wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 11:14 PM EDT

PDB ID	:	4AH6
Title	:	Human mitochondrial aspartyl-tRNA synthetase
Authors	:	Neuenfeldt, A.; Sissler, M.; Lorber, B.; Florentz, C.; Sauter, C.
Deposited on	:	2012-02-03
Resolution	:	3.70 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 (i) 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 (i)) were used in the production of this report:

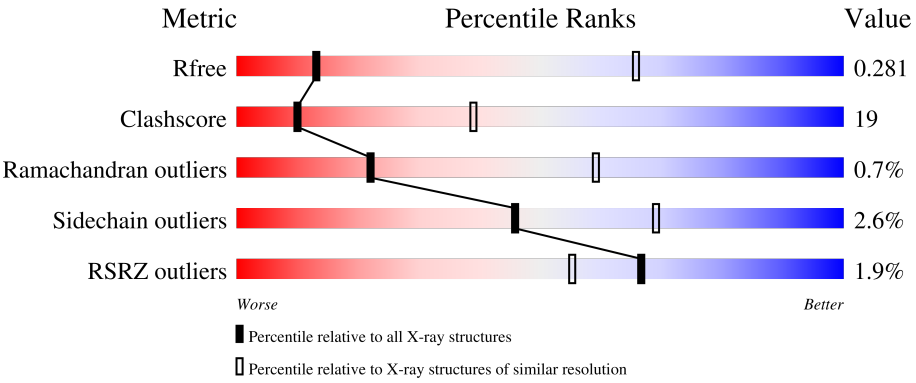
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)

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 <=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	617	2% 59% 35% • 5%
1	B	617	2% 58% 36% • 5%
1	C	617	% 58% 36% • 5%
1	D	617	2% 57% 37% • 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 18840 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 ASPARTATE–TRNA LIGASE, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	589	Total	C	N	O	S	0	0	0
			4710	3007	813	867	23			
1	B	589	Total	C	N	O	S	0	0	0
			4710	3007	813	867	23			
1	C	589	Total	C	N	O	S	0	0	0
			4710	3007	813	867	23			
1	D	589	Total	C	N	O	S	0	0	0
			4710	3007	813	867	23			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	MET	-	expression tag	UNP Q6PI48
A	646	VAL	-	expression tag	UNP Q6PI48
A	647	MET	-	expression tag	UNP Q6PI48
A	648	TYR	-	expression tag	UNP Q6PI48
A	649	LEU	-	expression tag	UNP Q6PI48
A	650	GLU	-	expression tag	UNP Q6PI48
A	651	HIS	-	expression tag	UNP Q6PI48
A	652	HIS	-	expression tag	UNP Q6PI48
A	653	HIS	-	expression tag	UNP Q6PI48
A	654	HIS	-	expression tag	UNP Q6PI48
A	655	HIS	-	expression tag	UNP Q6PI48
A	656	HIS	-	expression tag	UNP Q6PI48
B	40	MET	-	expression tag	UNP Q6PI48
B	646	VAL	-	expression tag	UNP Q6PI48
B	647	MET	-	expression tag	UNP Q6PI48
B	648	TYR	-	expression tag	UNP Q6PI48
B	649	LEU	-	expression tag	UNP Q6PI48
B	650	GLU	-	expression tag	UNP Q6PI48
B	651	HIS	-	expression tag	UNP Q6PI48
B	652	HIS	-	expression tag	UNP Q6PI48
B	653	HIS	-	expression tag	UNP Q6PI48

Continued on next page...

Continued from previous page...

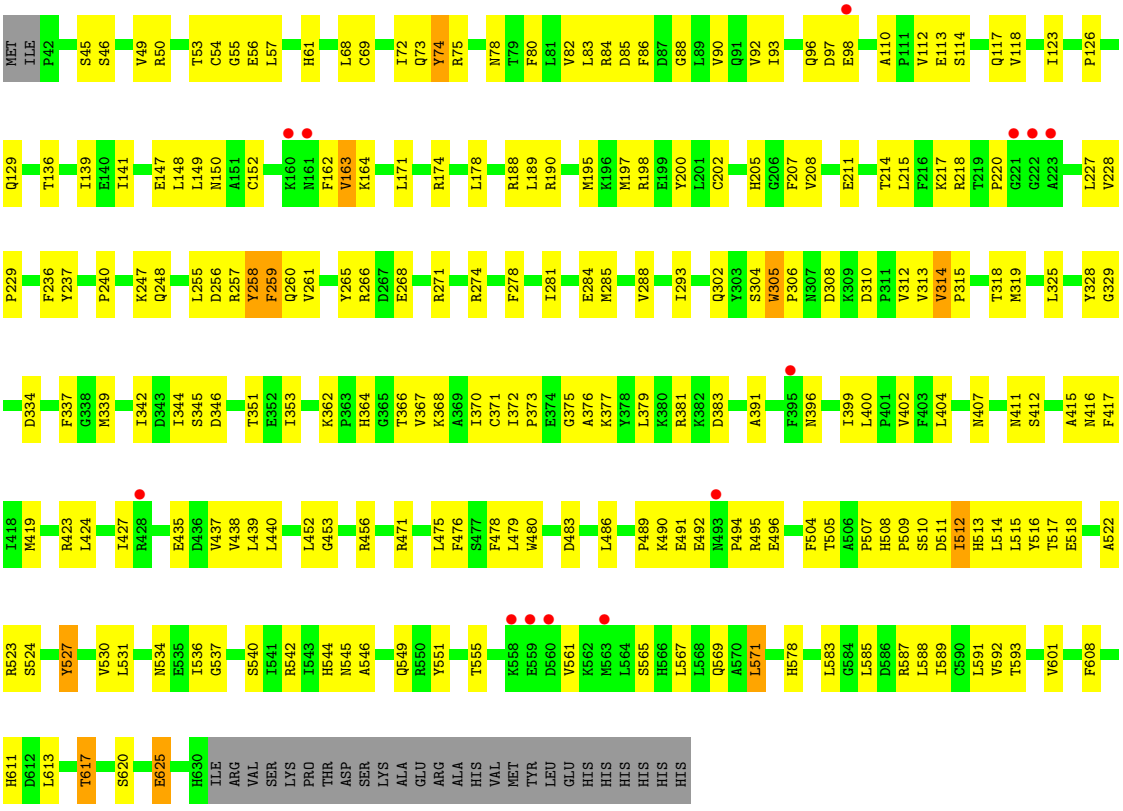
Chain	Residue	Modelled	Actual	Comment	Reference
B	654	HIS	-	expression tag	UNP Q6PI48
B	655	HIS	-	expression tag	UNP Q6PI48
B	656	HIS	-	expression tag	UNP Q6PI48
C	40	MET	-	expression tag	UNP Q6PI48
C	646	VAL	-	expression tag	UNP Q6PI48
C	647	MET	-	expression tag	UNP Q6PI48
C	648	TYR	-	expression tag	UNP Q6PI48
C	649	LEU	-	expression tag	UNP Q6PI48
C	650	GLU	-	expression tag	UNP Q6PI48
C	651	HIS	-	expression tag	UNP Q6PI48
C	652	HIS	-	expression tag	UNP Q6PI48
C	653	HIS	-	expression tag	UNP Q6PI48
C	654	HIS	-	expression tag	UNP Q6PI48
C	655	HIS	-	expression tag	UNP Q6PI48
C	656	HIS	-	expression tag	UNP Q6PI48
D	40	MET	-	expression tag	UNP Q6PI48
D	646	VAL	-	expression tag	UNP Q6PI48
D	647	MET	-	expression tag	UNP Q6PI48
D	648	TYR	-	expression tag	UNP Q6PI48
D	649	LEU	-	expression tag	UNP Q6PI48
D	650	GLU	-	expression tag	UNP Q6PI48
D	651	HIS	-	expression tag	UNP Q6PI48
D	652	HIS	-	expression tag	UNP Q6PI48
D	653	HIS	-	expression tag	UNP Q6PI48
D	654	HIS	-	expression tag	UNP Q6PI48
D	655	HIS	-	expression tag	UNP Q6PI48
D	656	HIS	-	expression tag	UNP Q6PI48

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.

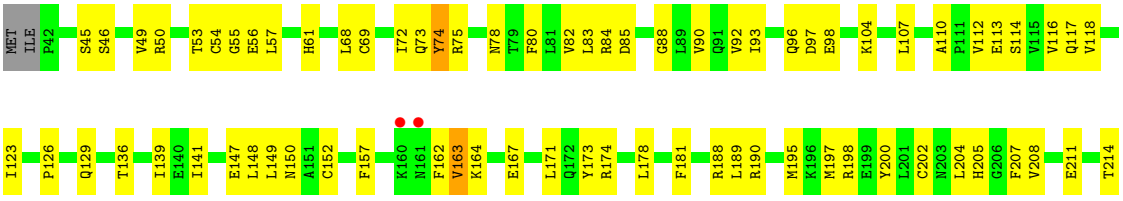
• Molecule 1: ASPARTATE–TRNA LIGASE, MITOCHONDRIAL

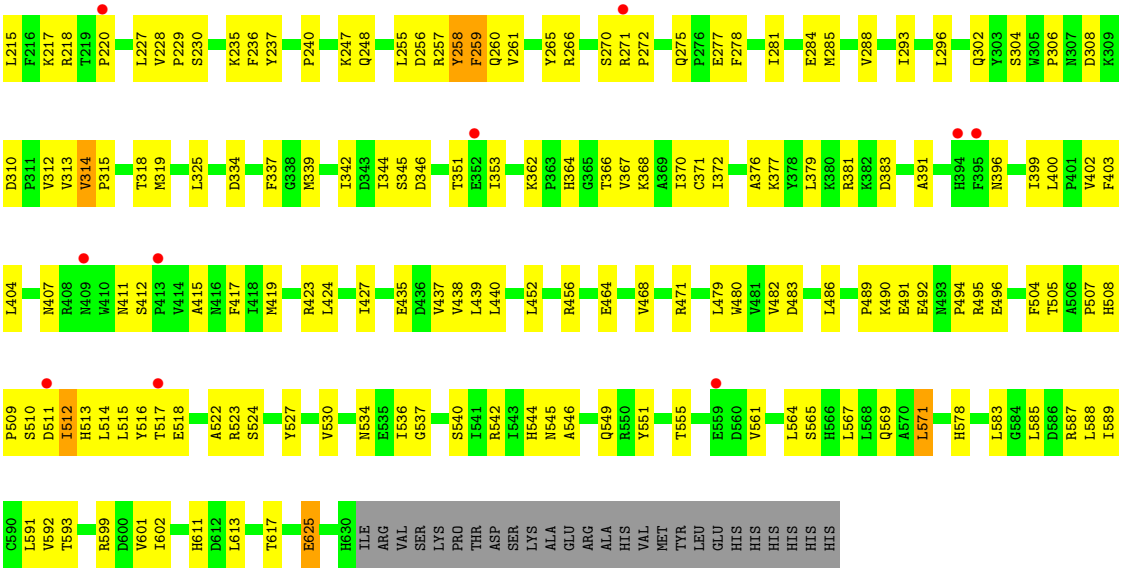
Chain A:



• Molecule 1: ASPARTATE–TRNA LIGASE, MITOCHONDRIAL

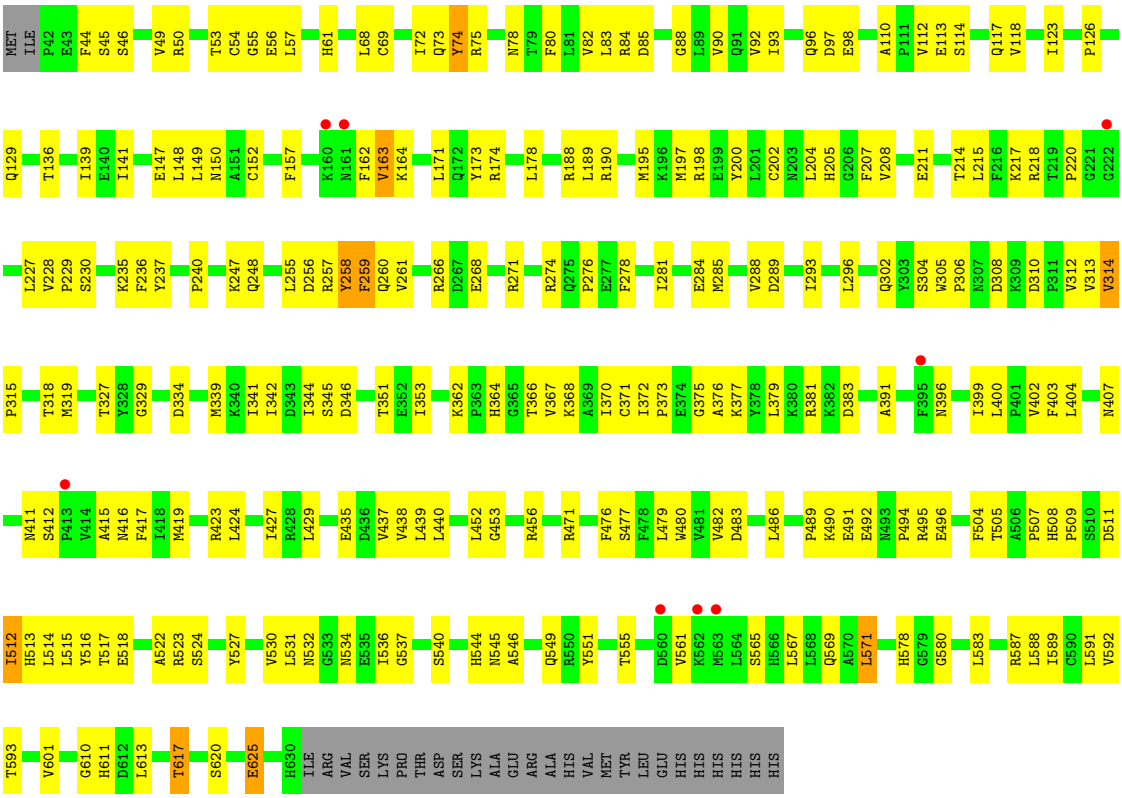
Chain B:





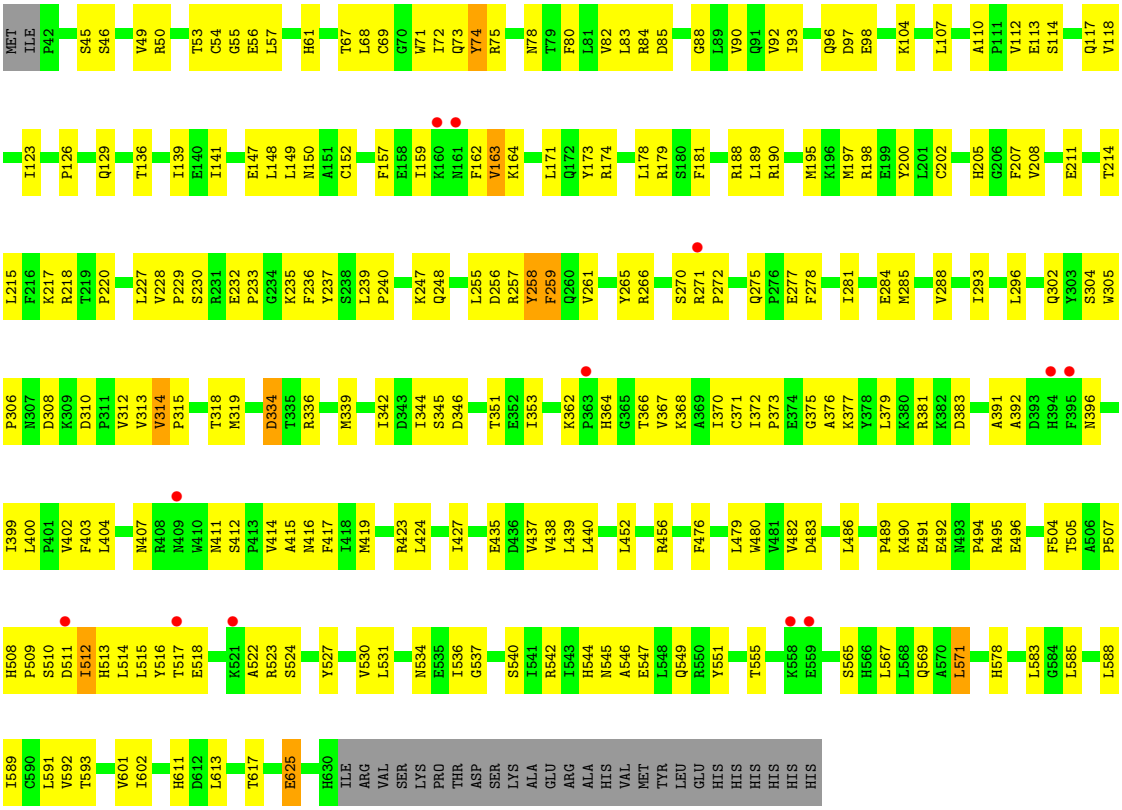
• Molecule 1: ASPARTATE–TRNA LIGASE, MITOCHONDRIAL

Chain C:



• Molecule 1: ASPARTATE–TRNA LIGASE, MITOCHONDRIAL

Chain D:


MET ILE P42 S45 S46 V49 R50 T53 C54 G55 E56 L57 H61 T67 L68 C69 G70 W71 I72 Q73 Y74 R75 N78 T79 F80 L81 V82 L83 R84 D85 G88 L89 V90 Q91 V92 I93 Q96 D97 E98 K104 L107 A110 P111 V112 E113 S114 Q117 V118
I123 P126 Q129 T136 I139 E140 I141 E147 L148 L149 N150 A151 C152 F157 E158 I159 K160 N161 F162 V163 K164 L171 Q172 Y173 R174 L178 R179 S180 F181 R188 L189 R190 M195 K196 M197 R198 E199 Y200 L201 C202 H205 G206 F207 V208 E211 T214
L215 F216 K217 R218 T219 P220 L227 V228 P229 S230 R231 E232 P233 G234 K235 F236 Y237 S238 L239 P240 K247 Q248 L255 D256 R257 Y258 F259 Q260 V261 Y265 R266 S270 R271 P272 Q275 P276 E277 F278 I281 E284 M285 V288 I293 L296 Q302 Y303 S304 W305
P306 N307 D308 K309 D310 P311 V312 V313 V314 P315 T318 M319 D334 T335 R336 M339 I342 D343 I344 S345 D346 T351 E352 I353 K362 P363 H364 G365 T366 V367 K368 A369 I370 C371 I372 P373 E374 G375 A376 K377 Y378 L379 K380 R381 K382 D383 A391 A392 D393 H394 F395 N396
I399 L400 P401 V402 F403 L404 N407 R408 N409 W410 N411 S412 P413 V414 A415 N416 F417 I418 M419 R423 L424 I427 E435 D436 V437 V438 L439 L440 L452 R456 F476 L479 W480 V481 V482 D483 L486 P489 K490 E491 E492 N493 P494 R495 E496 F504 T505 A506 P507
H508 P509 S510 D511 I512 H513 L514 L515 Y516 T517 E518 K521 A522 R523 S524 Y527 V530 L531 N534 E535 I536 G537 S540 I541 R542 I543 H544 N545 A546 E547 L548 Q549 R550 Y551 T555 K558 E559 S565 H566 L567 L568 Q569 A570 L571 H578 L583 G584 L585 L588
I589 C590 L591 V592 T593 V601 I602 H611 D612 L613 T617 E625 H630 ILE ARG VAL SER LYS PRO THR ASP SER LYS ALA GLU ARG ALA HIS VAL MET TYR LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	142.40Å 82.60Å 146.30Å 90.00° 100.40° 90.00°	Depositor
Resolution (Å)	29.92 – 3.70 29.92 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.8 (29.92-3.70) 98.8 (29.92-3.70)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 3.75Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.219 , 0.280 0.222 , 0.281	Depositor DCC
R_{free} test set	1782 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	88.0	Xtriage
Anisotropy	0.622	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 73.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.044 for l,-k,h	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	18840	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.29	0/4819	0.52	0/6529
1	B	0.29	0/4819	0.53	0/6529
1	C	0.29	0/4819	0.53	0/6529
1	D	0.29	0/4819	0.53	0/6529
All	All	0.29	0/19276	0.53	0/26116

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

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	4710	0	4729	187	0
1	B	4710	0	4729	205	0
1	C	4710	0	4729	192	1
1	D	4710	0	4729	202	1
All	All	18840	0	18916	716	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 19.

The worst 5 of 716 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:A:195:MET:CE	1:B:198:ARG:HG3	1.46	1.43
1:C:198:ARG:HG3	1:D:195:MET:CE	1.52	1.37
1:A:195:MET:HE1	1:B:198:ARG:CG	1.54	1.36
1:C:198:ARG:CG	1:D:195:MET:HE1	1.74	1.17
1:D:162:PHE:CE1	1:D:181:PHE:HB3	1.91	1.05

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:517:THR:OG1	1:D:547:GLU:OE2[1_545]	2.18	0.02

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	587/617 (95%)	512 (87%)	71 (12%)	4 (1%)	22	59
1	B	587/617 (95%)	510 (87%)	73 (12%)	4 (1%)	22	59
1	C	587/617 (95%)	512 (87%)	71 (12%)	4 (1%)	22	59
1	D	587/617 (95%)	511 (87%)	72 (12%)	4 (1%)	22	59
All	All	2348/2468 (95%)	2045 (87%)	287 (12%)	16 (1%)	22	59

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	512	ILE
1	B	512	ILE
1	C	512	ILE
1	D	512	ILE
1	A	163	VAL

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	525/551 (95%)	509 (97%)	16 (3%)	41	66
1	B	525/551 (95%)	514 (98%)	11 (2%)	53	74
1	C	525/551 (95%)	510 (97%)	15 (3%)	42	66
1	D	525/551 (95%)	512 (98%)	13 (2%)	47	70
All	All	2100/2204 (95%)	2045 (97%)	55 (3%)	46	69

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	80	PHE
1	C	334	ASP
1	D	625	GLU
1	D	334	ASP
1	C	162	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	545	ASN
1	D	422	GLN
1	D	73	GLN
1	D	205	HIS
1	D	545	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)

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

6.1 Protein, DNA and RNA chains

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^{th} 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	589/617 (95%)	-0.21	13 (2%)	62	50	55, 102, 161, 211	0
1	B	589/617 (95%)	-0.16	12 (2%)	65	53	54, 104, 163, 205	0
1	C	589/617 (95%)	-0.14	8 (1%)	75	64	53, 104, 162, 204	0
1	D	589/617 (95%)	-0.22	12 (2%)	65	53	53, 101, 160, 198	0
All	All	2356/2468 (95%)	-0.18	45 (1%)	66	55	53, 103, 162, 211	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	160	LYS	5.6
1	D	409	ASN	4.7
1	B	409	ASN	4.7
1	D	558	LYS	4.2
1	A	222	GLY	3.9

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.