Full wwPDB X-ray Structure Validation Report (i)

Dec 13, 2017 – 02:31 PM EST

PDB ID	:	5MIN
Title	:	Apo form of the soluble PQQ-dependent Glucose Dehydrogenase from Acinetobacter calcoaceticus
Authors	:	Stines-Chaumeil, C.; Mavre, F.; Limoges, B.; Kauffmann, B.; Mano, N.
Deposited on	:	unknown
Resolution	:	1.76 Å(reported)

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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.
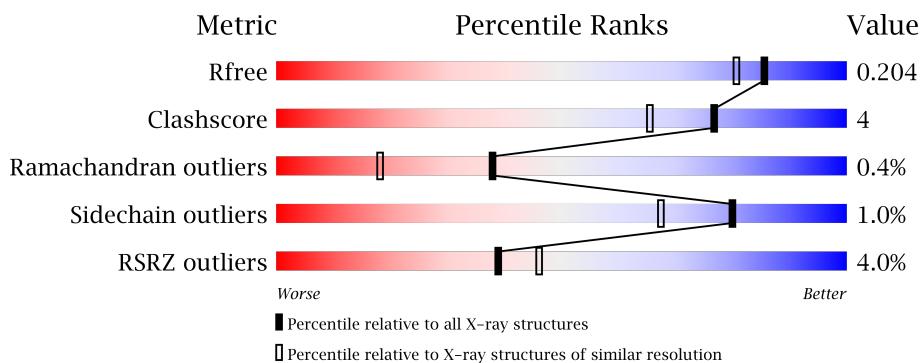
1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION
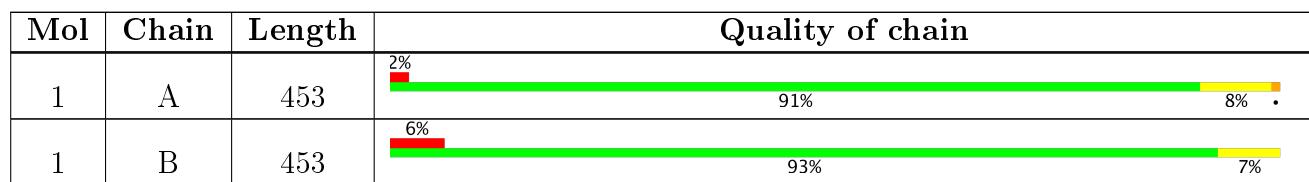
The reported resolution of this entry is 1.76 Å.

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	Percentile Ranks	Value
Rfree		0.204
Clashscore		4
Ramachandran outliers		0.4%
Sidechain outliers		1.0%
RSRZ outliers		4.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	100719	1762 (1.76-1.76)
Clashscore	112137	1889 (1.76-1.76)
Ramachandran outliers	110173	1868 (1.76-1.76)
Sidechain outliers	110143	1868 (1.76-1.76)
RSRZ outliers	101464	1770 (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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	453	2% / 91% / 8% ·
1	B	453	6% / 93% / 7%

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	Res	Chirality	Geometry	Clashes	Electron density
2	CA	B	503	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8241 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 Quinoprotein glucose dehydrogenase B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	452	Total 3575	C 2279	N 606	O 685	S 5	0	8	0
1	B	453	Total 3560	C 2268	N 603	O 684	S 5	0	3	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	109	ALA	GLU	conflict	UNP P13650
B	109	ALA	GLU	conflict	UNP P13650

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total 3	Ca 3	0	0
2	A	3	Total 3	Ca 3	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Cl 1	0	0
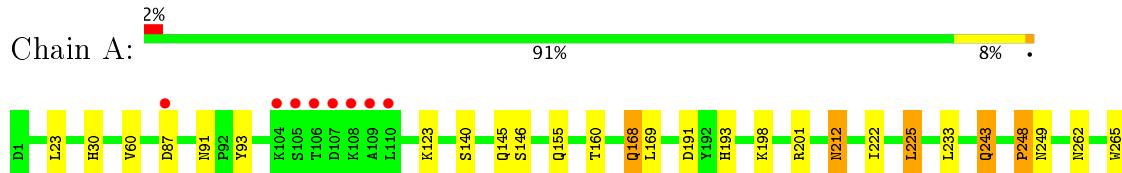
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	552	Total 552	O 552	0	0
4	B	547	Total 547	O 547	0	0
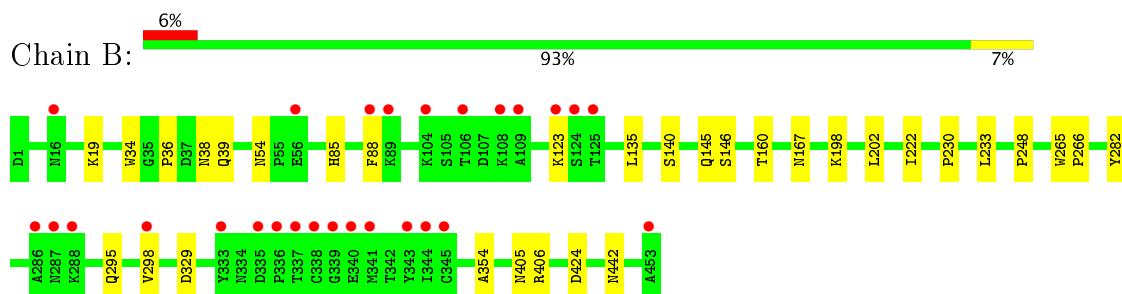
3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Quinoprotein glucose dehydrogenase B

• Molecule 1: Quinoprotein glucose dehydrogenase B

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	54.41Å 54.90Å 85.47Å 88.04° 81.58° 69.62°	Depositor
Resolution (Å)	25.73 – 1.76 25.73 – 1.76	Depositor EDS
% Data completeness (in resolution range)	91.7 (25.73-1.76) 91.3 (25.73-1.76)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.57 (at 1.76Å)	Xtriage
Refinement program	BUSTER-TNT 2.10.2, BUSTER	Depositor
R, R_{free}	0.178 , 0.214 0.169 , 0.204	Depositor DCC
R_{free} test set	4137 reflections (5.22%)	DCC
Wilson B-factor (Å^2)	14.8	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 55.0	EDS
L-test for twinning[2]	$<|L|> = 0.52, <L^2> = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8241	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.91% 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: CA, CL

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	0.50	0/3689	0.64	0/5023
1	B	0.50	0/3657	0.64	0/4982
All	All	0.50	0/7346	0.64	0/10005

There are no bond length outliers.

There are no bond angle outliers.

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	3575	0	3552	30	0
1	B	3560	0	3510	22	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	1	0	0	0	0
4	A	552	0	0	2	0
4	B	547	0	0	3	0
All	All	8241	0	7062	50	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.

All (50) 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:A:405:ASN:HD21	1:A:442:ASN:H	1.21	0.83
1:B:405:ASN:HD21	1:B:442:ASN:H	1.27	0.79
1:B:34:TRP:HE1	1:B:38:ASN:HD22	1.40	0.68
1:A:308:LYS:HB2	1:A:311:GLU:HG3	1.81	0.63
1:B:354:ALA:HB3	4:B:603:HOH:O	1.98	0.63
1:A:23:LEU:HD11	1:A:60[B]:VAL:HG21	1.81	0.61
1:A:328:GLN:HE22	1:B:329:ASP:H	1.51	0.58
1:A:225:LEU:HD11	1:A:262:ASN:ND2	2.19	0.58
1:B:295:GLN:O	1:B:298[B]:VAL:HG22	2.05	0.57
1:B:354:ALA:CB	4:B:603:HOH:O	2.56	0.53
1:A:249:ASN:H	1:A:296:ASN:HD21	1.57	0.53
1:A:243:GLN:C	1:A:243:GLN:HE21	2.13	0.52
1:A:405:ASN:HD21	1:A:442:ASN:N	1.98	0.52
1:B:405:ASN:HD21	1:B:442:ASN:N	2.03	0.51
1:A:30:HIS:HD2	4:A:816:HOH:O	1.94	0.51
1:A:249:ASN:H	1:A:296:ASN:ND2	2.09	0.51
1:A:265:TRP:CD2	1:A:266:PRO:HA	2.46	0.51
1:A:312:TRP:CZ2	1:A:314:GLY:HA3	2.45	0.51
1:B:265:TRP:CD2	1:B:266:PRO:HA	2.46	0.50
1:B:140:SER:HA	1:B:145:GLN:HE22	1.77	0.50
1:A:225:LEU:HD12	1:A:225:LEU:C	2.32	0.49
1:A:191:ASP:OD1	1:A:193:HIS:HD2	1.95	0.49
1:B:405:ASN:ND2	1:B:442:ASN:H	2.04	0.48
1:A:30:HIS:HE1	4:A:759:HOH:O	1.95	0.48
1:B:167:ASN:HD21	1:B:282:TYR:H	1.61	0.48
1:A:312:TRP:CE2	1:A:314:GLY:HA3	2.49	0.48
1:A:328:GLN:NE2	1:B:329:ASP:H	2.10	0.47
1:A:160:THR:HG23	1:A:233:LEU:HD22	1.95	0.47
1:A:201:ARG:HH11	1:A:212:ASN:HD22	1.62	0.47
1:B:160:THR:HB	1:B:230:PRO:HG2	1.95	0.47
1:A:87:ASP:HB3	1:A:91:ASN:HD22	1.81	0.46
1:A:201:ARG:HH11	1:A:212:ASN:ND2	2.14	0.45
1:B:135:LEU:HD22	1:B:202:LEU:HD21	1.98	0.45
1:B:36:PRO:HB3	1:B:88[B]:PHE:CG	2.53	0.44
1:A:23:LEU:HD21	1:A:60[B]:VAL:HG21	1.99	0.44
1:B:85:HIS:O	1:B:88[B]:PHE:CD2	2.71	0.44
1:B:406:ARG:HB2	1:B:424:ASP:HB2	2.00	0.43
1:A:407:TYR:CD1	1:A:421:VAL:HG11	2.54	0.43
1:A:91:ASN:HB3	1:A:93:TYR:CE2	2.54	0.43
1:A:405:ASN:ND2	1:A:442:ASN:H	2.03	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:168[B]:GLN:HG2	1:A:169:LEU:HG	2.02	0.42
1:B:198:LYS:HG2	1:B:222:ILE:HG12	2.01	0.42
1:A:198:LYS:HG2	1:A:222:ILE:HG23	2.02	0.41
1:B:298[B]:VAL:HG21	4:B:648:HOH:O	2.19	0.41
1:B:160:THR:HG23	1:B:233:LEU:HD22	2.01	0.41
1:B:39:GLN:NE2	1:B:54:ASN:HA	2.36	0.41
1:A:248:PRO:HG2	1:A:296:ASN:HD21	1.85	0.41
1:A:140:SER:HA	1:A:145:GLN:HE22	1.85	0.41
1:A:406:ARG:HB2	1:A:424:ASP:HB2	2.01	0.41
1:B:19:LYS:NZ	1:B:442:ASN:HD21	2.18	0.40

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	458/453 (101%)	444 (97%)	12 (3%)	2 (0%)	38	18
1	B	454/453 (100%)	437 (96%)	15 (3%)	2 (0%)	38	18
All	All	912/906 (101%)	881 (97%)	27 (3%)	4 (0%)	38	18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	248	PRO
1	B	248	PRO
1	A	146	SER
1	B	146	SER

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	397/389 (102%)	389 (98%)	8 (2%)	60	38
1	B	392/389 (101%)	391 (100%)	1 (0%)	94	91
All	All	789/778 (101%)	780 (99%)	9 (1%)	80	64

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	123	LYS
1	A	155	GLN
1	A	168[A]	GLN
1	A	168[B]	GLN
1	A	212	ASN
1	A	225	LEU
1	A	243	GLN
1	A	296	ASN
1	B	123	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (23) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	HIS
1	A	91	ASN
1	A	145	GLN
1	A	164	GLN
1	A	193	HIS
1	A	212	ASN
1	A	243	GLN
1	A	255	ASN
1	A	267	ASN
1	A	296	ASN
1	A	328	GLN
1	A	405	ASN
1	A	428	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	442	ASN
1	B	38	ASN
1	B	39	GLN
1	B	76	GLN
1	B	145	GLN
1	B	155	GLN
1	B	167	ASN
1	B	246	GLN
1	B	405	ASN
1	B	442	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 7 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, 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	452/453 (99%)	-0.20	9 (1%)	65	73	8, 15, 35, 70	0
1	B	453/453 (100%)	-0.04	27 (5%)	23	28	8, 16, 38, 63	0
All	All	905/906 (99%)	-0.12	36 (3%)	39	46	8, 16, 36, 70	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	109	ALA	9.7
1	A	106	THR	6.8
1	B	338	CYS	5.7
1	B	336	PRO	5.5
1	B	88[A]	PHE	5.5
1	A	108	LYS	5.1
1	B	340	GLU	5.1
1	B	286	ALA	4.6
1	A	107	ASP	4.4
1	B	341	MET	4.4
1	B	343	TYR	4.3
1	B	109	ALA	4.0
1	B	339	GLY	4.0
1	B	345	CYS	3.9
1	B	108	LYS	3.8
1	B	453	ALA	3.8
1	B	106	THR	3.5
1	B	344	ILE	3.5
1	B	16	ASN	3.5
1	B	124	SER	3.2
1	B	335	ASP	3.2
1	B	337	THR	3.1
1	B	56	GLU	3.0
1	B	89	LYS	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	104	LYS	2.7
1	A	105[A]	SER	2.7
1	A	87	ASP	2.6
1	A	110	LEU	2.6
1	B	123	LYS	2.5
1	B	288	LYS	2.3
1	B	104	LYS	2.3
1	B	287	ASN	2.1
1	A	340	GLU	2.1
1	B	298[A]	VAL	2.0
1	B	125	THR	2.0
1	B	333	TYR	2.0

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. 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	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	CA	B	503	1/1	0.97	0.19	8.92	41,41,41,41	0
2	CA	B	501	1/1	0.99	0.07	-0.24	10,10,10,10	0
2	CA	A	502	1/1	1.00	0.06	-1.25	11,11,11,11	0
2	CA	A	501	1/1	1.00	0.05	-1.82	8,8,8,8	0
2	CA	B	502	1/1	1.00	0.04	-2.54	11,11,11,11	0
2	CA	A	503	1/1	0.99	0.04	-3.10	27,27,27,27	0
3	CL	A	504	1/1	0.95	0.20	-	37,37,37,37	0

6.5 Other polymers (i)

There are no such residues in this entry.