



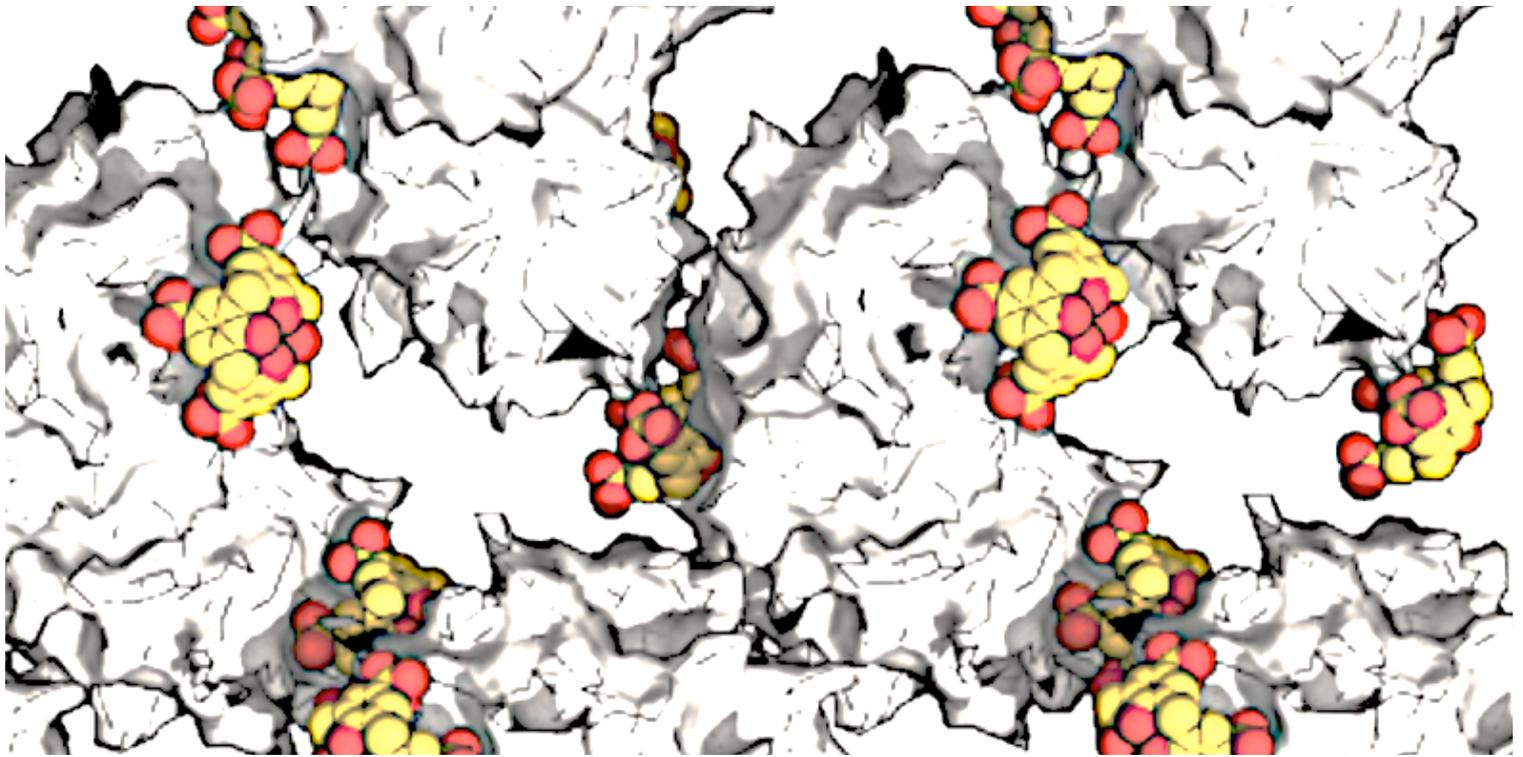
Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	Calixarene-mediated protein assembly
Author(s)	McGovern, Róise Ella
Publication Date	2014-05-23
Item record	http://hdl.handle.net/10379/4552

Downloaded 2024-04-19T06:08:13Z

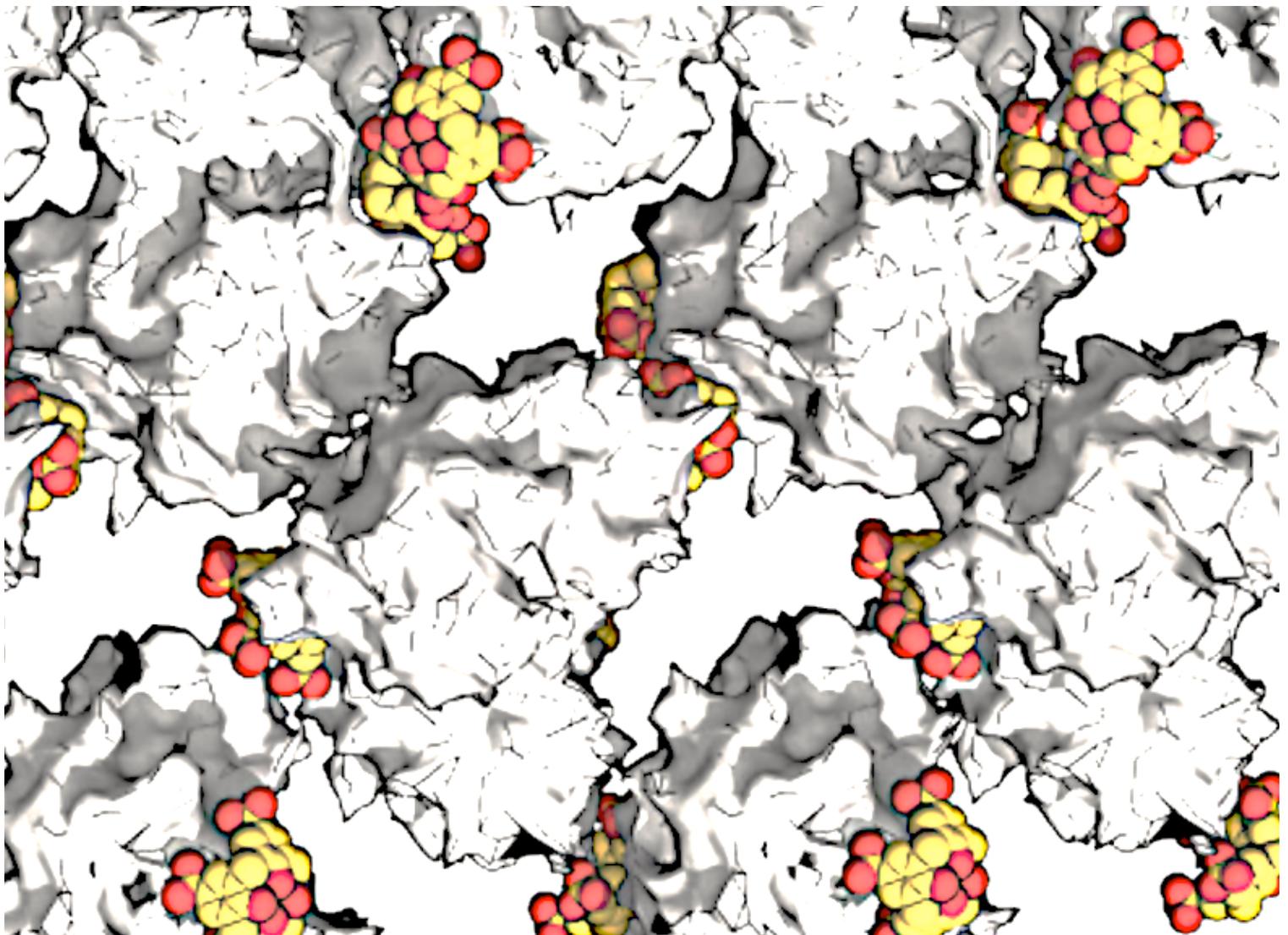
Some rights reserved. For more information, please see the item record link above.





Calixarene-Mediated Protein Assembly

Róise McGovern





Calixarene-Mediated Protein Assembly

Ph.D. Thesis

This thesis was prepared at the School of Chemistry, National University of
Ireland, Galway, from January 2010 to March 2014.

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: Impression of this thesis, calixarene-assisted protein assembly.

Róise Ella McGovern

Supervisor: Dr. Peter Crowley

External Examiner: Prof. Thomas Schrader

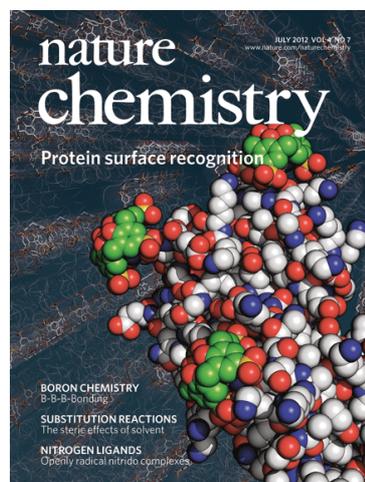
Internal Examiner: Prof. Corrado Santocanale

Chair: Prof. Paul Murphy

Viva voce: 13th May 2014

Summary of Contents

Solved the first crystal structure of a protein-calixarene complex (at 1.4 Å resolution). The results provide exact structural information on the amino acid preference of the anionic *p*-sulfonatocalix[4]arene (sclx₄) for the lysine-rich cytochrome *c*.¹ This work featured on the front cover of Nature Chemistry (July 2012 issue). The rich class of structurally well-defined calixarenes have been studied for several decades and some of their water-soluble derivatives are well known for their complexation ability of various proteins.² However, despite great efforts no crystal structure could be obtained until now. The detailed data analysis of an NMR titration of ¹⁵N-labelled cytochrome *c* with the calixarene revealed the presence of two or more binding sites, with mM affinities. The results from both techniques agree on the binding epitopes and the prediction of at least two binding sites on the protein surface. A cytochrome *c* mutant was also co-crystallised with sclx₄ for comparison and the large crystals were electrochemically characterised.³



In order to further substantiate the potential of sclx₄ as a protein surface binder, the crystal structures of the calixarene complexed to lysozyme in its native form and in a modified form were solved to 1.7 Å and 2.2 Å, respectively.⁴ In the native form, the calixarene generates a tetrameric protein assembly and exhibits a preference for arginine. The structure of the modified protein complexed with sclx₄ provides the first structural example of a synthetic molecule binding dimethyllysine. This interaction is significant as it resembles the naturally occurring aromatic protein cage, essential for gene expression. The results are important because the solved structures present valuable models for the structure-aided design of artificial receptors to target the protein surface.

References

- (1) McGovern *et al.* *Nat. Chem.* **2012**, 4, 527., (2) Perret *et al.* *Chem.Comm.* **2006**, 2425., (3) McGovern *et al.* (*Manuscript in preparation*) **2014**, (4) McGovern *et al.* (*Submitted*) **2014**

Contents

Introduction		1
Chapter 1	Co-crystallisation and X-ray Crystallography of Protein-Small Molecule Complexes	23
Chapter 2	The Complexes of Cytochrome <i>c</i> and <i>p</i> -sulfonatocalix[4]arene	41
Chapter 3	Electron Transfer in Cytochrome <i>c</i> -calixarene Crystals: An Electrochemical Study	63
Chapter 4	A Structural Study of <i>p</i> -sulfonatocalix[4]arene and Lysozyme Assemblies	81
Chapter 5	Crystal Structure of <i>p</i> -sulfonatocalix[4]arene and Lysozyme Bearing Dimethylated Lysines	97
Discussion		113
Bibliography		123

Introduction

Protein-protein interactions

This thesis focuses on protein surface binding by small molecules and its influence on protein self-assembly. Despite differences in the size and chemical composition of small molecules,¹ the molecular recognition features that influence protein-small molecule binding are similar to protein-protein complexation.²⁻⁴ Therefore, the interface features that govern protein-protein interactions are firstly addressed. The rationale and strategies for investigating protein surface binders that exploit these features are described.

Proteins are the primary molecular machines that work in concert to control the cell. Specific protein-protein interactions provide distinct biological functions as part of a vast array of cellular networks. For example, cytochrome *c* shuttles electrons from cytochrome reductase to cytochrome *c* oxidase resulting in the conversion of molecular oxygen to water, a crucial stage of the respiratory chain.⁵ Protein interactions with receptor tyrosine kinases mediate the signalling pathways that govern cell proliferation, differentiation and survival.⁶ Specific protein-protein interactions can also lead to disease. For example, the p53 protein, which functions in maintaining cellular homeostasis, is inhibited when it is bound by the MDM2 proto-oncoprotein.⁷ Subsequently, p53 induced apoptosis is prevented and cell proliferation proceeds, leading to tumour development.

Most proteins adopt a unique native conformation resulting in the decoration of their surfaces by various polar, charged, aromatic and aliphatic side chains. These diverse amalgamations of hydrophobic and hydrophilic surface patches are influential for protein-protein interactions.⁸⁻¹² For example, the composition of hydrophobic residues determines whether the complex will be transient or permanent and the arrangement of polar residues influence the specificity of the interaction.¹⁰ Some protein interactions involve binding to a defined cavity or groove, for example an α -helical recognition cleft (Figure 1A).¹³ However, the majority of protein-protein interactions occur via surface patches. The binding sites are primarily composed of a hydrophobic core and a hydrophilic rim.¹⁰ Transient protein complexes, such as redox protein complexes, involve a buried surface area of $\sim 500 \text{ \AA}^2$.^{9,14,15} The interfaces are flat, poorly packed and have low geometric complementarity, a feature that favours non-specific binding and fast dissociation. At the other extreme, permanent protein assemblies have much larger interfaces in the range of 2000-5000 \AA^2 (Figure 1B).¹⁶ The interfacial residues are closely packed, allowing the complexes to assemble more tightly resulting in high affinity binding. These proteins may undergo large conformational changes as a result of complexation. For example, CDK2

kinase undergoes domain and loop movements upon cyclin A complexation, resulting in the activation of the kinase.¹⁷ An α -helix belonging to CDK2 kinase rotates about its axis and contributes to the interaction interface, which has an area of 3400 \AA^2 . The “standard” buried surface areas for protein complexes is $\sim 1600 \text{ \AA}^2$, for which the interfacial residues are loosely packed and experience small conformational changes on binding. As the physicochemical composition of protein surfaces governs complex formation, these regions are important targets for manipulating protein complexes.

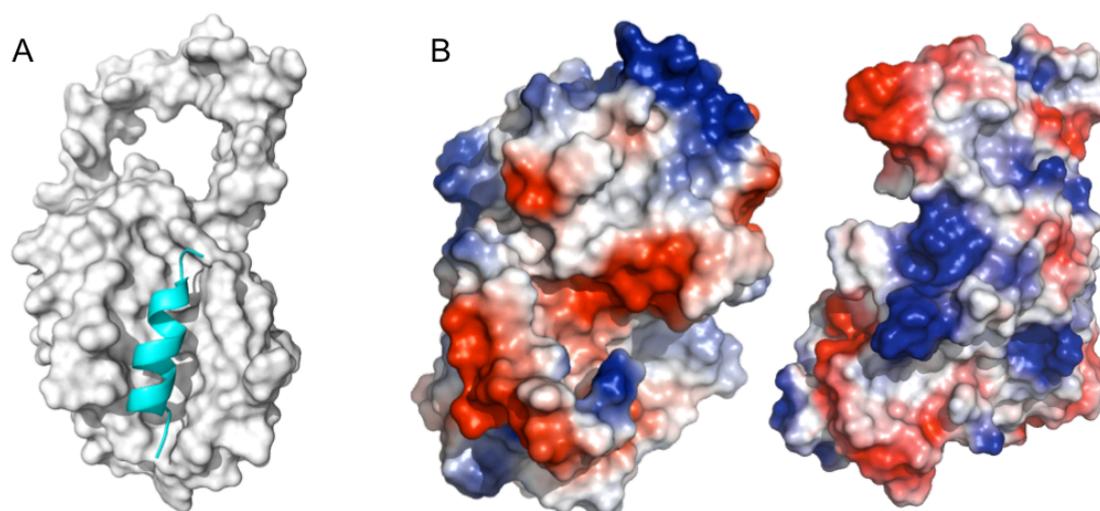


Figure 1. Examples of high affinity protein interfaces. **(A)** A surface representation of Bcl-xL protein in complex with Bak BH3 peptide (PDB ID 2lp8).¹⁸ The Bak BH3 peptide (cyan cartoon) binds at an α -helical recognition cleft. **(B)** The electrostatic surface representations of cyclin A fragment (left) and cyclin-dependent kinase (right) are shown (PDB ID 1fin).¹⁷ The complex of the cyclin A fragment with cyclin-dependent kinase was opened to reveal the large, continuous interfaces. The red, blue and white surfaces correspond to the anionic, cationic and uncharged residues, respectively.

Designed modulation of protein character

A balance between structure and noncovalent interactions determines protein behaviour and the global organisation of biomolecules.¹⁹ Modulating protein properties can lead to altered interaction behaviours. Covalent or noncovalent strategies are used to achieve this and the overall effects can influence the formation or inhibition of protein complexes. As a result, the designed modulation of protein properties has numerous applications in the fields of therapeutics,^{20,21} biosensors^{22,23} and nanomaterials.²⁴⁻²⁶ Protein solubility is a

good starting point for modification as it can be used to promote assembly or to transport and extract proteins. This will be discussed.

Covalent modification of proteins

Detailed studies of protein interfaces aid the understanding and prediction of molecular recognition events.^{27,28} This knowledge has facilitated the design and engineering of proteins that interact in a specific manner. For example, proteins were designed to self-assemble by introducing additional hydrophobic residues at the contact interface that would incur symmetry in the assembly.²⁶ Symmetry at the interface was important for the global structure design. The protein-protein interfaces were predominantly hydrophobic and the side-chain mobility was restricted to reduce the entropic cost of association and stabilise the assemblies. By using size exclusion chromatography and dynamic light scattering the designed global structures were observed. At near atomic resolution the crystal structures showed minute deviations ($< 3 \text{ \AA}$) from the designed template, suggesting that high precision is difficult to achieve.

The genetic fusion of protein subunits that self-assemble into well-ordered structures is another strategy developed to control and stabilise protein complexes.²⁹ The domains of trimeric bromoperoxidase protein and dimeric M1 virus matrix protein were rationally selected for such a study.³⁰ In the crystal structure the two protein domains, which were fused by a nine-residue helical linker, self-assembled as a three-dimensional, 12-subunit protein cage. This self-assembling cage displayed the designed tetrahedral geometry with minute deviations from perfect symmetry. Such structures have great potential as biomaterials with specific structural and functional properties.

In nature proteins are also covalently modified to alter interaction behaviours. For example, the posttranslational modification of histone proteins by arginine or lysine methylation alters the structure of chromatin and ultimately gene expression (Figure 2).³¹ Although lysine side chain methylation doesn't change the overall charge on the residue, it does increase the cationic propensity of the N^ϵ , promoting its ability to make cation- π interactions.³² Lysine methylation is routinely used as a rescue strategy to crystallise proteins that are otherwise intractable to crystallization.^{33,34} This modification can reduce protein solubility and, therefore, increase the likelihood of self-assembly which is essential for protein crystallization.

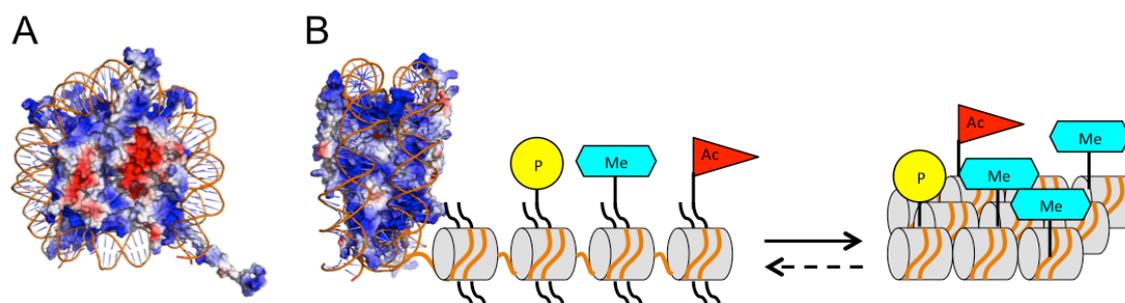


Figure 2. Posttranslational modifications of histones in nucleosome core particles alter interaction behaviour. **(A)** The nucleosome core particle (PDB ID 1aoi) consists of eight histone proteins (surface representation) around which a DNA superhelix is wrapped (ribbon traces for the 146-bp DNA phosphodiester backbones). Part of the histone tail peptide protruding out of the assembly is visible. **(B)** The nucleosome, rotated by 90°, is shown as part of a model of chromatin that condenses for chromosome formation. Engineered histone tail modifications are highlighted as methylated (Me), phosphorylated (P), and acetylated (Ac).

Cysteine is a good target for covalent modification owing to its strongly nucleophilic side chain sulfhydryl. Disulfide bonds occurring between two cysteine residues have been exploited to lock proteins into particular conformations and to contribute to protein stability.³⁵ Cysteine residues have also been used to cross-link proteins for crystallization as the larger protein complexes have been shown to crystallise more readily than their monomeric counterparts.³⁶⁻³⁸ A cysteine mutant of OmpF porin was covalently modified with a dibenzo-18-crown-6 derivative.³⁹ The crystal structure shows the cysteine-linked dibenzo-18-crown-6 derivative partly block the central pore, which otherwise serves as an ion conductance pathway.

Noncovalent modification of proteins by small molecules

Protein surface recognition by small molecules presents a valuable technique for altering protein properties and has great potential for therapeutic intervention.^{20,21,40} As a result, intense interest has been applied in the identification of synthetic molecules that target protein surfaces. X-ray analysis of co-crystals shows the locations for ligand binding on the target protein and the features of the binding sites that favour complexation. However, there is a paucity of high resolution crystal structures of the resulting complexes (Table 1). There are many different approaches used to design protein binders. Here, molecular mimicry is discussed as one such approach. This involves the design of ligands that mimic the interactions observed at specific protein-protein interfaces. Three approaches

used to design protein surface binders through molecular mimicry are described in context of the different ligands:

- Type 1 involves targeting recognition clefts (Figure 3).
- Type 2 involves binding small surface-displayed recognition sites (Figure 4).
- Type 3 involves binding large surface patches.

Table 1: Crystal structures of protein-small molecules complexes.

Ligand	Protein	Binding feature	PDB	Reference
Oligoamide foldamer	α -carbonic anhydrase	Active pocket and protein surface	4mtv	41
Cryptophane	α -carbonic anhydrase	Active pocket and protein surface	3cyu	42
Molecular tweezers	14-3-3 α	Lysine side chain	4hqw	43
Cucurbit[7]uril	Insulin	N-terminal phenylalanine	3q6e	44
Cyclam	Lysozyme	Tryptophan and aspartate	1yik	45
Xylyl-bicyclam	Lysozyme	Tryptophan and aspartate	2h9k	46
18-crown-6	RbmA	Lysine	4bei	47
Sclx ₄	Cytochrome <i>c</i>	Lysine	3tyi	48 Chapter 2
Sclx ₄	Cytochrome <i>c</i> R13E	Lysine	4n0k	Chapter 2
Sclx ₄	Lysozyme	Arginine	4prq	Chapter 3
Sclx ₄	Lysozyme-KMe ₂	Dimethyl lysine and arginine	4pru, 4noj	Chapter 4
Cavitein	Peptide	Covalent complex	-	49
Tetra(4-sulfonatophenyl)-porphyrin	Concanavalin A	Hydrophobic and polar residues	1jn2	50
Tetra(4-sulfonatophenyl)-porphyrin	Agglutinin	Polar residues Hydrophobic patches and/or polar residues	1pxd, 1rir, 1rit	51 52
Pyrene tetrasulfonic acid	Pyruvate kinase	Arginine	3is4	40
Suramin	Nucleocapsid protein	Lysine and arginine	4j4v	53

Type 1: Targeting recognition clefts using Foldamers

Certain protein folds present recognition clefts amenable to rational ligand design. For example, Bcl-xL has an α -helical recognition cleft on its surface that the α -helix of the Bak protein BH3 domain binds (Figure 1A).¹³ The interaction results in the inhibition of apoptosis and, therefore, presents an important target for therapeutic intervention. Based on this interaction site, an oligoamide foldamer was rationally designed to specifically bind the cleft.⁵⁴ The α -helix mimic possessed a rigid, preorganised structure containing functional groups that would promote protein interactions. The mimic was shown to interfere with Bcl-xL complexation by BH3, thereby suggesting a potential route for rational drug design.

α -carbonic anhydrase II is an important enzyme for drug intervention as its ability to interconvert CO_2 and HCO_3^- is crucial to a variety of physiological processes.⁵⁵ The enzyme's active site is located in a deep binding pocket that contains a catalytically essential Zn^{2+} coordinated by three histidine residues and is a target for inhibitor design. The affinity of helical aromatic oligoamide foldamers for α -carbonic anhydrase II have been investigated by circular dichroism.⁴¹ The foldamers were designed to bind deep inside the binding pocket via an N-terminal sulfonamide group that coordinates the Zn^{2+} . A crystal structure shows how the N-terminus portion of one particular foldamer binds in such a manner (Figure 3). The remainder of the foldamer protrudes from the protein surface, adopts a right-handed conformation and assembles with the second foldamer-protein complex. Foldamer-foldamer stacking contributes to stabilisation of dimer. Similarly, the crystal structure for α -carbonic anhydrase II complexed by ^{129}Xe -cryptophane shows the benzenesulfonamide tail of the ligand binding the active site.⁴² The bulky cryptophane moiety, which has a ^{129}Xe encapsulated for Nuclear Magnetic Spectroscopy (NMR) detection, is present nestled at the upper edge of the cavity mediating protein-protein interfaces.

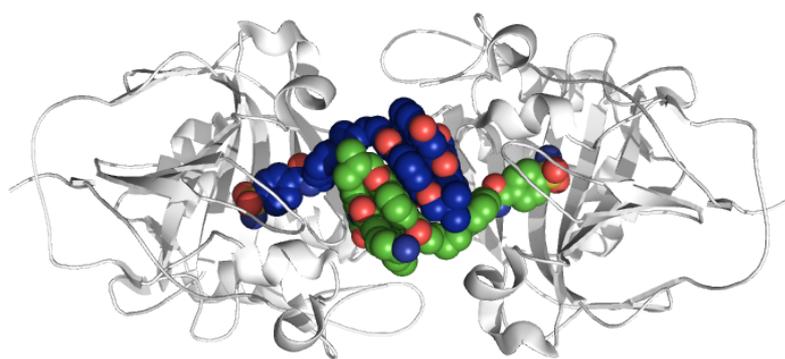


Figure 3. The crystal structure of the α -carbonic anhydrase-foldamer complex.⁴¹ One portion of the ligand binds the binding pocket and the remainder projects out and binds another foldamer-protein complex. The protein chains are grey cartoons and the foldamers are green or blue spheres.

Type 2: Binding small surface-displayed recognition sites

Ligands that exclusively target small binding sites in the form of solvent exposed residues also have great potential as interaction inhibitors. Arginine and lysine are particular good targets as are typically displayed on protein surfaces.^{56,57} They have long alkyl chains that can make hydrophobic interactions and positively charged functional groups that can hydrogen bond, make charge-charge and cation- π interactions. These features distinguish both cationic residues from other residues (Figure 4), making them available for rational ligand design. Furthermore, owing to their role at protein-protein interfaces^{58,59} and their susceptibility to posttranslational modification³¹ these residues are important targets for therapeutic intervention.

Type 2: Binding small surface-displayed recognition sites using Molecular Tweezers

A molecular tweezers having aromatic side walls and two peripheral anionic phosphonate groups was shown to specifically bind cationic amino acids.⁶⁰ The binding event involves threading a lysine or arginine side chain through the ligand cavity. A combination of van der Waals and electrostatic interactions are observed between the cationic residues and the molecular tweezers. Developing on this observation, the ligand was rationally applied to a system involving the 14-3-3 α adapter protein.⁴³ The protein-molecular tweezers complex inhibited 14-3-3 α protein interactions with two partner proteins, C-Raf and ExoS. Furthermore, the molecular tweezers was shown to specifically bind one lysine side chain on the protein by X-ray crystallography (Figure 4).

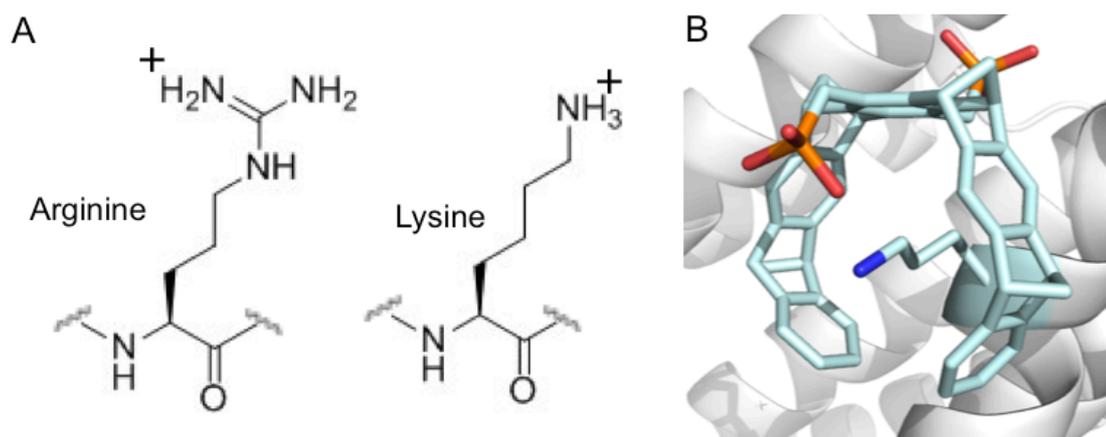


Figure 4. Small surface displayed targets for protein surface recognition. (A) Arginine and lysine residues. (B) The bisphosphonate molecular tweezers (illustrated) specifically binds one lysine on the 14-3-3 α adapter protein.⁴³

Type 2: Binding small surface-displayed recognition sites using Cucurbit[n]urils

Cucurbit[n]urils, the cyclic macrocycles possessing hydrophilic rims and hydrophobic cavities, can also bind small recognition sites. Isothermal titration calorimetry (ITC) showed cucurbit[7]uril bind insulin in a 1:1 stoichiometric ratio.⁴⁴ The N-terminal phenylalanine, as well as two phenylalanines and four tyrosine residues displayed on the protein surface are suitable binding sites. However, ITC measurements showed no measurable binding of cucurbit[7]uril to the insulin variant lacking N-terminal phenylalanine. The crystal structure of the cucurbit[7]uril-insulin complex shows the N-terminus unfolded and the ligand incorporating the N-terminal phenylalanine residue in its cavity. Furthermore, a two adjacent cucurbit[7]urils were observed at protein-protein interfaces. Cucurbit[7]uril also binds lysine and arginine amino acids and shows enhanced selectivity for their methylated counterparts, in particular trimethylated lysine (LysMe₃).⁶¹ Cucurbit[8]uril induces the self-assembly of caspase-9 by binding two N-terminal phenylalanines in its cavity.⁶² On cucurbit[8]uril-assisted homodimer formation, caspase-9 is enzymatically active and can be switched off by the addition of a PheGlyGly peptide, which competes for cucurbit[8]uril binding. Therefore, the macrocycle can be used as a molecular switch to control this enzyme's activity.

Type 2: Binding small surface-displayed recognition sites using Cyclams

Cyclams are fourteen-membered tetraamine macrocycles known for their ability to complex various cations including transition metals.⁶³ Xylyl-bicyclam, which consists of

two cyclam rings connected by phenylenebismethylene linker, is a potent inhibitor of human immunodeficiency virus (HIV) replication.⁶⁴ This anti-HIV activity is thought to involve the cyclams binding the carboxylate groups of aspartate and glutamate. The crystal structures of the Cu-cyclam and Cu₂-xylyl-bicyclam with lysozyme, show a portion of the cyclam ring between the indole side chains of two tryptophan residues.⁴⁵ For the Cu-cyclam complex, additional interactions involve hydrogen bonds between two cyclam NH's and the side chain of aspartate. For the Cu₂-xylyl-bicyclam complex additional interactions occur between the aspartate side chain and Cu.

Type 2: Binding small surface-displayed recognition sites using Crown Ethers

Another class of small molecules that target specific solvent exposed residues and modulate protein properties is the crown ethers.^{65,66} For example, 18-crown-6 transports the water-soluble cytochrome *c* into organic solvents.⁶⁷ Four 18-crown-6 molecules can simultaneously bind the protein through lysine residues.⁶⁸ Complexation of cytochrome *c* by 18-crown-6 in methanol allows its chemical activation at -40 °C.⁶⁹ Dicyclohexano-18-crown-6 also binds cytochrome *c* and transfers it into an organic, poly(ethylene glycol)-rich phase.⁷⁰ The crown ether derivative displays a preference for cytochrome *c* over the other cationic proteins that display significantly fewer lysines. This residue preference and resultant altered solubility property enabled cytochrome *c* separation from a mixture of proteins of a similar size and pI. 18-crown-6 co-crystallizes the RbmA protein almost immediately after experiment preparation.⁴⁷ The crystal structure of RbmA-ligand complex shows the 18-crown-6 at a protein-protein interface (Figure 5). The primary interaction involves the incorporation a lysine side chain in the central pore of the 18-crown-6 and this is conducive to protein self-assembly.

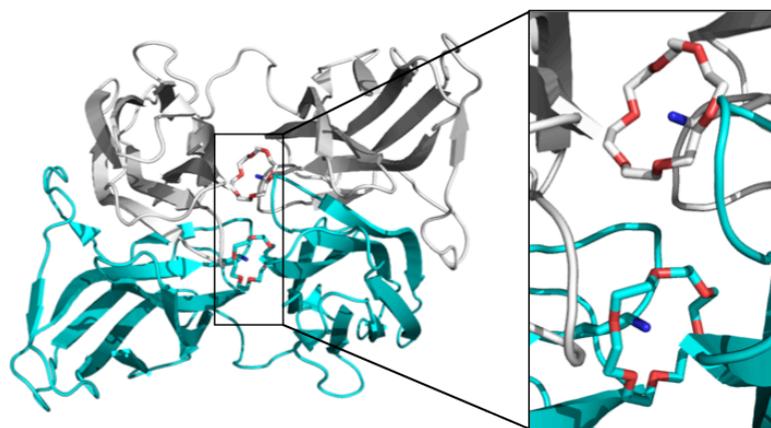


Figure 5: Crystal structure of RbmA with 18-crown-6.⁴⁷ Two 18-crown-6 molecules, nestled at the centre of the protein assembly, each entrap a lysine side chain. Chain A and B are represented as cartoons and coloured grey and cyan, respectively. The lysines and ligands are sticks.

Type 2: Binding small surface-displayed recognition sites using Calixarenes

Synthetic agents that present a scaffold amenable to modification can be tuned to complement small surface-displayed recognition sites. For example, the calix[n]arenes are cup-shaped molecules belonging to a family of macrocyclic compounds.⁷¹ The upper and lower rims of the calixarene are amenable to countless modifications and the cavity can incorporate guest molecules of appropriate size.⁷²⁻⁷⁷ A survey of the Cambridge Crystallographic Data Centre shows a wealth of structures involving the water soluble *p*-sulfonatocalix[n]arene (sclx_n). Significantly, these structures include sclx_n complexed with cationic functional groups, ammonium and guanidine (Figure 6). These complexes suggest that sclx_n is suitable for protein surface binding via lysine and arginine residues.

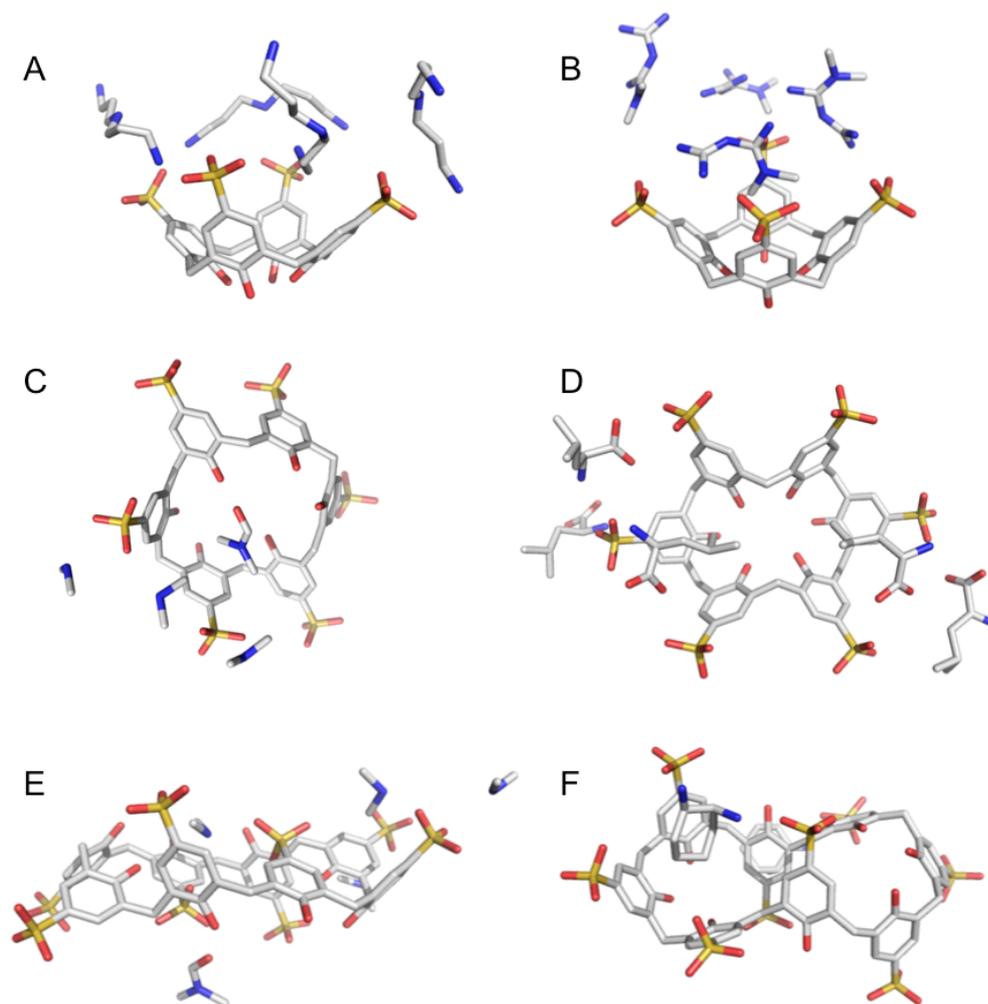


Figure 6. Crystal structures of sclx_n complexed with a variety of ammonium and guanidine cations. Sclx₄ with (A) norspermidine and (B) metformin.^{78,79} Sclx₆ with (C) dimethylammonium and dimethylformamide and (D) leucine.^{80,81} Sclx₈ with (E) dimethylammonium and dimethylformamide and (F) 1,2-cis-cyclohexanediammonium.^{82,83}

Sclx₄ is the most widely studied compound in the calixarene family.⁸⁴ Since the structure solution of the sclx₄-sodium salt, the calixarene has shown great potential for self-assembly.⁸⁵ Sclx₄ can self-assemble as bilayers in an antiparallel (up-down) fashion, such that the hydrophobic layer is separated from a hydrophilic layer containing the guest molecules. In an up-up arrangement mediated by Na⁺, pyridine *N*-oxide and lanthanide ions, sclx₄ self-assembly can be spherical or extend to helical tubular structures.⁸⁶ Crystal structures also show sclx₄ self-assembled with biorelevant molecules,⁸⁴ including acetylcholine,⁸⁷ oligoammonium cations⁷⁸ and chlorhexidine.⁸⁸ Crystal structures have been determined for sclx₄ complexed with the cationic amino acids histidine,⁸⁹ lysine,⁹⁰ and arginine⁹¹ and also the hydrophobic amino acids alanine, phenylalanine and tyrosine.⁸⁹ The upper rim sulfonates and lower rim hydroxyl substituents enable sclx₄ to interact through both polar and charge interactions. The aromatic nature of the calixarene core means that hydrophobic and cation- π interactions are also viable. These chemical properties explain why sclx₄ can bind amino acids, which fit the cavity. Sclx₄-amino acid complexes have also been studied in solution.⁹²⁻⁹⁴ In water (pH adjusted to 1 and 5) sclx₄ binds lysine and arginine in a 1:1 ratio.^{92,93} Significant perturbations for the ¹H on the C ^{δ} and C ^{ϵ} of lysine and also for the ¹H on the C ^{δ} of arginine suggest that this portion of the amino acids are included in calixarene cavity. The complexes can be inhibited by salt addition pointing to a strong electrostatic contribution to the interactions. Methylation of these side chains increases the binding affinity \sim 70-fold for lysine modified to trimethyllysine and 3-fold for arginine that has been dimethylated, in neutral phosphate-buffered solutions.⁹⁵ The increased tendency to cation- π , which is associated with the increasing number of methyl groups on the lysine, is expected to be the major contributor for sclx₄ complexation.

Interaction affinities are influenced by the symmetry of sclx₄.⁷⁶ For example, both NMR and ITC experiments show that a calixarene possessing three sulfonates and one phenyl group at the *para* position displays μ M affinity for the LysMe₃ amino acid in phosphate buffer, pH 7.4. The affinity is almost twice that of the parent compound, sclx₄. NMR spectroscopy showed significant upfield chemical shifts for the ¹H on the LysMe₃ C ^{α} indicating CH-aromatic contacts between this ¹H and the appended phenyl group. Introducing amino and carboxy substituents on this phenyl group significantly reduced LysMe₃ binding. These electron-withdrawing groups are considered to inhibit interactions through weakened CH- π interactions.

Sclx₄ alters the behaviour of bovine serum albumin. Complexation of the protein by sclx₄ results in aggregation.⁹⁶ Although the 67 kDa protein is acidic, it has 59 lysines, 23 arginines and 17 histidines.⁹⁷ Electrostatic interactions involving arginine- and lysine-rich surface patches are likely to be responsible for the binding event. Electrospray ionization mass spectrometry shows sclx₄ bind bovine serum albumin at three sites, while the larger sclx₆ and sclx₈ bind at one site.⁹⁸

The varying affinity of sclx₄ for different amino acids was exploited for a substrate-selective supramolecular tandem assay. The calixarene was used to monitor the hydrolysis of arginine to ornithine, which was catalyzed by arginase.⁹⁹ Sclx₄ has a higher affinity for the substrate, arginine ($K_d \sim 0.2$ mM), than the product, ornithine ($K_d \sim 2.0$ mM). Therefore, the depletion of the substrate over the course of the reaction allows the fluorescent dye to bind to the sclx₄ cavity ($K_d \sim 16.7$ μ M), which in turn decreases the fluorescence. During the enzymatic reaction, the sclx₄-dye complexes reduce the fluorescence response. Similarly, cucurbit[7]uril was applied in a supramolecular tandem assay to monitor the oxidation of cadaverine to 5-aminopentanal by diamine oxidase.⁹⁹

The combined readouts of different calix[n]arenes as dye-displacement sensors can also be exploited to discriminate posttranslational modifications of cationic residues on histone tail peptides.¹⁰⁰ Sclx₄, sclx₆ and a calix[4]arene, having three sulfonates and one Br at the *para* position, were used with two fluorescent dyes to produce a pattern of signals that is unique to each analyte and their concentrations. Solutions containing unmodified, trimethylated and acetylated Lys9 on the histone H3 tail peptides were screened. The unmodified and both modified Lys9 forms could be distinguished by virtue of their fluorescent fingerprints. The method was also used to resolve the location of trimethylated lysine residues in the peptide. These experiments demonstrate that small molecules that distinguish small solvent-exposed recognition sites are useful chemical sensors.

Lipid monolayers doped with different combinations of *para* substituted calix[4]arenes distinguished proteins based on their displayed surface residues.¹⁰¹ For example, both basic histone H1 and cytochrome *c* interacted with monolayers containing tetraphosphonate calix[4]arenes. However, when anilinium calixarene half-spheres were introduced into the monolayer, histone H1 preferentially interacted with the surface compared to cytochrome *c*. Carboxylic acid calix[6]arene derivatives were also incorporated into supported lipid films.¹⁰² Electrochemical impedance spectroscopy showed cytochrome *c* binding the film with K_d of 66.1 nM, however, films lacking the

calixarene showed no binding. Bovine serum albumin (pI ~5.6) made diminutive interactions highlighting the importance of complementing charge.

Similar to the crown ethers, a carboxylic acid derivative of calix[6]arene solubilised cytochrome *c* into chloroform.¹⁰³ In H₂O₂ supplemented chloroform, cytochrome *c* displayed enzymatic activity by catalysing the oxidation of 2,6-dimethoxyphenol. To demonstrate that the calixarene preferentially binds the lysine residues of cytochrome *c*, these side chains were guanidylated. The modified protein, possessing more guanidine groups than amino groups, was more difficult to solubilise than the native protein. In a related study, the calix[6]arene separated cytochrome *c* from solutions containing lysozyme.^{103,104} Calix[4]arenes with two or four carboxylphenyl groups at the *para* position bind cytochrome *c* in a 1:1 manner in dimethylformamide.¹⁰⁵ The affinity is greater for the *tetra*-substituted calix[4]arene than the *di*-substituted calix[4]arene, emphasizing the electrostatic contribution to association. Calix[4]arenes bearing one methylenebisphosphonic acid at the *para* position bind tyrosine phosphatase 1B via arginine and lysine residues.¹⁰⁶ The IC₅₀ value of 1.2 μM is indicative of a high affinity complex, however, increasing the number of methylenebisphosphonic substituents weakens the interaction.

Type 3: Binding large surface patches using Calixarenes

The large size and relatively flat nature of many protein interfaces requires a ligand design that matches these surfaces. Calixarenes that complement both the electrostatics and structural features of the target protein have great potential as modulators of protein behaviour. While sclx₄ and sclx₆ form 1:1 complexes with amino acids, the calix[8]arenes having a larger, more flexible cavity show great potential for targeting larger surface patches. The 1:2 complexes sclx₈ makes with amino acids lysine and arginine preliminarily suggests this.⁹³

Another approach for targeting bigger binding sites involves adorning calix[4]arenes with large substituents. For example, a calix[4]arene derivative consisting of a constrained anionic GlyAspGlyAsp loop at the *para* position was tested as a binder for cytochrome *c* (Figure 7).¹⁰⁷ The ligand design was inspired by the antibodies that complex protein surfaces, therefore, both the charge and interfacial area were factored. Formation of the protein-calixarene complex inhibited cytochrome *c* reduction by ascorbic acid. Furthermore, a fluorescence titration with the calix[4]arene derivative was

shown to disrupt the complex formed between cytochrome *c* and its natural redox partner, cytochrome *c* peroxidase.¹⁰⁸ The calixarene bound cytochrome *c* in a 1:1 binding mode with an affinity of ~ 0.3 nM in 10 mM phosphate buffer, pH 6.0. These experiments exemplify the potential of the calixarene as a modulator of both protein-protein interactions and protein function. Non-denaturing gel electrophoresis show that the calix[4]arene, which also complements the cationic surface of α -chymotrypsin, disrupts the protein from its 1:1 complex with trypsin.¹⁰⁹ However, the calix[4]arenes having GlyAspGlyTyr or GlyLysGlyLys loops show little or no inhibition of the α -chymotrypsin-trypsin complex, respectively.

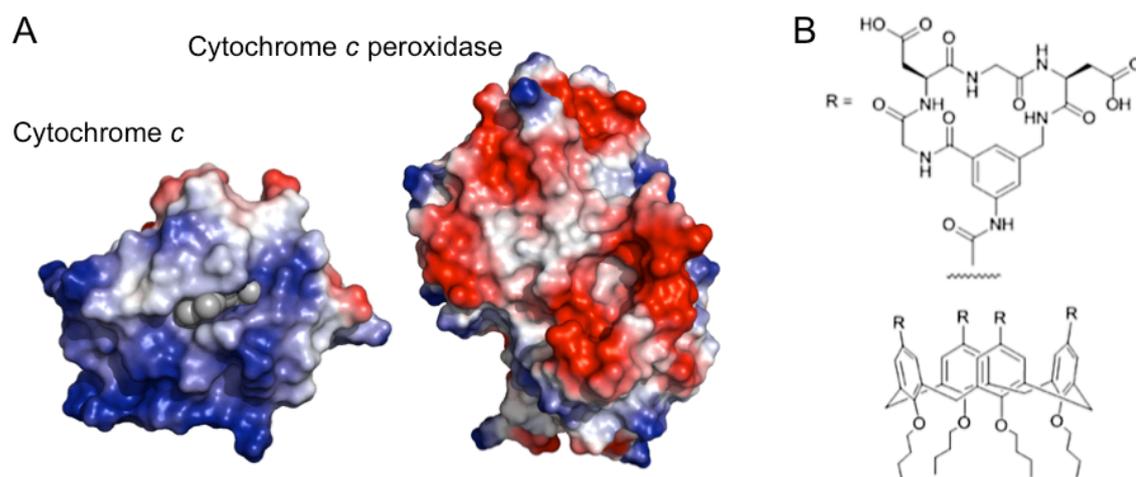


Figure 7. Designed molecular mimicry for binding large protein surfaces. (A) Electrostatic surface representation of the opened cytochrome *c*-cytochrome *c* peroxidase complex reveals complementary surfaces (PDB ID 2pcc).⁹ (B) Calix[4]arene decorated with GlyAspGlyAsp loops at the *para* position was designed to mimic the binding site of cytochrome *c* peroxidase.¹⁰⁷ The anionic ligand competes with the partner protein for cytochrome *c* complexation.¹⁰⁸

A calix[4]arene possessing four guanidiniomethyl substituents at the *para* position was designed to complement the “hot spot” glutamates of the p53 tumour suppressor protein.¹¹⁰ Using Molecular Dynamics (MD) and NMR the calix[4]arene was shown to bind the p53 R337H mutant. Interactions occurred between the guanidine substituents and glutamates, as well as the conical calix[4]arene core, which inserted into the hydrophobic cleft formed between two monomers (Figure 8). The calix[4]arene-protein interactions resulted in the stabilisation of an otherwise unstable p53 mutant in water, pH 7. Similarly, a calix[6]arene derivative having six imidazole groups was designed to bind the same p53

R337H glutamates.¹¹¹ *In vitro* assays show the calix[6]arene stabilise the protein and enhance the transcriptional activity under physiological conditions.

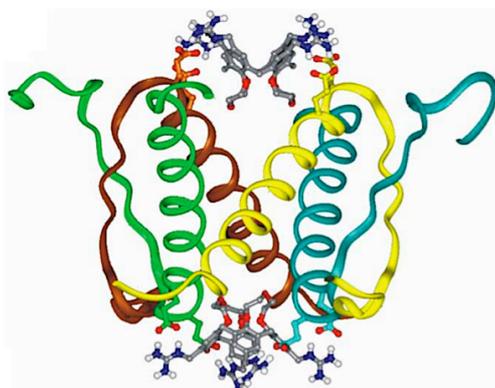


Figure 8. Guanidocalix[4]arene-p53 R337H complex determined by MD.¹¹⁰ Figure is copyright of (2008) National Academy of Sciences, USA. Two calixarenes bearing guanidine groups bind the tetramerisation domain, making charge-charge interactions with Glu336 and Glu339. The lower rim of the ligand fits into the hydrophobic pocket of the protein.

Voltage-dependent potassium channels consist of a tetrameric assembly with a central pore that conducts K^+ across the cell membrane.¹¹² MD and electrophysiology assays demonstrate that a calix[4]arene having guanidines at the *para* position bind these channels.¹¹³ The interactions were similar to the p53 R337H-calix[4]arene complex. For example, the cationic substituents of the calix[4]arene complemented the predominantly negatively charged surface patch of the channel. Moreover, the calix[4]arene core and lower rim hydroxyls bound the central, conical pore. As a result of the protein-calix[4]arene complex, the channel's function was inhibited.

Calix[4]arenes having lower rim substituents that mimic the antiparallel β -sheet structure and are terminated with cationic substituents have shown antitumor activities.¹¹⁴ The agglutination assay has shown that the calix[4]arene with tetra-amine terminated substituents inhibits cell proliferation in a dose-dependent manner.¹¹⁵ It does so by binding human galectin-1. NMR spectroscopy shows that the calixarene binds one region on the protein surface and this causes the perturbation of the residues located in the carbohydrate binding region, which prevents carbohydrate binding.

Type 3: Binding large surface patches using Porphyrins

Porphyrins, which are large, rigid, planar molecules can bind large surface patches and modulate protein behaviour. The porphyrin scaffold consists of a hydrophobic core that can be tuned at its periphery to match specific surfaces. For example, the tetra-phenylporphyrin scaffold bearing different anionic and hydrophobic groups, mimics the interaction interface of cytochrome *c* peroxidase that binds cytochrome *c* (Figure 9).¹¹⁶

Fluorescence quenching of the porphyrins, caused by cytochrome *c* binding in 5 mM phosphate buffer at pH 7.4, indicated the formation of 1:1 protein-porphyrin complexes. The porphyrin possessing 8 phenyl groups and the highest net charge (-8) displayed the greatest affinity, $K_d \sim 20$ nM, for cytochrome *c* in this study.

The interactions involving cytochrome *c* and a tetra-phenyl-porphyrin bearing four anionic TyrAsp substituents cause the protein's denaturation temperature to decrease.¹¹⁷ Acetylation of all the cytochrome *c* lysine N^ε greatly reduced the interactions indicating that the complementing surface charges are required for complex formation. The larger tetra-biphenyl-porphyrin bearing 16 carboxyl groups is a more potent binder ($K_d \sim 0.67$ nM) and denaturant for cytochrome *c*.¹¹⁸ Both porphyrins make 1:1 complexes with cytochrome *c*. However, the larger, more anionic porphyrin has a significantly higher affinity due to its increased size and charge. Proteins of a similar size and having a lower pI than cytochrome *c* experienced significantly weaker binding by the tetra-biphenyl-porphyrin. This suggests that complementary charges are required for complex formation. Additionally, cytochrome *c* denaturation could be reduced through salt addition, further highlighting the electrostatic contribution to protein binding.

NMR spectroscopy was used to map interactions between cytochrome *c* and two anionic porphyrins onto the protein surface, in 25 mM potassium phosphate and 50 mM NaCl, pH 6.0.¹¹⁹ The porphyrins possessed either four propionic acid or dicarboxylatophenyl substituents. Although the data for both porphyrins point to a single binding event, the interactions are dynamic and extend across a large patch that covers about half of the protein surface. The dicarboxylatophenyl substituted porphyrin displayed ~ 5 fold higher affinity for cytochrome *c* than its propionic acid counterpart.

Porphyrins can also be used to differentiate proteins based on their surface properties,¹²⁰ in a similar manner to the substrate-selective calix[n]arene assays. Owing to the spectroscopic properties of porphyrins, their association with protein causes fluorescence changes. Statistical treatment of the FRET-binding patterns enabled the identification of various proteins and protein mixtures. For example, the acidic ferredoxin can be distinguished from cytochrome *c*. The paramagnetic iron in both proteins show distinct fluorescence quenching when specific porphyrins strongly bind the protein surfaces. Lysozyme can be distinguished from the acidic α -lactalbumin as the association of specific porphyrins at the protein surfaces changes in the environment of the fluorophore and, in turn, its fluorescence.

Porphyrins have also been used to exploit the four-fold symmetry of potassium channels (Figure 9).¹²¹ For example, porphyrins bearing cationic substituents were designed to interact simultaneously with all four subunits of the channel via glutamate side chains, similar to the calix[4]arene. Competitive binding assays and electrophysiological assays showed that the interactions between a cationic porphyrin and the potassium channel partially blocked K^+ conductance and were reversible. Negatively charged porphyrins showed no interaction. The data provides a guide for the design of therapeutics that aim to hamper cell secretion by disrupting the function of potassium channels.

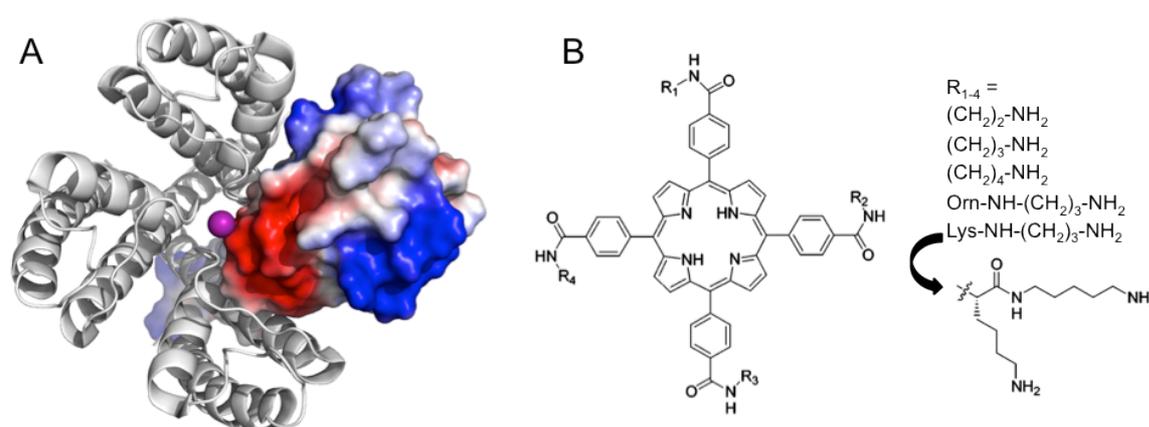


Figure 9. The exploitation of symmetry for molecular mimicry. (A) The crystal structure of the potassium channel tetramerisation domain. K^+ (purple sphere) binds the centre of the pore. Three protein chains are grey cartoons and the electrostatic surface of the fourth chain is shown. (B) The porphyrin matches the four-fold symmetry of the channel and complementary cationic substituents drive complexation.¹²¹

Type 3: Binding large surface patches using Polymers

Bisphosphonate co-polymers can distinguish protein surfaces by virtue of their cationic residue composition.¹²² For example, arginine-rich lysozyme binds the co-polymer containing bisphosphonate, alcohol and dodecyl substituents in a ratio of 3:1:1, with a 100-fold higher affinity (K_d value of 25 nM) than cytochrome *c*.¹²³ As the proteins share similar sizes and pI, the selectivity for lysozyme over cytochrome *c* is attributed to the polymer's preference for binding arginine side chains. Interactions between the solvent-exposed bisphosphonate head of the co-polymer and arginine residues lead to multiple protein molecules binding one polymer strand.

Type 3: Binding large surface patches using Dendrimers

Dendrimers are another group of well-defined molecules that are used for protein surface recognition. The ligands have an inner core that is repeatedly branched to complement to the size of protein surfaces. Dendrimers with bisphosphonate substituents bind the surfaces of basic proteins.¹²⁴ An octameric dendrimer forms strong complexes with small basic peptides and tightest binding, on the μM scale, is observed for histone H1. No interaction could be detected for the anionic ovalbumin or carbohydrate anhydrase. Carboxylate dendrimers that closely match the size of the natural interaction interfaces of cytochrome *c* and chymotrypsin have the greatest affinity for binding these protein surfaces and inhibit protein-protein interactions.¹²⁵

Subject of this thesis

There are numerous examples of protein-small molecule complexes showing altered protein properties, from solubility to modified interaction behaviours. The majority of these investigations lack structural data, which is required to thoroughly characterise precise binding interactions (Table 1). Although calixarenes are good protein binders and modulators of protein behaviour, there is no crystal structure showing exactly how a calixarene complexes proteins. Therefore, this thesis focused on the ability of *sclx*₄ to bind protein surfaces in three model systems. The goal was to establish the binding mode and precise structural information for protein complexation using X-ray crystallography. Protein-small molecule co-crystallization and crystallography were central to this thesis. Therefore, Chapter 1 describes the crystallization methods used and provides an overview of the data collection strategy right through to structure solution.

Model proteins

The two proteins used were lysine- or arginine-rich. The systems provided opportunities to identify, characterise and rationalise the observed interaction epitopes and finally to compare the results obtained for each system (Discussion).

In this work the first target protein used was *Saccharomyces cerevisiae* cytochrome *c* (cyt *c*). It is primarily known for its function in the mitochondrion supporting ATP synthesis, however, it also presents a crucial trigger for apoptosis.¹²⁶ Cyt *c* is a useful model protein for the ease at which it can be over-expressed and purified.¹²⁷

Moreover, the interactions with protein partners have been well studied and its crystal structure is known.^{9,128,129}

Cyt *c* consists of a 108 amino acids with a molecular weight of ~13 kDa. The protein is redox active by virtue of a single haem group covalently attached by Cys14 and Cys17. Importantly, cyt *c* is cationic (pI >10). Of the 34 protonatable residues, 16 are lysines, 3 are arginines and 4 are histidines.¹³⁰ Most of the cationic residues are located in a broad patch in the vicinity of the exposed haem edge of cyt *c* (Figure 10). This patch contributes to the binding site in protein–protein interactions through electrostatic steering and interactions with the complementing surfaces of its partner proteins.^{14,128,129,131-133} The complexes formed by cyt *c* and sclx₄ are described in Chapter 3. Both NMR and crystallography were used to study the complexes. Recalling that cyt *c* is lysine-rich, the system addressed three central points: (i) are lysines specifically involved in interactions with sclx₄, (ii) what is the contribution of electrostatic interactions, and (iii) what is the structure of the sclx₄ binding site(s)? An Arg13Glu (R13E) mutant of cyt *c* was also co-crystallized with sclx₄ for comparison (Chapter 2). The larger size of these crystals inspired their electrochemical characterisation (Chapter 3). Therefore, the crystals were grown on modified Au electrodes and characterised by using cyclic voltammetry. The project was carried out in collaboration with Prof. Lisdat and Dr. Sven Feifel in TH-Wildau.

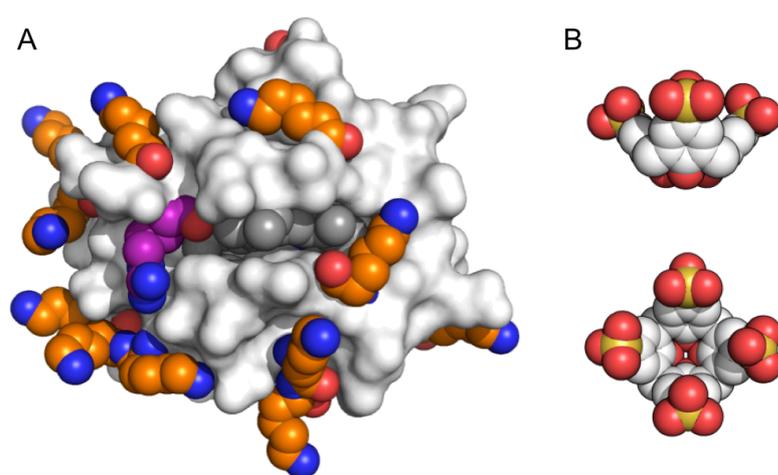


Figure 10. Small surface-displayed recognition sites for sclx₄ binding. (A) Surface representation of cyt *c* (grey) with lysine and arginine residues shown as orange and purple spheres, respectively. The exposed haem is shown as dark grey spheres. (B) Side on and top down views of sclx₄. Sclx₄ is represented as spheres.

Hen egg white lysozyme was the second model protein. Lysozyme catalyses bacterial cell wall degradation and specifically hydrolyses the β -(1,4)-linkages between *N*-acetylmuramic acid and *N*-acetyl-d-glucosamine in peptidoglycans.¹³⁴ It was

the first enzyme to have its three-dimensional structure solved by X-ray crystallography.¹³⁵ Although its biological function has been elucidated, lysozyme is better known for the ease at which it crystallizes. Therefore, it has been used in numerous systems that study protein crystallization^{37,136} and was engineered into the β_2 -adrenergic receptor to facilitate crystallization of the membrane protein.¹³⁷

Lysozyme consists of 129 amino acids and has a molecular weight of ~ 14 kDa. Of the 27 protonatable residues, 11 are arginines, 6 are lysines and 1 is histidine, inferring a net positive charge ($pI > 9$).¹³⁸ Lysozyme was chosen as a model protein as it is arginine-rich. Therefore, it is an interesting comparison for the lysine-rich cytochrome *c*. The lysozyme-sclx₄ complex is described in Chapter 4. Crystallography was used to study the complexes and identify the features that promote complexation of an arginine-rich protein by sclx₄.

Sclx₄ displays highest affinity for dimethyllysine, over lysine and arginine side chains.^{95,139} The lysine residues and N-terminus of lysozyme were dimethylated to yield lysozyme-KMe₂, the third model protein. Lysozyme-KMe₂ is arginine-rich and the modification doesn't change the net charge, therefore, it is a good model protein for testing the selectivity of sclx₄. The crystal structure of the protein and sclx₄ is described in Chapter 5. The system investigated three central points: (i) does sclx₄ have a preference for the dimethylated lysine residues, (ii) do other cationic side chains contribute to the association of lysozyme-KMe₂ and sclx₄ and (iii) do dimethylated lysine residues modify protein self-assembly? Overall, the three model systems have a similar size and charge but differ in the proportion of cationic residue types providing distinguishing interaction opportunities for sclx₄.

Chapter 1:

Co-crystallization and X-ray crystallography of
protein-small molecule complexes

Abstract

Model systems using cytochrome *c* or lysozyme are important for investigating protein-small molecule complexation. As small molecules can alter the solubility of proteins, the co-crystallization method, but moreover the crystal structure solution are valuable for understanding the recognition sites involved in self-assembly. In this chapter the application of crystallization and crystallographic methods for the analysis of a variety of protein-small molecule complexes is described. The data collection protocol employed at the European Synchrotron Radiation Facility, beamline ID23-2, right through to structure solution are described.

Introduction

X-ray crystallography is a powerful tool for visualizing protein structure at atomic resolution. The first step to determine an X-ray crystal structure of a protein is often deemed to be the most difficult. It requires the growth of sufficiently large, diffraction quality protein crystals. That is to say, the molecules must have an ordered arrangement that extends over a long range in 3 dimensions. Protein crystallization is burdened by an entropic penalty due to the incorporation of molecules into the crystal lattice.^{140,141} This involves the loss of rotational and translation freedom of the molecules. Flexible loops and side chains trapped at protein-protein interfaces during crystal contact formation also have reduced flexibility, thus presenting an additional entropic cost.

The tendency of proteins to avoid crystallization is attributed to evolutionary design.¹⁴² For example, globular proteins are believed to have evolved to avoid aggregation through a 'surface-entropy shield' consisting mostly of lysine and glutamate residues. Lysine residues are the most solvent-exposed residues in proteins with 68 % and 26 % being completely or partly exposed, respectively.⁵⁶ Due to the high conformational entropy of the solvent-exposed lysine residues, they are routinely mutated or covalently modified to promote the nucleation of proteins that are otherwise intractable to crystallization.^{33,34,143} Mutating this residue to small amino acids that have lower conformational entropy helps achieves this.¹⁴³ Lysine side chains are also regularly methylated to reduce protein solubility and promote protein crystallization.^{33,34} The use of small molecules as precipitants is also important for crystallization.⁴⁰ For example, 18-crown-6 promoted the crystallization of the RmbA protein by binding lysine side chains.⁴⁷ The small molecules tested for their ability to co-crystallise lysine-rich cytochrome *c* are discussed.

Principals of protein crystallization

The crystallization phenomenon is composed of two steps, nucleation and crystal growth.¹⁴⁴ Nucleation is the phase transition achieved by moving from an undersaturated solution to a supersaturated solution. Consequently, a cluster is formed creating two phases, the cluster and the mother solution. Crystal growth can subsequently occur via the mass transport of solute molecules towards the cluster surface followed by adsorption and incorporation into the lattice.

The varying degrees of saturation are distinguished by the phase diagram (Figure 1A).¹⁴⁴ The four unique zones are the undersaturation, metastable, nucleation and precipitation zones. The undersaturation zone is characterised by fully solubilised protein and is devoid of any precipitate. The area of lowest supersaturation, the metastable zone, is associated with crystal growth and stability. Spontaneous nucleation occurs in the nucleation or labile zone. The precipitation zone is the region of highest supersaturation in which undesirable protein precipitation or liquid-liquid phase separation occurs. Although supersaturation is important for crystal growth, the rate at which supersaturation occurs is also crucial to the process.¹⁴⁵ The crystallization method employed, be it vapour diffusion, batch or counter diffusion, governs the path taken through the solubility zones.^{146,147} The hanging drop vapour diffusion method is central to this thesis. In this technique μL quantities of the protein solution and the reservoir solution are combined in a drop that hangs from a cover slide (Figure 1B). Equilibration between the drop and the reservoir solution is the driving force for achieving supersaturation and, in turn, nucleation.

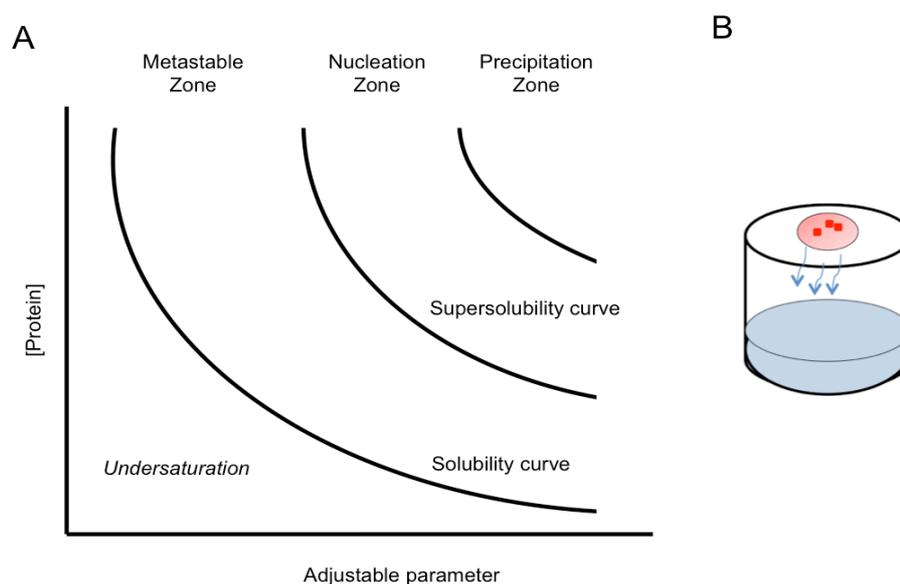


Figure 1. Protein crystallization involves a phase transition achieved by decreasing protein solubility. **(A)** A phase diagram depicting different solubility zones. Varying a parameter such as pH or precipitant concentration affects the protein solubility. **(B)** A scheme for the hanging drop vapour diffusion method. Water vapour diffuses out (arrows) of the drop allowing equilibration with the reservoir solution. As a result of this diffusion the protein concentration increases moving towards supersaturation and crystallization.

The effects of precipitants

The solutions used for crystallization typically consist of a homogenous protein sample in pure water or at low salt concentration. The protein solubility must be reduced to overcome natural protein-protein repulsions that exist in monodisperse protein solutions and for a new phase to appear. This phenomenon is characteristic of all protein crystallization experiments.¹⁴⁴

Intermolecular forces govern crystallization, therefore, the process is sensitive to physical and chemical properties. Precipitants are selected to drive the phase transition. For example, the ionic strength of the solution is a crucial factor for modifying protein solubility.¹⁴⁸⁻¹⁵⁰ At low ionic strength, termed salting-in, the solubility of the protein increases.¹⁴⁸ This is attributed to the association of the ions (from the salt) with the protein surface, which subsequently shields water molecules from those surface areas. Alternatively, increasing the ionic strength, or salting-out, involves the saturation of the protein surface by salt ions that interact with the solvent, resulting in decreased protein solubility.

pH is an important factor for crystallization as the protein's surface charge distribution and, in turn, the intermolecular interactions change with pH.¹⁵¹ The pH values that maintain the folded structure of the protein can only be used for crystallization. The protein's isoelectric point (pI) is usually the protein's lowest solubility. Therefore, buffers that operate in the pI range are useful for crystallising proteins. When the pH is less than the side chain's pK_a , the functional group is protonated and when pH is greater the side chain is then deprotonated. At $pH > 4$ acidic side chains are deprotonated and at $pH < 6$, 10 and 12 the side chains of histidine, lysine and arginine are protonated, respectively. For co-crystallization experiments the protein surface charge distribution should complement the charge of the ligand.

Organic polymers are also useful precipitants. For example, poly(ethylene glycol) (PEG) is recommended as an initial trial reagent for protein crystallization.¹⁵² McPherson has shown that screens prepared with varying concentrations of a given PEG size crystallized 13 out of 22 proteins. PEG-protein interactions are usually repulsive and both PEG and protein molecules compete for water. This leads to crowding effects that promote protein-protein interactions.¹⁵³

Metal ions are useful promoters of protein self-assembly.¹⁵⁴ As a result, they are used to promote crystal growth by bridging and stabilising intermolecular contacts.¹⁵⁵ For

example, crystals of plastocyanin could only be grown in conditions containing Zn^{2+} .¹⁵⁶ The crystal structure shows Zn^{2+} at protein interfaces coordinating a histidine residue on one molecule and two aspartates on another. Lysozyme and the maltose-binding protein were engineered to possess metal-binding sites.¹⁵⁷ The protein variants yielded crystals upon metal coordination. The location of the metal-binding sites was shown to affect the oligomeric assemblies and the crystal packing.

Small molecules that specifically bind protein surfaces can have substantial effects on protein solubility and subsequently, the crystallization behaviour.¹⁵⁸ For example, protein binding by small synthetic molecules can modify the surface charge distribution. This was demonstrated by the small molecule pyrenetetrasulfonic acid, which promoted the crystallization of *Leishmania mexicana* PYK.⁴⁰ The presence of the ligand in the crystallization buffer induced crystal growth. In the crystal structure, pyrenetetrasulfonic acid was observed stacked between the arginine side chains of two protein molecules acting as ‘molecular glue’. Charge-charge interactions were observed between the arginine side chains and the ligand sulfonates.

Protein-small molecule co-crystallization

The two main strategies for obtaining protein-small molecule crystals are the soak-in and co-crystallization methods.^{159,160} As coined by Emil Fischer, many protein-ligand complexes may be considered as a ‘lock-and-key’.¹⁶¹ This analogy highlights the specificity of a ligand binding an enzyme via insertion into a binding pocket. The soak-in method takes advantage of this phenomenon and is typically employed to screen a variety of ligands for one binding site. Crystals containing 4-anilinoquinazolines bound to cyclin-dependent kinase 2 were obtained using the soak-in method.¹⁶² Crystals of the apo-protein are a prerequisite for soak-in method and the crystals were soaked for 24 hours with an excess of each ligand prior to data collection. The solvent channels, accounting for the bulk solvent of crystals, present a route for the ligand to the binding site.^{163,164} Therefore, the binding site, which is typically well-characterised, must be solvent exposed in the structure. Ligand binding, however, can induce conformational changes that damage the crystal lattice and lead to crystal cracking or dissolution.

Tumour-necrosis factor (TNF) exists as a trimer and is a key factor in inflammatory responses.¹⁶⁵ Fragment screening identified a class of small molecules that bind and disrupt the TNF complex ($K_d \sim 13 \mu M$).²⁰ Preliminary efforts to soak-in one

small-molecule inhibitor of TNF into the crystals resulted in crystal cracking. These challenges were overcome using the co-crystallization method. For this method, the ligand and protein are premixed prior to preparing crystallization experiments. The structure of the TNF-inhibitor complex showed a TNF dimer with the small molecule bound at a protein-protein interface, a binding site that is otherwise completely buried in the TNF trimer complex.

Crystal structures for the protein ProX from *Archaeoglobus fulgidus*, show the conformational changes that occur on ligand binding.¹⁶⁶ The free ProX has an open conformation. Co-crystallization experiments were prepared for ProX with trimethylammonium, glycine betaine and proline betaine. The resultant crystal structures show a closed conformation and the box-type binding site is formed by the rearrangement of the two protein domains. Therefore, the co-crystallization method is useful as the protein-ligand complex has the ability to form prior to crystallization. As a result, the crystal structure is more likely to display the preferred conformation and interaction site(s), which may not have been available using the soak-in method.

Ligands selected for co-crystallization

The capacity of four synthetic ligands to alter the crystallization behaviour of one or two cationic proteins was tested. All the ligands are water soluble and possess features that complemented the protein surface of cytochrome *c* (cyt *c*).¹³⁰ For example, each ligand has a hydrophobic core and an anionic periphery. However, the size, structure and charge of the ligands vary (Figure 2). Therefore, the ligands have different tendencies to make noncovalent interactions and this was expected to influence complex formation.

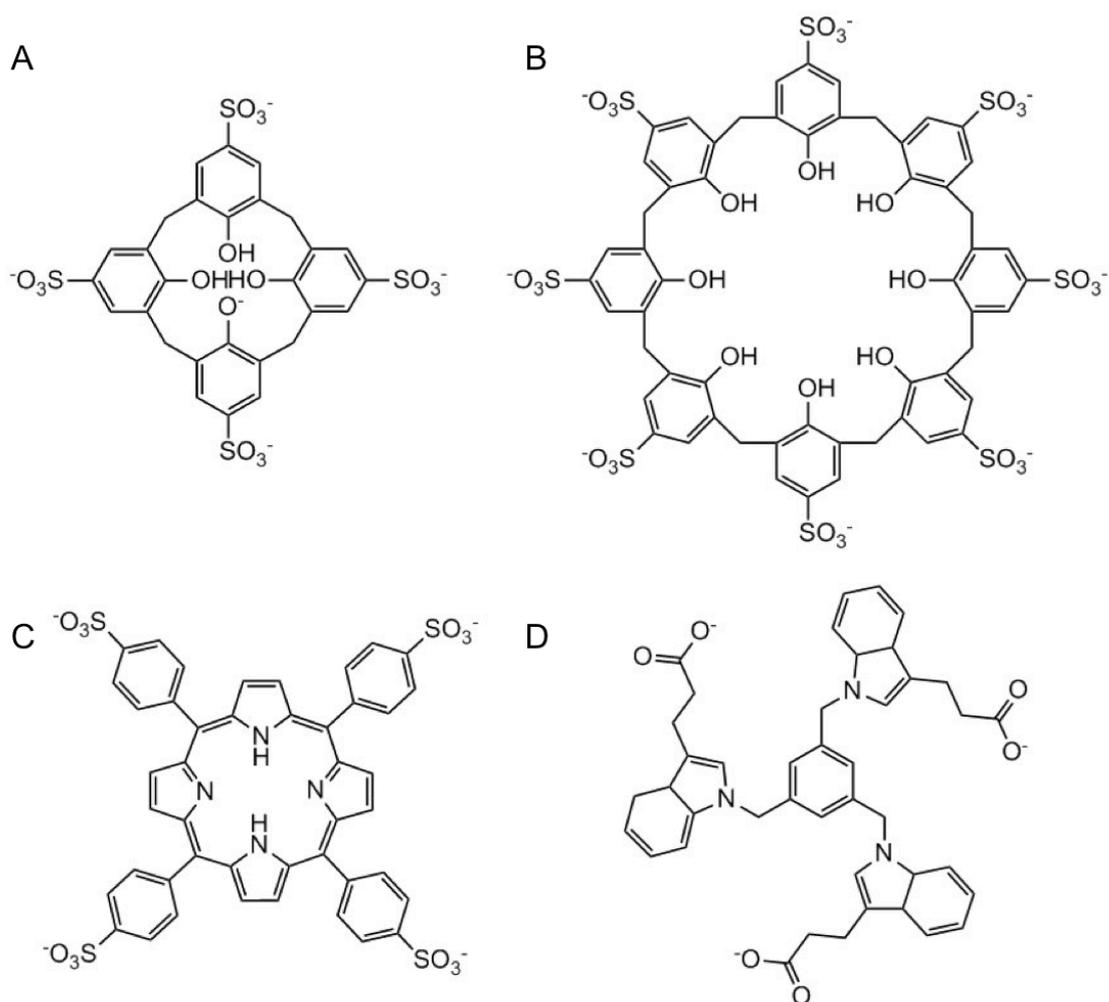


Figure 2. The ligands used for co-crystallization. (A) *p*-sulfonatocalix[4]arene, (B) *p*-sulfonatocalix[8]arene, (C) *meso*-tetra(4-sulfonatophenyl)-porphyrin and (D) tris-indole ligand.

p-sulfonatocalix[4]arene (sclx₄) was the first small molecule tested by co-crystallization experiments. The molecule was selected primarily for its ability to bind cationic residues and incorporate guest molecules in its cavity.^{90,93} *p*-sulfonatocalix[8]arene (sclx₈), which can bind two amino acids simultaneously,⁹³ was also chosen as a contender for co-crystallization experiments. The greater negative charge was expected to enhance *cyt c* complexation and the size was expected to complement a larger interface.

Owing to its water solubility and commercial availability tetra(4-sulfonatophenyl)-porphyrin (TPPS₄) is among the most commonly studied porphyrins. The four crystal structures in the PDB demonstrate the self-assembly propensity of TPPS₄.⁵⁰⁻⁵² In each structure the porphyrin pairs are sandwiched between protein monomers, mediating the crystal-lattice interface. TPPS₄ can also self-assemble in a cooperative manner with a

lysine-rich peptide, driven by complementary charges.¹⁶⁷ As TPPS₄ is a planar molecule, it was thought to present an interesting comparison for the cup-shaped molecule, sclx₄.

Small molecules implemented in molecular recognition studies of protein-ligand complexes are often designed taking inspiration from nature.³ For example, a tris-indole ligand was designed to mimic the aromatic cages that bind trimethylated lysines.¹⁶⁸ The host is a flexible, indole-based structure with mM affinity for trimethylated lysine in phosphate-buffered D₂O. There is no structural information showing the tris-indole ligand binding protein surfaces. Therefore, co-crystallization trials were prepared with the tris-indole ligand to study its capacity to bind the cationic surface of *cyt c*.

Standard protocols: From co-crystallization to crystallography and structure solution

Co-crystallization

The hanging drop vapour diffusion method was implemented for all crystallization experiments. All trays were incubated at 20 °C, unless otherwise stated. Two approaches were used for preparing co-crystallization experiment drops. One approach involved the addition of 1 μL of each protein and ligand to 1 μL of reservoir solution. The second approach involved premixing the protein and ligand solutions and allowing protein-ligand complexes to preform, similar to the TNF-inhibitor co-crystallization method.²⁰ The experiment drops were then prepared by adding 2 μL of the protein-ligand solution to 1 μL of reservoir solution. These two approaches for preparing crystallization drops can affect crystal growth. For example, premixing the protein and ligand solution prior to drop preparation improved crystal growth for a *cyt c* mutant with sclx₄ (Figure 3). For all initial co-crystallization experiments the ligand concentration was 10-fold higher than the protein concentration. Control drops were essential to confirm that the observed precipitates and crystals were a result of ligand-protein interactions, rather than a precipitant effect. Therefore, the control drops were prepared by substituting the ligand component of the crystallization drops with water.

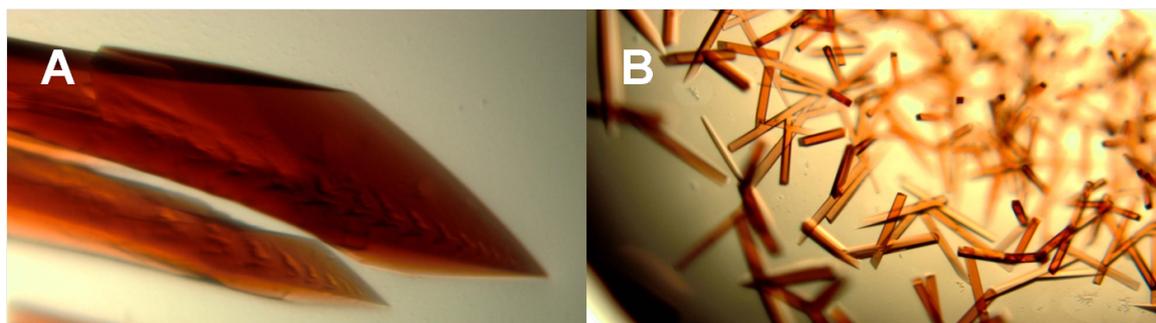


Figure 3. Cyt *c* R13E crystals that grew with 10-fold excess of sclx₄. **(A)** Depressed faces observed on the crystal surfaces are characteristic of the experiment drop preparation. The experiment drop was prepared by combining equal volumes (1 μ L) of reservoir solution, cyt *c* R13E and sclx₄. **(B)** Crystals that display smooth surfaces and many of the crystals are symmetrical. Mixing cyt *c* R13E and sclx₄ prior to addition to the reservoir solution significantly improved the appearance of the crystals.

Preliminary crystallization trials used PEG 8000 as the sole precipitant. The PEG concentrations were 0-30 % or 0-50 %. The crystallization trials typically incorporated either a salt or buffer over the same PEG range. Different combinations of buffers and salts were screened to identify conditions that favoured protein crystallization. Typically, the effects of 50 mM sodium cacodylate pH 6 and/or 50mM NaCl were tested in preliminary optimisation experiments. When sufficiently large crystals were obtained, an oil barrier was frequently employed to improve the appearance of the crystals, in anticipation of X-ray diffraction experiments. This slows down the rate of diffusion and, in turn, the approach of supersaturation and crystallization.¹⁴⁶ The rate of equilibration is influenced by the oil type and thickness. 200 μ L of paraffin oil gently placed above the reservoir solution was common practice. Crystal growth can also be improved by filtering the crystallization solution. When sufficiently large crystals were obtained (typically greater than 10x10x10 μ m) the crystals were brought to Trinity College Dublin and harvested in the Khan or Caffrey lab.

Crystal cryo-protection and harvesting

All crystals were cryo-protected to minimize ice formation during harvesting and to reduce radiation damage during data collection. The cryoprotectant solutions were prepared with up to 20 % of glycerol and 80 % reservoir solution. In advance of harvesting the cryoprotectant alone was tested in a loop under a 100 K liquid nitrogen stream for ice formation. If ice appeared, the concentration of glycerol was increased by

5 % and the test repeated. If glycerol wasn't compatible with the reservoir solution as a cryoprotectant, the glycerol was substituted for another antifreeze.

Typically, 1 μL of cryoprotectant agent was added to the crystallization drop, which was allowed to equilibrate for up to 1 minute. Introducing the cryoprotectant into the mother liquor can induce osmotic shock in the crystal, resulting in its defragmentation or dissolution. Therefore, the crystal behaviour was monitored during the equilibration process. Intact crystals or sufficiently large fragments were rapidly transferred to fresh μL aliquots of cryoprotectant and again allowed to equilibrate for up to 1 minute. The crystals were then harvested, ideally one crystal per nylon loop. The loop matched the crystal size. The crystal-containing loop was plunged into liquid nitrogen and transferred to a vial submerged in liquid nitrogen for storage.

Preliminary X-ray diffraction experiments

As mosaicity and twinning are limiting factors in data processing when possible the diffraction quality of the crystals was tested on an in-house X-ray source in the Khan lab. The crystals were flash-cooled under a stream of nitrogen gas at 100 K and diffracted at 1° and 180° with an exposure time of 2 minutes. The diffraction patterns were visually inspected and their suitability for further X-ray diffraction experiments assessed (Figure 4). The data were also integrated to make a preliminary assessment of the space group, the unit cell parameters and the resolution limit. A significant deviation in the unit cell parameters from the known protein structure indicated altered crystal lattice packing, hinting at a possible new protein-ligand complex. Isotropic crystals were saved for diffraction at synchrotron sources to obtain higher resolution data.

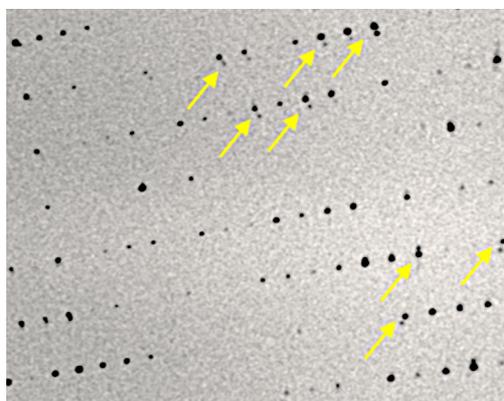


Figure 4. A region of an X-ray diffraction pattern of a protein crystal. Two patterns are observed overlapping. Below the intense rows of spots, weaker spots are visible (arrow), indicative of macroscopic twinning.

Data collection at the European Synchrotron Radiation Facility

Complete data sets were collected at synchrotron sources including the European Synchrotron Radiation Facility (ESRF), Diamond Light Source, SOLEIL, the Swiss Light Source and the Advanced Photon Source. As data collection protocols vary from user to user, the following focuses on the approach implemented at the ESRF beamline ID23-2 with Dr. Valerie Pye for diffracting crystals of proteins of known structure.

The **MxCube** package was used to select and mount a crystal from the basket and to centre the sample. Crystals that were not tested on an in-house X-ray source were rotated and diffracted at 1 ° and at the 90 ° offset angle for diffraction characterisation. The software **EDNA** was employed to determine the crystal symmetry, the unit cell parameters and the resolution limit. The signal/noise was also assessed. **EDNA** subsequently generated a data collection strategy to maximise both the resolution and completeness of the data set. Typically, extra data was collected to compensate for potential data lost due to spot overloads and overlaps.

Data processing and reduction

The **Collaborative Computational Project (CCP4)** is an integrated suite of programs used to determine macromolecular structures by X-ray crystallography, and other biophysical techniques.¹⁶⁹ The following programs, with the exception of **Molprobity**, were run through the **CCP4** suite.

MOSFLM was used to process the data sets and to produce a sorted MTZ file containing reflection indices, their intensities and standard deviations.¹⁷⁰ The first step, called Indexing, uses the Fourier-based data processing suite algorithm to generate a list of possible lattices. The calculation is based on two frames and the solution with the highest symmetry and lowest penalty was selected. The mosaicity estimate must be less than the degree of rotation. A good agreement between the “Spot finding” prediction and those observed on the diffraction pattern indicates the reliability of the assigned space group obtained from indexing. With no discrepancies or challenges a subset of frames from the dataset are used for the cell refinement step. If no significant problems are encountered the whole dataset is integrated (but unmerged) to generate an output file in the CCP4 mtz format for data reduction.

The program **Pointless** was used to score all of the possible Laue groups consistent with the crystal class.¹⁷¹ This involved matching potential symmetry equivalent reflections in the diffraction pattern.

Scala was used to scale symmetry-related intensities together and produce statistics on the data quality.¹⁷¹ Scaling tries to make symmetry-related and duplicate measurements of a reflection equal. It does so by modelling the diffraction experiment as a function of the incident and diffracted beam directions in the crystal. This makes the data internally consistent. The statistics detailed in the Scala log file indicate the overall quality of the data. The Scala output mtz file was used for further reduction.

The program **Matthews** was implemented to determine the solvent content of the unit cell.^{163,172} This information is crucial for molecular replacement, performed with **Phaser**.¹⁷³ The structures described in this thesis were all solved with Phaser, using a coordinate file. The PDB file was edited prior to use to contain only the chain that is in target structure. Water molecules, various ligands and additional chains, were deleted from the coordinate file. Phaser uses the initial model as the search model to be positioned in the unit cell using rotation and translation functions. Low resolution data is adequate to position the structure in the right place and high resolution data provides sufficient difference density to orient and refine the structure. An output coordinate file with the molecule correctly positioned was subsequently generated. In conjunction with this, the experimentally phased electron density map was released as an output mtz file. At this point the agreement between the maps and the crystal packing were manually inspected in **COOT** for the correct solution. The molecular replacement statistics detailed in the Phaser log file were also checked.

Structure refinement is an interactive and iterative process, the objective of which is to best fit a model to an electron density map. Using the molecular graphics program **COOT**, the maps were interpreted and the structure was rebuilt to best fit the map.¹⁷⁴ The modified structure was refined using **Refmac5** and a new map was calculated from that model.¹⁷⁵ More building into the new map was done and the refinement repeated. In addition to fitting the structure to the density map, a close inspection of the chemical environment of the residues is important during the building process. Therefore, a detailed residue-by-residue analysis was carried out. The atom occupancy, which is the fraction of the atom present at a specific site, is an important factor during structure refinement. For example, decreasing the occupancy of an atom increases its temperature factor. The effects can be observed in the mtz file as the amount of electron density at that

point lessens. The R-value and R-free indicate the model quality. The R-value provides an indication of the agreement between the observed and calculated structure factor and the R-free is calculated using a partial dataset that is not used in the refinement of a structure. Both values, expressed as a %, should be as low as possible and similar. Typically, as a last step in the refinement procedure, anisotropic data refinement was used for data better than 2 Å, to improve the R-value and R-free.

Structure validation was performed using **MolProbity**.¹⁷⁶ The program assesses covalent geometry, dihedral angles and chirality. The **MolProbity** results may indicate that structure rebuilding is required. The Ramachandran plot,¹⁷⁷ as generated in **COOT** or **MolProbity**, was also used to assess the structure quality. It shows the distribution of torsion angles of a protein structure and highlights the residues present in sterically disallowed conformations. The solved 3D structures were deposited in the Protein Data Bank.

Co-crystallization results

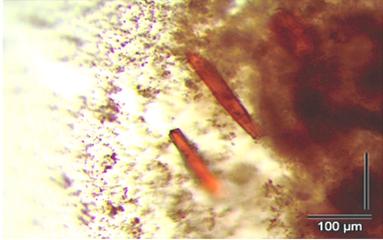
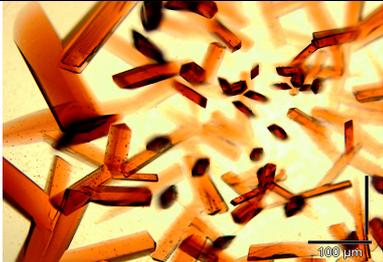
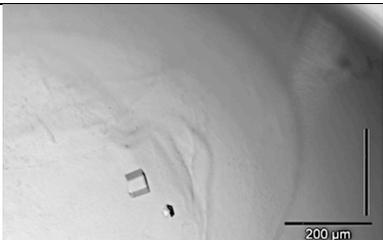
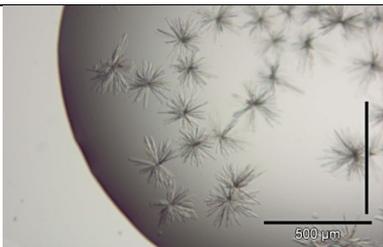
Saccharomyces cerevisiae *cyt c* was used as a model protein for co-crystallization experiments with *sclx*₄, *sclx*₈, *TPPS*₄ and the tris-indole ligand. Crystals were obtained for *cyt c* in the presence of three of the ligands.

Cyt c crystals grew readily in drops prepared with 10-fold excess *sclx*₄, for which PEG was the sole precipitant. Optimisation of these conditions yielded the growth of high quality diffraction crystals (Table 1). Subsequently, co-crystallization experiments extended to an Arg13Glu (R13E) mutant of *cyt c* (Figure 3). Co-crystallization trials of *cyt c* R13E and *sclx*₄ were prepared with the known crystallization conditions for the wild type. Significantly larger crystals of the mutant, which were visible by eye, grew rapidly (overnight) in a broader PEG range of 18-32 %. The size of the crystals and ease of crystallization suggested that the mutation and, in turn, the decrease in the net charge of the protein, facilitated assembly. The crystal structures for *cyt c* and *cyt c* R13E with *sclx*₄ were solved (Chapter 2).

*Sclx*₄ co-crystallization experiments were also carried out with hen egg white lysozyme (Table 1). High quality diffraction crystals of native lysozyme grew from drops containing similar conditions to *cyt c-sclx*₄ crystals and the structure was solved (Chapter 4). The lysine residues and the N-terminus were dimethylated in lysozyme, to yield lysozyme-KMe₂.¹⁷⁸ Crystallization trials were repeated with *sclx*₄ and lysozyme-KMe₂

crystals using the known *cyt c*-*sclx*₄ co-crystallization conditions. Crystals grew within minutes in a significantly broader PEG range, 0-35 %, and at different ionic strengths. The crystal structure was solved (Chapter 5).

Table 1. *Sclx*₄-protein co-crystals and the conditions used.

	<p>Cyt <i>c</i>-<i>sclx</i>₄: Crystals of <i>cyt c</i> grew in the presence of <i>sclx</i>₄. The crystallization drop was prepared by combining 1 μL of reservoir solution, 1 μL of 1.7 mM <i>cyt c</i> and 1 μL of 17 mM <i>sclx</i>₄. The reservoir solution contained 26 % PEG 8000, 50 mM NaCl, 100 mM MgCl₂ and 50 mM (CH₃)₂AsO₂Na pH 6. 200 μL of paraffin oil was placed on top of the reservoir solution.</p>
	<p>Cyt <i>c</i> R13E-<i>sclx</i>₄: Large crystals of <i>cyt c</i> R13E grew in the presence of <i>sclx</i>₄. The crystallization drop was prepared by combining 1 μL of reservoir solution and 2 μL of the complex solution, prepared by combining equal volumes of 1.8 mM <i>cyt c</i> R13E and 17 mM <i>sclx</i>₄. The reservoir solution contained 20 % PEG 8000, 50 mM NaCl, 100 mM MgCl₂ and 50 mM (CH₃)₂AsO₂Na pH 6. 200 μL of paraffin oil was placed on top of the reservoir solution.</p>
	<p>Lysozyme-<i>sclx</i>₄: Cubic crystals of lysozyme grew over 4 weeks. The crystallization drop was prepared by combining 1 μL of each the reservoir solution, 1.4 mM lysozyme and 17 mM <i>sclx</i>₄. The reservoir solution contained 24 % PEG 8000, 50 mM NaCl, 100 mM MgCl₂ and 50 mM (CH₃)₂AsO₂Na pH 6.3.</p>
	<p>Lysozyme-KMe₂-<i>sclx</i>₄: Plates grew in clusters overnight in similar conditions to the <i>cyt c</i> and native lysozyme-<i>sclx</i>₄ crystals. The crystallization drop was prepared by combining 1 μL of each the reservoir solution, 0.5 mM lysozyme and 4.3 mM <i>sclx</i>₄. The reservoir solution contained 18 % PEG 8000, 50 mM NaCl, 100 mM MgCl₂ and 50 mM (CH₃)₂AsO₂Na pH 6.3.</p>

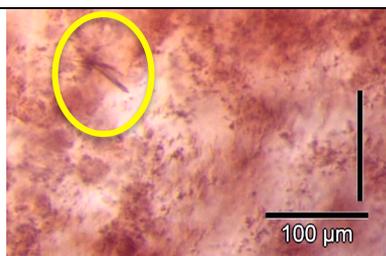
*Sclx*₈ did not co-crystallise *cyt c* in known crystallization conditions for *sclx*₄. Reducing the *sclx*₈ concentration to 5-fold excess of the protein concentration, to compensate for the net charge of the ligand, did not yield crystals. Increasing the ionic strength, changing the buffer and further decreasing the *sclx*₈ concentration, promoted

crystallization. Over several weeks, thin plates grew from a single nucleation site (Table 2).

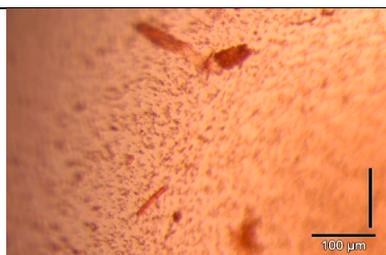
Crystal growth was most challenging with TPPS₄. Combining equal concentrations of the ligand and cyt *c* caused shock precipitation during crystallization drop setup. Despite changing the reservoir solution components, the ligand concentration had to be greatly reduced to prevent the precipitation of cyt *c* and favour crystallization (Table 2). However, defined protein crystals were not obtained.

The known cyt *c*-sclx₄ crystallization conditions were also used for preliminary co-crystallization trials for cyt *c* and tris-indole. However, the conditions yielded two distinct volumes. A red gel was observed at the centre of the drops surrounded by a halo of crystallization solution. Decreasing the ligand concentration resulted in microcrystalline precipitates. Changing the buffer and altering the salt components resulted in the growth of clusters of needles in drops containing 25-150 mM NaCl (Table 2). The needles did not diffract.

Table 2. Crystals and precipitates obtained for cyt *c* co-crystallized with sclx₈, TPPS₄ and tris indole.



Cyt *c*-sclx₈: Dense semi-crystalline precipitates appeared after ~48 hours from which small plates of cyt *c* grew over 6 weeks. The crystallization drop was prepared by combining 1 μL of reservoir solution and 2 μL of complex solution, prepared by combining equal volumes of 1.0 mM cyt *c* and 2.0 mM sclx₈. The reservoir solution contained 25 % PEG 8000, 50 mM NaCl, 200 mM MgCl₂ and 50 mM Tris-HCl pH 7.



Cyt *c*-TPPS₄: Crystalline precipitates obtained during co-crystallization trials. The crystallization drop was prepared by combining 1 μL of each the reservoir solution, 1.0 mM cyt *c* and 0.4 mM TPPS₄. The reservoir solution contained 50 mM NaCl and 50 mM Tris-HCl pH 8.



Cyt *c*-tris-indole ligand: Clusters of needles obtained after 2 weeks. The crystallization drop was prepared by combining 1 μL of reservoir solution and 2 μL of the complex solution. The complex solution contained equal volumes of 0.5 mM cyt *c* and 0.4 mM tris-indole ligand. The reservoir solution contained 40 % PEG 8000, 150 mM NaCl and 50 mM NaAc pH 4.5. 200 μL of paraffin oil was placed on top of the reservoir solution.

Discussion

Obtaining structures of protein-ligand complexes is challenging. The first task is to identify the conditions that yield crystals. These conditions then provide a starting point for optimizing crystal growth so that diffraction quality crystals can be obtained. The number of variable parameters for obtaining protein crystals is extensive and the greatest ambition is to arrive at crystal growth using the least amount of time and material. PEG is a useful reagent for preliminary crystallization trials¹⁵² and yielded *cyt c* crystals in the presence of *sclx*₄. Interestingly, the conditions that yielded high quality diffraction crystals of *sclx*₄ with *cyt c* also produced crystals of *sclx*₄ with *cyt c* R13E, lysozyme and lysozyme-KMe₂. Therefore, *sclx*₄ can crystallise proteins of similar size and pI (Table 3). Despite each ligand having complementary features for binding the surface of *cyt c*, the *cyt c*-*sclx*₄ co-crystallization conditions did not transfer to the three ligands and new co-crystallization conditions were required.

Table 3. Summary of the co-crystallization and structure solution for each ligand.

Protein	Ligand	Crystals	Diffraction Quality	Structure Solved
<i>Cyt c</i>	<i>Sclx</i> ₄	Yes	Yes	Yes
<i>Cyt c</i>	<i>Sclx</i> ₈	Yes	No	No
<i>Cyt c</i>	Tris-indole	Yes	No	No
<i>Cyt c</i>	TPPS ₄	No	No	No
<i>Cyt c</i> R13E	<i>Sclx</i> ₄	Yes	Yes	Yes
Lysozyme	<i>Sclx</i> ₄	Yes	Yes	Yes
Lysozyme-KMe ₂	<i>Sclx</i> ₄	Yes	Yes	Yes

All of the ligands have an aromatic core and are decorated with an acidic periphery. Crystals obtained for *sclx*₄ and the tris-indole ligand, were suitable for testing via X-ray diffraction (Table 3). *Sclx*₄ best supported crystal growth while TPPS₄ was the most challenging to co-crystallise. *Sclx*₄, *sclx*₈ and the tris-indole ligands all share a similar feature in that they possess, or can adopt, a cavity to contain a guest. *Sclx*₄ has the least flexibility. Therefore, co-crystallization with *sclx*₄ is believed to have the least entropic cost. Conversely, *sclx*₈ and the tris-indole ligand are extremely flexible. For example, the crystal structures of *sclx*₈ show two extreme conformations, one of which is an almost planar macrocyclic ring and the other displays an up-down double cone conformation.⁸³ TPPS₄ displays a high level of rigidity, however, the molecule lacks a cavity to entrap side chains. Therefore, both an acidic periphery and a pre-defined cavity are influential for co-crystallization with *cyt c*.

Chapter 2:

The complexes of cytochrome *c* and *p*-sulfonatocalix[4]arene

This work was published in part as McGovern, *et al.* Protein camouflage in cytochrome *c*-calixarene complexes. *Nature Chemistry* **2012**, 4, 527-533.

Abstract

Cationic residues of proteins present crucial markers in numerous cellular recognition events such as chromatin remodelling. As a result, efforts have focussed on the identification of synthetic molecules that target these residues. Calixarenes present a rich class of structurally well-defined molecules that can be conformationally fixed and functionalized in countless ways. Although these molecules have been studied for several decades now, with applications suggested in medicinal diagnostics and therapy, very little is known of their complexation ability with proteins. In this study we showed that the water-soluble *p*-sulfonatocalix[4]arene binds to cytochrome *c* by using NMR spectroscopy. The calixarene was observed sampling lysine-rich regions on the protein in a dynamic manner. X-ray crystallographic analysis of the co-crystal confirms the presence of distinct binding locations for the calixarene on the protein surface. Each site involved the entrapment of a lysine side chain. In addition to providing valuable information on protein recognition, the data also indicate that the calixarene is a mediator of protein–protein interactions. Thus, suggesting potential applications for generating assemblies and promoting crystallization.

Introduction

Ligands that mediate protein interactions are essential for a wealth of applications such as programming cells, biosensor design and promoting protein crystallization.^{120,179,180} Protein complexation by small molecules is particularly useful for identifying the point(s) of functionality, such as the sites that determine protein solubility. Studying these interactions at the molecular level allows a detailed assessment of the basic principles that trigger the recognition events.

Protein folds, such as α -helical recognition clefts, present a prominent feature that is readily targetable for the design of high-affinity ligands.⁵⁴ Saccharide binding by integral membrane proteins involves the saccharide burying into a binding pocket.¹⁸¹ Many protein surfaces, however, lack such binding pockets. These surfaces are large, flat and relatively featureless, thus rendering the design of specific, high affinity ligands challenging. Nonetheless, the difficulties associated with the identification of protein-surface binders can be overcome through the characterisation of protein-protein interfaces. For example, small-molecules that complement protein surface properties are emerging as important modulators of protein-protein interactions.^{3,108,119} Molecules that specifically target small epitopes, such as lysine and arginine side chains, also have potential as generic surface binders.^{43,122,182} Lysine recognition in particular, is important considering its susceptibility to posttranslational modification, crucial for chromatin remodelling.¹⁸³

Calix[n]arenes possess numerous protein binding possibilities via the upper and lower rims, in addition to being able to bind a guest in the cavity.^{89,94,110,113} Sclx₄ can complex various amino acids and peptides.^{89,93,100} The crystal structure of sclx₄ with lysine is of particular importance.⁹⁰ The sclx₄ complexes are defined by the positioning of the amino acid side chain in relation to the calixarene cavity. For example, the complexes formed by the inclusion of a lysine side chain in the cavity of sclx₄ are described *endo* (Figure 1, lysines coloured orange). Here, the *endo* bound molecule makes a salt bridge interaction with two sulfonates on the calixarene, which are 2.9-3.1 Å from the N^ε of lysine. The lysine carboxyl group makes van der Waals contact with three sulfonates, two of which belong to the same sclx₄. When the interaction occurs with the exterior of the calixarene, the binding is *exo*. In the same crystal structure, two lysine molecules interact *exo* to the calixarene. One lysine does so by traversing an aromatic ring of sclx₄ such that the N^ε makes a hydrogen bond with an sclx₄ hydroxyl group (Figure 1, lysines coloured

blue). The amino and carboxyl groups are each 3.2 and 4.3 Å to the nearest sulfonate. The side chain of the second lysine molecule is fully extended and is above the plane to two sulfonates on the sclx₄ (Figure 1, lysines coloured green). Now, the N^ε makes a salt bridge with two sclx₄ sulfonates and the terminal amino and carboxyl groups are 3.0-3.1 Å from another sulfonate.

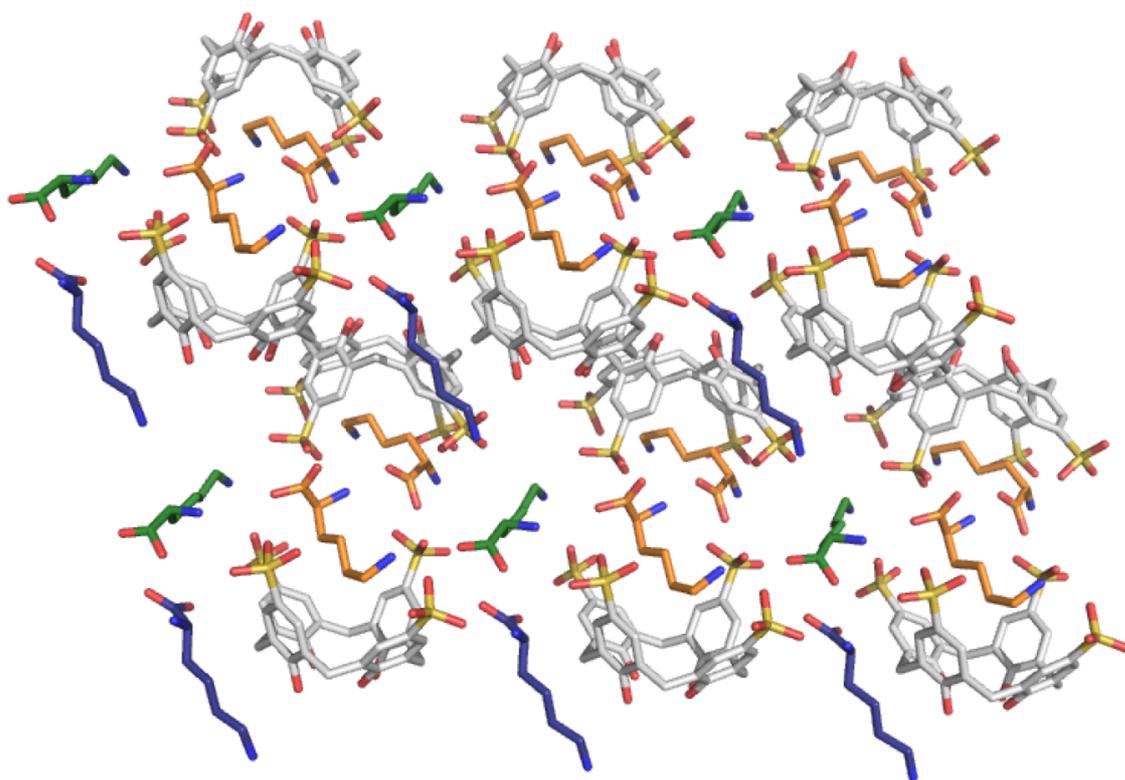


Figure 1. Sclx₄-lysine crystal structure showing amino acid binding *exo* and *endo* to the calixarene.⁹⁰ The *endo* interaction is characterised by the enclosure of a lysine side chain (orange) in the cavity of the calixarene. The lysine N^ε makes salt bridges with two sclx₄ sulfonates from within the cavity. Two lysine molecules bind *exo* to the calixarene. These amino acids are fully extended and traverse the plane of a sclx₄ aromatic ring (blue) or two adjacent sulfonates (green). One lysine N^ε makes a hydrogen bond with a sclx₄ hydroxyl group and the other makes a charge-charge interaction with the sulfonates on two sclx₄.

Although several labs have studied protein-calixarene interactions,^{101,104,105,110,184} a detailed solid-state characterisation of such a complex had not been reported. X-ray crystallography and NMR spectroscopy were used to investigate the complexes formed between sclx₄ and *Saccharomyces cerevisiae* cytochrome *c* (cyt *c*). The protein was chosen as a model protein due to its surface features, which are akin to histones. The lysine-rich surface of cyt *c* lacks features such as a binding cleft, which could accommodate a small molecule binder. The interaction properties of cyt *c* with proteins

have been extensively characterised on account of its importance for redox processes and apoptosis.^{14,128,185,186} Cyt *c* has mM- μ M affinity for different protein partners and, as a result, is considered a pseudospecific binder.¹⁴ The interactions of cyt *c* with small molecules have been extensively characterised.^{105,118,187,188} Anionic porphyrins make dynamic complexes with cyt *c* by having low specificity and affinity for the protein surface.¹¹⁹ Carboxylphenylcalix[4]arenes make 1:1 complexes with cyt *c* and the affinities that increase 16-fold for calixarenes possessing four rather than two carboxyl groups.¹⁰⁵ Here, we demonstrate at atomic resolution how the simple sclx₄ binds cyt *c* and camouflages the protein resulting in altered interaction properties.

Results and Discussion

Solution state characterisation of cyt *c*-sclx₄ complexes

Isotopically enriched cyt *c* was titrated with μ L aliquots of sclx₄ and monitored by 2D ¹H, ¹⁵N heteronuclear single quantum coherence (HSQC) spectroscopy. Notable chemical shift perturbations occurred during the titration with sclx₄. Increasing chemical shift perturbations as a function of the sclx₄ concentration were observed and indicated that ligand-free and ligand-bound forms of cyt *c* were in fast exchange on the NMR time scale. Of 37 resonances with a significant perturbation (¹H^N > 0.04 ppm or ¹⁵N > 0.4 ppm), nine are assigned to lysine residues, suggesting a preference for binding by this side chain. The largest perturbation (0.30 ppm at 50 mM NaCl) was observed for the ¹H^N of Lys87 (Figure 2A). Mapping the chemical shift changes to the surface of cyt *c* revealed a broad patch in the region of the known protein-binding site (Figure 2B).¹²⁸

In addition to providing structural information, the NMR data was used to calculate binding affinities. For resonances that experienced a significant chemical-shift change binding curves were prepared by plotting $\Delta\delta$ as a function of the ligand concentration for a total of sixteen ¹H^N resonances. Both hyperbolic and sigmoid curves were observed (Figure 2) and reliable fits were generated using models that accounted for cooperative binding.¹¹³ The Hill coefficients (*n*) were calculated to determine the degree of cooperativity between the two (or more) sites. The results revealed two regions, one centred on Lys4 and the second focused on Lys87. The average Hill coefficient and binding affinity at the Lys4 site was *n* = 1.62 (\pm 0.04) and *K*_d = 0.82 (\pm 0.04) mM, respectively. A larger Hill coefficient (*n* = 2.45 \pm 0.08) and lower affinity of 1.58 (\pm 0.08)

was measured for the Lys87 site. The differing affinities and degrees of cooperativity provide evidence that *cyt c* has two or more *sclx*₄ binding sites.¹⁸⁹

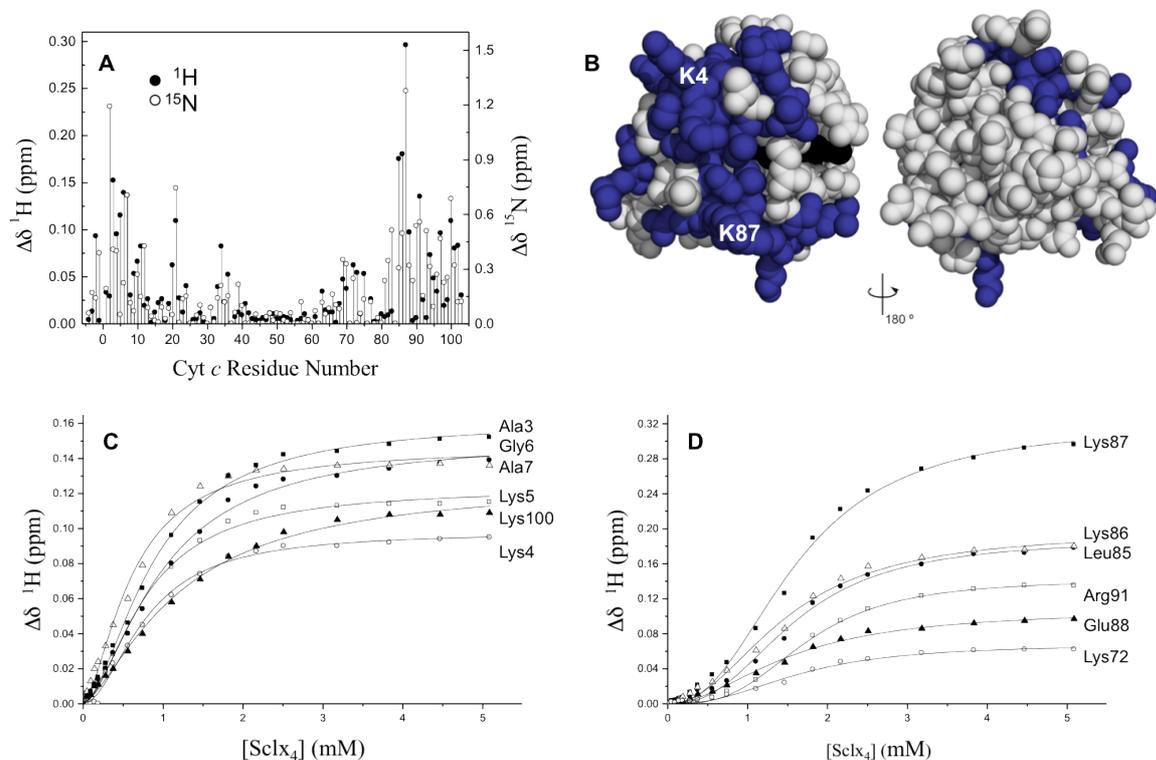


Figure 2. (A) Plot of the chemical shift perturbations measured for *cyt c* backbone amides in the presence of a 5-fold excess of *sclx*₄. *Cyt c* residues are numbered from -5 to 103. Blanks correspond to proline residues 25, 30, 71 and 76, and unassigned Gly84. (B) Space-filling representation of *cyt c* highlighting the *sclx*₄ binding surface. Residues for which the amide resonance had a significant chemical shift ($\Delta\delta$ ^1H > 0.04 or ^{15}N > 0.4 ppm) are coloured blue. The haem edge is black. Binding isotherms were obtained by plotting $\Delta\delta$ as a function of [sclx₄]. While all of the data were collected during a single titration there are clear differences in the curvature. (C) A cluster of residues in the vicinity of Lys4 have binding curves that appear hyperbolic while (D) a second cluster around Lys87 have sigmoid binding curves.

Co-crystallization and crystallography

The hanging drop vapour diffusion method was used for of *cyt c*-*sclx*₄ co-crystallization trials at 20 °C. Crystallization drops containing 5.7 mM *sclx*₄ in 10-fold excess of *cyt c* were prepared with poly(ethylene glycol) (PEG) as the sole precipitant. These conditions yielded red crystalline precipitates from which *cyt c* crystals grew in 26-30% PEG over two weeks (Figure 3). Preliminary diffraction experiments showed that the crystals diffracted (to 5 Å) and were anisotropic. The introduction of cacodylate buffer, NaCl and

MgCl₂ into the crystallization conditions helped the growth of more uniform crystals and an improvement in the physical appearance of the crystals was evident. Although the diffraction quality of the crystals improved, the crystals were still anisotropic. An oil barrier prepared with 200 μL of paraffin oil was required to obtain isotropic, diffraction quality crystals. Control drops lacking sclx₄ were also prepared. Importantly, cyt *c* remained fully soluble in the control, indicating that the calixarene was necessary in the crystallization event.



Figure 3. Cyt *c* crystals grown with 10-fold excess of sclx₄. (A) First hits were obtained in drops prepared with 30 % PEG 8000. The presence of (B) 50 mM sodium cacodylate pH 6-7, in addition to (C) 50 mM NaCl and (D) 100 mM MgCl₂ significantly improved the crystallization. (E) An oil barrier was necessary for obtaining diffraction quality crystals.

High quality diffraction data were obtained for the cyt *c* crystals and solved with the *P*2₁2₁2₁ space group to a resolution of 1.4 Å (Table 1). The unbiased electron density (calculated with phases prior to the addition of the ligand to the model) clearly showed the presence of sclx₄ (Figure 4). The structure was solved with two protein molecules and three sclx₄ (Figure 5). The two protein molecules are labelled chains A and B. The prominent feature of the crystal structure is that all three sclx₄ bind the cyt *c* surface via a lysine side chain *endo* to the cavity. This observation supports the NMR data showing the presence of two or more binding locations for the calixarene on the protein surface and a preference for lysine residues.

Cyt *c* R13E-sclx₄ crystallization and X-ray crystallography

After the identification of suitable crystallization conditions for cyt *c* with sclx₄, crystallization trials were prepared with the Arg13Glu (R13E) mutant of cyt *c*. Large crystals of the mutant, visible by eye, grew rapidly (overnight) in a broader PEG range of 18-32 %. The size of the crystals and ease of crystallization suggested that the decrease in the net charge of the protein facilitated assembly.

High quality diffraction data were also obtained for the *cyt c* R13E crystals, and a structure was solved to a resolution of 1.1 Å with the same space group as the wild type (Table 1). Similar to the wild type, the unbiased electron density, which was calculated with phases prior to the addition of *sclx*₄ to the model, indicates the presence of *sclx*₄ (Figure 4). Both *cyt c* and *cyt c* R13E share the same protein assembly and *sclx*₄ binding sites, despite the reduction in net positive charge (Figure 5).

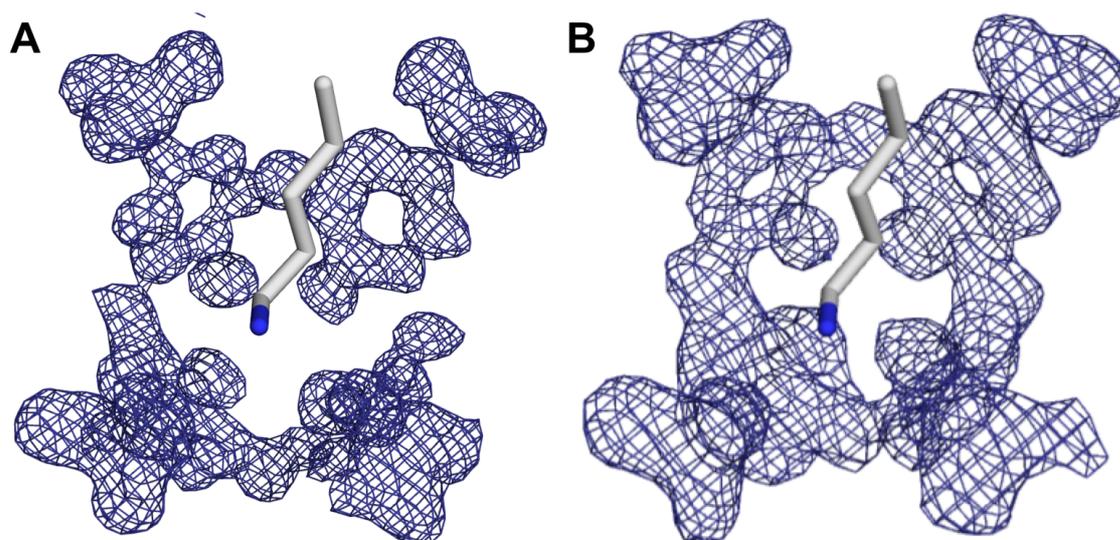


Figure 4. Unbiased electron density maps indicate the presence of a *sclx*₄ molecule at A.Lys89. **(A)** The maps were calculated at a resolution of 1.1 Å (for *cyt c* R13E) and **(B)** 1.4 Å (for the wild type) before the addition of *sclx*₄ to the *cyt c* model and contoured at 2σ . The side chain of Lys89 is shown as sticks.

Table 1. Crystallographic data for cyt *c*-sclx₄ complexes

	Cyt <i>c</i> -sclx ₄	Cyt <i>c</i> R13E-sclx ₄
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell constants	a = 36.04 Å	a = 36.11 Å
	b = 55.32 Å	b = 55.78 Å
	c = 117.42 Å	c = 119.08 Å
	α = β = γ = 90°	α = β = γ = 90°
Resolution, Å	30.19-1.40 (1.42-1.40)	15.39-1.05 (1.08-1.05)
Wavelength, Å	1.13	0.87
Unique reflections	47,260	113,051
Multiplicity	4.3 (2.3)	6.0 (5.4)
<i>I</i> /σ	15.8 (2.4)	9.1 (1.8)
Completeness, %	98.7 (90.6)	94.9 (100)
<i>R</i> _{merge} , ^b %	4.4 (29.9)	11.3 (96.2)
Solvent content, %	45.4	46.70
<i>R</i> _{factor} , %	17.7	15.70
<i>R</i> _{free} , %	19.9	17.98
rmsd ^c bonds, Å	0.01	0.01
rmsd ^c angles, °	1.24	1.27
# molecules in asymmetric unit		
Protein	2	2
Ligand	3	3
Solvent	395	348
Average <i>B</i> factors, Å ²		
Protein	14.32	9.38
Ligand	19.80	13.86
Solvent	30.19	21.00
Ramachandran analysis ^d		
% residues in favoured regions	98.1	97.64
% residues in allowed regions	100	100

^aValues in parentheses correspond to the highest resolution shell; ^b $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$; ^croot mean square deviation calculated from the *B* values of all non-hydrogen protein atoms; ^dCalculated with MolProbity.¹⁷⁶

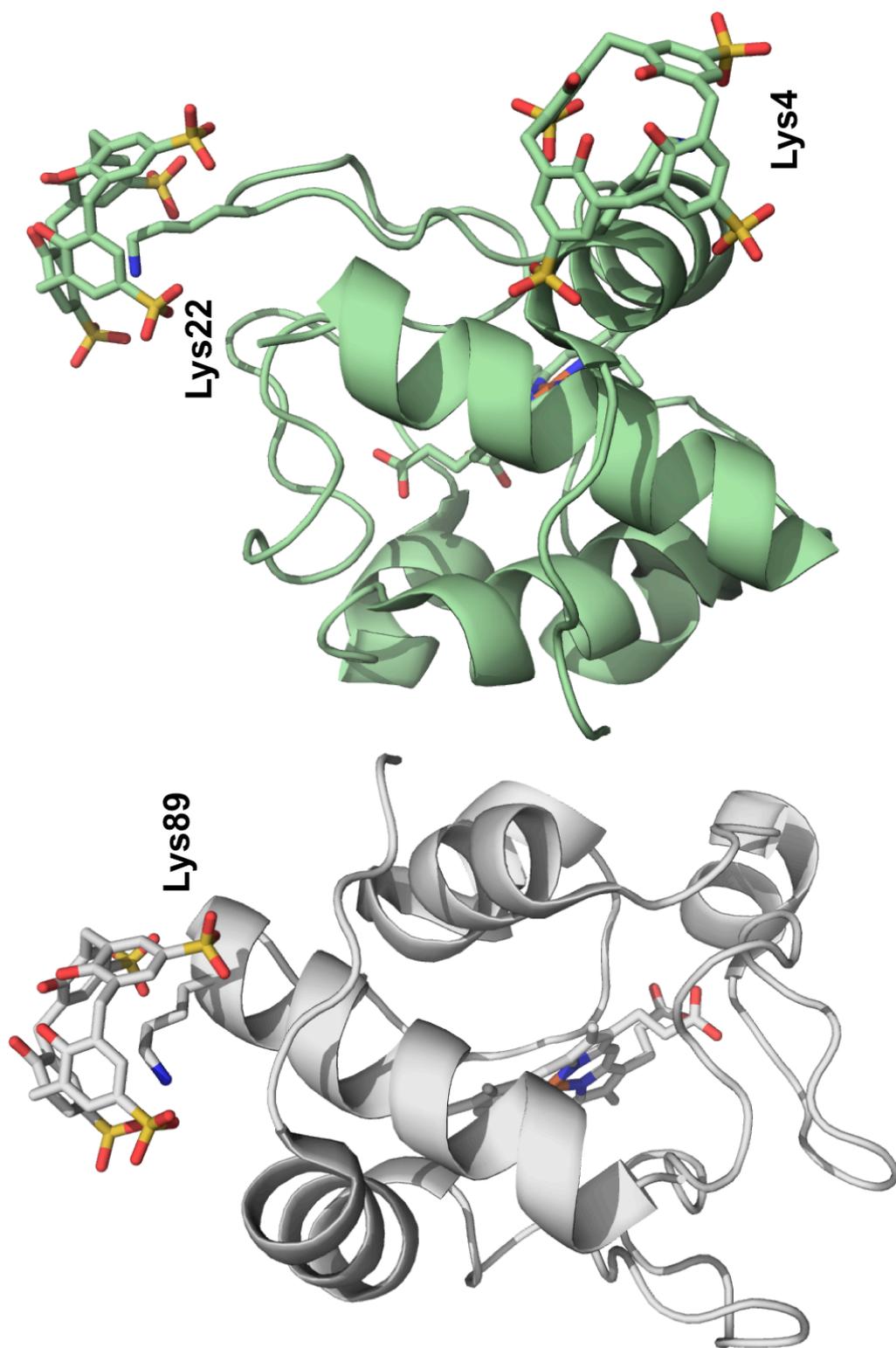


Figure 5. Crystal structure of the cyt *c*-sclx₄ complex. The asymmetric unit comprises two molecules of cyt *c* (chains A and B in grey and green, respectively) and three molecules of sclx₄. The haem groups are shown as sticks.

An overview of the lysine-sclx₄ binding sites

All three calixarene binding sites share similar features as they each entrap a lysine side chain. One sclx₄ binds Lys89 in chain A and the other two sclx₄ bind Lys4 and Lys22 in chain B (Figure 5). The C^γ, C^δ and C^ε atoms of the side chain are confined in the calixarene cavity while the amino group projects up into the plane of the sulfonates. The lysine amino group and the other polar groups of the protein make at least one noncovalent bond with the sulfonate bearing rim, and are described in detail below. The architecture of the sclx₄–lysine complex mimics the structural features of the solid state complex with the amino acid lysine (Figure 1) described in the introduction.⁹⁰ *P*2₁2₁2₁ symmetry operations were applied to the asymmetric unit of the cyt *c* crystal structure to generate assemblies and examine the contribution made by crystal packing neighbours. At each binding site sclx₄ was present at protein-protein interfaces making contacts with 2-4 cyt *c* molecules. In all binding sites at least two lysines are contributed by neighbouring molecules and bind *exo* to sclx₄.

Water molecules are observed in abundance at the sclx₄ binding sites and appear to play a role in the coordination environment of the lysine side chains. All three entrapped lysine side chains are partially solvated. A total of 16 water molecules make van der Waals contact with the sclx₄ sulfonates and hydroxyls. The water molecules play a significant role in bridging interactions between the polar groups of sclx₄ and cyt *c* via hydrogen-bonding. Overall, the cyt *c*-sclx₄ interfaces are charge rich and heavily solvated.

Lys89-sclx₄ binding site

The sclx₄ binding site at Lys89 in chain A (A.Lys89) involves numerous noncovalent interactions. In the bent conformation, Lys89 makes intermolecular salt bridges with two sclx₄ sulfonates (N^ε-O 2.8 and 2.9 Å) and an intramolecular salt bridge with Asp93 (Figure 6A). The backbone amide of Lys89 and the side chain -NH₂ of Asn92 each make a hydrogen bond with one sulfonate. Lys5 also makes a salt bridge with a sulfonate. Other noncovalent bonds include cation- π interactions with the C^ε of Lys89. The distances from the C^ε to all of the carbon atoms in the three closest phenyl rings are 3.6-4.5 Å. As part of gene expression, cation- π interactions are crucial to the binding of methylated lysine side chains by the aromatic pockets of chromodomains.^{190,191} Lysine binding by the calixarene displaces the water molecules that occupy the calixarene cavity.^{85,192} Therefore, the hydrophobic effect may be expected to contribute to binding. Furthermore, the entropy

gain of the liberated waters may be counterbalanced by the entropy loss of the confined lysine side chain. A total of three lysine side chains, contributed by chain A, make water-mediated salt bridges to the sulfonates. Alternative conformations are observed for the side chains of Lys-2 and Lys5. Both conformers of Lys5, but neither of the Lys-2 conformers, make a salt bridge with a sulfonate. These observations suggest a counterbalance between the different energetic contributions to binding.

Another interesting feature of this binding site is that the calixarene is positioned between two protein chains. Chain A of a crystal packing neighbour contributes Lys4 and Lys100 to the interaction site (Figure 6B). These lysine side chains traverse the aromatic region of sclx_4 , are within van der Waals distance and both N^ϵ are positioned at sulfonates allowing a direct or water-mediated salt bridge. The C^δ and C^ϵ of Tyr97 offer a partial positive charge to the A.Lys89 site. All four hydroxyls at the A.Lys89 binding site make van der Waals contacts with the methyl side chain of an alanine residue. This is an interesting feature as it might be expected that polar side chains would exclusively bind to the sclx_4 hydroxyl rim, which has one deprotonated hydroxyl group.⁸⁵ The crystal structure indicates otherwise.

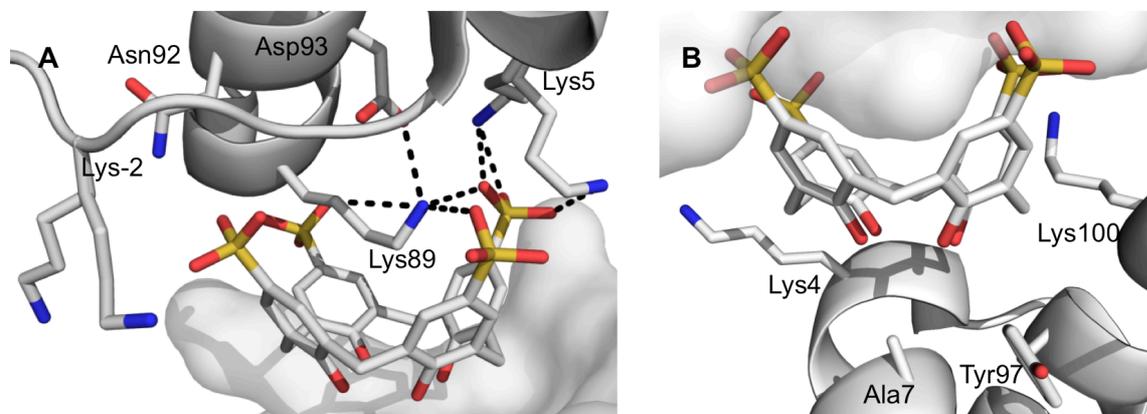


Figure 6. Two perspectives of the A.Lys89 binding site. **(A)** The intramolecular and intermolecular salt bridges Lys89 makes to Asp93 and to two sclx_4 sulfonates, respectively, are shown. Additional salt bridges with sclx_4 are shown. The crystal packing neighbour (also chain A) that surrounds the calixarene is shown as a transparent surface. **(B)** Residues (shown as sticks) from the chain A crystal packing neighbour (cartoon) offer additional noncovalent interactions to the site. Water molecules are omitted for clarity.

Lys4- and Lys22-sclx₄ binding sites

Sclx₄ binds B.Lys4 side chain in a similar manner to A.Lys89. The Lys4 N^ε makes a hydrogen bond to a water molecule, in addition to making a salt bridge to a sulfonate. The backbone amide of Lys4 makes a hydrogen bond to another sulfonate and the side chain -OH of Ser2. The interaction network is, in part, mediated by Ser2 which makes hydrogen bonds with a sulfonate and the backbone amide of Lys5 through its side chain -OH. Lys5 and Lys100 offer additional noncovalent interactions with the sulfonates. The side chain N^ε of Lys5 makes a direct salt bridge with a sulfonate and a water molecule, 3.3 Å from the side chain N^ε of Lys100, completes the contacts with two sulfonates (3.0 and 3.4 Å apart).

Three additional protein chains assemble around the sclx₄ (Figure 7A) such that approximately 30 % of the surface area of the calixarene is solvent accessible. These crystal packing neighbours contribute 5 lysine side chains to the calixarene binding site. Interestingly, Lys54 and Lys73 make van der Waals contact from the alkyl portions of the side chains and the sclx₄ phenyl rings. The C^ε of Lys54 makes a salt bridge with a sulfonate. The sclx₄ hydroxyl-bearing rim is heavily solvated and makes van der Waals contact with the carbonyls and one backbone amide of crystal packing neighbours. Although the side chain of crystal packing neighbours Lys55 is in the vicinity of the sclx₄ hydroxyls, the side chain N^ε makes a cation-π interaction with the phenyl side chain of Tyr74.

The noncovalent interactions are less numerous at B.Lys22-sclx₄, in comparison to the other two binding sites. This binding site involves a salt bridge between the lysine side chain N^ε and two sclx₄ sulfonates (3.3 and 3.6 Å apart). The side chain N^ε of Lys22 is also solvated by three water molecules and the backbone amide makes hydrogen bond with one sulfonate. Only His33 offers additional interaction with sclx₄. The His33 side chain N^{ε2} makes a salt bridge to a sulfonate. Crystal packing neighbours offer substantially more noncovalent bonds as sclx₄ interacts with three additional cyt *c* molecules (Figure 7B). Three lysine side chains interact via cation-π, hydrophobic effects or a water-mediated salt bridge. As observed at the A.Lys89 binding site, all sclx₄ hydroxyl groups make van der Waals contacts with the methyl side chain of an alanine residue at the B.Lys22 site. In general, the protein-calixarene contacts are fewer in chain B than in chain A, which has an interface ~100 Å² larger (Table 2).

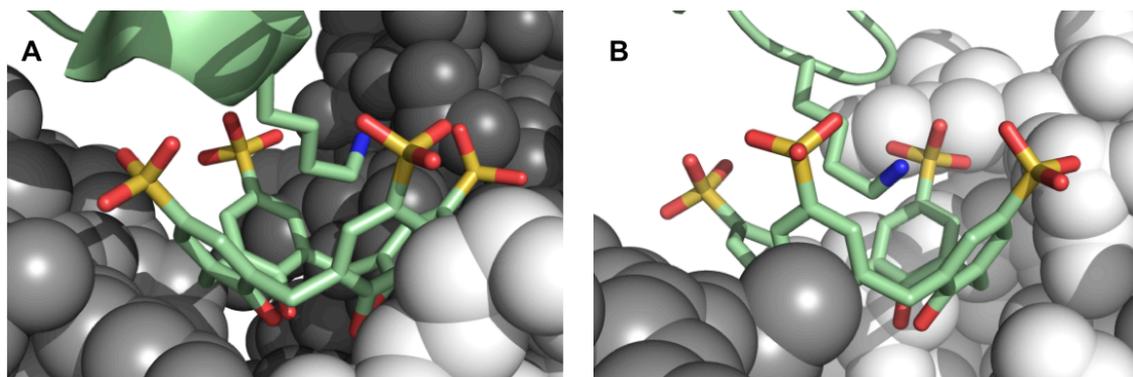


Figure 7. (A) Crystal packing neighbours at B.Lys4 involves three additional protein chains. (B) Crystal packing at B.Lys22 involves two additional protein chains. Chain B is represented as green cartoons and the lysine side chain and sclx₄ are sticks. Crystal packing neighbours are grey spheres.

Table 2. Interface size and polar interactions in cyt *c*-sclx₄ complexes.

Protein site*	sclx ₄	Interface	Total	Hydrogen bonds [§]	Salt bridges [¶]		Cation- π [†]	Lysine	Water molecules [#]
		area (\AA^2) [†]	area (\AA^2) [‡]		Direct	Water			
A.Lys89	1	320	520	2	3	2	2	4	16
A1		200		-	1	1		2	
B.Lys4	2	230	560	2	1	3	3	2	13
B1		160		1	-	-		1	
A2		140		-	-	-		1	
A3		30		-	-	-		-	
B.Lys22	3	210	590	-	2**	-	2	1	11
B2		200		1	-	-		2	
B3		180		-	-	1		1	

* $P2_12_12_1$ symmetry operations used to generate the neighbouring protein chains: A1, $x-1/2, -y+1/2, -z$; B1, $-x, y-1/2, -z-1/2$; A2, $-x, y-1/2, -z-1/2$; A3, $-x+1/2, -y, z-1/2$; B2, $x-1, y, z$; B3, $-x-1, y-1/2, -z-1/2$. [†]Surface area (reported to the nearest 10 \AA^2) of the protein or ligand that becomes inaccessible in the complex. [‡]Surface area of sclx₄ that becomes solvent inaccessible in the total complex, which includes all of the neighbouring protein chains. [§]Number of hydrogen bonds between sclx₄ and the protein. [¶]Defined as ‘Direct’ (Lysine N^ε to sulfonate O <3.5 \AA , unless otherwise stated) or ‘Water’ mediated, in which a water bridges the oppositely charged groups. [†]Total number of lysine side chains making cation- π with sclx₄. ^{||}Total number of lysine side chains in contact (including van der Waals) with sclx₄. [#]Number of water molecules hydrogen-bonded to sclx₄. ** One salt bridge involves His33 N^{ε2}.

Charge-charge interactions and the location of the sclx₄ binding sites

Charge-charge interactions were expected to have a significant part in the binding affinities as cyt *c* has a net positive charge (+11) and sclx₄ a net negative charge (-5) at pH 7. Despite a reduction in the net positive charge for cyt *c* R13E to +9 crystal growth was better for the mutant. Analysis of the crystal lattice shows Glu13 positioned at solvent channels, similar to Arg13 in the wild type. Leu9, Lys11, Thr12 and Phe82 surround the Arg13/Glu13 residue. The reduced positive charge of the mutation may facilitate the incorporation of cyt *c* R13E molecules into the lattice more easily than the wild type. The chemical shift perturbation experienced by Arg13 was below the assigned cut off in solution. Therefore, both the NMR data and the crystal structures suggest that the residue is not a major contributor for cyt *c* complexation by sclx₄.

NMR samples with cyt *c* were prepared in 20 mM phosphate buffer and 50 mM NaCl. Samples lacking NaCl displayed larger chemical-shift perturbations and lower dissociation constants, which had decreased by a factor of 1.6 (Table 3). Therefore, there is a charge-charge contribution to cyt *c*-sclx₄ complex formation. Sclx₄ sulfonates are known to coordinate Na⁺ ions⁸⁵ and this may have contributed to the effect. The incorporation of NaCl and MgCl₂ into the crystallization screen was essential for improving crystal growth. This may also be an electrostatic screening effect whereby the rate of protein complexation and calixarene-assisted self-assembly are reduced.

Although nine lysines, Arg91 and His33 had a significant perturbations, the acidic calixarene does not exclusively target basic side chains. In the crystal structure, A.Lys89 makes a salt bridge to Asp90 and is flanked by the acidic side chains of Glu88 and Asp93. However, Lys2, Lys4, Lys5 and Lys87 are within the Debye radius.¹⁹³ These side chains contribute an overall positive charge to the site. The Lys4 binding site in chain B offers the most positive charge owing to Lys5, Lys89, Lys100 and only one acidic residue (Asp93). The different charge distributions of the Lys4 and Lys87 binding sites correlates with the higher and lower binding affinities that were determined for each site, respectively.

Table 3: Salt dependence of the dissociation constant (mM).

Group	$^1\text{H}^{\text{N}}$ resonance	50 mM NaCl	0 mM NaCl
A	Ala3	0.91 (0.03)	0.55 (0.03)
	Lys4	0.78 (0.04)	0.53 (0.03)
	Ala7	0.60 (0.02)	0.35 (0.02)
	Phe10	0.75 (0.06)	0.43 (0.05)
	Lys100	1.08 (0.05)	0.65 (0.04)
B	Lys72	1.65 (0.15)	0.98 (0.12)
	Lys86	1.44 (0.05)	0.91 (0.03)
	Lys87	1.60 (0.03)	1.04 (0.04)
	Glu88	1.44 (0.09)	0.86 (0.08)
	Arg91	1.84 (0.08)	1.06 (0.06)

The K_d values (errors in parentheses) were determined according to Figure 2 for titrations performed in the presence or absence of NaCl. Groups A and B correspond to clusters of resonances with similar curves (Figure 2).

The location of the lysine residues is also expected to contribute to sclx_4 binding. Lys4 is positioned at the positive dipole of an α -helix. Although Lys5 and Lys100 immediately flank the residue, Lys4 may be described as being positioned at the apex of the three residues as it projects out furthest from the protein surface. The position of one of the sulfonates at the Lys4 site neatly coincides with a SO_4^{2-} anion observed in the original *cyt c* structure (Figure 8).¹³⁰ In both cases, a similar binding geometry is observed. Two hydrogen bonds are formed between the sulfonate/sulfate oxygens and the amide of Lys89 and the hydroxyl of Ser2. Additionally, the C^α of Ser2 makes a weak hydrogen bond with the nearest O atom on the sulfonate and sulfate, which are 3.6 Å and 3.4 Å apart, respectively. The backbone amide of Ala3 hydrogen bonds the sulfate and this interaction is lost the calixarene-binding site. Importantly, this sclx_4 /sulfate binding site is likely to be optimised for anion recognition as it possess the C^αNN motif that recognises sulfate ions.^{194,195} Lys87 is the first residue positioned at the positive dipole of its α -helix, similar to Lys4. Therefore, it may be considered as the most sclx_4 -accessible lysine side chain in the region. Although Lys-2, Lys5 and Lys89 flank Lys87, these residues can make intramolecular salt bridges with neighbouring Glu88, Asp09 and Asp93. A trade-off between intramolecular and intermolecular salt bridge formation may

be required to allow complexation by sclx_4 . Lys22 is flanked by His33, Arg38, Glu21 and C-terminal Glu103 and is positioned at a protein-protein interface in the original *cyt c* structure.¹³⁰ During the titration with sclx_4 Lys22 had a negligible perturbation (note that the measured chemical shift perturbation for Lys22 was below the cutoff; Figure 2). Therefore, binding at this site may be a feature of crystallization rather than a site optimised for anion recognition.

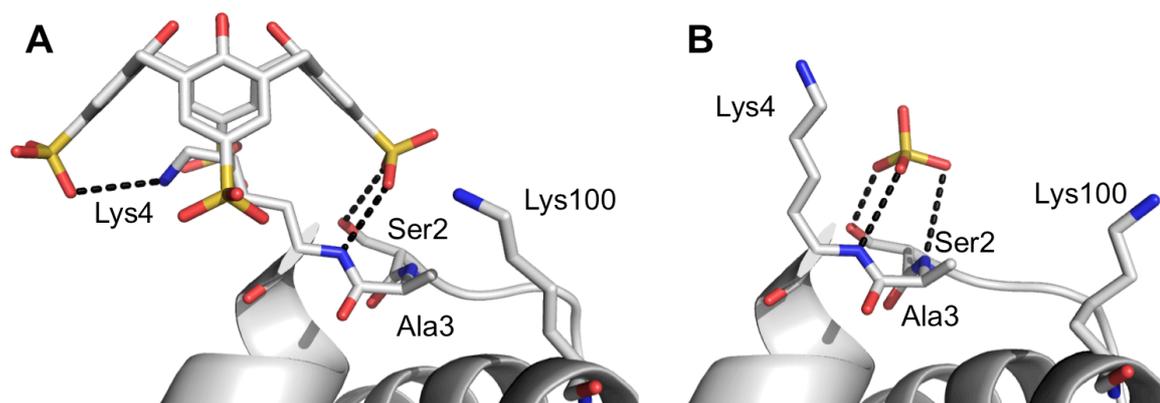


Figure 8. *Cyt c*-sulfate (PDB 1ycy) and sulfonate binding sites at Lys4 showing lysine side chain rearrangement. (A) One sclx_4 sulfonate at Lys4 coincides with the sulfate anion and makes two hydrogen bonds with the backbone amide of Lys89 and the hydroxyl of Ser2. Lys100 is involved in the binding, making a water-mediated salt bridge with the sulfonate. (B) The sulfate anion, 3.8 Å from the C^ε of Lys4, hydrogen bonds the hydroxyl of Ser2 and the backbone amides of Lys89 and Ala3. Lys100, pointing away from the site, is not involved in anion binding. The residues, sclx_4 and sulfate are represented as sticks.

Solid- and solution-state complexes.

Although the observations made by crystallography support the NMR data, there are notable differences in the experimental parameters as a result of the different techniques. Considering the dynamic nature and low interaction affinities, the sclx_4 -mediated *cyt c* assemblies, observed in the crystal lattice, cannot occur in solution. The contacts arose from crystallization conditions containing 26% PEG 8,000 and approximately 4-fold the ionic strength used for the NMR study. Therefore, as the protein concentration is less and there is no PEG, higher-order complexes are expected to have only a transient existence, if any, under the conditions used for NMR.

The presence of three different sclx_4 binding sites in the crystal structure raises the question of whether all sites can be occupied simultaneously in a single protein chain. Simultaneous binding of sclx_4 at Lys89 and Lys4 would involve van der Waals contact

between the sulfonate groups. Therefore, concurrent binding is hindered by electrostatic constraints. The Lys22 site can be occupied simultaneously with Lys89 or Lys4. From NMR analysis, the two clusters of residues around Lys4 and Lys87 indicate a cooperative process with tighter binding in the Lys4 region (Table 3). In the crystal structure the side chain -OH of Ser2 hydrogen bonds both the backbone amide of Lys4 and one sulfonate, possibly accounting for the high $\Delta\delta$ observed for this resonance. In solution it also appears that binding occurs at Lys87 rather than Lys89.

An analysis of all of the protein-sclx₄ contacts in the structure using crystal packing neighbours reveals a total of nine cyt *c*-sclx₄ complexes (Table 2, see Methods). Of these nine binding sites five calixarenes bind with contact areas of $\geq 200 \text{ \AA}^2$. This is consistent with the surface observed by NMR mapping. The broad patch suggests the dynamic nature whereby the sclx₄ might explore the surface of cyt *c* prior to adopting one or more stable configurations. This broad binding site is also observed in the cyt *c*-porphyrin complexes.¹¹⁹ Similar to the calixarene, 18-crown-6 interacts with lysine-rich peptides in dynamic manner moving between different binding sites on the peptide.¹⁹⁶

Conclusions

Small-molecule ligands that bind to and camouflage protein surfaces have considerable potential for modifying protein interaction properties. The molecular tweezers that binds a lysine side chain on the 14-3-3 α protein inhibits binding by two partner proteins.⁴³ Calixarenes and crown ethers alike have been used for protein separation by modifying protein surface properties.^{70,187} Cyt *c* can be solubilised in organic solvent by calixarene derivatives.^{103,104} Similarly, sclx₄ alters the solubility of cyt *c* allowing crystals to grow from PEG-only containing solutions. The crystal structure of the cyt *c*-sclx₄ complex shows the ligand binding the protein surface and, in turn, affecting its solubility. Similarly, calixarene derivatives are used as additives to promote membrane protein crystallization.¹⁹⁷

The presence of sclx₄ at protein-protein interfaces (Figure 7) supports the potential role as ‘molecular glue’. Crystals could not be obtained in drops lacking sclx₄. 18-crown-6, which is also found at protein-protein interfaces, entraps a lysine side chain, like the calixarene.⁴⁷ Pyrenetetrasulfonic acid and tetra(4-sulfonatophenyl)-porphyrin have previously been identified as having this ‘molecular glue’ role.^{40,50} The methylation of

lysine side chains or their mutation to alanine are two strategies used to promote the crystallization of otherwise tricky proteins.^{34,198} However, the noncovalent modification of protein surfaces to promote crystallization provides an alternative route for crystallising proteins.

The generation of multimolecular complexes via noncovalent interactions is important for the design of modulators of protein self-assembly. Sclx₄ is observed at protein-protein interfaces demonstrating the potential of the calixarene to generate protein assemblies. Similarly, phthalocyanine tetrasulfonate alters the aggregation of α -synuclein¹⁹⁹ and macrocycle cucurbituril induces tetramerisation of a fusion protein containing the N-terminal PheGlyGly motif.²⁰⁰

The calixarene makes low affinity interactions with two positively charged regions on the surface of cyt *c*. Using ¹H, ¹⁵N-HSQC experiments heparin sulfate was shown to bind the positively charged N-terminal region of the TSR domain of the TRAP protein.²⁰¹ In this system heparin is hypothesised to behave as a charged cluster and interacts with the protein surface with low affinity (0.5 mM). Sclx₄, which is conceptually related to the charged heparin cluster, might act as a molecular mimic for such heparin binding proteins.

Sclx₄ selectivity is expected to be higher on lysine-sparse surfaces or with proteins bearing methylated lysines, as they exhibit a greater affinity for sclx₄.⁹⁵ Reducing the symmetry of sclx₄ by substitution of one of the sulfonates has shown potential for altering guest affinities.⁷⁶ Reducing the sclx₄ symmetry may also be useful for increasing the specificity for protein surface binding. Alternatively, increasing the cavity size may enhance specificity. In this instance, the ligand might encapsulate a larger, more unique patch on the protein surface. The cyt *c*-sclx₄ complexes present a valuable template for tuning the specificity and affinities of protein-calixarene interactions, with applications extending from crystallization to histone recognition.

Experimental

Sclx₄ synthesis

The synthesis (by direct ipso-sulfonation of the tert-butyl-calix[4]arene) and characterisation of *p*-sulfonatocalix[4]arene⁷² was performed by Dr. Nicholas Power.

Protein production and purification

Saccharomyces cerevisiae cyt *c* and cyt *c* R13E were expressed in the cytosol of *Escherichia coli* BL21 (DE3), transformed with the vectors pBTR1 and pBTR1-R13E, respectively.^{127,202} The vectors contained the ampicillin resistance gene and expression cultures were supplemented with 75 µg/mL carbenicillin. Unlabelled cyt *c* and cyt *c* R13E were produced in Terrific Broth medium (12 g/L N-Z amine, 24 g yeast extract and 5 g/L NaCl) containing 20 mM sodium citrate and succinate salts.²⁰³ Cyt *c* and cyt *c* R13E were purified by ion exchange chromatography and size exclusion chromatography, according to published methods.^{127,204} The protein purity was assessed by SDS-PAGE and the concentration was determined by UV/Vis spectroscopy, assuming an extinction coefficient of $\epsilon_{550} = 27.5 \text{ mM}^{-1} \text{ cm}^{-1}$ for reduced cyt *c*. ¹⁵N-labelled cyt *c* was expressed and purified by Dr. Peter Crowley. Cyt *c* R13E was expressed and purified with Madeleine Mallon.

Protein crystallization

The hanging drop vapour diffusion method was used for all crystallization trials at 20 °C. Drops were prepared by combining 1 µL solutions of each oxidised cyt *c* (1.7 mM) and sclx₄ (17 mM) with the reservoir solution in a ratio of 1:1:1. Control drops were prepared by combining 1 µL solutions of oxidised cyt *c* (1.7 mM) and distilled water with the reservoir solution.

For the preliminary screen the reservoir solution contained only PEG 8000 and the concentration ranged from 10-50 %. Drops prepared with 30 % PEG yielded dense red crystalline precipitates, from which red crystals grew over a period of several days. The addition of 50 mM sodium cacodylate, pH 6 followed by 50 mM NaCl and 100 mM MgCl₂ to the crystallization solution significantly improved crystal growth and the diffraction quality of the crystals. A 200 µL oil barrier, prepared with paraffin oil, was applied to further improve crystal growth by slowing down the rate of diffusion, yielding isotropic, diffraction quality crystals.

Cyt *c* R13E-sclx₄ co-crystallization trials were prepared using the known crystallization conditions described above and 0-30 % PEG 8000. The experiment drops were prepared by combining 1 µL volumes of oxidised cyt *c* R13E (1.8 mM) and sclx₄ (17 mM) with the reservoir solution. Large red crystals, visible by eye, grew in drops containing ~20 % PEG and smaller crystals grew with ~30 % PEG. Combining equal

volumes of cyt *c* R13E (1.8 mM) and sclx₄ (17 mM) in a complex solution and subsequently preparing experiment drops with 1 μL of the reservoir solution and 2 μL of the complex solution, significantly improved the physical appearance of the crystals.

Data collection and X-ray structure determination

Cyt *c* crystals were transferred to a cryoprotectant solution containing 20 % glycerol and 80 % of a 26 % PEG 8000 solution and flash-frozen under a stream of nitrogen gas at 100 K (X-stream 2000). Intensity data were collected (φ scans of 1 ° over 200 °) to a resolution of 1.4 Å from a single crystal on beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France.

Cyt *c* R13E crystals were transferred to a cryoprotectant solution composed of 20 % glycerol and 80 % of the reservoir solution (18 % PEG 8000, 100 mM MgCl₂ and 50 mM sodium cacodylate pH 6) and flash-frozen under a stream of nitrogen gas at 100 K (X-stream 2000). Intensity data were collected (φ scans of 1 ° over 133 °) to a resolution of 0.9 Å from a single crystal on beamline ID23-2 at the ESRF, Grenoble, France.

All diffraction data was collected and processed with MOSFLM¹⁷⁰ and scaled with SCALA.¹⁷¹ Data collection and refinement statistics are given in Table 1. Initial phases for molecular replacement of the crystal structure were carried out using PHASER and the coordinates of cyt *c* (PBD ID 1ycc) were used as a search model. The coordinates of the chain A cyt *c* (PBD ID 3tyi) were used as a search model for cyt *c* R13E. REFMAC,¹⁷⁵ in the CCP4 suite, and COOT¹⁷⁵ were implemented for structure refinement and manual rebuilding, respectively. Crystal packing interfaces were analysed using the Protein Interfaces, Surfaces and Assemblies service (PISA) at the European Bioinformatics Institute²⁰⁵ and figures were prepared in PyMol.²⁰⁶

From the high quality diffraction data obtained for the cyt *c* crystals, sclx₄ is shown to bind Lys89 with 100 % occupancy and the other two binding sites with 85 %. Sclx₄ occupies all three sites of cyt *c* R13E with 100 % occupancy. This discrepancy is considered to be a product of the crystallization drop setup. Therefore, premixing the protein and sclx₄ prior to addition to the experiment drop, improves not only crystal growth but also the overall quality of the structure.

NMR spectroscopy

For NMR studies the typical cyt *c* concentration was 0.25 mM. For the high ionic strength studies 20 mM KPi, 50 mM NaCl, 0.1 mM sodium ascorbate (as a reductant) and 10 % D₂O at pH 6.0 were used, whilst for the low ionic strength studies 20 mM KPi, 0.1 mM sodium ascorbate and 10 % D₂O at pH 6.0 were used. Ascorbate was added to avoid paramagnetic effects of the haem iron in the NMR spectra. 2D ¹H, ¹⁵N HSQC spectra were obtained at 303 K with spectral widths of 14.1 ppm (¹H) and 40.0 ppm (¹⁵N) on a Varian 500 MHz NMR system spectrophotometer. The spectra were analysed in CARA (<http://www.nmr.ch/>) and the perturbations were normalized for each backbone resonance. The NMR experiments were designed and executed by Dr. Peter Crowley.

Binding isotherms

Binding isotherms were made by plotting the magnitude of the chemical shift change ($\Delta\delta$) as a function of the calixarene concentration. The data was fit (nonlinear least squares in Origin) to a one-site binding model, for which $\Delta\delta$ and [calixarene] were the dependent and independent variables, respectively. The dissociation constant and maximum chemical-shift change ($\Delta\delta_{\max}$) were the fit parameters. The curves were fit simultaneously to a single K_d value and $\Delta\delta_{\max}$ was varied for each resonance. For residues in which cooperative binding occurred Hill plots were made. Fitting was also performed using a Hill model,¹¹³ which includes a third parameter n , the Hill coefficient:

$$\Delta\delta = \Delta\delta_{\max} \times [\text{sclx}_4]^n / (K_d^n + [\text{sclx}_4]^n)$$

Chapter 3:

Electron transfer in cytochrome *c*-calixarene crystals: An electrochemical study

Manuscript in preparation

Abstract

The self-assembly of biomolecules, such as cytochrome *c*, on electrode surfaces has facilitated the development of biosensors. However, the greatest challenge is increasing the local protein concentration at the electrode surface to achieve higher sensitivity for electrochemical detection. Recently we showed that the “molecular glue” *p*-sulfonatocalix[4]arene is a useful promoter of protein crystallization. Here we present the data for large cytochrome *c* crystals grown by the calixarene on modified Au electrodes and are visible to the naked eye. The cytochrome *c* crystals were characterised by cyclic voltammetry and revealed profiles different to the conventional monolayer. An electrochemical cell solution that was similar to the mother liquor, in which the crystals grew, was required to ensure crystal stability during the measurements. The highest concentrations of electro-active cytochrome *c*, which ranged from 7200-13000 $\mu\text{mol}/\text{cm}^2$, were detected at the lowest scan rates. An analysis of the crystal structure provides evidence for interprotein electron transfer through the calixarene mediated cytochrome *c* assembly.

Introduction

Cytochrome *c* is an important mediator within the mitochondrial electron transfer chain and its interactions with redox partners are well characterised.^{9,14,129} Owing to its redox activity cytochrome *c* is routinely used for biosensor construction.^{23,207-209} This typically involves immobilising the protein on gold electrode surfaces functionalised with self-assembled monolayers (SAMs).²¹⁰ SAMs are densely packed ω -functionalised alkanethiols onto which cytochrome *c* adsorbs to form monolayers and electron transfer through these monolayers is influenced by the SAM composition.^{211,212} The ω -functional group influences the orientation of cytochrome *c* molecules on the SAM and this affects the electron transfer rate.²¹³ The alkanethiols chain length also affects the electron transfer rate.²¹⁴ Electron transfer through cytochrome *c* monolayers is diffusionless, however, the molecules possess some mobility that can allow subtle changes of the protein's orientation.²¹¹ This effect was apparent in the *c*-type cytochromes of *Geobacter sulfurreducens*.²¹⁵ The redox potentials of the bacterial cells was influenced by the SAM composition suggesting that electron transfer is governed by orientation changes of the *c*-type cytochromes.

The low protein concentration of the monolayer, typically ~ 10 pmol/cm², limits the sensitivity of electrochemical detection. Therefore, the construction of complex protein assemblies is required to improve biosensor design. The layer-by-layer adsorption technique offers a route to increasing the protein concentration in a controlled manner via the formation of polyelectrolyte mediated protein multilayers.^{216,217} Polyaniline sulfonic acid (PASA), DNA and silica nanoparticles have been used to generate cytochrome *c* multilayers on modified Au surfaces.²¹⁸⁻²²⁰ Up to 15 electro-active layers of cytochrome *c* were assembled by using PASA and the electrochemical response increased per layer as a result of direct interprotein electron transfer.²²

Protein crystals are a good source of highly concentrated protein assemblies. Despite the assembled molecules lacking mobility, electron transfer has been measured in redox protein crystals by using spectroscopic methods.²²¹⁻²²⁵ The contribution of the orientations and structures of cofactors to the rate of electron transfer has been examined.²²⁴ The electrochemical characterisation of protein crystals is limited with only one report made by using cyclic voltammetry (CV).²²⁶ The method involved transferring cytochrome *c* crystals to a fluid cell and fixing the crystals in a polypyrrole film.

Voltammograms for the immobilised crystal were similar to that of the cytochrome *c* monolayer prepared in the cell.

Previously, we demonstrated that *p*-sulfonatocalix[4]arene (sclx₄) is a useful molecular glue for cytochrome *c* crystallization.⁴⁸ Therefore, we sought to investigate if sclx₄-mediated cytochrome *c* assemblies, for which the crystal structure is known, could replace the PASA-cytochrome *c* multilayer systems used for electrochemical characterisation (Figure 1). The calixarene and PASA are similar because both have aromatic and sulfonated portions. Protein crystal structures are useful for gauging interprotein electron-transfer by analysing the distances between redox centres.²²⁷ The *Saccharomyces cerevisiae* cytochrome *c* Arg13Glu mutant (referred to as cyt *c* in this chapter) was used for co-crystallization experiments with the calixarene on electrodes. This mutant yields significantly large crystals, visible by eye, by comparison to the wild type. Co-crystallization with the latter is typically slower and yields smaller crystals. The cyt *c* crystals grown on SAM-modified Au electrodes were detected by using CV. A combination of the electrochemical cell solution and applied current influenced crystal stability. The standard electrochemical cell buffer (5 mM potassium phosphate, pH 7) was not suitable for scan rate dependency studies as it promoted crystal dissolution. Therefore, altered crystallization conditions that did not compromise the crystals were used for the CV measurements. A clear dependence of the redox signal intensity on scan rate is apparent and an analysis of the cyt *c*-sclx₄ crystal structure suggests that electron transfer is mediated by protein-protein interactions.

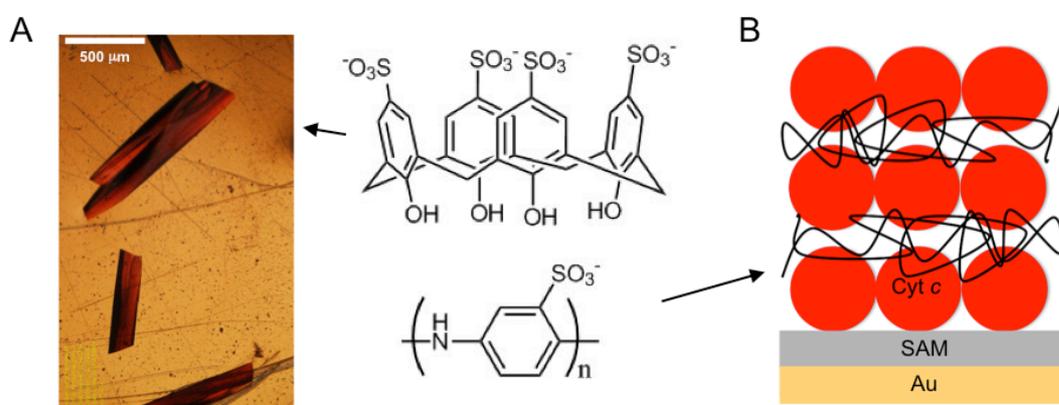


Figure 1: (A) Cyt *c*-sclx₄ crystals on a Au electrode and the calixarene used to grow the crystals. (B) An illustration showing the multilayer PASA-cyt *c* assemblies on a SAM-modified Au electrode. PASA is represented as lines mediating the cyt *c* layers.

Results and Discussions

Cyt *c*-sclx₄ crystallization on electrodes

Au wires with a diameter of 0.5 mm were used for preliminary crystallization and electrochemical characterisation (Figure 2). Both the bare and modified forms of the electrode surfaces were used. For the modified surfaces, a SAM was prepared consisting of 1:3 Mercaptoundecanoic acid/Mercaptoundecanol (MUA/MU) on the Au surface. On top of this a cyt *c* monolayer was prepared (Au-SAM-monolayer). Co-crystallization trials were prepared directly onto the electrode surface and a 10-fold excess of sclx₄ was required for crystal growth.

CV measurements of cyt *c*-sclx₄ crystals on Au wires

The electrochemical cell was set up by gently submerging a Au wire electrode containing the analyte (cyt *c* monolayer or crystals) in 1 mL of 5 mM potassium phosphate (KPi), pH 7. As protein crystals are fragile the transfer of the crystals from a PEG/salt solution to a phosphate-buffered solution is expected to induce osmotic shock, a precursor to crystal damage. As a result, crystal deterioration was to some extent unavoidable. The submersion of crystal-containing electrodes into the cell buffer resulted in some crystal displacement and/or dissolution.

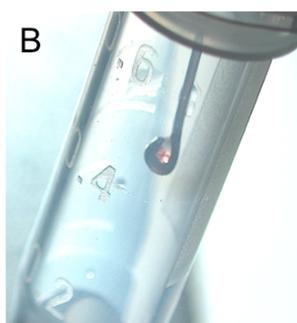
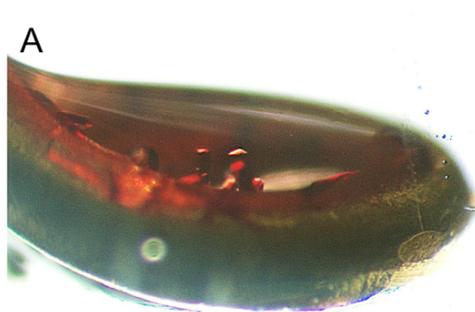


Figure 2: Cyt *c* crystallization on Au wires. (A) Cyt *c* crystals on a Au wire. (B) Crystals were grown on the loop-end of a Au wire, which was suspended over a 100 μ L reservoir solution in a sealed micro-centrifuge tube.

Initial experiments were on electrodes that had a SAM, a cyt *c* monolayer and crystals. Measurements made at 100 mV/s yielded a voltammogram consisting of a single pair of reduction-oxidation peaks, which were characteristic of cyt *c*. The reduction and oxidation peaks were shifted to the right and left, respectively. The cathodic-to-anodic peak separation was large (0.2 V) at 100 mV/s by comparison to the monolayer only, which was 0.05 V at 100 mV/s. The peak intensities for first measurement of the

electrode containing crystals were significantly large compared to the Au-SAM-monolayer electrode (Figure 3A). Repeating the measurement with the crystal-containing electrode over 100 scans showed a steady decline in peak heights and this is consistent with the crystal degradation observed over during the experiment (Figure 3B). No redox peaks were detected for crystals grown on bare Au wires, indicating no electro-activity. Therefore, a SAM was essential for the detection of electro-active crystals.

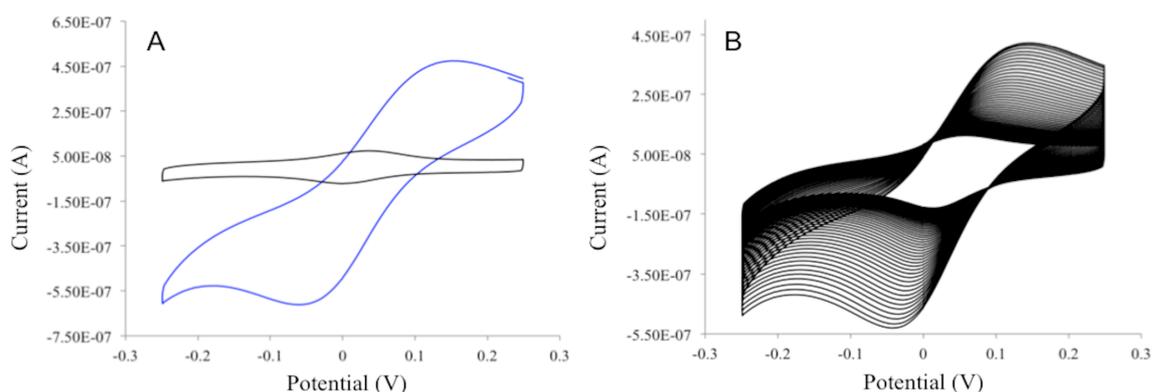


Figure 3: (A) Voltammograms of a cyt *c* monolayer (black) and crystals (blue) on modified Au electrodes. (B) The intensity of the redox signal decreased over 100 scans for the crystals. The scan rate was 100 mV/s.

Cyt *c*-sclx₄ crystallization on Au planar electrodes

After identifying that cyt *c* could be crystallized on Au wires and detected by CV, the method was transferred to Au planar electrodes with a diameter of ~ 0.7 cm. The advantage of using a planar surface meant that the surface coverage of the crystals could be determined. The protocol was further altered by excluding the cyt *c* monolayer on which crystallization was initially performed. The advantage of this was that the amount of detectable electro-active cyt *c* would correspond exclusively to the crystallized protein. The signals could not be confused with some phenomenon such as the rearrangement of the cyt *c* monolayer, which may occur during the crystallization process. The SAMs were essential for crystal detection. Therefore, crystallization experiments were prepared on SAM-modified Au electrodes (Au-SAM). Crystals grew directly on the modified electrodes in drops containing a 10-fold excess of sclx₄. The Au-SAM electrodes containing the crystals (Au-SAM-crystals) were transferred to an electrochemical cell containing 5 mL of 5 mM KPi, pH 7.

CV measurements of cyt *c*-sclx₄ crystals on Au planar electrodes

A Au-SAM-crystal electrode was used for preliminary experiments and the crystals covered 0.06 % of the surface. CV experiments were made in the range of 10-1000 mV/s and for each scan rate a single pair of reduction-oxidation peaks characteristic of cyt *c* was observed. The reduction and oxidation peaks were shifted to the right and left, respectively. The cathodic-to-anodic peak separation was largest (0.3 V) at a 1 V/s scan rate and smallest (0.1 V) for 10 mV/s. The first and final measurements were made at 10 mV/s and a comparison of each showed a decrease in signal by the last measurement (Figure 4A). The greatest amounts of detectable cyt *c* were at the lowest scan rates suggesting a slow electron transfer process (Table 1). In the final measurement, the concentration of electro-active cyt *c* was $\sim 17000 \text{ } \mu\text{mol/cm}^2$. This is a decrease of $\sim 50 \%$ with respect to the first measurement and is consistent with severe crystal degradation observed over during the experiments. Both the buffer and the electric field are expected to cause crystal deterioration. To investigate crystal dissolution, cyt *c* crystals were transferred to 5 mM KPi, pH 7 and monitored. Gradually, over the course of an hour, fractures were observed throughout the crystal leading to degradation and dissolution (Figure 4B). This process was accelerated under an electric field and is likely due to the rapid influx of ions through the crystal channels, resulting in lattice breakages. Therefore, efforts focused on overcoming this challenge by adapting the cell solution.

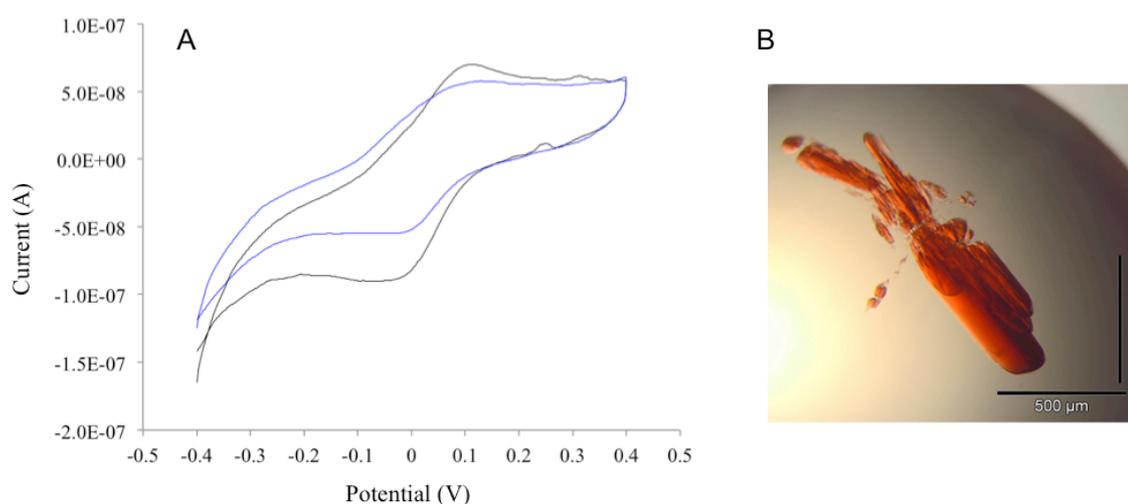


Figure 4: (A) Voltammograms of the first (black) and final (blue) measurements of the Au-SAM-crystal electrode at 10 mV/s. The redox peak intensities are less for the final measurement. (B) Photograph of a cyt *c* crystal showing fractures. The photograph was taken 30 minutes after transferring the crystal to 5 mM KPi, pH 7.

To improve crystal stability during the experiments a CV cell solution was prepared to match the crystallization solution. The electrochemical cell solution contained 50 mM NaCl, 100 mM MgCl₂ and 24 % PEG 4000 and was adjusted to pH 7. Preliminary experiments for an Au-SAM-crystal electrode involved scan rates of 10-1000 mV/s. As observed for the electrode characterised in KPi buffer, the greatest concentration was detected at the lowest scan rate (Table 1). A comparison of the first and final measurement at 10 mV/s showed an increase by almost 10 fold in the amount of electro-active species on the electrode surface (Figure 5). This suggested that the crystals required a period of equilibration prior to making measurements. Therefore, the Au-SAM-crystal electrodes were incubated for 10 minutes in the cell solution prior to making any measurements.

Table 1. Cyt *c* concentration (Γ) determined for Au-SAM-crystal electrodes at different scan rates.

Scan rate (mV/s)	Γ ($\mu\text{mol}/\text{cm}^2$)	
	5 mM KPi pH 7	Crystallization solution
10 (first scan)	32325.92	6187.55
1000	5481.26	1842.70
100	9726.42	7196.87
50	12049.76	6403.91
10 (final scan)	16865.60	58518.48

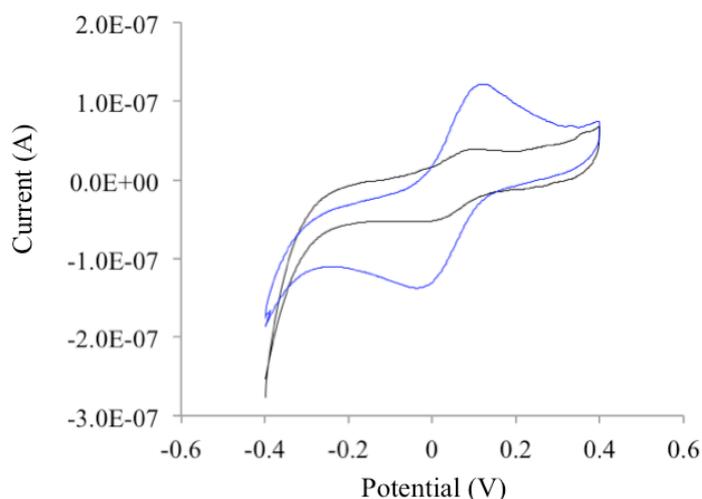


Figure 5: Voltammograms of the first (black) and final (blue) measurements of the Au-SAM-crystal electrode in 50 mM NaCl, 100 mM MgCl₂ and 24 % PEG 4000, pH 7. The crystal coverage was 0.12 %. Larger, better-defined redox peaks are observed for the final measurement. The scan rate for both measurements was 10 mV/s.

Reproducibility of scan rate dependency measurements

Three Au-SAM-crystal electrodes (electrode **I**, **II** and **III**) were subjected to a scan rate dependency study. Prior to making any CV measurements, photographs were taken of each electrode (Figure 6) and the surface coverage of the crystals was determined. Each Au-SAM-crystal electrode was subsequently incubated for 10 minutes in the cell solution. The scan rates used for the study were 2-1000 mV/s. Measurements were also made before and after the study at 10 mV/s. Photographs were taken of each electrode after the experiments to assess crystal intactness.

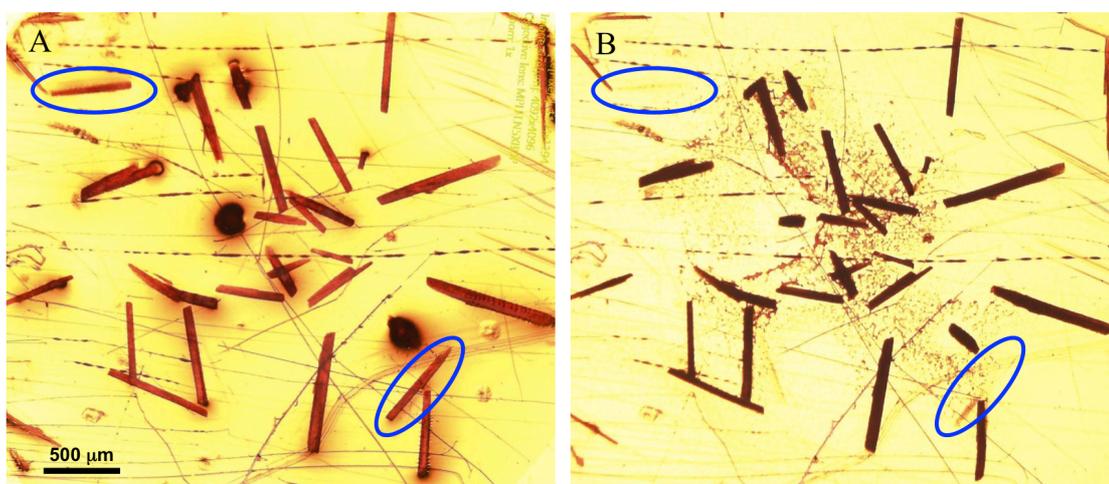


Figure 6: Au-SAM-crystal electrode (A) before and (B) after the scan rate dependency study. Crystals that dissolved during the experiment are circled.

The voltammograms for the Au-SAM-crystal electrodes at each scan rate showed a single pair of reduction-oxidation peaks, characteristic of cyt *c* (Figure 7A). The anodic peak potential was shifted to the right in each measurement. The anodic and cathodic peak currents measured for the redox couple increased linearly with increasing scan rate from 2 to 1000 mV/s (Figure 7B). The formal potential (E_f) of the $\text{Fe}^{2+}/\text{Fe}^{3+}$ haem couple of cyt *c* varied from electrode to electrode and scan rate to scan rate. Typically, the largest potentials were observed at 1 V/s. The potentials of the Au-SAM-crystal electrodes at 100 mV/s ($E_f = 0.05\text{-}0.04$ V) were higher than the E_f values of the Au-SAM with a cyt *c* monolayer measured in the same conditions and at the same scan rate ($E_f = 0.03\text{-}0.02$ V). This E_f difference between the crystals and the monolayers hints towards an electron transfer process more complicated than that observed in the monolayer. Additionally, the peak separations suggest slower electron transfer process than observed for the Au-SAM-

monolayer. The concentration of electro-active cyt c was determined for each electrode at each scan rate.

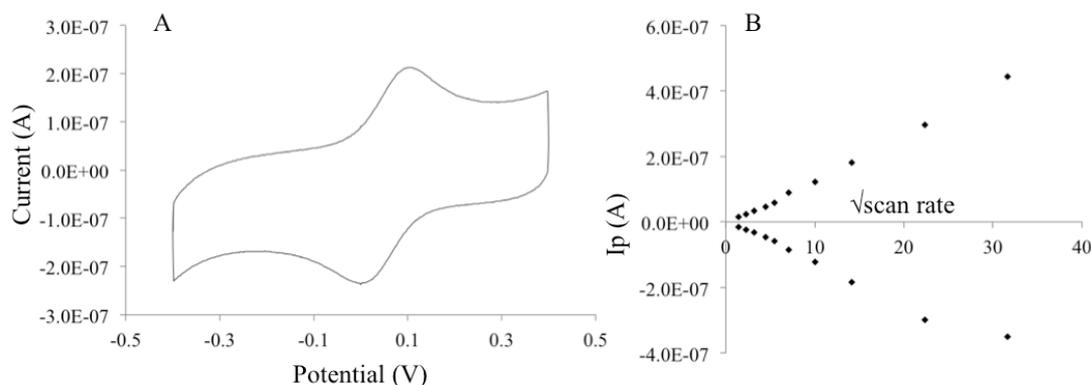


Figure 7: (A) Voltammogram measured for electrode **I** at 100 mV/s in the crystallization solution. (B) The dependence of the catalytic current on the $\sqrt{\text{scan rate}}$ for **I** (Inset).

A comparison of the first and final measurements shows an increase in the concentration of electrode **I** (Table 2). This is more substantial in electrode **II**. However, the inverse is observed for electrode **III**. A possible explanation for these deviations for each electrode is that no two crystallization drops can be made absolutely identical. For example, convection effects in the vapour diffusion method can cause differences in the kinetics to reach equilibrium. The crystal coverage difference per electrode also reflects the variability of the crystallization method.

Table 2. The cyt c concentration (Γ) determined for electrodes **I**, **II** and **III** at 10 mV/s.

Electrode	I (first)	I (final)	II (first)	II (final)	III (first)	III (final)
Γ ($\mu\text{mol}/\text{cm}^2$)	3206.4	3538.5	3857.9	6230.2	5440.1	3984.5

Overall, the results showed that the highest concentrations were detected at the lowest scan rates (Table 3). An increase in the scan rate reduced the amount of cyt c measured suggesting that the electron-transfer rate within the crystals is sensitive to the scanning speed. Similar effects were observed for the Au-SAM-crystals measured in KP buffer and a similar trend occurred in the PASA-cyt c multilayer systems, for which protein-protein electron transfer within the assembly is the ultimate pathway.²²⁸

Table 3. The cyt *c* concentration[†] (Γ), formal potential (E_f), peak-to-peak separation (ΔE_p) and the peak width at half peak height (δ_{avg}) determined for the three Au-SAM-crystal electrodes.

	Scan rate (mV/s)									
	1000	500	200	100	50	30	20	10	5	2
	Γ ($\mu\text{mol}/\text{cm}^2$)									
I	472.99	615.97	905.44	1197.31	1643.98	1580.76	1982.22	2881.46	4107.72	7311.48
II	1564.41	2037.72	2787.04	3289.31	3347.39	2804.22	3209.15	5306.59	8033.91	12783.42
III	1307.35	1448.93	1711.31	1902.65	2097.42	2741.13	3167.82	3286.42	4527.38	8525.37
	E_f (V vs. Ag/AgCl ₂)									
I	0.095	0.075	0.055	0.051	0.052	0.048	0.046	0.058	0.053	0.053
II	0.125	0.089	0.063	0.051	0.051	0.051	0.051	0.051	0.051	0.051
III	0.061	0.044	0.038	0.040	0.037	0.040	0.040	0.039	0.044	0.049
	ΔE_p (V)									
I	0.190	0.150	0.110	0.102	0.080	0.056	0.072	0.040	0.058	0.054
II	0.250	0.180	0.126	0.102	0.078	0.058	0.018	0.038	0.032	0.056
III	0.120	0.090	0.052	0.040	0.022	0.028	0.028	0.026	0.048	0.062
	δ_{avg} (V)									
I	0.19	0.19	0.18	0.17	0.17	0.15	0.15	0.14	0.14	0.12
II	0.24	0.23	0.19	0.16	0.14	0.13	0.12	0.12	0.09	0.23
III	0.16	0.14	0.10	0.13	0.12	0.13	0.12	0.15	0.15	0.16

[†]0.46 %, 0.16 % and 0.11 % of the surface was covered by crystals for electrode **I**, **II** and **III**, respectively.

Some protein dissolution from the crystals on the electrodes was expected to occur due to the transfer of the electrodes into the cell. As a result, it was essential to prove that the detected redox peaks corresponded to the crystals and not to protein that had dissolved from the crystals. To assess the effect of crystal dissolution the Au-SAM-crystal electrode was removed from the electrochemical cell solution after the scan rate dependency study and replaced by a Au-SAM. Importantly, the cell solution itself was not exchanged. Redox peaks characteristic of cyt *c* were barely visible in the voltammogram measured with the Au-SAM (Figure 8). This suggests that the cyt *c* molecules in the crystals are the key contributors to the redox signal observed in the Au-SAM-crystal voltammograms.

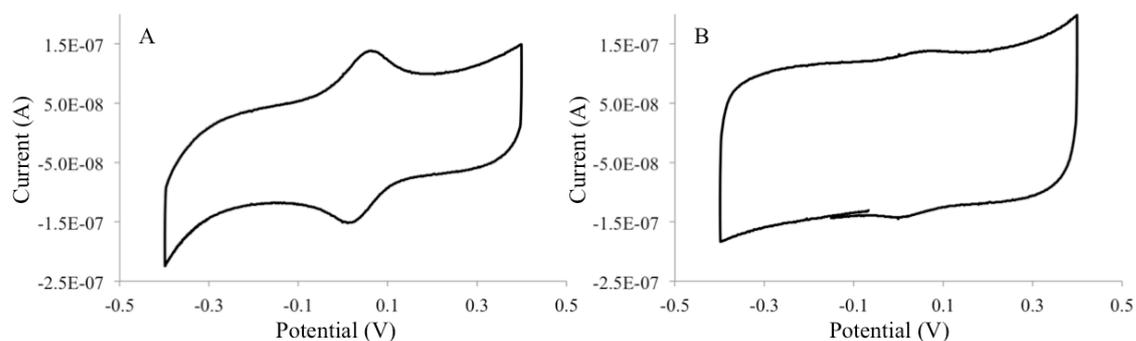


Figure 8. (A) Voltammogram of a Au-SAM-crystal electrode measured in crystallization solution. (B) Voltammogram of a Au-SAM electrode measured in the same solution after the scan rate dependency study with (A). Redox peaks are barely visible for the Au-SAM electrode. The scan rate was 100 mV/s.

Cyt *c* solution and monolayer profiles

As a standard experiment cyt *c* was characterised in the crystallization solution using a Au-SAM electrode. Operating in the 50-200 mV/s range a single pair of redox peaks was observed for 0.5 μM cyt *c*. However, at higher scan rates, 500-1000 mV/s, noise interfered with cyt *c* detection. Increasing the protein concentration improved the voltammogram resolution (Figure 9A). The highest concentration of electro-active protein was determined from the lowest scan rate. For 20 μM cyt *c* at 2 mV/s this was ~ 22 pmol/cm² and at 1 V/s only ~ 1 pmol/cm² was detected. Importantly, the addition of 0.5-100 μM sclx₄ to the solution did not affect the voltammograms. Furthermore, no redox peaks were visible in the assigned potential range for an Au-SAM electrode in the crystallization solution containing 200 μM sclx₄. Therefore, sclx₄ was not responsible for the Au-SAM-crystal voltammograms.

Cyt *c* monolayers were prepared on Au-SAM electrodes and characterised in 24 % PEG 4000, pH 7. This modified form of the crystallization solution was used as the presence of NaCl and MgCl₂ depleted the concentration of cyt *c* in the monolayer. This observation is a result of the competition between the SAM and the cell solution for cyt *c*, which is attributed to charge-charge interactions. A single pair of redox peaks was observed and the peak separations were ~ 0.04 V from which 7.1-9.0 pmol/cm² of electro-active cyt *c* was determined (Figure 9B).

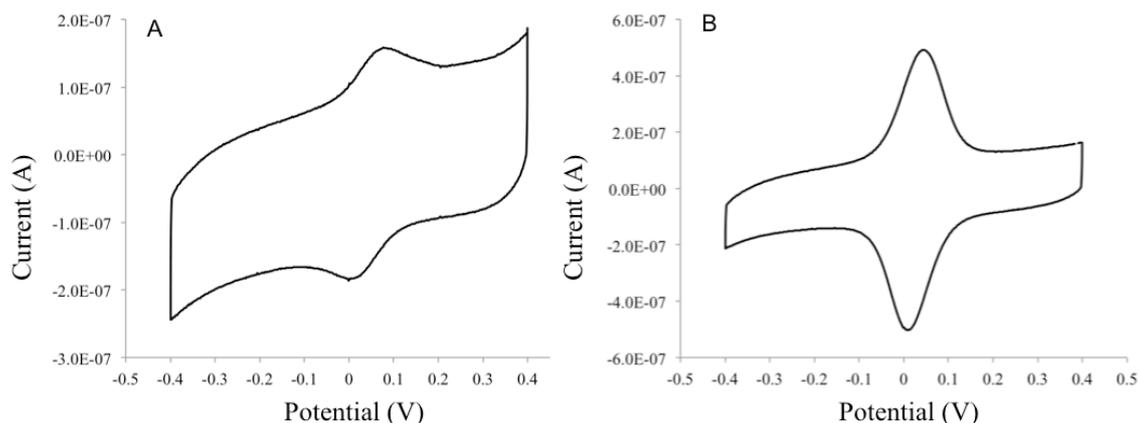


Figure 9: (A) Voltammogram of a cyt *c* solution. The measurement was made using a Au-SAM electrode and the crystallization solution in the electrochemical cell contained 20 μM cyt *c*. (B) Voltammogram of cyt *c* monolayer prepared on a Au-SAM and characterised in 24 % PEG 4000, pH 7. The measurements were made at 100 mV/s.

Communication with sulfite oxidase

In the presence of a substrate, high current outputs can be induced and detected by CV. The Au-SAM-crystal electrodes were tested as biosensors for sulfite detection, for which the crystals were mediators between an enzyme and the electrode. Sulfite oxidase (SO) is a molybdenum- and iron-containing enzyme that catalyses the oxidation of sulfite to sulfate at the molybdenum cofactor.²²⁹ From this reaction two electrons are transferred, one at a time, to the b_5 haem of sulfite oxidase and subsequently onto cyt *c*, its natural reaction partner. As a result of its sulfite conversion ability, direct electron transfer from SO to electrode surfaces can occur, meaning it has potential applications in biosensor design.^{23,208}

SO communication with Au-SAM-crystal electrodes was investigated in a cell solution containing of 24 % PEG 4000, pH 7 and the scan rate was 5 mV/s. 1 μM SO addition to the cell solution showed virtually no change in the voltammogram. However, the addition of 1 mM Na_2SO_3 resulted in a significant increase in the oxidation current, indicating the oxidation of sulfite to sulfate by SO. This is an interesting observation as cyt *c* mobility is diminished in the crystal. A similar trend was obtained with the Au-SAM-monolayer electrodes (Figure 10). Cyt *c* immobilised on a SAM consisting of poly(2-methoxyaniline-5-sulfonic acid)-*co*-aniline polymers also showed a current increase with the enzyme and substrate.²³⁰ Importantly, the results for the crystal-containing electrodes are preliminary and control experiments are required to investigate the validity of the reported voltammograms. Nonetheless, the data suggests that electron

transfer from the enzyme to immobilised cyt *c* molecules with reduced mobility can be achieved.

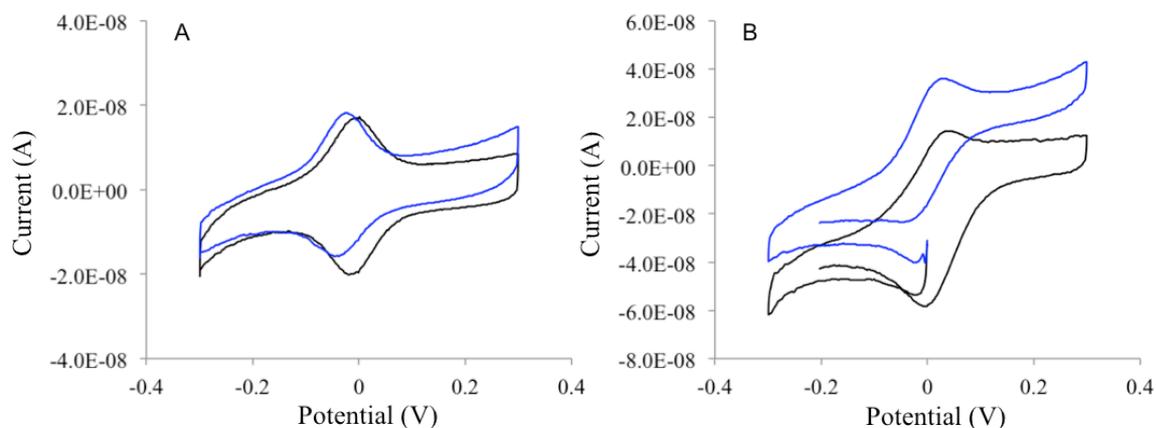


Figure 10: (A) Voltammograms of the Au-SAM-monolayer electrode in the presence of 1 μM SO (black) and after the addition of 1 mM Na_2SO_3 (blue) to the cell solution. (B) Voltammograms of the Au-SAM-crystal electrode with SO before (black) and after (blue) substrate addition. The scan rate was 5 mV/s.

Electron transfer through cyt *c* crystals

Electron transport from the electrode and into the crystals is apparent, however, the factors that control electron flow between the immobilised haem cofactors must be addressed. The conditions used to grow the crystals on Au-SAMs were the same as those used to grow cyt *c*-sclx₄ crystals for diffraction. Therefore, the orientations of the cyt *c* molecules in the crystals were believed to be similar to the solved structure (Chapter 2). The $P2_12_12_1$ symmetry operations were applied to the asymmetric unit of the cyt *c* crystal (PDB ID 4n0k) to generate assemblies and examine the contribution made by crystal packing neighbours. An analysis of the crystal packing neighbours shows that the shortest Fe-Fe distance is 23.7 Å (Figure 11). All other Fe-Fe distances in the crystal lattice are > 30 Å. It is well recognised that electron transfer can occur over distances < 25 Å.²³¹ Significantly, similar distances were observed in redox active crystals comprised of both native and Zn-substituted tuna cyt *c*. The shortest Zn-Fe distance was 24.1 Å, suggesting that the protein-protein interfaces in crystals mediated interprotein redox reactions. In these crystals the electron transfer rate was faster for the oxidation step (2000 s⁻¹) than the reduction step (320 s⁻¹).²²²

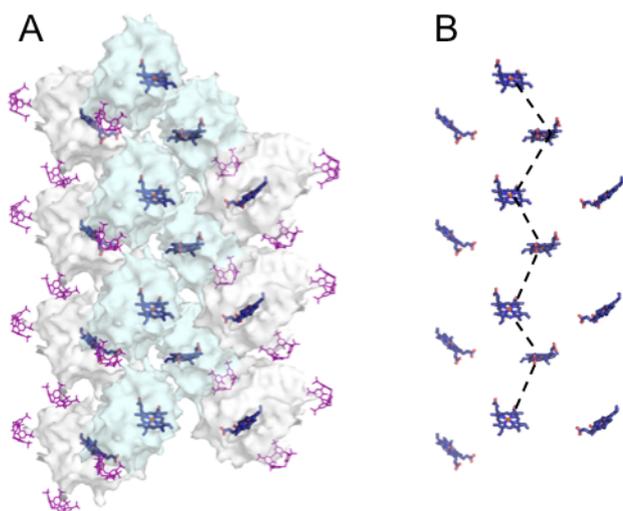


Figure 11: (A) Cyt *c*-sclx₄ assemblies generated by applying the $P2_12_12_1$ symmetry operations. Cyt *c* is represented as transparent surfaces and chains A and B are coloured cyan and grey, respectively. Haems are blue sticks and the calixarenes are purple lines. (B) The shortest distance between neighbouring Fe centres is 23.7 Å, represented by a dashed line. All other Fe centre distances are > 30 Å.

Conclusions

Protein crystallization on modified Au electrodes offers an alternative strategy to the layer-by-layer adsorption method for generating high protein concentrations. Cyt *c* crystals were grown and electrochemically detected on Au-SAM electrodes by using CV. The reported concentrations exceed those of multilayer assemblies, which are ~ 80 $\mu\text{mol}/\text{cm}^2$ for systems using poly(2-methoxyaniline-5-sulfonic acid)-*co*-aniline polymers or silica nanoparticles.^{220,232} The role of sclx₄ was to promote crystallization and importantly the calixarene afforded no additional electrochemical properties in the applied potential range. This is similar to the PASA polyelectrolyte, which also does not contribute to the electron transfer process as it shows no redox activity.²²

The concentration of electro-active cyt *c* in the crystals is dependent on the scan rate. At lower scan rates higher concentrations of cyt *c* contribute to the current. For example, a concentration of ~ 1600 $\mu\text{mol}/\text{cm}^2$ of cyt *c* was calculated from data collected at 1 V/s (Table 3). This value increased 8-fold when the scan rate was reduced to 2 mV/s. A similar observation was made with the cyt *c*-PASA multilayer systems and suggests that the rate of electron exchange between the cyt *c* molecules is slower than the applied scan rate.²² The distance between the metal centres influences electron transfer rates²³³ and metal centres separated by 20-25 Å have rates in the μs -ms range.^{222,234,235} Significantly, the Fe-Fe distance in the cyt *c*-sclx₄ crystal is shorter than 25 Å and this is similar to the metal-metal separation in tuna cyt *c* crystals.²²² The protein-protein interfaces in crystals mediate the interprotein redox reactions.

A major advantage of crystal systems is that the precise orientations and distances of the redox centres can be solved. The use of small molecules that act as molecular glue for redox protein crystallization presents a new perspective for generating highly ordered architectures for electrochemical characterisation. Therefore, the use of the co-crystallization method might inspire the development of custom-designed structures for biosensing.

Experimental

Preparation of SAM and cyt *c* monolayers on Au electrodes

Au wire and chip electrodes were incubated in 5 mM MUA/MU (1:3) in pure ethanol for 48 hours allowing SAM formation.²² The Au-SAM electrodes were washed by gently dipping the electrodes in pure ethanol and then in 5 mM KPi, pH 7 or water. Cyt *c* monolayers were subsequently prepared on the SAM-modified Au electrodes by incubating the electrodes in 30 μ M cyt *c*, 5 mM KPi pH 7 for 2 h. After cyt *c* monolayer formation, the electrodes were washed by gently dipping the electrodes in 5 mM KPi, pH 7 or water.

Cyt *c*-sclx₄ co-crystallization on electrodes

The hanging drop vapour diffusion method was used to grow crystals directly on the Au planar electrode surface, which was inverted over a 100 μ L reservoir solution. Each electrode was incubated at room temperature in a sealed chamber. For crystallization on Au wires, the end of the wire was fashioned into a loop to hold the drop. The Au wire was suspended in a sealed micro centrifuge tube over a 100 μ L reservoir solution (Figure 2). The reservoir solution contained 24 % PEG 8000, 50 mM NaCl and 100 mM MgCl₂. Equal volumes of 1.7 mM of cyt *c* and 17 mM sclx₄ were combined in a complex solution. Experiment drops were prepared by combining 2 μ L of the complex solution and 1 μ L of the reservoir solution. Control drops were prepared by combining 1 μ L solutions of each 1.7 mM cyt *c*, water and reservoir solution. Crystals grew on both the Au chip and Au wire electrodes. The control drops were devoid of any precipitate.

Electrochemistry

All electrochemical measurements were prepared using a homemade 1 mL cell using an Ag/AgCl/ 1 M KCl reference (Microelectrodes, Inc., Bedford, USA) and a Pt-wire counter electrode. The working electrodes were modified Au wires (diameter 0.5 mm, obtained from Goodfellow Bad Nauheim, Germany) or modified Au planar electrodes (diameter \sim 0.7 cm). The cell solution contained 24 % PEG 4000, 50 mM NaCl and 100 mM MgCl₂, pH 7, unless otherwise stated. Cyclic voltammetry experiments were carried out with a CH Instrument (Austin Texas, USA). The scan rates varied from 1000 to 2 mV/s. For the scan rate dependency study the scan rates were 1000 mV/s, 500 mV/s, 200 mV/s, 100 mV/s, 50 mV/s, 30 mV/s, 20 mV/s, 10 mV/s, 5 mV/s and 2 mV/s. A potential range of -0.4 and +0.4 V vs. Ag/AgCl was applied. Data analysis was performed using CHI 604d software. Integration of the CV peak areas provided the electro-active surface concentration (Γ).

Measurements with crystals, sulfite oxidase and Na₂SO₃

CV was performed on Au-SAM-monolayer and Au-SAM-crystal electrodes in 24 % PEG 4000, pH 7 at 5 mV/s. The buffer was supplemented with 1 μ M human sulfite oxidase and the electrochemical response was assessed by CV. The solution was then supplemented with 1 mM Na₂SO₃ and CV performed. All the CV measurements were made at 5 mV/s.

Chapter 4:

The crystal structure of lysozyme and *p*-sulfonatocalix[4]arene

Manuscript submitted

Abstract

Crystal structures of co-crystallized protein-ligand complexes allow the identification of epitopes involved in recognition events. Here, the first crystal structure of hen egg white lysozyme in complex with *p*-sulfonatocalix[4]arene is reported. The calixarene shows a preference for the cationic side chain of arginine. The structure provides new information on protein recognition by the calixarene. The structure also shows how the crystallization agents cooperate with the calixarene to stabilise the protein complex. The self-organising principles that govern this complex formation are discussed.

Introduction

The association of biological macromolecules into distinct levels of structural organisation is widespread in cellular systems. The size and shape of the protein, complementarities between surfaces, hydrophobicity, electrostatic interactions and conformational changes resulting from complex formation are important factors for governing assembly.¹⁹ These features have been exploited to gain control over protein self-assembly. For example, efforts have focused on designed protein-protein association via protein engineering.^{26,30} This was demonstrated by the introduction of additional hydrophobic residues at the contact interface of L-rhamnulose-1-phosphate aldolase, which promoted self-assembly and stabilised the complex.²⁶

Small molecules present a noncovalent approach for promoting protein self-assembly. Tetraarylporphyrin- β -cyclodextrin complexes have also been used to generate dimers of cytochrome *c*, by binding the protein surface.²³⁶ Cucurbit[8]uril induces protein tetramerisation by entrapping two phenylalanine side chains from the N-terminal PheGlyGly motif in its cavity.²⁰⁰ Small molecules that act as ‘molecular glue’ are also important for generating protein assemblies and promoting crystallization.^{40,47} Pyrenetetrasulfonic acid was required to grow diffraction quality crystals of PYK.⁴⁰ The crystal structure of the complex shows the ligand making salt bridges and cation- π interactions with an arginine side chain from two neighbouring protein chains. In this instance arginine may be considered as a valuable target for altering self-assembly.

Arginine is important at protein-protein interfaces and those involved in cation- π interactions are especially common.^{59,237} Arginine recognition by the aromatic cage of chromodomain 2 involves cation- π interactions and occurs as part of chromatin structure regulation.²³⁸ The complexation of arginine by various small synthetic molecules, typically decorated with acidic groups, have been explored owing to its role in protein interactions.^{60,239-242} For example, a cyclophane host, which has well-defined hydrophobic cavity and peripheral carboxyl groups, binds arginine.²⁴⁰ A combination of cation- π , hydrophobic and electrostatic interactions between the host and guest contribute to the high affinity ($-\Delta G^\circ = 5.0$ kcal/mol) of the binding event. The arginine cork is a small, rigid, planar molecule with two peripheral carboxylates and displays μ M affinity for arginine.²⁴¹ The crystal structure of the arginine cork with *N*-ethylguanidinium shows one cation in the receptor cleft and a second cation binding one carboxylate.

Calixarenes have played an important role in various strategies from protein extraction into organic phase to self-assembly.^{70,96} *p*-sulfonatocalix[4]arene (sclx₄) has demonstrated great potential as a protein surface binder via cationic amino acids.^{90,91,94,95} As described in Chapter 2, sclx₄ binds cytochrome *c* (cyt *c*) predominantly via lysine side chains. However, sclx₄ is also an arginine binder^{91,243} and similar to the complexes formed with lysine,⁹⁰ calixarene can bind arginine residues *endo* or *exo*. In aqueous solution at pH 8 sclx₄ makes 1:1 complexes with dipeptides and tripeptides bearing arginine residues.²⁴³ To substantiate the potential of sclx₄ as a protein surface binder and a mediator of protein-protein interactions we used X-ray crystallography. The structural consequences of the lysozyme-sclx₄ complex are described.

Results

Crystallization and structure determination

Lysozyme-sclx₄ co-crystallization trials were prepared using the conditions that yielded diffraction quality crystals for the cyt *c*-sclx₄ complex.⁴⁸ The conditions were chosen as lysozyme and cyt *c* have similar pI values in the range of 10.5-11.^{136,244} Cubic crystals of lysozyme grew in drops containing 24-28 % poly(ethylene glycol) (PEG) within three weeks (Figure 1A). Larger crystals of different shapes were observed after three months in a drop containing 22 % PEG (Figure 1B). Control drops lacking sclx₄ were devoid of precipitate and crystals.

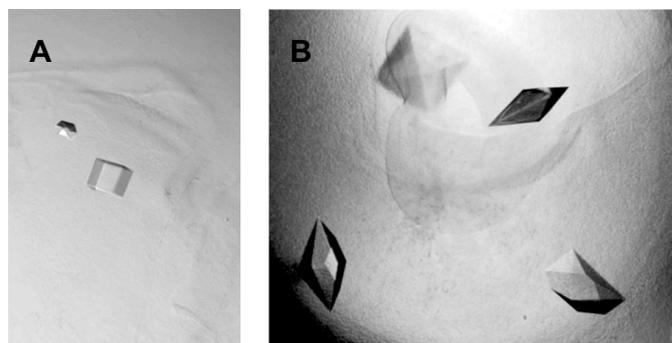


Figure 1: (A) A cubic crystal (10 μm length) which grew from 24 % PEG 8000, diffracted to 1.7 \AA and belonged to $P12_11$. (B) Other crystal forms belonged to $P4_12_12$ and $P4_32_12$. Although these crystals grew from almost identical conditions (a difference of 2-6 % PEG), two months were required for crystallization and the structures lacked sclx₄.

Structural characterisation was achieved by X-ray crystallography. The larger crystals belonged to space groups $P4_12_12$ and $P4_32_12$ and diffracted to 1.2 \AA , however, no

sclx₄ was present in the asymmetric unit. In contrast, the largest of the cubic crystals diffracted to 1.7 Å and belonged to the monoclinic space group *P*12₁1 (Table 1). For this crystal structure, the asymmetric unit is composed of four molecules of lysozyme and five molecules of sclx₄ (Figure 2). The four protein molecules are referred to as chains A, B, C and D. Five PEG fragments and three Mg²⁺ are also present in the asymmetric unit.

Table 1. Crystallographic data for the complex of sclx₄ with lysozyme.

Space group	<i>P</i> 12 ₁ 1
Cell constants	a = 44.01 Å b = 81.75 Å c = 72.10 Å $\alpha = \gamma = 90^\circ$ $\beta = 105.37^\circ$
Resolution, Å	81.75-1.72 (1.77-1.72)
Wavelength, Å	1.00792
Unique reflections-Corrected	1386759 (51142)
Multiplicity	3.0 (2.6)
<i>I</i> / σ	9.4 (1.8)
Completeness, %	98.7 (94.9)
<i>R</i> _{merge} ^b , %	7.3 (55.6)
Solvent content, %	43.86
<i>R</i> _{factor} , %	18.48
<i>R</i> _{free} , %	22.42
rmsd ^c bonds, Å	0.01
rmsd ^c angles, °	1.12
# molecules in asymmetric unit	
Protein	4
sclx ₄	5
PEG	5
Mg ²⁺	3
Solvent	322
Average <i>B</i> factors, Å ²	
Protein	21.51
sclx ₄	19.70
PEG	45.46
Mg ²⁺	28.88
solvent	28.44
Ramachandran analysis ^d	
% residues in favoured regions	98.8
% residues in allowed regions	100.0

^aValues in parentheses correspond to the highest resolution shell; ^b $R_{\text{merge}} = \frac{\sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$; ^croot mean square deviation calculated from the *B* values of all non-hydrogen protein atoms; ^dCalculated with MolProbity.¹⁷⁶

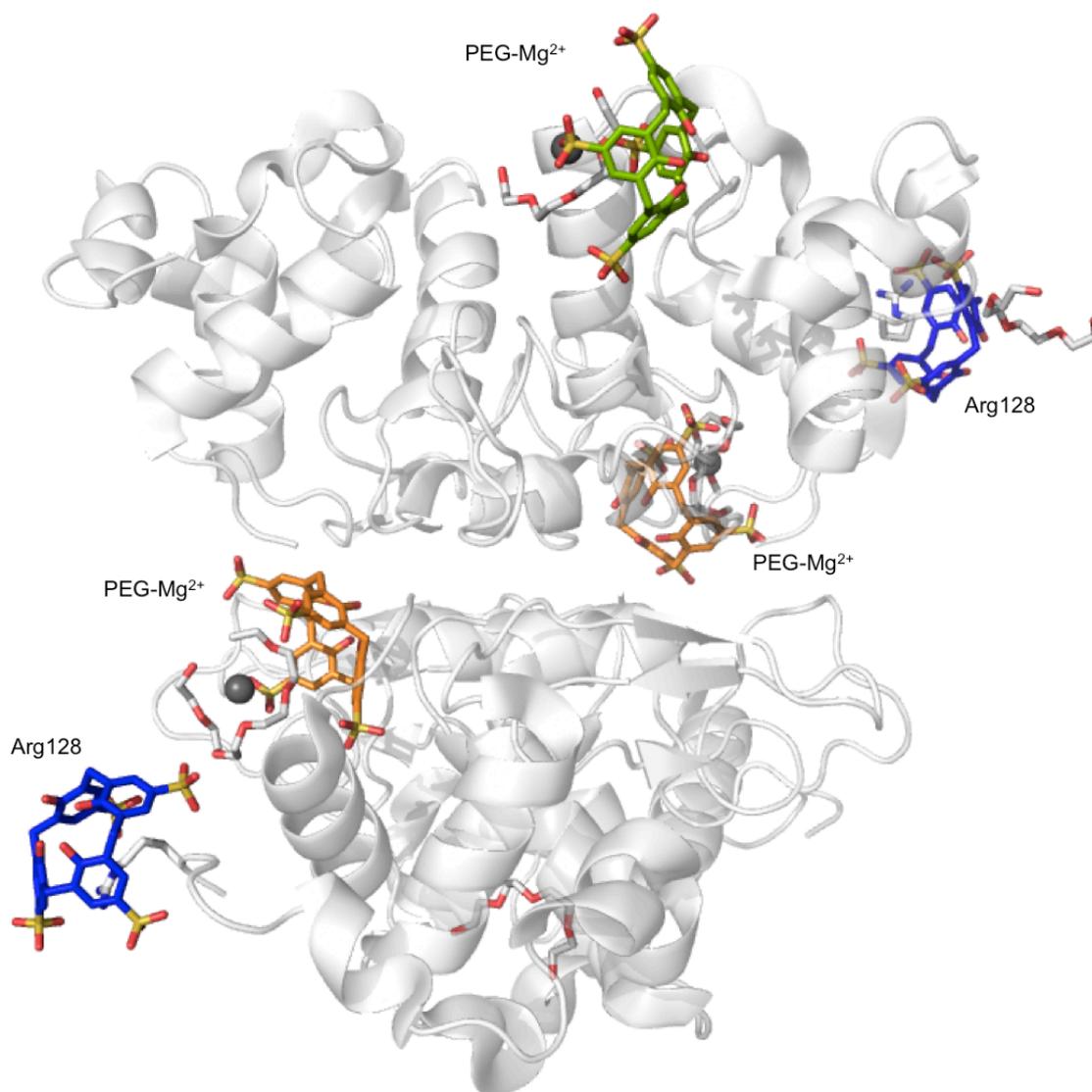


Figure 2. Crystal structure of the lysozyme–sclx₄ complex. The asymmetric unit comprises four molecules of lysozyme, five sclx₄, five PEG fragments and three Mg²⁺ ions. Lysozyme is represented as transparent grey cartoons. Sclx₄, PEG and arginine are shown as sticks and the Mg²⁺ ions are spheres. The sclx₄ are coloured by binding site. The two calixarenes that bind the C-terminal Arg128 are coloured blue and the two calixarenes that bind PEG-Mg²⁺ at the centre of the assembly are orange. The sclx₄ that binds PEG-Mg²⁺ near the active site is coloured green.

Molecular architecture of the lysozyme-sclx₄ assembly

The four lysozyme chains in the asymmetric unit assemble as a dimer of dimers (Figure 2). Each chain contributes a 3_{10} helix²⁴⁵ to the centre of the tetrameric assembly resulting in the formation of a channel. Pro79, Ser81, Ala82, Ser85, Ser86 and Asp87 protrude into the channel and either end is plugged by the hydroxyl-bearing rim of a pair of close-packed calixarenes. The sulfonated sclx₄ rims are surrounded by predominantly basic

protein surfaces. One of these sclx_4 encapsulates the side chain of C-terminal Arg128 from a neighbouring tetramer in the crystal packing environment. The adjacent sclx_4 is bound to a PEG fragment, which forms a crown ether-like complex around a Mg^{2+} cation. A fifth sclx_4 binds to a PEG fragment at the active site in chain B. In addition to protein- sclx_4 contacts, PEG- sclx_4 and protein-PEG contacts are widespread in the structure.

Sclx₄-sclx₄ packing

The sclx_4 pair plugs either end of the channel (Figure 3A and 3B). The packing motif consists of an up-up arrangement and is similar to but less pronounced than that observed in the sclx_4 -arginine crystal structure (Figure 3C).⁹¹ A phenyl ring from each sclx_4 is positioned 4.0-4.1 Å to the nearest methylene bridge of the neighbouring sclx_4 . Two water molecules hydrogen bond two sulfonates, one contributed by each sclx_4 . The side chains of two N-terminal lysines form salt bridges to a sulfonate on each sclx_4 . The N-terminal lysines and mediating water molecules work in concert to diffuse the dense negative charge of the sclx_4 pair. Water molecules are significant players in the crystal structure of sulfonated azo dyes with cationic amino acids.²⁴⁶ The crystal structure of the Orange G dianion with lysine shows a water molecule make hydrogen bonds with two sulfonates, one from each Orange G molecule. On the opposite side of the sulfonates, the N-terminal amine of a lysine amino acid makes salt bridges with the same two sulfonates. This noncovalent network involving waters and a lysine amino acid is similar to the interactions that occur with the adjacent sulfonates of two close-packed sclx_4 .

Arginine-sclx₄ binding sites

Both arginine- sclx_4 sites involve the entrapment of C-terminal Arg128 (Figure 3). The C^γ and C^δ of Arg128 are in the calixarene cavity and make van der Waals contact with three sclx_4 phenyl rings. The Arg128 guanidine group is projecting up, into the plane of the sulfonates. At the Arg128 binding site in chain A (A.Arg128), three sclx_4 sulfonates make a total of four salt bridges with Arg128. Three water molecules make hydrogen bonds with the guanidine group. The backbone amide of arginine makes hydrogen bonds with a water molecule and a sclx_4 sulfonate. Another water molecule hydrogen bonds both the backbone carbonyl and another sulfonate. These bridging water molecules interact with both the polar groups on the sclx_4 and the protein surface and were previously observed in the cyt *c*- sclx_4 structure.⁴⁸

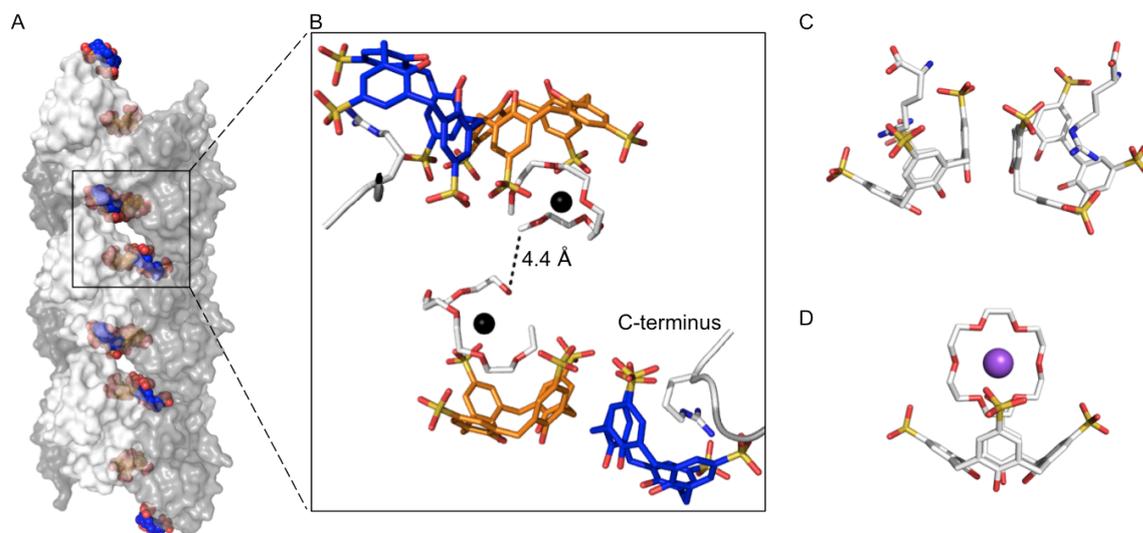


Figure 3. (A) An expansion of the crystal packing neighbours generated with the $P12_1$ symmetry operations. The protein chains are represented as grey surfaces and ligands are spheres that are coloured as before. (B) The opposing $sclx_4$ - $sclx_4$ pairs observed at the centre of the assembly are shown. The positioning of the $sclx_4$ -PEG complexes suggests one PEG fragment is entrapped by the two calixarenes. (C) The crystal structure of D-arginine and $sclx_4$ shows a similar packing arrangement for the calixarenes⁹¹ and (D) the crystal structure of $sclx_4$ with 18-crown-6 and Na^{+247} is similar to the $sclx_4$ -PEG- Mg^{2+} complex.

An analysis of crystal packing neighbours reveals two additional protein chains making noncovalent interactions with $sclx_4$. The $C^{\epsilon 1}$ of D.His15 hydrogen bonds the nearest sulfonate. The N^{ϵ} of D.Lys1 makes a salt bridge with two O atoms on one sulfonate and a third to the neighbouring $sclx_4$, which entraps a PEG fragment. The alkyl portion of the D.Lys1 side chain makes hydrophobic interactions with a phenyl ring from the adjacent $sclx_4$ -PEG- Mg^{2+} complex. These *exo* interactions involving a lysine side chain are also observed in the *cyt c*- $sclx_4$ crystal structure (Chapter 2). An interesting feature of this binding site is that the interactions with His15 and Lys1 are mirrored at the adjacent $sclx_4$ -PEG- Mg^{2+} complex. Other noncovalent interactions supplied by crystal packing neighbours to the Arg128- $sclx_4$ complex involve the N^{δ} of A.Asu74 and the backbone amide of D.Ile88, which hydrogen bond a sulfonate. The guanidine group D.Arg14 makes a salt bridge to one sulfonate. The C^{β} of A.Asu65 and D.Ser86 are within 4.2 Å of the nearest phenyl ring of and the C^{δ} of A.Pro79 is within 3.8 Å of a $sclx_4$ methylene bridge. All four $sclx_4$ hydroxyl groups are solvated by a water molecule. One of the water molecules also makes a hydrogen bond with the amino-terminus of B.Lys1. The interactions in chain A are comparable to the B.Arg128 and the 670 Å² interface area exists for both sites. This is the surface area of the $sclx_4$ binding site that becomes solvent

inaccessible in the total complex and includes all of the neighbouring protein chains generated using $P12_11$ symmetry operations.

PEG-sclx₄ binding sites

In the asymmetric unit three sclx₄ binding sites involve a PEG fragment bound within the sclx₄ cavity. One PEG-sclx₄ complex binds to the lysozyme active site (described in the next section).²⁴⁸ However, for the two other sites both calixarenes exist as part of the sclx₄ pairs. These sclx₄ are characterised by the submersion of a portion of the PEG fragment into the cavity, which is wrapped around a Mg²⁺. The PEG fragment forms a crown ether-like complex with Mg²⁺ in which the six O atoms coordinate Mg²⁺. The Mg-O bonds, in the 2.9-3.6 Å range, are considered relatively long in comparison to the crystal structure of 18-crown-6 with MgCl₂.²⁴⁹ The PEG-sclx₄ complex is similar the complex of sclx₄ with 18-crown-6 (Figure 3D).²⁴⁷ In the crystal structure two calixarenes form capsules that contain the 18-crown-6 ether, which coordinates a central Na⁺. Two water molecules make hydrogen bonds with Na⁺ and two sulfonates. This differs to the PEG-sclx₄ complex, for which the Mg²⁺ makes direct electrostatic interactions with two O atoms on a sulfonate.

The sclx₄-protein interactions involving Lys1, His15, Asn74, Pro79 and Ile88 are similar to those observed at the Arg128-sclx₄ binding site. The backbone amide of Val2 makes a hydrogen bond to one sulfonate. In the Arg128 binding site the interaction is mediated by a water molecule. The only PEG-protein interaction at this site involves the guanidine group of B.Arg14, which makes van der Waals contact with the PEG fragment. In the context of crystal packing neighbours, the two PEG-sclx₄ complexes oppose one another and may be anchoring the ends of one PEG molecule (Figure 3B). In doing so, order is induced in the two regions of the polymer.

PEG-sclx₄ binding at the active site

The third PEG-sclx₄ complex binds the surface of the lysozyme active site²⁴⁸ in chain B and is the only such site in the crystal structure. B.Arg61 and crystal packing neighbour A.Lys116 each make salt bridges with a sclx₄ sulfonate (Figure 4). One sulfonate is positioned 2.4 Å and 3.2 Å from the backbone amides of B.Gly104 and B.Asn103, respectively. Chain A is the only crystal packing neighbour at the binding site. The hydroxyl group of A.Tyr23 and the N^{δ2} of A.Asn106 make hydrogen bonds with a sclx₄

sulfonate. The phenyl ring of A.Tyr23 partly stacks above a sclx₄ phenyl. The lower sclx₄ rim is solvated by five water molecules. One water hydrogen bonds both the N^{δ2} of A.Asn27 and a sclx₄ hydroxyl.

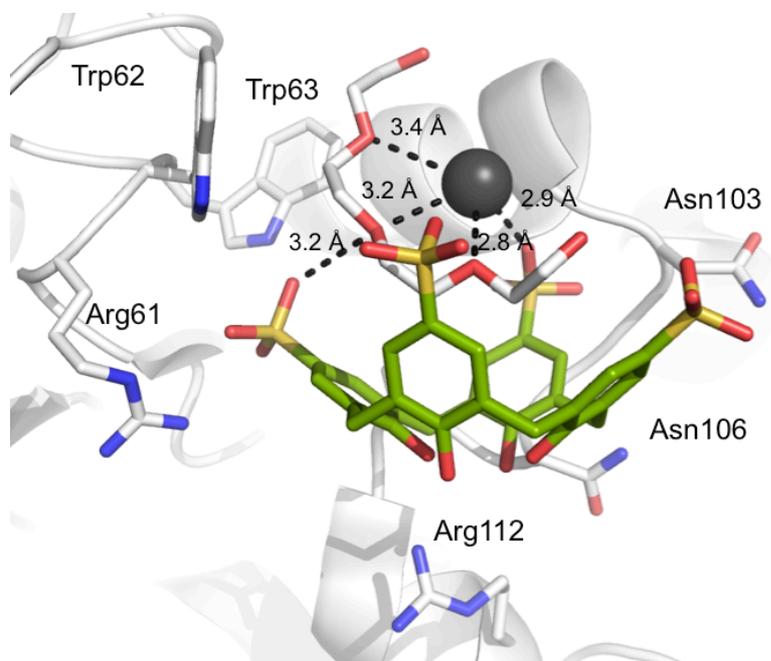


Figure 4. The sclx₄-PEG-Mg²⁺ complex near the lysozyme active site. A PEG fragment, partially in the sclx₄ cavity, adopts a crescent shape around Mg²⁺ (sphere), similar to the crown ether.²⁴⁷ The interactions involving the PEG fragment and both Mg²⁺ and sclx₄ are shown. The PEG also makes noncovalent interactions with the side chains of Trp62 and Trp63. The N^ε and N^{η2} of Arg61 make salt bridges with one

sulfonate.

The PEG fragment adopts a crescent shape around Mg²⁺ and, similar to the two other PEG-sclx₄ binding sites, Mg²⁺ interacts with three PEG O atoms 2.8-3.4 Å apart (Figure 4). Mg²⁺ also makes a salt bridge with a sclx₄ sulfonate and hydrogen bonds a water that also hydrogen bonds a sulfonate. A portion of the PEG fragment lies in the sclx₄ cavity and is within van der Waals distance of two sulfonates. Part of the PEG fragment is above the plane of the B.Trp62 and B.Trp63 side chains and makes van der Waals contacts with B.Trp62. Interestingly, an NMR titration of lysozyme with PEG suggests that PEG preferentially binds the interaction site centered on Trp62.²⁵⁰ Out of all the residues, Trp62 and Trp63 display the largest chemical shift change in the presence of PEG, with a *K_d* value of 80 mM.²⁵⁰ Arg61 and Arg73, which surround Trp62, also display a significant chemical shift change during the titration. PEG binds hydrophobic surface patches²⁵¹ and the side chains of B.Ile98, B.Trp62, B.Trp63 make the site hydrophobic in nature.

The residues Asn59, Trp62, Trp63, Ile98, Ala107 and Trp108 are central to lysozyme activity.²⁴⁸ The artificial, molecular switch containing three bisphosphonate dianions, one glucosamine moiety and one unpolar dodecyl tail binds lysozyme tightly

($K_d = 30$ nM) and inhibits its function.^{252,253} The glucosamine moiety was used as a substrate mimic to bind and block the active site. The bisphosphonate groups promote charge-charge interactions with the cationic protein surface similar to the sclx₄ sulfonates, and the dodecyl tail favours hydrophobic interactions similar to PEG. Lysozyme's function was restored by the addition of polyarginine, suggesting a strong electrostatic contribution to binding.

Lysozyme-anion binding sites

Coupled with the known cyt *c*-sclx₄ and the lysozyme-sclx₄ complexes, the structures deposited in Protein Data Bank (PDB) provide a guide for the design of surface binding ligands. Of the 446 HEWL structures deposited in the PDB, three are binding sulfates, seven are binding sulfonates, seven are binding phosphates and one is binding a phosphonate (Table 2). The strategy of co-crystallising lysozyme with proteins, such as antibodies, is useful for investigating protein surface recognition at the complex interface.²⁵⁴⁻²⁵⁷ Four of the lysozyme-protein complexes listed in Table 2 show one phosphate ion²⁵⁴⁻²⁵⁸ and one sulfate ion²⁵⁹ at an interaction interface in the crystal structure.

A superposition of the known cyt *c*-sulfate and cyt *c*-sclx₄ binding sites highlights a C^αNN anion binding motif²⁶⁰ that reorganises on sclx₄ sulfonate binding (Chapter 2). Similarly, overlaying the crystal structures of the lysozyme-sulfate²⁵⁹ and lysozyme-sclx₄ complexes suggest that the Arg128 site is optimised for anion recognition (Figure 5). At the sulfate binding site, the guanidine groups of both Arg125 and Arg128 each make two salt bridges with the anion. The Arg128-sclx₄ binding site involves the guanidine group of Arg128 only. Arg125, points away from the binding site and makes salt bridges with the side chain of Asp119 in a bidentate manner. The superposition of lysozyme-sclx₄ and another lysozyme-sulfate crystal structure shows that the interactions made by Arg14 and His15 to one sclx₄ sulfonate are echoed in the lysozyme-sulfate complex.²⁶¹ This suggests a second site optimised for anion recognition. The N^{ε1} of B.Trp63 makes a hydrogen bond with one sclx₄ sulfonate in another binding site. Similarly, trifluoromethanesulfonic acid is also reported to hydrogen bond the N^{ε1} of Trp63 in addition to the backbone amide of Asn59 and a water molecule via its sulfonate.²⁶² In the trifluoromethanesulfonic acid complex, Arg61 is pointing away from the binding site and Trp62 pointing into it. In the

presence of sclx₄ this is reversed and the Arg61 side chain points into the binding site, making three salt bridges to the sulfonate.

Table 2: Polar interactions in lysozyme-anion complexes.

PDB ID	Anion	Salt bridges [¶]	Hydrogen bonds [§]	Reference
4glv	Sulfate	N ⁿ¹ of E.Arg14	C ^{e1} of E.His15	††
2gv0	Sulfate	N ^e of A.Lys97	A.Gln124 N ^{e2} , A.Ala88 [†] and A.Gly1 [†] backbone amides	††
2i26	Sulfate 1	Q.Arg128 N ⁿ¹ , L.Arg125 N ^e	-	259
	Sulfate 2	Q.Arg125 N ⁿ¹	-	
	Sulfate 3	Q.Arg14 N ^e and N ⁿ²	C ^{e1} of Q.His15	
1ps5	Sulfate	A.Arg14 N ⁿ¹ , A.His15 N ^{e2}	-	261
4glv	HEPES	C.Arg21 N ⁿ¹	-	††
3b6l	Dodecyl sulfate	A.Arg73 N ⁿ²	A.Asn77 N ^{δ2} , A.Asn74 backbone amide	263
1b0d	<i>p</i> -toluene sulfonate	-	A.Asn44 N ^{δ2}	264
2hub	HEPES	-	A.Asp48 and A.Thr47 backbone amides	††
2q0m	Trifluoro- methanesulfonic acid	-	N ^{e1} of X.Trp63, X.Asn59 backbone amide	262
3e3d	HEPES	A.Arg21 N ⁿ¹ and N ⁿ²	A.Thr47 O ^{r1} and backbone amide	265
3rz4	HEPES	A.Arg21 N ⁿ¹ and N ⁿ²	A.Thr47 O ^{r1}	††
4glv	Phosphate	G.Arg21 N ⁿ¹	-	††
3m18	Phosphate	B.Lys33 N ^e	-	257
1zv5	Phosphate	L.Arg.45 N ⁿ¹ and N ⁿ²	-	256
1zvy	Phosphate	B.Arg128 N ⁿ¹	C ^α of Gly126 B.Gly126, B.Cys127 and B.Arg128 backbone amides	256
1rjc	Phosphate	B.Arg14 [†] N ^e	Q.His15 [†] C ^{e1}	258
1g7i	Phosphate	C.Lys13 N ^e	C.Leu129 C-terminal carboxyl	255
1a2y	Phosphate	C.Lys13 N ^e	C.Leu129 C-terminal carboxyl	254
4fjr	Ribose-5-phosphate	A.Arg125 N ⁿ¹ and N ⁿ²	-	††

[¶]Number of salt bridges between anion and the protein. Defined as anion O <3.8 Å. [§]Number of hydrogen bonds between the anion and protein. [†]Contacts observed by analysis of crystal packing neighbours. ^{††}Unpublished.

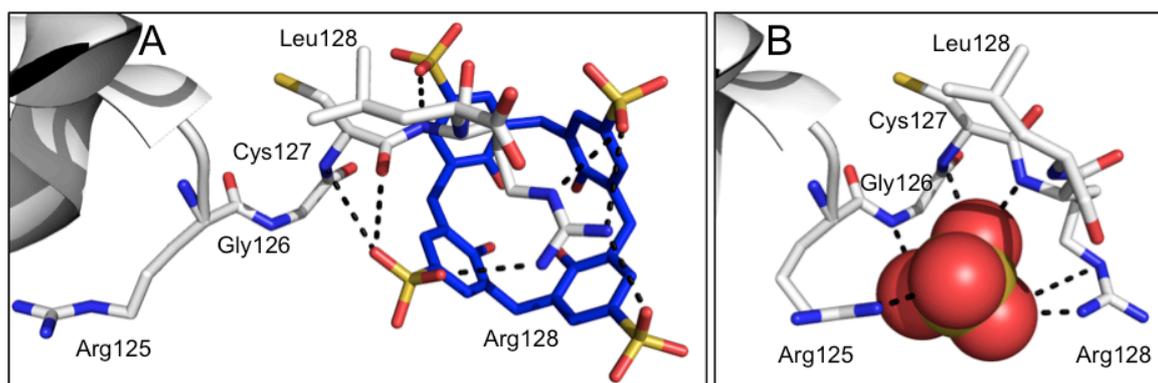


Figure 5. A comparison of the lysozyme-sclx₄ and lysozyme-sulfate binding sites. **(A)** The Arg128 side chain, entrapped by sclx₄, makes four salt bridges and one hydrogen bond with the sulfonates. The sulfonate that coincides with the sulfate makes two hydrogen bonds with the backbone amide and carbonyl of Cys127. **(B)** The sulfate makes hydrogen bonds with the backbone amides of Gly126, Cys127 and Arg128.²⁵⁹ Both Arg125 and Arg128 make salt bridges with the sulfate. The described residues (coloured grey) and calixarene (coloured blue) are shown as sticks and the sulfate is represented as spheres.

Conclusions

Similar to *cyt c*, lysozyme crystallized in PEG-containing solutions in the presence of sclx₄. In the asymmetric unit the calixarene was present at five binding sites on the protein surface and each site involved extensive interactions with lysozyme. Therefore, the presence of sclx₄ at protein interfaces validates its ‘molecular glue’ role. PEG and Mg²⁺ complexation by sclx₄ resulted in the ordering of a portion of the polymer, which is expected to have a high entropic cost. An expansion of the crystal packing neighbours shows two opposing sclx₄-PEG-Mg²⁺ complexes and the structure is reminiscent of the sclx₄ capsule comprised of two calixarenes encapsulating a crown ether-Na⁺ complex.²⁴⁷

Synthetic molecules have previously demonstrated the ability to bind N-terminal side chains. For example, cucurbit[7]uril selectively binds the N-terminus of human insulin.⁴⁴ Similarly, methyl- β -cyclodextrin binds a glucose-dependent insulinotropic polypeptide.²⁶⁶ In the lysozyme-sclx₄ crystal structure the calixarene binds C-terminal arginine. This is an interesting result for two reasons; firstly for the residue type (arginine) and secondly for the position near the C-terminus. We have previously shown with *cyt c* that sclx₄ displays a preference for lysines (Chapter 2). Instead of binding Lys1 of lysozyme, which is a highly cationic site due to the N-terminal amine, sclx₄ entrapped the second last residue Arg128. This raises the question, are these differences a question

of probability? The contribution of the arginine side chain to the binding free energy in protein–phosphate complexes is considered to be much larger than that of a lysine side chain.²⁶⁷ This phenomenon is likely to be due to the “chelate effect” of the guanidine group, which can similarly occur in arginine-sulfate complexes. Sclx₄ binds lysine side chains with a higher affinity than arginine in phosphate buffer solution.⁹⁵ Nonetheless, the abundance of arginine side chains renders arginine binding more accessible. The attractive forces that drive salt bridge formation and, in turn, sclx₄ complexation in this instance outweigh lysine entrapment.⁴⁸ Binding to and ordering the flexible C-terminus is expected to have a high entropic cost, however, the site appears to be optimised for anion recognition. For example, the crystal structure of lysozyme with the shark single-domain antibody shows sulfate binding at the same site (Figure 5).²⁵⁹ Furthermore, proteins co-crystallized with lysozyme in phosphate buffer also show a phosphate ion making a water-mediated salt bridge with lysozyme’s Arg128 (Table 2).^{251,252} This similarly hints at an anion recognition site.

Water-soluble, flexible, linear copolymers that target specific surface patches are useful protein surface binders as they are capable of induced-fit.²⁵² The bisphosphonate co-polymer, which complements both the surface and active site of lysozyme, acted as a molecular switch for protein function. The bisphosphonate and dodecyl portions may interact with the protein surface in a similar manner to the PEG-sclx₄ complex at the active site, which involved a combination of electrostatic and hydrophobic interactions.

Artificial symmetric scaffolds are useful regulators of protein interactions, extending to biosensor²⁶⁸⁻²⁷⁰ and chemical^{271,272} sensor design. The four-fold symmetry of calix[4]arenes has been exploited generate molecular capsules.⁷⁴ Oppositely charged upper rims drive the association of anionic and cationic calixarenes. Similarly, the symmetry of calix[4]arenes was exploited to bind multiple protein domains simultaneously.¹¹⁰ In the lysozyme-sclx₄ complex, the four-fold symmetry of sclx₄ was useful for protein packing. Close-packed sclx₄ pairs mediated the tetrameric interfaces (Figure 3A). The resultant lysozyme complex is similar to the proposed calix[4]arene-assisted p53 tetramerisation domain structure.¹¹⁰

Arginine is typically involved in cation- π interactions at protein-protein interfaces.^{59,237} However, attractive ionic interactions are influential for arginine complexation by particular ligands.^{40,60,239-241,273} Sclx₄ bound arginine resulting in the new crystal packing arrangement for lysozyme involving a dimer of dimers. In this instance,

salt-bridges and not the cation- π effect are observed between Arg128 and the calixarene. Knowledge of anion binding sites in the protein can be used to identify where the calixarene is likely to bind. The crystal structure of lysozyme with sclx₄ might guide designed ligand-assisted protein self-assembly.

Experimental

Lysozyme-sclx₄ co-crystallization

Hen white egg lysozyme, purchased from Sigma-Aldrich (62971 Fluka), was used without further purification. The hanging drop vapour diffusion method was used for crystallization trials at 20 °C. Drops were prepared by combining 1 μ L volumes of lysozyme (1.4 mM), sclx₄ (17 mM) and the reservoir solution. Control drops were prepared by substituting the addition of sclx₄ with 1 μ L of water. The crystallization conditions that yielded diffraction-quality crystals of the cytochrome *c*-sclx₄ complex were used. The lysozyme-sclx₄ structure was determined from a crystal grown in a solution containing 24 % PEG 8000, 50 mM NaCl, 100 mM MgCl₂ 50 mM sodium cacodylate pH 6.3.

Data collection and X-ray structure determination

Crystals were cryoprotected in μ L aliquots of the reservoir solution supplemented with 25 % glycerol and flash-frozen in liquid nitrogen (100 K). Diffraction data were collected on a Pilatus 6M detector, S/N 60-0106, Soleil, the French national Synchrotron Facility, Paris, France. Intensity data were collected (φ scans of 0.1° over 180°) to a resolution of 1.7 Å from a single crystal. Data processing and scaling were performed in MOSFLM and SCALA, respectively.^{170,171}

The data collection and refinement statistics are given in Table 1. The structure was determined by molecular replacement in PHASER and the refinement and manual rebuilding were performed in REFMAC5 (as implemented in CCP4) and COOT, respectively.^{174,274} Coordinates for sclx₄ were obtained from the PDB ID t3y and built into the model using COOT. Solvent molecules were placed automatically using ARP/wARP and refinement was continued until no features remained in the $F_o - F_c$ difference maps.²⁷⁵ Molprobit was used to check the qualities of the structures prior to deposition

in the Protein Data Bank [4prq].¹⁷⁶ Protein-protein and protein-ligand interfaces were analysed in COOT and PISA.²⁰⁵

Chapter 5:

Crystal Structure of dimethylated lysozyme and *p*-sulfonatocalix[4]arene

Manuscript submitted

Abstract

Lysine methylation results in significant structural reorganisation of chromatin. As a result, great efforts are focused on designing molecules that bind methylated lysine side chains. The goal is to identify highly specific small molecules that can be applied to alter chromatin self-assembly in a controlled manner. Lysozyme was subjected to reductive alkylation so that all of the lysines were dimethylated. Crystals of the modified lysozyme grew within minutes in the presence of *p*-sulfonatocalix[4]arene. The crystal structure shows remarkable structural reorganisation compared to the native lysozyme-sclx₄ complex. Significantly, this is the first crystal structure showing the complexation of a dimethylated lysine residue by calixarene. The structure also shows that the cation- π interaction plays a crucial role in the binding event.

Introduction

Chromatin is the biological material that is responsible for controlling all eukaryotic genetic information. Minor structural and chemical changes of histone proteins, via posttranslational modification, present triggers for chromodomain remodelling.^{31,276} The modifications including acetylation, methylation and phosphorylation are subtle. Nonetheless, they are crucial for promoting specific protein-protein interactions that are interpreted to provide numerous biological functions.^{238,277} Therefore, posttranslational modifications are the biological cues for chromatin function and as a result, much attention is focused on targeting these switches so that the ensuing cascade of reactions can be controlled.

Nucleosomes are the fundamental building blocks of chromatin and consist of a core histone octamer around which DNA is wrapped.²⁷⁸ The posttranslational modification of histone proteins affects chromatin assembly and occurs at the flexible, charged N-termini that protrude out from the nucleosome assembly. The enzymes that recognise the modifications are highly specific.^{279,280} For example, methylated Lys9 in histone H3 is recognised by heterochromatin protein 1 (HP1), which manages the binding of other proteins and thus controls chromatin structure and gene expression.²⁷⁹ The protein domains that recognise and bind *N*-methylated lysine are described as an “aromatic cage” and consist of an ordered enclosure of aromatic side chains into which the lysine side chain inserts (Figure 1A).^{190,191} Cation- π interactions between the methylated side chain and aromatic cage, rather than the hydrophobic effect, are the major driving force for binding.²⁸¹ Lysine side chain methylation shares a role with numerous cellular activities in addition to chromatin function. For example, the methylation of the C-terminus lysine of p53 tumour suppressor protein is important for regulating protein-protein interactions.^{7,282} Therefore, lysine residues bearing post translational modifications govern a variety of cellular activities and are important targets for therapeutic intervention.^{277,283}

Inhibitors that mimic the aromatic cage motif, owing to a predefined cavity and polar or charged periphery, have potential as modulators of protein behaviour. The molecular tweezers has mM affinity for lysine in 25 mM NaH₂PO₄, pH 7 (Figure 1B).⁶⁰ The affinity and selectivity of a synthetic receptor for trimethyllysine (LysMe₃) is comparable to the native HP1 chromodomain (Figure 1C).¹⁸² The receptor named *rac*-A₂B, exhibits increasing affinity for higher lysine methylation states. This observation is

attributed to differences in the magnitude of the cation- π interactions and LysMe₃ is expected to have an enhanced cation- π propensity over LysMe₂, LysMe and lysine. *p*-sulfonatocalix[4]arene (sclx₄) has a dissociation constant (K_d) of 0.8-1.7 mM for lysine residues of cytochrome *c* in phosphate buffered solution, pH 6.0 (Figure 1D).⁴⁸ Sclx₄ also displays ~ 30 μ M affinity for LysMe₃ as the free amino acid and also as part the Arg-LysMe₃-Ser-Thr peptide in phosphate buffered solution, pH 7.4.⁹⁵ The calixarene can discriminate between unmethylated, mono-, di-, and trimethylated lysines on a single histone tail sequence.¹⁰⁰ Cucurbit[7]uril is a neutral non-aromatic host molecule (Figure 1E) that exhibits large binding constants for LysMe₃ and selectivity for lysine.⁶¹ The hydrophobic effect and ion-dipole interactions influence cucurbit[7]uril binding. The molecular tweezer⁶⁰, *rac*-A₂B¹⁸² and cucurbit[7]uril⁶¹ involve side chain threading through the cavity while sclx₄ can entrap side chains⁴⁸. Cucurbit[7]uril lacks the ability to make cation- π and charge-charge interactions, nonetheless, it follows the same trend as sclx₄ and *rac*-A₂B for binding LysMe_n. The affinity for interaction is LysMe₃ > LysMe₂ > LysMe > Lys. Significantly, this is similar to the recognition pattern of LysMe_n protein readers such as the chromodomain HP1.¹⁸²

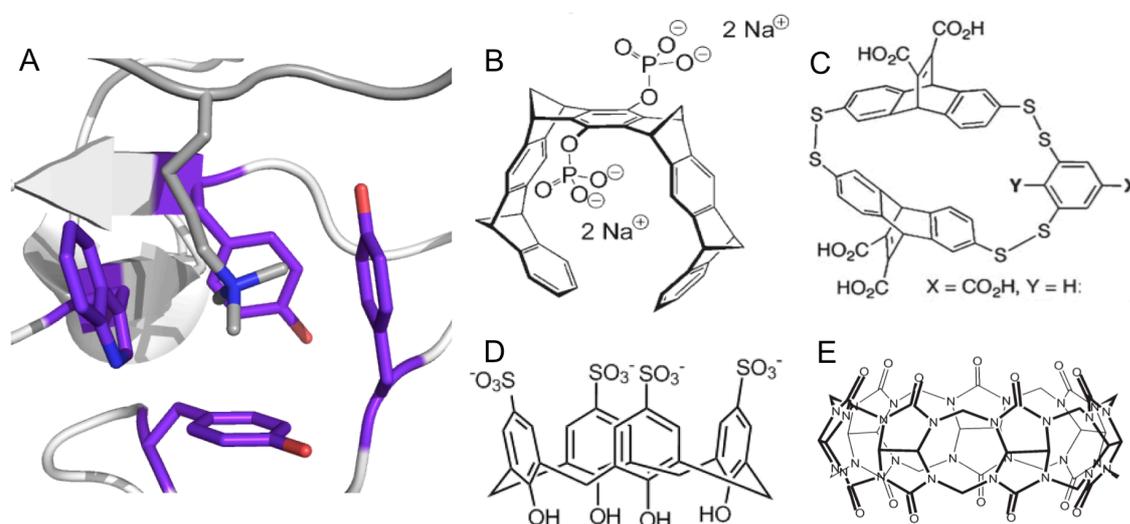


Figure 1: A naturally occurring aromatic cage, which favours LysMe₃ binding and the artificial mimics. (A) Structure of the BPTF PHD finger of NURF bound to the histone H3Lys4Me₃ peptide (dark grey) with the aromatic pocket shown as purple sticks (PDB ID 2fuu).¹⁸³ (B) Molecular tweezers,⁶⁰ (C) *rac*-A₂B¹⁸² and (D) sclx₄⁹⁵ can make cation- π interactions and salt bridge interactions with lysine side chains. (E) Cucurbit[7]uril binds LysMe₃ with an affinity of ~ 6 μ M and the hydrophobic effect contributes to the binding.⁶¹

The interactions of sclx₄ were structurally characterised with both cytochrome *c* and lysozyme, in Chapters 2 and 4, respectively. The crystal structures showed extensive charge-charge interactions between sclx₄ and the entrapped lysine and arginine side chains. The goal of this project was to substantiate the potential of sclx₄ as a protein surface binder in a system containing LysMe₂ and to further investigate its effect on protein self-assembly. X-ray crystallography was used to achieve this and lysozyme bearing LysMe₂ was the model system. The structural consequences for sclx₄-assisted assembly are described.

Results and Discussion

Rapid lysozyme-KMe₂ co-crystallization

Lysozyme was subjected to reductive alkylation resulting in the dimethylation of all the lysine residues and the N-terminus, yielding lysozyme-KMe₂.²⁸⁴ The hanging drop vapour diffusion method was used for lysozyme-KMe₂ co-crystallization with sclx₄. Crystallization drops were prepared using the conditions used to grow both cyt *c*- and lysozyme-sclx₄ crystals. Sclx₄ addition to lysozyme-KMe₂ resulted in immediate precipitation and crystals grew from these dense white precipitates. Rapid crystal growth occurred when the protein and ligand concentrations were decreased 10- and 100-fold, respectively. Furthermore, a precipitant such as PEG was not required for crystallization. The NMR and X-ray crystallographic data for the cyt *c*-calixarene complex suggested that sclx₄ can camouflage and modify the protein surface.⁴⁸ The impact of the protein surface camouflage scenario was tested using the crystallization setup with 20 μM lysozyme-KMe₂ and different sclx₄ concentrations. The effects were tested with the crystallization solution containing 20 mM salt (NaCl or Na₂SO₄), 10 mM sodium cacodylate pH 6. Crystals grew overnight in all drops except those containing eight-fold excess sclx₄. These drops initially yielded dense white precipitate from which crystals grew after 24 h. In NaCl, crystal growth improved when the sclx₄ concentration was increased from 1 to 20 μM (Figure 2). However, at two-fold the lysozyme-KMe₂ concentration, protein crystallization declined. This might indicate that the larger number of ligands competing and roaming the protein surface make self-assembly more difficult. A similar trend was observed in Na₂SO₄.

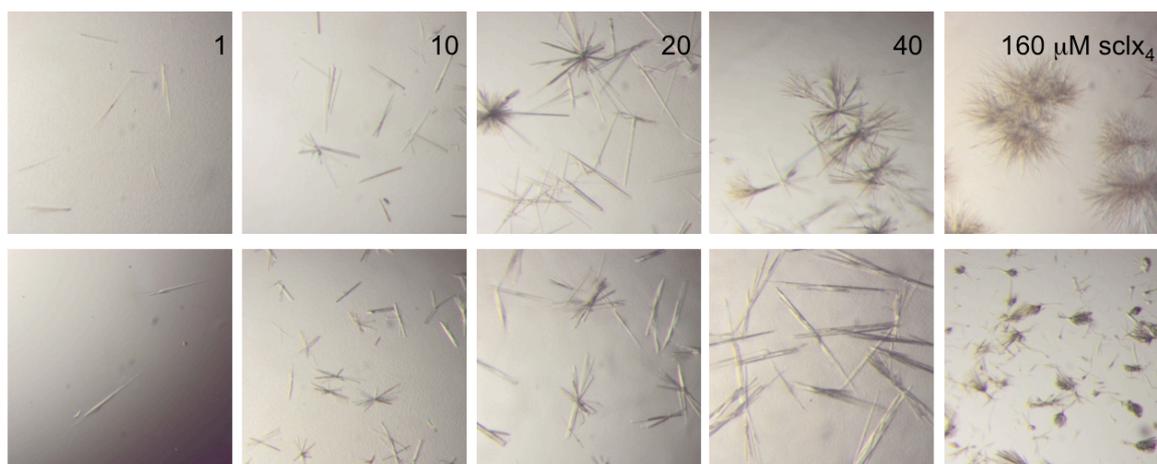


Figure 2: Calixarene concentration influenced lysozyme-KMe₂ crystal growth. The crystallization drops contained 20 μM lysozyme-KMe₂, 20 mM salt and 10 mM (CH₃)₂AsO₂Na, pH 6 and the sclx₄ concentration varied from 1-160 μM. The salt was NaCl (top panel) or Na₂SO₄ (bottom panel).

X-ray crystallography

X-ray crystallography was used for the structural characterisation of the lysozyme-KMe₂ crystals. The crystals belonged to space group *P12₁1* and diffracted to 1.8 and 2.1 Å. The 2.1 and 1.8 Å data sets were solved to 2.2 and 1.9 Å, respectively. For both crystal structures, the asymmetric unit is composed of two molecules of lysozyme-KMe₂ and four molecules of sclx₄ (Table 1). The two protein molecules are referred to as chains A and B. Two glycerol molecules are present in the asymmetric unit of the 2.2 Å structure. Five glycerol molecules, two Mg²⁺ ions, one Na⁺ and three Cl⁻ are also present in the asymmetric unit of the 1.9 Å structure.

Molecular architecture of the sclx₄-assisted lysozyme-KMe₂ assembly

The self-assembly of the two lysozyme-KMe₂ structures are similar and one of four sclx₄ mediates two lysozyme-KMe₂ chains in the asymmetric unit (Figure 3). Sclx₄ binds LysMe₂ at position 116 and arginine at position 14 *endo* in both chains A and B. Subtle differences between the 1.9 and 2.2 Å structures arise from a selection of residues that occupy alternate conformations (Asn59, Arg68, LysMe₂97, Asn103 and Arg114). For example, Arg68 makes hydrogen bonds from the N^{η1} to the backbone carbonyl of Ser100 and from the N^{η2} to the -OH of Tyr53 in the 1.9 Å structure. In 2.2 Å structure Arg68 makes a hydrogen bond from the N^{η1} to the backbone carbonyl of Asp101. Importantly, the residues that occupy alternate conformations do not affect the calixarene binding sites.

The following descriptions of the LysMe₂ and arginine binding sites are based on the 1.9 Å structure.

Table 1. Crystallographic data for the lysozyme-KMe₂ complexes with sclx₄ at 1.9 and 2.2 Å.^a

	1.9 Å structure	2.2 Å structure
Space group	<i>P</i> 12 ₁ 1	<i>P</i> 12 ₁ 1
Cell constants	a = 45.85 Å b = 30.20 Å c = 94.21 Å α = γ = 90° β = 96.31°	a = 44.04 Å b = 30.28 Å c = 95.01 Å α = γ = 90° β = 96.62°
Resolution, Å	31.32-1.90 (1.95-1.90)	39.41-2.20 (2.26-2.20)
Wavelength, Å	0.95372	1.033190
Unique reflections	20709 (19591)	44504 (13557)
Multiplicity	3.6 (3.5)	3.3 (3.3)
<i>I</i> /σ	9.2 (2.6)	7.9 (3.7)
Completeness, %	94.6 (90.1)	99.6 (98.9)
<i>R</i> _{merge} , ^b %	8.4 (41.1)	13.0 (73.4)
Solvent content, %	45.02	49.53
<i>R</i> _{factor} , %	16.77	18.00
<i>R</i> _{free} , %	21.04	23.52
rmsd ^c bonds, Å	• 0.01	0.01
rmsd ^c angles, °	1.19	1.53
# molecules in asymmetric unit	2	2
Protein	4	4
sclx ₄	1	-
Na ⁺	2	-
Mg ²⁺	3	-
Cl ⁻	5	2
Glycerol	191	97
Solvent		
Average <i>B</i> factors, Å ²		
Protein	13.12	27.57
sclx ₄	17.79	33.56
Na ⁺	16.85	-
Mg ²⁺	19.13	-
Cl ⁻	18.12	-
Glycerol	25.64	38.37
Solvent	23.75	30.65
Ramachandran analysis ^d		
% residues in favoured regions	97.95	98.0
% residues in allowed regions	100.0	98.5

^aValues in parentheses correspond to the highest resolution shell; ^b $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - [I(hkl)]|}{\sum_{hkl} \sum_i I_i(hkl)}$; ^croot mean square deviation calculated from the *B* values of all non-hydrogen protein atoms; ^dCalculated with Molprobitry.

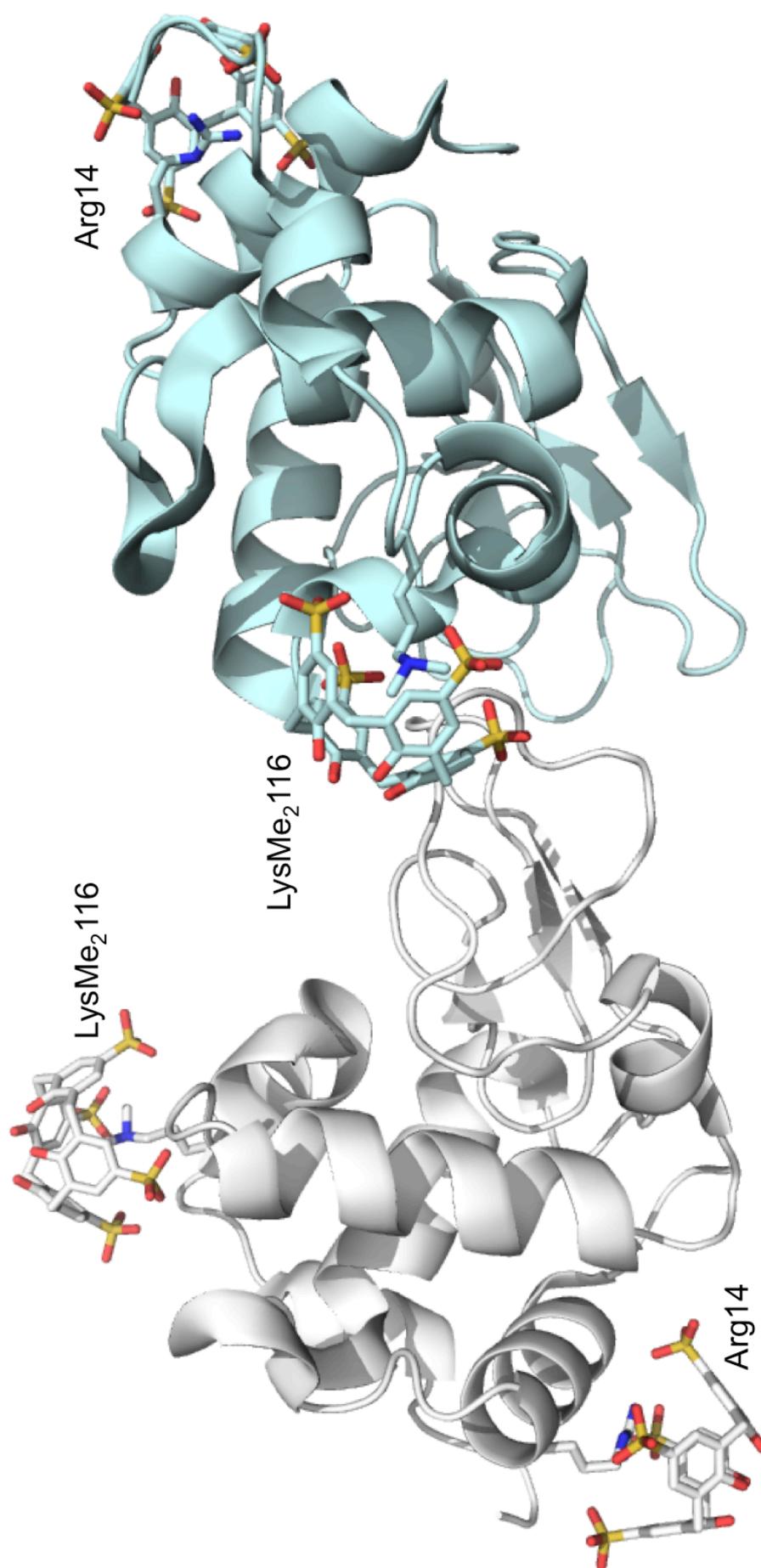


Figure 3: Crystal structure of the lysozyme-KMe₂ complex with sclx₄. The asymmetric unit comprises two molecules of lysozyme-KMe₂ (chains A and B in grey and cyan, respectively) and four molecules of sclx₄.

The LysMe₂ binding sites

The LysMe₂ binding sites are characterised by the inclusion of the side chain in the sclx₄ cavity. The N^ε of A.Lys116Me₂ (A.Lys116Me₂ corresponds to dimethylated Lys116 in chain A) is 4 Å from the nearest sclx₄ sulfonates. One methyl group is buried deep within the aromatic cage and is 3.8-4.1 Å from the centroids of each of the sclx₄ phenyl rings. Lysine methylation polarizes the C^ε-N^η bond, which increases the cationic character of the *N*-methylammonium group and its ability of the group to make CH⁺⋯π interactions.⁵⁸ The Lys116 methyl group makes cation-π interactions with each of the calixarene phenyl rings. The second methyl group is positioned between two sulfonates. The C^η1 makes salt bridge interactions with two adjacent sulfonates. A water molecule, 2.7 Å away, hydrogen bonds the N^ε of Lys116Me₂, in addition to an O atom on two sclx₄ sulfonates, completing the side chain-sclx₄ interaction network (Figure 4A). The Lys116Me₂-sclx₄ binding site, including the interaction with a water molecule, is reminiscent of the structural features and noncovalent interactions that are observed in the structure of the PHD finger and the tail peptide of histone H3Lys9Me₂ (Figure 4B).^{183,190} A glycerol molecule fills the space between chain A and sclx₄. Each of the glycerol -OH groups is involved in hydrogen bonds with a sclx₄ sulfonate or the protein.

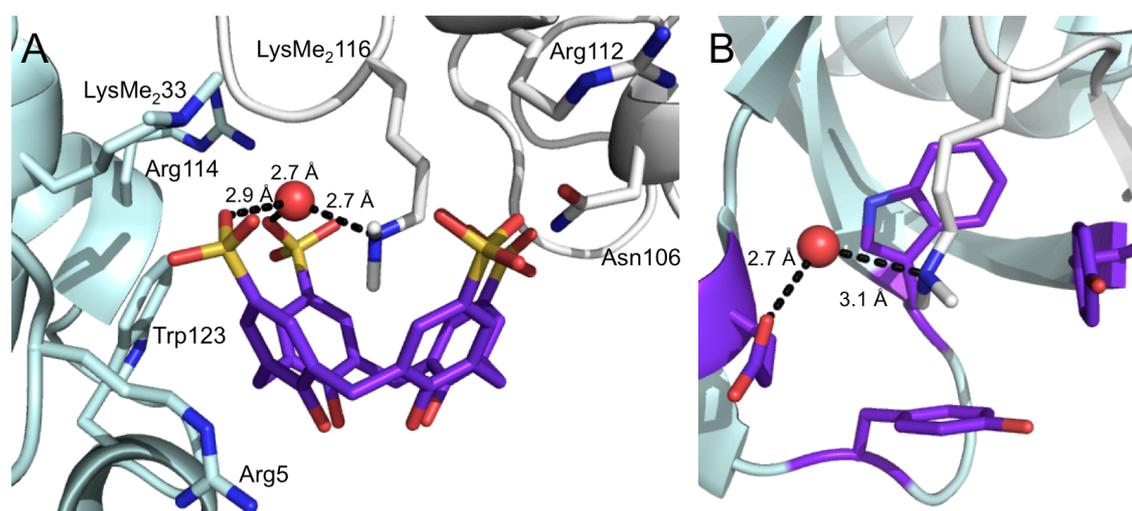


Figure 4: Aromatic cages that bind LysMe₂. **(A)** One LysMe₂ methyl group is buried in the sclx₄ cavity and the other is in the sulfonate plane. A water molecule (red sphere) hydrogen bonds both sclx₄ and LysMe₂. Sclx₄ is shown as purple sticks and the residues that make noncovalent interactions with sclx₄, including those from a close-packed chain, are also sticks. The protein chains are coloured as in Figure 3. **(B)** The crystal structure of PHD finger with H3Lys9Me₂ shows the histone LysMe₂ in the aromatic cage.¹⁸³ A water molecule coordinates both the glutamate and LysMe₂ side chains and a similar interaction occurs in the sclx₄-LysMe₂ complex. The cage motif and LysMe₂ are purple and grey sticks, respectively.

An expansion of the crystal packing neighbours shows two additional chains assembled around the calixarene (Figure 4). B.Lys33Me₂ makes a salt bridge with a sclx₄ sulfonate. The N^ε of B.Arg5 makes cation-π interactions with a sclx₄ phenyl ring. The backbone amide of B.Gly71 hydrogen bonds a sulfonate and the C^α is 4.0 Å to the centroid of the corresponding sclx₄ phenyl ring, making CH-π interactions. A sclx₄ methylene bridge is 4.5 Å from the centroid of the B.Trp123 phenyl ring. Two lower rim hydroxyl groups make a hydrogen bond with a water molecule.

The arginine binding sites

Sclx₄ binds Arg14 *endo* in both chains and the protein-sclx₄ interactions at both sites are similar. The Arg14 C^δ and guanidine group are in the plane of the sclx₄ sulfonates and three sulfonates make a total of four salt bridges with Arg14 (Figure 5). Two water molecules solvate the arginine side chain and hydrogen bond three sclx₄ sulfonates. The C^ε of His15 makes a water-mediated hydrogen bond with a sulfonate on the calixarene that also makes a salt bridge with the Lys1 N^ε. This interaction is similar to the native protein-calixarene structure, which involves a direct hydrogen bond between the His15 C^ε and the sulfonate (Chapter 4). Although Arg128 doesn't make contact with the sclx₄ at the Arg14 binding site, the N^{η1} is 6.0 Å from the nearest sclx₄ sulfonate. This feature was also observed at a phosphate binding site involving Arg14 and His15.^{254,255}

Two crystal packing neighbours are packed around the arginine-calixarene complex. The 3₁₀ helix²⁴⁵ that was the centre of the tetrameric assembly in the native lysozyme-sclx₄ structure is packed up against one side of the calixarene (Figure 5). From this 3₁₀ helix the side chain -OH of Ser81 and both the -OH and backbone amide of Ser86 each hydrogen bond a sulfonate. The second crystal packing neighbour contributes mostly cationic residues to the binding site. For example, one LysMe₂96 methyl group makes a cation-π interaction with a sclx₄ phenyl ring (the methyl is 4.1 Å from the centroid). The N^{η1} and N^{η2} of Arg21 make salt bridges with two lower rim hydroxyl groups. The C^ε of Tyr20 is 3.9 Å from two sclx₄ hydroxyls.

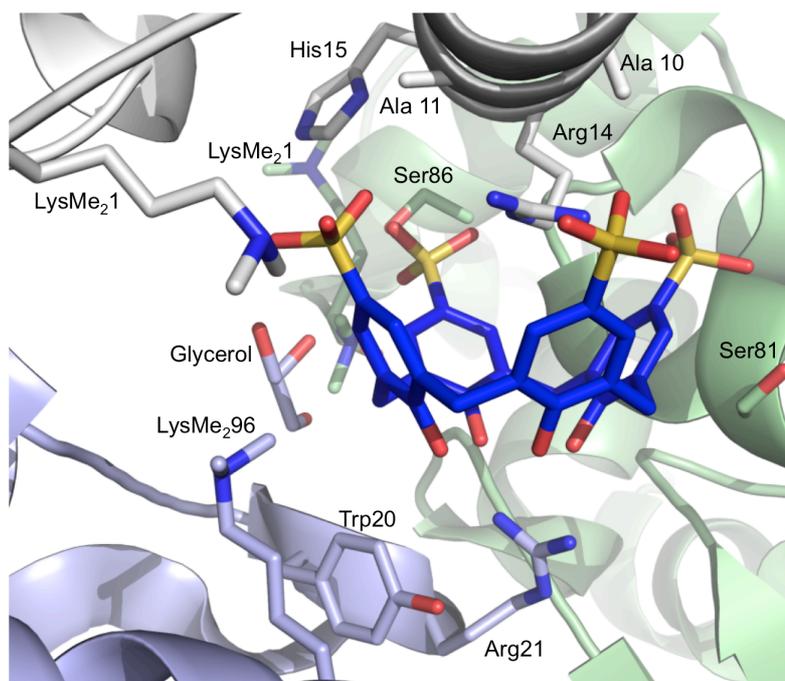


Figure 5: Two crystal packing neighbours are packed around the calixarene at the Arg14 binding site. The residues that interact with sclx₄ are shown as sticks. Glycerol was present at the site but made no direct contact with the calixarene. The crystal packing neighbours are coloured green and lilac.

Sclx₄ binds arginine side chains *endo* in both the native lysozyme and lysozyme-KMe₂ structures. However, different binding modes are observed for the arginine residues (Chapter 4). In the crystal structure of the native lysozyme-sclx₄ complex the C^γ to the guanidine of Arg128 are almost completely enclosed by the calixarene cavity, which has an elliptical cone conformation. The arginine C^δ to the guanidine is planar and sits into the long axis of the cone. The guanidine, pointing out of the plane of the sulfonates, makes salt bridge interactions and its upper face is solvent exposed. In the lysozyme-KMe₂ structure, sclx₄ binds Arg14 and in this instance the C^δ to the guanidine are engulfed. The calixarene has a symmetrical cone conformation. The guanidine is in the plane of the sulfonates and makes both direct and water-mediated salt bridge interactions. Interestingly, the methyl C atom of Ala10 makes van der Waals contact with the N^{η1} and N^{η2} of Arg14 and also sulfonate group. Ala11 makes the same contact with another sulfonate group. This di-alanine shield caps the Arg14-sclx₄ complex (Figure 5). This di-alanine shield is absent in the Arg128-sclx₄ binding site.

Surface grooves influence sclx₄ binding

Of the six LysMe₂ residues, the crystal structure shows that sclx₄ preferentially binds Lys116Me₂. An analysis of the lysozyme-KMe₂ structure provides some rationale for this observation. The side chains of LysMe₂1, LysMe₂13, LysMe₂96 and LysMe₂97 are

present in protein surface grooves (Figure 6) and are involved in intramolecular interactions. Closely similar conformations are observed for the lysine side chains in the native lysozyme structure. For example, the side chains of LysMe₂96 and LysMe₂97 expand across a continuous surface groove. LysMe₂96 is fully extended and the side chain N^ε makes a hydrogen bond with the backbone carbonyl of His15. The LysMe₂97 side chain is extended to a lesser degree than LysMe₂96. However, the side chain N^ε makes two salt bridge interactions with the O^{δ1} and O^{δ2} of Asp101. The side chain of LysMe₂33 points out from the protein surface. However, the side chains of Phe34, Phe38 and Trp123 conceal the alky portion of LysMe₂33 and only the dimethyl amino group is solvent exposed. Therefore, the steric accessibility of the LysMe₂33 side chain, similar to the other LysMe₂ side chains, is likely to be too low for calixarene binding. The side chain of LysMe₂116, which is flanked by Asn106, Arg112 and Gly117, projects out from the protein surface and is the most solvent-exposed. As a result, this residue is suitable for entrapment by sclx₄. The hydrogen bond from Lys116 to Asn106 in the native protein is absent in the dimethylated protein. At the LysMe₂116 binding site Asn106 hydrogen bonds two sclx₄ sulfonates. Importantly, the LysMe₂ intramolecular interactions in the lysozyme-KMe₂ structure and sclx₄ are echoed in the 1.8 Å crystal structure of lysozyme-KMe₂ (note that the side chain of Lys97 was not methylated).²⁸⁴

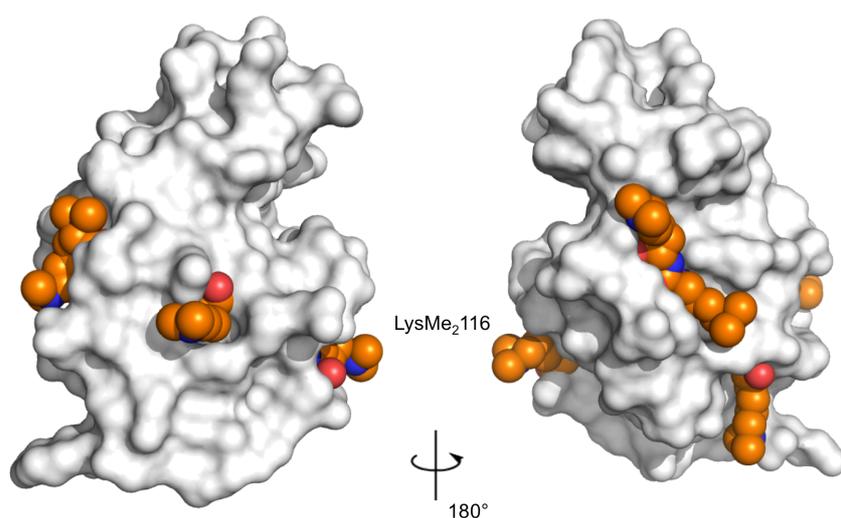


Figure 6: A surface representation of the lysozyme-KMe₂ crystal structure (coloured grey). The LysMe₂ side chains are present in surface cavities, with the exception of LysMe₂116 (labelled). The LysMe₂ residues are orange spheres and the sclx₄ are omitted.

Discussion

The observations made during the co-crystallization experiments hinted that the modified surface of lysozyme experienced different interactions with sclx₄ compared to the native protein. For example, sclx₄ caused the precipitation of lysozyme-KMe₂, followed by crystal growth. This rapid crystallization event was not observed in co-crystallization experiments with native lysozyme. Furthermore, high ionic strength solutions (containing 400 mM MgCl₂) did not hamper crystal growth. Therefore, charge-charge effects were not considered a major driving force for co-crystallization. Crystals, which grew in the absence of PEG, hinted at a new assembly, as the native lysozyme-sclx₄ complex involved PEG-sclx₄ complexes mediating protein-protein interactions.

The exciting feature of the lysozyme-KMe₂-calixarene crystal structure is the Lys116Me₂-sclx₄ binding site, which involves one lysine methyl group buried deep inside the calixarene cavity. Sclx₄ binds LysMe₂ with a higher affinity than arginine in phosphate buffered solution, therefore, LysMe₂ complexation was expected.⁹⁵ An analysis of the crystal structure indicates that Lys116Me₂ has the lowest steric hindrance out of all LysMe₂ residues and this is likely to favour calixarene binding. At the Arg14-sclx₄ binding site, interactions with the calixarene extend to His15 and LysMe₂1. These residues were involved in the calixarene binding sites in the native lysozyme structure, suggesting that sclx₄ preferentially binds this region.

The crystal structure of the lysozyme-KMe₂ complex with sclx₄ demonstrates how small molecule assisted self-assembly can be influenced by a subtle chemical modification of the protein surface. The methylation of lysine side chains results in the alteration of two chemical properties. In the first instance, there is an increased hydrophobic propensity attributed to the additional methyl groups, which promotes the tendency to make hydrophobic interactions with nonpolar groups. Secondly, lysine methylation polarizes the C^ε-N^ζ bond causing the *N*-methylamino group to assume an enhanced cationic character.^{58,285} The four-fold symmetry of sclx₄ was exploited to make four simultaneous cation- π interactions with one methyl group⁵⁸ (Figure 3). The cation- π component determines the affinity ($K_d = 10 \mu\text{M}$) and specificity LysMe₃ on the H3 peptide with the HP1 chromodomain.²⁸¹ Binding is concomitantly weaker for the LysMe₂ and LysMe peptides, while virtually no binding occurs with the unmethylated peptide. The peptide containing tert-butyl norleucine, instead of LysMe₃, has a significantly lower affinity ($K_d = 310 \mu\text{M}$), than its trimethylated lysine counterpart. This underlines the

preference of the HP1 chromodomain to make cation- π interactions with LysMe₃. It is likely that the energetic contributions that drive LysMe₂ complexation by sclx₄ are also dominated by cation- π interactions.

The quest to control chromatin assembly has resulted in numerous solutions that resemble the chromodomain cage motif.^{61,100,182} The described crystal structure shows a small molecule, reminiscent of the cage motifs, binding a protein containing dimethyllysine for the first time (Figure 3). The histone H3 tail peptide binds the aromatic cage of HP1 chromodomain via LysMe₂9, however, the interaction does not occur when Lys9 is unmethylated.¹⁷⁴ This is similar to unmethylated Lys116 that did not bind sclx₄ in the native lysozyme structure. The LysMe₂116-sclx₄ complex may benefit the design of receptors for histone tails.

Experimental

Crystallization

Hen white egg lysozyme, purchased from Sigma-Aldrich (62971 Fluka), was used without further purification. Formaldehyde, sodium borohydride and dimethylamine borane complex were obtained from Sigma-Aldrich. Reductive methylation of lysine residues was performed as per a previously published method.²⁸⁶ The degree of methylation was analysed by mass spectrometry, confirming that all sites, including the N-terminal amine, were dimethylated (Table 3). Trace amounts (less than 5 % abundance) of unmethylated or partially methylated species were also present.

Table 3: Mass spectrometry data for native lysozyme and the lysozyme-KMe₂.

Native lysozyme		Lysozyme-KMe ₂		Mass difference *	No. of sites
Measured mass (Da)	Abundance (%)	Measured mass (Da)	Abundance (%)		
14302.2	100	14498.0	100	195.8	7 sites dimethylated
		14411.9	< 5	109.7	4 sites dimethylated
		14302.3	< 5	0.1	No reaction

* Calculated by subtracting the measured mass of modified lysozyme from that of native lysozyme.

The hanging drop vapour diffusion method was used for crystallization trials at 20 °C. The crystallization conditions that yielded diffraction-quality crystals of cytochrome *c*-sclx₄ complex were implemented and the PEG range expanded. The 2.2 Å structure was determined from crystals grown in drops comprised of 1 µL volumes of lysozyme-KMe₂ (0.5 mM), sclx₄ (4.3 mM) and the reservoir solution. The reservoir solution contained 18 % PEG 8,000, 50 mM NaCl, 50 mM sodium cacodylate pH 6.3 and 100 mM MgCl₂. The 1.9 Å structure was determined from crystals grown in drops comprised of 1 µL volumes of lysozyme-KMe₂ (0.5 mM), sclx₄ (2.5 mM) and the reservoir solution. The reservoir solution contained 6 % PEG 8,000, 50 mM NaCl, 100 mM sodium cacodylate pH 6.1 and 100 mM MgCl₂. Control drops were prepared by substituting the addition of sclx₄ with 1 µL of water.

Data collection and X-ray structure determination

Crystals were transferred to the corresponding reservoir solution supplemented with 25 % glycerol and flash-frozen under a stream of nitrogen gas at 100 K. Diffraction data was collected using a 10 µm minibeam at the X10SA beam line of the Swiss Light Source, Villigen, at 100 K. Intensity data were collected (φ scans of 0.5° over 180°) to a resolution of 2.1 Å from a single crystal. Data processing and scaling were performed in Xia2²⁸⁷ using XDS,²⁸⁸ XSCALE and SCALA.¹⁷¹ Intensity data were also collected (φ scans of 1 ° over 180 °) to a resolution of 1.8 Å from a single crystal on beamline BM14 at the ESRF, Grenoble, France. Data processing and scaling were performed in MOSFLM and SCALA, respectively.^{170,171}

The data collection and refinement statistics are given in Table 1. The structures were determined by molecular replacement in PHASER and the refinement and manual rebuilding were performed in REFMAC5 (as implemented in CCP4) and COOT, respectively.^{174,274} Coordinates for sclx₄ were obtained from the PDB ID t3y and built into the model using COOT. Solvent molecules were placed automatically using ARP/wARP and refinement was continued until no features remained in the Fo – Fc difference maps.²⁷⁵ Molprobit was used to check the qualities of the structures prior to deposition in the Protein Data Bank [ID 4pru and 4noj].¹⁷⁶ Protein-protein and protein-ligand interfaces were analysed in COOT and PISA.²⁰⁵

Discussion

Overview

The co-crystallization capacity of several ligands (two calixarenes, a porphyrin and a tris-indole ligand) was tested. *p*-sulfonatocalix[4]arene (sclx₄) showed the greatest propensity to drive protein crystallization (Chapter 1). The ligands, which differ in size, structure, and chemical composition, are hypothesized to bind cytochrome *c* (cyt *c*) with different binding modes. From the analysis of the crystal structure of cyt *c* and sclx₄ in Chapter 2, the ability of the ligand to entrap a lysine residue in a pre-defined cavity was considered to be an important factor for co-crystallization.⁴⁸ Sclx₄ has a preformed cavity by virtue of its fixed cone conformation, however, the sulfonatophenyl-porphyrin, which is also decorated with an acidic periphery, lacks this feature. The tris-indole and sclx₈ can adopt different conformations in both the free and bound states and some of these structures contain cavities.^{82,168} The structural rearrangement of the tris-indole and calix[8]arene as well as the local rearrangement of flexible residues near the binding site(s) is expected to hamper co-crystallization as complexation is likely to have a high entropic cost. Other factors including the difference in the accessible surface areas are expected to contribute to the tendency of the ligands to co-crystallise cyt *c*.

The two proteins crystallized with sclx₄ have similar pIs and molecular weights.^{130,138} The types of cationic residues and their quantities differentiates the proteins. Arginine accounts for 61% of lysozyme's basic residues and lysine accounts for 70 % of cyt *c*'s basic residues. Typically, sclx₄ binds the cationic residues that are present in the highest concentration. The small surface area and the amino acid composition is the reason for the high charge density of cyt *c* and lysozyme. This is expected to facilitate molecular recognition, as it is easier for the ligand to find arginine or lysine residues. This was demonstrated for the bisphosphonate polymers that bind cyt *c* more tightly than the trypsin, which is also a cationic and lysine-rich protein but is 22 kDa in size.¹²²

In the first model system explored, NMR and X-ray crystallographic data were in good agreement for the complexation of cyt *c* by sclx₄ (Chapter 2). The NMR data point towards a dynamic ensemble, which involves sclx₄ binding the protein surface through multiple binding sites centred on lysines. Hill coefficients suggest cooperative binding between cyt *c* and sclx₄. Sclx₄ alters the solubility of cyt *c*, thus facilitating protein crystallization from PEG only solutions. This is similar to the carboxylic acid derivative of calix[6]arene, which transferred cyt *c* into organic phase.¹⁰³ Sclx₄ promoted protein

self-assembly by grabbing lysine side chains. This phenomenon is comparable to the rapid crystal growth of RbmA by the 18-crown-6, which also involved the incorporation a lysine side chain in the central ligand pore.⁴⁷ The R13E mutant of cyt *c* yielded large co-crystals with sclx₄, visible to the naked eye. The crystal structure of the complex shows that the calixarene interacts with the mutant via the same binding sites as the wild type complex. The R13E mutation, which is located at a solvent channel, may be conducive to self-assembly.

The cyt *c* R13E-sclx₄ crystals grown on SAM-modified Au electrodes were electrochemically characterised by using cyclic voltammetry (Chapter 3). The highest concentrations of electro-active cyt *c* were detected at the lowest scan rates. Similar results were obtained for multilayer cyt *c* assemblies.^{22,218} The analysis of the cyt *c* R13E-sclx₄ crystal structure showed that the closest haem groups were 23.7 Å apart suggesting that electron transfer required multiple steps and involved intervening amino acid side chains.²³¹ Protein crystallization on modified Au electrodes offers an alternative strategy for generating high protein concentrations and the calixarene is useful for promoting protein self-assembly in a controlled manner.

Co-crystallization experiments showed that sclx₄ altered the solubility of lysozyme in a similar manner to cyt *c*, by promoting crystal growth from PEG containing solutions. Although sclx₄ readily caused lysozyme precipitation, these precipitates did not yield crystals when the solutions lacked PEG. X-ray crystallography was used to investigate the lysozyme-sclx₄ complex (Chapter 4) and showed lysozyme assembled as a tetramer with four sclx₄ binding sites. A fifth site is also observed in one chain. At two binding sites the side chain of the C-terminal arginine is almost completely engulfed by the calixarene cavity. Two sclx₄ entrap a PEG-Mg²⁺ complex. An expansion of the asymmetric unit shows that the sclx₄ binding arginine is packed up against one of these two sclx₄-PEG-Mg²⁺ complexes. These close-packed calixarenes plug either side of the tetrameric assembly. Neighbouring tetramers are positioned such that two calixarene-PEG-Mg²⁺ complexes oppose each other. The short PEG-PEG distance (4.4 Å) suggests that one PEG fragment interlaces two tetrameric assemblies (Chapter 4, Figure 3). The fifth sclx₄ binds near the known lysozyme active site and a PEG fragment and Mg²⁺ are bound *endo*. Similar to the crystal structure of the cyt *c*-sclx₄ complex, the calixarene is nestled at protein-protein interactions.

Lysozyme possessing dimethyl lysine residues (lysozyme-KMe₂) was used as the final model system (Chapter 5). Lysozyme-KMe₂ crystallized rapidly (within minutes) in

the presence of sclx₄ and PEG was not a prerequisite for crystal growth. The crystal structure shows two lysozyme molecules and four calixarenes, which bind Lys116Me₂ and Arg14 in both chains. One methyl group of Lys116Me₂ is buried deep within the calixarene cavity making C-H $\cdots\pi$ interactions with the phenyl rings of sclx₄.²⁸⁹ An analysis of the crystal structure suggests that Lys116Me₂ is the most sterically accessible LysMe₂ residue. This data may benefit the design of receptors for histone tails and/or inhibitors of chromatin remodelling processes.

Sclx₄ interactions

Overall, the crystal structures of the protein-sclx₄ complexes reveal three types of binding sites that can be distinguished by the *endo* residue (lysine, LysMe₂ or arginine). The noncovalent interactions and the interface area vary at each binding site (Table 1). For example, the interface area of the sclx₄-Arg14 complex in the lysozyme-KMe₂ structure is 40 Å greater than that of the Arg128 complex in the native structure. A greater portion of the Arg14 side chain is engulfed by the calixarene cavity and the sclx₄ makes more noncovalent contacts with protein surface (Chapter 5). The interactions between the calixarene and the *endo* bound residue account for > 50 % of the protein surface area that becomes inaccessible in the complex, with the exception of the *cyt c* Lys89 binding site (Table 1).

In general, more salt bridges are involved in the arginine-calixarene complexes than those containing lysine. This can be rationalised by the planar geometry and diffuse charge distribution of the guanidine group. The arginine N^ε, N^{η1} and N^{η2} each make salt bridges with three calixarene sulfonates and this is similar to the sclx₄-arginine amino acid crystal structure.⁹¹ The charge-charge contribution to arginine complexation by sclx₄ was suggested by using NMR as complex formation was inhibited in > 100 mM salt.⁹² The lysine side chains and dimethylated side chains make salt bridges and cation- π interactions with the sclx₄. Interestingly, a cation- π interaction occurs between the lysine C^ε and sclx₄ in the *cyt c* structure and a similar interaction between the lysine C^ε and tryptophan side chain favours β -hairpin folding.²⁹⁰ The NMR titrations carried out in the presence of 50 mM NaCl showed a reduction in the affinity of sclx₄ for *cyt c*, which suggests that charge-charge interactions drive ligand binding in these complexes.⁴⁸ Interestingly, for lysozyme-KMe₂ the rapid crystal growth observed in drops containing

up to 400 mM MgCl₂ (in addition to 50 mM NaCl and 50 mM (CH₃)₂AsO₂Na, pH 6), suggest that the charge-charge contribution to protein-calixarene complexation is low. Isothermal calorimetry experiments show that the increased methylation of lysine is accompanied mostly by favourable enthalpy changes on sclx₄ complexation.⁹⁵

Table 1. Interface size and polar interactions in *endo* sclx₄ complexes.

Complex	Binding Site	Interface area (Å ²)		Hydrogen bonds [§]	Salt bridges [¶]		Cation-π [‡]
		Protein [†]	Residue ^{††}		Direct	Water	
Cyt <i>c</i>	Lys89	320	150	1	2	2	3
	Lys4	230	150	1	1	3	3
	Lys22	210	150	-	1	-	3
Lysozyme	Arg128	230	170	1	4	2 [*]	-
Lysozyme-KMe ₂	Arg14	270	150	-	4	4 ^{**}	-
	LysMe ₂ 116	230	140	-	2	2	4
Lysine amino acid ⁹⁰	Lys	-	170	-	3 ^{¶¶}	1	-
Arginine amino acid ⁹¹	Arg	-	160	-	6 ^{¶¶}	1	-

[†]Surface area (reported to the nearest 10 Å²) of the protein that becomes inaccessible in the complex.

^{††}Surface area of the residue or amino acid that becomes inaccessible in the complex. [§]Number of hydrogen bonds between sclx₄ and the residue. [¶]Defined as ‘Direct’ or ‘Water’ mediated, in which a H₂O bridges the oppositely charged groups. [‡]Total number of sclx₄ phenyl rings that make cation-π (Lysine N^ε to all C atoms in the phenyl ring < 4.7 Å). ^{*}Arg128 makes no water mediated salt bridges in chain B and ^{**}Arg14 makes 2 in chain B. ^{¶¶}One salt bridge is made with the backbone amine of the amino acid.

Prominent features of the sclx₄ binding sites

Several features of the calixarene binding sites are common to the crystal structures and may benefit the prediction of sclx₄ interaction sites in other proteins. Firstly, cationic side chains are observed inside the ligand cavity and frequently outside the cavity. For instance, sclx₄ binds Lys89 in the cyt *c* chain A and makes a salt bridge with the side chain of neighbouring residue Lys5. Secondly, the side chains must be sterically accessible for entrapment and this is likely to be the reason for the sclx₄-LysMe₂116 complex. Thirdly, knowledge of anion binding sites in the protein can be used to predict where sclx₄ is likely to bind. For example, Arg14 binding involves the insertion of a sclx₄ sulfonate into an adjacent anion binding motif, similar to the Arg128-sclx₄ complex. Significantly, this binding site coincides with a known sulfate binding pocket.²⁵⁹ Finally, in all the crystal structures, the C-terminus features prominently in sclx₄ binding. The N-terminus is also a noticeable feature of sclx₄ binding in the lysozyme and lysozyme-KMe₂

structures. This is similar to cucurbituril, which binds to the N-terminal phenylalanine of insulin.⁴⁴

An analysis of the crystal structures shows that *sclx*₄ binds the same sites in the chains A and B of both lysozyme and lysozyme-KMe₂. Note that this is discounting the 5th binding site near the active site in lysozyme-*sclx*₄ structure. Interestingly, there are three unique binding sites in the asymmetric unit of the *cyt c-sclx*₄ structure. The superposition of chain B onto chain A shows that the calixarenes that bind at A.Lys89 and B.Lys4 in close proximity (Figure 1). Unfavourable charge-charge interactions between the adjacent sulfonates (3.1 Å apart) might hamper simultaneous binding. Lys5 interacts with the calixarenes at both A.Lys89 and B.Lys4. In chain A two Lys5 conformers make salt bridges with a *sclx*₄ sulfonate and one conformer makes an additional salt bridge with Asp90. In chain B Lys5 has one conformation, the interaction with Asp90 is lost and the lysine makes a salt bridge with a sulfonate. It is possible that the side chain rearrangement required for both binding sites to be occupied simultaneously hinders concurrent binding. This may provide a molecular basis for cooperative binding observed in solution by using NMR. The Lys22 site can be occupied simultaneously with Lys89 or Lys4. However, A.Lys22 from crystal packing neighbours binds *sclx*₄-Lys89 complex *exo* and this may be the reason that *sclx*₄ does not bind Lys89 and Lys22 concomitantly.

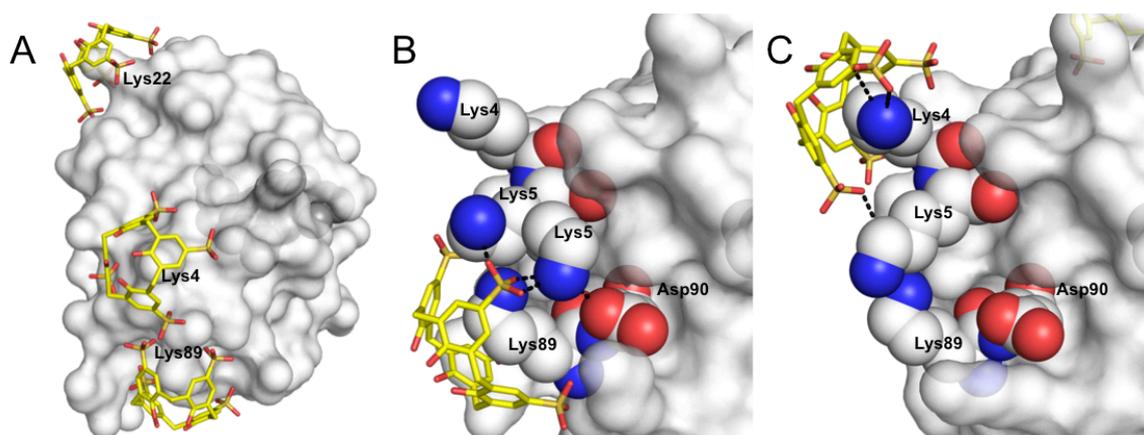


Figure 1. (A) The superposition of *cyt c* chain B onto chain A shows the proximity of the *sclx*₄ binding sites. (B) Lys5 has two conformations at the Lys89 binding site and makes salt bridges with the calixarene and Asp90. (C) One Lys5 conformation is observed when *sclx*₄ binds Lys4. Lys5 makes one salt bridge with the calixarene via its C^ε. *Cyt c* is represented as transparent grey surfaces. The calixarenes are yellow sticks. The discussed residues are shown as spheres and the salt bridges are dashed lines.

The protein-calixarene assemblies share several key characteristics. The most prominent feature is that sclx₄ is always present at protein-protein interfaces. For example, in the *cyt c* structure the calixarene that binds Lys4 is surrounded by three additional close-packed protein chains from crystal packing neighbours. Although the global structures of the sclx₄-mediated lysozyme and lysozyme-KMe₂ complexes are significantly different, several cationic residues were observed in both structures making noncovalent interactions with the calixarenes. These are Lys1, Arg14, His15, Arg61, Arg112, Lys116 and Arg128. The four-fold symmetry of the calixarene is useful for protein packing and the lysozyme-sclx₄ structure involves pairs of calixarenes at the tetramer interfaces. The complex is reminiscent of the structure proposed for the β -tryptase-calix[8]arene complex for which a cationic calix[8]arene binds the central pore of β -tryptase tetramer, inhibiting enzyme activity.²⁹¹

Sclx₄ as a low and high affinity protein surface binder

Protein surface interactions involving sulfated and sulfonated molecules are important for many biological applications and their affinities for protein surfaces vary. Sclx₄ weakly binds *cyt c*, which involves charge-charge interactions. Similar to the *cyt c*-sclx₄ complexes, the anionic polysaccharide heparin sulfate, which is prevalent in the regulation of numerous physiological processes,²⁹² weakly binds cationic protein surfaces ($K_d = 0.5$ mM).²⁰¹ The polyanionic suramin is a competitive inhibitor of heparin binding and both protein-ligand complexes involve interactions between the ligand sulfonates and lysine and arginine side chains.^{201,293} As sclx₄ is known to bind cationic surfaces it might have a potential application for systems involving heparin binding proteins. The analysis of the NMR data for *cyt c* and sclx₄ have inspired a new concept called protein camouflage, whereby transiently interacting small molecules associate to and hide the protein surface (Figure 2). This is feasible as sclx₄ has a low affinity for several binding sites on *cyt c*. Calixarenes camouflaging the cationic surface patches of *cyt c* are expected to hamper *cyt c* binding by natural protein partners, such as cytochrome *c* peroxidase. Electrostatic steering is conducive to *cyt c*-cytochrome *c* peroxidase complexation¹²⁸ and the altered surface charge of *cyt c*, as a result of sclx₄ binding, is expected to reduce the tendency of these proteins to associate.

Sulfonated ligands, such as heparin and phthalocyanine tetrasulfonate can affect the self-assembling ability of proteins.²⁹⁴ Sclx₄ can assist this self-assembly feature in specific model systems, whereby the protein surface properties influenced self-assembly and co-crystallization. The rapid crystal growth of lysozyme-KMe₂ by sclx₄, which occurred within minutes, suggests that the calixarene has a greater affinity for this modified protein over native lysozyme. In low salt solutions lysozyme-KMe₂ crystals grew with 20 μM protein when the sclx₄ concentration was as low as 1 μM (Chapter 5, Table 2). This suggests that the calixarene has an affinity for lysozyme-KMe₂ on the μM scale. Sclx₄ has binds methyllysine amino acids with high affinity⁹⁵ and recently its derivatives have been shown to disrupt H3K9Me₃ binding by the chromodomain of the PHD2 finger.²⁹⁵ We show structurally that the interactions between LysMe₂ and the calixarene are similar to those involving the natural aromatic cage motif (Chapter 5).

Protein-calixarene interactions in model systems are important for elucidating recognition events on a molecular level. These might inspire applications in physiologically relevant systems. Calixarenes can alter protein properties, including solubility, aggregation and self-assembly through surface recognition. We show that the simple and symmetrical sclx₄ is both a high and low affinity protein surface binder depending on the target surface and that it assists protein self-assembly. The solved crystal structures may help guide the next generation of anionic scaffolds that target specific protein surfaces and/or specific residues. One potential route could involve promoting the assembly of heterodimers by tuning the upper and lower calixarene rims to each complement a different target surface. This larger recognition capacity could lead to a range of possibilities for designing and generating complex protein assemblies mediated by calixarenes.

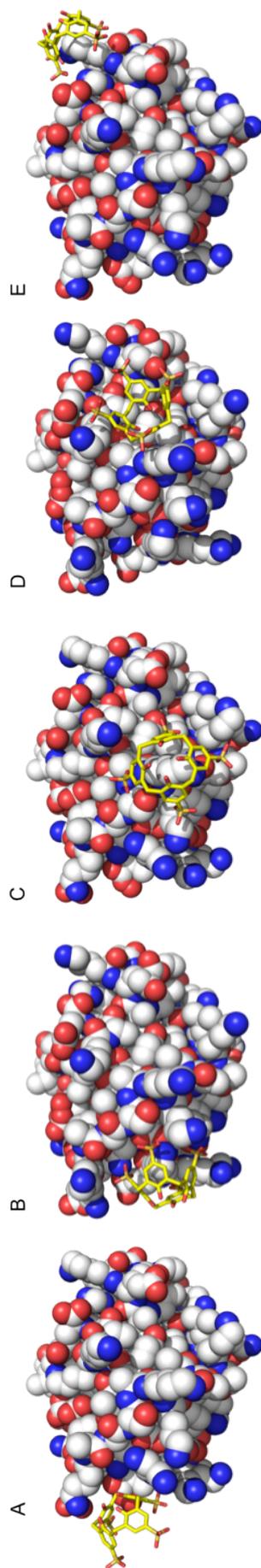


Figure 1. Roaming *sclx*, camouflages the surface of *cyt c*. A–E, From the crystal structure complexes B2 (a), A.Lys89 (b), B.Lys4 (c), A1 (d) and B.Lys22 (e), respectively (Chapter 2, Table 2). The sequence of images exposes the contiguous surface of *cyt c* that is explored by the calixarene. A similar binding patch was observed in solution by using NMR spectroscopy (Chapter 2, Figure 2). *Cyt c* is represented as spheres and the ligand as sticks, with carbons coloured yellow for contrast. Water molecules are omitted for clarity.

Bibliography

- 1 Lehn, J.-M. Supramolecular chemistry: Receptors, catalysts, and carriers. *Science* **1985**, 227, 849.
- 2 Yin, H. & Hamilton, A. D. Strategies for targeting protein-protein interactions with synthetic agents. *Angew. Chem. Int. Ed.* **2005**, 44, 4130.
- 3 Fletcher, S. & Hamilton, A. Targeting protein-protein interactions by rational design: mimicry of protein surfaces. *J. R. Soc. Interface* **2006**, 3, 215.
- 4 Schneider, H. J. Binding mechanisms in supramolecular complexes. *Angew. Chem. Int. Ed.* **2009**, 48, 3924.
- 5 Richter, O. M. H. & Ludwig, B. Cytochrome *c* oxidase - structure, function, and physiology of a redox-driven molecular machine. *Rev. Physiol. Biochem. Pharmacol.* **2003**, 147, 47.
- 6 Gschwind, A., Fischer, O. M. & Ullrich, A. Timeline - The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat. Rev. Cancer* **2004**, 4, 361.
- 7 Chen, L. H., Li, Z. Y., Zwolinska, A. K., Smith, M. A., Cross, B., Koomen, J., Yuan, Z. M., Jenuwein, T., Marine, J. C., Wright, K. L. & Chen, J. D. MDM2 recruitment of lysine methyltransferases regulates p53 transcriptional output. *EMBO J.* **2010**, 29, 2538.
- 8 Janin, J. & Chothia, C. The structure of protein-protein recognition sites. *J. Biol. Chem.* **1990**, 265, 16027.
- 9 Pelletier, H. & Kraut, J. Crystal structure of a complex between electron transfer partners, cytochrome *c* peroxidase and cytochrome *c*. *Science* **1992**, 258, 1748.
- 10 Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 13.
- 11 Nooren, I. M. A. & Thornton, J. M. Structural characterisation and functional significance of transient protein-protein interactions. *J. Mol. Biol.* **2003**, 325, 991.
- 12 Cho, K. I., Lee, K., Lee, K. H., Kim, D. & Lee, D. Specificity of molecular interactions in transient protein-protein interaction interfaces. *Proteins* **2006**, 65, 593.
- 13 Sattler, M., Liang, H., Nettlesheim, D., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H. S., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B. & Fesik, S. W. Structure of Bcl-xL-Bak Peptide Complex: Recognition between regulators of apoptosis. *Science* **1997**, 275, 983.

- 14 Crowley, P. B. & Carrondo, M. A. The architecture of the binding site in redox protein complexes: Implications for fast dissociation. *Proteins* **2004**, 55, 603.
- 15 Sato, K., Crowley, P. B. & Dennison, C. Transient homodimer interactions studied using the electron self-exchange reaction. *J. Biol. Chem.* **2005**, 280, 19281.
- 16 Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **1999**, 285, 2177.
- 17 Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. & Pavletich, N. P. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **1995**, 376, 313.
- 18 Wysoczanski, P., Mart, R. J., Loveridge, E. J., Williams, C., Whittaker, S. B. M., Crump, M. P. & Allemann, R. K. NMR solution structure of a photoswitchable apoptosis activating Bak peptide bound to Bcl-xL. *J. Am. Chem. Soc.* **2012**, 134, 7644.
- 19 Cochran, A. G. Antagonists of protein-protein interactions. *Chemistry & Biology* **2000**, 7, R85.
- 20 He, M. M., Smith, A. S., Oslob, J. D., Flanagan, W. M., Braisted, A. C., Whitty, A., Cancilla, M. T., Wang, J., Lugovskoy, A. A., Yoburn, J. C., Fung, A. D., Farrington, G., Eldredge, J. K., Day, E. S., Cruz, L. A., Cachero, T. G., Miller, S. K., Friedman, J. E., Choong, I. C. & Cunningham, B. C. Small-molecule inhibition of TNF- α . *Science* **2005**, 310, 1022.
- 21 Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, 450, 1001.
- 22 Beissenhirtz, M. K., Scheller, F. W., Stöcklein, W. F. M., Kurth, D. G., Möhwald, H. & Lisdat, F. Electroactive Cytochrome *c* multilayers within a polyelectrolyte assembly. *Angew. Chem. Int. Ed.* **2004**, 43, 4357.
- 23 Dronov, R., Kurth, D. G., Moehwald, H., Spricigo, R., Leimkuehler, S., Wollenberger, U., Rajagopalan, K. V., Scheller, F. W. & Lisdat, F. Layer-by-layer arrangement by protein-protein interaction of sulfite oxidase and cytochrome *c* catalyzing oxidation of sulfite. *J. Am. Chem. Soc.* **2008**, 130, 1122.
- 24 Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D. & Tirrell, D. A. Reversible hydrogels from self-assembling artificial proteins. *Science* **1998**, 281, 389.
- 25 Rajagopal, K. & Schneider, J. P. Self-assembling peptides and proteins for nanotechnological applications. *Curr. Opin. Struct. Biol.* **2004**, 14, 480.

- 26 Grueninger, D., Treiber, N., Ziegler, M., Koetter, J., Schulze, M. & Schulz, G. Designed protein-protein association. *Science* **2008**, 319, 206.
- 27 Prasad Bahadur, R., Chakrabarti, P., Rodier, F. & Janin, J. A Dissection of specific and non-specific protein-protein interfaces. *J. Mol. Biol.* **2004**, 336, 943.
- 28 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.
- 29 Sinclair, J. C., Davies, K. M., Venien-Bryan, C. & Noble, M. E. M. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* **2011**, 6, 558.
- 30 Lai, Y.-T., Cascio, D. & Yeates, T. O. Structure of a 16-nm cage designed by using protein oligomers. *Science* **2012**, 336, 1129.
- 31 Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **2001**, 293, 1074.
- 32 Ma, J. C. & Dougherty, D. A. The cation- π interaction. *Chem. Rev.* **1997**, 97, 1303.
- 33 Schubot, F. D. & Waugh, D. S. A pivotal role for reductive methylation in the de novo crystallization of a ternary complex composed of *Yersinia pestis* virulence factors YopN, SycN and YscB. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2004**, 60, 1981.
- 34 Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, Joanne E., Owens, R. J., Stuart, David I. & Grimes, J. M. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **2006**, 14, 1617.
- 35 Zhou, N. E., Kay, C. M. & Hodges, R. S. Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded α -helical coiled-coil. *Biochemistry* **1993**, 32, 3178.
- 36 Heinz, D. W. & Matthews, B. W. Rapid crystallization of T4 lysozyme by intermolecular disulfide cross-linking. *Protein Eng.* **1994**, 7, 301.
- 37 Banatao, D. R., Cascio, D., Crowley, C. S., Fleissner, M. R., Tienson, H. L. & Yeates, T. O. An approach to crystallizing proteins by synthetic symmetrization. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 16230.
- 38 Forse, G. J., Ram, N., Banatao, D. R., Cascio, D., Sawaya, M. R., Klock, H. E., Lesley, S. A. & Yeates, T. O. Synthetic symmetrization in the crystallization and structure determination of CelA from *Thermotoga maritima*. *Protein Sci.* **2011**, 20, 168.

- 39 Reitz, S., Cebi, M., Reiß, P., Studnik, G., Linne, U., Koert, U. & Essen, L.-O. On the function and structure of synthetically modified porins. *Angew. Chem. Int. Ed.* **2009**, 48, 4853.
- 40 Morgan, H. P., McNae, I. W., Hsin, K.-Y., Michels, P. A. M., Fothergill-Gilmore, L. A. & Walkinshaw, M. D. An improved strategy for the crystallization of *Leishmania mexicana* pyruvate kinase. *Acta Crystallogr. Sect. F.* **2010**, 66, 215.
- 41 Buratto, J., Colombo, C., Stupfel, M., Dawson, S. J., Dolain, C., Langlois d'Estaintot, B., 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.* **2013**, 53, 883.
- 42 Aaron, J. A., Chambers, J. M., Jude, K. M., Di Costanzo, L., Dmochowski, I. J. & Christianson, D. W. Structure of a ¹²⁹Xe-cryptophane biosensor complexed with human carbonic anhydrase II. *J. Am. Chem. Soc.* **2008**, 130, 6942.
- 43 Bier, D., Rose, R., Bravo-Rodriguez, K., Bartel, M., Ramirez-Anguila, J. M., Dutt, S., Wilch, C., Klarner, F.-G., Sanchez-Garcia, E., Schrader, T. & Ottmann, C. Molecular tweezers modulate 14-3-3 protein-protein interactions. *Nature Chem.* **2013**, 5, 234.
- 44 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.
- 45 Hunter, T. M., McNae, I. W., Liang, X., Bella, J., Parsons, S., Walkinshaw, M. D. & Sadler, P. J. Protein recognition of macrocycles: Binding of anti-HIV metalocyclams to lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 2288.
- 46 Hunter, T. M., McNae, I. W., Simpson, D. P., Smith, A. M., Moggach, S., White, F., Walkinshaw, M. D., Parsons, S. & Sadler, P. J. Configurations of nickel-cyclam antiviral complexes and protein recognition. *Chem. Eur. J.* **2007**, 13, 40.
- 47 Maestre-Reyna, M., Wu, W.-J. & Wang, A. Structural Insights into RbmA, a biofilm scaffolding protein of *V. Cholerae*. *PLoS ONE* **2013**, 8.
- 48 McGovern, R. E., Fernandes, H., Khan, A. R., Power, N. P. & Crowley, P. B. Protein camouflage in cytochrome *c*-calixarene complexes. *Nature Chem.* **2012**, 4, 527.
- 49 Freeman, J. O., Lee, W. C., Murphy, M. E. P. & Sherman, J. C. X-ray crystal analysis of a TASP: Structural insights of a cavitein dimer. *J. Am. Chem. Soc.* **2009**, 131, 7421.

- 50 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.
- 51 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**, 60, 281.
- 52 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.
- 53 Jiao, L., Ouyang, S., Liang, M., Niu, F., Shaw, N., Wu, W., Ding, W., Jin, C., Peng, Y., Zhu, Y., Zhang, F., Wang, T., Li, C., Zuo, X., Luan, C.-H., Li, D. & Liu, Z.-J. Structure of severe fever with thrombocytopenia syndrome virus nucleocapsid protein in complex with suramin reveals therapeutic potential. *Journal of Virology* **2013**, 87, 6829.
- 54 Ernst, J. T., Becerril, J., Park, H. S., Yin, H. & Hamilton, A. D. Design and application of an α -helix-mimetic scaffold based on an oligoamide-foldamer strategy: Antagonism of the Bak BH3/Bcl-xL complex. *Angew. Chem. Int. Ed.* **2003**, 42, 535.
- 55 Supuran, C. T. & Scozzafava, A. Carbonic anhydrases as targets for medicinal chemistry. *Biorg. Med. Chem.* **2007**, 15, 4336.
- 56 Baud, F. & Karlin, S. Measures of residue density in protein structures. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 12494.
- 57 Jones, S., Marin, A. & Thornton, J. M. Protein domain interfaces: Characterization and comparison with oligomeric protein interfaces. *Protein Eng.* **2000**, 13, 77.
- 58 Gallivan, J. P. & Dougherty, D. A. Cation- π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 9459.
- 59 Crowley, P. B. & Golovin, A. Cation- π interactions in protein–protein interfaces. *Proteins* **2005**, 59, 231.
- 60 Fokkens, M., Schrader, T. & Klärner, F.-G. A Molecular tweezer for lysine and arginine. *J. Am. Chem. Soc.* **2005**, 127, 14415.

-
- 61 Gamal-Eldin, M. A. & Macartney, D. H. Selective molecular recognition of methylated lysines and arginines by cucurbit[6]uril and cucurbit[7]uril in aqueous solution. *Org. Biomol. Chem.* **2013**, 11, 488.
- 62 Dang, D. T., Nguyen, H. D., Merckx, M. & Brunsveld, L. Supramolecular control of enzyme activity through cucurbit[8]uril-mediated dimerization. *Angew. Chem.* **2013**, 125, 2987.
- 63 Liang, X. & Sadler, P. J. Cyclam complexes and their applications in medicine. *Chem. Soc. Rev.* **2004**, 33, 246.
- 64 De Clercq, E., Yamamoto, N., Pauwels, R., Baba, M., Schols, D., Nakashima, H., Balzarini, J., Debyser, Z., Murrer, B. A., Schwartz, D., Thornton, D., Bridger, G., Fricker, S., Henson, G., Abrams, M. & Picker, D. Potent and selective inhibition of human immunodeficiency virus (HIV)-1 and HIV-2 replication by a class of bicyclams interacting with a viral uncoating event. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, 89, 5286.
- 65 Julian, R. R. & Beauchamp, J. L. Selective molecular recognition of arginine by anionic salt bridge formation with bis-phosphate crown ethers: Implications for gas phase peptide acidity from adduct dissociation. *J. Am. Soc. Mass. Spectrom.* **2004**, 15, 616.
- 66 Chen, Y. & Rodgers, M. T. Structural and energetic effects in the molecular recognition of protonated peptidomimetic bases by 18-crown-6. *J. Am. Chem. Soc.* **2012**, 134, 2313.
- 67 Odell, B. & Earlam, G. Dissolution of proteins in organic solvents using macrocyclic polyethers: association constants of a cytochrome *c*-[1,2-¹⁴C₂]-18-crown-6 complex in methanol. *J. Chem. Soc., Chem. Commun.* **1985**, 359.
- 68 Julian, R. R. & Beauchamp, J. L. Site specific sequestering and stabilization of charge in peptides by supramolecular adduct formation with 18-crown-6 ether by way of electrospray ionization. *Int. J. Mass spectrom.* **2001**, 210–211, 613.
- 69 Paul, D., Suzumura, A., Sugimoto, H., Teraoka, J., Shinoda, S. & Tsukube, H. Chemical activation of cytochrome *c* proteins via crown ether complexation: Cold-active synzymes for enantiomer-selective sulfoxide oxidation in methanol. *J. Am. Chem. Soc.* **2003**, 125, 11478.
- 70 Oshima, T., Suetsugu, A. & Baba, Y. Extraction and separation of a lysine-rich protein by formation of supramolecule between crown ether and protein in aqueous two-phase system. *Anal. Chim. Acta* **2010**, 674, 211.

- 71 Gutsche, C. D. Calixarenes. *Acc. Chem. Res.* **1983**, 16, 161.
- 72 Shinkai, S., Araki, K., Tsubaki, T., Arimura, T. & Manabe, O. New syntheses of calixarene-*p*-sulphonates and *p*-nitrocalixarenes. *J. Chem. Soc., Perkin Trans. 1* **1987**, 2297.
- 73 Shimizu, K. & Rebek, J. J. Synthesis and assembly of self-complementary calix[4]arenes. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, 92, 12403.
- 74 Corbellini, F., Fiammengo, R., Timmerman, P., Crego-Calama, M., Versluis, K., Heck, A. J. R., Luyten, I. & Reinhoudt, D. N. Guest encapsulation and self-assembly of molecular capsules in polar solvents via multiple ionic interactions. *J. Am. Chem. Soc.* **2002**, 124, 6569.
- 75 Kennedy, S., Dodgson, I. E., Beavers, C. M., Teat, S. J. & Dalgarno, S. J. Pyridine directed assembly of di-*o*-alkyl-tris-*p*-carboxylatocalix[4]arenes. *Cryst. Growth Des.* **2011**, 12, 688.
- 76 Daze, K. D., Ma, M. C. F., Pineux, F. & Hof, F. Synthesis of new trisulfonated calix[4]arenes functionalized at the upper rim, and their complexation with the trimethyllysine epigenetic mark. *Org. Lett.* **2012**, 14, 1512.
- 77 Shinkai, S., Araki, K., Matsuda, T., Nishiyama, N., Ikeda, H., Takasu, I. & Iwamoto, M. NMR and crystallographic studies of a *p*-sulfonatocalix[4]arene-guest complex. *J. Am. Chem. Soc.* **1990**, 112, 9053.
- 78 Danylyuk, O., Ghera, B. B., Lazar, A. N., Coleman, A. W. & Suwinska, K. The solid-state structures of *para*-sulfonatocalix[4]arene with the biologically active oligoammonium cations of norspermidine and triethyltetramine. *J. Mol. Struct.* **2008**, 891, 443.
- 79 Guo, D.-S., Zhang, H.-Q., Ding, F. & Liu, Y. Thermodynamic origins of selective binding affinity between *p*-sulfonatocalix[4,5]arenes with biguanidiniums. *Org. Biomol. Chem.* **2012**, 10, 1527.
- 80 Lazar, A. N., Danylyuk, O., Suwinska, K., Kassab, R. & Coleman, A. W. Stepped layers in the complexes of *para*-sulfonatocalix[6]arene with dimethylammonium and bis-6-aminohexylammonium cations. *New J. Chem.* **2008**, 32, 2116.
- 81 Atwood, J. L., Dalgarno, S. J., Hardie, M. J. & Raston, C. L. Selective single crystal complexation of L- or D-leucine by *p*-sulfonatocalix[6]arene. *Chem. Commun.* **2005**, 0, 337.

-
- 82 Danylyuk, O., Perret, F., Coleman, A. W. & Suwinska, K. The solid-state complex of *para*-sulphonato-calix[8]arene anion with dimethylammonium cations. *Open Crystallogr. J.* **2008**, 1, 18.
- 83 Perret, F., Bonnard, V., Danylyuk, O., Suwinska, K. & Coleman, A. W. Conformational extremes in the supramolecular assemblies of *para*-sulfonato-calix[8]arene. *New J. Chem.* **2006**, 30, 4.
- 84 Danylyuk, O. & Suwinska, K. Solid-state interactions of calixarenes with biorelevant molecules. *Chem. Commun.* **2009**, 0, 5799.
- 85 Coleman, A. W., Bott, S. G., Morley, S. D., Means, C. M., Robinson, K. D., Zhang, H. & Atwood, J. L. Novel layer structure of sodium calix[4]arenesulfonate complexes—a class of organic clay mimics? *Angew. Chem. Int. Ed.* **1988**, 27, 1361.
- 86 Orr, G. W., Barbour, L. J. & Atwood, J. L. Controlling molecular self-organization: Formation of nanometer-scale spheres and tubules. *Science* **1999**, 285, 1049.
- 87 Lehn, J.-M., Meric, R., Vigneron, J.-P., Cesario, M., Guilhem, J., Pascard, C., Asfari, Z. & Vicens, J. Binding of acetylcholine and other quaternary ammonium cations by sulfonated calixarenes. Crystal structure of a [choline-tetrasulfonated calix[4]arene] complex. *Supramol. Chem.* **1995**, 5, 97
- 88 Dupont, N., Lazar, A. N., Perret, F., Danylyuk, O., Suwinska, K., Navaza, A. & Coleman, A. W. Solid state structures of the complexes between the antiseptic chlorhexidine and three anionic derivatives of calix[4]arene. *CrystEngComm* **2008**, 10, 975.
- 89 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.
- 90 Selkti, M., Coleman, A. W., Nicolis, I., Douteau-Guevel, 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.
- 91 Lazar, A., Da Silva, E., Navaza, A., Barbey, C. & Coleman, A. W. A new packing motif for *para*-sulfonatocalix[4]arene: The solid state structure of the *para*-sulfonatocalix[4]arene D-arginine complex. *Chem. Commun.* **2004**, 2162.

- 92 Douteau-Guével, N., Coleman, A. W., Morel, J. P. & Morel-Desrosiers, N. Complexation of basic amino acids by water-soluble calixarene sulphonates as a study of the possible mechanisms of recognition of calixarene sulphonates by proteins. *J. Phys. Org. Chem.* **1998**, 11, 693.
- 93 Douteau-Guevel, N., Coleman, A. W., Morel, J. P. & Morel-Desrosiers, N. Complexation of the basic amino acids lysine and arginine by three sulfonatocalixarenes (n = 4, 6 and 8) in water: Microcalorimetric determination of the Gibbs energies, enthalpies and entropies of complexation. *J. Chem. Soc., Perkin Trans. 2* **1999**, 629.
- 94 Arena, G., Casnati, A., Contino, A., Magri, A., Sansone, F., Sciotto, D. & Ungaro, R. Inclusion of naturally occurring amino acids in water soluble calix[4]arenes: A microcalorimetric and ¹H NMR investigation supported by molecular modeling. *Org. Biomol. Chem.* **2006**, 4, 243.
- 95 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.
- 96 Memmi, L., Lazar, A. N., Brioude, A., Ball, V. & Coleman, A. W. Protein-calixarene interactions: Complexation of bovine serum albumin by sulfonatocalix[n]arenes. *Chem Commun (Camb)* **2001**, 2474.
- 97 Bujacz, A. Structures of bovine, equine and leporine serum albumin. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2012**, 68, 1278.
- 98 Silva, E. D., Rousseau, C. F., Zanella-Cléon, I., Becchi, M. & Coleman, A. W. Mass Spectrometric Determination of association constants of bovine serum albumin (BSA) with *para*-sulphonato-calix[n]arene derivatives. *J. Incl. Phenom. Macro.* **2006**, 54, 53.
- 99 Nau, W. M., Ghale, G., Hennig, A., Bakirci, H. s. & Bailey, D. M. Substrate-selective supramolecular tandem assays: Monitoring enzyme inhibition of arginase and diamine oxidase by fluorescent dye displacement from calixarene and cucurbituril macrocycles. *J. Am. Chem. Soc.* **2009**, 131, 11558.
- 100 Minaker, S. A., Daze, K. D., Ma, M. C. F. & Hof, F. Antibody-free reading of the histone code using a simple chemical sensor array. *J. Am. Chem. Soc.* **2012**, 134, 11674.
- 101 Zadmard, R. & Schrader, T. Nanomolar protein sensing with embedded receptor molecules. *J. Am. Chem. Soc.* **2004**, 127, 904.

- 102 Mohsin, M. A., Banica, F.-G., Oshima, T. & Hianik, T. Electrochemical impedance spectroscopy for assessing the recognition of cytochrome *c* by immobilized calixarenes. *Electroanalysis* **2011**, 23, 1229.
- 103 Oshima, T., Goto, M. & Furusaki, S. Complex formation of cytochrome *c* with a calixarene carboxylic acid derivative: A novel solubilization method for biomolecules in organic media. *Biomacromolecules* **2002**, 3, 438.
- 104 Oshima, T., Higuchi, H., Ohto, K., Inoue, K. & Goto, M. Selective extraction and recovery of cytochrome *c* by liquid–liquid extraction using a calix[6]arene carboxylic acid derivative. *Langmuir* **2005**, 21, 7280.
- 105 An, W. T., Jiao, Y., Sun, X. H., Zhang, X. L., Dong, C., Shuang, S. M., Xia, P. F. & Wong, M. S. Synthesis and binding properties of carboxylphenyl-modified calix[4]arenes and cytochrome *c*. *Talanta* **2009**, 79, 54.
- 106 Trush, V. V., Cherenok, S. O., Tanchuk, V. Y., Kukhar, V. P., Kalchenko, V. I. & Vovk, A. I. Calix[4]arene methylenebisphosphonic acids as inhibitors of protein tyrosine phosphatase 1B. *Bioorg. Med. Chem. Lett.* **2013**, 23, 5619.
- 107 Hamuro, Y., Calama, M. C., Park, H. S. & Hamilton, A. D. A calixarene with four peptide loops: An antibody mimic for recognition of protein surfaces. *Angew. Chem. Int. Ed.* **1997**, 36, 2680.
- 108 Wei, Y., McLendon, G. L., Hamilton, A. D., Case, M. A., Purring, C. B., Lin, Q., Park, H. S., Lee, C.-S. & Yu, T. Disruption of protein-protein interactions: design of a synthetic receptor that blocks the binding of cytochrome *c* to cytochrome *c* peroxidase. *Chem. Commun.* **2001**, 1580.
- 109 Park, H. S., Lin, Q. & Hamilton, A. D. Protein surface recognition by synthetic receptors: A route to novel submicromolar inhibitors for α -chymotrypsin. *J. Am. Chem. Soc.* **1998**, 121, 8.
- 110 Gordo, S., Martos, V., Santos, E., Menendez, M., Bo, C., Giralt, E. & de Mendoza, J. Stability and structural recovery of the tetramerization domain of p53-R337H mutant induced by a designed templating ligand. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 16426.
- 111 Kamada, R., Yoshino, W., Nomura, T., Chuman, Y., Imagawa, T. & Suzuki, T. Enhancement of transcriptional activity of mutant p53 tumor suppressor protein through stabilization of tetramer formation by calix[6]arene derivatives. *Bioorg. Med. Chem. Lett.* **2010**, 20, 4412.

- 112 Long, S. B., Campbell, E. B. & MacKinnon, R. Crystal structure of a mammalian voltage-dependent shaker family K⁺ channel. *Science* **2005**, 309, 897.
- 113 Martos, V., Bell, S. C., Santos, E., Isacoff, E. Y., Trauner, D. & de Mendoza, J. Calix[4]arene-based conical-shaped ligands for voltage-dependent potassium channels. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 10482.
- 114 Dings, R. P. M., Chen, X., Hellebrekers, D. M. E. I., van Eijk, L. I., Zhang, Y., Hoye, T. R., Griffioen, A. W. & Mayo, K. H. Design of nonpeptidic topomimetics of antiangiogenic proteins with antitumor activities. *J. Natl. Cancer. Inst.* **2006**, 98, 932.
- 115 Dings, R. P. M., Miller, M. C., Nesmelova, I., Astorgues-Xerri, L., Kumar, N., Serova, M., Chen, X., Raymond, E., Hoye, T. R. & Mayo, K. H. Antitumor agent calixarene 0118 targets human galectin-1 as an allosteric inhibitor of carbohydrate binding. *J. Med. Chem.* **2012**, 55, 5121.
- 116 Jain, R. K. & Hamilton, A. D. Protein surface recognition by synthetic receptors based on a tetraphenylporphyrin scaffold. *Org. Lett.* **2000**, 2, 1721.
- 117 Jain, R. K. & Hamilton, A. D. Designing protein denaturants: Synthetic agents induce cytochrome *c* unfolding at low concentrations and stoichiometries. *Angew. Chem. Int. Ed.* **2002**, 41, 641.
- 118 Aya, T. & Hamilton, A. D. Tetrabiphenylporphyrin-based receptors for protein surfaces show sub-nanomolar affinity and enhance unfolding. *Bioorg. Med. Chem. Lett.* **2003**, 13, 2651.
- 119 Crowley, P. B., Ganji, P. & Ibrahim, H. Protein surface recognition: Structural characterisation of cytochrome *c*-porphyrin complexes. *ChemBioChem* **2008**, 9, 1029.
- 120 Zhou, H., Baldini, L., Hong, J., Wilson, A. J. & Hamilton, A. D. Pattern recognition of proteins based on an array of functionalized porphyrins. *J. Am. Chem. Soc.* **2006**, 128, 2421.
- 121 Gradl, S. N., Felix, J. P., Isacoff, E. Y., Garcia, M. L. & Trauner, D. Protein surface recognition by rational design: Nanomolar ligands for potassium channels. *J. Am. Chem. Soc.* **2003**, 125, 12668.
- 122 Renner, C., Piehler, J. & Schrader, T. Arginine- and lysine-specific polymers for protein recognition and immobilization. *J. Am. Chem. Soc.* **2005**, 128, 620.
- 123 Koch, S. J., Renner, C., Xie, X. & Schrader, T. Tuning linear copolymers into protein-specific hosts. *Angew. Chem. Int. Ed.* **2006**, 45, 6352.

-
- 124 Arendt, M., Sun, W., Thomann, J., Xie, X. & Schrader, T. Dendrimeric bisphosphonates for multivalent protein surface binding. *Chemistry – An Asian Journal* **2006**, 1, 544.
- 125 Chiba, F., Chou, T., Twyman, L. J. & Wagstaff, M. Dendrimers as size selective inhibitors to protein-protein binding. *Chem. Commun.* **2008**, 4351.
- 126 Ow, Y. L. P., Green, D. R., Hao, Z. & Mak, T. W. Cytochrome *c*: Functions beyond respiration. *Nat. Rev. Mol. Cell Biol.* **2008**, 9, 532.
- 127 Pollock, W., Rosell, F., Twitchett, M., Dumont, M. & Mauk, A. Bacterial expression of a mitochondrial cytochrome *c*. Trimethylation of Lys72 in yeast iso-1-cytochrome *c* and the alkaline conformational transition. *Biochemistry* **1998**, 37, 6124.
- 128 Volkov, A. N., Worrall, J. A. R., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome *c* and cytochrome *c* peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 18945.
- 129 Lange, C. & Hunte, C. Crystal structure of the yeast cytochrome *bc1* complex with its bound substrate cytochrome *c*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99, 2800.
- 130 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.
- 131 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.
- 132 Wienk, H., Maneg, O., Lücke, C., Pristovšek, P., Löhr, F., Ludwig, B. & Rüterjans, H. Interaction of cytochrome *c* with cytochrome *c* oxidase: An NMR study on two soluble fragments derived from *Paracoccus denitrificans*. *Biochemistry* **2003**, 42, 6005.
- 133 Guo, M., Bhaskar, B., Li, H., Barrows, T. P. & Poulos, T. L. Crystal structure and characterization of a cytochrome *c* peroxidase–cytochrome *c* site-specific cross-link. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 5940.

- 134 Rupley, J. A. The hydrolysis of chitin by concentrated hydrochloric acid, and the preparation of low-molecular-weight substrate for lysozyme. *Biochim. Biophys. Acta* **1964**, 83, 245.
- 135 Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. Structure of hen egg-white lysozyme: A three-dimensional fourier synthesis at 2 angstrom resolution. *Nature* **1965**, 206, 757.
- 136 Hameed, M., Ahmad, B., Fazili, K. M., Andrabi, K. & Khan, R. H. Different molten globule-like folding intermediates of hen egg white lysozyme induced by high pH and tertiary butanol. *J. Biochem.* **2007**, 141, 573.
- 137 Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Kuhn, P., Weis, W. I., Kobilka, B. K. & Stevens, R. C. High-resolution crystal structure of an engineered human β 2-adrenergic G protein-coupled receptor. *Science* **2007**, 318, 1258.
- 138 Diamond, R. Real-space refinement of structure of hen egg-white lysozyme. *J. Mol. Biol.* **1974**, 82, 371.
- 139 Daze, K. D., Pinter, T., Beshara, C. S., Ibraheem, A., Minaker, S. A., Ma, M. C. F., Courtemanche, R. J. M., Campbell, R. E. & Hof, F. Supramolecular hosts that recognize methyllysines and disrupt the interaction between a modified histone tail and its epigenetic reader protein. *Chem. Sci.* **2012**.
- 140 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, 372, 774.
- 141 Schmit, J. D. & Dill, K. A. The stabilities of protein crystals. *J. Phys. Chem. B.* **2010**, 114, 4020.
- 142 Doye, J. P., Louis, A. A. & Vendruscolo, M. Inhibition of protein crystallization by evolutionary negative design. *Physical biology.* **2004**, 1, P9.
- 143 Derewenda, Z. S. Rational protein crystallization by mutational surface engineering. *Structure* **2004**, 12, 529.
- 144 Asherie, N. Protein crystallization and phase diagrams. *Methods* **2004**, 34, 266.
- 145 Schmit, J. D. & Dill, K. Growth rates of protein crystals. *J. Am. Chem. Soc.* **2012**, 134, 3934.
- 146 Chayen, N. E. & Saridakis, E. Protein crystallization: From purified protein to diffraction-quality crystal. *Nature Methods* **2008**, 5, 147.

-
- 147 Luft, J. R., Wolfley, J. R. & Snell, E. H. What's in a Drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Cryst. Growth Des.* **2012**, 11, 651.
- 148 Arakawa, T. & Timasheff, S. N. Mechanism of protein salting in and salting out by divalent cation salts: Balance between hydration and salt binding. *Biochemistry* **1984**, 23, 5912.
- 149 McPherson, A. A comparison of salts for the crystallization of macromolecules. *Protein Sci.* **2001**, 10, 418.
- 150 Dumetz, A. C., Snellinger-O'Brien, A. M., Kaler, E. W. & Lenhoff, A. M. Patterns of protein-protein interactions in salt solutions and implications for protein crystallization. *Protein Sci.* **2007**, 16, 1867.
- 151 Judge, R. A., Jacobs, R. S., Frazier, T., Snell, E. H. & Pusey, M. L. The effect of temperature and solution pH on the nucleation of tetragonal lysozyme crystals. *Biophys. J.* **1999**, 77, 1585.
- 152 McPherson, A. Crystallization of proteins from polyethylene glycol. *J. Biol. Chem.* **1976**, 251, 6300.
- 153 Bhat, R. & Timasheff, S. N. Steric exclusion is the principal source of the preferential hydration of proteins in the presence of polyethylene glycols. *Protein Sci.* **1992**, 1, 1133.
- 154 Salgado, E. N., Radford, R. J. & Tezcan, F. A. Metal-directed protein self-assembly. *Acc. Chem. Res.* **2010**, 43, 661.
- 155 McPherson, A. Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **1990**, 189, 1.
- 156 Bond, C. S., Bendall, D. S., Freeman, H. C., Guss, J. M., Howe, C. J., Wagner, M. J. & Wilce, M. C. The structure of plastocyanin from the cyanobacterium *Phormidium laminosum*. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **1999**, 55, 414.
- 157 Laganowsky, A., Zhao, M., Soriaga, A. B., Sawaya, M. R., Cascio, D. & Yeates, T. O. An approach to crystallizing proteins by metal-mediated synthetic symmetrization. *Protein Sci.* **2011**, 20, 1876.
- 158 McPherson, A. & Cudney, B. Searching for silver bullets: An alternative strategy for crystallizing macromolecules. *J. Struct. Biol.* **2006**, 156, 387.
- 159 Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter Iii, H. L., Deng, S. J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J.,

- Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R. & Shewchuk, L. M. Crystallization of protein-ligand complexes. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2006**, 63, 72.
- 160 Danley, D. Crystallization to obtain protein-ligand complexes for structure-aided drug design. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2006**, 62, 569.
- 161 Fischer, E. Einfluss der configuration auf die wirkung der enzyme. *Berichte der deutschen chemischen Gesellschaft* **1894**, 27, 2985.
- 162 Shewchuk, L., Hassell, A., Wisely, B., Rocque, W., Holmes, W., Veal, J. & Kuyper, L. F. Binding mode of the 4-anilinoquinazoline class of protein kinase inhibitor: X-ray crystallographic studies of 4-anilinoquinazolines bound to cyclin-dependent kinase 2 and p38 kinase. *J. Med. Chem.* **1999**, 43, 133.
- 163 Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, 33, 491.
- 164 Matthews, B. W. Determination of molecular-weight from protein crystals. *J. Mol. Biol.* **1974**, 82, 513.
- 165 Eck, M. J. & Sprang, S. R. The structure of tumor necrosis factor-alpha at 2.6 Å resolution. Implications for receptor binding. *J. Biol. Chem.* **1989**, 264, 17595.
- 166 Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J. Biol. Chem.* **2004**, 279, 48270.
- 167 Kuciauskas, D. & Caputo, G. A. Self-assembly of peptide-porphyrin complexes leads to pH-dependent excitonic coupling. *J. Phys. Chem. B* **2009**, 113, 14439.
- 168 Whiting, A. L., Neufeld, N. M. & Hof, F. A tryptophan-analog host whose interactions with ammonium ions in water are dominated by the hydrophobic effect. *Tetrahedron Lett.* **2009**, 50, 7035.
- 169 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2011**, 67, 235.
- 170 Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **1992**.
- 171 Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2006**, 62, 72.

- 172 Kantardjieff, K. A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci.* **2003**, 12, 1865.
- 173 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.
- 174 Emsley, P. & Cowtan, K. COOT: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2004**, 2126.
- 175 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**, 60, 2184.
- 176 Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S. & Richardson, D. C. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, 35, W375.
- 177 Ramachandran, G. N. & Sasisekharan, V. in *Adv. Protein Chem.* Vol. Volume 23 (eds M. L. Anson John T. Edsall C.B. Anfinsen & M. Richards Frederic) 283 (Academic Press, 1968).
- 178 Rayment, I. Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods Enzymol.* **1997**, 276, 171.
- 179 Zorn, J. A. & Wells, J. A. Turning enzymes ON with small molecules. *Nat. Chem. Biol.* **2010**, 6, 179.
- 180 Larson, S., Day, J., Cudney, R. & McPherson, A. A novel strategy for the crystallization of proteins: X-ray diffraction validation. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2007**, 63, 310.
- 181 Cao, Y., Jin, X., Levin, E. J., Huang, H., Zong, Y., Quick, M., Weng, J., Pan, Y., Love, J., Punta, M., Rost, B., Hendrickson, W. A., Javitch, J. A., Rajashankar, K. R. & Zhou, M. Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature* **2011**, 473, 50.
- 182 Ingerman, L. A., Cuellar, M. E. & Waters, M. L. A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity. *Chem. Commun.* **2010**, 46, 1839.

- 183 Li, H. T., Ilin, S., Wang, W. K., Duncan, E. M., Wysocka, J., Allis, C. D. & Patel, D. J. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **2006**, 442, 91.
- 184 Lin, Q. & Hamilton, A. D. Design and synthesis of multiple-loop receptors based on a calix[4]arene scaffold for protein surface recognition. *Comptes Rendus Chimie* **2002**, 5, 441.
- 185 Garrido, C. & Kroemer, G. Life's smile, death's grin: Vital functions of apoptosis-executing proteins. *Curr. Opin. Cell. Biol.* **2004**, 16, 639.
- 186 Hill, M. M., Adrain, C., Duriez, P. J., Creagh, E. M. & Martin, S. J. Analysis of the composition, assembly kinetics and activity of native Apaf-1 apoptosomes. *EMBO J.* **2004**, 23, 2134.
- 187 Oshima, T., Baba, Y., Shimojo, K. & Goto, M. Recognition of lysine residues on protein surfaces using calixarenes and its application. *Curr. Drug. Discov. Tech.* **2007**, 4, 220.
- 188 Assfalg, M., Bertini, I., Del Conte, R., Giachetti, A. & Turano, P. Cytochrome *c* and organic molecules: Solution structure of the *p*-aminophenol adduct. *Biochemistry* **2007**, 46, 6232.
- 189 Hunter, C. A. & Anderson, H. L. What is cooperativity? *Angew. Chem. Int. Ed.* **2009**, 48, 7488.
- 190 Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **2002**, 295, 2080.
- 191 Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V. & Laue, E. D. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **2002**, 416, 103.
- 192 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. & Steed, J. W. The structure of water in *p*-sulfonatocalix[4]arene. *Chem. Eur. J.* **2011**, 17, 10259.
- 193 Debye, P. *Ann. Phys. (Weinheim, Ger.)* **1915**, 3514, 809.
- 194 Watson, J. D. & Milner-White, E. J. A novel main-chain anion-binding site in proteins: the nest. A particular combination of ϕ, ψ values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. *J. Mol. Biol.* **2002**, 315, 171.

-
- 195 Denessiouk, K. A., Johnson, M. S. & Denesyuk, A. I. Novel C^αNN structural motif for protein recognition of phosphate ions. *J. Mol. Biol.* **2005**, 345, 611.
- 196 Weimann, D. P., Winkler, H. D. F., Falenski, J. A., Kokschi, B. & Schalley, C. A. Highly dynamic motion of crown ethers along oligolysine peptide chains. *Nature Chem.* **2009**, 1, 573.
- 197 Falson, P. G., Coleman, A. W., Matar-Merheb, R., Leydier, A. & Huché, F. Additive for the crystallization of proteins, use and process. (2012).
- 198 Derewenda, Z. S. & Vekilov, P. G. Entropy and surface engineering in protein crystallization. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2006**, 62, 116.
- 199 Lamberto, G. R., Binolfi, A., Orcellet, M. L., Bertoncini, C. W., Zweckstetter, M., Griesinger, C. & Fernández, C. O. Structural and mechanistic basis behind the inhibitory interaction of PcTS on α -synuclein amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 21057.
- 200 Dang, D. T., Schill, J. & Brunsveld, L. Cucurbit[8]uril-mediated protein homotetramerization. *Chem. Sci.* **2012**, 3, 2679.
- 201 Tossavainen, H., Pihlajamaa, T., Huttunen, T. K., Raulo, E., Rauvala, H., Permi, P. & Kilpelainen, I. The layered fold of the TSR domain of P-falciparum TRAP contains a heparin binding site. *Protein Sci.* **2006**, 15, 1760.
- 202 Crowley, P. B., Chow, E. & Papkovskaia, T. Protein interactions in the escherichia coli cytosol: An impediment to in-cell nmr spectroscopy. *Chembiochem* **2011**, 12, 1043.
- 203 Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Elsevier* **2005**, 207.
- 204 Morar, A. S., Kakouras, D., Young, G. B., Boyd, J. & Pielak, G. J. Expression of ¹⁵N-labeled eukaryotic cytochrome *c* in Escherichia coli. *J. Biol. Inorg. Chem.* **1999**, 4, 220.
- 205 Krissinel, E. & Henrick, K. Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute. *J. Mol. Biol.* **2007**, 372, 774.
- 206 De Lano, W. L. The PyMOL Molecular Graphics System **2002**.
- 207 Katz, E., Bückmann, A. F. & Willner, I. Self-powered enzyme-based biosensors. *J. Am. Chem. Soc.* **2001**, 123, 10752.
- 208 Spricigo, R., Dronov, R., Lisdat, F., Leimkühler, S., Scheller, F. & Wollenberger, U. Electrocatalytic sulfite biosensor with human sulfite oxidase co-immobilized

- with cytochrome *c* in a polyelectrolyte-containing multilayer. *Anal Bioanal Chem* **2009**, 393, 225.
- 209 Ma, X., Zhang, L.-H., Wang, L.-R., Xue, X., Sun, J.-H., Wu, Y., Zou, G., Wu, X., Wang, P. C., Wamer, W. G., Yin, J.-J., Zheng, K. & Liang, X.-J. Single-walled carbon nanotubes alter cytochrome *c* electron transfer and modulate mitochondrial function. *ACS Nano* **2012**, 6, 10486.
- 210 Song, S., Clark, R. A., Bowden, E. F. & Tarlov, M. J. Characterization of cytochrome *c*/alkanethiolate structures prepared by self-assembly on gold. *J. Phys. Chem.* **1993**, 97, 6564.
- 211 Avila, A., Gregory, B. W., Niki, K. & Cotton, T. M. An electrochemical approach to investigate gated electron transfer using a physiological model system: Cytochrome *c* immobilized on carboxylic acid-terminated alkanethiol self-assembled monolayers on gold electrodes. *J. Phys. Chem. B.* **2000**, 104, 2759.
- 212 Murgida, D. H. & Hildebrandt, P. Proton-coupled electron transfer of cytochrome *c*. *J. Am. Chem. Soc.* **2001**, 123, 4062.
- 213 Chen, X., Ferrigno, R., Yang, J. & Whitesides, G. M. Redox properties of cytochrome *c* adsorbed on self-assembled monolayers: A probe for protein conformation and orientation. *Langmuir* **2002**, 18, 7009.
- 214 Kranich, A., Ly, H. K., Hildebrandt, P. & Murgida, D. H. Direct observation of the gating step in protein electron transfer: Electric-field-controlled protein dynamics. *J. Am. Chem. Soc.* **2008**, 130, 9844.
- 215 Kuzume, A., Zhumaev, U., Li, J., Fu, Y., Füeg, M., Esteve-Nuñez, A. & Wandlowski, T. An in-situ surface electrochemistry approach toward whole-cell studies: Charge transfer between *Geobacter sulfurreducens* and electrified metal/electrolyte interfaces through linker molecules. *Electrochim. Acta* **2013**, 112, 933.
- 216 Kayushina, R., Lvov, Y., Stepina, N., Belyaev, V. & Khurgin, Y. Construction and X-ray reflectivity study of self-assembled lysozyme/polyion multilayers. *Thin Solid Films* **1996**, 284–285, 246.
- 217 Lvov, Y. M., Lu, Z., Schenkman, J. B., Zu, X. & Rusling, J. F. Direct electrochemistry of myoglobin and cytochrome p450_{cam} in alternate layer-by-layer films with DNA and other polyions. *J. Am. Chem. Soc.* **1998**, 120, 4073.

-
- 218 Wettstein, C., Möhwald, H. & Lisdat, F. Coupling of pyrroloquinoline quinone dependent glucose dehydrogenase to (cytochrome *c*/DNA)-multilayer systems on electrodes. *Bioelectrochemistry* **2012**, 88, 97.
- 219 Wegerich, F., Turano, P., Allegrozzi, M., Mohwald, H. & Lisdat, F. Electroactive multilayer assemblies of bilirubin oxidase and human cytochrome *c* mutants: Insight in formation and kinetic behavior. *Langmuir* **2011**, 27, 4202.
- 220 Feifel, S. & Lisdat, F. Silica nanoparticles for the layer-by-layer assembly of fully electro-active cytochrome *c* multilayers. *J. Nanobiotechnology* **2011**, 9, 59.
- 221 Crane, B. R., Di Bilio, A. J., Winkler, J. R. & Gray, H. B. Electron tunneling in single crystals of *pseudomonas aeruginosa* azurins. *J. Am. Chem. Soc.* **2001**, 123, 11623.
- 222 Tezcan, F. A., Crane, B. R., Winkler, J. R. & Gray, H. B. Electron tunneling in protein crystals. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, 98, 5002.
- 223 Ferrari, D., Merli, A., Peracchi, A., Di Valentin, M., Carbonera, D. & Rossi, G. L. Catalysis and electron transfer in protein crystals: The binary and ternary complexes of methylamine dehydrogenase with electron acceptors. *Biochim. Biophys. Acta* **2003**, 1647, 337.
- 224 Kang, S. A. & Crane, B. R. Effects of interface mutations on association modes and electron-transfer rates between proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 15465.
- 225 Cavaliere, C., Biermann, N., Vlasie, M. D., Einsle, O., Merli, A., Ferrari, D., Rossi, G. L. & Ubbink, M. Structural comparison of crystal and solution states of the 138 kDa complex of methylamine dehydrogenase and amicyanin from *paracoccus versutus*. *Biochemistry* **2008**, 47, 6560.
- 226 Acosta, F., Eid, D., Marín-García, L., Frontana-Urbe, B. A. & Moreno, A. From cytochrome *c* crystals to a solid-state electron-transfer device. *Cryst. Growth Des.* **2007**, 7, 2187.
- 227 Crowley, P. B., Matias, P. M., Mi, H., Firbank, S. J., Banfield, M. J. & Dennison, C. Regulation of protein function: Crystal packing interfaces and conformational dimerization. *Biochemistry* **2008**, 47, 6583.
- 228 Beissenhirtz, M. K., Scheller, F. W. & Lisdat, F. A superoxide sensor based on a multilayer cytochrome *c* electrode. *Anal. Chem.* **2004**, 76, 4665.

- 229 Rudolph, M. J., Johnson, J. L., Rajagopalan, K. V. & Kisker, C. The 1.2 Angstrom structure of the human sulfite oxidase cytochrome *b*(5) domain. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2003**, 59, 1183.
- 230 Sarauli, D., Xu, C., Dietzel, B., Stiba, K., Leimkuhler, S., Schulz, B. & Lisdat, F. Thin films of substituted polyanilines: Interactions with biomolecular systems. *Soft Matter* **2012**, 8, 3848.
- 231 Gray, H. B. & Winkler, J. R. Electron flow through metalloproteins. *Biochim. Biophys. Acta* **2010**, 1797, 1563.
- 232 Sarauli, D., Tanne, J., Xu, C., Schulz, B., Trnkova, L. & Lisdat, F. Insights into the formation and operation of polyaniline sulfonate/cytochrome *c* multilayer electrodes: Contributions of polyelectrolytes' properties. *PCCP* **2010**, 12, 14271.
- 233 Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. Nature of biological electron transfer. *Nature* **1992**, 355, 796.
- 234 Gray, H. B. & Winkler, J. R. Long-range electron transfer. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 3534.
- 235 Koshiyama, T., Shirai, M., Hikage, T., Tabe, H., Tanaka, K., Kitagawa, S. & Ueno, T. Post-crystal engineering of zinc-substituted myoglobin to construct a long-lived photoinduced charge-separation system. *Angew. Chem. Int. Ed.* **2011**, 50, 4849.
- 236 Kano, K. & Ishida, Y. Supramolecular complex of cytochrome *c* with a polyanionic β -cyclodextrin. *Angew. Chem. Int. Ed.* **2007**, 46, 727.
- 237 Janin, J., Miller, S. & Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **1988**, 204, 155.
- 238 Holdermann, I., Meyer, N. H., Round, A., Wild, K., Sattler, M. & Sinning, I. Chromodomains read the arginine code of post-translational targeting. *Nat. Struct. Mol. Biol.* **2012**, 19, 260.
- 239 Lehn, J. M., Vierling, P. & Hayward, R. C. Stable and selective guanidinium and imidazolium complexes of a macrocyclic receptor molecule. *J. Chem. Soc., Chem. Commun.* **1979**, 296.
- 240 Ngola, S. M., Kearney, P. C., Mecozzi, S., Russell, K. & Dougherty, D. A. A selective receptor for arginine derivatives in aqueous media. Energetic consequences of salt bridges that are highly exposed to water. *J. Am. Chem. Soc.* **1999**, 121, 1192.

- 241 Bell, T. W., Khasanov, A. B., Drew, M. G. B., Filikov, A. & James, T. L. A small-molecule guanidinium receptor: The arginine cork. *Angew. Chem. Int. Ed.* **1999**, 38, 2543.
- 242 Grawe, T., Schrader, T., Finocchiaro, P., Consiglio, G. & Failla, S. A new receptor molecule for lysine and histidine in water: Strong binding of basic amino acid esters by a macrocyclic host. *Org. Lett.* **2001**, 3, 1597.
- 243 Douteau-Guevel, N., Perret, F., Coleman, A. W., Morel, J.-P. & Morel-Desrosiers, N. Binding of dipeptides and tripeptides containing lysine or arginine by *p*-sulfonatocalixarenes in water: NMR and microcalorimetric studies. *Journal of the Chemical Society, Perkin Transactions 2* **2002**, 524.
- 244 Motonaga, K., Misaka, E., Nakajima, E., Ueda, S. & Nakanishi, K. Structure of yeast cytochrome *c*: I. monomer and dimer interconversion. *J. Biochem.* **1965**, 57, 22.
- 245 Haliloglu, T. & Bahar, I. Structure-based analysis of protein dynamics: comparison of theoretical results for hen lysozyme with X-ray diffraction and NMR relaxation data. *Proteins* **1999**, 37, 654.
- 246 Ojala, W. H., Sudbeck, E. A., Lu, L. K., Richardson, T. I., Lovrien, R. E. & Gleason, W. B. Complexes of lysine, histidine, and arginine with sulfonated azo dyes: Model systems for understanding the biomolecular recognition of glycosaminoglycans by proteins. *J. Am. Chem. Soc.* **1996**, 118, 2131.
- 247 Drljaca, A., Hardie, M. J., Raston, C. L. & Spiccia, L. Self-assembled superanions: Ionic capsules stabilized by polynuclear chromium(iii) aqua cations. *Chem. Eur. J.* **1999**, 5, 2295.
- 248 Hall, D. H., Grove, L. E., Yueh, C., Ngan, C. H., Kozakov, D. & Vajda, S. Robust identification of binding hot spots using continuum electrostatics: Application to hen egg-white lysozyme. *J. Am. Chem. Soc.* **2011**, 133, 20668.
- 249 Markies, P. R., Akkerman, O. S., Bickelhaupt, F., Smeets, W. J. J. & Spek, A. L. Complexation of bis(*p*-tert-butylphenyl)magnesium with 1,3-xylylene crown ethers and glymes. Vol. 13 (American Chemical Society, 1994).
- 250 Furness, E. L., Ross, A., Davis, T. P. & King, G. C. A hydrophobic interaction site for lysozyme binding to polyethylene glycol and model contact lens polymers. *Biomaterials* **1998**, 19, 1361.
- 251 Crowley, P. B., Brett, K. & Muldoon, J. NMR spectroscopy reveals cytochrome *c*-poly(ethylene glycol) interactions. *ChemBioChem* **2008**, 9, 685.

- 252 Wenck, K., Koch, S., Renner, C., Sun, W. & Schrader, T. A noncovalent switch for lysozyme. *J. Am. Chem. Soc.* **2007**, 129, 16015.
- 253 Ganguli, S., Yoshimoto, K., Tomita, S., Sakuma, H., Matsuoka, T., Shiraki, K. & Nagasaki, Y. Regulation of lysozyme activity based on thermotolerant protein/smart polymer complex formation. *J. Am. Chem. Soc.* **2009**, 131, 6549.
- 254 Dall'Acqua, W., Goldman, E. R., Lin, W., Teng, C., Tsuchiya, D., Li, H., Ysern, X., Braden, B. C., Li, Y., Smith-Gill, S. J. & Mariuzza, R. A. A Mutational analysis of binding interactions in an antigen–antibody protein–protein complex. *Biochemistry* **1998**, 37, 7981.
- 255 Sundberg, E. J., Urrutia, M., Braden, B. C., Isern, J., Tsuchiya, D., Fields, B. A., Malchiodi, E. L., Tormo, J., Schwarz, F. P. & Mariuzza, R. A. Estimation of the hydrophobic effect in an antigen–antibody protein–protein interface. *Biochemistry* **2000**, 39, 15375.
- 256 De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S. & Wyns, L. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 4586.
- 257 Deng, L., Velikovskiy, C. A., Xu, G., Iyer, L. M., Tasumi, S., Kerzic, M. C., Flajnik, M. F., Aravind, L., Pancer, Z. & Mariuzza, R. A. A structural basis for antigen recognition by the T cell-like lymphocytes of sea lamprey. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, 107, 13408.
- 258 De Genst, E., Silence, K., Ghahroudi, M. A., Decanniere, K., Loris, R., Kinne, J., Wyns, L. & Muyldermans, S. Strong in vivo maturation compensates for structurally restricted H3 loops in antibody repertoires. *J. Biol. Chem.* **2005**, 280, 14114.
- 259 Stanfield, R. L., Dooley, H., Verdino, P., Flajnik, M. F. & Wilson, I. A. Maturation of shark single-domain (IgNAR) antibodies: Evidence for induced-fit binding. *J. Mol. Biol.* **2007**, 367, 358.
- 260 Watson, J. D. & Milner-White, E. J. The conformations of polypeptide chains where the main-chain parts of successive residues are enantiomeric. Their occurrence in cation and anion-binding regions of proteins. *J. Mol. Biol.* **2002**, 315, 183.

- 261 Majeed, S., Ofek, G., Belachew, A., Huang, C.-c., Zhou, T. & Kwong, P. D. Enhancing protein crystallization through precipitant synergy. *Structure* **2003**, 11, 1061.
- 262 Razavet, M., Artero, V., Cavazza, C., Oudart, Y., Lebrun, C., Fontecilla-Camps, J. C. & Fontecave, M. Tricarbonylmanganese(I)-lysozyme complex: A structurally characterized organometallic protein. *Chem. Commun.* **2007**, 2805.
- 263 Michaux, C., Pouyez, J., Wouters, J. & Prive, G. G. Protecting role of cosolvents in protein denaturation by SDS: A structural study. *BMC Struct. Biol.* **2008**, 8.
- 264 Vaney, M. C., Broutin, I., Retailleau, P., Douangamath, A., Lafont, S., Hamiaux, C., Prange, T., Ducruix, A. & Ries-Kautt, M. Structural effects of monovalent anions on polymorphic lysozyme crystals. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2001**, 57, 929.
- 265 Beck, T., Krasauskas, A., Gruene, T. & Sheldrick, G. M. A magic triangle for experimental phasing of macromolecules. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2008**, 64, 1179.
- 266 Parthier, C., Kleinschmidt, M., Neumann, P., Rudolph, R., Manhart, S., Schlenzig, D., Fanghänel, J., Rahfeld, J.-U., Demuth, H.-U. & Stubbs, M. T. Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 13942.
- 267 Hirsch, A. K. H., Fischer, F. R. & Diederich, F. Phosphate recognition in structural biology. *Angew. Chem. Int. Ed.* **2007**, 46, 338.
- 268 Sansone, F., Dudič, M., Donofrio, G., Rivetti, C., Baldini, L., Casnati, A., Cellai, S. & Ungaro, R. DNA condensation and cell transfection properties of guanidinium calixarenes: Dependence on macrocycle lipophilicity, size, and conformation. *J. Am. Chem. Soc.* **2006**, 128, 14528.
- 269 Martos, V., Castreño, P., Valero, J. & de Mendoza, J. Binding to protein surfaces by supramolecular multivalent scaffolds. *Curr. Opin. Chem. Biol.* **2008**, 12, 698.
- 270 Dutt, S., Wilch, C. & Schrader, T. Artificial synthetic receptors as regulators of protein activity. *Chem. Commun.* **2011**, 47, 5376.
- 271 Leontiev, A. V., Jemmett, C. A. & Beer, P. D. Anion recognition and cation-induced molecular motion in a heteroditopic [2]rotaxane. *Chem. Eur. J.* **2011**, 17, 816.

- 272 Dionisio, M., Oliviero, G., Menozzi, D., Federici, S., Yebeutchou, R. M., Schmidtchen, F. P., Dalcanale, E. & Bergese, P. Nanomechanical Recognition of N-Methylammonium Salts. *J. Am. Chem. Soc.* **2012**, 134, 2392.
- 273 Dutt, S., Wilch, C., Gersthagen, T., Talbiersky, P., Bravo-Rodriguez, K., Hanni, M., Sánchez-García, E., Ochsenfeld, C., Klärner, F.-G. & Schrader, T. Molecular tweezers with varying anions: A comparative study. *J. Org. Chem.* **2013**, 78, 6721.
- 274 Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **1994**, 760.
- 275 van Asselt, E. J., Perrakis, A., Kalk, K. H., Lamzin, V. S. & Dijkstra, B. W. Accelerated X-ray structure elucidation of a 36 kDa muramidase/transglycosylase using wARP. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **1998**, 54, 58.
- 276 Briggs, S. D., Xiao, T., Sun, Z.-W., Caldwell, J. A., Shabanowitz, J., Hunt, D. F., Allis, C. D. & Strahl, B. D. Gene silencing: Trans-histone regulatory pathway in chromatin. *Nature* **2002**, 418, 498.
- 277 Kruidenier, L., Chung, C.-w., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., Eberhard, D., Hutchinson, S., Jones, E., Katso, R., Leveridge, M., Mander, P. K., Mosley, J., Ramirez-Molina, C., Rowland, P., Schofield, C. J., Sheppard, R. J., Smith, J. E., Swales, C., Tanner, R., Thomas, P., Tumber, A., Drewes, G., Oppermann, U., Patel, D. J., Lee, K. & Wilson, D. M. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **2012**, 488, 404.
- 278 Flaus, A. Principles and practice of nucleosome positioning in vitro. *Frontiers in Life Science* **2011**, 5, 5.
- 279 Nakayama, J.-i., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. S. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **2001**, 292, 110.
- 280 Zhang, X., Wen, H. & Shi, X. Lysine methylation: beyond histones. *Acta Biochim Biophys Sin* **2012**, 44, 14.
- 281 Hughes, R. M., Wiggins, K. R., Khorasanizadeh, S. & Waters, M. L. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 11184.

- 282 Shi, X. B., Kachirskaia, L., Yamaguchi, H., West, L. E., Wen, H., Wang, E. W., Dutta, S., Appella, E. & Gozani, O. Modulation of p53 function by SET8-mediated methylation at lysine 382. *Molecular Cell* **2007**, 27, 636.
- 283 Kubicek, S., Gilbert, J. C., Fomina-Yadlin, D., Gitlin, A. D., Yuan, Y., Wagner, F. F., Holson, E. B., Luo, T., Lewis, T. A., Taylor, B., Gupta, S., Shamji, A. F., Wagner, B. K., Clemons, P. A. & Schreiber, S. L. Chromatin-targeting small molecules cause class-specific transcriptional changes in pancreatic endocrine cells. *Proc. Natl. Acad. Sci. U.S.A.* **2012**.
- 284 Rypniewski, W. R., Holden, H. M. & Rayment, I. Structural consequences of reductive methylation of lysine residues in hen egg white lysozyme: An X-ray analysis at 1.8 Angstrom resolution. *Biochemistry* **1993**, 32, 9851.
- 285 Black, S. D. & Mould, D. R. Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal. Biochem.* **1991**, 193, 72.
- 286 Kim, Y. C., Quartey, P., Li, H., Volkart, L., Hatzos, C., Chang, C., Nocek, B., Cuff, M., Osipiuk, J., Tan, K. M., Fan, Y., Bigelow, L., Maltseva, N., Wu, R. Y., Borovilos, M., Duggan, E., Zhou, M., Binkowski, T. A., Zhang, R. G. & Joachimiak, A. Large-scale evaluation of protein reductive methylation for improving protein crystallization. *Nature Methods* **2008**, 5, 853.
- 287 Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **2010**, 43, 186.
- 288 Kabsch, W. XDS. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010**, 66, 125.
- 289 Dougherty, D. A. The cation- π interaction. *Acc. Chem. Res.* **2013**, 46, 885.
- 290 Tatko, C. D. & Waters, M. L. Comparison of C-H \cdots π and hydrophobic interactions in a β -hairpin peptide: Impact on stability and specificity. *J. Am. Chem. Soc.* **2004**, 126, 2028.
- 291 Mecca, T., Consoli, G. M. L., Geraci, C. & Cunsolo, F. Designed calix[8]arene-based ligands for selective tryptase surface recognition. *Biorg. Med. Chem.* **2004**, 12, 5057.
- 292 Capila, I. & Linhardt, R. J. Heparin-protein interactions. *Angew. Chem. Int. Ed.* **2002**, 41, 391.
- 293 Ganesh, V. K., Muthuvel, S. K., Smith, S. A., Kotwal, G. J. & Murthy, K. H. M. Structural basis for antagonism by suramin of heparin binding to vaccinia complement protein. *Biochemistry* **2005**, 44, 10757.

Bibliography

- 294 Taniguchi, S., Suzuki, N., Masuda, M., Hisanaga, S.-i., Iwatsubo, T., Goedert, M. & Hasegawa, M. Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins. *J. Biol. Chem.* **2005**, 280, 7614.
- 295 Allen, H. F., Daze, K. D., Shimbo, T., Lai, A., Musselman, C. A., Sims, J. K., Wade, P. A., Hof, F. & Kutateladze, T. G. Inhibition of histone binding by supramolecular hosts. *Biochem. J.* **2014**.

Since joining the Crowley lab, my Ph.D. experience has been productive and stimulating and would not have been possible without the support and guidance of many people. Therefore, I wish to thank the various people for their valuable contributions.

Firstly, I would like to thank **Dr Fawaz Aldabbagh** for making undergraduate chemistry exciting for me. Undoubtedly, Fawaz's lectures made an early positive impression on me and this sparked my first genuine interest in chemistry.

I express superior gratitude to my supervisor **Dr. Peter Crowley**. It has been an honour to be his first Ph.D. student. Peter's expert knowledge and research experience have been instrumental for the work involved in this thesis and also for the general development of my research skills. Moreover, the enthusiasm he has for research was contagious and motivational for me.

The National University of Ireland, Galway, **College of Science** fellowship funded my Ph.D, for which I am very grateful.

I am very grateful to the **Crowley group members** who provided useful discussions on my work.

Dr. Nicholas P. Power synthesized the calixarene, for which I am thankful.

I am grateful to the myriad of contributions made by all the staff and students at the School of Chemistry. The assistance of **Dr. Róisín Doohan**, **Marian Vignoles** and **Ger Fahy** for their help with various analytical methods is greatly appreciated.

I am also grateful to the staff and students in Department of Microbiology for the use of their resources. These are **Dr. Ger Wall** who allowed me for the extensive use of the microscope, **Malgorzata Wronska** who provided useful discussions on crystallization and **Michael Coughlan** for his advice and encouragement.

I am also extremely grateful to **Prof. Mike Hynes** and **Prof. Gary Pielak** for their words of wisdom and for encouraging and helping me to pursue a career in research.

I am very grateful to **Dr. Amir Khan** and **Prof. Martin Caffrey** from Trinity College Dublin provided their lab facilities for harvesting crystals. Furthermore, Amir allowed me to use of his lab space and equipment for experiments on several occasions in addition to assisting with crystal harvesting and data collection.

Dr. Andrew McCarthy, Dr. Humberto Fernandez, Dr. Valerie Pye and **Dr. Joseph Lyons** assisted with crystal harvesting and/or data collection and for this I am very appreciative. I am particularly thankful to Valerie for taking the time to show me the art of harvesting crystals and for her valuable discussions on crystallography.

I am grateful to **Prof. Fred Lisdat** and **Dr. Sven Feifel** from the Technische Hochschule Wildau. I am thankful for their time, guidance and efforts to make the collaboration on crystal biosensors a success. **COST Action TD1003** and the **Irish Research Council ‘New Foundations’ scheme** are gratefully acknowledged for the financial support, which allowed me to work in the Lisdat lab on three occasions.

Finally, I would like to thank to my *family* and *friends* for the endless patience and encouragement they have given me during this study. Their support enriched my Ph.D. experience and I am extremely grateful to them all, especially for celebrating every great moment with me.

Róise Ella McGovern was born in Galway, Ireland. In 2004 Róise enrolled in a Bachelor of Science degree at the National University of Ireland, Galway. During this period Róise worked as an inorganic and physical chemistry research scientist in Haldor Topsoe, Catalyst and Technology Company, in Demark (29.05.07-24.08.07). In 2008 Róise graduated with a 1st class honours chemistry degree. Following this, Róise worked as an analytical scientist in the development department of Wyeth Biotech, Grange Castle, Dublin (21.07.08-21.12.09). This involved the characterisation and analysis of recombinant glycoproteins, vaccines and small molecules in addition to method development. In 2010 Róise joined the Dr. Peter Crowley's lab at the NUI Galway to pursue a Ph.D. This work focused on the characterisation of protein-calixarene interactions and self-assembly. The results of this project are presented in this thesis.

Summary of activities

Collaborative work

Worked in Prof. Fred Lisdat's lab in Technische Hochschule Wildau, Berlin on three occasions (September 2013, December 2012 and December 2011).

Synchrotron trips

Data collected by X-ray crystallography at European Synchrotron Radiation Facility, Grenoble on two occasions (September 2012 and August 2011).

Oral presentations

Joint First Prize for BOC Researcher of the Year, School of Chemistry, National University of Ireland, Galway (April 2013).

First prize at the Roche Research in Life Sciences and Bioengineering, National University of Ireland, Galway (November 2012).

RSC Bioorganic Group Postgraduate Symposium in Cardiff University (April 2012).

Eli Lilly Postgraduate Chemistry Symposium, National University of Ireland, Galway (September 2011).

Chemistry Research Colloquium, University College Dublin (June 2011).

Poster presentations

Euromar, University College Dublin (May 2012).

International School on Biological Crystallization, Granada (May 2011).

Conferences

ChemComm-RSC Prizes and Awards Symposium, Trinity College Dublin (May 2013).

Inaugural Meeting of the Irish Crystallographic Association, Dublin (September 2012).

International Conference on the Crystallization of Biological Macromolecules, Trinity College Dublin (August 2010).

Training courses

BCA/CCP4 Summer School in Protein Crystallography, Oxfordshire (August 2012).

International School on Biological Crystallization, Granada. (May 2011).

List of publications

McGovern, R. E., Fernandes, H., Khan, A. R., Power, N. P. & Crowley, P. B. Protein camouflage in cytochrome *c*-calixarene complexes. *Nature Chem.* **2012**, 4, 527.

McGovern, R. E., Lyons, J.A., McCarthy, A. A., & Crowley, P. B. Calixarene-Mediated Protein Assembly and Selectivity of Cationic Side Chains. *Submitted.*

McGovern, R. E., Feifel, S.C., Lisdat, F., & Crowley, P. B. Electron Transfer in Cytochrome *c*-calixarene Crystals Characterised by Cyclic Voltammetry. *Manuscript in preparation.*

