



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 2, 2022 – 10:20 AM JST

PDB ID : 7V4Z  
Title : Structure of Horcolin native form  
Authors : Bobbili, K.B.; Sivaji, N.; Jayaprakash, N.G.; Narayanan, V.; Sekhar, A.; Suguna, K.; Surolia, A.  
Deposited on : 2021-08-16  
Resolution : 1.16 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](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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

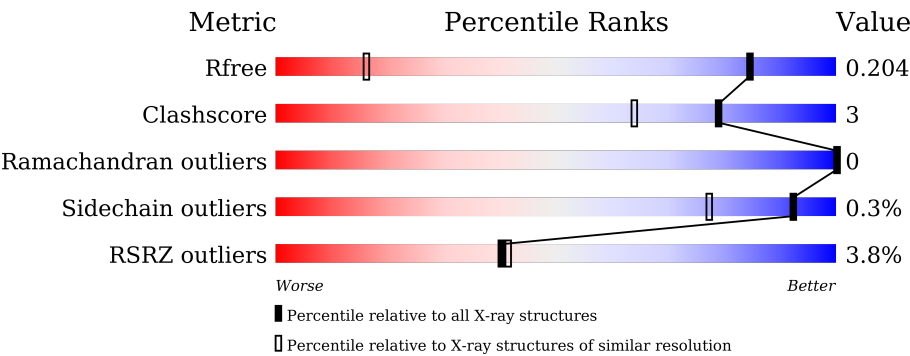
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.27  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

# 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 1.16 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	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)
RSRZ outliers	127900	1724 (1.20-1.12)

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	146	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>91%8%</div></div>
1	B	146	<div><div>13%</div><div></div><div></div><div></div><div></div></div> <div>94%</div>
1	C	146	<div><div>3%</div><div></div><div></div><div></div><div></div></div> <div>96%</div>
1	D	146	<div><div></div><div></div><div></div><div></div><div></div></div> <div>98%</div>
1	E	146	<div><div>3%</div><div></div><div></div><div></div><div></div></div> <div>95%</div>
1	F	146	<div><div>3%</div><div></div><div></div><div></div><div></div></div> <div>93%5%</div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	146	<div><div></div><div>3%</div><div>95%</div><div></div><div>..</div></div>
1	H	146	<div><div></div><div>%</div><div>91%</div><div>6%</div><div>..</div></div>
1	I	146	<div><div></div><div>5%</div><div>92%</div><div>5%</div><div>..</div></div>
1	J	146	<div><div></div><div>5%</div><div>93%</div><div>5%</div><div>.</div></div>
1	K	146	<div><div></div><div>2%</div><div>92%</div><div>5%</div><div>..</div></div>
1	L	146	<div><div></div><div>3%</div><div>94%</div><div></div><div>...</div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13981 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 Horcolin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	0	0
			1033	657	171	203	2			
1	C	144	Total	C	N	O	S	0	0	0
			1041	662	173	204	2			
1	B	143	Total	C	N	O	S	0	0	0
			1025	653	169	201	2			
1	D	144	Total	C	N	O	S	0	0	0
			1047	665	176	204	2			
1	E	143	Total	C	N	O	S	0	0	0
			1034	657	172	203	2			
1	F	143	Total	C	N	O	S	0	0	0
			1036	659	173	202	2			
1	G	144	Total	C	N	O	S	0	0	0
			1043	663	175	203	2			
1	H	143	Total	C	N	O	S	0	0	0
			1038	660	174	202	2			
1	I	143	Total	C	N	O	S	0	0	0
			1042	662	175	203	2			
1	J	144	Total	C	N	O	S	0	0	0
			1029	655	171	201	2			
1	K	143	Total	C	N	O	S	0	0	0
			1038	660	174	202	2			
1	L	143	Total	C	N	O	S	0	0	0
			1038	660	174	202	2			

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



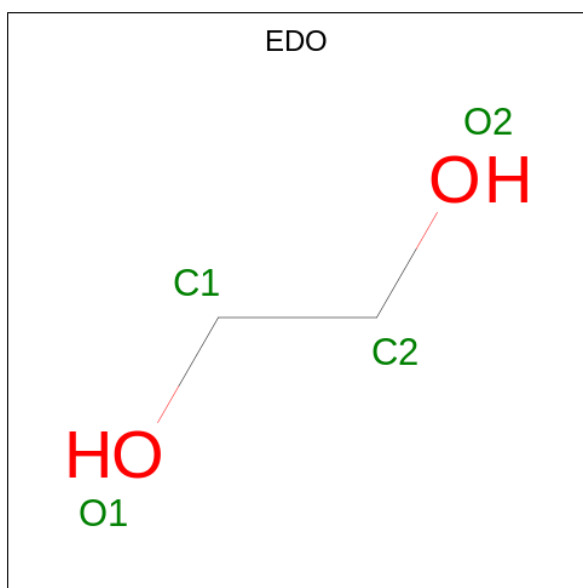
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total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

*Continued on next page...*

Continued from previous page...

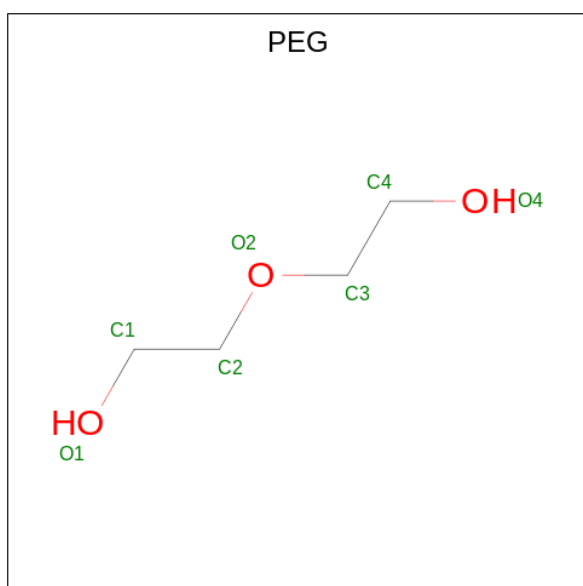
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	G	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		
2	I	1	Total	C	O	0	0
			6	3	3		
2	I	1	Total	C	O	0	0
			6	3	3		
2	J	1	Total	C	O	0	0
			6	3	3		
2	K	1	Total	C	O	0	0
			6	3	3		
2	K	1	Total	C	O	0	0
			6	3	3		
2	K	1	Total	C	O	0	0
			6	3	3		
2	L	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	K	1	Total C O 4 2 2	0	0
3	K	1	Total C O 4 2 2	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	1	Total C O 7 4 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	126	Total O 126 126	0	0
5	C	118	Total O 118 118	0	0
5	B	94	Total O 94 94	0	0
5	D	116	Total O 116 116	0	0
5	E	116	Total O 116 116	0	0

*Continued on next page...*

*Continued from previous page...*

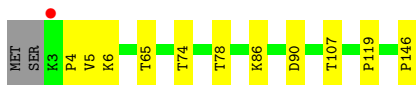
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	113	Total 113	O 113	0	0
5	G	124	Total 124	O 124	0	0
5	H	119	Total 119	O 119	0	0
5	I	107	Total 107	O 107	0	0
5	J	82	Total 82	O 82	0	0
5	K	129	Total 129	O 129	0	0
5	L	114	Total 114	O 114	0	0



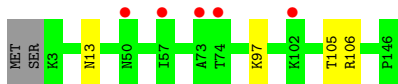
### 3 Residue-property plots [i](#)

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.

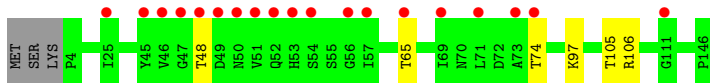
- Molecule 1: Horcolin



- Molecule 1: Horcolin



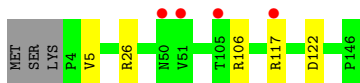
- Molecule 1: Horcolin



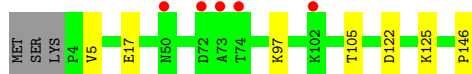
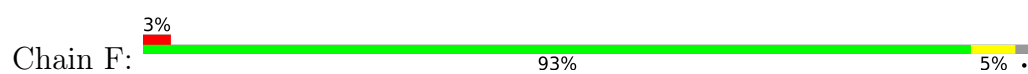
- Molecule 1: Horcolin



- Molecule 1: Horcolin



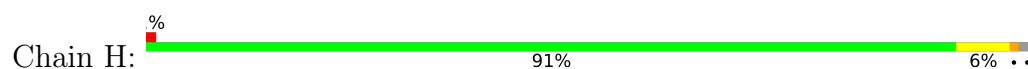
- Molecule 1: Horcolin



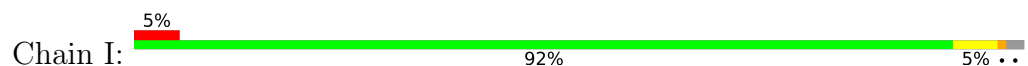
• Molecule 1: Horcolin



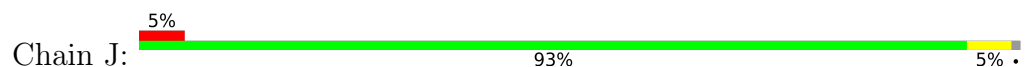
• Molecule 1: Horcolin



• Molecule 1: Horcolin



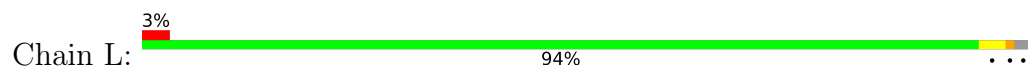
• Molecule 1: Horcolin



• Molecule 1: Horcolin



• Molecule 1: Horcolin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.18Å 94.42Å 238.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.87 – 1.16 55.80 – 1.16	Depositor EDS
% Data completeness (in resolution range)	96.7 (55.87-1.16) 96.7 (55.80-1.16)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 1.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.181 , 0.196 0.190 , 0.204	Depositor DCC
$R_{free}$ test set	25278 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	11.8	Xtrriage
Anisotropy	0.198	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 47.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	13981	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	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 38.22 % 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 3.8112e-04. 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: EDO, GOL, PEG

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	0.90	1/1055 (0.1%)	1.00	3/1436 (0.2%)
1	B	0.88	0/1047	0.97	1/1424 (0.1%)
1	C	0.87	0/1063	0.91	1/1445 (0.1%)
1	D	0.84	0/1069	0.88	0/1452
1	E	0.86	0/1056	0.98	2/1434 (0.1%)
1	F	0.85	1/1058 (0.1%)	0.93	3/1437 (0.2%)
1	G	0.85	0/1065	0.96	4/1447 (0.3%)
1	H	0.86	1/1060 (0.1%)	0.96	3/1439 (0.2%)
1	I	0.87	1/1064 (0.1%)	0.96	1/1444 (0.1%)
1	J	0.84	1/1051 (0.1%)	1.02	3/1431 (0.2%)
1	K	0.88	1/1060 (0.1%)	0.97	3/1439 (0.2%)
1	L	0.88	0/1060	0.99	4/1439 (0.3%)
All	All	0.87	6/12708 (0.0%)	0.96	28/17267 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	17	GLU	CD-OE1	7.99	1.34	1.25
1	I	66	GLU	CD-OE1	6.63	1.32	1.25
1	H	79	GLU	CD-OE1	-6.44	1.18	1.25
1	J	79	GLU	CD-OE1	-5.76	1.19	1.25
1	A	5	VAL	CB-CG2	-5.30	1.41	1.52
1	F	17	GLU	CD-OE2	-5.05	1.20	1.25

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	122	ASP	CB-CG-OD2	-9.67	109.60	118.30
1	J	18	ARG	NE-CZ-NH1	-9.19	115.70	120.30
1	E	122	ASP	CB-CG-OD1	9.08	126.47	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	117	ARG	NE-CZ-NH1	-8.73	115.94	120.30
1	A	146	PRO	CA-C-O	-7.89	101.27	120.20
1	A	90	ASP	CB-CG-OD1	-7.84	111.24	118.30
1	J	146	PRO	CA-C-O	-7.59	101.98	120.20
1	K	132	ARG	NE-CZ-NH1	7.22	123.91	120.30
1	L	146	PRO	CA-C-O	-6.84	103.79	120.20
1	G	117	ARG	NE-CZ-NH1	-6.79	116.91	120.30
1	G	18	ARG	NE-CZ-NH2	-6.64	116.98	120.30
1	G	106	ARG	NE-CZ-NH2	-6.48	117.06	120.30
1	C	106	ARG	NE-CZ-NH2	-6.33	117.14	120.30
1	H	146	PRO	CA-C-O	-6.23	105.25	120.20
1	E	106	ARG	NE-CZ-NH2	-6.06	117.27	120.30
1	H	122	ASP	CB-CG-OD2	-6.04	112.86	118.30
1	L	122	ASP	CB-CG-OD2	-5.93	112.97	118.30
1	F	122	ASP	CB-CG-OD2	-5.77	113.10	118.30
1	A	5	VAL	CA-CB-CG2	5.61	119.32	110.90
1	K	117	ARG	NE-CZ-NH1	-5.61	117.49	120.30
1	K	5	VAL	CG1-CB-CG2	5.48	119.67	110.90
1	L	117	ARG	CG-CD-NE	5.47	123.29	111.80
1	B	106	ARG	NE-CZ-NH2	-5.30	117.65	120.30
1	I	26	ARG	NE-CZ-NH2	-5.29	117.66	120.30
1	F	122	ASP	CB-CG-OD1	5.28	123.05	118.30
1	G	146	PRO	CA-C-O	-5.08	108.00	120.20
1	H	112	THR	CA-CB-CG2	5.06	119.48	112.40
1	F	146	PRO	CA-C-O	-5.01	108.17	120.20

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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	1033	0	1008	9	0
1	B	1025	0	1003	6	0
1	C	1041	0	1022	7	0
1	D	1047	0	1036	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1034	0	1017	3	0
1	F	1036	0	1022	6	0
1	G	1043	0	1030	1	0
1	H	1038	0	1029	7	0
1	I	1042	0	1035	8	0
1	J	1029	0	1000	4	0
1	K	1038	0	1029	12	0
1	L	1038	0	1029	8	0
2	A	12	0	16	0	0
2	B	6	0	8	0	0
2	C	24	0	32	0	0
2	D	18	0	24	0	0
2	E	12	0	16	0	0
2	F	6	0	8	0	0
2	G	24	0	32	2	0
2	H	12	0	16	0	0
2	I	12	0	16	0	0
2	J	6	0	8	0	0
2	K	18	0	24	0	0
2	L	6	0	8	0	0
3	A	8	0	12	2	0
3	K	8	0	12	2	0
4	G	7	0	10	3	0
5	A	126	0	0	5	0
5	B	94	0	0	3	0
5	C	118	0	0	1	0
5	D	116	0	0	0	0
5	E	116	0	0	1	0
5	F	113	0	0	4	0
5	G	124	0	0	3	0
5	H	119	0	0	4	0
5	I	107	0	0	1	0
5	J	82	0	0	3	0
5	K	129	0	0	4	0
5	L	114	0	0	5	0
All	All	13981	0	12502	71	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 3.

All (71) 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:C:97:LYS:HE3	1:C:105:THR:HG21	1.23	1.18
3:K:304:EDO:H12	5:K:427:HOH:O	1.46	1.13
1:K:125:LYS:HE2	5:K:458:HOH:O	1.46	1.13
1:J:57:ILE:HG13	5:J:414:HOH:O	1.55	1.06
1:C:97:LYS:CE	1:C:105:THR:HG21	1.84	1.05
1:I:13:ASN:HB2	5:I:483:HOH:O	1.57	1.03
1:C:97:LYS:HE3	1:C:105:THR:CG2	1.91	0.99
1:A:65:THR:HG21	1:D:13:ASN:HD21	1.28	0.98
1:J:112:THR:HB	5:J:427:HOH:O	1.67	0.94
1:C:13:ASN:HD21	1:B:65:THR:HG21	1.32	0.94
1:I:27:MET:H	1:I:125:LYS:HZ1	1.27	0.80
1:A:107:THR:HG21	5:B:437:HOH:O	1.83	0.78
1:C:97:LYS:CE	1:C:105:THR:CG2	2.57	0.73
5:C:504:HOH:O	1:B:65:THR:HG22	1.89	0.72
1:F:125:LYS:O	1:F:125:LYS:CD	2.40	0.70
1:F:125:LYS:CD	1:F:125:LYS:C	2.61	0.69
1:H:105:THR:HG22	5:H:439:HOH:O	1.93	0.68
1:L:102:LYS:HD2	5:L:406:HOH:O	1.91	0.68
1:I:117:ARG:NH2	1:I:119:PRO:HG3	2.10	0.67
1:K:25:ILE:O	1:K:125:LYS:NZ	2.31	0.63
1:C:97:LYS:HE2	1:C:105:THR:HG21	1.77	0.63
1:K:132:ARG:NH1	5:K:402:HOH:O	2.32	0.62
1:A:4:PRO:O	3:A:203:EDO:H22	2.00	0.62
1:H:65:THR:HG23	5:H:457:HOH:O	2.02	0.60
1:B:65:THR:HG23	5:B:453:HOH:O	2.02	0.58
3:K:304:EDO:C1	5:K:427:HOH:O	2.23	0.58
1:L:102:LYS:CE	5:L:406:HOH:O	2.51	0.58
1:E:117:ARG:NH1	5:E:401:HOH:O	2.32	0.58
1:K:26:ARG:CB	1:K:125:LYS:CE	2.82	0.57
1:I:27:MET:H	1:I:125:LYS:NZ	2.02	0.57
5:A:306:HOH:O	1:L:102:LYS:CD	2.52	0.57
1:B:48:THR:HG21	5:B:487:HOH:O	2.05	0.55
1:K:26:ARG:HB2	1:K:125:LYS:HE2	1.89	0.54
1:A:65:THR:HG23	5:A:343:HOH:O	2.07	0.53
5:A:306:HOH:O	1:L:102:LYS:HD2	2.07	0.53
1:H:57:ILE:N	1:H:57:ILE:HD12	2.23	0.53
1:I:27:MET:N	1:I:125:LYS:HZ1	2.02	0.53
1:L:102:LYS:CD	5:L:406:HOH:O	2.55	0.53
1:A:78:THR:HG21	5:L:489:HOH:O	2.07	0.53
1:I:26:ARG:HA	1:I:125:LYS:NZ	2.25	0.52
1:E:5:VAL:HG11	5:F:506:HOH:O	2.07	0.52
2:G:302:GOL:O3	4:G:303:PEG:H41	2.09	0.51

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:26:ARG:HB3	1:K:125:LYS:CE	2.41	0.51
1:H:57:ILE:HD12	1:H:57:ILE:H	1.75	0.51
1:A:65:THR:CG2	1:D:13:ASN:HD21	2.13	0.51
1:F:125:LYS:CD	5:F:401:HOH:O	2.58	0.50
2:G:302:GOL:O3	4:G:303:PEG:C4	2.59	0.50
1:L:117:ARG:NH1	5:L:402:HOH:O	2.45	0.50
1:H:115:PRO:HB3	5:H:444:HOH:O	2.11	0.50
1:K:26:ARG:CB	1:K:125:LYS:HE2	2.42	0.49
1:A:6:LYS:HB2	3:A:203:EDO:H21	1.96	0.48
1:J:63:GLY:HA3	1:K:86:LYS:HZ3	1.78	0.48
1:I:26:ARG:HA	1:I:125:LYS:HZ2	1.79	0.48
1:B:74:THR:HG21	1:I:9:PRO:HD3	1.95	0.47
4:G:303:PEG:H41	5:G:425:HOH:O	2.14	0.47
1:H:57:ILE:HD13	5:H:493:HOH:O	2.15	0.47
1:C:97:LYS:HE2	1:C:105:THR:CG2	2.40	0.46
1:J:63:GLY:HA3	1:K:86:LYS:NZ	2.31	0.46
1:H:94:THR:HB	1:H:112:THR:HG22	1.98	0.46
1:K:26:ARG:HB2	1:K:125:LYS:CE	2.46	0.45
1:E:26:ARG:CB	5:G:522:HOH:O	2.64	0.45
1:F:5:VAL:HG11	5:F:506:HOH:O	2.17	0.45
5:J:427:HOH:O	1:L:115:PRO:HG3	2.15	0.45
5:A:306:HOH:O	1:L:102:LYS:HD3	2.14	0.44
1:F:5:VAL:CG1	5:F:506:HOH:O	2.65	0.43
1:F:97:LYS:HE3	1:F:105:THR:HG21	2.01	0.43
1:K:26:ARG:HB3	1:K:125:LYS:NZ	2.33	0.43
1:B:97:LYS:HE3	1:B:105:THR:CG2	2.49	0.42
1:A:74:THR:HG21	1:K:9:PRO:HD3	2.02	0.41
1:G:57:ILE:HD12	5:G:501:HOH:O	2.20	0.41
1:A:86:LYS:CB	5:A:413:HOH:O	2.68	0.40

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	142/146 (97%)	137 (96%)	5 (4%)	0	100	100
1	B	141/146 (97%)	137 (97%)	4 (3%)	0	100	100
1	C	142/146 (97%)	139 (98%)	3 (2%)	0	100	100
1	D	142/146 (97%)	139 (98%)	3 (2%)	0	100	100
1	E	141/146 (97%)	138 (98%)	3 (2%)	0	100	100
1	F	141/146 (97%)	138 (98%)	3 (2%)	0	100	100
1	G	142/146 (97%)	139 (98%)	3 (2%)	0	100	100
1	H	141/146 (97%)	137 (97%)	4 (3%)	0	100	100
1	I	141/146 (97%)	138 (98%)	3 (2%)	0	100	100
1	J	142/146 (97%)	138 (97%)	4 (3%)	0	100	100
1	K	141/146 (97%)	137 (97%)	4 (3%)	0	100	100
1	L	141/146 (97%)	137 (97%)	4 (3%)	0	100	100
All	All	1697/1752 (97%)	1654 (98%)	43 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	107/113 (95%)	106 (99%)	1 (1%)	78	48
1	B	106/113 (94%)	106 (100%)	0	100	100
1	C	108/113 (96%)	108 (100%)	0	100	100
1	D	110/113 (97%)	110 (100%)	0	100	100
1	E	108/113 (96%)	108 (100%)	0	100	100
1	F	108/113 (96%)	108 (100%)	0	100	100
1	G	109/113 (96%)	109 (100%)	0	100	100
1	H	109/113 (96%)	108 (99%)	1 (1%)	78	48
1	I	110/113 (97%)	110 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	106/113 (94%)	105 (99%)	1 (1%)	78	48
1	K	109/113 (96%)	109 (100%)	0	100	100
1	L	109/113 (96%)	108 (99%)	1 (1%)	78	48
All	All	1299/1356 (96%)	1295 (100%)	4 (0%)	92	77

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	119	PRO
1	H	125	LYS
1	J	26	ARG
1	L	125	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	52	GLN
1	D	13	ASN
1	D	52	GLN
1	F	50	ASN
1	F	52	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

31 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	E	301	-	5,5,5	0.19	0	5,5,5	0.48	0
2	GOL	G	305	-	5,5,5	0.25	0	5,5,5	0.21	0
2	GOL	C	304	-	5,5,5	0.15	0	5,5,5	0.41	0
2	GOL	D	303	-	5,5,5	0.35	0	5,5,5	0.31	0
2	GOL	A	201	-	5,5,5	0.98	0	5,5,5	0.30	0
2	GOL	C	302	-	5,5,5	0.83	0	5,5,5	0.30	0
2	GOL	D	301	-	5,5,5	1.06	1 (20%)	5,5,5	0.62	0
2	GOL	G	301	-	5,5,5	0.12	0	5,5,5	0.35	0
2	GOL	E	302	-	5,5,5	0.44	0	5,5,5	0.62	0
3	EDO	A	204	-	3,3,3	0.19	0	2,2,2	0.33	0
2	GOL	I	301	-	5,5,5	0.18	0	5,5,5	0.40	0
2	GOL	I	302	-	5,5,5	0.60	0	5,5,5	0.49	0
2	GOL	K	303	-	5,5,5	0.17	0	5,5,5	0.35	0
2	GOL	J	301	-	5,5,5	0.20	0	5,5,5	0.28	0
2	GOL	K	301	-	5,5,5	0.15	0	5,5,5	0.33	0
2	GOL	F	301	-	5,5,5	0.24	0	5,5,5	0.29	0
2	GOL	D	302	-	5,5,5	0.28	0	5,5,5	0.46	0
2	GOL	A	202	-	5,5,5	0.53	0	5,5,5	0.36	0
2	GOL	C	303	-	5,5,5	0.18	0	5,5,5	0.15	0
2	GOL	H	302	-	5,5,5	0.21	0	5,5,5	0.61	0
2	GOL	L	301	-	5,5,5	0.22	0	5,5,5	0.28	0
3	EDO	K	305	-	3,3,3	0.23	0	2,2,2	0.34	0
2	GOL	C	301	-	5,5,5	0.17	0	5,5,5	0.83	0
2	GOL	H	301	-	5,5,5	0.20	0	5,5,5	0.34	0
3	EDO	K	304	-	3,3,3	0.36	0	2,2,2	0.19	0
2	GOL	G	304	-	5,5,5	0.30	0	5,5,5	0.43	0
3	EDO	A	203	-	3,3,3	0.12	0	2,2,2	0.38	0
2	GOL	K	302	-	5,5,5	0.59	0	5,5,5	0.54	0
4	PEG	G	303	-	6,6,6	0.53	0	5,5,5	0.76	0
2	GOL	B	301	-	5,5,5	0.18	0	5,5,5	0.48	0
2	GOL	G	302	-	5,5,5	0.80	0	5,5,5	0.60	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	E	301	-	-	0/4/4/4	-
2	GOL	G	305	-	-	0/4/4/4	-
2	GOL	C	304	-	-	4/4/4/4	-
2	GOL	D	303	-	-	0/4/4/4	-
2	GOL	A	201	-	-	0/4/4/4	-
2	GOL	C	302	-	-	0/4/4/4	-
2	GOL	D	301	-	-	0/4/4/4	-
2	GOL	G	301	-	-	0/4/4/4	-
2	GOL	E	302	-	-	0/4/4/4	-
3	EDO	A	204	-	-	1/1/1/1	-
2	GOL	I	301	-	-	0/4/4/4	-
2	GOL	I	302	-	-	0/4/4/4	-
2	GOL	K	303	-	-	0/4/4/4	-
2	GOL	J	301	-	-	0/4/4/4	-
2	GOL	K	301	-	-	0/4/4/4	-
2	GOL	F	301	-	-	0/4/4/4	-
2	GOL	D	302	-	-	0/4/4/4	-
2	GOL	A	202	-	-	0/4/4/4	-
2	GOL	C	303	-	-	0/4/4/4	-
2	GOL	H	302	-	-	0/4/4/4	-
2	GOL	L	301	-	-	0/4/4/4	-
3	EDO	K	305	-	-	0/1/1/1	-
2	GOL	C	301	-	-	0/4/4/4	-
2	GOL	H	301	-	-	0/4/4/4	-
3	EDO	K	304	-	-	1/1/1/1	-
2	GOL	G	304	-	-	0/4/4/4	-
3	EDO	A	203	-	-	1/1/1/1	-
2	GOL	K	302	-	-	0/4/4/4	-
4	PEG	G	303	-	-	1/4/4/4	-
2	GOL	B	301	-	-	0/4/4/4	-
2	GOL	G	302	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	301	GOL	O2-C2	2.36	1.50	1.43

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	304	GOL	C1-C2-C3-O3
2	C	304	GOL	O1-C1-C2-O2
2	C	304	GOL	O1-C1-C2-C3
2	C	304	GOL	O2-C2-C3-O3
3	A	204	EDO	O1-C1-C2-O2
3	A	203	EDO	O1-C1-C2-O2
4	G	303	PEG	C4-C3-O2-C2
3	K	304	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	K	304	EDO	2	0
3	A	203	EDO	2	0
4	G	303	PEG	3	0
2	G	302	GOL	2	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 ⓘ

### 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<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	144/146 (98%)	0.36	1 (0%) 87 88	7, 12, 20, 28	0
1	B	143/146 (97%)	0.84	19 (13%) 3 3	8, 17, 35, 45	0
1	C	144/146 (98%)	0.31	5 (3%) 44 45	7, 14, 21, 32	0
1	D	144/146 (98%)	0.22	2 (1%) 75 75	7, 14, 22, 25	0
1	E	143/146 (97%)	0.43	4 (2%) 53 53	7, 14, 26, 31	0
1	F	143/146 (97%)	0.41	5 (3%) 44 45	8, 15, 26, 40	0
1	G	144/146 (98%)	0.12	4 (2%) 53 53	9, 12, 19, 24	0
1	H	143/146 (97%)	0.25	2 (1%) 75 75	7, 13, 25, 30	0
1	I	143/146 (97%)	0.38	8 (5%) 24 25	9, 15, 24, 29	0
1	J	144/146 (98%)	0.49	7 (4%) 29 30	10, 17, 29, 35	0
1	K	143/146 (97%)	0.18	3 (2%) 63 63	8, 12, 20, 27	0
1	L	143/146 (97%)	0.30	5 (3%) 44 45	8, 13, 23, 31	0
All	All	1721/1752 (98%)	0.36	65 (3%) 40 41	7, 14, 26, 45	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	51	VAL	7.3
1	F	73	ALA	6.1
1	F	74	THR	5.2
1	B	49	ASP	4.5
1	B	52	GLN	4.4
1	J	104	VAL	4.0
1	I	13	ASN	3.6
1	B	46	VAL	3.6
1	B	50	ASN	3.5
1	B	57	ILE	3.5
1	I	74	THR	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	57	ILE	3.4
1	J	57	ILE	3.3
1	F	50	ASN	3.3
1	B	53	HIS	3.3
1	J	50	ASN	3.2
1	B	48	THR	3.1
1	I	73	ALA	3.1
1	I	50	ASN	2.9
1	D	69	ILE	2.9
1	A	3	LYS	2.8
1	E	50	ASN	2.8
1	G	74	THR	2.8
1	C	74	THR	2.7
1	E	117	ARG	2.7
1	L	74	THR	2.7
1	J	105	THR	2.7
1	B	54	SER	2.7
1	L	117	ARG	2.6
1	G	57	ILE	2.5
1	K	89	THR	2.5
1	B	111	GLY	2.5
1	H	50	ASN	2.5
1	B	73	ALA	2.4
1	F	102	LYS	2.4
1	E	51	VAL	2.4
1	B	74	THR	2.4
1	K	74	THR	2.4
1	I	89	THR	2.3
1	J	76	TYR	2.3
1	C	73	ALA	2.3
1	K	102	LYS	2.3
1	I	132	ARG	2.2
1	B	47	GLY	2.2
1	E	105	THR	2.2
1	B	45	TYR	2.2
1	I	15	GLY	2.2
1	F	72	ASP	2.2
1	J	74	THR	2.2
1	G	50	ASN	2.1
1	B	65	THR	2.1
1	D	21	GLN	2.1
1	B	25	ILE	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	69	ILE	2.1
1	I	102	LYS	2.1
1	L	104	VAL	2.1
1	B	56	GLY	2.1
1	B	71	LEU	2.1
1	H	57	ILE	2.0
1	G	73	ALA	2.0
1	C	102	LYS	2.0
1	L	73	ALA	2.0
1	L	57	ILE	2.0
1	C	50	ASN	2.0
1	J	13	ASN	2.0

## 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
4	PEG	G	303	7/7	0.72	0.21	23,24,29,31	0
3	EDO	A	204	4/4	0.80	0.20	21,27,28,29	0
2	GOL	A	201	6/6	0.80	0.22	29,30,31,35	0
2	GOL	C	304	6/6	0.83	0.17	29,38,41,42	0
3	EDO	A	203	4/4	0.83	0.20	29,30,30,33	0
2	GOL	G	302	6/6	0.86	0.14	21,24,27,30	0
3	EDO	K	304	4/4	0.91	0.10	15,20,25,27	0
2	GOL	H	302	6/6	0.91	0.10	18,19,20,22	0
3	EDO	K	305	4/4	0.92	0.12	20,21,22,22	0
2	GOL	C	303	6/6	0.93	0.11	17,18,18,18	0
2	GOL	C	302	6/6	0.93	0.11	15,20,20,24	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	K	302	6/6	0.93	0.13	17,17,18,22	0
2	GOL	E	302	6/6	0.93	0.10	16,17,18,19	0
2	GOL	D	301	6/6	0.94	0.10	18,20,21,24	0
2	GOL	C	301	6/6	0.94	0.07	16,18,19,20	0
2	GOL	I	301	6/6	0.95	0.08	19,23,23,24	0
2	GOL	H	301	6/6	0.96	0.08	11,13,14,14	0
2	GOL	B	301	6/6	0.96	0.08	12,13,14,15	0
2	GOL	G	304	6/6	0.96	0.16	16,19,20,21	0
2	GOL	I	302	6/6	0.96	0.09	17,23,25,25	0
2	GOL	J	301	6/6	0.96	0.07	15,17,17,18	0
2	GOL	K	301	6/6	0.96	0.08	15,17,17,17	0
2	GOL	L	301	6/6	0.97	0.07	12,13,14,15	0
2	GOL	G	301	6/6	0.97	0.07	16,18,18,18	0
2	GOL	D	303	6/6	0.97	0.07	13,13,14,14	0
2	GOL	E	301	6/6	0.97	0.08	10,11,12,12	0
2	GOL	D	302	6/6	0.97	0.06	14,17,18,19	0
2	GOL	F	301	6/6	0.97	0.06	14,16,17,17	0
2	GOL	G	305	6/6	0.98	0.07	12,14,15,15	0
2	GOL	A	202	6/6	0.98	0.06	10,11,11,12	0
2	GOL	K	303	6/6	0.98	0.06	12,14,15,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.