Atomic model building and refinement into high-resolution cryo-EM maps

Bruno Klaholz (klaholz@igbmc.fr)
Klaholz's Lab, Centre for Integrative Biology, Strasbourg, France
S. Kundhavai Natchiar
Klaholz's Lab, Centre for Integrative Biology, Strasbourg, France
Alexander Myasnikov
Klaholz's Lab, Centre for Integrative Biology, Strasbourg, France
Hanna Kratzat
Klaholz's Lab, Centre for Integrative Biology, Strasbourg, France (current address: Gene Center, Munich, Germany)
Isabelle Hazemann
Klaholz's Lab, Centre for Integrative Biology, Strasbourg, France

Method Article

Keywords: cryo electron microscopy, atomic model refinement, crystallography, structural biology, human ribosome, ribosome, RNA, RNA modifications

Posted Date: November 21st, 2017

DOI: https://doi.org/10.1038/protex.2017.122

Abstract

The interpretation of high-resolution cryo electron microscopy \(cryo-EM) maps relies on the creation of precise atomic models that help analyzing the detailed interactions within a macromolecular complex. In the accompanying paper to this protocol, we describe the high-resolution structure of the human ribosome, which includes numerous chemical modifications. Here we present a generally applicable protocol for atomic model building and refinement into high-resolution cryo-EM maps, which in addition includes the refinement of nucleotides and amino acids with chemical modifications.

Introduction

With the recent advent of high-resolution cryo electron microscopy \(cryo-EM) it becomes possible and necessary to refine atomic models to high-quality standards, i.e. with a good fit to the map and good geometric parameters like for crystal structures at equivalent resolution. In the past, cryo-EM structures were resolved at more moderate resolution and therefore atomic models underwent relatively simple rigid body or flexible fitting using a known crystal structure. However, this would not address the position of side-chains or their rotamer conformations, which need to be refined in detail into the map taking into account precise geometrical parameters and refine temperature \(B) factors. Here we provide a protocol for atomic model refinement into high-resolution cryo-EM maps. The description includes the method for refining chemically modified residues considering that the high-resolution cryo-EM structure of the human ribosome comprises numerous 2'-O-methylations and various base modifications of the rRNA as well as some amino acid methylations. Similarly to X-ray crystallography, interpreting high-resolution cryo-EM maps with atomic models involves iterative model building and refinement using geometrical constraints and restraints taking into account the geometrical and chemical properties of proteins and nucleic acids \(in particular bond length and distances between atoms within amino acids, DNA or RNA nucleotides and ligands, hydrogen bonding distances or van der Waals contacts). Throughout the structure refinement, the quality of the geometrical parameters is iteratively improved and monitored with Ramachandran plots and clash score statistics.

Reagents

Once a cryo-EM 3D reconstruction has been refined to the highest possible resolution, atomic model building and refinement against the cryo-EM map can be performed, as described in the following \(Fig. 1). See figure in Figures section. For this purpose, we used various tools available in the PHENIX software \(Afonine et al., 2012). First, the original cryo-EM map was Fourier transformed to obtain Fourier coefficients, which were calculated in a cubic box with dimensions of 544 Å in each axis with the grid size of 640. These Fourier coefficients were treated as observed diffraction data \(Data 1) in space group P1 and used for the part of the refinements carried out in reciprocal space. The human 80S ribosome atomic model \(PDB: 4UG0) was manually docked into the cryo-EM map of human 80S using CHIMERA and the initial rigid body fit and steric clashes were visually examined using the software Coot \(Emsley et al., 2010). Subsequently, the docked model underwent rigid body refinement against Data1, followed by real

space refinement against the original EM map. Real space refinement in PHENIX comprised five macro cycles, which include model morphing, local real-space fitting, global gradient-driven refinement, simulated annealing \(SA) and rigid-body refinements. Morphing is an iterative map-assisted fragmented rigid-body fit. Local real space fitting performs torsion angle restraints on individual residues, which fits the main chain and finds the correct side chain rotamers. The global gradient-driven refinement is a target function restraint, in which the model undergoes geometrical minimization with reference to standard restraints. During this refinement cycle, secondary structure restraints and Ramachandran restraints are also implemented. In each macro cycle, individual residue fits are validated with map correlation and absence of steric clashes with neighboring atoms. Because the cryo-EM map is resolved to near atomic resolution, it is possible to refine a large number of model parameters. For the following refinement procedure, atomic model refinements were carried out against Data 1. Refinements with restraints on atomic positions and grouped temperature factors were performed using reciprocal space refinement tools available in PHENIX. In order to move the atomic model refinement out of local minima we applied a SA procedure. SA is a standard procedure in macromolecular crystallography and introduces large atomic motions by simulating high temperature with starting velocities on the atoms, followed by a decrease of the temperature to bring the model into a deeper global ‘energy’ minimum while being restrained against the experimental cryo-EM map. This ‘crystallographic energy’ is a sum of inter- and intra-molecular energy terms completed by a term for similarity of amplitudes of the Fourier coefficients, experimental and those calculated from the model. The reduction in R-factor values allows monitoring the overall correctness of the convergence direction. Further map to model inconsistencies were corrected with a final cycle of real space refinement. Next, manual model inspection and atomic model building including RNA sequence register shift corrections were performed on this atomic model using Coot \(Emsley, 2010). The entire atomic model refinement procedure was repeated until the refinement converges. For the procedure discussed above, the refinement steps and various parameters used are detailed below.

Procedure

Step1: Conversion of the cryo-EM map into Fourier coefficients A set of Fourier coefficients is derived from the cryo-EM map using the command “phenix.map_to_structure_factor”. phenix.map_to_structure_factor 80S_map.ccp4 d_min=2.9 This will write the file “structure_factor.mtz” \ (Data 1) **Step 2: Real space refinement** After initial rigid body fit, fine tuning of the model inside the electron density was performed using “phenix.real_space_refine”. phenix.real_space_refine 80S.pdb 80S_map.ccp4 resolution=2.9 nproc=30 \
run=minimization_global+local_grid_search+simulated_annealing \
simulated_annealing.start_temperature=1500 \ output.file_name_prefix=80S **Step3: Positional, simulated and B-factor refinement** Further model perturbation was carried out with positional and B-factor refinement; prior to this, one cycle of SA refinement was performed with torsion angle restraints. phenix.refine 80S_real_sapce_refined.pdb structure_factor.mtz \ xray_data.low_resolution=60.0 xray_data.high_resolution=2.9 \ optimize_xyz_weight=true optimize_adp_weight=true \ strategy=individual_sites+group_adp \ simulated_annealing=true \ simulated_annealing.mode=first \

simulated_annealing.start_temperature=1000 \ xray_data.r_free_flags.generate=true \ xray_data.r_free_flags.fraction=0.05 \ xray_data.r_free_flags.max_free=20000 \ nproc=30 \ output.prefix=80S-RS \ main.number_of_macro_cycles=5 \ –unused_ok **Restraints for modified nucleotides** The 3D atomic model of modified nucleotides and their parameter files described through restraints libraries were prepared using JLigand available in Coot \(Emsley, 2010) or as a standalone version \(Lebedev et al., 2010). This new monomer library, which defines the moieties of the chemical modifications of the rRNA, can be used for atomic model refinement using PHENIX and also this should be imported into Coot for the model building. In order to maintain the geometry of the O3'-P/P-O3' bonds, as for standard nucleotide libraries, the "data_mod_list" and "data_link_p" records corresponding to DNA/RNA residues need to be added to the new monomer library. In addition, while building the atomic model, these newly built modified monomers were linked with covalent bonds to the canonical nucleic acid residues present in the atomic model with the following steps: i) delete atom OP3' in the modified nucleotide; \(ii) create link records \(O3'\(n-1)-P\(n)) with \(n-1)th and \(O3'\(n)-P\(n+1)) with \(n+1)th nucleotides for each additional nth modified nucleotide using the "Make Link" module available in Coot. **Magnesium ion identification** The identification of ions and water molecules in density maps has an important functional interest for the analysis and interpretation of three-dimensional structures. In X-ray crystallography, this also helps to account for the whole content of ordered molecules in the crystal and optimize the model fit with regards to the experimental diffraction amplitudes. In cryo-EM maps the model to data comparison is done locally and in real space. After fitting the ribosomal proteins and RNA chains to the density map, a significant number of residual densities near the ribosomal proteins and RNA chains were found; these possibly corresponded to water molecules and magnesium ions. Discrimination of magnesium ions from the water molecules involved systematic inspection of the residual density in the cryo-EM map, analysis of the proper coordination and the hydrogen bond potentials \(angle and distance). Magnesium ions were identified manually based on the coordination geometry, chemical environment and electron density distribution. Typical Mg^{2+} ions adopt hexa-coordination with octahedral geometry and the interatomic distance between Mg^{2+} ions and the neighbouring atoms varies between 1.85 and 2.3 Å with a bond angle close to 90º. Moreover, keeping in mind that the hydrated radius of Mg^{2+} ion is substantially larger, some peaks were annotated with Mg^{2+} ions, although they do not possess nearest neighbors for typical coordination. On the other hand, residual densities found at an environment with hydrogen bond geometry with the nearest polar atoms were annotated as water molecules.

Troubleshooting

For simulated annealing refinement, the starting temperature to which the atomic model is heated before slow cooling \(simulated_annealing.start_temperature) needs to be optimized to avoid inappropriate model distortions. In order to avoid large deviation from the input atomic model and the cryo-EM map at a given step of refinement, reference model restraints can be used while refining the atomic model. At final stages of refinement we also noticed that in some instances the tips of the side-chains were moving out of the density; setting rotamer restraints off allowed refining the side-chains properly \(notably for

residues that form hydrogen bonds and whose rotamer conformation might not be well represented in databases). Ramachandran outliers were checked individually and corrected manually in Coot; for this it was often helpful to flip peptides and refine the structure locally in Coot before running a full refinement of the entire atomic model in PHENIX, now with rotamers on when the backbone geometry was already optimized. Finally, in the last round of refinement, an energy minimization with grouped B-factor refinement was run to obtain the final atomic model.

References

S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann & B. P. Klaholz. Visualization of chemical modifications in the human 80S ribosome structure. Nature, 2017, in press. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 \(2010). Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367 \(2012). Lebedev, A. A. et al. JLigand: a graphical tool for the CCP4 template-restraint library. Acta Crystallogr. D Biol. Crystallogr. 68, 431–440 \(2012). Khatter, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P. Structure of the human 80S ribosome. Nature 520, 640–645 \(2015). O. von Loeffelholz, S. K. Natchiar, N. Djabeur, A. G. Myasnikov, H. Kratzat, J.-F. Ménétret, I. Hazemann & B. P. Klaholz. Focused classification and refinement in high-resolution cryo-EM structural analysis of ribosome complexes. Curr. Opin. Struct. Biol., 2017, 46, 140-148. B. P. Klaholz. The ribosome holds the RNA Polymerase on Track in Bacteria. Trends Biochem Sci., 2017, 42, 686-689. I. Orlov, A. G. Myasnikov, L. Andronov, S. K. Natchiar, H. Khatter, B. Beinsteiner, J-F. Ménétret, I. Hazemann, K. Mohideen, K. Tazibt, R. Tabaroni, H. Kratzat, N. Djabeur, T. Bruxelles, F. Raivoniaina, L. di Pompeo, M. Torchy, I. Billas, A. Urzhumtsev & B. P. Klaholz. The integrative role of cryo electron microscopy in molecular and cellular structural biology. Biol Cell., 2017, 109, 81-93. A. G. Myasnikov, S. K. Natchiar, M. Nebout, I. Hazemann, V. Imbert, H. Khatter, J.-F. Peyron & B. P. Klaholz. Structure-function insights reveal the human ribosome as a cancer target for antibiotics. Nat Commun., 2016, 7, 12856. B. P. Klaholz. Structure sorting of multiple macromolecular states in heterogeneous cryo-EM samples by 3D multivariate statistical analysis. Open Journal of Statistics, Special Issue on Multivariate Statistical Analysis, 2015, 5, 820-836. http://dx.doi.org/10.4236/ojs.2015.57081. B. Beinsteiner, J. Michalon & B. P. Klaholz. IBiSS, a versatile and interactive tool for integrated sequence and 3D structure analysis of large macromolecular complexes. Bioinformatics, 2015, 31, 3339-3344. H. Khatter, A. G. Myasnikov, L. Mastio, I. M. L. Billas, C. Birck, S. Stella & B. P. Klaholz. Purification, characterization and crystallization of the human 80S ribosome. Nucleic Acids Res., 2014, 42\(6), 1-11.

Acknowledgements

We thank Jonathan Michalon, Remy Fritz and Romaric David for IT support. This work was supported by CNRS, Association pour la Recherche sur le Cancer \(ARC), Institut National du Cancer \(INCa), Ligue nationale contre le cancer \(Ligue) and Agence National pour la Recherche \(ANR). The electron microscope facility was supported by the Alsace Region, the Fondation pour la Recherche Médicale \

(FRM), Inserm, CNRS and ARC, by Instruct-ULTRA as part of the European Union's Horizon 2020 grant ID 731005, the French Infrastructure for Integrated Structural Biology \(FRISBI) ANR-10-INSB-05-01 and by Instruct-ERIC.

Figures

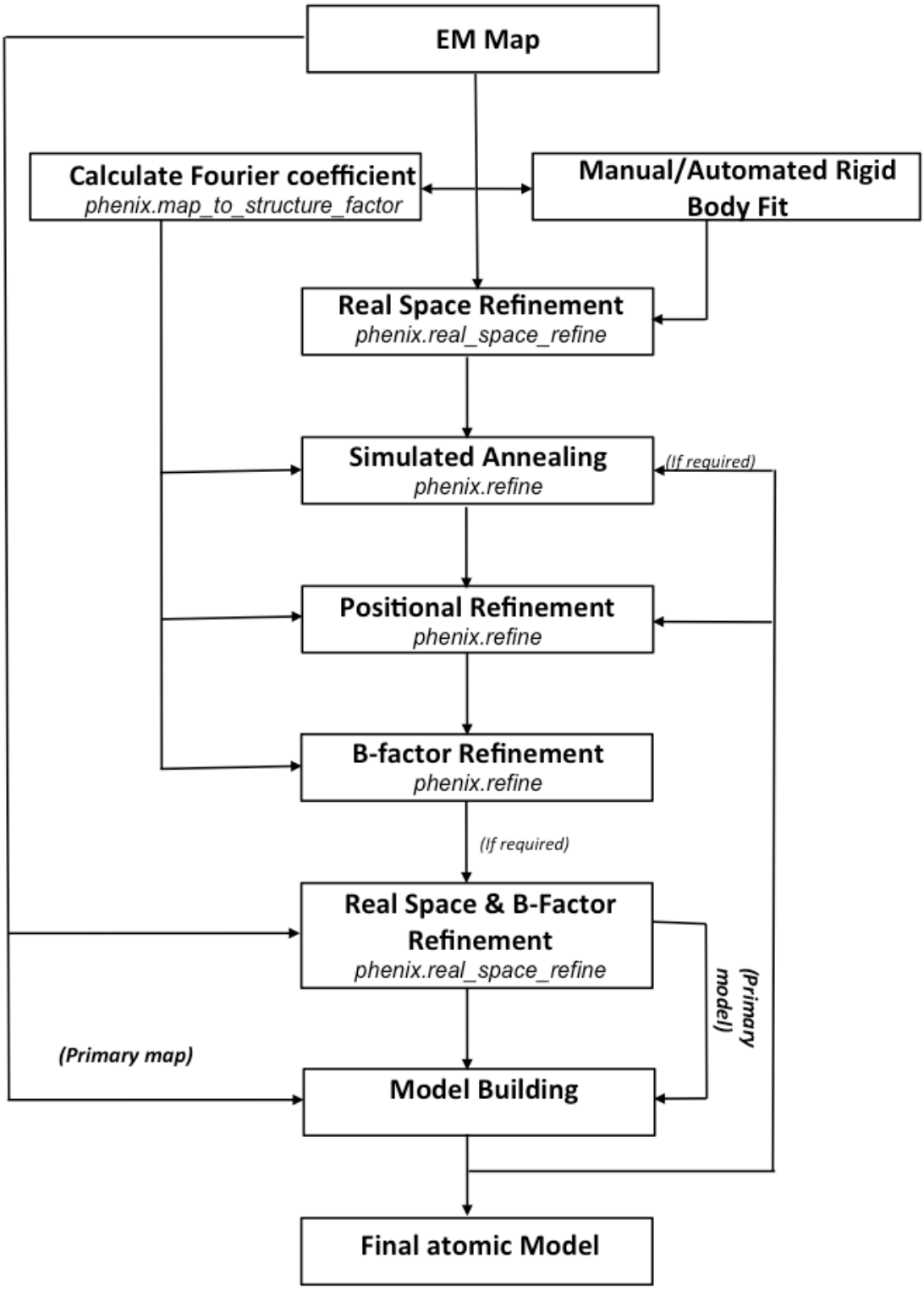

Figure 1

refinement scheme