



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 04:21 PM GMT

PDB ID : 4A2C
Title : Crystal structure of galactitol-1-phosphatedehydrogenase from Escherichia coli
Authors : Alvarez, Y.; Esteban-Torres, M.; Acebron, I.; De Las Rivas, B.; Munoz, R.; Mancheno, J.M.
Deposited on : 2011-09-26
Resolution : 1.87 Å(reported)

This is a full wwPDB validation 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>

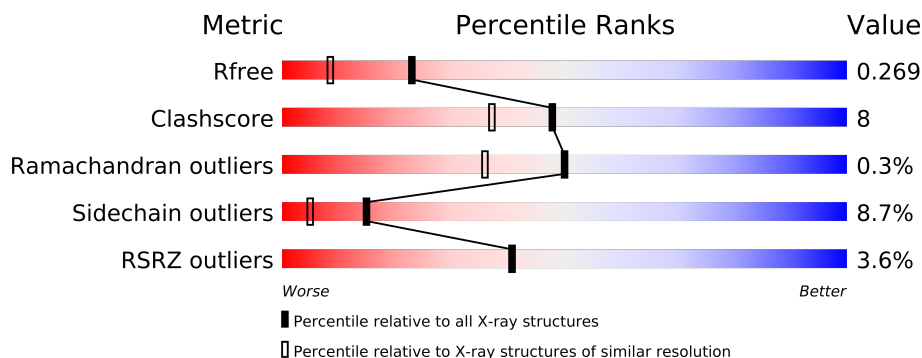
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
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 1.87 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}	66092	5260 (1.90-1.86)
Clashscore	79885	6268 (1.90-1.86)
Ramachandran outliers	78287	6195 (1.90-1.86)
Sidechain outliers	78261	6196 (1.90-1.86)
RSRZ outliers	66119	5262 (1.90-1.86)

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 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.

Mol	Chain	Length	Quality of chain
1	A	346	
1	B	346	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5923 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 GALACTITOL-1-PHOSPHATE5-DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	2	0
			2644	1684	454	491	15			
1	B	346	Total	C	N	O	S	0	5	0
			2666	1697	457	497	15			

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ni	0	0
			1	1		
2	A	2	Total	Ni	0	0
			2	2		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

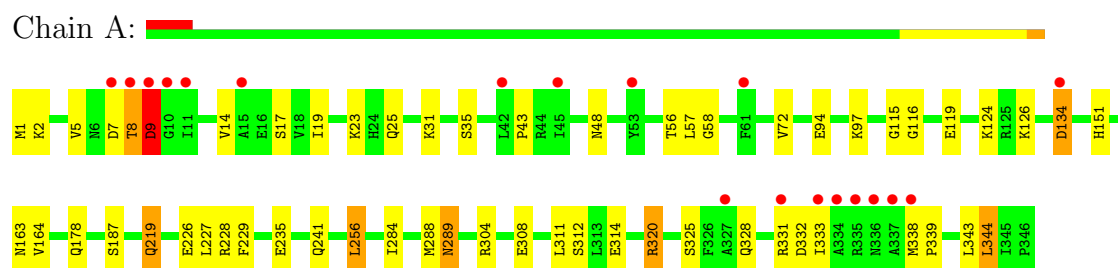
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	289	Total	O	0	0
			289	289		
4	B	319	Total	O	0	0
			319	319		

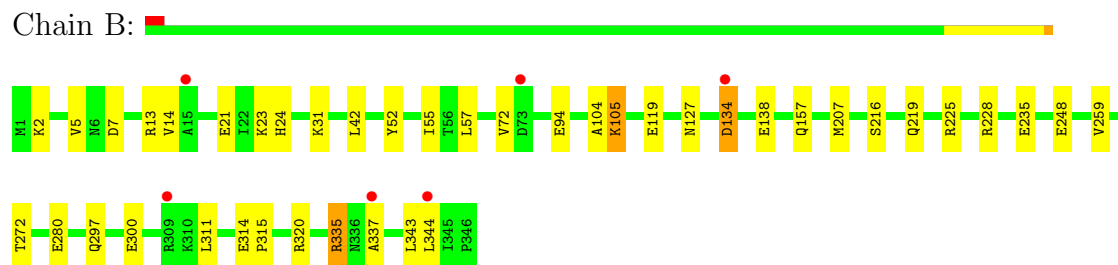
3 Residue-property plots

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.

• Molecule 1: GALACTITOL-1-PHOSPHATE5-DEHYDROGENASE



• Molecule 1: GALACTITOL-1-PHOSPHATE5-DEHYDROGENASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	43.72Å 77.14Å 107.51Å 90.00° 95.36° 90.00°	Depositor
Resolution (Å)	28.95 – 1.87 28.95 – 1.87	Depositor EDS
% Data completeness (in resolution range)	95.9 (28.95-1.87) 95.9 (28.95-1.87)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.86 (at 1.87Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.7.1.743)	Depositor
R, R_{free}	0.213 , 0.273 0.208 , 0.269	Depositor DCC
R_{free} test set	2861 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	21.7	Xtriage
Anisotropy	0.270	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 39.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 56448 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5923	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.02 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.6662e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NI, ZN

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	0.37	0/2699	0.54	1/3661 (0.0%)
1	B	0.39	0/2721	0.54	0/3692
All	All	0.38	0/5420	0.54	1/7353 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	9	ASP	N-CA-C	6.07	127.39	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	8	THR	Peptide

5.2 Close contacts

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	2644	0	0	28	0
1	B	2666	0	0	17	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	289	0	0	7	0
4	B	319	0	0	7	0
All	All	5923	0	0	42	0

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 8.

All (42) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:228:ARG:NH1	1:B:94:GLU:OE2	1.97	0.98
1:A:7:ASP:OD1	1:A:8:THR:N	2.13	0.82
1:A:151:HIS:NE2	1:B:280:GLU:OE2	2.26	0.68
1:B:2:LYS:N	1:B:119:GLU:OE2	2.29	0.66
1:A:1:MET:N	4:A:2001:HOH:O	2.28	0.65
1:A:325[B]:SER:O	4:A:2282:HOH:O	2.13	0.63
1:A:325[A]:SER:O	4:A:2282:HOH:O	2.16	0.63
1:B:24:HIS:CE1	4:B:2038:HOH:O	2.54	0.60
1:B:228:ARG:NH2	4:B:2234:HOH:O	2.36	0.59
1:A:134:ASP:N	1:A:134:ASP:OD1	2.35	0.59
1:A:314:GLU:OE2	1:A:320:ARG:NH1	2.37	0.58
1:B:104:ALA:O	4:B:2132:HOH:O	2.17	0.58
1:B:52:TYR:O	1:B:55:ILE:CG1	2.54	0.56
1:A:97:LYS:NZ	4:A:2105:HOH:O	2.39	0.56
1:A:124:LYS:NZ	4:A:2029:HOH:O	2.40	0.55
1:A:289:ASN:N	1:A:289:ASN:OD1	2.39	0.55
1:A:228:ARG:NH1	1:B:94:GLU:CD	2.62	0.53
1:B:157:GLN:NE2	4:B:2173:HOH:O	2.40	0.53
1:B:134:ASP:N	1:B:134:ASP:OD1	2.42	0.52
1:B:138:GLU:OE1	1:B:320:ARG:NH2	2.45	0.49
1:A:35:SER:O	1:A:344:LEU:N	2.47	0.48
1:B:335:ARG:NE	4:B:2302:HOH:O	2.45	0.48
1:A:2:LYS:N	1:A:119:GLU:OE2	2.46	0.47
1:A:331:ARG:NE	4:A:2014:HOH:O	2.47	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:9:ASP:OD2	1:A:48:ASN:N	2.49	0.46
1:A:304[B]:ARG:NH1	1:A:308:GLU:OE2	2.49	0.45
1:A:56:THR:O	1:A:115:GLY:N	2.49	0.45
1:A:328:GLN:NE2	1:A:332:ASP:OD1	2.49	0.45
1:A:219:GLN:OE1	4:A:2202:HOH:O	2.20	0.45
1:A:256:LEU:O	1:A:284:ILE:N	2.50	0.44
1:B:105:LYS:NZ	4:B:2133:HOH:O	2.49	0.43
1:A:178:GLN:OE1	1:A:312:SER:N	2.51	0.43
1:A:163:ASN:OD1	1:A:187:SER:OG	2.36	0.43
1:A:94:GLU:N	1:A:94:GLU:OE1	2.51	0.42
1:B:127:ASN:ND2	4:B:2141:HOH:O	2.52	0.42
1:A:228:ARG:NH1	1:A:229:PHE:CZ	2.88	0.42
1:A:43:PRO:O	1:A:48:ASN:O	2.38	0.41
1:B:21:GLU:O	1:B:23:LYS:NZ	2.54	0.41
1:B:248:GLU:OE2	1:B:272:THR:CG2	2.69	0.41
1:B:314:GLU:N	1:B:315:PRO:CD	2.84	0.40
1:A:7:ASP:CG	1:A:8:THR:N	2.74	0.40
1:A:58:GLY:O	1:A:116:GLY:N	2.54	0.40

There are no symmetry-related clashes.

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	346/346 (100%)	333 (96%)	12 (4%)	1 (0%)	50	36
1	B	349/346 (101%)	339 (97%)	9 (3%)	1 (0%)	50	36
All	All	695/692 (100%)	672 (97%)	21 (3%)	2 (0%)	50	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	337	ALA
1	A	339	PRO

5.3.2 Protein sidechains ⓘ

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	285/283 (101%)	258 (90%)	27 (10%)	12	4
1	B	288/283 (102%)	265 (92%)	23 (8%)	17	6
All	All	573/566 (101%)	523 (91%)	50 (9%)	15	5

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	VAL
1	A	9	ASP
1	A	14	VAL
1	A	17	SER
1	A	19	ILE
1	A	23	LYS
1	A	25	GLN
1	A	31	LYS
1	A	57	LEU
1	A	72	VAL
1	A	126	LYS
1	A	134	ASP
1	A	164	VAL
1	A	219	GLN
1	A	226	GLU
1	A	227	LEU
1	A	235	GLU
1	A	241	GLN
1	A	256	LEU
1	A	288	MET
1	A	289	ASN
1	A	311	LEU
1	A	320	ARG
1	A	333	ILE
1	A	338	MET
1	A	343	LEU
1	A	344	LEU
1	B	5	VAL
1	B	7	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	13	ARG
1	B	14	VAL
1	B	31	LYS
1	B	42	LEU
1	B	57	LEU
1	B	72	VAL
1	B	105	LYS
1	B	134	ASP
1	B	207	MET
1	B	216[A]	SER
1	B	216[B]	SER
1	B	219	GLN
1	B	225	ARG
1	B	235	GLU
1	B	259	VAL
1	B	297	GLN
1	B	300	GLU
1	B	311	LEU
1	B	335	ARG
1	B	343	LEU
1	B	344	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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.

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, 95th 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	346/346 (100%)	0.28	19 (5%) 24 24	12, 27, 55, 85	0
1	B	346/346 (100%)	0.09	6 (1%) 67 68	12, 25, 43, 64	0
All	All	692/692 (100%)	0.18	25 (3%) 41 41	12, 26, 50, 85	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	10	GLY	4.4
1	A	11	ILE	4.3
1	A	333	ILE	3.9
1	A	336	ASN	3.3
1	A	335	ARG	3.2
1	A	334	ALA	3.0
1	A	337	ALA	3.0
1	A	9	ASP	2.9
1	A	8	THR	2.8
1	A	15	ALA	2.8
1	B	15	ALA	2.6
1	A	7	ASP	2.6
1	A	42	LEU	2.5
1	B	134	ASP	2.4
1	A	61	PHE	2.4
1	A	327	ALA	2.3
1	A	45	ILE	2.2
1	B	344	LEU	2.2
1	B	73	ASP	2.2
1	B	309	ARG	2.2
1	B	337	ALA	2.1
1	A	53	TYR	2.1
1	A	331	ARG	2.1
1	A	338	MET	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	134	ASP	2.0

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	NI	B	1347	1/1	0.08	-1.08	34,34,34,34	0
2	NI	A	1347	1/1	0.07	-2.55	41,41,41,41	0
2	NI	A	1348	1/1	0.03	-2.82	25,25,25,25	0
3	ZN	B	1348	1/1	0.02	-3.30	21,21,21,21	0
3	ZN	A	1349	1/1	0.02	-5.25	19,19,19,19	0

6.5 Other polymers

There are no such residues in this entry.