Full wwPDB X-ray Structure Validation Report (i)

Sep 10, 2023 – 08:07 AM EDT

PDB ID	:	4IEF
Title	:	Complex of Porphyromonas gingivalis RgpB pro- and mature domains
Authors	:	de Diego, I.; Veillard, F.T.; Guevara, T.; Potempa, B.; Sztukowska, M.; Potempa, J.; Gomis-Ruth, F.X.
Deposited on	:	2012-12-13
Resolution	:	2.30 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:

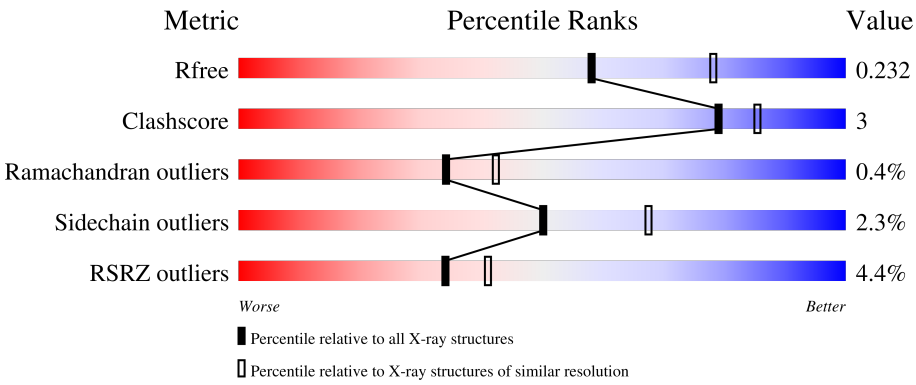
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	210	4%, 82%, 9%, 9%
1	C	210	7%, 85%, 8%, 7%
1	E	210	5%, 82%, 10%, 8%
1	G	210	9%, 84%, 6%, 10%
2	B	439	2%, 88%, 8%, •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	439	4% 88% 8%
2	F	439	5% 87% 9%
2	H	439	2% 88% 8%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 20046 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 Gingipain R2 Pro-Domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	192	1504	958	250	292	4	0	0	0
1	C	195	1527	972	256	295	4	0	0	0
1	E	194	1516	966	252	294	4	0	0	0
1	G	190	1492	951	248	289	4	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	GLY	-	expression tag	UNP P95493
A	21	PRO	-	expression tag	UNP P95493
A	22	LEU	-	expression tag	UNP P95493
A	23	GLY	-	expression tag	UNP P95493
A	24	SER	-	expression tag	UNP P95493
C	20	GLY	-	expression tag	UNP P95493
C	21	PRO	-	expression tag	UNP P95493
C	22	LEU	-	expression tag	UNP P95493
C	23	GLY	-	expression tag	UNP P95493
C	24	SER	-	expression tag	UNP P95493
E	20	GLY	-	expression tag	UNP P95493
E	21	PRO	-	expression tag	UNP P95493
E	22	LEU	-	expression tag	UNP P95493
E	23	GLY	-	expression tag	UNP P95493
E	24	SER	-	expression tag	UNP P95493
G	20	GLY	-	expression tag	UNP P95493
G	21	PRO	-	expression tag	UNP P95493
G	22	LEU	-	expression tag	UNP P95493
G	23	GLY	-	expression tag	UNP P95493
G	24	SER	-	expression tag	UNP P95493

- Molecule 2 is a protein called Gingipain R2 Mature Domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	423	3272	2058	548	649	17	0	0	0
2	D	424	3280	2062	550	651	17	0	0	0
2	F	424	3280	2062	550	651	17	0	0	0
2	H	424	3280	2062	550	651	17	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	663	HIS	-	expression tag	UNP P95493
B	664	HIS	-	expression tag	UNP P95493
B	665	HIS	-	expression tag	UNP P95493
B	666	HIS	-	expression tag	UNP P95493
B	667	HIS	-	expression tag	UNP P95493
B	668	HIS	-	expression tag	UNP P95493
D	663	HIS	-	expression tag	UNP P95493
D	664	HIS	-	expression tag	UNP P95493
D	665	HIS	-	expression tag	UNP P95493
D	666	HIS	-	expression tag	UNP P95493
D	667	HIS	-	expression tag	UNP P95493
D	668	HIS	-	expression tag	UNP P95493
F	663	HIS	-	expression tag	UNP P95493
F	664	HIS	-	expression tag	UNP P95493
F	665	HIS	-	expression tag	UNP P95493
F	666	HIS	-	expression tag	UNP P95493
F	667	HIS	-	expression tag	UNP P95493
F	668	HIS	-	expression tag	UNP P95493
H	663	HIS	-	expression tag	UNP P95493
H	664	HIS	-	expression tag	UNP P95493
H	665	HIS	-	expression tag	UNP P95493
H	666	HIS	-	expression tag	UNP P95493
H	667	HIS	-	expression tag	UNP P95493
H	668	HIS	-	expression tag	UNP P95493

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

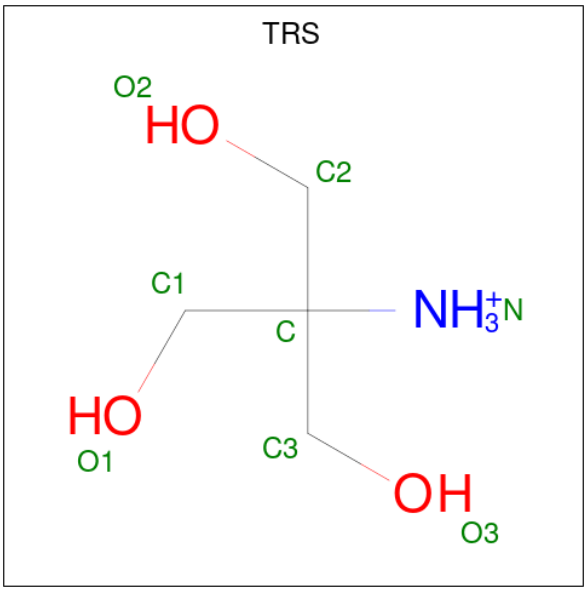
- Molecule 4 is BARIUM ION (three-letter code: BA) (formula: Ba).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ba 1 1	0	0
4	D	1	Total Ba 1 1	0	0
4	F	1	Total Ba 1 1	0	0
4	H	1	Total Ba 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	4	Total Ca 4 4	0	0
5	D	3	Total Ca 3 3	0	0
5	F	4	Total Ca 4 4	0	0
5	H	3	Total Ca 3 3	0	0

- Molecule 6 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	B	1	8	4	1	3	0	0

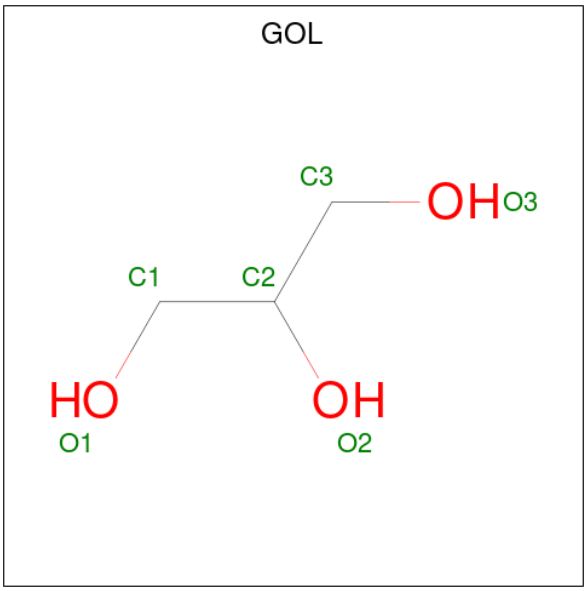
- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
7	F	3	3	3	0	0

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
8	H	1	1	1	0	0

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 10 is water.

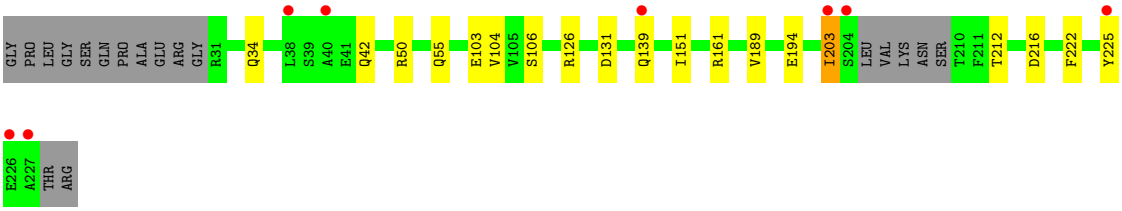
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
10	A	68	68	68	0	0
10	B	179	179	179	0	0
10	C	62	62	62	0	0
10	D	139	139	139	0	0
10	E	69	69	69	0	0
10	F	156	156	156	0	0
10	G	43	43	43	0	0
10	H	142	142	142	0	0

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.

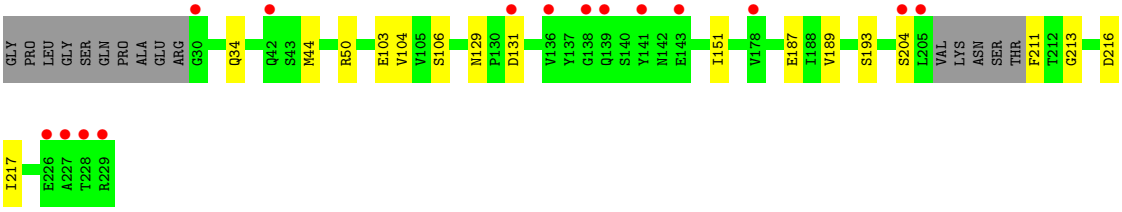
• Molecule 1: Gingipain R2 Pro-Domain

Chain A:



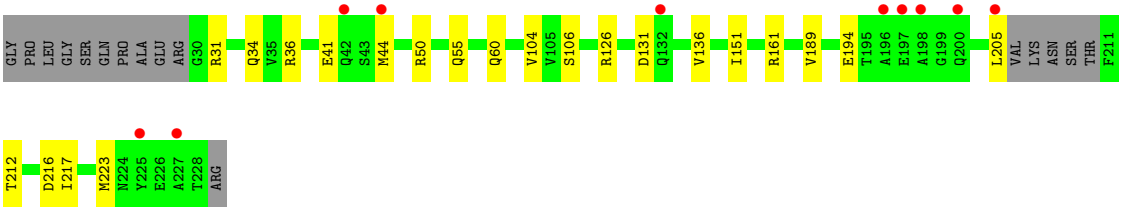
• Molecule 1: Gingipain R2 Pro-Domain

Chain C:



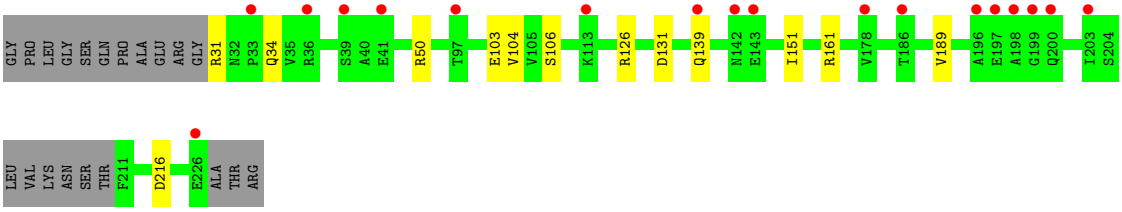
• Molecule 1: Gingipain R2 Pro-Domain

Chain E:



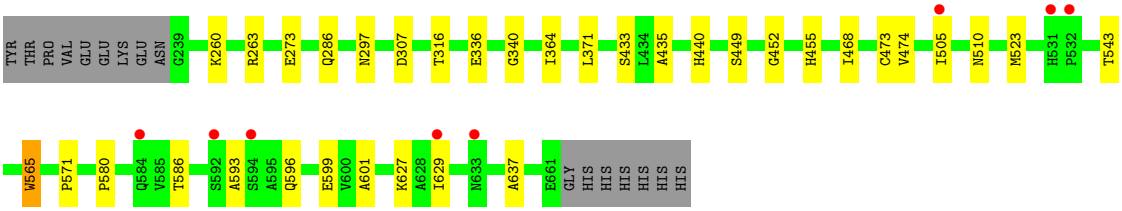
• Molecule 1: Gingipain R2 Pro-Domain

Chain G:



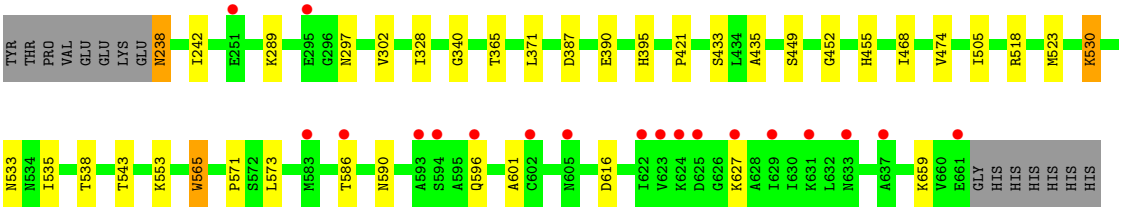
• Molecule 2: Gingipain R2 Mature Domain

Chain B:



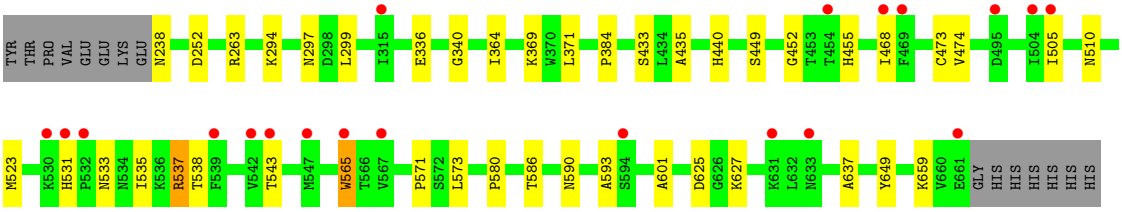
• Molecule 2: Gingipain R2 Mature Domain

Chain D:



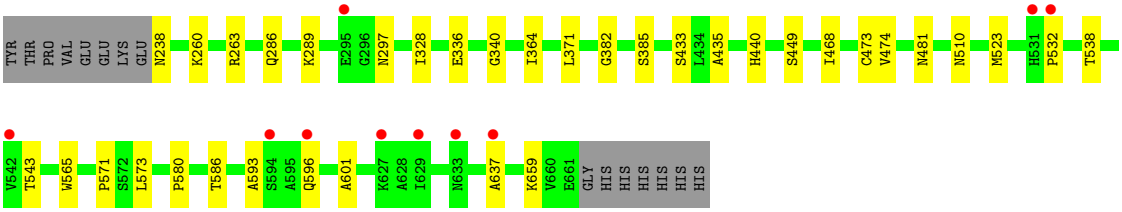
• Molecule 2: Gingipain R2 Mature Domain

Chain F:



• Molecule 2: Gingipain R2 Mature Domain

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	84.42Å 133.14Å 109.83Å 90.00° 90.52° 90.00°	Depositor
Resolution (Å)	33.67 – 2.30 33.61 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.2 (33.67-2.30) 99.3 (33.61-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.189 , 0.225 0.192 , 0.232	Depositor DCC
R_{free} test set	5323 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	32.5	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	20046	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: TRS, BA, CA, CL, CSD, NA, GOL, MG

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	0.40	0/1531	0.60	0/2075
1	C	0.40	0/1554	0.60	0/2105
1	E	0.39	0/1543	0.59	0/2091
1	G	0.39	0/1519	0.59	0/2058
2	B	0.41	0/3324	0.62	0/4508
2	D	0.40	0/3332	0.63	0/4519
2	F	0.41	0/3332	0.62	0/4519
2	H	0.40	0/3332	0.62	0/4519
All	All	0.40	0/19467	0.61	0/26394

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1504	0	1513	10	0
1	C	1527	0	1540	9	0
1	E	1516	0	1527	13	0
1	G	1492	0	1501	7	0
2	B	3272	0	3219	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	3280	0	3225	22	0
2	F	3280	0	3225	26	0
2	H	3280	0	3225	17	0
3	A	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	H	1	0	0	0	0
5	B	4	0	0	0	0
5	D	3	0	0	0	0
5	F	4	0	0	0	0
5	H	3	0	0	0	0
6	B	8	0	12	0	0
7	F	3	0	0	0	0
8	H	1	0	0	0	0
9	H	6	0	6	0	0
10	A	68	0	0	3	0
10	B	179	0	0	1	0
10	C	62	0	0	1	0
10	D	139	0	0	3	0
10	E	69	0	0	5	0
10	F	156	0	0	3	0
10	G	43	0	0	1	0
10	H	142	0	0	3	0
All	All	20046	0	18993	105	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 (105) 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:217:ILE:HD11	2:D:535:ILE:HG21	1.59	0.85
1:G:103:GLU:HG3	10:G:319:HOH:O	1.86	0.75
2:F:452:GLY:H	2:F:455:HIS:HD2	1.39	0.69
2:F:537:ARG:HD2	2:F:649:TYR:HB2	1.74	0.68
2:B:452:GLY:H	2:B:455:HIS:HD2	1.40	0.68
2:D:365:THR:HG23	10:D:852:HOH:O	1.93	0.68
2:D:452:GLY:H	2:D:455:HIS:HD2	1.41	0.68
1:G:31:ARG:HD3	10:H:813:HOH:O	1.93	0.66
2:B:627:LYS:HB2	2:F:627:LYS:HE3	1.77	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:E:351:HOH:O	2:F:384:PRO:HD2	1.98	0.63
2:F:369:LYS:NZ	10:F:860:HOH:O	2.29	0.63
2:B:599:GLU:HG3	2:F:625:ASP:O	2.00	0.61
1:E:223:MET:HA	10:E:345:HOH:O	2.00	0.60
1:C:217:ILE:HD11	2:D:535:ILE:CG2	2.30	0.60
1:A:126:ARG:O	2:B:440:HIS:ND1	2.34	0.59
1:C:103:GLU:HG3	10:C:307:HOH:O	2.00	0.59
1:E:212:THR:HG23	10:E:310:HOH:O	2.02	0.59
2:H:371:LEU:HA	2:H:433:SER:HB3	1.87	0.57
2:B:435:ALA:HB3	2:B:468:ILE:HG12	1.86	0.57
2:F:435:ALA:HB3	2:F:468:ILE:HG12	1.87	0.56
1:A:103:GLU:HG3	10:A:412:HOH:O	2.06	0.55
2:H:435:ALA:HB3	2:H:468:ILE:HG12	1.89	0.55
1:E:126:ARG:O	2:F:440:HIS:ND1	2.36	0.55
1:E:194:GLU:HG2	10:E:354:HOH:O	2.05	0.55
1:A:34:GLN:HB3	1:A:50:ARG:HB2	1.90	0.54
1:C:44:MET:HG3	1:C:193:SER:HB3	1.88	0.54
2:D:435:ALA:HB3	2:D:468:ILE:HG12	1.89	0.54
2:D:242:ILE:CG2	2:D:302:VAL:HG22	2.38	0.54
2:H:263:ARG:NH2	10:H:831:HOH:O	2.37	0.53
1:G:34:GLN:HB3	1:G:50:ARG:HB2	1.89	0.53
1:E:161:ARG:NH2	2:F:299:LEU:O	2.39	0.53
1:C:106:SER:HB3	1:C:189:VAL:HB	1.90	0.53
1:E:106:SER:HB3	1:E:189:VAL:HB	1.90	0.53
2:D:238:ASN:HB3	10:D:919:HOH:O	2.07	0.53
1:G:126:ARG:O	2:H:440:HIS:ND1	2.41	0.52
1:C:34:GLN:HB3	1:C:50:ARG:HB2	1.91	0.52
1:E:34:GLN:HB3	1:E:50:ARG:HB2	1.92	0.52
2:B:599:GLU:HB2	2:F:625:ASP:HB3	1.91	0.51
2:D:518:ARG:HD3	2:D:553:LYS:HD2	1.92	0.51
1:E:161:ARG:NH1	2:F:297:ASN:O	2.40	0.51
2:D:371:LEU:HA	2:D:433:SER:HB3	1.91	0.51
2:D:421:PRO:HD2	10:D:874:HOH:O	2.11	0.51
2:D:590:ASN:HD22	2:D:659:LYS:HB3	1.76	0.51
1:G:106:SER:HB3	1:G:189:VAL:HB	1.93	0.50
2:F:590:ASN:HD22	2:F:659:LYS:HB3	1.76	0.50
1:A:106:SER:HB3	1:A:189:VAL:HB	1.93	0.49
1:E:217:ILE:HD11	2:F:535:ILE:HG21	1.94	0.49
2:H:523:MET:HG2	2:H:543:THR:HA	1.95	0.48
2:B:523:MET:HG2	2:B:543:THR:HA	1.95	0.48
2:D:523:MET:HG2	2:D:543:THR:HA	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:586:THR:HB	2:B:601:ALA:HB3	1.97	0.47
2:B:629:ILE:HD12	2:F:625:ASP:O	2.15	0.47
2:D:387:ASP:OD1	2:D:518:ARG:HD2	2.14	0.47
2:D:242:ILE:HG22	2:D:302:VAL:HG22	1.96	0.47
2:F:263:ARG:NH1	10:F:806:HOH:O	2.36	0.47
2:F:523:MET:HG2	2:F:543:THR:HA	1.96	0.47
2:H:382:GLY:HA3	10:H:821:HOH:O	2.14	0.47
2:H:593:ALA:HB1	2:H:637:ALA:HA	1.96	0.47
2:B:505:ILE:HG12	2:B:565:TRP:CD1	2.50	0.47
2:D:586:THR:HB	2:D:601:ALA:HB3	1.97	0.47
1:A:194:GLU:HG2	10:A:461:HOH:O	2.14	0.46
2:H:586:THR:HB	2:H:601:ALA:HB3	1.98	0.46
2:F:593:ALA:HB1	2:F:637:ALA:HA	1.97	0.46
2:B:473:CSD:HB2	2:B:510:ASN:HD22	1.80	0.46
2:F:505:ILE:HG12	2:F:565:TRP:CD1	2.51	0.46
1:A:104:VAL:HG11	1:A:151:ILE:HD11	1.98	0.46
2:F:586:THR:HB	2:F:601:ALA:HB3	1.97	0.46
1:A:161:ARG:HA	2:B:336:GLU:HG3	1.98	0.45
2:B:340:GLY:HA3	2:B:571:PRO:HD3	1.99	0.45
2:D:340:GLY:HA3	2:D:571:PRO:HD3	1.99	0.45
2:B:593:ALA:HB1	2:B:637:ALA:HA	1.98	0.44
2:D:505:ILE:HG12	2:D:565:TRP:CD1	2.51	0.44
2:F:371:LEU:HA	2:F:433:SER:HB3	1.99	0.44
1:E:136:VAL:HG12	10:E:353:HOH:O	2.16	0.44
2:B:307:ASP:HB3	10:B:948:HOH:O	2.17	0.44
2:F:340:GLY:HA3	2:F:571:PRO:HD3	2.00	0.44
2:F:473:CSD:HB2	2:F:510:ASN:HD22	1.82	0.44
1:A:203:ILE:H	1:A:203:ILE:HG13	1.61	0.44
2:D:289:LYS:HG3	2:D:328:ILE:HG23	1.99	0.44
1:E:161:ARG:HA	2:F:336:GLU:HG3	1.99	0.44
1:C:104:VAL:HG11	1:C:151:ILE:HD11	1.99	0.43
2:H:289:LYS:HG3	2:H:328:ILE:HG23	2.00	0.43
2:H:340:GLY:HA3	2:H:571:PRO:HD3	2.00	0.43
2:B:371:LEU:HA	2:B:433:SER:HB3	2.00	0.43
1:C:213:GLY:O	2:D:535:ILE:HD12	2.18	0.43
2:D:538:THR:HA	2:D:573:LEU:O	2.18	0.43
1:C:216:ASP:CB	2:D:533:ASN:HB2	2.48	0.43
2:H:364:ILE:HG22	2:H:580:PRO:HG3	2.01	0.43
2:D:390:GLU:OE1	2:D:395:HIS:HA	2.19	0.43
2:H:473:CSD:HB2	2:H:510:ASN:HD22	1.83	0.43
1:E:104:VAL:HG11	1:E:151:ILE:HD11	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:260:LYS:NZ	2:H:263:ARG:HH11	2.17	0.43
2:H:538:THR:HA	2:H:573:LEU:O	2.20	0.42
2:F:538:THR:HA	2:F:573:LEU:O	2.20	0.42
2:H:371:LEU:HD23	2:H:433:SER:CB	2.50	0.42
1:E:41:GLU:HB2	1:E:44:MET:HB3	2.01	0.42
1:A:212:THR:HG22	10:A:467:HOH:O	2.20	0.41
2:B:260:LYS:NZ	2:B:263:ARG:HH21	2.18	0.41
2:B:364:ILE:HG22	2:B:580:PRO:HG3	2.02	0.41
2:F:364:ILE:HG22	2:F:580:PRO:HG3	2.03	0.41
1:A:222:PHE:HB2	1:A:225:TYR:HB2	2.03	0.41
2:H:371:LEU:HD23	2:H:433:SER:HB2	2.03	0.41
1:G:104:VAL:HG11	1:G:151:ILE:HD11	2.03	0.40
1:G:161:ARG:HA	2:H:336:GLU:HG3	2.03	0.40
2:F:537:ARG:HG2	10:F:951:HOH:O	2.22	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	188/210 (90%)	182 (97%)	6 (3%)	0	100	100
1	C	191/210 (91%)	186 (97%)	5 (3%)	0	100	100
1	E	190/210 (90%)	183 (96%)	7 (4%)	0	100	100
1	G	186/210 (89%)	180 (97%)	6 (3%)	0	100	100
2	B	420/439 (96%)	413 (98%)	5 (1%)	2 (0%)	29	35
2	D	421/439 (96%)	412 (98%)	6 (1%)	3 (1%)	22	26
2	F	421/439 (96%)	410 (97%)	9 (2%)	2 (0%)	29	35
2	H	421/439 (96%)	413 (98%)	5 (1%)	3 (1%)	22	26
All	All	2438/2596 (94%)	2379 (98%)	49 (2%)	10 (0%)	34	42

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	449	SER
2	B	474	VAL
2	D	449	SER
2	F	449	SER
2	F	474	VAL
2	H	449	SER
2	H	474	VAL
2	D	474	VAL
2	D	530	LYS
2	H	532	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	170/184 (92%)	164 (96%)	6 (4%)	36	50
1	C	172/184 (94%)	167 (97%)	5 (3%)	42	58
1	E	171/184 (93%)	164 (96%)	7 (4%)	30	43
1	G	169/184 (92%)	166 (98%)	3 (2%)	59	75
2	B	357/372 (96%)	351 (98%)	6 (2%)	60	76
2	D	358/372 (96%)	351 (98%)	7 (2%)	55	72
2	F	358/372 (96%)	351 (98%)	7 (2%)	55	72
2	H	358/372 (96%)	350 (98%)	8 (2%)	52	69
All	All	2113/2224 (95%)	2064 (98%)	49 (2%)	50	67

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	GLN
1	A	55	GLN
1	A	131	ASP
1	A	139	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	203	ILE
1	A	216	ASP
2	B	273	GLU
2	B	286	GLN
2	B	297	ASN
2	B	316	THR
2	B	565	TRP
2	B	596	GLN
1	C	129	ASN
1	C	131	ASP
1	C	187	GLU
1	C	204	SER
1	C	211	PHE
2	D	238	ASN
2	D	297	ASN
2	D	530	LYS
2	D	565	TRP
2	D	596	GLN
2	D	616	ASP
2	D	627	LYS
1	E	31	ARG
1	E	36	ARG
1	E	55	GLN
1	E	60	GLN
1	E	131	ASP
1	E	205	LEU
1	E	216	ASP
2	F	238	ASN
2	F	252	ASP
2	F	294	LYS
2	F	531	HIS
2	F	533	ASN
2	F	537	ARG
2	F	565	TRP
1	G	131	ASP
1	G	139	GLN
1	G	216	ASP
2	H	238	ASN
2	H	286	GLN
2	H	297	ASN
2	H	385	SER
2	H	481	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	565	TRP
2	H	596	GLN
2	H	659	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (22) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	455	HIS
2	B	475	ASN
2	B	510	ASN
1	C	76	ASN
1	C	129	ASN
1	C	132	GLN
2	D	455	HIS
2	D	475	ASN
2	D	510	ASN
2	D	590	ASN
2	F	282	ASN
2	F	455	HIS
2	F	464	GLN
2	F	475	ASN
2	F	510	ASN
2	F	590	ASN
1	G	55	GLN
1	G	60	GLN
1	G	139	GLN
1	G	142	ASN
2	H	475	ASN
2	H	510	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)

4 non-standard protein/DNA/RNA residues 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	CSD	B	473	2	3,7,8	0.76	0	1,8,10	0.72	0
2	CSD	F	473	2	3,7,8	0.84	0	1,8,10	0.18	0
2	CSD	H	473	2	3,7,8	0.82	0	1,8,10	0.43	0
2	CSD	D	473	2	3,7,8	0.81	0	1,8,10	0.20	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	CSD	B	473	2	-	0/2/6/8	-
2	CSD	F	473	2	-	0/2/6/8	-
2	CSD	H	473	2	-	0/2/6/8	-
2	CSD	D	473	2	-	0/2/6/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	473	CSD	1	0
2	F	473	CSD	1	0
2	H	473	CSD	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 25 ligands modelled in this entry, 23 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
9	GOL	H	706	5	5,5,5	0.38	0	5,5,5	0.58	0
6	TRS	B	706	-	7,7,7	0.26	0	9,9,9	1.14	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
9	GOL	H	706	5	-	2/4/4/4	-
6	TRS	B	706	-	-	6/9/9/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	706	TRS	C1-C-C2-O2
6	B	706	TRS	C3-C-C2-O2
6	B	706	TRS	N-C-C2-O2
9	H	706	GOL	C1-C2-C3-O3
6	B	706	TRS	C2-C-C1-O1
9	H	706	GOL	O2-C2-C3-O3
6	B	706	TRS	C3-C-C1-O1
6	B	706	TRS	N-C-C1-O1

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 (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, 95th 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	192/210 (91%)	0.09	8 (4%)	36	43	26, 38, 59, 97	0
1	C	195/210 (92%)	0.21	15 (7%)	13	17	23, 42, 68, 103	0
1	E	194/210 (92%)	0.27	10 (5%)	27	34	28, 43, 70, 101	0
1	G	190/210 (90%)	0.43	18 (9%)	8	11	31, 45, 67, 109	0
2	B	422/439 (96%)	-0.02	8 (1%)	66	73	19, 32, 55, 80	0
2	D	423/439 (96%)	0.16	19 (4%)	33	40	19, 34, 64, 87	0
2	F	423/439 (96%)	0.16	20 (4%)	31	38	21, 35, 56, 92	0
2	H	423/439 (96%)	-0.01	10 (2%)	59	66	20, 34, 57, 76	0
All	All	2462/2596 (94%)	0.13	108 (4%)	34	41	19, 37, 63, 109	0

All (108) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	196	ALA	7.0
1	A	227	ALA	5.8
2	D	594	SER	5.5
1	G	197	GLU	5.3
2	F	532	PRO	5.2
1	E	227	ALA	5.1
2	D	627	LYS	5.0
2	B	531	HIS	4.8
1	A	40	ALA	4.7
1	C	228	THR	4.7
1	C	42	GLN	4.5
1	E	197	GLU	4.3
1	C	229	ARG	4.2
2	D	624	LYS	4.1
1	A	225	TYR	4.1
1	G	142	ASN	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	G	178	VAL	3.8
1	G	226	GLU	3.8
2	D	602	CYS	3.8
1	C	178	VAL	3.7
2	F	530	LYS	3.6
1	C	205	LEU	3.6
2	F	633	ASN	3.6
2	H	531	HIS	3.6
2	H	627	LYS	3.5
2	D	629	ILE	3.5
2	F	505	ILE	3.4
2	F	594	SER	3.3
1	C	138	GLY	3.1
2	D	593	ALA	3.1
2	D	622	ILE	3.1
2	H	596	GLN	3.1
1	E	132	GLN	3.1
1	C	30	GLY	3.1
1	G	139	GLN	3.0
1	G	203	ILE	3.0
2	D	633	ASN	3.0
1	C	204	SER	3.0
1	G	198	ALA	3.0
2	D	295	GLU	2.9
2	D	605	ASN	2.9
1	C	227	ALA	2.9
2	H	594	SER	2.8
1	A	203	ILE	2.8
2	B	532	PRO	2.8
1	G	36	ARG	2.8
1	E	44	MET	2.8
2	H	629	ILE	2.8
1	C	136	VAL	2.7
2	F	468	ILE	2.7
2	F	531	HIS	2.7
2	H	637	ALA	2.7
2	D	623	VAL	2.7
2	B	629	ILE	2.6
2	D	586	THR	2.6
2	H	295	GLU	2.6
1	C	141	TYR	2.6
2	F	469	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	583	MET	2.6
2	F	543	THR	2.6
1	G	33	PRO	2.6
2	D	631	LYS	2.6
1	G	199	GLY	2.6
2	B	505	ILE	2.6
2	F	631	LYS	2.6
1	A	226	GLU	2.5
2	D	637	ALA	2.5
1	E	200	GLN	2.5
1	G	97	THR	2.5
2	D	625	ASP	2.5
1	C	143	GLU	2.4
2	F	567	VAL	2.4
2	F	504	ILE	2.4
1	E	198	ALA	2.4
2	B	592	SER	2.4
2	B	584	GLN	2.4
2	F	661	GLU	2.3
2	F	454	THR	2.3
1	A	204	SER	2.3
2	D	596	GLN	2.2
2	D	661	GLU	2.2
1	G	113	LYS	2.2
1	E	225	TYR	2.2
2	F	495	ASP	2.2
1	G	196	ALA	2.2
1	C	131	ASP	2.2
1	G	143	GLU	2.1
1	C	226	GLU	2.1
1	G	200	GLN	2.1
1	A	38	LEU	2.1
2	H	532	PRO	2.1
2	B	594	SER	2.1
1	C	139	GLN	2.1
1	G	41	GLU	2.1
1	G	39	SER	2.1
2	D	251	GLU	2.0
2	F	547	MET	2.0
2	F	542	VAL	2.0
2	F	539	PHE	2.0
2	F	565	TRP	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	542	VAL	2.0
2	B	633	ASN	2.0
2	H	633	ASN	2.0
1	E	42	GLN	2.0
1	E	205	LEU	2.0
2	F	315	ILE	2.0
1	A	139	GLN	2.0
1	G	186	THR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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, 95th 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	CSD	D	473	8/9	0.93	0.11	28,30,39,41	0
2	CSD	B	473	8/9	0.95	0.09	28,29,36,36	0
2	CSD	H	473	8/9	0.95	0.11	29,31,36,38	0
2	CSD	F	473	8/9	0.98	0.08	30,32,36,38	0

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, 95th 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
6	TRS	B	706	8/8	0.64	0.26	54,55,56,56	0
5	CA	B	704	1/1	0.78	0.13	51,51,51,51	0
9	GOL	H	706	6/6	0.91	0.16	39,42,46,49	0
8	MG	H	705	1/1	0.93	0.13	50,50,50,50	0
7	NA	F	708	1/1	0.94	0.34	42,42,42,42	0
7	NA	F	706	1/1	0.94	0.22	38,38,38,38	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	NA	F	707	1/1	0.94	0.11	63,63,63,63	0
5	CA	D	704	1/1	0.96	0.12	42,42,42,42	0
5	CA	B	705	1/1	0.97	0.24	44,44,44,44	0
5	CA	F	705	1/1	0.97	0.25	55,55,55,55	0
5	CA	F	704	1/1	0.98	0.18	54,54,54,54	0
5	CA	H	703	1/1	0.98	0.09	31,31,31,31	0
5	CA	B	702	1/1	0.99	0.11	27,27,27,27	0
5	CA	B	703	1/1	0.99	0.10	25,25,25,25	0
5	CA	H	702	1/1	0.99	0.03	35,35,35,35	0
3	CL	A	301	1/1	0.99	0.04	34,34,34,34	0
5	CA	H	704	1/1	0.99	0.15	44,44,44,44	0
4	BA	F	701	1/1	0.99	0.06	59,59,59,59	0
5	CA	D	702	1/1	0.99	0.07	28,28,28,28	0
5	CA	D	703	1/1	0.99	0.12	23,23,23,23	0
4	BA	H	701	1/1	0.99	0.05	67,67,67,67	0
5	CA	F	702	1/1	0.99	0.09	32,32,32,32	0
5	CA	F	703	1/1	0.99	0.05	36,36,36,36	0
4	BA	B	701	1/1	1.00	0.06	55,55,55,55	0
4	BA	D	701	1/1	1.00	0.05	63,63,63,63	0

6.5 Other polymers (i)

There are no such residues in this entry.