wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:21 AM GMT

PDB ID	:	3EB2
Title	:	Crystal structure of Dihydrodipicolinate Synthase from Rhodopseudomonas palustris at 2.0A resolution
Authors	:	Satyanarayana, L.; Eswaramoorthy, S.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2008-08-26
Resolution	:	2.04 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 (i) symbol.

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

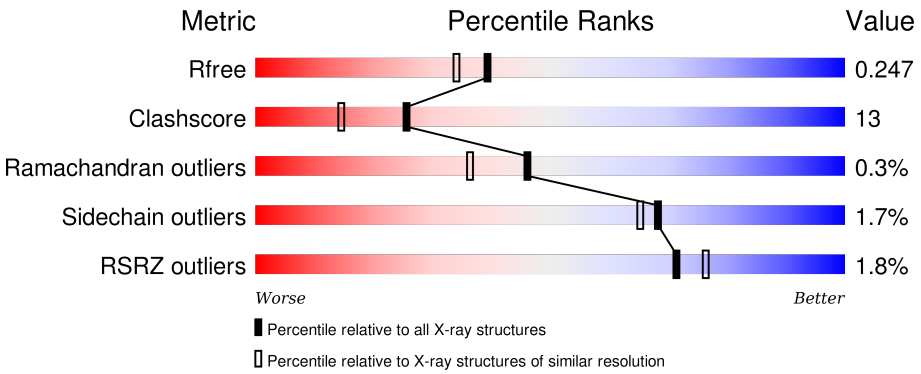
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.04 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	1192 (2.04-2.04)
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)
RSRZ outliers	91569	1194 (2.04-2.04)

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	300	75% 21% •
1	B	300	% 76% 20% • •
1	C	300	5% 74% 22% •
1	D	300	2% 75% 20% • •

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
2	PGE	A	888	-	-	-	X
2	PGE	B	999	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9399 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 Putative dihydrodipicolinate synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	288	2186	1393	384	400	9	0	0	0
			Total	C	N	O	S			
1	B	291	2205	1404	387	405	9	0	0	0
			Total	C	N	O	S			
1	C	288	2186	1393	384	400	9	0	0	0
			Total	C	N	O	S			
1	D	290	2197	1398	386	404	9	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

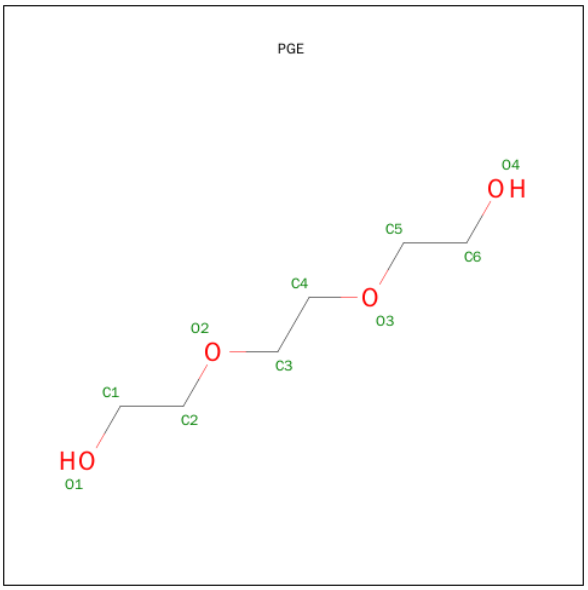
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q6N9H8
A	2	SER	-	expression tag	UNP Q6N9H8
A	3	LEU	-	expression tag	UNP Q6N9H8
A	293	GLU	-	expression tag	UNP Q6N9H8
A	294	GLY	-	expression tag	UNP Q6N9H8
A	295	HIS	-	expression tag	UNP Q6N9H8
A	296	HIS	-	expression tag	UNP Q6N9H8
A	297	HIS	-	expression tag	UNP Q6N9H8
A	298	HIS	-	expression tag	UNP Q6N9H8
A	299	HIS	-	expression tag	UNP Q6N9H8
A	300	HIS	-	expression tag	UNP Q6N9H8
B	1	MET	-	expression tag	UNP Q6N9H8
B	2	SER	-	expression tag	UNP Q6N9H8
B	3	LEU	-	expression tag	UNP Q6N9H8
B	293	GLU	-	expression tag	UNP Q6N9H8
B	294	GLY	-	expression tag	UNP Q6N9H8
B	295	HIS	-	expression tag	UNP Q6N9H8
B	296	HIS	-	expression tag	UNP Q6N9H8
B	297	HIS	-	expression tag	UNP Q6N9H8
B	298	HIS	-	expression tag	UNP Q6N9H8
B	299	HIS	-	expression tag	UNP Q6N9H8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	300	HIS	-	expression tag	UNP Q6N9H8
C	1	MET	-	expression tag	UNP Q6N9H8
C	2	SER	-	expression tag	UNP Q6N9H8
C	3	LEU	-	expression tag	UNP Q6N9H8
C	293	GLU	-	expression tag	UNP Q6N9H8
C	294	GLY	-	expression tag	UNP Q6N9H8
C	295	HIS	-	expression tag	UNP Q6N9H8
C	296	HIS	-	expression tag	UNP Q6N9H8
C	297	HIS	-	expression tag	UNP Q6N9H8
C	298	HIS	-	expression tag	UNP Q6N9H8
C	299	HIS	-	expression tag	UNP Q6N9H8
C	300	HIS	-	expression tag	UNP Q6N9H8
D	1	MET	-	expression tag	UNP Q6N9H8
D	2	SER	-	expression tag	UNP Q6N9H8
D	3	LEU	-	expression tag	UNP Q6N9H8
D	293	GLU	-	expression tag	UNP Q6N9H8
D	294	GLY	-	expression tag	UNP Q6N9H8
D	295	HIS	-	expression tag	UNP Q6N9H8
D	296	HIS	-	expression tag	UNP Q6N9H8
D	297	HIS	-	expression tag	UNP Q6N9H8
D	298	HIS	-	expression tag	UNP Q6N9H8
D	299	HIS	-	expression tag	UNP Q6N9H8
D	300	HIS	-	expression tag	UNP Q6N9H8

- Molecule 2 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	10	6	4	0	0
			Total	C	O		
2	B	1	10	6	4	0	0

- Molecule 3 is water.

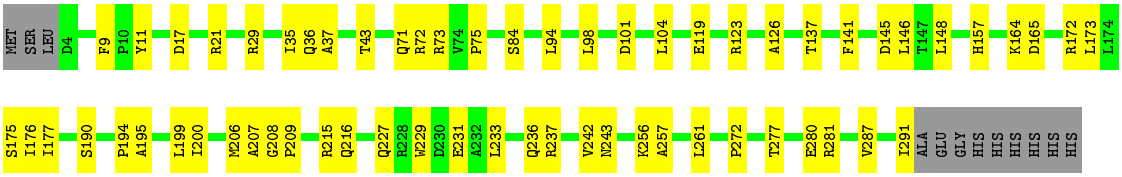
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	181	181	181	0	0
			Total	O		
3	B	154	154	154	0	0
			Total	O		
3	C	137	137	137	0	0
			Total	O		
3	D	133	133	133	0	0

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.

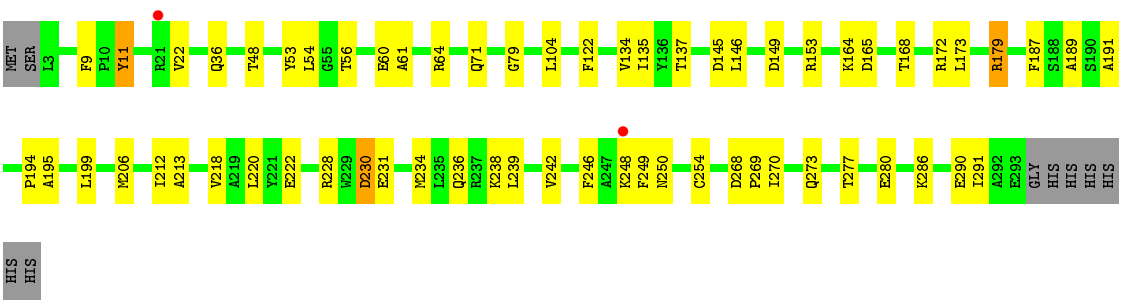
- Molecule 1: Putative dihydrodipicolinate synthetase

Chain A:



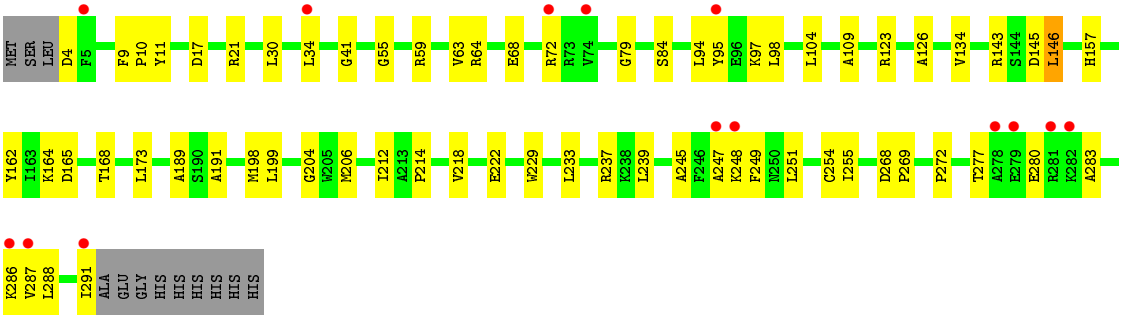
- Molecule 1: Putative dihydrodipicolinate synthetase

Chain B:



- Molecule 1: Putative dihydrodipicolinate synthetase

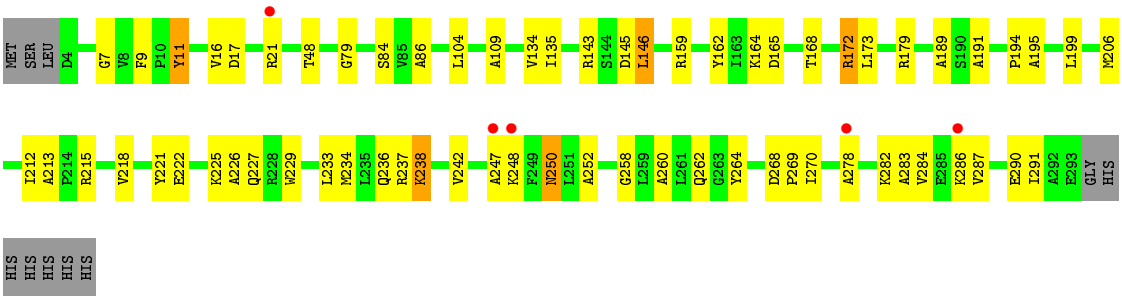
Chain C:



- Molecule 1: Putative dihydrodipicolinate synthetase

Chain D:

2%
75%
20%


MET SER LEU D4 G7 V8 F9 P10 Y11 V16 D17 R21 T48 G79 S84 V85 A86 L104 A109 V134 I135 R143 S144 D145 L146 R159 Y162 I163 K164 D165 T168 R172 L173 R179 A189 S190 A191 P194 A195 L199 M206
I212 A213 P214 R215 V218 Y221 E222 K225 A226 Q227 R228 W229 L233 M234 L235 Q236 R237 K238 V242 A247 K248 F249 N250 L251 A252 G258 L259 A260 L261 Q262 G263 Y264 D268 P269 I270 A278 K282 A283 V284 E285 K286 V287 E290 I291 A292 E293 GLY HIS
HIS HIS 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	79.75Å 120.85Å 135.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.04 39.21 – 2.04	Depositor EDS
% Data completeness (in resolution range)	95.4 (50.00-2.04) 95.5 (39.21-2.04)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.09 (at 2.05Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.212 , 0.248 0.210 , 0.247	Depositor DCC
R_{free} test set	3207 reflections (4.02%)	DCC
Wilson B-factor (Å^2)	21.7	Xtriage
Anisotropy	0.353	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 48.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 82883 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9399	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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 (i)

5.1 Standard geometry (i)

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

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	0/2227	0.59	0/3029
1	B	0.32	0/2246	0.58	0/3055
1	C	0.28	0/2227	0.56	0/3029
1	D	0.31	0/2238	0.58	0/3044
All	All	0.31	0/8938	0.58	0/12157

There are no bond length outliers.

There are no bond angle outliers.

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	2186	0	2220	71	0
1	B	2205	0	2233	53	0
1	C	2186	0	2220	48	0
1	D	2197	0	2222	54	0
2	A	10	0	14	1	0
2	B	10	0	14	0	0
3	A	181	0	0	15	0
3	B	154	0	0	9	0
3	C	137	0	0	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	133	0	0	6	0
All	All	9399	0	8923	222	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 13.

The worst 5 of 222 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:250:ASN:HD21	1:B:273:GLN:HE21	1.13	0.92
1:A:71:GLN:HB2	1:A:73:ARG:NH1	1.85	0.92
1:A:200:ILE:HG13	3:A:971:HOH:O	1.71	0.89
1:A:172:ARG:HG3	3:A:1029:HOH:O	1.76	0.85
1:A:164:LYS:HG2	3:A:1025:HOH:O	1.79	0.82

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	286/300 (95%)	282 (99%)	4 (1%)	0	100	100
1	B	289/300 (96%)	282 (98%)	7 (2%)	0	100	100
1	C	286/300 (95%)	282 (99%)	3 (1%)	1 (0%)	46	36
1	D	288/300 (96%)	284 (99%)	2 (1%)	2 (1%)	26	15
All	All	1149/1200 (96%)	1130 (98%)	16 (1%)	3 (0%)	46	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	248	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	247	ALA
1	C	247	ALA

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	221/231 (96%)	220 (100%)	1 (0%)	92	92
1	B	222/231 (96%)	218 (98%)	4 (2%)	66	62
1	C	221/231 (96%)	217 (98%)	4 (2%)	66	62
1	D	221/231 (96%)	215 (97%)	6 (3%)	52	45
All	All	885/924 (96%)	870 (98%)	15 (2%)	68	65

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	11	TYR
1	C	145	ASP
1	D	172	ARG
1	C	4	ASP
1	D	146	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	91	GLN
1	B	236	GLN
1	D	142	GLN
1	A	243	ASN
1	B	273	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)

2 ligands 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	$\#\|Z\| > 2$
2	PGE	A	888	-	9,9,9	0.60	0	8,8,8	0.92	0
2	PGE	B	999	-	9,9,9	0.60	0	8,8,8	0.91	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	PGE	A	888	-	-	0/7/7/7	0/0/0/0
2	PGE	B	999	-	-	0/7/7/7	0/0/0/0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	888	PGE	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 (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	288/300 (96%)	-0.27	0 100 100	10, 18, 29, 38	0
1	B	291/300 (97%)	-0.13	2 (0%) 89 91	13, 21, 35, 44	0
1	C	288/300 (96%)	0.29	14 (4%) 33 38	13, 27, 49, 64	0
1	D	290/300 (96%)	-0.03	5 (1%) 73 78	14, 23, 35, 43	0
All	All	1157/1200 (96%)	-0.04	21 (1%) 71 76	10, 22, 40, 64	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	282	LYS	3.8
1	C	248	LYS	3.4
1	C	279	GLU	3.3
1	D	247	ALA	3.2
1	C	74	VAL	2.8

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(Å^2)	Q<0.9
2	PGE	A	888	10/10	0.79	0.29	15.68	18,20,23,25	0
2	PGE	B	999	10/10	0.69	0.29	10.30	18,20,23,25	0

6.5 Other polymers (i)

There are no such residues in this entry.