Characterization of Structural and Functional Properties of Class IB Hydrophobins

by

Calem Austin Kenward

Submitted in partial fulfilment of the requirements

for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
March 2019

Dedication

To my mother Merlene,

Who taught me – *if you're going to do something, do it right.*

Table of Contents

List of Tables

List of Figures

Abstract

Hydrophobins are low molecular weight self-assembling proteins secreted by fungi and are active at hydrophobic-hydrophilic interfaces. Hydrophobins may undergo structural rearrangement and oligomerize to form rodlets, which are an insoluble functional amyloid. To better understand which sequence characteristics determine hydrophobin properties, I have characterized the structure and properties of class IB hydrophobins from various fungi: *Serpula lacrymans* (SL1), *Wallemia ichthyophaga* (WI1), and *Phanerochaete carnosa* (PC1). I determined the high-resolution structure of each hydrophobin using NMR spectroscopy. This revealed that these hydrophobins all share structural features despite their dissimilar sequences. The core conserved feature is a four strand anti-parallel β-sheet that is connected by three loop sequences (L1-L3). In all hydrophobins the β-sheet folds upon itself to form a β-barrel structure. Spectroscopic amyloid formation assays indicate that each hydrophobin has differing propensities to form rodlets. Overall, this work establishes a correlation between the sequence, structure, and self-assembly properties of hydrophobins.

List of Abbreviations and Symbols Used

δ:	Chemical shift
AFM:	Atomic force microscopy
BRMB:	Biological Magnetic Resonance Data Bank
CNS:	Crystallography and NMR system
COSY:	Homonuclear correlation spectroscopy
CSI:	Chemical shift index
dH2O:	Distilled water
D_2O:	Deuterium Oxide
DANGLE:	Dihedral ANgles from Global Likelihood Estimates
DewA:	Hydrophobin from *Aspergillus nidulans*
DSS:	2,2-dimethyl-2-silapentane-5-sulfonate
DTT:	Dithiothreitol
EAS:	Hydrophobin from *Neurospora crassa*
EM:	Electron microscopy
ESI-MS:	Electrospray ionization mass spectrometry
FPLC:	Fast protein liquid chromatography
GSH:	Glutathione-reduced
GSSH:	Glutathione-oxidized
HetNOE:	Heteronuclear NOE
HFBI/HFBII:	Hydrophobins from *Trichoderma reesei*
HGFI:	Hydrophobin from *Grifola frondosa*
HPLC:	high performance liquid chromatography

HOPG: Highly oriented pyrolytic graphite

HSQC: Heteronuclear single quantum coherence

IMAC: Immobilized metal affinity chromatography

IPTG: Isopropyl β-D-1-thiogalactopyranoside

MES: 2-(N-morpholino)ethanesulfonic acid

MPG1: Hydrophobin from *Magnaporthe griseaclass*

NC2: Hydrophobin-2 from *Neurospora crassa*

NMR: Nuclear magnetic resonance

NOE: Nuclear Overhauser effect

NOESY: Nuclear Overhauser effect spectroscopy

OD_{600}: Optical density at 600 nm

PC1: Hydrophobin from *Phanerochaete carnosa*

PDB: Protein Data Bank

r: Distance between two nuclei

RMSD: Root mean square deviation

RP-HPLC: Reverse-phase high performance liquid chromatography

SC16: Hydrophobin 16 from *Schizophyllum commune*

SC3: Hydrophobin 3 from *Schizophyllum commune*

SL1: Hydrophobin from *Serpula lacrymans*

TEV: Tobacco Etch Virus

TFA: Trifluoroacetic acid

Th: Thombin

ThT: Thioflavin T

TOCSY:	Total correlation spectroscopy
Tris:	Trisaminomethane
Trx:	Thioredoxin
UV/Vis:	Ultraviolet-visible spectrophotometry
WI1:	Hydrophobin from *Wallemia ichthyophaga*

Acknowledgements

I would first like the thank my supervisor Dr. David Langelaan, for his guidance throughout my degree. He has been incredibly supportive and patient during my time in the lab, and has given me the truly wonderful experience of being able to freely explore whatever research interests I had in mind. His passion for research inspires those around him, including me, and makes the lab a truly exciting place to work. Since joining the lab, I have felt as part of a team and I strongly believe that Dave has always has everyone's best interests in mind. I sincerely doubt I would have achieved even half of what I did if I had any other supervisor, so thank you.

I am grateful towards my committee members for their guidance throughout my degree, Dr. Vanya Ewart and Dr. Jan Rainey. I value the feedback on my thesis from my committee members and Dr. Andrew Roger for their thoughtful comments on this work during my defense. I would also like to acknowledge Jan's guidance throughout my master's and during my time I spent in his lab for my Honours. He has always been very approachable and his passion for structural biology is a major factor in why I took an interest in research in the first place.

I have received help from numerous other people. I thank everyone who has inspired me with their passion for science, taught me the skills I needed to succeed and been supportive during my undergraduate and graduate studies. I cannot include everyone but thank you to Dr. Kyungsoo Shin for his invaluable advice throughout the years. I greatly appreciate the extensive help with NMR which made this work possible: Dr. Mike Lumsden at the Nuclear Magnetic Resonance Research Resource at Dalhousie University, Ian Burton at the National Research Council – Halifax, and especially Dr. Tara Sprules at the Québec/Eastern Canada High Field NMR Facility. I thank Dr. Laurent Kreplak for his help with the AFM in this work. Thank you to all of the members of Rainey and Langelaan lab groups for their invaluable feedback during meetings and in the lab. I have received funding through the Beatrice Hunter Cancer Research Institute and Dalhousie University Faculty of Medicine, making graduate school possible for me.

Lastly, I am indebted to my family for being a rock solid foundation which I know I can always rely on. To my parents, Merlene and Gary for the love and support; to Briean, Tegan and Reilly, my sisters who inspire me; to Hadfield and Theodore for their support while writing; and to my partner Ms. Melanie MacArthur for her continued love, invaluable support, and for being the person who has always been there for me when I needed someone to listen. To everyone – *thank you!*

Chapter 1: Introduction

Hydrophobins are a large and diverse group of low-molecular-weight (5–20 kDa) proteins that are produced by fungi, self-assemble into larger structures, and are among the most surface active proteins known (as they tend to reduce the surface tension of a liquid in which they are dissolved).[1] Hydrophobins are an excellent example of functional amyloids, where the self-assembled amyloid form of hydrophobins has a biological role that is not provided by the monomeric form. These amyloid assemblies give hydrophobins unique functional properties, with many exciting potential applications, and are therefore of considerable interest for study.[2] However, despite extensive research into the mechanistic basis of hydrophobin properties, the exact process of self-assembly remains unresolved. Discovery, modification, and implementation of hydrophobins in industry is hindered by the lack of an understanding of hydrophobins function. The overall objective of this work was to better understand the hydrophobin self-assembly process and correlate the sequence, structure, and function of fungal hydrophobins. To do this, I used heterologous expression to produce several distinct hydrophobins (Chapter 2), determine the atomic-resolution structures of select hydrophobins (Chapter 3), and relate the self-assembly ability of hydrophobins to their structure and properties (Chapter 4).

In this introduction I provide a brief summary of the current understanding of hydrophobins. I first introduce fundamental fungal biology and describe hydrophobin expression in fungi. I then discuss the sequence composition and properties of different classes and subclasses of hydrophobins. This is followed by an examination of published atomic-resolution hydrophobin structures and how these structures relate to current models of rodlet formation. Lastly, I discuss how the features of hydrophobins make them excellent targets for commercialization and present some of their potential applications in industry and medicine.

1.1 The Biological Roles of Hydrophobins

Fungi are heterotrophic terrestrial eukaryotes, with two types of growth morphologies: a unicellular yeast form and multicellular filamentous forms. Filamentous fungi grow as tubular, elongated, and thread-like structures called hyphae, which contain multiple nuclei and extend by growing at their tips. This growth process describes the majority of what is commonly described as fungi. In most fungi these hyphae are the main mode of vegetative growth and are collectively called a mycelium. As the primary structure of fungi, hyphae interact with their surroundings, collect nutrients, and propagate. Many fungi have also evolved to produce upwardly projecting aerial hyphae and fruiting structures. Additionally, fungi have developed adaptations to spread spores to new locations and a capacity to withstand the extreme desiccation or near submersion encountered in terrestrial environments.[3] These adaptations allow filamentous fungi to thrive in a wide range of environments and ecological niches throughout their life cycle, necessitating increased metabolic plasticity and the ability of these fungi to modify or adapt to their local environment.

To manipulate their surroundings and adapt to environmental challenges, fungi produce and secrete self-assembling proteins named hydrophobins to tailor interfaces to fungal life. Hydrophobins are abundant and unique to filamentous fungi of the Ascomycota and Basidiomycota phyla. Ascomycetes, or sac fungi, are classified based on the production of sexual spore-bearing cells called asci which vary in shape from cylindrical to spherical.[3] The Ascomycota division includes a wide diversity of fungi that includes *Penicillium*, from which antibiotics are derived, to the commonly consumed morel mushrooms of the *Morchella* genus.[4] In contrast, the spores of basidiomycetes develop on projections that grow out from microscopic cells called basidia, rather than being enveloped within cells. This division includes common wood rot fungi such as *Serpula lacrymans* (Figure 1.1 A) and *Agaricus bisporus* or button mushrooms, which are commonly consumed around the world.[5]

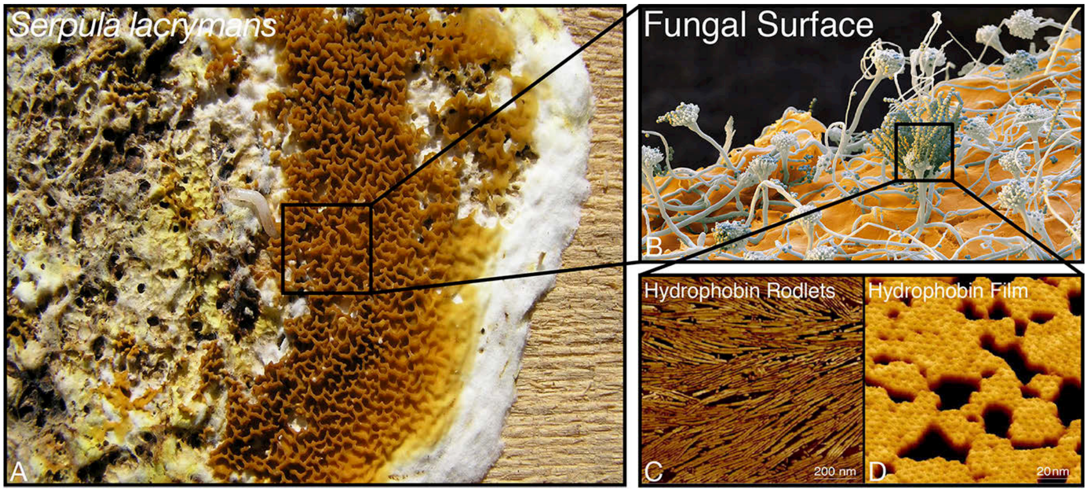


Figure 1.1. Hydrophobins assembled on a spore surface have distinct morphological organization at nanoscopic scales.
(A) *Serpula lacrymans* is found in the fungal division Basidiomycota and causes timber rot[6] (B) Spores of *Emericella nidulans* visible by SEM are covered by hydrophobin rodlets making them waterproof and more durable. Magnification 400 :1, 12 cm in width[7] (C) microscopic surface of the HGFI hydrophobin from *Grifola frondosa* shows the characteristic rodlet appearance of short but elongated repeating structures.[8] (D) Surface assembly of the hydrophobin HFBI, showing an organized honeycomb like film structure.[8]

Protein-based coatings were first observed on the surface of fungal spores in electron micrographs over 50 years ago.[9] However, the identification of hydrophobins forming a protective coating for spores occurred decades later.[10] Hydrophobins are often are secreted in abundance and a single fungus species can produce many different varieties in response to changing environmental conditions or life stages, including vegetative hyphae, sporulating cultures, and fruiting bodies such as mushrooms.[1] For example, the tomato pathogen *Cladosporium fulvum* has six hydrophobin genes identified by Lacroix et al.[11] These genes had different expression patterns, such as the hydrophobin HCf-6, which was observed only during infection of tomato plants and had a direct role in adhesion of the fungus to its growth substrate.[12]

Currently, the roles of hydrophobins in nature are expansive and applicable wherever surface modification is required by fungi. For example, the vegetative root-like hyphae of filamentous fungi growing in moist environments are hydrophilic and do not have hydrophobins on their surface, yet the aerial hyphae (Figure 1.1 B) and the asexual spores (conidia) are hydrophobic, due to the presence of hydrophobins which prevents wetting and aids dispersion into the air.[13] Hydrophobins have been observed in aerial structures,[10] at the interface between pathogenic fungi and host plants,[14] in fungal cell walls,[15] in protective waterproofing coats of fungal spores,[16] and in films on the surface and air cavities in fruiting bodies.[17]

Hydrophobins are secreted from fungi as soluble monomers. Upon reaching a hydrophobic-hydrophilic interface, they aggregate spontaneously to form an amphipathic monolayer, often adopting either a repeating “rodlet” or a honeycomb film morphology (Figure 1.1 C,D). This rodlet layer composed to short acts as a natural surfactant and reduces the surface tension of the surrounding medium, allowing the fungus to breach air-water interfaces and to more easily grow hyphae.[1,17,18] The films produced by hydrophobins on fungal structures can also provide protective coatings by making the surface hydrophobic. Coatings have been reported on the caps of mushrooms,[19] in aerial conidia (spores),[20,21] and on hyphae where a hydrophobic coating was proposed to have a role protecting against desiccation and wetting, and aiding in the dispersal of spores into the air [17,22,23] or in water.[24] Disruption of the rodA gene from *Aspergillus nidulans* and the EAS gene from *Neurospora crassa*, for example, resulted in mutants with spores which are easily wetted and lack the hydrophobic rodlet protein that normally coats the conidium.[20] Similarly, disruption of the SC3 gene caused formation of hydrophilic aerial hyphae and, under certain conditions, prevented formation of any aerial hyphae by *Schizophyllum commune*.[25–27]

Hydrophobins are also involved in fungal pathogenesis, mediating attachment of the fungal infection structures (appressorium) to their targets either as structural components of the appressorium cell wall or by binding to and modifying host surfaces such as plant hosts or insect cuticles.[28–30] Deletion of the hydrophobin MPG1 or MHP1 genes resulted in a mutant of *Magnaporthe grisea* (rice blast fungus) with reduced virulence and a reduced capacity to infect and colonize a susceptible rice cultivar.[31,32] Surface coatings of spores have also been shown to have a critical role in masking the immunogenicity of airborne fungal spores,[33] by ensuring that pathogen-associated molecular patterns are not recognized by innate and adaptive immune cells, thus preventing the activation of host immune system, inflammation, and host tissue damage.[34–36]

1.2 Sequence Diversity of Hydrophobins

Hydrophobins were first identified as mRNAs abundantly transcribed during developmental processes such as sporulation, fruiting body formation, or fungal infection of plants and insects in *Schizophyllum commune*, without knowing the identity and nature of the proteins.[10] Based on the deduced protein sequences, Wessels et al. introduced the name hydrophobin for these relatively small fungal proteins of about 10 kDa in size.[10] The first hydrophobin isolated was SC3 from *Schizophyllum commune* and remains one of the most heavily studied hydrophobins to date, with many of the defining features of hydrophobins being first described in studies examining SC3.[25,37–41] Since then, the hydrophobin family has expanded to include a myriad of predicted proteins,[42] with many being continually discovered in newly sequenced fungal genomes .[43] However, despite the thousands sequences that have been predicted, the roles of individual proteins remain largely unclear.[44]

Other than having high hydrophobic residue content, the only conserved feature among all hydrophobins is a pattern of four disulfide bonds formed by eight universally conserved cysteine amino acid residues (Figure 1.2).[8,42] The disulphide bonds form between C1–C6, C2–C5, C3–C4, and C7–C8, with C2 and C3 as well as C6 and C7 being adjacent to each other in the amino acid sequence.[45] Disruption of these bonds in SC3 results in insoluble protein, indicating their importance in proper hydrophobin folding.[46] The rest of the cysteine residues have varying numbers of intervening residues (Figure 1.2). The inter-cysteine sequences vary greatly in length between different hydrophobins, in some cases being nearly absent. Overall, hydrophobins as a family show exceptionally poor sequence conservation, especially outside of the core region spanning

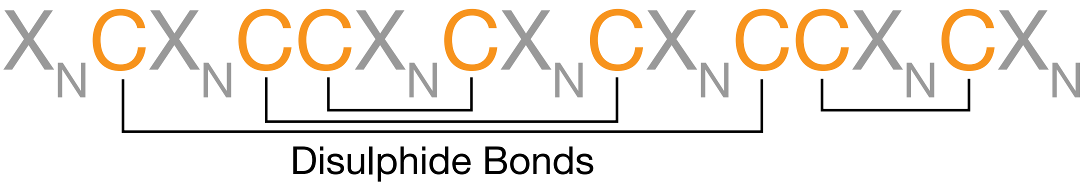


Figure 1.2. A distinct disulphide bonding pattern is the only universally conversed hydrophobin feature.
Hydrophobins contain a conserved pattern of eight cysteine residues that form four disulphide bonds. Inter-cysteine sequences can vary greatly in sequence composition and length, allowing hydrophobins to have diverse properties.

from the first to eighth cysteine residues.[47,48] The sequence characteristics of hydrophobins suggest that while cysteine residues are critical for hydrophobin function, it is possible that other residues vary substantially to generate hydrophobins with distinct properties.[44,49]

Hydrophobins can be divided into two classes, class I and class II, based on hydrophobic residue clustering, the classification of the organism of origin, and functional properties. In 1994, Wessels et al. proposed this class division after aligning the cysteine residues and examining patterns of clustering of hydrophobic and hydrophilic residues of the nine hydrophobins then known.[47] Functionally, these classes are very distinct, class I hydrophobins self-assemble at interfaces to form highly stable films that can only be disassociated by strong acids, such as trifluoroacetic acid or formic acid.[50–52] These films take on a characteristic fibrillar rodlet morphology (e.g. Figure 1.1 C) that is approximately 10 nm in height and has come to define the class.

These rodlet films are highly stable layers that can withstand detergents, organic solvents and high temperatures.[53,54] Class II assemblies are comparatively less stable, can be disassociated by detergent-alcohol mixtures or pressure, and form repeating films that are less well defined morphologically.[55–57] Despite the clear sequence and morphological differences, no obvious distinction between the functional role of class I and class II hydrophobins within the fungal life cycle has been determined.[58] To date, class II hydrophobins have been observed only in ascomycetes whereas class I hydrophobins are observed in both basidiomycetes and ascomycetes.[1]

Class I hydrophobins vary greatly in composition and length (Figure 1.3).[59] The intervening regions between cysteine residues are especially variable between class I members. For example, the sequence between cysteine three and cysteine four varies from four residues in HYD3 from *Gibberella moniliformis* to 44 residues AaPRI2 from *Agrocybe aegerita*. The sequence between cysteine four and cysteine five varies from eight residues in EAS from *Neurospora crassa* to 23 residues DewA from *Aspergillus nidulans*.[58] Yet, despite the low sequence similarity, class I hydrophobins from different fungal species could partially restore functionality to a *Magnaporthe griseaclass* mutant with a I hydrophobin gene (MPG1) disruption.[31] This suggests that regardless of low sequence similarity, hydrophobins constitute a closely related group proteins that have similar functions.

In contrast to class I, class II hydrophobins are generally smaller and have much higher sequence similarity (Figure 1.3). Unlike class I hydrophobins, the intervening regions between cysteine three and four and between cysteines four and five are fully conserved in class II hydrophobins, with the other inter-cysteine regions also having higher conservation (Figure 1.3 B).[1] Recently several proteins have been identified that do not neatly fit into either of the two classes, such as the DewA and DewD hydrophobins from *Aspergillus nidulans* due to the presence of features found in both class I and class II hydrophobins. Other intermediate forms of hydrophobins with distinct physicochemical characteristics may have also been overlooked due to their dissimilarity to established class I and class II members.[43,49,60]

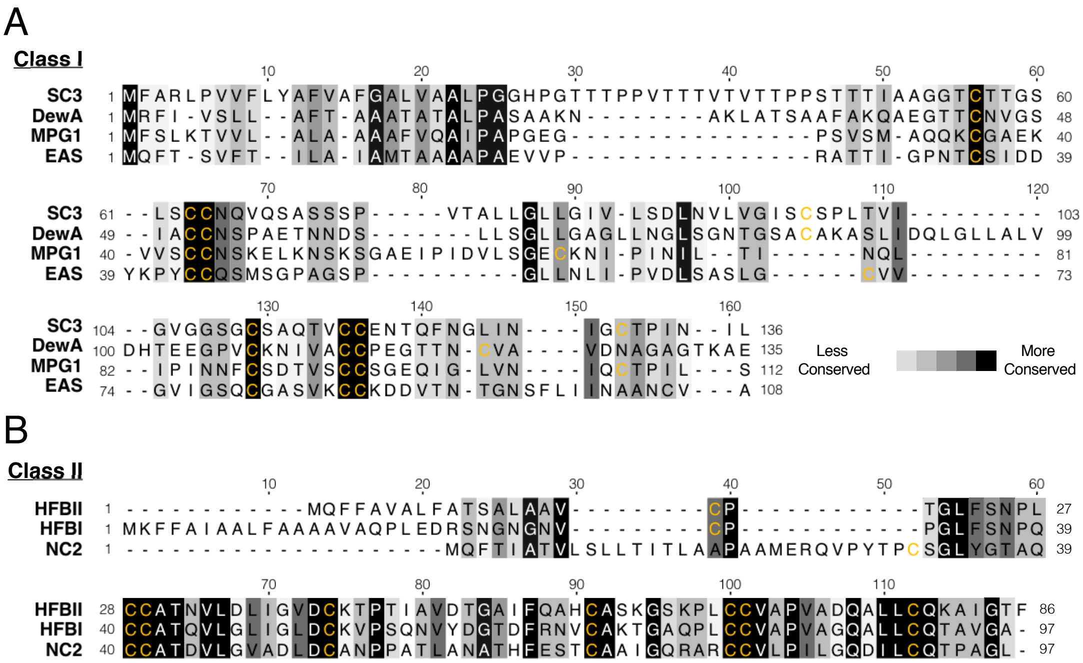


Figure 1.3. Class II hydrophobins have a higher degree of sequence conservation than class I hydrophobins.

Sequence alignment of several class I (A) and class II (B) hydrophobins shows that class I members have far greater sequence variability with large insertions and deletions. Cysteine residues are coloured yellow and conserved in all hydrophobins. Alignment and conservation values were generated using T-Coffee Expresso, which considers structural data to generate sequence alignments.[61,62] Published structures from the PDB for DewA[63], MPG1[64] and EAS[45] for class I, and HFBII, HFBI[65] and NC2[59] for class II, were used for each alignment respectively.

Difficulties determining class I hydrophobin features has led to the proposition for further subdivision of the class. In 2005, Linder et al. proposed the subdivision of class I into class IA and class IB on the basis of grouping those hydrophobins produced by ascomycetes and those form basidiomycetes fungi respectively.[1] The differences between class IA and class IB remained poorly described until 2017 when the subdivision was further supported by principal component analysis of hydrophobin sequences.[42] Using a database of confirmed and predicted hydrophobin sequences from the genomes and transcriptomes of 215 unique filamentous fungi (72% Basidiomycota, 28% Ascomycota), analysis of 1046 canonical sequences revealed that class IA and IB hydrophobins have distinct sequence features. The principal distinguishing feature of the class IA and IB subdivisions is the number of residues between cysteine residues. In class IA, regions between C3-C4, and C4-C5 residues are on average, 40 and 22 residues in length, while in class IB they are on average 32 and 13 residues.[42,66]

Class IB hydrophobins are reported to have a higher degree of sequence conservation in comparison to class IA. Sequence comparison of the class IB hydrophobins HGFI from *Grifola frondosa*, VMH2 from *Pleurotus ostreatus*, SC3 from *Schizophyllum commune*, and SC16 also from *Schizophyllum commune* confirmed that class IB hydrophobins contain similar length inter-cysteine sequences and a higher degree of sequence conservation beyond the conserved disulphide bonding pattern. For example, class IB SC3 has 43.8% and 31.5% sequence identity with Vmh2 and ABH1, respectively. This similar to the class IA homologues RodA, RodB and RodC which come from the same organism, *Aspergillus fumigatus* and share 36% of sequence identity in the region between the first and eighth cysteine residues. When comparing nonhomologous sequences such as those of DewA, EAS, MPG1, and RodA, even higher sequence and inter-cysteine spacing variability are observed, with DewA having only 17.6%, 4.2% and 11.5% identity to EAS, MPG1 and RodA respectively.[2]

1.3 Hydrophobin Structure

Similar to the variability observed in residue composition and inter-cysteine sequence lengths, the solution structures of monomeric hydrophobins are quite variable (Figure 1.4). To date, several high resolution hydrophobin structures have been published, including the class I hydrophobins EAS, DewA, MPG1, and RodA,[45,63,67,68] and the class II hydrophobins HFBI and HFBII from *Trichoderma reesei*, and NC2 from *Neurospora crassa*.[65,69–71] Structural characterization of hydrophobins is heavily biased towards those of ascomycete origin, and therefore class IA, despite hydrophobins from both phyla being heavily studied.

Class IA hydrophobins adopt diverse tertiary structures. The inter-cysteine regions connecting the β-sheet strands (L_1–L_3) often vary in sequence, length, structure, and dynamics. EAS, DewA and MPG1 all contain drastically different secondary structuring within in their respective loop regions, as summarized in Figure 1.5. For example, L_1 and L_3 of EAS are unstructured and L_2 contains a β-sheet,[45] while DewA and MPG1 contain α-helical elements in these loops that are absent in EAS.[63,64] Compared to class I hydrophobins, class II hydrophobins are more similar and generally contain similar structural elements including a more compact structure, with short L_1 and L_3 domains, and an α-helix in L_2.[59,69,70] The division of class I and class II is evident when comparing published protein structures as shown in Figure 1.4 and Figure 1.5. Both types of hydrophobin share a four-stranded β-barrel core region but large disordered loops between C3–C4 and C7–C8 (L_1 and L_3 regions) are found in class I hydrophobins but not class II.

Despite their structural heterogeneity, hydrophobins have some consistent general features. A core β-sheet region is always present with inter-cysteine sequences forming peripheral disordered loop regions denoted L_1-L_3.[13] The globular structures of hydrophobins are also stabilized by highly conserved disulphide bridges.[54] Two disulphide bonds are typically located within the β-barrel covalently linking β_1 to β_2 and β_3 to β_4 and the two remaining disulphide bonds connect the N-terminal tail to β_3 and β_1 to the second loop region. These bonds effectively cross-link the protein, pairing adjacent β_4 strands and peripheral loops together, resulting in efficient stabilization of the globular shape of the hydrophobin.[1]

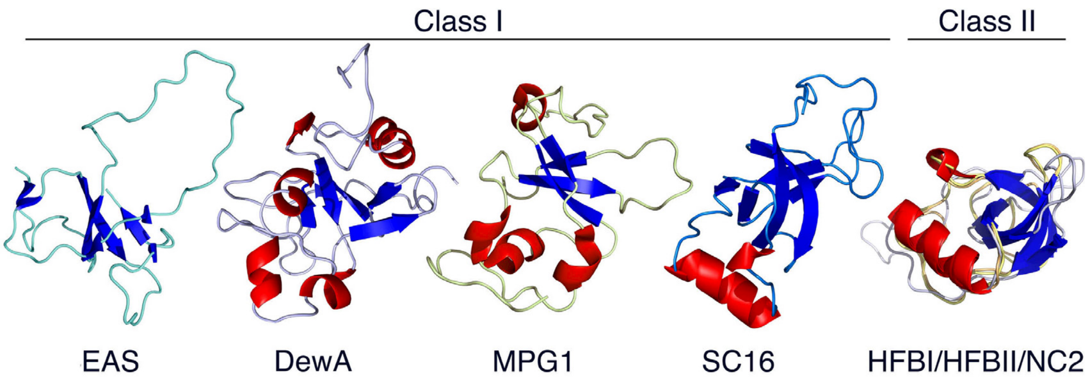


Figure 1.4. Class I hydrophobins have greater structural diversity than class II hydrophobins.
Ribbon representations of the class I hydrophobins EAS (PDB ID: 2FMC),[45] DewA (PDB ID: 2LSH),[63] MPG1 (PDB ID: 2N4O),[64] SC16 (PDB ID: 2NBH),[42] and the class II hydrophobins HFBI (PDB ID: 2FZ6),[65] HFBII (PDB ID: 2B97),[69] and NC2 (PDB ID: 4AOG),[59] with β-strands and α-helices shown in blue and red, respectively. Class I hydrophobins have much greater variety in structuring, particularly in loop sequences between β-strands (RMSD = 9.39). Class II hydrophobins have a high degree of structural conservation and their structures are overlaid to highlight similarities (RMSD = 1.04).

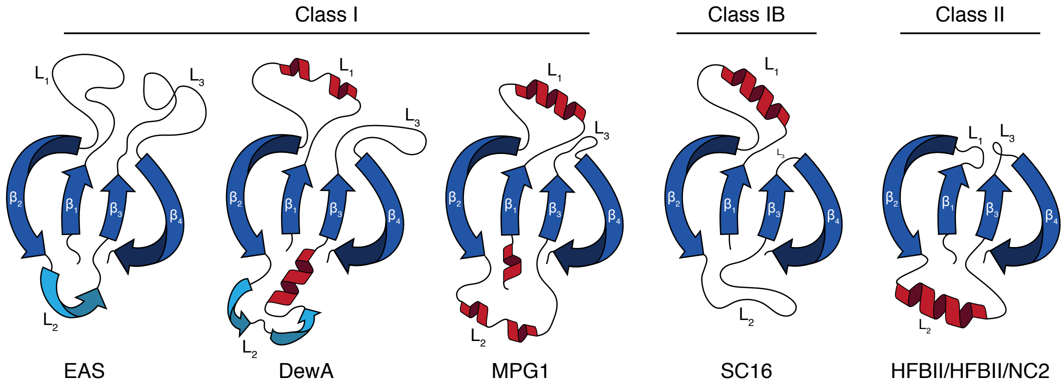


Figure 1.5. Summary of secondary structure features of hydrophobins.
A schematic of class I, class IB or class II hydrophobins that have been structurally characterized is shown. All hydrophobins share a β-sheet (blue) rich core but vary greatly in inter-cysteine loop regions which can contain α-helices (red) or additional β-sheets (teal). Class IB has elements similar to class II versus the disparate features observed in the loop regions of class I members. Figure adapted from similar comparisons by Gander et al.[42] and Wösten & Scholtmeijer.[54]

Class I and class II hydrophobins often have amphipathic character with pronounced charged or hydrophobic patches on the protein surface, including EAS, DewA, and NC2.[45,59,63] In EAS, only 8 of the 82 residues are charged, yet 6 of these are located on a single face of the protein forming distinct hydrophilic and hydrophobic faces.[58] In other hydrophobins, such as HFBI/HFBII, charged residues are more uniformly distributed throughout the surface of the proteins and there is not such a prominent separation of charge.[45,72] A large exposed hydrophobic patch is still present on these structures, such as HFBII, where approximately half of the hydrophobic aliphatic residues of HFBII are located at the surface.[70] The structures of hydrophobins are strongly stabilized by disulfide bonds, thereby compensating for the destabilizing effect of these exposed hydrophobic residues.[39,46]

Class IB hydrophobins have features reminiscent of both class I and class II hydrophobins. The first atomic-resolution structure of a class IB hydrophobin, and therefore originating from a Basidiomycota fungus, was published by Gandier et al. in 2017.[42] SC16 was chosen for characterization because it has assembly kinetics amenable to structural characterization by NMR spectroscopy and has 56% sequence identity with SC3 (measured from first to last cysteine residues), suggesting that it would have a similar structure to the extensively studied SC3. Interestingly, SC16 shares significant structural similarities with class II while retaining the rodlet forming ability of class I hydrophobins. The β-barrel structure of SC16 is more similar to the core structures of the class II hydrophobins HFBI and HFBII than to those of class I which commonly have open or irregular barrel structures (Figure 1.4). Class II hydrophobins and SC16 all consist of a β-barrel with an associated α-helix.

Overall, the loop regions of SC16 are most structurally similar to those of the class II hydrophobins. The inter-cysteine loops of class I hydrophobins EAS, DewA, and MPG1, are variable in terms of both structuring and dynamics, whereas class II hydrophobins all contain an α-helix within L_2 while L_1 and L_3 are very short . This is similar to the features observed in SC16, however the location of the α-helix differs between the two classes of hydrophobins. In SC16 an α-helix is located within L_1, while L_2 is disordered but of moderate length, and L_3 is significantly truncated.[42] This results in the α-helix being situated on opposite sides of the β-barrel in SC16 relative to class II hydrophobins (Figure 1.5). Unlike L_2, which is covalently linked to the β-sheet core in class II hydrophobins, the α-helix in the L_1 region of SC16 is stabilized by a hydrophobic patch between the inner face of the helix and adjacent β-sheet. The lack of a disulphide bond in the L_1 α-helix SC16 may allow a greater degree of conformational flexibility.[42]

1.4 Structural Basis for Rodlet and Film Assembly

A mechanism for rodlet and film assembly in hydrophobins remains unresolved, despite multiple proposed mechanisms and published hydrophobin structures. Early investigations into rodlet formation examined the role of the conserved cysteine residues. However, oligomerization driven by intermolecular disulphide bridging can be excluded as rodlet formation occurs in oxidizing conditions and all cysteines are already involved in disulphide bonding.[53] In addition, SC3 is still able to form rodlets after disulfide bond reduction and chemical blocking, indicating that the disulfide bonds are not required for self-assembly.[46] Disulphide bridging in SC3 may instead function to keep monomers in an active state by preventing self-assembly in solution, until they can adsorb to an interface and assemble into rodlets.[46] Recent observation of the formation of nanorods in solution by SC3 contradicts this finding, indicating that this effect is insufficient to counteract SC3 self-assembly.[73] Similarly, Kershaw et al. found that removal of disulphide bonds from the hydrophobin MPG1 did not affect its ability to self-assemble but hindered its extracellular secretion and localization in the fungal cell wall.[74] These results suggest that intermolecular disulphide bonds are unlikely the driver for rodlet assembly.[53]

The variably of structural features among class I hydrophobins makes it difficult to define which regions or structures are important for self-assembly. Rodlet formation is suggested to be linked with significant structural rearrangements of hydrophobin protein structuring. For the class IB hydrophobin SC3 rodlet formation is linked with rearrangements involving helical intermediates, with fully assembled rodlets containing amyloid-like characteristics and high β-sheet content.[39,75] Protease digestion and hydrogen–deuterium exchange experiments suggest that the conformational changes may be initiated within the L_1 region, which is presumed to adopt a helical structure when it initially adheres to a hydrophobic surface.[76] However, these models lack a precise description of a rodlet formation mechanism.

Kwan et al. in 2006 were the first to suggest a mechanism for hydrophobin self-assembly using the class I hydrophobin EAS as a model.[45] Based on the presented structure and X-ray fibre diffraction experiments, it was proposed that each EAS monomer aligns with a hydrophobic patch facing the hydrophilic-hydrophobic interface. The leading and trailing edges of each β-barrel core then stack end-to-end and form an elongated tube-like structure stabilized by backbone H-bonding. These changes do not result in significant structural perturbations to the core region of EAS.[45,77] Using site-directed mutagenesis to delete up to half of the largest loop (residues 29–35 or 27–37), it was shown that the disordered regions of EAS were not required for rodlet assembly since mutants could form native-like rodlets.[45]

This rodlet forming model was later refined using additional EAS deletion and point mutants[58] that indicated L_1 is not involved in rodlet formation. Simulations suggest instead that L_1 influences the kinetics of EAS localization to air-water interfaces and prevents aggregation within bulk solution.[78] At the air–water interface, the loop loses a significant amount of conformational entropy and can no longer act to prevent aggregation. In addition, mutagenesis of the L_3 region decreased the ability of EAS to form rodlets and when the L_3 sequence was incorporated into a class II hydrophobin, this rendered it amyloidogenic, suggesting that L_3 is crucial for self-assembly. In this revised model, the L_3 region undergoes a conformational change, unfolding to form an antiparallel β-sheet with the L_3 region of other EAS monomers. Many class I hydrophobins contain a sequence that is predicted to be amyloidogenic, and the position of this sequence varies: in DewA it was predicted to be present in L_2,[63] while in RodA it was predicted to be in L_3 like it is in EAS.[79]

For the class IB hydrophobin SC16, rodlet formation occurs upon aeration.[42] This is in contrast to SC3 but is similar to the assembly of MPG1 where agitation is necessary for rodlet assembly as it occurs via a surface driven mechanism. SC16 rodlet size measurements are consistent with the L_3 driven amyloidogenesis model observed in other class I members.[42] However, the L_3 of SC16 is a short 4 residue β-turn that connects β_3 and β_4, and is therefore unavailable for structural rearrangement. Only the L_1 and L_2 regions of SC16 are long enough to potentially undergo the conformational changes required for rodlet formation. Sequence analysis by the Waltz amyloid prediction algorithm[80] indicates that residues 40–47 of L_1 in SC16 are amyloidogenic, suggesting that they may be involved in structural rearrangements during rodlet formation. Structural rearrangements of this region are possible as the L_1 does not contain a disulphide bond, however, this may be unlikely as L_1 is folded into a stable α-helix, which is restrained by hydrophobic residues on the inner surface of the helix and adjacent β-sheet core. The solution structure of SC16 may be similar to transient α-helical structures which are formed by SC3 before undergoing further structural transitions to form a β-sheet rich rodlet.[40] The surface charge distribution of SC16 shows several acidic, basic, and uncharged patches, but these features are not as prominent for SC16 as for the other class I hydrophobins EAS, DewA and MPG1 further complicating structure-function correlations between hydrophobin classes.

In contrast, class II hydrophobins do not undergo conformational changes associated with self-assembly that are observed in class IA and class IB.[13] Instead, surface exposed hydrophobic patches are the driving force for their self-assembly, causing films to form at air-water interfaces.[81] However, class II hydrophobins are significantly less surface active than their class I counterparts. Unlike the class IB SC3,[46] disruption of disulfides in class II hydrophobins disrupts their structural stability, surface activity, and solution self-assembly.[82]

1.5 Properties of Hydrophobin Rodlets and Films

Prompted by the diverse ways hydrophobins are used by fungi, many studies have investigated the biophysical properties of hydrophobin assemblies. Hydrophobins are able to self-assemble on both hydrophilic and hydrophobic surfaces, reducing the surface tension of liquids and altering the wettability of surfaces, making hydrophilic surfaces hydrophobic and vice versa.[71,83] Class I hydrophobins assemble into amyloid-like rodlets, which are hundreds of nanometers long, packed into ordered lateral assemblies and do not exhibit an overall helical structure.[21,40,45,51] Rodlet assemblies are extremely durable, being resistant to a variety of organic solvents, detergents and high temperatures.[46,51] Class II hydrophobins on the other hand, lack the distinct morphology found in the assemblies of class I.[84,85] These films are less durable than class I assemblies, being soluble in some detergents and solvents but overall still show remarkable durability for proteins, being resistant to temperatures as high as 90 °C without any sign of denaturing.[55]

Driven by the presence of both hydrophilic and hydrophobic residues on the protein surface, hydrophobins can modify the properties of variety of surfaces and interfaces. Assembly of hydrophobins at air-water interfaces results in significant reduction of surface tension, For example, SC3 reduced the surface tension of water from 72 to 43 $mJ \cdot m^{-2}$ and HFBII reduced the surface tension of water to 28 $mJ \cdot m^{-2}$ at a concentration of 0.01 $mg \cdot ml^{-1}$ and 0.02 $mg \cdot ml^{-1}$ respectively.[70,86] Hydrophobins can also bind surfaces, which allow fungi to attach to insect cuticles or plant cellulose.[26,30] A variety of other inorganic surfaces can also be bound in vitro, such as treated or clean glass surfaces, polystyrene, or even Teflon.[27,40,51,87] SC3 assembly on Teflon has been shown to produce a surface even more hydrophobic than Teflon itself.[86]

Between hydrophobins of the same class, surface activity can vary significantly. Many hydrophobins require a hydrophobic-hydrophilic interface to induce morphological changes and self-assembly. MPG1, DewA and EAS all assemble at the interface in a process that is catalyzed by an interface and is impacted by pH, temperature, and protein concentration.[67] Highly surface active hydrophobins may not require these interfaces for assembly, for example SC3 can spontaneously assemble into nanorods in solution.[73] Similarly, the conversion of Vmh2 (from *Pleurotus ostreatus)* into the β-sheet rich forms occurs spontaneously without the need for an air water interface, and is promoted by high protein concentration, low pH, high temperature, and Ca^{2+} ions.[88]

This high surface activity and the propensity of hydrophobins to readily self-assemble has been a significant barrier to the atomic-resolution structural characterization of hydrophobins. Sample preparation remains difficult for nuclear magnetic resonance (NMR) spectroscopy as well as X-ray crystallography due to sample inhomogeneity, self-assembly, and protein precipitation. For example, despite SC3 being one of the most heavily studied hydrophobins, its atomic structure remains unknown due its ability to readily self-assemble at 100 μg/ml without significant agitation.[73]

1.6 Potential Applications for Hydrophobins

Due to the non-specific action of hydrophobins and their multifaceted functional roles as emulsifiers, surface modifiers and as protective coatings, the list of potential applications is long and diverse.[1,89–94] To date, the hydrophobins HFBI/HFBII from *Trichoderma reesei* and SC3 of S*chizophyllum commune* have been common targets for development, where they find use in both industry and medicine.[89,95] Hydrophobins are excellent emulsifiers, allowing them to be used to solubilize and increase the bioavailability of otherwise poorly soluble drugs (Figure 1.6).[96–98] This emulsification ability may also increase oil recovery from oil wells, increasing overall extraction efficiency.[99] Currently the largest use of hydrophobins in commercial settings is in the food industry where, due to their incredible foam stabilization abilities, they have been incorporated into products ranging from ice cream to margarine.[100] Bubble stability of HFBII was studied by Cox et al., who found that foams and bubbles of HFBII were stable for months to several years, where the amount of protein used was as low as only 0.1 wt% protein.[101] This makes hydrophobins an attractive target for inclusion in ice cream and whipped cream, where they can stabilize dispersed air bubbles.[92,102] However, large-scale applications of hydrophobins are currently limited by the production cost of recombinant proteins and difficulty in scaling-up protein production. Currently, only two modified hydrophobins are commercially available, H*ProteinA and H*ProteinB, which are produced at pilot-scale (i.e., on the order of the kg) by Baden Aniline and Soda Factory and marketed as a foam stabilizer.[103]

Hydrophobins find use in have medical applications, as they are not inherently toxic and prevent immune responses via formation of protective surface layers, making them a highly attractive material for coating medical devices.[90,104] For instance, variants of DewA have been applied in implant[105] and stent coatings.[106] Furthermore, surface modification properties can also be desirable; SC3 coatings of catheters had a 70–80% reduction in friction coefficient compared to non-coated catheters.[107]

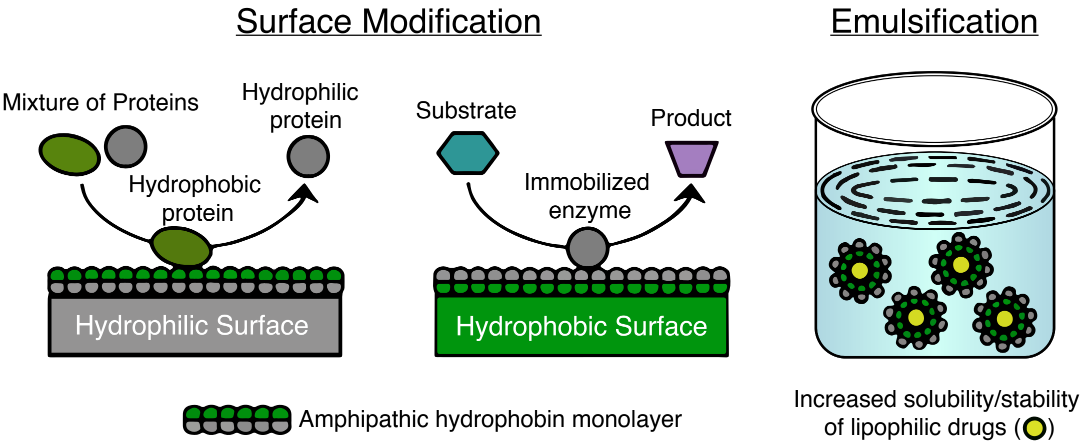


Figure 1.6. Hydrophobins have potential uses as surface modifiers and emulsifiers. Hydrophobins readily self-assemble at interfaces forming durable assemblies as rodlets or films. These assemblies bind to both hydrophobic and hydrophilic surfaces and reverse their polarity. Binding and purification of target hydrophobic peptides or immobilization of enzymes have both been proposed.[38] Emulsification of poorly soluble drugs and drug delivery by hydrophobins have also been explored.[108,109] Figure adapted from Bayry et al. 2012[13]

A more comprehensive understanding of how self-assembly occurs would facilitate quicker and more efficient implementation of hydrophobins in industrial and medical applications. Firstly, a residue-specific understanding of the rodlet formation mechanism would provide a roadmap to engineer hydrophobins with desirable properties beyond current efforts.[105,110,111] Secondly, a comprehensive understanding of hydrophobin structure would allow for modifications that would facilitate higher production yields, by potentially increasing solubility and reducing undesirable or premature self-assembly.

1.7 Rationale and Research Objectives

As detailed above, hydrophobins have unique properties that make them attractive targets for many commercial, industrial, and medical applications. They have a remarkable ability to self-assemble into durable rodlet structures and are an excellent example of functional amyloids. However, the mechanism of rodlet formation is unclear. Proposed models for hydrophobin self-assembly contain large discrepancies and no model is fully applicable to all members of either class I or class II. Understanding hydrophobin function is further complicated by the poor conservation and diversity of the origins, sequences, and structures of hydrophobins. However, several different attributes have been noted to be shared among hydrophobins. First, the conserved disulphide bonding pattern present in all hydrophobins stabilizes the core region and monomeric units prior to assembly. Second, an amyloidogenic loop is present in most Class I hydrophobins studied, although this region remains unresolved for class IB. Lastly, hydrophilic patches on the protein surface gives hydrophobins an affinity for hydrophobic-hydrophilic interfaces, with association potential resulting in a shift in oligomerization kinetics and the formation of hydrophobin assemblies. Determining how these structural features contribute to rodlet formation will provide fundamental insight into mechanisms of protein assembly and will aid in selecting hydrophobins with favourable features for commercial applications.

The recent description of class IB hydrophobins that have a more consistent sequence provides an opportunity to correlate hydrophobin sequence, structure, and function.[42] To this end, I compared the sequences, structure, and properties of three distinct class IB hydrophobins in an attempt to better understand what modulates hydrophobin self-assembly. My first objective was to express and purify each target hydrophobin (Chapter 2). My second objective was to determine the atomic resolution structure of each hydrophobin using NMR spectroscopy (Chapter 3). Finally, I confirmed rodlet forming ability of each target hydrophobin and determined the effects of solution conditions on self-assembly ability (Chapter 4).

Chapter 2: Heterologous Expression and Purification of Class IB Hydrophobins

2.1 Introduction

In Chapter 1, I provided an overview of the role of fungal hydrophobins in nature and a brief description of their many potential applications. I also outlined our current understanding of hydrophobin self-assembly as it relates to the sequence and structure of different classes of hydrophobins. Currently, the most rigorously described self-assembly mechanism supported by structural observations remains only narrowly applicable to a few hydrophobins of class I. To better understand structural contributions to this process, I aimed to characterize the structural and functional properties of several class IB hydrophobins with unique sequence properties and determine if they had shared structural features. Prior to structural characterization by NMR spectroscopy target hydrophobins were required to be expressed and purified, using isotopic labeling to facilitate the experiments detailed in the following chapters. In this chapter, I focus on preparation of class IB hydrophobins prior to further characterization. I first summarize previous methods used to express hydrophobins, then describe the expression and purification of target class IB hydrophobins.

2.1.1 Expression Systems Used in Recombinant Hydrophobin Production

No production method for hydrophobins has been universally established and a variety of host systems are used to produce hydrophobins. In some cases, hydrophobins are isolated form wild fungi, including the original work describing the first hydrophobin characterized, SC3.[51,112,113] Overproducing fungal strains are also commonly used, such as the system developed by VTT Technological Research Centre of Finland.[112] These overproducing strains contain multiple amplified copies hydrophobin genes of interest and deletions of other hydrophobin genes to achieve the highest possible yield of a single hydrophobin. For example, a modified strain of *Trichoderma reesei* containing three extra copies of the hfb2 gene results in yields of HFBII of up to 240 mg/L in lactose-enriched media. Yeast expression systems have also been used to achieve high yield hydrophobin production.[66,114–119] Similar to overproducing fungal strains, high yields of HFBII (260 mg/L) using yeast have been reported using a strain of *Saccharomyces cerevisiae* containing multiple copies of hfb2 gene.[120] *Pichia pastoris* can also produce high yields of hydrophobins, with the added benefit of secretion of the produced protein into the surrounding medium, simplifying subsequent purification.[66,115,116,119,121]

Recombinant bacterial expression can be used to produce hydrophobins, including production on an industrial-scale using a modified DewA construct.[45,103,117,122] Heterologous production of hydrophobins in *Escherichia coli* was first attempted with the characterization of EAS as described above, where growth in bacteria necessitated the oxidative refolding of EAS from inclusion bodies and confirmation of proper protein folding by reverse phase high performance liquid chromatography (RP-HPLC) analysis.[45]

The yield of class I hydrophobins has been reported to be 10 to 100-fold lower when expressed in bacteria compared to fungal or yeast sources.[63,122–124] However, using bacterial systems does have distinct advantages versus other expression systems. Simple vector construction, media preparation, high growth rates, and the use of a T7 promoter to control of protein expression make *Escherichia coli* the simplest system for producing recombinant proteins. Producing ^{15}N and ^{13}C isotopically labelled proteins required for triple-resonance NMR spectroscopy is simple with *E. coli*, as labeling can be achieved by using ^{15}N labelled NH_4Cl and ^{13}C-labelled glucose as the only sources of nitrogen and carbon in a minimal growth medium. Therefore, an E.coli expression host was chosen for this study due to cost-effectiveness, and simplicity in cultivation, and protein purification.

Several different strains of *E.coil* have been previously employed in the production of hydrophobins, including BL-21 (DE3),[123] Origami B (DE3),[42,124] and Rosetta (DE3).[105] The SHuffle T7 Express *E.coli* strain has been chosen for protein expression in this work, due to the established critical role of disulfide bonds in maintaining hydrophobin stability.[46] Due to the presence of numerous thiol reductases, glutathione, and small thiol reductants, cysteines are in their reduced state in the cytoplasm of wild-type *E. coli* and are not able to form stable disulfide bonds.[125] The SHuffle strain of *E. coli* is engineered to promote disulfide bond formation in the cytoplasm[126] by disrupting thioredoxin reductase and glutathione reductase genes. In addition, the insertion of a non-specific disulfide bond isomerase, DsbC, facilitates proper folding of proteins by disulphide shuffling.[126] Combined, these two modifications allow for increased folding and solubility in proteins that contain disulfide bonds.[127]

2.1.2 Selection of Target Class IB Constructs

Class IB hydrophobins have a higher degree of sequence conservation compared to most hydrophobins, but which features of class IB hydrophobins determine their properties and propensity for self-assembly is unknown. Based on previous work characterizing SC16,[42] we selected three class IB hydrophobins with distinct properties for study (Figure 2.1), SL1 from *Serpula lacrymans* (containing only one charged residue), WI1 from *Wallemia ichthyophaga* (containing many charged residues) and PC1 from *Phanerochaete carnosa* (containing no basic residues; Figure 2.1). The fungal sources of these hydrophobins are phylogenetically diverse, SL1 and PC1 originate from common dry rot fungus and a crust fungus, respectively. *Wallemia ichthyophaga* is one of the most halophilic fungi known, being native to the dead sea and growing in salt concentrations as high as 5 M.[128,129] Its high number of charged residues has been hypothesized to allow WI1 to function in such an extreme environment.[129]

As shown in Figure 2.1, SL1, PC1 and WI1 have a high degree of sequence similarity compared to SC16 and the well-studied hydrophobin, SC3. This contrasts with the poor sequence conservation observed between class I members (Figure 1.4 A) and is reminiscent of the high conservation observed between class II hydrophobins (Figure 1.4 B). Of particular note is the absence of large insertions or deletions among these constructs within the inter-cysteine sequences indicating that loop regions are of similar lengths, a feature observed in Class IB hydrophobins, but not Class I as a whole. Interestingly, the Waltz amyloid prediction algorithm does not predict the same regions to be amyloidogenic between class IB hydrophobins.[80] Residues 40–47 and 24-29 of L_1 in SC16 and SL1 respectively and residues 41-46 of β_2 in WI1 are suggested to be amyloidogenic. However, the same analysis predicts that no amyloidogenic sequences are present in PC1.

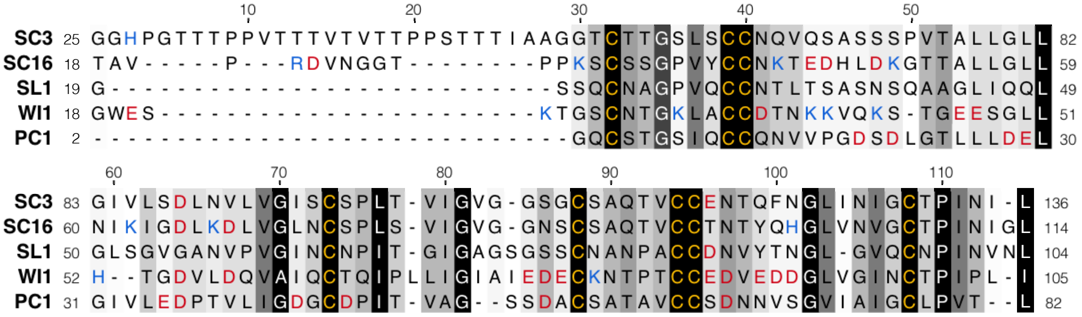


Figure 2.1. Target class IB hydrophobins have high sequence conservation despite large differences charged residue content.
Sequence alignment the class IB hydrophobins SC3, SC16, SL1, WI1, and PC1. Shading indicates the relative degree of conservation with black indicating the most highly conserved residues. Cysteine residues are highlighted in yellow and align completely between each hydrophobin. Acidic and basic residues are shown in red and blue respectively. Alignment and conservation values were generated using T-Coffee.[61]

2.2 Materials and Methods

2.2.1 Preparation of Class IB Expression Constructs

Synthetic genes coding for 53G-136L of the SC3 gene of *S. commune* (NCBI ID: 5334), 18T-116L of the Hyd1 (SC16) gene of *S. commune* (NCBI ID: 578458), 19G-104I of the slh4 (SL1) gene from *S. lacrymans* (NCBI ID: 578457), 8G-105I of the J056_002864 (WI1) gene from *W. ichthyophaga* (NCBI ID: 536056), or 21T-28I of the PHACADRAFT_78259 (PC1) gene from *P. carnosa* (NCBI ID: 650164), were purchased (BioBasic Inc.), These sequences correspond to full length hydrophobins sequences with the N-terminal signal peptide removed. A variety of plasmids coding for target hydrophobins were prepared to determine the optimal expression tag and cleavage system combination. Target genes were cloned into a modified pET21-b vector downstream of an expression tag using Bam HI and Xho I restriction enzymes. Multiple expression tags were used, including hexahistidine (His_6) and hexahistidine fused to thioredoxin (His_6-Trx) or Small Ubiquitin-like Modifier (His_6-SUMO). Two versions of His_6 tags were used, one with a thrombin protease recognition sequence (His_6-Th), or a tobacco etch virus recognition sequence (His_6-Trx). The His_6-Trx and His_6-SUMO constructs coded for enterokinase and SUMO protease recognition sequences, respectively. PC1 also has an appended N-terminal tryptophan to facilitate monitoring by UV-Vis spectroscopy. The final His_6-TEV fusion protein constructs used are as shown in Figure 2.2.

2.2.2 Bacterial Expression of Hydrophobins

Prepared constructs were confirmed by plasmid sequencing prior to transformation into both BL-21 (DE3) and SHuffle T7 Express *E.coli*. Electrocompetent *E. coli* were transformed via electroporation and cultured on an ampicillin plate overnight for 18 hours. Single bacterial colonies were grown in 2 L of LB media (10 g/L tryptone, 10 g/L NaCl, 10 g/L yeast extract, 10 mL glycerol) supplemented with 100 mg/L ampicillin at 37 °C until the culture reached an optical density at 600 nm (OD_{600}) of 0.6–0.8. Expression was then induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were cultured overnight (18-22 h) at 22 °C then harvested by centrifugation at 4000 × *g*

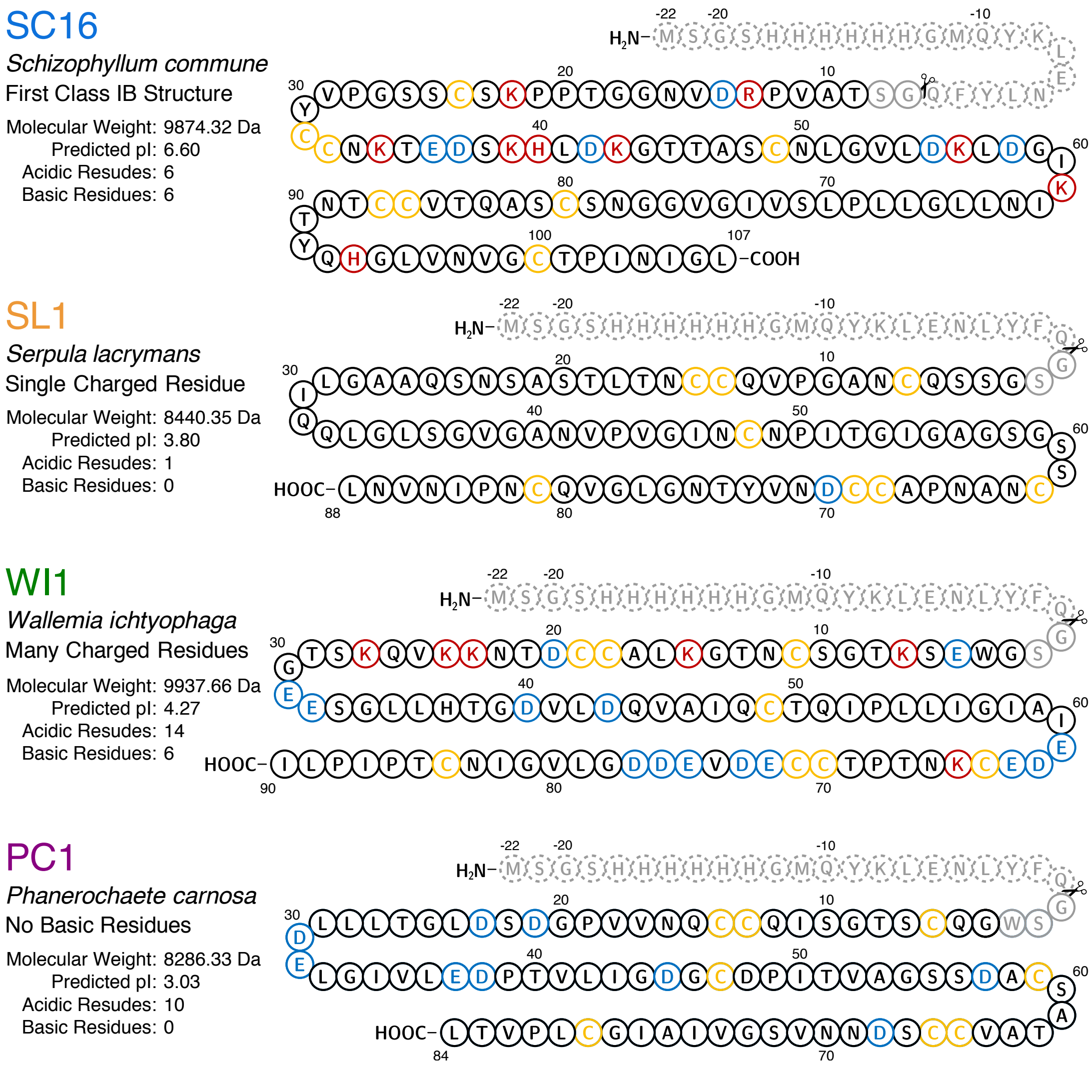


Figure 2.2. Hydrophobin proteins expressed in *E. coli.*
Schematics of SC16, SL1, WI1, and PC1 proteins. are C-terminal to a hexahistidine repeat which enables purification by immobilized metal ion affinity chromatography. The cleavage site of TEV protease is denoted by the pair of scissors. Physical properties of each construct are shown on the left,[130] with number of acidic (blue) and basic (red) residues shown. Cysteines are highlighted in yellow. Residues not present in native forms are shown in grey, including His_6-TEV fusion tag which is indicated by the dotted lines. Numbering is relative to final cleaved protein product, with the cleavage site at position zero.

For isotope labelled expression, 500 µL of a LB starter culture was used to inoculate 50 mL of M9 minimal medium[131,132] (2 g/L ^{13}C-glucose, 1 g/L $^{15}NH_4Cl$ (Cambridge Isotope Laboratories), 100 µM CaC_2, 1 mM $MgSO_4$ 10 µg/mL thiamine, 10 µg/mL biotin, 100 mg/L ampicillin, 60 mg/L $FeSO_4·7H_2O$, 60 mg/L $CaCl_2·2H_2O$, 120 mg/L $MnCl_2·4H_2O$, 8 mg/L $CoCl_2·6H_2O$, 7 mg/L $ZnSO_4·7H_2O$, 3 mg/L $CuCl_2·2H_2O$, 0.2 mg/L H_3BO_3, 2.5 mg/L $(NH_4)_6Mo_7O_{24}·4H_2O$, 50 mg/L EDTA) and grown overnight at 37 °C. This culture was pelleted and resuspended in 1 L of M9 media, incubated at 37 °C, and induced with 0.5 mM IPTG at an OD_{600} of 0.6-0.8. After induction cells were cultured overnight (18-22 h) at 22 °C then harvested by centrifugation at 4000 × *g* .

2.2.3 Purification by Ni^{2+} Affinity Column Chromatography

Cell pellets were resuspended in 50-100 mL denaturing Ni^{2+} binding buffer (20 mM trisaminomethane (Tris) pH 8.0, 250 mM NaCl, 8 M urea) and lysed until translucent (5 min) using sonication while on ice. The lysate was centrifuged at 25,000 × *g* for 20 min at 4 °C, then the supernatant was loaded onto a Ni^{2+}-charged Immobilized Metal Affinity Chromatography (IMAC) column (IMAC Sepharose 6 Fast Flow, GE Healthcare). The column washed with denaturing Ni^{2+} binding buffer containing 20 mM imidazole and protein was eluted with 7-10 mL of denaturing Ni^{2+} binding buffer containing 300 mM imidazole.

Ni^{2+} column elution fractions were dialysed against 20 mM Tris pH 8.0, 50 mM salt, 1 mM oxidized glutathione (GSSG) , and 2 mM reduced glutathione (GSH). If significant protein precipitation occurred, 1 M urea was added to this dialysis buffer. After three buffer refreshes, 0.4 mg of TEV protease was added to the dialysis bag and left overnight. The cleavage reaction was passed through a Ni^{2+} column (termed rNi^{2+}

purification). The flowthrough and wash fractions were collected and dialysed against distilled H_2O and further purified by FPLC or concentrated via ultrafiltration or lyophilization prior to purification by HPLC. A summary of this expression and purification procedure is shown in Figure 2.3. Expression and purification steps were monitored by UV/Vis absorbance spectroscopy and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in 15% acrylamide gels run at 170 volts for 50 min at room temperature. Protein was visualized with either Coomassie Brilliant Blue R-250 or silver staining methods.[133]

2.2.4 FPLC Purification of Hydrophobins

SL1, WI1, and PC1 were further purified by fast protein liquid chromatography (FPLC); ÄKTA FPLC, Amersham Biosciences) using a 1.7 mL QSepharose Fast Flow (GE Healthcare) anion exchange column. Hydrophobins were loaded onto the column that was equilibrated with 20 mM Tris pH 8.0, and eluted with a linear gradient of NaCl up to 400 mM over 20 column volumes at 4 mL/min. Fractions containing the desired protein were pooled and used immediately, or dialysed against distilled water, lyophilized, and stored at -20 °C.

2.2.5 HPLC Purification of Hydrophobins

After lyophilization, SL1, WI1 and PC1 were reconstituted into 5-10 mL of distilled water and centrifuged at 7000 × *g* for 10 minutes to remove precipitate prior to purification using reverse phase HPLC with a Prostar chromatography system equipped with a diode array detector (Varian Canada Inc.). Analytical HPLC was performed at a flow rate of 1 mL/min using a C_{18} matrix column (5 µm particle, 4.6 mm x 250 mm Hypersil-gold, Thermo Scientific), while semi-preparative HPLC used a C_8 matrix column (5 µm particle, 9.4 mm x 250 mm Zorbax, Aligent Technologies) at a flow rate of 3.5 mL/min. All HPLC was carried out using milliQ water and HPLC grade acetonitrile containing 0.05% trifluoroacetic acid. Analytical HPLC was carried out with a gradient of 10% - 60% acetonitrile over 25 minutes, while a gradient of 20%-50% acetonitrile over 15 minutes was used for semi-preparative HPLC. For PC1, a gradient of 30% to 60% acetonitrile over 15 minutes was used. Molecular masses of collected fractions were confirmed using positive mode electrospray ionization mass spectrometry (ESI+ MS) (Dalhousie Mass Spectrometry Laboratory, Halifax, NS). HPLC fractions containing desired protein were collected, lyophilized, and stored at -20 °C.

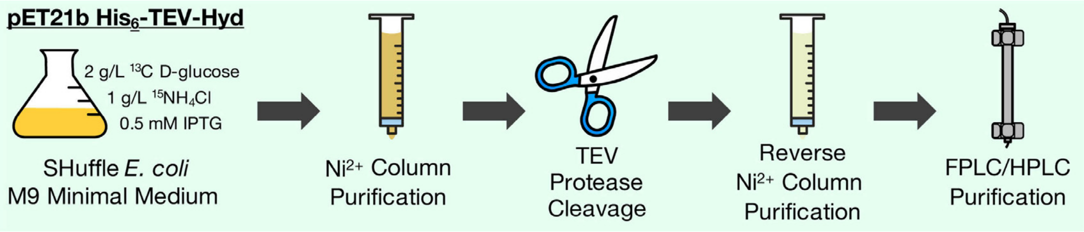


Figure 2.3. General overview of expression and purification of isotope-labelled class IB His$_6$-TEV-Hydrophobin constructs.
Expression and purification of either SC16, SL1, WI1 or PC1 begins with transformation of SHuffle T7 express *E. coli* cells. After selection with ampicillin plates, bacteria is cultured in either rich LB media or M9 minimal medium with added ^{15}N labelled NH_4Cl and ^{13}C-glucose to facilitate NMR spectroscopy experiments. The inclusion of a hexahistindine tag enables efficient purification of desired protein before and after cleavage. Final purification is performed using either FPLC, or more often HPLC prior to characterization experiments.

2.3 Results and Discussion

In this chapter I have outlined previous methods for producing fungal hydrophobins including fungal, yeast and bacterial sources. In this work, bacterial expression proved to be amenable to the production of class IB hydrophobins after optimization of several key steps. Due to the large number of potential hydrophobin constructs, optimization of hydrophobin expression and purification was usually carried out for one hydrophobin and then applied to the others. I initially attempted to express and purify SC3, with the goal of determining its atomic resolution structure. As such, my initial efforts to optimize expression conditions focus on SC3. However, issues with SC3 production became apparent, and I shifted my focus to production of WI1 and PC1, which were used for the bulk of method development. Once an optimized method was determined, this procedure (outlined in methods) was used to successfully express and purify SC16, SL1, WI1, and PC1.

2.3.1 Initial Attempts at Purification of SC3

To date, the atomic resolution structure SC3 remains unresolved and my initial goal was to prepare samples of SC3 suitable for characterization by NMR. Test expressions using BL21 (DE3) and SHuffle T7 express strains of *E.coli (*Henceforth called SHuffle *E. Coli*) were performed using SC3 alongside WI1 and PC1 (section 2.3.2 and 2.3.4) which served as a foundation of subsequent expressions. Two constructs of SC3:53-136 were used in the effort, one with a GB1 accessory fusion protein and one without, His_6-GB1-TEV-SC3 and His_6-TEV-SC3 respectively. These constructs were expressed and purified by Ni^{2+} affinity chromatography.

Figure 2.4 shows clear separation of SC3 from bulk contaminants present in lysate but significant losses were observed during dialysis and TEV protease cleavage of the expression tag, as visualized by SDS-PAGE gel analysis. GB1-SC3 showed less precipitation versus free SC3, likely due to the solubilizing effects of GB1, but had significantly less yield during expression. Interestingly, no change in apparent molecular weight is observed upon TEV protease cleavage, which would decrease the molecular weight from 13.0 kDa (His_6-TEV-SC3) to 11.2 kDa (SC3).

The lack of band shift upon TEV cleavage suggests that TEV protease cleavage did not occur. However, after cleavage, the SC3 did not bind the Ni^{2+} column, suggesting that cleavage of the His_6 tag was successful. Despite these initial encouraging results, SC3 has proven unsuitable for further characterization. Further purification attempts by either FPLC or HPLC were inconclusive. Using fractions containing still soluble SC3, purification by size exclusion chromatography showed that all available SC3 eluted with the void volume of the column, indicating a mass higher than 600 kDa. This suggests that, even in solution, SC3 is in an assembled form, which is consistent with findings from other groups.[42,73] Precluding the use of strong acids or organic solvents to prevent self-assembly, SC3 remains incompatible for characterization by solution state NMR spectroscopy.

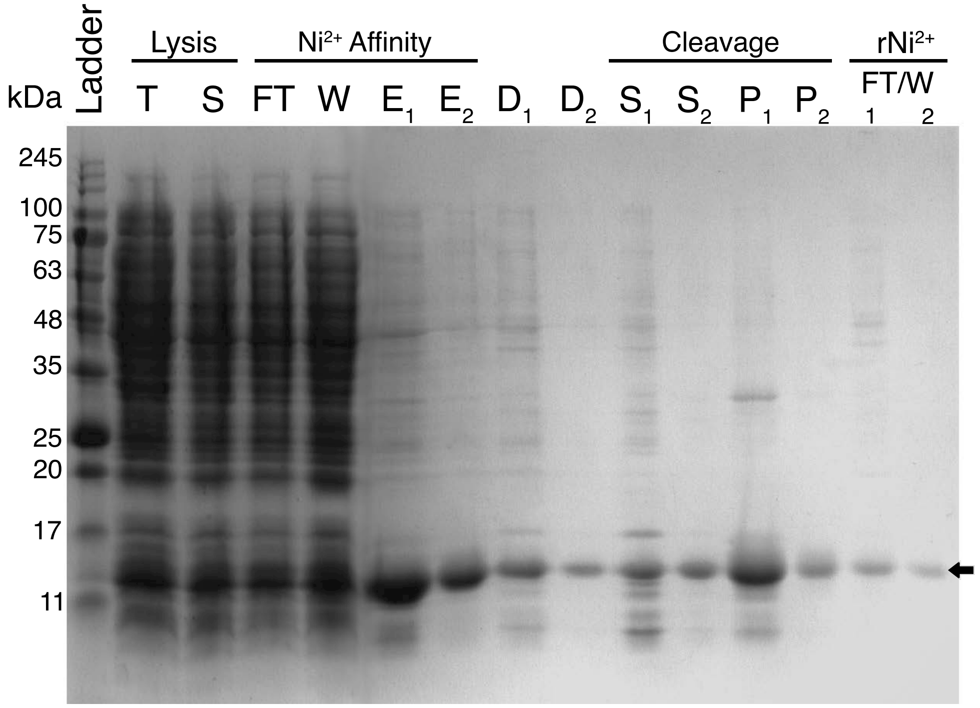


Figure 2.4. Expression and purification of SC3.
H_6-TEV-SC3 was expressed in BL21 (DE3) *E.coli* and purified. Fractions were resolved with SDS-PAGE and visualized with Coomassie blue. Arrow denotes bands of interest. Lysis total (T) and supernatant (S) and Ni^{2+} affinity column flow through (FT), wash (W) and elutions (E) dialysis (D), TEV protease cleavage (S) and precipitate (P) Combined flow through and wash (FT/W) from second Ni^{2+}-IMAC purification (rNi^{2+}) shows isolated SC3, but with a significant decrease in yield.

2.3.2 Optimizing *E. coli* Strain Choice and Expression Constructs

Initial expressions of SC3 were carried out with BL21 (DE3) *E. coli*, resulting in modest yields. Due to the conserved nature of cysteine residues and their reported effect in stabilizing and solubilizing hydrophobins,[46,82] it is likely that proper disulphide formation is important for high levels of recombinant hydrophobin expression. To test this, hydrophobin expression in SHuffle *E. coli* was compared against BL21 (DE3).

SHuffle *E. coli* expressed higher levels of WI1 and PC1 than BL21 (DE3) (Figure 2.5). This trend is especially clear for PC1, where expression is observed in SHuffle but not in BL21 (DE3). Similar results have also been observed for SL1 and SC3 constructs.

Fusion with proteins such as SUMO is commonly employed to enhance protein expression through their ability to increase stability and solubility of their attached fusion partner protein.[134–136] In order to determine if the creation of fusion proteins could increase hydrophobin expression, PC1 was chosen for optimization as it had the lowest observed yield among all targeted hydrophobins (Figure 2.5). His$_6$-TEV-PC1, His$_6$-Th PC1 and His$_6$-SUMO PC1 were cloned and transformed into *E. coli* and protein expression was induced overnight at 22°C. Since PC1 has low expression, to better quantify the yield of each construct, 1 L of cells were induced and the expressed protein was purified via Ni^{2+} column affinity (Figure 2.6). His$_6$-TEV-PC1 resulting in the highest yield before and after Ni^{2+} column purification. Based on these results, SHuffle *E.coli* and His$_6$-TEV constructs were used for subsequent hydrophobin expression

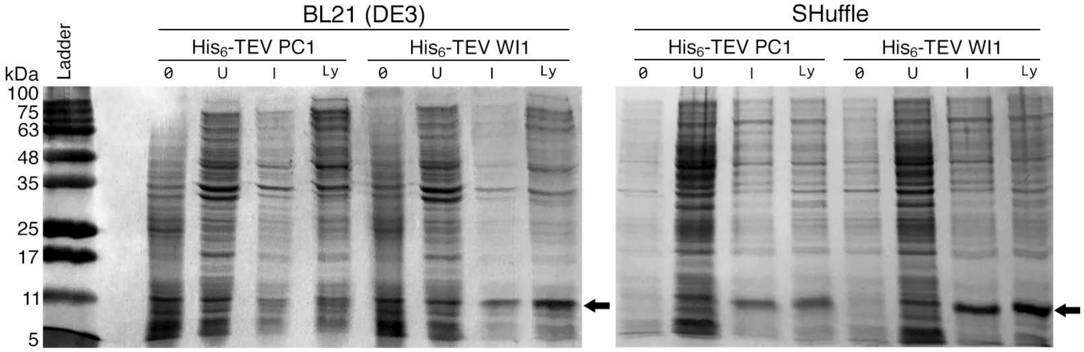


Figure 2.5. SHuffle *E. coli* allows for higher expression of target hydrophobins than BL21 (DE3) *E. coli*.
SDS-PAGE analysis of His$_6$-TEV-PC1 and His$_6$-TEV-WI1 expression in BL21 (DE3) and SHuffle strains of *E. coli.* Samples were collected at time of inducing expression (0), uninduced control (U), four hours at 37°C after induction (I), and lysis (Ly).

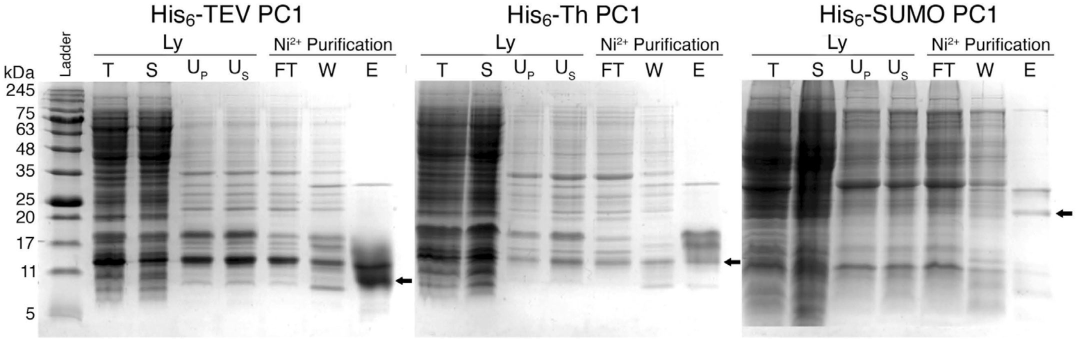


Figure 2.6. The His$_6$-TEV fusion tag results in the highest expression of PC1. His$_6$-TEV-PC1, His$_6$-Th-PC1 and His$_6$-SUMO PC1 were transformed into and expressed in SHuffle *E.coli* for 18 hours at 22 ºC. Cells were lysed (T) in Ni$_{2+}$-binding buffer (without urea), centrifuged, and the supernatant (S) was collected. The pellet was resuspended in Ni^{2+} binding buffer containing urea (U$_P$/U$_S$) and centrifuged again prior to Ni^{2+}-affinity purification, flowthrough (FT, wash (W) and elution (E) shown. Arrows denote expressed His$_6$-TEV-PC1 (8.3 kDa), His$_6$-Th PC1 (10.7 kDa), or His$_6$-SUMO PC1 (20.3 kDa).

2.3.3 Optimizing SHuffle *E. coli* Expression Conditions

To determine optimal hydrophobin expression conditions in rich LB medium, two factors were examined: expression temperature and carbon source. More H$_6$-TEV-WI1 was produced when protein expression occurred at 22 ºC for 18 hours compared to 37 ºC for 4 hours (Figure 2.7). Additionally, the presence of another carbon source in the form of glycerol resulted in higher WI1 expression, but only in cultures grown at 22 ºC for 18 hours. Cultures grown 37 ºC for 4 hours had higher cell density and thicker lysates, but this did not translate into higher WI1 expression. Addition of glycerol resulted in increased expression, but only at 22 ºC (Figure 2.7). In comparison, at 22 ºC all lanes had a much higher background, which reflects the higher cell density of the culture, but lower relative WI1 expression.

Cell culture growth conditions in the M9 minimal medium required for labelled protein expression was examined after several hydrophobin expressions resulted in little or no yield. Cultures in M9 medium were grown and protein expression induced at 22 ºC for 18 hours. Addition of extra glucose (4 g vs 2 g) did not increase hydrophobin expression but resulted in higher total cell density and increased background when resolved by SDS-PAGE (Figure 2.8). The method of inoculating cultures of M9 media used for full scale expression also affected hydrophobin expression. Cells grown to an OD$_{600}$ of 0.6-0.8 in 2 L of LB then resuspended in 1 L of M9 before induction resulted in poor expression relative to allowing the *E. coli* to grow in M9. To reduce the shock of resuspension, 50 mL M9 cultures were first inoculated to grow overnight, then resuspended in 1L M9 which was then grown to OD$_{600}$ of 0.6-0.8 prior to induction of expression.

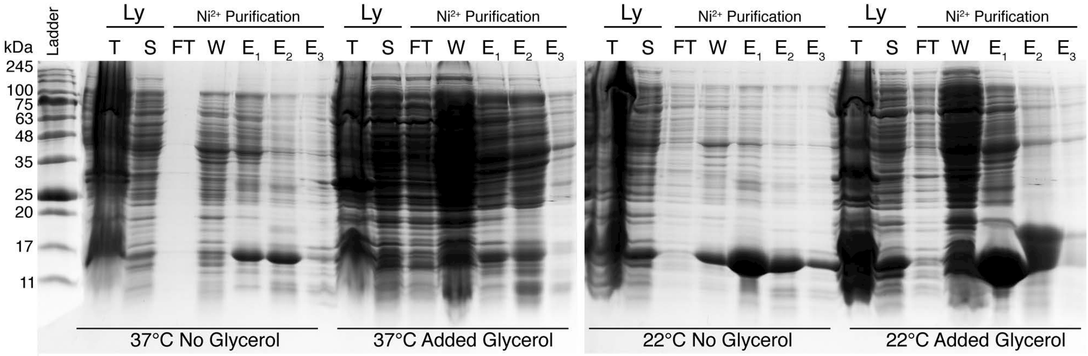
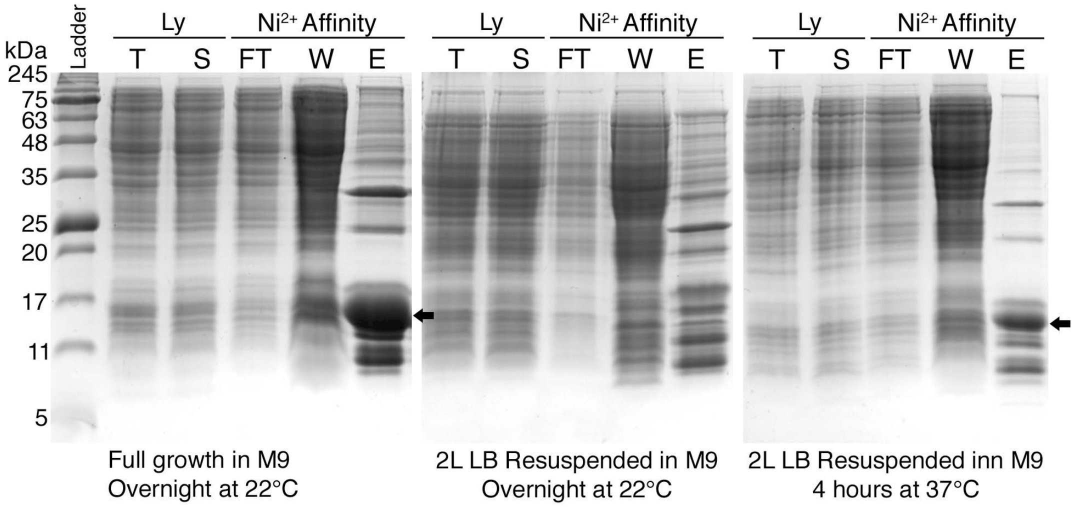


Figure 2.7. Optimization of media composition and temperature for His$_6$-TEV-WI1 expression.

1 L cultures of His$_6$-TEV-WI1 (9.4 kDa) were expressed at 22 ºC for 4 hours or 22 ºC for 18 hours overnight, and with or without added glycerol. Samples were lysed (Ly), clarified (total – T, supernatant – S), and purified by Ni affinity chromatography (flowthrough – FT , wash – W and elution – E) fractions were visualized by SDS-PAGE and Coomassie blue staining.



Full growth in M9
Overnight at 22°C

2L LB Resuspended in M9
Overnight at 22°C

2L LB Resuspended inn M9
4 hours at 37°C

Figure 2.8. Inoculation and growth in M9 medium results in highest His_6-TEV-WI1 expression.

Three 1 L LB cultures of His_6-TEV-WI1 (9.4kDa) were grown and expressed in different conditions. (Left) a 50 mL culture of M9 was used to inoculate 1L of M9 which was grown to an OD_{600} of 0.6-0.8 prior to induction and expression overnight at 22 ºC, resulting highest observed expression of the three conditions. (Middle) 2 L of LB was grown to OD_{600} of 0.6-0.8 then resuspended in 1 L of M9 and expressed overnight at 22 ºC. (Right) 2 L of LB was resuspended in 1 L of M9 and expressed overnight at 22 ºC. Cells pellets were lysed in Ni2+ binding buffer (Total lysate – T, supernatant – S) and purified by Ni^{2+} affinity chromatography (flowthrough – FT, wash – W, and elution – E).

2.3.4 Affinity Tag Removal and Refolding of Class IB Hydrophobins

In section 2.3.2 it was shown that the use of an His_6-TEV affinity tag resulted in the highest expression of WI1 and PC1. This affinity tag includes the sequence ENLYFQG/S, which TEV protease cleaves N-terminal to the QG with high specificity.[137] An advantage of TEV protease is that it is unaffected by many additives, including denaturing and reducing agents.[138] The majority of expression resulted in the expression of soluble hydrophobins. In protein preparations with significant precipitation, up to 4M urea has been added to cleavage buffer to keep cleaved hydrophobins soluble with no impact on refolding or cleavage observed by HPLC or SDS-PAGE.

HPLC was used to monitor general purity and the effect of TEV protease cleavage of the His_6-TEV affinity tag (Figure 2.9). Before TEV protease cleavage, a large collection of broad peaks is observed prior to cleavage, corresponding to multiple misfolded states instead of a single properly folded one as confirmed by MS analysis. Glutathione was added to each sample during cleavage to facilitate disulfide shuffling and protein refolding while maintaining TEV protease in the reduced state necessary for efficient cleavage. Upon addition of TEV-protease and a mixture of reduced : oxidized glutathione, a single sharp peak is observed, with many broad and smaller peaks trailing afterwards. This suggests that the reversible equilibrium of reduced and oxidized glutathione allows WI1 to fold into a single conformation. A similar pattern was observed with PC1, but SC16 and SL1 both show a single peak before and after cleavage by HPLC.

During initial purifications of WI1, 10:1 mM reduced to oxidized glutathione (GSH:GSSG) was used, but resulted in excessive protein precipitation, suggesting that the ratio of oxidized to reduced glutathione has an effect on proper folding. To investigate this, I refolded and cleaved His_6-TEV-WI1 in a variety of glutathione concentrations. Figure 2.10 shows that 1:1 mM GSH:GSSG results in optimal refolding with the largest folded peak (18 min) and smallest trailing broad peak corresponding to misfolded and uncleaved protein.

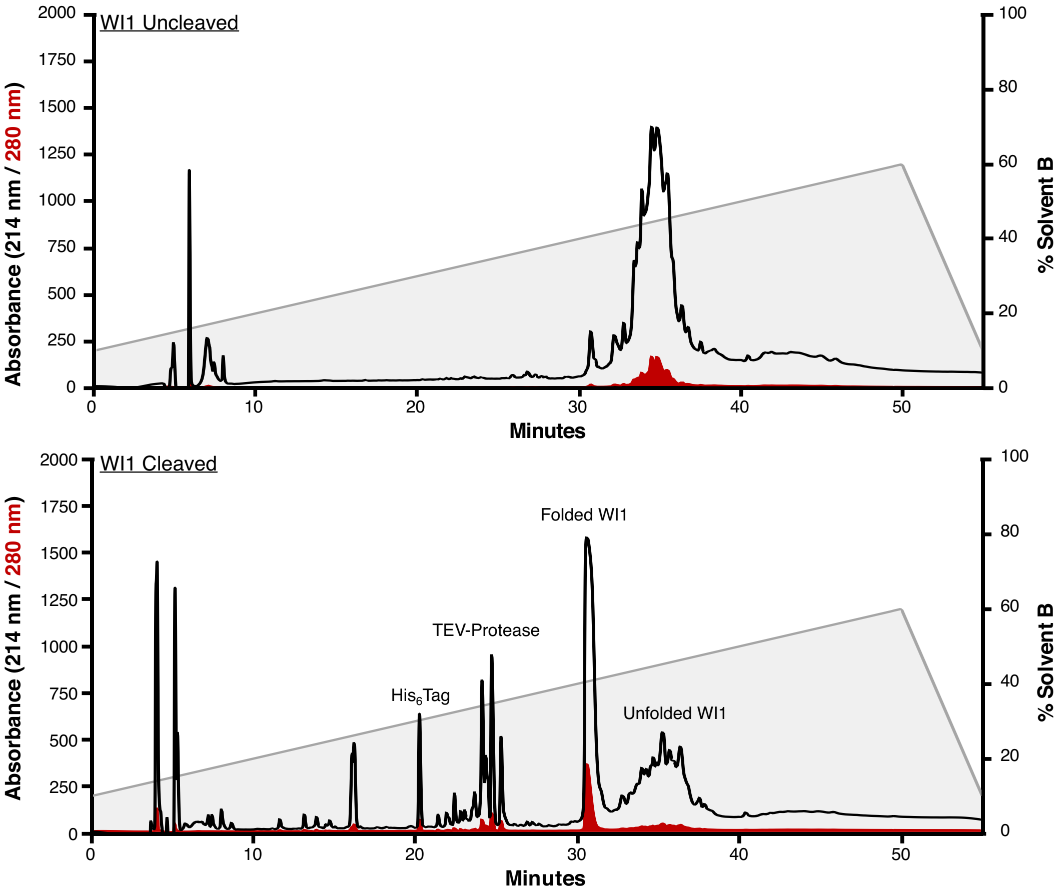


Figure 2.9. TEV protease cleavage of His_6-TEV-WI1 can be monitored by HPLC. (Top) Analytical HPLC trace with linear gradient of A:B (MilliQ dH_2O + 0.1% TFA : acetonitrile + 0.1% TFA) from 10-60% B over 50 minutes, of dialysed His_6-TEV-WI1 after Ni^{2+}-Affinity purification showing broad collection of peaks corresponding to multiple folded and misfolded states of uncleaved His_6-TEV-WI1. (Bottom) The same sample after cleavage with TEV-protease, showing free His_6-tag and the presence of a dominant peak at 30 min. Peak identity is annotated as confirmed using both ESI-MS (Dalhousie Mass Spectrometry Laboratory, Halifax, NS) and SDS-PAGE.

As measured by HPLC, TEV protease activity appears unaffected by glutathione concentration, as the front peak corresponding to free His_6 tag (16 min) appears even at 0:0 mM GSH:GSSG as well as all other conditions examined above. Interestingly, refolding appears to be dependent on the concentration of oxidized and reduced glutathione. Increasing total glutathione concentration affects folding efficiency. A 1:1 ratio of reduced and oxidized glutathione, with 1:1 mM GSH:GSSG appears optimal (Figure 2.11 A). Addition of 2:2 mM GSH:GSSG resulted in significant precipitation and lower yield of folded protein. SDS-PAGE analysis of cleavage reactions indicates that the best reaction conditions contained 2:1 mM GSH:GSSG, with excess GSH or GSSH giving a lower yield due to increased protein precipitation and lower refolding efficiency.

Several other conditions were found to impact refolding, including TEV protease concentration and pH (Figure 2.11 B,C). Increasing TEV protease concentration appears to result in lower overall yield, with an observed decrease in the intensity of the folded peak. However, at 4X TEV concentration the preceding shoulder peak is absent, possibly indicating more efficient cleavage. With decreasing pH, the folded peak is less intense in comparison to the trailing elutions, likely due to precipitation and lower refolding efficiency. As such, 2:1 mM GSH:GSSG and 1 μM TEV protease at pH 8.5 were found to result in optimal cleavage and refolding of WI1, resulting in the largest ratio of folded to trailing peaks while allowing for proper cleavage by TEV protease.

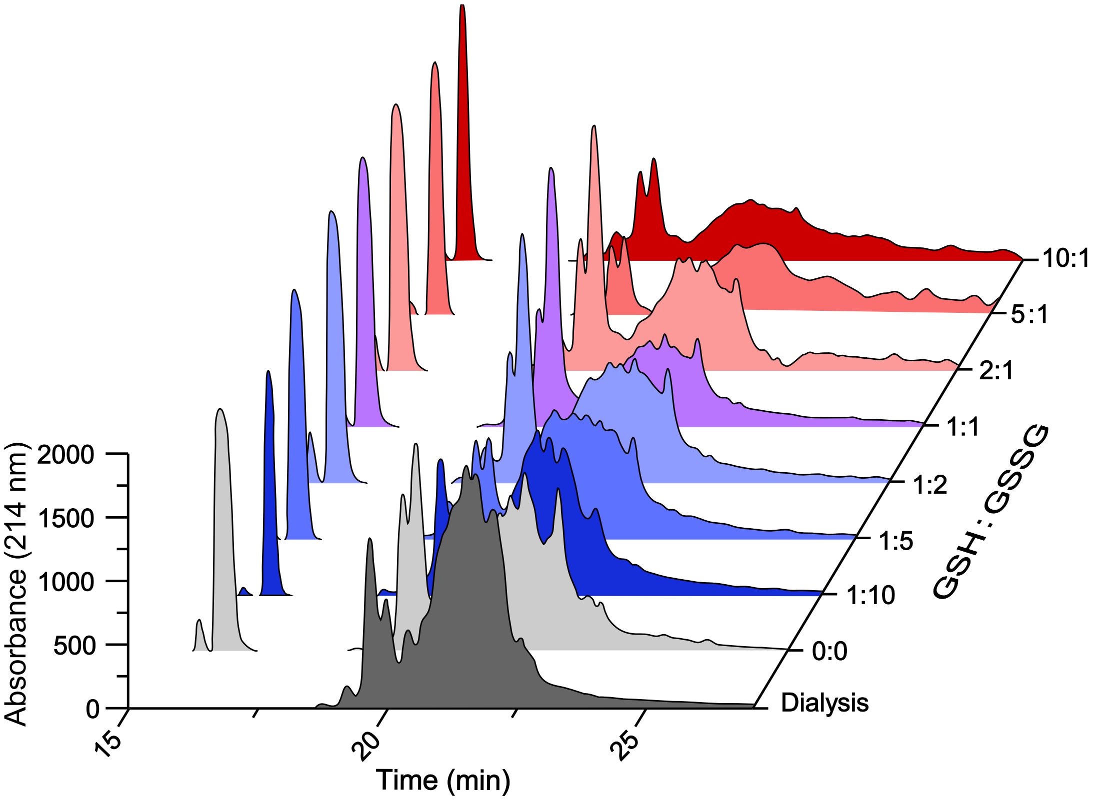


Figure 2.10. The balance of reduced and oxidized glutathione concentration affects hydrophobin refolding.

His_6-TEV-WI1 was cleaved with TEV protease and simultaneously refolded in buffers containing a series of glutathione mixtures from 10:1 mM to 1:10 mM GSH:GSSG and monitored by analytical HPLC with a linear gradient of A:B (MilliQ dH_2O + 0.1% TFA : acetonitrile + 0.1% TFA) from 10-60% B over 25 minutes. Increasing either GSH or GSSG resulted in significant protein precipitation and lower yield of folded WI1 (peak at 18 min), with 1 mM : 1 mM showing best refolding efficiency as determined by peak ratios.

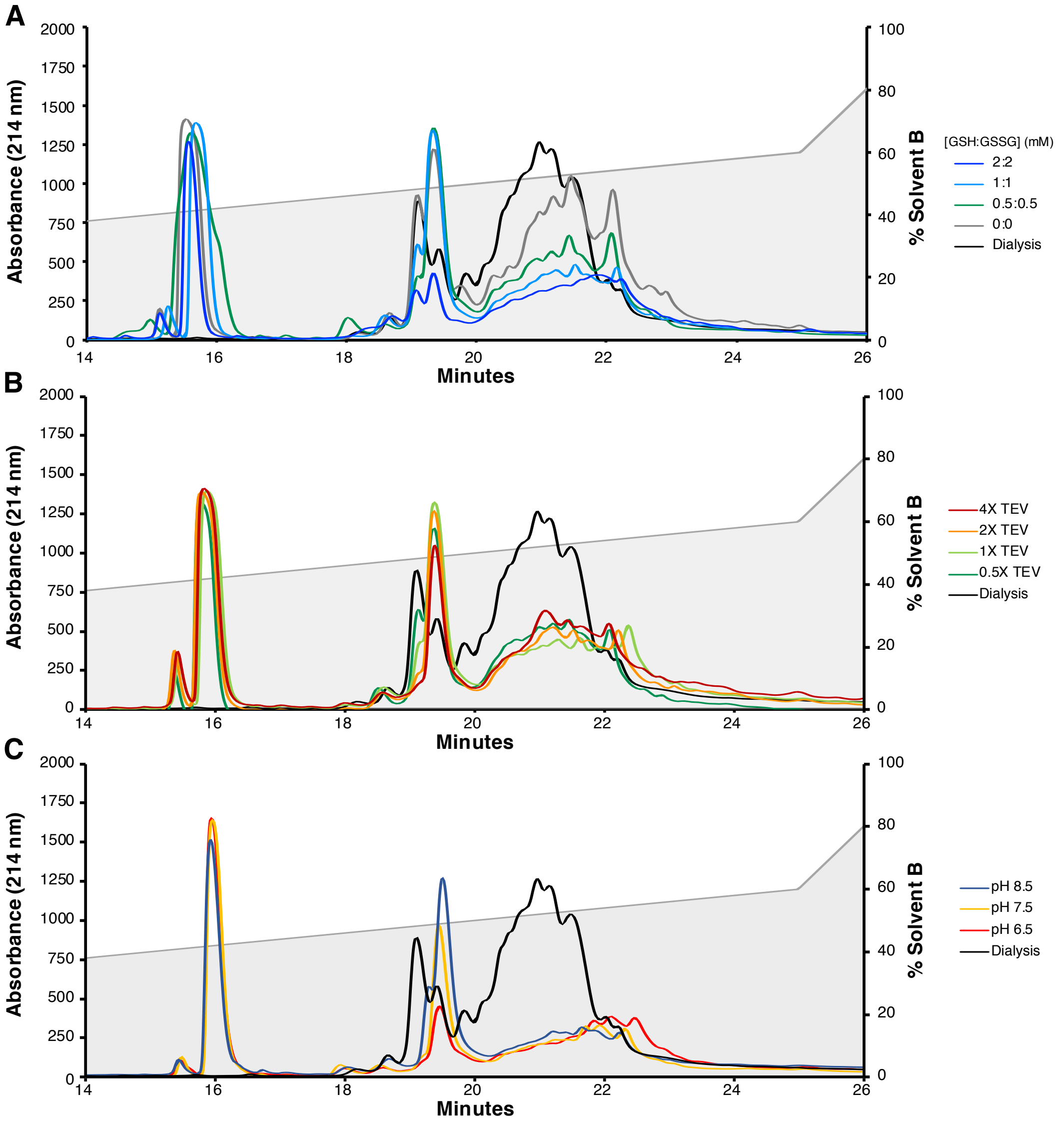


Figure 2.11. Glutathione concentration, protease concentration, and pH influence His_6-TEV-WI1 cleavage by TEV protease.

WI1 was cleaved by TEV protease and the reaction was analysed by analytical HPLC with a linear gradient of A:B (MilliQ dH_2O + 0.1% TFA : acetonitrile + 0.1% TFA) from 10-60% B over 25 minutes, The total concentration of GSSH:GSSH (A), the amount of TEV protease (B) and the pH (C) of the reaction were varied. A 1:1 mM ratio of GSSG:GSH and cleavage at pH 8.5 gave the best results while other conditions produce lower amounts of folded WI1.

2.3.5 Secondary Purification of Hydrophobins by FPLC

A combination of both FPLC and semi-preparative HPLC was used to purify samples used in the experiments detailed in the following chapters. The high purity of HPLC is advantageous when preparing samples that must survive weeks inside an NMR magnet. The use of HPLC in NMR sample preparation will be addressed in Chapter 3. However, the overall speed and high-throughput of FPLC makes it an attractive method for protein purification. For WI1, FPLC typically resulted in a chromatogram with three peaks (Figure 2.12). HPLC analysis of the pooled FPLC fractions reveals that folded WI1 is localized within the second peak (labelled B), with misfolded conformations present in the trailing region labelled C.

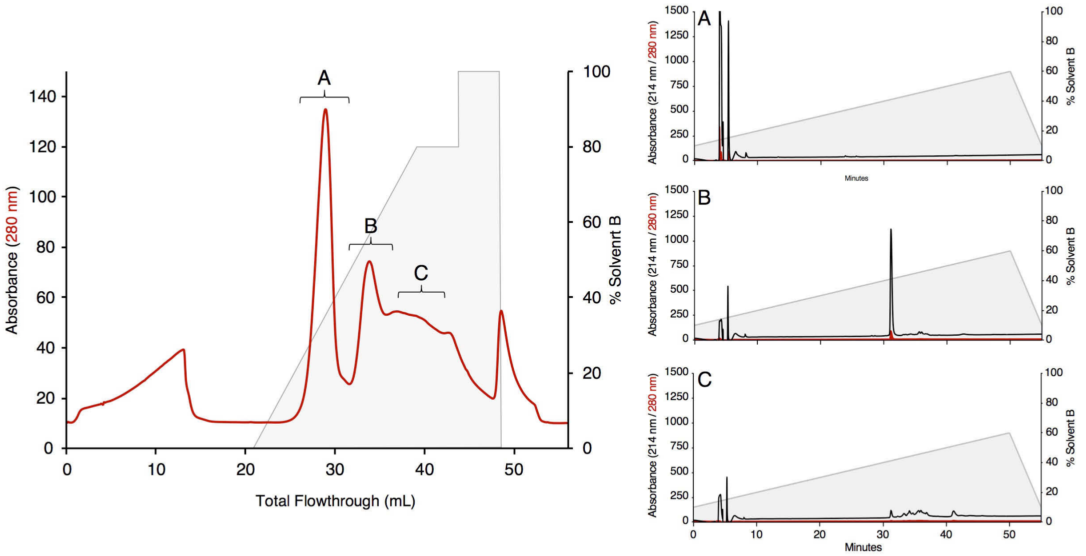


Figure 2.12. Anion exchange chromatography is able to efficiently purify WI1. (Left) FPLC chromatogram of unlabelled WI1 separated into three fractions, A, B and C. Sample was loaded in 20 mM Tris pH 8.0, and eluted with a liner gradient of NaCl up to 400 mM over 20 column volumes at 4 mL/min. (Right) The presence of WI1 in fraction B is confirmed by analytical HPLC with a linear gradient of A:B (MilliQ dH_2O + 0.1% TFA : acetonitrile + 0.1% TFA) from 10-60% B over 25 minutes was used. ESI-MS, and SDS-PAGE further confirmed that WI1 is present in fraction B.

Curiously, peak A, which is the largest in the FPLC chromatogram is unidentifiable but is a feature consistently seen in all hydrophobin purifications. HPLC analysis as well as FPLC experiments, ESI-MS, SDS-PAGE, and UV/Vis spectroscopy have all failed to detect anything in the fractions corresponding to peak A. Overall, FPLC is an effective means to purify hydrophobins, particularly with constructs that have a single folded state (as observed by HPLC) such as SC16 and SL1. However, for PC1, and WI1 there are a mix of folded and misfolded proteins present after cleavage and the superior resolution of HPLC is required to effectively purify these proteins.

2.4 Conclusion

The use of bacterial expression systems in the study of hydrophobins is well documented, as described in section 2.1, but I have developed a procedure to allow consistent production of purified class IB hydrophobins using SHuffle *E.coli* and a combination of Ni^{2+}-Affinity, FPLC, and HPLC. A summary of the optimized hydrophobin expression and purification methods used in this chapter is provided in Figure 2.3.

In spite of the initial work with SC3 proving unsuccessful, using the optimal conditions outlined above, heterologous expression and purification of SC16, SL1, WI1, and PC1 was achieved, as shown in Figure 2.13. SC16 expression showed the highest yield, resulting in approximately 10 mg of protein per 2 L expression. Moderate yields of SL1 and WI1 were obtained, of around 4 mg and 3 mg respectively per 2 L. PC1 consistently had the lowest yield, with 1-2 mg being produced per 2 L expression. Many factors contributed to these low yields, including precipitation (Figure 2.13 - SL1 and WI1) and incomplete cleavage (Figure 2.13 - SC16, SL1, and PC1). Despite the reverse

Ni^{2+} affinity flowthrough appearing pure by SDS-PAGE (Figure 2.13), HPLC indicates that this fraction contains both folded and misfolded protein (Figure 2.14). However, the use of semi-prep HPLC allows for efficient separation of this folded conformation, which is confirmed using ESI-MS to be the correct mass for the target hydrophobin containing four disulfide bonds (Appendix A).

Overall, final yield and refolding efficiency of hydrophobins remained low despite numerous attempts at optimization. Fusion protein construction, bacterial strains, expression temperature, and growth conditions were all examined for their effect on hydrophobin expression. However, many of these factors, such as length of expression and glucose concentration inconsistently impacted expression. Similar optimizations to cleavage and refolding conditions were performed in an effort to increase final yield to facilitate NMR spectroscopy experiments without the need for multiple costly isotope labelled protein preparations. No doubt there remain many details in hydrophobin protein expression and purification that could be improved to increase final protein yield; however, the new methods introduced here proved sufficient to allow for further characterization studies, as detailed in the following chapter.

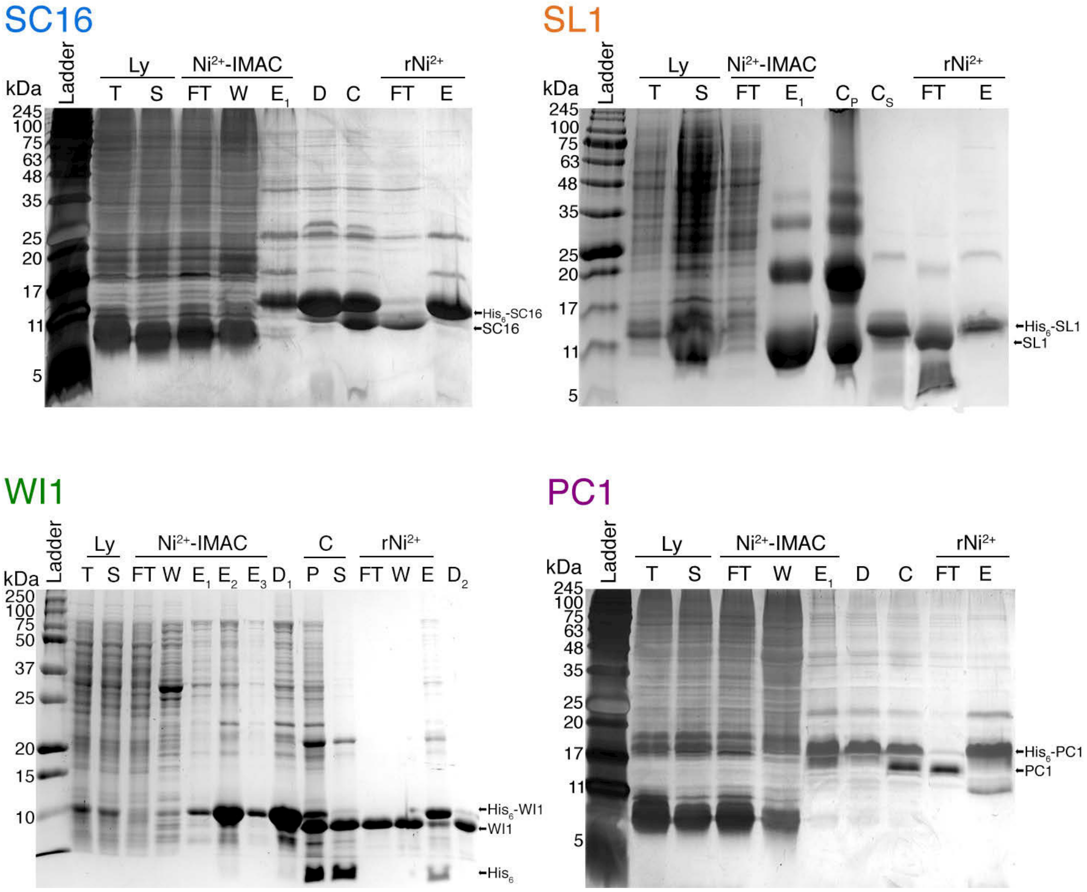


Figure 2.13. Expression and purification of target hydrophobins.
Using the optimized methods detailed in Chapter 2, adequate production of WI1, SL1, SC16, and PC1 was achieved. A summary of purification steps is shown visualized by SDS-PAGE stained with Coomassie blue (WI1) or silver stain (SC16/SL1/PC1). Overexpression with clear separation of each hydrophobin by Ni^{2+}-IMAC is shown. While a significant amount of protein precipitates during dialysis (D) and cleavage (C_P), presence of the free His_6-tag (lowermost band in WI1 and faintly in SC16) indicates successful cleavage. A second round of Ni^{2+}-IMAC purification successfully separates cleaved (FT) and uncleaved protein (E) before final purification by FPLC or HPLC.

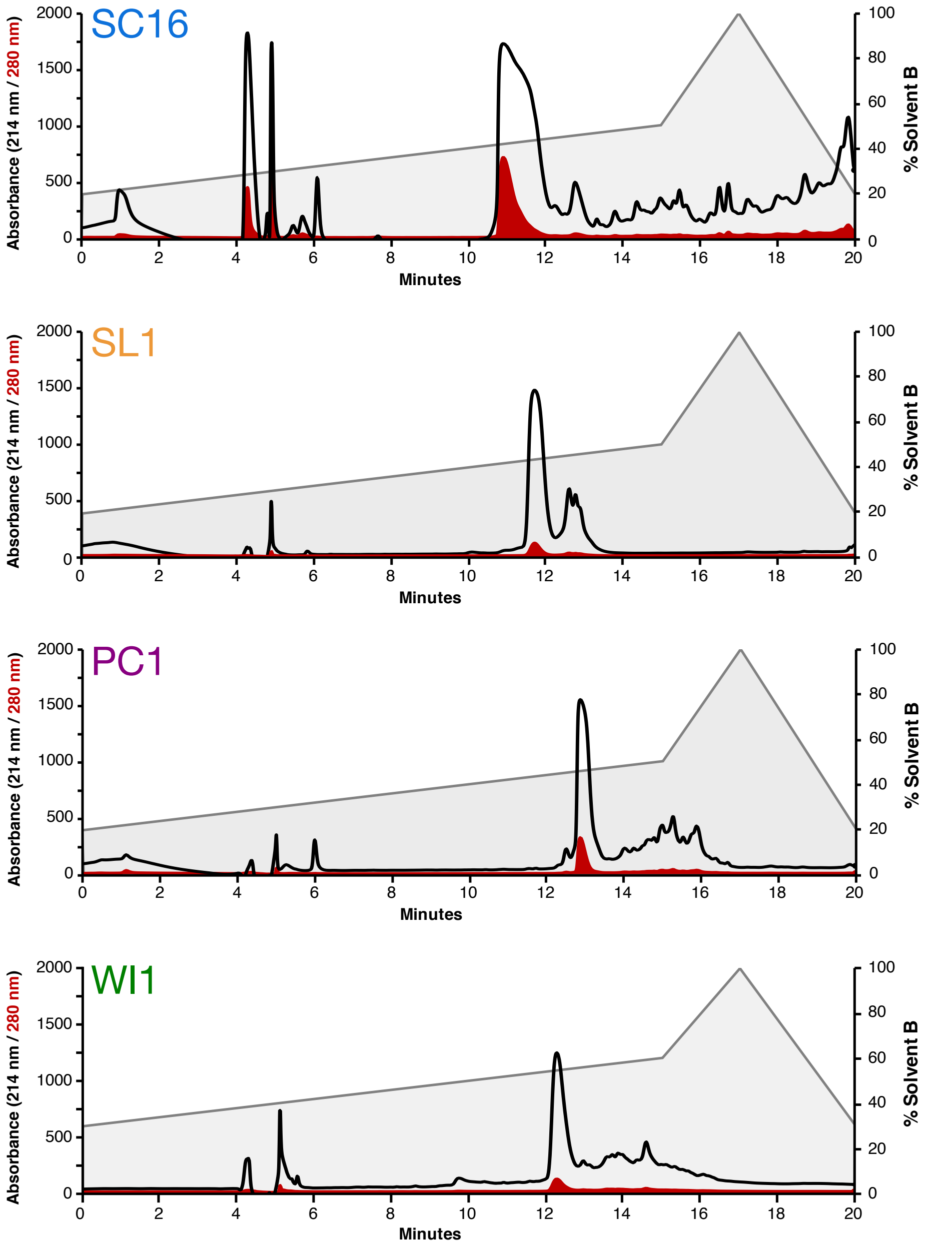


Figure 2.14. Semi-preparative HPLC purification of target hydrophobins. Semi-preparative HPLC with linear gradient of A:B (MilliQ dH_2O + 0.1% TFA : acetonitrile + 0.1% TFA) from 20-50% B (SC16/SL1/WI1) or 30-60% B (PC1) over 15 minutes, of each cleaved hydrophobin after rNi^{2+}-IMAC purification showing separation of folded protein (front peak) from broad collection of peaks corresponding to multiple folded and misfolded protein states.

Chapter 3: Determining the Atomic Resolution Structures of Class IB Hydrophobins

3.1 Introduction

In Chapter 1, the poor sequence and structure conservation among hydrophobins was established alongside the proposed Class IB subdivision, which has higher sequence conservation. To explore the proposed conserved features of this subclass, three Class IB targets were chosen for further analysis based on their high sequence conservation and unique sequence properties. In this chapter, I use the expression and purification procedures developed in Chapter 2 to optimize sample conditions for data collection and then determine the atomic resolution structures of Class IB fungal hydrophobins using NMR spectroscopy. I begin with a brief description of multidimensional NMR spectroscopy and the experiments used to explore protein structure and features. Due to the breadth of topics here, I focus on a more practical approach to interpreting NMR experiments and encourage the reader to refer to other texts for a more rigorous description of NMR spectroscopy.[139,140]

3.1.1 Using NMR Spectroscopy to Study Proteins in the Solution State

NMR spectroscopy has been instrumental in providing detailed structural information on molecules ranging from simple hydrocarbons to complex polypeptides.[139] From its origins studying water, ethanol, and wax to determine subatomic principles,[141] to the first protein structure solved by NMR,[142] to studying molecular dynamics of multisubunit protein complexes involved in disease,[143] NMR spectroscopy has become an invaluable tool in the field of structural biology. Many NMR experiments have been developed to provide insight into biologically significant events such as protein folding and unfolding,[144] enzymatic action,[145,146] protein interactions,[147,148] as well as aggregation and fibrillation in both soluble and membrane protein systems.[140]

NMR spectroscopy requires nuclei with a non-zero nuclear spin and an external magnetic field. Nuclei with spin > 1/2 are quadrupolar, leading to complex interactions with the external magnetic field of the NMR magnet and resulting in spectra that are more difficult to interpret. As such, using nuclei with spin 1/2 is preferred for solution-state NMR experiments as they provide the most straightforward spectra to interpret. Notably, some of the common nuclear isotopes in proteins have nuclear spin of 0 (e.g., ^{12}C and ^{16}O) or 1 (^{14}N) and are not visible or lead to complex spectra. Fortunately, ^{1}H, ^{13}C, and ^{15}N are all spin 1/2. ^{1}H is the most abundant hydrogen isotope (99.985%), while proteins can often be isotopically labelled with NMR-active ^{15}N (0.366% abundance) and ^{13}C (1.107% abundance), as described in Chapter 2.

In NMR spectroscopy, the chemical shift of a nucleus is a spectrometer-independent measurement of its magnetic environment, which is in turn determined by its chemical environment. Correspondingly, changes in chemical environment for a given nucleus give rise to changes in chemical shift. To enable specific study of a given molecule, the chemical shift of its nuclei must be first identified and assigned to a specific nucleus within the molecule. For simple organic molecules, this is a relatively straightforward process, as each chemical shift for each nucleus can be resolved in a 1 dimensional (1D) spectrum.

With increasing molecular size (e.g., larger peptides or proteins) the number of distinct NMR-active nuclei becomes considerably greater. This causes chemical shifts for a given type of NMR-active nucleus to overlap, preventing unambiguous assignment of all of the individual chemical shifts in a molecule and hampering its study. This limitation is often overcome through the use of multidimensional NMR spectroscopy, which allows for the deconvolution of these overlapping chemical shifts by correlating them to another nucleus (Figure 3.1). For proteins, correlations between ^{1}H, ^{13}C, or ^{15}N nuclei are commonly used to reduce signal overlap.

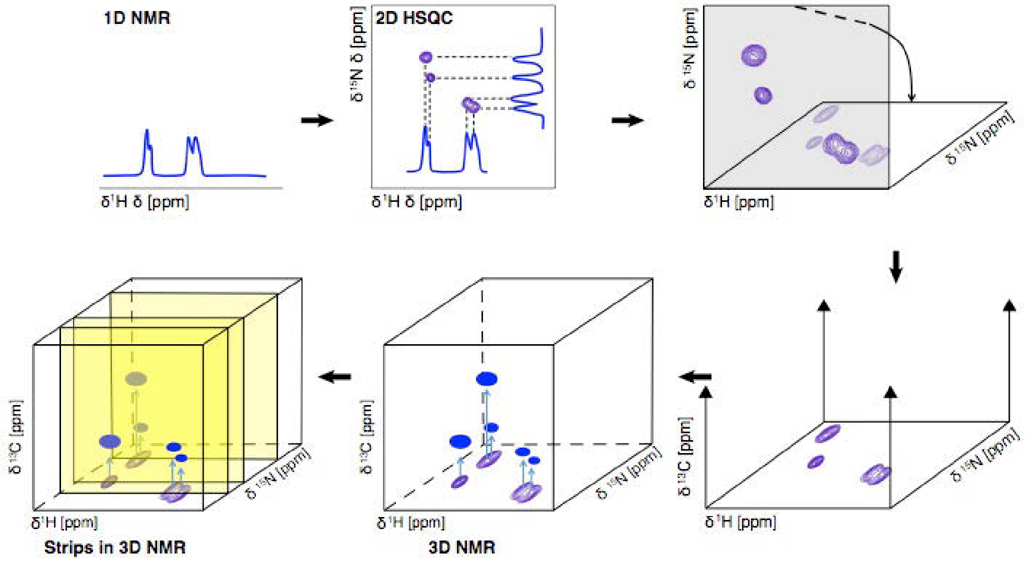


Figure 3.1. Multidimensional NMR allows for the deconvolution of overlapping chemical shifts.
In one dimensional NMR (top left) peaks are significantly overlapped, making accurate chemical shift measurement and assignment difficult. Addition of a second dimension (e.g. an ^{1}H-^{15}N HSQC) results in correlation of ^{1}H to ^{15}N nuclei, resulting in deconvolution of overlapped peaks. Signal overlap is further reduced by the addition of a third dimension (bottom middle) giving cross-peaks that are correlated to specific ^{1}H, ^{15}N, and ^{13}C nuclei. Planes of the resulting 3D spectra can be “stripped” to better visualize data on a computer monitor, in this case ^{1}H-^{13}C planes at different ^{15}N chemical shifts are shown, as is routinely done during backbone assignment (bottom left). Figure adapted from work presented by Dr. Kyungsoo Shin.

A common NMR experiment used to characterize polypeptides is the heteronuclear single quantum coherence (HSQC) experiment, in particular the ^{1}H-^{15}N HSQC. During this experiment, magnetisation from a ^{1}H nucleus is transferred to a directly-bonded ^{15}N nucleus, and then back to the ^{1}H for signal detection.[139] The ^{1}H-^{15}N HSQC provides a correlation between every proton that is chemically bonded (J-coupled with an appropriate coupling constant) to ^{15}N. Each measured signal has both the ^{1}H and ^{15}N chemical shift associated with it, therefore two dimensional plots are used to visualize data. A given ^{1}H-^{15}N bond will correspond to a cross-peak in this spectrum. In proteins each residue has an amide proton attached to a nitrogen as part of peptide bond, with the exception of proline. In proteins, the ^{1}H-^{15}N HSQC has a limited number of signals since only protons bonded to ^{15}N nuclei in proteins are observed and overlap is reduced by distributing the signal into two dimensions.

The resulting cross-peaks of an ^{1}H-^{15}N HSQC show chemical shift correlation for the amides of peptide bonds, asparagine and glutamine side chains, and of amines in basic residue side chains. By monitoring chemical shift differences, which correspond the local chemical environment, this experiment can be used to distinguish between disordered and folded proteins and observe changes in structural features. As signals are generated for nearly every residue with the HSQC experiment and are influenced by local chemical environment, the resulting spectra are excellent “fingerprints” of protein composition. If the protein is well-folded, the peaks are usually well-dispersed, and most of the individual peaks can be distinguished. In contrast, if peaks are severely overlapped and clustered around the middle of the spectrum, the presence of significant unstructured elements in the protein is likely.

3.1.2 The Backbone Walk

In order to determine the identity of each protein residue that gives rise to each peak in a ^1H-^{15}N HSQC, sequential NMR assignment (often termed as a 'backbone walk') is often used.[139] This is achieved by exploiting through-bond connections to build a map of connected resonances. Many experiments can be interpreted to determine these sequential connections, with the HNCO[149,150], HN(CA)CO[151,152], HN(CO)CACB[152]and HNCACB[153] experiments being the most commonly employed. These experiments are based on the transfer of magnetisation between ^1H, ^{15}N and ^{13}C nuclei that are directly bonded. The resulting ^1H and ^{15}N chemical shifts of residue i are modulated by the resonance frequencies of ^{13}C nuclei (C_{i-1} in HNCO and HN(CO)CACB; C_{i-1} and C_i in HN(CA)CO and HNCACB), giving rise to cross-peaks corresponding to each of the associated chemical shifts. These experiments can be grouped together in pairs, with one experiment in each pair detecting resonances for C_{i-1} and the other for both C_{i-1} and C_i (where C_i refers to a residue of interest, while C_{i-1} directly precedes it in sequence).

For example, for the HNCO experiment the initial magnetisation is passed from the H_N to N of residue i and then transferred specifically to carbonyl (C') of residue i-1, with a subsequent return back to H_N through N. The resulting cross-peak has three associated chemical shifts, ^1H and ^{15}N for HN and N of residue i, respectively, and ^{13}C for the C' of residue i-1. In the paired HN(CA)CO experiment, magnetisation is transferred from H_N and N through C_α of residue i to the C' of both the residue at position i and that at i-1. This results in two observable cross peaks having shared ^1H and ^{15}N frequencies from residue i but two different ^{13}C frequencies: one for the C' nuclei of residue i and one for the C' nuclei of residue i-1. Using these pairs of experiments, one for C' and another for C_α and C_β it is possible to sequentially walk through each connection in the backbone by taking "slices" in the ^1H and ^{15}N dimension (Figure 3.2).

The backbone walk forms the foundation from which assignment of all protein nuclei begins. Since each amino acid has expected chemical shift values for each nucleus,[154] it is possible to determine the identity of a given residue based on the chemical shift of its nuclei. With completion of the backbone walk, C', C_α, C_β and NH chemical shifts can be assigned for most observable peaks. Several other experiments can be used in conjunction with the backbone experiments to assign nuclei, such as the CCONH which correlates all carbon chemical shifts of the residue i-1 to the NH pair of residue i.[155] For side chain chemical shift assignment, proton based experiments are required, such as ^1H-^1H correlation spectroscopy (COSY)[156], ^1H-^1H total correlation spectroscopy (TOCSY)[157] or the HCCH- TOCSY.[158] Each of these experiments allows assignment of individual ^1H and ^{13}C nuclei. In ^1H-^1H COSY, magnetisation is transferred between protons that are connected to each other by three chemical bonds. Similarly, in TOCSY experiments magnetisation from a proton is transferred to all other neighboring protons so long as each intervening ^{13}C or ^{15}N has at least one proton attached to it. As the carbonyl carbon of each peptide bond has no attached protons, COSY and TOCSY magnetisation transfers cannot occur between amino acid residues, isolating protons from each residue into their own spin system. These proton experiments in combination with assignments from backbone experiments allow for the unambiguous assignment of all ^1H, ^{13}C, and ^{15}N chemical shifts in a protein, which is required to interpret through-space NOESY experiments.

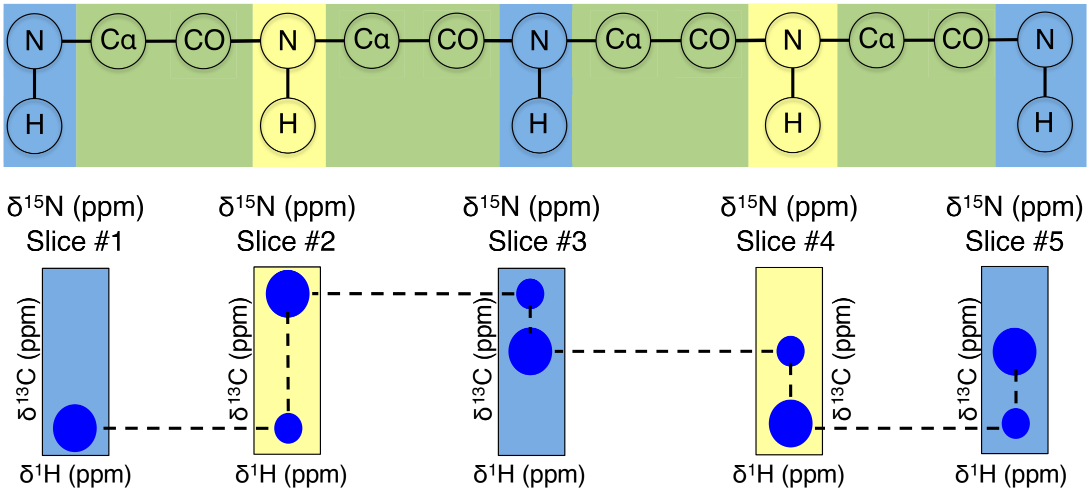


Figure 3.2. By using pairs of NMR experiments the sequence of a polypeptide can be determined.
(Top) The amide backbone of a theoretical protein is shown with constituent C', C_α and NH pairs. (Bottom) schematic of an HN(CA)CO experiment backbone walk shows how sequential NH parings can be linked through connecting ^{13}C chemical shifts in other NH slices. For example, during an HNCO experiment, in slice #5 the peak corresponds to C'_{i-1}, connecting to slice #4 in which the same resonance now corresponds to C'_i and a new C'_{i-1} is observed above. The process is repeated for each NH slice to build a network of through-bond connections, allowing assignment of specific resonances to peptide residues. Figure adapted with permission from work presented by Dr. Kyungsoo Shin.

3.1.3 Using NMR Spectroscopy to Measure Inter-Atomic Distances

Although through-bond connections were probed to assign resonances to nuclei in proteins, these experiments reveal very little about protein tertiary structure. To directly probe molecular structure, interatomic distances between nuclei can be measured using nuclear Overhauser effect (NOE) spectroscopy (NOESY). The NOE is due to dipole-dipole interactions between nuclei, which cause nuclei to interact with each other through space instead of through chemical bonds.[139,159] The amount of magnetisation transferred between nuclei is related to the distance (r) between them by a factor of r^{-6}, allowing for distances between nuclei to be estimated. Due to the large number of ^{1}H nuclei in a protein, proton-proton NOESY spectra suffer from signal overlap and can be difficult to analyse. In this work, 3D N-rooted (NNOESY) and C-rooted (CNOESY) spectra were collected instead. For these spectra, magnetisation is exchanged amongst proximal ^{1}H using the NOE, then correlated to their attached ^{15}N or ^{13}C nucleus. This results in a 3D spectrum with significantly less peak overlap (Figure 3.1).[160,161] If the ^{1}H, ^{13}C, and ^{15}N chemical shift assignments are known (e.g. by using the backbone walk), observed through space NOE contacts can be attributed to specific nuclei. A summary of the experiments used in this work for structural studies is given in Table 3.1.

Table 3.1. Summary of NMR experiments collected for structural characterization

Target	Experiment	Details / Assignments
Backbone Assignment (^{1}H-^{15}N HSQC)	HNCO	C'_{i-1} & C'_{i}
	HN(CA)CO	C'_{i}
	CBCA(CO)NH	$C\alpha_{i-1}$ / $C\beta_{i-1}$ & $C\alpha_{i}$ / $C\beta_{i}$
	HNCACB	$C\alpha_{i}$/$C\beta_{i}$
	CC(CO)NH	$C\alpha_{i-1}$ / $C\beta_{i-1}$ / $C\gamma_{i-1}$/ $C\delta_{i-1}$ & $C\alpha_{i}$ / $C\beta_{i}$ / $C\gamma_{i}$ / $C\delta_{i}$
Initial Proton Assignment	HBHA(CO)NH	$H\alpha_{i-1}$ / $H\beta_{i-1}$ & $H\alpha_{i}$ / $H\beta_{i}$
	H(CCO)NH	$H\alpha_{i-1}$ / $H\beta_{i-1}$ / $H\gamma_{i-1}$/ & $H\alpha_{i}$ / $H\beta_{i}$ / $H\gamma_{i}$
Assignment of ^{1}H-^{13}C HSQC	HCCH-COSY	Adjacent protons in residue spin system
	CCH-TOCSY	All protons in residue spin system
Structural and Dynamics Probes	D_2O Exchange ^{1}H-^{15}N HSQC	Hydrogen bonded amides
	^{15}N-NOESY HSQC	Inter-residue amide interactions within 6 Å
	^{13}C-NOESY HSQC	Inter-residue C-H bond interactions within 6 Å
	Interleaved {^{1}H}-^{15}N HetNOE HSQC	Per residue dynamics, simultaneous acquisition of reference ^{1}H-^{15}N HSQC

3.1.4 Protein Structure Calculation

Each peak in a NOESY spectrum can be used to define a specific distance restraint between two nuclei, where 5-6 Å is the maximum distance over which the NOE contributions can be observed.[139] Thousands of these NOE contacts can be detected and form the bulk of restraints used to calculate and refine a protein structure. The set of distance restraints is input into a molecular simulation program to generate a three-dimensional protein structure. In this work, ARIA (Ambiguous Restraints for Iterative Assignment)[162] was used for automated NOE assignment and NMR structure calculation, which in turn uses an external structure calculation engine (CNS) to generate molecular structures based on the provided restraints.[163] Structure calculation is performed using a simulated annealing protocol.[164,165] In simulated annealing, the structure of a protein randomly perturbed to generate a new structure. If this new structure fits known restraints better, then it may be kept as the new protein structure. This process is repeated thousands of times, generating a final structure with the best possible agreement with the experimental data.

In addition to NOESY-derived distance restraints, other restraints can be added such as disulphide bonds, hydrogen bonds, and dihedral angle restraints, which can be predicted by software such as DANGLE.[166] Since the refinement process is dependent on random changes to a starting structure no two generated structures will be identical, For this reason, NMR structure calculations on proteins are carried out multiple times, and the resulting structures are superimposed and represented as an ensemble of structures that all satisfy the experimental restraints.

3.1.5 Examining Protein Dynamics

Proteins are not entirely rigid structures, but instead are inherently dynamic and this conformational flexibility often enables the ability to perform specific functions or respond to environmental changes, ligand binding, and chemical modifications. This is especially true in hydrophobins, where global structural rearrangements[39] and highly dynamic inter-cysteine loop regions[58] have both been shown to play a critical role in rodlet self-assembly.

NMR spectroscopy is one of the few methods that can directly probe molecular motions on a range of timescales from ps – ms. Concerning this work, residue specific dynamics were measured using ^{1}H-^{15}N Heteronuclear NOE (HetNOE) experiments. As described above, the NOE is a cross-relaxation process arising from a dipole-dipole interaction between nuclei in sufficiently close proximity. The NOE effect is influenced both by the distance between nuclei and the rate at which nuclei move relative to each other. In the ^{1}H-^{15}N heteronuclear NOE experiment, the NOE is measured between the ^{1}H and ^{15}N nuclei of the amide bond, meaning any observed difference in NOE are due to changes in molecular motion, since the length of the amide bond is effectively constant. Molecular motion is due to both global tumbling of the molecule and local variation of backbone motion that corresponds to segments of flexibility or rigidity in the protein. If global tumbling is assumed to be relatively constant in a sample, then the NOE for each residue of a protein corresponds to the local motions of that residue.

3.2 Materials and Methods

3.2.1 NMR Sample Preparation

Uniformly $^{13}C/^{15}N$ labeled or ^{15}N labeled hydrophobins were purified by either FPLC (WI1 and SL1) or HPLC (PC1), as described in Chapter 2, and protein identity was confirmed using SDS-PAGE (Figure 2.13) and mass spectrometry (Appendix A). Lyophilized samples were reconstituted in NMR buffer containing 1.1 mM 4,4−dimethyl−4−silapentane−1−sulfonic acid (DSS), 1.1 mM NaN_3, 22.22 mM 2-(N-morpholino)ethanesulfonic acid (MES), 55.5 mM NaCl. This solution was concentrated by ultrafiltration with more NMR buffer being added to ensure proper solution composition and pH. Sample pH in all cases was confirmed to be 6.50 ± 0.05. This concentrated sample was then diluted with D_2O obtained from Cambridge Isotope Laboratories, to a final concentration of 10% D_2O. For $^{13}C/^{15}N$ labelled samples 350 µL of sample (SL1: 700 µM, WI1: 250 µM, PC1: 800 µM) was added to susceptibility matched tubes (Shigemi Inc). For ^{15}N labelled samples, standard 5 mm NMR tubes were used, with a concentration ranging from approximately 40 µM to 500 µM in 650 µL. For D_2O exchange experiments, after data collection, NMR samples were lyophilized and reconstituted with an equal volume of D_2O and incubated for 30-60 minutes before data collection.

3.2.2 NMR Spectroscopy

NMR spectra were collected on a variety of NMR spectrometers at 298.15KH-^{15}N HSQC experiments with sensitivity improvement and decoupling during acquisition (32 scans, sweep width 32/14 ppm in F1/F2, size of fid 128/2048 in F1/F2, recycle delay 1.2 s) were collected on a 500 MHz Avance spectrometer (Bruker Canada) using a double-resonance 5-mm BBFO SmartProbe located at the Nuclear Magnetic Resonance Research Resource (NMR^3; Dalhousie University, Halifax, NS). Instances for which increased sensitivity or resolution was desired, a 700 MHz Avance III spectrometer (Bruker Canada) with a triple-resonance 5-mm TCI CryoProbe was employed (Biomolecular Magnetic Resonance Facility (BMRF), NRC, Halifax, NS). Experiments for structural elucidation of SL1 were collected on a 600 MHz Varian INOVA NMR spectrometer equipped with a triple resonance probe (Queen’s University, Kingston, ON). Data for WI1 and PC1 were collected at the Québec/Eastern Canada High Field NMR Facility (McGill University, Montreal, QC) on 500 MHz Varian INOVA and 800 MHz Bruker Avance Neo NMR spectrometers equipped with triple-resonance room temperature and cryoprobes, respectively.

For all ^{13}C/^{15}N labeled samples, ^{1}H-^{15}N and ^{1}H-^{13}C HSQC experiments[167–169] (hsqcetfpgpsi2 and hsqcetgpsp pulse programs respectively) were collected to assess sample quality. HNCO,[169–171] HN(CA)CO,[171,172] HNCACB,[173,174] HN(CO)CACB[175,176] triple resonance experiments (hncogp3d, hncacogp3d, hncacbgp3d, and cbcaconhgp3d, pulse programs respectively), were collected and used to assign the backbone resonances. To assign sidechain resonances, CC(CO)NH, [177–181] HBHA(CO)NH,[175,176] H(CCO)NH,[177–181] HCCH-COSY,[182] and CCH-TOCSY[182] experiments (hccconhgp3d3, hbhaconhgpwg3d, hcchdigp3d2, hcchcogp3d, and hcchcogp3d, pulse programs respectively) were collected and analysed. Distance restraints were derived from ^{15}N-edited NOESY-HSQC[183–186] (noesyhsqcf3gpwg3d) and both aliphatic and aromatic ^{13}C-edited NOESY-HSQC[187] (noesyhsqcetgp3d) spectra with 100 ms (SL1) and 110 ms (WI1 and PC1) mixing times. ^{1}H-^{15}N heteronuclear NOE experiments[188] (hsqcnoef3gpsi) were collected on ^{15}N labeled samples with saturated and reference spectra collected in an interleaved manner with a saturation time of 4 sec for SL1 and 5 sec for WI1 and PC1. Spectra were processed using a combination of NMRPipe (9.9 Rev)[189] and Bruker Topspin (version 4.0.6) and analysed using CCPNMR Analysis (version 2.4).[190]

3.2.3 NMR Data Processing and Structure Calculation

Backbone and side-chain ^{1}H, ^{13}C, and ^{15}N chemical shift values, manually assigned NOESY peak lists, and dihedral angle restraints generated by Dihedral Angles from Global Likelihood Estimates (DANGLE)[166] were used as inputs to ARIA version 2.3[162] for automated NOE assignment and structure calculation. Due to the known critical role within hydrophobins, four disulphide bond restraints were added to the structure calculation. Hydrogen bond restraints were added for residues that had strong signals remaining in a ^{1}H-^{15}N HSQC after exchange with D_2O. Each ARIA structure calculation run was done in eight iterations followed by a final refinement step in water performed by CNS 1.21.[190,191] For iterations 1 to 7, the best 7 of 20 calculated structures were progressively used as the seed for the following iteration. For iteration 8, the 20 lowest energy structures out of 100 calculated were kept and used for the final water refinement step. Between each run, structural violations were noted and resolved manually. For calculation of the final structural ensemble every NOE restraint was manually verified in addition to any remaining structural violations.

The quality of the final ensemble of structures was assessed through Recall Precision and F-factor analysis[192], PROCHECK[193] and the protein structure validation suite.[194] The PyMOL Molecular Graphics System, (Version 2.0 Schrödinger, LLC.) with APBS electrostatics suite[195] was used for visualization, distance measurements, figure generation, and surface charge map generation. Chemical shifts and the final structural ensemble of SL1, WI1, and PC1 have been deposited into the Protein Data Bank (PDB) and the Biological Magnetic Resonance Data Bank (BRMB; Accession numbers 5W0Y, 6E9M, 6E98 and 30301, 30505, 30504, respectively).

3.3 Results and Discussion

3.3.1 Optimization of NMR Sample Conditions

Using the purification methods outlined in Chapter 2, $^{13}C/^{15}N$-labelled samples of SL1, WI1, and PC1 were prepared and resuspended in NMR buffer. Anion exchange chromatography was used as the final purification step for SL1 and WI1, while HPLC was required for PC1 because anion exchange chromatography was unable to separate the folded and misfolded states of PC1. PC1 should have approximately 80 backbone peaks (proline residues are not detected), although dynamic N-terminal residues are often poorly resolved. The superior spectral quality for HPLC-purified PC1 is shown in Figure 3.3; relative to HPLC, purification by FPLC results in a convoluted central region and more peaks than expected. HPLC purification also resulted in samples which were significantly more stable over the duration of NMR experiments. A complete dataset for structure determination can take ~2 weeks of NMR spectrometer experiment time and sample degradation complicates resonance assignment and data processing and may result in a dataset that is unsuitable for analysis.

Initial spectra of SL1 and PC1 were of high quality and required no further sample optimization, but instability was repeatedly observed for samples of WI1 (Figure 3.4). This was a significant barrier in structural characterization, especially since experiments could not be repeated if required. Samples purified by HPLC consistently showed much higher stability due to higher purity prior to NMR sample preparation. Buffer conditions were varied in an attempt to improve the spectral quality of WI, buffer composition and the addition of anti-aggregation additives[196,197] was examined to determine the effect on sample stability and peak separation for WI1. However, there was no significant improvement to spectral quality (Figure 3.5) with any of these modifications. Instead, spectra were of noticeably lower quality with increased line broadening upon addition of glycerol or more salt, in particular. Therefore, the same buffer (20 mM MES pH 6.5, 50 mM NaCl) used for data collection of SL1, WI1, and PC1.

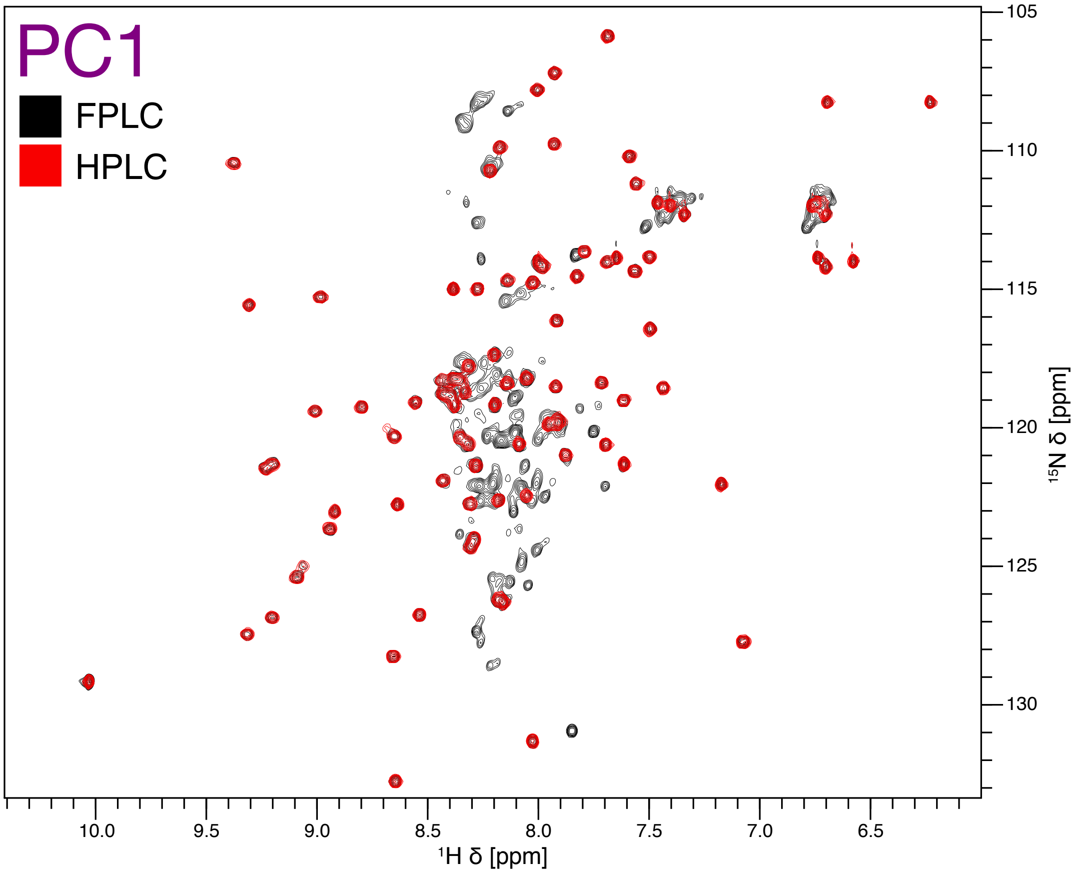


Figure 3.3. ^{1}H-^{15}N HSQC spectra of HPLC purified PC1 are high quality. ^{1}H-^{15}N HSQC spectra of PC1 purified by either FPLC or HPLC. Purification by FPLC results in a poorly resolved central region, additional peaks, and heterogeneous peak intensity, indicative of a heterogenous sample of PC1. HPLC purification results in a spectrum with few overlapped peaks and uniform peak intensity, suggestive of a homogenous sample.

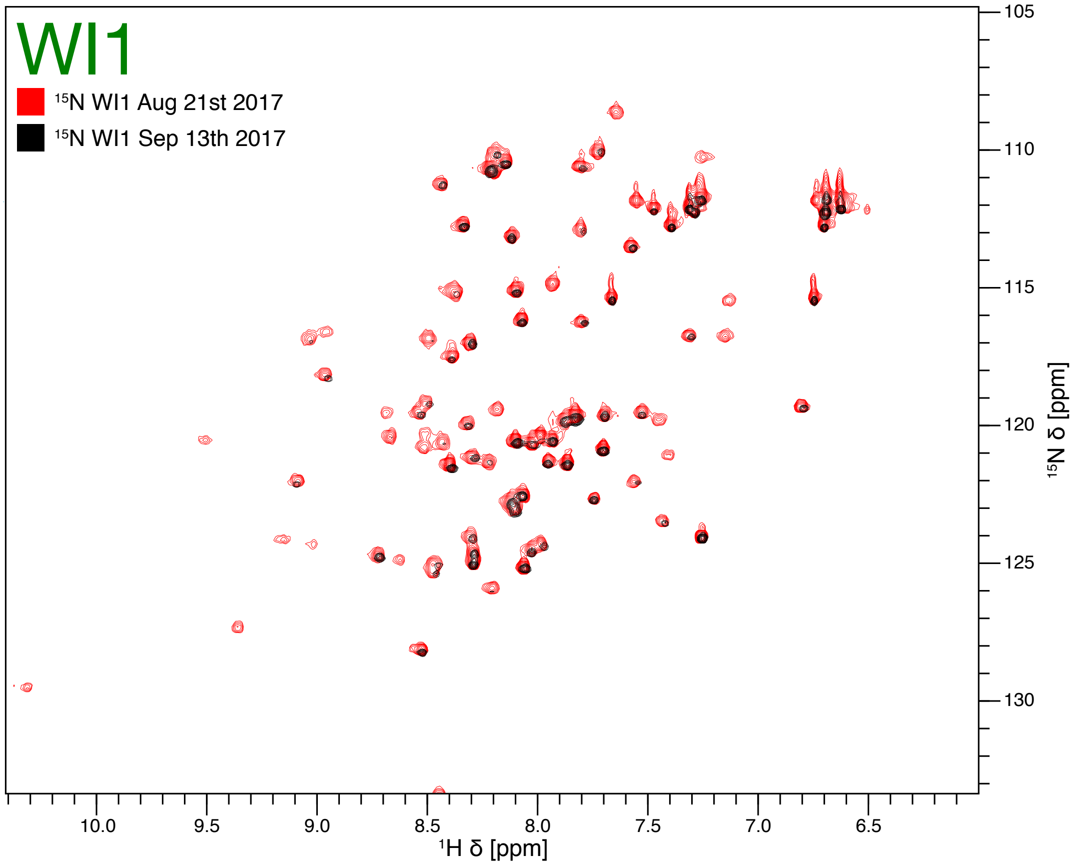


Figure 3.4. Instability of WI1 samples as observed by ^{1}H-^{15}N HSQC.

^{1}H-^{15}N HSQC spectra of WI1 purified by FPLC shows significant amount of non-uniform signal loss after only a few weeks at 25°C. No precipitation within the sample was visible at either timepoint. This instability prevents collection of accurate structural assignment and characterization, especially if experiments need to be repeated.

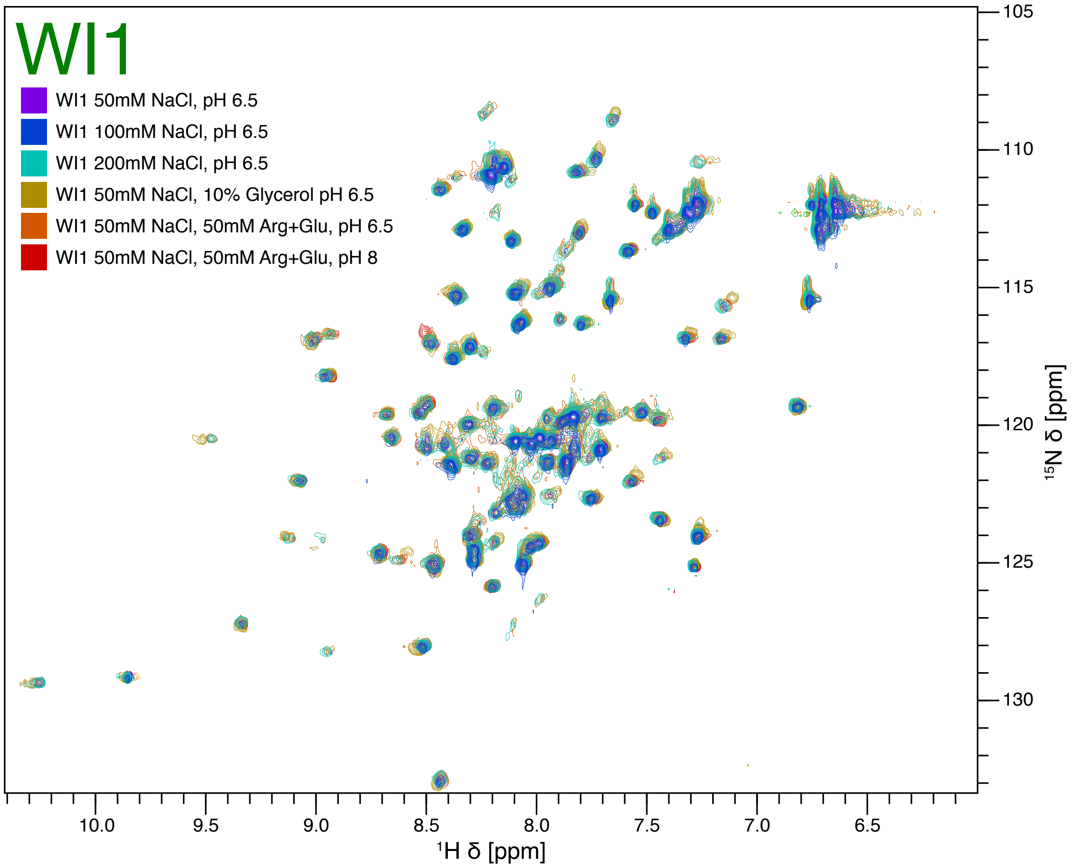


Figure 3.5. Varying buffer conditions does not improve WI1 ^{1}H-^{15}N HSQC spectral quality.

^{1}H-^{15}N HSQC spectra were collected of WI1 in 20mM MES pH 6.5, 10% D_2O with various additives. The convoluted core region of WI1 coupled with known sample instability makes peak assignment more challenging. Spectra quality was evaluated at different NaCl concentrations as well as after addition of glycerol, or arginine and glutamine salts at pH 6.5 and pH 8.0. No significant improvement in spectra quality was observed with addition of any additive shown.

3.3.2 SL1, WI1 And PC1 Adopt Folded Structures Based Upon ^{1}H-^{15}N HSQC Spectral Behaviour

High quality NMR data were collected for SL1, WI1, and PC1 in identical sample conditions. Figure 3.6 and Figure 3.7 show collected ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N/^{13}C labeled SL1, WI1, and PC1 along with previously published data for SC16 from Gandier et al.[42] With 78 , 97, and 91 peaks of an expected 81, 84, and 77peaks being assigned for SL1, WI1 and PC1, respectively. These spectra indicate that each protein adopts a single well-folded confirmation with stable tertiary structure. Spectra for SL1 and PC1 appear well resolved, with good dispersion and low peak overlap, and are of uniform intensity with the expected number of peaks being observed. WI1 spectra are of similar quality to SL1 and PC1, but contain less dispersed peaks that are fairly broad and a subset of extra weaker peaks are observed (final assignment included 91 of 119 observed in ^{1}H-^{15}N HSQC). The central region of the ^{1}H-^{15}N HSQC for WI1 contains several broad or poorly defined peaks, possibly due to lower sample homogeneity obtained from FPLC purification, or an uncharacterized slow assembly process in solution. However, this did not significantly impede further data analysis or structural characterization. Using a backbone walk approach, the high quality of the data allowed for 96%, 94%, and 97% of backbone, and of 100% of NMR assignable side chain ^{1}H, ^{15}N, and ^{13}C resonances to be identified for SL1, WI1, and PC1, respectively (Figure 3.6 and Figure 3.7).

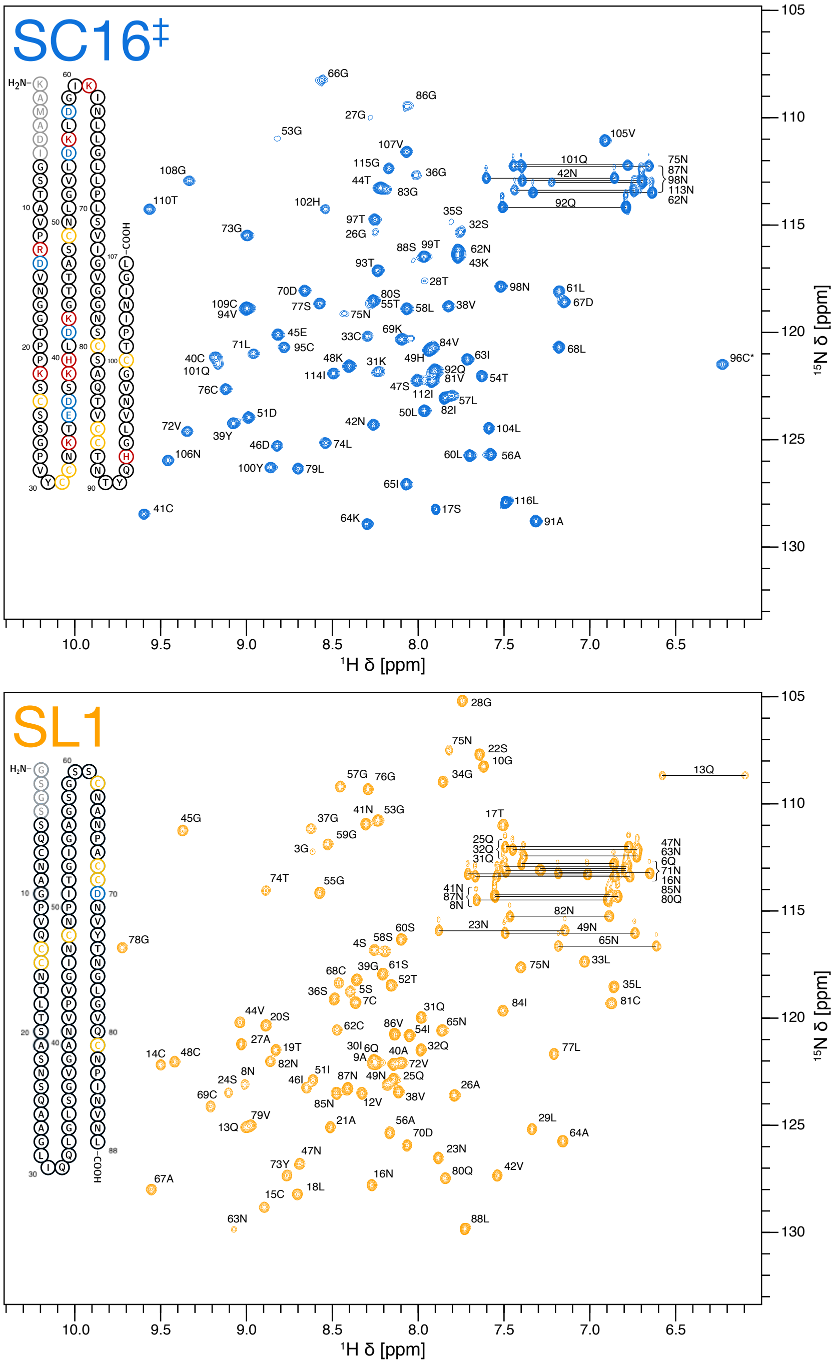


Figure 3.6. ^{1}H-^{15}N HSQC NMR spectra of ^{13}C/^{15}N-labelled SC16 and SL1. The ^{1}H-^{15}N HSQC spectrum of SC16 and SL1 is shown with peak assignments indicated. Sidechain amides are denoted by a horizontal line. The sequence of each protein is shown with cysteine residues highlighted in yellow, acidic residues in blue, and basic residues in red. ‡: Data for SC16, shown for here for comparison, was acquired by Gandier et al. as previously described.[42]

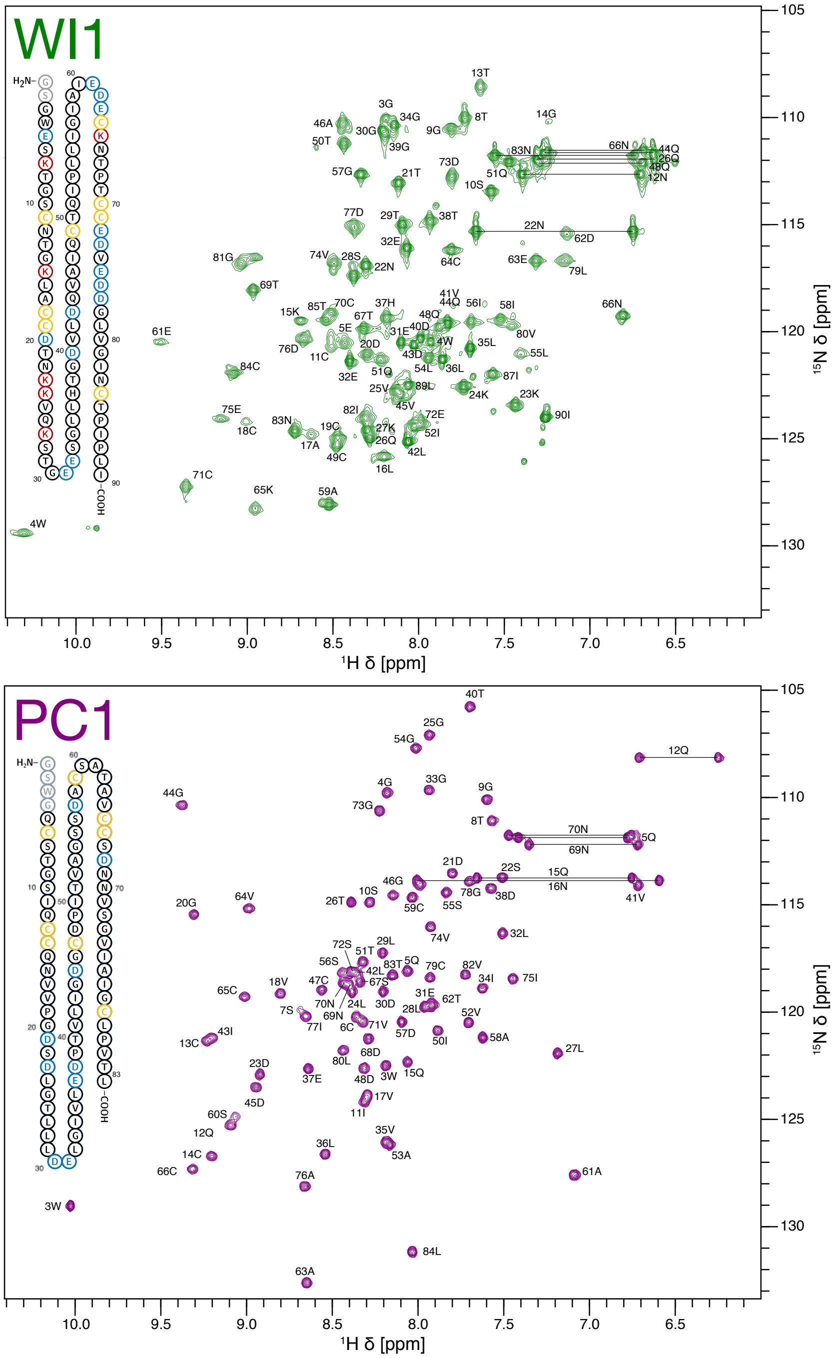


Figure 3.7. ^{1}H-^{15}N HSQC NMR spectra of ^{13}C/^{15}N-labelled WI1 and PC1. The ^{1}H-^{15}N HSQC spectrum of WI1 and PC1 is shown with the peak assignments indicated. Sidechain amides are denoted by a horizontal line. The sequence of each protein is shown with cysteines highlighted in yellow, acidic residues in blue, and basic residues in red.

For each protein, the first ARIA run was performed using only few unambiguous NOE contacts and known disulphide restraints. Disulphide bonds were confirmed using ^{13}C-NOESY to look for adjacent contacts between bonded Hβ nuclei (Figure 3.8). However several of these contacts were difficult to unambiguously assign due to weak signal or peak overlap (Appendix B). Therefore, cysteine connectivity is largely inferred from stronger contacts between one cysteine and the residues that neighbour the partnered cysteine. The presence of disulphide bonds is also indicated by the observed cysteine Cβ chemical shift values ranging from 50.06 to 38.87 ppm among SL1, WI1 and PC1 (Appendix B), where cysteine Cβ chemical shifts typically range from 30 – 24 ppm to 35 – 50 ppm for reduced and oxidized cysteine, respectively.[198] The formation of four disulphides is further supported by MS analysis (Appendix A), where the mass of eight protons (two for each disulphide bond) is consistently observed to be missing for each hydrophobin.

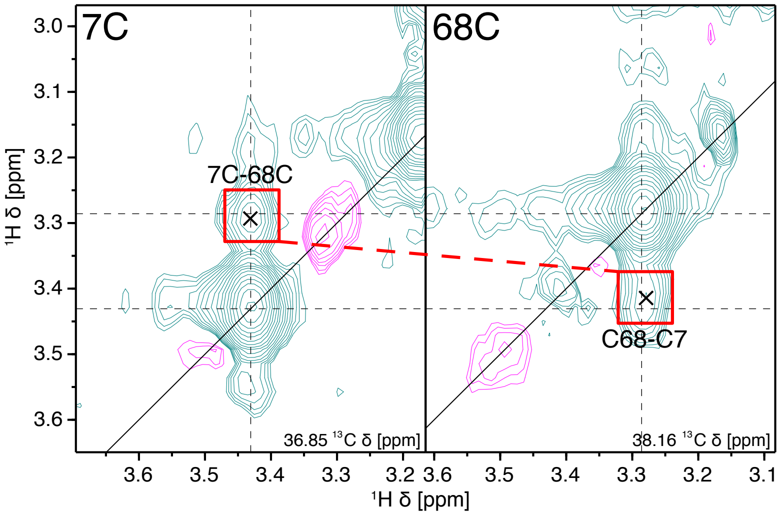


Figure 3.8. ^{13}C-edited NOESY-HSQC spectra confirm disulphide bonding of SL1. Disulphide bonding was confirmed ^{1}H-^{13}C NOESY experiments. Shown here is contact between 7C$^{H\beta}$-68C$^{H\beta}$ in SL1, where the marked cross peak in each strip of a 3D spectrum (36.85 ppm vs 38.16 ppm δ^{15}N) corresponds to a NOE contact between two Hβ nuclei of partnered cysteines.

As each structure converged, more restraints were manually added, including hydrogen bonds, additional assigned NOE contacts, and dihedral angle restraints. Protein hydrogen bonding was determined by resuspending the protein of interest in D_2O prior to a standard HSQC experiment. Deuterium is undetectable in a ^{1}H-^{15}N HSQC, therefore solvent exchange of backbone amide protons will result in signal loss. Comparison against a standard non-exchanged spectrum highlights which residues are unable to exchange with deuterium and, therefore, likely constrained in a hydrogen bond. ^{1}H-^{15}N HSQC experiments were collected for SL1, WI1, and PC1 after exchange with D_2O (Figure 3.9), and compared with ^{1}H-^{15}N HSQC experiments collected in H_2O. Although this approach only unambiguously determines the N-H pair involved in a hydrogen bond, the oxygen involved in the hydrogen bond became clear based on NOE contacts to neighbouring protons and through analysis of initial ensembles of structures. The number of restraints varied greatly, with 25, 1, and 16 pairs of hydrogen bonded nuclei (HN and C') being assigned in SL1, WI1, and PC1 respectively. The decreased number of hydrogen bond restraints observed for WI1 may be due to spectra quality (Section 3.3.1) or due to the increased dynamics of WI1 (Section 3.3.4).

NOE contacts and Chemical Shift Index (CSI)[199,200] suggest that SL1 and PC1 contain a helix in L_1, while WI1 does not. Canonical α-helical contacts (αN (i, i+3), α N (i,i+4) and αβ(i, i+3) are observed for residues 24-34 in SL1 and 22-32 in PC1, summarized in Figure 3.10. In combination with hydrogen bonding data, CSI, and DANGLE predictions, an α-helix can be confidently assigned to this region. In contrast, the same degree of NOE connectivity or hydrogen bonding is not observed within the L_1 region of WI1.

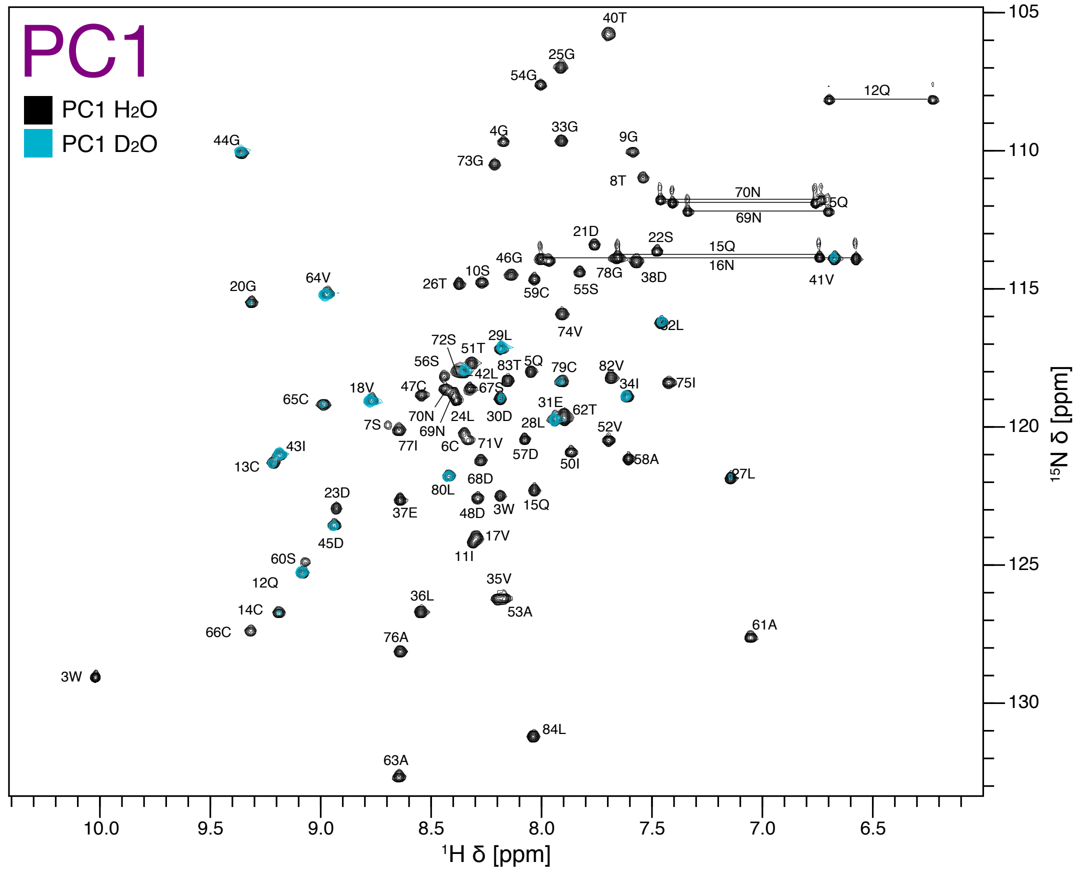


Figure 3.9. ^{1}H-^{15}N HSQC experiments collected after exchange of solvent for D_2O indicates hydrophobins have many residues protected from solvent exchange. ^{1}H-^{15}N HSQC spectra of PC1 were collected in 90% H_2O/10% D_2O (black), and 100% D_2O (teal) but otherwise as previously described. Peaks remaining in the D_2O data are protected from solvent exchange.

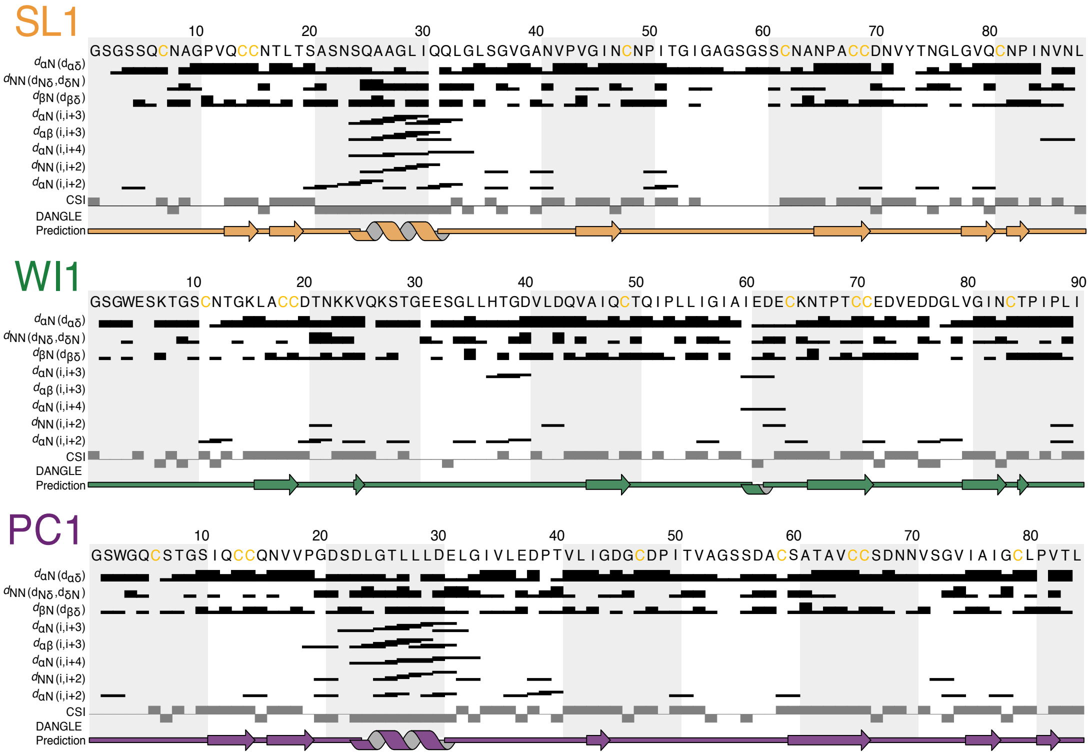


Figure 3.10. Summary of observed NOE contacts and CSI for SL1, WI1 and PC1. NOE contacts observed between the pairs of protons identified in subscripts and in brackets (between i and i+1 unless otherwise noted) indicated by horizontal bars. The thicker the horizontal line, the closer the two resonances are to each other. Chemical shift index (CSI)[199,200]and secondary structure predictions using the DANGLE algorithm[1] are shown. This figure was generated using CCPNMR Analysis[190].

3.3.3 The Dynamics of Inter-Cysteine Loop Regions is Not Conserved in Class IB Hydrophobins

To characterize dynamics on the ps-ns time-scale, I collected and analysed ^{1}H-^{15}N heteronuclear NOE spectra for SL1, SI1, and PC1, and compared this to previously published measurements for SC16[42] (Figure 3.11 and Figure 3.12) The majority of residues within class IB hydrophobins are ordered on the ps-ns timescale as indicated by a NOE > 0.7, excluding flexible N- and C-terminal residues. SL1, and to a lesser degree PC1, are similar to SC16 and have increased dynamics in the L_2 region. The flexibility of L_2 is consistent with the relative decrease in inter-residue contacts observed (Figure 3.13) and lack of a well-defined conformation is this region (Figure 3.14). In contrast to other class IB members, WI1 has a nearly opposite pattern of dynamics for its inter-cysteine loop regions (Figure 3.12): L_1 is very dynamic, with heteronuclear NOE values <0.5 between residues 25V and 45V, while L_2 is highly ordered. This is consistent with the lack of observed secondary structure features in L_1 (Figure 3.16). This increased flexibility and associated lower structural convergence in L_1 may be the result of the increased steric bulk or charge repulsion generated by the seven charged residues present in this region of WI1. The effect of losing helical structure in L_1 or additional charged residues in WI1 remains unclear, as the contribution of either property to rodlet formation is unknown.

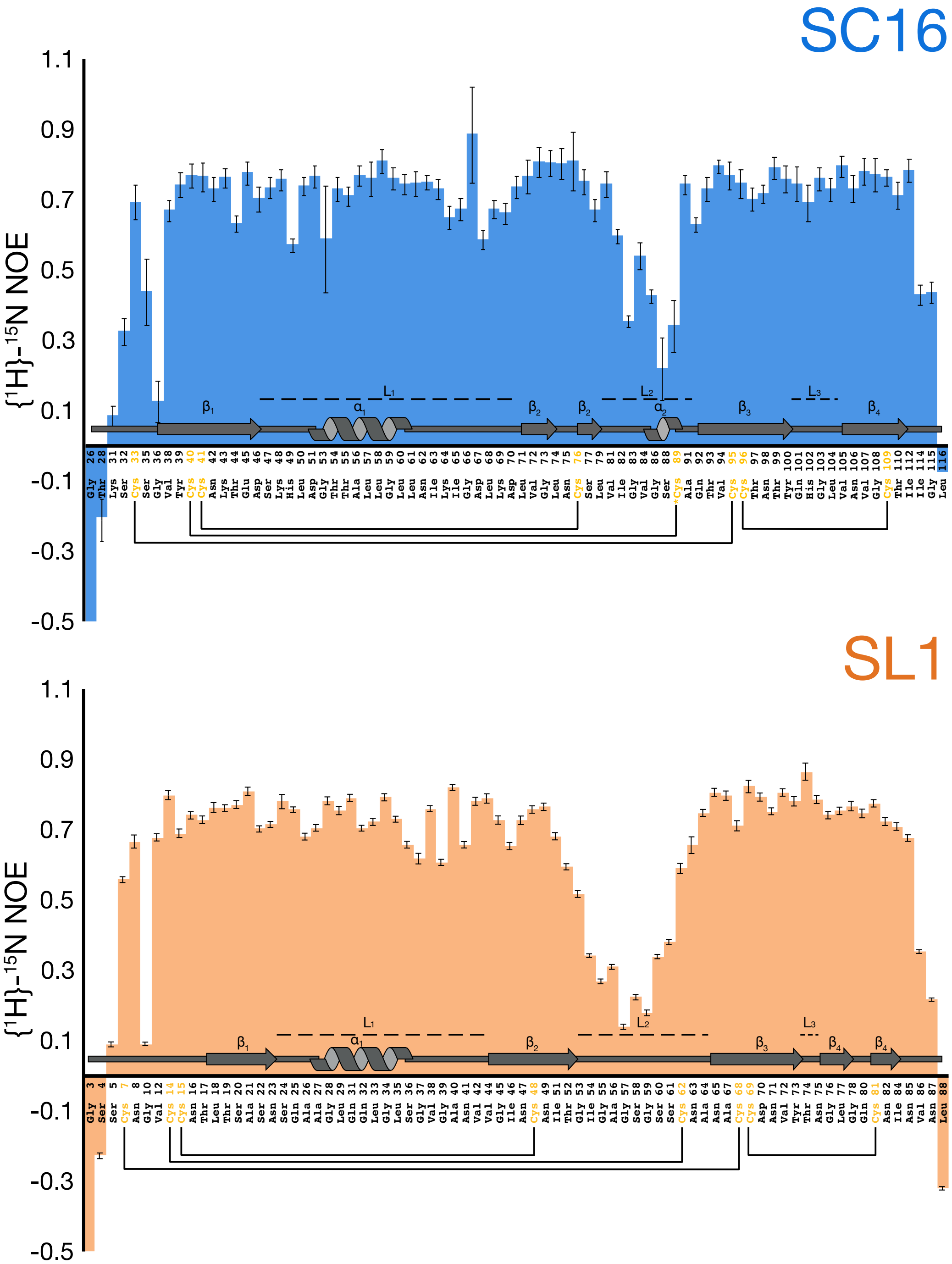


Figure 3.11. Per residue dynamics of SC16 and SL1.
{^{1}H}-^{15}N heteronuclear NOE measurements of ^{15}N-labelled SC16 and SL1. Secondary structure elements and disulphide bridges are annotated below . The majority of both constructs are ordered on the ps-ns timescale as indicated by a NOE > 0.7, excluding flexible N- and C-terminal residues. SL1 shows flexible L$_2$ domain similar to SC16 as reported by Gandier et al.[42]

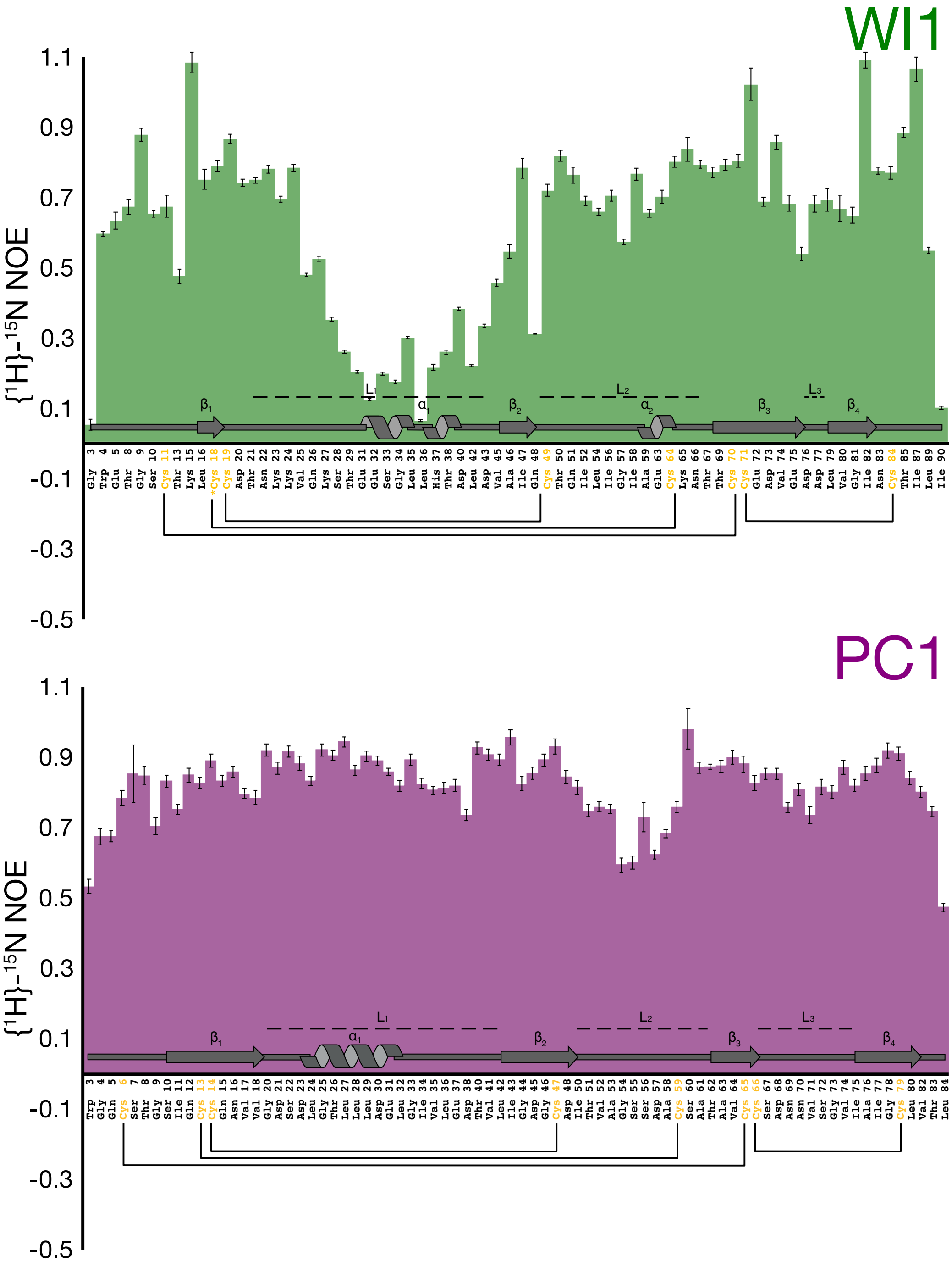


Figure 3.12. Per residue dynamics of WI1 and PC1.
{^{1}H}-^{15}N heteronuclear NOE of ^{15}N-labelled WI1 and PC1. Secondary structure elements and disulphide bridges are annotated. The majority of both constructs are ordered on the ps-ns timescale as indicated by a NOE > 0.7, excluding flexible N- and C-terminal residues.

The difference in dynamics of WI1 compared to SC16, SL1, and PC1 may be explained by the origin of WI1. *Wallemia ichthyophaga* requires 1.5 M salt to survive and readily grows suspended in the dead sea with salt concentrations reaching as high as 4.8 M.[201] As a result, the genome of *W. ichthyophaga* is greatly enriched in genes coding for hydrophobins, despite having one of the most reduced genomes of all sequenced basidiomycetes (9.6 Mbp, only 4884 predicted proteins).[129] These hydrophobins contain an unusually high proportion of acidic amino acids, likely as an adaptation to remain functional at such high concentrations of salt.[202] Acidic amino acids on a protein surface maintain soluble and active protein conformations in an environment that is generally detrimental to other proteins[203]

The flexibility of the peripheral loop regions in hydrophobins likely plays an important role in hydrophobin self-assembly. It was proposed that these intrinsically disordered regions play a role in recruitment of EAS to air-water interfaces but were not critical for rodlet self-assembly.[58] However, the exact role dynamic regions play in rodlet self-assembly remains unresolved and is complicated by the high variability of L_1, L_2, and L_3 in sequence, length, structure, and dynamics amongst hydrophobins. The differing dynamics between SL1 and WI1 further suggest that loop dynamics are highly variable between hydrophobins, even of the same class. It is likely that the flexibility of these regions is involved in tailoring individual hydrophobin characteristics rather than governing general hydrophobin function, although this remains to be rigorously established.

3.3.4 Shared Structural Features of Class IB Hydrophobins

^{15}N-edited and ^{13}C-edited NOESY spectra were assigned, resulting in 1583, 1294, and 1391 NOE-derived distance restraints for SL1, WI1, and PC1, respectively. However, the bulk of cross-peaks observed were typically between residues that neighbour each other in sequence, as they are also close in space. In particular, through-space connections are generally seen from the H^N of residue I to the H^α and H^β nuclei of the preceding i-1 residue. In contrast to this, when analysing NOESY spectra and generating restraints, long distance contacts are especially valuable for determining the protein structure. SL1 had the highest number of long range restraints (442, 5.02 per residue), while WI1 and PC1 were both considerably lower with 270 (3.00 per residue) and 288 (3.43 per residue), respectively. Figure 3.13 shows the number and type of NOE contacts assigned per residue in SL1, WI1 and PC1, with a noticeable decrease in the regions which were observed to be flexible by HetNOE (Figure 3.11 and Figure 3.12).

These combined restraints were used by ARIA to generate a final ensemble of 20 lowest energy structures for each class IB hydrophobin (Figure 3.14). The calculated structures of SL1, WI1, and PC1 have good overall statistics and converge with a root mean squared deviation of 0.3, 0.9 and 0.4 for backbone atoms, respectively (detailed statistics provided in Tables 3.2 – 3.4). No NOE violations over 0.5 Å were observed, and none of the hydrophobin ensembles have greater than 0.30 violations per structural ensemble between 0.2 – 0.5 Å. No dihedral angle violations were present for regions that were well-ordered (as defined by PSVS, Tables 3.2-3.4) Procheck-NMR was used to tabulate residues in favoured, generously allowed and disallowed areas of the Ramachandran plot. 94.6%, 86.1%, and 91.1% of residues for SL1, WI1, and PC1, were in the most favoured area of the Ramachandran plot respectively, while the rest were allowed. Finally, RPF analysis was used to calculate discriminating power (DP) scores that estimates how well each structure satisfies the data relative to a statistical random-coil structure. For SL1, WI1, and PC1, analysis indicates that overall fit between data and calculated structures is good with DP-scores scores of 0.865, 0.741, and 0.678 respectively.[192]

The structures calculated for SL1, WI1, and PC1 illustrate that a shared set of structural features are present among class IB hydrophobins when considered alongside the previously determined structure of SC16 (Figure 3.15).[42] Key features of these structures include a core β-sheet region, a β-hairpin in L_3, and an ordered, compact globular structure, comprising four β-strands and α-helical elements. The eight cysteines ubiquitous to all hydrophobins form disulphide bonds in a pattern consistent with other hydrophobins. Two disulphides located within the core β-barrel covalently link β_1 to β_2 and β_3 to β_4, while the two remaining disulphides connect the N-terminal tail to β_3 and β_1 to L_2 (Figure 3.15 & Figure 3.16). Within the L_1 region, SC16, SL1, and PC1 all contain an α-helix, while WI1 appears largely disordered with short helical motifs. β_3 and β_4 form a second antiparallel β-sheet connected by a β-hairpin motif within L_3. Between β_2 and β_3, L_2 forms an extended disordered loop.

Fewer restraints in the L_2 region (Figure 3.13) result in the lack of a well converged ensemble and coincides with the flexibility of this region in SC16, SL1, and PC1, as seen by ^{1}H-^{15}N HetNOE (Figure 3.11 and Figure 3.12). In comparison, WI1 appears less well converged, with the L_1 being particularly variable, reflected by the lower number of NOE contacts observed for this region and a decrease in structure quality statistics. The intrinsic flexible structural features of WI1 are confirmed by HetNOE experiments to not be due simply to the lack of constraints corresponding to that area (Figure 3.12). However, the decreased convergence of the structural ensemble of WI1 (Figure 3.14) is connected to sample instability, which negatively impacted the quality of data obtained.

Prior to this work, SC16 was the only class IB hydrophobin that has been structurally characterized. Overall, the structural features of SL1, PC1 and WI1 are consistent with those observed for SC16. SL1, PC1 and WI1 lack the long disordered N-terminal tail observed in SC16 (Figure 3.14), but otherwise are generally reminiscent of Class II hydrophobins. Both groups are compact and comprise a β-barrel with an associated α-helix (Figure 1.4, Figure 3.15 & Figure 3.16), but the location of the α-helix differs between the two classes of hydrophobin. For class IB it is located within L_1 while for the Class II hydrophobins it resides in L_2, where it is covalently linked to the core β-sheet region. As with SC16, the lack of a disulphide bond restraining the L_1 region permits more conformational flexibility in class IB hydrophobins and represents a possible structural basis for rodlet formation.

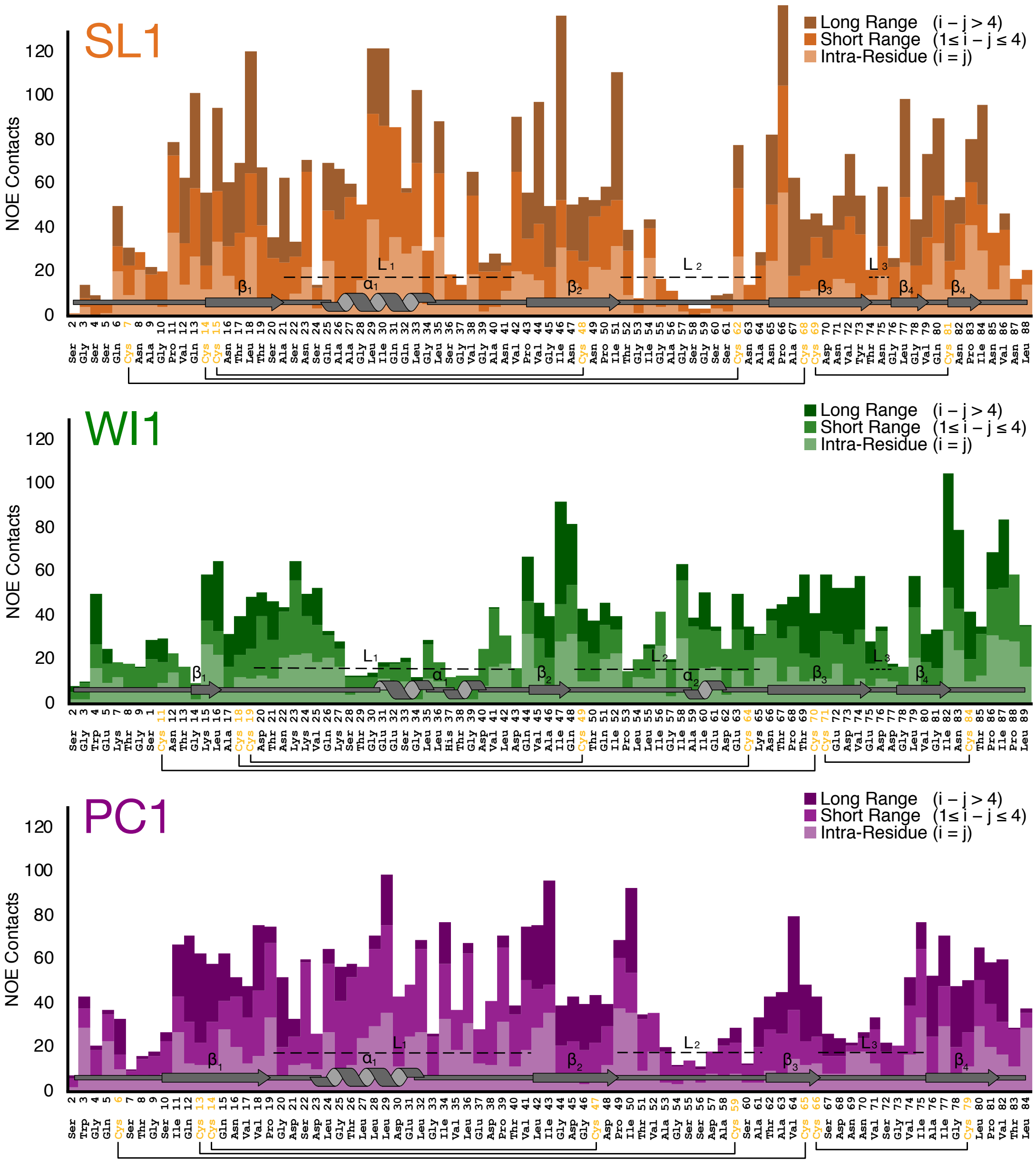


Figure 3.13. NOE contacts per residue of SL1, WI1 and PC1.

Assigned per residue NOE contacts used to calculate the final ensembles for SL1, WI1 and PC1. Inter-residue, short (including sequential) and long range NOE contacts observed in both ^{15}N-NOESY and ^{13}C-NOESY spectra are shown. Note that these plots include the ambiguous NOE contacts which may have been rejected by CNS during structure calculation. Regions which are less conformationally constrained, L_2 in SL1/PC1 and L_1 in WI1 have fewer contacts compared to other regions.

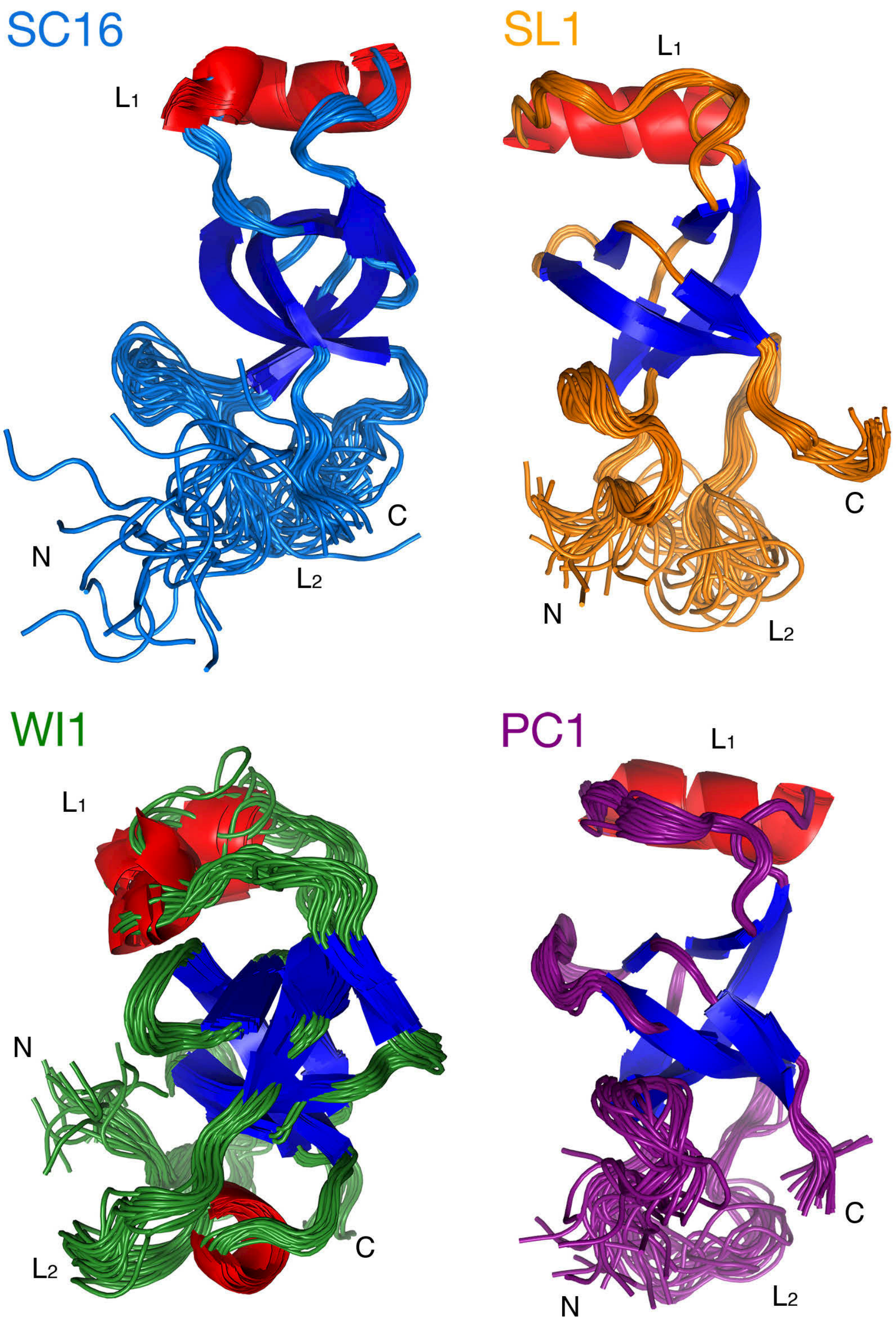


Figure 3.14. Ensembles of calculated hydrophobin structures.
Ribbon representation of the final ensemble of 20 lowest energy structures for SC16 (from Gandier et al.,[42] PDB ID: 2NBH), SL1, PC1, and WI1. Each of the hydrophobins superpose well over the central β-sheet, while only SC16, SL1, and PC1 have a well-converged L_1. Highlighted in red and blue are α-helix β-sheet features respectively. Labels denote loop regions of each hydrophobin.

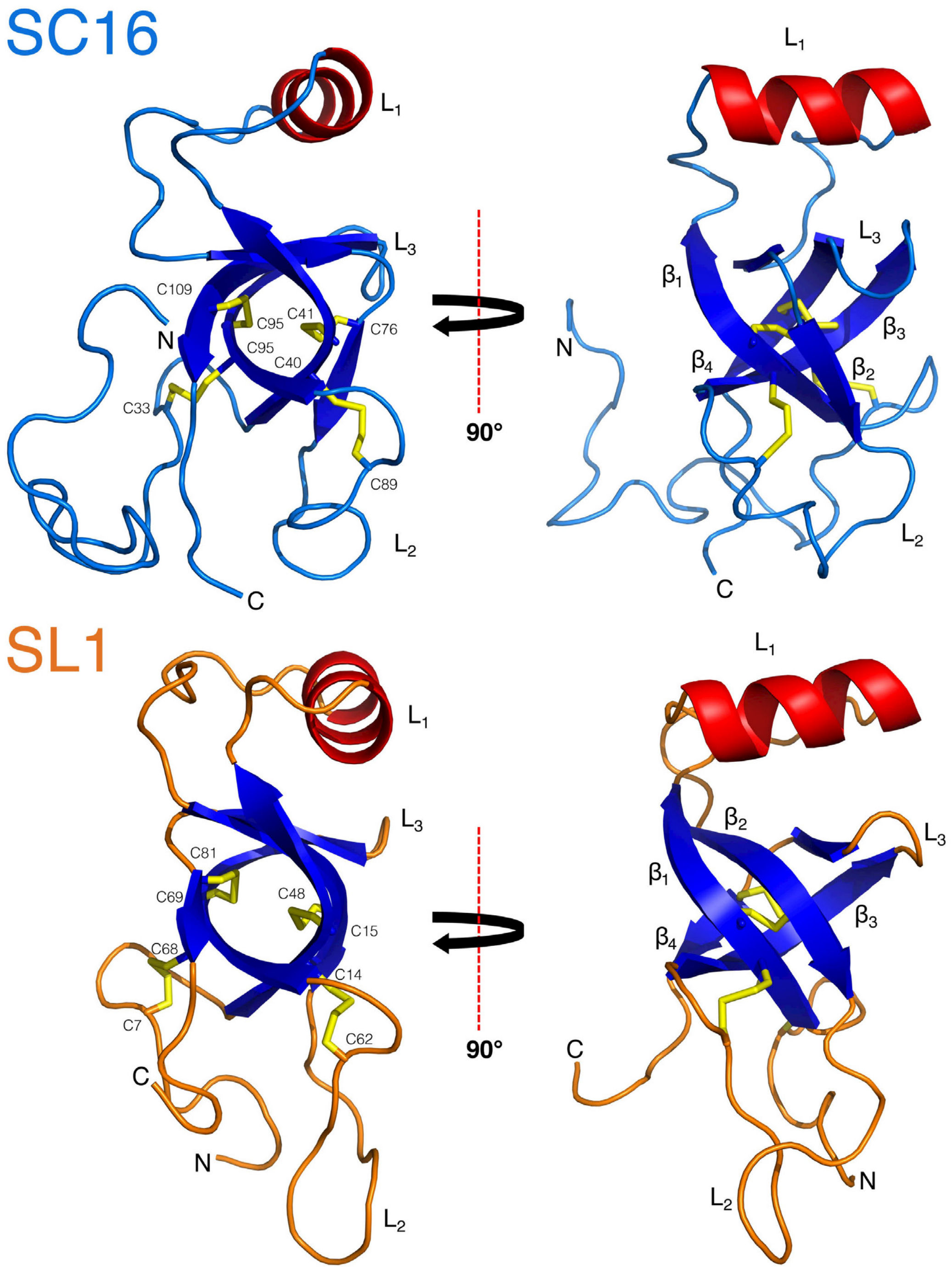


Figure 3.15. Cysteine bonding pattern and structural features of SC16 and SL1. A ribbon representation of the single member of the final calculated ensemble of lowest-energy structures of SC16 and SL1. Highlighted in red and blue are α-helix and β-sheet features respectively. Labels denotes loop regions of each hydrophobin. Two sets of two β-sheets (β1- β2 and β3- β4) are linked together by disulphide bonds while the two other disulphide bonds are responsible for linking peripheral loop regions to the core.

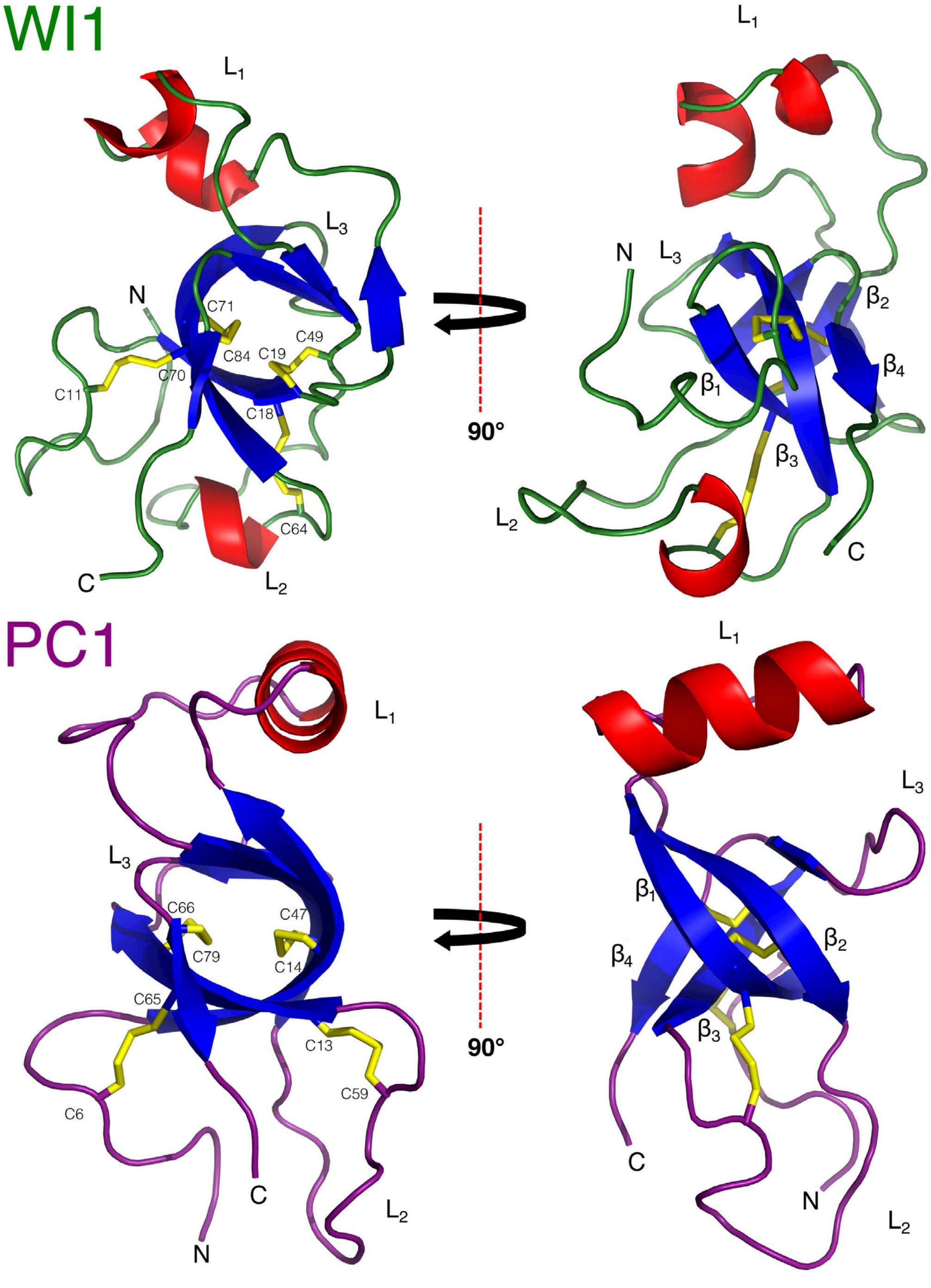


Figure 3.16. Cysteine bonding pattern and structural features of WI1 and PC1. A ribbon representation of the single member of the final calculated ensemble of lowest-energy structures of WI1 and PC1. Highlighted in red and blue are α-helix and β-sheet features respectively. Labels denotes loop regions of each hydrophobin. Two sets of two β-sheets (β1- β2 and β3- β4) are linked together by disulphide bonds while the two other disulphide bonds are responsible for linking peripheral loop regions to the core.

Table 3.2. Protein information and structure calculation statistics for SL1

Protein	**SL1**
Organism	*Serpula lacrymans*
ID (NCBI)	578457
Total No. Residues in Construct	88
Neg. Charged	1
Pos. Charged	0
Completeness of resonance assignments	
Backbone (%)	96.7
Side Chain (%)	96.2
Aromatic (%)	100
Conformationally-restricting restraints	
Distance Restraints	
Total	1583
intra-residue ($i = j$)	609
sequential ($\lvert i-j \rvert = 1$)	320
medium range ($1 < \lvert i - j \rvert < 5$)	161
long range ($\lvert i - j \rvert \geq 5$)	442
ambiguous	51
Dihedral Angle Restraints	124
Hydrogen Bond Restraints	50
No. of Restraints per Residue	17.99
No. of Long Range Restraint per Residue	5.02
Residual Restraint Violations:	
Average no. of distance violations per structure:	
0.1 – 0.2 Å	51.25
0.2 – 0.5 Å	0.25
> 0.5 Å	0
Average no. of dihedral angle violations per structure	
> 5°	0
Model Quality:	
RMSD backbone atoms (Å)	0.3
RMSD heavy atoms (Å)	0.6
RMSD bond lengths (Å)	0.006
RMSD bond angles (°)	0.8
Procheck Ramachandran statistics:	
most favored regions (%)	94.6
allowed regions (%)	5.4
disallowed regions (%)	0
Global quality scores (Raw / Z-score)	
Verify3D	0.44 / -0.32
ProsaII	0.60 / -0.21
ProCheck (phi-psi)	-0.49 / -1.61
ProCheck (all)	-0.51 / -3.02
MolProbity clash score	42.47 / -5.76
Ordered residue ranges	8-52,63-87
Total no. of residues in range	70
BMRB Accession Number:	30301
PDB ID:	5W0Y

Table 3.3. Protein information and structure calculation statistics for WI1

Protein	**WI1**
Organism	*Wallemia ichthyophaga*
ID (NCBI)	536056
Total No. Residues in Construct	90
Neg. Charged	14
Pos. Charged	6
Completeness of resonance assignments	
Backbone (%)	95.5
Side Chain (%)	84.6
Aromatic (%)	80.0
Conformationally-restricting restraints	
Distance Restraints	
Total	1294
intra-residue (i = j)	650
sequential (|i−j| = 1)	303
medium range (1 < |i − j| < 5)	71
long range (|i − j| ≥ 5)	270
ambiguous	32
Dihedral Angle Restraints	162
Hydrogen Bond Restraints	2
No. of Restraints per Residue	14.37
No. of Long Range Restraint per Residue	3.00
Residual Restraint Violations:	
Average no. of distance violations per structure:	
0.1 – 0.2 Å	38.75
0.2 – 0.5 Å	0.30
> 0.5 Å	0
Average no. of dihedral angle violations per structure > 5°	0.3
Model Quality:	
RMSD backbone atoms (Å)	0.9
RMSD heavy atoms (Å)	1.3
RMSD bond lengths (Å)	0.005
RMSD bond angles (°)	0.7
Procheck Ramachandran statistics:	
most favored regions (%)	86.1
allowed regions (%)	13.9
disallowed regions (%)	0
Global quality scores (Raw / Z-score)	
Verify3D	0.24 / -3.53
ProsaII	0.16 / -2.03
ProCheck (phi-psi)	-0.63 / -2.16
ProCheck (all)	-0.59 / -3.49
MolProbity clash score	29.05 / -3.46
Ordered residue ranges	4-9,11-24, 45-73, 81-87
Total no. of residues in range	56
BMRB Accession Number:	30505
PDB ID:	6E9M

Table 3.4. Protein information and structure calculation statistics for PC1

Protein	**PC1**
Organism	*Phanerochaete carnosa*
ID (NCBI)	650164
Total No. Residues in Construct	84
Neg. Charged	10
Pos. Charged	0
Completeness of resonance assignments	
Backbone (%)	98.5
Side Chain (%)	91.0
Aromatic (%)	100
Conformationally-restricting restraints	
Distance Restraints	
Total	1391
intra-residue (i = j)	650
sequential (|i–j| = 1)	332
medium range (1 < |i − j| < 5)	121
long range (|i − j| ≥ 5)	288
ambiguous	37
Dihedral Angle Restraints	134
Hydrogen Bond Restraints	34
No. of Restraints per Residue	16.55
No. of Long Range Restraint per Residue	3.43
Residual Restraint Violations:	
Average no. of distance violations per structure:	
0.1 – 0.2 Å	32.35
0.2 – 0.5 Å	0.30
> 0.5 Å	0
Average no. of dihedral angle violations per structure	
> 5°	0
Model Quality:	
RMSD backbone atoms (Å)	0.4
RMSD heavy atoms (Å)	0.7
RMSD bond lengths (Å)	0.005
RMSD bond angles (°)	0.7
Procheck Ramachandran statistics:	
most favored regions (%)	91.1
allowed regions (%)	8.9
disallowed regions (%)	0.0
Global quality scores (Raw / Z-score)	
Verify3D	0.37 / -1.44
ProsaII	0.52 / -0.54
ProCheck (phi-psi)	-0.50 / -1.65
ProCheck (all)	-0.61 / -3.61
MolProbity clash score	30.55 / -3.72
Ordered residue ranges	10-47,62-82
Total no. of residues in range	59
BMRB Accession Number:	30504
PDB ID:	6E98

3.3.5 Class IB Hydrophobins Lack Surface Features Observed in Other Hydrophobins

Interestingly, hydrophobic residues within class IB hydrophobins do not appear to be sequestered to the internal core of the β-barrel. Instead, many of these residues face outwards to form exposed hydrophobic patches or hydrophobic interfaces that may stabilize peripheral loop regions. Although they have a similar structure and surface hydrophobicity (Figure 3.17 & Figure 3.18), SC16 and the newly characterized SL1, PC1, and WI1 have different distribution of charge (Figure 3.19).

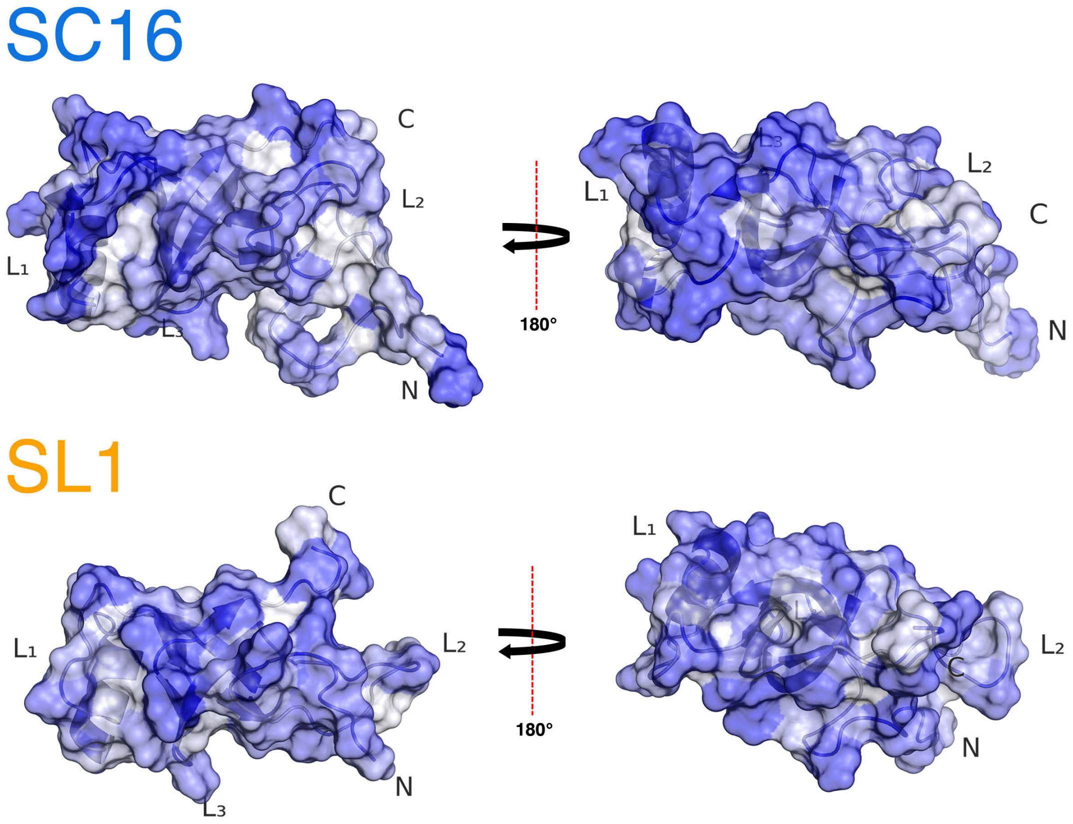


Figure 3.17. Surface hydrophobicity of SC16 and SL1.
A surface plot of SC16 and SL1 with underlying secondary structure features shown and with residues coloured blue according to hydrophilicity (hydrophobic residues are shown in grey).[204] SC16 and SL1 both lack a clear hydrophobic patch, with uniform hydrophobic residue distribution.

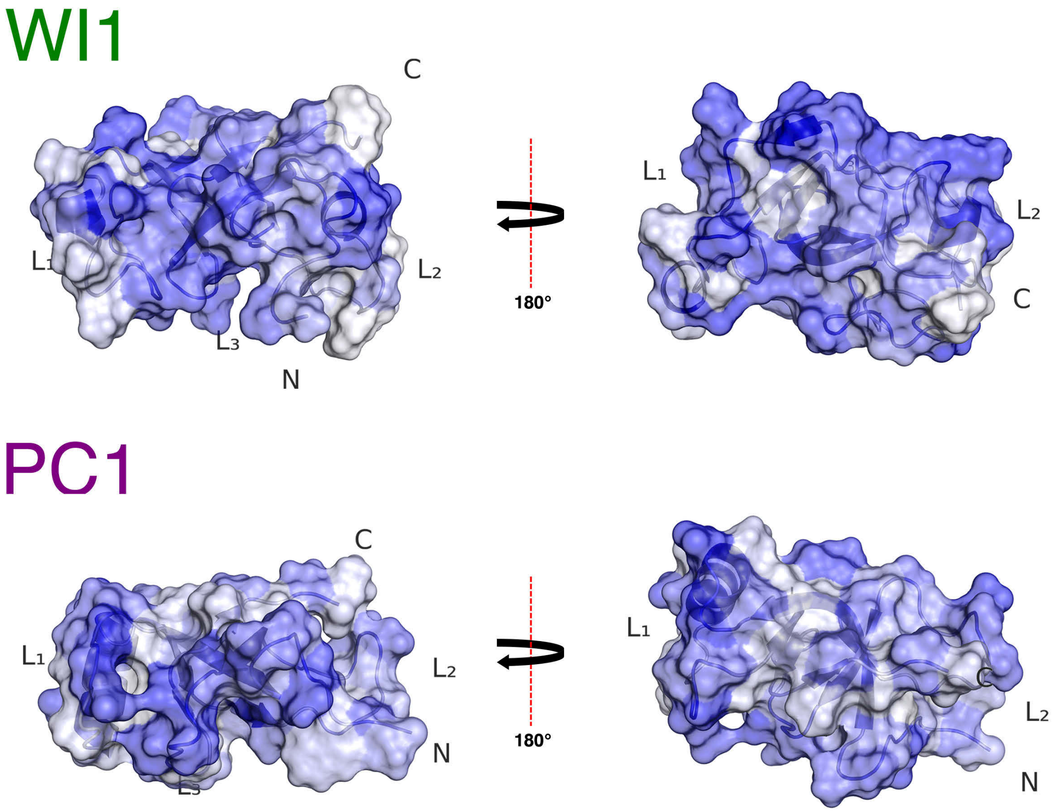


Figure 3.18. Surface hydrophobicity of WI1 and PC1.
A surface plot of WI1 and PC1 with underlying secondary structure features shown and with residues coloured blue according to hydrophilicity (hydrophobic residues are shown in grey).[204] Like SC16 and SL1, WI1 and PC1 also both lack a clear hydrophobic patch, with uniform hydrophobic residue distribution.

Analysis of SC16 revealed several acidic, basic, and uncharged patches on the surface of the protein.[42] In contrast, SL1, WI1 and PC1 appear to have a more uniform charge distribution and lack a distinct hydrophobic patch that is routinely observed in other hydrophobins. PC1 and WI1 appear significantly more charged, with a negatively charged patch in the L_1 region in Figure 3.18. In contrast, SC16 and SL1 have more sporadic surface charge distribution.

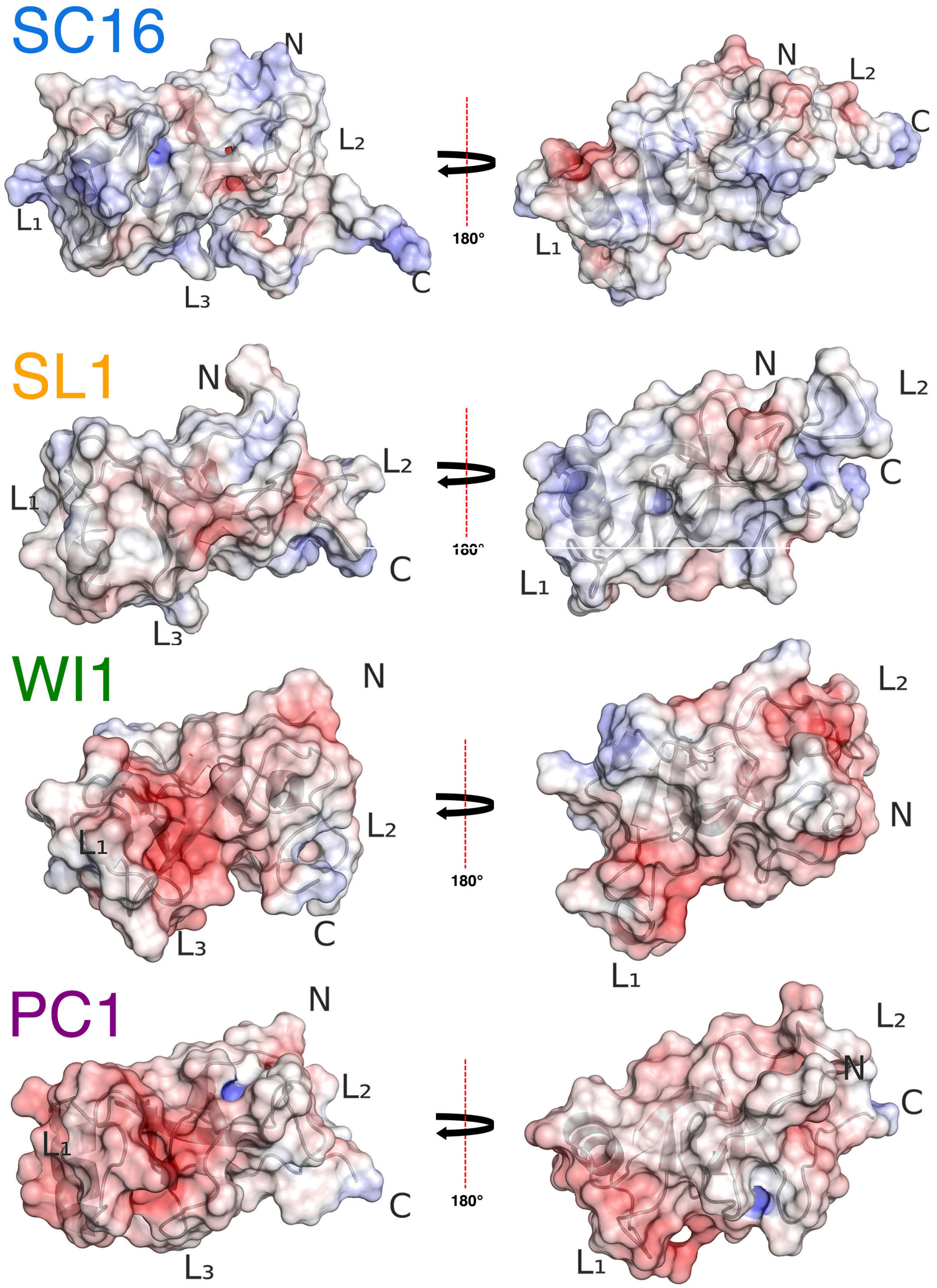


Figure 3.19. Surface electrostatics of class IB hydrophobins.
A surface plot of class IB hydrophobins with underlying secondary structure features shown with residues coloured red according to surface charge (red is negative, blue is positive) generated using the APBS electrostatics suite.[195]

SC16 appears to have distinct patches of basic residues, whereas SL1 appears to have very little surface charge, no doubt due to having only charged residue (Figure 2.2). The role that this charge localization plays in the functional properties of PC1 and WI1 versus SC16 and SL1 is unclear; however, the importance of hydrophobic patches on the recruitment of hydrophobins to interfaces and their role in rodlet assembly is established.[45,81]

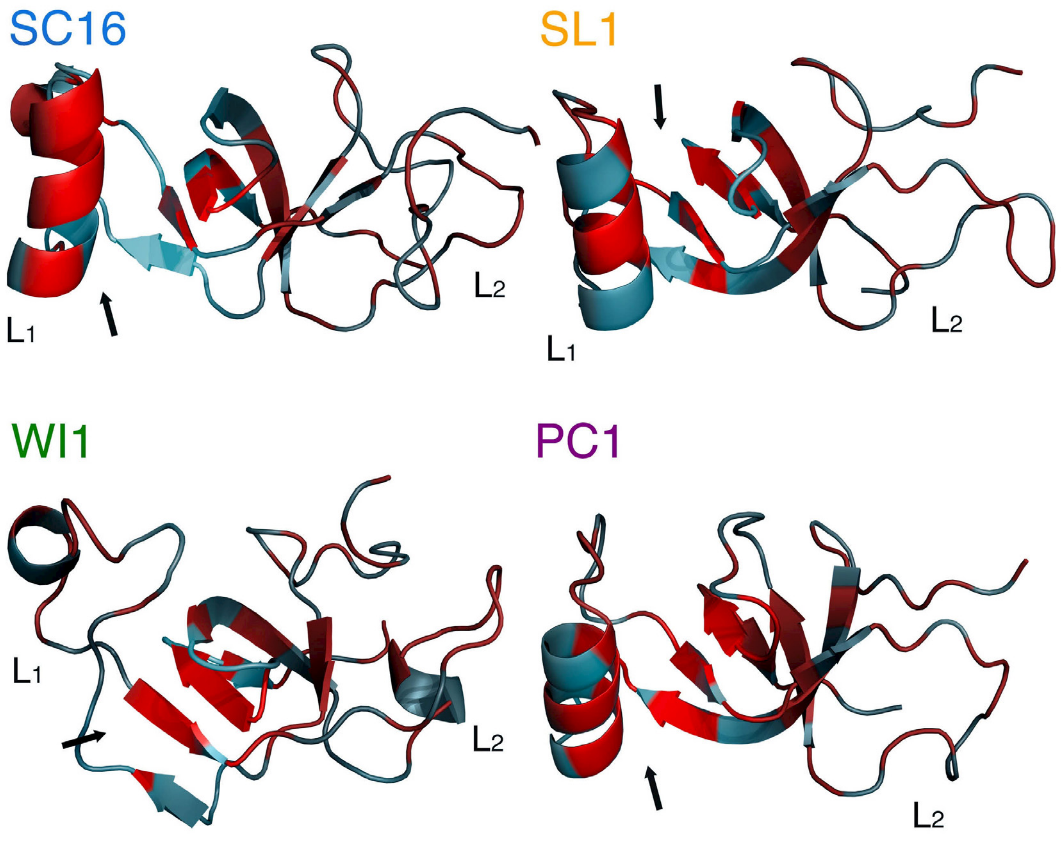


Figure 3.20. A hydrophobic interface between L₁ and core β-sheet region is present in all class IB hydrophobins.
Hydrophobic residues (Ile, Leu, Trp, Val, Ala, Gly) are shown highlighted in red on ribbon structures of SC16, SL1, WI1, and PC1. While these residues generally appear to be distributed throughout each structure, localization of these residues to one side of the α-helix and β-barrel is present. This forms a buried hydrophobic region, indicated by the arrow.

A buried hydrophobic patch between a hydrophobic face of the L_1 α-helix and the core β-barrel is present in all class IB hydrophobins. A similar concentration of hydrophobic residues is observed in PC1, and to a less effect, SL1 (Figure 3.20). In WI1 there appears to be a localization of hydrophobic residues on the exterior face of β-barrel oriented towards L_1, similar to the other class IB hydrophobins. However, as L_1 is structurally converged but dynamic in WI1, it is unlikely its hydrophobic patch is buried to the same extent. The flexibility of L_1 potentially allows for a more transiently buried hydrophobic face, where L_1 can easily unfold to expose the hydrophobic patch on the face of β_2 and β_4.

3.4 Conclusion

In this chapter I have used NMR spectroscopy to determine the atomic-resolution structures of SL1, WI1, and PC1, establishing that Class IB hydrophobins have a high degree of structural conservation. All of the structures are of high quality, with minimal NOE or dihedral angle violations, and low RMSD values. Using these structures, several shared key features are consistently observed between SC16, SL1, WI1, and PC1: a core region is formed by two sets of two β-sheets and linked together by disulphide bonds, a β-hairpin within the L_3 region is present, and a similar overall disulphide bonding pattern is observed between all class IB hydrophobins. Interestingly, several large differences in structural properties are noted as well. While all members are generally well structured, WI1 has ps-ns motions that are distinct from SC16, SL1, and PC1. The flexibility of the L_2 regions is only partially conserved and the role that peripheral loop dynamics play in Class IB functionality remains unclear.

Furthermore, the uniform distribution of surface charge and overall lack of hydrophobic patches on SL1, W11, and PC1 differs greatly from the previously characterized SC16 and other class I hydrophobins. However, a buried hydrophobic patch which could be exposed by rearrangement of L_1 was observed, providing a potential structural mechanism for rodlet assembly at hydrophobic-hydrophilic interfaces.

This work represents the first characterization of hydrophobins from disparate sources that show shared structural features. The structures of SC16, SL1, WI1, and PC1 will form the basis of experiments designed to probe the functional properties of class IB hydrophobins. In Chapter 4, I begin to characterize the functional properties of these hydrophobins and monitor their ability to self-assemble, in order to gain insight into the structural basis for rodlet self-assembly.

Chapter 4: Functional Characterization of Class IB Hydrophobins by Thioflavin T Fluorescence and Atomic Force Microscopy

4.1 Introduction

In this chapter, I discuss the morphology and functional properties of SC16, SL1, WI1, and PC1 assemblies. Using protein prepared as described in Chapter 2, the functional assays used in this chapter serve as a starting point for further studies that probe the structural changes that occur in response to changing environmental conditions. As SL1, PC1, and WI1 were until recently only predicted proteins based on sequence homology, the ability of each of these proteins to self-assemble was confirmed. I then used NMR spectroscopy-based titrations to associate features of hydrophobins with their function. The work presented here represents the final piece of my research into trying to understand how sequence, structure and function are correlated and contribute to rodlet self-assembly.

4.1.1 Visualizing Rodlet Self-Assembly

Detecting hydrophobins rodlets and films is typically done using either microscopy or fluorescence. Due to their nanometer scale size, direct observation of surface morphology typically uses either electron microscopy (EM)[31] or atomic force microscopy (AFM).[73] While both of these microscopy techniques have been extensively used, AFM studies allow for more detailed examination of rodlet structures with low nanometer resolution being possible.[42,70,205] Both techniques require distinct sample preparation prior to imaging. For EM, the spore surface or hydrophobins assembled onto a surface are sputter coated with silver or gold to increase conductivity and image quality.

Alternatively, for AFM, the sample must be deposited onto a substrate that acts as an atomically flat background, such as mica or highly oriented pyrolytic graphite (HOPG). In either case, rodlet assembly must be induced, then deposited in a uniform monolayer before drying to accurately determine surface features.

AFM is a type of scanning probe microscopy that can used to characterize sample morphology at the nanoscale or even the atomic-level.[206,207] AFM works by means of a reflective cantilever (probe) with a sharp and tip that is attached to a piezoelectric (PZT) actuator. This tip continually taps as it is then very carefully dragged across a prepared surface, deflecting whenever the tip encounters a topological feature (Figure 4.1A). When the cantilever is scanning a target sample, a continuous laser output is focused near its end which is reflected back to a position-sensitive photo detector allowing for measurement of cantilever deflection as a function of tip-sample interaction.[206,207]

While AFM can determine the surface features of rodlets, fluorescence assays using amyloid specific dyes, such as thioflavin T (ThT) or Congo red, have been extensively reported to detect rodlet formation.[42,73,84,208] For example, measuring ThT fluorescence at 485 nm was used to determine that EAS rodlet formation could be induced by vigorous vortexing, a process which maximizes air–water interface within the sample.[84] Both ThT and Congo red involve the intercalation of a small molecule fluorophore into stacking β-strands that are present upon rodlet formation.[58] When ThT associates with β-sheet rich structures, it has enhanced fluorescence and a characteristic red shift of its emission spectrum,[209] which can be easily monitored using a microplate reader or fluorometer.[210]

4.2 Materials and Methods

4.2.1 Atomic Force Microscopy of Hydrophobins

Lyophilized SC3, SC16, SL1, WI1 and PC1 were resuspended in MilliQ water at 80 μg/mL. This solution was used to prepare a 1 mL solution at 2 μg/mL in a 2 mL microcentrifuge tube. This solution was agitated gently overnight by end-over-end mixing to maximize contact between hydrophobins in solution and at the air-water interface. A 10 μL sample from the solution was then placed directly on freshly cleaved highly oriented pyrolytic graphite (HOPG; Mikromasch) and left on the bench to dry for 30 minutes before being placed in a desiccator overnight prior to imaging.

AFM imaging was performed in tapping mode at a scanning speed of 0.5-1 Hz with an Agilent 5500 (Agilent, Santa Barbara, CA, USA) ATM. High frequency (302-338 kHz) NTESP Antimony doped silicon cantilevers with a tip radius of 3.5–4.5μm (Bruker AFM Probes) were used. Images were captured at 512 × 512 pixel density a with cantilever drive frequency of between 305 and 325 kHz. AFM images were processed using Gwyddion,[211] which included the use of an unsharp filter with radius of 2 pixels, despeckle filter, and contact error correction.

4.2.2 Thioflavin T Assays

Lyophilized SC16, SL1, WI1, and PC1 were each resuspended in MilliQ water to prepare concentrated working stocks. The concentrations of these stocks was determined by UV/Vis spectroscopy and diluted with buffer (20 mM MES pH 5.5, 20 mM MES pH 6.5, 20 mM phosphate pH 7.5, or 20 mM Tris pH 8.5) to final concentrations of 80 μM. Stocks were also diluted into a buffer containing 20 mM MES, pH 6.5, with NaCl ranging from 50 mM to 5 M.

These solutions were left to incubate at ambient temperature for 16 hours either with or without gentle head-over-tail mixing at approximately 60 rotations per minute. After incubation, ThT was added to a concentration of 60 μM. Opaque 96-well plates were then prepared by loading 50 μL of each sample and 50 μL of the corresponding matched buffer containing 60 μM ThT. A serial dilution was then performed with matched buffer so that hydrophobin concentrations varied from 40 μM to 2.5 μM and ThT concentration remained constant at 60 μM. Emission at 485 nm was measured in triplicate with agitation between each read using a SpectraMax M3 with an excitation wavelength at 450 nm.

4.2.3 NMR Titrations

Two NMR samples of 200 μM ^{15}N-labeled WI1 were prepared, one in low salt buffer (50 mM) and the other in saturated salt buffer (5 M) with all other components being equal (20 mM MES pH 6.5, 1mM NaN_3, 10% D_2O). These samples were then sequentially mixed, resulting in NMR samples with varying amounts of NaCl. To mitigate the effect of high salt concentration in the sample on NMR pulse widths, a shaped NMR tube[212] was used. ^{1}H-^{15}N HSQC experiments were collected using a shaped NMR tube (Bruker Canada) on a 700 MHz Avance III spectrometer (BMRF facility, NRC, Halifax).

Four samples of ~100 μM ^{15}N-labeled PC1 were prepared as before (50 mM NaCl, 1 mM NaN_3, 10% D_2O) except each sample contained 20 mM of a different pH buffer: MES pH 5.5, MES pH 6.5, phosphate pH 7.5, or Tris pH 8.5. ^{1}H-^{15}N HSQC were collected on a 500 MHz Avance NMR spectrometer (NMR^3) as described in section

3.2.2. Chemical shift changes were analysed using CCPN Analysis, considering both ^{1}H and ^{15}N chemical shifts, using the following equation:

$$d = \sqrt{1.0\left(\delta\, {}^1H\right)^2 + 0.15\left(\delta\, {}^{15}N\right)^2}$$

Where d is the calculated chemical shift distance, δ^1H and $\delta\ ^{15}$N is the measured difference in chemical shift for ^{1}H and ^{15}N, respectively.[190]

4.3 Results and Discussion

4.3.1 Comparing Class IB Hydrophobin Assembly Morphology

SL1, WI1, and PC1 assembly morphology were all examined using AFM (Figure 4.1 B). In each case, assembled protein is visible on the surface; however, protein distribution is irregular and clustered into large groups. The protein also appears to be preferentially located along the ledges and boundaries naturally present on the HOPG surface. Furthermore, during imaging, the protein was not well adhered to the surface and could be dragged with the cantilever tip. If bonding between the HOPG and protein is weak, then it is possible that during deposition the protein was pulled along by the surface tension of the water as the drop evaporated, until the protein became physically lodged in a boundary. This indicates that the hydrophobins do not adhere to HOPG as well as expected.[42]

SC16 was included as a positive control, as it has been previously observed to form rodlets characteristic of class I hydrophobins.[42] The control sample of SC16 appears to have a somewhat rodlet-like appearance; however, the other class IB hydrophobins imaged formed either ill-defined clumps or adopted a film-like appearance in the case of PC1. The 3D view of the PC1 sample shown in Figure 4.1 C shows that the film is of relatively uniform height, and is reminiscent of the films formed by class II hydrophobins.[59,213]

Overall, these AFM results are inconclusive, with the required resolutions not being obtained during imaging. The quality of the AFM data obtained for SC16 does not match the results from Gandier et al.[42] When higher-resolution scans were attempted (0.25 μm x 0.25 μm), significant artefacts were observed, possibly due to the poor adhesion of hydrophobins to the HOPG substrate. In future experiments, sample preparation will be modified with the goal of forming a monolayer instead of discrete clumps. This may be achieved by using sonication or vortexing before sample deposition.

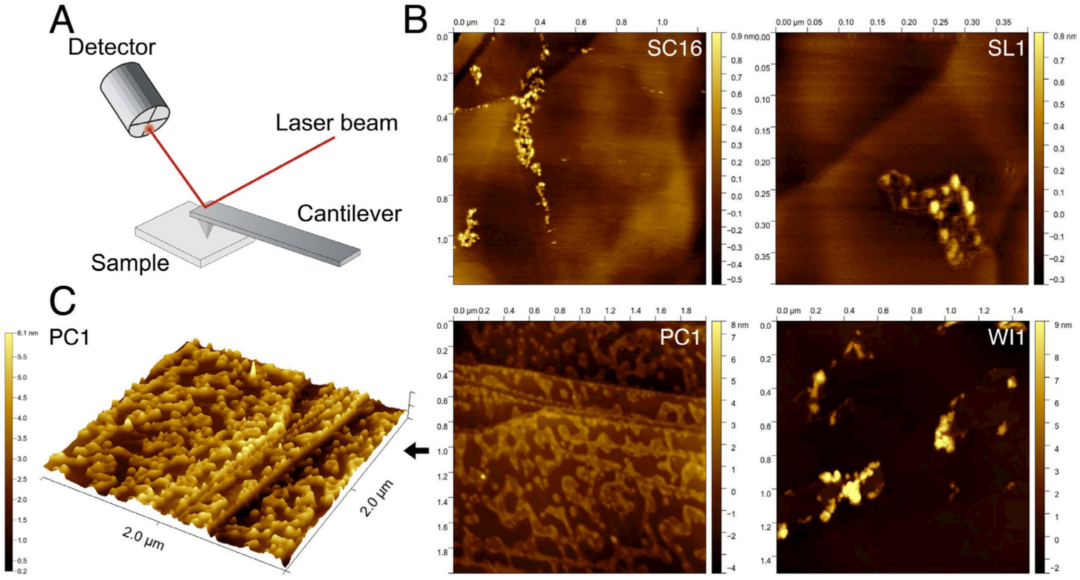


Figure 4.1. No clear rodlet morphology for SL1, WI1 and PC1 is observable by AFM.
(A) A Simplified schematic of the AFM experiment (B) 0.25 μm^2 to 4 μm^2 tapping mode height images of an HOPG surface coated with SC16, SL1, WI1, and PC1. (C) 3D view of the PC1 film shown in (B).

Additionally, the use of a desiccator might cause the sample drop to evaporate too quickly, pulling deposited protein into crevices on the surface of the HOPG. Future experiments could also investigate if mica substrates would result in better adhesion. Although their distinct morphology remains unresolved, these results show that SL1, PC1 and WI1 do self-assemble.

4.3.2 Hydrophobin Self Assembly is Dependent on Environment

The ability of Class IB hydrophobins to self-assemble was confirmed using ThT fluorescence assays. A variety of buffer conditions were examined to determine their effect on hydrophobin self-assembly. Figure 4.2 shows emission spectroscopy data obtained during a ThT assay of WI1 assembly. After excitation at 450 nm, an emission maximum is observed ranging from 470 nm to 485 nm, depending on protein concentration. These results indicate that WI1 is assembled after overnight gentle agitation as the unagitated sample shows much less fluorescence despite identical protein and ThT concentration.

It was observed that, after only gentle agitation, all samples turned opaque milky white (Figure 4.3). This change in appearance occurred relatively quickly in agitated samples, typically within 30 minutes, while it was never observed in unagitated samples. Even after centrifugation at 25 000 × *g* for 60 minutes, the supernatant of these samples remained opaque and induced ThT fluorescence, indicating the presence of assembled hydrophobins. This suggests that the opaque appearance of hydrophobin solutions upon agitation was not due solely to protein precipitation, but also a result of hydrophobin assembly.

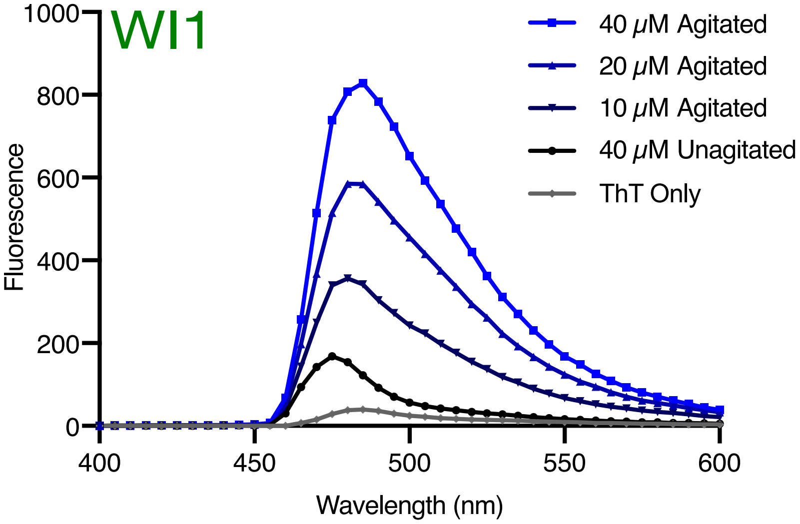


Figure 4.2. Increasing WI1 concentration results in increased ThT fluorescence. ThT emission scan from 400 nm to 600 nm, 5 nm steps, after excitation at 450 nm shown for different concentrations of agitated and unagitated WI1. Each sample was prepared in 20 mM MES pH 6.5 and 60 µM ThT.

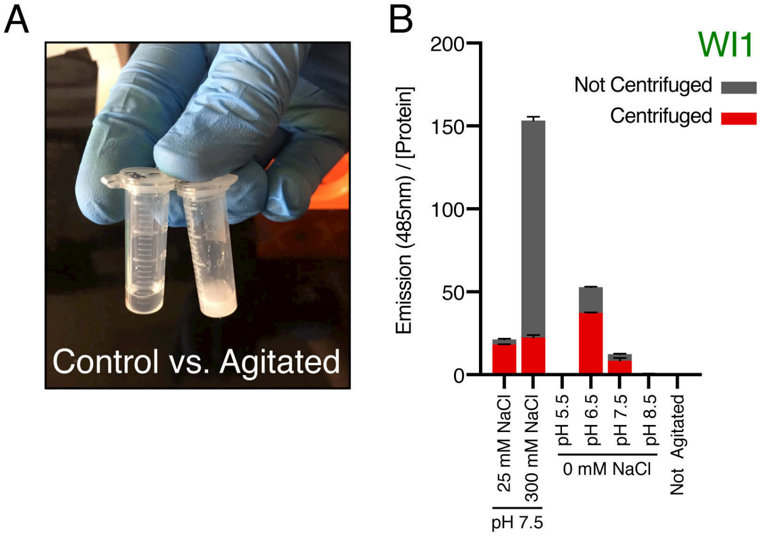


Figure 4.3. Agitation results in opaque suspension of assembled hydrophobins. (A) Gentle overnight agitation results in the formation of opaque suspension, suggestive of hydrophobin assembly. (B) Samples were centrifuged at 25,000 × *g* for 60 minutes resulting in a decrease in fluorescence, however sample supernatant remained cloudy and was still fluorescent during a ThT assay (N=3, error bars show standard deviation from mean). Non agitated sample was suspended in 25 mM NaCl, pH 7.5 buffer.

SC16, SL1, WI1 and PC1 were confirmed to have the ability to self-assemble as detected by a large increase in fluorescence after gentle agitation (Figure 4.4), and a linear dependence of enhanced ThT fluorescence with respect to protein concentration was observed. The propensity of each hydrophobin to self-assemble ranges, with SL1 having the largest increases in ThT fluorescence in most conditions tested. In order to determine if certain solution conditions are preferred for assembly, ThT assays were repeated using buffers with different pH and salt concentrations. The resulting data were processed (Appendix C) to obtain the average ThT fluorescence per unit of protein concentration (i.e., the slope of Figure 4.4), as summarized in Figure 4.5.

Interestingly, each hydrophobin has a unique preference for pH and salt conditions. SC16 appears to be the most salt tolerant, with optimal assembly occurring at pH 7.5. SL1 assembly increases with salt concentration, up to 1 M, but is unaffected by pH. Whereas PC1 is active in up to 4 M salt like SC16, but 1M appears optimal, and assembly is significantly reduced with increasing pH. Interestingly, WI1 appears not to show consistent trends, unlike the other class IB hydrophobins. Instead, WI1 appears highly active only at pH 6.5 and between 300 mM and 1 M salt. These trends likely have a basis in the sequence properties of each hydrophobin. Both SC16 and PC1 have a number of charged residues (12 and 20 respectively), which may allow them to function in changing salt conditions, but making them sensitive to pH. Furthermore, the differences between SC16 and PC1 assembly with increasing pH may be explained by the balance of charged residues within each protein. SC16 has an equal amount of acidic and basic residues (6 and 6), whereas PC1 has no basic residues at all. Therefore with increasing pH, PC1 becomes more charged which may ultimately hamper its ability to self-assemble. In contrast, SL1 has only a single charged residue and significantly less surface charge (Figure 3.19), making it active in a variety of pH conditions but sensitive to salt.

These results indicate that despite SC16, SL1, WI1, and PC1 having shared structural features and sequence similarities, each of these class IB hydrophobins has unique functional properties. This suggests that different hydrophobins were likely optimized to serve a specific purpose for the fungi and may explain why each individual fungal genus produces such a large diversity of hydrophobins. Between two and seven different hydrophobins per fungal species is frequently observed,[43] although in *Coprinopsis cinerea* there are 34 identified hydrophobins.[214]

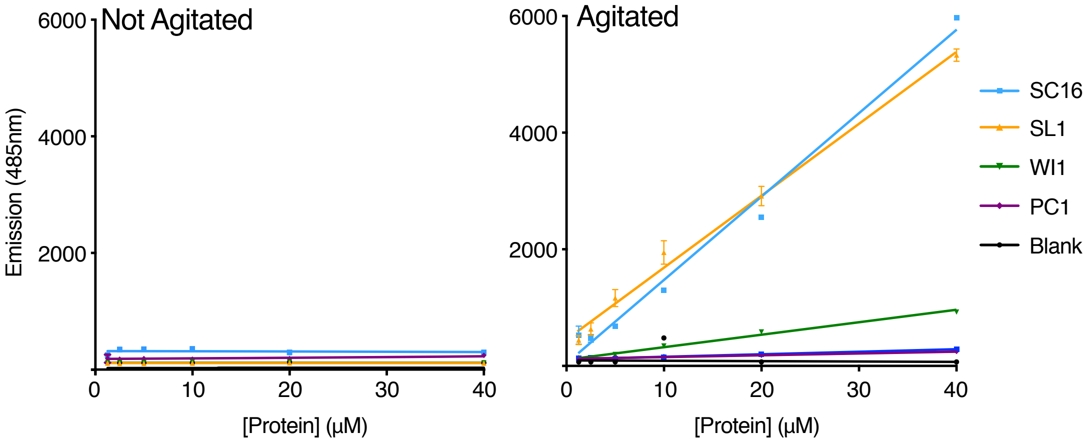


Figure 4.4. Agitation induces self-assembly in all hydrophobins tested.
Samples of SC16, SL1, WI1 and PC1 were left at room temperature overnight (Left) or agitated end over end at approximately 60 rpm overnight (Right) in 20 mM MES pH 6.5 and 50 mM NaCl. ThT emission was measured (N = 3) and plotted with standard deviation indicated.

The specificity of hydrophobins is also consistent with the variable expression of different hydrophobins[11] and the localization of unique hydrophobins to highly specialized organs within fungi.[12,24] For example, transcription of hydrophobin genes responds to changes in salinity in *Wallemia ichthyophaga.*[129]

Surprisingly, WI1, which is isolated from *W. ichthyophaga*, which thrives in high salt environments, does not form rodlets in high salt conditions. Instead WI1 shows low levels of assembly at low salt concentrations, moderate assembly at 300 mM and 1 M salt, and no assembly at 2 M NaCl and above.

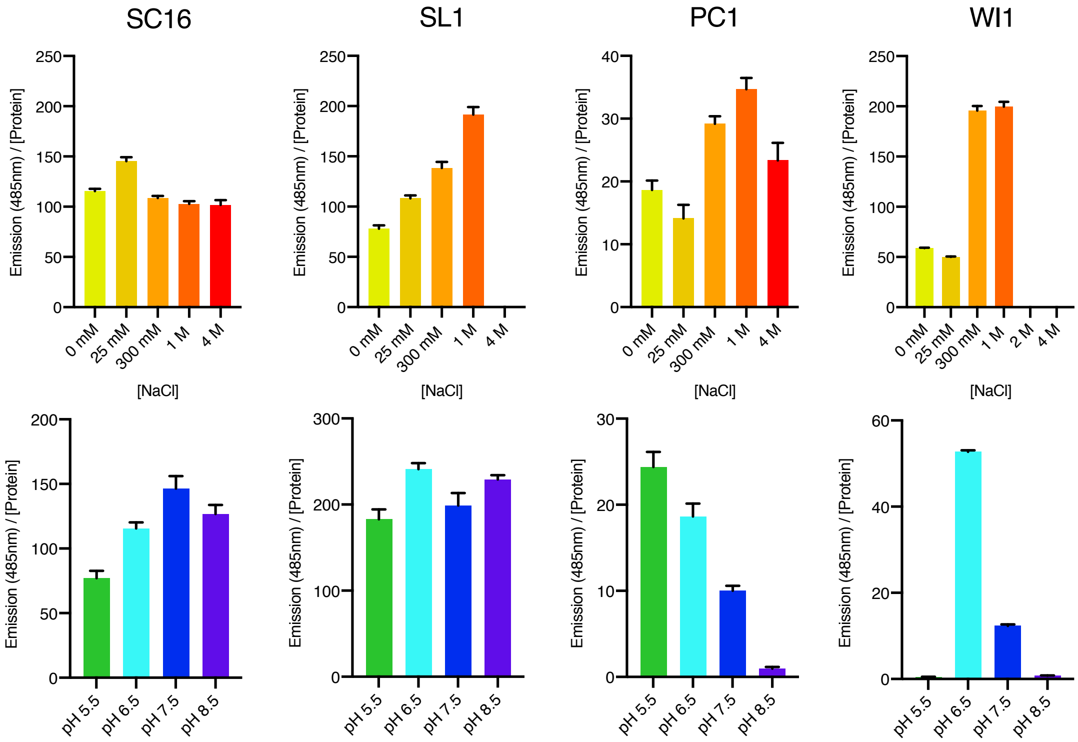


Figure 4.5. Hydrophobin self-assembly is influenced by pH and salt concentration. ThT assays were completed for Class IB hydrophobins while modifying buffer pH and salt concentration. The emission / protein ratio of the resulting hydrophobin samples is plotted. Calculated using linear regression of plots similar to the one shown in Figure 4.4. (Appendix C). For each condition error bars indicate standard deviation of measured values from the plotted linear fit.

It may be that WI1 is active in high salt conditions with a different composition, as the salt in the Dead Sea is not primarily NaCl (30.4%) but instead also contains $CaCl_2$ (14.4%), KCl (4.4%), and $MgCl_2$ (50.8%).[215] It has previously been shown that calcium ions can play a role in hydrophobin functionality[216] so it may be that different salt compositions will influence WI1 self-assembly. The narrow pH range which WI1 is functional is likely a result of its origin in the dead sea as well, where pH values range from 5.8 – 6.0.[217] It is surprising that WI1 appears completely inactive at pH 5.5, although the activity of WI1 at pH 5.5 in a high salt buffer remains to be determined.

4.3.3 Connecting Functional Properties to Structural Features

Since hydrophobin self-assembly is affected by buffer conditions, I examined if there is any coinciding structural change to the solution structure of hydrophobins that may explain this difference in functionality. Based on the results obtained from the ThT assays, I used NMR spectroscopy to monitor salt and pH titrations for WI1 and PC1, respectively (Figure 4.6 and Figure 4.7). As NMR pulse widths are sensitive to high salt concentration, a shaped NMR tube was used. Using this shaped tube, experiments did not suffer from long pulse widths typical of high salt NMR experiments, with proton pulse width remaining below 12.75 µs in 5 M NaCl relative to 5.1 us at 50 mM NaCl, allowing for the collection of high quality ^{1}H-^{15}N HSQC spectra.

By collecting a series of HSQC and comparing any differences in chemical shift values, changes within each hydrophobin structure can be localized. For WI1, increasing NaCl concentration had a global effect on peak dispersion and overall spectral quality. To align the WI1 spectra, each is centered on the C-terminal 90I residue, as this resulted in the highest degree of overlap for the entire spectrum. Many of the peaks move in a linear path with increasing salt concentration, indicating that these residues are exchanging between two states. Other peaks adopt a curved path, such as 37H or 23K as shown in the insert of Figure 4.6, possibly indicating that these residues are exchanging between three confirmations.[139,218]

With increasing salt concentration, many of the peaks for WI1 in Figure 4.6 begin to disappear. At high salt concentrations many of the peaks disappeared completely, with only 17 of 91 peaks being observed at 5 M NaCl. Peak disappearance could potentially be due to line-broadening as a result of exchange between monomer and dimer populations, where the missing peaks are associated with the dimer interface. This broadening results in peak disappearance at the lower sensitivity of spectra collected in high salt.

Increasing salt may also have an effect on the dynamics of WI1, with increasing dynamics resulting in the observed peak disappearance. To explore this hypothesis, I collected another ^{1}H-^{15}N HetNOE of WI1 at 1 M NaCl (Figure 4.8). In comparison to the ^{1}H-^{15}N HetNOE of WI1 at 50 mM NaCl (Figure 3.12), no significant differences in NOE values were observed between 50 mM and 1 M NaCl. This confirms that no pronounced change in protein dynamics is occurring with increasing salt concentrations, and therefore changes in dynamics would not explain the loss of WI1 rodlet forming ability at high NaCl concentrations.

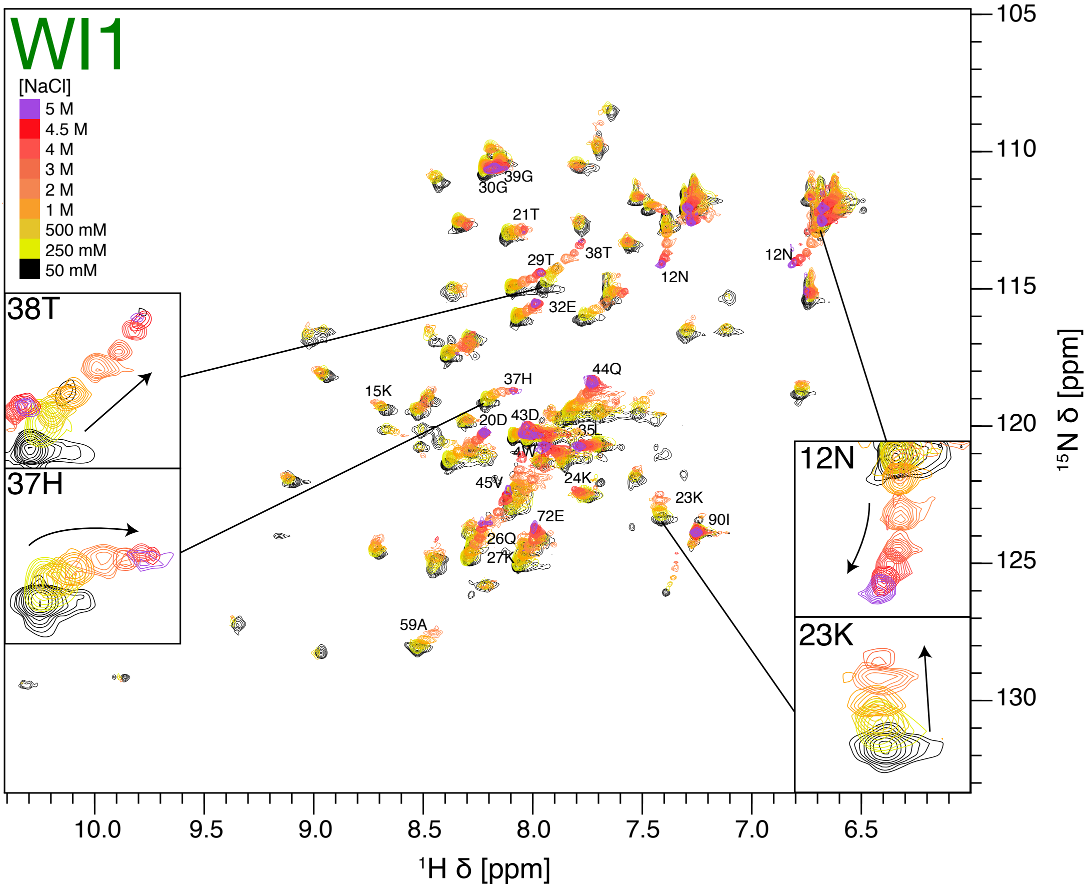


Figure 4.6. NMR-based salt titration of WI1.

Overlay of ^{1}H-^{15}N HSQC spectra collected of ^{15}N-labelled WI1 with varying NaCl concentration. Inserts in highlight the different types of peak responses observed. Peaks are labelled with previously determined assignment, with only residues that remain visible in 5 M NaCl being labeled for WI1.

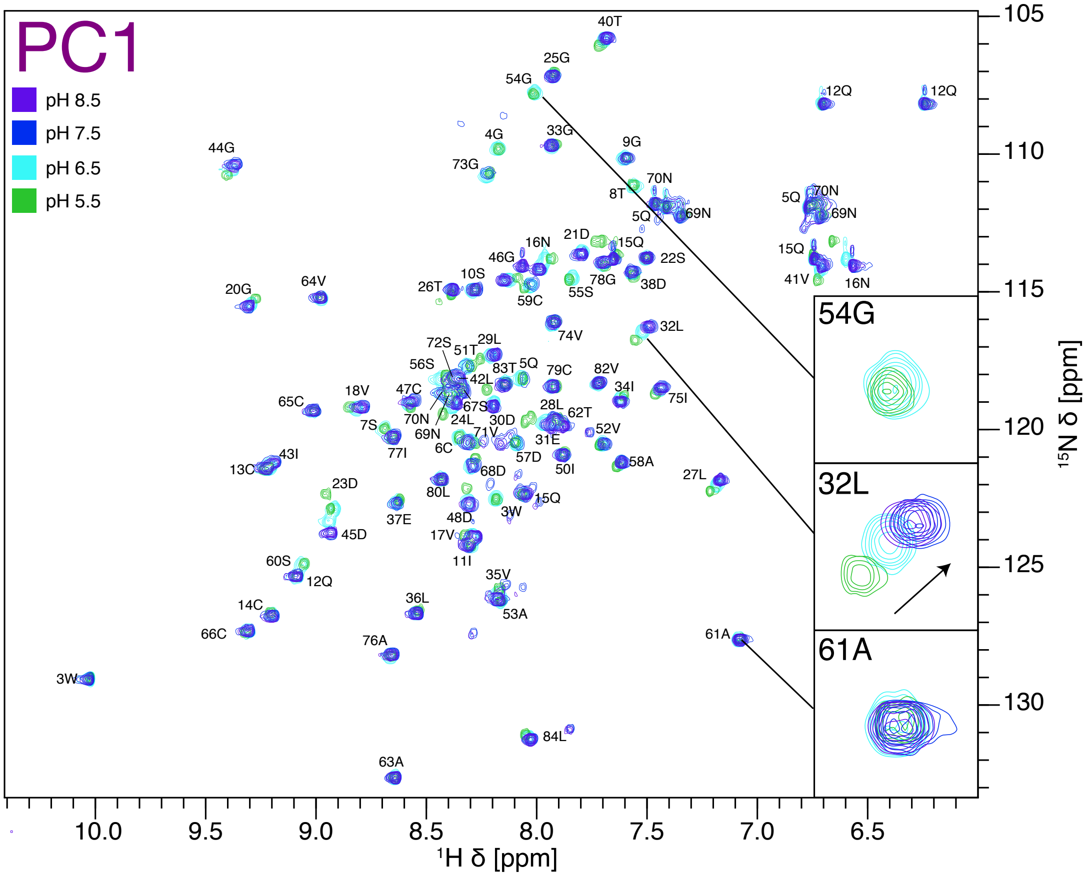


Figure 4.7. NMR-based pH titration of PC1.

Overlay of ^{1}H-^{15}N HSQC spectra collected of ^{15}N-labelled PC1 with varying pH. Inserts in highlight the different types of peak responses observed. Peaks are labelled with previously determined assignment.

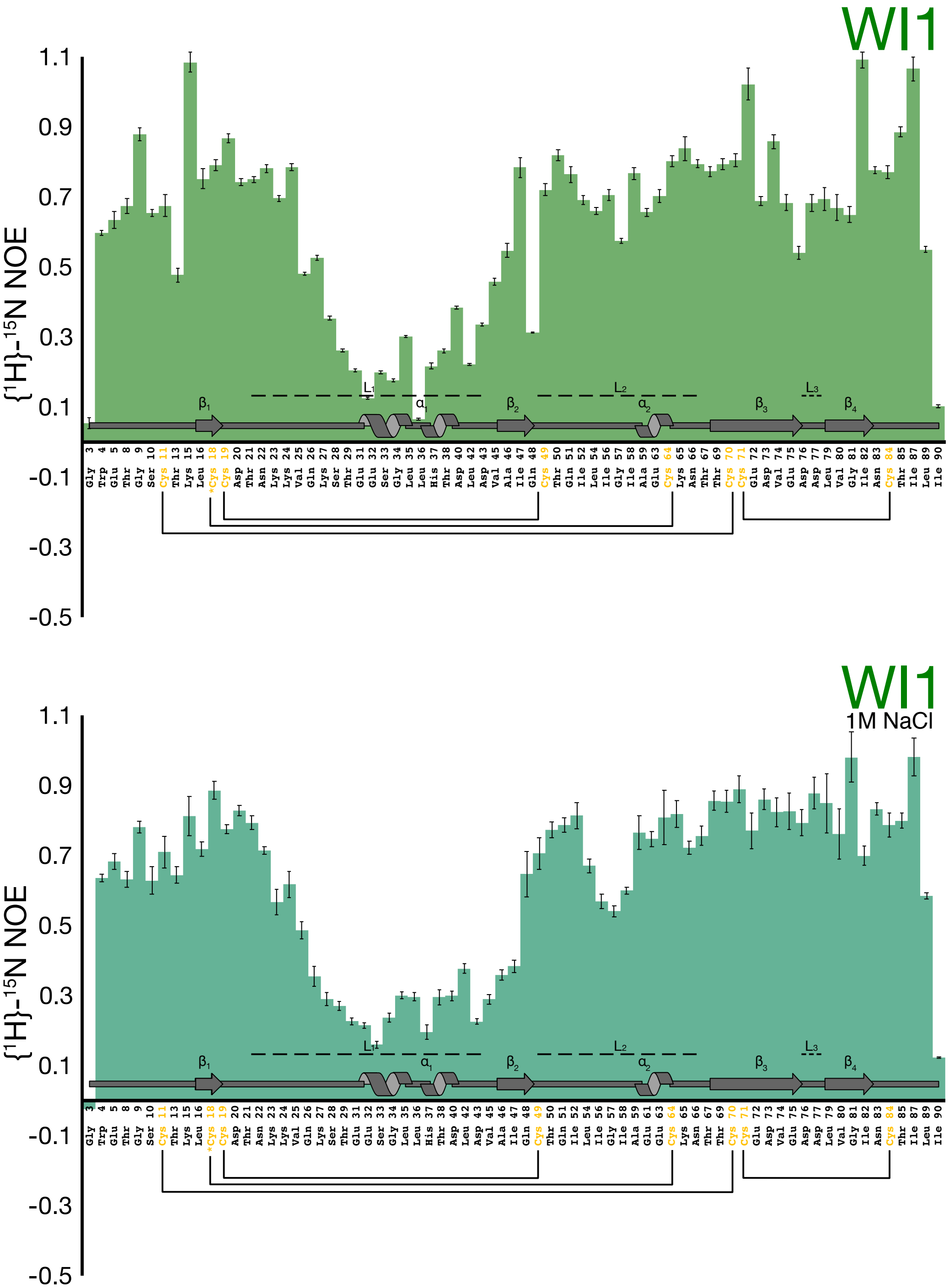


Figure 4.8. {^{1}H}-^{15}N heteronuclear NOE of ^{15}N-labelled WI1 in 1M salt.
The ps-ns timescale dynamics of WI1 appear largely unchanged upon addition of 1M NaCl (Bottom) in comparison to low salt 50 mM conditions (Top) (Figure 3.12).

Increasing pH has a detrimental effect on the ability of PC1 to form rodlets (Figure 4.5). When comparing PC1 ^{1}H-^{15}NHSQC spectra collected at different pH values, peak movement was observed (Figure 4.7), although to a lesser degree than for WI1 in salt. While some peaks disappeared in the PC1 spectra with changing pH due to amide protons more readily exchanging with water, the majority of the observed peaks do not move, such as 61A. Furthermore, the majority of peak movement occurs between pH 5.5 and pH 6.5, while few differences are observed between pH 7.5 and pH 8.5. This is inconsistent with the observed trend in ThT assays (Figure 4.5) where differences in self-assembly were consistently observed between pH 6.5 to pH 8.5. This could mean that PC1 does not undergo any further structural changes with increasing pH but, instead, suffers from decreased functionality due to changes in its surface charge. The increased surface charge of PC1 could prevent monomers from oligomerizing or associating with hydrophobic-hydrophilic interfaces instead of inhibiting structural changes involving L_1.

If a series of peaks corresponding to a specific region of the hydrophobin is perturbed between conditions, then it is likely that this region plays an important role in the differences in functionality observed. Per residue chemical shift changes were calculated for both WI1 salt and PC1 pH titrations to determine which regions are affected in either of these conditions (Figure 4.9). When the most significant shifts are mapped to the structure of WI1 and PC1 they localize to α-helix containing L_1 and β_2 (Figure 4.10). For WI1, a few other peaks within L_1 undergo significant chemical shift changes but ultimately disappear, namely 23K, 25V, and 27K. Figure 4.10 highlights the importance of L_1 and the buried hydrophobic patch on the self-assembly process.

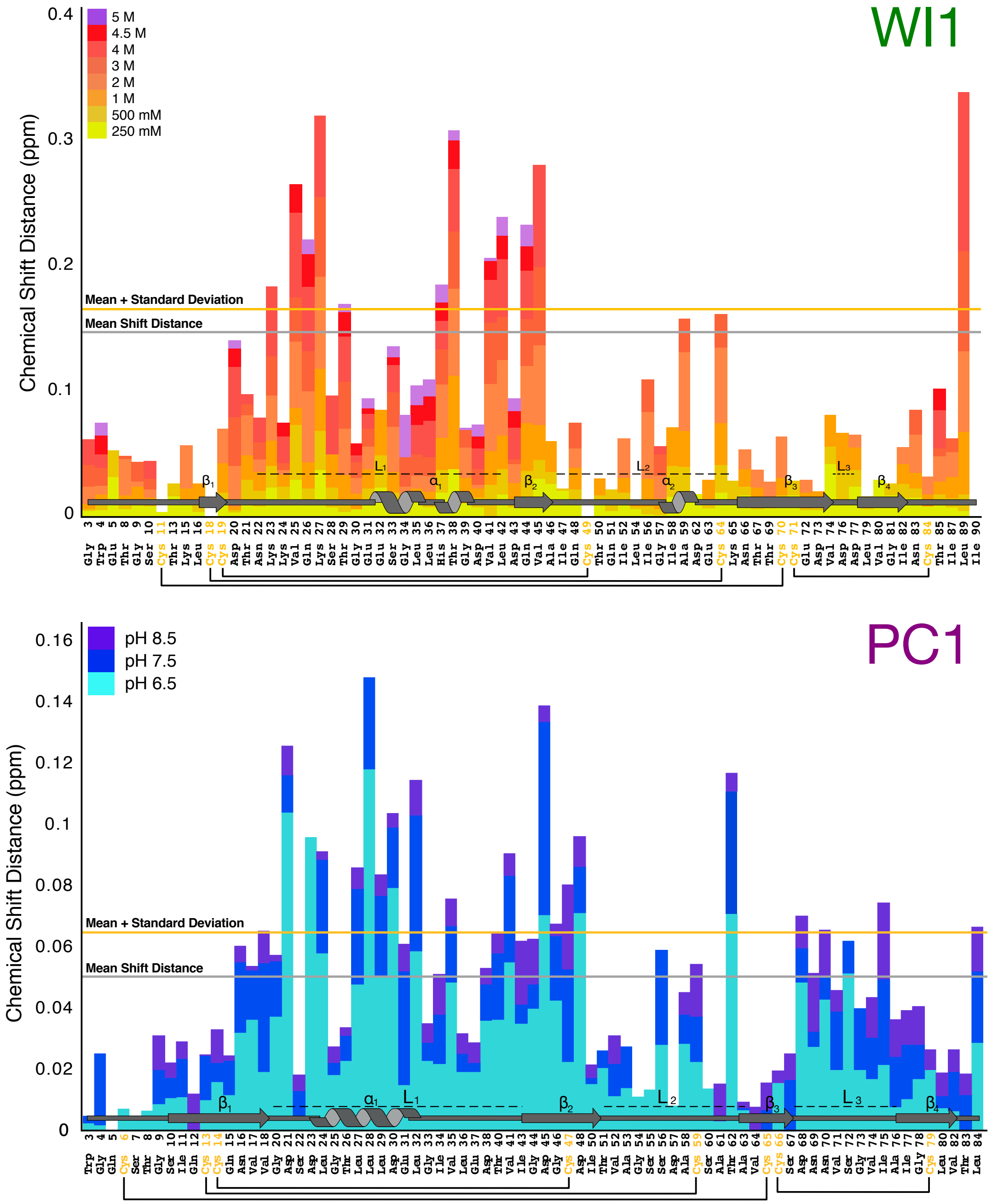


Figure 4.9. Chemical shift changes of WI1 and PC1 in response to pH and salt. Chemical shift changes were measured in ^{1}H-^{15}N HSQC spectra (Figure 4.7) and plotted on a per-residue basis for WI1 (top) and PC1 (bottom). The mean chemical shift distance and mean + standard deviation are shown. Hydrophobin secondary structure features and disulfide bonding patterns are summarized at the bottom of each chart. Chemical shift changes were determined using CCPNMR Analysis.[190]

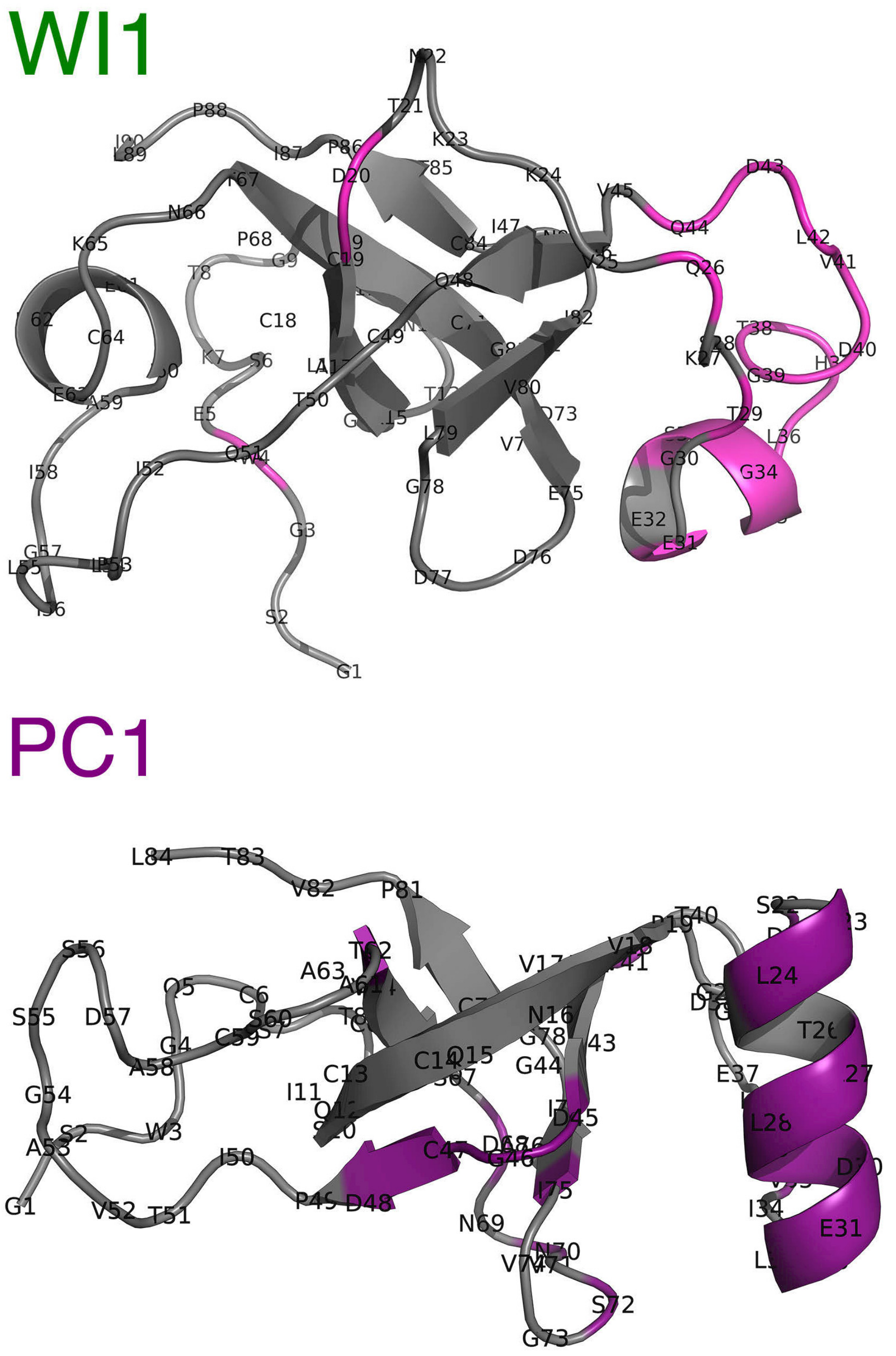


Figure 4.10. The L$_1$ region of WI1 and PC1 is perturbed upon changing salt concentration or pH.

A ribbon representation of WI1 (top) and PC1 (bottom). Residues that remain visible at 5 M salt for WI1 are shown in pink. For PC1, residues with chemical shift differences greater than mean chemical shift plus one standard deviation are coloured in purple.

In class IB hydrophobins the interface between L_1 and β_2 is hydrophobic and, if a structural change were to occur, this patch would become exposed. Hydrophobic patches are important for self-assembly in other classes of hydrophobins and the lack of a clear surface hydrophobic patch in SL1, WI1, and PC1 makes this buried patch of potentially critical importance. Any structural changes in response to salt concentration or pH within this region, such as L_1 unfolding to expose this patch (Figure 4.11), would therefore result in changes in rodlet forming ability. These results the suggest that the L_1 and β_2 regions are uniquely important for determining the function of SL1, WI1, and PC1, and is likely a common property of class IB hydrophobins.

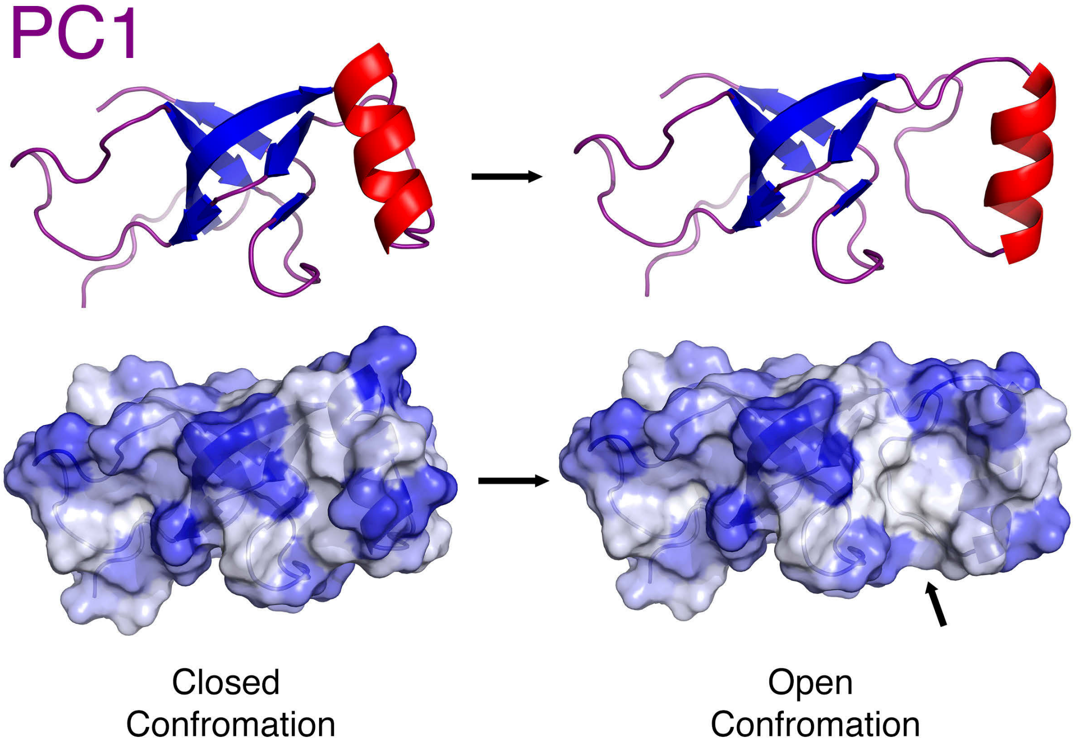


Figure 4.11. L_1 region opening in PC1 exposes a hydrophobic patch.
A ribbon representation (top) and surface plot (bottom) of PC1 with underlying secondary structure features shown and with residues coloured blue according to hydrophilicity (hydrophobic residues are shown in white). Proposed movement of the L_1 region exposes the buried hydrophobic patch marked by an arrow.

4.4 Conclusion

In this chapter, I probed the characteristics of and factors affecting class IB hydrophobin assembly. While AFM experiments were inconclusive due to issues with sample preparation, ThT fluorescence assays indicate that each target hydrophobin readily self-assembles in a surface dependent manner. Upon maximizing contact with air-water interfaces using gentle agitation, each hydrophobin formed an opaque white suspension. This suspension could not be separated by centrifugation, a property potentially reflective of the ability of hydrophobins to assemble at air water interfaces and modify surface tension.[13,219,220]

Further examination into how environmental conditions affect self-assembly ability determined that hydrophobin assembly is sensitive to local environment. Since the ability for WI1 and PC1 to self-assemble is sensitive to salt concentration and pH respectively, the effect these conditions have on WI1 and PC1 structure was examined. Comparing ^{1}H-^{15}N HSQC spectra for each condition reveals that in both cases, the L_1 and β_2 regions are affected the most upon increase salt concentration or pH. This may be connected to the presence of a buried hydrophobic patch in the region, and the established importance of hydrophobic patches on hydrophobin self-assembly.[70,85,220,221] In the model I propose movement of the L_1 region in response to environmental conditions exposes a hydrophobic patch (Figure 4.11), and could represent a structural basis for rodlet assembly. As such, the function of class IB hydrophobins is connected to this shared buried hydrophobic patch while their specific response to environmental conditions is likely dictated by other structural features and sequence properties.

Chapter 5: Conclusion

The aim of this work was to connect the sequence, structure and function of four novel hydrophobins that belong to the same subclass. With the results generated and presented in the previous chapters, I have established an optimized method for heterologous expression of recombinant hydrophobin constructs. Using samples prepared in this way, structural characterization using NMR spectroscopy showed that class IB hydrophobins share distinct structural features, including a typically ordered L_1 region and the presence of buried hydrophobic face between L_1 and β_2. Examination of self-assembly properties revealed that each hydrophobin has a unique response to its environment, which is connected to the number of charged residues within its sequence. Further structural analysis revealed that the buried hydrophobin face and the action of L_1 are critical to class IB hydrophobin self-assembly activity. Consequently, this work represents the first systematic characterization of hydrophobins from disparate sources which have been shown to share sequence and structural features.

5.1 Future Directions

Although detailed structural characterization of SL1, WI1 and PC1 was provided, a complete structural basis for hydrophobin functionality remains unresolved. This self-assembly activity could be better understood by further examining structure function correlations, as outlined in this work. For example, the structural effect of increasing pH on SL1 should be examined to determine if the same changes in L_1 are observed, despite no observed difference in rodlet formational ability. Despite the functional activity of class IB hydrophobins being confirmed in the previous chapter, characterization of assembly morphology remains to be resolved.

To test my hypothesis of the critical role that the buried hydrophobic patch plays in class IB hydrophobin assembly, mutagenesis experiments could be performed. By preparing constructs with extra disulphide bonds which constrict the conformation of the inter-cysteine loop regions, the role of each loop could be more concretely determined. For example, a construct could be prepared which contained an extra disulphide bond that linked L_1 and β_2, thereby preventing L_1 from unfolding and exposing the buried hydrophobic patch. If the functionality of this mutated construct was severely affected, the importance of L_1 unfolding and the role the hydrophobic patch plays in self-assembly would be confirmed.

The unique properties of WI1 also remains to be better understood. It was observed that WI1 did not display linear trends during the ThT assay as observed with other hydrophobins, but instead somewhat adopted an irregular relationship between NaCl concentration, pH and self-assembly ability. The effect of different types of salts, closer to those found in natural environments, could be examined to help better understand the observed activity of WI1.

Bibliography

(1) Linder, M. B., Szilvay, G. R., Nakari-Setälä, T., and Penttilä, M. E. (2005) Hydrophobins: the protein-amphiphiles of filamentous fungi. *FEMS Microbiol. Rev. 29*, 877–896.

(2) Pham, C. L. L., Rodríguez de Francisco, B., Valsecchi, I., Dazzoni, R., Pillé, A., Lo, V., Ball, S. R., Cappai, R., Wien, F., Kwan, A. H., *et al.* (2018) Probing Structural Changes during Self-assembly of Surface-Active Hydrophobin Proteins that Form Functional Amyloids in Fungi. *J. Mol. Biol. 430*, 3784–3801.

(3) Ainsworth, G. C. (2008) Ainsworth & Bisby's Dictionary of the Fungi. CABI.

(4) Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T., Kirk, P.M., Lücking, R., et al. (2007). A higher-level phylogenetic classification of the Fungi. Mycological Research *111*, 509–547.

(5) Kodani, S., Lodato, M. A., Durrant, M. C., Picart, F., and Willey, J. M. (2005) SapT, a lanthionine-containing peptide involved in aerial hyphae formation in the streptomycetes. *Mol. Microbiol. 58*, 1368–1380.

(6) File:Majavamm - Wikimedia Commons. Accessed : March 6, 2019. https://commons.wikimedia.org/wiki/File:Majavamm.jpg

(7) BASF. (2010) Proteins for surfaces.

(8) Linder, M. B. (2009) Hydrophobins: Proteins that self assemble at interfaces. *Curr. Opin. Colloid Interface Sci. 14*, 356–363.

(9) Hess, W. M., Sassen, M. M. A., and Remsen, C. C. (1968) Surface Characteristics of Penicillium Conidia. *Mycologia 60*, 290–303.

(10) Wessels, J. G., de Vries, O. M., Asgeirsdóttir, S. A., and Springer, J. (1991) The thn mutation of Schizophyllum commune, which suppresses formation of aerial hyphae, affects expression of the SC3 hydrophobin gene. *J. Gen. Microbiol. 137*, 2439–2445.

(11) Lacroix, H., and Spanu, P. D. (2009) Silencing of Six Hydrophobins in Cladosporium fulvum: Complexities of Simultaneously Targeting Multiple Genes. *Appl. Environ. Microbiol. 75*, 542–546.

(12) Lacroix, H., Whiteford, J. R., and Spanu, P. D. (2008) Localization of Cladosporium fulvum hydrophobins reveals a role for HCf-6 in adhesion. *FEMS Microbiol. Lett. 286*, 136–144.

(13) Bayry, J., Aimanianda, V., Guijarro, J. I., Sunde, M., and Latgé, J.-P. (2012) Hydrophobins—Unique Fungal Proteins. *PLOS Pathog. 8*, e1002700.

(14) Talbot, N. J., Ebbole, D. J., and Hamer, J. E. (1993) Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus Magnaporthe grisea. *Plant Cell 5*, 1575–1590.

(15) McCabe, P. M., and Van Alfen, N. K. (1999) Secretion of Cryparin, a Fungal Hydrophobin. *Appl. Environ. Microbiol. 65*, 5431–5435.

(16) Dubey, M. K., Jensen, D. F., and Karlsson, M. (2014) Hydrophobins are required for conidial hydrophobicity and plant root colonization in the fungal biocontrol agent Clonostachys rosea. *BMC Microbiol. 14*, 18.

(17) Wessels, J., De Vries, O., Asgeirsdottir, S., and Schuren, F. (1991) Hydrophobin Genes Involved in Formation of Aerial Hyphae and Fruit Bodies in Schizophyllum. *Plant Cell 3*, 793–799.

(18) Wösten, H. A. B., Richter, M., and Willey, J. M. (1999) Structural Proteins Involved in Emergence of Microbial Aerial Hyphae. *Fungal Genet. Biol. 27*, 153–160.

(19) De Groot, P. W., Schaap, P. J., Sonnenberg, A. S., Visser, J., and Van Griensven, L. J. (1996) The Agaricus bisporus hypA gene encodes a hydrophobin and specifically accumulates in peel tissue of mushroom caps during fruit body development. *J. Mol. Biol. 257*, 1008–1018.

(20) Stringer, M. A., Dean, R. A., Sewall, T. C., and Timberlake, W. E. (1991) Rodletless, a new Aspergillus developmental mutant induced by directed gene inactivation. *Genes Dev. 5*, 1161–1171.

(21) Beever, R. E., and Dempsey, G. P. (1978) Function of rodlets on the surface of fungal spores. *Nature 272*, 608.

(22) Lauter, F. R., Russo, V. E., and Yanofsky, C. (1992) Developmental and light regulation of eas, the structural gene for the rodlet protein of Neurospora. *Genes Dev. 6*, 2373–2381.

(23) Bell-Pedersen, D., Dunlap, J. C., and Loros, J. J. (1992) The Neurospora circadian clock-controlled gene, ccg-2, is allelic to eas and encodes a fungal hydrophobin required for formation of the conidial rodlet layer. *Genes Dev. 6*, 2382–2394.

(24) Whiteford, J. R., and Spanu, P. D. (2001) The Hydrophobin HCf-1 of Cladosporium fulvum Is Required for Efficient Water-Mediated Dispersal of Conidia. *Fungal Genet. Biol. 32*, 159–168.

(25) Wetter, M.-A. van, Schuren, F. H. J., Schuurs, T. A., and Wessels, J. G. H. (1996) Targeted mutation of the SC3 hydrophobin gene of Schizophyllum commune affects formation of aerial hyphae. *FEMS Microbiol. Lett. 140*, 265–269.

(26) Lugones, L. G., Wös, H. A. B., and Wessels, J. G. H. (1998) A hydrophobin (ABH3) specifically secreted by vegetatively growing hyphae of Agaricus bisporus (common white button mushroom). *Microbiology 144*, 2345–2353.

(27) Wösten, H. A., Schuren, F. H., and Wessels, J. G. (1994) Interfacial self-assembly of a hydrophobin into an amphipathic protein membrane mediates fungal attachment to hydrophobic surfaces. *EMBO J. 13*, 5848–5854.

(28) St. Leger, R. J., Staples, R. C., and Roberts, D. W. (1992) Cloning and regulatory analysis of starvation-stress gene, ssgA, encoding a hydrophobin-like protein from the entomopathogenic fungus, Metarhizium anisopliae. *Gene 120*, 119–124.

(29) Talbot, N. J., Kershaw, M. J., Wakley, G. E., De Vries, O., Wessels, J., and Hamer, J. E. (1996) MPG1 Encodes a Fungal Hydrophobin Involved in Surface Interactions during Infection-Related Development of Magnaporthe grisea. *Plant Cell 8*, 985–999.

(30) Temple, B., and Horgen, P. A. (2000) Biological Roles for Cerato-Ulmin, a Hydrophobin Secreted by the Elm Pathogens, Ophiostoma ulmi and O. novo-ulmi. *Mycologia 92*, 1–9.

(31) Kershaw, M. J., Wakley, G., and Talbot, N. J. (1998) Complementation of the mpg1 mutant phenotype in Magnaporthe grisea reveals functional relationships between fungal hydrophobins. *EMBO J. 17*, 3838–3849.

(32) Kim, S., Ahn, I.-P., Rho, H.-S., and Lee, Y.-H. (2005) MHP1, a Magnaporthe grisea hydrophobin gene, is required for fungal development and plant colonization. *Mol. Microbiol. 57*, 1224–1237.

(33) Aimanianda, V., Bayry, J., Bozza, S., Kniemeyer, O., Perruccio, K., Elluru, S. R., Clavaud, C., Paris, S., Brakhage, A. A., Kaveri, S. V., Romani, L., and Latgé, J.-P. (2009) Surface hydrophobin prevents immune recognition of airborne fungal spores. *Nature 460*, 1117–1121.

(34) Dagenais, T. R. T., Giles, S. S., Aimanianda, V., Latgé, J.-P., Hull, C. M., and Keller, N. P. (2010) Aspergillus fumigatus LaeA-mediated phagocytosis is associated with a decreased hydrophobin layer. *Infect. Immun. 78*, 823–829.

(35) Hohl, T. M., Van Epps, H. L., Rivera, A., Morgan, L. A., Chen, P. L., Feldmesser, M., and Pamer, E. G. (2005) Aspergillus fumigatus triggers inflammatory responses by stage-specific beta-glucan display. *PLoS Pathog. 1*, e30.

(36) Bruns, S., Kniemeyer, O., Hasenberg, M., Aimanianda, V., Nietzsche, S., Thywißen, A., Jeron, A., Latgé, J.-P., Brakhage, A. A., and Gunzer, M. (2010) Production of Extracellular Traps against Aspergillus fumigatus In Vitro and in Infected Lung Tissue Is Dependent on Invading Neutrophils and Influenced by Hydrophobin RodA. *PLOS Pathog. 6*, e1000873.

(37) Butko, P., Buford, J. P., Goodwin, J. S., Stroud, P. A., McCormick, C. L., and Cannon, G. C. (2001) Spectroscopic Evidence for Amyloid-like Interfacial Self-Assembly of Hydrophobin Sc3. *Biochem. Biophys. Res. Commun. 280*, 212–215.

(38) Corvis, Y., Brezesinski, G., Rink, R., Walcarius, A., Van der Heyden, A., Mutelet, F., and Rogalska, E. (2006) Analytical Investigation of the Interactions between SC3 Hydrophobin and Lipid Layers: Elaborating of Nanostructured Matrixes for Immobilizing Redox Systems. *Anal. Chem. 78*, 4850–4864.

(39) de Vocht, M. L., Reviakine, I., Ulrich, W.-P., Bergsma-Schutter, W., Wösten, H. A. B., Vogel, H., Brisson, A., Wessels, J. G. H., and Robillard, G. T. (2002) Self-assembly of the hydrophobin SC3 proceeds via two structural intermediates. *Protein Sci. Publ. Protein Soc. 11*, 1199–1205.

(40) de Vocht, M. L., Scholtmeijer, K., Van der Vegte, E. W., de Vries, O. M. H., Sonveaux, N., Wösten, H. A. B., Ruysschaert, J.-M., Hadziioannou, G., Wessels, J. G. H., and Robillard, G. T. (1998) Structural Characterization of the Hydrophobin SC3, as a Monomer and after Self-Assembly at Hydrophobic/Hydrophilic Interfaces. *Biophys. J. 74*, 2059–2068.

(41) Stroud, P. A., Goodwin, J. S., Butko, P., Cannon, G. C., and McCormick, C. L. (2003) Experimental Evidence for Multiple Assembled States of Sc3 from Schizophyllum commune. *Biomacromolecules 4*, 956–967.

(42) Gandier, J.-A., Langelaan, D. N., Won, A., O'Donnell, K., Grondin, J. L., Spencer, H. L., Wong, P., Tillier, E., Yip, C., Smith, S. P., and Master, E. R. (2017) Characterization of a Basidiomycota hydrophobin reveals the structural basis for a high-similarity Class I subdivision. *Sci. Rep. 7*, 45863.

(43) Littlejohn, K. A., Hooley, P., and Cox, P. W. (2012) Bioinformatics predicts diverse Aspergillus hydrophobins with novel properties. *Food Hydrocoll. 27*, 503–516.

(44) Mgbeahuruike, A. C., Kovalchuk, A., Chen, H., Ubhayasekera, W., and Asiegbu, F. O. (2013) Evolutionary analysis of hydrophobin gene family in two wood-degrading basidiomycetes, Phlebia brevispora and Heterobasidion annosum s.l. *BMC Evol. Biol. 13*, 240.

(45) Kwan, A. H. Y., Winefield, R. D., Sunde, M., Matthews, J. M., Haverkamp, R. G., Templeton, M. D., and Mackay, J. P. (2006) Structural basis for rodlet assembly in fungal hydrophobins. *Proc. Natl. Acad. Sci. U. S. A. 103*, 3621–3626.

(46) Vocht, M. L. de, Reviakine, I., Wösten, H. A. B., Brisson, A., Wessels, J. G. H., and Robillard, G. T. (2000) Structural and Functional Role of the Disulfide Bridges in the Hydrophobin SC3. *J. Biol. Chem. 275*, 28428–28432.

(47) Wessels, J. G. H. (1994) Developmental Regulation of Fungal Cell Wall Formation. *Annu. Rev. Phytopathol. 32*, 413–437.

(48) Sunde, M., Kwan, A. H. Y., Templeton, M. D., Beever, R. E., and Mackay, J. P. (2008) Structural analysis of hydrophobins. *Micron 39*, 773–784.

(49) Jensen, B. G., Andersen, M. R., Pedersen, M. H., Frisvad, J. C., and Søndergaard, I. (2010) Hydrophobins from Aspergillus species cannot be clearly divided into two classes. *BMC Res. Notes 3*, 344.

(50) Wösten, H. A. B., Ruardy, T. G., Van der Mei, H. C., Busscher, H. J., and Wessels, J. G. H. (1995) Interfacial self-assembly of a Schizophyllum commune hydrophobin into an insoluble amphipathic protein membrane depends on surface hydrophobicity. *Colloids Surf. B Biointerfaces 5*, 189–195.

(51) Wosten, H., Vries, O. D., and Wessels, J. (1993) Interfacial Self-Assembly of a Fungal Hydrophobin into a Hydrophobic Rodlet Layer. *Plant Cell 5*, 1567–1574.

(52) Wösten, H. A. B., van Wetter, M.-A., Lugones, L. G., Van der Mei, H. C., Busscher, H. J., and Wessels, J. G. H. (1999) How a fungus escapes the water to grow into the air. *Curr. Biol. 9*, 85–88.

(53) Vries, O. M. H. de, Fekkes, M. P., Wösten, H. A. B., and Wessels, J. G. H. (1993) Insoluble hydrophobin complexes in the walls of Schizophyllum commune and other filamentous fungi. *Arch. Microbiol. 159*, 330–335.

(54) Wösten, H. A. B., and Scholtmeijer, K. (2015) Applications of hydrophobins: current state and perspectives. *Appl. Microbiol. Biotechnol. 99*, 1587–1597.

(55) Askolin, S., Linder, M., Scholtmeijer, K., Tenkanen, M., Penttilä, M., de Vocht, M. L., and Wösten, H. A. B. (2006) Interaction and comparison of a class I hydrophobin from Schizophyllum commune and class II hydrophobins from Trichoderma reesei. *Biomacromolecules 7*, 1295–1301.

(56) Wösten, H. A. B. (2001) Hydrophobins: Multipurpose Proteins. *Annu. Rev. Microbiol. 55*, 625–646.

(57) Winandy, L., Hilpert, F., Schlebusch, O., and Fischer, R. (2018) Comparative analysis of surface coating properties of five hydrophobins from Aspergillus nidulans and Trichoderma reseei. *Sci. Rep. 8*, 12033.

(58) Kwan, A. H., Macindoe, I., Vukašin, P. V., Morris, V. K., Kass, I., Gupte, R., Mark, A. E., Templeton, M. D., Mackay, J. P., and Sunde, M. (2008) The Cys3–Cys4 Loop of the Hydrophobin EAS Is Not Required for Rodlet Formation and Surface Activity. *J. Mol. Biol. 382*, 708–720.

(59) Ren, Q., Kwan, A. H., and Sunde, M. (2014) Solution structure and interface-driven self-assembly of NC2, a new member of the Class II hydrophobin proteins. *Proteins Struct. Funct. Bioinforma. 82*, 990–1003.

(60) Grünbacher, A., Throm, T., Seidel, C., Gutt, B., Röhrig, J., Strunk, T., Vincze, P., Walheim, S., Schimmel, T., Wenzel, W., and Fischer, R. (2014) Six Hydrophobins Are Involved in Hydrophobin Rodlet Formation in Aspergillus nidulans and Contribute to Hydrophobicity of the Spore Surface. *PLOS ONE 9*, e94546.

(61) Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment11Edited by J. Thornton. *J. Mol. Biol. 302*, 205–217.

(62) Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res. 34*, W604–W608.

(63) Morris, V. K., Kwan, A. H., and Sunde, M. (2013) Analysis of the Structure and Conformational States of DewA Gives Insight into the Assembly of the Fungal Hydrophobins. *J. Mol. Biol. 425*, 244–256.

(64) Pham, C. L. L., Rey, A., Lo, V., Soulès, M., Ren, Q., Meisl, G., Knowles, T. P. J., Kwan, A. H., and Sunde, M. (2016) Self-assembly of MPG1, a hydrophobin protein from the rice blast fungus that forms functional amyloid coatings, occurs by a surface-driven mechanism. *Sci. Rep. 6*, 25288.

(65) Hakanpää, J., Szilvay, G. R., Kaljunen, H., Maksimainen, M., Linder, M., and Rouvinen, J. (2006) Two crystal structures of Trichoderma reesei hydrophobin HFBI—The structure of a protein amphiphile with and without detergent interaction. *Protein Sci. 15*, 2129–2140.

(66) Gandier, J.-A., and Master, E. R. (2018) *Pichia pastoris* is a Suitable Host for the Heterologous Expression of Predicted Class I and Class II Hydrophobins for Discovery, Study, and Application in Biotechnology. *Microorganisms 6*, 3.

(67) Morris, V. K., Linser, R., Wilde, K. L., Duff, A. P., Sunde, M., and Kwan, A. H. (2012) Solid-State NMR Spectroscopy of Functional Amyloid from a Fungal Hydrophobin: A Well-Ordered β-Sheet Core Amidst Structural Heterogeneity. *Angew. Chem. Int. Ed. 51*, 12621–12625.

(68) Rey, A. A., Hocher, A., Kwan, A. H., and Sunde, M. (2013) Backbone and sidechain ^{1}H, ^{13}C and ^{15}N chemical shift assignments of the hydrophobin MPG1 from the rice blast fungus Magnaporthe oryzae. *Biomol. NMR Assign. 7*, 109–112.

(69) Hakanpää, J., Linder, M., Popov, A., Schmidt, A., and Rouvinen, J. (2006) Hydrophobin HFBII in detail: ultrahigh-resolution structure at 0.75 Å. *Acta Crystallogr. D Biol. Crystallogr. 62*, 356–367.

(70) Hakanpää, J., Paananen, A., Askolin, S., Nakari-Setälä, T., Parkkinen, T., Penttilä, M., Linder, M. B., and Rouvinen, J. (2004) Atomic Resolution Structure of the HFBII Hydrophobin, a Self-assembling Amphiphile. *J. Biol. Chem. 279*, 534–539.

(71) Ren, Q., Kwan, A. H., and Sunde, M. (2013) Two forms and two faces, multiple states and multiple uses: Properties and applications of the self-assembling fungal hydrophobins. *Pept. Sci. 100*, 601–612.

(72) Linder, M., Selber, K., Nakari-Setälä, T., Qiao, M., Kula, M.-R., and Penttilä, M. (2001) The Hydrophobins HFBI and HFBII from Trichoderma reesei Showing Efficient Interactions with Nonionic Surfactants in Aqueous Two-Phase Systems. *Biomacromolecules 2*, 511–517.

(73) Zykwinska, A., Guillemette, T., Bouchara, J. P., and Cuenot, S. (2014) Spontaneous self-assembly of SC3 hydrophobins into nanorods in aqueous solution. *Biochim. Biophys. Acta 1844*, 1231–1237.

(74) Kershaw, M. J., Thornton, C. R., Wakley, G. E., and Talbot, N. J. (2005) Four conserved intramolecular disulphide linkages are required for secretion and cell wall localization of a hydrophobin during fungal morphogenesis. *Mol. Microbiol. 56*, 117–125.

(75) Wang, X., Permentier, H. P., Rink, R., Kruijtzer, J. A. W., Liskamp, R. M. J., Wösten, H. A. B., Poolman, B., and Robillard, G. T. (2004) Probing the Self-Assembly and the Accompanying Structural Changes of Hydrophobin SC3 on a Hydrophobic Surface by Mass Spectrometry. *Biophys. J. 87*, 1919–1928.

(76) Wang, X., Graveland-Bikker, J. F., De Kruif, C. G., and Robillard, G. T. (2004) Oligomerization of hydrophobin SC3 in solution: From soluble state to self-assembly. *Protein Sci. Publ. Protein Soc. 13*, 810–821.

(77) Macindoe, I., Kwan, A. H., Ren, Q., Morris, V. K., Yang, W., Mackay, J. P., and Sunde, M. (2012) Self-assembly of functional, amphipathic amyloid monolayers by the fungal hydrophobin EAS. *Proc. Natl. Acad. Sci. U. S. A. 109*, E804-811.

(78) De Simone, A., Kitchen, C., Kwan, A. H., Sunde, M., Dobson, C. M., and Frenkel, D. (2012) Intrinsic disorder modulates protein self-assembly and aggregation. *Proc. Natl. Acad. Sci. U. S. A. 109*, 6951–6956.

(79) Pille, A., Kwan, A. H., Cheung, I., Hampsey, M., Aimanianda, V., Delepierre, M., Latgé, J.-P., Sunde, M., and Guijarro, J. I. (2015) ^{1}H, ^{13}C and ^{15}N resonance assignments of the RodA hydrophobin from the opportunistic pathogen Aspergillus fumigatus. *Biomol. NMR Assign. 9*, 113–118.

(80) Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., de la Paz, M. L., Martins, I. C., Reumers, J., Morris, K. L., Copland, A., Serpell, L., Serrano, L., Schymkowitz, J. W. H., and Rousseau, F. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods 7*, 237–242.

(81) Kisko, K., Szilvay, G. R., Vainio, U., Linder, M. B., and Serimaa, R. (2008) Interactions of Hydrophobin Proteins in Solution Studied by Small-Angle X-Ray Scattering. *Biophys. J. 94*, 198–206.

(82) Sallada, N. D., Dunn, K. J., and Berger, B. W. (2018) A Structural and Functional Role for Disulfide Bonds in a Class II Hydrophobin. *Biochemistry 57*, 645–653.

(83) Sunde, M., Pham, C. L. L., and Kwan, A. H. (2017) Molecular Characteristics and Biological Functions of Surface-Active and Surfactant Proteins. *Annu. Rev. Biochem. 86*, 585–608.

(84) Lo, V. C., Ren, Q., Pham, C. L. L., Morris, V. K., Kwan, A. H., and Sunde, M. (2014) Fungal Hydrophobin Proteins Produce Self-Assembling Protein Films with Diverse Structure and Chemical Stability. *Nanomaterials 4*, 827–843.

(85) Szilvay, G. R., Paananen, A., Laurikainen, K., Vuorimaa, E., Lemmetyinen, H., Peltonen, J., and Linder, M. B. (2007) Self-Assembled Hydrophobin Protein Films at the Air−Water Interface: Structural Analysis and Molecular Engineering. *Biochemistry 46*, 2345–2354.

(86) Van der Vegt, W., Van der Mei, H. C., Wösten, H. A. B., Wessels, J. G. H., and Busscher, H. J. (1996) A comparison of the surface activity of the fungal hydrophobin SC3p with those of other proteins. *Biophys. Chem. 57*, 253–260.

(87) Linder, M., Szilvay, G. R., Nakari-Setälä, T., Söderlund, H., and Penttilä, M. (2002) Surface adhesion of fusion proteins containing the hydrophobins HFBI and HFBII from Trichoderma reesei. *Protein Sci. Publ. Protein Soc. 11*, 2257–2266.

(88) Gravagnuolo, A. M., Longobardi, S., Luchini, A., Appavou, M.-S., De Stefano, L., Notomista, E., Paduano, L., and Giardina, P. (2016) Class I Hydrophobin Vmh2 Adopts Atypical Mechanisms to Self-Assemble into Functional Amyloid Fibrils. *Biomacromolecules 17*, 954–964.

(89) Khalesi, M., Sylvie, D., Gebruers, K., Vissers, L., Verachtert, H., and Derdelinckx, G. (2012) Hydrophobins: Exceptional proteins for many applications in brewery environment and other bio-industries. *Cerevisia 37*, 3–9.

(90) Khalesi, M., Gebruers, K., and Derdelinckx, G. (2015) Recent Advances in Fungal Hydrophobin Towards Using in Industry. *Protein J. 34*, 243–255.

(91) Khalesi, M., Jahanbani, R., Riveros-Galan, D., Sheikh-Hassani, V., Sheikh-Zeinoddin, M., Sahihi, M., Winterburn, J., Derdelinckx, G., and Moosavi-Movahedi, A. A. (2016) Antioxidant activity and ACE-inhibitory of Class II hydrophobin from wild strain Trichoderma reesei. *Int. J. Biol. Macromol. 91*, 174–179.

(92) Tchuenbou-Magaia, F. L., Norton, I. T., and Cox, P. W. (2009) Hydrophobins stabilised air-filled emulsions for the food industry. *Food Hydrocoll. 23*, 1877–1885.

(93) Blesic, M., Dichiarante, V., Milani, R., Linder, M., and Metrangolo, P. (2018) Evaluating the potential of natural surfactants in the petroleum industry: the case of hydrophobins. *Pure Appl. Chem. 90*, 305–314.

(94) Singh, B. N., Singh, B. R., Gupta, V. K., Kharwar, R. N., and Pecoraro, L. (2018) Coating with Microbial Hydrophobins: A Novel Approach to Develop Smart Drug Nanoparticles. *Trends Biotechnol.*

(95) Scholtmeijer, K., Wessels, J., and Wösten, H. (2001) Fungal hydrophobins in medical and technical applications. *Appl. Microbiol. Biotechnol. 56*, 1–8.

(96) Reuter, L. J., Shahbazi, M.-A., Mäkilä, E. M., Salonen, J. J., Saberianfar, R., Menassa, R., Santos, H. A., Joensuu, J. J., and Ritala, A. (2017) Coating Nanoparticles with Plant-Produced Transferrin–Hydrophobin Fusion Protein Enhances Their Uptake in Cancer Cells. *Bioconjug. Chem. 28*, 1639–1648.

(97) Valo, H. K., Laaksonen, P. H., Peltonen, L. J., Linder, M. B., Hirvonen, J. T., and Laaksonen, T. J. (2010) Multifunctional Hydrophobin: Toward Functional Coatings for Drug Nanoparticles. *ACS Nano 4*, 1750–1758.

(98) Valo, H., Kovalainen, M., Laaksonen, P., Häkkinen, M., Auriola, S., Peltonen, L., Linder, M., Järvinen, K., Hirvonen, J., and Laaksonen, T. (2011) Immobilization of protein-coated drug nanoparticles in nanofibrillar cellulose matrices—Enhanced stability and release. *J. Controlled Release 156*, 390–397.

(99) Rocha-Pino, Z., Ramos-López, J. I., Gimeno, M., Barragán-Aroche, F., Durán-Valencia, C., López-Ramírez, S., and Shirai, K. (2018) Enhanced oil recovery by hydrophobins from Lecanicillium lecanii. *Fuel 224*, 10–16.

(100) Green, A. J., Littlejohn, K. A., Hooley, P., and Cox, P. W. (2013) Formation and stability of food foams and aerated emulsions: Hydrophobins as novel functional ingredients. *Curr. Opin. Colloid Interface Sci. 18*, 292–301.

(101) Cox, A. R., Aldred, D. L., and Russell, A. B. (2009) Exceptional stability of food foams using class II hydrophobin HFBII. *Food Hydrocoll. 23*, 366–376.

(102) Crilly, J. F., Russell, A. B., Cox, A. R., and Cebula, D. J. (2008) Designing Multiscale Structures for Desired Properties of Ice Cream. *Ind. Eng. Chem. Res. 47*, 6362–6367.

(103) Wohlleben, W., Subkowski, T., Bollschweiler, C., von Vacano, B., Liu, Y., Schrepp, W., and Baus, U. (2010) Recombinantly produced hydrophobins from fungal analogues as highly surface-active performance proteins. *Eur. Biophys. J. 39*, 457–468.

(104) Li, C., Qin, R., Liu, R., Miao, S., and Yang, P. (2018) Functional amyloid materials at surfaces/interfaces. *Biomater. Sci. 6*, 462–472.

(105) Boeuf, S., Throm, T., Gutt, B., Strunk, T., Hoffmann, M., Seebach, E., Mühlberg, L., Brocher, J., Gotterbarm, T., Wenzel, *et al.* (2012) Engineering hydrophobin DewA to generate surfaces that enhance adhesion of human but not bacterial cells. *Acta Biomater. 8*, 1037–1047.

(106) Weickert, U., Wiesend, F., Subkowski, T., Eickhoff, A., and Reiss, G. (2011) Optimizing biliary stent patency by coating with hydrophobin alone or hydrophobin and antibiotics or heparin: an in vitro proof of principle study. *Adv. Med. Sci. 56*, 138–144.

(107) Misra, R., Li, J., Cannon, G. C., and Morgan, S. E. (2006) Nanoscale Reduction in Surface Friction of Polymer Surfaces Modified with Sc3 Hydrophobin from Schizophyllum commune. *Biomacromolecules 7*, 1463–1470.

(108) Patravale, V. B., Date, A. A., and Kulkarni, R. M. (2004) Nanosuspensions: a promising drug delivery strategy. *J. Pharm. Pharmacol. 56*, 827–840.

(109) Haas Jimoh Akanbi, M., Post, E., Meter-Arkema, A., Rink, R., Robillard, G. T., Wang, X., Wösten, H. A. B., and Scholtmeijer, K. (2010) Use of hydrophobins in formulation of water insoluble drugs for oral administration. *Colloids Surf. B Biointerfaces 75*, 526–531.

(110) Scholtmeijer, K., Janssen, M. I., Gerssen, B., Vocht, M. L. de, Leeuwen, B. M. van, Kooten, T. G. van, Wösten, H. A. B., and Wessels, J. G. H. (2002) Surface Modifications Created by Using Engineered Hydrophobins. *Appl Env. Microbiol 68*, 1367–1373.

(111) Janssen, M. I., van Leeuwen, M. B. M., Scholtmeijer, K., van Kooten, T. G., Dijkhuizen, L., and Wösten, H. a. B. (2002) Coating with genetic engineered hydrophobin promotes growth of fibroblasts on a hydrophobic solid. *Biomaterials 23*, 4847–4854.

(112) Bailey, M., Askolin, S., Hörhammer, N., Tenkanen, M., Linder, M., Penttilä, M., and Nakari-Setälä, T. (2002) Process technological effects of deletion and amplification of hydrophobins I and II in transformants of Trichoderma reesei. *Appl. Microbiol. Biotechnol. 58*, 721–727.

(113) Askolin, S., Nakari-Setälä, T., and Tenkanen, M. (2001) Overproduction, purification, and characterization of the Trichoderma reesei hydrophobin HFBI. *Appl. Microbiol. Biotechnol. 57*, 124–130.

(114) Kottmeier, K., Günther, T. J., Weber, J., Kurtz, S., Ostermann, K., Rödel, G., and Bley, T. (2012) Constitutive expression of hydrophobin HFB1 from Trichoderma reesei in *Pichia pastoris* and its pre-purification by foam separation during cultivation. *Eng. Life Sci. 12*, 162–170.

(115) Niu, B., Huang, Y., Zhang, S., Wang, D., Xu, H., Kong, D., and Qiao, M. (2012) Expression and characterization of hydrophobin HGFI fused with the cell-specific peptide TPS in *Pichia pastoris*. *Protein Expr. Purif. 83*, 92–97.

(116) Niu, B., Wang, D., Yang, Y., Xu, H., and Qiao, M. (2012) Heterologous expression and characterization of the hydrophobin HFBI in *Pichia pastoris* and evaluation of its contribution to the food industry. *Amino Acids 43*, 763–771.

(117) Przylucka, A., Akcapinar, G. B., Bonazza, K., Mello-de-Sousa, T. M., Mach-Aigner, A. R., Lobanov, V., Grothe, H., Kubicek, C. P., Reimhult, E., and Druzhinina, I. S. (2017) Comparative physiochemical analysis of hydrophobins produced in *Escherichia coli* and *Pichia pastoris*. *Colloids Surf. B Biointerfaces 159*, 913–923.

(118) Stübner, M., Lutterschmid, G., Vogel, R. F., and Niessen, L. (2010) Heterologous expression of the hydrophobin FcHyd5p from Fusarium culmorum in *Pichia pastoris* and evaluation of its surface activity and contribution to gushing of carbonated beverages. *Int. J. Food Microbiol. 141*, 110–115.

(119) Wang, Z., Feng, S., Huang, Y., Li, S., Xu, H., Zhang, X., Bai, Y., and Qiao, M. (2010) Expression and characterization of a Grifola frondosa hydrophobin in *Pichia pastoris*. *Protein Expr. Purif. 72*, 19–25.

(120) Winterburn, J. B., Russell, A. B., and Martin, P. J. (2011) Integrated recirculating foam fractionation for the continuous recovery of biosurfactant from fermenters. *Biochem. Eng. J. 54*, 132–139.

(121) Niu, B., Li, B., Wang, H., Guo, R., Xu, H., Qiao, M., and Li, W. (2017) Investigation of the relationship between the rodlet formation and Cys3–Cys4 loop of the HGFI hydrophobin. *Colloids Surf. B Biointerfaces 150*, 344–351.

(122) Kirkland, B. H., and Keyhani, N. O. (2011) Expression and purification of a functionally active class I fungal hydrophobin from the entomopathogenic fungus Beauveria bassiana in *E. coli*. *J. Ind. Microbiol. Biotechnol. 38*, 327–335.

(123) Wang, Z., Feng, S., Huang, Y., Qiao, M., Zhang, B., and Xu, H. (2010) Prokaryotic expression, purification, and polyclonal antibody production of a hydrophobin from Grifola frondosa. *Acta Biochim. Biophys. Sin. 42*, 388–395.

(124) Kudo, T., Sato, Y., Hara, T., Joh, T., and Tasaki, Y. (2011) Heterogeneous expression and emulsifying activity of class I hydrophobin from Pholiota nameko. *Mycoscience 52*, 283–287.

(125) Berkmen, M. (2012) Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expr. Purif. 82*, 240–251.

(126) Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb. Cell Factories 11*, 753.

(127) Lebendiker, M., and Danieli, T. (2014) Production of prone-to-aggregate proteins. *FEBS Lett. 588*, 236–246.

(128) Zajc, J., Kogej, T., Galinski, E. A., Ramos, J., and Gunde-Cimerman, N. (2014) Osmoadaptation Strategy of the Most Halophilic Fungus, Wallemia ichthyophaga, Growing Optimally at Salinities above 15% NaCl. *Appl Env. Microbiol 80*, 247–256.

(129) Zajc, J., Liu, Y., Dai, W., Yang, Z., Hu, J., Gostinčar, C., and Gunde-Cimerman, N. (2013) Genome and transcriptome sequencing of the halophilic fungus Wallemia ichthyophaga: haloadaptations present and absent. *BMC Genomics 14*, 617.

(130) Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571–607. Humana Press, Totowa, NJ.

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

(132) Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif. 41*, 207–234.

(133) Chevallet, M., Luche, S., and Rabilloud, T. (2006) Silver staining of proteins in polyacrylamide gels. *Nat. Protoc. 1*, 1852–1858.

(134) LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., and McCoy, J. M. (1993) A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the *E. coli* Cytoplasm. *Bio/Technology 11*, 187.

(135) Davis, G. D., Elisee, C., Newham, D. M., and Harrison, R. G. (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol. Bioeng. 65*, 382–388.

(136) Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D., and Butt, T. R. (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics 5*, 75–86.

(137) Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A., and Dougherty, W. G. (1994) Release of Proteins and Peptides from Fusion Proteins Using a Recombinant Plant Virus Proteinase. *Anal. Biochem. 216*, 413–417.

(138) Sun, C., Liang, J., Shi, R., Gao, X., Zhang, R., Hong, F., Yuan, Q., and Wang, S. (2012) Tobacco etch virus protease retains its activity in various buffers and in the presence of diverse additives. *Protein Expr. Purif. 82*, 226–231.

(139) Cavanagh, J., Fairbrother, W., Palmer, A., Rance, M., and Skelton, N. (2006) Protein NMR spectroscopy, principles and practice. Academic Press.

(140) Levitt, M. (2008) Spin Dynamics: Basics of nuclear magnetic resonance. NJ: John Wiley & Sons, Chichester, England; Hoboken,.

(141) Becker, E. D. (1993) A brief history of nuclear magnetic resonance. *Anal. Chem. 65*, 295A-302A.

(142) Williamson, M. P., Havel, T. F., and Wüthrich, K. (1985) Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J. Mol. Biol. 182*, 295–315.

(143) Huang, R., Ripstein, Z. A., Rubinstein, J. L., and Kay, L. E. (2019) Cooperative subunit dynamics modulate p97 function. *Proc. Natl. Acad. Sci. 116*, 158–167.

(144) Balbach, J., Forge, V., Nuland, N. A. J. van, Winder, S. L., Hore, P. J., and Dobson, C. M. (1995) Following protein folding in real time using NMR spectroscopy. *Nat. Struct. Biol. 2*, 865.

(145) Brindle, K. M. (1988) NMR methods for measuring enzyme kinetics in vivo. *Prog. Nucl. Magn. Reson. Spectrosc. 20*, 257–293.

(146) Boehr, D. D., Dyson, H. J., and Wright, P. E. (2006) An NMR Perspective on Enzyme Dynamics. *Chem. Rev. 106*, 3055–3079.

(147) Zuiderweg, E. R. P. (2002) Mapping Protein−Protein Interactions in Solution by NMR Spectroscopy. *Biochemistry 41*, 1–7.

(148) Kenward, C., Shin, K., and Rainey, J. K. (2018) Mixed Fluorotryptophan Substitutions at the Same Residue Expand the Versatility of 19F Protein NMR Spectroscopy. *Chem. – Eur. J. 24*, 3391–3396.

(149) Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson. 1969 96*, 432–440.

(150) Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson. 1969 89*, 496–514.

(151) Bax, A., and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly^{15}N/^{13}C enriched proteins. *J. Biomol. NMR 1*, 99–104.

(152) Grzesiek, S., and Bax, A. (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson. 1969 99*, 201–207.

(153) Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc. 114*, 6291–6293.

(154) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135–140.

(155) Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in ^{13}C/^{15}N-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. *J. Magn. Reson. B 101*, 114–119.

(156) Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., and Kay, L. E. (1990) Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of ^{13}C-labeled proteins. *J. Magn. Reson. 1969 87*, 620–627.

(157) Griesinger, C., Otting, G., Wuethrich, K., and Ernst, R. R. (1988) Clean TOCSY for proton spin system identification in macromolecules. *J. Am. Chem. Soc. 110*, 7870–7872.

(158) Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson. 1969 88*, 425–431.

(159) Keeler, J. (2010) Understanding NMR spectroscopy 2nd ed. John Wiley and Sons, Chichester, U.K.

(160) Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J. Am. Chem. Soc. 111*, 1515–1517.

(161) Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry 28*, 6150–6156.

(162) Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T. E., and Nilges, M. (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinforma. Oxf. Engl. 23*, 381–382.

(163) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, *et al.* 1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr. 54*, 905–921.

(164) Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983) Optimization by simulated annealing. *Science 220*, 671–680.

(165) Aarts, E. H. L., and Laarhoven, P. J. M. van. (1989) Simulated annealing: An introduction. *Stat. Neerlandica 43*, 31–52.

(166) Cheung, M.-S., Maguire, M. L., Stevens, T. J., and Broadhurst, R. W. (2010) DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson. San Diego Calif 1997 202*, 223–233.

(167) Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson. 1969 93*, 151–170.

(168) Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc. 114*, 10663–10665.

(169) Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sørensen, O. W., and Griesinger, C. (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR 4*, 301–306.

(170) Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson. 96*, 432–440.

(171) Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H 2O Saturation. *J. Magn. Reson. 109*, 129–133.

(172) Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ^{1}HN, ^{15}N, and ^{13}C′ chemical shifts in ^{15}N-^{13}C-labelled proteins. *J. Magn. Reson. 1969 97*, 213–217.

(173) Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J. Magn. Reson. B 103*, 203–216.

(174) Wittekind, M., and Mueller, L. (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J. Magn. Reson. 101*, 201–205.

(175) Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N-enriched proteins. *J. Biomol. NMR 3*, 185–204.

(176) Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J. Magn. Reson. B 103*, 203–216.

(177) Carlomagno, T., Maurer, M., Sattler, M., Schwendinger, M. G., Glaser, S. J., and Griesinger, C. (1996) PLUSH TACSY: Homonuclear planar TACSY with two-band selective shaped pulses applied to Cα,C′ transfer and Cβ,Caromatic correlations. *J. Biomol. NMR 8*, 161–170.

(178) Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR 3*, 225–231.

(179) Lyons, B. A., and Montelione, G. T. (1993) An HCCNH Triple-Resonance Experiment Using Carbon-13 Isotropic Mixing for Correlating Backbone Amide and Side-Chain Aliphatic Resonances in Isotopically Enriched Proteins. *J. Magn. Reson. 101*, 206–209.

(180) Clowes, R. T., Boucher, W., Hardman, C. H., Domaille, P. J., and Laue, E. D. (1993) A 4D HCC(CO)NNH experiment for the correlation of aliphatic side-chain and backbone resonances in ^{13}C/^{15}N-labelled proteins. *J. Biomol. NMR 3*, 349–354.

(181) Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in ^{13}C/ ^{15}N-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. *J. Magn. Reson. 101*, 114–119.

(182) Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Formankay, J. D. (1993) A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain ^{1}H and ^{13}C Correlations in H 2O Samples of Proteins. *J. Magn. Reson. 101*, 333–337.

(183) Pervushin, K. V., Wider, G., Riek, R., and Wüthrich, K. (1999) The 3D NOESY-[^{1}H,^{15}N,^{1}H]-ZQ-TROSY NMR experiment with diagonal peak suppression. *Proc. Natl. Acad. Sci. 96*, 9607–9612.

(184) Zhu, G., Kong, X. M., and Sze, K. H. (1999) Gradient and sensitivity enhancement of 2D TROSY with water flip-back, 3D NOESY-TROSY and TOCSY-TROSY experiments. *J. Biomol. NMR 13*, 77–81.

(185) Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR 4*, 845–858.

(186) Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson. 1969 98*, 207–216.

(187) Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson. 1969 98*, 207–216.

(188) Palmer, A. G. (2004) NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev. 104*, 3623–3640.

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

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

(191) Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protoc. 2*, 2728–2733.

(192) Huang, Y. J., Powers, R., and Montelione, G. T. (2005) Protein NMR Recall, Precision, and F-measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics. *J. Am. Chem. Soc. 127*, 1665–1674.

(193) Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477–486.

(194) Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins Struct. Funct. Bioinforma. 66*, 778–795.

(195) Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G.-W., Holst, M. J., McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci. 27*, 112–128.

(196) Vagenende, V., Yap, M. G. S., and Trout, B. L. (2009) Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol. *Biochemistry 48*, 11084–11096.

(197) Hautbergue, G. M., and Golovanov, A. P. (2008) Increasing the sensitivity of cryoprobe protein NMR experiments by using the sole low-conductivity arginine glutamate salt. *J. Magn. Reson. 191*, 335–339.

(198) Zhang, H., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR 25*, 173–195.

(199) Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry 31*, 1647–1651.

(200) Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR 4*, 171–180.

(201) Kralj Kunčič, M., Kogej, T., Drobne, D., and Gunde-Cimerman, N. (2010) Morphological Response of the Halophilic Fungal Genus Wallemia to High Salinity. *Appl. Environ. Microbiol. 76*, 329–337.

(202) Madern, D., Ebel, C., and Zaccai, G. (2000) Halophilic adaptation of enzymes. *Extremophiles 4*, 91–98.

(203) Siglioccolo, A., Paiardini, A., Piscitelli, M., and Pascarella, S. (2011) Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct. Biol. 11*, 50.

(204) Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol. 179*, 125–142.

(205) Gunning, A. P., De Groot, P. W. J., Visser, J., and Morris, V. J. (1998) Atomic Force Microscopy of a Hydrophobin Protein from the Edible MushroomAgaricus bisporus. *J. Colloid Interface Sci. 201*, 118–126.

(206) Jalili, N., and Laxminarayana, K. (2004) A review of atomic force microscopy imaging systems: application to molecular metrology and biological sciences. *Mechatronics 14*, 907–945.

(207) Seo, Y., and Jhe, W. (2007) Atomic force microscopy and spectroscopy. *Rep. Prog. Phys. 71*, 016101.

(208) Mackay, J. P., Matthews, J. M., Winefield, R. D., Mackay, L. G., Haverkamp, R. G., and Templeton, M. D. (2001) The Hydrophobin EAS Is Largely Unstructured in Solution and Functions by Forming Amyloid-Like Structures. *Structure 9*, 83–91.

(209) Groenning, M. (2009) Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils—current status. *J. Chem. Biol. 3*, 1–18.

(210) Biancalana, M., and Koide, S. (2010) Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochim. Biophys. Acta 1804*, 1405–1412.

(211) Nečas, D., and Klapetek, P. (2012) Gwyddion: an open-source software for SPM data analysis. *Open Phys. 10*.

(212) Takeda, M., Hallenga, K., Shigezane, M., Waelchli, M., Löhr, F., Markley, J. L., and Kainosho, M. (2011) Construction and performance of an NMR tube with a sample cavity formed within magnetic susceptibility-matched glass. *J. Magn. Reson. San Diego Calif 1997 209*, 167–173.

(213) Paananen, A., Vuorimaa, E., Torkkeli, M., Penttilä, M., Kauranen, M., Ikkala, O., Lemmetyinen, H., Serimaa, R., and Linder, M. B. (2003) Structural Hierarchy in Molecular Films of Two Class II Hydrophobins. *Biochemistry 42*, 5253–5258.

(214) Stajich, J. E., Wilke, S. K., Ahrén, D., Au, C. H., Birren, B. W., Borodovsky, M., Burns, C., Canbäck, B., Casselton, L. A., Cheng, C. K., *et al.* (2010) Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom Coprinopsis cinerea (Coprinus cinereus). *Proc. Natl. Acad. Sci. U. S. A. 107*, 11889–11894.

(215) Steinhorn, I. (1983) In situ salt precipitation at the Dead Sea. *Limnol. Oceanogr. 28*, 580–583.

(216) Chang, H. J., Lee, M., and Na, S. (2019) Investigation of the role hydrophobin monomer loops using hybrid models via molecular dynamics simulation. *Colloids Surf. B Biointerfaces 173*, 128–138.

(217) Oren, A. (2006) Life at High Salt Concentrations, in *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry* (Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., and Stackebrandt, E., Eds.), pp 263–282. Springer New York, New York, NY.

(218) Arai, M., Dyson, H. J., and Wright, P. E. (2010) Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. *FEBS Lett. 584*, 4500–4504.

(219) Cox, A. R., Cagnol, F., Russell, A. B., and Izzard, M. J. (2007) Surface Properties of Class II Hydrophobins from Trichoderma reesei and Influence on Bubble Stability. *Langmuir 23*, 7995–8002.

(220) Szilvay, G. R., Kisko, K., Serimaa, R., and Linder, M. B. (2007) The relation between solution association and surface activity of the hydrophobin HFBI from Trichoderma reesei. *FEBS Lett. 581*, 2721–2726.

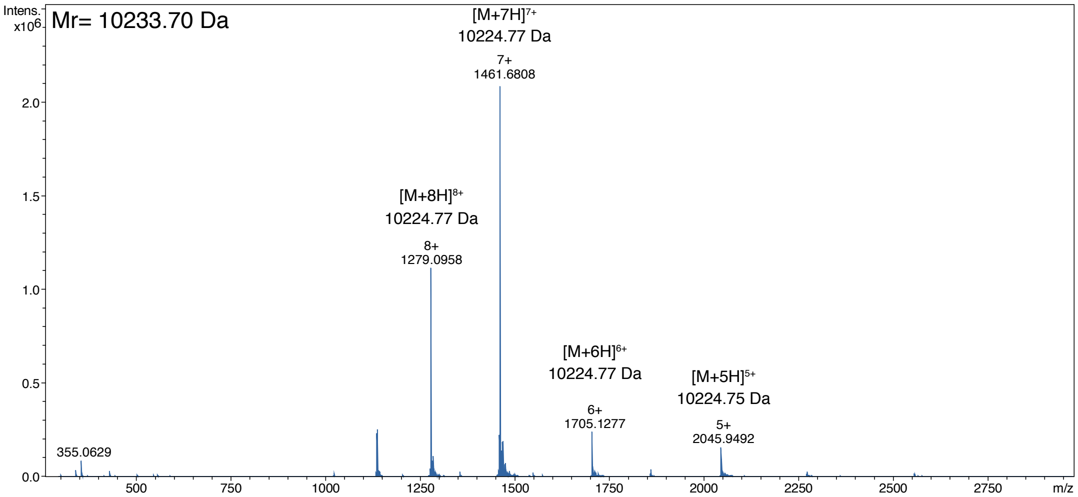
(221) Kallio, J. M., Linder, M. B., and Rouvinen, J. (2007) Crystal Structures of Hydrophobin HFBII in the Presence of Detergent Implicate the Formation of Fibrils and Monolayer Films. *J. Biol. Chem. 282*, 28733–28739.

Appendix A. Mass Spectrometry Data

SC16

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.0 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	200 °C
Scan Begin	300 m/z	Set End Plate Offset	-500 V	Set Dry Gas	5.0 l/min
Scan End	3000 m/z	Set Charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C



SL1

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.0 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	200 °C
Scan Begin	300 m/z	Set End Plate Offset	-500 V	Set Dry Gas	5.0 l/min
Scan End	3000 m/z	Set Charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C

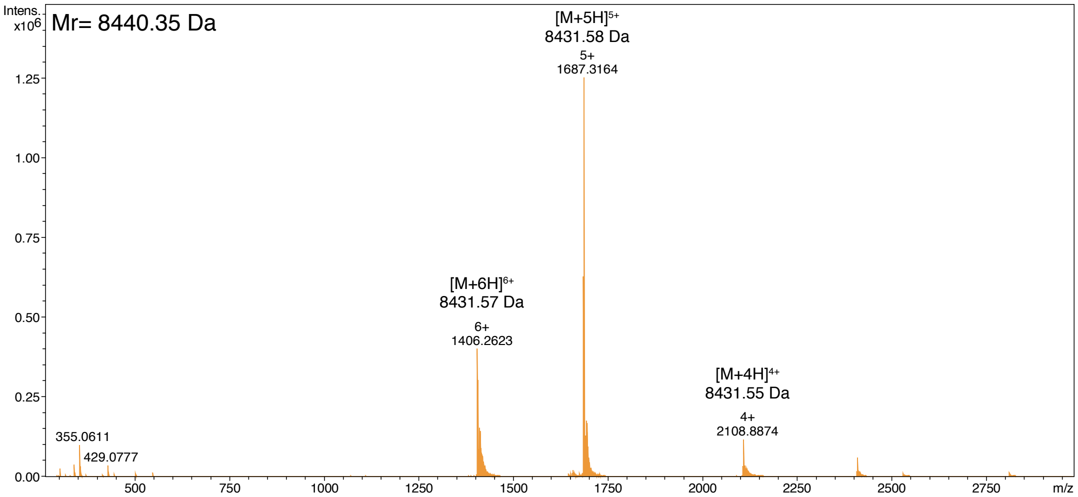
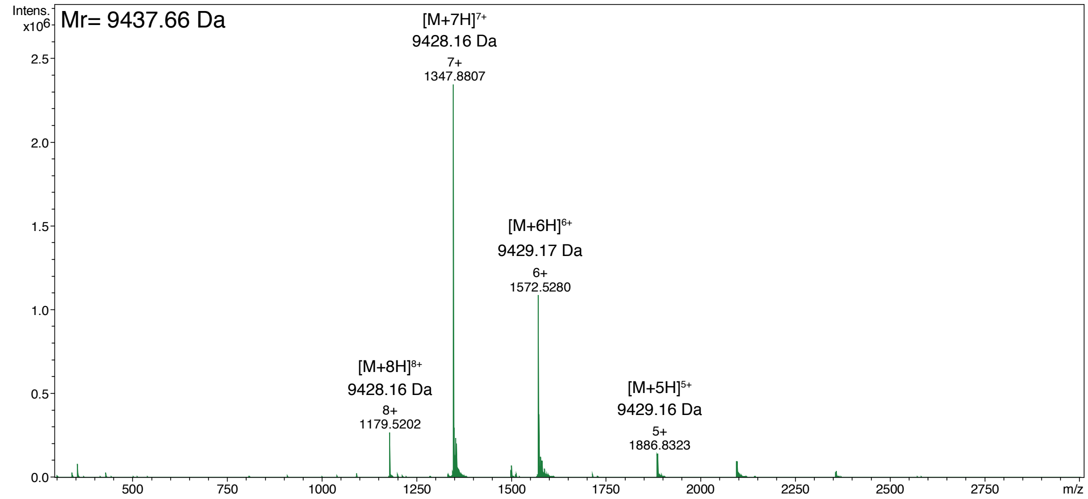


Figure A.1. Mass spectra of purified SC16 and SL1.
Masses SC16 (top) and SL1 (bottom) were confirmed after RP-HPLC purification using positive mode electrospray ionization mass spectrometry (ESI+ MS). (Dalhousie Mass Spectrometry Laboratory, Halifax, NS).

WI1

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.0 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	200 °C
Scan Begin	300 m/z	Set End Plate Offset	-500 V	Set Dry Gas	5.0 l/min
Scan End	3000 m/z	Set Charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C



PC1

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.0 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	200 °C
Scan Begin	300 m/z	Set End Plate Offset	-500 V	Set Dry Gas	5.0 l/min
Scan End	3000 m/z	Set Charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C

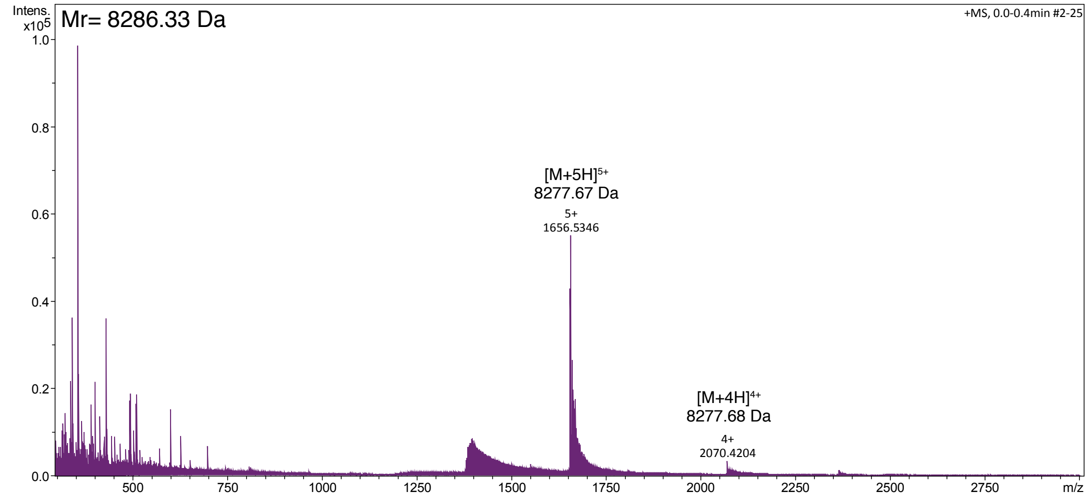


Figure A.2. Mass spectra of purified WI1 and PC1.

Masses WI1 (top) and PC1 (bottom) were confirmed RP-HPLC purification using positive mode electrospray ionization mass spectrometry (ESI+ MS) (Dalhousie Mass Spectrometry Laboratory, Halifax, NS).

Appendix B. Cysteine Chemical Shift Values and ^1H-^{13}C NOESY Contacts

Table B.1. Cβ chemical shift per cysteine residue in class IB hydrophobins

	Cβ Chemical Shift (ppm)		
Residue	PC1	SL1	WI1
C1	38.46	36.87	37.53
C2	38.79	38.82	38.80
C3	46.62	46.63	50.06
C4	45.33	43.98	47.68
C5	37.67	38.19	38.14
C6	39.62	38.18	38.37
C7	46.84	48.88	45.36
C8	45.37	45.62	45.79

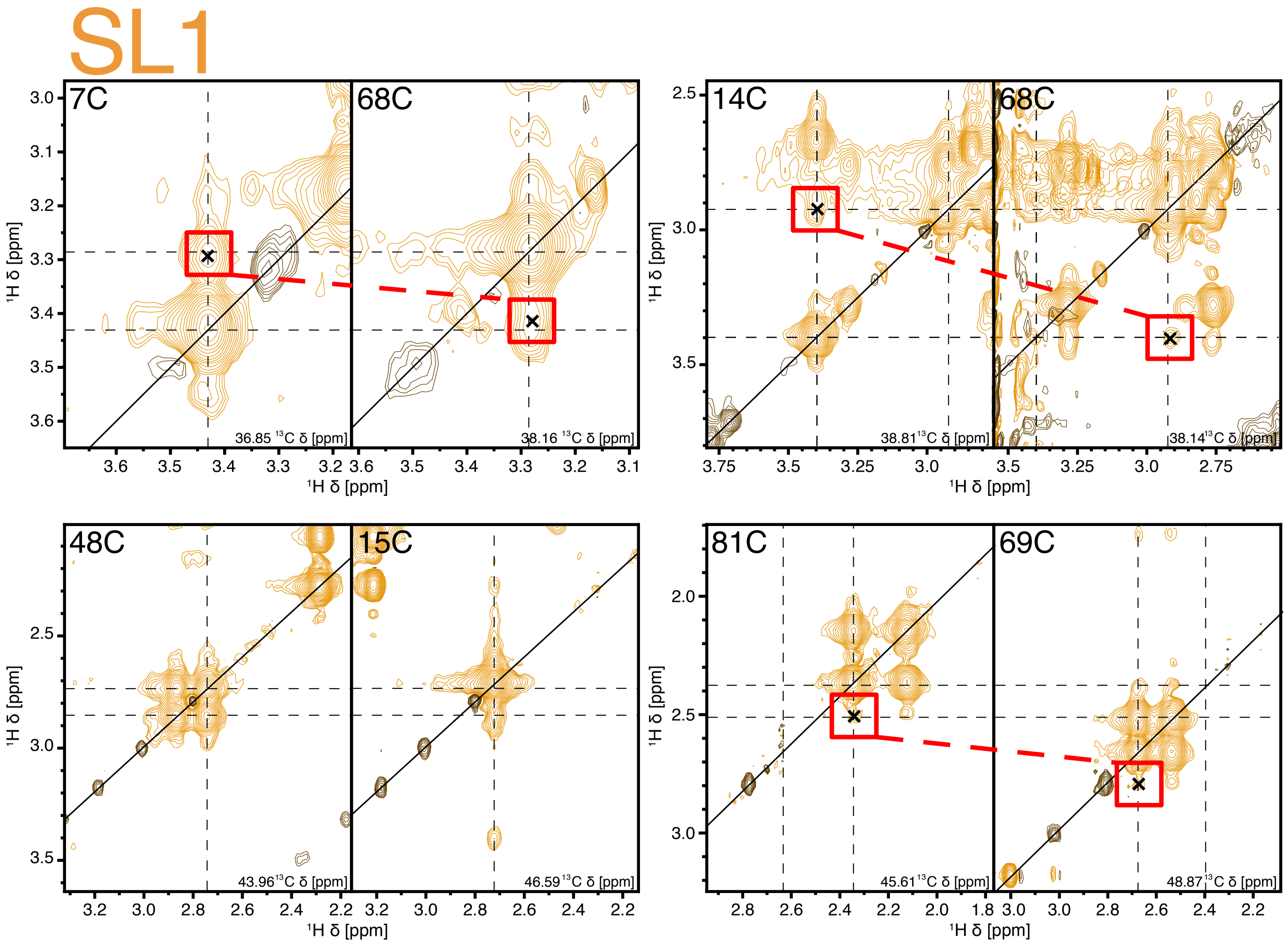


Figure B.1. Disulphide bridging in SL1 observed by $C^{H\beta}$ NOE contacts in ^{1}H-^{13}C NOESY experiments.
CNOESY experiments show contact between cysteine residues indicating disulphide bond formation. Expected resonances in each ^{1}H dimension indicated by dashed lines. Cross peaks corresponding to contact between two cysteine Hβ nuclei highlighted in red.

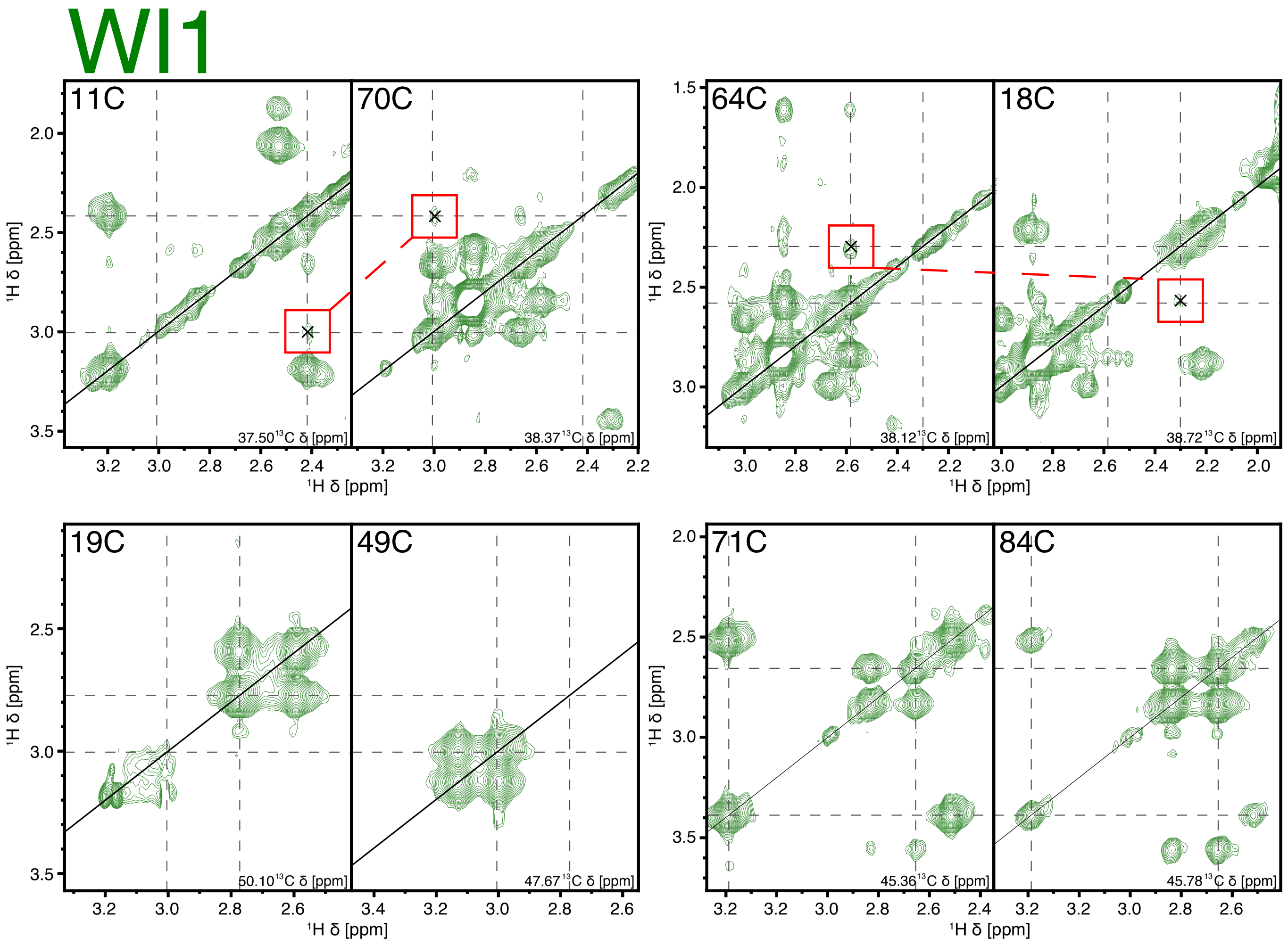


Figure B.2. Disulphide bridging in WI1 observed by $C^{H\beta}$ NOE contacts ^{1}H-^{13}C NOESY experiments.
CNOESY experiments show contact between cysteine residues indicating disulphide bond formation. Expected resonances in each ^{1}H dimension indicated by dashed lines. Cross peaks corresponding to contact between two cysteine Hβ nuclei highlighted in red.

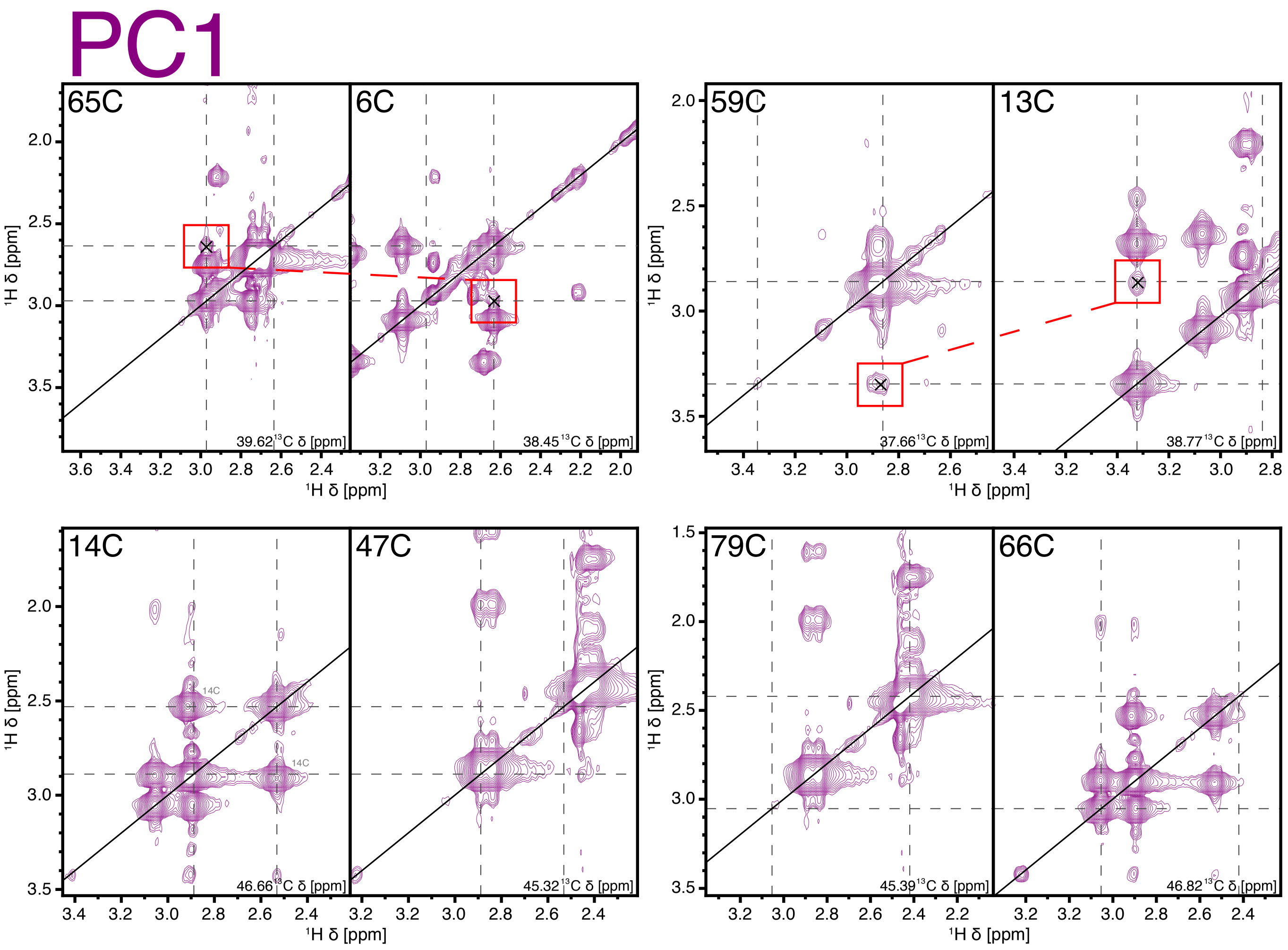


Figure B.3. Disulphide bridging in PC1 observed by $C^{H\beta}$ NOE contacts ^{1}H-^{13}C NOESY experiments.
CNOESY experiments show contact between cysteine residues indicating disulphide bond formation. Expected resonances in each ^{1}H dimension indicated by dashed lines. Cross peaks corresponding to contact between two cysteine Hβ nuclei highlighted in r

Appendix C. ThT Fluorescence Assay Data

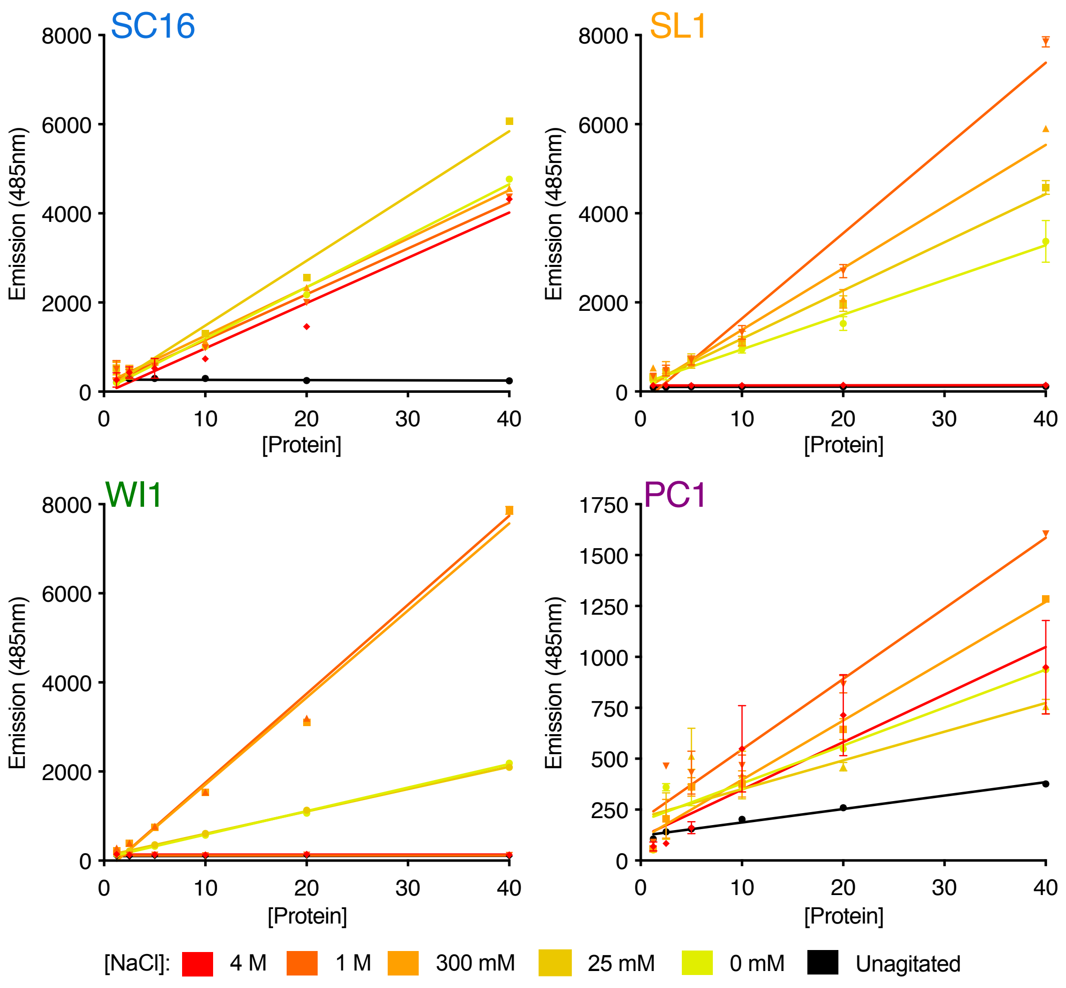


Figure C.1. ThT fluorescence of SC16, SL1, WI1 and PC1 is impacted by salt concentration.

Samples of SC16, SL1, WI1 and PC1 were agitated end over end at approximately 60 rpm overnight in 20 mM MES pH 6.5 and varying concentration of NaCl. Unagitated sample left at room temperature overnight in 20 mM MES pH 6.5, 25 mM NaCl. ThT emission was measured (N = 3) and plotted with standard deviation indicated (error bars not visible for all points due to scale).

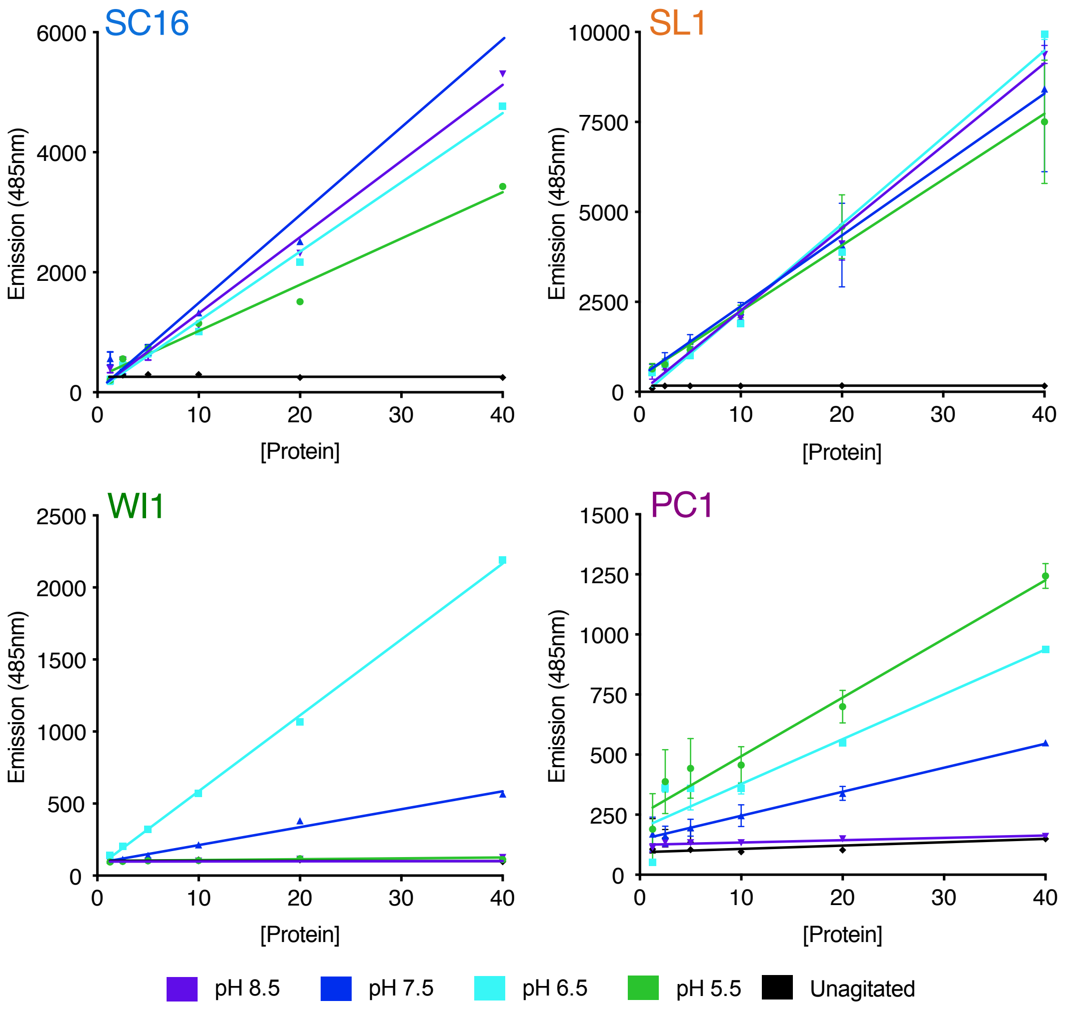


Figure C.2. ThT fluorescence of SC16, SL1, WI1 and PC1 is affected by sample pH. Samples of SC16, SL1, WI1 and PC1 were agitated end over end at approximately 60 rpm overnight in either 20 mM MES pH 5.5 or pH 6.5, 20 mM phosphate 7.5, or 20 mM Tris pH 8.5. Unagitated sample left at room temperature overnight in 20 mM MES pH 6.5. ThT emission was measured (N = 3) and plotted with standard deviation indicated (error bars not visible for all points due to scale).

Table C.1. Protein concentration dependant ThT fluorescence in varying salt conditions

Sample	SC16		SL1		WI1		PC1	
	I/[Protein]	R^2	I/[Protein]	R^2	I/[Protein]	R^2	I/[Protein]	R^2
0 mM	115.60 ± 2.19	0.992	77.94 ± 3.33	0.961	52.70 ± 0.38	0.999	18.64 ± 1.50	0.875
25 mM	145.40 ± 3.90	0.985	108.30 ± 2.87	0.985	50.15 ± 0.34	0.999	14.10 ± 2.17	0.668
300 mM	108.70 ± 1.99	0.993	138.40 ± 6.04	0.956	195.60 ± 4.60	0.988	29.17 ± 1.23	0.964
1 M NaCl	102.80 ± 2.73	0.985	191.60 ± 7.39	0.986	199.70 ± 4.76	0.988	34.68 ± 1.80	0.944
2 M NaCl					-0.02 ± 0.08	0.002		
4M NaCl	101.70 ± 4.92	0.951	0.08 ± 0.08	0.043	-0.02 ± 0.12	0.001	23.35 ± 2.81	0.579
Unagitated	-0.52 ± 0.52	0.044	0.36 ± 0.08	0.515	0.11 ± 0.07	0.100	6.61 ± 0.20	0.980

Values determined by linear regression of all data points in Figure C.1 (1.25 µM – 40 µM protein; N=3 measurements for each condition) with the resulting slope shown plus or minus the standard deviation of the fit. Low R^2 values are observed in samples with low overall intensity.

Table C.2. Protein concentration dependant ThT fluorescence in varying pH conditions

Sample	SC16		SL1		WI1		PC1	
	I/[Protein]	R^2	I/[Protein]	R^2	I/[Protein]	R^2	I/[Protein]	R^2
pH 5.5	77.00 ± 5.83	0.978	182.90 ± 11.54	0.920	0.43 ± 0.10	0.434	24.39 ± 1.74	0.899
pH 6.5	115.60 ± 4.72	0.993	241.00 ± 7.13	0.981	52.70 ± 0.38	0.999	18.64 ± 1.50	0.875
pH 7.5	146.40 ± 9.71	0.983	198.80 ± 14.62	0.894	12.34 ± 0.35	0.983	10.00 ± 0.59	0.923
pH 8.5	126.80 ± 6.89	0.988	229.10 ± 5.08	0.990	0.79 ± 0.04	0.947	0.96 ± 0.20	0.500
Unagitated	-0.52 ± 1.12	0.052	0.02 ± 0.06	0.005	-0.08 ± 0.08	0.043	1.18 ± 0.49	0.205

Values determined by linear regression of all data points in Figure C.2 (1.25 µM – 40 µM protein; N=3 measurements for each condition) with the resulting slope shown plus or minus the standard deviation of the fit. Low R^2 values are observed in samples with low overall intensity.