



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 09:01 am BST

PDB ID : 6AF3  
Title : Toxin-Antitoxin module from *Streptococcus pneumoniae*  
Authors : Kang, S.M.; Kim, D.H.; Jin, C.  
Deposited on : 2018-08-08  
Resolution : 2.80 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](mailto: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  
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

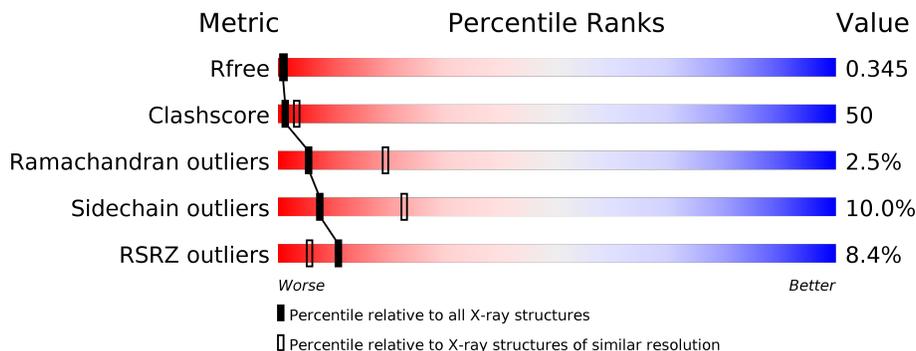
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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  $\geq 3$ , 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	141	
1	C	141	
1	E	141	
1	G	141	
2	B	97	
2	D	97	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	97	 <p>2% 46% 41% 5% 7%</p>
2	H	97	 <p>3% 44% 34% 8% 11%</p>

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 6717 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 HigB toxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	114	Total 968	C 619	N 174	O 171	Se 4	0	0	0
1	C	113	Total 959	C 614	N 173	O 168	Se 4	0	0	0
1	E	116	Total 983	C 627	N 179	O 173	Se 4	0	0	0
1	G	117	Total 991	C 633	N 180	O 174	Se 4	0	0	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MSE	-	initiating methionine	UNP A0A0H2UQ08
A	-18	GLY	-	expression tag	UNP A0A0H2UQ08
A	-17	SER	-	expression tag	UNP A0A0H2UQ08
A	-16	SER	-	expression tag	UNP A0A0H2UQ08
A	-15	HIS	-	expression tag	UNP A0A0H2UQ08
A	-14	HIS	-	expression tag	UNP A0A0H2UQ08
A	-13	HIS	-	expression tag	UNP A0A0H2UQ08
A	-12	HIS	-	expression tag	UNP A0A0H2UQ08
A	-11	HIS	-	expression tag	UNP A0A0H2UQ08
A	-10	HIS	-	expression tag	UNP A0A0H2UQ08
A	-9	SER	-	expression tag	UNP A0A0H2UQ08
A	-8	SER	-	expression tag	UNP A0A0H2UQ08
A	-7	GLY	-	expression tag	UNP A0A0H2UQ08
A	-6	LEU	-	expression tag	UNP A0A0H2UQ08
A	-5	VAL	-	expression tag	UNP A0A0H2UQ08
A	-4	PRO	-	expression tag	UNP A0A0H2UQ08
A	-3	ARG	-	expression tag	UNP A0A0H2UQ08
A	-2	GLY	-	expression tag	UNP A0A0H2UQ08
A	-1	SER	-	expression tag	UNP A0A0H2UQ08
A	0	HIS	-	expression tag	UNP A0A0H2UQ08
C	-19	MSE	-	initiating methionine	UNP A0A0H2UQ08

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-18	GLY	-	expression tag	UNP A0A0H2UQ08
C	-17	SER	-	expression tag	UNP A0A0H2UQ08
C	-16	SER	-	expression tag	UNP A0A0H2UQ08
C	-15	HIS	-	expression tag	UNP A0A0H2UQ08
C	-14	HIS	-	expression tag	UNP A0A0H2UQ08
C	-13	HIS	-	expression tag	UNP A0A0H2UQ08
C	-12	HIS	-	expression tag	UNP A0A0H2UQ08
C	-11	HIS	-	expression tag	UNP A0A0H2UQ08
C	-10	HIS	-	expression tag	UNP A0A0H2UQ08
C	-9	SER	-	expression tag	UNP A0A0H2UQ08
C	-8	SER	-	expression tag	UNP A0A0H2UQ08
C	-7	GLY	-	expression tag	UNP A0A0H2UQ08
C	-6	LEU	-	expression tag	UNP A0A0H2UQ08
C	-5	VAL	-	expression tag	UNP A0A0H2UQ08
C	-4	PRO	-	expression tag	UNP A0A0H2UQ08
C	-3	ARG	-	expression tag	UNP A0A0H2UQ08
C	-2	GLY	-	expression tag	UNP A0A0H2UQ08
C	-1	SER	-	expression tag	UNP A0A0H2UQ08
C	0	HIS	-	expression tag	UNP A0A0H2UQ08
E	-19	MSE	-	initiating methionine	UNP A0A0H2UQ08
E	-18	GLY	-	expression tag	UNP A0A0H2UQ08
E	-17	SER	-	expression tag	UNP A0A0H2UQ08
E	-16	SER	-	expression tag	UNP A0A0H2UQ08
E	-15	HIS	-	expression tag	UNP A0A0H2UQ08
E	-14	HIS	-	expression tag	UNP A0A0H2UQ08
E	-13	HIS	-	expression tag	UNP A0A0H2UQ08
E	-12	HIS	-	expression tag	UNP A0A0H2UQ08
E	-11	HIS	-	expression tag	UNP A0A0H2UQ08
E	-10	HIS	-	expression tag	UNP A0A0H2UQ08
E	-9	SER	-	expression tag	UNP A0A0H2UQ08
E	-8	SER	-	expression tag	UNP A0A0H2UQ08
E	-7	GLY	-	expression tag	UNP A0A0H2UQ08
E	-6	LEU	-	expression tag	UNP A0A0H2UQ08
E	-5	VAL	-	expression tag	UNP A0A0H2UQ08
E	-4	PRO	-	expression tag	UNP A0A0H2UQ08
E	-3	ARG	-	expression tag	UNP A0A0H2UQ08
E	-2	GLY	-	expression tag	UNP A0A0H2UQ08
E	-1	SER	-	expression tag	UNP A0A0H2UQ08
E	0	HIS	-	expression tag	UNP A0A0H2UQ08
G	-19	MSE	-	initiating methionine	UNP A0A0H2UQ08
G	-18	GLY	-	expression tag	UNP A0A0H2UQ08
G	-17	SER	-	expression tag	UNP A0A0H2UQ08

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-16	SER	-	expression tag	UNP A0A0H2UQ08
G	-15	HIS	-	expression tag	UNP A0A0H2UQ08
G	-14	HIS	-	expression tag	UNP A0A0H2UQ08
G	-13	HIS	-	expression tag	UNP A0A0H2UQ08
G	-12	HIS	-	expression tag	UNP A0A0H2UQ08
G	-11	HIS	-	expression tag	UNP A0A0H2UQ08
G	-10	HIS	-	expression tag	UNP A0A0H2UQ08
G	-9	SER	-	expression tag	UNP A0A0H2UQ08
G	-8	SER	-	expression tag	UNP A0A0H2UQ08
G	-7	GLY	-	expression tag	UNP A0A0H2UQ08
G	-6	LEU	-	expression tag	UNP A0A0H2UQ08
G	-5	VAL	-	expression tag	UNP A0A0H2UQ08
G	-4	PRO	-	expression tag	UNP A0A0H2UQ08
G	-3	ARG	-	expression tag	UNP A0A0H2UQ08
G	-2	GLY	-	expression tag	UNP A0A0H2UQ08
G	-1	SER	-	expression tag	UNP A0A0H2UQ08
G	0	HIS	-	expression tag	UNP A0A0H2UQ08

- Molecule 2 is a protein called HigA antitoxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	91	Total	C	N	O	Se	0	0	0
			697	435	119	139	4			
2	D	92	Total	C	N	O	Se	0	0	0
			706	441	121	140	4			
2	F	90	Total	C	N	O	Se	0	0	0
			688	430	118	136	4			
2	H	86	Total	C	N	O	Se	0	0	0
			664	416	112	132	4			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	12	Total	O	0	0
			12	12		
3	B	8	Total	O	0	0
			8	8		
3	C	7	Total	O	0	0
			7	7		
3	D	4	Total	O	0	0
			4	4		
3	E	7	Total	O	0	0
			7	7		

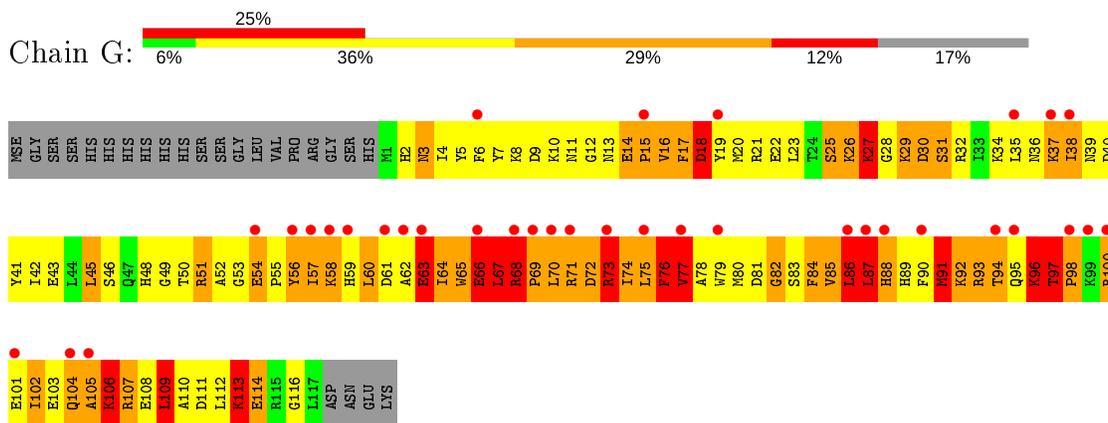
*Continued on next page...*

*Continued from previous page...*

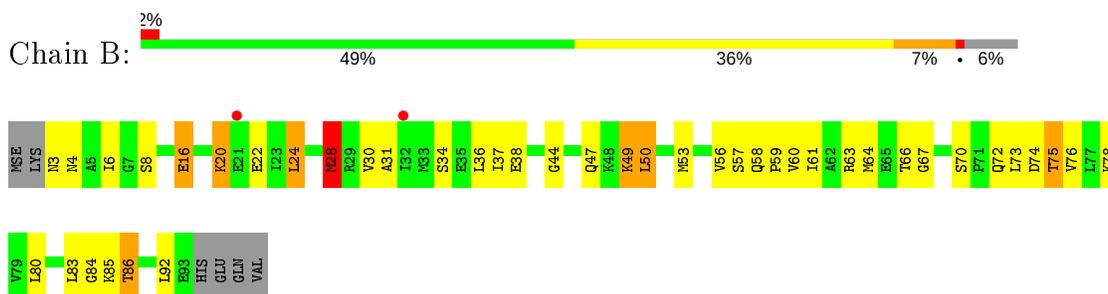
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	F	5	Total 5	O 5	0	0
3	G	7	Total 7	O 7	0	0
3	H	11	Total 11	O 11	0	0



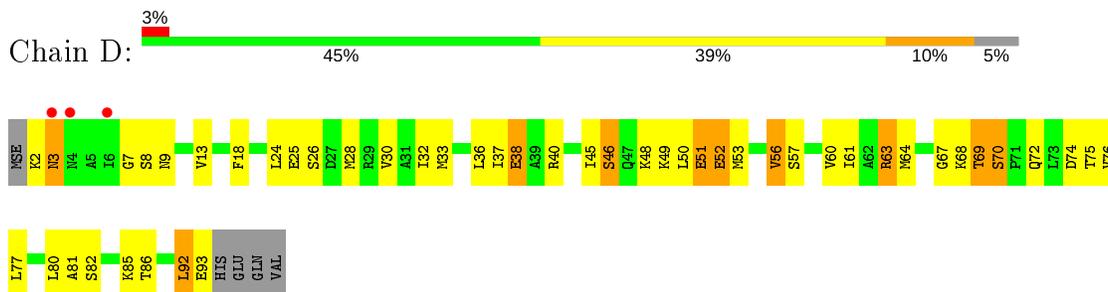
- Molecule 1: HigB toxin



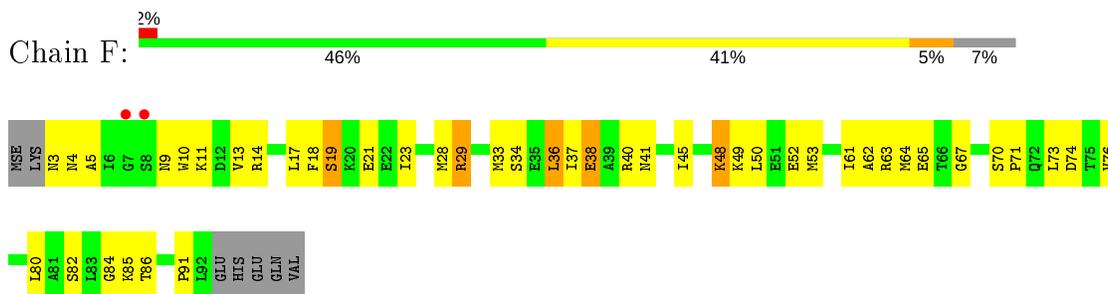
- Molecule 2: HigA antitoxin



- Molecule 2: HigA antitoxin

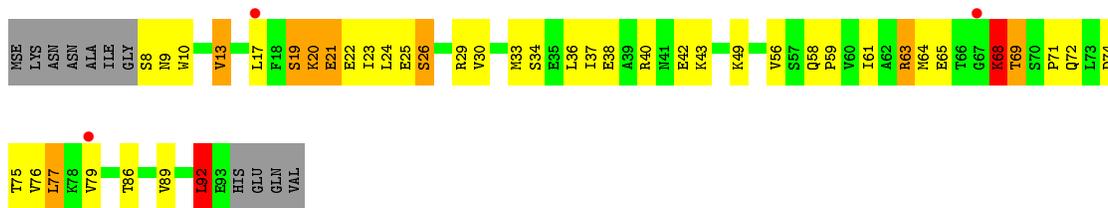


- Molecule 2: HigA antitoxin



- Molecule 2: HigA antitoxin





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.58Å 67.04Å 87.72Å 90.00° 94.22° 90.00°	Depositor
Resolution (Å)	49.80 – 2.80 49.79 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.80-2.80) 93.0 (49.79-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.77 (at 2.81Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.220 , 0.333 0.222 , 0.345	Depositor DCC
$R_{free}$ test set	1095 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	76.4	Xtrriage
Anisotropy	0.107	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 59.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6717	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.92% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	1.21	1/987 (0.1%)	1.32	9/1317 (0.7%)
1	C	1.26	4/978 (0.4%)	1.29	5/1305 (0.4%)
1	E	1.06	1/1002 (0.1%)	1.37	9/1336 (0.7%)
1	G	1.24	4/1010 (0.4%)	2.02	36/1347 (2.7%)
2	B	1.43	4/698 (0.6%)	1.49	4/931 (0.4%)
2	D	1.52	6/707 (0.8%)	1.43	7/942 (0.7%)
2	F	1.39	3/689 (0.4%)	1.35	6/919 (0.7%)
2	H	1.09	0/665	1.32	5/886 (0.6%)
All	All	1.27	23/6736 (0.3%)	1.48	81/8983 (0.9%)

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	A	0	2
1	C	0	1
1	E	0	8
1	G	0	24
2	F	0	1
2	H	0	3
All	All	0	39

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	38	GLU	CG-CD	9.11	1.65	1.51
1	G	77	VAL	CB-CG1	-8.74	1.34	1.52
1	C	56	TYR	CB-CG	-7.09	1.41	1.51
2	F	21	GLU	CG-CD	7.01	1.62	1.51
1	C	22	GLU	CG-CD	6.57	1.61	1.51

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	28	MSE	CA-CB-CG	-14.78	88.18	113.30
1	G	37	LYS	CD-CE-NZ	14.76	145.64	111.70
1	G	75	LEU	CB-CG-CD2	-12.86	89.15	111.00
1	A	32	ARG	NE-CZ-NH1	-11.54	114.53	120.30
1	G	70	LEU	CB-CG-CD2	10.66	129.13	111.00

There are no chirality outliers.

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	11	ASN	Peptide
1	A	82	GLY	Peptide
1	C	80	MSE	Peptide
1	E	5	TYR	Peptide
1	E	8	LYS	Peptide

## 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	968	0	978	58	1
1	C	959	0	972	72	0
1	E	983	0	994	134	0
1	G	991	0	1005	293	0
2	B	697	0	728	42	0
2	D	706	0	741	43	1
2	F	688	0	722	39	1
2	H	664	0	697	49	1
3	A	12	0	0	3	1
3	B	8	0	0	1	0
3	C	7	0	0	2	0
3	D	4	0	0	2	0
3	E	7	0	0	0	0
3	F	5	0	0	0	0
3	G	7	0	0	1	1
3	H	11	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6717	0	6837	670	3

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 50.

The worst 5 of 670 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:E:26:LYS:HG3	1:E:31:SER:HB3	1.27	1.14
1:G:6:PHE:HA	1:G:85:VAL:HG11	1.14	1.11
1:E:8:LYS:HD2	1:E:14:GLU:HA	1.26	1.10
1:G:37:LYS:HE2	1:G:73:ARG:NE	1.68	1.07
1:G:37:LYS:HE2	1:G:73:ARG:HE	1.04	1.05

All (3) 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)
2:D:48:LYS:NZ	2:F:52:GLU:OE2[2_454]	1.89	0.31
3:A:207:HOH:O	3:G:205:HOH:O[1_554]	2.08	0.12
1:A:71:ARG:NH1	2:H:42:GLU:OE1[2_554]	2.11	0.09

## 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	112/141 (79%)	100 (89%)	12 (11%)	0	100	100
1	C	111/141 (79%)	99 (89%)	12 (11%)	0	100	100
1	E	114/141 (81%)	94 (82%)	14 (12%)	6 (5%)	2	6
1	G	115/141 (82%)	76 (66%)	26 (23%)	13 (11%)	0	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	89/97 (92%)	83 (93%)	5 (6%)	1 (1%)	14	41
2	D	90/97 (93%)	86 (96%)	4 (4%)	0	100	100
2	F	88/97 (91%)	82 (93%)	6 (7%)	0	100	100
2	H	84/97 (87%)	81 (96%)	3 (4%)	0	100	100
All	All	803/952 (84%)	701 (87%)	82 (10%)	20 (2%)	5	19

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	11	ASN
1	E	26	LYS
1	G	66	GLU
1	G	67	LEU
1	G	85	VAL

### 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	104/122 (85%)	96 (92%)	8 (8%)	13	35
1	C	103/122 (84%)	96 (93%)	7 (7%)	16	42
1	E	105/122 (86%)	93 (89%)	12 (11%)	5	18
1	G	106/122 (87%)	88 (83%)	18 (17%)	2	6
2	B	79/80 (99%)	72 (91%)	7 (9%)	9	28
2	D	80/80 (100%)	71 (89%)	9 (11%)	6	18
2	F	78/80 (98%)	74 (95%)	4 (5%)	24	55
2	H	76/80 (95%)	68 (90%)	8 (10%)	7	20
All	All	731/808 (90%)	658 (90%)	73 (10%)	7	22

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	25	SER
1	E	81	ASP
2	H	26	SER
1	E	58	LYS
1	E	112	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	104	GLN
1	G	104	GLN
1	G	88	HIS
1	C	13	ASN
1	G	11	ASN

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	110/141 (78%)	-0.05	2 (1%) 68 61	63, 82, 111, 126	0
1	C	109/141 (77%)	0.08	5 (4%) 32 22	54, 84, 127, 135	0
1	E	112/141 (79%)	0.59	14 (12%) 3 2	69, 110, 149, 165	0
1	G	113/141 (80%)	1.80	35 (30%) 0 0	89, 159, 200, 227	0
2	B	87/97 (89%)	-0.10	2 (2%) 60 51	61, 77, 95, 113	0
2	D	88/97 (90%)	-0.03	3 (3%) 45 35	55, 68, 101, 139	0
2	F	86/97 (88%)	-0.06	2 (2%) 60 51	57, 77, 108, 133	0
2	H	82/97 (84%)	0.06	3 (3%) 41 31	70, 96, 132, 148	0
All	All	787/952 (82%)	0.33	66 (8%) 11 5	54, 89, 165, 227	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	75	LEU	15.9
1	G	62	ALA	13.3
1	G	69	PRO	12.3
1	G	100	ARG	11.2
1	G	99	LYS	8.6

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.