wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2023 – 02:56 pm GMT

PDB ID	:	2WFL
Title	:	Crystal structure of polyneuridine aldehyde esterase
Authors	:	Yang, L.; Hill, M.; Panjikar, S.; Wang, M.; Stoeckigt, J.
Deposited on	:	2009-04-08
Resolution	:	2.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

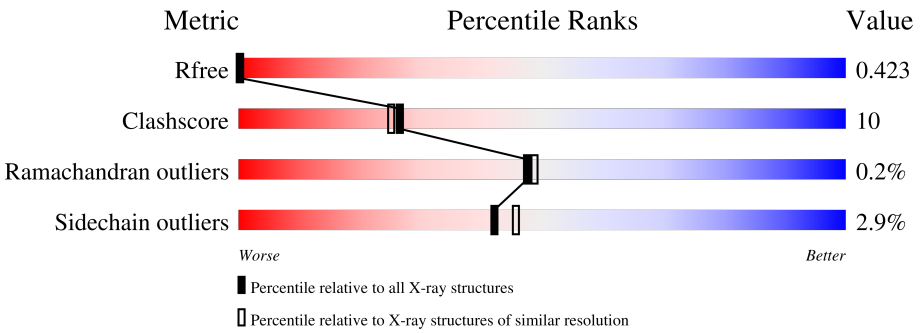
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.10 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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%

Mol	Chain	Length	Quality of chain
1	A	264	77% 17% • •
2	B	264	73% 21% • •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4347 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.

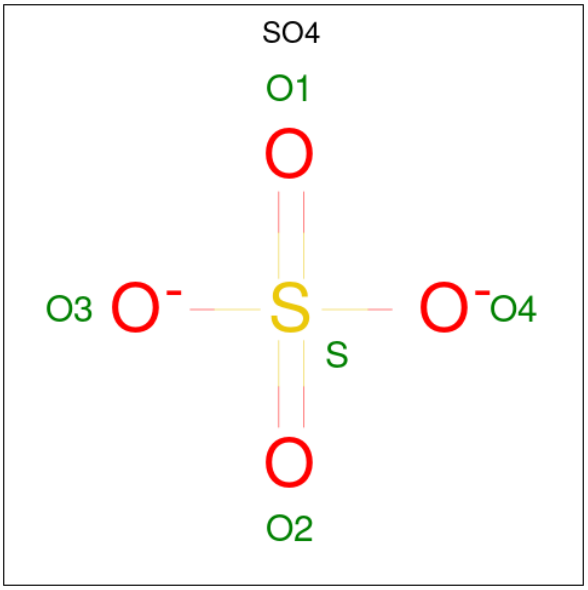
- Molecule 1 is a protein called POLYNEURIDINE-ALDEHYDE ESTERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	255	2015	1292	328	372	23	0	0	0

- Molecule 2 is a protein called POLYNEURIDINE-ALDEHYDE ESTERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	255	2019	1294	328	373	24	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
3	A	1	5	4	1	0	0
3	B	1	5	4	1	0	0

- Molecule 4 is water.

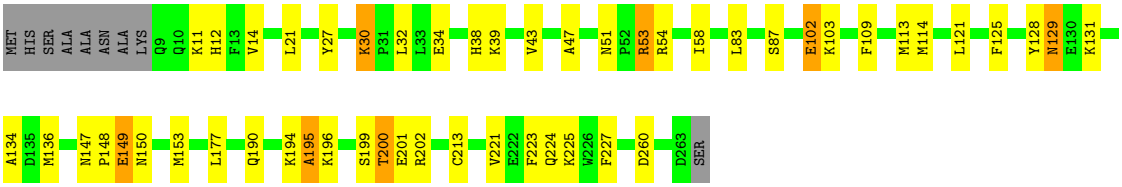
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	150	150	150	0	0
			Total	O		
4	B	153	153	153	0	0

3 Residue-property plots (i)

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. 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. 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.

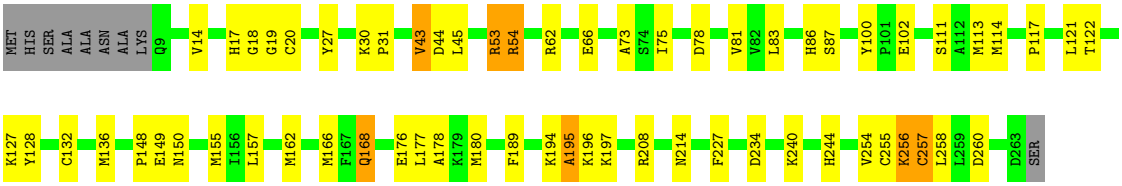
• Molecule 1: POLYNEURIDINE-ALDEHYDE ESTERASE

Chain A:



• Molecule 2: POLYNEURIDINE-ALDEHYDE ESTERASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	92.76Å 176.92Å 75.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10 19.83 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.10) 100.0 (19.83-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.98 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.4.0069	Depositor
R, R_{free}	0.182 , 0.231 0.385 , 0.423	Depositor DCC
R_{free} test set	1046 reflections (2.85%)	wwPDB-VP
Wilson B-factor (Å^2)	35.2	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 42.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.75	EDS
Total number of atoms	4347	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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: CME, SO4

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	0.94	7/2055 (0.3%)	0.89	5/2769 (0.2%)
2	B	0.99	5/2048 (0.2%)	0.92	7/2758 (0.3%)
All	All	0.96	12/4103 (0.3%)	0.90	12/5527 (0.2%)

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
2	B	0	3
All	All	0	4

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	53	ARG	CG-CD	-8.86	1.29	1.51
1	A	43	VAL	CB-CG2	-7.92	1.36	1.52
1	A	53	ARG	CG-CD	-7.77	1.32	1.51
1	A	102	GLU	CD-OE1	-6.91	1.18	1.25
2	B	53	ARG	CD-NE	-6.80	1.34	1.46

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	53	ARG	NE-CZ-NH1	9.36	124.98	120.30
2	B	234	ASP	CB-CG-OD2	8.06	125.56	118.30
1	A	53	ARG	NE-CZ-NH1	8.06	124.33	120.30
2	B	43	VAL	CG1-CB-CG2	-7.01	99.69	110.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	102	GLU	OE1-CD-OE2	-6.07	116.01	123.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	195	ALA	Peptide
2	B	195	ALA	Peptide
2	B	256	LYS	Mainchain
2	B	257	CME	Mainchain

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	2015	0	1991	40	70
2	B	2019	0	1994	50	72
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	150	0	0	2	7
4	B	153	0	0	5	4
All	All	4347	0	3985	83	77

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 10.

The worst 5 of 83 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:54:ARG:NH1	4:B:2035:HOH:O	1.87	1.08
1:A:131:LYS:HD2	1:A:131:LYS:O	1.55	1.05
2:B:114:MET:CE	2:B:208:ARG:HD2	1.87	1.04
1:A:21:LEU:HD12	4:A:2060:HOH:O	1.62	0.97
2:B:114:MET:HE1	2:B:208:ARG:HD2	1.49	0.93

The worst 5 of 77 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	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:153:MET:CG	2:B:78:ASP:OD2[6_564]	0.53	1.67
1:A:153:MET:CE	2:B:78:ASP:CB[6_564]	0.55	1.65
1:A:201:GLU:OE1	2:B:150:ASN:N[3_655]	0.57	1.63
1:A:134:ALA:CB	2:B:73:ALA:O[6_564]	0.58	1.62
1:A:201:GLU:CD	2:B:149:GLU:C[3_655]	0.66	1.54

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	252/264 (96%)	242 (96%)	9 (4%)	1 (0%)	34	32
2	B	251/264 (95%)	242 (96%)	9 (4%)	0	100	100
All	All	503/528 (95%)	484 (96%)	18 (4%)	1 (0%)	47	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	196	LYS

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	221/227 (97%)	213 (96%)	8 (4%)	35	36
2	B	220/226 (97%)	215 (98%)	5 (2%)	50	55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	441/453 (97%)	428 (97%)	13 (3%)	42 46

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	260	ASP
2	B	43	VAL
2	B	168	GLN
2	B	54	ARG
2	B	62	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	17	HIS
2	B	59	HIS
2	B	168	GLN
2	B	86	HIS
1	A	150	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)

3 non-standard protein/DNA/RNA residues 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	CME	A	255	1	8,9,10	0.59	0	5,9,11	1.41	0
2	CME	B	255	2	8,9,10	0.51	0	5,9,11	1.42	1 (20%)
2	CME	B	257	2	8,9,10	0.44	0	5,9,11	1.68	1 (20%)

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
1	CME	A	255	1	-	2/5/8/10	-
2	CME	B	255	2	-	1/5/8/10	-
2	CME	B	257	2	-	2/5/8/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	257	CME	CE-SD-SG	3.19	118.15	103.45
2	B	255	CME	CB-SG-SD	2.56	110.44	103.82

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	255	CME	CZ-CE-SD-SG
1	A	255	CME	SD-CE-CZ-OH
2	B	255	CME	CE-SD-SG-CB
2	B	257	CME	CZ-CE-SD-SG
2	B	257	CME	CA-CB-SG-SD

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SO4	A	1264	-	4,4,4	0.26	0	6,6,6	0.59	0
3	SO4	B	1264	-	4,4,4	0.79	0	6,6,6	1.37	1 (16%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1264	SO4	O4-S-O3	3.02	121.94	109.06

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.