



**This electronic thesis or dissertation has been
downloaded from Explore Bristol Research,
<http://research-information.bristol.ac.uk>**

Author:

Frank, Bettina

Title:

Effects of organic cosolvents on a de novo designed heme peroxidase

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>. This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Effects of organic cosolvents on a *de novo* designed heme peroxidase



Bettina Frank

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree Doctor of Philosophy in the Faculty of Life Sciences

School of Biochemistry, March 2020

Word Count: 46,291

Abstract

“What I cannot create, I do not understand” – Richard Feynman

While these famous words may stem from a physicist, they appear to have inspired scientists from a wider range of backgrounds: one of the key objectives of synthetic biologists is the design of functional proteins, a field which is said to have recently ‘come of age’ (1). The maquette approach provides a rational way of designing proteins without the complexity of their natural counterparts. The Anderson group have employed this method to produce a variety of four-helix bundle proteins, including C45, a heme-binding peroxidase with impressive thermostability (2).

This work focusses on the structural and functional impact of organic cosolvents on C45. The secondary structure was analysed by circular dichroism in concentrations of 10-80% acetonitrile, methanol, ethanol, isopropanol and TFE. A slight increase in helicity could be observed in up to 60% of any of the alcoholic cosolvents. In high (up to 80%) TFE concentrations, this was particularly pronounced, confirming the ability of the cosolvent to stabilise α -helices (3–5). NMR spectroscopy also suggested that TFE may have a stabilising effect on the enzyme.

Further, the effect of the cosolvents on the kinetics of the reaction between C45, ABTS and H_2O_2 was investigated under limiting peroxide concentration. While the general trend observed was a decrease in kinetic activity, in TFE, the activity increased. In the highest tested cosolvent concentration of 80% TFE, the k_{cat} showed a 7-fold increase compared to that in buffer. Total turnover numbers increased even more significantly, by a factor of 34. This was highly unusual. A smaller increase in

kinetic activity upon the addition of TFE is not unprecedented (6), however not commonly in concentrations above 30%.

This was investigated further using stopped flow spectrophotometry. The reaction did not reach saturation and no k_{cat} or K_{M} for H_2O_2 could be determined, however an estimate suggests the $k_{\text{cat}}/K_{\text{Ma}}/K_{\text{Mb}}$ increased 4-fold compared to the reaction in buffer. Further, it was observed that both in buffer and 80% TFE, $v_0/[E]_0$ increased significantly with increasing enzyme concentration. This very unusual observation was attributed to different levels of catalytic activity present in the dimer and monomer of C45. The presence of both of those was confirmed by analytical SEC as well as SEC-SAXS.

The mechanism of the formation of compound I, the radical intermediate in the catalytic cycle of heme peroxidases, was investigated using stopped-flow spectrophotometry of the reaction between C45 and H_2O_2 . First, the kinetic isotope effect was exploited to confirm which phase of this reaction corresponded to compound I formation. This process was shown to be significantly faster in 80% TFE than buffer. This leads to the conclusion that stabilisation of the radical intermediate compound I plays an important role in the TFE-mediated increase in catalytic activity of C45.

This work could provide a route to improving the catalytic activity of small heme peroxidases, which may lay the foundation of novel applications of maquette proteins, especially with the possibilities offered by the rapidly growing research field of enzyme engineering. This is especially relevant in light of the industrial interest in non-aqueous enzymology (7).

Acknowledgements

Thank you very much to Ross for supervising me and helping me learn and grow, for your kindness and understanding, your sense of humour, optimism, generosity and patience. I have thoroughly enjoyed (most of) my time in this lab and the wisdom you have imparted on me as well as your infectious enthusiasm for the wonders of science. I have also always enjoyed your very balanced and politely phrased political views. Thank you for the opportunity to learn from a wonderful scientist and person.

Secondly, I'd like to thank the BCFN for selecting and funding me and especially Annela who I've got the utmost respect for – I will never understand how you lead the BCFN, teach, do research and outreach and still always have time for your students. Thank you for your support and kind words of encouragement.

Als nächstes well ich mich bei Adrian bedanke. Adrian, janz villen Dank för dinge Hölp – ich han verdamp vill vun dir jeliert. Danke uch för de ville Verzällche, de Ermodijung un dass du nie de Jeduld verlore hatts. Un natörlich de jute Musik in et Bürro, besonders de Leeder öwer Nösse un Jemös. Danke och dafür, dass du immer ding Bier un dinge Schokolad deile däts. Villen, villen Dank!

Thanks to former and current Anderson group members, Adrian, Angelique, Claire, Dan, George, Jack, Jonnie, Katie, Kris, Paulina and Richard – it has been a pleasure to work with such kind, smart and helpful people. A special mention goes to my puta Richard whose scientific expertise, practical advice, IT support and friendship are very appreciated. I can't wait for a game of stickywhack on our beloved

Aalschokker. Thanks also to the Andersons' own Adrian Mulholland as well as The Crow for spreading ataraxy in our office

Thank you also to everyone else working in C101 – not only the fantastic students and postdocs, but also many thanks to our wonderful technicians. Thanks also go to Peter Wilson for his help with equipment and Graham Day and Kris Le Vay for their SAXS-pertise.

Thanks to my lovely and talented BCFN cohort for making first year a lot of fun and always having interesting, funny and helpful things to say, especially to Julia who has been a great friend supporting me in most ways (including physically, when meningitis rendered me unable to walk without help).

I would also like to thank the NHS with all its hardworking and dedicated staff and especially Dr Jenny Schaefer for helping me deal with a variety of health issues since the start of my PhD.

I am lucky enough to have some pretty great friends who have not only always been fun to be around, but also given practical advice, moral support, distractions, interesting conversations and some good laughs – I'd like to thank all the jugglers, hockey players, au pairs, kindergarden, school and uni friends and all the others who have kindly stood by my side and been there for me in the last few years.

Als letztes muss ich mich bei meiner wunderbaren Familie bedanken für moralische und finanzielle Unterstützung, für die Fürsorge während meiner Krankheit, für wunderschöne Osterurlaube und Besuche, für lange Telefongespräche, für Versorgungspäckchen, für meinen alljährlichen Adventskalender und Lebkuchen, für Bristmas 2019 und so viel mehr. Danke, dass ihr immer für mich da seid.

In loving memory of my wonderful grandparents.

An expert is [somebody] who has made all the mistakes which can be made, in a narrow field.

-Niels Bohr

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED:.....DATE:.....

Contents

Abstract.....	I
Acknowledgements.....	III
Author's Declaration.....	VI
Contents	VII
List of Abbreviations.....	XI
List of Figures	XV
List of Tables	XXI
1. Introduction	1
1.1. Historical perspective.....	1
1.2. Protein folding.....	3
1.2.1. Protein structure.....	3
1.2.2. Levinthal's paradox and Anfinsen's dogma.....	4
1.2.3. Molecular interactions governing protein folding	7
1.2.4. Thermodynamic <i>versus</i> kinetic control of protein folding.....	9
1.2.5. Different models of protein folding	10
1.2.6. Energy landscapes.....	12
1.2.7. Marginal stability and the effect of temperature	14
1.3. Enzymes in organic solvents	18

1.3.1.	Challenges in non-aqueous enzymology	20
1.3.2.	TFE as a cosolvent	24
1.4.	Enzyme catalysis.....	27
1.4.1.	Substrate binding	28
1.4.2.	Enzyme kinetics	30
1.5.	Cytochrome structure and chemistry	33
1.5.1.	Cytochromes as Peroxidases	35
1.6.	Enzyme Engineering	39
1.6.1.	Directed Evolution.....	39
1.6.2.	<i>De novo</i> enzyme design.....	40
1.6.3.	The maquette approach	42
1.7.	C45 – a <i>de novo</i> designed four-helix bundle	43
1.8.	Aims of this work	48
2.	Materials and Methods.....	49
2.1.	Materials	54
2.2.	Methods.....	54
2.2.1.	Protein Expression.....	55
2.2.2.	Cleavage of the His-Tag and further purification.....	56
2.2.3.	Size exclusion chromatography	57
2.2.4.	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS- PAGE).....	57
2.2.5.	UV-visible spectroscopy.....	57

2.2.6.	Pyridine Hemochrome.....	58
2.2.7.	Circular Dichroism spectroscopy.....	58
2.2.8.	Nuclear Magnetic Resonance	59
2.2.9.	Small Angle X-ray Scattering	59
2.2.10.	Steady-state kinetic analysis.....	60
2.2.11.	Kinetic analysis with varying substrate concentrations.....	60
2.2.12.	Data analysis.....	61
3.	Structural characterisation of C45 in organic cosolvents.....	63
3.1.	Introduction.....	63
3.2.	Results.....	69
3.2.1.	Expression and purification of C45.....	69
3.2.2.	Pyridine Hemochrome.....	73
3.2.3.	UV-visible spectroscopy.....	73
3.2.4.	Circular Dichroism of C45 at room temperature.....	76
3.2.5.	CD Thermal denaturation studies	79
3.3.	Discussion	106
4.	Functional characterisation of C45 in organic cosolvents	111
4.1.	Introduction.....	111
4.2.	Results.....	116
4.2.1.	Michaelis-Menten Kinetics of C45 with low substrate concentrations.....	116
4.3.	Discussion	131

5. Further kinetic and structural analysis of C45 in TFE	137
5.1. Introduction.....	137
5.2. Results.....	139
5.2.1. Kinetics of C45 with high substrate concentration.....	139
5.2.2. Mechanism of compound I formation	143
5.2.3. Steady-state kinetics of Horseradish Peroxidase with low substrate concentrations.....	148
5.2.4. Metallo-carbenoid formation by C45.....	151
5.2.5. Nuclear Magnetic Resonance (NMR) spectroscopy of C45	152
5.2.6. Small Angle X-ray Scattering of C45.....	155
5.2.7. Analytical Size Exclusion Chromatography of C45	164
5.3. Discussion	168
6. Conclusions.....	171
7. Appendix	175
7.1. C45 and ABTS batch to batch testing.....	175
7.2. SAXS parameters.....	177
8. References.....	179

List of Abbreviations

ϵ	Molar extinction coefficient
δ	Error in a variable
Δ	Quantitative difference or change
λ	Wavelength
λ_{\max}	Wavelength at which maximum absorbance occurs
θ	Molar residue ellipticity
1D, 2D, 3D	1-, 2- or 3-dimensional
ABTS	2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
ACN	Acetonitrile
AU	Arbitrary units
BCFN	Bristol Centre for Functional Nanomaterials
BL21	<i>E. coli</i> strain used for protein expression
C4	A <i>de novo</i> designed four-helix bundle protein binding a bis-histidine ligated heme C on the fourth helix
C45	A protein based on C46, with one mutation (H46F)
C46	A protein based on C4, with two mutations (H11F and H83F)
Ccm	The <i>E. coli</i> cytochrome <i>c</i> maturation System I
CD	Circular Dichroism Spectroscopy
cryo-EM	Cryogenic electron microscopy
CV	Column volume
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
$[E]_0$	Initial enzyme concentration
<i>E. coli</i>	<i>Escherichia coli</i>
EDA	Ethyl diazoacetate

EDTA	Ethylenediaminetetraacetic acid
EPSRC	Engineering and Physical Sciences Research Council
EtOH	Ethanol
G	Gibbs free energy
H	Enthalpy
His	Histidine
HMQC	Heteronuclear Multiple Quantum Coherence
HRP	Horseradish peroxidase
HSQC	Heteronuclear Single Quantum Coherence
IPA	Isopropanol
IPTG	Isopropyl β -d-1-thiogalactopyranoside
k_{cat}	Turnover number, i.e. maximum rate of a reaction
KIE	Kinetic isotope effect
K_M	Michaelis constant, i.e. substrate concentration at which the rate of a reaction is half of k_{cat}
LB	Lysogeny broth medium
MeOH	Methanol
MRE	Molar residue ellipticity
NMR	Nuclear magnetic resonance spectroscopy
OD ₆₀₀	Optical density measured at 600 nm
PDA	Photodiode array detector
PDB	Protein Data Bank
P	Product
Phe	Phenylalanine
pH	Acidity measurement scale based on the negative logarithm of proton concentration
ppm	Parts per million
RNA	Ribonucleic acid
rpm	revolutions per minute
S	Entropy

List of Abbreviations

s	Second(s)
SAXS	Small angle x-ray scattering
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SOFAST	Band-Selective Optimized Flip Angle Short Transient
St. Dev.	Standard deviation
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus
TFE	2,2,2-Trifluoroethanol
T	Temperature
T_m	Melting temperature
TROSY	Transverse relaxation-optimised spectroscopy
Trp	Tryptophan
TTN	Total turnover number
UoB	University of Bristol
UV	Ultraviolet
v_0	Initial rate of a reaction
v_{max}	Maximum rate of a reaction

List of Figures

Figure 1.1: Free energy diagram of protein folding.	6
Figure 1.2: The orientations of a water molecule in bulk water and in the proximity of a non-polar, hydrophobic group.	8
Figure 1.3: Different proposed mechanisms for protein folding.	12
Figure 1.4: Different protein energy landscapes.....	14
Figure 1.5: Cooperative thermal denaturation of a protein.	16
Figure 1.6: Diagram visualising the molecular memory effect of enzymes in anhydrous solvents owed to their higher rigidity.	22
Figure 1.7: Structural formula of TFE.	24
Figure 1.8: Free energy profiles of an uncatalysed reaction and a catalysed one...	28
Figure 1.9: The Michaelis-Menten model of enzyme kinetics.....	30
Figure 1.10: Michaelis-Menten curve showing reaction rate versus substrate concentration.....	31
Figure 1.11: Kinetic models for a reaction between an enzyme and two substrates.....	33
Figure 1.12: Chemical structures of heme B (left) and heme C (right).	35
Figure 1.13: An overview of classification of natural peroxidases.	37
Figure 1.14: Structure of horseradish peroxidase C1A.....	38
Figure 1.15: Schematic diagram of the protein sequence space.....	41
Figure 1.16: The design of heme-binding four helix bundle maquette protein C45.	45
Figure 1.17: Cartoon representation C45 with an inset showing the environment of the heme cofactor in detail	46

Figure 2.1: Size-based separation of three species by size exclusion chromatography.	50
Figure 2.2: Schematic representation of a SAXS experiment.	52
Figure 2.3: Schematic representation of a stopped-flow spectrometer.....	54
Figure 2.4: Sequence of C45.	56
Figure 3.1: Circularly polarised light and its interaction with optically active matter.	67
Figure 3.2: Example of CD spectra of proteins with different secondary structures.	68
Figure 3.3: Chromatogram from size exclusion chromatography showing the absorbance at 280 nm.....	69
Figure 3.4: Analysis of C45 at different stages of its purification and different chromatography fractions.....	72
Figure 3.5: Pyridine hemochrome of C45.....	73
Figure 3.6: UV-visible absorption spectra of C45 in aqueous-organic solutions.....	75
Figure 3.7: Wavelength of Soret peak for C45 in different cosolvents at increasing concentrations.....	76
Figure 3.8: Declining scattering of C45 in 80% ethanol compared to 80% TFE.	76
Figure 3.9: Room temperature CD spectra of C45 in aqueous solution with organic cosolvents.	78
Figure 3.10: Molar residue ellipticity of C45 in increasing cosolvent concentrations.....	79
Figure 3.11: CD thermal denaturation spectra of C45 in buffer.	80
Figure 3.12: CD scan of C45 in aqueous solution with acetonitrile during heating and cooling.....	82
Figure 3.13: Thermal denaturation curves of C45 in aqueous solution with acetonitrile.....	83

Figure 3.14: CD spectra of C45 in aqueous solution with acetonitrile during heating and re-cooling.	86
Figure 3.15: CD measurements of C45 in buffering solution with 70% acetonitrile.	87
Figure 3.16: CD scan of C45 in aqueous solution with methanol during heating and cooling.....	89
Figure 3.17: Thermal denaturation curves of C45 in aqueous solution with methanol.	90
Figure 3.18: CD spectra of C45 in aqueous solution with methanol.....	91
Figure 3.19: CD scan of C45 in aqueous solution with ethanol during heating and cooling.....	93
Figure 3.20: Thermal denaturation curves of C45 in aqueous solution with ethanol.	94
Figure 3.21: CD spectra of C45 in aqueous solution with ethanol.....	95
Figure 3.22: CD measurements of C45 in buffering solution with 80% ethanol.....	96
Figure 3.23: CD scan of C45 in aqueous solution with isopropanol during heating and cooling.....	98
Figure 3.24: Thermal denaturation curves of C45 in aqueous solution with isopropanol.....	99
Figure 3.25: CD spectra of C45 in aqueous solution with isopropanol.	100
Figure 3.26: CD measurements of C45 in buffering solution with 70% isopropanol.....	101
Figure 3.27: CD scan of C45 in aqueous solution with TFE during heating and cooling.....	103
Figure 3.28: Thermal denaturation curves of C45 in aqueous solution with TFE. .	104
Figure 3.29: CD spectra of C45 in aqueous solution with TFE.....	105

Figure 3.30: Molar residue ellipticity at 222 nm at 8 °C for C45 in various cosolvents.	108
Figure 3.31: Models for a four helix bundle protein with the same levels of helicity, but different tertiary structure.	109
Figure 3.32: Molar residue ellipticity at 222 nm of C45 in different cosolvents at 55 °C	110
Figure 4.1: The hypothesised catalytic cycle of C45 mediated oxidation of ABTS.	112
Figure 4.2: Kinetics of the oxidation of ABTS by C45 and H ₂ O ₂ at different pH values.....	113
Figure 4.3: The Michaelis-Menten model of enzyme kinetics.....	114
Figure 4.4: The C45-mediated oxidation of ABTS in the presence of hydrogen peroxide.	115
Figure 4.5: Michaelis-Menten kinetics of C45 catalysed ABTS oxidation in aqueous solution.....	117
Figure 4.6: Total turnover numbers for the oxidation of ABTS by C45 and 100 µM H ₂ O ₂ in aqueous solution.	119
Figure 4.7: Michaelis-Menten kinetics of ABTS ^{•+} formation in acetonitrile.....	122
Figure 4.8: Michaelis-Menten kinetics of ABTS ^{•+} formation in methanol.	124
Figure 4.9: Michaelis-Menten kinetics of ABTS ^{•+} formation in ethanol.	126
Figure 4.10: Michaelis-Menten kinetics of ABTS ^{•+} formation in isopropanol.....	128
Figure 4.11: Michaelis-Menten kinetics of ABTS ^{•+} formation in TFE.	130
Figure 4.12: k_{cat} and total turnover number (TTN) in TFE.	131
Figure 4.13: Total turnover number and k_{cat} for different cosolvents.	133
Figure 5.1: A Morse potential energy curve of C-H and C-D bonds.	139

Figure 5.2: Kinetics of ABTS and peroxide turnover by C45 in buffer and 80% TFE.	140
Figure 5.3: Enzyme concentration dependent $v_0/[E]_0$ for the reaction between C45, ABTS and H ₂ O ₂ in buffer (a) and 80% TFE (b).	142
Figure 5.4: Proposed mechanism of compound I formation from C45 and hydrogen peroxide.	143
Figure 5.5: Kinetics of compound I formation.	144
Figure 5.6: Peroxide concentration dependent rate constants for ABTS oxidation by C45 at 5 °C.	146
Figure 5.7: Deuterium substitution kinetic isotope effects for k_1 (KIE1) and k_2 (KIE 2).	146
Figure 5.8: Kinetics of ABTS oxidation by HRP and C45 in the absence and presence of TFE.	150
Figure 5.9: Mechanism of metallocarbenoid formation by reduced C45 (symbolised as Fe ^{II}).	151
Figure 5.10: Kinetics of C45 mediated metallocarbenoid formation at increasing EDA concentration in 75% TFE.	151
Figure 5.11: ¹ H NMR spectrum of C45 in buffer (red) and 80% TFE (blue).	153
Figure 5.12: 2D NMR spectra of C45 in aqueous solution and 80% TFE.	154
Figure 5.13: SEC-SAXS chromatogram of C45 in buffer.	155
Figure 5.14: SAXS scattering curves and P(r) distributions for C45 in buffer.	158
Figure 5.15: Typical shapes of density distribution functions for different particle sizes.	158
Figure 5.16: 3D models of C45 in buffer produced from scattering profiles.	159
Figure 5.17: SEC-SAXS chromatogram of C45 in 80% TFE.	160

Figure 5.18: SAXS scattering curves and P(r) distributions for C45 in 80% TFE. .	161
Figure 5.19: 3D models of C45 in 80% TFE produced from scattering profiles.	162
Figure 5.20: Normalised Kratky plots for C45 in buffer and 80% TFE.	163
Figure 5.21: Analytical size exclusion chromatogram of C45.	165
Figure 5.22: Analytical size exclusion chromatograms for C45 in buffer and 80% TFE.	166
Figure 5.23: Ratio of first eluting peak to second eluting peak for C45 in buffer (red) and 80% TFE (blue).	167
Figure 5.24: Rate of ABTS oxidation by the suspected dimer and monomer of C45.	168

List of Tables

Table 3.1: Outline of solvents used in this work and some of their key properties...	66
Table 4.1: Kinetic parameters for the oxidation of ABTS by C45 in aqueous solution.....	117
Table 4.2: Highest k_{cat} values for ABTS oxidation in different cosolvents.....	132
Table 5.1: Kinetic parameters for ABTS and peroxide turnover by C45 in buffer and 80% TFE.....	141
Table 5.2: Kinetic parameters of ABTS oxidation by C45 and horseradish peroxidase (HRP) in buffer and 80% TFE.....	150
Table 5.3: Summary of calculated parameters of the eluting species in SEC-SAXS of C45 in buffer or 80% TFE.....	164
Table 7.1: Table highlighting which batches of C45 and ABTS were used in which kinetic assay experiment in buffer for batch to batch comparisons for C45 and ABTS.....	175
Table 7.2: ABTS ^{•+} formation in buffer for different batches of C45 and ABTS.....	175
Table 7.3: Kinetics of ABTS ^{•+} formation in buffer for different batches of C45 and ABTS.....	176
Table 7.4: Total turnover numbers for kinetic assays of different batches of C45 and ABTS.....	176
Table 7.5: Key SAXS parameters. Table reproduced from (311).	177

1. Introduction

Enzymes are a class of proteins which have catalytic functions without which life would not be possible (8). A variety of intramolecular interactions allow them to spontaneously fold into their native state and carry out their functions (9–12). In nature, this generally occurs in aqueous solution, however, for some of their functions *in vitro*, organic media are highly advantageous, leading to the emergence of non-aqueous enzymology as a research field (7, 13–15). In many cases, structural stability and catalytic activity of enzymes decrease in organic solvents and especially in organic-aqueous mixtures (16). 2,2,2-trifluoroethanol (TFE) has been a notable exception to this, with some evidence showing it can stabilise helical proteins and potentially increase enzymatic activity (3–5). Our understanding of the molecular processes underlying the interactions of proteins and enzymes in particular with TFE and other cosolvents is only slowly improving, being impeded to some extent by the inherent complexity of natural enzymes. Evolutionary naïve maquette proteins provide a more simplistic alternative to this (17–20). This work aims to investigate the effect of different cosolvents on the structure and function of C45, a *de novo* designed peroxidase binding c-type heme (2, 21).

1.1. Historical perspective

Humanity has utilised enzymatic fermentation by microorganisms for a long time. While yeast and other microorganisms have been used in the production of bread or alcoholic beverages since as early as 6000 BC (22), the biochemical processes behind

such fermentations have not been understood until much more recently. This commenced in 1833, with the discovery of catalytic function in diastases, enzymes which catalyse the transformation of starch into sugars (8). 44 years later, Wilhelm Kühne coined the term 'enzyme', which stems from the Greek for 'within yeast' (23), reflecting the belief that fermentation was inseparably linked with the biological organism hosting the enzyme. Therefore, the extraction of enzymes from yeast and the reproduction of their function *in vitro* in "cell-free fermentation", was a significant breakthrough for which Eduard Buchner was awarded the 1907 Nobel Prize in chemistry (24, 25).

From these pioneering studies of enzymes in fermentation, the expanding field of enzymology has developed and expanded vastly, as the importance of enzymes in biology has become increasingly clear: the majority of reactions in biological organisms – including humans - would be too slow to allow life without catalysis (8). A factor that has contributed greatly to the advancement of enzymology is the development of various purification and characterisation techniques for biological macromolecules. In particular, the progress in x-ray crystallography has had a large impact on the field, facilitating the investigation of the three-dimensional structure at a high resolution. Initial successes were the discovery of the structural elements of α -helices and β -sheets by Linus Pauling, before entire protein structures were determined (26). A few years after DNA's helical structure was discovered using x-ray diffraction, Kendrew and Perutz employed the technique to study protein crystals, resolving the complete three-dimensional structures of sperm whale myoglobin and horse hemoglobin (25, 27, 28). In some cases, this technique even provided invaluable insights into the catalytic mechanism of an enzyme through co-crystallisation with a substrate bound to the active site (26).

Since our understanding of enzyme structure and function has evolved, a variety of applications has emerged. In medicine, for instance, enzymes can be used diagnostically in assays, for instance to mark tissue damage, monitor diabetes and for the early detection of cancer (23, 29–32). Additionally, enzymes have also generated considerable commercial interest with applications in many industries, such as the agriculture and food industries where they are often used to produce natural products on an industrial scale. The production of food additives for livestock (33), cycling of nutrients within soil (34), processing and refinement of sugars, dairy products or bread (35) require large volumes of naturally occurring enzymes. Enzymes owe their high demand for *in vitro* and *in vitro* applications to their highly desirable properties, such as stereoselectivity as well as their non-toxic and environmentally friendly nature. Furthermore, their high specificity and catalytic efficiency that typically increases the rate of a chemical reaction by a factor of $10^{10} - 10^{15}$ are difficult to parallel (9, 22, 36–38).

1.2. Protein folding

With the exception of ribozymes, RNA-molecules with catalytic function, enzymes are proteins (39). Enzymes can vary greatly in structure, shape and size. Barnase, for example, is very small with 110 amino acid residues and a mass of 12.4 kDa whereas an animal fatty acid synthase monomer can weigh more than 250 kDa (40–42). Nonetheless, enzymes exhibit the characteristic features of protein structure.

1.2.1. Protein structure

Proteins are polymers of α -amino acids, organic compounds that have a carboxyl and an amine functional group as well as an organic side chain attached to the α -carbon (9, 10). Twenty amino acids are naturally encoded by DNA, all of which are

levorotatory enantiomers (except for the non-chiral glycine). The overall topology of a protein is defined by its primary, secondary, tertiary and quaternary structure. Primary structure refers to the sequence of amino acids that are connected by peptide bonds. The secondary structure describes well-defined, smaller scale architectural motifs such as α -helices and β -sheets. α -helices are commonly found in DNA binding motifs, such as zinc finger proteins and leucine zippers (43, 44). The tertiary structure encompasses the spatial arrangement of the protein in its entirety. While this can be a non-defined topology unique to a single protein, there are also some structures conserved in protein families. Examples for this include β -barrel proteins, which possess a structure formed by multiple β -sheets and are often found in membrane proteins such as porins (45). Similarly, triosephosphate isomerase (TIM) barrels consist of repeated eight units of α -helices surrounded by parallel β -strands and have a variety of catalytic functions (46, 47). Lastly, quaternary protein structure, not exhibited by all proteins, refers to the assembly with other proteins into homo- or heteromeric multi-subunit complex. An example is hemoglobin: in human adults, the most common type of this oxygen carrier is a heteromeric tetramer consisting of two α - and two β -subunits (48).

1.2.2. Levinthal's paradox and Anfinsen's dogma

Protein folding refers to the conversion of unfolded amino acid sequences into a unique three-dimensional conformation characteristic for the specific protein. This process can be described as an equilibrium between the unfolded (U) and folded (F) proteins



where the equilibrium constant K_{folding} for the folding reaction can be described as

$$K_{folding} = \frac{[F]}{[U]} \quad 1.2$$

While we have come a long way since the publication of the first structure of a globular protein in 1958 – with over 150,000 further structures now published in atomistic detail – research into the mechanisms that govern protein folding continues to this day (12, 49).

In the late 1960s, Levinthal postulated that due to the large number of possible conformations, an unfolded polypeptide chain would take monumental amount of time to fold into its native state if arriving there purely by coincidence. Even a very small protein consisting of 101 amino acids would have 3^{100} possible configurations, assuming three rotational conformations for each of the 100 peptide bonds (50, 51). Even if configurational changes happen at picosecond timescales, an unbiased pathway sampling conformations at random would take longer than the age of the universe to fold correctly. Obviously, this is not how protein folding proceeds in nature, where it can be as fast as microseconds (12).

In spite of this paradox, Anfinsen demonstrated that the spontaneous self-assembly of a protein into its functionally active native state can, at least for some proteins, occur in solution without assistance (52–54). This indicates that the information about the three-dimensional structure the protein adopts is inherent to the amino acid chain under its normal physiological conditions, which Anfinsen named the “thermodynamic hypothesis” (55). The laws of thermodynamics state that a favourable reaction requires a negative free energy change ΔG . This is a dependent on changes in enthalpy H and entropy S as well as the temperature T :

$$\Delta G = \Delta H - T\Delta S \quad 1.3$$

Therefore, thermodynamically, the conformation with the lowest free energy is the most favourable, and if there is a unique structure at a significant energy minimum that is kinetically accessible, a protein will self-assemble into this conformation (**Figure 1.1**). This is, however, a thermodynamic trade-off between entropic and enthalpic factors. With the restriction of the conformational freedom of the protein chain, there is a significant loss in entropy, opposing the folding process. The spontaneously occurring folding of proteins into a specific three-dimensional structure therefore has significant enthalpic benefits to overcome this (12).

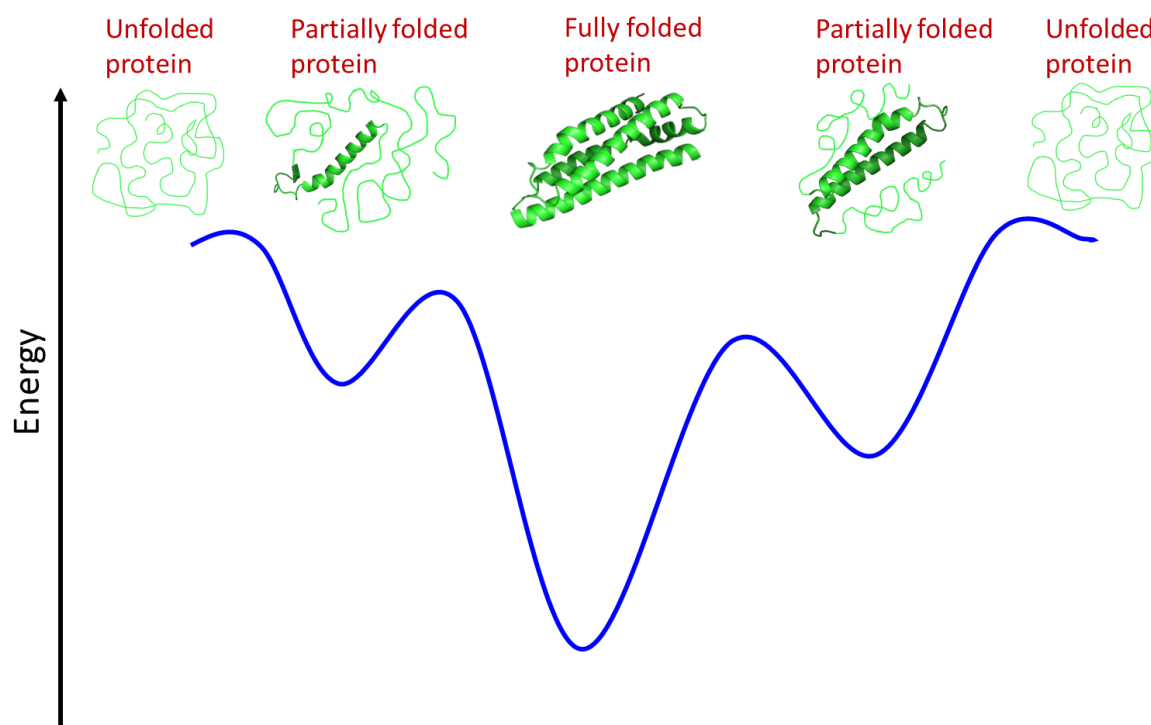


Figure 1.1: Free energy diagram of protein folding.

There are several local minima of free energy where the protein is partially folded and a real minimum for the fully folded protein in its native state. The unfolded chains have the highest free energy and are thus the least thermodynamically favourable states.

1.2.3. Molecular interactions governing protein folding

As previously described by Levinthal more than half a century ago, the protein backbone provides some sort of restriction upon a protein's conformational freedom. The different amino acids in the sequence have varying preferences and levels of flexibility of bond angles. Proline, for example, is significantly more rigid than glycine (1).

However, there are also a multitude of non-covalent interactions between non-neighbouring amino acids. First and foremost, the main driving force for protein folding is the hydrophobic effect, generally lending globular proteins a hydrophilic outer shell surrounding a hydrophobic centre, despite the large variation in protein topology (9, 37). This is due to the tendency of hydrophobic residues to reduce interactions with polar solvents (9, 10). In their natural, aqueous environment, proteins tend to position their hydrophobic amino acid residues on the inside, shielded by polar amino acids which interact with the solvent. This is an entropically driven process resulting from the directional nature of hydrogen bonds. In bulk water, a single water molecule can take any orientation and find a hydrogen bond acceptor. When placed in the proximity of a non-polar group, however, the water molecule will arrange itself so that the hydrogen atoms will face in the direction of polar groups rather than the non-polar group nearby (**Figure 1.2**).

This creates restrictions on the hydrogen atoms' orientation. This lowered degree of freedom results in decreased entropy, which is energetically unfavourable. In order to limit this decrease in entropy, polar and non-polar substances spontaneously separate, which in the case of intracellular proteins means the formation of a hydrophilic shell and a hydrophobic core. This has been exploited for protein design,

where patterns of polar and non-polar amino acids have been used to generate specific structural features, such as α -helices (56).

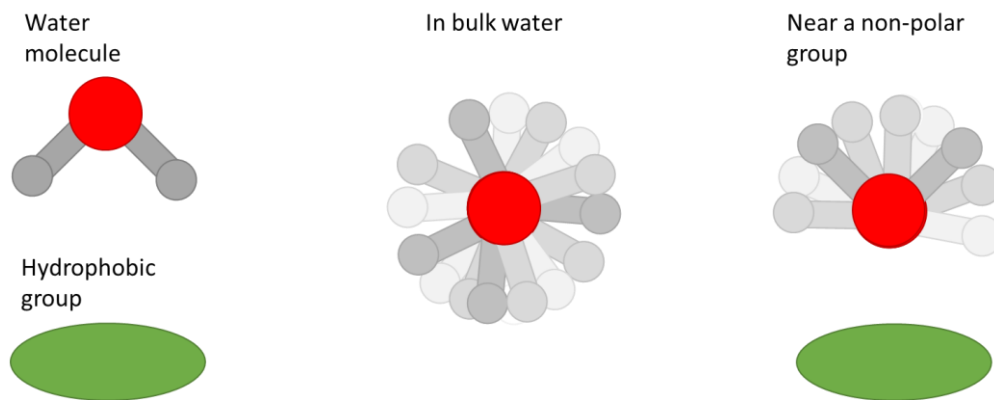


Figure 1.2: The orientations of a water molecule in bulk water and in the proximity of a non-polar, hydrophobic group.

In bulk water, the orientation of the water molecule is not restricted. Near a non-polar group, however, the water molecule has less conformational freedom due to the inability to hydrogen bond with the non-polar group and the resulting increased energy associated with orienting hydrogen atoms towards the non-polar group. This restriction in movement leads to a decrease in entropy. This is unfavourable and explains why proteins engulf non-polar amino acid residues within their core in aqueous solvents. Figure adapted from (10).

While the hydrophobic effect is the primary driving force of protein folding, there are some further intermolecular (or, in the case of macromolecules such as proteins, intramolecular) forces that promote folding giving rise to the structural diversity of proteins. Hydrogen bonds are particularly important, especially for the formation of secondary structural motifs. Both α -helices and β -sheets essentially arise from specific repetitive patterns of hydrogen bonds, with α -helices being based on shorter range interactions than β -sheets (12). Interactions between aromatic groups of delocalised π -electrons known as ' π -stacking' have been shown to be of crucial importance in organic self-assembly such as protein folding. This has also been exploited in *de novo* designed proteins or smaller peptide structures (57, 58). As folded proteins tend to be tightly packed, large molecules, there are also substantial van der Waals interactions

(12). These include dipole-dipole interactions as well as dispersion forces. Whilst generally, these non-directional forces are very weak, for a tightly packed macromolecule with the resulting interatomic contact, they do become significant (9). Especially in such close proximity, electrostatic interactions between positively and negatively charged amino acids also play a part (12). Finally, covalent disulfide bonds between non-neighbouring amino acids can contribute significantly to tertiary structure. Insulin, for example, consists of two different chains linked by disulfide bonds which are important for the structure and function of the hormone (59).

1.2.4. Thermodynamic *versus* kinetic control of protein folding

When Cyrus Levinthal first described the challenge of an amino acid chain successfully folding into a unique, native protein structure, it seemed mutually exclusive for a protein to reach the conformation which would be thermodynamically most stable and to do so in a biologically feasible timeframe (60–62). From this idea, the concept of thermodynamic versus kinetic control of protein folding arose. Thermodynamically controlled protein folding would imply that over time, the protein would reach the conformation at the global energy minimum, independent of the pathway taken to fold. In the case of thermodynamically controlled folding, no defined intermediate states could be observed, merely unfolded and folded protein. Kinetic control, on the other hand, would mean that the protein folds into the kinetically most accessible local minimum at a much quicker timescale. In this case, multiple populated transition states at local energy minima could be observed. These hypotheses motivated a search for intermediates, which would elucidate possible pathways much better.

For some proteins, especially small, monomeric ones, only unfolded and native state, but no intermediates could be observed (63–65). This means that in some cases proteins fold *via* a monophasic two-state pathway, a phenomenon known as cooperative folding. The first protein which was experimentally proven to unfold with complete cooperativity, i.e. *via* a two-state pathway, was barley chymotrypsin inhibitor 2 (CI2) in 1991 (66). This is a simple protein with only 64 amino acids in its sequence, but for most proteins over 100 amino acid residues, the two-state model of completely cooperative folding does not apply (65).

The presence of populated transition states during the folding of the majority of proteins suggests that protein folding is kinetically controlled in most cases. On the other hand, robust proteins have been designed merely on the basis of the computationally predicted stability of the native structure, which suggests a dominance of thermodynamic control (62, 67, 68). Clearly, protein folding is more convoluted and both kinetic and thermodynamic criteria play significant roles in protein folding.

1.2.5. Different models of protein folding

For those proteins that do fold *via* the formation of transition states, a number of folding models have been proposed that take both kinetic and thermodynamic factors into account to suggest different predefined pathways by which the folding processes for different proteins proceeds (**Figure 1.3**).

The framework model (69, 70), also known as the sequential protein folding model or the hierarchical model, is based on the formation of elements of secondary structure that is guided by local molecular interactions and occurs independent of tertiary structure. Elements of secondary structure grow and eventually merge, producing a unique, native structure (71).

Proposed by Karplus and Weaver in 1976 (72), the diffusion collision model describes the immediate formation of small hydrophobic microdomains, which coalesce upon diffusion mediated collision. Similar to the framework model, short range interactions are dominant at least in the initial steps of folding (73).

The hydrophobic collapse model also sees hydrophobic interactions as the initial impetus and main driving force of protein folding, however on a larger scale (61, 74, 75). According to this model, the unfolded molecule immediately forms the hydrophobic core characteristic for proteins by sequestering non-polar amino acid residues (76). This collapsed molecule, sometimes referred to as a 'molten globule', is partially folded and has a much smaller volume than the unfolded molecule. Within this special confinement, short range interactions drive the formation of secondary and tertiary structure.

The nucleation condensation model contains elements of all of the previous models (75, 77). It proposes a mechanism whereby the rate of folding is limited by the formation of a weak local 'nucleus' that consists of some secondary and tertiary structure equal to those in the folded protein. The rest of the unfolded molecule then condenses around this template which guides the structural self-assembly of the protein. Larger proteins can even fold according to this model in multiple smaller components or modules which finally dock to form the protein in its fully folded conformation (67).

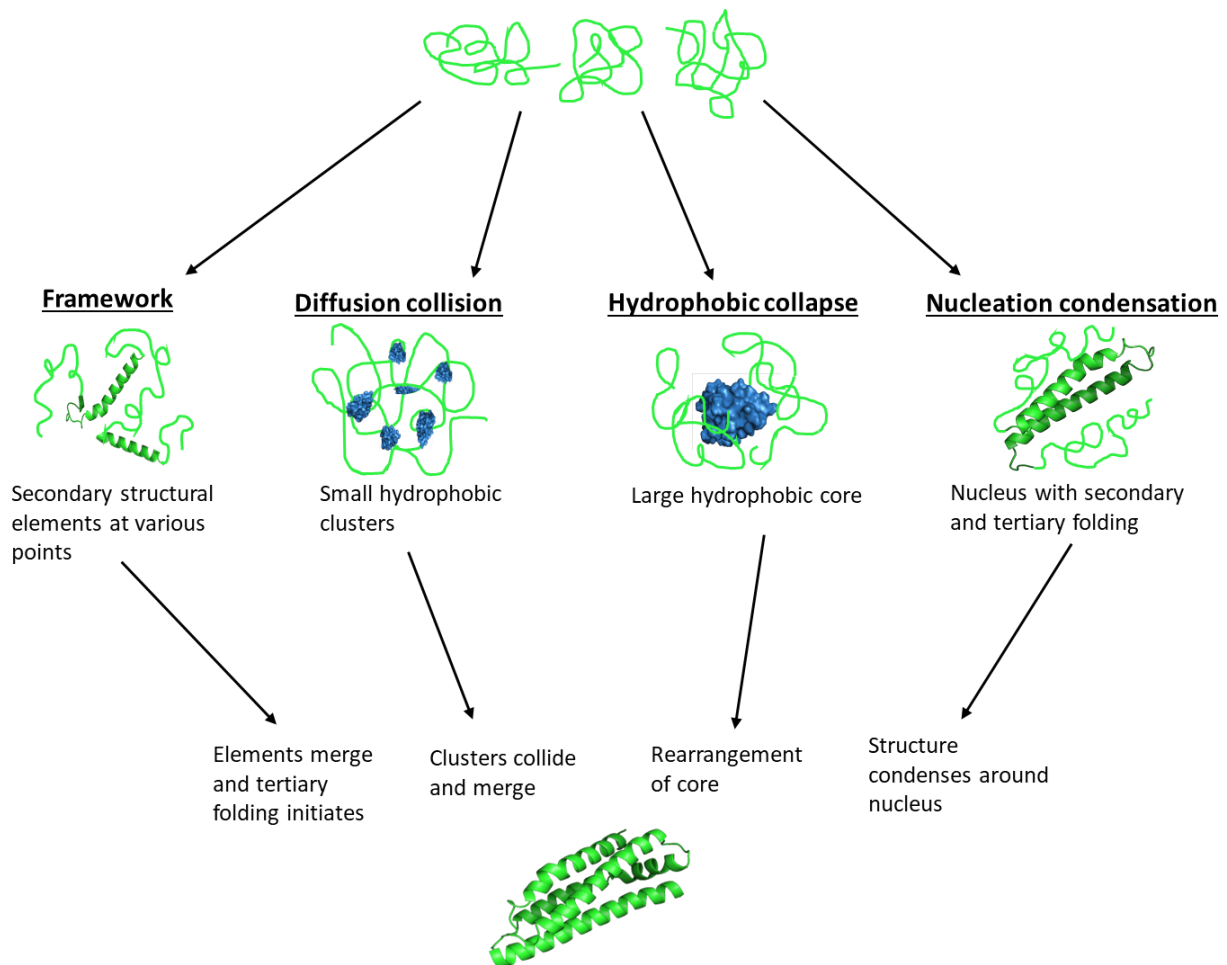


Figure 1.3: Different proposed mechanisms for protein folding.

The different steps involved in the formation of intermediates and the fully folded protein are highlighted in each case.

1.2.6. Energy landscapes

Whilst there are proteins that fold according to each of the above-mentioned pathways, a number of proteins have also been shown to fold according to a mixture of the different models (78). It has furthermore been observed both experimentally and in computational simulations that depending on their environment, proteins may alter their folding pathways and behave according to different models (79, 80). Therefore, the view of the distinct, pre-defined pathways for protein folding has been updated to some extent. Proteins are now believed to fold according to energy landscapes, which

show the different levels of free energy for the different conformations possible (81, 82). Each possible conformation is at a point on the map, with peaks and troughs showing particularly high or low energy conformations. The native structure – i.e. the thermodynamically most stable possible conformation - is at the global energy minimum, which is why this energetic surface is sometimes referred to as an “energy funnel”. Protein folding occurs from the edge of the funnel towards the thermodynamically most stable conformation. From each point on the map, a different energetic pathway may lead to the global energy minimum in the energy landscape’s centre. Every protein is considered to have a unique energy landscape for a specific set of environmental conditions. In a protein that exhibits cooperative folding, no energy barriers exist between the unfolded protein at the edge of the funnel and the native structure, meaning no intermediates exist. Other proteins may fold *via* multiple intermediates which are located in local energy minima. The energy landscape appears more rocky and uneven in this case. This theory explains why folding can proceed by any of the originally proposed models. Merely the energetic favourability of some conformations over others under a set of conditions is decisive and the energetic gradient on the route, rather than a specific folding pathway associated with any protein.

Figure 1.4 shows a variety of different possible energy landscapes, such as the one originally proposed by Levinthal (**Figure 1.4 a**) which is sometimes compared to a golf-course, where the landscape is entirely flat with the exception of the centrally located energy minimum. There is also the possibility of a flat landscape with a clearly defined path to the minimum (**Figure 1.4 b**), an energy landscape that appears like a smooth funnel where no local minima are present and protein folding would occur quickly and cooperatively (**Figure 1.4 c**). A rugged landscape with multiple kinetic

traps, i.e. local minima, is also shown, where proteins would fold with different intermediate states depending on the starting conformation (**Figure 1.4 d**).

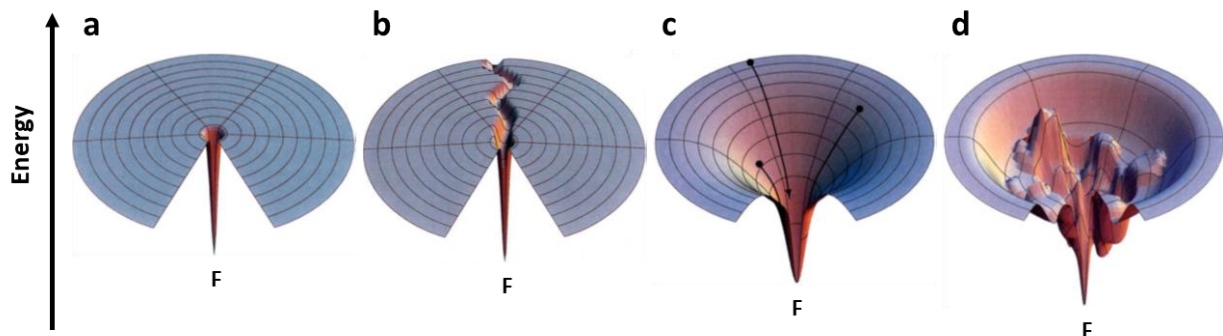


Figure 1.4: Different protein energy landscapes.

a – Levinthal’s originally proposed “golf course” landscape where a protein will eventually reach its fully folded state (F) by sampling conformations at random. **b** – a clear pathway exists from an unfolded to the folded state. **c** – No defined pathway exists, but any unfolded molecule can move down an energy gradient to the folded state. **d** – There are some energetic “traps”, semi-stable transition states, and kinetic barriers in the folding path of any unfolded molecule. Figure adapted from (60).

It is also worth noting that in many cases and especially for large proteins, folding is assisted by molecular chaperone proteins (83, 84). There are different classes, including proteins that prevent the newly synthesised amino acid chain from misfolding and larger, oligomeric structures that envelop unfolded proteins and assist in the folding process. These mechanisms are very important to further increase the rate of folding and prevent misfolding or aggregation in the highly crowded intracellular environment. Incorrectly folded proteins or aggregations can be toxic in biological organisms and have been shown to be linked to diseases such as Alzheimer’s disease (85) and cancer (86).

1.2.7. Marginal stability and the effect of temperature

Protein folding is a delicate thermodynamic balance. Firstly, there is the favourable enthalpic change arising from the stabilising intramolecular interactions. Secondly,

there is the energetically unfavourable loss of chain entropy which opposes the enthalpic gain. This refers to the restriction of conformational freedom of the protein chain by confining it to one specific fold rather than allowing it the myriad of possible conformations possible in an unfolded state. For protein folding to be thermodynamically favourable, of course the enthalpic benefit has to outweigh the high entropic cost, which is estimated to be approximately 70 kcal/mol for a protein with 100 amino acids (87, 88).

The majority of globular proteins have been shown to be marginally stable with a comparably small free energy of unfolding of approximately 5-10 kcal/mol (88, 89). It has been suggested that such marginal stability may facilitate cellular regulation of protein levels and turnover, allowing efficient degradation and recycling of proteins (90). Furthermore, excessive stability could lock proteins in one rigid conformation and thereby hinder conformational changes required for protein functionality such as allosteric activities (91). Similarly, external hydrophobic surfaces, which are thermodynamically unfavourable in aqueous solutions, have been demonstrated to provide some assistance in protein-protein binding (92).

Marginal stability implies that once fully folded, a change in environmental conditions may reverse folding. Temperature is directly related to the free energy of protein folding, as described in equation 1.3. Due to the loss of chain entropy associated with protein folding, ΔS is negative and a temperature increase will eventually render the folding process unfavourable. At the melting temperature T_M , ΔG between folded and unfolded states is 0 and equal amounts of the folded and unfolded protein chains are present (**Figure 1.5**).

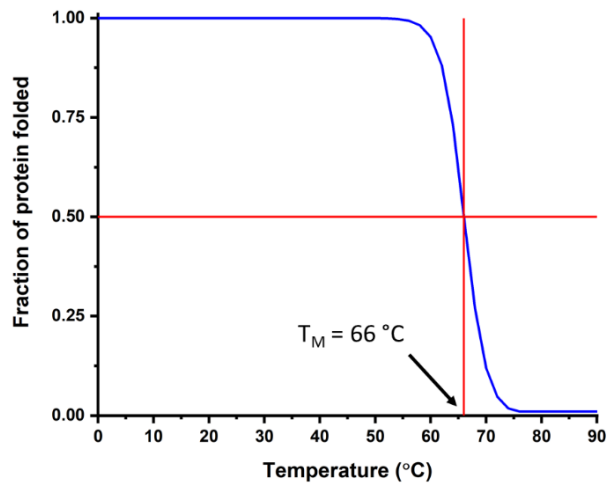


Figure 1.5: Cooperative thermal denaturation of a protein.

At the melting temperature T_M (in this case $66\text{ }^\circ\text{C}$), equal amounts of folded and unfolded protein exist.

When heated, non-covalent intramolecular interactions of proteins such as van der Waals forces and hydrogen bonds are disturbed and the flexibility of the chain increases (93). Some movement is important to the protein and often required for its function, so the native state is not entirely rigid: Even at low temperatures and in the crystalline state, some vibrations and rotations still take place (94). However, with temperature elevation, due to the strong entropic drive, these motions become excessive and less controlled. The protein starts to unfold which is accompanied with a loss of function such as catalytic activity in an enzyme.

In line with Anfinsen's theory that a protein chain will spontaneously fold into its native state, unfolding can be reversible when the solution is cooled and intramolecular forces reinstated, which is often observed for small proteins with simple folds (95). This reversible protein denaturation can occur either instantaneously, or more gradually, such as in multiple discrete steps, depending on the cooperativity of folding and unfolding events (13, 93).

On the other hand, unfolding can also be irreversible, which is particularly observed for larger proteins with complex folds involving many long-range interactions. The processes leading to irreversible denaturation can be monomolecular or polymolecular in nature. Monomolecular denaturation may involve a protein misfolding by adopting a conformation at a local minimum with a kinetic barrier that is too high to refold into its native state. The protein may also undergo covalent changes that prevent refolding. Furthermore, there is the possibility of unfolded proteins to interact with each other, inducing polymolecular irreversible denaturation. As the protein's surface becomes amphiphilic due to the hydrophobic core being exposed, it may aggregate with other unfolded proteins in order to increase solvent entropy, which is particularly likely in highly concentrated protein solutions.

In vivo, with the exception of proteins native to thermophilic organisms, proteins do not normally encounter temperatures above 40 °C. However, with applications of proteins *in vitro*, for example in bioreactors, there has been significant interest in discovering or engineering proteins that are structurally stable and functional in higher temperatures (96). Enzymes, for example, are used in brewing and fermentation processes, in material production and as detergents which makes thermostability a desirable attribute of these biological catalysts (97, 98).

Apart from *in vitro* applications of enzymes, research into thermostability of enzymes is motivated by scientific curiosity and the potential to acquire transferable insights into protein stability. It is also important to note that the effect of temperature on the catalytic activity of an enzyme is two-fold. Whilst denaturation of the protein is associated with loss of catalytic function, this relationship is more convoluted: Chemical reaction rates are correlated with temperature, meaning that an enzyme

catalysed reaction will occur at increasing rates with raised temperatures up to the point where enzyme denaturation occurs (99).

1.3. Enzymes in organic solvents

The chemical environment of an enzyme has a vast impact on its structure and function. Enzymes are highly adapted to their natural environments, where they are either in aqueous solutions, for instance the cytosol, or associated with cellular membranes. *In vitro*, enzymes are thus usually kept in buffered solutions that best mimic their natural environment.

The aqueous nature of the solvent is often of crucial importance as demonstrated by hydration studies (100) as it can allow the protein greater mobility (13). A protein is considered to be fully hydrated when the addition of further water would only cause dilution, but not change the protein's dynamics (100, 101). The water associated with the protein is referred to as the hydration shell, comprising of a small number of water molecules which are closely coordinated in specific locations of the protein such as active sites, the interior or metal ions associated with the protein as well as a monolayer enveloping the protein surface. This hydration shell is important for the dynamics of the protein. It acts as a form of molecular lubricating agent, enhancing the protein's conformational flexibility and allowing it to achieve its active conformation (13, 101–103). Furthermore, as discussed previously, the hydrophobic effect plays a key role in protein folding and in a non-aqueous setting, the formation of a hydrophobic core would be less favourable. This would strongly decrease protein stability.

With the important role of water molecules to proteins, which have, after all, evolved in aqueous environments, it is understandable the idea of non-aqueous enzymology was met with scepticism, initially, and most enzyme work thus far has

taken place in water-based buffering solutions (16, 104). In fact, Christian Anfinsen, 1972 Nobel laureate for his discoveries regarding the folding of amino acid chains into native proteins, said in his prize acceptance speech: “A protein molecule only makes stable, structural sense when it exists under conditions similar to [the] physiological state” (55).

However, there is significant motivation to use enzymes in non-aqueous media: firstly, there is the inherent scientific curiosity, with the possibility of gaining some valuable insights into enzymatic mechanisms and molecular dynamics. There is also some indication that there may be properties, such as specificity and selectivity, that can be modulated and controlled by finetuning the solvent in which enzymatic reactions take place (105–107). This has been shown for substrate selectivity (106, 108, 109), chemo- and regioselectivity (110, 111) and stereospecificity (112–115). It has been proposed that these properties may be related to the changed energetics of substrate binding in solvent: whilst the interactions between the substrate and the enzyme are likely unchanged if the protein retained its conformation, the exclusion of solvent molecules from the binding site will be associated with a different desolvation energy (106, 109).

Additionally, the prevention of contamination by microbial growths facilitates storage of enzymes in non-aqueous media (105). Most importantly, however, a much greater range of organic substrates insoluble in water could be used in biocatalytic reactions as the majority takes place in non-aqueous media. Furthermore, water decomposes some organic reagents or can promote side reactions (116). For these reasons, significant efforts have gone into the use of enzymes in organic solvents, a field which has grown significantly since the 1980s (15, 117–119).

1.3.1. Challenges in non-aqueous enzymology

In general, enzymes are less active in anhydrous organic solvents for a multitude of reasons, although they are not necessarily insurmountable (13, 16). Early work by Klibanov and colleagues showed that two proteases are 10^4 - 10^5 times more catalytically efficient in water than in octane which was the non-aqueous solvent they found to be the most pertinent to catalysis (14). Whilst significantly worse than in water, the enzyme remains catalytically active, with a rate acceleration of 10^{11} compared to that of the reaction in the absence of enzyme. One of the problems of the use of enzymes in organic solvents is the lower solubility of enzymes in organic solvents which leads to the formation of suspensions of enzymes in solvent, rather than solutions, as in the case of aqueous media (16, 120). Therefore, inaccessibility of enzymatic centres and diffusional limitations of the reaction substrates may hinder enzymatic activity, which is not uncommon in heterogeneous catalysis. Theoretical analysis suggested this may be a contributor. However, spectrophotometric titration of enzyme active centres with 1-trans-cinnamoyl imidazole (14) showed it can only be responsible for a small fraction of the activity loss in organic solvents (16, 121). There are, however, strategies to overcome the solubility hurdle: some groups have worked in inverse micelles, where the enzyme is separated from the organic phase in small aqueous pools. Unsurprisingly, the enzymes' observed structural and catalytic properties were not significantly different from those in aqueous solutions and technically, the reaction medium is not neat organic solvent (14, 122, 123). The other possibility is to dehydrate the protein entirely and resuspend the dry protein powder. This is usually achieved by means of lyophilisation, a process also known as freeze-drying, during which water is removed from frozen protein samples *via* sublimation and desorption (105, 124). With the importance of a stable temperature and hydration of

proteins, both freezing and drying are processes that are potentially denaturing of proteins. Generally, the structural motif of β -sheets appears to be more resilient to lyophilisation-induced denaturation than α -helices (105). Whilst protein denaturation by exposure to anhydrous organic solvents was one of the biggest hypothesised problems associated with non-aqueous enzymology, experimental data (16) show that denaturation is almost exclusively associated with lyophilisation rather than the subsequent resuspension. This denaturation can often be prevented with the use of lyoprotectants, such as inorganic salts (especially potassium chloride) (125), crown ethers (126) or sugars (127). Another form of protecting the enzyme's conformation during lyophilisation is the use of detergents or other amphiphilic molecules to form complexes with enzymes (128). These lyoprotection strategies have enhanced the activity of enzymes in organic solvents by up to four orders of magnitudes (16). An important aspect of lyophilisation is that it needs to be done in an aqueous solution of optimum pH for catalysis. Klibanov *et al.* found that the catalytic behaviour of resuspended enzymes mirrors that displayed at the pH of the last aqueous solution in which they were dissolved, a phenomenon sometimes referred to as 'pH memory' (14–16, 129). This is related to the observation that enzymes' ionogenic groups, such as charged amino acids, are less able to change their ionisation state in organic solvents than in aqueous ones (14). Furthermore, in non-aqueous solutions, enzymes have less conformational freedom for two reasons: firstly, the lower dielectric constants of non-aqueous solvents enhance intramolecular electrostatic interactions and thereby rigidify enzymes (130). Secondly, the lubricating properties of the hydration shell allow enzymes some mobility. Even during lyophilisation and resuspension in anhydrous solvents, some of the tightly bound water molecules are retained. In non-polar solvents, this is less likely to be stripped from the enzyme, which explains the observed

positive correlation between solvent hydrophobicity and enzymatic activity (131). Whilst this property is initially a challenge for non-aqueous enzyme catalysis, it has also been exploited to increase enzyme efficiency in anhydrous solvents by locking enzymes in their active conformation through lyophilisation of an enzyme-ligand complex as opposed to the enzyme by itself. Subtilisin, for instance, is up to two orders of magnitude as active in anhydrous solvents when lyophilised in the presence of competitive inhibitors (which are subsequently extracted) than in the absence of these ligands (132, 133). The principle of this ‘molecular memory’ is visualised in **Figure 1.6** (13).

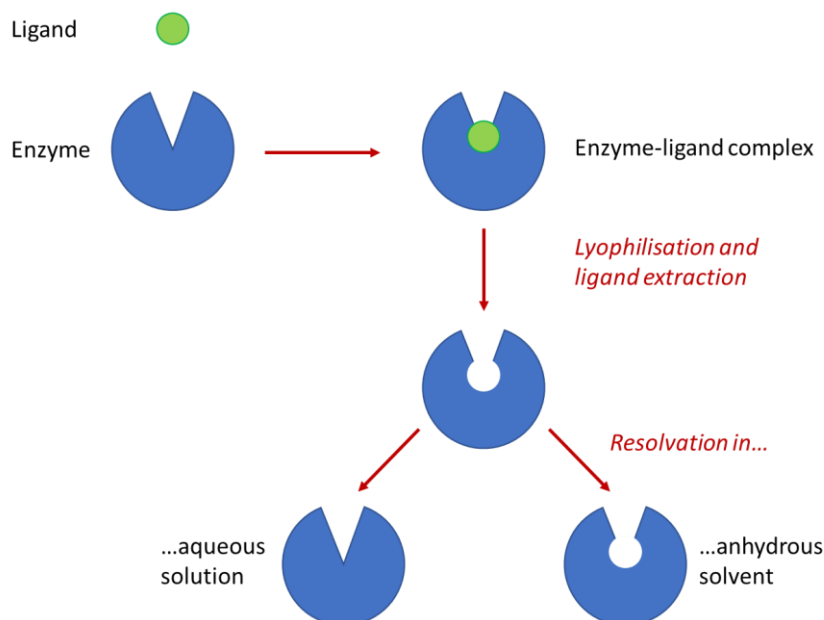


Figure 1.6: Diagram visualising the molecular memory effect of enzymes in anhydrous solvents owed to their higher rigidity.

A further obstacle to enzyme-mediated catalysis in non-polar organic solvents lies in the thermodynamics of substrate binding. The active centre of many enzymes is hydrophobic, and these enzymes tend to work most efficiently with hydrophobic substrates in aqueous solutions. The desolvation of these is much more energetically

favourable in aqueous or polar solvents than in hydrophobic solvents, where the ground state is stabilised. This higher desolvation energy increases the net binding energy which is a strong thermodynamic contributor to enzyme catalysis (16, 134). In one experiment with crystalline cross-linked subtilisin, the lowered desolvation energy in acetonitrile as opposed to water led to a k_{cat}/K_M reduction by a factor of 100 (135).

It is also worth noting that this applies only to anhydrous solvents. With the addition of very small amounts of water, this can be curbed significantly: in one case, in solvent with 0.5% water, the activity is approximately a quarter of that in aqueous solution (131) showing the important distinction between the use of neat organic solvents versus mixtures. Mixtures of organic solvents with aqueous buffers – apart from the aforementioned work in micelles - allow a smaller range of solvents, namely those that are miscible with water. Less attention has been paid to work of enzymes in aqueous-organic mixtures with more than 10% aqueous solution. However, for a few enzymes, some rate enhancements have been observed in cosolvent concentrations of up to 20% without any notable structural changes being observed (136), while others suggest organic-aqueous mixtures could be detrimental to enzyme structure and function (13).

In order to circumvent any possible complications arising from lyophilisation and to potentially observe a more gradual change in enzyme structure and function, the work presented herein investigates the structure and function of a *de novo* enzyme in mixtures of aqueous buffer with a variety of solvents. There are some important distinctions between various organic solvents, the first one being different polarities which in turn effect different levels of miscibility with water. Secondly, solvents can be divided into protic and aprotic, conveying a substance's ability – or inability – to act as hydrogen bond donors. Typically, protic solvents are much better at dissolving salts

than aprotic solvents, making them more suitable for enzymology work in organic aqueous mixtures (137). Depending on the stability of the solvent, a result of intramolecular interactions, they have a range of boiling points.

1.3.2. TFE as a cosolvent

A solvent that has proven particularly interesting for the study of helical proteins is TFE, shown in **Figure 1.7**. TFE possesses both a polar hydroxyl group and a triple halogenated hydrophobic group, making it an amphipathic molecule. This property makes it suitable for stabilising both the hydrophilic surface and the hydrophobic core of proteins (3, 138).

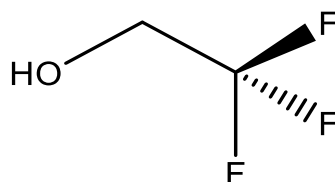


Figure 1.7: Structural formula of TFE.

A helix stabilising effect has been observed in polar alcohols previously (139), however this has been particularly strong in mixtures of water and TFE. Though the cosolvent can also stabilise β -sheets, it has been found to be particularly proficient at solubilising and stabilising α -helices (3–5, 140). In fact, it has been observed that the degree of helicity increases upon addition of TFE to aqueous solution, to the extent that regions which are non-helical in the native state may become helical in TFE-water mixtures. In proteins that are not almost exclusively helical, this can mean that tertiary structure is disturbed (141, 142).

The destruction of tertiary structure in favour of stabilisation of helices is manifestation of TFE strengthening short range interaction significantly more than long

range interactions (3). In line with this, the effects of TFE on proteins are very sequence dependent in that it is challenging to devise a specific scale of helical propensity for amino acids irrespective of their environment (143).

It appears that TFE propagates helices in a cooperative process rather than initiating helix formation, with longer helices particularly likely to be extended (139, 144, 145). TFE therefore does not form a site of helix nucleation, but merely exacerbates the intrinsic helical propensity of a protein, more than aqueous solvents. In 15% TFE, helical propensity of amino acids is still comparable to aqueous solvents, whereas in 40% TFE they are increased for all amino acids (143). This is especially prevalent in polar amino acids; glycine and proline, on the other hand, still act as “helix breakers” (3). 30% TFE concentration in water by volume appears to be an important threshold for protein stability for many proteins, which appear to have reached their maximal helicity at this cosolvent concentration and plateau as TFE concentration is raised beyond this point (4). In some proteins, the stability of helices even decrease upon further addition of TFE (146), although there have been cases where helicity continues to increase (147).

TFE also appears to have a strong effect on protein thermostability. In a study of the first 19 residues of ribonuclease A, the peptide displayed a 36% rise in helicity in 80% TFE compared to aqueous solution at 20 °C (144). This increased by a further 8% when the temperature was lowered to 3 °C. In line with this, temperature elevation has been demonstrated to cause rapid loss of helicity which – in contrast to the propagation of helices upon the addition of TFE – proceeded non-cooperatively with fraying at the termini of the helices (3, 139).

The mechanism for helix stabilisation by TFE is not yet completely understood and different models have been suggested. One idea is that TFE molecules directly bind the protein and exchange the solvation shell (139). No specific binding sites have been suggested, however, at 30% TFE by volume, the protein surface should be saturated, which would provide an explanation for the TFE concentration at which many proteins achieve their maximal helicity (4, 139, 148). Experimental data from ^1H and ^{19}F NMR confirm that TFE molecules associate with the surface of a protein and even penetrate the core (149). A disruption of the water structure would cause a reduction of the hydrophobic effect. This would beg the question whether protein folding – guided by the formation of a hydrophobic core – could take place in this medium or whether only the stabilisation of folded proteins is feasible in TFE (4). However, even with hydrophobic interactions weakened, clusters of hydrophobic sidechains have been observed (150) which are most likely stabilised by stacking of aromatic rings (3, 151).

Others have proposed a less direct interaction of TFE with the protein, a model in which hydrogen bonding is the main driving force for TFE-mediated stabilisation of helices (145). This effect would be two-fold: Firstly, solvent hydrogen bonds to the random coil formation would be decreased, thereby selectively destabilising this conformation (144). Secondly, the displacement of water with a large molecule with fewer hydrogen bond donor and acceptor sites would strengthen intrapeptide hydrogen bonding between carbonyl and amide groups (4, 5, 148). This is supported both experimentally and in molecular dynamics simulations (3).

Another theory was that this phenomenon was mostly dielectric, however this is unlikely as a primary driving force as experimental results show variation of TFE concentration has little impact on interactions between charged residues of the ribonuclease S-peptide (4, 152).

As no one mechanism would fully account for the diverse effects TFE has on proteins, it is likely that elements of different mechanisms occur simultaneously. One study did indeed suggest an intertwined mixture of the proposed mechanisms, where TFE molecules associate with peptides and displace water, providing a low dielectric environment that affects hydrogen bonds without disrupting hydrophobic interactions (5).

The structural stabilisation of proteins can in turn have beneficial effects on their function, such as an enhancement of catalytic activity in enzymes. For instance, the activity of horseradish peroxidase doubled upon the incubation of the enzyme in 5-25% TFE (6). At higher cosolvent concentrations, however, the enzyme denatures and activity declines; in 40% TFE, horseradish peroxidase is entirely inactive (6). Furthermore, increasing TFE concentrations have been shown to lead to an increase in the volume of horseradish peroxidase. This supports the theory that TFE-induced structural changes are largely related to diminishing hydrophobic interactions, which would drive the tight folding of a protein in an aqueous environment (6).

1.4. Enzyme catalysis

Enzymes are very impressive biocatalysts, able to speed up reaction rates significantly with great specificity and selectivity. Without them, most cellular reactions would proceed at rates not conducive to life. In a chemical reaction, the rate constant k depends on the transmission coefficient κ , the Boltzmann constant k_B , the temperature T , Planck's constant h , the universal gas constant R and the Gibbs energy of activation ΔG^\ddagger (153, 154), as shown by the Eyring equation:

$$k = \frac{\kappa k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}} \quad 1.4$$

As **Figure 1.8** shows, reaction rate is dependent on the free energy of the transition state relative to the ground state. Both a catalysed and an uncatalysed reaction will result in the same absolute change in free energy between reactants and products, ΔG° . However, in the catalysed reaction, the transition state is stabilised by the enzyme more than the ground state, resulting in higher kinetic accessibility and the resulting lower Gibbs energy of activation ΔG^\ddagger , thereby increasing the reaction rate.

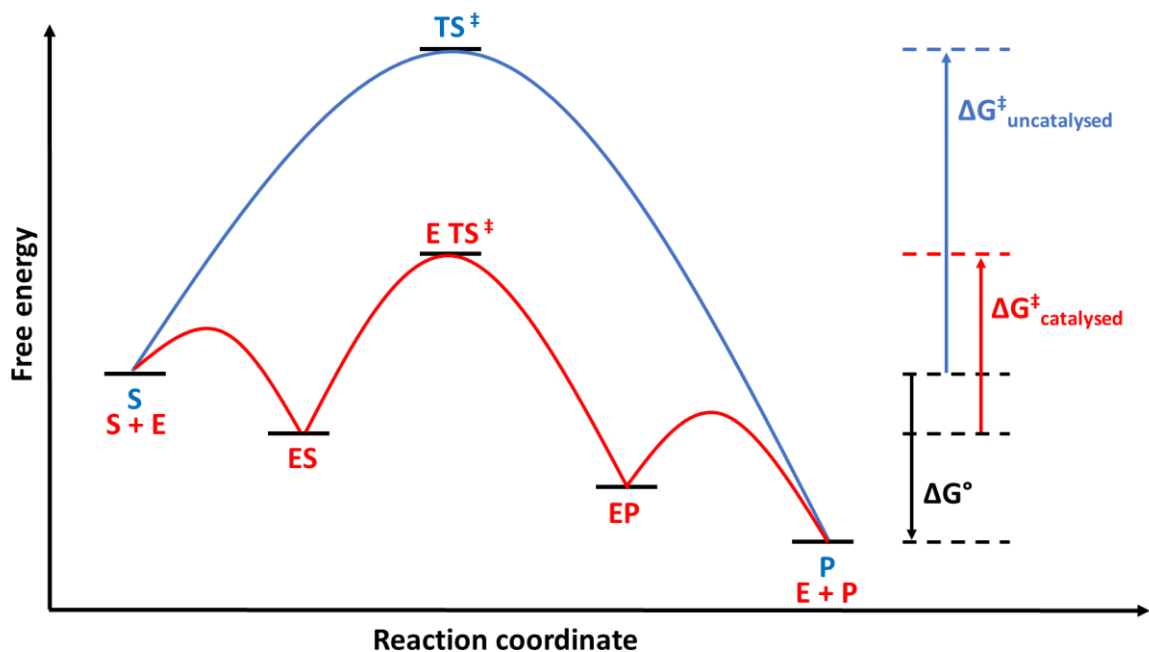


Figure 1.8: Free energy profiles of an uncatalysed reaction and a catalysed one.

In both the uncatalysed reaction (blue) and the catalysed one (red), the total change in free energy (ΔG°) between reactants (either substrate S or S and enzyme E) and products (product P or P and E) is equal. The catalysed reaction proceeds in several steps involving substrate binding, substrate conversion and product release. The transition state TS^\ddagger of the substrate is stabilised by the enzyme, which lowers the Gibbs energy of activation (ΔG^\ddagger) associated with the catalysed reaction compared to the uncatalysed one.

1.4.1. Substrate binding

An enzyme catalysed reaction proceeds in multiple steps (10, 23, 25, 134, 154), the first of which is the binding of its substrate. The structure of the enzyme is of

paramount importance for substrate binding. The high substrate specificity of enzymes is a result of binding pockets which are highly adapted to their substrates. This model for molecular recognition was originally suggested by Emil Fischer who compared this to a lock and key, where extreme specificity is also achieved by shape complementarity (134). This view was updated to account for the flexibility of enzymes, meaning that the binding site is dynamic and somewhat moulded by the substrate during binding, described as induced fit. Furthermore, while topology is important, the actual binding process requires intermolecular interactions between enzyme and substrate. These include electrostatic interactions between charged functional groups of the substrate and oppositely charged amino acid side chains. Of course, charges are very pH dependent, highlighting the importance of buffered solutions for enzyme function. Hydrogen bonding between substrate and enzyme also play a substantial role, as do van der Waals interactions. Lastly, the binding site is often located within a hydrophobic cleft of the enzyme and hydrophobic interactions between a substrate and an enzyme binding site are important in the aqueous environment. As these interactions are strongest at short range, proximity between substrate and enzyme is an important factor, highlighting the relevance of concentration to reaction rates.

While substrate binding is required in advance of enzymatic catalysis, it is impeding catalysis if this interaction is too strong, which is perhaps counterintuitive. However, if the enzyme-substrate complex is bound very tightly, this lowers the free energy, and the reaction is thermodynamically less favourable, lowering the reaction rate. The major driving force underlying catalysis is the stabilisation of the substrate's transition state by the enzyme, which is generally bound more tightly than the substrate (154). The release of the product or products and simultaneous recycling of enzyme is generally such a fast process that the catalytic mechanism between an

enzyme and a substrate is often viewed as consisting of only two steps, the binding of the substrate and the conversion of the substrate to form the product. For most enzymes, the rate limiting step is the conversion of the substrate to form the product. A perfect enzyme, however, could catalyse the reaction as quickly as it collides with the substrate, which is why this concept is referred as diffusion limited.

1.4.2. Enzyme kinetics

Kinetically, it is important to distinguish between single-substrate and multi-substrate reactions. For single substrate reactions, most reported enzymes obey the Michaelis-Menten model (155), although there are some cases where inhibitors or allosteric modulation change the kinetics of the reaction. The Michaelis-Menten model assumes two steps in a reaction between a single substrate and an enzyme, the binding and the conversion of the substrate (**Figure 1.9**).



Figure 1.9: The Michaelis-Menten model of enzyme kinetics.

The catalysis of the conversion of substrate S by enzyme E involves binding resulting in enzyme-substrate complex ES, followed by conversion of the substrate to product P and its dissociation.

During the course of the reaction, the concentration of product increases and the concentration of substrate, of course, decreases. At the commencement of the reaction, no product is present, meaning that the conversion of substrate to product is a unidirectional process and the velocity of the reaction is linear at this stage. A standard curve of a reaction catalysed by an enzyme following Michaelis-Menten kinetics is shown in **Figure 1.10**. The Michaelis-Menten equation can be used to calculate the relationship between this initial rate (v_0), the initial enzyme concentration

1. Introduction

($[E]_0$), the initial substrate concentration $[S]$ as well as the turnover number k_{cat} and the Michaelis constant K_M .

$$\frac{v_0}{[E]_0} = \frac{k_{cat} [S]}{K_M + [S]} \quad 1.5$$

where k_{cat} denotes the molecules of substrate each enzyme can convert per unit time when the reaction proceeds at its maximum velocity.

$$k_{cat} = \frac{v_{max}}{[E]_0} \quad 1.6$$

The Michaelis constant K_M is the substrate concentration where the reaction rate is half of its maximum v_{max} , as shown in **Figure 1.10**. K_M is inversely correlated with binding affinity.

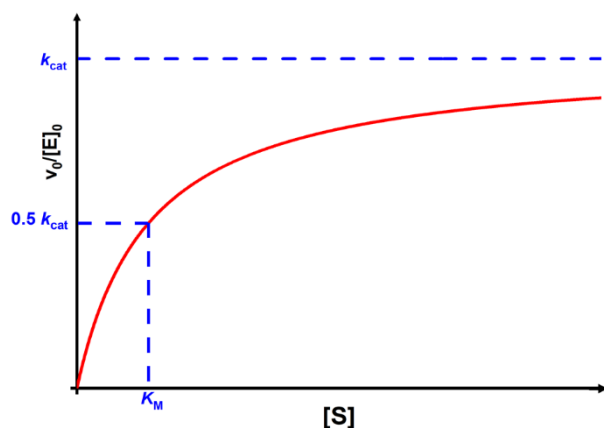


Figure 1.10: Michaelis-Menten curve showing reaction rate versus substrate concentration. The initial slope of the curve equals k_{cat}/K_M .

There are a variety of multi-substrate mechanisms (**Figure 1.11**), differing by two factors. Firstly, whether the substrates are bound and converted individually or whether a so-called ternary complex EAB is formed between the enzyme and both substrates. Secondly, the reaction *via* a ternary complex can either be random,

meaning the order of binding the substrates is irrelevant to the reaction, or sequential, where the reactions can only proceed if substrates are bound to the enzyme in a specific order (156).

This work will focus on the ping-pong mechanism (**Figure 1.11 c**), to which peroxidases typically adhere (2). In the ping-pong mechanism, two substrates sequentially react with an enzyme with the formation of an intermediate. Reaction of the enzyme (E) with the first substrate (A) proceeds *via* an enzyme-substrate complex (EA) and leads to the formation of an intermediate (E*) and release of the first product (P). A second substrate (B) then binds this intermediate and forms a complex of the substrate with the intermediate (E*B). As the second reaction product (P') is released (157), the enzyme is recycled back to its resting state.

The following equation describes ping-pong kinetics:

$$\frac{v_0}{[E]_0} = k_{cat} \frac{\left(\frac{[A]_0 [B]_0}{K_{M(A)} K_{M(B)}} \right)}{\frac{[A]_0}{K_{M(A)}} + \frac{[B]_0}{K_{M(B)}} + \frac{[A]_0 [B]_0}{K_{M(A)} K_{M(B)}}} \quad 1.7$$

This equation contains two separate values for K_M as the Michaelis constants for the two substrates are independent of each other. They can be determined individually by varying one substrate concentration with the other at saturating concentration.

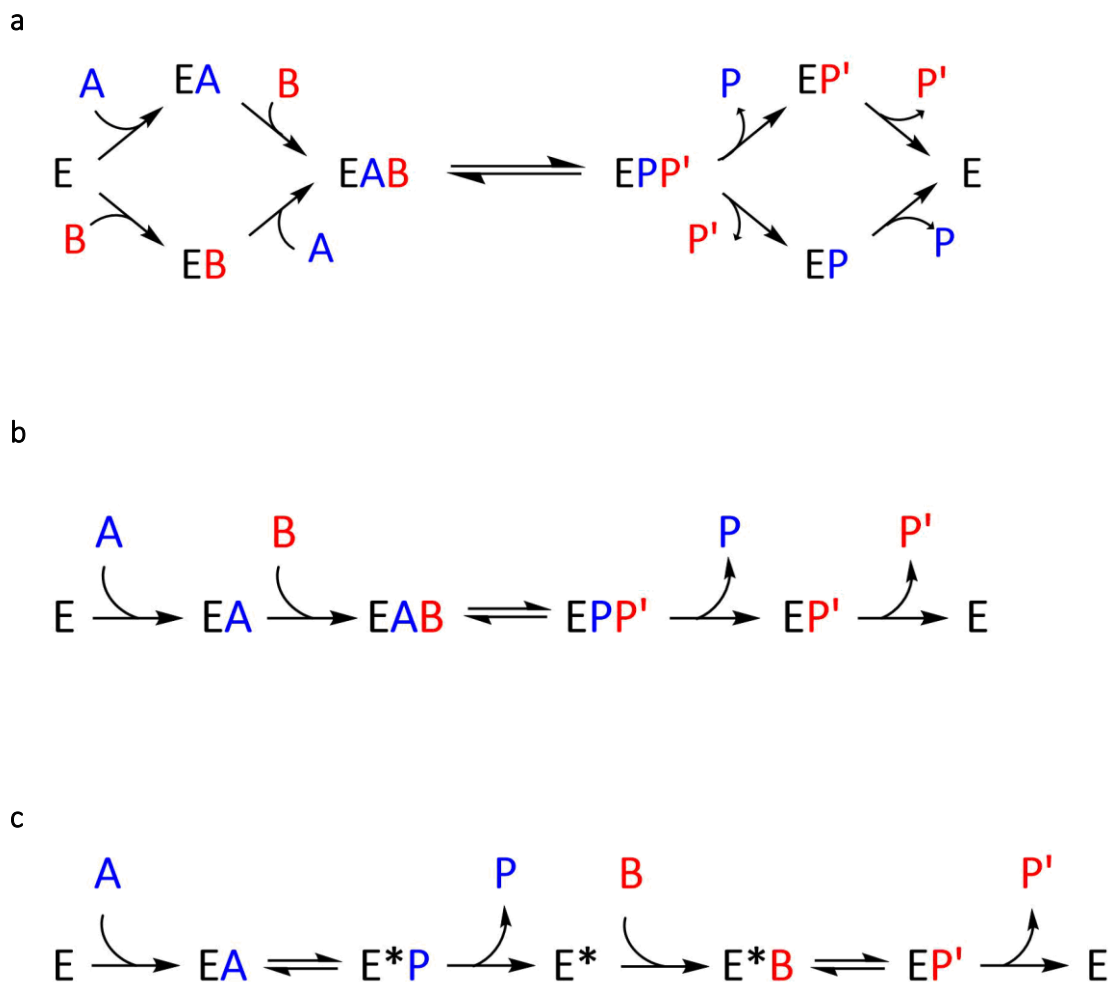


Figure 1.11: Kinetic models for a reaction between an enzyme and two substrates.

a – Random order ternary complex mechanism. The enzyme E binds substrate A and B in any order to form complex EAB which then releases products P and P', also in either possible order. **b** – Sequential ternary complex mechanism. E binds A and B in a specific order to form complex EAB which then releases the products in sequential order. **c** – Ping-pong mechanism. E first binds A and converts it to P, forming intermediate E* which can then convert B to P'.

1.5. Cytochrome structure and chemistry

A number of enzymes require coenzymes to carry out their catalytic function and bind metal ions or small organic molecules such as flavins as prosthetic groups. With biological macromolecules generally incorporating very few chemical elements, the introduction of metals – either in ionic form or as part of an organic cofactor – diversifies protein chemistry and enables novel functionalities (158). Molybdenum, for

example, plays an important role in some metabolic enzymes in humans and manganese in photosystem II is crucial in photosynthesis (159, 160).

The most prevalent transition metal in living organisms, however, is iron, which constitutes approximately 50 ppm of the human body by weight (161). It is commonly incorporated into enzymes as part of heme, a prosthetic group found in proteins in all three domains of life: Archaea, Bacteria and Eukarya (161, 162). Heme is a complex of the tetrapyrrole known as porphyrin coordinating the iron ion *via* four equatorial nitrogen atoms, and proteins binding this prosthetic group are called cytochromes.

Heme groups can be divided into several types, *b*- and *c*-type heme being the most common (**Figure 1.12**). Heme B is non-covalently coordinated to the protein whereas heme C covalently binds the protein. In most cases the heme C binding site is a CXXCH motif, where the sulfur-containing cysteine residues form thioether bonds with the porphyrin's vinyl groups. Histidine acts as an axial ligand. The thioether bonds potentially lend the protein more stability than the non-covalently bonded analogues and allow for efficient packing of multiple hemes in a protein (163). From a biosynthetic point of view, heme B is the precursor for heme C; in *E. coli*, the heme attachment is catalysed post-translationally by the cytochrome *c* maturation apparatus (Ccm) (164, 165). Structural analyses of apo- and holoproteins also suggest that in the majority of cytochromes, heme binding is followed by a small conformational change (166).

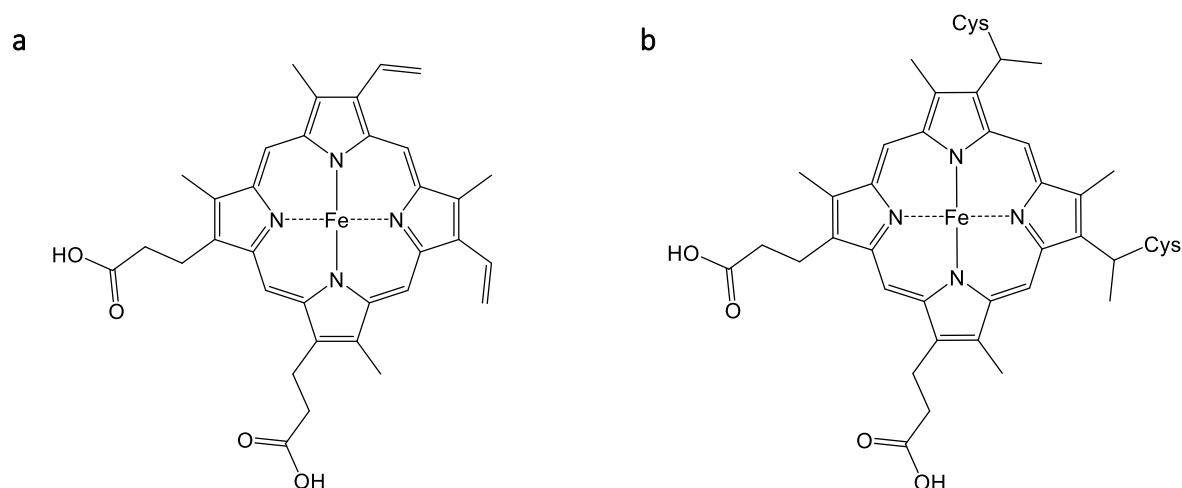
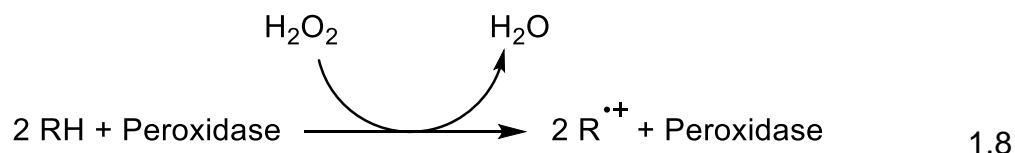


Figure 1.12: Chemical structures of heme B (left) and heme C (right).
Heme C is covalently bonded to a cysteine residue *via* thioether bonds.

Heme is a versatile prosthetic group and while a large variety of enzymes contain heme, not all cytochromes are catalytically active. The tetrapyrrole's conjugated system of delocalised π -electrons can stabilise different oxidation states of the coordinated metal ion, giving rise to a variety of activities. Oxidation states of +2, +3 and +4 are commonly accessible to the heme iron, which are referred to as ferrous, ferric and ferryl iron, respectively (167). Some heme-binding proteins are involved in other natural processes, such as storage, transport and sensing of oxygen, common functions of the α -helical family of heme proteins known as globins (164, 168). More typically, however, cytochromes are involved in electron transfer chains or function as oxidoreductases (161).

1.5.1. Cytochromes as Peroxidases

Whilst there is a vast array of functions for heme proteins, many natural cytochromes function as redox catalysts known as oxidoreductases. Among these, peroxidases are a very important family, which reduce hydrogen peroxide and simultaneously oxidise a range of substrates:



In nature, peroxidases are expressed in almost all organisms across all domains of life where they have a vast array of functions, for example in the metabolism of reactive oxygen species, innate immunity, cell wall synthesis and wound healing (169–175). Although heme is not essential for peroxidase function, over 80% of known natural peroxidases are cytochromes (172, 176). **Figure 1.13** shows the classification of natural peroxidases according to the RedoxiBase (176). It especially highlights where within the peroxidase family horseradish peroxidase falls, one of the first peroxidases to be discovered and studied, the structure of which is shown in **Figure 1.14**. After over a century of research into this enzyme extracted primarily from the plant's roots, it is a very well characterised enzyme that is produced on a relatively large scale for a variety of applications. Commercially available horseradish peroxidase is typically a composition of isoenzymes, structurally distinct proteins catalysing the same reaction (174, 175, 177).

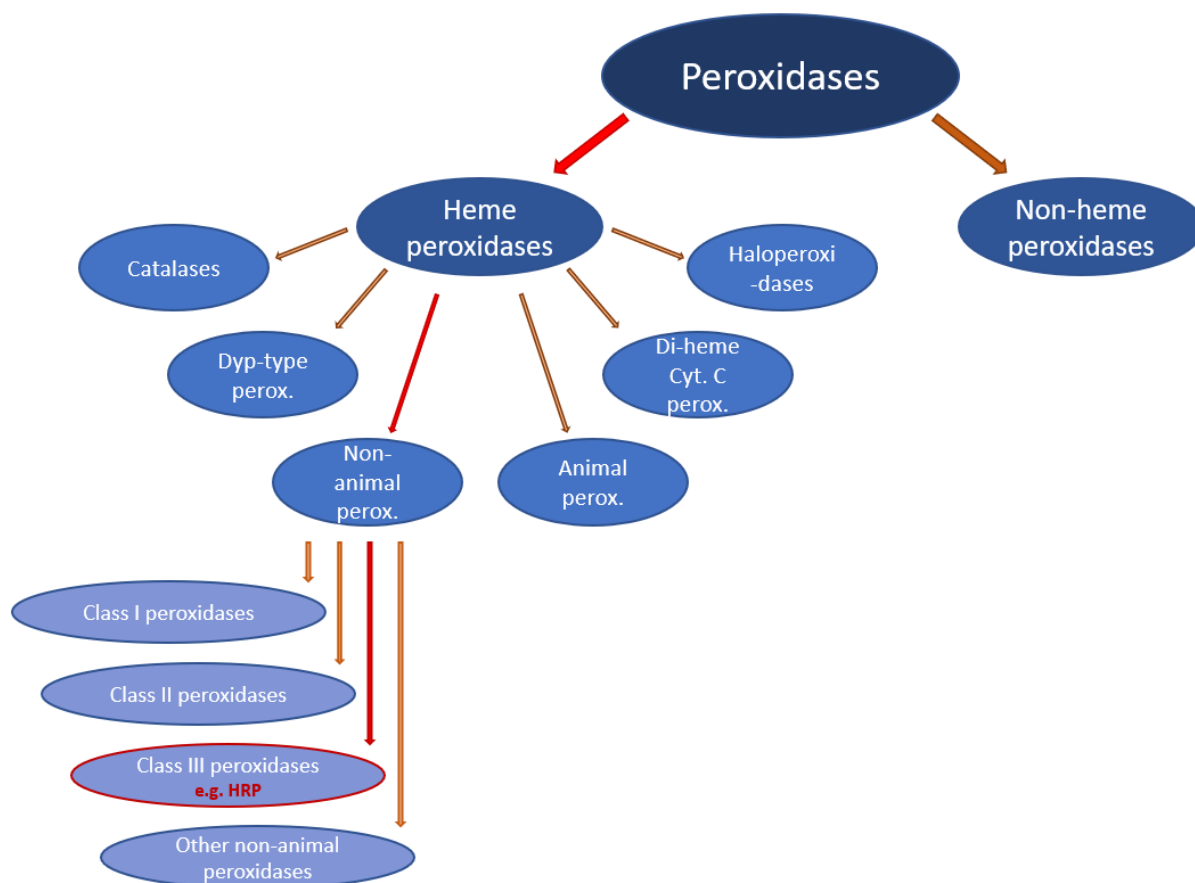


Figure 1.13: An overview of classification of natural peroxidases.

The situation of horseradish peroxidase (HRP) within the peroxidase family is highlighted in red.

In addition to the vast array of functions peroxidases have *in vivo*, they have been utilised *in vitro* for a variety of functions. Horseradish peroxidase, for example, has been extensively used in applications: as an antigen in immunoassays, in the biosynthesis of organic polymers, in cancer treatment, and in electrochemical sensors (173, 177).

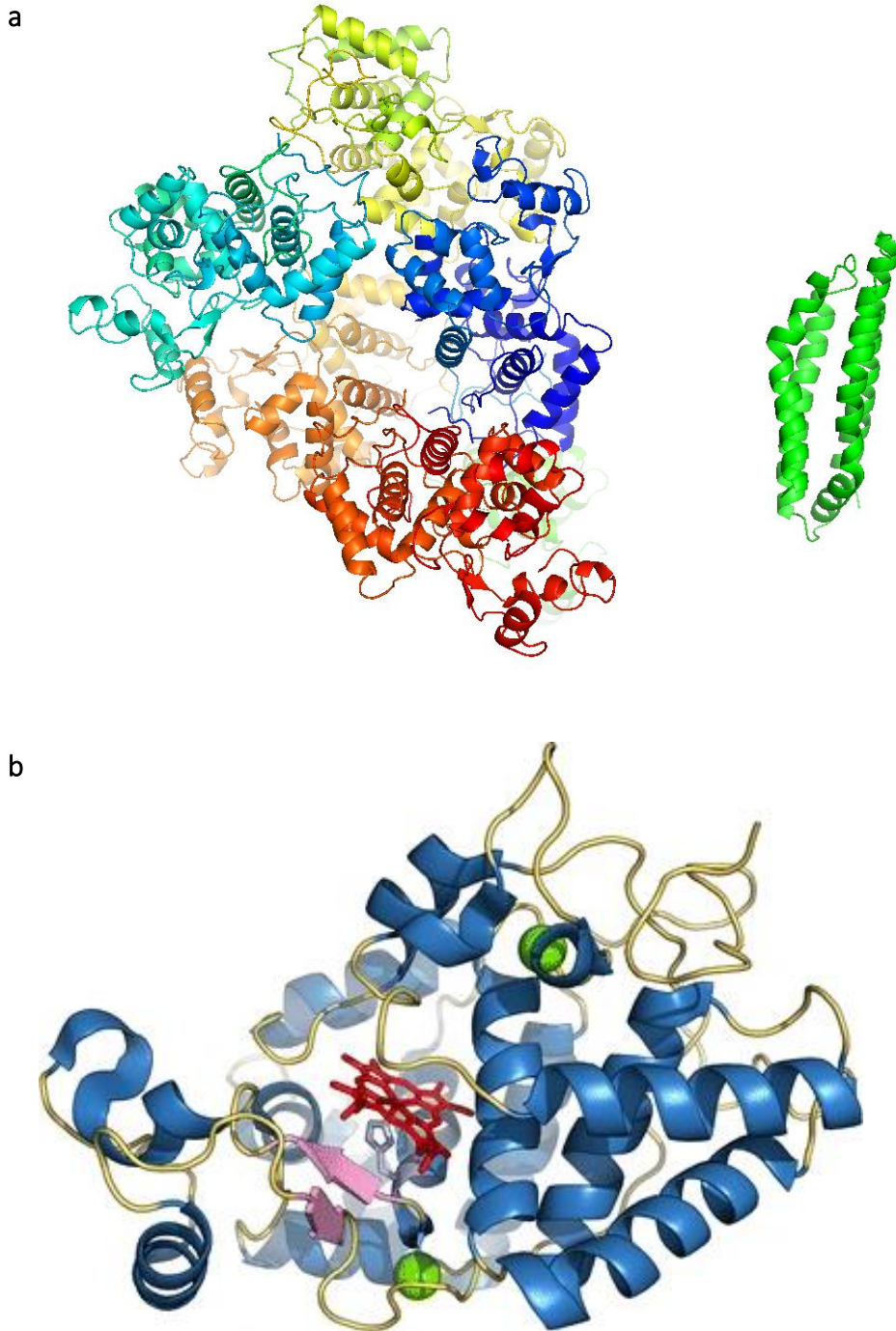


Figure 1.14: Structure of horseradish peroxidase C1A.

a - Cartoon representation of horseradish peroxidase (left), with C45 (right) shown for scale (178).

b - Detailed view of the heme environment in horseradish peroxidase, with heme shown in red, helices in blue, β -sheet regions in pink, loops in yellow and the endogenous Calcium ions as green spheres. The histidine residue in position 170 which coordinates to the heme iron is shown in light blue. Figure reproduced from (177).

1.6. Enzyme Engineering

Whilst enzymes have demonstrated impressive versatility for uses in biotechnology, medicine and chemistry outside of their natural environment, they remain highly evolved machines catalysing very specific reactions under very specific conditions. This means that the variety of applications is limited by the natural range of proteins that have evolved by natural means and any applications must be adapted to the proteins abilities and working conditions. Consequently, there has been considerable interest in modifying proteins to tailor them for specific applications. Of course, the idea of adapting, modifying, imitating and recreating natural processes and products is not restricted to proteins. The expanding field of synthetic biology has also successfully produced artificial protocells that allow for cell free gene expression systems, DNA origami and artificial viruses for more efficacious vaccines (179–183). As for producing man-made proteins, the two routes being explored are the modification of natural proteins and the creation of entirely novel proteins (which can then, of course, be subject to further modification). Proteins are very suitable for both of these as they are built from individual blocks – the twenty naturally occurring amino acids – that can individually be replaced to alter structural and functional characteristics (184).

1.6.1. Directed Evolution

Directed evolution has been used extensively to modify proteins; a small scale, laboratory method imitating Darwinian evolution where repeated mutations along with selective preservation and enhancement of the most advantageous traits for survival lead to amplification of desirable attributes over time. Partial randomisation of genes encoding a protein leads to a multitude of variants which can be screened for particular characteristics – such as specific ligand binding or enhanced enzyme kinetics – which

are then selectively amplified through PCR (185–187). However, the identification of molecules showing desired traits in the large number of proteins produced remains a challenge in specific protein modification through directed evolution. Mathematical modelling and computational biochemistry have been increasingly used to tackle this problem (184, 188–190).

Overall, directed evolution, though not specific to enzymes, has played an important role in contemporary enzymology. Its impressive accomplishments include enhanced enzyme specificity, increased catalytic efficiency and stereospecificity. Horseradish peroxidase has been subject to directed evolution to achieve higher thermostability and better longevity for the multitude of *in vitro* applications for which this well characterised cytochrome has been employed (175). The surface and the loop regions proved especially good targets for mutations, which is unusual in heme enzymes, where mutations are commonly focussed on the active site (175, 185).

A further achievement of directed evolution receiving a lot of attention in recent years is the attainment of enzyme promiscuity, so that they can react with unnatural substrates to perform stereospecific chemistry not possible in the natural environment *in vitro* (191, 192). Frances Arnold, Nobel laureate in Chemistry in 2018 for her work on directed evolution of enzymes, has shown a 9000-fold increase in catalytic efficiency for propane oxidation by a modified cytochrome P450 (193, 194).

1.6.2. *De novo* enzyme design

A significant limitation of directed evolution, however, is the very small fraction of the total possible sequence space that can be covered. There are 20^n possible amino acid sequences for a protein with n amino acid residues, typically in the hundreds. Natural evolution has only explored a minute fraction of the possible sequences, with

approximately 10^{12} distinct proteins being produced by living organisms (1, 184). Directed evolution widens the subset of sequence space that is open to investigation slightly by producing variants related to natural proteins. However, this still leaves a vast amount of potential sequence space that remains unexplored by both natural and directed evolution; the possibility of venturing into these novel territories has been the motivation behind *de novo* protein design, as shown in **Figure 1.15**.

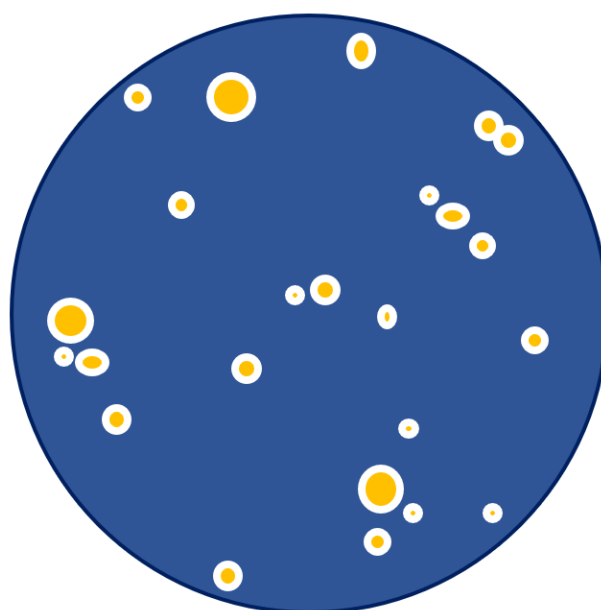


Figure 1.15: Schematic diagram of the protein sequence space.

Nature has only explored a tiny fraction of the possible protein sequences (blue), meaning native proteins (orange) only take up a small part of the sequence space. Directed evolution allows us to explore the space surrounding the “islands” of natural proteins (white), but cannot venture into the vastness of unexplored sequence space; *de novo* protein design, however, is free to explore any part of the protein sequence space that may be impossible to investigate otherwise. Figure adapted from (1).

De novo, a phrase meaning ‘from the beginning’ (or literally translated from Latin as ‘of new’) is an approach to protein engineering that seeks to design proteins from scratch by creating novel sequences from natural amino acids. Of course, this bottom up method requires a much deeper understanding of the biochemistry of protein stability, folding and functionality. A key advantage of directed evolution is that a

protein with a single point mutation is likely to take a similar three-dimensional shape as the native protein on which the design was based, resulting in predictability of protein structure and stability. A completely novel polypeptide, on the other hand, may not successfully fold into a stable or functional biomolecule at all. However, with a good understanding of protein biochemistry, *de novo* protein design could open up endless possibilities to tackle challenges within and outside the fields of biology, especially with the growing number of applications of proteins *in vitro*.

Much *de novo* protein design is based on computational methods where programmes such as Rosetta can be used to calculate the energy of an amino acid sequence and predict the three-dimensional protein structure from this (1). Using these *in silico* methods, both globular and membrane proteins have been successfully designed (195–197). Other groups have used the principles of protein engineering to produce multichain polypeptides that assemble into particles or vesicles (198–200).

1.6.3. The maquette approach

The first artificial four-helix bundle proteins were designed by DeGrado lab in the 1980s by mimicking the hydrophobic and polar amino acid patterning of natural four-helix bundles (17, 201–203). Building on this, the maquette approach developed by Dutton has been successful in designing a number of four-helix bundle proteins (18, 19, 204). This non-computational bottom-up method starts with a simple, generic protein sequence with minimal complexity encoding a bundle of helices with repeated amino acid sequences in a heptad of hydrophobic (H) and polar (P) amino acids in an HPPHHPP pattern. Subsequently, cofactors can be added to this protein scaffold to allow functionality, resulting in robust, operative proteins. This simplicity allows these proteins to be studied without the complexity of their natural counterparts that have

undergone millions of years of evolution, with the possibility of changing individual amino acids (205). These synthetic maquettes can not only incorporate a wide range of natural cofactors, such as metal ions or flavins but also self-assemble *in vivo* (19, 20). A diverse range of function can be imparted on the protein by choice of cofactor; maquettes have been produced with functions including light-harvesting (206), energy transfer (207), magnetosensing (208) and transmembrane electron transfer (209). Many maquettes are cytochromes, thus binding heme as a cofactor, an exceptionally diverse and common natural cofactor which is also bound by some naturally occurring four-helix bundle proteins (210, 211).

1.7. C45 – a *de novo* designed four-helix bundle

In the Anderson group, a variety of four-helix bundles have been designed using the maquette approach and assembled *in vivo* in *E. coli* (2, 20, 21, 212). The work herein will focus on C45, a cytochrome with peroxidase functionality developed in multiple rounds of mutagenesis. The initial maquette design from which C45 was developed includes four helices as well as flexible loops comprising of glycine and serine. The cellular cytochrome *c* maturation apparatus incorporates the heme cofactor by covalently linking it to the cystine side chains of the CXXCH binding motif *via* two thioether bonds. This heme C binding motif can be located on any of the four helices, producing different types of maquettes named C1, C2, C3 or C4, depending on the α -helix where the heme C binding motif is located. These four proteins were analysed biophysically using circular dichroism spectroscopy. With the highest helicity proportion and the second highest melting temperature of 51 °C, C4 seemed the most suitable of these four maquettes to further modify to incorporate peroxidase activity (2, 21).

C4 is a di-heme protein: firstly, the aforementioned CXXCH motif (in this case CIACH) is located on the fourth helix, which covalently binds *c*-type heme. In addition to the two thioether bonds to the tetrapyrrole, the heme is also ligated *via* a bis-histidine pair on helices two and four. A further bis-histidine pair on the other two helices serves as another heme binding site. This four-helix bundle was modified further (**Figure 1.16**) to produce a robust heme C binding redox catalyst. The first step of this process was the removal of the bis-histidine pair that form the non-covalent heme binding site on the first and third helix to construct a maquette called C46. The two histidine residues were replaced with phenylalanine to improve the packing of the protein's hydrophobic core and rigidify the protein. This enhanced the stability, as seen by the raised melting temperature (from 51 °C to 84 °C). The protein retained the covalent binding motif CXXCH on the fourth helix and the two histidine residues ligating the covalently bound heme. C46, however, did not show significant peroxidase activity, but instead bleached under addition of H₂O₂ and the common peroxidase substrate ABTS, which indicates heme degradation (2). Imitating natural peroxidases, which typically incorporate a mono-histidine ligated heme (213), the distal heme-ligating histidine residue on helix two was also substituted with phenylalanine. This resulted in C45, a mono-heme ligating four-helix bundle maquette protein, which retains the great thermostability of C46 (T_M of C45 is 86 °C). **Figure 1.17** shows a cartoon representation of C45, with special focus on the environment of the cofactor heme.

1. Introduction

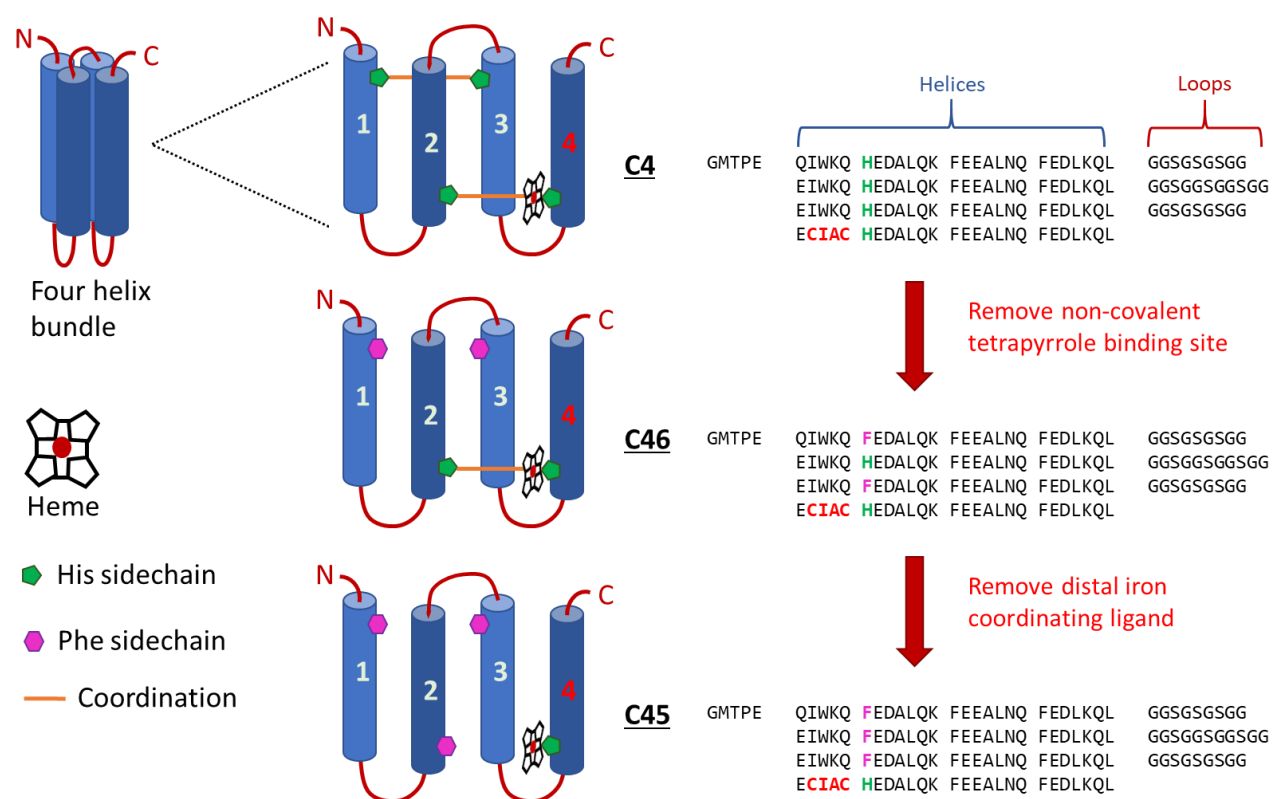


Figure 1.16: The design of heme-binding four helix bundle maquette protein C45.

The initial maquette protein is C4, a four helix bundle with a single CIACH motif on the fourth helix covalently binding heme C which is also ligated by a bis-histidine pair on helices 2 and 4. A bis-histidine pair on the other two helices forms another, non-covalent tetrapyrrole binding site. In the next step, the two histidine residues in these locations are replaced with phenylalanine and the non-covalent heme binding site thereby removed. This protein is called C46. This is further modified in a final step by removing the distal iron coordinating histidine residue and replacing it with phenylalanine. The resulting mono-histidine ligating maquette is C45. Adapted from (2).

In steady-state kinetic assays, C45 demonstrated the desired peroxidase activity. In contrast to C46, C45 reacted with H_2O_2 and ABTS (**Figure 4.4**) obeying ping-pong steady-state kinetics, like many of the natural peroxidases (2). With a $k_{\text{cat}}/K_{\text{M}}(\text{ABTS})$ of $3.2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ C45 showed high catalytic efficiency comparable to some other heme-containing *de novo* peroxidases (214, 215). This catalytic efficiency was even comparable to that of natural peroxidases, such as HRP, for which a $k_{\text{cat}}/K_{\text{M}}(\text{ABTS})$ value of $5.1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ has been measured (216). C45 showed great substrate

promiscuity, too, as well as catalytic promiscuity, being capable of oxidative dehalogenation of halogenated phenols with 5 times greater catalytic efficiency than a natural dehaloperoxidase (217). Furthermore, ferrous C45 could take part in interprotein electron transfer (218) with cytochrome *c* due to the complementary electrostatic surfaces. This is in agreement with other work with heme maquettes and natural cytochrome *c* peroxidases (213, 219). More recently, the enzyme was also shown to function as an efficient, stereoselective and promiscuous carbene transferase (220).

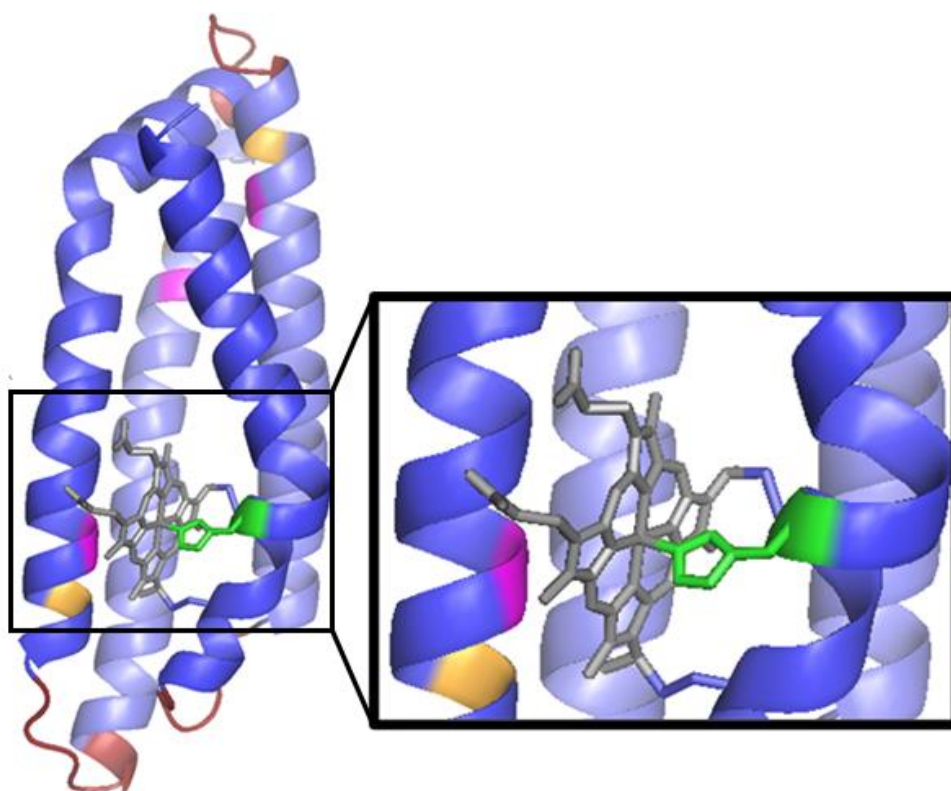


Figure 1.17: Cartoon representation C45 with an inset showing the environment of the heme cofactor in detail.

α -helices are shown in blue, red represents the loops between the helices, phenylalanine residues that were mutated from histidine in C4 and C46 are shown in pink and the histidine residue in the heme binding motif in the fourth alpha helix is shown in green, coordinating a single heme, shown in grey. Further, tryptophan residues are shown in orange as preliminary data from electron paramagnetic resonance spectroscopy confirmed the presence of a tryptophan radical during the enzyme's catalytic cycle. This is most likely W43, which is located closest to the heme cofactor and can be seen in the inset. Computational model (2) visualised using PyMol software.

C45 also exhibited remarkable resilience not only to high temperatures, as demonstrated by the high T_M and also the reversibility of thermal denaturation, but also to the presence of low concentrations of some alcohols. In temperatures of up to 70 °C, catalytic activity of the enzyme further increased. Limited work has been done regarding the stability of the enzyme in organic cosolvents. In concentrations up to 30% ethanol, ABTS oxidation proceeded with similar catalytic activity as in pure buffer (2). In higher ethanol concentration, activity declines significantly. In methanol concentrations of up to 50%, catalytic activity of dehalogenation of 2, 4, 6-trichlorophenol in the presence of H_2O_2 only decreases by about 25%. This thesis seeks to further explore this impressive resilience to such hostile environments

1.8. Aims of this work

Whilst a fair amount of work has been carried out investigating the behaviour of enzymes in organic solvents, this has mostly focused on simple, hydrolytic enzymes (13).

One of the primary advantages of maquettes is the ability to work with an evolutionary naïve protein that does not exhibit the complexity of its natural analogues. This makes C45 a suitable candidate to study the effect of organic cosolvents on the structure and function of a heme-binding oxidoreductase. To that end, chapters 3 and 4 describe the effects of acetonitrile, methanol, ethanol, isopropanol and TFE on the enzyme's structural integrity and catalytic activity, respectively. CD and UV-vis spectroscopy experiments are presented that probe the structure of the enzyme in cosolvent, whereas steady-state kinetics of the oxidation of ABTS by C45 and hydrogen peroxide are presented to elucidate changes in the enzyme's catalytic behaviour.

In these studies, a remarkable effect of TFE could be shown, which is highly beneficial for activity and increased the catalytic turnover number by a factor of 7. In order to elucidate this effect, further structural analyses are presented using SAXS, analytical size exclusion chromatography and NMR. These indicate that the effect of TFE is twofold. Firstly, it structurally stabilises C45, which reproduces its effects on other helical proteins. Secondly, the stabilising effect also influences the reactive catalytic intermediate compound I, thereby preventing the degradation of the protein and enhancing total turnover number.

This work could provide a model for other helical enzymes, which may exhibit similarly enhanced catalytic activity in the presence of TFE.

2. Materials and Methods

C45, the protein used in this work, was expressed in *E. coli* containing the plasmid vectors for the protein and the *E. coli* cytochrome *c* maturation apparatus, conferring resistance to carbenicillin and chloramphenicol (2, 221). The protein was then purified using nickel affinity chromatography and size exclusion chromatography, which is a method of separating mixed protein samples based on the size of individual species (**Figure 2.1**). Smaller species diffuse more into resin and pass through the porous network, thus having a longer path length through the column. Elution times are therefore inversely correlated with the size of a protein. Originally intended as a means of purifying proteins, this technique can also be used analytically to separate a sample by size of species present (25, 221). To confirm the presence and purity of C45, SDS-PAGE was employed, a commonly used method of analysing protein size, concentration and purity of samples of denatured and SDS-bound proteins based on distance travelled within a charged hydrogel (222).

The protein was further characterised using UV-visible spectroscopy. This technique can supply information about protein concentration, using the Beer-Lambert law:

$$A = \epsilon c l \tag{2.1}$$

where A is absorbance, ϵ is the extinction coefficient that is specific for a species measured at a specific wavelength, c the concentration of that species and l the path length.

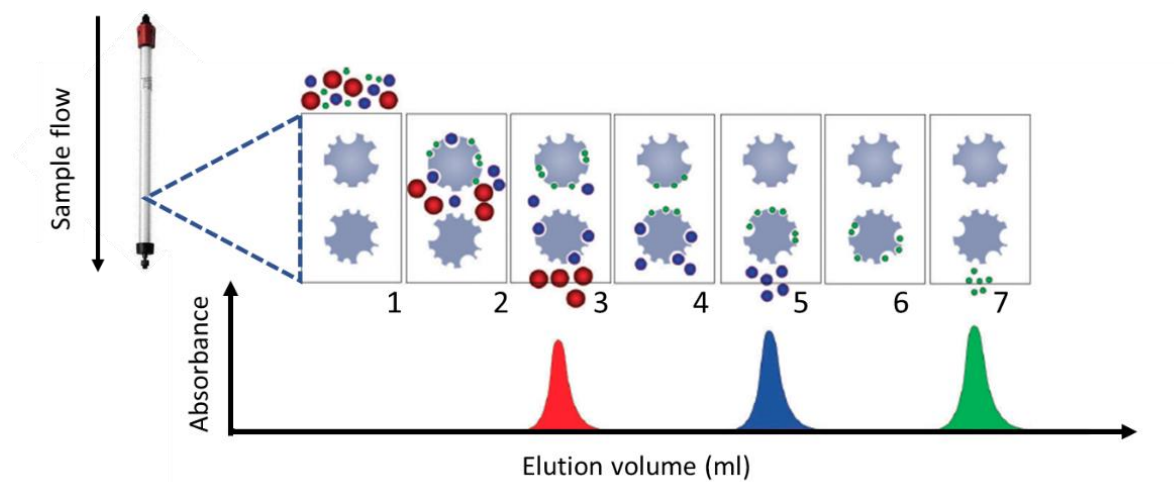


Figure 2.1: Size-based separation of three species by size exclusion chromatography.

The larger species (red) interacts less with the porous resin and elutes earlier, while smaller species (blue and green) pass through the porous network within the beads, thereby taking a less direct path and eluting later, as a larger fluid volume has passed through the column. There is therefore an inverse correlation between the species size and supplied fluid volume with which it elutes. The upper limit for elution volume is set by the column volume. Figure adapted from (223).

Furthermore, UV-visible spectroscopy is also very pertinent in providing insights into prosthetic groups and their chemical environments, such as the oxidation state of metal ions or metallic cofactors (224). In proteins, absorptions between 180 and 230 nm are associated with π - π^* and n - π^* transitions in the peptide bonds. The aromatic side chains of tryptophan, tyrosine and phenylalanine residues lead to absorption in the range of 230-300 nm similar to disulfide bonds that absorb at approximately 260 nm (225). These absorbances can be employed for determination of protein concentration. Additionally, a cofactor may also absorb UV-visible light. Heme, for instance, has a strong band of absorption in the region of 400 - 410 nm caused by π - π^* transitions known as the Soret band (226). The exact location of this band is dependent on the ligation, chemical environment as well as the oxidation state of the heme iron. In addition to the Soret band, Q-bands in the region of 520 – 550 nm can give more accurate information about the heme environment. UV-visible spectroscopy

can also be employed to acquire a pyridine hemochrome, as described in **section 2.2.6**. The concentration and identity can be worked out from the relative intensities and positions of the Q-bands located between 500 nm and 600 nm; Berry and Trumpower give extinction coefficients for both reduced and oxidised species of all three types of heme (227, 228).

A further technique employed for the characterisation of C45 is circular dichroism spectroscopy (CD). CD can be used to analyse secondary structural elements of a protein, such as α -helices or β -sheets. The technique is based on a protein sample absorbing left- and right-handed circularly polarised light differently owing to the chirality of amino acids. This results in a beam of elliptically polarised light where the degree of ellipticity can be measured. Different motifs of secondary structure exhibit distinct, characteristic CD spectra due to different arrangements of the chiral amino acids (229, 230). A benefit of CD is the ability to perform thermostability studies by continuously monitoring the sample at a specific wavelength or collecting full spectra at regular intervals as the sample is heated (231). An important limitation of the technique, however, is that CD can only be employed to investigate secondary structure and provide very little information regarding tertiary structure.

Nuclear magnetic resonance spectroscopy (NMR) provides some information on the chemical environment of specific nuclei, most importantly ^1H and ^2H (or D), ^{13}C as well as ^{15}N (232). It is based on certain nuclei possessing magnetic moment and their alignment with an external magnetic field. This alignment can be perturbed with a pulse of electromagnetic energy. The resonant frequency is dependent on the magnetic field observed at the nucleus, which is dependent not only on the strength of the external magnetic field, but also on the nature of the nucleus and its chemical environment, as electron density has a shielding effect (233).

Two-dimensional (2D) NMR is a method to gain more accurate information with better resolution about large molecules, such as proteins. In heteronuclear NMR focussing on ^1H and ^{15}N nuclei, magnetisation is transferred between those two nuclei. A multitude of different 2D NMR techniques involving these nuclei exists, which differ in the type of pulse sequence used to excite nuclei. In this work SOFAST-HMQC and TROSY-HSQC were used. The former had previously been employed for the characterisation of C45 while the latter was chosen due to its sensitivity for macromolecules with a molecular mass of up to 80 kDa (234, 235).

Small angle X-ray scattering (SAXS) can provide information on the size and shape of a protein. This method does not require crystallisation or freezing of the protein, but can be carried out in solution which greatly facilitates analysis. During this rapid technique, a monochromatic X-ray beam passes through a sample (**Figure 2.2**).

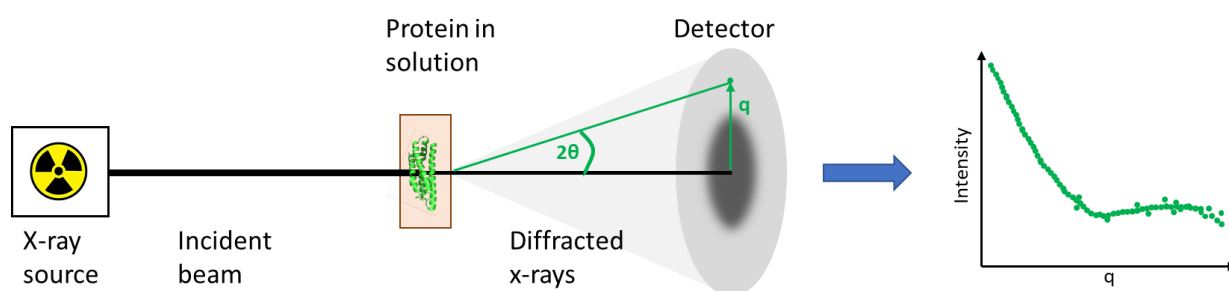


Figure 2.2: Schematic representation of a SAXS experiment.

A monochromatic x-ray beam passes through a solution of protein, which diffracts the beam. The scattered x-rays are detected and scattering angles measured. Figure adapted from (236).

Proteins in the sample solution scatter the beam before it reaches a two-dimensional detector. The angles by which the incident beam differs from the scattered beam (2θ) can be used to calculate the scattering vector q (237):

$$q = \frac{4\pi \sin \theta}{\lambda} \quad 2.2$$

To investigate the enzyme's catalytic activity, steady-state kinetic measurements were performed, referring to a stable concentration of the enzyme-substrate complex. This can be performed on a UV-visible spectrophotometer or a plate reader, which yields less precise data, but is a good method for the acquisition of a large number of repeat measurements. In order to investigate the initial stages of a reaction accurately, stopped flow spectrometry is most appropriate. During stopped flow spectrometry (**Figure 2.3**), an air pressure driven mechanism rapidly injects fixed volumes of fluid from two syringes into a mixing cell (in this case C45 and the substrates). Once this mixing cell is filled, flow is immediately halted, and the measurement triggered. An incident light beam passes through the chamber and the change in absorbance or fluorescence can be measured. A connected photomultiplier or photodiode array helps detect the absorbance changes and convert this measurement into a spectrum of light absorbance over time. Photodiode array detectors and photomultipliers differ in measuring multiple wavelengths or a single selected one, respectively (238). The advantage of stopped-flow spectrophotometry is the speed at which mixing and measurement can proceed, which makes it possible to accurately observe the kinetics of the first seconds of a reaction, which is not possible with a UV-visible spectrometer or plate reader

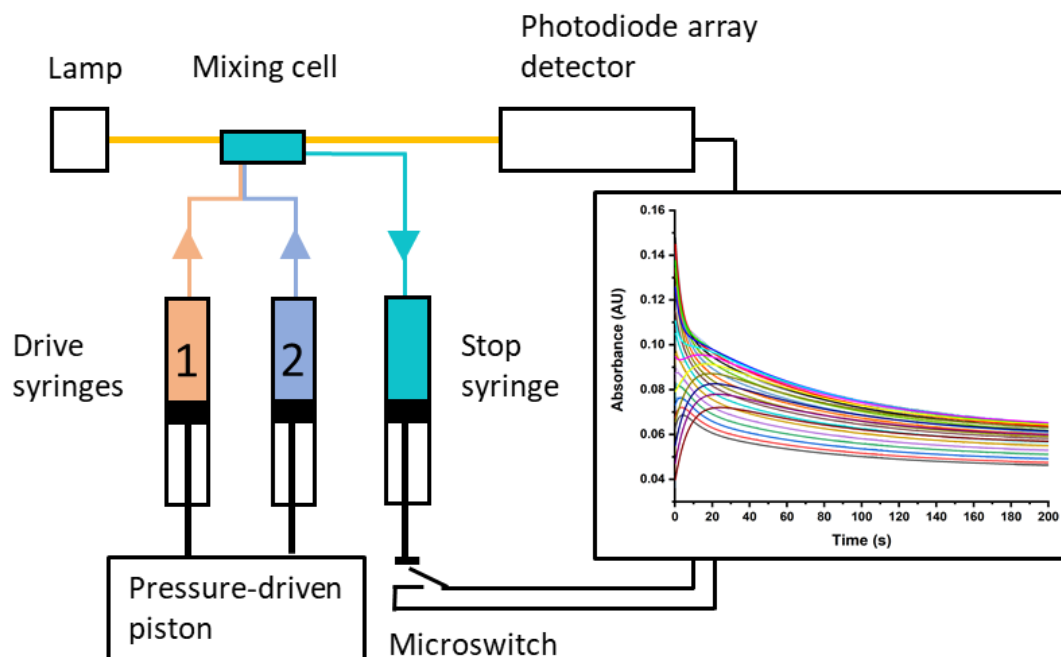


Figure 2.3: Schematic representation of a stopped-flow spectrometer.

A pressure driven piston initiates the rapid release of equal fluid volumes from syringes 1 (containing C45 solution) and 2 (containing substrate solution) into the mixing cell. A stop syringe also connected to the mixing cell acts as a backstop when the mixing cell is filled. A light beam passing through the sample is detected using a photodiode array detector. These results are recorded as a graph showing absorbance over time for all wavelengths measured. Figure reproduced and adapted from (238)

2.1. Materials

Chemicals used within this project were purchased from Sigma-Aldrich, ThermoFisher Scientific, Alfa Aesar and Apollo Scientific. Additionally, sterile and deionised water was used, generated using a Milli-Q Integral Water Purification System by Merck Millipore. The pSHT vector for C45 protein was produced by Dr Daniel Watkins (University of Bristol, School of Biochemistry).

2.2. Methods

Unless specifically mentioned otherwise, percentage mixtures of solutions always refer to volume by volume.

2.2.1. Protein Expression

C45 (**Figure 2.4**) was expressed in T7 express *E. coli* BL21 (DE3) co-transformed with vector pEC86 encoding for the c-type cytochrome maturation system Ccm. Cells from glycerol stock were grown in sterile LB broth (10 g tryptone, 5 g yeast extract, 10 g NaCl per litre purified water) treated with carbenicillin (50 µg/ml) and chloramphenicol (34 µg/ml) and incubated at 37 °C shaking at 200 rounds per minute. Protein expression was induced with the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 1mM) and the cultures were incubated for a further 4-5 hours. A cell pellet was collected by centrifugation (25 minutes, 4,500 × g, 8 °C) and resuspended in lysis buffer (300 mM NaCl, 50 mM sodium phosphate, 20 mM imidazole, pH 8.0).

The cell pellet was lysed by means of sonication (4 × 20 seconds, 23 kHz) and centrifugation (30 minutes, 40,000 × g, 4 °C). The filtered lysate was purified using nickel affinity column chromatography (GE Healthcare Life Sciences HisTrap HP 5ml, equilibrated in lysis buffer). A gradual increase in imidazole concentration was used to elute C45, with UV absorbance at 280 nm measured to monitor protein elution.

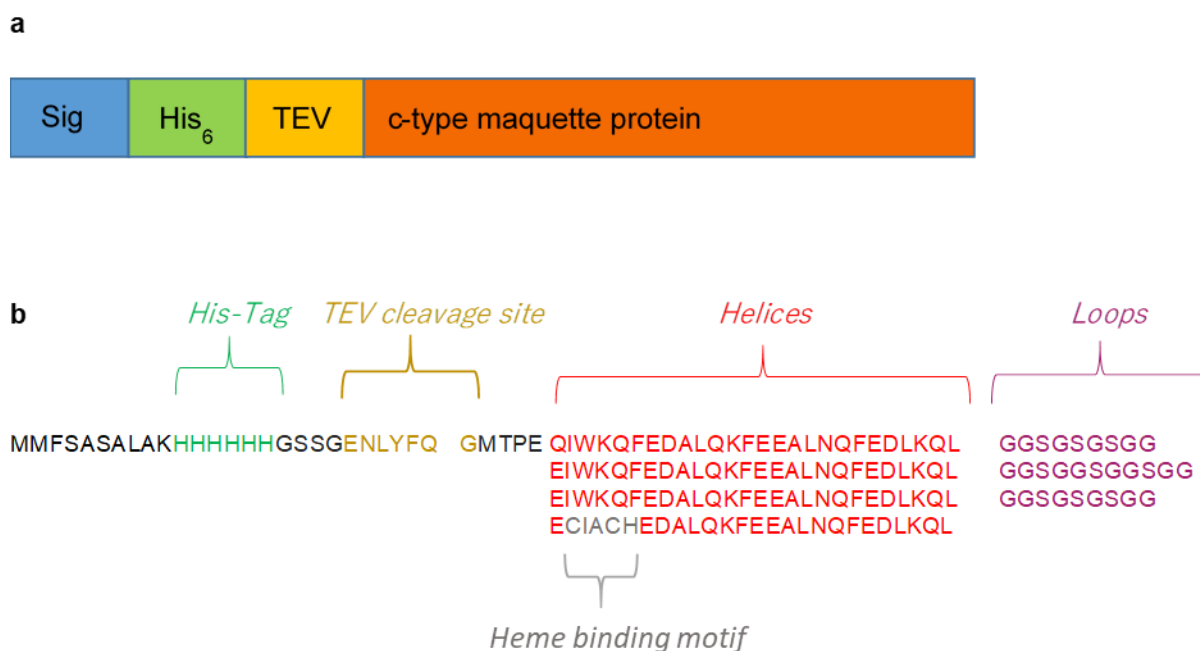


Figure 2.4: Sequence of C45.

a – Diagram of the different elements of the protein; the signal sequence at the N-terminus is cleaved when the protein is transported into the periplasm by the Sec translocon and serves as a recognition site. The His-Tag consisting of 6 histidines is used for purification using nickel affinity chromatography and may subsequently be cleaved from the protein using tobacco etch virus (TEV) protease. The plasmid was named pSHT due to the different elements of the protein (signal sequence, His-Tag, TEV protease recognition site). **b** - The amino acid sequence of C45. The protein can be cleaved between Q and G at the TEV cleavage site (yellow) using tobacco etch virus (TEV) protease to remove the hexahistidine tag (green). Alternating polar and non-polar amino acids results in self-assembling α -helices, here shown in red. These four helices are connected *via* three loops (purple). The CXXCH heme binding motif located on the fourth helix is shown in orange. Adaptation from (20).

2.2.2. Cleavage of the His-Tag and further purification

Cleavage of the 6-histidine tag was achieved in TEV cleavage buffer (20 mM Tris, 0.5 mM EDTA, pH 8.0) under anaerobic conditions with the addition of the reducing agent Tris(2-carboxyethyl)phosphine (TCEP, 1mM) and tobacco etch virus (TEV) protease (100 μ g per 1 L expression culture). After 6 hours, the cleaved C45 was purified using nickel affinity chromatography and size exclusion chromatography and analysed using SDS PAGE.

2.2.3. Size exclusion chromatography

Size exclusion chromatography was conducted at room temperature on ÄKTA systems by GE Healthcare (specifically the ÄKTA explorer and ÄKTA pure). A filtered buffer of pH 8.6 containing 20 mM CHES and 100 mM KCl was used. Different columns were utilised for analytical size exclusion chromatography (SEC) and SEC with the purpose of protein purification; the former was conducted on a Superdex 75 Increase 10/300 GL, the latter on a Superdex 75 pg HiLoad 16/600 column. Both column types were purchased from GE Healthcare; flow rates and pressure limits were set in accordance with manufacturer's instructions.

2.2.4. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-Page experiments were conducted using NuPage Tris-glycine precast gels with a 4-20% polyacrylamide content from Life Technologies and NuPage MES SDS Running Buffer (50 mM MES, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.3). 10 µl protein samples were mixed with 10 µl sample application buffer (50 mM Tris, 5% SDS, 30 % glycerol, 50 mM EDTA, 5% β-mercaptoethanol, 74.6 µM bromophenol blue, pH 7.0) before being denatured at 95 °C for 10 minutes and applied to the gels alongside PageRuler Prestained Protein Ladder (from Life Technologies) and run at 250 V for 30 minutes. Gels were stained using a solution of 25% ethanol, 10% acetic acid and 3 mM Coomassie Brilliant Blue R-250 before destaining in 20% ethanol, 5% acetic acid.

2.2.5. UV-visible spectroscopy

UV-visible spectrometry was conducted on a Cary 60 UV-visible spectrophotometer by Agilent Technologies. Quartz cuvettes with a path length of 10

mm were used. Spectra of buffer solutions (with cosolvent, if appropriate) were recorded and subtracted. Unless otherwise specified, all measurements were taken at room temperature.

2.2.6. Pyridine Hemochrome

A 120 μl sample of C45 with a concentration of 10-12 μM was mixed with 15 μl of 5M NaOH and 15 μl of pyridine (for protein denaturation and coordination of the heme, respectively). A UV-visible absorption spectrum was obtained and C45 subsequently reduced using sodium dithionite before another spectrum was obtained. This is an adaptation of a method described by Berry and Trumpower (227, 228).

2.2.7. Circular Dichroism spectroscopy

Samples for CD were prepared by mixing 4 nanomoles of C45 in buffer (so that the final concentration of C45 in 0.5 ml of solution would be 8 μM). This was then mixed with the desired volume of cosolvent to make up 0.5 μl , mixed thoroughly and transferred to a quartz cuvette with 1 mm path length. Protein concentration was determined using the absorbance at the heme Soret peak, where C45 has a known absorption coefficient of 147,300 $\text{M}^{-1} \text{cm}^{-1}$ (2). It is worth noting that this accounts only for holoprotein and not apoprotein. Therefore, any apoprotein present would lead to higher protein concentrations than calculated.

Circular dichroism (CD) experiments were carried out on a Jasco J-810 spectrophotometer under nitrogen atmosphere. Buffer solution (with cosolvent depending on the experiment) was recorded as a baseline and subtracted from the results. Scans were taken at a rate of 100 nm per minute and a temperature of 25 $^{\circ}\text{C}$.

Thermal denaturation studies were conducted with a ramp of 1 $^{\circ}\text{C}/\text{minute}$ (except where specified otherwise), starting at 8 $^{\circ}\text{C}$. Continuous measurements were taken at

222 nm, with additional full spectra taken at temperature intervals of 10-15 °C. When the top temperature was reached, a cooling cycle was immediately initiated, using the same conditions.

To normalise data for pathlength, concentration and number of peptide bonds in the protein (n), data were converted to mean residue molar ellipticity θ using the following equation (239):

$$\theta(\text{deg cm}^2 \text{ d mol}^{-1}) = \frac{\text{Ellipticity (mdeg)} \times 10^6}{\text{Pathlength (mm)} \times [\text{Protein}] (\mu\text{M}) \times n} \quad 2.3$$

2.2.8. Nuclear Magnetic Resonance

For heteronuclear 2D-NMR, ^{15}N -labelled protein was expressed in minimal media supplemented with 1 g of $^{15}\text{NH}_4\text{Cl}$ to 1 litre LB Broth. The addition was performed at the point of IPTG induction. Labelled protein was purified using the previously described expression protocol. NMR samples contained 250 μM ^{15}N -C45 in 100% redox buffer (20 mM CHES, 100 mM KCl, pH 8.6) containing 10% D_2O . A second sample was prepared with 80% deuterated TFE and 20% redox buffer with 10% D_2O .

All experiments were performed on a Bruker Avance II 700 MHz instrument equipped with a 1.7 mm TXI Z-gradient probe at 298 K.

2.2.9. Small Angle X-ray Scattering

1 mM C45 in redox buffer (20 mM CHES, 100 mM KCl, pH 8.6) and 0.7 mM C45 in 20% redox buffer and 80% TFE were used. Small angle x-ray scattering (SAXS) was conducted at 25 °C on a beam line directly connected to a Superdex 75 Increase 10/300 GL analytical size exclusion chromatography column. The experiments were conducted at Diamond Light Source in Didcot, using a bending magnet source beam

line which delivers approximately 3×10^{12} photons per second to the sample, over a cross-sectional area of 1.6 mm^2 . Data were fitted using ScÅtter and ATSAS software (240, 241).

2.2.10. Steady-state kinetic analysis

Whole spectra were taken on a Cary 60 UV-visible Spectrophotometer by Agilent Technologies with the Kinetics software by Agilent. Substrates and enzyme were added to a quartz cuvette immediately before the experiment and mixed by rapid pipetting.

The peroxidase activity at low substrate concentrations was observed on a Synergy neo2 multi-mode reader by BioTek Instruments using StarLab 96-well microplates with flat-bottomed, round wells. All reactants were preincubated in the experimental concentration of cosolvent. ABTS was injected into the wells and immediately prior to the measurement, H_2O_2 and enzyme were added, and the plate stirred. All additions of reactants were done using the instrument's built-in injector system to a total volume of $200 \mu\text{l}$.

2.2.11. Kinetic analysis with varying substrate concentrations

For rapid kinetic measurements with higher and varying substrate concentrations, an Applied Photophysics SX20 Stopped-Flow spectrometer was used. The instrument employed here was connected to a photodiode array detector. Prior to kinetic measurements, a baseline measurement of buffer with cosolvent was taken. All reactants were preincubated in cosolvent. Immediately prior to measurements, hydrogen peroxide and ABTS were mixed, and the two syringes filled with this mixture and with enzyme solution, respectively. Syringes and chamber were thoroughly flushed between measurements.

2.2.12. Data analysis

Data analysis and visualisation was performed using Microsoft Office programmes as well as Origin Lab Graphing Software, GraphPad Prism, ChemDraw Professional, PyMol, ScÅtter and ATSAS software (240–242).

3. Structural characterisation of C45 in organic cosolvents

In this chapter, the structure, especially secondary structure, and thermostability of C45 are analysed. The enzyme is monitored during the purification process and once fully purified, both in buffer and in a variety of aqueous-organic mixtures using UV-visible and CD spectroscopy. The protein's structural resilience to high temperatures is probed using CD temperature wavelength scans in buffer as well as aqueous-organic media.

3.1. Introduction

C45 consists of four α -helices connected by short loops. The fourth helix covalently binds a single *c*-type heme (19–21). Previous characterisation of C45 has largely focussed on the enzyme in conditions imitating those in which proteins have evolved in the natural environment. Typically this involves buffered solutions at neutral pH and room temperature (2, 21). Only limited data are available for the behaviour of C45 in hostile conditions such as organic solvents (2). These data indicate that the enzyme is highly thermostable and very catalytically efficient in temperatures up to 70 °C (and possibly higher). Furthermore, previous work suggests that peroxidase activity may increase slightly in low ethanol concentrations of up to 30%. Higher ethanol concentrations, however, lead to a strong decline in reaction rate most likely due to protein denaturation. Methanol does not appear to have any beneficial effect at low

concentrations, but the activity decline at concentrations above 30% is less pronounced than in ethanol.

Work with enzymes in organic solvents is not merely motivated by the curiosity-driven pursuit of a better understanding of enzymes, their chemistry and limitations. While this may have initiated research in this field, the intriguing results of improved thermostability and selectivity generated industrial interest in this field (13, 118, 243, 244). Furthermore, the ability to work in non-aqueous solvents facilitates the use of organic compounds insoluble in water for industrial biocatalysis (7).

In some instances, pure hydrophobic solvents have proved more beneficial to enzymatic activity than mixtures of organic solvents and aqueous media (13). This may be partly attributed to a pH memory effect – a term coined to describe the phenomenon of enzymes retaining their last ionisation state from aqueous media when dehydrated and dissolved in organic solvents (13). However, the use of organic solvents as media for proteins poses some challenges: Not only are dehydrated proteins often not particularly soluble in organic solvents, but the dehydration process, typically achieved by lyophilisation, can lead to protein denaturation as discussed in **section 1.3.1** (13). During lyophilisation, any water molecules tightly bound by the enzyme remain associated with it, potentially a contributing factor to the perhaps surprisingly good performances of some enzymes in neat, organic solvents.

In contrast to water immiscible organic solvents, miscible cosolvents may have a much stronger effect on an enzyme's hydration shell. Hydration is very important both structurally and functionally, and in homogeneous aqueous-organic mixtures, molecules from the hydration shell are more likely exchanged with the solution (13, 104). This can lead to the disruption of hydrogen bonding networks and unfolding of

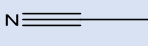
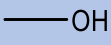
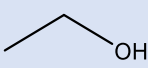
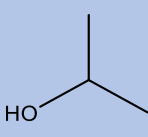
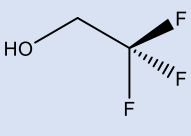
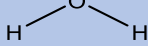
3. Structural characterisation of C45 in organic cosolvents

the protein in micelles, where the hydrophobic core of the protein is stabilised by the organic phase (245).

This work focusses on the characterisation of C45 in mixtures of aqueous buffering solutions and hydrophilic cosolvents. Their effect on the enzyme's structure and function will be probed by gradually increasing cosolvent concentration. The solvents used in this work are described in **Table 3.1**.

Methanol, ethanol and isopropanol are alcohols with increasing lengths of hydrocarbon chain, meaning an increasing proportion of the molecule exhibits hydrophobic properties. Acetonitrile has similar chemical properties to isopropanol in terms of polarity as measured by Lewis acidity (246), however it lacks a hydrogen bond donor. TFE may yield especially interesting results as it has been shown to stabilise helical proteins; in fact, TFE has been shown to alter the tertiary structure of some proteins to allow for more helicity than the native state possesses (4). The dielectric constant, which is a measure of polarity based on physical properties (246), has been demonstrated to be correlated to protein mobility (130). It is highest in water, meaning all cosolvents have significantly lower permittivity. These cosolvents are, however, not mixed with water, but with buffered solution containing salts raising solution permittivity. The negative impact of raising cosolvent concentration on the dielectric constant is therefore twofold: firstly, the cosolvent has a lower dielectric constant than water, secondly salt concentration is lower in higher cosolvent concentration. The table also shows $E_T(30)$ values, which is a common scale of solvent polarity from chemical properties, defined as the molar transition energy of betaine, a solvatochromic dye (246–248). Polarity is likely to affect the folding process which is strongly driven by the hydrophobic effect (249).

Table 3.1: Outline of solvents used in this work and some of their key properties.

Name	Abbreviation	Chemical formula	Structural formula	Boiling point	Dielectric constant	$E_T(30)$ (kcal mol ⁻¹)	pK _a
Acetonitrile	ACN	C ₂ H ₃ N		82 °C	38	47	25
Methanol	MeOH	CH ₄ O		65 °C	33	56	15.3
Ethanol	EtOH	C ₂ H ₆ O		78 °C	25	52	15.9
Isopropanol	IPA	C ₃ H ₈ O		82 °C	18	49	17.1
2,2,2-Trifluoroethanol	TFE	C ₂ H ₃ F ₃ O		74 °C	9	60	12.4
Water		H ₂ O		100 °C	80	63	14

Data obtained from (250–259).

The work presented in this chapter screens C45 for structural changes and thermostability in buffer and in a variety of solvents. First, C45 was analysed during the expression and purification process using UV-visible spectroscopy and SDS-PAGE.

CD spectroscopy can provide insights into the secondary structure of proteins in solution. During CD, a beam of circularly polarised light (**Figure 3.1 a**) is passed through a sample held in a quartz cuvette. Chiral molecules, such as amino acids, absorb left- and right-handed circularly polarised light differently. This effect does not only arise in a chiral centre, but also in species in an asymmetric environment, such as amino acid side chains (260–262).

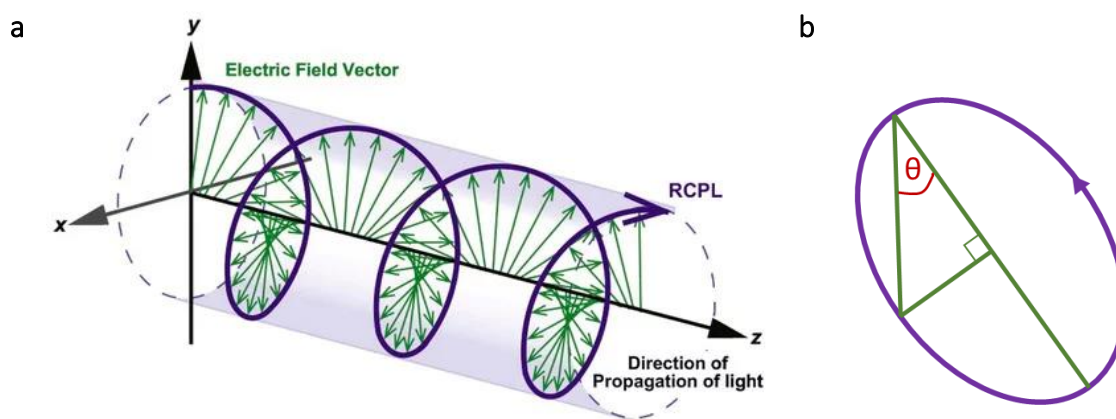


Figure 3.1: Circularly polarised light and its interaction with optically active matter.

a - Diagram visualising right-handed circular polarised light (RCPL). Figure taken from (263). **b** - The ellipticity angle θ , showing the rotation of the circular polarised light by an optically active sample in circular dichroism spectroscopy.

This leads to elliptically polarised light which has been rotated and distorted by the ellipticity angle θ (**Figure 3.1 b**). For comparability of CD results, the ellipticity is converted to molar residue ellipticity (θ) by normalising for sample concentration, protein sizes and path lengths of light using the following formula (264) where n refers to the number of peptide bonds:

$$\theta \text{ (deg cm}^2 \text{ dmol}^{-1}\text{)} = \frac{\text{Ellipticity (mdeg)}}{\text{Path length (mm)} * [\text{Protein}] \text{ (M)} * n} \quad 3.1$$

Absorption of the circularly polarised light in the far UV region (180-250 nm) is owed to excitation of electronic transitions in the amide groups arising from the peptide bonds in the protein's backbone (230). Different structural motifs of the peptide backbone, such as α -helices and β -sheets, have distinctive dihedral angles and patterns of hydrogen bonding which, in turn, result in distinctive CD spectra (264). Typical spectra for proteins in aqueous solution with different secondary structures are shown in **Figure 3.2**.

Proteins with α -helices, such as C45, exhibit a characteristically double-dip shaped spectrum with a strong positive band between 190 nm and 200 nm and two strong negative bands at 208 nm and 222 nm. The maximum as well as the minimum at 208 nm are attributed to $\pi \rightarrow \pi^*$ transitions and the strong negative band at 222 nm to $n \rightarrow \pi^*$ transitions (260, 265). In β -pleated proteins, a similar maximum between 190 nm and 200 nm can be observed, also arising from a $\pi \rightarrow \pi^*$ transition, with $n \rightarrow \pi^*$ transitions causing a minimum at approximately 215-218 nm (265, 266). The different spectra are owed to the dihedral angles in the protein backbone (α : -57° , -47° ; β : -120° , $+120^\circ$) which differ in magnitude and direction between the two secondary structural motifs (260, 267). Generally, signals in the range of 190 nm to 240 nm arise from transitions within the peptide bond (260).

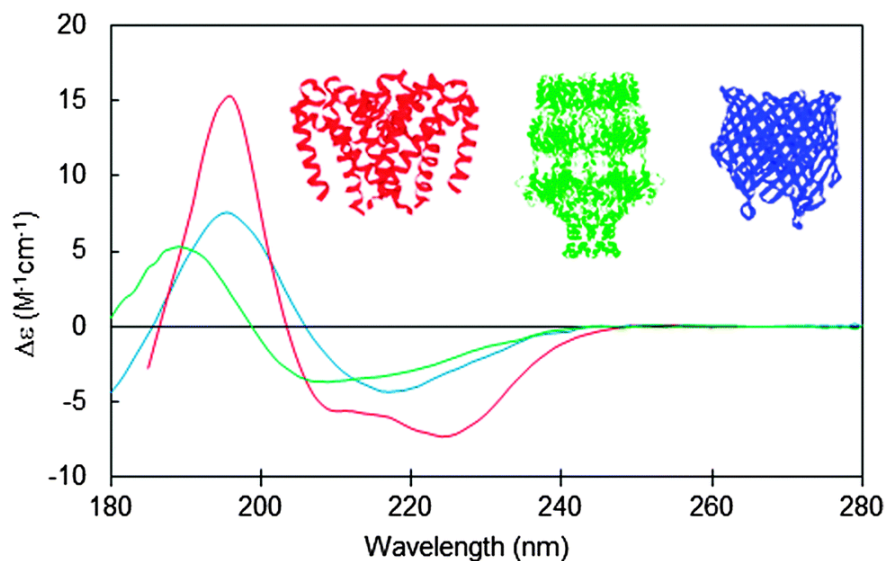


Figure 3.2: Example of CD spectra of proteins with different secondary structures.

The red spectrum is that of a predominantly α -helical bundle, the blue spectrum is of a β -barrel and the green spectrum belongs to a protein exhibiting mixtures of the aforementioned motifs as well as unordered structure. The structure of each protein is visualised in the same colour as the corresponding spectrum for reference. Figure taken from (268).

3.2. Results

3.2.1. Expression and purification of C45

C45 was expressed in *E. coli* with a 6-histidine tag and purified using nickel affinity chromatography, TEV cleavage, further nickel affinity chromatography and size exclusion chromatography. The latter type of chromatography can separate proteins by size and provide some information about the quaternary structure of the protein. The chromatogram indicating the elution of three peaks is shown in **Figure 3.3**. Fractions were taken at different points during the elution and analysed using UV-visible spectroscopy, which can indicate the samples' heme content. The UV-visible spectra shown in **Figure 3.4 a**, where each spectrum is shown in the colour that was also used to show the chromatography fraction it corresponds to in **Figure 3.3**.

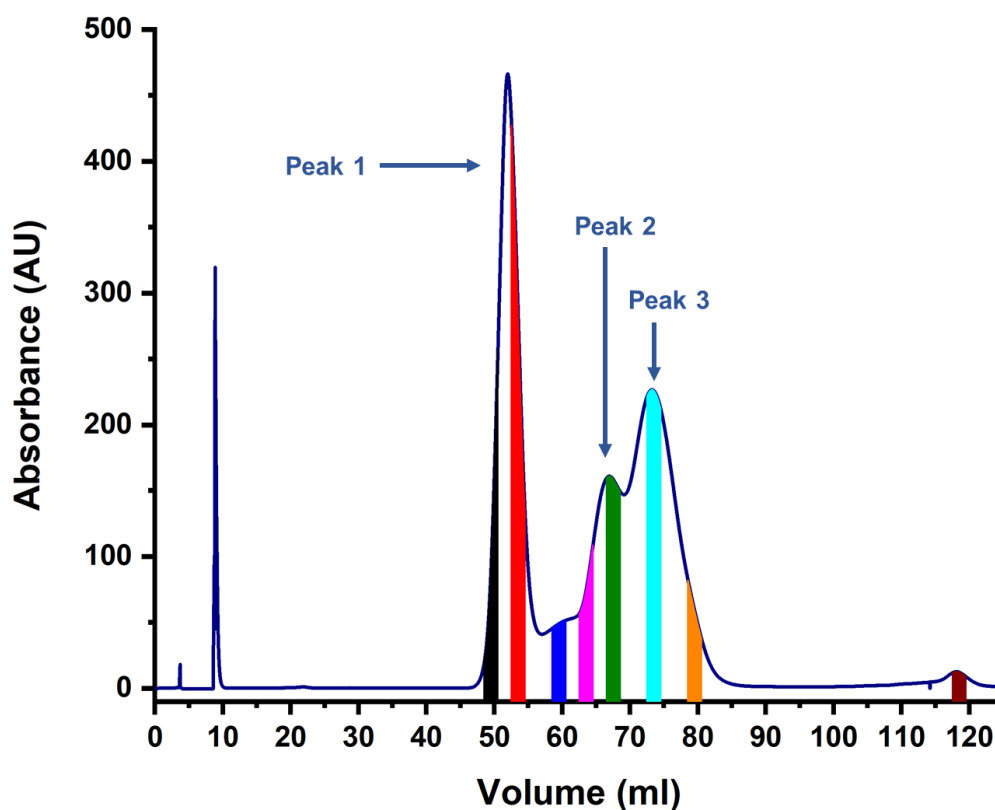


Figure 3.3: Chromatogram from size exclusion chromatography showing the absorbance at 280 nm. Coloured fractions were analysed *via* UV-visible spectroscopy.

The fractions from the first elution exhibit a small absorption in the region of the Soret band, compared to the strong peak at 280 nm. The low heme content in these fractions is likely due to apoprotein, suggesting incomplete heme incorporation by the cytochrome *c* maturation apparatus.

The lower absorbing region eluting between the first and the second peak on the chromatogram shows a clearly visible Soret band, potentially caused by C45 aggregates. The red shifted Soret band indicates a hexacoordinate heme iron (269). This could mean the metal coordinates some species present in the buffer, such as imidazole. It is unlikely, however, that this was bound to the heme sufficiently tightly to avoid removal by size exclusion chromatography. A more plausible alternative is that the protein is present in oligomeric form and heme bis-histidine ligated between two binding sites.

The second and third elution peaks contain C45. Since oligomerisation had previously been observed in maquette proteins (270), it is likely that the enzyme is present in dimeric (second peak) and monomeric (third peak) forms. All fractions from these elution peaks were pooled.

C45 was characterised during each step of the protocol using UV-visible spectroscopy and SDS-PAGE. Samples were taken of the cell pellet, the lysate, after the addition of TEV protease as well as after each run of chromatography. The gel (**Figure 3.4 b**) clearly demonstrates the effectiveness of the purification process: the cell pellet and lysate (Lanes A and B) contain a multitude of intracellular proteins, as indicated by the numerous blue bands. Nickel affinity chromatography removes the vast majority of these, which, in contrast to C45, do not bind tightly to the nickel resin (Lane C). However, multiple bands are still visible: the band at just over 15 kDa

3. Structural characterisation of C45 in organic cosolvents

corresponds to the His-tagged C45. After the addition of TEV protease (Lane D), this band is visible further down the gel, as the cleavage resulted in a smaller protein. The second round of nickel affinity chromatography removes TEV protease, the cleaved histidine tag and any protein impurities which bind the nickel resin as indicated by a lack of bands in Lane E. The last two lanes (F and G) correspond to peak 1 of the size exclusion chromatography and the pooled fractions of peak 2 and 3, respectively. The first peak to elute during the chromatography shows not a single clear band, but a mixture of low intensity bands of all masses. This band confirms the presence of several impurities. The final sample of C45 used for further protein work, on the other hand, clearly shows a single visible band at 15 kDa (Lane G).

Expressions yielded approximately 2.4 mg C45 per litre expression culture. This concentration was determined from absorption at 406 nm, where C45 has a known extinction coefficient of $147,300 \text{ M}^{-1} \text{ cm}^{-1}$ (2). Importantly, this will only account for holoprotein, as any apoprotein (where the heme cofactor was not incorporated) would not absorb at this wavelength. Consequently, it is possible that protein concentrations were larger than estimated, however, it is unlikely that apoprotein is catalytically active.

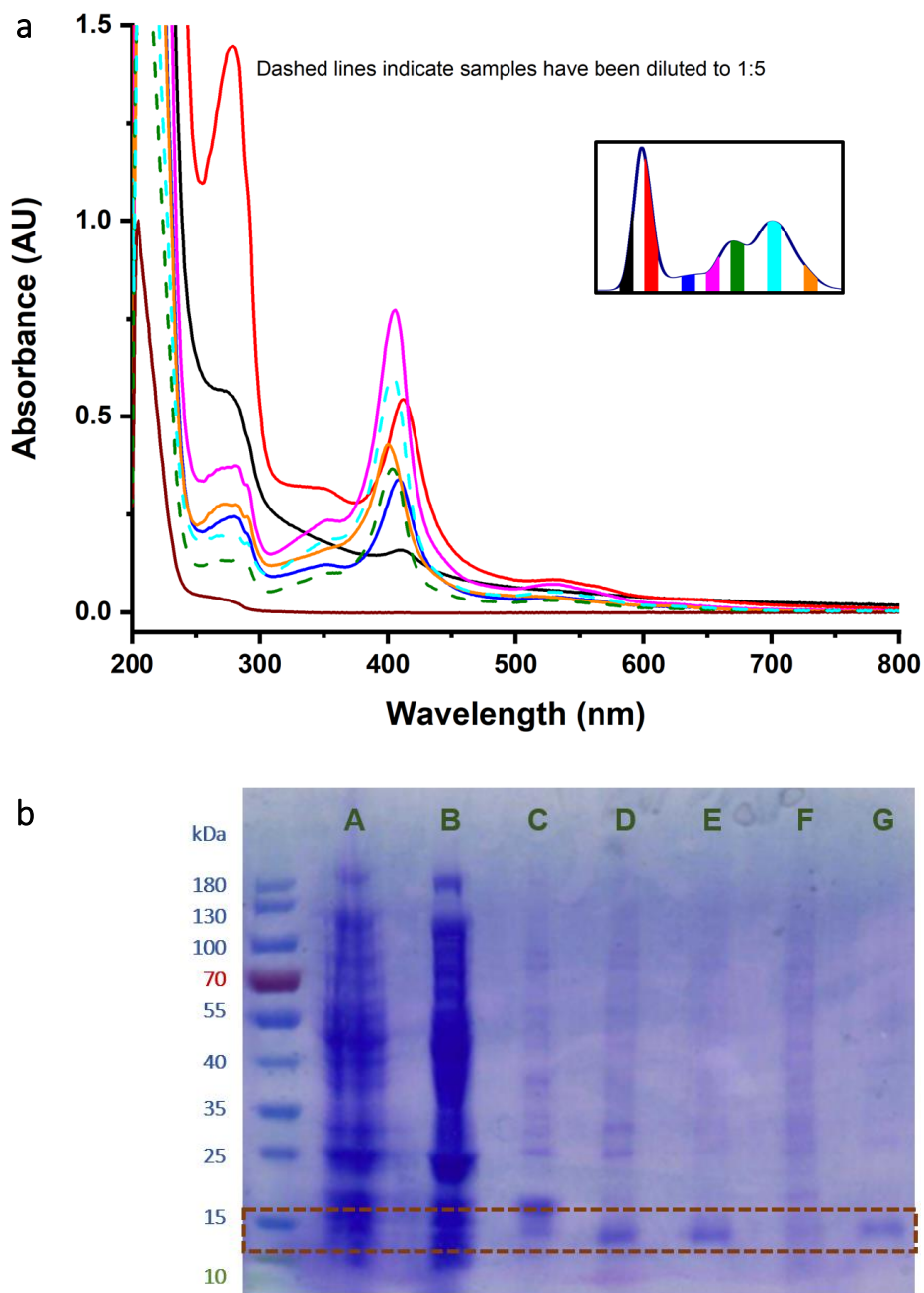


Figure 3.4: Analysis of C45 at different stages of its purification and different chromatography fractions.

a - UV-visible spectra of the fractions taken from size exclusion chromatography (see **Figure 3.3**). Colours correspond to the matching coloured fractions in the chromatogram, which is inset on the right side for reference. Samples were not concentrated or diluted, except for the two dashed spectra, for which the samples were diluted 1:5.

b - SDS-PAGE Gel showing C45 at different stages during the purification process. A – Cell pellet. B – Supernatant after sonication and centrifugation of cell pellet. C – post nickel affinity chromatography prior to TEV cleavage. D – post addition of TEV protease. E – post nickel affinity chromatography after TEV cleavage. F – peak one from size exclusion chromatography. G – 2nd and 3rd peak from size exclusion chromatography prior to concentration: The single band at 15 kDa corresponds to C45.

3.2.2. Pyridine Hemochrome

The pyridine hemochrome of C45 (**Figure 3.5**) confirmed the presence of cofactor heme, which corresponds to the Soret band seen at 409 nm and 413 nm in oxidised (ferric) and reduced (ferrous) protein, respectively. It reproduces the expected spectra for *c*-type cytochromes. The location of the Q-bands at 520 nm and 551 nm for ferrous heme are typical for a protein binding heme C (227).

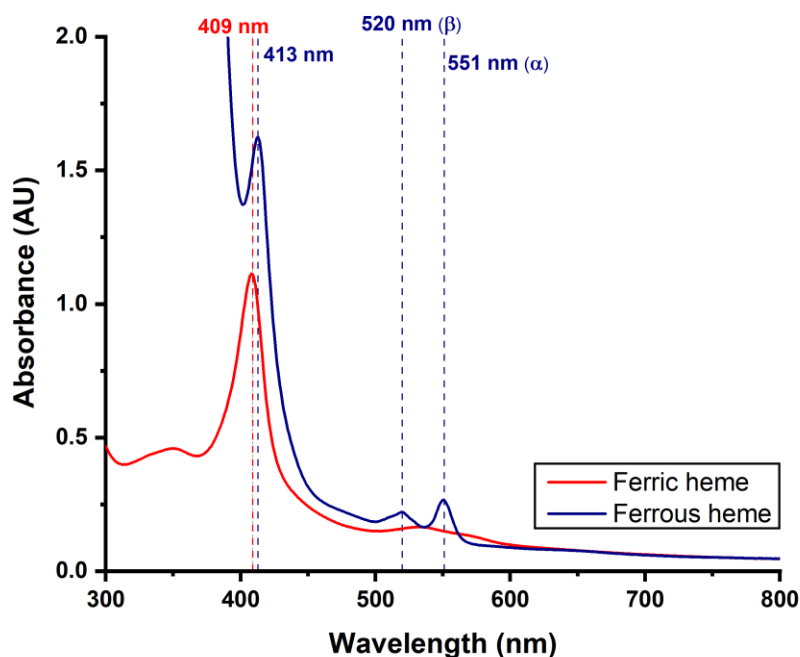


Figure 3.5: Pyridine hemochrome of C45.

3.2.3. UV-visible spectroscopy

C45 was analysed by UV-visible spectroscopy to probe the chemical environment of heme in increasing concentrations of cosolvents. The spectrum of C45 in buffer (**Figure 3.6 a**) indicates that the iron within the protein is primarily in the ferric (+3) oxidation state. The Soret band is clearly visible at 405 nm.

UV-visible absorption spectra were also taken of C45 in various solvents (**Figure 3.6 b-f**). The location of the Soret peak at different protein concentrations is shown in **Figure 3.7**. In methanol, ethanol and TFE, the Soret peak occurs at decreasing wavelengths as cosolvent concentration is raised. In isopropanol and acetonitrile no clear trend is observed, although repeat measurements may be beneficial. Changes in Soret band location are likely related to the electrostatic field the protein is exposed to, which changes with the addition of cosolvent (271). In all cosolvents, however, the Soret peak is located below 408 nm, implying mono-histidine ligation of heme.

In high (70-80%) concentrations of acetonitrile, ethanol and isopropanol, scattering was observed as indicated by the significant increase of baseline absorption. This also led to a slight shift in location of the Soret band. In line with the scattering observed, the solutions appeared turbid. Surprisingly, this effect diminishes over time (**Figure 3.8**). For instance, the spectrum of C45 in 80% TFE remains unchanged over an incubation period of several hours, whereas the corresponding spectrum in 80% ethanol shows significantly less scattering after five hours. This may indicate resolution of C45 in the aqueous-organic solution over time.

3. Structural characterisation of C45 in organic cosolvents

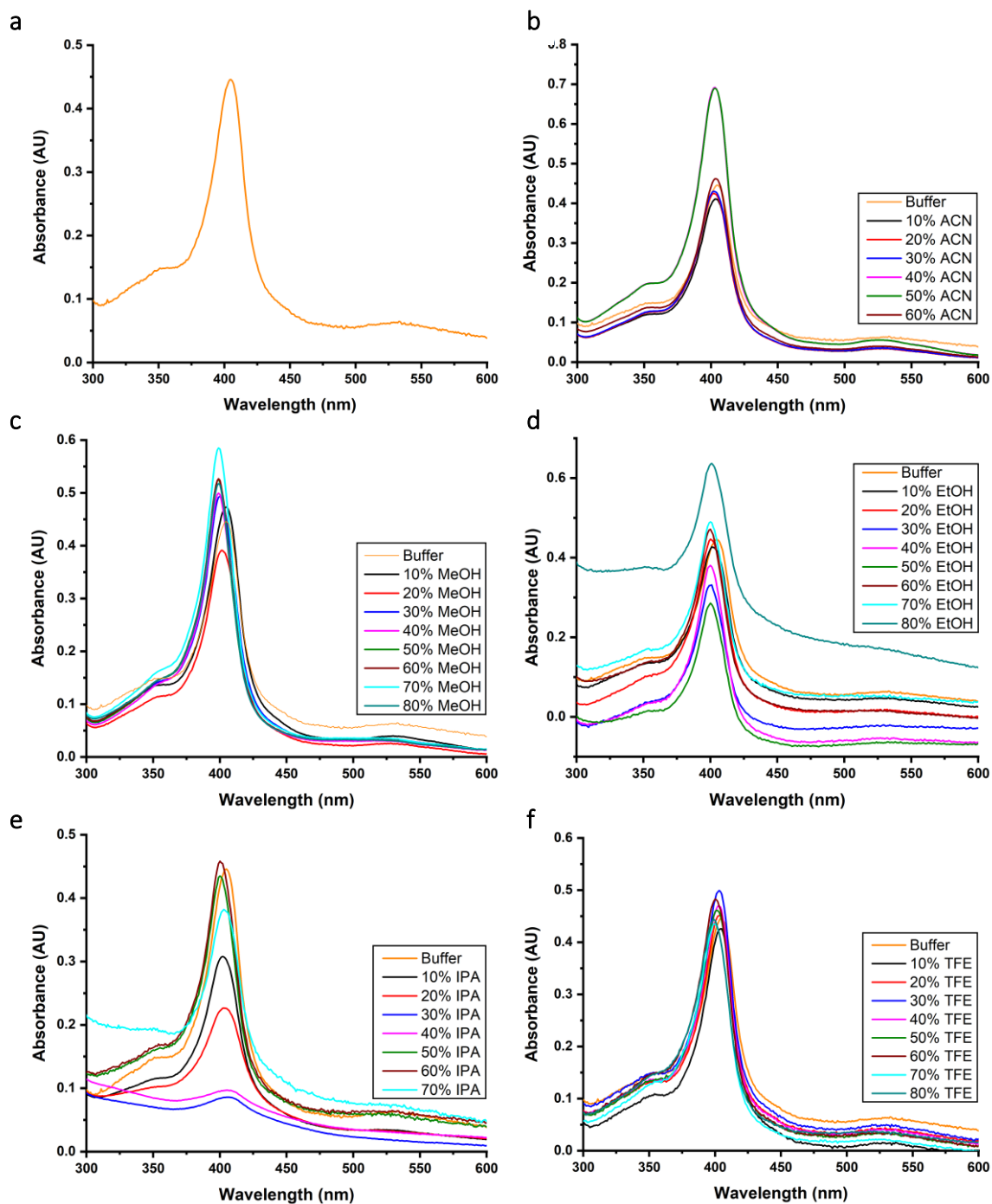


Figure 3.6: UV-visible absorption spectra of C45 in aqueous-organic solutions.

C45 in solutions containing no cosolvent (a) or varying amounts of acetonitrile (b), ethanol (c), isopropanol (d), methanol (e) and TFE (f). Results in 70-80% Acetonitrile and 80% IPA removed due to excessive scattering.

3.2. Results - Circular Dichroism of C45 at room temperature

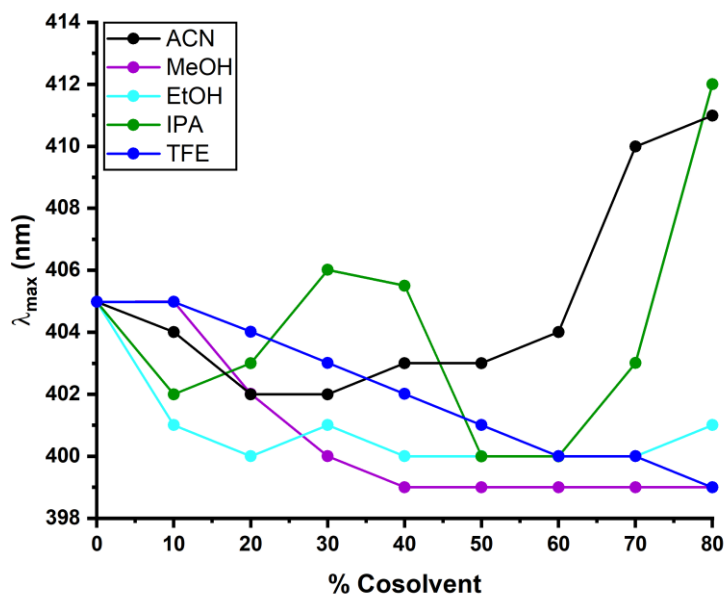


Figure 3.7: Wavelength of Soret peak for C45 in different cosolvents at increasing concentrations.

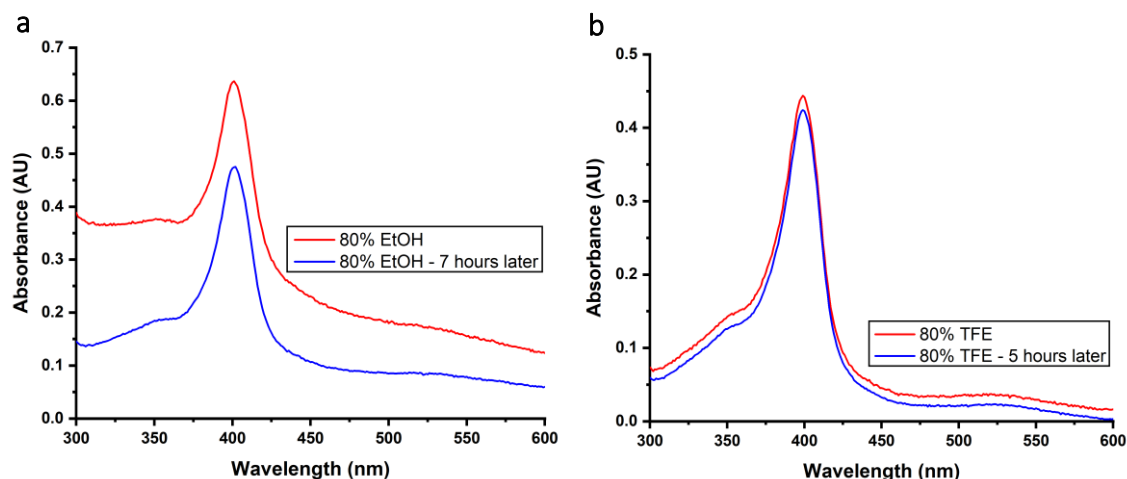


Figure 3.8: Declining scattering of C45 in 80% ethanol compared to 80% TFE.

UV-visible absorption spectra of C45 in 80% ethanol (a) and 80% TFE (b) immediately after mixing and after an incubation time, during which the scattering in the ethanol solution reduces and the spectrum in TFE remains.

3.2.4. Circular Dichroism of C45 at room temperature

Further analysis of the protein involved room temperature CD spectroscopy in buffer and cosolvents (Figure 3.9), which can indicate the presence of some motifs of secondary structure, such as α -helices or β -sheets. The spectrum of C45 in buffer

3. Structural characterisation of C45 in organic cosolvents

exhibits the features characteristic of α -helices, namely a maximum at approximately 197 nm and minima at 208 nm and 222 nm (44, 272). Generally, the data obtained at wavelengths below 200 nm are associated with higher errors due to large high-tension (HT) voltages at lower wavelengths. This leads to lower signal to noise ratios (273). Most of the analysis herein will therefore focus on the minimum at 208 and especially the minimum located at approximately 222 nm, with samples prepared and experiments performed as described in **section 2.2.7**.

In up to 60% of all cosolvents as well as 70% ethanol and up to 80% methanol and TFE, the CD spectra still clearly display this shape. This conveys that in these cosolvent concentrations, C45 is predominantly helical. In 70-80% acetonitrile, 80% ethanol and 70-80% isopropanol, however, this is not the case anymore and the protein appears to be denatured.

A comparison of the intensities of the peaks in different cosolvent concentrations (**Figure 3.10**) shows that the signal intensity for both minima is consistent in solutions with up to 60% acetonitrile. In all other cosolvents – the four alcohols – this signal increases along with cosolvent concentration for the non-denatured protein. This suggests a higher degree of helicity, which had previously been observed in alcohols and especially in TFE (3). This trend may, however, be slightly exaggerated here as the concentration of enzyme in the solution may be slightly larger in higher alcohol concentrations, as the addition of alcohol and water can lead to a smaller volume than the sum of the two volumes added (274).

3.2. Results - Circular Dichroism of C45 at room temperature

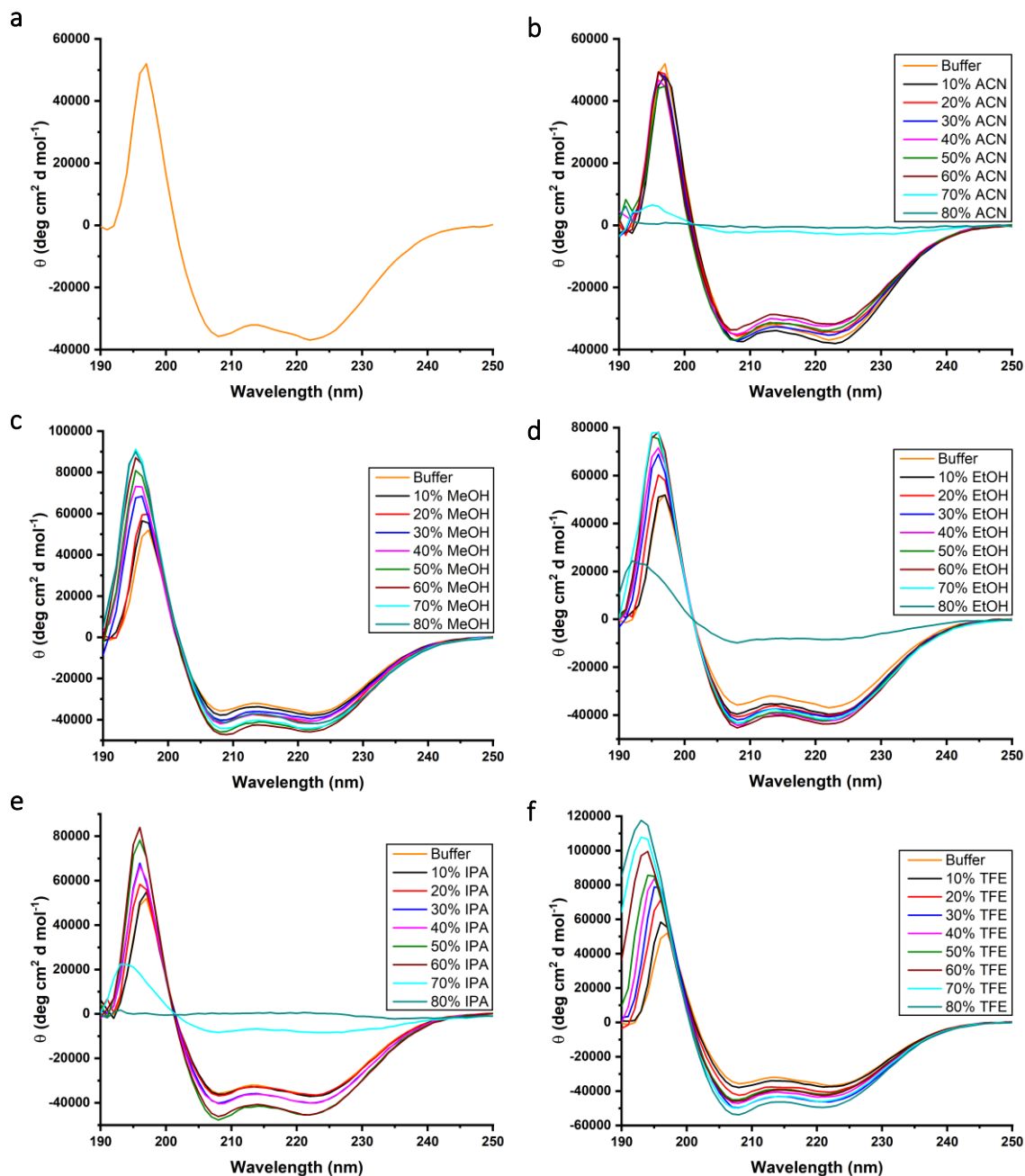


Figure 3.9: Room temperature CD spectra of C45 in aqueous solution with organic cosolvents. C45 in aqueous solution and (a) varying amounts of acetonitrile (b), methanol (c), ethanol (d), isopropanol (e) and TFE (f). The spectrum in buffer (a) indicates a well-folded α -helical protein. The other spectra show that this protein is stable in up to 60% isopropanol and in up to 70% ethanol and acetonitrile. No loss of secondary structure is observed in up to 80% TFE or methanol concentrations.

3. Structural characterisation of C45 in organic cosolvents

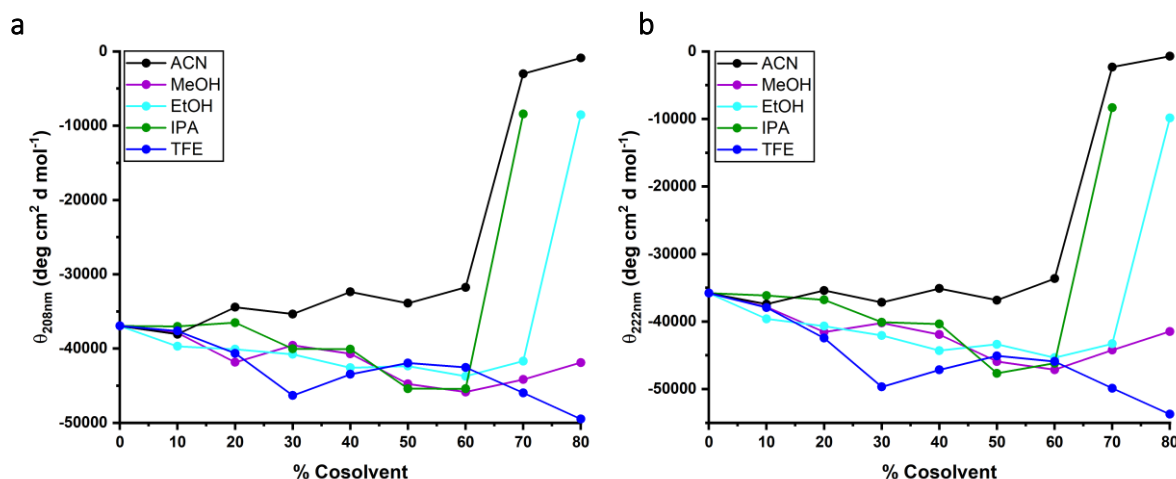


Figure 3.10: Molar residue ellipticity of C45 in increasing cosolvent concentrations.

Θ at (a) the minimum at approximately 208 nm and (b) the minimum at approximately 222 nm, characteristic for helical proteins. The signal intensifies with increasing concentrations of alcohol and is consistent in acetonitrile, except in cosolvent concentrations where the protein is denatured (70-80% acetonitrile, 70-80% isopropanol, 80% ethanol).

3.2.5. CD Thermal denaturation studies

3.2.5.1. Thermal denaturation of C45 in buffer

To probe the enzyme's thermostability, C45 was analysed using variable temperature circular dichroism (CD) spectroscopy. The protein was heated from 8 °C to 85 °C at a rate of 1 °C per minute, with ellipticity constantly monitored at 222 nm (**Figure 3.11 a**), a minimum for α -helical proteins, as seen in **Figure 3.9 a**.

The thermal denaturation curve (**Figure 3.11 a**) can be split into two sections: Initially, the increase in ellipticity is gradual, indicating a non-cooperative loss of helicity. In line with literature, this occurs most likely in the form of fraying at the termini of the α -helices (3). However, once elevated to above 64 °C, the steeper curve suggests that the protein denatures more cooperatively. This alludes to the enzyme simultaneously unfolding at multiple sites at higher temperatures, possibly around the active site situated between the four helices.

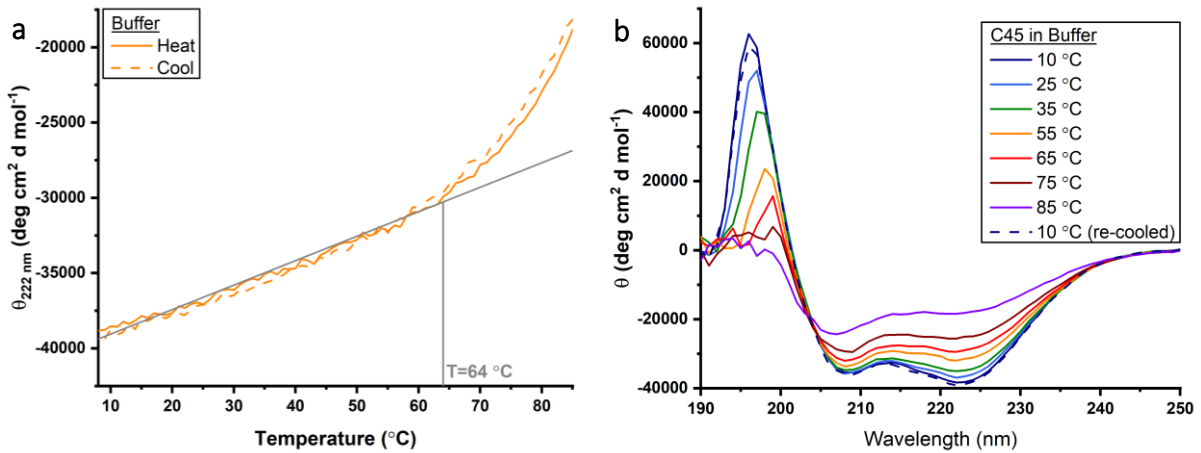


Figure 3.11: CD thermal denaturation spectra of C45 in buffer.

a - Molar residue ellipticity measured at 222 nm during the heating and subsequent re-cooling process. The overlapping curves indicate that thermal denaturation of the protein is reversible. Grey line superimposed for visualisation only. **b** - Full spectra of the sample at various, discrete temperatures showing good thermostability of the protein in buffer

A further interesting observation is the refolding profile of C45. After thermal denaturation, cooling was immediately initiated, and the refolding process assessed using CD. This trace, also monitored at 222 nm, is displayed as a dotted line in **Figure 3.11 a**, which overlaps with the heating curve, meaning the thermal denaturation of C45 is entirely reversible and the protein spontaneously refolds again.

CD spectra obtained at different points during the experiment (**Figure 3.11 b**) also clearly convey the thermal denaturation and refolding, as indicated by overlapping spectra prior to and after the experiment. Furthermore, the ratio between ellipticity values at 222 nm and 208 nm decreases. This is sometimes seen as an additional marker of helicity, with high ratios (>1) suggesting a tight, coiled structure and lower ratios isolated helices or even denatured helices (275, 276). The decreasing ratio further corroborates the loss of structural stability with increased temperature.

3. Structural characterisation of C45 in organic cosolvents

Unfortunately, no definite conclusions can be drawn from these results about folding mechanisms and cooperativity. It would be interesting to be able to heat the enzyme to a higher temperature to fully observe the denaturation. However, since ellipticity is heavily concentration dependent, it is crucial to avoid solvent evaporation which could lead to false or misleading data which is why heating was terminated at 85 °C and the cuvette sealed. An alternative technique to measure the unfolding of C45 would be CD in the presence of a denaturant or a titration with guanidine hydrochloride (277).

3.2.5.2. Thermal denaturation of C45 in Acetonitrile

CD thermal denaturation studies were also performed in 10-80% acetonitrile and ellipticity monitored at 222 nm (**Figure 3.12** and **Figure 3.13 a**). Grey lines are overlaid to show the initial, linear phase of unfolding, and the temperature at which folding becomes more cooperative.

Due the boiling point of acetonitrile being lower than that of water, the heating of C45 had to be terminated at a lower temperature, at 70 °C as opposed to 85 °C. The thermal denaturation curves of C45 in acetonitrile show that in 10% acetonitrile C45 unfolds similarly as it does in buffer, with some degree of cooperativity suggested by the steepening of the curve above 60 °C. In 20-30% acetonitrile, this change in curve gradient is still observable, but at lower temperatures (50 °C and 40 °C, respectively) suggesting that the enzyme loses thermostability in increasing cosolvent concentrations. In higher acetonitrile concentrations, unfolding is even more cooperative.

3.2. Results - CD Thermal denaturation studies

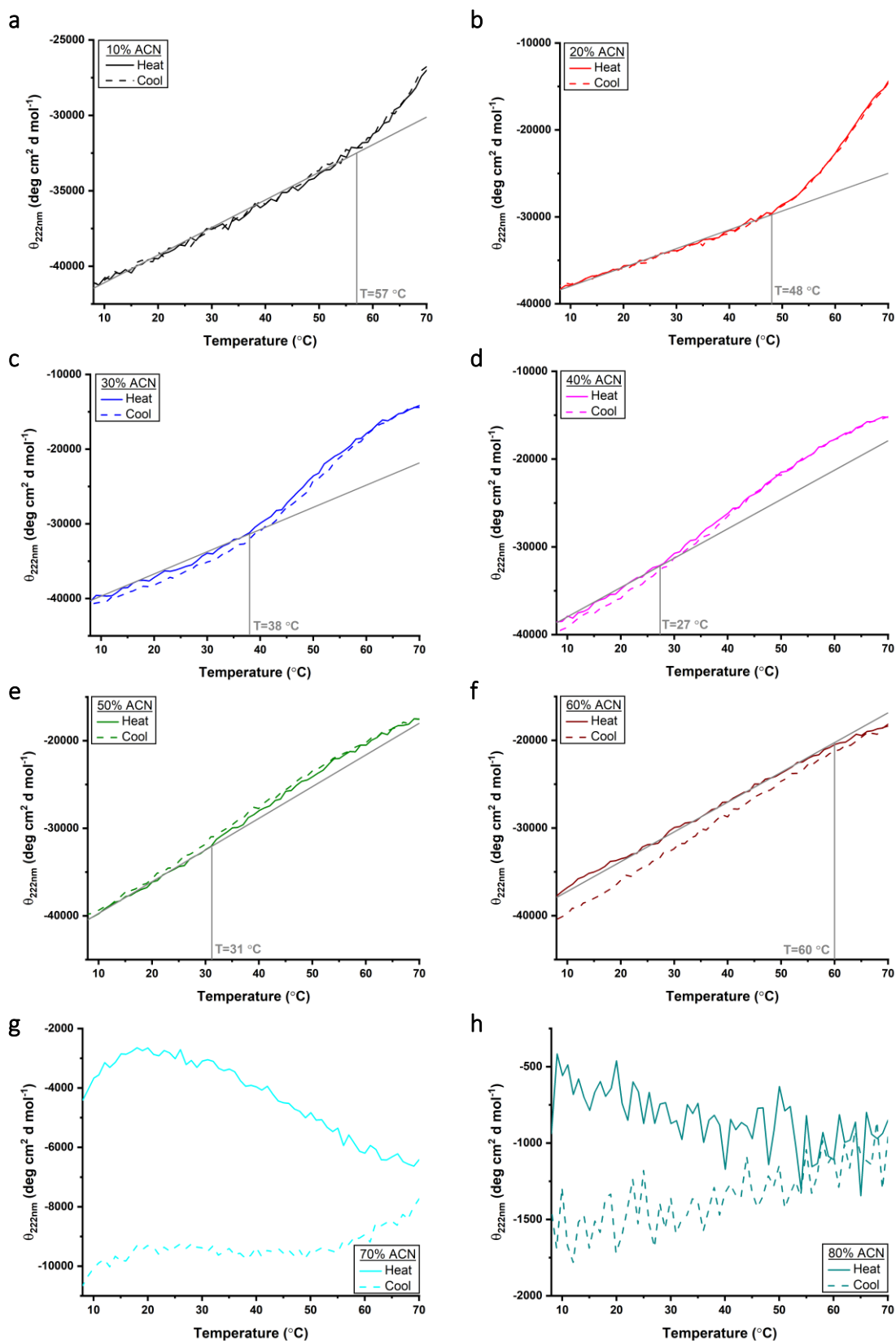


Figure 3.12: CD scan of C45 in aqueous solution with acetonitrile during heating and cooling. C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) acetonitrile. In up to 60% acetonitrile concentration, thermal denaturation is reversible. Grey lines superimposed for visualisation purposes only.

3. Structural characterisation of C45 in organic cosolvents

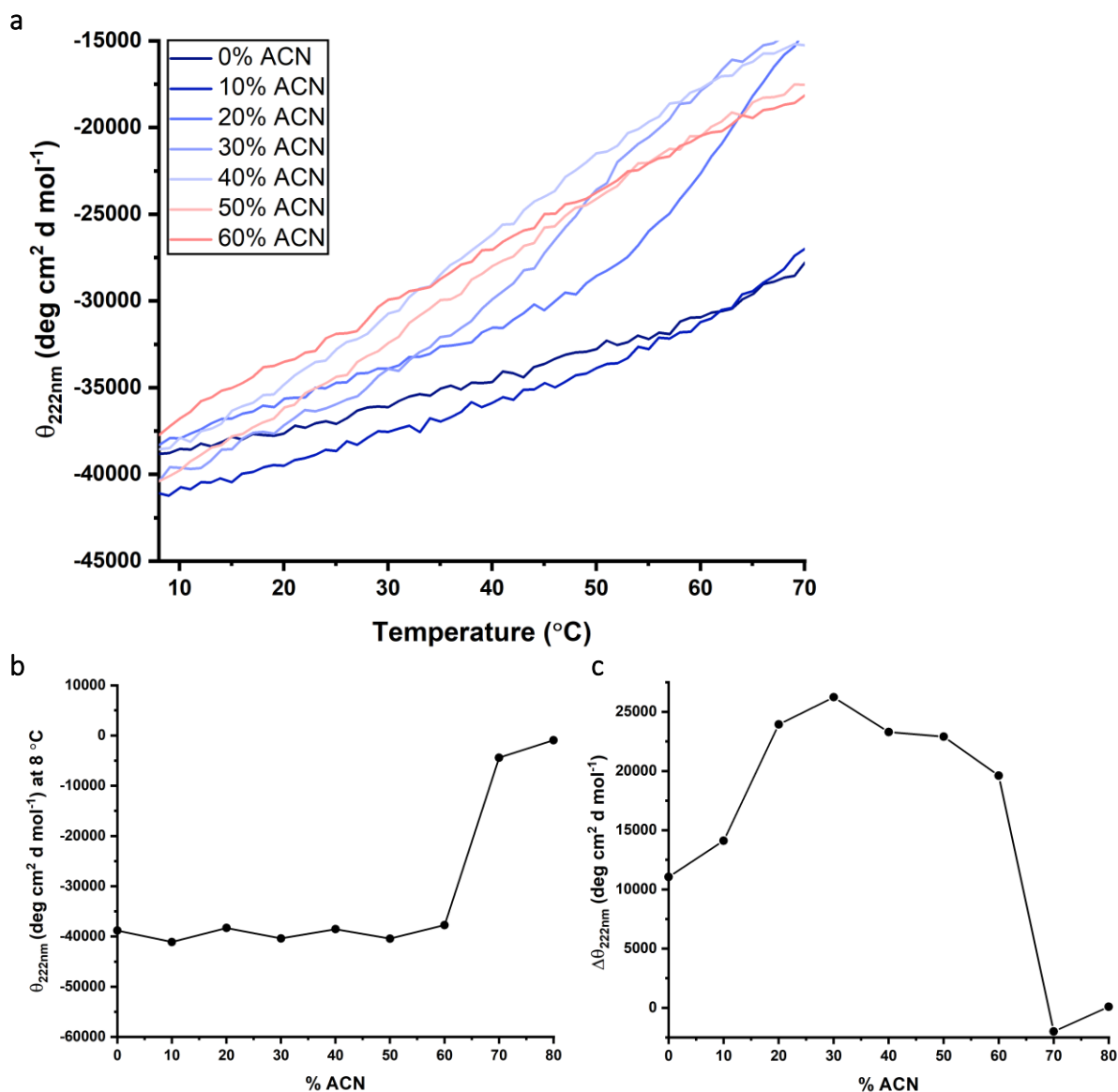


Figure 3.13: Thermal denaturation curves of C45 in aqueous solution with acetonitrile.

a – Overlay of thermal denaturation curves (monitored at 222 nm) in all acetonitrile concentrations tested. **b**– Ellipticity at 8 $^{\circ}\text{C}$ showing no change in acetonitrile concentrations of up to 60%. **c**–Difference in ellipticity 222 nm of C45 in increasing acetonitrile concentrations between 8 $^{\circ}\text{C}$ and 70 $^{\circ}\text{C}$.

In 40-60% acetonitrile, the change in gradient cannot be observed, but the overall shape of the curve is steeper, indicating more cooperative unfolding, though at the starting temperature, θ is similar (-37,000 to -42,000 $\text{deg cm}^2 \text{ d mol}^{-1}$), as shown in **Figure 3.13 b**. In all of these acetonitrile concentrations, C45 completely refolds. Furthermore, no melting temperature could be observed for acetonitrile concentrations

up to 60%, indicating C45 retains its secondary structure well even in cosolvent and at temperatures of up to 70 °C.

In 70% acetonitrile, molar residue ellipticity at 8 °C is at approximately $-4200 \text{ deg cm}^2 \text{ d mol}^{-1}$, suggesting that there is a significant loss of secondary structure. During the course of heating and the subsequent cooling, this value decreases to approximately $-11,000 \text{ deg cm}^2 \text{ d mol}^{-1}$. In 80% acetonitrile, the protein is entirely denatured with no secondary structure present.

Figure 3.13 c shows the difference in ellipticity at 222 nm between the initial and final temperatures of 8 °C and 70 °C indicating the loss of secondary structure between those temperatures. The data suggest that in 20-60% acetonitrile the loss of secondary structure is larger than in buffer or 10% acetonitrile, indicating that C45 is less thermostable in those concentrations. In higher acetonitrile concentrations, C45 is either denatured or precipitated or a mixture.

The theory that C45 loses thermostability in increasing acetonitrile concentrations is also supported by the ellipticity at 222 nm monitoring loss of helicity, as well as the full spectra obtained at various points throughout the heating process (Figure 3.14). In 10-20% acetonitrile, the spectrum at 70 °C does not exhibit the two minima typical of helical proteins, similar to the spectrum in buffer at 75 °C. The spectrum in buffer at 55 °C, however, does still display this typical shape with a maximum at 197 nm and minima at 208 nm and 222 nm. In acetonitrile concentrations of 30% or above, on the other hand, the spectra obtained at 50 °C do not follow this shape, suggesting some loss of helicity.

The general trend of declining thermostability of the protein in increasing acetonitrile concentration is not unanticipated. However, the decreasing ellipticity

during the heating of C45 in 70% acetonitrile was intriguing and the experiment was repeated multiple times (**Figure 3.15**). Once reproducibility of the data was established, this result was probed further. At the start of a melting experiment, the sample is usually cooled rapidly to the starting temperature. In order to probe whether this initial cooling of the sample at a high and uncontrolled rate was responsible for some form of denaturation of the protein, the initial cooling rate was set to 0.5 °C per minute. This was also used as the rate for heating and cooling to establish whether the increased helical signal after cooling is related to the heating rate. However, despite the slower heating of the sample, the phenomenon of increased helicity after heating and cooling was observed again. The same sample was then re-heated and re-cooled twice more at 0.5 °C per minute. While for the second and third cooling cycle the signal became weaker during the heating process, the trend of increased ellipticity upon complete re-cooling, as opposed to the initial measurement at the same temperature prior to heating was consistent and reproducible. The minor solvent loss (not above 10% volume) alone cannot account for the observed increase in molar residue ellipticity with heating and cooling of the protein solution.

Even though both the melting curves and the full CD spectra at various temperatures (**Figure 3.15 b**) seem to suggest an increase in helix stability upon heating and re-cooling, this is unlikely. It is more conceivable that some precipitation of enzyme occurred during the rapid addition of cosolvent which was at least partially reversed during the heating process, raising the concentration of C45 on which CD results are very dependent. This suggests that in 70% acetonitrile C45 retains at least some of its secondary structure, which however is only observable when the solution is thoroughly mixed without any precipitation.

3.2. Results - CD Thermal denaturation studies

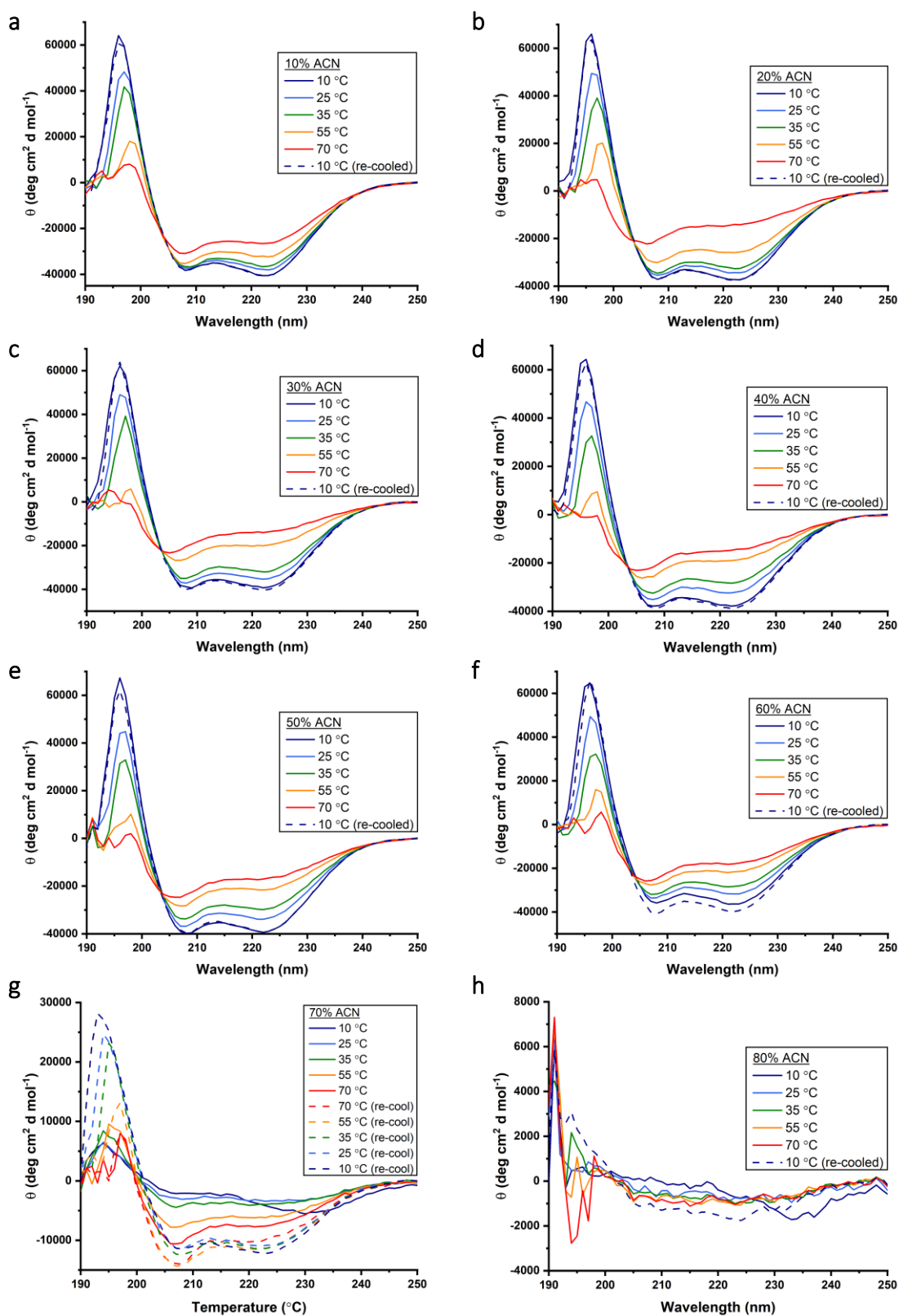


Figure 3.14: CD spectra of C45 in aqueous solution with acetonitrile during heating and re-cooling. C45 in aqueous solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) acetonitrile showing good folding and thermostability of C45 in up to 60% acetonitrile.

3. Structural characterisation of C45 in organic cosolvents

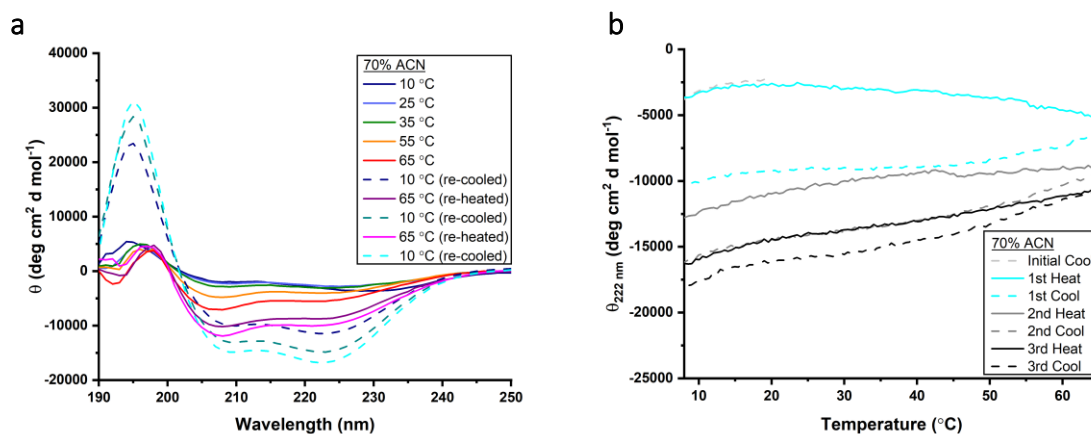


Figure 3.15: CD measurements of C45 in buffering solution with 70% acetonitrile.

a - θ at 222 nm during three cycles of heating and cooling at a rate of 0.5 °C per minute after initially being cooled from room temperature to 8 °C at the same rate (labelled “initial cool”). **b** - Full CD spectra taken during the first heating phase and when re-cooled and re-heated.

3.2.5.3. Thermal denaturation of C45 in Methanol

The same method was used to probe the effects of methanol, the smallest possible alcohol, on the structure and thermostability of C45, though due to methanol's lower boiling point, heating was terminated at 55 °C. The continuous scans at the minimum of 222 nm (**Figure 3.16**) shows that at the initial temperature of 8 °C, the ellipticity decreases from -37,000 deg cm² d mol⁻¹ to approximately -50,000 deg cm² d mol⁻¹ with increasing methanol concentration (**Figure 3.17**). This is likely to be indication of an increase in helicity, but is probably overestimated due to the decrease in volume of methanol-water mixtures which would effect a higher protein concentration in the sample and thus a stronger CD signal.

With increasing methanol concentration, the potentially more cooperative phase of unfolding occurs at earlier temperatures: While not observable up to the maximum measured temperature of 55 °C in up to 40% methanol, cooperative unfolding begins at 47 °C and 41 °C for C45 in 50% and 60% methanol, respectively. In higher methanol concentrations, any observed loss of secondary structure occurs more cooperatively

than in buffer or low methanol concentrations. This is particularly obvious in the overlay of the thermal denaturation curves (**Figure 3.17 a**). **Figure 3.17 b** and **c** show the general trend of decreasing ellipticity at the starting temperature (8 °C), but increasing differences in ellipticity between 8 °C and 55 °C with rising methanol concentration, highlighting the more cooperative nature of unfolding in higher methanol concentrations. Nonetheless, at 55 °C, molar residue ellipticity in higher methanol concentrations is lower, suggesting that even at this temperature, the cosolvent stabilises the protein's secondary structure, but not to the same extent as it does in low temperatures.

No melting temperature could be clearly observed in any methanol concentration. Furthermore, the protein refolds completely in all methanol concentrations.

3. Structural characterisation of C45 in organic cosolvents

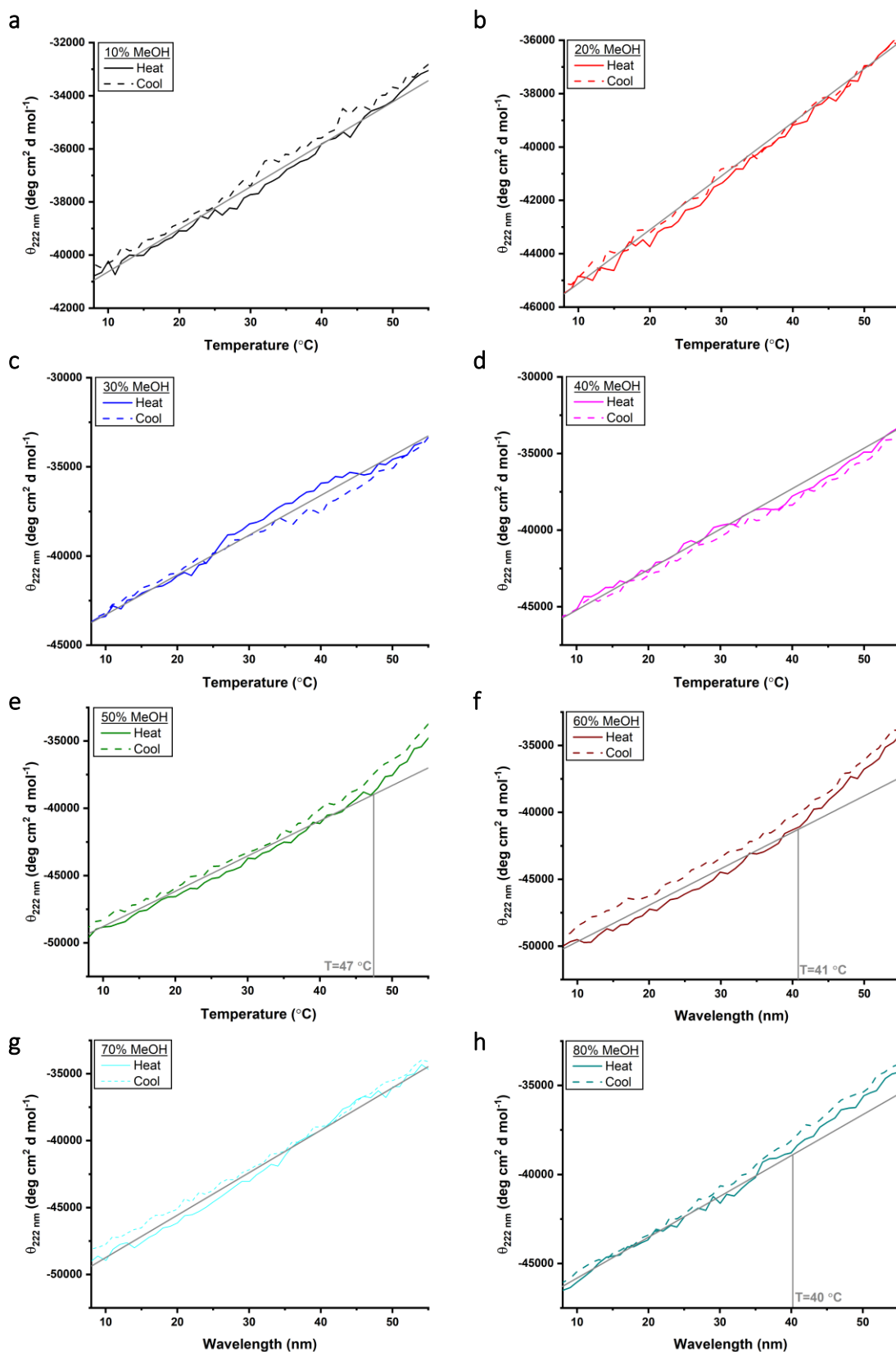


Figure 3.16: CD scan of C45 in aqueous solution with methanol during heating and cooling.

MRE (Θ) at 222 nm of C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) methanol showing that thermal denaturation of the enzyme is reversible. Grey lines superimposed for visualisation purposes only.

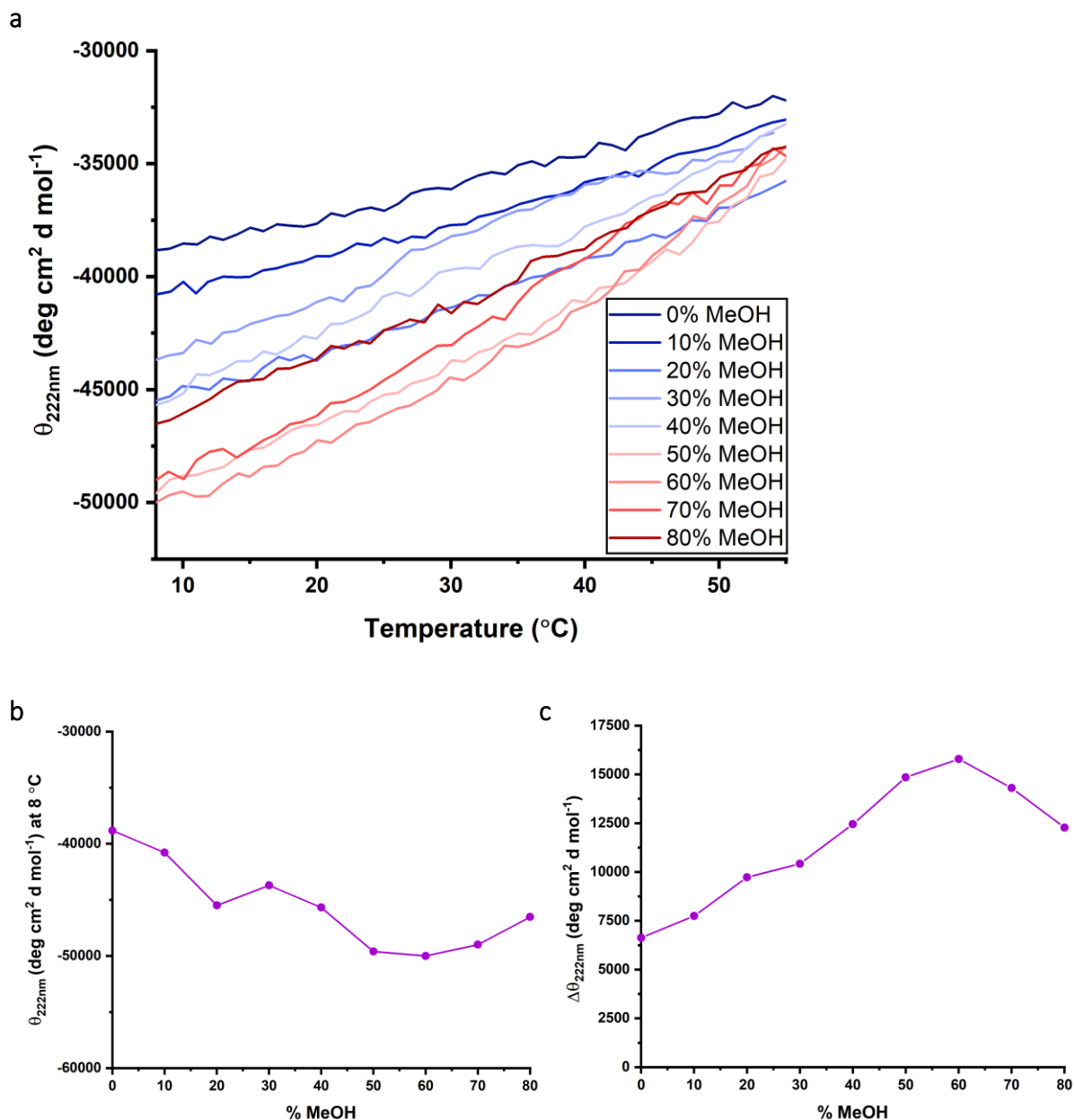


Figure 3.17: Thermal denaturation curves of C45 in aqueous solution with methanol.

a – Overlay of thermal denaturation curves (monitored at 222 nm) in all methanol concentrations tested. b – Ellipticity at 8 °C showing a decrease with increasing methanol concentration. c – Difference in ellipticity 222 nm of C45 in increasing methanol concentrations between 8 °C and 55 °C.

Full spectra of C45 in methanol-water mixtures taken at various points throughout the heating and cooling process seem to confirm that the protein only experiences minor loss of secondary structure which is entirely reversible (Figure 3.18).

3. Structural characterisation of C45 in organic cosolvents

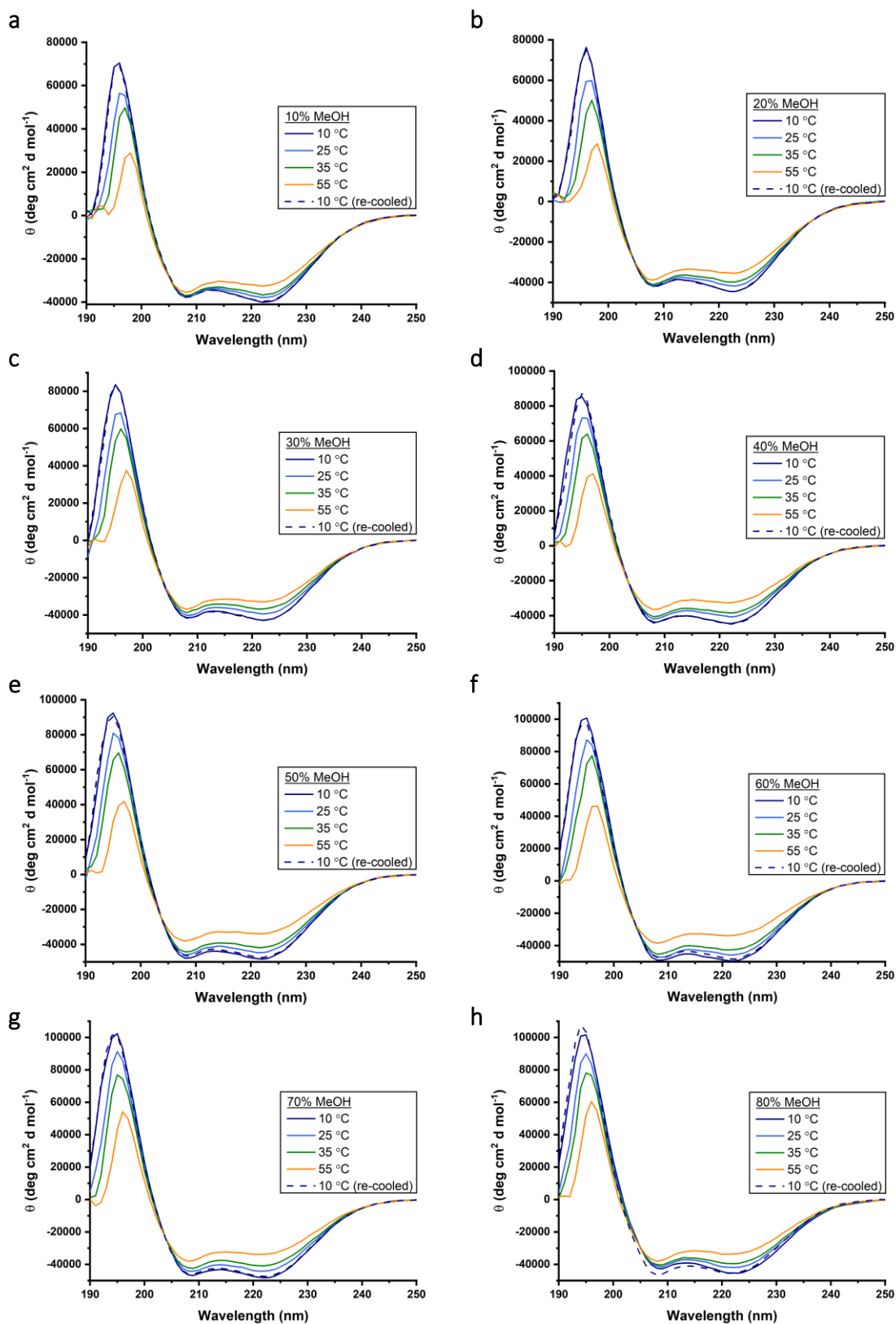


Figure 3.18: CD spectra of C45 in aqueous solution with methanol.

Full CD spectra at various temperatures during the heating and re-cooling of C45 in aqueous solution containing 10% (a) and 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) methanol showing only small loss of helicity.

3.2.5.4. Thermal denaturation of C45 in Ethanol

The same methods were used to investigate the enzyme's secondary structure in ethanol-water mixtures, with the slight exception that temperature was raised to 65 °C (**Figure 3.19** and **Figure 3.20 a**). Up to this temperature, the enzyme does not show complete thermal denaturation in up to 70% ethanol. The loss in helicity that is observed is entirely reversible upon cooling.

Similar to the results in methanol, the negative band at 222 nm appears to get stronger with increasing ethanol concentration at 8 °C (**Figure 3.20 b**). Like in methanol, this is likely not exclusively owed to increasing helicity, but also influenced by the volume of ethanol-water mixtures being smaller than the sums of the two volumes mixed (274). Furthermore, thermal denaturation curves become steeper as cosolvent concentration increases, suggesting a lowering of thermostability (**Figure 3.20 c**). It is also notable that the cooperative phase of unfolding begins at lower temperatures as ethanol concentration is raised, until the entirety of unfolding proceeds more cooperatively at 60% and 70% ethanol.

The exception to this trend is the scan in 80% ethanol, which is approximately -5,000 deg cm² dmol⁻¹ at 8 °C and exhibits decreasing molar residue ellipticity as temperature is raised. This experiment was repeated, with an increased number of heating and cooling cycles (**Figure 3.22 a** and **b**). In another repeat, the experiment was performed with a controlled, slow cooling before commencement of the experiment as well as a rate of heating of 0.5 °C per minute (**Figure 3.22 c** and **d**). The trend of decreasing ellipticity with ascending temperature is consistent. Akin to temperature wavelength scans in 70% acetonitrile, this was attributed to protein precipitation which was reversed during heating.

3. Structural characterisation of C45 in organic cosolvents

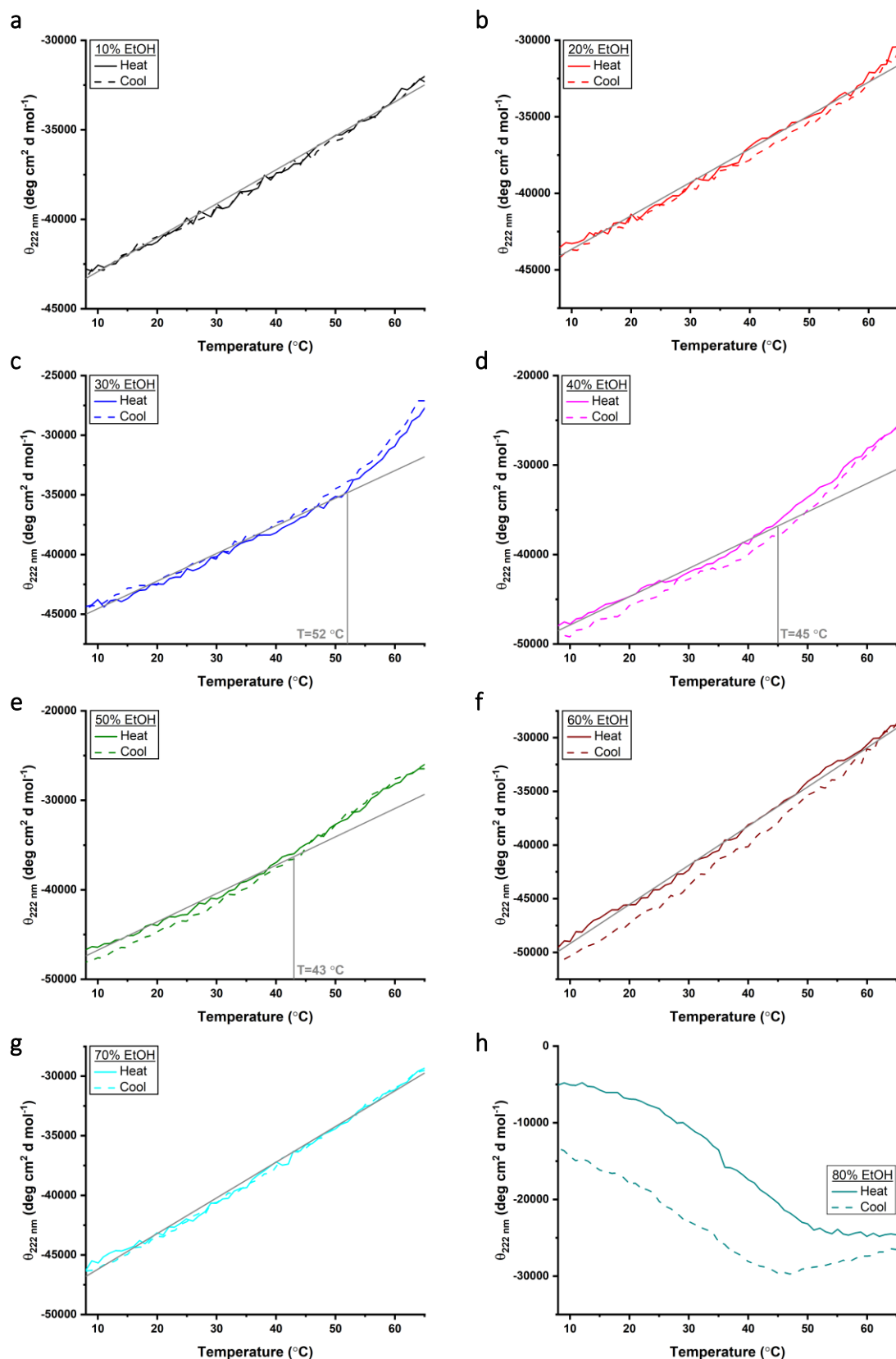


Figure 3.19: CD scan of C45 in aqueous solution with ethanol during heating and cooling.

Mean residue molar ellipticity (Θ) at 222 nm of C45 in aqueous solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) ethanol indicating that in solutions of up to 70% ethanol content, the thermal denaturation of C45 is reversible. Grey lines superimposed for visualisation only.

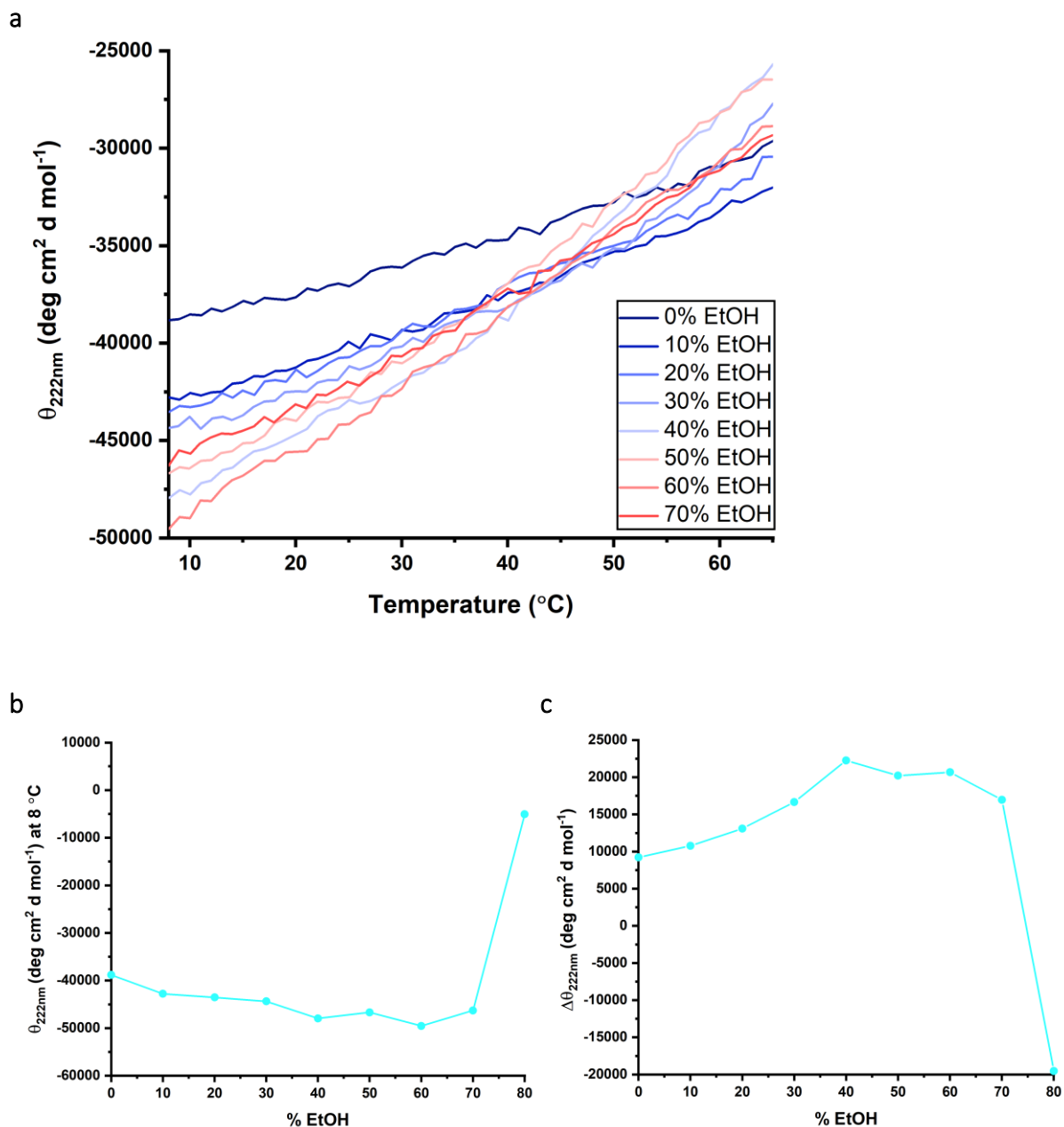


Figure 3.20: Thermal denaturation curves of C45 in aqueous solution with ethanol.

a – Overlay of thermal denaturation curves (monitored at 222 nm) in all ethanol concentrations tested. **b**– Ellipticity at 8 °C showing a decrease with increasing ethanol concentration. **c**–Difference in molar residue ellipticity at 222 nm of C45 in increasing ethanol concentrations between 8 °C and 65 °C.

Full CD spectra of C45 in increasing ethanol concentration (**Figure 3.21**), like the continuous scans at 222 nm, also suggest that in up to 70% ethanol, the enzyme only loses some helicity – less than in acetonitrile, but more than in methanol – and completely refolds.

3. Structural characterisation of C45 in organic cosolvents

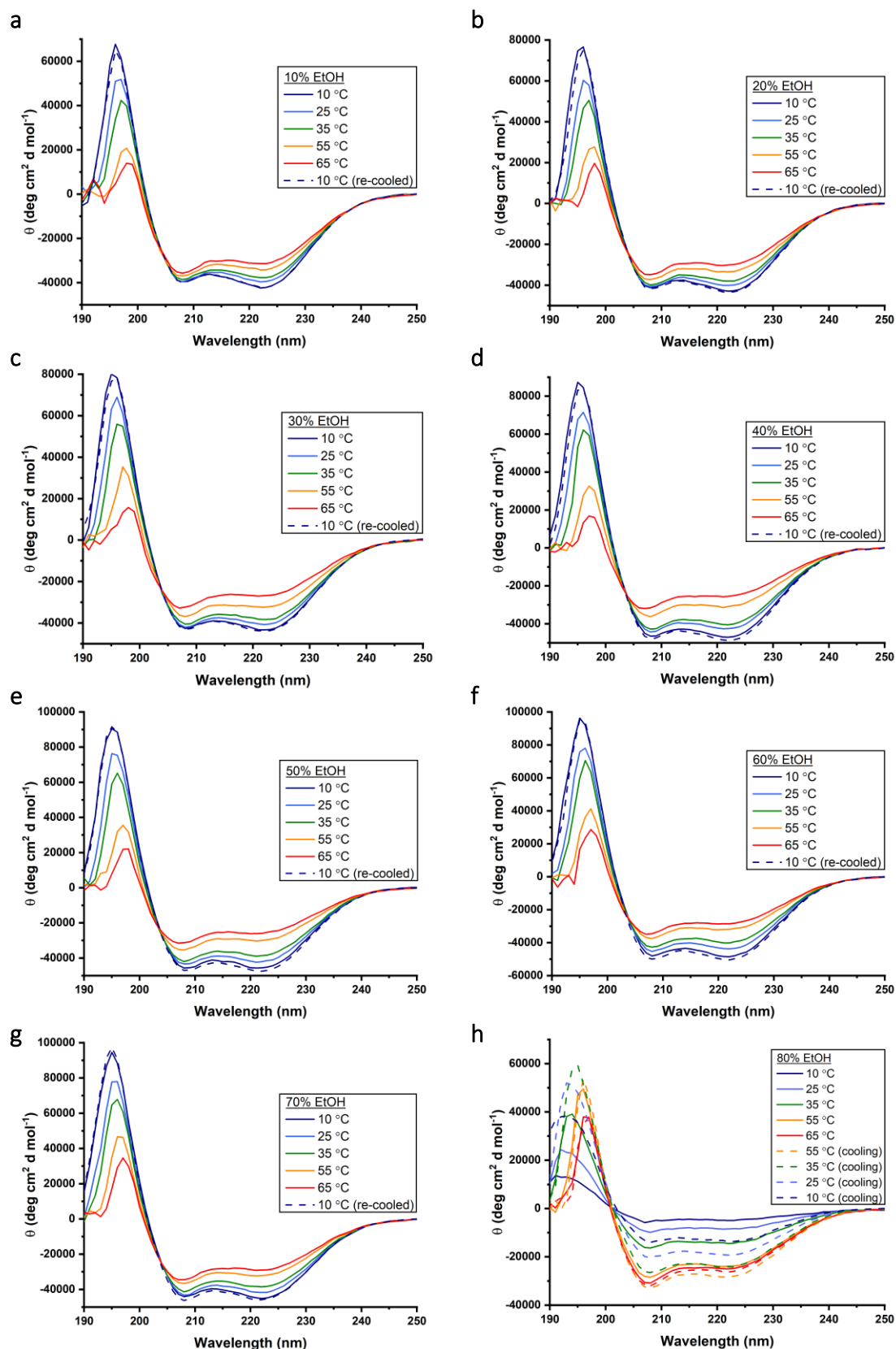


Figure 3.21: CD spectra of C45 in aqueous solution with ethanol.

Full CD spectra at various temperatures during the heating and re-cooling of C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) ethanol showing that C45 retains its helical fold in solutions of up to 70% ethanol.

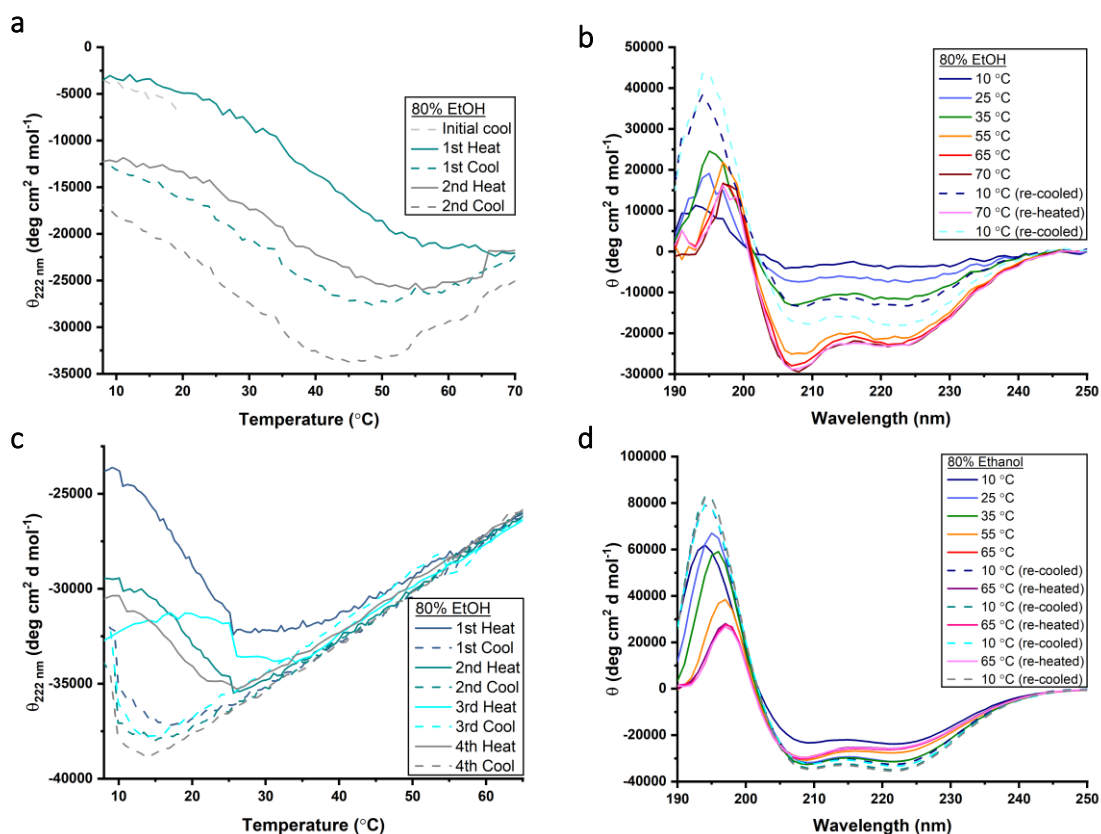


Figure 3.22: CD measurements of C45 in buffering solution with 80% ethanol.

a and c show the MRE at 222 nm during multiple cycles of heating and cooling at a rate of 0.5 $^{\circ}\text{C}$ per minute after initially being cooled from room temperature to 8 $^{\circ}\text{C}$ at the same rate. b and d show full CD spectra taken during the first heating phase and when re-cooled and re-heated.

3.2.5.5. Thermal denaturation of C45 in Isopropanol

To investigate the effects of a slightly bulkier alcohol on the folding and thermostability of C45, the enzyme was analysed in different concentrations of isopropanol using CD temperature scans up to 70 $^{\circ}\text{C}$ (Figure 3.23 and Figure 3.24 a). Figure 3.25 shows full CD spectra at various temperatures. In up to 60% isopropanol, scans suggest C45 possesses its native secondary structure at 8 $^{\circ}\text{C}$, with molar residue ellipticity values at 222 nm ranging from -37,000 to -52,000 $\text{deg cm}^2 \text{ d mol}^{-1}$ (Figure 3.24 b). As in methanol and ethanol, they decrease with increasing isopropanol concentration (up to 60%), likely for the same mixture of increasing helicity and volumetric change. In these isopropanol concentrations, some helicity is lost with

rising temperature – increasingly so at increasing isopropanol concentrations (**Figure 3.24 c**) – however not sufficiently so that a melting temperature could be calculated. As seen in other cosolvents, the cooperative phase of unfolding is initiated at lower temperatures with increasing isopropanol concentrations, until no change in cooperativity is observed anymore and the entire process appears to occur much more cooperatively than in buffer. Furthermore, C45 also refolds entirely upon cooling in all these isopropanol concentrations.

In 70% isopropanol, the scan showed similarities with those in 70% acetonitrile and 80% ethanol: initially, the molar residue ellipticity is significantly higher (approximately $-7,000 \text{ deg cm}^2 \text{ dmol}^{-1}$). This observation was attributed to precipitation and resolution, as in 70% acetonitrile and 80% ethanol. The experiment was repeated with a less steep temperature elevation ($0.5 \text{ }^\circ\text{C}$ per minute) and just like in 80% ethanol, this improved resolution (**Figure 3.26**).

In 80% isopropanol, the secondary structure appears to be denatured, however this is independent of temperature with a positive molar residue ellipticity at 222 nm at the initial temperature of $8 \text{ }^\circ\text{C}$.

3.2. Results - CD Thermal denaturation studies

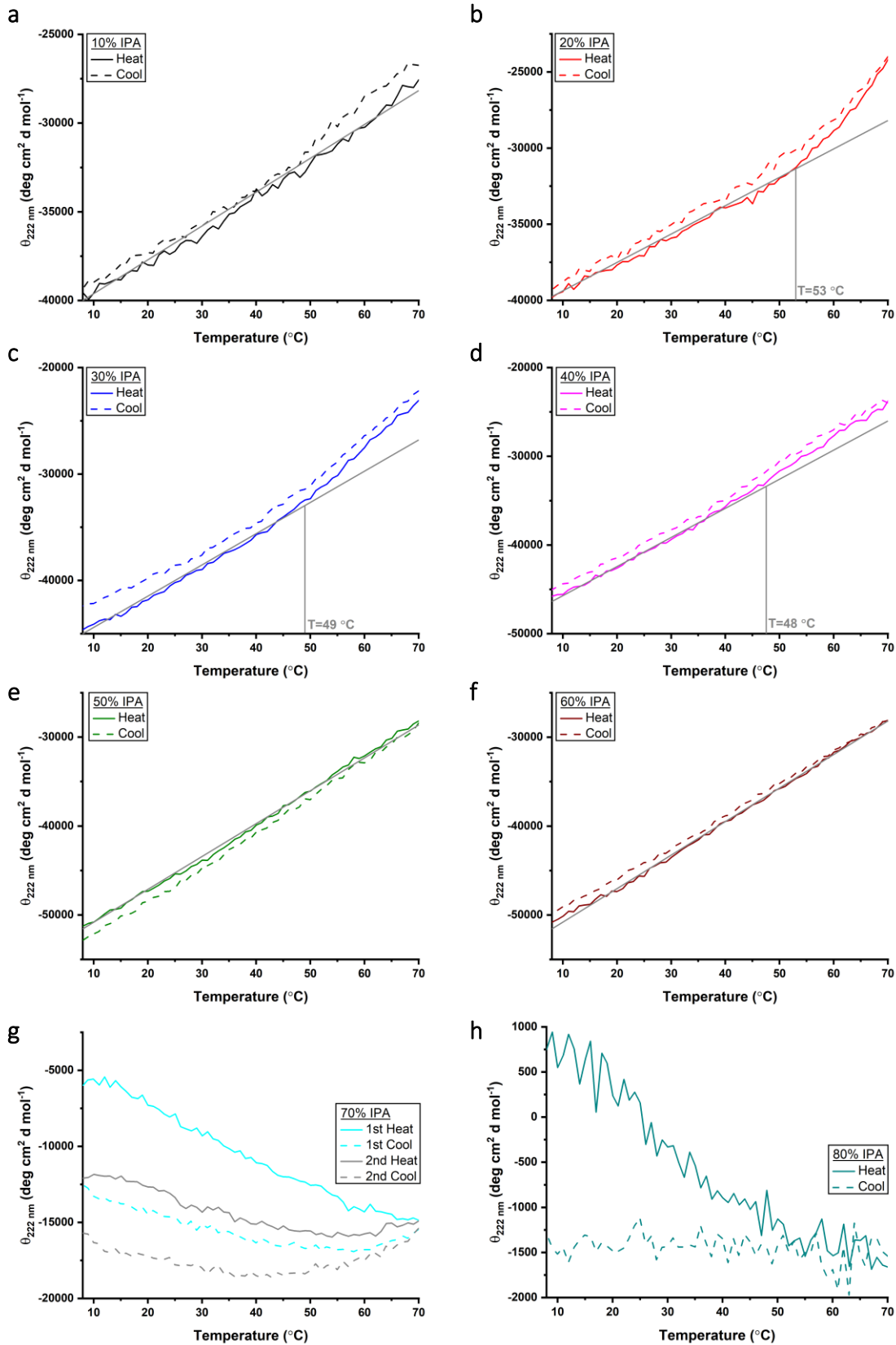


Figure 3.23: CD scan of C45 in aqueous solution with isopropanol during heating and cooling.

Mean residue molar ellipticity (θ) at 222 nm of C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) isopropanol conveying that in up to 60% isopropanol, thermal denaturation of C45 is entirely reversible. Grey lines superimposed for visualisation purposes only.

3. Structural characterisation of C45 in organic cosolvents

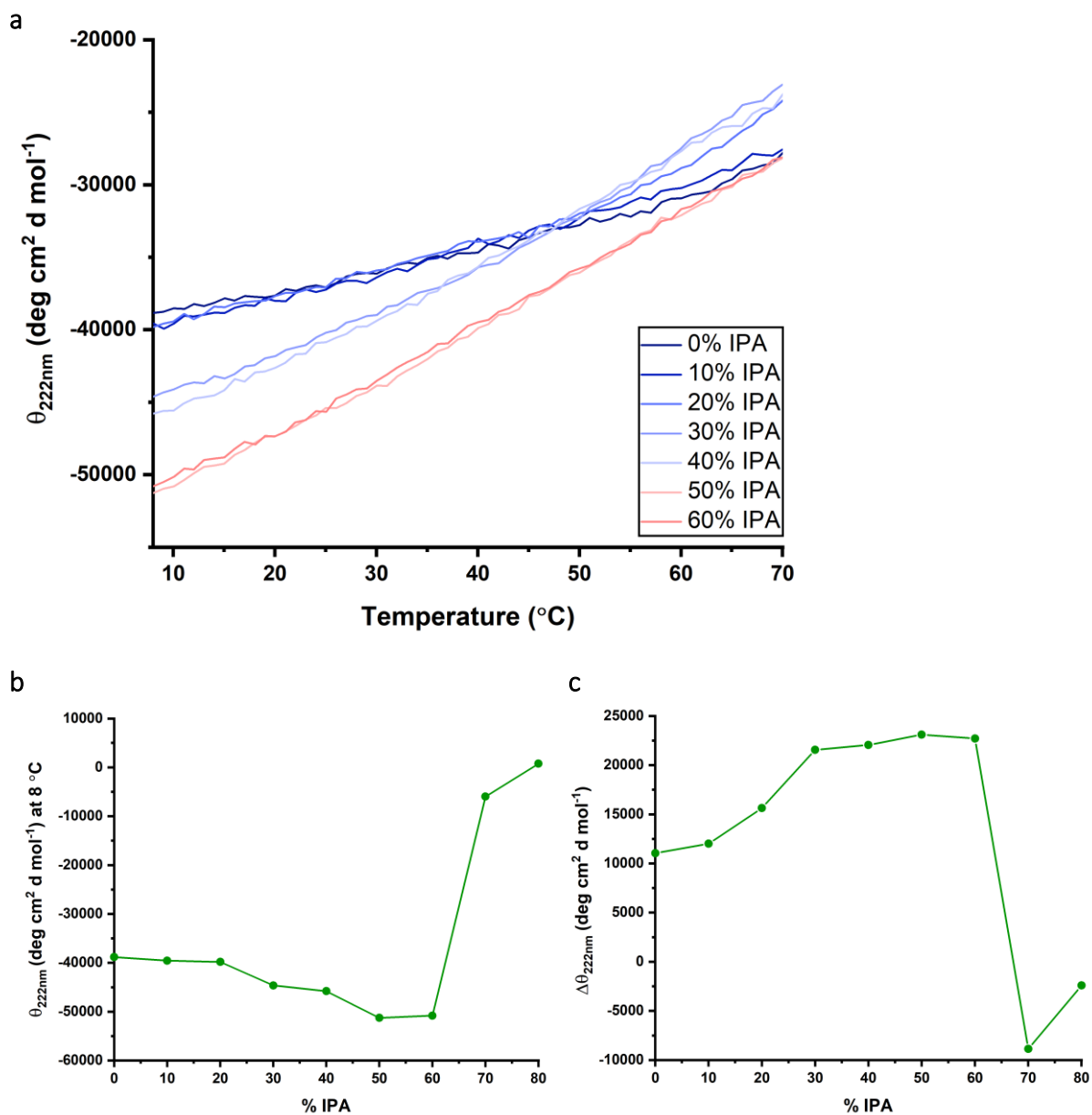


Figure 3.24: Thermal denaturation curves of C45 in aqueous solution with isopropanol.

a – Overlay of thermal denaturation curves (monitored at 222 nm) in all isopropanol concentrations tested. **b**– Molar residue ellipticity at 8 °C showing a decrease with increasing isopropanol concentration. **c**– Difference in MRE of C45 at 222 nm in increasing isopropanol concentrations between 8 °C and 70 °C.

3.2. Results - CD Thermal denaturation studies

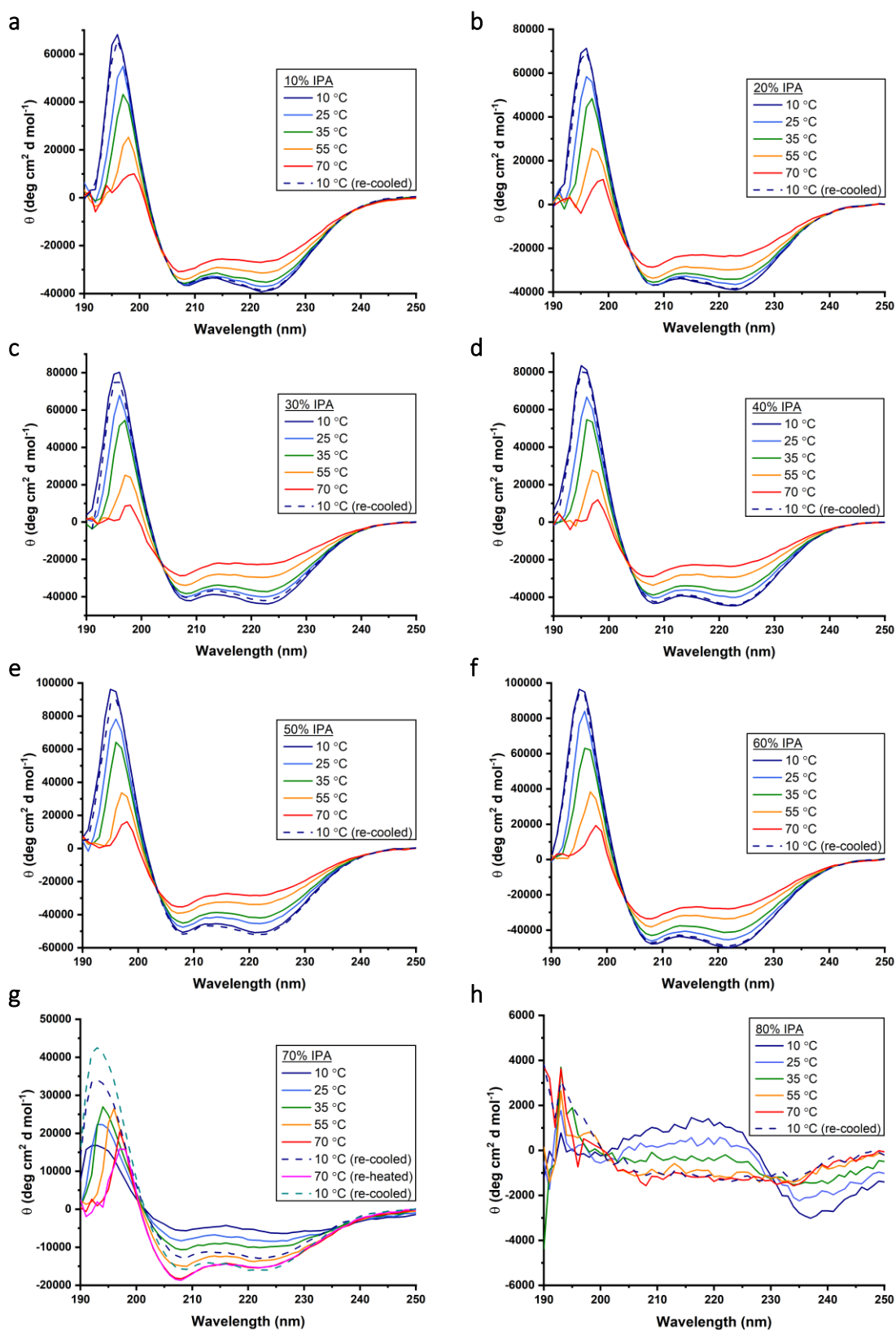


Figure 3.25: CD spectra of C45 in aqueous solution with isopropanol.

Full CD spectra at various temperatures during the heating and re-cooling of C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) isopropanol showing that the enzyme retains its helical fold and thermostability in up to 60% isopropanol.

3. Structural characterisation of C45 in organic cosolvents

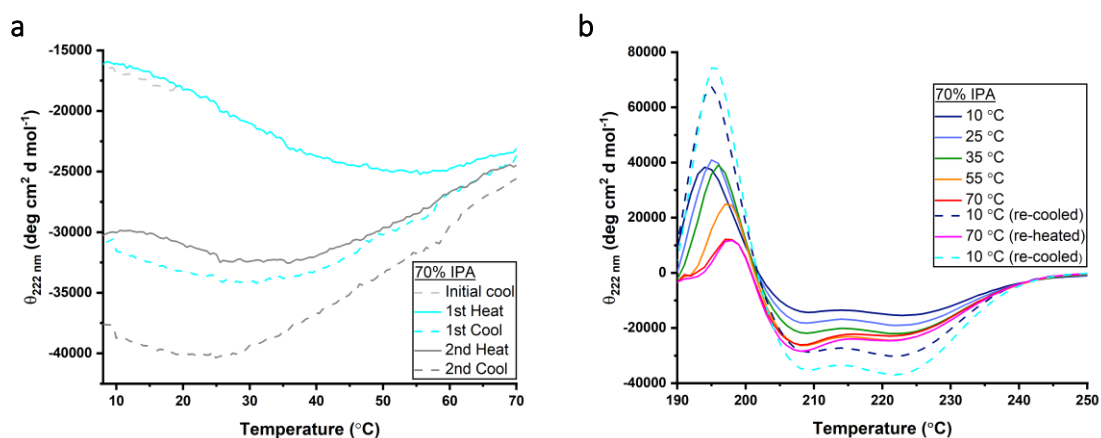


Figure 3.26: CD measurements of C45 in buffering solution with 70% isopropanol.

a – Molar residue ellipticity at 222 nm during multiple cycles of heating and cooling at a rate of 0.5 $^{\circ}\text{C}$ per minute after initially being cooled from room temperature to 8 $^{\circ}\text{C}$ at the same rate. **b** - Full CD spectra taken during the first heating phase and when re-cooled and re-heated.

3.2.5.6. Thermal denaturation of C45 in TFE

CD was also used to investigate the thermostability and folding of C45 in solutions of 10-80% TFE (**Figure 3.27**). Previous work on helical proteins has shown that TFE stabilises helices (4). The secondary structure of C45 was therefore expected to clearly exhibit a helical secondary structure in temperatures up to at least 55 $^{\circ}\text{C}$, similarly to aqueous solution. This was experimentally confirmed: up to a temperature of 70 $^{\circ}\text{C}$, C45 retained its secondary structure in solutions with up to 80% TFE. In all investigated TFE concentrations, the enzyme loses helicity with increasing temperatures, however no indication of different phases with distinct levels of cooperativity can be observed in the denaturation process, nor has the protein been observed to fully denature in temperatures up to 65 $^{\circ}\text{C}$. As in other cosolvents, any loss of secondary structure is entirely reversible in all TFE concentrations tested.

As predicted by the literature, the helicity of C45 increases with TFE concentrations, especially at the lowest temperature of 8 $^{\circ}\text{C}$ (**Figure 3.28 b**). Even at 65 $^{\circ}\text{C}$, the ellipticity is lowest for 80% TFE concentration, highlighting the helix

stabilising properties of the cosolvent. The signal increase with TFE addition is larger at lower temperatures, though, suggesting unfolding proceeds more cooperatively at higher TFE concentrations. This is in agreement with the literature, which states that the helix stabilising properties are known to be particularly observable at low temperatures (3).

It has been suggested that the ellipticity at the minimum at 222 nm should be approximately $-44,000 \text{ deg cm}^2 \text{ d mol}^{-1}$ for an infinite helix at $0 \text{ }^\circ\text{C}$ (148). At TFE concentrations above 30%, however, the experimentally measured minimum at 222 nm was lower for C45 at the starting temperature of $8 \text{ }^\circ\text{C}$ (**Figure 3.28 b**). Luo *et al.*, who proposed the aforementioned limit of $-44,000 \text{ deg cm}^2 \text{ d mol}^{-1}$, saw no further increase in protein helicity at TFE concentrations above 25%, which is very different from the results obtained for C45. Furthermore, their research did not involve work with TFE concentrations above 50%. It is therefore unclear how higher TFE concentrations affect the minimal value for ellipticity at 222 nm for a fully helical protein.

Nonetheless, the discrepancy between this value and the measurements for C45 in TFE may indicate significant experimental error, potentially due to an underestimation of protein concentration. This was calculated from absorbance in the heme region, signifying that any apoprotein would not be taken into account, resulting in an underestimation of protein concentration and thus an overestimation of the concentration-dependent ellipticity.

Spectra of C45 in TFE-aqueous mixtures obtained at increasing temperatures (**Figure 3.29**) confirm that the protein loses some helicity as temperature is elevated but refolds completely upon cooling. The protein appears to be helical in all TFE concentrations.

3. Structural characterisation of C45 in organic cosolvents

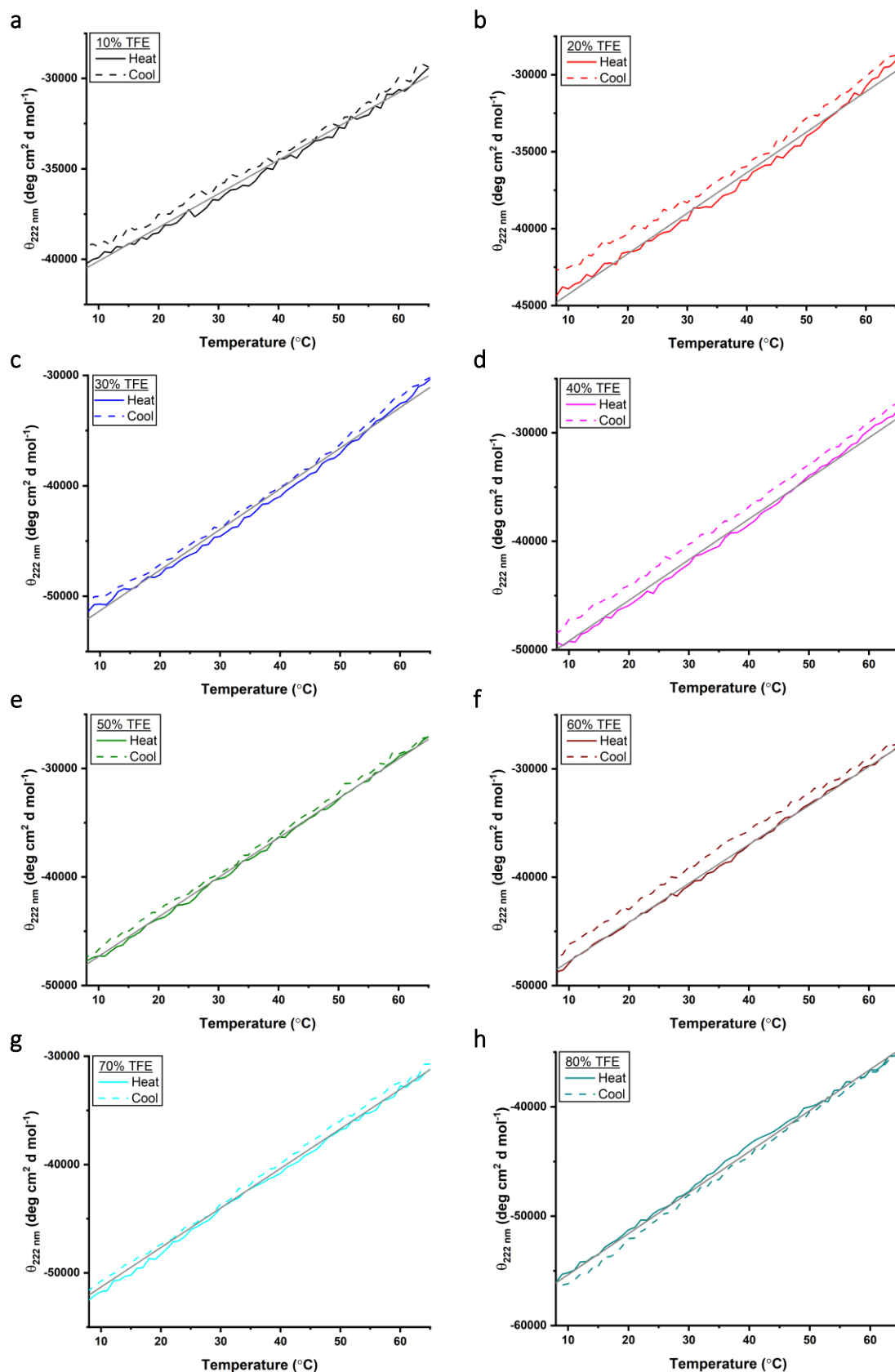


Figure 3.27: CD scan of C45 in aqueous solution with TFE during heating and cooling.

Mean residue molar ellipticity (Θ) at 222 nm during heating and cooling of C45 in aqueous solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) TFE showing that thermally denatured C45 in TFE refolds upon cooling. Grey lines superimposed for visualisation only.

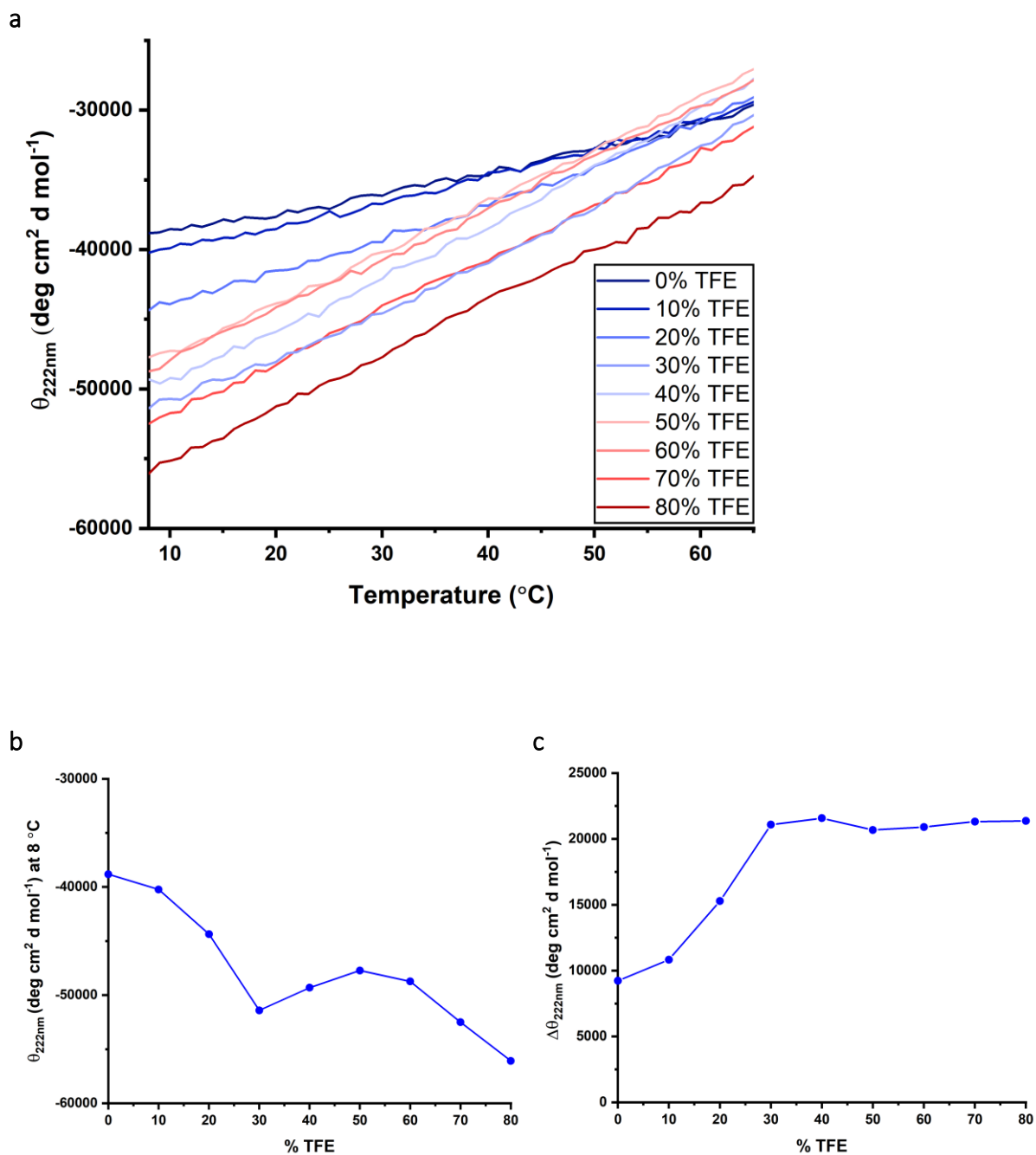


Figure 3.28: Thermal denaturation curves of C45 in aqueous solution with TFE.

a – Overlay of thermal denaturation curves (monitored at 222 nm) in all TFE concentrations tested. **b**– Ellipticity at 8 °C showing a decrease with increasing TFE concentration. **c**– Difference in ellipticity 222 nm of C45 in increasing TFE concentrations between 8 °C and 65 °C.

3. Structural characterisation of C45 in organic cosolvents

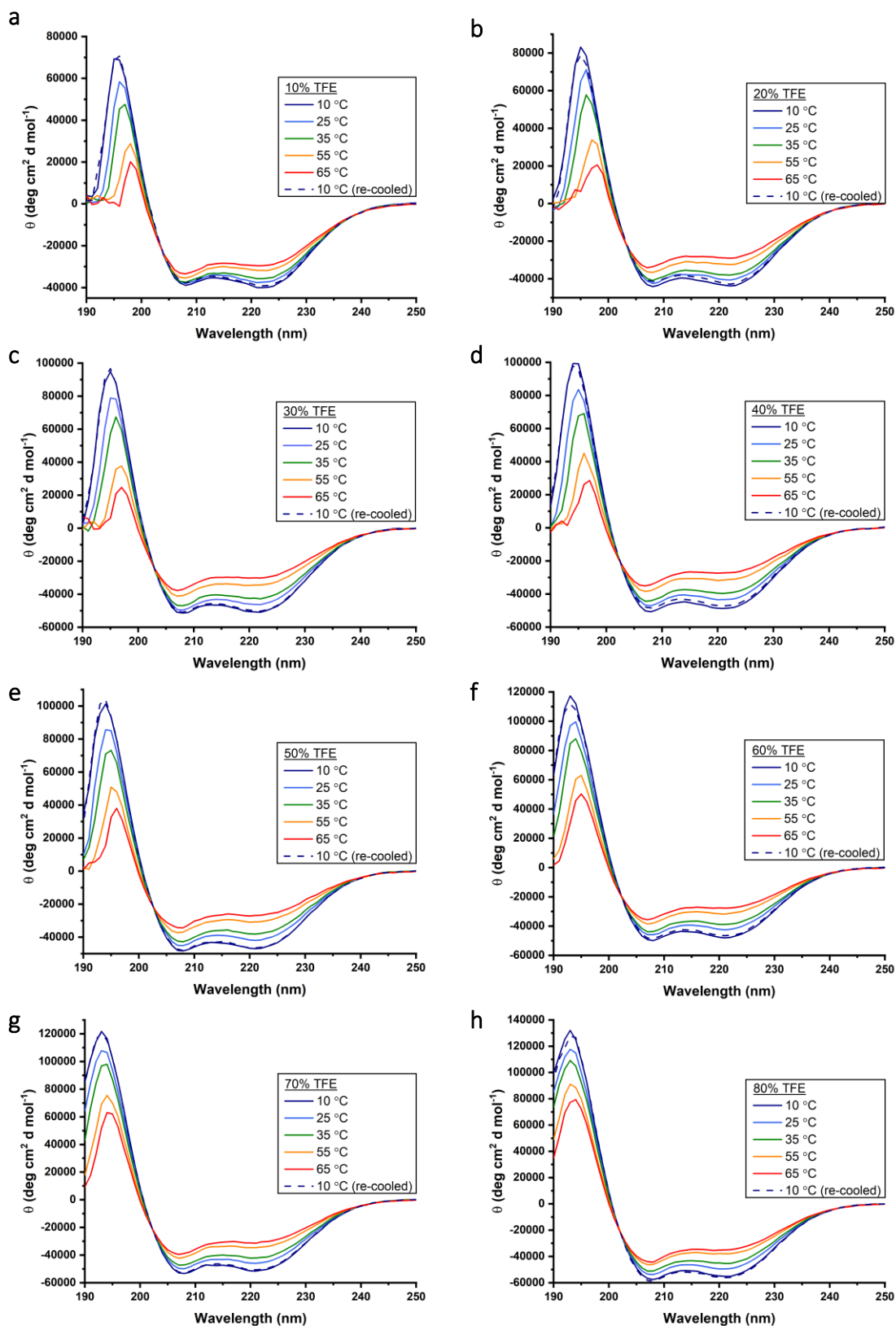


Figure 3.29: CD spectra of C45 in aqueous solution with TFE.

Full CD spectra at various temperatures during the heating and re-cooling of C45 in buffering solution containing 10% (a), 20% (b), 30% (c), 40% (d), 50% (e), 60% (f), 70% (g) and 80% (h) TFE indicating good thermostability of the protein.

3.3. Discussion

In this chapter the structural properties of C45 and the effects of different cosolvents on these were investigated. It has been shown that the C45 purification protocol is suitable to remove the vast majority of intracellular proteins, apoprotein, aggregates and other impurities to yield sufficiently pure C45. The protein elutes from size exclusion chromatography in two peaks on size exclusion chromatography, which suggests a degree of dimerisation. While it was confirmed with SDS-PAGE that both peaks correspond to C45, the technique is not capable of distinguishing between monomers and dimers. Techniques such as dynamic light scattering, small angle x-ray scattering or native PAGE would be better suited for this.

UV-visible spectroscopy of the purified protein clearly indicated the presence of heme, and a room temperature CD spectrum showed that the enzyme is strongly α -helical. This conveys that after the expression and purification, the protein is correctly folded and binding the catalytic heme cofactor, which is mono-histidine ligated. Furthermore, variable temperature CD indicated that the protein only partially unfolds at temperatures of up to 85 °C, implying good thermostability in agreement with the literature (2). This proved to be entirely reversible. The thermal denaturation curve showed that at temperatures above 65 °C, unfolding proceeds more cooperatively, which is very common in helical bundle proteins (277–279). However, this is neither pronounced strongly enough, nor can the protein be thermally unfolded completely in temperatures up to 85 °C, which would be necessary to reliably quantify this behaviour. Chemical unfolding using, for instance, guanidine hydrochloride may provide an alternative route to investigating the unfolding process of C45.

3. Structural characterisation of C45 in organic cosolvents

Further, analytical data are presented for C45 in varying concentrations of acetonitrile, methanol, ethanol, isopropanol and TFE. In concentrations up to 60% acetonitrile and isopropanol, up to 70% ethanol, and up to the highest tested concentration of 80% methanol and TFE, the protein appears to retain its monohistidine ligated heme as a cofactor as well as its secondary structure at room temperature. In higher concentrations of acetonitrile, ethanol and isopropanol, the protein appears to have lost its secondary structure and be entirely denatured. Whether there is a correlation between solvent polarity and cosolvent concentration, at which the protein denatures would require further testing with a larger range of cosolvents, however the data presented herein suggest this may be the case.

The loss of structural stability upon an increase in temperature as well as the presence of high concentrations of some cosolvents (70% isopropanol and 80% ethanol) can also be seen in the decreasing ratio of ellipticity at 222 nm and 208 nm. In all solvents, a clear decrease in this ratio can be seen, suggesting a less tightly packed and overall less helical protein upon application of heat.

Interestingly though, with the addition of the alcoholic cosolvents in lower concentrations than those causing denaturation, the degree of helicity appears to increase (**Figure 3.30**), especially with TFE. Alcohols as cosolvents – and TFE in particular - had previously been shown to stabilise helices (3). However, this effect may be exaggerated in the data presented here as volumetric reduction caused by the addition of alcohols to water may lead to underestimation of protein concentration values. Furthermore, the protein concentration was determined from absorbance at 406 nm, which is dependent on the successful incorporation of heme into the protein scaffold. The presence of any apoprotein, however, would not be mathematically taken into account, but would increase the CD signal, which is directly related to protein

concentration. It is, however, worth noting that the same quantity of C45 from the same expression batch was added to each of the experiments, so an underestimation of protein concentration would lead to an overestimation of ellipticity at 208 and 222 nm in all experiments and not exclusively those performed in high solvent concentrations.

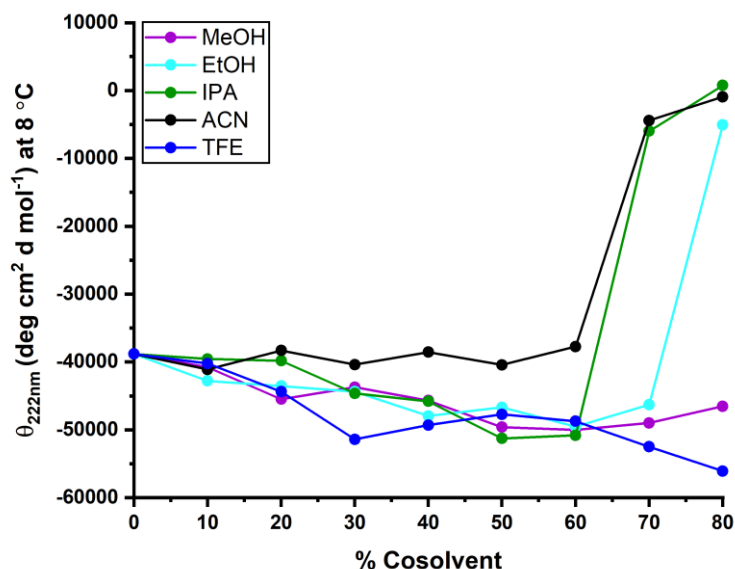


Figure 3.30: Molar residue ellipticity at 222 nm at 8 °C for C45 in various cosolvents.

TFE increases the helical propensity of amino acids. Nonetheless, glycine, which is a key component of the loops between the helices in C45, still acts as a so-called helix breaker (3). Therefore, it is less likely that helicity increases due to helical propagation through the loops than there being an equilibrium between the helical and the random coil states of the protein which is shifted towards the helix with the addition of TFE and other alcohols.

Furthermore, the distinction between secondary and tertiary structure is crucial. While CD can provide insights into the helicity of C45, it can only provide very limited insights into the tertiary structure. TFE has, for instance, been demonstrated to propagate helicity in proteins even at the cost of tertiary structure (141). It is possible,

3. Structural characterisation of C45 in organic cosolvents

that with the addition of cosolvents, the protein retains its helical structure or even increases in helicity, but is otherwise unfolded (**Figure 3.31**). Therefore, no conclusion can be drawn regarding the functionality of the enzyme from the helicity.

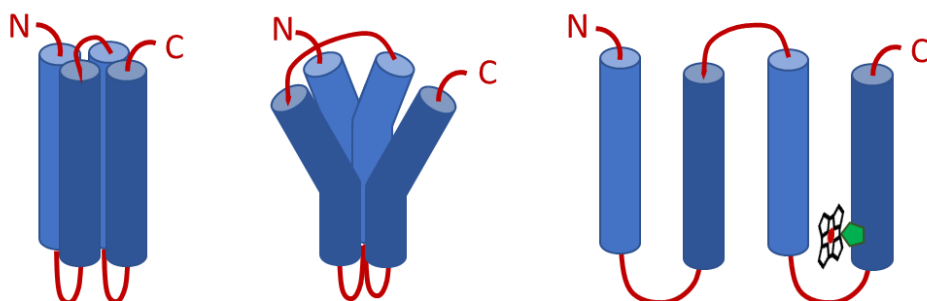


Figure 3.31: Models for a four helix bundle protein with the same levels of helicity, but different tertiary structure.

Variable temperature CD studies of C45 in cosolvents showed that C45 lost helicity with increasing temperature, however, not to the extent that the protein denatures completely. Therefore, very little can be concluded about the protein's unfolding profile. Repeat experiments with the addition of a chemical denaturant or titrations with guanidine hydrochloride may be pertinent to obtain an unfolding profile. Nonetheless, some insights can be gained from these results. Firstly, it is clear that unfolding is entirely reversible in all cosolvents. Secondly, the thermostability of C45 does not seem to diminish with the addition of alcohols as a cosolvent. Comparison of molar residue ellipticity values at 55 °C shows that in most alcoholic cosolvents (except for 70-80% isopropanol and 80% ethanol), C45 does not exhibit significantly less helicity than in buffer. In acetonitrile, however, ellipticity at this temperature rises with increasing cosolvent concentration. This suggests that out of the five cosolvents tested, acetonitrile is the only one that significantly decreases thermostability of the enzyme (**Figure 3.32**).

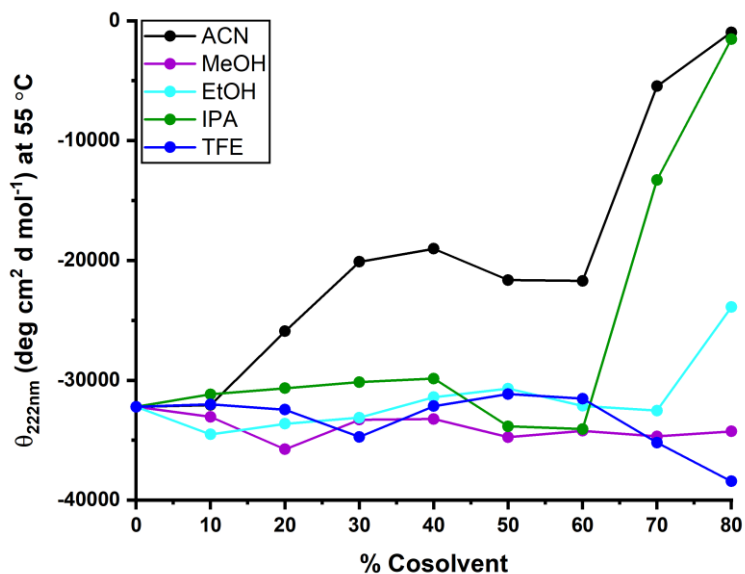


Figure 3.32: Molar residue ellipticity at 222 nm of C45 in different cosolvents at 55 °C

In 70-80% acetonitrile and isopropanol as well as 80% ethanol, the protein appears entirely denatured and the solutions appeared slightly turbid. Some of the CD spectra suggest that an increased degree of helicity is achieved by applying heat to the sample, however, this is more likely owed to resolution of precipitated protein. The fact that without the application of heat, the UV-visible spectrum also changed with the amount of scattering significantly reducing (indicating fewer aggregates in the solution) supports this hypothesis.

The main conclusion from the data presented in this chapter is that the protein appears structurally resilient to the presence of up to at least 60% cosolvent. The following chapter will address the impact on peroxidase kinetics that the presence of organic cosolvents has and whether catalytic activity is correlated with the enzyme's secondary structure.

4. Functional characterisation of C45 in organic cosolvents

In the previous chapter, the effects of acetonitrile, methanol, ethanol, isopropanol and TFE on the protein's secondary structure were explored. Generally, the addition of alcoholic cosolvents, especially TFE, appeared to cause an increase in helix stability in C45 whereas acetonitrile had the opposite effect. In high (70-80%) concentrations of acetonitrile and isopropanol as well as 80% ethanol, the protein appeared to have lost its structure. In lower cosolvent concentrations, however, C45 remained predominantly helical, although this does not necessarily mean that the tertiary structure remained unaffected. It will be interesting to observe whether the catalytic activity of the enzyme is similarly well retained in cosolvent and whether the stabilising alcohols will have a positive effect.

4.1. Introduction

The *c*-type heme cofactor of C45 gives rise to its peroxidase activity meaning it catalyses the oxidation of substrates using peroxide as the oxidising agent *via* the hypothesised mechanism shown in **Figure 4.1**.

In previous work with C45, the optimal pH for this reaction was determined to be 8.6, as shown in **Figure 4.2** (2). Subsequently, the kinetics of this reaction have been extensively tested in buffer at room temperature and above. The highest catalytic efficiencies were measured at high temperatures (70 °C), which is in good agreement with the high thermostability of C45 shown in chapter 3 of this work.

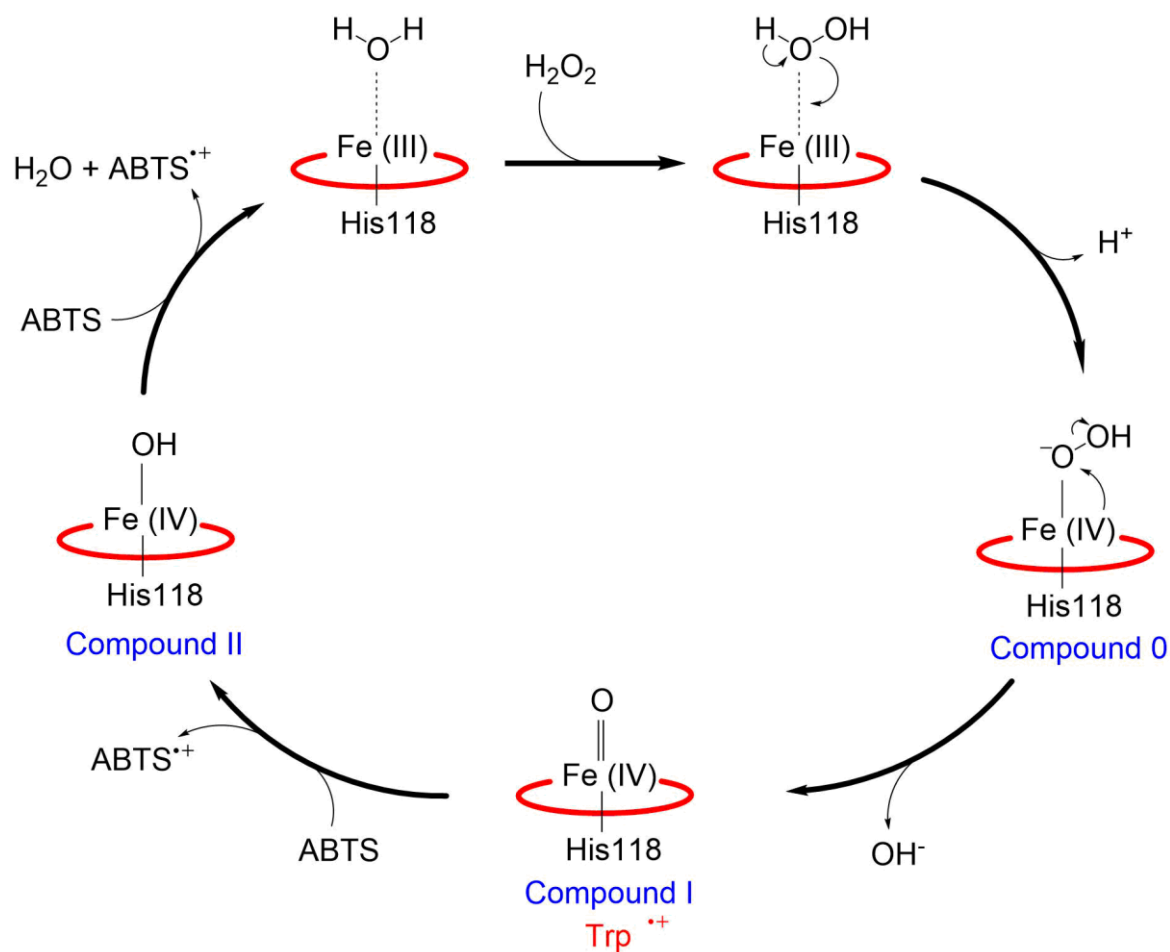


Figure 4.1: The hypothesised catalytic cycle of C45 mediated oxidation of ABTS.

The intermediate compound I is particularly reactive as it is a cationised radical. Preliminary electron paramagnetic resonance spectroscopy data have confirmed the presence of a radical tryptophan residue in Compound I, most likely W43, which is located very close to the heme (see **Figure 1.17**). The structure of ABTS is shown in **Figure 4.4**.

Analysis of two substrate reactions is often performed with one substrate at a constant concentration. Typically, saturating concentration is chosen to achieve pseudo-first order kinetics. In the case of C45, however, which not only has a very high K_M for peroxide but is also decomposed by the substrate, subsaturating peroxide concentration was chosen. Under these conditions, the reaction proceeds *via* pseudo-Michaelis-Menten kinetics.

4. Functional characterisation of C45 in organic cosolvents

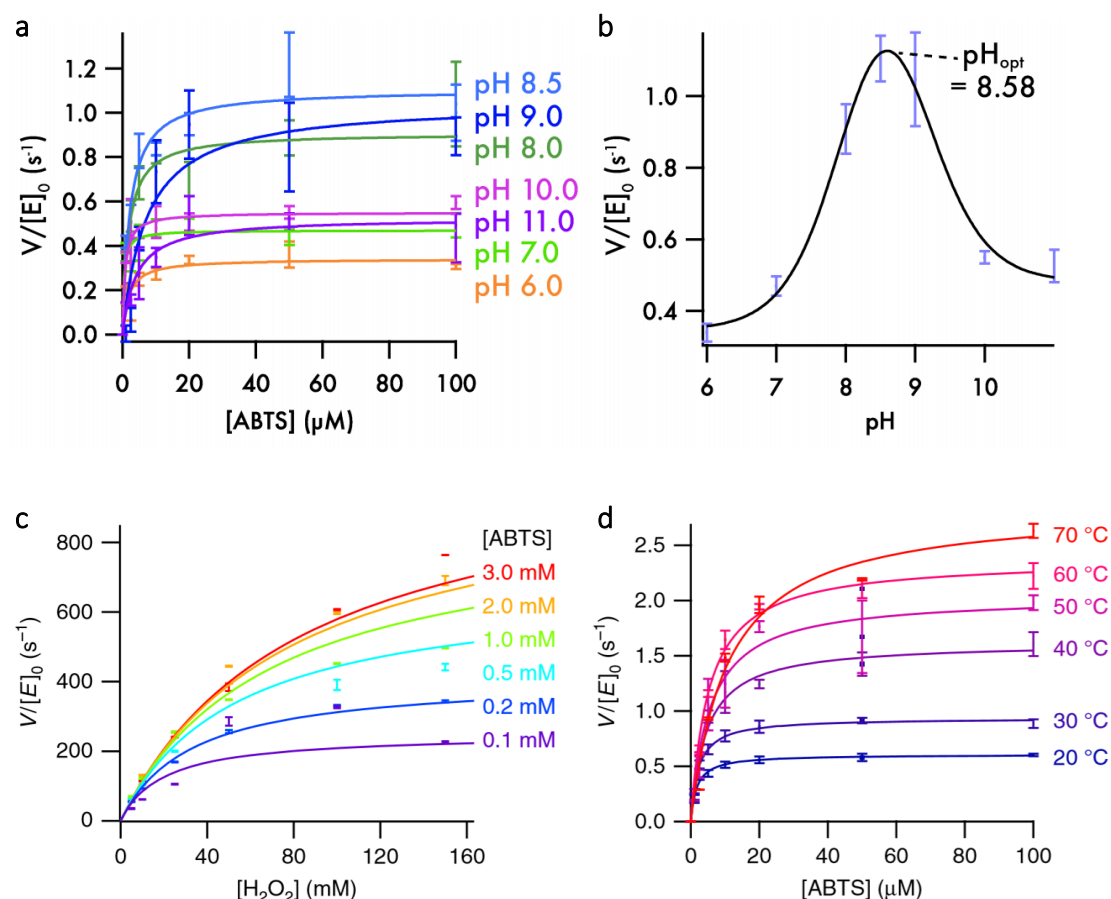


Figure 4.2: Kinetics of the oxidation of ABTS by C45 and H₂O₂ at different pH values.

a – Kinetics at 25 °C of the reaction of 250 nM C45, 100 μM H₂O₂ and 0-100 μM ABTS in buffer (100 mM KCl and 20 mM potassium phosphate, CHES or bis-tris for pH 6.0-8.0, 8.5-9.0 or 10.0-11.0 respectively) indicating that the activity is highest at a pH of 8.5. **b** – $v/[E]_0$ of the reactions at limiting peroxide and saturating ABTS concentrations (100 μM H₂O₂ and 100 μM ABTS) at different pH values fitted to a bell-shaped model, showing an optimum pH of this reaction of approximately 8.6 (2). **c** – Kinetic analysis of 100 nM C45 with varying concentrations of ABTS and H₂O₂ at pH 8.6, fitted to a ping-pong steady-state kinetic model. **d** – Kinetics of 100 nM C45 with 100 μM H₂O₂ and varying ABTS concentrations at pH 8.6 and increasing temperature, showing that activity increases with temperature.

Figure reproduced from (2).

In the Michaelis-Menten model (154), the formation of product from the enzyme and substrate is viewed as a two-step process, in which the enzyme (E) first forms a complex (ES) with the substrate (S), that subsequently undergoes the reaction to product (P) that is released from the enzyme (**Figure 4.3**). The second step is assumed to be irreversible as measurements are performed under initial velocities in

the absence of product. Furthermore, the reaction was performed under steady-state conditions, meaning that the concentration of the enzyme-substrate complex remains constant.



Figure 4.3: The Michaelis-Menten model of enzyme kinetics.

The diagram highlights the rate constants involved in the two-step reaction between enzyme E and substrate P to first form the enzyme-substrate complex ES and then split into enzyme and product P.

The Michaelis-Menten equation describes the relationship between initial reaction rate v_0 , substrate concentration $[S]$, Michaelis constant K_M and turnover number k_{cat} :

$$\frac{v_0}{[E]_0} = \frac{k_{cat} [S]}{K_M + [S]} \quad 4.1$$

where k_{cat} is equal to the maximum rate v_{max} divided by the initial enzyme concentration $[E]_0$. The turnover number k_{cat} refers to the number of substrate molecules an enzyme converts per unit time. The Michaelis constant K_M is the substrate concentration where the reaction rate is half of its maximum v_{max} . k_{cat}/K_M is referred to as the catalytic efficiency, with the diffusion limit of 10^8 - 10^9 $M^{-1} s^{-1}$ (280) setting the maximum.

ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) is a prominent model substrate for peroxidase assays. It forms an intensely turquoise coloured radical cation (281) upon oxidation (**Figure 4.4**), which allows the monitoring of peroxidase activity spectrophotometrically (20).

4. Functional characterisation of C45 in organic cosolvents

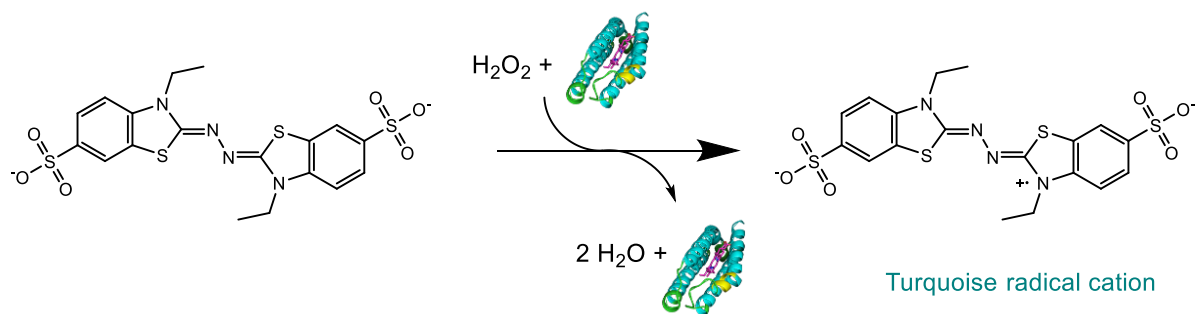


Figure 4.4: The C45-mediated oxidation of ABTS in the presence of hydrogen peroxide.

This can be performed on a UV-visible spectrometer, a stopped flow spectrometer or a plate reader. All three were utilised for this work and exploited for their respective strengths. A UV-visible spectrometer may be the most versatile, providing the options of continuous measurement at a single wavelength or full spectra at specific time intervals with a simple experimental setup and the possibility to conduct assays in small volumes. Absorbance measurements performed on a plate reader are less precise, though this method permits the analysis of many different reactions in parallel, rendering it ideal for obtaining kinetic data rapidly (282). The plate reader employed is equipped with an automated injection system, further facilitating the automation of kinetic assays. Neither setup is adequately equipped for the collection of accurate data of the start of the reaction (<10 seconds). A stopped flow spectrometer was used to measure kinetics at these short timescales.

This chapter seeks to understand the functional impact of the cosolvents used in chapter 3 on C45. To this end, catalytic activity of C45 is analysed in buffer as well as 10-80% acetonitrile, methanol, ethanol, isopropanol, and TFE. For this work, the optimal conditions established in previous work (2) of buffer containing 20 mM CHES and 100 mM KCl at pH 8.6 were used and measurements performed at room temperature.

4.2. Results

4.2.1. Michaelis-Menten Kinetics of C45 with low substrate concentrations

4.2.1.1. Peroxidase Kinetics of C45 in buffer

For the oxidation of ABTS by C45 in the presence of peroxide, the biggest increase in absorbance occurs at 405 nm. Whilst this is in the region of the Soret band, at protein concentrations of up to 100 nM the absorbance by heme should be below 0.015, which is negligible compared to the absorbance of the product. For this wavelength, ABTS^{•+} has a known extinction coefficient ϵ of 36,800 M⁻¹ cm⁻¹ (2, 283).

Similar to previous experiments, all kinetics in this chapter are performed under constant subsaturating H₂O₂ concentration of 100 μ M. The resulting k_{cat} values are thus smaller than the maximum k_{cat} expected for full saturation but can importantly be used in comparative fashion between different conditions at the same peroxide concentration. The concentration of C45 was set to 100 nM C45 and the ABTS concentrations used were 8, 16, 32, 64, 128 and 256 μ M. For each point, C45 and H₂O₂ were injected using a plate reader into plates containing various amounts of ABTS, shaken for five seconds and activity was measured in quadruplicates. The initial rate v_0 was determined and divided by the enzyme concentration $[E]_0$. These rates were plotted against ABTS concentration and the resulting graph fitted to Michaelis-Menten kinetic parameters using Origin Software.

To ensure accuracy and investigate batch to batch variation of C45 as well as ABTS, this experiment was conducted five times. The graph displayed in **Figure 4.5** uses the averaged values for $v_0/[E]_0$ from all experiments. For each experiment, Michaelis-Menten kinetics were fitted individually (**Table 7.3** in the appendix),

4. Functional characterisation of C45 in organic cosolvents

highlighting there is no significant variation between different batches of C45 or ABTS.

The k_{cat} and K_{M} values under limiting peroxide concentrations were calculated by averaging the results of the individually calculated parameters (**Table 4.1**).

Table 4.1: Kinetic parameters for the oxidation of ABTS by C45 in aqueous solution.

k_{cat} (s^{-1})	K_{M} (μM)	$k_{\text{cat}}/K_{\text{M}}$ ($\text{s}^{-1} \mu\text{M}^{-1}$)
0.76 ± 0.07	24 ± 7	0.033 ± 0.008

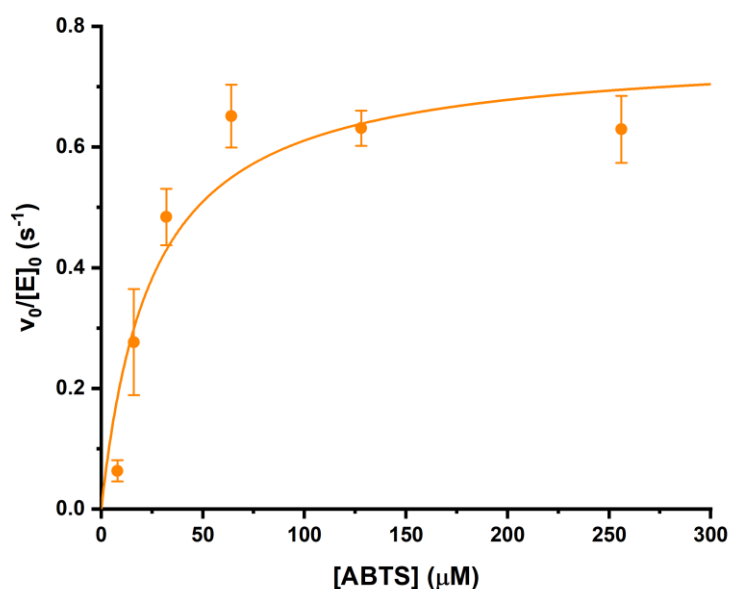


Figure 4.5: Michaelis-Menten kinetics of C45 catalysed ABTS oxidation in aqueous solution.

Kinetics of the oxidation of 8 - 256 μM ABTS by 100 nM C45 and 100 μM H_2O_2 at 25 $^\circ\text{C}$ in buffer fitted to Michaelis-Menten kinetics. Five experiments with four repeat measurements of each value were performed, results averaged with the standard deviation used as the error bar. Values were fitted to the Michaelis-Menten kinetic model.

The error for $k_{\text{cat}}/K_{\text{M}}$ was calculated using the Gaussian propagation of error (284). This states that random errors δx and δy associated with variables x and y used to calculate a further variable F propagate to the error δF according to the following formula:

$$\text{when } F = \frac{x}{y} \quad 4.2$$

$$\delta F = \sqrt{\left(\frac{\delta x}{y}\right)^2 + \left(\frac{x\delta y}{y^2}\right)^2} \quad 4.3$$

For 100 nM C45 (determined using the absorption at the Soret peak), 8 – 256 μM ABTS, and 100 μM H_2O_2 in buffer at room temperature, k_{cat} was $0.76 \pm 0.07 \text{ s}^{-1}$ and K_{M} $24 \pm 7 \mu\text{M}$. The resultant $k_{\text{cat}}/K_{\text{M}}$ is $0.03 \pm 0.01 \text{ s}^{-1} \mu\text{M}^{-1}$, all determined under limiting H_2O_2 concentration. This differs from the literature value of $1.16 \pm 0.05 \text{ s}^{-1}$ which was measured under the same conditions, except for the C45 concentration of 250 nM. In most cases, enzyme concentration can be mathematically corrected for in calculation of kinetic parameters to produce comparable values. However, in the case of C45, this appears not to be true. In chapter 5 a more thorough kinetic investigation can be found which shows that in the case of C45, enzyme concentration strongly influences kinetic behaviour. The effect of varying peroxide concentration is also probed further in chapter 5.

Generally, the error associated with the calculated k_{cat} is much smaller than that of K_{M} . This is due to the high reactivity of compound I leading to rapid inactivation of the enzyme. This was particularly evident at low substrate concentrations which predominantly affect K_{M} . A much larger number of repeat measurements would have to be performed to overcome this. Thus, all analyses will focus only on k_{cat} , while effects on K_{M} and $k_{\text{cat}}/K_{\text{M}}$ will not be discussed due to their substantial errors.

4. Functional characterisation of C45 in organic cosolvents

Total turnover numbers (**Figure 4.6**) represent the absolute number of turnovers each enzyme performed during the assay and are calculated by dividing the concentration of the totally formed product (derived from absorbance difference at 405 nm) by enzyme concentration. Excitingly, it was observed that total turnover numbers increase significantly with rising ABTS concentrations. This is likely related to the instability of the radical compound I which rapidly undergoes off-pathway reactions that destroy the enzyme (285, 286). Increasing the saturation with ABTS results in decreased binding times and favour the oxidation of ABTS over any competing side reactions, thereby increasing the total turnover of each enzyme.

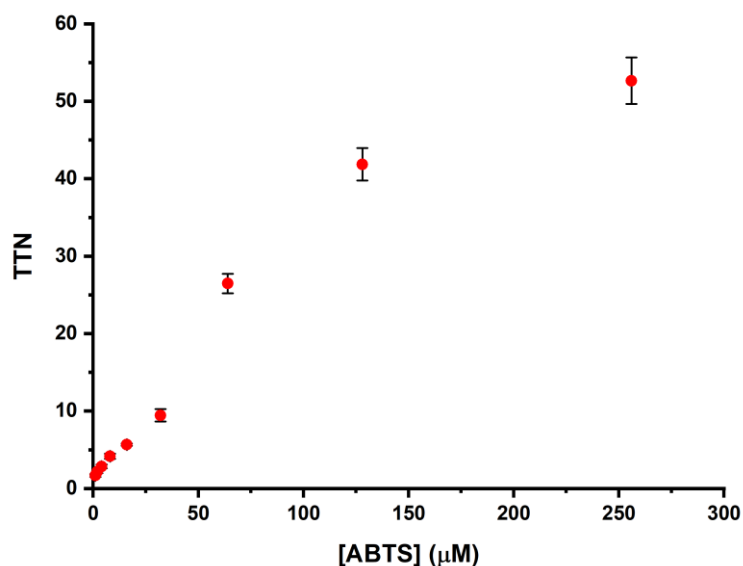


Figure 4.6: Total turnover numbers for the oxidation of ABTS by C45 and 100 μM H_2O_2 in aqueous solution.

The total turnover is larger for higher substrate concentrations.

4.2.1.2. Michaelis-Menten Kinetics of C45 in Acetonitrile

The characterisation of C45 in buffer was the basis for comparison of the enzyme's catalytic behaviour in different solvents. The experiment was repeated in various cosolvents at increasing concentrations. Reactants were mixed with 10-80% acetonitrile and tested under otherwise equal conditions as the previous experiments in buffer (**Figure 4.7 a**). Michaelis-Menten parameters were determined using Origin software for fitting. Resultant values are shown in **Figure 4.7 b-c**.

As the concentration of acetonitrile is raised, k_{cat} generally decreases. At concentrations of up to 40% acetonitrile, the decline in k_{cat} is limited to approximately 20%, whereas in 50-60% acetonitrile, the activity is halved. In 80% acetonitrile, the enzyme is entirely inactive which is in agreement with the CD results (**Figure 3.9**) showing complete denaturation of the enzyme in 80% acetonitrile. CD spectra of C45 in lower acetonitrile concentrations do not suggest significant loss of secondary structure, however. The functional effect of the cosolvent therefore appears to be more pronounced than any structural effects observed by CD. This may be related to the lower polarity of the environment which makes substrate desolvation - and thereby binding - less favourable.

There is, however, a clear trend of decreasing total turnover numbers at the maximum ABTS concentration of 256 μM with increasing acetonitrile concentration. This trend is consistent with the decreasing k_{cat} values. Viewing this alongside the structural analyses suggesting that C45 retains its secondary structure in up to 60% acetonitrile, this trend may suggest that not the entire enzyme denatures, but the active site does. To analyse this theory, further structural characterisation would be required, using methods such as NMR. There are likely also effects of the cosolvent on the reaction kinetics independent of the enzyme's structural integrity, such as

4. Functional characterisation of C45 in organic cosolvents

lowered stabilisation of an intermediate which may hinder the reaction, which could also account for the declining total turnover numbers. It is also likely that the cosolvent interferes with substrate binding.

4.2. Results - Michaelis-Menten Kinetics of C45 with low substrate concentrations

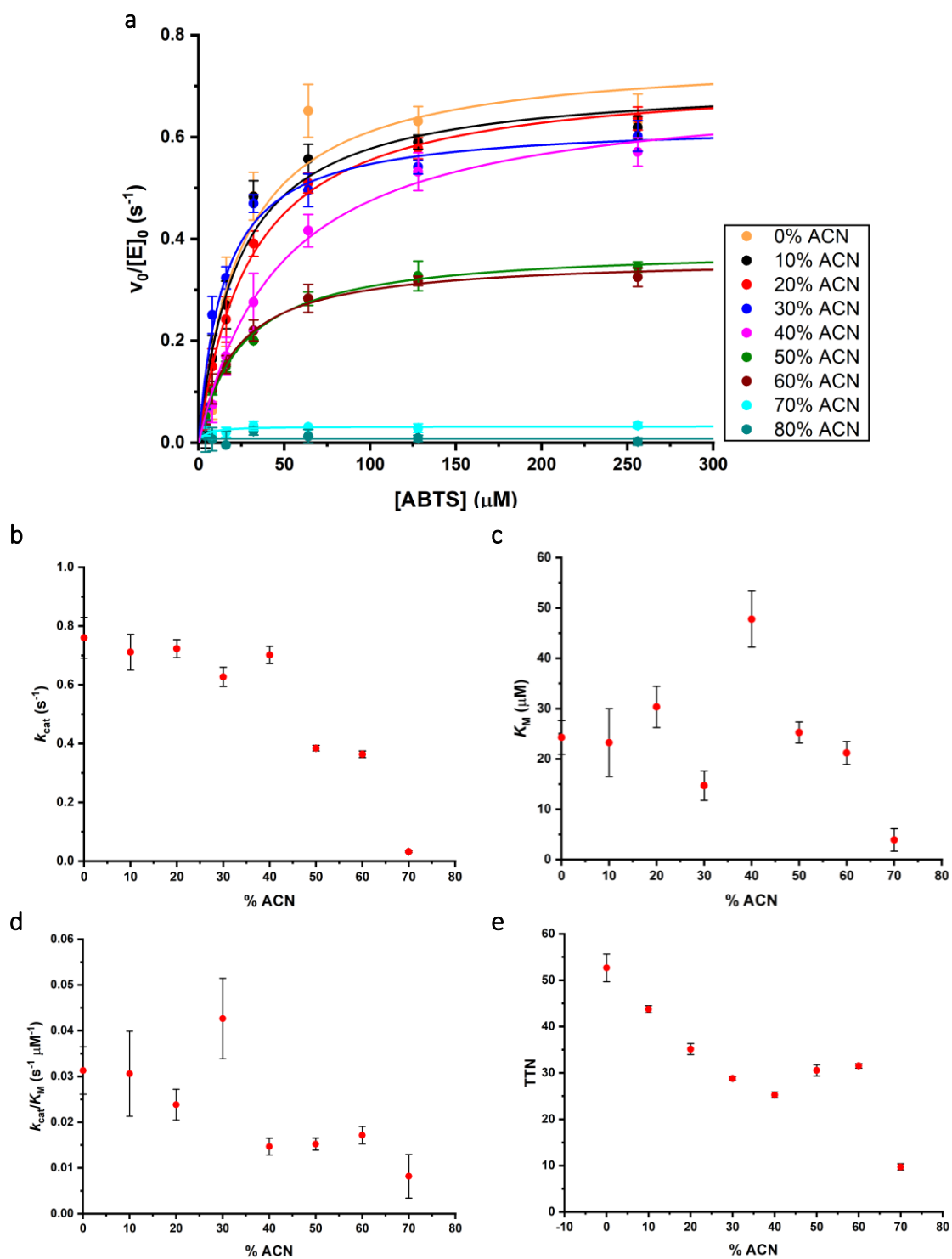


Figure 4.7: Michaelis-Menten kinetics of ABTS^{•+} formation in acetonitrile.

Kinetics of the oxidation of 4 - 256 μM ABTS by 100 nM C45 and 100 μM H₂O₂ at 25 °C in aqueous solution with 0 - 80% ACN. **a** - Four repeat measurements of each value were taken averaged, standard deviation was used as the error bar and the values fitted to the Michaelis-Menten kinetic model. Values calculated for k_{cat} (**b**), K_M (**c**) as well as k_{cat}/K_M (**d**) and total turnover numbers for 256 μM ABTS (**e**) were plotted against cosolvent concentration.

4.2.1.3. Michaelis-Menten Kinetics of C45 in Methanol

The same methods were used to analyse the effects of increasing methanol concentration on the kinetics of C45-mediated ABTS oxidation under otherwise equal conditions. Results are shown in **Figure 4.8**. k_{cat} values appear consistent in up to 40% methanol, with only a 13% loss between pure buffer (0.76 s^{-1}) and 40% methanol concentration (0.67 s^{-1}). In higher cosolvent concentrations, k_{cat} decreases rapidly. CD and UV-visible spectroscopy results do not indicate strong denaturation of the enzyme even in high methanol concentrations of 60-80%. This suggests that the decline in catalytic activity is either not entirely due to protein denaturation or that the denaturation that does occur is not observable using these spectroscopic methods.

Total turnover numbers decrease to 32.3 ± 0.8 in 80% methanol. However, they are consistent in up to 60% methanol: In buffer, this is 53 ± 3 , in 30-60% methanol between 50 and 60. In 60% methanol, the k_{cat} decreased by 65%, however total turnover only declined by 5% compared to the results in pure buffer. This means each enzyme performs an almost equal number of substrate conversions, however at a much lower rate. This suggests that the lowered catalytic activity may not be primarily caused by enzyme denaturation.

4.2. Results - Michaelis-Menten Kinetics of C45 with low substrate concentrations

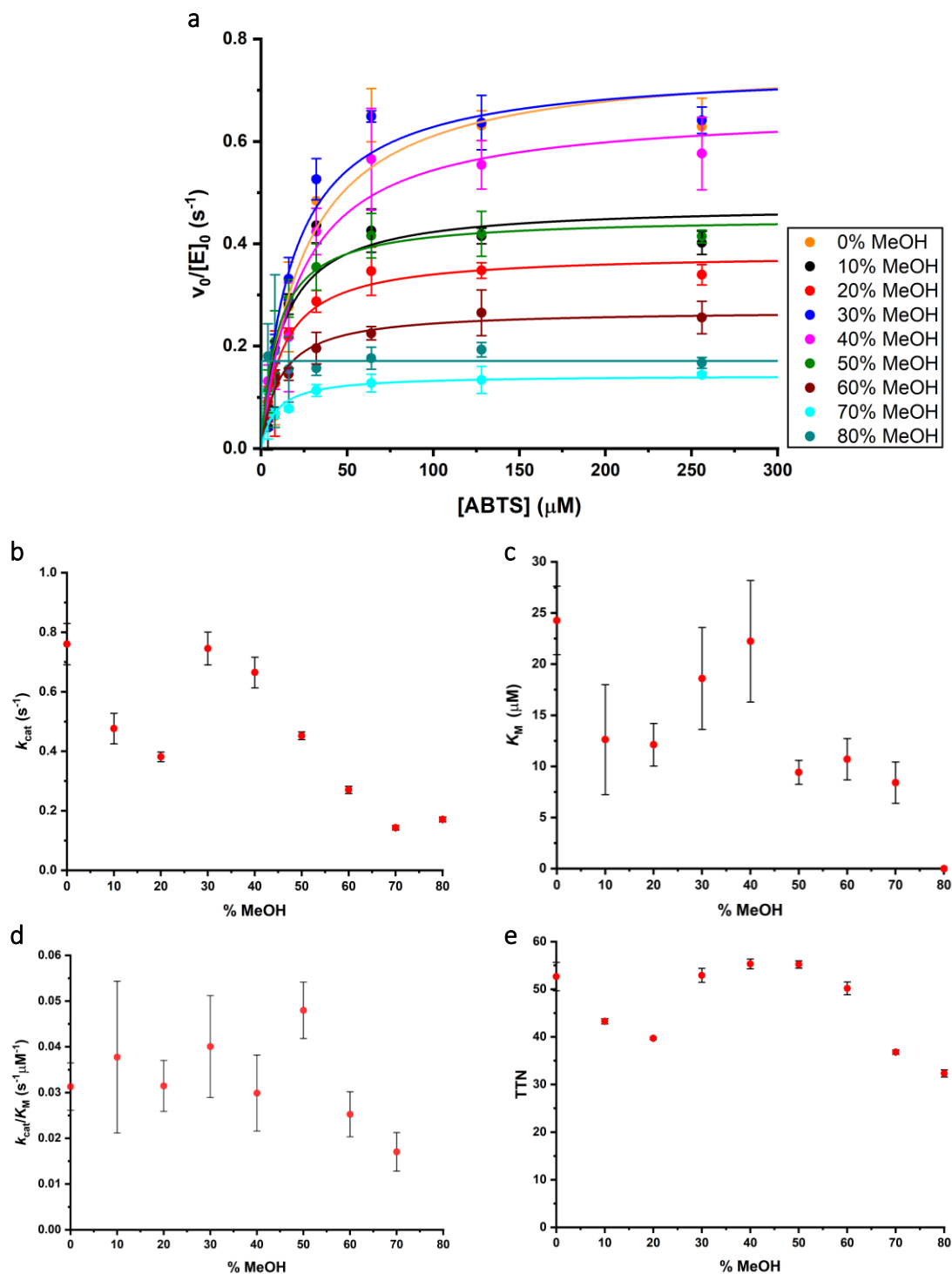


Figure 4.8: Michaelis-Menten kinetics of $ABTS^{+\bullet}$ formation in methanol.

Kinetics of the oxidation of 4- 256 μM ABTS by 100 nM C45 and 100 μM H_2O_2 at 25 $^{\circ}C$ in aqueous solution with 0 - 80% MeOH. **a** - Four repeat measurements of each value were taken averaged, standard deviation was used as the error bar and the values fitted to the Michaelis-Menten kinetic model. Values for k_{cat} (**b**), K_M (**c**) as well as k_{cat}/K_M (**d**) and total turnover numbers for 256 μM ABTS (**e**) were plotted against cosolvent concentration.

4.2.1.4. Michaelis-Menten Kinetics of C45 in Ethanol

The enzyme's catalytic behaviour in ethanol was probed using the same methods as in acetonitrile and methanol (**Figure 4.9**). The enzyme's k_{cat} in any ethanol concentration is much lower than that of C45 in pure buffer. In 10-30% ethanol, rates are somewhat consistent, with k_{cat} values between 0.31 s^{-1} and 0.41 s^{-1} compared to 0.76 s^{-1} in aqueous solution. In higher ethanol concentrations, k_{cat} values decline steadily. In 80% ethanol no catalytic activity is observed. Total turnover numbers mirror this pattern.

The lack of catalytic activity in 80% ethanol is not surprising when considering the denaturation observed using CD in chapter 3 (**Figure 3.9**). The strong decline in catalytic activity in lower ethanol concentrations, however, was not anticipated from structural analysis, similar to the decline in enzymatic activity of C45 in acetonitrile and methanol.

4.2. Results - Michaelis-Menten Kinetics of C45 with low substrate concentrations

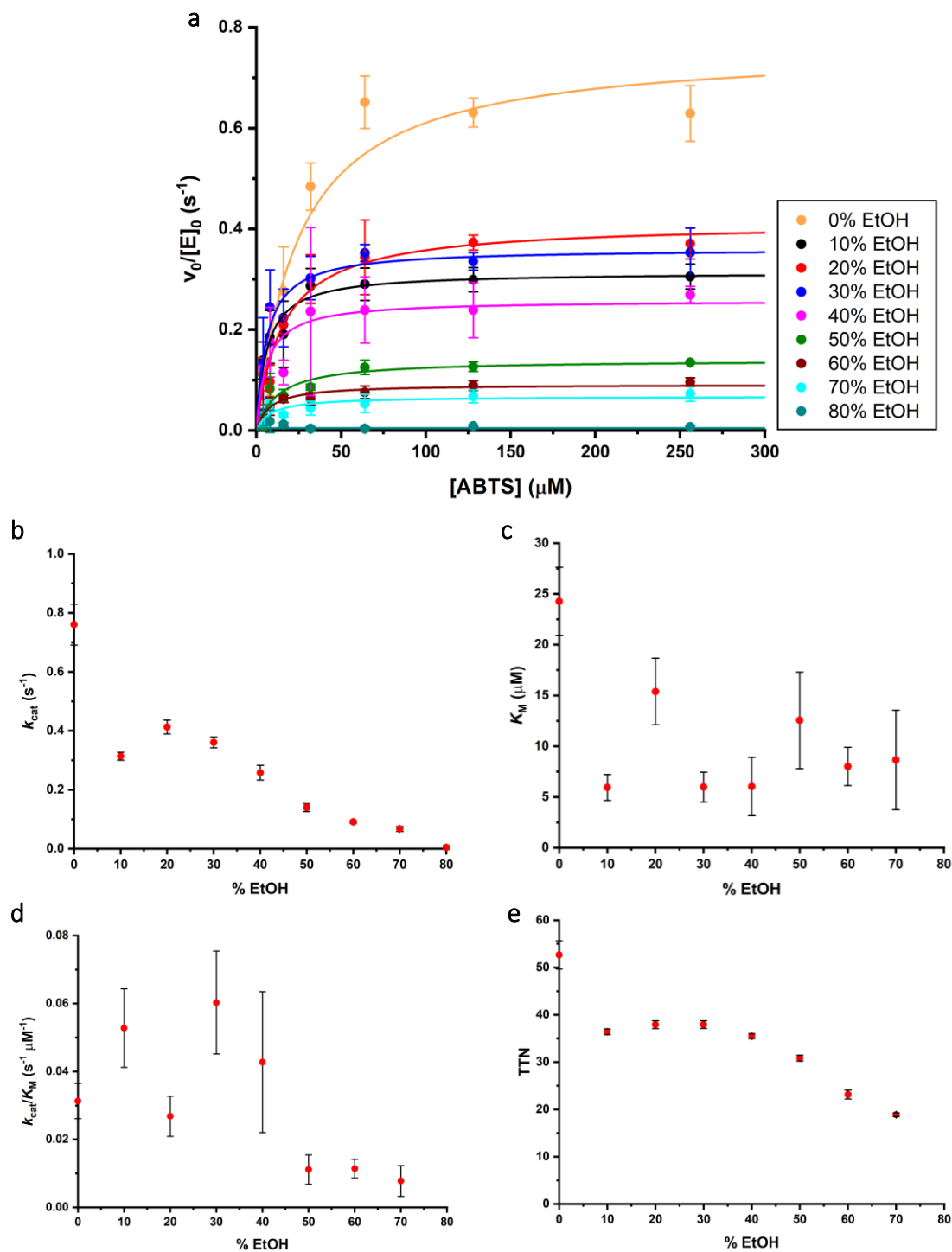


Figure 4.9: Michaelis-Menten kinetics of ABTS^{•+} formation in ethanol.

Kinetics of the oxidation of 4 - 256 μM ABTS by 100 nM C45 and 100 μM H₂O₂ at 25 °C in aqueous solution with 0 - 80% EtOH. **a** - Four repeat measurements of each value were taken averaged, standard deviation was used as the error bar and the values fitted to the Michaelis-Menten kinetic model. Values calculated for k_{cat} (**b**), K_M (**c**) as well as k_{cat}/K_M (**d**) and total turnover numbers for 256 μM ABTS (**e**) were plotted against cosolvent concentration.

4.2.1.5. Michaelis-Menten Kinetics of C45 in Isopropanol

In rising isopropanol concentrations, the decline in enzymatic activity is much steeper than in the other analysed cosolvents (**Figure 4.10**). In isopropanol concentrations of 30-40%, no enzymatic activity was observed. C45 was not tested in higher isopropanol concentrations than 40% as it is highly unlikely that the strong trend of decreasing catalytic activity with isopropanol addition is reversed at higher concentrations.

CD results, however, only indicated denaturation in much higher isopropanol concentrations, above 60%. Compared to the strong decline in catalytic activity in much lower isopropanol concentrations, this suggests that either further unfolding occurred that was not observable using the spectroscopic methods employed in this work - such as denaturation of the tertiary structure or highly localised denaturation at the active site - or that the decline in enzymatic activity is not caused by loss of protein structure.

4.2. Results - Michaelis-Menten Kinetics of C45 with low substrate concentrations

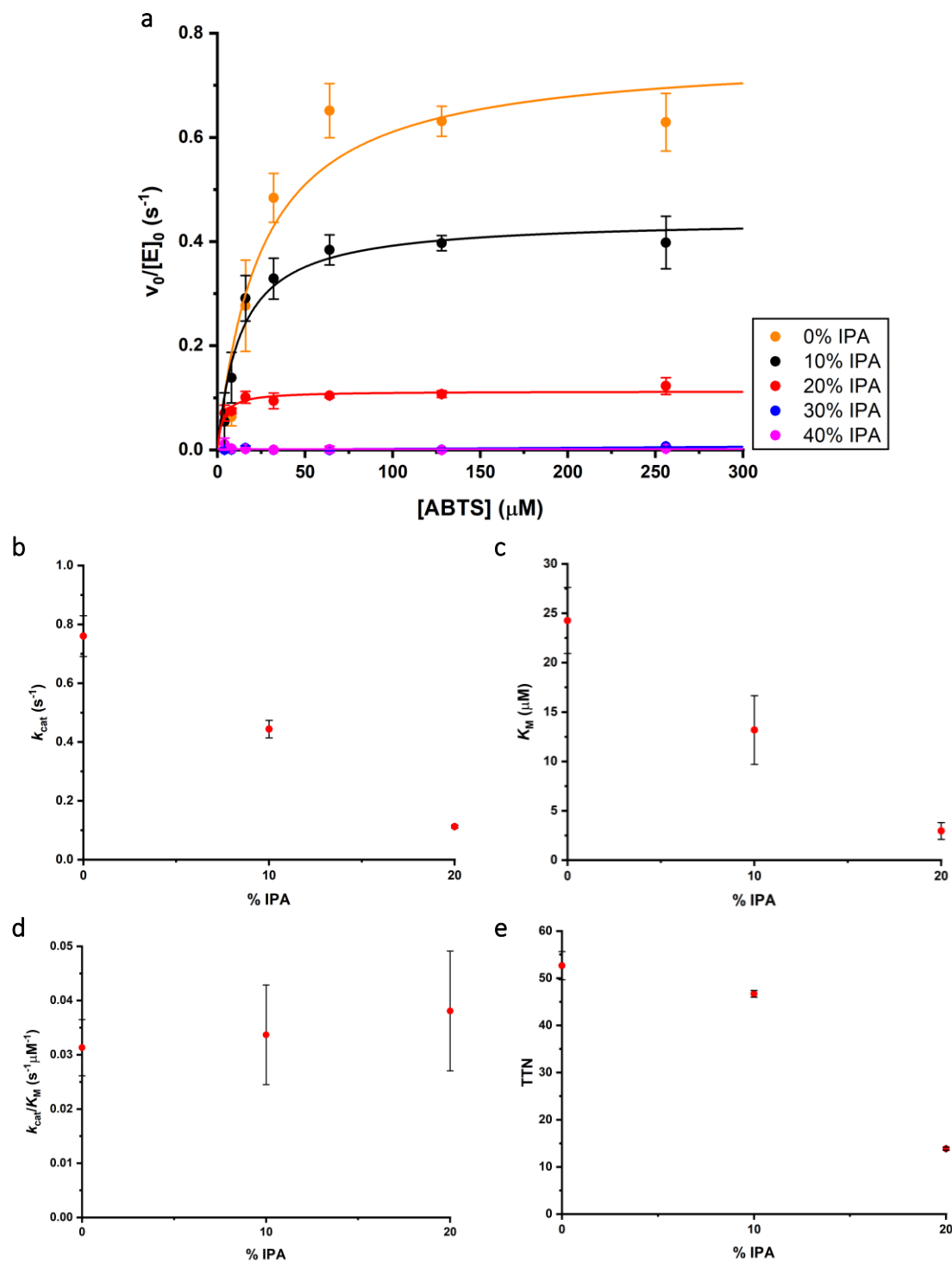


Figure 4.10: Michaelis-Menten kinetics of ABTS^{•+} formation in isopropanol.

Kinetics of the oxidation of 4- 256 μM ABTS by 100 nM C45 and 100 μM H₂O₂ at 25 °C in aqueous solution with 0 - 40% IPA. **a** - Four repeat measurements of each value were taken averaged, standard deviation was used as the error bar and the values fitted to the Michaelis-Menten kinetic model. Values calculated for k_{cat} (**b**), K_M (**c**) as well as k_{cat}/K_M (**d**) and total turnover numbers for 256 μM ABTS (**e**) were plotted against cosolvent concentration.

4.2.1.6. Michaelis-Menten Kinetics of C45 in TFE

With the increasing concentrations of acetonitrile, methanol, ethanol and isopropanol, the catalytic activity of C45 dropped strongly. In TFE, on the other hand, reaction rates and catalytic turnover of C45 actually increased (**Figure 4.11**).

Due to this increase, a wider range of substrate concentrations was used (2 – 1024 μM ABTS) to obtain more accurate data. However, the values for $\frac{v_0}{[E]_0}$ reached a plateau at ABTS concentrations above approximately 150 μM and did not increase significantly with higher substrate concentrations.

Furthermore, to ensure consistency of these results, the experiment in 80% TFE was conducted four times and results averaged. Due to the large increase in reaction rate and the associated strong absorbance, measurements were performed with a lower C45 concentration, which was accounted for in calculations of kinetic parameters.

In up to 30% TFE, k_{cat} remains at a consistent level, however, in 50-60% TFE, k_{cat} is approximately 60% higher than in aqueous solution, at $1.22 \text{ s}^{-1} \pm 0.08$. This rises further to $2.2 \text{ s}^{-1} \pm 0.3$ in 60% TFE. In 80% TFE, the highest cosolvent concentration tested, k_{cat} is $5.1 \text{ s}^{-1} \pm 0.3$, a 7-fold increase to the k_{cat} in aqueous solution.

The total turnover number was plotted for both 256 μM , as in the previous experiments, and 1024 μM , the maximum ABTS concentration used in TFE. Interestingly, the two data sets overlap strongly, suggesting that no further increase in total turnover number occurs at substrate concentrations above 256 μM . This implies that at 256 μM ABTS, the reaction is sufficiently saturated for maximum total turnover.

4.2. Results - Michaelis-Menten Kinetics of C45 with low substrate concentrations

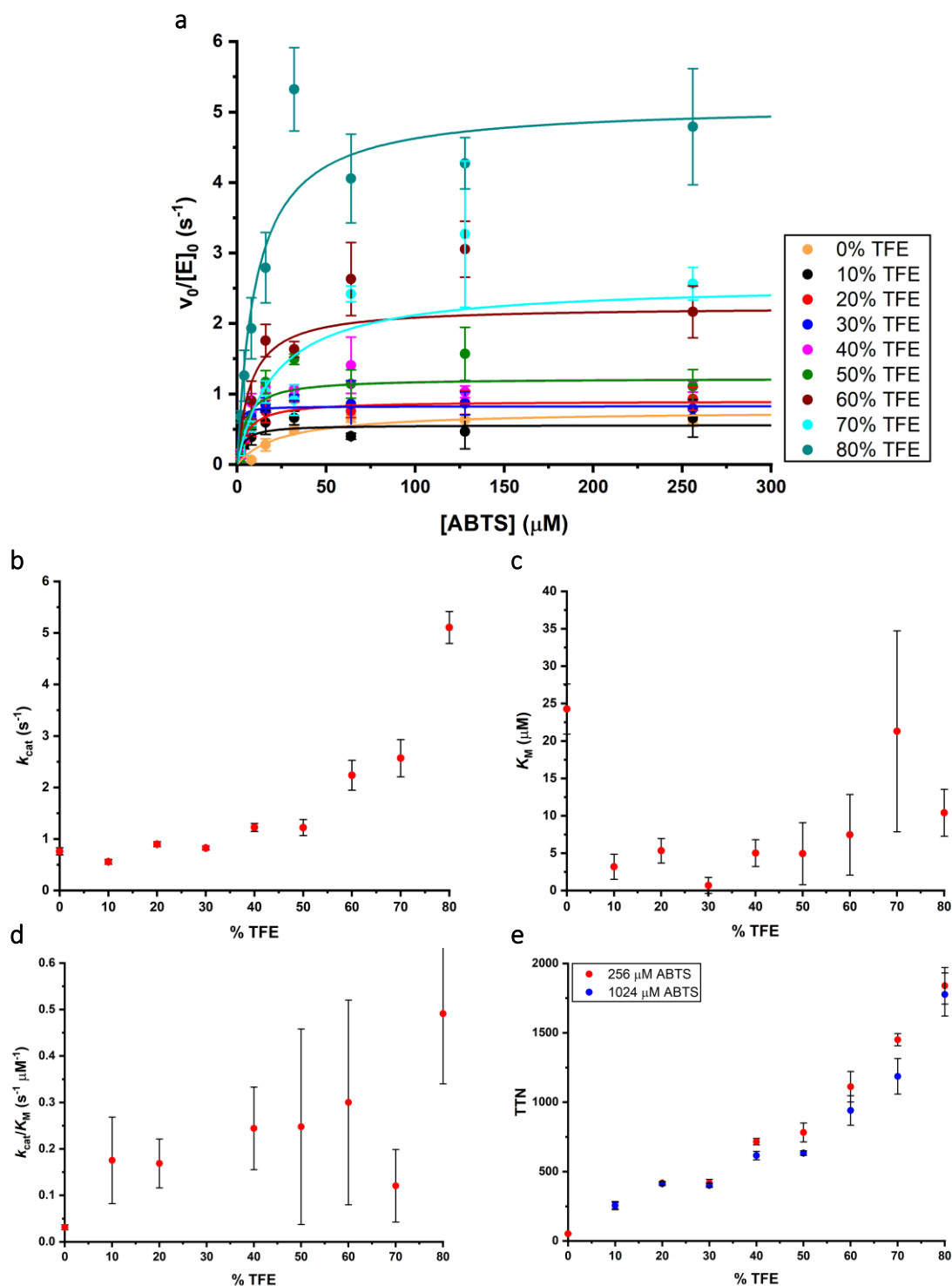


Figure 4.11: Michaelis-Menten kinetics of ABTS^{•+} formation in TFE.

Kinetics of the oxidation of 2 - 256 μM ABTS by 10, 25 and 50 nM C45 and 100 μM H₂O₂ at 25 °C in aqueous solution with 0 - 80% TFE. **a** - Four repeat measurements of each value were taken averaged, standard deviation was used as the error bar and the values fitted to the Michaelis-Menten kinetic model. For clarity, measurements at 512 and 1024 μM ABTS are not shown.

Values calculated for k_{cat} (**b**), K_M (**c**) as well as k_{cat}/K_M (**d**) and total turnover numbers for 256 μM and 1024 μM ABTS (**e**) were plotted against cosolvent concentration.

4. Functional characterisation of C45 in organic cosolvents

Whilst both total turnover number and k_{cat} increase with ascending TFE concentrations, the total turnover does this at a much steeper rate, as shown in **Figure 4.12**. The rising total turnover number indicates that each enzyme molecule performs a much larger number of conversions in 80% TFE than in aqueous solution. The rate at which these conversions occur is also much higher but does not increase by the same factor. There might be some stabilisation of an intermediate, possibly compound I, which positively impacts both k_{cat} and total turnover number.

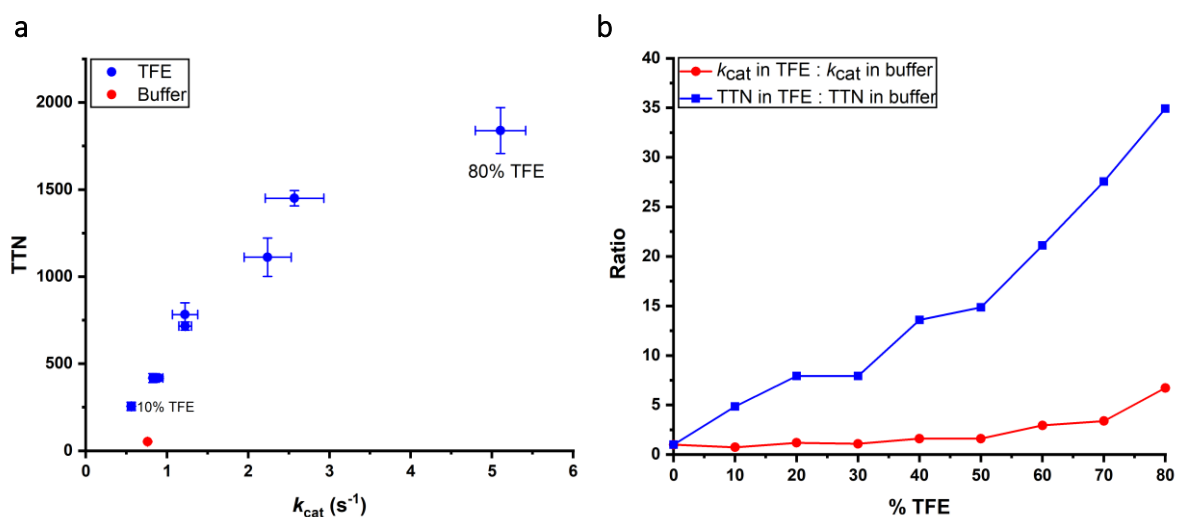


Figure 4.12: k_{cat} and total turnover number (TTN) in TFE.

a – TTN plotted against k_{cat} , showing both increase significantly in rising TFE concentrations. **b** – Ratios of k_{cat} and TTN in TFE to the same values in buffer, showing that the TTN increases at a much steeper rate than the k_{cat} as TFE concentration is raised.

4.3. Discussion

In this chapter, the impact of different cosolvents on the peroxidase activity of C45 was scrutinised. In buffer, this reaction proceeds with a k_{cat} of $0.76 \text{ s}^{-1} \pm 0.07$ under limiting concentrations of peroxide ($100 \text{ }\mu\text{M}$). Under these conditions and $256 \text{ }\mu\text{M}$ ABTS, the maximum concentration used, a total turnover number of 53 ± 3 was

determined for C45. The K_M of this reaction could not be adequately calculated as this requires accurate data for the reaction rate at low substrate concentrations. At low substrate concentrations, the activities are challenging to measure due to the reactive nature of intermediate compound I and more accurate data would require several replicates. Similarly, the total turnover number showed a strong dependency on the concentration of ABTS which is most likely also due to the reactivity of compound I.

In the organic cosolvents tested, with the exception of TFE, k_{cat} and total turnover number generally decreased with increasing cosolvent concentrations (**Figure 4.13**). As indicated in **Table 4.2**, cosolvent concentrations of 30% or lower already affected the peroxidase activity of C45. Excitingly though, in TFE, the highest cosolvent concentration of 80% yielded the greatest k_{cat} of $5.1 \text{ s}^{-1} \pm 0.3$ and total turnover number of 1840 ± 140 .

Table 4.2: Highest k_{cat} values for ABTS oxidation in different cosolvents.

	[Cosolvent]	$k_{cat} \text{ (s}^{-1}\text{)}$
Buffer		0.76 ± 0.07
ACN	20%	0.72 ± 0.03
MeOH	30%	0.75 ± 0.06
EtOH	20%	0.41 ± 0.03
IPA	10%	0.44 ± 0.03
TFE	80%	5.1 ± 0.3

Literature (3, 5, 139, 145, 287) as well as structural data presented in chapter 3 suggested that helical proteins such as C45 might be more stable in TFE. From this, an increase in enzymatic activity, manifested in a higher k_{cat} and total turnover number, could be expected. This was indeed the case, the catalytic activity of C45 increased strongly in TFE, however significantly more so than anticipated from the observed structural changes. This was particularly notable in high concentrations of TFE: In 80% cosolvent concentration, the k_{cat} measured was $5.1 \text{ s}^{-1} \pm 0.3$, a 7-fold increase on the

4. Functional characterisation of C45 in organic cosolvents

k_{cat} in aqueous solution. The total turnover number, however, increased by a much larger factor. At 1840 ± 140 , this is approximately 34 times larger than the total turnover number in water (both calculated at an ABTS concentration of $256 \mu\text{M}$). In other cosolvents, the relative change in k_{cat} and total turnover number are much more parallel. Generally, compared to the change observed in both k_{cat} and total turnover number in TFE, the change of catalytic activity in other cosolvents is small. This suggests that the intermediates in the reaction, especially the highly reactive radical compound I, are stabilised by TFE and the desired reaction with the substrate is favoured strongly over side reactions that would result in enzyme degradation.

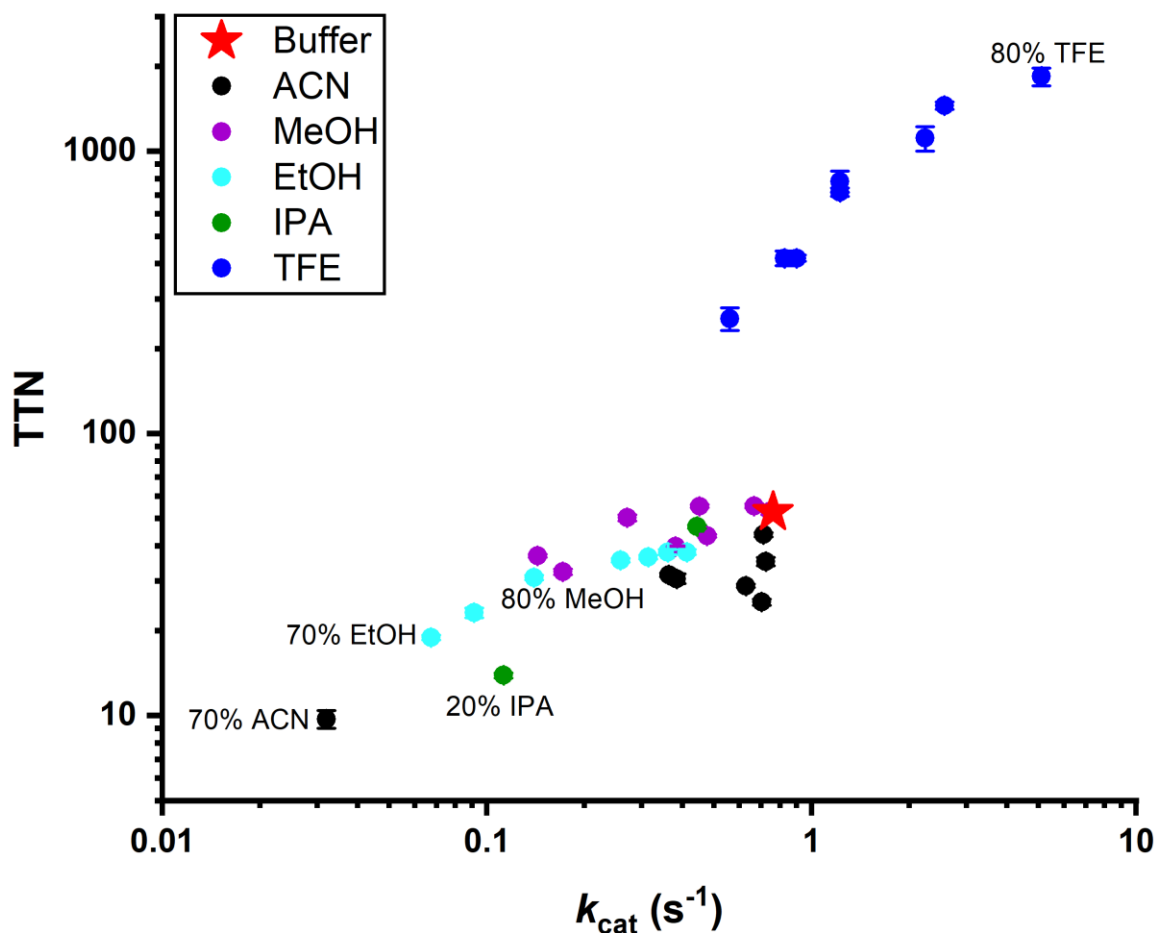


Figure 4.13: Total turnover number and k_{cat} for different cosolvents.

A decrease of catalytic activity can be seen as cosolvent concentration is increased for all cosolvents, except for TFE, in which both parameters increase strongly.

In other cosolvents, the general trend appeared to be an inverse relationship of cosolvent concentration and catalytic activity, which is common for enzymology in aqueous organic mixtures (13, 16). k_{cat} values in methanol and acetonitrile are very similar to each other, exhibiting resilience to cosolvent concentration of up to 40% before steadily decreasing with rising cosolvent concentration. In ethanol, k_{cat} values drop more precipitously as cosolvent concentration is elevated than in methanol or acetonitrile, and in isopropanol even more so, with total turnover numbers also reflecting this decline in enzymatic activity. To determine whether this is a generic trend of decreasing catalytic activity in alcohols of increasing carbon chain length, further analysis would be required. This observed behaviour may also be related to polarity – in particular, polarity as measured by the $E_{\text{T}}(30)$ scale (**Table 3.1**) - as less polar cosolvents tended to produce a greater decline in enzymatic activity, which may be indicative of some interference substrate-binding. The binding of organic substrates is often driven by hydrophobic interactions which may be disrupted by a less polar solvent. Again, this theory would require further testing with a wider range of cosolvents.

The general trend of decreasing catalytic activity with increasing organic cosolvent concentration is coherent with the loss of structure in high cosolvent concentrations that was demonstrated in chapter 3. However, the spectroscopic data presented in chapter 3 only indicate loss of structure in cosolvent concentrations of 60% or higher. The decline of enzymatic activity, however, commences at much lower cosolvent concentrations. This suggests one of two possibilities. Either this decline in enzymatic activity is not primarily caused by denaturation of the protein or there is more unfolding occurring than observable using CD and UV-visible spectroscopy. This could, for instance, be highly localised denaturation capable of deactivating the protein without

effecting sufficient loss of topology for these methods to expose. It may be worth analysing the protein structure in cosolvents using some different methods, such as NMR.

Generally, the negative impact of cosolvents on the function of C45 is not unanticipated, however the behaviour of C45 in TFE is intriguing. The difference in the impact of TFE and other cosolvent seems to be two-fold: Firstly, the increase rather than decline in enzymatic activity with the addition of TFE and secondly the possible stabilisation of intermediates. To gain further insights into the complete kinetics of C45 in TFE, especially at the initial stages of the reaction, and to understand the underlying mechanisms, this was probed further. A more thorough investigation of the catalytic activity of C45 in TFE can be found in chapter 5.

5. Further kinetic and structural analysis of C45 in TFE

In the previous chapters, the secondary structure and catalytic activity of C45 was studied in buffer, acetonitrile, methanol, ethanol, isopropanol and TFE. While secondary structure appeared to be resilient to the addition of up to 80% cosolvent and even increases in helicity with the addition of alcohols, catalytic activity generally declined significantly after the addition of 50% cosolvent. The exception was TFE, which significantly enhanced catalytic activity, especially at 80% cosolvent concentration. In this chapter, the peroxidase kinetics of C45 in buffer and 80% TFE are thus probed more thoroughly and steady-state as well as pre-steady-state kinetic data presented. Furthermore, a comparison with horseradish peroxidase is performed, as well as an investigation into the mechanism of compound I formation. Finally, the structural effects of TFE on C45 were probed using analytical SEC, SEC-SAXS and NMR.

5.1. Introduction

TFE is a water-miscible, amphiphilic alcohol that can stabilise both the polar exterior and the hydrophobic core of a protein. Mixtures of water and TFE have been shown to stabilise secondary structural motifs in proteins, in particular α -helices (3, 4, 139). The mechanism for this is not fully understood yet, however, displacement of the aqueous hydration shell with TFE likely leads to destabilisation of random coils and increased hydrogen bonding within the protein, which seem to play important roles in

such stabilisation (3, 5, 139, 145). Nonetheless, these effects primarily act on short range interactions, and have been shown to actually destabilise protein tertiary structure in order to increase the protein helical content in some instances (141, 142). This is unlikely in C45 though, as the protein comprises only helices and short loops of serine and glycine, which have been demonstrated to function as α -helix caps even in TFE (3). The observed increases in helicity in C45 are therefore more likely a result of a higher stability of the helical fold over random coil. Many enzymes, including horseradish peroxidase, exhibit a strong decline in catalytic activity upon the addition of over 30% TFE (6, 288–290), even if in some instances lower concentrations yielded a slight increase in catalytic activity. Therefore, the significant rise in enzymatic activity of C45 in a TFE concentration as high as 80% was surprising, sparking interest in a more thorough comparison of the kinetics of C45 in aqueous solution and 80% TFE.

Kinetic isotope effects may provide valuable insights into the kinetics and mechanism of a reaction (291), and are determined by the reaction rate ratios with a light and a heavy isotope. While this effect can be seen between isotopes of any element, the relative change in nuclear mass – and therefore the difference in zero-point energy which is inversely related to nuclear mass – is largest between hydrogen and deuterium (292). A deuterium bond has thus a the lower zero-point energy than a hydrogen bond, meaning that more energy is required to break a deuterium bond than a corresponding bond involving hydrogen (293). **Figure 5.1** shows a Morse potential energy curve of bonds between carbon and the two hydrogen isotopes. The diagram especially highlights the relation between zero-point energy of a bond and its dissociation energy.

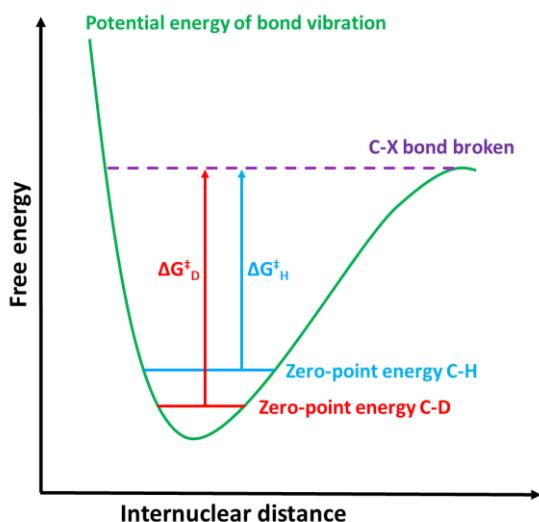


Figure 5.1: A Morse potential energy curve of C-H and C-D bonds.

This means that a reaction step where carbon-hydrogen bonds are broken can be measurably slowed down by substituting the hydrogen atom with its heavier isotope. This difference in reaction rate is known as the kinetic isotope effect (KIE) and can be described by the following ratio of rate constants.

$$KIE = \frac{k_H}{k_D} \quad 5.1$$

5.2. Results

5.2.1. Kinetics of C45 with high substrate concentration

As previously discussed, the reaction between C45 consumes hydrogen peroxide and ABTS to form the turquoise radical cation $ABTS^{+\bullet}$, following ping-pong kinetics. In chapter 4, this reaction was performed in various cosolvents under limiting peroxide concentrations. A more complete analysis of this reaction – with the concentrations of

both substrates varied and performed on a stopped-flow spectrometer – has been conducted in buffer and 80% TFE. This instrument can investigate much faster reactions and allowed the usage of higher substrate concentrations (0.1-3 mM ABTS and 5-200 mM H_2O_2). Twelve replicates were measured for each data point in two separate experiments and results averaged (**Figure 5.2**).

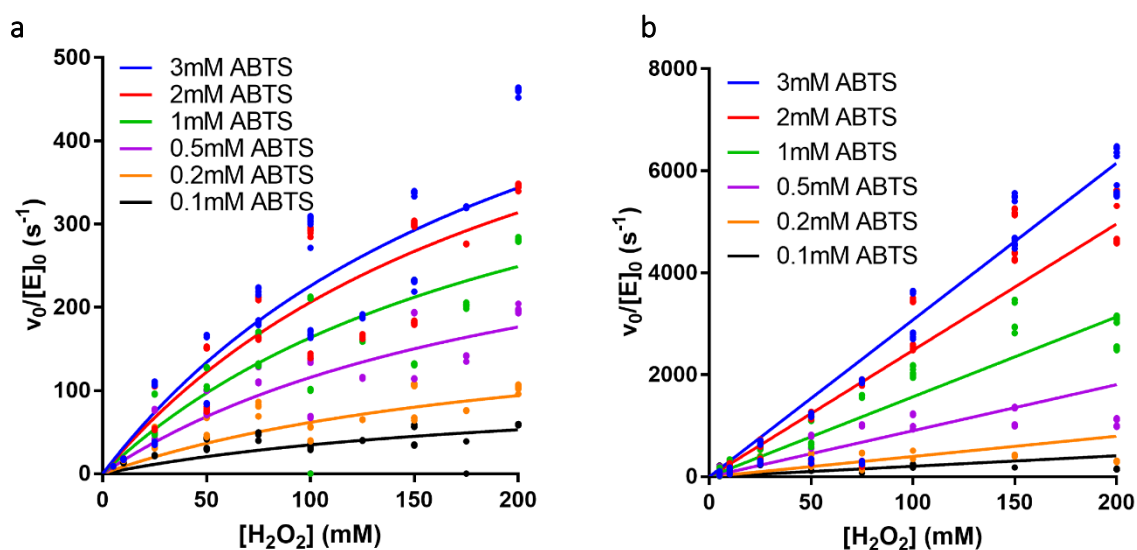


Figure 5.2: Kinetics of ABTS and peroxide turnover by C45 in buffer and 80% TFE. Data for ABTS oxidation by C45 in buffer (a) and 80% TFE (b) are fitted to a ping-pong kinetic model.

The data obtained for C45 in buffer were fitted to the ping-pong model, with fits suggesting a k_{cat} of $890 \pm 60 \text{ s}^{-1}$, a K_M for H_2O_2 of $220 \pm 30 \text{ mM}$, a K_M for ABTS of $0.70 \pm 0.04 \text{ mM}$ and a $k_{\text{cat}}/K_{M_a}/K_{M_b}$ of $5.8 \times 10^6 \text{ M}^{-1} \text{ M}^{-1} \text{ s}^{-1}$. It is, however, noteworthy that especially in higher ABTS concentrations (1 mM and above), the fits matched the experimental data very poorly and the errors associated with the data calculated from these fits are therefore likely larger.

The calculated k_{cat} of 890 s^{-1} was significantly smaller than the previously measured k_{cat} of 1200 s^{-1} , which had been obtained from measurements conducted in

5. Further kinetic and structural analysis of C45 in TFE

the same buffer and at the same temperature and pH, but with different C45 concentrations (2). As batch to batch variation was minimal (see **section 7.1** in the Appendix), the impact of enzyme concentration on the reaction is investigated in the following section.

Further, the data for the reaction in 80% TFE confirm that the presence of this cosolvent significantly increases the reaction rate to the extent that saturation was not achieved with the substrate concentrations used. Because of this, an accurate fit and calculations for a K_M for H_2O_2 and a k_{cat} could be performed.

In contrast to the k_{cat} and the K_M for H_2O_2 , the K_M for ABTS could be fitted, and was calculated to be 2.8 ± 0.05 mM. From the slope, the $k_{cat}/K_{Ma}/K_{Mb}$ was estimated to be $2.1 \times 10^7 M^{-1} M^{-1} s^{-1}$ which is a 4-fold increase on the catalytic efficiency in buffer. Neither k_{cat} nor K_M for H_2O_2 could be determined with the substrate concentrations used here, however. The kinetic parameters determined both in buffer and 80% TFE are summarised in **Table 5.1**.

Table 5.1: Kinetic parameters for ABTS and peroxide turnover by C45 in buffer and 80% TFE.

	k_{cat} (s^{-1})	K_M for H_2O_2 (mM)	K_M for ABTS (mM)	$k_{cat}/K_{Ma}/K_{Mb}$ ($M^{-1} M^{-1} s^{-1}$)
Buffer	890 ± 60	220 ± 30	0.70 ± 0.04	5.8×10^6
80% TFE	-	-	2.8 ± 0.05	2.1×10^7

5.2.1.1. Impact of C45 concentration

In light of the discrepancy between kinetic results obtained consistently within this work and those in the literature in the same buffer and at the same temperature and pH, but with different C45 concentrations, the impact of the enzyme concentration was investigated. C45 concentration was thus varied, with two sets of substrate of high (3

mM ABTS and 200 mM H₂O₂) and low (0.1 mM ABTS, 25 mM H₂O₂) concentrations (Figure 5.3).

Surprisingly, the enzyme concentration significantly influenced $v_0/[E]_0$. Rate increases are generally proportional to enzyme concentration and the initial rate normalised for enzyme concentration should be constant for all enzyme concentrations. For C45, however, this increases significantly with increasing enzyme concentration, regardless of substrate concentration and reaction medium (aqueous buffer or 80% TFE).

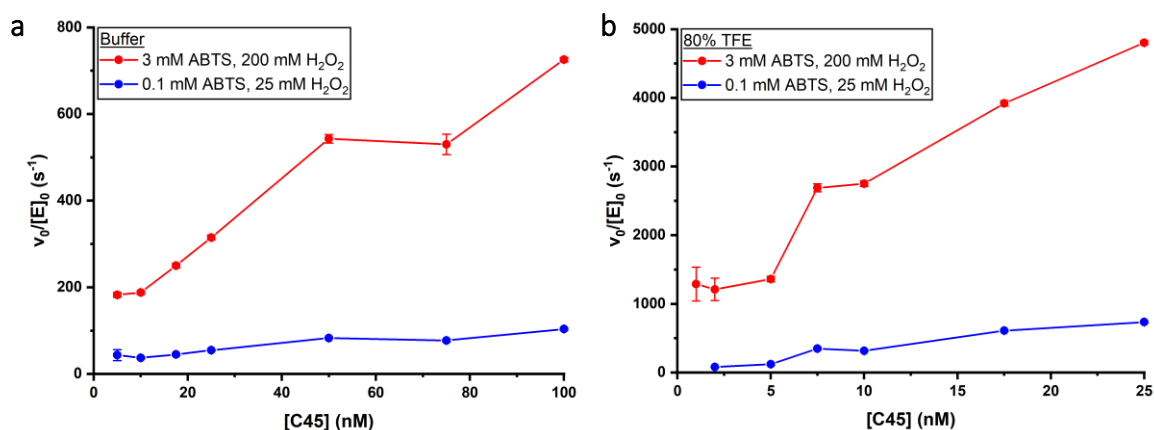


Figure 5.3: Enzyme concentration dependent $v_0/[E]_0$ for the reaction between C45, ABTS and H₂O₂ in buffer (a) and 80% TFE (b).

This is likely to be caused by oligomerisation. Chromatography results (see Figure 3.3 as well as SEC-SAXS and analytical SEC results presented in sections 5.2.6 and 5.2.7) show that C45 exists as a mixture of dimer and monomer. It is possible that the dimer has a higher catalytic activity than the monomer. As a higher proportion of dimer is likely to be present in higher concentrations of C45, this could explain the increase in $v_0/[E]_0$.

5.2.2. Mechanism of compound I formation¹

To better understand the mechanism of ABTS and peroxide turnover by C45 and the rate enhancing effect of TFE, the formation of compound I was investigated in both buffer and 80% TFE. According to the hypothesised reaction scheme for the oxidation of ABTS by C45 (**Figure 4.1**), the first step in this reaction is peroxide binding, followed by proton shuffling and elimination of water (**Figure 5.4**).

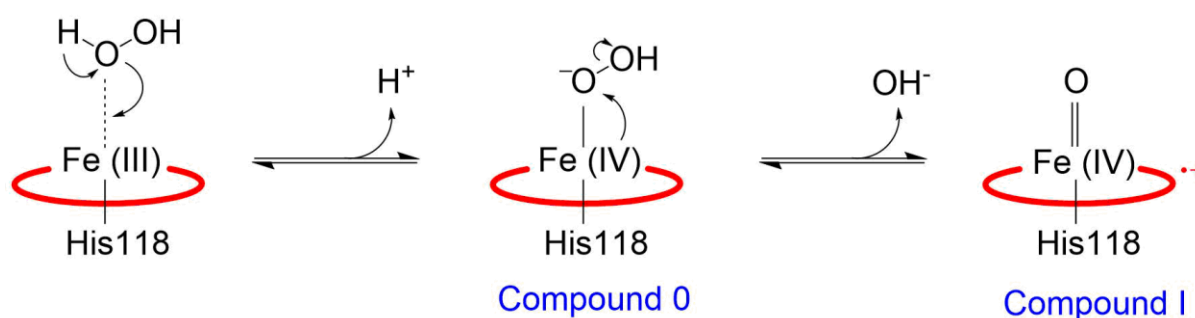


Figure 5.4: Proposed mechanism of compound I formation from C45 and hydrogen peroxide.

The reaction between $1\ \mu\text{M}$ C45 and $10\text{--}50\ \mu\text{M}$ peroxide was probed by stopped-flow to determine the rate-limiting step of the reaction – proton shuffling or water elimination. Reactions were performed with hydrogen peroxide in buffer with and without 80% D_2O (**Figure 5.5**). In addition, the effect of 80% TFE in non-deuterated buffer was probed to test the impact of the cosolvent on the reaction. In order to slow down the reaction sufficiently to observe compound I formation, kinetic assays were performed at $5\ ^\circ\text{C}$, and a dead time of 1 second was used to allow for an adequately long mixing period. Double-exponential kinetics with rates k_1 and k_2 were observed for the reaction in buffer and 80% D_2O , whereas the reaction in TFE followed triple-exponential kinetics.

¹ In collaboration with Dr Adrian Bunzel (University of Bristol, School of Biochemistry).

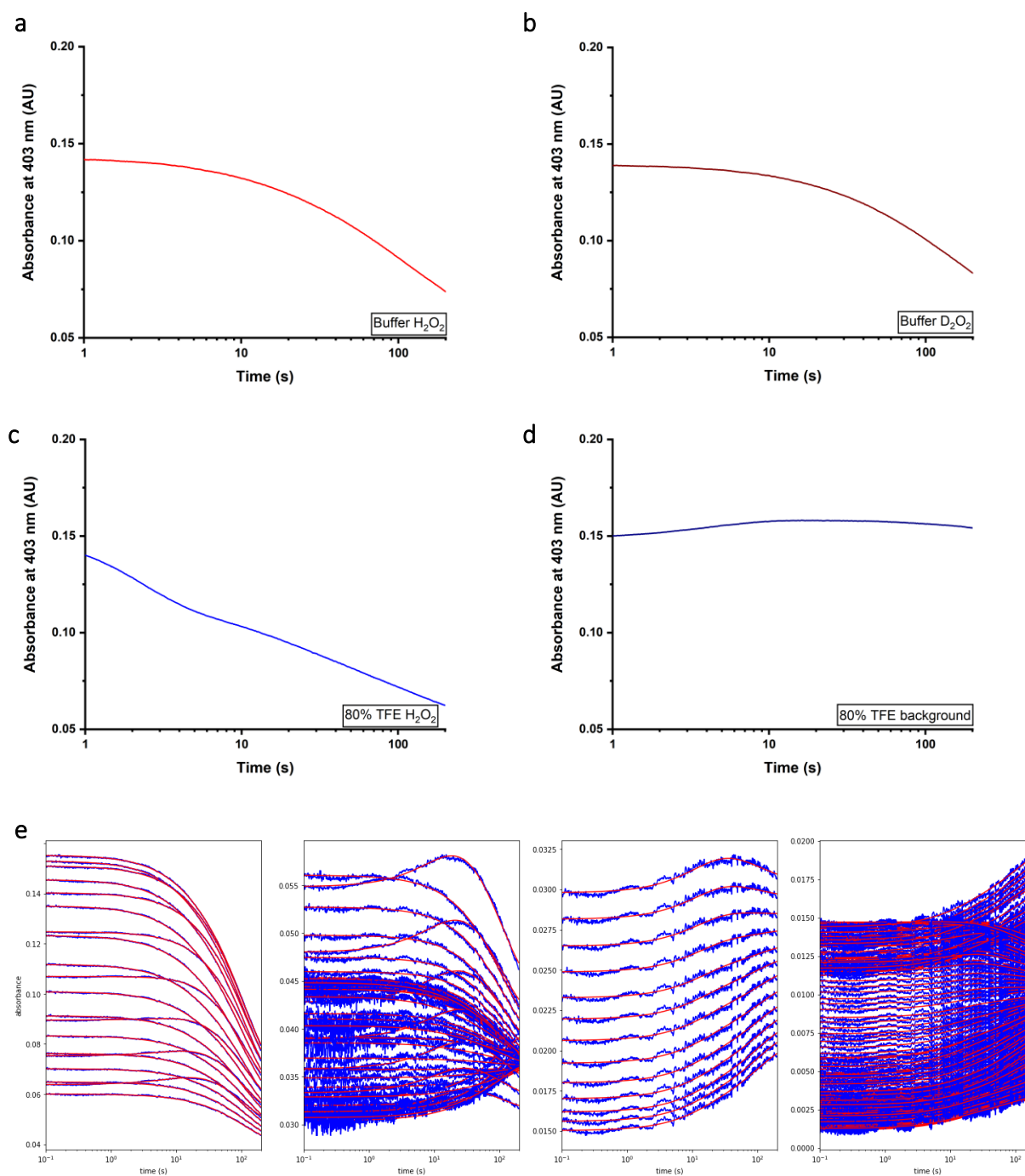


Figure 5.5: Kinetics of compound I formation.

a-d -Experimental measurement of absorbance at 403 nm over time for the reaction between 1 μM C45 and (a) 50 μM H_2O_2 in buffer, (b) 50 μM D_2O_2 in buffer, (c) 50 μM H_2O_2 in 80% TFE (d) 0 μM H_2O_2 in 80% TFE – this is an artefact attributed to the mixing of TFE and water and the accompanying rupture and formation of micelles.

e – Computational fit of absorbances measured for one experiment, ordered by initial absorbance intensity.

In the case of TFE, the additional exponential was also observed without the addition of peroxide. This process with the rate k_{bg} is likely a background artefact of TFE related to the rupture and reformation of micelles during mixing.

Since the reactions were monitored using a photodiode array, the data were globally fitted for all wavelengths between 350 nm and 600 nm with shared rate constants. For the reaction in buffer and 80% D₂O, a double exponential was used for fitting:

$$A = A_1e^{-k_1t} + A_2e^{-k_2t} + C \quad 5.2$$

In contrast, the data in 80% TFE were fitted to a triple exponential:

$$A = A_1e^{-k_1t} + A_2e^{-k_2t} + A_{bg}e^{-k_{bg}t} + C \quad 5.3$$

with k_{bg} being held constant to its previously determined value from fitting the mixing data ($0.26 \pm 0.02 \text{ s}^{-1}$) without peroxide to a single exponential:

$$A = A_{bg}e^{-k_{bg}t} + C \quad 5.4$$

Importantly, in buffer and 80% D₂O, no additional background phase was observed. The first process thus likely corresponds to compound I formation, and the second process possibly a variety of processes including solvent interactions and decomposition of compound I. **Figure 5.6** shows the two rate constants k_1 and k_2 for all three experiments.

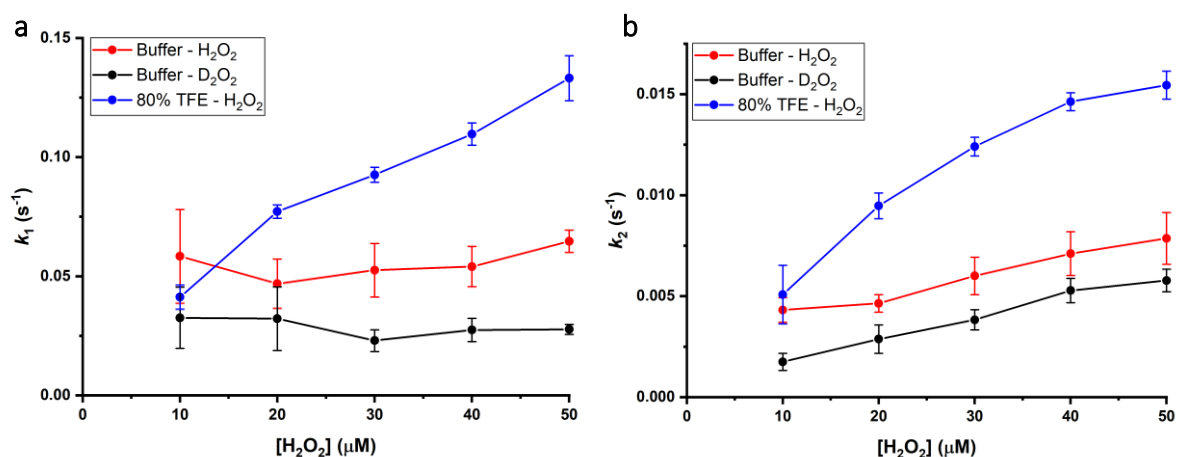


Figure 5.6: Peroxide concentration dependent rate constants for ABTS oxidation by C45 at 5 °C. **a** – k_1 , which has been assigned to compound I formation. **b** – k_2 which likely includes the degradation of compound I.

There is a clear trend for the rate to decrease when hydrogen peroxide is substituted with D₂O₂. The kinetic isotope effect was calculated at the maximum peroxide concentration of 50 μM, as this is associated with the lowest error. Values for the kinetic isotope effect for both rate constants are shown in **Figure 5.7**. For the first rate constant k_1 , this is greater than 2, whereas for the second rate constant k_2 , it is not significantly greater than 1.

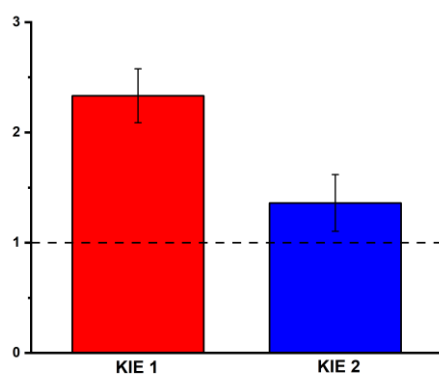


Figure 5.7: Deuterium substitution kinetic isotope effects for k_1 (KIE1) and k_2 (KIE 2).

The kinetic isotope effect on k_1 indicates that this process involves the partially rate-limiting cleavage of a proton-bond prior to compound I formation. This observation corroborates the hypothesis that proton-shuffling prior to compound I formation is rate limiting in C45. This is in line with the elevated K_M for H_2O_2 , and the absence of proton-shuffling bases such as a distal histidine often present in natural analogues (294).

The half-life $t_{1/2}$ of the formation of compound I (determined from the rate constant k_1) in H_2O_2 is 14 seconds.

$$t_{\frac{1}{2}} = \frac{\ln 2}{k} \quad 5.5$$

Unpublished data from electron paramagnetic resonance spectroscopy suggest that at 25 °C, C45 has a half-life of 1-2 seconds for compound I formation, and its decomposition has a half-life of 10-100 seconds, which fits well with this data as temperature has a strong effect on reaction rate.

Notably, a linear fit of k_1 versus peroxide concentration indicates a non-zero rate at 0 μM peroxide. This may be suggestive of a significant k_{-1} corresponding to the unbinding of peroxide. Along with the very high K_M values for hydrogen peroxide, this most likely indicates that C45 has only limited H_2O_2 affinity.

In terms of TFE, it was observed that k_1 and k_2 were significantly faster in 80% TFE, especially at the more accurately determined higher concentrations of peroxide. This supports the previous observation that activity is drastically increased in TFE, and suggests that this is due to a significant increase in k_1 , the rate of compound I formation. This is not only shown by the significantly higher k_1 , but also by the clear

positive slope observed for the linear fit of k_1 versus peroxide concentration. This slope, corresponding to $k_{\text{obs}} = 25.9 \pm 0.7 \text{ M}^{-1}\text{s}^{-1}$, is close to zero without TFE and indicates that turnover of peroxide is more efficient.

5.2.3. Steady-state kinetics of Horseradish Peroxidase with low substrate concentrations

In order to put the kinetic activity of C45 into perspective, it was compared to horseradish peroxidase (HRP) which has been subject to scientific investigation for over two centuries and is very well characterised (177). Steady-state kinetics were performed to analyse the HRP-mediated oxidation of ABTS to ABTS^{•+} in the presence of hydrogen peroxide in buffer and 80% TFE. The conditions chosen were similar to those previously used to measure activity under limiting peroxide concentrations: 25 °C, 100 μM H₂O₂ and redox buffer (20 mM CHES, 100 mM KCl, pH 8.6). The concentration of HRP used was 15-108 nM in buffer and 15-60 nM in 80% TFE.

The results (**Figure 5.8**) show that horseradish peroxidase is significantly more active in buffer than C45, with k_{cat} and total turnover number under limiting peroxide concentrations increasing 100-fold and 10-fold, respectively. Upon addition of 80% TFE, however, HRP is almost entirely deactivated and loses $\geq 99\%$ in k_{cat} , from 95 s⁻¹ to 0.066 s⁻¹. This stands in stark contrast to the 7-fold enhancement in k_{cat} exhibited by C45 when the aqueous reaction medium is replaced with 80% TFE. **Table 5.2** summarises the observed kinetic parameters for the reactions of C45 in buffer and 80% TFE. This is in agreement with the literature which states that high TFE concentration inactivates horseradish peroxidase (6).

It is not clear exactly what causes this difference in behaviour upon exposure to high TFE concentrations. However, a strongly contributing factor may be the excessive

stabilisation of helices over the native tertiary structure which has previously been observed in other enzymes (141). The native conformation of C45 is mostly helical, with connecting loops of serine and glycine, which will terminate the propagation of helices even in TFE (3). Horseradish peroxidase, in contrast, contains some β -sheets and more random coil. It is possible that the enzyme's native conformation, which supports its enzymatic function, becomes excessively helical at the cost of catalytic activity. CD could be employed to investigate this further.

Less likely is a loss of conformational flexibility associated with the addition of the TFE, which has been shown to rigidify proteins (3). This may interfere with the ability of horseradish peroxidase to bind substrates, as conformational fluctuations are important for accessibility of the active site to peroxide (295). In horseradish peroxidase, which is approximately thrice the size of C45, the active site is buried on the inside and accessible *via* a bottleneck entry (296). Therefore, if TFE reduces mobility of horseradish peroxidase, this would be detrimental to the enzyme's catalytic activity, an effect that would be less likely in C45.

5.2. Results – Steady-state kinetics of Horseradish Peroxidase

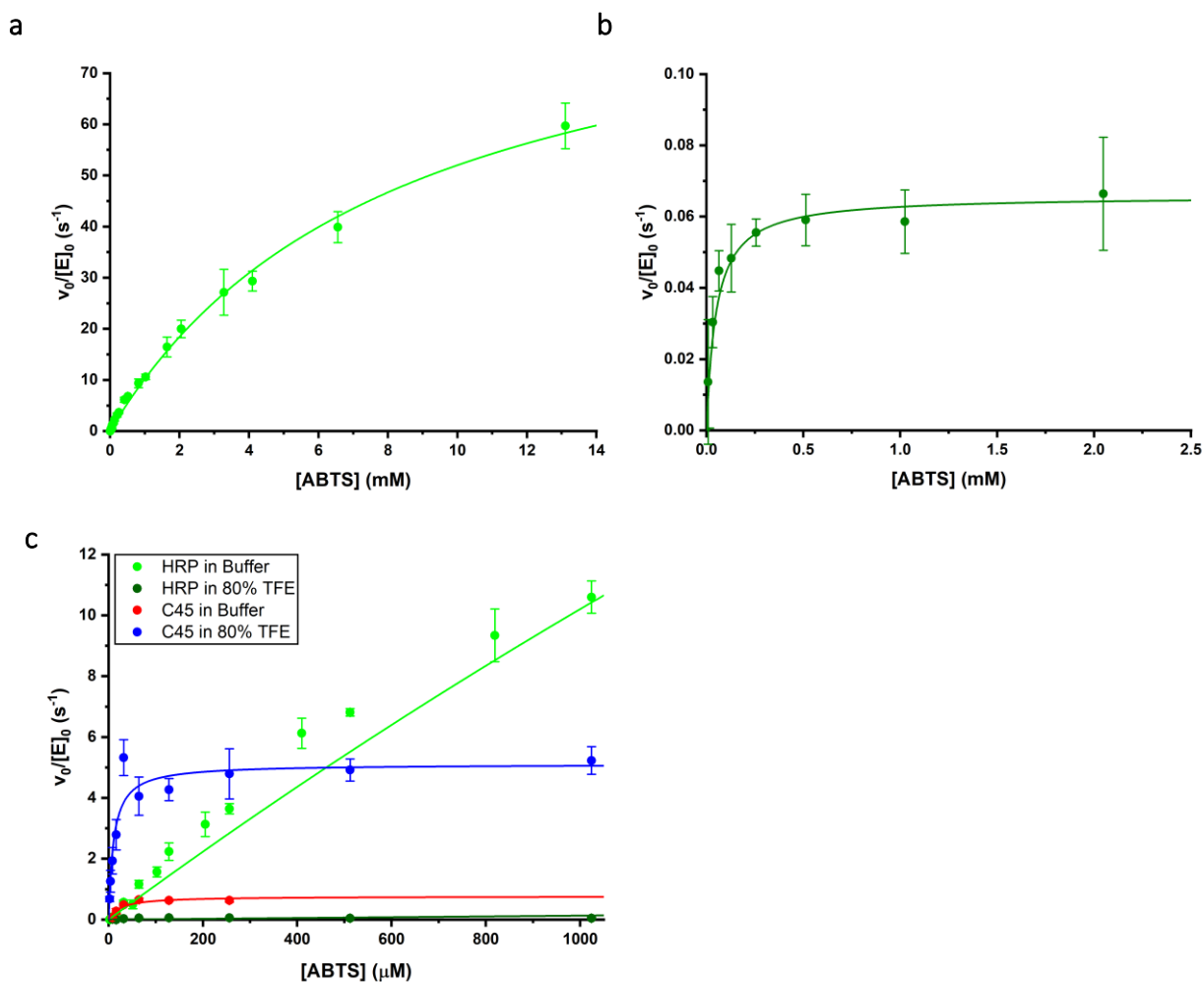


Figure 5.8: Kinetics of ABTS oxidation by HRP and C45 in the absence and presence of TFE.

a – Horseradish peroxidase, 100 μM H_2O_2 and varying ABTS in buffer (20 mM CHES, 100 mM KCl, pH 8.6).

b - Horseradish peroxidase, 100 μM H_2O_2 and varying ABTS in 80% TFE. **c** – Comparison of horseradish peroxidase and C45 with 100 μM H_2O_2 and varying ABTS in buffer and 80% TFE.

Table 5.2: Kinetic parameters of ABTS oxidation by C45 and horseradish peroxidase (HRP) in buffer and 80% TFE.

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1} \mu M^{-1}$)	TTN
C45 Buffer	0.76 ± 0.7	24 ± 7	0.033 ± 0.008	53 ± 3
C45 80% TFE	5.1 ± 0.3	10 ± 4	0.5 ± 0.2	1800 ± 200
HRP Buffer	95 ± 4	8400 ± 700	0.011 ± 0.001	606 ± 4
HRP 80% TFE	0.066 ± 0.006	50 ± 20	0.0013 ± 0.0005	29 ± 3

5.2.4. Metallocarbenoid formation by C45²

It has recently been demonstrated that C45 is a very versatile enzyme, with not just peroxidase functionality, but also the ability to act as a carbene transferase (220). A preliminary experiment was performed to confirm whether the formation of a metallocarbenoid (Figure 5.9) also proceeds in high TFE concentrations (75%). The reaction between C45 (reduced by sodium dithionite, Na₂S₂O₄) and ethyl diazoacetate (EDA, C₄H₆N₂O₂) was monitored under anaerobic conditions (Figure 5.10).

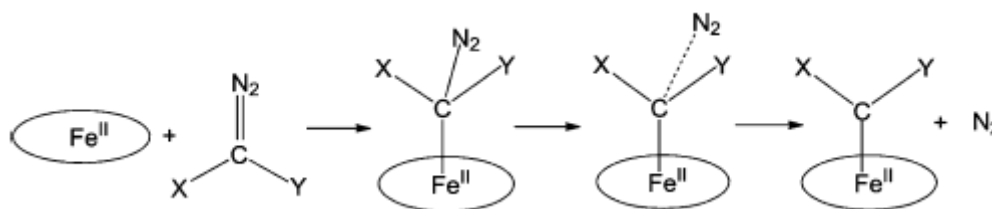


Figure 5.9: Mechanism of metallocarbenoid formation by reduced C45 (symbolised as Fe^{II}). Figure supplied by Richard Stenner.

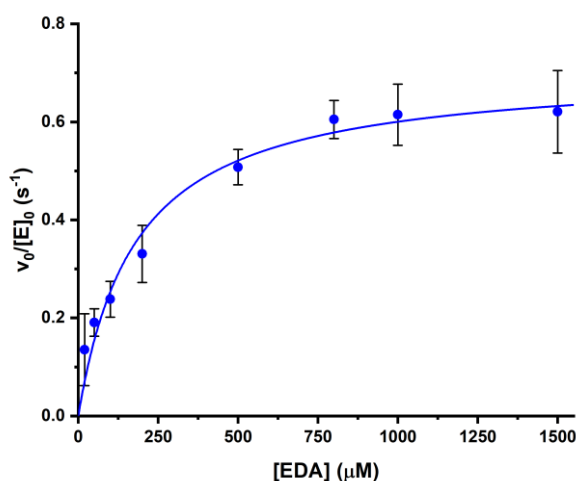


Figure 5.10: Kinetics of C45 mediated metallocarbenoid formation at increasing EDA concentration in 75% TFE.

² Performed in collaboration with Richard Stenner (University of Bristol, School of Biochemistry).

The observed limiting rate constant k_{lim} of this reaction is $0.71 \pm 0.04 \text{ s}^{-1}$, with a pseudo-Michaelis constant K_1 of $180 \pm 40 \text{ }\mu\text{M}$. The efficiency (k_{lim}/K_1) is thus $0.004 \pm 0.002 \text{ s}^{-1} \text{ }\mu\text{M}^{-1}$. In buffer, the values measured were $k_{lim} = 0.5 \text{ s}^{-1}$, $K_1 = 74$ and $k_{lim}/K_1 = 0.007$.

While this is a preliminary result that would require repeats to yield reliable results, it does suggest that TFE neither strongly improves metallocarbenoid formation nor does it hinder it. This is very different from the 7-fold rate enhancement that the cosolvent had on ABTS oxidation. These strongly different effects of TFE on reaction kinetics suggest that the cosolvent has a stabilising effect on compound I in the peroxidase reaction and the enhancement of peroxidase catalysis is not caused by the generic stabilisation of the enzyme.

5.2.5. Nuclear Magnetic Resonance (NMR) spectroscopy of C45³

To observe the effect of TFE on the protein further, NMR spectroscopy of C45 in buffer and in 80% TFE was performed. The ¹H NMR spectra (**Figure 5.11**) of C45 in buffer and 80% TFE are very comparable, suggesting the protein is not chemically denatured with the addition of TFE. The regions that the peaks are attributed to are highlighted in green. The spectrum in 80% TFE is, however, slightly better resolved, suggesting increased stability of the protein.

³ In collaboration with Dr Adrian Bunzel (University of Bristol, School of Biochemistry) and Dr Chris Williams (University of Bristol, School of Chemistry)

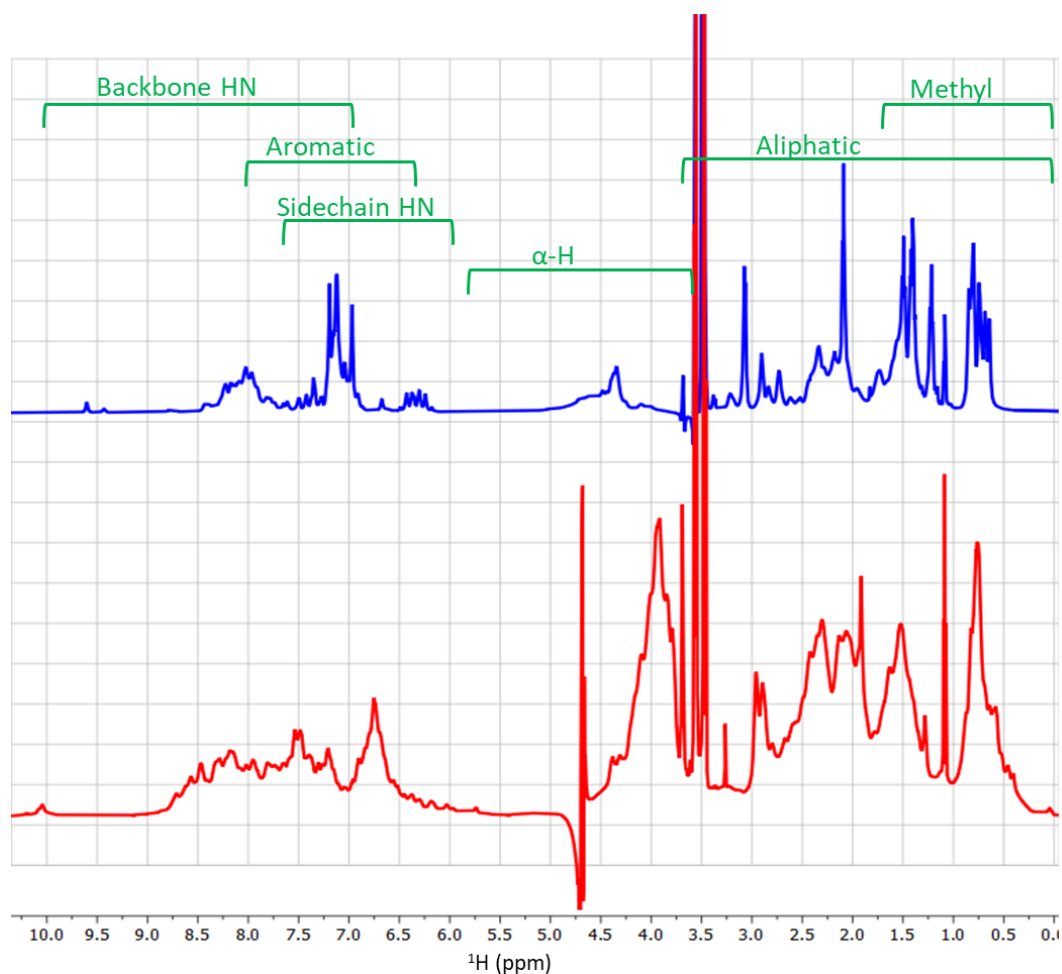


Figure 5.11: ^1H NMR spectrum of C45 in buffer (red) and 80% TFE (blue). The peak at 3.5 ppm is attributed to CHES present in the buffer.

For better resolution, 2D-NMR spectra were obtained (**Figure 5.12**). Generally, the spectra in buffer and 80% TFE appear to show the same peaks: peaks in the region characteristic for random coil serine and glycine can be identified, which correspond to the loop regions of the protein (2). Further, the peaks highlighted in green circles likely correspond to the amide groups in the side chains of glutamine and asparagine (297–299). The similarity of the spectra in buffer and 80% TFE is a good indicator that the protein structure is not significantly altered by the cosolvent. Nonetheless, signals observed for C45 in 80% TFE are slightly improved, which may indicate increased stability and rigidity of the protein which is in good agreement with the literature (3).

5.2. Results - Nuclear Magnetic Resonance (NMR) spectroscopy of C45

Other potential explanations for this result could be a lesser degree of self-association of the protein or reduced viscosity of the solution with 80% TFE. SOFAST-HMQC spectra had been obtained for C45 in buffer previously (2), with very comparable results, whereas TROSY-HSQC spectra were chosen as a complementary analysis exhibiting better resolution in different regions. Both types of 2D-spectra as well as the 1D ^1H spectrum suggest that the addition of TFE has a beneficial effect on the structural stability of the protein.

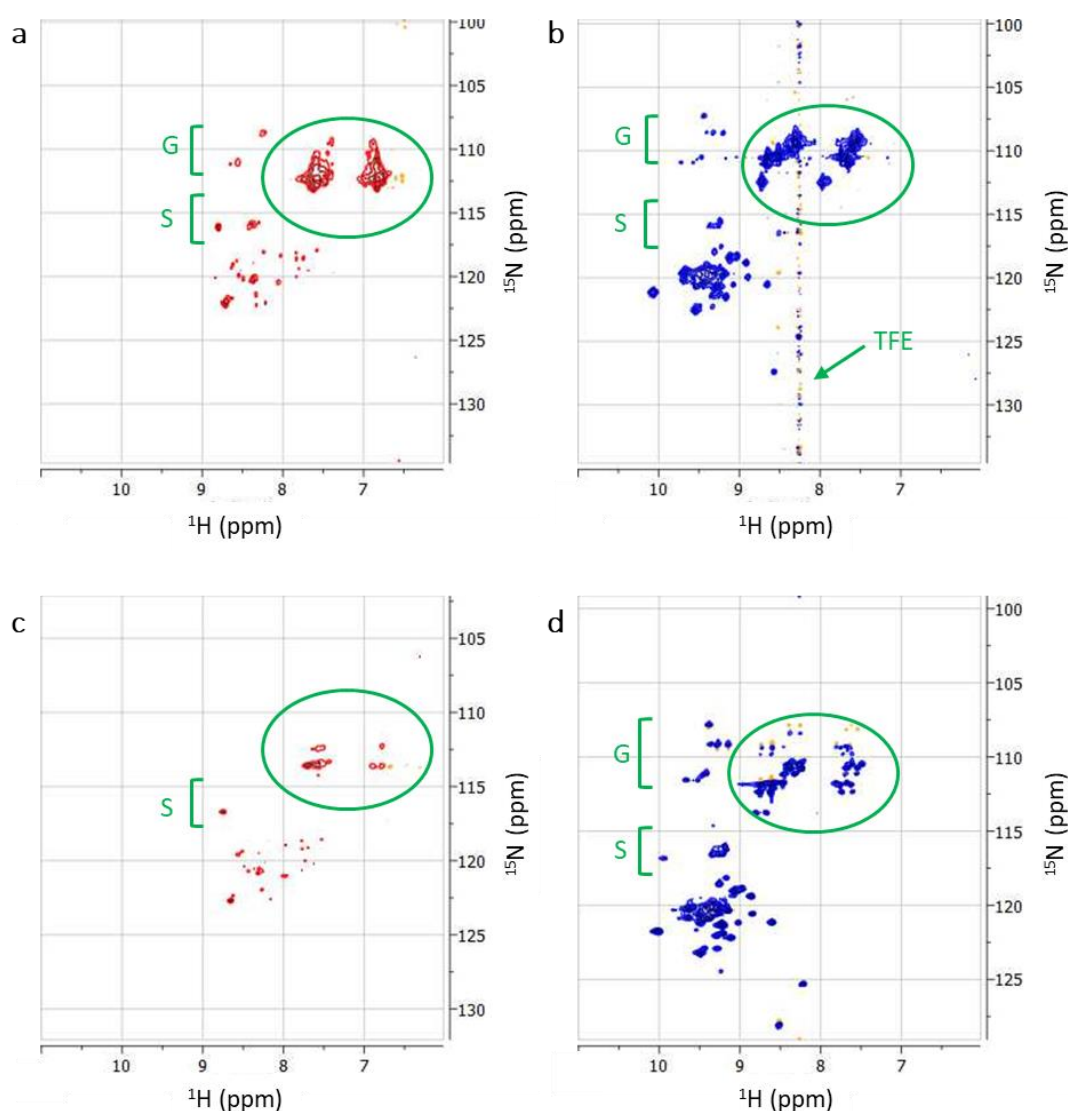


Figure 5.12: 2D NMR spectra of C45 in aqueous solution and 80% TFE.

^1H - ^{15}N SOFAST HMQC of C45 in buffer (a) and 80% TFE (b) and ^1H - ^{15}N TROSY-HSQC spectra in buffer (c) and 80% TFE (d).

5.2.6. Small Angle X-ray Scattering of C45⁴

In order to analyse the oligomerisation status of C45 both in buffer and in 80% TFE, the protein was subjected to size exclusion chromatography coupled Small Angle X-ray Scattering (SEC-SAXS). A sample of 1 mM C45 in buffer (20 mM CHES, 100 mM KCl, pH 8.6) was injected into an analytical size exclusion column and eluted in a double peak (**Figure 5.13**).

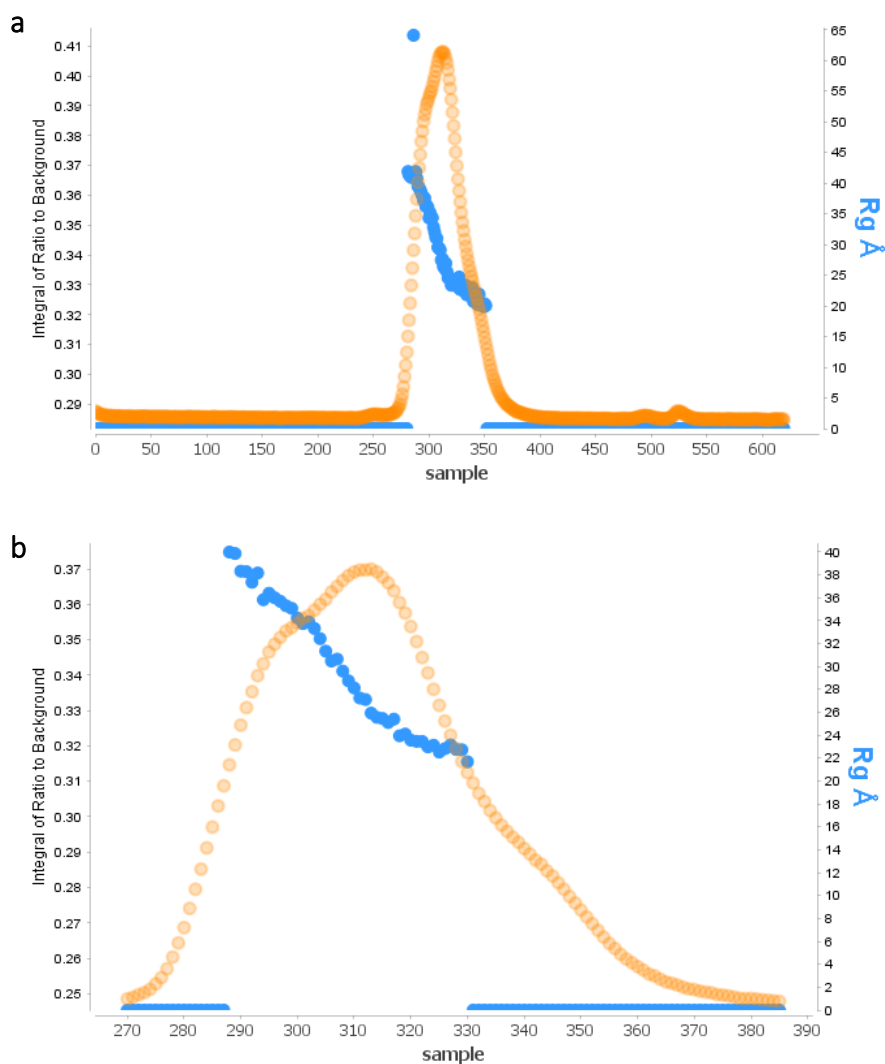


Figure 5.13: SEC-SAXS chromatogram of C45 in buffer.

Absorbance of the eluting solution at 280 nm (orange) overlaid with radius of gyration (blue) for the whole chromatogram (a) and a closer view of the elution peak of C45 (b).

⁴ In collaboration with Dr Adrian Bunzel (University of Bristol, School of Biochemistry).

While the peaks are poorly resolved, SAXS analysis of the eluting solution confirmed the presence of two species.

Scattering is presented as q , known as the scattering vector or momentum transfer, which can be calculated from the scattering angle (2θ) and beam wavelength λ using the following equation:

$$q = \frac{4\pi \sin \theta}{\lambda} \quad 5.6$$

The scattering vector and intensity I could be used to calculate a $P(r)$ distribution which describes the distance between a pair of points within a scattering particle (237, 300).

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty \frac{q^2 I(q) \sin(qr)}{qr} dq \quad 5.7$$

It can provide some information about shape of a particle as well as its radius of gyration (R_g). This parameter is a measure of particle size that is calculated as according to the following equation 5.8:

$$R_g^2 = \frac{\int_0^{D_{max}} r^2 p(r) dr}{2 \int_0^{D_{max}} p(r) dr} \quad 5.8$$

D_{max} is the maximum dimension of the particle.

ScÅtter software was used to fit a $p(r)$ distribution from the raw data and calculate radius of gyration R_g maximum dimension D_{max} . Furthermore, the Porod volume V_P was calculated by the programme.

$$V_P = \frac{2\pi^2 I_0}{Q} \quad 5.9$$

Where I_0 is the intensity in line with the beam and Q is the Porod invariant which is related to the surface to volume ratio of particles (301).

$$Q = \int_0^{\infty} [I(q) - K_4]q^2 dq \quad 5.10$$

K_4 is a constant which ensures proportionality between asymptotical intensity decay and q^{-4} at higher angles.

These calculations were performed computationally using the ATSAS programme (240, 241). Scattering curves as well as the distribution functions $p(r)$ for the two eluting peaks in buffer are shown in **Figure 5.14**.

The two species (eluting at approximately 290 ml and 325 ml) were calculated to have radii of gyration of 36 and 23 Å, respectively, with Porod volumes of 85,000 and 40,000 Å³, respectively. This corresponds roughly to a size ratio of 2:1, confirming the presence of a dimer and monomer. Furthermore, the shape of the curve suggests that the monomer is rod-shaped. The dimer is likely also somewhat oblong, but less cylindrical than the monomer. **Figure 5.15** shows the shapes density distribution function graphs typically adopt for differently shaped particle in small-angle scattering.

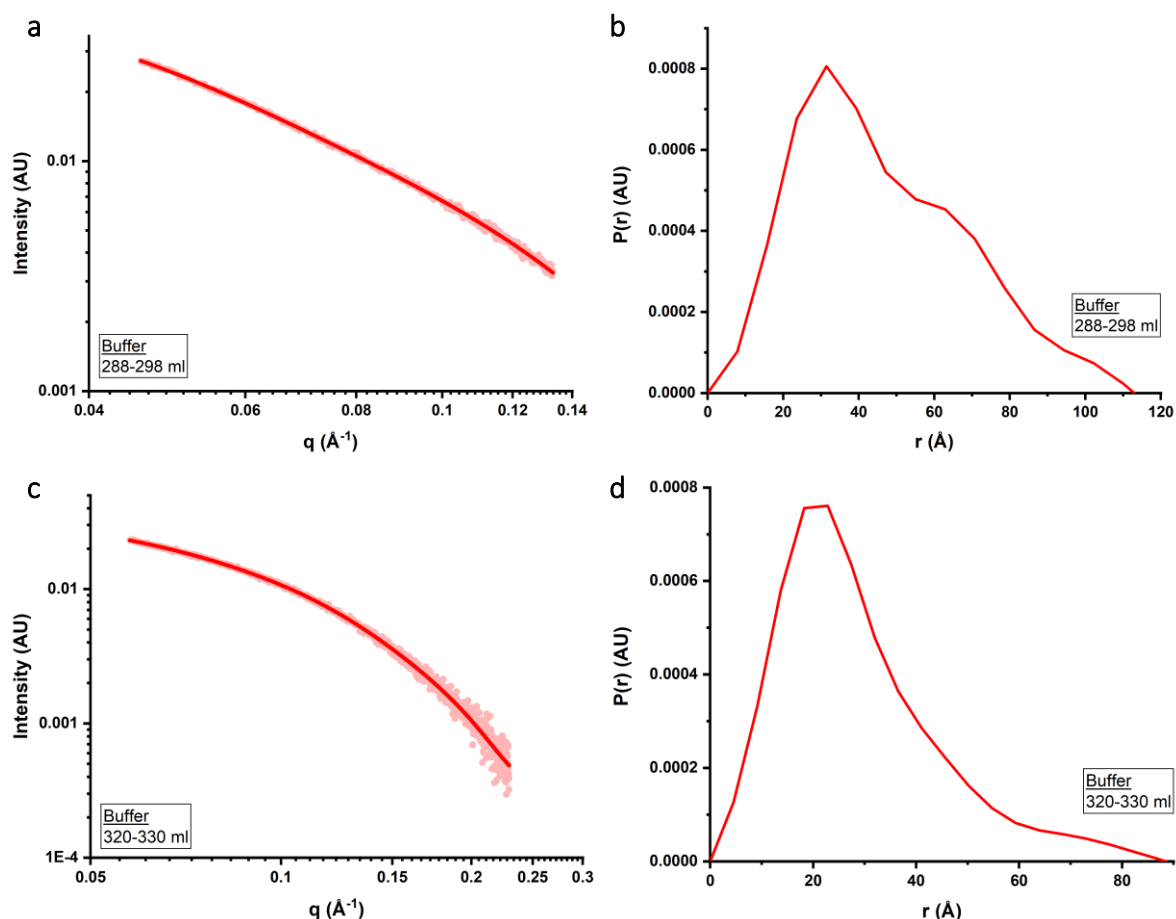


Figure 5.14: SAXS scattering curves and $P(r)$ distributions for C45 in buffer.

a, c – scattering curves of the solution eluting between 288-298 and 320-330 ml respectively. Light red dots show data points and the red line is the best fit line. **b, d** – $P(r)$ distributions of those species.

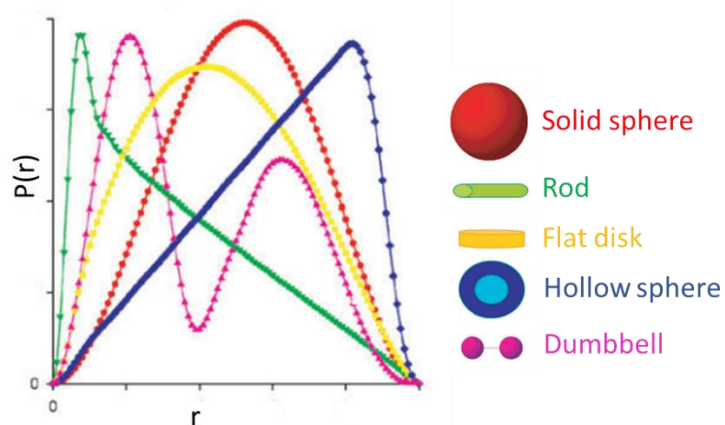


Figure 5.15: Typical shapes of density distribution functions for different particle sizes.

Colours of spectra correspond to colours of different particle shapes shown alongside. Figure taken from (302).

Scattering profiles were fitted to produce 3-dimensional models using SASREF and DAMMIF software by EMBL Hamburg (242). Three-dimensional models (**Figure 5.16**) suggest more complex shapes than predicted by the $P(r)$ distribution. Both resemble a slightly elongated cylinder with a small adjoined lobe, although of course the dimer is larger in size and more rod-like. The computationally produced model of C45 is shown for scale and comparison, it is notably more elongated and significantly smaller. This suggests that the protein might fold less tightly than expected, but also casts doubt on the accuracy of the SAXS models and highlights the importance of repeat experiments.

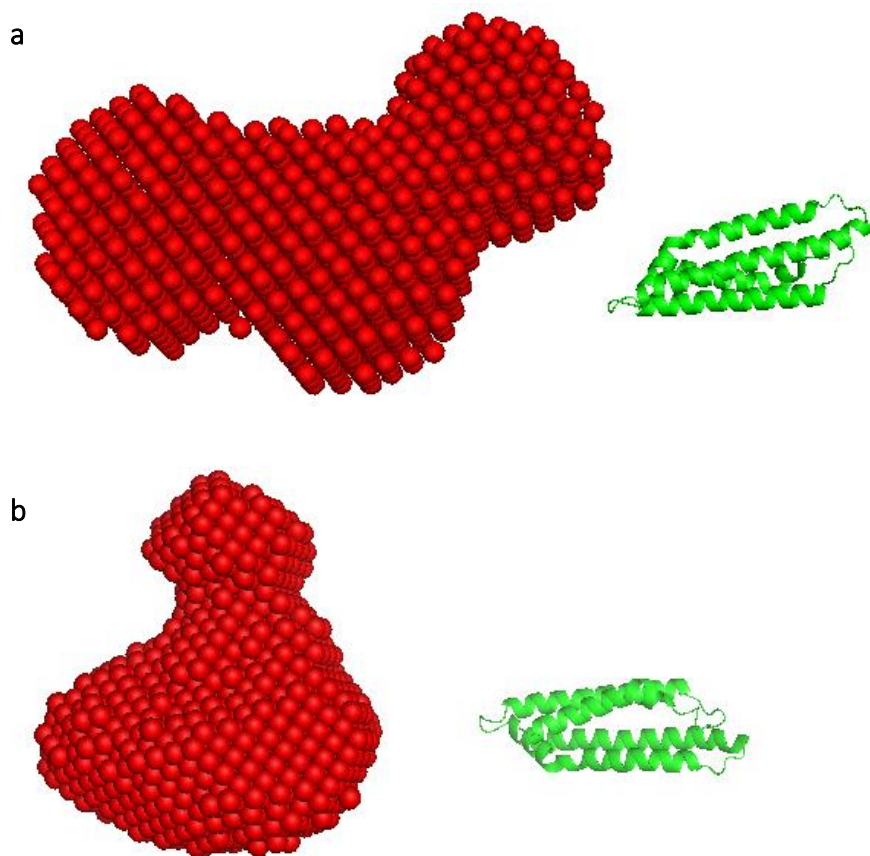


Figure 5.16: 3D models of C45 in buffer produced from scattering profiles.

Models of the predicted protein structure (2) shown for comparison.

a—suspected dimer (eluting at 288-298 ml) resolved to 40 Å. **b**—suspected monomer (eluting at 320-330 ml) resolved to 31 Å.

C45 in 80% TFE (at a concentration of 0.7 mM) was also subjected to SEC-SAXS and eluted as two species, poorly separated by chromatography (**Figure 5.17**).

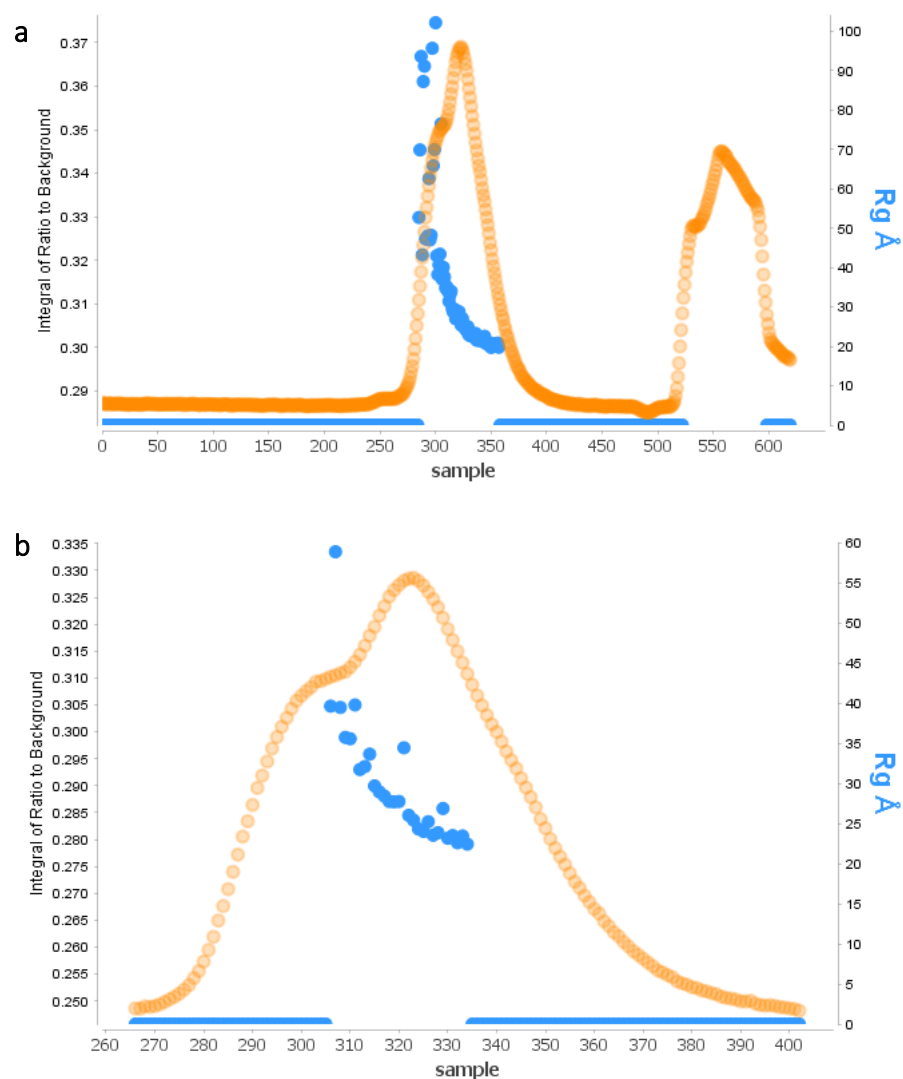


Figure 5.17: SEC-SAXS chromatogram of C45 in 80% TFE.

Absorbance of the eluting solution measured at 280 nm (orange) overlaid with radius of gyration (blue) for the whole chromatogram (**a**), including elution of TFE and re-equilibration of the column into buffer as seen at around 550 ml, and a closer view of the elution peak of C45 (**b**).

The two species eluting at approximately 295 ml and 325 ml were predicted to have a volume of 87,000 and 34,000 Å³, respectively, and a radius of gyration of 38 and 25 Å. The D_{\max} values were determined as 119 and 93 Å, respectively. The predicted radii of gyration, D_{\max} values and shape of the chromatogram are very similar

to those of C45 in buffer, suggesting that the species present here are also dimer and monomer. Scattering curves and density distribution functions for the two species of C45 in 80% TFE are presented in **Figure 5.18**.

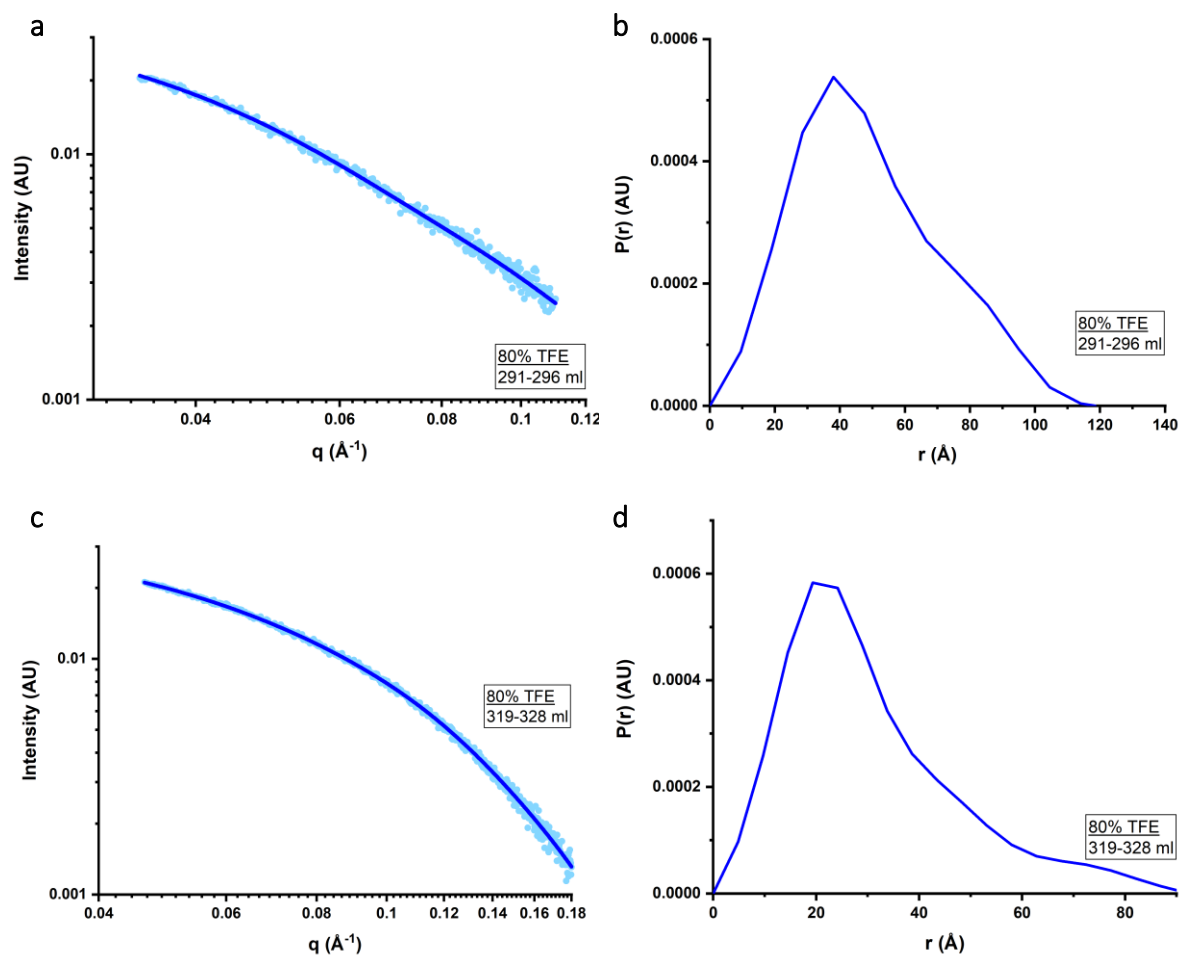


Figure 5.18: SAXS scattering curves and $P(r)$ distributions for C45 in 80% TFE.

a, c – scattering curves of the solution eluting between 291-296 and 319-328 ml respectively. Light blue dots show data points and the blue line is the best fit line. **b, d** – $P(r)$ distributions of those species.

The shapes of the density distributions are comparable to those observed of C45 in buffer, suggesting that like in buffer, in 80% TFE the C45 monomer is rod-shaped whereas the dimer is elongated, but not perfectly cylindrical. In the 3-dimensional models (**Figure 5.19**), the monomer appears to have a similar shape as in buffer, with the dimer being less elongated than in buffer. Both monomer and dimer appear to resemble a molten globule rather than rod-shaped protein. As with the models in

buffer, PyMol models are shown alongside for comparison, and similarly, both shape and size differ between computationally predicted structure and SAXS 3D models. To ensure accuracy of the SAXS models, further experiments would be required, ideally including x-ray crystallography.

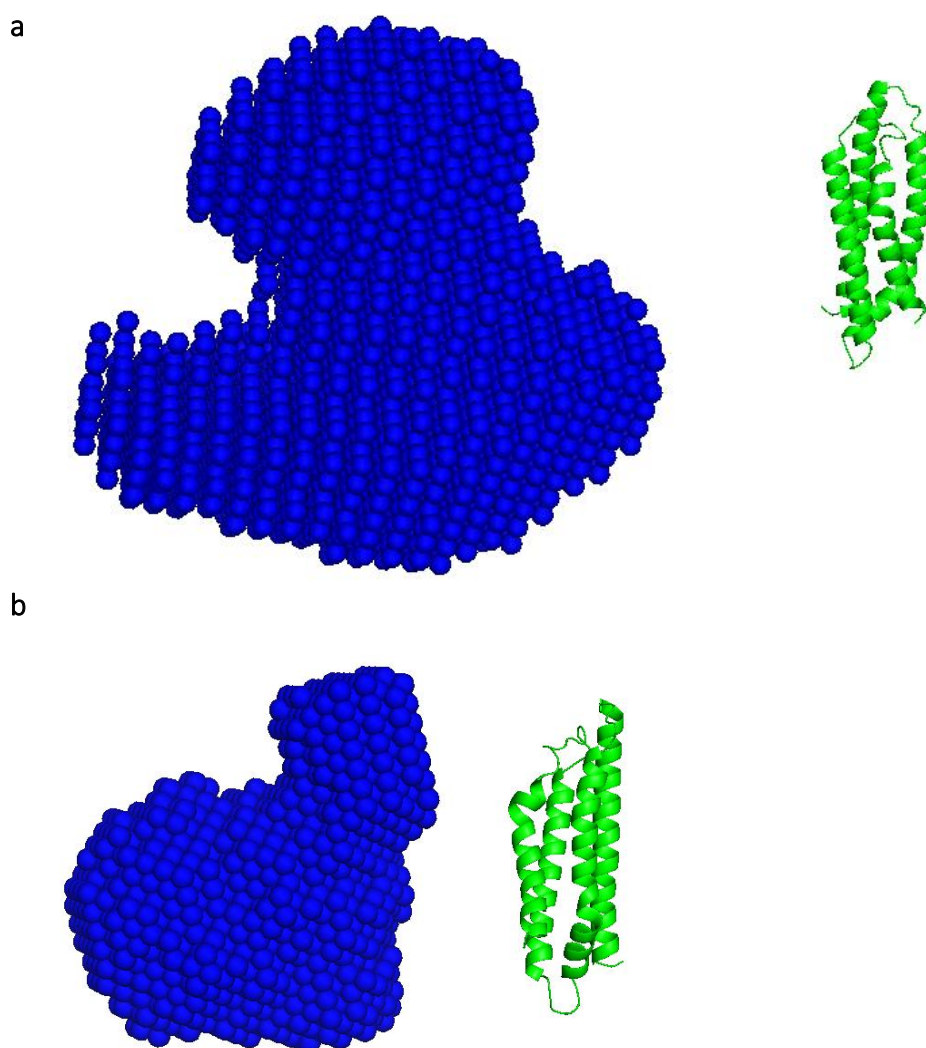


Figure 5.19: 3D models of C45 in 80% TFE produced from scattering profiles.

Models of the predicted protein structure (2) shown for comparison.

a– suspected dimer (eluting at 291-296 ml) resolved to 62 Å. **b**- suspected monomer (eluting at 319-328 ml) resolved to 32 Å.

The bell-shaped normalised Kratky plots (**Figure 5.20**) indicate that both monomer and dimer are folded in both environments (buffer and 80% TFE) (303).

5. Further kinetic and structural analysis of C45 in TFE

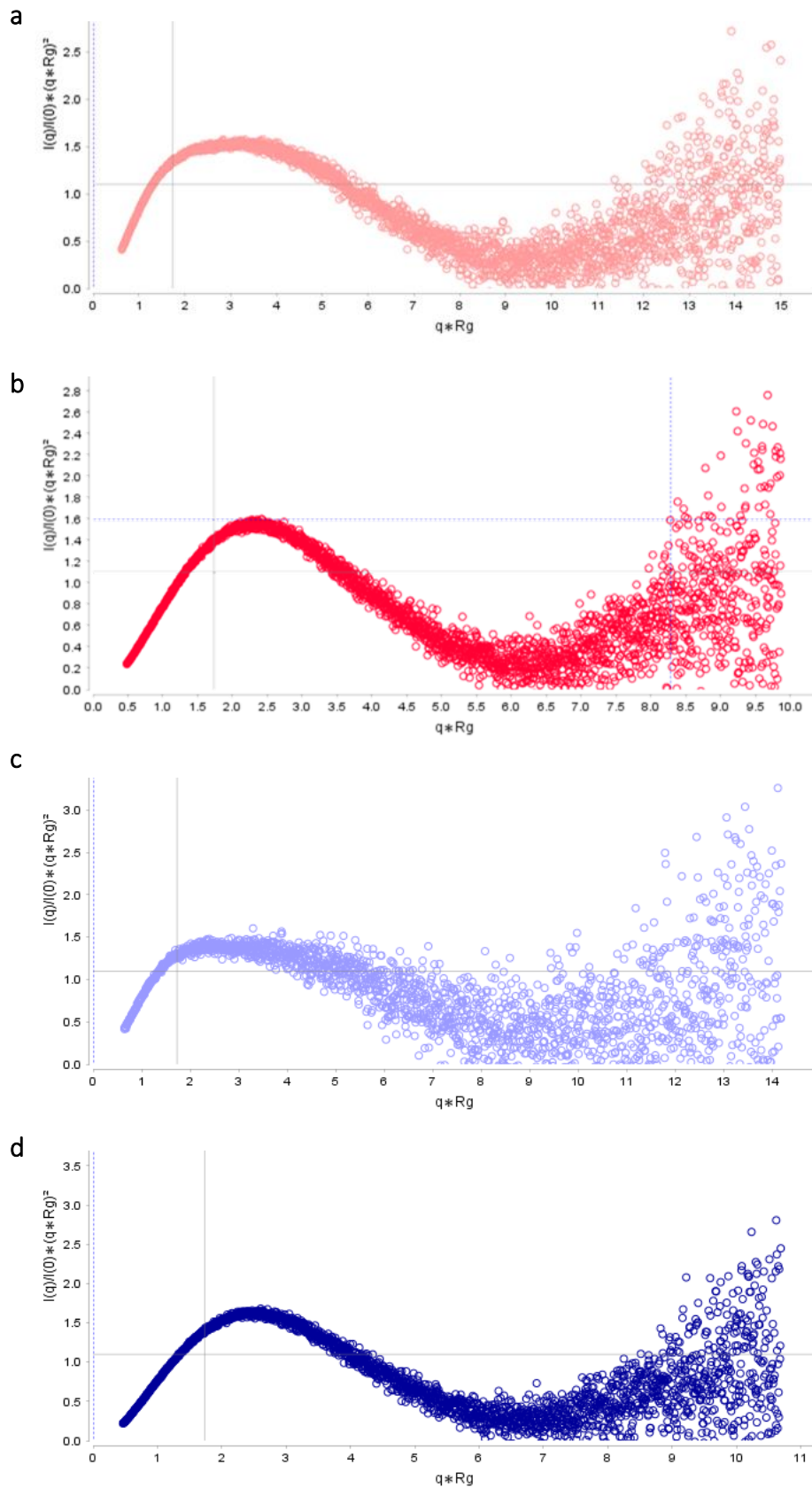


Figure 5.20: Normalised Kratky plots for C45 in buffer and 80% TFE. Suspected dimer in buffer (a) and 80% TFE (c), suspected monomer in buffer (b) and 80% TFE (d).

A summary of the calculated parameters of the species present in C45 solutions in buffer and 80% TFE is provided in **Table 5.3**. It shows that in buffer and in 80% TFE, two species elute, which are most likely a monomer and a dimer in each solution. This suggests that in 80% TFE, the quaternary structure of C45 is not significantly affected by the cosolvent at the timescales and concentrations analysed. It is, nonetheless, possible that over prolonged incubation times, the cosolvent stabilises the dimer over the monomer or *vice versa*.

Table 5.3: Summary of calculated parameters of the eluting species in SEC-SAXS of C45 in buffer or 80% TFE.

	Elution vol (ml)	V_P (Å ³)	D_{max} (Å)	R_g (Å)	χ^2
Buffer	288-298	85000	113	36	0.27
Buffer	320-330	40000	89	23	0.14
80%TFE	291-296	87000	119	38	0.19
80% TFE	319-328	34000	93	25	0.57

5.2.7. Analytical Size Exclusion Chromatography of C45

In order to analyse the behaviour of the monomer and dimer of C45 in buffer and in 80% TFE over prolonged periods of incubation, the protein was subjected to analytical size exclusion chromatography, with absorbance monitored at both 280 nm and 405 nm. A poorly resolved double peak eluted that was split into two fractions (**Figure 5.21 a**).

This was suspected to be C45 in dimeric and monomeric form. Absorbance at 405 nm was more intense, but absorbance at 280 nm was better resolved, so the latter was used to separate between the two peaks. Fractions were acquired of the shoulder regions of each peak to limit crossover. Immediately after elution, the collected fractions were diluted to a concentration of 2 μ M in buffer and 80% TFE and analysed using size exclusion chromatography once more (**Figure 5.21 b**). This analysis was

5. Further kinetic and structural analysis of C45 in TFE

repeated for each of the four samples at 12, 24, 48 and 72 hours after elution and separation (**Figure 5.22**).

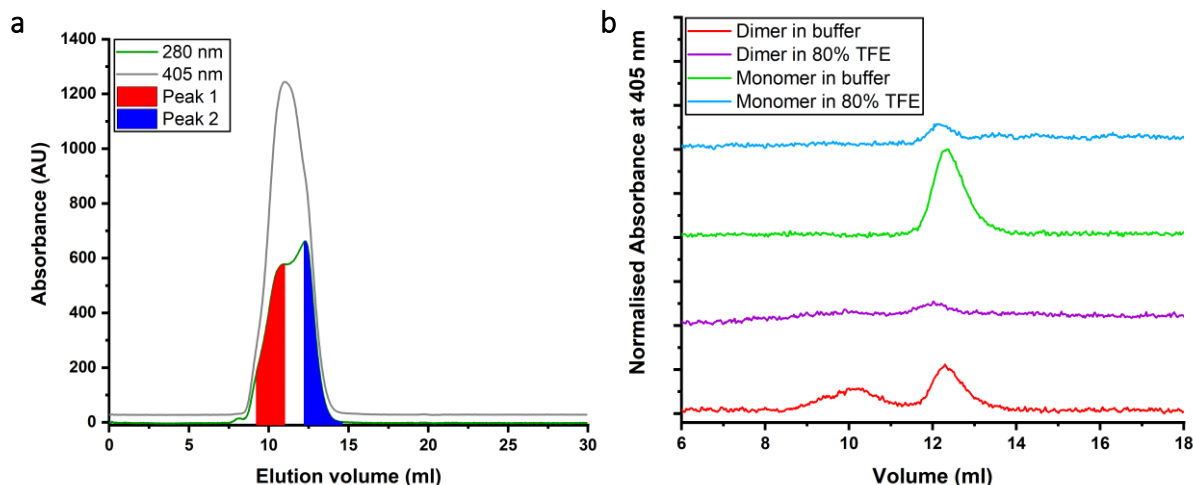


Figure 5.21: Analytical size exclusion chromatogram of C45.

a - A poorly resolved double peak elutes, with two fractions obtained (red shows the suspected dimer and blue the suspected monomer). **b** - Chromatogram of fractions from (a) diluted to 2 μ M in buffer or 80% TFE subjected to further SEC.

The subsequent chromatograms of the second fraction in both buffer and 80% TFE all exhibited a single peak with comparable intensity, indicating that the monomer is stable at room temperature in these conditions and does not spontaneously dimerise. The other fraction obtained, which was suspected to be dimeric C45, produced two peaks on subsequent chromatograms, suggesting the presence of dimer and monomer. While some of that is likely owed to the poor resolution of dimer and monomer in the initial experiment (**Figure 5.21 a**), further analysis of the two peaks showed the ratio of them to be changing with incubation time.

5.2. Results - Analytical Size Exclusion Chromatography of C45

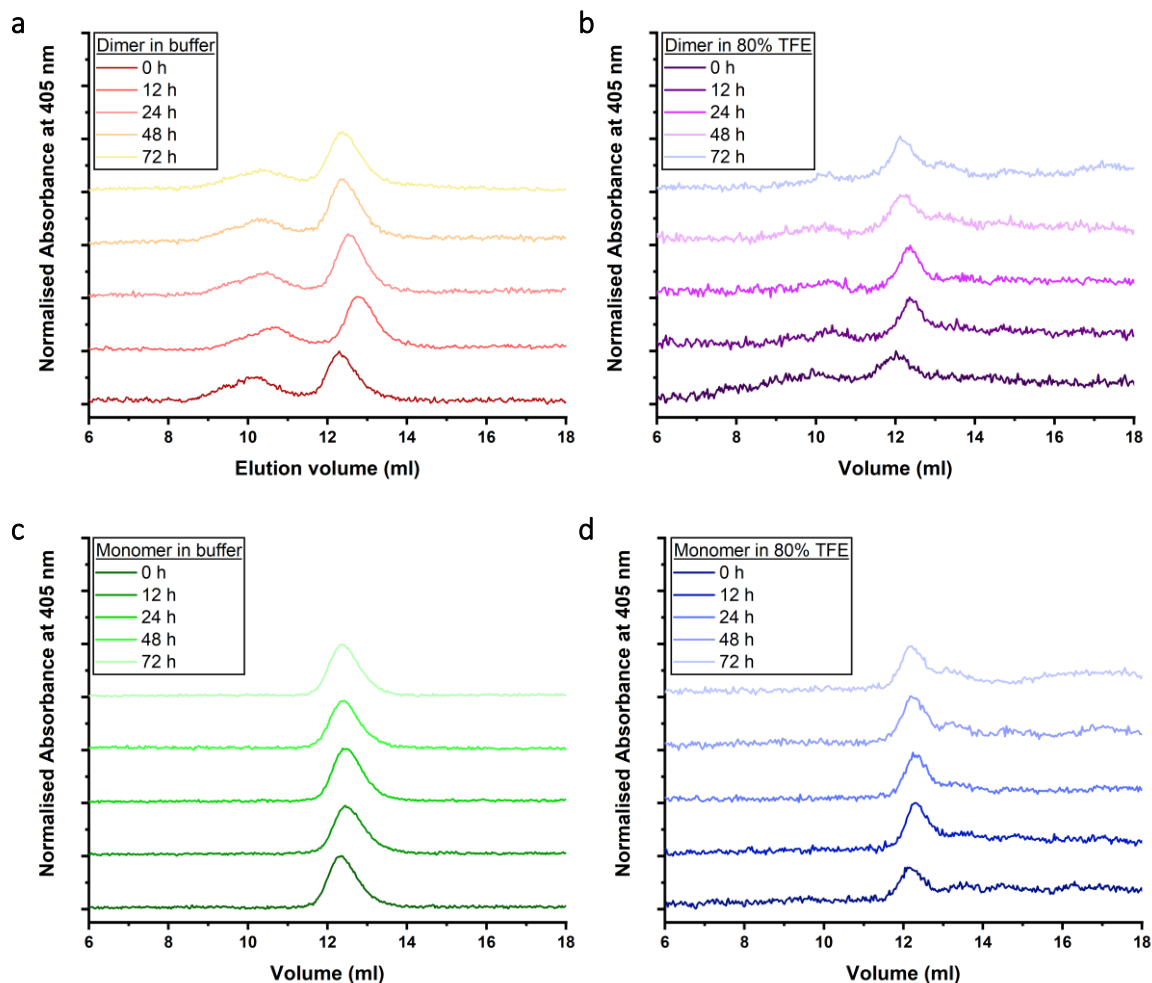


Figure 5.22: Analytical size exclusion chromatograms for C45 in buffer and 80% TFE.

a, b – Suspected dimer (fraction from peak 1 from initial analytical SEC) in buffer (**a**) and 80% TFE (**b**) at 0-72 hours after elution. **c, d** - Suspected monomer (fraction from peak 2 from initial analytical SEC) in buffer (**c**) and 80% TFE (**d**) at 0-72 hours after elution.

Absorbance was measured at 405 nm and 280 nm; however, spectra are only shown for 405 nm where absorbance was stronger.

For each subsequent run of the dimer fraction, the two maxima were calculated both at 280 nm and 405 nm and the minimum absorbance at the start of the run subtracted. Due to the high levels of noise, averages of at least six data points were taken for determination of maximum and minimum absorbance. The ratio of dimer peak to monomer peak decreases with incubation time, suggesting the dimer dissociates into monomers (**Figure 5.23**). This proportional decrease in dimeric C45

5. Further kinetic and structural analysis of C45 in TFE

is consistent between buffer and 80% TFE, suggesting the monomeric form is favourable at this protein concentration in both solutions. However, data were very noisy and the experiment would require repeats conducted with varying protein concentrations before being viewed as any more than a qualitative, preliminary result.

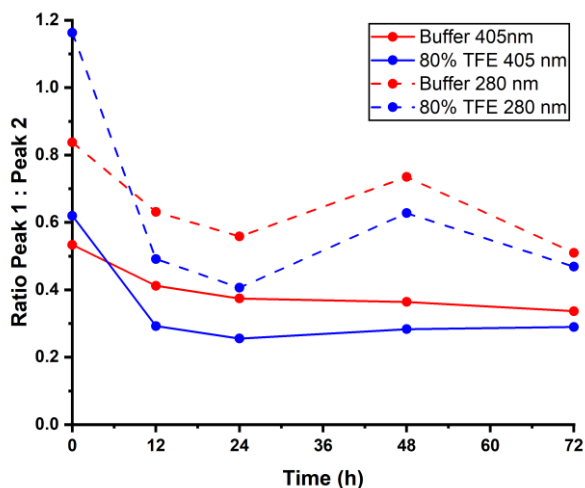


Figure 5.23: Ratio of first eluting peak to second eluting peak for C45 in buffer (red) and 80% TFE (blue).

Additionally, a single kinetic measurement of both the dimer and monomer was performed immediately after separation of the two fractions (**Figure 5.24**). The reaction between 500 nM C45, 100 μM H_2O_2 and 256 μM ABTS in buffer yielded $v_0/[E]_0$ values of 0.6 s^{-1} for the dimer peak and 1.2 s^{-1} for the monomer. This would suggest that the dimer has less catalytic activity than the monomer. This is different from the results showing the concentration dependence of $v_0/[E]_0$ on the enzyme concentration (**Figure 5.3**), which showed higher concentrations of C45 to be disproportionately more active, suggesting the dimer is more catalytically active than the monomer. This clearly needs more investigation, however, as the monomer and dimer were extremely poorly separated, no repeats of this were performed and only one measurement at a single set of concentrations was taken, this result should not be overinterpreted.

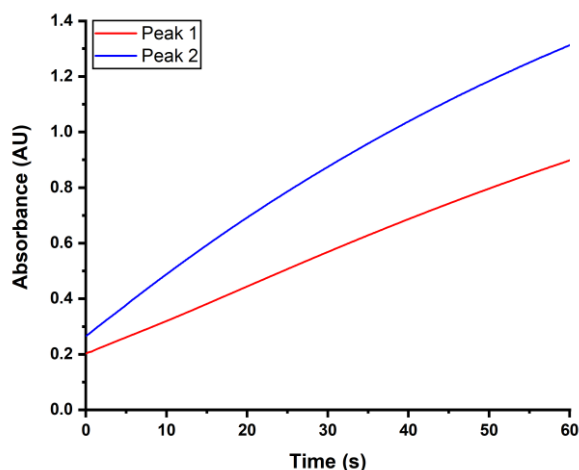


Figure 5.24: Rate of ABTS oxidation by the suspected dimer and monomer of C45.

Reaction between 500 nM of C45 from peak 1 (suspected dimer) and peak 2 (suspected monomer) from size exclusion chromatography, 100 μM H_2O_2 and 256 μM ABTS in aqueous solution.

5.3. Discussion

In the previous chapters, the generally observed trend was a decrease in structural stability and especially catalytic activity of C45 upon the addition of cosolvents. The notable exception was TFE, which yielded a significant increase in catalytic activity of C45, especially at the highest cosolvent concentration measured (80%). In this chapter, this intriguing behaviour is further probed.

A 4-fold increase in $k_{\text{cat}}/K_{\text{Ma}}/K_{\text{Mb}}$ was estimated (from $5.8 \times 10^6 \text{ M}^{-1} \text{ M}^{-1} \text{ s}^{-1}$ in buffer to $2.1 \times 10^7 \text{ M}^{-1} \text{ M}^{-1} \text{ s}^{-1}$ in 80% TFE). The k_{cat} measured in buffer was $890 \pm 60 \text{ s}^{-1}$, with a K_{M} of $220 \pm 30 \text{ mM}$ for peroxide. This was not in agreement with literature, where a k_{cat} of 1200 s^{-1} had been determined. Experimental conditions were very similar, the only discernible difference being the higher enzyme concentration used in the literature research. The good reproducibility of kinetic parameters for C45 between different batches ruled out batch to batch variation as a plausible explanation for this observation and the concentration dependence of this reaction was probed further.

Surprisingly, enzyme concentration was shown to have a significant impact on $v_0/[E]_0$, a value which is inherently normalised for enzyme concentration and therefore usually not affected by this. Here, however, it was shown that a higher enzyme concentration yielded higher reaction rates both in buffer and in 80% TFE. With this in mind, the TFE-induced 7-fold increase in k_{cat} at limiting peroxide concentration measured in chapter 4 may actually be higher, as this was measured with lower enzyme concentrations to allow for the plate reader's sensitivity limits.

The reason for this effect was proposed to be a difference in catalytic activity between the monomer and dimer of C45. SEC-SAXS and analytical SEC results have confirmed the presence of both species in C45. This does not seem to be affected by the presence of TFE. Curiously, when the kinetics of monomer and dimer were analysed after separation by analytical SEC, the monomer appeared to produce a reaction rate twice that of the dimer. This does not agree with the observation that in higher C45 concentrations, which are likely to shift the monomer-dimer equilibrium closer to the dimer, the highest $v_0/[E]_0$ values were observed. However, resolution of monomer and dimer peak by analytical SEC was very poor and no repeat measurements for this were performed. This raises suspicions that the observed increased catalytic activity of the monomer may be an erroneous measurement.

Furthermore, the mechanism of compound I formation was investigated. As ABTS is not involved in this step, the reaction between peroxide and C45 was performed both in 80% TFE and in buffer and fit to a double exponential (or in the case of TFE, triple exponential, due to the mixing of TFE with water). Kinetic isotope effects as well as comparison with unpublished EPR spectroscopy results confirmed that the first observed reaction phase was the formation of compound I, which involves breakage of an O-H bond. Especially due to the lack of a distal histidine in C45, which is common

in natural peroxidases and assists in this proton shuffling process, this step is likely to be rate-determining in the catalytic cycle of C45. In 80% TFE, this was shown to be significantly faster, suggesting that the cosolvent stabilises compound I.

The lack of an effect of TFE on the formation of a metallocarbenoid supports the hypothesis that the catalytic improvements seen in peroxidase kinetics were related to stabilisation of an intermediate of that reaction mechanism as opposed to stabilisation of C45.

Initially the inactivation of HRP by TFE concentrations above 30% may appear to be in conflict with the hypothesised stabilisation of compound I by TFE. However, HRP has a more complex structure (**Figure 1.14**) than the almost exclusively α -helical C45. Non-helical structural elements may be perturbed by the propensity of TFE to increase helicity. Furthermore, HRP requires a greater degree of conformational flexibility to bind hydrogen peroxide. Therefore, the hypothesis that TFE stabilises compound I is not invalidated by the inactivation of HRP in high concentrations of the cosolvent.

This chapter built on the previously observed increase in structural stability and catalytic activity of C45 in 80% TFE and explored the cosolvent's effect further. Both effects could be confirmed and it appears that the primary cause for the increase in catalytic activity is not the observed stabilisation of the enzyme's native state – which was in agreement with the literature (3, 5) – but the stabilisation of the radical intermediate compound I.

6. Conclusions

The field of enzyme engineering has picked up momentum over the past decade, with tailor-made proteins making up ground on their highly evolved natural counterparts (304, 305). Many enzymes have been created by directed evolution of their natural counterparts. This approach, however, limits the proportion of sequence space open to exploration drastically (1). *De novo* protein design, which has also made impressive advances in recent years (1, 306), allows the researcher more freedom. The maquette method has brought about some *de novo* enzymes with a range of functionalities (2, 19, 21, 220). This method involves the rational design of simplistic four-helix bundle proteins that are complemented with a cofactor. C45, a proficient peroxidase that is highly thermostable and has recently also been shown to be able to act as a carbene transferase (2, 220), has been designed this way. Little is known, however, about the enzyme's solvent tolerance. In perspective of the wide interest in non-aqueous enzymology (7), this work addresses the question whether the versatility of maquette enzymes in general, and C45 in particular, is also mirrored in its ability to carry out catalysis in organic cosolvents.

In this work, it has been shown that C45 remains active and stable in some organic cosolvents: In the four alcohols tested, including methanol, ethanol, isopropanol and the halogenated TFE, as well as acetonitrile, C45 maintained some catalytic activity in up to 20% isopropanol, 60% acetonitrile, 70% ethanol and 80% methanol. Generally, increasing cosolvent concentrations, and the concomitant increase in hydrophobicity, led to decreases in catalytic activity. Surprisingly, addition of TFE followed a different

trend and raised the k_{cat} of the reaction between C45, ABTS and hydrogen peroxide (under limiting peroxide concentration) by up to 7-fold at 80% TFE, the highest cosolvent concentration tested. Furthermore, the total turnover number was enhanced by a factor of 34.

It was particularly surprising that the largest increase in enzymatic activity was observed at such a high TFE concentration. TFE had previously been shown to improve the catalytic activity of peroxidases (6), but only at lower concentrations. Generally, TFE has the ability to stabilise and increase the helicity of proteins (3, 4), which is particularly relevant to C45. This effect is most likely caused by enhanced intra-protein hydrogen bonds, which can particularly stabilise helical folds (5). In agreement with this, CD and NMR spectroscopy did suggest an increase in helicity in C45 after addition of TFE. Nonetheless, the ability of TFE to increase helix stability of proteins (3) is not necessarily beneficial to all enzymes. Enzymes which are not predominantly α -helical, such as horseradish peroxidase, may experience a decrease in catalytic activity in high concentrations of TFE which was clearly observed here.

The positive effect of TFE on catalytic activity has mostly been attributed to the stabilisation of the highly reactive compound I, the formation of which proceeds significantly faster in TFE. The large increase in total turnover numbers in 80% TFE adds weight to the hypothesis that TFE stabilises compound I. This is further corroborated by the lack of any discernible effect of TFE on the rate of metallocarbenoid formation. This further, recently discovered, function of C45 does not involve peroxide mediated compound I formation (220). The observation that the presence of TFE did not affect the rate of this suggests that the main effect of TFE on the peroxidase kinetics of C45 is related to the stabilisation of compound I, rather than a generic structural effect on the enzyme.

6. Conclusions

Furthermore, it was discovered that both in 80% TFE and in buffer, $v_0/[E]_0$ was surprisingly not independent of the enzyme concentrations, as is usually the case. Activity was larger at higher enzyme concentrations. This likely suggests that oligomerisation of the enzyme, which is generally concentration dependent, has an effect on catalytic activity. Analytical SEC and SEC-SAXS results both confirm the presence of a dimer and monomer in C45 solutions. Thus, the increase in $v_0/[E]_0$ with increasing C45 concentrations suggests that the dimeric form of the enzyme may have a higher catalytic activity. It is worth noting that the presence of dimer and monomer was demonstrated both in buffer and 80% TFE and the cosolvent surprisingly did not have any significant effect on the relative proportions of these species.

Overall, TFE has shown some unforeseen effects on the peroxidase kinetics of C45. With the impressive beneficial effect TFE has as a cosolvent, it would be interesting to observe whether this could be further enhanced. For instance, catalysis by C45 may conceivably be even more enhanced in neat TFE. Here, an increase from 70% to 80% TFE concentration almost doubled the k_{cat} at limiting peroxide concentration (from $2.6 \pm 0.4 \text{ s}^{-1}$ to $5.1 \pm 0.4 \text{ s}^{-1}$). This does suggest that a further rise in TFE concentration may yield even more impressive raises in catalytic activity. A continuation of this work could begin with the analysis of C45 in higher TFE concentrations, as well as other helix-stabilising solvents such as hexafluoroisopropanol (HFIP), another halogenated alcohol (307).

A further avenue of interest would be the investigation of the resilience of C45 to a mixture of hostile conditions. The enzyme has previously been shown to be catalytically active up to temperatures of at least 70 °C (2), in fact significantly more so than at room temperature. With the strong stabilising effect of TFE on the enzyme, it would be interesting to observe whether the enzyme remains active at even higher

temperatures. With the positive effect high temperatures generally have on reaction rates, the catalytic activity of C45 may even increase in higher temperatures, if thermostability of the enzyme can be achieved. It would be particularly interesting to see the difference in total turnover numbers when the enzyme is used in 80% TFE at high temperatures. It is possible that an increase in temperatures would be a trade-off between rising reaction rates and declining total turnover, even in the presence of TFE. Further, the effects of high-pressure environments are worth exploring for the enzyme in buffer as well as TFE. This could lay the foundation for new applications of maquette enzymes in industry. High-pressure environments, which have been shown to lead to dissociation of oligomeric proteins (308, 309), may also provide an interesting route to further investigate the difference in catalytic activity between dimer and monomer. Of course, with the growing maquette family and the opportunities for enzyme optimisation that directed evolution and computational methods have provided, the effect of TFE could be investigated on other helical enzymes, for example multi-heme enzymes.

This thesis could provide the basis for a new generation of highly resilient enzymes tailor-made for non-aqueous catalysis. This could not only lead to expanded applications of peroxidases, for instance in pollutant removal (310), but also fulfil an important industrial need.

7. Appendix

7.1. C45 and ABTS batch to batch testing

For kinetic measurements of C45, different batches of both the enzyme and its substrate ABTS were compared to ensure results obtained using different batches were comparable. A series of experiments was conducted measuring the rate of formation of ABTS^{•+} in buffer at 25 °C. Table 7.1 outlines which batches of enzyme and substrate were used for which experiment in the series.

Table 7.1: Table highlighting which batches of C45 and ABTS were used in which kinetic assay experiment in buffer for batch to batch comparisons for C45 and ABTS..

Expt.	C45	ABTS
BF252	Batch 2	Batch 1
BF253	Batch 2	Batch 2
BF255	Batch 3	Batch 2
BF256	Batch 3	Batch 1
BF258	Batch 1	Batch 2

During each experiment, four repeat measurements were performed and averaged. The initial rates for these experiments are shown in **Table 7.2**.

Table 7.2: ABTS^{•+} formation in buffer for different batches of C45 and ABTS.

Rate of the oxidation in s⁻¹ of 0.1 - 256 μM ABTS by 100 nM C45 and 100 μM H₂O₂ at 25 °C in aqueous solution using different batches of enzyme and substrate as outlined in

Table 7.1.

[ABTS] (μM)	BF252	BF253	BF255	BF256	BF258	Average
8	0.0453	--	0.08152	0.0747	0.05208	0.06 ± 0.02
16	0.2061	0.23098	0.40082	0.2083	0.33741	0.28 ± 0.09
32	0.4529	0.4837	0.56205	0.4429	0.47826	0.48 ± 0.05
64	0.6238	0.65606	0.59291	0.7328	0.65082	0.65 ± 0.06
128	0.6082	0.64587	0.60527	0.6744	0.62177	0.63 ± 0.03
256	0.5694	0.6601	0.58386	0.7047	0.62807	0.63 ± 0.06

Rates were plotted using Origin software and fitted to Michaelis-Menten kinetics. Kinetic parameters for the different experiments are shown in **Table 7.3**, which shows that different batches of substrate and enzyme do not have significantly different kinetic properties and are thus comparable. For any further analysis or comparison within this work, the average values of these parameters were used as the parameters of ABTS oxidation of C45 in buffer.

Table 7.3: Kinetics of ABTS^{•+} formation in buffer for different batches of C45 and ABTS.

k_{cat} (s^{-1}), K_M (μM), k_{cat}/K_M ($s^{-1} \mu M^{-1}$) as calculated by the origin software for different experiments showing no significant difference in k_{cat} for different batches of enzyme and substrate. Different batches of enzyme and substrate were used as outlined in

Table 7.1.

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1} \mu M^{-1}$)
BF252	0.7 ± 0.2	28 ± 15	0.026
BF253	0.77 ± 0.09	20 ± 10	0.034
BF255	0.69 ± 0.09	15 ± 8	0.045
BF256	0.9 ± 0.2	33 ± 16	0.026
BF258	0.7 ± 0.1	22 ± 10	0.033
Average	0.76 ± 0.07	24 ± 7	0.033 ± 0.08

Finally, total turnover numbers were calculated for the different experiments (**Table 7.4**). Again, no significant differences can be seen between the batches.

Table 7.4: Total turnover numbers for kinetic assays of different batches of C45 and ABTS.

No significant differences can be seen between the total turnover of experiments using different batches of substrate and enzyme, as outlined in

Table 7.1.

[ABTS] (μM)	BF252	BF253	BF255	BF256	BF258	Average
4	2.8125	3.22011	2.67663	2.93478	2.59511	2.8 ± 0.3
8	4.23913	4.4837	3.84511	4.44293	3.79076	4.2 ± 0.4
16	5.58424	5.70652	5.91033	5.54348	5.54348	5.7 ± 0.2
32	9.11685	9.07609	10.8696	8.92663	9.2663	9.5 ± 0.8
64	25.9783	25.6522	27.6495	25.0815	27.962	26 ± 2
128	40	39.4565	43.4647	42.1467	44.2391	42 ± 2
256	49.2935	50.0543	53.9674	53.3288	56.6168	53 ± 3

7.2. SAXS parameters

Table 7.5: Key SAXS parameters. Table reproduced from (311).

Parameter	Formula	Range of data used and variable definitions	Comments
Radius of gyration (R_G): Guinier approximation	$\ln [I(q)] = \ln [I(0)] - \frac{q^2 R_G^2}{3}$	$qR_G < 1.3$ globular, $qR_G < 0.8$ elongated. $I(0)$: Intensity of the scattering profile extrapolated to $q=0$	Most common method of estimating R_G Measured via the slope of the plot $\ln[I(q)]$ vs. q^2
Radius of gyration (R_G): Debye approximation	$I(q) = \frac{2I(0)}{q^4 R_G^4} (q^2 R_G^2 - 1 + e^{-q^2 R_G^2})$	$qR_G < 1.4$ for elongated macromolecules	Particularly useful for elongated proteins where the Guinier approximation is valid over narrower range
Radius of gyration (R_G): defined by $P(r)$	$R_G^2 = \int_0^{D_{\max}} r^2 P(r) dr / \int_0^{D_{\max}} P(r) dr$	Entire q -range. D_{\max} : Maximum dimension of particle	Good consistency check for R_G , D_{\max} , and $P(r)$
Pair distribution function ($P(r)$)	$P(r) = \frac{r}{2\pi^2} \int_0^\infty I(q) q \sin(qr) dq$	Entire q -range	Indirect Fourier transform methods have been developed for calculating $P(r)$
Maximum dimension (D_{\max})	D_{\max} is the value of r at $P(r) = 0$ for large r	Requires data $q \leq \pi/D_{\max}$	Assignment of D_{\max} may be complicated by flexibility or multimerization
Particle volume (V): defined by Porod Invariant	$V = 2\pi^2 I_{\text{exp}}^2(0) / \left(\int_0^\infty I(q) q^2 dq \right)$	Entire q -range. $I_{\text{exp}}(0)$ is the experimental intensity at $q=0$ and does not require an absolute scale	The integral portion of this equation is known as the Porod invariant. Accuracy varies for shape and size; however absolute scale and concentration information are unnecessary
$I(0)$: Intensity at $q=0$ which is also proportional to mass and volume	$I(0) = 4\pi \left(\int_0^{D_{\max}} P(r) dr \right)$	Entire q -range	Calculation of M and V using this version of $I(0)$ is less susceptible to aggregation and inter-particle correlations than extrapolation of low q data
Mass (M)	$M = \frac{I(0)\mu^2}{N_A(1 - (\rho_s/\rho_p))^2}$	μ : Average mass per number of electrons. ρ_s : Solvent electron density ρ_p : Particle electron density N_A : Avagadro's number	$I(0)$ must be on an absolute scale and normalized by mass/volume and not molar concentration
Formulas for elongated or flexibly linked linear macromolecules			
Radius of gyration of cross-section (R_{XC})	$\ln [qI(q)] = \ln [qI(0)] - \frac{q^2 R_{XC}^2}{2}$	Intermediate q values	The slope of the linear portion of a plot of $\ln[qI(q)]$ vs. q^2 is R_{XC}^2 ; however, R_{XC}^2 goes to 0 as q goes to 0 in regimes where scattering is dominated by R_G
Length (L)	$L = (12(R_G^2 - R_{XC}^2))^{\frac{1}{2}}$	See R_G and R_{XC}	The co-axial length rather than the hypotenuse (D_{\max})

8. References

1. P. S. Huang, S. E. Boyken, D. Baker, The coming of age of de novo protein design. *Nature*. **537**, 320–327 (2016).
2. D. W. Watkins, J. M. X. Jenkins, K. J. Grayson, N. Wood, J. W. Steventon, K. K. Le Vay, M. I. Goodwin, A. S. Mullen, H. J. Bailey, M. P. Crump, F. MacMillan, A. J. Mulholland, G. Cameron, R. B. Sessions, S. Mann, J. L. R. Anderson, Construction and in vivo assembly of a catalytically proficient and hyperthermostable de novo enzyme. *Nat. Commun.* **8** (2017).
3. M. Buck, Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **31**, 297–355 (1998).
4. M. J. Bodkin, J. M. Goodfellow, Hydrophobic solvation in aqueous trifluoroethanol solution. *Biopolymers*. **39**, 43–50 (1998).
5. D. Roccatano, G. Colombo, M. Fioroni, A. E. Mark, Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12179–12184 (2002).
6. H. W. Zhou, Y. Xu, H. M. Zhou, Activity and conformational changes of horseradish peroxidase in trifluoroethanol. *Biochem. Cell Biol.* **80**, 205–213 (2002).
7. A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, Industrial biocatalysis today and tomorrow. *Nature*. **409**, 258–268 (2001).
8. S. Martínez Cuesta, S. A. Rahman, N. Furnham, J. M. Thornton, The Classification and Evolution of Enzyme Function. *Biophys. J.* **109**, 1082–1086 (2015).
9. D. Voet, J. Voet, *Biochemistry* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 4th edn., 2011).
10. J. Kuriyan, B. Konforti, D. Wemmer, *The Molecules of Life* (Garland Science, London, U.K., 2012).
11. K. A. Dill, Theory for the Folding and Stability of Globular Proteins. *Biochemistry*. **24**, 1501–1509 (1985).
12. K. A. Dill, J. L. Maccallum, The Protein-Folding Problem, 50 Years On. *Science*. **338**, 1042–1046 (2012).
13. A. M. Klibanov, Improving enzymes by using them in organic solvents. *Nature*. **409**, 241–246 (2001).
14. A. Zaks, A. Klibanov, Enzymatic catalysis in nonaqueous solvents. *J. Biol. Chem.* **263**, 3194–3201 (1988).

15. A. Zaks, A. M. Klibanov, Enzyme-catalyzed processes in organic solvents. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3192–3196 (1985).
16. A. Klibanov, Why are enzymes less active in organic solvents than in water? *Trends Biotechnol.* **15**, 97–101 (1997).
17. W. F. DeGrado, L. Regan, S. P. Ho, The design of a four-helix bundle protein. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 521–526 (1987).
18. B. M. Discher, R. L. Koder, C. C. Moser, P. L. Dutton, Hydrophilic to amphiphilic design in redox protein maquettes. *Curr. Opin. Chem. Biol.* **7**, 741–748 (2003).
19. B. R. Lichtenstein, T. A. Farid, G. Kodali, L. A. Solomon, J. L. R. Anderson, M. M. Sheehan, N. M. Ennist, B. A. Fry, S. E. Chobot, C. Bialas, J. A. Mancini, C. T. Armstrong, Z. Zhao, T. V Esipova, D. Snell, S. A. Vinogradov, B. M. Discher, C. C. Moser, P. L. Dutton, Engineering oxidoreductases: maquette proteins designed from scratch. *Biochem Soc Trans.* **40**, 561–566 (2013).
20. J. L. R. Anderson, C. T. Armstrong, G. Kodali, B. R. Lichtenstein, D. W. Watkins, J. A. Mancini, A. L. Boyle, T. A. Farid, M. P. Crump, C. C. Moser, P. L. Dutton, Constructing a man-made c-type cytochrome maquette in vivo: Electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chem. Sci.* **5**, 507–514 (2014).
21. D. W. Watkins, C. T. Armstrong, J. L. Beesley, J. E. Marsh, J. M. X. Jenkins, R. B. Sessions, S. Mann, J. L. Ross Anderson, A suite of de novo c-type cytochromes for functional oxidoreductase engineering. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 493–502 (2016).
22. R. Singh, M. Kumar, A. Mittal, P. K. Mehta, Microbial enzymes: industrial progress in 21st century. *3 Biotech.* **6**, 174–189 (2016).
23. P. K. Robinson, Enzymes: principles and biotechnological applications. *Essays Biochem.* **59**, 1–41 (2015).
24. Nobel Media, Eduard Buchner - Facts - NobelPrize.org, (available at <https://www.nobelprize.org/prizes/chemistry/1907/buchner/facts/>).
25. T. D. H. Bugg, The development of mechanistic enzymology in the 20th century. *Nat. Prod. Rep.* **18**, 465–493 (2001).
26. D. Eisenberg, The discovery of the α -helix and β -sheet, the principal structural features of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11207–11210 (2003).
27. J. C. Kendrew, R. E. Dickerson, B. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, Structure of myoglobin: A three-dimensional fourier synthesis at 2 Å resolution. *Nature.* **185**, 422–427 (1960).
28. M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T. North, Structure of Hæmoglobin: A three-dimensional fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature.* **185**, 416–422 (1960).
29. D. M. Goldberg, Clinical enzymology: An autobiographical history. *Clin. Chim. Acta.* **357**, 93–112 (2005).
30. N. V. Bhagavan, C.-E. Ha, in *Essentials of Medical Biochemistry* (Academic

- Press, Cambridge, MA, USA 2015), pp. 85–95.
31. O. Evliyaoğlu, E. Kibrisli, Y. Yildirim, O. Gökalp, L. Çolpan, Routine enzymes in the monitoring of type 2 diabetes mellitus. *Cell Biochem. Funct.* **29**, 506–512 (2011).
 32. S. L. Liang, D. W. Chan, Enzymes and related proteins as cancer biomarkers: A proteomic approach. *Clin. Chim. Acta.* **381**, 93–97 (2007).
 33. A. Gessesse, F. Mulaa, S. L. Lyantagaye, L. Nyina-Wamwiza, B. Mattiasson, A. Pandey, “Industrial Enzymes for Sustainable Bio-Economy: Large Scale Production and Application in Industry, Environment, and Agriculture in Eastern Africa” (Nairobi, Kenya, 2011), (available at <http://creativecommons.org/licenses/by/4.0/>).
 34. A. Piotrowska-Długosz, in *Biocatalysis* (Springer International Publishing, New York, NY, USA, 2019), pp. 277–308.
 35. S. Raveendran, B. Parameswaran, S. B. Ummalyma, A. Abraham, A. K. Mathew, A. Madhavan, S. Rebello, A. Pandey, Applications of microbial enzymes in food industry. *Food Technol. Biotechnol.* **56**, 16–30 (2018).
 36. O. Kirk, T. V. Borchert, C. C. Fuglsang, Industrial enzyme applications. *Curr. Opin. Biotechnol.* **13**, 345–351 (2002).
 37. E. Kuah, S. Toh, J. Yee, Q. Ma, Z. Gao, Enzyme Mimics: Advances and Applications. *Chem. - A Eur. J.* **22**, 8404–8430 (2016).
 38. R. Wolfenden, Y. Yuan, Monoalkyl sulfates as alkylating agents in water, alkylsulfatase rate enhancements, and the “energy-rich” nature of sulfate half-esters. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 83–86 (2007).
 39. A. Blanco, G. Blanco, in *Medical Biochemistry* (Academic Press, Cambridge, MA, USA, 2017), pp. 153–175.
 40. R. W. Hartley, J. R. Smeaton, On the reaction between the extracellular ribonuclease of *Bacillus amyloliquefaciens* (Barnase) and its intracellular inhibitor (Barstar). *J. Biol. Chem.* **248**, 5624–5626 (1973).
 41. W. Y. Huang, S. S. Chirala, S. J. Wakil, Amino-Terminal blocking group and sequence of the animal fatty acid synthase. *Arch. Biochem. Biophys.* **314**, 45–49 (1994).
 42. E. A. Doherty, J. A. Doudna, Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 457–475 (2001).
 43. M. L. Bulyk, X. Huang, Y. Choo, G. M. Church, Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7158–7163 (2001).
 44. E. K. O’Shea, R. Rutkowski, P. S. Kim, Evidence that the leucine zipper is a coiled coil. *Science.* **243**, 538–542 (1989).
 45. J. W. Fairman, N. Noinaj, S. K. Buchanan, The structural biology of β -barrel membrane proteins: A summary of recent reports. *Curr. Opin. Struct. Biol.* **21**, 523–531 (2011).
 46. R. K. Wierenga, The TIM-barrel fold: A versatile framework for efficient

- enzymes. *FEBS Lett.* **492**, 193–198 (2001).
47. K. T. Halloran, Y. Wang, K. Arora, S. Chakravarthy, T. C. Irving, O. Bilsel, C. L. Brooks, C. Robert Matthews, Frustration and folding of a TIM barrel protein. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16378–16383 (2019).
 48. R. Russo, S. Zucchelli, M. Codrich, F. Marcuzzi, C. Verde, S. Gustincich, Hemoglobin is present as a canonical $\alpha_2\beta_2$ tetramer in dopaminergic neurons. *Biochim. Biophys. Acta - Proteins Proteomics.* **1834**, 1939–1943 (2013).
 49. H. M. Berman, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
 50. C. Levinthal, How to fold graciously. *Mössbauer Spectrosc. Biol. Syst. Proc.* **67**, 22–24 (1969).
 51. R. Zwanzig, A. Szabo, B. Bagchi, Levinthal's paradox. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 20–22 (1992).
 52. C. B. Anfinsen, E. Haber, M. Sela, F. H. White, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1309–1314 (1961).
 53. H. Taniuchi, C. B. Anfinsen, An experimental approach to the study of the folding of staphylococcal nuclease. *J. Biol. Chem.* **244**, 3864–3875 (1969).
 54. Nobel Media, Press release: The 1972 Nobel Prize in Chemistry, (available at <https://www.nobelprize.org/prizes/chemistry/1972/press-release/>).
 55. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science.* **181**, 223–230 (1973).
 56. S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, Protein design by binary patterning of polar and nonpolar amino acids. *Science.* **262**, 1680–1685 (1993).
 57. K. M. Makwana, R. Mahalakshmi, Implications of aromatic-aromatic interactions: From protein structures to peptide models. *Protein Sci.* **24**, 1920–1933 (2015).
 58. M. Reches, E. Gazit, Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science.* **300**, 625–627 (2003).
 59. S. G. Chang, K. D. Choi, S. H. Jang, H. C. Shin, Role of disulfide bonds in the structure and activity of human insulin. *Mol. Cells.* **16**, 323–330 (2003).
 60. K. A. Dill, H. S. Chan, From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10–19 (1997).
 61. U. Ahluwalia, N. Katyal, S. Deep, Research Education Models of Protein Folding. *J P P J. Proteins Proteomics.* **3**, 85–93 (2012).
 62. D. Baker, What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 (2019).
 63. S. E. Jackson, How do small single-domain proteins fold? *Fold. Des.* **3**, 81–91 (1998).
 64. P. Malhotra, J. B. Udgaonkar, How cooperative are protein folding and unfolding

8. References

- transitions? *Protein Sci.* **25**, 1924–1941 (2016).
65. S. E. Radford, Protein folding: Progress made and promises ahead. *Trends Biochem. Sci.* **25**, 611–618 (2000).
 66. S. E. Jackson, A. R. Fersht, Folding of Chymotrypsin Inhibitor 2.1. Evidence for a Two-State Transition. *Biochemistry.* **30**, 10428–10435 (1991).
 67. A. R. Fersht, Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3–9 (1997).
 68. H. A. Scheraga, M. Khalili, A. Liwo, Protein-Folding Dynamics: Overview of Molecular Simulation Techniques. *Annu. Rev. Phys. Chem.* **58**, 57–83 (2007).
 69. O. B. Ptitsyn, A. A. Rashin, A model of myoglobin self-organization. *Biophys. Chem.* **3**, 1–20 (1975).
 70. P. S. Kim, R. L. Baldwin, Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Annu. Rev. Biochem.* **51**, 459–489 (1982).
 71. N. Go, The consistency principle in protein structure and pathways of folding. *Adv. Biophys.* **18**, 149–64 (1984).
 72. M. Karplus, D. L. Weaver, Protein-folding dynamics. *Nature.* **260**, 404–406 (1976).
 73. M. Karplus, D. L. Weaver, Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci.* **3**, 650–668 (1994).
 74. W. Kauzmann, Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* **14**, 1–63 (1959).
 75. B. Nölting, D. A. Agard, How general is the nucleation-condensation mechanism? *Proteins Struct. Funct. Genet.* **73**, 754–764 (2008).
 76. K. K. Sinha, J. B. Udgaonkar, Early events in protein folding. *Curr. Sci.* **96**, 1053–1070 (2009).
 77. V. I. Abkevich, A. M. Gutin, E. I. Shakhnovich, Specific Nucleus as the Transition State for Protein Folding: Evidence from the Lattice Model. *Biochemistry.* **33**, 10026–10036 (1994).
 78. V. Daggett, A. R. Fersht, Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **28**, 18–25 (2003).
 79. S. Gianni, M. Brunori, C. Travaglini-Allocatelli, Plasticity of the protein folding landscape: Switching between on- and off-pathway intermediates. *Arch. Biochem. Biophys.* **466**, 172–176 (2007).
 80. S. Wallin, K. B. Zeldovich, E. I. Shakhnovich, The Folding Mechanics of a Knotted Protein. *J. Mol. Biol.* **368**, 884–893 (2007).
 81. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct. Funct. Bioinforma.* **21**, 167–195 (1995).
 82. J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, Theory of protein folding: The

- Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
83. J. P. Hendrick, F. U. Hartl, The role of molecular chaperones in protein folding. *FASEB J.* **9**, 1559–1569 (1995).
84. F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature.* **475**, 324–332 (2011).
85. G. Ashraf, N. Greig, T. Khan, I. Hassan, S. Tabrez, S. Shakil, I. Sheikh, S. Zaidi, M. Akram, N. Jabir, C. Firoz, A. Naeem, I. Alhazza, G. Damanhouri, M. Kamal, Protein Misfolding and Aggregation in Alzheimer's Disease and Type 2 Diabetes Mellitus. *CNS Neurol. Disord. - Drug Targets.* **13**, 1280–1293 (2014).
86. L. Chen, M. D. Brewer, L. Guo, R. Wang, P. Jiang, X. Yang, Enhanced Degradation of Misfolded Proteins Promotes Tumorigenesis. *Cell Rep.* **18**, 3143–3154 (2017).
87. G. P. Brady, K. A. Sharp, Entropy in protein folding and in protein-protein interactions. *Curr. Opin. Struct. Biol.* **7**, 215–221 (1997).
88. A. Goldenzweig, S. J. Fleishman, Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **87**, 105–129 (2018).
89. D. M. Taverna, R. A. Goldstein, Why are proteins marginally stable? *Proteins Struct. Funct. Genet.* **46**, 105–109 (2002).
90. G. McLendon, E. Radany, Is protein turnover thermodynamically controlled? *J. Biol. Chem.* **253**, 6335–6337 (1978).
91. C. J. Tsai, R. Nussinov, A Unified View of “How Allostery Works.” *PLoS Comput. Biol.* **10** (2014).
92. I. Halperin, H. Wolfson, R. Nussinov, Protein-protein interactions: Coupling of structurally conserved residues and of hot spots across interfaces. Implications for docking. *Structure.* **12**, 1027–1038 (2004).
93. T. J. Ahern, A. M. Klibanov, in *Methods of Biochemical Analysis* (John Wiley and Sons Inc., Hoboken, NJ, USA 2006), vol. 33, pp. 91–127.
94. H. Frauenfelder, G. A. Petsko, D. Tsernoglou, Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature.* **280**, 558–563 (1979).
95. W. Pfeil, The problem of the stability of globular proteins. *Mol. Cell. Biochem.* **40**, 3–28 (1981).
96. H. P. Modarres, M. R. Mofrad, A. Sanati-Nezhad, Protein thermostability engineering. *RSC Adv.* **6**, 115252–115270 (2016).
97. D. Prakash, N. Nawani, M. Prakash, M. Bodas, A. Mandal, M. Khetmalas, B. Kapadnis, Actinomycetes: A repertory of green catalysts with a potential revenue resource. *Biomed Res. Int.* **2013** (2013).
98. F. Rigoldi, S. Donini, A. Redaelli, E. Parisini, A. Gautieri, Review: Engineering of thermostable enzymes for industrial applications. *APL Bioeng.* **2**, 11501–11517 (2018).

8. References

99. V. L. Arcus, E. J. Prentice, J. K. Hobbs, A. J. Mulholland, M. W. Van Der Kamp, C. R. Pudney, E. J. Parker, L. A. Schipper, On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry*. **55**, 1681-1688 (2016).
100. P. L. Poole, J. L. Finney, Sequential hydration of a dry globular protein. *Biopolymers*. **22**, 255–260 (1983).
101. G. T. J.A. Rupley, P.H. Yang, Water in polymers. *Am. Chem. Soc.* **127**, 111 (1980).
102. I. D. Kuntz, W. Kauzmann, Hydration of Proteins and Polypeptides. *Adv. Protein Chem.* **28**, 239–345 (1974).
103. J. A. Rupley, G. Careri, Protein Hydration and Function. *Adv. Protein Chem.* **41**, 37–172 (1991).
104. C. Mattos, D. Ringe, Proteins in organic solvents. *Curr. Opin. Struct. Biol.* **11**, 761–764 (2001).
105. I. Roy, M. N. Gupta, Freeze-drying of proteins: some emerging concerns. *Biotechnol. Appl. Biochem.* **39**, 165-177 (2004).
106. C. R. Wescott, A. M. Klivanov, The solvent dependence of enzyme specificity. *Biochim. Biophys. Acta* **1206**, 1–9 (1994).
107. G. Carrea, G. Ottolina, S. Riva, Role of solvents in the control of enzyme selectivity in organic media. *Trends Biotechnol.* **13**, 63–70 (1995).
108. A. Zaks, A. M. Klivanov, Substrate Specificity of Enzymes in Organic Solvents vs. Water Is Reversed. *J. Am. Chem. Soc.* **108**, 2767–2768 (1986).
109. C. R. Wescott, A. M. Klivanov, Solvent Variation Inverts Substrate Specificity of an Enzyme. *J. Am. Chem. Soc.* **115**, 1629–1631 (1993).
110. D. A. MacManus, E. N. Vulfson, Reversal of regioselectivity in the enzymatic acylation of secondary hydroxyl groups mediated by organic solvents. *Enzyme Microb. Technol.* **20**, 225–228 (1997).
111. E. Rubio, A. Fernandez-Mayorales, A. M. Klivanov, Effect of the Solvent on Enzyme Regioselectivity. *J. Am. Chem. Soc.* **113**, 695–696 (1991).
112. S. Ueji, S. Kurita, R. Fujino, N. Okubo, T. Miyazawa, M. Kitadani, A. Muromatsu, Solvent-induced inversion of enantioselectivity in Lipase-catalyzed esterification of 2-Phenoxypropionic acids. *Biotechnol. Lett.* **14**, 163–168 (1992).
113. S. Tawakit, A. M. Klivanov, Inversion of Enzyme Enantioselectivity Mediated by the Solvent. *J. Am. Chem. Soc.* **114**, 1882–1884 (1992).
114. C. R. Wescott, H. Noritomi, A. M. Klivanov, Rational control of enzymatic enantioselectivity through solvation thermodynamics. *J. Am. Chem. Soc.* **118**, 10365–10370 (1996).
115. S. H. Wu, F. Y. Chu, K. T. Wang, Reversible enantioselectivity of enzymatic reactions by media. *Bioorganic Med. Chem. Lett.* **1**, 339–342 (1991).
116. A. M. Klivanov, Asymmetric Transformations Catalyzed by Enzymes in Organic Solvents. *Acc. Chem. Res.* **23**, 114–120 (1990).

117. P. Halling, L. Kvittingen, Why did biocatalysis in organic media not take off in the 1930s? *Trends Biotechnol.* **17**, 343–344 (1999).
118. A. Zaks, A. M. Klivanov, Enzymatic catalysis in organic media at 100°C. *Science.* **224**, 1249–1251 (1984).
119. E. Antonini, G. Carrea, P. Cremonesi, Enzyme catalysed reactions in water - Organic solvent two-phase systems. *Enzyme Microb. Technol.* **3**, 291–296 (1981).
120. S. J. Singer, The Properties of Proteins in Nonaqueous Solvents. *Adv. Protein Chem.* **17**, 1–68 (1963).
121. S. Kamat, E. J. Beckman, A. J. Russell, Role of diffusion in nonaqueous enzymology. *Enzyme Microb. Technol.* **14**, 265–271 (1992).
122. P. L. Luisi, Enzymes Hosted in Reverse Micelles in Hydrocarbon Solution. *Angew. Chemie Int. Ed. English.* **24**, 439–450 (1985).
123. K. Martinek, A. M. Klivanov, V. S. Goldmacher, I. V. Berezin, The principles of enzyme stabilization I. Increase in thermostability of enzymes covalently bound to a complementary surface of a polymer support in a multipoint fashion. *BBA - Enzymol.* **485**, 1–12 (1977).
124. T. Arakawa, S. J. Prestrelski, W. C. Kenney, J. F. Carpenter, Factors affecting short-term and long-term stabilities of proteins. *Adv. Drug Deliv. Rev.* **46**, 307–326 (2001).
125. Y. L. Khmel'nitsky, S. H. Welch, J. S. Dordick, D. S. Clark, Salts dramatically enhance activity of enzymes suspended in organic solvents. *J. Am. Chem. Soc.* **116**, 2647–2648 (1994).
126. J. Broos, I. K. Sakodinskaya, J. F. J. Engbersen, W. Verboom, D. N. Reinhoudt, Large activation of serine proteases by pretreatment with crown ethers. *J. Chem. Soc. Chem. Commun.*, 255–256 (1995).
127. K. Debulis, A. M. Klivanov, Dramatic enhancement of enzymatic activity in organic solvents by lyoprotectants. *Biotechnol. Bioeng.* **41**, 566–571 (1993).
128. V. M. Paradkar, J. S. Dordick, Aqueous-like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents. *J. Am. Chem. Soc.* **116**, 5009–5010 (1994).
129. H. R. Costantino, K. Griebenow, R. Langer, A. M. Klivanov, On the pH memory of lyophilized compounds containing protein functional groups. *Biotechnol. Bioeng.* **53**, 345–348 (1997).
130. R. Affleck, C. A. Haynes, D. S. Clark, Solvent dielectric effects on protein dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5167–5170 (1992).
131. A. Zaks, A. M. Klivanov, The effect of water on enzyme action in organic media. *J. Biol. Chem.* **263**, 8017–8021 (1988).
132. J. O. Rich, J. S. Dordick, Controlling Subtilisin Activity and Selectivity in Organic Media by Imprinting with Nucleophilic Substrates. *J. Am. Chem. Soc.* **119**, 3245–3252 (1997).

8. References

133. A. J. Russell, A. M. Klibanov, Inhibitor-induced enzyme activation in organic solvents. *J. Biol. Chem.* **263**, 11624–11626 (1988).
134. A. R. Fersht, *Enzyme Structure and Mechanism* (W. H. Freeman and Company, New York, NY, USA, 2nd edn., 1984).
135. J. L. Schmitke, C. R. Wescott, A. M. Klibanov, The mechanistic dissection of the plunge in enzymatic activity upon transition from water to anhydrous solvents. *J. Am. Chem. Soc.* **118**, 3360–3365 (1996).
136. R. Batra, M. N. Gupta, Enhancement of enzyme activity in aqueous-organic solvent mixtures. *Biotechnol. Lett.* **16**, 1059–1064 (1994).
137. T. Sawamura, K. Takahashi, S. Inagi, T. Fuchigami, Sodium Salts Dissolution in an Aprotic Solvent by Coordination of Poly(ethylene glycol) for Effective Anodic Reactions of Organic Compounds. *Commun. Electrochem.* **81**, 365–367 (2013).
138. E. Candi, M. Paci, S. Oddi, A. Paradisi, P. Guerrieri, G. Melino, Ordered structure acquisition by the N- and C-terminal domains of the small proline-rich 3 protein. *J. Cell. Biochem.* **77**, 179–185 (2000).
139. A. Jasanoff, A. R. Fersht, Quantitative Determination of Helical Propensities from Trifluoroethanol Titration Curves. *Biochemistry.* **33**, 2129–2135 (1994).
140. S. Russ Lehrman, J. L. Tuls, M. Lund, Peptide α -Helicity in Aqueous Trifluoroethanol: Correlations with Predicted α -Helicity and the Secondary Structure of the Corresponding Regions of Bovine Growth Hormone. *Biochemistry.* **29**, 5590–5596 (1990).
141. M. Buck, S. E. Radford, C. M. Dobson, A Partially Folded State of Hen Egg White Lysozyme in Trifluoroethanol: Structural Characterization and Implications for Protein Folding. *Biochemistry.* **32**, 669–678 (1993).
142. K. Shiraki, K. Nishikawa, Y. Goto, Trifluoroethanol-induced Stabilization of the α -Helical Structure of β -Lactoglobulin: Implication for Non-hierarchical Protein Folding. *J. Mol. Biol.* **245**, 180–194 (1995).
143. J. K. Myers, C. Nick Pace, J. Martin Scholtz, Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. *Protein Sci.* **7**, 383–388 (2008).
144. R. W. Storrs, D. Truckses, D. E. Wemmer, Helix propagation in trifluoroethanol solutions. *Biopolymers.* **32**, 1695–1702 (1992).
145. A. Cammers-Goodwin, T. J. Allen, S. L. Oslick, K. F. McClure, J. H. Lee, D. S. Kemp, Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol. *J. Am. Chem. Soc.* **118**, 3082–3090 (1996).
146. A. Chakrabarty, T. Kortemme, S. Padmanabhan, R. L. Baldwin, Aromatic Side-Chain Contribution to Far-Ultraviolet Circular Dichroism of Helical Peptides and Its Effect on Measurement of Helix Propensities. *Biochemistry.* **32**, 5560–5565 (1993).
147. J. E. Brown, W. A. Klee, Helix-Coil Transition of the Isolated Amino Terminus of Ribonuclease. *Biochemistry.* **10**, 470–476 (1971).

148. P. Luo, R. L. Baldwin, Mechanism of Helix Induction by Trifluoroethanol: A Framework for Extrapolating the Helix-Forming Properties of Peptides from Trifluoroethanol/Water Mixtures Back to Water. *Biochemistry*. **36**, 8413–8421 (1997).
149. Y. Yang, S. Barker, Mu Jung Chen, K. H. Mayo, Effect of low molecular weight aliphatic alcohols and related compounds on platelet factor 4 subunit association. *J. Biol. Chem.* **268**, 9223–9229 (1993).
150. F. J. Blanco, A. R. Ortiz, L. Serrano, Role of a nonnative interaction in the folding of the protein G B1 domain as inferred from the conformational analysis of the α -helix fragment. *Fold. Des.* **2**, 123–133 (1997).
151. V. Rizzo, H. Jäckle, Side-Chain vs. Main-Chain Conformational Flexibility in Aromatic Dipeptides. *J. Am. Chem. Soc.* **105**, 4195–4205 (1983).
152. J. W. Nelson, N. R. Kallenbach, Stabilization of the ribonuclease S-peptide α -helix by trifluoroethanol. *Proteins Struct. Funct. Bioinforma.* **1**, 211–217 (1986).
153. K. J. Laidler, M. C. King, The Development of Transition-State Theory. *J. Phys. Chem.* **87**, 2657–2664 (1983).
154. T. D. H. Bugg, *Introduction to Enzyme and Coenzyme Chemistry* (John Wiley & Sons, Inc., Chichester, U.K., 3rd edn., 2012).
155. J. L. E. Monti, M. R. Montes, R. C. Rossi, Steady-state analysis of enzymes with non-Michaelis-Menten kinetics: The transport mechanism of Na⁺/K⁺-ATPase. *J. Biol. Chem.* **293**, 1373–1385 (2018).
156. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics* (Portland Press, London, U.K., 3rd edn., 2004).
157. N. N. Ulusu, Evolution of Enzyme Kinetic Mechanisms. *J. Mol. Evol.* **80**, 251–257 (2015).
158. R. David Britt, E. Raven, Editorial overview: Bioinorganic chemistry: Key roles for biological metal centers. *Curr. Opin. Chem. Biol.* **31**, 8–9 (2016).
159. M. Pérez-Navarro, F. Neese, W. Lubitz, D. A. Pantazis, N. Cox, Recent developments in biological water oxidation. *Curr. Opin. Chem. Biol.* **31**, 113–119 (2016).
160. G. Schwarz, Molybdenum cofactor and human disease. *Curr. Opin. Chem. Biol.* **31**, 179–187 (2016).
161. R. M. Roat-Malone, *Bioinorganic Chemistry: A Short Course* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2007).
162. S. N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, M. Guertin, J. Gough, S. Dewilde, L. Moens, J. R. Vanfleteren, Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11385–11389 (2005).
163. J. W. A. Allen, O. Daltrop, J. M. Stevens, S. J. Ferguson, C-type cytochromes: Diverse structures and biogenesis systems pose evolutionary problems. *Philos. Trans. R. Soc. B Biol. Sci.* **358**, 255–266 (2003).

8. References

164. S. E. J. Bowman, K. L. Bren, The Chemistry and Biochemistry of Heme c: Functional Bases for Covalent Attachment. *Nat. Prod. Rep.* **25**, 1118–1130 (2008).
165. J. W. A. Allen, N. Leach, S. J. Ferguson, The histidine of the c-type cytochrome CXXCH haem-binding motif is essential for haem attachment by the Escherichia coli cytochrome c maturation (Ccm) apparatus. *Biochem. J.* **389**, 587–592 (2005).
166. T. Li, H. L. Bonkovsky, J. T. Guo, Structural analysis of heme proteins: Implications for design and prediction. *BMC Struct. Biol.* **11** (2011).
167. G. Zamborlini, M. Jugovac, A. Cossaro, A. Verdini, L. Floreano, D. Lüftner, P. Puschnig, V. Feyer, C. M. Schneider, On-surface nickel porphyrin mimics the reactive center of an enzyme cofactor. *Chem. Commun.* **54**, 13423–13426 (2018).
168. S. N. Vinogradov, L. Moens, Diversity of globin function: Enzymatic, transport, storage, and sensing. *J. Biol. Chem.* **283**, 8773–8777 (2008).
169. V. Cannac-Caffrey, G. Hudry-Clergeon, Y. Pétillet, J. Gagnon, G. Zaccai, B. Franzetti, The protein sequence of an archaeal catalase-peroxidase. *Biochimie.* **80**, 1003–1011 (1998).
170. F. Passardi, N. Bakalovic, F. K. Teixeira, M. Margis-Pinheiro, C. Penel, C. Dunand, Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. *Genomics.* **89**, 567–579 (2007).
171. A. A. Khan, A. H. Rahmani, Y. H. Aldebasi, S. M. Aly, Biochemical and pathological studies on peroxidases -an updated review. *Glob. J. Health Sci.* **6**, 87–98 (2014).
172. V. P. Pandey, M. Awasthi, S. Singh, S. Tiwari, U. N. Dwivedi, A Comprehensive Review on Function and Application of Plant Peroxidases. *Biochem Anal Biochem.* **6**, 1 (2017).
173. G. Coussot, C. Faye, A. Le Postollec, M. Dobrijevic, One-step direct immunoassay with horseradish peroxidase as antigen for studying the functionality of antibody surfaces. *Talanta.* **178**, 922–927 (2018).
174. A. Bach, R. Chodat, Untersuchungen über die Rolle der Peroxyde in der Chemie der lebenden Zelle. IV. Ueber Peroxydase. *Berichte der Dtsch. Chem. Gesellschaft.* **36**, 600–605 (1903).
175. N. C. Veitch, Horseradish peroxidase: A modern view of a classic enzyme. *Phytochemistry.* **65**, 249–259 (2004).
176. B. Savelli, Q. Li, M. Webber, A. M. Jemmat, A. Robitaille, M. Zamocky, C. Mathé, C. Dunand, RedoxiBase: A database for ROS homeostasis regulated proteins. *Redox Biol.* **26**, 101247 (2019).
177. F. W. Krainer, A. Glieder, An updated view on horseradish peroxidases: recombinant production and biotechnological applications. *Appl. Microbiol. Biotechnol.* **99**, 1611–1625 (2015).
178. G. I. Berglund, G. H. Carlsson, A. T. Smith, H. Szöke, A. Henriksen, J. Hajdu,

- The catalytic pathway of horseradish peroxidase at high resolution. *Nature*. **417**, 463–468 (2002).
179. M. Li, D. C. Green, J. L. R. Anderson, B. P. Binks, S. Mann, In vitro gene expression and enzyme catalysis in bio-inorganic protocells. *Chem. Sci.* **2**, 1739–1745 (2011).
180. T. Y. Dora. Tang, D. Van Swaay, A. DeMello, J. L. Ross Anderson, S. Mann, In vitro gene expression within membrane-free coacervate protocells. *Chem. Commun.* **51**, 11429–11432 (2015).
181. J. F. Ross, A. Bridges, J. M. Fletcher, D. Shoemark, D. Alibhai, H. E. V. Bray, J. L. Beesley, W. M. Dawson, L. R. Hodgson, J. Mantell, P. Verkade, C. M. Edge, R. B. Sessions, D. Tew, D. N. Woolfson, Decorating Self-Assembled Peptide Cages with Proteins. *ACS Nano*. **11**, 7901–7914 (2017).
182. V. Thiel, Synthetic viruses—Anything new? *PLoS Pathog.* **14** (2018).
183. F. Benn, N. E. C. Haley, A. E. Lucas, E. Silvester, S. Helmi, R. Schreiber, J. Bath, A. J. Turberfield, Chiral DNA Origami Nanotubes with Well-Defined and Addressable Inside and Outside Surfaces. *Angew. Chemie Int. Ed.* **57**, 7687–7690 (2018).
184. C. Jäckel, P. Kast, D. Hilvert, Protein Design by Directed Evolution. *Annu. Rev. Biophys.* **37**, 153–173 (2008).
185. C. Zeymer, D. Hilvert, Directed Evolution of Protein Catalysts. *Annu. Rev. Biochem.* **87**, 131–157 (2018).
186. M. S. Packer, D. R. Liu, Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
187. J. D. Kaunitz, The Discovery of PCR: ProCuRement of Divine Power. *Dig. Dis. Sci.* **60**, 2230–2231 (2015).
188. R. E. Cobb, R. Chao, H. Zhao, Directed evolution: Past, present, and future. *AIChE J.* **59**, 1432–1440 (2013).
189. B. K. Shoichet, W. A. Baase, R. Kuroki, B. W. Matthews, A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 452–456 (1995).
190. N. Tokuriki, F. Stricher, L. Serrano, D. S. Tawfik, How Protein Stability and New Functions Trade Off. *PLoS Comput. Biol.* **4** (2008).
191. F. H. Arnold, The nature of chemical innovation: new enzymes by evolution. *Q. Rev. Biophys.* **48**, 404–410 (2019).
192. M. Goldsmith, D. S. Tawfik, Directed enzyme evolution: Beyond the low-hanging fruit. *Curr. Opin. Struct. Biol.* **22**, 406–412 (2012).
193. R. Fasan, Y. T. Meharena, C. D. Snow, T. L. Poulos, F. H. Arnold, Evolutionary History of a Specialized P450 Propane Monooxygenase. *J. Mol. Biol.* **383**, 1069–1080 (2008).
194. Nobel Media, All Nobel Prizes in Chemistry, (available at <https://www.nobelprize.org/prizes/lists/all-nobel-prizes-in-chemistry>).

8. References

195. B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science*. **302**, 1364–1368 (2003).
196. M. Mravic, J. L. Thomaston, M. Tucker, P. E. Solomon, L. Liu, W. F. DeGrado, Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science*. **363**, 1418–1423 (2019).
197. M. Mravic, H. Hu, Z. Lu, J. S. Bennett, C. R. Sanders, A. W. Orr, W. F. DeGrado, De novo designed transmembrane peptides activating the $\alpha 5\beta 1$ integrin. *Protein Eng. Des. Sel.* **31**, 181–190 (2018).
198. J. M. Fletcher, R. L. Harniman, F. R. H. Barnes, A. L. Boyle, A. Collins, J. Mantell, T. H. Sharp, M. Antognozzi, P. J. Booth, N. Linden, M. J. Miles, R. B. Sessions, P. Verkade, D. N. Woolfson, Self-assembling cages from coiled-coil peptide modules. *Science*. **340**, 595–599 (2013).
199. F. Boato, R. M. Thomas, A. Ghasparian, A. Freund-Renard, K. Moehle, J. A. Robinson, Synthetic Virus-Like Particles from Self-Assembling Coiled-Coil Lipopeptides and Their Use in Antigen Display to the Immune System. *Angew. Chemie Int. Ed.* **46**, 9015–9018 (2007).
200. A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady, D. N. Woolfson, Computational design of water-soluble α -helical barrels. *Science*. **346**, 485–488 (2014).
201. S. P. Ho, W. F. DeGrado, Design of a 4-Helix Bundle Protein: Synthesis of Peptides Which Self-Associate into a Helical Protein. *J. Am. Chem. Soc.* **109**, 6751–6758 (1987).
202. L. Regan, W. F. Degrado, Characterization of a helical protein designed from first principles. *Science*. **241**, 976–978 (1988).
203. C. Chothia, M. Levitt, D. Richardson, Structure of proteins: packing of α -helices and pleated sheets. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 4130–4134 (1977).
204. N. M. Ennist, J. A. Mancini, D. B. Auman, C. Bialas, M. J. Iwanicki, T. V. Esipova, B. M. Discher, C. C. Moser, P. Leslie Dutton, in *Photosynthesis and Bioenergetics* (World Scientific Publishing Co. Pte. Ltd., Singapore, 2017), pp. 1–34.
205. H. Kries, R. Blomberg, D. Hilvert, De novo enzymes by computational design. *Curr. Opin. Chem. Biol.* **17**, 221–228 (2013).
206. G. Kodali, J. A. Mancini, L. A. Solomon, T. V. Episova, N. Roach, C. J. Hobbs, P. Wagner, O. A. Mass, K. Aravindu, J. E. Barnsley, K. C. Gordon, D. L. Officer, P. L. Dutton, C. C. Moser, Design and engineering of water-soluble light-harvesting protein maquettes. *Chem. Sci.* **8**, 316–324 (2016).
207. J. A. Mancini, G. Kodali, J. Jiang, K. R. Reddy, J. S. Lindsey, D. A. Bryant, P. L. Dutton, C. C. Moser, Multi-step excitation energy transfer engineered in genetic fusions of natural and synthetic light-harvesting proteins. *J. R. Soc. Interface.* **14** (2017).
208. C. Bialas, L. E. Jarocha, K. B. Henbest, T. M. Zollitsch, G. Kodali, C. R. Timmel, S. R. Mackenzie, P. L. Dutton, C. C. Moser, P. J. Hore, Engineering an Artificial

- Flavoprotein Magnetosensor. *J. Am. Chem. Soc.* **138**, 16584–16587 (2016).
209. G. Goparaju, B. A. Fry, S. E. Chobot, G. Wiedman, C. C. Moser, P. Leslie Dutton, B. M. Discher, First principles design of a core bioenergetic transmembrane electron-transfer protein. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 503–512 (2016).
210. S. C. Patel, L. H. Bradley, S. P. Jinadasa, M. H. Hecht, Cofactor binding and enzymatic activity in an unevolved superfamily of de novo designed 4-helix bundle proteins. *Protein Sci.* **18**, 1388–1400 (2009).
211. C. J. Reedy, B. R. Gibney, Heme Protein Assemblies. *Chem. Rev.* **104**, 617–649 (2004).
212. D. W. Watkins, C. T. Armstrong, J. L. R. Anderson, De novo protein components for oxidoreductase assembly and biological integration. *Curr. Opin. Chem. Biol.* **19**, 90–98 (2014).
213. T. L. Poulos, Heme enzyme structure and function. *Chem. Rev.* **114**, 3919–3962 (2014).
214. D. A. Moffet, L. K. Certain, A. J. Smith, A. J. Kessel, K. A. Beckwith, M. H. Hecht, Peroxidase activity in heme proteins derived from a designed combinatorial library. *J. Am. Chem. Soc.* **122**, 7612–7613 (2000).
215. M. Faiella, O. Maglio, F. Nastri, A. Lombardi, L. Lista, W. R. Hagen, V. Pavone, De Novo Design, Synthesis and Characterisation of MP3, A New Catalytic Four-Helix Bundle Hemeprotein. *Chem. - A Eur. J.* **18**, 15960–15971 (2012).
216. M. I. Savenkova, J. M. Kuo, P. R. Ortiz de Montellano, Improvement of peroxygenase activity by relocation of a catalytic histidine within the active site of horseradish peroxidase. *Biochemistry.* **37**, 10828–10836 (1998).
217. P. Le, J. Zhao, S. Franzen, Correlation of heme binding affinity and enzyme kinetics of dehaloperoxidase. *Biochemistry.* **53**, 6863–6877 (2014).
218. G. Mclendon, R. Hake, Interprotein Electron Transfer. *Chem. Rev.* **92**, 481–490 (1992).
219. B. A. Fry, L. A. Solomon, P. Leslie Dutton, C. C. Moser, Design and engineering of a man-made diffusive electron-transport protein. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 513–521 (2016).
220. R. Stenner, J. W. Steventon, A. Seddon, J. L. R. Anderson, A de novo peroxidase is also a promiscuous yet stereoselective carbene transferase. *Proc. Natl. Acad. Sci.* **117**, 1419–1428 (2020).
221. H. Lodish, A. Berk, L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, in *Molecular Cell Biology* (W. H. Freeman and Company, New York, NY, USA, 4th edn., 2000).
222. M. Zarei, Application of nanocomposite polymer hydrogels for ultra-sensitive fluorescence detection of proteins in gel electrophoresis. *TrAC - Trends Anal. Chem.* **93**, 7–22 (2017).
223. D. Held, What are the differences between GPC, SEC, and GFC, and how do

8. References

- you get started with the technique? *The Column*. **14**, 2–8 (2018).
224. A. Hofmann, S. Clokie, Eds., *Principles and Techniques of Biochemistry and Molecular Biology* (Cambridge University Press, Cambridge, U.K., 8th edn., 2018).
225. J. M. Antosiewicz, D. Shugar, UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications. *Biophys. Rev.* **8**, 163–177 (2016).
226. M. K. Lykkegaard, A. Ehlerding, P. Hvelplund, U. Kadhane, M. B. S. Kirketerp, S. B. Nielsen, S. Panja, J. A. Wyer, H. Zettergren, A solet marker band for four-coordinate ferric heme proteins from absorption spectra of isolated Fe(III)-heme+ and Fe(III)-heme+ (his) ions in vacuo. *J. Am. Chem. Soc.* **130**, 11856–11857 (2008).
227. E. A. Berry, B. L. Trumpower, Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1–15 (1987).
228. I. Barr, F. Guo, Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *BIO-PROTOCOL*. **5** (2015).
229. N. J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2007).
230. A. Micsonai, F. Wien, L. Kernya, Y. H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3095–3103 (2015).
231. N. J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1**, 2527–2535 (2007).
232. J. Berg, J. Tymoczko, L. Stryer, *Biochemistry* (W. H. Freeman and Company, New York, NY, USA, 5th edn., 2002).
233. J. Clayden, N. Greeves, S. Warren, P. Wothers, *Organic Chemistry* (Oxford University Press, Oxford, U.K., 2001).
234. C. Fernández, G. Wider, TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.* **13**, 570–580 (2003).
235. A. Liu, L. Yao, Y. Li, H. Yan, TROSY of side-chain amides in large proteins. *J. Magn. Reson.* **186**, 319–326 (2007).
236. BIOSAXS GmbH, What is small angle X-ray scattering (SAXS) - BIOSAXS GmbH, (available at <http://biosaxs.com/technique.html>).
237. A. G. Kikhney, D. I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
238. C. R. Bagshaw, in *Encyclopedia of Biophysics* (Springer International Publishing, Heidelberg, Germany, 2013), pp. 2460–2466.
239. T. Sharpe, Circular Dichroism Unit Conversion. *Univ. Basel, Biozentrum*

- Biophys. Facil.* (2012), (available at https://www.biozentrum.unibas.ch/fileadmin/redaktion/05_Facilities/01_Technology_Platforms/BF/Protocols/CD_Unit_Conversion.pdf).
240. M. V. Petoukhov, P. V. Konarev, A. G. Kikhney, D. I. Svergun, ATSAS 2.1 - Towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40**, 223–228 (2007).
241. D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R. Hajizadeh, J. M. Franklin, C. M. Jeffries, D. I. Svergun, N. N. Semenov, ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* **50**, 1212–1225 (2017).
242. A. T. Tuukkanen, G. J. Kleywegt, D. I. Svergun, Resolution of ab initio shapes determined from small-angle scattering. *IUCrJ.* **3**, 440–447 (2016).
243. P. A. Mabrouk, The Use of Nonaqueous Media To Probe Biochemically Significant Enzyme Intermediates: The Generation and Stabilization of Horseradish Peroxidase Compound II in Neat Benzene Solution at Room Temperature. *Am. Chem. Soc.* **117**, 2141–2146 (1995).
244. M. Paravidino, M. J. Sorgedraeger, R. V. A. Orru, U. Hanefeld, Activity and enantioselectivity of the hydroxynitrile lyase MeHNL in dry organic solvents. *Chem. - A Eur. J.* **16**, 7596–7604 (2010).
245. F. C. Marhuenda-Egea, S. Piera-Velázquez, C. Cadenas, E. Cadenas, Reverse micelles in organic solvents: A medium for the biotechnological use of extreme halophilic enzymes at low salt concentration. *Archaea.* **1**, 105–111 (2002).
246. J. R. Chipperfield, *Non-aqueous Solvents* (Oxford Chemistry Primers, Oxford University Press, Oxford, U.K., 1999).
247. J. P. Cerón-Carrasco, D. Jacquemin, C. Laurence, A. Planchat, C. Reichardt, K. Sraïdi, Solvent polarity scales: Determination of new ET(30) values for 84 organic solvents. *J. Phys. Org. Chem.* **27**, 512–518 (2014).
248. F. Shakerizadeh-Shirazi, B. Hemmateenejad, A. M. Mehranpour, Determination of the empirical solvent polarity parameter ET(30) by multivariate image analysis. *Anal. Methods.* **5**, 891–896 (2013).
249. J. L. Finney, B. J. Gellatly, I. C. Golton, J. Goodfellow, Solvent effects and polar interactions in the structural stability and dynamics of globular proteins. *Biophys. J.* **32**, 17–33 (1980).
250. Sigma Aldrich, Properties of Solvents Table, (available at <https://www.sigmaaldrich.com/chemistry/stockroom-reagents/learning-center/technical-library/solvent-properties.html>).
251. Alfa Aesar, Trifluoroethanol, (available at <https://www.alfa.com/en/search/?q=trifluoroethanol>).
252. National Center for Biotechnology Information, PubChem Compound Summary for CID 6342, Acetonitrile. *PubChem Database*, (available at <https://pubchem.ncbi.nlm.nih.gov/compound/Acetonitrile>).

8. References

253. Louisiana State University, Dielectric Constant, (available at <https://macro.lsu.edu/HowTo/solvents/DielectricConstant.htm>).
254. University of Washington, Dielectric Constants of Common Solvents (2007), (available at https://depts.washington.edu/eoopic/linkfiles/dielectric_chart%5B1%5D.pdf).
255. C. Reichardt, T. Welton, *Solvents and Solvent Effects in Organic Chemistry* (Wiley-VCH Verlag, Weinheim, Germany, 4th edn., 2010).
256. Drugbank, Trifluoroethanol, (available at <https://www.drugbank.ca/drugs/DB03226>).
257. E. C. Meister, M. Willeke, W. Angst, A. Togni, P. Walde, Confusing quantitative descriptions of brønsted-lowry acid-base equilibria in chemistry textbooks - A critical review and clarifications for chemical educators. *Helv. Chim. Acta.* **97**, 1–31 (2014).
258. M. Mugnai, G. Cardini, V. Schettino, C. J. Nielsen, Ab initio molecular dynamics study of aqueous formaldehyde and methanediol. *Mol. Phys.* **105**, 2203–2210 (2007).
259. F. A. Carey, R. J. Sundberg, *Advanced Organic Chemistry Part A: Structure and Mechanisms* (Springer International Publishing, New York, NY, USA, 5th edn., 2007).
260. S. M. Kelly, T. J. Jess, N. C. Price, How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics.* **1751**, 119–139 (2005).
261. S. Kelly, N. Price, The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Protein Pept. Sci.* **1**, 349–384 (2005).
262. R. W. Woody, Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins. *Biopolymers.* **17**, 1451–1467 (1978).
263. G. Yang, S. Zhang, J. Hu, M. Fujiki, G. Zou, The chirality induction and modulation of polymers by circularly polarized light. *Symmetry.* **11**, 474–493 (2019).
264. N. J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1**, 2527–2535 (2007).
265. Y. Wei, A. A. Thyparambil, R. A. Latour, Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230 nm. *Biochim. Biophys. Acta - Proteins Proteomics.* **1844**, 2331–2337 (2014).
266. J. H. Bredehöft, N. C. Jones, C. Meinert, A. C. Evans, S. V. Hoffmann, U. J. Meierhenrich, Understanding photochirogenesis: Solvent effects on circular dichroism and anisotropy spectroscopy. *Chirality.* **26**, 373–378 (2014).
267. N. Sreerama, R. W. Woody, Computation and Analysis of Protein Circular Dichroism Spectra. *Methods Enzymol.* **383**, 318–351 (2004).
268. A. J. Miles, B. A. Wallace, Circular dichroism spectroscopy of membrane

- proteins. *Chem. Soc. Rev.* **45**, 4859–4872 (2016).
269. K. Ghosh, A. M. Thompson, R. A. Goldbeck, X. Shi, S. Whitman, E. Oh, Z. Zhiwu, C. Vulpe, T. R. Holman, Spectroscopic and biochemical characterization of heme binding to yeast Dap1p and mouse PGRMC1p. *Biochemistry*. **44**, 16729–16736 (2005).
270. S. S. Huang, R. L. Koder, M. Lewis, A. J. Wand, P. L. Dutton, The HP-1 maquette: From an apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5536–5541 (2004).
271. S. Franzen, S. G. Boxer, On the origin of heme absorption band shifts and associated protein structural relaxation in myoglobin following flash photolysis. *J. Biol. Chem.* **272**, 9655–9660 (1997).
272. Y. H. Chen, Determination of the helix and β -form of proteins in aqueous solution by circular dichroism. *Biochemistry*. **13**, 3350–3359 (1974).
273. J. M. Dinitto, J. M. Kenney, Noise characterization in circular dichroism spectroscopy. *Appl. Spectrosc.* **66**, 180–187 (2012).
274. P. Atkins, J. de Paula, *Atkins' Physical chemistry* (Oxford University Press, Oxford, U.K., 9th edn., 2010).
275. D. Wang, K. Chen, J. L. Kulp, P. S. Arora, Evaluation of biologically relevant short α -helices stabilized by a main-chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* **128**, 9248–9256 (2006).
276. R. O. Crooks, T. Rao, J. M. Mason, Truncation, randomization, and selection: Generation of a reduced length c-jun antagonist that retains high interaction stability. *J. Biol. Chem.* **286**, 29470–29479 (2011).
277. A. D. Nagi, L. Regan, An inverse correlation between loop length and stability in a four-helix-bundle protein. *Fold. Des.* **2**, 67–75 (1997).
278. K. Ghosh, K. A. Dill, Theory for protein folding cooperativity: Helix bundles. *J. Am. Chem. Soc.* **131**, 2306–2312 (2009).
279. M. Munson, L. Regan, R. O'Brien, J. M. Sturtevant, Redesigning the hydrophobic core of a four-helix-bundle protein. *Protein Sci.* **3**, 2015–2022 (1994).
280. A. Fersht, *Structure and mechanism in protein science* (W. H. Freeman and Company, New York, NY, USA, 3rd edn., 1999).
281. B. A. Otieno, C. E. Krause, J. F. Rusling, in *Methods in Enzymology* (Academic Press Inc., Cambridge, MA, USA, 2016), vol. 571, pp. 135–150.
282. F. Galgani, Y. Cadiou, G. Bocquene, Routine determination of enzyme kinetics using Plate reader. *Biotechnol. Bioeng.* **38**, 434–437 (1991).
283. R. E. Childs, W. G. Bardsley, The steady state kinetics of peroxidase with 2,2'-azino-di-(3-ethylbenzthiazoline-6-sulphonic acid) as chromogen. *Biochem. J.* **145**, 93–103 (1975).
284. R. A. Copeland, in *Evaluation of Enzyme Inhibitors in Drug Discovery* (John

8. References

- Wiley & Sons, Inc., Hoboken, NJ, USA, 2nd edn., 2013), pp. 493–495.
285. L. K. Folkes, L. P. Candeias, Interpretation of the reactivity of peroxidase compounds I and II with phenols by the Marcus equation. *FEBS Lett.* **412**, 305–308 (1997).
 286. A. N. P. Hiner, E. L. Raven, R. N. F. Thorneley, F. García-Cánovas, J. N. Rodríguez-López, Mechanisms of compound I formation in heme peroxidases. *J. Inorg. Biochem.* **91**, 27–34 (2002).
 287. D. P. Hong, M. Hoshino, R. Kuboi, Y. Goto, Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *J. Am. Chem. Soc.* **121**, 8427–8433 (1999).
 288. F. Pintus, A. Mura, A. C. Rinaldi, A. Contini, D. Spanò, R. Medda, G. Floris, Activity and structural changes of *Euphorbia characias* peroxidase in the presence of trifluoroethanol. *Protein J.* **27**, 434–439 (2008).
 289. F. Ataei, S. Hosseinkhani, Impact of trifluoroethanol-induced structural changes on luciferase cleavage sites. *J. Photochem. Photobiol. B Biol.* **144**, 1–7 (2015).
 290. Y. X. Zhang, S. L. Yan, H. M. Zhou, Inactivation and conformational changes of aminoacylase in trifluoroethanol solutions. *J. Protein Chem.* **15**, 631–637 (1996).
 291. J. P. Klinman, The power of integrating kinetic isotope effects into the formalism of the Michaelis-Menten equation. *FEBS J.* **281**, 489–497 (2014).
 292. A. Sen, A. Kohen, Enzymatic tunneling and kinetic isotope effects: Chemistry at the crossroads. *J. Phys. Org. Chem.* **23**, 613–619 (2010).
 293. K. B. Wiberg, The Deuterium Isotope Effect. *Chem. Rev.* **55**, 713–743 (1955).
 294. K. K. Khan, M. S. Mondal, L. Padhy, S. Mitra, The role of distal histidine in peroxidase activity of myoglobin. *Eur. J. Biochem.* **257**, 547–555 (1998).
 295. N. Mogharrab, H. Ghourchian, M. Amininasab, Structural stabilization and functional improvement of horseradish peroxidase upon modification of accessible lysines: Experiments and simulation. *Biophys. J.* **92**, 1192–1203 (2007).
 296. M. Khajepour, I. Rietveld, S. Vinogradov, N. V. Prabhu, K. A. Sharp, J. M. Vanderkooi, Accessibility of Oxygen With Respect to the Heme Pocket in Horseradish Peroxidase. *Proteins Struct. Funct. Genet.* **53**, 656–666 (2003).
 297. V. Higman, 1H-15N HSQC. *Protein NMR* (2012), (available at <https://www.protein-nmr.org.uk/solution-nmr/spectrum-descriptions/1h-15n-hsqc/>).
 298. J. Cavanagh, W. J. Fairbrother, A. G. Palmer, N. J. Skelton, M. Rance, *Protein NMR Spectroscopy* (Academic Press, Cambridge, MA, USA, 2nd edn., 2007).
 299. H.-S. Kang, M. Sattler, Capturing dynamic conformational shifts in protein–ligand recognition using integrative structural biology in solution. *Emerg. Top. Life Sci.* **2**, 107–119 (2018).
 300. T. M. Weiss, in *BioSAXS Workshop* (Stanford, CA, USA, 2016; <https://www->

- ssrl.slac.stanford.edu/~saxs/download/weiss_intro.pdf).
301. L. Boldon, F. Laliberte, L. Liu, Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev.* **6**, 25661 (2015).
 302. D. I. Svergun, M. H. J. Koch, Small-angle scattering studies of biological macromolecules in solution. *Reports Prog. Phys.* **66**, 1735–1782 (2003).
 303. D. Durand, C. Vivès, D. Cannella, J. Pérez, E. Pebay-Peyroula, P. Vachette, F. Fieschi, NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* **169**, 45–53 (2010).
 304. R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grütter, H. K. Privett, S. L. Mayo, D. Hilvert, Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature.* **503**, 418–421 (2013).
 305. H. Lechner, N. Ferruz, B. Höcker, Strategies for designing non-natural enzymes and binders. *Curr. Opin. Chem. Biol.* **47**, 67–76 (2018).
 306. K. J. Grayson, J. R. Anderson, The ascent of man(made oxidoreductases). *Curr. Opin. Struct. Biol.* **51**, 149–155 (2018).
 307. N. Hirota, K. Mizuno, Y. Goto, Cooperative α -helix formation of β -lactoglobulin and melittin induced by hexafluoroisopropanol. *Protein Sci.* **6**, 416–421 (1997).
 308. M. Ingr, E. Kutáľková, J. Hrnčířík, R. Lange, Equilibria of oligomeric proteins under high pressure – A theoretical description. *J. Theor. Biol.* **411**, 16–26 (2016).
 309. S. Marchal, A. Marabotti, M. Staiano, A. Varriale, T. Domaschke, R. Lange, S. D’Auria, Under Pressure That Splits a Family in Two. The Case of Lipocalin Family. *PLoS One.* **7**, e50489 (2012).
 310. N. Khatoon, A. Jamal, M. I. Ali, Polymeric pollutant biodegradation through microbial oxidoreductase: A better strategy to safe environment. *Int. J. Biol. Macromol.* **105**, 9–16 (2017).
 311. C. D. Putnam, M. Hammel, G. L. Hura, J. A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285 (2007).



EPSRC

Engineering and Physical Sciences
Research Council

