BIOZENTRUM
University of Basel
The Center for Molecular Life Sciences

SWISS-MODEL Homology Modelling Report

Model Building Report

This document lists the results for the homology modelling project "A6NFH5" submitted to SWISS-MODEL workspace on Feb. 28, 2018, 12:15 p.m..The submitted primary amino acid sequence is given in Table T1.

If you use any results in your research, please cite the relevant publications:

- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, W252-W258 (2014). doi
- Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009). doi
- Bienert, S., Waterhouse, A., de Beer, T.A., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017). doi
- Benkert, P., Biasini, M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27, 343-350 (2011). doi
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports 7 (2017). doi

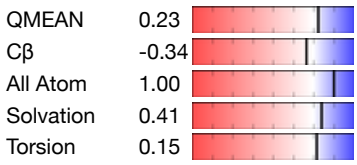
Results

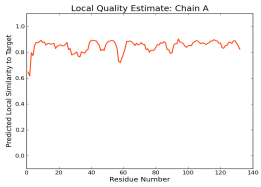
The SWISS-MODEL template library (SMTL version 2018-02-21, PDB release 2018-02-16) was searched with BLAST (Camacho et al.) and HHBlits (Remmert et al.) for evolutionary related structures matching the target sequence in Table T1. For details on the template search, see Materials and Methods. Overall 523 templates were found (Table T2).

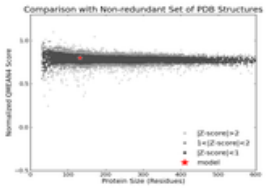
Models

The following model was built (see Materials and Methods "Model Building"):

Model #01	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 Version 1.1.0.	monomer	None	0.80	0.23







Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
4a1h.1.A	65.91	monomer	HHblits	X-ray	2.20Å	0.49	1 - 132	0.94	MYELIN P2 PROTEIN

Ligand	Added to Model	Description
CL	× - Not biologically relevant.	CHLORIDE ION
GOL	× - Not biologically relevant.	GLYCEROL
PLM	× - Binding site not conserved.	PALMITIC ACID

```
Target    MIDQLQGTWKSISCENSEDYMKELGIGRASRKLGRLAKPTVTISTDGDVITIKTKSIFKNNEISFKLGEEFEEITPGGHK
4a1h.1.A  MSNKFLGTWKLVSSENFDDYMKALGVGLATRKLGNLAKPTVIISKSGDIITIRTESTFKNTEISFKLGQEFEETTADNRK

Target    TKSKVTLDKESLIQVQDWDGKETTITRKLVDGKMVVESTVNSVICTRTYEKVSSNSVSNS
4a1h.1.A  TKSIVTLQRGSLNQVQRWDGKETTIKRKLVNGKMVAECKMKGVVCTRIYEKV--------
```

Materials and Methods

Template Search

Template search with BLAST and HHBlits has been performed against the SWISS-MODEL template library (SMTL, last update: 2018-02-21, last included PDB release: 2018-02-16).

The target sequence was searched with BLAST against the primary amino acid sequence contained in the SMTL. A total of 242 templates were found.

An initial HHblits profile has been built using the procedure outlined in (Remmert et al.), followed by 1 iteration of HHblits against NR20. The obtained profile has then be searched against all profiles of the SMTL. A total of 299 templates were found.

Template Selection

For each identified template, the template's quality has been predicted from features of the target-template alignment. The templates with the highest quality have then been selected for model building.

Model Building

Models are built based on the target-template alignment using ProMod3. Coordinates which are conserved between the target and the template are copied from the template to the model. Insertions and deletions are remodelled using a fragment library. Side chains are then rebuilt. Finally, the geometry of the resulting model is regularized by using a force field. In case loop modelling with ProMod3 fails, an alternative model is built with PROMOD-II (Guex et al.).

Model Quality Estimation

The global and per-residue model quality has been assessed using the QMEAN scoring function (Benkert et al.) . For improved performance, weights of the individual QMEAN terms have been trained specifically for SWISS-MODEL.

Ligand Modelling

Ligands present in the template structure are transferred by homology to the model when the following criteria are met: (a) The ligands are annotated as biologically relevant in the template library, (b) the ligand is in contact with the model, (c) the ligand is not clashing with the protein, (d) the residues in contact with the ligand are conserved between the target and the template. If any of these four criteria is not satisfied, a certain ligand will not be included in the model. The model summary includes information on why and which ligand has not been included.

Oligomeric State Conservation

The quaternary structure annotation of the template is used to model the target sequence in its oligomeric form. The method (Bertoni et al.) is based on a supervised machine learning algorithm, Support Vector Machines (SVM), which combines interface conservation, structural clustering, and other template features to provide a quaternary structure quality estimate (QSQE). The QSQE score is a number between 0 and 1, reflecting the expected accuracy of the interchain contacts for a model built based a given alignment and template. Higher numbers indicate higher reliability. This complements the GMQE score which estimates the accuracy of the tertiary structure of the resulting model.

References

- **BLAST**
  Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. BLAST+: architecture and applications. BMC Bioinformatics 10, 421-430 (2009). doi>
- **HHblits**
  Remmert, M., Biegert, A., Hauser, A., Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9, 173-175 (2012). doi>

Table T1:

Primary amino acid sequence for which templates were searched and models were built.

```
MIDQLQGTWKSISCENSEDYMKELGIGRASRKLGRLAKPTVTISTDGDVITIKTKSIFKNNEISFKLGEEFEEITPGGHKTKSKVTLDKESLIQVQDWDG
KETTITRKLVDGKMVVESTVNSVICTRTYEKVSSNSVSNS
```

Table T2:

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4a1h.1.A	65.91	monomer		HHblits	X-ray	2.20Å	0.49	0.94	MYELIN P2 PROTEIN
2wut.1.A	65.91	monomer		HHblits	X-ray	1.85Å	0.48	0.94	MYELIN P2 PROTEIN
4a1y.1.A	65.15	monomer		HHblits	X-ray	1.20Å	0.48	0.94	MYELIN P2 PROTEIN
4a8z.1.A	65.15	homo-dimer	0.46	HHblits	X-ray	1.80Å	0.48	0.94	MYELIN P2 PROTEIN
5n4p.1.A	65.15	monomer		HHblits	X-ray	1.53Å	0.48	0.94	Myelin P2 protein

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5n4q.1.A	65.15	monomer		HHblits	X-ray	1.72Å	0.48	0.94	Myelin P2 protein
5n4m.1.A	65.15	monomer		HHblits	X-ray	1.59Å	0.48	0.94	Myelin P2 protein
4d6b.1.A	65.15	homo-dimer	0.42	HHblits	X-ray	2.12Å	0.48	0.94	MYELIN P2 PROTEIN
5bvq.1.A	61.36	monomer		HHblits	X-ray	2.10Å	0.48	0.94	fatty acid-binding protein
5bvq.1.A	61.83	monomer		BLAST	X-ray	2.10Å	0.48	0.94	fatty acid-binding protein
1pmp.1.A	64.12	monomer		HHblits	X-ray	2.70Å	0.48	0.94	P2 MYELIN PROTEIN
1pmp.1.A	64.62	monomer		BLAST	X-ray	2.70Å	0.48	0.93	P2 MYELIN PROTEIN
1yiv.1.A	65.38	monomer		HHblits	X-ray	2.10Å	0.48	0.93	Myelin P2 protein
4a60.1.A	57.58	monomer		HHblits	X-ray	1.53Å	0.47	0.94	FATTY ACID-BINDING PROTEIN 9 TESTIS LIPID-BINDING PROTEIN, TLBP, TESTIS-TYPE FATTY ACID-BINDING PROTEIN, T-FABP
5hz9.1.A	51.13	monomer		HHblits	X-ray	2.30Å	0.45	0.95	Fatty acid-binding protein, heart
3wvm.1.A	51.13	monomer		HHblits	X-ray	0.88Å	0.45	0.95	Fatty acid-binding protein, heart
5b28.1.A	51.13	monomer		HHblits	X-ray	0.90Å	0.45	0.95	Fatty acid-binding protein, heart
3wbg.1.A	51.13	monomer		HHblits	X-ray	2.15Å	0.45	0.95	Fatty acid-binding protein, heart
3rsw.1.A	51.13	monomer		HHblits	X-ray	2.60Å	0.45	0.95	Fatty acid-binding protein, heart
5hz8.1.A	55.30	monomer		HHblits	X-ray	1.12Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5b27.1.A	50.38	monomer		HHblits	X-ray	1.02Å	0.45	0.95	Fatty acid-binding protein, heart
5bvt.1.A	57.25	monomer		BLAST	X-ray	2.31Å	0.46	0.94	Epidermal fatty acid-binding protein
5ce4.1.A	51.52	monomer		HHblits	neutron diff.	0.98Å	0.46	0.94	Fatty acid-binding protein, heart
2hnx.1.A	55.30	monomer		HHblits	X-ray	1.50Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5edb.1.A	55.30	monomer		HHblits	X-ray	1.18Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5edc.1.A	55.30	monomer		HHblits	X-ray	1.29Å	0.46	0.94	Fatty acid-binding protein, adipocyte
3p6c.1.A	55.30	monomer		HHblits	X-ray	1.25Å	0.46	0.94	Fatty acid-binding protein, adipocyte
3q6l.1.A	55.30	monomer		HHblits	X-ray	1.40Å	0.46	0.94	Fatty acid-binding protein, adipocyte
3fr4.1.A	55.30	monomer		HHblits	X-ray	2.16Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5hz8.1.A	55.73	monomer		BLAST	X-ray	1.12Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5c0n.1.A	54.55	monomer		HHblits	X-ray	3.00Å	0.45	0.94	Fatty acid-binding protein, adipocyte
5d8j.1.A	54.55	monomer		HHblits	X-ray	3.00Å	0.45	0.94	Fatty acid-binding protein, adipocyte
2qm9.1.A	54.55	homo-dimer	0.37	HHblits	X-ray	2.31Å	0.45	0.94	Fatty acid-binding protein, adipocyte
5hz9.1.A	51.91	monomer		BLAST	X-ray	2.30Å	0.46	0.94	Fatty acid-binding protein, heart
5bvt.1.A	55.30	monomer		HHblits	X-ray	2.31Å	0.45	0.94	Epidermal fatty acid-binding protein
5edc.1.A	55.73	monomer		BLAST	X-ray	1.29Å	0.46	0.94	Fatty acid-binding protein, adipocyte
3fr4.1.A	55.73	monomer		BLAST	X-ray	2.16Å	0.46	0.94	Fatty acid-binding protein, adipocyte
5c0n.1.A	54.96	monomer		BLAST	X-ray	3.00Å	0.45	0.94	Fatty acid-binding protein, adipocyte
5d8j.1.A	54.96	monomer		BLAST	X-ray	3.00Å	0.45	0.94	Fatty acid-binding protein, adipocyte
5ur9.2.A	50.38	monomer		HHblits	X-ray	2.20Å	0.44	0.95	Fatty acid-binding protein, epidermal
5ur9.6.A	50.38	monomer		HHblits	X-ray	2.20Å	0.44	0.95	Fatty acid-binding protein, epidermal
4azm.1.A	50.38	homo-dimer	0.46	HHblits	X-ray	2.75Å	0.44	0.95	FATTY ACID-BINDING PROTEIN, EPIDERMAL
5ur9.7.A	50.38	monomer		HHblits	X-ray	2.20Å	0.44	0.95	Fatty acid-binding protein, epidermal
5ur9.5.A	50.38	monomer		HHblits	X-ray	2.20Å	0.44	0.95	Fatty acid-binding protein, epidermal
1bwy.1.A	51.15	monomer		HHblits	NMR	NA	0.45	0.94	PROTEIN (HEART FATTY ACID BINDING PROTEIN)
1bwy.1.A	51.54	monomer		BLAST	NMR	NA	0.46	0.93	PROTEIN (HEART FATTY ACID BINDING PROTEIN)
1b56.1.A	50.76	monomer		HHblits	X-ray	2.05Å	0.44	0.94	FATTY ACID BINDING PROTEIN
1jjj.1.A	50.76	monomer		HHblits	NMR	NA	0.44	0.94	EPIDERMAL-TYPE FATTY ACID BINDING PROTEIN (E-FABP)

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5hz5.1.A	50.76	monomer		HHblits	X-ray	1.40Å	0.44	0.94	Fatty acid-binding protein, epidermal
4azo.1.A	49.62	homo-dimer	0.26	HHblits	X-ray	2.33Å	0.44	0.95	FATTY ACID-BINDING PROTEIN, EPIDERMAL

The table above shows the top 50 filtered templates. A further 267 templates were found which were considered to be less suitable for modelling than the filtered list.
2cbr.1.A, 5fen.1.A, 1hmr.1.A, 5u6g.5.A, 3rsw.1.A, 5l8i.2.A, 1fdq.1.A, 3wvm.1.A, 4gkc.1.A, 5b29.1.A, 4exz.1.A, 5b27.1.A, 4r80.2.A, 3fel.2.A, 1cbs.1.A, 3stn.1.A, 4zr2.1.B, 4zr2.1.A, 3fel.1.A, 1fe3.1.A, 5xn9.1.A, 5f6b.2.A, 4zcb.1.A, 1kqw.1.A, 1kzw.1.A, 3pp6.1.A, 4i3d.1.A, 1mdc.1.A, 1sa8.1.A, 3fa6.2.A, 1mx7.1.A, 1lpj.1.A, 1eio.1.A, 5ce4.1.A, 1zry.1.A, 1o8v.1.A, 1eal.1.A, 1lic.1.B, 4i3b.1.A, 4pww.1.A, 2frs.2.A, 4ykm.1.A, 5dpq.1.A, 2lfo.1.A, 4efg.1.A, 3fa9.2.A, 1p6p.1.A, 2qo6.1.A, 3jah.20.A, 5faz.2.A, 1ure.1.A, 4i3c.1.A, 1ggl.2.A, 4i3d.3.A, 5faz.1.A, 2f73.1.A, 4qgv.2.A, 4azm.1.A, 5fen.2.A, 2ft9.1.A, 4azp.1.A, 3cwk.1.A, 4i9r.1.A, 5b0u.2.A, 2ju8.1.A, 3hk1.1.A, 5ura.1.A, 2g7b.1.A, 4qgx.2.A, 2py1.1.A, 1jjj.1.A, 4yfp.1.A, 5ljd.1.A, 5hbs.1.A, 3fa7.1.A, 3fek.2.A, 5u6g.4.B, 4ruu.2.A, 4zgu.1.A, 1ifc.1.A, 2ju7.1.A, 1opa.1.A, 5l8i.3.A, 1t8v.1.A, 1alb.1.A, 5ur9.2.A, 3d96.2.A, 4zh6.1.A, 3q6l.1.A, 4r80.1.A, 1kgl.1.A, 5l8i.1.A, 4qyn.2.A, 5edb.1.A, 3f9d.2.A, 5b0u.1.A, 3fa6.1.A, 2jn3.1.A, 1lfo.1.A, 5lje.1.A, 5ffh.1.A, 1o1v.1.A, 2mm3.1.A, 3d95.2.A, 5gkb.1.A, 5dg4.2.A, 5h9a.1.A, 4qzu.4.A, 4zj0.2.A, 5ur9.5.A, 3fa8.2.A, 3ifb.1.A, 3p6c.1.A, 4zgu.4.A, 4i3c.2.A, 3em0.1.A, 1vyg.1.A, 5u6g.1.B, 5ur9.7.A, 3fr5.1.A, 1ael.1.A, 5hz5.1.A, 4m7m.1.A, 2mo5.1.A, 2fs7.1.A, 4ede.2.A, 3cr6.1.A, 1tw4.2.A, 3i17.1.A, 1jbh.1.A, 2poa.1.A, 3i17.1.B, 1ftp.1.A, 1blr.1.A, 2rct.1.A, 2ftb.1.A, 2fs6.2.A, 5hz6.1.A, 1crb.1.A, 4ruu.1.A, 1kzx.1.A, 3d96.1.A, 1mx8.1.A, 3jaj.14.A, 5u6g.3.A, 5ji4.1.A, 3js1.1.A, 2mji.1.A, 5ura.2.A, 3f8a.1.A, 1cbq.1.A, 2lkk.1.A, 1tvq.1.A, 2frs.1.A, 2flj.1.A, 4ybp.1.A, 2qo5.1.A, 2n93.1.A, 2qo4.1.A, 5f7g.1.A, 1xca.1.B, 3vg2.1.A, 1xca.1.A, 3d95.1.A, 1eii.1.A, 5h8t.1.A, 1g74.1.A, 5u6g.5.B, 1opb.2.A, 5hzq.2.A, 4qzu.3.A, 4qgv.1.A, 6at8.1.A, 4qzt.2.A, 4zcb.1.B, 4zh9.1.B, 5ur9.6.A, 3akn.1.A, 2l68.1.A, 4azn.1.A, 5xna.1.B, 3d97.1.A, 2l67.1.A, 4lkp.1.A, 4azr.1.A, 4azr.1.B, 4qgw.1.A, 3elz.3.A, 5ibo.1.A, 5f58.1.A, 5ljb.1.A, 4eej.1.A, 4qgx.1.A, 5ljc.1.A, 1kqx.1.A, 5kph.1.A, 3fa8.1.A, 5u6g.1.A, 1dc9.1.A, 2hnx.1.A, 3f9d.1.A, 3fen.1.A, 4yko.1.A, 4m6s.1.A, 5f6b.1.A, 5b28.1.A, 5u6g.6.A, 4i3d.4.A, 2f73.2.A, 3akm.1.A, 1mvg.1.A, 4zj0.1.A, 4i3b.2.A, 3em0.2.A, 3elx.1.A, 2ans.1.A, 3j79.1.X, 2g79.1.A, 1cbr.1.A, 1x4r.1.A, 2a0a.1.A, 4i9s.1.A, 2lba.1.A, 1b56.1.A, 3cbs.1.A, 1acd.1.A, 5dg4.1.A, 1b4m.1.A, 2k62.1.A, 3elz.1.A, 5hzq.1.A, 4lkt.1.A, 5u6g.6.B, 4zgu.2.A, 4ede.1.A, 3wbg.1.A, 4azo.1.A, 2ju3.1.A, 3fek.1.A, 4exz.2.A, 5u6g.4.A, 4lkt.3.A, 5f58.2.A, 3fa7.2.A, 3fa9.1.A, 5u6g.3.B, 1bm5.1.A, 3fen.2.A, 1icn.1.A, 1g5w.1.A, 1opb.1.A, 5u6g.2.A, 5dpq.1.B, 3d97.2.A, 1cbi.1.B, 1ggl.1.A, 1cbi.1.A, 4qgw.2.A, 1fdq.2.A, 3ppt.1.A, 1a57.1.A, 5ibo.2.A, 1o1u.1.A, 1jjx.1.A, 3nr3.1.A, 4i3d.2.A, 1ifb.1.A