



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 10, 2023 – 09:36 AM EDT

PDB ID : 4JTJ  
Title : Crystal structure of R117K 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 : 1.75 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](mailto: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 ⓘ 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 ⓘ](#)) 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

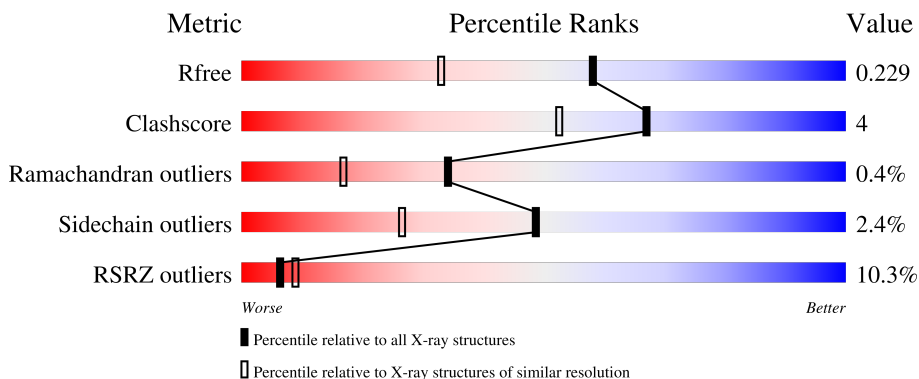
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	 12% 81% 9% 10%
1	B	280	 4% 79% 10% 10%
1	C	280	 6% 84% 7% 8%
1	D	280	 15% 84% 7% 8%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8274 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
			Total	C	N	O	S			
1	A	253	1963	1260	332	360	11	0	2	0
1	B	253	1981	1274	333	362	12	0	5	0
1	C	257	1991	1282	334	364	11	0	3	0
1	D	259	1997	1280	338	368	11	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	117	LYS	ARG	engineered mutation	UNP Q9JZ55
B	117	LYS	ARG	engineered mutation	UNP Q9JZ55
C	117	LYS	ARG	engineered mutation	UNP Q9JZ55
D	117	LYS	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	C	1	Total	Cl	0	0
			1	1		
2	D	1	Total	Cl	0	0
			1	1		

- 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		

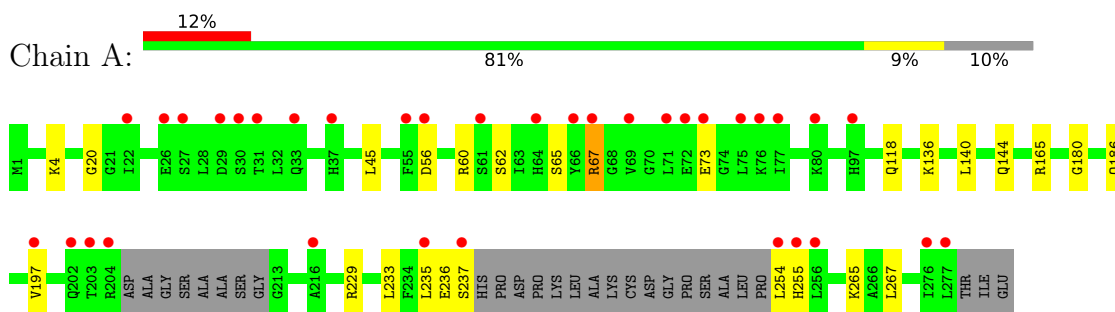
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	77	Total 77	O 77	0	0
4	B	96	Total 96	O 96	0	0
4	C	92	Total 92	O 92	0	0
4	D	73	Total 73	O 73	0	0

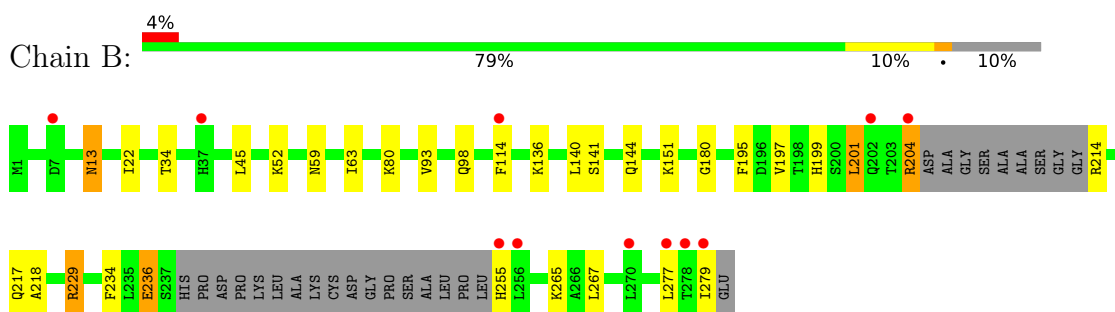
### 3 Residue-property plots [i](#)

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.

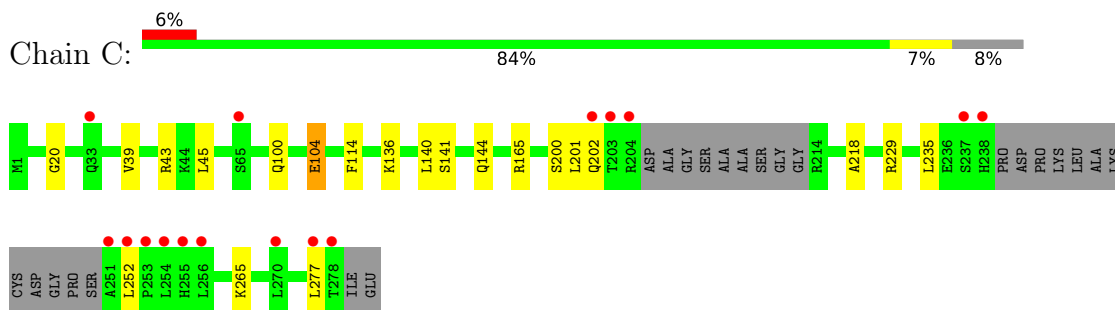
- Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase



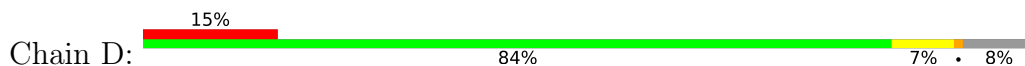
- Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase

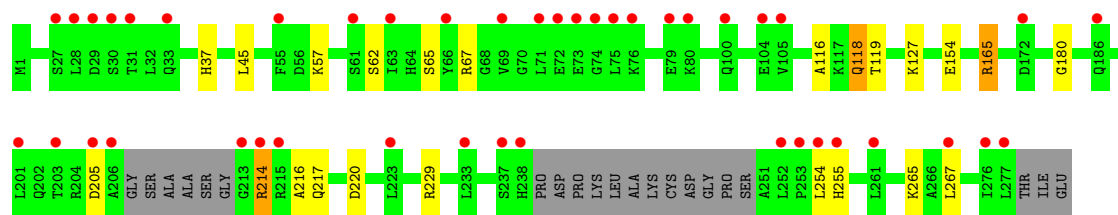


- Molecule 1: 2-dehydro-3-deoxyphosphooctonate aldolase



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





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.47Å 85.37Å 163.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.55 – 1.75 39.52 – 1.75	Depositor EDS
% Data completeness (in resolution range)	98.5 (39.55-1.75) 98.6 (39.52-1.75)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 1.75Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.197 , 0.226 0.203 , 0.229	Depositor DCC
$R_{free}$ test set	5687 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.5	Xtrriage
Anisotropy	0.328	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 38.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.004 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8274	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.80	0/1998	0.85	2/2698 (0.1%)
1	B	0.88	1/2025 (0.0%)	0.90	2/2735 (0.1%)
1	C	0.93	2/2034 (0.1%)	0.87	1/2749 (0.0%)
1	D	0.84	0/2034	0.89	3/2748 (0.1%)
All	All	0.86	3/8091 (0.0%)	0.88	8/10930 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	200	SER	CB-OG	-5.88	1.34	1.42
1	C	104	GLU	CB-CG	-5.80	1.41	1.52
1	B	236	GLU	CD-OE2	-5.58	1.19	1.25

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	136	LYS	CD-CE-NZ	5.86	125.18	111.70
1	C	136	LYS	CD-CE-NZ	5.77	124.97	111.70
1	D	67	ARG	NE-CZ-NH2	-5.45	117.57	120.30
1	D	165	ARG	NE-CZ-NH1	5.44	123.02	120.30
1	A	136	LYS	CD-CE-NZ	5.41	124.14	111.70

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	1963	0	2017	21	0
1	B	1981	0	2047	28	0
1	C	1991	0	2053	9	0
1	D	1997	0	2039	13	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	C	1	0	0	0	0
4	A	77	0	0	1	0
4	B	96	0	0	2	0
4	C	92	0	0	0	0
4	D	73	0	0	3	0
All	All	8274	0	8156	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:B:204:ARG:HH21	1:B:204:ARG:HG2	1.37	0.89
1:B:13:ASN:HD21	1:B:229:ARG:HH21	1.26	0.80
1:A:118:GLN:NE2	1:C:114:PHE:HE2	1.84	0.74
1:D:37:HIS:CD2	1:D:254:LEU:HG	2.25	0.71
1:A:118:GLN:NE2	1:C:114:PHE:CE2	2.61	0.69

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	249/280 (89%)	244 (98%)	4 (2%)	1 (0%)	34	17
1	B	252/280 (90%)	245 (97%)	6 (2%)	1 (0%)	34	17
1	C	254/280 (91%)	249 (98%)	4 (2%)	1 (0%)	34	17
1	D	254/280 (91%)	250 (98%)	3 (1%)	1 (0%)	34	17
All	All	1009/1120 (90%)	988 (98%)	17 (2%)	4 (0%)	34	17

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	229	ARG
1	B	229	ARG
1	A	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	218/236 (92%)	213 (98%)	5 (2%)	50	28
1	B	222/236 (94%)	217 (98%)	5 (2%)	50	28
1	C	222/236 (94%)	218 (98%)	4 (2%)	59	40
1	D	220/236 (93%)	212 (96%)	8 (4%)	35	13
All	All	882/944 (93%)	860 (98%)	22 (2%)	49	25

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	118	GLN
1	D	154[B]	GLU
1	D	154[A]	GLU
1	D	165	ARG
1	B	80	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15

such sidechains are listed below:

Mol	Chain	Res	Type
1	B	217	GLN
1	D	186	GLN
1	C	97	HIS
1	D	274	GLN
1	D	23	ASN

### 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 4 ligands modelled in this entry, 4 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	253/280 (90%)	0.79	35 (13%) <b>2</b> <b>4</b>	21, 35, 66, 80	1 (0%)
1	B	253/280 (90%)	0.37	11 (4%) 35 41	21, 29, 48, 75	0
1	C	257/280 (91%)	0.44	16 (6%) 20 26	21, 29, 48, 85	0
1	D	259/280 (92%)	0.97	43 (16%) <b>1</b> <b>2</b>	21, 36, 67, 75	0
All	All	1022/1120 (91%)	0.64	105 (10%) <b>6</b> <b>9</b>	21, 32, 64, 85	1 (0%)

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	277	LEU	11.7
1	D	277	LEU	9.1
1	A	66	TYR	8.2
1	C	278	THR	8.0
1	C	255	HIS	7.4

### 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 monosaccharides 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	CL	D	301	1/1	0.94	0.08	40,40,40,40	0
2	CL	C	301	1/1	0.96	0.06	41,41,41,41	0
3	NA	C	302	1/1	0.97	0.06	26,26,26,26	0
2	CL	A	301	1/1	0.98	0.08	34,34,34,34	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.