Full wwPDB X-ray Structure Validation Report (i)

Mar 9, 2026 – 09:10 PM UTC

PDB ID	:	3CVQ / pdb_00003cvq
Title	:	Structure of Peroxisomal Targeting Signal 1 (PTS1) binding domain of Trypanosoma brucei Peroxin 5 (TbPEX5)complexed to PTS1 peptide (7-SKL)
Authors	:	Sampathkumar, P.; Roach, C.; Michels, P.A.M.; Hol, W.G.J.
Deposited on	:	2008-04-18
Resolution	:	3.01 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

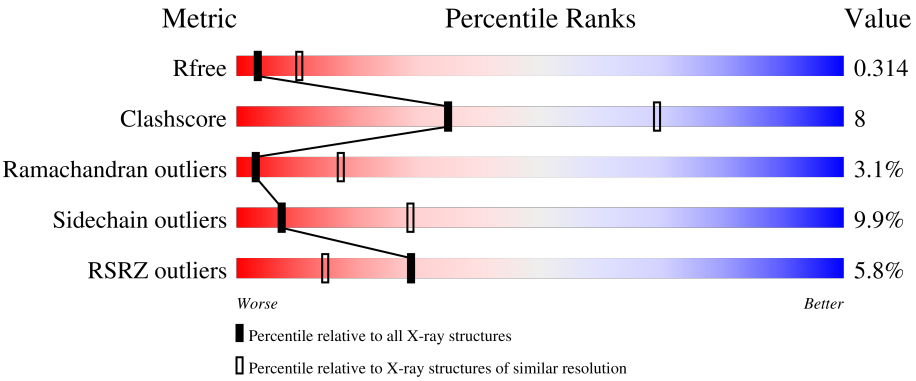
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.01 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	3131 (3.04-3.00)
Clashscore	190562	3444 (3.04-3.00)
Ramachandran outliers	187476	3319 (3.04-3.00)
Sidechain outliers	187428	3322 (3.04-3.00)
RSRZ outliers	180081	3130 (3.04-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	327	5% 68% 17% •• 12%
2	B	7	14% 71% 29%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2284 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Peroxisome targeting signal 1 receptor PEX5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	0	0	0
			2200	1381	382	425	12			

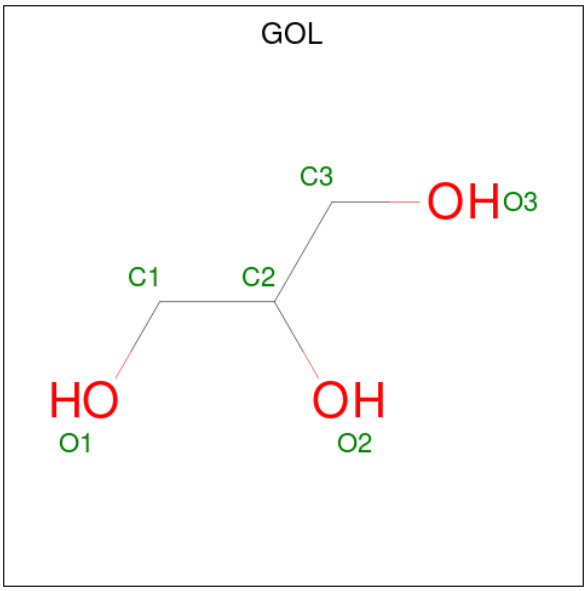
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	329	GLY	-	expression tag	UNP Q9U7C3
A	330	HIS	-	expression tag	UNP Q9U7C3
A	331	MET	-	expression tag	UNP Q9U7C3
A	411	ALA	MET	engineered mutation	UNP Q9U7C3
A	415	ALA	LYS	engineered mutation	UNP Q9U7C3

- Molecule 2 is a protein called PTS1 peptide 7-SKL (Ac-SNRWSKL).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	7	Total	C	N	O	0	0	0
			66	41	13	12			

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0

- Molecule 4 is water.

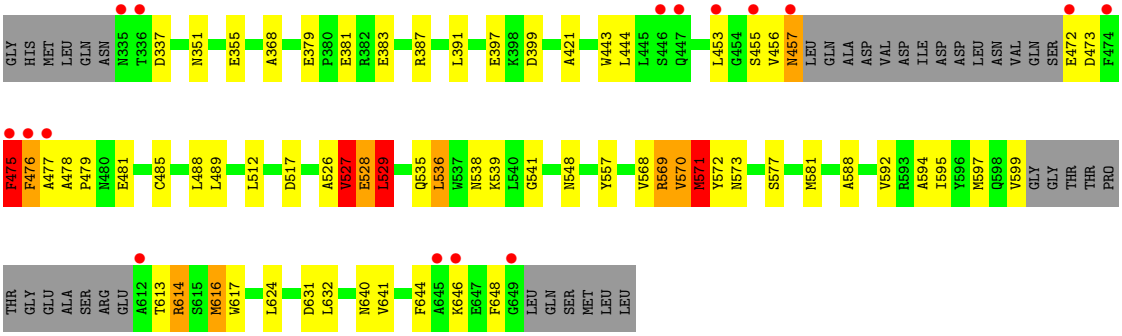
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	11	Total O 11 11	0	0
4	B	1	Total O 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

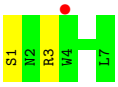
• Molecule 1: Peroxisome targeting signal 1 receptor PEX5

Chain A:



• Molecule 2: PTS1 peptide 7-SKL (Ac-SNRWSKL)

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	68.79Å 68.79Å 230.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.62 – 3.01 47.62 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.62-3.01) 99.8 (47.62-3.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.15 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.273 , 0.321 0.317 , 0.314	Depositor DCC
R_{free} test set	553 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å^2)	79.9	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 65.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	2284	wwPDB-VP
Average B, all atoms (Å^2)	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.01% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SAC

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.70	0/2246	1.08	8/3067 (0.3%)
2	B	0.91	0/58	0.91	0/75
All	All	0.71	0/2304	1.07	8/3142 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	476	PHE	N-CA-C	9.01	123.53	110.10
1	A	529	LEU	N-CA-C	-7.45	99.25	110.28
1	A	527	VAL	N-CA-C	-6.94	102.37	110.21
1	A	640	ASN	N-CA-C	6.81	118.95	109.15
1	A	569	ARG	N-CA-C	-6.74	101.47	110.24
1	A	568	VAL	N-CA-C	6.18	116.34	110.53
1	A	571	MET	N-CA-C	-5.50	105.19	111.07
1	A	475	PHE	N-CA-C	5.11	121.68	110.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	475	PHE	Peptide
1	A	477	ALA	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2200	0	2043	33	0
2	B	66	0	66	1	0
3	A	6	0	8	0	0
4	A	11	0	0	0	0
4	B	1	0	0	0	0
All	All	2284	0	2117	33	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (33) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:527:VAL:O	1:A:528:GLU:HB2	1.59	0.98
1:A:456:VAL:O	1:A:457:ASN:HB2	1.92	0.69
1:A:444:LEU:HD13	1:A:489:LEU:HD23	1.76	0.68
1:A:569:ARG:O	1:A:570:VAL:HB	1.94	0.66
1:A:571:MET:HB3	1:A:594:ALA:HB2	1.78	0.65
1:A:527:VAL:O	1:A:528:GLU:CB	2.36	0.65
1:A:570:VAL:HA	1:A:573:ASN:HB2	1.81	0.62
1:A:613:THR:OG1	1:A:617:TRP:NE1	2.36	0.58
1:A:613:THR:HG1	1:A:617:TRP:CD1	2.23	0.57
1:A:538:ASN:OD1	1:A:570:VAL:HG22	2.04	0.57
1:A:613:THR:OG1	1:A:617:TRP:CD1	2.62	0.53
1:A:456:VAL:O	1:A:457:ASN:CB	2.56	0.52
1:A:595:ILE:HG12	1:A:616:MET:HG2	1.92	0.51
1:A:577:SER:O	1:A:581:MET:HG3	2.11	0.51
1:A:572:TYR:OH	2:B:3:ARG:NH1	2.45	0.50
1:A:569:ARG:O	1:A:570:VAL:CB	2.59	0.48
1:A:541:GLY:HA3	1:A:557:TYR:CE1	2.48	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:527:VAL:HG12	1:A:536:LEU:HB3	1.96	0.47
1:A:472:GLU:HG3	1:A:473:ASP:H	1.80	0.47
1:A:485:CYS:HA	1:A:488:LEU:HB2	1.97	0.47
1:A:526:ALA:O	1:A:529:LEU:O	2.33	0.47
1:A:399:ASP:CG	1:A:569:ARG:HH22	2.24	0.46
1:A:481:GLU:H	1:A:481:GLU:CD	2.24	0.46
1:A:631:ASP:O	1:A:632:LEU:HB2	2.18	0.44
1:A:613:THR:O	1:A:614:ARG:CB	2.64	0.44
1:A:535:GLN:O	1:A:539:LYS:HG2	2.18	0.44
1:A:614:ARG:C	1:A:616:MET:N	2.76	0.43
1:A:588:ALA:O	1:A:592:VAL:HG23	2.17	0.43
1:A:614:ARG:C	1:A:616:MET:H	2.26	0.43
1:A:421:ALA:HB2	1:A:443:TRP:CD1	2.54	0.42
1:A:592:VAL:HG13	1:A:641:VAL:CG2	2.50	0.42
1:A:387:ARG:O	1:A:391:LEU:HB2	2.20	0.41
1:A:397:GLU:O	1:A:397:GLU:HG3	2.20	0.41

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	283/327 (86%)	250 (88%)	24 (8%)	9 (3%)	3 17
2	B	5/7 (71%)	5 (100%)	0	0	100 100
All	All	288/334 (86%)	255 (88%)	24 (8%)	9 (3%)	3 17

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	528	GLU
1	A	368	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	475	PHE
1	A	614	ARG
1	A	644	PHE
1	A	478	ALA
1	A	351	ASN
1	A	479	PRO
1	A	570	VAL

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	216/266 (81%)	194 (90%)	22 (10%)	7 27
2	B	6/6 (100%)	6 (100%)	0	100 100
All	All	222/272 (82%)	200 (90%)	22 (10%)	7 28

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	337	ASP
1	A	355	GLU
1	A	379	GLU
1	A	381	GLU
1	A	383	GLU
1	A	453	LEU
1	A	455	SER
1	A	457	ASN
1	A	476	PHE
1	A	512	LEU
1	A	517	ASP
1	A	527	VAL
1	A	529	LEU
1	A	536	LEU
1	A	548	ASN
1	A	571	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	597	MET
1	A	599	VAL
1	A	616	MET
1	A	624	LEU
1	A	646	LYS
1	A	648	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	396	ASN
1	A	447	GLN
1	A	490	HIS
1	A	515	ASN
1	A	522	ASN
1	A	564	ASN
1	A	640	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SAC	B	1	2	7,8,9	0.32	0	7,9,11	1.55	2 (28%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns.

'-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SAC	B	1	2	-	2/7/8/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	SAC	O-C-CA	-3.17	116.62	124.77
2	B	1	SAC	CB-CA-C	-2.43	106.00	111.35

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	SAC	C-CA-CB-OG
2	B	1	SAC	N-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	GOL	A	1	-	5,5,5	0.39	0	5,5,5	0.32	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	1	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1	GOL	O1-C1-C2-C3
3	A	1	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	289/327 (88%)	0.17	16 (5%) 30 15	2, 10, 15, 18	0
2	B	6/7 (85%)	1.04	1 (16%) 4 3	15, 17, 18, 21	0
All	All	295/334 (88%)	0.19	17 (5%) 29 14	2, 10, 15, 21	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	477	ALA	7.6
1	A	476	PHE	7.0
1	A	335	ASN	5.9
1	A	649	GLY	5.4
1	A	475	PHE	4.6
1	A	446	SER	4.2
1	A	457	ASN	4.1
1	A	472	GLU	4.0
1	A	646	LYS	3.7
1	A	474	PHE	3.1
1	A	645	ALA	2.8
1	A	447	GLN	2.7
1	A	612	ALA	2.6
1	A	336	THR	2.4
1	A	455	SER	2.2
1	A	453	LEU	2.1
2	B	4	TRP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SAC	B	1	9/10	0.78	0.16	15,15,16,16	0

6.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GOL	A	1	6/6	0.71	0.17	22,23,23,23	0

6.5 Other polymers (i)

There are no such residues in this entry.