anti-idiotypic antibody fragment mask were investigated, using three masks with different properties. Four main parameters were established as playing key roles. The efficacy of inactivation relies, first, on the extent of binding site overlap with the antigen, and second, on a relatively high association rate constant for mask and antibody. The ease of activation relies on the antibody having a lower affinity and higher dissociation rate constant for the mask than for the antigen. Fourth, the closer the affinity of the mask and the antigen for the antibody, the more disruptive the activation step needs to be.

Hydrogen-deuterium exchange mass spectrometry is a technique which probes molecular dynamics at high-resolution. To obtain high resolution, it is necessary to obtain a good peptide map. In order to understand the factors that limit the essential digest step and therefore the coverage in the peptide map, unfolding studies on an antibody were undertaken under the quench conditions used in HDX-MS experiments. Results provide confirmation that the rate of unfolding of an antibody domain limits resolution in these experiments.

ACKNOWLEDGEMENTS

I can't believe that this adventure has come to an end. Cambridge has been the most fabulous place to carry out my Ph.D.; I am very grateful to have matured in this incredibly stimulating environment, surrounded by wonderful international talents from such diverse disciplines. In these past four years, I have learned so many things and grown so much as a scientist, and this is thanks to the amazing supervisors that mentored me.

First, I would like to thank my supervisors Prof. Sophie Jackson, Dr Jonathan Phillips, Dr Nick Bond and Dr Peter Ravn.

Sophie, I don't have the words to thank you enough for all your support and help during these years. On top of being a brilliant scientist, you have been an incredible mentor. I really admire your way of supervising: you dedicate a lot of time to explaining us the theory, recommend avenues to explore, discuss the data directly with us, but also give us a lot of freedom to come up with our own ideas and research the things we are interested in. Thank you for cultivating a nurturing scientific environment, where everyone feels encouraged. I am very grateful to have had you as my supervisor in my major training as a scientist and hope we will be able to continue to collaborate.

JJ, I will always remember our first exchanges via email five years ago, when a young student from Paris contacted you regarding her interest in therapeutic antibodies and biophysical techniques, and you very kindly replied to her regarding the PhD offer you had set up and the exciting possibilities of research. Thank you for your supervision and training on the HDX-MS, for the times you hosted me at Exeter, and the numerous meetings we had remotely around exciting data. Thank you very much for having laid the foundations of the project in the proposal you imagined with Sophie.

Nick, I have been incredibly fortunate to have you as my industrial supervisor at AstraZeneca. I admire your scientific sharpness, but also your managerial skills, and how you balance both and keep track of the future plans. I really appreciated your support throughout my whole PhD, but also regarding my future career interests, it means a lot to me! Looking forward to continue collaborating.

Thank you, Peter, for your enthusiasm. Your bubbling scientific creativity is really inspiring; it was very stimulating to work with you. Thank you for having come up with the design of the second major project of my PhD, which led to a story even more exciting than expected. I have been very fortunate to work between Nick's team and yours, where I learned so much about mass spectrometry as well as antibody engineering and design. I am also very grateful for funding from the EPSRC and AstraZeneca which enabled me to undertake this research.

Thank you, Dr Matthew Edgeworth, for teaching me mass spectrometry from scratch! I am not an expert (yet!) but I definitely have a much more thorough understanding and practice now. Thanks for the good laughs and your enthusiasm for this complex class of instruments – you are a real expert!

I also wouldn't have been able to achieve so much in my second project without the help of Dr Manuela Bersellini – *grazie mille*! Thank you for teaching me how to express antibodies from scratch and for supervising me day-to-day on the masked antibody project.

This Ph.D. experience wouldn't have been the same without my two Part III students, Eli Morris in my second year and Jonny Jones in my fourth year. Thank you for your commitment and motivation: it is really exciting to teach excellent students and discuss the exciting science both of you generated. The supervision was at times challenging when I wasn't familiar enough with the techniques, especially with Eli's project, but being out of my comfort zone taught me a lot.

My passion for science started before my Ph.D. with the support from amazing mentors since high school, and during my studies at AgroParisTech in France. I remember sitting in Mr L'Helgoualc'h's biology classes in my two last years of secondary school and being absolutely fascinated by the complexity

of biological mechanisms, I always wanted to know more. *Merci M. L'Helgoualc'h pour avoir instillé la passion de la bio en moi ! Mille mercis* to Dr Claire Saulou-Bérion, for hosting me for my first internship in research at university; I remember our end-of-the-day conversations about my results, where I started discovering the intellectual gymnastics of research, reading the literature, doing some experiments, coming back to the literature… This first research experience definitely got me addicted to research! Many thanks to Dr Marine Bertin; my internship on CRISPR Cas9 under your supervision at Sanofi confirmed by interest for science and I am happy we built a nice friendship.

During my years in Cambridge, I have made amazing friends who have supported me all along. Thank you to my Peterhouse friends, Lukas, Pau, Erika, Jacopo, Peerapat, Iñigo, Eloise, Leo, Francesco, Daniel, the Freddys, Simon, Alistair, Miles, Simone, Monique, Alex, Jenny, Michael, for all the good times we shared, grad halls, events in the MCR, bike rides… Emily, I'll always cherish our gilrs' nights and good memories from the master's year at Peterhouse; Lauren, I love our heart-to-hearts; Jamie, thank you so much for your support, trust, and being an amazing friend; thank you so much David for helping us with the move! Thank you to my 40 WS housemates Sean, Olga, Theresa and Matt, it was so nice to live with friends when times were though, having a good time around nachos, thai food from the Cricketers and movie nights! Chiara, I am really grateful for our friendship that started in 290 and flourished over time, *grazie amica*! Katie, you make me laugh so much all the time; your energy and enthusiasm are incredibly contagious, and I am very glad for the friendship we got to build! (I'll be around to help you with the HDX-MS next year don't worry). Thank you to the Sensor CDT cohort, Tess, Francesco, and especially my two favourites Sammy and Elise; we've been having fun, helping each other and meeting up around our traditional Chinese takeaway since the beginning of our time in Cambridge. A very special thanks to Elise; for our laughs inside and outside the lab, for lifting me up when I was down, our endless scientific discussions until late into the night, the Cricketers dinners to relax after a long day (before pandemic started!), for all the love and support, *cette aventure n'aurait pas été pareille sans toi – merci du fond du coeur!*

Thank you to my wonderful AstraZeneca colleagues, Vassilis, Alistair, Owen, Lien, Piera, Sophie (Inman), Viv, Praveen, Giulia, Lydia, Sam, Paul, Wes, Nitin, Sara, Richard, Moni, Siobhan, James, Martin, Franco, Gareth, Guglielmo, Libby, Isabelle, Lorraine, David, Wes, Paul, Vidhi, Ed, Yasa, Jack, Kasia, Hilary, Maurizio and the DT lab. I am extremely grateful for all the collaborations and the help received with the mass spec and all the techniques I used at AZ. A special thanks to Eva, Chloé, Maria Laura and Elise, for our girls' nights and support in the lab and throughout the thesis, you are my little gems. Juhi, thanks our tea breaks, SEC-MALS discussions and so much more! Emily, you were one of the friendly faces in 290 and I am so glad I got to know you better as we worked together in the DT lab!

A big thank you to the Jackson Group for such a pleasant group to work with: Sophie, Eva, Freddy, Xinyang, Jack, Katie, Jonny, Eli – thanks for the nice atmosphere in the lab and pub trips! Many thanks as well to the Clarke group for "adopting" me at the very beginning of my PhD; a special thanks to Basile, Carol, Liza, Tristan, Quenton, Michael, Annette and Jane for your warm welcome; thanks for all your help on the practical side of the biophysical experiments, Basile, especially at the beginning of my Ph.D.! I was lucky to have friends from the department as well, Klara, Naunehal, Shianne, Oded, Rosie, Thomas. Thank you to the research technicians Kev and Naomi for running the facilities smoothly.

I would also like to thank the people at Peterhouse who have contributed to making the years so enjoyable: thank you Dr Chris Lester for your good humour, Peter and Simon for all our interactions when I was social secretary, the Master Bridget Kendall and Amanda Farnsworth for making me feel always welcome in this wonderful community. I would also like to thank my church friends here in Cambridge, thank you Amos and your family, Rebekah, Jordan, Tendai, Silvia, Verah, Rutendo and the Gospel Choir.

Last but not least, I am forever grateful for the support of my family, *Mamá, Papá y Naty*. Thank you for being my rock, thank for your support, for always having believed in me and motivated me to reach my goals, to never be afraid and always go for it. *Los quiero muchísimo.* Thank you to my friends from back home Maud, Sarita, Alice, Hélène. A very special thanks to my partner, Antoine, for all the love and

support throughout this journey, when things were going well and when things were tough; you're now almost as much of an expert as I am on HDX-MS and antibodies after carefully listening to all my science everyday! You've been a pillar during this Ph.D. and I have been so fortunate to have you by my side. Thank you God for this incredible adventure and all your blessings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ΔCp_{D-N}^{H2O}: difference in heat capacity between the native and denatured state in water

ΔG_{D-N}^{H2O} : difference in Gibbs free energy between the native and denatured state at equilibrium in water

ΔH_{D-N}^{H2O} : difference in enthalpy between the native and denatured state at equilibrium in water

ΔSASA: difference in solvent accessible surface area

2xSH: free thiol

2xCys: doubly cysteinylated

ADCC: antibody-dependent cell-mediated cytotoxicity

ADC: antibody-drug conjugate

AEW: average emission wavelength

BLI: biolayer interferometry

BSA: bovine serum albumin

CDC: complement-dependent cytotoxicity

CID: collision-induced dissociation

C_L: constant domain on the light chain

C_H1: 1st constant domains of the heavy chain

C_H2: 2nd constant domain of the heavy chain

C_H3: 3rd constant domains of the heavy chain

CHO: Chinese hamster ovarian

DAR: drug-to-antibody ratio

DPBS: Dulbecco's phosphate buffer saline

$[den]_{50\%\ D\text{-}N}$: midpoint of denaturation between the native and denatured states

ETD : electron transfer dissociation

Fc: crystallizable fragment (C_H2 and C_H3 domains)

Fab: antigen binding fragment (V_H, V_L, C_H1 and C_L domains)

FDA: US Food and Drug Administration

GdmCl: guanidinium chloride

GdmSCN: guanidinium thiocyanate

HDX-MS: hydrogen deuterium exchange mass spectrometry

HC: chain chain

iDSB: interchain disulfide bridge between the inserted cysteines after the position 239

IgG: immunoglobulin G

LC-MS: liquid chromatography mass spectrometry

LC: light chain

m_{D-N}: m-value of the unfolding transitions between the native and denature states, which is a constant of proportionality describing how much the ΔG_{X-Y} changes with denaturant concentration

mAb: monoclonal antibodies

MMP: matrix metalloproteinases

NEM: N-ethyl-maleimide

SEC-MALS: size-exclusion chromatography coupled to multi angle light scattering

scFv: single chain variable fragment

T : trastuzumab

T+DVP : trastuzumab conjugated to the DVP linkers

T+tetraDVP : trastuzumab conjugated to the tetraDVP linkers

TCEP: Tris(2-carboxyethyl)phosphine hydrochloride

TM: triple-mutation L234F/L235E/P331S in the C_H2 domain

TS: transition state

ppm: parts per million

V_L: variable domain of the light chain

V_H: variable domain of the heavy chain

WT: wild-type

1 INTRODUCTION

1.1 THERAPEUTIC ANTIBODIES

1.1.1 Monoclonal antibodies

Monoclonal antibodies have become a major class of therapeutics over the past 25 years and have proven to be promising alternatives to, and in combination with, conventional therapies for some cancers[1–3], and effective treatments for transplant rejection, autoimmune and infectious diseases to name a few[4]. The very early days of monoclonal antibodies date back to 1975, when Milstein and Köhler selectively produced monoclonal antibodies for the first time, i.e., coming from one specific clone (as opposed to polyclonal antibodies, i.e., a mixture of different antibodies which are secreted by different B cell lineages when an infection occurs). They managed to produce monoclonal antibodies by fusing myeloma and spleen cells, containing the B-cells of the clone of interest, of an immunized mouse[5]. This crucial invention, which earned them the Nobel Prize in Physiology or Medicine in 1984, opened the way to the engineering and production of monoclonal antibodies against chosen antigens[6]. Antibodies are produced for various detection techniques like ELISA assays, Western blots or pregnancy and COVID-19 antigenic tests and for many other applications, but also for therapeutic purposes. Since the 1980's, the increasing interest in antibodies as therapeutics has resulted in many advances and the production of a number of generations of antibodies over the years[6]: from murine antibodies[5] (-momab, 1975), where the genetic sequence comes entirely from the mouse, to chimeric[7,8] (-ximab, 1984), where the variable domains come from mice and the constant domains come from human genetic sequences, then humanized[9–11] (-zumab, 1986-1988), where the complementarity-determining regions, the fragments responsible for the recognition of the antigen by the variable domain, have a murine sequence, and the rest of the scaffold comes from a human genetic sequence, to finally fully human antibodies[12,13] (-mumab, 1994). Indeed, the early-generation therapeutic antibodies triggered strong immune responses against the antibody itself due to their murine origin, and did not interact very well with a patient's effector cells[14]. The vast majority of therapeutic antibodies are part of the IgG1 class.

There are a number of methods used in the development of antibodies against a specific antigen: directed engineering of antibodies can be undertaken using a transgenic mouse expressing the human genes for the antibodies, and exposing it to the antigen of interest[15]. Alternatively, phage display techniques can be used. These consist in

expressing a random combination of segments from the variable fragment in the heavy (V_H) and light (V_L) chain generated by PCR[16], to create a library with ~ 10^7 mutants[14]. These are then expressed on the surface of a phage, and clones are selected using rounds of affinity with the antigen of interest.

Monoclonal antibodies usually act with one of two modes of action: either by binding to cancer cells to act as a marker to incite the immune system to kill them (e.g. ofatumumab and alemtuzumab), or by preventing interactions between the cell and other ligands, often involved in signal silencing (trastuzumab, cetuximab, panitumumab and bevacizumab)[17]. In the former examples, antibodies are used as passive immunotherapy, and through the Fc domain, trigger the recruitment of effector cells in a process known as antibody-dependent cell-mediated cytotoxicity (ADCC), or by activation of the complement cascade leading to tumour cell lysis, called complement-dependent cytotoxicity (CDC)[17].

In April 2021, the US Food and Drug Administration (FDA) accepted the 100th therapeutic antibody[18], GlaxoSmithKline's PD1 blocker dostarlimab, thirty-five years after the first therapeutic antibody, orthoclone OKT3, to prevent kidney transplant rejection. These numbers include 80 canonical antibodies, 9 antibody-drug conjugates, 2 bispecifics, and 6 antibody fragments[18] aiming at treating pathologies ranging from cancer (the majority), cardio-vascular disorders, respiratory inflammation and autoimmunity to infectious diseases[19]. It is interesting to note that despite the large number of antibodies approved, they are directed to a rather restrained set of targets, as 42% of the approved antibodies are specific to 10 targets, counting ligands and their receptor pairs together (PD1/PDL1, CD20, TNF, HER2, CGRP/ CGRPR, VEGF/VEGFR, IL-6/IL-6R, IL-23 p19, EGFR, CD19)[18].

There are currently nearly 870 antibodies in clinical development, of which 85 ADCs and nearly 160 bispecifics and multispecifics in trials[18], according to the Antibody Society, but about 36% of these antibodies target a list of just ten receptors (PD1/PDL1, CD3, HER2, CTLA4, SARS-CoV-2, 4-1BB, LAG3, EGFr, CD20, CD47).

In 2019, antibodies accounted for 9 of the 20 top therapeutics by sales, which generated US$ 75 billion earnings that year[18]. Humira (Adalimumab), which targets TNF and is prescribed for rheumatoid arthritis, is the top selling antibody, engrossing 19.6 US$ billion in 2019[18].

1.1.2 Antibody-drug conjugates

In order to kill cancerous cells more effectively, the conjugation of drugs to antibodies has been investigated, with the objective to specifically deliver highly toxic payloads (like radioisotopes or cytotoxic drugs) to the targeted cells[20]. Their mode of action consists in recognizing the cancerous cell *via* the antibody, which then gets internalized, and the payload is then released, thereby, preventing mitotic activity leading to cell death (**Figure 1.1**).

The first antibody-drug conjugate (ADC) to be approved in 2000 by the FDA, gemtuzumab ozogamicin, Mylotarg®[21], was a humanized antibody conjugated to calicheamicin to treat acute leukemia. This drug was removed from the market in 2010 because of insufficient efficacy[22], and then reintroduced by the FDA for use with a different dosage. There are currently nine ADCs approved by the FDA[21,23–31], **Figure 1.2**, and more than 80 are currently in clinical development[32] (Clinicaltrials.gov).

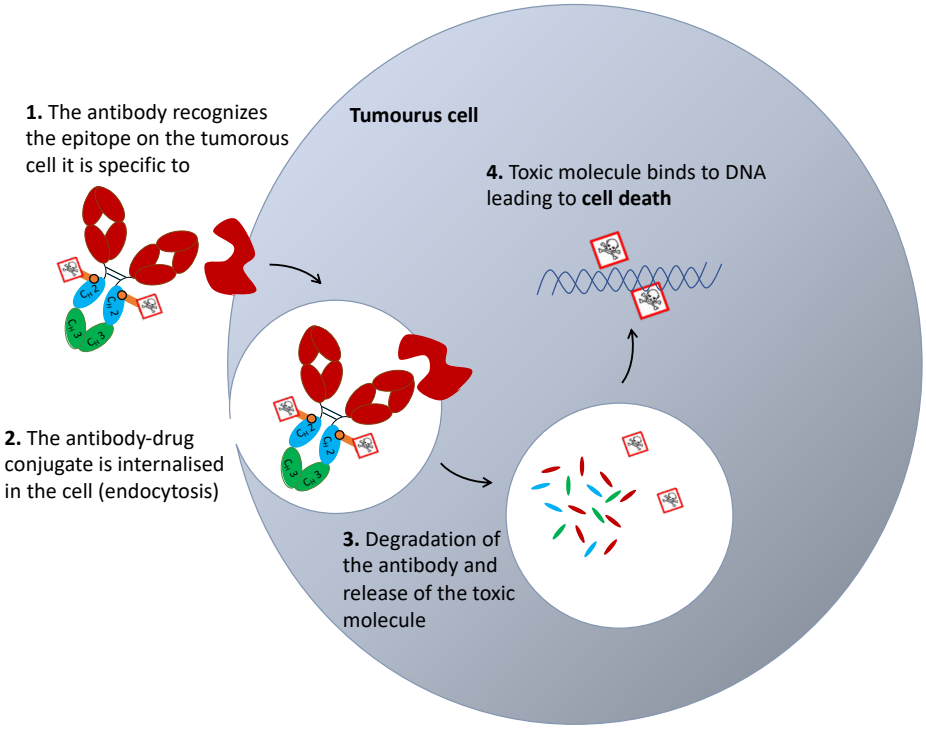


Figure 1.1: Mechanism of action of antibody-drug conjugates. The different domains of the antibody are represented in red, for the Fab domain, blue, for the C_H2 domain, and green, for the C_H3 domain. the conjugated drug is represented by the skull and crossbones. The receptor on the surface of the cell is also shown in red.

In the first generation of ADCs, the payload was conjugated to lysine side chains, which led to a distribution in the number and position of drugs attached, resulting in variable drug-to-antibody ratios (DARs)[33]. The DAR is an important contributor to the therapeutic index given it determines the dose range within which efficacy is achieved, which needs to be safe and tightly controlled. The field moved towards the conjugation of payloads to canonical cysteines[33] and then to strategies that enabled site-specific conjugation, by using point mutations to add unpaired cysteines for conjugation, either by substitution[34] or by insertion[35]. Other approaches, including the incorporation of non-natural amino acids[36–38], the use of enzymes such as formylglycine generating enzyme[39], transglutaminase[40,41] and sortase A[42], have also being investigated.

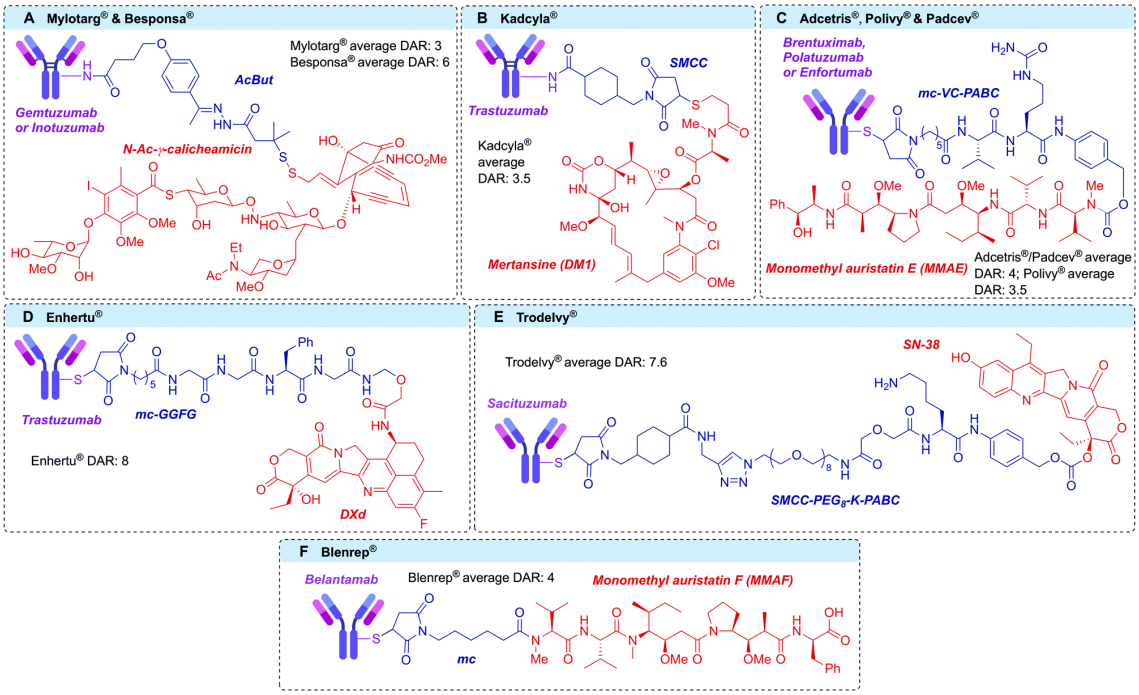


Figure 1.2: Structures of clinically-approved ADCs; the linkers are represented in blue and the payloads in red. **A.** Gemtuzumab ozogamicin (Mylotarg®, 2000)[21] and inotuzumab ozogamicin (Besponsa®, 2017)[24]. **B.** Trastuzumab emtansine (Kadcyla®, 2013)[23]. **C.** Brentuximab vedotin (Adcetris®, 2011)[25,31]; polatuzumab vedotin (Polivy®, 2019)[26] and enfortumab vedotin (Padcev®, 2019)[27]. **D.** trastuzumab deruxtecan (Enhertu®, 2019)[28]. **E.** Sacituzumab govitecan (Trodelvy®, 2020)[29]. **F.** Belantamab mafodotin (Blenrep®, 2020)[30]. AcBut = 4-(4-acetylphenoxy)butanoic acid, SMCC = succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate, MC = maleimidocaproyl, PABC = *p*-aminobenzyloxycarbonyl. Reproduced with permission from the journal *Chemical Society Reviews* (Walsh *et al.*, 2021[43])

In the next decade, the challenges are to develop new kinds of payloads with different activities, as today most belong to two families of antimitotic agents, auristatins and maytansinoids[44], which reflects the difficulty of finding cytotoxic molecules that are suitable as payloads of ADCs. The length of the linker is also a subject of investigation, as, in some cases, it needs to be stable while the antibody carries the drug in the

bloodstream, but it needs to be able to release it once in the cell. Conjugation on cysteines or lysines with malemeide-type linkers have shown to have limited stability in the blood as the thiosuccinimide adduct formed can deconjugate *via* a retro-Michael reaction[45] in the serum leading to off-target cytotoxicity[46]. Some novel approaches with very stable linkers have been developed[47–50] which ensure the cytotoxic payload is only released in the organ of interest.

As the objective of ADCs, ultimately, is to be administered to patients, the scaffold of the antibody-drug conjugates needs to be as stable as possible to retain its properties. As substitution or insertion of a cysteine residue and/or conjugation of the payload might lead to a destabilization of the native state, it is important to identify mutants which destabilize the antibody to a minimal extent.

1.1.3 Masked antibodies

Therapeutic antibodies are engineered to recognize very specifically the target of interest. However, antigens that are overexpressed in organs affected by a disease are also expressed at lower levels in healthy tissues. As the antibodies bind to their specific antigen wherever it is expressed in the body, they can cause side effects when they bind to healthy tissues[4,51–54]. These side effects can be diverse: immune reactions, including acute anaphylactic (IgE-mediated) and anaphylactoid reactions against the mAb (cetuximab[55], omalizumab[56–58], rituximab, natalizumab[59]), infections, thrombotic disorders (infliximab[60], efalizumab[60], rituximab[60], alemtuzumab[61], bevacizumab[62]), autoimmune diseases including lupus-like syndrome[63], thyroid disease[64] and autoimmune colitis[65,66] (alemtuzumab, ipilimumab), dermatitis[67] (cetuximab, panitumumab) cardiotoxicity[68,69] (trastuzumab). One of the challenges in the development of therapeutic antibodies is to find a target that is overexpressed in the disease tissue and almost not at all in healthy ones. However, despite considerable effort in this area, only a few antigens meet these criteria[4,17].

A possible strategy to control more tightly the activity of therapeutic antibodies consists of masking the antibody with another molecule, preventing it from recognizing and binding to the antigen, and removing the mask when the antibody is in the tissue of interest, relying on proteases that are overexpressed in tumours, for example. The idea of masking active compounds and activating them with proteases has been around for several years, and it started on small molecules[70]. The so-called “prodrugs” are

derivatives of drug molecules that are enzymatically or chemically activated *in vivo* to release the active parent drug[70]. Also using a protease digestion system, the very first engineered masked antibody-like fragment, published in 2009 by Donaldson *et al.*[71], consisted in a scFv masked by its own corresponding epitope. Erster *et al.*[72] from the University of California, Santa Barbara, then published in 2012 a format of masked antibody, or "proantibody", to detect protease activity *in vivo.* Their construct consists of a murine scFv-Fc fusion (instead of using a whole Fab) that targets vascular cell adhesion molecule 1 (VCAM-1), and a peptide masks the variable regions of the scFv until the linker is cleaved by matrix metalloproteinases (MMPs). With their setup, they managed to observe the activity of MMPs in aortic plaques in ApoE mice. Following that first construct, the first therapeutic masked antibody, referred to as "probody®", was successfully developed by Desnoyers *et al.*, 2013[73] from CytomX Therapeutic Inc. in collaboration with the University of California, Santa Barbara, and consisted in masking an antibody that targets the epidermal growth factor receptor (EGFR), cetuximab, with a peptide and activating it by proteolysis of the linker between the mask and the antibody using highly location-specific enzymes. Their results showed that the on-target off-tumour antibody toxicities had been considerably reduced, and that the probody was effectively activated in the region of interest. This first paper opened the path to masked antibody formats, which have been successfully developed in the past decade for several FDA-approved antibodies with known side effects[74–76]. Two main categories of masks have been developed so far: first, masks with affinity for the antibody-binding domains, including peptides [72,73,77] and antibody fragments[78–81], and second, non-binding masks relying on steric hinderance to block the antigen-binding site. This second category includes coiled-coil domains[82], protein M[83], peptide-DNA assemblies[84–86], and a hinge[23] region. Several different methods have been developed to restore the activity of the masked antibody once it is at its target site. The most widely developed approach has been activation by hydrolysis with enzymes that are overexpressed in the tumour or organ of interest [71,72,89,90,77,79–83,87,88]. Alternative options have also been investigated, such as photo-cleavable linkers[84], pH-dependent peptide-DNA locks[85] or pH-sensitive masking peptides[91]. Further opportunities remain, for example, to create a tuneable masking cap by specifically engineering its stability so it can be destabilised locally by adding energy.

1.1.4 Structure of antibodies

1.1.4.1 Domain structure

Therapeutic monoclonal antibodies belong to the class of immunoglobulin G proteins (IgG). These are large proteins of approximately 150 kDa, with two heavy chains, composed of four domains V_H, C_H1, C_H2, C_H3, and two light chains composed of the domains V_L and C_L (**Figure 1.3**). The two light chains are linked to the heavy chains by a disulfide bridge between the C_H1 and C_L, domains (**Figure 1.3 C**) and the two heavy chains are linked to each other by two disulfide bridges in the hinge region.

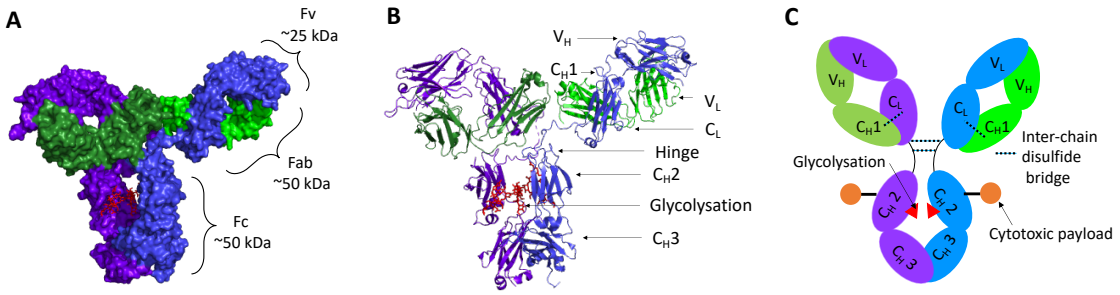


Figure 1.3: Structure of an IgG1. A. and **B.** Crystal structure of the intact human IgG B12 with broad and potent activity against primary HIV-1 isolates, coordinates from 1HZH[92] in the Protein DataBank (PDB). **C.** Schematic showing the different variable and constant domains. Purple and blue: heavy chains; light and dark green: light chains; red: glycosylation

The Fab fragment, composed of the V_H, C_H1, V_L and C_L domains does not interact with the Fc region composed of the C_H2 and C_H3 domains, due to the flexible hinge. Therefore, it is quite difficult to crystallize full-length IgGs and get an X-ray diffraction structure, so the Fab and Fc fragments are often crystallized separately.

The variable domains V_H and V_L contain three regions whose variability confers the unique affinity of the antibody for a specific antigen. These hypervariable regions, also called complementarity-determining regions, are referred to as CDR1, CDR2 and CDR3. The rest of the variable domain is constant.

1.1.5 Tertiary structure of the individual domains

IgGs are composed of only β-sheets. Specifically, each domain has a β-sandwich structure, similar to a Greek key barrel[93]. All constant domains (C_L, C_H1, C_H2 and C_H3) have the same characteristic arrangement of seven β-strands: four and three β-strands each form a β-sheet that packs together due to the interactions of hydrophobic side chains in the core; the interactions between residues are invariant within the different constant domains, and linked by the disulfide bridge between strand F and strand B[93] (**Figure 1.4 A**). The Greek key arrangement of the β-strands is shown in **Figure 1.4 B**. The loops between the β-strands often contain prolines [94].

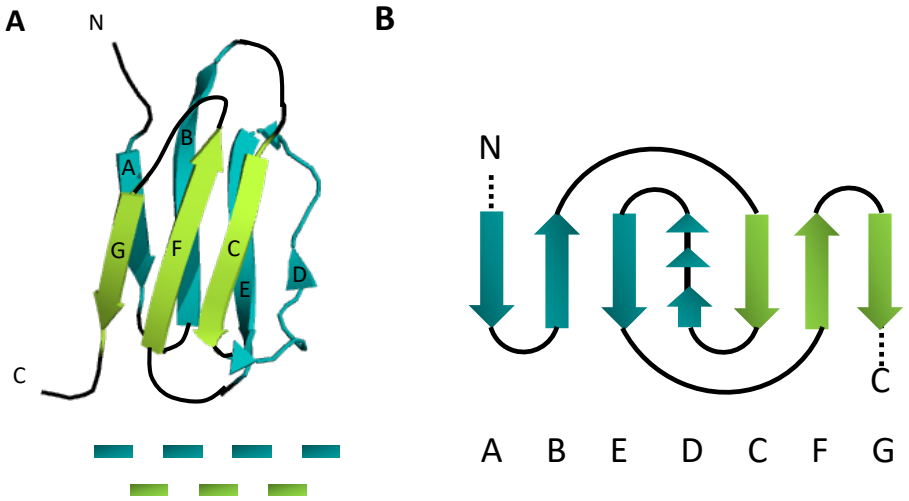


Figure 1.4: β-sandwich structure of the constant domain of an IgG. A. Cartoon adapted from the crystal structure of the C_H2 domain from the fucosylated Fc Fragment from human immunoglobulin G1, with the typical arrangement of constant domain with the four blue β-strands and three green β-strands, using the coordinates in 3AVE in the PDB. **B.** Detailed view of the connectivity between the β-strands showing the Greek key arrangement.

The variable domains are composed of nine β-strands. The two additional β-strands are located in the loop between β-strands C and D, and are noted as C' and C'' (**Figure 1.5**). The loops in the variable domain have important functions, as that is where the hypervariable domains CDR1, CDR2 and CDR3 are located, more specifically between residues 24-34, 50-56, and 89-97 in the light chain and 31-35, 50-65 and 95-102 in the heavy chain (**Figure 1.5**).[93]

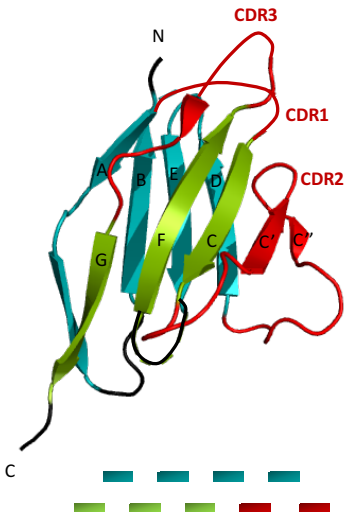


Figure 1.5: β-sandwich structure typical of the constant and variable domains of an IgG. Cartoon adapted from the crystal structure of the VL domain, using the coordinates from 4HJG in the PDB. The regions in red represent the additional β-strands not present in the constant domain and where the hypervariable domains are located. The rectangles at the bottom represent the arrangement of the two β-sheets: the first β-sheet is shown with four blue strands, and the second β-sheet is shown with three green strands with two additional β-strands shown in red.

Although the V_H and C_H1, and V_L and C_L domains are on the same chain, there are few interactions between these domains. However, there are interactions between the variable domains (V_H and V_L) and between the constant domains (C_H1 and C_L). The constant domains C_H1, C_L, and C_H3 associate by interaction with the four-stranded β-sheets of both domains at a right angle one to the other, whilst the three-stranded β-sheets face the exterior (**Figure 1.6 A**). The C_H2 domains interact only through glycosylation, which stabilizes them.[94,95,96]

The variable domains, however, interact differently. The domains face each other via the sheet with five β-strands, i.e., that has the two additional β-strands shown in red. This conformation brings the loops closer to where the CDR regions are, **Figure 1.6 B**. The orientation of the β-strands is relatively parallel to each other. The two five-stranded β-sheets do not pack tightly flatly against each other but are curved, forming therefore a barrel with four of the five β-strands of each sheet (**Figure 1.6 B**).

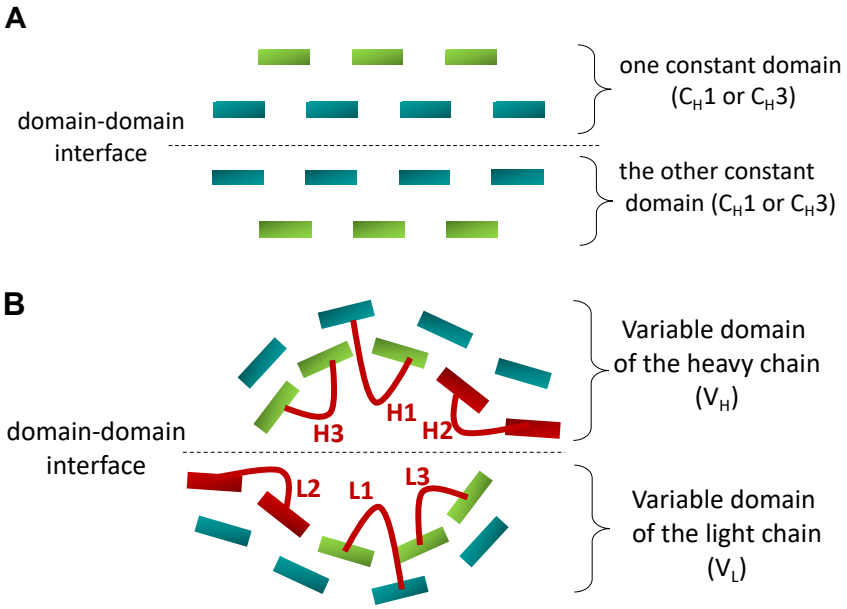


Figure 1.6: Schematic representation of the interaction of two constant and two variable domains. A. The constant domains interact by the two four-stranded β-sheets forming an interface. **B.** The variable domains interact via the five-stranded β-sheet, bringing the CDR regions closer together (L1, L2, L3: CDR1 to 3 on the light chain; H1, H2, H3: CDR1 to 3 on the heavy chain).

1.1.6 N-linked glycosylation

Today, therapeutic antibodies are mainly expressed in Chinese hamster ovarian (CHO) cells, as they produce the same glycosylation patterns as those in produced in human antibodies, i.e., predominantly G0F. Thus, this should ensure the lowest immune response against the therapeutic antibody. Some mutations in the C_H2 domain can lead to other glycosylation patterns, for example, G2SF2 was observed in the mutant Fc-C239i.

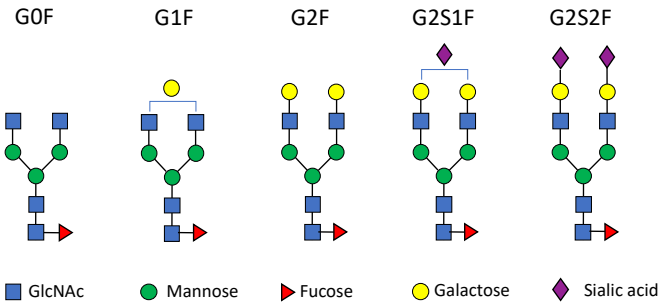


Figure 1.7: Major N-linked glycosylation patterns observed on therapeutic antibodies

1.2 BIOPHYSICAL CHARACTERISATION

1.2.1 Stability of the native state of proteins

Proteins form an incredibly modular class of biological molecules that accomplish a vast range of functions in living cells. They can have very different secondary, tertiary, and quaternary structures, and they almost always fold to the lowest energy state when they are in native conditions, the denatured state having the highest energy level under the same conditions (unless they are intrinsically disordered proteins, which are not discussed in this Thesis). The native states of proteins have different types of stability. The thermodynamic stability is a measure of the difference in Gibb's free energy between the native and denatured state at equilibrium in water (ΔG_{D-N}^{H2O}), consisting of a favourable enthalpic component (ΔH_{D-N}^{H2O}), and unfavourable entropic component (ΔS_{D-N}^{H2O}). The thermal stability, measured by the melting temperature T_{m}, also relates to the thermodynamic stability and, as well as depending upon the difference in enthalpy and entropy, also depends upon the difference in heat capacity between the native and the denatured states (ΔCp_{D-N}^{H2O}). The change in enthalpy between native and denatured states, is the difference between the sum of all the interactions that each atom makes with others within the protein and with the solvent in the native and denatured states. The native state has a lot of favorable interactions, which come from van der Waals', electrostatic, π-π, cation-π interactions and hydrogen bonds; therefore, the native state of a protein is enthalpically favorable compared to the denatured state. The change in entropy on folding has two major contributions. The first comes from the inherent flexibility of the backbone and side chains in the denatured state leading to an ensemble of states which are. close in energy but have different conformations. This large configurational entropy is mainly lost in the native state. The second entropic contribution comes from the solvent water molecules which are associated with hydrophobic residues in the denatured state and form what are sometimes called mini"icebergs" leading to a reduction in entropy. On folding, these water molecules are expelled into the bulk water increasing the overall entropy of the system. Overall, entropy decreases upon folding, i.e. folding is entropically unfavorable, mainly because of the loss of the configurational entropy of the polypeptide chain, but compensated for to some extent by the gain in solvent entropy when hydrophobic side chains are buried in the hydrophobic core of the protein. Without this favourable contribution from the solvent molecules, folding would not be possible, therefore, water (and the hydrophobic effect) plays a big role in the energetics of folding.

The kinetic stability, measured by the unfolding rate constant, informs on the difference in energy between native and transition states. Molecular dynamics in the native state is also an important measure of local stability and provides information on both local and global unfolding events.

1.2.2 Denaturation of proteins

To assess the stability of a protein, a measurable equilibrium between the native and denatured states needs to be established. Indeed, under native conditions, for a stable protein, there is only a very small proportion of denatured molecules (0.001% of the C_H2 domain, 0.0000001 % of the C_H3 domain and 0.0000001 % of the Fab fragment are denatured under native conditions, details of the calculations in Appendix IV) and therefore it is impossible to measure that population against a background of almost 100% of native proteins, unless using single-molecule techniques. It is therefore necessary to change the population of native (N) and denatured (D) states requiring some form of denaturation. Several factors favor the denaturation of proteins.

First, proteins can be unfolded by using a chaotropic agent. The chemical denaturants mainly used are urea, guanidinium chloride (GdmCl) and guanidinium thiocyanate (GdmSCN), which are shown in increasing order of denaturing strength[97–100]. The mechanism of action of such denaturants has been debated since the 1960s; although it is not exactly clear what the mechanism(s) are, there are two major models, the preferential binding and the preferential solvation models. In any case, it is empirically known that the chemical denaturants increase the solubility of both polar and non-polar side chains and the protein backbone, thus stabilising the denatured state compared to native state. As the chemical denaturant concentration increases, the energy of both the native and denatured state decrease. The denatured state lowers in energy more rapidly with increasing denaturant concentration than the native state as it has a larger solvent accessible surface area (ΔSASA, **Figure 1.8**).

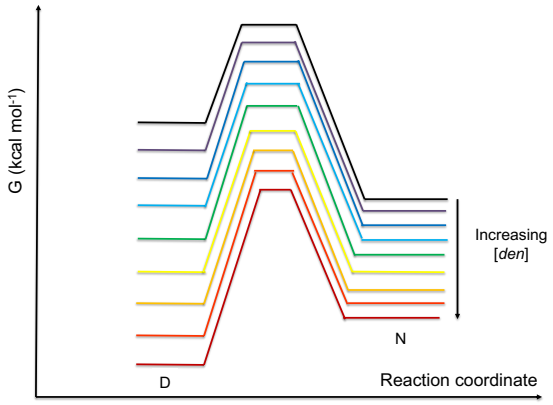


Figure 1.8: Free energy diagram illustrating how the energies of denatured and native states change with increasing concentration of chemical denaturant.

Second, high and low temperatures can unfold proteins when the unfavourable entropic component to the free energy dominates. However, thermal unfolding frequently leads to aggregation, due to the exposure of hydrophobic side chains.

Third, extreme pHs, acidic and basic, can also favour unfolding. Under acidic conditions, some side chains, like Asp and Glu, become protonated. This increases the overall net positive charge on the surface of the native state of the protein creating electrostatic repulsion. The result is unfolding to a state where residues are, on average, further apart than in the native state. There is also another factor driving the unfolding at low pH: considering an equilibrium between the native and denatured states at pH 4.5 and 7, $\Delta G_{D-N}{}^{pH7} > \Delta G_{D-N}{}^{pH4.5}$ because some favorable interactions, such as salt bridges and hydrogen bonds, are lost at the lower pH. As a consequence of these favourbale interactions in the native state, the pK_a of Asp and Glu side chains are higher in the denatured state. Thus, the negatively charged side chains of Asp and Glu protonate at higher pH values in the denatured state than in the native state, so the equilibrium is pushed towards the protonated denatured state. However, the force to drive the protonated state towards the protonated native is not that strong, as the $\Delta G_{D-N}{}^{pH4.5}$ is lower than at pH 7. Therefore, these states drive the equilibrium towards the protonated denatured state (this is acid-induced unfolding).

The equilibrium between native and denatured states, and thus the thermodynamic stability, is measured by probing the difference in fluorescence emission of tryptophan side chains, as the protein is incubated with increasing concentrations of

chemical denaturant. Upon excitation at 280 nm (maximum of absorption[101]), the maximal emission intensity is around 335 nm for buried tryptophan residues, whereas it is approx. 360 nm in the denatured state[101]. This is due to the fact that the excited state is lowered in energy in a polar environment compared to an apolar one. Thus, the difference in energy between ground and excited state is greater when the tryptophan is buried in a hydrophobic environment, as the wavelength is inversely related to the difference in energy ($E = h.\nu$ where h is the Plank constant, ν is the frequence, such that $\nu = \frac{1}{\lambda}$ where λ is the wavelength).

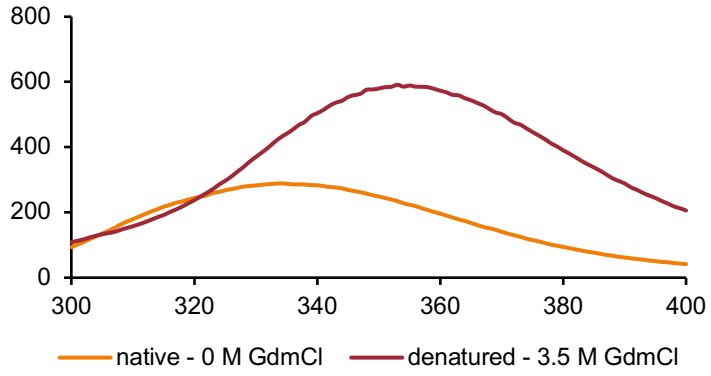


Figure 1.9: Typical fluorescence spectra for tryptophan in native (orange) and denatured (brown) states.

The difference in signal between the native and denatured state can also be measured by probing the ellipticity using far UV-CD data.

1.2.2.1 Unfolding following a two-state model

If a protein unfolds following a two-state model, only the native and denatured states are significantly populated:

$$N \overset{K_{D-N}}{\rightleftharpoons} D \qquad \textbf{Eq. 1.1}$$

And where:

$$K_{D-N} = \frac{[D]}{[N]} \qquad \textbf{Eq. 1.2}$$

In a two-state model, the intensity of fluorescence, F, **Figure 1.10**, at a certain concentration of denaturant can be expressed with respect to the denatured fraction, $f(D)$,

as follows:

$$\boldsymbol{f(D) = \frac{F - native\ baseline}{denatured\ baseline - native\ baseline}} \qquad \textbf{Eq. 1.3}$$

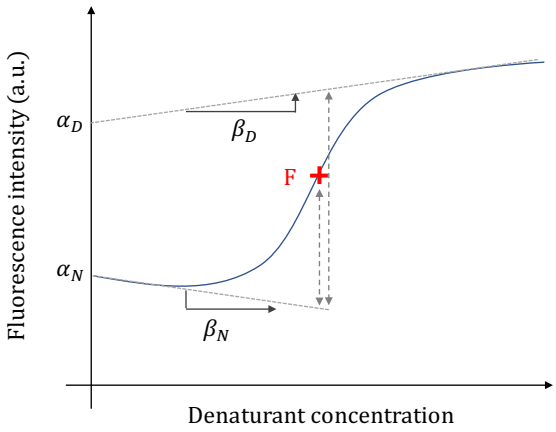


Figure 1.10: Denaturation curve for a protein that unfolds populating just two states. α_N is the intercept of the native baseline; β_N is the slope of the native baseline; α_D is the intercept of the denatured baseline and β_D is the slope of the denatured baseline.

Therefore:

$$\boldsymbol{f(D) = \frac{F - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}} \qquad \textbf{Eq. 1.4}$$

The equilibrium constant between the native and denatured state, K_{D-N}, is expressed as:

$$\boldsymbol{K_{D-N} = \frac{[D]}{[N]} = \frac{f(D)}{1 - f(D)}} \qquad \textbf{Eq. 1.5}$$

$$\big(1 - f(D)\big).K_{D-N} = f(D) \qquad \textbf{Eq. 1.6}$$

And,

$$\boldsymbol{\Delta G_{D-N}^{[den]} = -RT \ln(K_{D-N})} \qquad \textbf{Eq. 1.7}$$

Therefore:

$$\boldsymbol{K_{D-N} = \exp(\frac{-\Delta G_{D-N}^{[den]}}{RT})}$$ **Eq. 1.8**

Therefore, **Equation 1.6** becomes:

$$\boldsymbol{\left(1 - f(D)\right).\exp(\frac{-\Delta G_{D-N}^{[den]}}{RT}) = f(D)}$$ **Eq. 1.9**

Adding the definition of $f(D)$ from **Equation 1.4**, one obtains:

$$\boldsymbol{\left(1 - \frac{F - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}\right).\exp(\frac{-\Delta G_{D-N}^{[den]}}{RT}) = \frac{F - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}}$$ **Eq. 1.10**

$$\left(\frac{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den]) - F + (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}\right).\exp(\frac{-\Delta G_{D-N}^{[den]}}{RT}) = \frac{F - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}$$ **Eq. 1.11**

$$\left(\frac{(\alpha_D + \beta_D.[den]) - F}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}\right).\exp(\frac{-\Delta G_{D-N}^{[den]}}{RT}) = \frac{F - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])}$$ **Eq. 1.12**

$$\left((\alpha_D + \beta_D.[den]) - F\right).\exp(\frac{-\Delta G_{D-N}^{[den]}}{RT}) = F - (\alpha_N + \beta_N.[den])$$ **Eq. 1.13**

$$(\alpha_D + \beta_D.[den]).\exp\left(\frac{-\Delta G_{D-N}^{[den]}}{RT}\right) + (\alpha_N + \beta_N.[den]) = F(1 + \exp\left(\frac{-\Delta G_{D-N}^{[den]}}{RT}\right))$$ **Eq. 1.14**

$$F = \frac{(\alpha_N + \beta_N.[den]) + (\alpha_D + \beta_D.[den]).\exp\left(\frac{-\Delta G_{D-N}^{[den]}}{RT}\right)}{1 + \exp\left(\frac{-\Delta G_{D-N}^{[den]}}{RT}\right)}$$ **Eq. 1.15**

It has been empirically demonstrated that the free energy of unfolding of a protein is linearly dependent on the concentration of denaturant[102,103]. Therefore, the experimental determination of ΔG_{D-N}^{H2O} using chemical denaturation typically involves linear extrapolation of the empirically observed free energy changes in the transition region to the limit of zero molar denaturant concentration[97] (**Equation 1.16, Figure 1.11).**

$$\boldsymbol{\Delta G_{D-N}^{[den]} = \Delta G_{D-N}^{H2O} - m_{D-N}.[den]}$$ **Eq. 1.16**

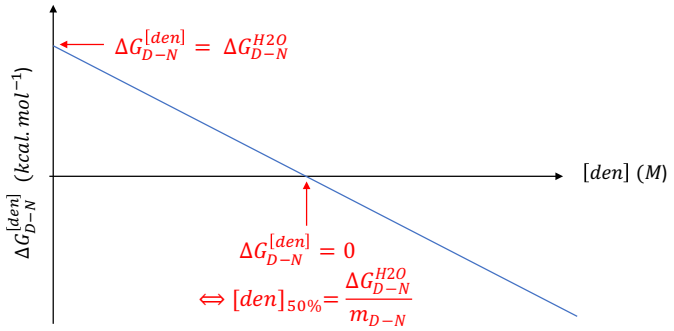


Figure 1.11: Graphical representation of the change in free energy of unfolding, $\Delta G_{\mathrm{D-N}}^{[den]}$, with denaturant concentration

The factor m_{D-N} depends on the strength of the denaturant (the stronger the denaturant, the higher the coefficient), and the change in solvent accessibility surface area between native and denatured states (the more buried regions, the larger the coefficient). Substituting **Equation 1.16** into **1.15** one obtains:

$$F = \frac{(\alpha_N + \beta_N.[den]) + (\alpha_D + \beta_D.[den]).\exp\left(\frac{m_{D-N}.[den] - \Delta G_{D-N}^{H2O}}{RT}\right)}{1 + \exp\left(\frac{m_{D-N}.[den] - \Delta G_{D-N}^{H2O}}{RT}\right)} \qquad \textbf{Eq. 1.17}$$

Equation 1.17 is used to fit the experimental data points of denaturation curves to a two-state model, to obtain the values of m_{D-N} and ΔG_{D-N}^{H2O}.

At a certain denaturation concentration, called the midpoint of denaturation, $[den]_{50\%}$, the difference in Gibbs free energy between the native and denatured state is zero ($\Delta G_{\mathrm{D-N}}^{[den]_{50\%}} = 0$, **Figure 1.11**). Therefore, according to **Equation 1.16**, $\Delta G_{D-N}^{H2O} = m_{D-N}.[den]_{50\%}$ i.e. $[den]_{50\%} = \frac{\Delta G_{D-N}^{H2O}}{m_{D-N}}$. It is therefore possible to write **Equation 1.17** in terms of $[den]_{50\%}$, as follows:

$$F = \frac{(\alpha_N + \beta_N.[den]) + (\alpha_D + \beta_D.[den]).\exp\left(\frac{m_{D-N}}{RT}.([den] - [den]_{50\%})\right)}{1 + \exp\left(\frac{m_{D-N}}{RT}.([den] - [den]_{50\%})\right)} \qquad \textbf{Eq. 1.18}$$

1.2.2.2 Unfolding following a three-state model

A three-state model has an intermediate between the native and denatured state stable enough to be populated and therefore observed. The equilibrium between the

different species is as follows:

$$\mathrm{N} \underset{\boldsymbol{K_{I-N}}}{\rightleftharpoons} \mathrm{I} \underset{\boldsymbol{K_{D-I}}}{\rightleftharpoons} \mathrm{D} \qquad \textbf{Eq. 1.19}$$

$$K_{I-N} = \frac{[I]}{[N]} \qquad \textbf{Eq. 1.20}$$

$$K_{D-I} = \frac{[D]}{[I]} \qquad \textbf{Eq. 1.21}$$

$$K_{D-N} = \frac{[D]}{[N]} = K_{I-N}.K_{D-I} \qquad \textbf{Eq. 1.22}$$

From Equations **1.20** and **1.22**, $[D]$ and $[I]$ can be expressed in terms of $[N]$:

$$[\boldsymbol{I}] = \boldsymbol{K_{I-N}}.[\boldsymbol{N}] \qquad \textbf{Eq. 1.23}$$

$$[D] = K_{I-N}.K_{D-I}.[N] \qquad \textbf{Eq. 1.24}$$

The fluorescence at a certain denaturation concentration can be expressed as the sum of the fractions weighted by the signal of each state, as follows:

$$\boldsymbol{F = Y_N f(N) + Y_I f(I) + Y_D f(D)} \qquad \textbf{Eq. 1.25}$$

$f(N)$ can be rewritten as:

$$\boldsymbol{f(N) = \frac{[N]}{[N] + [I] + [D]}} \qquad \textbf{Eq. 1.26}$$

Substituting **Equations 1.23** and **1.24** into **1.26**:

$$\boldsymbol{f(N) = \frac{[N]}{[N] + K_{I-N}.[N] + K_{I-N}.K_{D-I}.[N]}} \qquad \textbf{Eq. 1.27}$$

$$f(N) = \frac{1}{1 + K_{I-N} + K_{I-N}.K_{D-I}} \qquad \textbf{Eq. 1.28}$$

Likewise, for $f(I)$ and $f(D)$:

$$\boldsymbol{f(I) = \frac{K_{I-N}.[N]}{[N] + K_{I-N}.[N] + K_{I-N}.K_{D-I}.[N]}} \qquad \textbf{Eq. 1.29}$$

$$f(I) = \frac{K_{I-N}}{1 + K_{I-N} + K_{I-N}.K_{D-I}} \quad \textbf{Eq. 1.30}$$

$$f(D) = \frac{K_{I-N}.K_{D-I}.[N]}{[N] + K_{I-N}.[N] + K_{I-N}.K_{D-I}.[N]} \quad \textbf{Eq. 1.31}$$

$$f(D) = \frac{K_{I-N}.K_{D-I}}{1 + K_{I-N} + K_{I-N}.K_{D-I}} \quad \textbf{Eq. 1.32}$$

Therefore, substituting **Equations 1.28**, **1.30**, **1.32** into **1.25**:

$$\boldsymbol{F = \frac{Y_N + Y_I.K_{I-N} + Y_D K_{I-N}.K_{D-I}}{1 + K_{I-N} + K_{I-N}.K_{D-I}}} \quad \textbf{Eq. 1.33}$$

And given that:

$$\boldsymbol{Y_N = \alpha_N + \beta_N.[den]} \quad \textbf{Eq. 1.34}$$

$$Y_I = \alpha_I + \beta_I.[den] \quad \textbf{Eq. 1.35}$$

$$Y_D = \alpha_D + \beta_D.[den] \quad \textbf{Eq. 1.36}$$

F therefore becomes:

$$\boldsymbol{F = \frac{\alpha_N + \beta_N.[den] + (\alpha_I + \beta_I.[den]).K_{I-N} + (\alpha_D + \beta_D.[den])K_{I-N}.K_{D-I}}{1 + K_{I-N} + K_{I-N}.K_{D-I}}} \quad \textbf{Eq. 1.37}$$

And given that:

$$\boldsymbol{\Delta G_{I-N}^{[den]} = -RT \ln(K_{I-N})} \quad \textbf{Eq. 1.38}$$

$$K_{I-N} = \exp(\frac{-\Delta G_{I-N}^{[den]}}{RT}) \quad \textbf{Eq. 1.39}$$

$$K_{I-N} = \exp(\frac{m_{I-N}.[den] - \Delta G_{I-N}^{H2O}}{RT}) \quad \textbf{Eq. 1.40}$$

Likewise,

$$\boldsymbol{K_{D-I} = \exp(\frac{m_{D-I}.[den] - \Delta G_{D-I}^{H2O}}{RT})} \quad \textbf{Eq. 1.41}$$

Equation 1.37 becomes (obtaining $m_{D-I,}$ $m_{I-N,}$ and ΔG_{D-I}^{H2O}, ΔG_{I-N}^{H2O} fittings):

$$F = \frac{\alpha_N + \beta_N.[den] + (\alpha_I + \beta_I.[den]).\exp\left(\frac{m_{I-N}.[den] - \Delta G^{H2O}_{I-N}}{RT}\right) + (\alpha_D + \beta_D.[den])\ \exp\left(\frac{m_{I-N}.[den] - \Delta G^{H2O}_{I-N}}{RT}\right)\exp(\frac{m_{D-I}.[den] - \Delta G^{H2O}_{D-I}}{RT})}{1 + \exp\left(\frac{m_{I-N}.[den] - \Delta G^{H2O}_{I-N}}{RT}\right) + \exp\left(\frac{m_{I-N}.[den] - \Delta G^{H2O}_{I-N}}{RT}\right)\exp(\frac{m_{D-I}.[den] - \Delta G^{H2O}_{D-I}}{RT})} \quad \textbf{Eq. 1.42}$$

Equation 1.42 in terms of $[den]_{50\%}$, is expressed as follows:

$$F = \frac{\alpha_N + \beta_N.[den] + \alpha_I.\exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right) + (\alpha_D + \beta_D.[den])\ \exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right)\exp\left(\frac{m_{D-I}}{RT}.([den] - [den]_{50\%\,D-I})\right)}{1 + \exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right) + \exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right)\exp\left(\frac{m_{D-I}}{RT}.([den] - [den]_{50\%\,D-I})\right)} \quad \textbf{Eq. 1.43}$$

1.2.3 Kinetic stability

Kinetic data from unfolding or folding experiments can be obtained using rapid-mixing techniques and a stopped-flow spectrometer which measures the change in fluorescence signal on unfolding or folding. Kinetic data are fit to one or more exponential functions. Considering a system with only one kinetic phase, which is a simple first-order reaction:

$$N \overset{k_U}{\rightarrow} D \quad \textbf{Eq. 1.44}$$

Where the unfolding rate can be defined as the rate of increase in D as follows:

$$\frac{d[D]}{dt} = k_U.[N] \quad \textbf{Eq. 1.45}$$

Or the rate of decrease of N state:

$$\frac{d[N]}{dt} = -k_U.[N] \quad \textbf{Eq. 1.46}$$

Solving the differential equation, one obtains:

$$[N]_t = [N]_0.\exp(-k_U.t) \quad \textbf{Eq. 1.47}$$

Therefore, the kinetic data can be fitted with a single exponential, as follows:

$$A(t) = (A_{end} - A_0).(1 - \exp(-kt)) + A_0 \quad \textbf{Eq. 1.48}$$

Where $A(t)$ is the fluorescence signal at any time, t, $A_{end} - A_0$ is the amplitude (i.e., the change in fluorescence signal, k the unfolding rate constant and t the time. If the data shows several kinetic phases, the experimental data can be fit to a double or triple exponential, i.e., the sum of several exponentials, as follows:

$$A(t) = A_1.(1 - \exp(-k_1 t)) + A_2.(1 - \exp(-k_2 t)) + A_3.(1 - \exp(-k_3 t)) + c \qquad \textbf{Eq. 1.49}$$

where A_1, A_2 and A_3 are the amplitudes of the different phases, k_1, k_2 and k_3 the respective kinetic rate constants and c the offset. Fitting the experimental data with this equation provides the values of the amplitudes and the kinetic rate constants for each phase. The kinetics of unfolding are measured at different concentrations of denaturant, and the data transformed to create a plot in which the natural logarithm of the kinetic rate constants are plotted against the final denaturant concentration. These data can then be fit to a simple linear function. The y-axis intercept is $ln\, k_U^{H20}$ from which k_U^{H20} can be calculated, the unfolding kinetic rate constant in water, whilst the slope is m_{kU}

1.2.4 Relating thermodynamic and kinetic data

For mutants of proteins which fold with simple two-state kinetics, there can be a correlation between thermodynamic and kinetic stability: usually when the thermodynamic stability decreases, the kinetic stability decreases as well and inversely; for some mutations, the stabilities are not correlated. For example, a mutation may destabilize the protein but may not affect the unfolding kinetics. The energy diagram shown in **Figure 1.12** illustrates the thermodynamic and kinetic information that can be obtained from different techniques. This representation assumes that there is no change in the energy of the denatured state, which is generally the case; however, the free energy of the denatured state can change significantly though if disulfide bridges are either removed or introduced through a change in entropy.

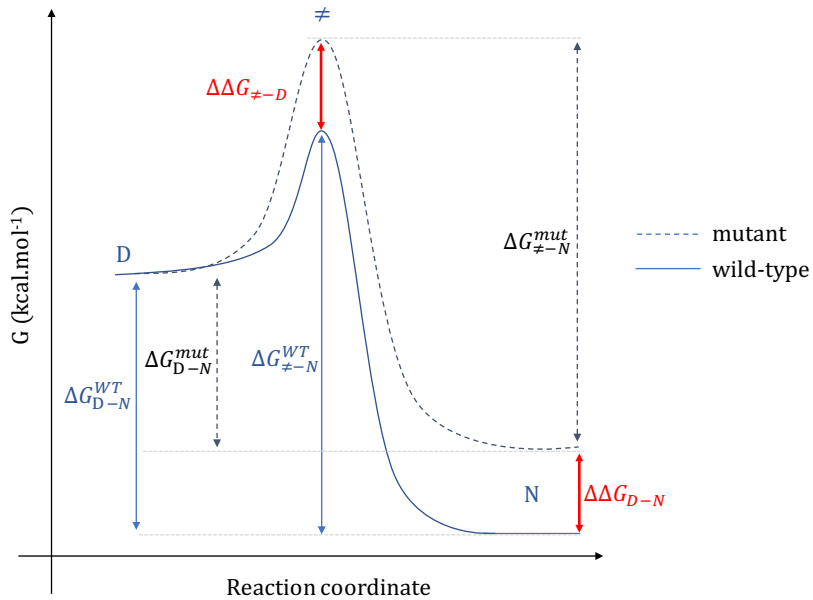


Figure 1.12: Free energy diagram for the unfolding of a two-state system. D: denatured state; N: native state; TS: transition state.

The Arrhenius law relates the unfolding kinetic rate constant (k_U) with the difference in Gibbs free energy between the transition state and the native state (ΔG_{TS-N}^{H2O}), as follows:

$$\boldsymbol{k_U = A.\exp\left(\frac{-Ea}{RT}\right)} \quad \textbf{Eq. 1.50}$$

$$k_U = A.\exp\left(\frac{-\Delta G_{TS-N}}{RT}\right) \quad \textbf{Eq. 1.51}$$

where A is the pre-exponential factor, Ea the activation energy equal to ΔG_{TS-N} in kcal mol^{-1}, i.e., the difference in Gibbs free energy between the transition state and the native state, R the gas constant and T the temperature. Considering the unfolding of a wild-type protein with a mutant, the ratio of their unfolding rate constants is given by:

$$\boldsymbol{\frac{k_U^{mut}}{k_U^{WT}} = \frac{A.\exp\left(\frac{-\Delta G_{TS-N}^{mut}}{RT}\right)}{A'.\exp\left(\frac{-\Delta G_{TS-N}^{WT}}{RT}\right)}} \quad \textbf{Eq. 1.52}$$

Assuming that the pre-exponential coefficients A and A' are equal for the unfolding of both wild-type and mutant, the equation can be simplified to:

$$\boldsymbol{\ln\left(\frac{k_U^{mut}}{k_U^{WT}}\right) = -\frac{\Delta G_{TS-N}^{mut}}{RT} + \frac{\Delta G_{TS-N}^{WT}}{RT}} \quad \textbf{Eq. 1.53}$$

$$-\mathrm{RT}.\ln\left(\frac{k_U^{mut}}{k_U^{WT}}\right) = \Delta G_{TS-N}^{mut} - \Delta G_{TS-N}^{WT} \quad \textbf{Eq. 1.54}$$

$$\mathrm{RT}.\ln\left(\frac{k_U^{mut}}{k_U^{WT}}\right) = \Delta\Delta G_{TS-N} \quad \textbf{Eq. 1.55}$$

From the energy diagram (**Figure 1.12**), it can be seen that:

$$\boldsymbol{\Delta\Delta G_{TS-D} + \Delta G_{TS-N}^{WT} = \Delta\Delta G_{D-N} + \Delta G_{TS-N}^{mut}} \quad \textbf{Eq. 1.56}$$

$$\Delta\Delta G_{TS-D} = \Delta\Delta G_{D-N} - (\Delta G_{TS-N}^{WT} - \Delta G_{TS-N}^{mut}) \quad \textbf{Eq. 1.57}$$

$$\Delta\Delta G_{TS-D} = \Delta\Delta G_{D-N} - \Delta\Delta G_{TS-N} \quad \textbf{Eq. 1.58}$$

Therefore, combining **Equations 1.55** and **1.58**:

$$\Delta\Delta G_{TS-D} = \Delta\Delta G_{D-N} - RT.\ln\left(\frac{k_U^{mut}}{k_U^{WT}}\right) \qquad \textbf{Eq. 1.59}$$

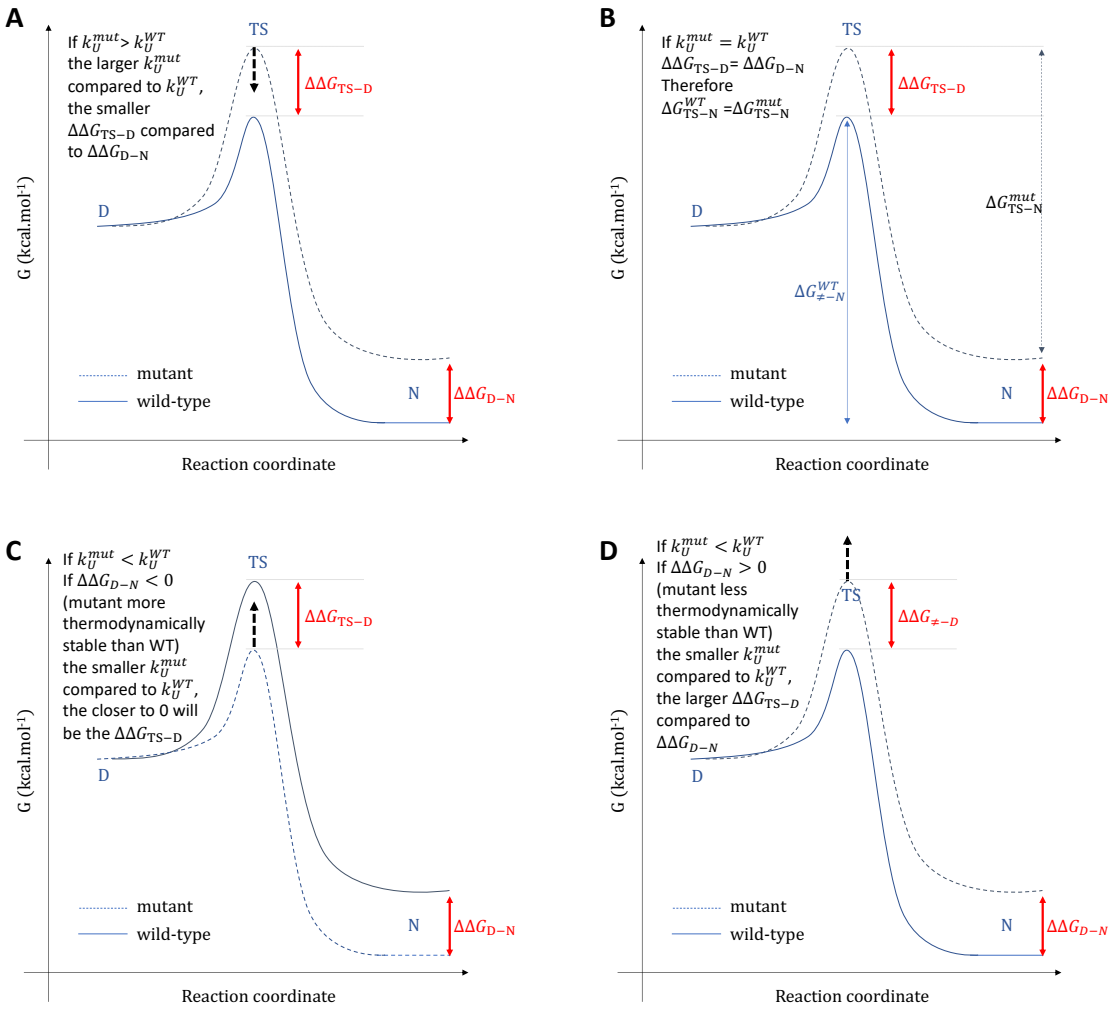


Figure 1.13: Diagram illustrating how the energy of the transition state of a mutant varies depending on the value of k_U^{mut} relative to k_U^{WT} A. $k_U^{mut} > k_U^{WT}$ **B.** $k_U^{mut} = k_U^{WT}$ **C.** $k_U^{mut} < k_U^{WT}$ and $\Delta\Delta G_{D-N} < 0$ **D.** $k_U^{mut} < k_U^{WT}$ and $\Delta\Delta G_{D-N} > 0$.

If $k_U^{mut} > k_U^{WT}$ (**Figure 1.13 A**), the mutant unfolds faster than the wild type. According to **Equation 1.59**, the larger k_U^{mut} compared to k_U^{WT}, the smaller $\Delta\Delta G_{TS-D}$ compared to $\Delta\Delta G_{D-N}$, therefore, the easier it is to overcome the energy barrier: the mutation decreases the kinetic stability of the protein.

If $k_U^{mut} = k_U^{WT}$ (**Figure 1.13 B**), the mutant unfolds at the same rate as the wild type. According to **Equation 1.59**, $\Delta\Delta G_{TS-D} = \Delta\Delta G_{D-N}$ so the transition state is as far from

the native state in both cases, i.e. the kinetic stability remains unchanged. An alternative way of viewing this is that the mutation destabilizes the transition state to the same degree as the native state.

If $k_U^{mut} < k_U^{WT}$, the mutant unfolds slower than the wild type; two cases are possible. In the most usual scenario, if the mutant is more thermodynamically stable than the wild type, then $\Delta\Delta G_{D-N}$ is negative ($\Delta G_{D-N}^{mut} > \Delta G_{D-N}^{WT}$, **Figure 1.13 C**). $\Delta\Delta G_{TS-D}$ is usually negative as well, and according to **Equation 1.59**, the smaller k_U^{mut} compared to k_U^{WT}, the larger $\Delta\Delta G_{TS-D}$ with respect to $\Delta\Delta G_{D-N}$, therefore the closer $\Delta\Delta G_{TS-D}$ would be to 0. Thus, the harder it is to overcome the energy barrier, i.e., the slower the protein unfolds: the mutation therefore increases kinetic stability. In the second scenario, the mutant is thermodynamically less stable than the wild type, i.e., $\Delta\Delta G_{D-N} > 0$, but kinetically more stable (**Figure 1.13 D**). In this case, the smaller k_U^{mut} compared to k_U^{WT}, the larger $\Delta\Delta G_{TS-D}$ with respect to $\Delta\Delta G_{D-N}$, i.e., the mutation destabilizes the transition state more than the native state.

1.2.5 Thermal stability

Amongst the different strategies to unfold a protein, it is possible to thermally unfold it using heat. The difference in Gibbs free energy is related to the difference in enthalpy and difference in entropy as follows:

$$\boldsymbol{\Delta G_{D-N} = \Delta H_{D-N} - T\Delta S_{D-N}} \qquad \textbf{Eq. 1.60}$$

At room temperature, ΔG_{D-N} is positive, i.e. $\Delta H_{D-N} > T\Delta S_{D-N}$, therefore the denaturation of the protein is not favored. When the temperature increases, $T\Delta S_{D-N}$ becomes larger than ΔH_{D-N}, so as ΔG_{D-N} becomes positive and the denatured state becomes favored. Both ΔH_{D-N} and ΔS_{D-N} depend on the temperature as follows:

$$\boldsymbol{\Delta H_{D-N}(T) = \Delta H_{D-N}(T_{ref}) + \int_{Tref}^{T} \Delta Cp.\, dT} \qquad \textbf{Eq. 1.61}$$

$$\Delta H_{D-N}(T) = \Delta H_{D-N}(T_{ref}) + \Delta Cp.(T - T_{ref}) \qquad \textbf{Eq. 1.62}$$

$$\Delta S_{D-N}(T) = \Delta S_{D-N}(T_{ref}) + \int_{Tref}^{T} \Delta Cp.\frac{dT}{T} \qquad \textbf{Eq. 1.63}$$

Where ΔCp is the difference in heat capacity, and $Tref$ is the temperature of reference.

Given that at T_{ref} (T_m), $\Delta G_{D-N}(T_m) = 0$, therefore:

$$\boldsymbol{T_{ref}.\Delta S_{D-N}(T_{ref}) = \Delta H_{D-N}(T_{ref})} \qquad \textbf{Eq. 1.64}$$

$$\Delta S_{D-N}(T) = \frac{\Delta H_{D-N}(T_{ref})}{T_{ref}} + \int_{Tref}^{T} \Delta Cp.\frac{dT}{T} \qquad \textbf{Eq. 1.65}$$

$$\Delta S_{D-N}(T) = \frac{\Delta H_{D-N}(T_{ref})}{T_{ref}} + \Delta Cp.\, ln(\frac{T}{T_{ref}}) \qquad \textbf{Eq. 1.66}$$

Therefore, combining **Equations 1.62** and **1.66** with **Equation 2.2** :

$$\boldsymbol{\Delta G_{D-N}(T) = \Delta H_{D-N}(T_{ref}) + \Delta Cp.(T - T_{ref}) - T(\frac{\Delta H_{D-N}(T_{ref})}{T_{ref}} + \Delta Cp.\, ln\left(\frac{T}{T_{ref}}\right))} \qquad \textbf{Eq. 1.67}$$

$$\boldsymbol{\Delta G_{D-N}(T) = \Delta H_{D-N}(T_{ref}).\left[1 - \frac{T}{T_{ref}}\right] - \Delta Cp.\left(T.\, ln\left(\frac{T}{T_{ref}}\right) - T + T_{ref}\right)} \qquad \textbf{Eq. 1.68}$$

If ΔCp is high, there is a significant change in ΔH and ΔS with temperature (**Figure 1.14**).

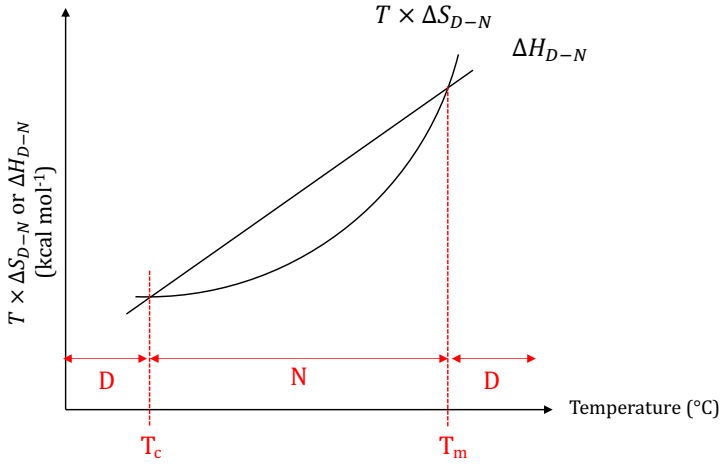


Figure 1.14: Diagram illustrating how thermodynamic parameters (T $\times$ ΔS_{D-N} and ΔH_{D-N}) vary with temperature. The hot and cold melting temperatures are the temperatures when T $\times$ ΔS_{D-N} = ΔH_{D-N}, i.e., $\Delta G_{D-N} = 0$.

Given the previous equations, ΔH_{D-N} changes linearly with temperature, whilst $T \times \Delta S_{D-N}$ does not. As $T \times \Delta S_{D-N}$ changes more rapidly than ΔH_{D-N} with temperature, this leads to a point where the term $T \times \Delta S_{D-N}$ dominates, and therefore the protein is no longer stable and unfolds. There is another intersection point at a lower temperature, called the cold denaturation temperature, at which the protein also unfolds. If the protein is incubated in temperatures between the cold unfolding and thermal melting temperature, it is in primarily in its native state. (**Figure 1.14**). There is a temperature of maximal stability, as represented in **Figure 1.15**

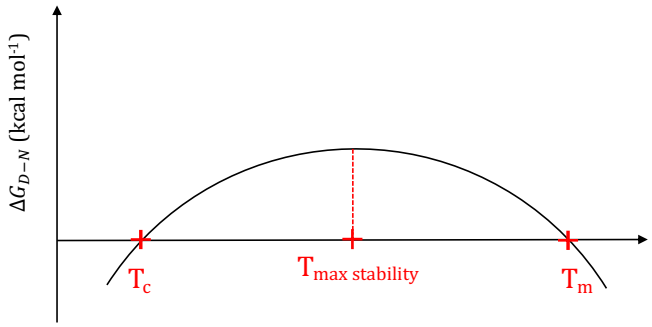


Figure 1.15: Temperature of maximal thermal stability

To calculate ΔC_p, one needs to perform the thermal unfolding at different pHs to obtain different values of T_m and $\Delta H_{D-N}(T_m)$. Plotting $\Delta H_{D-N}(T_m)$ versus T_m, the slope of the linear regression is ΔC_p (**Equation 1.62**).

1.2.6 Molecular dynamics: hydrogen-deuterium exchange mass spectrometry

1.2.6.1 Introduction on hydrogen-deuterium exchange

1.2.6.1.1 General principle

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) is a technique which probes the molecular dynamics of a native state of a protein at high-resolution. This technique is highly sensitive to the solvent accessibility and hydrogen-bonding of backbone amide groups and has been used successfully to investigate interactions between domains in multi-domain proteins as well as domain stability[104–107].

The HDX can be measured by mass spectrometry, by monitoring the increase in m/z over time, or by NMR, by monitoring the loss in ^{1}H NMR signal. In this Thesis, the focus is on HDX-MS.

1.2.6.1.2 Mass spectrometry

Mass spectrometry is an analytical technique that detects the mass of molecules, by first ionizing them into the gas phase and then sorting the ions by their m/z ratio. This enables the identification and quantification of different molecules in mixtures. All mass spectrometers are composed of an ionization source, where molecules are ionized, a mass analyzer, which sorts and separates ions according to their mass-to-charge ratio (m/z) and a detector, which measures the separated ions, the signal of which is then converted to an output. The fundamental components of a mass spectrometer are represented on **Figure 1.16**. Mass analyzers are kept at a constant high vacuum (low pressure) to prevent unwanted ion-molecule collisions and increase ion transmission and resolution.

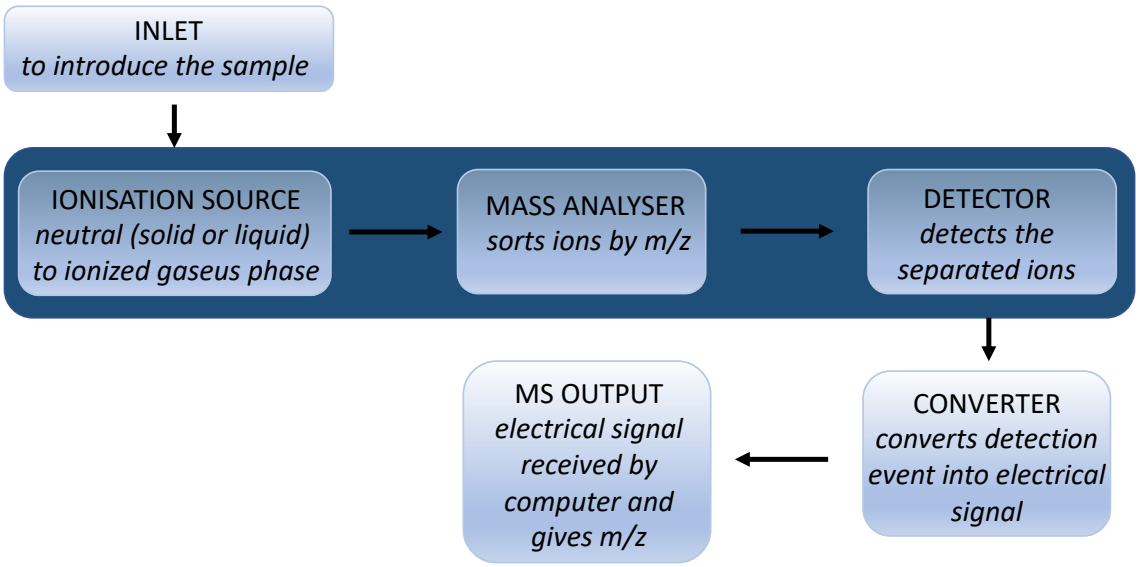


Figure 1.16: Schematic representation of the components of a mass spectrometer

1.2.6.2 Mechanism of hydrogen-deuterium exchange

1.2.6.2.1 Hydrogen-deuterium exchange rates are dependent on pH and temperature

This technique relies on the exchange of amide protons with deuterium, the heavier isotope of hydrogen (given it contains a neutron as well as a proton). Upon exposure to D_2O, hydrogens bound to nitrogen, oxygen or sulfur can exchange with deuterium. However, only hydrogens bound to the amide nitrogens in the protein

backbone exchange with deuterium on a timescale that can be probed by the HDX measurements. The exchange occurs via basic, acidic or water catalysis[108]. The intrinsic rate of exchange (k_{ch}) of amide hydrogens, taking into account the pH, follows the equation:

$$\boldsymbol{k_{ch} = k_{int,H}\,[H^+] + k_{int,OH}\,[OH^-] + k_{int,H2O}\,[H_2O]} \qquad \textbf{Eq. 1.69}$$

Where $k_{int,H}$ and $k_{int,OH}$ are the rate constants for the acid- and base-catalyzed exchange, and $k_{int,H2O}$ is the intrinsic coefficient of water-catalyzed proton transfer. This equation determines the exchange rate at different values of pH, and can be calculated using the rate constants published in Bai *et al.*,1993 [109,110]. **Figure 1.17 A** shows a representation of how the intrinsic rate of exchange depends on the pH.

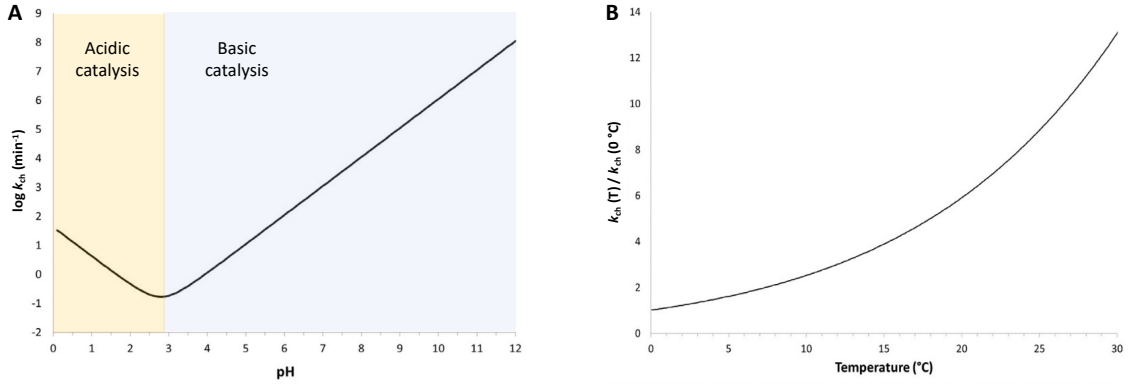


Figure 1.17: Dependence of the intrinsic rate of exchange on the pH of a solution and temperature. A. Chevron plot depicting exchange rates (log k_{ch}; 25 °C) for amide hydrogens with deuterium as a function of pH. The pH minimum corresponds to the lowest point of the curve and reflects the lowest exchange rate for amide backbone hydrogens. **B.** Plot depicting the intrinsic HDX rate of amide hydrogens (k_{ch}) as a function of temperature for base-catalysed exchange. k_{ch} was calculated based on **Equation 1.70** for temperature range from 0 to 30 °C and normalized to k_{ch} at 0 °C. Reproduced and adapted with permission from the journal *Methods* (Oganesyan *et al.,* 2018[111]).

Acid-catalyzed exchange is predominant at pH values lower than 2.5-3[110,112], whereas base-catalyzed exchange occurs at pH values higher than 3[110,112] (**Figure 1.17 A)**. Therefore, at neutral pH, the exchange is mainly driven by base catalysis. A schematic representation of the mechanism of acid and base catalysis can be found in **Figure 1.18**[113].

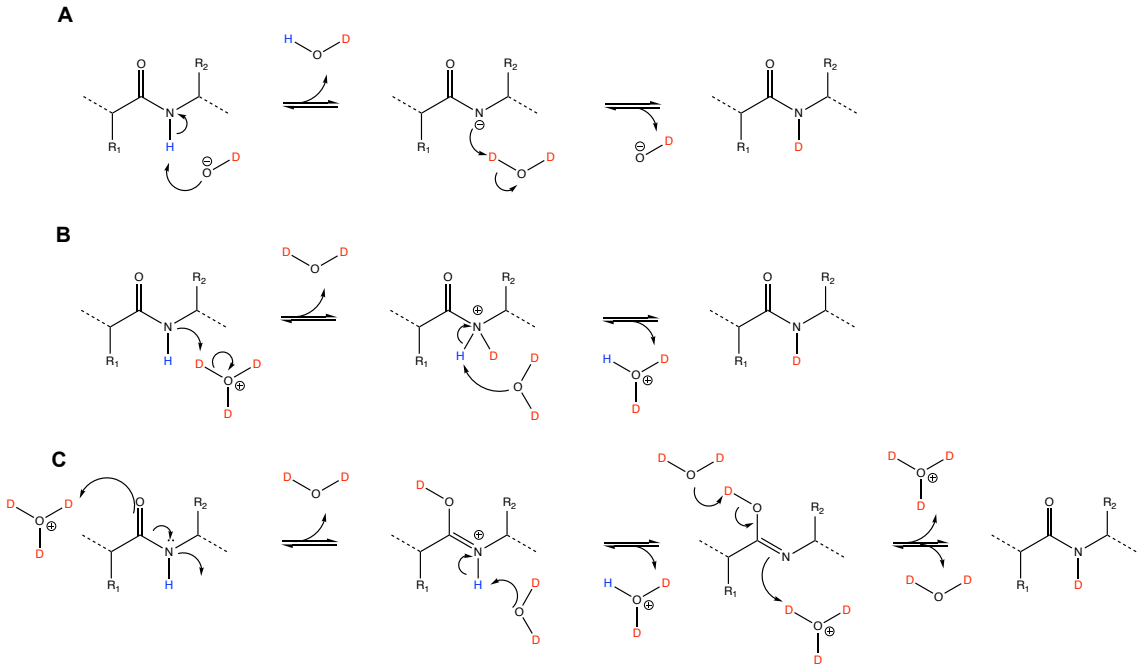


Figure 1.18: Hydrogen-deuterium exchange mechanism A. Basic catalysis **B.** Acidic catalysis occurs through protonation of the NH group. **C.** Acidic catalysis occurs through the protonation of the C=O group. Adapted with permission from the journal *Biomedicines* (Narang *et al.*, 2020[113]).

The intrinsic exchange rate is also sensitive to temperature, governed by the following equation[109]:

$$\boldsymbol{k_{ch} = k_{rc}(293) \times \exp\left(-\frac{E_a}{R}\left[\frac{1}{T} - \frac{1}{293}\right]\right)} \qquad \textbf{Eq. 1.70}$$

Where $k_{rc}(293)$ is the reference rate constant at 20 °C ($k_{int,acid}$ and/or $k_{int,base}$ and/or $k_{int,water}$ depending on the pH); E_a is the activation energy for acid-, base-, or water-catalyzed proton exchange ($E_{a,H+}$ =14 kcal mol^{-1}, $E_{a,OH}$ =17 kcal mol^{-1} and $E_{a,H2O}$ =19 kcal mol^{-1}); and R is the gas constant (0.001986 kcal mol^{-1} K^{-1}) [109,111]. The graph corresponding to this equation is represented in **Figure 1.17 B.**

1.2.6.2.2 Hydrogen-deuterium exchange rates dependent on protein folding

For a folded protein in the native state, the exchange of amide hydrogen atoms with deuterium is mainly dictated by the structure of the protein. Amide hydrogens involved in H-bonds, like in α-helices and β-sheets, or buried in hydrophobic cores, don't exchange with solvent. However, because of structural fluctuations in the native state

(thermally driven molecular dynamics), occasionally the bonds will transiently break and some buried regions will transiently become exposed to solvent. The amide hydrogens involved will then exchange with the solvent. This process is shown **Scheme 1.71.**

$$NH_{cl} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} NH_{op} \xrightarrow{k_{ch}} ND_{op} \underset{k_{op}}{\overset{k_{cl}}{\rightleftharpoons}} ND_{cl} \qquad \textbf{Sch. 1.71}$$

Therefore, the observed rate constant for HDX under native conditions can be described by **Equation 1.72**[114].

$$k_{HDX} = \frac{k_{op} \times k_{ch}}{k_{cl} + k_{ch}} \qquad \textbf{Eq. 1.72}$$

These open/closed fluctuations are probed by HDX-MS, by incubating samples in D_2O for increasing amounts of time. Two different amide exchange profiles can occur, EX1 and EX2, depending on the structural dynamics and therefore the relative values of k_{cl} and k_{ch}. Under the EX2 regime, the rate constant of closing is much faster than the chemical exchange rate ($k_{cl} >> k_{ch}$), and therefore the amide needs to undergo many brief unfolding events before hydrogen exchanges with deuterium occurs. Therefore, the exchange of the amide protons with the solvent happens iteratively when a H-bond transiently breaks or the hydrophobic core transiently opens (**Figure 1.19 A**). This is associated with a gradual shift in the isotopic distribution of peptides. This regime is the most common case. Under EX2 conditions, **Equation 1.72** simplifies to:

$$k_{HDX} = \frac{k_{op} \times k_{ch}}{k_{cl}} \qquad \textbf{Eq. 1.73}$$

Under the EX1 regime, the rate of closing is slower than that of the chemical exchange ($k_{cl} << k_{ch}$). In this case, a set of amide protons in a region within one protein molecule that transiently unfolds fully exchanges before refolding occurs. The proportion of the population of protons that have exchanged increases every time the region becomes exposed in a different molecule, leading to a bimodal distribution with the intensity of the non-deuterated distribution decreasing as the fully-deuterated one increases (**Figure 1.19 B).**

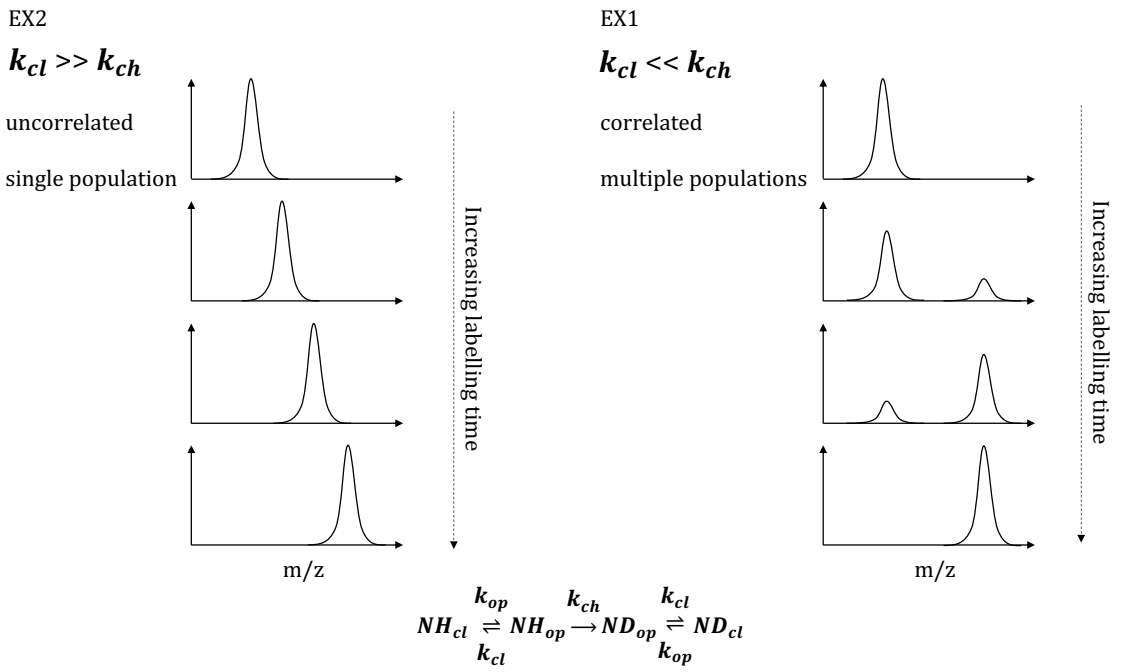


Figure 1.19: Schematic representation of EX2 and EX2 kinetics. Adapted with permission from the journal *Current Protocols in Protein Science* (Morgan and Engen, 2009[115]).

1.2.6.3 The workflow for HDX-MS experiments

1.2.6.3.1 Overall workflow

Several steps are needed to obtain high-resolution data on hydrogen-deuterium exchange monitored with mass spectrometry via a bottom-up approach (**Figure 1.20**), i.e., generating peptides before entering the mass spectrometer, as opposed to the top-down approach where the deuterated protein is electrospray ionized intact and is fragmented by electron transfer dissociation inside the MS, which does not scramble the deuteration. First, the sample is incubated in deuterated buffer for a fixed amount of time; several different incubation time points are measured, in triplicate, in order to obtain a time course of deuteration. The deuterated buffer is then quenched with a buffer with two aims. First, the quench buffer slows down the hydrogen-deuterium exchange back exchange as much as possible by decreasing the pH value to pH 2.5, which is the minimum intrinsic rate of exchange of amide hydrogens (**Figure 1.17 A**) together with performing the quenching step between 0 and 4 °C (**Figure 1.17 B**). The second objective of the quench buffer is to unfold the protein and reduce any potential disulfide bridges. Unfolding can occur by decreasing the pH on its own, however, frequently a chemical denaturant and reducing agent are used depending on the stability of the protein being

investigated. In this way, all the regions within the protein are accessible to the pepsin so they can be digested. The quenched, labelled protein is then digested via an online pepsin column to generate small peptides, and then the peptides bind to a trapping column which carries out a desalting step. The peptides are then eluted onto an analytical C18 column and electrospray ionized in the mass spectrometer, and analyzed with MS^1 only, to avoid scrambling the deuterium locations by collision-induced dissociation. The labelled peptide data is then analyzed for each peptide, and then per residue, and frequently data are compared to a reference dataset, in which case the difference in deuterium is plotted with colors onto an appropriate crystal structure.

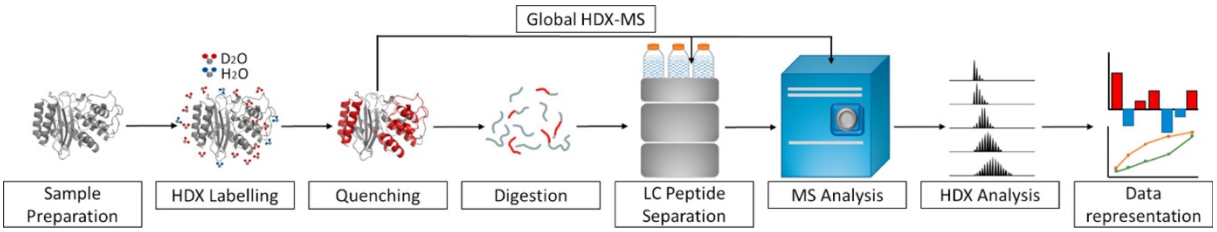


Figure 1.20: Typical overall work scheme for a HDX-MS experiment. Reproduced with permission from the journal *Methods* (Oganesyan *et al.,* 2018[111])

1.2.6.3.2 Generation of the peptide map

To obtain high-resolution data, a peptide map needs to be generated, to locate the peptides generated by the pepsin column onto the protein sequence and therefore structure. Ideally, the peptide map would have the highest coverage and redundancy possible, to be able to accurately assign deuterium exchange to single residues. The peptide map is generated with the same workflow as the labelled data, but a fragmentation method, called tandem mass spectrometry or MS^2, is used in the mass spectrometer, to guarantee high accuracy of the identification, i.e., that the two fragmented peptides generated still match the sequence. Several methods can be used to fragment the peptides. Low energy dissociation processes like collision-induced dissociation (CID, also called collisionally activated dissociation (CAD)), mainly breaks the thermally labile C-N amide bonds of the polypeptide chain, generating b and y ions[116–118] (**Figure 1.21**, **Figure 1.22**), and remains the most used approach. Electron-based activation methods like electron transfer dissociation (ETD) or electron capture dissociation (ECD) mainly generated c

and z ions as the bonds N-C_α are preferentially fragmented[116–118] (**Figure 1.21**). Higher energy activation methods generate a and y ions by cleaving C_α-C bonds, additionally to b-y and c-z ions[116–118] (**Figure 1.21**). The a, b and c ions retain the charge at the N-terminus of the fragmented peptide, whereas the x, y and z ions retain the charge at the C-terminus.

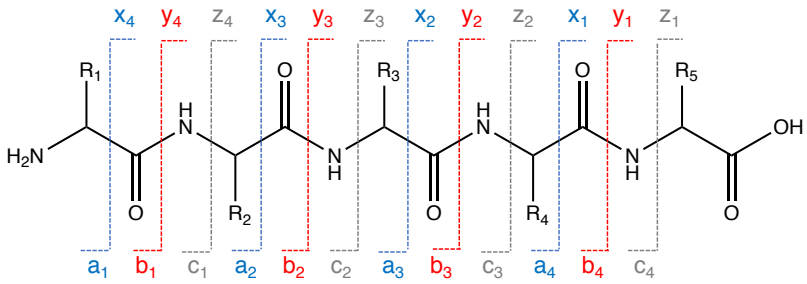


Figure 1.21: Cleavage sites depending on the fragmentation method. CID generates mainly b and y ions, ETD c and z, and higher energy activation methods generates a and x. a, b and c ions retain the charge at the N-terminus of the fragmented peptide, whereas the x, y and z ions retain the charge at the C-terminus. Nomenclature by Roepstorff, P. & Fohlman, 1984 [117], revised by Biemann, 1988 [118].

Collision-induced dissociation (CID) works by the energetic collisions between ions and non-reactive gas atoms, like helium, nitrogen or argon. Every time a collision occurs, part of the kinetic energy of the ion is converted to internal energy. The accumulation of internal energy in the ion results in the fragmentation between C-N amide bonds[116,119]. If the parent ion has one positive charge, the fragmentation of the C-N bonds generates an ionic and a neutral species. If the charge is retained at the N-terminal fragment, a b ion (with a C-terminal CO^+) and a neutral fragment are generated: the proton stays on the C-terminal fragment and neutralises it; if the charge remains on the C-terminal fragment, a neutral fragment and a y ion (N-terminal NH_3^+) are generated: the proton stays on the C-terminal fragment and an additional proton from the N-terminal fragment is transferred, creating a positive ion. Mechanistic details are represented in **Figure 1.22**[120].

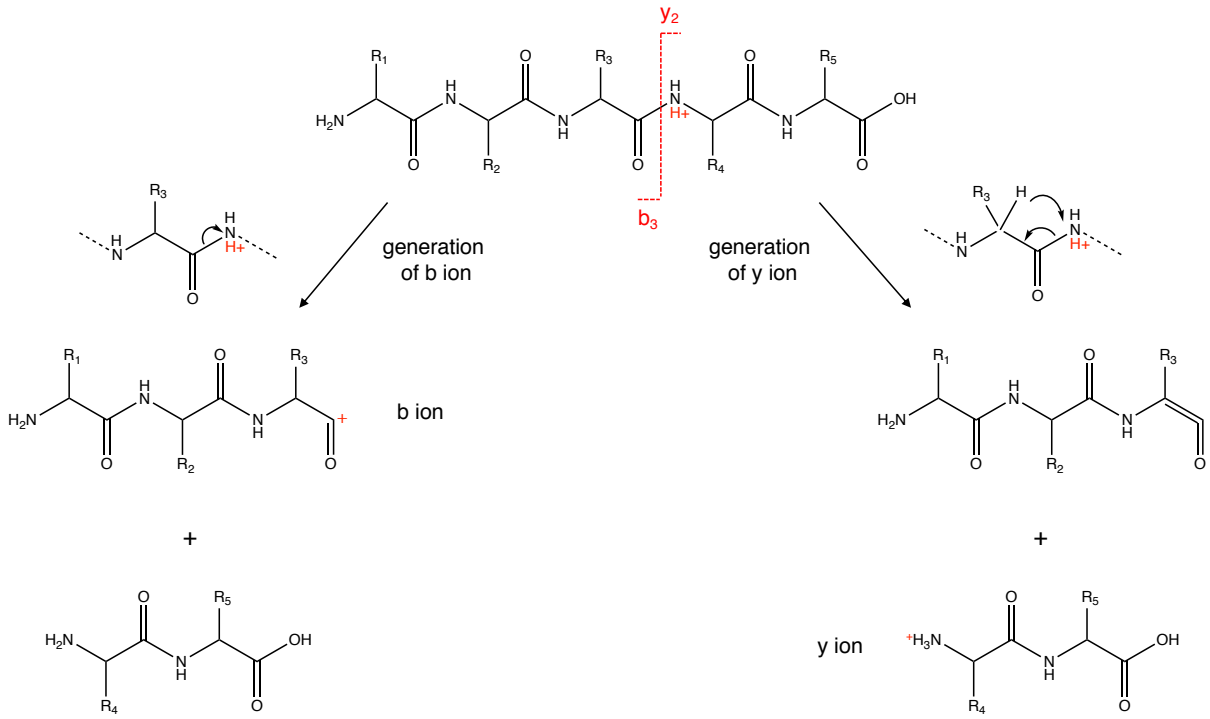


Figure 1.22: Details of the fragmentation process. If the charge stays on the N-terminal fragment, a b ion and a neutral C-terminal peptide are generated. If the charge remains on the C-terminal fragment, then a y ion and a neutral peptide are generated. Adapted from *Proceedings of National Academy of Science* (Hunt *et al.*, 1986 [120]).

1.2.6.3.3 Data Analysis

HDX-MS is often used to compare two or more different states of the same protein. The deuterium uptake of the peptides identified by the peptide map is calculated by specialized softwares based on undeuterated m/z values, considering the unlabelled and labelled peptides elute at very similar retention times in the analytical liquid chromatography step. The deuterium uptake of the reference is subtracted from that of the state of interest, and only significant differences are taken into account. The total deuterium incorporation per peptide is divided by the number of residues on the peptide that can exchange deuterium in order to obtain an average level of D incorporation per residue. The deuterium uptake of overlapping peptides is taken into account and the different incorporations for the same overlapping residue is averaged in order to obtain the most accurate labelling level possible. The deuterium uptake is then plotted onto a known crystal structure in order to visually identify the regions that are exchanging more or less then the reference.

1.2.7 Affinity measurements

If two molecules A and B have affinity for each other, the separate species and the associated complex are in equilibrium:

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \qquad \textbf{Eq. 1.74}$$

Where, k_{off} is the dissociation rate constant and k_{on} is the association rate constant. The affinity between these two molecules is defined by the inverse of the equilibrium dissociation constant, K_D, which can be calculated using the following equation:

$$K_D = \frac{k_{off}}{k_{on}} = \frac{[A].[B]}{[AB]} \qquad \textbf{Eq. 1.75}$$

because:

$$rate = \frac{d[A]}{dt} = k[A] \qquad \textbf{Eq. 1.76}$$

and:

$$k_{on}[A][B] = k_{off}[AB] \qquad \textbf{Eq. 1.77}$$

Where $k_{on}[A][B]$ and $k_{off}[AB]$ are respectively the association and dissociation rates.

To accurately measure k_{off} and k_{on} values, a range of concentrations of the antigen is usually and the data obtained fitted with using a global fitting strategy. Methods like surface plasmon resonance or biolayer interferometry can be used to obtain these values.

The data for the association phase is fit to a single exponential, as follows:

$$Y = Y_0 + A \times (1 - \exp(-k_{obs} \times t)) \qquad \textbf{Eq. 1.78}$$

As the observed rate constant depends on the analyte concentration and the dissociation constant, the k_{on} can be extracted using **Equation 1.79** :

$$k_{on} = \frac{k_{obs} - k_{off}}{[analyte]} \qquad \textbf{Eq. 1.79}$$

The data corresponding to the dissociation phase is fit to **Equation 1.80**

$$Y = Y_0 + A \times \exp(-k_{off} \times t) \qquad \textbf{Eq. 1.80}$$

1.2.8 Other stability measurement methods

Whilst the techniques that were used in this Thesis have been described in detail in earlier sections, there are other methods that inform on structure, stability, aggregation, or higher-order species formation.

Secondary structure can be investigated with far UV-circular dichroism (UV-CD), which depends on the difference in absorption of left-handed and right-handed circularly polarized light by the protein backbone[121,122]. In addition, Fourier-transform infra-red spectroscopy (FTIR), measures the intensity of the absorption of IR radiation of a protein as a function of wavelength. The IR excites vibrational transitions in the protein molecules which are influenced by the secondary structure and intramolecular and intermolecular effects[123]. Instruments like the AQS^3 pro from RedShiftBio use a similar technology. Structural information can be obtained by X-ray crystallography at high resolution[124], cryo-electron microscopy[125] for large complexes, or atomic force microscopy[126] which can provide details on the topology of a protein assembly if it is sufficiently large.

The presence of aggregates or higher-order structures can be probed by size-exclusion chromatography (SEC) liquid chromatography, if they are smaller than the 0.22 μm filter. Analytical ultra-centrifugation (AUC) is a technique which preserves the native state of a protein and presents a few advantages over SEC, namely the study of higher-order aggregates bigger than the 0.22 μm filter, and the absence of potential dilution of the higher-order species in the mobile phase. Using light scattering and sedimentation time, it informs on the native state, provides an approx. molecular weight, informs on impurities, reversible self-association[127] (if tested at different protein concentrations). Capillary isoelectric focusing (cEIF) is an analytical technique that allows the separation

of protein mixtures, protein glycoforms and other charge variants, based on their isoelectric point (pI)[128]. It is also used to determine the experimental pI of a protein. Reverse phase chromatography can be useful to observe impurities, which can be due to oxidation, fragmentation to name but a few examples. Mass spectrometry can be employed to undertake peptide mapping providing very precise information on the location of modification.

The identification of regions that are aggregation prone can be investigated by fast photochemical oxidation of proteins (FPOP), where hydroxyl radicals, liberated from the photolysis of hydrogen peroxide, covalently label solvent accessible amino acid side chains on the microsecond-millisecond timescale[129,130]. This enables the identification of regions that are exposed in the one state, e.g., a monomer, but buried in an aggregated state. HDX-MS can be similarly used to investigate aggregation prone regions.

More recently, manufacturers like Nanotemper have developed instruments, namely Prometheus, that can measure the thermal and thermodynamic stabilities, reversibility of unfolding as well as aggregation using embedded nano differential scanning fluorimetry (DSF) and dynamic light scattering (DLS).

2 MATERIAL AND METHODS

2.1 MATERIAL

2.1.1 Reagents

2.1.1.1 Reagents for molecular biology

The DNA sequences were ordered from Geneart and the vectors for mammalian expression are proprietary from AstraZeneca. The reagents used for the molecular biology experiments to assemble DNA strands and vectors were the following: OneTaq® 2X Master Mix with Standard Buffer (New England BioLabs), restriction enzymes ApaLI, BstEII, EcoRI, NgoMIV, PacI (New England BioLabs), NEBuilder® HiFi DNA Assembly Master Mix (New England BioLabs), Quick Ligation™ Kit (New England BioLabs), High Pure PCR Product Purification kit (Roche, Merck), Zymo Mix&Go! Competent DH5 alpha (Zymo Research), UltraPure™ agarose (Invitrogen by ThermoFisher Scientific), SYBR® Safe DNA gel stain (Invitrogen by ThermoFisher Scientific), 1 Kb Plus DNA Ladder (Invitrogen by ThermoFisher Scientific), ampicillin sodium salt (Duchefa Biochemie). TAE buffer was made internally at AstraZeneca using Tris base (Tris(hydroxymethyl)aminomethane, Molecular Biology Grade, Ultra Pure, Melford), acetic acid glacial (AnalaR NORMAPUR, ACS Reag. Ph. Eur., VWR, BDH Chemicals) and EDTA (Ethylenediaminetetraacetic acid disodium salt dihydrate, AnalaR NORMAPUR, Reag. Ph. Eur, ACS, VWR, BDH Chemicals). LB plates were made internally at AstraZeneca using BD Bacto™ Agar, sodium chloride (AnalR NORMAPUR, ACS, Reag. Ph. Eur., VWR, BDH Chemicals), Bacto™ Yeast Extract (Gibco), Bacto™ Tryptone (Pancreatic Digest of Casein, Gibco). The x6 loading dye was prepared internally at AstraZeneca using Bromphenol Blue sodium salt, for electrophoresis (Molecular Biology Tested, Sigma-Aldrich), sucrose (BioXtra ≥ 99.5% (GC), Sigma) and Tris base (Tris(hydroxymethyl) aminomethane, Molecular Biology Grade, Ultra Pure, Melford).

2.1.1.2 Buffers for biophysical experiments

Histidine buffer was prepared with the acidic form L-histidine monohydrochloride monohydrate ($C_6H_9N_3O_2.HCl.H_2O$, ≥98% (TLC)) and the basic form L-histidine ($C_6H_9N_3O_2$ ReagentPlus®, ≥99% (TLC)), both from Sigma-Aldrich, USA.

Tris buffer was prepared with the acidic form Tris hydrochloride ($NH_2C(CH_2OH)_3 \cdot HCl$) and the Trizma® base ($NH_2C(CH_2OH)_3$) both from Sigma-Aldrich, USA.

Potassium phosphate buffer was prepared with the acidic form potassium dihydrogen phosphate (KH_2PO_4) and the basic form dipotassium hydrogen phosphate (K_2HPO_4), both from Sigma-Aldrich, USA.

The quantities of the acid and the base were calculated using the Henderson-Hasselbalch equation:

$$\mathbf{pH = p}\boldsymbol{Ka} \mathbf{+ log_{10}(\frac{base}{acid})} \qquad \textbf{Eq. 2.1}$$

For low pH solutions where acid and base were not both available, the pH was adjusted with a pH meter (in the case of HDX-MS quench buffer).

Guanidinium chloride (GdmCl) was purchased from Invitrogen (UltraPureTM, ≥99% purity). Guanidine thiocyanate (GdmSCN) was purchased from Sigma-Aldrich (BioReagent, for molecular biology, ≥99%). Urea was purchased from Sigma (BioXtra, pH 7.5-9.5 (20 °C, 5 M in H_2O)). TCEP (Tris(2-carboxyethyl)phosphine hydrochloride, $C_9H_{15}O_6P{\cdot}HCl$) was purchased from Sigma-Aldrich.

2.1.1.3 Buffers for protein purification

For protein purification, DPBS was purchased from Sigma (D8537-500ML). Imidazole was purchased from ACROS Organics™ (99+%, crystalline).

2.1.1.4 Reagents for SDS-PAGE

Protein expression, purification and proteolysis were verified by SDS-PAGE, using the following reagents: NuPAGE™ MES SDS Running Buffer (20X, Invitrogen by ThermoFisher Scientific), NuPAGE™ LDS Sample Buffer (4X, Invitrogen by ThermoFisher Scientific), NuPAGE™ 4 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gel, 15-well (Invitrogen by ThermoFisher Scientific), Novex™ Sharp Pre-Stained Protein Standard (Invitrogen by ThermoFisher Scientific).

2.1.1.5 Buffers for activity assays, in vitro and on cells

The reagents used for activity assays (BioLayer Interferometry, flow cytometry and microscopy) were the following: DPBS (D8537-500ML, Sigma), BSA (bovine serum albumin solution, 30% in DPBS, sterile-filtered, BioXtra, Sigma-Aldrich), Tween 20 (TWEEN® 20 for molecular biology, Sigma-Aldrich), anti-human IgG Fc Capture (AHC) and streptavidin (SA) biosensors (Forte Bio), biotinylated HER2 (Her2/ERBB2

protein, human, recombinant (His & AVI Tag), Biotinylated, Sino Biological), Gibco™ StemPro™ Accutase™ cell Dissociation reagent (Gibco™ A1110501), DMEM 1x (Dubelcco's Modified Eagle Medium, Gibco), Pen Strep (Gibco™ penicillin-streptomycin (10,000 U/mL)), secondary antibody Alexa Fluor® 647 anti-human IgG Fc (BioLegend), paraformaldehyde (formaldehyde solution for molecular biology, 36.5-38% in H_2O), 96-well plates (PP V-bottom, Chimney Well, natural, Greiner Bio-one International), HBSS (Hank's balanced salt solution, Gibco), FBS (Fetal bovine serum, certified, heat inactivated, United States), HCS CellMask® Orange Stain (ThermoFisher), Hoechst stain (33342, Trihydrochloride, Trihydrate, ThermoFisher), Triton (Sigma-Aldrich), secondary antibody Alexa Fluor® 488 goat anti-Human IgG (ThermoFisher), 96-well plate (CellCarrier-96 Ultra Microplates, PDL-coated, black, 96-well, Perkin Elmer).

2.1.1.6 Enzymes used to cleave antibodies

Several enzymes were used to digest different parts of the antibodies and masks: TEV protease was purchased from Sigma-Aldrich (TEV protease T4455-1KU), Factor Xa protease from New England BioLabs, FabRICATOR and FabALACTICA from GENOVIS.

2.1.1.7 Buffers for the liquid chromatography mass spectrometry

2.1.1.7.1 Buffers for mobile phases for the liquid chromatography

The reagents used for the mobile phases of the liquid chromatography coupled to the mass spectrometry were the following: Water for chromatography (LC-MS grade) LiChrosolv®, (Supelco, Merck), HiPerSolv CHROMANORM acetonitrile for HPLC (VWR BDH Chemicals), formic acid, 99+% (ThermoFisher Scientific), TFA (trifluoroacetic acid, HPLC Grade, J.T.Baker™, Fisher Scientific, part of ThermoFisher Scientific).

2.1.1.7.2 Buffers for non-reduced peptide map sample preparation

For the non-reduced peptide map, the reagents used were: 8 M guanidine hydrochloride solution buffered, pH 8.5 (Sigma Aldrich, USA), N-ethylmaleimide BioUltra, ≥99.0%, HPLC grade (Sigma- Aldrich, USA), EDTA 0.5 M sterile solution biotechnology grade (Amresco, USA), Lysyl Endopeptidase® (Mass spectrometry grade, Fujifilm Wako Pure Chemical Corporation, Japan), water for chromatography LC-MS grade LiChrosolv® (Merck, Germany), acetonitrile for HPLC super gradient Reag. Ph.

Eur. USP, ACS water < 30 ppm (VWR, France), trifluoroacetic acid, Baker HPLC analyzed (Avantor Performance Materials, Poland), sodium phosphate monobasic (meets UPS testing specifications, anhydrous; Sigma-Aldrich, USA), sodium phosphate dibasic (puriss. p.a., ACS reagent, anhydrous, ≥99.0%, Sigma-Aldrich, USA) and sodium chloride NaCl, BioXtra, ≥99.5% (Sigma- Aldrich, USA).

2.1.2 Generation of antibody material

All the sequences of the antibodies investigated can be found in the Appendix I.

2.1.2.1 Generation of NIST mAb, TM S442C Fc, TM IP288 Fc (Chapter 3)

The full-length antibody NIST mAb IgG, and the Fc domains NIST mAb Fc, TM S442C Fc and TM Fc were a kind gift from AstraZeneca.

2.1.2.2 Generation of Fc-C239i

2.1.2.2.1 Expression and purification of Fc-C239i

The Fc domain of antibodies investigated in the interconversion of thiol states chapter (Chapter 4), Fc-C239i, were expressed and purified at AstraZeneca by Peter Ravn. NIST mAb Fc, the control protein, had been expressed and purified previously by AstraZeneca. The Fc-C239i construct, described in Dimasi *et al.*[35], has the exact same amino-acid sequence as NIST mAb Fc but with a cysteine inserted at position 239 in the C_H2 domain (sequences in Appendix I). The gene for the human Fc-C239i[35] was cloned into a mammalian expression vector. The Fc-C239i protein was transiently expressed in Chinese hamster ovary cells under serum-free conditions. Cleared culture supernatant was loaded directly onto MabSelect SuRe column equilibrated with PBS. Fc-C239i was eluted with 0.1 M glycine pH 2.7. Pooled fractions were buffer exchanged into PBS.

2.1.2.2.2 Enrichment in the three variants (2xSH, 2xCys and iDSB)

This experiment was performed by Paul Devine.

The inserted cysteine of Fc-C239i was shown to be present in three states: free-thiol, doubly-cysteinylated or forming an interchain disulfide bridge. Fc material was enriched in these three states by first diluting the material to 10 mg/mL in 10 mM TrisHCl, pH 8.0 and reducing the molecule with 2 mM DTT, incubating for 45 minutes at 37 °C. The sample was then buffer exchanged into 10 mM TrisHCl, using 10 kDa MWCO Amicon spin filters, to remove any remaining DTT and liberated cysteine. For generation of the free thiol (2xSH) variant of Fc-C239i, 1 mM of dehydroascorbic acid (dHAA) was

added and incubated for 2 hours at 25 °C. The material was buffer exchanged a final time to remove excess dHAA from the sample. Cysteinylated (2xCys) variants of Fc-C239i were generated through further manipulation of the free thiol variant. L-cysteine (Cys)/L-cystine (Ctn) was added to the sample, at a ratio of 1:4 Cys:Ctn, to a final concentration of 5 mM[131] and incubated at 37 °C for 3 hours. A buffer exchange step was conducted afterwards to remove free L-cysteine/L-cystine from the sample. Disulfide bonded variants (iDSB) were generated by incubating the free thiol variant at 50 °C for 12 hours overnight.

The free-thiol form was further capped with N-ethyl-maleimide (NEM), by adding 5 µg of N- ethylmaleimide (NEM) to 100 µg of reduced antibody, and incubated at room temperature for 20 min. The excess NEM was then removed by speed vacuum concentrator.

2.1.2.3 Generation of masked antibodies and masks in isolation

2.1.2.3.1 Expression and purification of masked antibodies and masks in isolation

The antibodies investigated in Chapter 6 are fusions of trastuzumab and an anti-idiotypic antibody fragment on the light chain, to prevent binding to HER2. The three anti-idiotypic antibody fragments investigated in this study were based on scFv40[132], scFv69[132] and a llama single domain[133], previously developed by the Navarro-Teulon's group[132,133]. Here, the order of the domains was inverted for scFv40 and scFv69, such that the N-terminal V_L domain was fused to the C-terminal V_H domain. An additional construct was also designed with a control scFv, scFvGipg013, with the sequence of the variable domains of Gipg013[134], which targets the GIP receptor and does not bind to trastuzumab, to investigate steric effects. The DNA strings encoding for scFv40, scFv69 and dAb were synthesized (GeneArt) and cloned as fusions to either a His-tag, for expression of the masks, or to the trastuzumab light chain[35] for expression of the masked antibodies (sequences in Appendix I) for expression of the masked antibodies, using standard molecular biology techniques. The vectors used were proprietary mammalian expression vectors. Successful cloning was confirmed by Sanger sequencing (SourceBioscience Sequencing). The antibodies with a single-chain mask were designed with two digestion sites: TEV protease in the linker between the antibody and the mask, and factor Xa protease in the linker between the two single-chain domains. The antibody

fused with a nanobody just had a TEV digestion site in the unique linker.

The antibody fragments scFv40, scFv69, dAb and the full-length antibody fusions T-scFv40, T-scFv69, T-dAb, T-scFvGipg013 were transiently expressed in Chinese hamster ovary (CHO) cells using proprietary medium[135]. The His-tagged proteins - scFv40, scFv69 and dAb were purified from culture supernatants using HisTrap™ Excel purification columns (Cytiva) equilibrated with DPBS (pH 7.4). An initial wash with DPBS and 20 mM imidazole pH 7.4 was performed to remove non-specific binding, and then the proteins were eluted with DPBS supplemented with 500 mM imidazole pH 7.4. Finally, the proteins were buffer exchanged into DPBS pH 7. 4 using a PD-10 desalting columns (Cytiva). For T, T-scFv40, T-scFv69, T-dAb, T-scFvGipg013, the cleared culture was loaded onto a MabSelect SuRe column (Cytiva), preequilibrated with DPBS pH 7.4. The proteins were eluted with 0.1 M glycine pH 2.7, and buffer exchanged into DPBS pH 7.4 using a PD-10 desalting columns (Cytiva).

2.1.2.3.2 Generation of masked antibody Fabs (single arm)

For SEC-MALS and X-ray crystallography, the full-length antibodies were digested by the enzyme FabALACTICA® (Genovis, Sweden), a protein that cleaves human IgG1 antibodies at a specific site above the hinge at the following site: …KSCDKT/HTCPPCP… and generates two single Fab arms and the Fc domain with hinge. The digestion was carried out overnight at 37 °C, using 2000 units to digest 2 mg of antibody. The digested Fab fragments were then purified with a CaptureSelect™ CH1-XL column (Life Technologies, ThermoFisher, Netherlands), run in DPBS and eluted with 0.1 M glycine pH 2.7. The recovered Fab domain pool was further purified by on SEC chromatography with DPBS (Superdex® 200 Increase 10/300 GL, Cytiva) to separate the single-armed Fabs from some undigested antibody. The fractions in glycine buffer were buffer-exchanged with PD-10 desalting columns (Cytiva).

2.1.2.3.3 Digestion of the masks by selected enzymes

Both the full-length IgG and the Fab domains of the antibody fusions were digested with TEV and Factor Xa, without further purification prior to the activity assays and the biophysical experiments. The TEV protease digestion (TEV protease T4455-1KU, Sigma-Aldrich) was carried out using 1 unit of enzyme for 20 µg of substrate overnight at 30 °C. The Factor Xa digestions (Factor Xa protease, NEB) were performed with 1 µg of enzyme per 50 µg of antibody, incubating the samples for three hours at 37 °C.

2.1.2.4 Generation of NIST mAb and trastuzumab Fab and Fc, and pepsin

NIST mAb was a kind gift from AstraZeneca, and trastuzumab was expressed as described previously in Section 2.1.2.3.1. To generate Fc and Fab domains from full-length IgGs, antibodies were digested with IdeS (FabRICATOR, Genovis), which cuts below the hinge (CPAPELLG/GPSVF), with a 1 unit : 20 µg ratio overnight at 37 °C. This generated Fab (two arms and hinge) and Fc domains (without hinge) which were separated affinity chromatography, i.e. CaptureSelect™ CH1-XL pre-packed column (Life Technologies, ThermoFisher Scientific), run in DPBS pH 7.4.

Pepsine was purchased from Sigma-Aldrich (Pepsin from porcine gastric mucosa, lyophilized powder, ≥ 2,500 units/mg protein (E1%/280)).

2.1.2.5 Generation of antibody linked to DVP and tetraDVP linker (Chapter 5)

The preparation of these samples was performed by Friederike Dannheim in the Spring Group. Trastuzumab, purchased from Roche, was modified with DVP and tetraDVP linker by an initial reduction of the disulfide bridges, the free cysteines generated were then reacted with the synthetic linkers. More specifically, to conjugate with the DVP and tetraDVP linkers, trastuzumab was reduced with TCEP, revealing eight free thiols. Subsequently, the reduced antibody was treated with a slight excess of linkers (~ 20 molar equiv. for DVP, ~ 2 molar equiv. for tetraDVP) for two hours at 37°C[47,49,50]. Removal of small molecule reagents by dialysis.

To generate Fc and Fab domains from full-length IgGs, the unmodified and tetraDVP-linked antibodies were digested with IdeS (FabRICATOR, Genovis), which cuts below the hinge (CPAPELLG / GPSVF), with a 1 unit : 20 µg ratio overnight at 37 °C. This generated Fab (two arms and hinge) and Fc domains (without hinge) which were separated by size-exclusion chromatography run in 50 mM Tris pH 7.5 (Superdex G75 30/100 GL and Superdex 200 30/100 GL, GE Healthcare).

2.2 METHODS

2.2.1 Quality control of the generated protein constructs using liquid chromatography mass spectrometry

The verification of the correct molecular weight was achieved by liquid chromatography mass spectrometry (LC-MS). The liquid chromatography desalting step was carried out on a reverse phase column (Acquity UPLC® BEH300 C4 1.7 µm 2.1 x

50 mm column, Waters Part #186004495), using an aqueous phase A (H_2O, 0.01% TFA and 0.1% formic acid) and an organic phase B (acetonitrile with 0.01% TFA and 0.1% formic acid), running at 0.15 mL/min. For the reduced method, 95% phase A was run for three minutes, followed by a one-minute gradient from 95 to 75% of phase A, then a ten-minute gradient from 75 to 55% phase A, and finally a one-minute gradient from 55 to 20% A (and maintained at 20% phase A for another minute). For the intact method, 95% phase was run for one min, followed by a two-minute gradient from 95 to 20% A, and then by a three-minute gradient from 20 to 5% B. The mass was analyzed either on the mass spectrometer Synapt G2 (Waters) or Q Exactive Orbitrap (ThermoFisher). Synapt G2 was operated in a positive polarity under sensitivity mode, with a capillary voltage of 3.4 kV, source temperature 120 °C, sampling cone at 60 V, desolvation temperature of 400 °C and desolvation gas flow of 800 L/hour. Q Exactive Orbitrap was operated in a positive polarity mode. For both the intact method and the reduced methods, MS data were acquired over a 500-4500 m/z range, using an AGC target of 3e6 and a maximum injection time of 200 ms, with a resolution of 15000. The capillary voltage was 3.5 kV, capillary temperature was 300 °C, auxiliary gas heater temperature was 250 °C. The intact method was run with a funnel RF level of 100, in-source CID of 50 eV, and the reduced method was run with a funnel RF level of 70 and in-source CID of 80 eV. Data acquired on the Synapt G2 was analysed using MassLynx and that acquired on the Q Exactive Orbitrap was analyzed using BioPharmaFinder. The difference between the theoretical and experimental masses was calculated (in ppm) as:

$$\mathbf{Difference} = \frac{(\boldsymbol{experimental\ average\ mass - theoretical\ average\ mass})}{\boldsymbol{theoretical\ average\ mass}} \times \mathbf{1000000} \qquad \textbf{Eq. 2.2}$$

The identification is considered correct if the ppm is between -100 and 100.

2.2.2 Biophysical stability

2.2.2.1 Measurement of the thermodynamic stability using chemical denaturation curves

2.2.2.1.1 Methods

Chemical unfolding and refolding curves in guanidinium chloride were performed. Each experiment was composed of forty-one points (120 µL total volume per point) of increasing concentrations of GdmCl from 0 to higher concentrations ranging from 3.5 to 6 M final depending on the highest $[den]_{50\%}$ of the antibody. 110 µL of

denaturant solution was mixed with 10 µL of protein to a final concentration of 1 µM. For the unfolding curves, the stock protein solution was made in 20 mM histidine buffer pH 5.5 (Chapters 3 and 4), in 55mM potassium phosphate pH 2.5 (Chapter 6), or in 50 mM Tris pH 7.5 (Chapter 7); for the refolding curves, the protein was first denatured in 5 M GdmCl for fifteen minutes at room temperature and then dispensed into the same forty-one-point denaturant solutions. The solutions were dispensed with a liquid handling robot (Microlab®500 Series, ML541C, Hamilton Company). The denaturant solutions mixed with the protein were incubated at 25 °C at different time points until they reached equilibrium (7 days). Each of the forty-one denaturation points were measured in a 100 µL quartz cuvette (Hellma, Precision Cell in Quartz SUPRASIL®, Typ No: 105.250-QS, light path: 10x2 mm, centre: 20 mm). The fluorescence was recorded with a Cary 400 Eclipse Fluorescence Spectrophotometer (Agilent Technologies) thermostatted at 25 °C controlled by a heat block. The samples were excited at 280 nm, the emission was recorded from 300 to 400 nm, with a scan rate of 300 nm min^{-1}, excitation and emission band passes were set at 10 nm.

2.2.2.1.2 Data Analysis

When the protein is denatured with increasing concentrations of chemical denaturant, the maximum of the fluorescence signal shifts towards red wavelengths. In the native state, the wavelength of maximum fluorescence intensity is at approximately 335 nm and in the denatured state, the wavelength of maximum fluorescence is around 360 nm. The data were analyzed using an average emission wavelength (AEW), which is the arithmetic mean of the wavelengths weighted by the fluorescence intensity at each wavelength. It is calculated as shown in **Equation 2.3.**

$$\boldsymbol{AEW = \frac{\sum_{i=1}^{N} F_i . \lambda_i}{\sum_{i=1}^{N} F_i}} \qquad \textbf{Eq. 2.3}$$

where F_i is the intensity of fluorescence at the wavelength i, and λ_i the wavelength.

A two-state model is used to fit data points of systems where only the native and denatured state are significantly populated. The equilibrium between the two states is as follows :

$$\mathbf{N} \underset{}{\overset{K_{D-N}}{\rightleftharpoons}} \mathbf{D} \qquad \textbf{Eq. 2.4}$$

The denaturation curves, using the AEW data, were then fitted with a two-state model:

$$F = \frac{(\alpha_N + \beta_N.[den]) + (\alpha_D + \beta_D.[den]).\exp\left(\frac{m_{D-N}}{RT}.([den] - [den]_{50\%})\right)}{1 + \exp\left(\frac{m_{D-N}}{RT}.([den] - [den]_{50\%})\right)} \qquad \textbf{Eq. 2.5}$$

where α_N, α_D are the fluorescence of the native and denatured states in H_2O respectively; β_N, β_D are the slopes of native and denatured baselines respectively; m_{D-N} is the m-value between the denatured and native state; ΔG_{D-N}^{H2O} is the difference in Gibbs free energy between the denatured and native states; T the temperature and R the gas constant.

A three-state model in which an intermediate state between the native and denatured state is sufficiently stable to be populated and observed was also used. The equilibrium between the different species is as follows:

$$\mathbf{N} \underset{}{\overset{K_{I-N}}{\rightleftharpoons}} \mathbf{I} \underset{}{\overset{K_{D-I}}{\rightleftharpoons}} \mathbf{D} \qquad \textbf{Eq. 2.6}$$

The denaturation curves, using the AEW data, were then fitted with a three-state model[136,137]:

$$F = \frac{\alpha_N + \beta_N.[den] + \alpha_I.\exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right) + (\alpha_D + \beta_D.[den])\,\exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right)\exp\left(\frac{m_{D-I}}{RT}.([den] - [den]_{50\%\,D-I})\right)}{1 + \exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right) + \exp\left(\frac{m_{I-N}}{RT}.([den] - [den]_{50\%\,I-N})\right)\exp\left(\frac{m_{D-I}}{RT}.([den] - [den]_{50\%\,D-I})\right)} \qquad \textbf{Eq. 2.7}$$

where $\alpha_N, \alpha_I, \alpha_D$ are the fluorescence of the native, intermediate and denatured states in H_2O respectively; β_N, β_D are the slopes of native and denatured baselines respectively; m_{I-N}, m_{D-I} are the m-values between the intermediate and native state, and denatured and intermediate states respectively; $\Delta G_{I-N}^{H2O}, \Delta G_{D-I}^{H2O}, \Delta G_{D-N}^{H2O}$ are the differences in Gibbs free energy between the intermediate and native states, denatured and intermediate states, and denatured and native states, respectively; T the temperature and R the gas constant.

The following equation can be used normalize the native and denatured baselines:

$$Norm = \frac{(AEW - (\alpha_N + \beta_N.[den])}{(\alpha_D + \beta_D.[den]) - (\alpha_N + \beta_N.[den])} \qquad \textbf{Eq. 2.8}$$

2.2.2.2 Measurement of the kinetic stability with stopped-flow

2.2.2.2.1 Methods

For the stopped-flow experiments, the native protein and the denaturant solutions were mixed in a 1:10 ratio respectively. Seven stock solutions of guanidinium chloride (GdmCl) were prepared in 20 mM histidine buffer pH 5.5 (Chapters 3 and 4) or in 50 mM Tris pH 7.5 (Chapter 5) so that the final concentrations range from 5.5 to 7.0 M GdmCl with an interval of 0.25 M. The stock solutions are 1.1-fold more concentrated than the final solutions, as solutions were diluted by a factor of 10/11 in the rapid mixing step, using 500 µL and 2.5 mL Hamilton syringes. The protein stock solutions were prepared between 7 and 11 µM to achieve a final concentration between 0.5 to 1 µM after mixing with the denaturant solutions.

To replicate HDX-MS quench conditions, native protein and HDX quench buffer (100 mM potassium phosphate, 4-8 M GdmCl or 1-2 M GdnSCN, pH 2.5) were mixed in a 1:1 ratio, using two 2.5 mL Hamilton syringes.

The unfolding kinetics were monitored with a SX20 stopped-flow spectrometer from Applied Photophysics (software: SX Spectrometer Control Panel Application version 2.2.27). The temperature of the water bath was set to 25 °C (Chapters 3, 4 and 7), and between 5 and 25 °C, the excitation wavelength was set to 280 nm, both slit widths were 2 mm. A cut off filter of 320 nm was used. For fast unfolding events, three short time traces were recorded with pressure hold (2 s), to accurately measure the fastest unfolding phase, and three longer time traces from 0-30 or 0-300 s were acquired to accurately measure the slower unfolding phase, both at each denaturant concentration.

2.2.2.2.2 Fitting of unfolding kinetics

The unfolding curves were fitted with the software Pro DataViewer version 4.2.27. Depending on the recorded traces and the number of domains unfolding, the fluorescence signal corresponding to the unfolding was fitted with a single (**Equation 2.9**), double-exponential (**Equation 2.10**) or triple-exponential function (**Equation 2.11**).

$$A(t) = A \cdot \exp(-kt) + c \qquad \textbf{Eq. 2.9}$$

$$A(t) = A_1 \cdot \exp(-k_1 t) + A_2 \cdot \exp(-k_2 t) + c \qquad \textbf{Eq. 2.10}$$

$$A(t) = A_1 . \exp(-k_1 t) + A_2 . \exp(-k_2 t) + A_3 . \exp(-k_3 t) + c \quad \textbf{Eq. 2.11}$$

where A_1 , A_2 and A_3 are the amplitudes, k_1 ,k_2 and k_3 the respective unfolding rate constants and c the offset.

The natural logarithm of the rate constants was then calculated and plotted versus the corresponding denaturant concentration, and the data fitted with **Equation 2.12** [138].

$$\ln k_U^{[den]} = \ln k_U^{H_2O} + m_{k_U}[den] \quad \textbf{Eq. 2.12}$$

Where $k_U^{[den]}$ is the observed unfolding rate constant at the denaturant concentration $[den]$, $k_U^{H_2O}$ is the unfolding rate constant in water and m_{k_U} is the slope of the plot of $\ln k_U^{[den]}$ versus denaturant concentration.

The percentage residual can be calculated with **Equation 2.13**:

$$\% residual = \frac{amplitude\ of\ residual}{total\ amplitude} \times 100 \quad \textbf{Eq. 2.13}$$

2.2.2.3 The half-life of unfolding can be calculated with the following equation:

$$t_{1/2} = \frac{ln(2)}{k} \quad \textbf{Eq. 2.14}$$

2.2.2.4 Thermal stability

2.2.2.4.1 Methods

The thermal denaturation of proteins was monitored by differential scanning calorimetry (DSC), with a Malvern MicroCal VP-DSC instrument. 500 µL of protein at 0.5-5 mg mL^{-1} in 20 mM histidine pH 5.5 (Chapter 3 and 4) or in 50 mM Tris pH 7.5 (Chapter 5) were used for each run. For each protein, the baseline was measured first, which consists of buffer in both cells (buffer versus buffer), and then the protein was run (buffer versus protein). Several clean-up cycles with water and suitability controls with lysozyme at 3 mg mL^{-1} in water were employed before, and after, the actual experiment with the antibody. Each protein sample was scanned twice to investigate the thermal reversibility. The temperature was ramped from 25 to 100 °C, increasing by 95 °C h^{-1}. The pre-scan thermostat was set to 2 min, no post-scan thermostat was employed.

2.2.2.4.2 Data Analysis

The data were processed with the Origin version 7.0 SR4 software. The baseline thermogram (buffer versus buffer) was subtracted from the thermogram of the protein (buffer versus protein). Two baselines, one at the beginning and one at the end of the thermogram, were placed to adjust the data, which was then normalized using the concentration of the protein. The unfolding peaks were selected and the thermogram was fitted to the "Non 2-state" model to obtain the melting temperatures (T_m) and the enthalpy of unfolding at the T_m, ΔH_{cal}.

2.2.2.5 Far UV-CD

The far UV-CD signal was measured on a Chirascan CD Spectrometer (Applied Photophysics), with a 350 µL quartz cuvette for the far UV-CD measurements (Hellma, Absorption Cell in Quartz Glass High Performance, Article No: 110-1-40, Optical path length: 1 mm, Centre height: 8.5-20 mm), thermostatted at 25 °C controlled by a heat block. The UV-CD absorbance was measured from 250 to 180 nm with a bandwidth of 2 nm, for 1 second per sample, in triplicates. The buffer baseline was also measured and subtracted from the protein signal.

2.2.3 Hydrogen deuterium exchange mass spectrometry

The HDX-MS experiments were employed in i) in Chapter 4, on 2xSH, 2xCys and iDSB Fc-C239i, as well as NIST mAb; ii) in Chapter 5 on full-length IgG unmodified, DVP and tetraDVP linked trastuzumab, iii) in Chapter 6, on full-length IgGs T, T-scFv40, T-scFv69, T-dAb, T-scFvGipg013 both intact and digested by TEV and/or Factor Xa, except for T-scFvGipg013, as well as on dAb, scFv40 and scFv69.

The peptide map was generated i) in Chapter 4, on 2xCys Fc-C239i in equilibration buffer 20 mM His pH 5.5, ii) and iii) in Chapters 5 and 6, on trastuzumab, in equilibration buffer 10 mM potassium phosphate pH 7.5 in H_2O, using a data-dependent acquisition (DDA) MS^2 approach. MS data were acquired using an Orbitrap Fusion (ThermoFisher) over a 300-2000 m/z range, using an AGC target of 200000 and a maximum injection time of 100 ms. MS^2 was acquired in the ion trap in centroid mode, using an AGC target of 10000 and a maximum injection time of 35 ms. Fragmentation was achieved by HCD using a collision energy of 30%.

The labelled data for Chapter 4 were recorded after after 1000, 6000, 30000, 60000, 300000, 600000, and 900000 ms of incubation at 20 °C in deuterated buffer (in 20 mM His pD 5.5 in D_2O), using the ms2 min HDX system[139] (patent WO2020074863A1): 10 µL of protein at 5 µM in equilibration buffer (20 mM His pH 5.5 in H_2O) were diluted 20-fold at 20°C into equilibration or labeling buffer (20 mM His pH 5.1 (pD 5.5) in D2O) to generate a peptide map or labelled data, respectively. The mixture was then diluted 1:1 with quench buffer at 2 °C (100 mM His, 8 M urea, 0.5 M TCEP, pH 2.5). The labelled data for Chapters 5 and 6 were recorded after 50, 500 and 5000 s of incubation at 20 °C in deuterated buffer (10 mM potassium phosphate pD 7.5 in D_2O). On the Leap robot, 7 µL of 10 µM protein in equilibration buffer (10 mM potassium phosphate pH 7.5 in H_2O) was diluted seven-fold at 20 °C into equilibration or labelling buffer (10 mM potassium phosphate pD 7.5 in D_2O), to generate a peptide map or labelled data, respectively. The mixture was then diluted 1:1 with quench buffer at 4 °C (100 mM potassium phosphate, 8 M urea, 0.5 M TCEP, pH 2.5). The quench solution was then injected into a Waters nanoAcquity UPLC system, flowing for four min at 400 µL/min onto the pepsin column at 20 °C for digestion (Waters Enzymate™ BEH Pepsin Column (2.1 x 30 mm, 5 µm)) to the trap (pushed by LC-MS grade H_2O with 0.2% formic acid, and then eluting from the C18 analytical column for 10 minutes, using a 5% to 40 % organic phase (ACN with 0.2% formic acid).

The mass spectrometry data were acquired with a MS^1 method (300-2000 m/z range, AGC target of 200000 and maximum injection time of 100 ms). All experiments were run in triplicate.

A peptide map was generated using BioPharmaFinder from equilibration data acquired with a MS^2 method. The recognition of the peptides was based in the fragmented b and y ions, to ensure confidence in the identification, and around the N-linked glycosylation in the C_H2 domain, the peptides including the asparagine at position 297 (in Eu numbering) were expected to carry the most abundant glycosylation expressed in CHO cells, i.e. G0F. The peptides obtained were filtered by confidence score higher than 80%. The exported csv file with the peptides as well as that same non-deuterated data were imported into HDExaminer, to operate a second filtration of the peptides: only the charge state with the highest intensity from the peptide map data was kept per peptide for comparative accuracy between the charge states, selected according to the highest intensity, the sharpest extracted ion chromatogram. After the peptide pool was curated,

the labelled data were added, and the D incorporation per peptide data was then exported as a csv file. To identify which deuterium incorporations were significant and to observe the deuterium exchange for each time point separately and overall, the data processing method used was first described by Dobson, Devine, Phillips *et al*., 2016[104]. Starting with a csv file containing the D incorporation data, this Matlab-coded method first assesses if the incorporation of deuterium per peptide is significant compared to the control sample (assessed by a t-test, if two-tailed p-value < 0.05 for Chapter 4, p-value < 0.01 for Chapters 5 and 6), then sums the significant time points per peptide, subsequently converts the peptide D incorporation to amino-acid incorporation by dividing the D incorporation by the maximal number of D that can be exchanged per peptide, subtracting the D incorporation from the control and dividing the amino-acid incorporation by the redundancy and finally normalizing it. The output is the deuterium incorporation for each individual amino acid. Crystal structures of the Fc wild-type (PDB: 3AVE), Fab domain of trastuzumab (PDB: 1N8Z), and the nanobody, solved by a colleague David Hargreaves, already existed, but no structures for scFv40 and scFv69 exist. These structures were generated by homology modelling with a sequence that had one of the highest similarities on PDB (PDB: 3AUV, 60.6 % with scFv40, 62.3 % with scFv69), using the software Maestro.

The crystal structures were coloured with a red-white-blue scale according to the relative incorporation of deuterium per amino acid, red representing deprotection and blue corresponding to protection, compared to the reference.

2.2.3.1 Non reduced peptide mapping

Four experiments were carried out using non-reduced peptide mapping in Chapter 4, with slightly different conditions for each. Two were undertaken with the standard non-reduced peptide mapping LC-MS method ((a) identification of iDSB and (b) quantification of the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times) and two were performed using the targeted non-reduced peptide mapping LC-MS method ((c) quantification of the enriched thiol states and (d) measure of the evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time). These experiments were done by Alistair Hines, Matthew Edgeworth and myself.

2.2.3.1.1 Preliminary sample incubations

To quantify the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times (b), and to measure the evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time (d), preliminary sample preparation was necessary.

2.2.3.1.1.1 Preliminary sample incubation for the quantification of the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times (b)

For each of the three variants, 100 µg were prepared in 34 µL, in 0 or 3.5 M GdmCl, in 20 mM His pH 5.5. They were incubated for 0 or 7 days at 25 °C.

2.2.3.1.1.2 Preliminary sample preparation to measure evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time (d)

The first set of experiments consisted in observing the effect of denaturant concentration after seven days of incubation at 25 °C on the inserted-cysteine states. For this first set of experiments, 100 µg of the free-thiol enriched and doubly-capped cysteine enriched variants were prepared in 0, 0.5, 1, 1.5, 2, 2.5, 3 and 3.5 M GdmCl, in 20 mM His pH 5.5 in a total volume of 34 µL. They were incubated seven days at 25 °C.

The second set of experiments aimed at assessing the effect of incubation time at 25 °C in 3.5 M GdmCl on the inserted-cysteine states. For the second set of experiments, 100 µg of the free-thiol enriched and doubly-capped cysteine enriched variants were prepared 3.5 M GdmCl, in 20 mM His pH 5.5 in 34 total µL. They were incubated for 0, 1, 2, 3 and 4 days at 25 °C.

2.2.3.1.2 Non-reduced peptide map preparation: NEM capping, denaturation and Lys-C digestion

2.2.3.1.2.1 Sample preparation for identification of iDSB (a) and quantification of the enriched thiol states (c)

50 µg of Fc-C239i variants were alkylated using 2.5 µg of N-ethyl maleimide (NEM) at room temperature for 20 minutes. Samples were then dried using a rotary evaporator to remove excess NEM. Protein was reconstituted in 45 µL of a denaturing solution (7.1 M GdmCl, 5.6 mM sodium phosphate pH 7.0,0.1 M NaCl) incubated at 37 °C for 30 minutes. The solution was diluted in 125 µL of 100 mM sodium phosphate Buffer, pH 7.0 and 0.5 µL of 40 mM EDTA. 2.5 µg of Endoproteinase Lys C were added to the sample and incubated at 37 °C for two hours. A further 2.5 µg of Endoproteinase Lys C were added and incubated for a final two hours.

2.2.3.1.2.2 Sample preparation for the quantification of the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times(b)

After the incubation, the samples were alkylated by adding 5 µg of N-ethylmaleimide (NEM), and incubated at room temperature for 20 min. Samples were then buffer exchanged three times into 7.1 M GdmCl, 5.6 mM phosphate pH 7.0 and 0.1 M NaCl with 10 kDa Amicon filters, and concentrated to a final volume of 90 µL. 250 µL of 100 mM phosphate buffer pH 7.0, 1 µL of 40 mM EDTA and 10 µL of 1 mg/mL of Endoproteinase Lys-C were added to each sample, and incubated for two hours at 37 °C. 10 µL at 1 mg/mL of Lys-C were added again, and incubated for another two hours at 37 °C. After digestion, the material was split in half. 5 µL of 500mM DTT was added to 45 µL of digested protein and incubated at room temperature to reduce. 5 µL of water was added to 45 µL of the same sample to act as a non-reduced control. Both samples were run side-by-side by LC-MS for comparative analysis.

2.2.3.1.2.3 Sample preparation to measure the evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time (d)

After the incubation, the samples were alkylated adding 5 µg of N-ethylmaleimide (NEM), and incubated at room temperature for 20 min. The samples were buffer exchanged three times into 7.1 M GdmCl, 5.6 mM sodium phosphate pH 7.0 and 0.1 M NaCl with 10 kDa Amicon filters, and concentrated to a final volume of 60 µL, and incubated at 37°C for 30 minutes. From this volume, 15 µL were taken (25 µg of antibody) and were then diluted four-fold in 100 mM $NaPO_4$ pH 7.0 containing 0.4% 40 mM EDTA, diluting the guanidinium chloride concentration to 1.8 M in preparation for digest. An aliquot of 10 µL (0.5 µg) of Lys-C was added to ~ 25 µg of thiol capped denatured protein at a 1:50 enzyme:protein ratio (10 and the mixture was incubated at 37 °C for two hours. A further 10 µL (0.5 µg) of Lys-C was added and incubated for a final two hours.

2.2.3.1.3 LC-MS parameters

2.2.3.1.3.1 Standard non-reduced peptide mapping LC-MS

The following parameters were used for the experiments (a) identification of iDSB and (b) quantification of the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times.

The digested samples were analysed by means of reverse-phase liquid chromatography (Acquity i-Class UPLC, Waters, Manchester, UK) coupled to mass spectrometry

(Orbitrap Fusion ThermoFisher mass spectrometer). Reduced and non-reduced samples were compared to identify disulfide-containing peptides, and identification was performed using a combination of MS1 and MS2. Peptides were separated using a Peptide BEH C18 column, 300 Å 1.7 µm, 2.1 mm x 150 mm (Waters, Manchester, UK) over a 76-minute linear gradient of 5-45% B (mobile phase A: 0.02 % TFA in water; mobile phase B: 0.02% TFA in ACN). MS data were acquired over a 250-2000 m/z range, using an AGC target of 200000 and a maximum injection time of 50 ms. MS2 was acquired in the Ion trap in Centroid mode, using an AGC target of 10000 and a maximum injection time of 35 ms. Fragmentation was achieved by collision induced dissociation (CID) using a collision energy of 35%. The data were then processed with the Qual Browser Thermo Xcalibur 3.0.63 software.

2.2.3.1.3.2 Targeted non reduced peptide map

The following parameters were used for the experiments (c) quantification of the enriched thiol states and (d) measure of the evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time.

Mobile phase A contained 0.1% TFA in water and mobile phase B contained 0.1% TFA in 100% Acetonitrile. The following LC conditions were used: flow rate 0.15 mL/minute, column temperature 55 °C throughout the separation with the autosampler maintained at 4 °C. Injections of 10 µL of 0.42 mg/mL peptide sample were separated using a UPLC Peptide CSH C18, 130Å pore size, 1.7 µm bead size, 2.1 mm X 150 mm column (Acquity). The gradient started at 0% B until a step up to 24% at four minutes and then gradually increased to 26% at eight minutes, followed by a step up to 80%. The gradient was dropped back to 100% A to equilibrated for the next injection. Total run time per samples was 12 minutes. On a Waters Xevo TQS, the MS desolvation temperature was 600 °C, source cone voltage was 40 V, desolvation gas flow 400 L/hr, capillary voltage 3 kV, Q3 collision energy 40 eV. The different variants were targeted according to the values referenced in **Table 2.1**.

Table 2.1: Chromatographic and m/z parameters for the identification of the variants

Compound	RT (minutes)	m/z precursor ion	m/z fragment ion
iDSB	5.70 +/- 0.20	1416.50	566.30
2xCys	5.50 +/- 0.20	1476.70	566.30
1xSH + 1 Cys	6.20 +/- 0.22	1478.20	566.30
2xSH	8.00 +/- 0.30	1479.70	566.30

2.2.3.1.4 Data analysis

2.2.3.1.4.1 Data analysis for the identification of iDSB (a)

The MS2 fragmentation data obtained by CID for the peptide in the hinge, corresponding to a +4 ion (1415.71 m/z) was analysed to find fragmented b and y ions. Peptides linked by a disulfide bridge between the inserted cysteines 239 were identified and the amino acid sequence was reconstituted amino acid by amino acid: they are composed of a neutral peptide truncated at the N-terminus only (resulting from a fragmentation that generated a y ion and a neutral peptide, as the charges were retained on the N-terminus fragmented peptide[1]), and a + 1 charged peptide truncated both at the N-terminus, as a neutral fragment (same explanation as above), as well as at the C-terminus as a b ion (after the fragmentation event, the charge was retained at the N-terminus).

2.2.3.1.4.2 Data analysis for the quantification of the thiol states of Fc-C239i enriched variants in specific denaturant concentrations and incubation times (b)

The peptides corresponding to the hinge region were searched (NEM-capped, cysteine-capped and disulfide bridged). The presence of the variants was quantitated by combining the most intense isotope on the 4+ m/z and 5+ m/z distribution, and then integrating the area under the obtained peaks on the total ion count chromatogram.

2.2.3.1.4.3 Analysis of targeted non reduced peptide map

This data analysis approach was used for the (c) quantification of the enriched thiol states and (d) measure of the evolution of thiol states upon increasing concentrations of chemical denaturant and incubation time.

A targeted MS2 analysis was performed on the 1415.71 (4+) ion, previously identified as the non-reduced 4+ hinge peptide (THTCPPCPAPELLG-GPSCVFLFPPKPKTHTCPPCPAPELLGGPSCVFLFPPKPK), and the different variants were quantified following the identification parameters in **Table 2.1**.

2.2.4 Size-exclusion chromatography couple to multi angle light scattering (SEC-MALS)

For SEC-MALS experiments, carried out by Wesley Howard, samples were prepared in PBS pH 7.4 at 0.25-0.5 mg/mL, and approx. 1-20 ug were injected onto a column equilibrated and run in PBS pH 7.4. The SEC column was a

TSKgel® G3000SWXL HPLC Column phase diol, L × I.D. 30 cm × 7.8 mm, 5 µm particle size (Tosoh Bioscience, Germany). The experiment was conducted using Agilent Technologies 1260/1290 Infinity and Wyatt DAWN-Optilab instruments. The refractive index and multiangle-light scattering data were analyzed using the Astra software version 7.3.2.

2.2.5 Activity assays

2.2.5.1 BioLayer Interferometry

2.2.5.1.1 Measurement of the affinity of the masks (scFv40, scFv69 and dAb) for trastuzumab by BioLayer Interferometry (BLI)

The affinity of the masks scFv40, scFv69 and dAb for trastuzumab were measured by BioLayer Interferometry (BLI) with the OctetRED384 (ForteBio). All samples were prepared in DPBS (Sigma-Aldrich) with 0.1 % BSA (bovine serum albumin solution, 30% BSA in DPBS, sterile-filtered, BioXtra, Sigma-Aldrich) and 0.02% Tween 20 (TWEEN® 20 for molecular biology, Sigma-Aldrich). Serial dilutions were prepared for each of them: for scFv40, 46, 92, 184, 369, 738 nM; for scFv69, 25, 100, 200, 300, 600 nM; for dAb, 1.7, 3.4, 6.9, 13.7, 27.5 nM. Anti-human IgG Fc Capture (AHC) biosensors (Forte Bio) were used to immobilise trastuzumab at 5 µg/mL Before starting the BLI experiment, the biosensors were incubated in DPBS with 0.1% BSA and 0.02% Tween 20 for ten min. For the measurements, the biosensors were first incubated for one minute in DPBS with 0.1% BSA and 0.02% Tween 20 for the first baseline, trastuzumab was then loaded onto the biosensors until reaching a displacement of 0.8-1 nm, the biosensors were then incubated in DPBS with 0.1% BSA and 0.02% Tween 20 for the second baseline for one minute, and subsequently moved to wells containing the serial dilutions of the different masks for the association step, and finally moved to DPBS with 0.1% BSA and 0.02% Tween 20 for the dissociation step. The association and dissociation results were fit to a 1:1 model using a global fitting.

2.2.5.1.2 Measurement of the binding of the intact and digested masked antibodies for HER2 by BioLayer Interferometry (BLI)

The binding of masked antibodies, intact and digested, to HER2 was measured by BLI. The antibodies T-scFv40, T-scFv69 were digested with either TEV or factor Xa, or both enzymes together. T-dAb was only digested by TEV. The effective digestion was verified by SDS-PAGE prior to the BLI experiments. T, T-scFv40, T-scFv69, T-scFvGipg013, NIP228 (used as a negative control[140]) were used in their intact form. All

samples were prepared at 2.5 µg mL^{-1} in in DPBS with 0.1% BSA and 0.02% Tween 20. Biotinylated HER2 (Her2/ERBB2 protein, human, recombinant (His & AVI Tag), Biotinylated, Sino Biological) was prepared at 5 µg mL^{-1} in DPBS with 0.1% BSA and 0.02% Tween 20.

Streptavidin (SA) biosensors (ForteBio) were used to immobilise the biotinylated HER2. The assay was carried out with the following steps: Baseline 1: incubation in DPBS with 0.1% BSA and 0.02% Tween 20 for 1 min; Loading: incubation in solution of biotinylated HER2 for binding to the biosensors for 1 min; Baseline 2: incubation of the biosensors in DPBS with 0.1% BSA and 0.02% Tween 20 for 1 min; Association: incubation of the biosensors with solutions containing intact and digested masked antibodies for 300 s; Dissociation: incubation of the biosensors DPBS with 0.1% BSA and 0.02% Tween 20 for 300 s.

2.2.5.2 Flow cytometry to measure the binding of the intact and digested masked antibodies for HER2 expressed on SK-BR-3 cells

SK-BR-3 cells, breast cancer cells overexpressing HER2, were grown in culture. At passage number 8 and 80% confluency the cells were harvested using standard accutase (Gibco) treatment. The cells were washed with DMEM (Gibco) 1% FBS (Fetal Bovine Serum heat inactivated, Gibco) 1% Pen Strep antibiotic (Penicillin-Streptomycin (10,000 U/mL), Gibco) and resuspended in DPBS 2% BSA. 100,000 cells/well were dispensed in 96-well plates (PP V-bottom, Chimney Well, natural, Greiner Bio-one International). The intact antibodies T, T-scFv40, T-scFv69, T-dAb, T-scFvGipg013, NIP228, as well as the digested antibodies T-scFv40 + TEV, T-scFv40 + Factor Xa, T-scFv40 + TEV + Factor Xa, T-dAb + TEV (and T +T EV, T + Factor Xa, T + TEV + Factor Xa as controls) were diluted to concentrations ranging from 0.4 to 1600 nM in DPBS 2% BSA, and incubated with the SK-BR-3 cells for 30 min at 4 °C. The cells were then washed with DPBS 2% BSA once, resuspended in 100 µL, and the secondary antibody Alexa Fluor® 647 anti-human IgG Fc (BioLegend) was added to the cells at 0.5 µL/100,000 cells for 30 min at 4 °C. The cells were washed once with DPBS 2% BSA and fixed with 4% paraformaldehyde in DPBS for 1 hour at room temperature. Finally, one additional washing step was performed prior to resuspension in DPBS 2% BSA. All experiments were conducted in triplicate.

The binding of the various antibodies was detected monitoring the fluorescence of the secondary antibody Alexa Fluor® 647 anti-human IgG Fc using a BD

LSRFortessa™ Flow Cytometer. The first gate (P1) selected singlets, based on the forward scatter height (FSC-H: forward scatter height) and the forward scatter area (FSC-A: forward scatter area). The second gate (P2) selected the viable cells according to the side scatter area (SSC-A) and the FSC-A. Subsequently, the fluorescence intensity of APC (channel to detect Alexa 647) was measured.

The data points were fitted to an agonist *versus* response fit using GraphPad Prism 9 and **Equation 2.15**:

$$Y = a + X \times (b - a)/(EC_{50} + X) \qquad \textbf{Eq. 2.15}$$

Where a and b are the upper and lower plateaux of the curves, X the concentration of antibody and EC_{50} the half maximal effective concentration.

2.2.5.3 High-thouput confocal microscopy to measure the binding of the intact and digested masked antibodies for HER2 expressed on SK-BR-3 cells

SK-BR-3 cells were plated at $2x10^4$ per well of 96-well plate (CellCarrier-96 Ultra Microplates, PDL-coated, black, 96-well, Perkin Elmer) in HBSS (Hank's Balanced Salt Solution, Gibco) 1% FBS (Fetal bovine serum) and were incubated over night at 37°C. The next day, cells were incubated with primary antibodies for 30 min at 4 °C: intact antibodies T, T-scFv40, T-scFv69, T-dAb, T-scFvGipg013, NIP228, as well as the digested antibodies T-scFv40 + TEV, T-scFv40 + factor Xa, T-scFv40 + TEV + factor Xa, T-dAb + TEV were diluted to concentrations 0.4, 1.2, 3.7, 11.1, 33.3, 100 nM in HBSS (Hank's Balanced Salt Solution, Gibco) with 1% FBS (fetal bovine serum). After washing to remove excess antibody, cells were fixed with 4% paraformaldehyde for one hour at room temperature. The cells were then permeabilised and stained by incubation with a mixture of HCS CellMask® Orange Stain (ThermoFisher) diluted to 1:25000, Hoechst stain (33342, Trihydrochloride, Trihydrate, ThermoFisher) diluted to 1:5000 and 0.1% Triton (Sigma-Aldrich) in HBSS 1% FBS, and incubated for 30 min at room temperature. After washing, the secondary antibody Alexa Fluor® 488 goat anti-Human IgG (ThermoFisher) was added to the cells for 30 min at room temperature (60 µL per well of 2 µg/mL), then washed and imaged. All experiments were done in duplicate, and one image was shown per condition.

The microscopy measurements were done by Lorraine Irving. High–throughput confocal microscopy was performed using the Opera (PerkinElmer) with filters, and

exposure times according to the manufacturer's instructions using 20X water objective. The same contrast was used for all images.

2.2.6 X-ray crystallography

The X-ray crystallography experiments were carried out by David Hargreaves. Crystals of T-dAb Fab were grown by sitting drop vapor diffusion. The crystallization reagent was 0.2 M ammonium sulphate, 0.1 M sodium acetate (pH 4.6) and 25% w/v PEG 4000. 200 nL of the protein solution of T-dAb Fab was mixed with 200 nL of the reagent and equilibrated over a reagent well containing 50 µl at 20°C. Crystals with a rod-like morphology (100 µm) appeared in the drops after 13 days. Crystals were cryo protected using 23% butane-2,3 diol added to the well solution and snap frozen in liquid nitrogen. X-ray data were collected on beamline IO4 at the Diamond Light Source. Data processing was performed using autoPROC[141] and STARANISO (Global Phasing Ltd) and showed acceptable data extending to 2.35 Å. Molecular replacement was achieved using PHASER[142] and a trastuzumab model (PDB 4HJG) with a nanobody model (PDB 6DBE). Coot[143], Refmac[144] and Buster[145] (Global Phasing Ltd) were used during model building and refinement. The final model was in space group P 2_1 2_1 2 with cell dimensions of a=160 Å.7, b= 59.9 Å, c=63.5Å α=β=γ=90° and showed a single T-dAb Fab fusion in the asymmetric unit.

3 STABILITY OF VARIOUS ANTIBODY MUTANTS

3.1 INTRODUCTION

Therapeutic antibodies trace their roots from the immune system and, for the vast majority, more specifically from the immunoglobulin G isotype, and the subclass IgG1[146]. They are therefore naturally highly stable, given they operate at 37 °C. Over the past 20 years, many mutants have been developed in order to confer new properties to the antibody. This includes enhancing biological properties for mechanism occurring *in vivo*, generating new functionalities on the antibody scaffold such as those needed for site-specific conjugation and high control of the drug to antibody ratio (DAR). One such mutant (TM), consisting of a triple-mutation of L234F/L235E/P331S in the Eu numbering, with all residues located in the C_H2 domain (**Figure 3.1**), results in a decrease in affinity for the Fc γ-receptor type IIIA (FcγRIIIA) and reduces antibody-dependent cellular cytotoxicity (ADCC)[106,147,148], i.e. can reduce immune clearance and avoid surrounding cell toxicity caused by the immune system triggered by binding to Fc -receptor type IIIA. In addition to these three mutations, another construct contains a substitution in the C_H3 domain, S442C, which was engineered to enable conjugation to a payload[34] (**Figure 3.1**). While these mutations are desirable and create a new functionality in the antibody, it is important to measure the cost of these mutations, if any, to the intrinsic stability of the antibody.

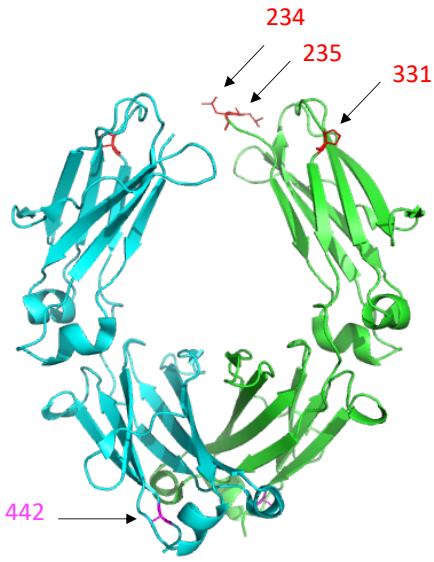


Figure 3.1: Structure of the Fc domain of NIST mAb Fc with the sites of the three mutations in TM Fc highlighted, as well as S442C. The figure was made using PyMol with the coordinates in the PDB file accession name 3AVE corresponding to NIST mAb Fc. The residues in the C_H2 domain that are mutated in TM Fc are represented in red. The two leucines in positions 234 and 235 are missing from the cyan chain as they are located in the very upper part of the C_H2 domain which is not resolved in the crystal structure. The substitution site of S442C in the C_H3 domain is represented in purple.

The objective of this Chapter is to assess the impact of various mutations on the stability of the antibody. Four antibody molecules were investigated: full-length IgG NIST mAb, and three Fc domains: NIST mAb Fc (the wild type), TM Fc (triple mutation in the C_H2 domain L234F/L235E/P331S) and TM S442C Fc (triple mutation in the C_H2 domain L234F/L235E/P331S as well as S442C substitution in the C_H3 domain).

To understand the impact of the mutations on the stability of the antibody, three different measures of stability were determined. The first was the thermal stability, measured by differential scanning calorimetry, which is influenced by the difference in enthalpy (ΔH_{D-N}), difference in entropy (ΔS_{D-N}) and difference in heat capacity (ΔC_p) between denatured and native states. The second is the thermodynamic stability, probed using chemical denaturants, which provides information on the difference in Gibbs free energy between the native and denatured state in water (ΔG_{D-N}), usually at 25 °C. It is also related to the difference in enthalpy and entropy between the native and denatured states, although these values are not measured directly in this experiment. The third is the kinetic stability, which measures the activation energy i.e., the difference in energy between the native and transition states (ΔG_{TS-N}).

3.2 RESULTS

3.2.1 Thermal stability

The thermal stability was measured using differential scanning calorimetry (DSC), which is a technique that probes the heat capacity of proteins and provides the melting temperatures (T_m) and the difference in enthalpy between the native and denatured states (ΔH_{D-N})[149]. The experiments were carried out on NIST mAb full-length IgG, as well as the three Fc domains: NIST mAb Fc, TM Fc and TM S442C Fc. The thermogram of NIST mAb IgG showed three peaks, corresponding to the thermal unfolding transitions of the C_H2, C_H3 and Fab domains in increasing order of melting temperature. The assignment of the unfolding of the C_H2 and C_H3 domains to the two peaks at lower temperatures comes from the studies of Tischenko *et al*.[150] The melting temperature of Fab domains can vary depending on the antibody, however, the unfolding of the Fab domain generally has a maximal C_p value about three times greater than those of the C_H2 and C_H3 domains due its size.[96] Consistent with this, the unfolding peak observed at the highest temperature was attributed to the Fab domain.

Melting temperatures and enthalpy differences obtained from the fitted data are shown in **Table 3.1**. All samples were denatured a first time up to 110 °C, then cooled down to 25 °C, and the difference in heat capacity was measured again, **Figure 3.3**. None of them showed reversible unfolding because the antibodies aggregated after unfolding, consistent with previous studies which reported irreversible aggregation of IgGs after heating[151]. Therefore, the melting temperatures obtained are lower than the actual value of T_m, and should be viewed as apparent measures of stability[152]. This is because the unfolded state irreversibly aggregates, and therefore the denatured population is depleted and this drives the N → D transition, which decreases the melting temperature.

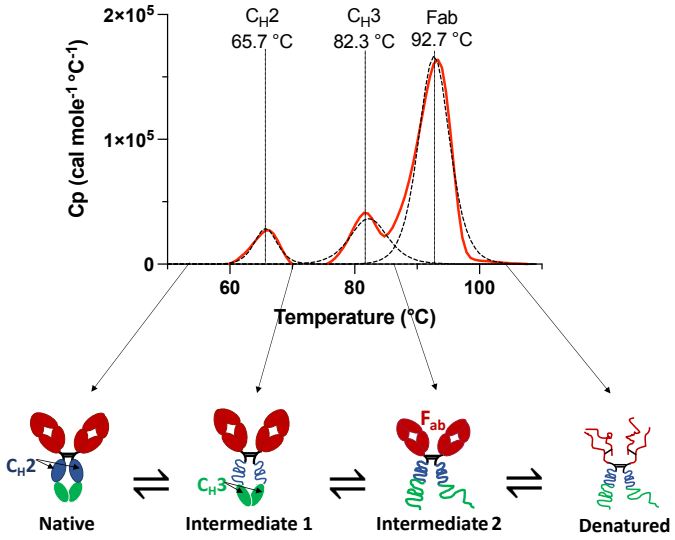


Figure 3.2: Thermogram of NIST mAb full-length IgG and schematic representation of the unfolding of the C_H2, C_H3 and Fab domains at pH 5.5. At the melting temperature (the apex of the peak), 50% of the domain of interest has unfolded. The DSC trace is shown in red and the fitting to the different peaks was plotted as black dashed lines. Underneath the thermogram is the schematic where the C_H2 domain is shown in green, C_H3 in blue and the domains within the Fab in red. The measurement was done in 20 mM His pH 5.5, at 0.5 mg/mL.

Table 3.1: Thermodynamic parameters obtained from the DSC measurements

Antibody mutants	T_m1 (°C)	T_m2 (°C)	T_m3 (°C)	$\Delta H1_{cal}$ (kcal mol^{-1})	$\Delta H2_{cal}$ (kcal mol^{-1})	$\Delta H3_{cal}$ (kcal mol^{-1})
NIST mAb IgG	65.7 ± 0.2	82.3 ± 0.3	92.7 ± 0.1	(1.4 ± 0.2) x 10^2	(3.0 ± 0.2) x 10^2	(11.1 ± 0.2) x 10^2
NIST mAb Fc	63.9 ± 0.1	79.6 ± 0.1		(2.4 ± 0.5) x 10^2	(1.72 ± 0.05) x 10^2	
TM Fc	62.7 ± 0.1	82.7 ± 0.1		(1.07 ± 0.03) x 10^2	(0.92 ± 0.03) x 10^2	
TM S442C Fc	58.6 ± 0.1	79.3 ± 0.1		(0.66 ± 0.02) x 10^2	(0.54 ± 0.02) x 10^2	

The errors are fitting errors, not standard deviations from multiple measurements. From other DSC data where repeats have been measured, standard deviations are typically similar to the fitting errors.

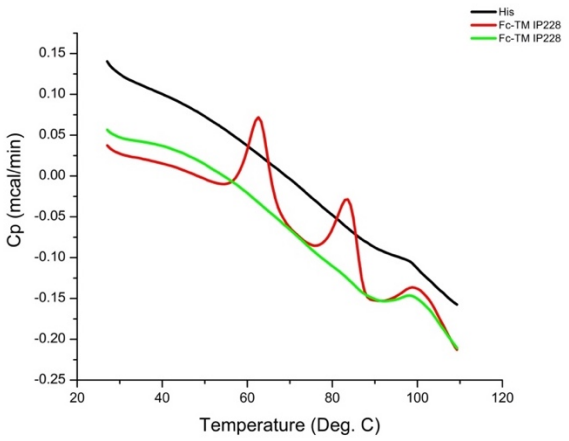


Figure 3.3: Raw data from the DSC experiment on the thermal unfolding of TM Fc. The sample was heated a first time (red), cooled and reheated (green). The black line is the buffer baseline, 20 mM His pH 5.5.

For NIST mAb IgG, the C_H2 domain is therefore the least thermally stable, with a melting temperature of approx. 66 °C, followed by the C_H3 domain ($T_m \approx 82$ °C) and the Fab domain the most stable ($T_m \approx 93$ °C, **Figure 3.2**, **Table 3.1**). The data establish that the Fab domain is extremely stable and that the four domains (V_H, V_L, C_H1 and C_L) all unfold cooperatively.

All Fc domains showed two thermal unfolding transitions, the first corresponding to the unfolding of the C_H2 domain and the second to the unfolding of the C_H3 domain (**Figure 3.4**). The lowest melting temperature for the C_H2 domain was observed for TM S442C Fc, whereas TM Fc had a T_m only slightly lower than that of NIST mAb Fc. This is rather surprising given that previous DSC studies[106] on TM Fc show that the melting temperature of the C_H2 domain decreases from 70.1 ± 0.7 to 64.1 ± 0.4 °C with the triple mutations (experiments done in 25 mM histidine, 7% sucrose, pH 6.0). These data suggest that the S442C mutation is more destabilizing than the TM mutations.

Regarding the thermal stability of the C_H3 domain (second unfolding peak), the melting temperatures of all four constructs measured are similar, and all within 2 ºC of 81 ºC, **Table 3.1**, however, they are not within error (fitting errors) of each other, which is surprising. It should be noted that although errors shown in **Table 3.1** are fitting errors, they are similar to standard deviations calculated from multiple runs. More replicates would be needed to confirm these results.

Overall, these results show that DSC provides good resolution in terms of determining

the stabilities of the different domains given their melting temperatures are sufficiently resolved to be fit individually. The data suggests that the triple mutation has slightly thermally destabilized the C_H2 domain under the conditions used, but that the S442C mutation has a more destabilizing effect. The data also indicate that none of the mutations affected the stability of the C_H3 domain, however, repeat measurements would be necessary to confirm this.

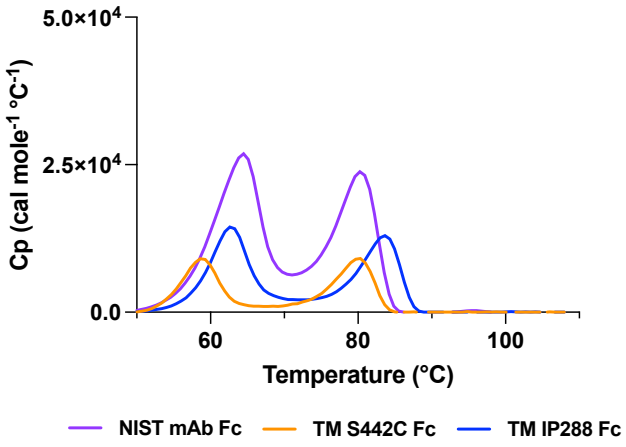


Figure 3.4: Thermograms of NIST mAb Fc, TM S442C Fc, TM Fc. The measurements were carried out in 20 mM His, pH 5.5. The concentration of the three proteins were respectively in 4.3, 0.5 and 2 mg/mL

3.2.2 Thermodynamic stability

3.2.2.1 The importance of reaching equilibrium for thermodynamic stability

Chemical denaturation curves provide information on the thermodynamic stability of proteins. The fluorescence derived from solvent-exposed tryptophan was measured at various concentrations of guanidinium chloride (GdmCl), after a certain incubation time. The data were analyzed using the average emission wavelength (AEW, **Equation 2.3**) and fitted to a two- or three-state model, depending on the number of unfolding transitions observed (**Equations 2.5 or 2.7**). Three thermodynamic parameters are obtained: ΔG_{X-Y}^{H2O}, the difference in Gibbs free energy between two states X and Y in water; m_{X-Y}, the m-value between the two states X and Y, which is a constant of proportionality describing how much the ΔG_{X-Y} changes with denaturant concentration; and $[den]_{50\%\,X-Y}$,the midpoint of denaturation between states X and Y. The data points from the curve are fitted to generate m_{X-Y} and $[den]_{50\%\,X-Y}$ and ΔG_{X-Y}^{H2O} was calculated by multiplying these two parameters together.

To obtain accurate data from chemical denaturation curves, the system being

measured needs to be at equilibrium and the unfolding needs to be reversible[103]. The fact that a denaturation curve has reached equilibrium was verified by measuring the same denaturation curve with increasing incubation times until the curves overlap and the fitted $[den]_{50\%\,X-Y}$ are within error. The intrinsic fluorescence of the full-length IgG NIST mAb was measured after 1, 2, 4, 9 and 14 days of incubation in increasing concentrations of guanidinium chloride, **Figure 3.5**. The data recorded were analyzed using the average emission wavelength (AEW, **Equation 2.3**) and normalized such that fully native signal has an AEW of zero, and the fully denatured state has a value of 1 (**Equation 2.8**). The chemical denaturation curves obtained showed two unfolding transitions and were therefore fit to a three-state model (**Equation 2.7**). The midpoint of denaturation ($[den]_{50\%}$) of the first transition didn't change very considerably with time, whereas the midpoint of denaturation of the second transition decreased over time and plateaued from nine days (**Figure 3.5**, **Figure 3.6**, **Table 3.2**).

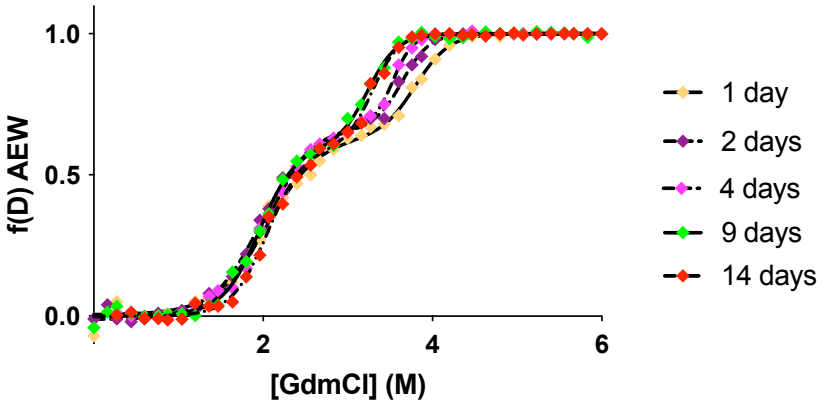


Figure 3.5: Chemical denaturation curves of NIST mAb IgG with increasing incubation times. 1 day (yellow), 2 days (purple), 4 days (pink), 9 days (green), 14 days (red).

Table 3.2: Thermodynamic parameters of NIST mAb IgG stability.

Incubation time (days)	m_{I-N} (kcal mol^{-1} M^{-1})	$[den]_{50\%\,I-N}$ (M)	ΔG^{H2O}_{I-N} (kcal mol^{-1})	m_{D-I} (kcal mol^{-1} M^{-1})	$[den]_{50\%\,D-I}$ (M)	ΔG^{H2O}_{D-I} (kcal mol^{-1})
1	1.9 ± 0.2	2.04 ± 0.03		2.9 ± 0.4	3.80 ± 0.03	
2	2.0 ± 0.1	1.97 ± 0.02		3.6 ± 0.4	3.65 ± 0.02	
4	2.2 ± 0.1	2.05 ± 0.01		4.9 ± 0.5	3.54 ± 0.02	
9	2.8 ± 0.2	1.96 ± 0.02	5.4 ± 0.4	3.4 ± 0.4	3.24 ± 0.03	11 ± 1
14	2.8 ± 0.3	2.06 ± 0.02	5.7 ± 0.5	3.8 ± 0.5	3.29 ± 0.03	12 ± 2

Parameters obtained from the fitting of the chemical denaturant unfolding curves. ΔG^{H2O}_{X-Y} was only calculated when the unfolding transitions reached equilibrium. The errors of m_{X-Y} and $[den]_{50\%\,X-Y}$ are fitting errors, while the errors of ΔG^{H2O}_{X-Y} were calculated as compound errors (see Appendix II).

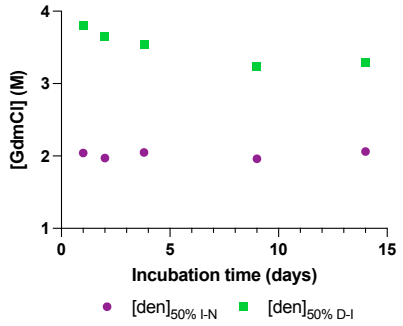


Figure 3.6: Evolution of the midpoints of chemical denaturation of the first and second unfolding transition of NIST mAb over time, in 20 mM His pH 5.5 at 25 °C. The first unfolding transition ($[den]_{50\%\ I\text{-}N}$) is shown in purple and the second unfolding transition ($[den]_{50\%\ D\text{-}I}$) is shown in green.

3.2.2.2 Comparison of the thermodynamic stabilities of the different antibody mutants

All antibodies were fitted to a three-state model, i.e., an intermediate is sufficiently populated to be measured and two unfolding transitions are observed. For the curves corresponding to the unfolding of the Fc fragments, a global fitting method, where $[den]_{50\%}$ was allowed to vary between mutants but where the *m*-value was shared for all mutants studied was performed. This assumes that the mutations don't affect the structure of the antibody and therefore the difference in solvent accessible surface area on unfolding (ΔSASA) is the same.

3.2.2.2.1 Identification of the unfolding transition of the Fab domain: comparison of the unfolding phases of the NIST mAb IgG and NIST mAb Fc

NIST mAb IgG and NIST mAb Fc differ only by the presence/absence of the Fab domain, as the rest of the sequence is identical. The chemical denaturation curve of the NIST mAb IgG showed two unfolding transitions (**Figure 3.5**), the first at $[den]_{50\%\ I-N}$ = 1.96 ± 0.02 M, and the second at $[den]_{50\%\ D-I}$ = 3.24 ± 0.03 M of GdmCl, after it reached equilibrium. According to the differential scanning calorimetry results of NIST mAb IgG (**Figure 3.2**), three unfolding events and two stable intermediates are expected. The fact that two transitions are observed instead of three suggests that two of the unfolding transitions likely overlap.

The unfolding and refolding curves were measured for NIST mAb Fc, and they reached equilibrium after 72 hours, given they overlap and the fitted parameters were within error (**Figure 3.8 A, Table 3.3**). The curves displayed two transitions which are rather close, with midpoints around 2 M GdmCl (**Table 3.3**). At this point, it is not yet

clear which of the two transition corresponds to the unfolding of the C_H2 domain and which to the unfolding of the C_H3 domain. Comparing the denaturation midpoints of NIST mAb Fc and IgG, it appears that the first unfolding transition observed for NIST mAb IgG is very close to those of NIST mAb Fc. This observation suggests that the first transition observed for NIST mAb IgG corresponds to both the C_H2 and C_H3 domains unfolding. The second unfolding transition of NIST mAb IgG ($[den]_{50\%\ D-I}$ = 3.24 ± 0.03 M GdmCl) then corresponds to the Fab domain (**Figure 3.7 A**). These observations agree with the thermal stability data, and also shows that the Fab does not seem to impact the stability of the Fc domain.

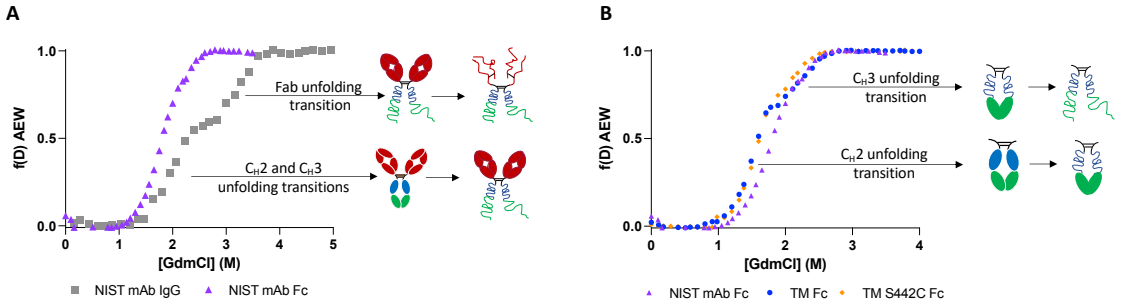


Figure 3.7: Normalized chemical denaturation curves of all antibodies at equilibrium. A. Comparison of NIST mAb IgG (9 days incubation) and NIST mAb Fc (72 h incubation). **B.** Comparison of NIST mAb Fc (72 h incubation), mean of triplicates of TM Fc (48 h incubation) and TM S442C Fc (48 h incubation). On the right of the chemical denaturation curves is a schematic showing which domains are unfolding in which transitions. Fab is shown in red, C_H2 in blue and the C_H3 domains in green.

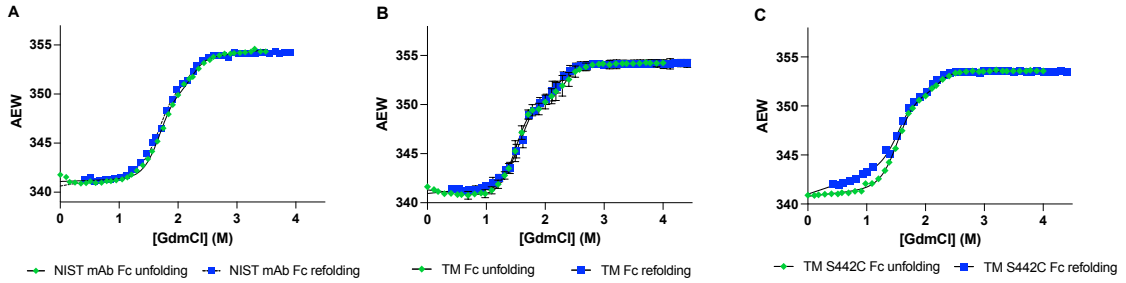


Figure 3.8: Unfolding and refolding curves for the Fc domains for all the mutants. All curves were measured in 20 mM His pH 5.5, incubated at 25 °C, and fitted to a three-state model. **A.** NIST mAb Fc at equilibrium (72-hour incubation, protein concentration was 1 µM). **B.** Mean of triplicate measurements for TM Fc at equilibrium, the error bars are the standard deviation of the triplicate measurements (48-hour incubation, protein concentration was 1 µM). **C.** TM S442C Fc at equilibrium (48-hour incubation, protein concentration was 2 µM). Green shows that data starting from a native, folded state (unfolding), whilst blue shows the data for refolding (starting from a fully denatured state).

Table 3.3: Thermodynamic parameters fitted from the unfolding curves at equilibrium for NIST mAb Fc, TM Fc and TM S442C Fc

		m_{I-N} (kcal mol^{-1} M^{-1})	$[den]_{50\%\,I-N}$ (M)	ΔG^{H2O}_{I-N} (kcal mol^{-1})	m_{D-I} (kcal mol^{-1} M^{-1})	$[den]_{50\%\,D-I}$ (M)	ΔG^{H2O}_{D-I} (kcal mol^{-1})
NIST mAb Fc	unfolding	3.9 ± 0.1	1.70 ± 0.02	6.6 ± 0.2	5.3 ± 0.4	2.28 ± 0.04	12.2 ± 0.9
	refolding	3.9 ± 0.1	1.69 ± 0.02	6.6 ± 0.2	5.3 ± 0.4	2.26 ± 0.03	12.1 ± 0.9
TM Fc	unfolding	3.9 ± 0.1	1.49 ± 0.01	5.8 ± 0.2	5.3 ± 0.4	2.35 ± 0.02	12.5 ± 0.9
	refolding	3.9 ± 0.1	1.58 ± 0.02	6.2 ± 0.2	5.3 ± 0.4	2.27 ± 0.03	12.1 ± 0.9
TM S442C Fc	unfolding	3.9 ± 0.1	1.54 ± 0.01	6.0 ± 0.2	5.3 ± 0.4	2.23 ± 0.04	11.9 ± 0.9
	refolding	3.9 ± 0.1	1.57 ± 0.02	6.1 ± 0.2	5.3 ± 0.4	2.16 ± 0.05	11.5 ± 0.9

The errors reported in the table are fitting errors.

3.2.2.2.2 Effect of a triple mutation in the C_H2 domain on thermodynamic stability: comparison of NIST mAb Fc and TM Fc

TM Fc differs from NIST mAb Fc by three mutations in the upper part of the C_H2 domain (**Figure 3.1**): L234F, L235E and P331S, which decrease the affinity for the Fc γ-receptor type IIIA[147,148]. While the denaturation midpoint was within error for the second transition corresponding o the unfolding of the C_H3 domain for both constructs, that of the first unfolding transition corresponding to the unfolding of the C_H2 domain was lower for TM Fc compared to NIST mAb Fc (**Figure 3.7 B, Figure 3.8**, **Table 3.3**). This suggests that the triple mutation destabilized the C_H2 domain thermodynamically consistent with the slight decrease in T_m observed here and the larger decrease in T_m seen in other studies[34].

3.2.2.2.3 Effect of the substitution of a serine by a cysteine in the C_H3 domain on the thermodynamic stability: comparison of TM Fc and TM S442C Fc

TM S442C Fc differs from TM Fc by the substitution of a serine by a cysteine at the C-terminal end of the C_H3 domain. For TM S442C Fc, two unfolding transitions and three states are observed in the chemical denaturation curve (**Figure 3.8 B**). The denaturation midpoint of the first transition was within error of that of TM Fc, and that of the second transition was very similar to both NIST mAb Fc and TM Fc (**Table 3.3**). This suggests that the S442C substitution hasn't affected the thermodynamic stability of the C_H3 domain, and that the stability of the C_H2 domain is also unaffected.

3.2.3 Kinetic stability

Unfolding kinetics were measured by rapidly diluting native protein into a range of high concentrations of chemical denaturant and monitoring the changes in intrinsic fluorescence on unfolding. For each concentration of denaturant, three repeats were

measured. The fluorescence signal corresponding to the unfolding of NIST mAb IgG was best fit with a triple exponential (**Equation 2.11**, **Figure 3.9 A&C**) whereas the fluorescence signals on unfolding for all Fc constructs were best fit with a double exponential (**Equation 2.10**, **Figure 3.9 B&D**).

Three independent measurements of the unfolding rate constants k_U were obtained for each protein at different concentrations of GdmCl. These values were averaged and then plotted as ln k_U *versus* the denaturant concentration, **Figure 3.10**. Data were then fitted to a linear equation (**Equation 2.12**) to obtain the value for ln k_U in water as well as m_{ku}, the slope of the plot (**Table 3.4**).

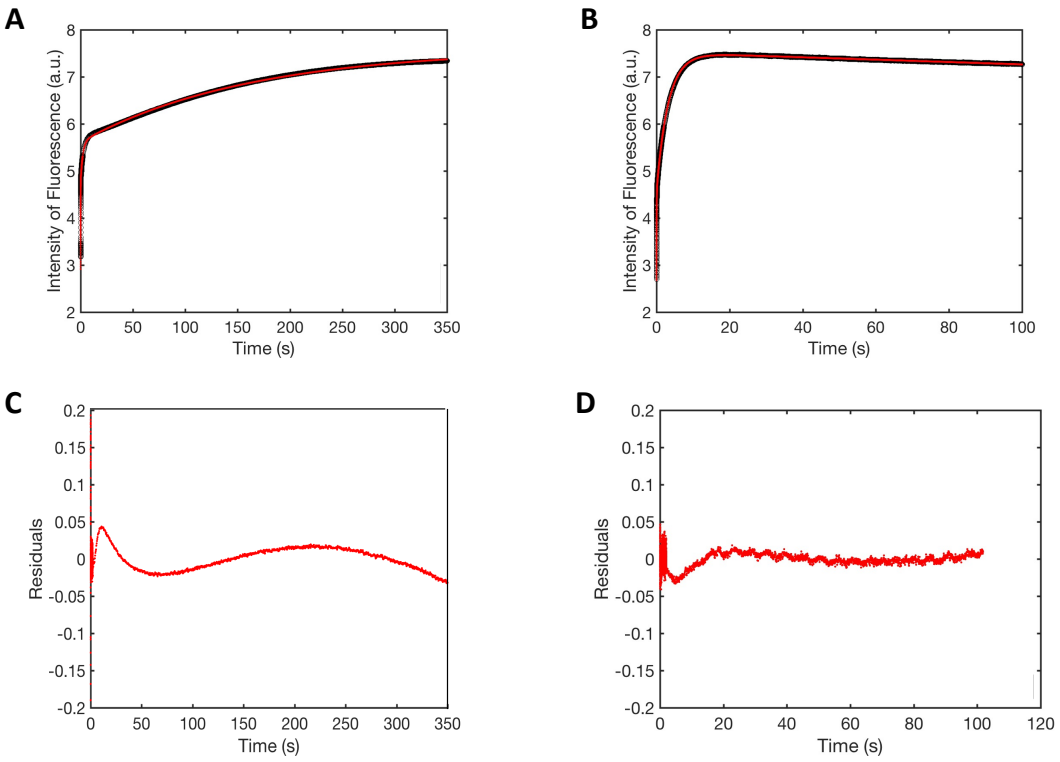


Figure 3.9: Typical kinetic unfolding data. A. Unfolding of NIST mAb IgG probed by intrinsic fluorescence (1 µM, black) showing the best fit of the data to a triple exponential (red line). **B.** Unfolding of NIST mAb Fc (1 µM, black). The best fit of the data to a double exponential function with slope (due to photobleaching) is shown as a red line. **C.** Residuals of the fitting of the NIST mAb IgG unfolding shown in A to a triple exponential function and **D.** Residuals of the fitting of NIST mAb Fc unfolding data shown in B to a double exponential plus drift. The percent residuals can be calculated using $\%residual = \frac{residual}{A\ total} \times 100$ (the residuals are just the absolute value of the largest value, so here 0.05). For the data shown in C, the %residuals are no higher than 1.25 % demonstrating a good fit.

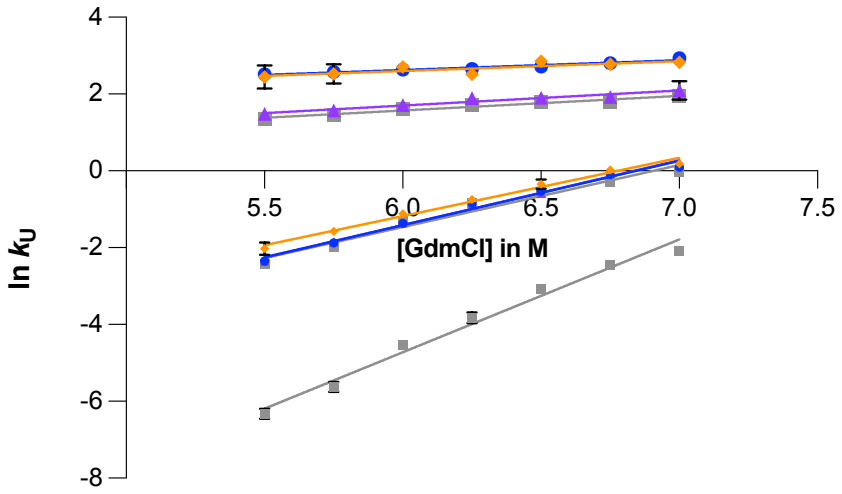


Figure 3.10: Unfolding kinetics of NIST mAb IgG (grey squares), NIST mAb Fc (purple triangles), TM Fc (blue dots) and TM S442C Fc (orange lozenges). The experiments were carried out at 25 °C in 20 mM His, pH 5.5, at 0.5 µM.

Table 3.4: Kinetic parameters from unfolding experiments

		m_{kU} (M^{-1})	ln k_U (H_2O)
NIST mAb IgG	ln($k1$ C_H2)	0.38 ± 0.04	-0.7 ± 0.2
	ln($k2$ C_H3)	1.6 ± 0.1	-11.1 ± 0.8
	ln($k3$ Fab)	2.9 ± 0.2	-22 ± 1
NIST mAb Fc	ln($k1$ C_H2)	0.39 ± 0.04	-0.7 ± 0.3
	ln($k2$ C_H3)	1.7 ± 0.1	-11.6 ± 0.4
TM Fc	ln($k1$ C_H2)	0.25 ± 0.03	1.1 ± 0.2
	ln($k2$ C_H3)	1.7 ± 0.1	-11.5 ± 0.6
TM S442C Fc	ln($k1$ C_H2)	0.3 ± 0.1	1.1 ± 0.5
	ln($k2$ C_H3)	1.5 ± 0.1	-10.3 ± 0.5

3.2.3.1.1 The errors reported are fitting errors.

3.2.3.2 Identification of the unfolding phase of the Fab domain: comparison of the unfolding kinetics of NIST mAb IgG and NIST mAb Fc

The fact that the NIST mAb IgG unfolding kinetic data can be best fitted with a triple exponential shows that there are three unfolding events happening (**Figure 3.10**). In contrast, all Fc constructs have just two unfolding phases corresponding to the fastest and the intermediate unfolding phases observed for the full-length protein (**Figure 3.10**). The rate constants for the fastest (k_1) and intermediate (k_2) phases are within error for NIST mAb IgG and Fc. Therefore, the slowest phase observed in the unfolding kinetics can be attributed to the unfolding of the Fab domain. Given the DSC data (**Figure 3.2**), it is likely that the fastest unfolding rate comes from the unfolding of the C_H2 domain and the intermediate rate comes from the unfolding of the C_H3 domain. It is also worth noting

that the presence or absence of the Fab domain does not influence the kinetic stability of the Fc domain.

3.2.3.3 Effect of a triple mutation in the C_H2 domain on kinetic stability: comparison of NIST mAb Fc and TM Fc

The two fastest unfolding rate constants (k_1 and k_2) corresponding to the unfolding of the C_H2 and C_H3 domain are larger for the triple mutant TM Fc compared to the control NIST mAb Fc (**Figure 3.10**, **Table 3.4**), which shows that the C_H2 domain has been kinetically destabilized by the triple mutation. In addition, the m_{k_U} value for the fastest phase of TM Fc is smaller than that of NIST mAb Fc. According to the Hammond effect, as the slope m_{k_U} is directly related to the change in solvent accessible surface area between the native and transition state by the equation $m_{TS-N} = RTm_{k_U}$, where m_{TS-N} is in kcal mol^{-1} M^{-1} and m_{k_U} in M^{-1}, then the smaller the value, the more similar the structures of the native and the transition states are. Therefore, the mutations not only kinetically destabilize the C_H2 domain but they shift the transition state towards a more native-like structure (in the energy landscape, x-axis). From only the unfolding data, it is possible to deduce that the difference in energy between the native and transition state has decreased. According to the thermodynamic data, the native state has increased in energy, i.e., has been destabilized consistent with this.

3.2.3.4 Effect of the substitution of a serine by a cysteine in the C_H3 domain on kinetic stability: comparison of TM Fc and TM S442C Fc

TM Fc and TM S442C Fc differ by a single substitution of a serine by a cysteine in the C_H3 domain. The unfolding kinetic rate constants for both unfolding phases for TM Fc and TM S442C Fc are within error (**Figure 3.10**, **Table 3.4**) which suggests that the substitution at the studied position does not impact the stability of the C_H3 domain nor of the C_H2 domain.

3.2.3.5 General observation on the m_{k_U} values

One can notice that generally the slopes of the three unfolding phases, given by the m_{k_U} values, are different: the fastest unfolding phase (unfolding of the C_H2 domain) has the shallowest slope whereas the slowest unfolding phase (unfolding of the Fab domain) has the steepest slope, which means the unfolding rate is more sensitive to denaturant concentration. According to the Hammond effect, it appears therefore that the difference in structure between the native and unfolding transition state of the C_H2 domain is rather smaller than for the C_H3 domain. This may be because only a few interactions

require breaking in the C_H2 domain to overcome the unfolding energy barrier, consistent with the fact that the C_H2 domain is less stable than the C_H3 and Fab domains.

3.3 DISCUSSION AND CONCLUSION

In this Chapter, a combination of several approaches to probe the stability of wild-type and mutant antibodies was employed to investigate the impact of different mutations on the stability of a therapeutic antibody scaffold. The relative stabilities of the three domains is in agreement with previous studies, which demonstrated that, thermally[150], thermodynamically[153] and kinetically[153], for wild-type antibodies, the C_H2 domain is less stable than the C_H3 domain, which is less stable than the Fab. In addition, consistent with previous work, the Fab fragment has the largest *Cp* value, about three times greater that the C_H2 and C_H3 domains[96].

Overall, the results from the different techniques were in good agreement and demonstrated that the triple mutation L234F/L235E/P331S destabilizes the C_H2 domain thermodynamically and kinetically. However, the destabilization was not too significant, and is comparable or lower to other formats used in the clinic[34,106,107]. In addition, the triple mutation did not affect the stability of the C_H3 domain. The S442C substitution in the C_H3 domain engineered into the antibody to allow conjugation, didn't destabilize any domain either. Results on the Fab domain of NIST mAb IgG showed that it is remarkably stable (crystal structure of NIST Fab: PDB 5K8A[154]); the stability of this domain can be very variable as the different light chain types and CDRs have a significant impact on stability[96]. It is a little less clear how the mutations affected the thermal stability, and the experiments require repeating to get standard deviations in T_m values rather than using fitting errors, to establish what differences are significant and which are not.

In conclusion, the TM S442C Fc is an excellent antibody-drug conjugate scaffold, as the substitution at the end of the C_H3 domain does not have a significant impact on the stability. However, the stability of the protein is not the only parameter that will determine how suitable TM S442C Fc is for clinical studies; the ease of conjugation, the stability of the conjugated payload (both in terms of antibody stability as well as whether the payload remains attached *in vivo* until it is at its target site) and the final drug to antibody ratio will all be very important as well.

4 INTERCONVERSION OF UNEXPECTED THIOL STATES AFFECTS THE STABILITY, STRUCTURE AND DYNAMICS OF ANTIBODY ENGINEERED FOR SITE-SPECIFIC CONJUGATION

This chapter was published in *Bioconjugate Chemistry* as:

Orozco, C. T., Edgeworth M. J., Devine P. W. A, Hines A. R., Cornwell O., Thompson C., Wang X., Phillips J. J., Ravn P., Jackson S. E., Bond N. J. Interconversion of unexpected thiol states affects the stability, structure, and dynamics of antibody engineered for site-specific conjugation. *Bioconjug. Chem.* **32**, 1834–1844 (2021)

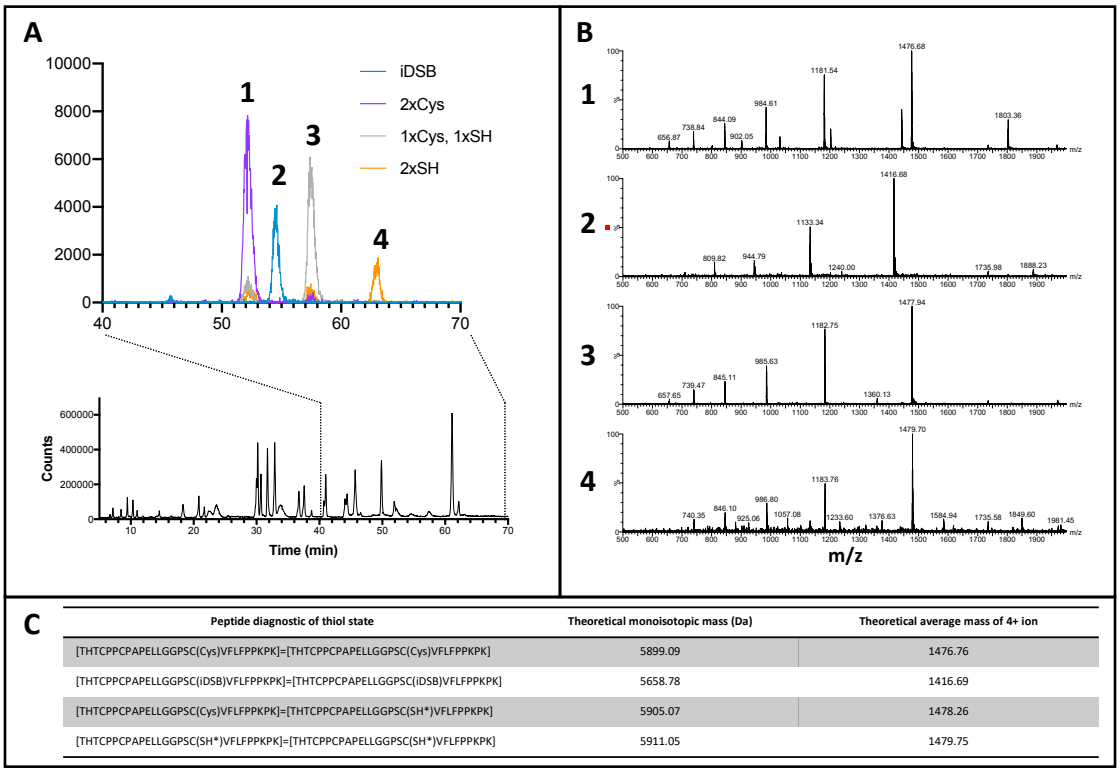
4.1 INTRODUCTION

Antibody-drug conjugates (ADCs) have become very promising therapeutics in oncology by combining the high specificity of a tumour-recognising monoclonal antibody (mAb) with the potency of a chemotherapeutic small molecule (payload)[155]. Combining two therapeutic molecules into a single agent reduces the systemic toxicity of small-molecule chemotherapy whilst facilitating the use of more potent cytotoxic agents which, if administered alone, would have significant dose limitation due to toxicity[156–160]. ADCs represent a huge area of research and there are currently nine FDA-approved ADCs on the market, including five within the past year, and more than 60 ADCs are being clinically evaluated in more than 200 active or recently completed clinical trials[161] (Clinicaltrials.gov). In the first generation of ADCs, the payload was conjugated to lysines, which lead to a distribution in number and position of drugs attached, resulting in variable drug-to-antibody ratios (DARs)[33]. The DAR is an important contributor to the therapeutic index: the dose range within which efficacy is achieved with an acceptable safety profile, and it must be tightly controlled. To do this, the field has iterated towards the conjugation of payloads to canonical cysteines and then to strategies that enabled site-specific conjugation, including non-natural amino acids[36–38], the use of enzymes such as formylglycine generating enzyme[39], transglutaminase[40,41] and sortase A[42], as well as point mutations to add unpaired cysteines for conjugation, either by substitution[34] or by insertion[35]. In particular, the addition of a cysteine near the hinge region of an IgG1 scaffold has been investigated[34,35,162], and both the substitution at position 239 (Eu mAb numbering) in the heavy chain (S239C) and the insertion after that position (C239i) have given promising results[34,35]: both cysteines are easily conjugated, provide stability of the payload over time, decrease FcγR binding and do not affect the binding to the neonatal Fc receptor (FcRn), which should ensure a similar half-life to the wild-type scaffold[34,35].

4.2 BACKGROUND

During the large-scale manufacture (50 to 500 L) of several C239i antibodies, it was observed that the inserted cysteine could adopt three distinguishable chemical states. Using non-reduced peptide mapping, these were characterised as: free thiol (2xSH); capped with cysteine (cysteinylated, 2xCys, forming commonly on free cysteines in an oxidising environment where cysteines are present in the media[163]; this is a common modification during manufacturing); or forming an additional disulfide bridge between both C239i residues (iDSB) (**Figure 4.1**, **Figure 4.2**). A fourth state, 1xCys + 1xSH

(containing both free thiol and cysteinylated C239i), was also observed, although in much smaller proportions. Further analysis revealed that the proportions of these different thiol states vary from lot-to-lot and across different antibodies (**Figure 4.3 A**).



Peptide diagnostic of thiol state	Theoretical monoisotopic mass (Da)	Theoretical average mass of 4+ ion
[THTCPPCPAPELLGGPSC(Cys)VFLFPPKPK]=[THTCPPCPAPELLGGPSC(Cys)VFLFPPKPK]	5899.09	1476.76
[THTCPPCPAPELLGGPSC(iDSB)VFLFPPKPK]=[THTCPPCPAPELLGGPSC(iDSB)VFLFPPKPK]	5658.78	1416.69
[THTCPPCPAPELLGGPSC(Cys)VFLFPPKPK]=[THTCPPCPAPELLGGPSC(SH*)VFLFPPKPK]	5905.07	1478.26
[THTCPPCPAPELLGGPSC(SH*)VFLFPPKPK]=[THTCPPCPAPELLGGPSC(SH*)VFLFPPKPK]	5911.05	1479.75

Figure 4.1: Non-reduced peptide mapping of Fc-C239i antibody. A. Total ion chromatogram (bottom) and zoom of extracted ion chromatograms (top) for iDSB, 2xCys, 1xCys & 1xSH and 2xSH thiol states (1 to 4). **B.** Mass spectrum for each extracted ion (1 to 4). **C.** Theoretical MW for peptide diagnostic of thiol state; * denotes modification by N-ethylmaleimide. Experiments were carried out and analysed by Dr Paul Devine and Alistair Hines.

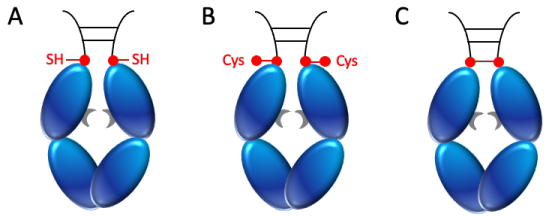


Figure 4.2: Schematic of the Fc domain containing the inserted cysteine after position 239 (Fc-C239i), in the three predominant forms: A. free thiol (2xSH), **B.** doubly-cysteinylated (2xCys) and **C.** forming an additional disulfide bridge between both C239i residues (iDSB). Drawn in grey: glycosylation. A partially cysteinylated state, 1xCys + 1xSH (containing both free thiol and cysteinylated C239i), was also observed during manufacture, although in much smaller proportions and was not enriched for further characterisation.

Interestingly, oxidised thiol states (2xCys and iDSB) predominate during expression in Chinese Hamster Ovarian (CHO) cells under fed-batch culture (**Figure 4.3 B**), and it is upon harvest that reduced forms are introduced: 2xSH and 1xCys + 1xSH.

The thiol state appears to be largely unaffected by the subsequent purification steps (**Figure 4.3 C**) and whilst the proportion of iDSB appears to vary between antibodies and from lot-to-lot, the iDSB was always present.

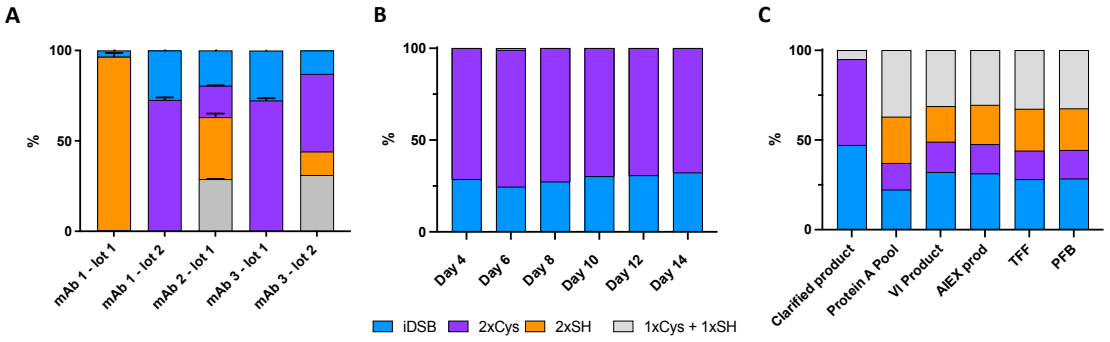
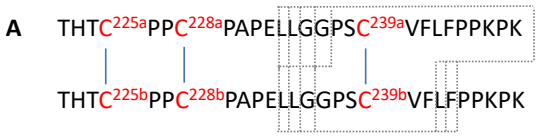
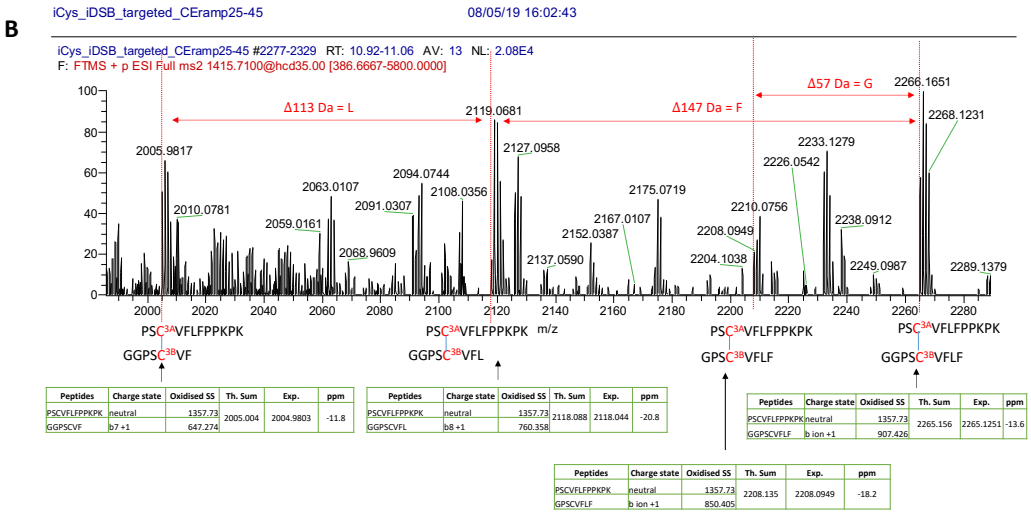


Figure 4.3: Thiol states observed for different C239i (full length) antibodies during large-scale manufacture. A. Proportions of thiol states for different antibodies and different lots. Variability is observed in the proportions of thiol states; iDSB is present in all lots. **B.** Variation of thiol states over the course of expression in CHO cells for another C239i antibody, mAb 4 lot 1. The proportions of thiol states are consistent throughout the expression. **C.** Evolution of thiol states for antibody mAb 4 lot 2 during downstream purification (VI: viral inactivation, AIEX: anion exchange, TFF: tangential flow filtration; PFB: pre-formulated bulk). Introduction of reduced thiol states occurs during harvest after which the proportions of thiol states remain consistent. These thiol states were also observed in small-scale cultures (15 mL – data not shown). Experiments were carried out and analysed by Dr Paul Devine and Alistair Hines.

The additional disulfide bridge (iDSB) was confirmed by tandem mass spectrometry (MS/MS by collision induced dissociation) to be an additional interchain disulfide bond between the two inserted cysteines, downstream of the two canonical disulfide bonds that covalently bond IgG1 heavy chains. The main fragment ions identified support the formation of three parallel disulfide bonds. Whilst it is possible to form alternative configurations of disulfide linkage, no diagnostic ions were found. The peptide in the hinge, corresponding to a +4 ion (1415.71 m/z) was isolated and fragmented, and the obtained peptides were identified. Peptides linked by a disulfide bridge between the inserted cysteines 239 were identified and the amino-acid sequence was reconstituted amino acid by amino acid (**Figure 4.4 B**): they are composed of a neutral peptide truncated at the N-terminus only (resulting from a fragmentation that generated a y ion and a neutral peptide, as the charges were retained on the N-terminus fragmented peptide), and a + 1 charged peptide truncated both at the N-terminus, as a neutral fragment (same explanation as above), as well as at the C-terminus as a b ion (after the fragmentation event, the charge was retained at the N-terminus). These findings demonstrate that the additional disulfide bridge is indeed an interchain disulfide bridge

between the two 239 inserted cysteine linking the two heavy chains, as the canonical disulfide bridges do. No evidence of scrambling was observed (**Figure 4.4**). The presence of the additional interchain disulfide bridge is surprising since the α-carbons of the amino acids at position 239 on each heavy chain are 17.2 Å apart from each other in the crystal structure (PDB: 3AVE), whilst the distance between the two α-carbons of the cysteines involved in a disulfide bridge is 6.4 Å (canonical disulfide bridges in 3AVE). This raises questions regarding how the iDSB forms and the resulting structural changes necessary to accommodate it.



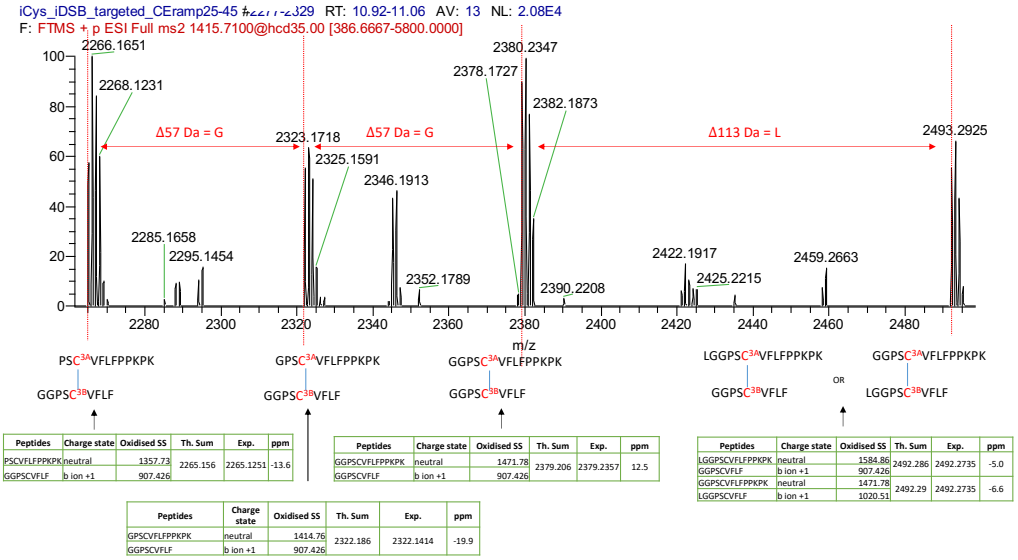


Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
PSCVFLFPPKPK	neutral	1357.73	2005.004	2004.9803	-11.8
GGPSCVF	b7 +1	647.274			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
PSCVFLFPPKPK	neutral	1357.73	2118.088	2118.044	-20.8
GGPSCVFL	b8 +1	760.358			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
PSCVFLFPPKPK	neutral	1357.73	2265.156	2265.1251	-13.6
GGPSCVFL	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
PSCVFLFPPKPK	neutral	1357.73	2208.135	2208.0949	-18.2
GPSCVFLF	b ion +1	850.405			

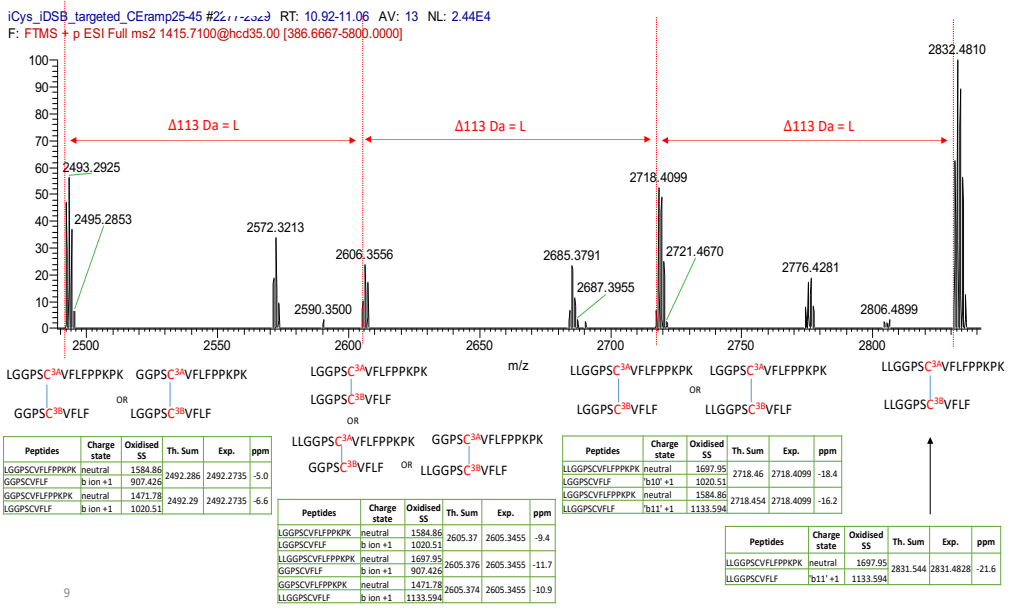


Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
PSCVFLFPPKPK	neutral	1357.73	2265.156	2265.1251	-13.6
GGPSCVFLF	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
GGPSCVFLFPPKPK	neutral	1471.78	2379.206	2379.2357	12.5
GGPSCVFLF	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
LGGPSCVFLFPPKPK	neutral	1584.86	2492.286	2492.2735	-5.0
GGPSCVFLF	b ion +1	907.426			
GGPSCVFLFPPKPK	neutral	1471.78	2492.29	2492.2735	-6.6
LGGPSCVFLF	b ion +1	1020.51			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
GPSCVFLFPPKPK	neutral	1414.76	2322.186	2322.1414	-19.9
GGPSCVFLF	b ion +1	907.426			



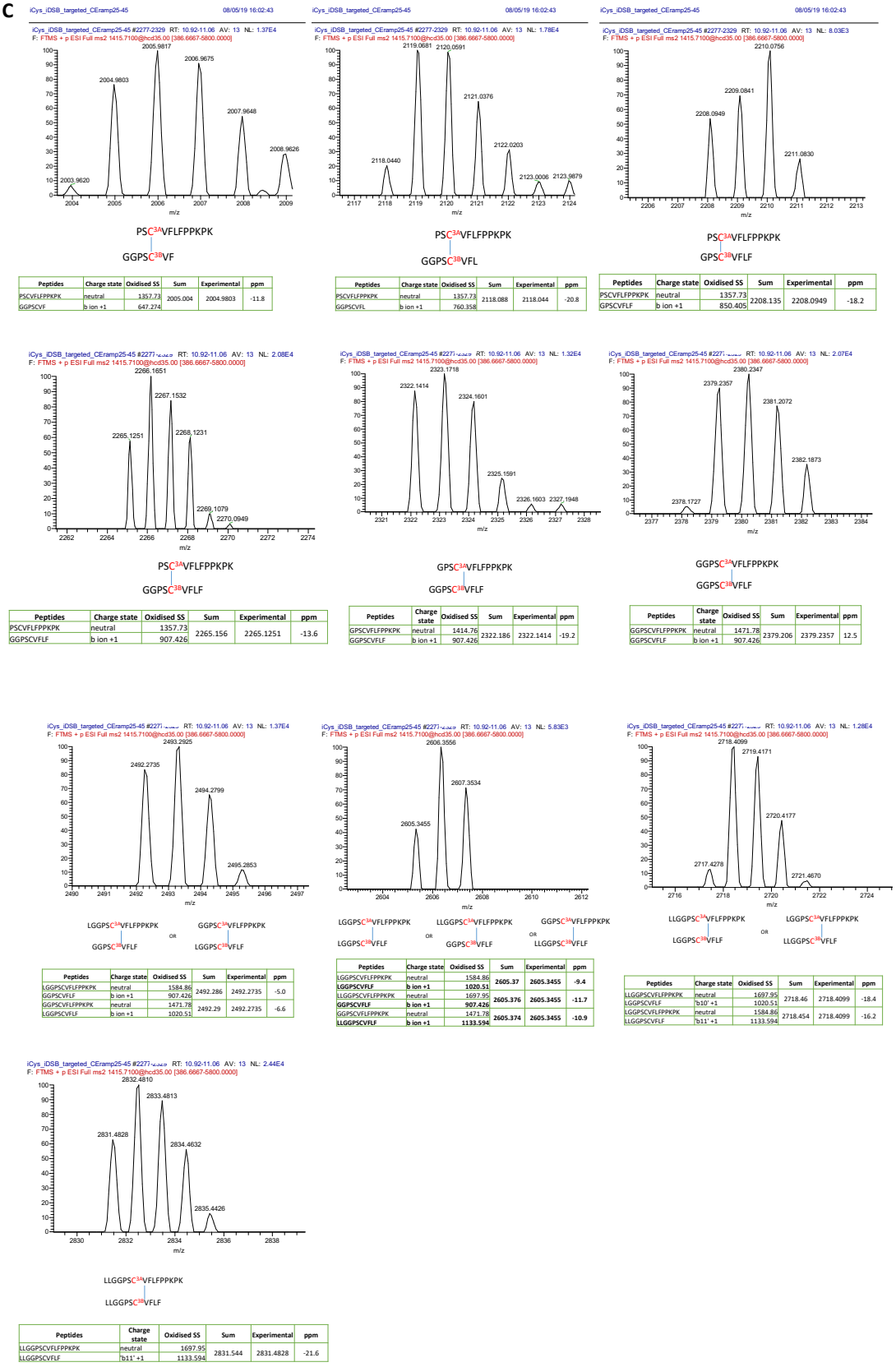
Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
LGGPSCVFLFPPKPK	neutral	1584.86	2492.286	2492.2735	-5.0
GGPSCVFLF	b ion +1	907.426			
GGPSCVFLFPPKPK	neutral	1471.78	2492.29	2492.2735	-6.6
LGGPSCVFLF	b ion +1	1020.51			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
LGGPSCVFLFPPKPK	neutral	1584.86	2605.37	2605.3455	-9.4
LGGPSCVFLF	b ion +1	1020.51			
LLGGPSCVFLFPPKPK	neutral	1697.95	2605.376	2605.3455	-11.7
GGPSCVFLF	b ion +1	907.426			
GGPSCVFLFPPKPK	neutral	1471.78	2605.374	2605.3455	-10.9
LLGGPSCVFLF	b ion +1	1133.594			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
LLGGPSCVFLFPPKPK	neutral	1697.95	2718.46	2718.4099	-18.4
LGGPSCVFLF	b10' +1	1020.51			
LGGPSCVFLFPPKPK	neutral	1584.86	2718.454	2718.4099	-16.2
LLGGPSCVFLF	b11' +1	1133.594			

Peptides	Charge state	Oxidised SS	Th. Sum	Exp.	ppm
LLGGPSCVFLFPPKPK	neutral	1697.95	2831.544	2831.4828	-21.6
LLGGPSCVFLF	b11' +1	1133.594			

C



Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
PSCVFLFPPKPK	neutral	1357.73	2005.004	2004.9803	-11.8
GGPSCVF	b ion +1	647.274			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
PSCVFLFPPKPK	neutral	1357.73	2118.088	2118.044	-20.8
GGPSCVFL	b ion +1	760.358			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
PSCVFLFPPKPK	neutral	1357.73	2208.135	2208.0949	-18.2
GPSCVFLF	b ion +1	850.405			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
PSCVFLFPPKPK	neutral	1357.73	2265.156	2265.1251	-13.6
GGPSCVFLF	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
GPSCVFLFPPKPK	neutral	1414.76	2322.186	2322.1414	-19.2
GGPSCVFLF	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
GGPSCVFLFPPKPK	neutral	1471.78	2379.206	2379.2357	12.5
GGPSCVFLF	b ion +1	907.426			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
LGGPSCVFLFPPKPK	neutral	1584.86	2492.286	2492.2735	-5.0
GGPSCVFLF	b ion +1	907.426			
GGPSCVFLFPPKPK	neutral	1471.78	2492.29	2492.2735	-6.6
LGGPSCVFLF	b ion +1	1020.51			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
LGGPSCVFLFPPKPK	neutral	1584.86	**2605.37**	**2605.3455**	**-9.4**
LGGPSCVFLF	**b ion +1**	**1020.51**			
LLGGPSCVFLFPPKPK	neutral	1697.95	**2605.376**	**2605.3455**	**-11.7**
GGPSCVFLF	**b ion +1**	**907.426**			
GGPSCVFLFPPKPK	neutral	1471.78	**2605.374**	**2605.3455**	**-10.9**
LLGGPSCVFLF	**b ion +1**	**1133.594**			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
LLGGPSCVFLFPPKPK	neutral	1697.95	2718.46	2718.4099	-18.4
LGGPSCVFLF	'b10' +1	1020.51			
LGGPSCVFLFPPKPK	neutral	1584.86	2718.454	2718.4099	-16.2
LLGGPSCVFLF	'b11' +1	1133.594			

Peptides	Charge state	Oxidised SS	Sum	Experimental	ppm
LLGGPSCVFLFPPKPK	neutral	1697.95	2831.544	2831.4828	-21.6
LLGGPSCVFLF	'b11' +1	1133.594			

Figure 4.4: Identification of iDSB by tandem-mass spectrometry. A. Sequence of the hinge peptide fragmented by tandem-mass spectrometry, with the fragmented ions identified. **B.** Reconstitution of the sequence fragmented by MS^2 amino by amino acid. **C.** Detailed m/z spectra for each main peak. Experiments were carried out by Dr Matthew Edgeworth and analysed by Carolina Orozco.

4.3 RESULTS AND DISCUSSION

In this study, the structural consequences of inserting cysteine at position C239i and the associated thiol states observed during manufacture of the antibody intermediate were investigated in detail. To investigate this systematically, Fc constructs enriched for each of the three thiol states were generated by Dr Paul Devine (2xSH, 2xCys and iDSB Fc-C239i), verified, and quantified by mass spectrometry by Alistair Hines (**Figure 4.5**, **Figure 4.6**, **Table 4.1**). The Fc constructs were used as surrogates for C239i IgG since it has been well established that the antibody binding fragment (Fab) does not affect the stability of the C_H2 and C_H3 domains[164,165]. The enriched variants were a kind gift from AstraZeneca.

Chemical denaturation experiments, differential scanning calorimetry and millisecond HDX mass spectrometry[139] were used to probe the stability, dynamics and structures of the C_H2 and C_H3 domains in each of the enriched states. Together, these studies provide insight into the conditions under which different thiol states interconvert, their relative stabilities and the conformational changes that occur upon formation of an additional disulfide bond. Finally, the effect of iDSB on the subsequent conjugation of the antibody intermediate was explored.

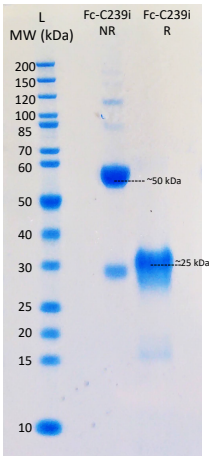


Figure 4.5: SDS-PAGE of 2xCys Fc-C239i. NR: non reduced. R: reduced. L: ladder: Page Ruler™ unstained ladder, Thermo Scientific.

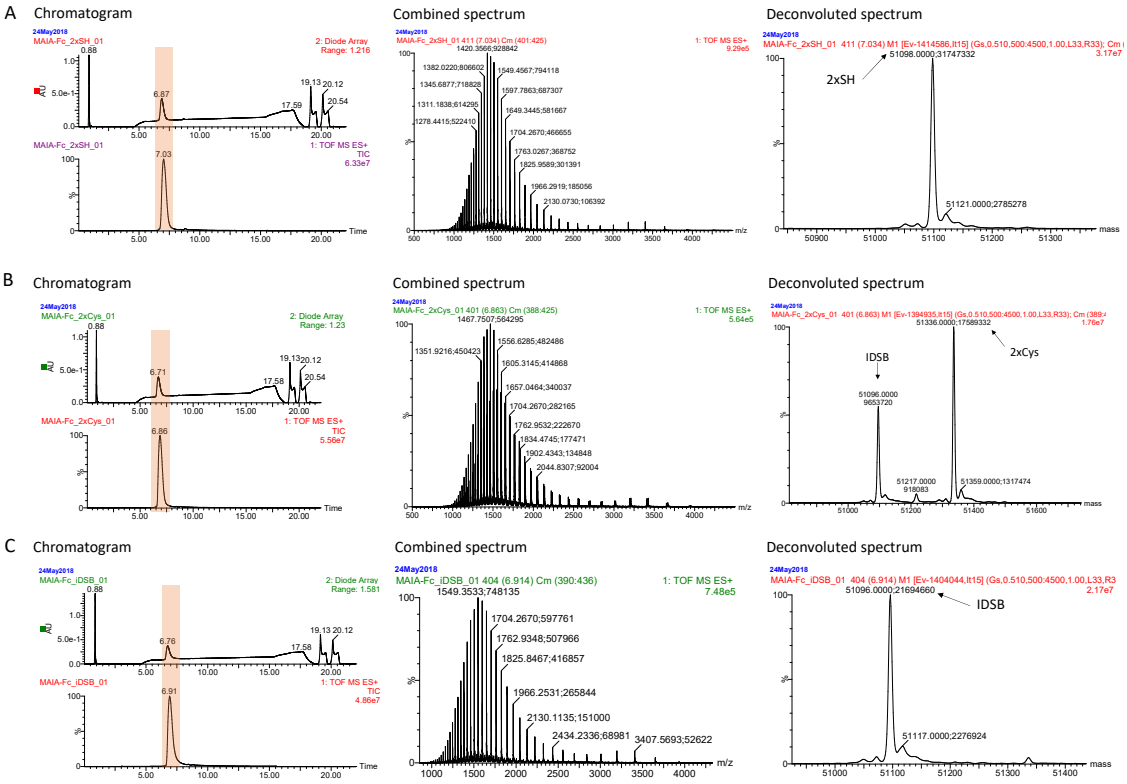


Figure 4.6: Chromatograms, combined spectra and deconvoluted spectra (LC-MS) of the enriched thiol states. A. 2xSH Fc-C239i; **B.** 2xCys Fc-C239i; **C.** iDSB Fc-C239i enriched variants. The highlighted peak on the chromatogram is the portion that was combined. The measured molecular weights confirmed that the samples were the expected variants, and were then subjected to further investigations. Experiments were carried out by Dr Paul Devine and analysed by Carolina Orozco.

Table 4.1: Quality control of the Fc-C239i variants after enrichment

	2xSH Fc-C239i (%)	2xCys Fc-C239i (%)	iDSB Fc-C239i (%)
iDSB	3.6	33.4	91.3
2xCys	0.2	54.4	8.4
1xCys + 1xSH	0.4	10.6	0.0
2xSH	95.8	1.0	0.0
2x GSH	0.0	0.3	0.1
1xGSH + 1xCys	0.0	0.1	0.1
1xGSH + 1xSH	0.0	0.0	0.0

Experiments and analysis were performed by Alistair Hines.

4.3.1 Interconversion of unexpected thiol states

To understand how the interconversion of the thiol states can occur, stress conditions that could induce unfolding and increase the molecular dynamics were

investigated. These stresses can be heat, low pH, shear stress, chemical denaturation to name a few. For this study, chemical denaturant stress was selected. The evolution of the thiol states was monitored over time using non-reducing peptide mapping. Under native conditions and incubation for seven days at 25 °C, the relative proportion of all enriched Fc variants did not change significantly (**Figure 4.7 A, Table 4.2**). However, in the presence of chemical denaturant (3.5 M guanidinium chloride (GdmCl)), both 2xSH and 2xCys converted into the iDSB Fc-C239i (**Figure 4.7 A**). Moreover, slower conversion of 2xCys to iDSB Fc-C239i was observed suggesting that interconversion might not be direct.

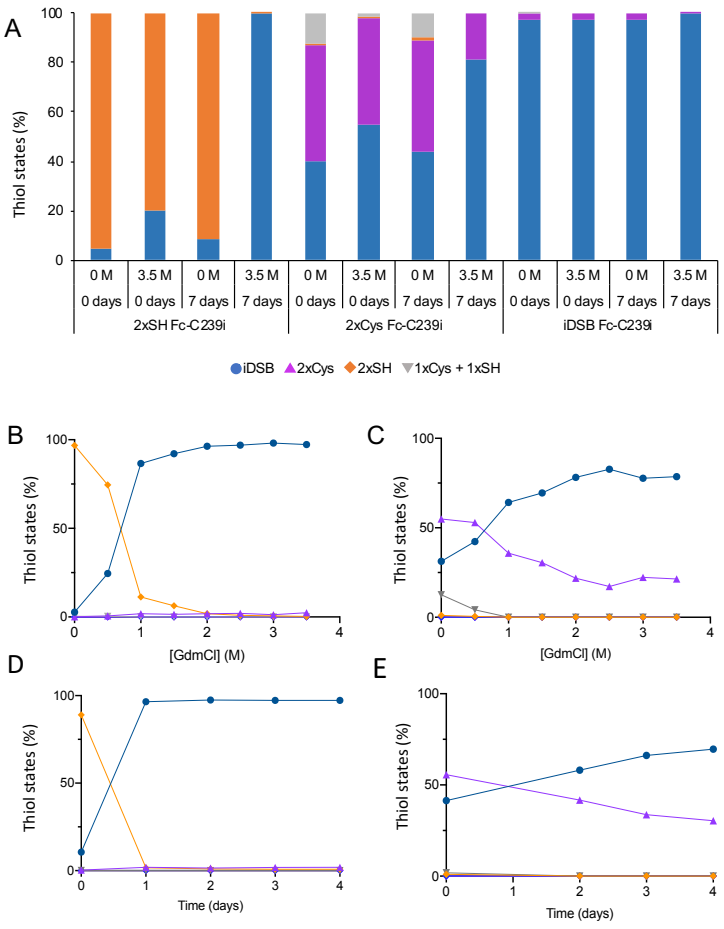


Figure 4.7: Variation of the thiol state of C239i with denaturant concentration and time. A. Effect of the concentration of guanidinium chloride and incubation time at 25 °C in 20 mM His pH 5.5 on the proportion of the thiol states for each of the three initially enriched variants (2xSH, 2xCys and iDSB Fc-C239i); raw data are shown in **Table 4.2. B.** Effect of the concentration of denaturant after 7 days of incubation at 25 °C on 2xSH Fc-C239i and **C.** 2xCys Fc-C239i enriched starting material; raw data are shown in **Table 4.3** and **Table 4.4**. **D.** Effect of the incubation time at 25 °C in 3.5 M GdmCl on 2xSH Fc-C239i enriched starting material; raw data are shown in **Table 4.5**. **E.** Effect of the incubation time at 25 °C in 3.5 M GdmCl on 2xCys Fc-C239i enriched starting material; raw data are shown in **Table 4.6**. The calculated error is the standard deviation of a double injection for each sample, to take into account the technical reproducibility. The error bars are too small to be visible. Experiment carried out and analysed by Carolina Orozco and Alistair Hines.

Table 4.2: Effect of the incubation time and the concentration of chemical denaturant on the proportion of the inserted-cysteine states

%	2xSH Fc-C239i				2xCys Fc-C239i				iDSB Fc-C239i			
Incubation time (days)	0 days	0 days	7 days	7 days	0 days	0 days	7 days	7 days	0 days	0 days	7 days	7 days
[GdmCl] (M)	0 M	3.5 M	0 M	3.5 M	0 M	3.5 M	0 M	3.5 M	0 M	3.5 M	0 M	3.5 M
iDSB	4.7	19.8	8.3	99.7	39.8	55.1	43.9	81.1	97.0	97.5	97.4	99.7
2xCys	0.0	0.0	0.0	0.0	46.8	42.6	45.3	18.9	2.6	2.5	2.6	0.3
2xSH	95.3	80.2	91.7	0.3	1.0	0.6	0.8	0.0	0.0	0.0	0.0	0.0
1xCys	0.0	0.0	0.0	0.0	12.3	1.7	10.0	0.0	0.4	0.0	0.0	0.0

All experiments conducted at 25 °C on the three initially enriched variants (2xSH Fc-C239i, 2xCys Fc-C239i, iDSB Fc-C239i), carried out and analysed by Carolina Orozco and Alistair Hines.

To understand the interconversion in more detail, the evolution of the thiol states was monitored over a greater range of denaturant concentrations and time points, again by non-reduced peptide mapping. The conversion from 2xSH to iDSB Fc-C239i was almost complete at denaturant concentrations over 1 M GdmCl (**Figure 4.7 B**, **Table 4.3**). For 2xCys Fc-C239i under equivalent conditions, the conversion to iDSB progressed but not to completion, reinforcing the hypothesis that the doubly cysteinylated form needs to convert into an intermediate, most probably the singly cysteinylated form, before converting into the iDSB (**Figure 4.7 C**, **Table 4.4**). These data are further supported by the observation that rapid and relatively slow formation of iDSB from 2xSH and 2xCys Fc-C239i, respectively, occur over time at a fixed denaturant concentration (1-4 days at 3.5 M GdmCl, **Figure 4.7 D&E**, and **Table 4.6**).

Table 4.3: Effect of the concentration of denaturant after 7 days of incubation on the final thiol states of Fc-C239i 2xSH enriched variant

[GdmCl] (M)	**iDSB (mean ± SD)**	**2xCys (mean ± SD)**	**1xCys + 1xSH (mean ± SD)**	**2xSH (mean ± SD)**	**2x GSH (mean ± SD)**	**1xGSH + 1xCys (mean ± SD)**	**1xGSH + 1xSH (mean ± SD)**
0	2.70 ± 0.01	0.080 ± 0.003	0.41 ± 0.03	96.80 ± 0.02	0 ± 0	0 ± 0	0 ± 0
0.5	24.5 ± 0.4	0.6 ± 0.01	0.272 ± 0.002	74.5 ± 0.4	0 ± 0	0.07 ± 0.02	0 ± 0
1.0	86.6 ± 0.4	1.8 ± 0.1	0 ± 0	11.3 ± 0.5	0 ± 0	0.25 ± 0.04	0 ± 0
1.5	92.05 ± 0.02	1.5 ± 0.1	0 ± 0	6.3 ± 0.1	0 ± 0	0 ± 0	0 ± 0
2.0	96.3 ± 0.1	1.8 ± 0.1	0 ± 0	1.8 ± 0.1	0 ± 0	0.06 ± 0.03	0 ± 0
2.5	96.9 ± 0.2	2.0 ± 0.1	0 ± 0	0.9 ± 0.1	0.1 ± 0.0	0 ± 0	0 ± 0
3.0	98.1 ± 0.1	1.3 ± 0.1	0 ± 0	0.56 ± 0.01	0 ± 0	0 ± 0	0 ± 0
3.5	97.3 ± 0.7	2.4 ± 0.7	0 ± 0	0.32 ± 0.01	0 ± 0	0 ± 0	0 ± 0

Experiments were performed at 25 °C. Thiol states with glutathione (GSH) were also monitored. All the values are percentages. Experiments carried out and analysed by Carolina Orozco and Alistair Hines.

Table 4.4: Effect of the concentration of denaturant after 7 days of incubation on the final thiol states of the Fc-C239i 2xCys enriched variant

[GdmCl] (M)	iDSB (mean ± SD)	2xCys (mean ± SD)	1xCys + 1xSH (mean ± SD)	2xSH (mean ± SD)	2x GSH (mean ± SD)	1xGSH + 1xCys (mean ± SD)	1xGSH + 1xSH (mean ± SD)
0	31.3 ± 0.4	54.9 ± 0.2	12.6 ± 0.2	1 ± 0	0.08 ± 0.01	0 ± 0	0 ± 0
0.5	42.3 ± 0.2	52.9 ± 0.1	4.2 ± 0.2	0.46 ± 0.02	0 ± 0	0 ± 0	0 ± 0
1.0	64.2 ± 0.8	35.8 ± 0.9	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
1.5	69.5 ± 0.1	30.5 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
2.0	78.2 ± 0.6	21.8 ± 0.6	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
2.5	82.7 ± 0.3	17.2 ± 0.3	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
3.0	77.71 ± 0.02	22.26 ± 0.01	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
3.5	78.6 ± 0.6	21.4 ± 0.6	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Experiments were performed at 25 °C. Thiol states with glutathione (GSH) were also monitored. Experiments carried out and analysed by Carolina Orozco and Alistair Hines.

Table 4.5: Effect of the incubation time in 3.5 M GdmCl on the final thiol states of the Fc-C239i 2xSH enriched variant

Incubation time (days)	iDSB (mean ± SD)	2xCys (mean ± SD)	1xCys + 1xSH (mean ± SD)	2xSH (mean ± SD)	2x GSH (mean ± SD)	1xGSH + 1xCys (mean ± SD)	1xGSH + 1xSH (mean ± SD)
0	10.6 ± 0.3	0 ± 0	0 ± 0	89.0 ± 0.3	0 ± 0	0 ± 0	0 ± 0
1	96.5 ± 0.2	1.9 ± 0.1	0 ± 0	1.6 ± 0.1	0 ± 0	0 ± 0	0 ± 0
2	97.5 ± 0.1	2 ± 0	0 ± 0	1 ± 0	0 ± 0	0 ± 0	0 ± 0
3	97.3 ± 0.6	1.8 ± 0.6	0 ± 0	1 ± 0	0 ± 0	0 ± 0	0 ± 0
4	97.3 ± 0.2	1.9 ± 0.1	0 ± 0	0.8 ± 0.1	0 ± 0	0 ± 0	0 ± 0

Experiments were performed at 25 °C. Thiol states with glutathione (GSH) were also monitored. Experiments carried out and analysed by Carolina Orozco and Alistair Hines.

Table 4.6: Effect of the incubation time in 3.5 M GdmCl on the final thiol states of the Fc-C239i 2xCys enriched variant

Incubation time (days)	iDSB (mean ± SD %)	2xCys (mean ± SD %)	1xCys + 1xSH (mean ± SD %)	2xSH (mean ± SD %)	2x GSH (mean ± SD %)	1xGSH + 1xCys (mean ± SD %)	1xGSH + 1xSH (mean ± SD %)
0	41.4 ± 0.3	55.6 ± 0.4	1.80 ± 0.02	1 ± 0	0.09 ± 0.01	0 ± 0	0 ± 0
2	58.1 ± 0.1	41.7 ± 0.1	0 ± 0	0 ± 0	0.09 ± 0.01	0 ± 0	0 ± 0
3	66.2 ± 0.1	33.7 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
4	69.6 ± 0.1	30.36 ± 0.03	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Experiments were performed at 25 °C. Thiol states with glutathione (GSH) were also monitored. Experiments carried out and analysed by Carolina Orozco and Alistair Hines.

To explain why significant interconversion only occurs in the presence of denaturant (≥ 1 M GdmCl), chemical denaturation unfolding experiments were performed and the thermodynamic stability of both C_H2 and C_H3 domains was determined for each of the Fc-C239i enriched variants. Tryptophan fluorescence spectra were measured after seven days of incubation in various concentrations of GdmCl and the data analyzed using the average emission wavelength (AEW, **Equation 2.3**) and fitted to a three-state model (**Equation 2.7**). Three thermodynamic parameters were obtained: $\Delta G_{X-Y}^{H_2O}$, the difference in Gibbs free energy between two states X and Y in water; m_{X-Y}, the m-value between the two states X and Y, a constant of proportionality describing how much the ΔG_{X-Y} changes upon denaturant concentration; and $[den]_{50\%\,X-Y}$, the midpoint of denaturation between states X and Y. All Fc variants showed two unfolding transitions (**Figure 4.12 A**), the first, at lower denaturant concentrations, corresponding to unfolding of the C_H2 domain and the second, at higher denaturant concentrations, corresponding to the unfolding of the C_H3 domain[153]. The midpoint of denaturation observed for C_H2 domain unfolding in the 2xSH, 2xCys and iDSB Fc-C239i were 0.58, 0.91 and 0.75 M GdmCl, respectively (**Table 4.8**). These low denaturation midpoints explain why the conversion from 2xSH and 2xCys to iDSB Fc-C239i starts to occur at 0.5 M and becomes significant over 1 M GdmCl, and suggest the C_H2 domain has unfolded, at least in part, under these conditions thereby reducing steric constraints and increasing the frequency of disulfide bond formation at the inserted cysteine.

Reproducibility was verified as the thermodynamic parameters obtained for the chemical denaturation curves performed in triplicates (NIST mAb Fc, 2xSH and 2xCys Fc-C239i) or in duplicate (iDSB Fc-C239i) are very similar or within error (**Figure 4.8 A-D, Table 4.7).**

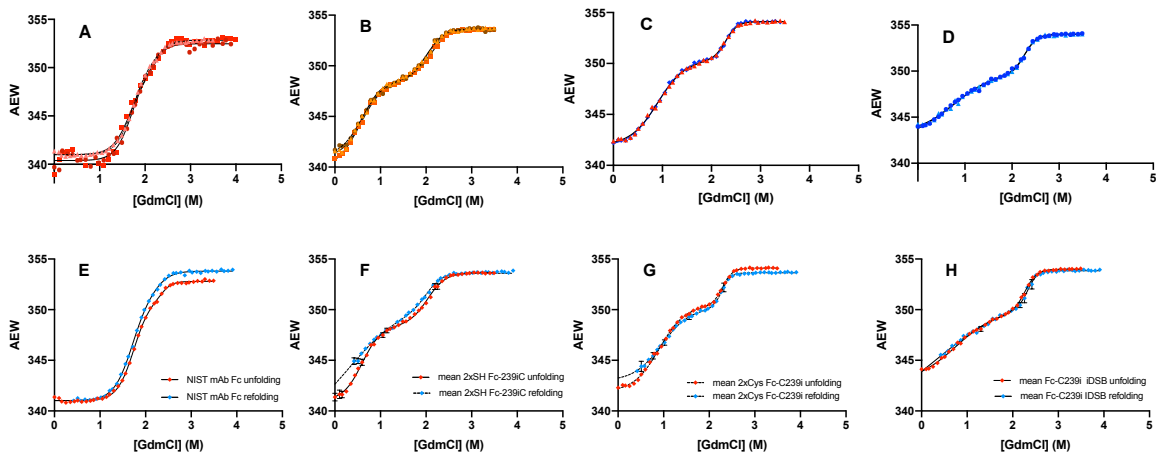


Figure 4.8: Chemical denaturation curves for the unfolding/refolding of enriched variants of Fc-C239i. All the denaturation curves were done at 1 μM final protein concentration in 20 mM His pH 5.5 and samples were incubated at 25 °C for seven days. **A.** Triplicates of the unfolding curves of NIST mAb Fc (WT) after seven days of incubation. **B.** Triplicates of the unfolding curves of 2xSH Fc-C239i enriched variant after seven days of incubation. **C.** Duplicates of the unfolding curves of 2xCys Fc-C239i enriched variant after seven days of incubation. **D.** Duplicates of the unfolding curves of iDSB Fc-C239i enriched variant after seven days of incubation. **E.** Unfolding (red – one curve) and refolding (blue – one curve) curves of NIST mAb Fc (WT) after 7 days of incubation. **F.** Unfolding (red - mean of triplicates) and refolding (blue - mean of triplicates) curves of 2xSH Fc-C239i. **G.** Unfolding (red – one curve) and refolding (blue - mean of triplicates) curves of 2xCys Fc-C239i. **H.** Unfolding (red – mean of duplicates) and refolding (blue - mean of duplicates) curves of iDSB Fc-C239i.

Table 4.7: Thermodynamic parameters from the unfolding/refolding curves

		m_{I-N} (kcal mol^{-1} M^{-1})	$[den]_{50\% I-N}$ (M)	ΔG_{I-N} (kcal mol^{-1})	m_{D-I} (kcal mol^{-1} M^{-1})	$[den]_{50\% D-I}$ (M)	ΔG_{D-I} (kcal mol^{-1})
2xSH Fc-C239i	unfolding	2.5 ± 0.2	0.58 ± 0.02	1.5 ± 0.2	3.1 ± 0.1	2.05 ± 0.04	6.4 ± 0.3
	refolding	1.4 ± 0.3	0.52 ± 0.03	0.7 ± 0.2	3.0 ± 0.1	1.98 ± 0.03	5.9 ± 0.3
2xCys Fc-C239i	unfolding	2.099 ± 0.003	0.9060 ± 0.0001	1.902 ± 0.002	5.3 ± 0.6	2.30 ± 0.01	12 ± 1
	refolding	1.94 ± 0.01	0.96 ± 0.08	1.9 ± 0.2	6.3 ± 0.1	2.28 ± 0.03	14.4 ± 0.3
iDSB Fc-C239i	unfolding	1.3 ± 0.1	0.75 ± 0.05	1.0 ± 0.1	4.5 ± 0.2	2.272 ± 0.002	10.3 ± 0.5
	refolding	1.0 ± 0.2	0.6 ± 0.08	0.7 ± 0.1	5.5 ± 0.3	2.31 ± 0.02	12.8 ± 0.8
NIST mAb Fc	unfolding	3.2 ± 0.1	1.69 ± 0.02	5.5 ± 0.2	3.9 ± 0.5	2.25 ± 0.04	9 ± 1
	refolding	3.0 ± 0.1	1.69 ± 0.01	5.0 ± 0.2	3.2 ± 0.4	2.24 ± 0.02	7.1 ± 0.8

The values for NIST mAb Fc refolding are from a single run and the error shown is a fitting error. In other cases, the mean value and standard deviations are shown for 2xCys Fc-C239 and iDSB Fc-C239i unfolding and refolding (run in duplicate), and 2xSH Fc-C239i unfolding and refolding, 2xCys Fc-C239i refolding and NIST mAb Fc unfolding (run in triplicate).

The reversibility of chemical denaturation of the Fc constructs was investigated by recording refolding curves, where the protein has been unfolded at the start, then diluted into a range of denaturant concentrations and finally measured once equilibrium has been reached. For Fc-C239i enriched in the free thiol form, the m-value of the first unfolding transition between the native and intermediate states (m_{I-N}) is quite different to that obtained from the unfolding curve data. This might be due to the fact that there is no native baseline in the refolding curves as data collection starts at 0.4 M GdmCl, meaning it is difficult to get an accurate m-value. Nevertheless, the denaturation midpoints are very similar, which establishes unfolding is fully reversible. For the other mutants, the thermodynamic parameters are within error, which proves that the denaturation is reversible and at equilibrium for these time points. Replicate refolding denaturation curves are shown in **Figure 4.9**.

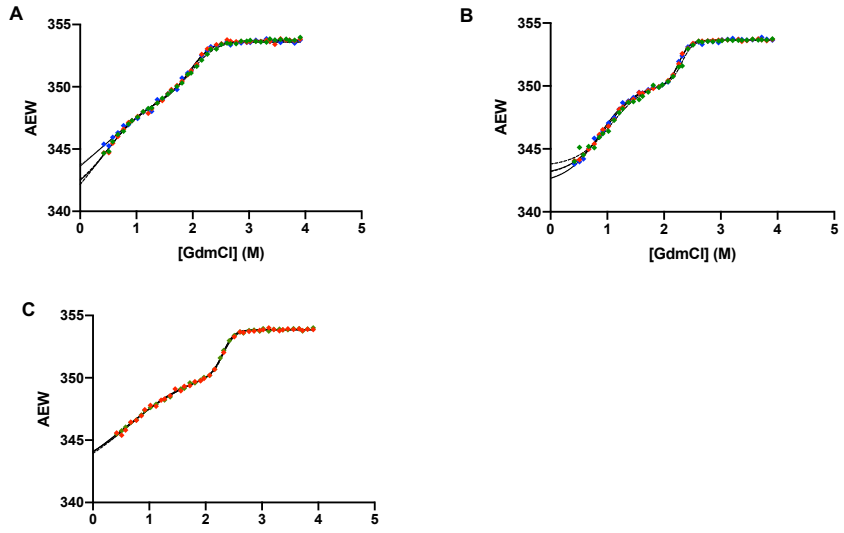


Figure 4.9: Repeats of chemical renaturation curves of the Fc-C239i formats. A. Triplicates (red, green, blue) of the refolding curves of 2xSH Fc-C239i. **B.** Triplicates (red, green, blue) of the refolding curves of 2xCys Fc-C239i. **C.** Duplicates (red, green) of iDSB Fc-C239i. All these denaturation curves were performed at a 1 µM final concentration of protein in 20 mM His pH 5.5 incubated at 25 °C for seven days.

Formation of iDSB also occurs in the absence of chemical denaturant: for example, when 2xSH IgG-C239i was incubated at 37 °C (in 10 mM Tris, pH 8.0) for 13 days the proportion of iDSB increased from 20% to 100%. The rate of iDSB formation increased further under heat stress with almost complete (98%) conversion after incubation at 50 °C for 24 hours (**Figure 4.10**). Increasing the temperature may affect both the proportion of molecules that are in the denatured state (even below the T_m value) as well as increase protein dynamics and thereby potentially increasing the number of local as well as global unfolding events. It is not possible to deconvolute these two effects without further experiments, therefore it is unclear whether the iDSB forms in a fully unfolded state resulting from global unfolding of the C_H2 domain, or a transiently populated partially unfolded state. What is clear is that iDSB formation occurs under physiologically relevant conditions and increases under any conditions where the stability of the native state is decreased or there is an increase in dynamics.

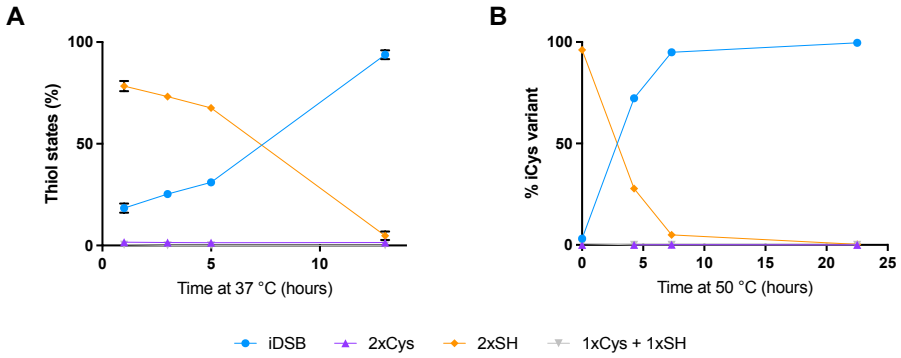


Figure 4.10: Effect of heat on the thiol state evolution over time starting from Fc-C239i 2xSH enriched material: A. at 37 °C; **B.** at 50 °C. Experiment carried out and analysed by Dr Paul Devine and Alistair Hines.

A scheme for the interconversion of the different thiol states is shown in **Figure 4.11**, assuming 1xCys + 1xSH can interconvert into iDSB Fc-C239i by nucleophilic attack [166,167]. 2xCys can interconvert into 1xCys+1xSH which can interconvert into 2xSH from either the native or denatured states, and both 2xSH and 1xCys+1xSH can interconvert into iDSB via the denatured state or with increased molecular dynamics. The 1xCys+1xSH form could interconvert to iDSB by nucleophilic attack of the free thiol to the capped cysteine.

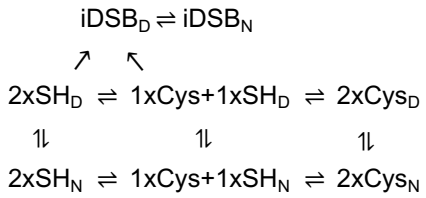


Figure 4.11: Interconversion network between the 2xSH, 2xCys, iDSB Fc-C239i variants. N: C_H2 domain in the native state. D: C_H2 domain in the denatured state or with increased molecular dynamics.

4.3.2 Effect of the cysteine insertion on the biophysical stability of the antibody

To determine if the thermodynamic stability of the antibody had been affected by insertion of the cysteine or the resultant thiol variants, thermodynamic parameters were measured from the chemical denaturation unfolding experiments. All enriched variants showed lower denaturation midpoints for the first transition compared to the wild type (**Table 4.8**), suggesting that the C_H2 domain was destabilized by the insertion of the cysteine in the upper C_H2 domain. The $\Delta G_{I-N}^{H_2O}$ values, which correspond to the thermodynamic stability of the C_H2 domain, demonstrate that this domain in iDSB Fc-C239i is the least stable, followed by 2xSH Fc-C239i and finally 2xCys Fc-C239i, which

are all significantly less thermodynamically stable than the wild type (**Table 4.8**). Since the m-values are correlated with the difference in solvent accessible surface area[168] ($\Delta SASA$) between native and denatured states, they provide useful information on the effect of the insertion on the structure. All three variants have a lower m_{I-N} value than the wild type, showing that the C_H2 domain has a more solvent accessible surface area in the native state (if there are no effects on the denatured state) suggesting that conformational changes have occurred. Interestingly, the iDSB Fc-C239i has the lowest m_{I-N} value, demonstrating that either the C_H2 domain has more solvent exposed surface area in the native state, or that the denatured state is more structured. The midpoint of the second unfolding transition is similar for all constructs, suggesting that the stability of the C_H3 domain is unaffected by the insertion in the C_H2 domain.

Differential scanning calorimetry (DSC) was employed to measure the thermal stability of the different variants. All constructs showed two unfolding transitions, the first, at lower temperatures, corresponding to unfolding of the C_H2 domain and the second, at higher temperatures, to unfolding of the C_H3 domain (**Figure 4.12 B**). The thermal stability of the C_H2 domain was lower for all Fc-C329i variants compared to the wild type, and the relative thermal stabilities were the same as those determined using chemical denaturation: iDSB < 2xSH< 2xCys < 2xSH with NEM< NIST mAb Fc (**Figure 4.12 B, Table 4.9**). The NEM capping of 2xSH prevents the interconversion of 2xSH into other states during the stability measurements, therefore, it is important to keep in mind that in the absence of NEM, the enriched states 2xSH and 2xCys will convert to the iDSB during the measurements. Therefore, in the stability measurements, although they start as enriched in one thiol state (either 2xSH or 2xCys), their fraction of iDSB present will rise as denaturant concentration and temperature increase. The only data sets with the same thiol state throughout the experiments are the iDSB-enriched variant and the 2xSH NEM capped state. Thermal unfolding was found to be irreversible, as the unfolding trace of a sample that had already been heated and cooled did not overlap with the initial one (**Figure 4.13**). ΔH_{cal}, corresponding to the area under an unfolding peak, is the calorimetric determination of the change in enthalpy of unfolding and is a measure of the favourable interactions that must be overcome to unfold the protein. ΔH_{cal} for the first transition, corresponding to the unfolding of the C_H2 domain, for the iDSB-enriched Fc-C239i variant was significantly lower than for all the other Fc thiol states and wild type (**Table 4.9**), indicating that favourable interactions have been lost in the native state of

iDSB Fc-C239i. In addition, this peak was also broader for the iDSB Fc-C239i (**Figure 4.12 B**), than for other states/wild type suggestive of a lower ΔC_p . This indicates a lower $\Delta SASA$[168], which is consistent with the lower m-value observed for iDSB Fc-C239i in the GdmCl denaturation experiments. Both results suggest that either the native state of the C_H2 domain is more solvent accessible in the iDSB variant than in wild type, or that the denatured state may be more structured. The T_m1 of Fc-C239i increased when the free thiols were capped with NEM suggesting the apparent reduction in relative thermal stability of Fc-C239i 2xSH may be due to it converting into iDSB at elevated temperatures, consistent with previous observations (**Figure 4.10**). The melting temperatures of the C_H3 domain for all constructs were very similar, confirming the insertion has no effect on the stability of this domain.

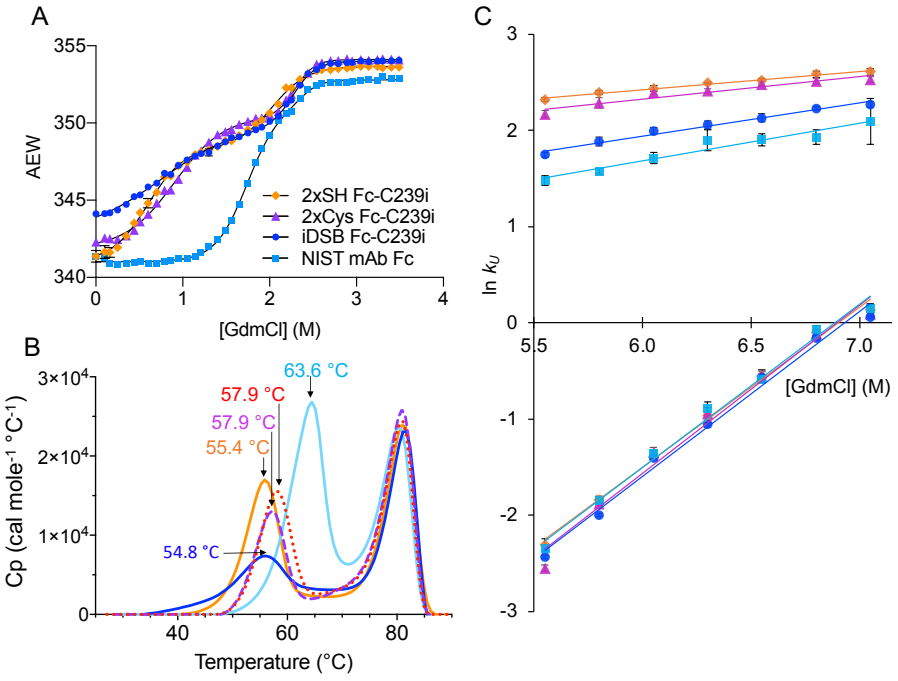


Figure 4.12: Biophysical experiments on the unfolding of the enriched variants of Fc-C239i. A. Unfolding curves of 2xSH Fc-C239i (orange lozenge), 2xCys Fc-C239i (purple triangles), iDSB Fc-C239i (dark blue circles) and NIST mAb Fc (light blue squares). In all cases, the values shown are the average from multiple experiments. **B.** Results from DSC experiments. Thermal stabilities of 2xSH Fc-C239i (orange, solid line), 2xSH NEM-capped Fc-C239i (red, dotted line), 2xCys Fc-C239i (purple, dashed line), iDSB Fc-C239i (dark blue, solid line), NIST mAb Fc (light blue, solid line). The experiment was performed in triplicate but the data shown is only for one experiment. The T_m values shown are the mean value from multiple measurements. **C.** Unfolding kinetics 2xSH Fc-C239i (orange lozenge), 2xCys Fc-C239i (purple triangles), iDSB Fc-C239i (dark blue circles) and NIST mAb Fc (light blue squares), measured in triplicate. The fastest unfolding phases correspond to the C_H2 domain unfolding, whereas the slower unfolding phases correspond to the unfolding of the C_H3 domain. The solid line shows the best fit of the data to **Equation 2.12**. Error bars represent the standard deviation from triplicate measurements. In many cases, they are not visible because they are smaller than the size of the data point. All biophysical experiments were all carried out in 20 mM His pH 5.5.

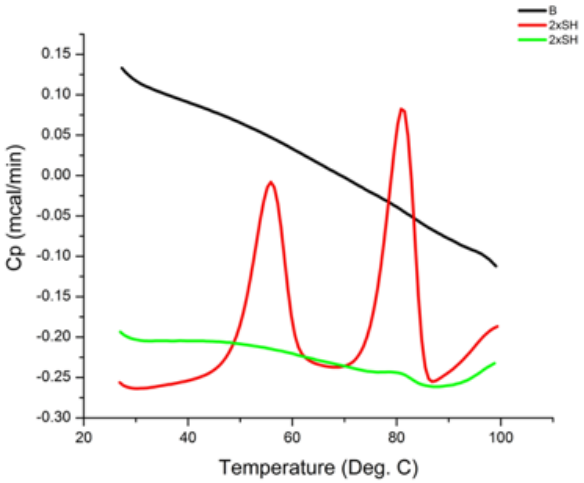


Figure 4.13: Raw data of the thermogram from the DSC experiments on the 2xSH-enriched Fc-C239i variant. The sample was heated a first time (red), cooled and reheated (green). The black line is the buffer baseline, 20 mM His pH 5.5.

Table 4.8: Thermodynamic parameters from the chemical denaturation experiments of the Fc fragment of enriched variants of C239i and wild type

Proteins	m_{I-N} (kcal mol^{-1} M^{-1})	$[den]_{50\% I-N}$ (M)	ΔG_{I-N} (kcal mol^{-1})	m_{D-I} (kcal mol^{-1} M^{-1})	$[den]_{50\% D-I}$ (M)	ΔG_{D-I} (kcal mol^{-1})
2xSH Fc-C239i	2.5 ± 0.2	0.58 ± 0.02	1.5 ± 0.2	3.1 ± 0.1	2.05 ± 0.04	6.4 ± 0.3
2xCys Fc-C239i	2.099 ± 0.003	0.9060 ± 0.0001	1.902 ± 0.002	5.3 ± 0.6	2.30 ± 0.01	12 ± 1
iDSB Fc-C239i	1.3 ± 0.1	0.75 ± 0.05	1.0 ± 0.1	4.5 ± 0.2	2.272 ± 0.002	10.3 ± 0.5
NIST mAb Fc	3.2 ± 0.1	1.69 ± 0.02	5.45 ± 0.2	3.9 ± 0.5	2.25 ± 0.04	9 ± 1

The values shown are the average of duplicate (2xCys Fc-C239i, iDSB Fc-C239i) and triplicate (2xSH Fc-C239i, NIST mAb Fc) measurements and the error is the standard deviation between measurements.

Table 4.9: Melting temperatures corresponding to the thermal unfolding of C_H2 and C_H3 domains of the antibody C239i variants

Antibody variant	$\boldsymbol{T_m1}$ (°C) C_H2 domain	$\boldsymbol{T_m2}$ (°C) C_H3 domain	$\Delta\boldsymbol{H1_{cal}}$ (kcal mol^{-1}) C_H2 domain	$\Delta\boldsymbol{H2_{cal}}$ (kcal mol^{-1}) C_H3 domain
NIST mAb Fc	63.6 ± 0.2	79.7 ± 0.1	(1.7 ± 0.5) x 10^2	(1.4 ± 0.2) x 10^2
2xSH Fc-C239i	55.4 ± 0.1	80.0 ± 0.3	(1.3 ± 0.3) x 10^2	(1.36 ± 0.06) x 10^2
1 2xSH NEM capped Fc-C239i	57.9 ± 0.2	80.1 ± 0.1	(1.4 ± 0.3) x 10^2	(1.27 ± 0.01) x 10^2
2xCys Fc-C239i	56.6 ± 0.3	80.2 ± 0.1	(1.3 ± 0.4) x 10^2	(1.2 ± 0.1) x 10^2
iDSB Fc-C239i	54.8 ± 0.3	80.6 ± 0.2	(0.5 ± 0.2) x 10^2	(1.4 ± 0.1) x 10^2

1The Fc-C239i 2xSH NEM capped format was run just twice. The errors are the standard deviations from repeated measurements. $\Delta\boldsymbol{H1_{cal}}$ was determined via a calorimetric experiment (area under the curve).

Table 4.10: Kinetic parameters from unfolding experiments

	$k_U^{H_2O}$ (s^{-1})	m_{k_U} (M^{-1})
ln(k1) NIST mAb Fc	0.5 ± 0.1	0.40 ± 0.04
ln(k1) 2xSH	3.6 ± 0.3	0.19 ± 0.01
ln(k1) 2xCys	2.6 ± 0.5	0.23 ± 0.03
ln(k1) iDSB	0.9 ± 0.1	0.34 ± 0.02
ln(k2) NIST mAb Fc	(9 ± 4) x 10^{-6}	1.69 ± 0.07
ln(k2) 2xSH	(11 ± 4) x 10^{-6}	1.67 ± 0.05
ln(k2) 2xCys	(6 ± 4) x 10^{-6}	1.8 ± 0.1
ln(k2) iDSB	(8 ± 4) x 10^{-6}	1.71 ± 0.08

The kinetic parameters shown are from the best fit of the data shown in **Figure 4.12 C** to **Equation 2.12.**

Unfolding kinetics provide information on the rate (and therefore frequency) at which a domain unfolds. Natively folded protein was rapidly diluted into a range of chemical denaturant concentrations and the change in intrinsic fluorescence on unfolding measured. The kinetic data for all Fc variants (**Figure 4.12 C, Table 4.10**) were fit with a double exponential function (**Equation 2.10**) that best describes both a fast and slower unfolding phase. The unfolding rate constants of the slower phase were very similar for all variants (**Figure 4.12 C, Table 4.10**) suggesting that this phase corresponds to the unfolding of the C_H3 domain and it, therefore, does not change significantly with insertion of the cysteine at position 239 in the C_H2 domain. This is in agreement with the results on thermodynamic stability described earlier, as well as the relative kinetic stabilities reported by Sumi and Hamaguchi[153] for a different antibody. The unfolding rates of the C_H2 domain for 2xSH and 2xCys Fc-C239i were very similar, and faster than wild type (**Figure 4.12 C**), suggesting that the insertion of a cysteine in the C_H2 domain kinetically destabilizes it. Interestingly, the rate of unfolding of iDSB was slower than the 2xSH and 2xCys Fc-C239i variants indicating a higher kinetic stability. This is in contrast to the thermodynamic measurements and is likely due to the unfolding transition state being more structured than for 2xSH and 2xCys Fc-C239i. A more detailed representation of the energy levels for the native (N) and transition states (TS) for the C_H2 domain only of each variants is provided in **Figure 4.14**. When comparing 2xSH and 2xCys Fc-C239i to wild type (**Figure 4.14 A**), the energy of the native state and to a lesser extent the TS, has increased as favourable interactions have been lost. The energy of the denatured state is assumed to be the same, therefore, $\Delta G_{I\text{-}N}$ (WT) > $\Delta G_{I\text{-}N}$ (2xSH, 2xCys Fc-C239i), i.e. the C_H2 domain of WT is thermodynamically more stable than the inserted cysteine mutant in either the free thiol or cysteinylated forms. From the kinetic data, $\Delta G_{TS\text{-}N}$ (WT) > $\Delta G_{TS\text{-}}$

$_N$ (2xSH, 2xCys Fc-C239i), i.e., the rate constant for unfolding of the C_H2 domain is smaller for WT than the inserted cysteine mutant in either the free thiol or cysteinylated forms and therefore the WT is more kinetically stable than the mutant. When comparing the iDSB Fc-C239i to the wild type (**Figure 4.14 B**), the energy of the native state (and to a lesser extent the TS) has increased as favourable interactions have been lost. In this case, the energy of the denatured state is higher as the additional disulfide bond restricts movement and decreases entropy.

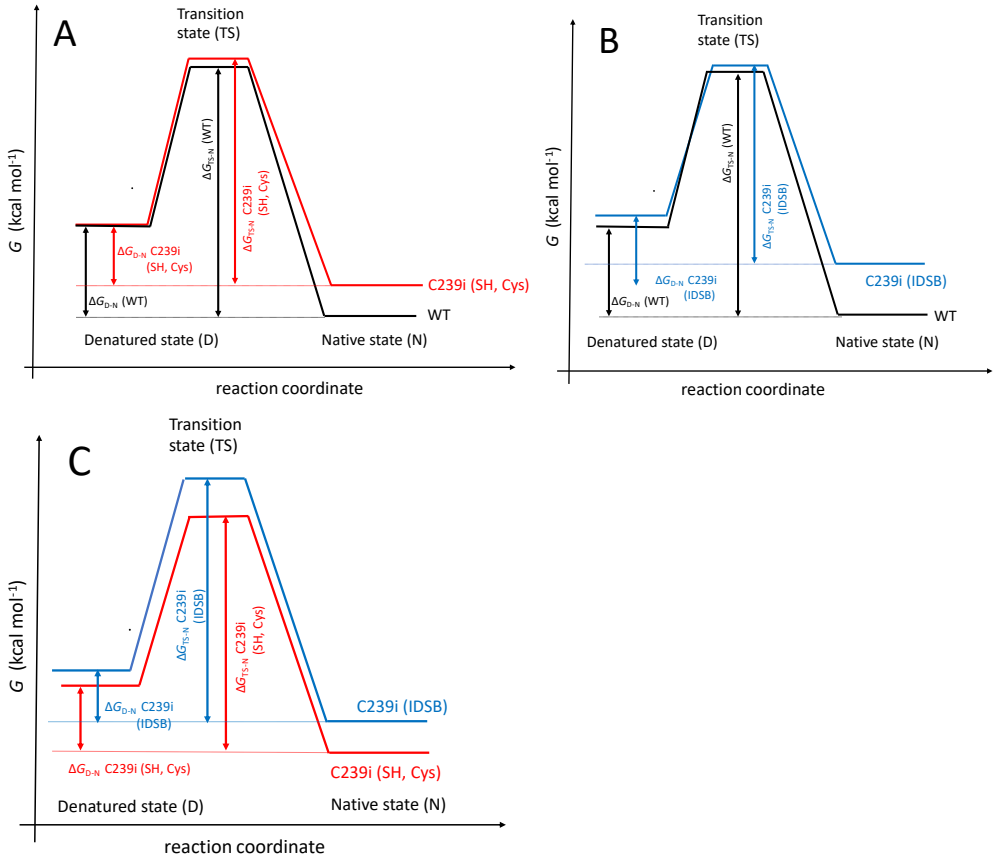


Figure 4.14: Free energy diagrams combining the results from thermodynamic and kinetic stability experiments for the C_H2 domain of Fc-C239i variants and wild type. A. Free energy diagrams of 2xSH and 2xCys Fc-C239i (red) compared with wild type (black). **B.** Free energy diagrams of iDSB Fc-C239i (blue) compared with wild type (black). **C.** Free energy diagrams of iDSB Fc-C239i (blue) compared with 2xSH and 2xCys Fc-C239i (red).

From the thermodynamic data, it is observed that the WT is thermodynamically more stable than the inserted cysteine mutant in the iDSB form consistent with the fact that formation of the iDSB greatly disrupts the structure and interactions in the native state to a much larger degree that it increases the energy of the denatured state by restricting entropy. From the kinetic data, it is observed that $\Delta G_{TS\text{-}N}$ (WT) > $\Delta G_{TS\text{-}N}$ (iDSB Fc-C239i) i.e., the rate constant for unfolding is smaller for WT than the iDSB variant and therefore the WT is more kinetically stable than the mutant. However, when comparing the

thermodynamic data for iDSB Fc-C239i with 2xSH and 2xCys Fc-C239i (**Figure 4.14 C**), Fc-C239i 2xSH and 2xCys are more stable than the iDSB. From the kinetic data: ΔG_{TS-N} (iDSB Fc-C239i) > ΔG_{TS-N} (2xSH, 2xCys Fc-C239i), i.e., the rate constant for unfolding is smaller for iDSB Fc-C239i than the 2xSH and 2xCys Fc-C239i) and therefore the iDSB is more kinetically stable than the free thiol, cysteinylated forms even though it is less thermodynamically stable.

4.3.3 Effect of the cysteine insertion on structural dynamics

To investigate whether the thermodynamic and kinetic differences associated with the formation of an intramolecular disulfide bond were related to local changes in structure or dynamics, and to identify the regions affected, hydrogen-deuterium exchange mass spectrometry (HDX-MS) experiments were conducted. This implementation of HDX-MS uses a novel fully-automated system capable of measuring hydrogen-deuterium exchange in the millisecond to minutes time scale[139]. Post labelling, online pepsin digestion was conducted and complete sequence coverage was achieved in the resulting peptide map (**Figure 4.15**). The samples were labelled in D_2O for various incubation times (1000, 6000, 30000, 60000, 300000, 600000, and 900000 ms at 20 °C) and from the peptide deuteration levels, the normalised differences per residue were calculated, between the Fc-C239i enriched samples and the wild type (**Figure 4.16**).

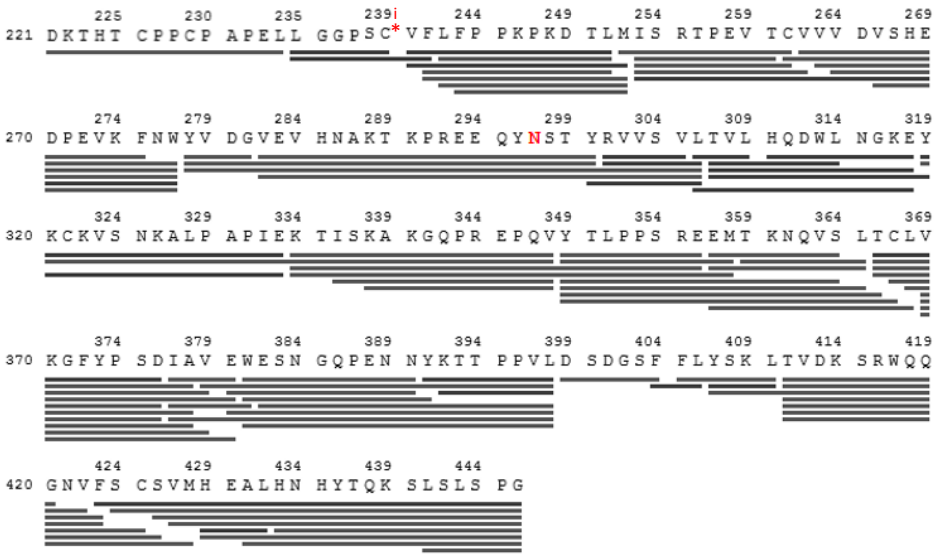


Figure 4.15: Peptide coverage of the sequence of 2xCys Fc-C239i in the HDX-MS experiments (100% coverage, 93 peptides, redundancy: 4.8). The red asterisk shows the location of the inserted cysteine. The canonical numbering of residues of an IgG1 was kept despite the insertion, therefore serine is at position 239 and the inserted cysteine is referred to as 239i. Red N (Asn297) denotes N-glycan modification site.

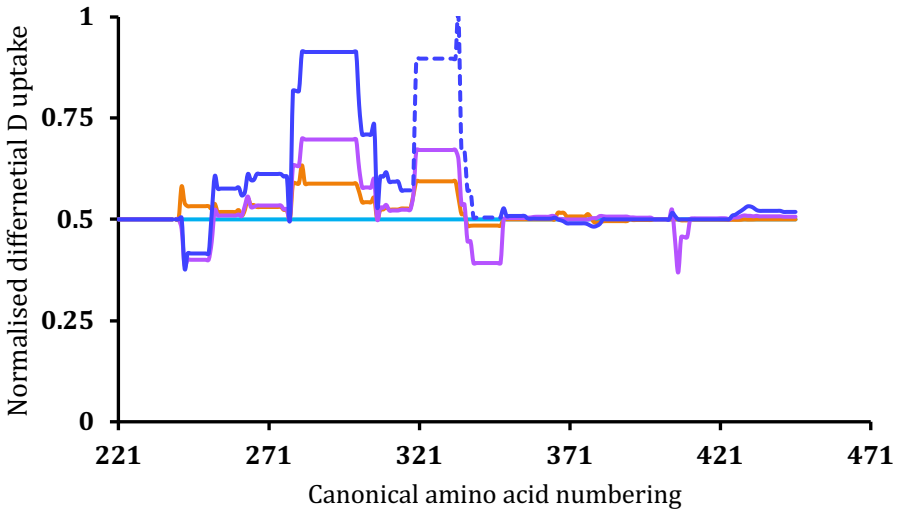


Figure 4.16: Normalized difference plot obtained after summing the significant D-incorporations of the enriched Fc-C239i states relative to the wild type. Light blue: wild type (NIST mAb Fc) baseline. Orange: iDSB enriched Fc-C239i minus WT. Purple: 2xCys enriched Fc-C239i minus WT. Dark blue: iDSB enriched Fc-C239i minus WT. The dashed blue line corresponds to the region showing mixed EX1 and EX2 kinetics.

Interrogation of differences in the rate of deuterium uptake revealed that insertion of a cysteine at position 239 of the heavy chain (Fc-C239i) changes the structure and/or conformational dynamics of the C_H2 domain, which is further, and substantially, disrupted by subsequent formation of an interchain disulfide bond (**Figure 4.17 A**). Significant increases (t-test, p-value < 0.05) in deuterium exchange consistent with either increased solvent exposure and/or reduced intramolecular hydrogen-bonding were observed in three locations: a small increase in exchange was observed between residues Ser254 and Trp277 that are involved in two β-strands facing each other (strands B and C in **Figure 4.17 B&C**); a larger increase was observed in residues located between Val279 and Leu306 which are in a β-strand at the edge of the β-sandwich and beginning of the next β-strand (strands D and beginning of E in **Figure 4.17 B&C**); finally, an increase in exchange was also observed to take place in the two last β-strands of the C_H2 domain (strands F and G, **Figure 4.17 B&C**), where mixed EX1 and EX2 kinetics were observed for the iDSB (**Figure 4.18**). EX1 exchange kinetics are seldom observed and indicative of a local unfolding event. This correlated local unfolding can be precisely localised to amino acids 319-333 in the C_H2 domain: EX1 kinetics are observed in peptide 319-348 and in the shorter overlapping peptide 319-333 (**Figure 4.18**), but not in the other overlapping peptide 334-348 nor other adjacent regions. This is consistent with the conclusion that these anti-parallel strands undergo local unfolding, i.e., unzipping and rezipping without breaking the surrounding H-bonds of the β-sheet. In addition, the H-

bonding network involving β-strand D has been greatly destabilized and partly broken, and the β-sheet composed of the β -strands A, B and E has also been destabilized but to a lesser extent. The mixed EX1/EX2 kinetics suggest that there is a mixed population of folded and unfolded forms of the C_H2 domain for the iDSB variant ($iDSB_D \rightleftharpoons iDSB_N$).

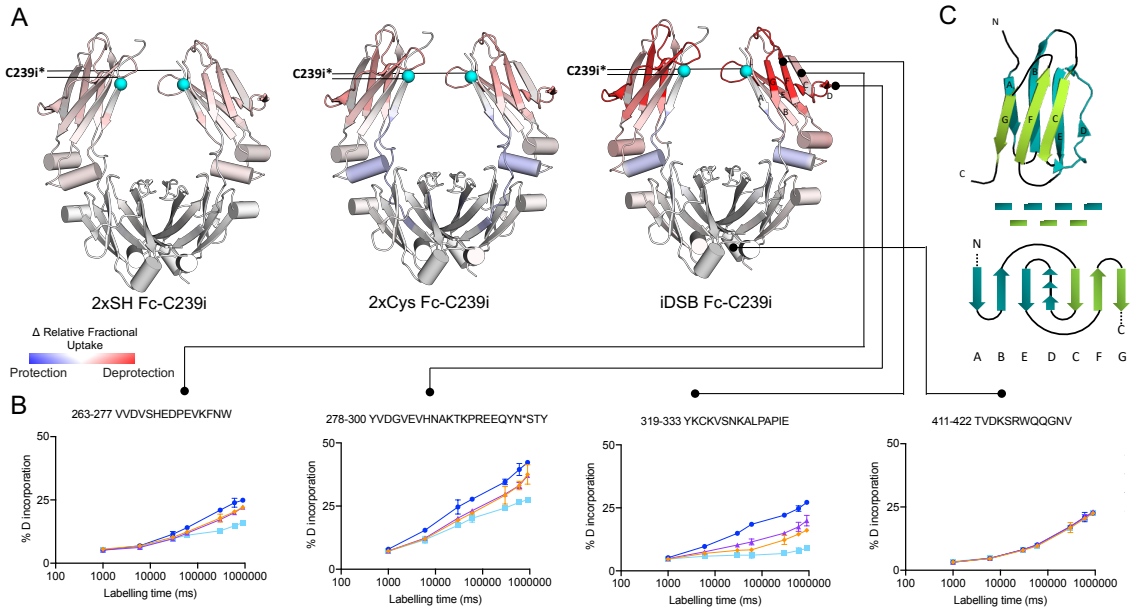


Figure 4.17: HDX-MS results on the various enriched thiol states of Fc-C239i. **A.** Crystal structure (3AVE) showing the relative change in fractional deuterium exchange represented as a color scale: reduced exchange (blue), no change (white), increased exchange (red), for the three enriched Fc-C239i variants compared to wild type and normalized to iDSB Fc-C239i which has the largest differences in exchange rate for all three varinats. **B.** Uptake plots for four peptides covering different regions of the Fc domain which demonstrate different behaviours in terms of changes in HDX uptake: 2xSH Fc-C239i (orange lozenge), 2xCys Fc-C239i (purple triangles), iDSB Fc-C239i (dark blue circles) and NIST mAb Fc (light blue squares). The error bars represent a Student's t-distribution 95% confidence interval, n=3. **C.** Simplified representation of the β–sandwich structure of the C_H2 domain, indicating the arrangement of the β–strands within the β-sheets and tertiary structure. *The crystal structure 3AVE corresponds to the wild type and does not have an inserted cysteine at position 239; the annotation shows where the inserted cysteine is located.

In previous studies[106,169], it had been observed that β-strand G in the C_H2 domain was the most destabilized by other mutations in the C_H2 domain, demonstrating a lower stability in that specific β-strand. Modest protection was observed in the region of the C_H2 domain closest to the C_H3 domain for 2xCys and iDSB, which suggests it comes from the iDSB form (as the 2xCys states also contains some iDSB) and not the insertion itself, potentially as a structural compensation of the distortion in the upper C_H2 domain. Overall, our data suggest that the additional disulfide bridge at the beginning of the C_H2 domain applies strain to a larger region, i.e., across the C_H2 domain, distorting it and opening it up compared to the wild type. The 2xSH and 2xCys Fc-C239i show an increased exchange in the same regions but to a lesser extent. The C_H3 domain does not show any significant

increase in deuterium exchange for any of the enriched variants when compared to the wild type (**Figure 4.17 A&C**). Gallagher *et al.*[162] showed by X-ray crystallography that the inserted cysteine after the 239th residue replaces the position of the serine 239 and affects the preceding residues in the hinge, resulting in a one-residue upward-shift towards the N-terminus of Ser239, Pro238 and Gly237, with residues after the insertion maintaining their wild-type positions. These are rather minimal perturbations that are not consistent with their HDX-MS results. However, the more detailed HDX-MS findings presented here are in agreement with their HDX data, perhaps reflecting that the crystal structure may not represent the true structure in solution due to the restraints imposed by the crystal lattice.

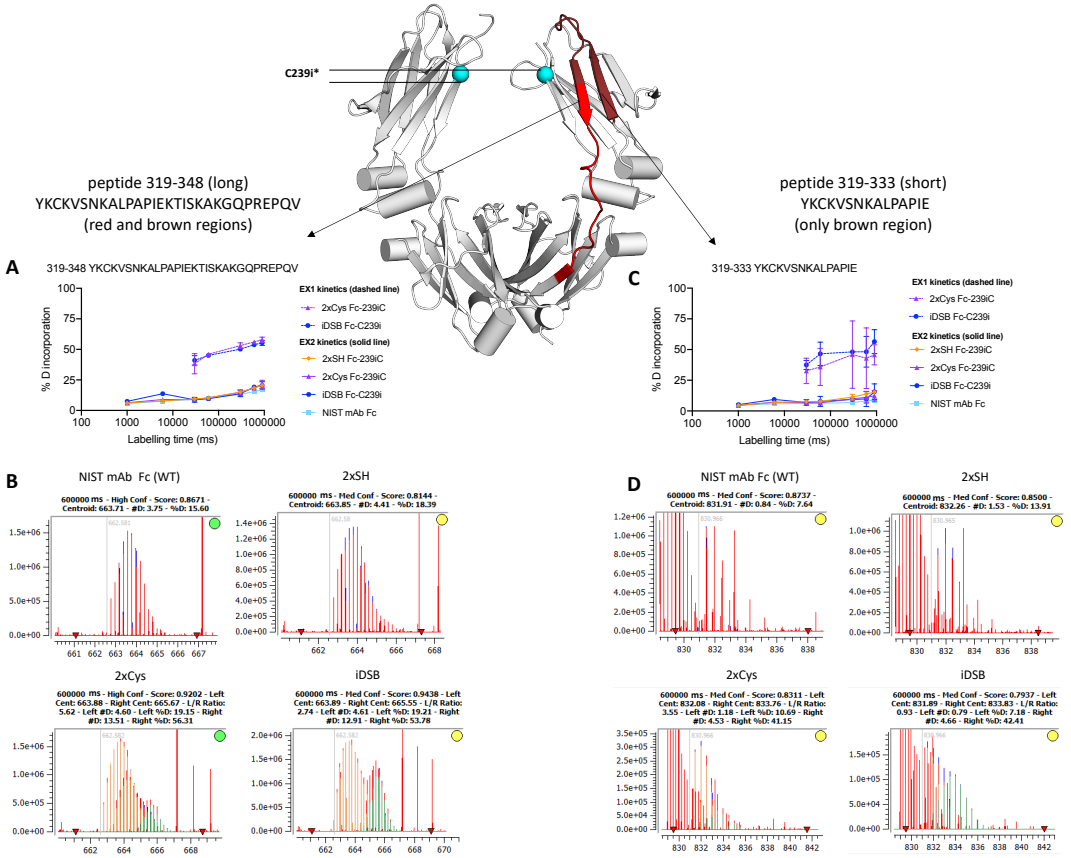


Figure 4.18: Details of mixed EX1/EX2 kinetics observed for HDX in peptides in the last two β-sheets of the C_H2 domain (colored on the crystal structure). A. Uptake plot representing the percentage of deuterium incorporation for the EX1 exchange (dashed line), and EX2 exchange (solid line) for the peptide 319-348 (YKCKVSNKALPAPIEKTISKAKGQPREPQV: brown and red regions on the crystal structure). **B.** Raw spectra representing the unimodal or bimodal distributions for the deuterium uptake at 600 s. In the case of a bimodal distribution, the EX2 distribution is represented in orange and the EX1 distribution is in green. The EX1 distribution is three times more abundant for iDSB than for 2xCys at 600 s: the EX1 kinetics signal shifts by the same ratio as the amount of residual iDSB in the 2xCys enriched sample (33.4%, **Table 4.1**), which supports the hypothesis that the the EX1 kinetics comes from the iDSB form. **C.** Uptake plot representing the percentage of deuterium incorporation for the EX1 exchange (dashed line), and EX2 exchange (solid line) for the peptide 319-333 (YKCKVSNKALPAPIE: red region on the crystal structure). **D.** Raw spectra representing the unimodal or bimodal distributions for the deuterium uptake at 600 s. Analysis by Dr Owen Cornwell and Carolina Orozco.

A second method, developed by Cornwell *et al.*[170], where the deuteration is analyzed time point by time point rather than by addition like the previous method, showed very similar results (**Figure 4.19)**. Both methods agree on the regions of greatest deuterium exchange, which occur in the β-sheet composed of the strands C, F and G (green β -sheet, **Figure 4.17 C, Figure 4.19**) as well as with β-strand D. They also agree on the low protection in the region of the C_H2 domain closest to the C_H3 domain. Both analysis methods are in excellent agreement, highlighting the fact that different approaches in calculating fractional exchange across time points and the statistical method employed to determine significance led to the same result.

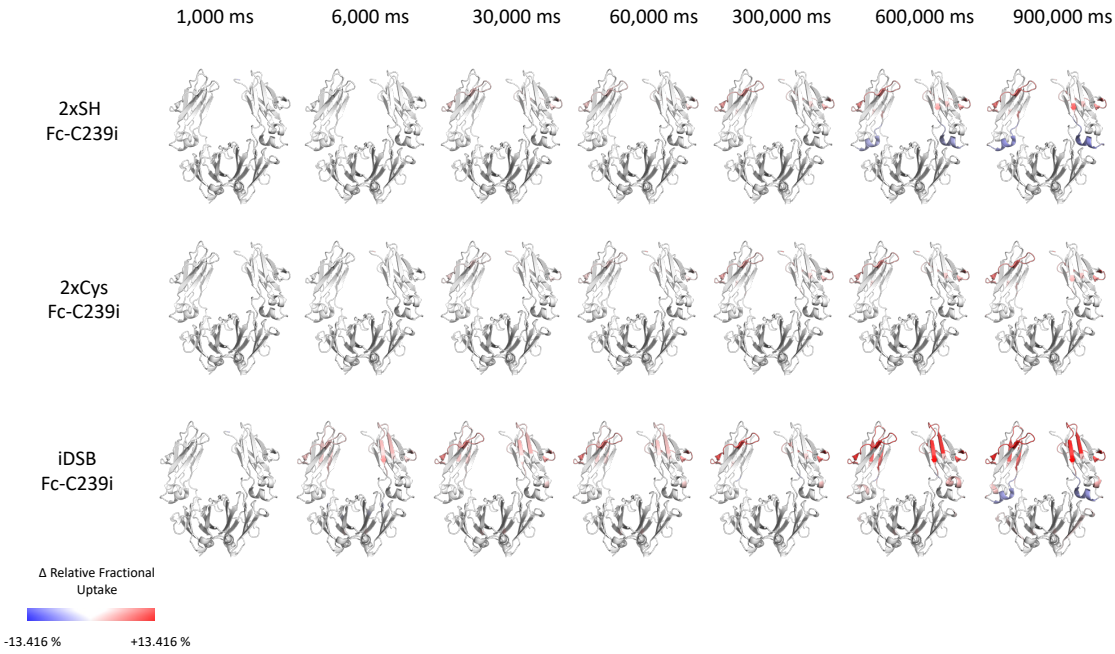


Figure 4.19: **Crystal structure (3AVE) showing the changes in deuterium exchange**, with the color scale blue (less exchange), white (no difference), red (more exchange) per time point obtained with the second analysis method for the three enriched Fc-C239i variants compared to wild type. iDSB Fc-C239i is the variant that exchanges deuterium the most and each variant is normalized to this level. Analysis by Dr Owen Cornwell.

4.3.4 Impact of the additional disulfide bridge on the drug to antibody ratio

Additional experiments were carried out by a collaborator, Christopher Thompson, based in AstraZeneca, Gaithersburg (USA), to assess the impact of the iDSB thiol state on the efficacy of the conjugation and the final DAR. The conjugation process involves preparatory redox steps: a partial reduction, to free the engineered cysteines, a mild reoxidation and the conjugation (**Figure 4.20**). IgG-C239i was conjugated to a payload before and after these preparatory redox steps to compare the conjugation

efficiencies. When the enriched variants were conjugated directly without reduction/reoxidation, the conjugation efficiency on 2xCys and iDSB IgG C239i was very low (**Table 4.11**). In contrast, when the partial reduction and reoxidation were carried out on the enriched variants prior to conjugation, the conjugation results showed consistent conjugation efficiencies leading to DARs of 1.57, 1.54, and 1.51 for 2xSH, 2xCys, and iDSB Fc-C239i, respectively (**Table 4.11**). Non-reduced peptide mapping on these samples showed that there were ca. 20% of iDSB remaining in the final conjugated sample (**Figure 4.21**), which would prevent the DAR from reaching the value of 2.

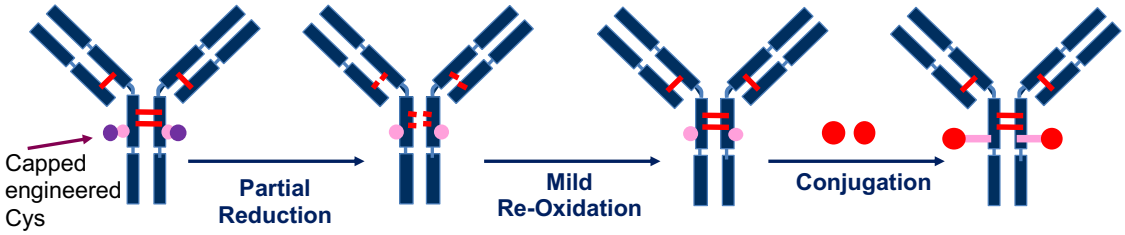


Figure 4.20: Schematic summarising the operations typically employed to conjugate a payload to IgG-C239i. Shown as red bars are the native disulfide bonds, in pink circles the inserted cysteine residue, purple circles represent a capped inserted cysteine, and the payload as a larger red circle.

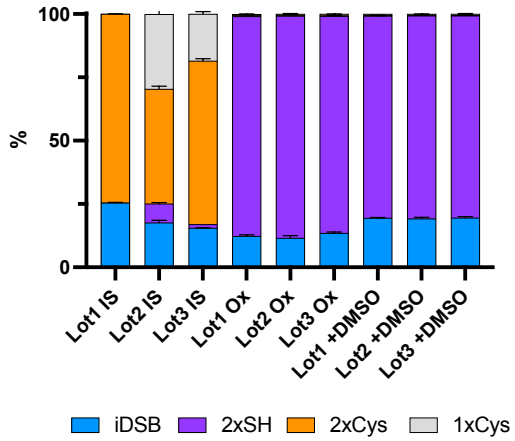


Figure 4.21: Proportion of thiol states in IgG-C239i monitored by non-reduced peptide mapping in their initial state (IS), after partial reduction and mild reoxidation (Ox) and after conditions used for conjugation but in the absence of the payload (DMSO). Approximately 20% of cysteines were in the iDSB state after exposure to conditions used for conjugation. Experiment carried out and analysed by Dr Paul Devine and Alistair Hines.

Table 4.11: Drug antibody ratio after conjugation of the different enriched variants, going through the reduction/oxidation steps

Enriched material	DAR after direct conjugation	DAR after reduction and reoxidation prior to conjugation
2xSH	1.75	1.57
2xCys	0.12	1.54
iDSB	0.08	1.51

4.4 DISCUSSION AND CONCLUSION

Overall, these studies demonstrate that the C_H2 domain in all the enriched variants of Fc-C239i has been destabilized by the cysteine insertion, whereas the C_H3 domain is unaffected. This is in agreement with previous work on antibodies containing either a cysteine insertion C239i[35] or substitution S239C[34] which showed that the thermal stability of the C_H2 domain was not affected by the substitution, but was decreased by the insertion.

Multiple states of the Fc-C239i were observed including one in which both inserted cysteines were present as unmodified thiols (2xSH), and another where the thiols of the cysteines residue had both reacted with a cysteine amino acid in the culture media to form a doubly cysteinylated state (2xCys). Surprisingly, another naturally occurring state of the C239i mutant was found containing an additional disulfide bridge (iDSB) formed between the inserted cysteines on each of the two heavy chains of the antibody. This was unexpected as the iDSB can form only when the two the two heavy chains can come close enough to form a disulfide bridge (the sulfur atoms need to be 2 Å apart[171]), yet these side chains are far apart in the crystal structure of the Fc domain of the parent antibody[172,173] (PDB: 3AVE). Further experiments demonstrated that the three enriched Fc-C239i variants can interconvert, with 2xSH and 2xCys states forming the iDSB over time, the 2xSH state converting more rapidly than the 2xCys state.

The conditions under which the iDSB forms were investigated in detail. It was known that some iDSB initially forms during expression and purification, and further mechanistic studies were undertaken to establish the effect of stress conditions, under which there is either an increase in molecular dynamics or an increased proportion of Fc where the C_H2 domain is in the denatured state, on iDSB formation. Both of which could lead to increased iDSB formation. Chemical denaturant and heat were used as proxies for any stress that the antibody may experience in its lifetime, which could also be shear stress or pressure during the expression and purification steps.

A series of biophysical studies were also undertaken to assess the impact of the inserted cysteine, and more specifically the iDSB thiol state on the antibody's structure and stability. Chemical denaturation curves and thermal unfolding showed, first, that the C_H2 domain is less stable than the C_H3 domain, and second, that for all Fc-C239i enriched variants, the C_H2 domain was destabilized by the insertion of the cysteine at position 239

while the stability of the C_H3 domain wasn't affected. When the additional disulfide bridge was present, it was found to apply an additional destabilizing strain on the C_H2 domain. This was illustrated by the significant increase in hydrogen/deuterium exchange especially in β-strands C, F, G and in β-strand D due to H-bonding breakage and an increase in solvent accessible surface area in the native state. This was confirmed by the lower m_{I-N}, ΔCp and ΔH_{cal} values obtained for the iDSB compared to 2xSH and 2xCys Fc-C239i in thermal unfolding and chemical denaturation studies all indicators that the native state had been disrupted and was less stable.

Interestingly, although a decrease in thermodynamic stability would generally lead to a decrease in kinetic stability, unfolding kinetic experiments showed that the iDSB Fc-C239i has greater kinetic stability than the other variants. This unexpected result is likely to be due to the fact that the formation of an additional interchain disulfide bridge not only changes the structure and stability of the native state but also the unfolding transition state. All these findings are consistent with a model in which there is destabilisation brought about by a distortion in the structure of C_H2 domain as a direct consequence of the formation of an additional disulfide bridge in the native state.

Additional results from collaborators showed that the formation and reformation (after the reduction step) of the iDSB reduces the efficiency of the final conjugation step, while still enabling efficient and consistent conjugation with DAR in a similar range of values as those observed for other constructs.

Taken together, these data demonstrate the utility of combining biophysical techniques with high-resolution mass spectrometry to provide detailed characterization of the structure and dynamics of biopharmaceuticals and their variants. Similar observations are anticipated for other engineered antibodies where cysteines are introduced close to the hinge, and the insight provided here serves to guide future cysteine or site-specific engineering strategies.

5 STABILITY OF DVP- AND TETRA DVP-CONJUGATED ANTIBODIES

5.1 INTRODUCTION

Antibody-drug conjugates (ADCs) have become very promising therapeutics in oncology by combining the high specificity of a tumor-recognizing monoclonal antibody (mAb) with the potency of a chemotherapeutic small molecule (payload)[155]. Whilst the first generation of ADCs with lysine conjugation led to variable drug-to-antibody ratios due a distribution in number and position of drugs attached, the field has progressed towards more site-specific conjugation strategies using cysteines, either canonical or substituted/inserted ones. In these cases, the most widely used linker contains a maleimide which reacts with the cysteine side chain reacts. However, the thiosuccinimide adduct formed can deconjugate *via* a retro-Michael reaction[45], in which case the cytotoxic payload is released from the antibody, and *in vivo* is lost in circulation.

To overcome this problem, the Spring Group has developed two linkers which irreversibly react with cysteines, thus avoiding the uncontrolled release of cytotoxic payload. The first, developed by Walsh *et al.*, 2019[47], is a divinylpyrimidine (DVP) linker which can be conjugated to one[47] or two[48] payloads and which reacts with two cysteines (**Figure 5.1**). In the case of conjugation to canonical cysteines, the interchain disulfide bridges in the antibody are first reduced and then four DVP linkers react with the resultant cysteine side chains, linking the cysteines in the Fab of the constant domain of the light chain (C_L) with the constant domain 1 of the heavy chain (C_H1) and the cysteines located in the hinge. The product of the DVP reaction leads to a heterogeneous mixture of interchain (between two heavy chains) or intrachain linkages (on the same heavy chain); the biophysical experiments described in this Chapter were carried out on the mixture. The details of the chemical reactions are shown in **Figure 5.2**. The second linker, developed by Friederike Dannheim and collaborators [49,50], is a tetradivinylpyrimidine (tetraDVP) linker, where four moieties within the tetraDVP linker can be conjugated with up to four payloads depending on the spacer (**Figure 5.1**). A single molecule of tetraDVP linker reacts with all eight canonical cysteines after reduction of the interchain disulfide bridges. Which of the cysteine-reactive groups in the linker reacts with which cysteine residue in the antibody is random, therefore there is likely to be a mixture of products with subtly different connectivity. However, there is no half-antibody formation, i.e. the covalent connection between the polypeptide chains is fully restored.

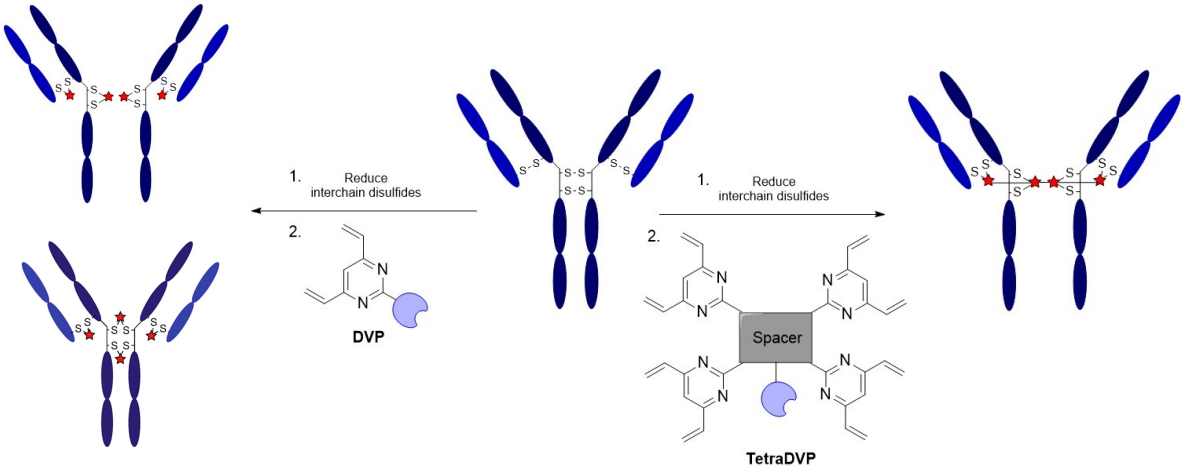


Figure 5.1: Modification of a therapeutic antibody with DVP and tetraDVP linkers. On the left of the figure shows the reaction of trastuzumab (centre) with four molecules of DVP to produce a modified antibody with four DVP molecules attached, forming intrachain linkages (top left) or interchain linkages (bottom left). On the right, the reaction of trastuzumab (centre) with a single molecule of tetraDVP is shown to produce a tetraDVP-linked antibody (right). Figure adapted with permission from Walsh *et al.* [47] and Dannheim *et al*. [49,50]. The figure illustrates how the linkers react with the canonical cysteines which in the wild-type antibody form interchain disulfide bridges in the hinge and between the heavy and light chains. The red stars represent the payloads that can be attached to the linkers.

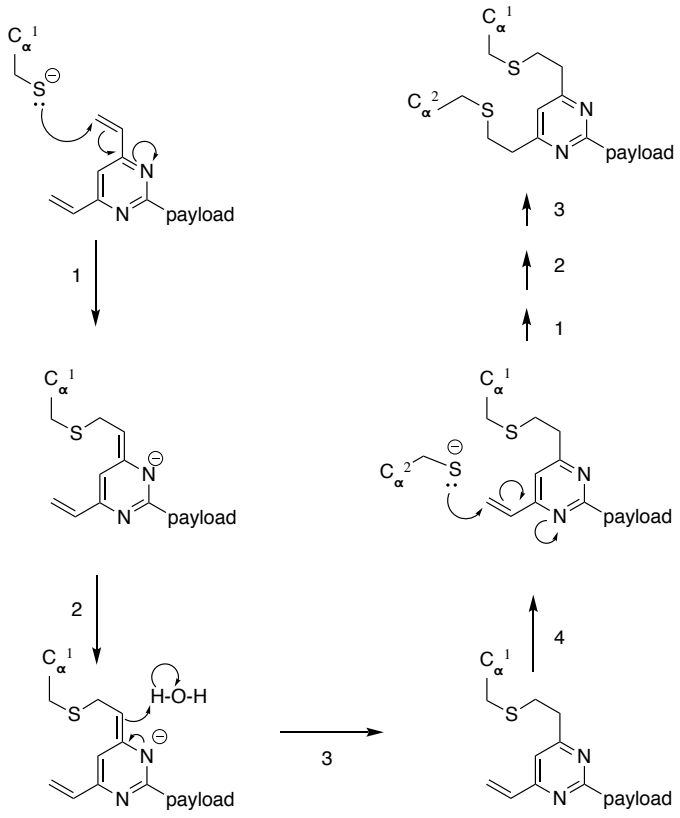


Figure 5.2: Schematic of the reaction of the divinylpyrimidine with the thiols, carried out under slightly basic conditions (pH 8).

The clear advantages of these linkers in terms of stability of the linker and the

biological function of the antibodies they are conjugated to has been demonstrated[47,49,50] and this novel approach therefore represents a very promising alternative to the maleimide-conjugated linkers currently employed. In this Chapter, the stability of the DVP- or tetraDVP- conjugated antibody was investigated to understand the impact that the conjugation to these linkers has on the thermodynamic and kinetic stability and dynamics of the native state of the modified antibody.

5.2 RESULTS

5.2.1 Background

The therapeutic antibody trastuzumab was modified by Friederike Dannheim with the DVP and tetraDVP linkers, using published protocols[47], and unpublished protocols developed in her PhD[50] research by Friederike Dannheim, respectively (**Figure 5.1**). The unmodified as well as the conjugated antibodies were a kind gift from the Spring group. The experiments on the DVP-linked trastuzumab were carried out on the heterogeneous mixture previously mentioned, composed of more intrachain (on the same heavy chain) than interchain (between two heavy chains) linkages, whereas the experiments on the tetraDVP-linked trastuzumab were done on all reformed full-length IgG (**Figure 5.1**).

To understand the impact of the conjugation of the DVP and tetraDVP linkers on the stability of the antibody, four different measures of the stability were carried out for both the unmodified and conjugated antibodies.

The first was the thermodynamic stability, which provides information on the difference in Gibbs free energy between the native and denatured state in water, usually at 25 °C (ΔG_{D-N}), which is measured by chemical denaturation. It is related to the difference in enthalpy (ΔH_{D-N}) and entropy (ΔS_{D-N}) between the native and denatured states, although these values are not measured directly in this experiment. The second is thermal stability which is also influenced by the change in enthalpy and entropy on unfolding but also the difference in heat capacity (ΔC_p). The third is the kinetic stability, which measured the activation energy (ΔG_{TS-N}), i.e., the difference in energy between the native and transition states. The fourth approach is less a measure of stability and more of the dynamics in the native state, and was measured by HDX-MS. Under these conditions, the exchange rates are limited by the global unfolding rates.

The experiments were carried out on full-length IgG unmodified trastuzumab, full-length trastuzumab linked by DVP and tetraDVP, as well as on the Fab (two arms and hinge) and Fc domains (without hinge), generated by digestion of the full-length antibody with the IdeS enzyme (FabRICATOR, Genovis), which cuts below the hinge (CPAPELLG / GPSVF).

5.2.2 Thermodynamic stability

The thermodynamic stability was probed by chemical denaturation curves monitored by using the fluorescence of tryptophan. The thermodynamic stability of each domain was determined for the unmodified, tetraDVP- and DVP-linked trastuzumab. The change in tryptophan fluorescence was measured after five days of incubation in various concentrations of GdmCl and the data analyzed using the average emission wavelength (AEW, **Equation 2.3**) and fitted to a two- or three-state model (**Equations 2.5 and 2.7**). Three thermodynamic parameters were obtained: $\Delta G_{X-Y}^{H_2O}$, the difference in Gibbs free energy between two states X and Y in water; m_{X-Y}, the m-value between the two states X and Y, a constant of proportionality describing how much ΔG_{X-Y} changes with denaturant concentration; and $[den]_{50\%\,X-Y}$, the midpoint of denaturation between states X and Y.

The stability measurements on the unmodified antibody showed a single cooperative unfolding transition for the Fab domain, and two unfolding transitions for the Fc domain corresponding to unfolding of the C_H2 and C_H3 domains[153] (**Figure 5.3, Table 5.1**). However, the chemical denaturation curve of the full-length IgG showed less than three transitions, because of overlap between unfolding transitions (**Figure 5.3, Table 5.1**).

Table 5.1: Thermodynamic parameters fitted from the unfolding curves

	$m_{\text{I-N}}$ (kcal mol^{-1} M^{-1})	$[\text{den}]_{50\%\,\text{I-N}}$ (M)	$m_{\text{D-I}}$ (kcal mol^{-1} M^{-1})	$[\text{den}]_{50\%\,\text{D-I}}$ (M)
T IgG	2.5 ± 0.5	2.5 ± 0.2	8.6 ± 1.6	2.73 ± 0.03
T+tetraDVP IgG	2.3 ± 0.6	2.0 ± 0.2	4.6 ± 0.4	2.49 ± 0.03
T+DVP IgG	2.9 ± 0.3	2.03 ± 0.05	5.9 ± 0.8	2.57 ± 0.03
T Fab			6.9 ± 0.5	2.66 ± 0.01
T+tetraDVP Fab	1.5 ± 0.6	2.1 ± 0.5	5.8 ± 0.8	2.56 ± 0.02
T Fc	3.5 ± 0.1	1.93 ± 0.01	5 ± 1	2.80 ± 0.03
T+tetraDVP Fc	3.8 ± 0.3	1.84 ± 0.01	13 ± 26*	2.69 ± 0.09

The errors reported in this table are fitting errors.*As the two transitions of T tetraDVP Fc are very close, the software doesn't distinguish them easily, leading to a fitted value for $m_{\text{D-I}}$ which is not accurate.

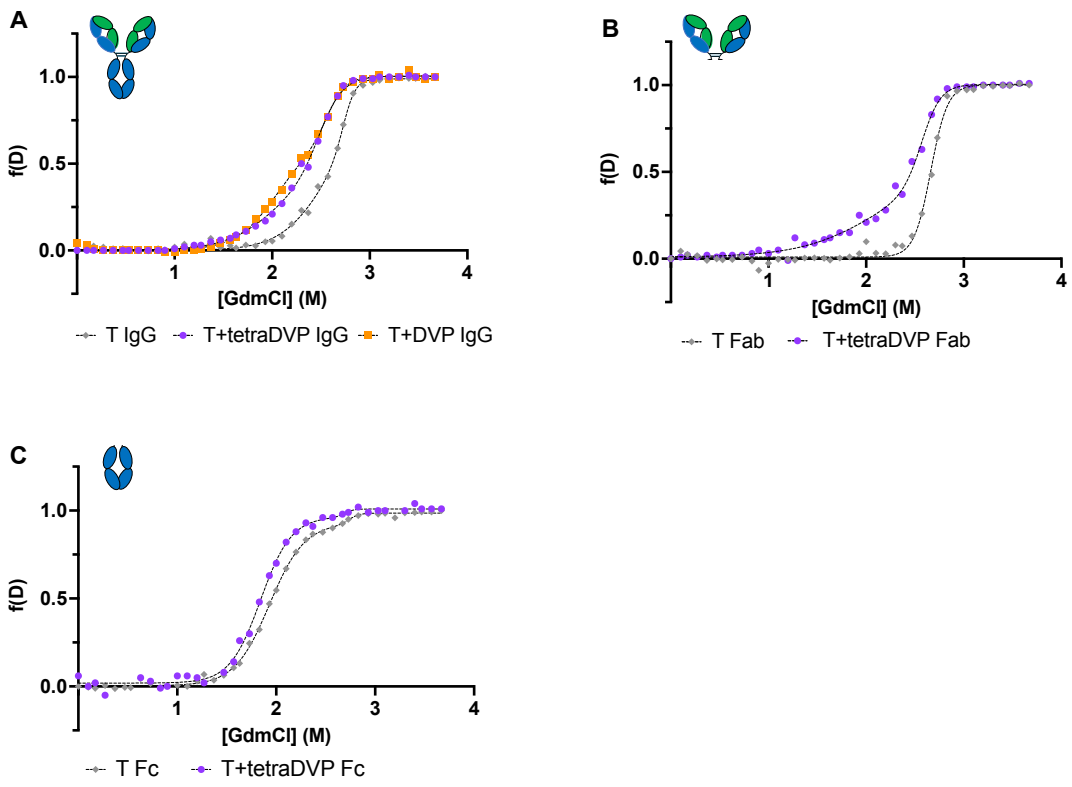


Figure 5.3: Chemical denaturation curves of unmodified and modified trastuzumab. Trastuzumab (grey lozenges), tetraDVP-linked trastuzumab (purple circles), DVP-linked trastuzumab (orange squares). A. full-length IgG. **B.** Fab domain with hinge (two arms). **C.** Fc domain without hinge. The samples were denatured in guanidinium chloride, 50 mM Tris pH 7.5, at 25 °C until equilibrium was reached (at least 5 days).

The chemical denaturation curves of the full-length unmodified trastuzumab and tetraDVP- and DVP-linked trastuzumab clearly show that the additional multivalent linker has destabilized the protein, the modified variants unfolding at lower GdmCl concentrations than unmodified trastuzumab, **Figure 5.3 A**. This is reflected in, and can be attributed to, the destabilization of the native state of the Fab domain on conjugation, **Figure 5.3 B**. In this case, an additional unfolding transition appeared at a lower concentration of GdmCl whilst the other unfolding transition at higher GdmCl concentrations remains unchanged as shown by the $[den]_{50\%}$ and the m-value which are within error of the values calculated for the single transition observed with the non-modified Fab (**Table 5.1**). Given that the tetraDVP is covalently linked to the four canonical cysteines in the hinge and also to the two cysteines in each Fab arm originally forming the intrachain disulfide bridge between the heavy and light chains (**Figure 5.1**), it is unlikely that the linker has affected the stability of the variable domains which are distant from the conjugation sites. Thus, these data suggest that the tetraDVP linker has destabilised the constant domains within the Fab, i.e. the C_L and C_H1 domains. Therefore,

the unfolding transition observed at lower GdmCl concentrations for the modified Fab corresponds to unfolding of the C_L and C_H1 domains, while the unfolding of the variable domains, V_L and V_H, which are the same in both modified and unmodified trastuzumab correspond to the transition observed at higher GdmCl concentrations (**Figure 5.3 B, Table 5.1**).

The unfolding of the Fc domains of the unmodified and tetraDVP-linked trastuzumab are very similar, as expected as they are identical molecules as they don't contain any of the modified cysteine sites (**Figure 5.3 C**). Therefore, the unfolding of the Fc domains should be exactly the same, however, a small difference is observed, which is almost certainly due to the fact that the protein concentrations in the two experiments were different (T Fc: 12 μM, T+tetraDVP Fc: 3 μM) and the Fc fragment without the hinge is a dimer so concentration will affect the $[\mathrm{den}]_{50\%}$. One would expect a lower midpoint for the T+tetraDVP Fc as it was measured at a lower concentration[174] and this was indeed observed. One could easily prove this by measuring the stability of the Fc for both modified and unmodified at the same protein concentration.

5.2.3 Kinetic stability

Unfolding kinetics provide information on the rate at which a native state of a domain unfolds. Native, folded protein was rapidly diluted into a range of chemical denaturant concentrations and the protein unfolding was monitored by the change in intrinsic fluorescence on a stopped-flow spectrophotometer. The data for all Fab domains (unmodified and tetraDVP-linked trastuzumab) were best fit to a single-exponential function (**Equation 2.9**), and the data for the unmodified trastuzumab Fc domains was fit with a double-exponential function (**Equation 2.10**) which corresponds to a faster and a slower unfolding phase (**Figure 5.4**).

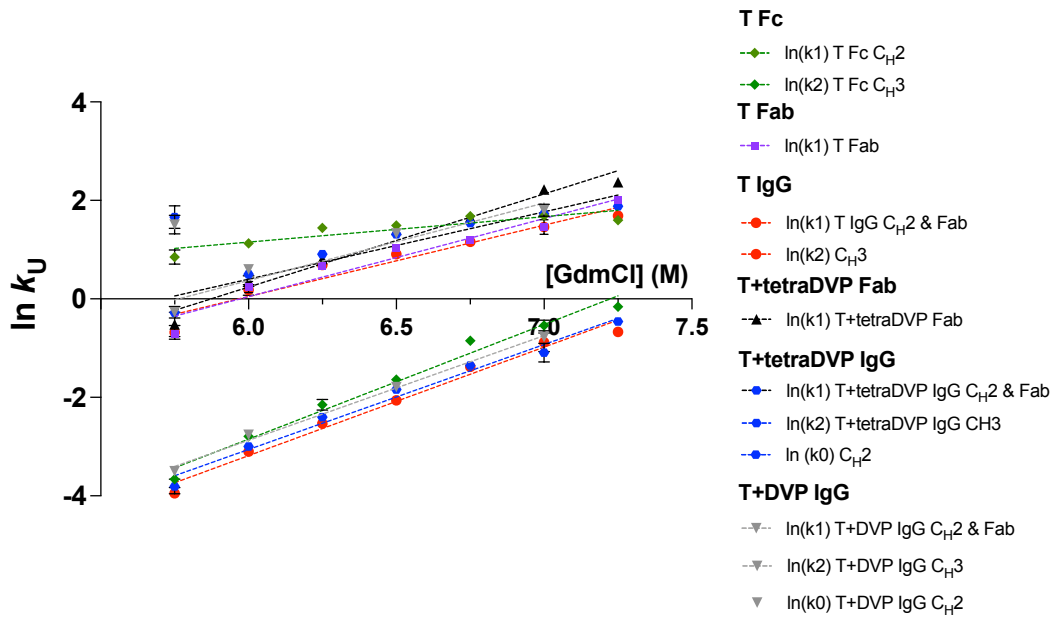


Figure 5.4: Unfolding kinetics of unmodified and modified forms of trastuzumab. Trastuzumab Fc (green circles), trastuzumab Fab (purple squares), trastuzumab IgG (red circles), tetraDVP-linked trastuzumab Fab (black triangles), tetraDVP-linked trastuzumab IgG (blue circles), DVP-linked trastuzumab IgG (grey triangles). The dotted lines show the best fit of the data to **Equation 2.12**. Error bars represent the standard deviation from triplicate measurements. In many cases, they are not visible because they are smaller than the size of the data points.

Table 5.2: Unfolding kinetic parameters

		m_{kU} (M^{-1})	ln k_U (H_2O)
T Fc	ln($k1$) T Fc C_H2	0.5 ± 0.1	-1.9 ± 0.8
	ln($k2$) T Fc C_H3	2.3 ± 0.1	-16.8 ± 0.9
T Fab	ln($k1$) T Fab	1.6 ± 0.2	-9.5 ± 1.2
T IgG	ln($k1$) T IgG C_H2 & Fab	1.4 ± 0.2	-8.6 ± 1.2
	ln($k2$) T IgG C_H3	2.2 ± 0.1	-16.4 ± 0.9
T+tetraDVP Fab	ln($k1$) T+tetraDVP Fab	1.9 ± 0.2	-11.1 ± 1.0
T+tetraDVP IgG	ln($k1$) T+tetraDVP IgG C_H2 & Fab	1.4 ± 0.2	-7.8 ± 1.1
	ln($k2$) T+tetraDVP IgG C_H3	2.1 ± 0.1	-15.8 ± 0.8
T+DVP IgG	ln($k1$) T+DVP IgG C_H2 & Fab	1.6 ± 0.3	-9.1 ± 1.9
	ln(k_2) T+DVP IgG C_H3	2.1 ± 0.1	-15.6 ± 0.7

The data points from **Figure 5.4** were fitted to **Equation 2.12**. The errors are the standard deviation of triplicate measurements.

It has been previously reported that the fastest unfolding phase corresponds to the unfolding of the C_H2 domain and the slower phase comes from unfolding of the C_H3 domain[153]. Interestingly, the unmodified, tetraDVP- and DVP-linked full-length IgGs showed just two unfolding phases at most of the GdmCl concentrations used, most

probably because two of the unfolding rate constants are similar. The kinetic parameters obtained from the fit of the data to **Equation 2.12** showed that, on the one hand, the unfolding rate constants of the Fab domain and the fastest of the two phases observed for the full-length IgGs, and, on the other hand, the C_H3 domain and the slowest of the two phases of the full-length IgGs, are within error (**Table 5.2**). This suggests that the fastest of the two unfolding transitions observed with the full-length IgGs represents the unfolding of both the C_H2 and Fab domains. Further evidence for this hypothesis is provided by the percentage amplitudes of the two phases observed (**Table 5.3, Table 5.4, Table 5.5**). For all full-length IgGs, the percentage amplitude of the fast-unfolding phase (% a1) is larger than that of the slow unfolding phase (% a2), whereas the opposite is the case for the Fc domains. This observation confirms that the fastest unfolding phase of the full-length IgGs contains the signal corresponding to the unfolding of both the C_H2 and Fab domains.

Interestingly, at the lowest GdmCl concentration (5.75 M), the unfolding data for the tetraDVP- and DVP-linked full-length IgGs were best fit to a triple-exponential function (**Equation 2.11**). The fact that this additional transition is observed for the tetraDVP- and DVP-linked full-length IgGs, but not the unmodified form, suggests that it could either be that the C_H2 domain was destabilized by both modifications in the hinge, or that the C_H1 and C_L domains are destabilised as showed in the thermodynamics. This third, rapid phase is not observed at higher GdmCl concentrations as the difference in kinetic stability between the C_H2 and the Fab domains becomes too small to be de convoluted.

Table 5.3: Amplitudes associated with the various unfolding phases of full-length IgG and the Fc domain of unmodified trastuzumab

[GdmCl] (M)	a1 (Fab and C_H2)	a2 (C_H3)	% a1	% a2
5.75	-1.4 ± 0.3	-0.7 ± 0.2	67	33
6	-1.3 ± 0.1	-0.8 ± 0.1	62	38
6.25	-1.3 ± 0.1	-0.8 ± 0.1	61	39
6.5	-1.4 ± 0.2	-0.7 ± 0.3	67	33
6.75	-1.4 ± 0.1	-0.9 ± 0.1	60	40
7	-1.4 ± 0.1	-0.9 ± 0.1	63	37
7.25	-1.3 ± 0.2	-1.0 ± 0.1	57	43

[GdmCl] (M)	a1 (C_H2)	a2 (C_H3)	% a1	%a2
5.75	-1.1 ± 0.1	-1.6 ± 0.1	41	59
6	-1.4 ± 0.1	-1.9 ± 0.1	42	58
6.25	-0.9 ± 0.1	-1.50 ± 0.05	38	62
6.5	-1.1 ± 0.1	-1.64 ± 0.02	40	60
6.75	-1.1 ± 0.1	-1.6 ± 0.1	41	59
7	-1.18 ± 0.04	-1.8 ± 0.1	40	60
7.25	-1.22 ± 0.04	-1.7 ± 0.1	41	59

Percentage amplitudes of the fastest phase (a1) and the slowest phase (a2) for full-length IgG (left) and Fc domain (right) of unmodified trastuzumab. The errors reported in this table are the standard deviations of triplicate measurements.

Table 5.4: Percentage amplitudes of the unfolding phases of full-length IgG and Fc domain of tetraDVP-linked trastuzumab.

[GdmCl] (M)	a1 (Fab and C_H2)	a2 (C_H3)	% a1	% a2
5.75	-1.0 ± 0.1	-0.7 ± 0.1	59	41
6	-1.3 ± 0.1	-1.0 ± 0.1	57	43
6.25	-1.36 ± 0.04	-1.2 ± 0.1	54	46
6.5	-1.3 ± 0.1	-1.0 ± 0.2	57	43
6.75	-1.3 ± 0.1	-1.3 ± 0.2	51	49
7	-1.3 ± 0.1	-1.3 ± 0.2	51	49
7.25	-1.1 ± 0.1	-1.28 ± 0.04	46	54

[GdmCl] (M)	a1 (C_H2)	a2 (C_H3)	% a1	% a2
5.75	-0.7 ± 0.2	-1.1 ± 0.4	40	61

Fastest phase (a1) and the slowest phase (a2) for full-length IgG (left) and Fc domain (right) of tetraDVP-linked trastuzumab. The errors reported in this table are the standard deviations of triplicate measurements.

Table 5.5: Percentage amplitudes of the unfolding phases of full-length IgG of DVP-linked trastuzumab.

[GdmCl] (M)	a1 (Fab and C_H2)	a2 (C_H3)	% a1	% a2
5.75	-0.48 ± 0.08	-0.34 ± 0.07	59	41
6	-0.84 ± 0.02	-0.68 ± 0.06	55	45
6.5	-1.1 ± 0.1	-0.7 ± 0.1	60	40
7	-1.1 ± 0.1	-0.8 ± 0.2	59	41

Fastest phase (a1) and the slowest phase (a2) for full-length IgG of DVP-linked trastuzumab. The errors reported in this table are the standard deviations of triplicate measurements.

Overall, the unfolding kinetics data show that the modifications do not cause any significant change in kinetic stability, despite a clear decrease in the thermodynamic stability of the Fab as seen in the equilibrium unfolding experiments (for the Fab domain: $\Delta G_{D-N}^{T} > \Delta G_{D-N}^{DVP/tetraDVP\ T}$, **Figure 5.5**). These results suggest the modifications have destabilized the transition state for unfolding of the Fab domain to the same degree as the native state, hence $\Delta G_{TS-N}^{T} = \Delta G_{TS-N}^{DVP/tetraDVP\ T}$ (**Figure 5.5**).

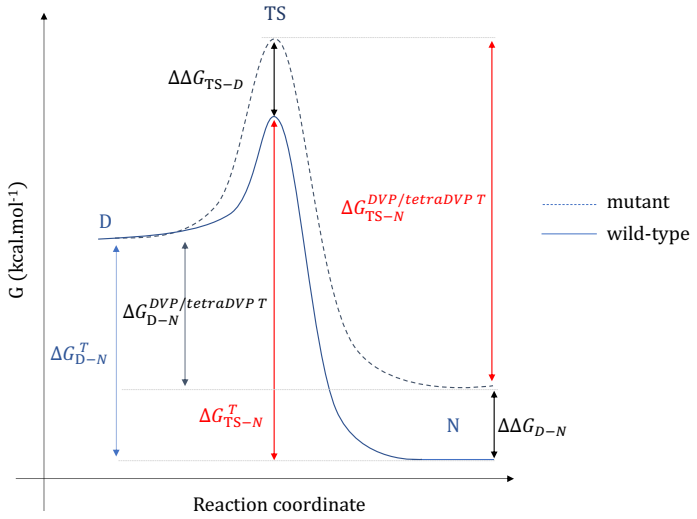


Figure 5.5: Energy landscape of Fab domain of unmodified and DVP- and tetraDVP-linked trastuzumab.

5.2.4 Thermal stability

The thermal stability was measured by differential scanning calorimetry (DSC). The experiments were carried out on full-length IgG, as well as on Fab and Fc domains of the unmodified and DVP- and tetraDVP linked trastuzumab. The results from the full-length IgGs showed two peaks, the first corresponding to the unfolding of the C_H2 domain, and the second to the unfolding of the C_H3 and Fab domains which overlap[149,150,175]. The melting temperatures of the first unfolding peak show that the C_H2 domain was destabilized by the tetraDVP-linker, and to a greater extent the DVP linker (**Figure 5.6 A**, **Table 5.6**). The melting temperatures of the second peak for the modified trastuzumab, corresponding to the unfolding of the Fab and C_H3 domains, were very similar to that of the unmodified, however, a small shoulder has appeared at a lower temperature. The analysis of the DSC curves of the Fab domains confirms this observation (**Figure 5.6 B).** It suggests that the constant domain of the light chain may have been destabilized by the conjugation of the DVP and tetraDVP linkers. Furthermore, the decrease in ΔH_{cal}, i.e., the area under the unfolding peak which is a measure of the favorable interactions that must be overcome to unfold, suggests that some favorable interactions were lost in the native state of the DVP- and tetraDVP-linked trastuzumab. (**Table 5.6,** $\Delta H2_{cal}$). The peaks corresponding to unfolding of the C_H2 unfolding were broader for the conjugated trastuzumab compared to the unmodified form, which may be due to a smaller ΔCp and therefore smaller difference in solvent accessible surface area ($\Delta SASA$)[168]. The results from the Fc domain, which were generated by digestion under the hinge and therefore don't have the strain of the linkers anymore, show very similar

curves for three constructs as expected, as they are the same.

Overall, the results from the DSC experiments suggest that conjugation of trastuzumab to the linkers has destabilized the C_H2 domain, and also potentially the constant domains in the Fab fragment.

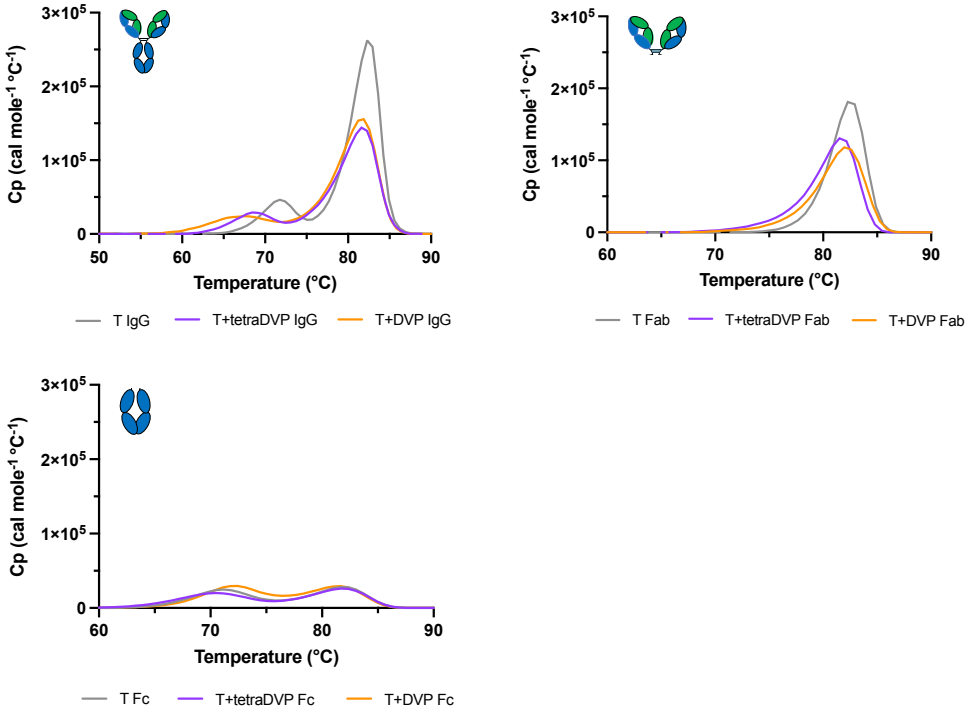


Figure 5.6: Differential scanning calorimetry results of unmodified and DVP- and tetraDVP-linked trastuzumab A. full-length IgG B. Fab domain C. Fc domain. The experiments were carried out in 50 mM Tris pH 7.5.

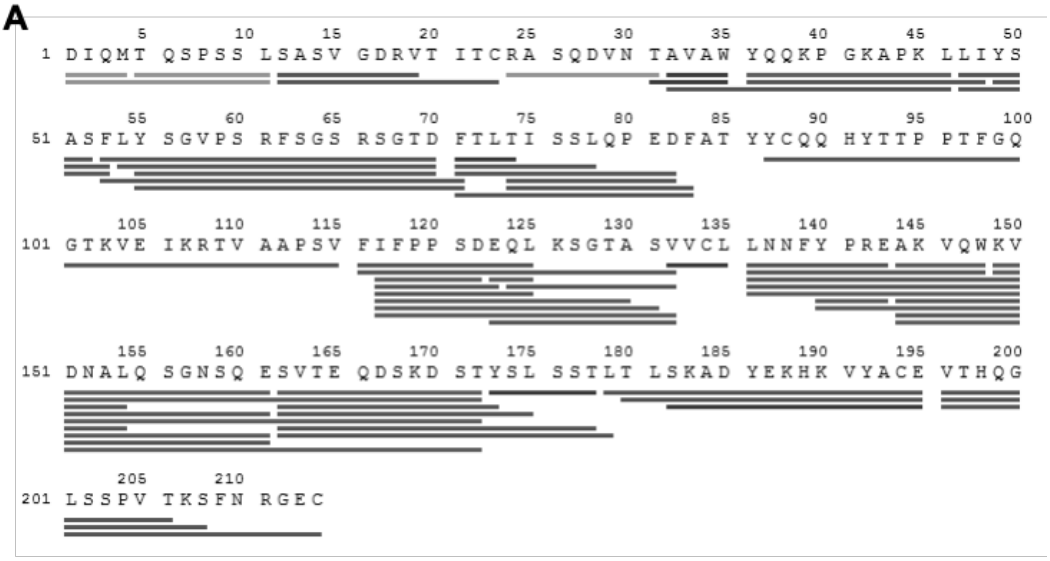
Table 5.6: Melting temperatures of the antibody variants from the DSC experiments

Antibody variant	$\boldsymbol{T_m1}$ (°C)	$\Delta \boldsymbol{H1_{cal}}$ (kcal mol $^{-1}$)	$\boldsymbol{T_m2}$ (°C)	$\Delta \boldsymbol{H2_{cal}}$ (kcal mol^{-1})
T IgG	71.8*	(3.2 ± 0.3) x 10^2	82.05 ± 0.03	(11.9 ± 0.2) x 10^2
T DVP IgG	67.4*	(3.2 ± 0.3) x 10^2	81.10 ± 0.05	(9.3 ± 0.2) x 10^2
T+tetraDVP IgG	69.0*	(2.7 ± 0.2) x 10^2	81.19 ± 0.05	(8.6 ± 0.2) x 10^2
T Fab	-	-	82.25 ± 0.03	(7.6 ± 0.1) x 10^2
T+tetraDVP Fab	-	-	81.14 ± 0.05	(6.8 ± 0.1) x 10^2
T DVP Fab	-	-	81.68 ± 0.04	(6.4 ± 0.1) x 10^2
T Fc	71.20 ± 0.05	(1.8 ± 0.3) x 10^2	81.55 ± 0.04	(1.7 ± 0.3) x 10^2
T+tetraDVP Fc	70.52 ± 0.06	(1.82 ± 0.03) x 10^2	81.52 ± 0.04	(1.62 ± 0.03) x 10^2
T+DVP Fc	72.24 ± 0.05	(2.29 ± 0.04) x 10^2	81.22 ± 0.04	(1.76 ± 0.03) x 10^2

The errors are the fitting errors. *These melting temperatures don't have an error as they were fixed based on the raw data to facilitate the fitting.

5.2.5 Molecular Dynamics in the native state determined by HDX-MS experiments

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) is a technique which probes the molecular dynamics of the native state of proteins, in particular the exposure of different parts of a protein to solvent. In this experiment, the hydrogen/deuterium exchange of the DVP and tetraDVP-linked trastuzumab was compared to the refence unmodified trastuzumab. To obtain high-resolution information of the deuterium exchange per residue, a peptide map was acquired using online pepsin digestion on the unmodified trastuzumab. High coverage (light chain: 98.6%, heavy chain 90.9%) was obtained (**Figure 5.7, Table 5.7**), however, a large region of the C_H1 domain (from residues 201 to 237) was missing from the peptide map. All the HDX-MS labelling experiments were done on full-length IgG constructs and the deuterium uptake of the control unmodified trastuzumab was subtracted from the data obtained for the different conjugated antibodies.



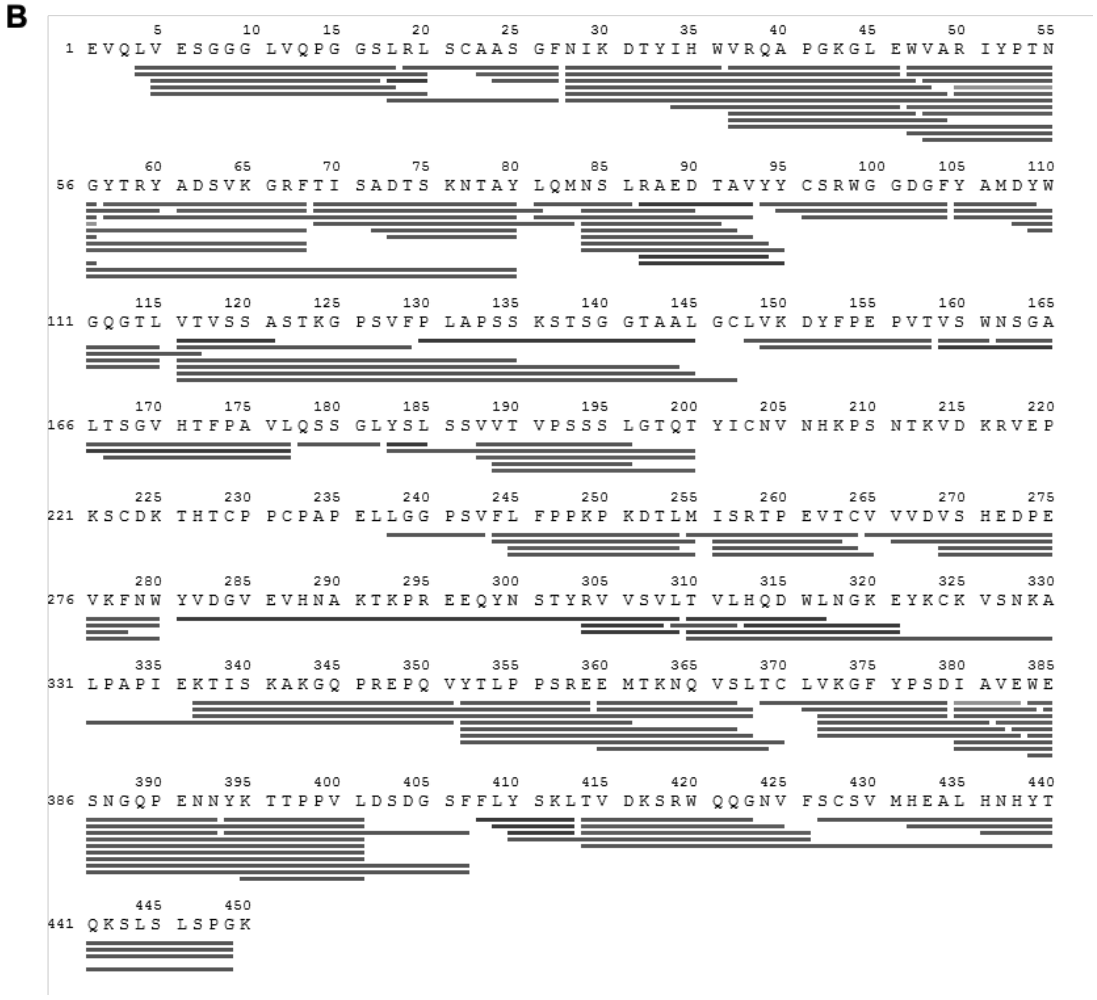


Figure 5.7: Peptide map of the light chain (A) and heavy chain (B) of trastuzumab

The analyzed data showed mainly significant increases in deuterium exchange in various regions and a small significant decrease in exchange in the constant domain of the light chain (**Figure 5.8**). These differences are represented on the crystal structures (**Figure 5.9 B**) and the details of the uptake plots were also reported (**Figure 5.9 C**). Although the absolute maximum increase in deuterium exchange was low, a scale enhancing the differences was used for more clarity of the results. TetraDVP- and DVP-linked trastuzumab both showed some increase in exchange in a β-strand of the V_H and in the CDR2 of the light chain in the V_L. A very low level of deprotection was observed in the C_L, and, as the C_H1 was missing 37 residues in the peptide map, it is not possible to conclude anything on effect of the linkers on the exchange of this region; this region was colored in yellow on **Figure 5.9 B**. For the DVP-linked trastuzumab, the largest difference in exchange took place in the C_H2 domain, in the region of the C_H2 domain closest to the C_H3 domain, highlighting a zone of destabilization. The tetraDVP-linked

trastuzumab was destabilized to a lesser extent on a β-strand at the heart of the C_H2 domain, suggesting that the two linkers impose different strains on the C_H2 domain, and that the DVP linker is more destabilizing than the tetraDVP linker. Very low levels of deprotection were observed in some regions of the C_H3 domain for the DVP-linked trastuzumab.

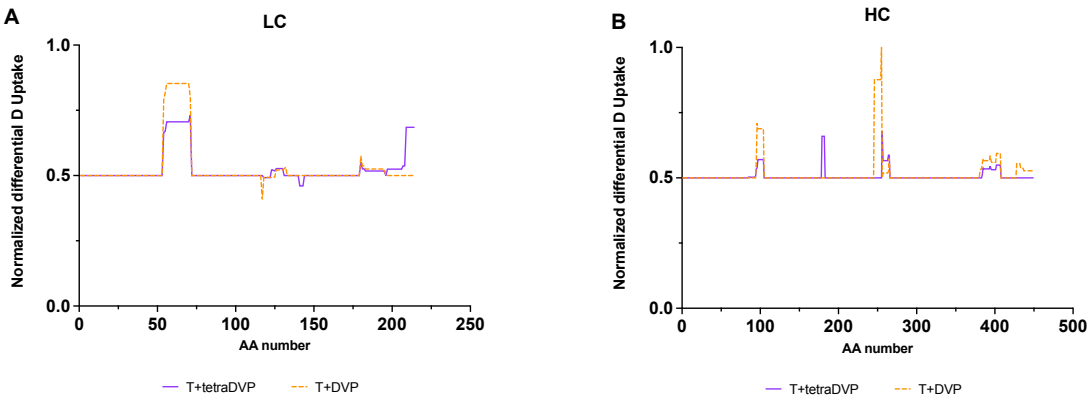


Figure 5.8: Difference plots showing the regions of trastuzumab that change dynamics after linkage with DVP and tetraDVP. Normalized difference plots of A. the light chain and **B.** the heavy chain, obtained by summing the significant D-incorporations relative to the reference, and subtracting the reference exchange to that of the T+tetraDVP and T+DVP. The light and heavy chains were normalized with the same maxima.

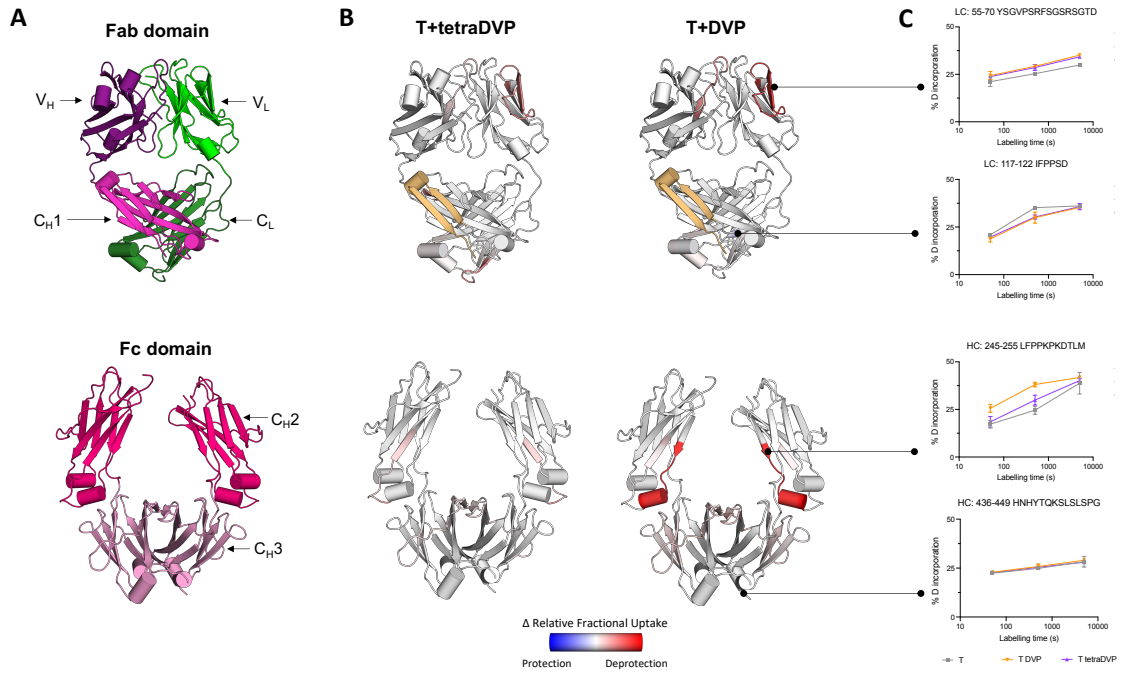


Figure 5.9: HDX-MS results. A. Representation of the different domains in the Fab and Fc domains on the crystal structure of the Fab (1N8Z) and Fc (3AVE). **B.** Crystal structures colored according to the relative change in fractional deuterium exchange for T+tetraDVP and T+DVP compared to T; the differences plotted are significant differences assessed by a t-test, with p-value < 0.01. The yellow region represents the residues that were missing in the peptide map. **C.** Details of the uptake plots for selected regions; the error bars represent a Student's t distribution with 95% confidence interval based on triplicates. Reduced exchange is shown in blue, no change in white, and increased exchange in red.

Overall, the HDX-MS data showed that the C_H2 domain was destabilized by the DVP linker more than the tetraDVP linker, compared to the unmodified trastuzumab, most probably due to a distortion of that domain because of the strain applied by the conjugation. Given that the C-terminal region of the C_H1 domain unfortunately wasn't covered in the peptide map, it was not possible to probe any difference in deuterium exchange in that region, although there might be some, as suggested by the thermodynamic and thermal stability data.

Table 5.7: Summary of HDX mass spectrometry experimental details

Dataset	Unmodified trastuzumab (10 µM) –trastuzumab expressed at AZ DVP-linked trastuzumab (10 µM) –commercial trastuzumab tetraDVP-linked trastuzumab (10 µM) – commercial trastuzumab
HDX reaction details	Equilibration in H_2O: 10 mM potassium phosphate pH 7.5 Labelling in D_2O: 10 mM potassium phosphate pD 7.5 Both were done at 20 °C Quench buffer in H_2O: 100 mM potassium phosphate, 0.5 M TCEP, 8 M urea pH 2.5
HDX time course	50 s, 500 s, 5000 s
HDX controls	0 s
Number of peptides	LC: 63 HC: 140
Sequence coverage	LC: 98.6 % HC: 90.9 %
Average peptide length	LC: 12.8 AA HC: 13.4 AA
Replicates	Technical replicates: 3
Repeatability	Average Student's t-distribution 95% confidence interval for all peptides, all time points and states (#D error): LC: 0.1389 HC: 0.1432
Significant differences in HDX	The incorporation of deuterium per peptide is significant compared to the wild-type if p-value < 0.01 (t-test)

5.3 DISCUSSION AND CONCLUSION

In this study, the impact of the conjugation of the DVP and tetraDVP linkers on the stability of a monoclonal antibody, trastuzumab, was investigated. The results from the different biophysical techniques provided different complementary information on different types of stability of the conjugated antibody. The thermal unfolding and HDX-MS data on full-length IgG confirmed that the linkers destabilize the C_H2 domain, the DVP linker to a greater extent than the tetraDVP linker. The thermodynamic and thermal stability results suggest that the linkers also destabilize the Fab domain, most probably the constant domains C_L and C_H1 by applying a strain to them. However, incorporation of the linkers did not destabilize any domain kinetically suggesting that the linkers have a similar destabilizing effect on the transition state as well as the native state. These most probably are due to the strain that the linkers are apply onto the tertiary/quaternary structure of the antibody, forcing a slight change in conformation compared to the unmodified format as shown by the HDX-MS experiments

Despite some loss in intrinsic stability compared to the unmodified antibody, the stability of both DVP- and tetraDVP-linked antibodies are still high, as the melting temperature of the least stable domain is 67.4 °C, considerably higher than body temperature. Destabilization of the scaffold of antibody drug conjugates has been previously observed. For example, the C_H2 domain of trastuzumab conjugated via lysine residues to the drug DM1, had a lower melting temperature compared to the non-conjugated form[176]; the DSC traces obtained in this study look very similar to those reported here with a lower T_m and broader peak. Other kinds of scaffolds with inserted cysteines like C239i also lose some stability in the C_H2 domain prior to the conjugation[35,107]. Overall, the DVP- and tetraDVP-modified antibodies are very promising scaffolds for the development of future secure antibody-drug conjugates.

6 MECHANISTIC INSIGHTS INTO THE RATIONAL DESIGN OF MASKED ANTIBODIES WITH ANTI-IDIOTYPIC ANTIBODY FRAGMENTS

The results presented in this Chapter have been submitted as a paper to *mAbs:*

Orozco C. T., Bersellini M., Irving L. M., Howard W.W., Hargreaves D., Devine P. W. A., Siouve E., Browne G. J., Bond N. J., Phillips J. J., Ravn P., Jackson S. E. Mechanistic insights into the rational design of masked antibodies

6.1 INTRODUCTION

Monoclonal antibodies have become a major class of therapeutics over the past twenty-five years and have proven to be promising as alternatives to, and in combination with, conventional therapies for some cancers[1–3]. Despite their clear potential, they can cause side effects, mainly due to on-target, off-tumour activity[4]. In the past decade, strategies have emerged to improve the selectivity of therapeutic antibodies at the tissue level, by masking the paratope of an antibody, thus injecting them in an inactive form and only unmasking them when they are in the target region[74–76]. Two main categories of masks have been developed so far: first, masks with affinity for the antigen-binding domains, including peptides [72,73,77] and antibody fragments[78–81], and second, non-binding masks relying on steric hinderance to block the antigen-binding site. This second category includes coiled-coil domains[82], protein M[83], peptide-DNA assemblies[84–86], and a hinge[23] region. Several different methods have been developed to restore the activity of the masked antibody once it is at the target site. The most widely developed approach has been activation by hydrolysis with enzymes that are overexpressed in the tumour or organ of interest [71,72,89,90,77,79–83,87,88]. Alternatives have also been investigated, such as photo-cleavable linkers[84], pH-dependent peptide-DNA locks[85] or pH-sensitive masking peptides[91]. A number of masked antibodies have been successfully developed which effectively decrease binding to the target, when not at the tumour site/in the correct organ, and improve therapeutic index (i.e., the ratio of maximal tolerated dose to therapeutic dose). However, in the case of affinity masks, it is not yet well understood how to optimally design a masking moiety that has sufficient affinity to prevent the binding before activation, but that can efficiently dissociate after cleavage to allow binding to the target antigen and activation.

The objective of the research presented in this Chapter is to understand what parameters impact the inactivation efficacy of a mask on an antibody, and how the relative affinities of the antibody for the mask *versus* the antigen dictate the level of activation required. In this study, masked antibodies were developed based on the humanized antibody trastuzumab (Herceptin, by Roche/Genentech[19,177], approved by the FDA in 1998), prescribed to treat HER2-positive breast cancer, but which can provoke cardiotoxicity given that some HER2 is expressed, albeit at lower levels, in cardiomyocytes[68,69,178]. Anti-idiotypic antibody fragments, which had been previously developed as vaccines for breast cancer by mimicking HER2/NEU[132,133], were fused to

trastuzumab to prevent binding to HER2. In addition, an antibody fragment that did not bind to the CDR region of trastuzumab, was also employed to probe the effect of steric hinderance alone. While other excellent studies have shown successful masking of therapeutic antibodies, no clear basic set of rules for designing effective masks have been established. Here, by studying a series of masks with different affinities and association/dissociation rate constants, the factors that are critical for both efficient masking and activation are established, generating fundamental principles that should be widely applicable to other masking strategies and studies.

6.2 RESULTS

6.2.1 Design of a panel of masked antibodies with variable affinities

Three antibody fragments with affinity for trastuzumab were selected as anti-idiotypic masks. These were based on two single-chain variable fragments (scFv40 and scFv69) and one nanobody (dAb), previously developed and characterized by the Navarro-Teulon group[132,133]. The three masks were successfully expressed and purified in isolation, as well as fusion proteins with trastuzumab on the light chain (LC) to generate three masked antibodies (**Figure 6.1 A**). An additional construct was also designed (T-scFvGipg013), using a scFv mask derived from the Gipg013 antibody, which targets the GIP receptor[134], to investigate the steric effects of a non-binding mask and to act as a control (**Figure 6.1 A**). The antibody NIP228 was used as a negative control as it does not bind to HER2[140]. The antibodies with a single-chain mask were designed with two digestion sites: one is TEV protease-specific, located in the linker between the antibody and the mask, and the other is Factor Xa protease-specific, in the linker between the two single-chain domains (**Figure 6.1 B**). The antibody fused with dAb just had a TEV digestion site in the linker (**Figure 6.1 C**). The sequences of each construct can be found in Appendix I.

After expression and purification following the methods described in Section 2.1.2.3, all constructs were quality controlled using SDS-PAGE (**Figure 6.2**) and liquid chromatography mass spectrometry (LC-MS), **Figure 6.3** to **Figure 6.10**). A comparison between the theoretical and experimental masses determined using LC-MS can be found in **Table 6.1**.

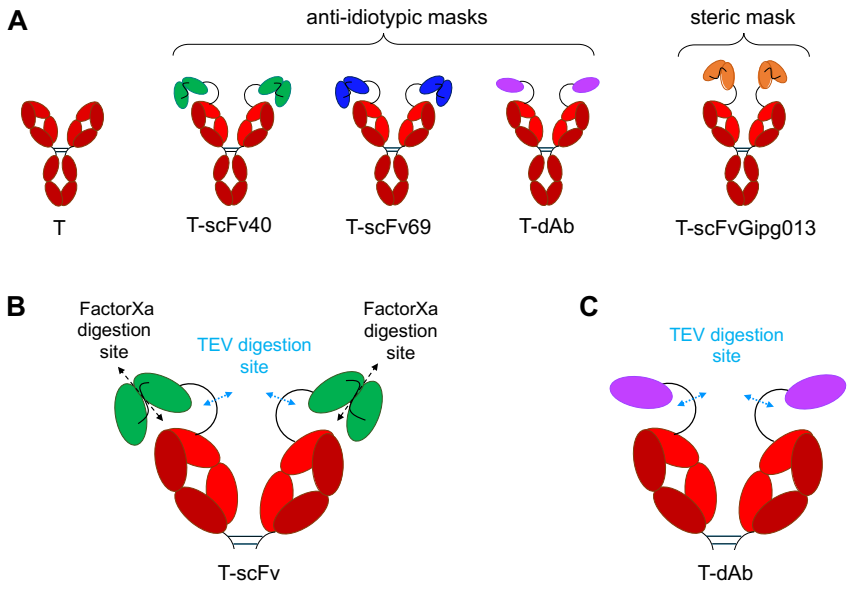


Figure 6.1: **Schematic of the antibodies developed for this study and the location of the digestion site(s). A.** "naked" trastuzumab (T) serves as a positive control; T-scFv40, T-scFv69 and T-dAb are fusions of trastuzumab with various anti-idiotypic antibody fragments; T-scFvGipg013 is a fusion of trastuzumab with a scFv which has no affinity for trastuzumab. **B.** Detailed representation of the digestion sites on the T-scFv constructs: TEV protease cleaves the linker between the antibody and the mask, and Factor Xa protease digests the linker between the two single-chain domains in the T-scFv constructs. **C.** Detailed representation of the digestion site on the T-dAb fusion: TEV protease cleaves the linker between the antibody and the mask. For clarity, panels B and C show the Fab arms of trastuzumab only.

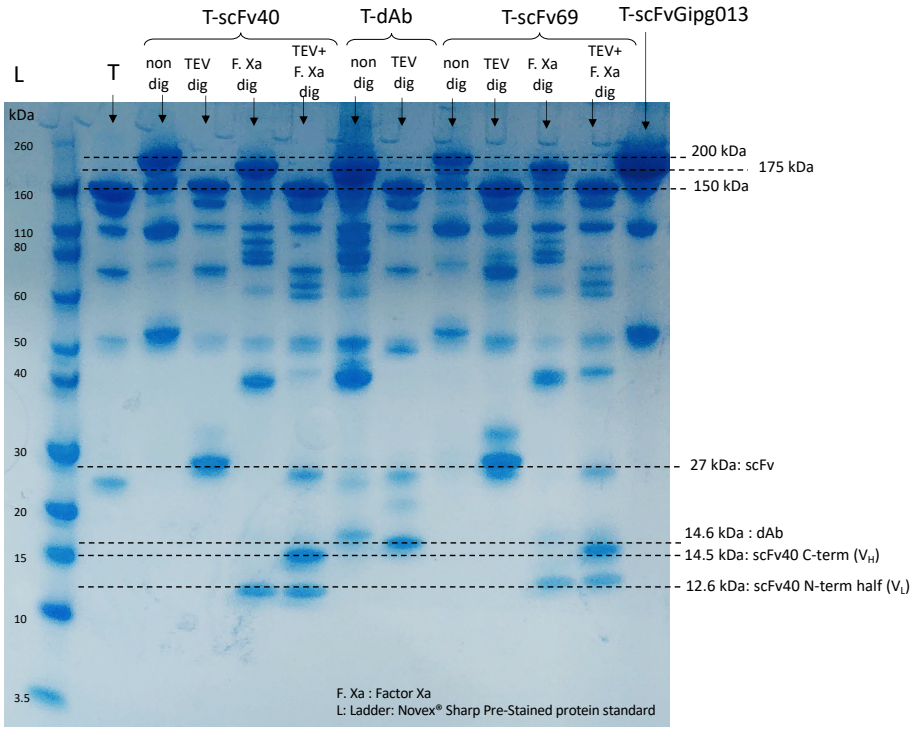


Figure 6.2: SDS-PAGE of the purified antibodies and scFvs and verification of the digestion of the masked antibodies (IgG) with TEV and/or Factor Xa protease. T: trastuzumab; T-dAb: dAb fused to trastuzumab on LC; T-scFv40: scFv40 fused to trastuzumab on LC; T-scFv69: scFv69 fused to trastuzumab on LC. L: Ladder: Novex® Sharp pre-stained protein standards (ThermoFisher).

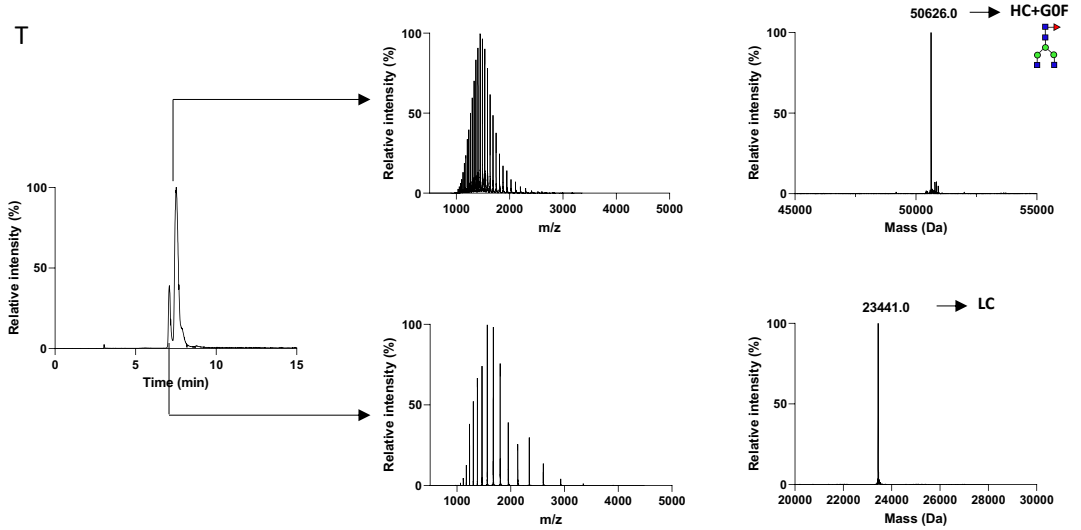


Figure 6.3: Verification of the correct molecular weight of trastuzumab IgG by LC-MS in its glycosylated reduced form (chromatogram, lock mass corrected combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Synapt G2 (Waters). The deconvoluted mass of the heavy chain corresponded to the sum of the protein molecular weight and the glycosylation G0F (shown in the top right-hand corner).

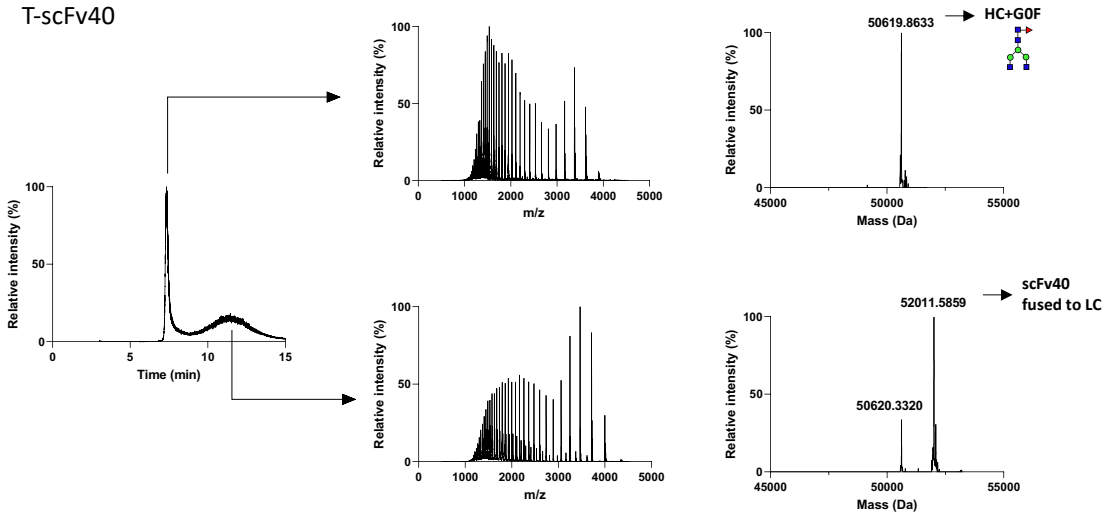


Figure 6.4: Verification of the correct molecular weight of trastuzumab IgG masked by scFv40 (T-scFv40) by LC-MS in its glycosylated reduced form (chromatogram, combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Q Exactive Orbitrap (ThermoFisher). The deconvoluted mass of the heavy chain corresponded to the sum of the protein molecular weight and the glycosylation G0F (shown in the top right-hand corner).

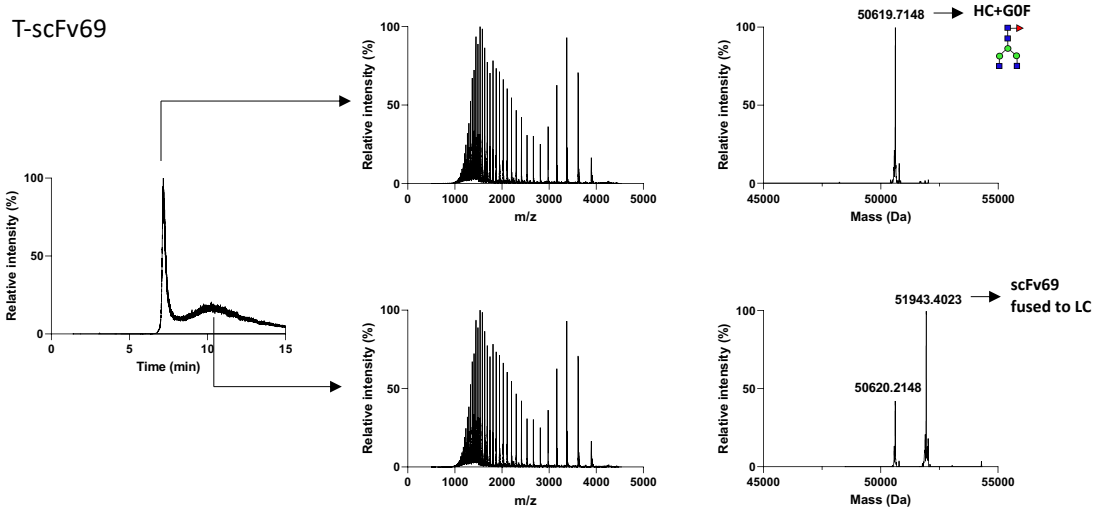


Figure 6.5: Verification of the correct molecular weight of trastuzumab IgG masked by scFv69 (T-scFv69) by LC-MS in its glycosylated reduced form (chromatogram, combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Q Exactive Orbitrap (ThermoFisher). The deconvoluted mass of the heavy chain corresponded to the sum of the protein molecular weight and the glycosylation G0F (shown in the top right-hand corner).

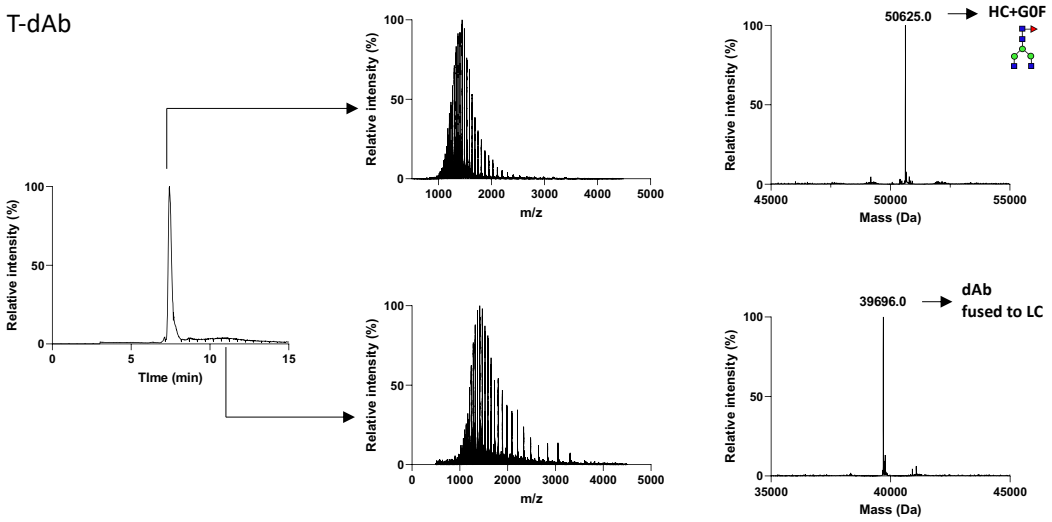


Figure 6.6: Verification of the correct molecular weight of trastuzumab IgG masked by dAb (T-dAb) by LC-MS in its glycosylated reduced form (chromatogram, combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Synapt G2 (Waters). The deconvoluted mass of the heavy chain corresponded to the sum of the protein molecular weight and the glycosylation G0F (shown in the top right-hand corner).

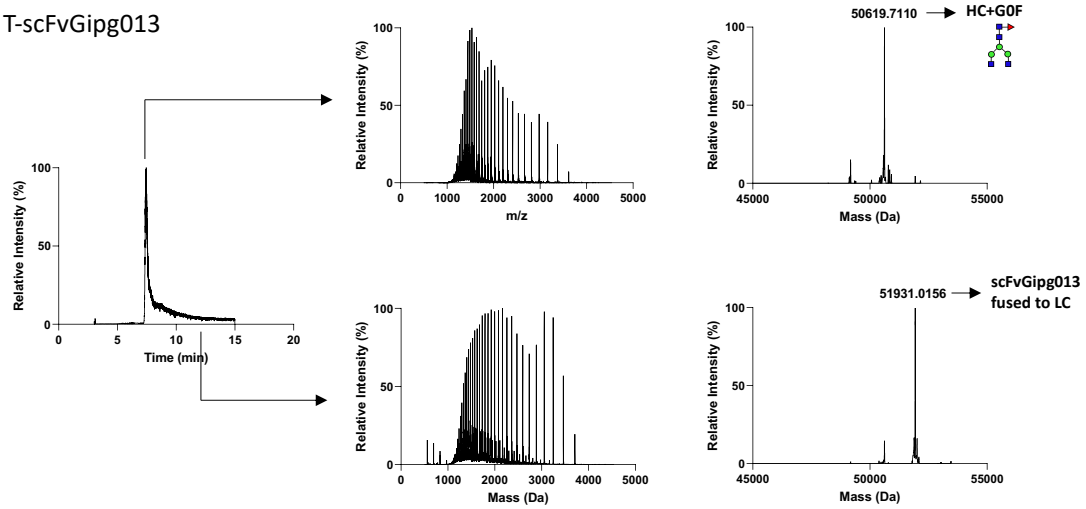


Figure 6.7: Verification of the correct molecular weight of trastuzumab IgG masked by scFvGipg013 (T-scFvGipg013) by LC-MS in its glycosylated reduced form (chromatogram, combined spectra and deconvoluted mass). Mass spectrometry data acquired on a QExactive Orbitrap (ThermoFisher). The deconvoluted mass of the heavy chain corresponded to the sum of the protein molecular weight and the glycosylation G0F (shown in the top right-hand corner).

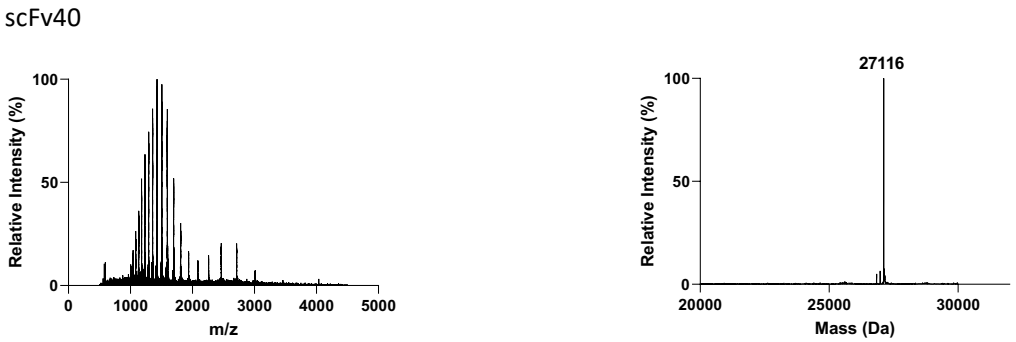


Figure 6.8: Verification of the correct molecular weight of scFv40 by LC-MS (combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Synapt G2 (ThermoFisher). The protein is a single chain resulting in a single peak corresponding to the protein in the chromatogram (data not shown).

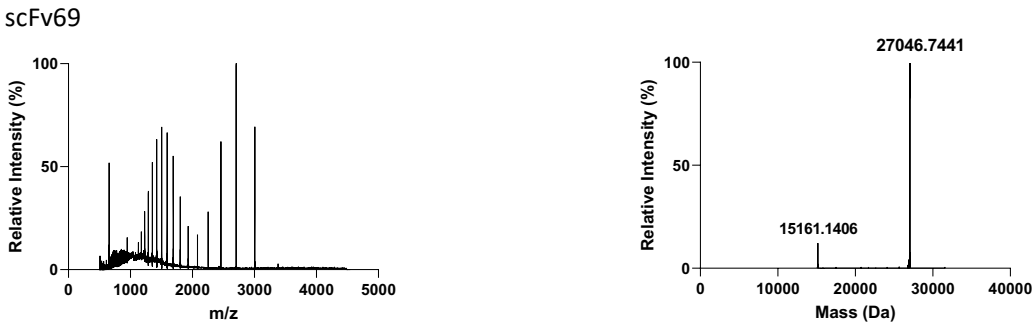


Figure 6.9: Verification of the correct molecular weight of scFv69 by LC-MS (combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Q Exactive Orbitrap (ThermoFisher). The protein is a single chain resulting in a single peak corresponding to the protein in the chromatogram (data not shown).

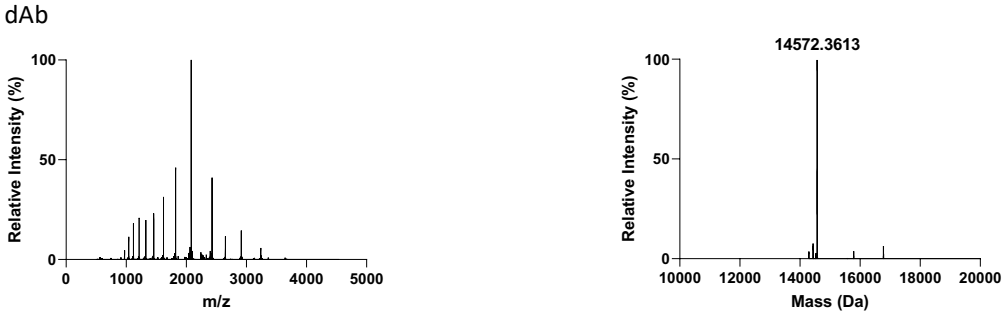


Figure 6.10: Verification of the correct molecular weight of dAb by LC-MS (combined spectrum and deconvoluted mass). Mass spectrometry data acquired on a Q Exactive Orbitrap (ThermoFisher). The protein is a single chain resulting in a single peak corresponding to the protein in the chromatogram (data not shown).

Table 6.1: Comparison between the theoretical and experimental masses

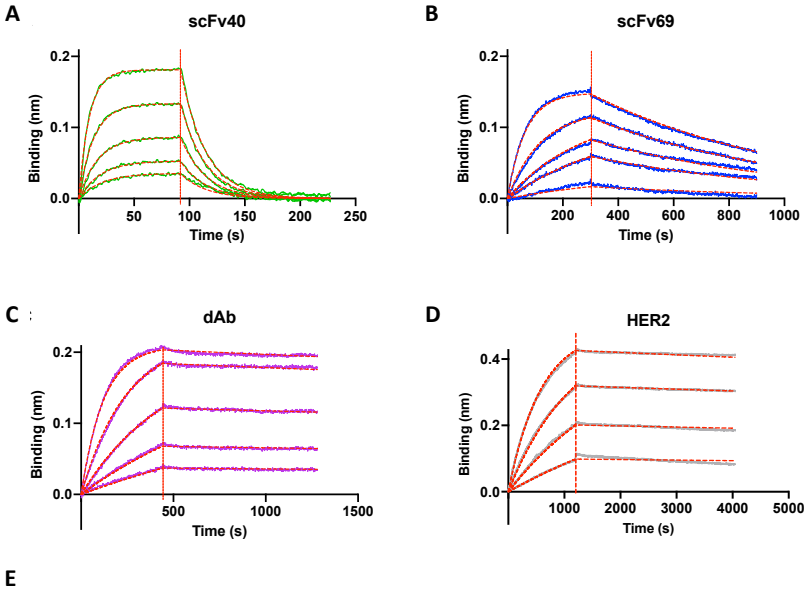
		Theoretical Average Mass (Da)	Experimental Mass (Da)	Difference (ppm)
T	HC*	50621.16	50626	96
	LC	23438.79	23441	94
T-scFv40	HC*	50621.16	50619.86	-26
	LC	52012.68	52011.59	-20.99
T-scFv69	HC*	50621.16	50619.72	-28.48
	LC	51944.65	51943.40	-23.98
T-dAb	HC*	50621.16	50625	75.93
	LC	39692.26	39696	94.28
T-scFvGipg013	HC*	50621.16	50619.71	-28.56
	LC	51931.56	51931.02	-10.44
scFv40	-	27115.43	27116	21.02
scFv69	-	27047.40	27046.74	-24.25
dAb	-	14572.77	14572.36	-28.05

*The theoretical average mass for the heavy chain corresponds to the mass with the C-terminal K clipped and the G0F glycosylation. The samples measured on the Synapt G2 were deconvoluted to the nearest Dalton, whereas those measured on the Q Exactive had nine significant figures, seven are shown in the Table. The difference is routinely calculated at AstraZeneca to ensure confidence in the identification; if the ppm is between -100 and 100, the identification is considered correct. It is calculated as the difference between the theoretical and experimental masses, divided by the theoretical mass, multiplied by 1,000,000 (**Equation 2.2**).

1.2.1 Affinities of the different binding masks

The affinity of the anti-idiotypic antibody fragments (on their own) scFv40, scFv69 and dAb for trastuzumab was measured by biolayer interferometry (BLI). Data

acquired at different mask concentrations were globally fit to obtain high confidence values for the dissociation (k_{off}) and association (k_{on}) rate constants, and the ratio of the two, i.e., the equilibrium dissociation constant (K_D), **Figure 6.11 E**. The measured affinities of the masks for trastuzumab were 540 ± 10 nM for scFv40, 56.5 ± 0.03 nM for scFv69 and 0.192 ± 0.003 nM for dAb, which are lower (scFv40 and scFv69) or similar (dAb) to the affinity of HER2 for trastuzumab (K_D = 0.071 ± 0.001 nM). Although the exact values differ from those previously reported (likely due to the inversion of the V_H and V_L domains), overall, the three antibody fragments showed the same trends as previously reported[132,133] : scFv40 has a large k_{on} but also a large k_{off}, dAb has a large k_{on} and a very low k_{off}, with values very similar to those of HER2 (**Figure 6.11**), and scFv69 has the lowest k_{on} while its k_{off} value is between those of scFv40 and dAb. The selected masks form a comprehensive panel with different affinities, ranging from 0.2 to 540 nM, as well as different association and dissociation rate constants.



E

	K_D (M)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})
scFv40	(5.4 ± 0.1) x 10^{-7}	(8.1 ± 0.2) x 10^{4}	(4.40 ± 0.02) x 10^{-2}
scFv69	(5.65 ± 0.03) x 10^{-8}	(2.37 ± 0.01) x 10^{4}	(1.340 ± 0.003) x 10^{-3}
dAb	(1.92 ± 0.03) x 10^{-10}	(2.993 ± 0.007) x 10^{5}	(5.74 ± 0.08) x 10^{-5}
HER2	(7.07 ± <0.1) x 10^{-11}	(2.483 ± 0.005) x 10^{5}	(1.76 ± 0.02) x 10^{-5}

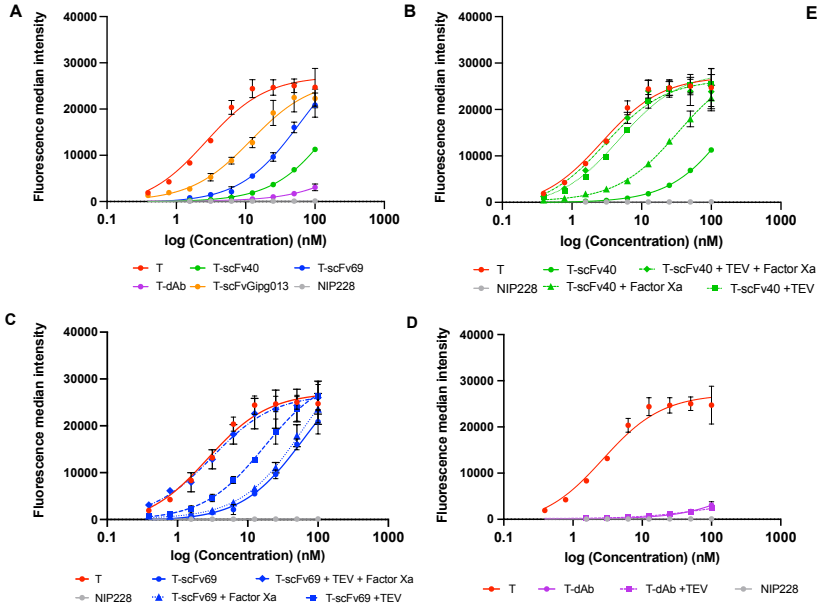
Figure 6.11: Binding curves showing the association and dissociation of different masks to trastuzumab. A. scFv40 (green: 46, 92, 184, 369, 738 nM). **B.** scFv69 (blue: 25, 100, 200, 300, 600 nM). **C.** dAb (purple: 1.7, 3.4, 6.9, 13.7, 27.5 nM). **D.** HER2 (grey: 0.9, 1.7, 3.4, 6.9 nM). In all cases, the dotted red line shows the best fit of the data to 1:1 binding model. **E.** Kinetic and thermodynamic parameters (K_D, k_{on}, k_{off}) obtained from the best global fit of the association/dissociation data to a 1:1 binding model. The errors given are fitting errors from the global fitting.

6.2.2 Efficacy of masks against HER2 binding while tethered to the antibody

The masking efficacy of scFv40, scFv69, dAb and scFvGipg013 fused to trastuzumab against HER2 was assessed using three different methods: i) biolayer interferometry, where a fixed concentration of the biotinylated antigen HER2 was first bound to streptavidin biosensors and subsequently, the binding of the antibodies to HER2 was measured; ii) flow cytometry, where the binding of a serial dilution of antibody to HER2, overexpressed on the surface of SK-BR-3 cells, was measured, and the half maximal effective concentration (EC_{50}) extracted; iii) high content imaging (HCI) monitored by confocal microscopy, which shows the amount and the homogeneity of the antibody binding to the surface of SK-BR-3 cells.

The results from the three assays agreed, and showed the same trends (**Figure 6.12 A, Figure 6.13, Figure 6.14**). Naked trastuzumab had the strongest binding to HER2, i.e., lowest EC_{50} of 2.7 ± 0.5 nM, whilst T-scFvGipg013 significantly bound to HER2 although to a lower extent than naked trastuzumab. Its EC_{50} value was an order of magnitude higher (EC_{50}: 12 ± 2 nM, **Figure 6.12 E**), suggesting that the control scFv causes some steric effect and reduces the binding somewhat, but does not suppress it. Much lower binding to HER2 was observed for T-scFv40, T-scFv69 and T-dAb at the concentration used for BLI (12.5 μM, **Figure 6.13**), as well as for the lowest concentrations probed by flow cytometry and HCI (≤ 6.25 μM, **Figure 6.12 A, Figure 6.14**), which highlights that affinity of the covalently linked antibody fragments for the antibody is necessary to achieve effective masking.

The flow cytometry and confocal microscopy provided further insight into the binding of the masked antibodies to HER2 at higher concentrations. Although the anti-idiotypic masks strongly reduced the binding of trastuzumab to HER2, none of them totally prevented it, most probably due to the competition with an antigen with a higher affinity for the antibody. T-dAb showed the least binding to HER2 even at high concentrations (weak binding, **Figure 6.12 E**), indicative of very effective masking, followed by T-scFv40 (EC_{50}: 84 ± 16 nM, **Figure 6.12 E**), and T-scFv69 (EC_{50}: 30 ± 9 nM, **Figure 6.12 A, Figure 6.14, Figure 6.13**).



E

IgG constructs	EC_{50} (nM)
T	2.7 ± 0.5
T-scFv40	84 ± 16
T-scFv40 +TEV	4.8 ± 0.7
T-scFv40 +factorXa	27 ± 3
T-scFv40 +TEV +factorXa	2.9 ± 0.7
T-dAb	weak binding
T-dAb +TEV	weak binding
T-scFv69	30 ± 9
T-scFv69 +TEV	15 ± 1
T-scFv69 +factorXa	28 ± 9
T-scFv69 +TEV +factorXa	3.2 ± 0.4
T-scFvGipg013	12 ± 2
NIP228	no binding

Figure 6.12: Fluorescence median intensity measured by flow cytometry with channel detecting the AlexaFluor 647 fluorophore signal from a secondary antibody. All experiments were performed in triplicate, mean and standard deviation are plotted. The data in Panel A show the masking efficacy of the masked antibodies and control, whilst the data in panels B-D show the degree to which the masked antibodies can be activated by cleavage with proteases. **A.** Fluorescence median intensity of the secondary antibody showing the binding of the masked intact antibodies against HER2. **B.** Fluorescence median intensity of the secondary antibody showing the binding of the T-scFv40 constructs cleaved with TEV and/or Factor Xa to HER2, illustrating the degree of activation of the T-scFv40 upon cleavage. **C.** Fluorescence median intensity of a secondary antibody showing the binding of T-scFv69 to HER2 after cleavage with TEV and/or Factor Xa proteases. **D.** Fluorescence median intensity of the secondary antibody showing the binding of T-dAb to HER2 after cleavage of the linker with TEV protease. **E.** EC_{50} parameters obtained from the flow cytometry measurements, data fitted to **Equation 2.15.**

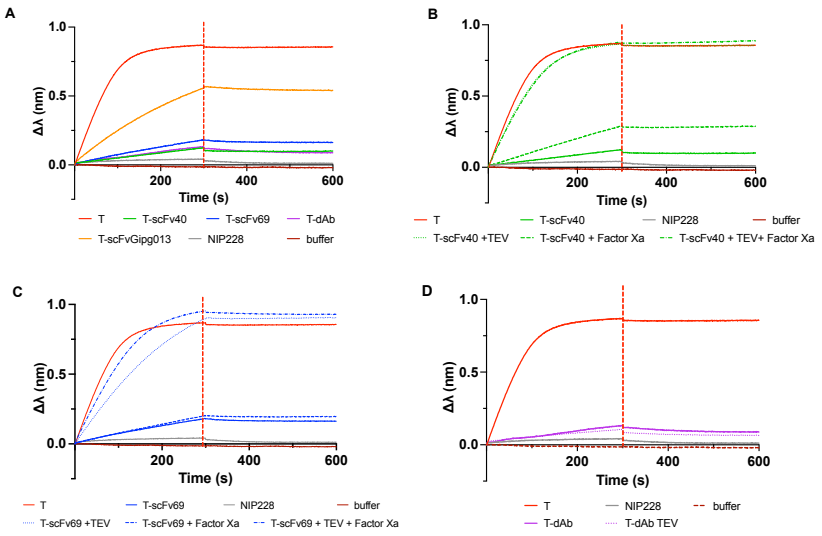


Figure 6.13: Biolayer interferometry: measurement of the wavelength shift due to binding of the trastuzumab constructs to HER2. The first half of the time course up to 500 s corresponds to the association phase, whilst the second half corresponds to the dissociation phase. **A** shows the masking efficacy, whilst **B-D** show the activation efficiency. **A.** Binding of the intact covalently linked masked antibodies and controls to HER2. **B.** Binding of T-scFv40 after cleavage of the linker(s) with TEV and/or Factor Xa to HER2 showing activation efficiency. **C.** Binding of T-scFv69 after cleavage of the linker(s) with TEV and/or Factor Xa to HER2, showing activation efficiency. **D.** Binding of T-dAb after cleavage of the linker against with TEV to HER2.

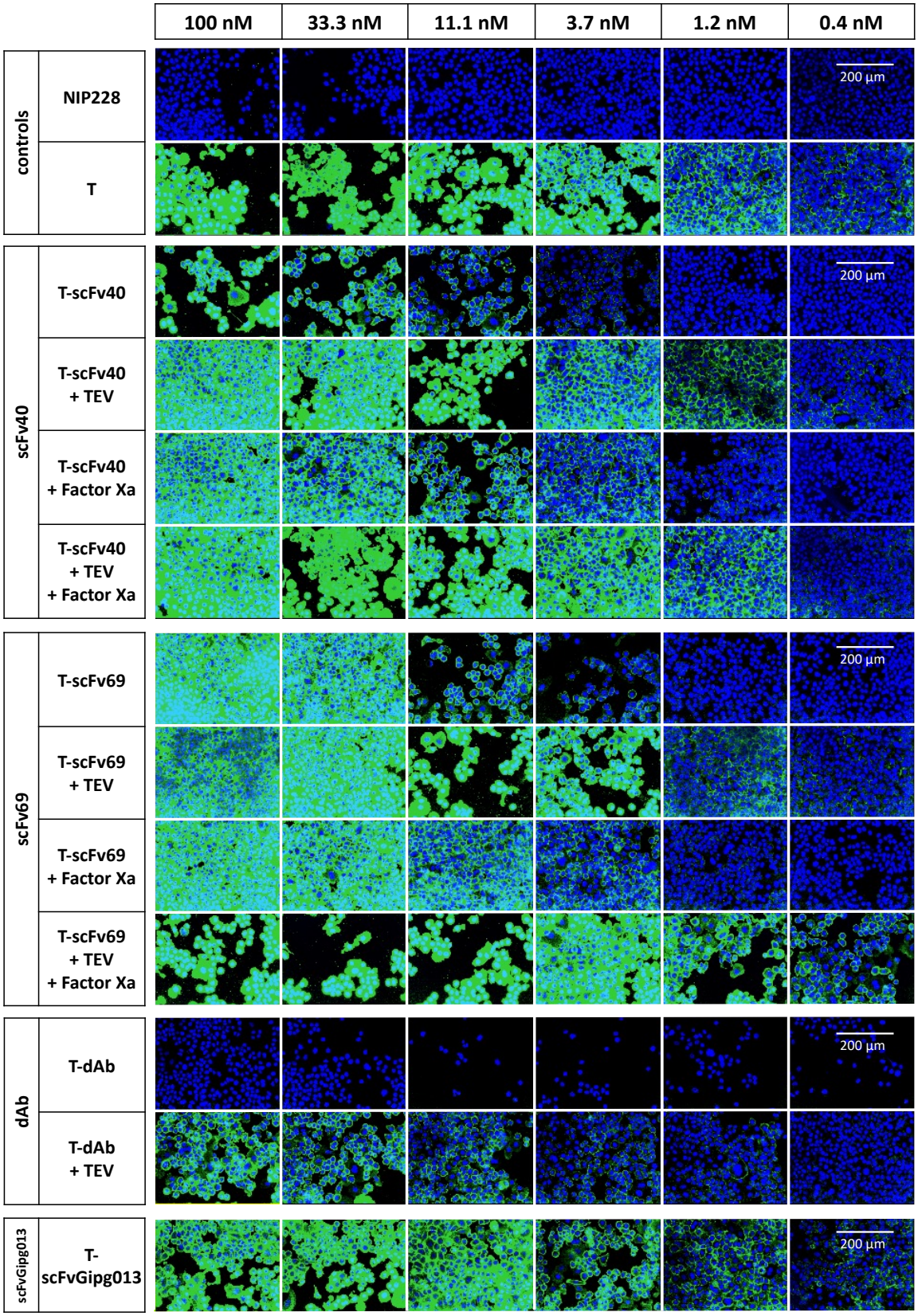


Figure 6.14: Antibody binding to HER2 on SK-BR-3 cells, measured by immunofluorescence before and after cleavage of the linker between the mask and trastuzumab. T alone is a positive control whilst NIP228 alone is a negative control; T-scFv40 and T-scFv69, intact, digested by TEV, Factor Xa, and both enzymes; T-dAb intact and digested by TEV; T-scFvGipg013 intact. All antibodies were detected with Alexa488 anti-human secondary (green) and nuclear Hoechst stain (blue). The confocal microscopy data was imaged by Lorraine Irving.

These differences in masking capabilities can be explained by the strength of the competition between the different masks and HER2. At all concentrations, there is a dynamic equilibrium between two different forms of the fusion protein. In one form, the mask is sitting on top of the antigen-binding site, thus masking it. In the other, it has transiently dissociated from the antigen-binding site but remains covalently tethered through the linker (**Figure 6.15**). dAb has a very slow dissociation rate constant from trastuzumab (k_{off} = (5.76 ± 0.08) x 10^{-5} s^{-1}, **Figure 6.11 E**), and therefore comes off very rarely (half-life of dissociation is 3.3 hours), which does not provide many opportunities for HER2 to bind, especially over the time course of the incubation with SK-BR-3 cells.

Additionally, dAb and HER2 have very similar association rate constants, which means that both proteins compete for trastuzumab when the dAb transiently dissociates. Interestingly, scFv40 prevents binding of trastuzumab to HER2 more effectively than scFv69, despite scFv69 having a stronger affinity for trastuzumab and slower dissociation constant than scFv40 (**Figure 6.12 A**). This observation suggests that other parameters must come into play. Importantly, there is a correlation between the k_{on} values and the EC_{50} values of the different masks, but not with K_D nor $k_{off,}$ indicating that the critical parameter may be k_{on}. This can be rationalised by considering the following mechanism: the k_{on} for scFv40 is larger than for scFv69 (despite the weaker affinity of scFv40 for trastuzumab), meaning that an unmasked T-scFv40 will re-mask itself more quickly than an unmasked T-scFv69 would. Because T-scFv40 has a much shorter dwell time in the unmasked state, an unmasked T-scFv69 has a higher likelihood of binding to a HER2 molecule. Once the unmasked T-scFv69 binds a HER2 molecule, its tighter affinity for its mask is irrelevant, because HER2 is now occupying the mask site and will likely not dissociate during the experimental time course as its dissociation rate is extremely slow (**Figure 6.11**). A schematic representation of this mechanism is illustrated in **Figure 6.15**. An alternative explanation to scFv69 being less effective than scFv40 in its ability to mask despite its higher affinity to trastuzumab, is that it might not bind to the exact same location in the complementarity-determining regions (CDRs) of trastuzumab as scFv40, suggesting that scFv69 covers a smaller region of the binding domains.

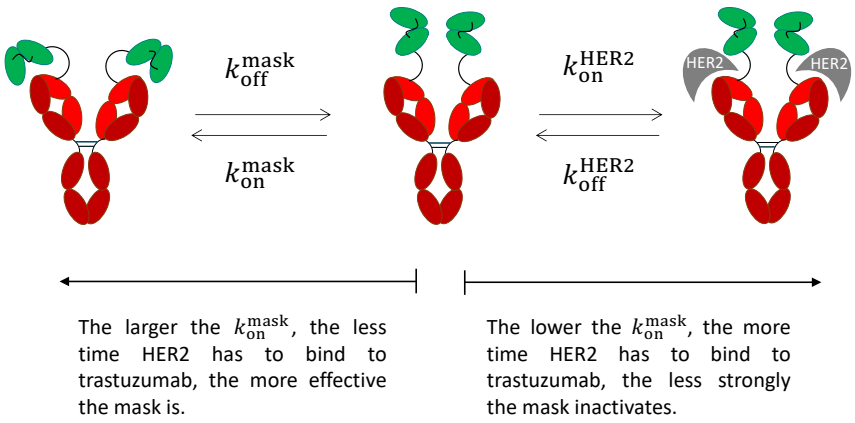


Figure 6.15: Schematic representation of the effect of k_{on} of the mask on the binding of HER2 to the transiently dissociated trastuzumab-mask complex: an ideal mask should have a k_{on}^{mask} which is lower than the antigen's, so it can bind when the linker is cleaved and the mask is unbound, but it shouldn't be too low so that when it is tethered, the mask can compete with the antigen and prevent it from binding.

In addition to providing information on the degree of binding, the confocal microscopy images showed that the binding was homogeneous across the surface of the cell membrane. Internalisation was investigated over a short time course, which showed that trastuzumab did not internalise after 3 hours at 37 °C in the SK-BR-3 cells (**Figure 6.16**), consistent with data previously published[179].

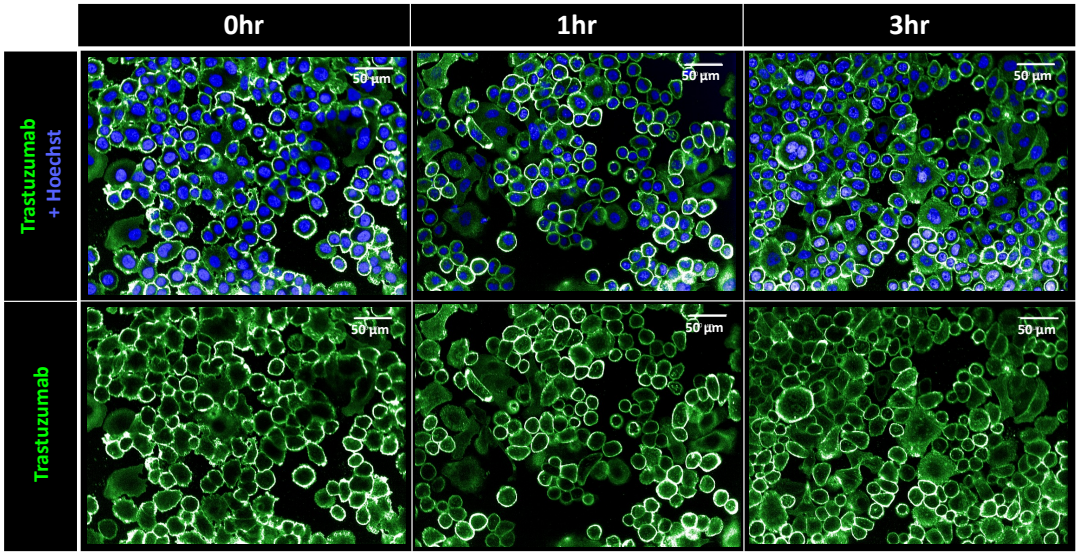


Figure 6.16: High content imaging of trastuzumab (2 nM) by AlexaFluo488 secondary human antibody on breast cancer SK-BR-3 cells overexpressing HER2 after 0, 1 and 3 hours of incubation at 37 °C. The binding of trastuzumab is shown in green, with or without the nucleus staining (shown in blue). No internalization was observed over 3 hours, consistent with previously published data[179]. The confocal microscopy data was imaged by Lorraine Irving.

Overall, all three masked antibody formats (T-dAb, T-scFv40, T-scFv69) showed significantly increased EC_{50} values compared to trastuzumab and therefore meet the first requirement to decrease the binding of trastuzumab to HER2.

Some Hook effect was observed at high concentrations (full raw data are presented in **Figure 6.17**). This is caused by excessively high concentrations of the primary antibody simultaneously saturating both HER2 and the secondary antibodies. The high-dose hook effect occurs mostly (but not exclusively) in one-step immunometric (sandwich) assays, giving a decrease in signal at very high concentration of primary antibody[180–182]. Therefore, the signal obtained is probably caused to insufficient washes between the incubation of the primary and secondary antibodies. The higher concentration points were excluded from the fitting as they are not representative of the actual binding to HER2 and are artefacts of the protocol followed. The parameters used for the flow cytometer can be found in **Figure 6.18**

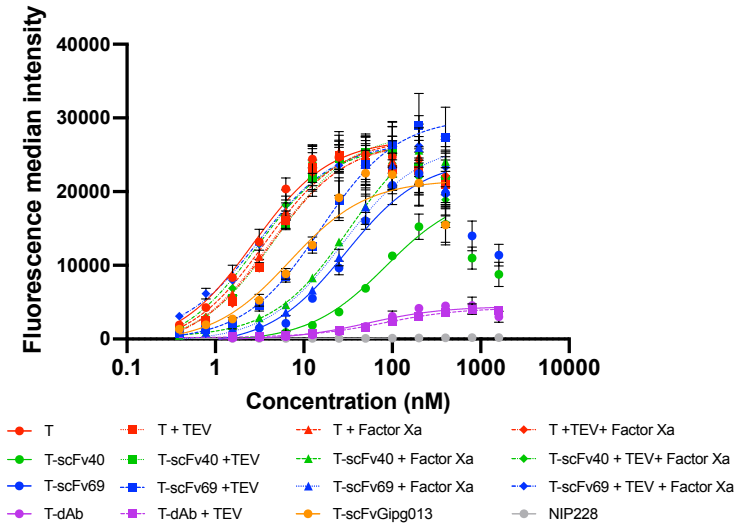


Figure 6.17: Flow cytometry (all data) showing some Hook effect at higher protein concentrations. The fitting shown is that reported in **Figure 6.12 E** and **Table 6.2**.

Table 6.2: EC_{50} values for trastuzumab and digested controls

	EC_{50} (nM)
T	2.7 ± 0.6
T +TEV	4 ± 1
T + Factor Xa	3.6 ± 0.8
T +TEV+ Factor Xa	4 ± 1

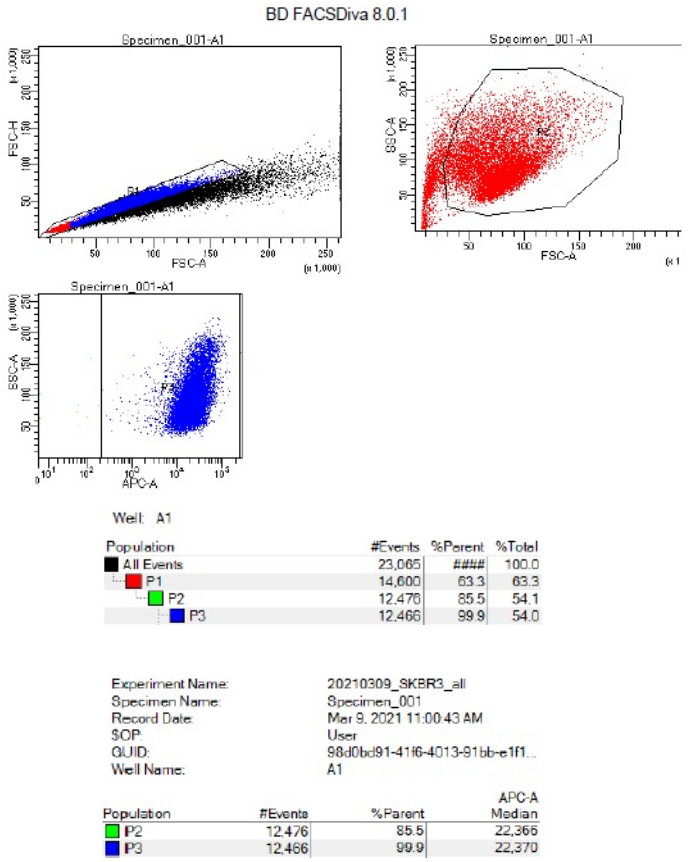


Figure 6.18: Flow cytometry raw data obtained for cells incubated with 400 nM protein, details of the gating strategy used to analyse and evaluate the binding to HER2 on SK-BR-3 cells. In the first gate (P1), singlets were selected (FSC-H: forward scatter height; FSC-A: forward scatter area). In the second gate (P2), the viable cells are selected (SSC-A: side scatter area). In the third gate (P3), the fluorescence on cells above a certain threshold was measured with the APC channel to detect Alexa Fluor 647. The median fluorescence intensity in the APC channel for the P2 population was reported.

6.2.3 Unmasking efficacy: HER2 binding measurements after digestion of the linker(s)

Following these observations, the activation of the antibody upon cleavage of the linker between the mask and trastuzumab or the linker between the two domains of the scFv masks was assessed. In particular, for the antibodies masked with scFvs, the aim was to understand if having two different cleavage sites, one between the antibody and the scFv, the other between the two domains of the scFv (**Figure 6.1 B and C**), has a synergistic effect on activation. To evaluate this possibility, BLI, flow cytometry and high-content imaging were used following the same protocol as described above, this time after digesting T-scFv40, T-scFv69 and T-dAb with TEV, Factor Xa, or both enzymes prior to the measurements. The completion of the digestion was verified by SDS-PAGE (**Figure 6.2**).

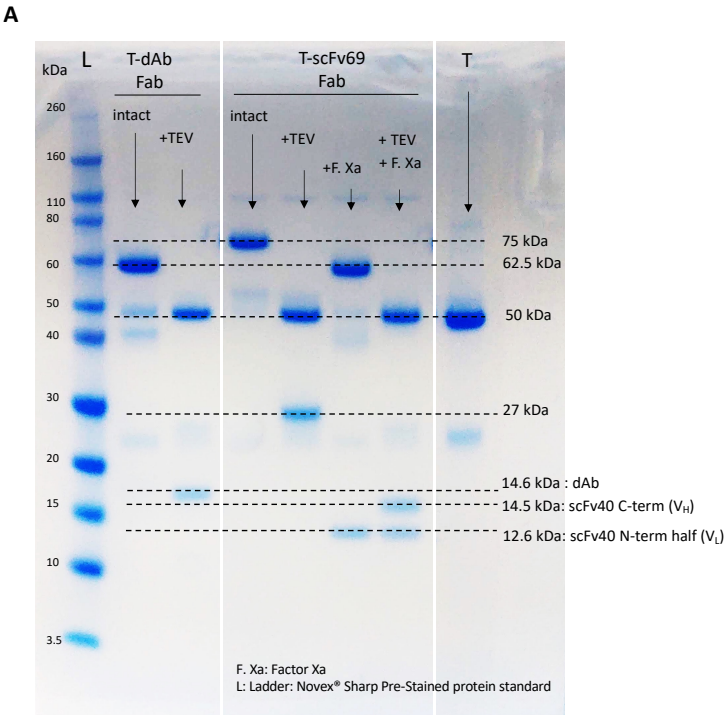
The results of the three assays showed that, for T-scFv40, cleaving the linker between the two scFv domains (Factor Xa digestion) only slightly increases its binding to HER2, but the digestion of the linker between antibody and the mask alone (TEV digestion) fully activated the antibody. As expected, cleaving both digestion sites also fully activated the antibody (**Figure 6.12 B, Figure 6.14, Figure 6.13 B**). In contrast, for T-scFv69, the cleavage of the linker between the two scFv domains by Factor Xa did not activate the antibody, whilst the digestion of the linker connecting the scFv to the antibody (TEV digestion) only partially activated the antibody. The digestion of both linkers, however, fully restored the binding of trastuzumab to HER2 (**Figure 6.12 C, Figure 6.14, Figure 6.13 C**). For T-dAb, the cleavage of the linker between the antibody and the mask by TEV protease did not result in the activation of the antibody (**Figure 6.12 D, Figure 6.14, Figure 6.13 D**).

These different levels of activation highlight the importance of the affinity of the mask for trastuzumab and the location of the digestion sites. scFv40 is the mask with the weakest affinity, in the micromolar range (**Figure 6.11**); when it is covalently linked to trastuzumab it increases the EC_{50} by two orders of magnitude compared to naked trastuzumab for binding to HER2 on SK-BR-3 cells (**Figure 6.12 E**). In this case, simply cleaving the linker between the antibody and the mask is sufficient to activate the antibody in the presence of HER2. This suggests two mechanisms: first, that the covalent linkage of scFv40 to the antibody increases the local concentration of the mask and therefore applies a stronger competition with HER2 for the variable domains of trastuzumab; second, that in the absence of any covalent linkage between the antibody and the mask, trastuzumab preferentially binds to the very high affinity antigen HER2, over the lower affinity svFv40. The scFv69 mask acts somewhat differently. The affinity of scFv69 for trastuzumab is stronger than that of scFv40, and in particular, the dissociation constant of scFv69 is lower than that of scFv40. In this case, the fact that the digestion of the linker between the antibody and the mask is not sufficient to fully activate it shows that if the mask has a high affinity, it requires the hydrolysis of both digestion sites to fully activate the antibody. Finally, the dAb has a very similar affinity, and dissociation constant compared to the antigen HER2, and therefore the cleavage of the linker has no effect as the mask effectively remains bound on the antibody and doesn't dissociate over the time course of the experiment.

6.2.4 Interaction of the masks with trastuzumab after digestion of the linker

All the results presented previously were measured under conditions where there is competition between the mask and the antigen HER2 for trastuzumab binding. To understand to what extent the masks interact with the CDRs of trastuzumab after digestion, in the absence of competition with HER2, intact and digested samples were run on size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) by Wesley Howard, a collaborator from AstraZeneca (Gaithersburg, USA).

SEC-MALS is a technique in which there is first a separation step by SEC which provides a retention volume depending on the hydrodynamic radius of the species, followed by MALS which can extract the species' molecular weight (MW) when coupled to a concentration detector (refractive index or UV). The Fabs of all the masked antibodies (T, T-dAb, T-scFv40, T-scFv69, T-scFvGipg013) were generated to ensure the best resolution possible given the optimal molecular weight range of the column and the protein standards used. Samples of intact Fabs were run alongside those digested by TEV and/or Factor Xa. Cleavage was verified by SDS-PAGE before the SEC-MALS experiment (**Figure 6.19**).



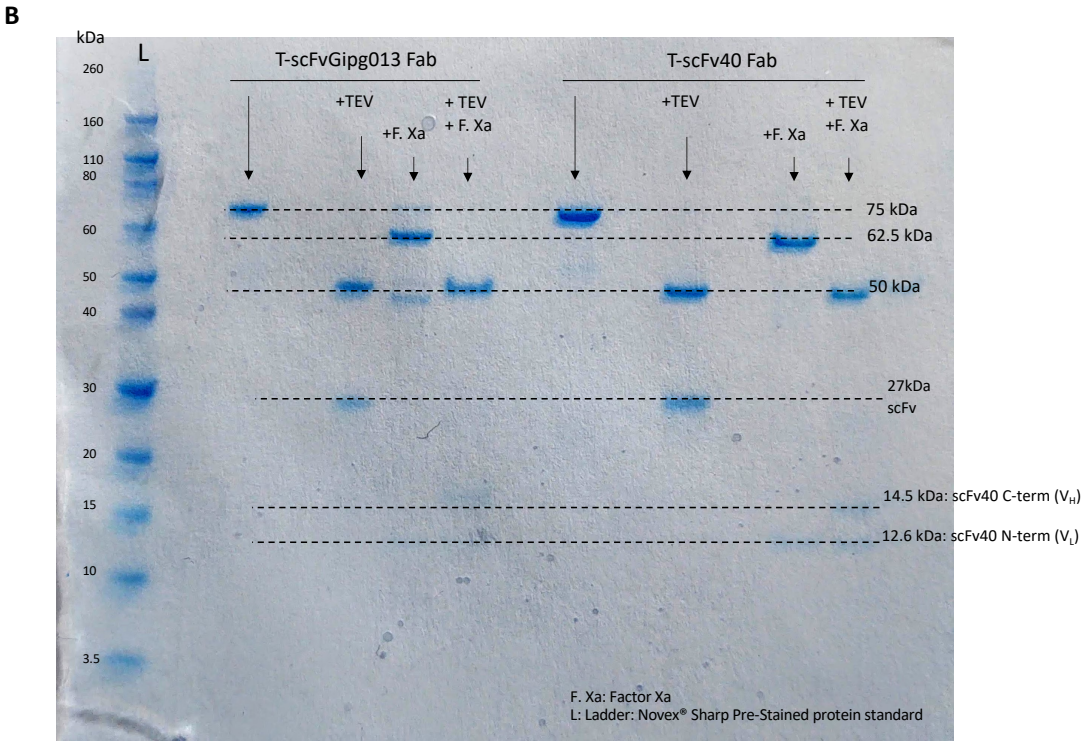


Figure 6.19: Verification of the digestion of the masked antibodies (Fab) with TEV and Factor Xa protease by SDS-PAGE, in preparation for the SEC-MALS experiments.

Table 6.3: Molecular weights of intact and digested mask-Fab fusions measured by SEC-MALS

Fab constructs	Refractive Index MW (kDa)
T	46 ± 2
T-dAb	61 ± 3
T-dAb + TEV	63 ± 3
T-scFv40	74 ± 4
T-scFv40 + TEV*	41 ± 2* 47 ± 2*
T-scFv40 + Factor Xa	71 ± 4
T-scFv40 + TEV + Factor Xa	39 ± 2
T-scFv69	73 ± 4
T-scFv69 + TEV	68 ± 3
T-scFv69 + Factor Xa	71 ± 4
T-scFv69 + TEV + Factor Xa	67 ± 3
T-scFvGipg013	71 ± 4
T-scFvGipg013 + TEV	43 ± 2
T-scFvGipg013 + Factor Xa	53 ± 3
T-scFvGipg013 + TEV + Factor Xa	43 ± 2

The error values shown above are based on the average of fitting errors of duplicate measurements of BSA which gave an error of 5%.

*Two peaks were observed for T-scFv40 cleaved with TEV.

Consistent with the fact that the mask on T-scFvGipg013 has no affinity for trastuzumab, the results of the TEV cleavage of the linker for this construct are within error of the naked trastuzumab Fab, showing that there is no non-specific binding after cleavage (**Table 6.3).** When the linker between the two domains of scFvGipg013 was cleaved by Factor Xa, two peaks were observed, one corresponding to trastuzumab linked to the C-terminal domain of scFvGipg013, and the other to the intact construct (**Figure 6.20 E**). This result suggests that there might be some non-covalent interactions at the interface between the two scFv domains[183]. For T-scFv40 Fab digested by TEV, two peaks, both with MWs close to that of the naked trastuzumab Fab, were observed. While both species appear to be naked Fab, the second is probably naked Fab that was not bound to the mask, while the species eluting earlier might be Fab that had been bound to mask at the start of the experiment, but then dissociated during the chromatography step. The fact that scFv40 has high association and dissociation rates supports this hypothesis. The MW and retention time obtained for T-scFv40 Fab digested with Factor Xa show that the N-term of scFv40 (V_L) which has been cleaved is either still interacting with trastuzumab, or with the C-term of scFv40 (V_H), similar to scFvGip013. When T-scFv40 Fab is digested by both enzymes, only one peak is observed, and the MW measured is closest to naked trastuzumab. For T-scFv69 Fab, none of the digestions triggered the total unmasking of trastuzumab as all MWs are within error of the intact T-scFv69 Fab (**Table 6.3)** apart from a small peak observed with digestion by TEV alone and TEV combined with Factor Xa, which elutes close to naked trastuzumab (**Figure 6.20 C**). Finally, for T-dAb Fab, despite the digestion of the linker between the antibody and the mask, dAb still interacts with trastuzumab (**Table 6.3, Figure 6.20 D**). For all these experiments, the scFvs that were released were not observed as they are smaller than the Fab and the concentration was too low to be quantitatively assessed.

Overall, these results show that the fate of the masks upon cleavage of the various linkers depends on the affinity of the mask for trastuzumab and, where relevant, the interactions between one domain of an scFv and the other: the different masks remaining bound to different degrees to trastuzumab. Thus, activation of the masked antibodies by proteolytic cleavage relies on two factors: the k_{off} of the mask, and the competition between the anti-idiotypic antibody fragments and the targeted antigen.

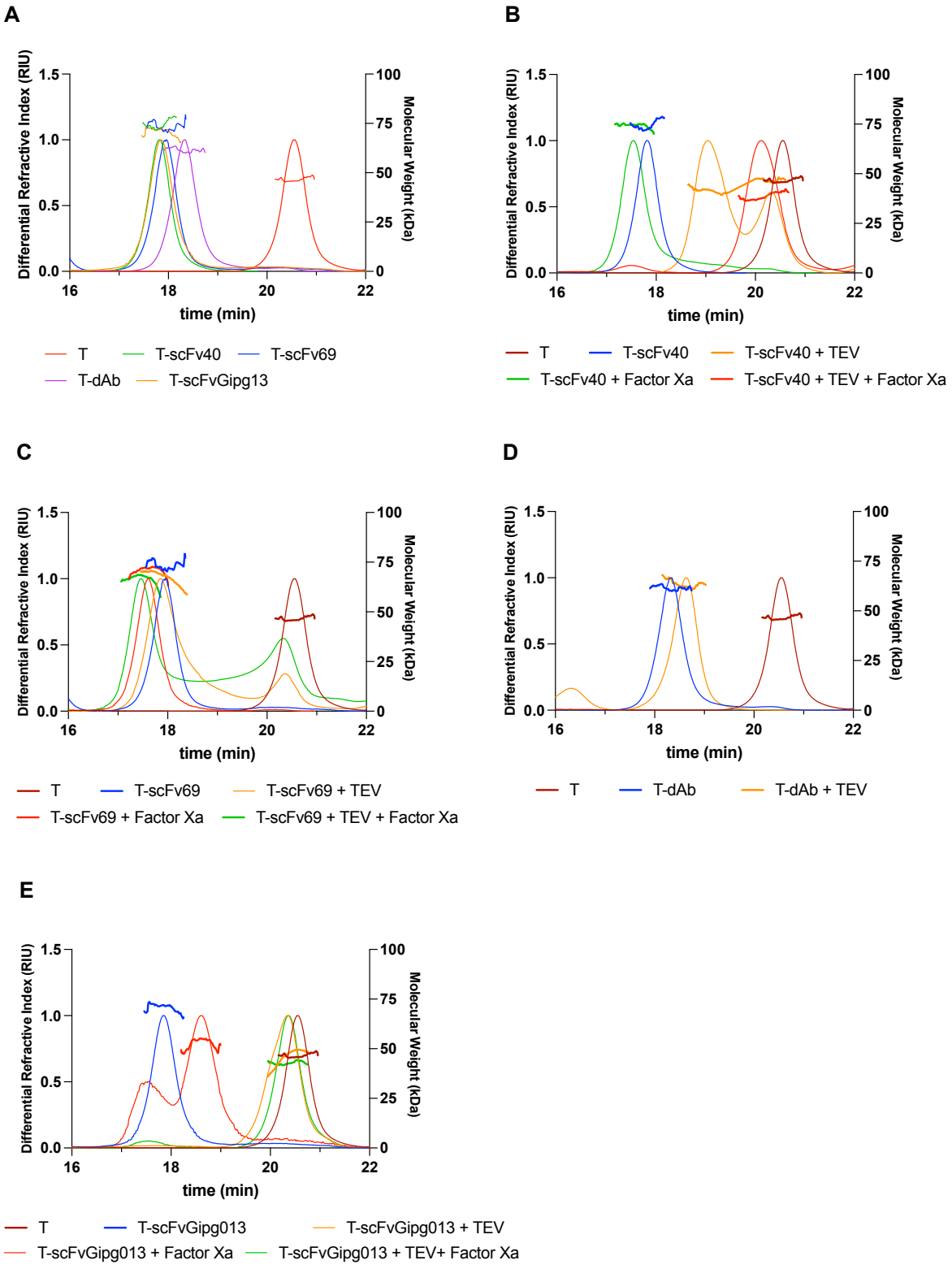


Figure 6.20: SEC-MALS traces for the different masked Fabs digested by TEV and/or Factor Xa proteases. The chromatograms are plotted on the left y-axis, and the MW (the horizontal line above or below the peaks of the chromatograms) on the right y-axis. **A.** All intact masked antibodies. **B.** T-scFv40. **C.** T-scFv69. **D.** T-dAb. **E.** T-scFvGipg013. The SEC-MALS data was acquired by Wesley Howard.

6.2.5 Investigation of the binding sites of the masks on trastuzumab

To measure the interaction surface experimentally in solution, hydrogen-deuterium exchange mass spectrometry (HDX-MS) was employed. This technique is highly sensitive to the solvent accessibility and hydrogen-bonding of backbone amide groups in proteins and has been used successfully to investigate interactions of antibody domains[104,105].

To obtain high-resolution information of the deuterium exchange per residue, a peptide map was acquired using online pepsin digestion. Peptide maps which have a very high coverage of the regions of interest are required to obtain the most accurate results. The MS peptide mapping experiment was performed on full-length naked trastuzumab but only data corresponding to the V_L and V_H domains were analysed as these are the regions of interest. Very high coverage (light chain: 94.8%, heavy chain 95.9%) was obtained (**Figure 6.21**, **Table 6.4**), and importantly all the complementarity-determining regions (CDRs) were fully covered, ensuring reliable information on the CDRs. All the HDX-MS labelling experiments were done on full-length IgG constructs including naked trastuzumab, T-scFv40, T-scFv69, T-dAb intact and cleaved by TEV and/or Factor Xa and T-scFvGipg013 intact, and the deuterium uptake of the control naked trastuzumab was subtracted from the data obtained for the different antibody formats (**Figure 6.22**). The uptake plots are represented in **Figure 6.23** and **Figure 6.24**.

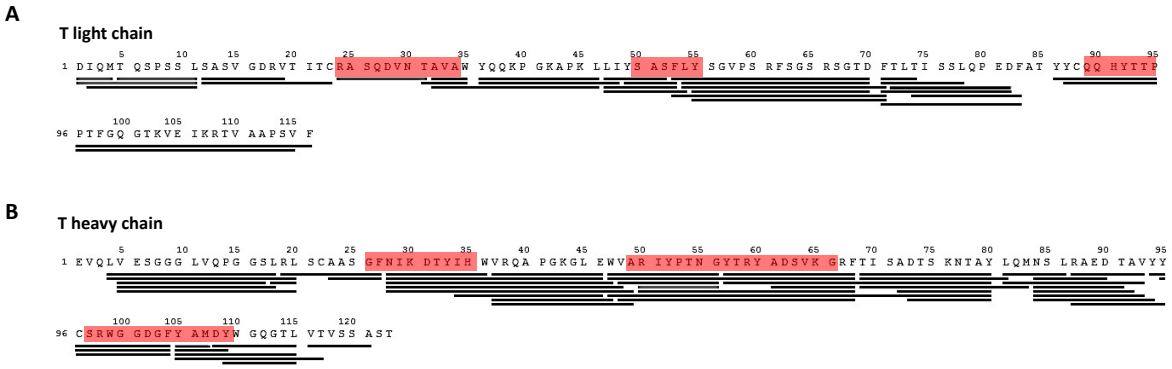


Figure 6.21: Peptide coverages of trastuzumab. **A.** Trastuzumab V_L (94.8 % coverage, 27 peptides). **B.** Trastuzumab V_H (95.9% coverage, 51 peptides), obtained from full-length IgG trastuzumab sample. All the complementarity-determining regions (CDRs) were fully covered (highlighted in red), ensuring reliable information was obtained on the CDRs.

Table 6.4: Summary of HDX mass spectrometry experimental details

Dataset	T (10 µM) T-scFv40 (10 µM) T-scFv40 + TEV (10 µM) T-scFv40 + Factor Xa (10 µM) T-scFv40 + TEV + Factor Xa (10 µM) T-scFv69 (10 µM) T-scFv69 + TEV (10 µM) T-scFv69 + Factor Xa (10 µM) T-scFv69 + TEV + Factor Xa (10 µM) T-dAb (10 µM) T-dAb + TEV (10 µM) T-scFvGipg013 (10 µM)
HDX reaction details	Equilibration in H_2O: 10 mM potassium phosphate pH 7.5 Labelling in D_2O: 10 mM potassium phopshate pD 7.5 Both were done at 20 °C Quench buffer in H_2O: 100 mM potassium phosphate, 0.5 M TCEP, 8 M urea pH 2.5
HDX time course	50 s, 500 s, 5000 s
HDX controls	0 s
Number of peptides	T V_L: 27 T V_H: 51
Sequence coverage	T V_L: 94.8% T V_H: 95.9%
Average peptide length	T V_L: 11.9 AA T V_H: 12.1 AA
Replicates	Technical replicates: 3
Significant differences in HDX	The incorporation of deuterium per peptide is significant compared to the wild-type if p-value < 0.01 (t-test)
Repeatability	Average Student's t-distribution 95% confidence interval for all peptides, all time points and states (#D error): T V_L: 0.339 T V_H: 0.124

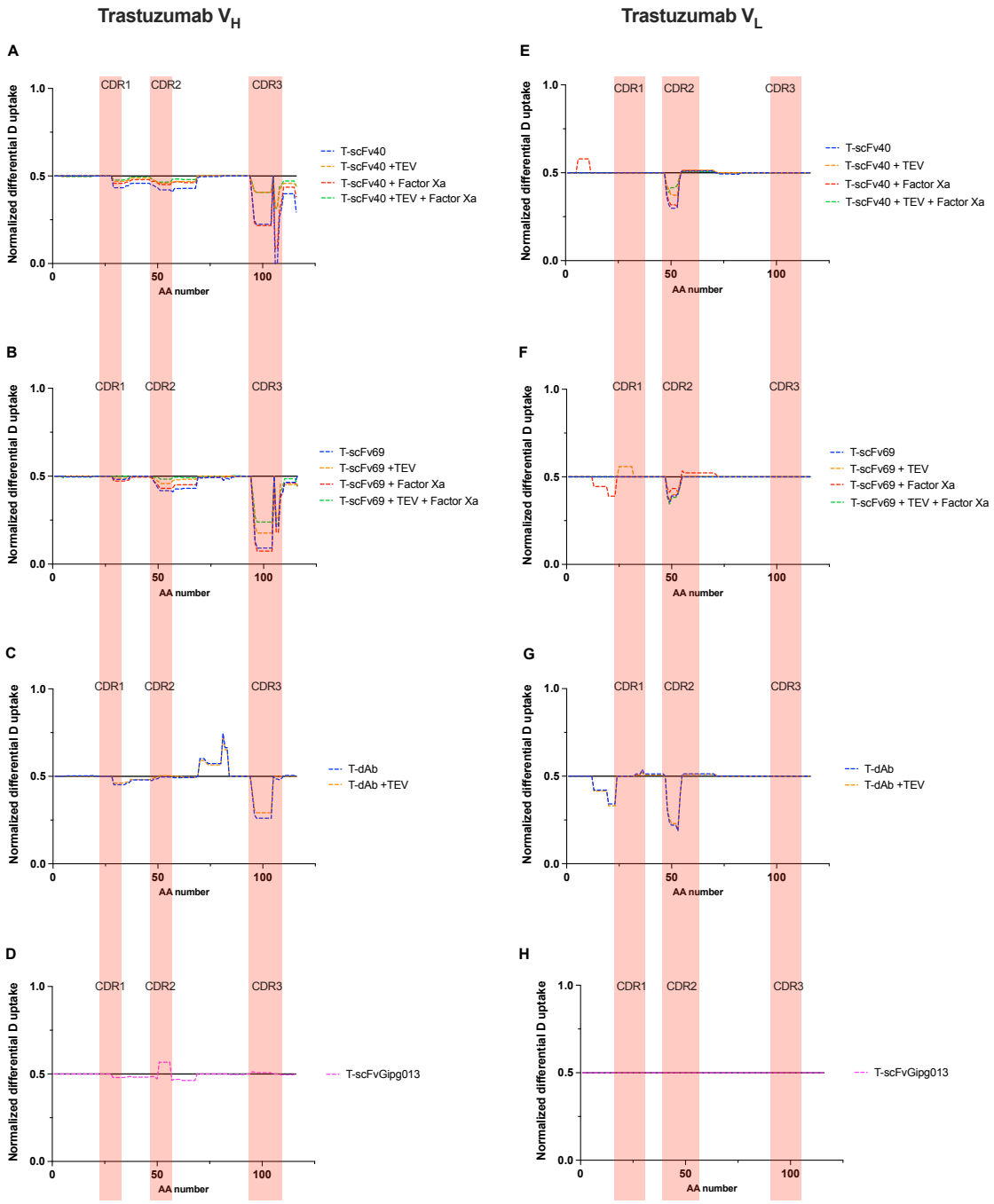


Figure 6.22: Normalized difference plot obtained by summing the significant D-incorporations relative to the reference, and subtracting the reference exchange to that of the different states. T V_L, T V_H were normalized with the same maxima. **A.** Normalized differential deuterium uptake between trastuzumab variable heavy domain (V_H) and T-scFv40, T-scFv40 + TEV, T-scFv40 + Factor Xa. **B.** Normalized differential deuterium uptake between trastuzumab variable heavy domain (V_H) and T-scFv69, T-scFv69 + TEV, T-scFv69 + Factor Xa, T-scFv69 + TEV + Factor Xa. **C.** Normalized differential deuterium uptake between trastuzumab variable heavy domain (V_H) and T-dAb, T-dAb + TEV. **D.** Normalized differential deuterium uptake between trastuzumab variable heavy domain (V_H) and T-scFvGipg013. **E.** Normalized differential deuterium uptake between trastuzumab variable light domain (V_L) and T-scFv40, T-scFv40 + TEV, T-scFv40 + Factor Xa. **F.** Normalized differential deuterium uptake between trastuzumab variable light domain (V_L) and T-scFv69, T-scFv69 + TEV, T-scFv69 + Factor Xa, T-scFv69 + TEV + Factor Xa. **G.** Normalized differential deuterium uptake between trastuzumab variable light domain (V_L) and T-dAb, T-dAb + TEV. **H.** Normalized differential deuterium uptake between trastuzumab variable light domain (V_L) and T-scFvGipg013.

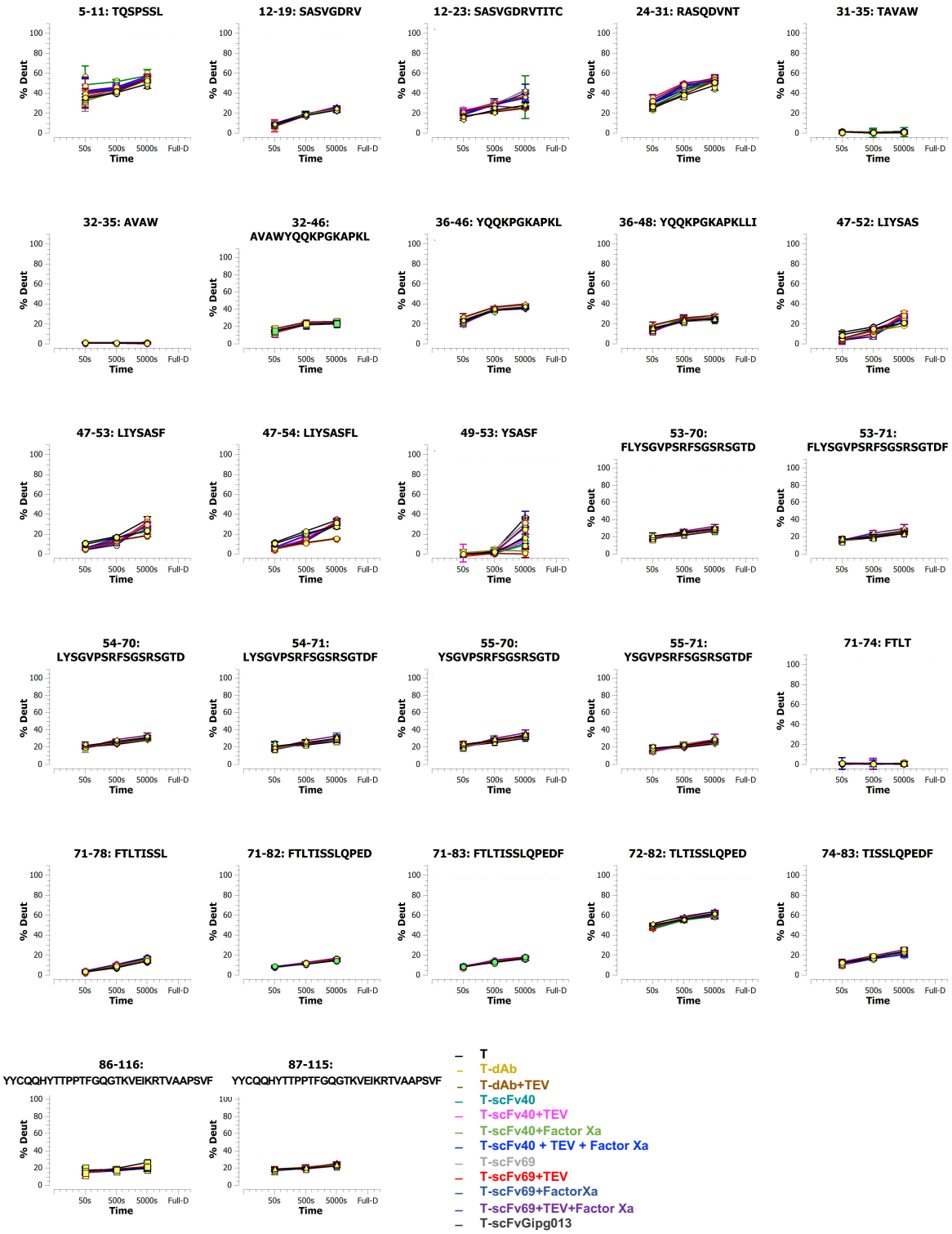


Figure 6.23: HDX-MS uptake plots per peptide mapped on the light chain for the intact and digested masked antibodies (T, T-dAb, T-dAb + TEV, T-scFv40, T-scFv40 + TEV, T-scFv40 + Factor Xa, T-scFv40 + TEV + Factor Xa, T-scFv69, T-scFv69 + TEV, T-scFv69 + Factor Xa, T-scFv69 + TEV + Factor Xa, T-scFvGipg013).

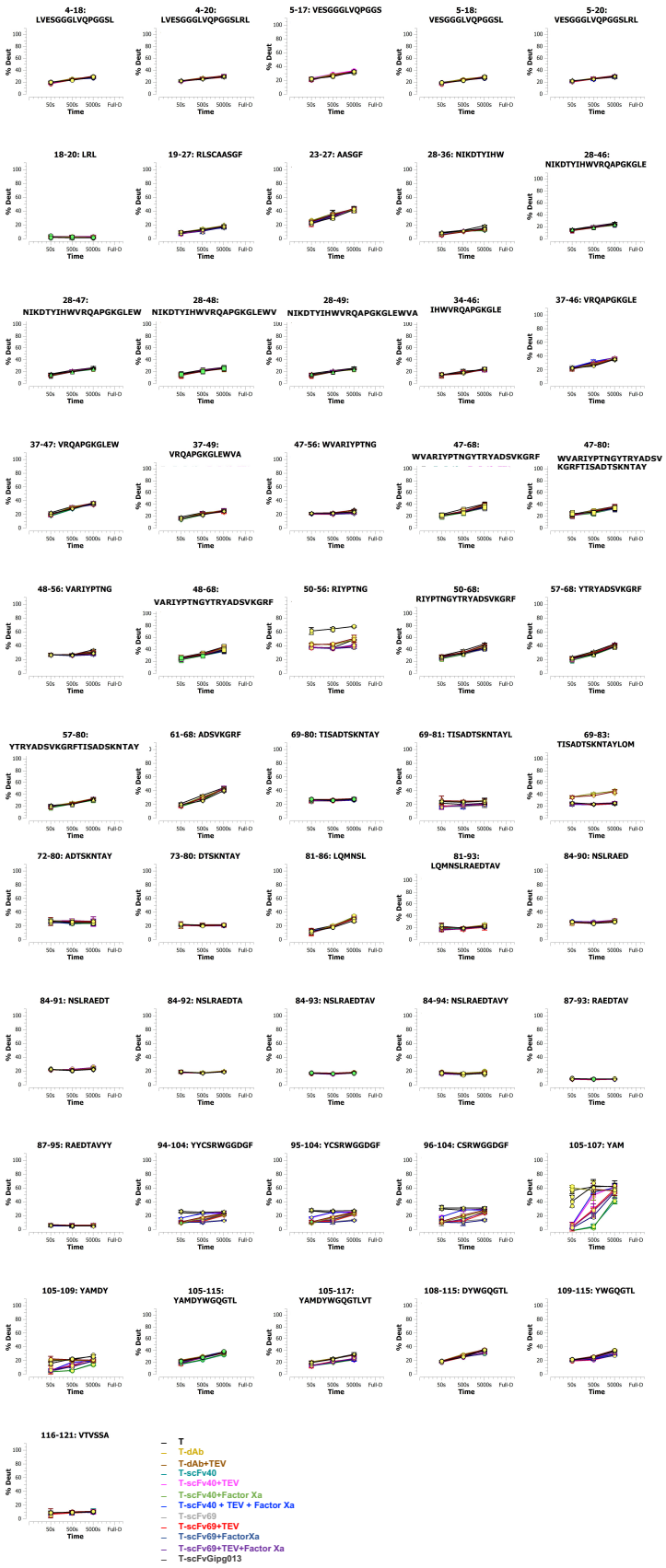


Figure 6.24: HDX-MS uptake plots per peptide mapped on the heavy chain for the intact and digested masked antibodies (T, T-dAb, T-dAb + TEV, T-scFv40, T-scFv40 + TEV, T-scFv40 + Factor Xa, T-scFv40 + TEV + Factor Xa, T-scFv69, T-scFv69 + TEV, T-scFv69 + Factor Xa, T-scFv69 + TEV + Factor Xa, T-scFvGipg013).

The first objective was to identify the binding sites of the masks to trastuzumab when they are covalently linked by probing the differences in the protection of amide protons in the CDRs of trastuzumab. The crystal structure of HER2 in complex with trastuzumab Fab solved by Cho *et al*.[184] showed that HER2 mainly interacts with CDR3 in both the LC and HC (CDRs L3 and H3). The HDX-MS data acquired showed that the highest levels of protection, when the masks are covalently linked to trastuzumab, were observed on CDR L2 and the CDR H3, while the other CDRs did not show a large degree of protection (**Figure 6.25**). More specifically, scFv40 and dAb provided substantial coverage of both CDR L2 and CDR H3 of trastuzumab, while scFv69 primarily covers CDR H3 and only minimally occluded the CDR L2 (**Figure 6.25**). As expected, scFvGipg013, which does not have affinity for trastuzumab, did not provide any protection to any of trastuzumab's CDRs (**Figure 6.25**, **Figure 6.22**). The flow cytometry data showed that scFv69 was less effective than scFv40 in reducing the binding of trastuzumab to HER2, despite having a higher affinity (**Figure 6.12**). The HDX-MS data support the hypothesis that scFv69 covers less well the HER2 paratope of trastuzumab than scFv40 (**Figure 6.25**), and therefore inactivates the antibody less efficiently, although this finding does not exclude the possible role, proposed earlier, of the importance of the kinetic parameters. The HDX-MS results suggest that the binding to both CDR L2 and CDR H3 is more effective in preventing the binding of HER2 to trastuzumab than the coverage of CDR L2 alone, hence explaining why scFv69 is not an ideal mask against HER2, despite its high affinity. This observation highlights the importance of the binding location to generate an effective anti-idiotypic mask.

To elucidate which specific regions of the dAb and trastuzumab contribute to the measured affinities and HDX-MS protection, a crystal structure was determined at 2.35 Å resolution. The model shows a single T-dAb Fab fusion in the asymmetric unit. Examination of the interface reveals the main T-dAb:Fab interactions are located in CDR L2 and CDR H3 on trastuzumab and CDR1 and CDR3 in the dAb. CDR H1, CDR H2, CDR L1 and CDR L3 of trastuzumab are not involved in the interaction (**Figure 6.26 B**). The co-ordinates were used to calculate the buried surface area of the trastuzumab-dAb interface using PISA[185] which is 405.8 $Å^2$ and the complex (without the linker) was predicted to be stable in solution. Therefore, the X-ray structure confirms the observations from the HDX-MS experiments that show that dAb mainly protects the CDR L2 and CDR H3 of trastuzumab from exchange (**Figure 6.25 C**, T-dAb intact).

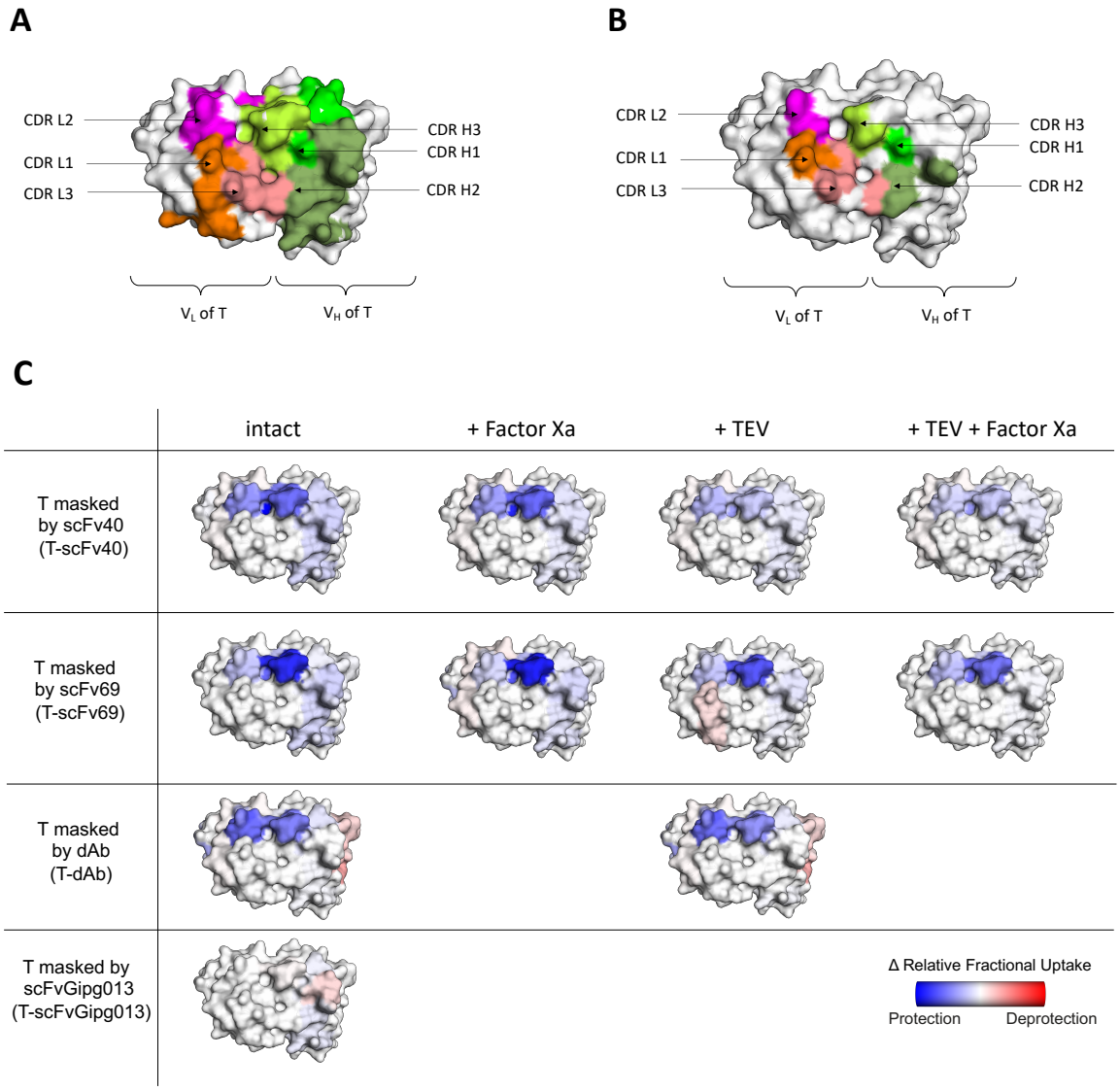


Figure 6.25: Schematic of the results of the HDX-MS experiments superimposed on the structure of trastuzumab. Crystal structure (1N8Z) of the variable domains of trastuzumab observed from the top. Shown in colour and labelled are the CDRs of both the LC and HC showing the possible binding sites. **A.** CDRs of the V_L and V_H of trastuzumab. **B.** Paratope of trastuzumab when interacting with HER2, generated by highlighting the residues which have atoms that are within 4 Å of HER2; the CDRs L3 and H3 are the main contributors to the binding of HER2 according to Cho *et al*., 2003[184]. **C.** Relative change in fractional deuterium exchange represented as a colour scale: reduced exchange (blue), no change (white), increased exchange (red), for trastuzumab interacting with the different masks scFv40, scFv69, dAb and scFvGipg013, with the linkers intact or digested by TEV and/or Factor Xa.

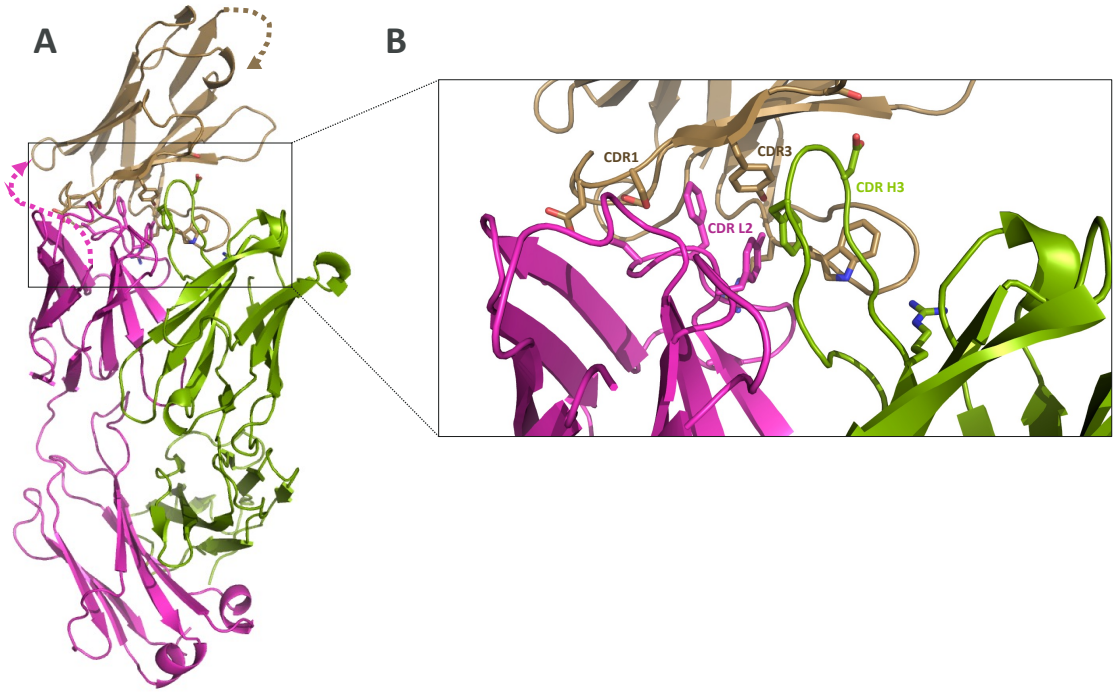


Figure 6.26: Crystal structure of T-dAb showing the interactions between trastuzumab variable Fab fusion and the dAb domain (PDB deposition 7PKL). **A.** Shows the overall fold of the protein. Trastuzumab heavy chain is shown in green, the light chain in magenta and the fusion dAb in wheat. Key side residue:side chains interactions are shown as sticks. The 32-residue linker between the trastuzumab light chain and the dAb was not visible in the electron density (marked with dotted lines). **B.** Close up view of key residues in the CDR. CDR H1, H2, L1 and L3 are not involved in the interaction (not labelled). The bulk of the interaction is derived from CDR L2 and CDR H3 side chain interactions with the dAb CDR1 and CDR3. The X-ray crystallography was generated by Dr David Hargreaves.

6.2.6 Investigation of the deprotection of the CDRs upon cleavage of the linkers

The second objective of the HDX-MS experiments was to determine how the protection levels in trastuzumab and the masked domains changed after digestion, in the absence of HER2. For trastuzumab masked by scFv40, cleaving the linker between the two scFv domains by Factor Xa showed the same level of protection in CDR L2 and CDR H3. The cleavage of the linker between the antibody and the mask by TEV protease, or the cleavage of both linkers by TEV and Factor Xa, resulted in less protection than in the intact construct (**Figure 6.25 C, Figure 6.22**), consistent with the mask at least partially dissociating from trastuzumab. It is worth noting that even after cleavage by TEV and Factor Xa, some protection relative to naked trastuzumab was still observed showing that the mask still interacts transiently with trastuzumab's CDRs, corroborated by the SEC-MALS. These results demonstrate that the covalent linkage of scFv40 to trastuzumab effectively increases the local concentration of the mask which therefore increased the percentage bound, hence its protection.

When trastuzumab is masked by scFv69, the CDR H3 is afforded extensive protection even when the linkers are cleaved by either TEV or Factor Xa. Only the combination of TEV and Factor Xa gave rise to a lower level of protection in CDR H3 (**Figure 6.25 C**). Finally, when trastuzumab is masked by dAb, the digestion did not affect the protection, consistent with other results which strongly suggest that the dAb remains bound to trastuzumab under the conditions used due to its high affinity and very slow dissociation rate.

6.2.7 DISCUSSION

In this study, the properties that impact the efficacy of an anti-idiotypic antibody fragment mask were investigated in order to i) reduce the antibody binding to its antigen when it is not at its target site, and ii) enable effective activation of the antibody upon cleavage of the linker(s). Three masked antibodies were successfully developed based on trastuzumab with covalently linked anti-idiotypic antibody fragments[132,133], and these were shown to be successful at preventing HER2 binding to different degrees. Together, the data acquired from the biological activity assays and the biophysical techniques provide complementary information which greatly aids in our understanding of the parameters that are key in determining the properties of an ideal mask and which are summarized in **Figure 6.27**.

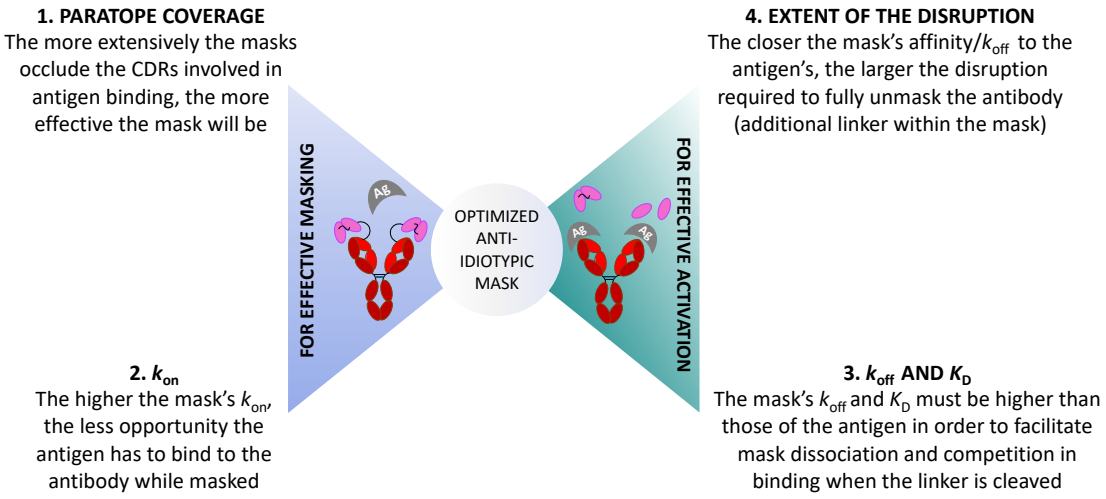


Figure 6.27: Summarized representation of the four key factors contributing interdependently to the optimal design of an anti-idiotypic mask: two define the effective masking of an antibody to avoid on-target off-tumour effects (1 and 2) and two others influence the effective activation of the masked antibody upon proteolytic cleavage (3 and 4).

The results on the intact constructs provides information on how well the masks cover the antigen-binding region of trastuzumab and the location of its binding interface, thus establishing their effectiveness in inactivating the antibody. The flow cytometry and microscopy results showed that dAb is the most effective mask, followed by scFv40 and finally scFv69. As this order does not simply follow that of the relative affinities of the different masks for trastuzumab, other factors must also be at play. One possible factor is purely structural: the HDX-MS results showed that scFv40 and dAb both clearly cover the CDR H3 and the CDR L2, whereas scFv69 primarily protects the CDR H3, explaining the higher binding of T-scFv69 observed in the flow cytometry experiments. From structural studies by Cho *et al.* [184], it is known that HER2 interacts mainly with the CDR3s of the HC and LC. This highlights the fact that, for trastuzumab, steric hinderance of the CDR L3 may be induced by binding of the mask to CDR L2, additionally to the CDR H3. This effectively prevents HER2 from binding to trastuzumab. The other possible factor is kinetic, and not thermodynamic, given that the experiments were not performed at equilibrium: the BLI measurements reveal that transiently dissociated scFv40 tends to reassociate and remask trastuzumab more rapidly than scFv69, indicated by the difference in their k_{on} values. Therefore, a transiently dissociated scFv69 gives HER2 a wider time window in which to capture the unmasked trastuzumab. In contrast, after dissociation, the fickler scFv40 gives HER2 much less time to capture trastuzumab in its unmasked state. The EC_{50} results for the various masks correlate with their k_{on} values, and not with their affinities or k_{off} values, supporting this hypothesis. Therefore, how effectively a covalently-linked mask inactivates an antibody and increases the EC_{50} depends on its binding site and its association kinetics, compared to those of the antigen.

The EC_{50} values show that scFv69 increases the EC_{50} by 11-fold, scFv40 by 31-fold and dAb by an even higher factor (**Figure 6.12 E**), similar to what has been observed in other studies. Desnoyers *et al.*[73] developed a probody, consisting of an EGFR-targeting antibody tethered to a short peptide mask, which increased the EC_{50} by 47-fold. Lu and Chuang *et al.* [87] have developed a pro-Infliximab rheumatoid arthritis therapy for higher selectivity and safety, where the mask reduced the antigen-binding affinity by 395-fold. Chen *et al.*[89], designed a EGFR antibody masked by a latency-associated peptide whose EC_{50} was boosted by 7.5-fold.

In order to be useful as a therapeutic agent, effective masking of the antigen-binding site must be combined with efficient activation of the antibody when it reaches

its target. Thus, the cleavage and subsequent dissociation of the mask is as equally important as efficient initial masking. Our results show that both affinity and dissociation rate are critical in the activation process. Results from the BLI, flow cytometry and microscopy experiments show that for T-scFv40, cleaving the linker between the antibody and the mask is sufficient to activate the antibody, whereas for T-scFv69, the cleavage of both linkers is necessary. This observation agrees with the results from the SEC-MALS and HDX-MS experiments which show that scFv69 interacts more strongly with trastuzumab than scFv40. The fact that scFv40 and scFv69 still transiently bind to trastuzumab after cleavage in the absence of HER2, does not prevent the binding of trastuzumab to HER2 in their presence. This reflects that the activation of the antibody is driven by the association and dissociation kinetics of the mask. Additional evidence of this is seen with dAb, for which cleavage did not lead to activation over any relevant timescale because its k_{off} value is orders of magnitude less than those of the other two masks. In this case, there is no opportunity for competition between the mask and the antigen. The relative affinities of antigen and mask may also influence activation: HER2 has a much stronger affinity than either scFv40 and scFv69, which were both successfully activated. In contrast, masks with affinities close to that of the antigen such as the dAb may not be so useful, because the mask does not dissociate and therefore the antibody is not activated.

Therefore, to facilitate activation of the antibody, i) as the mask needs to come off, its k_{off} can't be too small and needs to be higher than the antigen's, ii) the antigen needs to be able to bind to antibody faster than it rebinds the mask when it has been cleaved and dissociated, so the mask's k_{on} needs to be smaller than that of the antigen. This leads to the fact that the mask has to have a higher K_D and therefore lower affinity compared to the antigen.

6.2.8 CONCLUSION

In this study, the impact of three different anti-idiotypic antibody fragment masks on the HER2-binding activity of trastuzumab were investigated. The results provide valuable information that will guide future design of therapeutic antibodies that can be effectively inactivated by a covalently tethered anti-idiotypic mask, and subsequently efficiently activated by protease cleavage. The work establishes that activity assays, such as flow cytometry, in addition to biophysical characterisation with HDX-

MS, forms a very powerful strategy for assessing the suitability of any potential mask. Moreover, the agreement observed in this study between results from HDX-MS and X-ray crystallography illustrates how, in the absence of a high-resolution structure, HDX-MS can be used as a convenient substitute allowing an initial analysis of whether a mask may be effective, and to characterise the molecular interactions important to stabilise the mask-antibody complex.

Taken together, this work reveals that to successfully engineer an effective anti-idiotypic mask, four major parameters need to be optimized (Figure 6.27). First, the mask must bind to either the same paratope as the antigen or a nearby site that effectively blocks binding of the antigen. Second, the higher the k_{on} of the mask for the antibody, the less opportunity the antigen has to bind to the antibody when it is masked; however, it needs to be smaller than the antigen's k_{on} value to enable it to bind to the antibody when the mask is released. Third, the mask's k_{off} and K_D must not be lower than, nor the same order of magnitude as, those of the antigen to favour the antigen binding by competition with the mask after cleavage. Fourth, higher affinity masks can be employed if the interaction between the mask and the antibody can be disrupted more, through the cleavage of linkers located between the mask and the antibody but also within the mask. It is anticipated that these conclusions can be applied broadly to many different masking strategies.

7 UNFOLDING OF MAB UNDER HDX-MS QUENCH CONDITIONS

7.1 INTRODUCTION

Hydrogen deuterium exchange mass spectrometry is a state-of-the-art technique which probes the molecular dynamics of a native state of a protein at high-resolution. This technique is highly sensitive to the solvent accessibility and hydrogen-bonding of backbone amide groups and has been used successfully to investigate interactions between domains in multi-domain proteins as well as domain stability[104–107]. One of the strengths of this technique relies on the high-resolution data, which informs exactly on the level of deuterium incorporation for each residue of the protein of interest. To obtain this high resolution when using the HDX-MS in a bottom-up workflow (i.e., generating peptides before entering the mass spectrometer, as opposed to top-down, where the peptides are generated by electron transfer dissociation (ETD) fragmentation inside the mass spectrometer), it is necessary to digest the protein, and to reconstitute a peptide map, i.e., locate the digested peptides in the original sequence. Obtaining a good peptide map, the closest to total sequence coverage, and with high redundancy, i.e. having many overlapping peptides in a stacked manner, is ideal as the deuterium exchange can then be measured accurately and located to specific residue. In the HDX-MS workflow, prior to digestion, the deuterated protein is "quenched" by a buffer, which has two functions. On the one hand, it reduces the back exchange of deuterium to hydrogen as much as possible, with low pH[109] and low temperature (4 °C). In addition, it unfolds the protein as much as possible to facilitate the access of the pepsin to the polypeptide chain for the digestion step, either only by the action of the acidic pH, or with addition of a chemical denaturant if the protein is highly stable. If the proteins have disulfide bridges, TCEP is added to the buffer to reduce them. Extensive unfolding of the protein prior to digestion is crucial as the protein cannot be digested properly if it is not denatured. The HDX-MS workflow therefore relies on a compromise between low temperature, which ensures minimal back exchange, but which slows down the rate of unfolding and reduction, and therefore may limit unfolding prior to digestion.

Antibodies are very stable proteins, which do not unfold in acidic conditions at pH 2.5 which therefore necessitates the addition of a chemical denaturant to unfold the protein in the quench buffer. The most commonly used quench buffer for antibodies, before mixing 1:1 with the native protein, includes 0.1 M phosphate, 4 M GdmCl, 0.5 M TCEP at pH ~2.3. However, for some proteins, this buffer does not lead to high peptide coverage and redundancy. A peptide map of NIST mAb IgG obtained by mixing the

protein 1:1 with 100 mM His, 8 M urea (equivalent to approx. 4 M GdmCl in terms of denaturing strength[168]), 0.5 M TCEP, pH 2.3, resulted in rather low peptide coverage (73% for the LC and 66 % for the HC) and redundancy (35 peptides for LC, 27 peptides for HC) were obtained in the Fab domain, compared to the Fc domain (79 peptides, 95% coverage), **Figure 7.1**.

Figure 7.1: Peptide map of NIST mAb IgG. A. Light chain**. B.** Heavy chain. The quench buffer was made of 100 mM His, 8 M urea and 0.5 M TCEP , pH 2.3 (before mixing), and no quench hold was used (waiting time after the equilibrated/labelled protein and the quench are mixed at 4 °C to leave more time to the unfolding and the reduction, before the mix is injected onto the pepsin column). The experiments were run on a Synapt G2-Si (Waters), and analyzed with PLGS.

The biophysical studies on NIST mAb IgG described in Chapter 3 revealed a remarkably high stability of the Fab domain, thermodynamically ($\Delta \boldsymbol{G}_{\boldsymbol{D-N}}^{\boldsymbol{H2O}}$= 11 kcal mol^{-1}), thermally (T_m = 92.7 °C), and kinetically (half-life of unfolding in water = 79.6 years). These data showed the Fab fragment has a much higher stability than that of the C_H3 or C_H2 domains. These observations raise the question of whether the poor peptide coverage obtained for the Fab is related to the very high stability of the Fab, i.e. that the Fab does not have the time to unfold enough in the quench buffer to be efficiently digested by the pepsin. While studies have been conducted to optimize the digestion conditions[186,187], by packing pepsin on a column, using high pressure (~ 10,000-15,000 psi) in the trapping phase with the Enzymate BEH pepsin column[188,189], which will help unfold the proteins by shear stress[190], optimization of the unfolding of the protein in the quench buffer hasn't

been investigated to the best of our knowledge. The objective of this Chapter is to understand the impact of quench buffer composition, more specifically the denaturant concentration, on the unfolding of the proteins in the HDX-MS workflow. The insights obtained for Fab of NIST mAb can be applied to other very highly stable proteins.

7.2 RESULTS

7.2.1 Unfolding kinetics of NIST mAb IgG in 2 M GdmCl, pH 2.5

To understand if the rather low coverage obtained for the Fab domain of NIST mAb is due to insufficient unfolding, the unfolding kinetics of NIST mAb full-length IgG, Fab and Fc were measured with a stopped-flow spectrophotometer. The Fab and Fc domains were generated by digestion by FabRICATOR which cuts below the hinge (CPAPELLG/GPSVF) and purified by affinity chromatography with a CH1-XL column which binds to the Fab fragment only. The samples obtained were therefore two-armed Fab with a hinge and the Fc domains without a hinge. For one experiment, NIST mAb Fc with hinge, generated by colleagues in AstraZeneca, was used. The effect of the chemical denaturant, TCEP and D_2O were probed separately.

First, the impact of the concentration of chemical denaturant was investigated. To replicate the mixing conditions taking place when the equilibrated or labelled samples are quenched, native protein in 10 mM potassium phosphate buffer pH 7.0 was rapidly mixed 1:1 with 100 mM potassium phosphate 4 M GdmCl, pH ~2.3, so that the resulting solution had a pH of 2.5 (at 4 °C) and contains 2 M GdmCl. From now on, only the final concentration after 1:1 mixing will be reported. The unfolding kinetics were measured from 25 °C down to 5 °C, to reliably extrapolate the unfolding rates constants between 0 and 4 °C, the temperature range usually used for the quenching step. The unfolding kinetic data on NIST mAb IgG discussed in Chapter 3 showed that three unfolding rates were observed, corresponding to unfolding of the C_H2, the C_H3 and Fab domains (fastest to slowest).

The results from the unfolding kinetics of NIST mAb Fc at 2 M GdmCl, pH 2.5 was best fit to a double exponential (**Equation 2.10**), revealing two unfolding transitions, the fastest most likely corresponding to the unfolding of the C_H2 domain, given it is the least stable domain, and the second to unfolding of the C_H3 domain (**Figure 7.2**). From the fitted parameters, the half-lives of unfolding of the C_H2 and C_H3 domains under these

quench unfolding conditions (extrapolated to 0 °C) were 8 and 30 seconds, respectively ($t_{1/2\ 0\ °C}$, **Table 7.1**) whilst under the same conditions at 25 °C, they were 0.3 and 1.4 seconds for unfolding of the C_H2 and C_H3 domains, respectively ($t_{1/2\ 25\ °C}$, **Table 7.1**). It is interesting to note, that as expected the antibody is less stable and unfolds faster at pH 2.5 compared to pH 5.5 at 25 °C, under which conditions the half-lives are 0.7 and 1906 seconds for unfolding of the C_H2 and C_H3 domains, respectively (extrapolated from the NIST mAb Fc, **Table 3.4**, using **Equation 2.14**). These results also show that the C_H3 is destabilized considerably more than the C_H2 domain on lowering the pH to 2.5.

Surprisingly, the unfolding kinetics of NIST mAb full-length IgG also showed the same two unfolding transitions, and no slower unfolding phase was observed over the 15-minute longest time frame used.

Table 7.1: Kinetic parameters for the unfolding of the different domains of mAb IgG in 2M GdmCl, at pH 2.5

	m_T (°C^{-1})	ln k_U 2M GdmCl 0 °C	$t_{1/2}$ at 0 °C (s)	$t_{1/2}$ at 25°C (s)
ln($k_{1\ CH2}$) NIST mAb IgG in H_2O	0.129 ± 0.005	-2.46 ± 0.08	8.09 ± 0.05	0.32 ± 0.05
ln($k_{2\ CH3}$) NIST mAb IgG in H_2O	0.124 ± 0.004	-3.78 ± 0.06	30.40 ± 0.04	1.37 ± 0.04
ln ($k_{3\ Fab}$) NIST mAb Fab in H_2O	0.16 ± 0.03	-11.1 ± 0.5	45789.0 ± 0.4	796.5 ± 0.4
ln(k_0) NIST mAb IgG in D_2O	0.12 ± 0.04	2.3 ± 0.3	0.1 ± 0.2	0.004 ± 0.2
ln($k_{1\ CH2}$) NIST mAb IgG in D_2O	0.121 ± 0.005	-2.36 ± 0.07	7.32 ± 0.05	0.36 ± 0.05
ln($k_{2\ CH3}$) NIST mAb IgG in D_2O	0.126 ± 0.004	-3.67 ± 0.07	27.12 ± 0.05	1.16 ± 0.05
ln($k_{1\ CH2}$) NIST mAb Fc (with hinge)	0.131 ± 0.003	-2.51 ± 0.04	8.49 ± 0.03	0.33 ± 0.03
ln($k_{2\ CH3}$) NIST mAb Fc (with hinge)	0.113 ± 0.002	-3.83 ± 0.03	32.04 ± 0.02	1.92 ± 0.02
ln($k_{1\ CH2}$) NIST mAb Fc (without hinge)	0.102 ± 0.005	-1.84 ± 0.07	4.36 ± 0.05	0.34 ± 0.05
ln($k_{2\ CH3}$) NIST mAb Fc (without hinge)	0.088 ± 0.005	-3.35 ± 0.07	19.84 ± 0.05	2.22 ± 0.05
ln($k_{1\ CH2}$) NIST mAb Fc (without hinge) + TCEP	0.15 ± 0.01	-2.4 ± 0.2	7.4 ± 0.1	0.2 ± 0.1
ln($k_{2\ CH3}$) NIST mAb Fc (without hinge) + TCEP	0.145 ± 0.008	-3.6 ± 0.1	26.66 ± 0.08	0.72 ± 0.08

The errors of m_{kU} and lnk_U are fitting errors. The errors of the half-lives were calculated as compound errors. The natural logarithm of the unfolding rate constants was fitted to: $\ln k_U = \ln k_{U\ at\ 2M\ GdmCl}^{0\ °C} + m_T \times temperature$.

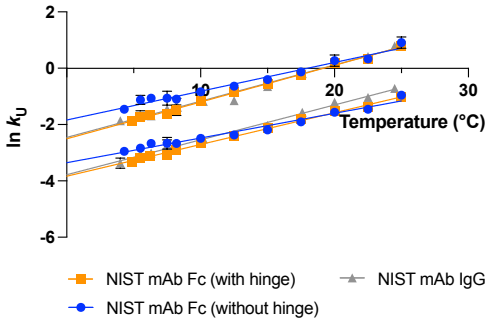


Figure 7.2: Unfolding kinetics of NIST mAb IgG and NIST mAb Fc with and without hinge, in 2 M GdmCl, 55 mM phosphate, pH 2.5 final, over a range of temperatures. Measurements were performed using a stopped-flow spectrometer at a final protein concentration of 0.25 μM.

Consistent with the observations on NIST mAb IgG, the unfolding kinetics of NIST mAb Fab in 2 M GdmCl pH 2.5 did not generate any change in fluorescence signal over the time scales measured by the stopped-flow (max. 15 minutes). This suggests that these conditions maybe insufficient to unfold the Fab domain quickly enough. In order to better understand the stability of the antibody at low pH, a chemical denaturation curve of NIST mAb Fab was recorded at 25 °C, pH 2.5 in the final conditions of the quench unfolding mix. The fluorescence signal obtained from the excitation of tryptophans at 280 nm, processed by average emission wavelength (**Equation 2.3**) revealed a flat line (**Figure 7.3**). This suggests that, either the Fab domain is quite sensitive to acidic conditions and has already unfolded at pH 2.5 in the absence of chemical denaturant, or that incubation at pH 2.5 and up to 2 M GdmCl at 25 °C are insufficient to unfold it. To investigate this further, a far UV-CD spectrum of NIST mAb Fab at pH 2.5 was recorded and compared to NIST mAb Fab at pH 7.0. The results showed that NIST mAb Fab still has native-like secondary structure at pH 2.5 compared to pH 7.0 (**Figure 7.4**). These data indicate that the midpoint of denaturation of NIST mAb Fab is considerably higher than 2 M GdmCl at pH 2.5, consistent with the very high stability observed in Chapter 3 at pH 5.5 ($[den]_{50\%\ D\text{-}N\ NIST\ mAb\ Fab}$ = 3.2 M GdmCl in 20 mM His pH 5.5, **Table 3.2**), and that at 2 M GdmCl, the protein did not unfold in the stopped-flow experiments.

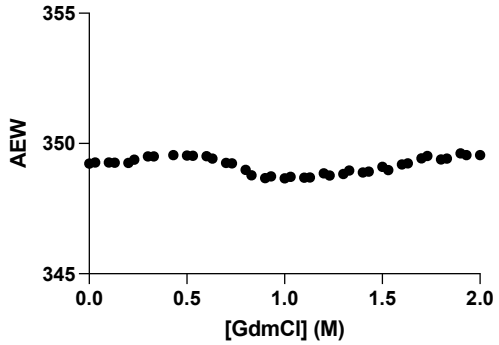


Figure 7.3: Chemical denaturation curve of NIST mAb Fab at pH 2.5. Protein concentration was 0.5 µM at 25 °C.

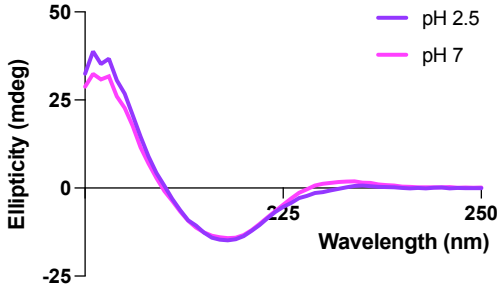


Figure 7.4: Far UV-CD spectra of NIST mAb Fab at pH 2.5 and pH 7.0. The baseline recorded with the buffers only was substracted. The proteins were measured at 5 µM at 25 °C.

The unfolding kinetics of NIST mAb Fab in 2 M GdmCl was measured on a fluorimeter using manual mixing and much longer timescales (3- 12 hours). The intensity of the fluorescence emission was measured at 335, 345 and 360 nm and the experiment was repeated at different temperatures. A very slow unfolding phase was observed (data were fit to a single exponential, **Equation 2.9** to obtain unfolding rate constants). Using an extrapolation, these data show that the unfolding of the Fab domains has a half-life of approx. 13 hours at 0 °C ($t_{1/2\ 0\ °C}$, **Table 7.1**, **Figure 7.6 A**), consistent with the fact that no signal is observed for the unfolding of the Fab domains in the stopped-flow experiment.

7.2.2 Influence of D_2O and TCEP on the unfolding kinetics

The influence of D_2O and TCEP on the unfolding kinetics were also measured, to assess if they made any difference to the rates observed. The unfolding of NIST mAb Fc was carried out starting with the protein in D_2O, as it would be when it is labelled in the HDX experiment, and this was mixed 1:1 with the aqueous (H_2O) quench buffer. The results showed unfolding rate constants and fittings within error of those measured in H_2O, thus results from experiments measured in H_2O can be used as a reliable measure of those that would be observed in D_2O (**Figure 7.5 A**, **Table 7.1**).

Similarly, the unfolding kinetics of NIST mAb Fc in the presence of 0.25 M TCEP showed two unfolding phases similar to those observed without TCEP (**Figure 7.5**, **Table 7.1**). This confirms that the results obtained from measuring the unfolding rates without reducing agent can be used as good guides for the unfolding rates in the presence of TCEP. Ideally, the experiments would have been repeated for the Fab domain, in case the reduction of the disulfide bridges between the heavy and light chain affected the cooperative unfolding of this fragment, however, there was insufficient time to complete these experiments.

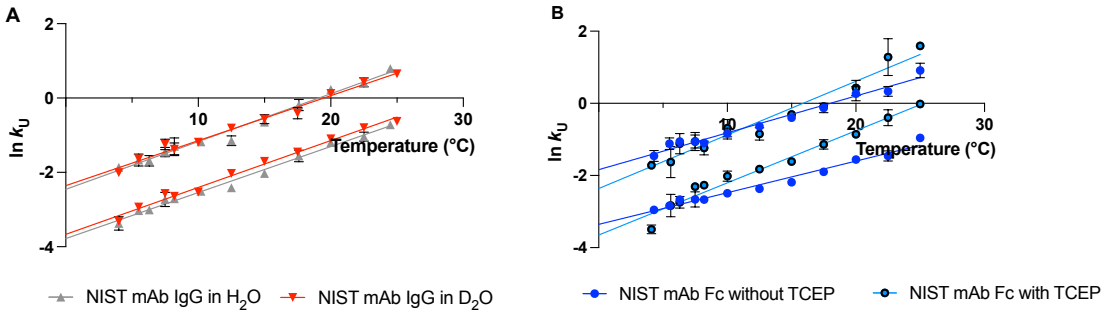


Figure 7.5: Unfolding kinetics in 2 M GdmCl, over a range of temperatures. **A.** NIST mAb IgG in H_2O and D_2O. **B.** NIST mAb Fc (without hinge), in the presence and absence of 0.25 M TCEP. All proteins were measured at a final concentration of 0.25 µM.

Therefore, collectively these data confirm that the NIST mAb Fab unfolds too slowly in the conditions investigated to be measured by stopped-flow. This indicates that the NIST mAb Fab reaches the pepsin in a folded state after it is quenched with a buffer composed of 100 mM phosphate, 4 M GdmCl, 0.5 M TCEP, and explains why the coverage of the Fab is much poorer than that of the Fc fragment, which unfolds rapidly in the quench buffer.

7.2.3 Measurement of the unfolding rate constants at higher denaturant concentrations

In order to estimate how much faster the unfolding of NIST mAb Fab becomes at higher concentrations of chemical denaturants, thereby establishing the conditions required to fully unfold the Fab fragment before the pepsin column in the HDX experiments, the unfolding kinetics of NIST mAb Fab and Fc were measured in 2.5, 3, 3.5, and 4 M GdmCl, pH 2.5.

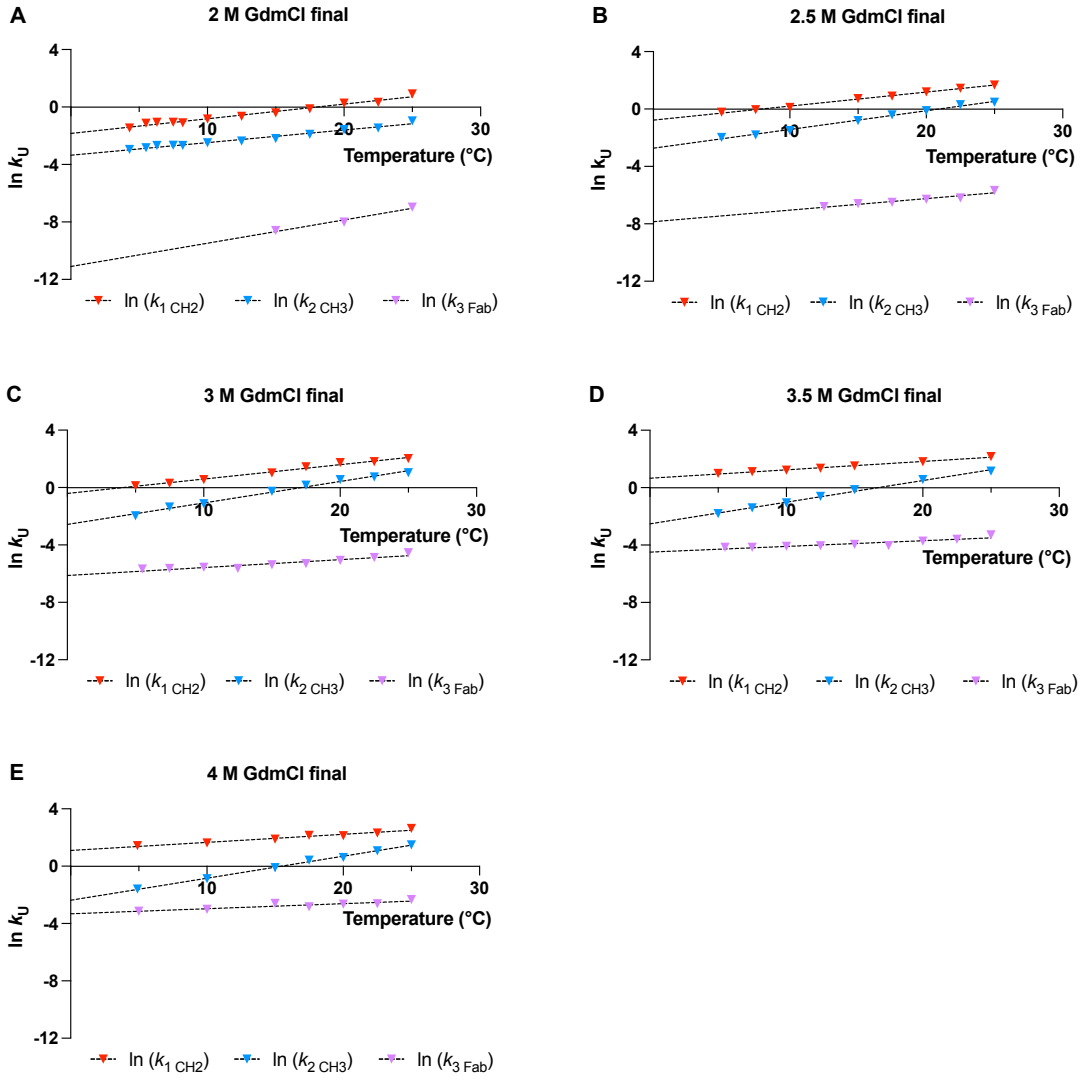


Figure 7.6: Temperature dependence of the unfolding kinetics of NIST mAb Fc and NIST mAb Fab at different final concentrations of GdmCl at pH 2.5. The slowest phase corresponds to unfolding of the Fab fragment, the second slowest to the unfolding of the C_H3 domain, and the fastest to the unfolding of the C_H2 domain. These data were obtained mixing 1:1 native protein in 10 mM phosphate pH 7, with **A.** 4 M GdmCl pH 2.3 (2 M final, pH 2.5). **B.** 5 M GdmCl pH 2.3 (2.5 M final, pH 2.5). **C.** 6 M GdmCl pH 2.3 (3 M GdmCl final, pH 2.5). **D.** 7 M GdmCl pH 2.3 (3.5 M GdmCl final, pH 2.5). **E.** 8 M GdmCl pH 2.3 (4 M GdmCl final, pH 2.5).

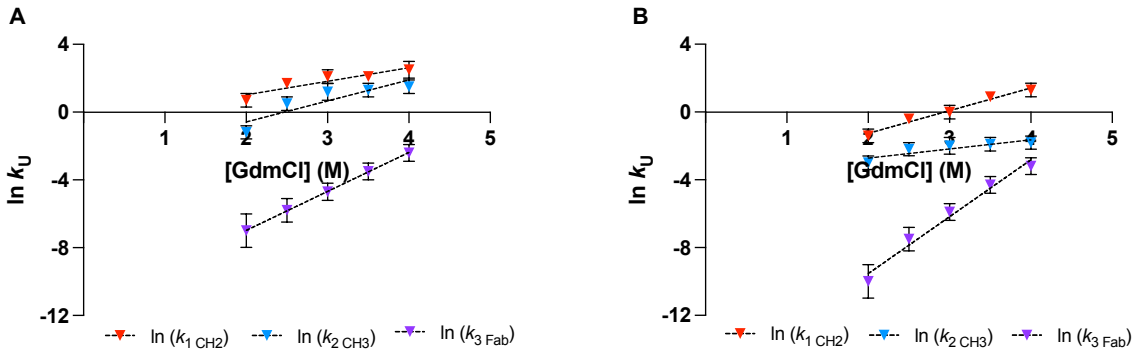


Figure 7.7: Unfolding of NIST mAb Fc and NIST mAb Fab at pH 2.5 over different concentrations of guanidinium chloride at **A.** 25 °C **B.** 4 °C. The values were calculated from the fitting of the lnk_U *versus* temperature reported in **Table 7.2**, the results obtained are shown in **Table 7.3**.

Table 7.2: Kinetic parameters for the unfolding of Fc and Fab domains of NIST mAb at higher concentrations of GdmCl at pH 2.5

[GdmCl] in quench buffer (M)	Final [GdmCl] (M)		m_T (°C^{-1})	ln (k_U) 0 °C	$t_{1/2}$ at 0 °C (s)	$t_{1/2}$ at 4 °C (s)	$t_{1/2}$ at 10 °C (s)
4	2	ln ($k_{1\ CH2}$)	0.102 ± 0.005	-1.84 ± 0.07	4.36 ± 0.05	2.90 ± 0.05	1.57 ± 0.06
		ln ($k_{2\ CH3}$)	0.088 ± 0.005	-3.35 ± 0.07	19.84 ± 0.05	13.97 ± 0.05	8.26 ± 0.06
		ln ($k_{3\ Fab}$)	0.16 ± 0.03	-11.1 ± 0.5	45789.0 ± 0.4	23945.5 ± 0.4	9055.6 ± 0.4
5	2.5	ln ($k_{1\ CH2}$)	0.098 ± 0.003	-0.77 ± 0.04	1.50 ± 0.03	1.02 ± 0.03	0.57 ± 0.03
		ln ($k_{2\ CH3}$)	0.130 ± 0.004	-2.73 ± 0.06	10.59 ± 0.04	6.29 ± 0.05	2.88 ± 0.05
		ln ($k_{3\ Fab}$)	0.08 ± 0.01	-7.9 ± 0.2	1775.9 ± 0.2	1287.2 ± 0.2	794.3 ± 0.2
6	3	ln ($k_{1\ CH2}$)	0.100 ± 0.005	-0.41 ± 0.08	1.04 ± 0.06	0.70 ± 0.06	0.38 ± 0.06
		ln ($k_{2\ CH3}$)	0.150 ± 0.006	-2.6 ± 0.1	9.05 ± 0.07	4.97 ± 0.07	2.03 ± 0.08
		ln ($k_{3\ Fab}$)	0.055 ± 0.007	-6.1 ± 0.1	317.09 ± 0.08	254.33 ± 0.09	182.7 ± 0.1
7	3.5	ln ($k_{1\ CH2}$)	0.059 ± 0.002	0.66 ± 0.03	0.36 ± 0.02	0.28 ± 0.02	0.20 ± 0.02
		ln ($k_{2\ CH3}$)	0.151 ± 0.005	-2.52 ± 0.07	8.62 ± 0.05	4.71 ± 0.05	1.91 ± 0.06
		ln ($k_{3\ Fab}$)	0.040 ± 0.007	-4.45 ± 0.1	61.76 ± 0.08	52.70 ± 0.08	41.53 ± 0.09
8	4	ln ($k_{1\ CH2}$)	0.056 ± 0.005	1.10 ± 0.09	0.2 ± 0.1	0.19 ± 0.07	0.13 ± 0.07
		ln ($k_{2\ CH3}$)	0.154 ± 0.004	-2.37 ± 0.07	7.44 ± 0.05	4.03 ± 0.05	1.60 ± 0.05
		ln ($k_{3\ Fab}$)	0.035 ± 0.007	-3.3 ± 0.1	19.2 ± 0.1	16.64 ± 0.09	13.5 ± 0.1

The errors of m_{kU} and lnk_U are fitting errors. The errors of the half-lives were calculated as compound errors. The natural logarithm of the unfolding rate constants was fitted to: $\ln k_U = \ln k_{U\ at\ 4M\ GdmCl}^{0\ °C} + m_T \times temperature$.

Employing higher concentrations of denaturant enables the unfolding of GdmCl to occur within tens of seconds ($t_{1/2\ Fab\ 8M\ 0\ °C}$ = 19 s; $t_{1/2\ Fab\ 7M\ 0\ °C}$ = 62 s), which becomes compatible with the timelines used in the HDX-MS workflow (**Figure 7.6**, **Figure 7.7**). **Figure 7.6** shows that unfolding of the C_H3 domain is the most sensitive to temperature, whereas the unfolding of the C_H2 and Fab domains slow down to a lesser extent as the temperature decreases. **Figure 7.7** reveals that the sensitivity of the unfolding rates of the different domains and fragments to denaturant concentration varies and is also temperature dependent. The Fab domain is the most sensitive to chemical denaturant concentration, which agrees with the results shown in Chapter 3.

Table 7.3: Unfolding kinetic parameters from the fitting shown in Table 7.2

[GdmCl] in quench buffer (M)	Final [GdmCl] (M)		ln k_U at 4 °C	ln k_U at 25 °C
4	2	ln ($k_{1\ CH2}$)	-1.4 ± 0.4	0.7 ± 0.4
		ln ($k_{2\ CH3}$)	-3.0 ± 0.4	1.2 ± 0.4
		ln ($k_{3\ Fab}$)	-10 ± 1	-7 ± 1
5	2.5	ln ($k_{1\ CH2}$)	-0.4 ± 0.3	1.7 ± 0.3
		ln ($k_{2\ CH3}$)	-2.2 ± 0.4	0.5 ± 0.4
		ln ($k_{3\ Fab}$)	-7.5 ± 0.7	-5.8 ± 0.7
6	3	ln ($k_{1\ CH2}$)	0.0 ± 0.4	2.1 ± 0.4
		ln ($k_{2\ CH3}$)	-2.0 ± 0.5	1.2 ± 0.5
		ln ($k_{3\ Fab}$)	-5.9 ± 0.5	-4.7 ± 0.5
7	3.5	ln ($k_{1\ CH2}$)	0.9 ± 0.2	2.1 ± 0.2
		ln ($k_{2\ CH3}$)	-1.9 ± 0.4	1.3 ± 0.4
		ln ($k_{3\ Fab}$)	-4.3 ± 0.5	-3.5 ± 0.5
8	4	ln ($k_{1\ CH2}$)	1.3 ± 0.4	2.5 ± 0.5
		ln ($k_{2\ CH3}$)	-1.8 ± 0.4	1.5 ± 0.4
		ln ($k_{3\ Fab}$)	-3.5 ± 0.5	-2.4 ± 0.5

These measurements at different denaturant concentrations and temperatures are very informative with respect to the optimal conditions required to unfold the Fab fragment of NIST mAb on a timescale compatible with the HDX workflow. However, the results establish that higher concentrations of denaturant than are currently used are required, and therefore it is now essential to verify the resistance of the pepsin to this chemical denaturant stress.

7.2.4 Measurement of the stability of pepsin

To assess the stability of the pepsin, the Enzymate BEH pepsin column from Waters was taken as a reference and the stability of its pepsin[188] (porcine gastric mucosa, part #P7012) was investigated. Unfolding and refolding curves were measured in guanidinium chloride (GdmCl) and guanidinium thiocyanate (GdmSCN), a chemical denaturant approximately twice as strong as GdmCl[98–100]. The experiments were carried out in 55 mM phosphate, pH 2.5 at 20 °C, which is the temperature at which the digestion by the pepsin occurs. Results showed that unfolding are not reversible, and surprisingly showed two trends. All curves were measured after 24 hours, and also after a few more days (6 days for GdmCl, 2 days for GdmSCN), but the denaturation midpoints did not change over time (**Table 7.4**). In guanidinium chloride, the denaturation midpoint of the

unfolding curve was higher than that of the refolding curve, whereas in guanidinium thiocyanate, the denaturation midpoint of the unfolding curve was lower than that of the refolding curve.

In GdmSCN, the unfolding seems to occur via an intermediate (follows a three-state model), although it is not very populated so the curve was fit to a two-state model, whereas the refolding curve showed a single transition. In GdmCl, the unfolding curve also seems to show two transitions, although the denatured baseline was not reached, and the refolding seems to be a single transition, although it is not very clear (**Figure 7.8**).

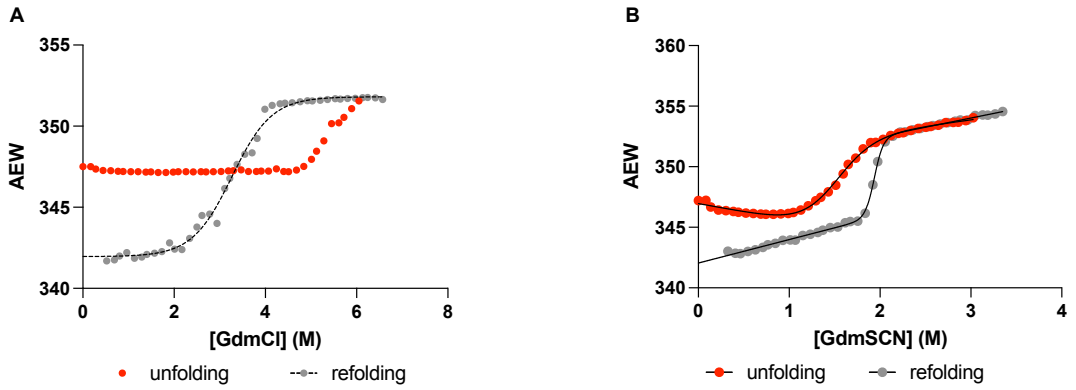


Figure 7.8: Unfolding and refolding denaturation curves of pepsin A. In GdmCl. **B.** In GdmSCN. The experiments were carried out in 55 mM phosphate pH 2.5, incubated at 20 °C for 48 hours. The concentrations of pepsin were 2.25 and 1.7 µM for unfolding and refolding curves, respectively.

Table 7.4: Thermodynamic parameters from the fitting of the unfolding and refolding curves of pepsin in GdmCl and GdmSCN

		$m_{D\text{-}N}$ (M^{-1})	$[den]_{50\%\ D\text{-}N}$ (M)
GdmCl	refolding - 48h	11.9 ± 0.5	1.938 ± 0.003
	refolding - 72h	11.9 ± 0.5	1.940 ± 0.003
GdmSCN	unfolding - 48h	2.9 ± 0.2	1.49 ± 0.02
	unfolding - 72h	2.3 ± 0.2	1.51 ± 0.02
	refolding - 48h	11.9 ± 0.5	1.938 ± 0.003
	refolding - 72h	11.9 ± 0.5	1.940 ± 0.003

The errors are fitting errors.

Taking into account only the unfolding data, it seems that the pepsin is very

stable, with the first denaturation midpoint around 5.25 M GdmCl. This would suggest that using a quench buffer with a final concentration of 4 M GdmCl, should not unfold the pepsin. Guanidinium thiocyanate is approximately twice as strong as guanidinium chloride, but the midpoint of denaturation seems even lower than expected.

7.3 DISCUSSION AND CONCLUSION

The HDX MS workflow relies on a compromise between low back exchange after the labelling step (ensured by low pH and low temperatures), and maximum peptide map coverage (ensured by the unfolding and therefore effective digestion of the labelled protein, which is favored by high chemical denaturant concentrations, high temperature and low pH). For many proteins studied by HDX-MS, like the Fc fragment of NIST mAb, the conditions most widely employed which use 2 M GdmCl, are suitable to unfold the protein sufficiently rapidly to obtain a very good peptide map, higher than 85%. However, for some very stable proteins, such as the Fab fragment of NIST mAb investigated here, the unfolding is slow under these conditions and therefore insufficient to obtain good peptide coverage.

The preliminary results presented in this Chapter provide information on the two avenues that could be explored further which would favor faster unfolding or more complete unfolding. First, both a higher temperature and a higher chemical denaturant concentration would increase the rate of unfolding. For example, the investigation of the unfolding of NIST mAb Fab, which is a very stable domain (see Chapter 3), revealed that the unfolding becomes much more rapid by an increase in chemical denaturant concentration of 1 M, e.g., between 2 and 3 M GdmCl the unfolding half-life of the Fab decreases from 12.7 hours to less than an hour approx. 250 s at 4 °C, **Table 7.2**. Second, it is also possible to hold the quenching step for a longer time (usually it is up to 5 minutes), which would provide more time for more of the protein molecules to unfold, but which needs to be weighed against more deuterium/hydrogen back exchange. Some investigation on this point was carried out by Waters and presented in an application note[189]. The unfolding of NIST mAb Fab seems to be less sensitive to temperature than to chemical denaturant concentration (half-lives in **Table 7.2**), and therefore it might not be ideal to compromise on the back exchange for slightly quicker unfolding. The impact of the chemical denaturant concentration is much stronger than the impact of the

temperature on the unfolding for NIST mAb Fab.

Employing higher concentrations of chemicals denaturant, however, needs to take into account the stability of the pepsin, more specifically its denaturation midpoint in the conditions used in HDX-MS, because if the final chemical denaturant concentration also unfolds the pepsin, the peptide coverage would still be poor despite very rapid unfolding of the protein of interest. The investigation of the thermodynamic stability of pepsin in guanidinium chloride and thiocyanate, however, gave somewhat surprising results which should be repeated to establish reproducibility. However, the preliminary results suggest that it might be possible for the pepsin to withstand 3 M GdmCl rather than the 2 M typically used currently.

An extensive study on the Fab domain of NIST mAb was conducted by Hudgens *et al.,* 2019[191] where they compared the HDX-MS results obtained from fifteen different laboratories. Although a kit with all the buffers was sent, slightly different quench buffers were used, as reported in **Table 7.5**. The quench buffers contained from 4 to 8 M GdmCl, which were diluted, in most cases, 1:1 with the labelled protein and therefore the concentration was halved when the mixed solution reached the pepsin column. It is not possible to draw any straightforward correlations between the denaturant concentration and the peptide coverage and redundancy in **Table 7.5** as many parameters were changed from lab-to-lab, e.g. the digestion temperature and duration of the quench step, the mass spectrometers and the methods used (for example on Synapt instruments, activating the ion-mobility separation will enable better resolution between the different peptides eluting at the same time from the liquid chromatography and generate a richer peptide map). However, the lowest peptide coverages and number of peptides (for labs 4, 7, 8, 9, 13, 14 and 15, **Table 7.5**), were similar to the results observed in our experiments (**Figure 7.1**). In contrast, the use of higher denaturant concentration led to very high coverage and redundancy, which confirms our findings.

Table 7.5: Quench buffers, methods used and peptide coverage obtained for the HDX-MS of NIST mAb Fab, from Hudgens *et al*., *Anal. Chem.*, 2019[191]

Lab	Quench buffer composition	Labelled protein: quench ratio	Digestion tempe-rature	Digestion time	Protease	Peptide coverage	Number of peptides	Mass Spectro-meter
1	7.04 M GdmCl 352 mM phosphate 0.5 M TCEP pH 2.5	1:3	10 °C	*not reported*	Poroszyme™ Immobilized Pepsin Cartridge (Thermo Scientific)	HC: 99% LC: 98%	HC: 76 LC: 52	Synapt G2 (Waters)
2	8 M GdmCl 400 mM phosphate 0.5 M TCEP pH 2.5	1:1	20 °C	240 sec	Enzymate BEH Pepsin (Waters)	HC: 81% LC: 83%	HC: 87 LC: 53	Synapt G2 High Definition (Waters)
3	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	1 °C	180 sec	Enzymate BEH Pepsin (Waters)	HC: 98% LC: 93%	HC: 50 LC: 39	Synapt G2-Si (Waters)
4	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	20 °C	180 sec	Enzymate BEH Pepsin (Waters)	HC: 67% LC: 75%	HC: 24 LC: 22	Synapt G2-Si (Waters)
5	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	18 °C	120 sec	IDEX guard column (2mm x 2cm) packed with immobilized pepsin beads (#20343, Thermo)	HC: 91% LC: 60%	HC: 97 LC: 95	Synapt G2-S (Waters)
6	6.4 M GdmCl 0 mM phosphate 1 M TCEP pH 2.5	6:9	1.5 °C	48 sec	pepsin column prepared in lab with aldehyde linkers	HC: 97% LC: 95%	HC: 89 LC: 42	Orbitrap Elite (Thermo Scientific)
7	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	0 °C	180 sec	pepsin column prepared in lab	HC: 81% LC: 71%	HC: 48 LC: 43	6530 Quadrupole / ToF (Agilent)
8	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	20 °C	18 sec	Poroszyme™ Immobilized Pepsin Cartridge (Thermo Scientific)	HC: 77% LC: 87%	HC: 76 LC: 50	Synapt G2 (Waters)
9	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	room tempe-rature	180 sec	Enzymate BEH Pepsin (Waters)	HC: 74% LC: 64%	HC: 37 LC: 30	Orbitrap Elite (Thermo Scientific)
10	6.6 M GdmCl 330 mM phosphate 1 M TCEP pH 2.5	6:25	3 °C	30 sec	pepsin column prepared in lab	HC: 92% LC: 93%	HC: 92 LC: 77	Orbitrap Q-Exactive (Thermo Scientific)
11	7.1 M GdmCl 350 mM phosphate 0.45 M TCEP pH 2.5	1:1	15 °C	180 sec	Poroszyme™ Immobilized Pepsin Cartridge (Thermo Scientific)	HC: 75% LC: 93%	HC: 54 LC: 41	Synapt G2-S (Waters)
12	8 M urea 100 mM phopshate 1 M TCEP pH 2.5	1:1	11 °C	5 sec	Pepsin + Type XIII protease from Aspergillus (NovaBioassay)	HC: 81% LC: 91%	HC: 83 LC: 92	Synapt G2-Si (Waters)
13	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	0 °C	130 sec	Poroszyme™ Immobilized Pepsin Cartridge (Thermo Scientific)	HC: 72% LC: 69%	HC: 48 LC: 53	Orbitrap Elite (Thermo Scientific)
14	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	15 °C	150 sec	Enzymate BEH Pepsin (Waters)	HC: 67% LC: 67%	HC: 41 LC: 29	Orbitrap Q-Exactive (Thermo Scientific)
15	4 M GdmCl 200 mM phosphate 0.5 M TCEP pH 2.5	1:1	8 °C	180 sec	Enzymate BEH Pepsin (Waters)	HC: 56% LC: 77%	HC: 19 LC: 59	Orbitrap Elite (Thermo Scientific)

Other approaches to increase the peptide coverage has included top-down approaches, where the protein is not digested by pepsin and the peptides are generated by ETD fragmentation inside the mass spectrometer. Pan *et al.*[192] managed to obtain 100% coverage of the LC by a top-down approach (74 % by bottom-up), but 50% for the HC (87% by bottom-up).

7.4 FUTURE OUTLOOK

The results discussed in this Chapter provide interesting insight into the unfolding of antibodies, including the very stable Fab fragment, under HDX-MS quench conditions, which, to the best of our knowledge, had not previously been investigated. However, the full exploration of an alternative protocol was not achieved because of lack of time at the end of the Ph.D.

To obtain further information on the stability of the pepsin, in particular its kinetic stability, it would be useful to measure the unfolding kinetics of pepsin with increasing concentrations of guanidinium chloride at pH 2.5 and 4 ºC to probe how long pepsin takes to unfold under these conditions. The trapping step, during which digestion on the pepsin column happens, usually last 4 minutes. It would be interesting to know if this time window is sufficiently small to avoid significant unfolding of the pepsin at the higher denaturant concentrations that would favor rapid antibody unfolding. It might also be relevant to test denaturant concentrations close to, but below, the denaturation midpoint, measured in the unfolding curves, thus maximizing the denaturing power while avoiding too much unfolding of the pepsin.

Guanidinium thiocyanate is the strongest chemical denaturant commonly used in protein research, and is approximately twice as strong as guanidinium chloride. Its use in quench buffers, rather than GdmCl, would reduce the amount of ions going into the source, which are non-negligeable despite the removal of most of the chemical denaturant in the trapping step. This would enable better detection of peptide ions. Jonathan Phillips and Laura Mitchell (MedImmune/University of Exeter) have shown that using 8 M urea rather than 4 M GdmCl leads to a better peptide map as urea is not a salt and therefore does not compete with the ionized peptides (personal communication).

Once some potentially optimal parameters have been chosen based on the results

shown in this Chapter, then an actual peptide coverage map would need to be generated to establish whether the improvement in parameters do result in improved coverage. In addition, if sufficient data is gathered on several mutants of antibodies (different variable domains, mutants in the C_H2 and C_H3 domains…), given they have the same structure as immunoglobulin domains, it may be possible to correlate the unfolding half-lives under a specific set of conditions with the melting temperatures, obtained respectively by stopped-flow and DSC. This would be useful as pharmaceutical companies routinely use DSC to probe antibody stability but rarely, if ever, measure unfolding kinetics. Thus, in the future it would be possible to establish optimal conditions for HDX MS for a new antibody based on DSC alone.

8 CONCLUSION AND FUTURE OUTLOOK

8.1 CONLUSION

In this Ph.D. thesis, the stability of different formats of therapeutic antibodies was investigated: from 'naked' monoclonal antibodies to antibody drug conjugates to masked antibodies. Biophysical tools, informing on the thermodynamic, kinetic, and thermal stability, as well as molecular dynamics and affinity, were used and the results from the different studies provided significant insight into many different aspects of the development of therapeutic antibodies. The research will play an important role and have impact in the future rational design of novel formats of antibody therapeutics.

In Chapter 3, the stability of a panel of antibody mutants, engineered to introduce new biological features or conjugation to a drug, was investigated. The thermal stability of these two mutants had been previously measured; more thorough biophysical characterization was carried out in this Chapter under equilibrium and non-equilibrium conditions using chemical denaturant. Both results agreed on the fact that the triple substitution in the C_H2 domain destabilized this domain only, and the substitution in the C_H3 domain had little impact on stability. More generally, the relative stabilities of the different domain measured in this Chapter, established through the identification of the different unfolding transitions observed, were in agreement with results previously reported in the literature (i.e. C_H2 less stable than C_H3 domain, and the stability of the Fab domain depends on the variable domains). Furthermore, the characterisation of the NIST mAb IgG revealed a remarkably high stability for the Fab domain. This observation laid the foundation for the optimisation of the HDX-MS quench unfolding conditions for very stable proteins discussed in Chapter 7.

Antibody-drug conjugates have become a major class of therapeutics for cancer treatments thanks to the combination of the high specificity of antibodies, and the highly potent nature of the cytotoxic payloads conjugated. Several strategies have been developed to conjugate the payload to the antibody, and the most reliable so far to ensure a reproducible drug-to-antibody ratio is conjugation to cysteines. In Chapter 4, an antibody-drug conjugate scaffold, where a cysteine was inserted after position 239 in the upper C_H2 domain to enable conjugation to a drug, was investigated. The inserted cysteine was found to be in one of three different states: a free-thiol (2xSH), doubly-cysteinylated (2xCys), or linked together by a third interchain disulfide bridge (iDSB), additional to the two canonical ones in the hinge region. This third state was unexpected given that in the crystal structure of an Fc domain, the residues at position 239 are too far apart to form a

covalent link. Biophysical characterisation revealed that the formation of the iDSB is favoured under conditions with increased molecular dynamics or unfolding of the C_H2 domain. Moreover, the formation of this third disulfide bridge led to a decrease in stability and even a local unfolding, most probably due to distortion of the C_H2 domain because of the newly applied strain. Even given the aforementioned destabilisation, the antibody is suitable for therapeutic purposes since it remains folded at body temperature. Furthermore, the conjugation to this scaffold generated a reproducible drug-to-antibody ratio DAR which makes it reliable in terms of manufacturing. The results from this Chapter were published in *Bioconjugate Chemistry*.

Many antibody-drug conjugates which have reproducible drug-to-antibody ratios due to the conjugation of cysteines (either canonical or inserted/substituted) have been successfully developed. The most widely used technology relies on maleimide linkers which react to the cysteine thiol groups. However, the linker can be deconjugated via a retro Michael reaction and the cytotoxic payload can be released in circulation and affect tissues that were not targeted. The focus of Chapter 5 was on a new technology of linkers which conjugate cytotoxic payloads to antibodies but which have stable linkages. The DVP and tetraDVP linkers, developed by the Spring Group (Department of Chemistry, University of Cambridge), react with the canonical cysteines and have the advantage of being extremely stable chemically, therefore ensuring the good carriage of the drug in the body until the antibody reaches its targeted organ. DVP- and tetraDVP-linked, as well as unmodified trastuzumab, were characterized to assess the impact of the linkers on the stability of the antibody. The results from the biophysical investigation showed that, although these linkers did destabilise the antibody, the effects were not large, which is encouraging for their future use as therapeutics. More specifically, the results for the tetraDVP linker suggest it destabilizes the constant domains of the Fab, whilst the DVP linker likely destabilizes the C_H2 domain and to a greater extent than the tetraDVP linker.

Although antibodies have become promising alternatives in combination with cancer therapies, they can cause some side effects, due to the recognition of their target in healthy organs where it is expressed at low levels. In the past eight years, several modalities have been successfully developed to initially inactivate the antibodies by masking the variable domains, releasing the mask only at the site of interest, usually by proteolytic cleavage of the linker with an enzyme overexpressed in the tumour/tissue.

However, it is not yet well understood how to design an ideal mask that has affinity for the therapeutic antibody, ensuring effective masking when it is tethered to the antibody, and also effective activation when the linker between the mask and antibody is cleaved. In Chapter 6, a panel of three masked antibodies, where the masks were anti-idiotypic antibody fragments (scFv and nanobody) with different affinities for trastuzumab, was generated and characterized to understand how different parameters determine the effectiveness of the mask. Digestion sites, which enable antibody activation upon proteolytic cleavage, were included i) in each linker between mask and antibody, and ii) between domains in the mask that are scFvs. The combination of activity assays (flow cytometry, biolayer interferometry, confocal microscopy) with biophysical techniques (SEC-MALS, HDX-MS and X-ray crystallography) revealed four main parameters that are key and which need to be considered in the future design of affinity masks: two of these optimise the inactivation of the antibody, whilst the other two guarantee the recovery of function once the mask is cleaved. The efficacy of the inactivation relies, first, on the extent of binding site overlap between the mask and the antigen, and second, on a relatively high k_{on} for the mask to the antibody to prevent the antigen from binding when the mask in tethered and not yet activated. As far as the efficacy of activation is concerned, the antibody needs a lower affinity and higher k_{off} for the mask than for the antigen, and finally, the closer the affinity of the mask and the antigen for the antibody, the more disruptive the activation step needs to be, e.g., multiple cleavage sites. The findings presented in Chapter 6 provide a toolbox which should prove very useful for future engineering of therapeutic antibodies with affinity masks. The content of this Chapter was recently submitted to *Nature Communications.*

The stability of the masked antibodies was also measured, Appendix III. The results showed that the dAb mask confers more stability on the Fab fragment than the svFv40 when they are covalently linked to trastuzumab. These preliminary data are very informative for future engineering experiments on these masks in order to ultimately develop alternative and more controlled ways to activate the antibody, for example by thermal unfolding of the mask at the tumour site/location of interest.

Finally, Chapter 7 tackled a technical issue observed while carrying out some HDX-MS experiments which showed rather poor coverage of the peptide map for very stable proteins or domains of proteins. As mentioned previously, the Fab domain of NIST mAb is remarkably stable, much more than other Fabs ($\Delta \boldsymbol{G}_{\boldsymbol{D-N}}^{\boldsymbol{H2O}}$= 11 kcal mol^{-1}; T_m = 92.7

°C; half-life of unfolding in water = 79.6 years). The unfolding kinetics of the C_H2, C_H3 and Fab domains of NIST mAb IgG, carried out under the same conditions as the quench step in the HDX experiments (i.e., same mixing ratio, concentration of chemical denaturant, use of TCEP and D_2O) were measured. The results showed that the concentration of chemical denaturant in the most commonly used buffer for antibodies, 4 M GdmCl, leading to a final GdmCl concentration of 2 M, is not enough to unfold the Fab domain of NIST mAb within the timeframe of seconds to minutes used in the quench step (at 0 °C in 2 M GdmCl final, the half-life is 12.7 hours). The unfolding kinetics of NIST mAb Fab were investigated using higher concentrations of GdmCl, up to 4 M final. The findings revealed that at that concentration, the Fab can unfold much more rapidly within 19.2 seconds at 0 °C. To understand if it is possible to use a much higher concentration of chemical denaturant in the quench buffer, while ensuring optimal activity of the pepsin column, the thermodynamic stability of porcine pepsin, used in the Enzymate BEH pepsin column from Waters, was measured. The results suggest that the denaturation midpoint was higher than 4 M GdmCl and therefore the use of this much higher concentration, which would enable unfolding of the Fab domain within seconds, may be possible. The effect of sheer stress on the pepsin during the trapping phase, which was not measured, should be ideally taken into account when measuring the resistance to denaturation. If more time had been available, the unfolding kinetics of the pepsin under the conditions of the mix of the buffer and labelled protein would have been measured, at different concentrations of GdmCl, to understand how long the pepsin would take to unfold. A preliminary investigation was also carried out using guanidinium thiocyanate, a chemical denaturant approximately twice as strong as GdmCl. Using it in HDX-MS workflow would reduce by half the amount of salt present in the sample and therefore ease the trapping step and interfere less with the electrospray ionisation, likely leading to a larger number of peptides identified and thus a better peptide coverage. Ultimately, the experiments need to be extended and the results need to be validated by measuring the peptide map on the mass spectrometer, with higher concentrations of the denaturant in the quench buffer.

Therapeutic antibodies are mainly developed and investigated by pharmaceutical companies and, although a lot of effort is put into engineering the complementarity determining regions so they are specific to the target, the intrinsic stability is generally not investigated apart from thermal stability using differential scanning calorimetry. The thorough characterization of the different antibodies presented in this Ph.D. Thesis brings

a more thorough and fundamental understanding of the factors that impact antibody stability with particular focus on novel formats. The results should greatly facilitate the design of future therapeutics.

8.2 FUTURE OUTLOOK

The studies carried out in this Thesis open up many avenues for future research. The Spring Group has developed several different formats of linkers with which to conjugate cytotoxic payloads to antibodies. These are stable and, in contrast to the maleimide linkers, the payload does not fall off during circulation of the ADC in the bloodstream, as discussed in Chapter 5. Given that the design of their linkers is quite flexible, it would be interesting to test the stability of antibodies conjugated with different linkers in order to find the least destabilizing one, or even one that does not affect the stability at all, or even stabilises the antibody. The conjugation of these novel linkers to NIST mAb IgG, rather than trastuzumab, could facilitate the interpretation of the biophysical results. Given that the midpoints of denaturation of the C_H2, C_H3 and the Fab domains of trastuzumab are very close, and the melting temperature of the C_H3 and the Fab domains are the same (T_m = 82 °C in 50 mM Tris pH 7.5), the interpretation of the data and the identification of potential changes in the stability of the different domains can be complicated. On the other hand, the very high stability of the Fab domain of NIST mAb IgG (T_m = 92 °C in 20 mM His pH 5.5) compared to the C_H2 and C_H3 domains means that distinct unfolding transitions can be observed, making the interpretation of changes in stability of the different domains much easier (see Chapter 3).

Regarding the studies on the masked antibodies (Chapter 6), there are a number of ways in which this project should be taken forward. First, it would be interesting to confirm the effectiveness of the masks in preventing side effects whilst maintaining efficacy *in vivo* and, in order to do this, the enzyme cleavage site(s) need to be changed to an enzyme(s) that are overexpressed in breast cancer. This could be MMP-2, MMP-9 or matriptase. The same set of biophysical experiments should be undertaken before a suitable *in vivo* model is used.

Thanks to the information obtained from the crystal structure of the trastuzumab masked by dAb (T-dAb, Chapter 6), it is now possible to engineer the dAb, which is

currently not a good masking moiety, in order to have an additional mask that has the optimal properties for inactivation/activation. For example, a mask which increases the EC_{50} of trastuzumab to a larger extent than the scFv40, but which can also be efficiently activated by cleavage of a linker. As an initial step, one could undertake an alanine scan on all residues in the CDRs which interact with trastuzumab. Such a project would serve two purposes, first of all it would generate an additional mask which could ultimately be tested *in vivo*. Second, it would serve to confirm and extend the results presented in Chapter 6, and to test the hypothesis that in addition to binding site and affinity, that association and dissociation kinetics also play a key role in the inactivation/activation steps.

It would also be possible to lower the stability of the dAb, which has a very high T_m of 67 ºC (Appendix III), by introducing substitutions in the hydrophobic core; this would potentially enable the activation step to be carried out by means other than proteolytic cleavage only. As many enzymes that are overexpressed in tumours are also present in healthy organs as well albeit at a lower level, there may also be the problem of activation of the masked antibodies at unwanted sites. There is, therefore, a need to investigate other and even tighter activation mechanisms, for instance, by thermal unfolding of the mask, which may be achieved under conditions generating heat, e.g. plasmonic nanoparticles. This would enable the very specific control of the activation of the masked antibody *in situ*. The engineering of this kind of format would necessitate a thorough biophysical understanding of the whole antibody (thermodynamic, kinetic, and thermal stability) for optimal design of the mask. This could be useful for the next generation of therapeutic antibodies as a way of preventing side effects but, more interestingly, antibodies that have failed in the clinical trials because they triggered severe reactions like cytokine storms and therefore the development of the antibody had to be stopped, could be revisited if an effective inactivation/activation mechanism could be engineered into these systems.

Regarding the investigation of the unfolding of antibodies under HDX quench conditions (Chapter 7), it would be interesting to finish the study by testing the unfolding kinetics of the pepsin with different concentrations of guanidinium chloride under the conditions of the quench, to fully characterise the resistance of pepsin to chemical denaturant. If the pepsin was found to be more thermodynamically and kinetically stable than currently thought (as outlined by the guides on using pepsin columns supplied by the

manufacturers), then more highly denaturing conditions could be used during the pepsin digest step and then, hopefully, improved coverage of the peptide map of NIST mAb Fab would be attained, thus establishing a better methodology to use with highly stable proteins. This would make a technical paper which could be very useful for the HDX-MS community, especially for the analysis of various stable proteins including some antibodies.

8.3 CONTRIBUTIONS TO THE THESIS

All the experiments presented in this Thesis were carried out by myself unless otherwise stated. In Chapter 3, the antibodies were a kind gift from AstraZeneca and I did all the experiments presented in the Chapter. Chapter 4 involved the most collaborative project, where I generated 60% of the results, and the other 40% were carried out by my collaborators at AstraZeneca (enrichment of the variants, quantification of the states by mass spectrometry and generation of ADCs by Dr Matthew Edgeworth, Dr Paul Devine, Alistair Hines, Christopher Thompson, Dr Owen Cornwell). In Chapter 5, the DVP and tetra DVP linkers were prepared by the Spring group (Department of Chemistry) and I generated all the results presented. In Chapter 6, I generated 80% of the results, and the AstraZeneca collaborators generated 20% (SEC-MALS, confocal microscopy and X-ray crystallography by Wesley Howard, Dr Lorraine Irving and Dr David Hargreaves, respectively). In Chapter 7, NIST mAb was a kind gift from AstraZeneca and I performed all the experiments.

9 References

1. Khalil, D. N., Smith, E. L., Brentjens, R. J. & Wolchok, J. D. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nat. Rev. Clin. Oncol.* **13**, 273–290 (2016).

2. Zahavi, D. & Weiner, L. Monoclonal antibodies in cancer therapy. *Antibodies* **9**, 34 (2020).

3. Kintzing, J. R., Filsinger Interrante, M. V. & Cochran, J. R. Emerging strategies for developing next-generation protein therapeutics for cancer treatment. *Trends Pharmacol. Sci.* **37**, 993–1008 (2016).

4. Hansel, T. T., Kropshofer, H., Singer, T., Mitchell, J. A. & George, A. J. T. The safety and side effects of monoclonal antibodies. *Nat. Rev. Drug Discov.* **9**, 325–338 (2010).

5. Köhler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495–497 (1975).

6. Reichert, J. M., Rosensweig, C. J., Faden, L. B. & Dewitz, M. C. Monoclonal antibody successes in the clinic. *Nat. Biotechnol.* **23**, 1073–1078 (2005).

7. Morrison, S. L., Johnson, M. J., Herzenberg, L. A. & Oi, V. T. Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6851–6855 (1984).

8. Boulianne, G. L., Hozumi, N. & Shulman, M. J. Production of functional chimaeric mouse/human antibody. *Nature* **312**, 643–646 (1984).

9. Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* **321**, 522–525 (1986).

10. Verhoeyen, M., Milstein, C. & Winter, G. Reshaping human antibodies: grafting an antilysozyme activity. *Science* **239**, 1534–1536 (1988).

11. Riechmann, L., Clark, M., Waldmann, H. & Winter, G. Reshaping human antibodies for therapy. *Nature* **332**, 323–327 (1988).

12. Green, L. L. *et al.* Antigen–specific human monoclonal antibodies from mice

engineered with human Ig heavy and light chain YACs. *Nat. Genet.* **7**, 13–21 (1994).

13. Lonberg, N. *et al.* Antigen-specific human antibodies from mice comprising four distinct genetic modifications. *Nature* **368**, 856–859 (1994).

14. Carter, P. J. Potent antibody therapeutics by design. *Nat. Rev. Immunol.* **6**, 343–357 (2006).

15. Lonberg, N. Human antibodies from transgenic animals. *Nat. Biotechnol.* **23**, 1117–1125 (2005).

16. Marks, J. D. *et al.* By-passing immunization: human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* **222**, 581–597 (1991).

17. Even-Desrumeaux, K., Baty, D. & Chames, P. State of the art in tumor antigen and biomarker discovery. *Cancers (Basel).* **3**, 2554–2596 (2011).

18. Mullard, A. FDA approves 100th monoclonal antibody product. *Nat. Rev. Drug Discov.* **20**, 491–495 (2021).

19. Lu, R.-M. *et al.* Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.* **27**, 1 (2020).

20. Wu, A. M. & Senter, P. D. Arming antibodies: prospects and challenges for immunoconjugates. *Nat. Biotechnol.* **23**, 1137–1146 (2005).

21. Bross, P. F. *et al.* Approval summary: gemtuzumab ozogamicin in relapsed acute myeloid leukemia. *Clin. Cancer Res.* **7**, 1490–1496 (2001).

22. Teicher, B. A. Antibody drug conjugates. *Curr. Opin. Oncol.* **26**, 476–483 (2014).

23. Ballantyne, A. & Dhillon, S. Trastuzumab emtansine: first global approval. *Drugs* **73**, 755–765 (2013).

24. Lamb, Y. N. Inotuzumab Ozogamicin: first global approval. *Drugs* **77**, 1603–1610 (2017).

25. Richardson, N. C. *et al.* FDA approval summary: Brentuximab vedotin in first-line treatment of peripheral T-cell lymphoma. *Oncologist* **24**, e180–e187 (2019).

26. Deeks, E. D. Polatuzumab vedotin: first global approval. *Drugs* **79**, 1467–1475 (2019).

27. Hanna, K. S. Clinical overview of enfortumab vedotin in the management of locally advanced or metastatic urothelial carcinoma. *Drugs* **80**, 1–7 (2020).

28. Keam, S. J. Trastuzumab deruxtecan: first approval. *Drugs* **80**, 501–508 (2020).

29. Syed, Y. Y. Sacituzumab govitecan: first approval. *Drugs* **80**, 1019–1025 (2020).

30. Tzogani, K. *et al.* EMA Review of belantamab mafodotin (Blenrep) for the treatment of adult patients with relapsed/refractory multiple myeloma. *Oncologist* **26**, 70–76 (2021).

31. Senter, P. D. & Sievers, E. L. The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma. *Nat. Biotechnol.* **30**, 631–637 (2012).

32. Gauzy-Lazo, L., Sassoon, I. & Brun, M.-P. Advances in antibody-drug conjugate design: current clinical landscape and future innovations. *SLAS Discov. Adv. life Sci. R D* **25**, 843–868 (2020).

33. Beck, A., Goetsch, L., Dumontet, C. & Corvaïa, N. Strategies and challenges for the next generation of antibody–drug conjugates. *Nat. Rev. Drug Discov.* **16**, 315–337 (2017).

34. Thompson, P. *et al.* Rational design, biophysical and biological characterization of site-specific antibody-tubulysin conjugates with improved stability, efficacy and pharmacokinetics. *J. Control. Release* **236**, 100–116 (2016).

35. Dimasi, N. *et al.* Efficient preparation of site-specific antibody–drug conjugates using cysteine insertion. *Mol. Pharm.* **14**, 1501–1516 (2017).

36. Kim, C. H., Axup, J. Y. & Schultz, P. G. Protein conjugation with genetically encoded unnatural amino acids. *Curr. Opin. Chem. Biol.* **17**, 412–419 (2013).

37. Lang, K. & Chin, J. W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **114**, 4764–4806 (2014).

38. Tian, F. *et al.* A general approach to site-specific antibody drug conjugates. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1766–1771 (2013).

39. Rabuka, D., Rush, J. S., deHart, G. W., Wu, P. & Bertozzi, C. R. Site-specific chemical protein conjugation using genetically encoded aldehyde tags. *Nat. Protoc.* **7**, 1052–1067 (2012).

40. Jeger, S. *et al.* Site-specific and stoichiometric modification of antibodies by bacterial transglutaminase. *Angew. Chemie Int. Ed.* **49**, 9995–9997 (2010).

41. Strop, P. *et al.* Location matters: site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates. *Chem. Biol.* **20**, 161–167 (2013).

42. Madej, M. P. *et al.* Engineering of an anti-epidermal growth factor receptor antibody to single chain format and labeling by sortase A-mediated protein ligation. *Biotechnol. Bioeng.* **109**, 1461–1470 (2012).

43. Walsh, S. J. *et al.* Site-selective modification strategies in antibody–drug conjugates. *Chem. Soc. Rev.* **50**, 1305–1353 (2021).

44. Beck, A. & Reichert, J. M. Antibody-drug conjugates. *MAbs* **6**, 15–17 (2014).

45. Akkapeddi, P. *et al.* Construction of homogeneous antibody–drug conjugates using site-selective protein chemistry. *Chem. Sci.* **7**, 2954–2963 (2016).

46. Dere, R. *et al.* PK assays for antibody–drug conjugates: case study with ado-trastuzumab emtansine. *Bioanalysis* **5**, 1025–1040 (2013).

47. Walsh, S. J. *et al.* A general approach for the site-selective modification of native proteins, enabling the generation of stable and functional antibody–drug conjugates. *Chem. Sci.* **10**, 694–700 (2019).

48. Walsh, S. J. *et al.* General dual functionalisation of biomacromolecules *via* a cysteine bridging strategy. *Org. Biomol. Chem.* **18**, 4224–4230 (2020).

49. Dannheim, F. M., Walsh, S. J. & Spring, D. R. Conjugating reagents and conjugates thereof, Pat Appl GB2110726.3. (2021).

50. Dannheim, F. Development of a novel bioconjugation platform for the generation

of homogenous antibody-drug conjugates. (University of Cambridge, 2021).

51. Giezen, T. J. *et al.* Safety-related regulatory actions for biologicals approved in the United States and the European Union. *JAMA* **300**, 1887–1896 (2008).

52. Ji-Liang Li, A. M. J. & A. L. H. Targeting DLL4 in tumors shows preclinical activity but potentially significant toxicity. *Futur. Oncol.* **6**, 1099–1103 (2010).

53. Tijink, B. M. *et al.* A phase I dose escalation study with anti-CD44v6 bivatuzumab mertansine in patients with incurable squamous cell carcinoma of the head and neck or esophagus. *Clin. Cancer Res.* **9**, 6064–6072 (2006).

54. Haluska, P. *et al.* Safety, tolerability, and pharmacokinetics of the anti-IGF-1R monoclonal antibody figitumumab in patients with refractory adrenocortical carcinoma. *Cancer Chemother. Pharmacol.* **65**, 765–773 (2010).

55. Chung, C. H. *et al.* Cetuximab-induced anaphylaxis and IgE specific for galactose-α-1,3-galactose. *N. Engl. J. Med.* **358**, 1109–1117 (2008).

56. Cox, L. *et al.* American academy of allergy, asthma & immunology/American college of allergy, asthma and immunology joint task force report on omalizumab-associated anaphylaxis. *J. Allergy Clin. Immunol.* **120**, 1373–1377 (2007).

57. Corren, J. *et al.* Safety and tolerability of omalizumab. *Clin. Exp. Allergy* **39**, 788–797 (2009).

58. Cox, L. S. How safe are the biologicals in treating asthma and rhinitis? *Allergy, Asthma Clin. Immunol.* **5**, 4 (2009).

59. Ransohoff, R. M. Natalizumab for multiple sclerosis. *N. Engl. J. Med.* **356**, 2622–2629 (2007).

60. Aster, R. H. & Bougie, D. W. Drug-induced immune thrombocytopenia. *N. Engl. J. Med.* **357**, 580–587 (2007).

61. Haider, I. & Cahill, M. Fatal thrombocytopaenia temporally related to the administration of alemtuzumab (MabCampath) for refractory CLL despite early discontinuation of therapy. *Hematology* **9**, 409–411 (2004).

62. Scappaticci, F. A. *et al.* Arterial thromboembolic events in patients with metastatic carcinoma treated with chemotherapy and bevacizumab. *J. Natl. Cancer Inst.* **99**, 1232–1239 (2007).

63. Ramos-Casals, M. *et al.* Autoimmune diseases induced by TNF-targeted therapies. *Medicine (Baltimore).* **86**, 242–251 (2007).

64. The CAMMS223 Trial Investigators. Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. *N. Engl. J. Med.* **359**, 1786–1801 (2008).

65. Peggs, K. S., Quezada, S. A., Korman, A. J. & Allison, J. P. Principles and use of anti-CTLA4 antibody in human cancer immunotherapy. *Curr. Opin. Immunol.* **18**, 206–213 (2006).

66. Weber, J. Review: anti–CTLA-4 antibody ipilimumab: case studies of clinical response and immune-related adverse events. *Oncologist* **12**, 864–872 (2007).

67. Peréz-Soler, R. & Saltz, L. Cutaneous adverse effects with HER1/EGFR-targeted agents: is there a silver lining? *J. Clin. Oncol.* **23**, 5235–5246 (2005).

68. Force, T., Krause, D. S. & Van Etten, R. A. Molecular mechanisms of cardiotoxicity of tyrosine kinase inhibition. *Nat. Rev. Cancer* **7**, 332–344 (2007).

69. Guglin, M., Cutro, R. & Mishkin, J. D. Trastuzumab-induced cardiomyopathy. *J. Card. Fail.* **14**, 437–444 (2008).

70. Rautio, J. *et al.* Prodrugs: design and clinical applications. *Nat. Rev. Drug Discov.* **7**, 255–270 (2008).

71. Donaldson, J. M., Kari, C., Fragoso, R. C., Rodeck, U. & Williams, J. C. Design and development of masked therapeutic antibodies to limit off-target effects: Application to anti-EGFR antibodies. *Cancer Biol. Ther.* **8**, 2147–2152 (2009).

72. Erster, O. *et al.* Site-specific targeting of antibody activity in vivo mediated by disease-associated proteases. *J. Control. Release* **161**, 804–812 (2012).

73. Desnoyers, L. R. *et al.* Tumor-specific activation of an EGFR-targeting probody enhances therapeutic index. *Sci. Transl. Med.* **5**, 207ra144 (2013).

74. Polu, K. R. & Lowman, H. B. Probody therapeutics for targeting antibodies to diseased tissue. *Expert Opin. Biol. Ther.* **14**, 1049–1053 (2014).

75. Kavanaugh, W. M. Antibody prodrugs for cancer. *Expert Opin. Biol. Ther.* **20**, 163–171 (2020).

76. Lin, W.-W., Lu, Y.-C., Chuang, C.-H. & Cheng, T.-L. Ab locks for improving the selectivity and safety of antibody drugs. *J. Biomed. Sci.* **27**, 76 (2020).

77. Stroh, M. *et al.* Model-informed drug development of the masked anti-PD-L1 antibody CX-072. *Clin. Pharmacol. Ther.* **109**, 383–393 (2021).

78. Dettling, D. *et al.* Abstract 557: COBRA: A novel conditionally active bispecific antibody that regresses established solid tumors in mice. in *Immunology* **79**, 557–557 (2019).

79. Onuoha, S. C., Ferrari, M., Sblattero, D. & Pitzalis, C. Rational design of antirheumatic prodrugs specific for sites of inflammation. *Arthritis Rheumatol.* **67**, 2661–2672 (2015).

80. Pai, C.-C. S. *et al.* Tumor-conditional anti-CTLA4 uncouples antitumor efficacy from immunotherapy-related toxicity. *J. Clin. Invest.* **129**, 349–363 (2018).

81. Geiger, M. *et al.* Protease-activation using anti-idiotypic masks enables tumor specificity of a folate receptor 1-T cell bispecific antibody. *Nat. Commun.* **11**, 3196 (2020).

82. Trang, V. H. *et al.* A coiled-coil masking domain for selective activation of therapeutic antibodies. *Nat. Biotechnol.* **37**, 761–765 (2019).

83. Kim, H., Gaynor, A. & Chen, W. Tunable modulation of antibody-antigen interaction by protease cleavage of protein M. *Biotechnol. Bioeng.* **116**, 2834–2842 (2019).

84. Wouters, S. F. A., Wijker, E. & Merkx, M. Optical control of antibody activity by using photocleavable bivalent peptide–DNA locks. *ChemBioChem* **20**, 2463–2466 (2019).

85. Engelen, W. *et al.* Programmable bivalent peptide–DNA locks for pH-based

control of antibody activity. *ACS Cent. Sci.* **6**, 22–31 (2020).

86. Janssen, B. M. G. *et al.* Reversible blocking of antibodies using bivalent peptide–DNA conjugates allows protease-activatable targeting. *Chem. Sci.* **4**, 1442 (2013).

87. Lu, Y.-C. *et al.* Specific activation of pro-Infliximab enhances selectivity and safety of rheumatoid arthritis therapy. *PLOS Biol.* **17**, e3000286 (2019).

88. Metz, S. *et al.* Bispecific antibody derivatives with restricted binding functionalities that are activated by proteolytic processing. *Protein Eng. Des. Sel.* **25**, 571–580 (2012).

89. Chen, I.-J. *et al.* Selective antibody activation through protease-activated pro-antibodies that mask binding sites with inhibitory domains. *Sci. Rep.* **7**, 11587 (2017).

90. Gunnoo, S. B. *et al.* Creation of a gated antibody as a conditionally functional synthetic protein. *Nat. Commun.* **5**, 4388 (2014).

91. Ngambenjawong, C., Sylvestre, M., Gustafson, H. H., Pineda, J. M. B. & Pun, S. H. Reversibly switchable, pH-dependent peptide ligand binding via 3,5-diiodotyrosine substitutions. *ACS Chem. Biol.* **13**, 995–1002 (2018).

92. Saphire, E. O. *et al.* Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. *Science* **293**, 1155–1159 (2001).

93. Branden, C. & Tooze, J. *Introduction to Protein Structure - Second Edition*. (Garland Science, 1998).

94. Feige, M. J., Walter, S. & Buchner, J. Folding mechanism of the CH2 antibody domain. *J. Mol. Biol.* **344**, 107–118 (2004).

95. Feige, M. J. *et al.* Structure of the murine unglycosylated IgG1 Fc fragment. *J. Mol. Biol.* **391**, 599–608 (2009).

96. Garber, E. & Demarest, S. J. A broad range of Fab stabilities within a host of therapeutic IgGs. *Biochem. Biophys. Res. Commun.* **355**, 751–757 (2007).

97. Greene, R. F. & Pace, C. N. Urea and Guanidine Hydrochloride Denaturation of

Ribonuclease, Lysozyme, α-Chymotrypsin, and b-Lactoglobulin. *J. Biol. Chem.* **249**, 5388–5393 (1974).

98. von Hippel, P. H. & Wong, K.-Y. On the conformational stability of globular proteins: the effects of various electrolytes and nonelectrolytes on the thermal ribonuclease transition. *J. Biol. Chem.* **240**, 3909–3923 (1965).

99. Castellino, F. J. & Barker, R. The denaturing effectiveness of guanidinium, carbamoylguanidinium, and guanylguanidinium salts. *Biochemistry* **7**, 4135–4138 (1968).

100. Lapanje, S. Denaturation of globular proteins by guanidine thiocyanate I. Optical rotation in aqueous guanidine thiocyanate solutions. *Biochim. Biophys. Acta - Protein Struct.* **243**, 349–356 (1971).

101. Schmid, F. X. Spectral probes of conformation. in *Protein structure: a practical approach* (ed. Creighton, T. E.) 251–285 (IRL Press at Oxford University Press, 1989).

102. Tanford, C. Protein Denaturation. *Adv. Protein Chem.* **23**, 121–282 (1968).

103. Pace, C. N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280 (1986).

104. Dobson, C. L. *et al.* Engineering the surface properties of a human monoclonal antibody prevents self-association and rapid clearance in vivo. *Sci. Rep.* **6**, 38644 (2016).

105. Kim, M. *et al.* Antibody mechanics on a membrane-bound HIV segment essential for GP41-targeted viral neutralization. *Nat. Struct. Mol. Biol.* **18**, 1235–1243 (2011).

106. Edgeworth, M. J. *et al.* Global and local conformation of human IgG antibody variants rationalizes loss of thermodynamic stability. *Angew. Chemie Int. Ed.* **54**, 15156–15159 (2015).

107. Orozco, C. T. *et al.* Interconversion of unexpected thiol states affects the stability, structure, and dynamics of antibody engineered for site-specific conjugation.

Bioconjug. Chem. **32**, 1834–1844 (2021).

108. Eigen, M. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: Elementary processes. *Angew. Chemie Int. Ed. English* **3**, 1–19 (1964).

109. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 75–86 (1993).

110. Smith, D. L., Deng, Y. & Zhang, Z. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* **32**, 135–146 (1997).

111. Oganesyan, I., Lento, C. & Wilson, D. J. Contemporary hydrogen deuterium exchange mass spectrometry. *Methods* **144**, 27–42 (2018).

112. Dempsey, C. E. Hydrogen exchange in peptides and proteins using NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **39**, 135–170 (2001).

113. Narang, D., Lento, C. & J. Wilson, D. HDX-MS: An analytical tool to capture protein motion in action. *Biomedicines* **8**, 224 (2020).

114. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521 (1983).

115. Morgan, C. R. & Engen, J. R. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr. Protoc. Protein Sci.* **58**, 17.6.1-17.6.17 (2009).

116. Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **88**, 30–51 (2016).

117. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biol. Mass Spectrom.* **11**, 601–601 (1984).

118. Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biol. Mass Spectrom.* **16**, 99–111 (1988).

119. McLuckey, S. A. Principles of collisional activation in analytical mass

spectrometry. *J. Am. Soc. Mass Spectrom.* **3**, 599–614 (1992).

120. Hunt, D. F., Yates, J. R., Shabanowitz, J., Winston, S. & Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 6233–7 (1986).

121. Böhm, G., Muhr, R. & Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *"Protein Eng. Des. Sel.* **5**, 191–195 (1992).

122. Clarke, D. T. Circular Dichroism and Its Use in Protein-Folding Studies. in 59–72 (Humana Press, Totowa, NJ, 2011). doi:10.1007/978-1-60327-223-0_5

123. Yang, H., Yang, S., Kong, J., Dong, A. & Yu, S. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nat. Protoc.* **10**, 382–396 (2015).

124. Huxford, T. X-Ray Crystallography. *Brenner's Encycl. Genet.* 366–368 (2013). doi:10.1016/B978-0-12-374984-0.01657-0

125. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).

126. Kreplak, L. Introduction to Atomic Force Microscopy (AFM) in Biology. *Curr. Protoc. Protein Sci.* **85**, 17.7.1-17.7.21 (2016).

127. Cole, J. L., Lary, J. W., P. Moody, T. & Laue, T. M. Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium. *Methods Cell Biol.* **84**, 143–179 (2008).

128. Shimura, K. Capillary Isoelectric Focusing. *Capill. Electromigr. Sep. Methods* 167–187 (2018). doi:10.1016/B978-0-12-809375-7.00007-1

129. Johnson, D. T., Di Stefano, L. H. & Jones, L. M. Fast photochemical oxidation of proteins (FPOP): A powerful mass spectrometry-based structural proteomics tool. *J. Biol. Chem.* **294**, 11969–11979 (2019).

130. Cornwell, O., Radford, S. E., Ashcroft, A. E. & Ault, J. R. Comparing Hydrogen Deuterium Exchange and Fast Photochemical Oxidation of Proteins: a Structural Characterisation of Wild-Type and ΔN6 β_2 -Microglobulin. *J. Am. Soc. Mass*

Spectrom. **29**, 2413–2426 (2018).

131. Chen, X. (Nancy), Nguyen, M., Jacobson, F. & Ouyang, J. Charge-based analysis of antibodies with engineered cysteines. *MAbs* **1**, 563–571 (2009).

132. Coelho, M. *et al.* Isolation and characterisation of a human anti-idiotypic scFv used as a surrogate tumour antigen to elicit an anti-HER-2/neu humoral response in mice. *Br. J. Cancer* **90**, 2032–2041 (2004).

133. Alvarez-Rueda, N. *et al.* A llama single domain anti-idiotypic antibody mimicking HER2 as a vaccine: Immunogenicity and efficacy. *Vaccine* **27**, 4826–4833 (2009).

134. Ravn, P. *et al.* Structural and pharmacological characterization of novel potent and selective monoclonal antibody antagonists of glucose-dependent insulinotropic polypeptide receptor. *J. Biol. Chem.* **288**, 19760–19772 (2013).

135. Daramola, O. *et al.* A high-yielding CHO transient system: coexpression of genes encoding EBNA-1 and GS enhances transient protein expression. *Biotechnol. Prog.* **30**, 132–141 (2014).

136. Scholtz, J. M., Grimsley, G. R. & Pace, C. N. Solvent denaturation of proteins and interpretations of the m value. *Methods Enzymol.* **466**, 549–565 (2009).

137. Gani, K., Bhambure, R., Deulgaonkar, P., Mehta, D. & Kamble, M. Understanding unfolding and refolding of the antibody fragment (Fab). I. In-vitro study. *Biochem. Eng. J.* **164**, 1077642 (2020).

138. Jackson, S. E. & Fersht, A. R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991).

139. Kish, M. *et al.* Allosteric regulation of glycogen phosphorylase solution phase structural dynamics at high spatial resolution. *bioRxiv* 654665 (2019).

140. Webster, C. I. *et al.* Enhanced delivery of IL-1 receptor antagonist to the central nervous system as a novel anti-transferrin receptor-IL-1RA fusion reverses neuropathic mechanical hypersensitivity. *Pain* **158**, 660–668 (2017).

141. Vonrhein, C. *et al.* Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 293–302 (2011).

142. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

143. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. & IUCr. Features and development of *Coot*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).

144. Murshudov, G. N. *et al. REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).

145. Bricogne, G. *et al.* BUSTER version 2.8.0. *Cambridge, United Kindgom Glob. Phasing Ltd* (2009).

146. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: from structure to effector functions. *Front. Immunol.* **5**, 520 (2014).

147. Oganesyan, V. *et al.* Structural characterization of a human Fc fragment engineered for lack of effector functions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 700–704 (2008).

148. Sondermann, P., Huber, R., Oosthuizen, V. & Jacob, U. The 3.2-Å crystal structure of the human IgG1 Fc fragment–FcγRIII complex. *Nature* **406**, 267–273 (2000).

149. Johnson, C. M. Differential scanning calorimetry as a tool for protein folding and stability. *Arch. Biochem. Biophys.* **531**, 100–109 (2013).

150. Tischenko, V. M., Vyacheslav M. Abramov, A. & Zav'yalov, V. P. Investigation of the cooperative structure of Fc fragments from myeloma immunoglobulin G. *Biochemistry* **37**, 5576–5581 (1998).

151. Vermeer, A. W. P. & Norde, W. The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophys. J.* **78**, 394–404 (2000).

152. Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Cortijo, M. & Mateo, P. L. Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. *Biochemistry* **27**, 1648–1652 (1988).

153. Sumi, A. & Hamaguchi, K. Denaturation by guanidine hydrochloride of the Fc(t)

and pFc' fragments of human immunoglobulin G. *J. Biochem.* **92**, 823–33 (1982).

154. Gallagher, D. T., Karageorgos, I., Hudgens, J. W. & Galvin, C. V. Data on crystal organization in the structure of the Fab fragment from the NIST reference antibody, RM 8671. *Data Br.* **16**, 29–36 (2018).

155. Gébleux, R. & Casi, G. Antibody-drug conjugates: Current status and future perspectives. *Pharmacol. Ther.* **167**, 48–59 (2016).

156. Alley, S. C. *et al.* The pharmacologic basis for antibody-auristatin conjugate activity. *J. Pharmacol. Exp. Ther.* **330**, 932–8 (2009).

157. Lutz, R. J. & Whiteman, K. R. Antibody-maytansinoid conjugates for the treatment of myeloma. *MAbs* **1**, 548–551 (2009).

158. Ikeda, H. *et al.* The monoclonal antibody nBT062 conjugated to cytotoxic maytansinoids has selective cytotoxicity against CD138-positive multiple myeloma cells in vitro and in vivo. *Clin. Cancer Res.* **7**, 1490–1496 (2009).

159. Bardia, A. *et al.* Sacituzumab govitecan-hziy in refractory metastatic triple-negative breast cancer. *N. Engl. J. Med.* **380**, 741–751 (2019).

160. Modi, S. *et al.* Trastuzumab deruxtecan in previously treated HER2-positive breast cancer. *N. Engl. J. Med.* **382**, 610–621 (2020).

161. Birrer, M. J., Moore, K. N., Betella, I. & Bates, R. C. Antibody-drug conjugate-based therapeutics: state of the science. *JNCI J. Natl. Cancer Inst.* **111**, 538–549 (2019).

162. Gallagher, D. T. *et al.* Structure and dynamics of a site-specific labeled Fc fragment with altered effector functions. *Pharmaceutics* **11**, 546 (2019).

163. Raju, T. S. Cysteinylation of proteins. in *Co- and Post-Translational Modifications of Therapeutic Antibodies and Proteins* 35–38 (John Wiley & Sons, Inc., 2019).

164. Wu, H., Kroe-Barrett, R., Singh, S., Robinson, A. S. & Roberts, C. J. Competing aggregation pathways for monoclonal antibodies. *FEBS Lett.* **588**, 936–941 (2014).

165. Ionescu, R. M., Vlasak, J., Price, C. & Kirchmeier, M. Contribution of variable

domains to the stability of humanized IgG1 monoclonal antibodies. *J. Pharm. Sci.* **97**, 1414–1426 (2008).

166. Alegre-Cebollada, J., Kosuri, P., Rivas-Pardo, J. A. & Fernández, J. M. Direct observation of disulfide isomerization in a single protein. *Nat. Chem.* **3**, 882–887 (2011).

167. Gilbert, H. F. Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods Enzymol.* **251**, 8–28 (1995).

168. Myers, J. K., Pace, C. N. & Scholtz, J. M. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–48 (1995).

169. Walters, B. T. *et al.* Conformational destabilization of Iimmunoglobulin G increases the low pH binding affinity with the neonatal Fc receptor. *J. Biol. Chem.* **291**, 1817–1825 (2016).

170. Cornwell, O., Radford, S. E., Ashcroft, A. E. & Ault, J. R. Comparing hydrogen deuterium exchange and fast photochemical oxidation of proteins: a structural characterisation of wild-type and ΔN6 β2-microglobulin. *J. Am. Soc. Mass Spectrom.* **29**, 2413–2426 (2018).

171. Bhattacharyya, R., Pal, D. & Chakrabarti, P. Disulfide bonds, their stereospecific environment and conservation in protein structures. *Protein Eng. Des. Sel.* **17**, 795–808 (2004).

172. Matsumiya, S. *et al.* Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J. Mol. Biol.* **368**, 767–779 (2007).

173. Matsumiya, S. *et al.* Corrigendum to “Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1” [J. Mol. Biol. 386/3 (2007) 767–779]. *J. Mol. Biol.* **408**, 1001 (2011).

174. Doyle, C. M. *et al.* Energetics of oligomeric protein folding and association. *Arch. Biochem. Biophys.* **531**, 44–64 (2013).

175. Garber, E. & Demarest, S. J. A broad range of Fab stabilities within a host of

therapeutic IgGs. *Biochem. Biophys. Res. Commun.* **355**, 751–757 (2007).

176. Wakankar, A. A. *et al.* Physicochemical stability of the antibody−drug conjugate trastuzumab-DM1: changes due to modification and conjugation processes. *Bioconjug. Chem.* **21**, 1588–1595 (2010).

177. Hudziak, R. M. *et al.* p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor. *Mol. Cell. Biol.* **9**, 1165–1172 (1989).

178. Seidman, A. *et al.* Cardiac Dysfunction in the Trastuzumab Clinical Trials Experience. *J. Clin. Oncol.* **20**, 1215–1221 (2002).

179. Ram, S., Kim, D., Ober, R. J. & Ward, E. S. The level of HER2 expression is a predictor of antibody-HER2 trafficking behavior in cancer cells. *MAbs* **6**, 1211–1219 (2014).

180. Schiettecatte, J., Anckaert, E. & Smitz, J. Interferences in Immunoassays. in *Advances in Immunoassay Technology* (InTech, 2012).

181. Tate, J. & Ward, G. Interferences in immunoassay. *Clin. Biochem. Rev.* **25**, 105–20 (2004).

182. Fernando, S. A. & Wilson, G. S. Studies of the 'hook' effect in the one-step sandwich immunoassay. *J. Immunol. Methods* **151**, 47–66 (1992).

183. Wörn, A. & Plückthun, A. Different equilibrium stability behavior of scFv fragments: identification, classification, and improvement by protein engineering. *Biochemistry* **38**, 8739–8750 (1999).

184. Cho, H.-S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421**, 756–760 (2003).

185. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

186. Wang, L., Pan, H. & Smith, D. L. Hydrogen exchange-mass spectrometry: optimization of digestion conditions. *Mol. Cell. Proteomics* **1**, 132–138 (2002).

187. Busby, S. A., Chalmers, M. J. & Griffin, P. R. Improving digestion efficiency under H/D exchange conditions with activated pepsinogen coupled columns. *Int. J. Mass Spectrom.* **259**, 130–139 (2007).

188. Ahn, J., Jung, M. C., Wyndham, K., Yu, Y. Q. & Engen, J. R. Pepsin immobilized on high-strength hybrid particles for continuous flow online digestion at 10,000 psi. *Anal. Chem.* **84**, 7256–62 (2012).

189. Fang, J., Qing Yu, Y. & Chen, W. Pressurized online pepsin digestion of mAb IgG2 for hydrogen deuterium exchange mass spectrometry. Available at: https://www.waters.com/webassets/cms/library/docs/720005693en.pdf. (Accessed: 13th September 2021)

190. Silva, J. L. & Weber, G. Pressure Stability of Proteins. *Annu. Rev. Phys. Chem.* **44**, 89–113 (1993).

191. Hudgens, J. W. *et al.* Interlaboratory comparison of hydrogen–deuterium exchange mass spectrometry measurements of the Fab fragment of NISTmAb. *Anal. Chem.* **91**, 7336–7345 (2019).

192. Pan, J., Zhang, S. & Borchers, C. H. Comparative higher-order structure analysis of antibody biosimilars using combined bottom-up and top-down hydrogen-deuterium exchange mass spectrometry. *Biochim. Biophys. Acta - Proteins Proteomics* **1864**, 1801–1808 (2016).

10 APPENDICES

APPENDIX I: SEQUENCE OF THE ANTIBODIES INVESTIGATED

NIST mAb IgG

Light chain:
DIQMTQSPSTLSASVGDRVTITCSASSRVGYMHWYQQKPGKAPKLLIYDTSKLASGVPS
RFSGSGSGTEFTLTISSLQPDDFATYYCFQGSGYPFTFGGGTKVEIKRTVAAPSVFIFPPS
DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Heavy chain:
QVTLRESGPALVKPTQTLTLTCTFSGFSLSTAGMSVGWIRQPPGKALEWLADIWWDDK
KHYNPSLKDRLTISKDTSKNQVVLKVTNMDPADTATYYCARDMIFNFYFDVWGQGTT
VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFP
AVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCP
APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREP
QVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF
LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

NIST mAb Fc

THTCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG
VEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPASIEKTISKA
KGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVL
DSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

TM Fc

THTCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG
VEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPASIEKTISKA
KGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVL
DSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG

TM S442C Fc

THTCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG
VEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPASIEKTISKA
KGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVL
DSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLCLSPGK

Fc-C239i

THTCPPCPAPELLGGPSCVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD
GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP

VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Trastuzumab

Light chain:
DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVP
S RFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST
LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Heavy chain:

EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYT
RYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGT
LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTF
PAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCP
APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREP
QVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF
LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

T-scFv40: trastuzumab linked to scFv40, with two TEV and one Factor Xa digestion sites in the linkers (same heavy chain as trastuzumab)

Light chain:

SELTQDPAVSVALGQTVRITCQGDSLRSYYASWYQQKPGQAPVLVIYGKNNRPSGIPD
RFSGSSSGNTASLTITGAQAEDEADYYCNSSDPDQLLVVFGGGTKLTVLGGGGGAGAS
GIEGRGGGGAGAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLE
WVSAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKNYQIHP
FDYWGQGTLVTVSSGSSGAGSGSAENLYFQGSGSAENLYFQGSGGADIQMTQSPSSLS
ASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDF
TLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASV
VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK
VYACEVTHQGLSSPVTKSFNRGEC

T-scFv69: trastuzumab linked to scFv69, with two TEV and one Factor Xa digestion sites in the linkers (same heavy chain as trastuzumab)

Light chain:
SELTQDPAVSVALGQTVRITCQGDSLRSYYASWYQQKPGQAPVLVIYGKNNRPSGIPD
RFSGSSSGNTASLTITGAQAEDEADYYCNSSEPTPPRVVFGGGTKLTVLGGGGGAGASG
IEGRGGGGAGAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLE
WVSAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKNVHIQP
FDYWGQGTLVTVSSGSSGAGSGSAENLYFQGSGSAENLYFQGSGGADIQMTQSPSSLS
ASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDF
TLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASV
VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK
VYACEVTHQGLSSPVTKSFNRGEC

T-dAb: trastuzumab linked to dAb, with two TEV digestion site in the linkers (same heavy chain as trastuzumab)

Light chain:

EVQLVESGGGLVQAGDSLTLSCAASGRTFSSVAMGWFRQAPGKERKFVANISWNGDS
TYYTDSVKGRFTISRDNAKNTVYLQMSSLKPEDTAVYYCAADVRWTGDGHRADYWG
QGTQVTVSSGSSGAGSGSAENLYFQGSGSAENLYFQGSGGADIQMTQSPSSLSASVGDR
VTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSL
QPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNN
FYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACE
VTHQGLSSPVTKSFNRGEC

T-scFvGipg013: trastuzumab linked to scFvGipg013, with two TEV and Factor Xa digestion sites in the linkers

Light chain:

SYVLTQPPSASGTPGQRVAISCSGSNSNIGSNTVHWYQQLPGAAPKLLIYSNNQRPSGVP
DRFSGSNSGTSASLAISRLQSEDEADYYCAAWDDSLNGVVFGGGTKVTVLGGGGGAG
ASGIEGRGGGGAGAQVQLQQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQ
GLEWMGGIIPTFGTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCAQGPIV
GAPTDYWGKGTLVTVSSGSSGAGSGSAENLYFQGSGSAENLYFQGSGGADIQMTQSPS
SLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSG
TDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGT
ASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
KHKVYACEVTHQGLSSPVTKSFNRGEC

APPENDIX II: CALCULATION OF COMPOUND ERRORS

For all the following error calculations, the two variables are such that A ± a and B ± b.

1. Error of A+B or A-B

$$err_{A+B\ or\ A-B} = \sqrt{a^2 + b^2}$$

2. Error of A×B or A/B (with fractional compound error)

$$err_{A\times B} = A.B.\sqrt{(\frac{a}{A})^2 + (\frac{b}{B})^2}$$

$$err_{A/B} = \frac{A}{B}.\sqrt{(\frac{a}{A})^2 + (\frac{b}{B})^2}$$

($err = \sqrt{(\frac{\mathrm{a}}{\mathrm{A}})^2 + (\frac{\mathrm{b}}{\mathrm{B}})^2}$ is called the fractional compound error)

3. Error of ln A

$$err_{\ln A} = \frac{a}{A}$$

4. Error of exp (ln *A*) given ln *A* ± y

$$err_{\exp(ln\ A)} = \exp(\ln A) \times y$$

5. Error of A×c or A/c where c is a constant

$$err_{A\times c} = a \times c$$

$$err_{A/c} = \frac{a}{c}$$

6. Error of c/A, where c is a constant

$$err_{c/A} = \frac{c}{A}.\sqrt{(\frac{0}{c})^2 + (\frac{a}{A})^2}$$

$$err_{c/A} = \frac{c}{A}.\frac{a}{A}$$

$$err_{c/A} = \frac{c.a}{A}$$

APPENDIX III: STABILITY OF MASKED ANTIBODIES (INVESTIGATED IN CHAPTER 6)

The thermal stability of the masked antibodies was investigated. First, the thermal stabilities of the domains of naked trastuzumab were measured. The Fab domain shows a single unfolding transition at 83 °C (**Figure III.1 B**), and the Fc domain shows two unfolding transitions at 74 and 83 °C. As the Fab and C_H3 domains have the same thermal stability, the thermal unfolding of full-length IgG has only two unfolding transitions, the first (at approx. 74 °C) corresponding to unfolding of the C_H2 domain, and the second (at approx. 83 °C) corresponds to the unfolding of both the C_H3 and Fab domains.

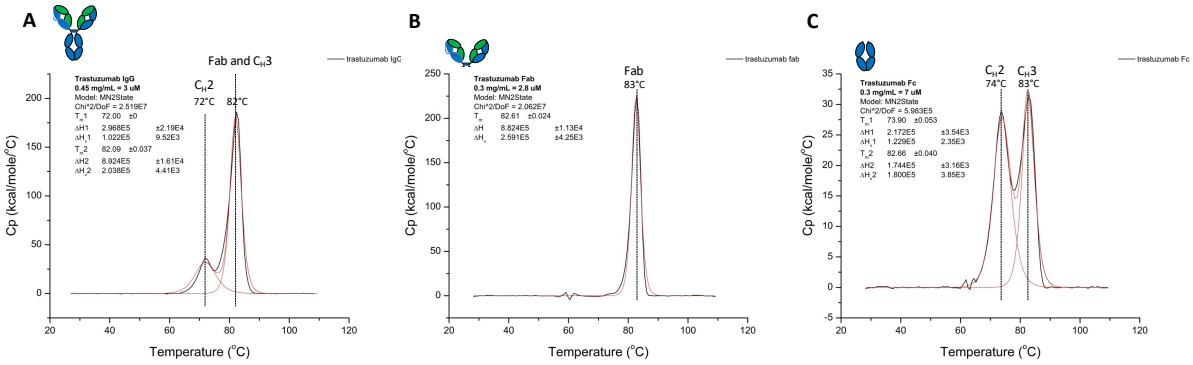


Figure III.1: Thermograms of the unfolding of different domains of naked trastuzumab. A. Full-length IgG. **B.** Fab fragment. **C.** Fc fragment**.** The fragments were generated by digestion of full-length IgG by FabRICATOR which cleaves below the hinge, generating a two-armed Fab species and an Fc fragment without the hinge region. The experiments were carried out in 50 mM Tris pH 7.5.

The thermal stability was also measured for the masked antibodies. The melting temperature of dAb showed a stability of 67 °C (**Figure III.2 A**), however, when the dAb was covalently tethered to the trastuzumab Fab, the melting temperature of the mask increased by 15 °C, to 82 °C, while that of the Fab increased to 88 °C. This observation suggests that the very high affinity of dAb for trastuzumab, which is in the nanomolar range, stabilises both proteins significantly, leading to higher melting temperatures. It is worth noting that the stabilities of the C_H2 and C_H3 domains were not affected by the presence of dAb as expected.

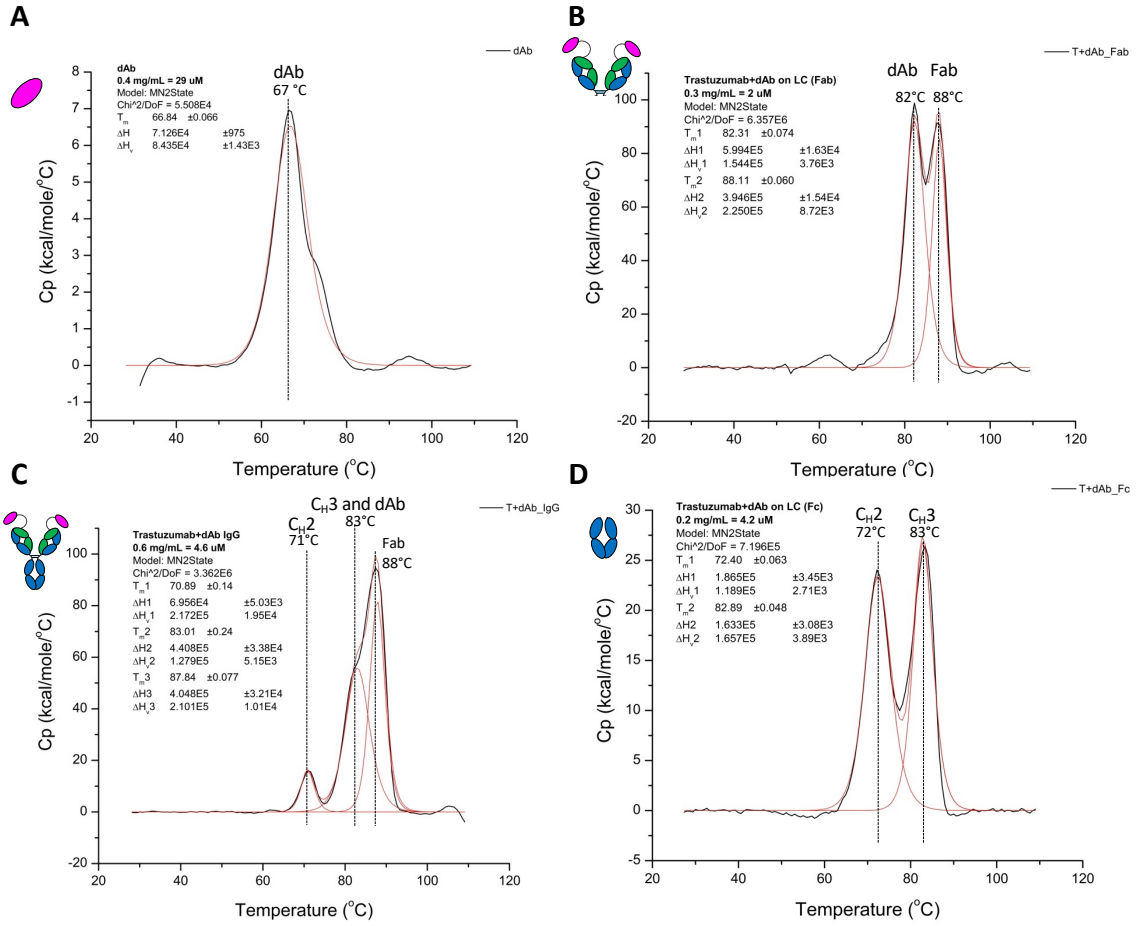


Figure III.2: Thermograms of the unfolding of different constructs of dAb and T-dAb. A. dAb in isolation. **B.** Fab domain covalently linked to dAb. **C.** Full-length IgG T-dAb**. D.** Fc domain**.** The proteins were generated by digestion of full-length T-dAb by FabRICATOR which cleaves below the hinge, generating a two-armed Fab-dAb species and an Fc fragment without hinge. The experiments were carried out in 50 mM Tris pH 7.5.

Regarding the scFv40, the thermal stability of the mask on its own was 57 °C (**Figure III.3 A**), however, when the scFv40 was covalently tethered to the trastuzumab Fab, the melting temperature of the mask increased to 65 °C, and that of the Fab was 78 °C. This observation confirms that the micromolar affinity of scFv40 for trastuzumab stabilised the protein, leading to a higher melting temperature. Surprisingly, the melting temperature of the Fab decreases slightly, this is probably due to the fact that at the higher temperature required to unfold the Fab fragment, the covalently linked scFv has already unfolded and therefore no longer confirms any additional stability on the Fab.

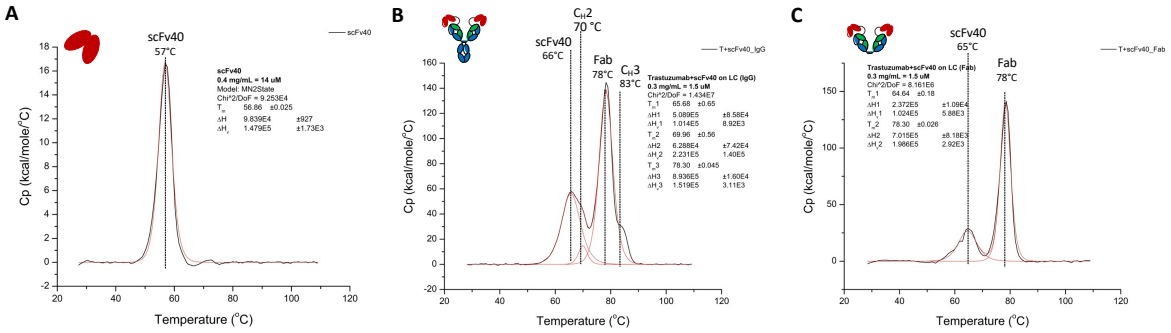


Figure III.3: Thermograms of the unfolding of different *domains* of T-scFv40. A. scFv40 in isolation. **B.** Full-length IgG T-scFv40**. C.** Fab domain covalently linked to scFv40. The proteins were generated by digestion by FabRICATOR which cleaves below the hinge, generating a two-armed Fab-scFv40 species and an Fc fragment without hinge. The experiments were carried out in 50 mM Tris pH 7.5.

The thermodynamic stabilities of the scFv40 and dAb were also measured using chemical denaturants. Both proteins unfolded with a single reversible unfolding transition and data were fit to a two-state model. scFv40 had a lower midpoint of denaturation ($[den]_{50\%}$ = 1.9 M) than dAb ($[den]_{50\%}$ = 2.5 M), however, scFv40 had a higher *m*-value than dAb (**Table III.1**); consequently, both proteins have similar overall thermodynamic stabilities ($\Delta G_{D\text{-}N}$), scFv40 is even a little higher than the dAb. In addition, measurements of the curves at 37 °C (**Figure III.4**, **Table III.1**) suggest that the stability of dAb at 37 °C is very similar to that at 25 °C.

It is interesting to note that the scFv40 is more thermodynamically stable at 25 °C than dAb, but dAb unfolds at a higher temperature. This may be due to the fact that 25 °C might not be the maximal stability temperature for dAb, and that scFv40 and dAb may well differ in the temperature at which they have maximal stability. Furthermore, given scFv40 and dAb have different structures (one is an assembly of two Ig-like domains (scFv40), whilst the other is a single domain), they therefore may have differences in solvent accessible surface area between native and denatured state which would lead to a difference in their *m*-values, as well as differences in the heat capacity change on unfolding (ΔC_p). The latter affects the rate at which the thermodynamic stability changes with temperature which may also explain the observed results.

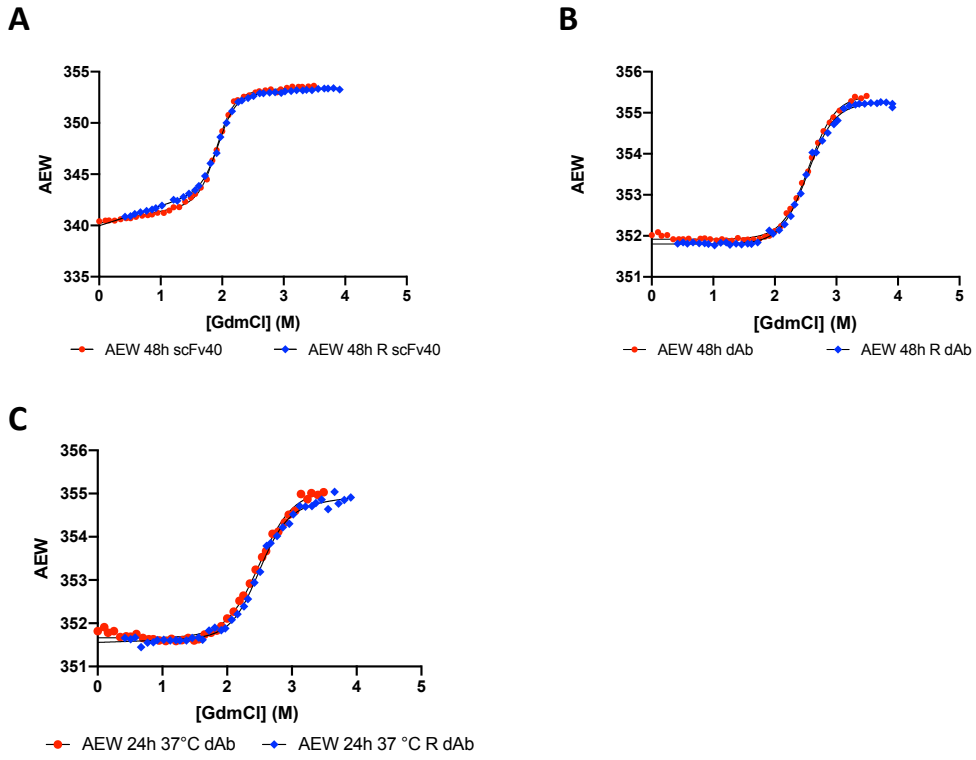


Figure III.4: Chemical denaturant unfolding and refolding curves of dAb and scFv40 masks in isolation. A. Unfolding and refolding curves of scFv40 at 25 °C after 48 h of incubation. **B.** Unfolding and refolding curves of dAb at 25 °C after 48 h of incubation. **C.** Unfolding and refolding curves of dAb at 37 °C after 24 h of incubation. The experiments were carried out in 50 mM Tris pH 7.5.

Table III.1: Thermodynamic parameters from the fitting of the chemical denaturant unfolding and refolding curves of scFv40 and dAb

		$m_{D\text{-}N}$ (kcal mol^{-1} M^{-1})	$[den]_{50\%\ D\text{-}N}$ (M)	$\Delta G_{D\text{-}N}$ (kcal mol^{-1})
scFv40 (25 °C)	unfolding	4.1 ± 0.2	1.91 ± 0.01	7.9 ± 0.3
	refolding	4.2 ± 0.1	1.96 ± 0.01	8.3± 0.3
dAb (25 °C)	unfolding	2.74 ± 0.09	2.55 ± 0.01	7.0 ± 0.2
	refolding	2.70 ± 0.09	2.53 ± 0.01	6.9 ± 0.2
dAb (37 °C)	unfolding	2.3 ± 0.1	2.50± 0.02	5.7 ± 0.3
	refolding	2.7 ± 0.2	2.48± 0.03	6.8 ± 0.6

APPENDIX IV: CALCULATION OF THE FRACTION DENATURED

The denatured fraction is calculated as:

$$f([D]) = \frac{[D]}{[D] + [N]}$$

$$f([D]) = \frac{1}{1 + \frac{[N]}{[D]}}$$

And as $K_{D-N} = \frac{[D]}{[N]}$

$$f([D]) = \frac{1}{1 + \frac{1}{K_{D-N}}}$$

Given that $\Delta G_{D-N} = -RTlnK_{D-N}$

Therefore: $exp\left(-\frac{\Delta G_{D-N}}{RT}\right) = K_{D-N}$

$$f([D]) = \frac{1}{1 + exp\left(\frac{\Delta G_{D-N}}{RT}\right)}$$

From the fittings in Chapter 3, for NIST mAb, $\Delta G_{D-N}^{CH2} = 6.6 \pm 0.2$ kcal mol^{-1}, $\Delta G_{D-N}^{CH3} = 12.2 \pm 0.9$ kcal mol^{-1} and $\Delta G_{D-N}^{Fab} = 12 \pm 2$ kcal mol^{-1}, therefore:

$$f([D])_{CH2} = \frac{1}{1 + exp\left(\frac{6.6}{0.592}\right)} = 1.4 \times 10^{-5}$$

$$f([D])_{CH3} = \frac{1}{1 + exp\left(\frac{12.2}{0.592}\right)} = 1.1 \times 10^{-9}$$

$$f([D])_{Fab} = \frac{1}{1 + exp\left(\frac{12}{0.592}\right)} = 1.6 \times 10^{-9}$$

Thus, 1.4×10^{-3} % of C_H2 domain, 1.1×10^{-7} % of the C_H3 domain and 1.6×10^{-7} % of the Fab fragment are denatured under native conditions.