Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 02:29 AM GMT

PDB ID	:	2HEK
Title	:	Crystal structure of O67745, a hypothetical protein from Aquifex aeolicus at 2.0 A resolution.
Authors	:	Oganesyan, V.; Jancarik, J.; Adams, P.D.; Kim, R.; Kim, S.H.; Berkeley Structural Genomics Center (BSGC)
Deposited on	:	2006-06-21
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:

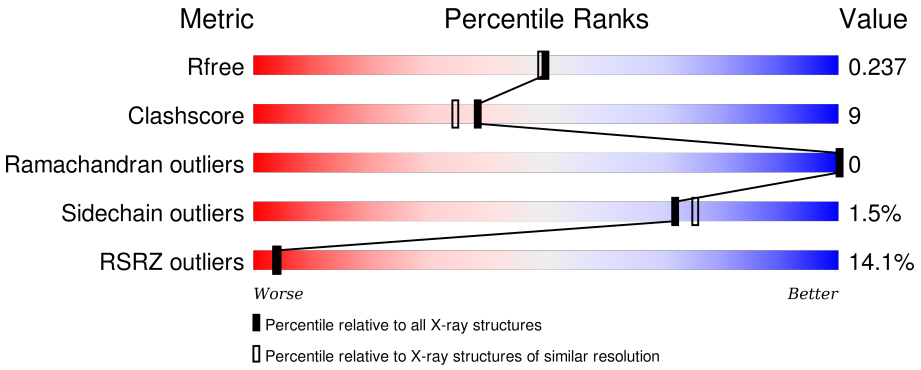
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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	371	19% 85% 13% ••
1	B	371	9% 81% 18% ••

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
2	PO4	B	504	-	-	-	X
6	GDP	A	402	-	-	-	X
7	GOL	A	403	-	-	-	X
7	GOL	A	406	-	-	-	X
7	GOL	A	407	-	-	-	X
7	GOL	A	410	-	-	X	-
7	GOL	B	405	-	-	-	X
7	GOL	B	409	-	-	-	X

2 Entry composition

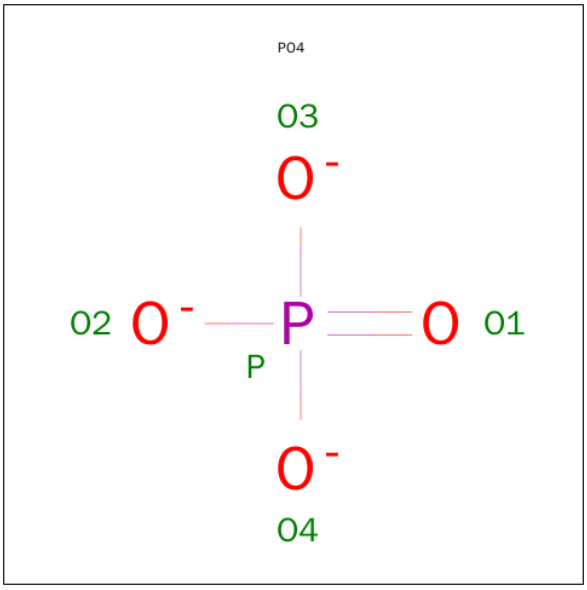
There are 8 unique types of molecules in this entry. The entry contains 6705 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Hypothetical protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	369	Total	C	N	O	S	0	0	0
			3109	2013	526	565	5			
1	B	369	Total	C	N	O	S	0	0	0
			3109	2013	526	565	5			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		
2	B	1	Total	O	P	0	0
			5	4	1		
2	B	1	Total	O	P	0	0
			5	4	1		
2	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

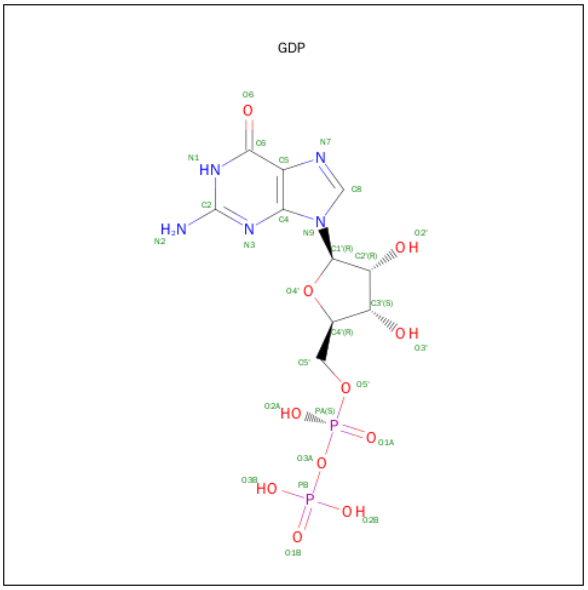
- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	3	Total Cl 3 3	0	0
4	A	2	Total Cl 2 2	0	0

- Molecule 5 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Br 1 1	0	0

- Molecule 6 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



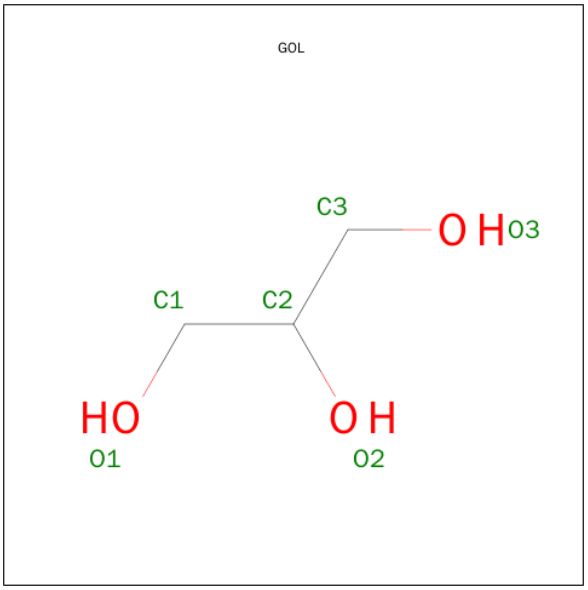
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O P 28 10 5 11 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 8 is water.

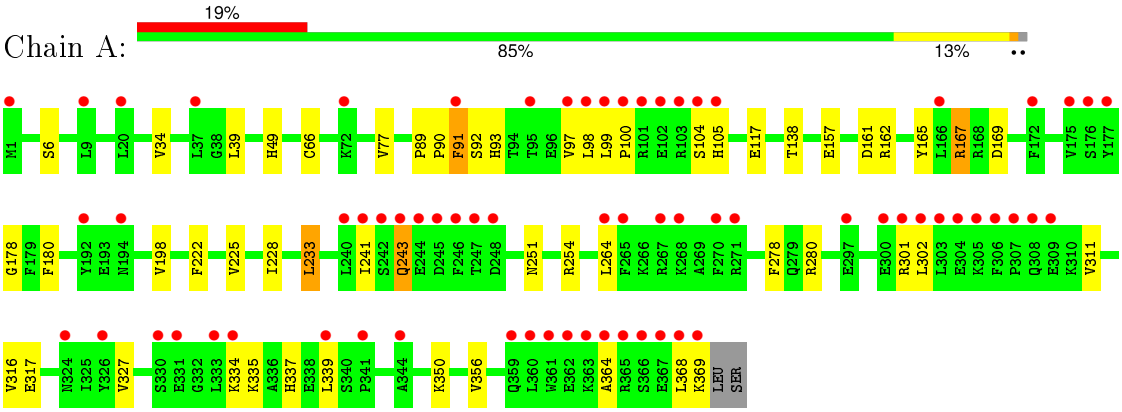
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	179	Total O 179 179	0	0
8	B	176	Total O 176 176	0	0

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.

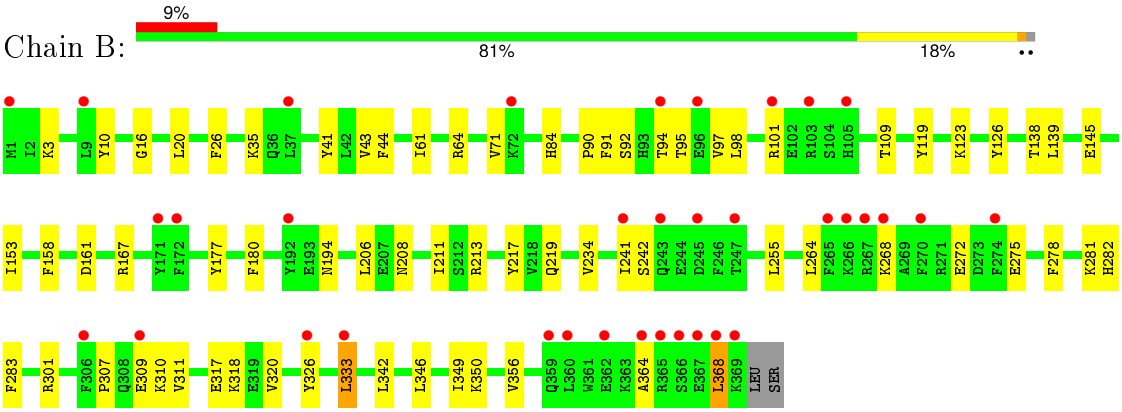
- Molecule 1: Hypothetical protein

Chain A:



- Molecule 1: Hypothetical protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	107.69Å 144.41Å 144.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 2.00 43.19 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.3 (12.00-2.00) 99.3 (43.19-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.175 , 0.220 0.196 , 0.237	Depositor DCC
R_{free} test set	3807 reflections (5.30%)	DCC
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.458	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 61.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Outliers	0 of 75962 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6705	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.03% 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: GDP, GOL, ZN, CL, PO4, BR

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	0.76	1/3178 (0.0%)	0.73	3/4273 (0.1%)
1	B	0.74	0/3178	0.70	0/4273
All	All	0.75	1/6356 (0.0%)	0.72	3/8546 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	167	ARG	CG-CD	6.21	1.67	1.51

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	91	PHE	CB-CG-CD2	6.88	125.61	120.80
1	A	91	PHE	CB-CG-CD1	-6.63	116.16	120.80
1	A	233	LEU	CA-CB-CG	-5.37	102.95	115.30

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	3109	0	3126	64	0
1	B	3109	0	3126	51	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	5	0	0	1	0
2	B	15	0	0	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	0	0
4	B	3	0	0	1	0
5	B	1	0	0	0	0
6	A	56	0	24	3	0
7	A	36	0	48	10	0
7	B	12	0	16	1	0
8	A	179	0	0	1	1
8	B	176	0	0	5	0
All	All	6705	0	6340	116	1

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 (116) 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:167:ARG:NH1	7:A:410:GOL:H11	1.59	1.18
1:A:89:PRO:HG3	1:A:98:LEU:HD12	1.40	1.03
1:B:177:TYR:HD1	1:B:219:GLN:NE2	1.58	1.02
1:A:180:PHE:HB2	7:A:406:GOL:H32	1.45	0.98
1:A:167:ARG:NH1	7:A:410:GOL:C1	2.32	0.93
1:A:167:ARG:HH12	7:A:410:GOL:C1	1.84	0.90
1:A:89:PRO:HB3	1:A:98:LEU:HD11	1.61	0.82
1:A:241:ILE:O	1:A:241:ILE:HG13	1.80	0.82
1:B:177:TYR:HD1	1:B:219:GLN:HE22	0.85	0.81
1:A:167:ARG:HH11	7:A:410:GOL:H11	1.45	0.81
1:B:161:ASP:OD2	2:B:502:PO4:O2	2.03	0.77
1:B:311:VAL:HG12	1:B:356:VAL:HG12	1.71	0.71
1:B:364:ALA:O	1:B:368:LEU:HB2	1.91	0.71
1:A:104:SER:O	1:A:105:HIS:HB2	1.92	0.70
1:A:39:LEU:HG	1:A:91:PHE:CE2	2.29	0.68
1:A:167:ARG:HH12	7:A:410:GOL:H12	1.57	0.68
1:A:89:PRO:CG	1:A:98:LEU:HD12	2.22	0.67
1:A:89:PRO:CB	1:A:98:LEU:HD11	2.25	0.67
1:B:317:GLU:HG2	1:B:350:LYS:HG2	1.76	0.67
1:B:101:ARG:HE	1:B:241:ILE:CD1	2.08	0.66
1:A:39:LEU:HG	1:A:91:PHE:CD2	2.31	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:180:PHE:HB2	7:A:406:GOL:C3	2.21	0.66
1:A:311:VAL:HG12	1:A:356:VAL:HG12	1.78	0.65
1:A:98:LEU:HD23	1:A:241:ILE:HD13	1.79	0.64
1:A:364:ALA:O	1:A:368:LEU:HG	1.98	0.64
1:B:177:TYR:CD1	1:B:219:GLN:NE2	2.47	0.63
7:A:408:GOL:H12	1:B:26:PHE:HD1	1.62	0.63
1:A:89:PRO:HG3	1:A:98:LEU:CD1	2.21	0.63
1:A:222:PHE:CD1	1:A:316:VAL:HG11	2.33	0.63
1:A:161:ASP:OD2	2:A:501:PO4:O4	2.17	0.62
6:A:401:GDP:O1A	1:B:3:LYS:NZ	2.33	0.61
1:B:326:TYR:HB3	1:B:333:LEU:HG	1.83	0.61
1:B:92:SER:O	1:B:94:THR:HG23	2.01	0.61
1:A:89:PRO:HB3	1:A:98:LEU:CD1	2.31	0.60
1:B:95:THR:HG23	8:B:750:HOH:O	2.02	0.60
1:A:49:HIS:CD2	1:A:49:HIS:H	2.18	0.60
1:B:264:LEU:HD23	1:B:278:PHE:HE1	1.67	0.60
1:A:49:HIS:HE1	1:A:169:ASP:OD1	1.85	0.60
1:B:35:LYS:NZ	8:B:667:HOH:O	2.33	0.59
1:B:126:TYR:OH	2:B:504:PO4:O3	2.18	0.59
1:A:39:LEU:CD2	1:A:91:PHE:HE2	2.14	0.59
1:A:39:LEU:CD2	1:A:91:PHE:CE2	2.86	0.58
1:B:64:ARG:NH1	4:B:608:CL:CL	2.74	0.58
1:B:272:GLU:OE2	8:B:712:HOH:O	2.17	0.58
1:A:104:SER:O	1:A:105:HIS:CB	2.53	0.56
1:B:282:HIS:HD2	1:B:283:PHE:O	1.88	0.56
1:B:119:TYR:CE2	1:B:123:LYS:HE3	2.41	0.56
1:B:94:THR:HG22	1:B:234:VAL:HG22	1.87	0.55
1:B:301:ARG:HB2	1:B:368:LEU:HD11	1.87	0.55
1:B:211:ILE:HG13	1:B:342:LEU:HD21	1.87	0.55
1:A:222:PHE:CE1	1:A:316:VAL:HG11	2.43	0.53
1:B:217:TYR:CE2	1:B:318:LYS:HD2	2.44	0.53
6:A:402:GDP:O2A	1:B:282:HIS:HE1	1.92	0.53
1:A:243:GLN:HA	1:A:243:GLN:OE1	2.09	0.53
1:A:222:PHE:CE1	1:A:316:VAL:CG1	2.92	0.52
1:A:251:ASN:OD1	1:A:254:ARG:NH2	2.43	0.52
1:A:117:GLU:HG2	8:A:763:HOH:O	2.08	0.52
1:A:92:SER:O	1:A:93:HIS:HB2	2.10	0.51
1:A:302:LEU:HD13	1:A:368:LEU:HD11	1.93	0.50
1:A:89:PRO:CG	1:A:98:LEU:CD1	2.86	0.50
1:A:89:PRO:CB	1:A:98:LEU:CD1	2.89	0.49
1:B:41:TYR:HA	1:B:44:PHE:O	2.12	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:208:ASN:OD1	7:B:409:GOL:H32	2.12	0.48
1:B:301:ARG:HE	1:B:368:LEU:HD13	1.78	0.48
1:A:91:PHE:CD1	1:A:233:LEU:HD21	2.48	0.48
1:A:317:GLU:HG2	1:A:350:LYS:HG2	1.94	0.48
1:A:98:LEU:HD23	1:A:241:ILE:CD1	2.44	0.48
1:A:66:CYS:SG	1:A:77:VAL:HG11	2.54	0.47
1:B:158:PHE:HE2	1:B:206:LEU:HD13	1.79	0.47
1:A:233:LEU:O	1:A:233:LEU:HG	2.14	0.47
1:A:34:VAL:HG13	1:A:90:PRO:HA	1.95	0.46
1:A:97:VAL:HG12	1:A:241:ILE:HD11	1.97	0.46
6:A:401:GDP:N3	6:A:401:GDP:H2'	2.31	0.46
1:B:320:VAL:HG12	1:B:346:LEU:HD13	1.98	0.46
1:B:16:GLY:O	1:B:20:LEU:HG	2.16	0.46
1:A:39:LEU:HD21	1:A:91:PHE:CE2	2.51	0.45
1:A:92:SER:HB3	1:A:93:HIS:CD2	2.51	0.45
1:A:301:ARG:HG3	1:A:368:LEU:HD22	1.98	0.45
1:A:39:LEU:CG	1:A:91:PHE:CE2	2.97	0.45
1:B:264:LEU:HD23	1:B:278:PHE:CE1	2.49	0.45
1:A:92:SER:CB	1:A:93:HIS:CD2	3.00	0.45
1:B:307:PRO:HG2	1:B:310:LYS:HD3	1.98	0.45
1:B:84:HIS:CD2	1:B:138:THR:O	2.70	0.44
1:A:264:LEU:HD23	1:A:278:PHE:HE1	1.82	0.44
1:A:6:SER:OG	7:A:403:GOL:H32	2.18	0.44
1:B:95:THR:CG2	8:B:750:HOH:O	2.64	0.43
1:A:91:PHE:CE1	1:A:233:LEU:HD11	2.53	0.43
1:B:97:VAL:HG21	1:B:234:VAL:HG13	2.01	0.43
1:A:49:HIS:HD2	1:A:49:HIS:H	1.64	0.42
1:A:49:HIS:CD2	1:A:49:HIS:N	2.87	0.42
1:B:71:VAL:HG21	1:B:153:ILE:HD11	2.01	0.42
1:A:334:LYS:HG3	1:A:339:LEU:HD21	2.01	0.42
1:A:178:GLY:HA3	1:A:180:PHE:CE1	2.54	0.42
1:A:66:CYS:SG	1:A:77:VAL:CG1	3.08	0.42
1:A:301:ARG:CG	1:A:368:LEU:HD22	2.50	0.42
1:B:101:ARG:HG3	1:B:241:ILE:HD12	2.01	0.42
1:B:109:THR:HB	1:B:139:LEU:HD21	2.02	0.42
1:B:275:GLU:HB3	1:B:281:LYS:HG3	2.02	0.41
1:B:10:TYR:CE2	1:B:61:ILE:HD11	2.56	0.41
1:B:317:GLU:HA	1:B:349:ILE:O	2.20	0.41
1:B:90:PRO:HB2	1:B:91:PHE:CD1	2.56	0.41
1:B:213:ARG:CD	8:B:714:HOH:O	2.69	0.41
1:A:157:GLU:O	1:A:162:ARG:HD2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:145:GLU:HB3	1:B:194:ASN:ND2	2.35	0.41
1:A:225:VAL:HA	1:A:228:ILE:HG22	2.03	0.41
1:B:97:VAL:HG12	1:B:241:ILE:CD1	2.51	0.41
1:B:307:PRO:HB2	1:B:309:GLU:OE1	2.21	0.41
1:A:92:SER:HB2	1:A:93:HIS:HD2	1.86	0.41
7:A:403:GOL:H31	1:B:43:VAL:O	2.21	0.41
1:B:167:ARG:HG3	1:B:180:PHE:CE2	2.56	0.40
1:A:39:LEU:HD23	1:A:91:PHE:HE2	1.86	0.40
1:B:84:HIS:HD2	1:B:138:THR:O	2.04	0.40
1:A:198:VAL:O	1:A:327:VAL:HA	2.22	0.40
1:B:101:ARG:CZ	1:B:242:SER:HA	2.51	0.40
1:A:99:LEU:N	1:A:100:PRO:HD2	2.36	0.40
1:A:335:LYS:HB3	1:A:337:HIS:CE1	2.57	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:A:702:HOH:O	8:A:771:HOH:O[3_555]	2.18	0.02

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	367/371 (99%)	356 (97%)	11 (3%)	0	100	100
1	B	367/371 (99%)	358 (98%)	9 (2%)	0	100	100
All	All	734/742 (99%)	714 (97%)	20 (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	342/344 (99%)	337 (98%)	5 (2%)	72 75
1	B	342/344 (99%)	337 (98%)	5 (2%)	72 75
All	All	684/688 (99%)	674 (98%)	10 (2%)	72 75

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	138	THR
1	A	165	TYR
1	A	243	GLN
1	A	280	ARG
1	A	369	LYS
1	B	98	LEU
1	B	255	LEU
1	B	268	LYS
1	B	333	LEU
1	B	368	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	48	GLN
1	A	49	HIS
1	A	84	HIS
1	A	93	HIS
1	A	194	ASN
1	B	48	GLN
1	B	84	HIS
1	B	243	GLN
1	B	282	HIS
1	B	308	GLN
1	B	337	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 22 ligands modelled in this entry, 8 are monoatomic - leaving 14 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	$\#\|Z\| > 2$
6	GDP	A	401	-	23,30,30	1.17	2 (8%)	30,47,47	1.73	6 (20%)
6	GDP	A	402	-	23,30,30	1.28	2 (8%)	30,47,47	1.84	7 (23%)
7	GOL	A	403	-	5,5,5	0.74	0	5,5,5	1.00	0
7	GOL	A	404	-	5,5,5	0.61	0	5,5,5	0.53	0
7	GOL	A	406	-	5,5,5	0.34	0	5,5,5	1.12	0
7	GOL	A	407	-	5,5,5	0.52	0	5,5,5	0.85	0
7	GOL	A	408	-	5,5,5	0.45	0	5,5,5	0.42	0
7	GOL	A	410	-	5,5,5	0.82	0	5,5,5	0.53	0
2	PO4	A	501	3	4,4,4	0.64	0	6,6,6	0.28	0
7	GOL	B	405	-	5,5,5	0.46	0	5,5,5	0.73	0
7	GOL	B	409	-	5,5,5	0.36	0	5,5,5	0.34	0
2	PO4	B	502	3	4,4,4	0.50	0	6,6,6	0.34	0
2	PO4	B	503	-	4,4,4	0.30	0	6,6,6	0.28	0
2	PO4	B	504	-	4,4,4	0.60	0	6,6,6	0.30	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
6	GDP	A	401	-	-	0/12/32/32	0/3/3/3
6	GDP	A	402	-	-	0/12/32/32	0/3/3/3
7	GOL	A	403	-	-	0/4/4/4	0/0/0/0
7	GOL	A	404	-	-	0/4/4/4	0/0/0/0
7	GOL	A	406	-	-	0/4/4/4	0/0/0/0
7	GOL	A	407	-	-	0/4/4/4	0/0/0/0
7	GOL	A	408	-	-	0/4/4/4	0/0/0/0
7	GOL	A	410	-	-	0/4/4/4	0/0/0/0
2	PO4	A	501	3	-	0/0/0/0	0/0/0/0
7	GOL	B	405	-	-	0/4/4/4	0/0/0/0
7	GOL	B	409	-	-	0/4/4/4	0/0/0/0
2	PO4	B	502	3	-	0/0/0/0	0/0/0/0
2	PO4	B	503	-	-	0/0/0/0	0/0/0/0
2	PO4	B	504	-	-	0/0/0/0	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	401	GDP	C5-C4	2.42	1.46	1.40
6	A	401	GDP	C6-C5	3.31	1.47	1.41
6	A	402	GDP	C6-C5	3.69	1.48	1.41
6	A	402	GDP	C5-C4	3.74	1.48	1.40

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	401	GDP	N3-C2-N1	-4.04	121.30	127.44
6	A	401	GDP	C6-C5-C4	-3.90	116.24	120.90
6	A	402	GDP	C5-C6-N1	-3.82	118.36	123.59
6	A	402	GDP	N3-C2-N1	-3.57	122.01	127.44
6	A	402	GDP	C6-C5-C4	-3.51	116.70	120.90
6	A	402	GDP	PA-O3A-PB	-2.83	123.18	132.67
6	A	402	GDP	C4-C5-N7	-2.80	106.91	109.48
6	A	401	GDP	C5-C6-N1	-2.76	119.82	123.59
6	A	401	GDP	O3'-C3'-C4'	-2.35	103.99	111.05
6	A	401	GDP	PA-O3A-PB	-2.05	125.80	132.67
6	A	402	GDP	C4'-O4'-C1'	2.27	112.21	109.72
6	A	401	GDP	C6-N1-C2	4.18	121.74	115.94
6	A	402	GDP	C6-N1-C2	5.18	123.13	115.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	401	GDP	2	0
6	A	402	GDP	1	0
7	A	403	GOL	2	0
7	A	406	GOL	2	0
7	A	408	GOL	1	0
7	A	410	GOL	5	0
2	A	501	PO4	1	0
7	B	409	GOL	1	0
2	B	502	PO4	1	0
2	B	504	PO4	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	369/371 (99%)	1.01	69 (18%) 2 2	22, 41, 89, 111	0
1	B	369/371 (99%)	0.65	35 (9%) 10 11	23, 38, 76, 109	0
All	All	738/742 (99%)	0.83	104 (14%) 4 4	22, 40, 84, 111	0

All (104) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	366	SER	7.6
1	A	91	PHE	6.9
1	A	101	ARG	6.4
1	A	103	ARG	6.4
1	A	364	ALA	6.0
1	A	368	LEU	6.0
1	A	98	LEU	5.7
1	A	302	LEU	5.7
1	A	1	MET	5.2
1	A	265	PHE	5.2
1	A	243	GLN	4.9
1	A	363	LYS	4.9
1	A	104	SER	4.8
1	A	366	SER	4.8
1	A	301	ARG	4.7
1	A	360	LEU	4.7
1	A	305	LYS	4.7
1	A	105	HIS	4.6
1	A	241	ILE	4.5
1	B	268	LYS	4.5
1	B	192	TYR	4.5
1	B	241	ILE	4.3
1	A	192	TYR	4.0
1	A	303	LEU	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	368	LEU	4.0
1	B	103	ARG	4.0
1	A	97	VAL	3.8
1	B	270	PHE	3.6
1	B	360	LEU	3.5
1	B	101	ARG	3.4
1	A	306	PHE	3.3
1	A	365	ARG	3.3
1	A	244	GLU	3.3
1	A	307	PRO	3.3
1	A	308	GLN	3.2
1	B	243	GLN	3.2
1	A	9	LEU	3.2
1	B	306	PHE	3.2
1	B	333	LEU	3.2
1	A	267	ARG	3.1
1	B	172	PHE	3.1
1	A	304	GLU	3.1
1	A	361	TRP	3.1
1	A	37	LEU	3.1
1	B	105	HIS	3.0
1	B	96	GLU	3.0
1	A	242	SER	3.0
1	A	177	TYR	3.0
1	A	99	LEU	3.0
1	A	172	PHE	2.9
1	A	359	GLN	2.9
1	A	245	ASP	2.9
1	A	331	GLU	2.9
1	A	100	PRO	2.8
1	B	265	PHE	2.8
1	B	326	TYR	2.8
1	A	72	LYS	2.8
1	B	266	LYS	2.7
1	A	341	PRO	2.7
1	B	365	ARG	2.7
1	B	367	GLU	2.7
1	A	369	LYS	2.7
1	A	102	GLU	2.7
1	A	246	PHE	2.7
1	B	274	PHE	2.6
1	A	297	GLU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	247	THR	2.6
1	A	334	LYS	2.6
1	A	362	GLU	2.6
1	A	268	LYS	2.6
1	A	339	LEU	2.5
1	B	359	GLN	2.5
1	B	245	ASP	2.4
1	A	324	ASN	2.4
1	A	344	ALA	2.4
1	A	330	SER	2.4
1	A	309	GLU	2.4
1	B	72	LYS	2.4
1	A	270	PHE	2.4
1	A	300	GLU	2.4
1	B	37	LEU	2.4
1	A	367	GLU	2.3
1	B	267	ARG	2.3
1	B	362	GLU	2.3
1	A	176	SER	2.2
1	B	1	MET	2.2
1	A	326	TYR	2.2
1	A	175	VAL	2.2
1	A	240	LEU	2.2
1	B	94	THR	2.2
1	A	20	LEU	2.2
1	A	271	ARG	2.2
1	B	9	LEU	2.2
1	B	247	THR	2.2
1	A	333	LEU	2.1
1	B	171	TYR	2.1
1	A	166	LEU	2.1
1	B	364	ALA	2.1
1	B	369	LYS	2.0
1	A	194	ASN	2.0
1	A	95	THR	2.0
1	A	264	LEU	2.0
1	B	309	GLU	2.0
1	A	248	ASP	2.0

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

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

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
7	GOL	A	406	6/6	0.80	0.28	4.72	50,59,60,60	0
2	PO4	B	504	5/5	0.91	0.45	4.31	61,63,66,67	0
7	GOL	A	407	6/6	0.88	0.30	3.32	35,49,53,54	0
6	GDP	A	402	28/28	0.69	0.26	2.97	67,85,115,115	0
7	GOL	B	409	6/6	0.84	0.21	2.26	49,60,61,63	0
7	GOL	A	403	6/6	0.89	0.20	2.23	39,49,52,55	0
7	GOL	B	405	6/6	0.90	0.22	2.01	34,49,53,56	0
7	GOL	A	404	6/6	0.94	0.19	1.73	33,48,51,52	0
7	GOL	A	410	6/6	0.90	0.20	0.77	53,54,56,57	0
7	GOL	A	408	6/6	0.85	0.20	0.74	70,72,73,74	0
2	PO4	B	503	5/5	0.90	0.16	0.62	90,90,91,91	0
6	GDP	A	401	28/28	0.92	0.14	-0.06	32,42,69,72	0
5	BR	B	603	1/1	0.99	0.09	-0.68	40,40,40,40	1
2	PO4	A	501	5/5	0.97	0.12	-0.86	61,61,66,67	0
4	CL	A	604	1/1	0.99	0.07	-0.96	39,39,39,39	0
2	PO4	B	502	5/5	0.94	0.11	-1.03	54,59,60,63	0
4	CL	A	606	1/1	1.00	0.06	-1.43	40,40,40,40	0
3	ZN	A	602	1/1	0.96	0.10	-1.55	59,59,59,59	1
4	CL	B	605	1/1	0.99	0.09	-1.92	36,36,36,36	0
3	ZN	B	601	1/1	0.99	0.05	-3.27	41,41,41,41	1
4	CL	B	607	1/1	0.99	0.06	-	50,50,50,50	0
4	CL	B	608	1/1	0.98	0.12	-	66,66,66,66	0

6.5 Other polymers (i)

There are no such residues in this entry.