



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 01:47 AM UTC

PDB ID : 9CQE / pdb_00009cqe
Title : CRYSTAL STRUCTURE OF APO C-TERMINAL HIS-TAG DOG
HSP47(36-418) IN A P 1 21 1 CRYSTAL FORM
Authors : Sheriff, S.
Deposited on : 2024-07-19
Resolution : 3.18 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

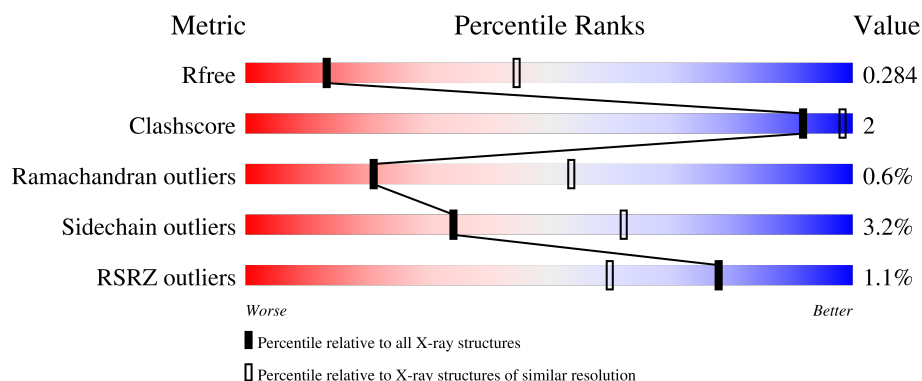
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.18 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}	180053	2001 (3.20-3.16)
Clashscore	190562	2119 (3.20-3.16)
Ramachandran outliers	187476	2070 (3.20-3.16)
Sidechain outliers	187428	2069 (3.20-3.16)
RSRZ outliers	180081	2001 (3.20-3.16)

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	392	% 90% 6% ..
1	B	392	89% 9% .
1	D	392	% 86% 10% ..
1	E	392	% 89% 9% .
1	F	392	% 91% 6% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	392	% 89% 8% .
1	H	392	% 89% 8% ..
1	I	392	2% 91% 6% .

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 45775 atoms, of which 22579 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 Serpin H1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	381	Total	C	H	N	O	S	2884	0	0
			5807	1853	2884	512	544	14			
1	B	382	Total	C	H	N	O	S	2887	0	0
			5821	1862	2887	513	545	14			
1	D	382	Total	C	H	N	O	S	2897	0	0
			5824	1863	2897	509	541	14			
1	E	381	Total	C	H	N	O	S	2786	0	0
			5660	1826	2786	500	535	13			
1	F	381	Total	C	H	N	O	S	2784	0	0
			5665	1828	2784	499	540	14			
1	G	384	Total	C	H	N	O	S	2805	0	0
			5708	1838	2805	511	541	13			
1	H	382	Total	C	H	N	O	S	2851	0	0
			5762	1848	2851	509	540	14			
1	I	381	Total	C	H	N	O	S	2685	0	0
			5528	1794	2685	499	537	13			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	35	MET	-	initiating methionine	UNP C7C419
A	419	LEU	-	expression tag	UNP C7C419
A	420	GLU	-	expression tag	UNP C7C419
A	421	HIS	-	expression tag	UNP C7C419
A	422	HIS	-	expression tag	UNP C7C419
A	423	HIS	-	expression tag	UNP C7C419
A	424	HIS	-	expression tag	UNP C7C419
A	425	HIS	-	expression tag	UNP C7C419
A	426	HIS	-	expression tag	UNP C7C419
B	35	MET	-	initiating methionine	UNP C7C419
B	419	LEU	-	expression tag	UNP C7C419
B	420	GLU	-	expression tag	UNP C7C419
B	421	HIS	-	expression tag	UNP C7C419

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	422	HIS	-	expression tag	UNP C7C419
B	423	HIS	-	expression tag	UNP C7C419
B	424	HIS	-	expression tag	UNP C7C419
B	425	HIS	-	expression tag	UNP C7C419
B	426	HIS	-	expression tag	UNP C7C419
D	35	MET	-	initiating methionine	UNP C7C419
D	419	LEU	-	expression tag	UNP C7C419
D	420	GLU	-	expression tag	UNP C7C419
D	421	HIS	-	expression tag	UNP C7C419
D	422	HIS	-	expression tag	UNP C7C419
D	423	HIS	-	expression tag	UNP C7C419
D	424	HIS	-	expression tag	UNP C7C419
D	425	HIS	-	expression tag	UNP C7C419
D	426	HIS	-	expression tag	UNP C7C419
E	35	MET	-	initiating methionine	UNP C7C419
E	419	LEU	-	expression tag	UNP C7C419
E	420	GLU	-	expression tag	UNP C7C419
E	421	HIS	-	expression tag	UNP C7C419
E	422	HIS	-	expression tag	UNP C7C419
E	423	HIS	-	expression tag	UNP C7C419
E	424	HIS	-	expression tag	UNP C7C419
E	425	HIS	-	expression tag	UNP C7C419
E	426	HIS	-	expression tag	UNP C7C419
F	35	MET	-	initiating methionine	UNP C7C419
F	419	LEU	-	expression tag	UNP C7C419
F	420	GLU	-	expression tag	UNP C7C419
F	421	HIS	-	expression tag	UNP C7C419
F	422	HIS	-	expression tag	UNP C7C419
F	423	HIS	-	expression tag	UNP C7C419
F	424	HIS	-	expression tag	UNP C7C419
F	425	HIS	-	expression tag	UNP C7C419
F	426	HIS	-	expression tag	UNP C7C419
G	35	MET	-	initiating methionine	UNP C7C419
G	419	LEU	-	expression tag	UNP C7C419
G	420	GLU	-	expression tag	UNP C7C419
G	421	HIS	-	expression tag	UNP C7C419
G	422	HIS	-	expression tag	UNP C7C419
G	423	HIS	-	expression tag	UNP C7C419
G	424	HIS	-	expression tag	UNP C7C419
G	425	HIS	-	expression tag	UNP C7C419
G	426	HIS	-	expression tag	UNP C7C419
H	35	MET	-	initiating methionine	UNP C7C419

Continued on next page...

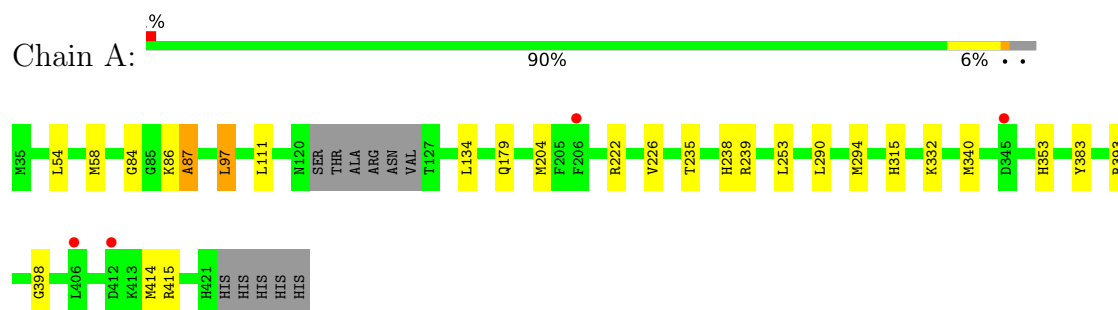
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	419	LEU	-	expression tag	UNP C7C419
H	420	GLU	-	expression tag	UNP C7C419
H	421	HIS	-	expression tag	UNP C7C419
H	422	HIS	-	expression tag	UNP C7C419
H	423	HIS	-	expression tag	UNP C7C419
H	424	HIS	-	expression tag	UNP C7C419
H	425	HIS	-	expression tag	UNP C7C419
H	426	HIS	-	expression tag	UNP C7C419
I	35	MET	-	initiating methionine	UNP C7C419
I	419	LEU	-	expression tag	UNP C7C419
I	420	GLU	-	expression tag	UNP C7C419
I	421	HIS	-	expression tag	UNP C7C419
I	422	HIS	-	expression tag	UNP C7C419
I	423	HIS	-	expression tag	UNP C7C419
I	424	HIS	-	expression tag	UNP C7C419
I	425	HIS	-	expression tag	UNP C7C419
I	426	HIS	-	expression tag	UNP C7C419

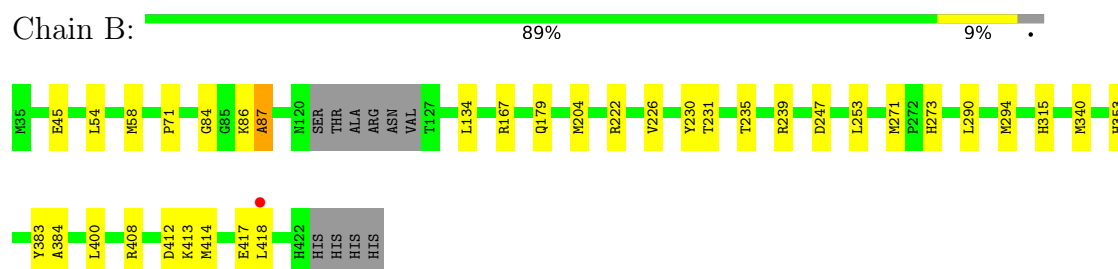
3 Residue-property plots [i](#)

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.

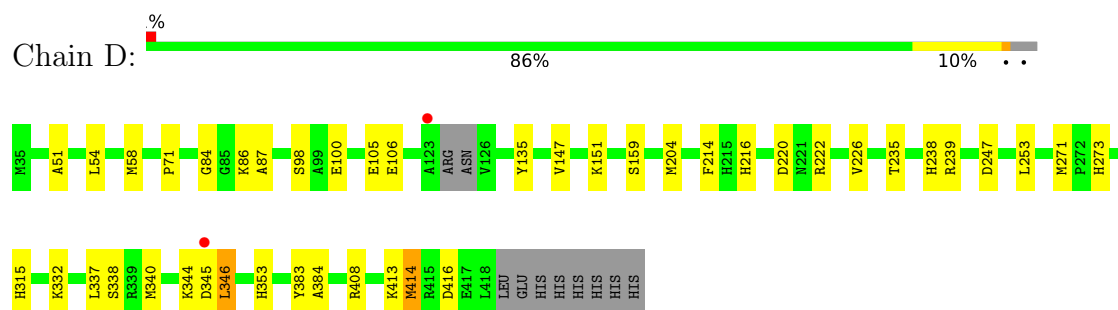
- Molecule 1: Serpin H1



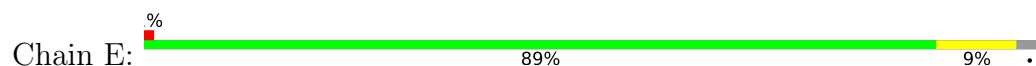
- Molecule 1: Serpin H1



- Molecule 1: Serpin H1

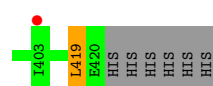
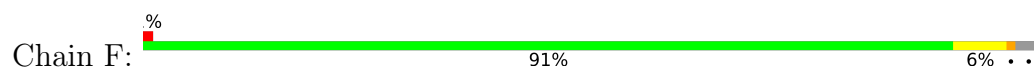


- Molecule 1: Serpin H1

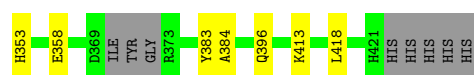
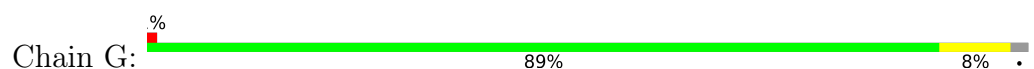




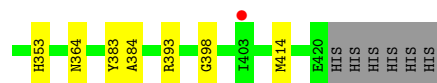
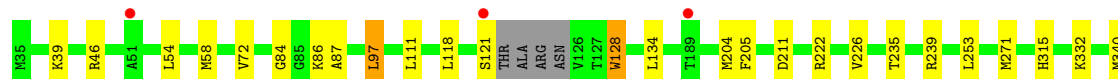
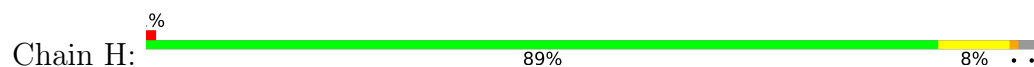
• Molecule 1: Serpin H1



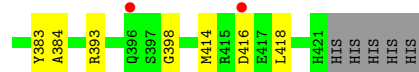
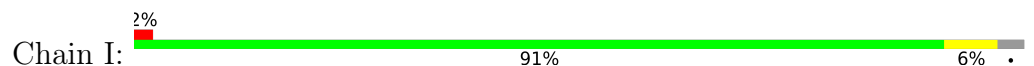
• Molecule 1: Serpin H1



• Molecule 1: Serpin H1



• Molecule 1: Serpin H1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	140.18Å 75.31Å 163.87Å 90.00° 110.80° 90.00°	Depositor
Resolution (Å)	45.85 – 3.18 45.85 – 3.18	Depositor EDS
% Data completeness (in resolution range)	73.1 (45.85-3.18) 73.0 (45.85-3.18)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 3.19Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
R, R_{free}	0.263 , 0.288 0.262 , 0.284	Depositor DCC
R_{free} test set	2024 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	82.2	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 34.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	45775	wwPDB-VP
Average B, all atoms (Å ²)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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

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	0.70	1/2980 (0.0%)	0.97	3/4026 (0.1%)
1	B	0.72	1/2993 (0.0%)	0.96	1/4046 (0.0%)
1	D	0.75	1/2985 (0.0%)	1.05	5/4033 (0.1%)
1	E	0.71	0/2932	0.96	1/3972 (0.0%)
1	F	0.72	2/2939 (0.1%)	0.97	2/3983 (0.1%)
1	G	0.74	2/2961 (0.1%)	0.99	4/4009 (0.1%)
1	H	0.71	1/2969 (0.0%)	0.96	1/4017 (0.0%)
1	I	0.70	1/2901 (0.0%)	0.98	4/3932 (0.1%)
All	All	0.72	9/23660 (0.0%)	0.98	21/32018 (0.1%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	371	TYR	CA-C	5.59	1.58	1.52
1	D	87	ALA	CA-C	5.57	1.61	1.53
1	G	262	HIS	CA-C	5.26	1.59	1.52
1	G	87	ALA	CA-C	5.16	1.59	1.52
1	I	87	ALA	CA-C	5.15	1.59	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	414	MET	CA-C-N	6.84	134.01	121.70
1	A	414	MET	C-N-CA	6.84	134.01	121.70
1	E	226	VAL	N-CA-C	-6.55	104.37	110.53
1	I	345	ASP	CA-CB-CG	5.87	118.47	112.60
1	G	120	ASN	CA-C-N	5.78	132.58	121.54

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	2923	2884	2884	9	0
1	B	2934	2887	2885	14	0
1	D	2927	2897	2894	14	0
1	E	2874	2786	2782	11	0
1	F	2881	2784	2784	6	0
1	G	2903	2805	2802	9	0
1	H	2911	2851	2851	11	0
1	I	2843	2685	2683	7	0
All	All	23196	22579	22565	75	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 2.

The worst 5 of 75 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:F:244:ASN:OD1	1:F:299:LYS:HG3	1.74	0.88
1:B:230:TYR:OH	1:B:408:ARG:NH2	2.20	0.74
1:D:51:ALA:HB2	1:D:71:PRO:HB3	1.78	0.65
1:D:247:ASP:OD1	1:D:273:HIS:NE2	2.32	0.61
1:D:337:LEU:HD12	1:D:346:LEU:O	2.02	0.59

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	377/392 (96%)	357 (95%)	18 (5%)	2 (0%)	24	57
1	B	378/392 (96%)	359 (95%)	18 (5%)	1 (0%)	36	66
1	D	378/392 (96%)	355 (94%)	20 (5%)	3 (1%)	16	47
1	E	377/392 (96%)	359 (95%)	16 (4%)	2 (0%)	24	57
1	F	377/392 (96%)	355 (94%)	20 (5%)	2 (0%)	24	57
1	G	380/392 (97%)	356 (94%)	20 (5%)	4 (1%)	11	41
1	H	378/392 (96%)	356 (94%)	21 (6%)	1 (0%)	36	66
1	I	377/392 (96%)	359 (95%)	16 (4%)	2 (0%)	24	57
All	All	3022/3136 (96%)	2856 (94%)	149 (5%)	17 (1%)	21	53

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	416	ASP
1	F	419	LEU
1	G	121	SER
1	G	261	ALA
1	G	263	LYS

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	305/337 (90%)	297 (97%)	8 (3%)	40	66
1	B	306/337 (91%)	297 (97%)	9 (3%)	37	64
1	D	304/337 (90%)	293 (96%)	11 (4%)	31	60
1	E	292/337 (87%)	282 (97%)	10 (3%)	32	61
1	F	295/337 (88%)	286 (97%)	9 (3%)	35	62
1	G	294/337 (87%)	284 (97%)	10 (3%)	32	61
1	H	301/337 (89%)	290 (96%)	11 (4%)	30	59
1	I	282/337 (84%)	274 (97%)	8 (3%)	38	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2379/2696 (88%)	2303 (97%)	76 (3%)	34 62

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	39	LYS
1	I	239	ARG
1	H	86	LYS
1	H	253	LEU
1	I	418	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	353	HIS
1	I	158	HIS
1	F	353	HIS
1	I	353	HIS
1	G	396	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	381/392 (97%)	-0.15	4 (1%) 79 62	19, 38, 64, 77	0
1	B	382/392 (97%)	-0.03	1 (0%) 90 81	33, 51, 78, 90	0
1	D	382/392 (97%)	-0.13	2 (0%) 87 75	18, 35, 56, 78	0
1	E	381/392 (97%)	0.01	4 (1%) 79 62	30, 52, 79, 95	0
1	F	381/392 (97%)	0.04	4 (1%) 79 62	29, 47, 74, 82	0
1	G	384/392 (97%)	0.06	5 (1%) 75 55	31, 52, 76, 91	0
1	H	382/392 (97%)	-0.11	4 (1%) 79 62	22, 39, 65, 80	0
1	I	381/392 (97%)	0.15	9 (2%) 59 39	30, 55, 93, 105	0
All	All	3054/3136 (97%)	-0.02	33 (1%) 78 60	18, 46, 77, 105	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	198	ALA	4.6
1	D	345	ASP	3.6
1	E	122	THR	3.3
1	I	198	ALA	3.3
1	F	345	ASP	3.3

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 oligosaccharides 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.