



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:58 PM GMT

PDB ID : 1Z9A
Title : Crystal Structure Of The Asn-309 To Asp Mutant Of Candida Tenuis Xylose Reductase (Akr2B5) Bound To Nad+
Authors : Kratzer, R.; Leitgeb, S.; Wilson, D.K.; Nidetzky, B.
Deposited on : 2005-04-01
Resolution : 2.40 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

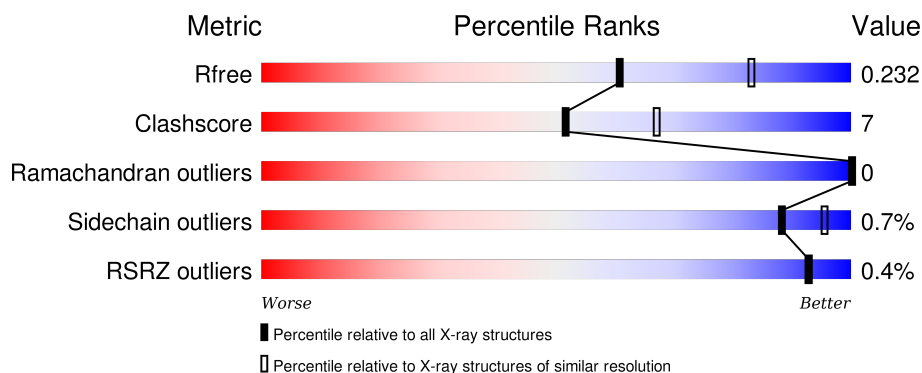
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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}	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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 $\leq 5\%$. 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	321	<div> <div></div> <div>82%16%..</div> </div>
1	B	321	<div> <div></div> <div>85%15%.</div> </div>
1	C	321	<div> <div></div> <div>83%16%..</div> </div>
1	D	321	<div> <div></div> <div>89%11%.</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11093 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 NAD(P)H-dependent D-xylose reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2532	1645	418	465	4			
1	B	319	Total	C	N	O	S	0	0	0
			2532	1645	418	465	4			
1	C	319	Total	C	N	O	S	0	0	0
			2532	1645	418	465	4			
1	D	319	Total	C	N	O	S	0	0	0
			2532	1645	418	465	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	310	ASP	ASN	ENGINEERED	UNP O74237
B	310	ASP	ASN	ENGINEERED	UNP O74237
C	310	ASP	ASN	ENGINEERED	UNP O74237
D	310	ASP	ASN	ENGINEERED	UNP O74237

- Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C₂₁H₂₇N₇O₁₄P₂).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	171	Total	O	0	0
			171	171		
3	B	185	Total	O	0	0
			185	185		
3	C	190	Total	O	0	0
			190	190		
3	D	243	Total	O	0	0
			243	243		

- Molecule 1: NAD(P)H-dependent D-xylose reductase



Q283	I288	F309	I320	F321	V322
------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	180.63Å 128.40Å 80.00Å 90.00° 90.70° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40 29.79 – 2.40	Depositor EDS
% Data completeness (in resolution range)	90.0 (30.00-2.40) 89.9 (29.79-2.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.49 (at 2.39Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.176 , 0.231 0.176 , 0.232	Depositor DCC
R_{free} test set	3160 reflections (4.93%)	DCC
Wilson B-factor (Å ²)	19.2	Xtriage
Anisotropy	0.484	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 40.5	EDS
Estimated twinning fraction	0.034 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 64135 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11093	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: NAD

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	0.38	0/2598	0.60	1/3528 (0.0%)
1	B	0.40	0/2598	0.61	0/3528
1	C	0.39	0/2598	0.60	0/3528
1	D	0.42	0/2598	0.62	0/3528
All	All	0.40	0/10392	0.61	1/14112 (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	B	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	113	ILE	N-CA-C	-5.14	97.12	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	217	TYR	Sidechain
1	C	217	TYR	Sidechain

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	2532	0	2523	45	0
1	B	2532	0	2523	31	0
1	C	2532	0	2523	42	0
1	D	2532	0	2523	22	0
2	A	44	0	26	8	0
2	B	44	0	26	3	0
2	C	44	0	26	1	0
2	D	44	0	26	2	0
3	A	171	0	0	2	0
3	B	185	0	0	5	0
3	C	190	0	0	4	0
3	D	243	0	0	3	0
All	All	11093	0	10196	140	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 7.

The worst 5 of 140 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:C:241:ALA:HA	1:C:246:LYS:HE3	1.51	0.92
1:C:294:LYS:O	1:C:298:GLU:HG3	1.85	0.76
1:A:294:LYS:O	1:A:298:GLU:HG3	1.87	0.74
1:C:149:GLU:HG3	3:C:3433:HOH:O	1.89	0.72
1:C:223:GLN:O	1:C:226:VAL:HG22	1.90	0.71

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	317/321 (99%)	304 (96%)	13 (4%)	0	100	100
1	B	317/321 (99%)	308 (97%)	9 (3%)	0	100	100
1	C	317/321 (99%)	308 (97%)	9 (3%)	0	100	100
1	D	317/321 (99%)	311 (98%)	6 (2%)	0	100	100
All	All	1268/1284 (99%)	1231 (97%)	37 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	268/269 (100%)	264 (98%)	4 (2%)	72	87
1	B	268/269 (100%)	267 (100%)	1 (0%)	93	98
1	C	268/269 (100%)	266 (99%)	2 (1%)	88	95
1	D	268/269 (100%)	267 (100%)	1 (0%)	93	98
All	All	1072/1076 (100%)	1064 (99%)	8 (1%)	88	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	303	LEU
1	D	217	TYR
1	C	274	LYS
1	A	246	LYS
1	B	256	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	34	GLN
1	C	91	ASN
1	D	229	ASN
1	B	229	ASN
1	D	91	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](#)

4 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	$\# Z > 2$
2	NAD	A	1350	-	38,48,48	1.91	10 (26%)	47,73,73	1.70	7 (14%)
2	NAD	B	2350	-	38,48,48	1.85	9 (23%)	47,73,73	1.89	9 (19%)
2	NAD	C	3350	-	38,48,48	1.95	10 (26%)	47,73,73	1.83	8 (17%)
2	NAD	D	4350	-	38,48,48	1.95	10 (26%)	47,73,73	1.88	9 (19%)

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
2	NAD	A	1350	-	-	0/22/62/62	0/5/5/5
2	NAD	B	2350	-	-	0/22/62/62	0/5/5/5
2	NAD	C	3350	-	-	0/22/62/62	0/5/5/5
2	NAD	D	4350	-	-	0/22/62/62	0/5/5/5

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2350	NAD	PA-O1A	-3.07	1.40	1.51
2	C	3350	NAD	PA-O1A	-2.71	1.41	1.51
2	A	1350	NAD	PA-O1A	-2.64	1.41	1.51
2	D	4350	NAD	PA-O1A	-2.63	1.41	1.51
2	D	4350	NAD	C5N-C4N	2.01	1.43	1.38

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2350	NAD	N3A-C2A-N1A	-7.88	122.86	128.89
2	C	3350	NAD	N3A-C2A-N1A	-7.74	122.97	128.89
2	A	1350	NAD	N3A-C2A-N1A	-7.57	123.09	128.89
2	D	4350	NAD	N3A-C2A-N1A	-7.57	123.10	128.89
2	C	3350	NAD	C1B-N9A-C4A	-3.56	121.57	126.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1350	NAD	8	0
2	B	2350	NAD	3	0
2	C	3350	NAD	1	0
2	D	4350	NAD	2	0

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 [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	319/321 (99%)	-0.49	2 (0%) 90 90	11, 23, 46, 67	0
1	B	319/321 (99%)	-0.57	1 (0%) 94 94	10, 20, 32, 39	0
1	C	319/321 (99%)	-0.56	2 (0%) 90 90	8, 21, 43, 52	0
1	D	319/321 (99%)	-0.82	0 100 100	7, 15, 28, 36	0
All	All	1276/1284 (99%)	-0.61	5 (0%) 93 93	7, 20, 40, 67	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	229	ASN	3.0
1	C	253	ASN	2.4
1	C	4	SER	2.4
1	B	138	ASN	2.3
1	A	279	GLU	2.3

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(\AA^2)	Q<0.9
2	NAD	A	1350	44/44	0.93	0.15	0.23	29,44,53,54	0
2	NAD	B	2350	44/44	0.98	0.13	-0.06	11,16,21,23	0
2	NAD	C	3350	44/44	0.96	0.12	-0.37	30,35,39,42	0
2	NAD	D	4350	44/44	0.98	0.10	-0.50	7,12,17,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.