W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 02:36 am BST

PDB ID : 3HT4

Title : Crystal Structure of the Q81A77_BACCR Protein from Bacillus cereus. Northeast Structural Genomics Consortium Target BcR213

Authors : Vorobiev, S.; Lew, S.; Seetharaman, J.; Wang, H.; Foote, E.; Ciccosanti, C.; Janjua, H.; Xiao, R.; Mao, L.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)

Deposited on : 2009-06-11

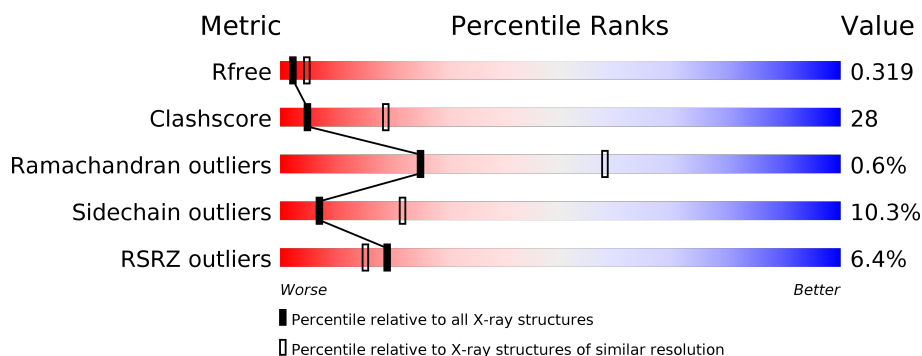Resolution : 2.90 Å(reported)

---

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 2.90 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	Percentile Ranks	Value
Rfree		0.319
Clashscore		28
Ramachandran outliers		0.6%
Sidechain outliers		10.3%
RSRZ outliers		6.4%

Worse ... *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	431	5% — 45% / 47% / 5% ·
1	B	431	4% — 49% / 41% / 5% / 6%
1	C	431	6% — 50% / 40% / 5% / 5%
1	D	431	4% — 53% / 38% / · / 5%
1	E	431	5% — 51% / 40% / · / 5%
1	F	431	5% — 49% / 41% / 5% / 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	431	7% ▐ 53% 38% 5%
1	H	431	11% ▐ 48% 43% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 25070 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 Aluminum resistance protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	S	Se	0	0	0
			3151	2015	517	605	5	9			
1	B	407	Total	C	N	O	S	Se	0	0	0
			3112	1994	509	596	5	8			
1	C	411	Total	C	N	O	S	Se	0	0	0
			3123	2001	513	596	5	8			
1	D	408	Total	C	N	O	S	Se	0	0	0
			3117	1997	510	597	5	8			
1	E	409	Total	C	N	O	S	Se	0	0	0
			3117	1997	511	596	5	8			
1	F	410	Total	C	N	O	S	Se	0	0	0
			3118	1998	512	595	5	8			
1	G	408	Total	C	N	O	S	Se	0	0	0
			3102	1986	510	593	5	8			
1	H	407	Total	C	N	O	S	Se	0	0	0
			3104	1989	508	594	5	8			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	424	LEU	-	EXPRESSION TAG	UNP Q81A77
A	425	GLU	-	EXPRESSION TAG	UNP Q81A77
A	426	HIS	-	EXPRESSION TAG	UNP Q81A77
A	427	HIS	-	EXPRESSION TAG	UNP Q81A77
A	428	HIS	-	EXPRESSION TAG	UNP Q81A77
A	429	HIS	-	EXPRESSION TAG	UNP Q81A77
A	430	HIS	-	EXPRESSION TAG	UNP Q81A77
A	431	HIS	-	EXPRESSION TAG	UNP Q81A77
B	424	LEU	-	EXPRESSION TAG	UNP Q81A77
B	425	GLU	-	EXPRESSION TAG	UNP Q81A77
B	426	HIS	-	EXPRESSION TAG	UNP Q81A77
B	427	HIS	-	EXPRESSION TAG	UNP Q81A77
B	428	HIS	-	EXPRESSION TAG	UNP Q81A77

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	429	HIS	-	EXPRESSION TAG	UNP Q81A77
B	430	HIS	-	EXPRESSION TAG	UNP Q81A77
B	431	HIS	-	EXPRESSION TAG	UNP Q81A77
C	424	LEU	-	EXPRESSION TAG	UNP Q81A77
C	425	GLU	-	EXPRESSION TAG	UNP Q81A77
C	426	HIS	-	EXPRESSION TAG	UNP Q81A77
C	427	HIS	-	EXPRESSION TAG	UNP Q81A77
C	428	HIS	-	EXPRESSION TAG	UNP Q81A77
C	429	HIS	-	EXPRESSION TAG	UNP Q81A77
C	430	HIS	-	EXPRESSION TAG	UNP Q81A77
C	431	HIS	-	EXPRESSION TAG	UNP Q81A77
D	424	LEU	-	EXPRESSION TAG	UNP Q81A77
D	425	GLU	-	EXPRESSION TAG	UNP Q81A77
D	426	HIS	-	EXPRESSION TAG	UNP Q81A77
D	427	HIS	-	EXPRESSION TAG	UNP Q81A77
D	428	HIS	-	EXPRESSION TAG	UNP Q81A77
D	429	HIS	-	EXPRESSION TAG	UNP Q81A77
D	430	HIS	-	EXPRESSION TAG	UNP Q81A77
D	431	HIS	-	EXPRESSION TAG	UNP Q81A77
E	424	LEU	-	EXPRESSION TAG	UNP Q81A77
E	425	GLU	-	EXPRESSION TAG	UNP Q81A77
E	426	HIS	-	EXPRESSION TAG	UNP Q81A77
E	427	HIS	-	EXPRESSION TAG	UNP Q81A77
E	428	HIS	-	EXPRESSION TAG	UNP Q81A77
E	429	HIS	-	EXPRESSION TAG	UNP Q81A77
E	430	HIS	-	EXPRESSION TAG	UNP Q81A77
E	431	HIS	-	EXPRESSION TAG	UNP Q81A77
F	424	LEU	-	EXPRESSION TAG	UNP Q81A77
F	425	GLU	-	EXPRESSION TAG	UNP Q81A77
F	426	HIS	-	EXPRESSION TAG	UNP Q81A77
F	427	HIS	-	EXPRESSION TAG	UNP Q81A77
F	428	HIS	-	EXPRESSION TAG	UNP Q81A77
F	429	HIS	-	EXPRESSION TAG	UNP Q81A77
F	430	HIS	-	EXPRESSION TAG	UNP Q81A77
F	431	HIS	-	EXPRESSION TAG	UNP Q81A77
G	424	LEU	-	EXPRESSION TAG	UNP Q81A77
G	425	GLU	-	EXPRESSION TAG	UNP Q81A77
G	426	HIS	-	EXPRESSION TAG	UNP Q81A77
G	427	HIS	-	EXPRESSION TAG	UNP Q81A77
G	428	HIS	-	EXPRESSION TAG	UNP Q81A77
G	429	HIS	-	EXPRESSION TAG	UNP Q81A77
G	430	HIS	-	EXPRESSION TAG	UNP Q81A77

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	431	HIS	-	EXPRESSION TAG	UNP Q81A77
H	424	LEU	-	EXPRESSION TAG	UNP Q81A77
H	425	GLU	-	EXPRESSION TAG	UNP Q81A77
H	426	HIS	-	EXPRESSION TAG	UNP Q81A77
H	427	HIS	-	EXPRESSION TAG	UNP Q81A77
H	428	HIS	-	EXPRESSION TAG	UNP Q81A77
H	429	HIS	-	EXPRESSION TAG	UNP Q81A77
H	430	HIS	-	EXPRESSION TAG	UNP Q81A77
H	431	HIS	-	EXPRESSION TAG	UNP Q81A77

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	19	Total 19	O 19	0	0
2	B	16	Total 16	O 16	0	0
2	C	12	Total 12	O 12	0	0
2	D	24	Total 24	O 24	0	0
2	E	18	Total 18	O 18	0	0
2	F	18	Total 18	O 18	0	0
2	G	10	Total 10	O 10	0	0
2	H	9	Total 9	O 9	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 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.
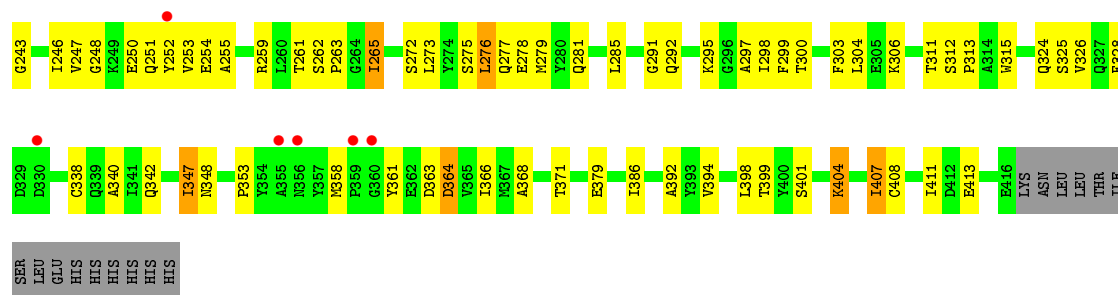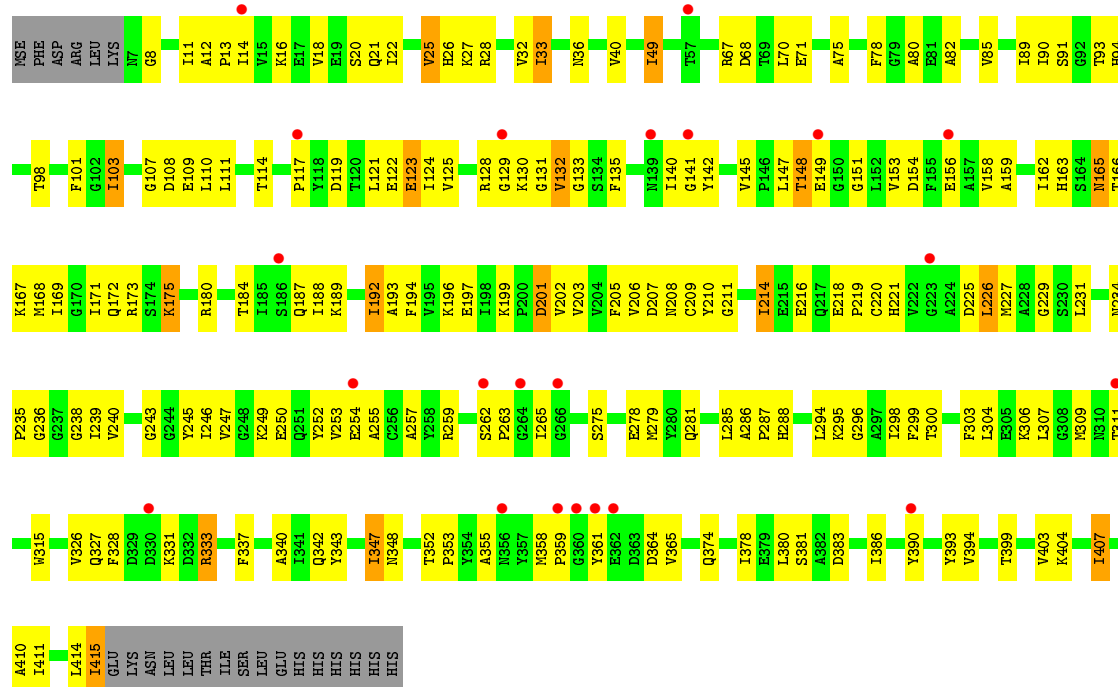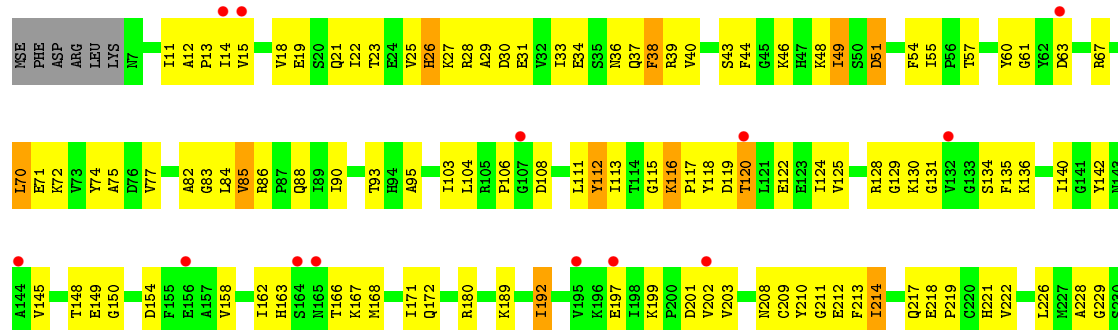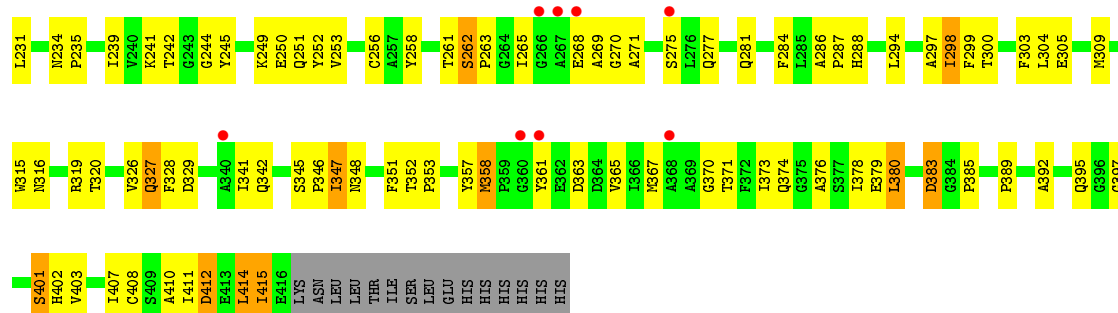
- Molecule 1: Aluminum resistance protein

- Molecule 1: Aluminum resistance protein

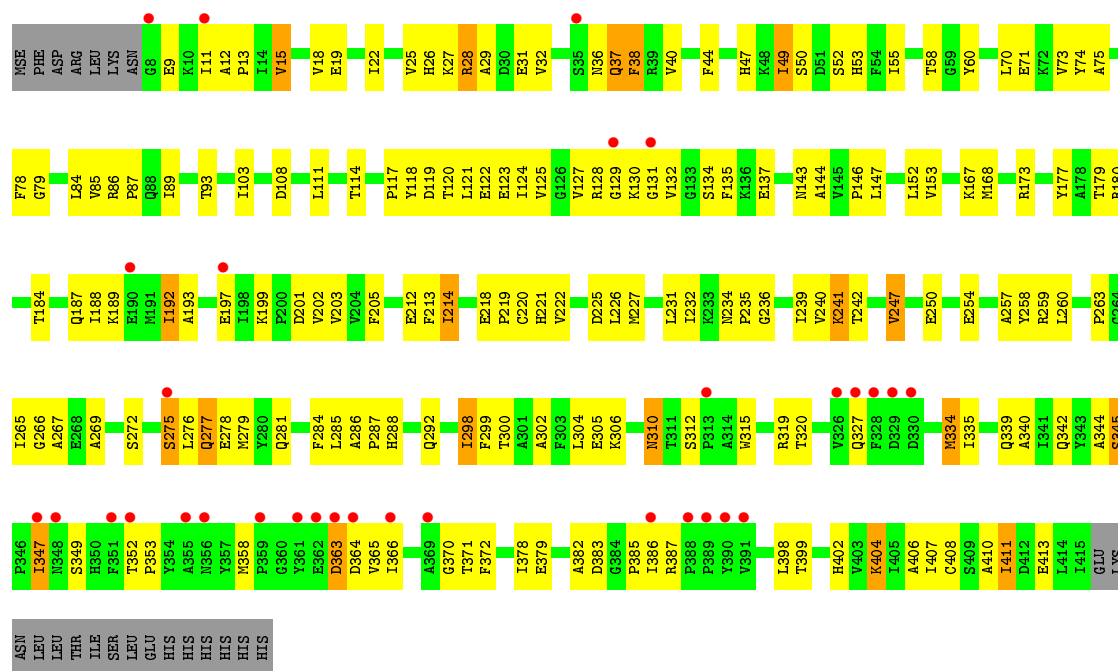• Molecule 1: Aluminum resistance protein
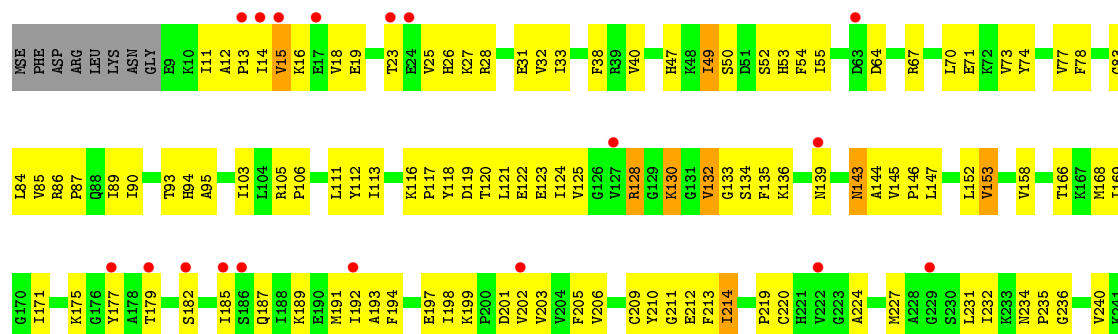

Chain C:

• Molecule 1: Aluminum resistance protein

Chain D:

- Molecule 1: Aluminum resistance protein

Chain E:

- Molecule 1: Aluminum resistance protein

Chain F:

• Molecule 1: Aluminum resistance protein
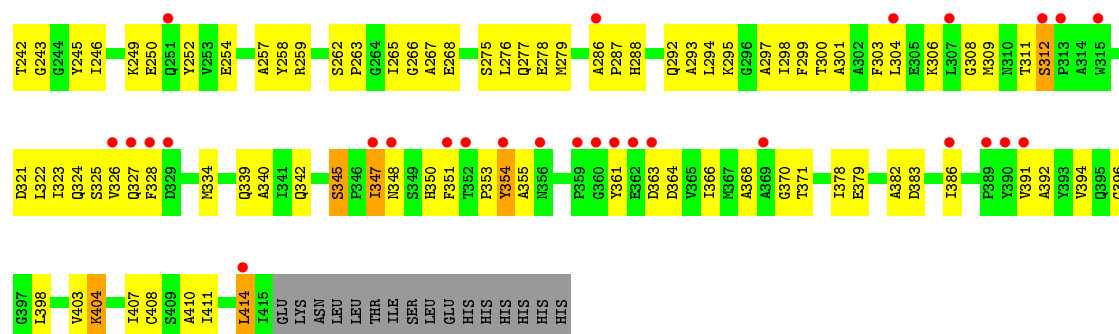
Chain G:

• Molecule 1: Aluminum resistance protein

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	91.41Å 144.99Å 131.97Å 90.00° 106.24° 90.00°	Depositor
Resolution (Å)	46.19 – 2.90 46.19 – 2.90	Depositor EDS
% Data completeness (in resolution range)	75.0 (46.19-2.90) 91.4 (46.19-2.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.86 (at 2.91Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.274 , 0.299 0.291 , 0.319	Depositor DCC
R_{free} test set	6339 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	34.2	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 40.0	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.80	EDS
Total number of atoms	25070	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.75 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.7445e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 (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	0.22	0/3213	0.36	0/4346
1	B	0.22	0/3174	0.36	0/4291
1	C	0.22	0/3185	0.37	0/4307
1	D	0.23	0/3179	0.37	0/4298
1	E	0.22	0/3179	0.36	0/4298
1	F	0.22	0/3180	0.36	0/4300
1	G	0.22	0/3164	0.37	0/4278
1	H	0.21	0/3166	0.38	0/4282
All	All	0.22	0/25440	0.37	0/34400

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	3151	0	3073	205	0
1	B	3112	0	3066	180	0
1	C	3123	0	3067	170	0
1	D	3117	0	3068	158	0
1	E	3117	0	3067	181	0
1	F	3118	0	3065	183	0
1	G	3102	0	3043	177	0
1	H	3104	0	3051	189	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	19	0	0	1	0
2	B	16	0	0	0	0
2	C	12	0	0	2	0
2	D	24	0	0	4	0
2	E	18	0	0	3	0
2	F	18	0	0	4	0
2	G	10	0	0	1	0
2	H	9	0	0	1	0
All	All	25070	0	24500	1364	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 28.

The worst 5 of 1364 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:H:327:GLN:HG2	1:H:391:VAL:HG22	1.33	1.10
1:E:414:LEU:O	1:E:415:ILE:HG13	1.51	1.09
1:E:8:GLY:HA2	1:E:11:ILE:HG22	1.36	1.04
1:E:184:THR:H	1:E:187:GLN:HE21	1.03	1.02
1:B:49:ILE:HD11	1:C:346:PRO:O	1.60	1.01

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	414/431 (96%)	356 (86%)	53 (13%)	5 (1%)	13	40
1	B	405/431 (94%)	347 (86%)	56 (14%)	2 (0%)	29	61
1	C	409/431 (95%)	361 (88%)	45 (11%)	3 (1%)	22	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	406/431 (94%)	350 (86%)	55 (14%)	1 (0%)	47	78
1	E	407/431 (94%)	361 (89%)	45 (11%)	1 (0%)	47	78
1	F	408/431 (95%)	356 (87%)	51 (12%)	1 (0%)	47	78
1	G	406/431 (94%)	344 (85%)	60 (15%)	2 (0%)	29	61
1	H	405/431 (94%)	335 (83%)	65 (16%)	5 (1%)	13	40
All	All	3260/3448 (94%)	2810 (86%)	430 (13%)	20 (1%)	25	58

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	265	ILE
1	A	166	THR
1	A	351	PHE
1	B	89	ILE
1	C	67	ARG

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	325/340 (96%)	285 (88%)	40 (12%)	4	14
1	B	326/340 (96%)	287 (88%)	39 (12%)	5	15
1	C	324/340 (95%)	291 (90%)	33 (10%)	7	22
1	D	326/340 (96%)	295 (90%)	31 (10%)	8	26
1	E	325/340 (96%)	296 (91%)	29 (9%)	9	29
1	F	324/340 (95%)	280 (86%)	44 (14%)	3	11
1	G	322/340 (95%)	295 (92%)	27 (8%)	11	31
1	H	324/340 (95%)	299 (92%)	25 (8%)	13	35
All	All	2596/2720 (95%)	2328 (90%)	268 (10%)	7	22

5 of 268 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	112	TYR
1	E	103	ILE
1	H	49	ILE
1	D	123	GLU
1	D	364	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	163	HIS
1	E	21	GLN
1	H	187	GLN
1	D	187	GLN
1	D	292	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 carbohydrates 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(Å²)	Q<0.9
1	A	407/431 (94%)	0.58	23 (5%)	23	19	15, 33, 51, 62	0
1	B	399/431 (92%)	0.52	17 (4%)	35	31	14, 30, 50, 66	0
1	C	403/431 (93%)	0.63	27 (6%)	17	13	15, 32, 55, 69	0
1	D	400/431 (92%)	0.51	18 (4%)	33	29	10, 28, 57, 71	0
1	E	401/431 (93%)	0.53	22 (5%)	25	21	8, 29, 56, 67	0
1	F	402/431 (93%)	0.53	21 (5%)	27	23	14, 30, 53, 64	0
1	G	400/431 (92%)	0.71	32 (8%)	12	9	17, 36, 71, 84	0
1	H	399/431 (92%)	0.83	46 (11%)	4	3	20, 37, 67, 81	0
All	All	3211/3448 (93%)	0.60	206 (6%)	19	15	8, 32, 60, 84	0

The worst 5 of 206 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	360	GLY	12.6
1	D	359	PRO	10.7
1	E	360	GLY	7.3
1	H	362	GLU	6.2
1	G	362	GLU	6.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.