



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 10:45 pm GMT

PDB ID : 5MST
Title : Structure of the A domain of carboxylic acid reductase (CAR) from *Segniliparus rugosus* in complex with AMP and a co-purified carboxylic acid
Authors : Gahloth, D.; Leys, D.
Deposited on : 2017-01-05
Resolution : 1.72 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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (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

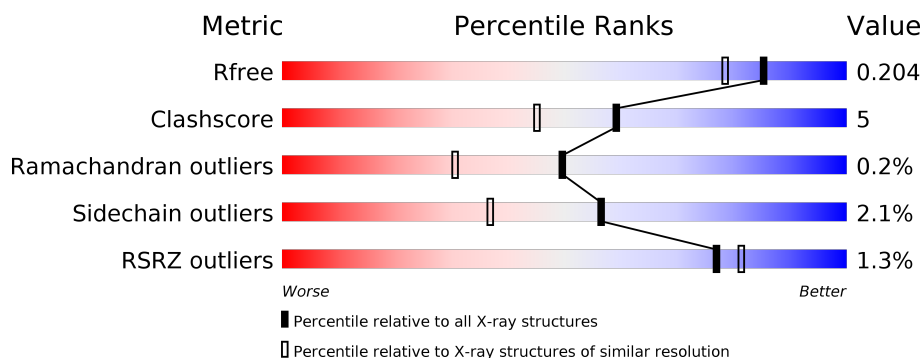
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.72 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	4894 (1.74-1.70)
Clashscore	122126	5296 (1.74-1.70)
Ramachandran outliers	120053	5219 (1.74-1.70)
Sidechain outliers	120020	5219 (1.74-1.70)
RSRZ outliers	108989	4804 (1.74-1.70)

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	1188	<div> <div>%</div> <div> <div></div> <div>48%</div> <div>5%</div> <div>47%</div> </div> </div>
1	B	1188	<div> <div>%</div> <div> <div></div> <div>48%</div> <div>5%</div> <div>47%</div> </div> </div>

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
3	FUM	B	1203	-	-	-	X

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 11374 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.

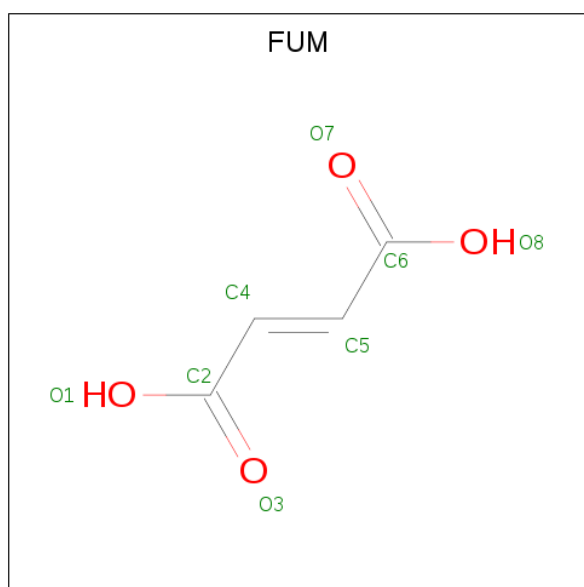
- Molecule 1 is a protein called Thioester reductase domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	631	Total	C	N	O	S	0	0	0
			4809	3050	817	935	7			
1	B	632	Total	C	N	O	S	0	0	0
			4819	3055	819	938	7			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

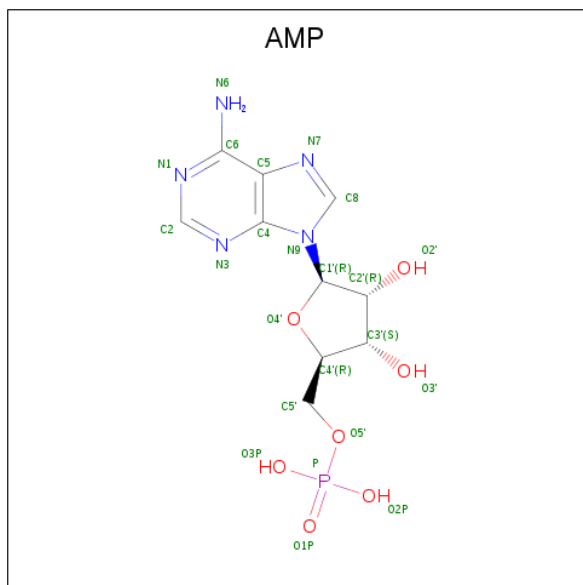
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Ca	0	0
			2	2		
2	A	2	Total	Ca	0	0
			2	2		

- Molecule 3 is FUMARIC ACID (three-letter code: FUM) (formula: C₄H₄O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	4	4		
3	B	1	Total	C	O	0	0
			8	4	4		

- Molecule 4 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$).




Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
4	B	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

- Molecule 5 is water.

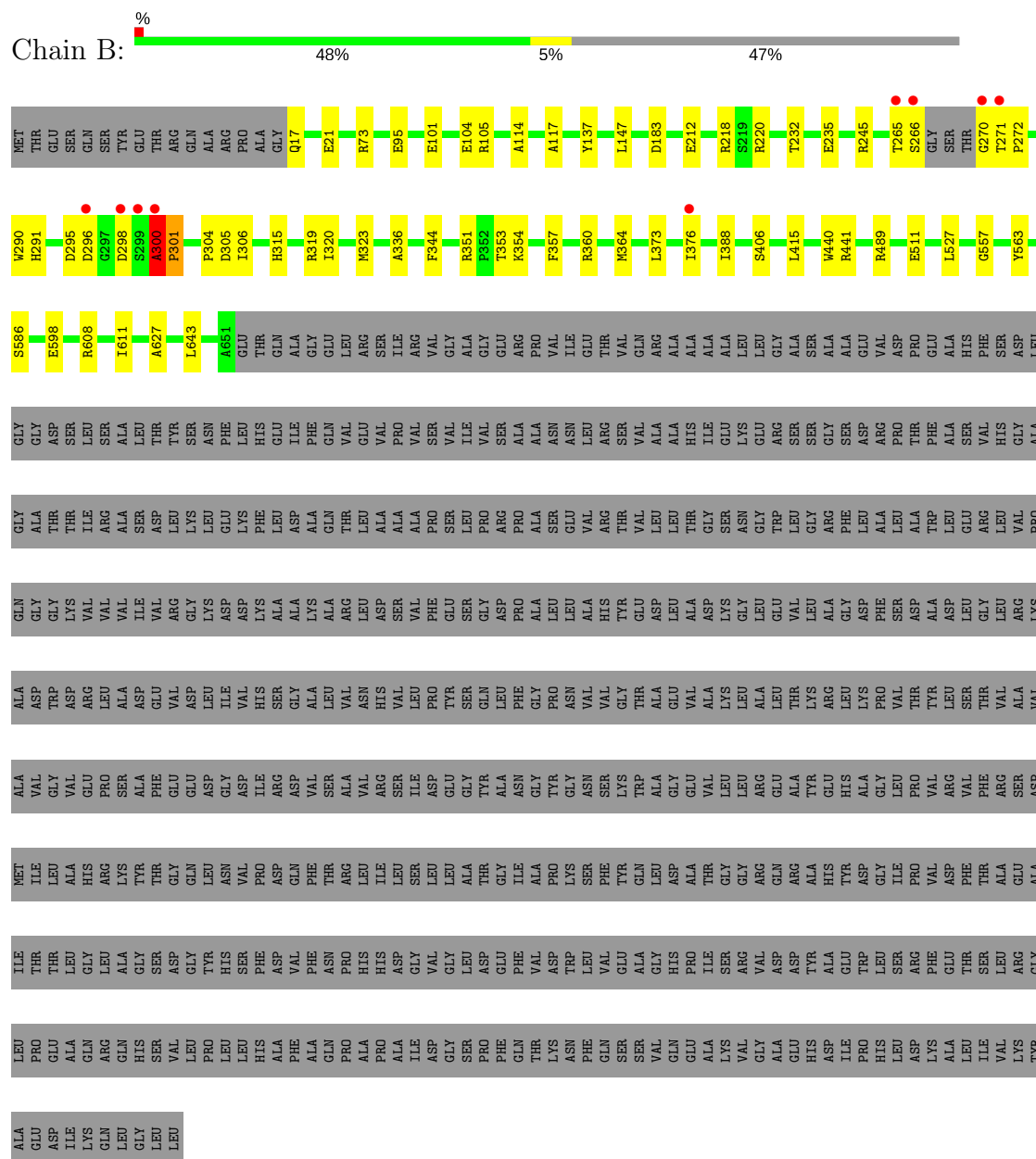
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	849	Total	O	0	0
			849	849		
5	B	831	Total	O	0	0
			831	831		

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.

- Chain A: 

Position	Residue	Category	Score
1	ALA	Core	1.00
2	LEU	Core	1.00
3	THR	Core	1.00
4	ILE	Core	1.00
5	LEU	Core	1.00
6	VAL	Core	1.00
7	LYS	Core	1.00
8	ARG	Core	1.00
9	GLY	Core	1.00
10	THR	Core	1.00
11	ALA	Core	1.00
12	TYR	Core	1.00
13	LEU	Core	1.00
14	GLY	Core	1.00
15	PRO	Core	1.00
16	THR	Core	1.00
17	GLN	Core	1.00
18	GLY	Core	1.00
19	ILE	Core	1.00
20	ASP	Core	1.00
21	THR	Core	1.00
22	GLY	Core	1.00
23	ASP	Core	1.00
24	GLY	Core	1.00
25	GLN	Core	1.00
26	ARG	Core	1.00
27	LEU	Core	1.00
28	PRO	Core	1.00
29	GLN	Core	1.00
30	LEU	Core	1.00
31	GLY	Core	1.00
32	THR	Core	1.00
33	ASP	Core	1.00
34	GLY	Core	1.00
35	ASP	Core	1.00
36	GLY	Core	1.00
37	GLN	Core	1.00
38	ARG	Core	1.00
39	LEU	Core	1.00
40	ASP	Core	1.00
41	GLY	Core	1.00
42	THR	Core	1.00
43	ASN	Core	1.00
44	GLN	Core	1.00
45	ALA	Core	1.00
46	THR	Core	1.00
47	ARG	Core	1.00
48	THR	Core	1.00
49	PHE	Core	1.00
50	GLN	Core	1.00
51	VAL	Core	1.00
52	PRO	Core	1.00
53	ASP	Core	1.00
54	GLY	Core	1.00
55	ASP	Core	1.00
56	GLY	Core	1.00
57	ASP	Core	1.00
58	GLY	Core	1.00
59	ASP	Core	1.00
60	GLY	Core	1.00
61	ASP	Core	1.00
62	GLY	Core	1.00
63	ASP	Core	1.00
64	GLY	Core	1.00
65	ASP	Core	1.00
66	GLY	Core	1.00
67	ASP	Core	1.00
68	GLY	Core	1.00
69	ASP	Core	1.00
70	GLY	Core	1.00
71	ASP	Core	1.00
72	GLY	Core	1.00
73	ASP	Core	1.00
74	GLY	Core	1.00
75	ASP	Core	1.00
76	GLY	Core	1.00
77	ASP	Core	1.00
78	GLY	Core	1.00
79	ASP	Core	1.00
80	GLY	Core	1.00
81	ASP	Core	1.00
82	GLY	Core	1.00
83	ASP	Core	1.00
84	GLY	Core	1.00
85	ASP	Core	1.00
86	GLY	Core	1.00
87	ASP	Core	1.00
88	GLY	Core	1.00
89	ASP	Core	1.00
90	GLY	Core	1.00
91	ASP	Core	1.00
92	GLY	Core	1.00
93	ASP	Core	1.00
94	GLY	Core	1.00
95	ASP	Core	1.00
96	GLY	Core	1.00
97	ASP	Core	1.00
98	GLY	Core	1.00
99	ASP	Core	1.00
100	GLY	Core	1.00
101	ASP	Core	1.00
102	GLY	Core	1.00
103	ASP	Core	1.00
104	GLY	Core	1.00
105	ASP	Core	1.00
106	GLY	Core	1.00
107	ASP	Core	1.00
108	GLY	Core	1.00
109	ASP	Core	1.00
110	GLY	Core	1.00
111	ASP	Core	1.00
112	GLY	Core	1.00
113	ASP	Core	1.00
114	GLY	Core	1.00
115	ASP	Core	1.00
116	GLY	Core	1.00
117	ASP	Core	1.00
118	GLY	Core	1.00

- 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.30Å 91.00Å 113.78Å 90.00° 92.84° 90.00°	Depositor
Resolution (Å)	113.64 – 1.72 113.64 – 1.72	Depositor EDS
% Data completeness (in resolution range)	98.6 (113.64-1.72) 98.6 (113.64-1.72)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.60 (at 1.72Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.176 , 0.204 0.177 , 0.204	Depositor DCC
R_{free} test set	7340 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	21.6	Xtriage
Anisotropy	0.408	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.088 for h,-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11374	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: AMP, CA, FUM

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	0.38	0/4905	0.55	0/6673
1	B	0.37	0/4915	0.58	2/6686 (0.0%)
All	All	0.37	0/9820	0.57	2/13359 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	300	ALA	C-N-CD	-11.86	94.50	120.60
1	B	300	ALA	C-N-CA	7.59	153.86	122.00

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	4809	0	4790	46	0
1	B	4819	0	4803	49	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	8	0	2	0	0
3	B	8	0	2	3	0
4	A	23	0	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	23	0	12	3	0
5	A	849	0	0	23	3
5	B	831	0	0	21	3
All	All	11374	0	9621	98	6

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 5.

The worst 5 of 98 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)
3:B:1203:FUM:C2	4:B:1204:AMP:O3P	2.01	1.08
1:B:300:ALA:HB3	1:B:301:PRO:O	1.55	1.06
1:A:296:ASP:OD2	5:A:1301:HOH:O	1.80	0.97
1:A:384:GLN:OE1	5:A:1302:HOH:O	1.89	0.91
3:B:1203:FUM:O1	4:B:1204:AMP:O3P	1.89	0.90

The worst 5 of 6 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)
5:B:1487:HOH:O	5:B:1526:HOH:O[2_454]	1.89	0.31
5:A:1440:HOH:O	5:A:1866:HOH:O[2_455]	1.95	0.25
5:B:1859:HOH:O	5:B:2086:HOH:O[2_454]	2.04	0.16
5:B:1796:HOH:O	5:B:1827:HOH:O[2_444]	2.14	0.06
5:A:1464:HOH:O	5:A:1498:HOH:O[2_445]	2.16	0.04

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	627/1188 (53%)	620 (99%)	6 (1%)	1 (0%)	49 30
1	B	628/1188 (53%)	618 (98%)	8 (1%)	2 (0%)	43 25
All	All	1255/2376 (53%)	1238 (99%)	14 (1%)	3 (0%)	49 30

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	380	ASP
1	B	300	ALA
1	B	301	PRO

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	503/947 (53%)	494 (98%)	9 (2%)	62 44
1	B	505/947 (53%)	493 (98%)	12 (2%)	52 31
All	All	1008/1894 (53%)	987 (98%)	21 (2%)	56 37

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	137	TYR
1	B	266	SER
1	B	376	ILE
1	A	577	LYS
1	B	415	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 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	FUM	A	1203	-	1,7,7	0.46	0	0,8,8	0.00	-
4	AMP	A	1204	-	22,25,25	1.04	2 (9%)	23,38,38	1.26	2 (8%)
3	FUM	B	1203	-	1,7,7	0.27	0	0,8,8	0.00	-
4	AMP	B	1204	-	22,25,25	0.85	2 (9%)	23,38,38	1.37	4 (17%)

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	FUM	A	1203	-	-	0/0/5/5	0/0/0/0
4	AMP	A	1204	-	-	0/6/26/26	0/3/3/3
3	FUM	B	1203	-	-	0/0/5/5	0/0/0/0
4	AMP	B	1204	-	-	0/6/26/26	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1204	AMP	C8-N9	2.04	1.39	1.36
4	B	1204	AMP	C2-N3	2.19	1.35	1.32

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1204	AMP	C8-N9	2.56	1.40	1.36
4	A	1204	AMP	C2-N3	2.87	1.36	1.32

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1204	AMP	O3P-P-O5'	-2.98	98.81	106.73
4	B	1204	AMP	C4-C5-N7	2.02	111.36	109.41
4	B	1204	AMP	O2P-P-O1P	2.21	119.23	110.60
4	A	1204	AMP	O3P-P-O2P	2.62	117.95	107.59
4	A	1204	AMP	N3-C2-N1	3.02	131.45	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1203	FUM	3	0
4	B	1204	AMP	3	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 [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	631/1188 (53%)	-0.18	8 (1%) 77 81	14, 22, 38, 56	0
1	B	632/1188 (53%)	-0.17	9 (1%) 75 80	14, 23, 40, 71	0
All	All	1263/2376 (53%)	-0.18	17 (1%) 77 81	14, 22, 40, 71	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	299	SER	5.9
1	B	298	ASP	5.1
1	B	266	SER	4.5
1	A	300	ALA	4.4
1	B	300	ALA	4.3

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(\AA^2)	Q<0.9
3	FUM	B	1203	8/8	0.50	0.44	24,27,31,34	8
3	FUM	A	1203	8/8	0.58	0.28	19,24,25,29	8
4	AMP	A	1204	23/23	0.94	0.09	13,17,25,29	0
4	AMP	B	1204	23/23	0.95	0.09	14,18,25,30	0
2	CA	A	1202	1/1	0.97	0.05	30,30,30,30	0
2	CA	B	1201	1/1	0.97	0.05	28,28,28,28	0
2	CA	B	1202	1/1	0.99	0.03	23,23,23,23	0
2	CA	A	1201	1/1	1.00	0.07	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.