Full wwPDB X-ray Structure Validation Report (i)

May 23, 2024 – 03:49 PM EDT

PDB ID	:	4PXL
Title	:	Structure of Zm ALDH2-3 (RF2C) in complex with NAD
Authors	:	Morera, S.; Vigouroux, A.; Kopecny, D.
Deposited on	:	2014-03-24
Resolution	:	2.25 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

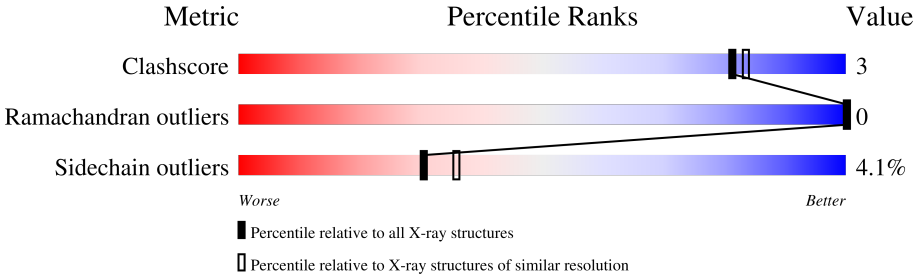
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.25 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	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 7811 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 Cytosolic aldehyde dehydrogenase RF2C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	486	3699	2366	622	694	17	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-14	MET	-	INITIATING METHIONINE	UNP Q8S531
A	-13	GLY	-	EXPRESSION TAG	UNP Q8S531
A	-12	SER	-	EXPRESSION TAG	UNP Q8S531
A	-11	SER	-	EXPRESSION TAG	UNP Q8S531
A	-10	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-9	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-8	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-7	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-6	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-5	HIS	-	EXPRESSION TAG	UNP Q8S531
A	-4	SER	-	EXPRESSION TAG	UNP Q8S531
A	-3	GLN	-	EXPRESSION TAG	UNP Q8S531
A	-2	ASP	-	EXPRESSION TAG	UNP Q8S531
A	-1	PRO	-	EXPRESSION TAG	UNP Q8S531
A	0	ASN	-	EXPRESSION TAG	UNP Q8S531
A	1	SER	-	EXPRESSION TAG	UNP Q8S531
A	57	ASP	GLY	SEE REMARK 999	UNP Q8S531
A	?	-	ASN	DELETION	UNP Q8S531

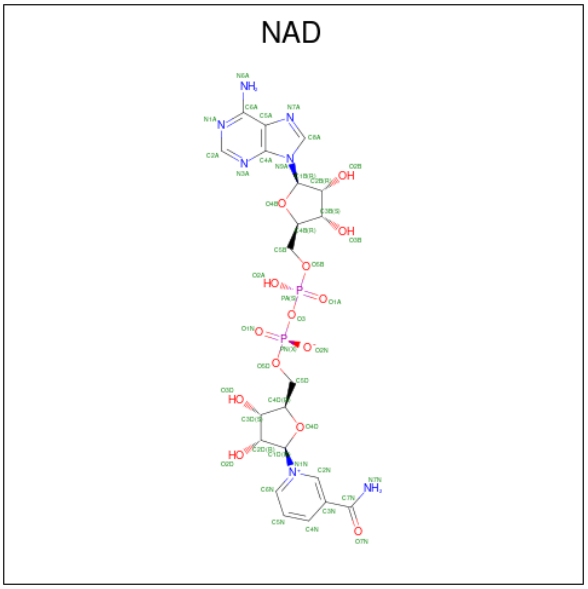
- Molecule 2 is a protein called Cytosolic aldehyde dehydrogenase RF2C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	486	3698	2366	622	693	17	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

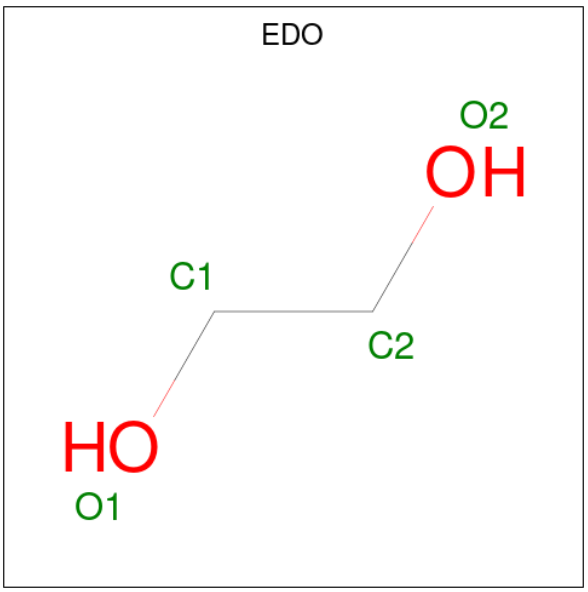
Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	MET	-	INITIATING METHIONINE	UNP Q8S531
B	-13	GLY	-	EXPRESSION TAG	UNP Q8S531
B	-12	SER	-	EXPRESSION TAG	UNP Q8S531
B	-11	SER	-	EXPRESSION TAG	UNP Q8S531
B	-10	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-9	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-8	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-7	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-6	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-5	HIS	-	EXPRESSION TAG	UNP Q8S531
B	-4	SER	-	EXPRESSION TAG	UNP Q8S531
B	-3	GLN	-	EXPRESSION TAG	UNP Q8S531
B	-2	ASP	-	EXPRESSION TAG	UNP Q8S531
B	-1	PRO	-	EXPRESSION TAG	UNP Q8S531
B	0	ASN	-	EXPRESSION TAG	UNP Q8S531
B	1	SER	-	EXPRESSION TAG	UNP Q8S531
B	57	ASP	GLY	SEE REMARK 999	UNP Q8S531
B	?	-	ASN	DELETION	UNP Q8S531

- Molecule 3 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: $C_{21}H_{27}N_7O_{14}P_2$).



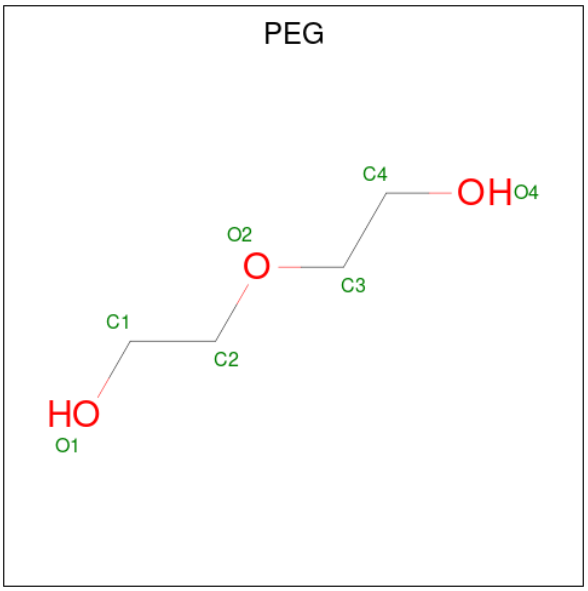
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	44	21	7	14	2	0	0
			Total	C	N	O	P		
3	B	1	44	21	7	14	2	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 7 4 3	0	0
5	A	1	Total C O 7 4 3	0	0
5	A	1	Total C O 7 4 3	0	0
5	A	1	Total C O 7 4 3	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Na 1 1	0	0
6	B	1	Total Na 1 1	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Ca 1 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	135	135	135	0	0
			Total	O		
8	B	124	124	124	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	109.86Å 126.08Å 78.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.10 – 2.25	Depositor
% Data completeness (in resolution range)	99.9 (49.10-2.25)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 2.24Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.171 , 0.205	Depositor
Wilson B-factor (Å^2)	36.0	Xtriage
Anisotropy	0.555	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7811	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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.

4 Model quality (i)

4.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, NAD, NA, PEG, CSO, CA

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	0.51	0/3773	0.68	0/5118
2	B	0.50	0/3780	0.68	0/5129
All	All	0.51	0/7553	0.68	0/10247

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	3699	0	3686	20	0
2	B	3698	0	3686	22	0
3	A	44	0	26	0	0
3	B	44	0	26	0	0
4	A	20	0	30	1	0
4	B	16	0	24	6	0
5	A	28	0	40	2	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	1	0	0	0	0
8	A	135	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	124	0	0	0	0
All	All	7811	0	7518	39	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 3.

All (39) 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)
2:B:441:ASN:HB2	4:B:602:EDO:H12	1.71	0.73
2:B:79:ARG:HE	4:B:605:EDO:H22	1.53	0.72
2:B:436:ASN:HD22	2:B:439:VAL:H	1.34	0.72
1:A:436:ASN:HD22	1:A:439:VAL:H	1.35	0.71
2:B:441:ASN:HB2	4:B:602:EDO:C1	2.29	0.62
2:B:362:LYS:HE2	2:B:368:LEU:HD22	1.83	0.59
2:B:79:ARG:HE	4:B:605:EDO:C2	2.16	0.59
2:B:165:HIS:HD2	2:B:243:SER:HB3	1.68	0.58
1:A:256:ARG:HG3	2:B:256:ARG:HG3	1.90	0.53
1:A:242:VAL:CG1	1:A:266:VAL:HG12	2.39	0.52
2:B:115:PHE:CZ	2:B:119:LYS:HD2	2.44	0.52
1:A:115:PHE:CZ	1:A:119:LYS:HD2	2.45	0.52
1:A:429:ALA:O	4:A:604:EDO:H21	2.10	0.51
1:A:252:ARG:NH2	2:B:261:SER:O	2.44	0.49
1:A:329:LYS:NZ	5:A:603:PEG:H21	2.28	0.49
1:A:261:SER:O	2:B:252:ARG:NH2	2.45	0.47
2:B:438:ASP:HA	4:B:602:EDO:H12	1.96	0.47
1:A:167:VAL:HG23	1:A:194:PRO:HA	1.96	0.46
2:B:79:ARG:NE	4:B:605:EDO:H22	2.24	0.45
1:A:450:GLY:HA3	1:A:467:GLY:O	2.17	0.45
2:B:302:ILE:HD12	2:B:305:ALA:HB2	1.99	0.45
2:B:240:ASP:O	2:B:265:PRO:HD2	2.16	0.44
2:B:450:GLY:HA3	2:B:467:GLY:O	2.17	0.44
1:A:165:HIS:HD2	1:A:243:SER:HB3	1.82	0.44
2:B:436:ASN:ND2	2:B:439:VAL:H	2.10	0.43
2:B:351:TYR:HB2	2:B:380:TYR:HB3	2.01	0.43
1:A:179:LYS:NZ	8:A:828:HOH:O	2.49	0.43
1:A:110:ASP:OD2	1:A:198:THR:HA	2.19	0.42
1:A:241:LYS:HG3	1:A:265:PRO:HG2	2.01	0.42
1:A:275:PRO:HA	1:A:308:ARG:O	2.19	0.42
1:A:297:THR:O	1:A:302:ILE:HD11	2.21	0.41
2:B:275:PRO:HA	2:B:308:ARG:O	2.19	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:242:VAL:HG12	1:A:266:VAL:HG12	2.02	0.41
2:B:436:ASN:HD21	2:B:438:ASP:HB2	1.84	0.41
2:B:295:THR:HG21	2:B:406:MET:HB2	2.02	0.41
1:A:436:ASN:HD21	1:A:438:ASP:HB2	1.85	0.41
1:A:165:HIS:CD2	1:A:179:LYS:HB3	2.56	0.41
2:B:165:HIS:CD2	2:B:179:LYS:HB3	2.56	0.40
1:A:329:LYS:HZ2	5:A:603:PEG:H21	1.87	0.40

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	483/517 (93%)	474 (98%)	9 (2%)	0	100	100
2	B	484/517 (94%)	476 (98%)	8 (2%)	0	100	100
All	All	967/1034 (94%)	950 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	381/407 (94%)	366 (96%)	15 (4%)	32	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	382/408 (94%)	366 (96%)	16 (4%)	30	34
All	All	763/815 (94%)	732 (96%)	31 (4%)	30	36

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	ARG
1	A	40	GLU
1	A	52	SER
1	A	110	ASP
1	A	121	ARG
1	A	193	LYS
1	A	211	GLU
1	A	269	GLU
1	A	299	LYS
1	A	302	ILE
1	A	329	LYS
1	A	377	ASP
1	A	402	PHE
1	A	481	ASP
1	A	495	PRO
2	B	20	ARG
2	B	40	GLU
2	B	52	SER
2	B	110	ASP
2	B	160	VAL
2	B	193	LYS
2	B	211	GLU
2	B	269	GLU
2	B	284	LEU
2	B	285	ASP
2	B	289	ASN
2	B	299	LYS
2	B	302	ILE
2	B	402	PHE
2	B	481	ASP
2	B	488	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	99	HIS
1	A	350	GLN
1	A	436	ASN
2	B	99	HIS
2	B	128	HIS
2	B	436	ASN
2	B	488	GLN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 18 ligands modelled in this entry, 3 are monoatomic - leaving 15 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.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates (i)

EDS failed to run properly - this section is therefore empty.

5.4 Ligands (i)

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers (i)

EDS failed to run properly - this section is therefore empty.