



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 12, 2017 – 10:03 pm GMT

PDB ID : 2FHN  
Title : Avidin related protein AVR4 (C122S, K109I mutant) in complex with BNA  
Authors : Livnah, O.; Prizant, M.  
Deposited on : 2005-12-26  
Resolution : 1.30 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](mailto: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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

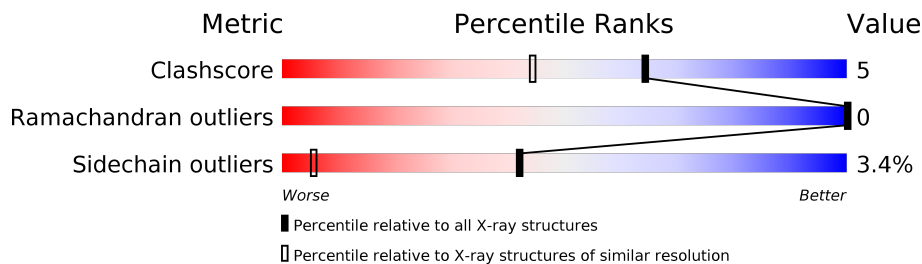
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.30 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	1185 (1.32-1.28)
Ramachandran outliers	110173	1138 (1.32-1.28)
Sidechain outliers	110143	1138 (1.32-1.28)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	X	126	 82% 11% • 6%
1	Y	126	 79% 13% • 6%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2141 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.

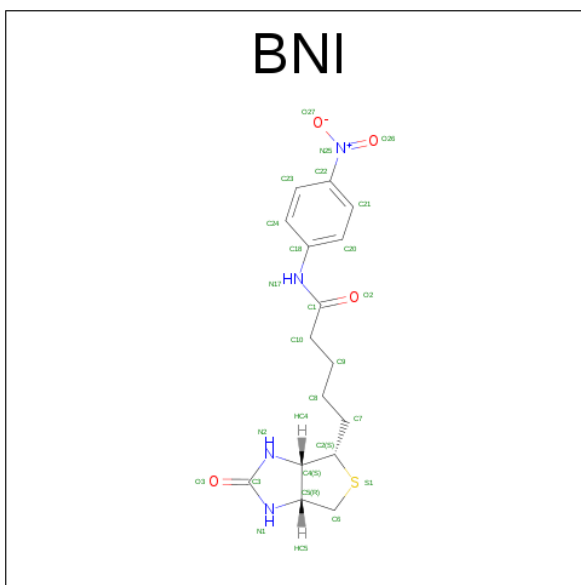
- Molecule 1 is a protein called Avidin-related protein 4/5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	118	Total	C	N	O	S	0	0	0
			926	588	161	173	4			
1	Y	118	Total	C	N	O	S	0	0	0
			920	585	158	173	4			

There are 12 discrepancies between the modelled and reference sequences:

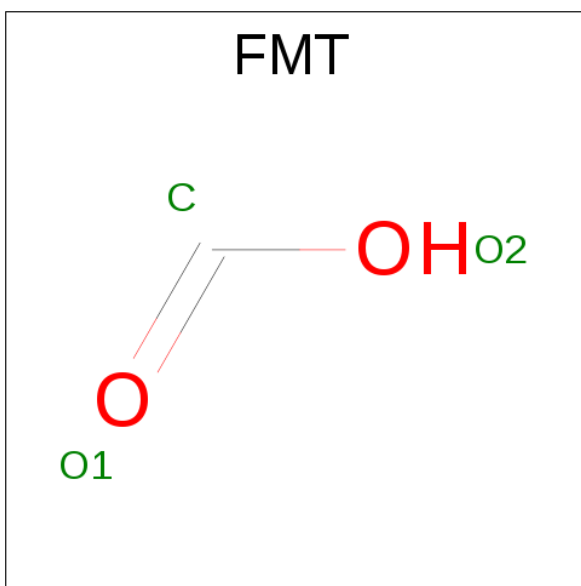
Chain	Residue	Modelled	Actual	Comment	Reference
X	109	ILE	LYS	ENGINEERED	UNP P56734
X	122	SER	-	CLONING ARTIFACT	UNP P56734
X	123	THR	-	CLONING ARTIFACT	UNP P56734
X	124	VAL	-	CLONING ARTIFACT	UNP P56734
X	125	GLU	-	CLONING ARTIFACT	UNP P56734
X	126	GLU	-	CLONING ARTIFACT	UNP P56734
Y	309	ILE	LYS	ENGINEERED	UNP P56734
Y	322	SER	-	CLONING ARTIFACT	UNP P56734
Y	323	THR	-	CLONING ARTIFACT	UNP P56734
Y	324	VAL	-	CLONING ARTIFACT	UNP P56734
Y	325	GLU	-	CLONING ARTIFACT	UNP P56734
Y	326	GLU	-	CLONING ARTIFACT	UNP P56734

- Molecule 2 is 5-(2-OXO-HEXAHYDRO-THIENO[3,4-D]IMIDAZOL-6-YL)-PENTANOIC ACID (4-NITRO-PHENYL)-AMIDE (three-letter code: BNI) (formula: C<sub>16</sub>H<sub>20</sub>N<sub>4</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	X	1	Total 25	C 16	N 4	O 4	S 1	0	0
2	Y	1	Total 25	C 16	N 4	O 4	S 1	0	0

- Molecule 3 is FORMIC ACID (three-letter code: FMT) (formula:  $\text{CH}_2\text{O}_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	Y	1	Total C O 3 1 2	0	0
3	Y	1	Total C O 3 1 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	X	1	Total	C	O	0	0
			3	1	2		
3	Y	1	Total	C	O	0	0
			3	1	2		
3	Y	1	Total	C	O	0	0
			3	1	2		
3	Y	1	Total	C	O	0	0
			3	1	2		

- Molecule 4 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	X	121	Total	O	0	0
			121	121		
4	Y	106	Total	O	0	0
			106	106		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.


Note EDS was not executed.

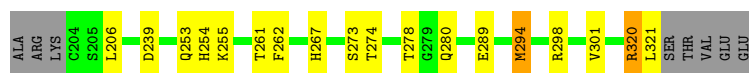
- Molecule 1: Avidin-related protein 4/5

Chain X:  82% 11% • 6%



- Molecule 1: Avidin-related protein 4/5

Chain Y:  79% 13% • 6%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.59Å 77.59Å 109.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.30	Depositor
% Data completeness (in resolution range)	99.8 (40.00-1.30)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.151 , 0.166	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2141	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FMT, BNI

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	X	0.46	0/948	0.81	2/1290 (0.2%)
1	Y	0.46	0/942	0.80	3/1283 (0.2%)
All	All	0.46	0/1890	0.80	5/2573 (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	X	1	0
1	Y	2	0
All	All	3	0

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	X	94	MET	CG-SD-CE	-10.46	83.46	100.20
1	Y	294	MET	CG-SD-CE	-7.17	88.73	100.20
1	Y	320	ARG	NE-CZ-NH1	6.45	123.53	120.30
1	Y	239	ASP	CB-CG-OD2	5.65	123.39	118.30
1	X	103	ASP	CB-CG-OD2	5.22	123.00	118.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	X	61	THR	CB
1	Y	261	THR	CB
1	Y	278	THR	CB



There are no planarity outliers.

## 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	X	926	0	900	13	0
1	Y	920	0	889	16	0
2	X	25	0	20	0	0
2	Y	25	0	20	0	0
3	X	3	0	1	1	0
3	Y	15	0	5	1	0
4	X	121	0	0	1	0
4	Y	106	0	0	0	0
All	All	2141	0	1835	18	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 5.

The worst 5 of 18 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:X:80:GLN:HE22	1:Y:301:VAL:H	1.37	0.72
1:X:101:VAL:H	1:Y:280:GLN:HE22	1.37	0.71
1:X:61:THR:HG23	1:Y:274:THR:OG1	1.93	0.69
1:X:74:THR:OG1	1:Y:261:THR:HG23	1.92	0.69
1:X:80:GLN:HE21	1:Y:298:ARG:HH21	1.42	0.65

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	X	116/126 (92%)	115 (99%)	1 (1%)	0	100	100
1	Y	116/126 (92%)	116 (100%)	0	0	100	100
All	All	232/252 (92%)	231 (100%)	1 (0%)	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	X	102/110 (93%)	98 (96%)	4 (4%)	37	3
1	Y	101/110 (92%)	98 (97%)	3 (3%)	46	7
All	All	203/220 (92%)	196 (97%)	7 (3%)	42	5

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	121	LEU
1	Y	321	LEU
1	Y	255	LYS
1	X	55	LYS
1	Y	294	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	X	80	GLN
1	Y	280	GLN
1	Y	254	HIS
1	X	67	HIS
1	Y	253	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

8 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	BNI	X	501	-	26,27,27	1.82	8 (30%)	35,37,37	1.18	3 (8%)
3	FMT	X	903	-	0,2,2	0.00	-	0,1,1	0.00	-
2	BNI	Y	502	-	26,27,27	1.89	9 (34%)	35,37,37	1.27	5 (14%)
3	FMT	Y	901	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	Y	902	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	Y	904	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	Y	905	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	Y	906	-	0,2,2	0.00	-	0,1,1	0.00	-

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	BNI	X	501	-	-	0/13/36/36	0/3/3/3
3	FMT	X	903	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BNI	Y	502	-	-	0/13/36/36	0/3/3/3
3	FMT	Y	901	-	-	0/0/0/0	0/0/0/0
3	FMT	Y	902	-	-	0/0/0/0	0/0/0/0
3	FMT	Y	904	-	-	0/0/0/0	0/0/0/0
3	FMT	Y	905	-	-	0/0/0/0	0/0/0/0
3	FMT	Y	906	-	-	0/0/0/0	0/0/0/0

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	501	BNI	C18-N17	-2.31	1.37	1.41
2	Y	502	BNI	C18-N17	-2.23	1.37	1.41
2	X	501	BNI	C24-C23	2.06	1.42	1.38
2	X	501	BNI	C23-C22	2.11	1.43	1.38
2	Y	502	BNI	C24-C23	2.16	1.42	1.38

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Y	502	BNI	C5-C6-S1	-2.26	104.62	106.24
2	X	501	BNI	C4-C5-N1	2.06	104.39	102.37
2	Y	502	BNI	C6-C5-C4	2.16	110.50	108.70
2	Y	502	BNI	C10-C1-N17	2.26	118.58	114.53
2	Y	502	BNI	N2-C3-N1	2.29	110.62	108.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	X	903	FMT	1	0
3	Y	902	FMT	1	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](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.