ELSEVIER

Combining electron microscopy and comparative protein structure modeling
Maya Topf and Andrej Sali

Recently, advances have been made in methods and applications that integrate electron microscopy density maps and comparative modeling to produce atomic structures of macromolecular assemblies. Electron microscopy can benefit from comparative modeling through the fitting of comparative models into electron microscopy density maps. Also, comparative modeling can benefit from electron microscopy through the use of intermediate-resolution density maps in fold recognition, template selection and sequence-structure alignment.

Addresses
Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, and California Institute for Quantitative Biomedical Research, University of California San Francisco, San Francisco, CA 94143, USA

Corresponding author: Sali, Andrej (sali@salilab.org)

Introduction

Atomic high-resolution structures of complete macromolecular complexes, such as viruses, ion channels, ribosomes and proteasomes, are needed for studying their assembly, function, regulation and evolution [1–3]. Although some assembly structures have been determined by X-ray crystallography [4] and NMR spectroscopy [3,4], their number is quite small compared to the thousands of different complexes in a typical cell. Thus, other methods for the structural characterization of assemblies at near atomic resolution are needed.

Electron cryo-microscopy (cryo-EM) can image complexes in their physiological environment and does not require large quantities of the sample [5–10]. Structures of macromolecular complexes with molecular weights greater than ∼150 kDa can now be visualized in different functional states at intermediate resolution (∼5–15 Å) [11,12]. The corresponding cryo-EM maps are generally still insufficient for atomic structure determination on their own. However, they do enable accurate fitting of atomic-resolution structures of assembly components (e.g. domains, proteins and subcomplexes) into the lower-resolution density of the whole assembly [13–17]. Such fitting produces an atomic structure of the entire complex. In addition to rigid-body fitting [18–22,23•], a growing number of 'flexible' fitting methods distort the component structures in order to optimize their fit into the density map [24–31].

Unfortunately, experimentally determined atomic-resolution structures of assembly components are frequently not available. Even when they are available, induced fit may severely limit their usefulness in the reconstruction of the complex. In such cases, it may be possible to get useful models of the components by comparative protein structure modeling (homology modeling) [32–35]. In comparative modeling, the structure of a target protein sequence is predicted by finding one or more related proteins of known structure (i.e. templates), aligning the target sequence to the template structures, building a model based primarily on the alignment from the previous step and assessing the model. Currently, ∼1.1 million of the ∼1.9 million known protein sequences [36] have at least one domain that can be modeled based on its similarity to one or more of the ∼30,000 known protein structures [37]. Thus, the number of models that can be constructed with useful accuracy, at least comparable to the resolution of the cryo-EM maps, is almost two orders of magnitude greater than the number of available experimentally determined structures. Moreover, comparative modeling is becoming increasingly applicable and accurate. These improvements result partly from structural genomics initiatives, which aim to solve the structure of a member of most protein families by X-ray crystallography or NMR spectroscopy, such that most of the remaining proteins can be modeled with useful accuracy based on their similarity to the known structure [37–39]. They are also due to the growth in sequence databases, faster computers and improved comparative modeling methods [34,35,40].

In this review, we list recent advances in methods and applications that involve integrating comparative modeling and EM density maps to produce atomic structures of macromolecular assemblies. We start by describing how EM can benefit from comparative modeling through the fitting of comparative models into EM density maps. Next, we show how comparative modeling can benefit from EM through the use of intermediate-resolution density maps in fold recognition, template selection

and sequence-structure alignment. Finally, we outline our approach to combined comparative modeling and fitting into a density map.

Fitting comparative models into a density map

It is becoming increasingly common to fit comparative models into EM density maps to produce atomic structures of macromolecular assemblies and yield insight into their functions [41]. A recent study showed that fitting a comparative model based on a remotely related template structure is generally better than fitting the template itself [23•]. Specifically, an accurate comparative model of a protein based on an experimentally determined atomic structure of a remotely related homolog generally has a higher correlation with its intermediate-resolution density map than the homolog itself.

Assembly structures that have recently been determined at low resolution (worse than 15 Å), which involved, at least partly, the fitting of comparative models into EM maps, include the actin–fibrin bundle [41], the membrane-bound coagulation factor FVIII–FIXa complex [42], the skeletal muscle Ca^{2+} release channel [43], the yeast exosome [44], smooth muscle α-actinin [45], the cytosolic chaperonin CCF complexed with phosducin-like protein [46] and the karyopherin CRM1–exportin1 complex [47•]. In the study of CRM1–exportin1, integrating sequence analysis, comparative modeling, a low-resolution density map (i.e. a 22 Å resolution 'negative-stain' EM map) and mutagenesis data suggested cooperativity during formation of the nuclear export complex. This study demonstrates how the hybrid approach can sometimes lead to important biological insights, even when the resolution of the density map is worse than 20 Å. Low-resolution maps, however, may not have sufficiently distinctive features for the unambiguous placement of the high-resolution assembly components [18,30,48]. In such cases, information about component proximity, derived from experiments such as affinity purification or immuno-EM [49], can be combined with the shape information encoded by the EM density to localize the components [50].

At intermediate resolution, the effectiveness of combining cryo-EM maps and comparative modeling is demonstrated by several ribosome studies [15,51–54,55••]. Due to the conservation of the ribosome across species, it is frequently possible to build relatively accurate comparative models of many ribosomal proteins. For instance, comparative models and crystal structures of ribosomal RNAs and proteins were fitted into the 12 Å resolution cryo-EM density map of a targeting complex, consisting of a mammalian signal recognition particle (SRP) bound to an active 80S ribosome carrying a signal sequence [53]. As a result, a molecular model of SRP in its functional state was obtained, showing how its S-domain contacts the large ribosomal subunit at the nascent chain exit site

in order to bind the signal sequence. In another example, the ribosome-bound RACK1 protein was first localized on the ribosome using cryo-EM density maps at 15.6 Å resolution with and without the protein [55••]. Next, a comparative model of RACK1 was fitted into the 11.7 Å resolution map of the *Saccharomyces cerevisiae* 80S ribosome (Figure 1). The resulting model structure, supplemented by a representation of the charges on its surface, mutational analysis, and an atomic model of the 40S subunit derived by cryo-EM and comparative modeling, revealed the interactions of RACK1 with the ribosome and clarified its role in the control of eukaryotic translation.
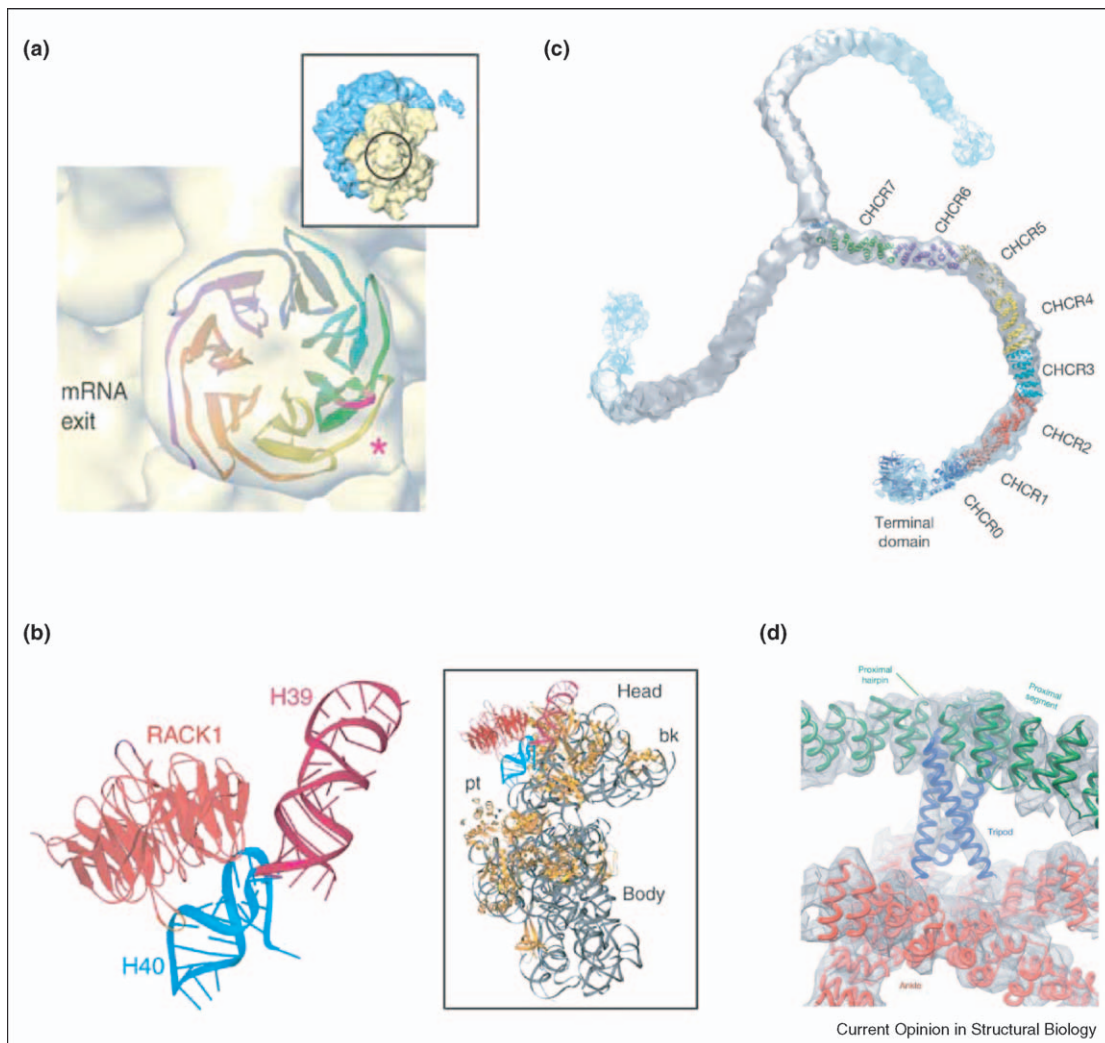
In the subnanometer-resolution range, most examples of fitting comparative models into cryo-EM maps involve highly symmetric biological machines. The usefulness of this approach was recently demonstrated by constructing an atomic model of a complete clathrin lattice [56••] (Figure 1). Fitting crystallographic structures of two components and comparative models of the other seven components into the density map at ~8 Å resolution allowed the tracing of most of the 1675-residue clathrin heavy chain. The atomic structure explains how clathrin-coated vesicles adapt to cargos of different shapes and sizes.

Fold assignment and template selection

When the resolution of the density map is better than ~12 Å, it is possible to use it for fold assignment of the constituting domains (Figure 2). Such fold assignments are particularly helpful when a structural homolog of the target component cannot be detected by sequence-based or threading search methods [34,35]. Recent examples include viral [57] and membrane [58] proteins. At ~12 Å resolution, it is usually possible to recognize boundaries between the individual components in the complex. Secondary structure features, such as long α helices and large β sheets, can begin to be identified at ~10 Å resolution, and short helices and individual strands at ~4 Å [59]. When the fold is known, the density maps can be useful in selecting template structures for comparative modeling and fitting [23•]. Furthermore, even if a crystallographic structure of the isolated component exists, the density map may still be useful in selecting a template whose conformation better resembles the conformation of the target protein in the context of the assembly compared to the isolated component.

There are two classes of methods for identifying the structural motifs in a given density map. The methods in the first class identify secondary structure segments and their arrangement in space without recourse to known protein structures. Examples include a method for identifying the location of helices consisting of three or more turns (*Helixhunter* [19]); a method for assigning transmem-
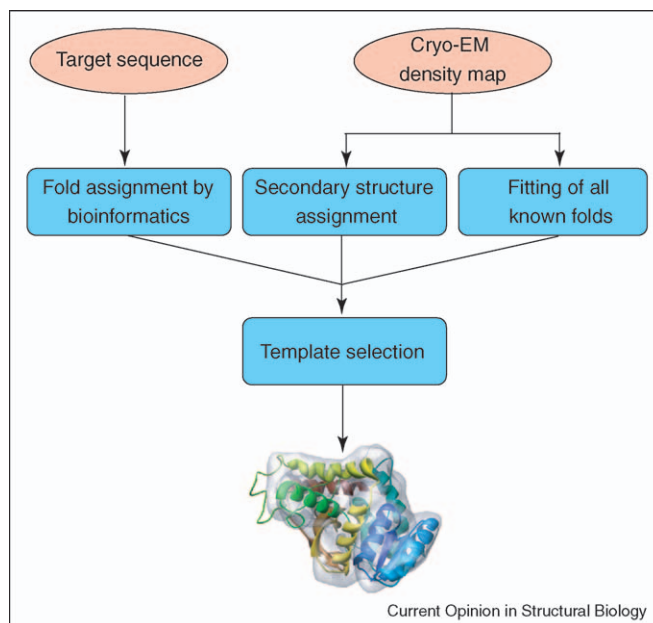
Figure 1

Fitting comparative models into cryo-EM density maps. **(a)** The RACK1 model (β propeller) fitted into the *S. cerevisiae* 80S ribosome density map (40S subunit in yellow, 60S subunit in blue) [55**,79], showing a protruding bulge, marked with an asterisk. **(b)** In combination with the atomic model of the 40S subunit (bk, beak; pt, platform), the fit reveals RACK1 interactions with 18S rRNA helices 39 and 40. **(c)** Crystallographic structures and comparative models fitted into the clathrin triskelion density map [56**]. **(d)** The fit reveals the atomic structure of the hub assembly: the tripod helix (blue) connects the C-terminal part of the CHCR7 segments (green) with the ankle segments from triskelions centered two vertices away (red).

brane helices in the sequences of integral membrane proteins to the approximate locations of the helices in the density map [60]; methods for locating regions belonging to β sheets (*Sheetminer* [61]) and for building Cα models of β sheets (*Sheettracer* [62]); a deconvolution method for enhancing secondary structure features in density maps [62]; and a method for determining protein topology from skeletons of secondary structure segments [63•]. These methods can be combined with secondary structure prediction based on sequence only [64–66] to improve the secondary structure assignments [67,68, 69•,70]. In addition, the assigned relative spatial arrangement of helices can be matched against known protein structures to identify remote structural relationships [71–73]. This approach was used successfully for detecting previously determined structures similar to the rice dwarf virus outer shell protein RDV P8 [67] and the herpes simplex virus capsid protein VP5 (middle domain) [69•].

The methods in the second class are based on the fitting of structural units larger than secondary structure segments, such as whole folds. Recently, a method for fold assignment at intermediate resolution, SPI-EM, has been developed [74•]. The method combines the fitting of many high-resolution domains from the CATH database

Figure 2

Integration of fold assignment and template selection with cryo-EM information, as exemplified by a number of methods [56••,67,68,69•,70]. Information derived from cryo-EM density maps (fold assignment by identifying secondary structure segments [19,60–62] or by fitting known structures [74•]) is integrated with bioinformatics methods (fold assignment from sequence [34], threading [35] and secondary structure prediction [64–66]) to assign the fold of the target sequence. Finally, the candidate templates are fitted into the density map and the best template is selected by the quality of its fit.

[75] into a given cryo-EM density map at intermediate resolution with a probabilistic analysis to determine the superfamily of the component. However, as the resolution decreases, some superfamily folds become indistinguishable. In such cases, the method will require additional constraints from other sources, such as bioinformatics or experiment.
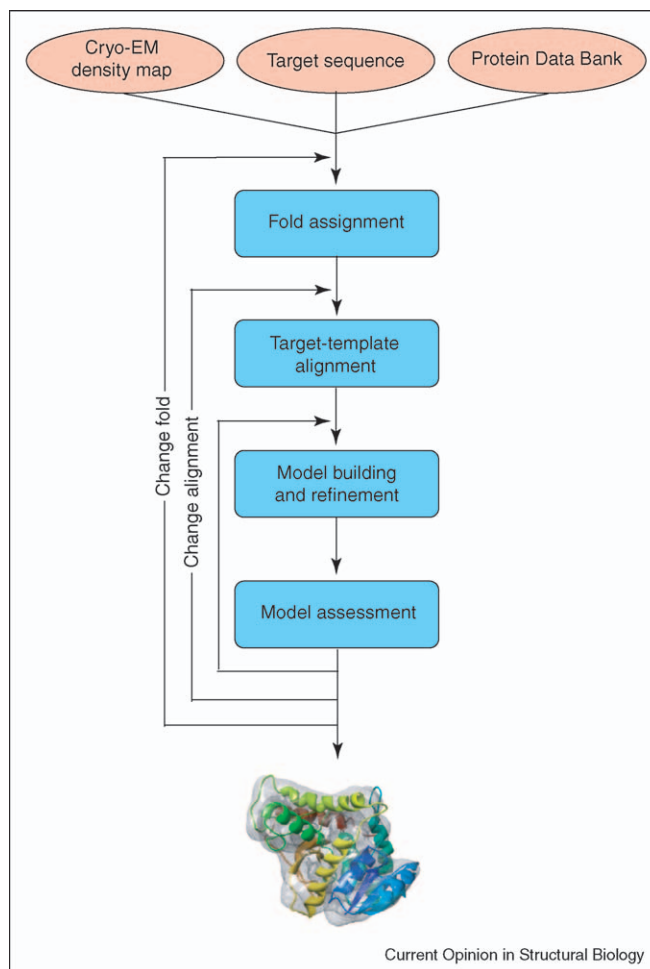
Improving comparative models

It was recently shown that intermediate-resolution cryo-EM density maps are helpful for improving the accuracy of comparative protein structure modeling [23•]. More specifically, the cross-correlation between a comparative model and the corresponding density map is highly correlated with the accuracy of the model; in other words, a more accurate model fits the EM density map more tightly. Therefore, cryo-EM maps are useful not only for deriving atomic models of assemblies, but also for reducing comparative modeling errors.

The largest errors in comparative models result from incorrect sequence alignment and fold assignment, especially in models of sequences that are remotely related to their templates (i.e. less than 30% sequence identity). Other errors include rigid-body shifts, errors in the modeling of loops and errors in sidechain packing [34]. It is usually possible to use different programs and options to generate a set of models based on alternative templates and alignments that vary in the orientation of domains, the packing of secondary structure elements and the conformation of loops. Selecting the most accurate model from a model set can then be accomplished by combining [23•] density fitting with various methods of model assessment, such as statistical potentials of mean force [76,77].

An example of such an approach relies on the previously developed 'moulding' protocol, which iterates over alignment, model building and model assessment [78] (Figure 3). The underlying genetic algorithm depends on the operators that evolve the alignments by mutations and cross-overs, as well as on a fitness function that corresponds to model assessment by statistical potentials. A combined comparative modeling and density fitting protocol (M Topf *et al.*, unpublished) is implemented by adding the cryo-EM density fitting score to the fitness function. A benchmark shows that intermediate-resolution density maps help reduce the Cα root mean square error of comparative models by ∼27%. Further improvements of combined comparative modeling and cryo-EM fitting are likely to be achieved by exploring the conformational and configurational degrees of freedom using flexible fitting, elastic deformation and real-space refinement methods [24–30].

Figure 3

Iterative model building, refinement and assessment relying on cryo-EM density fitting. Alternative target-template alignments and the corresponding models are generated. The models are then evaluated by the combination of a statistical potentials score and a density fitting score. The final structure is consistent with template considerations, molecular mechanics energy functions and the cryo-EM density map.

Conclusions

Our broad objective is to maximize the coverage, accuracy, resolution and efficiency of the structural characterization of macromolecular assemblies [2,3]. This aim is likely to be achieved by integrating techniques that consider various types of information, including density maps from cryo-EM, atomic structures from crystallography and NMR spectroscopy, and atomic models from protein structure prediction. A major class of such hybrid methods involves the fitting of molecular components into cryo-EM density maps of large assemblies. Given the increasing number of molecular machines under investigation using cryo-EM techniques, it is likely that fitting comparative models into intermediate-resolution density maps will increase significantly in the near future.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Alberts B: **The cell as a collection of protein machines: preparing the next generation of molecular biologists**. *Cell* 1998, **92**:291-294.

2. Sali A, Glaeser R, Earnest T, Baumeister W: **From words to literature in structural proteomics**. *Nature* 2003, **422**:216-225.

3. Russell RB, Alber F, Aloy P, Davis FP, Korkin D, Pichaud M, Topf M, Sali A: **A structural perspective on protein-protein interactions**. *Curr Opin Struct Biol* 2004, **14**:313-324.

4. Dutta S, Berman HM: **Large macromolecular complexes in the Protein Data Bank: a status report**. *Structure* 2005, **13**:381-388.

5. Frank J: **Single-particle imaging of macromolecules by cryo-electron microscopy**. *Annu Rev Biophys Biomol Struct* 2002, **31**:303-319.

6. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH: **Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy**. *J Mol Biol* 1990, **213**:899-929.

7. Nogales E, Wolf SG, Downing KH: **Structure of the alpha beta tubulin dimer by electron crystallography**. *Nature* 1998, **391**:199-203.

8. Yonekura K, Maki-Yonekura S, Namba K: **Building the atomic model for the bacterial flagellar filament by electron cryomicroscopy and image analysis**. *Structure* 2005, **13**:407-412.

9. van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M *et al.*: **Single-particle electron cryo-microscopy: towards atomic resolution**. *Q Rev Biophys* 2000, **33**:307-369.

10. Ranson NA, Farr GW, Roseman AM, Gowen B, Fenton WA, Horwich AL, Saibil HR: **ATP-bound states of GroEL captured by cryo-electron microscopy**. *Cell* 2001, **107**:869-879.

11. Henderson R: **Realizing the potential of electron cryo-microscopy**. *Q Rev Biophys* 2004, **37**:3-13.

12. Chiu W, Baker ML, Jiang W, Dougherty M, Schmid MF: **Electron cryomicroscopy of biological machines at subnanometer resolution**. *Structure* 2005, **13**:363-372.

13. Rossmann MG, Morais MC, Leiman PG, Zhang W: **Combining X-ray crystallography and electron microscopy**. *Structure* 2005, **13**:355-362.

14. Davis JA, Takagi Y, Kornberg RD, Asturias FA: **Structure of the yeast RNA polymerase II holoenzyme: mediator conformation and polymerase interaction**. *Mol Cell* 2002, **10**:409-415.

15. Gao H, Sengupta J, Valle M, Korostelev A, Eswar N, Stagg SM, Roey PV, Agrawal RK, Harvey SC, Sali A *et al.*: **Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement**. *Cell* 2003, **113**:789-801.

16. Holmes KC, Angert I, Kull FJ, Jahn W, Schroder RR: **Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide**. *Nature* 2003, **425**:423-427.

17. Ludtke SJ, Chen DH, Song JL, Chuang DT, Chiu W: **Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy**. *Structure* 2004, **12**:1129-1136.

18. Roseman AM: **Docking structures of domains into maps from cryo-electron microscopy using local correlation**. *Acta Crystallogr D Biol Crystallogr* 2000, **56**:1332-1340.

19. Jiang W, Baker ML, Ludtke SJ, Chiu W: **Bridging the information gap: computational tools for intermediate resolution structure interpretation**. *J Mol Biol* 2001, **308**:1033-1044.

20. Chacon P, Wriggers W: **Multi-resolution contour-based fitting of macromolecular structures**. *J Mol Biol* 2002, **317**:375-384.

21. Navaza J, Lepault J, Rey FA, Alvarez-Rua C, Borge J: **On the fitting of model electron densities into EM reconstructions: a reciprocal-space formulation**. *Acta Crystallogr D Biol Crystallogr* 2002, **58**:1820-1825.

22. Volkmann N, Hanein D: **Docking of atomic models into reconstructions from electron microscopy**. *Methods Enzymol* 2003, **374**:204-225.

23. Topf M, Baker ML, John B, Chiu W, Sali A: **Structural
• characterization of components of protein assemblies by comparative modeling and electron cryo-microscopy**. *J Struct Biol* 2005, **149**:191-203.
A new method for fitting comparative models into intermediate-resolution cryo-EM density maps (Mod-EM, a module in MODELLER) is used to demonstrate how such maps can be helpful in improving the accuracy of models. It is shown that one of the most accurate models among alternative comparative models can be identified by the quality of its fit into the density.

24. Tama F, Wriggers W, Brooks CL III: **Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory**. *J Mol Biol* 2002, **321**:297-305.

25. Keskin O, Durell SR, Bahar I, Jernigan RL, Covell DG: **Relating molecular flexibility to function: a case study of tubulin**. *Biophys J* 2002, **83**:663-680.

26. Kong Y, Ming D, Wu Y, Stoops JK, Zhou ZH, Ma J: **Conformational flexibility of pyruvate dehydrogenase complexes: a computational analysis by quantized elastic deformational model**. *J Mol Biol* 2003, **330**:129-135.

27. Chacon P, Tama F, Wriggers W: **Mega-Dalton biomolecular motion captured from electron microscopy reconstructions**. *J Mol Biol* 2003, **326**:485-492.

28. Ma J: **Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes**. *Structure* 2005, **13**:373-380.

29. Gao H, Frank J: **Molding atomic structures into intermediate-resolution cryo-EM density maps of ribosomal complexes using real-space refinement**. *Structure* 2005, **13**:401-406.

30. Fabiola F, Chapman MS: **Fitting of high-resolution structures into electron microscopy reconstruction images**. *Structure* 2005, **13**:389-400.

31. Wriggers W: **Spanning the length scales of biomolecular simulation**. *Structure* 2004, **12**:1-2.

32. Blundell TL, Sibanda BL, Sternberg MJ, Thornton JM: **Knowledge-based prediction of protein structures and the design of novel molecules**. *Nature* 1987, **326**:347-352.

33. Sali A, Blundell TL: **Comparative protein modelling by satisfaction of spatial restraints**. *J Mol Biol* 1993, **234**:779-815.

34. Madhusudhan MS, Marti-Renom MA, Eswar N, John B, Pieper U, Karchin R, Shen M, Sali A: **Comparative protein structure modeling**. In *The Proteomics Protocols Handbook.* Edited by Walker JM. Humana Press Inc; 2005:831-860.

35. Ginalski K, Grishin NV, Godzik A, Rychlewski L: **Practical lessons from protein structure prediction**. *Nucleic Acids Res* 2005, **33**:1874-1891.

36. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M *et al.*: **The Universal Protein Resource (UniProt)**. *Nucleic Acids Res* 2005, **33**:D154-D159.

37. Pieper U, Eswar N, Braberg H, Madhusudhan MS, Davis FP, Stuart AC, Mirkovic N, Rossi A, Marti-Renom MA, Fiser A *et al.*: **MODBASE, a database of annotated comparative protein structure models, and associated resources**. *Nucleic Acids Res* 2004, **32**:D217-D222.

38. Sali A, Kuriyan J: **Challenges at the frontiers of structural biology**. *Trends Cell Biol* 1999, **9**:M20-M24.

39. Baker D, Sali A: **Protein structure prediction and structural genomics**. *Science* 2001, **294**:93-96.

40. Wallner B, Elofsson A: **All are not equal: a benchmark of different homology modeling programs**. *Protein Sci* 2005, **14**:1315-1327.

41. Volkmann N, DeRosier D, Matsudaira P, Hanein D: **An atomic model of actin filaments cross-linked by fimbrin and its implications for bundle assembly and function**. *J Cell Biol* 2001, **153**:947-956.

42. Stoilova-McPhie S, Villoutreix BO, Mertens K, Kemball-Cook G, Holzenburg A: **3-dimensional structure of membrane-bound coagulation factor VIII: modeling of the factor VIII heterodimer within a 3-dimensional density map derived by electron crystallography**. *Blood* 2002, **99**:1215-1223.

43. Baker ML, Serysheva II, Sencer S, Wu Y, Ludtke SJ, Jiang W, Hamilton SL, Chiu W: **The skeletal muscle Ca²⁺ release channel has an oxidoreductase-like domain**. *Proc Natl Acad Sci USA* 2002, **99**:12155-12160.

44. Aloy P, Ciccarelli FD, Leutwein C, Gavin AC, Superti-Furga G, Bork P, Boettcher B, Russell RB: **A complex prediction: three-dimensional model of the yeast exosome**. *EMBO Rep* 2002, **3**:628-635.

45. Liu J, Taylor DW, Taylor KA: **A 3-D reconstruction of smooth muscle alpha-actinin by cryoEM reveals two different conformations at the actin-binding region**. *J Mol Biol* 2004, **338**:115-125.

46. Martin-Benito J, Bertrand S, Hu T, Ludtke PJ, McLaughlin JN, Willardson BM, Carrascosa JL, Valpuesta JM: **Structure of the complex between the cytosolic chaperonin CCT and phosducin-like protein**. *Proc Natl Acad Sci USA* 2004, **101**:17410-17415.

47. Petosa C, Schoehn G, Askjaer P, Bauer U, Moulin M,
• Steuerwald U, Soler-Lopez M, Baudin F, Mattaj IW, Muller CW: **Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex**. *Mol Cell* 2004, **16**:761-775.
A structural model of the human karyopherin CRM1 is suggested based on a combination of X-ray crystallography, EM and comparative modeling. The architecture of CRM1 resembles that of the import receptor transportin1, with 19 HEAT repeats and a large loop implicated in the binding of the small GTPase Ran. The integrated structural information provides insights into karyopherin-mediated nuclear export signal recognition.

48. Wriggers W, Chacon P: **Modeling tricks and fitting techniques for multiresolution structures**. *Structure* 2001, **9**:779-788.

49. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT: **The yeast nuclear pore complex: composition, architecture, and transport mechanism**. *J Cell Biol* 2000, **148**:635-651.

50. Alber F, Kim MF, Sali A: **Structural characterization of assemblies from overall shape and subcomplex compositions**. *Structure* 2005, **13**:435-445.

51. Beckmann R, Spahn CM, Eswar N, Helmers J, Penczek PA, Sali A, Frank J, Blobel G: **Architecture of the protein-conducting channel associated with the translating 80S ribosome**. *Cell* 2001, **107**:361-372.

52. Tung CS, Joseph S, Sanbonmatsu KY: **All-atom homology model of the *Escherichia coli* 30S ribosomal subunit**. *Nat Struct Biol* 2002, **9**:750-755.

53. Halic M, Becker T, Pool MR, Spahn CM, Grassucci RA, Frank J, Beckmann R: **Structure of the signal recognition particle interacting with the elongation-arrested ribosome**. *Nature* 2004, **427**:808-814.

54. Wild K, Halic M, Sinning I, Beckmann R: **SRP meets the ribosome**. *Nat Struct Mol Biol* 2004, **11**:1049-1053.

55. Sengupta J, Nilsson J, Gursky R, Spahn CM, Nissen P, Frank J:
•• **Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM**. *Nat Struct Mol Biol* 2004, **11**:957-962.
Using an 11.7 Å resolution cryo-EM map of the *S. cerevisiae* 80S ribosome, the RACK1 protein was localized to the head region of the 40S subunit in the immediate vicinity of the mRNA exit channel. A detailed analysis of the fit of the RACK1 comparative model into the density reveals how the protein interacts with rRNA and how it can recruit proteins to the ribosome via its WD-repeat platform.

56. Fotin A, Cheng Y, Sliz P, Grigorieff N, Harrison SC,
•• Kirchhausen T, Walz T: **Molecular model for a complete clathrin lattice from electron cryomicroscopy**. *Nature* 2004, **432**:573-579.
The architecture of a complete clathrin lattice is resolved by fitting X-ray structures and comparative models into its subnanometer-resolution cryo-EM density map (1675 residues). The structure explains how clathrin-coated vesicles adapt to cargos of different shapes and sizes.

57. Zhou ZH, Chiu W: **Determination of icosahedral virus structures by electron cryomicroscopy at subnanometer resolution**. *Adv Protein Chem* 2003, **64**:93-124.

58. Stahlberg H, Engel A, Philippsen A: **Assessing the structure of membrane proteins: combining different methods gives the full picture**. *Biochem Cell Biol* 2002, **80**:563-568.

59. Chiu W, Baker ML, Jiang W, Zhou ZH: **Deriving folds of macromolecular complexes through electron cryomicroscopy and bioinformatics approaches**. *Curr Opin Struct Biol* 2002, **12**:263-269.

60. Enosh A, Fleishman SJ, Ben-Tal N, Halperin D: **Assigning transmembrane segments to helices in intermediate-resolution structures**. *Bioinformatics* 2004, **20(suppl 1)**: I122-I129.

61. Kong Y, Ma J: **A structural-informatics approach for mining beta-sheets: locating sheets in intermediate-resolution density maps**. *J Mol Biol* 2003, **332**:399-413.

62. Kong Y, Zhang X, Baker TS, Ma J: **A Structural-informatics approach for tracing beta-sheets: building pseudo-C(alpha) traces for beta-strands in intermediate-resolution density maps**. *J Mol Biol* 2004, **339**:117-130.

63. Wu Y, Chen M, Lu M, Wang Q, Ma J: **Determining protein
• topology from skeletons of secondary structures**. *J Mol Biol* 2005, **350**:571-586.
A new method for determining protein topology by defining loop connectivity based on skeletons of secondary structure segments obtained from intermediate-resolution cryo-EM density maps. The method involves a knowledge-based geometry filter and an energy-based evaluation. It is tested successfully on a wide range of protein architectures.

64. Rost B, Sander C, Schneider R: **PHD-an automatic mail server for protein secondary structure prediction**. *Comput Appl Biosci* 1994, **10**:53-60.

65. Baldi P, Brunak S, Frasconi P, Soda G, Pollastri G: **Exploiting the past and the future in protein secondary structure prediction**. *Bioinformatics* 1999, **15**:937-946.

66. McGuffin LJ, Bryson K, Jones DT: **The PSIPRED protein structure prediction server**. *Bioinformatics* 2000, **16**:404-405.

67. Zhou ZH, Baker ML, Jiang W, Dougherty M, Jakana J, Dong G, Lu G, Chiu W: **Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus**. *Nat Struct Biol* 2001, **8**:868-873.

68. Jiang W, Li Z, Zhang Z, Baker ML, Prevelige PE Jr, Chiu W: **Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions**. *Nat Struct Biol* 2003, **10**:131-135.

69. Baker ML, Jiang W, Bowman BR, Zhou ZH, Quiocho FA, Rixon FJ,
• Chiu W: **Architecture of the herpes simplex virus major capsid protein derived from structural bioinformatics**. *J Mol Biol* 2003, **331**:447-456.
Structural analysis of the 8.5 Å resolution cryo-EM map of the major capsid protein (VP5) of herpes simplex virus type 1 is combined with sequence-based informatics to give its architectural model. The model provides insights into the strategies used to achieve viral capsid stability.

70. Zhang W, Chipman PR, Corver J, Johnson PR, Zhang Y, Mukhopadhyay S, Baker TS, Strauss JH, Rossmann MG, Kuhn RJ: **Visualization of membrane protein domains by cryo-electron microscopy of dengue virus**. *Nat Struct Biol* 2003, **10**:907-912.

71. Kleywegt GJ, Jones TA: **Detecting folding motifs and similarities in protein structures**. In *Methods in Enzymology*, vol 277. Edited by Charles W, Carter J, Sweet RM. Academic Press; 1997:525-545.

72. Kinoshita K, Kidera A, Go N: **Diversity of functions of proteins with internal symmetry in spatial arrangement of secondary structural elements**. *Protein Sci* 1999, **8**:1210-1217.

73. Mizuguchi K, Go N: **Comparison of spatial arrangements of secondary structural elements in proteins**. *Protein Eng* 1995, **8**:353-362.

74. Velazquez-Muriel JA, Sorzano CO, Scheres SH, Carazo JM:
• **SPI-EM: towards a tool for predicting CATH superfamilies in 3D-EM maps**. *J Mol Biol* 2005, **345**:759-771.
SPI-EM is a new method for determining the fold of a protein domain using its EM density map at intermediate resolution. This aim is achieved by fitting protein structural domains representative of the Protein Data Bank into an EM map of the specimen, followed by a probabilistic analysis of the results. The approach is demonstrated with the aid of simulated and experimental multidomain density maps.

75. Orengo CA, Pearl FM, Thornton JM: **The CATH domain structure database**. *Methods Biochem Anal* 2003, **44**:249-271.

76. Sippl MJ: **Recognition of errors in three-dimensional structures of proteins**. *Proteins* 1993, **17**:355-362.

77. Melo F, Sanchez R, Sali A: **Statistical potentials for fold assessment**. *Protein Sci* 2002, **11**:430-448.

78. John B, Sali A: **Comparative protein structure modeling by iterative alignment, model building and model assessment**. *Nucleic Acids Res* 2003, **31**:3982-3992.

79. Spahn CM, Gomez-Lorenzo MG, Grassucci RA, Jorgensen R, Andersen GR, Beckmann R, Penczek PA, Ballesta JP, Frank J: **Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation**. *EMBO J* 2004, **23**:1008-1019.