Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 10:35 PM GMT

PDB ID	:	5D3G
Title	:	Structure of HIV-1 Reverse Transcriptase Bound to a Novel 38-mer Hairpin Template-Primer DNA Aptamer
Authors	:	Miller, M.T.; Tuske, S.; Das, K.; Arnold, E.
Deposited on	:	2015-08-06
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
http://wwpdb.org/validation/2016/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:

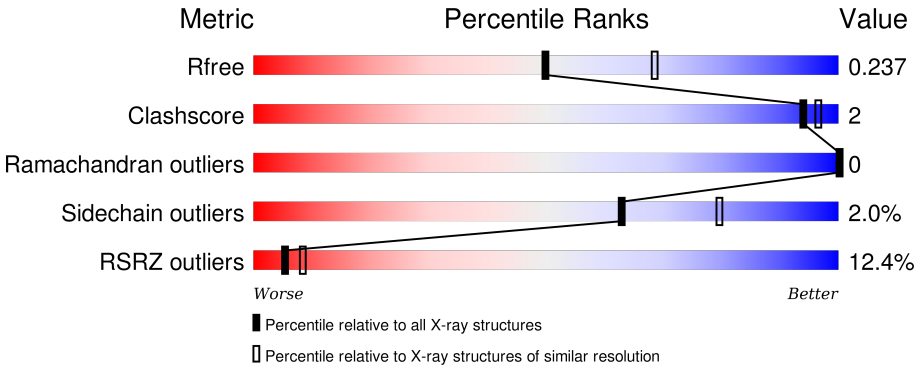
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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.



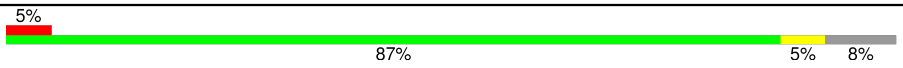
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	555	14% 91% 8% •
1	C	555	12% 93% 6% •
2	B	444	8% 89% 5% 7%
2	D	444	16% 86% 5% 8%
3	E	38	87% 5% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	38	

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
5	GOL	A	602	-	-	-	X
5	GOL	B	2006	-	-	-	X
5	GOL	D	502	-	-	-	X
5	GOL	D	504	-	-	-	X
5	GOL	F	101	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 17868 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 HIV-1 REVERSE TRANSCRIPTASE P66 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	0	0	0
			4504	2916	749	832	7			
1	C	553	Total	C	N	O	S	0	0	0
			4504	2916	749	832	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	SER	CYS	engineered mutation	UNP P03366
C	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE P51 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	415	Total	C	N	O	S	0	1	0
			3440	2243	568	622	7			
2	D	410	Total	C	N	O	S	0	2	0
			3409	2225	562	615	7			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	MET	-	initiating methionine	UNP P03366
B	-14	ALA	-	expression tag	UNP P03366
B	-13	HIS	-	expression tag	UNP P03366
B	-12	HIS	-	expression tag	UNP P03366
B	-11	HIS	-	expression tag	UNP P03366
B	-10	HIS	-	expression tag	UNP P03366
B	-9	HIS	-	expression tag	UNP P03366
B	-8	HIS	-	expression tag	UNP P03366
B	-7	ALA	-	expression tag	UNP P03366
B	-6	LEU	-	expression tag	UNP P03366

Continued on next page...

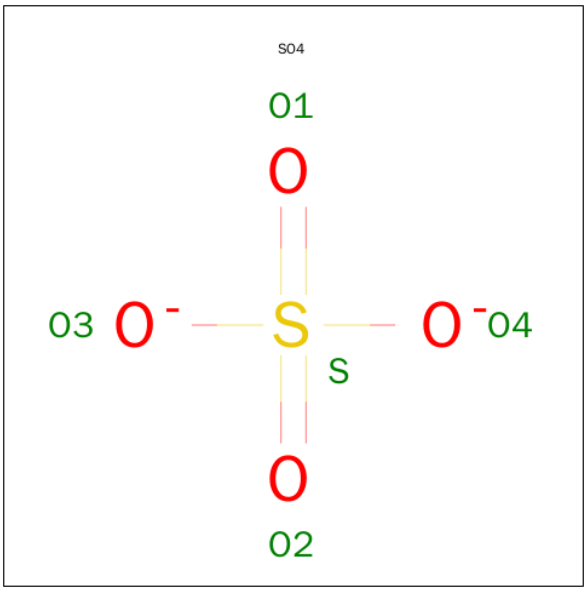
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	GLU	-	expression tag	UNP P03366
B	-4	VAL	-	expression tag	UNP P03366
B	-3	LEU	-	expression tag	UNP P03366
B	-2	PHE	-	expression tag	UNP P03366
B	-1	GLN	-	expression tag	UNP P03366
B	0	GLY	-	expression tag	UNP P03366
B	280	SER	CYS	engineered mutation	UNP P03366
D	-15	MET	-	initiating methionine	UNP P03366
D	-14	ALA	-	expression tag	UNP P03366
D	-13	HIS	-	expression tag	UNP P03366
D	-12	HIS	-	expression tag	UNP P03366
D	-11	HIS	-	expression tag	UNP P03366
D	-10	HIS	-	expression tag	UNP P03366
D	-9	HIS	-	expression tag	UNP P03366
D	-8	HIS	-	expression tag	UNP P03366
D	-7	ALA	-	expression tag	UNP P03366
D	-6	LEU	-	expression tag	UNP P03366
D	-5	GLU	-	expression tag	UNP P03366
D	-4	VAL	-	expression tag	UNP P03366
D	-3	LEU	-	expression tag	UNP P03366
D	-2	PHE	-	expression tag	UNP P03366
D	-1	GLN	-	expression tag	UNP P03366
D	0	GLY	-	expression tag	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is a DNA chain called DNA aptamer (38-MER).

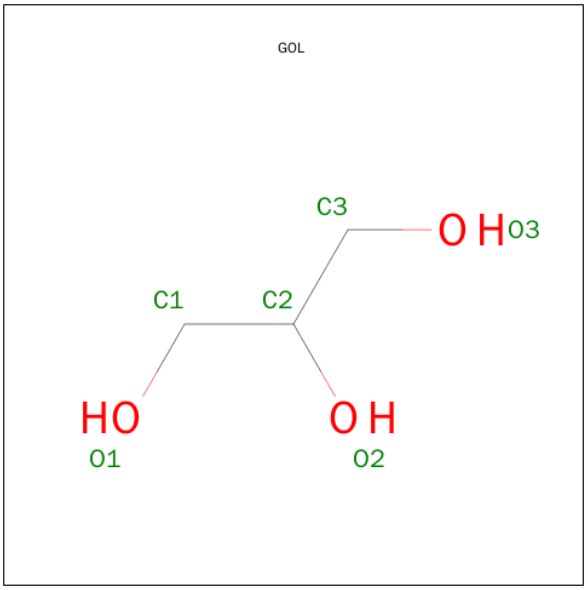
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	35	Total	C	N	O	P	0	0	0
			721	340	130	216	35			
3	E	35	Total	C	N	O	P	0	0	0
			721	340	130	216	35			

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



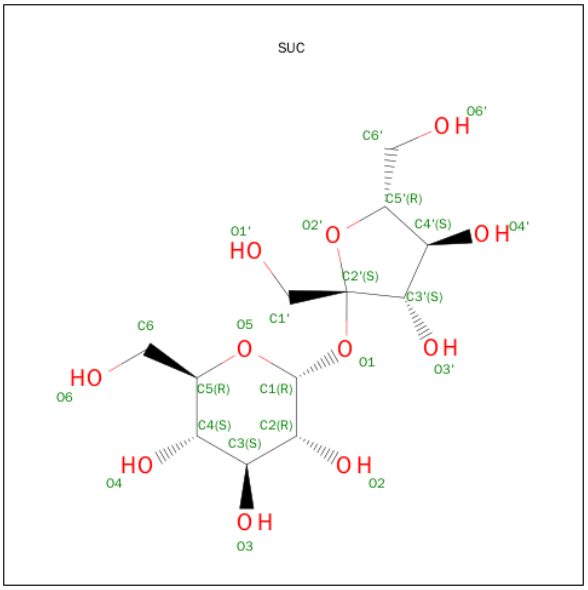
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	F	1	Total C O 6 3 3	0	0

- Molecule 6 is SUCROSE (three-letter code: SUC) (formula: $C_{12}H_{22}O_{11}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C O 23 12 11	0	0
6	D	1	Total C O 23 12 11	0	0

- Molecule 7 is water.

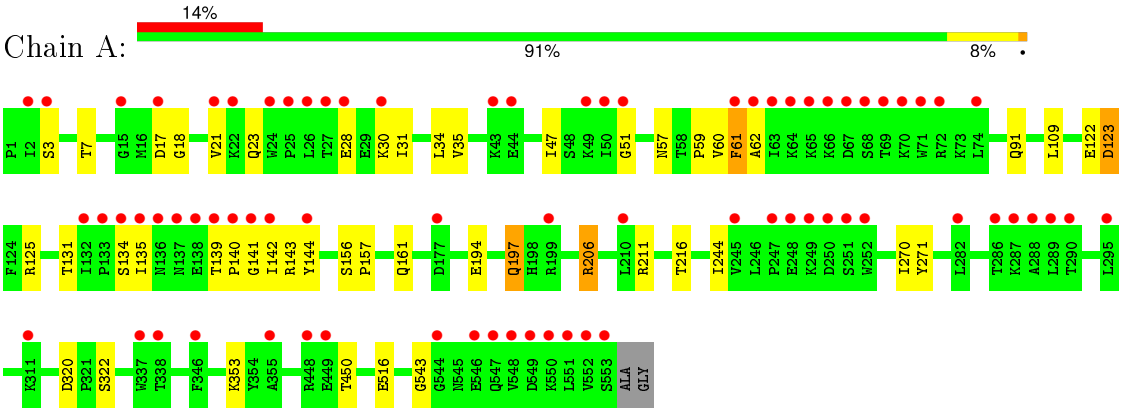
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	100	Total O 100 100	0	0
7	B	124	Total O 124 124	0	0
7	C	114	Total O 114 114	0	0
7	D	85	Total O 85 85	0	0
7	F	11	Total O 11 11	0	0
7	E	13	Total O 13 13	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

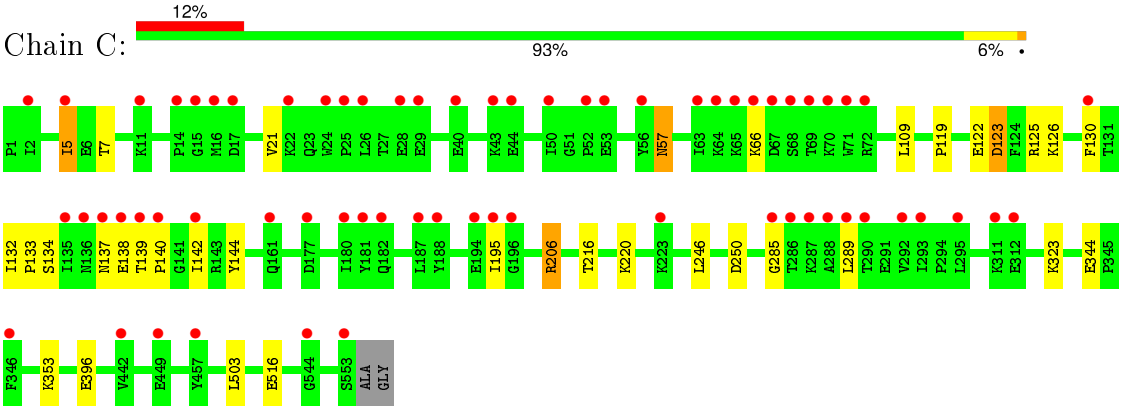
- Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 subunit

Chain A:



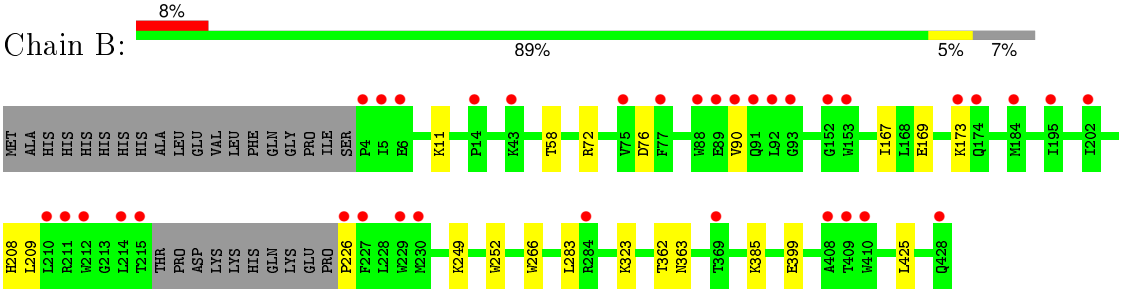
- Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 subunit

Chain C:



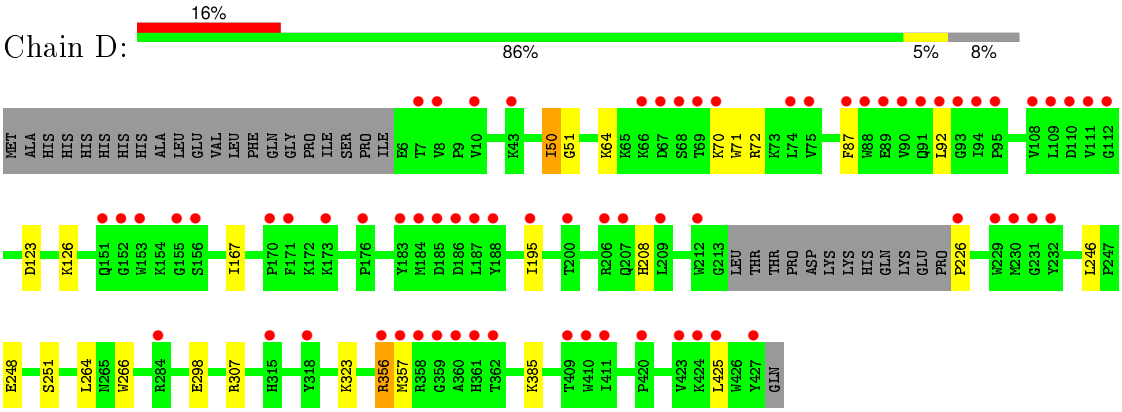
- Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 subunit

Chain B:



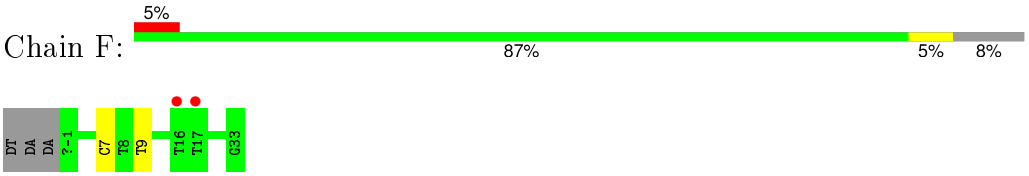
- Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 subunit

Chain D:



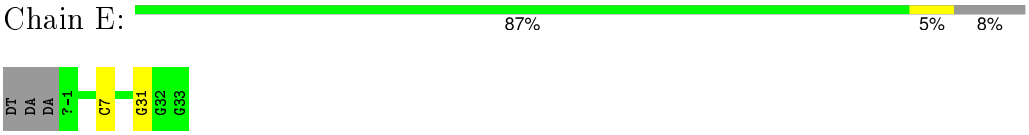
- Molecule 3: DNA aptamer (38-MER)

Chain F:



- Molecule 3: DNA aptamer (38-MER)

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	90.02Å 128.91Å 132.62Å 90.00° 101.34° 90.00°	Depositor
Resolution (Å)	29.78 – 2.30 29.78 – 2.27	Depositor EDS
% Data completeness (in resolution range)	99.3 (29.78-2.30) 96.3 (29.78-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.26Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.191 , 0.225 0.207 , 0.237	Depositor DCC
R_{free} test set	1958 reflections (1.55%)	DCC
Wilson B-factor (Å^2)	43.5	Xtriage
Anisotropy	0.185	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 58.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 133424 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	17868	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: TMP, GOL, OMC, SUC, SO4

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	0.25	0/4622	0.46	0/6279
1	C	0.25	0/4622	0.45	0/6279
2	B	0.24	0/3543	0.47	0/4812
2	D	0.24	0/3515	0.48	0/4775
3	E	0.52	0/737	0.94	1/1136 (0.1%)
3	F	0.53	0/737	0.93	0/1136
All	All	0.28	0/17776	0.52	1/24417 (0.0%)

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
2	D	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	31	DG	O4'-C4'-C3'	-5.12	102.45	104.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	356	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	4504	0	4565	25	0
1	C	4504	0	4565	14	0
2	B	3440	0	3474	11	0
2	D	3409	0	3438	14	0
3	E	721	0	394	1	0
3	F	721	0	395	2	0
4	A	5	0	0	0	0
4	C	5	0	0	0	0
5	A	6	0	8	0	0
5	B	30	0	40	1	0
5	C	6	0	8	0	0
5	D	18	0	24	0	0
5	F	6	0	8	0	0
6	B	23	0	22	0	0
6	D	23	0	22	0	0
7	A	100	0	0	0	0
7	B	124	0	0	0	0
7	C	114	0	0	0	0
7	D	85	0	0	0	0
7	E	13	0	0	0	0
7	F	11	0	0	0	0
All	All	17868	0	16963	63	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 2.

All (63) 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:21:VAL:O	1:C:57:ASN:ND2	2.17	0.77
1:A:3:SER:OG	1:A:211:ARG:NH1	2.28	0.67
2:D:356:ARG:HA	2:D:357:MET:HB2	1.76	0.67
1:C:134:SER:OG	1:C:139:THR:O	2.07	0.67
2:B:167:ILE:O	2:B:208:HIS:NE2	2.35	0.59
2:D:72:ARG:HG3	2:D:226:PRO:HB3	1.85	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:30:LYS:O	1:A:34:LEU:HG	2.05	0.57
2:D:356:ARG:N	2:D:357:MET:HG3	2.20	0.57
1:A:34:LEU:HD21	1:A:62:ALA:HB2	1.87	0.56
1:A:91:GLN:NE2	1:A:161:GLN:OE1	2.35	0.56
1:A:21:VAL:HB	1:A:59:PRO:HD3	1.87	0.56
2:B:399:GLU:O	5:B:2003:GOL:O3	2.25	0.55
1:C:353:LYS:NZ	3:F:7:DC:OP1	2.41	0.53
1:A:543:GLY:N	2:B:283:LEU:O	2.41	0.53
2:D:167:ILE:O	2:D:208:HIS:NE2	2.39	0.53
1:A:51:GLY:O	1:A:143:ARG:NH1	2.42	0.52
2:D:87:PHE:O	2:D:92:LEU:N	2.42	0.52
1:A:123:ASP:N	1:A:123:ASP:OD1	2.43	0.52
1:A:320:ASP:OD1	1:A:322:SER:OG	2.20	0.52
2:D:50:ILE:HG22	2:D:51:GLY:H	1.75	0.52
2:B:72:ARG:HG2	2:B:226:PRO:HB3	1.92	0.52
1:A:28:GLU:HA	1:A:31:ILE:HD12	1.92	0.51
1:A:353:LYS:NZ	3:E:7:DC:OP1	2.43	0.51
2:D:248:GLU:OE1	2:D:307:ARG:NH2	2.42	0.50
1:A:17:ASP:OD1	1:A:18:GLY:N	2.41	0.49
1:A:131:THR:HG22	1:A:143:ARG:HG2	1.95	0.49
1:C:123:ASP:N	1:C:123:ASP:OD1	2.46	0.48
1:C:285:GLY:N	3:F:9:DT:OP1	2.47	0.47
2:D:266:TRP:CE3	2:D:425:LEU:HD22	2.49	0.47
1:A:206:ARG:NH2	1:A:216:THR:O	2.47	0.47
1:A:134:SER:OG	1:A:135:ILE:N	2.47	0.47
1:C:206:ARG:NH2	1:C:216:THR:O	2.42	0.47
1:C:139:THR:HB	1:C:140:PRO:HD2	1.96	0.46
1:C:323:LYS:NZ	1:C:344:GLU:OE2	2.48	0.46
1:C:132:ILE:HG21	1:C:142:ILE:HG12	1.98	0.46
1:A:23:GLN:NE2	1:A:60:VAL:O	2.46	0.45
1:A:134:SER:HB3	1:A:141:GLY:HA2	1.99	0.45
1:A:156:SER:HB2	1:A:157:PRO:HD3	1.98	0.44
1:A:61:PHE:CD1	1:A:61:PHE:N	2.86	0.44
1:C:396:GLU:OE1	1:C:396:GLU:N	2.44	0.44
2:B:362:THR:HG23	2:B:363:ASN:N	2.34	0.43
2:B:266:TRP:CZ3	2:B:425:LEU:HD21	2.53	0.43
1:A:194:GLU:HB3	1:A:197:GLN:HG3	2.01	0.43
2:D:123:ASP:O	2:D:126:LYS:NZ	2.47	0.42
2:D:298:GLU:OE1	2:D:298:GLU:N	2.51	0.42
2:B:249:LYS:HB2	2:B:252:TRP:CE2	2.54	0.42
2:D:323:LYS:O	2:D:385:LYS:NZ	2.52	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:323:LYS:O	2:B:385:LYS:NZ	2.53	0.42
2:B:90:VAL:HG12	2:B:90:VAL:O	2.20	0.42
1:C:122:GLU:HA	1:C:125:ARG:HE	1.84	0.42
1:C:130:PHE:CZ	1:C:144:TYR:HB2	2.55	0.42
2:D:246:LEU:HD11	2:D:264:LEU:HD21	2.01	0.41
1:A:47:ILE:CG2	1:A:144:TYR:HB3	2.50	0.41
1:A:270:ILE:HG23	1:A:271:TYR:N	2.35	0.41
1:C:133:PRO:HB3	1:C:137:ASN:ND2	2.36	0.41
2:D:266:TRP:CZ3	2:D:425:LEU:HD22	2.55	0.41
1:C:5:ILE:HG13	1:C:119:PRO:HD2	2.02	0.41
2:D:64:LYS:HE3	2:D:71:TRP:CE2	2.56	0.41
1:A:139:THR:HB	1:A:140:PRO:HD2	2.03	0.41
1:A:57:ASN:OD1	1:A:131:THR:N	2.54	0.40
2:B:72:ARG:CG	2:B:226:PRO:HB3	2.50	0.40
1:A:122:GLU:HA	1:A:125:ARG:HE	1.86	0.40
2:B:58:THR:HG23	2:B:76:ASP:O	2.21	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	551/555 (99%)	539 (98%)	12 (2%)	0	100	100
1	C	551/555 (99%)	540 (98%)	11 (2%)	0	100	100
2	B	412/444 (93%)	405 (98%)	7 (2%)	0	100	100
2	D	408/444 (92%)	396 (97%)	12 (3%)	0	100	100
All	All	1922/1998 (96%)	1880 (98%)	42 (2%)	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	495/495 (100%)	484 (98%)	11 (2%)	60	77
1	C	495/495 (100%)	479 (97%)	16 (3%)	46	62
2	B	378/403 (94%)	374 (99%)	4 (1%)	80	90
2	D	374/403 (93%)	370 (99%)	4 (1%)	80	90
All	All	1742/1796 (97%)	1707 (98%)	35 (2%)	63	79

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	THR
1	A	35	VAL
1	A	61	PHE
1	A	109	LEU
1	A	123	ASP
1	A	142	ILE
1	A	197	GLN
1	A	206	ARG
1	A	244	ILE
1	A	450	THR
1	A	516	GLU
2	B	11	LYS
2	B	169	GLU
2	B	173	LYS
2	B	209	LEU
1	C	5	ILE
1	C	7	THR
1	C	57	ASN
1	C	66	LYS
1	C	109	LEU
1	C	123	ASP
1	C	126	LYS
1	C	138	GLU
1	C	195	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	206	ARG
1	C	220	LYS
1	C	246	LEU
1	C	250	ASP
1	C	289	LEU
1	C	503	LEU
1	C	516	GLU
2	D	50	ILE
2	D	70	LYS
2	D	195	ILE
2	D	251	SER

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

Mol	Chain	Res	Type
1	A	137	ASN
1	C	545	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
3	OMC	E	2	3	15,22,23	3.04	6 (40%)	20,31,34	1.49	1 (5%)
3	OMC	E	4	3	15,22,23	3.05	6 (40%)	20,31,34	1.51	1 (5%)
3	OMC	F	2	3	15,22,23	3.02	6 (40%)	20,31,34	1.51	1 (5%)
3	OMC	F	4	3	15,22,23	3.02	6 (40%)	20,31,34	1.50	1 (5%)

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	OMC	E	2	3	-	0/5/27/28	0/2/2/2
3	OMC	E	4	3	-	0/5/27/28	0/2/2/2
3	OMC	F	2	3	-	0/5/27/28	0/2/2/2
3	OMC	F	4	3	-	0/5/27/28	0/2/2/2

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	4	OMC	C4-N4	2.29	1.41	1.35
3	E	2	OMC	C4-N4	2.30	1.41	1.35
3	F	4	OMC	C4-N4	2.30	1.41	1.35
3	F	2	OMC	C4-N4	2.30	1.41	1.35
3	F	4	OMC	C5-C4	2.43	1.46	1.41
3	F	2	OMC	C5-C4	2.44	1.46	1.41
3	E	4	OMC	C5-C4	2.57	1.47	1.41
3	E	2	OMC	C5-C4	2.59	1.47	1.41
3	E	2	OMC	C2-N3	4.42	1.47	1.38
3	E	4	OMC	C2-N3	4.57	1.47	1.38
3	F	4	OMC	C2-N3	4.58	1.47	1.38
3	F	2	OMC	C2-N3	4.65	1.47	1.38
3	F	4	OMC	C6-C5	4.89	1.48	1.38
3	F	2	OMC	C6-C5	4.91	1.48	1.38
3	E	4	OMC	C6-C5	4.98	1.48	1.38
3	E	2	OMC	C6-C5	5.01	1.49	1.38
3	F	4	OMC	C4-N3	5.14	1.44	1.35
3	E	2	OMC	C4-N3	5.15	1.44	1.35
3	F	2	OMC	C4-N3	5.16	1.44	1.35
3	E	4	OMC	C4-N3	5.17	1.44	1.35
3	F	2	OMC	C6-N1	7.25	1.45	1.35
3	F	4	OMC	C6-N1	7.33	1.45	1.35
3	E	2	OMC	C6-N1	7.38	1.45	1.35
3	E	4	OMC	C6-N1	7.41	1.45	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	2	OMC	C6-C5-C4	5.61	119.64	117.44
3	E	2	OMC	C6-C5-C4	5.73	119.68	117.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	4	OMC	C6-C5-C4	5.74	119.68	117.44
3	E	4	OMC	C6-C5-C4	6.00	119.79	117.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

15 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	$\#\|Z\|>2$
4	SO4	A	601	-	4,4,4	0.23	0	6,6,6	0.07	0
5	GOL	A	602	-	5,5,5	0.39	0	5,5,5	0.18	0
6	SUC	B	2001	-	24,24,24	0.43	0	36,36,36	0.67	0
5	GOL	B	2002	-	5,5,5	0.31	0	5,5,5	0.40	0
5	GOL	B	2003	-	5,5,5	0.43	0	5,5,5	0.17	0
5	GOL	B	2004	-	5,5,5	0.38	0	5,5,5	0.24	0
5	GOL	B	2005	-	5,5,5	0.34	0	5,5,5	0.37	0
5	GOL	B	2006	-	5,5,5	0.37	0	5,5,5	0.33	0
4	SO4	C	601	-	4,4,4	0.23	0	6,6,6	0.07	0
5	GOL	C	602	-	5,5,5	0.37	0	5,5,5	0.09	0
6	SUC	D	501	-	24,24,24	0.43	0	36,36,36	0.59	0
5	GOL	D	502	-	5,5,5	0.38	0	5,5,5	0.23	0
5	GOL	D	503	-	5,5,5	0.41	0	5,5,5	0.53	0
5	GOL	D	504	-	5,5,5	0.36	0	5,5,5	0.31	0
5	GOL	F	101	-	5,5,5	0.35	0	5,5,5	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
4	SO4	A	601	-	-	0/0/0/0	0/0/0/0
5	GOL	A	602	-	-	0/4/4/4	0/0/0/0
6	SUC	B	2001	-	-	0/12/51/51	0/2/2/2
5	GOL	B	2002	-	-	0/4/4/4	0/0/0/0
5	GOL	B	2003	-	-	0/4/4/4	0/0/0/0
5	GOL	B	2004	-	-	0/4/4/4	0/0/0/0
5	GOL	B	2005	-	-	0/4/4/4	0/0/0/0
5	GOL	B	2006	-	-	0/4/4/4	0/0/0/0
4	SO4	C	601	-	-	0/0/0/0	0/0/0/0
5	GOL	C	602	-	-	0/4/4/4	0/0/0/0
6	SUC	D	501	-	-	0/12/51/51	0/2/2/2
5	GOL	D	502	-	-	0/4/4/4	0/0/0/0
5	GOL	D	503	-	-	0/4/4/4	0/0/0/0
5	GOL	D	504	-	-	0/4/4/4	0/0/0/0
5	GOL	F	101	-	-	0/4/4/4	0/0/0/0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	2003	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	553/555 (99%)	0.82	75 (13%) 4 6	29, 71, 143, 210	0
1	C	553/555 (99%)	0.68	66 (11%) 6 9	29, 75, 142, 198	0
2	B	415/444 (93%)	0.56	35 (8%) 14 19	27, 61, 113, 157	0
2	D	410/444 (92%)	0.83	69 (16%) 2 3	33, 68, 130, 187	0
3	E	32/38 (84%)	-0.01	0 100 100	53, 80, 101, 116	0
3	F	32/38 (84%)	0.17	2 (6%) 23 31	52, 79, 113, 150	0
All	All	1995/2074 (96%)	0.70	247 (12%) 5 8	27, 69, 133, 210	0

All (247) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	90	VAL	14.6
1	C	24	TRP	12.9
2	D	358	ARG	9.6
2	D	88	TRP	8.6
2	B	88	TRP	8.6
1	C	67	ASP	8.6
1	A	24	TRP	8.6
1	A	134	SER	8.1
1	C	71	TRP	7.7
1	A	139	THR	7.6
1	A	50	ILE	7.6
2	D	232	TYR	7.6
2	B	4	PRO	7.5
2	B	90	VAL	7.4
1	A	140	PRO	7.1
1	A	63	ILE	7.0
1	A	553	SER	6.7
2	D	359	GLY	6.6
2	D	361	HIS	6.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	67	ASP	6.4
1	A	30	LYS	6.3
1	A	135	ILE	6.2
2	D	92	LEU	6.2
1	C	68	SER	6.1
2	B	214	LEU	6.1
2	D	91	GLN	6.0
1	C	25	PRO	5.8
1	C	136	ASN	5.7
1	A	133	PRO	5.6
1	A	26	LEU	5.4
1	A	49	LYS	5.4
1	A	136	ASN	5.4
2	D	89	GLU	5.3
1	C	66	LYS	5.3
2	D	360	ALA	5.3
1	A	70	LYS	5.2
2	B	215	THR	5.2
1	C	137	ASN	5.2
1	A	251	SER	5.2
1	A	71	TRP	5.2
2	D	409	THR	5.0
2	D	186	ASP	5.0
2	D	184	MET	4.9
1	A	2	ILE	4.8
2	B	227	PHE	4.8
2	D	356	ARG	4.7
1	A	27	THR	4.7
2	D	357	MET	4.6
1	A	550	LYS	4.6
2	B	89	GLU	4.5
2	D	427	TYR	4.4
1	A	137	ASN	4.4
1	A	138	GLU	4.3
1	C	52	PRO	4.3
1	A	286	THR	4.3
1	C	287	LYS	4.3
1	A	64	LYS	4.2
2	D	67	ASP	4.2
2	D	87	PHE	4.2
1	A	62	ALA	4.2
1	A	15	GLY	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	289	LEU	4.2
2	B	284	ARG	4.1
1	A	66	LYS	4.1
1	A	141	GLY	4.1
1	C	286	THR	4.0
1	C	15	GLY	4.0
1	A	28	GLU	4.0
2	B	229	TRP	4.0
2	D	173	LYS	4.0
2	D	226	PRO	4.0
1	C	63	ILE	3.9
2	B	226	PRO	3.9
1	A	22	LYS	3.9
2	D	187	LEU	3.9
1	C	28	GLU	3.9
2	D	230	MET	3.9
2	D	68	SER	3.8
1	C	56	TYR	3.8
1	C	64	LYS	3.8
1	C	140	PRO	3.8
1	A	25	PRO	3.8
2	D	318[A]	TYR	3.7
1	C	135	ILE	3.7
1	A	68	SER	3.7
2	D	93	GLY	3.7
1	C	346	PHE	3.7
1	C	69	THR	3.7
1	C	195	ILE	3.7
2	D	188	TYR	3.7
1	C	70	LYS	3.6
1	A	290	THR	3.6
1	C	292	VAL	3.5
2	D	111	VAL	3.5
2	D	212	TRP	3.5
2	D	156	SER	3.5
1	C	29	GLU	3.5
1	C	43	LYS	3.4
1	C	177	ASP	3.4
2	B	43	LYS	3.4
2	B	92	LEU	3.4
2	B	212	TRP	3.4
1	C	72	ARG	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	91	GLN	3.4
1	C	139	THR	3.4
2	D	185	ASP	3.4
2	B	14	PRO	3.3
2	D	362	THR	3.3
1	C	288	ALA	3.3
2	D	152	GLY	3.3
2	D	423	VAL	3.3
2	B	184	MET	3.3
1	A	43	LYS	3.3
2	D	183	TYR	3.2
1	C	311	LYS	3.2
2	B	173	LYS	3.2
1	A	65	LYS	3.2
1	C	65	LYS	3.2
1	A	547	GLN	3.1
1	A	346	PHE	3.1
2	D	69	THR	3.1
1	A	549	ASP	3.1
2	D	229	TRP	3.1
3	F	16	DT	3.1
1	A	132	ILE	3.0
1	C	130	PHE	3.0
1	A	287	LYS	3.0
1	C	285	GLY	3.0
2	B	210	LEU	3.0
1	A	449	GLU	3.0
1	A	74	LEU	3.0
2	D	109	LEU	3.0
2	D	43	LYS	3.0
2	D	410	TRP	3.0
2	D	8	VAL	2.9
2	D	66	LYS	2.9
1	A	548	VAL	2.9
2	D	75	VAL	2.9
1	A	311	LYS	2.9
1	A	177	ASP	2.9
2	D	176	PRO	2.9
1	A	144	TYR	2.8
2	B	174	GLN	2.8
2	B	428	GLN	2.8
2	B	75	VAL	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	295	LEU	2.8
1	C	2	ILE	2.8
2	B	211	ARG	2.8
2	B	6	GLU	2.8
1	C	14	PRO	2.8
1	A	3	SER	2.8
1	A	17	ASP	2.8
2	D	209	LEU	2.8
1	C	553	SER	2.8
2	D	110	ASP	2.8
2	D	207	GLN	2.8
1	C	293	ILE	2.7
1	A	247	PRO	2.7
1	C	44	GLU	2.7
1	A	69	THR	2.7
2	D	153	TRP	2.7
1	A	51	GLY	2.7
1	A	21	VAL	2.7
1	A	249	LYS	2.6
2	B	77	PHE	2.6
2	B	152	GLY	2.6
2	B	230	MET	2.6
2	D	424	LYS	2.6
2	D	95	PRO	2.6
1	C	22	LYS	2.6
1	A	551	LEU	2.6
1	A	248	GLU	2.6
1	C	26	LEU	2.6
1	A	355	ALA	2.6
1	C	181	TYR	2.6
1	C	138	GLU	2.6
1	C	182	GLN	2.5
1	A	199	ARG	2.5
3	F	17	DT	2.5
2	B	202	ILE	2.5
2	D	200	THR	2.5
1	A	61	PHE	2.5
2	D	94	ILE	2.5
1	A	250	ASP	2.5
1	A	72	ARG	2.4
1	A	288	ALA	2.4
2	D	112	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	44	GLU	2.4
1	C	194	GLU	2.4
1	C	289	LEU	2.4
2	D	170	PRO	2.4
2	D	171	PHE	2.4
2	D	420	PRO	2.4
1	C	17	ASP	2.4
2	B	408	ALA	2.4
1	C	180	ILE	2.4
1	A	245	VAL	2.4
1	C	312	GLU	2.4
1	C	442	VAL	2.4
2	D	195	ILE	2.3
2	D	74	LEU	2.3
2	D	151	GLN	2.3
1	A	448	ARG	2.3
2	B	5	ILE	2.3
1	A	295	LEU	2.3
2	B	410	TRP	2.3
1	C	53	GLU	2.3
1	C	449	GLU	2.3
2	D	70	LYS	2.3
1	A	142	ILE	2.3
1	A	252	TRP	2.3
2	D	10	VAL	2.3
1	A	544	GLY	2.3
1	C	5	ILE	2.3
1	C	187	LEU	2.3
2	D	231	GLY	2.3
1	A	337	TRP	2.3
1	C	11	LYS	2.3
2	B	409	THR	2.2
2	D	206	ARG	2.2
1	A	546	GLU	2.2
2	B	369	THR	2.2
1	C	457	TYR	2.2
2	D	108	VAL	2.2
1	C	16	MET	2.2
1	A	338	THR	2.2
1	A	552	VAL	2.2
2	D	7	THR	2.2
1	C	142	ILE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	195	ILE	2.2
1	C	188	TYR	2.1
1	C	50	ILE	2.1
2	D	155	GLY	2.1
1	C	223	LYS	2.1
1	C	290	THR	2.1
1	C	544	GLY	2.1
2	D	315	HIS	2.1
2	D	284	ARG	2.1
2	B	93	GLY	2.1
1	C	40	GLU	2.0
1	A	282	LEU	2.0
2	D	411	ILE	2.0
2	D	425	LEU	2.0
1	C	196	GLY	2.0
2	B	153	TRP	2.0
1	A	210	LEU	2.0
1	C	161	GLN	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
3	OMC	E	4	21/22	0.97	0.11	-	43,49,56,58	0
3	OMC	F	4	21/22	0.97	0.12	-	48,55,62,71	0
3	OMC	F	2	21/22	0.96	0.13	-	66,70,79,92	0
3	OMC	E	2	21/22	0.96	0.13	-	53,60,74,78	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
5	GOL	F	101	6/6	0.83	0.27	5.23	71,80,86,91	0
5	GOL	D	504	6/6	0.71	0.26	5.18	68,75,77,81	0
5	GOL	A	602	6/6	0.74	0.29	3.75	42,78,87,92	0
5	GOL	D	502	6/6	0.82	0.24	3.56	69,79,86,88	0
5	GOL	B	2006	6/6	0.81	0.26	2.30	53,70,73,84	0
5	GOL	B	2003	6/6	0.91	0.22	1.95	65,77,82,93	0
4	SO4	C	601	5/5	0.93	0.15	1.94	80,106,113,122	0
5	GOL	C	602	6/6	0.88	0.21	1.21	49,61,70,73	0
5	GOL	B	2004	6/6	0.96	0.16	1.16	46,49,59,60	0
6	SUC	B	2001	23/23	0.90	0.20	1.12	40,75,87,95	0
6	SUC	D	501	23/23	0.94	0.17	0.99	48,73,82,85	0
4	SO4	A	601	5/5	0.91	0.15	0.86	82,103,107,116	0
5	GOL	B	2005	6/6	0.91	0.19	0.47	34,47,52,58	0
5	GOL	D	503	6/6	0.95	0.17	0.37	44,49,53,55	0
5	GOL	B	2002	6/6	0.92	0.12	-0.36	45,53,62,67	0

6.5 Other polymers

There are no such residues in this entry.