

Full wwPDB X-ray Structure Validation Report (i)

May 18, 2020 – 06:29 am BST

PDB ID : 4WPG

> Title Group A Streptococcus GacA is an essential dTDP-4-dehydrorhamnose

> > reductase (RmlD)

Authors : van der Beek, S.L.; Le Breton, Y.; Ferenbach, A.T.; van Aalten, D.M.F.;

Navratilova, I.; McIver, K.; van Sorge, N.M.; Dorfmueller, H.C.

Deposited on 2014-10-18

1.10 Å(reported) Resolution

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 https://www.wwpdb.org/validation/2017/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

> 1.8.5 (274361), CSD as541be (2020) Mogul

Xtriage (Phenix) 1.13 EDS 2.11

20191225.v01 (using entries in the PDB archive December 25th 2019) Percentile statistics

> Refmac 5.8.0158

CCP4 7.0.044 (Gargrove) Ideal geometry (proteins) Engh & Huber (2001)

Ideal geometry (DNA, RNA) Parkinson et al. (1996)

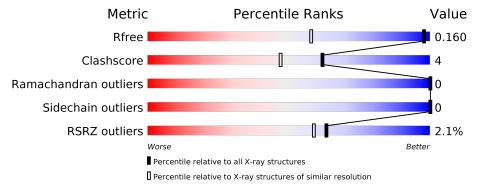
Validation Pipeline (wwPDB-VP) 2.11

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 1.10 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	$\begin{array}{c} \text{Whole archive} \\ (\#\text{Entries}) \end{array}$	$egin{aligned} ext{Similar resolution} \ (\# ext{Entries}, ext{resolution range}(\mathring{A})) \end{aligned}$		
R_{free}	130704	1619 (1.14-1.06)		
Clashscore	141614	1671 (1.14-1.06)		
Ramachandran outliers	138981	1615 (1.14-1.06)		
Sidechain outliers	138945	1613 (1.14-1.06)		
RSRZ outliers	127900	1588 (1.14-1.06)		

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 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	· ·						
			2%							
1	A	289	88%	11%	•					

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 criteria:

Mol	Type	Chain	m Res	Chirality	Geometry	Clashes	Electron density
4	PO4	A	303	-	-	X	-



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5062 atoms, of which 2259 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.

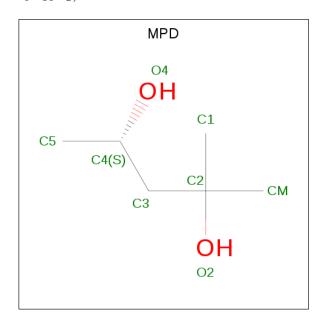
• Molecule 1 is a protein called dTDP-4-dehydrorhamnose reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	${f AltConf}$	Trace	
1	A	287	Total 4545	C 1472	H 2231	N 381	O 452	S 9	54	6	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	ain Residue Modelled		Actual	Comment	Reference
A	-4	GLY	_	expression tag	UNP T0BV37
A	-3	PRO	-	expression tag	
A	-2	LEU	_	expression tag	UNP T0BV37
A	-1	LEU	-	expression tag	UNP T0BV37
A	0	GLU	-	expression tag	UNP T0BV37

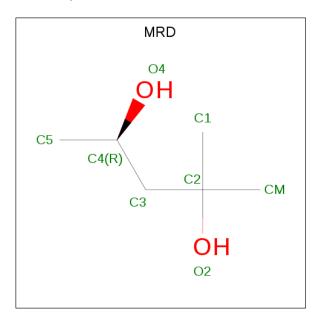
• Molecule 2 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).





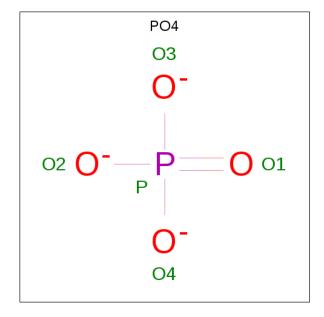
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	Δ	1	Total	С	Η	О	1	0
	Λ	1	22	6	14	2	1	

• Molecule 3 is (4R)-2-METHYLPENTANE-2,4-DIOL (three-letter code: MRD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	Λ	1	Total	С	Н	О	0	0
J	Λ	1	22	6	14	2	0	0

• Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	P 1	0	0

• Molecule 5 is water.

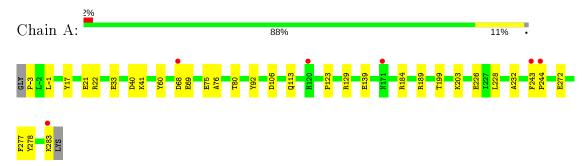
Mol	Chain	Residues	${f Atoms}$	ZeroOcc	AltConf
5	A	468	Total O 468 468	0	0



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 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: dTDP-4-dehydrorhamnose reductase





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants	36.48Å 45.98Å 48.87Å	Depositor
a, b, c, α , β , γ	66.30° 81.21° 95.27°	Depositor
Resolution (Å)	43.77 - 1.10	Depositor
resolution (A)	23.99 - 1.10	EDS
% Data completeness	94.0 (43.77-1.10)	Depositor
(in resolution range)	94.0 (23.99-1.10)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	4.88 (at 1.10Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
P. P.	0.129 , 0.155	Depositor
R, R_{free}	0.138 , 0.160	DCC
R_{free} test set	5445 reflections $(5.03%)$	wwPDB-VP
Wilson B-factor (\mathring{A}^2)	12.2	Xtriage
Anisotropy	0.292	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.42, 37.1	EDS
L-test for twinning ²	$ < L > = 0.49, < L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	5062	wwPDB-VP
Average B, all atoms $(Å^2)$	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 8.48% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: MRD, PO4, MPD

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	Bo	nd lengths	Bond angles		
			RMSZ	# Z > 5	RMSZ	# Z >5	
	1	Α	1.18	13/2367~(0.5%)	1.12	$16/3219 \ (0.5\%)$	

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\operatorname{\AA})$	$\operatorname{Ideal}(\text{\AA})$
1	A	75	GLU	CD-OE1	13.13	1.40	1.25
1	A	21	GLU	CD-OE2	8.75	1.35	1.25
1	A	139	GLU	CG-CD	8.40	1.64	1.51
1	A	228	LEU	C-O	7.72	1.38	1.23
1	A	139	GLU	CD-OE1	6.44	1.32	1.25
1	A	75	GLU	CG-CD	6.09	1.61	1.51
1	A	75	GLU	CB-CG	5.70	1.62	1.52
1	A	272	GLU	CD-OE1	5.44	1.31	1.25
1	A	203	LYS	CD-CE	-5.36	1.37	1.51
1	A	277	PHE	CG-CD2	5.32	1.46	1.38
1	A	184	ARG	CZ-NH2	-5.28	1.26	1.33
1	A	123	PRO	N-CD	5.25	1.55	1.47
1	A	226	GLU	CG-CD	5.21	1.59	1.51

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^o)$
1	A	-1	LEU	CB-CG-CD1	-10.94	92.41	111.00
1	A	17	TYR	CB-CG-CD1	7.61	125.57	121.00
1	A	184	ARG	NE-CZ-NH2	7.10	123.85	120.30
1	A	139	GLU	OE1-CD-OE2	-7.01	114.88	123.30
1	A	22	ARG	NE-CZ-NH2	6.73	123.67	120.30
1	A	75	GLU	CB-CG-CD	6.66	132.19	114.20
1	A	106[A]	ASP	CB-CG-OD2	-5.67	113.20	118.30
1	A	106[B]	ASP	CB-CG-OD2	-5.67	113.20	118.30

Continued on next page...



Continued	t_{mom}	marcata care	maaa
-	110111	DIEUIUUS	Duue
0 0 10001000000	J . \circ \circ	r	r

Mol	Chain	Res	Type	${f Atoms}$	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
1	A	189	ARG	NE-CZ-NH1	5.43	123.01	120.30
1	A	17	TYR	CB-CG-CD2	-5.34	117.79	121.00
1	A	92	TYR	CZ-CE2-CD2	5.29	124.56	119.80
1	A	92	TYR	CB-CG-CD1	5.29	124.17	121.00
1	A	272	GLU	CG-CD-OE2	-5.26	107.79	118.30
1	A	40	ASP	CB-CG-OD1	5.18	122.96	118.30
1	A	278	TYR	CB-CG-CD1	-5.12	117.93	121.00
1	A	60	TYR	CB-CG-CD1	5.05	124.03	121.00

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	Α	2314	2231	2213	17	0
2	A	8	14	14	2	0
3	A	8	14	14	0	0
4	A	5	0	0	2	0
5	A	468	0	0	12	1
All	All	2803	2259	2241	20	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 4.

All (20) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{array}{l} ext{Interatomic} \ ext{distance} \ (ext{Å}) \end{array}$	Clash overlap (Å)
1:A:283:LYS:HB2	5:A:777:HOH:O	1.69	0.92
1:A:113:GLN:OE1	5:A:401:HOH:O	1.94	0.84
1:A:41:LYS:HD2	5:A:843:HOH:O	1.82	0.78
4:A:303:PO4:O2	5:A:834:HOH:O	2.03	0.77
1:A:199:THR:OG1	2:A:301:MPD:H13	1.87	0.74
1:A:41:LYS:CD	5:A:843:HOH:O	2.35	0.73
4:A:303:PO4:P	5:A:834:HOH:O	2.51	0.67

Continued on next page...



~ · · · · · · · · · · · · · · · · · · ·	e		
Continued	trom	nremous	naae
-	110116	picolous	puyc

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${f distance}({ m \AA})$	overlap (Å)
1:A:41:LYS:HG3	5:A:843:HOH:O	1.97	0.65
1:A:129:ARG:NH1	5:A:862:HOH:O	2.36	0.59
1:A:232[B]:ALA:HA	5:A:711:HOH:O	2.02	0.58
2:A:301:MPD:O4	2:A:301:MPD:H11	2.05	0.57
1:A:-3:PRO:N	5:A:404:HOH:O	2.39	0.55
1:A:41:LYS:CG	5:A:843:HOH:O	2.52	0.53
1:A:243:PHE:CG	1:A:244:PRO:HD2	2.44	0.52
1:A:243:PHE:CD1	1:A:244:PRO:HD2	2.49	0.47
1:A:33:GLU:O	1:A:41:LYS:HE2	2.14	0.47
1:A:68:ASP:HB3	1:A:69:GLU:OE1	2.15	0.47
1:A:76:ALA:O	1:A:80:THR:HB	2.17	0.45
1:A:113:GLN:NE2	5:A:405:HOH:O	2.40	0.44
1:A:33:GLU:O	1:A:41:LYS:CE	2.68	0.42

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			$egin{aligned} ext{Clash} \ ext{overlap } (ext{Å}) \end{aligned}$
5:A:526:HOH:O	5:A:528:HOH:O[1_545]	1.98	0.22

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	291/289 (101%)	285 (98%)	6 (2%)	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	Analysed Rotameric		Percer	itiles
1	A	241/240 (100%)	241 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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	Tuno	Chain	Dog	Link	B	ond leng	$_{ m gths}$	В	ond ang	gles
10101	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	MPD	A	301	-	7,7,7	0.43	0	9,10,10	1.57	3 (33%)
3	MRD	A	302	-	7,7,7	0.91	0	9,10,10	1.27	1 (11%)
4	PO4	A	303	-	4,4,4	1.10	0	6,6,6	1.44	1 (16%)



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	MPD	A	301	-	-	0/5/5/5	_
3	MRD	A	302	_	_	0/5/5/5	_

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^o)$
2	A	301	MPD	O2-C2-CM	2.73	116.83	108.08
2	A	301	MPD	CM-C2-C1	-2.55	105.26	110.57
4	A	303	PO4	O4-P-O3	2.48	115.92	107.97
3	A	302	MRD	O4-C4-C3	2.22	120.33	111.36
2	A	301	MPD	O2-C2-C1	-2.05	101.50	108.08

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	MPD	2	0
4	A	303	PO4	2	0

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^{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		$\mathbf{OWAB}(\mathbf{\mathring{A}}^2)$	Q < 0.9
1	A	287/289 (99%)	-0.03	6 (2%) 63	59	9, 15, 26, 36	0

All (6) RSRZ outliers are listed below:

Mol	Chain	${f Res}$	\mathbf{Type}	RSRZ
1	A	244	PRO	5.1
1	A	283	LYS	4.5
1	A	243	PHE	3.3
1	A	68	ASP	2.6
1	A	120	HIS	2.2
1	A	171	ASN	2.2

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. 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	${f B\text{-factors}}({f \AA}^2)$	Q < 0.9
4	PO4	A	303	5/5	0.87	0.16	21,22,29,29	5

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	${f B\text{-factors}}({f \AA}^2)$	Q < 0.9
3	MRD	A	302	8/8	0.95	0.09	15,17,19,20	22
2	MPD	A	301	8/8	0.96	0.06	11,16,18,27	1

6.5 Other polymers (i)

There are no such residues in this entry.

