



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 28, 2024 – 11:26 PM EDT

PDB ID : 2QGV
Title : Crystal structure of hydrogenase-1 operon protein hyaE from *Shigella flexneri*.
Northeast Structural Genomics Consortium Target Sfr170
Authors : Forouhar, F.; Su, M.; Seetharaman, J.; Benach, J.; Wang, D.; Fang, Y.; Cunningham, K.; Ma, L.-C.; Xiao, R.; Liu, J.; Baran, M.C.; Acton, T.B.; Rost, B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2007-06-29
Resolution : 2.70 Å (reported)

This is a wwPDB X-ray Structure Validation Summary 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

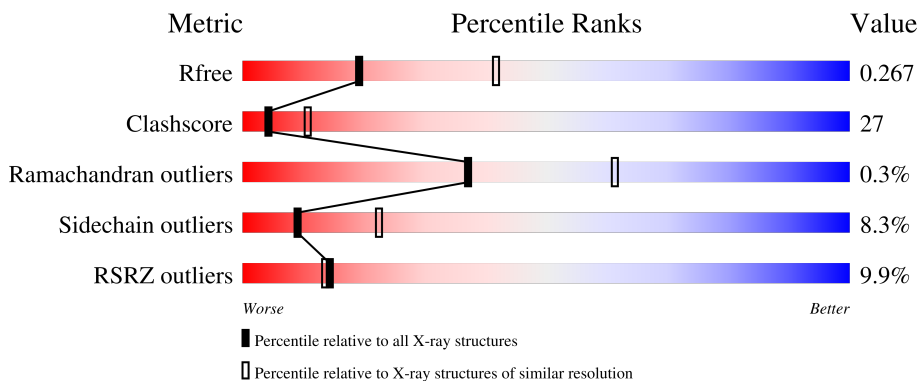
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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 of 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 $\leq 5\%$. 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	140	
1	B	140	
1	C	140	
1	D	140	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	140	
1	F	140	
1	G	140	
1	H	140	
1	I	140	
1	J	140	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9610 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 Hydrogenase-1 operon protein hyaE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	121	961	617	163	178	3	0	0	0
1	B	121	961	617	163	178	3	0	0	0
1	C	121	961	617	163	178	3	0	0	0
1	D	121	961	617	163	178	3	0	0	0
1	E	121	961	617	163	178	3	0	0	0
1	F	121	961	617	163	178	3	0	0	0
1	G	121	961	617	163	178	3	0	0	0
1	H	121	961	617	163	178	3	0	0	0
1	I	121	961	617	163	178	3	0	0	0
1	J	121	961	617	163	178	3	0	0	0

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP Q83RW6
A	14	MSE	MET	modified residue	UNP Q83RW6
A	57	MSE	MET	modified residue	UNP Q83RW6
A	120	MSE	MET	modified residue	UNP Q83RW6
A	133	LEU	-	cloning artifact	UNP Q83RW6
A	134	GLU	-	cloning artifact	UNP Q83RW6
A	135	HIS	-	cloning artifact	UNP Q83RW6
A	136	HIS	-	cloning artifact	UNP Q83RW6
A	137	HIS	-	cloning artifact	UNP Q83RW6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	HIS	-	cloning artifact	UNP Q83RW6
A	139	HIS	-	cloning artifact	UNP Q83RW6
A	140	HIS	-	cloning artifact	UNP Q83RW6
B	1	MSE	MET	modified residue	UNP Q83RW6
B	14	MSE	MET	modified residue	UNP Q83RW6
B	57	MSE	MET	modified residue	UNP Q83RW6
B	120	MSE	MET	modified residue	UNP Q83RW6
B	133	LEU	-	cloning artifact	UNP Q83RW6
B	134	GLU	-	cloning artifact	UNP Q83RW6
B	135	HIS	-	cloning artifact	UNP Q83RW6
B	136	HIS	-	cloning artifact	UNP Q83RW6
B	137	HIS	-	cloning artifact	UNP Q83RW6
B	138	HIS	-	cloning artifact	UNP Q83RW6
B	139	HIS	-	cloning artifact	UNP Q83RW6
B	140	HIS	-	cloning artifact	UNP Q83RW6
C	1	MSE	MET	modified residue	UNP Q83RW6
C	14	MSE	MET	modified residue	UNP Q83RW6
C	57	MSE	MET	modified residue	UNP Q83RW6
C	120	MSE	MET	modified residue	UNP Q83RW6
C	133	LEU	-	cloning artifact	UNP Q83RW6
C	134	GLU	-	cloning artifact	UNP Q83RW6
C	135	HIS	-	cloning artifact	UNP Q83RW6
C	136	HIS	-	cloning artifact	UNP Q83RW6
C	137	HIS	-	cloning artifact	UNP Q83RW6
C	138	HIS	-	cloning artifact	UNP Q83RW6
C	139	HIS	-	cloning artifact	UNP Q83RW6
C	140	HIS	-	cloning artifact	UNP Q83RW6
D	1	MSE	MET	modified residue	UNP Q83RW6
D	14	MSE	MET	modified residue	UNP Q83RW6
D	57	MSE	MET	modified residue	UNP Q83RW6
D	120	MSE	MET	modified residue	UNP Q83RW6
D	133	LEU	-	cloning artifact	UNP Q83RW6
D	134	GLU	-	cloning artifact	UNP Q83RW6
D	135	HIS	-	cloning artifact	UNP Q83RW6
D	136	HIS	-	cloning artifact	UNP Q83RW6
D	137	HIS	-	cloning artifact	UNP Q83RW6
D	138	HIS	-	cloning artifact	UNP Q83RW6
D	139	HIS	-	cloning artifact	UNP Q83RW6
D	140	HIS	-	cloning artifact	UNP Q83RW6
E	1	MSE	MET	modified residue	UNP Q83RW6
E	14	MSE	MET	modified residue	UNP Q83RW6
E	57	MSE	MET	modified residue	UNP Q83RW6

Continued on next page...

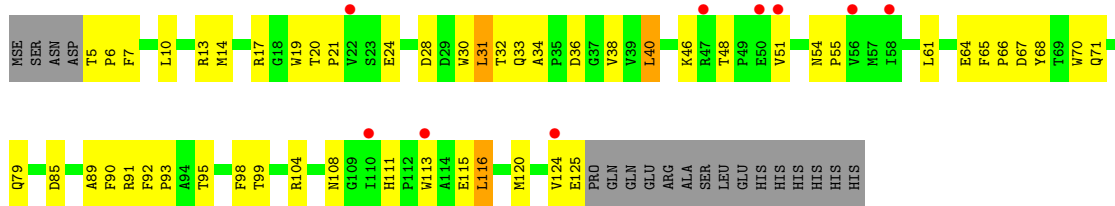
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	120	MSE	MET	modified residue	UNP Q83RW6
E	133	LEU	-	cloning artifact	UNP Q83RW6
E	134	GLU	-	cloning artifact	UNP Q83RW6
E	135	HIS	-	cloning artifact	UNP Q83RW6
E	136	HIS	-	cloning artifact	UNP Q83RW6
E	137	HIS	-	cloning artifact	UNP Q83RW6
E	138	HIS	-	cloning artifact	UNP Q83RW6
E	139	HIS	-	cloning artifact	UNP Q83RW6
E	140	HIS	-	cloning artifact	UNP Q83RW6
F	1	MSE	MET	modified residue	UNP Q83RW6
F	14	MSE	MET	modified residue	UNP Q83RW6
F	57	MSE	MET	modified residue	UNP Q83RW6
F	120	MSE	MET	modified residue	UNP Q83RW6
F	133	LEU	-	cloning artifact	UNP Q83RW6
F	134	GLU	-	cloning artifact	UNP Q83RW6
F	135	HIS	-	cloning artifact	UNP Q83RW6
F	136	HIS	-	cloning artifact	UNP Q83RW6
F	137	HIS	-	cloning artifact	UNP Q83RW6
F	138	HIS	-	cloning artifact	UNP Q83RW6
F	139	HIS	-	cloning artifact	UNP Q83RW6
F	140	HIS	-	cloning artifact	UNP Q83RW6
G	1	MSE	MET	modified residue	UNP Q83RW6
G	14	MSE	MET	modified residue	UNP Q83RW6
G	57	MSE	MET	modified residue	UNP Q83RW6
G	120	MSE	MET	modified residue	UNP Q83RW6
G	133	LEU	-	cloning artifact	UNP Q83RW6
G	134	GLU	-	cloning artifact	UNP Q83RW6
G	135	HIS	-	cloning artifact	UNP Q83RW6
G	136	HIS	-	cloning artifact	UNP Q83RW6
G	137	HIS	-	cloning artifact	UNP Q83RW6
G	138	HIS	-	cloning artifact	UNP Q83RW6
G	139	HIS	-	cloning artifact	UNP Q83RW6
G	140	HIS	-	cloning artifact	UNP Q83RW6
H	1	MSE	MET	modified residue	UNP Q83RW6
H	14	MSE	MET	modified residue	UNP Q83RW6
H	57	MSE	MET	modified residue	UNP Q83RW6
H	120	MSE	MET	modified residue	UNP Q83RW6
H	133	LEU	-	cloning artifact	UNP Q83RW6
H	134	GLU	-	cloning artifact	UNP Q83RW6
H	135	HIS	-	cloning artifact	UNP Q83RW6
H	136	HIS	-	cloning artifact	UNP Q83RW6
H	137	HIS	-	cloning artifact	UNP Q83RW6

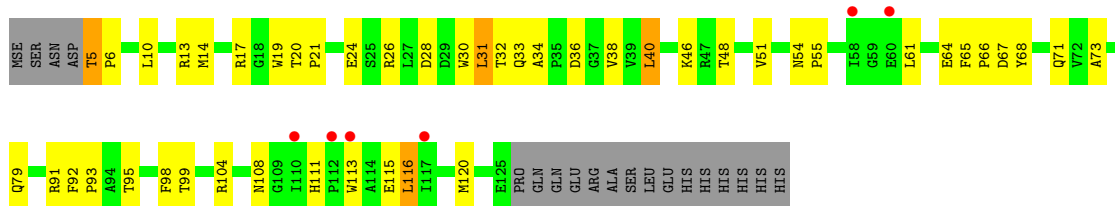
Continued on next page...

Continued from previous page...

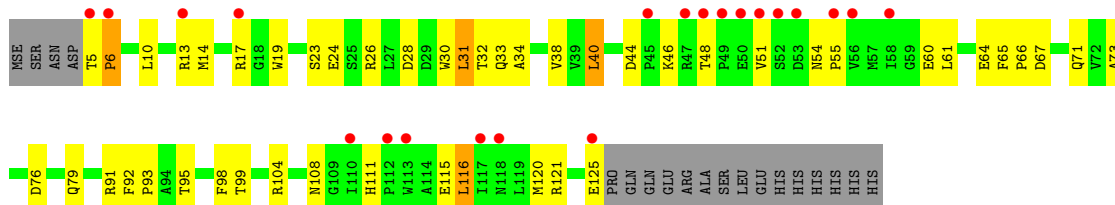
Chain	Residue	Modelled	Actual	Comment	Reference
H	138	HIS	-	cloning artifact	UNP Q83RW6
H	139	HIS	-	cloning artifact	UNP Q83RW6
H	140	HIS	-	cloning artifact	UNP Q83RW6
I	1	MSE	MET	modified residue	UNP Q83RW6
I	14	MSE	MET	modified residue	UNP Q83RW6
I	57	MSE	MET	modified residue	UNP Q83RW6
I	120	MSE	MET	modified residue	UNP Q83RW6
I	133	LEU	-	cloning artifact	UNP Q83RW6
I	134	GLU	-	cloning artifact	UNP Q83RW6
I	135	HIS	-	cloning artifact	UNP Q83RW6
I	136	HIS	-	cloning artifact	UNP Q83RW6
I	137	HIS	-	cloning artifact	UNP Q83RW6
I	138	HIS	-	cloning artifact	UNP Q83RW6
I	139	HIS	-	cloning artifact	UNP Q83RW6
I	140	HIS	-	cloning artifact	UNP Q83RW6
J	1	MSE	MET	modified residue	UNP Q83RW6
J	14	MSE	MET	modified residue	UNP Q83RW6
J	57	MSE	MET	modified residue	UNP Q83RW6
J	120	MSE	MET	modified residue	UNP Q83RW6
J	133	LEU	-	cloning artifact	UNP Q83RW6
J	134	GLU	-	cloning artifact	UNP Q83RW6
J	135	HIS	-	cloning artifact	UNP Q83RW6
J	136	HIS	-	cloning artifact	UNP Q83RW6
J	137	HIS	-	cloning artifact	UNP Q83RW6
J	138	HIS	-	cloning artifact	UNP Q83RW6
J	139	HIS	-	cloning artifact	UNP Q83RW6
J	140	HIS	-	cloning artifact	UNP Q83RW6



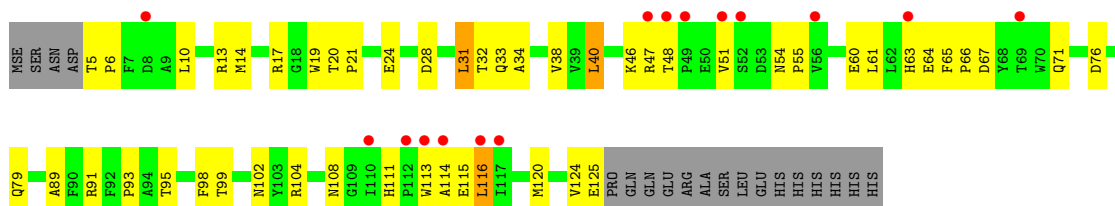
• Molecule 1: Hydrogenase-1 operon protein hyaE



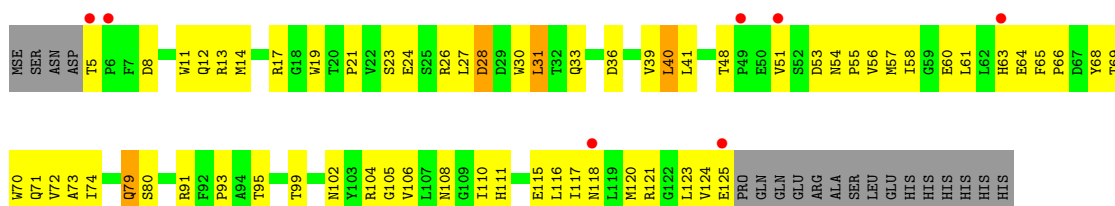
• Molecule 1: Hydrogenase-1 operon protein hyaE



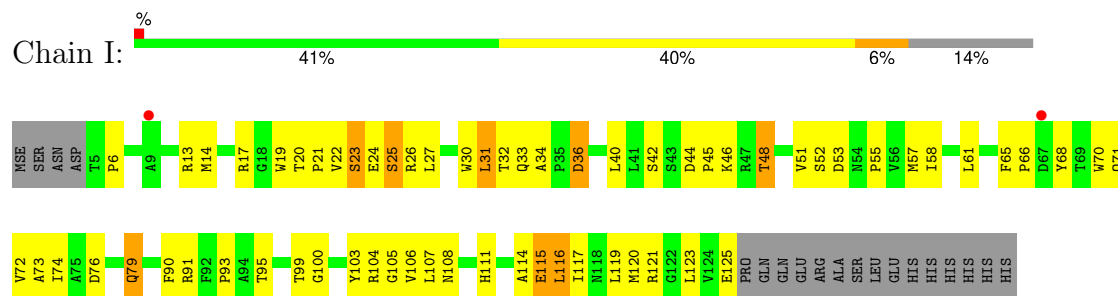
• Molecule 1: Hydrogenase-1 operon protein hyaE



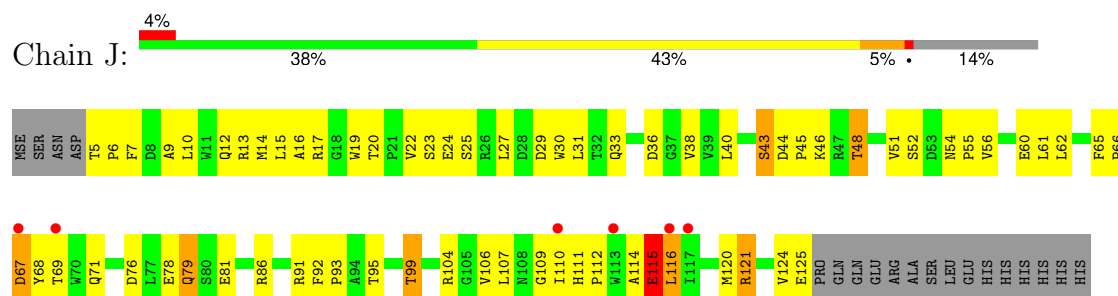
• Molecule 1: Hydrogenase-1 operon protein hyaE



- Molecule 1: Hydrogenase-1 operon protein hyaE



- Molecule 1: Hydrogenase-1 operon protein hyaE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.17Å 215.78Å 64.07Å 90.00° 94.80° 90.00°	Depositor
Resolution (Å)	48.99 – 2.70 48.99 – 2.70	Depositor EDS
% Data completeness (in resolution range)	78.0 (48.99-2.70) 86.4 (48.99-2.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.41 (at 2.58Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.253 , 0.292 0.272 , 0.267	Depositor DCC
R_{free} test set	3653 reflections (9.63%)	wwPDB-VP
Wilson B-factor (Å ²)	50.4	Xtrriage
Anisotropy	0.681	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 36.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	9610	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.46	0/986	0.58	0/1343
1	B	0.42	0/986	0.57	0/1343
1	C	0.44	0/986	0.57	0/1343
1	D	0.44	0/986	0.57	0/1343
1	E	0.43	0/986	0.58	0/1343
1	F	0.44	0/986	0.57	0/1343
1	G	0.42	0/986	0.56	0/1343
1	H	0.50	0/986	0.69	0/1343
1	I	0.54	0/986	0.69	0/1343
1	J	0.56	0/986	0.72	0/1343
All	All	0.47	0/9860	0.61	0/13430

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [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	961	0	926	62	0
1	B	961	0	926	55	1
1	C	961	0	926	39	0
1	D	961	0	926	48	0
1	E	961	0	926	45	0
1	F	961	0	926	46	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	961	0	926	59	1
1	H	961	0	926	62	0
1	I	961	0	926	58	0
1	J	961	0	926	79	0
All	All	9610	0	9260	502	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 27.

The worst 5 of 502 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:I:14:MSE:HE3	1:I:55:PRO:HG3	1.26	1.09
1:H:14:MSE:HE2	1:H:40:LEU:HD23	1.40	1.04
1:H:14:MSE:HE3	1:H:55:PRO:HB3	1.42	0.99
1:H:61:LEU:HD21	1:H:120:MSE:HG3	1.42	0.98
1:A:113:TRP:CE2	1:B:13:ARG:HG2	1.99	0.97

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)
1:B:91:ARG:NH2	1:G:76:ASP:OD1[1_655]	2.12	0.08

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	119/140 (85%)	113 (95%)	6 (5%)	0	100 100
1	B	119/140 (85%)	114 (96%)	5 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	119/140 (85%)	114 (96%)	5 (4%)	0	100	100
1	D	119/140 (85%)	115 (97%)	4 (3%)	0	100	100
1	E	119/140 (85%)	112 (94%)	7 (6%)	0	100	100
1	F	119/140 (85%)	114 (96%)	4 (3%)	1 (1%)	16	38
1	G	119/140 (85%)	114 (96%)	5 (4%)	0	100	100
1	H	119/140 (85%)	107 (90%)	11 (9%)	1 (1%)	16	38
1	I	119/140 (85%)	108 (91%)	10 (8%)	1 (1%)	16	38
1	J	119/140 (85%)	103 (87%)	15 (13%)	1 (1%)	16	38
All	All	1190/1400 (85%)	1114 (94%)	72 (6%)	4 (0%)	37	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	25	SER
1	J	115	GLU
1	F	6	PRO
1	H	70	TRP

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	102/116 (88%)	95 (93%)	7 (7%)	13	31
1	B	102/116 (88%)	95 (93%)	7 (7%)	13	31
1	C	102/116 (88%)	95 (93%)	7 (7%)	13	31
1	D	102/116 (88%)	94 (92%)	8 (8%)	10	26
1	E	102/116 (88%)	94 (92%)	8 (8%)	10	26
1	F	102/116 (88%)	95 (93%)	7 (7%)	13	31
1	G	102/116 (88%)	95 (93%)	7 (7%)	13	31
1	H	102/116 (88%)	91 (89%)	11 (11%)	5	13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	I	102/116 (88%)	92 (90%)	10 (10%)	6 16
1	J	102/116 (88%)	89 (87%)	13 (13%)	3 9
All	All	1020/1160 (88%)	935 (92%)	85 (8%)	9 22

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	69	THR
1	I	116	LEU
1	H	102	ASN
1	I	36	ASP
1	J	31	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	54	ASN
1	I	108	ASN
1	H	63	HIS
1	H	108	ASN
1	J	33	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	118/140 (84%)	1.05	22 (18%) 4 4	34, 59, 96, 103	0
1	B	118/140 (84%)	0.82	16 (13%) 8 7	37, 53, 90, 96	0
1	C	118/140 (84%)	0.57	13 (11%) 12 11	29, 54, 81, 89	0
1	D	118/140 (84%)	0.64	9 (7%) 21 19	35, 57, 86, 92	0
1	E	118/140 (84%)	0.74	6 (5%) 34 32	32, 57, 84, 91	0
1	F	118/140 (84%)	1.02	21 (17%) 4 5	34, 56, 92, 103	0
1	G	118/140 (84%)	0.82	15 (12%) 9 8	39, 57, 95, 97	0
1	H	118/140 (84%)	0.34	7 (5%) 29 27	33, 47, 79, 92	0
1	I	118/140 (84%)	0.22	2 (1%) 69 68	31, 51, 73, 85	0
1	J	118/140 (84%)	0.43	6 (5%) 34 32	33, 53, 83, 90	0
All	All	1180/1400 (84%)	0.66	117 (9%) 14 13	29, 54, 87, 103	0

The worst 5 of 117 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	51	VAL	9.6
1	G	113	TRP	8.1
1	A	113	TRP	7.5
1	A	51	VAL	7.3
1	F	52	SER	5.5

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.