wwPDB X-ray Structure Validation Summary Report (i)

Sep 10, 2023 – 07:26 PM EDT

PDB ID	:	4JTK
Title	:	Crystal structure of R117Q 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	:	1.86 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 (1)) were used in the production of this report:

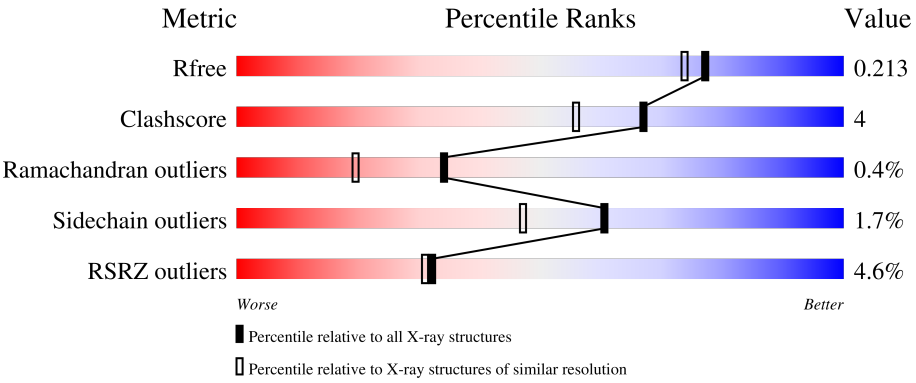
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.86 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	7% 82% 8% 10%
1	B	280	% 80% 10% • 10%
1	C	280	4% 82% 9% 9%
1	D	280	5% 84% 7% 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8384 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	252	Total	C	N	O	S	0	0	0
			1941	1246	328	356	11			
1	B	253	Total	C	N	O	S	0	0	0
			1952	1251	330	360	11			
1	C	256	Total	C	N	O	S	0	3	0
			1993	1279	337	366	11			
1	D	257	Total	C	N	O	S	0	2	0
			1988	1277	335	365	11			

There are 4 discrepancies between the modelled and reference sequences:

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

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

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

- Molecule 4 is water.

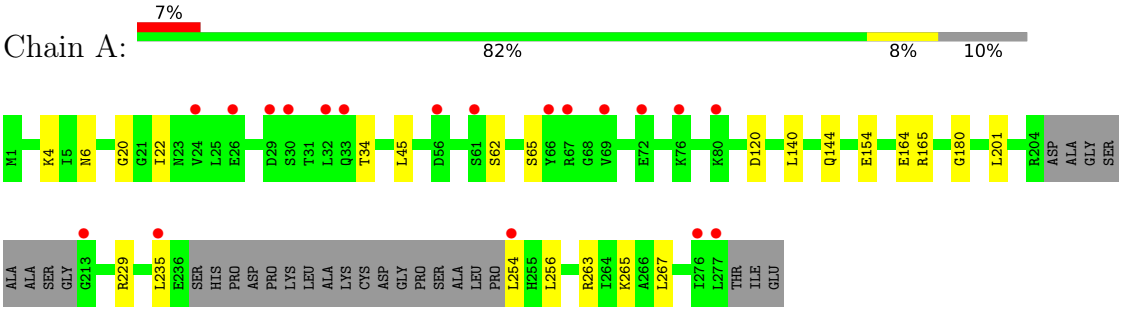
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	109	Total O 109 109	0	0
4	B	145	Total O 145 145	0	0
4	C	128	Total O 128 128	0	0
4	D	123	Total O 123 123	0	0

3 Residue-property plots

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.

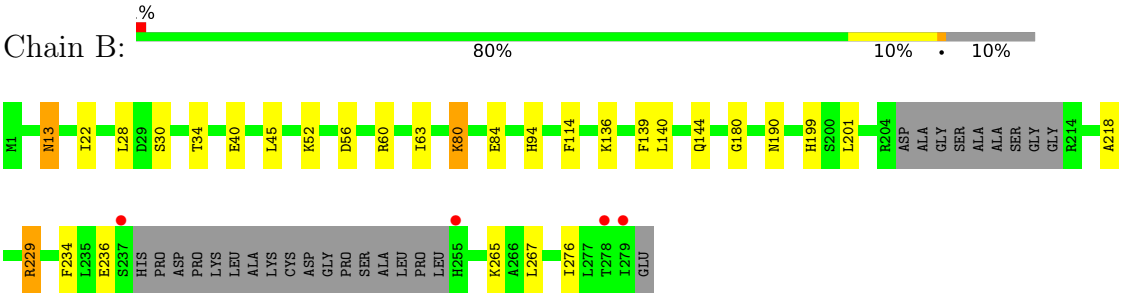
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

Chain A:



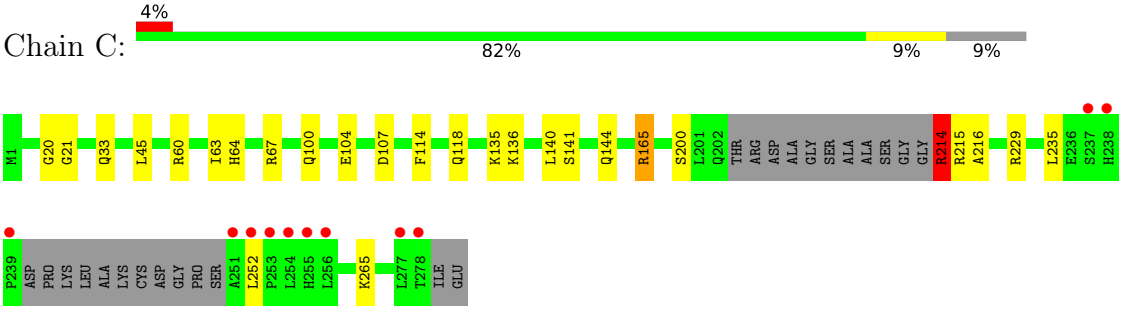
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

Chain B:



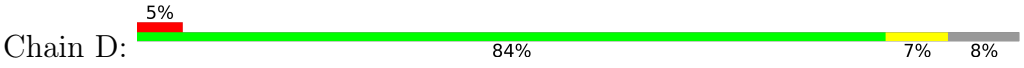
• Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

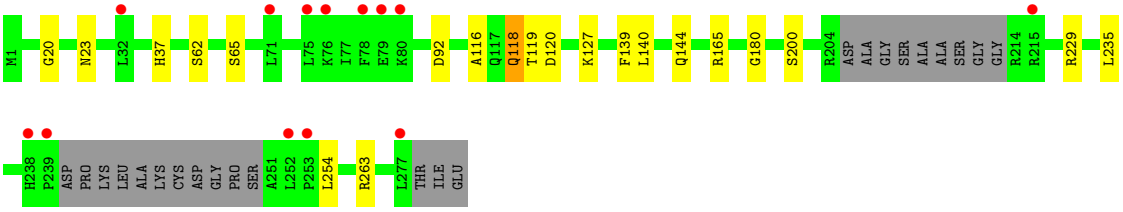
Chain C:



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

Chain D:




M1
G20
N23
L32
H37
S62
S65
L71
L75
K76
I77
F78
E79
K80
D92
A116
Q117
Q118
T119
D120
K127
F139
L140
Q144
R165
G180
S200
R204
ASP
ALA
GLY
SER
ALA
ALA
SER
GLY
GLY
R214
R215
R229
L235

H238
P239
ASP
PRO
LYS
LEU
ALA
LYS
CYS
ASP
GLY
PRO
SER
A251
L252
P253
L254
R263
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	81.47Å 85.42Å 162.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.31 – 1.86 41.31 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.9 (41.31-1.86) 99.9 (41.31-1.86)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.31 (at 1.86Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.176 , 0.206 0.184 , 0.213	Depositor DCC
R_{free} test set	4796 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	23.6	Xtriage
Anisotropy	0.533	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 42.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.017 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8384	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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: CL, 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	0.86	2/1973 (0.1%)	0.84	2/2665 (0.1%)
1	B	0.94	0/1984	0.88	1/2680 (0.0%)
1	C	0.95	2/2028 (0.1%)	0.89	3/2743 (0.1%)
1	D	0.84	1/2026 (0.0%)	0.89	3/2741 (0.1%)
All	All	0.90	5/8011 (0.1%)	0.88	9/10829 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	200	SER	CB-OG	-5.98	1.34	1.42
1	D	200	SER	CB-OG	-5.84	1.34	1.42
1	A	164	GLU	CD-OE2	-5.17	1.20	1.25
1	C	21	GLY	N-CA	-5.16	1.38	1.46
1	A	164	GLU	CD-OE1	-5.06	1.20	1.25

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	252	LEU	CA-CB-CG	5.98	129.05	115.30
1	D	263	ARG	NE-CZ-NH2	-5.90	117.35	120.30
1	A	263	ARG	NE-CZ-NH2	-5.82	117.39	120.30
1	C	214	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	C	107	ASP	CB-CG-OD2	5.35	123.12	118.30

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	1941	0	1987	14	0
1	B	1952	0	1995	25	0
1	C	1993	0	2029	21	0
1	D	1988	0	2028	12	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	2	0	0	0	0
3	C	1	0	0	0	0
4	A	109	0	0	1	0
4	B	145	0	0	4	0
4	C	128	0	0	1	0
4	D	123	0	0	5	0
All	All	8384	0	8039	62	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 4.

The worst 5 of 62 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:C:165[A]:ARG:HH11	1:C:165[A]:ARG:HG3	1.42	0.84
1:C:165[B]:ARG:HH21	1:C:165[B]:ARG:HG3	1.41	0.84
1:C:165[A]:ARG:HH11	1:C:165[A]:ARG:CG	2.02	0.72
1:A:120:ASP:OD1	1:C:67:ARG:NH1	2.24	0.70
1:C:165[B]:ARG:HG3	1:C:165[B]:ARG:NH2	2.02	0.68

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	246/280 (88%)	241 (98%)	4 (2%)	1 (0%)	34	19
1	B	247/280 (88%)	240 (97%)	6 (2%)	1 (0%)	34	19
1	C	253/280 (90%)	246 (97%)	6 (2%)	1 (0%)	34	19
1	D	253/280 (90%)	249 (98%)	3 (1%)	1 (0%)	34	19
All	All	999/1120 (89%)	976 (98%)	19 (2%)	4 (0%)	34	19

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	229	ARG
1	A	229	ARG
1	B	229	ARG
1	C	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	214/236 (91%)	212 (99%)	2 (1%)	78	72
1	B	216/236 (92%)	211 (98%)	5 (2%)	50	34
1	C	220/236 (93%)	214 (97%)	6 (3%)	44	29
1	D	220/236 (93%)	217 (99%)	3 (1%)	67	55
All	All	870/944 (92%)	854 (98%)	16 (2%)	60	45

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	127	LYS
1	D	118	GLN
1	C	60	ARG
1	C	214	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	33	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	100	GLN
1	C	186	GLN
1	D	23	ASN
1	C	217	GLN
1	B	138	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 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

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

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	252/280 (90%)	0.26	19 (7%)	14	14	15, 30, 60, 87	1 (0%)
1	B	253/280 (90%)	-0.20	4 (1%)	72	72	14, 23, 42, 98	0
1	C	256/280 (91%)	-0.18	11 (4%)	35	33	15, 24, 46, 76	0
1	D	257/280 (91%)	0.21	13 (5%)	28	26	16, 31, 59, 79	0
All	All	1018/1120 (90%)	0.02	47 (4%)	32	31	14, 26, 57, 98	1 (0%)

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	277	LEU	8.8
1	B	279	ILE	7.0
1	D	277	LEU	6.6
1	C	239	PRO	6.2
1	A	213	GLY	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
3	NA	C	302	1/1	0.97	0.07	27,27,27,27	0
2	CL	D	302	1/1	0.98	0.03	33,33,33,33	0
2	CL	C	301	1/1	0.98	0.04	40,40,40,40	0
2	CL	B	301	1/1	0.99	0.04	34,34,34,34	0
2	CL	D	301	1/1	1.00	0.04	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.