CHAPTER 12

HYBRID METHODS FOR PROTEIN STRUCTURE PREDICTION

DMITRI MOURADOV and BOSTJAN KOBE
The University of Queensland
School of Chemistry and Molecular Biosciences
QLD, Australia

NICHOLAS E. DIXON
School of Chemistry
University of Wollongong
NSW, Australia

THOMAS HUBER
The University of Queensland
School of Chemistry and Molecular Biosciences
QLD, Australia

12.1. INTRODUCTION

Structural bioinformatics is a highly cost-efficient solution for accelerated determination of the three-dimensional (3D) structures of proteins. Purely computational prediction methods, such as advanced fold recognition (Chapters 9 and 10), composite approaches (Chapter 12), *ab initio* fragment assembly [1,2], and molecular docking [3] are routinely applied today to extend our knowledge of protein structures, how they interact and what their functional roles are in a biological context. Very often, however, predicted protein structures are not given the same trust as their experimental counterparts. This comes mainly from the need for extensive expertise to produce high-quality models, generally high rates of false predictions, and the difficulty to measure the confidence that can be associated with a structure "solved" algorithmically.

Introduction to Protein Structure Prediction: Methods and Algorithms,
Edited by Huzefa Rangwala and George Karypis

Hybrid approaches are a means to overcome these shortcomings; by incorporating limited experimental measurements, reliable structural models can be computed and unlikely predictions eliminated. Hybrid approaches take advantage of data derived from a range of very different biochemical and biophysical methods, most of which are becoming routinely available in many laboratories. These methods are of increasing interest in view of increasingly easy access to analytical instruments, such as high-resolution mass spectrometers and high-frequency electron paramagnetic resonance (EPR) spectrometers. Similarly, small-angle neutron scattering and small-angle X-ray scattering (SANS/SAXS) data are becoming routinely accessible through advanced neutron and synchrotron light sources. In addition, recent developments in nuclear magnetic resonance (NMR) spectroscopy make large (>100 kDa) protein systems amenable to analysis and, in combination with site-specific isotope labeling, have opened unprecedented possibilities to obtain sparse structural data on selected regions within an entire system.

Moreover, hybrid approaches have shown great promise in complementing high-resolution structural biology. To fully characterize function in dynamically interacting assemblies where both the components and their structures may vary throughout a complex multistep process, structures need to be determined at each step. By using model structures, it is possible to design and analyze new hypothesis-driven experiments and thus significantly speed up high-resolution structure determination.

12.2. SOURCES OF LIMITED STRUCTURAL DATA

A variety of biophysical and biochemical techniques exist that can rapidly give a wealth of information on shape, local structure, residue proximities, and residue environment in macromolecular systems (Table 12.1). These include in-solution scattering measurements where the angular distribution of SAXS and SANS can be fit to yield global information on the structural envelope of a protein in solution. The reliability of SAXS data and computational analysis tools has recently improved dramatically [4,5]. SANS has the added advantage that contrast matching of small-angle and buffers selectively renders parts of the system invisible (to neutrons), and shapes of proteins in a larger assembly can be determined individually [5]. A complementary way to map protein surfaces is by chemical modification (CM) or hydrogen/deuterium (H/D) exchange (DX). After various times of exposure to D_2O or a CM agent that targets side chain functional groups, the protein system is digested with proteases and liquid chromatography-mass spectrometry (LC-MS)/MS fingerprinting is then used to determine where and to what degree CM has occurred. CM and H/D exchange of amide protons is generally several orders of magnitude slower in residues that are buried and/or part of regular secondary structure, thus providing a quantitative measurement of protein structure. Circular dichroism (CD) spectroscopy constitutes a reliable method for mea-

TABLE 12.1 Examples of Methods That Produce Sparse Structural Data and Have Been used in Combination with Molecular Modeling to Compute Structure

Method	Data measured	Structural data generated	Example application
SAXS/SANS	Scattering intensity as a function of momentum transfer	Pair distribution function; shape envelope	[21,22]
CD	Mean residue ellipticity as a function of wavelength	Secondary structure content	[23]
FRET	Yield of fluorescence energy transfer	Distance between donor and acceptor pair	[24]
EPR	Dipole-dipole coupling between electron spins	Spin label environment and distance between pairs of spin labels	[25]
Deuterium exchange-mass spectrometry (DXMS)	Rate constant of H/D exchange	Solvent exposure	[26]
Radical footprinting	Rate constant from dose-response curve	Solvent exposure	[27]
Chemical cross-linking	Mass/charge ratio of joint peptides and fragmentation	Upper limit on pair distance between reacted groups	[28,29]

suring the secondary structure content of a protein. But while secondary structure content information by itself is generally of very limited use for computing structure, in combination with residue by residue secondary structure prediction it can provide important local structure restraints.

Arguably the most powerful constraints in protein structure modeling are measured distances between pairs of residues. Molecular probe techniques, such as fluorescence resonance energy transfer (FRET) [6] and EPR [7], can provide selective distances between specifically labeled parts of a molecule. Paramagnetic or fluorescent labels are generally attached either *via* disulfides or other CM to engineered cysteine mutants, or as more recently demonstrated, can be selectively and efficiently incorporated in the form of non-canonical amino acids into proteins using a cell-free expression system [8]. A clear advantage of these probe techniques is their ability to provide long-range structural information. Among all the lanthanide ions, Gd^{3+} is most highly paramagnetic, and recent work has shown that spin-spin interactions between pairs of Gd^{3+} ions can be used in high-field pulse EPR experiments to accurately determine distances up to 40Å [9]. Similarly, new developments in

NMR spectroscopy employs paramagnetic ions to induce the anisotropic pseudo contact shift (PCS) in NMR active nuclei of a protein molecule [10]. Because of the anisotropy in the PCS, the electron-proton interaction decaying slowly and the gyromagnetic ratio of a free electron being nearly three orders of magnitude larger than that of a proton, the PCS effect provides ample long-range structural information, reporting on interactions over distances up to 40Å.

The chemical analog of these biophysical methods to measure inter-residue distances is chemical cross-linking [11,12], which attempts to covalently connect functional groups with a molecular spacer. A link can only be formed if the functional groups are within the reach of the spacer, thus providing a chemical means to measure (upper bound) distances between groups. This conceptually simple chemical approach has a long history in protein science where it proved to be a useful tool to determine intermolecular interactions [10] and topological information in multi-protein assemblies [13]. More recently, chemical cross-linking has attracted much interest in more detailed structural studies, spurred by improved structural modeling capabilities and rapid advances in high-accuracy multi-dimensional MS, which today allows not only the reliable identification of the chemically modified molecules, but also determination of the exact cross-linker insertion point [14–17]. By combining enzymatic digestion with MS, the cross-linking technique is able to remove size limitations imposed by other techniques as only proteolytic fragments are analyzed. A key advantage of chemical cross-linking over FRET and EPR probe techniques is that multiple distance constraints between several groups within a protein or protein complex can be obtained in a single experiment and no labels need to be incorporated. However, exactly this advantage has also proven to be a major technical challenge, because cross-linked peptides must be identified in an abundance of proteolytically digested native peptides. In the past few years, innovations in cross-linker design and peptide separation methods combined with multi-dimensional MS and new analysis techniques have greatly improved identification [18–20].

12.3. TRANSLATION INTO STRUCTURAL RESTRAINTS

In combination, these approaches yield overlapping and distinct structural information on solvent exposure, inter-amino acid distances, and protein shape. For structure calculations, all this information is gathered in a pseudo-energy function, the global minimum of which corresponds to the structure that best satisfies all restraints:

$$E^{\exp} = \sum_{i}^{N} w_i \left(Q_i^{calc} - Q_i^{\exp} \right)^2 \tag{12.1}$$

where Q_i^{calc} is the i^{th} experimental property calculated from the candidate structure, $Q_i^{\exp}$ is the i^{th} measured property, and N is the total number of

experimental data. Different weightings w_i account for the experimental error in each datum.

Depending on the nature of the measurement, Q_i^{calc} itself is a more or less complicated function with respect to the structure. In the case of chemical cross-linking, for example, it is a Boolean function (functional groups in the structure are either within the cross-linker spanning distance, or they are not), rendering the quadratic term into a simple step function. When using different experiments, it appears to be appealing to first transform all measurements into a corresponding structural metric, such as a distance, and then restrain the structures with respect to this new metric. This should be avoided, because the error in the measurement is not necessarily Gaussian with respect to the new transformed metric and correct weighting of each datum becomes difficult.

12.4. USE OF LIMITED EXPERIMENTAL DATA TO ELUCIDATE STRUCTURE

Molecular models are a common way to represent experimental structural measurements.

The classical approach in high-resolution structure determination is to measure sufficient experimental restraints to define a 3D model by the system's 3N Cartesian coordinates, where N is the total number of atoms in the system. What makes hybrid approaches special is that they require only a limited number of explicit experimental constraints to calculate a structure (Fig. 12.1).

Explicit constraints can be used to simply filter through theoretical models, allowing exclusion of models that do not meet the constraints. Taking this a step further, the measurements can be used to refine or even compute models, with direct application of the constraint data during calculations. When

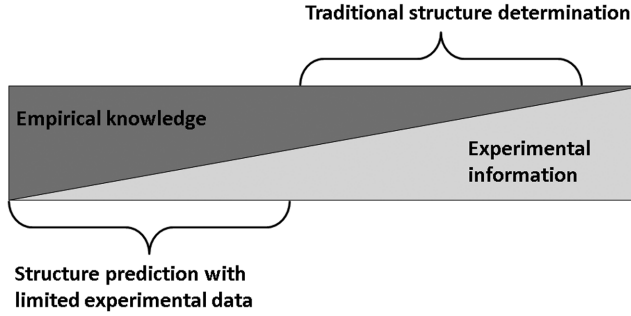

FIGURE 12.1 Limited experimental information is required for structure prediction using hybrid techniques as opposed to traditional techniques for structure determination such as X-ray crystallography and NMR.

combined with molecular modeling/docking [28–31], a limited number of constraints may be sufficient to determine the orientation of components in a protein complex or even of domains in a multi-domain protein.

12.4.1. Re-Scoring Models (Filtering)

At the most basic level, Equation 12.1 can be used to rank a set of theoretical models according to how well they are in agreement with all measurements. Albrecht et al. [32] conducted a theoretical study of how effective distance constraints from chemical cross-linking are at improving the success of fold recognition by threading. The analysis was carried out on 81 single-domain proteins (Hobohm96-25 database) whose pairwise sequence identity does not exceed 25%. Hypothetical cross-linking constraints were generated for all aspartate, glutamate, and lysine residues separated by between 8Å and 12Å based on the known structures of the proteins. These hypothetical cross-links were then used to re-rank theoretical models computed with the 123D threading program [33]. Various ranking functions on validity of distance constraints were applied, including simply counting the number of satisfied constraints and using a more complex scoring function that gives higher scores for satisfied constraints that are conserved among members of the same fold class. The results show that employing sparse constraints from cross-linking studies to re-rank models from fold recognition can improve success rates from 54–65% to 58–73%, depending on the quality of the initial alignment.

A limitation of this study was its use of hypothetical cross-links, equivalent to the outcomes of optimal experiments. Due to large differences in group reactivities, reagent accessibility and competitive suppression of some low-abundance cross-linked peptides in the MS analysis, generally only a small subset of all theoretically possible cross-links can be observed in real experiments. However, studies using data from experimental cross-linking have shown similar improvements to that reported by Albrecht et al. [32]. For example, Young et al. [30] used a similar post-filtering approach on a fibroblast growth factor-2 (FGF-2) protein where cross-links were experimentally identified using MS techniques. Again using the 123D software, the protein was initially incorrectly categorized as belonging to the beta-clip fold family. However, re-ranking of the top threading models by a simple scoring function based on the number of satisfied constraints resulted in the first, second, and fourth ranked structures all correctly identifying the FGF-2 structure as a member of the beta-trefoil family.

12.4.2. Structure Refinement

One inherent limitation of X-ray crystallography is that various proteins, such as membrane channels, may adopt multiple stable conformations that cannot be observed in static crystal structures. In such a case limited experimental constraints can be used to refine an existing crystal structure to show an

alternate stable conformational state. FRET spectroscopy has been used to demonstrate the concept by modeling the conformational change involved in channel gating in MscL, a multimeric membrane protein important in releasing pressure during hypo-osmotic stress [34]. A crystal structure had been solved showing a closed conformation of MscL comprised of five identical subunits surrounding a central pore. Site-directed mutagenesis, specifically the insertion of cysteine residues, was used to insert different fluorescent probes into identical sites in all five subunits. Measuring fractions of energy transferred between donor and acceptor probes before and after induction of channel opening correlated with a radius increase of 8Å. As the protein volume remains constant, channel activation must trigger an opening of a large pore, as inferred also by previous studies using paramagnetic resonance spectroscopy and site-directed spin labeling. This radius change was used to model an open-gate conformation of MscL, showing one of the largest conformational changes recorded by any membrane protein. This approach paves the way for probing conformational changes of membrane proteins *in situ.*

12.4.3. Structure Calculations (Docking with Constraints)

A more challenging problem is to directly compute molecular models that satisfy a given set of (sparse) constraints. Xu et al. [35] detail such an approach that uses NMR Nuclear Overhauser Effect (NOE) data to improve threading performance. They employed a divide and conquer strategy, which divides the structures into substructures (cores) each comprised of only one secondary structure element, then optimally aligns substructures with subsequences. Two conditions must always be met to incorporate the distance constraints: (i) a link must be present for a constraint to be aligned to two cores and (ii) linked cores must not be aligned to sequence positions that violate constraints. Results show that even a small number of NOEs were sufficient to improve threading success in difficult to predict proteins. Even though NOEs provide a medium density network of distance restraints, the same approach can also be applied to the more sparse constraints derived from chemical cross-linking.

12.4.4. Multi-Domain Proteins and Multimeric Assemblies

The success of structural genomics has brought about the systematic determination of many structures of individual proteins and protein domains. This has led frequently to the situation where although individual domains in multidomain proteins are structurally characterized, their relative orientations are not. Similarly, structures of constituents of many multimeric protein assemblies have been solved by X-ray crystallography or NMR spectroscopy, but due to technical difficulties the whole assemblies could not be structurally determined using these techniques. While using purely computational techniques to predict relative orientations of domains and proteins in complexes often results in an inaccurate conformation [36], hybrid approaches have been

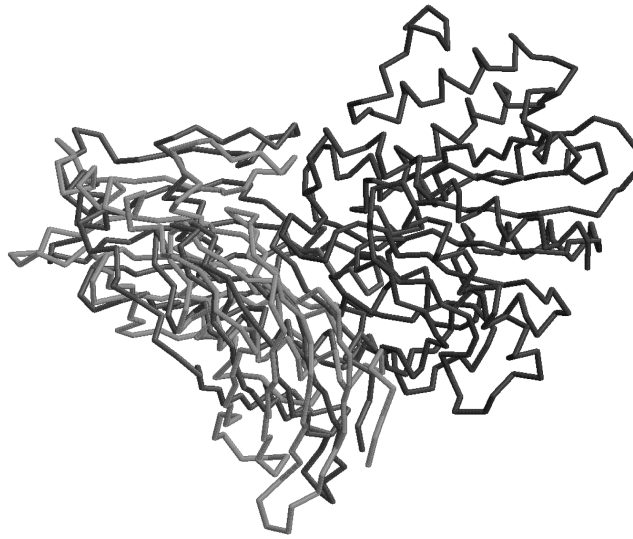

FIGURE 12.2 The crystal structure of the latexin (green)—CPA1 (blue) complex, overlaid with the orientation of the latexin molecule revealed by the top-scoring docked structure (red). (See color insert.)

shown to complement high-resolution structure determination in these cases. In combination with molecular docking, they can provide sufficient additional inter-domain or inter-protein constraints to allow the positioning of components in an overall structure at high resolution.

As an example, we demonstrated this concept by combining molecular modeling with a very small number of experimental distance restraints from chemical cross-linking to establish the structure of the protein complex between a mammalian carboxypeptidase A (CPA1) and its protein inhibitor latexin [28]. Three distance constraints were identified using a combination of chemical cross-linking and MS technology. Rigid body docking with a simple scoring function was employed to calculate the structure of latexin:CPA1 to within a C_α root-mean-square deviation (RMSD) of 3.74Å relative to the crystal structure (Fig. 12.2). The elucidated structure defines the interface between the two molecules accurately enough to guide mutagenesis experiments to probe the contribution of interacting residues and to provide reagents for use to probe the cellular functions of the proteins.

Structure determination of multi-domain proteins using the hybrid approach was assessed using acyl-CoA thioesterase (Acot7), which contains two hotdog fold domains, both of which are required to catalyse the hydrolysis of fatty acyl-CoA into CoA and free fatty acids [29]. While the two hotdog domains were able to be crystallized and their structures solved separately, no diffraction quality crystals were obtained for full-length Acot7. Using two separate techniques, molecular docking and homology modeling, together with identification of seven inter-domain cross-links, the orientation of the two hotdog

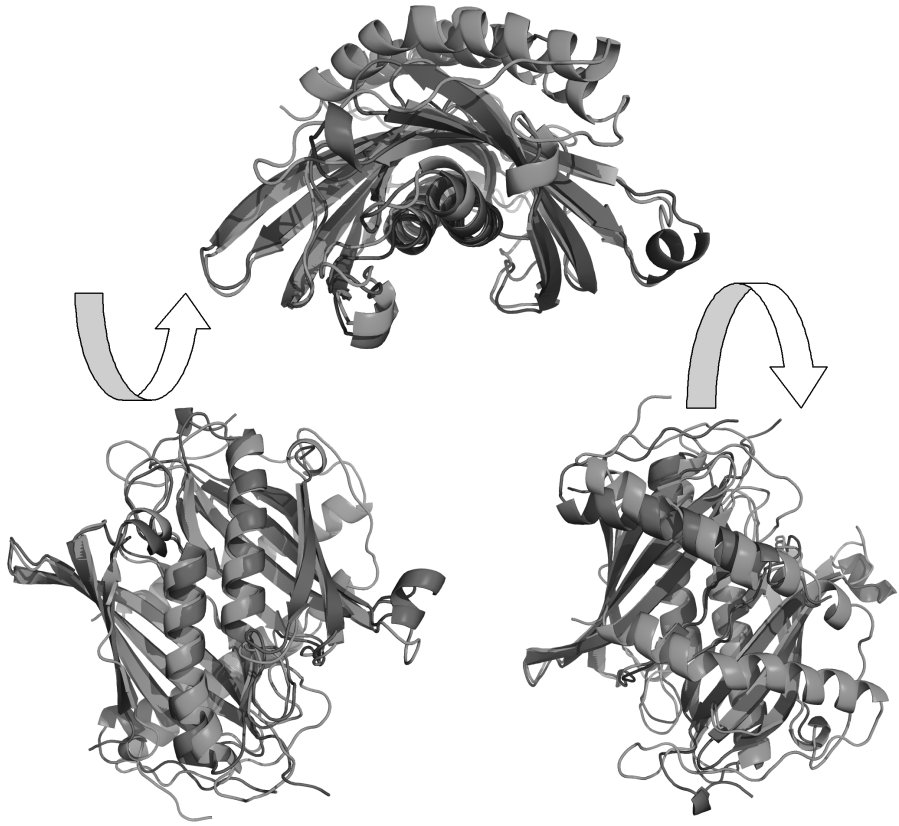

FIGURE 12.3 Multiple views of the superimposition of the predicted Acot7 model using chemical cross-linking and molecular modeling (green) with the crystal structure of ACOT12 (red). The structure suggests that the two hotdog fold domains adopt a similar structure to a dimer of single hotdog domains from other species. (See color insert.)

domains in the full-length Acot7 was predicted. A consistent model of a double-hotdog structure was observed using both docking and modeling. Superimposition of the model onto the a recently solved crystal structure of the human homolog ACOT12 shows remarkable structure similarity with C_α RMSD of 1.6Å over 242 amino acids (Fig. 12.3).

12.5. CONCLUSIONS AND FUTURE OUTLOOK

A clear advantage of hybrid approaches is that they offer means to combine very different sources of structural information and to integrate them iteratively into a single structural model (but not necessarily a single structure). Even when the amount of experimental information is still insufficient to determine a structure, it will already provide structural models that enable

new hypotheses to be tested experimentally, for example, better paramagnetic labeling sites will be identified, and putative interaction surfaces will be identified at which labels for distance measurements can be site-specifically incorporated. This is quite different from structure determination with X-ray crystallography, where a high-resolution structure is almost guaranteed if one is able to grow well-diffracting protein crystals, but is left in despair if this is not the case. While the hybrid technique cannot attain the high-resolution structures observed in X-ray crystallography, it is currently seen as a rescue strategy for protein systems where traditional structure characterization techniques do not provide a solution.

Despite the clear appeal of hybrid approaches and the enthusiasm of its proponents, they have not been widely used. For example, no protein structure has been determined completely *de novo* using this method. This is partly due to various experimental obstacles and the need for combining a diverse range of expertise to derive explicit experimental constraints. However, development of novel low-cost techniques has the potential to quickly change this. Specifically, the incorporation of subsequences of amino acids as fluorophores, EPR labels, and cross-linkable groups has the potential to drastically simplify the process of deriving experimental constraints. When coupled with software capable of automated data analysis, the hybrid techniques can yield rapid and low-cost structural information. As the hybrid approach allows the use of a wide range of sources of structural information, one must consider the strengths, weaknesses, and peculiarities of each method on a case-specific basis to identify which method will yield the most promising experimental constraints.

The hybrid method's strongest point may lie in its potential to determine structures of protein complexes through a bottom-up integration of atomic-detail crystallographic or NMR structures with explicit experimental constraints to obtain missing pieces of information, for example, the relative orientation of the individual components as we demonstrated with the complex of CPA1 and latexin. This approach is not limited to high-affinity complexes. By introducing strong bonds formed on contact it is in principle possible to trap short-lived protein-protein interactions in the cross-linking process. Detailed structural information about such transient complexes would be invaluable for our functional understanding of biological process, since it is currently not directly accessible by any other methods. Similarly this technique is amenable to accurately define global orientation of structural domains in large molecular assemblies, where the structures of individual domains have already been determined, for example, by X-ray crystallography, NMR, or molecular modeling.

ACKNOWLEDGEMENT

This work was supported by an Australian Research Council (ARC) grant to NED and TH. BK is an ARC Federation Fellow and a National Health and

Medical Research Council Honorary Research Fellow, and NED is an ARC Professorial Fellow.

REFERENCES

1. K.T. Simons, C. Kooperberg, E. Huang, and D. Baker. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *Journal of Molecular Biology*, 268:209–225, 1997.
2. Z. Yang, K.A. Adrian, and S. Jeffrey. TASSER: An automated method for the prediction of protein tertiary structures in CASP6. *Proteins: Structure, Function, and Bioinformatics*, 61:91–98, 2005.
3. G.R. Smith and M.J.E. Sternberg. Prediction of protein-protein interactions by docking methods. *Current Opinion in Structural Biology*, 12:28–35, 2002.
4. C.D. Putnam, M. Hammel, G.L. Hura, and J.A. Tainer. X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics*, 40:191–285, 2007.
5. C. Neylon. Small angle neutron and X-ray scattering in structural biology: Recent examples from the literature. *European Biophysics Journal*, 37:531–541, 2008.
6. E.R. Goedken, S.L. Kazmirski, G.D. Bowman, M. O'Donnell, and J. Kuriyan. Mapping the interaction of DNA with the *Escherichia coli* DNA polymerase clamp loader complex. *Nature Structural & Molecular Biology*, 12:183–190, 2005.
7. S. Popp, L. Packschies, N. Radzwill, K.P. Vogel, H.-J. Steinhoff, and J. Reinstein. Structural dynamics of the DnaK-peptide complex. *Journal of Molecular Biology*, 347:1039–1052, 2005.
8. K. Ozawa, M.J. Headlam, D. Mouradov, S.J. Watt, J.L. Beck, K.J. Rodgers, R.T. Dean, T. Huber, G. Otting, and N.E. Dixon. Translational incorporation of L-3,4-dihydroxyphenylalanine into proteins. *FEBS Journal*, 272:3162–3171, 2005.
9. A.M. Raitsimring, C. Gunanathan, A. Potapov, I. Efremenko, J.M.L. Martin, D. Milstein, and D. Goldfarb. Gd3+ complexes as potential spin labels for high field pulsed EPR distance measurements. *Journal of the American Chemical Society*, 129:14138–14139, 2007.
10. M.J. Hunter and M.L. Ludwig. Reaction of imidoesters with proteins and related small molecules. *Journal of the American Chemical Society*, 84:3491–3504, 1962.
11. J.B. Swaney, J. Segrest, P., and J.J. Albers. Use of cross-linking reagents to study lipoprotein structure. *Methods in Enzymology*, 128:613–626, 1986.
12. P. Friedhoff. Mapping protein–protein interactions by bioinformatics and cross-linking. *Analytical and Bioanalytical Chemistry*, 381:78–80, 2005.
13. S.C. Liu and J. Palek. Metabolic dependence of protein arrangement in human erythrocyte membranes. II. Crosslinking of major proteins in ghosts from fresh and ATP-depleted red cells. *Blood*, 54:1117–1130, 1979.
14. K.M. Pearson, L.K. Pannell, and H.M. Fales. Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Communications in Mass Spectrometry*, 16:149–159, 2002.

15. G.H. Dihazi and A. Sinz. Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. *Rapid Communications in Mass Spectrometry*, 17:2005–2014, 2003.
16. G.H. Kruppa, J. Schoeniger, and M.M. Young. A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry. *Rapid Communications in Mass Spectrometry*, 17:155–162, 2003.
17. X.H. Chen, Y.H. Chen, and V.E. Anderson. Protein cross-links: Universal isolation and characterization by isotopic derivatization and electrospray ionization mass spectrometry. *Analytical Biochemistry*, 273:192–203, 1999.
18. G.J. King, A. Jones, B. Kobe, T. Huber, D. Mouradov, D.L. Hume, and I.L. Ross. Identification of disulfide-containing chemical cross links in proteins using MALDI-TOF/TOF-mass spectrometry. *Analytical Chemistry*, 80:5036–5043, 2008.
19. D.R. Muller, P. Schindler, H. Towbin, U. Wirth, H. Voshol, S. Hoving, and M.O. Steinmetz. Isotope tagged cross linking reagents. A new tool in mass spectrometric protein interaction analysis. *Analytical Chemistry*, 73:1927–1934, 2001.
20. J.W. Back, V. Notenboom, L.J. de Koning, A.O. Muijsers, T.K. Sixma, C.G. de Koster, and L.Z. de Jong. Identification of cross-linked peptides for protein interaction studies using mass spectrometry and ^{18}O labeling. *Analytical Chemistry*, 74:4417–4422, 2002.
21. W. Zheng and S. Doniach. Protein structure prediction constrained by solution X-ray scattering data and structural homology identification. *Journal of Molecular Biology*, 316:173–187, 2002.
22. W. Zheng and S. Doniach. Fold recognition aided by constraints from small angle X-ray scattering data. *Protein Engineering, Design and Selection*, 18:209–219, 2005.
23. J. Lees and R. Janes. Combining sequence-based prediction methods and circular dichroism and infrared spectroscopic data to improve protein secondary structure determinations. *BMC Bioinformatics*, 9:24, 2008.
24. G.F. Schröder and H. Grubmüller. FRETsg: Biomolecular structure model building from multiple FRET experiments. *Computer Physics Communications*, 158:150–158, 2004.
25. N. Alexander, A. Al-Mestarihi, M. Bortolus, H. McHaourab, and J. Meiler. De novo high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure*, 16:181–196, 2008.
26. Y. Hamuro, L.L. Burns, J.M. Canaves, R.C. Hoffman, S.S. Taylor, and V.L. Woods. Domain organization of -AKAP2 revealed by enhanced deuterium exchange-mass spectrometry (DXMS). *Journal of Molecular Biology*, 321:703–716, 2002.
27. A.J.K. Kamal and M.R. Chance. Modeling of protein binary complexes using structural mass spectrometry data. *Protein Science*, 17:79–94, 2008.
28. D. Mouradov, A. Craven, J.K. Forwood, J.U. Flanagan, R. Garcia-Castellanos, F. X. Gomis-Ruth, D.A. Hume, J.L. Martin, B. Kobe, and T. Huber. Modelling the structure of latexin-carboxypeptidase A complex based on chemical cross-linking and molecular docking. *Protein Engineering, Design and Selection*, 19:9–16, 2006.
29. J.K. Forwood, A.S. Thakur, G. Guncar, M. Marfori, D. Mouradov, W. Meng, J. Robinson, T. Huber, S. Kellie, J.L. Martin, D.A. Hume, and B. Kobe. Structural basis for recruitment of tandem hotdog domains in acyl-CoA thioesterase 7 and

its role in inflammation. *Proceedings of the National Academy of Sciences U S A*, 104:10382–10387, 2007.

30. M.M. Young, N. Tang, J.C. Hempel, C.M. Oshiro, E.W. Taylor, I.D. Kuntz, B.W. Gibson, and G. Dollinger. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proceedings of the National Academy of Sciences U S A*, 97:5802–5806, 2000.
31. D.M. Schulz, S. Kalkhof, A. Schmidt, C. Ihling, C. Stingl, K. Mechtler, O. Zschoernig, and A. Sinz. Annexin A2/P11 interaction: New insights into annexin A2 tetramer structure by chemical crosslinking, high-resolution mass spectrometry, and computational modeling. *Proteins: Structure Function, and Bioinformatics*, 69:254–269, 2007.
32. M. Albrecht, D. Hanisch, R. Zimmer, and T. Lengauer. Improving fold recognition of protein threading by experimental distance constraints. *In Silico Biology*, 2:325–337, 2002.
33. A.N. Nickolai, N. Ruth, and Z.M. Ralf. Fast protein fold recognition via sequence to structure alignment and contact capacity potentials. In *Pacific Symposium on Biocomputing '96*, pp. 53–72. Singapore: World Scientific Publishing Co, 1996.
34. B. Corry, P. Rigby, Z.-W. Liu, and B. Martinac. Conformational changes involved in MscL channel gating measured using FRET spectroscopy. *Biophysical Journal*, 89:L49–L51, 2005.
35. Y. Xu, D. Xu, O.H. Crawford, J.R. Einstein, and E. Serpersu. Protein structure determination using protein threading and sparse NMR data (extended abstract). *Proceedings of the Fourth Annual International Conference on Computational Molecular Biology*. Tokyo, Japan: ACM Press, 2000.
36. R.X. Wang, Y.P. Lu, and S.M. Wang. Comparative evaluation of 11 scoring functions for molecular docking. *Journal of Medicinal Chemistry*, 46:2287–2303, 2003.