Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 01:01 PM GMT

PDB ID	:	3SI8
Title	:	Human DNA polymerase eta - DNA ternary complex with the 5'T of a CPD in the active site (TT2)
Authors	:	Biertumpfel, C.; Zhao, Y.; Kondo, Y.; Ramon-Maiques, S.; Gregory, M.; Lee, J.Y.; Masutani, C.; Lehmann, A.R.; Hanaoka, F.; Yang, W.
Deposited on	:	2011-06-17
Resolution	:	2.15 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:

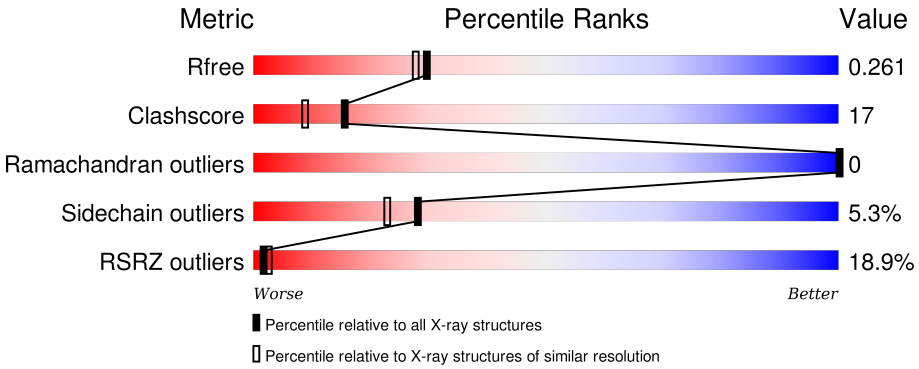
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.15 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	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	435	18%, 70%, 25%
2	T	12	17%, 33%, 50%, 17%
3	P	9	11%, 67%, 33%

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	437	-	-	-	X
5	GOL	A	441	-	-	X	X
5	GOL	A	442	-	-	-	X
6	3D1	A	438	-	-	-	X
7	DTT	A	444	-	-	-	X
9	MG	A	450	-	-	-	X

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 3972 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 DNA polymerase eta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3309	2077	596	613	23			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q9Y253
A	-1	PRO	-	EXPRESSION TAG	UNP Q9Y253
A	0	HIS	-	EXPRESSION TAG	UNP Q9Y253

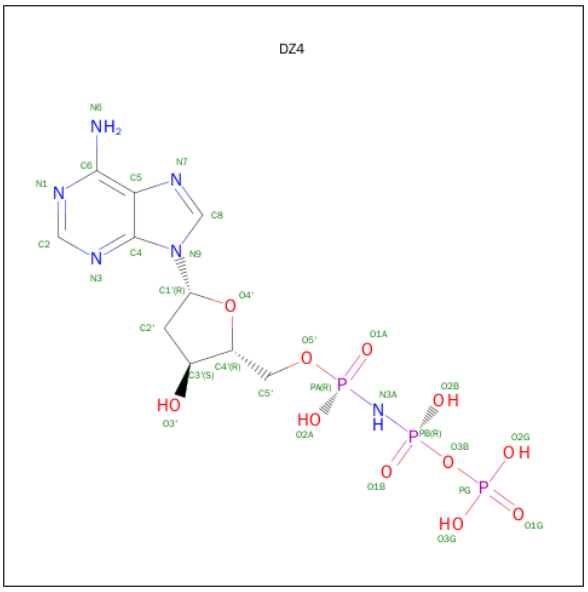
- Molecule 2 is a DNA chain called DNA (5'-D(*TP*A*AP*CP*(TTD)P*AP*TP*GP*AP*CP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	10	Total	C	N	O	P	0	0	0
			221	107	40	64	10			

- Molecule 3 is a DNA chain called DNA (5'-D(*TP*GP*CP*GP*TP*CP*AP*TP*A)-3').

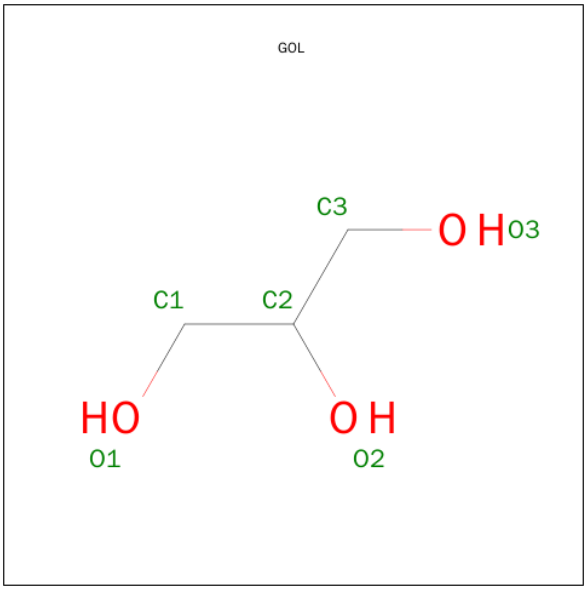
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	9	Total	C	N	O	P	0	0	0
			181	88	32	53	8			

- Molecule 4 is 2'-DEOXY-5'-O-[(R)-HYDROXY{[(R)-HYDROXY(PHOSPHONOOXY)PHOSPHORYL]AMINO}PHOSPHORYL]ADENOSINE (three-letter code: DZ4) (formula: $C_{10}H_{17}N_6O_{11}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			30	10	6	11	3		

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



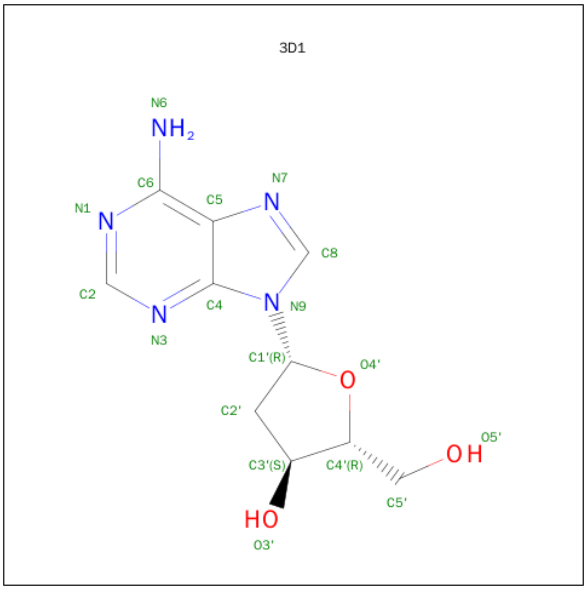
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

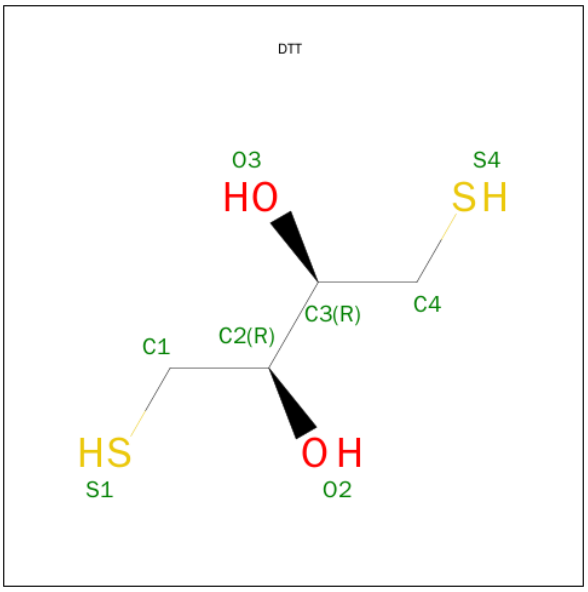
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0

- Molecule 6 is (2R,3S,5R)-5-(6-AMINO-9H-PURIN-9-YL)-TETRAHYDRO-2-(HYDROXY METHYL)FURAN-3-OL (three-letter code: 3D1) (formula: $C_{10}H_{13}N_5O_3$).



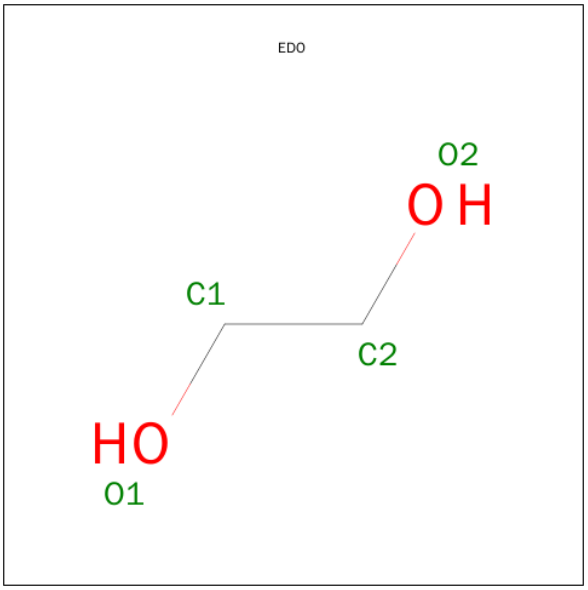
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O 18 10 5 3	0	0

- Molecule 7 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula: $C_4H_{10}O_2S_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O S 8 4 2 2	0	0
7	A	1	Total C O S 8 4 2 2	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	2	Total Mg 2 2	0	0

- Molecule 10 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	1	Total Co 1 1	0	0

- Molecule 11 is water.

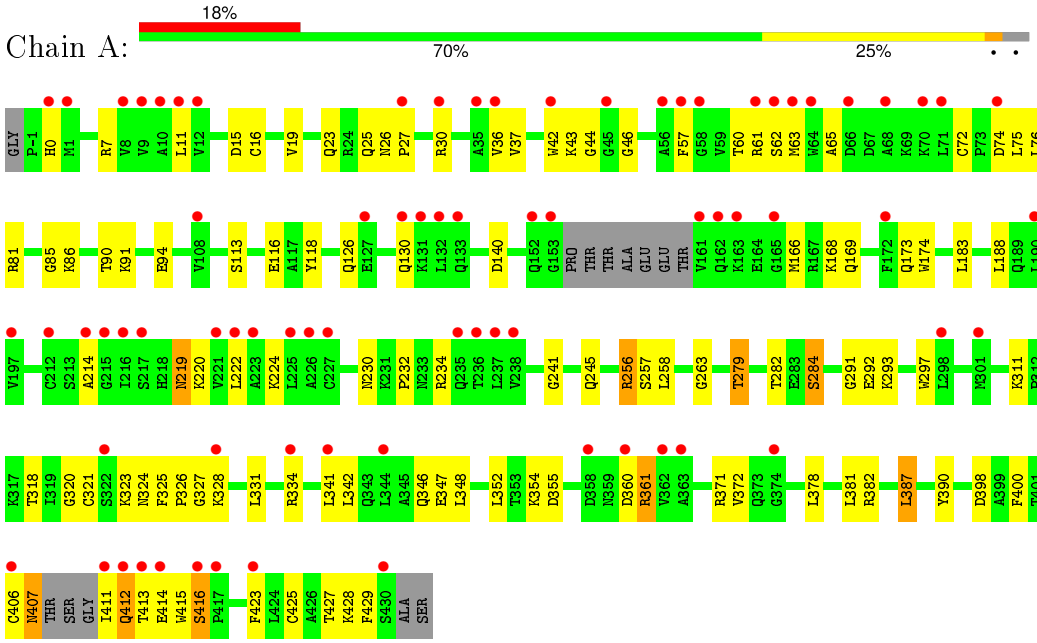
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	112	Total O 112 112	0	0
11	T	7	Total O 7 7	0	0
11	P	7	Total O 7 7	0	0

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.

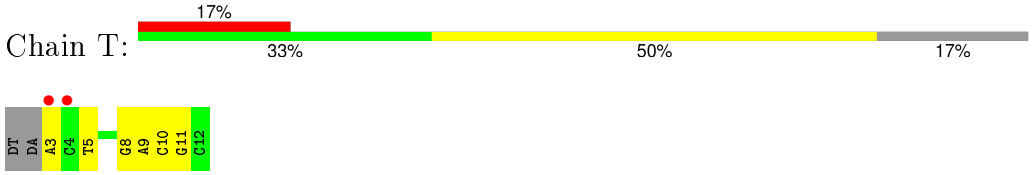
- Molecule 1: DNA polymerase eta

Chain A:



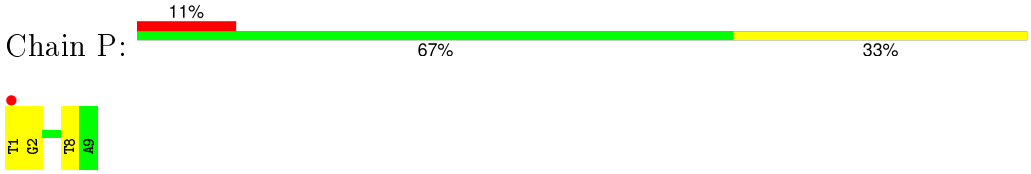
- Molecule 2: DNA (5'-D(*TP*A*AP*CP*(TTD)P*AP*TP*GP*AP*CP*GP*C)-3')

Chain T:



- Molecule 3: DNA (5'-D(*TP*GP*CP*GP*TP*CP*AP*TP*A)-3')

Chain P:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.24Å 80.26Å 139.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.28 – 2.15 26.28 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.6 (26.28-2.15) 98.7 (26.28-2.15)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.83 (at 2.15Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.241 , 0.258 0.243 , 0.261	Depositor DCC
R_{free} test set	2896 reflections (7.68%)	DCC
Wilson B-factor (Å^2)	49.4	Xtriage
Anisotropy	0.632	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 67.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 37836 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3972	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% 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: GOL, MG, CO, 3D1, EDO, DZ4, DTT, TTD

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	0.37	0/3366	0.63	0/4542
2	T	0.40	0/202	0.84	0/307
3	P	0.36	0/202	0.70	0/310
All	All	0.37	0/3770	0.65	0/5159

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	3309	0	3338	124	0
2	T	221	0	126	12	0
3	P	181	0	103	2	0
4	A	30	0	12	0	0
5	A	48	0	62	13	0
6	A	18	0	12	3	0
7	A	16	0	20	5	0
8	A	20	0	30	1	0
9	A	2	0	0	0	0
10	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	A	112	0	0	11	0
11	P	7	0	0	0	0
11	T	7	0	0	1	0
All	All	3972	0	3703	130	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 17.

All (130) 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:27:PRO:HA	1:A:30:ARG:HE	1.19	1.03
1:A:25:GLN:HE22	1:A:91:LYS:HD2	1.33	0.90
1:A:284:SER:HB3	5:A:435:GOL:H31	1.53	0.89
1:A:168:LYS:HZ3	5:A:441:GOL:H31	1.40	0.86
1:A:168:LYS:NZ	5:A:441:GOL:H31	1.90	0.84
1:A:292:GLU:H	7:A:443:DTT:H12	1.43	0.83
1:A:27:PRO:HG3	1:A:30:ARG:HH21	1.46	0.79
1:A:326:PRO:HB2	2:T:3:DA:H3'	1.65	0.79
1:A:25:GLN:NE2	1:A:91:LYS:HD2	1.97	0.78
1:A:168:LYS:HZ1	5:A:441:GOL:H11	1.48	0.78
1:A:342:LEU:O	1:A:346:GLN:HG3	1.85	0.77
1:A:27:PRO:HA	1:A:30:ARG:NE	2.00	0.76
1:A:27:PRO:HG3	1:A:30:ARG:NH2	2.02	0.74
1:A:326:PRO:HB2	2:T:3:DA:C3'	2.18	0.73
1:A:241:GLY:HA2	5:A:434:GOL:H11	1.70	0.73
1:A:60:THR:O	1:A:63:MET:HG2	1.89	0.73
1:A:354:LYS:HE2	11:A:550:HOH:O	1.89	0.72
1:A:361:ARG:HD3	1:A:429:PHE:CD1	2.25	0.71
1:A:219:ASN:ND2	1:A:222:LEU:H	1.89	0.70
1:A:323:LYS:HE3	11:A:482:HOH:O	1.91	0.70
1:A:292:GLU:N	7:A:443:DTT:H12	2.07	0.70
1:A:334:ARG:HH22	1:A:404:LYS:CE	2.05	0.69
1:A:90:THR:O	1:A:94:GLU:HG3	1.94	0.68
1:A:19:VAL:O	1:A:23:GLN:HG3	1.94	0.67
1:A:25:GLN:HE22	1:A:91:LYS:CD	2.05	0.67
1:A:334:ARG:HH22	1:A:404:LYS:NZ	1.94	0.65
1:A:326:PRO:HB2	2:T:3:DA:C2'	2.27	0.65
1:A:74:ASP:O	1:A:76:LEU:HD12	1.97	0.65
1:A:323:LYS:HG2	1:A:325:PHE:CZ	2.34	0.62
1:A:352:LEU:HB3	1:A:390:TYR:CE1	2.35	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:61:ARG:O	1:A:62:SER:HB3	2.00	0.61
1:A:284:SER:HB3	5:A:435:GOL:C3	2.29	0.61
1:A:113:SER:OG	1:A:116:GLU:HG2	2.01	0.61
2:T:5:TTD:H2'	2:T:5:TTD:O5R	2.01	0.60
1:A:86:LYS:HB2	11:T:86:HOH:O	2.03	0.58
5:A:442:GOL:H11	8:A:447:EDO:O1	2.03	0.57
1:A:324:ASN:OD1	1:A:371:ARG:NH2	2.35	0.57
1:A:341:LEU:HD11	1:A:403:ILE:HD11	1.87	0.56
1:A:85:GLY:O	1:A:323:LYS:NZ	2.38	0.56
1:A:297:TRP:CD1	6:A:438:3D1:H2'1	2.41	0.55
1:A:411:ILE:HG13	1:A:413:THR:HG22	1.87	0.55
1:A:126:GLN:NE2	1:A:126:GLN:HA	2.21	0.55
1:A:360:ASP:OD1	1:A:360:ASP:O	2.24	0.55
1:A:36:VAL:HA	1:A:78:ALA:O	2.06	0.55
1:A:15:ASP:O	1:A:16:CYS:C	2.45	0.54
1:A:411:ILE:O	1:A:412:GLN:HB2	2.06	0.54
1:A:230:ASN:O	1:A:234:ARG:HB2	2.08	0.54
1:A:311:LYS:HE3	11:A:485:HOH:O	2.07	0.53
1:A:313:ARG:NH2	2:T:9:DA:OP1	2.42	0.53
1:A:371:ARG:NH1	1:A:378:LEU:O	2.41	0.53
1:A:411:ILE:N	1:A:414:GLU:O	2.42	0.53
1:A:257:SER:N	3:P:8:DT:OP1	2.42	0.53
1:A:37:VAL:HA	1:A:46:GLY:O	2.09	0.52
1:A:320:GLY:HA2	1:A:348:LEU:HD11	1.90	0.52
2:T:10:DC:H2"	2:T:11:DG:C8	2.45	0.52
1:A:174:TRP:HZ3	7:A:444:DTT:S4	2.33	0.52
1:A:11:LEU:HD22	1:A:224:LYS:HA	1.90	0.51
1:A:355:ASP:OD2	1:A:361:ARG:HD2	2.10	0.51
1:A:219:ASN:HD21	1:A:222:LEU:H	1.58	0.51
1:A:37:VAL:HG23	1:A:79:GLN:HG3	1.91	0.51
1:A:411:ILE:O	1:A:412:GLN:CB	2.59	0.51
1:A:279:THR:HG22	11:A:505:HOH:O	2.11	0.51
1:A:168:LYS:HZ1	5:A:441:GOL:H31	1.74	0.50
1:A:25:GLN:OE1	1:A:91:LYS:HD2	2.12	0.50
1:A:327:GLY:HA3	2:T:3:DA:H5'	1.94	0.50
1:A:423:PHE:CE1	1:A:425:CYS:HB2	2.46	0.50
1:A:423:PHE:CZ	1:A:425:CYS:HB2	2.47	0.50
1:A:293:LYS:HE3	6:A:438:3D1:O4'	2.10	0.50
1:A:11:LEU:CD2	1:A:224:LYS:HA	2.42	0.50
1:A:411:ILE:CG2	1:A:416:SER:OG	2.61	0.49
1:A:387:LEU:HD12	1:A:387:LEU:C	2.33	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:42:TRP:O	1:A:43:LYS:HB2	2.12	0.49
1:A:398:ASP:O	1:A:402:VAL:HG13	2.11	0.49
1:A:25:GLN:CD	1:A:91:LYS:HD2	2.32	0.49
1:A:140:ASP:OD2	5:A:441:GOL:O3	2.30	0.49
1:A:130:GLN:NE2	11:A:528:HOH:O	2.45	0.48
1:A:25:GLN:HE22	1:A:91:LYS:CG	2.27	0.48
1:A:174:TRP:HZ3	7:A:444:DTT:HS2	1.58	0.48
1:A:297:TRP:HD1	6:A:438:3D1:H2'1	1.77	0.48
1:A:411:ILE:O	1:A:411:ILE:HG13	2.13	0.47
1:A:166:MET:HE2	1:A:166:MET:HA	1.96	0.47
1:A:407:ASN:N	11:A:500:HOH:O	2.43	0.47
2:T:5:TTD:H2'	2:T:5:TTD:O4R	2.14	0.46
1:A:7:ARG:HD3	11:A:465:HOH:O	2.15	0.46
1:A:334:ARG:HH22	1:A:404:LYS:HE3	1.77	0.46
1:A:72:CYS:O	1:A:75:LEU:HB2	2.14	0.46
1:A:361:ARG:HD3	1:A:429:PHE:HD1	1.77	0.46
1:A:11:LEU:O	1:A:214:ALA:HA	2.16	0.45
1:A:318:THR:HG22	1:A:427:THR:OG1	2.15	0.45
1:A:355:ASP:OD2	1:A:361:ARG:NE	2.50	0.45
1:A:355:ASP:OD2	1:A:361:ARG:CD	2.65	0.45
1:A:328:LYS:NZ	11:A:546:HOH:O	2.50	0.45
2:T:5:TTD:H1'	2:T:5:TTD:O4R	2.16	0.45
1:A:245:GLN:HB2	11:A:456:HOH:O	2.17	0.45
1:A:428:LYS:HE2	1:A:428:LYS:HB3	1.74	0.45
1:A:315:LEU:HB2	11:A:481:HOH:O	2.17	0.45
1:A:282:THR:HB	5:A:435:GOL:H2	1.99	0.44
1:A:44:GLY:HA3	1:A:65:ALA:HB3	2.00	0.44
1:A:11:LEU:HD21	5:A:437:GOL:H12	1.99	0.44
1:A:37:VAL:HG11	1:A:65:ALA:HB2	2.00	0.44
1:A:11:LEU:HD13	1:A:118:TYR:CE2	2.53	0.44
1:A:258:LEU:O	1:A:263:GLY:HA3	2.18	0.44
1:A:355:ASP:CG	1:A:361:ARG:HE	2.20	0.43
1:A:118:TYR:CD2	1:A:220:LYS:HG2	2.53	0.43
1:A:57:PHE:CB	1:A:72:CYS:HB2	2.48	0.43
1:A:411:ILE:HG23	1:A:416:SER:HB3	2.01	0.43
1:A:256:ARG:HE	1:A:257:SER:HB3	1.84	0.43
1:A:406:CYS:O	1:A:407:ASN:CB	2.66	0.43
1:A:404:LYS:HE2	11:A:514:HOH:O	2.18	0.43
1:A:126:GLN:HE21	1:A:126:GLN:HA	1.84	0.42
1:A:407:ASN:OD1	1:A:407:ASN:N	2.51	0.42
1:A:26:ASN:C	1:A:26:ASN:OD1	2.57	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:P:1:DT:H2"	3:P:2:DG:C8	2.54	0.42
1:A:168:LYS:HZ1	5:A:441:GOL:C1	2.26	0.42
1:A:326:PRO:CB	2:T:3:DA:H3'	2.45	0.42
1:A:168:LYS:NZ	5:A:441:GOL:C3	2.74	0.42
1:A:74:ASP:O	1:A:76:LEU:CD1	2.68	0.42
1:A:291:GLY:HA2	7:A:443:DTT:H12	2.01	0.41
1:A:326:PRO:HB2	2:T:3:DA:H2'	1.98	0.41
1:A:334:ARG:HG2	1:A:400:PHE:HZ	1.84	0.41
1:A:86:LYS:HE3	1:A:321:CYS:HB2	2.03	0.41
1:A:381:LEU:HD12	1:A:381:LEU:C	2.41	0.41
1:A:86:LYS:HD3	1:A:347:GLU:OE1	2.20	0.41
1:A:316:PRO:HA	2:T:8:DG:OP1	2.21	0.41
1:A:219:ASN:HD22	1:A:219:ASN:C	2.23	0.40
1:A:411:ILE:HG13	1:A:413:THR:CG2	2.51	0.40
1:A:406:CYS:O	1:A:407:ASN:HB2	2.21	0.40
1:A:169:GLN:O	1:A:173:GLN:HG3	2.21	0.40
1:A:334:ARG:HG3	1:A:334:ARG:HH11	1.86	0.40
1:A:318:THR:HG22	1:A:427:THR:HG1	1.87	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	416/435 (96%)	406 (98%)	10 (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	360/372 (97%)	341 (95%)	19 (5%)	28 23

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	HIS
1	A	81	ARG
1	A	183	LEU
1	A	188	LEU
1	A	219	ASN
1	A	232	PRO
1	A	256	ARG
1	A	279	THR
1	A	284	SER
1	A	331	LEU
1	A	361	ARG
1	A	372	VAL
1	A	382	ARG
1	A	387	LEU
1	A	402	VAL
1	A	407	ASN
1	A	412	GLN
1	A	415	TRP
1	A	416	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	126	GLN
1	A	169	GLN
1	A	219	ASN
1	A	287	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	TTD	T	5	2	40,45,46	1.30	4 (10%)	58,74,77	3.24	19 (32%)

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	TTD	T	5	2	-	0/22/109/110	0/3/6/6

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	5	TTD	C5-C4	-3.65	1.44	1.51
2	T	5	TTD	C5T-C4T	-3.06	1.45	1.51
2	T	5	TTD	C5T-C6T	3.72	1.59	1.55
2	T	5	TTD	C1R-N1T	4.18	1.51	1.45

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	5	TTD	C5T-C6T-N1T	-6.77	106.12	115.70
2	T	5	TTD	O4T-C4T-C5T	-6.57	117.56	122.92
2	T	5	TTD	C5T-C6T-C6	-5.30	80.31	89.27
2	T	5	TTD	C4T-N3T-C2T	-5.27	117.35	126.84
2	T	5	TTD	C5T-C5-C6	-4.93	82.13	88.37
2	T	5	TTD	O4-C4-C5	-4.42	119.32	122.92
2	T	5	TTD	C5-C5T-C4T	-3.47	101.27	113.09
2	T	5	TTD	O2-C2-N1	-2.64	119.38	123.36
2	T	5	TTD	O4T-C4T-N3T	-2.62	116.07	120.50
2	T	5	TTD	C6-C5-C4	-2.10	108.28	114.27

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	5	TTD	N3T-C2T-N1T	3.41	120.23	116.82
2	T	5	TTD	N3-C2-N1	3.54	120.37	116.82
2	T	5	TTD	C5-C6-C6T	3.96	95.97	89.27
2	T	5	TTD	C5-C5T-C6T	4.31	93.82	88.37
2	T	5	TTD	C5T-C5-C4	4.35	127.90	113.09
2	T	5	TTD	C5M-C5T-C4T	5.30	115.98	108.41
2	T	5	TTD	C5-C4-N3	6.92	122.02	116.01
2	T	5	TTD	C5-C6-N1	8.48	127.71	115.70
2	T	5	TTD	C5T-C4T-N3T	11.93	126.37	116.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	5	TTD	3	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 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	DZ4	A	433	9	27,32,32	5.92	11 (40%)	32,50,50	2.56	9 (28%)
5	GOL	A	434	-	5,5,5	0.21	0	5,5,5	0.26	0
5	GOL	A	435	-	5,5,5	0.30	0	5,5,5	0.16	0
5	GOL	A	436	-	5,5,5	0.27	0	5,5,5	0.28	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	A	437	-	5,5,5	0.18	0	5,5,5	0.24	0
6	3D1	A	438	-	16,20,20	1.72	3 (18%)	17,29,29	3.66	10 (58%)
5	GOL	A	439	-	5,5,5	0.27	0	5,5,5	0.30	0
5	GOL	A	440	-	5,5,5	1.45	1 (20%)	5,5,5	0.68	0
5	GOL	A	441	-	5,5,5	0.17	0	5,5,5	0.22	0
5	GOL	A	442	-	5,5,5	1.35	1 (20%)	5,5,5	0.58	0
7	DTT	A	443	-	7,7,7	1.54	2 (28%)	4,8,8	2.16	2 (50%)
7	DTT	A	444	-	7,7,7	1.23	1 (14%)	4,8,8	1.99	1 (25%)
8	EDO	A	445	-	3,3,3	0.47	0	2,2,2	0.44	0
8	EDO	A	446	-	3,3,3	0.45	0	2,2,2	0.40	0
8	EDO	A	447	-	3,3,3	0.45	0	2,2,2	0.48	0
8	EDO	A	448	-	3,3,3	0.63	0	2,2,2	0.23	0
8	EDO	A	449	-	3,3,3	0.46	0	2,2,2	0.42	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	DZ4	A	433	9	-	0/13/34/34	0/3/3/3
5	GOL	A	434	-	-	0/4/4/4	0/0/0/0
5	GOL	A	435	-	-	0/4/4/4	0/0/0/0
5	GOL	A	436	-	-	0/4/4/4	0/0/0/0
5	GOL	A	437	-	-	0/4/4/4	0/0/0/0
6	3D1	A	438	-	-	0/2/18/18	0/3/3/3
5	GOL	A	439	-	-	0/4/4/4	0/0/0/0
5	GOL	A	440	-	-	0/4/4/4	0/0/0/0
5	GOL	A	441	-	-	0/4/4/4	0/0/0/0
5	GOL	A	442	-	-	0/4/4/4	0/0/0/0
7	DTT	A	443	-	-	0/8/8/8	0/0/0/0
7	DTT	A	444	-	-	0/8/8/8	0/0/0/0
8	EDO	A	445	-	-	0/1/1/1	0/0/0/0
8	EDO	A	446	-	-	0/1/1/1	0/0/0/0
8	EDO	A	447	-	-	0/1/1/1	0/0/0/0
8	EDO	A	448	-	-	0/1/1/1	0/0/0/0
8	EDO	A	449	-	-	0/1/1/1	0/0/0/0

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	433	DZ4	PB-O1B	-17.76	1.25	1.46
4	A	433	DZ4	PA-O1A	-16.36	1.27	1.46
4	A	433	DZ4	PB-O2B	-11.24	1.25	1.56
4	A	433	DZ4	PA-O2A	-6.68	1.38	1.56
4	A	433	DZ4	PG-O2G	-5.88	1.33	1.54
4	A	433	DZ4	PG-O1G	-5.84	1.31	1.51
4	A	433	DZ4	PG-O3G	-5.39	1.35	1.54
4	A	433	DZ4	PB-O3B	-5.26	1.52	1.59
4	A	433	DZ4	O3'-C3'	-5.07	1.32	1.43
6	A	438	3D1	O3'-C3'	-4.42	1.33	1.43
4	A	433	DZ4	O4'-C1'	-3.33	1.34	1.42
6	A	438	3D1	O4'-C1'	-3.30	1.34	1.42
6	A	438	3D1	O4'-C4'	-3.09	1.37	1.45
5	A	440	GOL	O2-C2	-3.06	1.34	1.43
4	A	433	DZ4	O4'-C4'	-3.01	1.38	1.45
5	A	442	GOL	O2-C2	-2.97	1.34	1.43
7	A	444	DTT	O3-C3	2.12	1.48	1.43
7	A	443	DTT	O2-C2	2.40	1.48	1.43
7	A	443	DTT	O3-C3	2.43	1.48	1.43

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	438	3D1	N3-C2-N1	-10.63	120.76	128.89
4	A	433	DZ4	N3-C2-N1	-9.74	121.44	128.89
4	A	433	DZ4	O1B-PB-N3A	-3.76	106.14	111.90
7	A	444	DTT	C3-C4-S4	-3.48	108.14	113.91
7	A	443	DTT	C3-C4-S4	-3.25	108.52	113.91
6	A	438	3D1	C1'-N9-C4	-3.06	121.97	127.16
7	A	443	DTT	C2-C1-S1	-2.85	109.19	113.91
4	A	433	DZ4	C2'-C1'-N9	-2.55	107.95	114.16
6	A	438	3D1	C4-C5-N7	-2.24	107.42	109.48
6	A	438	3D1	O3'-C3'-C2'	2.01	117.39	110.74
6	A	438	3D1	O5'-C5'-C4'	2.13	118.37	111.33
6	A	438	3D1	C2-N1-C6	2.14	122.59	118.77
6	A	438	3D1	O4'-C4'-C3'	2.19	111.18	105.67
6	A	438	3D1	C3'-C2'-C1'	2.20	107.69	102.40
4	A	433	DZ4	C3'-C2'-C1'	2.45	108.30	102.40
4	A	433	DZ4	O4'-C4'-C3'	2.46	111.86	105.67
4	A	433	DZ4	O3'-C3'-C2'	2.71	119.72	110.74
4	A	433	DZ4	O5'-C5'-C4'	2.86	119.65	109.12
4	A	433	DZ4	O4'-C1'-N9	3.46	113.71	107.72
6	A	438	3D1	O4'-C4'-C5'	4.00	117.84	109.17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	433	DZ4	O2A-PA-O1A	4.72	119.86	110.00
6	A	438	3D1	O4'-C1'-N9	6.86	119.60	107.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	434	GOL	1	0
5	A	435	GOL	3	0
5	A	437	GOL	1	0
6	A	438	3D1	3	0
5	A	441	GOL	7	0
5	A	442	GOL	1	0
7	A	443	DTT	3	0
7	A	444	DTT	2	0
8	A	447	EDO	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	422/435 (97%)	0.99	80 (18%) 2 2	40, 60, 101, 111	0
2	T	9/12 (75%)	1.58	2 (22%) 1 2	65, 75, 127, 132	0
3	P	9/9 (100%)	0.20	1 (11%) 7 12	51, 60, 92, 116	0
All	All	440/456 (96%)	0.98	83 (18%) 2 2	40, 61, 102, 132	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	161	VAL	7.9
2	T	3	DA	7.7
1	A	42	TRP	7.0
1	A	216	ILE	5.7
1	A	153	GLY	5.6
1	A	315	LEU	5.4
1	A	190	LEU	5.0
1	A	132	LEU	4.5
1	A	217	SER	4.5
1	A	133	GLN	4.4
1	A	417	PRO	4.4
1	A	413	THR	4.3
2	T	4	DC	4.3
1	A	162	GLN	4.2
1	A	222	LEU	4.2
1	A	152	GLN	4.2
1	A	108	VAL	4.1
1	A	328	LYS	4.1
1	A	412	GLN	4.1
1	A	45	GLY	4.0
1	A	71	LEU	3.9
1	A	214	ALA	3.9
1	A	9	VAL	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	223	ALA	3.8
1	A	411	ILE	3.7
1	A	74	ASP	3.7
1	A	62	SER	3.7
1	A	58	GLY	3.5
1	A	238	VAL	3.5
1	A	237	LEU	3.4
1	A	194	ALA	3.3
1	A	197	VAL	3.2
3	P	1	DT	3.2
1	A	314	GLN	3.2
1	A	374	GLY	3.2
1	A	221	VAL	3.2
1	A	10	ALA	3.2
1	A	68	ALA	3.1
1	A	36	VAL	3.1
1	A	11	LEU	3.1
1	A	358	ASP	3.1
1	A	416	SER	3.0
1	A	70	LYS	2.9
1	A	172	PHE	2.9
1	A	215	GLY	2.9
1	A	0	HIS	2.8
1	A	30	ARG	2.8
1	A	236	THR	2.8
1	A	226	ALA	2.8
1	A	414	GLU	2.8
1	A	57	PHE	2.7
1	A	363	ALA	2.7
1	A	131	LYS	2.7
1	A	27	PRO	2.6
1	A	360	ASP	2.6
1	A	406	CYS	2.5
1	A	341	LEU	2.5
1	A	227	CYS	2.5
1	A	63	MET	2.5
1	A	163	LYS	2.5
1	A	225	LEU	2.4
1	A	423	PHE	2.4
1	A	56	ALA	2.4
1	A	1	MET	2.3
1	A	193	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	212	CYS	2.3
1	A	8	VAL	2.3
1	A	12	VAL	2.3
1	A	64	TRP	2.3
1	A	344	LEU	2.3
1	A	334	ARG	2.2
1	A	165	GLY	2.2
1	A	430	SER	2.2
1	A	301	MET	2.2
1	A	235	GLN	2.1
1	A	298	LEU	2.1
1	A	322	SER	2.1
1	A	130	GLN	2.1
1	A	35	ALA	2.1
1	A	127	GLU	2.1
1	A	66	ASP	2.1
1	A	61	ARG	2.0
1	A	362	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	TTD	T	5	40/41	0.86	0.16	-	68,75,112,113	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
6	3D1	A	438	18/18	0.45	0.58	9.82	98,101,106,107	0
5	GOL	A	442	6/6	0.84	0.31	3.28	104,105,106,107	0
5	GOL	A	441	6/6	0.73	0.40	2.95	96,96,97,98	0
7	DTT	A	444	8/8	0.56	0.40	2.74	110,111,113,117	0
9	MG	A	450	1/1	0.97	0.19	2.59	40,40,40,40	0
5	GOL	A	437	6/6	0.82	0.30	2.04	76,77,77,78	0
5	GOL	A	440	6/6	0.75	0.28	1.66	99,100,101,101	0
7	DTT	A	443	8/8	0.84	0.21	1.49	101,102,104,108	0
8	EDO	A	446	4/4	0.92	0.18	0.26	72,73,73,78	0
10	CO	A	452	1/1	0.99	0.14	-0.57	43,43,43,43	0
4	DZ4	A	433	30/30	0.97	0.12	-0.68	41,51,54,57	0
5	GOL	A	439	6/6	0.95	0.11	-1.34	54,57,62,63	0
9	MG	A	451	1/1	0.98	0.13	-	40,40,40,40	0
8	EDO	A	448	4/4	0.64	0.30	-	80,82,82,83	0
5	GOL	A	434	6/6	0.65	0.55	-	108,108,108,109	0
8	EDO	A	445	4/4	0.89	0.19	-	89,89,89,90	0
8	EDO	A	447	4/4	0.85	0.14	-	86,86,87,88	0
8	EDO	A	449	4/4	0.77	0.18	-	94,94,94,96	0
5	GOL	A	436	6/6	0.61	0.46	-	99,99,100,100	0
5	GOL	A	435	6/6	0.27	0.77	-	109,110,110,110	0

6.5 Other polymers (i)

There are no such residues in this entry.