Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 05:47 am BST

PDB ID	:	3AVI
Title	:	Crystal structures of novel allosteric peptide inhibitors of HIV integrase in the LEDGF binding site
Authors	:	Peat, T.S.; Deadman, J.J.; Newman, J.; Rhodes, D.I.
Deposited on	:	2011-03-05
Resolution	:	1.70 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

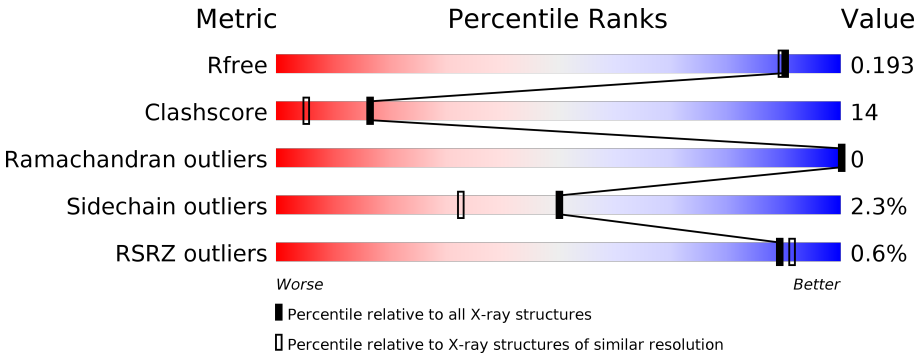
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.70 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}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 respectively. 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	183	% 69% 10% • 19%
1	B	183	% 64% 14% • 19%
2	D	8	100%
2	F	8	100%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	ACY	A	213	-	-	X	-
5	ACY	A	3	-	-	X	-
5	ACY	B	2	-	-	X	-
5	ACY	B	4	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 2864 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 Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	15	0
			1256	807	214	231	4			
1	B	149	Total	C	N	O	S	0	14	0
			1249	801	214	230	4			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	EXPRESSION TAG	UNP P12497
A	31	GLY	-	EXPRESSION TAG	UNP P12497
A	32	SER	-	EXPRESSION TAG	UNP P12497
A	33	SER	-	EXPRESSION TAG	UNP P12497
A	34	HIS	-	EXPRESSION TAG	UNP P12497
A	35	HIS	-	EXPRESSION TAG	UNP P12497
A	36	HIS	-	EXPRESSION TAG	UNP P12497
A	37	HIS	-	EXPRESSION TAG	UNP P12497
A	38	HIS	-	EXPRESSION TAG	UNP P12497
A	39	HIS	-	EXPRESSION TAG	UNP P12497
A	40	SER	-	EXPRESSION TAG	UNP P12497
A	41	SER	-	EXPRESSION TAG	UNP P12497
A	42	GLY	-	EXPRESSION TAG	UNP P12497
A	43	LEU	-	EXPRESSION TAG	UNP P12497
A	44	VAL	-	EXPRESSION TAG	UNP P12497
A	45	PRO	-	EXPRESSION TAG	UNP P12497
A	46	ARG	-	EXPRESSION TAG	UNP P12497
A	47	GLY	-	EXPRESSION TAG	UNP P12497
A	48	SER	-	EXPRESSION TAG	UNP P12497
A	49	HIS	-	EXPRESSION TAG	UNP P12497
A	56	SER	CYS	ENGINEERED MUTATION	UNP P12497
A	139	ASP	PHE	ENGINEERED MUTATION	UNP P12497
A	185	HIS	PHE	ENGINEERED MUTATION	UNP P12497
B	30	MET	-	EXPRESSION TAG	UNP P12497
B	31	GLY	-	EXPRESSION TAG	UNP P12497

Continued on next page...

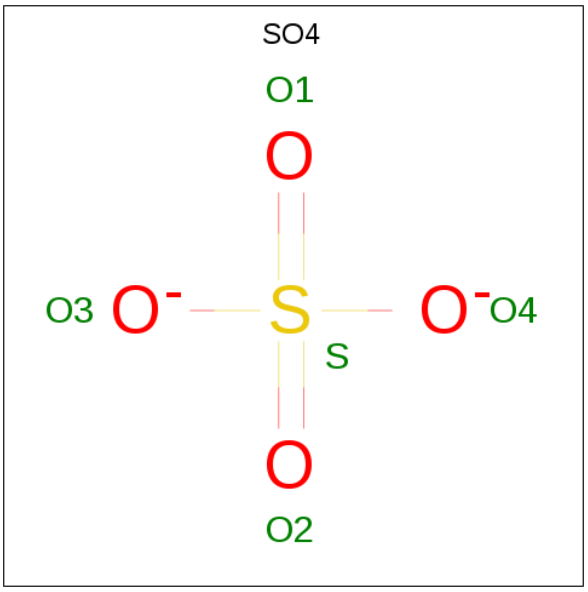
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	32	SER	-	EXPRESSION TAG	UNP P12497
B	33	SER	-	EXPRESSION TAG	UNP P12497
B	34	HIS	-	EXPRESSION TAG	UNP P12497
B	35	HIS	-	EXPRESSION TAG	UNP P12497
B	36	HIS	-	EXPRESSION TAG	UNP P12497
B	37	HIS	-	EXPRESSION TAG	UNP P12497
B	38	HIS	-	EXPRESSION TAG	UNP P12497
B	39	HIS	-	EXPRESSION TAG	UNP P12497
B	40	SER	-	EXPRESSION TAG	UNP P12497
B	41	SER	-	EXPRESSION TAG	UNP P12497
B	42	GLY	-	EXPRESSION TAG	UNP P12497
B	43	LEU	-	EXPRESSION TAG	UNP P12497
B	44	VAL	-	EXPRESSION TAG	UNP P12497
B	45	PRO	-	EXPRESSION TAG	UNP P12497
B	46	ARG	-	EXPRESSION TAG	UNP P12497
B	47	GLY	-	EXPRESSION TAG	UNP P12497
B	48	SER	-	EXPRESSION TAG	UNP P12497
B	49	HIS	-	EXPRESSION TAG	UNP P12497
B	56	SER	CYS	ENGINEERED MUTATION	UNP P12497
B	139	ASP	PHE	ENGINEERED MUTATION	UNP P12497
B	185	HIS	PHE	ENGINEERED MUTATION	UNP P12497

- Molecule 2 is a protein called LEDGF peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	8	Total	C	N	O	S	0	0	0
			63	38	10	14	1			
2	F	8	Total	C	N	O	S	0	0	0
			63	38	10	14	1			

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

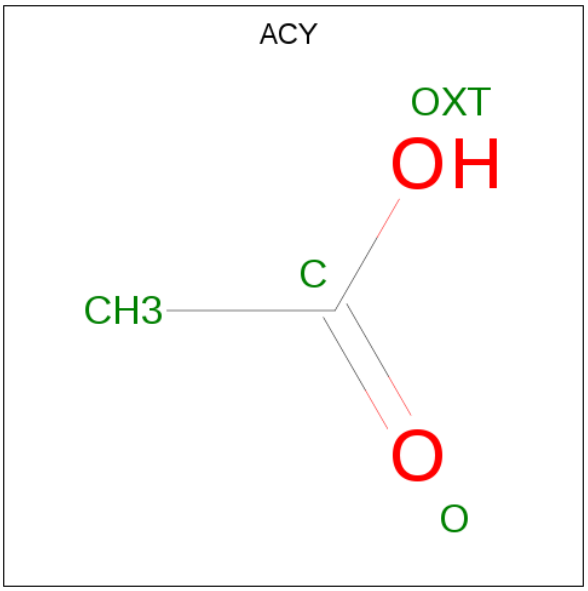


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is ACETIC ACID (three-letter code: ACY) (formula: $C_2H_4O_2$).



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

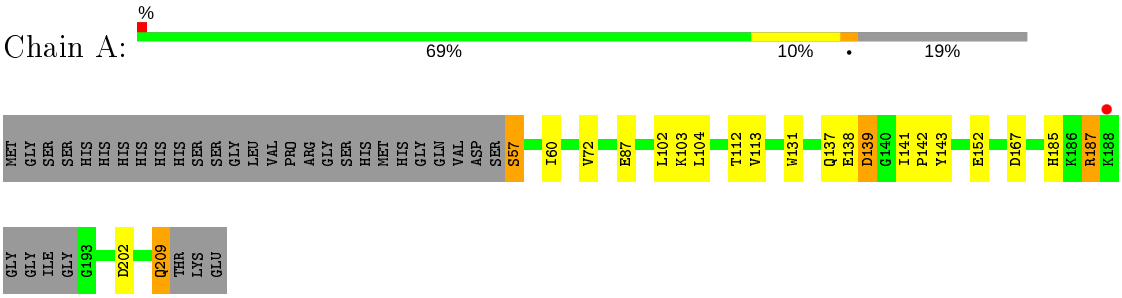
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	91	Total O 91 91	0	0
6	B	79	Total O 79 79	0	0
6	D	3	Total O 3 3	0	0
6	F	3	Total O 3 3	0	0

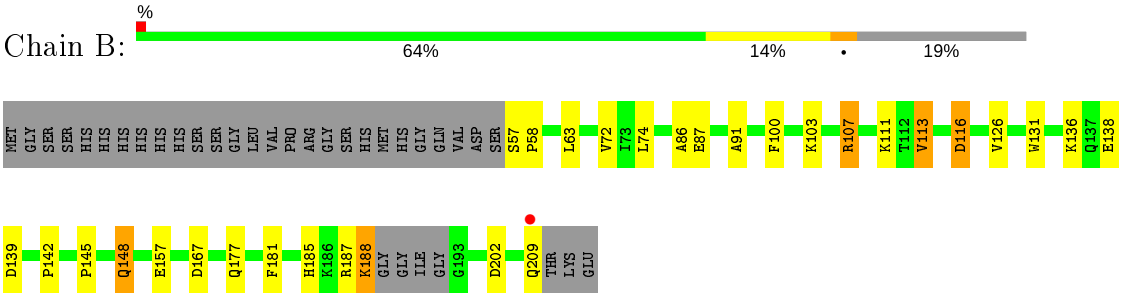
3 Residue-property plots (i)

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



- Molecule 1: Integrase



- Molecule 2: LEDGF peptide

There are no outlier residues recorded for this chain.

- Molecule 2: LEDGF peptide

There are no outlier residues recorded for this chain.

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	71.00Å 71.00Å 67.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	61.48 – 1.70 45.30 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (61.48-1.70) 99.9 (45.30-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.6.0062	Depositor
R, R_{free}	0.159 , 0.194 0.157 , 0.193	Depositor DCC
R_{free} test set	2086 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	19.5	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 38.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l 0.476 for h,-h-k,-l 0.021 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2864	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.91% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: ACY, SO4, CL

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	1.36	3/1305 (0.2%)	1.13	2/1770 (0.1%)
1	B	1.39	8/1298 (0.6%)	1.15	4/1758 (0.2%)
2	D	1.13	0/62	1.05	0/81
2	F	0.99	0/62	1.20	0/81
All	All	1.36	11/2727 (0.4%)	1.14	6/3690 (0.2%)

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	72	VAL	CB-CG1	-7.09	1.38	1.52
1	B	100	PHE	CD1-CE1	6.46	1.52	1.39
1	A	138	GLU	CB-CG	6.08	1.63	1.52
1	B	113[A]	VAL	CB-CG1	5.76	1.65	1.52
1	B	113[B]	VAL	CB-CG1	5.76	1.65	1.52
1	A	72	VAL	CB-CG1	-5.62	1.41	1.52
1	B	131	TRP	CZ3-CH2	5.59	1.49	1.40
1	B	91	ALA	CA-CB	5.55	1.64	1.52
1	A	131	TRP	CZ3-CH2	5.31	1.48	1.40
1	B	157	GLU	CG-CD	5.14	1.59	1.51
1	B	126	VAL	CB-CG2	5.08	1.63	1.52

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	116	ASP	CB-CG-OD1	6.38	124.04	118.30
1	B	107	ARG	NE-CZ-NH2	-5.55	117.53	120.30
1	B	111	LYS	CD-CE-NZ	-5.39	99.30	111.70
1	A	104	LEU	CB-CG-CD2	-5.30	101.99	111.00
1	B	139	ASP	CB-CG-OD2	-5.14	113.67	118.30
1	A	139	ASP	CB-CG-OD2	-5.05	113.76	118.30

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	1256	0	1287	35	0
1	B	1249	0	1278	29	0
2	D	63	0	63	0	0
2	F	63	0	63	0	0
3	A	20	0	0	0	0
3	B	20	0	0	1	0
4	A	1	0	0	0	0
5	A	8	0	6	14	0
5	B	8	0	6	10	0
6	A	91	0	0	7	0
6	B	79	0	0	3	0
6	D	3	0	0	0	0
6	F	3	0	0	0	0
All	All	2864	0	2703	77	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 14.

All (77) 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:60[B]:ILE:HD13	1:A:112[B]:THR:CG2	1.65	1.25
5:B:2:ACY:C	5:B:4:ACY:CH3	2.24	1.14
1:A:187:ARG:HH11	1:A:187:ARG:HG2	1.14	1.12
1:B:57:SER:HA	6:B:274:HOH:O	1.50	1.09
1:A:87[B]:GLU:OE1	6:A:8:HOH:O	1.72	1.05
1:A:141[B]:ILE:HD12	1:A:142:PRO:HD2	1.09	1.02
1:A:141[B]:ILE:HD12	1:A:142:PRO:CD	1.90	1.01
5:B:2:ACY:C	5:B:4:ACY:H1	1.91	0.98
1:A:60[B]:ILE:CD1	1:A:112[B]:THR:CG2	2.41	0.97
5:A:213:ACY:C	5:A:3:ACY:CH3	2.43	0.96

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:60[B]:ILE:CD1	1:A:112[B]:THR:HG21	1.97	0.94
5:A:213:ACY:OXT	5:A:3:ACY:CH3	2.18	0.92
1:A:141[B]:ILE:HG23	1:A:143:TYR:CE1	2.06	0.91
5:B:2:ACY:OXT	5:B:4:ACY:H1	1.71	0.91
5:A:213:ACY:C	5:A:3:ACY:H1	2.02	0.89
1:A:141[B]:ILE:CD1	1:A:142:PRO:HD2	2.02	0.88
5:B:2:ACY:OXT	5:B:4:ACY:CH3	2.22	0.87
1:A:139:ASP:OD1	1:A:141[B]:ILE:HG22	1.74	0.86
1:A:60[B]:ILE:HD13	1:A:112[B]:THR:HG22	1.56	0.86
1:B:167:ASP:OD1	6:B:283:HOH:O	1.94	0.86
1:B:148:GLN:HE21	1:B:148:GLN:H	1.26	0.83
1:A:187:ARG:NH1	1:A:187:ARG:HG2	1.93	0.79
1:A:187:ARG:HH11	1:A:187:ARG:CG	1.96	0.78
1:A:103[B]:LYS:NZ	1:B:87[B]:GLU:OE1	2.15	0.77
6:A:282:HOH:O	1:B:107:ARG:HG3	1.85	0.77
5:B:2:ACY:H1	6:B:214:HOH:O	1.82	0.77
5:B:2:ACY:O	5:B:4:ACY:CH3	2.32	0.77
1:B:145:PRO:HA	1:B:148:GLN:HE22	1.50	0.76
5:B:2:ACY:O	5:B:4:ACY:H2	1.87	0.74
5:A:213:ACY:C	5:A:3:ACY:H3	2.18	0.73
1:A:167:ASP:OD2	6:A:287:HOH:O	2.08	0.71
1:B:145:PRO:O	1:B:148:GLN:NE2	2.24	0.71
5:A:213:ACY:OXT	5:A:3:ACY:H3	1.89	0.71
1:A:60[B]:ILE:HD13	1:A:112[B]:THR:HG23	1.69	0.70
5:B:2:ACY:C	5:B:4:ACY:H2	2.19	0.70
1:A:185:HIS:CE1	5:A:213:ACY:H2	2.28	0.69
1:A:57:SER:HB3	6:A:284:HOH:O	1.93	0.68
1:A:202[B]:ASP:OD1	1:B:209:GLN:NE2	2.28	0.67
1:A:152:GLU:OE2	6:A:229:HOH:O	2.13	0.67
1:B:86:ALA:C	1:B:87[B]:GLU:HG2	2.16	0.64
1:A:60[B]:ILE:HD12	1:A:112[B]:THR:HG21	1.82	0.62
1:A:185:HIS:CG	5:A:213:ACY:H2	2.35	0.61
5:B:2:ACY:C	5:B:4:ACY:H3	2.27	0.61
1:A:185:HIS:CD2	5:A:213:ACY:H2	2.36	0.61
1:A:187:ARG:NH1	1:A:187:ARG:CG	2.61	0.60
1:A:87[B]:GLU:OE2	1:B:103[B]:LYS:NZ	2.34	0.60
1:A:141[B]:ILE:HG23	1:A:143:TYR:CZ	2.37	0.59
1:B:57:SER:HB3	1:B:58:PRO:HD3	1.85	0.59
1:B:187:ARG:O	1:B:188:LYS:HG2	2.03	0.58
6:A:282:HOH:O	1:B:107:ARG:CG	2.49	0.57
1:B:86:ALA:O	1:B:87[B]:GLU:HG2	2.07	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:213:ACY:O	5:A:3:ACY:CH3	2.53	0.54
5:A:213:ACY:O	5:A:3:ACY:H1	2.06	0.54
1:B:136[B]:LYS:HE2	1:B:138:GLU:OE1	2.07	0.54
1:A:209:GLN:NE2	1:B:202[B]:ASP:OD1	2.41	0.53
1:B:181:PHE:O	1:B:185:HIS:HD2	1.92	0.53
5:B:2:ACY:OXT	5:B:4:ACY:H3	2.05	0.52
1:A:141[B]:ILE:CG2	1:A:143:TYR:CZ	2.94	0.51
1:B:148:GLN:NE2	1:B:148:GLN:H	2.02	0.50
1:B:63[B]:LEU:HD11	1:B:74:LEU:HD11	1.93	0.50
5:A:213:ACY:OXT	5:A:3:ACY:H2	2.08	0.49
5:A:213:ACY:H3	6:A:214:HOH:O	2.13	0.49
1:B:142:PRO:HB2	1:B:148:GLN:NE2	2.28	0.48
1:A:141[B]:ILE:CG2	1:A:143:TYR:CE1	2.91	0.47
1:B:63[B]:LEU:HD23	1:B:116:ASP:OD1	2.14	0.47
1:A:103[B]:LYS:HG2	1:B:177:GLN:HG3	1.98	0.46
1:B:57:SER:CB	1:B:58:PRO:HD3	2.47	0.45
1:A:202[B]:ASP:CG	1:B:209:GLN:HE22	2.21	0.45
1:B:187:ARG:C	1:B:188:LYS:HG2	2.38	0.44
1:A:185:HIS:NE2	5:A:213:ACY:H2	2.32	0.44
1:B:187:ARG:O	1:B:188:LYS:CG	2.66	0.43
1:A:185:HIS:ND1	5:A:213:ACY:H2	2.34	0.42
1:B:145:PRO:CA	1:B:148:GLN:HE22	2.25	0.42
1:B:103[B]:LYS:HE2	3:B:8:SO4:S	2.60	0.40
1:A:102[B]:LEU:HD23	1:A:102[B]:LEU:HA	1.88	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	160/183 (87%)	159 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	159/183 (87%)	158 (99%)	1 (1%)	0	100	100
2	D	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
2	F	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
All	All	331/382 (87%)	327 (99%)	4 (1%)	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	138/150 (92%)	134 (97%)	4 (3%)	42	23
1	B	137/150 (91%)	135 (98%)	2 (2%)	65	51
2	D	8/8 (100%)	8 (100%)	0	100	100
2	F	8/8 (100%)	8 (100%)	0	100	100
All	All	291/316 (92%)	285 (98%)	6 (2%)	50	36

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	SER
1	A	137	GLN
1	A	187	ARG
1	A	209	GLN
1	B	148	GLN
1	B	188	LYS

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

Mol	Chain	Res	Type
1	A	137	GLN
1	A	171	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	146	GLN
1	B	148	GLN
1	B	185	HIS
1	B	209	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

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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	$\#\|Z\| > 2$
5	ACY	A	213	-	1,3,3	2.02	1 (100%)	0,3,3	0.00	-
3	SO4	B	5	-	4,4,4	0.39	0	6,6,6	0.70	0
3	SO4	B	1	-	4,4,4	0.57	0	6,6,6	0.73	0
5	ACY	B	4	-	1,3,3	2.10	1 (100%)	0,3,3	0.00	-
3	SO4	A	7	-	4,4,4	0.53	0	6,6,6	1.62	2 (33%)
3	SO4	B	3	-	4,4,4	0.24	0	6,6,6	0.32	0
5	ACY	A	3	-	1,3,3	1.89	0	0,3,3	0.00	-
3	SO4	B	8	-	4,4,4	0.44	0	6,6,6	1.27	1 (16%)
3	SO4	A	4	-	4,4,4	0.31	0	6,6,6	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	ACY	B	2	-	1,3,3	1.19	0	0,3,3	0.00	-
3	SO4	A	6	-	4,4,4	0.22	0	6,6,6	0.55	0
3	SO4	A	2	-	4,4,4	0.35	0	6,6,6	1.31	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	4	ACY	CH3-C	-2.10	1.46	1.48
5	A	213	ACY	CH3-C	-2.02	1.46	1.48

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	7	SO4	O3-S-O2	-2.80	94.68	109.31
3	B	8	SO4	O4-S-O1	-2.67	95.38	109.31
3	A	7	SO4	O4-S-O3	2.42	119.38	109.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	213	ACY	14	0
5	B	4	ACY	9	0
5	A	3	ACY	8	0
3	B	8	SO4	1	0
5	B	2	ACY	10	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	149/183 (81%)	-0.55	1 (0%) 87 90	12, 18, 42, 58	17 (11%)
1	B	149/183 (81%)	-0.57	1 (0%) 87 90	11, 18, 40, 60	18 (12%)
2	D	8/8 (100%)	-0.67	0 100 100	18, 24, 30, 30	0
2	F	8/8 (100%)	-0.44	0 100 100	19, 24, 30, 30	0
All	All	314/382 (82%)	-0.56	2 (0%) 89 91	11, 19, 42, 60	35 (11%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	209	GLN	2.3
1	A	188	LYS	2.3

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	B-factors(Å^2)	Q<0.9
5	ACY	B	4	4/4	0.80	0.17	32,33,33,34	4
5	ACY	A	3	4/4	0.90	0.18	30,31,33,33	4
5	ACY	A	213	4/4	0.91	0.12	44,44,46,56	0
5	ACY	B	2	4/4	0.94	0.15	46,51,53,60	0
3	SO4	A	7	5/5	0.95	0.11	23,34,38,38	5
4	CL	A	1	1/1	0.96	0.06	48,48,48,48	0
3	SO4	B	3	5/5	0.97	0.08	25,33,37,46	5
3	SO4	B	8	5/5	0.98	0.10	25,34,37,39	5
3	SO4	A	4	5/5	0.98	0.07	25,30,33,43	5
3	SO4	B	1	5/5	0.98	0.08	26,30,40,45	5
3	SO4	B	5	5/5	0.99	0.13	14,18,19,20	5
3	SO4	A	6	5/5	0.99	0.13	14,18,18,19	5
3	SO4	A	2	5/5	0.99	0.06	26,30,39,40	5

6.5 Other polymers (i)

There are no such residues in this entry.