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 "CD63 WT" submitted to SWISS-MODEL workspace on Oct. 19, 2019, 2:38 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:

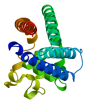
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1), W296-W303 (2018). doi>
- Guex, N., Peitsch, 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.P., 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>

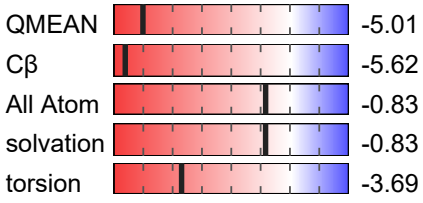
Results

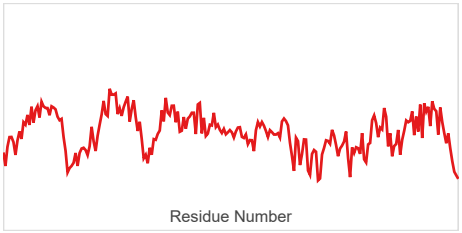
The SWISS-MODEL template library (SMTL version 2019-10-17, PDB release 2019-10-11) was searched with BLAST (Camacho et al.) and HHBlits (Remmert et al.) for evolutionary related structures matching the target sequence in Table T1. For details on the template search, see Materials and Methods. Overall 38 templates were found (Table T2).

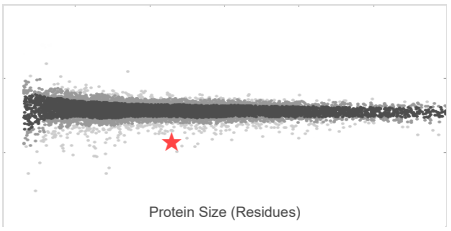
Models

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

Model #01	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 2.0.0	monomer	None	0.60	-5.01







Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
5tcx.1.A	24.43	monomer	0.00	HHblits	X-ray	2.96Å	0.33	6 - 233	0.93	CD81 antigen

Excluded ligands

Ligand Name.Number	Reason for Exclusion	Description
CLR.1	Binding site not conserved.	CHOLESTEROL

```
Target    MAVEGGMKCVKFLLYVLLLAFCACAVGLIAVGVGAQLVLSQTII----QG----ATPGSLLPVVIIAVGVFLFLVAFVGC
5tcx.1.A  -GVEGSTKSIKYLLFVFNFVFWLAGGVILGVALWLRHDPQTTNLLYLELGDKPAPNTFYVGIYILIAVGAVMMFVGFLGC

Target    CGACKENYCLMITFAIFLSLIMLVEVAAAIAGYVFRDKVMSEFNNNFRQQMENYP---KNNHTASILDRMQADFKCCGAA
```

```
5tcx.1.A  YGAIQESQCLLGTFFTCLVILFACEVAAGIWGFVNKDQIAKDVKQFYDQALQQAVVDDDANNAKAVVKTFHETLDCCGSS

Target    NYTDWEKIPSMSKNRVPDSCCINVTVGCGINFNEKAIHKEGCVEKIGGWLRKNVLVVAAAALGIAFVEVLGIVFACCLVK
5tcx.1.A  TLTALTTS------VLKNNLCPSG---SN---IISNLFKEDCHQKIDDLFSGKLYLIGIAAIVVAVIMIFEMILSMVLSS

Target    SIRSGYEVM
5tcx.1.A  GIRNS----
```

Materials and Methods

Template Search

Template search with BLAST and HHBlits has been performed against the SWISS-MODEL template library (SMTL, last update: 2019-10-17, last included PDB release: 2019-10-11).

The target sequence was searched with BLAST 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.), followed by 1 iteration of HHblits against NR20. The obtained profile has then be searched against all profiles of the SMTL. A total of 42 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.

```
MAVEGGMKCVKFLLYVLLLAFCACAVGLIAVGVGAQLVLSQTIIQGATPGSLLPVVIIAVGVFLFLVAFVGCCGACKENYCLMITFAIFLSLIMLVEVAA
AIAGYVFRDKVMSEFNNNFRQQMENYPKNNHTASILDRMQADFKCCGAANYTDWEKIPSMSKNRVPDSCCINVTVGCGINFNEKAIHKEGCVEKIGGWLR
KNVLVVAAAALGIAFVEVLGIVFACCLVKSIRSGYEVM
```

Table T2:

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5tcx.1.A	24.43	monomer	-	HHblits	X-ray	2.96Å	0.33	0.93	CD81 antigen
5m3d.1.A	17.24	homo-dimer	0.06	HHblits	X-ray	2.38Å	0.32	0.37	CD81 antigen
5m2c.1.B	17.24	homo-dimer	0.10	HHblits	X-ray	1.96Å	0.32	0.37	CD81 antigen
5dfv.1.A	17.24	homo-dimer	-	HHblits	X-ray	2.80Å	0.32	0.37	CD81 antigen
5dfv.1.B	17.24	homo-dimer	-	HHblits	X-ray	2.80Å	0.32	0.37	CD81 antigen
6ejg.1.A	17.05	homo-dimer	0.04	HHblits	X-ray	2.82Å	0.32	0.37	CD81 antigen
3x0g.1.A	13.33	homo-dimer	-	HHblits	X-ray	1.90Å	0.30	0.38	CD81
6ejm.1.A	17.05	homo-dimer	0.10	HHblits	X-ray	2.15Å	0.32	0.37	CD81 antigen
3x0f.2.A	16.09	monomer	-	HHblits	X-ray	1.47Å	0.31	0.37	CD81 antigen
3x0f.1.A	16.09	monomer	-	HHblits	X-ray	1.47Å	0.31	0.37	CD81 antigen
6ejm.1.B	17.05	homo-dimer	0.10	HHblits	X-ray	2.15Å	0.32	0.37	CD81 antigen
2m7z.1.A	23.08	monomer	-	HHblits	NMR	NA	0.35	0.33	CD63-like protein Sm-TSP-2
6cfw.1.I	23.08	monomer	-	HHblits	EM	3.70Å	0.29	0.11	Monovalent cation/H+ antiporter subunit G
4dbg.1.B	8.00	monomer	-	HHblits	X-ray	2.71Å	0.26	0.11	RING finger protein 31
5x0w.1.A	8.00	monomer	-	HHblits	X-ray	3.00Å	0.26	0.11	E3 ubiquitin-protein ligase RNF31
5x0w.2.A	8.00	monomer	-	HHblits	X-ray	3.00Å	0.26	0.11	E3 ubiquitin-protein ligase RNF31
5x0w.3.A	8.00	monomer	-	HHblits	X-ray	3.00Å	0.26	0.11	E3 ubiquitin-protein ligase RNF31
5x0w.4.A	8.00	monomer	-	HHblits	X-ray	3.00Å	0.26	0.11	E3 ubiquitin-protein ligase RNF31
5y3t.1.B	8.00	monomer	-	HHblits	X-ray	2.40Å	0.25	0.11	E3 ubiquitin-protein ligase RNF31

The table above shows the top 19 filtered templates. A further 19 templates were found which were considered to be less suitable for modelling than the filtered list.
6ek2.1.B, 5m3d.1.B, 5m33.1.A, 5dfw.1.A, 5m33.1.B, 1g8q.1.A, 5m4r.3.A, 6ek2.1.A, 1iv5.1.A, 1iv5.1.B, 5m4r.2.A, 5m3d.2.A, 3x0e.1.B, 5m3t.1.B, 5m3t.1.A, 5m4r.1.A, 5m4r.1.B, 3x0e.1.A, 1g8q.1.B