Full wwPDB X-ray Structure Validation Report ⓘ

Jan 4, 2024 – 11:59 pm GMT

PDB ID	:	5K31
Title	:	Crystal structure of Human fibrillar procollagen type I C-propeptide Homo-trimer
Authors	:	Sharma, U.; Hulmes, D.J.S.; Aghajari, N.
Deposited on	:	2016-05-19
Resolution	:	2.20 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 ⓘ 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:

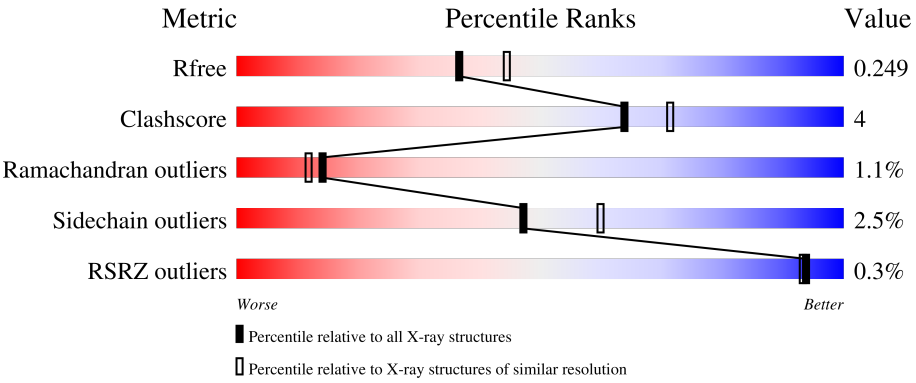
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	255	85% 9% • 5%
1	B	255	84% 8% • 6%
1	C	255	83% 9% • 7%
1	D	255	% 85% 8% 6%
1	E	255	84% 8% •• 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	255	

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
3	GOL	B	303	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11547 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 Collagen alpha-1(I) chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	0	0	0
			1866	1164	319	370	13			
1	B	239	Total	C	N	O	S	0	0	0
			1847	1155	313	366	13			
1	C	238	Total	C	N	O	S	0	0	0
			1844	1151	315	365	13			
1	D	239	Total	C	N	O	S	0	0	0
			1840	1151	311	365	13			
1	E	239	Total	C	N	O	S	0	0	0
			1834	1145	314	362	13			
1	F	240	Total	C	N	O	S	0	0	0
			1857	1159	316	369	13			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	GLU	-	expression tag	UNP P02452
A	-7	THR	-	expression tag	UNP P02452
A	-6	GLY	-	expression tag	UNP P02452
A	-5	HIS	-	expression tag	UNP P02452
A	-4	HIS	-	expression tag	UNP P02452
A	-3	HIS	-	expression tag	UNP P02452
A	-2	HIS	-	expression tag	UNP P02452
A	-1	HIS	-	expression tag	UNP P02452
A	0	HIS	-	expression tag	UNP P02452
A	147	GLN	ASN	engineered mutation	UNP P02452
A	216	SER	THR	variant	UNP P02452
B	-8	GLU	-	expression tag	UNP P02452
B	-7	THR	-	expression tag	UNP P02452
B	-6	GLY	-	expression tag	UNP P02452
B	-5	HIS	-	expression tag	UNP P02452
B	-4	HIS	-	expression tag	UNP P02452
B	-3	HIS	-	expression tag	UNP P02452

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	expression tag	UNP P02452
B	-1	HIS	-	expression tag	UNP P02452
B	0	HIS	-	expression tag	UNP P02452
B	147	GLN	ASN	engineered mutation	UNP P02452
B	216	SER	THR	variant	UNP P02452
C	-8	GLU	-	expression tag	UNP P02452
C	-7	THR	-	expression tag	UNP P02452
C	-6	GLY	-	expression tag	UNP P02452
C	-5	HIS	-	expression tag	UNP P02452
C	-4	HIS	-	expression tag	UNP P02452
C	-3	HIS	-	expression tag	UNP P02452
C	-2	HIS	-	expression tag	UNP P02452
C	-1	HIS	-	expression tag	UNP P02452
C	0	HIS	-	expression tag	UNP P02452
C	147	GLN	ASN	engineered mutation	UNP P02452
C	216	SER	THR	variant	UNP P02452
D	-8	GLU	-	expression tag	UNP P02452
D	-7	THR	-	expression tag	UNP P02452
D	-6	GLY	-	expression tag	UNP P02452
D	-5	HIS	-	expression tag	UNP P02452
D	-4	HIS	-	expression tag	UNP P02452
D	-3	HIS	-	expression tag	UNP P02452
D	-2	HIS	-	expression tag	UNP P02452
D	-1	HIS	-	expression tag	UNP P02452
D	0	HIS	-	expression tag	UNP P02452
D	147	GLN	ASN	engineered mutation	UNP P02452
D	216	SER	THR	variant	UNP P02452
E	-8	GLU	-	expression tag	UNP P02452
E	-7	THR	-	expression tag	UNP P02452
E	-6	GLY	-	expression tag	UNP P02452
E	-5	HIS	-	expression tag	UNP P02452
E	-4	HIS	-	expression tag	UNP P02452
E	-3	HIS	-	expression tag	UNP P02452
E	-2	HIS	-	expression tag	UNP P02452
E	-1	HIS	-	expression tag	UNP P02452
E	0	HIS	-	expression tag	UNP P02452
E	147	GLN	ASN	engineered mutation	UNP P02452
E	216	SER	THR	variant	UNP P02452
F	-8	GLU	-	expression tag	UNP P02452
F	-7	THR	-	expression tag	UNP P02452
F	-6	GLY	-	expression tag	UNP P02452
F	-5	HIS	-	expression tag	UNP P02452

Continued on next page...

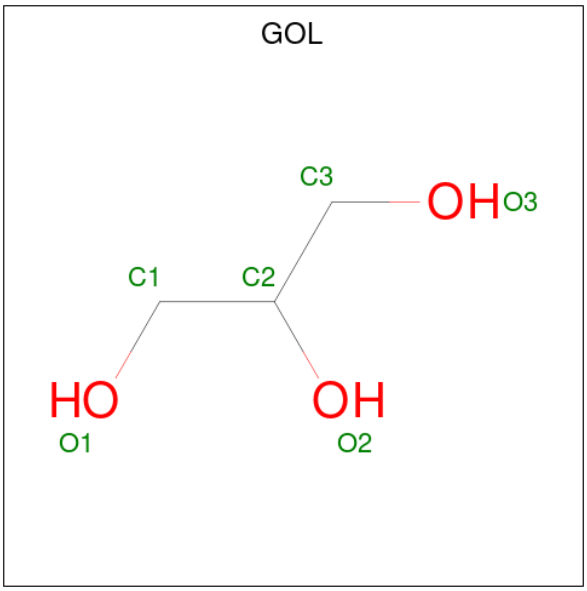
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-4	HIS	-	expression tag	UNP P02452
F	-3	HIS	-	expression tag	UNP P02452
F	-2	HIS	-	expression tag	UNP P02452
F	-1	HIS	-	expression tag	UNP P02452
F	0	HIS	-	expression tag	UNP P02452
F	147	GLN	ASN	engineered mutation	UNP P02452
F	216	SER	THR	variant	UNP P02452

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0
2	C	1	Total Ca 1 1	0	0
2	D	1	Total Ca 1 1	0	0
2	E	1	Total Ca 1 1	0	0
2	F	1	Total Ca 1 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	2	Total Cl 2 2	0	0
4	D	1	Total Cl 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	71	Total O 71 71	0	0
5	B	83	Total O 83 83	0	0
5	C	70	Total O 70 70	0	0
5	D	70	Total O 70 70	0	0
5	E	76	Total O 76 76	0	0
5	F	50	Total O 50 50	0	0

3 Residue-property plots (i)

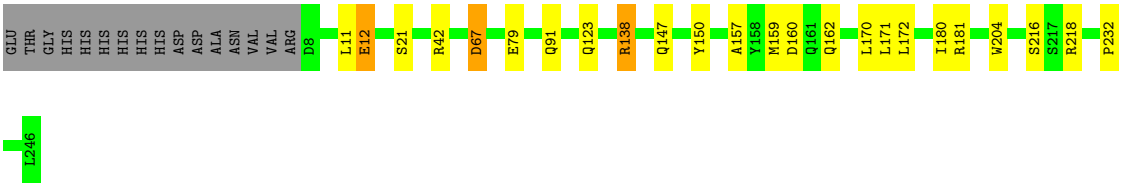
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Collagen alpha-1(I) chain

Chain A:

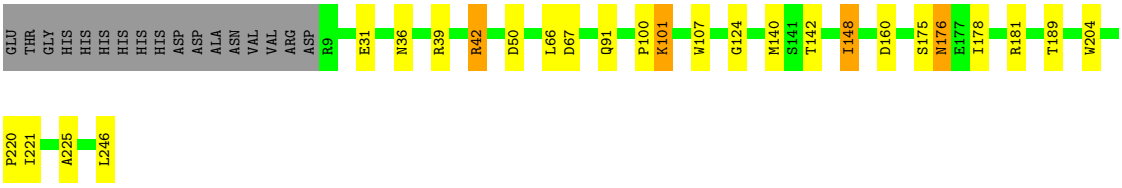
• Molecule 1: Collagen alpha-1(I) chain

Chain B:



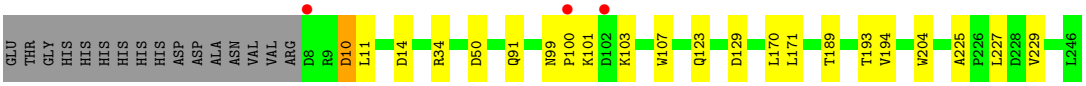
• Molecule 1: Collagen alpha-1(I) chain

Chain C:



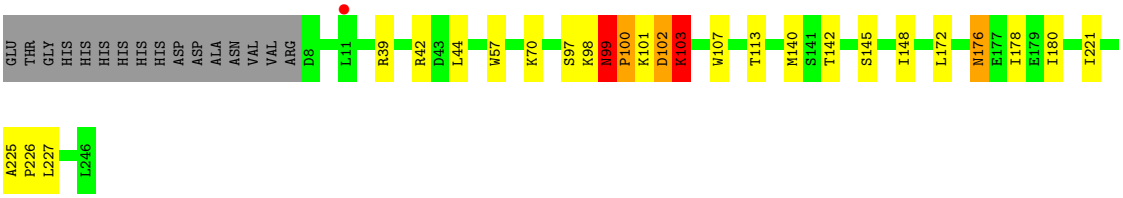
• Molecule 1: Collagen alpha-1(I) chain

Chain D:



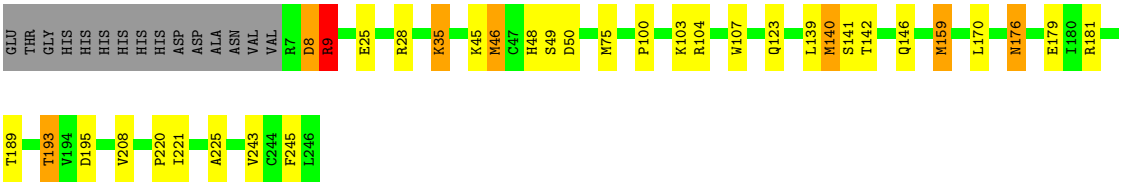
• Molecule 1: Collagen alpha-1(I) chain

Chain E:



• Molecule 1: Collagen alpha-1(I) chain

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.82Å 149.63Å 105.95Å 90.00° 101.68° 90.00°	Depositor
Resolution (Å)	47.07 – 2.20 47.07 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.07-2.20) 99.7 (47.07-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, R_{free}	0.196 , 0.238 0.207 , 0.249	Depositor DCC
R_{free} test set	5780 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.483	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 38.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11547	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.78% 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: CL, CA, 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	0.70	0/1907	0.87	3/2590 (0.1%)
1	B	0.75	0/1888	0.98	9/2563 (0.4%)
1	C	0.73	0/1885	0.90	7/2558 (0.3%)
1	D	0.69	0/1881	0.83	2/2555 (0.1%)
1	E	0.72	0/1875	0.89	3/2548 (0.1%)
1	F	0.68	0/1898	0.85	4/2577 (0.2%)
All	All	0.71	0/11334	0.89	28/15391 (0.2%)

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	B	0	2
1	C	0	1
1	D	0	1
1	E	0	4
1	F	0	2
All	All	0	10

There are no bond length outliers.

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	138	ARG	NE-CZ-NH2	-14.32	113.14	120.30
1	B	138	ARG	NE-CZ-NH1	12.47	126.53	120.30
1	B	42	ARG	NE-CZ-NH1	11.77	126.19	120.30
1	A	138	ARG	NE-CZ-NH2	-10.93	114.84	120.30
1	B	42	ARG	NE-CZ-NH2	-10.66	114.97	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	42	ARG	NE-CZ-NH1	10.33	125.47	120.30
1	C	42	ARG	NE-CZ-NH2	-10.10	115.25	120.30
1	A	138	ARG	NE-CZ-NH1	9.27	124.94	120.30
1	E	42	ARG	NE-CZ-NH1	7.69	124.14	120.30
1	E	42	ARG	NE-CZ-NH2	-7.17	116.71	120.30
1	A	159	MET	CG-SD-CE	6.60	110.76	100.20
1	C	176	ASN	N-CA-CB	-6.33	99.21	110.60
1	F	140	MET	CG-SD-CE	-6.13	90.39	100.20
1	C	148	ILE	CB-CA-C	-5.76	100.09	111.60
1	C	160	ASP	CB-CG-OD1	5.68	123.42	118.30
1	B	67	ASP	CB-CG-OD1	5.67	123.40	118.30
1	B	218	ARG	NE-CZ-NH2	-5.54	117.53	120.30
1	D	129	ASP	CB-CG-OD1	5.46	123.21	118.30
1	D	34	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	C	42	ARG	CD-NE-CZ	5.33	131.06	123.60
1	F	46	MET	CG-SD-CE	-5.31	91.71	100.20
1	B	138	ARG	CD-NE-CZ	5.29	131.01	123.60
1	E	176	ASN	N-CA-CB	-5.19	101.26	110.60
1	F	159	MET	CG-SD-CE	5.17	108.48	100.20
1	B	160	ASP	CB-CG-OD1	5.09	122.88	118.30
1	B	218	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	C	67	ASP	CB-CG-OD1	5.03	122.83	118.30
1	F	9	ARG	NE-CZ-NH1	5.03	122.81	120.30

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	11	LEU	Peptide
1	B	12	GLU	Peptide
1	C	124	GLY	Peptide
1	D	10	ASP	Peptide
1	E	103	LYS	Peptide
1	E	176	ASN	Peptide
1	E	98	LYS	Peptide
1	E	99	ASN	Peptide
1	F	49	SER	Peptide
1	F	9	ARG	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	1866	0	1757	14	0
1	B	1847	0	1743	19	0
1	C	1844	0	1741	14	0
1	D	1840	0	1730	7	0
1	E	1834	0	1716	17	0
1	F	1857	0	1747	25	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	6	0	8	0	0
3	B	18	0	24	6	0
3	E	6	0	8	1	0
4	B	2	0	0	0	0
4	D	1	0	0	0	0
5	A	71	0	0	0	0
5	B	83	0	0	0	0
5	C	70	0	0	0	0
5	D	70	0	0	0	0
5	E	76	0	0	0	0
5	F	50	0	0	3	0
All	All	11547	0	10474	90	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 4.

All (90) 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:E:102:ASP:CB	1:E:103:LYS:CB	2.27	1.12
1:D:193:THR:HG22	1:D:194:VAL:HG23	1.62	0.81
1:E:99:ASN:C	1:E:101:LYS:H	1.94	0.71
1:E:101:LYS:HA	1:E:102:ASP:CB	2.21	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:102:ASP:CB	1:E:103:LYS:CA	2.73	0.66
1:B:67:ASP:N	1:F:46:MET:HE1	2.12	0.65
1:E:103:LYS:HA	1:E:227:LEU:CD2	2.29	0.63
1:B:67:ASP:HB3	1:F:46:MET:HE1	1.80	0.62
1:E:103:LYS:HA	1:E:227:LEU:HD21	1.81	0.62
1:B:67:ASP:HB3	1:F:46:MET:CE	2.28	0.62
1:A:91:GLN:HB3	1:A:204:TRP:CH2	2.36	0.61
1:C:148:ILE:HD11	1:C:221:ILE:HD12	1.84	0.60
1:C:176:ASN:HB3	1:C:178:ILE:HG12	1.85	0.58
1:E:97:SER:HB3	1:E:99:ASN:HB3	1.85	0.58
1:A:97:SER:C	1:A:99:ASN:H	2.09	0.56
1:C:246:LEU:HD23	1:C:246:LEU:C	2.26	0.55
1:B:138:ARG:HD3	1:B:216:SER:O	2.08	0.54
1:B:67:ASP:CB	1:F:46:MET:HE1	2.37	0.54
1:C:39:ARG:NH1	1:C:140:MET:CE	2.72	0.53
1:A:94:TRP:HB3	1:A:112:MET:HE3	1.89	0.53
1:C:39:ARG:NH1	1:C:140:MET:HE3	2.23	0.53
1:E:39:ARG:NE	1:E:140:MET:SD	2.82	0.52
1:D:170:LEU:HD12	1:D:170:LEU:C	2.29	0.52
1:B:171:LEU:HD23	3:B:303:GOL:H32	1.91	0.52
1:F:176:ASN:N	5:F:401:HOH:O	2.09	0.52
1:A:107:TRP:CE3	1:A:225:ALA:HB2	2.44	0.52
1:D:107:TRP:CE3	1:D:225:ALA:HB2	2.46	0.51
1:D:91:GLN:HB3	1:D:204:TRP:CH2	2.45	0.51
1:A:94:TRP:CB	1:A:112:MET:CE	2.89	0.50
1:A:94:TRP:HB3	1:A:112:MET:CE	2.42	0.50
1:A:138:ARG:HD3	1:A:216:SER:O	2.11	0.50
1:F:8:ASP:O	1:F:9:ARG:CB	2.59	0.50
1:F:140:MET:CE	5:F:437:HOH:O	2.59	0.50
1:C:100:PRO:O	1:C:101:LYS:CB	2.59	0.50
1:F:107:TRP:CE3	1:F:225:ALA:HB2	2.47	0.50
1:E:99:ASN:C	1:E:101:LYS:N	2.63	0.49
1:F:45:LYS:HD2	1:F:75:MET:O	2.12	0.49
1:B:171:LEU:HD23	3:B:303:GOL:C3	2.41	0.49
1:F:9:ARG:HB3	1:F:9:ARG:CZ	2.41	0.49
1:B:79:GLU:OE2	3:B:305:GOL:C1	2.61	0.49
1:C:107:TRP:CE3	1:C:225:ALA:HB2	2.48	0.49
1:A:107:TRP:CD2	1:A:225:ALA:HB2	2.47	0.49
1:C:91:GLN:HB3	1:C:204:TRP:CH2	2.48	0.48
1:F:8:ASP:O	1:F:9:ARG:HB2	2.13	0.48
1:F:179:GLU:OE1	1:F:181:ARG:NH2	2.47	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:91:GLN:HB3	1:B:204:TRP:CH2	2.48	0.47
1:F:146:GLN:OE1	1:F:221:ILE:HD12	2.14	0.47
1:F:48:HIS:C	1:F:50:ASP:H	2.18	0.47
1:E:100:PRO:C	1:E:102:ASP:HA	2.35	0.47
1:A:113:THR:O	1:A:114:ASP:HB2	2.15	0.47
1:F:193:THR:HG22	1:F:208:VAL:HB	1.95	0.47
1:B:79:GLU:OE2	3:B:305:GOL:O1	2.22	0.47
1:D:103:LYS:HA	1:D:227:LEU:HD21	1.96	0.46
1:B:171:LEU:CD2	3:B:303:GOL:H32	2.45	0.46
1:A:96:ILE:HA	1:A:230:GLY:O	2.15	0.46
1:F:141:SER:OG	1:F:245:PHE:HB3	2.16	0.46
1:C:39:ARG:HD3	1:C:140:MET:HE1	1.98	0.46
1:D:99:ASN:HB3	1:D:101:LYS:H	1.81	0.45
1:E:57:TRP:CZ2	1:E:70:LYS:HD3	2.52	0.45
1:F:181:ARG:O	1:F:189:THR:HA	2.17	0.45
1:C:39:ARG:CZ	1:C:140:MET:HE3	2.47	0.45
1:A:126:ASP:C	1:A:126:ASP:OD1	2.56	0.44
1:C:42:ARG:HD2	5:F:427:HOH:O	2.18	0.44
1:C:181:ARG:O	1:C:189:THR:HA	2.17	0.44
1:F:35:LYS:N	1:F:35:LYS:HD3	2.33	0.44
1:A:94:TRP:HB2	1:A:112:MET:CE	2.47	0.44
1:F:107:TRP:CD2	1:F:225:ALA:HB2	2.52	0.44
1:D:50:ASP:OD1	1:D:50:ASP:N	2.49	0.44
1:E:148:ILE:HD11	1:E:221:ILE:HD12	1.99	0.44
1:B:67:ASP:CA	1:F:46:MET:HE1	2.48	0.43
1:B:181:ARG:CB	3:B:303:GOL:H32	2.48	0.43
1:C:31:GLU:HG2	1:C:36:ASN:OD1	2.19	0.43
1:B:12:GLU:OE1	1:B:12:GLU:O	2.36	0.43
1:B:150:TYR:CD2	1:B:157:ALA:HB2	2.55	0.42
1:B:172:LEU:HB2	1:B:180:ILE:HB	2.01	0.42
1:B:138:ARG:CD	1:B:216:SER:O	2.68	0.41
1:F:170:LEU:C	1:F:170:LEU:HD12	2.40	0.41
1:C:175:SER:OG	1:C:220:PRO:HD2	2.21	0.41
1:B:170:LEU:C	1:B:170:LEU:HD12	2.40	0.41
1:E:101:LYS:CA	1:E:102:ASP:CB	2.95	0.41
1:B:67:ASP:HB3	1:F:46:MET:HE3	2.02	0.41
1:E:172:LEU:HB2	1:E:180:ILE:HB	2.02	0.41
1:F:220:PRO:HB3	1:F:243:VAL:HB	2.02	0.41
1:A:170:LEU:C	1:A:170:LEU:HD12	2.42	0.41
1:E:107:TRP:CD2	1:E:225:ALA:HB2	2.56	0.40
1:F:159:MET:HB3	1:F:195:ASP:OD2	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:25:GLU:OE1	1:F:28:ARG:NH2	2.49	0.40
1:E:145:SER:N	3:E:302:GOL:O2	2.54	0.40
1:A:39:ARG:HD3	1:A:140:MET:SD	2.62	0.40
1:E:99:ASN:CG	1:E:100:PRO:CD	2.90	0.40

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	239/255 (94%)	231 (97%)	7 (3%)	1 (0%)	34	37
1	B	237/255 (93%)	230 (97%)	6 (2%)	1 (0%)	34	37
1	C	236/255 (92%)	230 (98%)	5 (2%)	1 (0%)	34	37
1	D	237/255 (93%)	222 (94%)	13 (6%)	2 (1%)	19	19
1	E	237/255 (93%)	224 (94%)	8 (3%)	5 (2%)	7	4
1	F	238/255 (93%)	223 (94%)	9 (4%)	6 (2%)	5	3
All	All	1424/1530 (93%)	1360 (96%)	48 (3%)	16 (1%)	14	12

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	123	GLN
1	E	99	ASN
1	E	100	PRO
1	F	8	ASP
1	F	9	ARG
1	F	100	PRO
1	F	103	LYS
1	F	123	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	98	LYS
1	B	123	GLN
1	C	101	LYS
1	F	104	ARG
1	E	102	ASP
1	E	103	LYS
1	E	178	ILE
1	D	100	PRO

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	204/223 (92%)	198 (97%)	6 (3%)	42	54
1	B	202/223 (91%)	197 (98%)	5 (2%)	47	60
1	C	202/223 (91%)	199 (98%)	3 (2%)	65	78
1	D	200/223 (90%)	194 (97%)	6 (3%)	41	53
1	E	198/223 (89%)	194 (98%)	4 (2%)	55	69
1	F	203/223 (91%)	197 (97%)	6 (3%)	41	53
All	All	1209/1338 (90%)	1179 (98%)	30 (2%)	47	60

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	ASN
1	A	39	ARG
1	A	47	CYS
1	A	139	LEU
1	A	142	THR
1	A	212	LYS
1	B	21	SER
1	B	147	GLN
1	B	159	MET
1	B	162	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	232	PRO
1	C	50	ASP
1	C	66	LEU
1	C	142	THR
1	D	10	ASP
1	D	11	LEU
1	D	14	ASP
1	D	171	LEU
1	D	189	THR
1	D	229	VAL
1	E	44	LEU
1	E	113	THR
1	E	142	THR
1	E	226	PRO
1	F	9	ARG
1	F	35	LYS
1	F	139	LEU
1	F	142	THR
1	F	176	ASN
1	F	193	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	48	HIS
1	F	62	GLN

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)

Of 14 ligands modelled in this entry, 9 are monoatomic - leaving 5 for Mogul analysis.

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	B	303	-	5,5,5	0.41	0	5,5,5	1.09	0
3	GOL	B	305	-	5,5,5	0.28	0	5,5,5	1.14	0
3	GOL	B	304	-	5,5,5	0.52	0	5,5,5	0.61	0
3	GOL	E	302	-	5,5,5	0.40	0	5,5,5	0.78	0
3	GOL	A	302	-	5,5,5	0.74	0	5,5,5	0.75	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	B	303	-	-	4/4/4/4	-
3	GOL	B	305	-	-	4/4/4/4	-
3	GOL	B	304	-	-	0/4/4/4	-
3	GOL	E	302	-	-	4/4/4/4	-
3	GOL	A	302	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	302	GOL	O1-C1-C2-C3
3	E	302	GOL	C1-C2-C3-O3
3	E	302	GOL	O2-C2-C3-O3
3	A	302	GOL	C1-C2-C3-O3
3	B	303	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	B	303	GOL	C1-C2-C3-O3
3	B	305	GOL	O1-C1-C2-C3
3	B	305	GOL	C1-C2-C3-O3
3	E	302	GOL	O1-C1-C2-C3
3	A	302	GOL	O1-C1-C2-O2
3	B	305	GOL	O1-C1-C2-O2
3	B	303	GOL	O1-C1-C2-O2
3	E	302	GOL	O1-C1-C2-O2
3	B	303	GOL	O2-C2-C3-O3
3	B	305	GOL	O2-C2-C3-O3
3	A	302	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	303	GOL	4	0
3	B	305	GOL	2	0
3	E	302	GOL	1	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^{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	241/255 (94%)	-0.22	1 (0%) 92 91	25, 39, 66, 89	0
1	B	239/255 (93%)	-0.24	0 100 100	23, 36, 60, 97	0
1	C	238/255 (93%)	-0.20	0 100 100	24, 39, 61, 84	0
1	D	239/255 (93%)	-0.16	3 (1%) 77 75	24, 40, 78, 101	0
1	E	239/255 (93%)	-0.24	1 (0%) 92 91	25, 39, 65, 93	0
1	F	240/255 (94%)	-0.18	0 100 100	28, 43, 69, 95	0
All	All	1436/1530 (93%)	-0.21	5 (0%) 94 93	23, 39, 67, 101	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	11	LEU	2.7
1	D	102	ASP	2.6
1	D	100	PRO	2.4
1	A	6	VAL	2.4
1	D	8	ASP	2.2

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

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

6.3 Carbohydrates

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^{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	302	6/6	0.76	0.23	46,51,54,59	0
3	GOL	B	304	6/6	0.85	0.34	35,46,55,63	0
4	CL	B	302	1/1	0.85	0.06	65,65,65,65	0
3	GOL	E	302	6/6	0.91	0.38	47,50,55,59	0
3	GOL	B	303	6/6	0.94	0.27	27,34,35,41	0
3	GOL	B	305	6/6	0.95	0.36	52,56,57,62	0
2	CA	B	301	1/1	0.97	0.06	19,19,19,19	0
2	CA	C	301	1/1	0.97	0.09	17,17,17,17	0
4	CL	B	306	1/1	0.97	0.09	31,31,31,31	0
2	CA	F	301	1/1	0.98	0.04	16,16,16,16	0
2	CA	A	301	1/1	0.98	0.05	16,16,16,16	0
4	CL	D	302	1/1	0.98	0.06	25,25,25,25	0
2	CA	E	301	1/1	0.99	0.08	18,18,18,18	0
2	CA	D	301	1/1	0.99	0.06	14,14,14,14	0

6.5 Other polymers (i)

There are no such residues in this entry.