



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 09:28 am BST

PDB ID : 1T6J
Title : Crystal Structure of Phenylalanine Ammonia Lyase from *Rhodospiridium toruloides*
Authors : Calabrese, J.C.; Jordan, D.B.
Deposited on : 2004-05-06
Resolution : 2.10 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 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 : 2.11
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

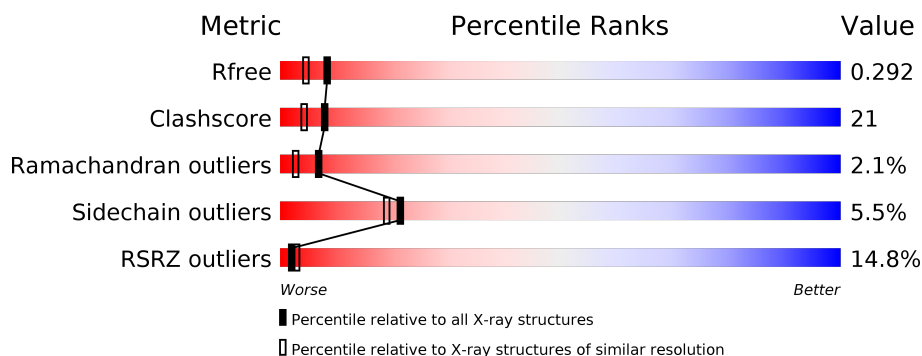
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 $\leq 5\%$. 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	714	
1	B	714	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	175	A	211	X	-	-	-
1	175	B	211	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10300 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 phenylalanine ammonia-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	647	Total	C	N	O	S	Se	0	0	0
			4939	3104	878	940	4	13			
1	B	647	Total	C	N	O	S	Se	0	0	0
			4939	3104	878	940	4	13			

There are 36 discrepancies between the modelled and reference sequences:

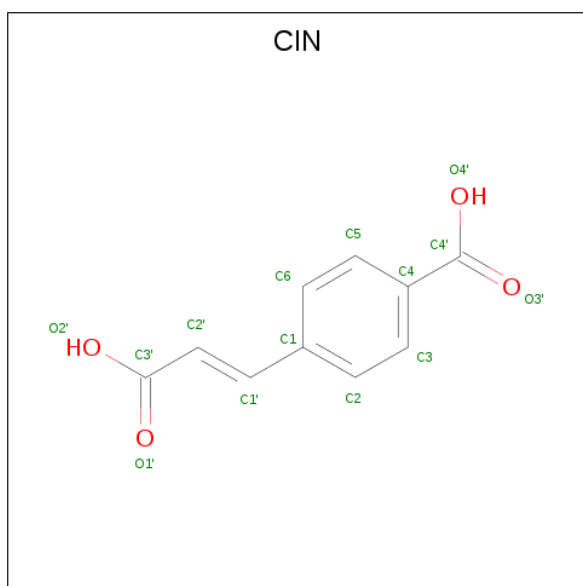
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	211	175	ALA	SEE REMARK 999	UNP P11544
A	211	175	SER	SEE REMARK 999	UNP P11544
A	211	175	GLY	SEE REMARK 999	UNP P11544
A	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	211	175	ALA	SEE REMARK 999	UNP P11544
B	211	175	SER	SEE REMARK 999	UNP P11544
B	211	175	GLY	SEE REMARK 999	UNP P11544

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	714	MSE	MET	MODIFIED RESIDUE	UNP P11544

- Molecule 2 is 4-CARBOXYCINNAMIC ACID (three-letter code: CIN) (formula: C₁₀H₈O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	C	O	0	0
			11	9	2		

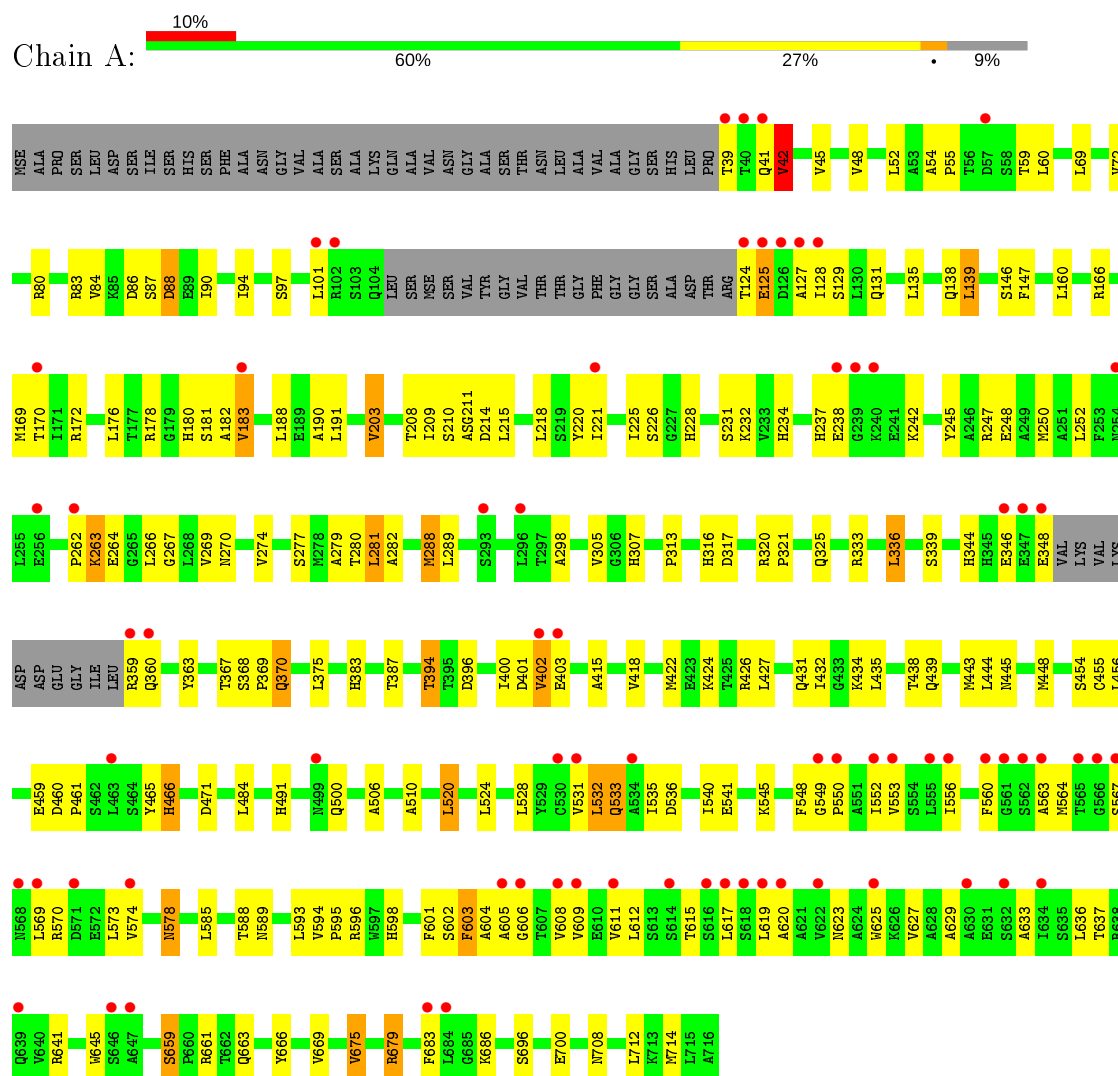
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	245	Total	O	0	0
			245	245		
3	B	166	Total	O	0	0
			166	166		

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: phenylalanine ammonia-lyase



- Molecule 1: phenylalanine ammonia-lyase





4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	107.86Å 107.86Å 204.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 2.10 24.77 – 2.10	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.10) 93.6 (24.77-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.42 (at 2.10Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.243 , 0.298 0.246 , 0.292	Depositor DCC
R_{free} test set	1934 reflections (2.56%)	wwPDB-VP
Wilson B-factor (Å ²)	31.2	Xtriage
Anisotropy	0.701	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10300	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 175, CIN

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	0.33	0/4995	0.64	0/6756
1	B	0.32	0/4995	0.61	0/6756
All	All	0.32	0/9990	0.63	0/13512

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	A	2	1
1	B	2	1
All	All	4	2

There are no bond length outliers.

There are no bond angle outliers.

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	211	175	C3,C2
1	B	211	175	C3,C2

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	210	SER	Peptide
1	B	210	SER	Peptide

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	4939	0	4998	187	0
1	B	4939	0	4998	244	0
2	B	11	0	6	0	0
3	A	245	0	0	6	0
3	B	166	0	0	17	0
All	All	10300	0	10002	427	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 21.

The worst 5 of 427 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:247:ARG:HB3	1:B:247:ARG:HH11	1.11	1.08
1:A:279:ALA:HA	1:A:422:MSE:HE2	1.42	1.01
1:A:422:MSE:HE3	1:A:510:ALA:HA	1.43	0.99
1:A:593:LEU:HD11	1:A:637:THR:HG21	1.46	0.94
1:A:169:MSE:HE3	1:A:221:ILE:HG23	1.51	0.93

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	640/714 (90%)	587 (92%)	44 (7%)	9 (1%)	11 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	640/714 (90%)	571 (89%)	51 (8%)	18 (3%)	5 1
All	All	1280/1428 (90%)	1158 (90%)	95 (7%)	27 (2%)	7 3

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	VAL
1	B	238	GLU
1	B	360	GLN
1	B	454	SER
1	A	88	ASP

5.3.2 Protein sidechains ⓘ

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	538/573 (94%)	516 (96%)	22 (4%)	30 31
1	B	538/573 (94%)	501 (93%)	37 (7%)	15 12
All	All	1076/1146 (94%)	1017 (94%)	59 (6%)	21 19

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	169	MSE
1	B	270	ASN
1	B	596	ARG
1	B	172	ARG
1	B	200	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	180	HIS
1	B	284	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	578	ASN
1	B	254	ASN
1	B	294	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	175	A	211	1	12,14,15	1.36	2 (16%)	9,19,21	5.34	3 (33%)
1	175	B	211	1	12,14,15	1.28	1 (8%)	9,19,21	5.31	4 (44%)

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
1	175	A	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	B	211	1	2/2/5/7	0/3/25/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	211	175	C3-C5	-2.98	1.47	1.52
1	B	211	175	C3-C5	-2.88	1.48	1.52
1	A	211	175	O6-C5	2.34	1.26	1.22

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	211	175	C4-C3-N2	13.21	130.64	114.09
1	A	211	175	C4-C3-N2	13.07	130.47	114.09
1	B	211	175	C4-C3-C5	7.60	130.68	112.19
1	A	211	175	C4-C3-C5	7.58	130.65	112.19
1	A	211	175	C-CA-N0	4.02	119.49	112.65

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	211	175	C3
1	A	211	175	C2
1	B	211	175	C3
1	B	211	175	C2

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

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	$\# Z > 2$
2	CIN	B	1001	-	8,11,14	1.53	2 (25%)	10,13,18	0.35	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
2	CIN	B	1001	-	-	0/3/5/9	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1001	CIN	C2-C1	2.28	1.43	1.39
2	B	1001	CIN	C6-C1	2.15	1.43	1.39

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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	633/714 (88%)	0.75	72 (11%) 5 6	23, 42, 88, 103	0
1	B	633/714 (88%)	1.00	116 (18%) 1 1	23, 48, 87, 101	0
All	All	1266/1428 (88%)	0.88	188 (14%) 2 3	23, 45, 88, 103	0

The worst 5 of 188 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	127	ALA	9.4
1	A	124	THR	8.6
1	A	569	LEU	6.2
1	B	254	ASN	6.2
1	B	239	GLY	6.2

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	175	A	211	14/15	0.89	0.15	36,40,44,44	0
1	175	B	211	14/15	0.90	0.14	37,43,44,45	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands

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(\AA^2)	Q<0.9
2	CIN	B	1001	11/14	0.60	0.35	83,84,85,85	0

6.5 Other polymers

There are no such residues in this entry.