



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:12 PM GMT

PDB ID : 1SJA
Title : X-ray structure of o-Succinylbenzoate Synthase complexed with N-acetylmethionine
Authors : Thoden, J.B.; Taylor-Ringia, E.A.; Garrett, J.B.; Gerlt, J.A.; Holden, H.M.; Rayment, I.
Deposited on : 2004-03-03
Resolution : 2.30 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
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

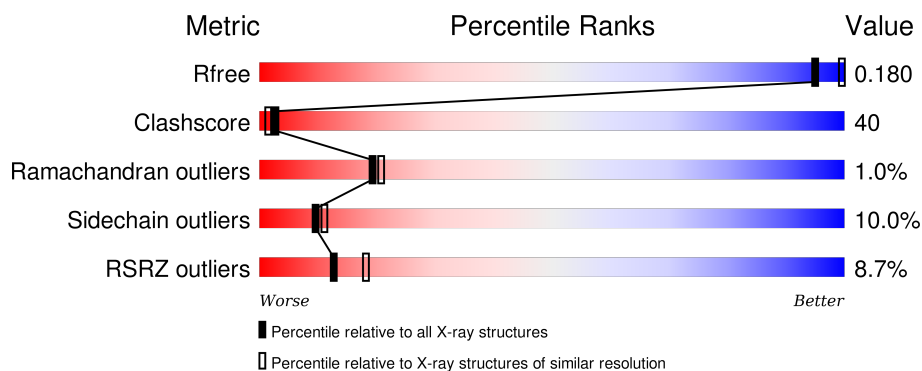
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	368	<div> <div>12%</div> <div>37%</div> <div>52%</div> <div>10%</div> </div>
1	B	368	<div> <div>2%</div> <div>48%</div> <div>42%</div> <div>10%</div> </div>
1	C	368	<div> <div>7%</div> <div>43%</div> <div>45%</div> <div>11%</div> </div>
1	D	368	<div> <div>14%</div> <div>39%</div> <div>48%</div> <div>11%</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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	AME	A	600	-	-	X	X
3	AME	B	700	-	-	X	X
3	AME	C	800	-	-	X	X
3	AME	D	900	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11704 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.

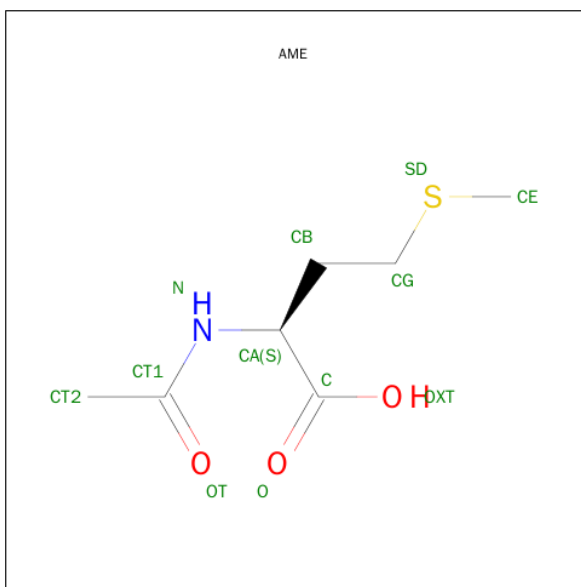
- Molecule 1 is a protein called N-acylamino acid racemase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	368	Total	C	N	O	S	0	2	0
			2785	1769	487	517	12			
1	B	368	Total	C	N	O	S	0	3	0
			2796	1775	487	522	12			
1	C	368	Total	C	N	O	S	0	0	0
			2771	1761	481	517	12			
1	D	368	Total	C	N	O	S	0	1	0
			2778	1765	484	517	12			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

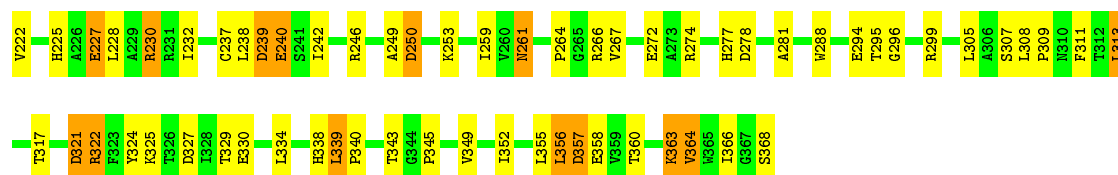
- Molecule 3 is N-ACETYLMETHIONINE (three-letter code: AME) (formula: C₇H₁₃NO₃S).



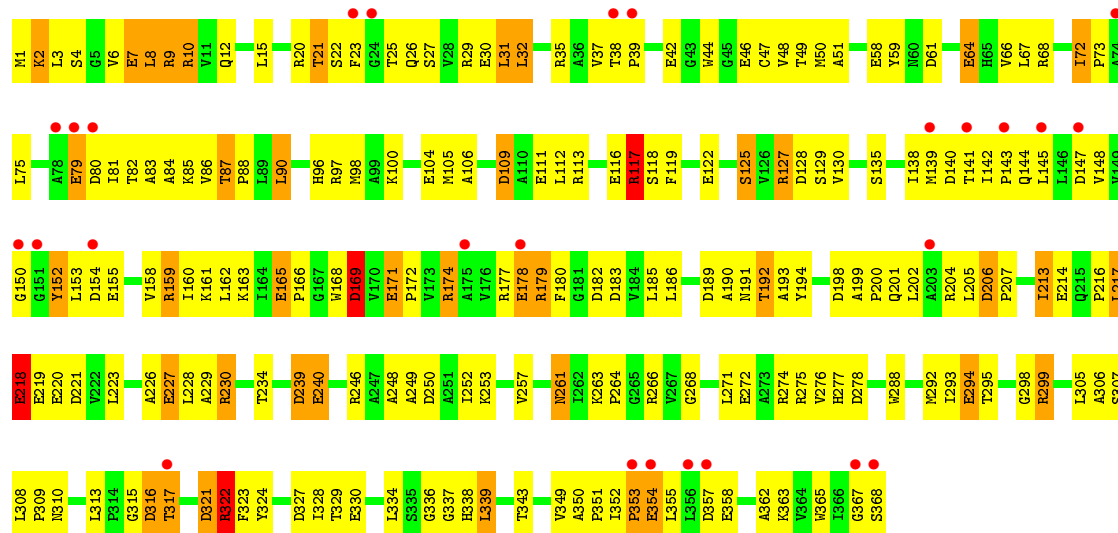
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 12	C 7	N 1	O 3	S 1	0	0
3	B	1	Total 12	C 7	N 1	O 3	S 1	0	0
3	C	1	Total 12	C 7	N 1	O 3	S 1	0	0
3	D	1	Total 12	C 7	N 1	O 3	S 1	0	0

- Molecule 4 is water.

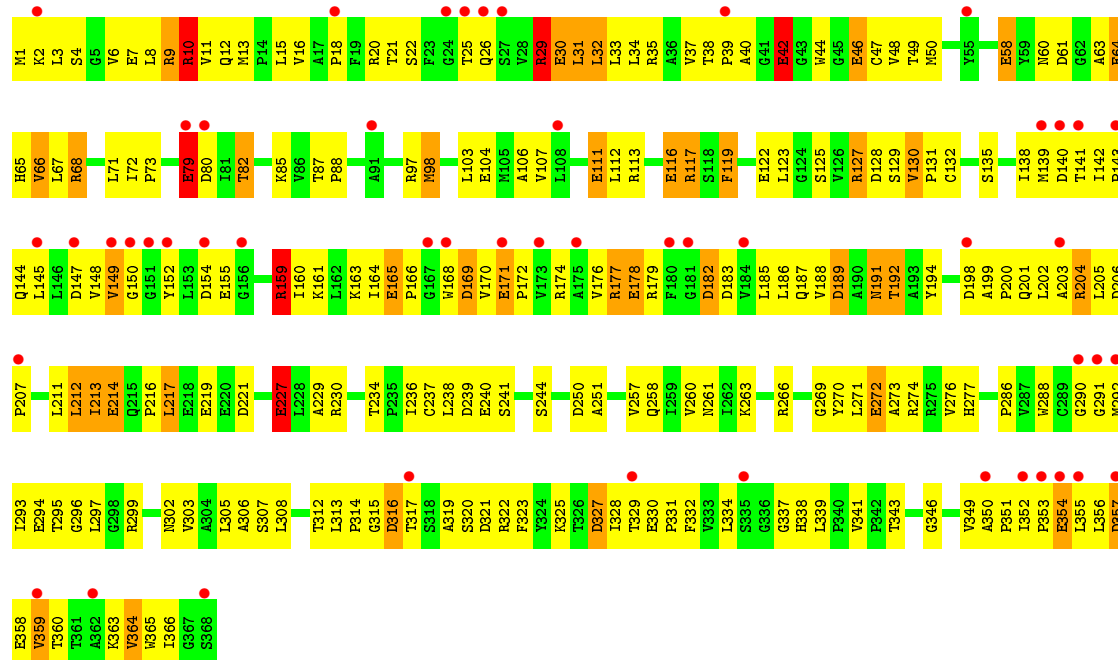
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	135	Total O 135 135	0	0
4	B	153	Total O 153 153	0	0
4	C	117	Total O 117 117	0	0
4	D	117	Total O 117 117	0	0



• Molecule 1: N-acylamino acid racemase



• Molecule 1: N-acylamino acid racemase



4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	215.30Å 215.30Å 259.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.30 29.86 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.2 (20.00-2.30) 97.7 (29.86-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.24Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.211 , 0.280 0.178 , 0.180	Depositor DCC
R_{free} test set	10015 reflections (11.14%)	DCC
Wilson B-factor (Å ²)	13.7	Xtriage
Anisotropy	0.232	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 395.2	EDS
Estimated twinning fraction	0.217 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.216 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.207 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.13$	Xtriage
Outliers	0 of 106040 reflections	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	11704	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.29% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, AME

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	1.00	25/2849 (0.9%)	1.35	41/3878 (1.1%)
1	B	1.04	26/2868 (0.9%)	1.40	41/3902 (1.1%)
1	C	1.03	23/2827 (0.8%)	1.44	41/3850 (1.1%)
1	D	1.00	23/2838 (0.8%)	1.43	44/3864 (1.1%)
All	All	1.02	97/11382 (0.9%)	1.40	167/15494 (1.1%)

The worst 5 of 97 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	165	GLU	CD-OE2	8.79	1.35	1.25
1	D	178	GLU	CD-OE2	7.80	1.34	1.25
1	D	79	GLU	CD-OE2	7.77	1.34	1.25
1	A	116	GLU	CD-OE2	7.66	1.34	1.25
1	C	171	GLU	CD-OE2	7.62	1.34	1.25

The worst 5 of 167 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	299	ARG	NE-CZ-NH1	11.01	125.80	120.30
1	D	80	ASP	CB-CG-OD2	-9.03	110.17	118.30
1	C	147	ASP	CB-CG-OD2	-8.31	110.82	118.30
1	B	206	ASP	CB-CG-OD2	-8.29	110.84	118.30
1	C	321	ASP	CB-CG-OD2	-8.20	110.92	118.30

There are no chirality outliers.

There are no planarity outliers.

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	2785	0	2828	268	0
1	B	2796	0	2834	197	0
1	C	2771	0	2810	220	0
1	D	2778	0	2819	226	2
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	12	0	12	10	0
3	B	12	0	12	8	0
3	C	12	0	12	17	0
3	D	12	0	12	13	0
4	A	135	0	0	17	0
4	B	153	0	0	12	0
4	C	117	0	0	12	0
4	D	117	0	0	16	0
All	All	11704	0	11339	910	2

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 40.

The worst 5 of 910 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:A:117:ARG:NH1	1:A:121:ALA:HB1	1.41	1.34
1:A:58:GLU:HB2	1:A:98:MET:CE	1.67	1.23
1:C:38:THR:HB	1:C:39:PRO:HD2	1.17	1.15
1:C:35:ARG:NH1	1:C:42:GLU:OE2	1.80	1.14
1:C:141:THR:CG2	1:C:143:PRO:HD2	1.76	1.13

All (2) 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:D:325:LYS:CD	1:D:325:LYS:CD[6_556]	1.08	1.12
1:D:325:LYS:CD	1:D:325:LYS:CE[6_556]	1.84	0.36

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	368/368 (100%)	334 (91%)	29 (8%)	5 (1%)	14	13
1	B	370/368 (100%)	343 (93%)	26 (7%)	1 (0%)	46	57
1	C	366/368 (100%)	337 (92%)	25 (7%)	4 (1%)	17	18
1	D	367/368 (100%)	315 (86%)	47 (13%)	5 (1%)	14	13
All	All	1471/1472 (100%)	1329 (90%)	127 (9%)	15 (1%)	19	21

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	317	THR
1	C	218	GLU
1	A	191	ASN
1	C	306	ALA
1	D	212	LEU

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	293/291 (101%)	264 (90%)	29 (10%)	10	11
1	B	295/291 (101%)	264 (90%)	31 (10%)	8	9
1	C	291/291 (100%)	261 (90%)	30 (10%)	9	10
1	D	292/291 (100%)	261 (89%)	31 (11%)	8	9
All	All	1171/1164 (101%)	1050 (90%)	121 (10%)	9	10

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	325	LYS
1	C	31	LEU
1	D	177	ARG
1	B	356	LEU
1	C	8	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	12	GLN
1	C	65	HIS
1	D	65	HIS
1	B	69	HIS
1	C	338	HIS

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

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	AME	A	600	2	8,11,11	1.63	2 (25%)	6,13,13	1.28	1 (16%)
3	AME	B	700	2	8,11,11	1.65	2 (25%)	6,13,13	1.74	2 (33%)
3	AME	C	800	2	8,11,11	1.66	2 (25%)	6,13,13	1.34	1 (16%)
3	AME	D	900	2	8,11,11	1.63	2 (25%)	6,13,13	1.45	2 (33%)

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	AME	A	600	2	-	0/8/12/12	0/0/0/0
3	AME	B	700	2	-	0/8/12/12	0/0/0/0
3	AME	C	800	2	-	0/8/12/12	0/0/0/0
3	AME	D	900	2	-	0/8/12/12	0/0/0/0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	900	AME	CT1-N	3.13	1.46	1.34
3	D	900	AME	OT-CT1	3.15	1.30	1.23
3	B	700	AME	CT1-N	3.15	1.46	1.34
3	A	600	AME	CT1-N	3.15	1.46	1.34
3	A	600	AME	OT-CT1	3.17	1.30	1.23

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	900	AME	OT-CT1-CT2	-2.28	117.88	122.06
3	A	600	AME	OT-CT1-CT2	-2.25	117.94	122.06

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	700	AME	OT-CT1-CT2	-2.08	118.24	122.06
3	C	800	AME	CG-CB-CA	-2.06	107.02	113.06
3	D	900	AME	CT2-CT1-N	2.17	120.26	116.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 48 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	600	AME	10	0
3	B	700	AME	8	0
3	C	800	AME	17	0
3	D	900	AME	13	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	368/368 (100%)	0.67	43 (11%) 6 10	23, 55, 88, 100	0
1	B	368/368 (100%)	0.23	9 (2%) 62 71	20, 40, 77, 100	0
1	C	368/368 (100%)	0.48	26 (7%) 19 26	24, 48, 84, 100	0
1	D	368/368 (100%)	0.79	50 (13%) 4 6	24, 58, 93, 100	0
All	All	1472/1472 (100%)	0.54	128 (8%) 13 18	20, 50, 87, 100	0

The worst 5 of 128 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	80	ASP	5.5
1	A	141	THR	5.3
1	D	24	GLY	5.2
1	C	143	PRO	5.0
1	D	359	VAL	5.0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	AME	A	600	12/12	0.88	0.33	3.44	62,88,100,100	0
3	AME	B	700	12/12	0.86	0.23	3.39	36,63,80,97	0
3	AME	C	800	12/12	0.93	0.22	2.34	26,97,100,100	0
3	AME	D	900	12/12	0.84	0.26	1.23	40,83,100,100	0
2	MG	A	601	1/1	0.98	0.13	-1.15	41,41,41,41	0
2	MG	D	901	1/1	0.99	0.12	-1.23	54,54,54,54	0
2	MG	B	701	1/1	0.97	0.13	-1.50	35,35,35,35	0
2	MG	C	801	1/1	0.94	0.10	-1.77	49,49,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.