



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 08:01 AM UTC

PDB ID : 2E2G / pdb_00002e2g
Title : Crystal structure of archaeal peroxiredoxin, thioredoxin peroxidase from *Aeropyrum pernix* K1 (pre-oxidation form)
Authors : Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T.
Deposited on : 2006-11-13
Resolution : 2.40 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

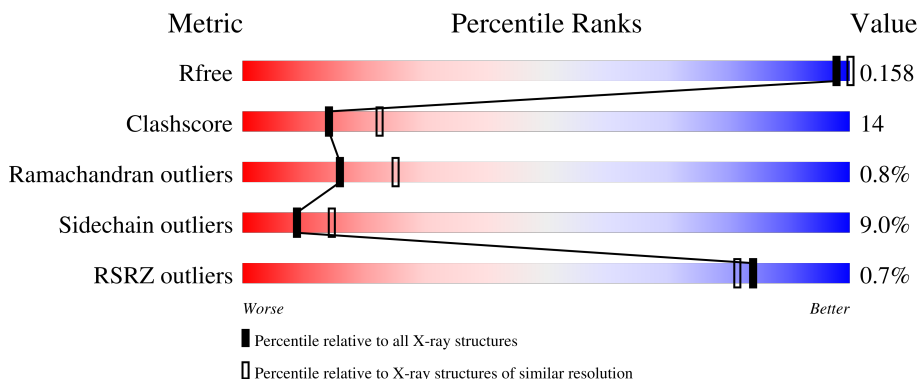
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	250	
1	B	250	
1	C	250	
1	D	250	
1	E	250	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	250	 66% 22% 6% • 5%
1	G	250	%  65% 25% 5% 5%
1	H	250	%  63% 25% 6% • 5%
1	I	250	 64% 26% 5% 5%
1	J	250	%  65% 26% • 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 19878 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 Probable peroxiredoxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	238	1932	1244	339	343	6	0	0	0
1	B	239	1942	1250	342	344	6	0	0	0
1	C	239	1938	1247	340	345	6	0	0	0
1	D	238	1932	1244	339	343	6	0	0	0
1	E	238	1932	1244	339	343	6	0	0	0
1	F	238	1932	1244	339	343	6	0	0	0
1	G	238	1932	1244	339	343	6	0	0	0
1	H	238	1932	1244	339	343	6	0	0	0
1	I	238	1932	1244	339	343	6	0	0	0
1	J	238	1932	1244	339	343	6	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	SER	CYS	engineered mutation	UNP Q9Y9L0
B	207	SER	CYS	engineered mutation	UNP Q9Y9L0
C	207	SER	CYS	engineered mutation	UNP Q9Y9L0
D	207	SER	CYS	engineered mutation	UNP Q9Y9L0
E	207	SER	CYS	engineered mutation	UNP Q9Y9L0
F	207	SER	CYS	engineered mutation	UNP Q9Y9L0
G	207	SER	CYS	engineered mutation	UNP Q9Y9L0
H	207	SER	CYS	engineered mutation	UNP Q9Y9L0
I	207	SER	CYS	engineered mutation	UNP Q9Y9L0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	207	SER	CYS	engineered mutation	UNP Q9Y9L0

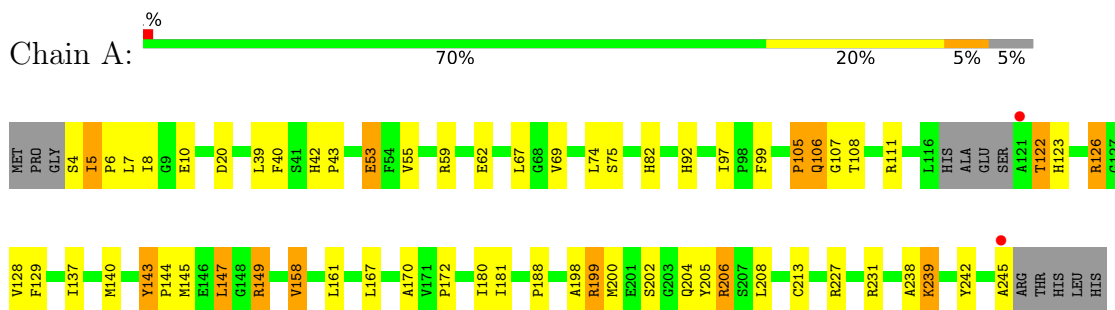
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	57	Total O 57 57	0	0
2	B	60	Total O 60 60	0	0
2	C	39	Total O 39 39	0	0
2	D	57	Total O 57 57	0	0
2	E	65	Total O 65 65	0	0
2	F	59	Total O 59 59	0	0
2	G	61	Total O 61 61	0	0
2	H	65	Total O 65 65	0	0
2	I	30	Total O 30 30	0	0
2	J	49	Total O 49 49	0	0

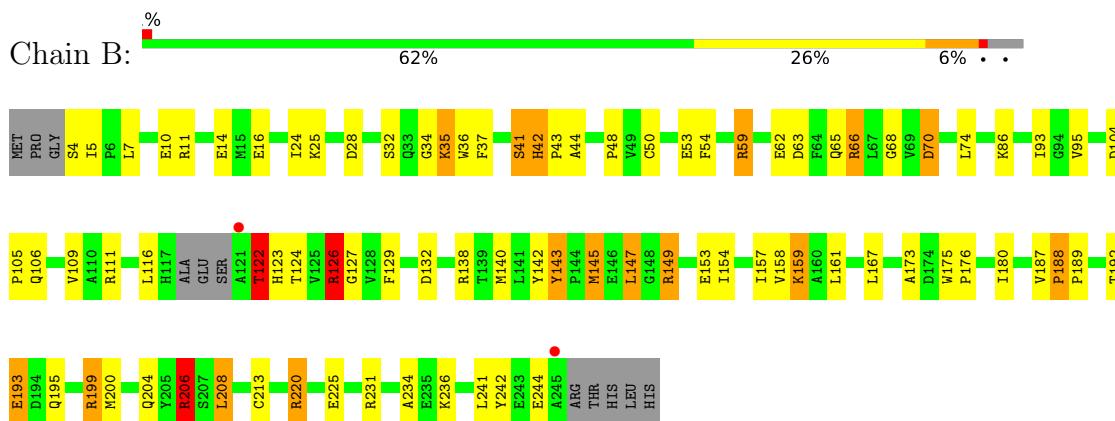
3 Residue-property plots [i](#)

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.

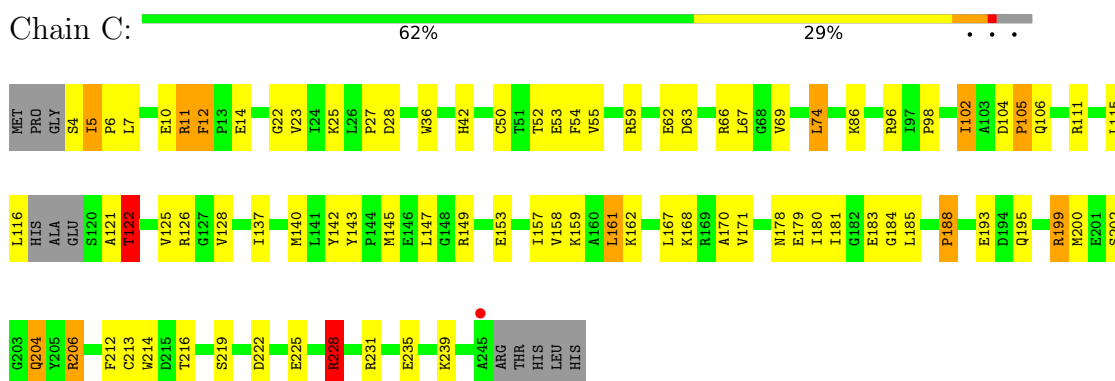
- Molecule 1: Probable peroxiredoxin



- Molecule 1: Probable peroxiredoxin

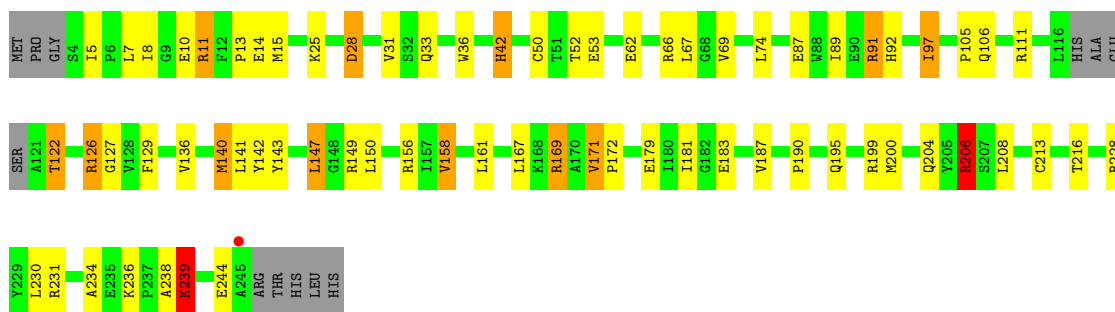


- Molecule 1: Probable peroxiredoxin



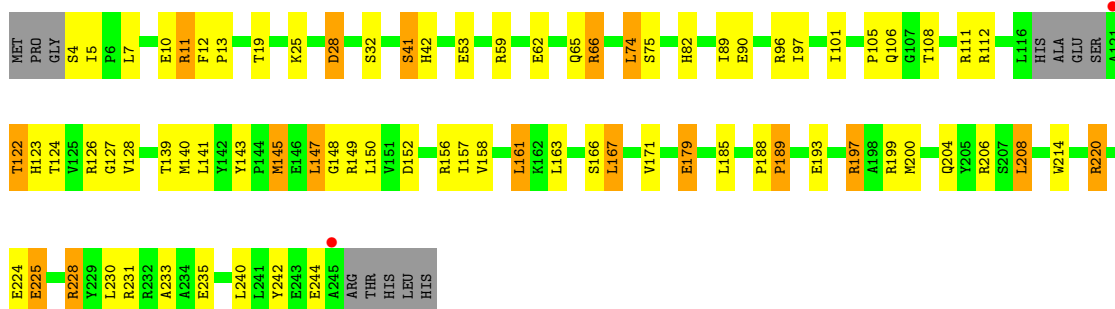
- Molecule 1: Probable peroxiredoxin

Chain D:  67% 22% 5% • 5%



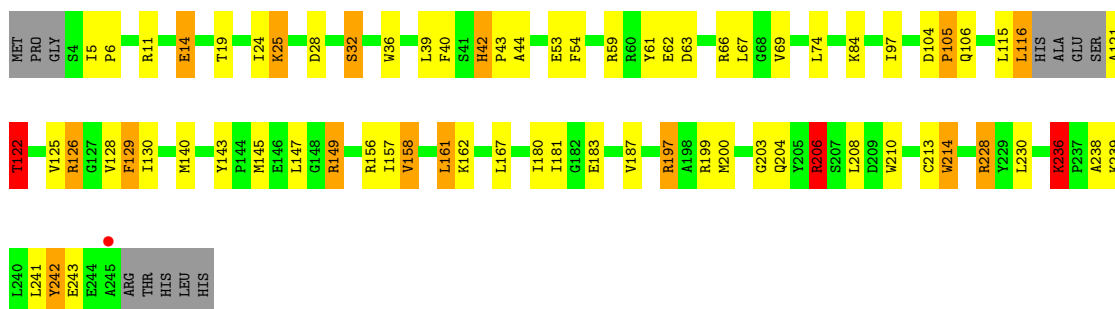
- Molecule 1: Probable peroxiredoxin

Chain E:  64% 24% 7% 5%



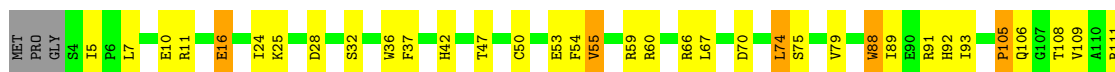
- Molecule 1: Probable peroxiredoxin

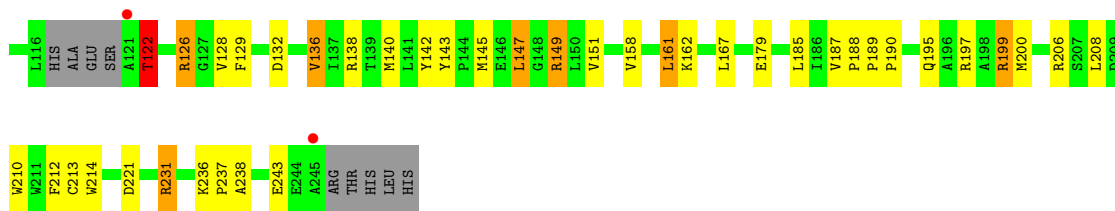
Chain F:  66% 22% 6% • 5%



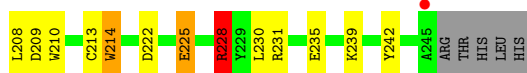
- Molecule 1: Probable peroxiredoxin

Chain G:  65% 25% 5% 5%

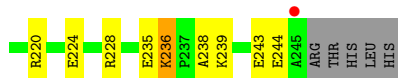




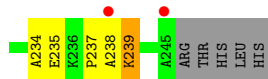
• Molecule 1: Probable peroxiredoxin



• Molecule 1: Probable peroxiredoxin



• Molecule 1: Probable peroxiredoxin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	76.20Å 103.35Å 104.63Å 105.79° 105.19° 92.68°	Depositor
Resolution (Å)	19.99 – 2.40 19.99 – 2.40	Depositor EDS
% Data completeness (in resolution range)	86.8 (19.99-2.40) 87.2 (19.99-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.84 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.153 , 0.229 0.160 , 0.158	Depositor DCC
R_{free} test set	5083 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	23.5	Xtrriage
Anisotropy	0.222	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 45.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	19878	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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	1.32	1/1984 (0.1%)	1.35	10/2696 (0.4%)
1	B	1.36	7/1995 (0.4%)	1.36	19/2711 (0.7%)
1	C	1.37	5/1990 (0.3%)	1.38	14/2704 (0.5%)
1	D	1.33	5/1984 (0.3%)	1.37	18/2696 (0.7%)
1	E	1.41	11/1984 (0.6%)	1.36	11/2696 (0.4%)
1	F	1.46	7/1984 (0.4%)	1.42	17/2696 (0.6%)
1	G	1.38	5/1984 (0.3%)	1.35	12/2696 (0.4%)
1	H	1.41	12/1984 (0.6%)	1.33	7/2696 (0.3%)
1	I	1.33	6/1984 (0.3%)	1.34	14/2696 (0.5%)
1	J	1.32	4/1984 (0.2%)	1.31	10/2696 (0.4%)
All	All	1.37	63/19857 (0.3%)	1.36	132/26983 (0.5%)

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	189	PRO	CA-C	10.86	1.57	1.51
1	B	189	PRO	CA-C	10.43	1.57	1.51
1	E	171	VAL	CA-CB	7.79	1.61	1.54
1	C	102	ILE	CA-CB	-7.58	1.45	1.54
1	I	189	PRO	CA-C	7.53	1.56	1.51

The worst 5 of 132 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	42	HIS	CA-C-N	9.41	129.63	119.28
1	F	42	HIS	C-N-CA	9.41	129.63	119.28
1	C	204	GLN	N-CA-C	7.89	122.94	113.16
1	B	126	ARG	CA-C-N	-7.53	106.65	121.41
1	B	126	ARG	C-N-CA	-7.53	106.65	121.41

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	1932	0	1924	60	0
1	B	1942	0	1931	74	0
1	C	1938	0	1929	64	0
1	D	1932	0	1924	51	0
1	E	1932	0	1924	64	0
1	F	1932	0	1924	56	1
1	G	1932	0	1924	67	0
1	H	1932	0	1924	64	1
1	I	1932	0	1924	57	0
1	J	1932	0	1924	47	0
2	A	57	0	0	0	0
2	B	60	0	0	2	0
2	C	39	0	0	4	0
2	D	57	0	0	4	0
2	E	65	0	0	9	0
2	F	59	0	0	6	0
2	G	61	0	0	5	0
2	H	65	0	0	6	0
2	I	30	0	0	2	0
2	J	49	0	0	1	0
All	All	19878	0	19252	549	1

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 14.

The worst 5 of 549 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:206:ARG:HG3	1:F:206:ARG:HH11	1.10	1.15
1:H:200:MET:HE3	1:H:210:TRP:HA	1.26	1.12
1:A:206:ARG:HG3	1:A:206:ARG:HH11	1.18	1.07
1:F:129:PHE:CE2	1:F:140:MET:HE2	1.90	1.07
1:I:129:PHE:CE2	1:I:140:MET:HE3	1.91	1.06

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:228:ARG:NH1	1:H:228:ARG:CB[1_544]	2.19	0.01

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	234/250 (94%)	224 (96%)	7 (3%)	3 (1%)	9	14
1	B	235/250 (94%)	228 (97%)	6 (3%)	1 (0%)	30	43
1	C	235/250 (94%)	219 (93%)	15 (6%)	1 (0%)	30	43
1	D	234/250 (94%)	219 (94%)	13 (6%)	2 (1%)	14	22
1	E	234/250 (94%)	223 (95%)	11 (5%)	0	100	100
1	F	234/250 (94%)	224 (96%)	8 (3%)	2 (1%)	14	22
1	G	234/250 (94%)	217 (93%)	15 (6%)	2 (1%)	14	22
1	H	234/250 (94%)	221 (94%)	10 (4%)	3 (1%)	9	14
1	I	234/250 (94%)	221 (94%)	10 (4%)	3 (1%)	9	14
1	J	234/250 (94%)	211 (90%)	21 (9%)	2 (1%)	14	22
All	All	2342/2500 (94%)	2207 (94%)	116 (5%)	19 (1%)	16	25

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	THR
1	C	121	ALA
1	F	32	SER
1	F	122	THR
1	I	122	THR

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	206/216 (95%)	190 (92%)	16 (8%)	11	20
1	B	207/216 (96%)	184 (89%)	23 (11%)	6	9
1	C	207/216 (96%)	191 (92%)	16 (8%)	12	20
1	D	206/216 (95%)	188 (91%)	18 (9%)	9	16
1	E	206/216 (95%)	185 (90%)	21 (10%)	7	11
1	F	206/216 (95%)	187 (91%)	19 (9%)	8	14
1	G	206/216 (95%)	192 (93%)	14 (7%)	14	25
1	H	206/216 (95%)	185 (90%)	21 (10%)	7	11
1	I	206/216 (95%)	190 (92%)	16 (8%)	11	20
1	J	206/216 (95%)	184 (89%)	22 (11%)	6	10
All	All	2062/2160 (96%)	1876 (91%)	186 (9%)	9	15

5 of 186 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	32	SER
1	H	199	ARG
1	G	108	THR
1	H	65	GLN
1	I	24	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	106	GLN
1	I	195	GLN
1	E	106	GLN
1	F	106	GLN
1	F	204	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](#)

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	238/250 (95%)	-0.78	2 (0%) 82 80	12, 22, 45, 72	0
1	B	239/250 (95%)	-0.68	2 (0%) 82 80	14, 24, 44, 62	0
1	C	239/250 (95%)	-0.68	1 (0%) 88 86	15, 25, 44, 71	0
1	D	238/250 (95%)	-0.68	1 (0%) 88 86	13, 25, 46, 76	0
1	E	238/250 (95%)	-0.65	2 (0%) 82 80	13, 25, 44, 65	0
1	F	238/250 (95%)	-0.71	1 (0%) 88 86	11, 22, 41, 65	0
1	G	238/250 (95%)	-0.67	2 (0%) 82 80	14, 24, 44, 59	0
1	H	238/250 (95%)	-0.65	2 (0%) 82 80	13, 24, 45, 63	0
1	I	238/250 (95%)	-0.56	1 (0%) 88 86	18, 30, 49, 68	0
1	J	238/250 (95%)	-0.62	3 (1%) 75 71	15, 27, 48, 63	0
All	All	2382/2500 (95%)	-0.67	17 (0%) 84 81	11, 25, 46, 76	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	245	ALA	3.5
1	F	245	ALA	3.3
1	H	121	ALA	3.2
1	H	245	ALA	3.1
1	J	238	ALA	3.0

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 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.