



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 06:57 AM GMT

PDB ID : 1KEP
Title : The crystal structure of dTDP-D-glucose 4,6-dehydratase (RmlB) from *Streptococcus suis* with dTDP-xylose bound
Authors : Allard, S.T.M.; Beis, K.; Giraud, M.-F.; Hegeman, A.D.; Gross, J.W.; Whitfield, C.; Graninger, M.; Messner, P.; Allen, A G.; Naismith, J.H.
Deposited on : 2001-11-16
Resolution : 1.80 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

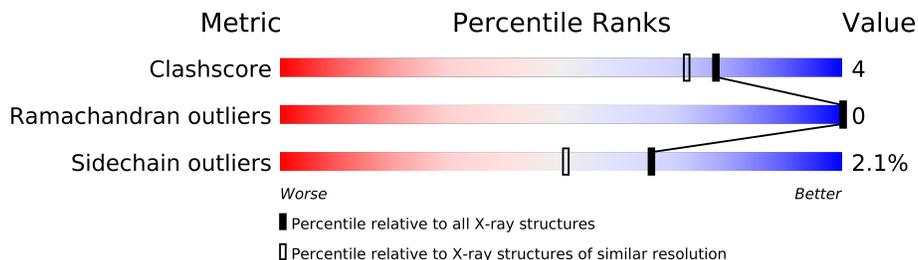
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 21963
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.80 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))
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	348	
1	B	348	

2 Entry composition (i)

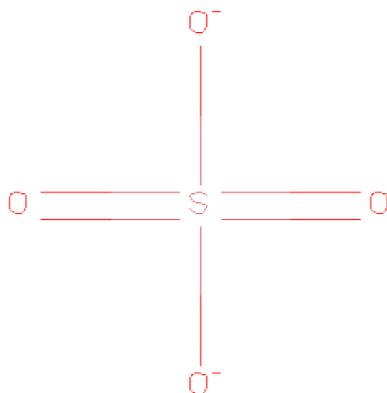
There are 5 unique types of molecules in this entry. The entry contains 6431 atoms, of which 0 are hydrogen and 0 are deuterium.

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 dTDP-D-glucose 4,6-dehydratase.

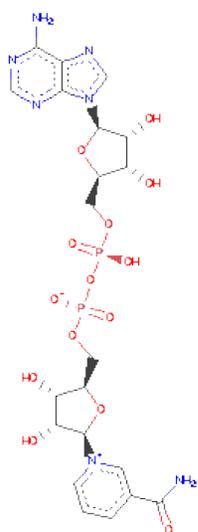
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	346	2741	1736	468	534	3	0	0	0
1	B	346	2737	1733	467	534	3	0	0	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



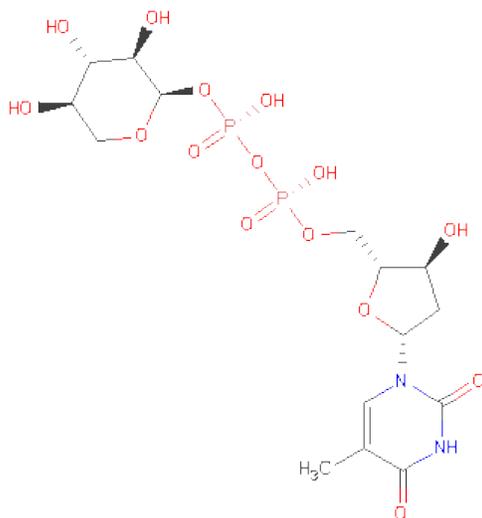
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0

- Molecule 3 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C₂₁H₂₇N₇O₁₄P₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
3	A	1	44	21	7	14	2	0	0
3	B	1	44	21	7	14	2	0	0

- Molecule 4 is THYMIDINE-5'-DIPHOSPHO-BETA-D-XYLOSE (three-letter code: TDX) (formula: $C_{15}H_{24}N_2O_{15}P_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	12	5	7	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	B	1	34	15	2	15	2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	438	Total	O	0	0
			438	438		
5	B	366	Total	O	0	0
			366	366		

3 Residue-property plots

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.

Note EDS was not executed.

- Molecule 1: dTDP-D-glucose 4,6-dehydratase

Chain A:



- Molecule 1: dTDP-D-glucose 4,6-dehydratase

Chain B:



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	62.45Å 97.18Å 185.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.29 – 1.80	Depositor
% Data completeness (in resolution range)	96.8 (91.29-1.80)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	0.08	Depositor
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.145 , 0.172	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	6431	wwPDB-VP
Average B, all atoms (Å ²)	8.0	wwPDB-VP

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: TDX, SO4, NAD

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	0.96	4/2809 (0.1%)	1.09	16/3817 (0.4%)
1	B	0.86	2/2805 (0.1%)	0.95	11/3814 (0.3%)
All	All	0.91	6/5614 (0.1%)	1.02	27/7631 (0.4%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	163	SER	CA-CB	9.68	1.67	1.52
1	A	131	ASP	CB-CG	-9.29	1.32	1.51
1	B	131	ASP	CB-CG	-9.01	1.32	1.51
1	A	131	ASP	CA-CB	7.99	1.71	1.53
1	A	344	GLN	CB-CG	-7.61	1.31	1.52

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	131	ASP	CB-CG-OD1	-27.39	93.65	118.30
1	A	131	ASP	CB-CG-OD2	22.08	138.18	118.30
1	B	131	ASP	CB-CG-OD1	-17.98	102.11	118.30
1	B	131	ASP	CB-CG-OD2	15.20	131.98	118.30
1	A	163	SER	CA-CB-OG	-9.73	84.92	111.20

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2741	0	2627	20	0
1	B	2737	0	2621	21	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
3	A	44	0	26	0	0
3	B	44	0	26	1	0
4	A	12	0	9	0	0
4	B	34	0	22	2	0
5	A	438	0	0	14	1
5	B	366	0	0	11	0
All	All	6431	0	5331	41	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 41 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:265:GLU:HG2	5:B:6269:HOH:O	1.64	0.95
1:A:19:ASN:OD1	5:A:6433:HOH:O	1.94	0.85
1:B:2:SER:HA	5:B:6039:HOH:O	1.79	0.82
1:A:19:ASN:HB2	5:A:6100:HOH:O	1.82	0.79
1:B:2:SER:OG	5:B:6359:HOH:O	2.02	0.77

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	Distance(Å)	Clash(Å)
5:A:6295:HOH:O	5:A:6298:HOH:O[4_556]	2.17	0.03

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	344/348 (99%)	331 (96%)	13 (4%)	0	100	100
1	B	344/348 (99%)	333 (97%)	11 (3%)	0	100	100
All	All	688/696 (99%)	664 (96%)	24 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	289/291 (99%)	283 (98%)	6 (2%)	66	51
1	B	289/291 (99%)	283 (98%)	6 (2%)	66	51
All	All	578/582 (99%)	566 (98%)	12 (2%)	66	51

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	269	GLU
1	B	131	ASP
1	B	262	GLU
1	A	246	MET
1	B	191	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	260	ASN
1	B	32	HIS
1	B	260	ASN
1	A	231	ASN
1	B	231	ASN

5.3.3 RNA [i](#)

There are no RNA chains 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)

7 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
3	NAD	A	1400	-	48,48,48	1.25	4 (8%)	73,73,73	1.44	6 (8%)
4	TDX	A	3500	-	10,10,36	1.10	1 (10%)	13,14,55	0.89	1 (7%)
2	SO4	A	6001	-	4,4,4	0.63	0	6,6,6	0.62	0
2	SO4	A	6002	-	4,4,4	0.58	0	6,6,6	0.41	0
3	NAD	B	2400	-	48,48,48	1.35	4 (8%)	73,73,73	1.74	12 (16%)
4	TDX	B	4500	-	36,36,36	1.07	2 (5%)	50,55,55	1.24	4 (8%)
2	SO4	B	6003	-	4,4,4	0.44	0	6,6,6	0.57	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
3	NAD	A	1400	-	-	0/30/62/62	0/3/5/5
4	TDX	A	3500	-	-	0/0/17/50	0/1/1/3
2	SO4	A	6001	-	-	0/0/0/0	0/0/0/0
2	SO4	A	6002	-	-	0/0/0/0	0/0/0/0
3	NAD	B	2400	-	-	0/30/62/62	0/3/5/5
4	TDX	B	4500	-	-	0/18/50/50	0/3/3/3
2	SO4	B	6003	-	-	0/0/0/0	0/0/0/0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2400	NAD	O7N-C7N	5.66	1.37	1.24
3	A	1400	NAD	O7N-C7N	5.16	1.36	1.24
3	A	1400	NAD	C2N-N1N	3.13	1.39	1.35
4	A	3500	TDX	O1'-C1'	3.02	1.44	1.25
3	A	1400	NAD	C2A-N3A	2.86	1.37	1.32

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2400	NAD	N3A-C2A-N1A	-8.03	122.00	128.71
3	A	1400	NAD	N3A-C2A-N1A	-6.49	123.28	128.71
3	B	2400	NAD	C3N-C7N-N7N	5.67	124.22	117.77
4	B	4500	TDX	N3-C2-N1	5.38	120.46	115.97
3	A	1400	NAD	C3N-C7N-N7N	3.73	122.01	117.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.