wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 08:47 pm BST

PDB ID	:	1XK6
Title	:	Crystal Structure- P1 form- of Escherichia coli Crotonobetainyl-CoA: carnitine CoA Transferase (CaiB)
Authors	:	Rangarajan, E.S.; Li, Y.; Iannuzzi, P.; Cygler, M.; Matte, A.
Deposited on	:	2004-09-27
Resolution	:	1.85 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 following versions of software and data (see references (1)) were used in the production of this report:

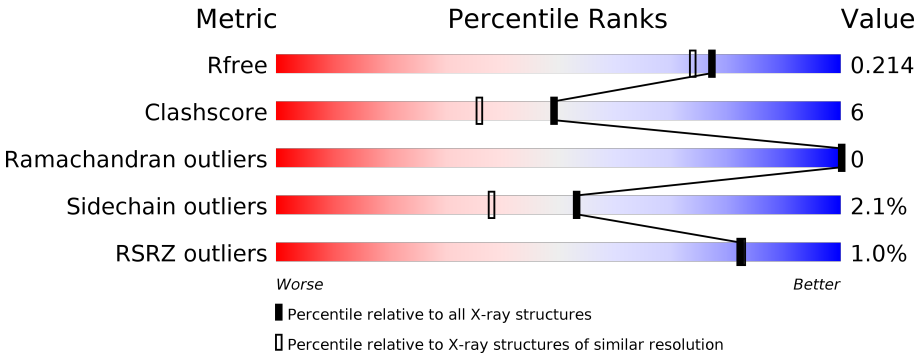
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 on 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	408	88% 9%
1	B	408	86% 12%
1	C	408	90% 7%
1	D	408	88% 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14136 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 Crotonobetainyl-CoA:carnitine CoA-transferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	Se	0	0	0
			3144	2014	522	585	7	16			
1	B	402	Total	C	N	O	S	Se	0	0	0
			3144	2014	522	585	7	16			
1	C	402	Total	C	N	O	S	Se	0	0	0
			3144	2014	522	585	7	16			
1	D	402	Total	C	N	O	S	Se	0	0	0
			3144	2014	522	585	7	16			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	CLONING ARTIFACT	UNP P31572
A	-1	SER	-	CLONING ARTIFACT	UNP P31572
A	0	HIS	-	CLONING ARTIFACT	UNP P31572
A	6	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	32	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	86	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	161	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	200	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	204	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	207	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	212	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	213	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	221	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	225	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	248	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	346	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	357	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	371	MSE	MET	MODIFIED RESIDUE	UNP P31572
A	376	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	-2	GLY	-	CLONING ARTIFACT	UNP P31572
B	-1	SER	-	CLONING ARTIFACT	UNP P31572

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	HIS	-	CLONING ARTIFACT	UNP P31572
B	6	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	32	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	86	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	161	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	200	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	204	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	207	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	212	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	213	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	221	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	225	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	248	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	346	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	357	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	371	MSE	MET	MODIFIED RESIDUE	UNP P31572
B	376	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	-2	GLY	-	CLONING ARTIFACT	UNP P31572
C	-1	SER	-	CLONING ARTIFACT	UNP P31572
C	0	HIS	-	CLONING ARTIFACT	UNP P31572
C	6	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	32	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	86	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	161	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	200	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	204	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	207	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	212	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	213	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	221	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	225	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	248	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	346	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	357	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	371	MSE	MET	MODIFIED RESIDUE	UNP P31572
C	376	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	-2	GLY	-	CLONING ARTIFACT	UNP P31572
D	-1	SER	-	CLONING ARTIFACT	UNP P31572
D	0	HIS	-	CLONING ARTIFACT	UNP P31572
D	6	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	32	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	86	MSE	MET	MODIFIED RESIDUE	UNP P31572

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	161	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	200	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	204	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	207	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	212	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	213	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	221	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	225	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	248	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	346	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	357	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	371	MSE	MET	MODIFIED RESIDUE	UNP P31572
D	376	MSE	MET	MODIFIED RESIDUE	UNP P31572

- Molecule 2 is water.

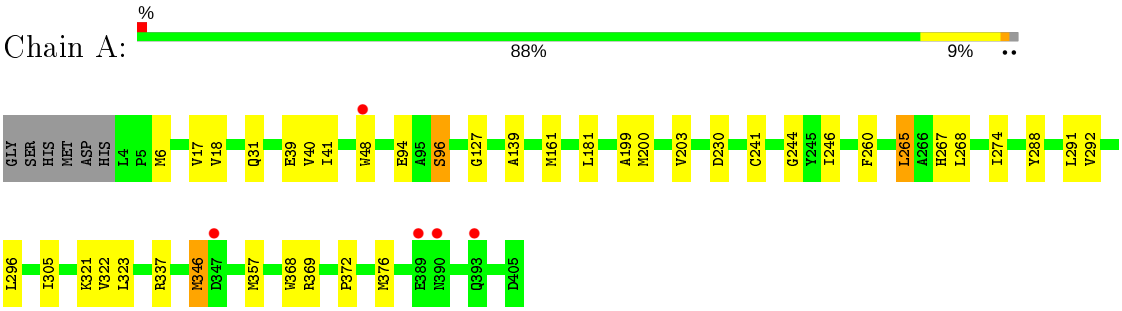
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	351	Total O 351 351	0	0
2	B	360	Total O 360 360	0	0
2	C	423	Total O 423 423	0	0
2	D	426	Total O 426 426	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

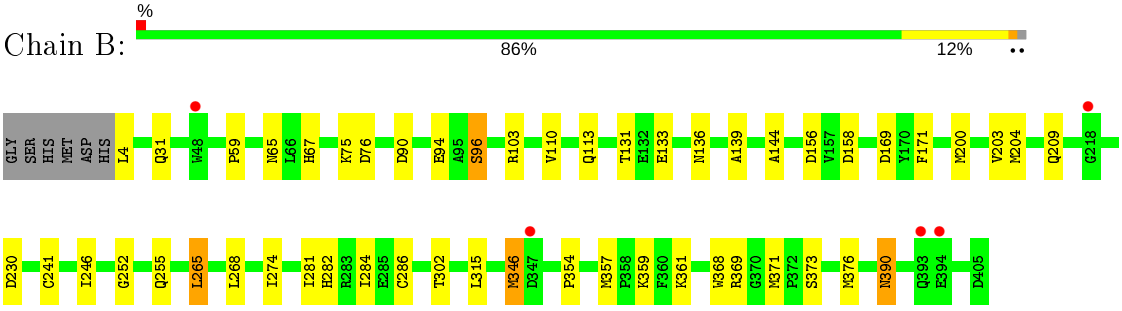
- Molecule 1: Crotonobetainyl-CoA:carnitine CoA-transferase

Chain A:



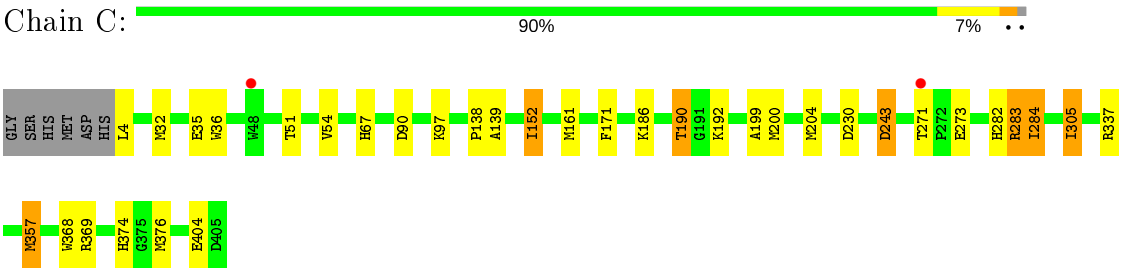
- Molecule 1: Crotonobetainyl-CoA:carnitine CoA-transferase

Chain B:



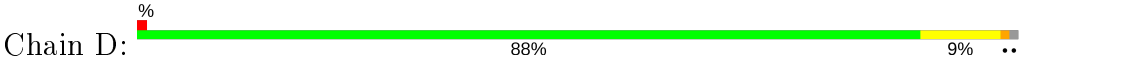
- Molecule 1: Crotonobetainyl-CoA:carnitine CoA-transferase

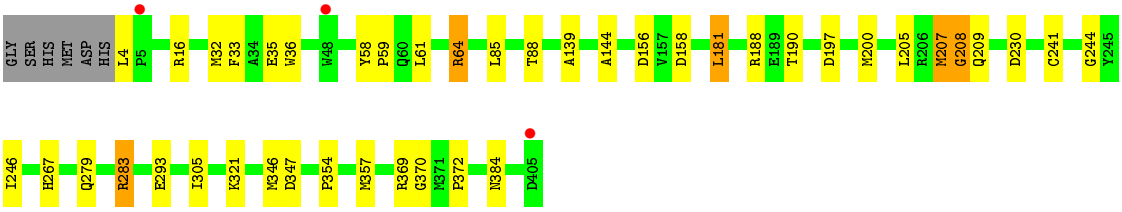
Chain C:



- Molecule 1: Crotonobetainyl-CoA:carnitine CoA-transferase

Chain D:




GLY SER HIS MET ASP HIS L4 P5 R16 M32 F33 A34 E35 W36 W48 Y58 P59 Q60 L61 R64 L85 T88 A139 A144 D156 V157 D158 L181 R188 E189 T190 D197 M200 L205 R206 M207 G208 Q209 D230 C241 G244 Y245
I246 H267 Q279 R283 E293 I305 K321 M346 D347 P354 M357 R369 G370 M371 P372 N384 D405

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	70.84Å 83.17Å 89.94Å 100.39° 110.01° 100.03°	Depositor
Resolution (Å)	48.54 – 1.85 48.54 – 1.85	Depositor EDS
% Data completeness (in resolution range)	95.3 (48.54-1.85) 95.4 (48.54-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	3.00 (at 1.86Å)	Xtriage
Refinement program	CNS 1.1, REFMAC 5.1.24	Depositor
R, R_{free}	0.177 , 0.215 0.179 , 0.214	Depositor DCC
R_{free} test set	14911 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å^2)	17.3	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14136	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 56.89 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.5620e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.37	0/3207	0.63	1/4324 (0.0%)
1	B	0.38	0/3207	0.64	6/4324 (0.1%)
1	C	0.39	0/3207	0.64	3/4324 (0.1%)
1	D	0.39	0/3207	0.68	6/4324 (0.1%)
All	All	0.38	0/12828	0.65	16/17296 (0.1%)

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
1	D	0	2

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	347	ASP	CB-CG-OD2	7.10	124.69	118.30
1	D	156	ASP	CB-CG-OD2	6.23	123.91	118.30
1	D	283	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	C	283	ARG	NE-CZ-NH2	-5.51	117.55	120.30
1	B	90	ASP	CB-CG-OD2	5.32	123.09	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	207	MSE	Peptide
1	D	208	GLY	Peptide

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	3144	0	3087	42	0
1	B	3144	0	3087	41	0
1	C	3144	0	3087	48	0
1	D	3144	0	3087	38	0
2	A	351	0	0	6	0
2	B	360	0	0	5	0
2	C	423	0	0	3	0
2	D	426	0	0	2	0
All	All	14136	0	12348	148	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 6.

The worst 5 of 148 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:171:PHE:CB	1:C:204:MSE:HE1	1.76	1.15
1:C:171:PHE:HB2	1:C:204:MSE:HE1	1.33	1.09
1:A:346:MSE:HE1	1:A:372:PRO:HA	1.37	1.06
1:B:203:VAL:HG12	1:B:204:MSE:CE	1.89	1.03
1:A:199:ALA:HB2	1:C:337:ARG:HH12	1.27	0.98

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	400/408 (98%)	387 (97%)	13 (3%)	0	100	100
1	B	400/408 (98%)	390 (98%)	10 (2%)	0	100	100
1	C	400/408 (98%)	390 (98%)	10 (2%)	0	100	100
1	D	400/408 (98%)	388 (97%)	12 (3%)	0	100	100
All	All	1600/1632 (98%)	1555 (97%)	45 (3%)	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	329/318 (104%)	322 (98%)	7 (2%)	53	38
1	B	329/318 (104%)	319 (97%)	10 (3%)	41	24
1	C	329/318 (104%)	324 (98%)	5 (2%)	65	53
1	D	329/318 (104%)	323 (98%)	6 (2%)	59	45
All	All	1316/1272 (104%)	1288 (98%)	28 (2%)	53	38

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	315	LEU
1	B	361	LYS
1	D	279	GLN
1	B	346	MSE
1	B	357	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	209	GLN
1	C	45	ASN
1	D	374	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	374	HIS
1	B	384	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	386/408 (94%)	-0.12	5 (1%)	77 78	8, 19, 32, 41	0
1	B	386/408 (94%)	-0.15	5 (1%)	77 78	8, 18, 29, 39	0
1	C	386/408 (94%)	-0.25	2 (0%)	91 91	8, 17, 27, 35	0
1	D	386/408 (94%)	-0.24	3 (0%)	86 86	8, 17, 27, 36	0
All	All	1544/1632 (94%)	-0.19	15 (0%)	82 82	8, 18, 29, 41	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	48	TRP	3.1
1	D	5	PRO	3.0
1	C	271	THR	3.0
1	B	48	TRP	2.9
1	A	390	ASN	2.8

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 carbohydrates 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.