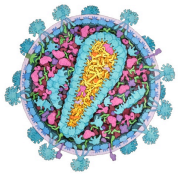
Summary of Chimera Molecular Assemblies Proposal

Grant renewal for NCRR Resource for Biocomputing, Visualization and Informatics

1) Analyzing atomic models of molecular assemblies.



Develop software to make "composite models" e.g. HIV, or a muscle sarcomere, or a transcription factory, using all available structures. Clicking on the 3-d model components will give direct access to structures, multiple sequence alignments and EM maps. This is a 3-d version of David Goodsell's HIV painting:

http://www.rcsb.org/pdb/static.do?p=education_discussion/educational_resources/hiv-animation.html

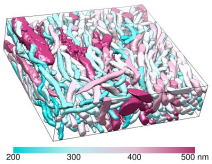
2) Analyzing single-particle EM maps.

Real-time fitting methods mostly for low resolution maps. A) Optimize correlation with full symmetry. B) Iterative fitting and masking away neighbors to get good packing. C) Energy minimization for turns when fitting secondary structure in 5-10 A maps. D) Progressive global search fitting. E) Best fit using NMR or homology ensembles.

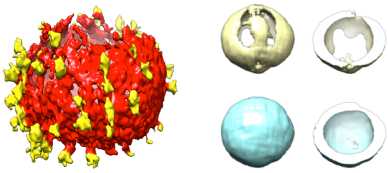
Efficient handling of large maps. F) Display, analyze, save only a few asymmetric units. G) Hidden surface removal (e.g. layers inside viruses). H) Fast transparency rendering. I) Fourier space representation for fast variable resolution display and analysis.

Small-angle x-ray scattering. J) Compute SAXS profiles directly from EM maps for validation.

3) Analyzing structurally heterogeneous assemblies.



Segmentation. A) Multi-resolution watershed and flood-fill. B) Connect two marked locations through highest density (filaments). C) Auto-size and center spheres on vesicles and virus particles.

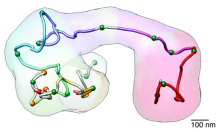


Single-particle tomography: D) Solvent-flattening missing wedge correction. E) Alignment with missing wedge correction.

4) Enabling researchers to share and archive analysis (models and maps).

A) Automatically create an HTML log of Chimera analysis (see example on next page). Includes images, links to data sets (PDB files, maps, sequence alignments, residue lists), Chimera session files, movies, quantitative measurements (correlation, buried area). Organizes hundreds of files of a research project.

B) Document HDF5 map, segmentation, coarse-grain model standards and encourage use by other software.

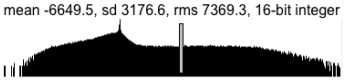


C) 3-dimensional movie recording. BluRay, YouTube.

Rous sarcoma virus capsid compared to HIV

Thu Apr 7 17:48 PDT 2011

1 Opened emd_1862.map (#0), size 255^3, grid spacing 1.27 Å
contour level -363, minimum -17270, maximum 14818
mean -6649.5, sd 3176.6, rms 7369.3, 16-bit integer



2 Opened 1em9.pdb (#1), 2 chains A B, 2257 atoms, x-ray 2.05 Angstroms
rous sarcoma virus capsid protein: n-terminal domain
weight 30.04 KDa

3 Deleted 1091 atoms (chain B), 1em9.pdb (#1)

4 Opened 1eoq.pdb (#2), 1 chain A, 1209 atoms, NMR
rous sarcoma virus capsid protein: c-terminal domain
weight 8.497 KDa

5 Molecular weight 24.04 KDa, 2375 atoms (1em9.pdb (#1), 1eoq.pdb (#2)).

6 Contour level 1239, emd_1862.map (#0)

7 Enclosed volume 1.78e6 Å^3, area 454.6e3 Å^3, emd_1862.map (#0)

8 **Note**: Set contour level to enclose volume 60 * 24 KDa * 1.2 Å^3/Da

9 Fit 1em9.pdb (#1) in emd_1862.map (#0)
correlation 0.7883, average density 4850, 173 of 1166 atoms outside contour
simulated map 10.4 Å, envelope 0.95 mass enclosed, optimized overlap

rock movie
session9.py

10 Segmented emd_1862.map (#0)
regions 360, smoothing steps 3, step size 1 voxel

11 Segmentation grouping emd_1862.map (#0)
regions 120, smoothing level 4 voxels

12 Fit to segment 1eoq.pdb (#2) in region 1966, emd_1862.seg (#3)
correlation 0.8737, masked to region
simulated map 10.4 Å, envelope 0.95 mass enclosed, optimized overlap

13 Spin movie 1em9.pdb (#1), 1eoq.pdb (#2), emd_1862.seg (#3)

spin session13.py

14 Combined 1em9.pdb (#1), 1eoq.pdb (#2) producing combination (#4)

15 Saved ca_fit.pdb (#4) relative to emd_1862.map (#0).

16 Symmetry copies ca_fit.pdb (#4) producing (#5.1-59)
icosahedral 222r, coordinate system emd_1862.map (#0).

17 Simulated map ca_fit.pdb (#4), (#5.1-59) producing molmap res 10.4 (#6)
resolution 10.4, grid spacing 1.27 Å

session17.py

18 Saved map ca_fit_60_r10.4.mrc (#6)

19 Measure correlation ca_fit_60_r10.4.mrc (#6), emd_1862.map (#0).
corr 0.9403, corr about mean 0.6705, within contour level 0.19 (#6).

20 Resampled map ca_fit_60_r10.4.mrc (#6) on grid emd_1862.map (#0) producing (#7).

21 Saved map ca_fit.mrc (#7).

22 Shifted map emd_1862.map (#0) by 6649 producing (#9).

23 Scaled map (#9) by 2.63e-5 producing map (#10).

24 Saved map emd_1862_normal.mrc (#10)

25 Morph maps emd_1862_normal.mrc (#10) to ca_fit.mrc (#7)
producing map (#11).

morph25.mov

26 Difference map ca_fit.mrc (#7), emd_1862_normal.mrc (#10).
Minimum RMS scaling. Contours -0.19 blue, 0.19 red.

spin26.mov

27 **Note**: 2 pentamers

session26.py

28 Molecular surfaces ca_fit.pdb (#4), (#5.1)

rock28.mov

29 Buried area ca_fit.pdb (#4), (#5.1)
solvent accessible area 486.4 Å^2, solvent excluded area 138.0 Å^2.

29 Buried residues ca_fit.pdb (#4), (#5.1)
11 residues, solvent accessible area cutoff 10 Å^2.
Total buried SAS area 303.0 Å^2, average 27.5 Å^2.

residue list

30 Opened 3p05.pdb (#12), x-ray 2.50 Angstroms
5 chains A B C D E, 7918 atoms, 1043 residues
pentameric hiv-1 ca
weight 105 KDa

31 Fit 3p05.pdb (#12) in emd_1862.map (#0)
correlation 0.53, average density 3548, 2160 of 7918 atoms outside contour
simulated map 10.4 Å, envelope 0.95 mass enclosed, optimized overlap

32 **Note**: Superposition of HIV and RSV pentamers.
HIV 3p05.pdb (light blue and pink)
RSV ca_fit.pdb (blue and red)
Helices align. HIV squeezed toward center.

spin32.mov
session32.py

Fri Apr 8 00:15 PDT 2011

Figure: Illustration of web log file fitting Rous sarcoma virus (RSV) map and comparing HIV virus capsid pentamer. Records analysis done in Chimera fitting of N-terminal and C-terminal capsid protein domains using segmentation, applying symmetry, visualizing residual density, measuring C-terminal dimerization domain buried area, finding dimer interface residues, and aligning and comparing HIV pentamer. Links to maps, PDB files, the segmentation, session snapshots, lists of residues, spin, rock and morph movies, and larger images are provided. Quantitative results are recorded. The proposed Chimera logging tool will create this type of web page, semi-automatically, to record interactive analysis sessions. This example was mocked up by hand.

Possible ideas to add to the proposal

1) EM model validation tools
2) Animation production tools
3) Single-particle segmentation, impose symmetry
4) Atomic model building in 3 – 4.5 Angstrom EM maps (Gorgon).