STRUCTURE-BASED DRUG DESIGN

will crystallisation 'chaperones' make all proteins crystallisable?

Macromolecular x-ray crystallography is usually an integral part of small molecule drug discovery for soluble targets. Recent advances in the crystallography of membrane proteins, such as ion channels and G-protein-coupled receptors, have opened up the prospect of extending structure-based methods into these well validated target classes. The use of chaperone-assisted crystallisation has been a key feature of this success. This article explores the growing repertoire of crystallisation chaperones, examples of their success and how they may enhance the general structural tractability of many protein targets.

By Dr Chun-wa Chung

Protein crystallography has become an invaluable tool embedded within small molecule drug discovery programmes. It often provides important insights into the mode of action of the compounds and stimulates ideas on how their affinities, specificities and physical chemical properties may be optimised to make a drug-like molecule. To determine x-ray complexes of a target protein with bound ligands it is necessary to fulfil a number of steps that can be simplistically summarised as protein production, protein crystallisation and x-ray structure determination. Advances in recombinant protein technologies and automation, coupled with wide access to intense in-house and external x-ray sources and improved computational methods, now allow rapid structure determination on a scale and speed previously unimaginable. In ideal cases, it has even been possible to use x-ray crystallography to screen small fragment libraries for lead generation efforts producing hundreds, if not thousands of datasets and structures.

In reality, there are many factors that limit the number of targets for which this ideal scenario can be achieved. Attrition can occur at every step on the road to establishing a robust crystal system but the acknowledged bottleneck is protein crystallisation. Proteins are irregular objects where flexibility and motion can be key to their physiological function. Packing these complex molecules together to give a regular three dimensional lattice is difficult. Therefore, it is not surprising that finding a crystallisation regime that supports crystal nucleation and growth remains an empirical, trial and error process requiring the screening of many crystallisation conditions. For structure-guided drug design, details of the protein-ligand and solvent interactions should be well resolved. This requires the packing uniformity within the crystal to be nearly perfect. To realise this criterion, techniques to rigidify the proteins and to make them conformationally as well as covalently homogeneous are widely employed. Small molecule ligands or native

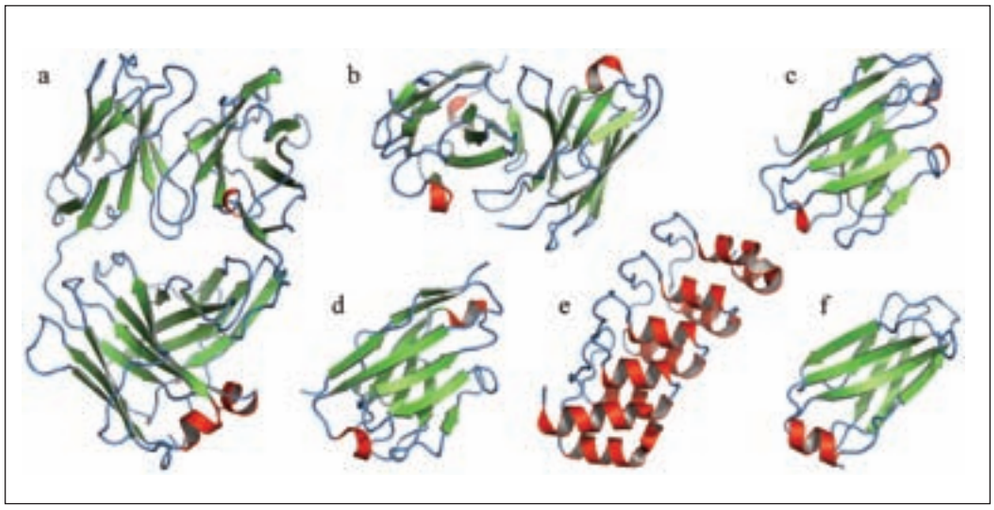

Figure 1
Representative structures of crystallisation chaperones
1a: Fab antibody from complex with β2AR (2r4r)
1b: Fv antibody fragment from complex with bacterial cytochrome c oxidase (1bgy)
1c:VHH nanobody from complex with EspI: EpsJ (3cfi)
1d: dAb chaperone from complex with MDM4 (2vyr)
1e: DARPIN chaperone from complex with AcrB (2j8s)
1f: FN3 chaperone from complex with ERalpha ligand binding domain (2ocf)

protein partners can be added to freeze the protein into a unique conformational state. Flexible regions can be removed and mutations made to encourage better packing interactions, for example by reducing the surface entropy, or by altering the chemical nature of the surface. These well-established 'tricks of the trade' have increased the proportion of soluble targets accessible to structure-based drug design, but there is still much room for improvement. The recent successful use of crystallisation chaperones for integral membrane proteins (IMPs) may highlight one possible avenue to do this.

Chaperone-assisted crystallography is the term that has been coined to describe the use of an auxiliary protein that binds to the target of interest thereby enhancing its ability to form well-ordered crystals[1]. This concept is not new and was first successfully applied over a decade ago to the structure solution of bacterial membrane cytochrome c oxidase using an antibody Fv fragment. At the time, its promise was compromised by the expense, time and effort necessary to identify and produce a suitable antibody, as this could only be achieved by animal immunisation and hybridoma technologies. The growth in biopharmaceutical research has fuelled advances that make the generation of antibody-like molecules and a growing number of synthetic binders much more accessible. A by-product of this effort is that these same technologies can be adopted to find crystallisation chaperones. At present the methods are not routinely available in crystallography laboratories. However, if the promise revealed by the successes with membrane proteins continues, it is conceivable that chaperones would not only be used for recalcitrant soluble proteins and previously intractable membrane proteins but will become regarded as another standard 'trick of the trade'.

Chaperone-assisted crystallography

Crystallisation chaperones help proteins to crystallise in two ways. Firstly, they lock an otherwise flexible target into a single conformation which packs more regularly and reduces the entropy lost on crystallisation. Secondly, they provide additional surfaces to form productive interactions, while masking less productive regions on the target. Normally the auxiliary protein forms a non-covalent complex with the target, but covalently fused chaperones have also been tried. In a biological context, proteins often exist as part of large protein assemblies, so the ideal chaperone for a given protein may be its native protein partner. For broader utility, a number of generic classes of crystallisation chaperones have been suggested. This growing repertoire includes: native monoclonal antibodies (mAb) and their fragments (Fv, Fab), single chain antibodies (sFv), human domain antibodies (dAb), camelid nanobodies (VHH), DARPINs and Fibronectin domains (**Figure 1**).

Many of these protein classes are being developed in their own right as potential protein therapeutics that neutralise target antigens for disease modification. This confers a number of convenient attributes on their generation and properties that make them well suited as crystallisation tools. Firstly, the chaperones have stable and relatively constant core structural elements that allow the display of hypervariable features that can be tuned to recognise a host of protein targets. This characteristic is exemplified by natural antibodies, where

the framework of constant and variable immunoglobin domains support the six hypervariable complementarity determining regions (CDRs) which recognise diverse antigens. Secondly, it is possible to generate libraries of these chaperones and pan for binders to the target for crystallisation. This can be accomplished using *in vivo* inoculation (eg mAb), *in vitro* selection methods such as ribosome display (eg DARPINs) or phage display (eg dAb, sFv), or a combination of both (eg VHH). Finally, as the production of therapeutics must be scalable and cost-effective, the generation of large amounts of highly pure and soluble chaperone protein for co-crystallisation studies is often routine.

Antibody chaperones

The most studied class of chaperone is derived from natural antibodies. Intact monoclonal antibodies are multi-domain molecules not ideally suited for crystallisation due to their flexible linker regions. Smaller Fab or Fv fragments retain the antigen interaction domains and are better matched to aid crystallisation. This type of chaperone has been used in the crystallisation of G-protein-coupled receptors (GPCRs)[2].

GPCRs form a large family of structurally related transmembrane proteins. Their role is to sense the presence of a chemically diverse array of physiological ligands within the extracellular environment and to transmit this information to the interior of a cell. This function requires them to be highly dynamic proteins able to adopt multiple distinct conformations depending on the bound ligand and the associated signalling partner, eg G-proteins and arrestins. Approximately 40-50% of marketed drugs target this family of proteins, so they represent a very important area for pharmaceutical intervention. Until 2008, the only GPCR for which high resolution structural information was available was rhodopsin. This protein is unusual in its high natural abundance and packing density, as a two-dimensional array within the eye. Other GPCRs have been less amenable to crystallography, with crystallisation proving a key stumbling point even when sufficient protein is available.

Within the last two years the inactive state of three GPCRs have been captured by crystallography (β2AR, β1AR, and A2A). The two human β2AR structures have both used crystallisation chaperones that target a flexible intercellular loop (ICL3) involved in G-protein coupling. These chaperones confine the GPCR into a single inactive state, helped by the addition of a small molecule inverse agonist, and form most of the contacts that hold the crystal together. Integral membrane proteins often struggle to build up a sufficiently extensive network of interactions to allow crystal growth, as much of the protein surface is masked by detergent and lipid molecules that mimic the membrane raft.

A question often posed about chaperone protein structures is whether this 'unnatural' binder distorts the target protein, forcing it into an unrepresentative conformation. So far, the evidence suggests that it is more energetically likely that the target protein is captured in a relatively abundant natural state and there have been no reports of significant distortions. However, since one of the roles of a chaperone is to restrain the conformational flexibility of its target, the structure obtained cannot represent the entire ensemble of equilibrium states normally sampled. This was clearly demonstrated by a recent computational docking study using the inactive structure of β2AR. The success of this experiment was that six new ligands of affinities from 4uM to 9nM could be identified, illustrating the value of experimental x-ray structures even for a well studied system. The limitation was that all of these ligands were shown to be inverse agonists, just like the compound used to stabilise the protein for crystallisation. As compounds of diverse pharmacology have proven useful drugs for GPCRs, it will be necessary to capture structural snapshots for a variety of conformational states before structure-based drug design (SBDD) will be broadly applicable to this target class. Even then it is likely that the full complexities of these targets will not be revealed by crystallography alone, but by the coherent application of biophysical, pharmacological and structural characterisation.

Single domain antibody chaperones (camelid VHH and dAb)

Minimal antigen-binding domains of camelid (VHH) or human origin (dAb) are alternative crystallisation chaperones to conventional antibody fragments. These are small, approximately 110-140 residue, variable domains based on the heavy (VH) or light (VL) chain of an antibody. Investigators using camelid VHHs, or nanobodies, require *in vivo* immunisation to obtain a suitable cDNA library for *in vitro* selection via phage display, in contrast to dAbs which are selected from large recombinant libraries without the need for animal inoculation.

The application of *in vitro* display technologies easily allows a panel rather than a single binder to be selected and produced. This increases the likelihood and rate of success in the unpredictable area of protein crystallisation. An illustration of this

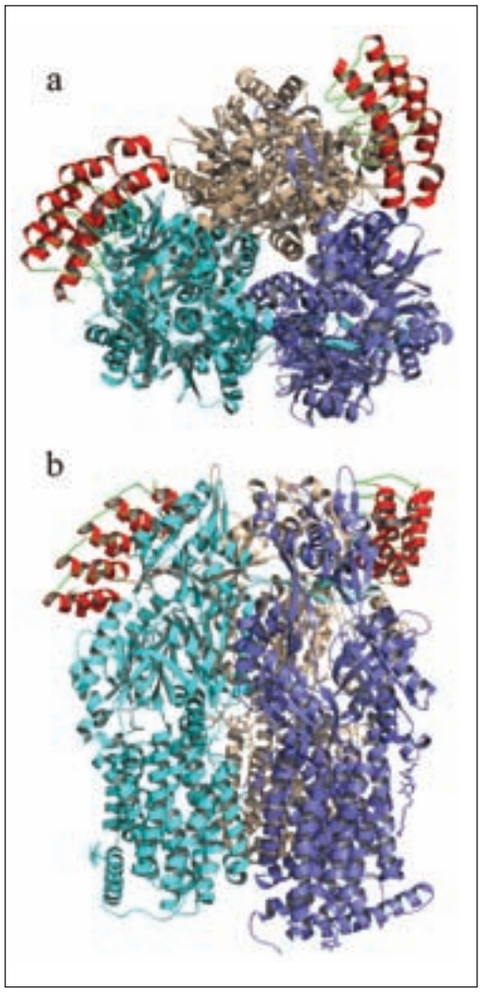

Figure 2
Asymmetric structure of trimeric AcrB exporter with two DARPINs (2j8s at 2.54Å) DARPINs coloured in red and green according to secondary structure.
AcrB subunits coloured in wheat, cyan and violet.
2a: View looking down into channel
2b: View at 90° to View 2a

comes from the structure determination of the EpsI:EpsJ pseudophilin heterodimer by conventional and chaperone approaches[3]. This dimer is a periplasmic component of the type 2 secretin system (T2SS) from *Vibrio vulnifius*, which is related to cholera. The causative agent in cholera, the ~85kDa cholera toxin, is translocated as a folded protein into the lumen of the gut by T2SS, so blocking this excretion machinery would be highly beneficial. The EpsI:EpsJ complex has been successfully solved using a combination of the standard optimisation approaches, specifically surface mutation, N and C-terminal truncation, and the use of species variation. To overcome the crystallisation difficulties encountered with the wild type EpsI:EpsJ required a total of 17 protein variants and took 10 months. In contrast, initial crystal screening of complexes of wild-type N-terminal truncated EpsI:EspJ with four VHH nanobodies gave crystals of sufficient quality to determine the x-ray structure without optimisation. This shortened the crystal optimisation period to the 15 days required for the initial crystals to grow. Not all systems will yield success so readily and this timing does not include nanobody selection and production, but this example provide good reasons to add chaperones to our box of tools for soluble as well as membrane protein crystallisation.

A more convenient and maybe quicker way to obtain chaperones is to use large and diverse recombinant libraries rather than relying on animal immunisation. This method avoids issues with targets that do not elicit a strong immune response. These problematic targets include: pathogenic proteins evolutionarily designed to evade the immune system; native, or close to native proteins; or antigens that may not survive long after inoculation, such as weakly associated complexes or proteins that are intrinsically or metabolically unstable. For these reasons an *in vitro* display method can provide greater opportunities to tailor the selection conditions for an individual protein and yield binders where *in vivo* methods fail. The structure determination of a human domain antibody (dAb) with the N-terminal domain of MDM4, an important negative regulator of tumour activity, demonstrates this point[4]. Isolated MDM4 readily aggregates but is stabilised by the presence of its protein partner p53. Screening for dAb binders to MDM4 was therefore performed *in vitro* in the presence of a 120nM p53 peptide, yielding the VH9 antibody. Attempts to crystallise a complex of MDM4, p53 peptide and VH9 resulted in a complex composed only of MDM4 and VH9, as the VH9 was found to bind in the p53 groove.

One concern over the use of such small chaperones is whether they have sufficient bulk to significantly improve the crystallisation potential of a large target. Studies on protein-protein interactions suggest that exposed β-strands have a natural propensity to form extended sheet structures and self assemble. Therefore the β-sheet nature of domain antibodies may render them 'chaperone efficient', acting co-operatively both in promoting crystal nucleation and growth. The VHH complexes determined to date, including the six structures of VHH-RNaseA, certainly lend support to this idea, as intermolecular β-sheet formation between chaperone molecules is found within the crystal lattice of many of these structures.

Non-antibody chaperones

Crystallisation chaperones need not be restricted to immunoglobin domains. *In vitro* technologies allow a variety of non-antibody based scaffolds to

References

1 Hunte et al. Crystallisation of membrane proteins mediated by antibody fragments. Current Opinion in Structural Biology, (2002) 12: 503-508.

2 Kobilka et al. New G-protein-coupled receptor crystal structures: insights and limitations. Trends in Pharmacological Sciences (2008) 29, 79-83.

3 Lam et al. Nanobody-aided structure determination of the EpsI;EpsJ pseudophilin heterodimer from Vibrio vulnificus. J. Structural Biology, (2009) 166: 8-15.

4 Yu et al. Structure of Human MDM4 N-Terminal Domain Bound to a Single-Domain Antibody. J. Mol. Biol.,(2009) 385:1578-1589.

5 Gaby et al. Chaperone-Assisted Crystallography with DARPins. Structure (2008) 16: 1443-1453.

6 Koide et al. High-affinity single-domain binding G-proteins with a binary-code interface. Proc. Nat. Acad. Sci. USA (2007) 104:6632-6637.

7 Tereshko et al. Towards chaperone-assisted crystallography: Protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. Protein Science (2008) 17: 1175-1187.

be created that increase the structural diversity of the binder. Designed ankyrin repeat proteins (DARPINs) are binders based on the natural ankyrin repeat protein fold with randomised surface residues allowing specific binding to a large variety of proteins. They are the only helical proteins routinely used as chaperones. At least five co-crystal structures of DARPINs with a variety of targets (kinase, protease and membrane protein) have been solved, including the *E. coli* multidrug exporter AcrB, a drug resistance target[5]. Isolated AcrB could be crystallised due to its large periplasmic domain, but the DARPIN complex gave a significantly higher resolution structure. The export mechanism of AcrB involves an asymmetric trimer, which is highlighted by the binding of DARPINs to only two of the possible three subunits (**Figure 2**).

Another potential non-antibody chaperone is the fibronectin type III (FN3) module. This has a β-fold similar to that of an antibody domain, but lacks the presence of disulphide bonds. Fusion constructs concatenating the FN3 with its crystallisation target can therefore be produced regardless of the redox environment. Using a fusion protein has the advantage of increasing the local concentration of the binder and ensuring a 1:1 stoichiometry. It also eliminates the need to generate and purify the chaperone separately. Crystal structures of proteins with FN3 have been reported and include both covalent and non-covalent chaperone complexes with maltose-binding protein (MBP), estrogen receptor ligand binding domain and SH2 domain of Abl1[6].

Recombinant crystallisation chaperones as phasing tools

In addition to acting as a crystallisation aid, the presence of a well-structured, easily expressed and recombinant chaperone provides an opportunity to use this molecule to provide phasing information to solve the phase problem in difficult structures[7]. The phase problem is a well known issue with x-ray structure determination. When x-rays strike a crystal lattice, information about the nature of the lattice and molecular contents are encoded in the scattering or diffraction pattern. To decipher the code, the spatial distribution, intensity and relative phase of the diffracted waves must be known. Unfortunately, x-ray detectors only record an intensity pattern and the lost phase information results in the 'phase' problem. Increasingly powerful computational methods together with a larger database of homologous structures allow us to overcome this issue to some extent, by providing a good starting model of the true structure that can be iteratively optimised. However, there remains a need for experimental methods such as single-wavelength and multi-wavelength anomalous dispersion (SAD, MAD) which give an unbiased starting point. In a proof of concept study, a methionine enriched VHH binder was readily labelled with selenomethionine (SeMet) in *E.coli*, complexed with ribonuclease A and used for SAD experiments. The selenium atoms in the chaperone provided ample phasing power to solve the entire complex structure by SAD without recourse to SeMet labelling of the target itself. This is an especially attractive proposition for difficult to express or non-recombinant proteins where introduction of SeMet may not be possible.

Conclusion

Will crystallisation chaperones make all proteins crystallisable? Clearly, the answer must be no, there is no panacea to this technical bottleneck. A more appropriate question may be: have advances in technology progressed sufficiently that these chaperones can be used as part of our routine crystallographic tool box? For the most challenging proteins, such as integral membrane proteins, the need to exploit all the 'tricks of the trade' will drive the use of chaperones to help tackle these targets. For other proteins, they will probably be part of the tool box, but on the second tier, to be accessed once more standard tools have failed, as these methodologies are not widely available in most laboratories today. The introduction of commercial crystallisation chaperone generation and selection kits may change this balance and enable this methodology to fulfil its full potential.

Acknowledgments

Thanks to the many colleagues that kindly took time to read and make useful comments on this article. DDW

Dr Chun-wa Chung *leads a Biophysics and Structural Analysis team at GlaxoSmithKline R&D in Stevenage, UK. She joined Glaxo as a post-doc after completing a PhD in NMR methods development and a BA in Natural Sciences (Chemistry) at Cambridge University. Her current position combines the two major interests of her career at Glaxo, GlaxoWellcome and now GlaxoSmithKline, namely biophysical methods for characterising protein-ligand interactions and mode of action studies and structural activities in both the x-ray and NMR arena.*