



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 20, 2024 – 08:54 pm GMT

PDB ID : 7AVR  
Title : The tetrameric structure of haloalkane dehalogenase DpaA from *Paraglaciecola agarilytica* NO2  
Authors : Mazur, A.; Kolenko, P.; Prudnikova, T.; Grinkevich, P.; Kuta Smatanova, I.  
Deposited on : 2020-11-05  
Resolution : 2.00 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 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

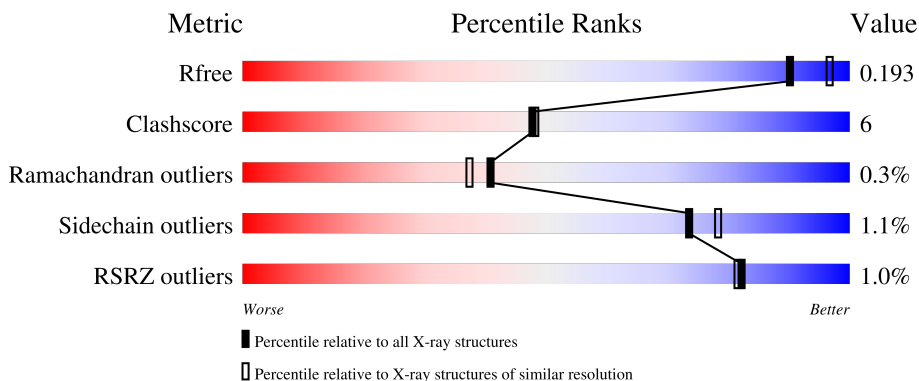
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	307	
1	B	307	
1	C	307	
1	D	307	
1	E	307	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	307	<p>2% 84% 13% .</p>
1	G	307	<p>2% 86% 11% ..</p>
1	H	307	<p>2% 88% 9% .</p>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CL	F	404	-	-	X	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 20821 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 Haloalkane dehalogenase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	299	2355	1513	380	444	18	0	1	0
1	B	298	2341	1503	379	441	18	0	0	0
1	C	299	2354	1511	381	444	18	1	1	0
1	D	297	2333	1498	378	440	17	1	0	0
1	E	298	2358	1513	380	447	18	0	3	0
1	F	298	2348	1508	381	442	17	0	1	0
1	G	299	2354	1511	381	444	18	0	1	0
1	H	299	2352	1512	381	441	18	0	1	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP K6XNL5
A	2	THR	-	expression tag	UNP K6XNL5
A	302	HIS	-	expression tag	UNP K6XNL5
A	303	HIS	-	expression tag	UNP K6XNL5
A	304	HIS	-	expression tag	UNP K6XNL5
A	305	HIS	-	expression tag	UNP K6XNL5
A	306	HIS	-	expression tag	UNP K6XNL5
A	307	HIS	-	expression tag	UNP K6XNL5
B	1	MET	-	initiating methionine	UNP K6XNL5
B	2	THR	-	expression tag	UNP K6XNL5
B	302	HIS	-	expression tag	UNP K6XNL5
B	303	HIS	-	expression tag	UNP K6XNL5
B	304	HIS	-	expression tag	UNP K6XNL5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	305	HIS	-	expression tag	UNP K6XNL5
B	306	HIS	-	expression tag	UNP K6XNL5
B	307	HIS	-	expression tag	UNP K6XNL5
C	1	MET	-	initiating methionine	UNP K6XNL5
C	2	THR	-	expression tag	UNP K6XNL5
C	302	HIS	-	expression tag	UNP K6XNL5
C	303	HIS	-	expression tag	UNP K6XNL5
C	304	HIS	-	expression tag	UNP K6XNL5
C	305	HIS	-	expression tag	UNP K6XNL5
C	306	HIS	-	expression tag	UNP K6XNL5
C	307	HIS	-	expression tag	UNP K6XNL5
D	1	MET	-	initiating methionine	UNP K6XNL5
D	2	THR	-	expression tag	UNP K6XNL5
D	302	HIS	-	expression tag	UNP K6XNL5
D	303	HIS	-	expression tag	UNP K6XNL5
D	304	HIS	-	expression tag	UNP K6XNL5
D	305	HIS	-	expression tag	UNP K6XNL5
D	306	HIS	-	expression tag	UNP K6XNL5
D	307	HIS	-	expression tag	UNP K6XNL5
E	1	MET	-	initiating methionine	UNP K6XNL5
E	2	THR	-	expression tag	UNP K6XNL5
E	302	HIS	-	expression tag	UNP K6XNL5
E	303	HIS	-	expression tag	UNP K6XNL5
E	304	HIS	-	expression tag	UNP K6XNL5
E	305	HIS	-	expression tag	UNP K6XNL5
E	306	HIS	-	expression tag	UNP K6XNL5
E	307	HIS	-	expression tag	UNP K6XNL5
F	1	MET	-	initiating methionine	UNP K6XNL5
F	2	THR	-	expression tag	UNP K6XNL5
F	302	HIS	-	expression tag	UNP K6XNL5
F	303	HIS	-	expression tag	UNP K6XNL5
F	304	HIS	-	expression tag	UNP K6XNL5
F	305	HIS	-	expression tag	UNP K6XNL5
F	306	HIS	-	expression tag	UNP K6XNL5
F	307	HIS	-	expression tag	UNP K6XNL5
G	1	MET	-	initiating methionine	UNP K6XNL5
G	2	THR	-	expression tag	UNP K6XNL5
G	302	HIS	-	expression tag	UNP K6XNL5
G	303	HIS	-	expression tag	UNP K6XNL5
G	304	HIS	-	expression tag	UNP K6XNL5
G	305	HIS	-	expression tag	UNP K6XNL5
G	306	HIS	-	expression tag	UNP K6XNL5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	307	HIS	-	expression tag	UNP K6XNL5
H	1	MET	-	initiating methionine	UNP K6XNL5
H	2	THR	-	expression tag	UNP K6XNL5
H	302	HIS	-	expression tag	UNP K6XNL5
H	303	HIS	-	expression tag	UNP K6XNL5
H	304	HIS	-	expression tag	UNP K6XNL5
H	305	HIS	-	expression tag	UNP K6XNL5
H	306	HIS	-	expression tag	UNP K6XNL5
H	307	HIS	-	expression tag	UNP K6XNL5

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0
2	B	1	Total Cl 1 1	0	0
2	C	3	Total Cl 3 3	0	0
2	D	1	Total Cl 1 1	0	0
2	E	2	Total Cl 2 2	0	0
2	F	6	Total Cl 6 6	0	0
2	G	3	Total Cl 3 3	0	0
2	H	1	Total Cl 1 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	264	Total O 264 264	0	0
3	B	244	Total O 244 244	0	0
3	C	291	Total O 291 291	0	0
3	D	217	Total O 217 217	0	0
3	E	264	Total O 264 264	0	0

*Continued on next page...*

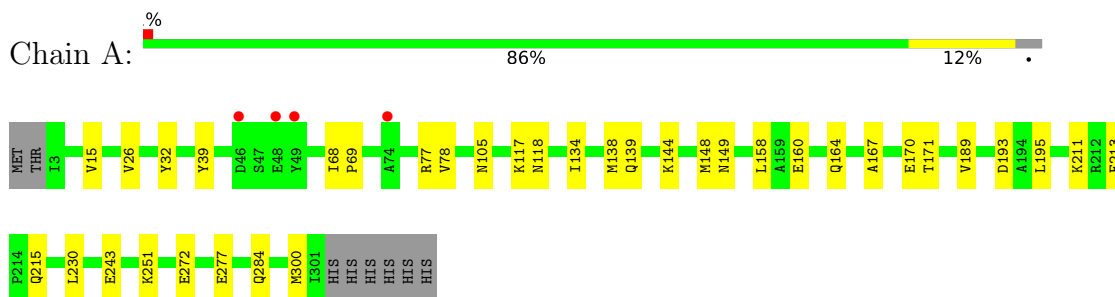
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	F	298	Total O 298 298	0	0
3	G	231	Total O 231 231	0	0
3	H	199	Total O 199 199	0	0

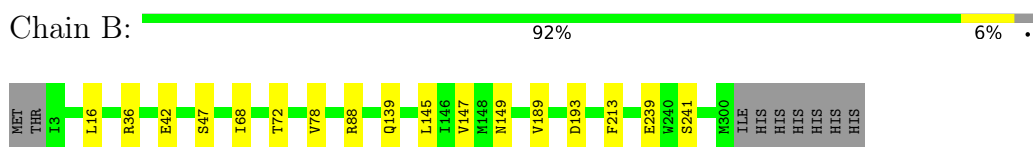
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

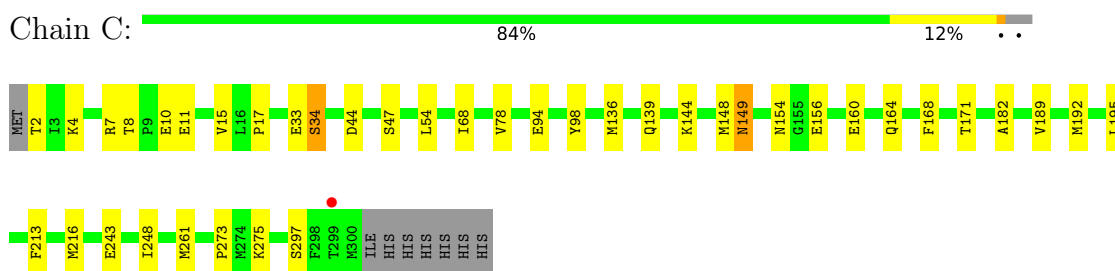
- Molecule 1: Haloalkane dehalogenase 1



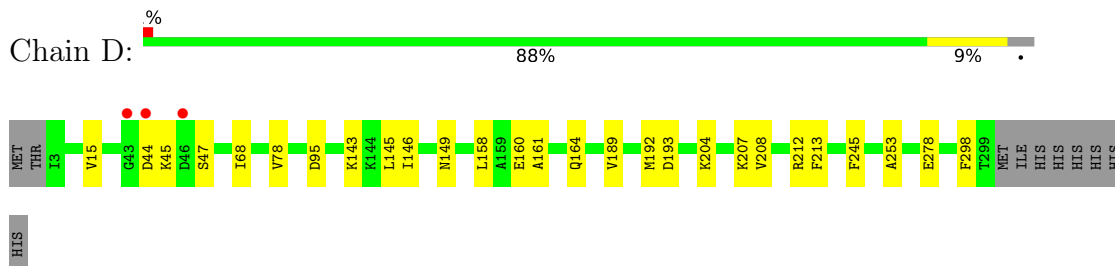
- Molecule 1: Haloalkane dehalogenase 1



- Molecule 1: Haloalkane dehalogenase 1

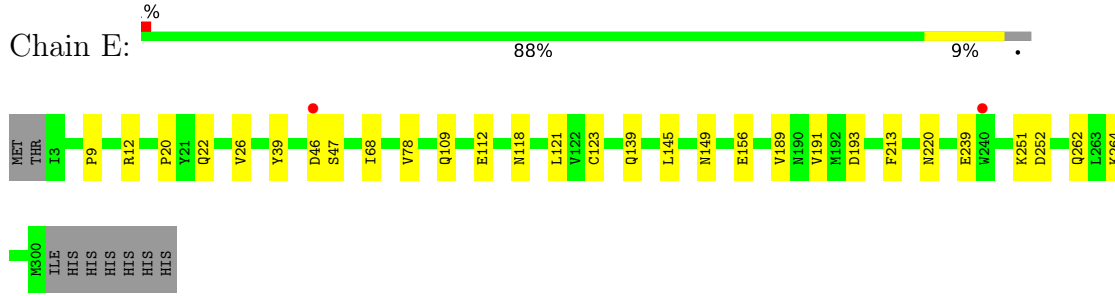


- Molecule 1: Haloalkane dehalogenase 1

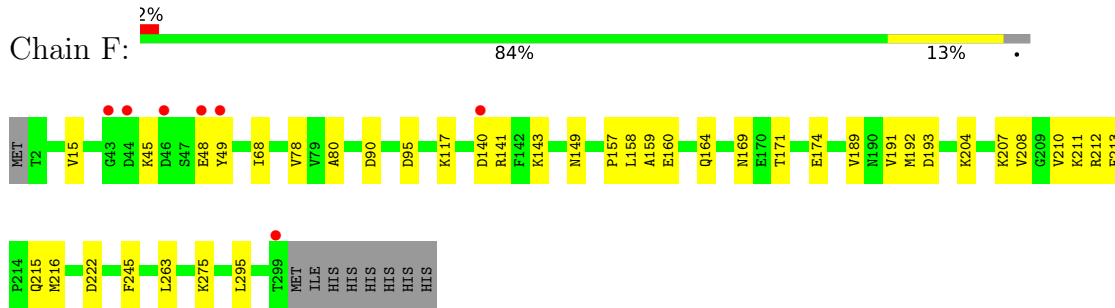




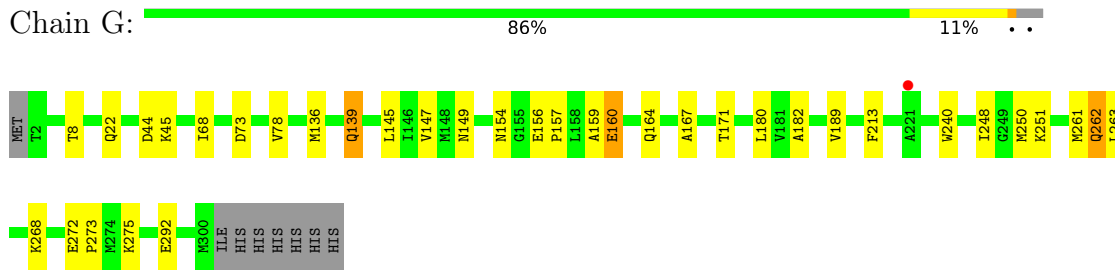
- Molecule 1: Haloalkane dehalogenase 1



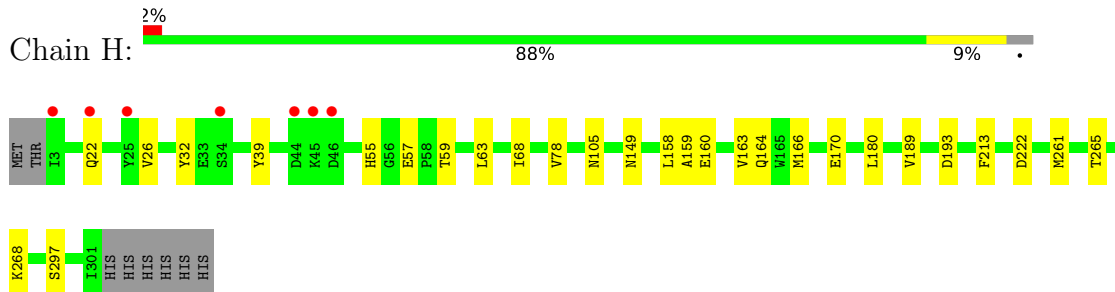
- Molecule 1: Haloalkane dehalogenase 1



- Molecule 1: Haloalkane dehalogenase 1



- Molecule 1: Haloalkane dehalogenase 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.39Å 155.51Å 155.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.44 – 2.00 47.44 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (47.44-2.00) 99.8 (47.44-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.22 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0189	Depositor
R, $R_{free}$	0.176 , 0.206 0.171 , 0.193	Depositor DCC
$R_{free}$ test set	9496 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtrriage
Anisotropy	0.023	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 55.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.105 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	20821	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 63.14 % 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 1.0126e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	1.02	0/2417	0.90	0/3279
1	B	0.99	0/2400	0.89	0/3256
1	C	1.03	0/2416	0.89	0/3278
1	D	0.98	0/2392	0.86	0/3246
1	E	0.96	0/2423	0.89	0/3288
1	F	1.01	0/2407	0.86	0/3266
1	G	1.02	0/2416	0.87	0/3279
1	H	0.94	0/2414	0.87	0/3276
All	All	1.00	0/19285	0.88	0/26168

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	2355	0	2291	35	0
1	B	2341	0	2274	14	0
1	C	2354	0	2289	40	0
1	D	2333	0	2265	23	0
1	E	2358	0	2288	25	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2348	0	2281	35	0
1	G	2354	0	2289	41	0
1	H	2352	0	2291	21	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	3	0	0	1	0
2	D	1	0	0	0	0
2	E	2	0	0	1	0
2	F	6	0	0	2	0
2	G	3	0	0	1	0
2	H	1	0	0	1	0
3	A	264	0	0	18	0
3	B	244	0	0	4	0
3	C	291	0	0	13	0
3	D	217	0	0	4	0
3	E	264	0	0	9	0
3	F	298	0	0	19	0
3	G	231	0	0	9	0
3	H	199	0	0	9	0
All	All	20821	0	18268	226	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 6.

The worst 5 of 226 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:G:156:GLU:HG2	1:G:157:PRO:HD2	1.37	1.03
1:G:182:ALA:HA	1:G:189:VAL:HG21	1.59	0.83
1:E:251:LYS:HZ3	1:G:251:LYS:HD2	1.44	0.83
1:C:182:ALA:HA	1:C:189:VAL:HG21	1.63	0.79
1:G:156:GLU:HG2	1:G:157:PRO:CD	2.11	0.79

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	298/307 (97%)	284 (95%)	13 (4%)	1 (0%)	41	37
1	B	296/307 (96%)	283 (96%)	12 (4%)	1 (0%)	41	37
1	C	298/307 (97%)	287 (96%)	10 (3%)	1 (0%)	41	37
1	D	295/307 (96%)	284 (96%)	10 (3%)	1 (0%)	41	37
1	E	299/307 (97%)	288 (96%)	10 (3%)	1 (0%)	41	37
1	F	297/307 (97%)	285 (96%)	11 (4%)	1 (0%)	41	37
1	G	298/307 (97%)	286 (96%)	11 (4%)	1 (0%)	41	37
1	H	298/307 (97%)	284 (95%)	13 (4%)	1 (0%)	41	37
All	All	2379/2456 (97%)	2281 (96%)	90 (4%)	8 (0%)	41	37

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	149	ASN
1	A	149	ASN
1	B	149	ASN
1	D	149	ASN
1	E	149	ASN

### 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	250/257 (97%)	248 (99%)	2 (1%)	81	86

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	248/257 (96%)	245 (99%)	3 (1%)	71	76
1	C	250/257 (97%)	247 (99%)	3 (1%)	71	76
1	D	247/257 (96%)	246 (100%)	1 (0%)	91	93
1	E	251/257 (98%)	248 (99%)	3 (1%)	71	76
1	F	248/257 (96%)	247 (100%)	1 (0%)	91	93
1	G	250/257 (97%)	245 (98%)	5 (2%)	55	58
1	H	249/257 (97%)	244 (98%)	5 (2%)	55	58
All	All	1993/2056 (97%)	1970 (99%)	23 (1%)	73	76

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	139	GLN
1	G	262	GLN
1	G	213	PHE
1	H	22[A]	GLN
1	C	34	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	HIS
1	B	164	GLN
1	H	76	HIS

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry

Of 18 ligands modelled in this entry, 18 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	299/307 (97%)	-0.09	4 (1%) 77 76	13, 24, 43, 52	2 (0%)
1	B	298/307 (97%)	-0.16	0 100 100	14, 27, 44, 67	1 (0%)
1	C	299/307 (97%)	-0.08	1 (0%) 94 93	14, 26, 41, 51	1 (0%)
1	D	297/307 (96%)	-0.16	3 (1%) 82 81	15, 26, 42, 64	1 (0%)
1	E	298/307 (97%)	0.09	2 (0%) 87 87	17, 29, 41, 61	0
1	F	298/307 (97%)	0.01	7 (2%) 60 59	15, 27, 42, 66	0
1	G	299/307 (97%)	-0.13	1 (0%) 94 93	14, 24, 40, 54	1 (0%)
1	H	299/307 (97%)	-0.04	7 (2%) 60 59	17, 26, 44, 65	0
All	All	2387/2456 (97%)	-0.07	25 (1%) 82 81	13, 26, 42, 67	6 (0%)

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	44	ASP	4.6
1	F	140	ASP	4.4
1	A	46	ASP	4.3
1	A	49	TYR	3.8
1	H	3	ILE	3.5

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	CL	F	401	1/1	0.93	0.12	58,58,58,58	0
2	CL	C	403	1/1	0.94	0.12	47,47,47,47	0
2	CL	F	403	1/1	0.94	0.07	42,42,42,42	0
2	CL	E	402	1/1	0.97	0.08	53,53,53,53	0
2	CL	F	405	1/1	0.97	0.05	33,33,33,33	0
2	CL	C	402	1/1	0.98	0.09	34,34,34,34	0
2	CL	F	404	1/1	0.98	0.07	35,35,35,35	0
2	CL	A	401	1/1	0.98	0.09	18,18,18,18	0
2	CL	F	406	1/1	0.98	0.13	50,50,50,50	0
2	CL	H	401	1/1	0.98	0.08	21,21,21,21	0
2	CL	G	401	1/1	0.99	0.07	22,22,22,22	0
2	CL	G	402	1/1	0.99	0.07	30,30,30,30	0
2	CL	G	403	1/1	0.99	0.06	29,29,29,29	0
2	CL	F	402	1/1	0.99	0.07	19,19,19,19	0
2	CL	E	401	1/1	1.00	0.05	24,24,24,24	0
2	CL	B	401	1/1	1.00	0.09	17,17,17,17	0
2	CL	C	401	1/1	1.00	0.09	21,21,21,21	0
2	CL	D	401	1/1	1.00	0.04	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.