



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 08:14 pm GMT

PDB ID : 5FAW
Title : T502A mutant of choline TMA-lyase
Authors : Funk, M.A.; Drennan, C.L.
Deposited on : 2015-12-12
Resolution : 1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtrriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

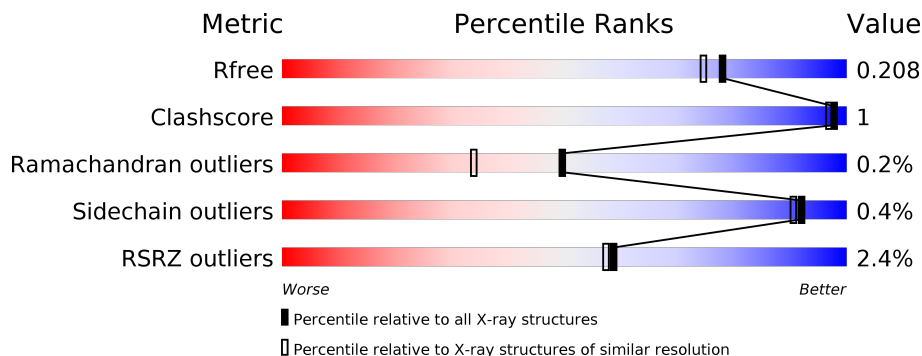
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	111664	2111 (1.86-1.86)
Clashscore	122126	2258 (1.86-1.86)
Ramachandran outliers	120053	2234 (1.86-1.86)
Sidechain outliers	120020	2234 (1.86-1.86)
RSRZ outliers	108989	2075 (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. 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	815	 96%
1	B	815	 97%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 27642 atoms, of which 12501 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 Choline trimethylamine-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	803	12590	4026	6235	1082	1201	46	0	16	0
1	B	806	12580	4025	6226	1082	1201	46	0	9	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	MET	-	initiating methionine	UNP Q30W70
A	33	GLY	-	expression tag	UNP Q30W70
A	34	SER	-	expression tag	UNP Q30W70
A	35	SER	-	expression tag	UNP Q30W70
A	36	HIS	-	expression tag	UNP Q30W70
A	37	HIS	-	expression tag	UNP Q30W70
A	38	HIS	-	expression tag	UNP Q30W70
A	39	HIS	-	expression tag	UNP Q30W70
A	40	HIS	-	expression tag	UNP Q30W70
A	41	HIS	-	expression tag	UNP Q30W70
A	42	SER	-	expression tag	UNP Q30W70
A	43	SER	-	expression tag	UNP Q30W70
A	44	GLY	-	expression tag	UNP Q30W70
A	45	LEU	-	expression tag	UNP Q30W70
A	46	VAL	-	expression tag	UNP Q30W70
A	47	PRO	-	expression tag	UNP Q30W70
A	48	ARG	-	expression tag	UNP Q30W70
A	49	GLY	-	expression tag	UNP Q30W70
A	50	SER	-	expression tag	UNP Q30W70
A	51	HIS	-	expression tag	UNP Q30W70
A	52	MET	-	expression tag	UNP Q30W70
A	502	ALA	THR	engineered mutation	UNP Q30W70
B	32	MET	-	initiating methionine	UNP Q30W70
B	33	GLY	-	expression tag	UNP Q30W70
B	34	SER	-	expression tag	UNP Q30W70

Continued on next page...

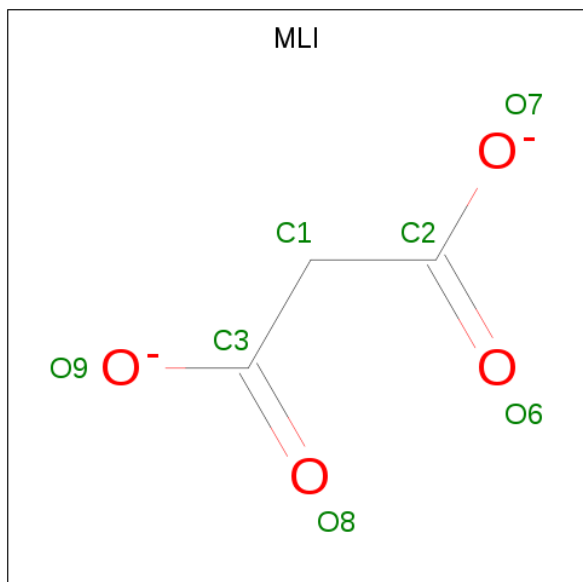
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	35	SER	-	expression tag	UNP Q30W70
B	36	HIS	-	expression tag	UNP Q30W70
B	37	HIS	-	expression tag	UNP Q30W70
B	38	HIS	-	expression tag	UNP Q30W70
B	39	HIS	-	expression tag	UNP Q30W70
B	40	HIS	-	expression tag	UNP Q30W70
B	41	HIS	-	expression tag	UNP Q30W70
B	42	SER	-	expression tag	UNP Q30W70
B	43	SER	-	expression tag	UNP Q30W70
B	44	GLY	-	expression tag	UNP Q30W70
B	45	LEU	-	expression tag	UNP Q30W70
B	46	VAL	-	expression tag	UNP Q30W70
B	47	PRO	-	expression tag	UNP Q30W70
B	48	ARG	-	expression tag	UNP Q30W70
B	49	GLY	-	expression tag	UNP Q30W70
B	50	SER	-	expression tag	UNP Q30W70
B	51	HIS	-	expression tag	UNP Q30W70
B	52	MET	-	expression tag	UNP Q30W70
B	502	ALA	THR	engineered mutation	UNP Q30W70

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

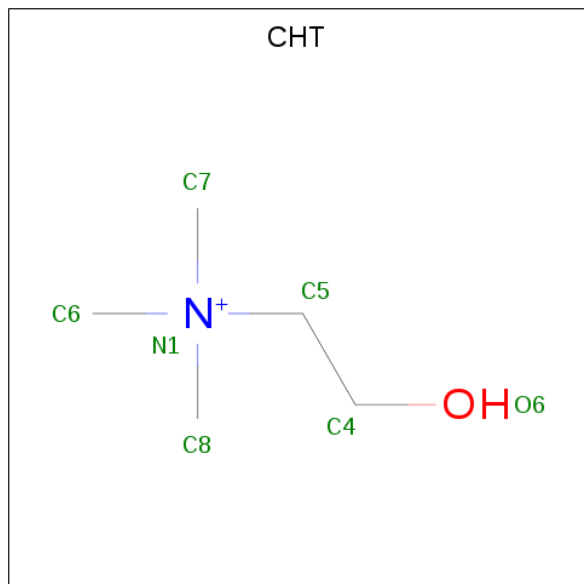
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Na 2 2	0	0

- Molecule 3 is MALONATE ION (three-letter code: MLI) (formula: C₃H₂O₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	B	1	Total	C	H	O	0	0
			9	3	2	4		

- Molecule 4 is CHOLINE ION (three-letter code: CHT) (formula: C₅H₁₄NO).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			21	5	14	1	1		
4	B	1	Total	C	H	N	O	0	0
			21	5	14	1	1		

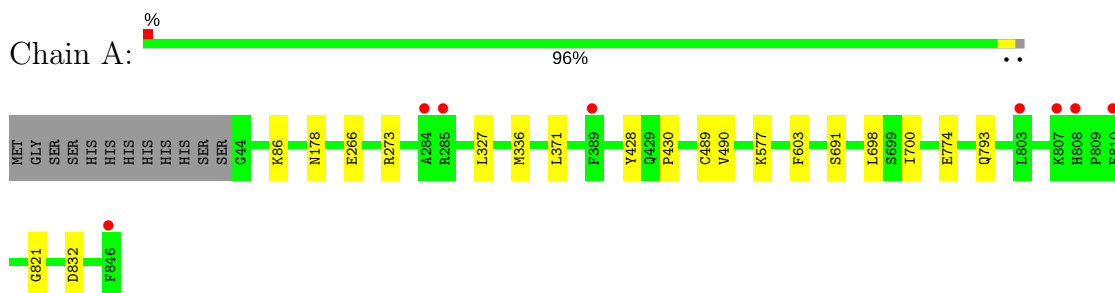
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1245	Total	O	0	0
			1245	1245		
5	B	1129	Total	O	0	0
			1129	1129		

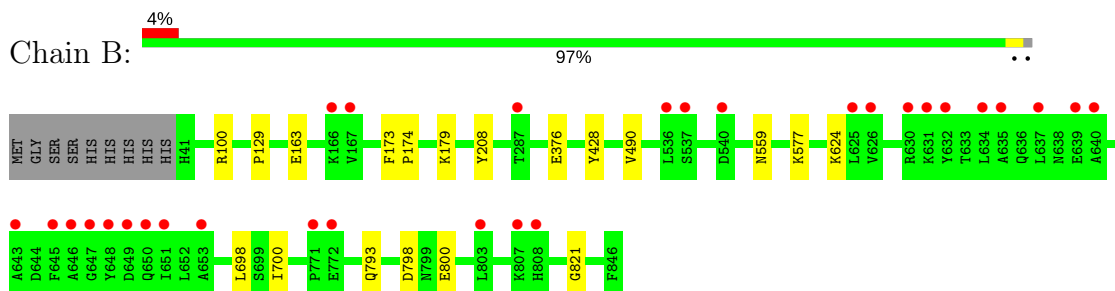
3 Residue-property plots [i](#)

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: Choline trimethylamine-lyase



- Molecule 1: Choline trimethylamine-lyase



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	229.57Å 229.57Å 78.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.48 – 1.85 49.48 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.48-1.85) 98.1 (49.48-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.63 (at 1.86Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.174 , 0.208 0.176 , 0.208	Depositor DCC
R_{free} test set	5327 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å ²)	16.8	Xtrriage
Anisotropy	0.290	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 59.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	27642	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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: NA, MLI, CHT

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	0.43	0/6538	0.53	0/8851
1	B	0.42	0/6515	0.53	1/8821 (0.0%)
All	All	0.43	0/13053	0.53	1/17672 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	100	ARG	NE-CZ-NH2	-5.46	117.57	120.30

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	6355	6235	6129	11	0
1	B	6354	6226	6169	8	1
2	A	2	0	0	0	0
3	A	35	10	10	0	0
3	B	7	2	2	0	0
4	A	7	14	14	1	0
4	B	7	14	14	0	0
5	A	1245	0	0	6	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	1129	0	0	5	2
All	All	15141	12501	12338	20	4

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 1.

All (20) 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:B:559:ASN:ND2	5:B:1003:HOH:O	2.17	0.77
1:A:266[B]:GLU:OE2	5:A:1001:HOH:O	2.05	0.75
1:A:178:ASN:OD1	5:A:1002:HOH:O	2.11	0.69
1:B:793:GLN:HE22	1:B:821:GLY:H	1.40	0.68
1:B:208:TYR:OH	5:B:1002:HOH:O	2.13	0.65
1:A:793:GLN:HE22	1:A:821:GLY:H	1.47	0.63
1:A:86:LYS:NZ	5:A:1010:HOH:O	2.31	0.58
1:B:163:GLU:OE2	5:B:1004:HOH:O	2.18	0.56
1:B:624:LYS:NZ	5:B:1041:HOH:O	2.45	0.50
1:B:798:ASP:OD2	1:B:800:GLU:HB2	2.15	0.46
1:A:273[B]:ARG:NH2	5:A:1026:HOH:O	2.41	0.45
1:A:832:ASP:OD2	5:A:1003:HOH:O	2.19	0.43
1:A:774:GLU:OE2	5:A:1004:HOH:O	2.20	0.43
1:A:489:CYS:SG	4:A:908:CHT:HC42	2.59	0.43
1:B:129:PRO:HB3	1:B:376:GLU:HG2	2.02	0.42
1:A:371:LEU:HD22	1:A:430:PRO:HD2	2.02	0.41
1:A:327:LEU:HB3	1:A:336:MET:SD	2.60	0.41
1:B:173:PHE:HB2	1:B:174:PRO:HD3	2.02	0.41
1:A:603:PHE:HA	1:A:691:SER:OG	2.22	0.40

All (4) 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:B:179:LYS:NZ	1:B:179:LYS:NZ[7_555]	1.81	0.39
5:A:1547:HOH:O	5:A:1584:HOH:O[1_556]	2.14	0.06
5:B:1244:HOH:O	5:B:1730:HOH:O[7_556]	2.19	0.01
5:B:1622:HOH:O	5:B:1788:HOH:O[7_556]	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	809/815 (99%)	790 (98%)	17 (2%)	2 (0%)	49	34
1	B	807/815 (99%)	784 (97%)	21 (3%)	2 (0%)	49	34
All	All	1616/1630 (99%)	1574 (97%)	38 (2%)	4 (0%)	49	34

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	700	ILE
1	B	700	ILE
1	B	490	VAL
1	A	490	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	682/686 (99%)	679 (100%)	3 (0%)	92	90
1	B	680/686 (99%)	677 (100%)	3 (0%)	92	90
All	All	1362/1372 (99%)	1356 (100%)	6 (0%)	92	90

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	428	TYR
1	A	577	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	698	LEU
1	B	428	TYR
1	B	577	LYS
1	B	698	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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	MLI	A	902	-	0,6,6	0.00	-	0,7,7	0.00	-
3	MLI	A	903	-	0,6,6	0.00	-	0,7,7	0.00	-
3	MLI	A	904	-	0,6,6	0.00	-	0,7,7	0.00	-
3	MLI	A	905	-	0,6,6	0.00	-	0,7,7	0.00	-
3	MLI	A	906	-	0,6,6	0.00	-	0,7,7	0.00	-
4	CHT	A	908	-	6,6,6	1.40	0	8,8,8	0.46	0
3	MLI	B	901	-	0,6,6	0.00	-	0,7,7	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	CHT	B	902	-	6,6,6	1.50	1 (16%)	8,8,8	0.74	0

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	MLI	A	902	-	-	0/0/4/4	0/0/0/0
3	MLI	A	903	-	-	0/0/4/4	0/0/0/0
3	MLI	A	904	-	-	0/0/4/4	0/0/0/0
3	MLI	A	905	-	-	0/0/4/4	0/0/0/0
3	MLI	A	906	-	-	0/0/4/4	0/0/0/0
4	CHT	A	908	-	-	0/4/4/4	0/0/0/0
3	MLI	B	901	-	-	0/0/4/4	0/0/0/0
4	CHT	B	902	-	-	0/4/4/4	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	902	CHT	C6-N1	-2.04	1.44	1.50

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	908	CHT	1	0

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, 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	803/815 (98%)	-0.12	8 (0%) 82 82	11, 19, 32, 43	0
1	B	806/815 (98%)	0.06	30 (3%) 41 39	11, 20, 35, 48	0
All	All	1609/1630 (98%)	-0.03	38 (2%) 59 57	11, 20, 33, 48	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	630	ARG	4.2
1	B	632	TYR	3.4
1	A	285	ARG	3.1
1	A	808	HIS	3.0
1	B	648	TYR	3.0
1	B	651	ILE	2.9
1	B	645	PHE	2.9
1	B	771	PRO	2.8
1	B	625	LEU	2.8
1	B	646	ALA	2.8
1	A	807	LYS	2.7
1	B	643	ALA	2.7
1	B	634	LEU	2.7
1	B	536	LEU	2.6
1	B	639	GLU	2.6
1	B	808	HIS	2.6
1	B	167	VAL	2.5
1	B	637	LEU	2.5
1	B	166	LYS	2.4
1	B	540	ASP	2.3
1	A	810	GLU	2.3
1	B	640	ALA	2.3
1	A	803	LEU	2.3
1	B	650	GLN	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	649	ASP	2.3
1	B	653	ALA	2.2
1	B	631	LYS	2.2
1	B	537	SER	2.2
1	B	635	ALA	2.1
1	A	846	PHE	2.1
1	B	626	VAL	2.1
1	B	647	GLY	2.1
1	B	807	LYS	2.1
1	B	772	GLU	2.1
1	A	389	PHE	2.1
1	A	284	ALA	2.1
1	B	803	LEU	2.0
1	B	287	THR	2.0

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 carbohydrates 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
3	MLI	B	901	7/7	0.73	0.29	60,72,87,87	0
3	MLI	A	906	7/7	0.79	0.20	46,52,64,64	0
3	MLI	A	904	7/7	0.83	0.18	44,58,70,72	0
3	MLI	A	903	7/7	0.90	0.14	28,45,54,55	0
3	MLI	A	905	7/7	0.91	0.12	31,40,48,48	0
2	NA	A	907	1/1	0.91	0.21	57,57,57,57	0
4	CHT	A	908	7/7	0.93	0.24	23,27,32,32	0
4	CHT	B	902	7/7	0.94	0.27	20,34,44,44	0
3	MLI	A	902	7/7	0.94	0.14	27,44,53,54	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	NA	A	901	1/1	0.99	0.08	24,24,24,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.