Journal of Structural Biology 185 (2014) 427–439

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Combined approaches to flexible fitting and assessment in virus capsids undergoing conformational change

Arun Prasad Pandurangan [a,1], Shabih Shakeel [b,1], Sarah Jane Butcher [b], Maya Topf [a,*]

[a] Institute of Structural and Molecular Biology, Department of Biological Sciences/Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, United Kingdom
[b] Institute of Biotechnology, P.O. Box 65 (Viikinkaari 1), FIN-00014 University of Helsinki, Helsinki, Finland

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form 28 November 2013
Accepted 6 December 2013
Available online 12 December 2013

Keywords:
Coxsackievirus A7
Picornaviridae
Flexible fitting
Electron cryo-microscopy
Model assessment

ABSTRACT

Fitting of atomic components into electron cryo-microscopy (cryoEM) density maps is routinely used to understand the structure and function of macromolecular machines. Many fitting methods have been developed, but a standard protocol for successful fitting and assessment of fitted models has yet to be agreed upon among the experts in the field. Here, we created and tested a protocol that highlights important issues related to homology modelling, density map segmentation, rigid and flexible fitting, as well as the assessment of fits. As part of it, we use two different flexible fitting methods (Flex-EM and iMODfit) and demonstrate how combining the analysis of multiple fits and model assessment could result in an improved model. The protocol is applied to the case of the mature and empty capsids of Coxsackievirus A7 (CAV7) by flexibly fitting homology models into the corresponding cryoEM density maps at 8.2 and 6.1 Å resolution. As a result, and due to the improved homology models (derived from recently solved crystal structures of a close homolog – EV71 capsid – in mature and empty forms), the final models present an improvement over previously published models. In close agreement with the capsid expansion observed in the EV71 structures, the new CAV7 models reveal that the expansion is accompanied by ~5° counterclockwise rotation of the asymmetric unit, predominantly contributed by the capsid protein VP1. The protocol could be applied not only to viral capsids but also to many other complexes characterised by a combination of atomic structure modelling and cryoEM density fitting.

1. Introduction

In recent years, electron cryo-microscopy (cryoEM) has become one of the most prominent techniques for visualising macromolecular assemblies (Orlova and Saibil, 2011; Sali et al., 2003). However, the vast majority of density maps resulting from the various cryoEM reconstruction techniques are not of atomic or near-atomic resolution (even for icosahedral viruses) but rather belong to the so-called intermediate resolution zone (~5–20 Å) (Baker et al., 1999; Beck et al., 2011), where a detailed interpretation of the map can only be achieved by docking (or fitting) into it an atomic model. Docking of atomic models (from X-ray crystallography, NMR or structure prediction methods) into EM maps has become common practice with a rapidly increasing number of atomic models associated with EM maps deposited in the PDB (currently over 460) (Lawson et al., 2011; Patwardhan et al., 2012).

Due to the differences between the conformations of the atomic model being fitted and the EM map, modifying the conformation of the atomic structure during the fitting process, referred to as flexible fitting, is often needed (Beck et al., 2011). The variety of flexible fitting approaches is currently large. Common to all is the limited sampling of conformational degrees of freedom. Therefore, they are usually applied to components that are first placed into the density map by rigid fitting, whereby a global search of the fit is performed on the atomic model as a single component in six translation/rotation degrees of freedom (Ahmed et al., 2012; Beck et al., 2011). Both rigid and flexible fitting result in a "pseudo-atomic" model for which the quality assessment is not trivial. Approaches that begin to address this issue include the use of confidence intervals and quantifying the best-fitting model relative to a distribution of different fits (Henderson et al., 2012; Tung et al., 2010; Volkmann, 2009; Roseman, 2000; Rossmann et al., 2005; Vasishtan and Topf, 2011). Additionally, if the models are calculated by different methods a question arises regarding their consensus. A recent paper pioneered the issue of consensus among different flexible fitting approaches and proposed to use this

* Corresponding author.
E-mail address: m.topf@cryst.bbk.ac.uk (M. Topf).
[1] Equal contribution.

1047-8477
http://dx.doi.org/10.1016/j.jsb.2013.12.003

information to improve the quality of the fitted models (Ahmed et al., 2012).

Here, we developed a protocol to aid flexible fitting and assessment of virus capsids into cryoEM maps at sub-nanometer resolution. The protocol is designed to use multiple flexible fitting programs, compare and assess the quality of the fit locally, at the level of individual secondary structure elements (SSEs). It also highlights the possibility of producing an improved fit guided by the comparison of multiple independent programs. The protocol is generic and could also be used for systems other than virus capsids.

First, in order to demonstrate the effects of modelling errors on flexible fitting we fitted a homology model of an actin subunit into a density map simulated from a known actin crystal structure in different conformation. Second, to address the challenge of fitting a structure in one conformation into a corresponding EM map in a different conformation, we fitted the crystal structure of EV71 mature (full) capsid into the procapsid map of EV71 strain 1095 (Cifuente et al., 2013). Finally, we applied the protocol to characterise the conformational states of the mature (full) and empty capsid of Coxsackievirus A7 (CAV7). We had previously calculated homology models of the same virus and fitted them into the sub-nanometer resolution cryoEM maps representing the empty (6.09 Å) and full (8.23 Å) CAV7 capsids (Seitsonen et al., 2012).

CAV7 belongs to the *Human enterovirus A* species within the *Picornaviridae* family (Oberste et al., 2004). It is an important pathogen with different strains varying in their pathogenicity and tropism (Seitsonen et al., 2012). The CAV7-USSR strain is associated with flaccid paralysis (Voroshilova and Chumakov, 1959) whereas CAV7-275/58 causes aseptic meningitis (Richter et al., 1971). Our original models were based on remote homologs to the virus (Seitsonen et al., 2012) and were refined within the corresponding cryoEM maps using a single flexible fitting method (Flex-EM) (Topf et al., 2008). Here, to improve our original models, we used as templates, recently published crystal structures of the empty and full capsids of the much closer homolog, EV71 (Plevka et al., 2012; Wang et al., 2012) with capsid protein sequence identity of 60% for VP1, 84% for VP2, and 76% for VP3. This time we refined the homology models using two flexible fitting programs, Flex-EM (Topf et al., 2008) and iMODfit (Lopez-Blanco and Chacon, 2013). The different fits were assessed and compared, and new hybrid pseudo-atomic models were generated using the results from both programs. Finally, the conformational changes between the empty and full capsids were characterised based on the new models.

2. Methods

We describe a protocol for modelling and fitting of virus capsids into the cryoEM maps at intermediate resolution using two different flexible fitting programs (Fig. 1). The main feature of the protocol is its ability to compare and assess the quality of the fits produced by independent programs. This approach allows the identification of reliable local fits as well as those that could be further improved by additional stages of refinement. The assessment/refinement protocol can also be applied to systems other than virus capsids. Below we describe the various steps involved in the protocol.

2.1. Data preparation

2.1.1. Density map segmentation

The capsid of a mature CAV7 and EV71 virion (full) is made of icosahedrally-arranged viral proteins VP1, VP2, VP3 and VP4 with encapsidated RNA. The empty capsid is also icosahedral but lacks VP4 and RNA. The five-fold vertex is composed of VP1 whereas the three- and the two-fold symmetry axes are made of alternating VP2 and VP3. VP4, a small protein characterised by an extended chain (possibly with a small helix in the middle), is present below the shell of VP1, VP2 and VP3. To help the initial rigid fitting of the asymmetric unit of CAV7 we used the manually segmented maps of the individual viral proteins VP1–VP3 from the density of both empty and full capsids, as described in our previous study (Seitsonen et al., 2012). In the CAV7 full map, VP4 could not be segmented unambiguously and therefore we decided that there were not enough density features to accurately model it. For fitting the EV71 full capsid, the procapsid map was segmented around the asymmetric unit using the fit deposited in PDB (PDB ID: 3VBU; EMD-5557) (Cifuente et al., 2013; Wang et al., 2012).

2.1.2. Homology modelling

CAV7 modelling: From the three target sequences of CAV7-USSR, homology models of the capsid proteins (VP1–VP3) were built using the I-TASSER server (Roy et al., 2010). For a given sequence, I-TASSER builds fragments of template proteins using threading and/or *ab initio* techniques. The fragments are assembled and refined into a complete model using replica-exchange Monte Carlo simulation (Roy et al., 2010). The template structures used for the modelling were the respective viral proteins in the enterovirus 71 (EV71) crystal structures of empty (PDB ID: 3VBO) and full (PDB ID: 3VBF) capsid forms (Wang et al., 2012). The server generated five different models for each of the two conformations of the three capsid proteins (30 in total) and we selected the model with the top I-TASSER score (out of the five) for further analysis (six models in total). Additionally, the qualitative model energy analysis (QMEAN) scores (Benkert et al., 2008) were used to evaluate both the global and local quality of the selected models and were compared with the previously published models (Seitsonen et al., 2012). Briefly, the QMEAN score for a given protein model is calculated using a combination of the geometrical structural descriptors that include the torsion angle, pairwise residue and solvation potentials. The best I-TASSER models for the three capsid proteins (VP1, VP2 and VP3) obtained using the template structure of the empty capsid (PDB ID: 3VBO) were assembled into an empty capsid asymmetric unit (“empty asymmetric unit”) by superposing the individual VP proteins onto their respective VP proteins in the template structures. Similarly, a full capsid asymmetric unit (“full asymmetric unit”) was assembled using the I-TASSER model (VP1, VP2 and VP3) obtained using the full capsid (mature virus) as the template (PDB ID: 3VBF). The superposition was done using the *superpose* command in Chimera (Pettersen et al., 2004).

Actin modelling: a homology model of actin was generated from the actin sequence (UniProt: P68135) with MODELLER (Sali and Blundell, 1993) based on the crystal structure of actin-related protein 3 from the Arp2/3 complex, (PDB ID: 1K8K: A) (Robinson et al., 2001). The two proteins share sequence identity of ~38%.

Below, we describe the general procedure we used for fitting the models into the density maps.

2.2. Rigid fitting and re-segmentation

The actin model was rigidly fitted into the simulated map of the native structure with the Chimera *fit_in_map* tool (Goddard et al., 2007). For EV71 test case, the initial rigid fit was obtained by superposing the asymmetric unit onto the asymmetric fit deposited in PDB (PDB ID: 3VBU; EMD-5557).

In real-case scenarios of virus capsids, however, a rigid fit can be obtained by fitting individual subunits or the whole asymmetric unit into the density (either of the whole virus or segmented around the asymmetric unit). The former approach is followed when the arrangements of the subunits within the asymmetric unit is unknown. The latter approach is more appropriate when the knowledge of the intra-subunit interactions within the