Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2016 – 08:13 PM GMT

PDB ID	:	4XIF
Title	:	Human OGT in complex with UDP-5S-GlcNAc and substrate peptide (keratin-7)
Authors	:	Schimpl, M.; van Aalten, D.M.F.
Deposited on	:	2015-01-06
Resolution	:	3.20 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:

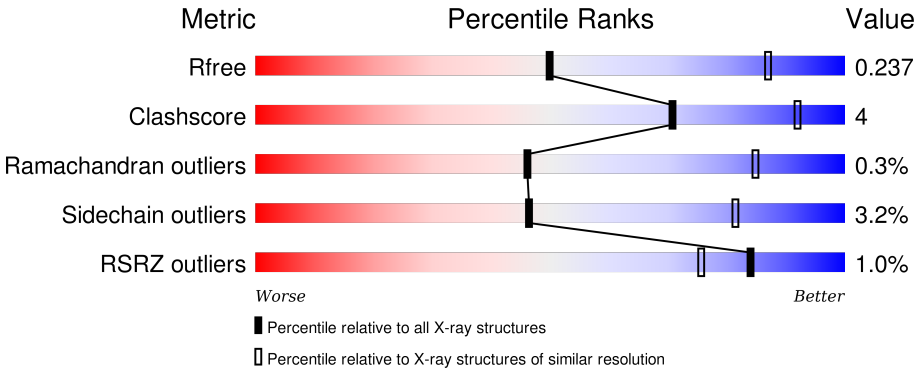
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.20 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.



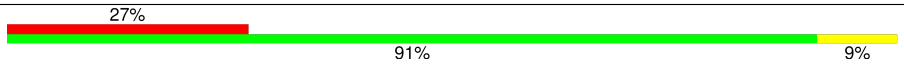
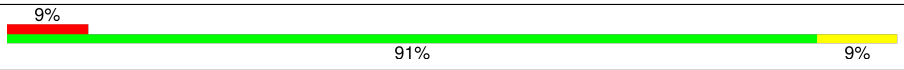
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	723	86% 11%
1	B	723	87% 9%
1	C	723	85% 11%
1	D	723	86% 10%
2	E	11	9% 82% 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	11	
2	G	11	
2	H	11	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 22608 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 UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	702	Total	C	N	O	S	0	0	0
			5535	3511	968	1017	39			
1	B	702	Total	C	N	O	S	0	0	0
			5535	3511	968	1017	39			
1	C	702	Total	C	N	O	S	0	0	0
			5535	3511	968	1017	39			
1	D	702	Total	C	N	O	S	0	0	0
			5535	3511	968	1017	39			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	309	GLY	-	expression tag	UNP O15294
A	310	PRO	-	expression tag	UNP O15294
A	311	GLY	-	expression tag	UNP O15294
A	312	SER	-	expression tag	UNP O15294
B	309	GLY	-	expression tag	UNP O15294
B	310	PRO	-	expression tag	UNP O15294
B	311	GLY	-	expression tag	UNP O15294
B	312	SER	-	expression tag	UNP O15294
C	309	GLY	-	expression tag	UNP O15294
C	310	PRO	-	expression tag	UNP O15294
C	311	GLY	-	expression tag	UNP O15294
C	312	SER	-	expression tag	UNP O15294
D	309	GLY	-	expression tag	UNP O15294
D	310	PRO	-	expression tag	UNP O15294
D	311	GLY	-	expression tag	UNP O15294
D	312	SER	-	expression tag	UNP O15294

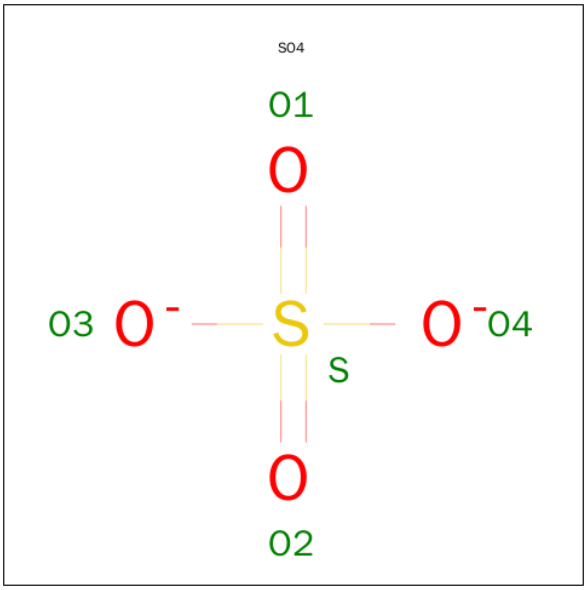
- Molecule 2 is a protein called Keratin, type II cytoskeletal 7.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	11	Total	C	N	O	0	0	0
			73	45	14	14			
2	F	11	Total	C	N	O	0	0	0
			73	45	14	14			
2	G	11	Total	C	N	O	0	0	0
			73	45	14	14			
2	H	11	Total	C	N	O	0	0	0
			73	45	14	14			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	2007	GLY	-	expression tag	UNP P08729
E	2017	GLY	-	expression tag	UNP P08729
F	2007	GLY	-	expression tag	UNP P08729
F	2017	GLY	-	expression tag	UNP P08729
G	2007	GLY	-	expression tag	UNP P08729
G	2017	GLY	-	expression tag	UNP P08729
H	2007	GLY	-	expression tag	UNP P08729
H	2017	GLY	-	expression tag	UNP P08729

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



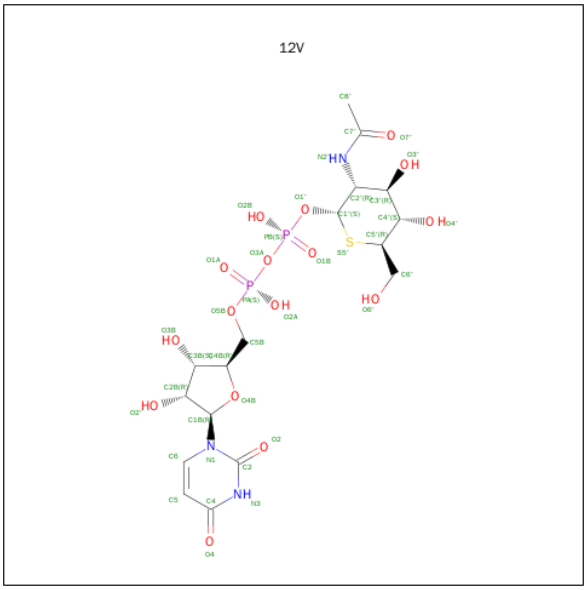
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	H	1	Total O S 5 4 1	0	0

- Molecule 4 is (2S,3R,4R,5S,6R)-3-(acetylamino)-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-thiopyran-2-yl [(2R,3S,4R,5R)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3,4-dihydroxytetrahydrofuran-2-yl]methyl dihydrogen diphosphate (three-letter code: 12V) (formula: $C_{17}H_{27}N_3O_{16}P_2S$).



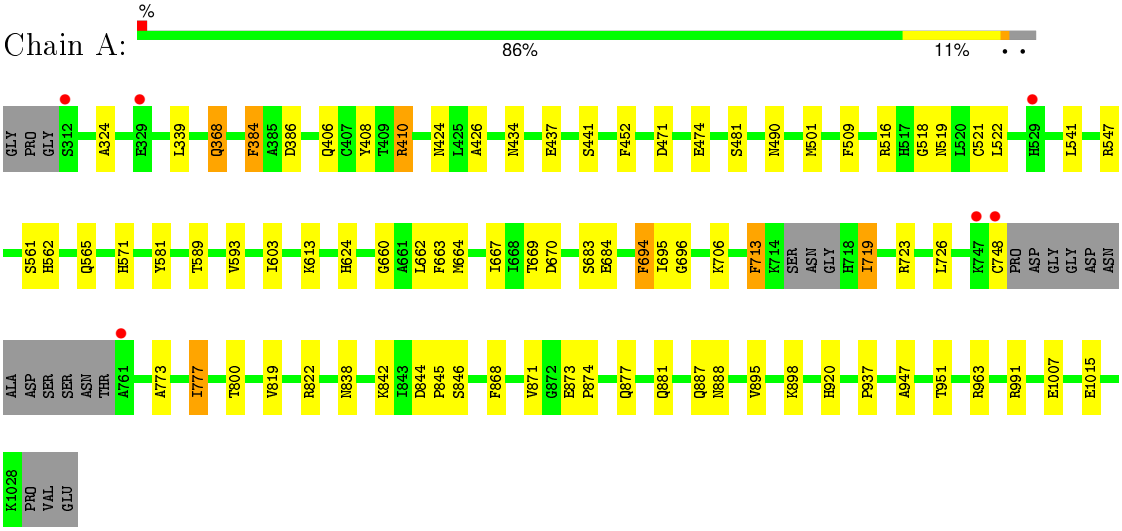
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P S 39 17 3 16 2 1	0	0
4	B	1	Total C N O P S 39 17 3 16 2 1	0	0
4	C	1	Total C N O P S 39 17 3 16 2 1	0	0
4	D	1	Total C N O P S 39 17 3 16 2 1	0	0

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.

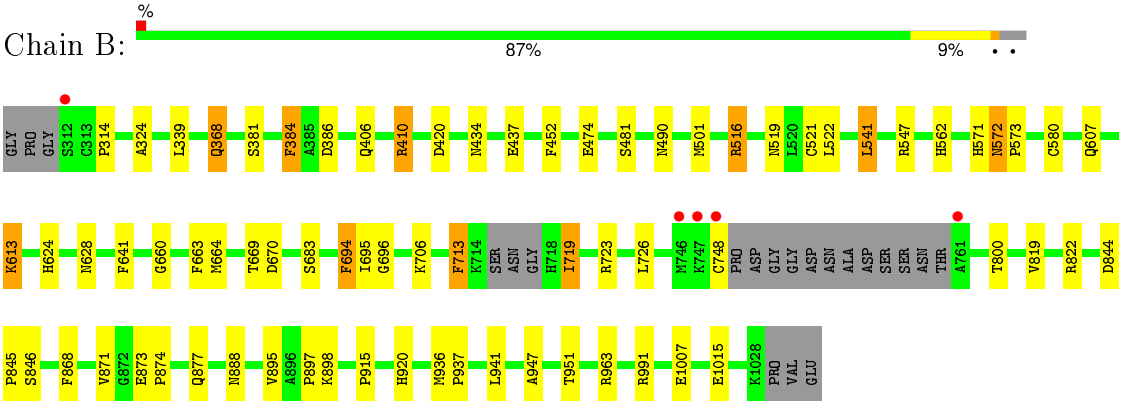
• Molecule 1: UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit

Chain A:



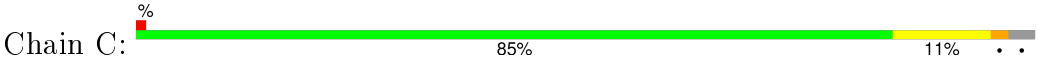
• Molecule 1: UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit

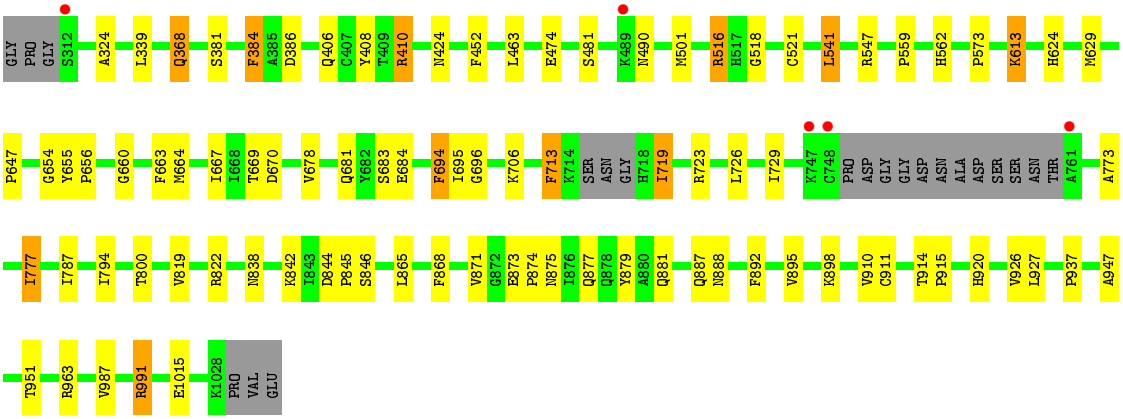
Chain B:



• Molecule 1: UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit

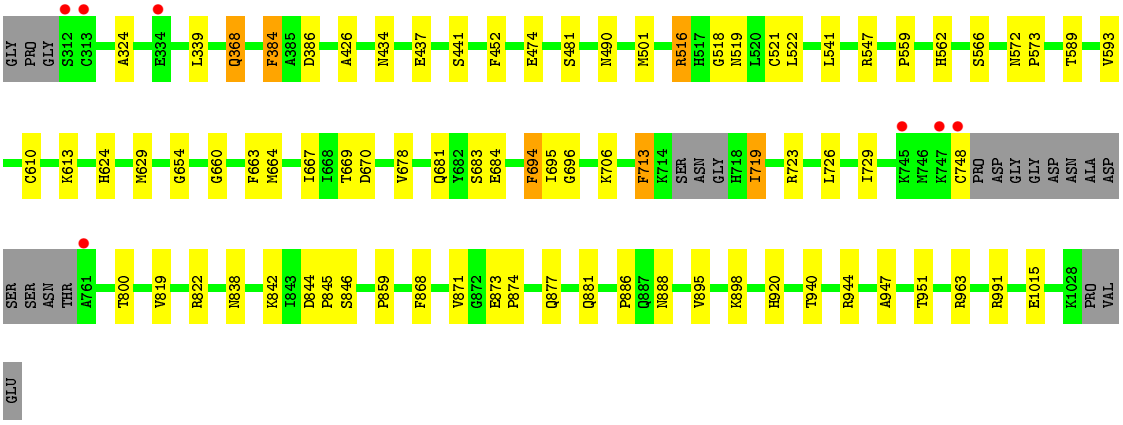
Chain C:





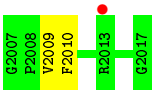
• Molecule 1: UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit

Chain D:



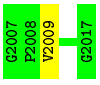
• Molecule 2: Keratin, type II cytoskeletal 7

Chain E:



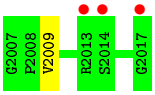
• Molecule 2: Keratin, type II cytoskeletal 7

Chain F:



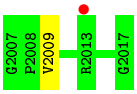
• Molecule 2: Keratin, type II cytoskeletal 7

Chain G:



- Molecule 2: Keratin, type II cytoskeletal 7

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	275.14Å 275.14Å 143.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 3.20 24.98 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (25.00-3.20) 99.0 (24.98-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.05 (at 3.17Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.216 , 0.239 0.210 , 0.237	Depositor DCC
R_{free} test set	1015 reflections (1.01%)	DCC
Wilson B-factor (Å^2)	59.1	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 4.5	EDS
Estimated twinning fraction	0.023 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 101319 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	22608	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% 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: SO4, 12V

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	0.70	0/5662	0.69	1/7679 (0.0%)
1	B	0.71	2/5662 (0.0%)	0.69	2/7679 (0.0%)
1	C	0.68	0/5662	0.68	2/7679 (0.0%)
1	D	0.71	1/5662 (0.0%)	0.68	2/7679 (0.0%)
2	E	1.23	0/74	0.80	0/99
2	F	1.14	0/74	0.75	0/99
2	G	1.26	0/74	0.78	0/99
2	H	1.11	0/74	0.81	0/99
All	All	0.71	3/22944 (0.0%)	0.69	7/31112 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	580	CYS	CB-SG	-7.53	1.69	1.82
1	B	521	CYS	CB-SG	-5.88	1.72	1.81
1	D	610	CYS	CB-SG	5.71	1.92	1.82

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	963	ARG	NE-CZ-NH1	10.09	125.34	120.30
1	A	963	ARG	NE-CZ-NH1	9.33	124.97	120.30
1	D	963	ARG	NE-CZ-NH1	9.17	124.88	120.30
1	C	963	ARG	NE-CZ-NH1	7.96	124.28	120.30
1	D	963	ARG	NE-CZ-NH2	-5.82	117.39	120.30
1	C	963	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	B	963	ARG	NE-CZ-NH2	-5.08	117.76	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5535	0	5503	52	0
1	B	5535	0	5503	47	0
1	C	5535	0	5503	56	0
1	D	5535	0	5503	46	0
2	E	73	0	70	3	0
2	F	73	0	70	2	0
2	G	73	0	70	2	0
2	H	73	0	70	2	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	H	5	0	0	0	0
4	A	39	0	25	1	0
4	B	39	0	25	0	0
4	C	39	0	25	1	0
4	D	39	0	25	1	0
All	All	22608	0	22392	196	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 4.

All (196) 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:B:547:ARG:H	1:B:624:HIS:HD2	1.10	0.98
1:A:547:ARG:H	1:A:624:HIS:HD2	1.08	0.98
1:C:547:ARG:H	1:C:624:HIS:HD2	1.14	0.94
1:D:547:ARG:H	1:D:624:HIS:HD2	1.10	0.88
1:A:547:ARG:H	1:A:624:HIS:CD2	1.92	0.87
1:D:547:ARG:H	1:D:624:HIS:CD2	1.93	0.85
1:C:547:ARG:H	1:C:624:HIS:CD2	1.99	0.79
1:B:547:ARG:H	1:B:624:HIS:CD2	1.98	0.79
1:A:368:GLN:HA	1:A:368:GLN:NE2	1.98	0.78
1:D:368:GLN:NE2	1:D:368:GLN:HA	1.99	0.76
1:B:368:GLN:HA	1:B:368:GLN:NE2	2.06	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:368:GLN:HA	1:A:368:GLN:HE21	1.58	0.68
1:D:474:GLU:OE1	1:D:474:GLU:HA	1.92	0.67
1:C:368:GLN:HA	1:C:368:GLN:NE2	2.08	0.67
1:D:719:ILE:HG22	1:D:719:ILE:O	1.94	0.67
1:A:663:PHE:CD1	1:A:664:MET:HE2	2.30	0.67
1:C:719:ILE:HG22	1:C:719:ILE:O	1.95	0.66
1:A:547:ARG:N	1:A:624:HIS:HD2	1.89	0.65
1:A:719:ILE:HG22	1:A:719:ILE:O	1.95	0.65
1:B:663:PHE:CD1	1:B:664:MET:HE2	2.33	0.64
1:C:873:GLU:HB3	1:C:874:PRO:HD3	1.79	0.64
1:C:663:PHE:CD1	1:C:664:MET:HE2	2.33	0.63
1:A:695:ILE:HG13	1:A:696:GLY:N	2.14	0.62
1:B:368:GLN:HA	1:B:368:GLN:HE21	1.64	0.62
1:A:410:ARG:NH2	1:B:410:ARG:NH2	2.48	0.61
1:D:547:ARG:N	1:D:624:HIS:HD2	1.91	0.61
1:D:368:GLN:HE21	1:D:368:GLN:HA	1.65	0.61
1:A:895:VAL:HG11	2:E:2009:VAL:HG21	1.83	0.60
1:D:559:PRO:HB2	4:D:1200:12V:H6'	1.83	0.60
1:D:895:VAL:HG11	2:H:2009:VAL:HG21	1.84	0.59
1:B:719:ILE:O	1:B:719:ILE:HG22	2.01	0.59
1:C:877:GLN:HA	1:C:877:GLN:OE1	2.01	0.59
1:C:474:GLU:HA	1:C:474:GLU:OE1	2.03	0.59
1:B:562:HIS:ND1	1:B:898:LYS:HE3	2.18	0.58
1:B:726:LEU:HD22	1:B:819:VAL:HG22	1.86	0.58
1:C:895:VAL:HG11	2:G:2009:VAL:HG21	1.83	0.58
1:D:562:HIS:ND1	1:D:898:LYS:HE3	2.19	0.57
1:A:562:HIS:ND1	1:A:898:LYS:HE3	2.18	0.57
1:D:895:VAL:HG11	2:H:2009:VAL:CG2	2.35	0.57
1:B:695:ILE:HG13	1:B:696:GLY:N	2.20	0.57
1:C:667:ILE:HG22	1:C:684:GLU:HB2	1.87	0.56
1:B:474:GLU:HA	1:B:474:GLU:OE1	2.06	0.55
1:B:844:ASP:HB2	1:B:845:PRO:CD	2.37	0.54
1:B:895:VAL:HG11	2:F:2009:VAL:HG21	1.87	0.54
1:C:368:GLN:HA	1:C:368:GLN:HE21	1.72	0.54
1:A:1015:GLU:OE1	1:A:1015:GLU:HA	2.08	0.54
1:C:562:HIS:ND1	1:C:898:LYS:HE3	2.22	0.54
1:A:410:ARG:HH22	1:B:410:ARG:NH2	2.02	0.54
1:D:663:PHE:CD1	1:D:664:MET:HE2	2.42	0.54
1:A:669:THR:OG1	1:A:670:ASP:N	2.41	0.54
1:A:713:PHE:N	1:A:713:PHE:CD2	2.75	0.54
1:B:947:ALA:O	1:B:951:THR:HG23	2.08	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:895:VAL:HG11	2:G:2009:VAL:CG2	2.37	0.53
1:A:406:GLN:O	1:A:410:ARG:HB2	2.09	0.53
1:B:723:ARG:HD2	1:B:822:ARG:HD2	1.91	0.53
1:A:947:ALA:O	1:A:951:THR:HG23	2.09	0.52
1:D:713:PHE:CD2	1:D:713:PHE:N	2.78	0.52
1:C:695:ILE:HG13	1:C:696:GLY:N	2.25	0.51
1:A:663:PHE:HD1	1:A:664:MET:HE2	1.72	0.51
1:B:873:GLU:HB3	1:B:874:PRO:HD3	1.91	0.51
1:A:660:GLY:HA2	1:A:683:SER:HB3	1.92	0.51
1:A:895:VAL:HG11	2:E:2009:VAL:CG2	2.40	0.51
1:B:541:LEU:H	1:B:541:LEU:HD12	1.75	0.51
1:C:947:ALA:O	1:C:951:THR:HG23	2.10	0.51
1:D:947:ALA:O	1:D:951:THR:HG23	2.11	0.50
1:C:629:MET:O	1:C:654:GLY:HA3	2.11	0.50
1:C:937:PRO:HG3	1:C:947:ALA:HB2	1.92	0.50
1:A:474:GLU:HA	1:A:474:GLU:OE1	2.10	0.50
1:D:844:ASP:HB2	1:D:845:PRO:CD	2.41	0.50
1:A:662:LEU:HD21	1:D:886:PRO:HA	1.92	0.50
1:C:723:ARG:HD2	1:C:822:ARG:HD2	1.94	0.50
1:D:519:ASN:HA	1:D:522:LEU:HD12	1.94	0.50
1:D:719:ILE:CG2	1:D:719:ILE:O	2.60	0.49
1:A:490:ASN:HA	1:A:516:ARG:NH2	2.27	0.49
1:A:726:LEU:HD22	1:A:819:VAL:HG22	1.93	0.49
1:D:873:GLU:HB3	1:D:874:PRO:HD3	1.94	0.49
1:B:877:GLN:OE1	1:B:877:GLN:HA	2.13	0.49
1:D:726:LEU:HD22	1:D:819:VAL:HG22	1.93	0.49
1:D:695:ILE:HG13	1:D:696:GLY:N	2.27	0.49
1:A:937:PRO:HG3	1:A:947:ALA:HB2	1.94	0.49
1:A:873:GLU:HB3	1:A:874:PRO:HD3	1.95	0.49
1:A:368:GLN:CA	1:A:368:GLN:NE2	2.74	0.49
1:A:667:ILE:HG22	1:A:684:GLU:HB2	1.94	0.49
1:C:713:PHE:N	1:C:713:PHE:CD2	2.79	0.49
1:A:723:ARG:HD2	1:A:822:ARG:HD2	1.95	0.49
1:D:434:ASN:HB3	1:D:437:GLU:OE1	2.13	0.48
1:B:406:GLN:O	1:B:410:ARG:HB2	2.13	0.48
1:D:669:THR:OG1	1:D:670:ASP:N	2.44	0.48
1:A:773:ALA:O	1:A:777:ILE:HG12	2.14	0.48
1:B:713:PHE:CD2	1:B:713:PHE:N	2.82	0.48
1:B:547:ARG:N	1:B:624:HIS:HD2	1.93	0.48
1:D:490:ASN:HA	1:D:516:ARG:NH2	2.28	0.48
1:C:660:GLY:HA2	1:C:683:SER:HB3	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:865:LEU:O	1:C:892:PHE:HA	2.14	0.48
1:B:324:ALA:HB2	1:B:339:LEU:HB2	1.95	0.48
1:C:408:TYR:CZ	1:C:424:ASN:HB3	2.49	0.48
1:A:368:GLN:CA	1:A:368:GLN:HE21	2.24	0.48
1:C:559:PRO:HB2	4:C:1102:12V:H6'	1.95	0.48
1:D:723:ARG:HD2	1:D:822:ARG:HD2	1.96	0.47
1:C:490:ASN:HA	1:C:516:ARG:NH2	2.28	0.47
1:B:663:PHE:HD1	1:B:664:MET:HE2	1.78	0.47
1:D:877:GLN:O	1:D:881:GLN:OE1	2.32	0.47
1:C:406:GLN:O	1:C:410:ARG:HB2	2.14	0.47
1:A:898:LYS:HE2	4:A:1102:12V:O3B	2.14	0.47
1:B:434:ASN:HB3	1:B:437:GLU:OE1	2.15	0.47
1:C:663:PHE:HD1	1:C:664:MET:HE2	1.79	0.47
1:C:844:ASP:HB2	1:C:845:PRO:CD	2.45	0.47
1:B:660:GLY:HA2	1:B:683:SER:HB3	1.97	0.46
1:B:519:ASN:HA	1:B:522:LEU:HD12	1.97	0.46
1:A:581:TYR:CE1	1:A:603:ILE:HD13	2.51	0.46
1:A:719:ILE:CG2	1:A:719:ILE:O	2.62	0.46
1:A:877:GLN:HA	1:A:877:GLN:OE1	2.15	0.46
1:A:434:ASN:HB3	1:A:437:GLU:OE1	2.16	0.46
1:D:368:GLN:HE21	1:D:368:GLN:CA	2.28	0.46
1:B:628:ASN:HB2	1:B:641:PHE:CZ	2.51	0.46
1:D:729:ILE:HD13	1:D:729:ILE:HA	1.88	0.46
1:D:572:ASN:HA	1:D:573:PRO:HD3	1.83	0.45
1:A:384:PHE:CE2	1:A:386:ASP:HB3	2.51	0.45
1:C:719:ILE:CG2	1:C:719:ILE:O	2.63	0.45
1:C:624:HIS:O	1:C:647:PRO:HG2	2.16	0.45
1:B:694:PHE:CZ	1:B:920:HIS:HB3	2.52	0.45
1:D:368:GLN:NE2	1:D:368:GLN:CA	2.75	0.45
1:A:881:GLN:HE22	1:A:887:GLN:HG2	1.82	0.45
1:D:660:GLY:HA2	1:D:683:SER:HB3	1.97	0.45
1:C:694:PHE:CZ	1:C:920:HIS:HB3	2.52	0.45
1:B:314:PRO:HG2	2:E:2010:PHE:HZ	1.81	0.45
1:C:881:GLN:HE22	1:C:887:GLN:HG2	1.82	0.45
1:A:324:ALA:HB2	1:A:339:LEU:HB2	1.98	0.45
1:C:875:ASN:O	1:C:879:TYR:HD2	2.00	0.45
1:D:518:GLY:O	1:D:521:CYS:HB2	2.17	0.45
1:C:547:ARG:N	1:C:624:HIS:HD2	1.97	0.44
1:B:613:LYS:H	1:B:613:LYS:HG2	1.52	0.44
1:C:773:ALA:O	1:C:777:ILE:HG12	2.16	0.44
1:A:844:ASP:HB2	1:A:845:PRO:CD	2.47	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:838:ASN:HB3	1:A:842:LYS:HD2	1.99	0.44
1:B:937:PRO:HG3	1:B:947:ALA:HB2	1.99	0.44
1:B:941:LEU:HD23	1:B:941:LEU:C	2.38	0.44
1:A:695:ILE:HG13	1:A:696:GLY:H	1.83	0.44
1:B:895:VAL:HG11	2:F:2009:VAL:CG2	2.47	0.44
1:D:1015:GLU:OE1	1:D:1015:GLU:HA	2.17	0.44
1:C:838:ASN:HB3	1:C:842:LYS:HD2	2.00	0.44
1:C:669:THR:OG1	1:C:670:ASP:N	2.49	0.44
1:A:518:GLY:O	1:A:521:CYS:HB2	2.18	0.44
1:C:1015:GLU:OE1	1:C:1015:GLU:HA	2.17	0.44
1:D:940:THR:O	1:D:944:ARG:HG3	2.18	0.44
1:C:777:ILE:HG12	1:C:777:ILE:H	1.63	0.43
1:C:729:ILE:HA	1:C:729:ILE:HD13	1.89	0.43
1:B:490:ASN:HA	1:B:516:ARG:NH2	2.34	0.43
1:B:1015:GLU:HA	1:B:1015:GLU:OE1	2.17	0.43
1:C:910:VAL:HG12	1:C:911:CYS:N	2.32	0.43
1:D:629:MET:O	1:D:654:GLY:HA3	2.19	0.43
1:D:426:ALA:HB2	1:D:441:SER:HB2	2.01	0.43
1:C:914:THR:HA	1:C:915:PRO:HD3	1.89	0.43
1:D:384:PHE:CE2	1:D:386:ASP:HB3	2.54	0.43
1:B:669:THR:OG1	1:B:670:ASP:N	2.51	0.42
1:A:408:TYR:CZ	1:A:424:ASN:HB3	2.53	0.42
1:B:719:ILE:O	1:B:719:ILE:CG2	2.67	0.42
1:A:571:HIS:HA	1:A:1007:GLU:OE2	2.18	0.42
1:B:410:ARG:HH22	1:C:410:ARG:NH2	2.18	0.42
1:C:324:ALA:HB2	1:C:339:LEU:HB2	2.01	0.42
1:D:694:PHE:CZ	1:D:920:HIS:HB3	2.54	0.42
1:B:571:HIS:HA	1:B:1007:GLU:OE2	2.20	0.42
1:D:713:PHE:HD2	1:D:713:PHE:N	2.18	0.42
1:A:777:ILE:H	1:A:777:ILE:HG12	1.63	0.42
1:C:678:VAL:O	1:C:681:GLN:HG2	2.20	0.42
1:A:561:SER:O	1:A:565:GLN:HB3	2.20	0.42
1:C:910:VAL:CG1	1:C:911:CYS:N	2.82	0.42
1:B:572:ASN:HA	1:B:573:PRO:HD3	1.80	0.42
1:A:509:PHE:HE1	1:D:859:PRO:HG3	1.85	0.41
1:B:381:SER:O	1:B:384:PHE:HB2	2.20	0.41
1:D:838:ASN:HB3	1:D:842:LYS:HD2	2.02	0.41
1:C:927:LEU:HA	1:C:927:LEU:HD23	1.89	0.41
1:B:607:GLN:HB2	1:B:607:GLN:HE21	1.69	0.41
1:A:589:THR:O	1:A:593:VAL:HG23	2.19	0.41
1:C:541:LEU:HD12	1:C:541:LEU:H	1.85	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:897:PRO:O	1:B:898:LYS:C	2.59	0.41
1:C:518:GLY:O	1:C:521:CYS:HB2	2.20	0.41
1:A:694:PHE:CZ	1:A:920:HIS:HB3	2.56	0.41
1:D:877:GLN:OE1	1:D:877:GLN:HA	2.21	0.41
1:C:726:LEU:HD22	1:C:819:VAL:HG22	2.02	0.41
1:C:381:SER:O	1:C:384:PHE:HB2	2.20	0.41
1:C:787:ILE:HG13	1:C:794:ILE:HB	2.02	0.41
1:C:384:PHE:CE2	1:C:386:ASP:HB3	2.55	0.41
1:D:324:ALA:HB2	1:D:339:LEU:HB2	2.02	0.41
1:C:463:LEU:HA	1:C:463:LEU:HD23	1.94	0.41
1:C:987:VAL:O	1:C:991:ARG:HG2	2.21	0.41
1:D:589:THR:O	1:D:593:VAL:HG23	2.20	0.40
1:C:655:TYR:HA	1:C:656:PRO:HD3	1.96	0.40
1:D:678:VAL:O	1:D:681:GLN:HG2	2.21	0.40
1:B:384:PHE:CE2	1:B:386:ASP:HB3	2.56	0.40
1:A:426:ALA:HB2	1:A:441:SER:HB2	2.04	0.40
1:B:936:MET:HA	1:B:937:PRO:HD3	1.91	0.40
1:C:613:LYS:H	1:C:613:LYS:HG2	1.54	0.40
1:B:844:ASP:HB2	1:B:845:PRO:HD3	2.04	0.40
1:A:519:ASN:HA	1:A:522:LEU:HD12	2.02	0.40
1:D:667:ILE:HG22	1:D:684:GLU:HB2	2.03	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	696/723 (96%)	659 (95%)	35 (5%)	2 (0%)	46	85
1	B	696/723 (96%)	654 (94%)	40 (6%)	2 (0%)	46	85
1	C	696/723 (96%)	652 (94%)	42 (6%)	2 (0%)	46	85
1	D	696/723 (96%)	657 (94%)	37 (5%)	2 (0%)	46	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	9/11 (82%)	9 (100%)	0	0	100	100
2	F	9/11 (82%)	9 (100%)	0	0	100	100
2	G	9/11 (82%)	9 (100%)	0	0	100	100
2	H	9/11 (82%)	8 (89%)	1 (11%)	0	100	100
All	All	2820/2936 (96%)	2657 (94%)	155 (6%)	8 (0%)	46	85

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	888	ASN
1	C	888	ASN
1	D	719	ILE
1	B	719	ILE
1	B	888	ASN
1	D	888	ASN
1	A	719	ILE
1	C	719	ILE

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	601/618 (97%)	582 (97%)	19 (3%)	46	81
1	B	601/618 (97%)	580 (96%)	21 (4%)	43	80
1	C	601/618 (97%)	581 (97%)	20 (3%)	45	81
1	D	601/618 (97%)	583 (97%)	18 (3%)	48	82
2	E	7/7 (100%)	7 (100%)	0	100	100
2	F	7/7 (100%)	7 (100%)	0	100	100
2	G	7/7 (100%)	7 (100%)	0	100	100
2	H	7/7 (100%)	7 (100%)	0	100	100
All	All	2432/2500 (97%)	2354 (97%)	78 (3%)	46	81

All (78) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	368	GLN
1	A	384	PHE
1	A	410	ARG
1	A	452	PHE
1	A	471	ASP
1	A	481	SER
1	A	501	MET
1	A	541	LEU
1	A	613	LYS
1	A	694	PHE
1	A	706	LYS
1	A	713	PHE
1	A	748	CYS
1	A	777	ILE
1	A	800	THR
1	A	846	SER
1	A	868	PHE
1	A	871	VAL
1	A	991	ARG
1	B	368	GLN
1	B	384	PHE
1	B	410	ARG
1	B	420	ASP
1	B	452	PHE
1	B	481	SER
1	B	501	MET
1	B	516	ARG
1	B	541	LEU
1	B	572	ASN
1	B	613	LYS
1	B	694	PHE
1	B	706	LYS
1	B	713	PHE
1	B	748	CYS
1	B	800	THR
1	B	846	SER
1	B	868	PHE
1	B	871	VAL
1	B	915	PRO
1	B	991	ARG
1	C	368	GLN
1	C	384	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	410	ARG
1	C	452	PHE
1	C	481	SER
1	C	501	MET
1	C	516	ARG
1	C	541	LEU
1	C	573	PRO
1	C	613	LYS
1	C	694	PHE
1	C	706	LYS
1	C	713	PHE
1	C	777	ILE
1	C	800	THR
1	C	846	SER
1	C	868	PHE
1	C	871	VAL
1	C	926	VAL
1	C	991	ARG
1	D	368	GLN
1	D	384	PHE
1	D	452	PHE
1	D	481	SER
1	D	501	MET
1	D	516	ARG
1	D	541	LEU
1	D	566	SER
1	D	613	LYS
1	D	694	PHE
1	D	706	LYS
1	D	713	PHE
1	D	748	CYS
1	D	800	THR
1	D	846	SER
1	D	868	PHE
1	D	871	VAL
1	D	991	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (35) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	321	ASN
1	A	368	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	413	GLN
1	A	507	HIS
1	A	607	GLN
1	A	624	HIS
1	A	763	ASN
1	A	781	ASN
1	A	786	GLN
1	A	881	GLN
1	B	321	ASN
1	B	368	GLN
1	B	413	GLN
1	B	607	GLN
1	B	624	HIS
1	B	763	ASN
1	B	786	GLN
1	B	881	GLN
1	C	368	GLN
1	C	413	GLN
1	C	507	HIS
1	C	607	GLN
1	C	620	GLN
1	C	624	HIS
1	C	763	ASN
1	C	786	GLN
1	C	881	GLN
1	D	321	ASN
1	D	368	GLN
1	D	507	HIS
1	D	607	GLN
1	D	624	HIS
1	D	763	ASN
1	D	786	GLN
1	D	881	GLN

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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

8 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	$\#\|Z\| > 2$
3	SO4	A	1101	-	4,4,4	0.16	0	6,6,6	0.41	0
4	12V	A	1102	-	32,41,41	1.69	7 (21%)	41,62,62	1.92	6 (14%)
3	SO4	B	1101	-	4,4,4	0.22	0	6,6,6	0.12	0
4	12V	B	1102	-	32,41,41	1.57	4 (12%)	41,62,62	1.91	8 (19%)
3	SO4	C	1101	-	4,4,4	0.16	0	6,6,6	0.27	0
4	12V	C	1102	-	32,41,41	1.38	4 (12%)	41,62,62	1.73	5 (12%)
4	12V	D	1200	-	32,41,41	1.75	6 (18%)	41,62,62	1.92	6 (14%)
3	SO4	H	2101	-	4,4,4	0.21	0	6,6,6	0.18	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
3	SO4	A	1101	-	-	0/0/0/0	0/0/0/0
4	12V	A	1102	-	-	0/21/63/63	0/3/3/3
3	SO4	B	1101	-	-	0/0/0/0	0/0/0/0
4	12V	B	1102	-	-	0/21/63/63	0/3/3/3
3	SO4	C	1101	-	-	0/0/0/0	0/0/0/0
4	12V	C	1102	-	-	0/21/63/63	0/3/3/3
4	12V	D	1200	-	-	0/21/63/63	0/3/3/3
3	SO4	H	2101	-	-	0/0/0/0	0/0/0/0

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1200	12V	C4'-C5'	-3.69	1.50	1.53
4	D	1200	12V	C5'-S5'	-3.47	1.76	1.82
4	D	1200	12V	C2B-C1B	-3.43	1.48	1.53
4	C	1102	12V	C5'-S5'	-3.03	1.77	1.82
4	A	1102	12V	C2B-C1B	-2.96	1.48	1.53
4	A	1102	12V	C5'-S5'	-2.52	1.78	1.82
4	B	1102	12V	C5'-S5'	-2.39	1.78	1.82
4	B	1102	12V	C2B-C1B	-2.29	1.50	1.53
4	A	1102	12V	C2B-C3B	-2.23	1.47	1.53
4	C	1102	12V	C2B-C1B	-2.12	1.50	1.53
4	D	1200	12V	O4B-C1B	2.32	1.44	1.41
4	A	1102	12V	C4-N3	2.96	1.38	1.33
4	A	1102	12V	C6'-C5'	3.54	1.55	1.52
4	C	1102	12V	C4-N3	3.56	1.39	1.33
4	A	1102	12V	O4B-C1B	3.61	1.46	1.41
4	D	1200	12V	C4-N3	3.82	1.39	1.33
4	C	1102	12V	C6-N1	4.03	1.41	1.35
4	A	1102	12V	C6-N1	4.43	1.41	1.35
4	B	1102	12V	C4-N3	4.44	1.41	1.33
4	D	1200	12V	C6-N1	4.67	1.41	1.35
4	B	1102	12V	C6-N1	4.95	1.42	1.35

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1102	12V	C4B-O4B-C1B	-8.23	100.92	109.64
4	D	1200	12V	C6'-C5'-S5'	-5.75	87.07	109.66
4	B	1102	12V	C1'-C2'-N2'	-5.50	99.88	111.17
4	C	1102	12V	C1'-C2'-N2'	-3.62	103.74	111.17
4	C	1102	12V	C6'-C5'-S5'	-3.43	96.18	109.66
4	B	1102	12V	C6'-C5'-S5'	-3.24	96.91	109.66
4	D	1200	12V	C4B-O4B-C1B	-3.07	106.39	109.64
4	B	1102	12V	O3'-C3'-C4'	-2.83	103.98	110.36
4	D	1200	12V	C1'-C2'-N2'	-2.70	105.63	111.17
4	B	1102	12V	C4B-O4B-C1B	-2.63	106.86	109.64
4	C	1102	12V	C5-C4-N3	-2.26	117.73	123.28
4	A	1102	12V	C5B-C4B-C3B	-2.24	106.52	115.20
4	A	1102	12V	C1'-C2'-N2'	-2.24	106.58	111.17
4	B	1102	12V	C3'-C2'-N2'	2.22	115.29	110.67
4	B	1102	12V	O3'-C3'-C2'	2.45	114.66	109.57
4	B	1102	12V	O1'-C1'-C2'	3.11	112.47	107.39
4	A	1102	12V	O1'-C1'-C2'	3.33	112.81	107.39
4	A	1102	12V	O4B-C1B-N1	3.41	114.59	108.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	1102	12V	O4B-C1B-N1	3.46	114.69	108.10
4	D	1200	12V	O3A-PB-O1'	3.83	108.89	102.05
4	D	1200	12V	O1'-C1'-C2'	4.04	113.98	107.39
4	A	1102	12V	C4-N3-C2	4.20	118.64	114.21
4	C	1102	12V	C4-N3-C2	6.17	120.71	114.21
4	D	1200	12V	C4-N3-C2	6.33	120.88	114.21
4	B	1102	12V	C4-N3-C2	6.74	121.31	114.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1102	12V	1	0
4	C	1102	12V	1	0
4	D	1200	12V	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	702/723 (97%)	-0.43	6 (0%) 85 78	35, 51, 78, 113	0
1	B	702/723 (97%)	-0.47	5 (0%) 89 83	35, 51, 78, 113	0
1	C	702/723 (97%)	-0.52	5 (0%) 89 83	36, 51, 78, 113	0
1	D	702/723 (97%)	-0.38	7 (0%) 84 75	35, 51, 78, 113	0
2	E	11/11 (100%)	1.12	1 (9%) 11 6	49, 69, 83, 85	0
2	F	11/11 (100%)	0.92	0 100 100	49, 69, 83, 85	0
2	G	11/11 (100%)	1.68	3 (27%) 1 0	49, 69, 83, 85	0
2	H	11/11 (100%)	1.10	1 (9%) 11 6	49, 69, 83, 85	0
All	All	2852/2936 (97%)	-0.42	28 (0%) 84 75	35, 52, 79, 113	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	312	SER	5.7
1	D	312	SER	4.7
1	D	748	CYS	4.3
1	D	747	LYS	4.0
1	B	748	CYS	3.8
1	A	761	ALA	3.6
1	A	747	LYS	3.5
1	A	748	CYS	3.3
1	B	747	LYS	3.3
1	C	312	SER	3.0
1	C	748	CYS	2.7
1	A	312	SER	2.6
1	D	313	CYS	2.5
2	G	2017	GLY	2.5
1	D	761	ALA	2.5
1	B	761	ALA	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	2013	ARG	2.4
2	G	2014	SER	2.3
1	A	529	HIS	2.3
1	D	745	LYS	2.3
1	C	761	ALA	2.3
1	C	747	LYS	2.2
1	D	334	GLU	2.1
1	B	746	MET	2.1
2	G	2013	ARG	2.1
2	E	2013	ARG	2.0
1	A	329	GLU	2.0
1	C	489	LYS	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
3	SO4	A	1101	5/5	0.93	0.25	0.13	92,92,93,93	0
3	SO4	B	1101	5/5	0.96	0.27	0.10	98,98,99,99	0
3	SO4	H	2101	5/5	0.94	0.21	-0.53	93,93,94,94	0
3	SO4	C	1101	5/5	0.90	0.21	-0.65	106,107,107,108	0
4	12V	C	1102	39/39	0.98	0.11	-0.70	56,58,63,65	0
4	12V	A	1102	39/39	0.97	0.12	-0.96	38,43,47,49	0
4	12V	B	1102	39/39	0.98	0.12	-0.98	37,39,44,47	0
4	12V	D	1200	39/39	0.98	0.11	-1.25	37,39,48,51	0

6.5 Other polymers (i)

There are no such residues in this entry.