



Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

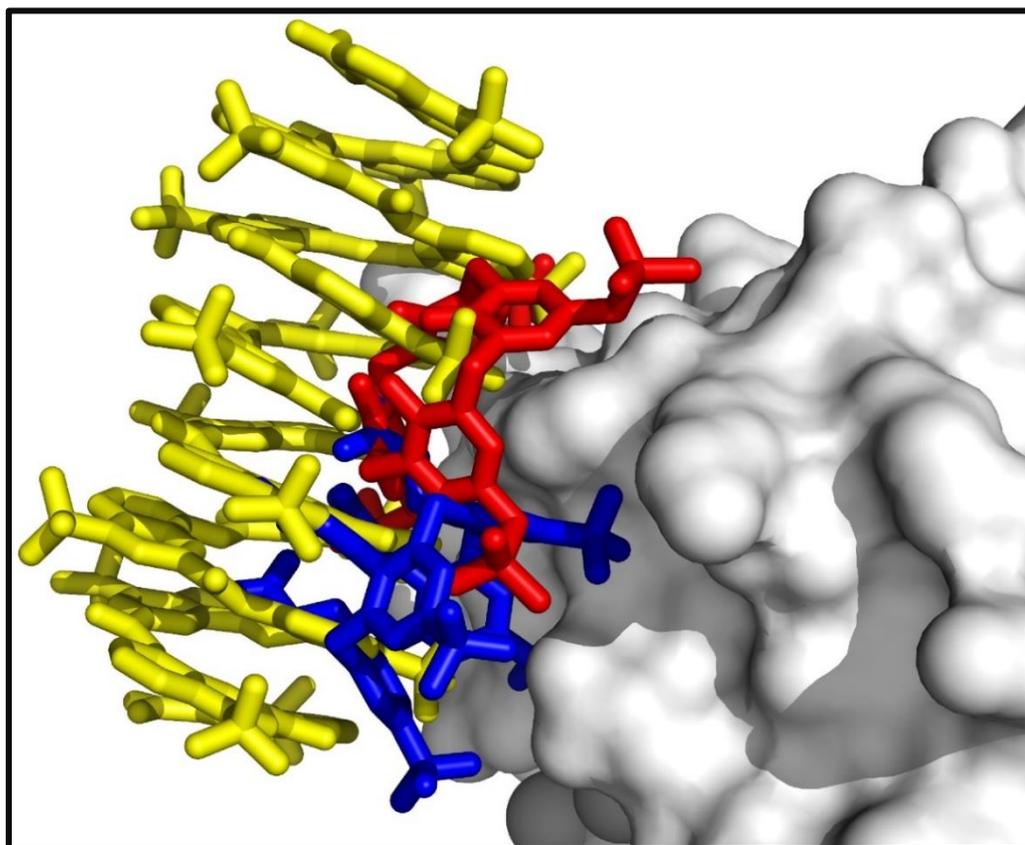
|                  |   |
|------------------|---|
| Title            | Supramolecular 'Glues' for protein recognition and assembly                       |
| Author(s)        | Alex, Jimi  |
| Publication Date | 2019-12-06  |
| Publisher        | NUI Galway  |
| Item record      | <a href="http://hdl.handle.net/10379/15631">http://hdl.handle.net/10379/15631</a> |

Downloaded 2024-04-27T03:12:03Z

Some rights reserved. For more information, please see the item record link above.



**Supramolecular 'Glues' for  
Protein Recognition and Assembly**



Jimi Marin Alex, M. Pharm

School of Chemistry  
National University of Ireland, Galway

PhD, 2019





# Supramolecular 'Glues' for Protein Recognition and Assembly

PhD Thesis

This thesis was prepared at the School of Chemistry, National University of Ireland, Galway, from September 2015 to September 2019.

I declare that the work included in this thesis is entirely my own work and has not been previously submitted for a degree to this or any other academic institution.

---

**Cover Art:** A cationic patch on cytochrome *c* can be recognized by different anionic receptors, calixarenes (red and blue) or a foldamer (yellow).

Jimi Marin Alex

Born in Kerala, India 1987

Supervisor: Prof. Peter B. Crowley

External Examiner: Prof. Tanja Weil (Max Planck Institute for Polymer Research)

Internal Examiner: Dr. Constantina Papatriantafyllopoulou

*Viva voce*: 2<sup>nd</sup> December 2019





## Summary of Contents

---

Supramolecules such as calixarenes, cucurbiturils and foldamers are rapidly emerging as tools for protein assembly. This thesis builds on past successes with some of these scaffolds. Previously, sulfonato-calix[4]arene was shown to recognize and assemble the lysine-rich cytochrome *c* (cytc).<sup>1</sup> Here, we tested calix[4]arenes with different upper or lower rims. A phosphonomethyl derivative demonstrated enhanced selectivity for lysine<sup>2</sup> whereas a lower rim oxomethylcarboxylate yielded a honeycomb network of cytc.<sup>3</sup> A comparison of these structures highlighted how different substituents alter the recognition and the assembly-inducing behaviour. In related experiments a co-crystal structure revealed a surprising assembly of sulfonato-calix[8]arene and a cationic porphyrin.<sup>4</sup>

Inspired by the interactions of calixarenes with cytc, the sulfonato-calix[4]arene (n= 4, 6, 8) series was co-crystallized with another cationic protein – the small antifungal protein from *Penicillium* (PAF). It was an exciting and a rewarding project to solve the crystal structure of PAF for the first time.<sup>5</sup> X-ray and solution-state studies enabled us to compare the influence of increasing calixarene size and charge on recognition and assembly.

Another interesting project was to elucidate crystallographically the protein binding of a tether-free foldamer.<sup>6</sup> The complex of a quinoline foldamer and cytc yielded a remarkable biohybrid assembly with chiral resolution of the foldamer helix handedness. NMR and CD experiments suggested differences in the solution state recognition processes.

### References

- (1) McGovern *et al.* *Nat. Chem.* **2012**, *4*, 527.
- (2) Chapter 1, Alex *et al.* *Cryst. Growth Des.* **2018**, *18*, 2467.
- (3) Chapter 2, Alex *et al.* *Org. Biomol. Chem.* (Accepted).
- (4) Chapter 3, Alex *et al.* *CrystEngComm.* **2019**. *In print*
- (5) Chapter 4, Alex *et al.* *IUCrJ.* **2019**, *6*, 238.
- (6) Chapter 5, Alex *et al.* *Chem. Commun.* **2019**, *55*, 11087.



# Contents

---

|                         |   |     |
|-------------------------|---|-----|
| Introduction            |   | 1   |
| Chapter 1               | Phosphonated calixarene as a “molecular glue” for protein crystallization           | 27  |
| Chapter 2               | Protein frameworks <i>via</i> co-crystallization with an octa-anionic calix[4]arene | 45  |
| Chapter 3               | Supramolecular stacking in a high <i>Z'</i> calix[8]arene – porphyrin assembly      | 61  |
| Chapter 4               | Calixarene-mediated assembly of a small antifungal protein                          | 75  |
| Chapter 5               | Crystal structure of a protein – aromatic foldamer composite                        | 101 |
| Discussion              |   | 127 |
| Bibliography            |   | 138 |
| Acknowledgements        |   | 173 |
| <i>Curriculum Vitae</i> |   | 176 |
| Appendix                | Availability of molecular glues   | 182 |
|                         | Calixar kit (Molecular Dimensions)  | 183 |
|                         | Preliminary NMR of protein – dye complexes  | 184 |



# **Introduction**

---

Proteins are ubiquitous and the most versatile building block across all life forms.<sup>1,2</sup> Their role in modulating dynamic cellular processes, as well as maintaining functional and structural integrity, is imperative for the living cell.<sup>3,4</sup> Nature employs a bottom-up approach to constitute proteins (transcription and translation) in which usually the primary structure determines the tertiary structure,<sup>5,6</sup> and higher order quaternary structures occur *via* homo- (collagen) or hetero- (hemoglobin) oligomeric assembly.<sup>7,8</sup> Non-covalent interactions between the monomers serve to stabilize these structures.<sup>2,9</sup> Compared to a single protein unit, self-assembled structures exhibit improved stability, mechanical strength, and better regulation of function / biological activity.<sup>2,3</sup> Thus, strategies to achieve a similar level of complexity and sophistication in fabricated protein nanostructures has been a major research focus of recent decades.

Owing to their biocompatibility and safety, proteins are attractive candidates for biomedical applications.<sup>10-12</sup> Various protein-based materials have been designed that rely on protein engineering / fusions,<sup>13-16</sup> metal-coordination<sup>17</sup> or coiled-coil subunits<sup>18</sup> that yield unique architectures. The use of supramolecular ligands to fabricate protein assemblies is a rapidly emerging field. The ease of synthesis and availability of a range of customizations offer vast possibilities in structural design that can be explored with different proteins. For example, synthetic receptors including calixarenes,<sup>19,20</sup> cucurbiturils,<sup>21</sup> pillararenes,<sup>22</sup> and molecular tweezers<sup>23</sup> as well as ligands such as foldamers<sup>24</sup> are useful building blocks to direct different types of protein assembly.

### **Protein structure determination**

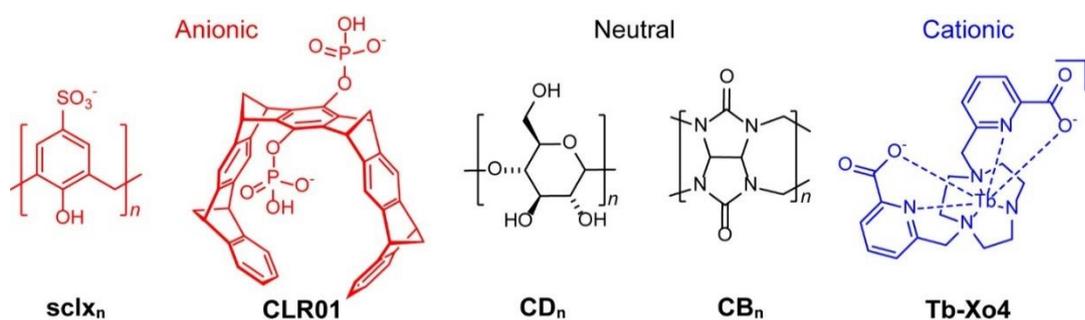
Advances in bioinformatics, biotechnology and analytical tools have enhanced the understanding of protein structures, folding and interactions.<sup>25,26</sup> However, protein structure determination remains pivotal to elucidate how proteins mediate their function.<sup>5,27</sup> X-ray crystallography,<sup>28</sup> solution- and solid-state NMR,<sup>29,30</sup> as well

as cryo-Electron Microscopy<sup>31</sup> have significantly promoted structure determination of numerous proteins. X-ray crystallography alone accounts for ~90 % of the structures in the Protein Data Bank (PDB).<sup>5</sup> However, the stochasticity of obtaining diffraction-quality crystals<sup>27</sup> renders the crystallization process a major bottleneck to structure determination.<sup>32,33</sup> Numerous strategies have been developed to overcome the crystallization bottleneck. Protein truncation,<sup>34</sup> lysine mutagenesis / methylation,<sup>35</sup> protein fusion,<sup>36</sup> and chaperones<sup>37</sup> have improved crystallization of various recalcitrant proteins. High-throughput gene synthesis and automated protein purification and crystallization (structural genomics) has facilitated the discovery of 'crystallisable' protein constructs.<sup>38-41</sup> Sparse matrix screens have made it possible to screen hundreds of crystallization conditions within a short time span.<sup>42</sup> Classic seeding or microseed matrix screening (MMS) technique have also been used to promote / optimize crystal growth.<sup>43,44</sup> In parallel to these approaches, the use of additives or small molecules ("silver bullets") in the crystallization mix has had some success in protein crystallization.<sup>45-47</sup>

### **'Molecular glues'**

Crystallization additives are typically small compounds bearing multiple polar / charged groups, which can interact (hydrogen bonds / salt bridges) with two or more neighbouring proteins.<sup>45,46</sup> In recent years, this approach was extended to synthetic receptors / ligands that are significantly larger than the original additives. For example, calixarenes,<sup>19,20,48-57</sup> cucurbiturils,<sup>21,58,59</sup> molecular tweezers<sup>23,60,61</sup> and others have been investigated (Figure 1; Table 1). Their structural framework (hydrophobic cavity and polar substituents) allows the receptors to establish a wide range of non-covalent interactions (protein recognition) and intermolecular contact formation (assembly / crystallization) with the protein. These ligand-bound regions on the protein surface behave as 'sticky patches' to form interfaces with the surrounding proteins and dictate assembly, thus giving rise to the name 'molecular glue'.<sup>62-64</sup> These 'glues' can be employed to explore new host-guest chemistries,<sup>21,48,61</sup> recognition modes,<sup>65</sup> supramolecular assemblies,<sup>24</sup> porous

frameworks,<sup>19</sup> or to improve resolution.<sup>66</sup> Further, functionalization of the scaffold offers the benefit of tailored interactions with the protein of interest. These ligands can be easily synthesized and/or are commercially available (Table A1; Appendix). While the differences in structural framework, conformation, charge, size, and solubility diversify these compounds, their ability to promote protein-protein interaction / crystal growth unifies them.



**Figure 1.** Representative molecular glues highlighting the diversity of chemical structures.

Here, 15 categories of glues and their representative examples are discussed briefly (Table 1). The focus is on X-ray crystal structures in the Protein Data Bank (PDB) where the ligand was co-crystallized with the protein. Protein – ligand complexes obtained by soaking have been omitted.

**Table 1:** X-ray crystal structures of molecular glue – protein assemblies.

| Ligand and MW (Da)                    | Ligand code | Protein                   | PDB code | Ref |
|---------------------------------------|-------------|---------------------------|----------|-----|
| <b>calixarenes</b>                    |             |                           |          |     |
| sulfonato-calix[4]arene (745)         | T3Y         | cytc                      | 3TYI     | 48  |
|                                       |             | lysozyme                  | 4PRQ     | 49  |
|                                       |             | lysozyme-Kme <sub>2</sub> | 4N0J     | 50  |
|                                       |             | PAF                       | 6HA4     | 55  |
| bromo-sulfonato-calix[4]arene (743)   | 6VB         | cytc                      | 5LFT     | 52  |
| phenyl-sulfonato-calix[4]arene (741)  | 6VJ         | cytc                      | 5KPF     | 52  |
| phosphonatomethyl-calix[4]arene (800) | 8TE         | cytc                      | 5NCV     | 53  |

|   |     |                                  |      |     |
|---|-----|----------------------------------|------|-----|
| Oxomethylcarboxylate-sulfonato-calix[4]arene (1076) | LVT | cytc                             | 6SUY | 57  |
| phosphonato-calix[6]arene (1117)                    | 7AZ | cytc                             | 5LYC | 51  |
| sulfonato-calix[6]arene (1117)                      | FWQ | PAF                              | 6HAH | 55  |
|   |     | cytc                             | 6RGI | 56  |
| sulfonato-calix[8]arene (1489)                      | EVB | cytc                             | 6GD9 | 19  |
|   |     | PAF                              | 6HAJ | 55  |
| sulfonato-calix[4]arene mono-PEG (1508)             | B4T | cytc                             | 6EGY | 54  |
| sulfonato-calix[4]arene di-PEG (1918)               | B4X | cytc                             | 6EGZ | 54  |
| <b>18-crown-6</b> (264)                             | O4B | Pin1                             | 3WHO | 75  |
|   |     | hemoglobin                       | 3WHM | 75  |
|   |     | DNA mimic protein DMP19          | 3WUR | 75  |
|   |     | farnesyl diphosphate synthase    | 5HXP | 76  |
|   |     | PI3-kinase C2 domain             | 6BTZ | 77  |
|   |     | biofilm scaffolding protein RbmA | 4BEI | 78  |
| <b>cryptophane A</b> (1423)                         | 1CR | carbonic anhydrase II            | 3CYU | 82  |
| <b>cucurbituril</b>                                 |     |                                  |      |     |
| cucurbit[7]uril (1163)                              | QQ7 | insulin                          | 3Q6E | 58  |
|   |     | lectin RSL-Kme <sub>2</sub>      | 6F7W | 21  |
| cucurbit[8]uril (1329)                              | C8L | adapter 14-3-3 and peptide       | 5N10 | 59  |
| <b>cyclodextrins</b>                                |     |                                  |      |     |
| $\alpha$ -cyclodextrin (973)                        | ACX | starch binding SusE              | 4FEM | 93  |
| $\beta$ -cyclodextrin (1135)                        | BCD | starch binding SusD              | 3CK7 | 92  |
| <b>foldamers</b>                                    |     |                                  |      |     |
| quinoline oligoamide (1337)                         | Q4I | carbonic anhydrase II            | 4LP6 | 100 |
|   |     |                                  | 5LVS | 101 |
|   |     |                                  | 6Q9T | 102 |
| sulfonated quinoline oligoamide (2060)              | L0T | cytc                             | 6S8Y | 24  |
| <b>fullerene</b>                                    |     |                                  |      |     |
| C60 pyrrolidine tris-acid                           | 60C | designed coiled coil             | 5ET3 | 103 |
| <b>lanthanide complexes</b>                         |     |                                  |      |     |
| Eu-dipicolinate                                     | PDC | lysozyme                         | 2PC2 | 104 |
|   |     | thaumatin                        | 2PE7 | 105 |
| Tb-crystallophore (556)                             | 7MT | bacteriophage tail protein pb9   | 5MF2 | 108 |
|   |     | diiron flavoprotein (FprA)       | 6FRN | 66  |
|   |     | Lysozyme                         | 6F2I | 66  |
|   |     | protease 1                       | 6F2F | 66  |
| <b>molecular tweezers</b> (727)                     | 9SZ | adapter 14-3-3                   | 5OEG | 60  |
|   |     | adapter 14-3-3 and peptide       | 5M36 | 61  |
| <b>naphthalene sulfonates</b>                       |     |                                  |      |     |
| orange G (408)                                      | ORA | amyloid- $\beta$ peptide         | 3OVJ | 110 |
| ponceau S (673)                                     | QV7 | pyruvate kinase                  | 3QV7 | 112 |
| suramin (1297)                                      | SVR | myotoxin II                      | 1Y4L | 113 |

## Introduction

|                                      |     |                                 |      |     |
|--------------------------------------|-----|---------------------------------|------|-----|
|                                      |     | sirtuin homolog 5               | 2NYR | 114 |
|                                      |     | RNA polymerase (RdRp)           | 3URO | 115 |
|                                      |     | ecarpholin S                    | 3BJW | 116 |
|                                      |     | chromodomain CBX7               | 4X3U | 117 |
|                                      |     | myotoxin I                      | 6CE2 | 118 |
| <b>polyoxometalates</b>              |     |                                 |      |     |
| 12-polytungstate (2848)              | E43 | hydrolase NTPDase1              | 4BVO | 123 |
| Zr(IV)-substituted Keggin (2769)     | ZKG | lysozyme                        | 4XYX | 125 |
| hexatungstotellurate (1615)          | TEW | mushroom tyrosinase             | 4OUA | 129 |
|                                      |     | aurone synthase                 | 4Z12 | 130 |
|                                      |     | HSP70 nucleotide binding domain | 6G3R | 131 |
| <b>polyphosphates</b>                |     |                                 |      |     |
| inositol-hexakisphosphate (660)      | IHP | HIV-1 Gag protein               | 6BHR | 134 |
|                                      |     | PH-TH module of Btk             | 4Y94 | 137 |
|                                      |     | arrestin-3                      | 5TV1 | 138 |
|                                      |     | DEAD-box ATPase                 | 3RRN | 139 |
| <b>porphyrins</b>                    |     |                                 |      |     |
| tetrasulfonato-phenylporphyrin (939) | SFP | concanavalin A                  | 1JN2 | 142 |
|                                      |     | jacalin                         | 1PXD | 143 |
|                                      |     | peanut lectin                   | 1R1R | 144 |
| uroporphyrinogen III (831)           | UP3 | uroporphyrinogen III synthase   | 3D8N | 145 |
| <b>pyrene derivatives</b>            |     |                                 |      |     |
| pyrene-tetrasulfonic acid (522)      | PTK | pyruvate kinase                 | 3IS4 | 148 |
| methoxypyrene-trisulfonic acid (472) | 2M9 | antibody (Fab)                  | 4NJ9 | 159 |
| pyrenyl-acetamide (273)              | JPY | adenylate kinase                | 3X2S | 150 |
| <b>rhodamine-sugar conjugates</b>    |     |                                 |      |     |
| rhodamine-mannose conjugate (820)    | R3M | Concanavalin A                  | 4P9W | 151 |

**Calix[n]arenes**<sup>67</sup> are well-established as supramolecular building blocks.<sup>68,69</sup> The possibility of customizing these molecules makes them excellent candidates for molecular recognition.<sup>70-74</sup> In particular, the anionic water soluble sulfonatocalix[n]arenes (**sclx<sub>n</sub>**)<sup>19,20,48-50,52,54-56</sup> have widely demonstrated their potential to recognize lysine and arginines on the protein surfaces and induce assembly. The first complex of **sclx<sub>4</sub>** with cytc highlighted the participation of the ligand in numerous non-covalent interactions at the protein – protein interface.<sup>48</sup> Different calixarene functionalizations have been explored since then to determine the effect of substituents on recognition and assembly. For example, replacement of one sulfonate group with bromo or a phenyl group caused no significant changes in

recognition.<sup>52</sup> On the other hand, a new binding site on cytc was identified when the upper rim was functionalized with phosphonomethyl groups (**Chapter 1**)<sup>53</sup> or the **sclx<sub>4</sub>** lower rim was PEGylated.<sup>54</sup> Compared to these calix[4]arenes, an octa-anionic derivative of **sclx<sub>4</sub>** with oxomethylcarboxylates at the lower rim induced a honeycomb network of cytc (~75 % solvent content; **Chapter 2**)<sup>57</sup> which resembled that generated by **sclx<sub>6</sub>** in complex with cytc (~65 % solvent content).<sup>56</sup> In the latter case, the crystal structure revealed the ability of **sclx<sub>6</sub>** to stabilize a partially unfolded loop. This feature suggests a role for calixarenes as a trap to capture unfolded or intrinsically disordered protein. In contrast to **sclx<sub>6</sub>**, phosphonato-calix[6]arene (**pclx<sub>6</sub>**) facilitated a C<sub>2</sub> symmetric dimer of cytc.<sup>51</sup> Compared to these ligands the conformationally flexible **sclx<sub>8</sub>** behaved quite distinctly with cytc (Figure 2D and 3D).<sup>19</sup> It was able to auto-regulate the assembly and disassembly of a tetramer without the requirement of a competitive guest molecule. Remarkably, **sclx<sub>8</sub>** was able to induce a high-porosity crystalline framework (~85 % solvent content). Furthermore, floppiness of the ligand enables it to mould to the protein surface and form large interfaces (~550 Å<sup>2</sup>) as a demonstration of its ‘molecular glue’ property. Also, it was interesting to learn that these porous cytc – **sclx<sub>8</sub>** assemblies could be modulated by the presence of an effector ligand.<sup>20</sup>

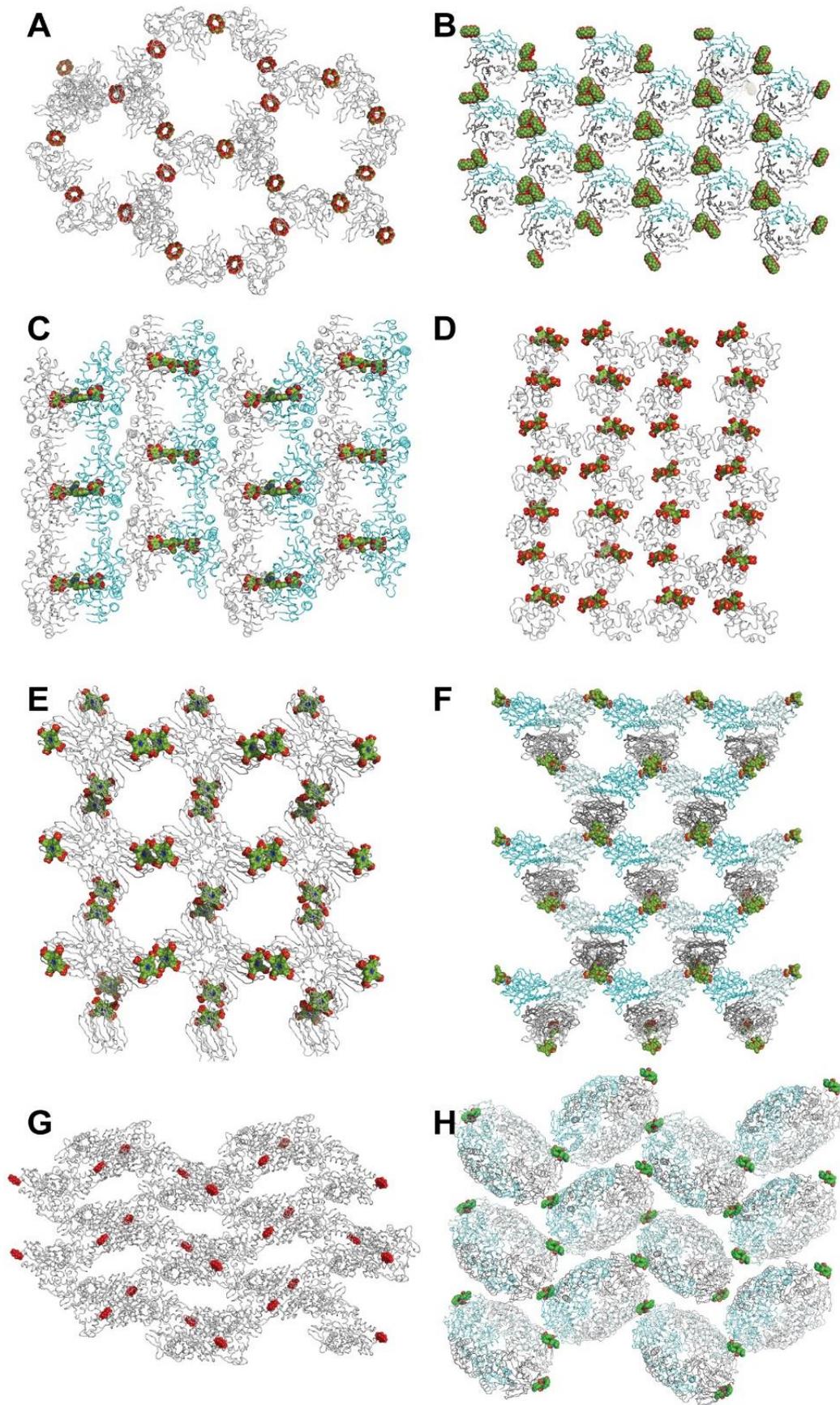
In addition to cytc, **sclx<sub>4</sub>** mediated recognition and assembly of other cationic proteins including lysozyme (tetramer)<sup>49</sup> and dimethylated lysozyme (dimer).<sup>50</sup> Recently, **sclx<sub>4</sub>**, **sclx<sub>6</sub>** and **sclx<sub>8</sub>** demonstrated their ability to facilitate the crystal packing in a small antifungal protein from *Penicillium* (PAF) for which no crystal structure was reported (**Chapter 4**).<sup>55</sup> Pure PAF failed to yield diffraction-quality crystals. These crystal structures highlight the role of substituents, charge, and ring size in influencing protein recognition and assembly.

**Crown ethers** are small cyclic polyethers used as an additive to enhance protein rigidity, and stability. For example, **18-crown-6** has been employed for promoting the crystal growth of a wide range of proteins.<sup>75-78</sup> Similar compounds such as **cyclen**<sup>79</sup>

(aza analogue of 12-crown-4) and **cyclam**<sup>80</sup> have also been utilized as crystallization agents.

**Cryptophanes** as 'containers' of xenon ion is used for biosensing.<sup>81</sup> The complex of this ligand with human carbonic anhydrase (HCA) illustrates its ability to target the active site of the protein. Also, its role in crystal packing was evident by its participation in establishing crystal contact.<sup>82</sup>

**Cucurbit[n]urils (CBn)** are donut-shaped molecules composed of glycoluril units ( $n = 5$  to  $8$ ) linked by methylene bridges (Figure 1).<sup>83</sup> The cyclic urea (of the glycoluril) and the methylene groups form the hydrophobic cavity, whereas the carbonyl groups decorate both the portals. While the low aqueous solubility of **CB6** restricts its broad application, **CB7** and **CB8**<sup>84,85</sup> have been used extensively.<sup>86</sup> Considering molecular recognition, the small cavity of **CB6** favours entrapment of an alkyl chain (1:1), **CB7** is able to host one aromatic group (1:1) whereas **CB8** can host up to 2 aromatic guests (1:1:1 or 1:2).<sup>86,87</sup> For example, **CB7** encaged a N-terminal phenylalanine of insulin<sup>58</sup> whereas the larger **CB8** in complex with 14-3-3 adapter protein facilitated dimerization. Moreover, with a larger cavity (relative to **CB7**) **CB8** accommodated two N-terminal phenylalanine groups.<sup>59</sup> In contrast, the complexation of **CB7** to a dimethylated RSL illustrated **CB** binding to a surface-exposed lysine located on a loop region (Figure 2B and 3B).<sup>21</sup> Here, **CB7** facilitated a cage-like assembly similar to that of ferretin. In these cases, shape and charge complementarity between the host and the guest play an essential role in protein binding.<sup>88</sup> Perhaps, the entrapped waters within the **CB** cavity also contribute towards protein – ligand interaction as the release of waters (on guest complexation) yields complexes with higher affinities.<sup>89</sup>



**Figure 2.** Supramolecular protein assemblies exemplified by crystal structures of **(A)**  $\alpha$ -cyclodextrin and a starch-binding protein (PDB 4FEM),<sup>93</sup> **(B)** cucurbit[7]uril clusters with a trimeric lectin (PDB 6F7W),<sup>21</sup> **(C)** suramin and sitruin homolog 5 (PDB 2NYR),<sup>114</sup> **(D)** sulfonato-calix[8]arene and cytochrome *c* (PDB 6GD6),<sup>19</sup> **(E)** porphyrin dimers with jacalin (PDB 1PXD),<sup>143</sup> **(F)** dimers of a rhodamine-mannose conjugate and concanavalin A (PDB 4P9W),<sup>151</sup> **(G)** polyoxotungstate with HSP70 nucleotide binding domain (PDB 6G3R).<sup>131</sup> **(H)** crystallophore dimer with a di-iron flavoprotein (PDB 6FRN),<sup>66</sup> Independent protein chains in the asymmetric unit are rendered as grey, dark grey, cyan, or pale cyan ribbons. Ligands are shown as spheres.

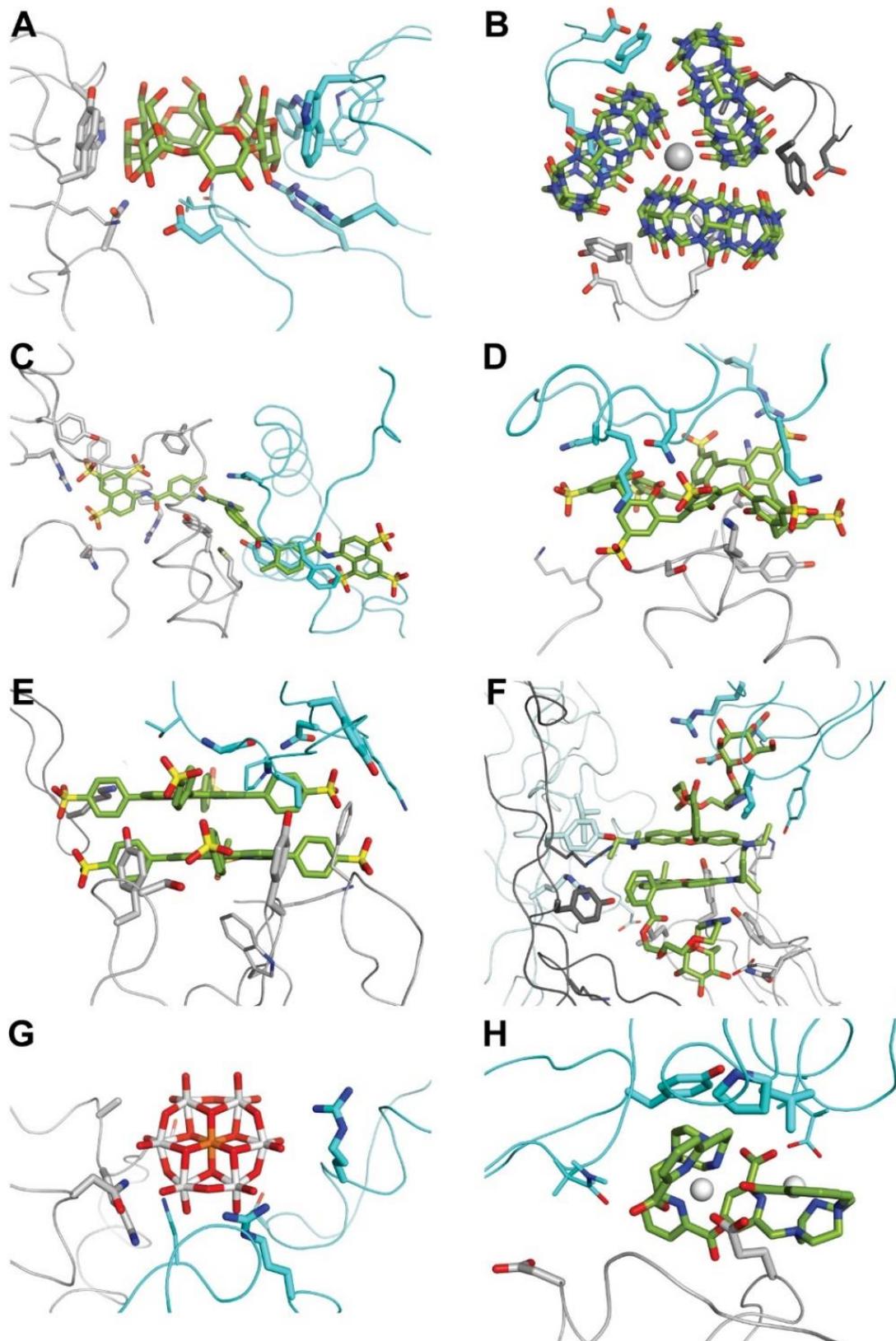
**Cyclodextrins (CD)**, are donut-shaped molecules classified as  $\alpha$ ,  $\beta$  and  $\gamma$ -**CD** depending on the number of  $\alpha$ -(1,4) linked glucopyranose subunits (Figure 1).<sup>90</sup> They are similar to **CBs** in their shape and size but structural features distinguish them.<sup>91</sup> For example, OH groups in **CD** mainly interact *via* hydrogen bonds, whereas in **CBs** the carbonyl groups participate in charge-dipole interaction in addition to hydrogen bonds. Furthermore, in contrast to **CDs**, **CB** are symmetrical owing to the identical opening. Currently, cyclodextrins are not employed as nucleating agents. However, the assembly facilitated by  $\alpha$  or  $\beta$ -**CD** at the interface of starch utilizing systems suggests their future potential as interface 'glues' (Figure 2A).<sup>92,93</sup> In these structures, hydrogen bond interactions are the apparent mode of recognition (Figure 3A).

**Foldamers** are supramolecules with a propensity to form stably-folded architectures in solution.<sup>94-96</sup> This feature combined with their resemblance to the  $\alpha$ -helix make them attractive candidates for protein recognition.<sup>97-99</sup> The helical aromatic oligoamides have become popular owing to their stability, fold predictability and ease of synthesis / functionalization. Aromatic foldamers based on 8-amino-2-quinoline carboxylic acid have been customized with amino acid analogues to enhance biocompatibility.<sup>100</sup> Also, a tether functionality has been utilized to accomplish protein recognition.<sup>100-102</sup> Recently, binding of a tether-free quinoline oligoamide to cytc was achieved (**Chapter 5**).<sup>24</sup> Charge-charge interaction appeared

to facilitate protein interactions. This complex between the globular positively charged protein and a helical anionic synthetic ligand suggested the potential of quinoline-based foldamers to generate biohybrid assembly.

**Fullerene (C<sub>60</sub>)** directs the assembly of a coiled coil peptide to yield a honeycomb architecture. The dominant interaction in the complex appeared to be mediated by the fullerene wedged between two tyrosine residues provided by the adjacent tetramers.<sup>103</sup>

**Lanthanide complexes.** Anionic organic chelates, rich in oxygen donor groups are able to entrap lanthanide cations (Ln<sup>3+</sup>) and rigidify the chelating ligand. A complex of tris(dipicolinate) and europium lanthanide was demonstrated to interact with proteins *via* electrostatic interactions.<sup>104-106</sup> Functionalization of the tris(dipicolinate) with alkyl groups resulted in stronger protein interaction, and in one case identified a new binding mode.<sup>107</sup> Recent works have highlighted the use of the mono-cationic crystallaphore (**Tb-Xo<sub>4</sub>**) for mediating protein recognition and assembly.<sup>66,108</sup> This ligand comprised a triazacyclononane macrocyclic scaffold functionalized with dipicolinate groups and coordinated to a terbium ion (Figure 1). Presence of the lanthanide ion make this compound a powerful phasing agent for crystal structure determination. **Tb-Xo<sub>4</sub>** improved crystallization probability and/or resolution in complex with various proteins including bacteriophage tail proteins (pb6, pb9), lysozyme, malate dehydrogenase, and a flavoenzyme FprA (Figure 2G and 3G).



**Figure 3.** Protein – ligand molecular recognition details of **(A-H)** structures as per Figure 2. Independent protein chains are rendered in different colours to highlight the molecular glue effect. Ligands and key side chains are shown as sticks. **(B)** Sodium and **(H)** terbium ions are represented as spheres.

**Molecular tweezers** comprising alternatively fused norbornadiene and benzene rings hosts a belt-like electron-rich cavity (Figure 1).<sup>23,60,61,109</sup> The central hydroquinone ring of the molecule is functionalized with phosphonate groups to enable the otherwise hydrophobic molecule to participate in electrostatic interactions. The lysine side chains (arginine to a lesser extent) can thread through the tweezer cavity and interact with the phosphonates and the aromatic rings can establish hydrophobic contacts. They are mainly known to modulate protein-protein interactions. For example, **CLR01** binding either inhibited<sup>60</sup> the interaction of the 14-3-3 adapter protein with one partner protein or stabilized<sup>61</sup> the interaction with another.

**Naphthalene sulfonates** have been utilized successfully to improve protein stability and crystallization. For example, the azo-based **orange G** dye acts as a filler between the amyloid  $\beta$ - or SOD1 peptide sheets and improved the crystal quality.<sup>110,111</sup> Also, **ponceau S** dye has been observed to mediate crystal packing of pyruvate kinase.<sup>112</sup> An extensively investigated naphthalene sulfonate for its assembly-inducing behaviour in protein is **suramin**. It is an anionic, conformationally flexible linear molecule. As observed with **sclx<sub>8</sub>**,<sup>19,20</sup> the ‘floppiness’ enables it form large protein – ligand interfaces (Figure 2C and 3C).<sup>113-118</sup> The first crystal structure of a protein – **suramin** complex reported more than a decade ago illustrated its potential to direct assembly.<sup>113</sup> In addition to X-ray crystal structures, solution-state ITC or NMR studies also highlight their propensity to facilitate protein dimerization.<sup>114,116-118</sup>

**Polyoxometalates (POM)** are octahedral anionic compounds comprising clusters of transition metal oxides (MO) such as vanadium, molybdenum and tungsten in their

high oxidation states.<sup>119-121</sup> The versatility offered by their structure and chemistry has led to broad spectrum of applications.<sup>120,122,123</sup> In particular they have been widely employed as phasing agents for X-ray structure determination owing to the presence of heavy atoms which result in strong anomalous scattering.<sup>124</sup> The participation of different POMs for protein binding suggests their growing popularity as facilitators of protein packing.<sup>65,119,120</sup> For example, Keggin,<sup>125</sup> Wells-Dawson<sup>126,127</sup> and Anderson-Evans type (**TEW**).<sup>128</sup> While Keggin and Wells-Dawson type have been exclusively studied in complex with lysozyme, the recognition and assembly-inducing properties of **TEW** have been illustrated with variety of proteins. For example, promoting heterogenous crystallization of an active and latent form of the enzyme mushroom tyrosinase.<sup>129</sup> Also, it improved the crystallization behaviour of aurone synthase<sup>130</sup> and HSP70 (Figure 2G and 3G).<sup>131</sup> An interesting feature that aids it in recognition is its compact disc-shape which allows it to bind in the narrow protein cleft or pass through narrow channels.

**Polyphosphates** such as the inositol phosphates (**IP**)<sup>132-134</sup> function as phosphorylating agents in various signalling cascades.<sup>135,136</sup> Particularly, **IP<sub>6</sub>** is known to induce assembly to achieve enzyme activation. For example, **IP<sub>6</sub>** facilitated a dimer (PH-TH module),<sup>137</sup> trimer (arrestin)<sup>138</sup> or a hexamer (Gag construct of HIV-1)<sup>134</sup> to regulate enzyme activity. This ligand can also function as a tether between two proteins (PDB 3RRN).<sup>139</sup> Similarly, polysulfonates such as heparin and heparin mimics (sucrose octasulfate) have also demonstrated their tendency to induce protein assembly.<sup>140,141</sup>

**Porphyryns** such as sulfonatoporphyrins have been mainly investigated for their ability to facilitate assembly in lectins – concanavalin A,<sup>142</sup> jacalin (Figure 2E and 3E),<sup>143</sup> and peanut.<sup>144</sup> Interestingly, its complex with jacalin resulted in a high porous structure (~65 % solvent content). In all three porphyrin complexes the ligand formed dimers interacting *via*  $\pi$ - $\pi$  stacking interactions and thus mediated crystal contacts. Other porphyryns highlighted in the PDB include uroporphyrinogen III (facilitating

dimerization)<sup>145</sup> and the cationic tetra *N*-methylpyridinium porphyrin (TMPyP; recognition of telomeric<sup>146</sup> or DNA quadruplexes<sup>147</sup>).

**Pyrene** and its derivatives have been employed to facilitate crystal packing *via* non-covalent interaction<sup>148,149</sup> or covalent conjugation.<sup>150</sup>

**Rhodamine-sugar conjugates** comprise of a monosaccharide (glucose or mannose) tethered to rhodamine dye *via* an oligoethylene glycol (3 or 4 units) chain.<sup>151</sup> Complexation with a lectin concanavalin A gave rise to a highly porous crystalline framework (Figure 2F and 3F). The length of the linker determined the extent of porosity.<sup>151,152</sup> For example, triethylene glycol linker resulted in a more porous framework (~69 % solvent content) compared to tetraethylene glycol (~40 % solvent content). In this case, it is noteworthy how a 'simple' structural modifications can be used to modulate protein frameworks. In a previous similar work a bivalent bismannopyranoside was utilized to direct the concanavalin A assembly into a pre-designed diamond-like porous architectures.<sup>153</sup>

### Scope of Thesis

X-ray crystallography is considered the technique-of-choice for protein structure determination.<sup>5</sup> However, the pre-requisite of obtaining diffraction-quality crystals renders protein crystallization a serious bottleneck. Previously, protein modification approaches were employed to overcome the issue.<sup>34-37</sup> An alternate strategy to promote crystallization of recalcitrant proteins is to use additives/small molecules in the crystallization mix.<sup>45-47</sup> To this end, advancements in supramolecular chemistry have availed numerous ligands / receptors (Table 1). In a protein – ligand complex, these compounds are able to act as a ‘glue’ or a ‘bridge’ to interlink the neighbouring proteins and drive crystal packing. In many instances, such complexations have resulted in noteworthy assemblies.<sup>20,21,24,143,151</sup>

The work presented in this thesis investigated the proclivity of five different calix[n]arenes and a sulfonated quinoline foldamer for protein recognition and inducing assembly. The phosphonomethyl-calix[4]arene (**Chapter 1**) and the oxomethylcarboxylate-sulfonato-calix[4]arene (**Chapter 2**) exemplified the influence of ligand customizations on protein recognition and assembly. The complexation between a calix[8]arene and porphyrin revealed an alternating stacking assembly and the potential to generate other such hybrid architectures (**Chapter 3**). Further work evaluated the tendency of the **sclx<sub>n</sub>** series to act as a ‘molecular glue’ for facilitating crystal packing of the *Penicillium* antifungal (PAF) protein (**Chapter 4**). The success of various sulfonate-bearing compounds (**sclx<sub>n</sub>**; **suramin**) in promoting protein assembly motivated us to test the crystallization propensity of a tether-free sulfonated quinoline foldamer with cytc (**Chapter 5**). This project was interesting as the previous protein – foldamer complexes relied on a tether functionality to achieve protein assembly.

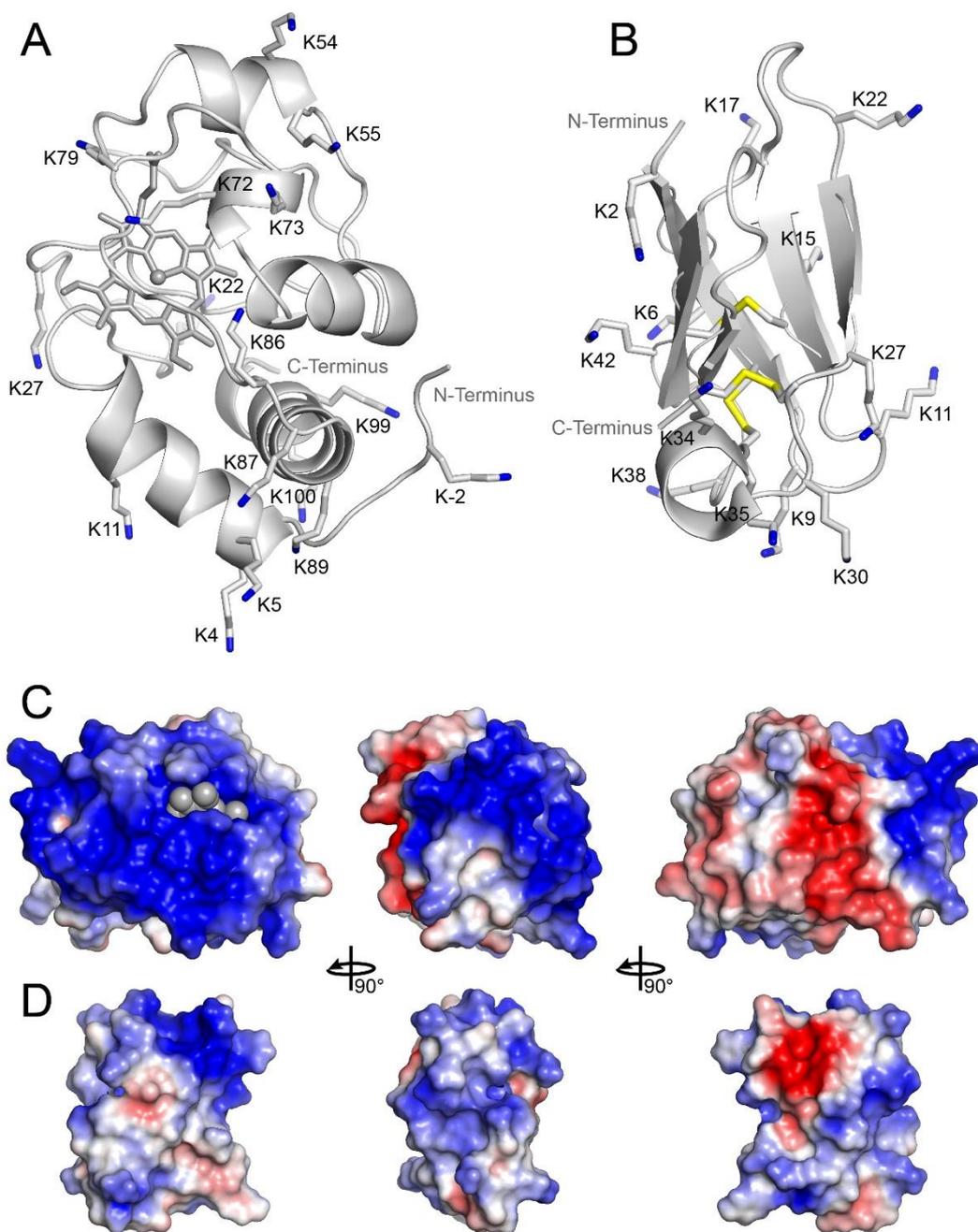
## Model Proteins

Cytochrome *c* and *Penicillium* antifungal protein were the two proteins used in the research work. The crystal structures enabled an in-depth analysis of the similarities as well as differences in the molecular recognition of the two Lys-rich cationic proteins.

### Cytochrome *c* (cytc)

The *Saccharomyces cerevisiae* cytc is a highly water-soluble heme-containing metalloprotein located in the mitochondrial intermembrane space (Figure 4A).<sup>154</sup> Here, it acts as a carrier to transfer electrons from complex III to complex IV, which is a part of electron-proton energy transduction pathway.<sup>155</sup> In this way, cytc contributes towards ATP synthesis.<sup>154</sup> Also, cytc has been illustrated to bind inositol-1,4,5-triphosphate receptors, which amplifies calcium signalling in apoptosis.<sup>154,156</sup> Thus, it acts as an intermediate in the apoptotic process. Its interactions with anionic partners is also well characterized.<sup>19,20,48,51-54,56,155-158</sup> Interestingly, cytc is classified as a moonlighting protein whose structure and function are determined by the conditions / localizations within the cell.<sup>154</sup>

Cytc is a small ~13 kDa (108 residues) cationic protein (pI 9.5; Figure 4C). The cytc fold consists of five  $\alpha$ -helices interconnected by extended  $\Omega$  loops, in addition to two-stranded anti-parallel  $\beta$ -sheets (Figure 4A).<sup>159</sup> The heme group is covalently attached *via* thio-ether linkages of Cys14 and Cys17, and is completely buried into the hydrophobic pocket. Only the heme edge is partially solvent exposed (Figure 4C). Cytc has 34 protonatable residues: Asp (4), Glu (7), Lys (16), Arg (3), and His (4). Majority of the cationic residues are located in the large patch surrounding the heme edge, which can establish contacts with anionic ligands. In addition, the electron transfer capabilities has led to its application in fabricating biosensors.<sup>160,161</sup>



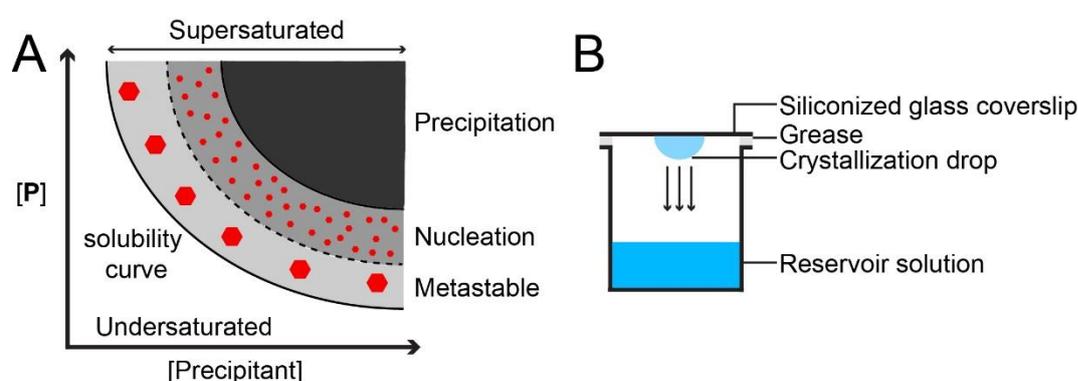
**Figure 4. Two cationic proteins.** Cartoon representations of **(A)** cytc and **(B)** PAF displaying all 16 and 13 Lys residues (sticks), respectively. The disulfide bonds in PAF are represented in yellow. The electrostatic potential surface of **(C)** cytc and **(D)** PAF showing the cationic (blue) and the anionic (red) patches (ABPS Electrostatics, PyMOL). The heme edge of cytc **(C)** is shown as grey spheres.

***Penicillium* antifungal protein (PAF)**

We were motivated to study PAF<sup>162</sup> owing to its similarity with cytc with respect to charge (cationic) and function (apoptosis).<sup>163,164</sup> PAF is a ~6.2 kDa (55 residue) highly cationic ( $pI \sim 9.0$ ; Figure 4B and D) and water-soluble protein. They are potent agent against *Aspergillus* species and dermatophytes. It is postulated that PAF exhibits the antifungal activity *via* interaction with anionic partners on the cell membrane, resulting in cell permeation, and eventually cell death.<sup>163-165</sup> PAF has 21 protonatable residues: Asp (7), Glu (1), Lys (13). The NMR structure<sup>166,167</sup> had revealed that PAF consists of five  $\beta$ -strands connected by three small loops involving  $\beta$ -turn motifs and a large loop. Each  $\beta$ -sheet is composed of five antiparallel  $\beta$ -strands and stabilised by three disulfide bridges (C7 – C36, C14 – C43, C28 – C54; Figure 4B).

### Protein – ligand co-crystallization

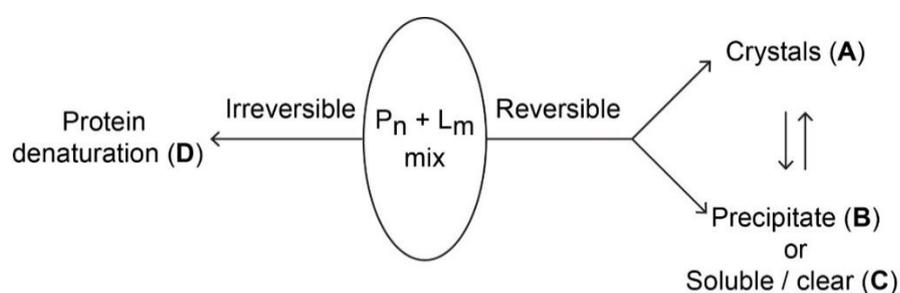
In a crystallization process,<sup>168</sup> nucleation is the pre-requisite to crystal growth.<sup>169-171</sup> However, excess nucleation may lead to growth of numerous small crystals, which may not be of diffraction-quality (Figure 5A). Metastable zone is the optimal region for the growth of bigger diffraction-quality crystals (Figure 5A and 6A) and no further nucleation occurs in this region. In contrast, if the supersaturation is too high, then the system proceeds towards the precipitation zone resulting in formation of disordered structures (Figure 6D). Crystal growth will not occur when protein is undersaturated.



**Figure 5. (A)** A crystallization phase diagram depicting the metastable (favourable for crystal growth), nucleation (numerous small crystals) and precipitation (unfavourable for crystal growth) zones. The protein remains completely soluble in the undersaturated zone. [P] indicates the protein concentration **(B)** A scheme depicting the hanging drop vapour diffusion method.

In the hanging drop vapour diffusion method, an aqueous drop contains a low concentration of the protein and the crystallization agents (and ligand). As water vapour diffuses out of the drop, the drop equilibrates with the reservoir solution (Figure 5B). Subsequently, the concentration of the protein in the drop increases resulting in supersaturation, and eventually crystal growth.

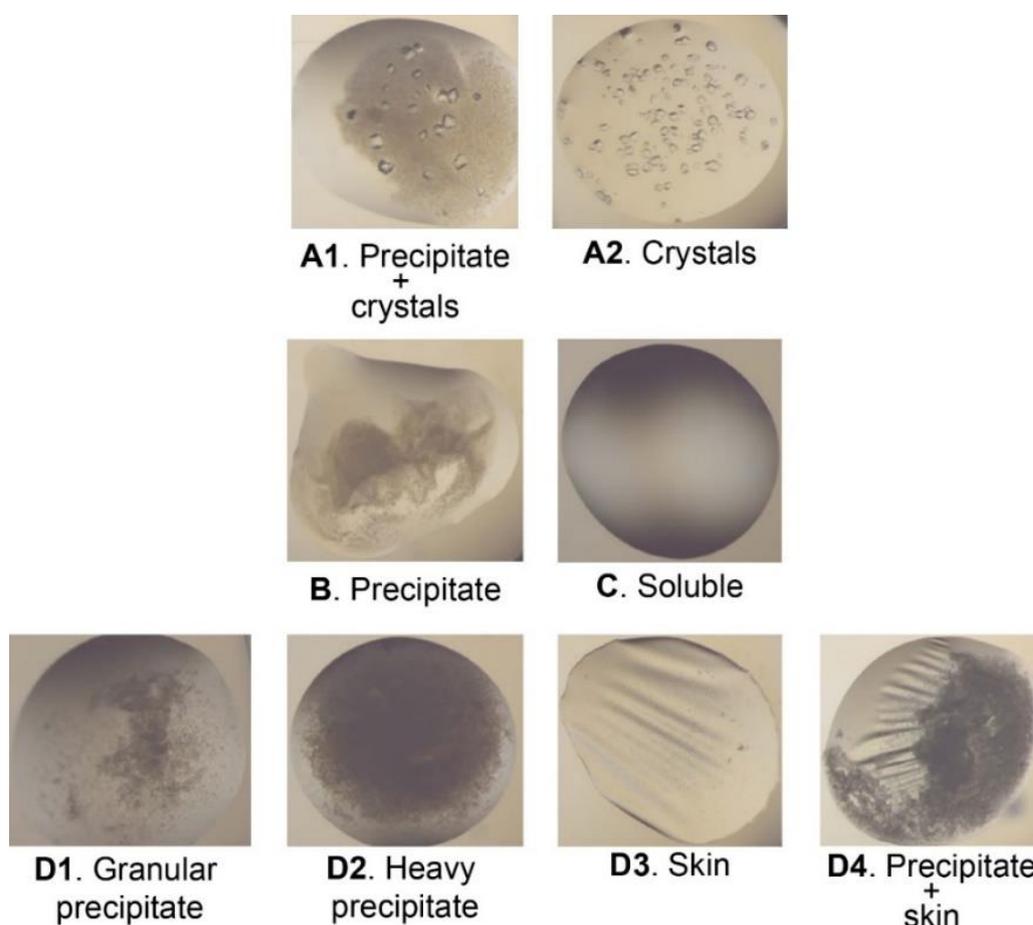
Compared to the preparation of a phase diagram, an easier approach is to carry out a small-scale precipitation test. Though not as descriptive as a phase diagram, it helps to determine the number of ligand equivalents that is suited for co-crystallization at a defined protein concentration ( $P$ ; Scheme 1). On the one hand, sufficient ligand is required to induce complexation (Figure 6A). On the other hand excessive protein precipitation is best avoided (Scheme 1; Figure 6D). In some interesting cases, low ligand equivalents may lead to protein precipitation while high equiv. may promote protein solubility.<sup>19</sup>



| n | m | Representative examples (PDB depositions) |
|---|---|---|
| 1 | 1 | 6GD6, <sup>19</sup> 6RGI <sup>56</sup>    |
| 1 | 2 | 1PXD <sup>143</sup>                       |
| 1 | 3 | 5LFT, <sup>52</sup> 6GD9 <sup>19</sup>    |
| 2 | 1 | 6CE2, <sup>118</sup> 6HAJ4 <sup>55</sup>  |
| 2 | 3 | 3TYI, <sup>48</sup> 5NCV <sup>53</sup>    |
| 4 | 3 | 6GD8 <sup>19</sup>                        |
| 4 | 5 | 4PRQ <sup>49</sup>                        |

**Scheme 1.** A protein – ligand co-crystallization experiment may result in the formation of co-crystals under favourable concentration of the protein and the ligand (Figure 6. A1-A2). Crystal growth may also eventually result from a clear or lightly precipitated drop (Figure 6. B1-B2). Protein denaturation is an irreversible phenomenon and should be avoided (Figure 6. D1-D4).

A light to medium precipitation can be an indicator of protein-ligand interaction preceding crystallization (Figure 7). However, persistent heavy protein precipitate over a range of [L] might be slow / unyielding to crystal generation. Furthermore, a prior knowledge of the 'safe' [L] becomes especially important for robot crystallization trials which usually requires 32  $\mu$ L of a protein-ligand mixture. A random selection of [L] might cause protein precipitation / denaturation resulting in undesirable sample loss (Figure 6. D1-D4). In addition, a precipitated sample for crystallization robot (Oryx 8; Douglas Instrument) trials poses the risk of blocking the sample channels.



**Figure 6.** Protein-Ligand Drops. **(A1)** Lysozyme – crystallophore crystals growing from a precipitate (PDB 6f2i).<sup>66</sup> **(A2)** Lysozyme – sulfonatocalix[4]arene crystals (PDB 4prq)<sup>49</sup> **(B)** A light precipitate (cloudy/turbid appearance) on addition of the ligand may eventually lead to crystal growth. **(C)** A soluble or clear drop suggesting absence

of protein-ligand interaction may produce crystals over time as the drop equilibrates with the reservoir condition. A bad precipitate suggests protein denaturation and can be characterized as **(D1)** dry granular **(D2)** heavy / clumpy / brown-coloured **(D3)** skin over the crystallization drop or **(D4)** a combination of precipitate and skin.

Consider, **P** mM and **L** mM stock solutions of protein and ligand, respectively.

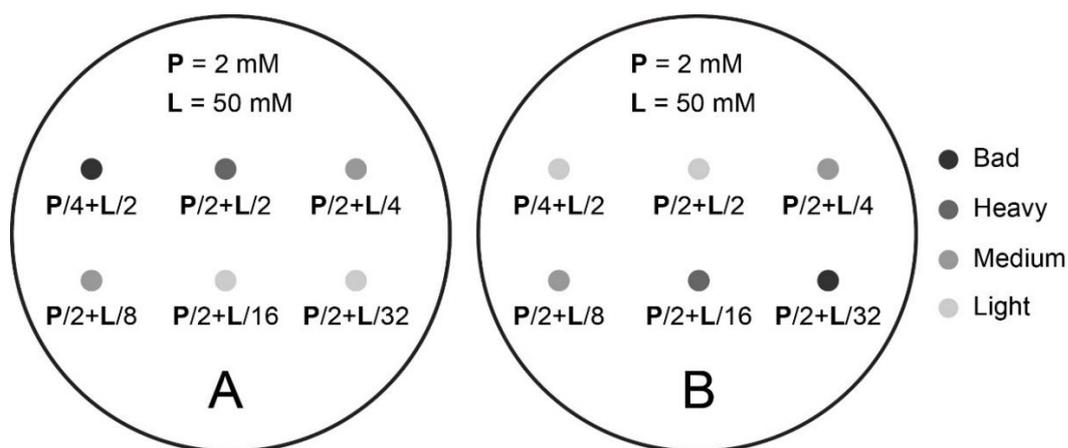
Example, **P** = 4 mM and **L** = 50 mM

- Tests are carried out using 1-2  $\mu\text{L}$  (final volume) of protein – ligand mixture.
- A constant **P** is tested against varying **L** or vice versa.
- The drops are examined under a microscope to determine solubility or any precipitation.
- Bear in mind that **P** and **L** decrease by 2-fold when mixed in a 1:1 ratio --  
 $1 \mu\text{L } \mathbf{P} \text{ mM} + 1 \mu\text{L } \mathbf{L} \text{ mM} = 2 \mu\text{L mixture of } \mathbf{P}/2 \text{ and } \mathbf{L}/2$

Step 1. Prepare a 2-fold serial dilution of protein – **P/2**, **P/4**, **P/8**, **P/16** mM.

Step 2. Similarly prepare a 2-fold serial dilution of the ligand stock, **L** mM - **L/2**, **L/4**,  
**L/8**, **L/16**, **L/32**

Step 3. It is advisable to test range of ligand equivalents such as 1 equiv., 25 equiv., 50 equiv. etc. prior to selecting the experimental **P** and **L** (Figure 7).



**Figure 7.** Protein-Ligand Mixtures. **(A)** Extent of precipitation (grey scale) is dependent on the ligand concentration. **(B)** A counter-intuitive result is the appearance of heavy precipitate at lower ligand concentration and soluble /clear drops at higher concentration of the ligand.

#### Case A

1. In this example,  $P/2$  and  $L/16$  or  $L/32$  can be used for the crystallization trials.
2. Higher  $[L]$  of 50 equiv. and 25 equiv. should be avoided owing to the tendency to result in heavy precipitation. However, these equivalents can be tested in manual crystallization trials where the ligand and protein solutions are not pre-mixed. The heavy precipitate (if any) may dissolve over weeks to yield crystals (Figure 6. A1 and A2).

Similarly, other protein and ligand concentrations can also be explored.

#### Case B

The protein is soluble at higher  $[L]$  but starts precipitating as  $[L]$  is lowered (Figure 7B). This result is counter-intuitive as normally the extent of precipitation decreases as  $[L]$  is decreased (Figure 7A). Interestingly, this observation might suggest the occurrence of higher assembly state at lower  $[L]$  while a higher  $[L]$  disrupts the oligomeric state.<sup>19</sup>

In this case, both the highest [L] and the lowest [L] which do not cause precipitation can be tested in robot crystallization trials. For example,  $P/4$  or  $P/2$  with  $L/2$



## Chapter 1

---

### Phosphonated calixarene as a “molecular glue” for protein crystallization

The material in this chapter was published as:

**Alex J. M.**, Rennie M. L., Volpi S., Sansone F., Casnati A., Crowley P. B.

Phosphonated calixarene as a “molecular glue” for protein crystallization.

*Cryst. Growth Des.* **2018**, *18*, 2467-2473.

Alessandro Casnati's Lab (Università degli Studi di Parma, Italy) –

- Synthesis of the phosphonomethyl-calix[4]arene (**pmclx<sub>4</sub>**)

### Abstract

Protein crystallization remains a serious bottleneck to structure determination by X-ray diffraction methods. Compounds acting as ‘molecular glue’ provide a promising strategy to overcome this bottleneck. Such molecules interact *via* non-covalent bonds with two or more protein surfaces to promote lattice formation. Here, we report a 1.5 Å resolution crystal structure of lysine-rich cytochrome *c* complexed with *p*-phosphonomethyl-calix[4]arene (**pmclx<sub>4</sub>**). Evidence for complex formation in solution was provided by NMR studies. Similar to *p*-sulfonato-calix[4]arene (**sclx<sub>4</sub>**), the cavity of **pmclx<sub>4</sub>** entrapped a single lysine side chain. Interesting features of protein recognition by the phosphonate substituents were identified in the crystal structure. A new calixarene binding site was identified at Lys54. The electron density at this site indicated two distinct calixarene conformers, suggesting a degree of ligand mobility. The role of **pmclx<sub>4</sub>** in protein crystal packing (‘molecular glue’ and patchy particle model) as well as differences in protein-binding with respect to **sclx<sub>4</sub>** are discussed.

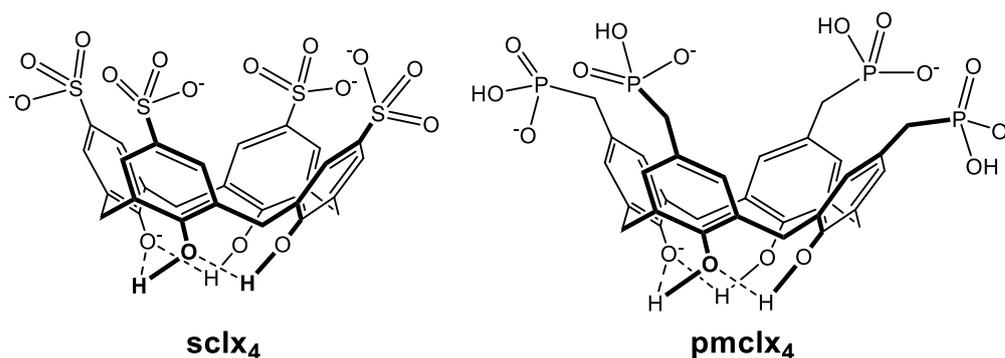
## Introduction

X-ray crystallography is the principal technique for protein structure determination at atomic resolution and the need for diffraction-quality crystals renders protein crystallization a serious bottleneck.<sup>28,32,171</sup> Protein crystals are finding applications also in catalysis and sensing.<sup>160,172,173</sup> Considering that the protein is the main variable in crystallization,<sup>174</sup> various methods such as truncation, site-directed mutagenesis and lysine methylation (surface entropy reduction) have been utilized to achieve crystal growth of hard-to-crystallize proteins.<sup>175-178,34</sup>

An alternative strategy to covalent modification of the protein is the use of additives that prompt assembly and crystallization by interacting with protein surfaces. In view of this approach, a repertoire of compounds have been employed to achieve crystallization of recalcitrant proteins. Examples include dicarboxylic acids / diamino derivatives,<sup>45-57</sup> pyrene-tetrasulfonic acid,<sup>112</sup> polyoxometalates,<sup>65,123</sup> and most recently the crystallophore.<sup>66,108</sup> We have been tackling the problem by using *p*-sulfonato-calix[4]arene (**sclx<sub>4</sub>**, Figure 1), a highly water soluble supramolecular building block. Similar to the molecular tweezers,<sup>61,179</sup> this anionic, bowl-shaped compound provides a hydrophobic cavity that can accommodate amino acids,<sup>180</sup> in particular, the cationic side chains of lysine and arginine.<sup>70,71,181</sup> Crystal structures of **sclx<sub>4</sub>** in complex with the cationic proteins cytochrome *c*<sup>48</sup> (cytc) and lysozyme,<sup>49,50</sup> suggest that the calixarene acts as a ‘molecular glue’ to drive protein crystallization.

A great variety of calix[4]arenes have been synthesized with different functional groups at the upper rim to facilitate host-guest supramolecular chemistry.<sup>182,183</sup> Considering the utility of sulfonated-calixarenes for protein binding<sup>48-50,52,70,71,181-184</sup> we were motivated to investigate the ‘molecular glue’ property of other anionic calixarenes. Phosphonate-containing calixarenes<sup>185</sup> have been developed for protein binding<sup>186-188</sup> and recently, we reported the structure and assembly of cytc in complex with phosphonato-calix[6]arene (**pclx<sub>6</sub>**).<sup>51</sup> Here, we demonstrate co-crystallization between *p*-phosphonomethyl-calix[4]arene<sup>189,190</sup>

(**pmclx<sub>4</sub>**, 800 Da, Figure 1) and cytc. In contrast to **sclx<sub>4</sub>** with the sulfonate anion fixed in the plane of the aromatic ring, the phosphonomethyl substituent of **pmclx<sub>4</sub>** can switch between distinct conformations. We were interested to study the impact of this different substituent (and charge distribution) on the protein binding capacity of the ligand.



**Figure 1.** Molecular structures of *p*-sulfonato-calix[4]arene (**sclx<sub>4</sub>**) and *p*-phosphonomethyl-calix[4]arene (**pmclx<sub>4</sub>**). Calix[4]arenes adopt a cone conformation due to hydrogen bonding between the phenolic hydroxyls, one of which is deprotonated.

A 1.5 Å resolution crystal structure provided a detailed view of how **pmclx<sub>4</sub>** binds to cytc. Complex formation was observed at Lys86, which enabled a direct comparison with the structure of a bromo-functionalized sulfonato-calixarene<sup>52</sup> that also binds this residue. A new binding site was observed at Lys54, though this feature was not evident in NMR solution studies. The molecular recognition capacity and binding affinity imparted by the phosphonomethyl substituent is discussed. The crystal structure and solution-state studies illustrate the two-sided role of **pmclx<sub>4</sub>** as a lysine recognition agent and as molecular glue. Furthermore, we suggest that protein surface decoration by calixarenes may provide additional ‘sticky patches’ for crystallization.<sup>34</sup>

## Experimental

**Materials.** The ligand 5,11,17,23-tetra-(methylphosphonic acid)-25,27,26,28-tetrahydroxy calix [4]arene (**pmclx<sub>4</sub>**) was (1) used as presented in the protein crystallization additive kit CALIXAR (Molecular Dimensions; Figure A1, Appendix) or (2) synthesized<sup>189,191</sup> and prepared as a 50 mM solution in water at pH 6.0. Samples of <sup>15</sup>N-labelled and unlabelled *Saccharomyces cerevisiae* cytc C102T were produced by established methods.<sup>48,191</sup>

**Protein-ligand co-crystallization.** Cytc – **pmclx<sub>4</sub>** co-crystals were grown at 20° C using the hanging drop vapor diffusion method. Hanging drops were prepared with 2 μL of 1.5 mM oxidized cytc, 0.4 μL of 3.0 or 5.0 mM **pmclx<sub>4</sub>** and 1.6 μL of the reservoir. Crystals were obtained at 17-25 % of PEG 3350, 8000 or 10000, in 50 mM NaCl and 50 mM sodium acetate at pH 5.6 (Figure 2). Control drops lacking **pmclx<sub>4</sub>** remained devoid of crystals.

**Data collection.** Crystals of ~150 μm dimension (Figure 2D) were cryo-protected in the reservoir solution supplemented with 20 % glycerol and cryo-cooled in liquid nitrogen. Diffraction data were collected at 100° K at beamline 24-ID-C (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) with a Pilatus 6M-F detector. The dataset was collected from a single crystal using  $\phi$  scans of 0.1° over 200°.

**Structure determination.** The observed reflections were reduced, merged, and scaled with XDS.<sup>192</sup> POINTLESS<sup>193</sup> and AIMLESS<sup>194</sup> revealed anisotropic diffraction extending to 1.2 Å along 0.57 h + 0.82 l and k axes and 1.5 Å along -0.47 h + 0.88 l axis as indicated by CC<sub>1/2</sub> (> 0.3). Data extending to 1.5 Å was used for model building and refinement (Table 1). The structure was determined by molecular replacement in PHASER<sup>195</sup> with 1ycc as the search model. The coordinates for **pmclx<sub>4</sub>** were built in JLigand<sup>196</sup> and added to the model in COOT.<sup>197</sup> Iterative cycles of manual

model building and refinement (in REFMAC5 as implemented in CCP4<sup>198</sup>) was continued until no further improvements in  $R_{\text{free}}$  or electron density were obtained. The structure was validated with MolProbity<sup>199</sup> and deposited in the Protein Data Bank (PDB 5NCV).

**Accessible surface area (ASA) calculations.** The effect of **pmclx<sub>4</sub>** on the ASA of cytc residues in the crystal structure was determined in ArealMol,<sup>200</sup> according to equations 1 and 2. These calculations take into account the ASA of residues in the crystal packing environment in the absence ( $A_{\text{pro}}$ , Å<sup>2</sup>) and presence ( $A_{\text{pro+lig}}$ , Å<sup>2</sup>) of the ligand. We define  $Y_{\text{lig}}$  (Å<sup>2</sup>), the ASA of each residue in the presence of ligand in the crystal packing environment and  $X_{\text{free}}$  (Å<sup>2</sup>), the ASA of residues in the free protein (ignoring ligand and crystal packing).  $Z_{\text{lig}}$  (Å<sup>2</sup>) is the area that remains accessible.

$$Y_{\text{lig}} = A_{\text{pro}} - A_{\text{pro+lig}} \quad (1)$$

$$Z_{\text{lig}} = X_{\text{free}} - Y_{\text{lig}} \quad (2)$$

In Figure 7,  $X_{\text{free}}$  and  $Z_{\text{lig}}$  are plotted as black and grey bars, respectively.

**NMR spectroscopy.** The typical sample composition was 0.1 mM <sup>15</sup>N cytc in 20 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NaCl, 1 mM sodium ascorbate, 10 % D<sub>2</sub>O at pH 6.0. NMR titrations were performed at 303 K using 0.5-32 μL aliquots of a 50 mM **pmclx<sub>4</sub>** stock. <sup>1</sup>H-<sup>15</sup>N HSQC spectra were acquired with spectral widths of 16 ppm (<sup>1</sup>H) and 40 ppm (<sup>15</sup>N) on a 600 MHz Varian spectrometer equipped with a HCN coldprobe.<sup>48</sup> Analysis of the ligand-induced chemical shift perturbations was performed in CCPN.<sup>201</sup>

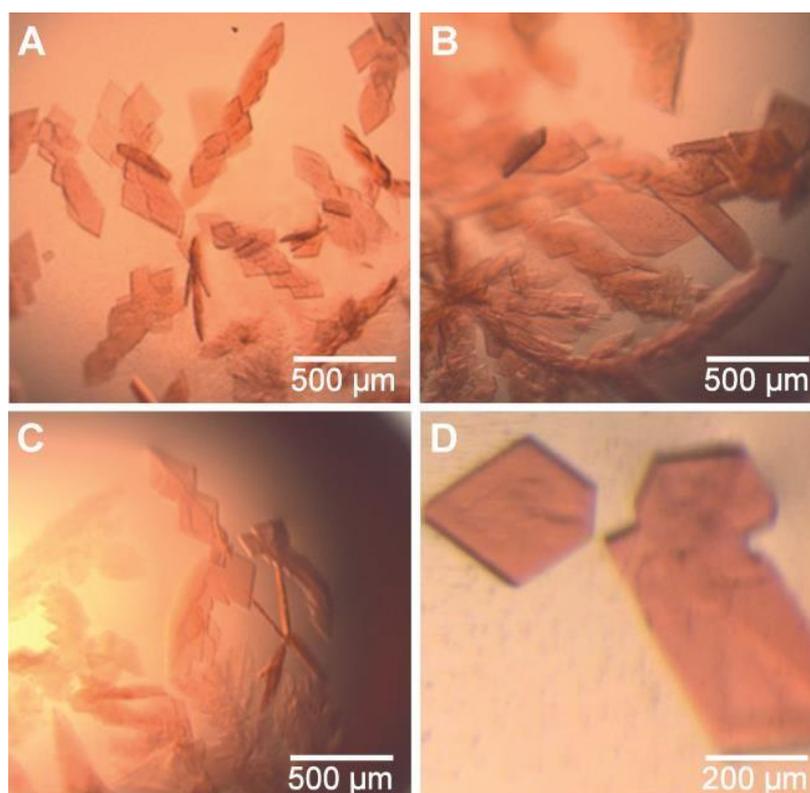
**NMR-derived binding curves.** Binding isotherms were obtained by plotting the chemical shift changes ( $\Delta\delta$ ) as a function of the ligand concentration. The data were fit globally in Origin by using a 1:1 binding model (Equation 3).

$$\Delta\delta = \Delta\delta_{\text{max}} \frac{[A_T] + [B_T] + [K_d] - \sqrt{([A_T] + [B_T] + [K_d])^2 - 4[A_T][B_T]}}{2[B_T]} \quad (3)$$

Where,  $[A_T]$  and  $[B_T]$  were the ligand and protein concentrations, respectively.

## Results and discussion

**Cytc – pmclx<sub>4</sub> co-crystallization.** pmclx<sub>4</sub> was identified initially as a ligand for cytc crystallization by using the CALIXAR Kit (Molecular Dimensions). Co-crystals of pmclx<sub>4</sub> and cytc were obtained from conditions similar to those reported for sclx<sub>4</sub>.<sup>48</sup> While 10 equiv. of sclx<sub>4</sub> were required to achieve crystal growth, <1 equiv. of ligand were sufficient to obtain cytc – pmclx<sub>4</sub> co-crystals. Crystals grew reproducibly in a range of PEG molecular weights and concentrations (17-25 % PEG 3350, 8000 or 10000) in 50 mM NaCl and 50 mM sodium acetate pH 5.6 (Figure 2). In the absence of pmclx<sub>4</sub> there was no crystal growth, indicating that the ligand was necessary for crystallization under these conditions.



**Figure 2.** Co-crystals of cytc and pmclx<sub>4</sub> grown in hanging drops that contained 17 % (A) PEG 3350, (B) PEG 8000, (C) PEG 10000 or (D) 20 % PEG 8000 in 50 mM NaCl and 50 mM sodium acetate pH 5.6. Crystals were obtained at 0.75 mM protein and 0.3 or 0.5 mM ligand. A crystal from (D) diffracted to 1.5 Å.

**X-ray structure of cytc – pmclx<sub>4</sub>.** A single crystal (Figure 2D) resulted in anisotropic diffraction (1.2 Å resolution in two directions and 1.5 Å in the third direction). Structure determination in PHASER<sup>195</sup> (Table 1) yielded two protein molecules (chains A and B), and three calixarenes in the asymmetric unit (Figure 3A).

**Table 1.** X-ray data collection, processing and refinement statistics

| <i>Data collection</i>             |  |
|------------------------------------|--|
| Light source                       | APS  |
| Beamline                           | 24-ID-C  |
| Wavelength (Å)                     | 0.97910  |
| Space group                        | <i>P</i> 1 2 <sub>1</sub> 1  |
| Cell constants                     | $a = 28.79 \text{ \AA}$ , $b = 79.90 \text{ \AA}$ , $c = 47.73 \text{ \AA}$<br>$\alpha = \gamma = 90^\circ$ , $\beta = 93^\circ$ |
| Resolution (Å)                     | 47.46-1.50 (1.55-1.50)   |
| # unique reflections               | 34591 (3452)   |
| Multiplicity                       | 5.5 (5.4)  |
| $I/\sigma(I)$                      | 18.7 (5.1)   |
| Completeness (%)                   | 99.2 (99.4)  |
| $R_{\text{meas}}^b$ (%)            | 6.1 (37.6)   |
| $R_{\text{pim}}^c$ (%)             | 2.6 (15.8)   |
| CC <sub>1/2</sub>                  | 0.999 (0.966)  |
| Solvent content (%)                | 46.7   |
| <i>Refinement</i>                  |  |
| $R_{\text{work}}$ (%)              | 0.1707   |
| $R_{\text{free}}$ (%)              | 0.2012   |
| rmsd bonds (Å)                     | 0.0113   |
| rmsd angles (°)                    | 1.6080   |
| asymmetric unit composition        |  |
| protein                            | 2  |
| <b>pmclx<sub>4</sub></b>           | 3  |
| water                              | 281  |
| Avg. B-factor (Å <sup>2</sup> )    | 20.6   |
| Ramachandran analysis <sup>d</sup> |  |
| % residues in favoured regions     | 98.0   |
| % residues in allowed regions      | 100.0  |
| PDB code                           | 5ncv   |

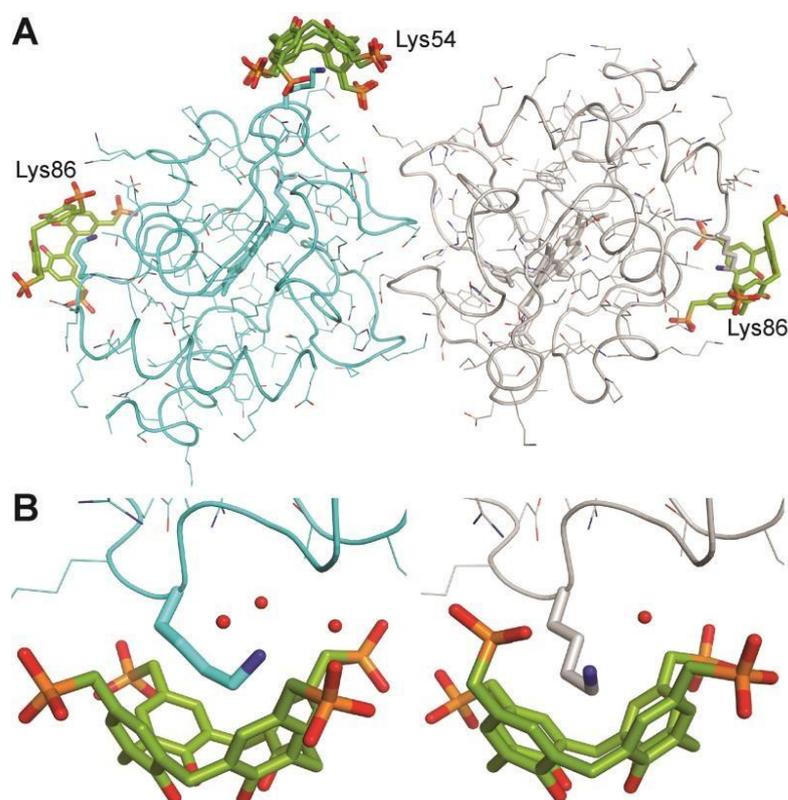
<sup>a</sup>Values in parentheses correspond to the highest resolution shell;

<sup>b</sup> $R_{\text{meas}} = \frac{\sum_{hkl} \sqrt{(n/n-1) \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}}{\sum_{hkl} \sum_i I_i(hkl)}$ ;

<sup>c</sup> $R_{\text{pim}} = \frac{\sum_{hkl} \sqrt{(1/n-1) \sum_{i=1}^n |I_i(hkl) - \langle I(hkl) \rangle|}}{\sum_{hkl} \sum_i I_i(hkl)}$ ;

<sup>d</sup>Calculated in MolProbity.

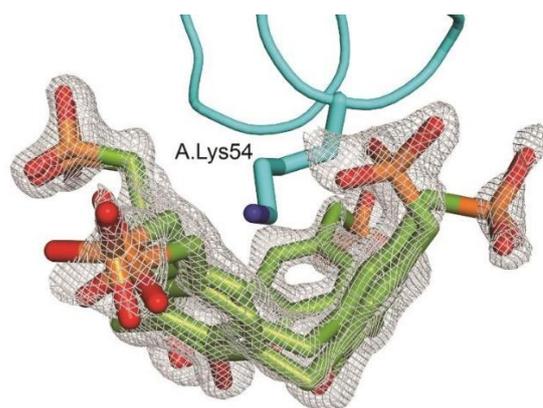
There was no electron density for the four N-terminal residues (including Lys-2) in either protein chain, indicative of disorder. The presence of two cytc molecules in the asymmetric unit was similar to other calixarene-bound structures.<sup>48,51,52</sup> All three calixarenes were modeled at sites where they encapsulated a single lysine residue, Lys86 in chains A and B and Lys54 in chain A only.



**Figure 3.** (A) The asymmetric unit of the cytc – **pmclx<sub>4</sub>** complex comprised two molecules of cytc (chains A and B in blue and grey), and three molecules of **pmclx<sub>4</sub>** (PDB 5NCV). The proteins are shown as ribbons with side chains as lines and the entrapped lysines as sticks. (B) The Lys86 binding site in chains A and B differ by the absence or presence of a salt bridge to one phosphonate substituent. Note the altered lysine conformation and hydration, with waters indicated as red spheres.

In the protein-bound form, the phosphonate substituents of **pmclx<sub>4</sub>** occurred in two distinct conformations. Either all four substituents were splayed outwards

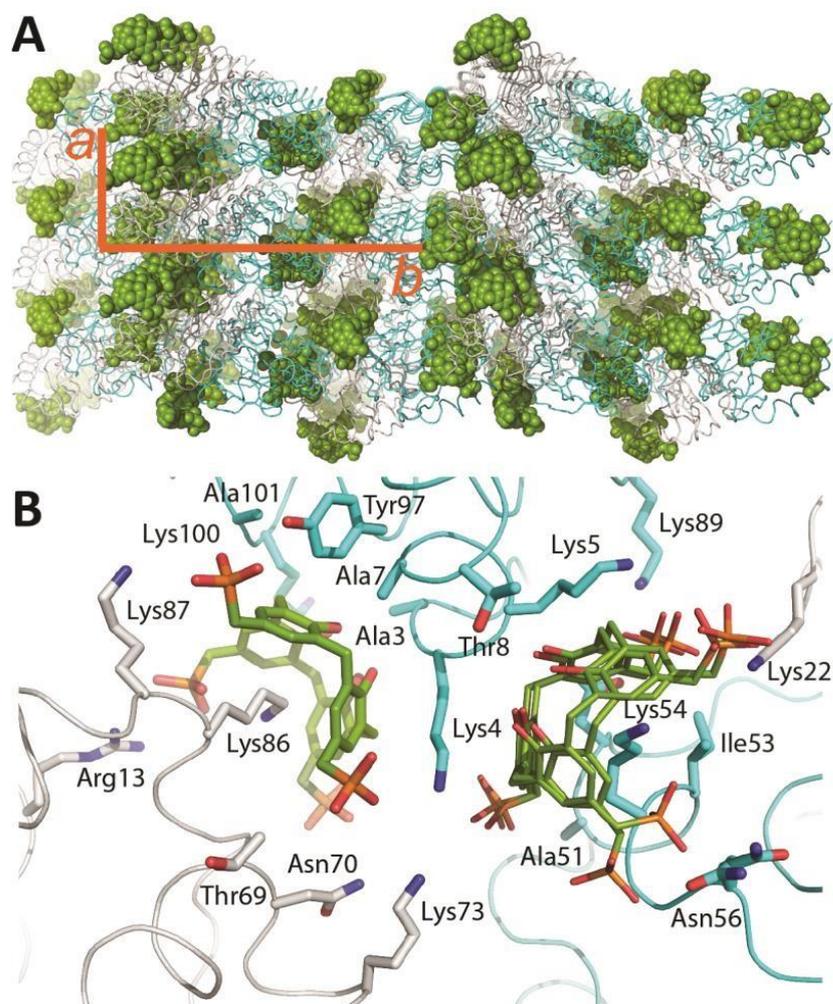
(away from the cavity) or one was turned inwards (toward the cavity). While both conformations permitted complete encapsulation of one lysine side chain they resulted in markedly different protein-calixarene interactions (Figure 3B). For example, at Lys86 in chain A (A.Lys86) there was no salt bridge interaction between the lysine ammonium group and the phosphonates. Instead, the ammonium group was solvated by 3 water molecules. In contrast, at Lys86 in Chain B (B.Lys86) one of the phosphonomethyl groups was rotated into the cavity permitting salt bridge formation to the lysine. In this case, the ammonium was solvated by one water molecule. Similar features occurred at Lys54 in chain A (A.Lys54) site. Here, it was necessary to model two calixarenes (70:30 occupancy) to account for the electron density (Figure 4 and 5B).



**Figure 4.**  $2F_o - F_c$  map (at 1.5 Å resolution and contoured to 1.0  $\sigma$ ) showing the electron density at A.Lys54, requiring two **pmclx<sub>4</sub>** conformers.

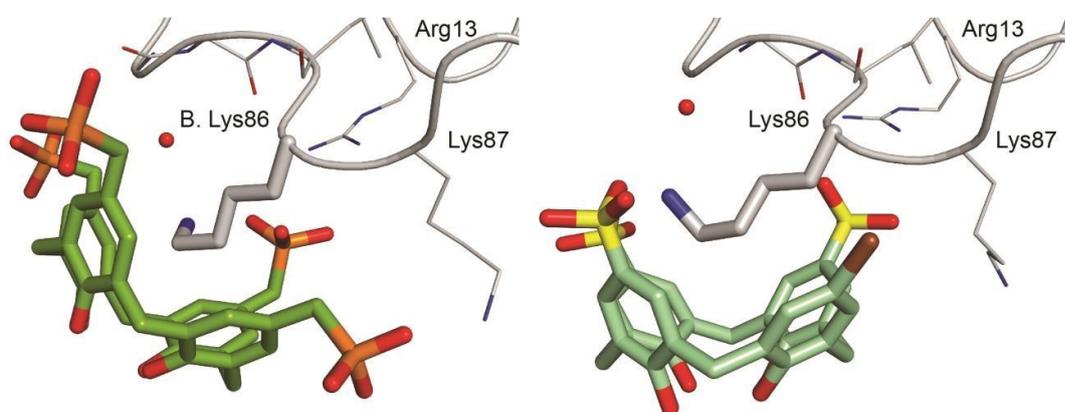
The two models differed most by a rotation of one phosphonomethyl group and the higher occupancy conformer had a salt bridge to the lysine ammonium ( $N^{\zeta} \cdots O-P = 3.1$  Å). This phosphonate also hydrogen bonded with the backbone carbonyl of Lys54 ( $CO \cdots O-P = 2.8$  Å). The requirement for two conformations at this site suggested that ligand binding was less well defined. Interestingly, this site was not detected in NMR experiments (*vide infra*). Lys54 is flanked by hydrophobic side chains, Ala51 and Ile53, both of which were in van der Waals contact with an upper

rim  $-\text{CH}_2-$  of **pmclx<sub>4</sub>**. Polar side chains contributed also with hydrogen bonds from Asp50 and Asn56 to the phosphonates.



**Figure 5. (A)** Crystal packing in the cytc – **pmclx<sub>4</sub>** complex. Proteins and ligands are rendered as ribbons and spheres, respectively. The unit cell axes *a*, *b* are indicated. **(B)** ‘Molecular glue’ contacts between two molecules of **pmclx<sub>4</sub>** and four protein chains (color scheme as per Figure 3). Alternate conformations of the ligand at Lys54 (shown as light sticks) were apparent in the electron density (Figure 4). All side chains within van der Waals contact of the ligands are shown as sticks. Lys4 is sandwiched between two **pmclx<sub>4</sub>** ligands at A.Lys54 and B.Lys86. In addition to Lys-phosphonate salt bridges, numerous other features (labeled side chains) contributed to binding.

Lysine encapsulation by **pmclx<sub>4</sub>** had some similarities with **sclx<sub>4</sub>** including cation- $\pi$  and CH- $\pi$  bonds with the Lys-C <sup>$\epsilon$</sup> .<sup>48</sup> However, the complexes formed by these ligands differed crucially in the degree of salt bridge formation with the lysine ammonium group. In **pmclx<sub>4</sub>** a single phosphonate alternated in/out (Figures 3B and 5B), while in **sclx<sub>4</sub>** (with the anion fixed in the plane of the aromatic ring) salt bridges were formed with one or two sulfonates. The sulfonate also had a tendency to hydrogen bond with backbone amide NHs. No such interactions were observed with **pmclx<sub>4</sub>**. Similarities with the bromine-functionalized sulfonato-calixarene **Br.sclx**, should be noted, as this ligand bound also to Lys86 (Figure 6).<sup>52</sup>



**Figure 6.** The Lys86 binding site of cytc in complex with **pmclx<sub>4</sub>** or **Br.sclx** (PDB 5LFT<sup>52</sup>). The proteins are oriented identically to highlight differences in the calixarene-protein contacts.

**Crystal packing interactions.** Applying the symmetry operations to the asymmetric unit revealed a close-packed structure (Figure 5A) and nine cytc – **pmclx<sub>4</sub>** interfaces, ranging in size from 30-300 Å<sup>2</sup> (Table 2). Interestingly, the A.Lys54 and B.Lys86 complexes were positioned adjacent to each other and sandwiched Lys4 (Figure 5B). Together the two ligands masked ~90 % of this Lys4 (Figure 5B) *via* multiple interactions including salt bridges (N <sup>$\zeta$</sup> ...O-P ~3.0 Å), and CH- $\pi$  / cation- $\pi$  bonds (Lys4-C <sup>$\epsilon$</sup> /C <sup>$\delta$</sup> ...centroid ~4.0 Å).

Crystal packing also involved direct or water-mediated salt bridges, for example, A.Lys54 **pmclx<sub>4</sub>** with Lys5, Lys22, Lys73 and Lys89. And B.Lys86 **pmclx<sub>4</sub>** formed a cation- $\pi$  bond to Lys100 and an edge-to-face  $\pi\cdots\pi$  interaction with Tyr97 ( $C^{\epsilon}\cdots$ centroid = 3.8 Å). The  $C^{\beta}$  atoms of Ala7 and Ala101 (which flank Tyr97) were in van der Waals contact with methylene bridges of **pmclx<sub>4</sub>**. Interestingly, this hydrophobic binding patch was identified previously in the complexes with a phenyl-substituted sulfonato-calixarene (**Ph.sclx**)<sup>52</sup> and with **pclx<sub>6</sub>**.<sup>51</sup> Two other interactions merit mention here. It has been observed previously that the calixarene lower rim phenolic hydroxyls make van der Waals' contact with the  $C^{\beta}$  of alanine.<sup>48,52</sup> This same feature was evident at B.Lys86 **pmclx<sub>4</sub>**. At A.Lys54 **pmclx<sub>4</sub>** the phenolic hydroxyls were hydrogen bonded to a water molecule and to the hydroxyl of Thr8.

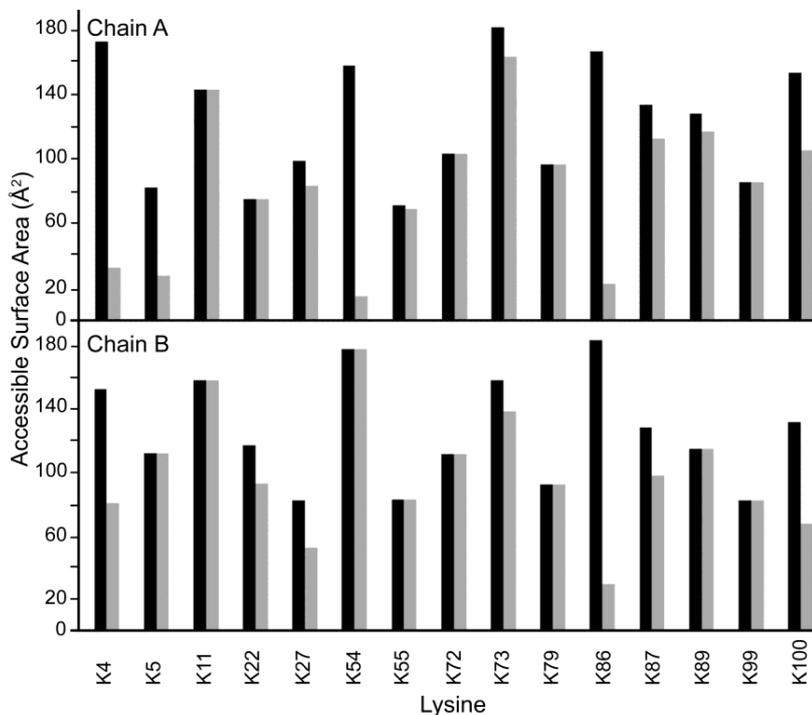
**Table 2.** Interface analysis of the cytc – **pmclx<sub>4</sub>** binding sites.

| Protein Site <sup>a</sup> | <b>pmclx<sub>4</sub></b> | Interface Area (Å <sup>2</sup> ) | Total Area (Å <sup>2</sup> ) |
|---------------------------|--------------------------|----------------------------------|------------------------------|
| <b>A.Lys54</b>            | 1                        | 270                              |                              |
| A1                        |                          | 210                              | 540                          |
| B                         |                          | 30                               |                              |
| B1                        |                          | 30                               |                              |
| <b>A.Lys86</b>            | 2                        | 260                              | 550                          |
| B2                        |                          | 290                              |                              |
| <b>B.Lys86</b>            | 3                        | 300                              | 630                          |
| A2                        |                          | 265                              |                              |
| B3                        |                          | 65                               |                              |

<sup>a</sup> $P12_1$  symmetry operations used to generate symmetry mates: **A1**, -x, y-1/2, -z+1; **B1**, x-1, y, z; **B2**, -x, y-1/2, -z+1; **A2**, -x-1, y-1/2, -z; **B3**, x-1, y, z.

The varying contributions of all lysines to ligand binding were evident from ASA calculations (methods and Figure 7). Of the most accessible residues; Lys4, Lys11, Lys54, Lys73 and Lys86 (ASA > 150 Å<sup>2</sup>), **pmclx<sub>4</sub>** selected Lys86. On the other hand Lys55, Lys79 and Lys99 were least accessible (ASA < 100 Å<sup>2</sup>) and did not contribute to ligand binding. The substantially different contributions from each

protein chain are also evident, as discussed for Lys54. Similarly, Lys5 from chain A was largely involved in ligand binding, while in chain B it did not interact with **pmclx<sub>4</sub>**.



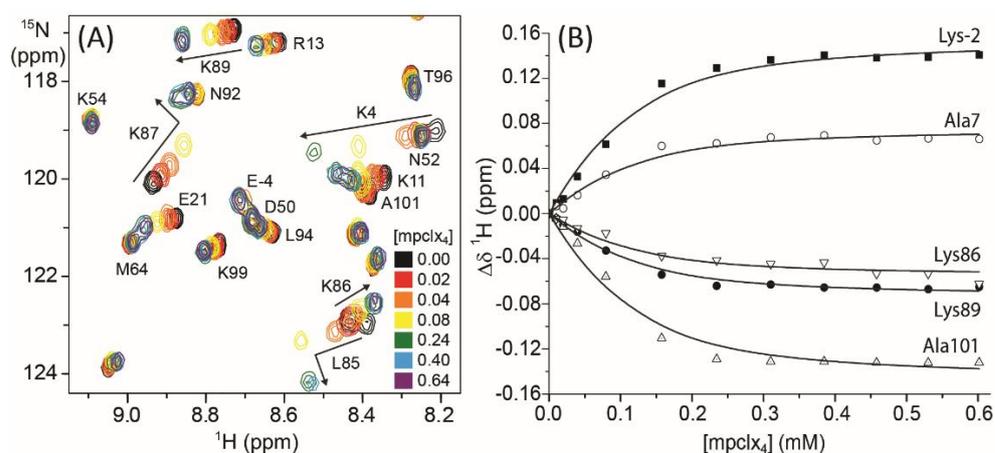
**Figure 7.** Accessible surface area (ASA) of lysines in the ligand-free (black bars) and -bound (grey) states of cytc. The varying contributions of different lysines to **pmclx<sub>4</sub>** binding in chains A and B are evident. Lys-2 is not shown as this residue was disordered in the crystal structure.

Lysine encapsulation by **pmclx<sub>4</sub>** resulted in  $\sim 300 \text{ \AA}^2$  of protein surface coverage (Table 2, similar to **sclx<sub>4</sub>**<sup>48</sup>). In addition to masking this surface area, the calixarene contributed a new surface of  $\sim 600 \text{ \AA}^2$  to the particle. This patch size is equivalent to a protein crystal packing interface<sup>202</sup> and, as discussed, provides a surface that can participate in numerous noncovalent interactions with adjacent proteins in the crystal (Figure 5B) consistent with a “molecular glue” functionality.<sup>48,51,52</sup> Notably, the exposed surface of the calixarene is equivalent to  $\sim 10 \%$  of the cytc surface area. Therefore, it appears that protein surface decoration by calixarenes may be a special case of the patchy particle model of protein crystallization.<sup>34,62-64</sup>

**NMR characterization and comparison with the solid state.**

Protein-ligand titrations were performed by the addition of microliter aliquots of 50 mM **pmclx<sub>4</sub>** to <sup>15</sup>N-labelled cytc, which was monitored by <sup>1</sup>H-<sup>15</sup>N HSQC. The overlaid HSQC spectra (Figure 8) revealed increasing chemical shift perturbations ( $\Delta\delta$ ) as a function of the **pmclx<sub>4</sub>** concentration, indicative of fast-intermediate exchange between the ligand-free and -bound states. However, the resonances of Leu85 and Lys87 exhibited strongly biphasic chemical shift perturbations and some resonances were broadened beyond detection at ~4 (Lys4) or ~6 equiv. (Lys11 and Leu85) of **pmclx<sub>4</sub>**. Spectral broadening might indicate intermediate exchange or multiple conformations adopted by these residues. Significant perturbations (<sup>1</sup>H<sup>N</sup>  $\geq$  0.04 ppm or <sup>15</sup>N  $\geq$  0.4 ppm) were observed for 37 resonances, including Lys-2, Lys4, Lys5, Lys73, Lys86, Lys87, Lys89 and Lys100. These residues (except Lys-2), participated in multiple non-covalent interactions at protein-calixarene interfaces in the crystal structure (Figures 4B and 5). While the large NMR effects for Leu85, Lys86 and Lys87 are consistent with the crystal structure (complex formation at Lys86, Figure 3B) there were also some contrasting results. For example, Lys11 did not contribute to complex formation in the crystal structure (Figure 5) yet this resonance was broadened beyond detection. Conversely, Lys54 did not exhibit a significant  $\Delta\delta$  yet complex formation was observed at this residue in the crystal (*vide supra*, Figure 3).

A similar observation was made in the cytc – **sclx<sub>4</sub>** complex in which Lys22, a binding site observed in the crystal structure, was not affected in the NMR.<sup>48</sup> Overall, these results suggest that complex formation occurred with slight differences in the solution and solid states. While binding to Lys54 may be the result of the crystallization conditions/environment other differences may be attributed to the fluxionality of calixarene binding, which likely involves ligand hopping between lysine side chains.<sup>48,52</sup>



**Figure 8.** NMR characterization of **pmclx<sub>4</sub>** binding to cytc. **(A)** Region from overlaid <sup>1</sup>H-<sup>15</sup>N HSQC spectra of 0.1 mM pure cytc (black contours) and in the presence of 0.02–0.64 mM **pmclx<sub>4</sub>** (colored scale). Notes: negligible change to Lys54; biphasic shifts for Leu85 and Lys87; and Lys4, Lys11 and Leu85 broadened beyond detection. **(B)** Binding curves for individual resonances were fit globally to a 1:1 binding model to yield  $K_d \sim 0.04$  mM.

*Apparent* dissociation constants were calculated from titration data collected up to ~6 equiv. of **pmclx<sub>4</sub>**. The binding isotherms (Figure 8B) were hyperbolic and fit to a 1:1 binding model (Equation 3). The average *apparent*  $K_d$  of  $0.04 (\pm 0.01)$  mM was at least 10-fold stronger than for the complex with **sclx<sub>4</sub>**.<sup>48</sup> The tighter binding of **pmclx<sub>4</sub>** agrees with the observation that protein crystallization occurred at approximately ten-fold lower equivalents of this ligand compared to **sclx<sub>4</sub>**. Features such as the increased hydrogen-bonding capacity of phosphonate versus sulfonate, the mobility of the anionic substituent, as well as the increased hydrophobicity of the cavity (upper rim –CH<sub>2</sub>– substituents) likely contribute to the tighter binding of **pmclx<sub>4</sub>**.

## Conclusions

The cytc – **pmclx**<sub>4</sub> crystal structure illustrates the utility of the phosphonomethyl substituent as a contributor to stronger, higher specificity ‘molecular glue’ properties compared to the sulfonate.<sup>48</sup> Cytc – **pmclx**<sub>4</sub> crystal growth was achieved with ~10-fold less ligand than was required to obtain cytc – **sclx**<sub>4</sub> co-crystals. While **sclx**<sub>4</sub> occupied three different sites (Lys4, Lys22 and Lys89) on the protein surface, **pmclx**<sub>4</sub> bound to two different sites (Lys54 and Lys86) suggesting increased specificity of **pmclx**<sub>4</sub> in the solid state.

Interestingly, lysine encapsulation occurred with the presence or absence of a salt bridge to one phosphonomethyl substituent that could rotate into the cavity. In contrast, the sulfonate substituents always formed one or more salt bridges to the lysine. The phosphonomethyl functionalized calixarenes may prove beneficial for achieving crystal growth of cationic proteins, in particular nucleic acid binding proteins. More generally, protein surface decoration by calixarenes may provide the necessary ‘sticky patches’ that facilitate packing contacts and crystal growth.<sup>34,62-64</sup>



## Chapter 2

---

### Protein Frameworks *via* Co-crystallization with an Octa-anionic Calix[4]arene

The material in this chapter is accepted for publication:

**Alex J. M.**, Brancatelli G., Volpi S., Bonaccorso C., Casnati A., Geremia S., Crowley P.B.

Probing the determinants of porosity in protein frameworks: Co-crystals of cytochrome *c* and an octa-anionic calix[4]arene.

*Org. Biomol. Chem.* **2019**. (Accepted)

Alessandro Casnati's Lab (Università degli Studi di Parma, Italy) –

- Synthesis of the oxomethylcarboxylate sulfonato-calix[4]arene (**sclx<sub>4</sub>mc**).

Silvano Geremia's Lab (University of Trieste, Italy) –

- Structure of the horse cytc – **sclx<sub>4</sub>mc** complex.

### **Abstract**

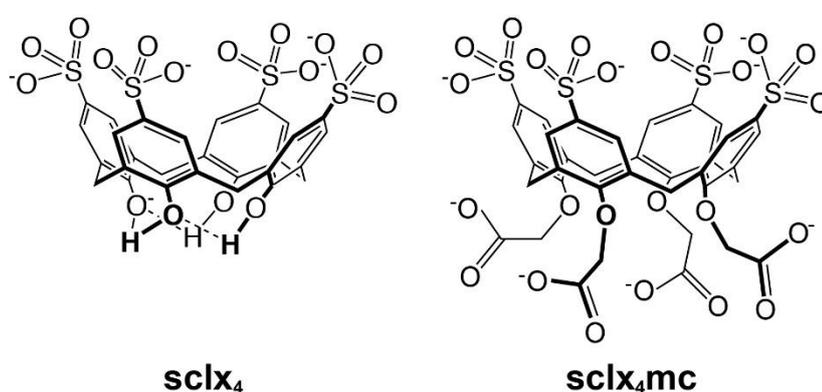
Sulfonato-calix[n]arenes (**sclx<sub>n</sub>**) are promising tools to generate crystalline protein frameworks. Two crystal structures of an octa-ionic calix[4]arene (**sclx<sub>4</sub>mc**) in complex with yeast or horse heart cytochrome *c* (cytc) are reported. The calixarene bound to a similar site on each protein but different assemblies occurred, with a honeycomb arrangement of yeast cytc and a tubular assembly of horse cytc. The yeast cytc assembly was more porous (~75 % solvent content) than the horse cytc case (~55 %). High porosity structures have been obtained previously with **sclx<sub>8</sub>** and **sclx<sub>6</sub>** but not with **sclx<sub>4</sub>**, suggesting that the charge of the ligand is a contributing factor to porous architectures. We discuss how the choice of (customized) calixarene may dictate protein assembly.

## Introduction

Protein-based materials have the potential to transform society as they are biodegradable, biocompatible, and have tunable functions.<sup>2,13-15,203-205</sup> Consequently, there is a growing interest in the design and fabrication of protein architectures that can be harnessed for nanoscale applications.<sup>2,13,205</sup> Porous protein assemblies are particularly appealing as reaction vessels (bionanoreactors) or delivery systems.<sup>2,203</sup> Protein architectures can be engineered by a great variety of strategies, including protein fusion,<sup>13</sup> disulfide bridges,<sup>14,204</sup> metal-coordination<sup>17</sup> or coiled coil appendages.<sup>18</sup> Assembly can be mediated also by charge-charge interactions, as demonstrated recently with super-charged forms of green fluorescent protein (highly cationic or anionic).<sup>15</sup> Highly charged synthetic ligands are also proving instrumental to dictate protein assembly.<sup>19,20,48-56,143</sup> For example, anionic calixarenes have been used as scaffolds to assemble cationic proteins such as cytochrome *c* (cytc),<sup>19,20,48,51-54</sup> lysozyme<sup>49,50</sup> and PAF.<sup>55</sup> These supramolecular scaffolds function like ‘molecular glue’ by forming interfaces with two or more proteins.

The sulfonato-calix[*n*]arenes (**sclx<sub>n</sub>**) are highly water soluble anionic receptors, with a hydrophobic core and an anionic rim, suited to encapsulation of Arg or Lys residues.<sup>70,182</sup> While calix[4]arenes are rigid cones, the calix[6,8]arenes are flexible.<sup>182</sup> Owing to their ‘floppiness’ these ligands can mould to the protein surface and form large interfaces. Both **sclx<sub>6</sub>** and **sclx<sub>8</sub>** can induce highly porous cytc assemblies. While **sclx<sub>6</sub>** yielded a honeycomb arrangement (~65 % solvent content),<sup>56</sup> **sclx<sub>8</sub>** mediated a high-porosity crystalline framework (~85 % solvent content).<sup>19,20</sup> Interestingly, it was recently found that the presence of spermine modulated the porosity of cytc – **sclx<sub>8</sub>** assembly.<sup>20</sup> Earlier, it was observed that the length of the oligoethylene glycol linker in a sugar–dye conjugate influenced the extent of porosity in crystalline frameworks of concanavalin A.<sup>151</sup> Also, other macrocyclic ligands such as pillararenes have illustrated their potential to induce porous architecture.<sup>22</sup>

Here, the assembly-inducing behaviour of **sclx<sub>4</sub>mc** was investigated. This compound is a **sclx<sub>4</sub>** derivative with four carboxylate functionalities at the lower rim (Figure 1).<sup>206-209</sup> **sclx<sub>4</sub>mc** has been studied in complex with cationic porphyrins but not with proteins.<sup>210-212</sup> Its structural similarity to **sclx<sub>4</sub>** may aid in understanding how the lower rim substituents influence protein recognition and assembly.<sup>48</sup> The presence of four chelating oxomethylcarboxylate (O-CH<sub>2</sub>-COO<sup>-</sup>) units confers the ability of these podand-like calixarenes to encapsulate metal ions at the lower rim.<sup>211</sup>



**Figure 1.** Molecular structures of *p*-sulfonatocalix[4]arene (**sclx<sub>4</sub>**), and the carboxylate derivative (**sclx<sub>4</sub>mc**).

Metal complexation rigidifies the cone structure (prevents “breathing of the calix”)<sup>182</sup> and enhances the binding of small neutral/cationic guests in the calix[4]arene cavity.<sup>206-209</sup> Furthermore, the presence of charged groups at both rims of the calix[4]arene gives rise to a bolaamphiphile structure (Figure 1). This feature can potentially increase the ‘molecular glue’ activity of the calixarene. In this direction, **sclx<sub>4</sub>mc** was co-crystallized with yeast or horse heart cytc. These highly cationic proteins (*pI* ~9.5) share 65 % sequence identity (Figure 2).<sup>159,213</sup>

```

Yeast cytc  AEFKA  1  GSAKKGATLF  10  KTRCLQCHTV  20  EKGGPHKVGP  30
Horse cytc  ----- GDVEKGGKIF  VQKCAQCHTV  EKGGKHKHTGP
              *   **   *           * * * * *   * * * * *   * * * *
              *   * *   *           * * * * *   * * * * *   * * * *

Yeast cytc  NLHGIFGRHS  40  GQAEGYSYTD  50  ANIKKNVLWD  60
Horse cytc  NLHGLFGRKT  GQAPGFTYTD  ANKNKGITWK
              * * * * *   * * * *   * * * *   * *   *   *

Yeast cytc  ENNMSEYLTN  70  PKKYIPGTKM  80  AFGGLKKEKD  90
Horse cytc  EETLMEYLEN  PKKYIPGTKM  IFAGIKKTE
              *   * * * *   * * * * * * * * * *   *   *   *

Yeast cytc  RNDLITYLKK  100  ATE-
Horse cytc  REDLIAYLKK  ATNE
              * * * *   * * * *   * *
    
```

**Figure 2.** Alignment of the yeast and horse heart cytc primary structures (MUSCLE).<sup>214</sup> Conserved residues are highlighted with an asterisk.

## Experimental

**Sample preparation.** **sclx<sub>4</sub>mc** was synthesized, as described.<sup>206-209</sup> Millimolar stock solutions of the ligand were prepared in water at pH 6.0. Unlabelled *Saccharomyces cerevisiae* (yeast) cytc (C102T) was produced as reported,<sup>48,191</sup> and horse heart cytc was from Sigma-Aldrich.

**Co-crystallization trials.** An Oryx 8 Robot (Douglas Instruments) and a sparse matrix screen (JCSG++, Jena Bioscience) was used for co-crystallization of yeast cytc and **sclx<sub>4</sub>mc** at 20° C. Protein and calixarene were tested at ratios of 1:1, 1:5 and 1:25. Crystals grew only in 1:1 ratio in condition F2 comprising 3.15 M ammonium sulfate and 0.1 M sodium citrate pH 5.0. Horse cytc and **sclx<sub>4</sub>mc** were co-crystallized by the hanging drop vapour diffusion method at 20° C, from 52-62 % PEG 3350 and 0.05 M sodium cacodylate at pH 5.5. Drops were prepared by mixing 1 µL of 1.7 mM protein, 0.5 µL of 17 mM ligand and 1.7 mM gadolinium chloride, with 1 µL of reservoir solution. Crystal growth occurred in 56 % PEG 3350 after 4 days.

**X-ray data collection.** Crystals of ~100  $\mu\text{m}$  dimension were cryo-protected in the reservoir solution supplemented with 20 % glycerol and cryo-cooled in liquid nitrogen. Diffraction data were collected to 1.7  $\text{\AA}$  resolution for the yeast cytc – **sclx<sub>4</sub>mc** crystal using  $\phi$  scans of 0.1° over 360° with an Eiger X 9M detector (PROXIMA-2A, SOLEIL synchrotron). A dataset extending to 2.5  $\text{\AA}$  resolution was collected for the horse cytc – **sclx<sub>4</sub>mc** crystal using  $\phi$  scans of 0.5° over 100° with a PILATUS detector (XRD1 beamline, Elettra Synchrotron).

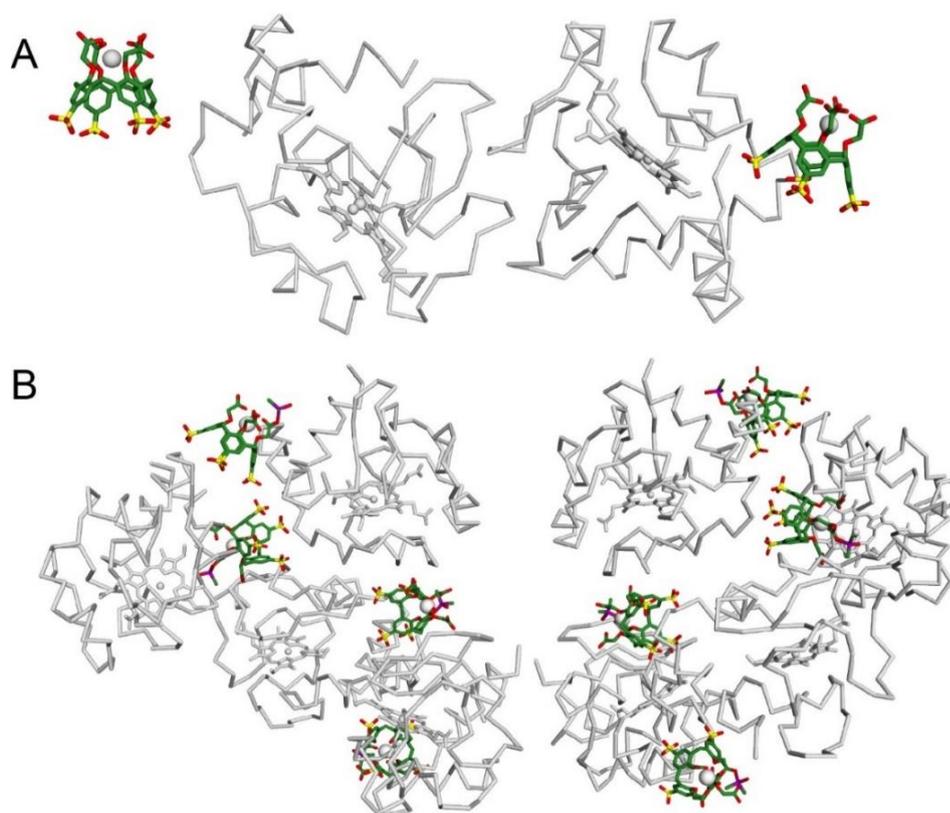
**X-ray structure determination.** The observed reflections were processed with the autoPROC pipeline<sup>215</sup> (yeast cytc – **sclx<sub>4</sub>mc**) or XDS<sup>192</sup> (horse cytc – **sclx<sub>4</sub>mc**). In both cases the data were scaled using POINTLESS<sup>193</sup> and AIMLESS.<sup>194</sup> *Xtriage* (PHENIX)<sup>216</sup> analysis of the horse cytc – **sclx<sub>4</sub>mc** dataset indicated a perfect merohedral twin and a twin law of h, -k, -l was required for refinement. The two structures were solved by molecular replacement in PHASER<sup>195</sup> using 5lyc (yeast cytc) or 1hrc (horse cytc) as the search models. The calixarene coordinates and restraints were generated in JLigand.<sup>196</sup> Iterative cycles of model building in COOT<sup>197</sup> and refinement were performed with BUSTER<sup>217</sup> (yeast cytc – **sclx<sub>4</sub>mc**) or REFMAC5<sup>198</sup> (horse cytc – **sclx<sub>4</sub>mc**) until no further improvements in the  $R_{\text{free}}$  or electron density were obtained (Table 1). The final structures were validated with MolProbity<sup>199</sup> and deposited in the Protein Data Bank as PDB 6suy (yeast cytc – **sclx<sub>4</sub>mc**) and PDB 6suv (horse cytc – **sclx<sub>4</sub>mc**).

## Results and Discussion

**Cytc – sclx<sub>4</sub>mc co-crystallization.** A sparse matrix screen (Jena Bioscience, JCSG++) was used for co-crystallization trials of yeast cytc and **sclx<sub>4</sub>mc**. Compared to 10 equiv. **sclx<sub>4</sub>** required for crystal growth,<sup>48</sup> 1 equiv. **sclx<sub>4</sub>mc** was sufficient. The crystals grew in condition F2, 3.15 M ammonium sulfate and 0.1 M sodium citrate pH 5.0. Previously, highly porous cytc – **sclx<sub>8</sub>** crystals (PDB 6gd9)<sup>19</sup> were obtained with ~2 M ammonium sulfate whereas cytc – **sclx<sub>6</sub>** crystals (PDB 6rgi)<sup>56</sup> were obtained in 10 %

PEG 8000 and 0.1 M imidazole pH 8. The horse cytc – **sclx<sub>4</sub>mc** co-crystals were obtained with a 1:10 ratio of protein:ligand in 56 % PEG 3350 and 0.05 M sodium cacodylate pH 5.5.

**Data collection and model building.** Datasets extending to 1.7 Å or 2.5 Å resolution were collected from the yeast cytc – **sclx<sub>4</sub>mc** (SOLEIL synchrotron) or the horse cytc – **sclx<sub>4</sub>mc** (Elettra synchrotron) crystals, respectively. The former belonged to the trigonal *P*<sub>3</sub><sup>2</sup><sub>1</sub> and the latter to the tetragonal *P*<sub>4</sub><sub>3</sub> space groups (Table 1). The structures were solved by molecular replacement with asymmetric units comprising 2 yeast or 8 horse cytc molecules (Figure 3).



**Figure 3.** The asymmetric units of (A) yeast cytc – **sclx<sub>4</sub>mc** and (B) horse cytc – **sclx<sub>4</sub>mc** comprising 2 or 8 proteins, respectively. The protein chains, calixarenes and sodium ions are shown in grey, green, and white.

**Table 1.** X-ray data collection, processing and refinement statistics

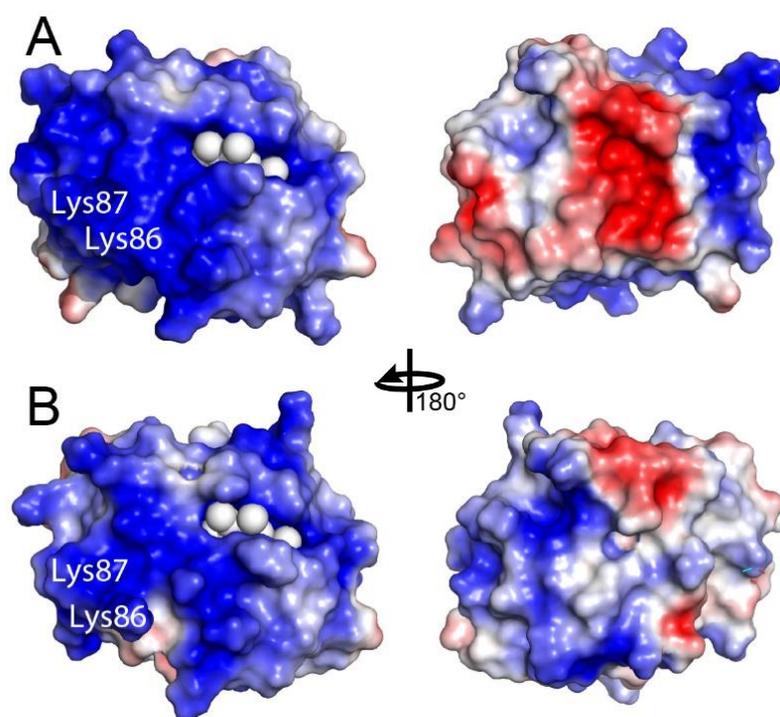
| <i>Data Collection</i>                            |  |  |
|---|--|--|
|   | Yeast cytc – <b>sclx<sub>4</sub>mc</b>   | Horse cytc – <b>sclx<sub>4</sub>mc</b>   |
| Light source                                      | SOLEIL, PROXIMA-2A   | Elettra, XRD1  |
| Wavelength (Å)                                    | 0.97624  | 1.0000   |
| Space group                                       | <i>P</i> 3 <sub>2</sub> 21   | <i>P</i> 4 <sub>3</sub>  |
| Cell constants                                    | <i>a</i> = <i>b</i> = 102.48 Å<br><i>c</i> = 180.00 Å<br>$\alpha = \beta = 90^\circ, \gamma = 120^\circ$ | <i>a</i> = <i>b</i> = 65.59 Å<br><i>c</i> = 250.69 Å<br>$\alpha = \beta = \gamma = 90^\circ$ |
| Resolution (Å)                                    | 56.00-1.74 (1.77-1.74)   | 46.38 – 2.50 (2.60-2.50)   |
| # reflections                                     | 904004 (42248)   | 94255 (9021)   |
| # unique reflections                              | 44640 (2207)   | 34903 (3775)   |
| Multiplicity                                      | 20.3 (19.1)  | 2.7 (2.4)  |
| <i>I</i> / $\sigma$ ( <i>I</i> )                  | 15.7 (2.2)   | 5.9 (2.6)  |
| Completeness (%)                                  | 100 (100)  | 96.3 (92.6)  |
| <i>R</i> <sub>pim</sub> <sup>b</sup> (%)          | 3.4 (44.0)   | 10.1 (24.3)  |
| CC <sub>1/2</sub>                                 | 99.8 (77.9)  | 94.3 (73.9)  |
| Solvent content (%)                               | 73   | 57   |
| <i>Refinement</i>                                 |  |  |
| <i>R</i> <sub>work</sub>                          | 0.1665   | 0.1754   |
| <i>R</i> <sub>free</sub>                          | 0.1790   | 0.2365   |
| rmsd bonds (Å)                                    | 0.0101   | 0.0114   |
| rmsd angles (°)                                   | 1.0858   | 3.6120   |
| asymmetric unit composition                       |  |  |
| protein   | 2  | 8  |
| <b>sclx<sub>4</sub>mc</b>                         | 2  | 8  |
| sodium ion  | 2  | 8  |
| acetate   | 0  | 8  |
| cacodylate  | 0  | 8  |
| sulfates  | 5  | 0  |
| water   | 280  | 423  |
| Ave. B-factor (Å <sup>2</sup> )                   | 31.19  | 25.67  |
| Clashscore  | 1.09   | 6.40   |
| Ramachandran analysis, <sup>c</sup> % residues in |  |  |
| favoured regions                                  | 98.1   | 97.9   |
| allowed regions                                   | 100  | 100  |
| PDB code  | 6suy   | 6suv   |

<sup>a</sup>Values in parentheses correspond to the highest resolution shell

<sup>b</sup> $R_{pim} = \frac{\sum_{hkl} \sqrt{(1/n-1) \sum_{i=1}^n |I_i(hkl) - \langle I(hkl) \rangle|}}{\sum_{hkl} \sum_i I_i(hkl)}$

<sup>c</sup>Determined in MolProbity

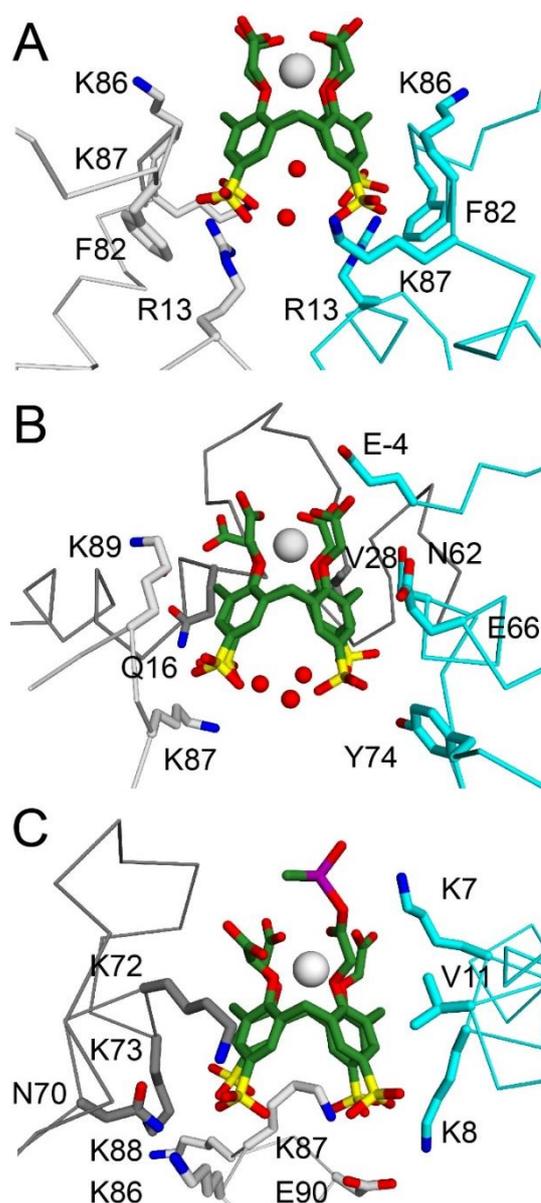
**Cytc – sclx<sub>4</sub>mc crystal structure.** In both structures, the calixarene was bound to a cationic patch on Lys87 (Figures 4 and 5). A sodium ion chelated the lower rim of the calixarene.<sup>206-212</sup> Additionally, a cacodylate was bound to a carboxylate substituent in the horse cytc – sclx<sub>4</sub>mc complex (Figure 5C).



**Figure 4.** The electrostatic surface potentials showing the cationic (blue) and anionic (red) patches of **(A)** yeast cytc and **(B)** horse cytc (APBS Electrostatics, PyMOL). The conserved residues Lys86 and Lys87 which comprise the binding site in both cytc variants are indicated. Heme edge is shown as spheres

Encapsulation of an Arg or Lys side chain is the usual protein recognition mode of anionic calix[4]arenes.<sup>48,-50,52-54</sup> In the yeast cytc – sclx<sub>4</sub>mc crystal structure, the calixarene did not entrap any side chain. Instead, 2-3 waters occupied the cavity and formed hydrogen bonds with the sulfonates or a  $\pi$ -hydrogen bond with a phenyl ring (Figures 5A and B).<sup>218</sup> Two types of protein – sclx<sub>4</sub>mc complex were evident in the structure, with distinct binding site features (Figures 5A and B). One calixarene

was present at a  $C_2$  symmetric interface, flanked by four residues (Arg13, Phe82, Lys86 and Lys87) from each protein (Figure 5A) whereas the other ligand bound to three neighbouring proteins (Figure 5B). Each molecule of **sclx<sub>4</sub>mc** accounted for cytc surface coverage of  $\sim 550 \text{ \AA}^2$  (PDBePISA<sup>219</sup>).

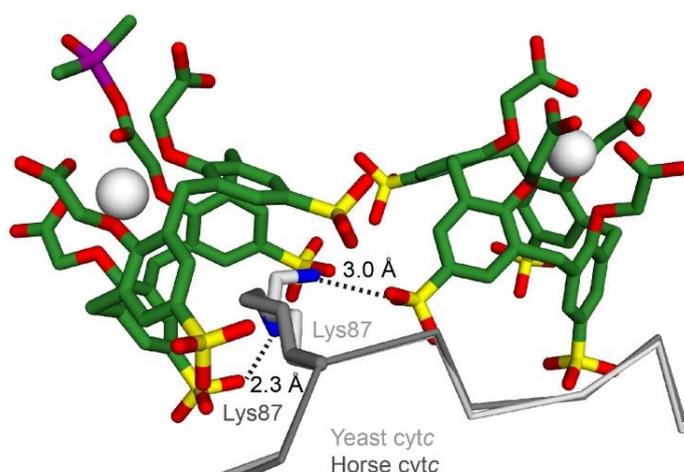


**Figure 5.** The interfaces formed by **sclx<sub>4</sub>mc** highlight the ‘molecular glue’ property. In the yeast cytc case one calixarene mediated (A) a  $C_2$  symmetric interface and (B) a second calixarene interacted with three protein chains. The hydrophobic cavity was

occupied by water molecules (red spheres) and did not encapsulate any residue. (C) In the horse cytc complex the calixarene encapsulated Lys87 and was surrounded by three neighbouring proteins. Protein chains are indicated as light grey, dark grey or cyan ribbon. Calixarenes and side chains are shown as sticks. Sodium ions are shown as spheres.

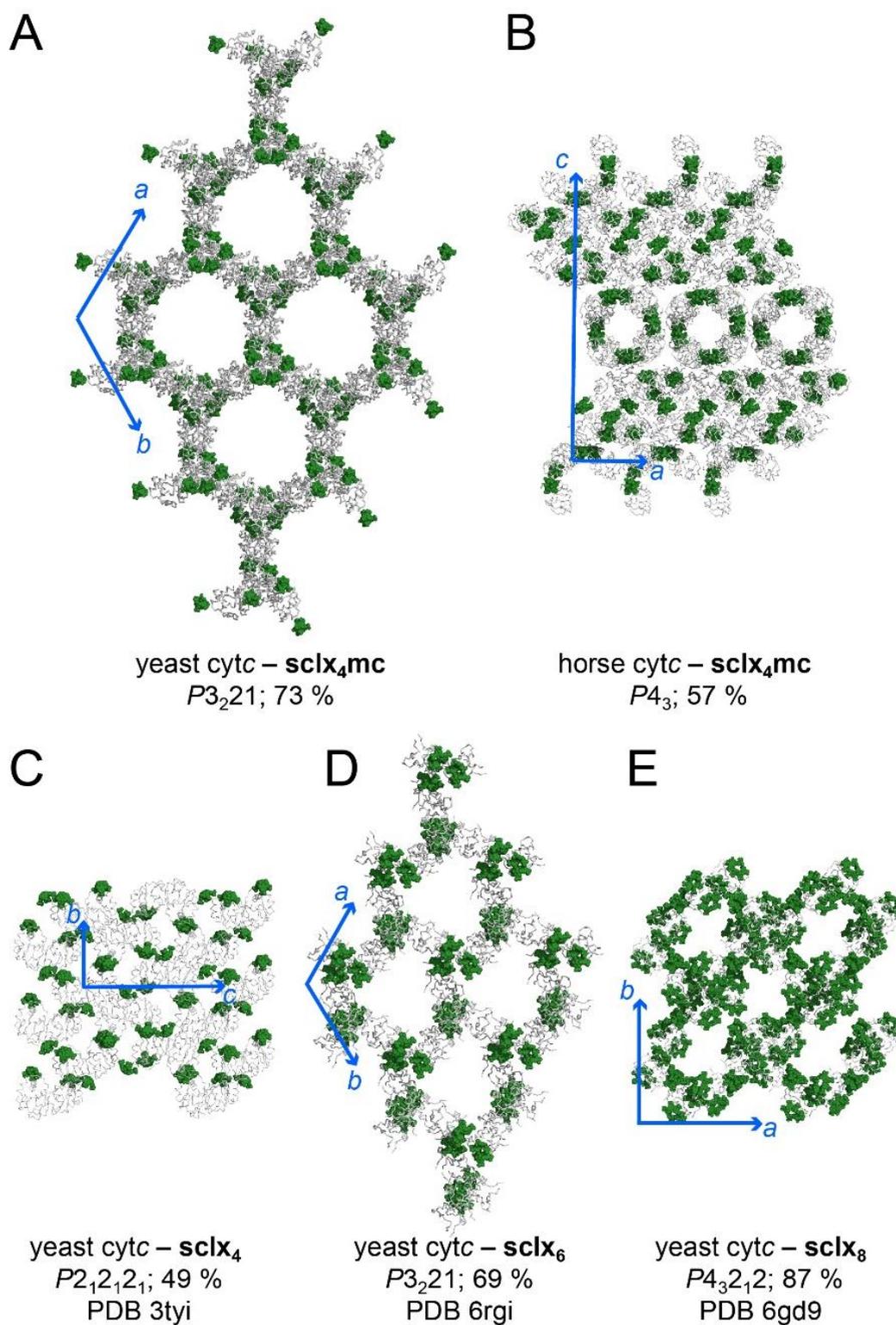
Recognition of yeast cytc by **sclx<sub>4</sub>mc** differed markedly from that by **sclx<sub>4</sub>**.<sup>48</sup> In the latter case, the calixarene encapsulated a single lysine side chain (Lys4, Lys22 or Lys89) and formed a salt bridge with the ammonium group. In contrast, **sclx<sub>4</sub>mc** did not encapsulate any of the binding site residues (Figures 5A and B). Unlike **sclx<sub>4</sub>**, this ligand formed only weak CH- $\pi$  interactions with the neighbouring Lys86 or Lys89 and no cation- $\pi$  bonds. Also, the **sclx<sub>4</sub>** methylene groups (-CH<sub>2</sub>-) established contacts with the C <sup>$\beta$</sup>  of Ala3, Ala7 and Ala101, as well as the ring edge of Tyr97. These residues comprise a hydrophobic patch on yeast cytc that formed interfaces with other calixarenes.<sup>48,52,54</sup> However, this interaction was absent in the cytc – **sclx<sub>4</sub>mc** complex. Instead, mainly Phe82 interacted with this ligand.

In the horse cytc – **sclx<sub>4</sub>mc** structure the mode of recognition varied significantly from the yeast cytc complex (Figures 5C and 6). Here, the Lys87 side chain was encapsulated by the calixarene, which formed three interfaces of  $\sim 710 \text{ \AA}^2$  comprising Lys7-Lys8, Lys72-Lys73, or Lys86-Lys87-Lys88 (Figure 5C). Dilysine motifs of yeast cytc are well-known calixarene-binding sites. For example, Lys4-Lys5,<sup>53,19</sup> Lys72-Lys73<sup>20</sup> and Lys86-Lys87.<sup>48,53,20</sup>



**Figure 6.** Superposition of yeast (light grey) and horse cytc (dark grey) highlights the binding of **sclx<sub>4</sub>mc** to a similar site on both proteins. While Lys87 is encapsulated by the calixarene in the horse cytc, it interacted laterally with the ligand in the yeast cytc.

**Sclx<sub>4</sub>mc** has an available surface area of  $\sim 900 \text{ \AA}^2$ . In both the yeast and horse cytc complexes the interaction with a single protein masked  $220\text{--}270 \text{ \AA}^2$  of the ligand surface. Of the remaining  $630\text{--}680 \text{ \AA}^2$  calixarene surface, 50 % (in yeast cytc) or 70 % (in horse cytc) participates in forming interfaces. The larger surface coverage in case of horse cytc may be attributed to the encapsulation of the Lys residue. The size similarity of the binding site patches to a typical protein – protein interface signifies their role as a ‘molecular glue’ for mediating intermolecular crystal contacts with the surrounding proteins (Figure 7).<sup>19,20,48-56,202</sup>



**Figure 7.** Crystal packing in cytc – calixarene complexes. **(A)** Honeycomb arrangement of yeast cytc – **sclx<sub>4</sub>mc**. **(B)** Tubular assembly of horse cytc – **sclx<sub>4</sub>mc**. **(C)** A layer assembly of yeast cytc – **sclx<sub>4</sub>**.<sup>48</sup> **(D)** Honeycomb arrangement of yeast cytc – **sclx<sub>6</sub>**.<sup>56</sup>

(E) Framework of yeast cytc – **sclx<sub>8</sub>** with all crystal contacts *via* the ligand.<sup>19</sup> The space groups and % solvent contents are indicated. Proteins, calixarenes and unit cell axes are grey, green and blue, respectively.

Considering the packing of cytc by calix[4]arenes,<sup>48-50,52,53</sup> the porous architectures induced by **sclx<sub>4mc</sub>** were a surprising result (Figure 3). A tubular assembly ( $P4_3$ , ~55 % solvent content)<sup>220</sup> was obtained with the horse cytc. The tube structure is based on eight cytc – **sclx<sub>4mc</sub>** complexes arranged in a ring, with an internal diameter of ~2 nm. These tubes are arranged in orthogonal layers. The assembly of yeast cytc ( $P3_221$ , ~75 % solvent content)<sup>220</sup> was a honeycomb network with ~7 nm diameter pores. In each case, the assemblies involved edges that are one protein/calixarene unit thick but the yeast cytc packing was more porous. Interestingly, **sclx<sub>6</sub>** also mediated a honeycomb arrangement of yeast cytc ( $P3_221$ , ~65 % solvent content).<sup>56</sup> Despite this resemblance, the pores were ~1.8 times narrower compared to that in the yeast cytc – **sclx<sub>4mc</sub>** structure (~3.8 nm diameter; Figure 3D). Compared to these ligands, **sclx<sub>8</sub>** behaved quite distinctly. In one of the yeast cytc – **sclx<sub>8</sub>** structures, all of the crystal contacts were mediated by the calixarene leading to a crystalline framework ( $P4_32_12$ ) with exceptionally high porosity of ~85 % (Figure 3E).<sup>19</sup>

A common feature of **sclx<sub>4mc</sub>**, **sclx<sub>6</sub>** and **sclx<sub>8</sub>** is the high formal charge, even though the conformations are quite different. Also, the co-crystallization conditions may have played a part in facilitating the porous assemblies. For example, at the crystallization pH of 5.0 the phenolic hydroxyls of **sclx<sub>4mc</sub>** are likely to be deprotonated.<sup>210</sup> Two of the phenolic OH of **sclx<sub>6</sub>** and **sclx<sub>8</sub>** will be deprotonated at the co-crystallization pH of 8 (cytc – **sclx<sub>6</sub>**)<sup>56</sup> or 7 (cytc – **sclx<sub>8</sub>**),<sup>19</sup> resulting in a charge of -8 and -10, respectively.<sup>221,222</sup> The similarity of the total anion charge of **sclx<sub>4mc</sub>** and **sclx<sub>6</sub>** may have been instrumental in facilitating similar porous assemblies. Earlier, crystalline protein frameworks (~70 % solvent content) were obtained with

concanavalin A and dye-sugar conjugates.<sup>151</sup> Similarly, porous assemblies of the jacalin lectin were also induced by an anionic porphyrin.<sup>143</sup> Here, also electrostatic interactions between the protein – ligand interfaces appear to be an important contributor towards such assemblies. The significance of charge – charge interaction in yielding porous crystalline framework was also observed in the complex of cationic pillar[5]arene and anionic apoferritin protein cages.<sup>22</sup>

In view of the various instances of ligand-mediated protein assemblies, probably the charge modification and subsequent electrostatic interactions at the protein – calixarene interfaces facilitate crystal porosity (Figure 4).<sup>223</sup> However, crystal structures of the ~6.2 kDa cationic PAF protein (*pI* ~8.9) with **sclx<sub>4</sub>**, **sclx<sub>6</sub>** or **sclx<sub>8</sub>** all involved tightly packed sheet-like assemblies.<sup>55</sup> Thus, it remains to be seen whether this property of the calix[4,6,8]arenes is restricted to cytc or transferable to other proteins.

## Conclusions

This work demonstrates the use of bolaamphiphilic **sclx<sub>4</sub>mc** to mediate honeycomb (yeast cytc) or tubular (horse cytc) architectures. This contrast with the assembly mediated by other calix[4]arenes but the similarity with that dictated by **sclx<sub>6</sub>** and **sclx<sub>8</sub>** might suggest a role for the high anionic charge of these molecular glues. In this respect, the carboxylate functionality may have been useful as it enhanced the anionic charge of the calix[4]arene, while the rigid cone structure obtained upon sodium ion coordination might also play a role. **Sclx<sub>4</sub>mc** adds to the existing library of supramolecular building blocks, which can be utilized to facilitate unique protein architectures. Furthermore, the details of molecular recognition presented for the two complexes may provide structural clues for functionalizing calixarenes (substituents and formal charge) to target the desired protein. Finally, there is scope to explore the potential of binding polyanionic calixarenes with other proteins, including neutral / acidic candidates, to generate porous structures.



## Chapter 3

---

### Supramolecular Stacking in a High Z' Calix[8]arene – Porphyrin Assembly

The material in this chapter was published as

**Alex J. M.**, McArdle P., Crowley P. B.

Supramolecular stacking in a high Z' calix[8]arene – porphyrin assembly.

*CrystEngComm.* **2019.** *In print.*

Dr. Patrick McArdle (NUI Galway) helped with the ShelxL refinement.

**Abstract**

A co-crystal structure of sulfonato-calix[8]arene (**sclx<sub>8</sub>**) and trimethylanilinium-porphyrin (**tmap**) at 1.0 Å resolution is reported. The oppositely charged macrocycles form a stacked assembly of alternating calixarene and porphyrin. Crystal packing was completed by additional porphyrins that interlinked neighbouring stacks. This structure provide insights into macrocycle-based assemblies.

## Introduction

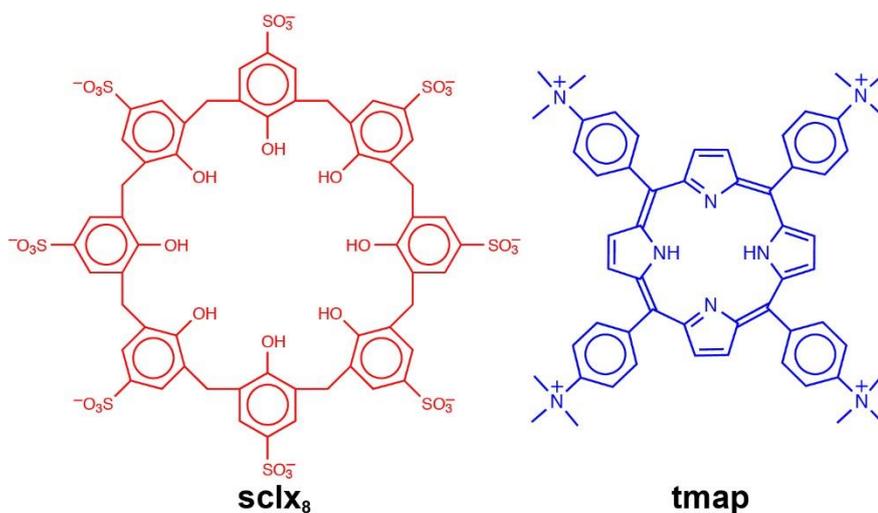
Calixarenes<sup>73,74</sup> and porphyrins<sup>224,225</sup> are versatile supramolecular building blocks. The vast possibilities for customization make these macrocycles useful molecular recognition agents. For example, the upper and lower rims of calixarenes can be functionalized to favour host-guest interactions with biomolecules.<sup>182</sup> Similarly, the periphery of the porphyrin ring (pyrrolic  $\beta$  sites or meso-positions) can be substituted with diverse groups.<sup>224,225</sup> This possibility of structural manipulation has afforded a wide-range of applications including sensing,<sup>71,226</sup> catalysis,<sup>227,228</sup> and light harvesting systems.<sup>229-231</sup>

These molecules have been utilized for fabricating unique architectures *via* self-assembly.<sup>7</sup> For example, sheet-like,<sup>68</sup> columnar,<sup>232</sup> or layered<sup>233</sup> calixarene structures and flowers,<sup>234</sup> rings<sup>235</sup> or pillars<sup>236</sup> of porphyrins. Furthermore, porphyrin – calixarene complexation has led to generation of hybrid assemblies including porous frameworks.<sup>210-212,237-241</sup> In these complexes, calixarenes act as a templating agent to aid the porphyrin assembly. For example, the interaction of tetra(*N*-methyl-4-pyridyl)-porphyrin (TMPyP) with a calix[4]arene derivative resulted in complexes with varying calixarene:porphyrin ratios.<sup>210-212</sup> Moreover, these calixarene – porphyrin architectures could be modulated by pH<sup>211</sup> or by the presence of metal ions.<sup>212,240</sup>

Additionally, both these synthetic ligands are known to facilitate assembly of proteins. Solution-state NMR<sup>242</sup> and X-ray crystallographic<sup>19,20,48-54,143,144</sup> studies have revealed their ability to act as ‘molecular glues’. When bound to a protein, they are able to establish non-covalent contacts and form interfaces with the surrounding protein molecules. For example, the highly water soluble anionic sulfonatocalix[8]arene (**sclx<sub>8</sub>**) have been demonstrated to form large interfaces with cytochrome *c* (cytc) and mediate highly porous crystalline frameworks of cytc (70 -85 % solvent content).<sup>19,20</sup>

We sought to investigate the ability of **sclx<sub>8</sub>** to bind to two cationic partners – cytc and tetra(4-N,N,N-trimethylanilinium) porphyrin (**tmap**; Figure 1). There are at least ten structures of **sclx<sub>8</sub>** illustrating its binding with different guests.<sup>243-250</sup> Though, no X-ray data is available for **tmap**, crystal structures of the structurally-related tetra(4-aminophenyl)porphyrin<sup>251,252</sup> highlight their potential to generate interesting assemblies.

Here, we report a 1.0 Å resolution crystal structure that revealed a remarkable stacking architecture of **sclx<sub>8</sub>** and **tmap**. Interestingly, the porphyrin occurred in two distinct conformations depending on its position/function within the stack. The molecular recognition features that appear to stabilize the assembly are discussed. The current study may prove useful in understanding the behaviour of larger calixarenes with porphyrins with an aim to fabricate hybrid structures.



**Figure 1.** Molecular structures of the ions used in this study.

## Experimental

**Materials.** Samples of unlabelled *Saccharomyces cerevisiae* cytc C102T were produced by the established methods.<sup>48,191</sup> Stock solutions of **tmap** (Frontier Scientific T973) and **sclx<sub>8</sub>** (TCI Chemicals S0471) at pH 6.0 were prepared in 10 mM NaOH or water, respectively.

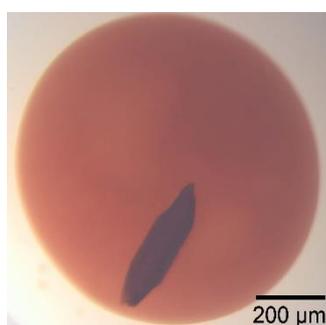
**Co-crystallization Trials.** A sparse matrix screen (JCSG++ Jena Biosciences) and an Oryx 8 robot (Douglas Instruments) were used for the crystallization experiments. Protein – ligand mixture were prepared by combining 1 mM cytc, 6.5 mM **tmap** and 5 – 20 mM **sclx<sub>8</sub>**. A single crystal grew in condition A11 (50 % 2-methyl-2,4-pentanediol, 0.1 M TRIS-HCl pH 8.5 and 0.2 M ammonium di-hydrogen phosphate) at 10 mM **sclx<sub>8</sub>**.

**Data collection.** Crystals of ~400 µm dimension in the mother liquor were cryo-cooled in liquid nitrogen. Diffraction data were collected at SOLEIL synchrotron (France) to 1.0 Å using  $\phi$  scans of 0.1° over 360° using an Eiger X 9M detector.

**X-ray structure determination.** The observed reflections were processed with the autoPROC pipeline<sup>215</sup> and scaled using POINTLESS<sup>193</sup> and AIMLESS.<sup>194</sup> *ab initio* phasing in ACORN (CCP4 suite) was used to generate the map,<sup>253</sup> with unambiguous density for both **tmap** and **sclx<sub>8</sub>**. The coordinates and restraints for **tmap** and **sclx<sub>8</sub>** were generated using the Grade Web Server. Iterative cycles of model building in COOT<sup>197</sup> and refinement (full-matrix least square on  $F^2$ ) were performed in ShelXL (Shelx suite).<sup>254</sup> The PLATON/SQUEEZE<sup>255</sup> procedure was implemented to remove waters which led to an improvement in the refinement statistics. Crystallographic data for the structure with and without water is deposited with the Cambridge Crystallographic Data Centre (CCDC) with deposit numbers 1956108 and 1956128, respectively.

## Results and Discussion

**Cytc – sclx<sub>8</sub> – tmap co-crystallization trials.** A sparse matrix screen (JCSG++ Biosciences) and an Oryx 8 Robot (Douglas Instrument) were used for cytc – sclx<sub>8</sub> – tmap co-crystallization trials. 1 mM cytc, 10 mM sclx<sub>8</sub> and 6.5 mM tmap yielded a single crystal in condition A11 comprising 50 % 2-methyl-2,4-pentanediol (mpd), 0.1 M TRIS pH 8.5 and 0.2 M ammonium di-hydrogen phosphate (Figure 2).



**Figure 2.** A single ~400 μm crystal grew in a protein-rich phase (red color).

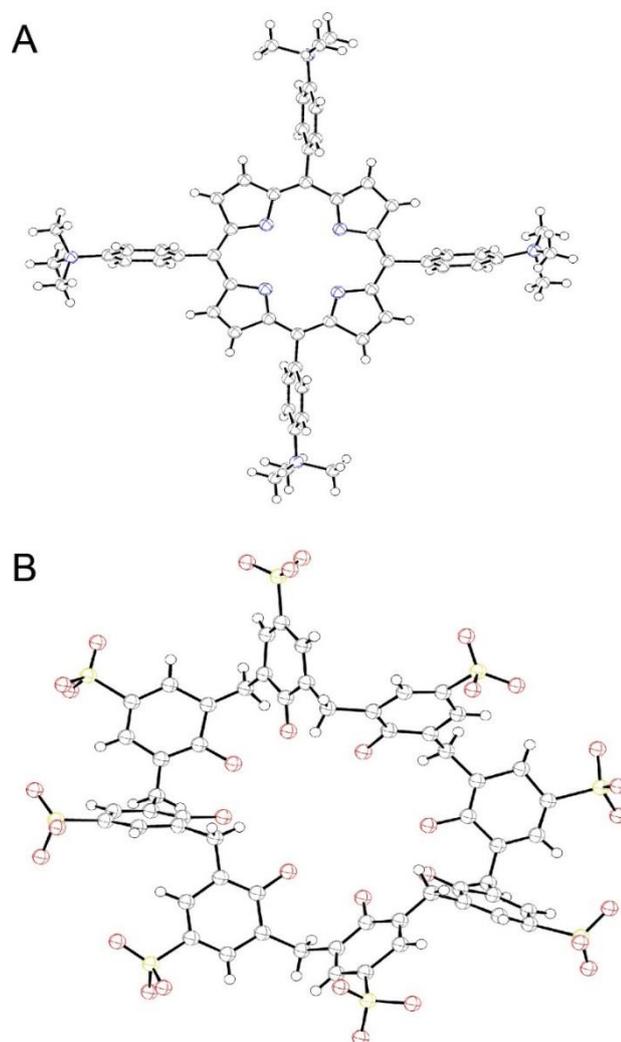
**Data collection and model building.** A dataset extending to 1.0 Å resolution was collected at SOLEIL synchrotron. The crystal belonged to the monoclinic  $P12_11$  space group (Tables 1 and 2). The electron density map was generated *ab initio* using ACORN (CCP4 suite).<sup>253</sup> Unambiguous electron density was evident for the two macrocycles. Interestingly, the asymmetric unit comprised four sclx<sub>8</sub>, six tmap and ten mpd molecules and indicated that the complex crystallized with a high  $Z' = 22$  (Figures 3 and 4).<sup>256-258</sup> Considering the small number of structures with  $Z' \geq 16$  that are well-characterized, the recognition features of the complex may provide knowledge about conditions / interactions that lead to other such structures.<sup>256,257</sup>

**Table 1.** X-ray data collection and refinement statistics for **sclx<sub>8</sub> – tmap** (with waters).

|                                       |  |
|---------------------------------------|--|
| Empirical formula                     | C <sub>313</sub> H <sub>317</sub> N <sub>24</sub> O <sub>149.50</sub> S <sub>16</sub>  |
| Formula weight                        | 7319.85  |
| Temperature (K)                       | 293(2)   |
| Wavelength (Å)                        | 0.82656 Å  |
| Crystal system                        | monoclinic   |
| Space group                           | <i>P</i> 2 <sub>1</sub>  |
| Unit Cell dimensions                  | a = 33.621(17) Å<br>b = 38.411(13) Å<br>c = 40.234(15) Å<br>α = γ = 90° β = 102.56(4)° |
| V (Å <sup>3</sup> )                   | 50715 (37)   |
| ρ <sub>cal</sub> (g/cm <sup>3</sup> ) | 0.959  |
| μ (mm <sup>-1</sup> )                 | 0.205  |
| Z                                     | 4  |
| F(000)                                | 15260  |
| Crystal size (mm <sup>3</sup> )       | 0.20 x 0.05 x 0.05   |
| Index ranges                          | -37 ≤ h ≤ 38; -40 ≤ k ≤ 40, -42<br>≤ l ≤ 40  |
| θ                                     | 0.603 to 28.132°   |
| # reflections collected               | 367899   |
| # independent reflections             | 119907 [R(int) = 0.0572]   |
| Resolution range (Å)                  | 39.27-1.05 (1.07-1.05)   |
| Completeness (%) to θ = 28.132°       | 78.8   |
| Data/restraints/parameters            | 119907/15387/8137  |
| Goodness of fit                       | 1.627  |
| Final R indices [I > 2σ(I)]           | R <sub>1</sub> = 0.1837, wR <sub>2</sub> = 0.4178                                      |
| R indices (all data)                  | R <sub>1</sub> = 0.2282, wR <sub>2</sub> = 0.4651                                      |
| Absolute structure parameter          | 0.497(16)  |
| CCDC                                  | 1956108  |

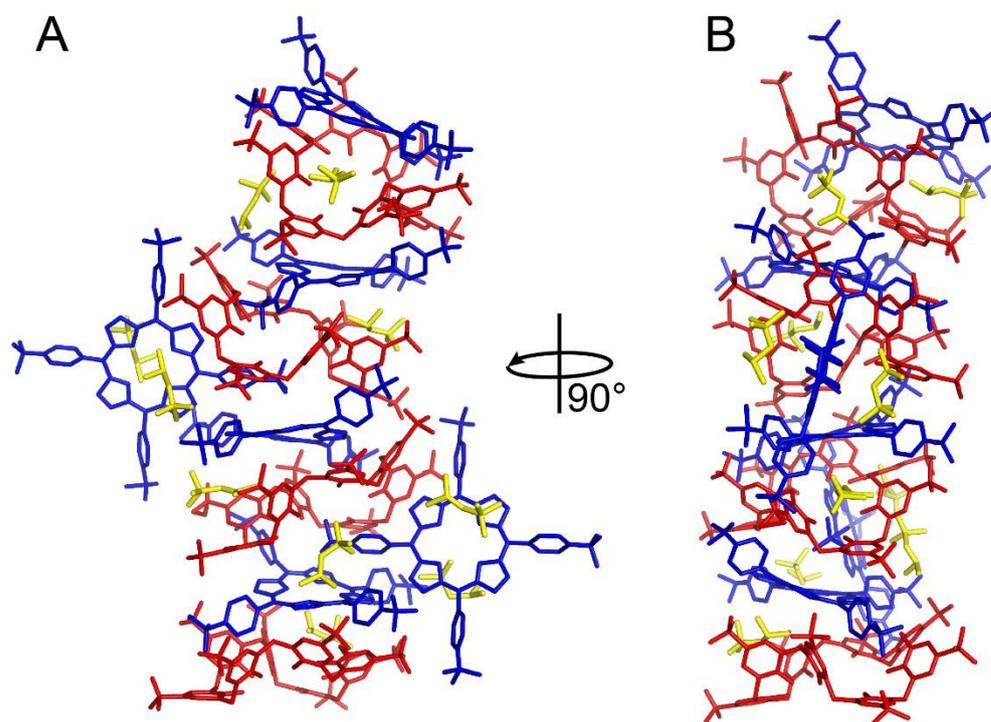
**Table 2.** X-ray data collection and refinement statistics for **sclx<sub>8</sub> – tmap** (squeeze).

|  |  |
|--|--|
| Empirical formula                            | C <sub>313</sub> H <sub>317</sub> N <sub>24</sub> O <sub>77</sub> S <sub>16</sub>                                  |
| Formula weight                               | 6159.85  |
| Temperature (K)                              | 293 (2)  |
| Wavelength (Å)                               | 0.82656 Å  |
| Crystal system                               | monoclinic   |
| Space group                                  | <i>P</i> 2 <sub>1</sub>  |
| Unit Cell dimensions                         | a = 33.621(17) Å<br>b = 38.411(13) Å<br>c = 40.234(15) Å<br>$\alpha = \gamma = 90^\circ$ $\beta = 102.56(4)^\circ$ |
| V (Å <sup>3</sup> )                          | 50715 (37)   |
| $\rho_{\text{cal}}$ (mg/cm <sup>3</sup> )    | 0.807  |
| $\mu$ (mm <sup>-1</sup> )                    | 0.177  |
| Z  | 4  |
| F(000)                                       | 12940  |
| Crystal size (mm <sup>3</sup> )              | 0.20 x 0.05 x 0.05   |
| Index ranges                                 | -37 ≤ h ≤ 38; -40 ≤ k ≤ 40, -42<br>≤ l ≤ 40  |
| $\theta$                                     | 0.603 to 28.132°   |
| # reflections collected                      | 367899   |
| # independent reflections                    | 119907 [R(int) = 0.0572]   |
| Resolution range (Å)                         | 39.27-1.05 (1.07-1.05)   |
| Completeness (%) to $\theta = 28.132^\circ$  | 78.8   |
| Data/restraints/parameters                   | 119907/15357/6455  |
| Goodness of fit                              | 0.945  |
| Final R indices [ <i>I</i> > 2σ( <i>I</i> )] | R <sub>1</sub> = 0.0836, wR <sub>2</sub> = 0.2297  |
| R indices (all data)                         | R <sub>1</sub> = 0.1196, wR <sub>2</sub> = 0.2828  |
| Absolute structure parameter                 | 0.495(15)  |
| CCDC   | 1956128  |

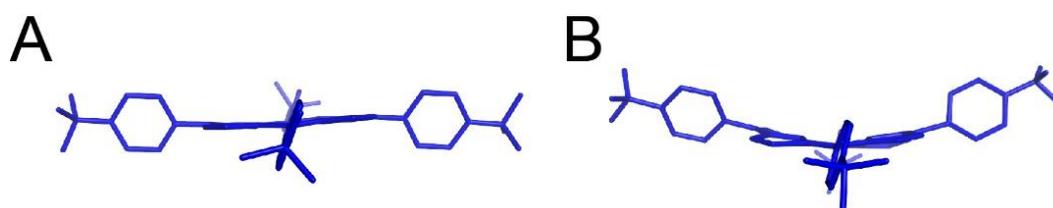


**Figure 3.** Representative ORTEP diagram of **(A) tmap** and **(B) sclx<sub>8</sub>**

**sclx<sub>8</sub> – tmap supramolecular architecture.** The crystal structure revealed a stacked assembly in which four **tmap** molecules alternated with four **sclx<sub>8</sub>** (Figure 4A). The remaining two porphyrins bound the stack peripherally (Figure 4). The stacking porphyrins were 75 % more puckered (root mean square deviation from plane= ~0.2 Å; ORTEX module in Oscale)<sup>259</sup> compared to the peripheral planar porphyrins (root mean square deviation from plane= ~0.05 Å; Figure 5).<sup>225</sup> The **sclx<sub>8</sub>** conformation consisted of two ‘calix[3]arene’ cavities which hosted either a trimethylanilinium group or an **mpd** (Figure 4).



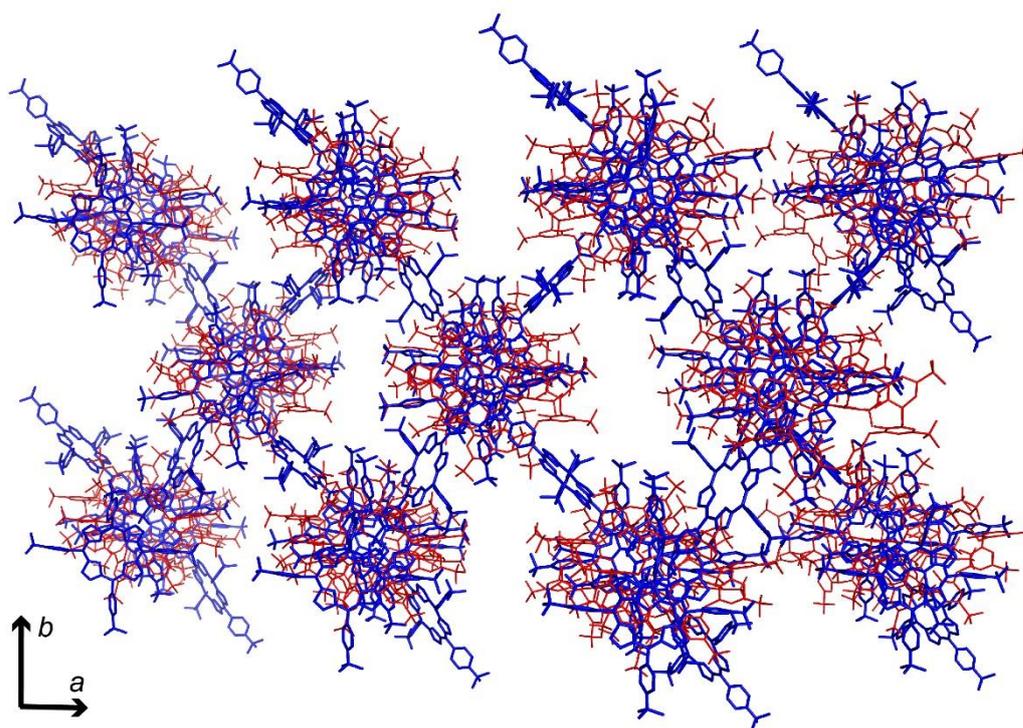
**Figure 4.** (A) The crystal structure revealed a stack of four **sclx<sub>8</sub>** (red), six **tmap** (blue), and ten **mpd** (yellow). The four porphyrins alternating with **sclx<sub>8</sub>** were puckered. The peripheral porphyrins were planar. The puckered conformation of **tmap** facilitated its interaction with the two flanking calixarenes. The **sclx<sub>8</sub>** ‘calix[3]arene’ cavities accommodated a trimethylanilinium group or **mpd** (yellow).



**Figure 5.** In the **sclx<sub>8</sub> – tmap** stack (A) the peripheral porphyrins were planar whereas (B) the stacking porphyrins were distorted or non-planar.

Porphyrin – calixarene co-crystals have been characterized previously.<sup>210-212,240</sup> These crystal structures comprised a stack of porphyrins to which calixarenes bound peripherally. In contrast, in the **sclx<sub>8</sub>** – **tmap** complex both the macrocycles adopted conformations that facilitated their participation in the stack, enhancing the network of non-covalent interactions (Figure 4). This binding mode appears to have been favoured by the charge and shape complementarity of the trimethylanilinium with the **sclx<sub>8</sub>** conformation. For example, in a **sclx<sub>8</sub>** – **tmap** – **sclx<sub>8</sub>** sandwich, trimethylanilinium groups interacted with the phenolsulfonate units of the flanking calixarenes *via* weak edge-to-face (anilinium and phenol rings) and/or charge-charge interactions (partial positively charged anilinium methyls with sulfonates;  $-^+N-C\cdots O-S = 2.8-3.6 \text{ \AA}$ ). A trimethylanilinium subunit hosted within a ‘calix[3]arene’ cavity of **sclx<sub>8</sub>** participated also in cation- $\pi$  interaction ( $-^+N-C\cdots \text{Centroid} = 3.5 - 3.8 \text{ \AA}$ ) in addition to the weak edge-to-face and salt bridge ( $-^+N-C\cdots O-S = 3.3 - 3.7 \text{ \AA}$ ) contacts. Furthermore, it is noteworthy that 3 pyrrole rings of each stacking porphyrin is involved in CH- $\pi$  interaction ( $-C\cdots \text{Centroid} = 3.8 - 4.1 \text{ \AA}$ ) with the methylene groups ( $-CH_2-$ ) of the flanking calixarenes. These interactions may have led to the puckering of the porphyrin ring.

The conformation adopted by the two macrocycles in the stack also facilitated interaction with the peripheral porphyrins. Every alternating calixarene – porphyrin pair in the stack established contacts with one of the trimethylanilinium groups *via* cation –  $\pi$  (partially positive anilinium methyls and the pyrrole ring), edge-to-face (anilinium and pyrrole rings) and charge-charge interactions (anilinium methyl and phenolic OH). Interactions involving the pyrrole rings were limited in this porphyrin. Electrostatic interactions were also observed with the anilinium methyls and the calixarene sulfonates. Another trimethylanilinium group of the same **tmap** formed similar contacts in the neighbouring stack thus acting as a ‘bridge’ or a ‘connector’ to mediate the calixarene – porphyrin supramolecular architecture (Figure 6).



**Figure 6.** Each calixarene – porphyrin stack interacts with the neighbouring stack *via* the peripheral porphyrins that act as a ‘bridge’ or a ‘connector’ to facilitate the supramolecular architecture. Calixarenes, porphyrins, and the unit cells axes are red, blue and black, respectively.

To determine how well the macrocycles pack with each other in the stack (asymmetric unit), the number of intermolecular van der Waals contacts between them were measured in the ORTEX module (Table 3).<sup>259</sup> The variations in the number of contacts between **sclx<sub>8</sub> – tmap / tmap – sclx<sub>8</sub>** is indicative of the asymmetric nature of the interaction. Such differences arise due to the different environment of each atom of the molecule (in a symmetric interaction the number of atoms in contacts will be the same). In the stack, approximately 40 - 60 % atoms of the stacking **tmap** were in contact with the flanking **sclx<sub>8</sub>**. On the other hand, 30 - 33 % atoms of the peripheral **tmap** (trimethylanilinium groups of 1\* and 5\*; Table 3) were in contact with the neighbouring calixarenes (1 and 3) and the stacking porphyrins (4 and 2).

**Table 3.** Number of atoms in van der Waals contacts between **tmap** and **sclx<sub>8</sub>** within the stack

| #                                   | <b>tmap<sup>a</sup></b> |    |    |    |    |    | <b>sclx<sub>8</sub><sup>b</sup></b> |    |    |    |
|-------------------------------------|-------------------------|----|----|----|----|----|-------------------------------------|----|----|----|
|                                     | 1*                      | 2  | 3  | 4  | 5* | 6  | 1                                   | 2  | 3  | 4  |
| <b>tmap<sup>a</sup></b>             |                         |    |    |    |    |    |                                     |    |    |    |
| 1*                                  |                         |    |    | 30 |    |    | 43                                  |    |    |    |
| 2                                   |                         |    |    |    | 29 |    | 75                                  |    | 51 |    |
| 3                                   |                         |    |    |    |    |    |                                     |    | 74 | 51 |
| 4                                   | 28                      |    |    |    |    |    | 50                                  | 75 |    |    |
| 5*                                  |                         | 34 |    |    |    |    |                                     |    | 44 |    |
| 6                                   |                         |    |    |    |    |    |                                     |    |    | 77 |
| <b>sclx<sub>8</sub><sup>b</sup></b> |                         |    |    |    |    |    |                                     |    |    |    |
| 1                                   | 42                      | 66 |    | 43 |    |    |                                     |    |    |    |
| 2                                   |                         |    |    | 58 |    |    |                                     |    |    |    |
| 3                                   |                         | 39 | 59 |    | 48 |    |                                     |    |    |    |
| 4                                   |                         |    | 44 |    |    | 68 |                                     |    |    |    |

\*indicates the peripheral porphyrin in the stack

<sup>a</sup> number of atoms in **tmap** – 126 (excluding the 2 protons of pyrrole nitrogen)

<sup>b</sup> number of atoms in **sclx<sub>8</sub>** – 128 (excluding the 8 protons of sulfonates and hydroxyls)

## Conclusions

In conclusion, we have illustrated an alternating stacking assembly of **sclx<sub>8</sub>** and **tmap**. Interestingly, weak edge-to-face, cation- $\pi$ , and/or charge-charge interactions stabilize the complex. The conformations adopted by the macrocycles played a major role in facilitating these non-covalent interactions. In particular, considering the ‘floppiness’ of the larger calix[6,8]arenes, they may be utilized in generating stacking calixarene – porphyrin hybrid architectures. Furthermore, these complexes may have applications in the development of sensors<sup>74,226</sup> or light-harvesting systems.<sup>229-231</sup> Finally, this complex exemplifies a high Z' structure and may provide insight into understanding the factors that contribute towards such unique crystal packing.<sup>256-258</sup>



## Chapter 4

---

### Calixarene-mediated Assembly of a Small Antifungal Protein

The material in this chapter was published as:

**Alex J. M.**, Rennie M. L., Engilberge S., Lehoczki G., Dorottya H., Fizil Á., Batta G., Crowley P. B.

Calixarene-mediated assembly of a small antifungal protein.

*IUCrJ*, **2019**, *6*, 238-247.

Gyula Batta's Lab (University of Debrecen, Hungary) –

- PAF protein production and purification
- NMR and ITC data collection

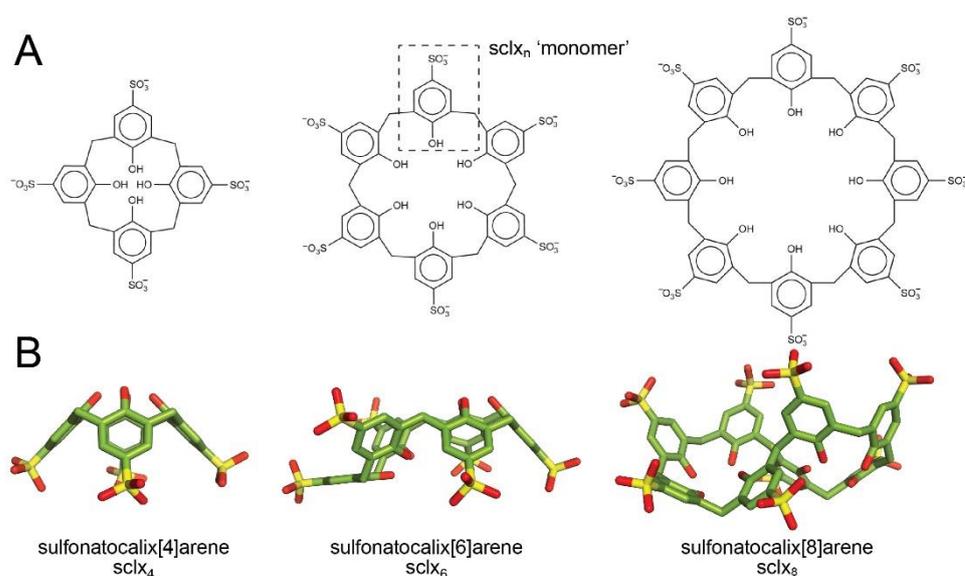
The raw data from these experiments were processed and analysed by JMA and MLR.

### Abstract

Synthetic macrocycles such as calixarenes and cucurbiturils are increasingly applied as mediators of protein assembly and crystallization. The macrocycle can facilitate assembly by providing a surface on which two or more proteins bind simultaneously. This work explores the capacity of the sulfonato-calix[n]arene (**sclx<sub>n</sub>**) series to effect crystallization of PAF, a small, cationic antifungal protein. Co-crystallization with **sclx<sub>4</sub>**, **sclx<sub>6</sub>** or **sclx<sub>8</sub>** led to high-resolution crystal structures. In the absence of **sclx<sub>n</sub>**, diffraction-quality crystals of PAF were not obtained. Interestingly, all three **sclx<sub>n</sub>** were bound to a similar patch on PAF. The largest and most flexible variant, **sclx<sub>8</sub>**, yielded a dimer of PAF. Complex formation was evident in solution via NMR and ITC experiments, showing more pronounced effects with increasing macrocycle size. In agreement with the crystal structure, the ITC data suggested that **sclx<sub>8</sub>** acts as a bidentate ligand. The contributions of calixarene size/conformation to protein recognition and assembly are discussed. Finally, it is suggested that the conserved binding site for anionic calixarenes implicates this region of PAF in membrane binding, which is a prerequisite for antifungal activity.

## Introduction

There is growing interest in the use of synthetic macrocycles as mediators of protein assembly.<sup>4</sup> The special case of protein crystallization<sup>46</sup> has benefitted from ‘molecular glues’ such as calixarenes and cucurbiturils that promote crystal packing.<sup>19-21</sup> The sulfonato-calix[n]arenes (**sclx<sub>n</sub>**, Figure. 1) are highly water-soluble, anionic macrocycles with diverse applications in the biosciences.<sup>182,70,260</sup> The hydrophobic core and the anionic rim of the calixarene can facilitate protein recognition, in particular, *via* the entrapment of arginine or lysine side chains.<sup>19,20,48-54,182</sup> Consequently, **sclx<sub>4</sub>** and related compounds readily co-crystallize with the highly cationic cytochrome *c* (cytc) and lysozyme.<sup>48-50,52,53</sup> With increasing calixarene size there tends to be more pronounced effects; for example, phosphonato-calix[6]arene (**pclx<sub>6</sub>**) has an approximately ten-fold increase in affinity (with respect to **sclx<sub>4</sub>**) and prompts dimerization of cytc in solution.<sup>51</sup> Sulfonato-calix[8]arene (**sclx<sub>8</sub>**) on the other hand induces a tetramer of cytc.<sup>19</sup> Furthermore, while calix[4]arene is locked in a bowl conformation, the larger calixarenes are flexible and adopt various conformations (Figure. 1)<sup>19,20,51,67,244,245,248,260-262</sup> Accordingly, **sclx<sub>8</sub>** can bind to cytc either *via* an extended ‘pleated loop’ or a collapsed ‘double cone’ conformation, as shown using X-ray crystallography.<sup>19,20</sup>



**Figure 1.** Sulfonato-calix[n]arenes. **(A)** Molecular structures and **(B)** cone (**sclx<sub>4</sub>**), double partial cone (**sclx<sub>6</sub>**) and double cone (**sclx<sub>8</sub>**) conformations.

We were motivated to characterize the **sclx<sub>n</sub>** series with a single protein and thus investigate systematically how the calixarene size and flexibility influence protein recognition and assembly. Furthermore, we were interested in studying a protein for which a crystal structure was not available. Acknowledging the tendency of **sclx<sub>n</sub>** to complex cationic proteins we chose the *Penicillium* antifungal protein (PAF) as a test case.<sup>162,263</sup> PAF is a small (6.2 kDa, 55 residues) lysine-rich protein (13 Lys, *pI*~9) and a potent agent against *Aspergillus* species and dermatophytes.<sup>163,164,264</sup> The NMR structure is a twisted-barrel composed of five antiparallel  $\alpha$ -strands and stabilized by three disulfide bridges.<sup>156,157,265</sup> Lys30, Phe31, Lys34, Lys35 and Lys38 (loop 3) belong to a conserved region of PAF that is important for antifungal activity.<sup>165,166,266</sup> Similar to defensins, the mechanism of antifungal action is postulated to require interaction with anionic components on the cell membrane.<sup>164,165,267</sup> Recent X-ray crystal structures have revealed how defensin–phospholipid binding leads to oligomerization, suggesting a mechanism for membrane permeation.<sup>268-271</sup> These observations provided further motivation to characterize PAF binding with anionic receptors.

Here, we report three PAF – **sclx<sub>n</sub>** crystal structures, demonstrating the fitness of calixarenes as crystallization agents. Interestingly, all three calixarenes were bound to PAF, mainly at the conserved loop 3. A similar interaction site was determined by NMR studies; these results suggest that loop 3 is favoured for recognition by anionic receptors. The largest calixarene **sclx<sub>8</sub>** mediated a PAF dimer that was observed both crystallographically and in solution. The thermodynamics of PAF – **sclx<sub>n</sub>** interactions were characterized by isothermal titration calorimetry, providing further evidence of PAF dimerization *via* **sclx<sub>8</sub>**. The results are discussed in the context of protein assembly and membrane binding. Finally, insights into protein complexation by flexible calixarenes are provided, including the role of PEG fragments at the protein – calixarene interface.

## Experimental

**Materials.** PAF was produced as described.<sup>166,263</sup> The calixarenes were purchased from TCI Chemicals. Stock solutions of **sclx<sub>4</sub>**, **sclx<sub>6</sub>** and **sclx<sub>8</sub>** were prepared in water and the pH was adjusted to 6.0.

**Crystallization trials.** Co-crystallization experiments were performed by the hanging-drop vapour-diffusion method at 20 °C. The reservoir solution was 20-30% PEG 3350 and 50 mM sodium acetate, pH 5.6. A range of protein (0.7–7.0 mM PAF) and ligand (5 – 40 mM **sclx<sub>4</sub>**) concentrations were tested for PAF – **sclx<sub>4</sub>** co-crystallization. Drops were prepared by combining sequentially 1 mL each of reservoir solution, protein and **sclx<sub>4</sub>**. Crystals grew at 7 mM PAF and 40 mM **sclx<sub>4</sub>**. In the case of PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>**, the protein–ligand solutions were premixed before combining with the reservoir solution. Co-crystals were obtained with 10 mM **sclx<sub>6</sub>** and 40 mM **sclx<sub>8</sub>**. Crystals grew in 4–5 days (**sclx<sub>4</sub>**), 2–3 weeks (**sclx<sub>6</sub>**) or 6–8 weeks (**sclx<sub>8</sub>**).

The crystallization of ligand-free PAF (7 mM) was performed with an Oryx 8 Robot (Douglas Instruments) and a sparse matrix screen (JCSG++, Jena Bioscience).

Spherulites were obtained in C6 (40 % PEG 300, 0.1 M potassium phosphate citrate pH 4.2) and needles grew in D7 (40 % PEG 400, 0.1 M Tris-HCl pH 8.5, 0.2 M lithium sulfate). Manual crystallization trials under these conditions did not yield suitable crystals.

**X-ray data collection.** Crystals were cryo-protected in reservoir solution supplemented with 20% glycerol and cryo-cooled in liquid nitrogen. Diffraction data were collected at the SOLEIL synchrotron (France) to 1.30, 1.45 and 1.50 Å resolution for PAF – **sclx<sub>4</sub>**, PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>**, respectively. Datasets were collected using  $\phi$  scans of 0.1° over 200 (PAF – **sclx<sub>4</sub>**), 180 (PAF – **sclx<sub>6</sub>**) and 110 (PAF – **sclx<sub>8</sub>**) using an EIGER X 9M detector. In the case of pure PAF, a dataset extending to 3.0 Å was collected for the spherulites (condition C6), but was difficult to index/integrate in both XDS and iMOSFLM. The needle-like crystals (condition D7) did not diffract.

**Structure determination.** The observed reflections for PAF – **sclx<sub>4</sub>** were processed with XDS,<sup>192</sup> whereas iMOSFLM<sup>272</sup> was used for the PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>** datasets. In all cases, the data were scaled using POINTLESS<sup>193</sup> and AIMLESS.<sup>194</sup> Xtriage (PHENIX)<sup>216</sup> suggested pseudo-merohedral twinning for the PAF – **sclx<sub>4</sub>** data with twin law -h, -k, -h -l, and estimated twin fractions of 0.025 (Britton analyses), 0.066 (H-test) and 0.022 (maximum-likelihood method). The structure was determined by molecular replacement in PHASER<sup>195</sup> by using the NMR structure (PDB reference 2mhv, conformer 1)<sup>166</sup> as the search model. A satisfactory solution (LLG, 134; TFZ, 7.4) was obtained with a search model in which residues 1–2, 17–24 and 47–49 were deleted and all six cysteines were replaced by alanine. The coordinates and restraints for **sclx<sub>4</sub>** (ligand ID T3Y) were added in COOT.<sup>197</sup> Twin refinement did not result in any significant improvement in the electron density. No twinning was indicated for the PAF – **sclx<sub>6</sub>** or PAF – **sclx<sub>8</sub>** data. The structures were solved by molecular replacement using the structure of PAF – **sclx<sub>4</sub>** (devoid of **sclx<sub>4</sub>**) as the search model. The coordinates for **sclx<sub>6</sub>** and **sclx<sub>8</sub>** were built in JLigand.<sup>196</sup> High mosaic spread (0.3–0.9) in the PAF – **sclx<sub>8</sub>** dataset made it difficult to obtain better R values.

Truncating the images with high mosaicity did not help in this respect. Iterative cycles of manual model building in COOT<sup>197</sup> and refinement in BUSTER<sup>217</sup> were carried out until no further improvements in  $R_{\text{free}}$  and electron density were observed. The final structures were validated with MolProbity<sup>199</sup> and deposited in the Protein Data Bank as PAF – **sclx**<sub>4</sub> (PDB 6ha4), PAF – **sclx**<sub>6</sub> (PDB 6hah) and PAF – **sclx**<sub>8</sub> (PDB 6haj).

**Accessible surface area calculations.** The effect of **sclx**<sub>4</sub>, **sclx**<sub>6</sub> and **sclx**<sub>8</sub> on the accessible surface area (ASA) of PAF residues in the crystal packing environments was determined in ArealMol as described previously.<sup>200</sup>

**NMR spectroscopy.** The sample conditions were 0.3 or 0.5 mM <sup>15</sup>N-PAF in 10 mM sodium phosphate buffer at pH 6.0. NMR titrations were performed at 298 K using 0.5–1 ml aliquots of 50 mM stocks of **sclx**<sub>4</sub>, **sclx**<sub>6</sub> or **sclx**<sub>8</sub>. <sup>1</sup>H-<sup>15</sup>N HSQC spectra were acquired with spectral widths of 12 p.p.m. (<sup>1</sup>H) and 19 p.p.m. (<sup>15</sup>N) using two scans and 128 increments on a Bruker Avance-II-500 NMR spectrometer. Ligand-induced chemical-shift perturbations were analysed in CCPN.<sup>201</sup>

**Isothermal titration calorimetry and data fitting.** PAF samples were dissolved in 10 mM sodium phosphate pH 6.0. The same buffer was used to dilute stocks of **sclx**<sub>4</sub> (7.1 mM, PAF 0.5 mM), **sclx**<sub>6</sub> (3.6 mM, PAF 0.5 mM) and **sclx**<sub>8</sub> (2.5 mM, PAF 0.3 mM) to the required concentration. Samples were degassed prior to the titration. Measurements were made at 25 °C using a Microcal ITC-200 instrument. Titrations were performed in duplicate with similar trends between each replicate. A single replicate from each calixarene was used for model fitting. Separate titrations of each calixarene into buffer confirmed that the heats of dilution were small, exothermic and approximately constant.

NITPIC<sup>273</sup> was used for baseline correction and integration of the thermograms. Pytc<sup>274</sup> was used to perform model fitting and parameter estimation. The system of equations relating the independent variables of the model (total

concentrations) to the experimental observations (heat generated during injections) for the single-site and bidentate-ligand models are as follows.

### Single-site model

$$[P_T]_i = [P]_i + [PL]_i \quad (1)$$

$$[L_T]_i = [L]_i + [PL]_i$$

$$[PL]_i = K[P]_i[L]_i \quad (2)$$

$$q_i = V_{\text{cell}}\Delta H^\circ([PL]_i - [PL]_{i-1}(1 - v_i/V_{\text{cell}})) + q_{\text{dil}} \quad (3)$$

Where,

$[P_T]_i$  is the total cell concentration of protein at the  $i^{\text{th}}$  injection (independent variable)

$[L_T]_i$  is the total cell concentration of ligand at the  $i^{\text{th}}$  injection (independent variable)

$K_1$ , is the equilibrium association constant (fit parameter)

$\Delta H$  is the enthalpy (fit parameter) associated with  $K$

$V_{\text{cell}}$  is the volume of the cell

$v_i$  is the volume of the  $i^{\text{th}}$  injection

$q_i$  is the heat generated from the  $i^{\text{th}}$  injection (dependent variable)

$q_{\text{dil}}$  is the heat of dilution (fit parameter, assumed to be constant)

### Bidentate Ligand model

$$[P_T]_i = [P]_i + [PL]_i + 2[P_2L]_i \quad (4)$$

$$[L_T]_i = [L]_i + [PL]_i + [P_2L]_i$$

$$[PL]_i = 2K_1[P]_i[L]_i \quad (5)$$

$$[P_2L]_i = K_1K_2[P]_i^2[L]_i$$

$$q_i = V_{\text{cell}}(\Delta H_1^\circ([PL]_i - [PL]_{i-1}(1 - v_i/V_{\text{cell}})) + (\Delta H_1^\circ + \Delta H_2^\circ)([P_2L]_i - [P_2L]_{i-1}(1 - v_i/V_{\text{cell}}))) + q_{\text{dil}} \quad (6)$$

Where,

$K_1, K_2$  are the microscopic equilibrium association constants (fit parameters)

$\Delta H_1, \Delta H_2$  are the enthalpies (fit parameters) associated with  $K_1, K_2$ , respectively.

The expressions for mass balance of the protein and ligand is represented by equations 1 or 4. Equations 2 or 5 define the equilibrium constants. For the bidentate ligand model equation 5 was solved numerically (Levenberg-Marquardt algorithm) to yield the free protein ( $[P]_i$ ) and free ligand concentrations ( $[L]_i$ ). The free concentrations were used to compute the concentrations of the other states *via* the equilibrium equations. The heat generated from a given injection was determined by equations 3 or 6. Parameters were constrained to physically reasonable bounds (*e.g.*  $K_1, K_2$  between  $10^2$  and  $10^{10} \text{ M}^{-1}$ ) and best-fits were obtained by maximum likelihood starting from a range of initial estimates. Parameter errors and correlations were estimated using a Bayesian approach (Markov chain Monte Carlo simulations). The error for each integrated heat was determined using NITPIC.<sup>273</sup>

## Results and discussion

**PAF – sclx<sub>n</sub> co-crystallization.** Pure PAF proved to be recalcitrant to crystallization. A sparse-matrix screen yielded spherulites or needle-like crystals only (see experimental). In contrast, PAF – **sclx<sub>4</sub>** mixtures were crystallized readily from solutions containing PEG and sodium acetate. PAF – **sclx<sub>4</sub>**, PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>** co-crystals were obtained at 28–30% PEG 3350 and 50 mM sodium acetate pH 5.6 (Figure 2).

**Data collection and model building.** Datasets extending to 1.30, 1.45 and 1.50 Å resolution were collected from monoclinic ( $P12_11$ ) PAF – **sclx<sub>4</sub>**, PAF – **sclx<sub>6</sub>** and hexagonal ( $P6_1$ ) PAF – **sclx<sub>8</sub>** co-crystals, respectively (Table 1).

**Table 1.** X-ray data collection, processing and refinement statistics

| <i>Crystallization</i>                    |   |   |  |
|---|---|---|--|
|   | PAF – sclx <sub>4</sub>   | PAF – sclx <sub>6</sub>   | PAF – sclx <sub>8</sub>  |
| [PAF] / [sclx <sub>n</sub> ] (mM)         | 7 / 40  | 7 / 10  | 7 / 40   |
| PEG 3350 (%)                              | 30  | 30  | 28   |
| Buffer                                    | 0.05 M sodium acetate pH 5.6  |   |  |
| <i>Data Collection</i>                    |   |   |  |
| Light source                              | Soleil, PROXIMA 2A  |   |  |
| Wavelength (Å)                            | 0.980105  |   |  |
| Space group                               | <i>P</i> 12 <sub>1</sub> 1  | <i>P</i> 12 <sub>1</sub> 1  | <i>P</i> 6 <sub>1</sub>  |
| Cell constants                            | <i>a</i> = 22.71 Å<br><i>b</i> = 37.58 Å<br><i>c</i> = 29.95 Å<br>$\alpha = \gamma = 90^\circ$<br>$\beta = 111^\circ$ | <i>a</i> = 24.74 Å<br><i>b</i> = 38.59 Å<br><i>c</i> = 29.93 Å<br>$\alpha = \gamma = 90^\circ$<br>$\beta = 112^\circ$ | <i>a</i> = 24.30 Å<br><i>b</i> = 24.30 Å<br><i>c</i> = 313.69 Å<br>$\alpha = \beta = 90^\circ$<br>$\gamma = 120^\circ$ |
| Resolution (Å)                            | 27.82-1.33<br>(1.37-1.33)   | 27.08-1.45<br>(1.48-1.45)   | 21.06-1.50<br>(1.60-1.50)  |
| # unique reflections                      | 10142 (842)   | 8592 (397)  | 16146 (2346)   |
| Multiplicity                              | 3.4 (2.9)   | 2.9 (2.5)   | 4.8 (4.0)  |
| <i>I</i> / $\sigma$ ( <i>I</i> )          | 6.1 (1.3)   | 11.5 (5.1)  | 8.4 (1.0)  |
| Completeness (%)                          | 92.2 (75.7)   | 94.7 (90.3)   | 96.8(97.7)   |
| <i>R</i> <sub>meas</sub> <sup>b</sup> (%) | 12.1 (61.1)   | 7.4 (33.9)  | 8.9 (98.2)   |
| <i>R</i> <sub>pim</sub> <sup>c</sup> (%)  | 6.2 (34.2)  | 4.2 (20.4)  | 3.7 (47.4)   |
| CC <sub>1/2</sub>                         | 0.986 (0.742)   | 0.992 (0.760)   | 0.998 (0.226)  |
| Solvent content (%)                       | 35  | 41  | 43   |
| <i>Refinement</i>                         |   |   |  |
| <i>R</i> <sub>work</sub>                  | 0.185   | 0.200   | 0.217  |
| <i>R</i> <sub>free</sub>                  | 0.217   | 0.236   | 0.244  |
| rmsd bonds (Å)                            | 0.009   | 0.011   | 0.011  |
| rmsd angles (°)                           | 1.250   | 1.41  | 1.450  |
| # molecules in asymmetric unit            |   |   |  |
| Protein                                   | 1   | 1   | 2  |
| Ligand                                    | 1   | 1   | 1  |
| Water                                     | 55  | 57  | 77   |
| Ave. B-factor (Å <sup>2</sup> )           | 20.82   | 23.86   | 30.01  |
| Ramachandran analysis <sup>d</sup>        |   |   |  |
| % residues in                             |   |   |  |
| favoured regions                          | 100.0   | 100.0   | 95.3   |
| allowed regions                           |   |   | 3.76   |
| PDB code                                  | 6ha4  | 6hah  | 6haj   |

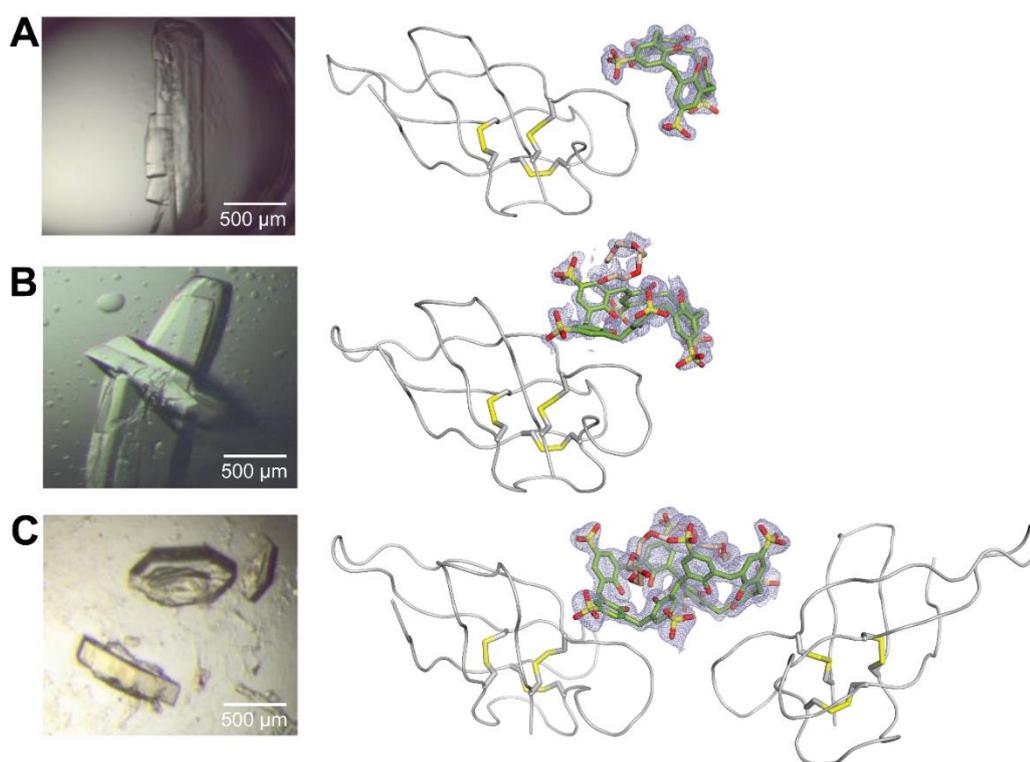
<sup>a</sup>Values in parentheses correspond to the highest resolution shell

<sup>b</sup> $R_{\text{meas}} = \frac{\sum_{hkl} \sqrt{(n/n-1) \sum_l |I_i(hkl) - \langle I(hkl) \rangle|}}{\sum_{hkl} \sum_l I_i(hkl)}$ ;

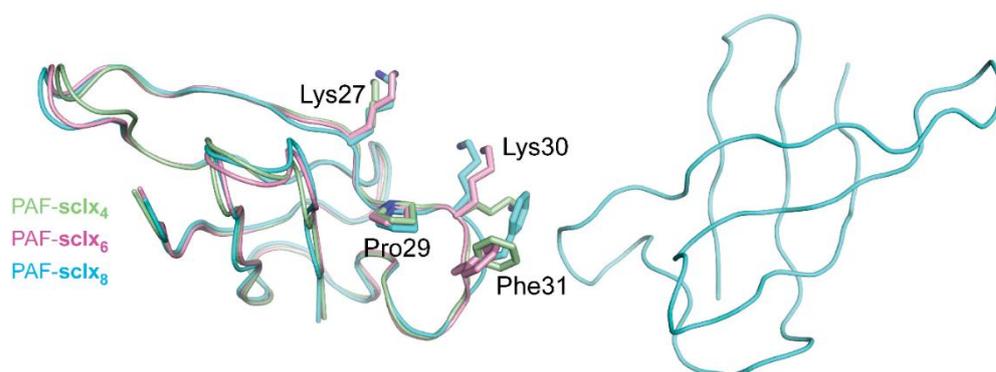
<sup>c</sup> $R_{\text{pim}} = \frac{\sum_{hkl} \sqrt{(1/n-1) \sum_{i=1}^n |I_i(hkl) - \langle I(hkl) \rangle|}}{\sum_{hkl} \sum_l I_i(hkl)}$ ;

<sup>d</sup>Calculated in MolProbity.

The PAF – **sclx<sub>4</sub>** structure was determined using the NMR coordinates (PDB reference 2mhv)<sup>166</sup> as the search model. To obtain a satisfactory solution it was necessary to delete two loops and replace all six cysteines with alanines. After several rounds of model building and refinement a complete PAF structure was obtained. This model was used to solve the PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>** structures. The PAF fold and the three disulfide bridges in the X-ray structures were consistent with the NMR model.<sup>166,167,265</sup> Interestingly, the fold was altered slightly in response to **sclx<sub>n</sub>** binding (Figure 3). Superposition of the three structures revealed a C<sup>α</sup> rmsd of 0.54 Å (PAF – **sclx<sub>6</sub>**) and 0.78 Å (PAF – **sclx<sub>8</sub>**) relative to PAF – **sclx<sub>4</sub>**, with the largest differences at loops 2, 3 and 4.



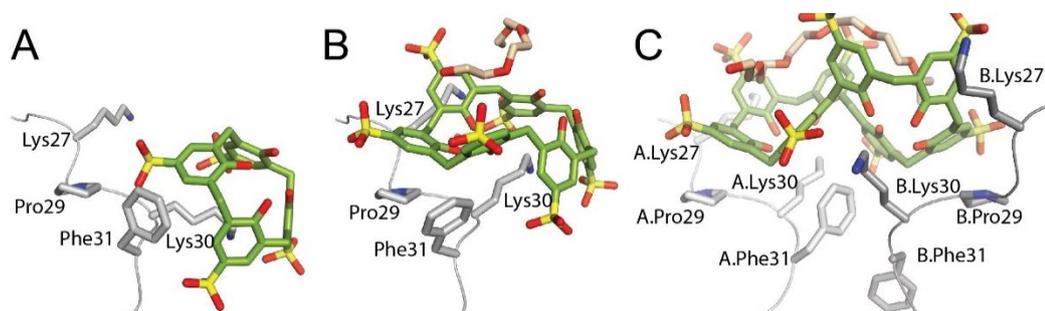
**Figure 2.** PAF – **sclx<sub>n</sub>** crystals and X-ray structures. Diffraction-quality co-crystals and the asymmetric units of PAF with (A) **sclx<sub>4</sub>** (B) **sclx<sub>6</sub>** and (C) **sclx<sub>8</sub>**. The asymmetric unit comprised one protein and one ligand (A,B) or two proteins and one ligand (C). The unbiased  $2F_o - F_c$  electron density map (contoured at 1.0  $\sigma$ ) showing calixarenes, and PEG fragments in (B and C).



**Figure 3.** PAF backbone flexibility and binding site conformations. Superposition of the three crystal structures highlight changes in the PAF backbone. The side chains of Lys27, Lys30 and Phe31 have different conformations, while Pro29 provides a rigid platform for hydrophobic interactions with **sclx<sub>6</sub>** and **sclx<sub>8</sub>**. The calixarene coordinates are not shown.

In contrast to the PAF – **sclx<sub>n</sub>** crystals, the spherulites and needles of pure PAF failed to provide a usable dataset. The needles did not diffract and the spherulites yielded a 3.0 Å resolution dataset, which proved difficult to index and integrate. The difficulty in obtaining suitable crystals of pure PAF suggests that the calixarene facilitates protein assembly and crystallization.<sup>19,20,48-50,52-54</sup>

**Different calixarene, similar binding site.** The asymmetric unit of the PAF – **sclx<sub>n</sub>** complexes comprised one (in the case of PAF – **sclx<sub>4</sub>** and PAF – **sclx<sub>6</sub>**) or two (PAF – **sclx<sub>8</sub>**) molecules of PAF. Each structure contained one calixarene, as shown by the  $2F_o - F_c$  electron-density maps (Figures 2 and 4). Additional electron density adjacent to **sclx<sub>6</sub>** and **sclx<sub>8</sub>** was modelled as a PEG fragment equivalent to tetraethylene glycol (EG4) and heptaethylene glycol (EG7), respectively (Figure 4 and 5). **Sclx<sub>4</sub>**, locked in the cone conformation, encapsulates the side chain of a single lysine (Lys30), as observed previously in different protein-clx<sub>4</sub> complexes.<sup>48-50,52-54</sup> The larger flexible **sclx<sub>6</sub>** and **sclx<sub>8</sub>** adopted distinct conformations and bound at least two lysines. **Sclx<sub>6</sub>** was in the double partial-cone conformation,<sup>243,262</sup> with three sulfonates pointed upwards and three pointed downwards (Figures 1B and 4B). **Sclx<sub>8</sub>** adopted the double cone conformation,<sup>244,245,248</sup> with each half of the molecule acting like a calix[4]arene to bind one PAF molecule, thus mediating a crystallographic dimer (Figure 4C).



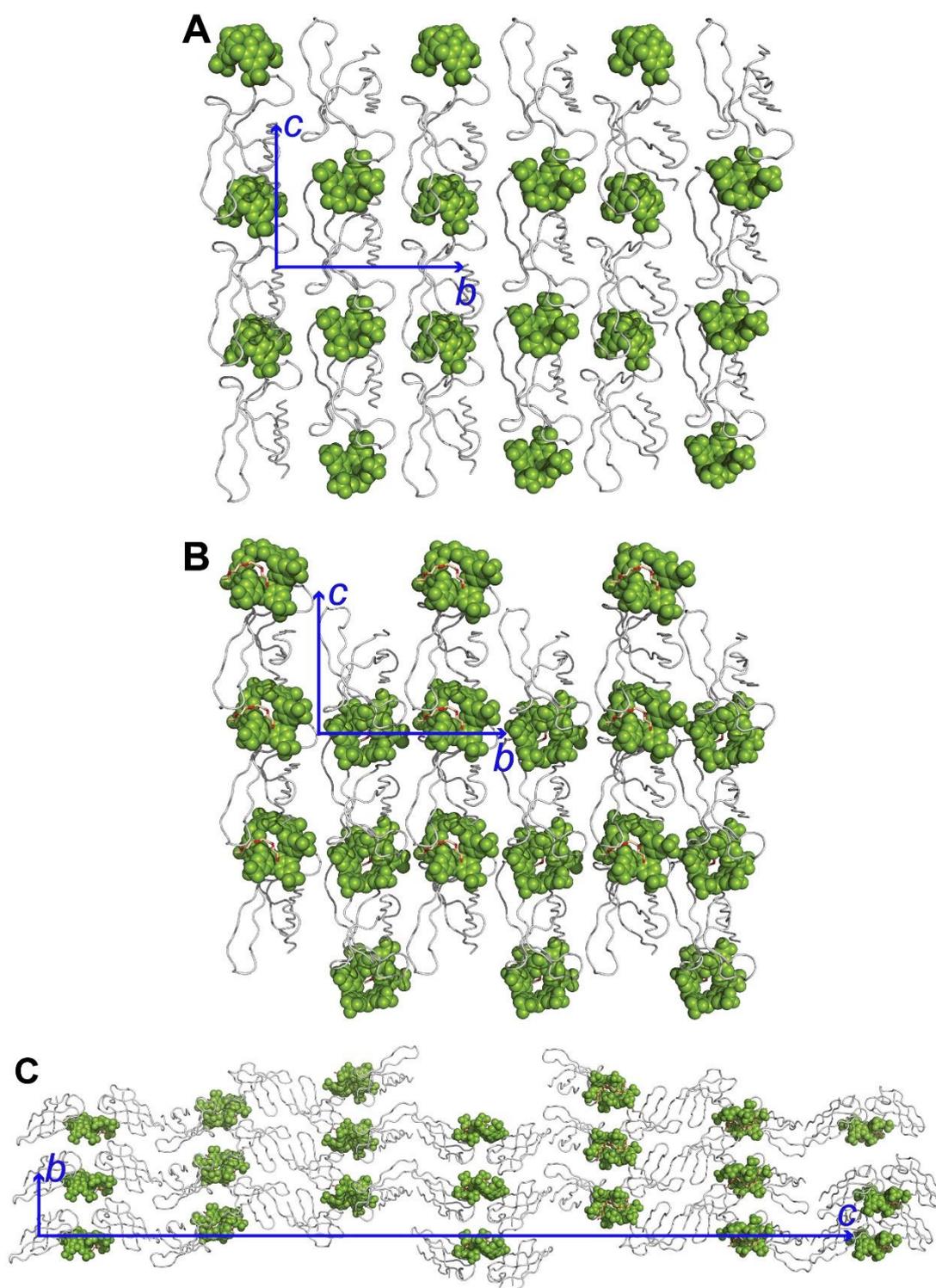
**Figure 4.** Binding site interactions in PAF – **sclx<sub>n</sub>**. (A) **sclx<sub>4</sub>**, (B) **sclx<sub>6</sub>** and (C) **sclx<sub>8</sub>** bind to PAF at Lys30. Note the altered conformations of Lys30 and Phe31 in each structure, while Pro29 provides a rigid hydrophobic surface for face-to-face interaction with **sclx<sub>6</sub>** and **sclx<sub>8</sub>**. In PAF-**sclx<sub>8</sub>**, two protein chains interact with the calixarene. PEG fragments equivalent to tetraethylene glycol and heptaethylene glycol were bound to **sclx<sub>6</sub>** and **sclx<sub>8</sub>**, respectively.

All three calixarenes bound to Lys30, while interacting also with neighbouring residues as well as other proteins (symmetry mates) in the crystal packing. Depending on the ligand size/conformation, the noncovalent contacts varied in their type and multiplicity. The PAF – **sclx<sub>4</sub>** complex (Figure 4A) was similar to cytochrome *c* – **sclx<sub>4</sub>**,<sup>48</sup> involving a salt bridge and CH- $\pi$  /cation- $\pi$  bonds with the encapsulated lysine. Hydrogen bonds to the backbone amide NHs of Lys30, Phe31, and Asp32 were evident and the aromatic ring of Phe31 was in van der Waals contact with a **sclx<sub>4</sub>** methylene bridge. Considering symmetry mates (Figure 5A), **sclx<sub>4</sub>** formed substantial interfaces (>150 Å<sup>2</sup>) with three proteins. Interestingly, a salt bridge was formed with the N $^{\alpha}$  of Ala1. Salt bridges occurred also with Lys2, Lys17, Lys22 and Lys35, emphasizing a large charge-charge component to complexation. In total, the protein – **sclx<sub>4</sub>** interfaces buried ~660 Å<sup>2</sup> of protein.

**sclx<sub>6</sub>** (1.5 times larger than **sclx<sub>4</sub>**) also completely encaged Lys30 (Figure 4B). However, one wall of the calixarene cage was composed of three phenolic groups. The phenolic oxygens were in van der Waals contact with the C $^{\beta}$ , C $^{\gamma}$  and C $^{\delta}$  of Lys30, indicative of CH...O hydrogen bonding and the Lys30-N $^{\alpha}$  was hydrogen bonded to a phenolic OH (rather than to a sulfonate). Other differences, with respect to **sclx<sub>4</sub>**, were water-mediated salt bridges between Lys30-N $^{\zeta}$  and two sulfonates and a weak  $\pi$ - $\pi$  interaction with Phe31 (Figure 4B). Adjacent residue Pro29 was also important for calixarene-binding (*vide infra*). In terms of crystal packing (Figure 5B), the larger **sclx<sub>6</sub>** was nestled between five proteins and formed numerous salt bridges (Lys6, Lys9, Lys11, Lys27, Lys38, Lys42). The resulting protein-ligand contacts mask ~970 Å<sup>2</sup> of protein surface. Compared to **sclx<sub>4</sub>**, the more extensive interactions exhibited by **sclx<sub>6</sub>** may explain why 4 times less ligand was required to achieve crystal growth (Experimental and Table 1).

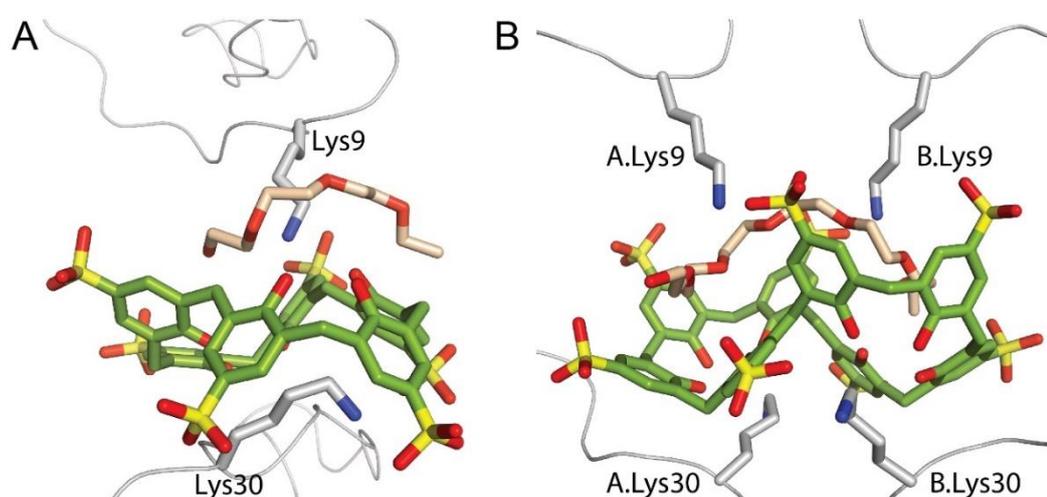
The interactions of **sclx<sub>8</sub>** with PAF were similar to those observed with **sclx<sub>6</sub>**, though less extensive. At twice the size of **sclx<sub>4</sub>** it might be expected that **sclx<sub>8</sub>** would mask a larger protein surface; however, **sclx<sub>8</sub>** formed a PAF dimer (Figures 4C and 5C)

resulting in a total protein surface coverage of 950 Å<sup>2</sup>. The double-cone conformation (compared with the 'pleated loop')<sup>19</sup> adopted by **sclx<sub>8</sub>** minimized its contact with protein surfaces. Salt-bridge interactions involved up to three lysines from each monomer. Here, again a hydrogen bond was formed between the Lys30 N<sup>α</sup> and a phenolic OH. In one of the protein chains Phe31 formed an edge-to-face interaction with an **sclx<sub>8</sub>** phenolic ring. In protein chain B, Phe31 was disordered (Figures 5C).



**Figure 5.** Calixarenes as molecular glues. The crystal packing is dominated by PAF-sclx<sub>n</sub> interactions in (A) PAF – sclx<sub>4</sub>, (B) PAF – sclx<sub>6</sub>, and (C) PAF – sclx<sub>8</sub>. This observation suggests that the calixarene is a molecular glue for protein assembly.

In complex with PAF, **sclx<sub>4</sub>**, **sclx<sub>6</sub>** and **sclx<sub>8</sub>** contributed an additional surface of 550, 850 and 1290 Å<sup>2</sup> respectively (calculated for a single protein). The exposed calixarene surface is a relatively homogenous ‘mask’ that is conducive to forming noncovalent bridges with other proteins. Apparently, the calixarene acts as molecular glue (Figure 5) by providing a patch that mediates protein assembly (subsequently driving protein crystallization) in a special case of the ‘patchy particle model’.<sup>34,53,62-64</sup>

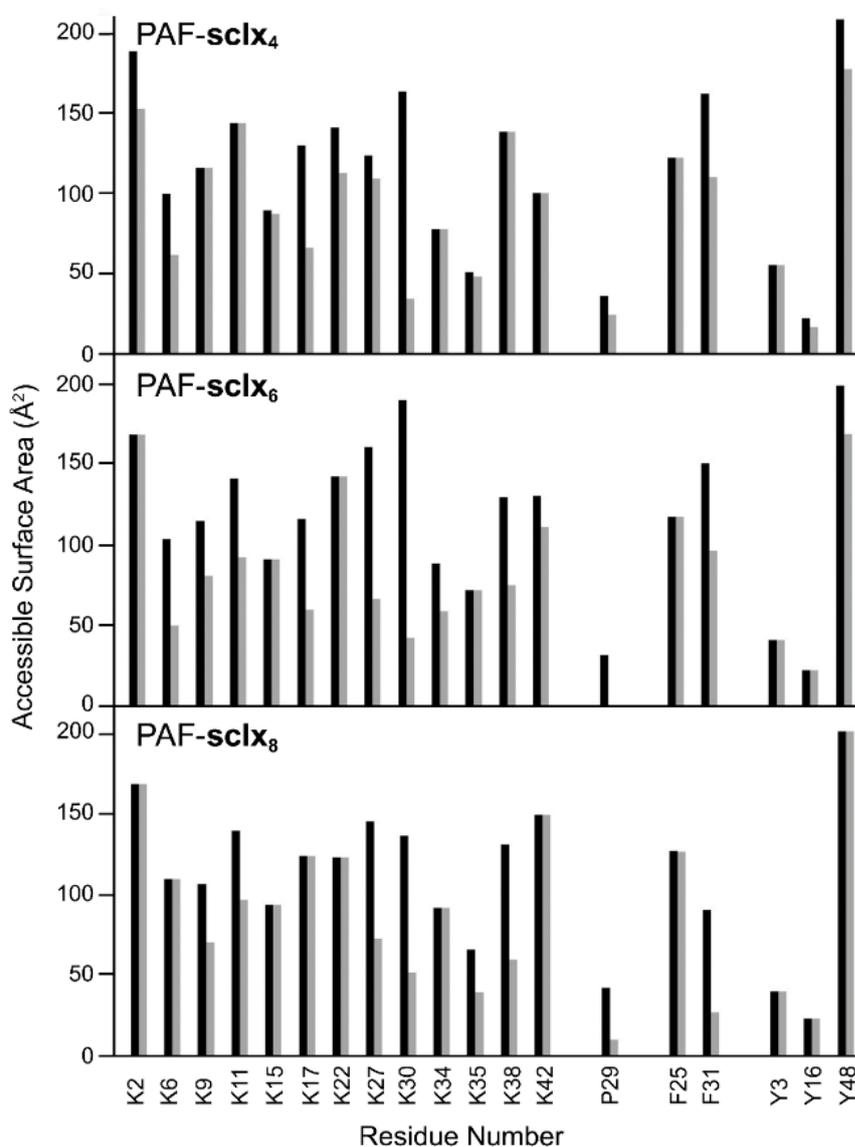


**Figure 6.** Protein-PEG-calixarene interfaces. The protein-calixarene interfaces are completed by a PEG fragment in **(A)** PAF – **sclx<sub>6</sub>** and **(B)** PAF – **sclx<sub>8</sub>**. Lys9-N<sup>ζ</sup> simultaneously forms ion-dipole bonds to the PEG (crown-ether like complex) and a salt bridge to one sulfonate. CH- $\pi$  and lone pair- $\pi$  bonds occur also between PEG and the calixarene phenolic rings.

The presence of PEG fragments (EG4 and EG7) markedly distinguished the PAF – **sclx<sub>6</sub>** and PAF – **sclx<sub>8</sub>** complexes (Figure 6). The PEG-calixarene interaction involved lone pair- $\pi$ <sup>275</sup> and CH- $\pi$  bonds, while the PEG-protein contacts included hydrogen bonds between the oxygen lone pairs and Lys9 (Lys9-N<sup>ζ</sup>...O-PEG = 3.0-3.3 Å). This crown ether like Lys9-PEG interaction resembles the binding of lysine to 18-crown-6 (PDB entry 3wur).<sup>75</sup> A heptaethylene glycol fragment has been observed

bound to an antibody (PDB entry 2ajs),<sup>276</sup> where it adopted a crown-ether like conformation, compared with the extended conformation in PAF – **sclx<sub>8</sub>**. In addition, a crystal structure of an SH3 domain (PDB entry 5xg9)<sup>277</sup> revealed various PEG fragments at protein–protein interfaces. These examples suggest that the role of PEG is as an interface ‘filler’ and possibly the PEG fragments (Figure 6) contribute towards calixarene conformation selection/stability.

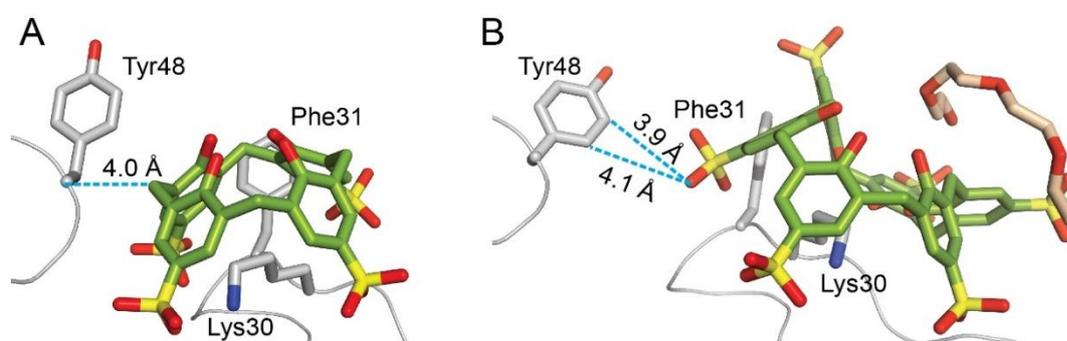
**Selectivity of PAF – sclx<sub>n</sub> complexation, why Lys30?** Considering that PAF contains 13 lysines the question arises as to why Lys30 was selected by **sclx<sub>n</sub>**. ASA calculations were used to probe the selectivity of **sclx<sub>n</sub>** for the Pro29-Lys30-Phe31 patch over other possible binding sites (Figure 7). The calculations accounted for contributions from symmetry mates in the crystal packing.<sup>53</sup> The effect of ligand binding on the ASA of all Lys, Pro, Phe and Tyr residues is plotted in Figure 7. At least half of the lysines, including Lys30, are highly exposed (ASA 125 Å<sup>2</sup>) in each structure in the absence of **sclx<sub>n</sub>**. This observation suggests that steric accessibility<sup>49</sup> was not the determining factor in **sclx<sub>n</sub>** selectivity. For example, Lys2 (>150 Å<sup>2</sup>) was significantly masked (ASA≥15%) by binding with **sclx<sub>4</sub>** only. Perhaps a salt-bridge interaction with Asp46 reduced the availability of Lys2 in the other complexes. In contrast, Lys30 was strongly affected by all three calixarenes (ASA up to 80 %). Adjacent residue Lys27 was also strongly affected in the complexes with **sclx<sub>6</sub>** and **sclx<sub>8</sub>**. The differences in the degree of masking can be attributed to the calixarene sizes (small, **sclx<sub>4</sub>**) and conformations (‘double cone’, **sclx<sub>8</sub>**). However, **sclx<sub>8</sub>** had more in common with **sclx<sub>6</sub>** than **sclx<sub>4</sub>**. For example, Lys9, Lys11 and Lys38 were 30–50% buried by **sclx<sub>6</sub>** or **sclx<sub>8</sub>**, while **sclx<sub>4</sub>** had no effect on these residues. Overall, calixarene binding resulted in significant masking of five (**sclx<sub>4</sub>**), eight (**sclx<sub>6</sub>**) and six (**sclx<sub>8</sub>**) lysines.



**Figure 7.** Accessible surface area (ASA) plots. Accessibility of Lys, Pro, Phe and Tyr residues in ligand-free (black) and -bound (grey) PAF. The PAF – **sclx<sub>8</sub>** data correspond to chain A.

PAF has five aromatic residues, Phe25, Phe31, Tyr3, Tyr16 and Tyr48 (Figure 7); the latter is highly solvent exposed ( $200 \text{ \AA}^2$ ) and might be expected to interact with **sclx<sub>n</sub>**. However, only minor contributions were evident (Figure 8). Phe31 was the dominant aromatic residue for **sclx<sub>n</sub>** complexation. The adjacent Lys30, Lys34 and

Lys35 may facilitate (*via* charge–charge interactions) calixarene binding here, compared with Tyr48, which is proximal to Lys2 only. The contribution of Pro29 merits special attention as it completes the binding site for both **sclx<sub>6</sub>** and **sclx<sub>8</sub>** *via* face-to-face hydrophobic stacks with a phenolic ring (Figures 4B and 4C). These interactions are reminiscent of polyphenol binding to proline-rich proteins.<sup>278-280</sup> The rigid pyrrolidine ring appears to provide a stable platform for binding the ‘floppy’ **sclx<sub>6</sub>** or **sclx<sub>8</sub>**. Thus, it is perhaps unsurprising that the only proline residue in PAF was involved at the binding site.



**Figure 8.** Tyr48-calixarene Interactions. **(A)** Tyr48-C $\alpha$  in van der Waals contact with a methylene bridge of **sclx<sub>4</sub>** **(B)** Tyr48-C $\delta^2$  forms a weak anion-quadrupole bond with a sulfonate of **sclx<sub>6</sub>**.

As such, it appears to be the combination of the Pro29-Lys30-Phe31 motif and adjacent lysines (charge–charge interactions) that stabilize **sclx<sub>n</sub>** binding and impart selectivity. This region has been implicated in PAF function, with decreased antifungal activity when Phe31, Lys35 or Lys38 were mutated to Asn or Ala.<sup>165,166,266</sup> The selectivity of the anionic calixarenes for this site suggests that it may be involved in cell membrane binding and permeation as required for antifungal activity.

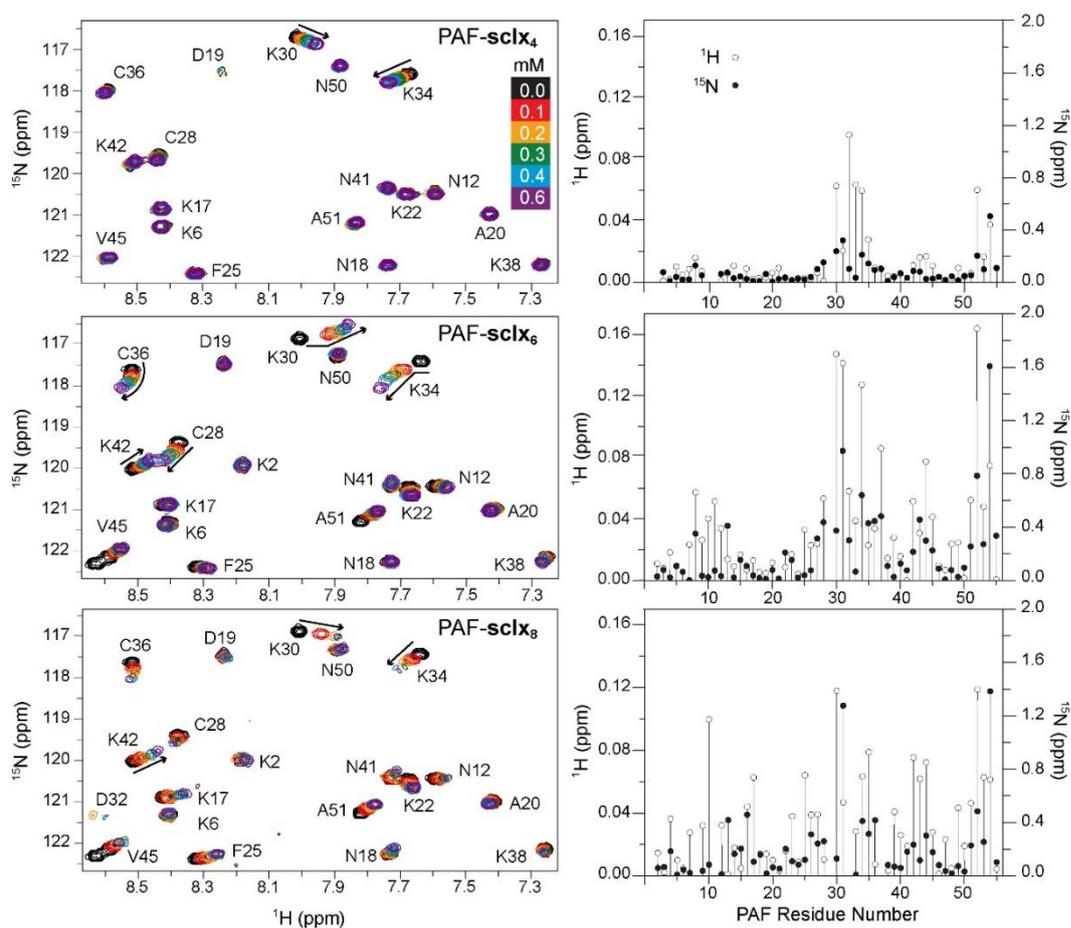
**NMR characterization and comparison with the solid state.** PAF – calixarene binding in solution was assessed by NMR spectroscopy. Titrations were performed by the addition of microlitre aliquots of **sclx<sub>n</sub>** to <sup>15</sup>N-labelled PAF, which was monitored by

$^1\text{H}$ - $^{15}\text{N}$  HSQC spectroscopy.<sup>48,191</sup> The overlaid spectra (Figure 9) revealed increasing chemical-shift perturbations ( $\Delta\delta$ ) as a function of **sclx<sub>4</sub>** or **sclx<sub>6</sub>** concentration, indicative of fast to intermediate exchange between the ligand-free and ligand-bound states. Some biphasic shifts were evident for **sclx<sub>6</sub>**. Severe broadening effects were observed with 0.3 eq **sclx<sub>8</sub>**, indicative of a slow-exchange process and suggesting the possibility of ligand-mediated oligomerization.<sup>19,51,52,184,281</sup>

The  $\Delta\delta$  plot (Figure 9) shows a clear selectivity for **sclx<sub>4</sub>** binding to Lys30 and neighbouring residues 31–36. In the crystal structure, all of these residues occurred in the vicinity of **sclx<sub>4</sub>**. Significant shifts were observed also for the C-terminal Val52 and Cys54, which are further from the crystallographic binding site. However, both of these residues are adjacent to Pro29, and Cys54 is hydrogen bonded to Lys34, suggesting a mechanism for how these resonances sense ligand binding. In the presence of **sclx<sub>6</sub>**, the plot again shows a preference for binding around Lys30 as well as effects at the C-terminus (Val52 N $^{\alpha}$  is hydrogen bonded to **sclx<sub>6</sub>**). However, compared with **sclx<sub>4</sub>**, the shifts are 2–4 times larger and other segments of the primary structure (residues 6–13 and 42–45) were also affected. These two regions correspond to additional **sclx<sub>6</sub>** binding sites evident in the crystal packing. Therefore, the NMR data suggests that the PAF – **sclx<sub>6</sub>** interaction fluctuates, with the calixarene exploring different patches on the protein surface, as observed previously for cytc – **sclx<sub>4</sub>** complexes.<sup>48,52</sup> Judging from the magnitude of the shifts, binding to Lys30 is preferred while a weaker interaction occurred at a patch involving Lys6 and Lys42.

The titrations with **sclx<sub>8</sub>** resulted in different effects. In addition to pronounced perturbations of Lys30 and neighbours, substantial broadening effects occurred. Cys28, Lys30, Lys34 and Cys36 broadened at 0.3 mM, and Thr8, Lys11, Asp32 and Thr37 broadened beyond detection at 0.6 mM **sclx<sub>8</sub>**. These eight residues are located at the crystallographically defined binding site. Thus, the broadening effects may be indicative of PAF dimerization, consistent with the **sclx<sub>8</sub>**-mediated

dimer in the crystal structure (Figure 4C). Previously, we observed a complete loss of the HSQC spectrum of cytc in complex with **pclx<sub>6</sub>**, which also yielded a dimer in the solid state.<sup>51</sup>



**Figure 9.** NMR characterization of PAF – sclx<sub>n</sub> complexation. **(A)** Region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of pure PAF (black contours) and in the presence of 0.1-0.6 mM ligand (coloured scale). Biphasic shifts occurred for resonances K30, K34 and C36 in the presence of sclx<sub>6</sub>. Resonances K17, C28, K30, K34, and C36 were broadened at 0.3 mM sclx<sub>8</sub> while resonances T8, K11, D32, T37 were broadened beyond detection at 0.6 mM sclx<sub>8</sub>. **(B)** Plots of chemical shift perturbations measured for PAF backbone amides in the presence of 0.6 mM sclx<sub>4</sub>, sclx<sub>6</sub> or sclx<sub>8</sub>. Blanks correspond to Pro29 and undetectable resonances (due to broadening).

**Thermodynamics of PAF – sclx<sub>n</sub> complexation.** Isothermal titration calorimetry was used to characterize the PAF – sclx<sub>n</sub> binding affinities and stoichiometries (Figure 10). The data were fitted to a single-site or a bidentate-ligand model. The latter model describes a bidentate ligand that can bind two protein molecules and was necessary to describe the obviously biphasic data for sclx<sub>8</sub>. The choice of this model is supported by the observation of a PAF – sclx<sub>8</sub> – PAF dimer in the crystal structure, and by the spectral broadening in the NMR experiments. All of the fit parameters were well determined by the data (Table 2), with parameter errors assessed by Bayesian methods.<sup>282</sup>

**Table 2.** Thermodynamic of PAF – sclx<sub>n</sub> complexation determined by ITC.\*

| [Ligand] <sup>a</sup><br>( $\mu\text{M}$ )       | [PAF] <sup>b</sup><br>( $\mu\text{M}$ ) | $K_d$<br>( $\mu\text{M}$ ) | $\Delta H$<br>(kJ/mol) | $T\Delta S$<br>kJ/mol) |
|--|---|----------------------------|------------------------|------------------------|
| PAF – sclx <sub>4</sub> (Single Site Model)      |   |                            |                        |                        |
| 7143<br>(1248)                                   | 500 (412)                               | 107 (0.0, 0.0)             | -16.9 (0.1, 0.1)       | -5.6 (0.2, 0.2)        |
| PAF – sclx <sub>6</sub> (Single Site Model)      |   |                            |                        |                        |
| 3623 (633)                                       | 500 (412)                               | 15.4 (0.0, 0.0)            | -28.2 (0.2, 0.1)       | -0.7 (0.2, 0.2)        |
| PAF – sclx <sub>6</sub> (Bidentate Ligand Model) |   |                            |                        |                        |
| 3623 (633)                                       | 500 (412)                               | 47.8 (0.0, 0.0)            | -9.2 (0.1, 0.2)        | -15.4 (0.4, 0.3)       |
|  |   | 45.8 (3.5, 4.1)            | -20.1 (0.4, 0.4)       | -4.6 (0.6, 0.6)        |
| PAF – sclx <sub>8</sub> (Bidentate Ligand Model) |   |                            |                        |                        |
| 2500 (437)                                       | 300 (247)                               | 10.6 (1.3, 1.4)            | -3.5 (0.3, 0.3)        | -24.8 (0.3, 0.3)       |
|  |   | 33.5 (4.8, 6.5)            | -36.0 (1.7, 1.4)       | 10.5 (1.8, 2.1)        |

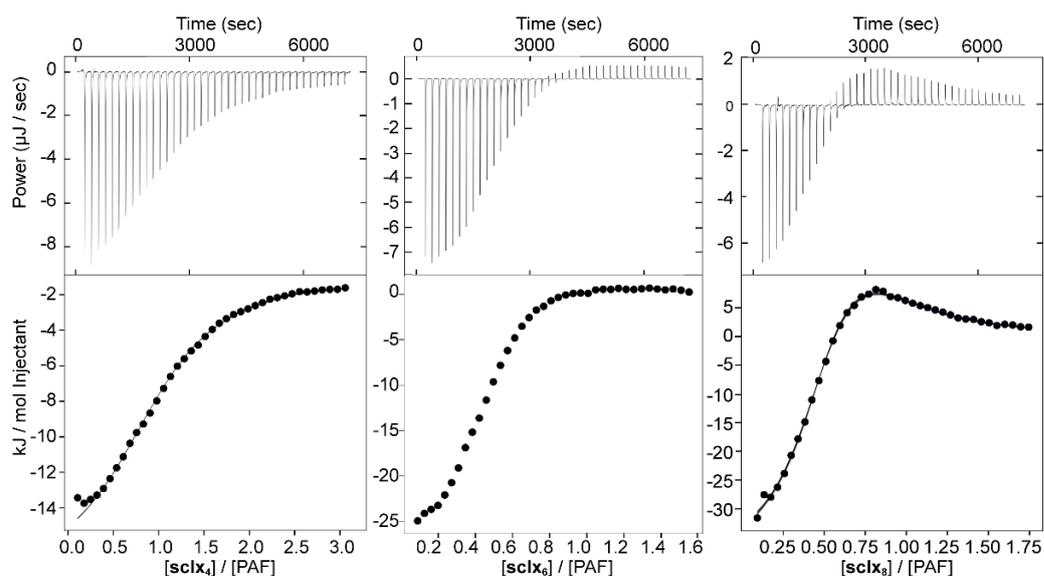
\*Fit values are median (2.5% quantile, 97.5% quantile) from Markov chain Monte Carlo. In the case of sclx<sub>6</sub> the fit parameters for both models are shown.

<sup>a</sup>Calixarene concentrations in the syringe, final concentration indicated in parenthesis.

<sup>b</sup>Protein concentration in the cell, final concentration indicated in parenthesis.

The isotherms for sclx<sub>4</sub> injected into PAF were fitted to a single-site binding model with  $K_d$  110 mM. In contrast, the isotherms for sclx<sub>8</sub> were biphasic<sup>283</sup> and fitted to a bidentate ligand model with  $K_d$  values of 10 and 30 mM, for binding the first and second molecule of PAF, respectively. The isotherms for sclx<sub>6</sub> were intermediate

between **sclx<sub>4</sub>** and **sclx<sub>8</sub>**, suggesting that this ligand may exhibit weak bidentate binding. A satisfactory fit for this data was not obtained with either model. The ITC data demonstrate an increasing affinity for PAF as the calixarene size increases and a switch in binding mode from the small, rigid **sclx<sub>4</sub>** (single site) to the large, flexible **sclx<sub>8</sub>** (bidentate).



**Figure 10.** ITC analysis of PAF - **sclx<sub>n</sub>** complexation. Top panels show the baseline corrected thermograms for injections of **sclx<sub>4</sub>**, **sclx<sub>6</sub>** or **sclx<sub>8</sub>** into PAF. Bottom panels are the observed heats (data points) and the fits (solid line) for single site (**sclx<sub>4</sub>**) and bidentate ligand (**sclx<sub>8</sub>**) models. Note, the titration heats for **sclx<sub>6</sub>** were intermediate between **sclx<sub>4</sub>** and **sclx<sub>8</sub>**.

## Conclusions

Using a combination of X-ray crystallography and NMR spectroscopy it was demonstrated that the **sclx<sub>n</sub>** series binds selectively to the highly cationic PAF. Despite the varying size and conformational flexibility, **sclx<sub>4</sub>**, **sclx<sub>6</sub>** and **sclx<sub>8</sub>** bound similarly the Pro29-Lys30-Phe31 motif in loop3. The selectivity of the anionic calixarenes for this motif, and the role of loop 3 in antifungal activity, suggests that this region may be required for membrane binding. In addition to charge–charge

interactions (showed by numerous lysine-to-sulfonate salt bridges), other noncovalent bonds including CH- $\pi$  and  $\pi$ - $\pi$  (*via* Pro29 and Phe31, respectively) participated in ligand stabilization. The presence of PEG fragments at the protein – **sclx**<sub>6</sub> and protein – **sclx**<sub>8</sub> interfaces suggests that PEG acts as a ‘filler’ to complete the binding site, potentially reinforcing the calixarene conformation.

The structures of all three PAF – **sclx**<sub>n</sub> co-crystals highlight the potential of calixarenes as a ‘sticky patch’ on the protein surface that facilitates assembly and crystallization. In the case of the **sclx**<sub>4</sub> and **sclx**<sub>6</sub> co-crystals (*P*12<sub>1</sub>1), it is evident that the calixarene is a dominant contributor to the crystal packing (Fig. 4). Similarly in the **sclx**<sub>8</sub> structure (*P*6<sub>1</sub>), the packing involves substantial protein – calixarene contacts, and the structure is interesting as **sclx**<sub>8</sub> mediates a PAF dimer. Previously, we found that **sclx**<sub>8</sub> mediates a tetramer of cytc.<sup>19</sup> Generally, it seems that calixarene-mediated protein crystallization may be a special case of the patchy particle model for protein assembly.<sup>34,53,62-64</sup> Considering that PAF alone did not yield diffraction-quality crystals, we conclude that co-crystallization with **sclx**<sub>n</sub> was beneficial. Anionic calixarenes may generally facilitate crystallization and structure determination of small cationic proteins.

The binding surfaces observed in the NMR experiments were consistent with the X-ray data. However, the NMR effects were more pronounced with increasing calixarene size, suggesting that the larger calixarenes mask a greater portion of the protein surface and/or lead to assembly in solution. Similarly, the ITC experiments revealed tighter affinities and more complex effects with increasing calixarene size. In particular, **sclx**<sub>8</sub> behaved as a bidentate ligand that facilitated PAF dimerization. These data add to the growing evidence of calixarene-mediated protein assembly in solution.<sup>51,52</sup> In terms of the biological relevance of these data it is noted that defensin oligomerization (upon phospholipid binding) has implications for anti-fungal activity.<sup>268-271</sup> Perhaps calixarenes can be used to modulate the activity of PAF and related proteins.



## Chapter 5

---

### Crystal Structure of a Protein – Aromatic Foldamer Composite: Macromolecular Chiral Resolution

The material in this chapter was published as:

**Alex J. M.**, Corvaglia V., Hu X., Engilberge S., Huc I., Crowley P. B.

Crystal structure of a protein – aromatic foldamer composite: macromolecular chiral resolution.

*Chem Commun.* **2019**, 55, 11087-11090.

Ivan Huc's Lab (Ludwig-Maximilians-Universität, Germany) –

- Synthesis and characterization of foldamers **1** and **2**.
- Circular dichroism studies of cytc – foldamer **2** complex.

**Abstract**

Oligoamide foldamers are being developed as synthetic ligands for protein surface recognition. Recent work with quinoline-based helical foldamers has relied on a tether functionality to direct protein binding. Here, we characterized complex formation between a ~2 kDa tether-free sulfonated foldamer and the ~13 kDa lysine-rich cytochrome *c*. Co-crystallization led to the spontaneous chiral resolution of the foldamer helix handedness. A 2.1 Å crystal structure revealed foldamer stacking and charge-charge interactions with the protein. Zinc ions were important also for the crystalline assembly. By mass, the structure comprised ~1:3 foldamer:protein, hinting at a biohybrid material. Despite this composition the individual foldamer – protein interfaces were small ( $\leq 180 \text{ \AA}^2$ ). Complexation in solution was studied by NMR and identified a different binding patch to those observed by X-ray. Circular dichroism experiments also highlighted differences in helix handedness preference, with respect to the crystal structure. Foldamer – protein recognition and assembly are discussed with suggestions for future foldamer designs.

## Introduction

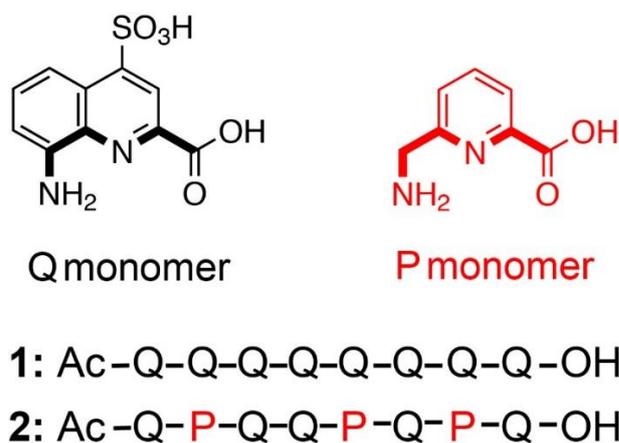
Supramolecular chemistry provides a repertoire of abiotic synthetic receptors and ligands for protein recognition and assembly.<sup>2,4</sup> These properties have been illustrated for calixarenes,<sup>19,20,55,56</sup> cucurbiturils,<sup>21,59</sup> aromatic foldamers,<sup>98-102,284,285</sup> suramin<sup>117,118</sup> and tweezers,<sup>23,61</sup> all of which complement medchem products and peptidomimetics. Among the supramolecular components, foldamers are unique in that they form stably-folded architectures in solution.<sup>94-96</sup> This feature, arising from their resemblance to biopolymers, makes them attractive candidates for protein recognition.<sup>286,287</sup> Currently, there is growing interest in helical aromatic oligoamides.<sup>288,289</sup> The stability, fold predictability and ease of synthesis / functionalization of these foldamers have led to a wide range of applications.<sup>95,289-294</sup>

Aromatic foldamers can be customized to enhance biocompatibility. For example, oligoamides of 8-amino-2-quinoline carboxylic acid (Q, Scheme 1) were decorated with proteinogenic side chains.<sup>100</sup> Additionally, they can be endowed with a tether functionality that confines the foldamer to a specific region of the protein surface, enabling the identification of weak binding interactions.<sup>100,101,295</sup> For example, foldamers with a benzenesulfonamide were anchored to the active site of human carbonic anhydrase II (HCA), thus facilitating recognition and assembly,<sup>100-102</sup> while others were linked to cyclophilin A or interleukin 4.<sup>295</sup> The *ab initio* design of foldamers (without a tether) that can bind proteins is a challenging task owing to the dearth of molecular recognition information. Taking inspiration from the charge-charge multivalent complexation between a cationic quinoline oligoamide and a DNA G-quadruplex,<sup>296</sup> the present work was carried out with the objective to demonstrate tether-free foldamer binding to a protein.

We explored the recognition properties of an anionic helical quinoline oligoamide with a lysine-rich protein. The foldamer was a ~2 kDa octamer spanning three helix turns functionalized with sulfonate groups (**1**, Scheme 1) that could be

synthesized readily on solid phase using previously reported procedures.<sup>297</sup> The choice of sulfonic acid was motivated by its suitability to interact with cationic residues.<sup>14,39,43,46,47,109</sup> Cytochrome *c* (cytc, 13 kDa, *pI* ~9) was selected as its interactions with anionic receptors such as sulfonato-,<sup>19,20,48,52,56</sup> phosphonomethyl-<sup>53</sup> and phosphonato-calix[*n*]arenes are well-established.<sup>51</sup>

A 2.1 Å resolution crystal structure provided a detailed view of protein – foldamer binding. Remarkably, the co-crystal resulted in the chiral resolution of the *P* and *M* helices of **1** the interconversion of which is kinetically hampered. NMR studies were performed to characterize complexation in solution. Furthermore, circular dichroism (CD) experiments with **2**, an analogue of **1** containing flexible aminomethyl-pyridine units (*P*, Scheme 1) that make helix handedness inversion possible, provided insight into helix handedness induction by the protein.<sup>298</sup> Finally, the influence of the ~2 kDa helical ligand (**1**) on recognition and assembly of a ~13 kDa globular protein is discussed and suggests new possibilities of composite bio-materials.<sup>299</sup>



**Scheme 1.** General structures of amino acid quinoline (Q) and pyridine (P) monomers, and sequences of helical oligoamide foldamers **1** and **2**. The inner rim of the helix is marked in bold.

## Experimental

**Sample preparation.** Millimolar stock solutions of **1** were prepared in 1 M sodium acetate or water whereas foldamer **2** was prepared in water. <sup>15</sup>N-labelled and unlabelled *Saccharomyces cerevisiae* cytochrome *c* (C102T) was prepared as described previously.<sup>48,191</sup>

**Co-crystallization trials.** Trials were performed with an Oryx 8 Robot (Douglas Instruments) and a sparse matrix screen (JCSG++, Jena Bioscience) at 20° C. Cytochrome *c* (1 mM) and **1** (in sodium acetate) were tested in ratios of 1:1 and 1:4. Crystal growth was achieved with 4 mM of **1** only in condition C7, 10 % PEG 3000, 0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate. These rod-shaped crystals were reproducible by the hanging drop method in 5 - 25 % PEG 3350. Trials with **1** in water yielded thin plates in conditions A9 (20 % PEG 3350, 0.2 M ammonium chloride pH 6.3), G7 (15 % PEG 3350, 0.1 M di-sodium succinate pH 7.0) and G8 (20 % PEG 3350, 0.1 M di-sodium DL-malate pH 7.0). Thin plates were obtained with **2** (in water) in conditions G7 and H6 (17 % PEG 10,000, 0.1 M Bis-TRIS pH 5.5, 0.1 M ammonium acetate).

Co-crystallization trials with 8 other foldamers resulted in either poor diffraction or ligand free structure (Table 1)

**X-ray data collection.** Crystals of ~100 μm dimension were cryo-protected in the reservoir solution supplemented with 20 % glycerol and cryo-cooled in liquid nitrogen. The rod-shaped crystals of **1** at 5 % PEG 3350 diffracted to 2.1 Å resolution at beamline PROXIMA-2A (SOLEIL synchrotron). Datasets were collected using  $\phi$  scans of 0.1° over 360° with an Eiger X 9M detector. The thin plates with **1** or **2** diffracted poorly.

**X-ray structure determination.** Diffraction data were integrated and scaled using the autoPROC pipeline (Table 2).<sup>215</sup> The structure was solved by molecular replacement in PHASER, with 5lyc as the search model.<sup>195</sup> JLigand was used to generate the foldamer coordinates and refinement restraints.<sup>196</sup> Iterative cycles of model building in COOT<sup>197</sup> and refinement in BUSTER<sup>217</sup> were performed until no further improvements in the  $R_{\text{free}}$  or electron density were obtained. The final structure was validated with MolProbity<sup>199</sup> and deposited in the Protein Data Bank as PDB 6s8y.

**Circular Dichroism.** CD spectra were recorded on a Jasco J-810 spectrometer using quartz cells with a 1 mm path length.<sup>298</sup> Scans were acquired at 20 °C, over the 300-500 nm range, with a 1 s response time and a 100 nm/min scan speed. The CD data (Figure 7) are an average of three scans. Samples contained 90  $\mu\text{M}$  cytochrome *c* and 1 eq. **2** in 25 mM sodium acetate, 25 mM sodium chloride at pH 5.4 and were equilibrated for a minimum of 24 h. Data were baseline-corrected for signal contributions due to the buffer and protein.

**NMR spectroscopy.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC monitored titrations were performed at 30°C on a 600 MHz Varian spectrometer equipped with a HCN coldprobe, as described previously.<sup>48</sup> Samples of 0.1 mM  $^{15}\text{N}$ -labelled cytochrome *c* were titrated with 3 - 24  $\mu\text{L}$  aliquots of 8 mM ligand (in water) in 20 mM potassium phosphate, 50 mM sodium chloride at pH 6.0. Spectra were processed in NMRPipe and chemical shift perturbations were determined in CCPN.<sup>300</sup>

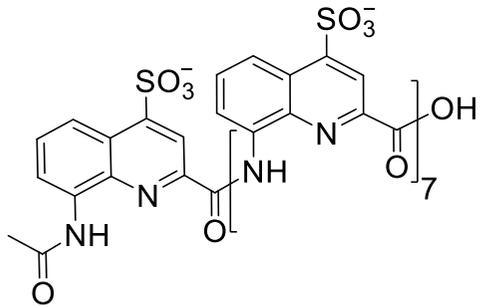
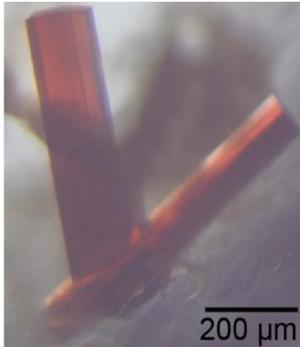
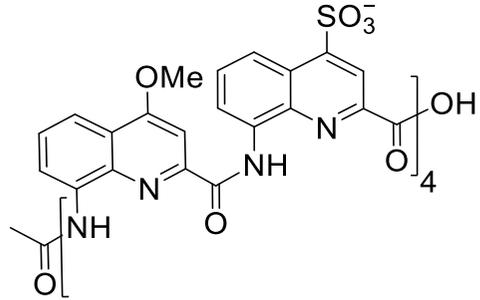
## Results and Discussion

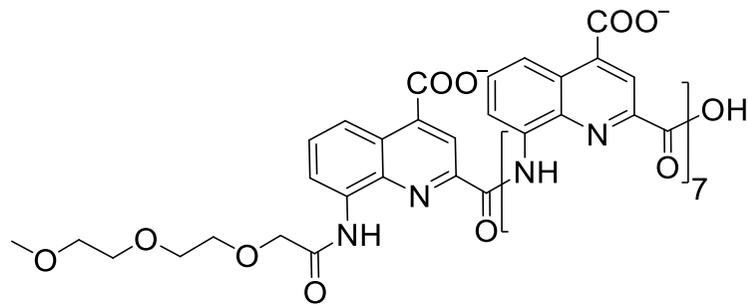
**Cytochrome *c* – 1 co-crystallization.** Initial crystallization trials of cytc – **1** mixtures were performed with a sparse matrix screen (Jena Bioscience, JCSG++) on an Oryx 8 robot (Douglas Instruments). Crystals grew only in condition C7, containing 10 % PEG 3000, 0.1 M sodium acetate pH 4.5 and 0.2 M zinc acetate. Similar conditions B7 (8 % PEG 4000 and 0.1 M sodium acetate pH 4.6) and D4 (30 % PEG 8000, 0.1 M sodium

acetate pH 4.5 and 0.2 M lithium sulfate) lacking zinc, did not yield crystals. Rod-like crystals (Table 1) were reproducible with 5 – 25 % PEG 3350 and 4 eq. of **1** prepared in 1 M sodium acetate. Thin plates co-crystals were obtained also from a stock of **1** or **2** in pure water.

Co-crystallization trials with other quinoline-based foldamers failed to yield diffraction-quality crystals (Table 1).

**Table 1.** Cytc – foldamer co-crystallization trials

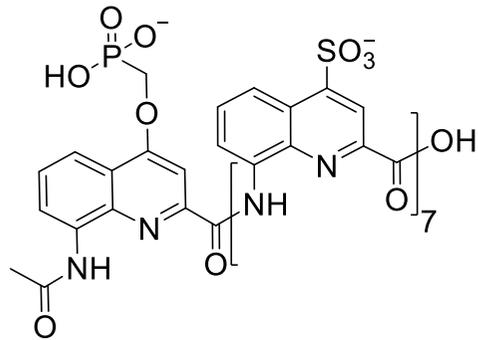
| Molecular structure  | Molecular weight | Representative crystals   | Diffraction / Resolution                                     |
|--|------------------|---|--|
|   | 2061.89          |  | 2.1 Å<br>Good diffraction;<br>structure solved<br>(PDB 6s8y) |
|  | 1861.74          | –No–  | –No–   |



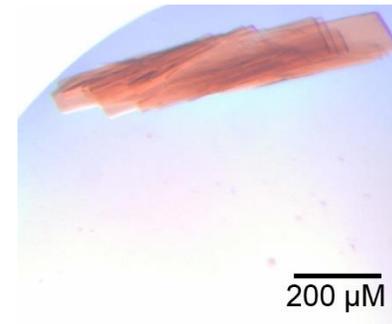
2003.08

-No-

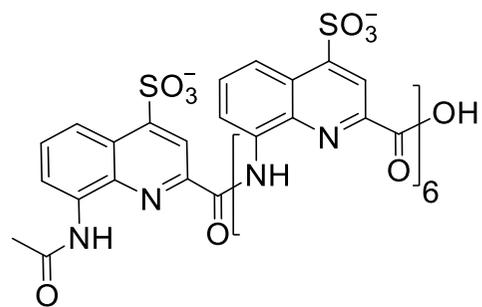
-No-



2090.07



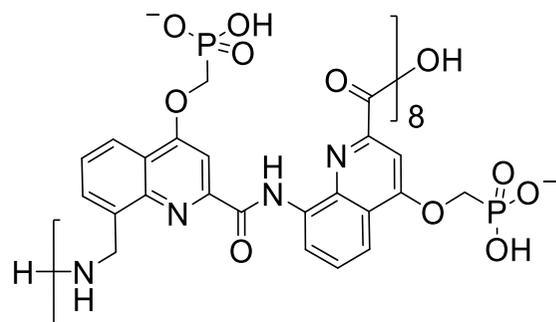
1.7 Å  
Difficult to solve



1810.05



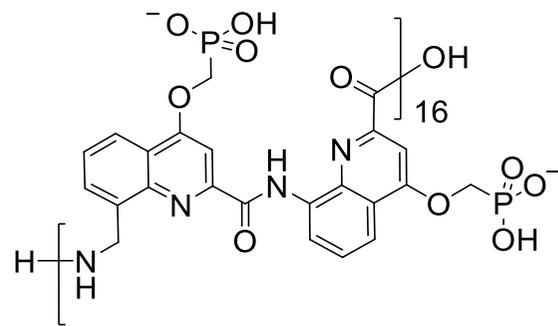
2.8 Å  
Poor diffraction



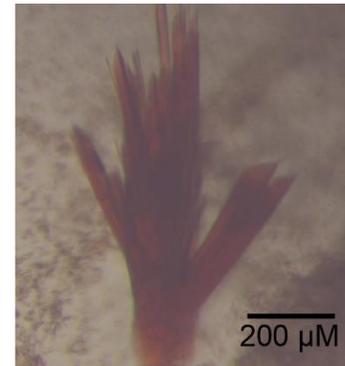
4610.53

-No-

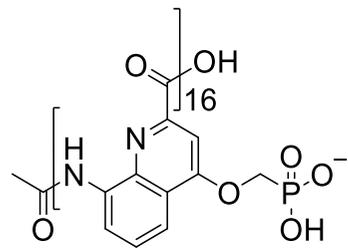
-No-



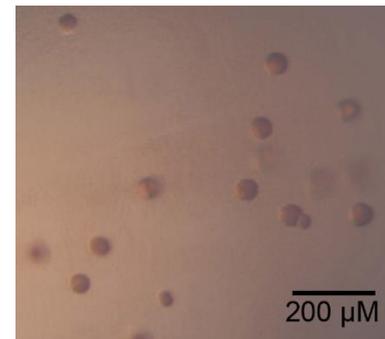
9203.05



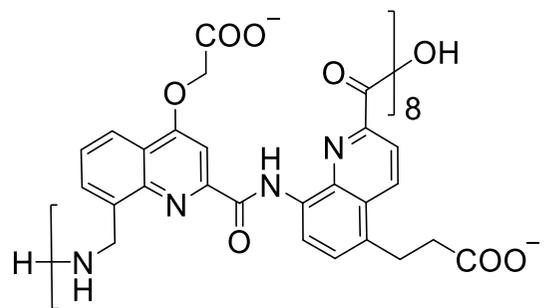
1.6 Å  
Devoid of ligand



4540.41



No diffraction

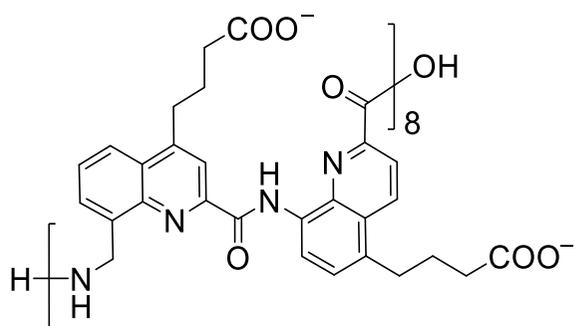


4019.07

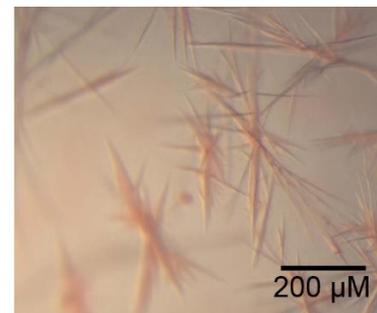
-No-

-No-

---



4227.49



No diffraction

---

**Data collection and model building.** The plate crystals with **1** or **2** diffracted poorly to  $> 4$  Å resolution datasets. Data extending to 2.1 Å resolution was collected (SOLEIL synchrotron) from the rod-like cytc – **1** crystals (Table 1), which belonged to the trigonal space group  $P3_221$  (Table 2). The structure was solved by molecular replacement with one cytc in the asymmetric unit.

**Table 2.** X-ray data collection, processing and refinement statistics.

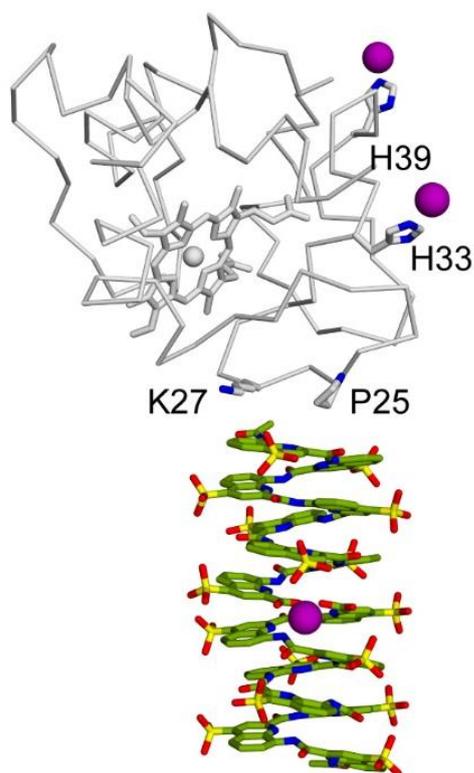
| <i>Data Collection</i>                            |  |
|---|--|
| Light source                                      | SOLEIL, PROXIMA-2A   |
| Wavelength (Å)                                    | 0.98009  |
| Space group                                       | $P3_221$   |
| Cell constants                                    | $a = b = 39.13 \text{ \AA}$ , $c = 187.36 \text{ \AA}$<br>$\alpha = \beta = 90^\circ$ , $\gamma = 120^\circ$ |
| Resolution (Å)                                    | 62.45-2.09 (2.12-2.09)   |
| # reflections                                     | 199568 (10386)   |
| # unique reflections                              | 10590 (521)  |
| Multiplicity                                      | 18.8 (19.9)  |
| $I/\sigma(I)$                                     | 19.0 (2.3)   |
| Completeness (%)                                  | 100 (100)  |
| $R_{\text{meas}}^b$ (%)                           | 8.2 (12.7)   |
| $R_{\text{pim}}^c$ (%)                            | 2.0 (28.3)   |
| $CC_{1/2}$  | 99.7 (85.6)  |
| Solvent content (%)                               | 50   |
| <i>Refinement</i>                                 |  |
| $R_{\text{work}}$                                 | 0.2290   |
| $R_{\text{free}}$                                 | 0.2780   |
| rmsd bonds (Å)                                    | 0.0101   |
| rmsd angles (°)                                   | 1.5249   |
| Asymmetric unit composition                       |  |
| protein   | 1  |
| foldamer  | 2  |
| zinc ion  | 3  |
| water   | 26   |
| Ave. B-factor (Å <sup>2</sup> )                   | 67.41  |
| Clashscore  | 2.71   |
| Ramachandran analysis, <sup>d</sup> % residues in |  |
| favoured regions                                  | 97.1   |
| allowed regions                                   | 100  |
| PDB code  | 6s8y   |

<sup>a</sup>Values in parentheses correspond to the highest resolution shell;

$$^b R_{\text{meas}} = \frac{\sum_{hkl} \sqrt{(n/n-1)} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$$

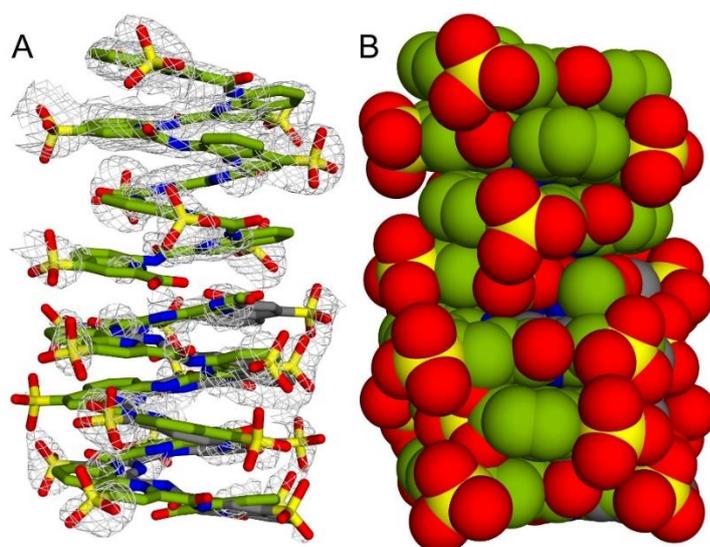
$$^c R_{\text{pim}} = \frac{\sum_{hkl} \sqrt{(1/n-1)} \sum_{i=1}^n |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)},$$

<sup>d</sup>Determined in MolProbity



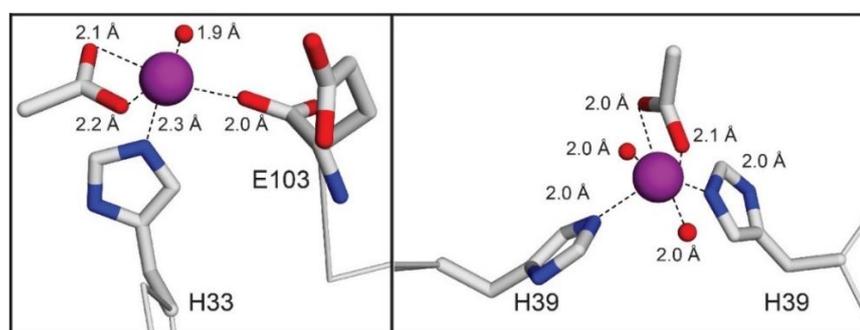
**Figure 1.** The asymmetric unit of cytc – **1** comprised one cytc and a stack of two foldamers (PDB id 6s8y). Purple spheres indicate zinc ions.

Electron density was evident for two molecules of **1** with *M*-helicity. One foldamer, adjacent to Pro25 and Lys27 was clearly defined, while the second foldamer was modelled in both head-to-head and head-to-tail orientations relative to the well-defined foldamer (Figure 1 and 2).

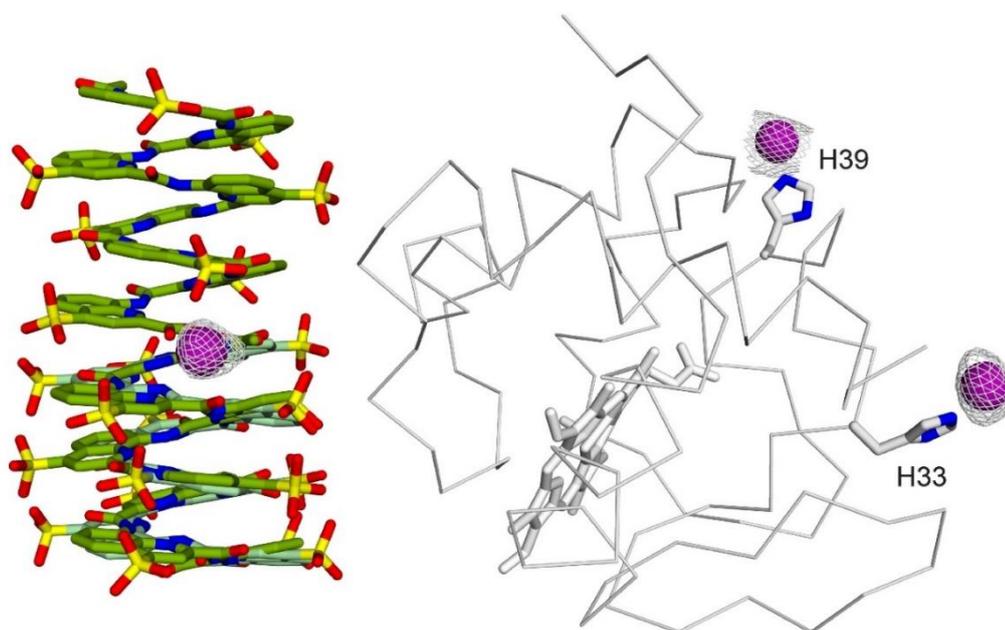


**Figure 2.** (A) Omit map showing the  $2F_o - F_c$  electron density (contoured at  $1.0 \sigma$ ) of the foldamers. The grey foldamer corresponds to the head-to-head orientation. (B) Spacefill representation of the foldamer stack depicting the cylinder with an anionic surface (hydrophobic quinoline rings are buried).

Zinc ions were modelled at three separate sites involving His33, His39 or the foldamer – foldamer interface. The zinc bound at His39 facilitated crystal packing (Figures 3 and 4).



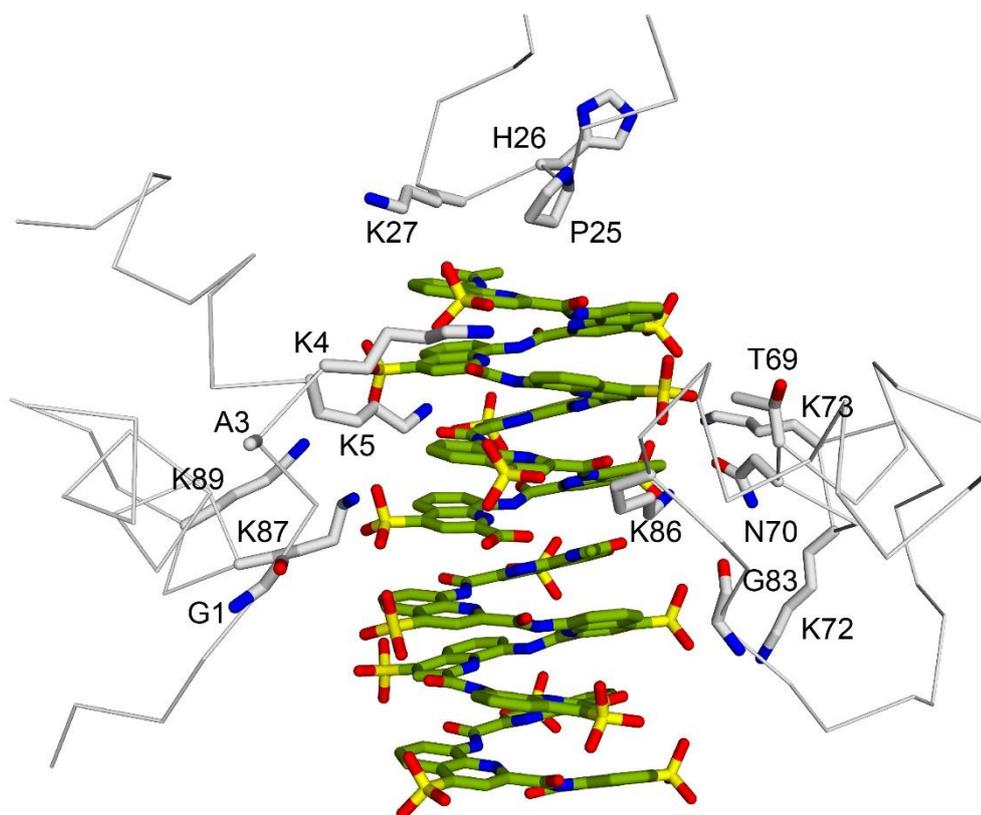
**Figure 3.** Zinc ions coordinated at His33 or His39 with acetate ions and waters completing the coordination spheres. At His39, the zinc is a bridging ion that contributes to protein assembly in the crystal.



**Figure 4.** Fourier map (contoured at  $3.0 \sigma$ ) confirmed the presence and position of three zinc ions.

**Foldamer bound, not tethered.** In the crystal structure, two symmetry unrelated molecules of oligoamide **1** were found to be arranged in a stack yielding a cylinder of  $\sim 2.3$  nm height, similar in dimension to cytc (Figure 1B). Left-handedness (*M*) was unambiguously assigned to both helices from the electron density map. The curved surface of the cylinder is anionic owing to the sulfonates whereas the ends are mostly hydrophobic (Figure 2B). In the stack, one foldamer orientation is well-defined with the N-terminal quinoline packed against the protein at Pro25 and Lys27 (Figure 1B). The other foldamer was modelled in opposite orientations (50 % occupancy; Figure 2), arranged head-to-head or head-to-tail relative to the well-defined foldamer. Previously, it was observed that in a stack of three foldamers the central foldamer (lacking a tether for HCA binding) had negligible protein interactions and was present in two orientations.<sup>101</sup> In the current structure the well-ordered foldamer participated in several protein interactions, which probably stabilized a single orientation and contributed to the selection of a single helix handedness.

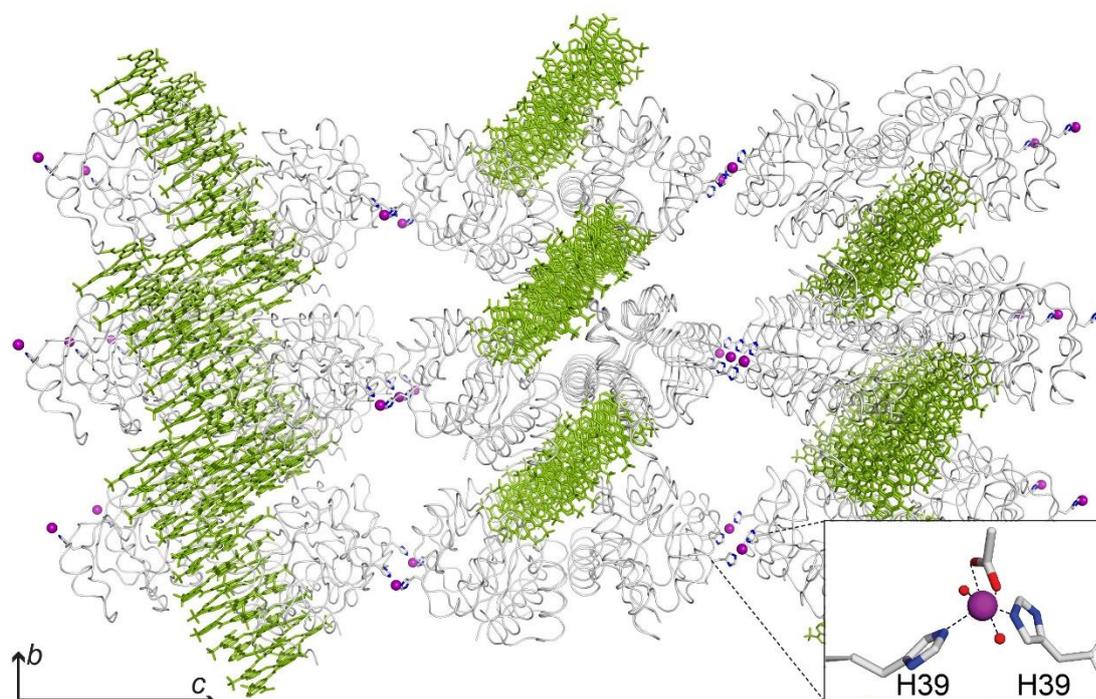
We considered the protein – foldamer interactions in terms of their interface areas. The ~2 kDa foldamer **1** with a ~1300 Å<sup>2</sup> surface area might be expected to form large interfaces with the protein. However, in the crystal, the well-defined foldamer formed small interfaces, ranging in size from 100 to 180 Å<sup>2</sup> with four neighbouring proteins (Figure 5 and 6). The disordered foldamer formed only two significant interfaces of ~165 or 190 Å<sup>2</sup> with cytc, which apparently were insufficient to dictate its orientation. With < 20 % of its surface involved, this foldamer was bound weakly. All of these interfaces were smaller than the main protein – protein (~270 Å<sup>2</sup>) as well as the foldamer – foldamer interfaces (~220 Å<sup>2</sup>). The similarity with typical protein crystal packing interfaces<sup>191</sup> suggests that the protein – foldamer contacts are weak and non-specific. Furthermore, the interfaces were comparable to those formed in protein – calix[4]arene complexes (~200 Å<sup>2</sup>).<sup>48,53,53</sup> The considerably smaller calix[4]arenes (~0.8 kDa) bear a cavity that encapsulate individual lysine residues leading to larger interface areas compared to the foldamer.



**Figure 5.** Molecular recognition details of the three largest protein – foldamer interfaces. The limited interactions of the second foldamer are evident, consistent with the orientational disorder. Protein, foldamer and zinc ions are depicted as grey ribbon, sticks and purple spheres, respectively.

Comparisons with sulfonatocalix[8]arene (**sclx<sub>8</sub>**, ~1.5 kDa)<sup>19</sup> and suramin (~1.3 kDa)<sup>118</sup> are also informative. While the largest cytc – **1** interface is 180 Å<sup>2</sup>, cytc – **sclx<sub>8</sub>** interfaces range from 400 to 550 Å<sup>2</sup>, as the calixarene moulds to the protein surface. The even more flexible suramin forms interfaces up to ~750 Å<sup>2</sup>. Although smaller than foldamer **1**, the calixarenes and suramin form larger interfaces, facilitating their ‘molecular glue’ activity. The compactness and relative rigidity of the helical foldamer may diminish this ability. Nevertheless, the foldamer nestles between four protein chains where its anionic surface complements the cationic patches of the surrounding proteins (Figure 5 and 6). The foldamer stack comprises 4 molecules in total, resulting in a ~4.6 nm long cylinder with a formal net charge of

36 (including the terminal carboxylates). The surrounding 8 molecules of cytc are bound loosely (small interfaces) and the overall structure has large channels and a 50 % solvent content (Figure 6). The role of zinc is noteworthy also as it bridges cytc molecules *via* His39 coordination (Figures 3 and 6). Zinc is prevalent in protein crystal structures as a mediator of assembly.<sup>301-302</sup>



**Figure 6.** Crystal packing involved a ~5 nm cylinder comprising 4 foldamers. A bridging zinc ion at His39 mediated assembly (inset).

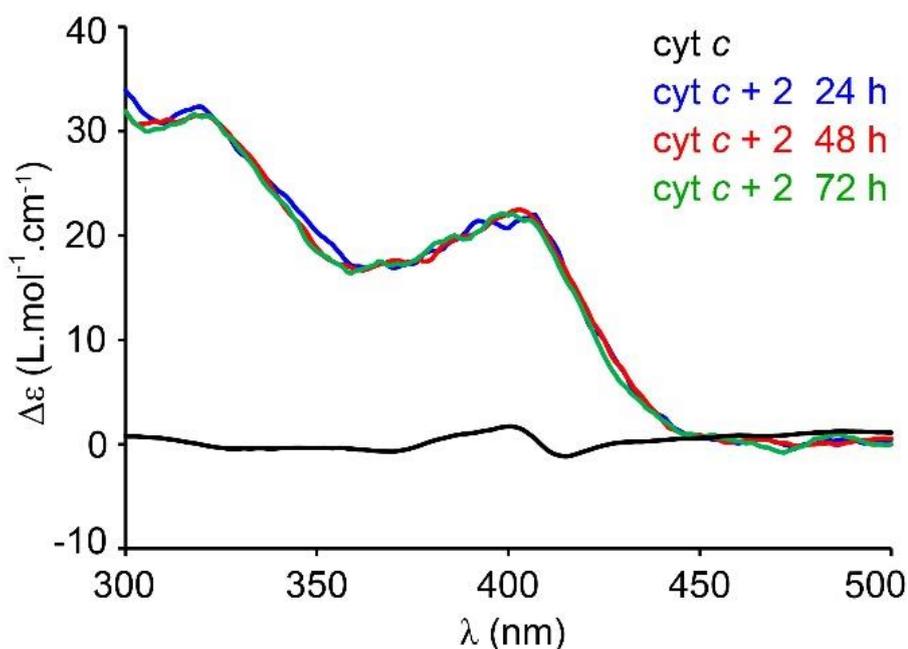
Notwithstanding the 2.1 Å resolution, close inspection of the crystal packing reveals multiple salt bridge interactions. In the well-defined foldamer, seven of the eight sulfonates make salt bridges to five lysine side chains ( $N^{\epsilon} \cdots O-S = 2.8 - 3.4 \text{ \AA}$ ) in four neighbouring proteins (Figure 5). Lys5 and Lys73 are salt bridged to two sulfonates while Lys86, Lys87 and Lys100 each interact with one sulfonate. Interestingly, Lys5 binds the N-terminal quinoline (the cylinder base) in combination with Pro25 and Lys27 from a neighbouring protein (Figure 5). These contacts among

others may have stabilized **1** in a single orientation and more importantly, selected the *M*-helix.

**Helix handedness selection and induction.** Compound **1** lacks stereogenic centres and is synthesized as a racemic mixture of *P* and *M* conformers. Helix handedness inversion in water is extremely slow for such quinoline-based octamers.<sup>298</sup> Therefore, the exclusive occurrence of the *M* conformer of **1** in the crystal structure amounts to a spontaneous chiral resolution. The resolution of racemic mixtures of chiral acids upon crystallization with a chiral base is well-established.<sup>303</sup> However, chiral resolution between molecules as large as a protein and a large helical foldamer is, to the best of our knowledge, unprecedented. It reflects diastereoselective interactions in the solid state that presumably are preceded by similar interactions in solution. We attempted to observe handedness selection in solution by circular dichroism using analogue **2**. In this sequence, synthesized by using reported procedures,<sup>297,298</sup> three quinoline monomers were substituted by aminomethyl-pyridine units (*P*, Scheme 1). These flexible units enhance the kinetics of helix handedness inversion making it possible to observe handedness bias in solution due to preferential interactions between one conformer and the protein.<sup>298</sup> *P* monomers are isosteric to *Q* (*i.e.* the same inner rim) but have additional rotatable bonds (increased flexibility), and lack side chain functionality.

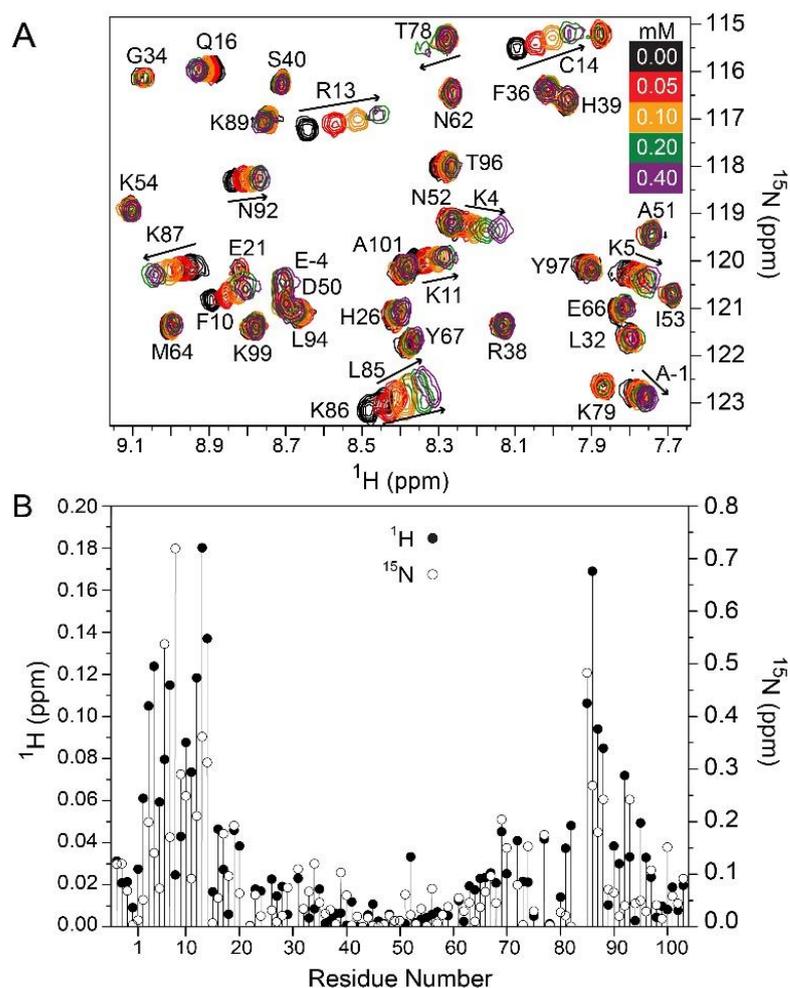
It was not possible to monitor solution interactions between cytc and **2** in the buffer that was used to grow the co-crystals with **1**. Under these conditions, co-precipitation with **2** occurred. Raising the pH to 5.4 and the addition of 25 mM sodium chloride made the mixture soluble. When **2** was equilibrated in the presence of cytc the CD spectrum revealed a positive band in the quinoline absorption region indicative of *P*-helicity, which remained unchanged over 72 h (Figure 7). This band confirms diastereoselective interactions between **2** and cytc in solution. In contrast to earlier investigations,<sup>91,92</sup> this selection occurs in the absence of any tether between the foldamer and the protein. However, the preferred handedness is

opposite to that of **1** in the crystal structure. This change may be the result of the different conditions used (*e.g.* pH, zinc acetate) or structural differences between **1** and **2**. Another explanation may be that the strongest interactions in solution are not those that favoured crystal growth most, *i.e.* different binding modes in solution and in the solid state.



**Figure 7.** Circular dichroism spectra of the cyt c – **2** mixture after equilibration at 20 °C. The sample contained 90  $\mu$ M protein – foldamer complex in 25 mM sodium acetate, 25 mM sodium chloride at pH 5.4.

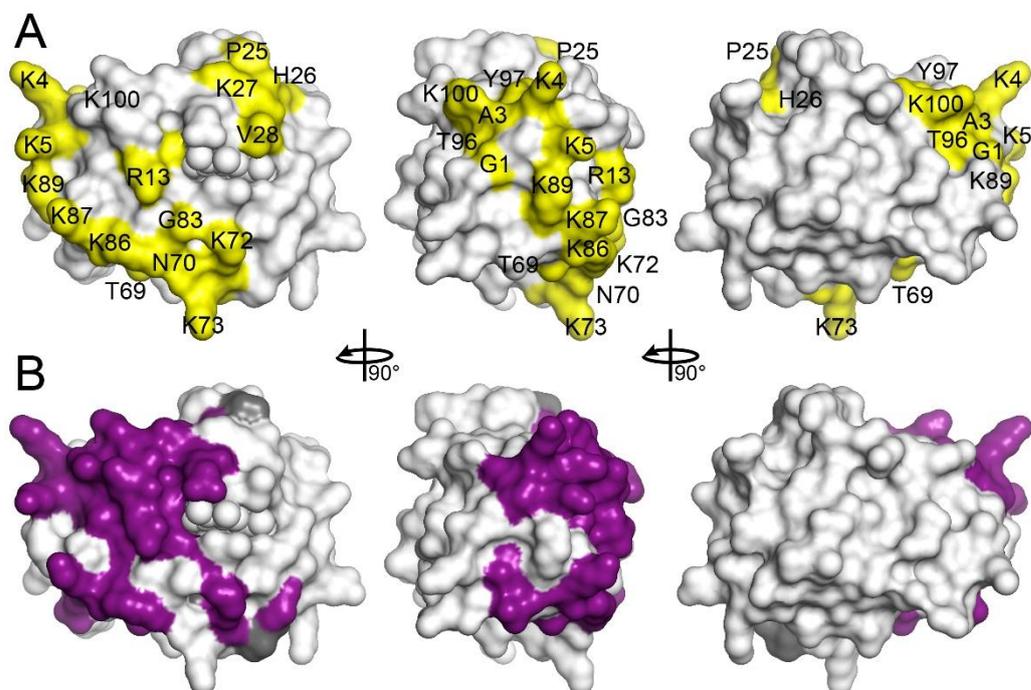
**NMR characterization and comparison with the solid-state.** NMR spectroscopy was used to gain further insight into cyt c recognition by foldamer **1** in solution. <sup>1</sup>H-<sup>15</sup>N HSQC-monitored titrations revealed increasing chemical shift perturbations ( $\Delta\delta$ ) and/or broadening as a function of the foldamer concentration, consistent with fast-to-intermediate exchange (Figure 8A).



**Figure 8. NMR characterization.** (A) Region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of pure cytochrome *c* (black contours) in the presence of **1** (coloured scale) at 30 °C. Samples contained 100  $\mu\text{M}$  protein and the indicated foldamer concentration in 20 mM potassium phosphate, 50 mM sodium chloride at pH 6.0. (B) Plot of chemical shift perturbations measured for cytochrome *c* backbone amides in the presence of 0.4 mM foldamer. Blanks correspond to prolines, unassigned or broadened resonances.

The binding site in solution was a contiguous patch of  $\sim 20$  residues, which was substantially different to the patches observed in the crystal structure. Lys4 and Lys5 are the only N-terminal residues that establish foldamer contacts in the crystal, whereas in solution the entire N-terminal  $\alpha$ -helix was affected (see  $\Delta\delta$  plot, Figures

8B and 9). For example, Ala3, Lys11, Thr12 and Arg13 do not bind **1** in the crystal structure but their amide resonances were shifted and broadened at 2 eq. foldamer.



**Figure 9.** (A) The crystallographic binding sites with residues coloured yellow are represented on cytc surface if at least one atom was  $\leq 4.5 \text{ \AA}$  from foldamer **1**, and (B) the NMR binding site with residues coloured purple  $\Delta\delta \text{ } ^1\text{H}^{\text{N}} \geq 0.04$  and/or  $^{15}\text{N} \geq 0.4$ , at 0.4 mM foldamer **1**. Prolines are dark grey.

Differences in the NMR and X-ray binding sites (Figure 9) are likely due to reorganization during crystallization, in favour of the site around Lys27, and facilitated by the zinc-mediated assembly (Figure 6). Broadening effects were observed also for Leu85 and Lys86. Most of the binding site identified by NMR was not involved in crystal contacts implying that significant reorganization occurs during crystallization. Surprisingly, the NMR-derived binding site for the foldamer is similar to a typical cytc – calix[4]arene complexes.<sup>39,43,44</sup> This observation suggests that the sulfonate – lysine interactions direct the complexation event (electrostatic steering). The lysine-rich patch of cytc is also prominent in binding protein partners in solution.<sup>158,304,305</sup>

**Cytochrome *c* – 1 as a bio-hybrid material.** The cytochrome *c* – 1 crystal structure is a mixture of two entities that differ strikingly in their chemical nature, size and shape. One is a highly cationic and globular biopolymer, the other is an anionic and helical synthetic ligand. The crystal composition of two foldamers (~4 kDa total) per protein (~13 kDa), results in a 1:3 mass ratio of synthetic to biological molecules. Charge-charge complementarity appears to facilitate the protein – foldamer interaction, highlighting the ‘molecular glue’ potential of foldamers to mediate new interfaces and assembly. In view of these features, it is suggested that cytochrome *c* – 1 complex may be considered as a composite bio-material.<sup>299</sup>

### Conclusions

Tether-free, charge-based complexation of a sulfonated foldamer with lysine-rich cytochrome *c* was demonstrated. In the crystal, a cylindrical stack of four foldamers interacted with eight proteins *via* small interfaces (100 - 180 Å<sup>2</sup>), consistent with weak interactions. Each end of the stack was packed against Pro25 and Lys27 from one protein and Lys5 from another. Interestingly, CD experiments with **2** identified the *P*-helix and NMR studies revealed a different binding patch suggesting an alternative binding mode in solution. The relatively compact and smooth surface of the foldamer precludes inter-digitation with the protein side chains. More extensive interactions with the protein surface may be achieved by a foldamer with alkyl (–R) or alkoxy (–OR; R = –CH<sub>3</sub> or –C<sub>2</sub>H<sub>5</sub>) linkers to the sulfonate substituents.

The property of **1** to mediate protein assembly by interacting with several surfaces hints at its application as a ‘molecular glue’.<sup>19,48,51-53,5,56,66</sup> Owing to the ease of functionalization, foldamers can be tailored according to the chemistry of the target protein<sup>100-102,295</sup> For example, the replacement of sulfonate with amino substituents<sup>92</sup> may facilitate recognition and assembly of anionic proteins. Finally, this study highlights the potential of quinoline oligoamides to generate protein-based composite materials, for applications in therapeutics, sensors and nanoarchitectures.<sup>100-102,293,295</sup>



## **Discussion**

---

Recent years have witnessed a rapid increase in supramolecules that can facilitate recognition and assembly of the desired protein (Table1, Introduction). Such complexation have led to the generation of interesting architectures (Figure 2, Introduction). The remarkable structures open up avenues to explore different ligand functionalities that will influence new modes of protein binding/packing. For example, **pclx<sub>6</sub>** facilitated a cytc dimerization<sup>51</sup> whereas **sclx<sub>6</sub>** yielded a porous honeycomb network<sup>56</sup> (~69 % solvent content). Similarly, in a rhodamine-sugar conjugate the length of the oligoethylene glycol linker determined the porosity of concanvalin A crystalline frameworks.<sup>151</sup>

When bound to a protein, these molecules form interfaces with two or more surrounding protein molecules. A range of non-covalent contacts established by the ligands at the binding site serves to stabilize the intermolecular interactions and drive crystal packing. Owing to a 'glue' like property at the interfaces these compounds are known as 'molecular or interface glues'. The change in the chemical environment of the protein accompanying a ligand-binding event is particularly advantageous where the neighbouring proteins experiences repulsive forces, and resist interaction.<sup>57</sup> In this respect, using a 'glue' may offer an alternate route to protein engineering (truncation, methylation) strategies. Additional features include stabilization of a disordered protein loop (**sclx<sub>6</sub>**,<sup>56</sup> **CB8**,<sup>59</sup> polyoxometallates<sup>65</sup>), auto-regulation of protein assembly (**sclx<sub>8</sub>**),<sup>19,20,145</sup> modulating of protein interaction with its protein partner (molecular tweezers)<sup>60,61</sup> and improving crystal quality.<sup>66,75</sup>

These molecules are available in varying sizes (200 - 2000 kDa), shapes (cone, donut, helical, linear, tweezers) and chemistries (hydrophobic, polar, neutral, charged; Figure 1, Introduction). Thus, a suitable 'glue' can be selected from a vast array of ligands (Table 1, Introduction). Numerous instances in the literature demonstrating the remarkable contribution of supramolecules towards protein recognition and assembly motivated this work. The crystal structures illustrated in

the thesis revealed interesting details of molecular recognition. This information can be harnessed for further application in designing biohybrid materials,<sup>24</sup> constructing porous architectures,<sup>56</sup> or understanding the biological relevance of a protein.<sup>55</sup>

### Sheet-like *versus* porous assembly (Chapter 1 and 2)

The effect of **sclx<sub>4</sub>** on protein recognition and assembly of cationic proteins is well established.<sup>48-50</sup> To determine the influence of different substituents on protein binding, two different calix[4]arenes were co-crystallized with cytc. These derivatives had the same lower (**pmclx<sub>4</sub>**)<sup>53</sup> or upper (**sclx<sub>4</sub>mc**) rim substitutions as **sclx<sub>4</sub>**. **pmclx<sub>4</sub>** had upper rim phosphonomethyl functionality (Chapter 1) whereas **sclx<sub>4</sub>mc** had oxomethylcarboxylate groups at the lower rim (Chapter 2). Cyt<sub>c</sub>-complexes with the three ligands highlight the effect of substituents on crystallization, recognition and assembly.

In crystallization trials, compared to 10 equiv. **sclx<sub>4</sub>**,<sup>48</sup> <1 equiv. **pmclx<sub>4</sub>**<sup>53</sup> and 1 equiv. **sclx<sub>4</sub>mc** was sufficient to achieve crystal growth of the cytc complex. The crystal structures suggested that this difference in the required ligand concentration may be owing to the enhanced Lys selectivity. **sclx<sub>4</sub>** bound to Lys4, Lys22 and Lys89 whereas **pmclx<sub>4</sub>** and **sclx<sub>4</sub>mc** exhibited selectivity towards Lys86 or Lys87, respectively. This increased selectivity of **pmclx<sub>4</sub>** was also evident from the 10-fold stronger *apparent*  $K_d$  (NMR studies) compared to the cytc – **sclx<sub>4</sub>** complex.<sup>48,53</sup> The chemistry offered by the flexible phosphonomethyl substituents (hydrophobicity, hydrogen bond acceptor/donor) may have been instrumental in this aspect.

Crystal packing also highlighted the contribution of these functionalities. **sclx<sub>4</sub>** and **pmclx<sub>4</sub>** yielded sheets or layered protein – calixarene units,<sup>48,53</sup> whereas **sclx<sub>4</sub>mc** induced a highly porous honeycomb like assembly (~73 % solvent content). Similarly, **sclx<sub>6</sub>** mediated a honeycomb arrangement of cytc (*P3<sub>2</sub>21*; ~65 % solvent content).<sup>56</sup> In contrast, **sclx<sub>8</sub>** dictated a high-porosity crystalline framework with ~85 %

porosity.<sup>19,20</sup> The contrast with calix[4]arenes in this regard but similarity with the **sclx<sub>6</sub>** and **sclx<sub>8</sub>** may suggest a role for high charge of the ligands. Additionally, the co-crystallization pH of **sclx<sub>4mc</sub>** (pH 5.0), **sclx<sub>6</sub>** (pH 8.0), and **sclx<sub>8</sub>** (pH 7.0) may have aided in this direction. Under these condition, all the phenolic OH of **sclx<sub>4mc</sub>** are likely to be deprotonated (charge = -8), whereas two phenolic OH of **sclx<sub>6</sub>** and **sclx<sub>8</sub>** will be deprotonated resulting in a charge of -8 and -10, respectively. Probably, the crystal porosity is facilitated by the charge-charge interactions enabled by these polyanionic calixarene at the protein interfaces.

### **Sclx<sub>8</sub> recognition – selection between cytc or porphyrin (Chapter 3)**

**sclx<sub>8</sub>** was observed to successfully co-crystallize with cytc under various crystallization condition to yield three different assemblies (space groups – *H3*, *P3<sub>1</sub>*, *P4<sub>3</sub>2<sub>1</sub>2*).<sup>14</sup> In these structures, the inherent flexibility enables **sclx<sub>8</sub>** to mould to the protein surface and form large interfaces (~550 Å<sup>2</sup>). To investigate the binding of **sclx<sub>8</sub>** with two cationic partners, we co-crystallized this receptor with cytc and a cationic trimethylanilinium porphyrin (**tmap**). As only calix[4]arene and porphyrin complexations have been explored, this data might aid in understanding how the larger calixarene behaves with porphyrin.

A 1.0 Å resolution revealed an alternating assembly of four **sclx<sub>8</sub>** and six **tmap** in a supramolecular stack. The macrocycles adopted conformations that would enable them to stabilize the stack. For example, the **sclx<sub>8</sub>** conformation comprised two ‘calix[3]arene’ cavities that accommodated a trimethylanilinium group or a solvent molecule. Porphyrins were present in two arrangements – planar or non-planar – depending on their function within the stack. The non-planar **tmap** were involved in forming the stack whereas the planar ones bound to the stack peripherally. These peripheral **tmap** molecules behave as a ‘bridge’ or a ‘connector’

that links the neighbouring stacks, and facilitates a **sclx<sub>8</sub>** – **tmap** supramolecular architecture.

In contrast, the calix[4]arene – porphyrin crystal structures revealed a stack comprising of porphyrin to which the calixarenes bound laterally.<sup>210-212,237-241</sup> With respect to this arrangement, it is suggestive that in the present structure ‘floppiness’ of **sclx<sub>8</sub>** enables it to adopt a conformation that favours its participation within the stack. Interestingly, this complex exemplifies a structure with high  $Z' = 22$  and may be useful in understanding the factors that govern the generation of crystal structures with a  $Z' \geq 16$ .<sup>256-258</sup> Finally, the details of molecular recognition may be helpful for generating hybrid structures involving larger calixarenes and porphyrins

#### **Sclx<sub>n</sub> and two cationic proteins (Chapter 4)**

The propensity of **sclx<sub>n</sub>** to facilitate crystallization by acting as ‘molecular glues’ was investigated in the *Penicillium* antifungal (PAF)<sup>166</sup> protein for which the crystal structure was not previously elucidated. While the pure protein failed to yield diffraction-quality crystals, co-crystallization with **sclx<sub>4</sub>**, **sclx<sub>6</sub>**, and **sclx<sub>8</sub>** resulted in high-resolution crystal structures (~1.5 Å resolution). Interestingly, the PAF – **sclx<sub>n</sub>** crystal structures enables a direct comparison with cytc – **sclx<sub>n</sub>** complexes in respect to protein recognition and assembly.<sup>19,48,56</sup>

**Sclx<sub>4</sub>** and **sclx<sub>6</sub>** bound a PAF monomer whereas **sclx<sub>8</sub>** facilitated a dimer.<sup>46</sup> In complex with cytc, **sclx<sub>4</sub>** facilitated a crystallographic dimer,<sup>48</sup> **sclx<sub>6</sub>** a monomer<sup>56</sup> and **sclx<sub>8</sub>** a monomer or a tetramer.<sup>19</sup> While all three calixarenes bound a Pro29-Lys30-Phe31 motif (loop 3) on PAF, these ligands recognized different binding sites on cytc. The selectivity of calixarenes for this site may be attributed to a stable platform provided by Pro29. The hydrophobic stacking interaction between pyrrolidine and phenolsulfonate ring may have stabilized the binding, especially of the ‘floppy’ **sclx<sub>6</sub>** and **sclx<sub>8</sub>**. Perhaps, the surface offered by the ~13 kDa cytc (compared to 6.2 kDa PAF) enables the ligand to explore a larger area for favourable protein interactions.

Considering crystal packing in these complexes interesting observations were made. In contrast to the porous cytc assemblies induced by **sclx<sub>6</sub>**<sup>56</sup> and **sclx<sub>8</sub>**,<sup>19</sup> the **sclx<sub>n</sub>** series yielded tightly packed sheet-assembly of PAF which resembled that of cytc – **sclx<sub>4</sub>**. Interestingly, the solution-state studies of calixarene complexes with cytc (**sclx<sub>4</sub>** / **sclx<sub>8</sub>**) and PAF (**sclx<sub>4</sub>** / **sclx<sub>6</sub>** / **sclx<sub>8</sub>**) illustrated pronounced effects on recognition and assembly as the ring size and charge increases. For example, **sclx<sub>8</sub>** caused spectral broadening of resonances in both proteins but not with **sclx<sub>4</sub>** (cytc and PAF) or **sclx<sub>6</sub>** (PAF only). However, **pclx<sub>6</sub>** was observed to cause severe spectral broadening of cytc resonances and was in agreement with dimerization revealed in crystal structure.

Finally, the recognition of PAF by the anionic calixarenes, especially the selectivity for the Pro29-Lys30-Phe31 motif may have implications for biological relevance. Previously, X-ray crystal structures have revealed how defensin-phospholipid binding leads to oligomerization, suggesting a mechanism for membrane permeation.<sup>268-271</sup> In view of this, perhaps this site on the antifungal protein is involved in cell membrane binding and the subsequent membrane permeation.

### **A sulfonated quinoline foldamer – bound, not tethered (Chapter 5)**

Foldamers are attractive candidates for protein recognition owing to their ability to adopt stably folded conformations.<sup>94-96</sup> Among the aromatic foldamers, oligoamides of 8-amino-2-quinoline carboxylic acid have become important for investigating protein recognition.<sup>100-102</sup> Previously, these ligands relied on a tether functionality to confine the foldamer to a specific region on the protein surface.<sup>101,102,295</sup> With an aim to explore the recognition properties of a tether-free anionic sulfonated foldamer, we co-crystallized it with cytc. The crystal structure revealed a ~5 nm stack of two foldamer bound to one cytc molecule. While one foldamer was well-ordered, the other foldamer was modelled in opposite orientations (each 50 % occupancy). In the well-defined foldamer seven of the eight sulfonates are salt bridged to five Lys side

chains in four surrounding proteins (100-180 Å<sup>2</sup>). The N-terminal quinoline interacts with Lys5 from one chain in combination with Pro25 and Lys27 from a neighbouring protein. These contacts among others may have served to stabilize this foldamer in a single orientation and thus, favoured the selection of *M*-helix. In contrast, the poorly ordered foldamer formed interfaces with only two proteins (165 or 190 Å<sup>2</sup>) which may have been insufficient to dictate its orientation.

Considering the surface area (~2kDa; ~1300 Å<sup>2</sup>), the interfaces mediated by the foldamer are rather small, and comparable to calix[4]arenes (~0.8 kDa). These ligands bear a cavity that can encapsulate individual Lys residues leading to larger interface areas. In contrast to this oligoamide, **sclx**<sub>8</sub> (~1.5 kDa)<sup>19</sup> or **suramin** (~1.3 kDa)<sup>118</sup> accounted for protein surface coverage of up to 550 Å<sup>2</sup> and 750 Å<sup>2</sup>, respectively. Owing to the inherent conformational flexibility these ligands mould to the protein surface, thus forming large interfaces. This ability exemplifies them as effective 'molecular glue'. Probably, the compact and the rigid nature of the helical foldamer diminished this property to some extent.

This foldamer lacks stereogenic centre and was synthesized as a racemic mixture of *P*- and *M*-conformers. Also, the helix handedness inversion is slow for such quinoline-based octamers. Under these circumstances, the exclusive occurrence of the *M*-conformer is suggestive of spontaneous chiral resolution. Furthermore, due to these structural and chemical features, they are not amenable for circular dichroism (CD) experiments. Therefore, a flexible analogue containing aminomethyl-pyridine units was employed. A positive band was obtained in the quinoline absorption region on equilibration of cytc with this foldamer indicating *P*-helicity. This contrast from the crystal structure might suggest a different binding mode in solution. Similarly, solution-state NMR studies also exhibited variation from the solid-state. For example, the binding site in solution was a contiguous patch of ~20 residues, which is significantly different to the patches observed in the crystal structure.

Interestingly, this binding site was similar to a typical cytc – calix[4]arene complex highlighting a role for sulfonate – Lys interaction in complexation.

The cytc – foldamer complex can be considered as a biohybrid material and illustrates the potential of quinoline-based foldamers to fabricate such assemblies.<sup>301</sup> This complex is composed of two chemically and structurally different components – a cationic, globular biopolymer and an anionic, helical synthetic ligand. The crystal composition of two foldamers (~4 kDa total) per protein (~13 kDa), results in a 1:3 mass ratio of synthetic to biological molecules. Charge-charge complementarity appears to facilitate the protein – foldamer interaction highlighting the ‘molecular glue’ property of the foldamers to direct assembly. With respect to the molecular recognition details provided by the complex, foldamer with alkyl (-R) or alkoxy (-OR; R = -CH<sub>3</sub> or -C<sub>2</sub>H<sub>5</sub>) linkers to the sulfonate substituents can be investigated for enhanced interactions. Also, the replacement of sulfonate with amino substituents<sup>296</sup> may facilitate recognition and assembly of anionic proteins.

### **Suitability of a ‘glue’**

Despite the plethora of examples in literature, it is difficult to predict which ligand will be suitable for crystallizing a desired protein. The representative examples (Table 1, Introduction) may aid in drawing some inferences. For instance, polyoxometallates can bind anionic / cationic / neutral proteins. Crystallophores particularly recognize glutamates / aspartates on anionic proteins. Anionic calixarenes, suramin or tweezers are likely to bind proteins with accessible lysines or arginines. Various dyes have also been successful in facilitating protein assembly including orange G, tetrasulfonato-phenyl porphyrin, pyrene tetrasulfonate. However, attempts to co-crystallize these compounds with cytc were unsuccessful (Appendix). In contrast, the preliminary NMR solution-state studies suggested their tendency (in some cases to a small extent) for cytc recognition (Table A2 and Figures A2 – A6; Appendix).

### What lies ahead?

Previously, the ability of **sclx<sub>8</sub>** to auto-regulate cytc assembly was demonstrated.<sup>19</sup> While at ~1 equiv. **sclx<sub>8</sub>** formed cytc tetramer, at >3 equiv. the tetramer disassembled to form yield cytc monomers. Interestingly, this latter higher equiv. was able to induce a highly porous crystalline framework in which all the contacts were mediated by **sclx<sub>8</sub>**. Recently, with the same complex it was illustrated how an effector molecule, in this case a polycationic spermine, was able to modulate porous protein architecture.<sup>20</sup> Such frameworks have immense scope for generating biomaterials with applications in catalysis and drug delivery systems.<sup>2,13-15,203-205</sup> Furthermore, protein assembly is imperative for enabling the protein/enzyme to exhibit its action. For example, polyphosphate-mediated dimers,<sup>137</sup> trimers,<sup>138</sup> and hexamers<sup>134</sup> of different proteins appear to be important in regulating enzyme activity.

Though the ability to have control over protein assembly in semi-synthetic systems is in stages of infancy, it is nevertheless a (small) step towards achieving the level of complexity that is designed by Nature. In this direction, it would be interesting to apply the knowledge gleaned from the crystal structures reported in the thesis. We have demonstrated / highlighted the usefulness of calixarenes in fabricating porous architectures. Additionally, the PAF - calixarene complexes illustrate the potential of these synthetic receptors for crystallizing recalcitrant proteins. Also, we have revealed the bio-hybrid assembly generated by the tether-free foldamer.

In this thesis, protein – ligand complexes were elucidated crystallographically, and the details of molecular recognition obtained were confirmed by solution-state NMR studies. Several other characterizations techniques are also available including TEM, AFM, ITC, SEC-MALS, SAXS, SPR or UV.<sup>2,5,7,19,20,25,26,145</sup> These approaches are commonly used to investigate kinetics, molecular weight, oligomerization state or the dynamic nature of a complex. Their utilization at the crystallization stage though challenging may help in understanding the role of different components (ligand,

buffer or salt) that contribute towards mediating assembly. For example, CD and fluorescence studies were used to demonstrate that initially rhodamine-sugar conjugate bound ConA, and dimerization of the ligand occurred as a second step.<sup>151</sup> Also, application of these approaches to a protein-ligand co-crystal/drop may help provide clues regarding the fitness of a 'glue' for a target protein.



## **Bibliography**

---

1. Alberts, B. The cell as a collection overview of protein machines: preparing the next generation of molecular biologists. *Cell*, **1998**, *92*, 291–294.
2. Kuan, S. L.; Bergamini, F. R. G.; Weil, T. Functional protein nanostructures: A chemical toolbox. *Chem. Soc. Rev.* **2018**, *47*, 9069–9105.
3. Jain, A.; Liu, R.; Ramani, B.; Arauz, E.; Ishitsuka, Y.; Ragnathan, K.; Park, J.; Chen, J.; Xiang, Y. K.; Ha, Taekjip. Probing cellular protein complexes using single-molecule pull-down, *Nature* **2011**, *473*, 484–488.
4. van Dun, S.; Ottmann, C.; Milroy, L. G.; L. Brunsveld. Supramolecular chemistry targeting proteins. *J. Am. Chem. Soc.* **2017**, *139*, 13960–13968.
5. Pieters, B. J. G. E.; Eldijk, M. B. van; Noltea, R. J. M.; Mecinović, J. Natural supramolecular protein assemblies. *Chem. Soc. Rev.* **2016**, *45*, 24–39.
6. Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. Protein assembly: versatile approaches to construct highly ordered nanostructures. *Chem. Rev.* **2016**, *116*, 13571–13632.
7. Marsh, J. A.; Teichmann, S. A. Structure, dynamics, assembly, and evolution of protein complexes. *Annu. Rev. Biochem.* **2015**, *84*, 5.1–5.25.
8. Okesola B. O.; Mata, A. Multicomponent self-assembly as a tool to harness new properties from peptides and proteins in material design. *Chem. Soc. Rev.* **2018**, *47*, 3721–3736.
9. Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein–protein interactions: what are the preferred ways for proteins to interact? *Chem. Rev.* **2008**, *108*, 1225–1244.
10. Abbas, M.; Zou, Q.; Li, S.; Yan, X. Self-assembled peptide- and protein-based nanomaterials for antitumor photodynamic and photothermal therapy. *Adv. Mater.* **2017**, *29*, e1605021.
11. Bhaskar, S.; Lim, S. Engineering protein nanocages as carriers for biomedical applications. *NPG Asia Mater.* **2017**, *9*, e371.

12. Spicer, C. D.; Jumeaux, C.; Gupta, B.; Stevens, M. M. Peptide and protein nanoparticle conjugates: versatile platforms for biomedical applications. *Chem. Soc. Rev.* **2018**, *47*, 3574–3620.
13. Yeates, T. O. Geometric principles for designing highly symmetric self-assembling protein nanomaterials. *Annu. Rev. Biophys.* **2017**, *46*, 23–42.
14. Kawakami, N.; Kondo, H.; Matsuzawa, Y.; Hayasaka, K.; Nasu, E.; Sasahara, K.; Arai, R.; Miyamoto, K. Design of hollow protein nanoparticles with modifiable interior and exterior surfaces. *Angew. Chem. Int. Ed.* **2018**, *57*, 12400–12404.
15. Simon, A. J.; Zhou, Y.; Ramasubramani, V.; Glaser, J.; Pothukuchy, A.; Gollihar, J.; Gerberich, J. C.; Leggere, J. C.; Morrow, B. R.; Jung, C.; Glotzer, S. C.; Taylor, D. W.; Ellington, A. D. Supercharging enables organized assembly of synthetic biomolecules. *Nat. Chem.* **2019**, *11*, 204–212.
16. Hernández, N. E.; Hansen, W. A.; Zhu, D.; Shea, M. E.; Khalid, M.; Manichev, V.; Putnins, M.; Chen, M.; Dodge, A. G.; Yang, L.; Marrero-Berríos, I.; Banal, M.; Rechani, P.; Gustafsson, T.; Feldman, L. C.; Lee, S.-H.; Wackett, L. P.; Dai, W.; Khare, S. D. Stimulus-responsive self-assembly of protein-based fractals by computational design. *Nat. Chem.* **2019**, *11*, 605–614.
17. Malay, A. D.; Miyazaki, N.; Biela, A.; Chakraborti, S.; Majsterkiewicz, K.; Stupka, I.; Kaplan, C. S.; Kowalczyk, A.; Piette, B. M. A. G.; Hochberg, G. K.A.; Wu, D.; Wrobel, T. P.; Fineberg, A.; Kushwah, M. S.; Kelemen, M.; Vavpetič, P.; Pelicon, P.; Kukura, P.; Benesch, J. L. P.; Iwasaki, K.; Heddle, J. G. An ultra-stable gold-coordinated protein cage displaying reversible assembly. *Nature* **2019**, *569*, 438–442.
18. Ross, J. F.; Wildsmith, G. C.; Johnson, M.; Hurdiss, D. L.; Hollingsworth, K.; Thompson, R. F.; Mosayebi, M.; Trinh, C. H.; Paci, E.; Pearson, A. R.; Webb, M. E.; Turnbull, W. B. Directed assembly of homopentameric cholera toxin b-subunit proteins into higher-order structures using coiled-coil appendages. *J. Am. Chem. Soc.* **2019**, *141*, 5211–5219.

19. Rennie, M. L.; Fox, G. C.; Pérez J.; Crowley, P. B. Auto-regulated protein assembly on a supramolecular scaffold. *Angew. Chem. Int. Ed.* **2018**, *57*, 13764–13769.
20. Engilberge, S.; Rennie, M. L.; Dumont, E.; Crowley, P. B. Tuning protein frameworks *via* auxiliary supramolecular interactions. *ACS Nano*, **2019**, *13*, 10343–10350.
21. Guagnini, F.; Antonik, P. M.; Rennie, M. L.; O'Byrne, P.; Khan, A. R.; Pinalli, R.; Dalcanale, E.; Crowley, P. B. Cucurbit[7]uril-dimethyllysine recognition in a model protein. *Angew. Chem. Int. Ed.* **2018**, *57*, 7126–7130.
22. Beyeh, N. K.; Nonappa, Liljeström, V.; Mikkilä, J.; Korp, A.; Bochicchio, D.; Pavan, G. M.; Ikkala, O.; Ras, R. H. A.; Kostianen, M. A. Crystalline cyclophane–protein cage frameworks. *ACS Nano* **2018**, *12*, 8029–8036.
23. Heid, C.; Sowislok, A.; Schaller, T.; Niemeyer, F.; Klärner F.-G.; Schrader, T. Molecular tweezers with additional recognition sites. *Chem. Eur. J.* **2018**, *24*, 11332–11343.
24. Alex, J. M.; Corvaglia, V.; Hu, X.; Engilberge, S.; Huc, I.; Crowley, P. B. Crystal structure of a protein–aromatic foldamer composite: macromolecular chiral resolution. *Chem. Commun.* **2019**, *55*, 11087–11090.
25. Garcia-Seisdedos, H.; Empereur-Mot, C.; Elad, N.; Levy, E. D. Proteins evolve on the edge of supramolecular self-assembly. *Nature* **2017**, *548*, 244–247.
26. Schneider, M.; Belsom, A.; Rappsilber, J. Protein tertiary structure by crosslinking/ mass spectrometry, *Trends Biochem. Sci.* **2018**, *43*, 157–169.
27. Maita, N. Crystal structure determination of ubiquitin by fusion to a protein that forms a highly porous crystal lattice. *J. Am. Chem. Soc.* **2018**, *140*, 13546–13549.
28. Garman, E. F. Developments in X-ray crystallographic structure determination of biological macromolecules. *Science* **2014**, *343*, 1102–1108.
29. Kwon, B.; Lee, M.; Waring, A. J.; Hong, M. Oligomeric structure and three-dimensional fold of the HIV gp41 membrane-proximal external region and

- transmembrane domain in phospholipid bilayers. *J. Am. Chem. Soc.* **2018**, *140*, 8246–8259.
30. Huang, C.; Kalodimos, C. G. Structures of large protein complexes determined by nuclear magnetic resonance spectroscopy. *Annu. Rev. Biophys.* **2017**, *46*, 317–336.
31. Cheng, Y.; Glaeser, R. M.; Nogales E. How Cryo-EM became so hot. *Cell* **2017**, *171*, 1229–1231.
32. Schubert, R.; Meyer, A.; Baitan, D.; Dierks, K.; Perbandt, M.; Betzel, C. Real-time observation of protein dense liquid cluster evolution during nucleation in protein crystallization. *Cryst. Growth Des.* **2017**, *17*, 954–958.
33. Holcomb, J.; Spellmon, N.; Zhang, Y.; Doughan, M.; Li, C.; Yang, Z. Protein crystallization: eluding the bottleneck of X-ray crystallography. *AIMS Biophys.* **2017**, *4*, 557–575.
34. Derewenda, Z. S.; Godzik, A. The “sticky patch” model of crystallization and modification of proteins for enhanced crystallizability. *Methods Mol. Biol.* **2017**, *1607*, 77–115.
35. Nowotny, P.; Hermann, J.; Li, J.; Krautenbacher, A.; Klöpfer, K.; Hekmat, D.; Weuster-Botz, D. Rational crystal contact engineering of *Lactobacillus brevis* alcohol dehydrogenase to promote technical protein crystallization. *Cryst. Growth Des.* **2019**, *19*, 2380–2387.
36. Zou, Y.; Weis, W. I.; Kobilka, B. K. N-Terminal T4 Lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PLoS ONE*, **2012**, *7*, e0046039.
37. McIlwain, B. C.; Newstead, S.; Stockbridge, R. B. Cork-in-bottle occlusion of fluoride ion channels by crystallization chaperones. *Structure* **2018**, *26*, 635–639.
38. Ng, J. T.; Dekker, C.; Reardon, P.; von Delft, F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2016**, *D72*, 224–235. Lessons from ten years of crystallization experiments at the SGC.
39. Deller, M. C.; Rupp, B. Approaches to automated protein crystal harvesting. *Acta Crystallogr. F, Struct. Biol. Commun.* **2014**, *70*, 133–155.

40. Sigdel, M.; Pusey, M. L.; Aygun, R. S. CrystPro: Spatiotemporal analysis of protein crystallization images. *Cryst. Growth Des.* **2015**, *15*, 5254–5262.
41. Lin, Y. What's Happened over the last five years with high-throughput protein crystallization screening? *Expert Opin. Drug Discov.* **2018**, *13*, 691–695.
42. Jancarik, J.; Kim, S.-H. Sparse matrix sampling: A Screening method for crystallization of proteins. *J. Appl. Cryst.* **1991**, *24*, 409–411.
43. D'Arcy, A.; Bergfors, T.; Jacob, S.W. C.; Marsh, M. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr. F, Struct. Biol. Commun.* **2014**, *70*, 1117–1126.
44. Gorrec, F. Protein crystallization screens developed at the MRC laboratory of molecular biology. *Drug Discov. Today* **2016**, *21*, 819–825.
45. McPherson, A.; Nguyen, C.; Cudney, R.; Larson, S. B. The role of small molecule additives and chemical modification in protein crystallization. *Cryst. Growth Des.* **2011**, *11*, 1469–1474.
46. Fischer, E. S.; Park, E.; M. J. Eck.; Thomä, N. H. SPLINTS: Small-molecule protein ligand interface stabilizers. *Curr. Opin. Struct. Biol.* **2016**, *37*, 115–122.
47. McPherson, A.; Cudney, B. Searching for silver bullets: An alternative strategy for crystallizing macromolecules. *J. Struct. Biol.* **2006**, *156*, 387–406.
48. McGovern, R. E., Fernandes, H., Khan, A. R., Power, N. P., Crowley, P. B. Protein camouflage in cytochrome *c*-calixarene complexes. *Nat. Chem.* **2012**, *4*, 527–533.
49. McGovern, R. E., McCarthy, A. A., Crowley, P. B. Protein assembly mediated by sulfonato-calix[4]arene. *Chem. Commun.* **2014**, *50*, 10412–10415.
50. McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. Structural study of a small molecule receptor bound to dimethyllysine in lysozyme. *Chem. Sci.* **2015**, *6*, 442–449.
51. Rennie, M. L.; Doolan, A. M.; Raston, C. L.; Crowley, P. B. Protein dimerization on a phosphonated calix[6]arene disc. *Angew. Chem. Int. Ed.* **2017**, *56*, 5517–5521.

52. Doolan, A. M.; Rennie, M. L.; Crowley, P. B. Protein recognition by functionalized sulfonato-calix[4]arenes. *Chem. Eur. J.* **2018**, *24*, 984–991.
53. Alex, J. M.; Rennie, M. L.; Volpi, S.; Sansone, F.; Casnati, A.; Crowley, P. B. Phosphonated calixarene as a “molecular glue” for protein crystallization. *Cryst. Growth Des.* **2018**, *18*, 2467–2473.
54. Mummidivarapu, V. V. S.; Rennie, M. L.; Doolan, A. M.; Crowley, P. B. Noncovalent pegylation *via* sulfonatocalix[4]arene—A crystallographic proof. *Bioconjugate Chem.* **2018**, *29*, 3999–4003.
55. Alex, J. M.; Rennie, M. L.; Engilberge, S.; Lehoczki, G.; Dorotyya, H.; Fizil, Á.; Batta, G.; Crowley, P. B. Calixarene-mediated assembly of a small antifungal protein. *IUCrJ* **2019**, *6*, 238–247.
56. Engilberge, S.; Rennie, M. L.; Crowley, P. B. Calixarene-capture of partially unfolded cytochrome *c*. *FEBS Letters* **2019**, *593*, 2112–2117.
57. Alex, J. M.; Brancatelli, G.; Volpi, S.; Bonaccorso, C.; Casnati, A.; Geremia, S.; Crowley, P. B. Porous protein architectures *via* co-crystallization with an octa-anionic bolaamphiphile calix[4]arene. *Org. Biomol. Chem.* 2019. (*Accepted*).
58. Chinai, J. M.; Taylor, A. B.; Ryno, L. M.; Hargreaves, N. D.; Morris, C. A.; Hart, P. J.; Urbach, A. R. Molecular recognition of insulin by a synthetic receptor. *J. Am. Chem. Soc.* **2011**, *133*, 8810–8813.
59. de Vink, P. J.; Briels, J. M.; Schrader, T.; Milroy, L.-G.; Brunsveld, L.; Ottmann, C. A binary bivalent supramolecular assembly platform based on cucurbit[8]uril and dimeric adapter protein 14-3-3. *Angew. Chem. Int. Ed.* **2017**, *56*, 8998–9002.
60. Bier, D.; Rose, R.; Bravo-Rodriguez, K.; Bartel, M.; Ramirez-Anguita, J. M.; Dutt, S.; Wilch, C.; Klärner, F.-G.; Sanchez-Garcia, E.; Schrader, T.; Ottmann, C. Molecular tweezers modulate 14–3–3 protein-protein interactions. *Nat. Chem.* **2013**, *5*, 234–239.
61. Bier, D.; Mittal, S.; Bravo-Rodriguez, K.; Sowislok, A.; Guillory, X.; Briels, J.; Heid, C.; Bartel, M.; Wettig, B.; Brunsveld, L.; Sanchez-Garcia, E.; Schrader, T.;

- Ottmann, C. The molecular tweezer CLR01 stabilizes a disordered protein–protein interface. *J. Am. Chem. Soc.* **2017**, *139*, 16256–16263.
62. Fusco, D.; Headd, J. J.; De Simone, A.; Wang, J.; Charbonneau, P. Characterizing protein crystal contacts and their role in crystallization: rubredoxin as a case study. *Soft Matter* **2014**, *10*, 290–302.
63. James, S.; Quinna, M. K.; McManus, J. J. The self-assembly of proteins: probing patchy protein interactions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 5413–5420.
64. Staneva, I.; Frenkel, D. The role of non-specific interactions in a patchy model of protein crystallization. *J. Chem. Phys.* **2015**, *143*, 194511–194516.
65. Bijelic, A.; Rompel, A. Ten good reasons for the use of the tellurium-centered anderson-evans polyoxotungstate in protein crystallography. *Acc. Chem. Res.* **2017**, *50*, 1441–1448.
66. Engilberge, S.; Riobé, F.; Wagner, T.; Pietro, S. D.; Breyton, C.; Franzetti, B.; Shima, S.; Girard, E.; Dumont, E.; Maury, O. Unveiling the binding modes of the crystallophore, a terbium-based nucleating and phasing molecular agent for protein crystallography. *Chem. Eur. J.* **2018**, *24*, 9739–9746.
67. Gutsche, C. D.; Bauer, L. J. Calixarenes. 13. The conformational properties of calix[4]arenes, calix[6]arenes, calix[8]arenes, and oxacalixarenes. *J. Am. Chem. Soc.* **1985**, *107*, 6052–6059.
68. Shetty, D.; Skorjanc, T.; Raya, J.; Sharma, S. K.; Jahovic, I.; Polychronopoulou, K.; Asfari, Z.; Han, D. S.; Dewage, S.; Olsen, J.-C.; Jagannathan, R.; Kirmizialtin, S.; Trabolsi, A. Calix[4]arene-based porous organic nanosheets. *ACS Appl. Mater. Interfaces* **2018**, *23*, 17359–17365.
69. Wang, S.; Gao, X.; Hang, X.; Zhu, X.; Han, H.; Li, X.; Liao, W.; Chen, W. Calixarene-based {Ni<sub>18</sub>} coordination wheel: Highly efficient electrocatalyst for the glucose oxidation and template for the homogenous cluster fabrication. *J. Am. Chem. Soc.* **2018**, *140*, 6271–6277.
70. Lazar, A.; Da Silva, E.; Navaza, A.; Barbey, C.; Coleman, A. W. A new packing motif for *p*-sulfonatocalix[4]arene: the solid state structure of the para-

- sulfonatocalix[4]arene D-arginine complex. *Chem. Commun.* **2004**, *19*, 2162–2163.
71. Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. A simple calixarene recognizes post-translationally methylated lysine. *ChemBioChem.* **2010**, *11*, 63–66.
72. Mareeswaran, P. M.; Prakash, M.; Subramanian, V.; Rajagopal, S. Recognition of aromatic amino acids and proteins with *p*-sulfonatocalix[4]arene – a luminescence and theoretical approach. *J. Phys. Org. Chem.* **2012**, *25*, 1217–1227.
73. Guo, D.-S.; Y. Liu, Supramolecular chemistry of *p*-sulfonatocalix[*n*]arenes and its biological applications. *Acc. Chem. Res.* **2014**, *47*, 1925–1934.
74. Kumar, R.; Sharma, A.; Singh, H.; Suating, P.; Kim, H. S.; Sunwoo, K.; Shim, I.; Gibb, B. C.; Kim, J. S. Revisiting fluorescent calixarenes: from molecular sensors to smart materials. *Chem. Rev.* **2019**, *119*, 9657–9721.
75. Lee, C. C.; Maestre-Reyna, M.; Hsu, K. C.; Wang, H. C.; Liu, C. I.; Jeng, W. Y.; Lin, L. L.; Wood, R.; Chou, C. C.; Yang, J. M.; Wang, A. H. Crowning proteins: modulating the protein surface properties using crown ethers. *Angew. Chem. Int. Ed.* **2014**, *53*, 13054–13058.
76. Chan, Y.-T.; Ko, T.-P.; Yao, S.-H.; Chen, Y.-W.; Lee, C.-C.; Wang, A. H.-J. Crystal structure and potential head-to-middle condensation function of a Z, Z-farnesyl diphosphate synthase. *ACS Omega* **2017**, *2*, 930–936.
77. Chen, K.-E.; Tillu, V. A.; Chandra, M.; Collins, B. M. Molecular basis for membrane recruitment by the PX and C2 domains of class II phosphoinositide 3-Kinase-C2 $\alpha$ . *Structure* **2018**, *26*, 1612–1625.
78. Maestre-Reyna, M.; Wu, W.-J.; Wang, A. H.-J.; Structural insights into RbmA, a biofilm scaffolding protein of *V. Cholerae*. *PLoS One* **2013**, *8*, e82458.
79. Rosnizeck, I. C.; Graf, T.; Spoerner, M.; Tränkle, J.; Filchtinski, D.; Herrmann, C.; Gremer, L.; Vetter, I. R.; Wittinghofer, A.; König, B.; Kalbitzer, H. R. Stabilizing a weak binding state for effectors in the human Ras protein by cyclen complexes. *Angew. Chem. Int. Ed.* **2010**, *49*, 3830–3833.

- 
80. Gerlach, L. O.; Skerlj, R. T.; Bridger, G. J.; Schwartz, T. W. Molecular interactions of cyclam and bicyclam non-peptide antagonists with the CXCR4 chemokine receptor. *J. Biol. Chem.* **2001**, *276*, 14153–14160.
81. Khan, N. S.; Riggle, B. A.; Seward, G. K.; Bai, Y.; Dmochowski, I. J. Cryptophane-folate biosensor for  $^{129}\text{Xe}$  NMR. *Bioconjugate chem.* **2015**, *26*, 101–109.
82. Aaron, J. A.; Chambers, J. M.; Jude, K. M.; Di Costanzo, L.; Dmochowski, I. J.; Christianson, D. W. Structure of a  $^{129}\text{Xe}$ -cryptophane biosensor complexed with human carbonic anhydrase II. *J. Am. Chem. Soc.* **2008**, *130*, 6942–6943.
83. Freeman, W. A.; Mock, W. L.; Shih, N. Y. Cucurbituril. *J. Am. Chem. Soc.* **1981**, *103*, 7367–7368.
84. Uhlenheuer, D. A.; Young, J. F.; Nguyen, H. D.; Scheepstra, M.; Brunsveld, L. Cucurbit[8]uril induced heterodimerization of methylviologen and naphthalene functionalized proteins. *Chem. Commun.* **2011**, *47*, 6798–6800.
85. Hou, C.; Li, J.; Zhao, L.; Zhang, W.; Luo, Q.; Dong, Z.; Xu, J.; Liu, J. Construction of protein nanowires through cucurbit[8]uril-based highly specific host–guest interactions: an approach to the assembly of functional proteins. *Angew. Chem. Int. Ed.* **2013**, *52*, 5590–5593.
86. Barrow, S. J.; Kasera, S.; Rowland, M. J.; Barrio, J. d.; Scherman, O. A.; Cucurbituril-based molecular recognition. *Chem. Rev.* **2015**, *115*, 12320–12406.
87. Lee, J. W.; Samal, S.; Selvapalam, N.; Kim, H.-J.; Kim, K. Cucurbituril homologues and derivatives: new opportunities in supramolecular chemistry. *Acc. Chem. Res.* **2003**, *36*, 621–630.
88. Mecozi, S.; Rebek, Jr. J. The 55% Solution: A formula for molecular recognition in the liquid state. *Chem. Eur. J.* **1998**, *4*, 1016–1022.
89. Biedermann, F.; Nau, W. M.; Schneider, H.-J. The hydrophobic effect revisited - studies with supramolecular complexes imply high-energy water as a noncovalent driving force. *Angew. Chem. Int. Ed.* **2014**, *53*, 11158–11171.
90. Valle, E. M. M. D. Cyclodextrins and their uses: A review. *Process Biochem.* **2004**, *39*, 1033–1046.

91. Basilio, N.; García-Río, L.; Moreira, J. A.; Pessêgo, M. Supramolecular catalysis by cucurbit[7]uril and cyclodextrins: similarity and differences. *J. Org. Chem.* **2010**, *75*, 848–855.
92. Koropatkin, N. M.; Martens, E. C.; Gordon, J. I.; Smith, T. J. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* **2008**, *16*, 1105–1115.
93. Cameron, E. A.; Maynard, M. A.; Smith, C. J.; Smith, T. J.; Koropatkin, N. M.; Martens, E. C. Multidomain carbohydrate-binding proteins involved in bacteroides thetaiotaomicron starch metabolism. *J. Biol. Chem.* **2012**, *287*, 34614–34625.
94. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A field guide to foldamers. *Chem. Rev.* **2001**, *101*, 3893–4012.
95. Guichard, G.; Huc, I. Synthetic foldamers. *Chem. Commun.* **2011**, *47*, 5933–5941.
96. Checco, J. W.; Gellman, S. H. Targeting recognition surfaces on natural proteins with peptidic foldamers. *Curr. Opin. Struct. Biol.* **2016**, *39*, 96–105.
97. Liu, S.; Cheloha, R. W.; Watanabe, T.; Gardella, T. J.; Gellman, S. H. Receptor selectivity from minimal backbone modification of a polypeptide agonist. *Proc. Natl. Acad. Sci.* **2018**, *115*, 12383–12388.
98. Maity, D.; Kumar, S.; Curreli, F.; Debnath, A. K.; Hamilton, A. D.  $\alpha$ -Helix-mimetic foldamers for targeting HIV-1 TAR RNA. *Chem. Eur. J.* **2019**, *25*, 7265–7269.
99. Hegedus, Z.; Grison, C. M.; Miles, J. A.; Rodriguez-Marin, S.; Warriner, S. L.; Webb, M. E.; Wilson, A. J. A catalytic protein–proteomimetic complex: using aromatic oligoamide foldamers as activators of RNase S. *Chem. Sci.* **2019**, *10*, 3956–3962.
100. Buratto, J.; Colombo, C.; Stupfel, M.; Dawson, S. J.; Dolain, C.; d'Estaintot, B. L.; Fischer, L.; Granier, T.; Laguerre, M.; Gallois, B.; Huc, I. Structure of a complex formed by a protein and a helical aromatic oligoamide foldamer at 2.1 Å resolution. *Angew. Chem. Int. Ed.* **2014**, *53*, 883–887.

101. Jewginski, M.; Granier, T.; d'Estaintot, B. L.; Fischer, L.; Mackereth, C. D.; Huc, I. Self-assembled protein–aromatic foldamer complexes with 2:3 and 2:2:1 stoichiometries. *J. Am. Chem. Soc.* **2017**, *139*, 2928–2931.
102. Reddy, P. S.; d'Estaintot, B. L.; Granier, T.; Mackereth, C. D.; Fischer, L.; Huc, I. Structure elucidation of helical aromatic foldamer-protein complexes with large contact surface areas. *Chem. Eur. J.* **2019**, *25*, 11042–11047.
103. Kim, K.-H.; Ko, D.-K.; Kim, Y.-T.; Kim, N. H.; Paul, J.; Zhang, S.-Q.; Murray, C. B.; Acharya, R.; DeGrado, W. F.; Kim, Y. H.; Grigoryan, G. Protein-directed self-assembly of a fullerene crystal. *Nat. Commun.* **2016**, *7*, e11429.
104. Pompidor, G.; D'Aléo, A.; Vicat, J.; Toupet, L.; Giraud, N.; Kahn, R.; Maury, O. Protein crystallography through supramolecular interactions between a lanthanide complex and arginine. *Angew. Chem. Int. Ed.* **2008**, *47*, 3388–3391.
105. Pompidor, G.; Maury, O.; Vicat, J.; Kahn, R. A dipicolinate lanthanide complex for solving protein structures using anomalous diffraction. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 762–769.
106. Dumont, E.; Pompidor, G.; D'Aléo, A.; Vicat, J.; Toupet, L.; Kahn, R.; Girard, E.; Maury, O.; Giraud, N. Exploration of the supramolecular interactions involving tris-dipicolinate lanthanide complexes in protein crystals by a combined biostructural, computational and NMR study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18235–18242.
107. Talon, R.; Nauton, L.; Canet, J.-L.; Kahn, R.; Girard, E.; Gautier, A. Clicked europium dipicolinate complexes for protein X-ray structure determination. *Chem. Commun.* **2012**, *48*, 11886–11888.
108. Engilberge, S.; Riobé, F.; Di Pietro, S.; Lassalle, L.; Coquelle, N.; Arnaud, C.-A.; Pitrat, D.; Mulatier, J.-C.; Madern, D.; Breyton, C.; Maury, O.; Girard, E. Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* **2017**, *8*, 5909–5917.

109. Schrader, T.; Bitan, G.; Klärner, F.-G. Molecular tweezers for lysine and arginine – powerful inhibitors of pathologic protein aggregation. *Chem. Commun.* **2016**, *52*, 11318–11334.
110. Landau, M.; Sawaya, M. R.; Faull, K. F.; Laganowsky, A.; Jiang, L.; Sievers, S. A.; Liu, J.; Barrio, J. R.; Eisenberg, D. Towards a pharmacophore for amyloid. *PLoS Biol.* **2011**, *9*, e1001080.
111. Sangwan, S.; Sawaya, M. R.; Murray, K. A.; Hughes, M. P.; Eisenberg, D. S. Atomic structures of corkscrew-forming segments of SOD1 reveal varied oligomer conformations. *Protein Sci.* **2018**, *27*, 1231–1242.
112. Morgan, H. P.; McNae, I. W.; Nowicki, M. W.; Zhong, W.; Michels, P. A. M.; Auld, D. S.; Fothergill-Gilmore, L. A.; Walkinshaw, M. D. The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **2011**, *286*, 31232–31240.
113. Murakami, M. T.; Arruda, E. Z.; Melo, P. A.; Martinez, A. B.; Calil-Eliás, S.; Tomaz, M. A.; Lomonte, B.; Gutiérrez, J. M.; Arni, R. K. Inhibition of myotoxic activity of *Bothrops asper* myotoxin II by the anti-trypanosomal drug suramin. *J. Mol. Biol.* **2005**, *350*, 416–426.
114. Schuetz, A.; Min, J.; Antoshenko, T.; Wang, C.-L.; Allali-Hassani, A.; Dong, A.; Loppnau, P.; Vedadi, M.; Bochkarev, A.; Sternglanz, R.; Plotnikov, A. N. structural basis of inhibition of the human NAD<sup>+</sup>-dependent deacetylase Sirt5 by suramin. *Structure* **2007**, *15*, 377–389.
115. Mastrangelo, E.; Pezzullo, M.; Tarantino, D.; Petazzi, R.; Germani, F.; Kramer, D.; Robel, I.; Rohayem, J.; Bolognesi, M.; Milani, M. Structure-based inhibition of norovirus RNA-dependent RNA polymerases. *J. Mol. Biol.* **2012**, *419*, 198–210.
116. Zhou, X.; Tan, T.-C.; Valiyaveetil, S.; Go, M. L.; Kini, R. M.; Velazquez-Campoy, A.; Sivaraman, J. Structural characterization of myotoxic ecarpholin S from *Echis carinatus* venom. *Biophys J.* **2008**, *95*, 3366–3380.
117. Ren, C.; Morohashi, K.; Plotnikov, A. N.; Jakoncic, J.; Smith, S. G.; Li, J.; Zeng, L.; Rodriguez, Y.; Stojanoff, V.; Walsh, M.; Zhou, M.-M. Small-molecule modulators

- of methyl-lysine binding for the CBX7 chromodomain. *Chem. Biol.* **2015**, *22*, 161–168.
118. Salvador, G. H. M.; Dreyer, T. R.; Gomes, A. A. S.; Cavalcante, W. L. G.; Dos, S. J. I.; Gandin, C. A.; de Oliveira, N. M.; Gallacci, M.; Fontes, M. R. M. Structural and functional characterization of suramin-bound MjTX-I from *Bothrops moojeni* suggests a particular myotoxic mechanism. *Sci Rep.* **2018**, *8*, e10317.
119. Long, D.-L.; Tsunashima R.; Cronin, L. Polyoxometalates: building blocks for functional nanoscale systems. *Angew. Chem. Int. Ed.* **2010**, *49*, 1736–1758.
120. Rompuy, L. S. V.; Parac-Vogt, T. N. Interactions between polyoxometalates and biological systems: from drug design to artificial enzymes. *Curr. Opin. Biotechnol.* **2019**, *58*, 92–99.
121. Bijeli, A.; Rompel, A. The use of polyoxometalates in protein crystallography – an attempt to widen a well-known bottleneck. *Coordin. Chem. Rev.* **2015**, *299*, 22–38.
122. Gumerova, N. I.; Rompel, A. Synthesis, structures and applications of electron-rich polyoxometalates. *Nat. Rev. Chem.* **2018**, *2*, e0112.
123. Zebisch, M.; Krauss, M.; Schäfer, P.; Straeter, N. Structures of *Legionella pneumophila* NTPDase1 in complex with polyoxometallates. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *D70*, 1147–1154.
124. Tocilj, A.; Schlünzen, F.; Janell, D.; Glühmann, M.; Hansen, H. A. S.; Harms, J.; Bashan, A.; Bartels, H.; Agmon, I.; Franceschi, F.; Yonath, A. The small ribosomal subunit from *Thermus thermophilus* at 4.5 Å resolution: Pattern fittings and the identification of a functional site. *Proc. Natl. Acad. Sci.* **1999**, *96*, 14252–14257.
125. Sap, A.; De Zitter, E.; Meervelt, L. V.; Parac-Vogt, T. N. Structural Characterization of the complex between hen egg-white lysozyme and ZrIV-substituted keggin polyoxometalate as artificial protease. *Chem. Eur. J.* **2015**, *21*, 11692–11695.
126. Vandebroek, L.; De Zitter, E.; Ly, H. G. T.; Conić, D.; Mihaylov, T.; Sap, A.; Proost, P.; Pierloot, K.; Meervelt, L. V.; Parac-Vogt, T. N. Protein-assisted formation and

- stabilization of catalytically active polyoxometalate species. *Chem. Eur. J.* **2018**, *24*, 10099–10108.
127. Vandebroek, L.; Mampaey, Y.; Antonyuk, S.; Meervelt, L. V.; Parac-Vogt, T. N. Noncovalent complexes formed between metal-substituted polyoxometalates and hen egg white lysozyme. *Eur. J. Inorg. Chem.* **2019**, *2019*, 506–511.
128. Bijelic, A.; Molitor, C.; Mauracher, S. G.; Al-Oweini, R.; Kortz, U.; Rompel, A. Hen egg-white lysozyme crystallisation: protein stacking and structure stability enhanced by a tellurium(VI)-centred polyoxotungstate. *ChemBioChem* **2015**, *16*, 233–241.
129. Mauracher, S. G.; Molitor, C.; Al-Oweini, R.; Kortz, U.; Rompel, A. Latent and active abPPO4 mushroom tyrosinase cocrystallized with hexatungstotellurate(VI) in a single crystal. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *D70*, 2301–2315.
130. Molitor, C.; Mauracher, S. G.; Rompel, A. Aurone synthase is a catechol oxidase with hydroxylase activity and provides insights into the mechanism of plant polyphenol oxidases. *Proc. Natl. Acad. Sci.* **2016**, *113*, E1806–E1815.
131. Mac Sweeney, A.; Chambovey, A.; Wicki, M.; Müller, M.; Artico, N.; Lange, R.; Bijelic, A.; Breibeck, J.; Rompel, A. The crystallization additive hexatungstotellurate promotes the crystallization of the HSP70 nucleotide binding domain into two different crystal forms. *PLoS One* **2018**, *13*, e0199639.
132. Ceccarelli, D. F. J.; Blasutig, I. M.; Goudreault, M.; Li, Z.; Ruston, J.; Pawson, T.; Sicheri, F. Non-canonical interaction of phosphoinositides with pleckstrin homology domains of Tiam1 and ArhGAP9. *J. Biol. Chem.* **2007**, *282*, 13864–13874.
133. Cash, J. N.; Davis, E. M.; Tesmer, J. J. G. Structural and biochemical characterization of the catalytic core of the metastatic factor P-Rex1 and its regulation by PtdIns(3,4,5)P3. *Structure* **2016**, *24*, 730–740.
134. Dick, R. A.; Zadrozny, K. K.; Xu, C.; Schur, F. K. M.; Lyddon, T. D.; Ricana, C. L.; Wagner, J. M.; Perilla, J. R.; Ganser-Pornillos, B. K.; Johnson, M. C.; Pornillos, O.;

- Vogt, V. M. Inositol phosphates are assembly co-factors for HIV-1. *Nature* **2018**, *560*, 509–512.
135. Best, M.; Zhang, D. H.; Prestwich, G. D. Inositol polyphosphates, diphosphoinositol polyphosphates and phosphatidylinositol polyphosphate lipids: structure, synthesis, and development of probes for studying biological activity. *Nat. Prod. Rep.* **2010**, *27*, 1403–1430.
136. Kim, E.; Ahn, H.; Kim, M. G.; Lee, H.; Kim, S. Expanding significance of inositol polyphosphate multikinase as a signaling hub. *Mol. Cells.* **2017**, *40*, 315–321.
137. Wang, Q.; Vogan, E. M.; Nocka, L. M.; Rosen, C. E.; Zorn, J. A.; Harrison, S. C.; Kuriyan, J. Autoinhibition of Bruton's tyrosine kinase (Btk) and activation by soluble inositol hexakisphosphate. *eLife* **2015**, *4*, e06074.
138. Chen, Q.; Perry, N. A.; Vishnivetskiy, S. A.; Berndt, S.; Gilbert, N. C.; Zhuo, Y.; Singh, P. K.; Tholen, J.; Ohi, M. D.; Gurevich, E. V.; Brautigam, C. A.; Klug, C. S.; Gurevich, V. V.; Iverson, T. M. Structural basis of arrestin-3 activation and signalling. *Nat. Commun.* **2017**, *8*, e1427.
139. Montpetit, B.; Thomsen, N. D.; Helmke, K. J.; Seeliger, M. A.; Berger, J. M.; Weis, K. A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **2011**, *472*, 238–242.
140. Xue, Y.; Lee, S.; Ha, Y. Crystal structure of amyloid precursor-like protein 1 and heparin complex suggests a dual role of heparin in E2 dimerization. *Proc. Natl. Acad. Sci.* **2011**, *108*, 16229–16234.
141. Desai, B. J.; Boothello, R. S.; Mehta, A. Y.; Scarsdale, J. N.; Wright, H. T.; Desai, U. R. Interaction of thrombin with sucrose octasulfate. *Biochemistry* **2011**, *50*, 6973–6982.
142. Goel, M.; Jain, D.; Kaur, K. J.; Kenoth, R.; Maiya, B. G.; Swamy M. J.; Salunke, D. M. Functional equality in the absence of structural similarity - An added dimension to molecular mimicry. *J. Biol. Chem.* **2001**, *276*, 39277–39281.
143. Goel, M.; Anuradha, P.; Kaur, K. J.; Maiya, B. G.; Swamy, M. J.; Salunke, D. M. Porphyrin binding to jacalin is facilitated by the inherent plasticity of the

- carbohydrate-binding site: novel mode of lectin-ligand interaction. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 281–288.
144. Goel, M.; Damai, R. S.; Sethi, D. K.; Kaur, K. J.; Maiya, B. G.; Swamy M. J.; Salunke, D. M. Crystal structures of the PNA–porphyrin complex in the presence and absence of lactose: Mapping the conformational changes on lactose binding, interacting surfaces, and supramolecular aggregations. *Biochemistry* **2005**, *44*, 5588–5596.
145. Schubert, H. L.; Phillips, J. D.; Heroux, A.; Hill, C. P. Structure and mechanistic implications of a uroporphyrinogen III synthase–product complex. *Biochemistry* **2008**, *47*, 8648–8655.
146. Parkinson, G. N.; Ghosh, R.; Neidle, S. Structural basis for binding of porphyrin to human telomeres. *Biochemistry* **2007**, *46*, 2390–2397.
147. Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human Myc promoter. *Nat. Chem. Biol.* **2005**, *1*, 167–173.
148. Morgan, H. P.; McNae, I. W.; Hsin, K.-Y.; Michels, P. A.; Fothergill-Gilmore, L. A.; Walkinshaw, M. D. An Improved strategy for the crystallization of *leishmania mexicana* pyruvate kinase. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010**, *F66*, 215–218.
149. Adhikary, R.; Yu, W.; Oda, M.; Walker, R. C.; Chen, T.; Stanfield, R. L.; Wilson, I. A.; Zimmermann, J.; Romesberg, F. E. Adaptive mutations alter antibody structure and dynamics during affinity maturation. *Biochemistry* **2015**, *54*, 2085–2093.
150. Fujii, A.; Sekiguchi, Y.; Matsumura, H.; Inoue, T.; Chung, W.-S.; Hirota, S.; Matsuo, T. Excimer emission properties on pyrene-labeled protein surface: correlation between emission spectra, ring stacking modes, and flexibilities of pyrene probes. *Bioconjug. Chem.* **2015**, *26*, 537–548.
151. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M. Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nat. Commun.* **2014**, *5*, e4634.

152. Yang, G.; Ding, H.-m.; Kochovski, Z.; Hu, R.; Lu, Y.; Ma, Y.-q.; Chen, G.; Jiang, M.; Highly ordered self-assembly of native proteins into 1D, 2D, and 3D structures modulated by the tether length of assembly-inducing ligands. *Angew. Chem. Int. Ed.* **2017**, *56*, 10691–10695.
153. Dotan, N.; Arad, D.; Frolow, F.; Freeman, A. Self-assembly of a tetrahedral lectin into predesigned diamondlike protein crystals. *Angew. Chem. Int. Ed.* **1999**, *38*, 2363–2366.
154. Alvarez-Paggi, D.; Hannibal, L.; Castro, M. A.; Oviedo-Rouco, S.; Demicheli, V.; Tórtora, V.; Tomasina, F.; Radi, R.; Murgida, D. H. Multifunctional cytochrome c: learning new tricks from an old dog. *Chem. Rev.* **2017**, *117*, 13382–13460.
155. Solmaz, S. R.; Hunte, C. Structure of complex III with bound cytochrome c in reduced state and definition of a minimal core interface for electron transfer. *J. Biol. Chem.* **2008**, *283*, 17542–17549.
156. Hannibal, L.; Tomasina, F.; Capdevila, D. A.; Demicheli, V.; Tórtora, V.; Alvarez-Paggi, D.; Jemmerson, R.; Murgida, D. H.; Radi, R. Alternative conformations of cytochrome c: structure, function, and detection. *Biochemistry*, **2016**, *55*, 407–428.
157. Shimada, S.; Shinzawa-Itoh, K.; Baba, J.; Aoe, S.; Shimada, A.; Yamashita, E.; Kang, J.; Tateno, M.; Yoshikawa, S.; Tsukihara, T. Complex structure of cytochrome c-cytochrome c oxidase reveals a novel protein-protein interaction mode. *EMBO J* 2017, *36*, 291–300.
158. Volkov, A. N.; Bashir, Q.; Worrall, J. A. R.; Ullmann, G. M.; Ubbink, M. Shifting the equilibrium between the encounter state and the specific form of a protein complex by interfacial point mutations. *J. Am. Chem. Soc.* **2010**, *132*, 11487–11495.
159. Louie, G. V.; Brayer, G. D. High-resolution refinement of yeast iso-1-cytochrome c and comparisons with other eukaryotic cytochromes c. *J. Mol. Biol.* **1990**, *214*, 527–555.

160. McGovern, R. E.; Feifel, S.C.; Lisdat, F.; Crowley, P. B. Microscale crystals of cytochrome c and calixarene on electrodes: interprotein electron transfer between defined sites. *Angew. Chem. Int. Ed.* **2015**, *54*, 6356–6359.
161. Manickam, P.; Kaushik, A.; K, C.; Bhansali, S. Recent advances in cytochrome c biosensing technologies. *Biosens. and Bioelectron.* **2017**, *87*, 654–668.
162. Marx, F.; Haas, H.; Reindl, M.; Stöffler, G.; Lottspeich, F.; Redl, B. Cloning, structural organization and regulation of expression of the *Penicillium chrysogenum* PAF gene encoding an abundantly secreted protein with antifungal activity. *Gene* **1995**, *167*, 167–171.
163. Leiter, É.; Szappanos, H.; Oberparleiter, C.; Kaiserer, L.; Csernoch, L.; Pusztahelyi, T.; Emri, T.; Pócsi, I.; Salvenmoser, W.; Marx, F. Antifungal protein PAF severely affects the integrity of the plasma membrane of *Aspergillus nidulans* and induces an apoptosis-like phenotype. *Antimicrob. Agents Chemother.* **2005**, *49*, 2445–2453.
164. Binder, U.; Oberparleiter, C.; Meyer, V.; Marx, F. The antifungal protein PAF interferes with PKC/MPK and cAMP/PKA signalling of *Aspergillus nidulans*. *Mol. Microbiol.* **2010**, *75*, 294–307.
165. Garrigues, S.; Gandía, M.; Borics, A.; Marx, F.; Manzanares, P.; Marcos J. F. Mapping and identification of antifungal peptides in the putative antifungal protein AfpB from the filamentous fungus *Penicillium digitatum*. *Front. Microbiol.* **2017**, *8*, 1–13.
166. Batta, G.; Barna, T.; Gaspari, Z.; Sandor, S.; Kövér, K. E.; Binder, U.; Sarg, B.; Kaiserer, L.; Chhillar, A. K.; Eigentler, A. Functional aspects of the solution structure and dynamics of PAF - A highly-stable antifungal protein from *Penicillium chrysogenum*. *FEBS J.* **2009**, *276*, 2875–2890.
167. Fizil, Á.; Sonderegger, C.; Czajlik, A.; Fekete, A.; Komáromi, I.; Hajdu, D.; Marx, F.; Batta, G. Calcium binding of the antifungal protein PAF: Structure, dynamics and function aspects by NMR and MD simulations. *PLoS One* **2018**, *13*, e0204825.

168. McPherson, A.; Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr., F Struct. Biol. Commun.* **2014**, *70*, 1–20.
169. Chayen, N. E. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallogr., D Struct. Biol.* **1998**, *D54*, 8–15.
170. Asherie, N. Protein crystallization and phase diagrams. *Methods* **2004**, *34*, 266–272.
171. Sauter, A.; Roosen-Runge, F.; Zhang, F.; Lotze, G.; Jacobs, R. M. J.; Schreiber, F. Real-time observation of nonclassical protein crystallization kinetics. *J. Am. Chem. Soc.* **2015**, *137*, 1485–1491.
172. Abe, S.; Maity, B.; Ueno, T. Functionalization of protein crystals with metal ions, complexes and nanoparticles. *Curr. Opin. Chem. Biol.* **2017**, *43*, 68–76.
173. Lopez, S.; Rondot, L.; Leprêtre, C.; Marchi-Delapierre, C.; Ménage, S.; Cavazza, C. Cross-linked artificial enzyme crystals as heterogeneous catalysts for oxidation reactions. *J. Am. Chem. Soc.* **2017**, *139*, 17994–18002.
174. Longenecker, K. L.; Garrard, S. M.; Sheffield, P. J.; Derewenda, Z. S. Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 679–688.
175. Walter, T. S.; Meier, C.; Assenberg, R.; Au, K. F.; Ren, J.; Verma, A.; Nettleship, J. E.; Owens, R. J.; Stuart, D. I.; Grimes, J. M. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **2006**, *14*, 1617–1622.
176. Cooper, D. R.; Boczek, T.; Grelewska, K.; Pinkowska, M.; Sikorska, M.; Zawadzki, M.; Derewenda, Z. Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 636–645.
177. Price II, W. N.; Chen, Y.; Handelman, S. K.; Neely, H.; Manor, P.; Karlin, R.; Nair, R.; Liu, J.; Baran, M.; Everett, J.; Tong, S. N.; Forouhar, F.; Swaminathan, S. S.; Acton, T.; Xiao, R.; Luft, J. R.; Lauricella, A.; DeTitta, G. T.; Rost, B.; Montelione,

- G. T.; Hunt, J. F. Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat. Biotech.* **2009**, *27*, 51–57.
178. Sledz, P.; Zheng, H.; Murzyn, K.; Chruszcz, M.; Zimmerman, M. D.; Chordia, M. D.; Joachimiak, A.; Minor, W. New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Sci.* **2010**, *19*, 1395–1404.
179. Trusch, F.; Kowski, K.; Bravo-Rodriguez, K.; Beuck, C.; Sowislok, A.; Wettig, B.; Matena, A.; Sanchez-Garcia, E.; Meyer, H.; Schrader, T.; Bayer, P. Molecular tweezers target a protein-protein interface and thereby modulate complex formation. *Chem. Commun.* **2016**, *52*, 14141–14144.
180. Atwood, J. L.; Ness, T.; Nichols, P. J.; Raston, C. L. Confinement of amino acids in tetra-*p*-sulfonated calix[4]arene bilayers. *Cryst. Growth Des.* **2002**, *2*, 171–176.
181. Selkti, M.; Coleman, A. W.; Nicolis, I.; Douteau-Guével, N.; Villain, F.; Tomas, A.; de Rango, C. The first example of a substrate spanning the calix[4]arene bilayer: the solid state complex of *p*-sulfonatocalix[4]arene with L-lysine. *Chem. Commun.* **2000**, 161–162.
182. Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. Moulding calixarenes for biomacromolecule targeting. *Chem. Commun.* **2015**, *51*, 14140–14159.
183. Atwood, J. L. *Comprehensive Supramolecular Chemistry II*, 2nd Edition, **2017**, Elsevier.
184. Mallon, M.; Dutt, S.; Schrader, T.; Crowley, P. B. Protein camouflage: supramolecular anion recognition by ubiquitin. *ChemBioChem* **2016**, *17*, 774–783.
185. Martin, A. D.; Raston, C. L. Multifunctional *p*-phosphonated calixarenes. *Chem. Commun.* **2011**, *47*, 9764–9772.
186. Zadnani, R.; Schrader, T. Nanomolar protein sensing with embedded receptor molecules. *J. Am. Chem. Soc.* **2005**, *127*, 904–915.

- 
187. Kolusheva, S.; Zadmard, R.; Schrader, T.; Jelinek, R. Color fingerprinting of proteins by calixarenes embedded in lipid/ polydiacetylene vesicles. *J. Am. Chem. Soc.* **2006**, *128*, 13592–13598.
188. Trush, V. V.; Kharchenko, S. G.; Tanchuk, V. Y.; Kalchenko, V.I.; Vovk, A. I. Phosphonate monoesters on a thiacalix[4]arene framework as potential inhibitors of protein tyrosine phosphatase 1B. *Org. Biomol. Chem.* **2015**, *13*, 8803–8806
189. Almi, M.; Arduini, A.; Casnati, A.; Pochini, A.; Ungaro, R. Chloromethylation of calixarenes and synthesis of new water soluble macrocyclic hosts. *Tetrahedron*, **1989**, *45*, 2177–2182. pmclx4 was obtained after salification with 4 equiv. of 0.1 M NaOH and crystallized as a beige solid upon addition of ethanol.
190. Mourer, M.; Psychogios, N.; Laumond, G.; Aubertin, A.-M.; Regnouf-de-Vains, J.-B. Synthesis and anti-HIV evaluation of water-soluble calixarene-based bithiazolyl podands. *Bioorg. Med. Chem.* **2010**, *18*, 36–45.
191. Crowley P. B.; Ganji, P; Ibrahim, H. Protein surface recognition: structural characterisation of cytochrome c-porphyrin Complexes. *ChemBioChem* **2008**, *9*, 1029–1033.
192. Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 125–132.
193. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *D67*, 282–292.
194. Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *D69*, 1204–1214.
195. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
196. Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *D68*, 431–440.

197. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of COOT. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 486–501.
198. Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2184–2195.
199. Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 12–21.
200. Lee, B.; Richards, F. M.; The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379–400.
201. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **1995**, *6*, 277-293.
202. Janin, J.; Bahadur, R. P.; Chakrabarti, P. Q. Protein-protein interaction and quaternary structure. *Rev. Biophys.* **2008**, *41*, 133–180.
203. Edwardson, T. G. W.; Mori T.; Hilvert, D. Rational engineering of a designed protein cage for siRNA delivery. *J. Am. Chem. Soc.* **2018**, *140*, 10439–10442.
204. Nguyen, T. K.; Negishi, H.; Abe, S.; Ueno, T. Construction of supramolecular nanotubes from protein crystals. *Chem. Sci.* **2019**, *10*, 1046–1051.
205. Schreiber, A.; Stühn, L. G.; Huber, M. C.; Geissinger, S. E.; Rao A.; Schiller, S. M. Self-assembly toolbox of tailored supramolecular architectures based on an amphiphilic protein library. *Small* **2019**, *15*, e1900163.
206. Casnati, A.; Ting, Y.; Berti, D.; Fabbi, M.; Pochini, A.; Ungaro, R.; Sciotto, D.; Lombard, G. Synthesis of water soluble molecular receptors from calix[4]arenes fixed in the cone conformation. *Tetrahedron*, **1993**, *49*, 9815–9822.
207. Arena, G.; Casnati, A.; Contino, A.; Lombardo, G. G.; Sciotto D.; Ungaro, R. Water-Soluble calixarene hosts that specifically recognize the

- trimethylammonium group or the benzene ring of aromatic ammonium cations: A combined H-1 NMR, calorimetric, and molecular mechanics investigation. *Chem. Eur. J.* **1999**, *5*, 738–744.
208. Arena, G.; Casnati, A.; Contino, A.; Magrì, A.; Sansone, F.; Sciotto, D.; Ungaro, R. Inclusion of naturally occurring amino acids in water soluble calix[4]arenes: A microcalorimetric and <sup>1</sup>H NMR investigation supported by molecular modelling. *Org. Biomol. Chem.* **2006**, *4*, 243–249.
209. D'Urso, A.; Brancatelli, G.; Hickey, N.; Farnetti, E.; Zorzi, R. De; Bonaccorso, C.; Purrello, R.; Geremia, S. Interactions of a water-soluble calix[4]arene with spermine: Solution and solid-state characterisation. *Supramol. Chem.* **2016**, *28*, 499–505.
210. Costanzo, L. D.; Geremia, S.; Randaccio, L.; Purrello, R.; Lauceri, R.; Sciotto, D.; Gulino, F. G.; Pavone, V. Calixarene - porphyrin supramolecular complexes: pH-tuning of the complex stoichiometry. *Angew. Chem. Int. Ed.* **2001**, *40*, 4245–4247.
211. Zorzi, R. De.; Guidolin, N.; Randaccio, L.; Purrello, R.; Geremia, S. Nanoporous Crystals of calixarene/porphyrin supramolecular complex functionalized by diffusion and coordination of metal ions. *J. Am. Chem. Soc.* **2009**, *131*, 2487–2489.
212. Brancatelli, G.; Zorzi, R. De.; Hickey, N.; Siega, P.; Zingone, G.; Geremia, S. New multicomponent porous architecture of self-assembled porphyrins/calixarenes driven by nickel ions. *Cryst. Growth Des.* **2012**, *12*, 5111–5117.
213. Bushnell, G. W.; Louie, G. V.; Brayer, G. D. High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* **1990**, *214*, 585–595.
214. Madeira, F.; Park, Y. m.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, 636–641.
215. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *D67*, 293–302.

216. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J. R.; Grosse-Kunstleve, W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 213–221.
217. Smart, O. S.; Womack, T. O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *D68*, 368–380.
218. Fucke, K.; Anderson, K. M.; Filby, M. H.; Henry, M.; Wright, J.; Mason, S. A.; Gutmann, M. J.; Barbour, L. J.; Oliver, C.; Coleman, A. W.; Atwood, J. L.; Howard, J. A. K.; J. W. Steed. The structure of water in *p*-sulfonatocalix[4]arene. *Chem. Eur. J.* **2011**, *17*, 10259–10271.
219. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797.
220. Juers, D. H.; Ruffin, J. MAP\_CHANNELS: A computation tool to aid in the visualization and characterization of solvent channels in macromolecular crystals. *J. Appl. Crystallogr.* **2014**, *47*, 2105–2108.
221. Arena, G.; Contino, A.; Lombardo, G. G.; Sciotto, D. Water-soluble calix[6]arenes. characterization of 5,11,17,23,29,35-hexasulphonate-37,38,39,40,41,42-hexahydroxycalix[6]arene and thermodynamic study of proton complex formation. *Thermochim. Acta* **1995**, *264*, 1–11.
222. Suga, K.; Ohzono, T.; Negishi, M.; Deuchi, K. Effect of various cations on the acidity of *p*-sulfonatocalixarenes. *Supramol. Sci.* **1998**, *5*, 9–14.
223. March, M. D.; Demitri, N.; Zorzi, R. D.; Casini, A.; Gabbiani, C.; Guerri, A.; Messori, L.; Geremia, S. Nitrate as a probe of cytochrome *c* surface: crystallographic identification of crucial "hot spots" for protein-protein recognition. *J. Inorg. Biochem.* **2014**, *135*, 58–67.

224. Auwärter, W.; Écija, D.; Klappenberger F.; Barth, J. V. Porphyrins at interfaces. *Nat. Chem.* **2015**, *7*, 105–120.
225. Kielmann, M.; Senge, M. O. Molecular Engineering of Free-base porphyrins as ligands—the N···H···X binding motif in tetrapyrroles. *Angew. Chem. Int. Ed.* **2019**, *58*, 418–441.
226. D. Masih, V. Chernikova, O. Shekhah, M. Eddaoudi and O. F. Mohammed, Zeolite-like metal–organic framework (MOF) encaged Pt(ii)-porphyrin for anion-selective sensing. *ACS Appl. Mater. Interfaces*, **2018**, *101*, 11399-11405.
227. C. Schöttle, E. Guan, A. Okrut, N. A. Grosso-Giordano, A. Palermo, A. Solovyov, B. C. Gates and A. Katz, Bulky calixarene ligands stabilize supported iridium pair-site catalysts. *J. Am. Chem. Soc.*, **2019**, *141*, 4010-4015.
228. Smith, P. T.; Benke, B. P.; Cao, Z.; Kim, Y.; Nichols, E. M.; Kim, K; Chang, C. J. Iron porphyrins embedded into a supramolecular porous organic cage for electrochemical CO<sub>2</sub> reduction in water. *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 9684–9688.
229. Kubát, P; Šebera, J.; Zálíš, S.; Langmaier, J.; Fuciman, M.; Polívka, T.; Lang, K. Charge transfer in porphyrin–calixarene complexes: ultrafast kinetics, cyclic voltammetry, and DFT calculations *Phys. Chem. Chem. Phys.* **2011**, *13*, 6947–6954.
230. Otsuki, J. Supramolecular approach towards light-harvesting materials based on porphyrins and chlorophylls. *J. Mater. Chem. A* **2018**, *6*, 6710–6753.
231. Pathak, P.; Yao, W.; Hook, K. D.; Vik, R.; Winnerdy, F. R.; Brown, J. Q.; Gibb, B. C.; Pursell, Z. F.; Phan, A. T.; Jayawickramarajah, J. Bright G-quadruplex nanostructures functionalized with porphyrin lanterns. *J. Am. Chem. Soc.*, 2019, **141**, 12582–12591.
232. Yamasaki, Y.; Sekiya R.; Haino, T. Hexameric assembly of 5,17-di-substituted calix[4]arene in the solid state. *CrystEngComm.* **2017**, *19*, 6744–6751.
233. Jiang, H.; Zhang, L.; Chen, J.; Liu, M. Hierarchical self-assembly of a porphyrin into chiral macroscopic flowers with superhydrophobic and enantioselective property. *ACS Nano* **2017**, *11*, 12453–12460.

234. Haver, R.; Tejerina, L.; Jiang, H.-W.; Rickhaus, M.; Jirasek, M.; Grübner, I.; Eggimann, H. J.; Herz, L. M.; Anderson, H. L. Tuning the circumference of six-porphyrin nanorings. *J. Am. Chem. Soc.* **2019**, *141*, 7965–7971.
235. Moradi, M.; Opara, N. L.; Tulli, L. G.; Wäckerlin, C.; Dalgarno, S. J.; Teat, S. J.; Baljovic, M.; Popova, O.; van Genderen, E.; Kleibert, A.; Stahlberg, H.; Abrahams, J. P.; Padeste, C.; Corvini, P. F.-X.; Jung, T. A.; Shahgaldian, P. Supramolecular architectures of molecularly thin yet robust free-standing layers. *Sci. Adv.* **2019**, *5*, e4489.
236. Martinez-Bulit, P.; O’Keefe, C. A.; Zhu, K.; Schurko, R. W.; Loeb, S. J. Solvent and steric influences on rotational dynamics in porphyrinic metal–organic frameworks with mechanically interlocked pillars. *Cryst. Growth Des.* **2019**. In print.
237. Moschetto, G.; Lauceri, R.; Gulino, F. G.; Sciotto, D.; Purrello, R. Non-covalent synthesis in aqueous solution of discrete multi-porphyrin aggregates with programmable stoichiometry and sequence. *J. Am. Chem. Soc.* **2002**, *124*, 14536–14537.
238. Gulino, F. G.; Lauceri, R.; Frish, L.; Evan-Salem, T.; Cohen, Y.; Zorzi, R. D.; Geremia, S.; Di Costanzo, L.; Randaccio, L.; Sciotto, D.; Purrello, R. Noncovalent synthesis in aqueous solution and spectroscopic characterization of multi-porphyrin complexes. *Chem. Eur. J.* **2006**, *12*, 2722–2729.
239. D’Urso, A.; Cristaldi, D. A.; Fragalà, M. E.; Gattuso, G.; Pappalardo, A.; Villari, V.; Micali, N.; Pappalardo, S.; Parisi, M. F.; Purrello, R. Sequence, stoichiometry, and dimensionality control in porphyrin/bis-calix[4]arene self-assemblies in aqueous solution. *Chem. Eur. J.* **2010**, *16*, 10439–10446
240. Zorzi, R. De.; Guidolin, N.; Randaccio L.; Geremia, S. A bifunctionalized porous material containing discrete assemblies of copper-porphyrins and calixarenes metallated by ion diffusion. *CrystEngComm* **2010**, *12*, 4056–4058.
241. D’Urso, A.; Marino, N.; Gaeta, M.; Rizzo, M. S.; Cristaldi, D. A.; Fragalá, M. E.; Pappalardo, S.; Gattuso, G.; Notti, A.; Parisi, M. F.; Pisagatti I.; Purrello, R. Porphyrin stacks as an efficient molecular glue to induce chirality in hetero-

- component calixarene–porphyrin assemblies. *New J. Chem.* **2017**, *41*, 8078–8083.
242. Kokhan, O.; Ponomarenko, N.; Pokkuluri, P. R.; Schiffer M.; Tiede, D. M. Multimerization of solution-state proteins by tetrakis(4-sulfonatophenyl)porphyrin. *Biochemistry* **2014**, *53*, 5070–5079.
243. Dalgarno, S. J.; Hardie, M. J.; Makha, M.; Raston, C. L. Controlling the conformation and interplay of *p*-sulfonatocalix[6]arene as lanthanide crown ether complexes. *Chem. Eur. J.* **2003**, *9*, 2834–2839.
244. Perret, F.; Bonnard, V.; Danylyuk, O.; Suwinska, K.; Coleman, A. W. Conformational extremes in the supramolecular assemblies of *p*-sulfonato-calix[8]arene. *New J. Chem.* **2006**, *30*, 987–990.
245. Smith, C. B.; Barbour, L. J.; Makha, M.; Raston, C. L.; Soboleva, A. N. Unlocking the elusive binding cavity in *p*-sulfonatocalix[8]arene. *New J. Chem.* **2006**, *30*, 991–996.
246. Danylyuk, O.; Perret, F.; Coleman, A. W.; Suwinska, K. The solid-state complex of para-sulphonato-calix[8]arene anion with dimethylammonium cations. *The Open Crystallo. J.* **2008**, *1*, 18–23.
247. He, W.; Bi, Y.; Liao, W.; Li, D. A ternary supramolecular compound of *p*-sulfonatocalix[8]arene with 1D channels. *J. Mol. Struct.* **2009**, *937*, 95–99.
248. Liu, Y.; Liao, W.; Bi, Y.; Wang, M.; Wu, Z.; Wang, X.; Sub, Z.; Zhang, H. 1,2,3,4-Alternate double cone conformational extreme in the supramolecular assemblies of *p*-sulfonatocalix[8]arene. *CrystEngComm* **2009**, *11*, 1803–1806.
249. Lesniewska, B.; Perret, F.; Suwinska, K.; Coleman, A. W. Structural characterization of inclusion complexes of para-sulphonato-calix[8]arene with 1,2-bis(4-pyridyl)-ethane and 1,3-bis(4-pyridyl)-propane. new ‘double cone’ and ‘up–flat–down’ conformations of para-sulphonato-calix[8]arene. *CrystEngComm* **2014**, *16*, 4399–4405.
250. Danylyuka, O.; Butkiewicz, H.; Coleman, A. W.; Suwinska, K. Host-guest complexes of local anesthetics with cucurbit[6]uril and para-sulphonatocalix[8]arene in the solid state. *J. Mol. Struct.*, **2017**, *1150*, 28–36.

251. Meng, W.; Breiner, B.; Rissanen, K.; Thoburn, J. D.; Clegg, J. K.; Nitschke, J. R. A Self-assembled  $M_8L_6$  cubic cage that selectively encapsulates large aromatic guests. *Angew. Chem. Int. Ed.* **2011**, *50*, 3479–3483.
252. Park, Y.; Koo, J. Y.; Choi, H. C. Additive-free morphology control of organic polyhedral molecular crystals by the antisolvent molecular geometry: from rod, disk, to cube. *Cryst. Growth Des.* **2018**, *18*, 7239–7243.
253. Dodson, E. J.; Woolfson, M. M. ACORN2: New Developments of the ACORN concept. *Acta Crystallogr., D Biol Crystallogr.* **2009**, *D65*, 881–891.
254. Sheldrick, G.M. A short history of SHELX. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *D64*, 112–122.
255. Spek, A.L. Structure validation in chemical crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **2009**, *D65*, 148–155.
256. Steed, K. M.; Steed, J. W. Packing problems: high  $Z'$  crystal structures and their relationship to cocrystals, inclusion compounds, and polymorphism. *Chem. Rev.* **2015**, *115*, 2895–2933
257. Clegg, W. Crystallographic curiosities: polymorphism and structures with  $Z' > 1$ . *Acta Crystallogr., C Struct. Chem.*, **2019**, *C75*, 833–834.
258. Alvarenga, M. E.; Valdo, A. K. S. M.; Ribeiro, L.; Neto, J. A. Do N.; Araujo, D. P. De; Silva, C. M. Da; Fátima, A. De; Martins, F. T. *Acta Crystallogr., C Struct. Chem.*, **2019**, *C75*, 667–677. Exploring the structural landscape of 2-(thiophen-2-yl)-1,3-benzothiazole: high- $Z'$  packing polymorphism and cocrystallization with calix[4]tube.
259. McArdle, P. Oscal, A program package for small-molecule single-crystal crystallography with crystal morphology prediction and molecular modelling. *J. Appl. Crystallogr.* **2017**, *50*, 320–326.
260. Baldini, L.; Casnati, A.; Sansone, F. in *Comprehensive Supramolecular Chemistry II*, **2017**, *4*, 371–408. Editors. Atwood, J., Gokel, G. W. & Barbour, L. Elsevier.
261. Wang, K.; Guo, D. S.; Zhao, M. Y.; Liu, Y. A supramolecular vesicle based on the complexation of p-sulfonatocalixarene with protamine and its trypsin-triggered controllable-release properties. *Chem. Eur. J.* **2016**, *22*, 1475–1483.

262. Atwood, J. L.; Clark, D. L.; Juneja, R. K.; Orr, G. W.; Robinson, K. D.; Vincent, R. L. Double partial cone conformation for Na<sub>8</sub>{calix[6]arene sulfonate}·20.5H<sub>2</sub>O and its parent acid. *J. Am. Chem. Soc.* **1992**, *114*, 7558–7559.
263. Marx, F.; Binder, U.; Leiter, É.; Pócsi, I. The *Penicillium chrysogenum* antifungal protein PAF, a promising tool for the development of new antifungal therapies and fungal cell biology studies. *Cell. Mol. Life Sci.* **2008**, *65*, 445–454.
264. Palicz, Z.; Gáll, T.; Leiter, É.; Kollár, S.; Kovács, I.; Miszti-Blasius, K.; Pócsi, I.; Csernoch, L.; Szentesi, P. Application of a low molecular weight antifungal protein from *Penicillium chrysogenum* (PAF) to treat pulmonary aspergillosis in mice. *Emerg. Microbes Infect.* **2016**, *5*, 1–9.
265. Fizil, Á.; Gáspári, Z.; Barna, T.; Marx, F.; Batta, G. “Invisible” conformers of an antifungal disulfide protein revealed by constrained cold and heat unfolding, CEST-NMR experiments, and molecular dynamics calculations. *Chem. Eur. J.* **2015**, *21*, 5136–5144.
266. Sonderegger, C.; Galgóczy, L.; Garrigues, S.; Fizil, Á.; Borics, A.; Manzanares, P.; Hegedüs, N.; Huber, A.; Marcos, J. F.; Batta, G.; Marx, F. A *Penicillium chrysogenum*-based expression system for the production of small, cysteine-rich antifungal proteins for structural and functional analyses. *Microb. Cell Fact.* **2016**, *15*, 1–14.
267. Silva, P. M.; Gonçalves, S.; Santos, N. C. Defensins: antifungal lessons from eukaryotes. *Front. Microbiol.* **2014**, *5*, 1–17.
268. Poon, I. K. H.; Baxter, A. A.; Lay, F. T.; Mills, G. D.; Adda, C. G.; Payne, J. A.E.; Phan, T. K.; Ryan, G. F.; White, J. A.; Veneer, P. K.; van der Weerden, N. L.; Anderson, M. A.; Kvensakul, M.; Hulett, M. D. Phosphoinositide-mediated oligomerization of a defensin induces cell lysis. *eLife* **2014**, *3*, e01808.
269. Kvensakul, M.; Lay, F. T.; Adda, C. G.; Veneer, P. K.; Baxter, A. A.; Phan, T. K.; Poon, I. K. H.; Hulett, M. D. Nsd7 binds PA to form oligomeric complexes. *Proc. Natl. Acad. Sci.* **2016**, *113*, 11202–11207.
270. Cools, T.L.; Vriens, K.; Struyfs, C.; Verbandt, S.; Ramada, M.H.S.; Brand, G. D.; Bloch, C. Jr.; Koch, B.; Traven, A.; Drijfhout, J. W.; Demuyser, L.; Kucharíková,

- S.; Van, D. P.; Spasic, D.; Lammertyn, J.; Cammue, B. P. A.; Thevissen, K. The antifungal plant defensin HsAFP1 is a phosphatidic acid-interacting peptide inducing membrane permeabilization. *Front. Microbiol.* **2017**, *8*, 1–13.
271. Järvå, M.; Lay, F. T.; Phan, T. K.; Humble, C.; Poon, I. K. H.; Bleackley, M. R.; Anderson, M. A.; Hulett, M. D.; Kvensakul, M. X-ray structure of a carpet-like antimicrobial defensin-phospholipid membrane disruption complex. *Nat. Commun.* **2018**, *9*, 1–10.
272. Battye, T. G. G.; Kontogiannis, L.; Johnson, O., Powell, H. R.; Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with iMOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 271–281.
273. Keller, S.; Vargas, C.; Zhao, H.; Piszczek, G.; Brautigam, C. A.; Schuck, P. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal. Chem.* **2012**, *84*, 5066–5073.
274. Duvvuri, H.; Wheeler L. C.; Harms M. J. Pytc: Open-source python software for global analyses of isothermal titration calorimetry data. *Biochemistry*, **2018**, *57*, 2578–2583.
275. Jain, A.; Ramanathan, V.; Sankararamakrishnan, R. Lone pair- $\pi$  interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds. *Protein Sci.* **2009**, *18*, 595–605.
276. Zhu, X.; Dickerson, T. J.; Rogers, C. J.; Kaufmann, G. F.; Mee, J. M.; McKenzie, K. M.; Janda, K. D.; Wilson, I. A. Complete reaction cycle of a cocaine catalytic antibody at atomic resolution. *Structure* **2006**, *14*, 205–216.
277. Gautam, G.; Rehman, S. A. A.; Pandey, P.; Gourinath, S. Crystal structure of the PEG-bound SH3 domain of myosin IB from *Entamoeba histolytica* reveals its mode of ligand recognition. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2017**, *73*, 672–682.
278. Baxter, N. J.; Lilley, T. H.; Haslam, E.; Williamson, M. P. Multiple interactions between polyphenols and a salivary proline-rich protein repeat result in complexation and precipitation. *Biochemistry* **1997**, *36*, 5566–5577.

279. Charlton, A. J.; Haslam, E.; Williamson, M. P. Multiple conformations of the proline-rich protein/epigallocatechin gallate complex determined by time-averaged nuclear Overhauser effects. *J. Am. Chem. Soc.* **2002**, *124*, 9899–9905.
280. Quideau, S.; Deffieux, D.; Douat-Casassus, C.; Pouységu, L. Plant polyphenols: chemical properties, biological activities, and synthesis. *Angew. Chem. Int. Ed.* **2011**, *50*, 586–621.
281. Fonseca-Ornelas, L.; Schmidt, C.; Camacho-Zarco, A. R.; Fernandez, C. O.; Becker, S.; Zweckstetter, M. Small-molecule-induced soluble oligomers of  $\alpha$ -synuclein with helical structure. *Chem. Eur. J.* **2017**, *23*, 13010–13014.
282. Patil, A.; Huard, D.; Fonnesbeck, C. J. PyMC: Bayesian stochastic modelling in python. *J. Stat. Soft.* **2010**, *35*, 1–81.
283. Brautigam, C. A. Fitting two- and three-site binding models to isothermal titration calorimetric data. *Methods.* **2015**, *76*, 124–136.
284. Ernst, J. T.; Becerril, J.; Park, H. S.; Yin, H.; Hamilton, A. D. Design and application of an alpha-helix-mimetic scaffold based on an oligoamide-foldamer strategy: antagonism of the Bak BH3/Bcl-xL complex. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 535–539.
285. Kumar, S.; Birol, M.; Schlamadinger, D. E.; Wojcik, S. P.; Rhoades, E.; Miranker, A. D. Foldamer-mediated manipulation of a pre-amyloid toxin. *Nat. Commun.* **2016**, *25*, e11412.
286. Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nat. Chem. Biol.* **2007**, *3*, 252–262.
287. Checco, J. W.; Kreitler, D. F.; Thomas, N. C.; Belair, D. G.; Rettko, N. J.; Murphy, W. L.; Forest, K. T.; Gellman, S. H. Targeting diverse protein–protein interaction interfaces with  $\alpha/\beta$ -peptides derived from the z-domain scaffold. *Proc. Natl. Acad. Sci.* **2015**, *112*, 4552–4557.
288. Zhang, D. W.; Zhao, X.; Hou, J. L.; Li, Z. T. Aromatic amide foldamers: structures, properties, and functions. *Chem. Rev.* **2012**, *112*, 5271–5316.

289. Ferrand, Y.; Huc, I. Designing helical molecular capsules based on folded aromatic amide oligomers. *Acc. Chem. Res.* **2018**, *51*, 970–977.
290. Tsiamantas, C.; Dawson, S. J.; Huc, I. Solid phase synthesis of oligoethylene glycol-functionalized quinolinecarboxamide foldamers with enhanced solubility properties. *C. R. Chimie* **2016**, *19*, 132–142.
291. Lockhart, Z.; Knipe, P. C. Conformationally programmable chiral foldamers with compact and extended domains controlled by monomer structure. *Angew. Chem. Int. Ed.* **2018**, *130*, 8614–8618.
292. Lang, C.; Deng, X.; Yang, F.; Yang, B.; Wang, W.; Qi, S.; Zhang, X.; Zhang, C.; Dong, Z.; Liu, J. Highly selective artificial potassium ion channels constructed from pore-containing helical oligomers. *Angew. Chem. Int. Ed.* **2017**, *56*, 12668–12671.
293. Gole, B.; Kauffmann, B.; Maurizot, V.; Huc, I.; Ferrand, Y. Light-controlled conformational switch of an aromatic oligoamide foldamer. *Angew. Chem. Int. Ed.* **2019**, *58*, 8063–8067.
294. Rogers, J. M.; Kwon, S.; Dawson, S. J.; Mandal, P. K.; Suga, H.; Huc, I. Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids. *Nat. Chem.* **2018**, *10*, 405–412.
295. Vallade, M.; Jewginski, M.; Fischer, L.; Buratto, J.; Bathany, K.; Schmitter, J. M.; Stupfel, M.; Godde, F.; Mackereth, C. D.; Huc, I. Assessing interactions between helical aromatic oligoamide foldamers and protein surfaces: a tethering approach. *Bioconjug Chem.* **2018**, *30*, 54–62.
296. Mandal, P. K.; Baptiste, B.; d'Estaintot, B. L.; Kauffmann, B.; Huc, I. Multivalent interactions between an aromatic helical foldamer and a DNA G-quadruplex in the solid state. *ChemBioChem* **2016**, *17*, 1911–1914.
297. Hu, X.; Dawson, S. J.; Mandal, P. K.; de Hatten, X.; Baptiste, B.; Huc, I. Optimizing side chains for crystal growth from water: a case study of aromatic amide foldamers. *Chem Sci.* **2017**, *1*, 3741–3749.
298. Hu, X.; Cebe, P.; Weiss, A. S.; Omenetto, F.; Kaplan, D. L. Protein-based composite materials. *Mater. Today* **2012**, *15*, 208–215.

- 
299. Vallade, M.; Reddy, P. S.; Fischer, L.; Huc, I. Enhancing aromatic foldamer helix dynamics to probe interactions with protein surfaces. *Eur. J. Org. Chem.* **2018**, *2018*, 5489–5498.
300. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005**, *59*, 687–696.
301. Crowley, P. B.; Matias, P. M.; Khan, A. R.; Roessle, M.; Svergun, D. I. Metal-mediated self-assembly of a  $\beta$ -sandwich protein. *Chem. Eur. J.* **2009**, *15*, 12672–12680.
302. Churchfield, L. A.; Alberstein, R. G.; Williamson, L. M.; Tezcan, F. A. Determining the structural and energetic basis of allostery in a de novo designed metalloprotein assembly. *J. Am. Chem. Soc.* **2018**, *140*, 10043–10053.
303. Jacques, J.; Collet, A.; Wilen, S. H. in *Enantiomers, racemates and resolutions*, 3rd ed., Krieger, Malabar, **1994**.
304. Crowley, P. B.; Rabe, K. S.; Worrall, J. A. R.; Canters, G. W.; Ubbink, M. The ternary complex of cytochrome *f* and cytochrome *c*: Identification of a second binding site and competition for plastocyanin binding. *ChemBioChem* **2002**, *3*, 526–533.
305. Crowley, P. B.; Carrondo, M. A. The architecture of the binding site in redox protein complexes: Implications for fast dissociation. *Proteins* **2004**, *55*, 603–612.



## Acknowledgements

---

My journey over the last four years has been a huge learning experience in all aspects. PhD has not only helped in inculcating in me the disciplines of being a researcher/scientist but has helped me to grow as a person. A number of people and situations contributed towards making this a successful and an enjoyable (as much as it can be) endeavour.

Every PhD researcher is indebted to their supervisor for their successes and so am I. Big time! I would like to thank Prof. Peter B. Crowley for all his inputs, suggestions, and creative ideas to propel me on this journey. An excellent scientist, dedicated, passionate, zealous with everything that he hopes to achieve. These infectious 'traits' motivated me to perform at my level best (I choose to believe that). I would like to thank him for his patience (sometime endurance) for my mistakes, encouragement for various tasks, and for believing in me when I did not. I cannot thank him enough for the projects and opportunities he had given me, especially the PAF and foldamer project. Last but not the least, I thank him for giving me the opportunity to come here and pursue research in his lab otherwise there wouldn't have been a trip to Ireland or to other countries.

I am grateful to our collaborators – Prof. Alessandro Casnati (University of Parma, Italy), Prof. Gyula Batta (University of Debrecen, Hungary) and Prof. Ivan Huc (LMU, Germany) – for entrusting us with their ideas / projects. I especially thank Dr. Corvaglia (Huc Lab) for helpful discussions in various experiments.

I would like thank the Hardiman Research Scholarship for enabling me to pursue my studies and funding my PhD research.

I owe my crystallography experiences to Martin. You were a motivation and an inspiration to learn more about the various softwares. Thank you! Sylvain, similarly I am thankful to you for teaching the other new crystallographic softwares and all your help and encouragement.

I thank Dr. Patrick McArdle, NUI Galway for his help and suggestions with Shelx refinement.

I would like to thank all the technical staff especially Dr. Roisín Doohan for the help with NMR training and data collection.

I thank Aishling, Ciara, Gosia and Srinu for their moral support and encouragement at numerous times during my time here.

I thank Kiefer, Francesca, Pawel, Tomasz for all the wonderful times and laughter. Will miss you guys!

I would like to thank my father and amma for allowing me to come to Ireland to pursue my higher studies when most Indian parents would just want to get their daughters married off. Thanks for all your prayers, support, encouragement, and patience. I would like to thank my brother who would leave his cubicle to attend my calls – every single time – to hear me rant for the thousandth time about random lab stuff. I would also like to thank all the other people who have held me in their prayers.

Lastly, I thank my Father Lord, for holding it all together for me – all my success and all my failures. Thank you Father for being faithful to me even when I wasn't. Thank you for your grace, which unfolded to hold me every minute of the last four years. Thank you for your presence, love, peace, and joy in my life.



## Curriculum Vitae

---

### Education

- 2015 - 2019    PhD with Prof. Peter Crowley at NUI Galway  
Thesis Title: "Supramolecular 'glues' for protein recognition and assembly"  
**Hardiman Research Scholarship**
- 2011 - 2013    MPharm with Dr. Raj Kumar at Central University of Punjab, India  
Thesis Title: "Anticancer potential of *N*-acetyl pyrazoline derivatives"  
**Gold Medal**
- 2006 - 2011    BPharm at Pushpagiri College of Pharmacy, Kerala  
Thesis Title: "Nanochromatography and nanocapillary electrophoresis"

### PhD Publications

1. **J. M. Alex**, G. Brancatelli, S. Volpi, C. Bonaccorso, A. Casnati, S. Geremia, P.B. Crowley  
Probing the determinants of porosity in protein frameworks: Co-crystals of cytochrome *c* and an octa-anionic calix[4]arene.  
*Org. Biomol. Chem.* **2019**. (Accepted)
2. **J. M. Alex**, P. McArdle, P. B. Crowley.  
Supramolecular Stacking in a high *Z'* a calix[8]arene – porphyrin assembly  
*CrystEngComm.* **2019**. *In print*.
3. **J. M. Alex**, V. Corvaglia, X. Hu, S. Engilberge, I. Huc, P. B. Crowley  
Crystal structure of a protein – aromatic foldamer composite: macromolecular chiral resolution.  
*Chem Commun.* **2019**, *55*, 11087-11090.
4. **J. M. Alex**, M. L. Rennie, S. Engilberge, G. Lehoczki, H. Dorottya, Á. Fizil, G. Batta, P. B. Crowley  
Calixarene-mediated assembly of a small antifungal protein.  
*IUCrJ.* **2019**, *6*, 238-247.
5. **J. M. Alex**, M. L. Rennie, S. Volpi, A. Casnati, P. B. Crowley  
Phosphonated calixarene as a "molecular glue" for protein crystallization.  
*Cryst. Growth Des.* **2018**, *18*, 2467-2473.

## PDB Depositions

| PDB id | Protein                               | Ligand                                   |
|--------|---------------------------------------|--|
| 5ncv   | Cytochrome <i>c</i>                   | phosphonomethyl-calix[4]arene            |
| 6suy   | Cytochrome <i>c</i>                   | tetracarboxylate sulfonato-calix[4]arene |
| 6s8y   | Cytochrome <i>c</i>                   | sulfonated quinoline foldamer            |
| 6ha4   | <i>Penicillium</i> antifungal protein | sulfonatocalix[4]arene                   |
| 6hah   | <i>Penicillium</i> antifungal protein | sulfonatocalix[6]arene                   |
| 6haj   | <i>Penicillium</i> antifungal protein | sulfonatocalix[8]arene                   |

## CCDC Deposition

| CCDC#   | Complex   |
|---------|---|
| 1956108 | sulfonatocalix[8]arene – trimethylanilinium porphyrin |
| 1956128 |   |

## Oral Presentations

- 1. Eli Lilly (Third Prize)**  
Title: Supramolecular glues for protein recognition and assembly  
Venue: Aras Moyola, NUI Galway.  
Date: 21<sup>st</sup> February 2019
- 2. Supramol Workshop**  
Title: Supramolecular glues for protein recognition and assembly  
Venue: Orbsen Building, NUI Galway.  
Date: 7<sup>th</sup> Feb 2019
- 3. Workshop Protein-ligand interaction in biocrystallography**  
Title: Complexes of cytochrome *c* with anionic calixarenes and foldamers  
Venue: Laboratoire de Chimie – ENS de Lyon  
Date: 11<sup>th</sup> December 2019
- 4. Science Foundation Ireland (SFI) Review**  
Title: Calixarenes – probably the best glue in the protein world  
Venue: Orbsen Building, NUI Galway  
Date: 9<sup>th</sup> June 2017
- 5. BCA-CCP4 Protein X-ray Crystallography Summer School**  
Title: Calixarene-based ligands for protein recognition and assembly  
Venue: Diamond Light House, Oxfordshire.  
Date: 21 – 26 August 2017
- 6. 6th NUIG-UL Alliance Postgraduate Research Day (Thesis in Three)**  
Title: Calixarene-based ligands for protein recognition and assembly  
Venue: University of Limerick  
Date: 29<sup>th</sup> April 2016

## Poster Presentations

1. **Supr@Lyon Conference**  
Title: Calixarene-mediated assembly of small antifungal protein  
Venue: Claude Bernard University Lyon 1, France.  
Date: 12 – 14 December 2018
2. **19th International Union of Pure and Applied Biophysics (IUPAB) and 11th European Biophysical Societies' Association (EBSA) Congress 2017**  
Title: Calixarene “glue” facilitates crystal structure determination of *penicillium antifungal* protein  
Venue: Edinburgh International Convention Centre, Edinburgh, Scotland  
Date: 16 – 20 July 2017
3. **The 69<sup>th</sup> Irish University Research Colloquium 2017**  
Title: “Molecular glues” for assembly and structure determination of *penicillium antifungal* protein  
Venue: Dublin City University, School of Chemical Sciences  
Date: 22 – 23 June 2017
4. **6<sup>th</sup> NUIG-UL Alliance Postgraduate Research Day 2016**  
Title: Calixarene-based ligands for protein recognition and assembly  
Venue: University of Limerick  
Date: 29 April 2016

## MPharm Publications

1. B. S. Gill, **J. M. Alex**, Navgeet, S. Thakur  
Missing link between microRNA and prostate cancer.  
*Tumour Biol.* **2016**, *37*, 5683-5704.
2. M. Chauhan, A. Rana, **J. M. Alex**, A. Negi, S. Singh, R. Kumar  
Design, Microwave-mediated synthesis and biological evaluation of novel 4-aryl(alkyl)amino-3-nitroquinoline and 2,4-Diaryl(dialkyl)amino-3-nitroquinolines as anticancer agents.  
*Bioorg. Chem.* **2015**, *58*, 1-10.
3. A. Negi, **J. M. Alex**, S. Singh, A. T. Bhaviskar, U. C. Banerjee, R. Kumar  
Imine/amide–imidazole conjugates derived from 5-amino-4-cyano-N1-substituted benzyl imidazole: Microwave-assisted synthesis and anticancer activity via selective topoisomerase-II-a inhibition.  
*Bioorg. Med. Chem.* **2015**, *23*, 5654-5661.
4. M. Garg, M. Chauhan, P. K. Singh, **J. M. Alex**, R. Kumar  
Pyrazoloquinazolines: synthetic strategies and bioactivities.  
*Eur. J Med. Chem.* **2015**, *97*, 444-461.

5. A. Rana, **J. M. Alex**, M. Chauhan, G. Joshi, R. Kumar  
A review on pharmacophoric designs of antiproliferative agents.  
*Med. Chem. Res.* **2015**, *24*, 903-920.
6. **J. M. Alex**, S. Singh and R. Kumar  
1-Acetyl-3,5-diaryl-4,5-dihydro(1H)pyrazoles: Exhibiting anticancer activity through intracellular ROS scavenging and the mitochondria-dependent death pathway.  
*Archive der Pharmazie Chem. Life Sci.* **2014**, *347*, 717-727.
7. P. Singh, **J. M. Alex** and F. Bast  
Insulin receptor (IR) and insulin-like growth factor receptor 1(IGF-1R) signaling systems: novel treatment strategies for cancer.  
*Med. Oncol.* **2014**, *31*, e805.
8. **J. M. Alex** and R. Kumar  
4, 5-Dihydro-1H-pyrazole: an indispensable scaffold.  
*J Enzyme Inhib. Med. Chem.* **2013**, *29*, 427-442.



## **Appendix**

---

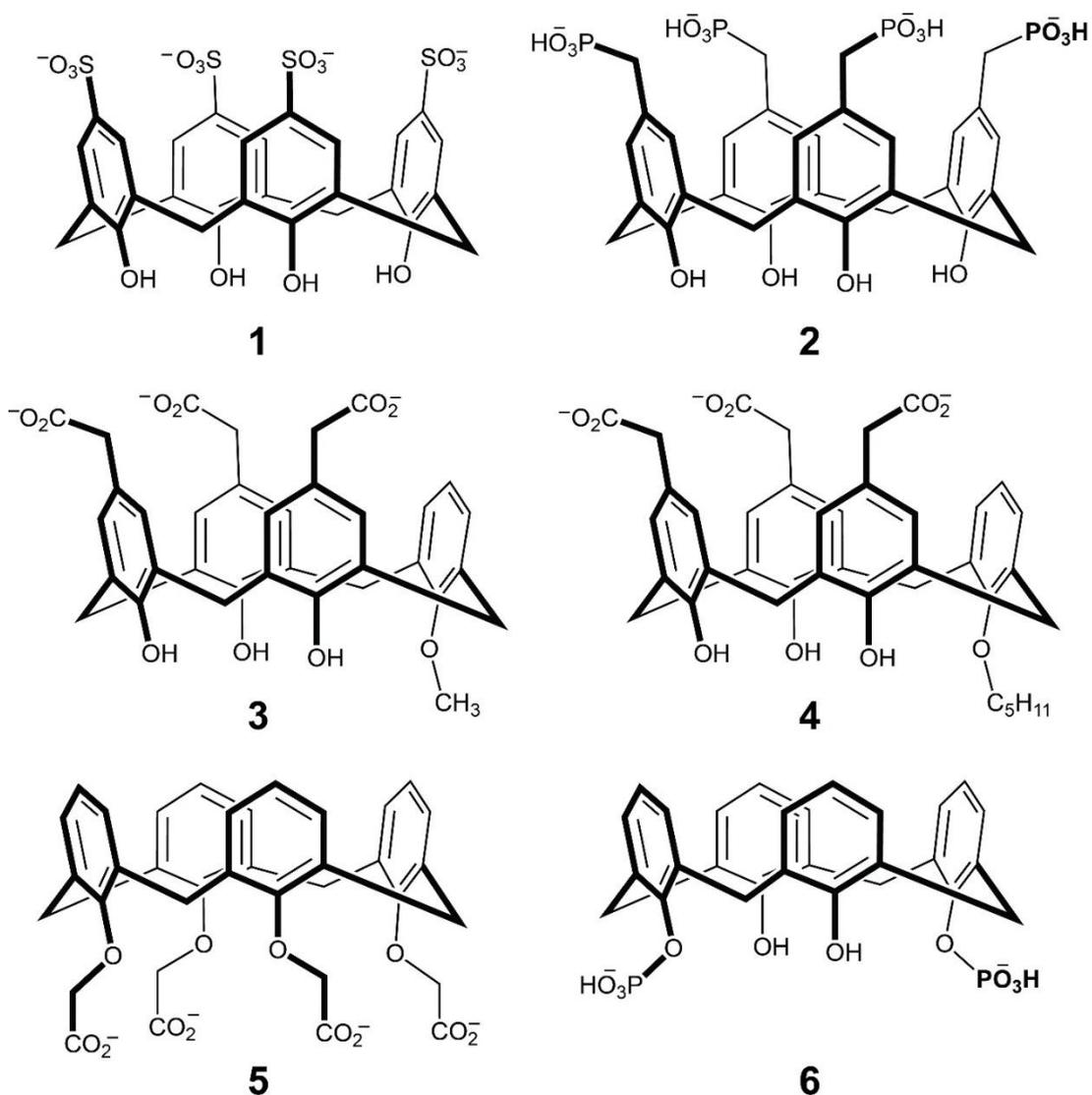
**Table A1.** Commercial source of the ligands used as 'molecular glues'

| # | Ligand  | Commercial Source                  |
|---|---|------------------------------------|
| 1 | Phosphonomethyl calix[4]arene ( <b>pmclx<sub>4</sub></b> )  | Molecular dimensions (CALIXAR kit) |
| 2 | Sulfonatocalix[4]arene ( <b>sclx<sub>4</sub></b> )          | TCI chemicals                      |
| 3 | Sulfonatocalix[6]arene ( <b>sclx<sub>6</sub></b> )          | TCI chemicals                      |
| 4 | Sulfonatocalix[8]arene ( <b>sclx<sub>8</sub></b> )          | TCI chemicals                      |
| 5 | tetra(4-N,N,N-trimethylanilinium) porphyrin ( <b>tmap</b> ) | Frontier Scientific                |

Note:

**pmclx<sub>4</sub>** (Chapter 1) and **sclx<sub>4</sub>mc** (Chapter 2): synthesized in Prof. Casnati's Lab (Università degli Studi di Parma, Italy).

Foldamers: synthesized in Prof. Huc's Lab (Ludwig-Maximilians-Universität, Germany; Chapter 5).



**Figure A1. Calixar Kit (Molecular Dimensions).** All compounds in the calixar kit were tested for their crystallization propensity with cytc and lysozyme. Compound **2** is **sclx<sub>4</sub>** which has been reported previously.<sup>39</sup> Only compound **1** (**pmclx<sub>4</sub>**) yielded co-crystals with cytc (Chapter 1).<sup>44</sup> Compound **5** is similar to **sclx<sub>4</sub>mc** (without upper rim sulfonates; Chapter 2) but did not yield crystals with either proteins.

Preliminary protein NMR experiments were performed as an initial test to investigate binding of azo dyes (orange G and chromotropic acid), pyrene, porphyrin, and suramin to the cationic cytc (cytc; Table A2).

Protein: N<sup>15</sup>-labelled cytc; Stock concentration: 1.0 mM

Sample Preparation:

400  $\mu$ L of NMR Buffer (50mM KPi + 20mM NaCl pH 6.0)

50  $\mu$ L of cytc (1.0 mM stock)

50  $\mu$ L of D<sub>2</sub>O

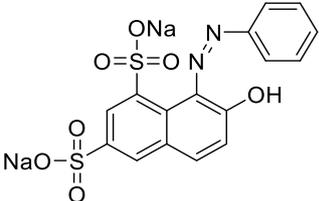
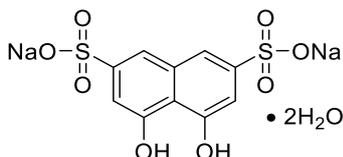
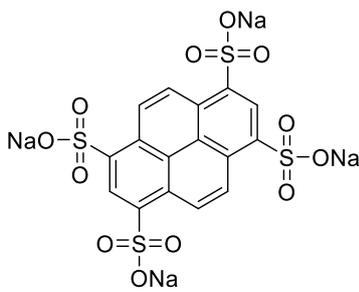
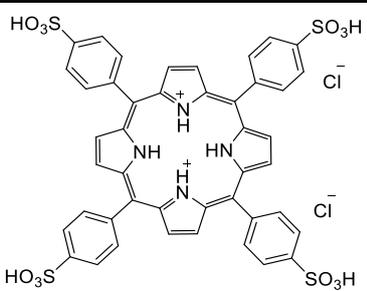
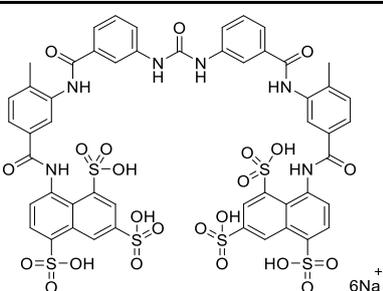
0.5  $\mu$ L of ascorbic acid (1M stock)

Final protein concentration: 0.1mM

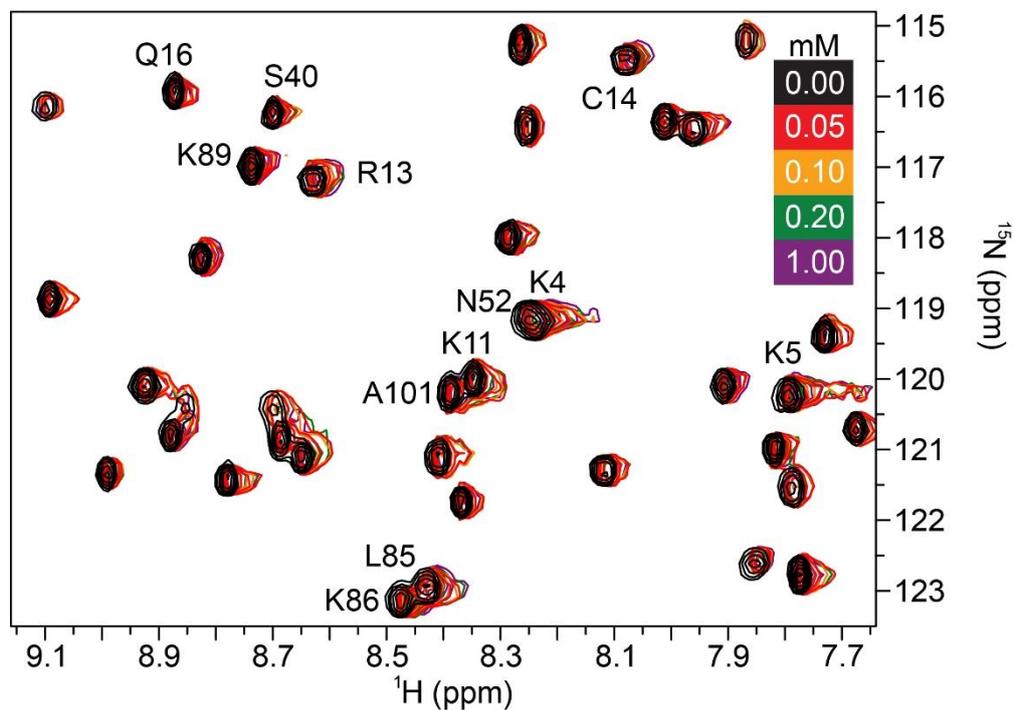
### **NMR spectroscopy**

<sup>1</sup>H-<sup>15</sup>N HSQC monitored titrations were performed at 30°C on a 600 MHz Varian spectrometer equipped with a HCN coldprobe, as described previously.<sup>39</sup> Samples of 0.1 mM <sup>15</sup>N-labelled cytc were titrated with required aliquots of ligand stock in water) in 20 mM potassium phosphate, 50 mM sodium chloride at pH 6.0. Spectra were processed in NMRPipe and chemical shift perturbations were determined in CCPN.<sup>303</sup>

**Table A2.** Molecular structures of the dyes used in the study

| Ligands  | Structure   | Molecular Weight (Da) | [Stock]; pH |
|--|---|-----------------------|-------------|
| Orange G   |    | 452.37                | 50mM; 5.96  |
| Chromotropic acid disodium salt dihydrate              |    | 400.29                | 50mM; 5.99  |
| 1,3,6,8-Pyrene tetrasulfonic acid tetrasodium salt     |   | 610.43                | 50mM; 6.00  |
| Meso-tetra(4-sulfonatophenyl) porphine dihydrochloride |  | 1007.92               | 10mM; 6.00  |
| Suramin sodium salt                                    |  | 1429.17               | 50mM; 5.97  |

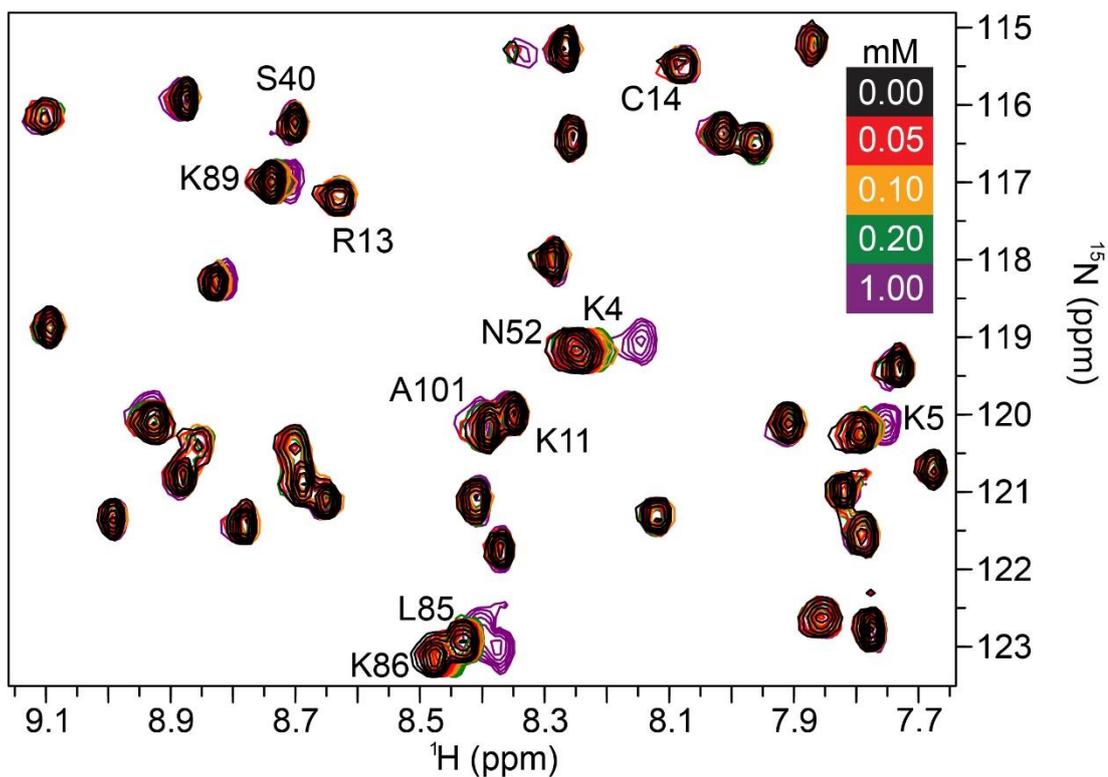
Cytc – orange G



**Figure A2.** Spectral region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 0.1 mM  $^{15}\text{N}$  labelled cytc in the presence of 0.05–1.0 mM orangeG (coloured scale).

In solution, orange G appears to interact weakly with cytc as evident from the overlaid spectra. Note the slight broadening of the K4 and K5 resonances at 1.0 mM (10 equiv.). Co-crystallization trials of the complex yielded soluble drops.

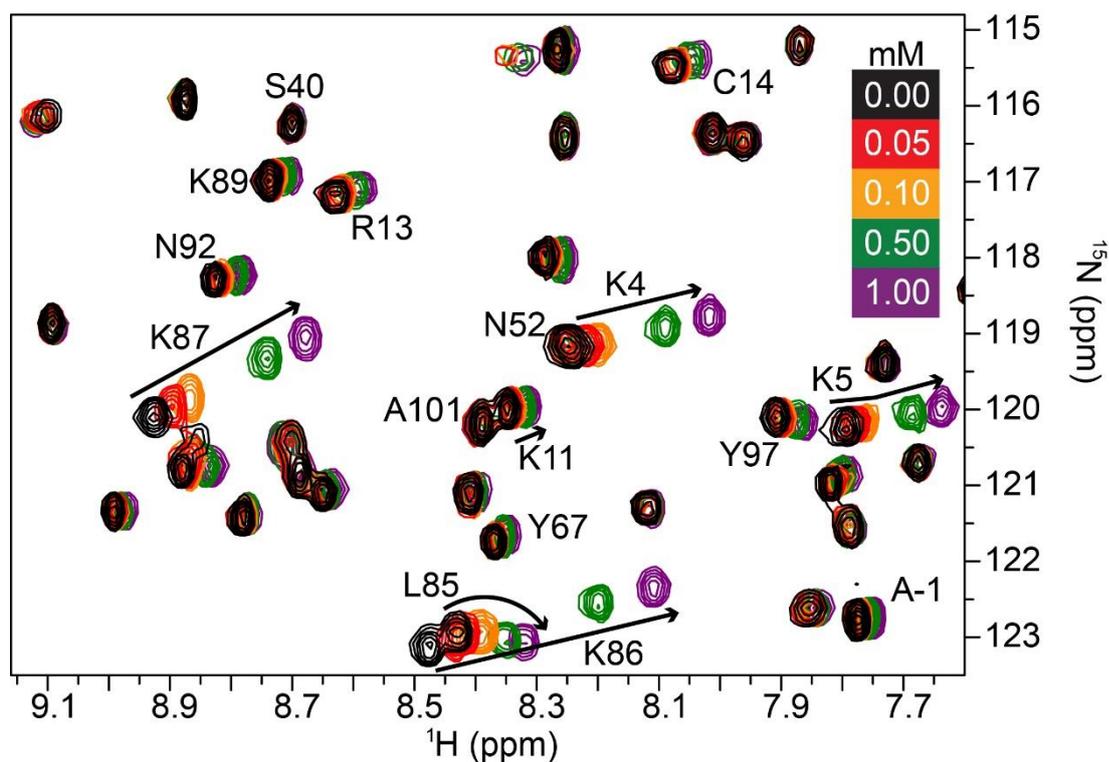
## Cytc – chromotropic acid disodium salt dihydrate



**Figure A3.** Spectral region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 0.1 mM  $^{15}\text{N}$  labelled cytc in the presence of 0.05–1.0 mM chromotropic acid (coloured scale).

The chemical shift perturbations observed here are similar to that observed in the presence of orange G including the perturbations at K4 and K5 at 10 equiv. of the ligand (1.0 mM). The co-crystallization drops remained clear.

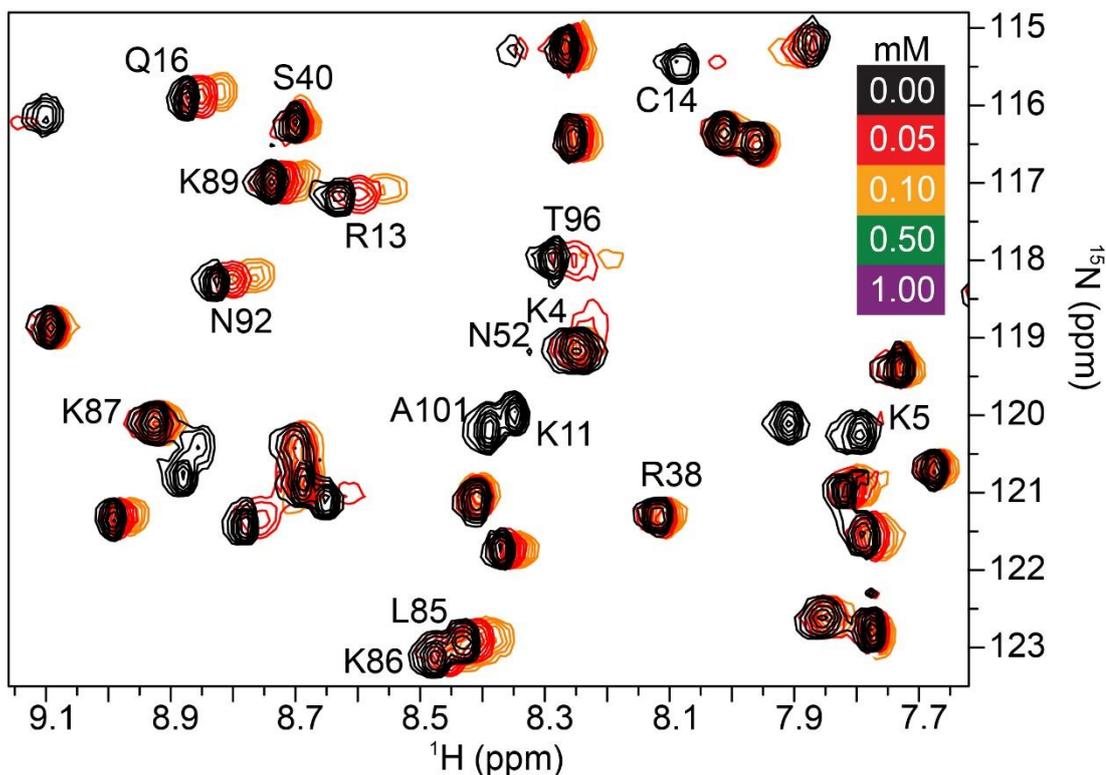
**Cytc – 1,3,6,8-pyrene tetrasulfonic acid tetrasodium salt**



**Figure A4.** Spectral region from overlaid <sup>1</sup>H-<sup>15</sup>N HSQC spectra of 0.1 mM <sup>15</sup>N labelled cytc in the presence of 0.05–1.0 mM pyrene tetrasulfonate (coloured scale).

In contrast to the significant chemical shift perturbation ( $\Delta\delta \text{ } ^1\text{H}^{\text{N}} \geq 0.04$  and/or  $^{15}\text{N} \geq 0.4$ ) observed here especially around 0.5 mM (5 equiv.) of the ligand, co-crystallization trials at 5, 10, 15 and 25 equiv. of the compound yielded soluble or phase-separated drops. This difference probably suggests a different mode of interaction in the solution-state as compared to the solid state.

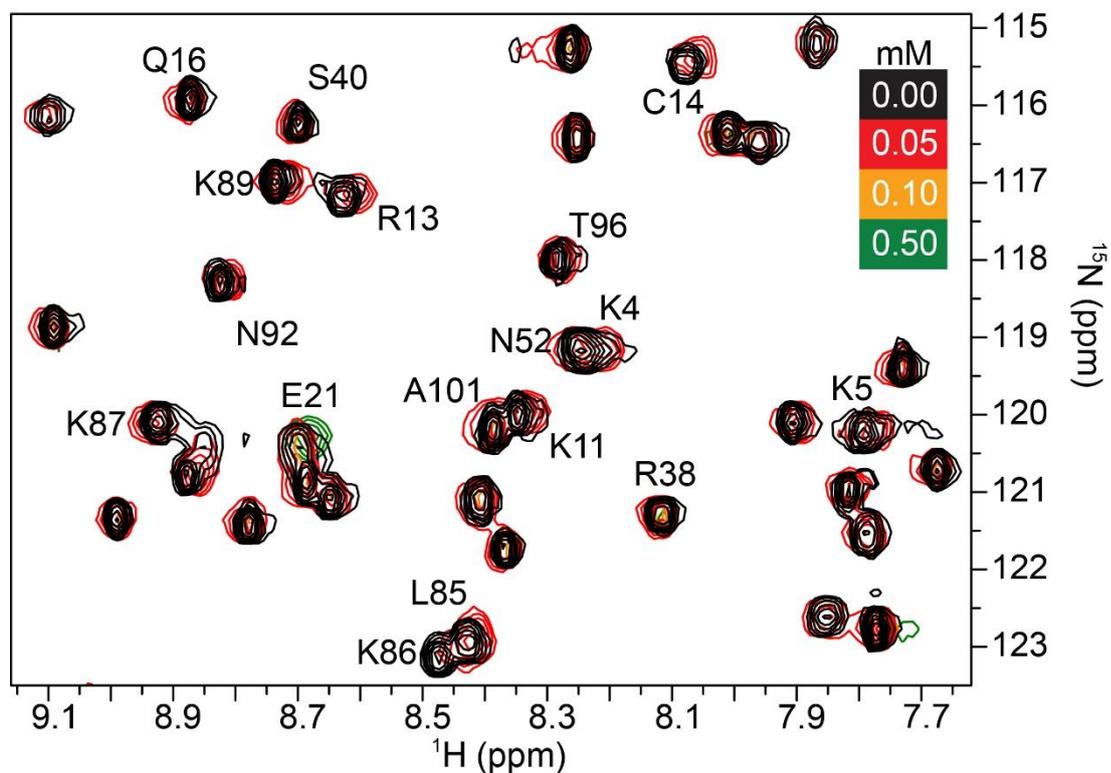
## Cytc – meso-tetra(4-sulfonatophenyl) porphine dihydrochloride (TPPS)



**Figure A5.** Spectral region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 0.1 mM  $^{15}\text{N}$  labelled cytc in the presence of 0.05–1.0 mM TPPS (coloured scale).

Previously, formation of multimeric structures were observed in SAXS studies of cytc – TPPS complex.<sup>242</sup> Also, in these preliminary protein NMR studies the resonances for most of the residues are broadened away at 0.5 mM of the ligand probably indicating formation of higher ordered structures. However, no precipitation was observed in co-crystallization trials and the drops remained clear perhaps suggesting a weaker binding in the solid-state.

Cytc – suramin sodium salt



**Figure A6.** Spectral region from overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 0.1 mM  $^{15}\text{N}$  labelled cytc in the presence of 0.05–1.0 mM suramin (coloured scale).

Similar to the observations of cytc – TPPS complex, spectral broadening was observed in this complex as well at 0.1 mM or 1 equiv. of the ligand. In contrast, co-crystallization trials yielded clear/soluble drops. While the inherent conformational flexibility of suramin may have favoured broadening (indicative of higher assemblies), the floppiness probably impedes the interaction during crystallization.





