Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 11:26 AM GMT

PDB ID	:	3OYR
Title	:	Crystal structure of polyprenyl synthase from Caulobacter crescentus CB15 complexed with calcium and isoprenyl diphosphate
Authors	:	Patskovsky, Y.; Toro, R.; Rutter, M.; Sauder, J.M.; Burley, S.K.; Poulter, C.D.; Gerlt, J.A.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2010-09-23
Resolution	:	2.00 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

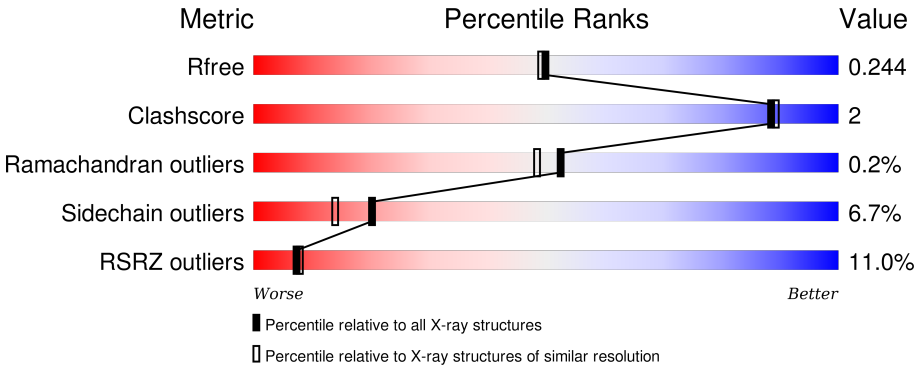
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 <=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	345	13% 75% 9% • 15%
1	B	345	7% 88% 7% • 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4996 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.

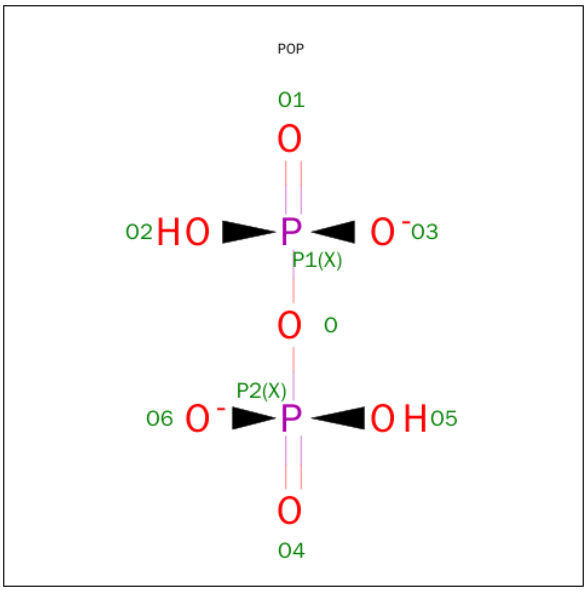
- Molecule 1 is a protein called Trans-isoprenyl diphosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	8	0
			2228	1410	393	417	8			
1	B	328	Total	C	N	O	S	0	8	0
			2518	1585	452	473	8			

There are 18 discrepancies between the modelled and reference sequences:

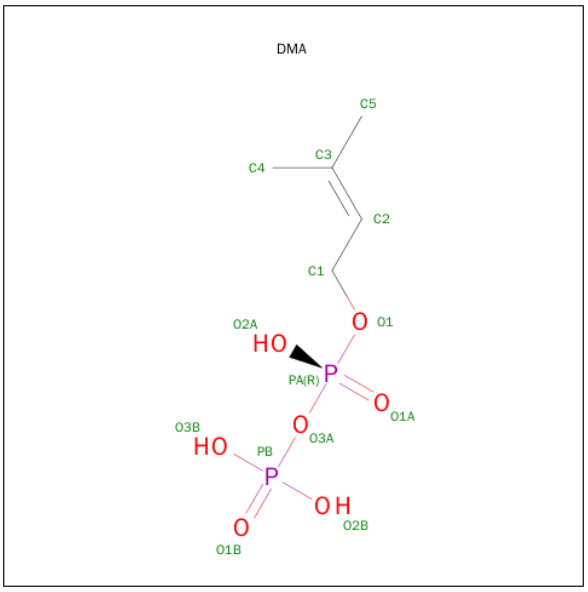
Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	EXPRESSION TAG	UNP Q9A6I1
A	-8	ALA	-	EXPRESSION TAG	UNP Q9A6I1
A	-7	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-6	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-5	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-4	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-3	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-2	HIS	-	EXPRESSION TAG	UNP Q9A6I1
A	-1	SER	-	EXPRESSION TAG	UNP Q9A6I1
B	-9	MET	-	EXPRESSION TAG	UNP Q9A6I1
B	-8	ALA	-	EXPRESSION TAG	UNP Q9A6I1
B	-7	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-6	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-5	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-4	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-3	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-2	HIS	-	EXPRESSION TAG	UNP Q9A6I1
B	-1	SER	-	EXPRESSION TAG	UNP Q9A6I1

- Molecule 2 is PYROPHOSPHATE 2- (three-letter code: POP) (formula: $H_2O_7P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 9 7 2	0	0

- Molecule 3 is DIMETHYLALLYL DIPHOSPHATE (three-letter code: DMA) (formula: $C_5H_{12}O_7P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O P 14 5 7 2	0	0
3	B	1	Total C O P 14 5 7 2	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 5 is water.

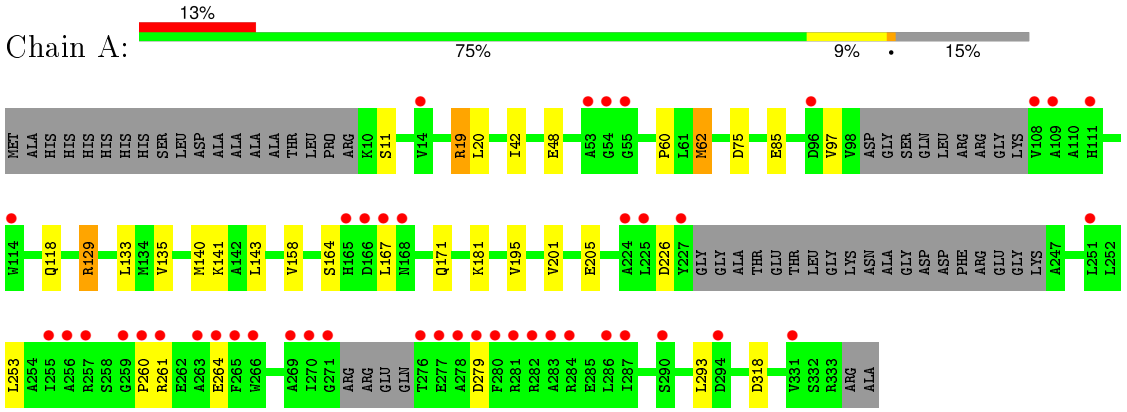
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	59	Total O 59 59	0	0
5	B	152	Total O 152 152	0	0

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 errors 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.

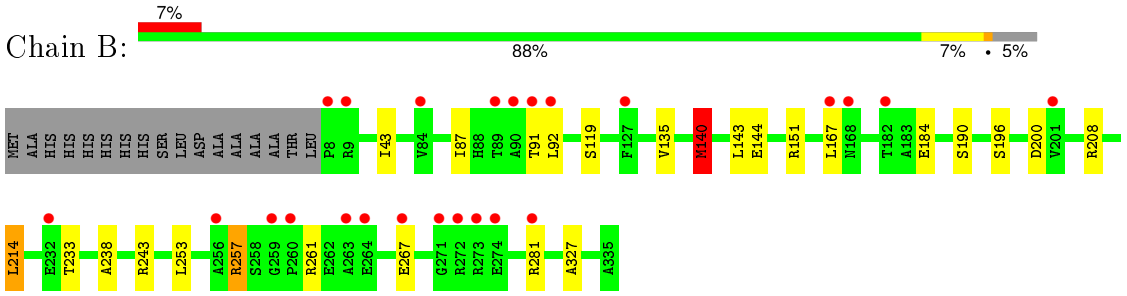
- Molecule 1: Trans-isoprenyl diphosphate synthase

Chain A:



- Molecule 1: Trans-isoprenyl diphosphate synthase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.89Å 72.44Å 125.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.00 36.36 – 2.00	Depositor EDS
% Data completeness (in resolution range)	90.0 (40.00-2.00) 90.1 (36.36-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.197 , 0.247 0.197 , 0.244	Depositor DCC
R_{free} test set	1283 reflections (3.22%)	DCC
Wilson B-factor (Å^2)	34.1	Xtriage
Anisotropy	0.496	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 60.8	EDS
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	2 of 41152 reflections (0.005%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4996	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.10% 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, 0.375 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: DMA, CA, POP

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	0.50	0/2279	0.65	0/3081
1	B	0.52	0/2576	0.64	1/3478 (0.0%)
All	All	0.51	0/4855	0.64	1/6559 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	140	MET	CG-SD-CE	5.11	108.38	100.20

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	2228	0	2257	16	0
1	B	2518	0	2560	13	0
2	A	9	0	0	0	0
3	B	28	0	18	1	0
4	B	2	0	0	0	0
5	A	59	0	0	0	0
5	B	152	0	0	0	0
All	All	4996	0	4835	24	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 2.

All (24) 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:19[A]:ARG:HH11	1:A:19[A]:ARG:CG	1.65	1.08
1:A:19[A]:ARG:HH11	1:A:19[A]:ARG:HG3	0.90	1.06
1:A:19[A]:ARG:NH1	1:A:19[A]:ARG:HG3	1.68	1.00
1:A:135:VAL:HG11	1:B:144:GLU:HB2	1.68	0.76
1:A:140:MET:HE1	1:B:135:VAL:HG22	1.70	0.74
1:A:140:MET:CE	1:B:135:VAL:HG22	2.23	0.69
1:B:140:MET:HA	1:B:140:MET:CE	2.27	0.65
1:B:184:GLU:OE2	1:B:208:ARG:NH2	2.29	0.64
1:B:253:LEU:O	1:B:257:ARG:HG2	2.06	0.56
1:B:140:MET:HA	1:B:140:MET:HE2	1.88	0.56
1:A:141:LYS:HB3	1:A:195[B]:VAL:HG11	1.89	0.53
1:A:261:ARG:HA	1:A:264:GLU:HB2	1.92	0.51
1:A:19[A]:ARG:NH1	1:A:19[A]:ARG:CG	2.39	0.49
1:B:87:ILE:O	1:B:91:THR:HG23	2.14	0.48
1:B:92:LEU:HD23	3:B:339:DMA:H52	1.95	0.48
1:A:60:PRO:HG3	1:A:85:GLU:HB2	1.96	0.47
1:A:158:VAL:HG11	1:B:43:ILE:HG13	2.00	0.44
1:A:129:ARG:HD3	1:B:151:ARG:HG3	2.00	0.43
1:A:20:LEU:HD11	1:A:318:ASP:HB3	2.00	0.43
1:B:214:LEU:HD12	1:B:327:ALA:HA	2.00	0.42
1:A:62[B]:MET:HA	1:A:62[B]:MET:HE3	2.02	0.41
1:B:238:ALA:O	1:B:243:ARG:NH2	2.54	0.41
1:A:171:GLN:HE21	1:A:253:LEU:HD21	1.86	0.41
1:A:201:VAL:O	1:A:205:GLU:HB2	2.22	0.40

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	292/345 (85%)	282 (97%)	9 (3%)	1 (0%)	46 41
1	B	334/345 (97%)	330 (99%)	4 (1%)	0	100 100
All	All	626/690 (91%)	612 (98%)	13 (2%)	1 (0%)	52 48

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	260	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	219/253 (87%)	200 (91%)	19 (9%)	13 7
1	B	248/253 (98%)	234 (94%)	14 (6%)	26 20
All	All	467/506 (92%)	434 (93%)	33 (7%)	20 12

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	SER
1	A	19[A]	ARG
1	A	19[B]	ARG
1	A	42	ILE
1	A	48	GLU
1	A	62[A]	MET
1	A	62[B]	MET
1	A	75	ASP
1	A	97	VAL
1	A	118	GLN
1	A	129	ARG
1	A	133	LEU
1	A	143	LEU
1	A	164	SER
1	A	167	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	181	LYS
1	A	226	ASP
1	A	279	ASP
1	A	293	LEU
1	B	119	SER
1	B	140	MET
1	B	143	LEU
1	B	167	LEU
1	B	190	SER
1	B	196	SER
1	B	200	ASP
1	B	214	LEU
1	B	233[A]	THR
1	B	233[B]	THR
1	B	257	ARG
1	B	261	ARG
1	B	267	GLU
1	B	281	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	ASN
1	A	111	HIS
1	A	171	GLN
1	B	171	GLN

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 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\#\|Z\|>2$
2	POP	A	336	-	8,8,8	0.64	0	13,13,13	1.16	1 (7%)
3	DMA	B	336	-	11,13,13	0.56	0	16,19,19	1.70	5 (31%)
3	DMA	B	339	4	11,13,13	0.63	0	16,19,19	1.46	1 (6%)

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	POP	A	336	-	-	0/6/6/6	0/0/0/0
3	DMA	B	336	-	-	0/13/13/13	0/0/0/0
3	DMA	B	339	4	-	0/13/13/13	0/0/0/0

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	336	DMA	O1-C1-C2	-3.24	95.07	109.68
2	A	336	POP	P2-O-P1	-2.86	124.70	132.73
3	B	336	DMA	C1-C2-C3	-2.74	122.22	126.71
3	B	339	DMA	C1-C2-C3	-2.52	122.57	126.71
3	B	336	DMA	O3A-PA-O1	-2.32	96.78	102.94
3	B	336	DMA	O3B-PB-O2B	2.25	115.94	107.38
3	B	336	DMA	C5-C3-C4	2.34	120.40	114.64

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
3	B	339	DMA	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, 95^{th} 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	292/345 (84%)	0.69	44 (15%) 3 3	23, 46, 122, 139	0
1	B	328/345 (95%)	0.18	24 (7%) 18 19	21, 37, 85, 112	0
All	All	620/690 (89%)	0.42	68 (10%) 7 8	21, 40, 104, 139	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	165	HIS	9.1
1	A	167	LEU	7.9
1	A	283	ALA	7.7
1	A	276	THR	7.6
1	A	271	GLY	7.1
1	A	114	TRP	6.8
1	A	280	PHE	5.9
1	B	8	PRO	5.8
1	A	278	ALA	5.6
1	A	260	PRO	5.5
1	A	263	ALA	5.3
1	A	270	ILE	5.0
1	A	269	ALA	4.9
1	A	108	VAL	4.8
1	A	279	ASP	4.4
1	A	255	ILE	4.1
1	A	277	GLU	4.1
1	A	261	ARG	4.0
1	A	286	LEU	3.9
1	A	265	PHE	3.8
1	A	224	ALA	3.8
1	A	256	ALA	3.7
1	B	271	GLY	3.6
1	A	251	LEU	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	281	ARG	3.6
1	A	111	HIS	3.6
1	A	266	TRP	3.5
1	A	225	LEU	3.5
1	A	227	TYR	3.4
1	B	273	ARG	3.4
1	A	264	GLU	3.4
1	A	294	ASP	3.4
1	A	54	GLY	3.2
1	B	256	ALA	3.2
1	B	259	GLY	3.2
1	B	264	GLU	3.2
1	A	109	ALA	3.1
1	A	259	GLY	3.1
1	A	168	ASN	3.1
1	A	55	GLY	3.0
1	A	287	ILE	2.9
1	B	9	ARG	2.9
1	B	272	ARG	2.9
1	B	267	GLU	2.8
1	B	260	PRO	2.8
1	A	282	ARG	2.8
1	A	257	ARG	2.8
1	A	281	ARG	2.8
1	B	232	GLU	2.7
1	A	53	ALA	2.7
1	B	182	THR	2.7
1	B	92	LEU	2.7
1	B	90	ALA	2.6
1	B	89	THR	2.6
1	B	274	GLU	2.5
1	B	168	ASN	2.5
1	B	127[A]	PHE	2.5
1	A	331	VAL	2.5
1	A	14	VAL	2.3
1	A	290	SER	2.3
1	B	167	LEU	2.3
1	B	91	THR	2.2
1	A	96	ASP	2.2
1	B	263	ALA	2.1
1	B	201	VAL	2.1
1	B	84	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	284	ARG	2.0
1	A	166	ASP	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. 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, 95^{th} 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(Å^2)	Q<0.9
3	DMA	B	336	14/14	0.98	0.18	0.34	24,34,40,44	0
2	POP	A	336	9/9	0.89	0.20	0.27	107,112,114,116	0
3	DMA	B	339	14/14	0.98	0.15	0.03	32,37,47,55	0
4	CA	B	337	1/1	0.99	0.11	-0.01	33,33,33,33	0
4	CA	B	338	1/1	0.99	0.10	-0.86	39,39,39,39	0

6.5 Other polymers (i)

There are no such residues in this entry.