WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 17, 2023 – 06:18 PM EST

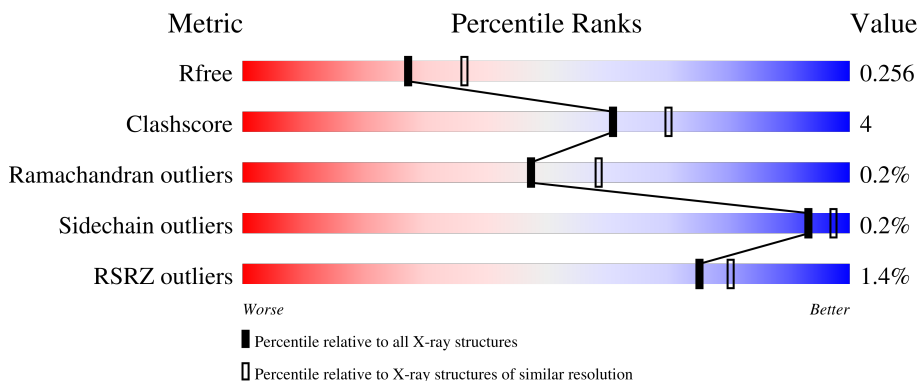|		
PDB ID	:	2GB5
Title	:	Crystal structure of NADH pyrophosphatase (EC 3.6.1.22) (1790429) from Escherichia coli K12 at 2.30 A resolution
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2006-03-09
Resolution	:	2.30 Å(reported)

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 2.30 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		0.256
Clashscore		4
Ramachandran outliers		0.2%
Sidechain outliers		0.2%
RSRZ outliers		1.4%

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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 >=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
1	A	269	88% 10% ·
1	B	269	89% 7% ·

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4458 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 NADH pyrophosphatase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	265	Total	C	N	O	S	Se	0	4	0
			2163	1377	384	388	6	8			
1	B	261	Total	C	N	O	S	Se	1	1	0
			2053	1315	352	372	6	8			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MSE	-	expression tag	UNP P32664
A	-10	GLY	-	expression tag	UNP P32664
A	-9	SER	-	expression tag	UNP P32664
A	-8	ASP	-	expression tag	UNP P32664
A	-7	LYS	-	expression tag	UNP P32664
A	-6	ILE	-	expression tag	UNP P32664
A	-5	HIS	-	expression tag	UNP P32664
A	-4	HIS	-	expression tag	UNP P32664
A	-3	HIS	-	expression tag	UNP P32664
A	-2	HIS	-	expression tag	UNP P32664
A	-1	HIS	-	expression tag	UNP P32664
A	0	HIS	-	expression tag	UNP P32664
A	1	MSE	MET	modified residue	UNP P32664
A	65	MSE	MET	modified residue	UNP P32664
A	105	MSE	MET	modified residue	UNP P32664
A	114	MSE	MET	modified residue	UNP P32664
A	176	MSE	MET	modified residue	UNP P32664
A	201	MSE	MET	modified residue	UNP P32664
A	205	MSE	MET	modified residue	UNP P32664
A	251	MSE	MET	modified residue	UNP P32664
B	-11	MSE	-	expression tag	UNP P32664
B	-10	GLY	-	expression tag	UNP P32664
B	-9	SER	-	expression tag	UNP P32664
B	-8	ASP	-	expression tag	UNP P32664
B	-7	LYS	-	expression tag	UNP P32664

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	ILE	-	expression tag	UNP P32664
B	-5	HIS	-	expression tag	UNP P32664
B	-4	HIS	-	expression tag	UNP P32664
B	-3	HIS	-	expression tag	UNP P32664
B	-2	HIS	-	expression tag	UNP P32664
B	-1	HIS	-	expression tag	UNP P32664
B	0	HIS	-	expression tag	UNP P32664
B	1	MSE	MET	modified residue	UNP P32664
B	65	MSE	MET	modified residue	UNP P32664
B	105	MSE	MET	modified residue	UNP P32664
B	114	MSE	MET	modified residue	UNP P32664
B	176	MSE	MET	modified residue	UNP P32664
B	201	MSE	MET	modified residue	UNP P32664
B	205	MSE	MET	modified residue	UNP P32664
B	251	MSE	MET	modified residue	UNP P32664

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	134	Total O 134 134	0	0
3	B	106	Total O 106 106	0	0

3 Residue-property plots ⓘ

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: NADH pyrophosphatase
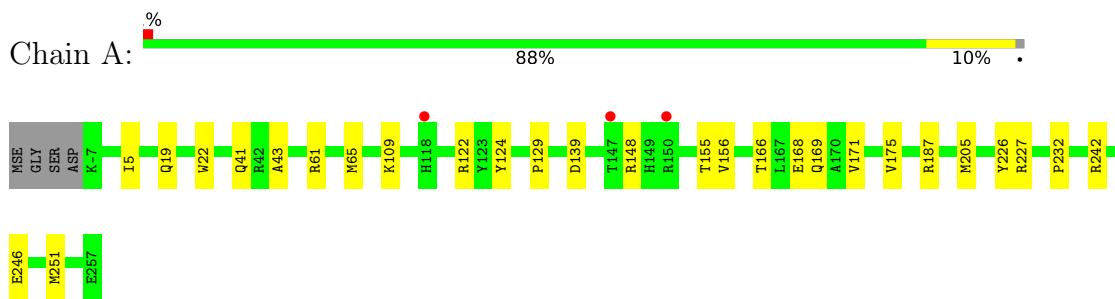
• Molecule 1: NADH pyrophosphatase
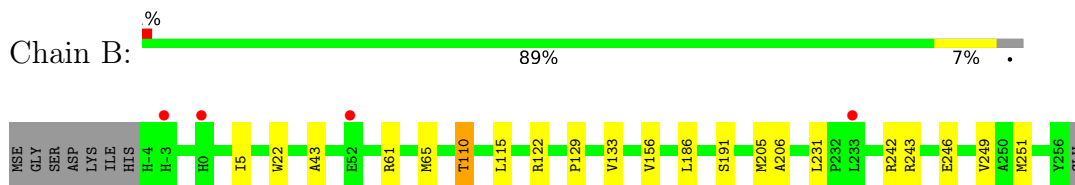
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	99.29Å 103.55Å 57.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.61 – 2.30 49.61 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.61-2.30) 99.0 (49.61-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.198 , 0.252 0.204 , 0.256	Depositor DCC
R_{free} test set	1345 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	40.5	Xtriage
Anisotropy	0.538	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 35.7	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.023 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4458	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.64% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.53	0/2219	0.64	0/3005
1	B	0.53	1/2102 (0.0%)	0.60	0/2849
All	All	0.53	1/4321 (0.0%)	0.63	0/5854

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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	110	THR	C-O	-8.36	1.07	1.23

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	110	THR	Mainchain

5.2   Too-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 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	2163	0	2064	23	0
1	B	2053	0	1926	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	134	0	0	2	0
3	B	106	0	0	0	0
All	All	4458	0	3990	37	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 37 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:61:ARG:HB2	1:A:65:MSE:HE3	1.54	0.89
1:B:231:LEU:HD11	1:B:249:VAL:HG21	1.61	0.83
1:B:61:ARG:HB2	1:B:65:MSE:HE3	1.63	0.79
1:A:5[A]:ILE:HD11	1:A:43:ALA:HB2	1.78	0.65
1:B:22:TRP:CD1	1:B:65:MSE:HE1	2.35	0.62

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	267/269 (99%)	257 (96%)	9 (3%)	1 (0%)	34	42
1	B	260/269 (97%)	250 (96%)	10 (4%)	0	100	100
All	All	527/538 (98%)	507 (96%)	19 (4%)	1 (0%)	47	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	ARG

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	225/223 (101%)	225 (100%)	0	100	100
1	B	204/223 (92%)	203 (100%)	1 (0%)	88	95
All	All	429/446 (96%)	428 (100%)	1 (0%)	93	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	243	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	-4	HIS
1	A	17	HIS
1	B	185	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

wwPDB X-ray Structure Validation Summary Report

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 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 ⓘ

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	257/269 (95%)	-0.01	3 (1%)	79	83	28, 40, 54, 68	0
1	B	253/269 (94%)	0.10	4 (1%)	72	77	29, 39, 54, 68	1 (0%)
All	All	510/538 (94%)	0.04	7 (1%)	75	80	28, 40, 54, 68	1 (0%)

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	52	GLU	2.7
1	A	150	ARG	2.5
1	B	-3	HIS	2.4
1	B	233	LEU	2.2
1	B	0	HIS	2.1

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 monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
2	ZN	A	301	1/1	0.98	0.03	44,44,44,44	0
2	ZN	B	301	1/1	1.00	0.04	43,43,43,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.