Full wwPDB X-ray Structure Validation Report

Dec 20, 2016 – 05:41 PM EST

PDB ID	:	5F4X
Title	:	Fructose-1,6-bisphosphate aldolase K229M mutant from rabbit muscle
Authors	:	LowKam, C.; Arthus-Cartier, G.
Deposited on	:	2015-12-03
Resolution	:	1.84 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:

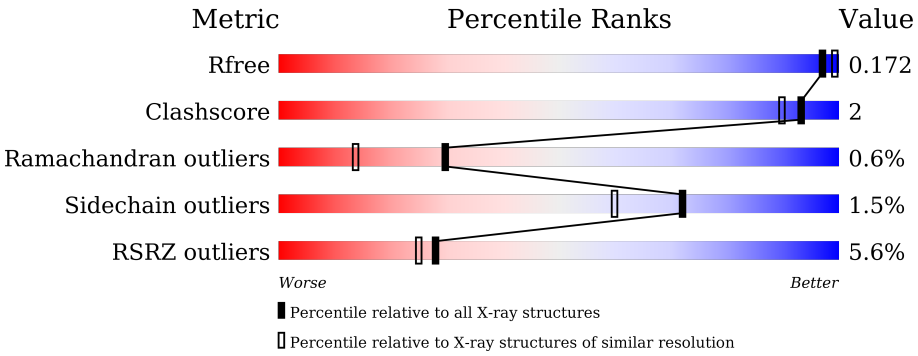
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

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 1.84 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	2634 (1.86-1.82)
Clashscore	102246	2862 (1.86-1.82)
Ramachandran outliers	100387	2831 (1.86-1.82)
Sidechain outliers	100360	2832 (1.86-1.82)
RSRZ outliers	91569	2639 (1.86-1.82)

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	363	5% 94% 5% ·
1	B	363	5% 96% · ·
1	C	363	5% 96% ·
1	D	363	7% 94% 6%

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
2	GOL	B	401[A]	-	-	-	X
2	GOL	B	401[B]	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24064 atoms, of which 11152 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.

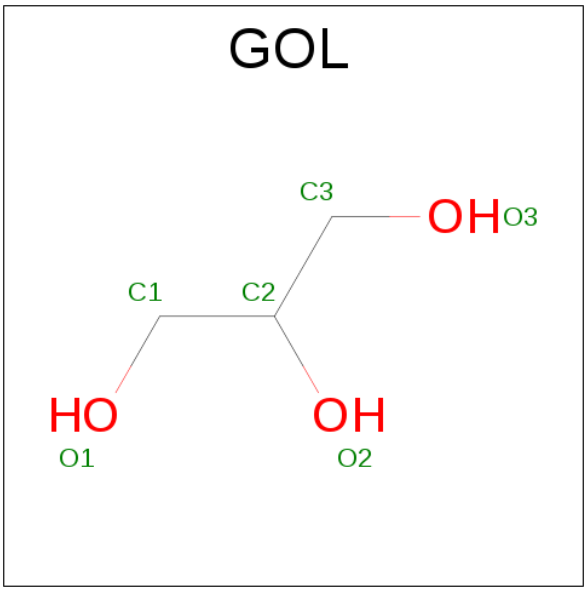
- Molecule 1 is a protein called Fructose-bisphosphate aldolase A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	363	Total	C	H	N	O	S	0	2	0
			5547	1737	2782	489	526	13			
1	B	363	Total	C	H	N	O	S	0	1	0
			5546	1737	2781	489	526	13			
1	C	363	Total	C	H	N	O	S	0	2	0
			5547	1737	2782	489	526	13			
1	D	363	Total	C	H	N	O	S	0	2	0
			5548	1737	2783	489	526	13			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	229	MET	LYS	engineered mutation	UNP P00883
B	229	MET	LYS	engineered mutation	UNP P00883
C	229	MET	LYS	engineered mutation	UNP P00883
D	229	MET	LYS	engineered mutation	UNP P00883

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C H O 28 6 16 6	0	1
2	D	1	Total C H O 14 3 8 3	0	0

- Molecule 3 is water.

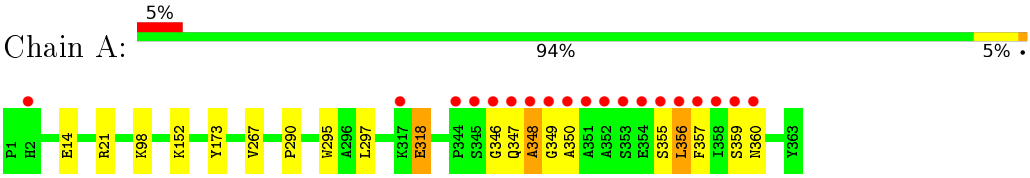
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	481	Total O 481 481	0	0
3	B	470	Total O 470 470	0	0
3	C	474	Total O 474 474	0	0
3	D	409	Total O 409 409	0	0

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.

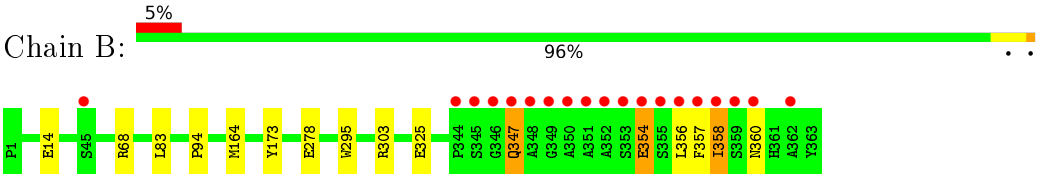
- Molecule 1: Fructose-bisphosphate aldolase A

Chain A:



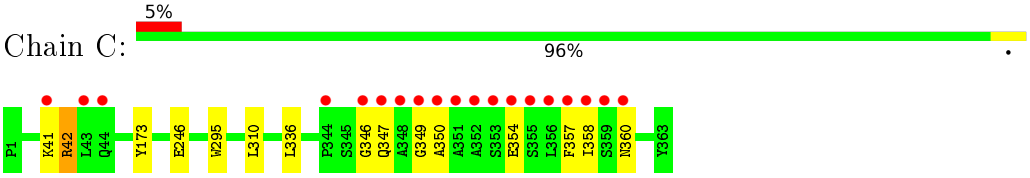
- Molecule 1: Fructose-bisphosphate aldolase A

Chain B:



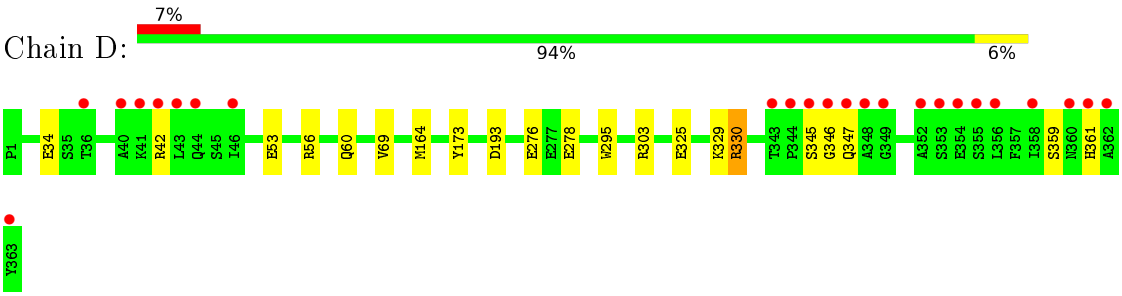
- Molecule 1: Fructose-bisphosphate aldolase A

Chain C:



- Molecule 1: Fructose-bisphosphate aldolase A

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	83.48Å 103.19Å 84.53Å 90.00° 98.66° 90.00°	Depositor
Resolution (Å)	35.20 – 1.84 48.36 – 1.84	Depositor EDS
% Data completeness (in resolution range)	99.4 (35.20-1.84) 95.9 (48.36-1.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.47 (at 1.84Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.146 , 0.174 0.145 , 0.172	Depositor DCC
R_{free} test set	1922 reflections (1.63%)	DCC
Wilson B-factor (Å^2)	17.6	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 58.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.014 for l,-k,h	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	24064	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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

5.1 Standard geometry

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

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	0.36	0/2828	0.53	0/3831
1	B	0.36	0/2819	0.54	0/3819
1	C	0.36	0/2828	0.52	0/3831
1	D	0.35	0/2828	0.50	0/3831
All	All	0.36	0/11303	0.52	0/15312

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

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	2765	2782	2775	12	0
1	B	2765	2781	2780	12	0
1	C	2765	2782	2775	5	0
1	D	2765	2783	2775	14	0
2	B	12	16	16	0	0
2	D	6	8	8	0	0
3	A	481	0	0	6	1
3	B	470	0	0	8	1
3	C	474	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	409	0	0	6	0
All	All	12912	11152	11129	43	1

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 (43) 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:D:56:ARG:NH1	3:D:501:HOH:O	1.88	1.04
1:B:278:GLU:OE1	3:B:501:HOH:O	2.08	0.71
1:D:329:LYS:NZ	3:D:503:HOH:O	2.21	0.71
1:D:42:ARG:NH1	3:D:506:HOH:O	2.28	0.67
1:A:21:ARG:NH1	3:A:405:HOH:O	2.24	0.67
1:D:60:GLN:OE1	3:D:501:HOH:O	2.13	0.67
1:B:278:GLU:OE2	3:B:502:HOH:O	2.13	0.66
1:A:318:GLU:OE1	3:A:401:HOH:O	2.13	0.65
1:B:164[A]:MET:SD	3:B:819:HOH:O	2.55	0.63
1:B:68:ARG:NH1	3:B:507:HOH:O	2.31	0.59
1:A:152:LYS:NZ	3:A:406:HOH:O	2.24	0.59
1:A:14:GLU:OE2	3:A:402:HOH:O	2.17	0.58
1:B:357:PHE:CE2	1:B:358:ILE:HG13	2.40	0.57
1:B:354:GLU:O	3:B:503:HOH:O	2.17	0.57
1:C:336:LEU:HD21	1:C:347:GLN:OE1	2.05	0.56
1:D:346:GLY:HA2	1:D:347:GLN:HB2	1.87	0.55
1:D:53:GLU:OE2	1:D:56:ARG:NH2	2.41	0.54
1:C:246:GLU:OE1	3:C:401:HOH:O	2.19	0.53
1:D:34:GLU:CD	1:D:42:ARG:HH22	2.13	0.52
1:A:98:LYS:NZ	3:A:409:HOH:O	2.40	0.50
1:B:14:GLU:OE2	3:B:504:HOH:O	2.20	0.50
1:A:360:ASN:N	3:A:411:HOH:O	2.44	0.49
1:B:357:PHE:CD2	1:B:358:ILE:HG13	2.47	0.49
1:A:290:PRO:HG3	1:A:357:PHE:HB2	1.93	0.49
1:D:325:GLU:O	1:D:329:LYS:HG3	2.13	0.49
1:B:68:ARG:NH2	1:B:325:GLU:OE2	2.41	0.48
1:B:347:GLN:NE2	3:B:513:HOH:O	2.46	0.48
1:D:42:ARG:HD3	1:D:303:ARG:HD3	1.94	0.48
1:D:164[A]:MET:SD	3:D:803:HOH:O	2.61	0.45
1:A:349:GLY:O	1:A:350:ALA:HB3	2.17	0.45
1:A:355:SER:O	1:A:356:LEU:HG	2.15	0.45
1:B:83:LEU:HD12	1:B:94:PRO:HG3	1.99	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:349:GLY:O	1:C:350:ALA:HB3	2.18	0.44
1:A:347:GLN:O	1:A:348:ALA:HB2	2.18	0.44
1:D:346:GLY:CA	1:D:347:GLN:HB2	2.49	0.43
1:A:267:VAL:HB	1:A:297:LEU:HD23	2.00	0.43
1:C:42:ARG:HG3	1:C:310:LEU:CD2	2.49	0.43
1:D:278:GLU:OE1	3:D:502:HOH:O	2.21	0.42
1:D:359:SER:HB2	1:D:361:HIS:NE2	2.35	0.42
1:D:276:GLU:HB3	1:D:330:ARG:HD2	2.03	0.41
1:B:303:ARG:HD3	3:B:552:HOH:O	2.20	0.41
1:C:357:PHE:O	1:C:358:ILE:HG13	2.21	0.41
1:A:357:PHE:HE2	1:A:359:SER:HB2	1.86	0.40

All (1) 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)
3:A:506:HOH:O	3:B:549:HOH:O[1_655]	2.15	0.05

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	363/363 (100%)	345 (95%)	15 (4%)	3 (1%)	24 8
1	B	362/363 (100%)	346 (96%)	14 (4%)	2 (1%)	30 14
1	C	363/363 (100%)	348 (96%)	13 (4%)	2 (1%)	30 14
1	D	363/363 (100%)	353 (97%)	9 (2%)	1 (0%)	46 29
All	All	1451/1452 (100%)	1392 (96%)	51 (4%)	8 (1%)	30 14

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	348	ALA
1	B	347	GLN
1	A	346	GLY
1	B	358	ILE
1	A	356	LEU
1	C	346	GLY
1	C	354	GLU
1	D	345	SER

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%)	290 (99%)	3 (1%)	82	75
1	B	292/291 (100%)	287 (98%)	5 (2%)	68	54
1	C	293/291 (101%)	288 (98%)	5 (2%)	68	54
1	D	293/291 (101%)	288 (98%)	5 (2%)	68	54
All	All	1171/1164 (101%)	1153 (98%)	18 (2%)	72	60

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	TYR
1	A	295	TRP
1	A	318	GLU
1	B	173	TYR
1	B	295	TRP
1	B	354	GLU
1	B	356	LEU
1	B	360	ASN
1	C	41	LYS
1	C	42	ARG
1	C	173	TYR
1	C	295	TRP
1	C	360	ASN
1	D	69	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	173	TYR
1	D	193	ASP
1	D	295	TRP
1	D	330	ARG

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)

3 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	GOL	B	401[A]	-	5,5,5	0.39	0	5,5,5	0.65	0
2	GOL	B	401[B]	-	5,5,5	0.35	0	5,5,5	0.20	0
2	GOL	D	401	-	5,5,5	0.42	0	5,5,5	0.56	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	GOL	B	401[A]	-	-	0/4/4/4	0/0/0/0
2	GOL	B	401[B]	-	-	0/4/4/4	0/0/0/0
2	GOL	D	401	-	-	0/4/4/4	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.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	363/363 (100%)	0.11	19 (5%)	31	28	10, 19, 52, 143	8 (2%)
1	B	363/363 (100%)	0.20	19 (5%)	31	28	9, 18, 58, 169	9 (2%)
1	C	363/363 (100%)	0.26	19 (5%)	31	28	10, 18, 53, 175	8 (2%)
1	D	363/363 (100%)	0.21	24 (6%)	22	19	11, 22, 60, 135	2 (0%)
All	All	1452/1452 (100%)	0.19	81 (5%)	28	25	9, 19, 59, 175	27 (1%)

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	355	SER	28.7
1	A	353	SER	23.0
1	B	356	LEU	22.3
1	C	351	ALA	21.2
1	B	351	ALA	20.6
1	B	352	ALA	18.9
1	C	357	PHE	17.1
1	B	357	PHE	16.9
1	C	358	ILE	16.5
1	C	346	GLY	15.8
1	B	358	ILE	15.4
1	A	355	SER	15.1
1	C	353	SER	15.0
1	B	353	SER	13.6
1	D	363	TYR	13.6
1	B	349	GLY	12.9
1	A	356	LEU	12.9
1	A	348	ALA	12.4
1	A	357	PHE	12.2
1	D	361	HIS	12.2
1	C	349	GLY	12.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	354	GLU	11.9
1	B	348	ALA	11.8
1	C	352	ALA	11.8
1	C	350	ALA	11.5
1	C	356	LEU	11.4
1	D	346	GLY	11.0
1	D	362	ALA	10.9
1	A	352	ALA	10.9
1	C	348	ALA	10.8
1	A	345	SER	10.6
1	D	345	SER	10.5
1	C	355	SER	10.0
1	B	350	ALA	9.9
1	A	350	ALA	9.2
1	D	348	ALA	8.9
1	A	346	GLY	8.2
1	B	354	GLU	7.6
1	A	351	ALA	7.5
1	C	347	GLN	7.4
1	B	346	GLY	7.4
1	A	358	ILE	6.8
1	A	359	SER	6.7
1	A	347	GLN	6.3
1	D	41	LYS	5.7
1	B	345	SER	5.0
1	A	344	PRO	4.6
1	B	359	SER	4.5
1	B	347	GLN	4.4
1	C	359	SER	4.4
1	C	354	GLU	4.4
1	D	42	ARG	4.3
1	D	349	GLY	4.2
1	A	349	GLY	3.7
1	C	44	GLN	3.6
1	D	347	GLN	3.4
1	D	358	ILE	3.3
1	B	344	PRO	3.1
1	D	343	THR	3.0
1	A	360	ASN	3.0
1	B	360	ASN	3.0
1	D	354	GLU	2.8
1	D	355	SER	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	360	ASN	2.8
1	C	344	PRO	2.7
1	D	360	ASN	2.7
1	A	2	HIS	2.7
1	D	344	PRO	2.7
1	D	40	ALA	2.6
1	D	36	THR	2.6
1	D	353	SER	2.5
1	A	317	LYS	2.4
1	D	356	LEU	2.4
1	B	45	SER	2.3
1	D	44	GLN	2.2
1	B	362	ALA	2.2
1	D	352	ALA	2.2
1	D	46	ILE	2.1
1	D	43	LEU	2.1
1	C	41	LYS	2.1
1	C	43	LEU	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
2	GOL	B	401[A]	6/6	0.88	0.20	7.20	11,17,23,26	0
2	GOL	B	401[B]	6/6	0.88	0.20	7.12	13,16,21,23	14
2	GOL	D	401	6/6	0.93	0.14	1.27	14,20,25,25	0

6.5 Other polymers (i)

There are no such residues in this entry.