STRUCTURAL
BIOLOGY

Volume 72 (2016)

Supporting information for article:

PRISM-EM: template interface-based modelling of multi-protein complexes guided by cryo-electron microscopy density maps

Guray Kuzu, Ozlem Keskin, Ruth Nussinov and Attila Gursoy

Table S1 Scores obtained and templates used in homology modeling

Protein	PDB ID	C-score	Templates
Antibody m36	-	1.47	3ux9B, 3u2sA, 2kh2B, 1igtB, 1yc8A, 1dlfH
Acetylcholine-binding protein type 1	4aodA	1.46	3sq6A, 2qc1B, 4aq5B, 2qc1A, 4aoeA
Heavy chain of antibody VRC-PG04	3se9H	0.96	1dfbA, 3mlrH, 1y0lH, 3se8A, 4hsH, 2r56H
Light chain of antibody VRC-PG04	3se9L	0.94	3se8L
Lidless Mm-cpn	3iyfA	1.86	1q3qA, 4aolH, 1a6dA, 3j1bA
Conjugal transfer protein TrwB	1e9rA	1.53	4ag5A
Circadian clock protein KaiC	3dvlA	-1.79	2dr3D, 2ztsC, 1tf7A

Table S2 Evaluation of interfaces predicted by *Multifit*

PDB ID	PDB, interface	Model	Model, interface	IS-score	Result
3cre	AB	3cre-model	AB	0.07	incorrect
	BC		BC	0.07	incorrect
4a6j	AC	4a6j-3mer-model	AB	0.10	incorrect
			BC	0.23	correct
3se9	GH	3se9-model	GH	-	-
	HL		HL	0.12	acceptable
1e9r	AB	1e9r-model	AB	0.08	incorrect
			BC	0.08	incorrect
			CD	0.08	incorrect
			DE	0.08	incorrect
			EF	0.08	incorrect
			FA	0.08	incorrect
3dvl	AB	3dvl-model	AB	0.36	correct
			BC	0.36	correct
			CD	0.36	correct
			DE	0.36	correct
			EF	0.36	correct
			FA	0.36	correct

Table S3 Evaluation of interfaces predicted by *Segger*

PDB ID	PDB, interface	Model	Model, interface	IS-score	Result
4aod	AB	4aod-model	AB	0.08	incorrect
			BC	0.08	incorrect
			CD	0.08	incorrect
			DE	0.08	incorrect
			EA	0.08	incorrect
3iyf	AB	3iyf-model	AB	-	-
			BC	-	-
			CD	0.33	correct
			DE	0.34	correct
			EF	-	-
			FG	-	-
			GH	-	-
			HA	-	-

Table S4 Evaluation of interfaces in the *Multifit* benchmark set

For the cases 1z5sBD predicted by our method and 1z5sAD predicted by *Multifit*, the minimum interface residue number was decreased to 10 (default value: 15) in IS-score calculations.

PDB ID	Interface	Our predictions		*Multifit* predictions	
		IS-score	Result	IS-score	Result
1qu9	AB	0.84	correct	0.73	correct
	AC	0.84	correct	0.73	correct
	BC	0.79	correct	0.73	correct
1urz	AB	0.76	correct	0.85	correct
	AC	0.69	correct	0.80	correct
	BC	0.70	correct	0.83	correct
1z5s	AB	0.27	correct	-	
	AC	0.60	correct	0.11	incorrect
	AD	0.88	correct	0.04	incorrect
	BD	0.34	correct	-	
1e6v	AB	0.84	correct	0.15	acceptable
	AC	0.84	correct	0.23	correct
	AD	0.89	correct	0.34	correct
	AE	0.85	correct	0.09	incorrect
	BC	0.90	correct	0.34	correct
	BD	0.86	correct	0.55	correct
	BE	0.82	correct	-	
	DE	0.87	correct	0.09	incorrect
	DF	0.85	correct	-	
	EF	0.91	correct	-	
1mty	BC	0.84	correct	0.74	correct
	BD	0.92	correct	0.66	correct
	BG	0.82	correct	0.49	correct
	CE	0.89	correct	0.49	correct
	CH	0.82	correct	0.46	correct
	DE	0.85	correct	0.28	correct
	DG	0.86	correct	0.39	correct
	EH	0.89	correct	0.43	correct
1tyq	AB	0.48	correct	-	
	AD	0.84	correct	0.21	correct
	AE	0.76	correct	-	
	BF	0.86	correct	-	
	BG	0.55	correct	-	
	CF	0.77	correct	-	
	CG	0.64	correct	-	
	DF	0.87	correct	0.10	incorrect
	FG	0.72	correct	-	
1oel	AB	0.78	correct	0.51	correct

	BC	0.79	correct	0.56	correct
	CD	0.66	correct	0.52	correct
	DE	0.88	correct	0.51	correct
	EF	0.80	correct	0.54	correct
	FG	0.75	correct	0.52	correct
	GA	0.80	correct	0.51	correct
1gru	AB	0.80	correct	0.35	correct
	BC	0.81	correct	0.34	correct
	CD	0.78	correct	0.33	correct
	DE	0.79	correct	0.33	correct
	EF	0.71	correct	0.33	correct
	FG	0.75	correct	0.33	correct
	GA	0.49	correct	0.35	correct

Table S5 Predictions of HADDOCK-EM set

Complex	**Template**	**Query Proteins**		**Fiberdock energy**
1avx	1avwAB	2a31A	1avwB	-67.78
2oul	3e1zAB	2nnrA	3bpfA	-45.59
1ay7	1b27AD	1lniA	3da7C	-40.49
7cei	2jb0AB	1ayiA	1m08A	-22.33
2oob	2oobAB	2ooaB	4xofA	-15.79
2fd6	3bt2HU	2fatH	3bt2U	-17.80
1ak4	1m9eBC	2cplA	4j93A	-30.50
1b6c	3h9rAB	1iasA	1d6oA	-34.03
1bgx	1bgxHT	1ay1H	1tauA	-20.20
1r6q	1r6oAC	1mg9B	1mbvB	-40.8
1m10	1u0nAD	1fnsA	4c2aB	-18.45
1acb	1h9hEI	1p2nA	1eglA	-17.29
1jk9	1pu0CD	1qupB	1b4tA	-15.24
1bkd	1nvuRS	4nymR	4us0S	-103.86
1jmo	1tbqHR	1dm4B	1jmjA	-13.20

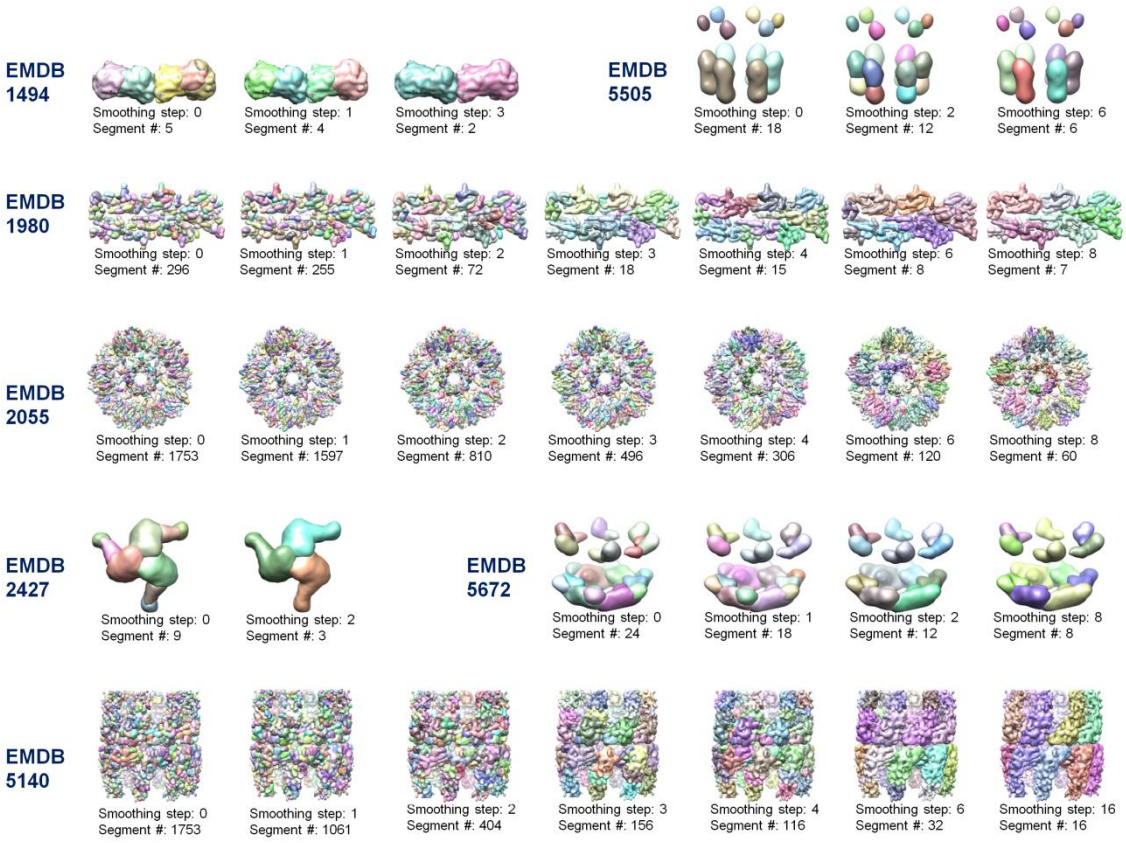

Figure S1 Results of segmentation by *Segger*. *Segger* gives different segmentation for different smoothing step number. Smoothing step and segment numbers are given below each figure. Figures obtained with different smoothing step number but giving the same result are not shown here; for example, smoothing step 2 gave the same result as smoothing step 1. We could obtain proper segmentation for EMDB IDs 2055 and 5140. We could obtain desired number of segments for EMDB IDs 1980, 5505 and 5672, but segmentation was not properly. Using different contour level of density maps did not change the segmentation.

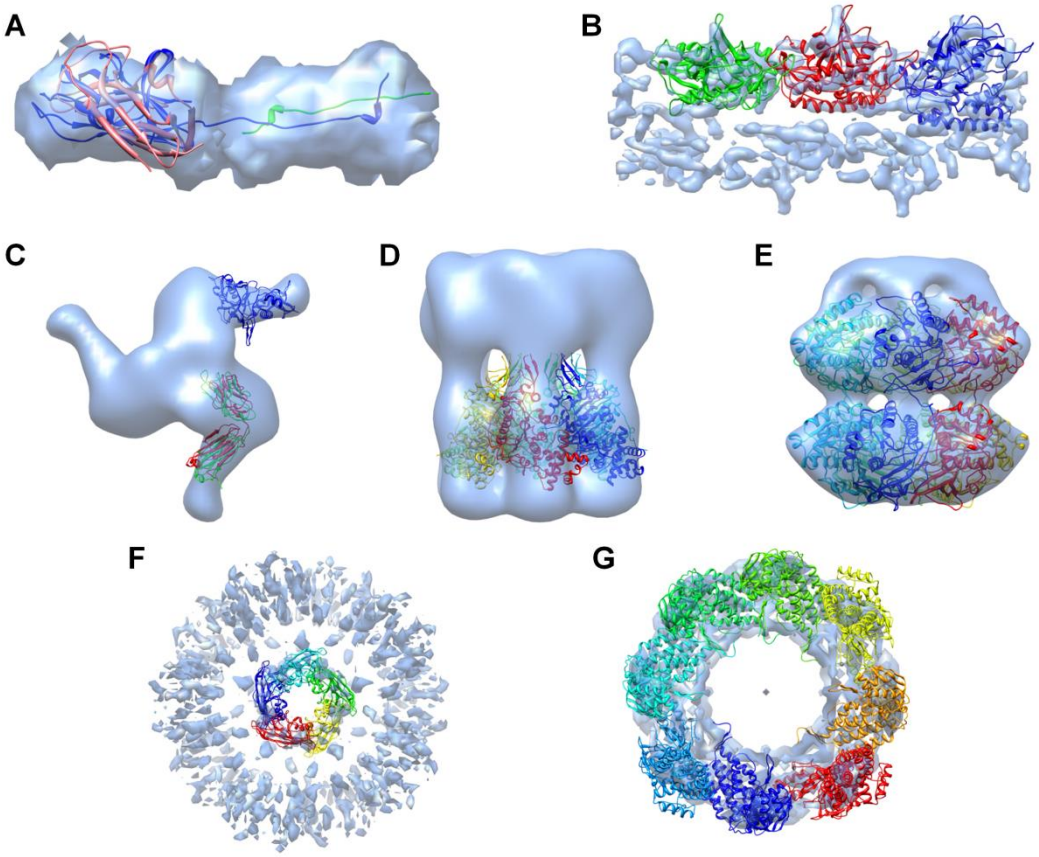

Figure S2 Models obtained by using *Multifit* and *Segger*. Models are fitted into EM density maps. Chains are given in different colors. (A) *Multifit* model of Saf Pilus fitted into EMDB ID 1494. (B) *Multifit* model of ParM filament fitted into EMDB ID 1980. (C) *Multifit* model of HIV-1 gp120/VRC-PG04 complex fitted into EMDB ID 2427. (D) *Multifit* model of conjugal transfer protein TrwB fitted into EMDB ID 5505. (E) *Multifit* model of circadian clock protein KaiC fitted into EMDB ID 5672. (F) *Segger* model of acetylcholine-binding protein type 1 fitted into EMDB ID 2055. (G) *Segger* model of lidless Mm-cpn fitted into EMDB ID 5140.