



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2023 – 10:05 AM EDT

PDB ID : 1TR0  
Title : Crystal Structure of a boiling stable protein SP1  
Authors : Almog, O.; Gonzalez, A.; Sofer, O.; Dgany, O.; Shoseyov, O.  
Deposited on : 2004-06-18  
Resolution : 1.80 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

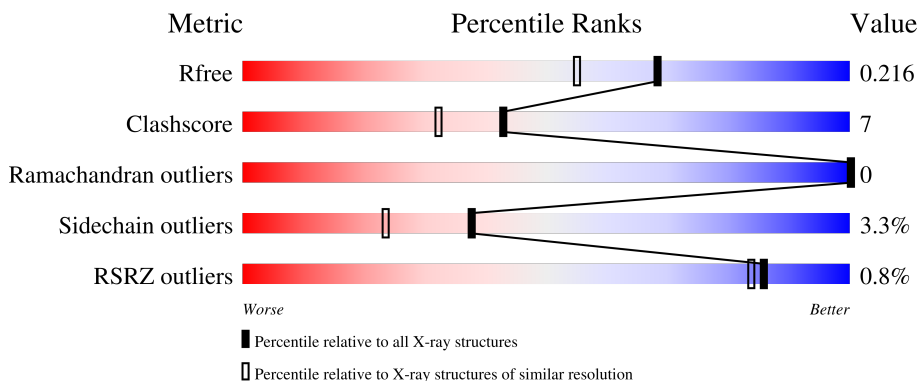
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	108	 83% 15% .
1	B	108	 88% 9% ..
1	C	108	 86% 11% ..
1	D	108	 83% 13% ..
1	E	108	 83% 13% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	F	108	87%	11%
1	G	108	81%	15%
1	H	108	85%	11%
1	I	108	82%	14%
1	J	108	89%	7%
1	K	108	86%	11%
1	L	108	83%	13%
1	M	108	83%	13%
1	N	108	86%	12%
1	O	108	86%	10%
1	P	108	84%	11%
1	R	108	83%	14%
1	S	108	87%	10%
1	T	108	86%	12%
1	U	108	81%	15%
1	V	108	87%	9%
1	W	108	85%	9%
1	X	108	83%	10% 5%
1	Y	108	83%	13%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	H	4820	-	-	X	-
2	GOL	L	5320	-	-	X	-
2	GOL	U	6820	-	-	X	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 24095 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 stable protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	106	867	555	139	171	2	0	1	0
1	B	106	867	555	139	171	2	0	1	0
1	C	106	867	555	139	171	2	0	1	0
1	D	106	863	553	139	169	2	0	0	0
1	E	106	867	555	139	171	2	0	1	0
1	F	106	867	555	139	171	2	0	1	0
1	G	106	863	553	139	169	2	0	0	0
1	H	106	863	553	139	169	2	0	0	0
1	I	106	867	555	139	171	2	0	1	0
1	J	106	867	555	139	171	2	0	1	0
1	K	106	867	555	139	171	2	0	1	0
1	L	106	863	553	139	169	2	0	0	0
1	M	106	863	553	139	169	2	0	0	0
1	N	106	867	555	139	171	2	0	1	0
1	O	106	863	553	139	169	2	0	0	0
1	P	106	863	553	139	169	2	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			
1	S	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			
1	T	106	Total	C	N	O	S	0	0	0
			863	553	139	169	2			
1	U	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			
1	V	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			
1	W	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			
1	X	106	Total	C	N	O	S	0	0	0
			863	553	139	169	2			
1	Y	106	Total	C	N	O	S	0	1	0
			867	555	139	171	2			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total 6	C 3	O 3	0	0
2	E	1	Total 6	C 3	O 3	0	0
2	F	1	Total 6	C 3	O 3	0	0
2	G	1	Total 6	C 3	O 3	0	0
2	H	1	Total 6	C 3	O 3	0	0
2	I	1	Total 6	C 3	O 3	0	0
2	J	1	Total 6	C 3	O 3	0	0
2	K	1	Total 6	C 3	O 3	0	0
2	L	1	Total 6	C 3	O 3	0	0
2	M	1	Total 6	C 3	O 3	0	0
2	N	1	Total 6	C 3	O 3	0	0
2	O	1	Total 6	C 3	O 3	0	0
2	P	1	Total 6	C 3	O 3	0	0
2	R	1	Total 6	C 3	O 3	0	0
2	S	1	Total 6	C 3	O 3	0	0
2	T	1	Total 6	C 3	O 3	0	0
2	U	1	Total 6	C 3	O 3	0	0
2	V	1	Total 6	C 3	O 3	0	0
2	W	1	Total 6	C 3	O 3	0	0
2	X	1	Total 6	C 3	O 3	0	0
2	Y	1	Total 6	C 3	O 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	121	Total 121	O 121	0	0
3	B	127	Total 127	O 127	0	0
3	C	124	Total 124	O 124	0	0
3	D	140	Total 140	O 140	0	0
3	E	135	Total 135	O 135	0	0
3	F	131	Total 131	O 131	0	0
3	G	129	Total 129	O 129	0	0
3	H	148	Total 148	O 148	0	0
3	I	149	Total 149	O 149	0	0
3	J	142	Total 142	O 142	0	0
3	K	152	Total 152	O 152	0	0
3	L	131	Total 131	O 131	0	0
3	M	133	Total 133	O 133	0	0
3	N	132	Total 132	O 132	0	0
3	O	143	Total 143	O 143	0	0
3	P	138	Total 138	O 138	0	0
3	R	138	Total 138	O 138	0	0
3	S	105	Total 105	O 105	0	0
3	T	117	Total 117	O 117	0	0
3	U	121	Total 121	O 121	0	0
3	V	124	Total 124	O 124	0	0

*Continued on next page...*

*Continued from previous page...*

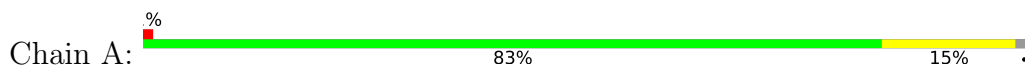
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	W	133	Total 133	O 133	0	0
3	X	144	Total 144	O 144	0	0
3	Y	122	Total 122	O 122	0	0



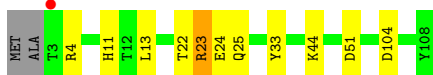
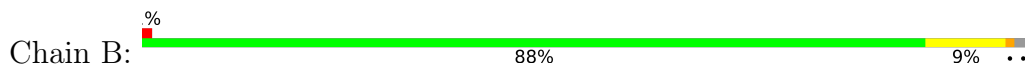
### 3 Residue-property plots [i](#)

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.

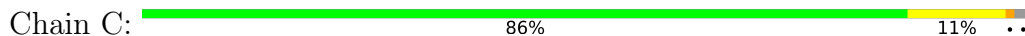
- Molecule 1: stable protein 1



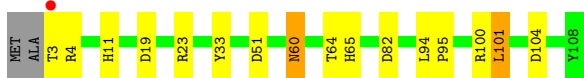
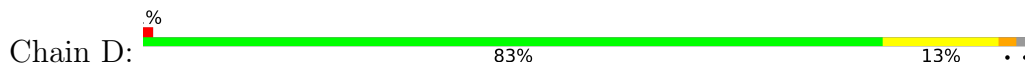
- Molecule 1: stable protein 1



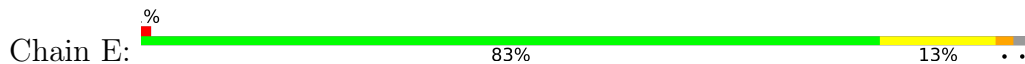
- Molecule 1: stable protein 1



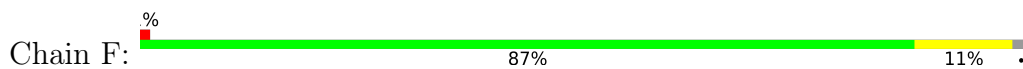
- Molecule 1: stable protein 1



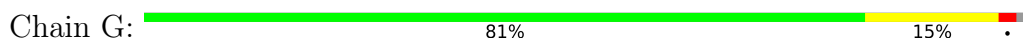
- Molecule 1: stable protein 1



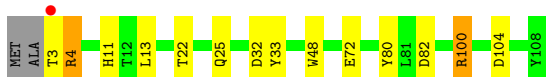
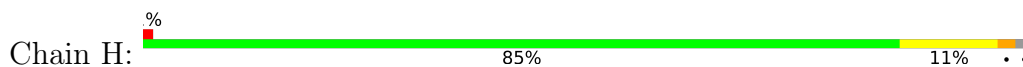
- Molecule 1: stable protein 1



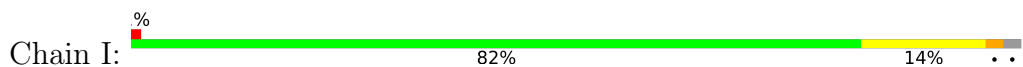
- Molecule 1: stable protein 1



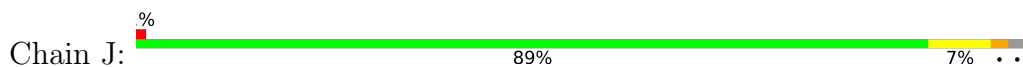
- Molecule 1: stable protein 1



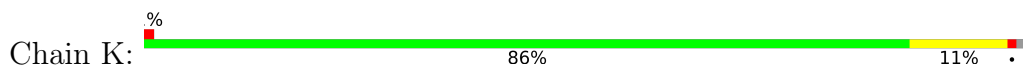
- Molecule 1: stable protein 1



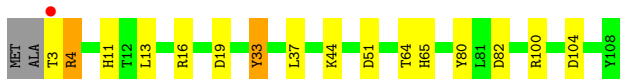
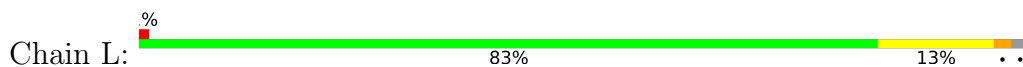
- Molecule 1: stable protein 1



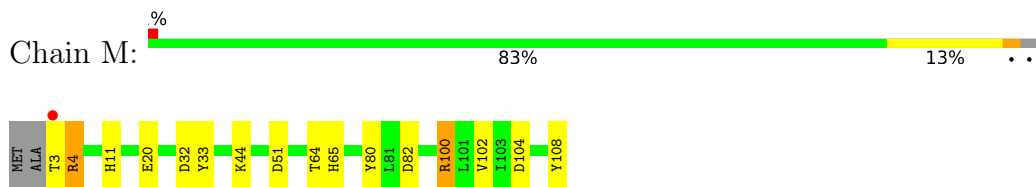
- Molecule 1: stable protein 1



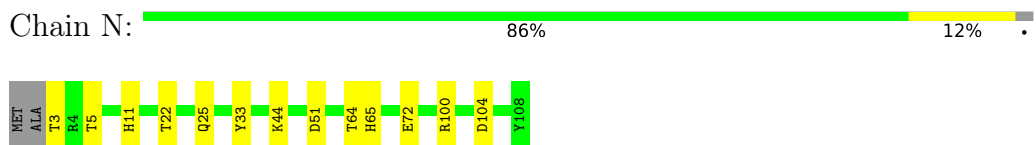
- Molecule 1: stable protein 1



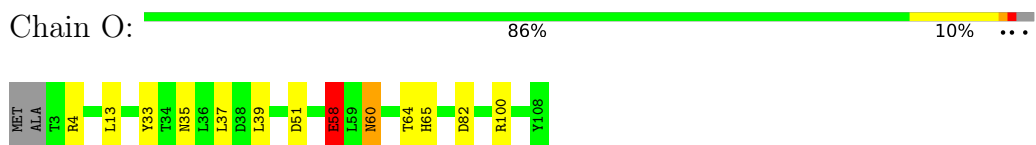
- Molecule 1: stable protein 1



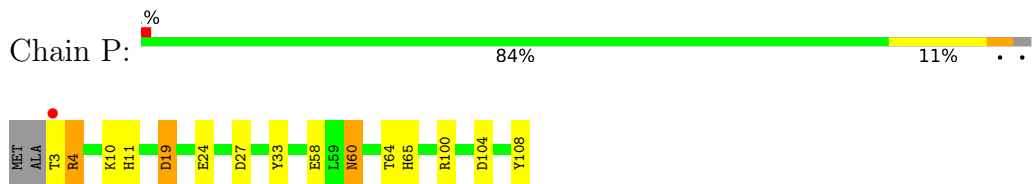
- Molecule 1: stable protein 1



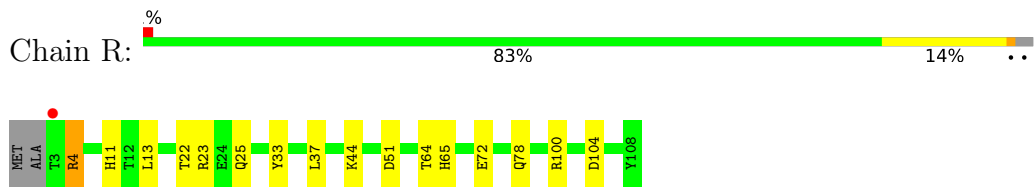
- Molecule 1: stable protein 1



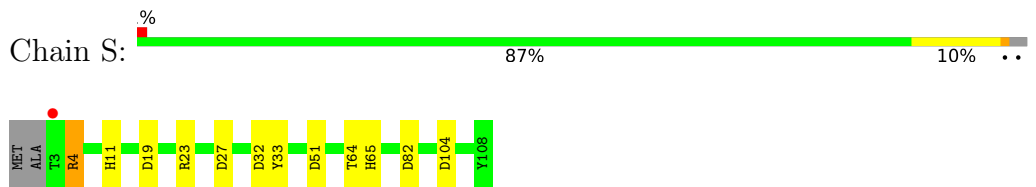
- Molecule 1: stable protein 1



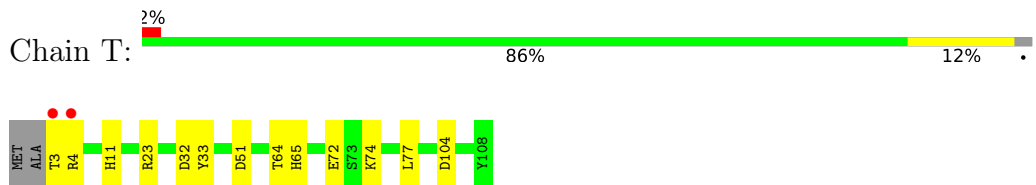
- Molecule 1: stable protein 1



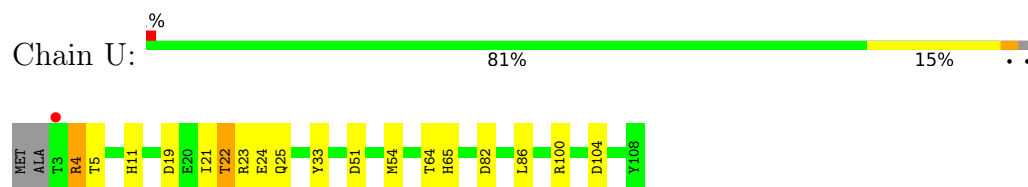
- Molecule 1: stable protein 1



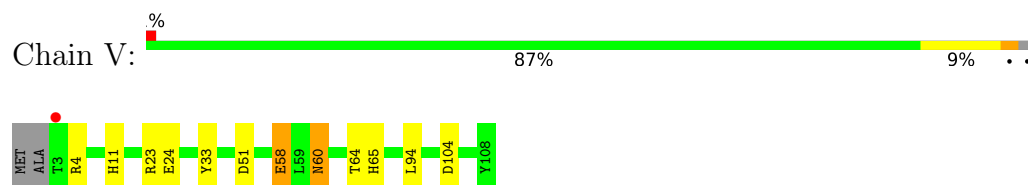
- Molecule 1: stable protein 1



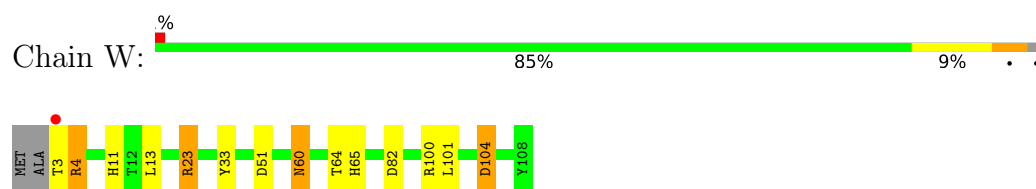
- Molecule 1: stable protein 1



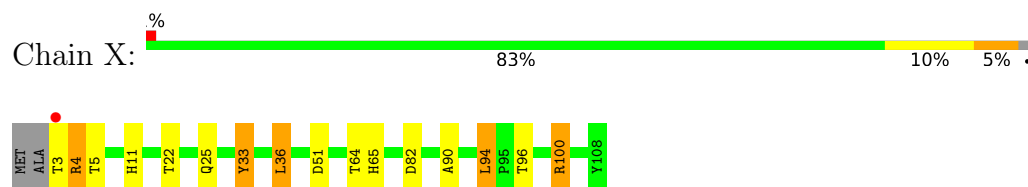
- Molecule 1: stable protein 1



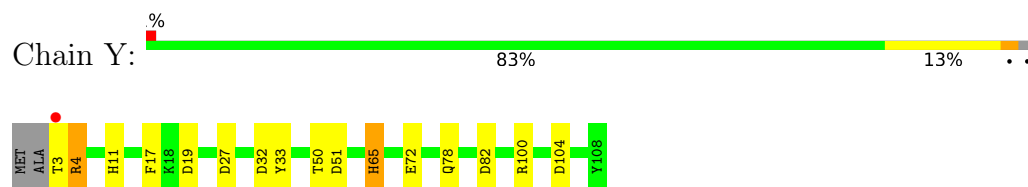
- Molecule 1: stable protein 1



- Molecule 1: stable protein 1



- Molecule 1: stable protein 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.03Å 94.75Å 168.03Å 90.00° 90.11° 90.00°	Depositor
Resolution (Å)	48.50 – 1.80 48.51 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.50-1.80) 92.3 (48.51-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.70 (at 1.79Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.162 , 0.202 0.182 , 0.216	Depositor DCC
$R_{free}$ test set	13492 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.2	Xtrriage
Anisotropy	0.201	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 67.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l 0.000 for -k,-h,-l 0.037 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	24095	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	10.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 67.50 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.0968e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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

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	0.79	0/891	1.01	6/1205 (0.5%)
1	B	0.76	0/891	0.92	1/1205 (0.1%)
1	C	0.75	0/891	0.94	5/1205 (0.4%)
1	D	0.77	0/882	0.98	5/1193 (0.4%)
1	E	0.82	0/891	0.96	3/1205 (0.2%)
1	F	0.83	0/891	0.99	4/1205 (0.3%)
1	G	0.87	0/882	1.00	9/1193 (0.8%)
1	H	0.86	0/882	0.95	4/1193 (0.3%)
1	I	0.89	1/891 (0.1%)	1.08	8/1205 (0.7%)
1	J	0.82	0/891	0.92	3/1205 (0.2%)
1	K	0.79	0/891	0.98	4/1205 (0.3%)
1	L	0.79	0/882	0.93	4/1193 (0.3%)
1	M	0.77	0/882	0.92	5/1193 (0.4%)
1	N	0.83	0/891	0.94	2/1205 (0.2%)
1	O	0.82	1/882 (0.1%)	0.97	4/1193 (0.3%)
1	P	0.81	0/882	0.91	3/1193 (0.3%)
1	R	0.80	0/891	0.92	1/1205 (0.1%)
1	S	0.78	0/891	0.98	3/1205 (0.2%)
1	T	0.79	0/882	0.92	4/1193 (0.3%)
1	U	0.80	0/891	0.93	3/1205 (0.2%)
1	V	0.81	0/891	0.97	1/1205 (0.1%)
1	W	0.80	0/891	0.93	3/1205 (0.2%)
1	X	0.77	0/882	0.93	3/1193 (0.3%)
1	Y	0.82	0/891	0.96	5/1205 (0.4%)
All	All	0.81	2/21303 (0.0%)	0.96	93/28812 (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	L	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	58	GLU	CG-CD	-6.73	1.41	1.51
1	I	16	ARG	NE-CZ	5.49	1.40	1.33

The worst 5 of 93 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	16	ARG	NE-CZ-NH1	13.18	126.89	120.30
1	A	51	ASP	CB-CG-OD2	10.37	127.64	118.30
1	B	51	ASP	CB-CG-OD2	9.73	127.06	118.30
1	F	51	ASP	CB-CG-OD2	9.50	126.85	118.30
1	H	100	ARG	NE-CZ-NH1	9.44	125.02	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	L	3	THR	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	867	0	839	5	0
1	B	867	0	839	10	0
1	C	867	0	839	12	0
1	D	863	0	837	8	0
1	E	867	0	839	16	0
1	F	867	0	839	9	0
1	G	863	0	837	17	0
1	H	863	0	837	12	0
1	I	867	0	839	10	0
1	J	867	0	839	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	K	867	0	839	9	0
1	L	863	0	837	13	0
1	M	863	0	837	15	0
1	N	867	0	839	9	0
1	O	863	0	837	9	0
1	P	863	0	837	15	0
1	R	867	0	839	15	0
1	S	867	0	839	13	0
1	T	863	0	837	10	0
1	U	867	0	839	19	0
1	V	867	0	839	10	0
1	W	867	0	839	14	0
1	X	863	0	837	12	0
1	Y	867	0	839	13	0
2	A	6	0	8	1	0
2	B	6	0	8	1	0
2	C	6	0	8	1	0
2	D	6	0	8	2	0
2	E	6	0	8	3	0
2	F	6	0	8	2	0
2	G	6	0	8	1	0
2	H	6	0	8	4	0
2	I	6	0	8	3	0
2	J	6	0	8	1	0
2	K	6	0	8	0	0
2	L	6	0	8	6	0
2	M	6	0	8	3	0
2	N	6	0	8	1	0
2	O	6	0	8	2	0
2	P	6	0	8	1	0
2	R	6	0	8	3	0
2	S	6	0	8	1	0
2	T	6	0	8	1	0
2	U	6	0	8	4	0
2	V	6	0	8	1	0
2	W	6	0	8	3	0
2	X	6	0	8	2	0
2	Y	6	0	8	2	0
3	A	121	0	0	2	0
3	B	127	0	0	5	0
3	C	124	0	0	1	0
3	D	140	0	0	3	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	135	0	0	12	0
3	F	131	0	0	2	0
3	G	129	0	0	7	0
3	H	148	0	0	3	0
3	I	149	0	0	2	0
3	J	142	0	0	2	0
3	K	152	0	0	3	0
3	L	131	0	0	4	0
3	M	133	0	0	4	0
3	N	132	0	0	3	0
3	O	143	0	0	8	0
3	P	138	0	0	8	0
3	R	138	0	0	2	0
3	S	105	0	0	10	0
3	T	117	0	0	9	0
3	U	121	0	0	8	0
3	V	124	0	0	2	0
3	W	133	0	0	7	0
3	X	144	0	0	3	0
3	Y	122	0	0	3	0
All	All	24095	0	20310	280	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 7.

The worst 5 of 280 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:V:4:ARG:HD3	3:X:7265:HOH:O	1.62	0.99
1:P:19:ASP:HB3	3:P:6524:HOH:O	1.61	0.99
1:U:22:THR:HB	1:U:25:GLN:HE21	1.31	0.96
1:E:22:THR:H	1:E:25:GLN:HE21	1.17	0.91
1:X:11:HIS:NE2	2:X:7220:GOL:H32	1.87	0.89

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	B	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	C	105/108 (97%)	105 (100%)	0	0	100	100
1	D	104/108 (96%)	103 (99%)	1 (1%)	0	100	100
1	E	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	F	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	G	104/108 (96%)	104 (100%)	0	0	100	100
1	H	104/108 (96%)	104 (100%)	0	0	100	100
1	I	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	J	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	K	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	L	104/108 (96%)	103 (99%)	1 (1%)	0	100	100
1	M	104/108 (96%)	104 (100%)	0	0	100	100
1	N	105/108 (97%)	105 (100%)	0	0	100	100
1	O	104/108 (96%)	103 (99%)	1 (1%)	0	100	100
1	P	104/108 (96%)	103 (99%)	1 (1%)	0	100	100
1	R	105/108 (97%)	105 (100%)	0	0	100	100
1	S	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	T	104/108 (96%)	104 (100%)	0	0	100	100
1	U	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	V	105/108 (97%)	104 (99%)	1 (1%)	0	100	100
1	W	105/108 (97%)	105 (100%)	0	0	100	100
1	X	104/108 (96%)	103 (99%)	1 (1%)	0	100	100
1	Y	105/108 (97%)	105 (100%)	0	0	100	100
All	All	2511/2592 (97%)	2496 (99%)	15 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	95/95 (100%)	90 (95%)	5 (5%)	22 9
1	B	95/95 (100%)	92 (97%)	3 (3%)	39 25
1	C	95/95 (100%)	93 (98%)	2 (2%)	53 42
1	D	94/95 (99%)	89 (95%)	5 (5%)	22 9
1	E	95/95 (100%)	92 (97%)	3 (3%)	39 25
1	F	95/95 (100%)	94 (99%)	1 (1%)	73 68
1	G	94/95 (99%)	90 (96%)	4 (4%)	29 14
1	H	94/95 (99%)	92 (98%)	2 (2%)	53 42
1	I	95/95 (100%)	90 (95%)	5 (5%)	22 9
1	J	95/95 (100%)	93 (98%)	2 (2%)	53 42
1	K	95/95 (100%)	92 (97%)	3 (3%)	39 25
1	L	94/95 (99%)	91 (97%)	3 (3%)	39 25
1	M	94/95 (99%)	92 (98%)	2 (2%)	53 42
1	N	95/95 (100%)	93 (98%)	2 (2%)	53 42
1	O	94/95 (99%)	91 (97%)	3 (3%)	39 25
1	P	94/95 (99%)	89 (95%)	5 (5%)	22 9
1	R	95/95 (100%)	92 (97%)	3 (3%)	39 25
1	S	95/95 (100%)	93 (98%)	2 (2%)	53 42
1	T	94/95 (99%)	92 (98%)	2 (2%)	53 42
1	U	95/95 (100%)	91 (96%)	4 (4%)	30 15
1	V	95/95 (100%)	91 (96%)	4 (4%)	30 15
1	W	95/95 (100%)	91 (96%)	4 (4%)	30 15
1	X	94/95 (99%)	90 (96%)	4 (4%)	29 14
1	Y	95/95 (100%)	92 (97%)	3 (3%)	39 25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2271/2280 (100%)	2195 (97%)	76 (3%)	38 23

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	U	4	ARG
1	X	36	LEU
1	U	22	THR
1	W	4	ARG
1	Y	65	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 72 such sidechains are listed below:

Mol	Chain	Res	Type
1	U	25	GLN
1	Y	11	HIS
1	V	11	HIS
1	W	65	HIS
1	H	28	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

24 ligands are 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	GOL	O	6320	-	5,5,5	0.50	0	5,5,5	0.58	0
2	GOL	P	6420	-	5,5,5	0.66	0	5,5,5	1.05	1 (20%)
2	GOL	R	6520	-	5,5,5	0.43	0	5,5,5	0.87	0
2	GOL	E	4520	-	5,5,5	0.79	0	5,5,5	0.94	0
2	GOL	Y	7320	-	5,5,5	0.57	0	5,5,5	0.71	0
2	GOL	C	4320	-	5,5,5	0.61	0	5,5,5	0.86	0
2	GOL	B	4220	-	5,5,5	0.50	0	5,5,5	0.68	0
2	GOL	F	4620	-	5,5,5	0.53	0	5,5,5	0.95	0
2	GOL	L	5320	-	5,5,5	0.76	0	5,5,5	1.08	0
2	GOL	S	6620	-	5,5,5	0.45	0	5,5,5	0.71	0
2	GOL	T	6720	-	5,5,5	0.55	0	5,5,5	1.06	0
2	GOL	N	6220	-	5,5,5	0.49	0	5,5,5	0.73	0
2	GOL	K	5220	-	5,5,5	0.42	0	5,5,5	0.94	0
2	GOL	W	7120	-	5,5,5	0.45	0	5,5,5	1.00	0
2	GOL	M	6120	-	5,5,5	0.62	0	5,5,5	0.94	0
2	GOL	I	4920	-	5,5,5	0.55	0	5,5,5	0.77	0
2	GOL	G	4720	-	5,5,5	0.70	0	5,5,5	1.15	0
2	GOL	D	4420	-	5,5,5	0.63	0	5,5,5	0.97	0
2	GOL	H	4820	-	5,5,5	0.58	0	5,5,5	0.93	0
2	GOL	A	4120	-	5,5,5	0.46	0	5,5,5	0.50	0
2	GOL	V	6920	-	5,5,5	0.53	0	5,5,5	0.97	0
2	GOL	X	7220	-	5,5,5	0.68	0	5,5,5	1.57	2 (40%)
2	GOL	J	5120	-	5,5,5	0.46	0	5,5,5	1.14	0
2	GOL	U	6820	-	5,5,5	0.56	0	5,5,5	0.66	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
2	GOL	O	6320	-	-	4/4/4/4	-
2	GOL	P	6420	-	-	2/4/4/4	-
2	GOL	R	6520	-	-	2/4/4/4	-
2	GOL	E	4520	-	-	2/4/4/4	-
2	GOL	Y	7320	-	-	2/4/4/4	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	C	4320	-	-	3/4/4/4	-
2	GOL	B	4220	-	-	2/4/4/4	-
2	GOL	F	4620	-	-	2/4/4/4	-
2	GOL	L	5320	-	-	2/4/4/4	-
2	GOL	S	6620	-	-	2/4/4/4	-
2	GOL	T	6720	-	-	2/4/4/4	-
2	GOL	N	6220	-	-	4/4/4/4	-
2	GOL	K	5220	-	-	2/4/4/4	-
2	GOL	W	7120	-	-	4/4/4/4	-
2	GOL	M	6120	-	-	2/4/4/4	-
2	GOL	I	4920	-	-	4/4/4/4	-
2	GOL	G	4720	-	-	2/4/4/4	-
2	GOL	D	4420	-	-	4/4/4/4	-
2	GOL	H	4820	-	-	4/4/4/4	-
2	GOL	A	4120	-	-	4/4/4/4	-
2	GOL	V	6920	-	-	2/4/4/4	-
2	GOL	X	7220	-	-	2/4/4/4	-
2	GOL	J	5120	-	-	2/4/4/4	-
2	GOL	U	6820	-	-	2/4/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	X	7220	GOL	O3-C3-C2	2.50	122.18	110.20
2	X	7220	GOL	C3-C2-C1	-2.05	103.74	111.70
2	P	6420	GOL	C3-C2-C1	-2.04	103.77	111.70

There are no chirality outliers.

5 of 63 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	4120	GOL	O1-C1-C2-C3
2	A	4120	GOL	C1-C2-C3-O3
2	B	4220	GOL	C1-C2-C3-O3
2	C	4320	GOL	C1-C2-C3-O3
2	D	4420	GOL	O1-C1-C2-C3

There are no ring outliers.

23 monomers are involved in 49 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	O	6320	GOL	2	0
2	P	6420	GOL	1	0
2	R	6520	GOL	3	0
2	E	4520	GOL	3	0
2	Y	7320	GOL	2	0
2	C	4320	GOL	1	0
2	B	4220	GOL	1	0
2	F	4620	GOL	2	0
2	L	5320	GOL	6	0
2	S	6620	GOL	1	0
2	T	6720	GOL	1	0
2	N	6220	GOL	1	0
2	W	7120	GOL	3	0
2	M	6120	GOL	3	0
2	I	4920	GOL	3	0
2	G	4720	GOL	1	0
2	D	4420	GOL	2	0
2	H	4820	GOL	4	0
2	A	4120	GOL	1	0
2	V	6920	GOL	1	0
2	X	7220	GOL	2	0
2	J	5120	GOL	1	0
2	U	6820	GOL	4	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	106/108 (98%)	-0.54	1 (0%) 84 82	3, 6, 15, 38	0
1	B	106/108 (98%)	-0.39	1 (0%) 84 82	3, 6, 14, 38	0
1	C	106/108 (98%)	-0.46	0 100 100	3, 6, 14, 38	0
1	D	106/108 (98%)	-0.52	1 (0%) 84 82	3, 6, 14, 38	0
1	E	106/108 (98%)	-0.58	1 (0%) 84 82	3, 7, 15, 38	0
1	F	106/108 (98%)	-0.52	1 (0%) 84 82	3, 7, 15, 38	0
1	G	106/108 (98%)	-0.46	0 100 100	4, 7, 15, 38	0
1	H	106/108 (98%)	-0.36	1 (0%) 84 82	4, 7, 14, 38	0
1	I	106/108 (98%)	-0.26	1 (0%) 84 82	3, 6, 13, 38	0
1	J	106/108 (98%)	-0.44	1 (0%) 84 82	3, 6, 14, 39	0
1	K	106/108 (98%)	-0.54	1 (0%) 84 82	3, 6, 14, 38	0
1	L	106/108 (98%)	-0.52	1 (0%) 84 82	3, 6, 13, 38	0
1	M	106/108 (98%)	-0.68	1 (0%) 84 82	2, 5, 12, 38	0
1	N	106/108 (98%)	-0.64	0 100 100	2, 5, 14, 37	0
1	O	106/108 (98%)	-0.57	0 100 100	2, 6, 13, 38	0
1	P	106/108 (98%)	-0.62	1 (0%) 84 82	2, 6, 14, 38	0
1	R	106/108 (98%)	-0.55	1 (0%) 84 82	3, 6, 14, 38	0
1	S	106/108 (98%)	-0.35	1 (0%) 84 82	3, 6, 14, 38	0
1	T	106/108 (98%)	-0.40	2 (1%) 66 63	3, 6, 14, 38	0
1	U	106/108 (98%)	-0.55	1 (0%) 84 82	3, 6, 14, 37	0
1	V	106/108 (98%)	-0.50	1 (0%) 84 82	2, 5, 13, 38	0
1	W	106/108 (98%)	-0.64	1 (0%) 84 82	2, 5, 13, 38	0
1	X	106/108 (98%)	-0.69	1 (0%) 84 82	2, 5, 13, 37	0
1	Y	106/108 (98%)	-0.61	1 (0%) 84 82	2, 4, 13, 38	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	2544/2592 (98%)	-0.52	21 (0%) 86 84	2, 6, 14, 39	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	3	THR	4.6
1	H	3	THR	4.3
1	I	3	THR	4.0
1	D	3	THR	4.0
1	X	3	THR	4.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	GOL	W	7120	6/6	0.73	0.25	23,37,40,42	0
2	GOL	E	4520	6/6	0.78	0.23	23,32,34,36	0
2	GOL	K	5220	6/6	0.79	0.19	23,34,34,38	0
2	GOL	V	6920	6/6	0.81	0.19	25,36,37,37	0
2	GOL	L	5320	6/6	0.82	0.21	21,34,36,36	0
2	GOL	X	7220	6/6	0.82	0.24	24,36,37,37	0
2	GOL	F	4620	6/6	0.83	0.18	26,37,40,40	0
2	GOL	C	4320	6/6	0.83	0.21	23,33,34,37	0
2	GOL	M	6120	6/6	0.84	0.20	23,36,39,40	0
2	GOL	O	6320	6/6	0.84	0.17	24,35,36,38	0
2	GOL	S	6620	6/6	0.84	0.21	28,37,39,42	0
2	GOL	T	6720	6/6	0.85	0.15	29,35,35,38	0
2	GOL	B	4220	6/6	0.85	0.24	25,38,41,44	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GOL	U	6820	6/6	0.86	0.17	25,34,35,38	0
2	GOL	G	4720	6/6	0.86	0.18	23,30,32,36	0
2	GOL	H	4820	6/6	0.87	0.15	24,34,36,36	0
2	GOL	J	5120	6/6	0.87	0.17	21,33,35,37	0
2	GOL	D	4420	6/6	0.87	0.21	24,33,36,38	0
2	GOL	P	6420	6/6	0.88	0.20	23,35,36,39	0
2	GOL	A	4120	6/6	0.89	0.15	25,37,39,40	0
2	GOL	N	6220	6/6	0.89	0.17	23,35,37,39	0
2	GOL	R	6520	6/6	0.91	0.18	26,36,38,39	0
2	GOL	Y	7320	6/6	0.92	0.17	16,32,34,36	0
2	GOL	I	4920	6/6	0.94	0.13	14,28,30,34	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.