SWISS-MODEL Homology Modelling Report

Model Building Report

This document lists the results for the homology modelling project "HRTV GP-Heterodimer" submitted to SWISS-MODEL workspace on April 16, 2025, 7:46 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 FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T
  SWISS-MODEL: homology modelling of protein structures and complexes.
  Nucleic Acids Res 46, W296-W303. (2018) PMID 29788355 DOI 10.1093/nar/gky427
- Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T
  The SWISS-MODEL Repository - new features and functionality.
  Nucleic Acids Res 45, D313-D319. (2017) PMID 27899672 DOI 10.1093/nar/gkw1132
- Studer G, Tauriello G, Bienert S, Biasini M, Johner N, Schwede T
  ProMod3 - A versatile homology modelling toolbox.
  PLOS Comp Biol 17(1), e1008667. (2021) PMID 33507980 DOI 10.1371/journal.pcbi.1008667
- Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T
  QMEANDisCo - distance constraints applied on model quality estimation.
  Bioinformatics 36, 1765-1771. (2020) PMID 31697312 DOI 10.1093/bioinformatics/btz828
- 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) PMID 28874689 DOI 10.1038/s41598-017-09654-8

Results

The SWISS-MODEL template library (SMTL version 2025-04-16, PDB release 2025-04-11) was searched with BLAST (Camacho et al.) and HHblits (Steinegger et al.) for evolutionary related structures matching the target sequences in Table T1. For details on the template search, see Materials and Methods. Overall 94 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	QMEANDisCo Global
	PDB	ProMod3 3.4.1	hetero-1-1-mer	None	0.73	0.64 ± 0.05

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
8ilq.1	64.33	hetero-1-1-mer	0.58	BLAST	EM	-	0.52	A: 23-523 B: 1-508	0.98	Envelopment polyprotein Envelopment polyprotein

Excluded ligands

Ligand Name.Number	Reason for Exclusion	Description
NAG.2	Clashing with protein.	2-acetamido-2-deoxy-beta-D-glucopyranose
NAG.3	Not in contact with model.	2-acetamido-2-deoxy-beta-D-glucopyranose
NAG.4	Clashing with protein.	2-acetamido-2-deoxy-beta-D-glucopyranose

Ligand Name.Number	Reason for Exclusion	Description
NAG.5	Binding site not conserved.	2-acetamido-2-deoxy-beta-D-glucopyranose
NAG-NAG.1	Binding site not conserved.	2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

```
Target    MIVPIVLFLTLCPSELSAWGSPGDPIVCGVRTETNKSIQIEWKEGRSEKLCQIDRLGHVTSWLRNHSSFQGLIGQVKGRP
8ilq.1.A  ------------------GGDSGPIICAGPIHSNKSANIPHLLGYSEKICQIDRLIHVSSWLRNHSQFQGYVGQRGGRS

Target    SVSYFPEGASYPRWSGLLSPCDAEWLGLIAVSKAGDTDMIVPGPTYKGKIFVERPTYNGYKGWGCADGKSLSHSGTYCET
8ilq.1.A  QVSYYPAENSYSRWSGLLSPCDADWLGMLVVKKAKGSDMIVPGPSYKGKVFFERPTFDGYVGWGCSSGKSRTESGELCSS

Target    DSSVSSGLIQGDRVLWVGEVVCQRGTPVPEDVFSELVSLSQSEFPDVCKIDGVALNQCEQESIPQPLDVAWIDVGRSHKV
8ilq.1.A  DSGTSSGLLPSDRVLWIGDVACQPMTPIPEETFLELKSFSQSEFPDICKIDGVVFNQCEGESLPQPFDVAWMDVGHSHKI

Target    LMREHKTKWVQESSAKDFVCFKVGQGPCSKQEEDDCMSKGNCHGDEVFCRMAGCSARMQDNQEGCRCELLQKPGEIIVNY
8ilq.1.A  IMREHKTKWVQESSSKDFVCYKAGTGPCSESEEKTCKTSGSCRGDMQFCKVAGCEHGEEASEAKCRCSLVHKPGEVVVSY

Target    GGVSVRPTCYGFSRMMATLEVHKPDRELTGCTGCHLECIEGGVKIVTLTSELRSATVCASHFCASAKGGSKTTDILFHTG
8ilq.1.A  GGMRVRPKCYGFSRMMATLEVNPPEQRTGQCTGCHLECINGGVRLITLTSELKSATVCASHFCSSATSGKKSTEIQFHSG

Target    ALVGPNSIRITGQLLDGSKFSFDGHCIFPDGCMALDCTFCKEFLRNPQCYPVKKWLFLVVVVMCCYCALMLLTNILRAIG
8ilq.1.A  SLVGRTAIHVKGALVDGTEFTFEGSCMFPDGCDAVDCTFCREFLKNPQCYPAKKWLFIIIVILLGYAGLMLLTNVLKAIG

Target    VWGTWVFAPIKLALALGLRLAKLSKKGLVAVVTRGQMIVNDELHQ
8ilq.1.A  IWGSWVIAPVKLIFAIIKKLMRTVSCLMGKLMDRGRQVIHEEI--


Target    SGCDELVHAESKSITCKSASGNEKECSVTGRALLPAVNPGQEACLHFSMPGSPDSKCLKIKVKSINLRCKQASSYYVPEA
8ilq.1.B  ESCDEMVHADSKLVSCRQGSGNMKECVTTGRALLPAVNPGQEACLHFTAPGSPDSKCLKIKVKRINLKCKKSSSYFVPDA

Target    KARCTSVRRCRWAGDCQSGCPTYFSSNSFSDDWANRMDRAGLGMSGCSDGCGGAACGCFNAAPSCIFWRKWVENPSNRVW
8ilq.1.B  RSRCTSVRRCRWAGDCQSGCPPHSTSNSFSDDWAGKMDRAGLGFSGCSDGCGGAACGCFNAAPSCIFWRKWVENPHGIIW

Target    KVSPCASWVLAATIELTLPSGEVKTLEPVTGQATQMFKGVAITYLGSSIEIVGMTRLCEMKEMGTGIMALAPCNDPGHAI
8ilq.1.B  KVSPCAAWVPSAVIELTMPSGEVRTFHPMSGIPTQVFKGVSVTYLGSDMEVSGLTDLCEIEELKSKKLALAPCNQAGMGV

Target    MGNVGEIQCSSIESAKHIRSDGCIWNADLVGIELRVDDAVCFSKLTSVEAVANFSKIPATISGVRFDQGNHGESRIYGSP
8ilq.1.B  VGKVGEIQCSSEESARTIKKDGCIWNADLVGIELRVDDAVCYSKITSVEAVANYSAIPTTIGGLRFERSHDSLGKISGSP

Target    LDITRVSGEFSVSFRGMRLKLSEISASCTGEITNVSGCYSCMTGASVSIKLHSSKNTTGHLKCDSDETAFSVMEGTHTYR
8ilq.1.B  LDITAIRGSFSVNYRGLRLSLSEITATCTGEVTNVSGCYSCMTGAKVSIKLHSSKNSTAHVRCKGDETAFSVLEGVHSYT

Target    PHMSFDKAVIDEECVLNCGGHSSKLLLKGSLVFMDVPRFVDGSYVQTYHSKVPAGGRVPNPVDWLNALFGDGITRWILGI
8ilq.1.B  VSLSFDHAVVDEQCQLNCGGHESQVTLKGNLIFLDVPKFVDGSYMQTYHSSVPTGANIPSPTDWLNALFGNGLSRWILGV

Target    IGVLLACVMLFVVVVAITRRLIKGLT-QRAKVA
8ilq.1.B  IGVLLGGLALFFLIMSLFKLGTKQVFRSRTKL-
```

Materials and Methods

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 (Studer et al.). 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.

Model Quality Estimation

The global and per-residue model quality has been assessed using the QMEAN scoring function (Studer et al.).

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

- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL
  BLAST+: architecture and applications.
  BMC Bioinformatics, 10, 421-430. (2009) 20003500 10.1186/1471-2105-10-421
- Steinegger M, Meier M, Mirdita M, Vöhringer H, Haunsberger SJ, Söding J
  HH-suite3 for fast remote homology detection and deep protein annotation.
  BMC Bioinformatics 20, 473. (2019) 31521110 10.1186/s12859-019-3019-7

Table T1:

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

```
MIVPIVLFLTLCPSELSAWGSPGDPIVCGVRTETNKSIQIEWKEGRSEKLCQIDRLGHVTSWLRNHSSFQGLIGQVKGRPSVSYFPEGASYPRWSGLLSP
CDAEWLGLIAVSKAGDTDMIVPGPTYKGKIFVERPTYNGYKGWGCADGKSLSHSGTYCETDSSVSSGLIQGDRVLWVGEVVCQRGTPVPEDVFSELVSLS
QSEFPDVCKIDGVALNQCEQESIPQPLDVAWIDVGRSHKVLMREHKTKWVQESSAKDFVCFKVGQGPCSKQEEDDCMSKGNCHGDEVFCRMAGCSARMQD
NQEGCRCELLQKPGEIIVNYGGVSVRPTCYGFSRMMATLEVHKPDRELTGCTGCHLECIEGGVKIVTLTSELRSATVCASHFCASAKGGSKTTDILFHTG
ALVGPNSIRITGQLLDGSKFSFDGHCIFPDGCMALDCTFCKEFLRNPQCYPVKKWLFLVVVVMCCYCALMLLTNILRAIGVWGTWVFAPIKLALALGLRL
AKLSKKGLVAVVTRGQMIVNDELHQ

SGCDELVHAESKSITCKSASGNEKECSVTGRALLPAVNPGQEACLHFSMPGSPDSKCLKIKVKSINLRCKQASSYYVPEAKARCTSVRRCRWAGDCQSGC
PTYFSSNSFSDDWANRMDRAGLGMSGCSDGCGGAACGCFNAAPSCIFWRKWVENPSNRVWKVSPCASWVLAATIELTLPSGEVKTLEPVTGQATQMFKGV
AITYLGSSIEIVGMTRLCEMKEMGTGIMALAPCNDPGHAIMGNVGEIQCSSIESAKHIRSDGCIWNADLVGIELRVDDAVCFSKLTSVEAVANFSKIPAT
ISGVRFDQGNHGESRIYGSPLDITRVSGEFSVSFRGMRLKLSEISASCTGEITNVSGCYSCMTGASVSIKLHSSKNTTGHLKCDSDETAFSVMEGTHTYR
PHMSFDKAVIDEECVLNCGGHSSKLLLKGSLVFMDVPRFVDGSYVQTYHSKVPAGGRVPNPVDWLNALFGDGITRWILGIIGVLLACVMLFVVVVAITRR
LIKGLTQRAKVA
```

Table T2:

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
8ilq.1	64.33	hetero-1-1-mer	0.58	BLAST / HHblits	EM	NA	0.52	0.98	Envelopment polyprotein; Envelopment polyprotein
7x72.1	63.10	hetero-5-5-mer	0.56	HHblits	EM	7.20Å	0.51	0.99	Envelopment polyprotein; Envelopment polyprotein
7x72.1	63.59	hetero-5-5-mer	0.56	BLAST / HHblits	EM	7.20Å	0.51	0.99	Envelopment polyprotein; Envelopment polyprotein
7x6w.1	63.59	hetero-6-6-mer	0.46	BLAST / HHblits	EM	NA	0.51	0.99	Envelopment polyprotein; Envelopment polyprotein

The table above shows the top 4 filtered templates. A further 9 templates were found which were considered to be less suitable for modelling than the filtered list.
6f9b.1, 6f9c.1, 6f9d.1, 6f9e.1, 6f9f.1, 7x6w.1, 7x72.1, 8i4t.1, 8ilq.1