Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 10:02 PM GMT

PDB ID	:	1RRV
Title	:	X-ray crystal structure of TDP-vancosaminyltransferase GtfD as a complex with TDP and the natural substrate, desvancosaminyl vancomycin.
Authors	:	Mulichak, A.M.; Lu, W.; Losey, H.C.; Walsh, C.T.; Garavito, R.M.
Deposited on	:	2003-12-09
Resolution	:	2.00 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

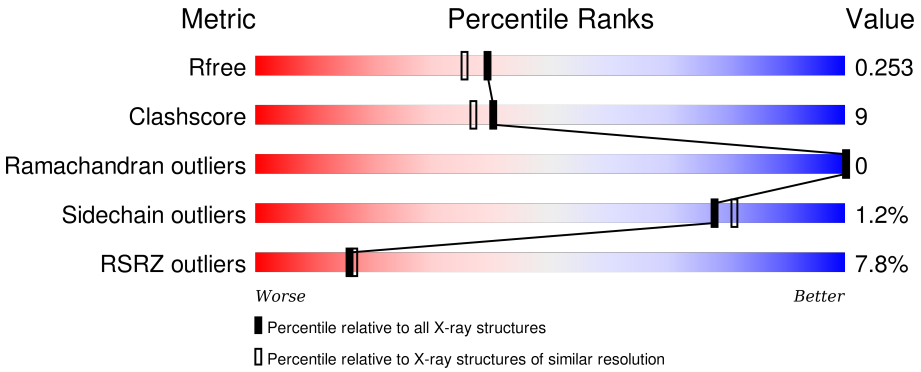
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.00 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	416	5%, 82%, 13%, ·
1	B	416	10%, 77%, 19%, ·
2	C	7	71%, 29%
2	D	7	57%, 43%

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	1006	-	-	-	X
5	GOL	B	1007	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6567 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 GLYCOSYLTRANSFERASE GTFD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	401	Total	C	N	O	S	0	0	0
			2915	1848	519	536	12			
1	B	400	Total	C	N	O	S	0	0	0
			2900	1838	515	535	12			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	LEU	CYS	CLONING ARTIFACT	UNP Q9AFC7
A	196	LEU	TRP	CLONING ARTIFACT	UNP Q9AFC7
A	197	ALA	PRO	CLONING ARTIFACT	UNP Q9AFC7
A	409	LEU	-	CLONING ARTIFACT	UNP Q9AFC7
A	410	GLU	-	CLONING ARTIFACT	UNP Q9AFC7
A	411	HIS	-	EXPRESSION TAG	UNP Q9AFC7
A	412	HIS	-	EXPRESSION TAG	UNP Q9AFC7
A	413	HIS	-	EXPRESSION TAG	UNP Q9AFC7
A	414	HIS	-	EXPRESSION TAG	UNP Q9AFC7
A	415	HIS	-	EXPRESSION TAG	UNP Q9AFC7
A	416	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	195	LEU	CYS	CLONING ARTIFACT	UNP Q9AFC7
B	196	LEU	TRP	CLONING ARTIFACT	UNP Q9AFC7
B	197	ALA	PRO	CLONING ARTIFACT	UNP Q9AFC7
B	409	LEU	-	CLONING ARTIFACT	UNP Q9AFC7
B	410	GLU	-	CLONING ARTIFACT	UNP Q9AFC7
B	411	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	412	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	413	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	414	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	415	HIS	-	EXPRESSION TAG	UNP Q9AFC7
B	416	HIS	-	EXPRESSION TAG	UNP Q9AFC7

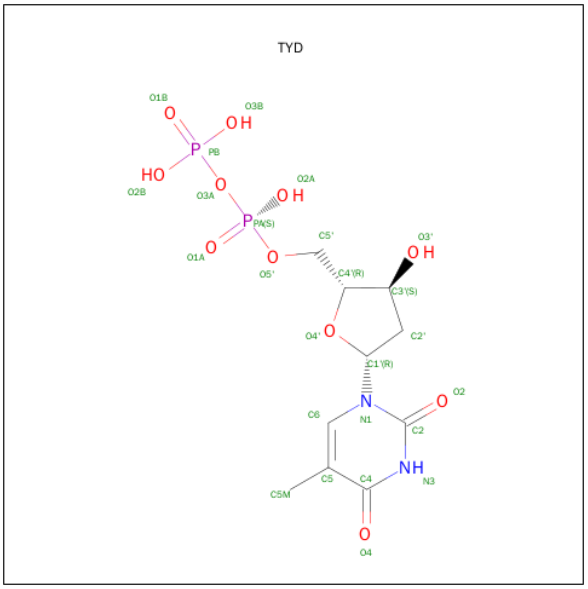
- Molecule 2 is a protein called DESVANCOSAMINYL VANCOMYCIN.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	C	7	Total C Cl N O 80 53 2 8 17	0	0	0
2	D	7	Total C Cl N O 80 53 2 8 17	0	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

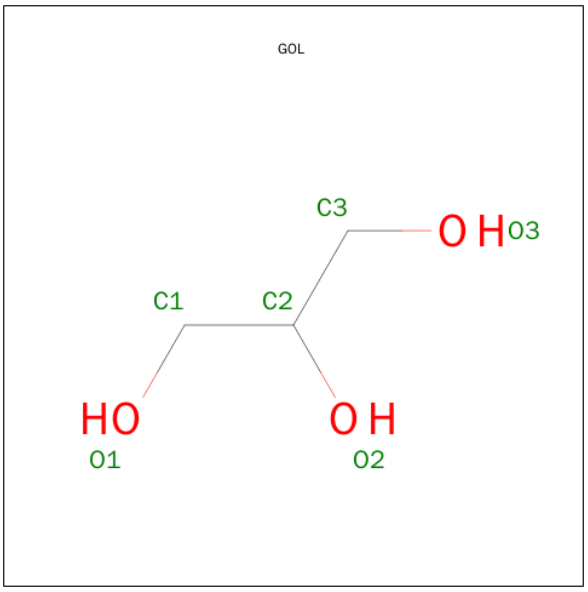
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total K 1 1	0	0
3	A	1	Total K 1 1	0	0

- Molecule 4 is THYMIDINE-5'-DIPHOSPHATE (three-letter code: TYD) (formula: $C_{10}H_{16}N_2O_{11}P_2$).



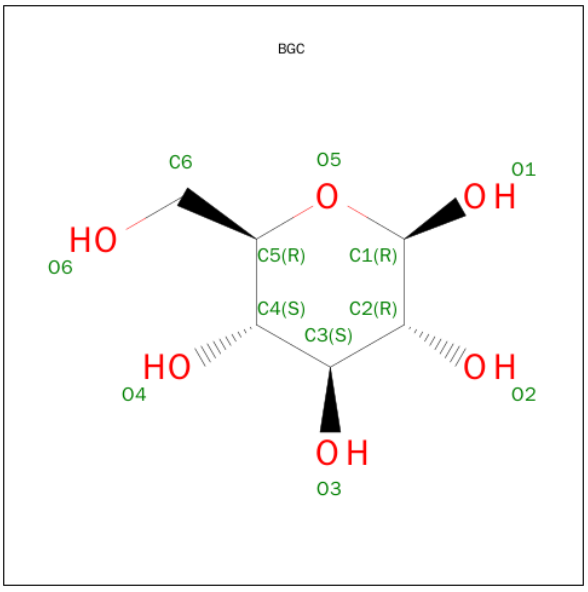
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P 25 10 2 11 2	0	0
4	B	1	Total C N O P 25 10 2 11 2	0	0

- 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	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SUGAR (BETA-D-GLUCOSE) (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			11	6	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			11	6	5		

- Molecule 7 is water.

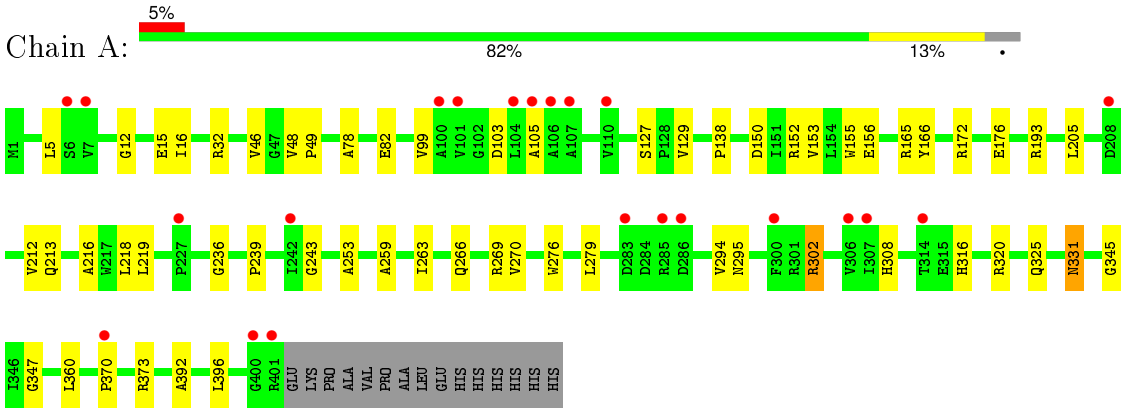
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	303	Total	O	0	0
			303	303		
7	B	177	Total	O	0	0
			177	177		
7	C	13	Total	O	0	0
			13	13		
7	D	7	Total	O	0	0
			7	7		

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.

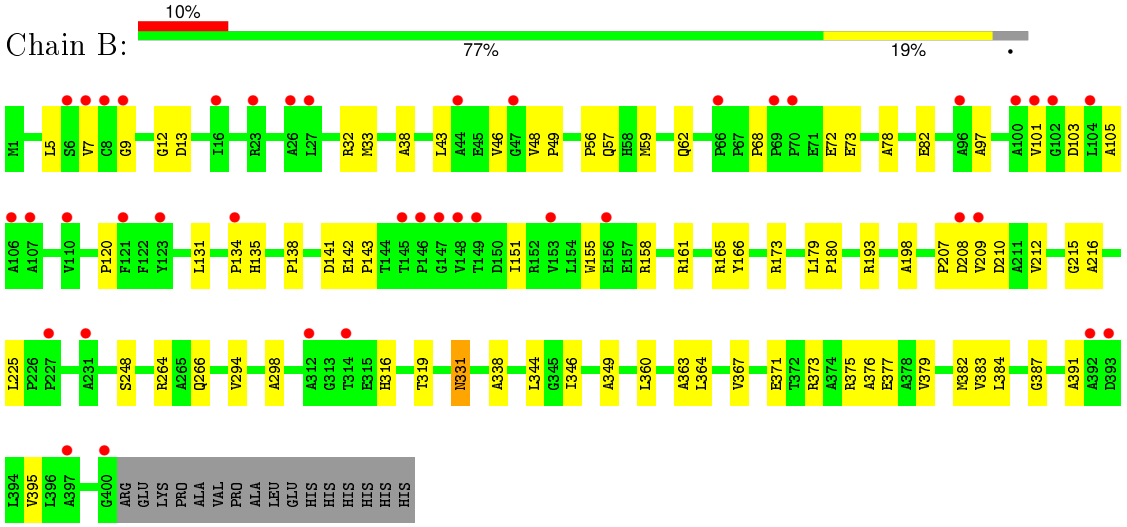
- Molecule 1: GLYCOSYLTRANSFERASE GTFD

Chain A:



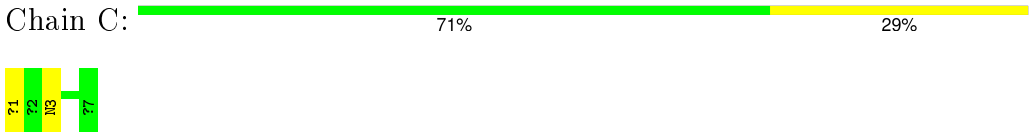
- Molecule 1: GLYCOSYLTRANSFERASE GTFD

Chain B:



- Molecule 2: DESVANCOSAMINYL VANCOMYCIN

Chain C:



- Molecule 2: DESVANCOSAMINYL VANCOMYCIN

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.67Å 64.12Å 144.08Å 90.00° 91.73° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 28.93 – 2.08	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.00) 99.6 (28.93-2.08)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	5.14 (at 2.08Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.210 , 0.252 0.212 , 0.253	Depositor DCC
R_{free} test set	2668 reflections (4.82%)	DCC
Wilson B-factor (Å^2)	31.9	Xtriage
Anisotropy	0.477	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 53.3	EDS
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 55519 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6567	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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: GOL, BGC, GHP, TYD, K, 3FG, MLU, OMZ, OMY

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	0.32	0/2985	0.58	0/4093
1	B	0.28	0/2970	0.56	0/4075
2	C	1.05	0/7	0.83	0/8
2	D	1.14	0/7	0.90	0/8
All	All	0.31	0/5969	0.57	0/8184

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

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	2915	0	2891	47	0
1	B	2900	0	2867	57	0
2	C	80	0	44	1	0
2	D	80	0	44	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	25	0	13	3	0
4	B	25	0	13	2	0
5	A	12	0	14	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	6	0	6	0	0
6	C	11	0	10	0	0
6	D	11	0	10	0	0
7	A	303	0	0	9	0
7	B	177	0	0	2	0
7	C	13	0	0	0	0
7	D	7	0	0	0	0
All	All	6567	0	5912	104	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 9.

All (104) 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:A:205:LEU:H	1:A:213:GLN:HE22	1.28	0.80
1:B:12:GLY:HA2	4:B:1004:TYD:H5'1	1.70	0.73
1:A:345:GLY:HA2	7:A:2274:HOH:O	1.91	0.70
1:B:166:TYR:OH	2:D:1:MLU:HCN3	1.93	0.69
1:B:319:THR:HG22	1:B:383:VAL:HG21	1.77	0.66
1:B:193:ARG:HD2	1:B:212:VAL:HG23	1.79	0.65
1:B:373:ARG:O	1:B:377:GLU:HG3	1.97	0.64
1:B:12:GLY:HA2	4:B:1004:TYD:C5'	2.28	0.64
1:B:346:ILE:HG22	1:B:376:ALA:HA	1.78	0.64
1:B:46:VAL:HG23	1:B:48:VAL:HG23	1.81	0.63
1:B:379:VAL:HA	1:B:382:MET:CE	2.30	0.62
1:A:166:TYR:OH	2:C:1:MLU:HCN3	1.99	0.61
1:A:12:GLY:O	1:A:16:ILE:HG12	2.01	0.61
1:B:135:HIS:NE2	1:B:207:PRO:HG2	2.15	0.61
1:B:173:ARG:HD2	7:B:2067:HOH:O	2.00	0.60
1:A:276:TRP:HZ3	5:A:1005:GOL:H12	1.67	0.60
1:A:78:ALA:O	1:A:82:GLU:HG3	2.02	0.59
1:B:216:ALA:HB2	1:B:316:HIS:CD2	2.37	0.59
1:B:32:ARG:HA	1:B:49:PRO:HG2	1.85	0.59
1:A:12:GLY:HA2	4:A:1003:TYD:H5'1	1.85	0.59
1:A:12:GLY:HA2	4:A:1003:TYD:C5'	2.33	0.58
1:A:129:VAL:HG22	7:A:2120:HOH:O	2.03	0.58
1:B:384:LEU:HD21	7:B:2144:HOH:O	2.05	0.56
1:A:127:SER:HB2	7:A:2120:HOH:O	2.05	0.56
1:A:46:VAL:HG23	1:A:48:VAL:HG23	1.87	0.55
1:A:32:ARG:HA	1:A:49:PRO:HG2	1.87	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:259:ALA:HA	1:A:360:LEU:CD2	2.38	0.54
1:A:150:ASP:OD2	1:A:153:VAL:HG23	2.08	0.54
1:A:259:ALA:HA	1:A:360:LEU:HD21	1.91	0.53
1:B:266:GLN:HG3	1:B:364:LEU:HD23	1.92	0.52
1:A:172:ARG:O	1:A:176:GLU:HG3	2.09	0.52
1:B:379:VAL:HA	1:B:382:MET:HE3	1.91	0.52
1:B:43:LEU:HD22	1:B:48:VAL:HB	1.92	0.52
1:A:331:ASN:C	1:A:331:ASN:HD22	2.14	0.51
1:B:379:VAL:HA	1:B:382:MET:HE2	1.92	0.50
1:A:205:LEU:H	1:A:213:GLN:NE2	2.03	0.50
1:B:135:HIS:CD2	1:B:207:PRO:HG2	2.46	0.50
1:A:152:ARG:O	1:A:156:GLU:HG3	2.11	0.50
1:A:138:PRO:HG3	1:A:155:TRP:CD2	2.46	0.50
1:B:338:ALA:O	1:B:349:ALA:HB2	2.12	0.50
1:B:56:PRO:HG2	1:B:59:MET:HG2	1.94	0.50
1:B:165:ARG:HD3	2:D:1:MLU:HCN1	1.94	0.50
1:A:219:LEU:C	1:A:219:LEU:HD12	2.32	0.49
1:B:193:ARG:NH1	1:B:210:ASP:OD2	2.45	0.49
1:A:16:ILE:HD11	7:A:2256:HOH:O	2.12	0.49
1:B:207:PRO:O	1:B:208:ASP:HB2	2.12	0.49
1:A:193:ARG:HD2	1:A:212:VAL:HG23	1.95	0.48
1:A:392:ALA:O	1:A:396:LEU:HG	2.14	0.48
1:A:236:GLY:HA3	1:A:269:ARG:CZ	2.44	0.48
1:A:219:LEU:O	1:A:219:LEU:HD12	2.14	0.47
1:B:383:VAL:HG23	1:B:383:VAL:O	2.15	0.47
1:B:68:PRO:HG2	1:B:73:GLU:HG2	1.97	0.47
1:A:302:ARG:HH11	1:A:302:ARG:HB3	1.79	0.47
1:B:225:LEU:HD21	1:B:298:ALA:HB3	1.97	0.47
1:B:62:GLN:HG2	1:B:248:SER:HB2	1.95	0.47
1:B:225:LEU:HD22	1:B:294:VAL:CG1	2.45	0.47
1:A:263:ILE:HG21	1:A:270:VAL:HG22	1.96	0.47
1:A:103:ASP:OD1	1:A:105:ALA:HB3	2.14	0.47
1:B:193:ARG:HD2	1:B:212:VAL:CG2	2.46	0.46
1:B:141:ASP:OD1	1:B:142:GLU:HG2	2.15	0.46
1:A:243:GLY:O	1:A:308:HIS:HA	2.16	0.46
1:A:259:ALA:O	1:A:263:ILE:HG13	2.15	0.45
1:A:138:PRO:HG3	1:A:155:TRP:CE2	2.51	0.45
1:A:325:GLN:O	1:A:347:GLY:HA3	2.16	0.45
1:B:78:ALA:O	1:B:82:GLU:HG3	2.16	0.45
1:B:143:PRO:HD3	2:D:7:3FG:O	2.17	0.45
1:B:5:LEU:O	1:B:33:MET:HA	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:LEU:HD23	7:A:2208:HOH:O	2.17	0.45
1:B:68:PRO:HB2	1:B:72:GLU:HB3	1.99	0.44
1:A:370:PRO:HG2	7:A:2290:HOH:O	2.18	0.44
1:A:216:ALA:HB2	1:A:316:HIS:CD2	2.52	0.44
1:B:131:LEU:HD13	1:B:209:VAL:HG11	2.00	0.44
1:B:9:GLY:HA3	1:B:13:ASP:OD2	2.18	0.44
1:B:331:ASN:HD22	1:B:331:ASN:C	2.21	0.44
1:A:239:PRO:HB3	1:A:269:ARG:HB2	2.00	0.44
1:A:294:VAL:CG1	1:A:295:ASN:N	2.81	0.44
1:A:15:GLU:HG3	7:A:2208:HOH:O	2.19	0.43
7:A:2288:HOH:O	1:B:264:ARG:HB3	2.19	0.43
1:B:138:PRO:HG3	1:B:155:TRP:CE2	2.54	0.43
1:A:320:ARG:HD3	7:A:2253:HOH:O	2.18	0.43
1:A:253:ALA:HA	1:A:279:LEU:HD22	2.00	0.43
1:B:344:LEU:O	1:B:375:ARG:NH1	2.52	0.43
1:B:193:ARG:HG3	1:B:193:ARG:HH11	1.83	0.43
1:A:5:LEU:HD23	1:A:99:VAL:HB	2.01	0.43
1:B:198:ALA:O	1:B:215:GLY:HA2	2.19	0.42
1:B:179:LEU:HA	1:B:180:PRO:HD3	1.88	0.42
1:B:158:ARG:HA	1:B:161:ARG:NH2	2.34	0.42
1:B:103:ASP:OD1	1:B:105:ALA:HB3	2.20	0.42
1:A:316:HIS:O	1:A:320:ARG:HG2	2.19	0.42
1:A:12:GLY:HA2	4:A:1003:TYD:H5'2	2.02	0.42
1:B:97:ALA:CB	1:B:120:PRO:HG2	2.49	0.41
1:B:391:ALA:O	1:B:395:VAL:HG23	2.19	0.41
1:B:363:ALA:O	1:B:367:VAL:HG23	2.20	0.41
1:B:138:PRO:HG3	1:B:155:TRP:CD2	2.55	0.41
1:B:215:GLY:O	1:B:387:GLY:HA3	2.21	0.41
1:B:360:LEU:HD12	1:B:360:LEU:HA	1.86	0.41
1:A:294:VAL:HG12	1:A:295:ASN:N	2.35	0.41
1:B:7:VAL:HG22	1:B:101:VAL:HB	2.01	0.41
1:B:38:ALA:HB2	1:B:57:GLN:HB2	2.03	0.41
1:B:375:ARG:HA	1:B:375:ARG:HD2	1.85	0.41
1:A:165:ARG:HG3	1:A:165:ARG:HH11	1.86	0.41
1:A:276:TRP:CZ3	5:A:1005:GOL:H12	2.52	0.40
1:B:134:PRO:O	1:B:151:ILE:HD12	2.22	0.40
1:B:371:GLU:H	1:B:371:GLU:CD	2.25	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	399/416 (96%)	390 (98%)	9 (2%)	0	100	100
1	B	398/416 (96%)	384 (96%)	14 (4%)	0	100	100
2	C	1/7 (14%)	1 (100%)	0	0	100	100
2	D	1/7 (14%)	1 (100%)	0	0	100	100
All	All	799/846 (94%)	776 (97%)	23 (3%)	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	289/314 (92%)	285 (99%)	4 (1%)	74	77
1	B	287/314 (91%)	286 (100%)	1 (0%)	94	96
2	C	1/1 (100%)	0	1 (100%)	0	0
2	D	1/1 (100%)	0	1 (100%)	0	0
All	All	578/630 (92%)	571 (99%)	7 (1%)	78	81

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	266	GLN
1	A	302	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	331	ASN
1	A	373	ARG
1	B	331	ASN
2	C	3	ASN
2	D	3	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	213	GLN
1	A	331	ASN
1	B	266	GLN
1	B	331	ASN
2	C	3	ASN
2	D	3	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)

12 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	MLU	C	1	2	7,8,9	0.94	1 (14%)	4,9,11	1.13	1 (25%)
2	OMZ	C	2	2	13,14,15	1.82	5 (38%)	17,19,21	0.97	1 (5%)
2	GHP	C	4	2,6	10,11,12	1.65	2 (20%)	12,14,16	1.11	1 (8%)
2	GHP	C	5	2	10,11,12	1.48	3 (30%)	12,14,16	2.37	5 (41%)
2	OMY	C	6	2	13,14,15	2.04	4 (30%)	17,19,21	1.66	1 (5%)
2	3FG	C	7	2	9,13,13	2.80	5 (55%)	14,18,18	1.33	1 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	MLU	D	1	2	7,8,9	0.84	0	4,9,11	1.08	1 (25%)
2	OMZ	D	2	2	13,14,15	1.94	5 (38%)	17,19,21	1.07	1 (5%)
2	GHP	D	4	2,6	10,11,12	1.75	4 (40%)	12,14,16	1.09	1 (8%)
2	GHP	D	5	2	10,11,12	1.67	2 (20%)	12,14,16	2.24	4 (33%)
2	OMY	D	6	2	13,14,15	2.35	6 (46%)	17,19,21	1.69	1 (5%)
2	3FG	D	7	2	9,13,13	3.29	6 (66%)	14,18,18	1.38	1 (7%)

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	MLU	C	1	2	-	0/4/8/10	0/0/0/0
2	OMZ	C	2	2	-	0/8/10/12	0/1/1/1
2	GHP	C	4	2,6	-	0/4/6/8	0/1/1/1
2	GHP	C	5	2	-	0/4/6/8	0/1/1/1
2	OMY	C	6	2	-	0/8/10/12	0/1/1/1
2	3FG	C	7	2	-	0/4/8/8	0/1/1/1
2	MLU	D	1	2	-	0/4/8/10	0/0/0/0
2	OMZ	D	2	2	-	0/8/10/12	0/1/1/1
2	GHP	D	4	2,6	-	0/4/6/8	0/1/1/1
2	GHP	D	5	2	-	0/4/6/8	0/1/1/1
2	OMY	D	6	2	-	0/8/10/12	0/1/1/1
2	3FG	D	7	2	-	0/4/8/8	0/1/1/1

All (43) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	6	OMY	ODE-CB	-2.22	1.38	1.42
2	C	5	GHP	C6-C1	2.04	1.42	1.39
2	D	6	OMY	CD1-CG	2.06	1.42	1.39
2	D	6	OMY	CD1-CE1	2.07	1.42	1.38
2	C	2	OMZ	OH-CZ	2.20	1.41	1.36
2	C	7	3FG	CG1-CD1	2.22	1.42	1.39
2	C	1	MLU	CA-N	2.22	1.51	1.47
2	D	7	3FG	CG2-CB	2.23	1.42	1.39
2	D	4	GHP	C1-CA	2.23	1.54	1.52
2	C	7	3FG	CG2-CD2	2.27	1.42	1.39
2	D	4	GHP	O4-C4	2.27	1.42	1.37
2	C	5	GHP	CA-C	2.30	1.57	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2	OMZ	CD1-CG	2.32	1.42	1.39
2	C	4	GHP	C1-CA	2.34	1.54	1.52
2	C	2	OMZ	CE2-CZ	2.41	1.44	1.39
2	C	2	OMZ	CD2-CG	2.41	1.43	1.39
2	C	2	OMZ	CD1-CG	2.46	1.43	1.39
2	D	5	GHP	C3-C4	2.47	1.43	1.38
2	D	4	GHP	C5-C4	2.57	1.44	1.38
2	C	6	OMY	CD2-CG	2.58	1.43	1.39
2	C	5	GHP	C3-C4	2.58	1.44	1.38
2	D	6	OMY	OCZ-CZ	2.60	1.41	1.36
2	C	7	3FG	CZ-CD1	2.62	1.43	1.39
2	D	2	OMZ	CD2-CG	2.64	1.43	1.39
2	D	2	OMZ	CE2-CZ	2.85	1.44	1.39
2	D	6	OMY	CD2-CG	2.85	1.43	1.39
2	D	5	GHP	CA-C	2.90	1.59	1.50
2	C	6	OMY	CA-CB	2.91	1.59	1.54
2	D	2	OMZ	OH-CZ	2.92	1.42	1.36
2	D	7	3FG	CG2-CD2	2.94	1.43	1.39
2	D	7	3FG	CG1-CD1	2.95	1.43	1.39
2	D	4	GHP	C3-C2	2.95	1.44	1.38
2	D	6	OMY	CA-CB	3.06	1.59	1.54
2	D	7	3FG	CZ-CD1	3.13	1.44	1.39
2	C	4	GHP	C3-C2	3.42	1.44	1.38
2	D	2	OMZ	CA-CB	3.89	1.61	1.54
2	C	2	OMZ	CA-CB	4.23	1.62	1.54
2	C	6	OMY	CZ-CE1	4.47	1.43	1.39
2	C	7	3FG	CG1-CB	4.80	1.47	1.39
2	C	7	3FG	CB-CA	5.09	1.57	1.52
2	D	7	3FG	CG1-CB	5.28	1.47	1.39
2	D	6	OMY	CZ-CE1	5.31	1.44	1.39
2	D	7	3FG	CB-CA	5.98	1.58	1.52

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	6	OMY	CG-CB-CA	-6.12	103.40	111.96
2	C	6	OMY	CG-CB-CA	-6.09	103.44	111.96
2	C	5	GHP	C2-C1-CA	-3.01	115.51	120.70
2	D	5	GHP	C2-C1-CA	-2.57	116.25	120.70
2	D	2	OMZ	O-C-CA	-2.57	118.66	125.44
2	C	2	OMZ	O-C-CA	-2.55	118.70	125.44
2	D	5	GHP	C5-C6-C1	-2.47	118.66	121.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	5	GHP	C5-C6-C1	-2.46	118.68	121.20
2	D	5	GHP	C2-C3-C4	-2.40	117.10	119.87
2	C	1	MLU	O-C-CA	-2.18	119.68	125.44
2	D	1	MLU	O-C-CA	-2.05	120.02	125.44
2	C	5	GHP	C2-C3-C4	-2.02	117.53	119.87
2	C	5	GHP	C6-C1-CA	2.21	124.52	120.70
2	C	4	GHP	C5-C6-C1	2.29	123.56	121.20
2	D	4	GHP	C5-C6-C1	2.32	123.59	121.20
2	D	7	3FG	C-CA-CB	2.74	115.05	111.31
2	C	7	3FG	C-CA-CB	2.94	115.32	111.31
2	D	5	GHP	C1-CA-N	5.37	125.19	112.54
2	C	5	GHP	C1-CA-N	5.81	126.22	112.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	MLU	1	0
2	D	1	MLU	2	0
2	D	7	3FG	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	$\#\|Z\|>2$
4	TYD	A	1003	-	19,26,26	1.08	3 (15%)	27,40,40	3.34	3 (11%)
5	GOL	A	1005	-	5,5,5	0.21	0	5,5,5	0.21	0
5	GOL	A	1006	3	5,5,5	0.28	0	5,5,5	0.30	0
4	TYD	B	1004	-	19,26,26	1.08	2 (10%)	27,40,40	3.30	3 (11%)
5	GOL	B	1007	3	5,5,5	0.22	0	5,5,5	0.19	0
6	BGC	C	8	2	11,11,12	1.13	0	14,15,17	1.84	3 (21%)
6	BGC	D	8	2	11,11,12	1.15	0	14,15,17	1.86	4 (28%)

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	TYD	A	1003	-	-	0/12/28/28	0/2/2/2
5	GOL	A	1005	-	-	0/4/4/4	0/0/0/0
5	GOL	A	1006	3	-	0/4/4/4	0/0/0/0
4	TYD	B	1004	-	-	0/12/28/28	0/2/2/2
5	GOL	B	1007	3	-	0/4/4/4	0/0/0/0
6	BGC	C	8	2	-	0/2/19/22	0/1/1/1
6	BGC	D	8	2	-	0/2/19/22	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1003	TYD	C6-C5	-2.05	1.34	1.40
4	A	1003	TYD	C6-N1	2.16	1.38	1.35
4	B	1004	TYD	C6-N1	2.35	1.38	1.35
4	A	1003	TYD	C4-N3	2.78	1.38	1.33
4	B	1004	TYD	C4-N3	2.79	1.38	1.33

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1003	TYD	C5-C4-N3	-8.96	115.16	125.14
4	B	1004	TYD	C5-C4-N3	-8.95	115.17	125.14
6	C	8	BGC	C2-C3-C4	-2.47	106.85	111.04
6	D	8	BGC	C2-C3-C4	-2.32	107.10	111.04
6	D	8	BGC	O5-C1-C2	-2.07	107.50	110.86
4	A	1003	TYD	C5M-C5-C6	2.10	122.84	118.62
6	D	8	BGC	C1-C2-C3	2.12	112.04	109.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	1004	TYD	C5M-C5-C6	2.14	122.94	118.62
6	C	8	BGC	C1-C2-C3	2.55	112.56	109.54
6	C	8	BGC	O6-C6-C5	4.94	127.64	111.33
6	D	8	BGC	O6-C6-C5	5.10	128.18	111.33
4	B	1004	TYD	C4-N3-C2	13.90	127.26	115.25
4	A	1003	TYD	C4-N3-C2	14.11	127.44	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1003	TYD	3	0
5	A	1005	GOL	2	0
4	B	1004	TYD	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{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	401/416 (96%)	0.28	22 (5%) 29 30	18, 31, 52, 60	0
1	B	400/416 (96%)	0.53	41 (10%) 9 9	29, 40, 58, 70	0
2	C	1/7 (14%)	-0.51	0 100 100	24, 24, 24, 24	0
2	D	1/7 (14%)	0.34	0 100 100	37, 37, 37, 37	0
All	All	803/846 (94%)	0.40	63 (7%) 16 17	18, 37, 55, 70	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	101	VAL	5.4
1	B	400	GLY	4.5
1	A	306	VAL	3.9
1	A	100	ALA	3.7
1	B	153	VAL	3.6
1	A	307	ILE	3.5
1	B	7	VAL	3.4
1	A	101	VAL	3.4
1	B	27	LEU	3.3
1	B	69	PRO	3.3
1	B	102	GLY	3.2
1	B	148	VAL	3.2
1	A	286	ASP	3.1
1	B	209	VAL	2.9
1	B	70	PRO	2.9
1	B	16	ILE	2.9
1	B	208	ASP	2.9
1	B	145	THR	2.9
1	B	47	GLY	2.9
1	B	397	ALA	2.9
1	B	26	ALA	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	110	VAL	2.7
1	B	96	ALA	2.7
1	B	104	LEU	2.6
1	B	123	TYR	2.6
1	A	242	ILE	2.6
1	B	393	ASP	2.6
1	B	146	PRO	2.5
1	B	44	ALA	2.5
1	B	23	ARG	2.5
1	A	400	GLY	2.5
1	B	110	VAL	2.5
1	A	227	PRO	2.5
1	A	314	THR	2.5
1	B	106	ALA	2.4
1	B	156	GLU	2.4
1	A	285	ARG	2.4
1	A	208	ASP	2.4
1	A	283	ASP	2.4
1	A	106	ALA	2.4
1	B	149	THR	2.4
1	B	312	ALA	2.3
1	A	6	SER	2.3
1	A	104	LEU	2.3
1	B	231	ALA	2.3
1	B	147	GLY	2.3
1	B	100	ALA	2.2
1	B	107	ALA	2.2
1	B	66	PRO	2.2
1	A	401	ARG	2.2
1	A	300	PHE	2.2
1	B	121	PHE	2.1
1	A	105	ALA	2.1
1	B	134	PRO	2.1
1	A	107	ALA	2.1
1	B	314	THR	2.1
1	B	9	GLY	2.1
1	A	370	PRO	2.1
1	B	227	PRO	2.1
1	B	392	ALA	2.1
1	B	6	SER	2.0
1	B	8	CYS	2.0
1	A	7	VAL	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
2	MLU	D	1	9/10	0.88	0.18	-	44,46,51,53	0
2	3FG	C	7	13/13	0.93	0.12	-	27,29,41,41	0
2	MLU	C	1	9/10	0.94	0.15	-	24,38,40,40	0
2	OMY	D	6	14/15	0.70	0.16	-	31,34,40,40	0
2	GHP	C	5	11/12	0.95	0.11	-	17,23,27,27	0
2	3FG	D	7	13/13	0.85	0.16	-	42,46,53,54	0
2	OMZ	D	2	14/15	0.68	0.21	-	35,37,39,40	0
2	OMZ	C	2	14/15	0.77	0.17	-	18,22,26,27	0
2	OMY	C	6	14/15	0.72	0.17	-	18,23,26,30	0
2	GHP	D	5	11/12	0.87	0.13	-	37,41,43,44	0
2	GHP	C	4	11/12	0.97	0.12	-	19,22,25,27	0
2	GHP	D	4	11/12	0.94	0.15	-	31,33,36,38	0

6.3 Carbohydrates (i)

There are no carbohydrates 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. 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	A	1006	6/6	0.88	0.27	4.45	32,43,45,49	0
5	GOL	B	1007	6/6	0.84	0.30	2.81	37,44,46,50	0
6	BGC	C	8	11/12	0.90	0.23	1.71	30,32,34,39	0
6	BGC	D	8	11/12	0.95	0.27	1.58	41,44,46,51	0
5	GOL	A	1005	6/6	0.97	0.11	0.51	36,38,38,40	0
4	TYD	B	1004	25/25	0.97	0.17	0.27	31,37,38,40	0
4	TYD	A	1003	25/25	0.97	0.15	-0.24	21,32,35,37	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	K	B	417	1/1	0.97	0.09	-1.11	38,38,38,38	0
3	K	A	417	1/1	0.99	0.06	-3.85	25,25,25,25	0

6.5 Other polymers (i)

There are no such residues in this entry.