



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 2, 2024 – 10:09 AM EST

PDB ID : 1SYR
Title : Initial Structural Analysis of Plasmodium falciparum thioredoxin
Authors : Robien, M.A.; Hol, W.G.J.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)
Deposited on : 2004-04-01
Resolution : 2.95 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 : 4.02b-467
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

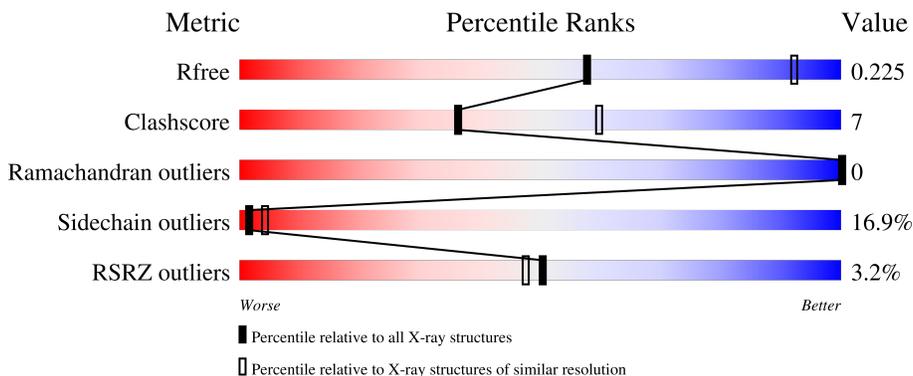
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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}	164625	1044 (2.98-2.94)
Clashscore	180529	1097 (2.98-2.94)
Ramachandran outliers	177936	1049 (2.98-2.94)
Sidechain outliers	177891	1049 (2.98-2.94)
RSRZ outliers	164620	1044 (2.98-2.94)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	112	 9% 66% 22% • 8%
1	B	112	 4% 68% 21% • 9%
1	C	112	 66% 25% • 8%
1	D	112	 61% 27% • 9%
1	E	112	 2% 62% 27% • 9%
1	F	112	 68% 18% 5% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	112	<p>66% 21% 9%</p>
1	H	112	<p>61% 25% 12%</p>
1	I	112	<p>59% 30% 9%</p>
1	J	112	<p>67% 21% 9%</p>
1	K	112	<p>57% 30% 9%</p>
1	L	112	<p>60% 28% 9%</p>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9662 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 thioredoxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	103	Total 815	C 523	N 124	O 162	S 6	0	0	0
1	B	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	C	103	Total 812	C 521	N 124	O 162	S 5	0	0	0
1	D	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	E	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	F	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	G	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	H	98	Total 772	C 498	N 117	O 152	S 5	0	0	0
1	I	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	J	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	K	102	Total 807	C 518	N 123	O 161	S 5	0	0	0
1	L	102	Total 807	C 518	N 123	O 161	S 5	0	0	0

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q7KQL8
A	2	ALA	-	expression tag	UNP Q7KQL8
A	3	HIS	-	expression tag	UNP Q7KQL8
A	4	HIS	-	expression tag	UNP Q7KQL8
A	5	HIS	-	expression tag	UNP Q7KQL8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	HIS	-	expression tag	UNP Q7KQL8
A	7	HIS	-	expression tag	UNP Q7KQL8
A	8	HIS	-	expression tag	UNP Q7KQL8
B	1	MET	-	expression tag	UNP Q7KQL8
B	2	ALA	-	expression tag	UNP Q7KQL8
B	3	HIS	-	expression tag	UNP Q7KQL8
B	4	HIS	-	expression tag	UNP Q7KQL8
B	5	HIS	-	expression tag	UNP Q7KQL8
B	6	HIS	-	expression tag	UNP Q7KQL8
B	7	HIS	-	expression tag	UNP Q7KQL8
B	8	HIS	-	expression tag	UNP Q7KQL8
C	1	MET	-	expression tag	UNP Q7KQL8
C	2	ALA	-	expression tag	UNP Q7KQL8
C	3	HIS	-	expression tag	UNP Q7KQL8
C	4	HIS	-	expression tag	UNP Q7KQL8
C	5	HIS	-	expression tag	UNP Q7KQL8
C	6	HIS	-	expression tag	UNP Q7KQL8
C	7	HIS	-	expression tag	UNP Q7KQL8
C	8	HIS	-	expression tag	UNP Q7KQL8
D	1	MET	-	expression tag	UNP Q7KQL8
D	2	ALA	-	expression tag	UNP Q7KQL8
D	3	HIS	-	expression tag	UNP Q7KQL8
D	4	HIS	-	expression tag	UNP Q7KQL8
D	5	HIS	-	expression tag	UNP Q7KQL8
D	6	HIS	-	expression tag	UNP Q7KQL8
D	7	HIS	-	expression tag	UNP Q7KQL8
D	8	HIS	-	expression tag	UNP Q7KQL8
E	1	MET	-	expression tag	UNP Q7KQL8
E	2	ALA	-	expression tag	UNP Q7KQL8
E	3	HIS	-	expression tag	UNP Q7KQL8
E	4	HIS	-	expression tag	UNP Q7KQL8
E	5	HIS	-	expression tag	UNP Q7KQL8
E	6	HIS	-	expression tag	UNP Q7KQL8
E	7	HIS	-	expression tag	UNP Q7KQL8
E	8	HIS	-	expression tag	UNP Q7KQL8
F	1	MET	-	expression tag	UNP Q7KQL8
F	2	ALA	-	expression tag	UNP Q7KQL8
F	3	HIS	-	expression tag	UNP Q7KQL8
F	4	HIS	-	expression tag	UNP Q7KQL8
F	5	HIS	-	expression tag	UNP Q7KQL8
F	6	HIS	-	expression tag	UNP Q7KQL8
F	7	HIS	-	expression tag	UNP Q7KQL8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	8	HIS	-	expression tag	UNP Q7KQL8
G	1	MET	-	expression tag	UNP Q7KQL8
G	2	ALA	-	expression tag	UNP Q7KQL8
G	3	HIS	-	expression tag	UNP Q7KQL8
G	4	HIS	-	expression tag	UNP Q7KQL8
G	5	HIS	-	expression tag	UNP Q7KQL8
G	6	HIS	-	expression tag	UNP Q7KQL8
G	7	HIS	-	expression tag	UNP Q7KQL8
G	8	HIS	-	expression tag	UNP Q7KQL8
H	1	MET	-	expression tag	UNP Q7KQL8
H	2	ALA	-	expression tag	UNP Q7KQL8
H	3	HIS	-	expression tag	UNP Q7KQL8
H	4	HIS	-	expression tag	UNP Q7KQL8
H	5	HIS	-	expression tag	UNP Q7KQL8
H	6	HIS	-	expression tag	UNP Q7KQL8
H	7	HIS	-	expression tag	UNP Q7KQL8
H	8	HIS	-	expression tag	UNP Q7KQL8
I	1	MET	-	expression tag	UNP Q7KQL8
I	2	ALA	-	expression tag	UNP Q7KQL8
I	3	HIS	-	expression tag	UNP Q7KQL8
I	4	HIS	-	expression tag	UNP Q7KQL8
I	5	HIS	-	expression tag	UNP Q7KQL8
I	6	HIS	-	expression tag	UNP Q7KQL8
I	7	HIS	-	expression tag	UNP Q7KQL8
I	8	HIS	-	expression tag	UNP Q7KQL8
J	1	MET	-	expression tag	UNP Q7KQL8
J	2	ALA	-	expression tag	UNP Q7KQL8
J	3	HIS	-	expression tag	UNP Q7KQL8
J	4	HIS	-	expression tag	UNP Q7KQL8
J	5	HIS	-	expression tag	UNP Q7KQL8
J	6	HIS	-	expression tag	UNP Q7KQL8
J	7	HIS	-	expression tag	UNP Q7KQL8
J	8	HIS	-	expression tag	UNP Q7KQL8
K	1	MET	-	expression tag	UNP Q7KQL8
K	2	ALA	-	expression tag	UNP Q7KQL8
K	3	HIS	-	expression tag	UNP Q7KQL8
K	4	HIS	-	expression tag	UNP Q7KQL8
K	5	HIS	-	expression tag	UNP Q7KQL8
K	6	HIS	-	expression tag	UNP Q7KQL8
K	7	HIS	-	expression tag	UNP Q7KQL8
K	8	HIS	-	expression tag	UNP Q7KQL8
L	1	MET	-	expression tag	UNP Q7KQL8

Continued on next page...

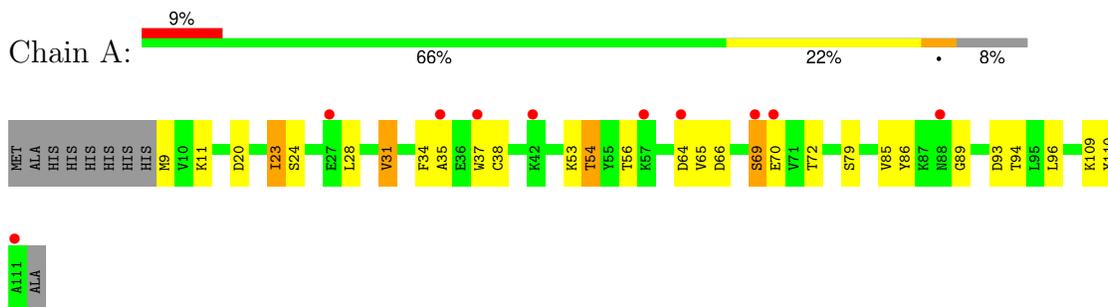
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	2	ALA	-	expression tag	UNP Q7KQL8
L	3	HIS	-	expression tag	UNP Q7KQL8
L	4	HIS	-	expression tag	UNP Q7KQL8
L	5	HIS	-	expression tag	UNP Q7KQL8
L	6	HIS	-	expression tag	UNP Q7KQL8
L	7	HIS	-	expression tag	UNP Q7KQL8
L	8	HIS	-	expression tag	UNP Q7KQL8

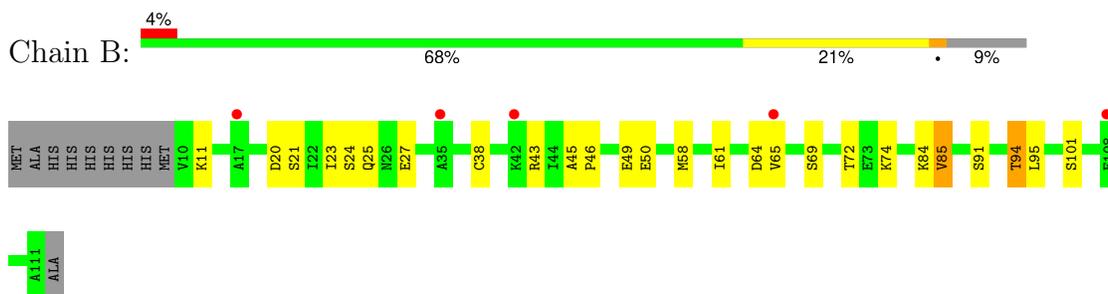
3 Residue-property plots [i](#)

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. 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. 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: thioredoxin



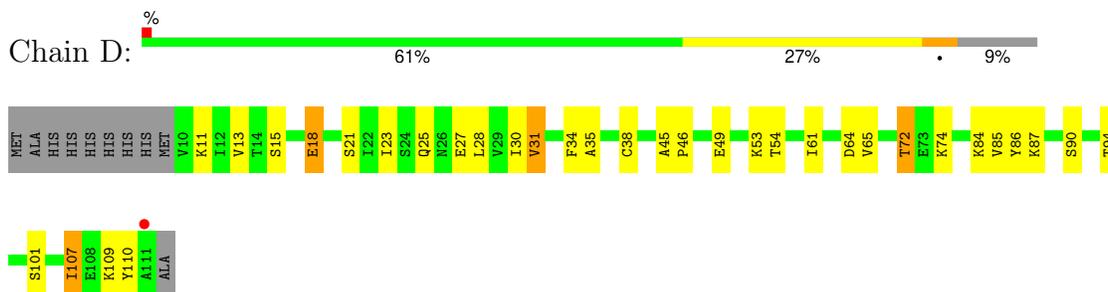
- Molecule 1: thioredoxin



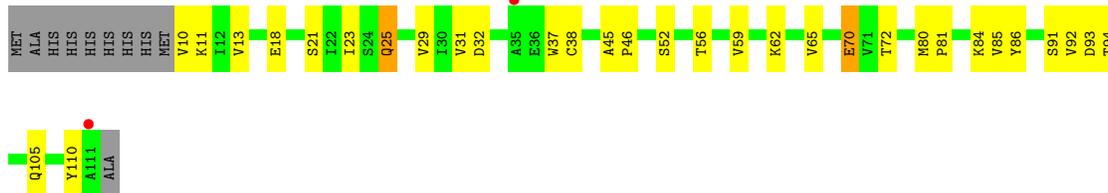
- Molecule 1: thioredoxin



- Molecule 1: thioredoxin



- Molecule 1: thioredoxin



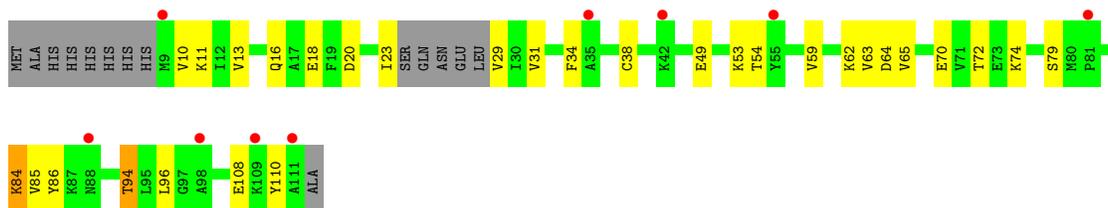
- Molecule 1: thioredoxin



- Molecule 1: thioredoxin



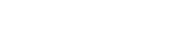
- Molecule 1: thioredoxin



- Molecule 1: thioredoxin



- Molecule 1: thioredoxin

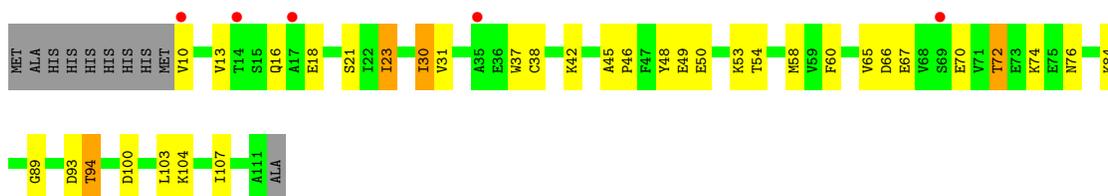




- Molecule 1: thioredoxin



- Molecule 1: thioredoxin



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	108.19Å 201.31Å 200.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.95 30.00 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.3 (30.00-2.95) 99.2 (30.00-2.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 2.95Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.197 , 0.236 0.191 , 0.225	Depositor DCC
R_{free} test set	2330 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	85.8	Xtrriage
Anisotropy	0.053	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 69.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9662	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% 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 [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	0.47	0/829	0.72	4/1120 (0.4%)
1	B	0.44	0/821	0.72	2/1110 (0.2%)
1	C	0.45	0/826	0.70	1/1117 (0.1%)
1	D	0.46	0/821	0.70	0/1110
1	E	0.43	0/821	0.71	2/1110 (0.2%)
1	F	0.47	0/821	0.71	3/1110 (0.3%)
1	G	0.50	0/821	0.72	3/1110 (0.3%)
1	H	0.50	0/785	0.74	2/1060 (0.2%)
1	I	0.46	0/821	0.73	3/1110 (0.3%)
1	J	0.47	0/821	0.72	1/1110 (0.1%)
1	K	0.49	0/821	0.74	3/1110 (0.3%)
1	L	0.45	0/821	0.71	3/1110 (0.3%)
All	All	0.47	0/9829	0.72	27/13287 (0.2%)

There are no bond length outliers.

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	64	ASP	CB-CG-OD2	6.13	123.82	118.30
1	H	64	ASP	CB-CG-OD2	6.00	123.70	118.30
1	J	66	ASP	CB-CG-OD2	5.83	123.55	118.30
1	L	93	ASP	CB-CG-OD2	5.81	123.53	118.30
1	K	20	ASP	CB-CG-OD2	5.78	123.50	118.30

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	815	0	810	15	0
1	B	807	0	801	8	0
1	C	812	0	803	10	0
1	D	807	0	801	15	0
1	E	807	0	801	12	0
1	F	807	0	801	6	0
1	G	807	0	801	13	0
1	H	772	0	766	16	0
1	I	807	0	801	16	0
1	J	807	0	801	11	0
1	K	807	0	801	13	0
1	L	807	0	801	15	0
All	All	9662	0	9588	136	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 7.

The worst 5 of 136 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:L:30:ILE:HD11	1:L:60:PHE:CD2	2.16	0.81
1:H:65:VAL:HA	1:H:72:THR:HG21	1.65	0.76
1:G:65:VAL:HA	1:G:72:THR:HG21	1.67	0.75
1:H:31:VAL:CG1	1:H:86:TYR:HE1	2.00	0.75
1:H:13:VAL:CG2	1:H:63:VAL:HG22	2.15	0.75

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	101/112 (90%)	96 (95%)	5 (5%)	0	100	100
1	B	100/112 (89%)	96 (96%)	4 (4%)	0	100	100
1	C	101/112 (90%)	97 (96%)	4 (4%)	0	100	100
1	D	100/112 (89%)	96 (96%)	4 (4%)	0	100	100
1	E	100/112 (89%)	96 (96%)	4 (4%)	0	100	100
1	F	100/112 (89%)	95 (95%)	5 (5%)	0	100	100
1	G	100/112 (89%)	95 (95%)	5 (5%)	0	100	100
1	H	94/112 (84%)	91 (97%)	3 (3%)	0	100	100
1	I	100/112 (89%)	96 (96%)	4 (4%)	0	100	100
1	J	100/112 (89%)	95 (95%)	5 (5%)	0	100	100
1	K	100/112 (89%)	94 (94%)	6 (6%)	0	100	100
1	L	100/112 (89%)	96 (96%)	4 (4%)	0	100	100
All	All	1196/1344 (89%)	1143 (96%)	53 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	94/101 (93%)	82 (87%)	12 (13%)	3	10
1	B	93/101 (92%)	80 (86%)	13 (14%)	3	8
1	C	93/101 (92%)	78 (84%)	15 (16%)	2	5
1	D	93/101 (92%)	77 (83%)	16 (17%)	1	4
1	E	93/101 (92%)	79 (85%)	14 (15%)	2	6
1	F	93/101 (92%)	72 (77%)	21 (23%)	1	1
1	G	93/101 (92%)	78 (84%)	15 (16%)	2	5
1	H	88/101 (87%)	74 (84%)	14 (16%)	2	5
1	I	93/101 (92%)	80 (86%)	13 (14%)	3	8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	93/101 (92%)	76 (82%)	17 (18%)	1	3
1	K	93/101 (92%)	71 (76%)	22 (24%)	0	0
1	L	93/101 (92%)	77 (83%)	16 (17%)	1	4
All	All	1112/1212 (92%)	924 (83%)	188 (17%)	1	4

5 of 188 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	70	GLU
1	J	74	LYS
1	H	108	GLU
1	I	87	LYS
1	K	10	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	16	GLN
1	J	88	ASN
1	L	88	ASN
1	E	25	GLN
1	D	25	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 oligosaccharides 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](#)

Warning: The R factor obtained from EDS is 0.2587, which does not match the depositor's R factor of 0.19677. Please interpret the results in this section carefully.

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	103/112 (91%)	0.66	10 (9%) 15 16	32, 41, 54, 61	0
1	B	102/112 (91%)	0.43	5 (4%) 36 32	31, 41, 53, 61	0
1	C	103/112 (91%)	-0.37	0 100 100	31, 41, 54, 61	0
1	D	102/112 (91%)	-0.18	1 (0%) 79 77	29, 41, 53, 62	0
1	E	102/112 (91%)	0.15	2 (1%) 64 63	32, 42, 54, 62	0
1	F	102/112 (91%)	-0.01	1 (0%) 79 77	32, 41, 54, 62	0
1	G	102/112 (91%)	-0.21	0 100 100	33, 41, 53, 61	0
1	H	98/112 (87%)	1.01	9 (9%) 16 17	32, 41, 55, 61	0
1	I	102/112 (91%)	-0.31	1 (0%) 79 77	32, 41, 53, 61	0
1	J	102/112 (91%)	0.38	3 (2%) 54 50	33, 42, 53, 62	0
1	K	102/112 (91%)	0.26	2 (1%) 64 63	33, 41, 53, 61	0
1	L	102/112 (91%)	0.56	5 (4%) 36 32	32, 41, 53, 61	0
All	All	1222/1344 (90%)	0.19	39 (3%) 50 47	29, 41, 54, 62	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	88	ASN	5.0
1	A	88	ASN	3.9
1	J	35	ALA	3.6
1	H	35	ALA	3.4
1	A	27	GLU	3.1

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.