Validation

Pavel Afonine

phenix-online.org

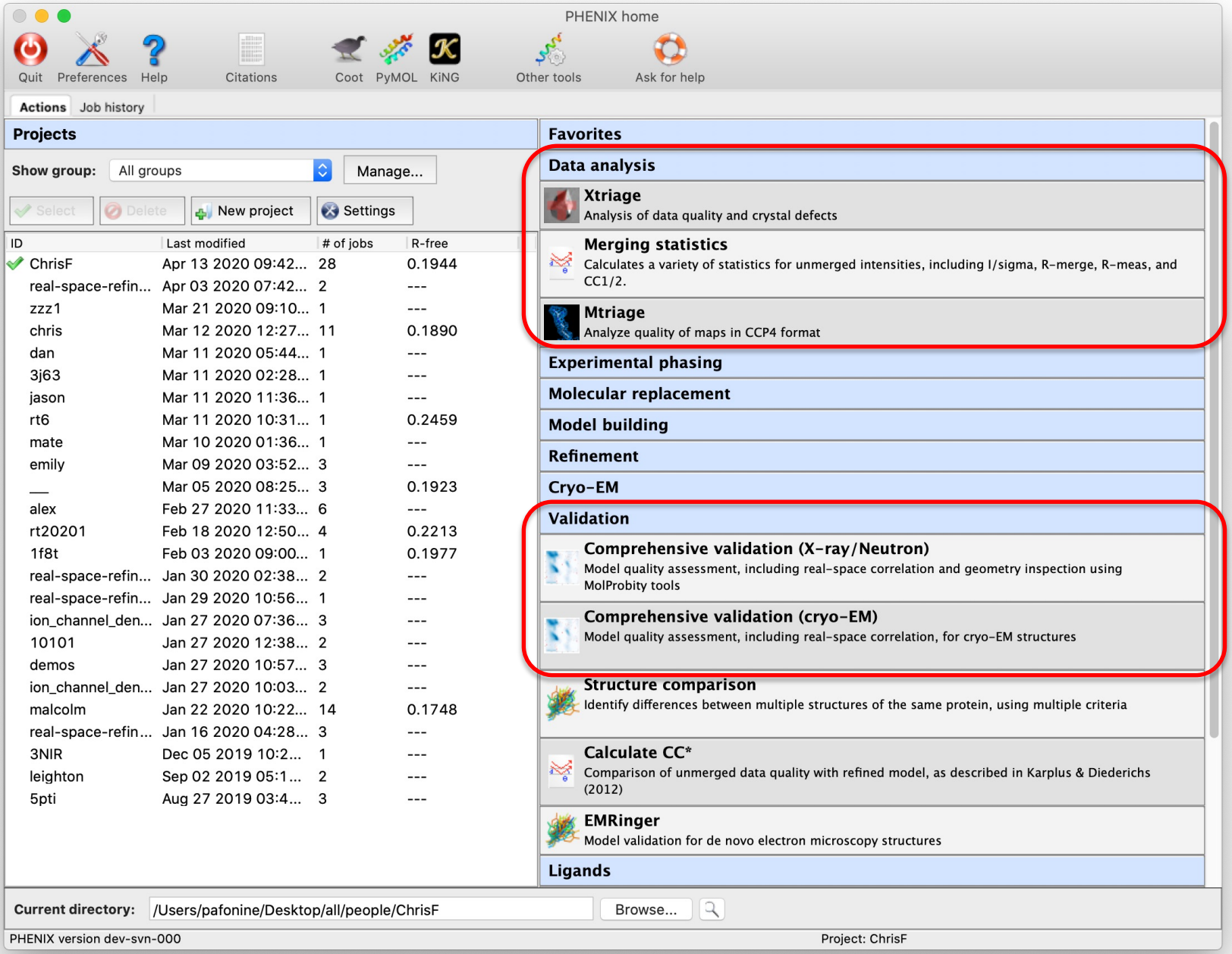
lbl.gov

qrefine.com

Hosted by the Oklahoma COBRE in Structural Biology

March 19th 2024

Validation tools in Phenix



Validation (cryo-EM)

Comprehensive validation (cryo-EM) (Project: real-space-refine-5ljv_0)

Preferences | Help | Run | Abort | Ask for help

Input/Output | ValidationCryoEM_1

Run status | Summary | MolProbity | Model vs. Data | Data

Open in Coot

Export Table 1

Model

Composition (#)	
Chains	2
Atoms	2500 (Hydrogens: 0)
Residues	Protein: 325 Nucleotide: 0
Water	0
Ligands	MG: 1 ADP: 1
Bonds (RMSD)	
Length (Å) (# > 4σ)	0.029 (146)
Angles (°) (# > 4σ)	2.853 (122)
MolProbity score	3.14
Clash score	19.06
Ramachandran plot (%)	
Outliers	3.10
Allowed	7.12
Favored	89.78
Rotamer outliers (%)	11.57
Cβ outliers (%)	3.68
Peptide plane (%)	
Cis proline/general	5.6/0.0
Twisted proline/general	11.1/0.7
CaBLAM outliers (%)	2.18
ADP (B-factors)	
Iso/Aniso (#)	2500/0
min/max/mean	
Protein	30.26/493.42/109.69
Nucleotide	---
Ligand	57.57/99.69/75.15
Water	---
Occupancy	
Mean	1.00
occ = 1 (%)	100.00
0 < occ < 1 (%)	0.00
occ > 1 (%)	0.00

Data

Box		
Lengths (Å)	50.92, 68.34, 83.08	
Angles (°)	90.00, 90.00, 90.00	
Supplied Resolution (Å)	3.6	
Resolution Estimates (Å)	Masked	Unmasked
d FSC (half maps; 0.143)	---	---
d 99 (full/half1/half2)	3.7/---/---	3.1/---/---
d model	3.7	3.7
d FSC model (0/0.143/0.5)	3.4/3.5/3.6	3.4/3.6/3.9
Map min/max/mean	-0.42/0.80/0.03	

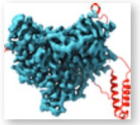
Model vs. Data

CC (mask)	0.83
CC (box)	0.55
CC (peaks)	0.34
CC (volume)	0.83
Mean CC for ligands	0.86

Idle | Project: real-space-refine-5ljv_0

RESEARCH PAPERS

Acta Cryst. (2018). D**74**, 814-840
https://doi.org/10.1107/S2059798318009324
Cited by 71
OPEN



New tools for the analysis and validation of cryo-EM maps and atomic models

P. V. Afonine, B. P. Klaholz, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C. Terwilliger, P. D. Adams and A. Urzhumtsev

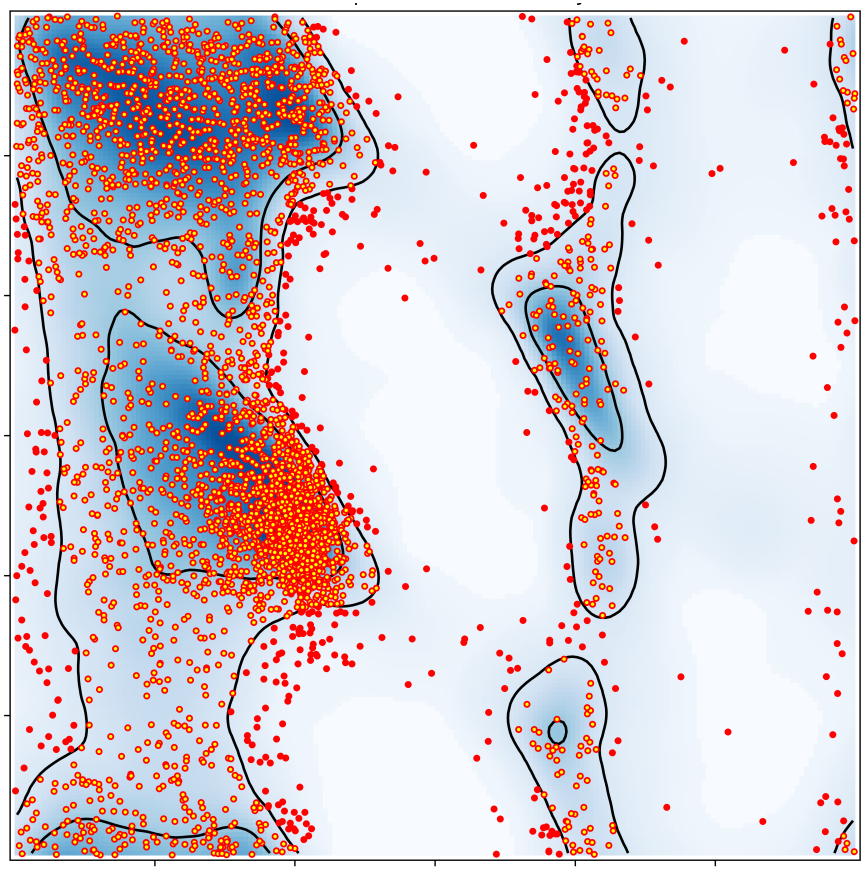
Recent advances in the field of electron cryomicroscopy (cryo-EM) have resulted in a rapidly increasing number of atomic models of biomacromolecules that have been solved using this technique and deposited in the Protein Data Bank and the Electron Microscopy Data Bank. Similar to macromolecular crystallography, validation tools for these models and maps are required. While some of these validation tools may be borrowed from crystallography, new methods specifically designed for cryo-EM validation are required. Here, new computational methods and tools implemented in *PHENIX* are discussed, including d_{99} to estimate resolution, *phenix.auto_sharpen* to improve maps and *phenix.mtriage* to analyze cryo-EM maps. It is suggested that cryo-EM half-maps and masks should be deposited to facilitate the evaluation and validation of cryo-EM-derived atomic models and maps. The application of these tools to deposited cryo-EM atomic models and maps is also presented.

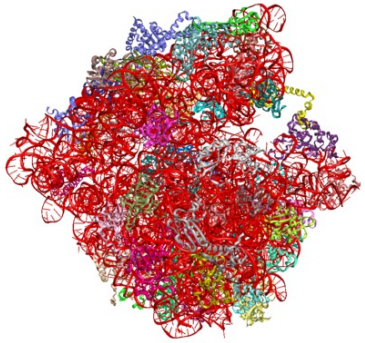
Keywords: cryo-EM; **atomic models**; **model quality**; **data quality**; **validation**; **resolution**.

Read article **Similar articles**

Validation: why to do?

(2019) Nature 570: 400-404 | PDB: 6o9j | EMDB: 0661 | 3.9Å



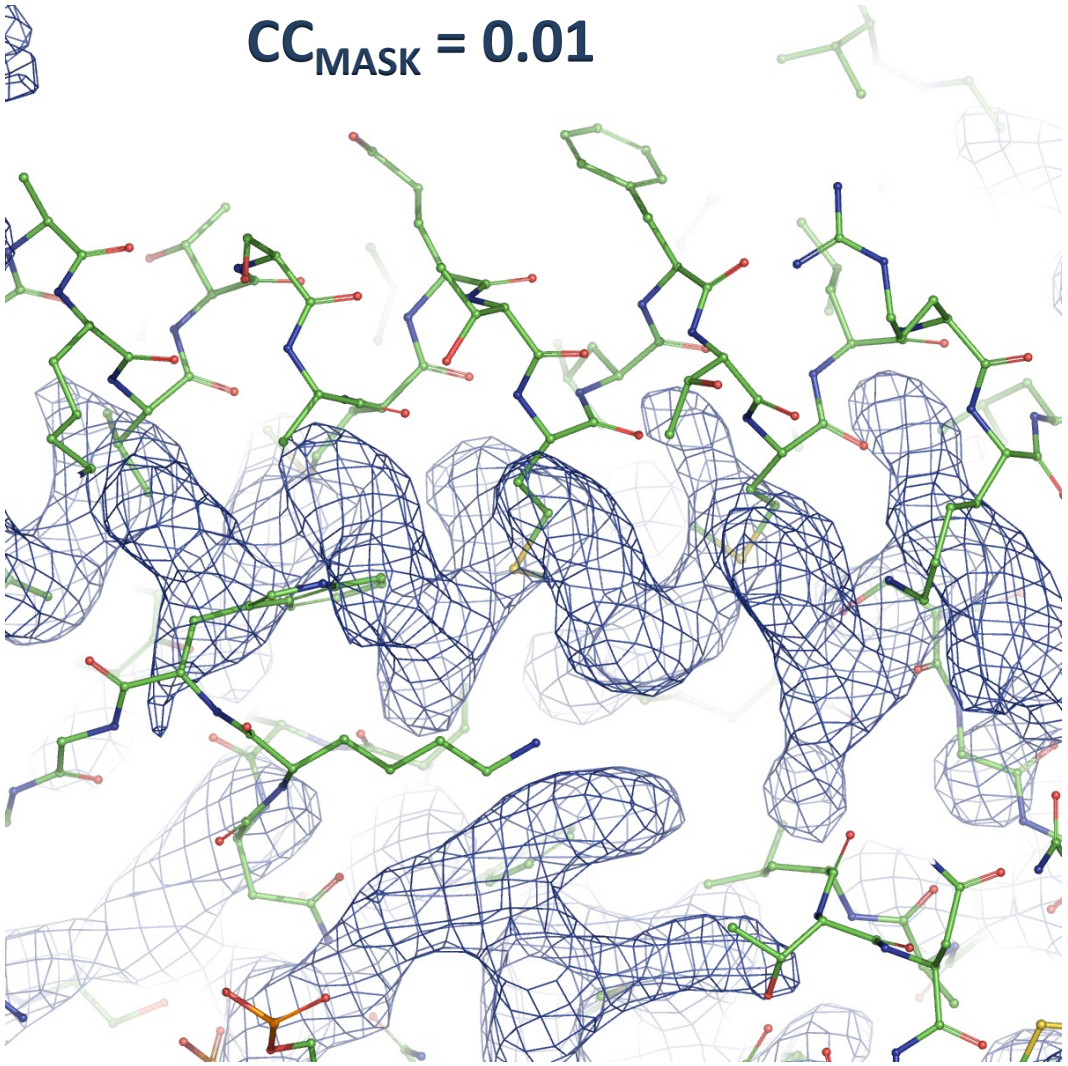


Metric	6o9j	Expected
Clashscore	70	Less than 10
Ramachandran favored, %	59	More than 98
Ramachandran outliers, %	15	0
Rotamer outliers, %	23	0
C_β deviations, %	0.5	0

Examples of very poor model-to-map fit

PDB: 8gwb | EMDB: 34308 | 2.8 Å | Cell (2022) 185: 4347-4360

Chain	CC_{MASK}
A	0.01
B	0.02
C	0
D	0.01
I	0.04
J	0
F	0.12
E	0.08
G	0.1
M	0.16
A	0
F	-0.13
E	0.16
A	0.1
G	0.15
M	0.19

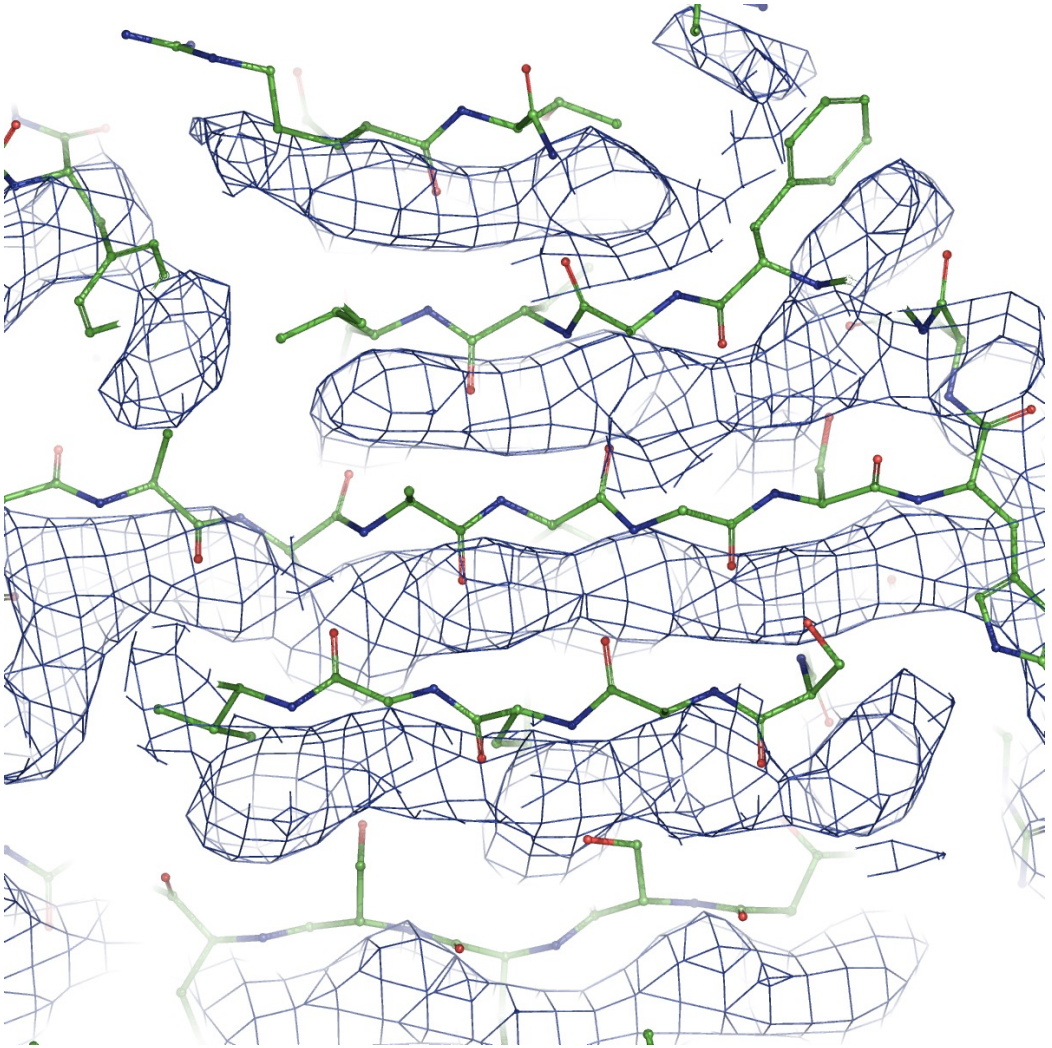


Examples of very poor model-to-map fit

PDB: 7xov | EMDB: 33360 | 3 Å | Cell Discov (2022) 8: 55-55

Chain	CC_{MASK}
A	0.04
B	-0.01
G	0.18
N	0.06
R	0.03
R	-0.02

CC_{MASK} = 0.02

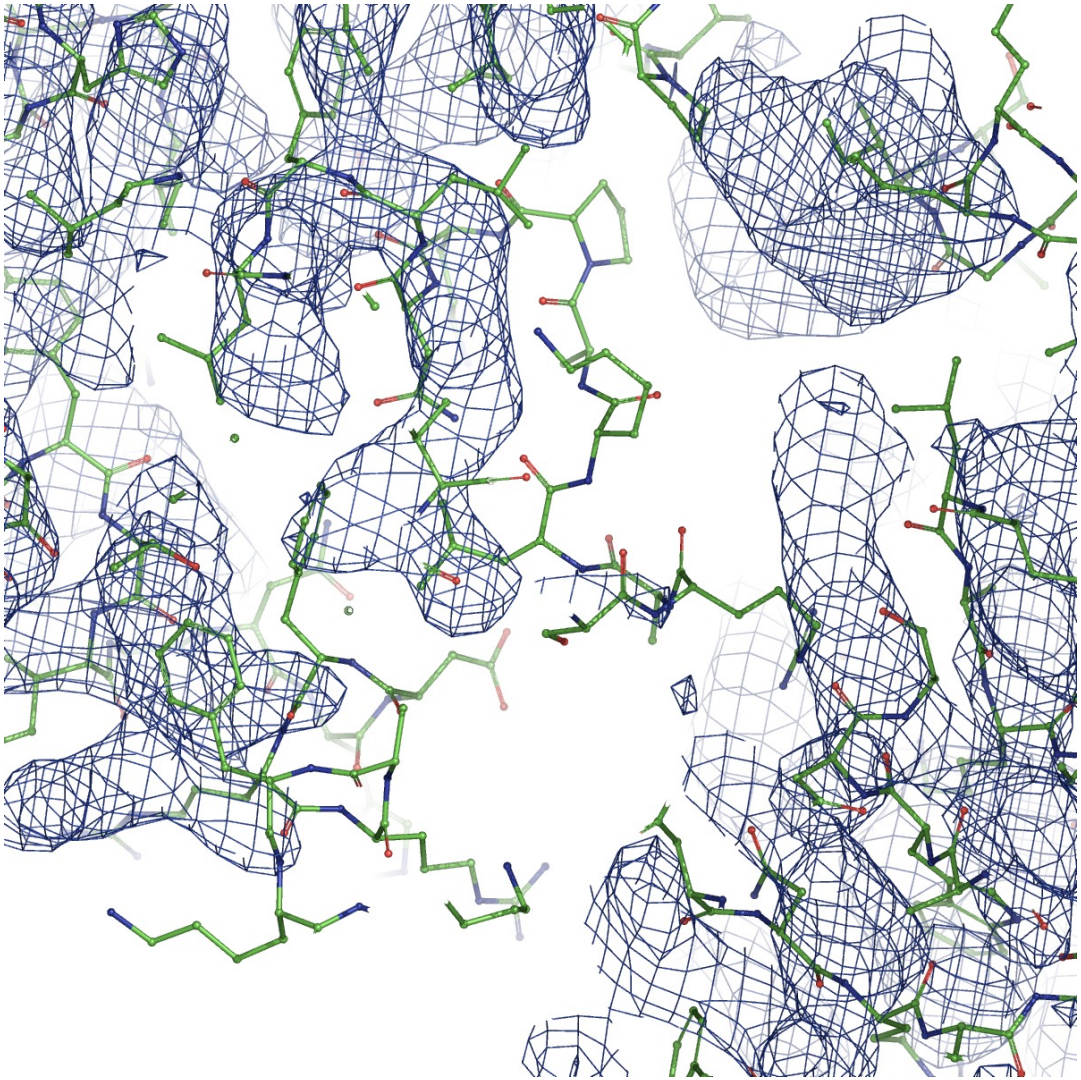


Examples of very poor model-to-map fit

PDB: 7w6p | EMDB: 32331 | 3.5 Å | Science (2022) 377: 7065-7065

Chain	CC_{MASK}
A	0.09
B	0.11
G	0.12
H	0.07
R	0.16
R	-0.08

CC_{MASK} = 0.1

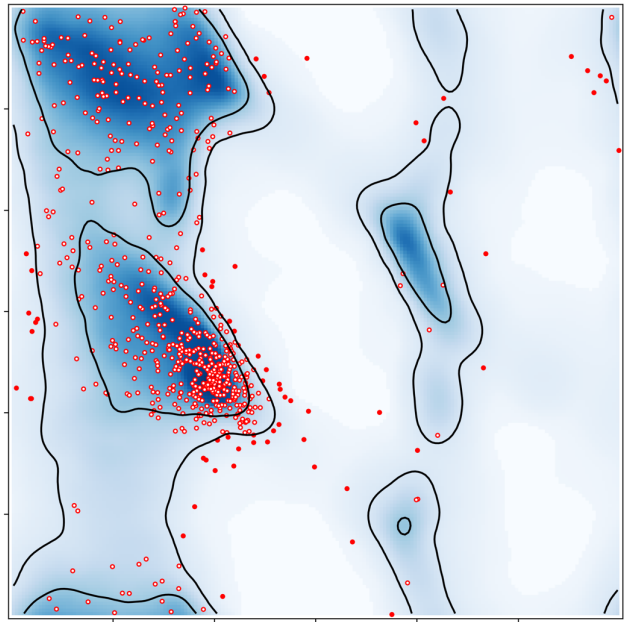


Model

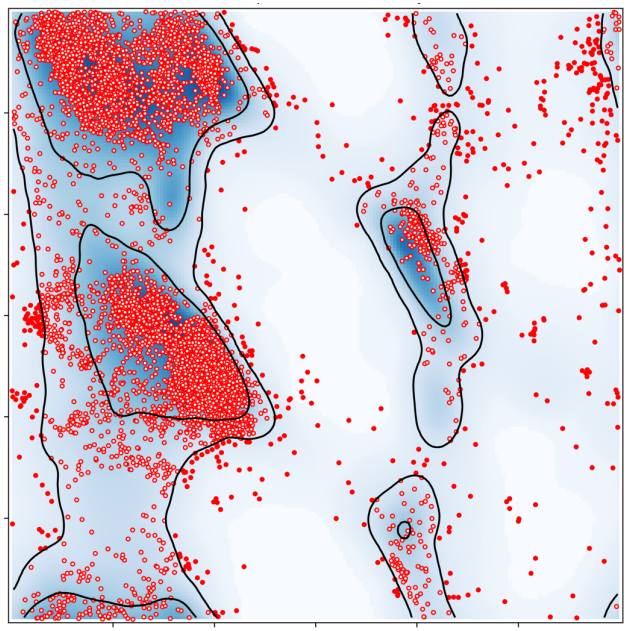
	6kio	6kiq	6o9j	7ase
Bond/angle	0.04/3.4	0.04/3.7	0.01/1.3	0.02/2.2
Clashscore	11	12	55	9
Rotamer outl., %	8	15	23	3
Cb deviations, %	5	16	0.5	1.4
Ramachandran, % favored outliers	 74 7	 70 11	 59 15	 79 7
Resolution (Å)	3.9	3.6	3.9	3.3
Published in	Nature Comm.	Nature Comm.	Nature	Cell
Year	2020	2020	2019	2020

Model

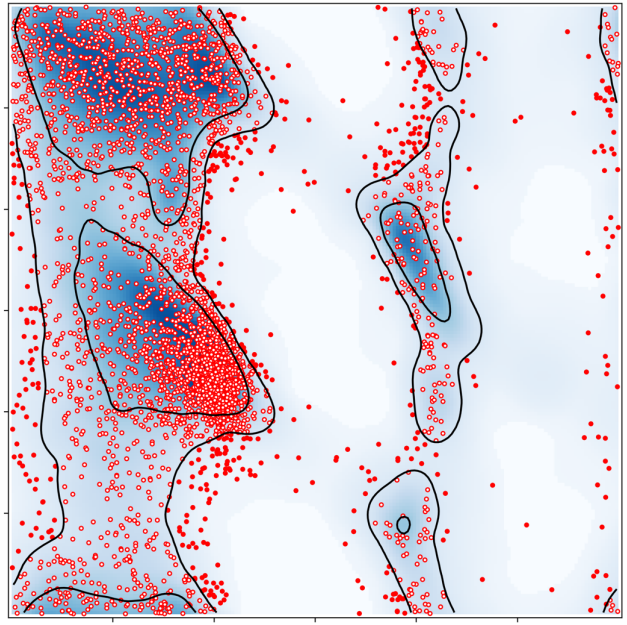
6kiq



7ase



6o9j



6kio