



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:19 AM GMT

PDB ID : 3AJ1
Title : The structure of AxCeSD octamer (N-terminal HIS-tag) from Acetobacter xylinum
Authors : Hu, S.Q.; Tajima, K.; Zhou, Y.; Tanaka, I.; Yao, M.
Deposited on : 2010-05-20
Resolution : 2.50 Å(reported)

This is a wwPDB validation summary 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/ValidationPDFNotes.html>

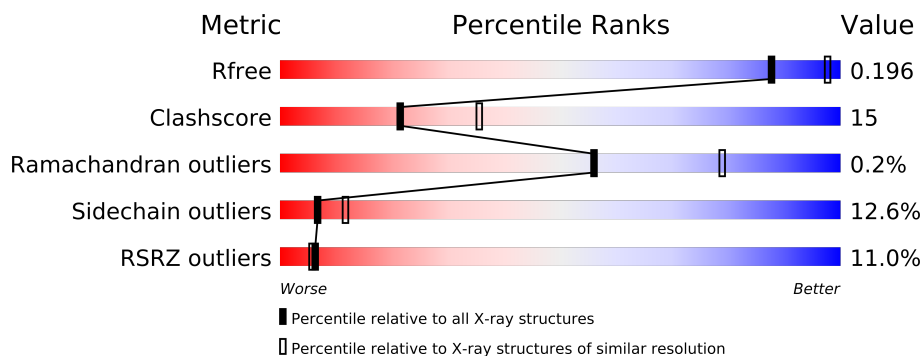
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.50 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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	167	
1	B	167	
1	C	167	
1	D	167	
1	E	167	
1	F	167	
1	G	167	
1	H	167	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9931 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cellulose synthase operon protein D.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	151	Total	C	N	O	S	Se	0	0	0
			1177	745	202	225	1	4			
1	B	156	Total	C	N	O	S	Se	0	0	0
			1220	774	207	233	1	5			
1	C	151	Total	C	N	O	S	Se	0	0	0
			1178	745	202	226	1	4			
1	D	151	Total	C	N	O	S	Se	0	0	0
			1177	745	202	225	1	4			
1	E	155	Total	C	N	O	S	Se	0	0	0
			1212	769	206	232	1	4			
1	F	151	Total	C	N	O	S	Se	0	0	0
			1177	745	202	225	1	4			
1	G	151	Total	C	N	O	S	Se	0	0	0
			1177	745	202	225	1	4			
1	H	156	Total	C	N	O	S	Se	0	0	0
			1221	774	207	234	1	5			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MSE	-	EXPRESSION TAG	UNP P37719
A	-9	ARG	-	EXPRESSION TAG	UNP P37719
A	-8	GLY	-	EXPRESSION TAG	UNP P37719
A	-7	SER	-	EXPRESSION TAG	UNP P37719
A	-6	HIS	-	EXPRESSION TAG	UNP P37719
A	-5	HIS	-	EXPRESSION TAG	UNP P37719
A	-4	HIS	-	EXPRESSION TAG	UNP P37719
A	-3	HIS	-	EXPRESSION TAG	UNP P37719
A	-2	HIS	-	EXPRESSION TAG	UNP P37719
A	-1	HIS	-	EXPRESSION TAG	UNP P37719
A	0	GLY	-	EXPRESSION TAG	UNP P37719
B	-10	MSE	-	EXPRESSION TAG	UNP P37719
B	-9	ARG	-	EXPRESSION TAG	UNP P37719

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	EXPRESSION TAG	UNP P37719
B	-7	SER	-	EXPRESSION TAG	UNP P37719
B	-6	HIS	-	EXPRESSION TAG	UNP P37719
B	-5	HIS	-	EXPRESSION TAG	UNP P37719
B	-4	HIS	-	EXPRESSION TAG	UNP P37719
B	-3	HIS	-	EXPRESSION TAG	UNP P37719
B	-2	HIS	-	EXPRESSION TAG	UNP P37719
B	-1	HIS	-	EXPRESSION TAG	UNP P37719
B	0	GLY	-	EXPRESSION TAG	UNP P37719
C	-10	MSE	-	EXPRESSION TAG	UNP P37719
C	-9	ARG	-	EXPRESSION TAG	UNP P37719
C	-8	GLY	-	EXPRESSION TAG	UNP P37719
C	-7	SER	-	EXPRESSION TAG	UNP P37719
C	-6	HIS	-	EXPRESSION TAG	UNP P37719
C	-5	HIS	-	EXPRESSION TAG	UNP P37719
C	-4	HIS	-	EXPRESSION TAG	UNP P37719
C	-3	HIS	-	EXPRESSION TAG	UNP P37719
C	-2	HIS	-	EXPRESSION TAG	UNP P37719
C	-1	HIS	-	EXPRESSION TAG	UNP P37719
C	0	GLY	-	EXPRESSION TAG	UNP P37719
D	-10	MSE	-	EXPRESSION TAG	UNP P37719
D	-9	ARG	-	EXPRESSION TAG	UNP P37719
D	-8	GLY	-	EXPRESSION TAG	UNP P37719
D	-7	SER	-	EXPRESSION TAG	UNP P37719
D	-6	HIS	-	EXPRESSION TAG	UNP P37719
D	-5	HIS	-	EXPRESSION TAG	UNP P37719
D	-4	HIS	-	EXPRESSION TAG	UNP P37719
D	-3	HIS	-	EXPRESSION TAG	UNP P37719
D	-2	HIS	-	EXPRESSION TAG	UNP P37719
D	-1	HIS	-	EXPRESSION TAG	UNP P37719
D	0	GLY	-	EXPRESSION TAG	UNP P37719
E	-10	MSE	-	EXPRESSION TAG	UNP P37719
E	-9	ARG	-	EXPRESSION TAG	UNP P37719
E	-8	GLY	-	EXPRESSION TAG	UNP P37719
E	-7	SER	-	EXPRESSION TAG	UNP P37719
E	-6	HIS	-	EXPRESSION TAG	UNP P37719
E	-5	HIS	-	EXPRESSION TAG	UNP P37719
E	-4	HIS	-	EXPRESSION TAG	UNP P37719
E	-3	HIS	-	EXPRESSION TAG	UNP P37719
E	-2	HIS	-	EXPRESSION TAG	UNP P37719
E	-1	HIS	-	EXPRESSION TAG	UNP P37719
E	0	GLY	-	EXPRESSION TAG	UNP P37719

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-10	MSE	-	EXPRESSION TAG	UNP P37719
F	-9	ARG	-	EXPRESSION TAG	UNP P37719
F	-8	GLY	-	EXPRESSION TAG	UNP P37719
F	-7	SER	-	EXPRESSION TAG	UNP P37719
F	-6	HIS	-	EXPRESSION TAG	UNP P37719
F	-5	HIS	-	EXPRESSION TAG	UNP P37719
F	-4	HIS	-	EXPRESSION TAG	UNP P37719
F	-3	HIS	-	EXPRESSION TAG	UNP P37719
F	-2	HIS	-	EXPRESSION TAG	UNP P37719
F	-1	HIS	-	EXPRESSION TAG	UNP P37719
F	0	GLY	-	EXPRESSION TAG	UNP P37719
G	-10	MSE	-	EXPRESSION TAG	UNP P37719
G	-9	ARG	-	EXPRESSION TAG	UNP P37719
G	-8	GLY	-	EXPRESSION TAG	UNP P37719
G	-7	SER	-	EXPRESSION TAG	UNP P37719
G	-6	HIS	-	EXPRESSION TAG	UNP P37719
G	-5	HIS	-	EXPRESSION TAG	UNP P37719
G	-4	HIS	-	EXPRESSION TAG	UNP P37719
G	-3	HIS	-	EXPRESSION TAG	UNP P37719
G	-2	HIS	-	EXPRESSION TAG	UNP P37719
G	-1	HIS	-	EXPRESSION TAG	UNP P37719
G	0	GLY	-	EXPRESSION TAG	UNP P37719
H	-10	MSE	-	EXPRESSION TAG	UNP P37719
H	-9	ARG	-	EXPRESSION TAG	UNP P37719
H	-8	GLY	-	EXPRESSION TAG	UNP P37719
H	-7	SER	-	EXPRESSION TAG	UNP P37719
H	-6	HIS	-	EXPRESSION TAG	UNP P37719
H	-5	HIS	-	EXPRESSION TAG	UNP P37719
H	-4	HIS	-	EXPRESSION TAG	UNP P37719
H	-3	HIS	-	EXPRESSION TAG	UNP P37719
H	-2	HIS	-	EXPRESSION TAG	UNP P37719
H	-1	HIS	-	EXPRESSION TAG	UNP P37719
H	0	GLY	-	EXPRESSION TAG	UNP P37719

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	52	Total O 52 52	0	0
2	B	47	Total O 47 47	0	0
2	C	45	Total O 45 45	0	0

Continued on next page...

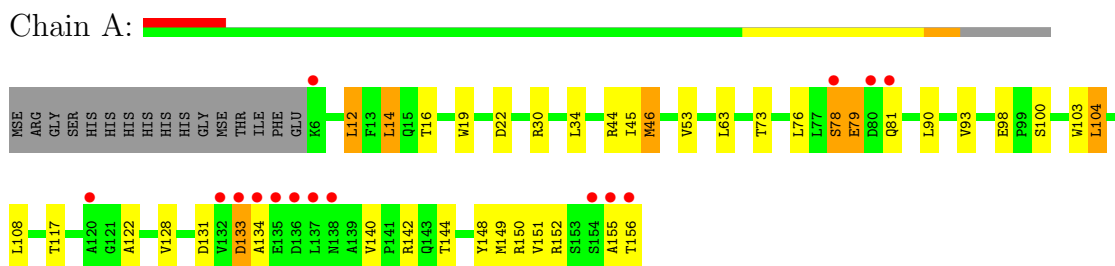
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	54	Total 54	O 54	0	0
2	E	48	Total 48	O 48	0	0
2	F	52	Total 52	O 52	0	0
2	G	42	Total 42	O 42	0	0
2	H	52	Total 52	O 52	0	0

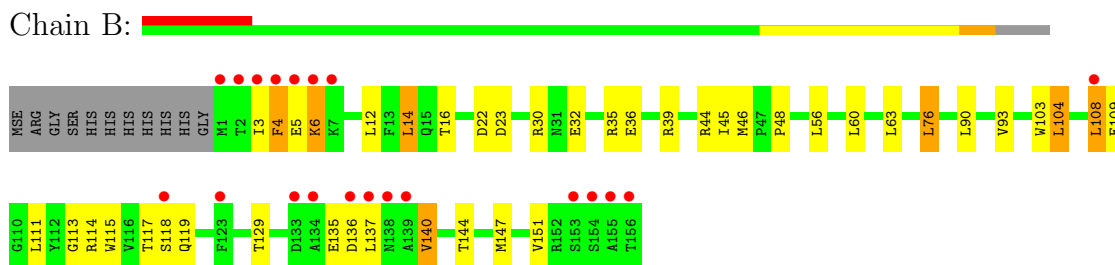
3 Residue-property plots

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.

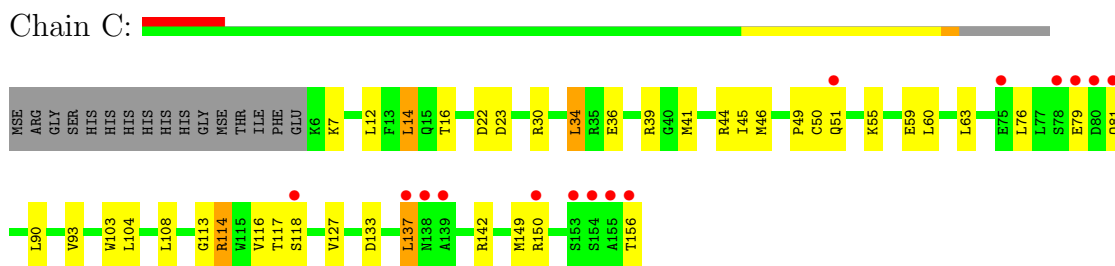
- Molecule 1: Cellulose synthase operon protein D



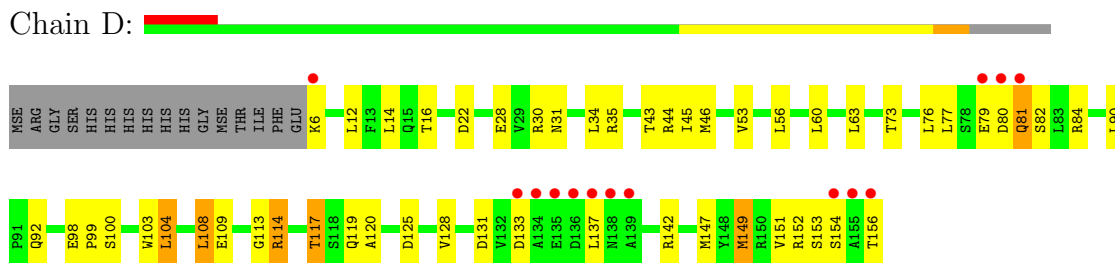
- Molecule 1: Cellulose synthase operon protein D



- Molecule 1: Cellulose synthase operon protein D

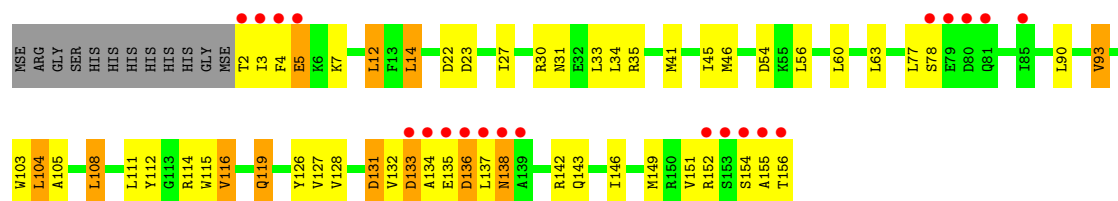


- Molecule 1: Cellulose synthase operon protein D



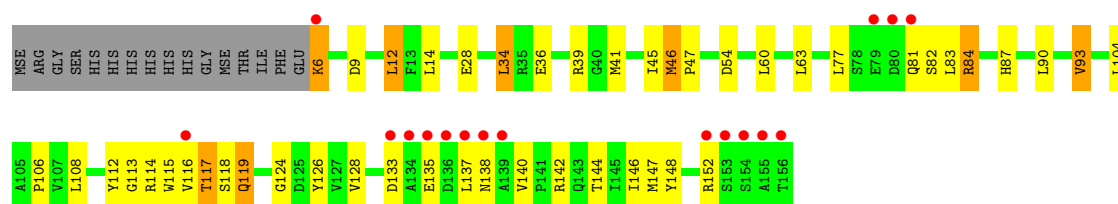
- Molecule 1: Cellulose synthase operon protein D

Chain E: 



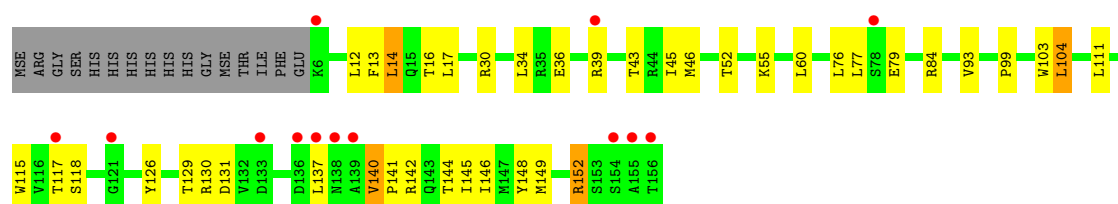
- Molecule 1: Cellulose synthase operon protein D

Chain F: 



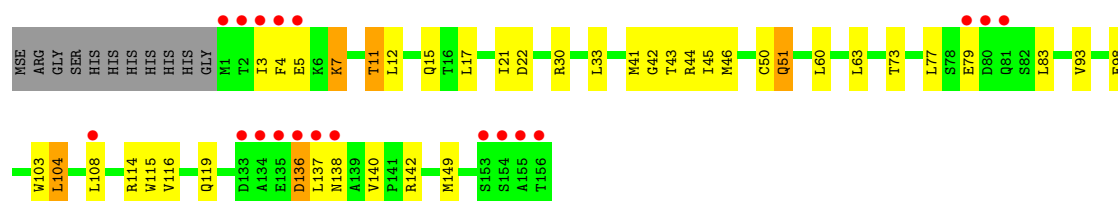
- Molecule 1: Cellulose synthase operon protein D

Chain G: 



- Molecule 1: Cellulose synthase operon protein D

Chain H: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	77.70Å 77.70Å 213.92Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.87 – 2.50 19.87 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.4 (19.87-2.50) 99.3 (19.87-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.203 , 0.257 0.208 , 0.196	Depositor DCC
R_{free} test set	4986 reflections (11.20%)	DCC
Wilson B-factor (Å ²)	43.9	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 52.3	EDS
Estimated twinning fraction	0.016 for -h,-k,l 0.074 for h,-h-k,-l 0.031 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 49522 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9931	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.50	0/1195	0.68	0/1621
1	B	0.54	0/1239	0.68	0/1680
1	C	0.53	0/1196	0.68	0/1621
1	D	0.60	1/1195 (0.1%)	0.69	0/1621
1	E	0.51	0/1231	0.69	0/1670
1	F	0.51	0/1195	0.66	0/1621
1	G	0.47	0/1195	0.63	0/1621
1	H	0.57	0/1240	0.71	1/1680 (0.1%)
All	All	0.53	1/9686 (0.0%)	0.68	1/13135 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	E	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	153	SER	CB-OG	5.46	1.49	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	104	LEU	CA-CB-CG	5.11	127.05	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	138	ASN	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1177	0	1180	33	0
1	B	1220	0	1225	34	0
1	C	1178	0	1180	36	0
1	D	1177	0	1180	47	0
1	E	1212	0	1213	46	0
1	F	1177	0	1180	40	0
1	G	1177	0	1180	37	0
1	H	1221	0	1225	42	0
2	A	52	0	0	5	0
2	B	47	0	0	1	0
2	C	45	0	0	2	0
2	D	54	0	0	4	0
2	E	48	0	0	4	0
2	F	52	0	0	4	0
2	G	42	0	0	2	0
2	H	52	0	0	2	0
All	All	9931	0	9563	292	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

The worst 5 of 292 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:116:VAL:CG1	1:E:149:MSE:HE3	1.38	1.52
1:E:116:VAL:HG13	1:E:149:MSE:CE	1.53	1.37
1:H:50:CYS:H	1:H:119:GLN:NE2	1.45	1.15
1:F:46:MSE:HE2	1:F:47:PRO:HD2	1.19	1.14
1:D:81:GLN:HA	1:D:81:GLN:HE21	0.98	1.10

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	149/167 (89%)	142 (95%)	6 (4%)	1 (1%)	30	50
1	B	154/167 (92%)	146 (95%)	7 (4%)	1 (1%)	33	55
1	C	149/167 (89%)	142 (95%)	7 (5%)	0	100	100
1	D	149/167 (89%)	143 (96%)	6 (4%)	0	100	100
1	E	153/167 (92%)	144 (94%)	9 (6%)	0	100	100
1	F	149/167 (89%)	145 (97%)	3 (2%)	1 (1%)	30	50
1	G	149/167 (89%)	144 (97%)	5 (3%)	0	100	100
1	H	154/167 (92%)	147 (96%)	7 (4%)	0	100	100
All	All	1206/1336 (90%)	1153 (96%)	50 (4%)	3 (0%)	56	79

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	5	GLU
1	F	152	ARG
1	A	152	ARG

5.3.2 Protein sidechains ⓘ

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	129/137 (94%)	116 (90%)	13 (10%)	11	20
1	B	134/137 (98%)	118 (88%)	16 (12%)	8	13
1	C	129/137 (94%)	117 (91%)	12 (9%)	13	23
1	D	129/137 (94%)	108 (84%)	21 (16%)	3	6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	133/137 (97%)	106 (80%)	27 (20%)	2	3
1	F	129/137 (94%)	111 (86%)	18 (14%)	5	9
1	G	129/137 (94%)	117 (91%)	12 (9%)	13	23
1	H	134/137 (98%)	121 (90%)	13 (10%)	12	21
All	All	1046/1096 (95%)	914 (87%)	132 (13%)	7	12

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	137	LEU
1	E	63	LEU
1	H	11	THR
1	D	156	THR
1	E	12	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	81	GLN
1	H	119	GLN
1	F	119	GLN
1	D	138	ASN
1	F	138	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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, 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	151/167 (90%)	0.46	15 (9%) 8 7	38, 53, 86, 92	0
1	B	156/167 (93%)	0.66	20 (12%) 4 4	41, 52, 87, 104	0
1	C	151/167 (90%)	0.56	15 (9%) 8 7	43, 54, 90, 97	0
1	D	151/167 (90%)	0.46	14 (9%) 9 8	43, 54, 85, 90	0
1	E	155/167 (92%)	0.72	21 (13%) 4 3	46, 58, 89, 101	0
1	F	151/167 (90%)	0.63	17 (11%) 6 5	47, 62, 92, 101	0
1	G	151/167 (90%)	0.62	13 (8%) 11 10	50, 61, 90, 98	0
1	H	156/167 (93%)	0.70	19 (12%) 5 4	43, 55, 91, 108	0
All	All	1222/1336 (91%)	0.60	134 (10%) 6 6	38, 57, 89, 108	0

The worst 5 of 134 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	156	THR	14.2
1	A	156	THR	11.6
1	H	4	PHE	10.8
1	E	2	THR	10.5
1	C	155	ALA	10.2

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.