wwPDB X-ray Structure Validation Summary Report (i)

Nov 12, 2024 – 06:58 AM JST

PDB ID	:	8X4Q
Title	:	Apo structure of L-tryptophan specific decarboxylase PsiD
Authors	:	Meng, C.Y.; Wen, Y.; Guo, W.T.; Wu, B.X.
Deposited on	:	2023-11-15
Resolution	:	2.55 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 (1)) were used in the production of this report:

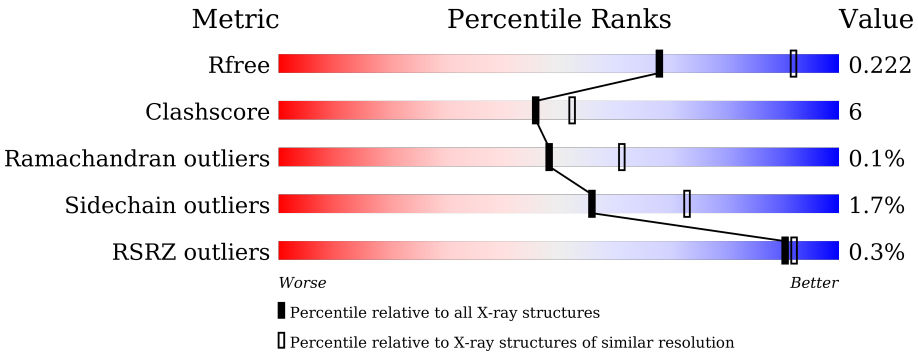
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.55 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	1004 (2.54-2.54)
Clashscore	180529	1055 (2.54-2.54)
Ramachandran outliers	177936	1048 (2.54-2.54)
Sidechain outliers	177891	1048 (2.54-2.54)
RSRZ outliers	164620	1004 (2.54-2.54)

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	402	78% 9% 13%
1	C	402	74% 12% 14%
1	E	402	75% 10% 15%
1	G	402	73% 12% 14%
1	I	402	73% 13% 14%
1	K	402	73% 13% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	37	78% 19%
2	D	37	78% 16% 5%
2	F	37	73% 22% 5%
2	H	37	65% 30% 5%
2	J	37	3% 86% 14%
2	L	37	70% 24% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 19055 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 L-tryptophan decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	350	Total	C	N	O	S	0	1	0
			2808	1801	473	519	15			
1	C	346	Total	C	N	O	S	0	0	0
			2758	1771	461	511	15			
1	E	342	Total	C	N	O	S	0	0	0
			2743	1763	458	507	15			
1	G	346	Total	C	N	O	S	0	0	0
			2767	1777	463	512	15			
1	I	346	Total	C	N	O	S	0	0	0
			2759	1771	462	511	15			
1	K	347	Total	C	N	O	S	0	0	0
			2776	1782	468	511	15			

- Molecule 2 is a protein called L-tryptophan decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	35	Total	C	N	O	S	0	0	0
			263	168	46	48	1			
2	J	37	Total	C	N	O	S	0	0	0
			274	174	48	51	1			
2	H	35	Total	C	N	O	S	0	0	0
			257	164	42	50	1			
2	F	35	Total	C	N	O	S	0	0	0
			267	170	46	50	1			
2	D	35	Total	C	N	O	S	0	0	0
			257	165	43	48	1			
2	B	36	Total	C	N	O	S	0	0	0
			258	166	44	47	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	501	PYR	SER	modified residue	UNP P0DPA6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	501	PYR	SER	modified residue	UNP P0DPA6
H	501	PYR	SER	modified residue	UNP P0DPA6
F	501	PYR	SER	modified residue	UNP P0DPA6
D	501	PYR	SER	modified residue	UNP P0DPA6
B	501	PYR	SER	modified residue	UNP P0DPA6

- Molecule 3 is water.

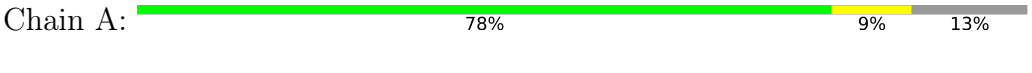
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	158	Total O 158 158	0	0
3	C	139	Total O 139 139	0	0
3	E	146	Total O 146 146	0	0
3	G	143	Total O 143 143	0	0
3	I	93	Total O 93 93	0	0
3	K	130	Total O 130 130	0	0
3	L	6	Total O 6 6	0	0
3	J	12	Total O 12 12	0	0
3	H	10	Total O 10 10	0	0
3	F	12	Total O 12 12	0	0
3	D	10	Total O 10 10	0	0
3	B	9	Total O 9 9	0	0

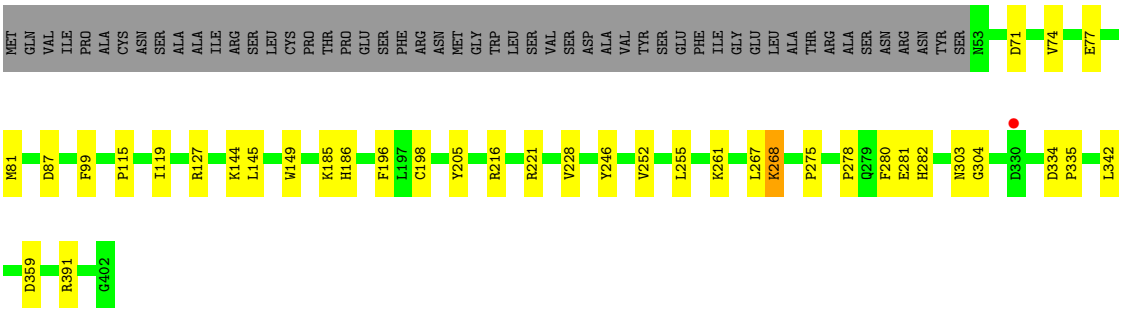
3 Residue-property plots (i)

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: L-tryptophan decarboxylase

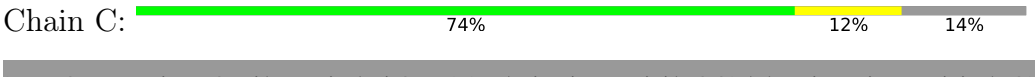
Chain A:

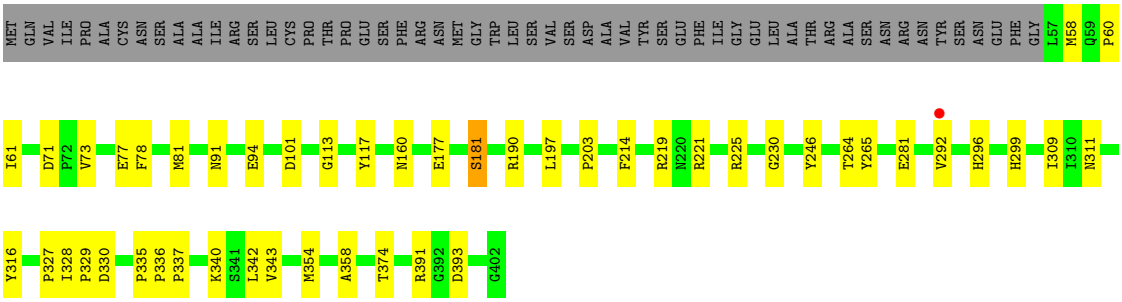




• Molecule 1: L-tryptophan decarboxylase

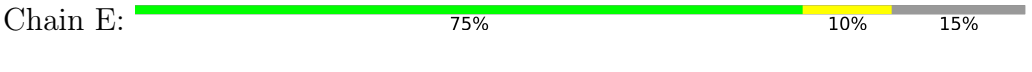
Chain C:

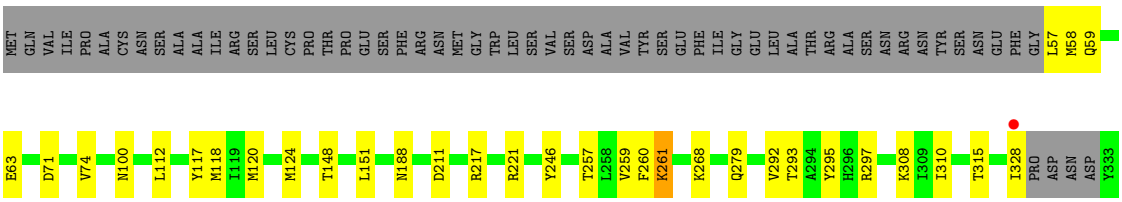


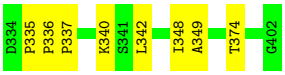


• Molecule 1: L-tryptophan decarboxylase

Chain E:

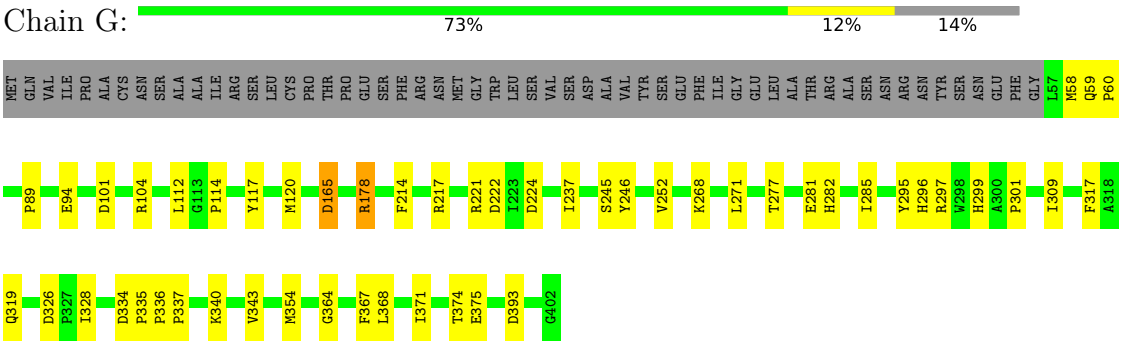






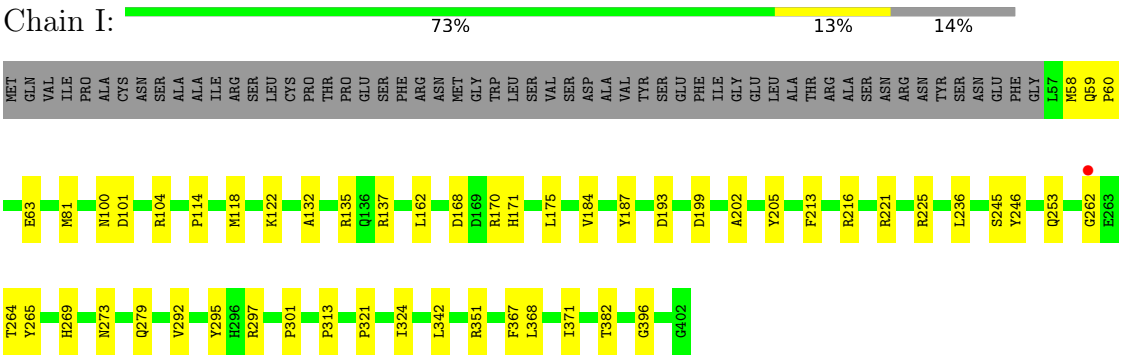
• Molecule 1: L-tryptophan decarboxylase

Chain G:



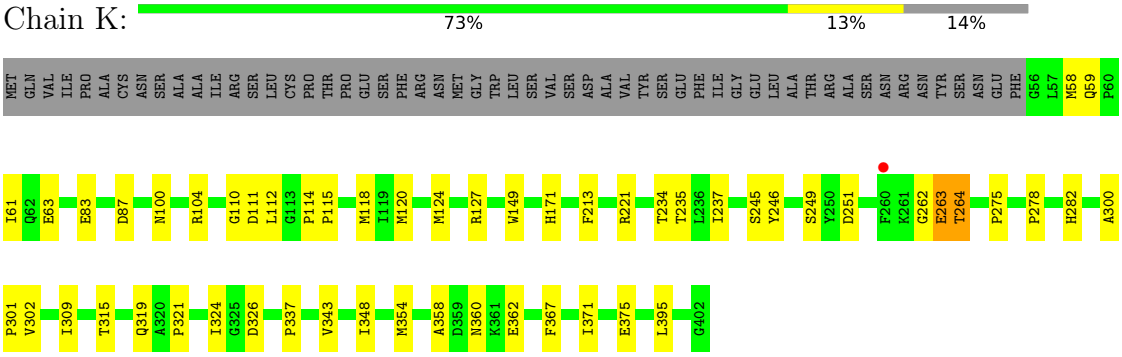
• Molecule 1: L-tryptophan decarboxylase

Chain I:



• Molecule 1: L-tryptophan decarboxylase

Chain K:

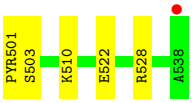


• Molecule 2: L-tryptophan decarboxylase

Chain L:

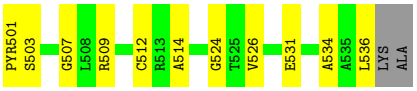
• Molecule 2: L-tryptophan decarboxylase

Chain J:



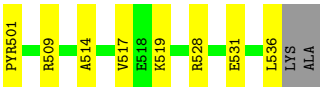
• Molecule 2: L-tryptophan decarboxylase

Chain H:



• Molecule 2: L-tryptophan decarboxylase

Chain F:

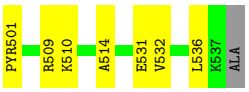


• Molecule 2: L-tryptophan decarboxylase

Chain D:

• Molecule 2: L-tryptophan decarboxylase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	82.28Å 120.76Å 128.46Å 90.00° 99.43° 90.00°	Depositor
Resolution (Å)	29.37 – 2.55 29.37 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.37-2.55) 99.9 (29.37-2.55)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 2.54Å)	Xtriage
Refinement program	PHENIX 1.21rc1_5156	Depositor
R, R_{free}	0.172 , 0.222 0.172 , 0.222	Depositor DCC
R_{free} test set	4094 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	37.5	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 35.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	19055	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% 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

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

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	0.42	0/2886	0.60	0/3916
1	C	0.41	0/2835	0.58	0/3850
1	E	0.44	0/2818	0.58	0/3821
1	G	0.43	0/2844	0.61	0/3860
1	I	0.41	1/2836 (0.0%)	0.59	0/3850
1	K	0.42	0/2853	0.57	0/3870
2	B	0.48	0/255	0.68	0/345
2	D	0.45	0/254	0.63	0/343
2	F	0.49	0/264	0.66	0/355
2	H	0.40	0/254	0.60	0/344
2	J	0.49	0/271	0.64	0/364
2	L	0.43	0/260	0.65	0/350
All	All	0.42	1/18630 (0.0%)	0.59	0/25268

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
2	D	0	1
2	F	0	1
2	H	0	1
2	J	0	1
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	122	LYS	CE-NZ	5.85	1.63	1.49

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	501	PYR	Mainchain
2	D	501	PYR	Mainchain
2	F	501	PYR	Mainchain
2	H	501	PYR	Mainchain
2	J	501	PYR	Mainchain

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	2808	0	2665	23	0
1	C	2758	0	2612	34	0
1	E	2743	0	2615	24	0
1	G	2767	0	2631	37	0
1	I	2759	0	2614	33	0
1	K	2776	0	2650	35	0
2	B	258	0	259	6	0
2	D	257	0	261	4	0
2	F	267	0	276	5	0
2	H	257	0	254	9	0
2	J	274	0	279	2	0
2	L	263	0	272	7	0
3	A	158	0	0	4	1
3	B	9	0	0	3	0
3	C	139	0	0	11	0
3	D	10	0	0	0	0
3	E	146	0	0	5	0
3	F	12	0	0	1	0
3	G	143	0	0	7	1
3	H	10	0	0	2	0
3	I	93	0	0	10	0
3	J	12	0	0	0	0
3	K	130	0	0	5	0
3	L	6	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	19055	0	17388	201	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 6.

The worst 5 of 201 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:C:101:ASP:OD2	3:C:502:HOH:O	1.82	0.98
1:C:203:PRO:O	3:C:501:HOH:O	1.80	0.96
1:I:313:PRO:O	3:I:502:HOH:O	1.86	0.91
1:G:328:ILE:O	3:G:501:HOH:O	1.95	0.83
1:A:334:ASP:O	3:A:501:HOH:O	1.96	0.82

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)
3:A:567:HOH:O	3:G:513:HOH:O[2_546]	2.03	0.17

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	349/402 (87%)	338 (97%)	11 (3%)	0	100 100
1	C	344/402 (86%)	330 (96%)	14 (4%)	0	100 100
1	E	338/402 (84%)	327 (97%)	10 (3%)	1 (0%)	37 46
1	G	344/402 (86%)	330 (96%)	14 (4%)	0	100 100
1	I	344/402 (86%)	331 (96%)	13 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	K	345/402 (86%)	331 (96%)	13 (4%)	1 (0%)	37 46
2	B	34/37 (92%)	33 (97%)	1 (3%)	0	100 100
2	D	33/37 (89%)	32 (97%)	1 (3%)	0	100 100
2	F	33/37 (89%)	33 (100%)	0	0	100 100
2	H	33/37 (89%)	32 (97%)	1 (3%)	0	100 100
2	J	35/37 (95%)	34 (97%)	1 (3%)	0	100 100
2	L	33/37 (89%)	32 (97%)	1 (3%)	0	100 100
All	All	2265/2634 (86%)	2183 (96%)	80 (4%)	2 (0%)	48 61

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	261	LYS
1	K	263	GLU

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	302/350 (86%)	298 (99%)	4 (1%)	65 79
1	C	296/350 (85%)	293 (99%)	3 (1%)	73 84
1	E	296/350 (85%)	294 (99%)	2 (1%)	81 90
1	G	298/350 (85%)	291 (98%)	7 (2%)	45 63
1	I	296/350 (85%)	288 (97%)	8 (3%)	40 57
1	K	299/350 (85%)	294 (98%)	5 (2%)	56 73
2	B	25/29 (86%)	25 (100%)	0	100 100
2	D	26/29 (90%)	25 (96%)	1 (4%)	28 41
2	F	28/29 (97%)	28 (100%)	0	100 100
2	H	26/29 (90%)	25 (96%)	1 (4%)	28 41
2	J	27/29 (93%)	25 (93%)	2 (7%)	11 15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	L	27/29 (93%)	26 (96%)	1 (4%)	29 42
All	All	1946/2274 (86%)	1912 (98%)	34 (2%)	56 73

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	264	THR
2	L	531	GLU
2	H	531	GLU
1	G	178	ARG
1	G	165	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	166	GLN
1	K	59	GLN
1	K	232	ASN
1	C	296	HIS
1	C	311	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 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

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	350/402 (87%)	-0.64	1 (0%) 90 92	19, 31, 42, 64	1 (0%)
1	C	346/402 (86%)	-0.59	1 (0%) 90 92	21, 33, 51, 63	0
1	E	342/402 (85%)	-0.64	1 (0%) 90 92	22, 30, 46, 58	0
1	G	346/402 (86%)	-0.60	0 100 100	22, 31, 43, 50	0
1	I	346/402 (86%)	-0.41	1 (0%) 90 92	24, 38, 54, 75	0
1	K	347/402 (86%)	-0.48	1 (0%) 90 92	24, 36, 49, 59	0
2	B	35/37 (94%)	-0.60	0 100 100	24, 33, 51, 63	0
2	D	34/37 (91%)	-0.64	0 100 100	20, 30, 49, 52	0
2	F	34/37 (91%)	-0.68	0 100 100	22, 28, 41, 53	0
2	H	34/37 (91%)	-0.53	0 100 100	23, 34, 50, 55	0
2	J	36/37 (97%)	-0.41	1 (2%) 55 59	23, 34, 61, 72	0
2	L	34/37 (91%)	-0.19	0 100 100	30, 43, 60, 63	0
All	All	2284/2634 (86%)	-0.56	6 (0%) 90 92	19, 33, 50, 75	1 (0%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	330	ASP	2.5
1	C	292	VAL	2.5
2	J	538	ALA	2.5
1	E	328	ILE	2.3
1	K	260	PHE	2.3

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 (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.