wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 11:15 AM EDT

PDB ID	:	3KRA
Title	:	Mint heterotetrameric geranyl pyrophosphate synthase in complex with magnesium
Authors	:	Chang, T.-H.; Ko, T.-P.; Hsieh, F.-L.; Wang, A.H.-J.
Deposited on	:	2009-11-18
Resolution	:	1.90 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

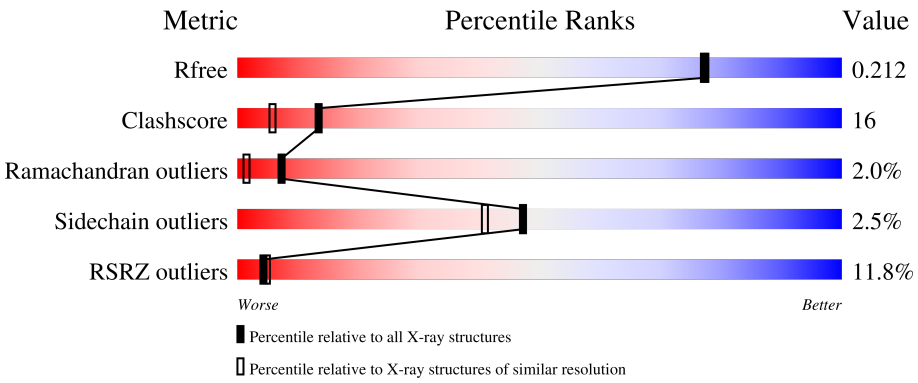
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 of 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	295	5%; 83%; 9%; • 5%
1	D	295	16%; 64%; 26%; • 6%
2	B	274	13%; 77%; 16%; •• 5%
2	C	274	11%; 73%; 23%; • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9540 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 Geranyl diphosphate synthase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	279	2120	1338	372	392	18	0	0	0
1	D	277	2102	1326	368	390	18	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP Q9SBR3
D	1	MET	-	EXPRESSION TAG	UNP Q9SBR3

- Molecule 2 is a protein called Geranyl diphosphate synthase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	259	1954	1233	339	368	14	0	0	0
2	C	271	2056	1297	361	384	14	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	EXPRESSION TAG	UNP Q9SBR4
B	267	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	268	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	269	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	270	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	271	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	272	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	273	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	274	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	1	MET	-	EXPRESSION TAG	UNP Q9SBR4

Continued on next page...

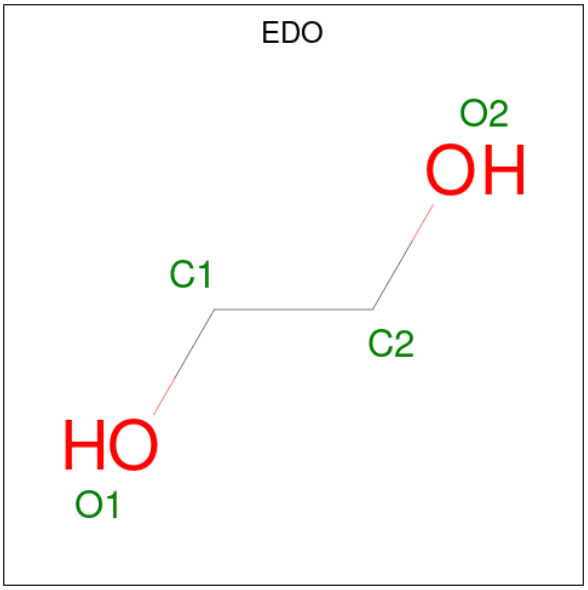
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	267	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	268	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	269	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	270	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	271	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	272	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	273	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	274	HIS	-	EXPRESSION TAG	UNP Q9SBR4

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	C	1	4	2	2	0	0

- Molecule 5 is water.

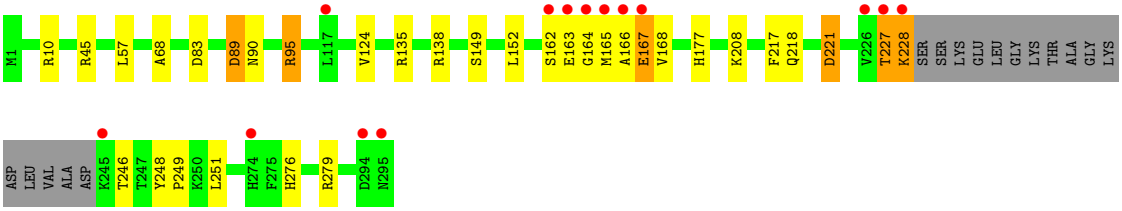
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	383	383	383	0	0
			Total	O		
5	B	319	319	319	0	0
			Total	O		
5	C	287	287	287	0	0
			Total	O		
5	D	312	312	312	0	0

3 Residue-property plots (i)

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

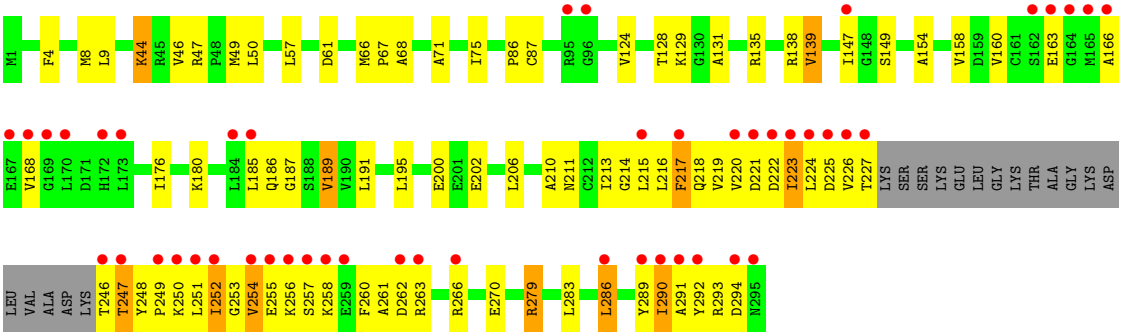
• Molecule 1: Geranyl diphosphate synthase large subunit

Chain A:



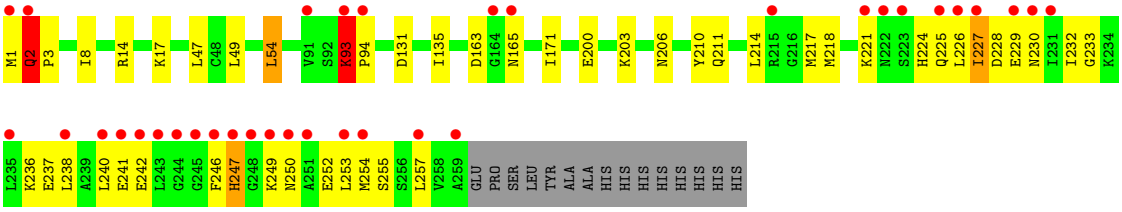
• Molecule 1: Geranyl diphosphate synthase large subunit

Chain D:



• Molecule 2: Geranyl diphosphate synthase small subunit

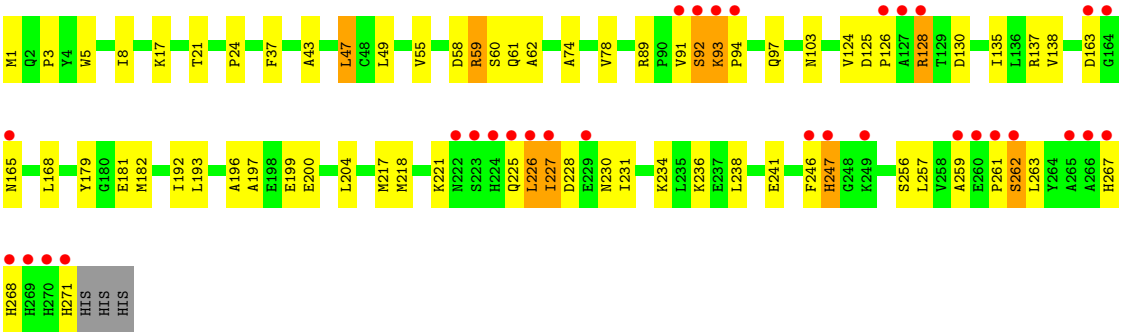
Chain B:



• Molecule 2: Geranyl diphosphate synthase small subunit

Chain C:

11%
73%
23%


M1 Q2 P3 Y4 W5 I8 K17 T21 P24 F37 A43 L47 C48 L49 V55 D58 R59 S60 Q61 A62 A74 V78 R89 P90 V91 S92 K93 P94 Q97 N103 V124 D125 P126 A127 R128 T129 D130 I135 L136 R137 V138 D163 G164
N165 L168 Y179 G180 E181 M182 I192 L193 A196 A197 E198 E199 E200 L204 M217 M218 K221 N222 S223 H224 Q225 L226 I227 D228 E229 N230 I231 K234 L235 K236 E237 L238 E241 F246 H247 G248 K249 S256 L257 V258 A259 E260 P261 S262 L263 Y264 A265 A266 H267
H268 H269 H270 H271 HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.28Å 109.25Å 182.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 29.84 – 1.90	Depositor EDS
% Data completeness (in resolution range)	93.4 (30.00-1.90) 93.2 (29.84-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.99 (at 1.89Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.176 , 0.216 0.171 , 0.212	Depositor DCC
R_{free} test set	4074 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	23.5	Xtriage
Anisotropy	0.622	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 55.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9540	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.*

[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 (i)

5.1 Standard geometry (i)

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

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	0.97	0/2155	0.87	1/2903 (0.0%)
1	D	0.88	1/2137 (0.0%)	0.85	3/2881 (0.1%)
2	B	0.81	0/1993	0.83	1/2695 (0.0%)
2	C	0.82	1/2102 (0.0%)	0.80	2/2845 (0.1%)
All	All	0.87	2/8387 (0.0%)	0.84	7/11324 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	138	VAL	CB-CG2	5.42	1.64	1.52
1	D	189	VAL	CB-CG1	-5.06	1.42	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	93	LYS	C-N-CD	5.73	140.44	128.40
1	D	286	LEU	CA-CB-CG	-5.71	102.16	115.30
1	A	45	ARG	NE-CZ-NH2	-5.44	117.58	120.30
1	D	61	ASP	CB-CG-OD2	-5.34	113.49	118.30
2	C	182	MET	CG-SD-CE	5.31	108.70	100.20

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	2120	0	2146	35	0
1	D	2102	0	2120	88	0
2	B	1954	0	1948	75	0
2	C	2056	0	2031	69	0
3	A	2	0	0	0	0
3	D	1	0	0	0	0
4	C	4	0	6	0	0
5	A	383	0	0	10	1
5	B	319	0	0	8	0
5	C	287	0	0	13	1
5	D	312	0	0	10	1
All	All	9540	0	8251	266	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 16.

The worst 5 of 266 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:138:ARG:HD2	5:A:862:HOH:O	1.39	1.20
2:B:200:GLU:HG3	5:B:890:HOH:O	1.46	1.16
1:D:286:LEU:HD13	5:D:1297:HOH:O	1.46	1.11
2:C:236:LYS:HD2	5:C:1191:HOH:O	1.53	1.07
2:B:54:LEU:HG	2:B:250:ASN:HB3	1.41	0.99

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)
5:D:1138:HOH:O	5:D:1229:HOH:O[1_455]	2.16	0.04
5:A:861:HOH:O	5:C:297:HOH:O[4_555]	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	275/295 (93%)	267 (97%)	5 (2%)	3 (1%)	14	5
1	D	273/295 (92%)	250 (92%)	16 (6%)	7 (3%)	5	1
2	B	257/274 (94%)	243 (95%)	9 (4%)	5 (2%)	8	1
2	C	269/274 (98%)	249 (93%)	13 (5%)	7 (3%)	5	1
All	All	1074/1138 (94%)	1009 (94%)	43 (4%)	22 (2%)	7	1

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	93	LYS
2	B	227	ILE
2	B	247	HIS
2	C	227	ILE
2	C	247	HIS

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	222/234 (95%)	216 (97%)	6 (3%)	44	38
1	D	220/234 (94%)	213 (97%)	7 (3%)	39	30
2	B	201/214 (94%)	197 (98%)	4 (2%)	55	51
2	C	211/214 (99%)	207 (98%)	4 (2%)	57	53
All	All	854/896 (95%)	833 (98%)	21 (2%)	47	41

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

Mol	Chain	Res	Type
1	D	9	LEU
1	D	139	VAL
1	D	279	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	263	ARG
1	D	86	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	230	ASN
2	C	250	ASN
1	D	218	GLN
1	D	178	HIS
1	D	211	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
4	EDO	C	2001	-	3,3,3	0.88	0	2,2,2	0.50	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
4	EDO	C	2001	-	-	0/1/1/1	-

There are no bond length outliers.

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	279/295 (94%)	-0.07	14 (5%)	28	32	15, 23, 44, 78	0
1	D	277/295 (93%)	0.68	48 (17%)	1	1	16, 28, 90, 100	0
2	B	259/274 (94%)	0.47	35 (13%)	3	3	18, 28, 103, 117	0
2	C	271/274 (98%)	0.51	31 (11%)	5	5	16, 31, 72, 101	0
All	All	1086/1138 (95%)	0.39	128 (11%)	4	5	15, 28, 89, 117	0

The worst 5 of 128 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	226	VAL	11.1
1	D	165	MET	10.3
2	B	246	PHE	9.3
1	D	224	LEU	9.3
1	D	166	ALA	9.2

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 monosaccharides 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	MG	D	3001	1/1	0.75	0.10	69,69,69,69	0
3	MG	A	3003	1/1	0.82	0.28	58,58,58,58	0
3	MG	A	3002	1/1	0.97	0.12	37,37,37,37	0
4	EDO	C	2001	4/4	0.98	0.17	16,17,19,19	0

6.5 Other polymers (i)

There are no such residues in this entry.