SWISS-MODEL Homology Modelling Report

Model Building Report

This document lists the results for the homology modelling project "Q93VI0" submitted to SWISS-MODEL workspace on March 6, 2017, 5:07 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:

Marco Biasini; Stefan Bienert; Andrew Waterhouse; Konstantin Arnold; Gabriel Studer; Tobias Schmidt; Florian Kiefer; Tiziano Gallo Cassarino; Martino Bertoni; Lorenza Bordoli; Torsten Schwede. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research (1 July 2014) 42 (W1): W252-W258; doi: 10.1093/nar/gku340.
Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics, 22, 195-201.
Benkert, P., Biasini, M. and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics, 27, 343-350

Results

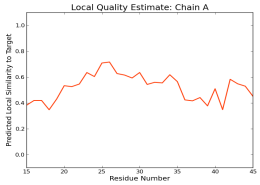
The SWISS-MODEL template library (SMTL version 2017-02-08, PDB release 2017-02-03) was searched with Blast (Altschul et al., 1997) and HHBlits (Remmert, et al., 2011) for evolutionary related structures matching the target sequence in Table T1. For details on the template search, see Materials and Methods. Overall 57 templates were found (Table T2).

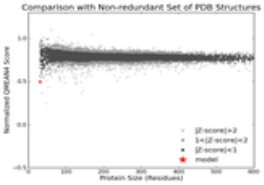
Models

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

Model #01	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 Version 1.0.2.	MONOMER (matching prediction)	None	0.16	-2.43

QMEAN	-2.43
Cβ	-2.18
All Atom	-0.46
Solvation	0.14
Torsion	-2.32





Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
2lcx.1.B	36.67	homo-dimer	HHblits	NMR	NA	0.36		0.41	Receptor tyrosine-protein kinase erbB-4

The template contained no ligands.

```
Target    MAAEFDGKIESKGLNPG-LIVLLVIGGLLLTFLVGNFI-LYTYAQKNLPPRKKKPVSKKKMKKEKMKQGVQVPGE
2lcx.1.B  --------------ARTPLIAAGVIGGLFILVIVGLTFAVYV-RRKS----------------------------
```

Materials and Methods

Template Search

Template search with Blast and HHBlits has been performed against the SWISS-MODEL template library (SMTL, last update: 2017-02-08, last included PDB release: 2017-02-03).

The target sequence was searched with BLAST (Altschul et al., 1997) against the primary amino acid sequence contained in the SMTL.

An initial HHblits profile has been built using the procedure outlined in (Remmert, et al., 2011), followed by 1 iteration of HHblits against NR20. The obtained profile has then be searched against all profiles of the SMTL. A total of 62 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., 1997).

Model Quality Estimation

The global and per-residue model quality has been assessed using the QMEAN scoring function (Benkert, et al., 2011) . 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 (Gallo -Casserino, to be published): (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

Homo-oligomeric structure of the target protein is predicted based on the analysis of pairwise interfaces of the identified template structures. For each relevant interface between polypeptide chains (interfaces with more than 10 residue-residue interactions), the QscoreOligomer (Mariani et al., 2011) is predicted from features such as similarity to target and frequency of observing this interface in the identified templates (Kiefer, Bertoni, Biasini, to be published). The prediction is performed with a random forest regressor using these features as input parameters to predict the probability of conservation for each interface. The QscoreOligomer of the whole complex is then calculated as the weight-averaged QscoreOligomer of the interfaces. The oligomeric state of the target is predicted to be the same as in the template when QscoreOligomer is predicted to be higher or equal to 0.5.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res, 25, 3389-3402.
Remmert, M., Biegert, A., Hauser, A. and Soding, J. (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods, 9, 173-175.
Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis, 18, 2714-2723.
Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol, 234, 779-815.
Benkert, P., Biasini, M. and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics, 27, 343-350.
Mariani, V., Kiefer, F., Schmidt, T., Haas, J. and Schwede, T. (2011) Assessment of template based protein structure predictions in CASP9. Proteins, 79 Suppl 10, 37-58.

Table T1:

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

```
MAAEFDGKIESKGLNPGLIVLLVIGGLLLTFLVGNFILYTYAQKNLPPRKKKPVSKKKMKKEKMKQGVQVPGE
```

Table T2:

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage	Description
2lcx.1.A	36.67	homo-dimer	HHblits	NMR	NA	0.36	0.41	Receptor tyrosine-protein kinase erbB-4
2lcx.1.B	36.67	homo-dimer	HHblits	NMR	NA	0.36	0.41	Receptor tyrosine-protein kinase erbB-4
2l2t.1.A	36.67	homo-dimer	HHblits	NMR	NA	0.36	0.41	Receptor tyrosine-protein kinase erbB-4
2l2t.1.B	36.67	homo-dimer	HHblits	NMR	NA	0.36	0.41	Receptor tyrosine-protein kinase erbB-4
4wo1.2.B	27.59	homo-dimer	HHblits	X-ray	2.14Å	0.36	0.40	TYRO protein tyrosine kinase-binding protein
2l34.1.A	27.59	homo-dimer	HHblits	NMR	NA	0.36	0.40	TYRO protein tyrosine kinase-binding protein
4wo1.2.A	27.59	homo-dimer	HHblits	X-ray	2.14Å	0.36	0.40	TYRO protein tyrosine kinase-binding protein

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage	Description
2l34.1.B	27.59	homo-dimer	HHblits	NMR	NA	0.36	0.40	TYRO protein tyrosine kinase-binding protein
4wol.1.C	27.59	homo-trimer	HHblits	X-ray	1.77Å	0.36	0.40	TYRO protein tyrosine kinase-binding protein
4wol.1.B	27.59	homo-trimer	HHblits	X-ray	1.77Å	0.36	0.40	TYRO protein tyrosine kinase-binding protein
2l35.1.B	27.59	hetero-oligomer	HHblits	NMR	NA	0.36	0.40	TYRO protein tyrosine kinase-binding protein
2n9y.1.A	22.58	hetero-oligomer	HHblits	NMR	NA	0.31	0.42	Integrin alpha-IIb
2mfr.1.A	27.59	monomer	HHblits	NMR	NA	0.34	0.40	Insulin receptor
3m78.1.A	9.38	homo-trimer	HHblits	X-ray	2.60Å	0.25	0.44	Tellurite resistance protein tehA homolog
2bg9.1.C	20.69	hetero-oligomer	HHblits	EM	4.00Å	0.32	0.40	ACETYLCHOLINE RECEPTOR PROTEIN, DELTA CHAIN
2knc.1.A	20.69	hetero-oligomer	HHblits	NMR	NA	0.31	0.40	Integrin alpha-IIb
3m74.1.A	6.45	homo-trimer	HHblits	X-ray	1.65Å	0.24	0.42	Tellurite resistance protein tehA homolog
3m73.1.A	6.45	homo-trimer	HHblits	X-ray	1.15Å	0.24	0.42	Tellurite resistance protein tehA homolog
3m76.1.A	6.45	homo-trimer	HHblits	X-ray	1.50Å	0.24	0.42	Tellurite resistance protein tehA homolog
3m7e.1.A	6.45	homo-trimer	HHblits	X-ray	1.80Å	0.24	0.42	Tellurite resistance protein tehA homolog
2k9j.1.A	21.43	hetero-oligomer	HHblits	NMR	NA	0.31	0.38	Integrin alpha-IIb light chain
2k1a.1.A	21.43	monomer	HHblits	NMR	NA	0.31	0.38	Integrin alpha-IIb
2ks1.1.B	29.63	hetero-oligomer	HHblits	NMR	NA	0.33	0.37	Epidermal growth factor receptor
2m0b.1.A	29.63	homo-dimer	HHblits	NMR	NA	0.33	0.37	Epidermal growth factor receptor
2m0b.1.B	29.63	homo-dimer	HHblits	NMR	NA	0.33	0.37	Epidermal growth factor receptor
3m75.1.A	6.67	homo-trimer	HHblits	X-ray	1.60Å	0.24	0.41	Tellurite resistance protein tehA homolog
2l35.1.A	32.00	hetero-oligomer	HHblits	NMR	NA	0.38	0.34	DAP12-NKG2C_TM
2m3e.1.A	34.62	monomer	HHblits	NMR	NA	0.32	0.36	Integrin alpha-L
2lx0.1.A	63.64	monomer	HHblits	NMR	NA	0.44	0.30	Membrane fusion protein p14
2n2a.1.A	25.93	homo-dimer	HHblits	NMR	NA	0.28	0.37	Receptor tyrosine-protein kinase erbB-2
2n2a.1.B	25.93	homo-dimer	HHblits	NMR	NA	0.28	0.37	Receptor tyrosine-protein kinase erbB-2
3eh4.1.C	29.17	hetero-oligomer	HHblits	X-ray	2.90Å	0.34	0.33	Cytochrome c oxidase polypeptide 2A
1ehk.1.C	29.17	hetero-oligomer	HHblits	X-ray	2.40Å	0.34	0.33	BA3-TYPE CYTOCHROME-C OXIDASE
1xme.1.C	29.17	hetero-oligomer	HHblits	X-ray	2.30Å	0.34	0.33	Cytochrome c oxidase polypeptide IIA
2n5s.1.A	29.17	monomer	HHblits	NMR	NA	0.32	0.33	Epidermal growth factor receptor
2l8s.1.A	29.17	monomer	HHblits	NMR	NA	0.32	0.33	Integrin alpha-1
2knc.1.B	36.36	hetero-oligomer	HHblits	NMR	NA	0.35	0.30	Integrin beta-3
2m20.1.A	31.82	homo-dimer	HHblits	NMR	NA	0.34	0.30	Epidermal growth factor receptor
2lzl.1.A	21.74	homo-dimer	HHblits	NMR	NA	0.28	0.32	Fibroblast growth factor receptor 3
2lzl.1.B	21.74	homo-dimer	HHblits	NMR	NA	0.28	0.32	Fibroblast growth factor receptor 3
2rmz.1.A	35.00	monomer	HHblits	NMR	NA	0.35	0.27	Integrin beta-3

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage	Description
2rn0.1.A	35.00	monomer	HHblits	NMR	NA	0.35	0.27	Integrin beta-3
2k9j.1.B	35.00	hetero-oligomer	HHblits	NMR	NA	0.35	0.27	Integrin beta-3
2l91.1.A	35.00	monomer	HHblits	NMR	NA	0.35	0.27	Integrin beta-3
5djq.1.D	33.33	hetero-oligomer	HHblits	X-ray	3.20Å	0.34	0.25	Putative uncharacterized protein
3jcu.1.J	35.29	hetero-oligomer	HHblits	EM	NA	0.35	0.23	Photosystem II reaction center protein J
5h1s.1.T	31.25	hetero-oligomer	HHblits	EM	NA	0.40	0.22	50S ribosomal protein L32, chloroplastic
4ub8.2.I	23.53	hetero-oligomer	HHblits	X-ray	1.95Å	0.35	0.23	Photosystem II reaction center protein J
4fby.1.I	23.53	hetero-oligomer	HHblits	X-ray	6.56Å	0.35	0.23	Photosystem II reaction center protein J
2axt.1.I	23.53	hetero-oligomer	HHblits	X-ray	3.00Å	0.35	0.23	Photosystem II reaction center J protein
3a0h.1.I	23.53	hetero-oligomer	HHblits	X-ray	4.00Å	0.35	0.23	Photosystem II reaction center protein J
3a0b.1.I	23.53	hetero-oligomer	HHblits	X-ray	3.70Å	0.35	0.23	Photosystem II reaction center protein J
1s5l.1.I	23.53	hetero-oligomer	HHblits	X-ray	3.50Å	0.35	0.23	Photosystem II reaction center J protein
5li0.1.j	11.76	hetero-oligomer	HHblits	EM	NA	0.33	0.23	50S ribosomal protein L32
4yuu.1.I	33.33	hetero-oligomer	HHblits	X-ray	2.77Å	0.39	0.21	Photosystem II reaction center protein J
4il6.1.I	26.67	hetero-oligomer	HHblits	X-ray	2.10Å	0.36	0.21	Photosystem II reaction center protein J
5e7c.1.I	26.67	hetero-oligomer	HHblits	X-ray	4.50Å	0.36	0.21	Photosystem II reaction center protein J