Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 06:34 pm BST

PDB ID	:	5NPW
Title	:	Structure of human ATG5-ATG16L1(ATG5BD) complex (C2)
Authors	:	Archna, A.; Scrima, A.
Deposited on	:	2017-04-19
Resolution	:	3.10 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 ⓘ) were used in the production of this report:

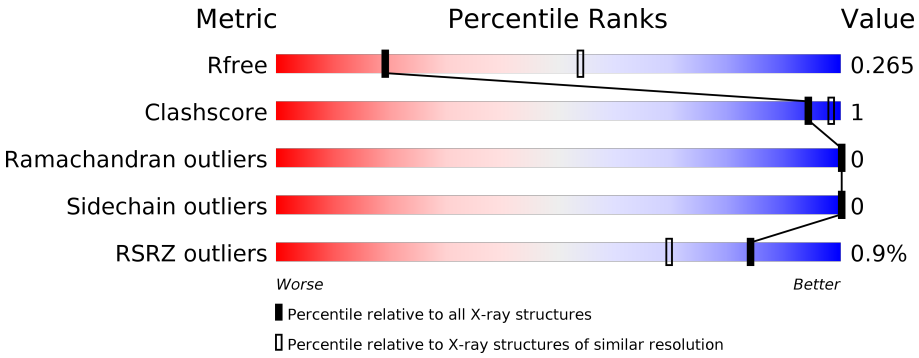
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	282	90% 5% 6%
1	C	282	% 92% 5%
1	E	282	% 91% 6%
1	G	282	89% 7%
2	B	301	13% 87%
2	D	301	12% 87%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	301	
2	H	301	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 19441 atoms, of which 9489 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 Autophagy protein 5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	266	Total	C	H	N	O	S	0	0	0
			4205	1406	2049	346	392	12			
1	C	268	Total	C	H	N	O	S	0	0	0
			4253	1417	2076	351	396	13			
1	E	266	Total	C	H	N	O	S	0	0	0
			4131	1388	1999	347	385	12			
1	G	261	Total	C	H	N	O	S	0	0	0
			4084	1364	1984	342	381	13			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP Q9H1Y0
A	-5	ALA	-	expression tag	UNP Q9H1Y0
A	-4	HIS	-	expression tag	UNP Q9H1Y0
A	-3	MET	-	expression tag	UNP Q9H1Y0
A	-2	SER	-	expression tag	UNP Q9H1Y0
A	-1	GLY	-	expression tag	UNP Q9H1Y0
A	0	ARG	-	expression tag	UNP Q9H1Y0
C	-6	GLY	-	expression tag	UNP Q9H1Y0
C	-5	ALA	-	expression tag	UNP Q9H1Y0
C	-4	HIS	-	expression tag	UNP Q9H1Y0
C	-3	MET	-	expression tag	UNP Q9H1Y0
C	-2	SER	-	expression tag	UNP Q9H1Y0
C	-1	GLY	-	expression tag	UNP Q9H1Y0
C	0	ARG	-	expression tag	UNP Q9H1Y0
E	-6	GLY	-	expression tag	UNP Q9H1Y0
E	-5	ALA	-	expression tag	UNP Q9H1Y0
E	-4	HIS	-	expression tag	UNP Q9H1Y0
E	-3	MET	-	expression tag	UNP Q9H1Y0
E	-2	SER	-	expression tag	UNP Q9H1Y0
E	-1	GLY	-	expression tag	UNP Q9H1Y0
E	0	ARG	-	expression tag	UNP Q9H1Y0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-6	GLY	-	expression tag	UNP Q9H1Y0
G	-5	ALA	-	expression tag	UNP Q9H1Y0
G	-4	HIS	-	expression tag	UNP Q9H1Y0
G	-3	MET	-	expression tag	UNP Q9H1Y0
G	-2	SER	-	expression tag	UNP Q9H1Y0
G	-1	GLY	-	expression tag	UNP Q9H1Y0
G	0	ARG	-	expression tag	UNP Q9H1Y0

- Molecule 2 is a protein called Autophagy-related protein 16-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	40	Total	C	H	N	O	0	0	0
			652	209	320	65	58			
2	D	40	Total	C	H	N	O	0	0	0
			700	216	352	72	60			
2	F	39	Total	C	H	N	O	0	0	0
			673	212	335	67	59			
2	H	41	Total	C	H	N	O	0	0	0
			736	225	374	76	61			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	7	GLY	-	expression tag	UNP Q676U5
B	8	GLY	-	expression tag	UNP Q676U5
B	9	GLY	-	expression tag	UNP Q676U5
B	10	ARG	-	expression tag	UNP Q676U5
D	7	GLY	-	expression tag	UNP Q676U5
D	8	GLY	-	expression tag	UNP Q676U5
D	9	GLY	-	expression tag	UNP Q676U5
D	10	ARG	-	expression tag	UNP Q676U5
F	7	GLY	-	expression tag	UNP Q676U5
F	8	GLY	-	expression tag	UNP Q676U5
F	9	GLY	-	expression tag	UNP Q676U5
F	10	ARG	-	expression tag	UNP Q676U5
H	7	GLY	-	expression tag	UNP Q676U5
H	8	GLY	-	expression tag	UNP Q676U5
H	9	GLY	-	expression tag	UNP Q676U5
H	10	ARG	-	expression tag	UNP Q676U5

- Molecule 3 is water.

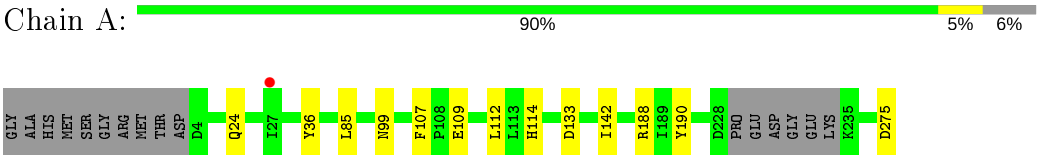
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	C	2	Total O 2 2	0	0
3	E	1	Total O 1 1	0	0
3	G	2	Total O 2 2	0	0

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.

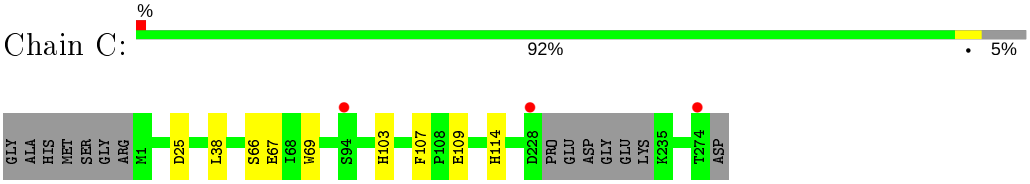
- Molecule 1: Autophagy protein 5

Chain A:



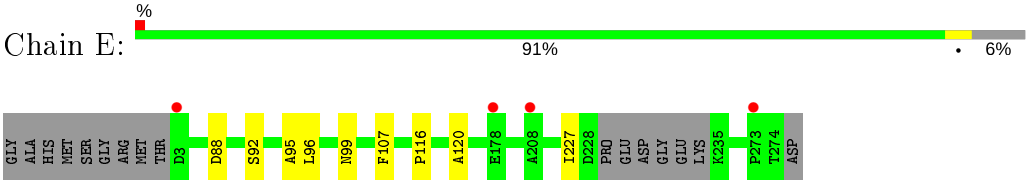
- Molecule 1: Autophagy protein 5

Chain C:



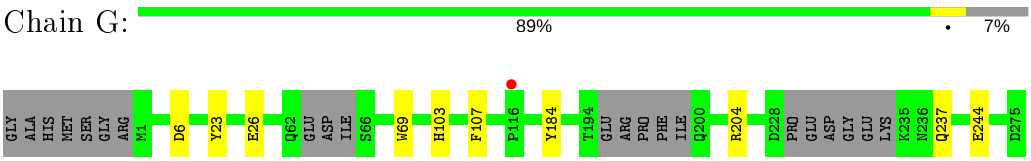
- Molecule 1: Autophagy protein 5

Chain E:



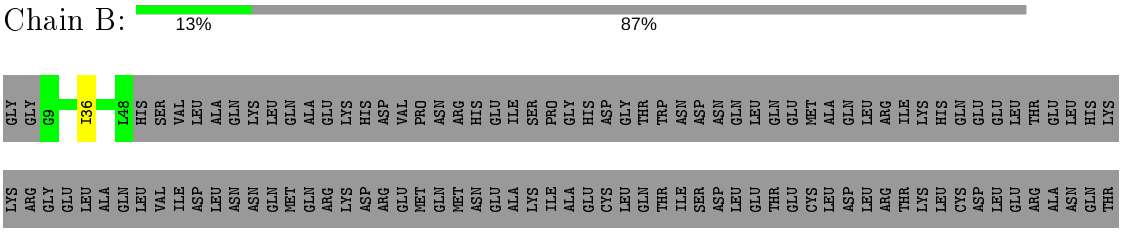
- Molecule 1: Autophagy protein 5

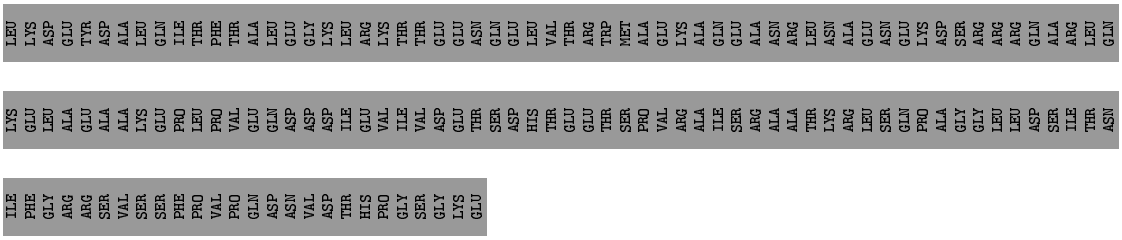
Chain G:



- Molecule 2: Autophagy-related protein 16-1

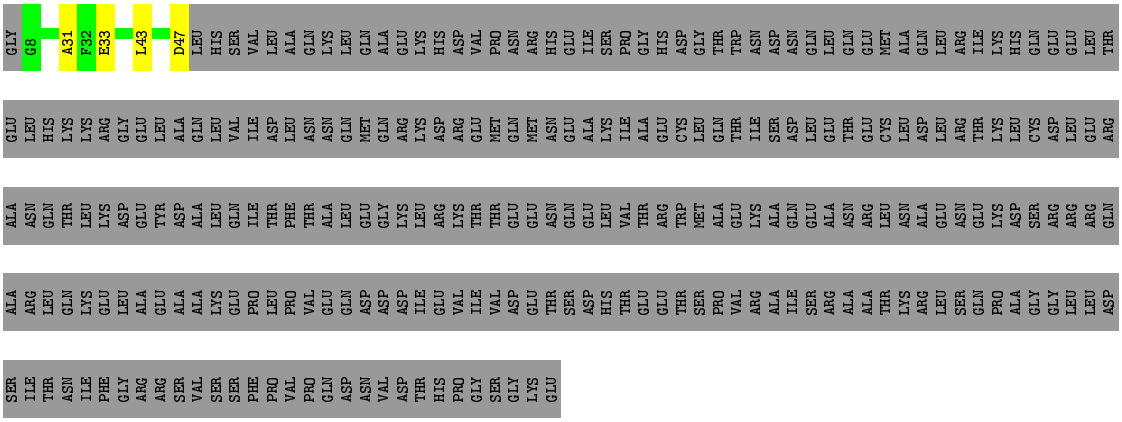
Chain B:





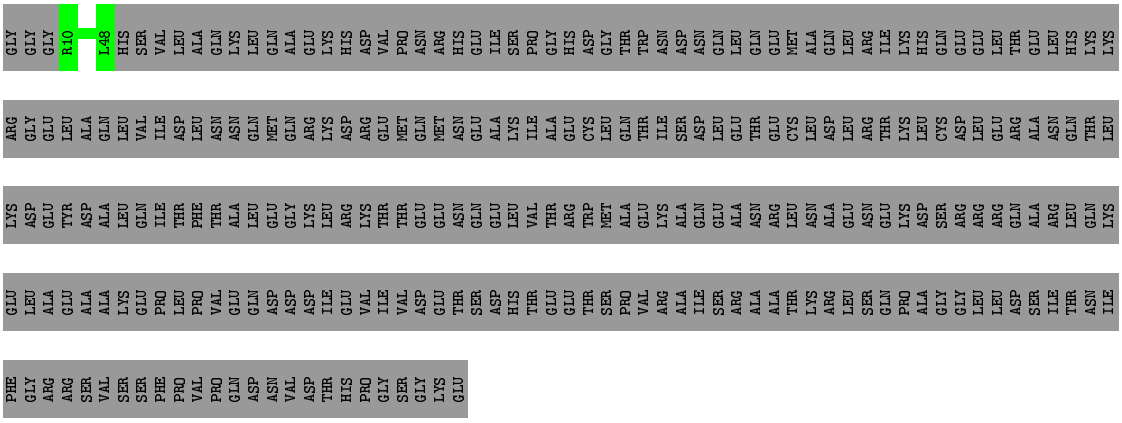
- Molecule 2: Autophagy-related protein 16-1

Chain D:



- Molecule 2: Autophagy-related protein 16-1

Chain F:



- Molecule 2: Autophagy-related protein 16-1

Chain H:

THR LEU LYS ASP GLU TYR ASP ALA LEU GLN ILE THR PHE THR ALA LEU GLU GLY LYS LEU ARG LYS THR THR GLU GLU ASN GLN GLU LEU VAL THR ARG TRP MET ALA GLU LYS ALA GLN GLU ALA ASN ARG LEU ASN ALA GLU ASN GLU LYS ASP SER ARG ARG ARG GLN ALA ARG LEU

GLN LYS GLU LEU ALA GLU ALA ALA LYS GLU PRO LEU PRO VAL GLU GLN ASP ASP ASP ILE GLU VAL ILE VAL ASP GLU THR SER ASP HIS THR GLU GLU THR SER PRO VAL ARG ALA ILE SER ARG ALA ALA THR LYS ARG LEU SER GLN PRO ALA GLY GLY LEU LEU ASP SER ILE THR

ASN ILE PHE GLY ARG ARG SER VAL SER SER PHE PRO VAL PRO GLN ASP ASN VAL ASP THR HIS PRO GLY SER GLY LYS GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	200.86Å 75.99Å 142.44Å 90.00° 131.41° 90.00°	Depositor
Resolution (Å)	48.32 – 3.10 48.39 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.32-3.10) 99.6 (48.39-3.10)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.217 , 0.264 0.220 , 0.265	Depositor DCC
R_{free} test set	1472 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	62.2	Xtriage
Anisotropy	0.502	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 55.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	19441	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.64% 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

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	0.29	1/2222 (0.0%)	0.41	0/3026
1	C	0.31	1/2243 (0.0%)	0.42	0/3055
1	E	0.30	1/2198 (0.0%)	0.41	0/3001
1	G	0.31	1/2162 (0.0%)	0.41	0/2943
2	B	0.23	0/337	0.35	0/451
2	D	0.23	0/353	0.39	0/469
2	F	0.23	0/343	0.40	0/459
2	H	0.23	0/367	0.37	0/487
All	All	0.29	4/10225 (0.0%)	0.41	0/13891

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	107	PHE	C-N	9.36	1.52	1.34
1	C	107	PHE	C-N	8.50	1.50	1.34
1	A	107	PHE	C-N	7.82	1.49	1.34
1	E	107	PHE	C-N	7.75	1.49	1.34

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	2156	2049	2045	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2177	2076	2073	5	0
1	E	2132	1999	1998	6	0
1	G	2100	1984	1981	5	0
2	B	332	320	320	1	0
2	D	348	352	352	3	0
2	F	338	335	335	0	0
2	H	362	374	374	2	0
3	A	2	0	0	2	0
3	C	2	0	0	0	0
3	E	1	0	0	2	0
3	G	2	0	0	0	0
All	All	9952	9489	9478	26	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 1.

All (26) 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:275:ASP:OD2	2:H:26:ARG:NH1	2.10	0.85
1:E:99:ASN:OD1	3:E:301:HOH:O	2.06	0.73
1:A:24:GLN:OE1	1:A:24:GLN:N	2.23	0.71
1:A:99:ASN:ND2	3:A:302:HOH:O	2.24	0.66
1:G:6:ASP:OD1	2:H:10:ARG:NH2	2.39	0.55
1:G:237:GLN:NE2	1:G:244:GLU:OE2	2.41	0.54
1:E:95:ALA:O	3:E:301:HOH:O	2.18	0.53
1:A:112:LEU:O	3:A:301:HOH:O	2.19	0.51
1:E:227:ILE:O	1:E:227:ILE:HG22	2.13	0.49
1:A:188:ARG:NH1	1:C:25:ASP:OD2	2.49	0.46
1:A:133:ASP:OD2	1:A:142:ILE:N	2.38	0.46
1:C:69:TRP:CE2	1:C:103:HIS:HB2	2.52	0.45
1:E:116:PRO:HB2	1:E:120:ALA:HB3	2.00	0.43
1:A:109:GLU:OE1	1:A:114:HIS:NE2	2.51	0.43
1:G:184:TYR:OH	1:G:204:ARG:NH2	2.50	0.43
1:A:85:LEU:HD21	1:A:190:TYR:CZ	2.54	0.43
2:D:43:LEU:O	2:D:47:ASP:HB2	2.20	0.42
1:G:69:TRP:CE2	1:G:103:HIS:HB2	2.54	0.42
1:A:275:ASP:OD2	1:A:275:ASP:O	2.38	0.42
1:C:109:GLU:OE1	1:C:114:HIS:NE2	2.53	0.42
1:C:38:LEU:HD11	2:D:33:GLU:HA	2.02	0.42
2:D:31:ALA:HA	1:E:96:LEU:O	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:36:TYR:HB3	2:B:36:ILE:CD1	2.50	0.41
1:E:88:ASP:HA	1:E:92:SER:HB3	2.02	0.41
1:G:23:TYR:HB3	1:G:26:GLU:HG2	2.03	0.41
1:C:66:SER:OG	1:C:67:GLU:N	2.54	0.41

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	262/282 (93%)	258 (98%)	4 (2%)	0	100	100
1	C	264/282 (94%)	258 (98%)	6 (2%)	0	100	100
1	E	262/282 (93%)	255 (97%)	7 (3%)	0	100	100
1	G	253/282 (90%)	245 (97%)	8 (3%)	0	100	100
2	B	38/301 (13%)	36 (95%)	2 (5%)	0	100	100
2	D	38/301 (13%)	36 (95%)	2 (5%)	0	100	100
2	F	37/301 (12%)	36 (97%)	1 (3%)	0	100	100
2	H	39/301 (13%)	37 (95%)	2 (5%)	0	100	100
All	All	1193/2332 (51%)	1161 (97%)	32 (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	228/258 (88%)	228 (100%)	0	100	100
1	C	232/258 (90%)	232 (100%)	0	100	100
1	E	223/258 (86%)	223 (100%)	0	100	100
1	G	222/258 (86%)	222 (100%)	0	100	100
2	B	32/267 (12%)	32 (100%)	0	100	100
2	D	36/267 (14%)	36 (100%)	0	100	100
2	F	35/267 (13%)	35 (100%)	0	100	100
2	H	38/267 (14%)	38 (100%)	0	100	100
All	All	1046/2100 (50%)	1046 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

There are no ligands in this entry.

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	266/282 (94%)	-0.23	1 (0%) 92 84	31, 57, 108, 151	0
1	C	268/282 (95%)	-0.25	3 (1%) 80 64	33, 55, 92, 121	0
1	E	266/282 (94%)	-0.10	4 (1%) 73 54	36, 65, 110, 171	0
1	G	261/282 (92%)	-0.19	1 (0%) 92 84	35, 59, 110, 160	0
2	B	40/301 (13%)	-0.12	0 100 100	40, 65, 107, 119	0
2	D	40/301 (13%)	-0.29	0 100 100	41, 54, 89, 99	0
2	F	39/301 (12%)	0.03	0 100 100	35, 62, 102, 110	0
2	H	41/301 (13%)	-0.06	2 (4%) 29 14	41, 58, 104, 126	0
All	All	1221/2332 (52%)	-0.18	11 (0%) 84 69	31, 59, 108, 171	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	116	PRO	4.3
1	E	3	ASP	4.0
1	C	274	THR	3.9
2	H	8	GLY	2.9
1	A	27	ILE	2.7
1	E	208	ALA	2.6
2	H	9	GLY	2.5
1	C	228	ASP	2.4
1	E	273	PRO	2.2
1	E	178	GLU	2.2
1	C	94	SER	2.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.