



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 7, 2017 – 01:19 AM EST

PDB ID : 4JNZ
Title : Crystal structure of PutA86-630 mutant D370N complexed with L-Tetrahydr
o-2-furoic acid
Authors : Tanner, J.J.
Deposited on : unknown
Resolution : 1.85 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20030345

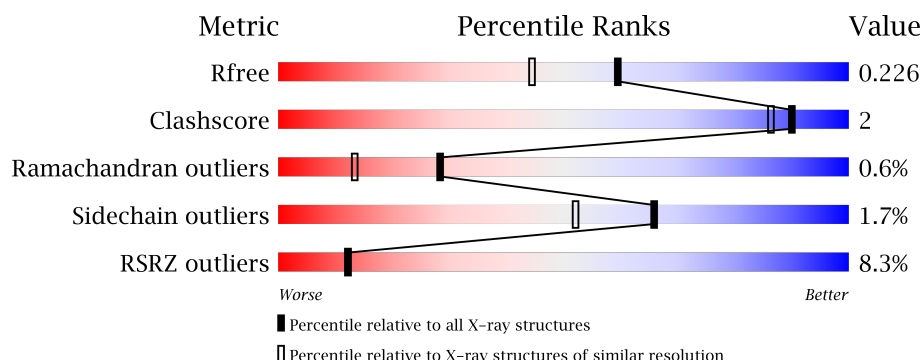
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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}	100719	1923 (1.86-1.86)
Clashscore	112137	2083 (1.86-1.86)
Ramachandran outliers	110173	2060 (1.86-1.86)
Sidechain outliers	110143	2060 (1.86-1.86)
RSRZ outliers	101464	1932 (1.86-1.86)

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	602	

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
4	1PE	A	2004	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4073 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 Bifunctional protein PutA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	493	3728	2371	652	687	18	0	3	0

There are 19 discrepancies between the modelled and reference sequences:

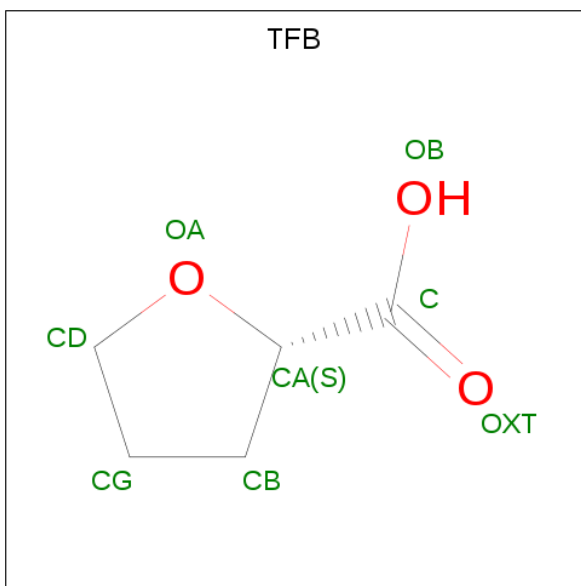
Chain	Residue	Modelled	Actual	Comment	Reference
A	370	ASN	ASP	ENGINEERED MUTATION	UNP P09546
A	670	SER	-	EXPRESSION TAG	UNP P09546
A	671	SER	-	EXPRESSION TAG	UNP P09546
A	672	SER	-	EXPRESSION TAG	UNP P09546
A	673	VAL	-	EXPRESSION TAG	UNP P09546
A	674	ASP	-	EXPRESSION TAG	UNP P09546
A	675	LYS	-	EXPRESSION TAG	UNP P09546
A	676	LEU	-	EXPRESSION TAG	UNP P09546
A	677	ALA	-	EXPRESSION TAG	UNP P09546
A	678	ALA	-	EXPRESSION TAG	UNP P09546
A	679	ALA	-	EXPRESSION TAG	UNP P09546
A	680	LEU	-	EXPRESSION TAG	UNP P09546
A	681	GLU	-	EXPRESSION TAG	UNP P09546
A	682	HIS	-	EXPRESSION TAG	UNP P09546
A	683	HIS	-	EXPRESSION TAG	UNP P09546
A	684	HIS	-	EXPRESSION TAG	UNP P09546
A	685	HIS	-	EXPRESSION TAG	UNP P09546
A	686	HIS	-	EXPRESSION TAG	UNP P09546
A	687	HIS	-	EXPRESSION TAG	UNP P09546

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is TETRAHYDROFURAN-2-CARBOXYLIC ACID (three-letter code: TFB) (formula: $C_5H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	5	3		

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $\text{C}_{10}\text{H}_{22}\text{O}_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	10	6		
4	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	258	Total	O	0	0
			258	258		

i

- Molecule 1: Bifunctional protein PutA



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	72.80 Å 140.55 Å 146.95 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.64 – 1.85 40.64 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.64-1.85) 99.5 (40.64-1.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.87 (at 1.86 Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.192 , 0.225 0.192 , 0.226	Depositor DCC
R_{free} test set	3232 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	25.5	Xtriage
Anisotropy	0.695	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 56.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.003 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4073	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.46% 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, 0.333 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: 1PE, FAD, TFB

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	0.34	0/3806	0.51	0/5172

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

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	3728	0	3633	14	0
2	A	53	0	31	4	0
3	A	8	0	7	0	0
4	A	26	0	34	0	0
5	A	258	0	0	0	0
All	All	4073	0	3705	16	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 2.

All (16) 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:267:GLU:O	1:A:271:ASN:ND2	2.20	0.74
1:A:540:TYR:HB2	2:A:2001:FAD:HM72	1.86	0.58
1:A:159:SER:O	1:A:162:GLY:N	2.43	0.51
1:A:546:HIS:CE1	1:A:547:GLU:HG3	2.45	0.51
1:A:570:ILE:HG13	1:A:580:LEU:HD21	1.94	0.48
1:A:154:GLU:C	1:A:156:SER:H	2.17	0.47
1:A:111:LEU:HA	1:A:111:LEU:HD12	1.89	0.43
1:A:214:LEU:HD23	1:A:214:LEU:HA	1.83	0.43
1:A:485:ALA:HB1	2:A:2001:FAD:C8	2.49	0.43
1:A:287:LEU:HD21	1:A:542:PRO:HB3	2.01	0.42
1:A:101:ARG:NH2	1:A:109:MET:HE1	2.34	0.42
2:A:2001:FAD:H9	2:A:2001:FAD:H1'1	1.83	0.42
1:A:106:ALA:O	1:A:110:LEU:HG	2.20	0.41
1:A:522:GLU:O	1:A:532:LYS:HE2	2.20	0.41
1:A:373:GLU:H	1:A:373:GLU:CD	2.24	0.41
2:A:2001:FAD:H4'	2:A:2001:FAD:H1'1	1.81	0.41

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	488/602 (81%)	474 (97%)	11 (2%)	3 (1%)	28 13

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	GLU
1	A	155	PHE
1	A	159	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	360/490 (74%)	354 (98%)	6 (2%)	66	52

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	99	TYR
1	A	243	LEU
1	A	254	LEU
1	A	361	ARG
1	A	406	TYR
1	A	574	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	FAD	A	2001	-	51,58,58	1.98	14 (27%)	54,89,89	2.21	10 (18%)
3	TFB	A	2002	-	5,8,8	2.44	1 (20%)	5,10,10	0.50	0
4	1PE	A	2003	-	15,15,15	0.54	0	14,14,14	1.45	0
4	1PE	A	2004	-	9,9,15	0.51	0	8,8,14	1.40	0

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	FAD	A	2001	-	-	0/28/50/50	0/6/6/6
3	TFB	A	2002	-	-	0/0/11/11	0/1/1/1
4	1PE	A	2003	-	-	0/13/13/13	0/0/0/0
4	1PE	A	2004	-	-	0/7/7/13	0/0/0/0

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2002	TFB	CB-CA	-4.94	1.45	1.53
2	A	2001	FAD	C2B-C1B	-3.49	1.48	1.53
2	A	2001	FAD	O2'-C2'	-2.68	1.37	1.43
2	A	2001	FAD	O4B-C4B	-2.32	1.39	1.45
2	A	2001	FAD	PA-O5B	-2.24	1.49	1.59
2	A	2001	FAD	C2A-N3A	2.55	1.36	1.32
2	A	2001	FAD	C2-N1	2.66	1.43	1.38
2	A	2001	FAD	C4-N3	2.68	1.37	1.33
2	A	2001	FAD	C6A-N6A	2.77	1.45	1.34
2	A	2001	FAD	C5X-N5	2.95	1.39	1.35
2	A	2001	FAD	C10-N1	3.81	1.38	1.33
2	A	2001	FAD	C9A-N10	3.94	1.43	1.38
2	A	2001	FAD	C4-C4X	4.19	1.49	1.41
2	A	2001	FAD	C4X-N5	4.61	1.40	1.33
2	A	2001	FAD	O4-C4	4.87	1.36	1.24

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	FAD	N3A-C2A-N1A	-11.03	119.25	128.86
2	A	2001	FAD	C4-C4X-C10	-4.01	116.72	119.96
2	A	2001	FAD	C4X-C4-N3	-2.04	120.57	123.48
2	A	2001	FAD	O5'-P-O1P	2.08	117.64	109.25
2	A	2001	FAD	C2A-N1A-C6A	2.16	122.56	118.77
2	A	2001	FAD	C1'-N10-C9A	2.21	120.38	118.35
2	A	2001	FAD	C4X-N5-C5X	2.51	119.42	116.76
2	A	2001	FAD	C5X-C9A-N10	3.31	120.12	117.66
2	A	2001	FAD	C4-C4X-N5	3.43	122.44	118.68
2	A	2001	FAD	C4-N3-C2	5.90	120.31	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	FAD	4	0

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 ⓘ

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	493/602 (81%)	0.32	41 (8%) 12 12	18, 33, 81, 93	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	155	PHE	6.9
1	A	228	LEU	5.6
1	A	151	LEU	5.4
1	A	244	ILE	5.2
1	A	226	ALA	5.2
1	A	209	ALA	5.1
1	A	215	PHE	5.1
1	A	159	SER	4.8
1	A	243	LEU	4.8
1	A	205	PHE	4.8
1	A	148	VAL	4.3
1	A	232	LEU	4.0
1	A	152	LEU	4.0
1	A	206	VAL	3.9
1	A	233	ASN	3.9
1	A	236	ILE	3.8
1	A	156	SER	3.6
1	A	87	PRO	3.6
1	A	204	LEU	3.6
1	A	230	ARG	3.5
1	A	208	ALA	3.4
1	A	235	ILE	3.3
1	A	153	GLN	3.1
1	A	575	LEU	3.1
1	A	229	SER	3.1
1	A	185	ILE	3.1
1	A	227	SER	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	179	ALA	2.9
1	A	231	SER	2.7
1	A	403	ILE	2.6
1	A	184	LEU	2.6
1	A	158	SER	2.5
1	A	342	TYR	2.5
1	A	402	VAL	2.4
1	A	164	ALA	2.4
1	A	609	ARG	2.4
1	A	187	ASP	2.4
1	A	168	LEU	2.1
1	A	247	GLY	2.0
1	A	157	LEU	2.0
1	A	186	ARG	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
4	1PE	A	2004	10/16	0.84	0.21	3.81	45,50,57,60	0
4	1PE	A	2003	16/16	0.76	0.16	1.36	61,64,68,70	0
3	TFB	A	2002	8/8	0.98	0.18	0.92	23,24,27,28	0
2	FAD	A	2001	53/53	0.97	0.13	-0.05	18,21,23,26	0

6.5 Other polymers [i](#)

There are no such residues in this entry.