



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 30, 2024 – 06:37 PM EST

PDB ID : 1HR6
Title : Yeast Mitochondrial Processing Peptidase
Authors : Taylor, A.B.; Smith, B.S.; Kitada, S.; Kojima, K.; Miyaura, H.; Otwinowski, Z.; Ito, A.; Deisenhofer, J.
Deposited on : 2000-12-21
Resolution : 2.50 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

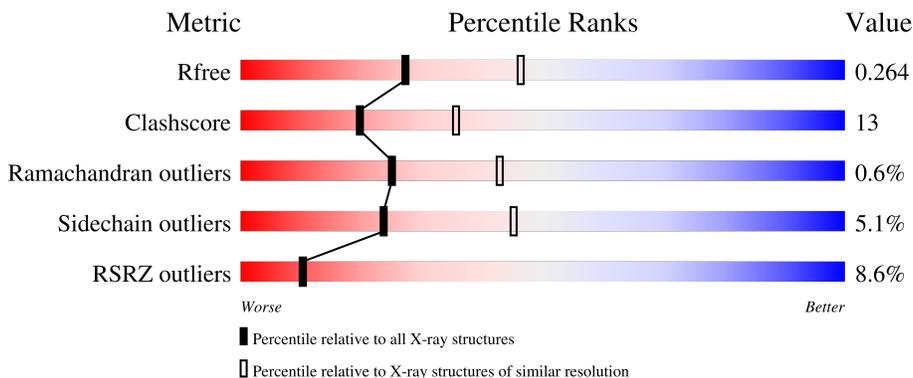
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	475	 6% 72% 22% . .
1	C	475	 4% 72% 21% . 5%
1	E	475	 6% 73% 21% . .
1	G	475	 6% 70% 23% . .
2	B	443	 4% 72% 25% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	443	
2	F	443	
2	H	443	

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 28061 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 MITOCHONDRIAL PROCESSING PEPTIDASE ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	457	3531	2233	599	680	19	0	0	0
1	C	453	3503	2218	592	674	19	0	0	0
1	E	455	3519	2227	597	676	19	0	0	0
1	G	457	3531	2233	599	680	19	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	177	GLY	GLU	SEE REMARK 999	UNP P11914
A	217	GLY	GLU	SEE REMARK 999	UNP P11914
A	483	HIS	-	expression tag	UNP P11914
A	484	HIS	-	expression tag	UNP P11914
A	485	HIS	-	expression tag	UNP P11914
A	486	HIS	-	expression tag	UNP P11914
A	487	HIS	-	expression tag	UNP P11914
A	488	HIS	-	expression tag	UNP P11914
C	177	GLY	GLU	SEE REMARK 999	UNP P11914
C	217	GLY	GLU	SEE REMARK 999	UNP P11914
C	483	HIS	-	expression tag	UNP P11914
C	484	HIS	-	expression tag	UNP P11914
C	485	HIS	-	expression tag	UNP P11914
C	486	HIS	-	expression tag	UNP P11914
C	487	HIS	-	expression tag	UNP P11914
C	488	HIS	-	expression tag	UNP P11914
E	177	GLY	GLU	SEE REMARK 999	UNP P11914
E	217	GLY	GLU	SEE REMARK 999	UNP P11914
E	483	HIS	-	expression tag	UNP P11914
E	484	HIS	-	expression tag	UNP P11914

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	485	HIS	-	expression tag	UNP P11914
E	486	HIS	-	expression tag	UNP P11914
E	487	HIS	-	expression tag	UNP P11914
E	488	HIS	-	expression tag	UNP P11914
G	177	GLY	GLU	SEE REMARK 999	UNP P11914
G	217	GLY	GLU	SEE REMARK 999	UNP P11914
G	483	HIS	-	expression tag	UNP P11914
G	484	HIS	-	expression tag	UNP P11914
G	485	HIS	-	expression tag	UNP P11914
G	486	HIS	-	expression tag	UNP P11914
G	487	HIS	-	expression tag	UNP P11914
G	488	HIS	-	expression tag	UNP P11914

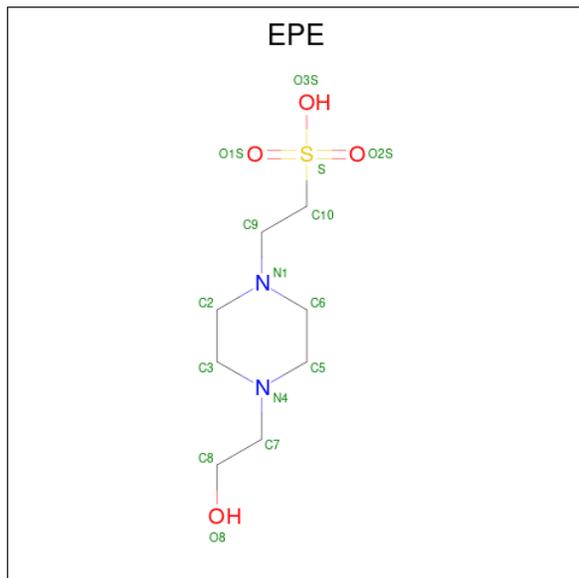
- Molecule 2 is a protein called MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	439	Total 3414	C 2148	N 590	O 669	S 7	0	0	0
2	D	443	Total 3442	C 2165	N 595	O 675	S 7	0	0	0
2	F	443	Total 3442	C 2165	N 595	O 675	S 7	0	0	0
2	H	440	Total 3422	C 2154	N 591	O 670	S 7	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	ALA	-	cloning artifact	UNP P10507
B	84	PRO	SER	SEE REMARK 999	UNP P10507
B	350	ARG	GLN	SEE REMARK 999	UNP P10507
D	20	ALA	-	cloning artifact	UNP P10507
D	84	PRO	SER	SEE REMARK 999	UNP P10507
D	350	ARG	GLN	SEE REMARK 999	UNP P10507
F	20	ALA	-	cloning artifact	UNP P10507
F	84	PRO	SER	SEE REMARK 999	UNP P10507
F	350	ARG	GLN	SEE REMARK 999	UNP P10507
H	20	ALA	-	cloning artifact	UNP P10507
H	84	PRO	SER	SEE REMARK 999	UNP P10507
H	350	ARG	GLN	SEE REMARK 999	UNP P10507

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	15	8	2	4	1	0	0
3	G	1	15	8	2	4	1	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	B	1	1	1	0	0
4	D	1	1	1	0	0
4	F	1	1	1	0	0
4	H	1	1	1	0	0

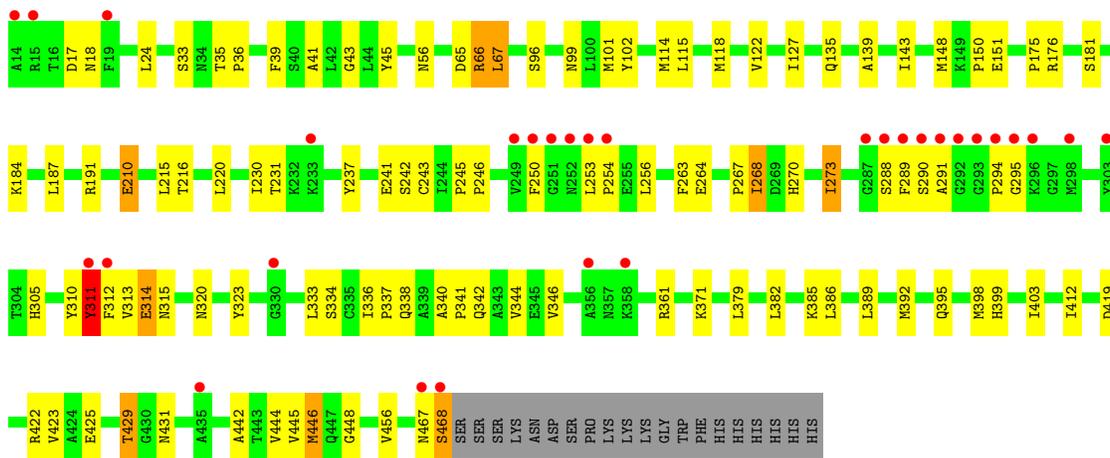
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	47	47	47	0	0
5	B	54	54	54	0	0

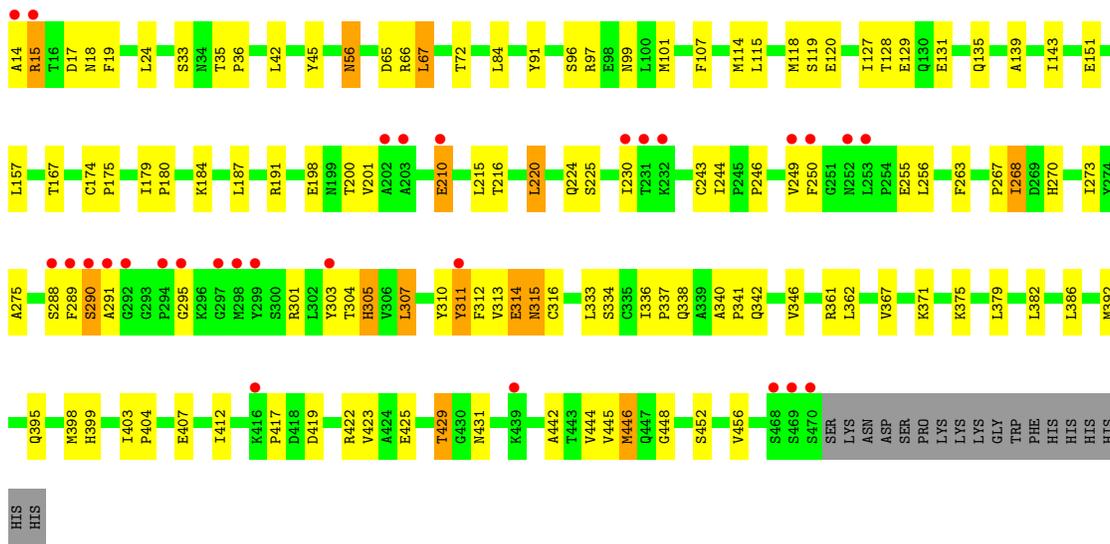
Continued on next page...

Continued from previous page...

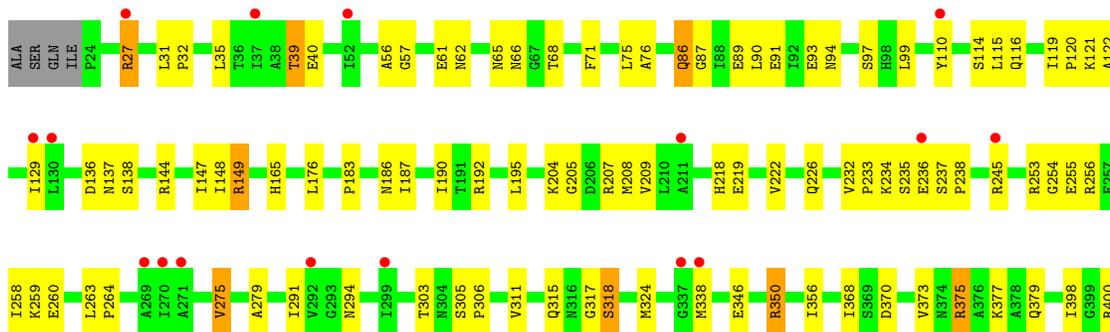
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	28	Total O 28 28	0	0
5	D	21	Total O 21 21	0	0
5	E	35	Total O 35 35	0	0
5	F	10	Total O 10 10	0	0
5	G	25	Total O 25 25	0	0
5	H	3	Total O 3 3	0	0



• Molecule 1: MITOCHONDRIAL PROCESSING PEPTIDASE ALPHA SUBUNIT

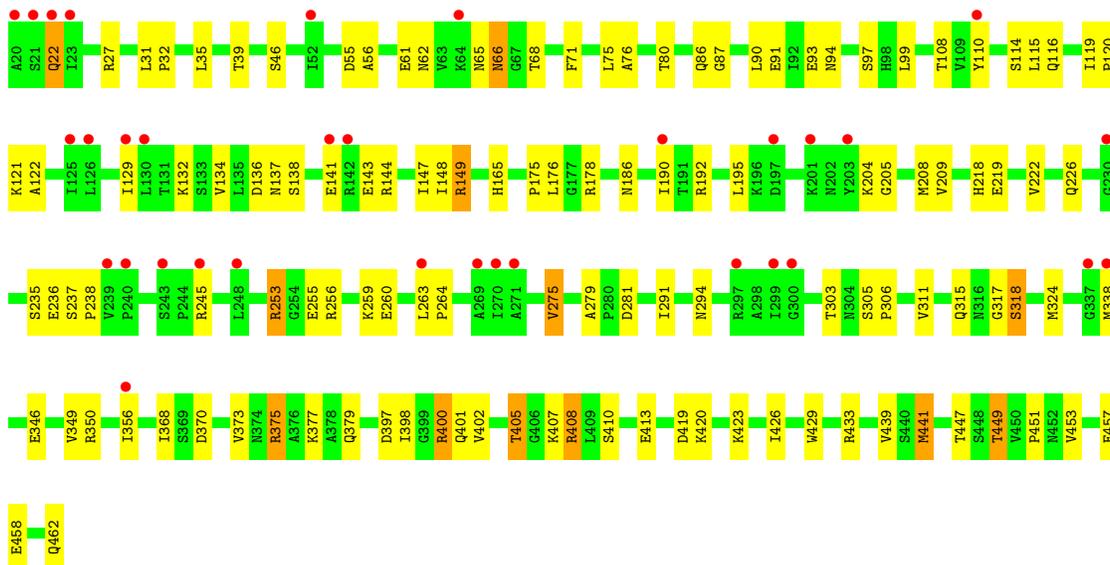


• Molecule 2: MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT

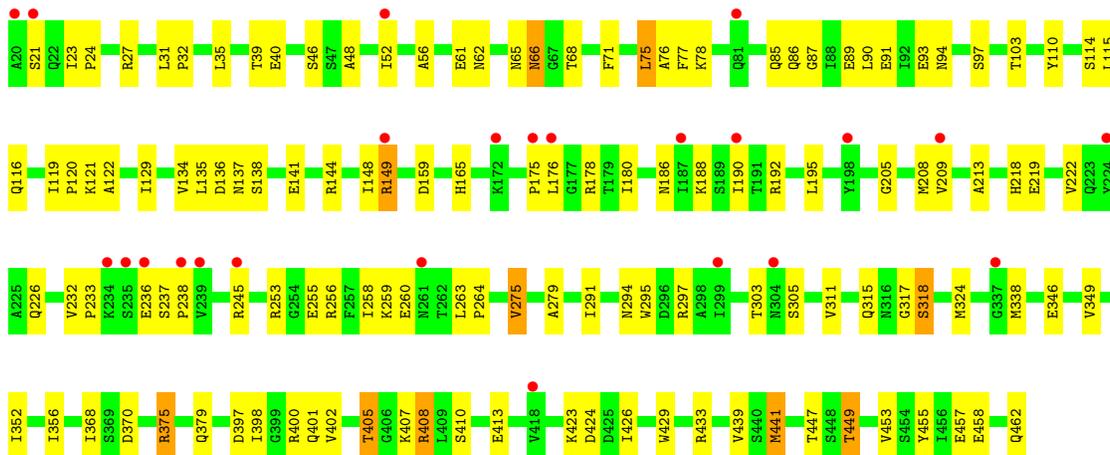




● Molecule 2: MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT

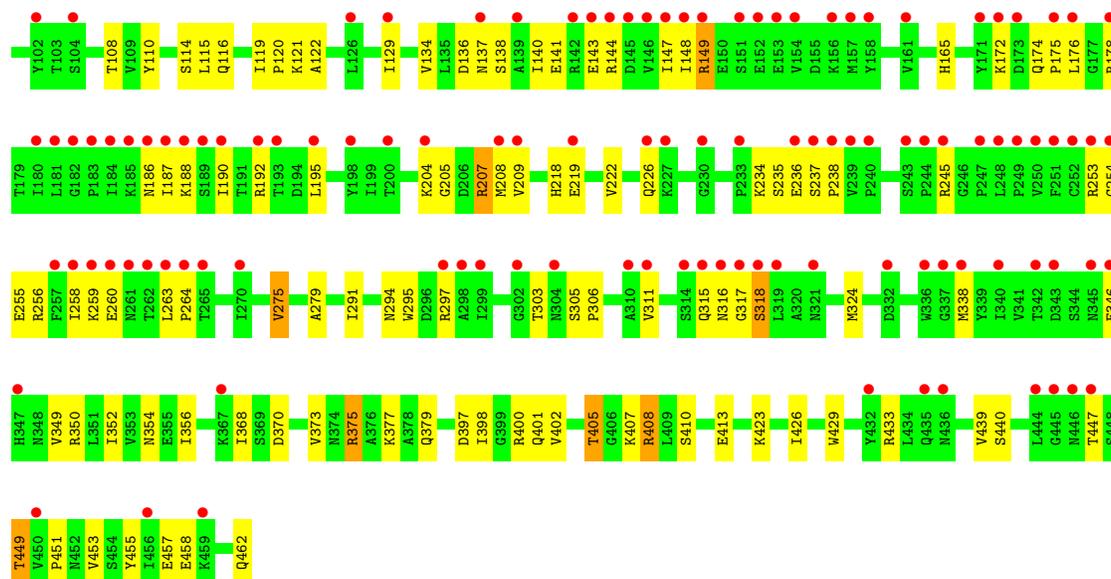


● Molecule 2: MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT



● Molecule 2: MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	134.18Å 178.50Å 202.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.96 – 2.50 29.96 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.4 (29.96-2.50) 99.1 (29.96-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.68 (at 2.51Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.245 , 0.279 0.234 , 0.264	Depositor DCC
R_{free} test set	2029 reflections (1.22%)	wwPDB-VP
Wilson B-factor (Å ²)	53.1	Xtrriage
Anisotropy	0.297	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 44.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	28061	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.32% 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](#)

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

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	4/3604 (0.1%)	0.77	3/4877 (0.1%)
1	C	0.66	1/3576 (0.0%)	0.76	1/4840 (0.0%)
1	E	0.63	2/3592 (0.1%)	0.83	4/4861 (0.1%)
1	G	0.62	0/3604	0.76	3/4877 (0.1%)
2	B	0.62	0/3478	0.80	5/4720 (0.1%)
2	D	0.57	0/3506	0.84	6/4759 (0.1%)
2	F	0.53	0/3506	0.75	4/4759 (0.1%)
2	H	0.56	2/3486 (0.1%)	0.87	7/4732 (0.1%)
All	All	0.61	9/28352 (0.0%)	0.80	33/38425 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	467	ASN	CG-OD1	-7.06	1.08	1.24
1	E	315	ASN	CG-OD1	-7.03	1.08	1.24
1	A	90	ASN	CG-ND2	-6.81	1.15	1.32
1	A	90	ASN	CG-OD1	-6.73	1.09	1.24
1	A	467	ASN	CG-ND2	-6.59	1.16	1.32

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	149	ARG	NE-CZ-NH2	-17.06	111.77	120.30
2	D	253	ARG	NE-CZ-NH1	-16.95	111.82	120.30
2	D	253	ARG	NE-CZ-NH2	16.31	128.45	120.30
1	E	422	ARG	NE-CZ-NH1	-16.20	112.20	120.30
2	H	149	ARG	NE-CZ-NH1	16.11	128.35	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	3531	0	3488	85	0
1	C	3503	0	3460	84	0
1	E	3519	0	3478	70	0
1	G	3531	0	3488	97	0
2	B	3414	0	3412	97	0
2	D	3442	0	3440	92	0
2	F	3442	0	3440	90	0
2	H	3422	0	3422	113	0
3	A	15	0	17	0	0
3	G	15	0	17	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	H	1	0	0	0	0
5	A	47	0	0	3	0
5	B	54	0	0	2	0
5	C	28	0	0	1	0
5	D	21	0	0	0	0
5	E	35	0	0	3	0
5	F	10	0	0	0	0
5	G	25	0	0	1	0
5	H	3	0	0	0	0
All	All	28061	0	27662	703	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 13.

The worst 5 of 703 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:33:SER:HB3	1:E:392:MET:HE1	1.28	1.16
2:H:294:ASN:HD22	2:H:324:MET:HA	1.11	1.11
2:B:294:ASN:HD22	2:B:324:MET:HA	1.14	1.07
1:A:33:SER:HB3	1:A:392:MET:HE1	1.42	0.99
1:C:315:ASN:HD22	1:C:316:CYS:N	1.61	0.99

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	455/475 (96%)	436 (96%)	17 (4%)	2 (0%)	34	54
1	C	451/475 (95%)	433 (96%)	16 (4%)	2 (0%)	34	54
1	E	453/475 (95%)	435 (96%)	16 (4%)	2 (0%)	34	54
1	G	455/475 (96%)	434 (95%)	17 (4%)	4 (1%)	17	31
2	B	437/443 (99%)	423 (97%)	11 (2%)	3 (1%)	22	39
2	D	441/443 (100%)	426 (97%)	12 (3%)	3 (1%)	22	39
2	F	441/443 (100%)	425 (96%)	13 (3%)	3 (1%)	22	39
2	H	438/443 (99%)	423 (97%)	11 (2%)	4 (1%)	17	31
All	All	3571/3672 (97%)	3435 (96%)	113 (3%)	23 (1%)	25	43

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	250	PHE
2	B	57	GLY
1	C	250	PHE
1	C	290	SER
2	D	318	SER

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	384/401 (96%)	364 (95%)	20 (5%)	23	44
1	C	381/401 (95%)	360 (94%)	21 (6%)	21	41
1	E	382/401 (95%)	360 (94%)	22 (6%)	20	38
1	G	384/401 (96%)	361 (94%)	23 (6%)	19	37
2	B	376/379 (99%)	359 (96%)	17 (4%)	27	51
2	D	379/379 (100%)	362 (96%)	17 (4%)	27	51
2	F	379/379 (100%)	361 (95%)	18 (5%)	26	49
2	H	377/379 (100%)	360 (96%)	17 (4%)	27	51
All	All	3042/3120 (98%)	2887 (95%)	155 (5%)	24	45

5 of 155 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	15	ARG
2	H	75	LEU
1	G	67	LEU
1	G	315	ASN
2	H	375	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	223	GLN
2	H	186	ASN
1	E	406	ASN
2	H	165	HIS
2	H	374	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	EPE	G	489	-	15,15,15	1.40	2 (13%)	18,20,20	1.02	0
3	EPE	A	489	-	15,15,15	1.27	2 (13%)	18,20,20	1.03	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EPE	G	489	-	-	2/9/19/19	0/1/1/1
3	EPE	A	489	-	-	2/9/19/19	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	489	EPE	C10-S	2.97	1.81	1.77
3	A	489	EPE	C10-S	2.50	1.81	1.77
3	G	489	EPE	C6-N1	2.36	1.53	1.46
3	A	489	EPE	C6-N1	2.01	1.52	1.46

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	489	EPE	C7-N4-C5	-2.08	105.91	111.23

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	489	EPE	C10-C9-N1-C2
3	G	489	EPE	C10-C9-N1-C6
3	A	489	EPE	C10-C9-N1-C2
3	A	489	EPE	C10-C9-N1-C6

There are no ring outliers.

No monomer is involved in short contacts.

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	457/475 (96%)	0.21	28 (6%) 21 22	30, 46, 85, 101	0
1	C	453/475 (95%)	0.15	17 (3%) 40 43	31, 49, 82, 101	0
1	E	455/475 (95%)	0.34	30 (6%) 18 19	33, 53, 85, 101	0
1	G	457/475 (96%)	0.33	29 (6%) 20 21	35, 54, 89, 101	0
2	B	439/443 (99%)	0.13	16 (3%) 42 46	28, 49, 75, 96	0
2	D	443/443 (100%)	0.38	33 (7%) 14 15	34, 58, 84, 101	0
2	F	443/443 (100%)	0.47	24 (5%) 25 27	45, 68, 88, 100	0
2	H	440/443 (99%)	1.49	133 (30%) 0 0	54, 78, 93, 100	0
All	All	3587/3672 (97%)	0.43	310 (8%) 10 10	28, 58, 88, 101	0

The worst 5 of 310 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	187	ILE	8.9
2	H	184	ILE	8.7
1	C	289	PHE	8.1
2	F	21	SER	7.9
1	C	290	SER	7.7

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, 95th 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 ²)	Q<0.9
4	ZN	F	503	1/1	0.83	0.06	84,84,84,84	0
3	EPE	G	489	15/15	0.91	0.26	70,78,84,85	0
4	ZN	H	504	1/1	0.94	0.13	101,101,101,101	0
4	ZN	D	502	1/1	0.95	0.06	79,79,79,79	0
3	EPE	A	489	15/15	0.96	0.22	68,73,80,81	0
4	ZN	B	501	1/1	0.98	0.04	62,62,62,62	0

6.5 Other polymers [i](#)

There are no such residues in this entry.