wwPDB X-ray Structure Validation Summary Report (i)

Sep 10, 2023 – 10:45 AM EDT

PDB ID	:	4JTG
Title	:	Crystal structure of F114R/R117A mutant of 3-deoxy-D-manno-octulosonate 8-phosphate synthase (KDO8PS) from Neisseria meningitidis
Authors	:	Allison, T.M.; Cochrane, F.C.; Jameson, G.B.; Parker, E.J.
Deposited on	:	2013-03-23
Resolution	:	1.85 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:

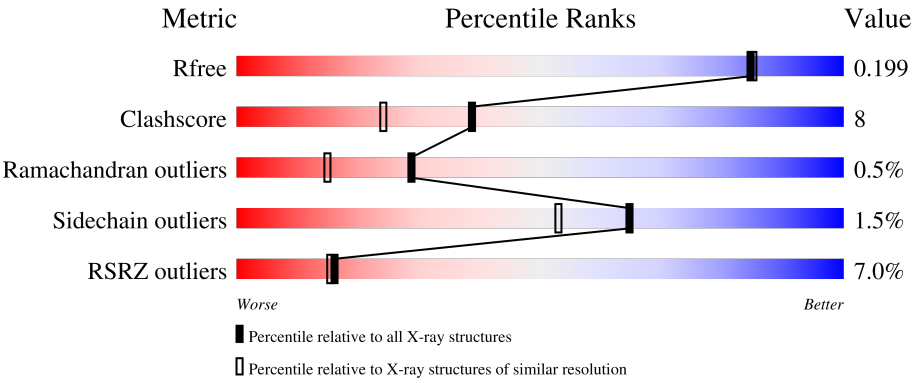
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	280	9% 78% 11% • 10%
1	B	280	5% 76% 13% • 9%
1	C	280	5% 84% 7% • 7%
1	D	280	7% 84% 8% • 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8876 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 2-dehydro-3-deoxyphosphooctonate aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	253	Total	C	N	O	S	0	9	0
			2015	1290	344	370	11			
1	B	255	Total	C	N	O	S	0	9	0
			2008	1289	339	368	12			
1	C	260	Total	C	N	O	S	0	8	0
			2051	1311	347	382	11			
1	D	257	Total	C	N	O	S	0	4	0
			1999	1281	339	368	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	114	ARG	PHE	engineered mutation	UNP Q9JZ55
A	117	ALA	ARG	engineered mutation	UNP Q9JZ55
B	114	ARG	PHE	engineered mutation	UNP Q9JZ55
B	117	ALA	ARG	engineered mutation	UNP Q9JZ55
C	114	ARG	PHE	engineered mutation	UNP Q9JZ55
C	117	ALA	ARG	engineered mutation	UNP Q9JZ55
D	114	ARG	PHE	engineered mutation	UNP Q9JZ55
D	117	ALA	ARG	engineered mutation	UNP Q9JZ55

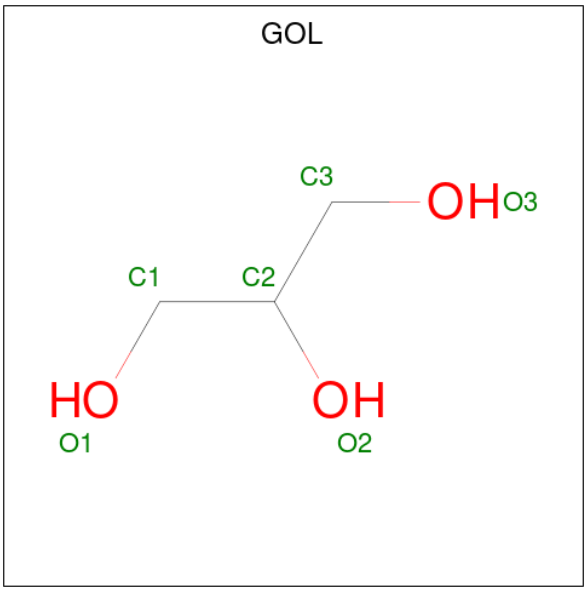
- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total Na 1 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total C O 6 3 3	0	0

- Molecule 5 is water.

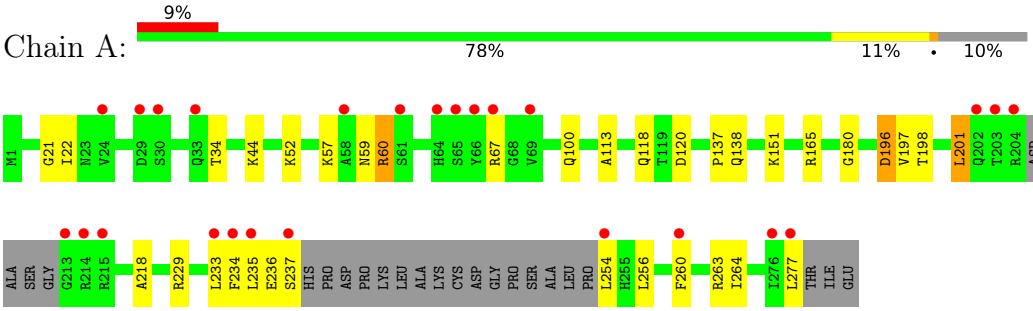
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	179	Total O 179 179	0	0
5	B	207	Total O 207 207	0	0
5	C	203	Total O 203 203	0	0
5	D	201	Total O 201 201	0	0

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.

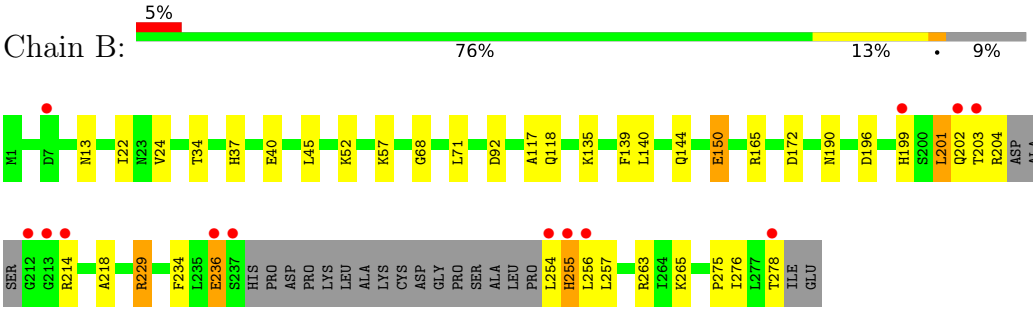
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

Chain A:



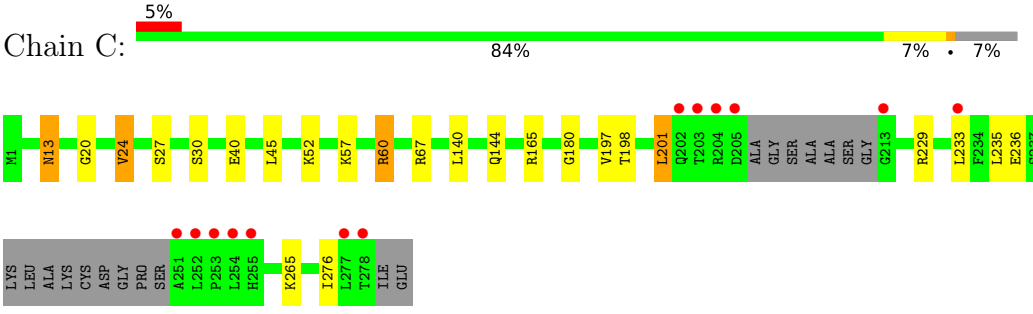
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

Chain B:



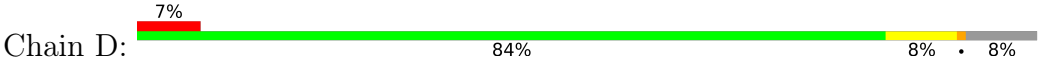
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

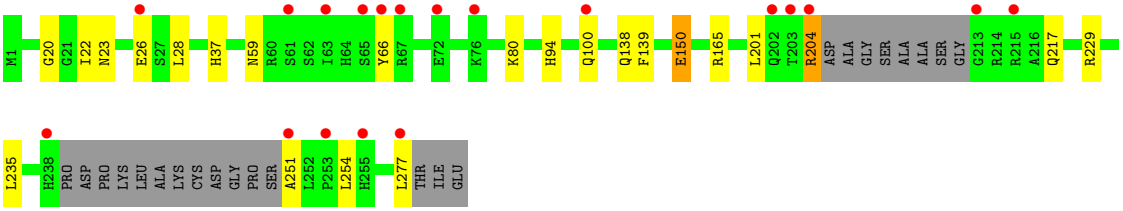
Chain C:



• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

Chain D:




M1
G20
G21
I22
N23
E26
S27
L28
H37
N59
R60
S61
S62
I63
H64
S65
Y66
R67
E72
K76
K80
H94
Q100
Q138
F139
E150
R165
L201
Q202
T203
R204
ASP
ALA
GLY
SER
ALA
ALA
SER
GLY
G213
R214
R215
A216
Q217
R229
L235
H238
PRO
ASP
PRO
LYS
LEU
ALA
LYS
CYS
ASP
GLY
PRO
SER
A251
L252
P253
L254
H255
L277
THR
ILE
GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.59Å 85.37Å 163.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.58 – 1.85 33.58 – 1.85	Depositor EDS
% Data completeness (in resolution range)	97.6 (33.58-1.85) 97.6 (33.58-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.93 (at 1.85Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.161 , 0.191 0.172 , 0.199	Depositor DCC
R_{free} test set	4779 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	20.4	Xtriage
Anisotropy	0.265	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 47.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8876	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% 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: CL, GOL, NA

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	0.92	2/2052 (0.1%)	0.89	3/2770 (0.1%)
1	B	1.00	1/2051 (0.0%)	0.94	5/2772 (0.2%)
1	C	0.95	0/2088	0.88	2/2824 (0.1%)
1	D	0.90	2/2035 (0.1%)	0.87	2/2750 (0.1%)
All	All	0.94	5/8226 (0.1%)	0.89	12/11116 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	21	GLY	N-CA	-5.96	1.37	1.46
1	D	217	GLN	CG-CD	5.68	1.64	1.51
1	D	150	GLU	CD-OE2	5.50	1.31	1.25
1	B	150	GLU	CG-CD	5.29	1.59	1.51
1	A	100	GLN	CG-CD	5.14	1.62	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	201	LEU	CB-CA-C	-5.86	99.06	110.20
1	B	135	LYS	CD-CE-NZ	-5.80	98.36	111.70
1	A	67	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	B	214	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	C	201	LEU	CB-CA-C	-5.58	99.60	110.20

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	2015	0	2071	45	0
1	B	2008	0	2068	46	0
1	C	2051	0	2088	19	0
1	D	1999	0	2048	20	0
2	A	1	0	0	1	0
2	B	1	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	1	0
3	C	1	0	0	0	0
4	D	6	0	8	0	0
5	A	179	0	0	13	0
5	B	207	0	0	12	0
5	C	203	0	0	2	0
5	D	201	0	0	10	0
All	All	8876	0	8283	123	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 8.

The worst 5 of 123 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:235:LEU:HD11	1:A:260:PHE:CE1	1.61	1.34
1:A:263[A]:ARG:NH1	5:A:577:HOH:O	1.70	1.18
1:A:118[A]:GLN:NE2	5:A:554:HOH:O	1.78	1.11
1:D:80:LYS:HD2	5:D:532:HOH:O	1.52	1.09
1:B:37:HIS:HE2	1:B:254:LEU:N	1.57	1.01

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	256/280 (91%)	250 (98%)	5 (2%)	1 (0%)	34 19
1	B	258/280 (92%)	251 (97%)	5 (2%)	2 (1%)	19 7
1	C	262/280 (94%)	258 (98%)	3 (1%)	1 (0%)	34 19
1	D	255/280 (91%)	250 (98%)	4 (2%)	1 (0%)	34 19
All	All	1031/1120 (92%)	1009 (98%)	17 (2%)	5 (0%)	29 15

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	229	ARG
1	B	229	ARG
1	B	255	HIS
1	C	229	ARG
1	D	229	ARG

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	223/235 (95%)	219 (98%)	4 (2%)	59 45
1	B	223/235 (95%)	221 (99%)	2 (1%)	78 72
1	C	228/235 (97%)	224 (98%)	4 (2%)	59 45
1	D	221/235 (94%)	218 (99%)	3 (1%)	67 55
All	All	895/940 (95%)	882 (98%)	13 (2%)	65 53

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	40	GLU
1	C	60	ARG
1	D	277	LEU
1	D	165	ARG
1	D	204	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	23	ASN
1	C	147	ASN
1	C	118	GLN
1	C	100	GLN
1	C	144	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 7 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	GOL	D	303	-	5,5,5	1.18	0	5,5,5	1.79	2 (40%)

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	GOL	D	303	-	-	0/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	303	GOL	C3-C2-C1	2.52	121.50	111.70
4	D	303	GOL	O2-C2-C1	-2.04	100.12	109.12

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	253/280 (90%)	0.28	25 (9%)	7	7	12, 25, 62, 80	1 (0%)
1	B	255/280 (91%)	-0.10	13 (5%)	28	26	11, 20, 43, 77	0
1	C	260/280 (92%)	-0.06	15 (5%)	23	22	12, 20, 48, 75	0
1	D	257/280 (91%)	0.16	19 (7%)	14	14	12, 24, 58, 74	0
All	All	1025/1120 (91%)	0.07	72 (7%)	16	15	11, 22, 57, 80	1 (0%)

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	254	LEU	9.5
1	B	212	GLY	6.0
1	B	256	LEU	5.5
1	A	277	LEU	5.4
1	A	254	LEU	5.4

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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, 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	B-factors(Å^2)	Q<0.9
4	GOL	D	303	6/6	0.85	0.18	30,35,38,40	0
2	CL	C	303	1/1	0.97	0.04	32,32,32,32	0
2	CL	A	301	1/1	0.98	0.08	33,33,33,33	0
2	CL	B	301	1/1	0.99	0.05	21,21,21,21	0
2	CL	D	301	1/1	0.99	0.03	34,34,34,34	0
2	CL	D	302	1/1	0.99	0.05	27,27,27,27	0
3	NA	C	302	1/1	0.99	0.04	19,19,19,19	0
2	CL	C	301	1/1	0.99	0.07	22,22,22,22	0

6.5 Other polymers (i)

There are no such residues in this entry.