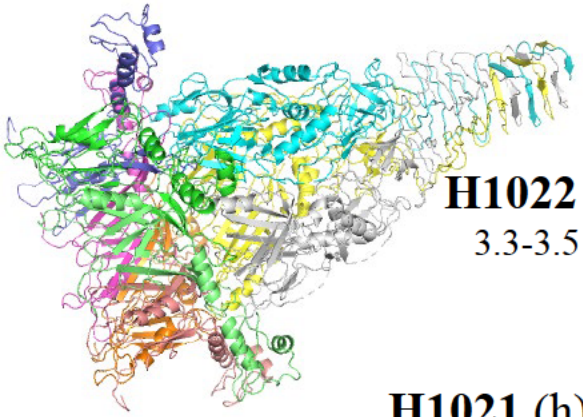
Cryo-EM Targets in CASP13

CASP:
Image redacted

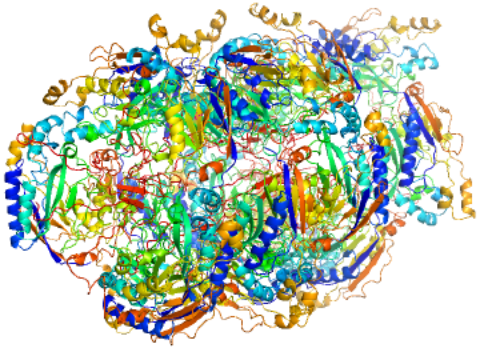
T0990 (h)
4.0 Å



H1022 (h)
3.3-3.5 Å

90

T0995 (e)
3.15 Å

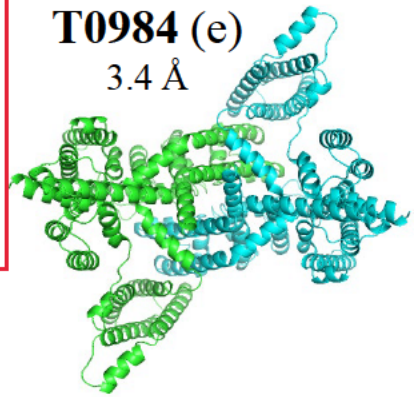


90

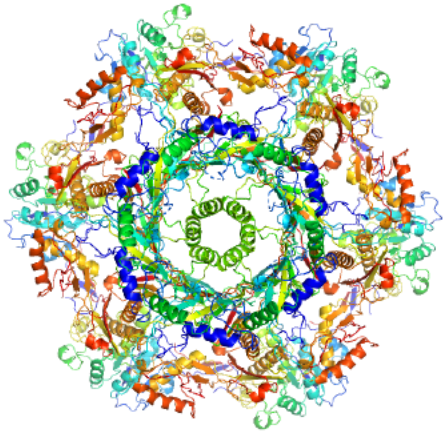
CASP:
Image redacted

T1020 (e)
3.3 Å

T0984 (e)
3.4 Å



H1021 (h)
variable



Andriy Kryshtafovych, Bohdan Monastyrskyy *(UC Davis)*
Maya Topf, Sony Malhotra *(Birkbeck, U London)*

Cryo-EM Targets

T0984 (Xiaochen Bai, UT Southwestern Medical Center, Dallas)

T0990 (Hong Zhou, UCLA)

T0995 (Bryan Trevor Sewell, University of Cape Town)

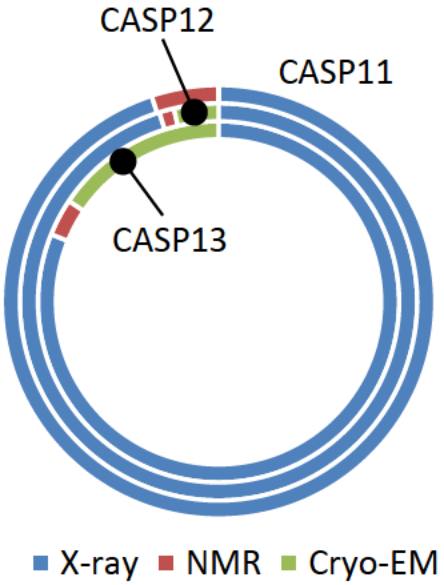
T0996 (Damian Ekiert, former UCSF, now at Skirball Institute, New York)

T1020 (Oliver Clarke, Columbia University)

H1021 (Ambroise Desfosses, Institut de Biologie Structurale, Grenoble)

H1022 (Ambroise Desfosses, Institut de Biologie Structurale, Grenoble)

H1023 (Adam Frost, UCSF) - no structure

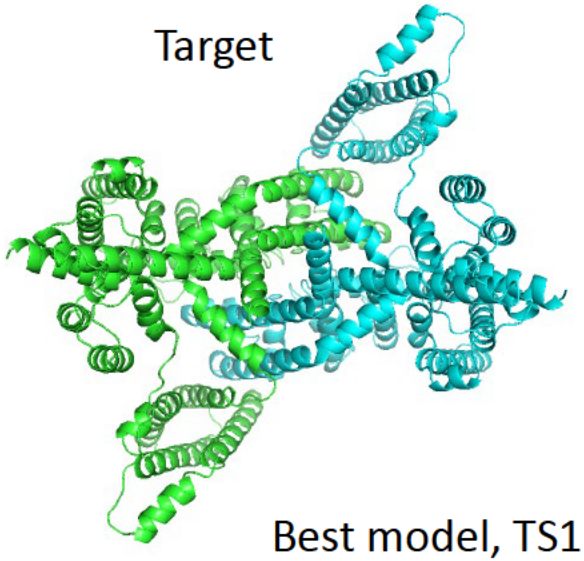


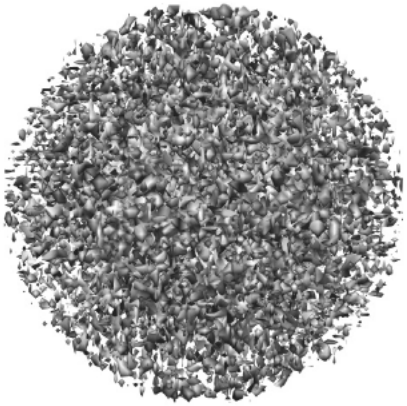
Target T0984(o)

a TM protein (11 helices), dimer, easy, 752 res. in each chain

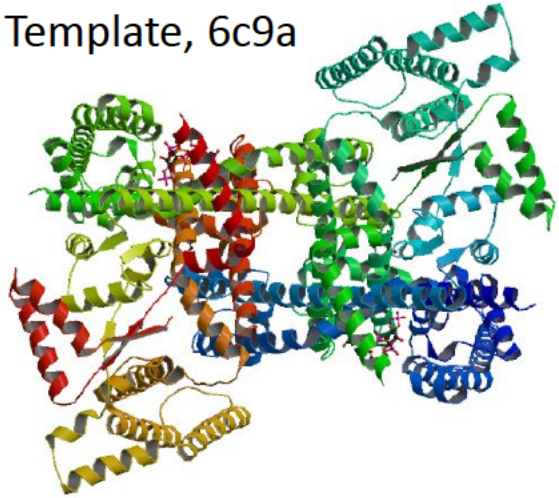
Density map : 3.4 Å

Target



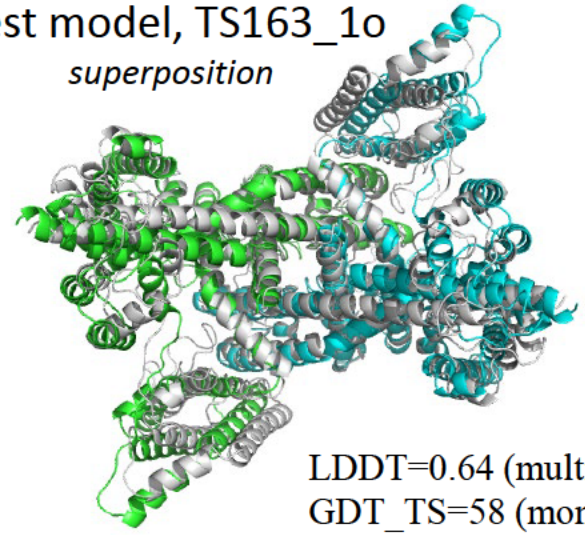


Template, 6c9a



Best model, TS163_1o

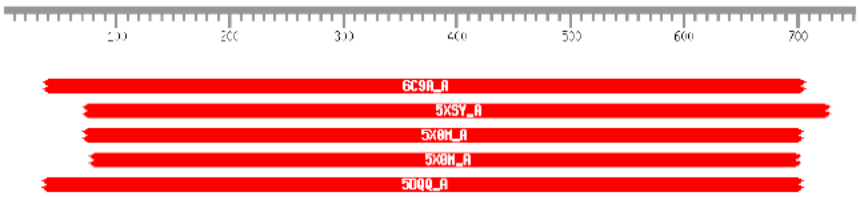
superposition



LDDT=0.64 (multimeric)

GDT_TS=58 (monomeric)

HHSearch alignment



Target T0996(o)

hexamer, medium difficulty, 848 res. in 1 subunit

Density maps (12 different): 3-3.5 Å

CASP:
Images redacted

Target T0996(o)

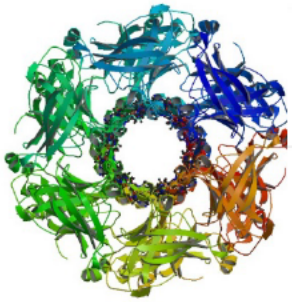
hexamer, medium difficulty, 848 res. in 1 subunit

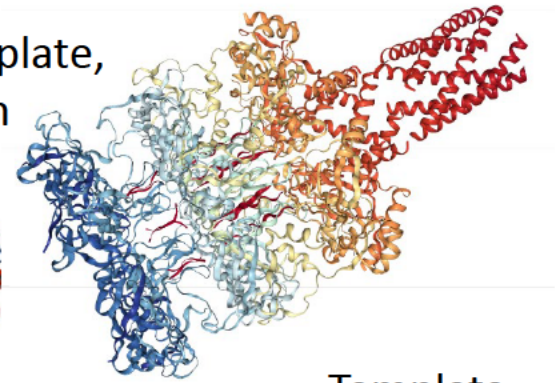
90

Target

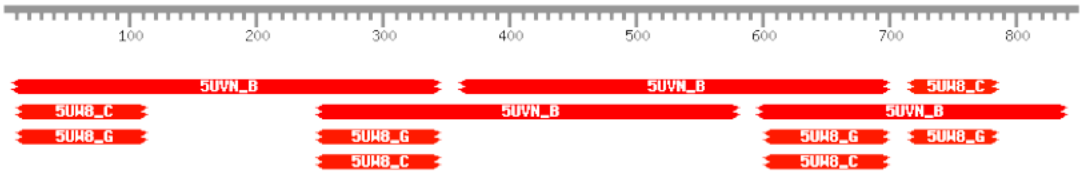
CASP:
Images redacted

Template,
5uvn

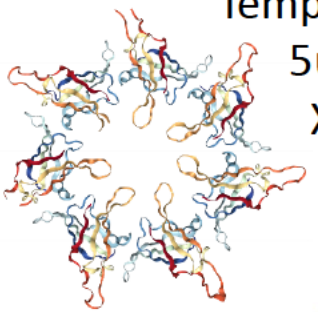




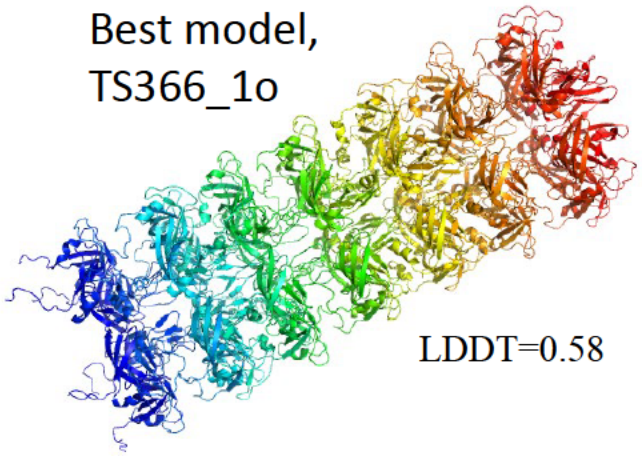
HHSearch alignment



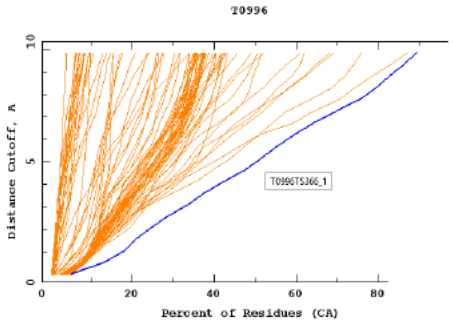
Template,
5uw8,
X-ray



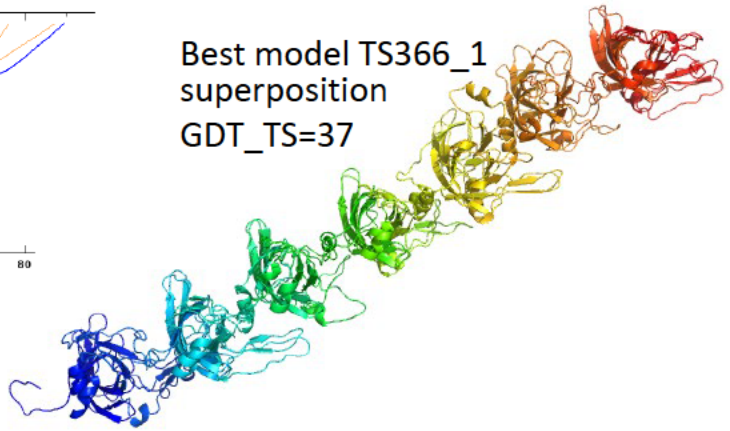
Best model,
TS366_1o



LDDT=0.58

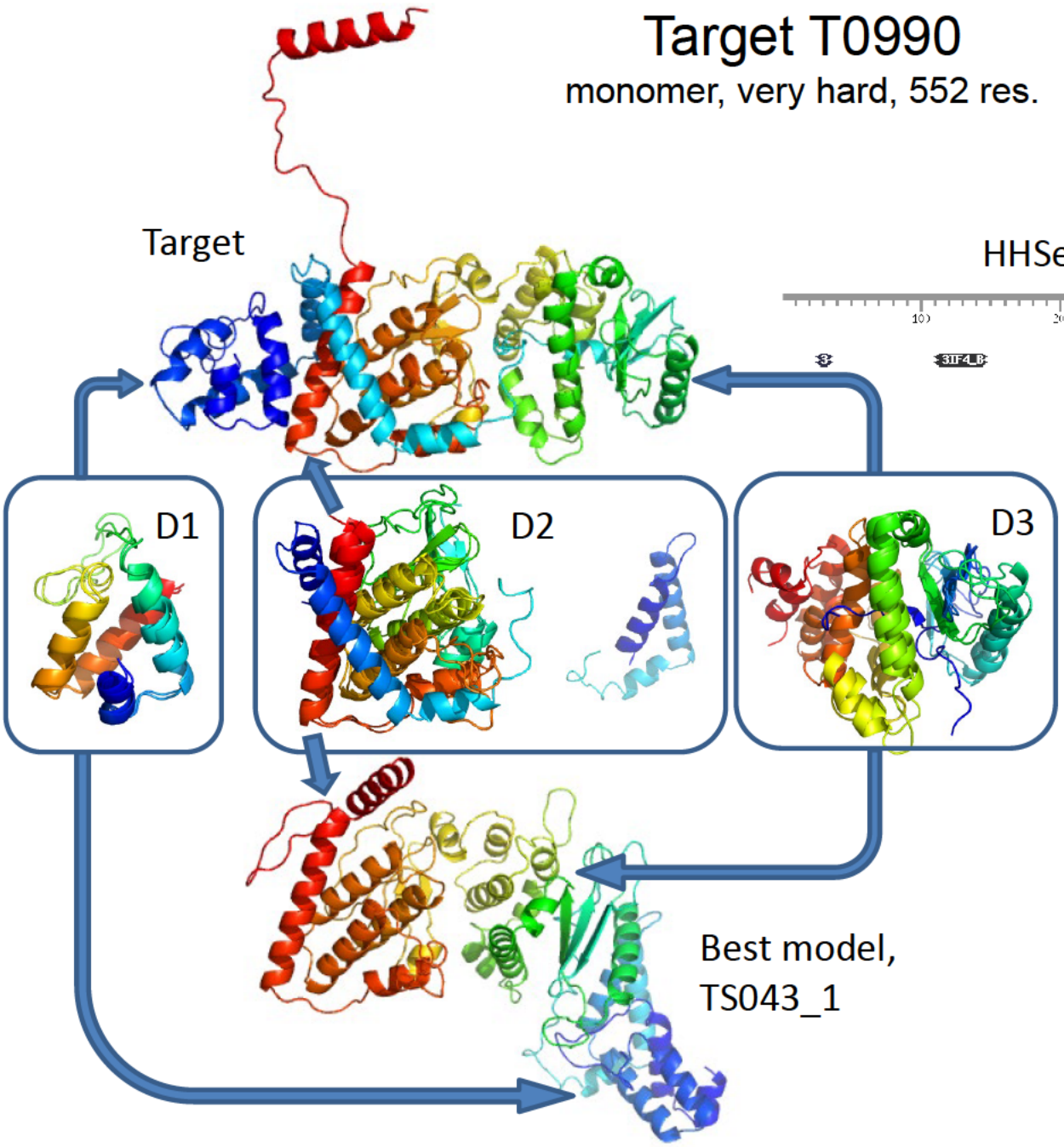


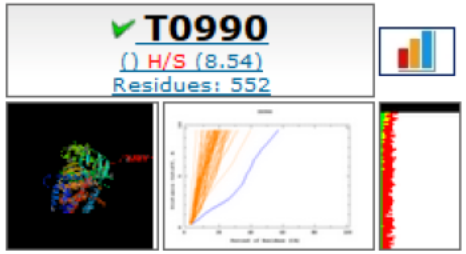
Best model TS366_1
superposition
GDT_TS=37



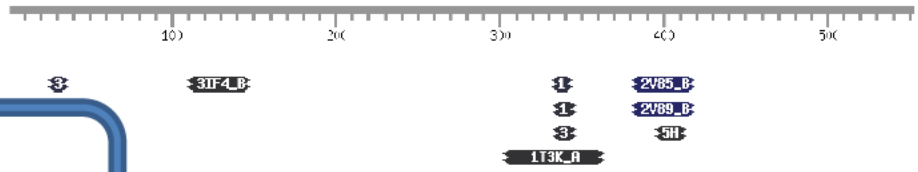
Target T0990

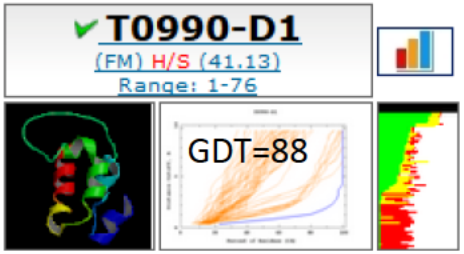
monomer, very hard, 552 res.

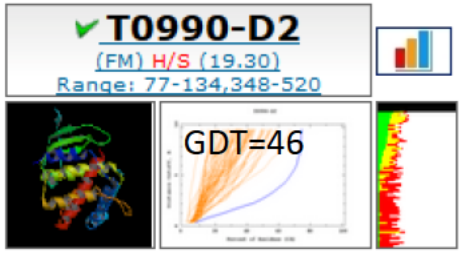


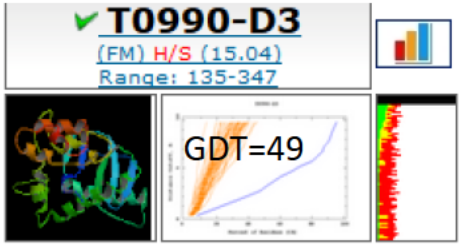


HHSearch alignment

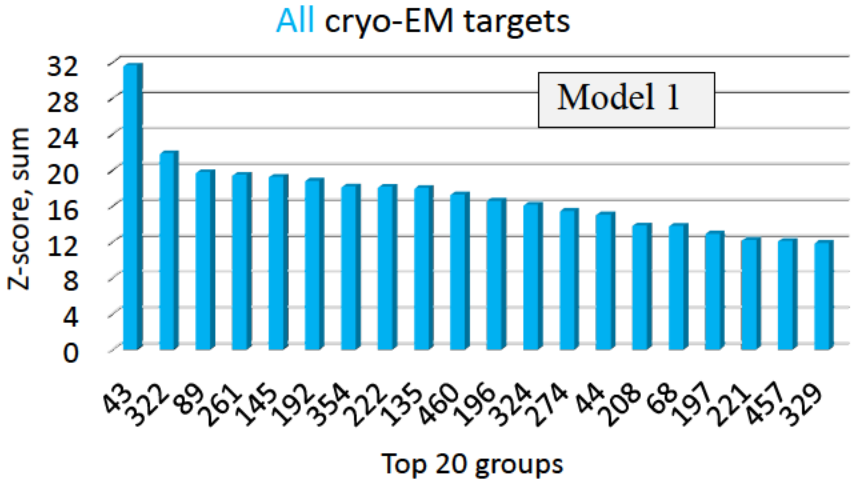


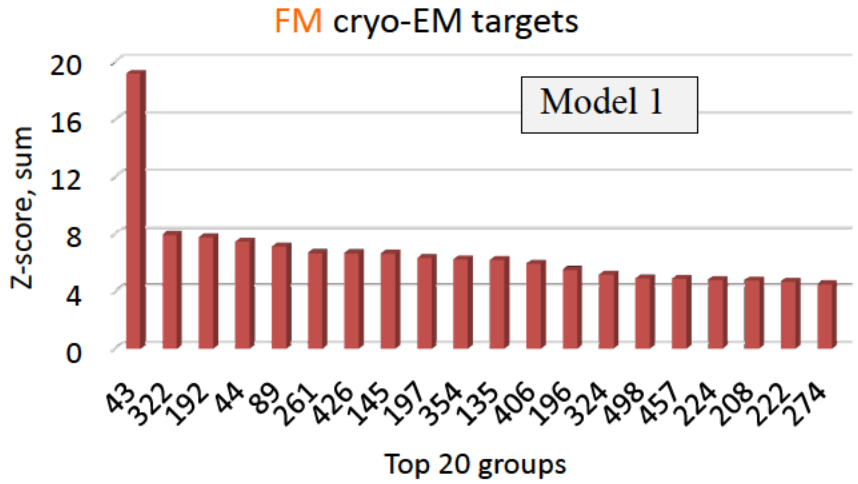


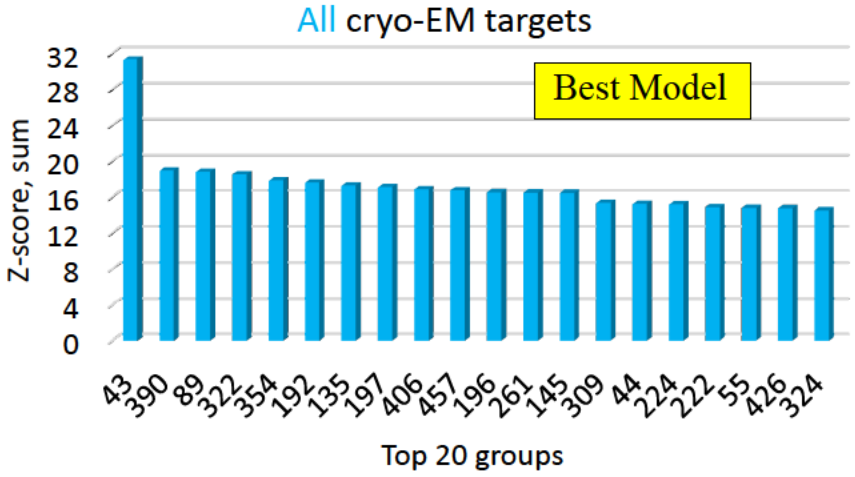


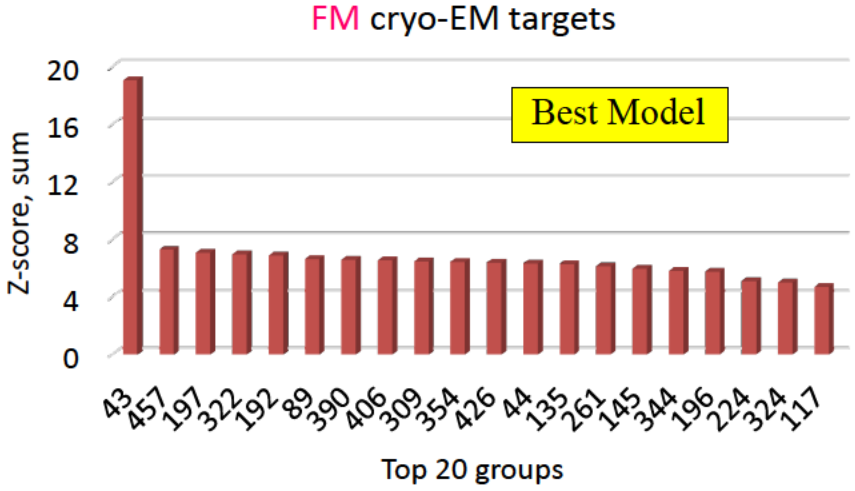


CASP Models on Cryo-EM targets
Evaluation vs Reference Structures
Tertiary structure (3GDT_TS + LDDT + CADaa + SG)

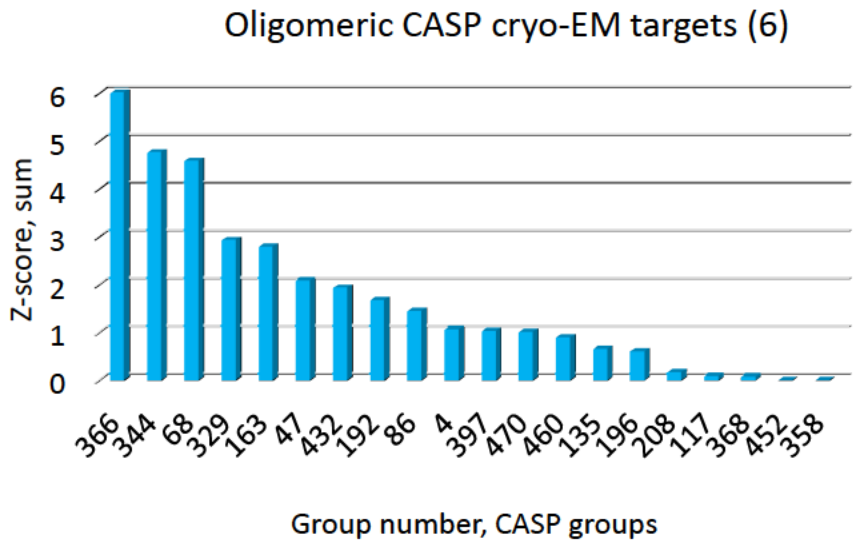


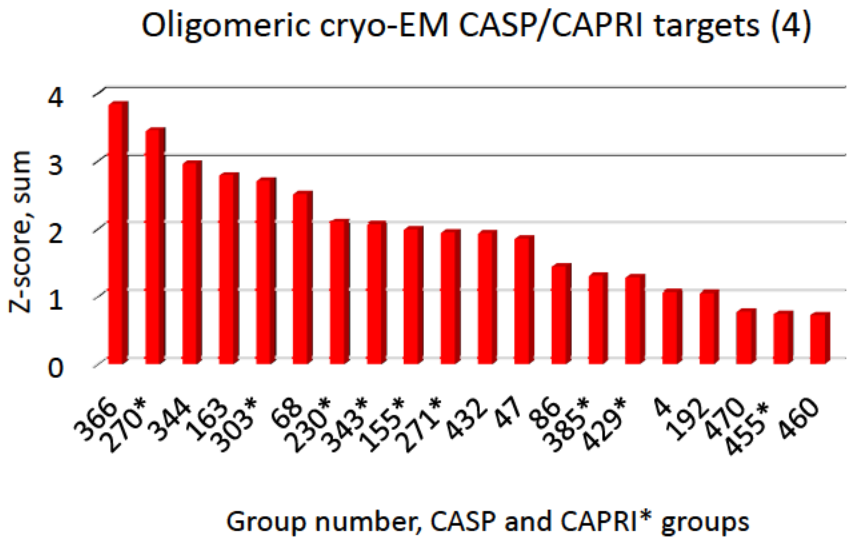






CASP Models on Cryo-EM targets
Evaluation vs Reference Structures
Quaternary structure (QSglob + LDDTo + F1 + JaccCoef)





Model 1

CASP Models on Cryo-EM targets
Evaluation vs Maps

Placing models in map's frame of reference
(phenix.dock_in_map, Tom Terwilliger
UCSF Chimera)

CASP:
Images redacted

CASP Models on Cryo-EM targets
Evaluation vs Maps
Model to map fit

TEMPy (Agnel Joseph, Maya Topf)
CCC, LAP, MI, SMOC, MI_overlap, CCC_overlap

PHENIX.model_vs_map (Paul Adams)
3 variants of cross-correlation scores

PHENIX.chain_compare (Tom Terwilliger)

EMRinger (Ben Barad, James Fraser)
Local and global EMRinger scores

CASP Models on Cryo-EM targets
Evaluation vs Maps
Model to map fit

Contents lists available at ScienceDirect

Journal of Structural Biology

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

This article is part of the Special Issue on the 2016 CryoEM Challenges

Evaluation system and web infrastructure for the second cryo-EM model challenge☆

Andriy Kryshtafovych[a,*], Paul D. Adams[b,c], Catherine L. Lawson[d], Wah Chiu[e,f]

[a] Genome Center, University of California, Davis, 451 Health Sciences Drive, Davis, CA 95616, USA
[b] Molecular Biophysics & Integrated Bioimaging, LBNL, CA 94720, USA
[c] Department of Bioengineering, University of California Berkeley, CA 94720, USA
[d] Institute for Quantitative Biomedicine and Research Collaboratory for Structural Bioinformatics, Rutgers, The State University of New Jersey, 174 Frelinghuysen Road, Piscataway, NJ 08854, USA
[e] Departments of Bioengineering and Microbiology & Immunology, Stanford University, Stanford, CA 94305-5447, USA
[f] Division of CryoEM and Bioimaging, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

ARTICLE INFO

Keywords:
Cryo-EM
Model challenge
Protein structure modeling
Protein structure verification

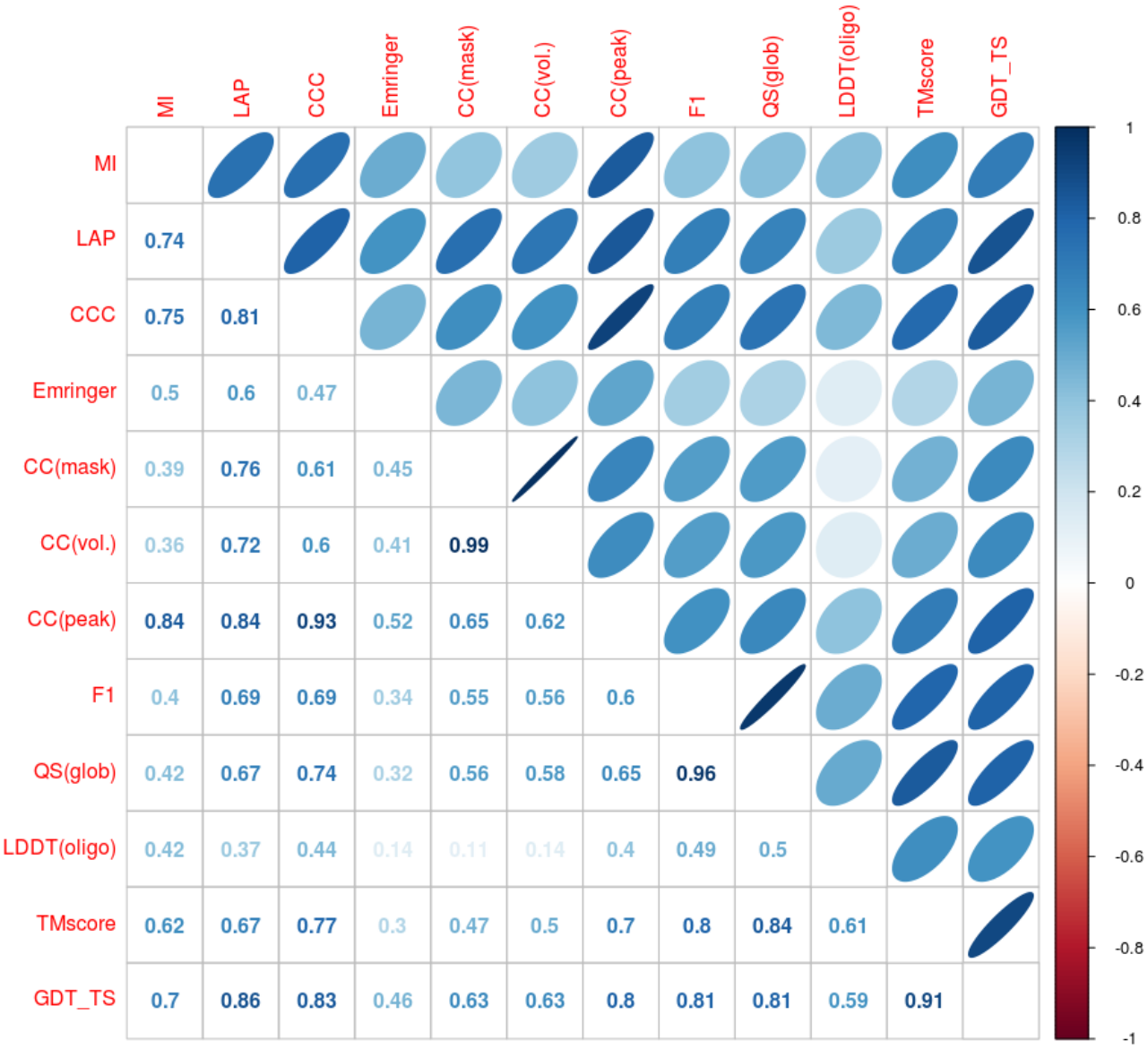
ABSTRACT

An evaluation system and a web infrastructure were developed for the second cryo-EM model challenge. The evaluation system includes tools to validate stereo-chemical plausibility of submitted models, check their fit to the corresponding density maps, estimate their overall and per-residue accuracy, and assess their similarity to reference cryo-EM or X-ray structures as well as other models submitted in this challenge. The web infrastructure provides a convenient interface for analyzing models at different levels of detail. It includes interactively sortable tables of evaluation scores for different subsets of models and different sublevels of structure organization, and a suite of visualization tools facilitating model analysis. The results are publicly accessible at http://model-compare.emdatabank.org.

0. Introduction

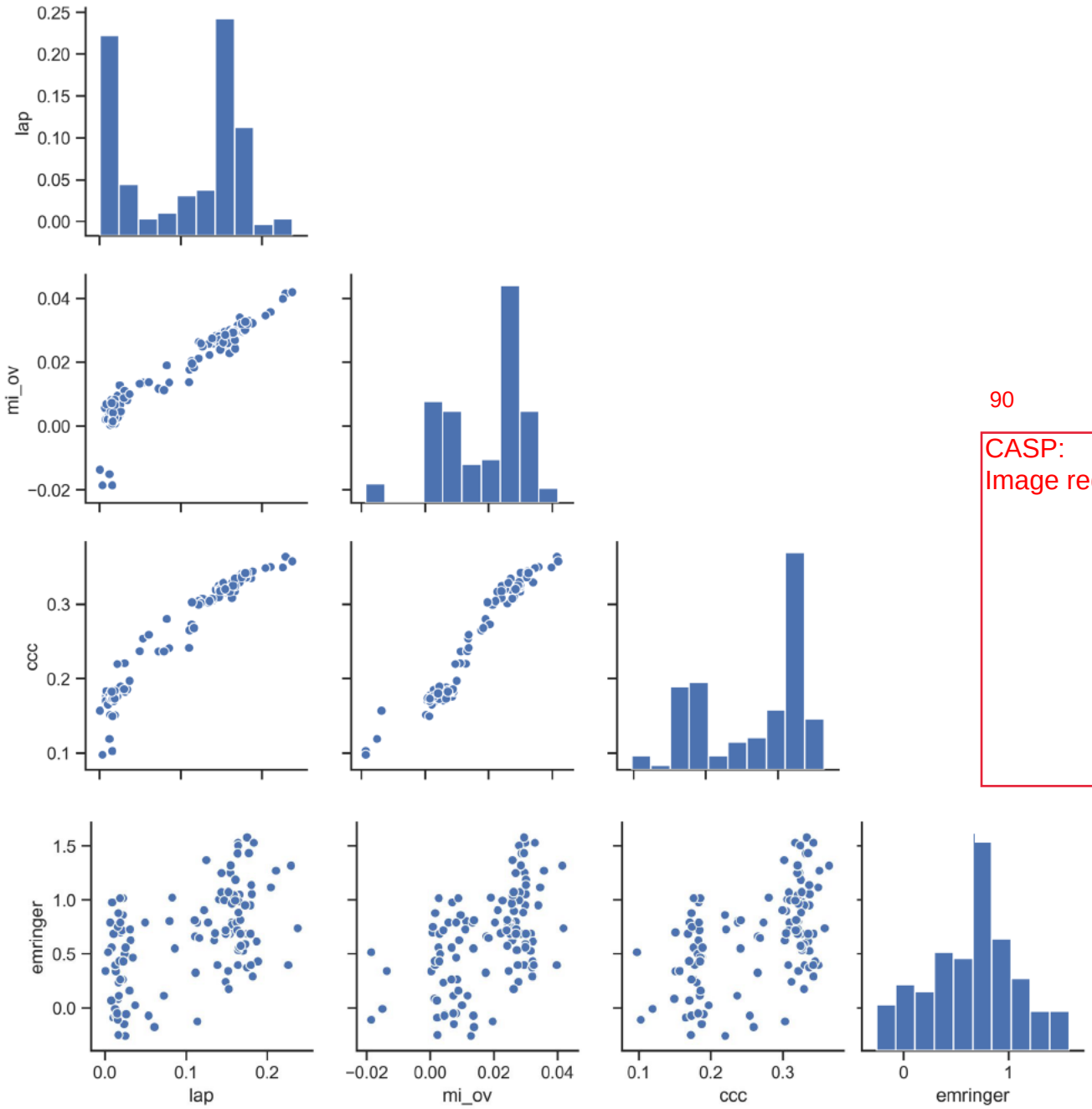
The second cryo-Electron Microscopy Model Challenge (EMMC) was organized to discuss preliminary outcomes. The challenge culminated in a joint participants, assessors and organizers meeting in October 2017, where the results were reviewed and discussed, and

Overall correlation



	MI	LAP	CCC	Emringer	CC(mask)	CC(vol.)	CC(peak)	F1	QS(glob)	LDDT(oligo)	TMscore	GDT_TS
MI												
LAP	0.74											
CCC	0.75	0.81										
Emringer	0.5	0.6	0.47									
CC(mask)	0.39	0.76	0.61	0.45								
CC(vol.)	0.36	0.72	0.6	0.41	0.99							
CC(peak)	0.84	0.84	0.93	0.52	0.65	0.62						
F1	0.4	0.69	0.69	0.34	0.55	0.56	0.6					
QS(glob)	0.42	0.67	0.74	0.32	0.56	0.58	0.65	0.96				
LDDT(oligo)	0.42	0.37	0.44	0.14	0.11	0.14	0.4	0.49	0.5			
TMscore	0.62	0.67	0.77	0.3	0.47	0.5	0.7	0.8	0.84	0.61		
GDT_TS	0.7	0.86	0.83	0.46	0.63	0.63	0.8	0.81	0.81	0.59	0.91	

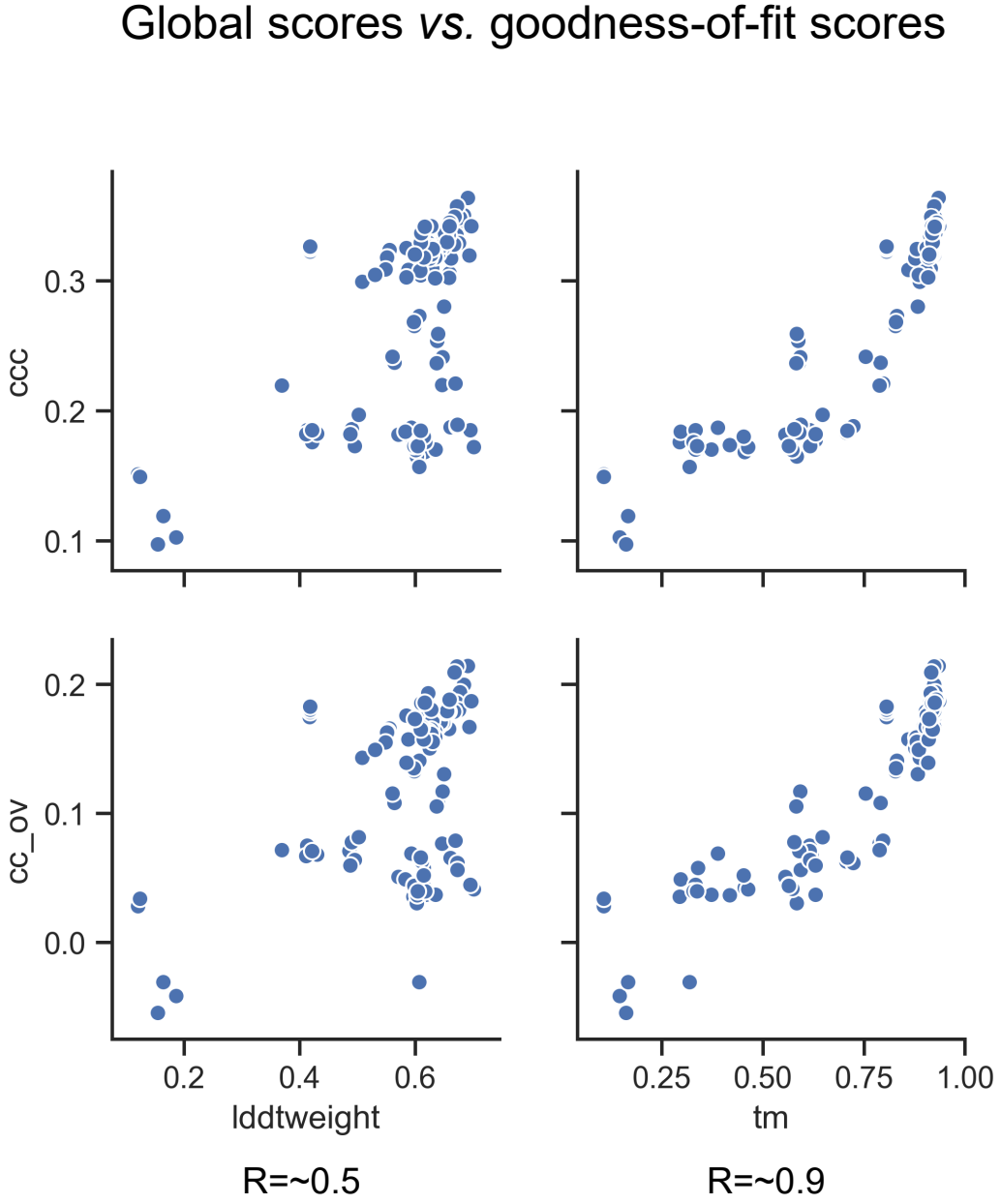
T1020o - 3.3 Å resolution



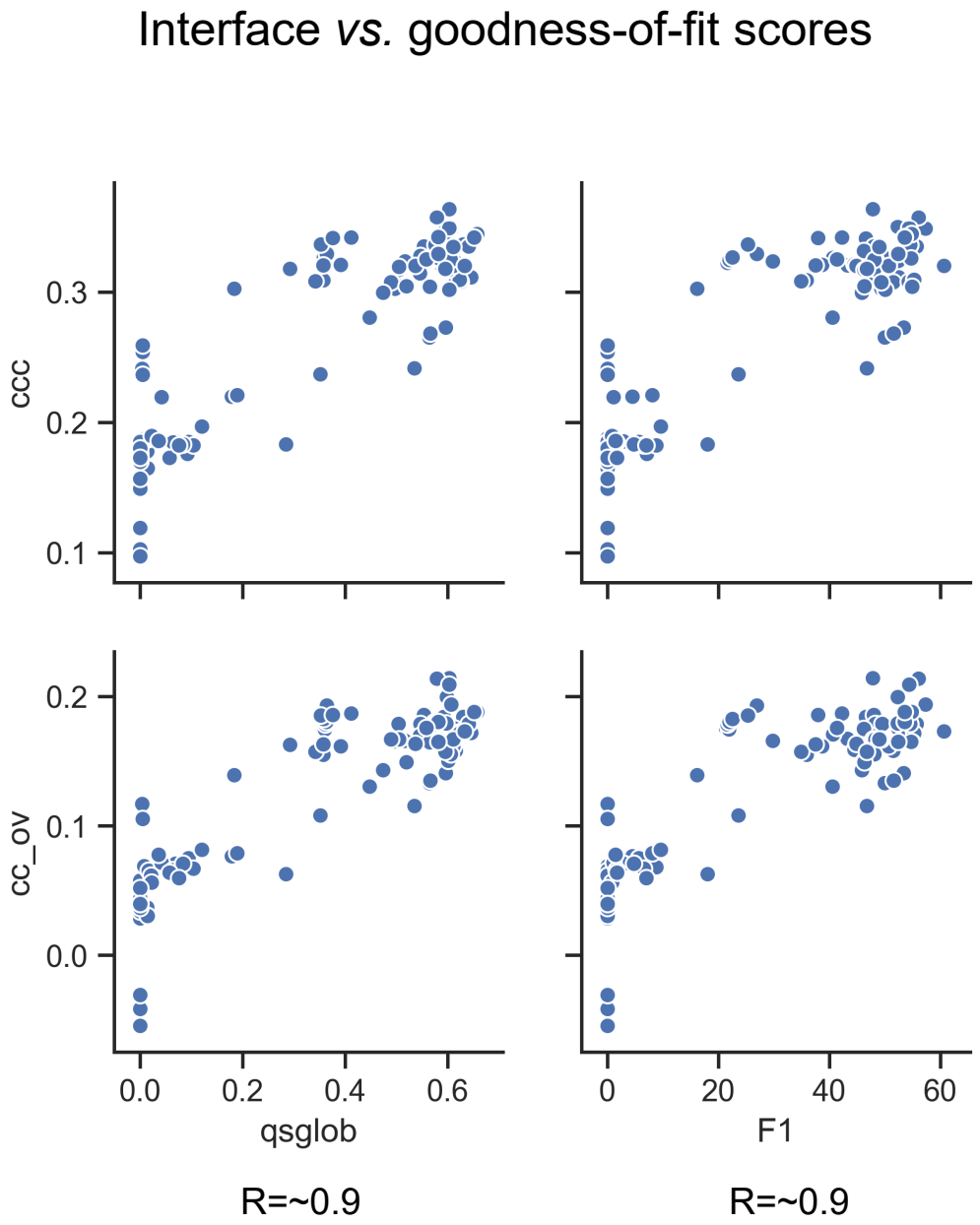
90

CASP:
Image redacted

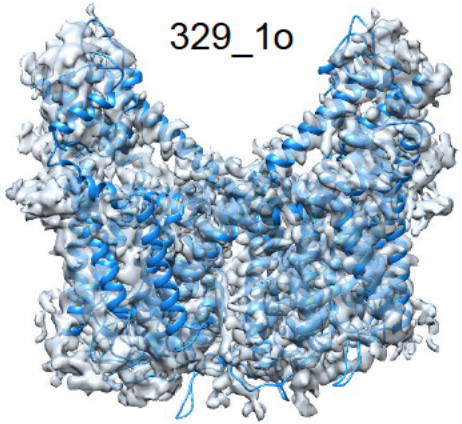
T1020o


Global scores *vs.* goodness-of-fit scores
ccc
0.3
0.2
0.1
cc_OV
0.2
0.1
0.0
0.2
0.4
0.6
lddtweight
0.25
0.50
0.75
1.00
tm
R=~0.5
R=~0.9

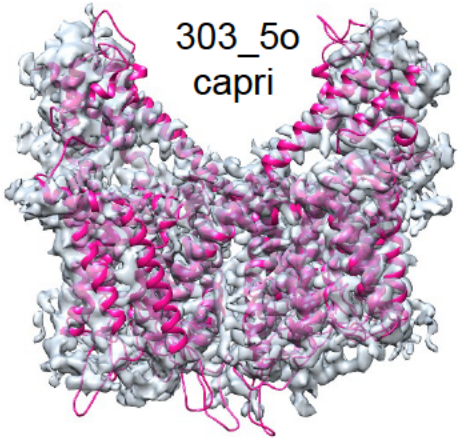
T1020o


Interface *vs.* goodness-of-fit scores
ccc
0.3
0.2
0.1
cc_ov
0.2
0.1
0.0
0.0
0.2
0.4
0.6
qsglob
0
20
40
60
F1
R=~0.9
R=~0.9

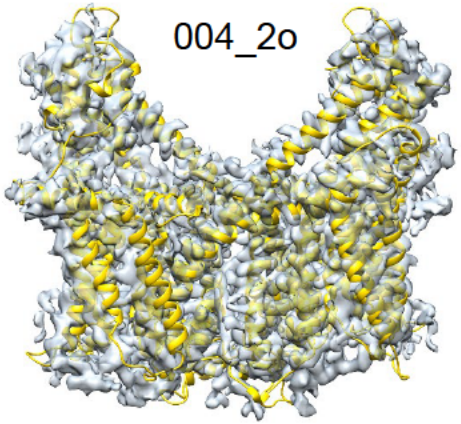
T0984 – 3.4 Å resolution



1st CCC
16th TM
0.338, 0.86

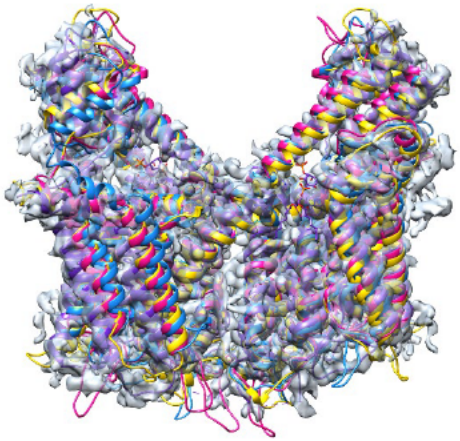


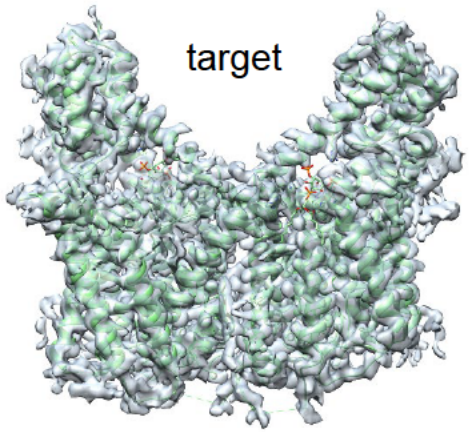
1st GDT/TM
9th CCC
0.89, 0.328



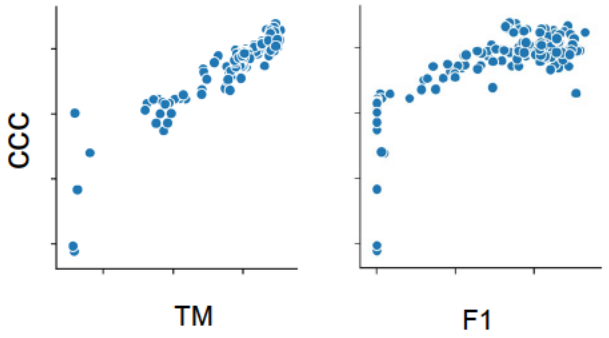
1st TM & CCC

0.87, 0.337

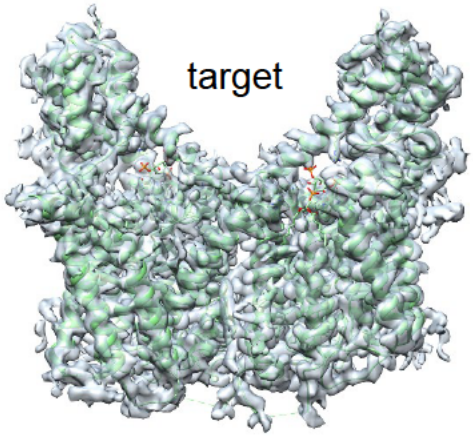


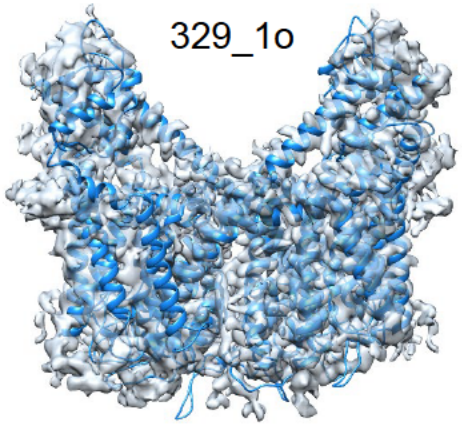


CCC=0.47

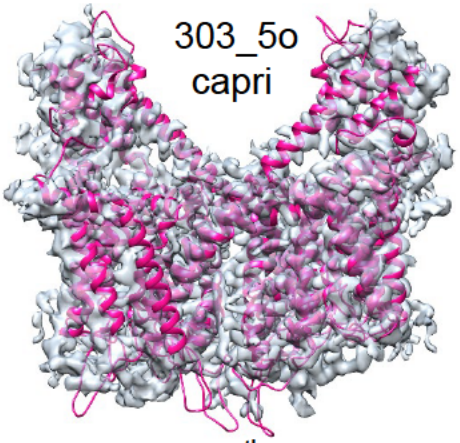


TS0984 – local fit to the density (per residue)

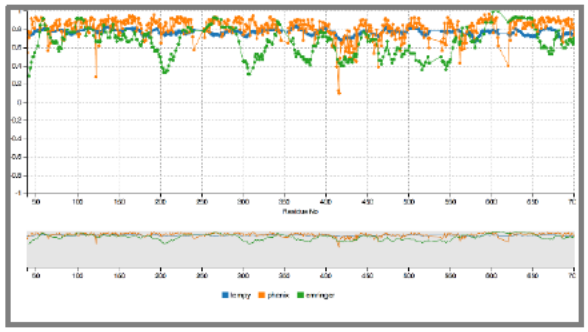


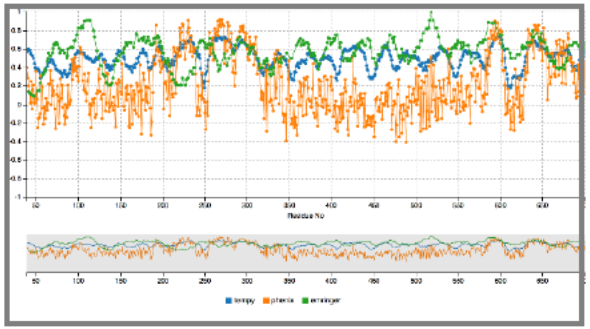


CCC (16th by TM)



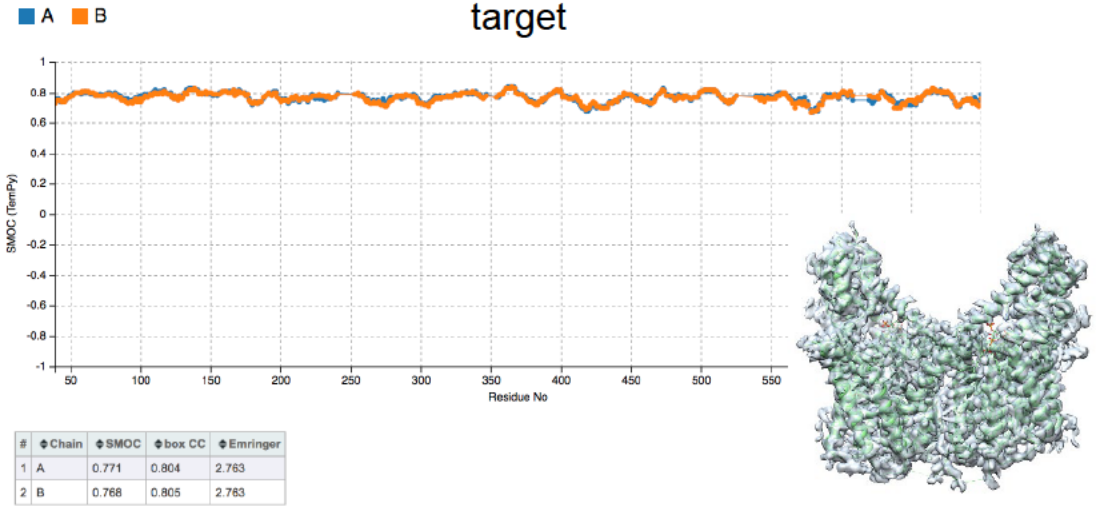
GDT/TM (9th by CCC)



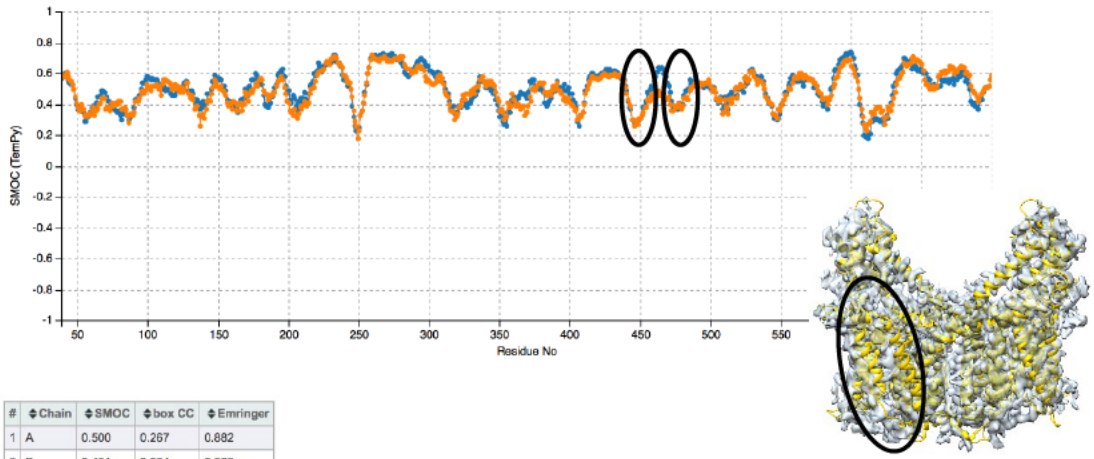


Local scores such as SMOC come to rescue!

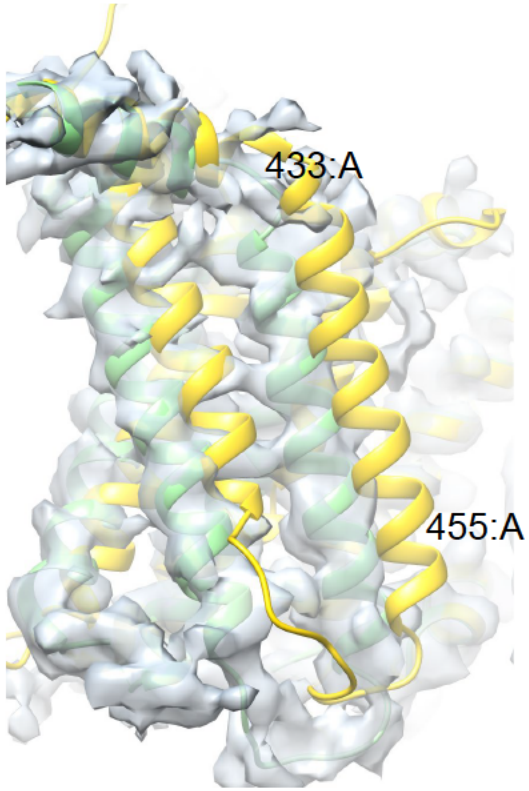
TS0984



#	Chain	SMOC	box CC	Emringer
1	A	0.771	0.804	2.763
2	B	0.768	0.805	2.763

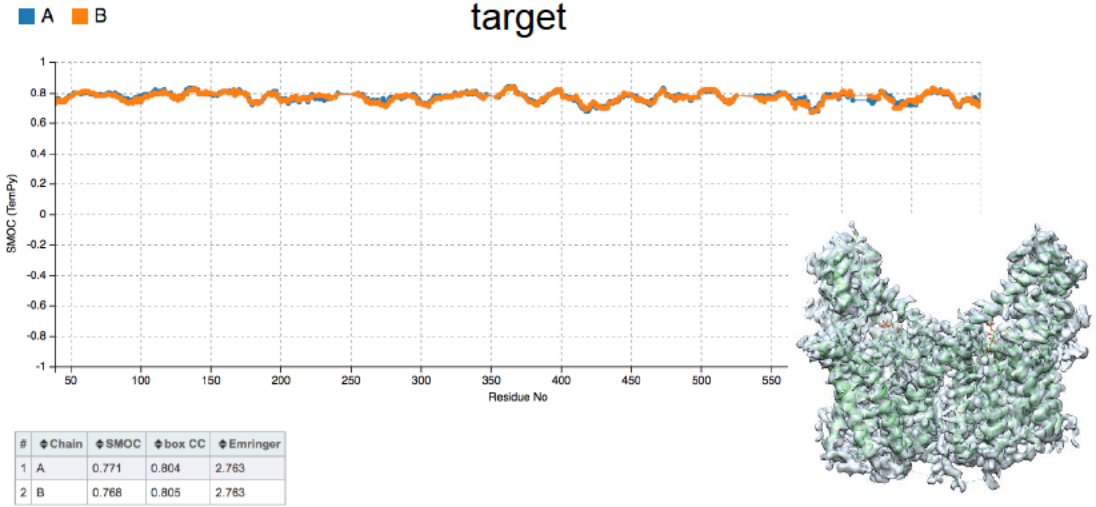


#	Chain	SMOC	box CC	Emringer
1	A	0.500	0.267	0.882
2	B	0.491	0.254	0.882

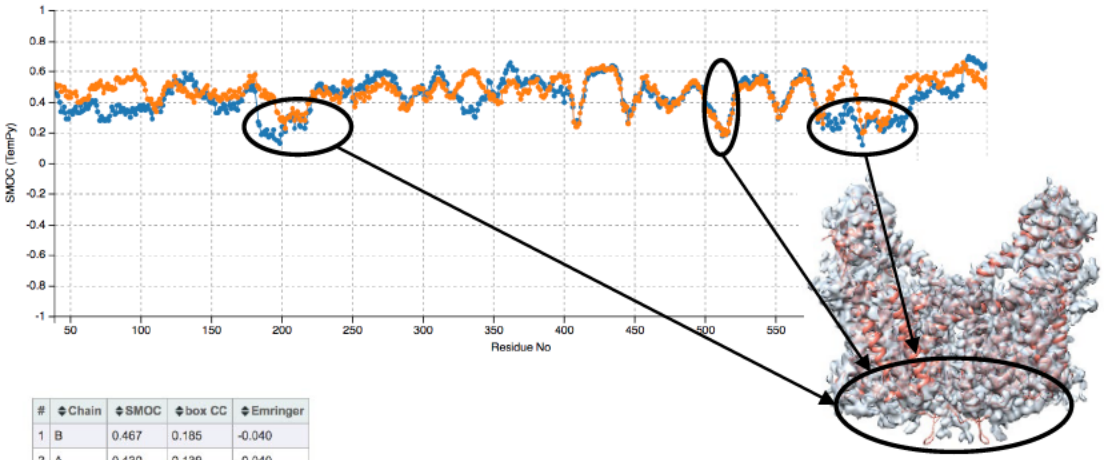


004_2o: tm=0.87, ccc= 0.337, F1= 35

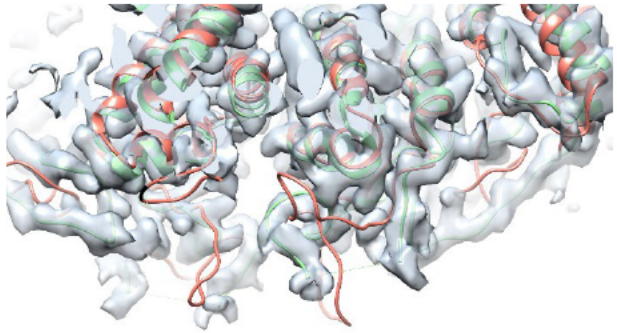
TS0984



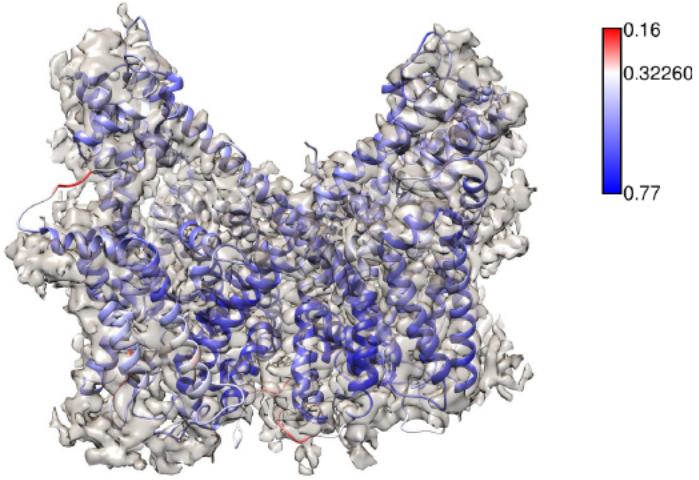
#	Chain	SMOC	box CC	Emringer
1	A	0.771	0.804	2.763
2	B	0.768	0.805	2.763



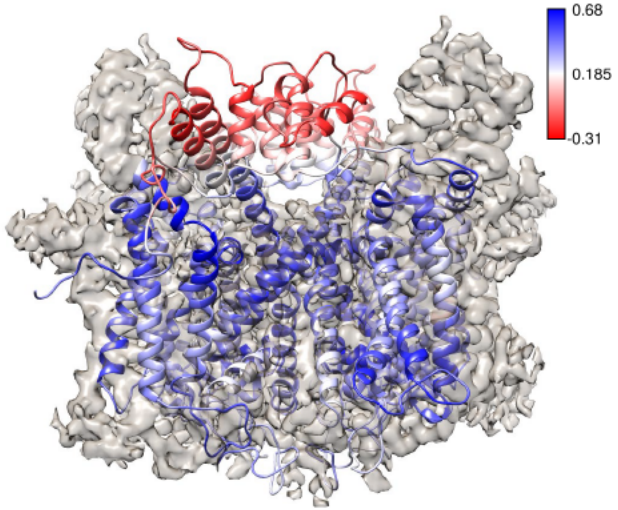
#	Chain	SMOC	box CC	Emringer
1	B	0.467	0.185	-0.040
2	A	0.430	0.139	-0.040



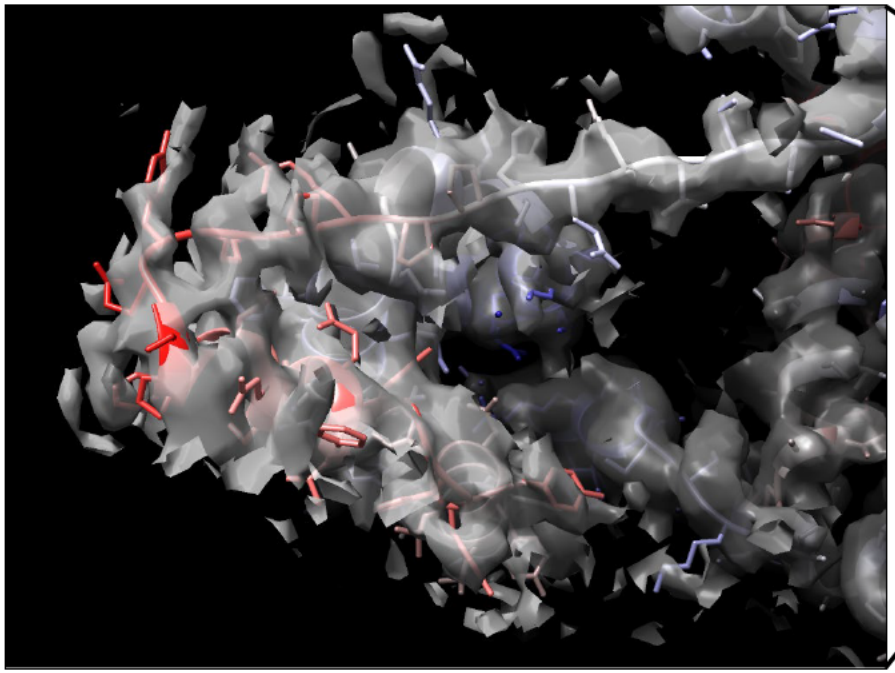
343_4: tm=0.88, ccc= 0.313, F1= 50



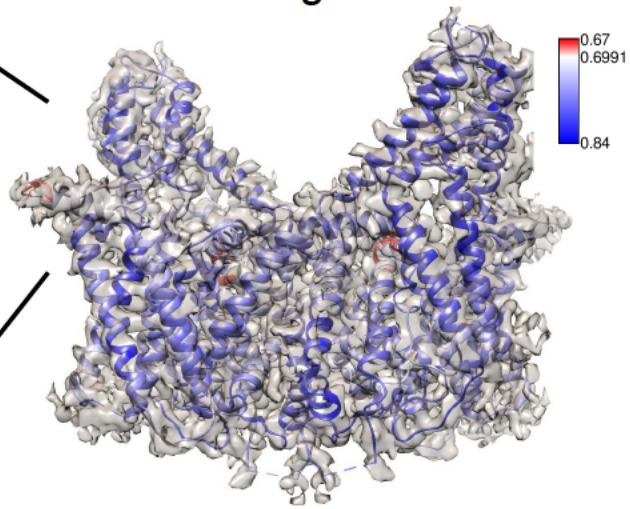
329_1o: best by CCC (TEMPy)



460_1o: best by lDDT



target



Conclusions

- At 3-4 Å useful to assess models against map – help identifying local errors in SSE, sidechain, and mis-orientation of sub-structures
- TM and GDT are highly correlated with global cross-correlation scores: a good indicator of overall quality of the model
- The global EMRinger may not be a reliable measure to evaluate CASP models, especially poor models
- Some regions in the target structure may have low accuracy so assessing the models against the map is more reliable in those regions
- Protein structure prediction methods that are assessed in CASP can be adapted for generating starting models prior to EM-based refinement