



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 29, 2024 – 06:30 AM EDT

PDB ID : 3SZH  
Title : Crystal structure of apo shwanavidin (P1 form)  
Authors : Livnah, O.; Meir, A.  
Deposited on : 2011-07-19  
Resolution : 1.07 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

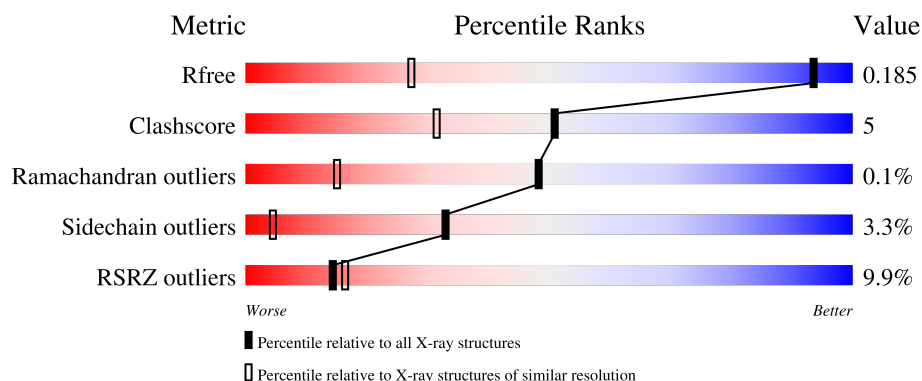
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.07 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))
$R_{free}$	164625	1502 (1.10-1.06)
Clashscore	180529	1658 (1.10-1.06)
Ramachandran outliers	177936	1614 (1.10-1.06)
Sidechain outliers	177891	1611 (1.10-1.06)
RSRZ outliers	164620	1502 (1.10-1.06)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	122	<div> <div>7%</div> <div>84%</div> <div>8%</div> <div>6%</div> </div>
1	B	122	<div> <div>11%</div> <div>82%</div> <div>10%</div> <div>5%</div> </div>
1	C	122	<div> <div>11%</div> <div>81%</div> <div>13%</div> <div>5%</div> </div>
1	D	122	<div> <div>10%</div> <div>82%</div> <div>11%</div> <div>5%</div> </div>
1	E	122	<div> <div>7%</div> <div>88%</div> <div>7%</div> <div>5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	122	<div><div></div><div>11%</div><div>84%</div><div>11%</div><div></div><div></div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6361 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/streptavidin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	115	Total	C	N	O	S	3	7	0
			902	574	146	179	3			
1	B	116	Total	C	N	O	S	4	7	0
			906	577	146	179	4			
1	C	117	Total	C	N	O	S	6	10	0
			931	594	148	184	5			
1	D	116	Total	C	N	O	S	5	11	0
			927	593	147	183	4			
1	E	116	Total	C	N	O	S	7	8	0
			910	579	146	181	4			
1	F	117	Total	C	N	O	S	9	13	0
			934	597	148	185	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q12QS6
A	2	ALA	-	expression tag	UNP Q12QS6
B	1	MET	-	expression tag	UNP Q12QS6
B	2	ALA	-	expression tag	UNP Q12QS6
C	1	MET	-	expression tag	UNP Q12QS6
C	2	ALA	-	expression tag	UNP Q12QS6
D	1	MET	-	expression tag	UNP Q12QS6
D	2	ALA	-	expression tag	UNP Q12QS6
E	1	MET	-	expression tag	UNP Q12QS6
E	2	ALA	-	expression tag	UNP Q12QS6
F	1	MET	-	expression tag	UNP Q12QS6
F	2	ALA	-	expression tag	UNP Q12QS6

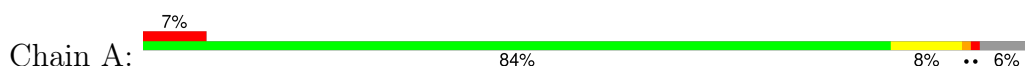
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	155	Total 155	O 155	0	0
2	B	140	Total 140	O 140	0	0
2	C	136	Total 136	O 136	0	0
2	D	137	Total 137	O 137	0	0
2	E	136	Total 136	O 136	0	0
2	F	147	Total 147	O 147	0	0

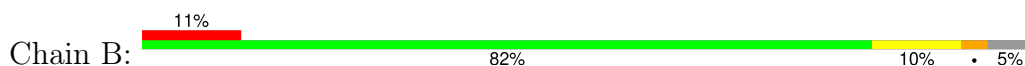
### 3 Residue-property plots [i](#)

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 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.

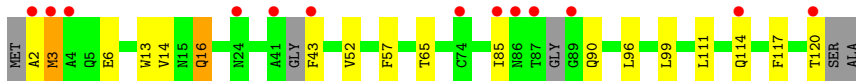
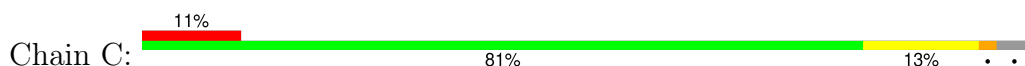
- Molecule 1: Avidin/streptavidin



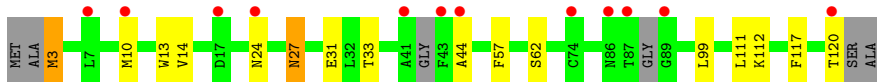
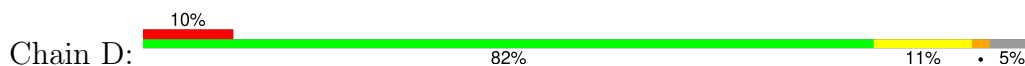
- Molecule 1: Avidin/streptavidin



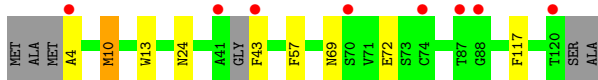
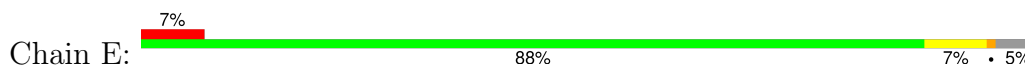
- Molecule 1: Avidin/streptavidin



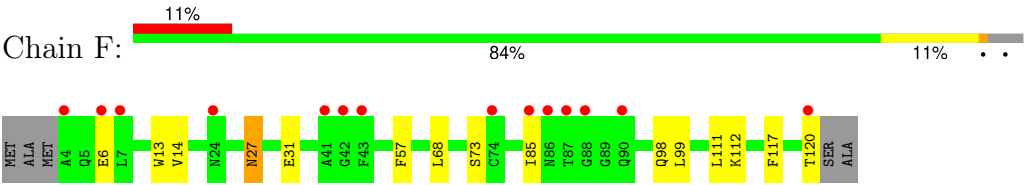
- Molecule 1: Avidin/streptavidin



- Molecule 1: Avidin/streptavidin



- Molecule 1: Avidin/streptavidin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.33Å 63.50Å 67.75Å 103.83° 107.28° 103.84°	Depositor
Resolution (Å)	60.97 – 1.07 60.96 – 1.07	Depositor EDS
% Data completeness (in resolution range)	92.5 (60.97-1.07) 92.5 (60.96-1.07)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 1.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.158 , 0.185 0.158 , 0.185	Depositor DCC
$R_{free}$ test set	12985 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	10.7	Xtriage
Anisotropy	0.477	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6361	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.70% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0.97	5/943 (0.5%)	0.79	1/1285 (0.1%)
1	B	1.34	4/947 (0.4%)	0.81	4/1289 (0.3%)
1	C	0.79	4/980 (0.4%)	0.80	2/1333 (0.2%)
1	D	0.52	0/979	0.73	0/1332
1	E	0.65	2/954 (0.2%)	0.73	0/1298
1	F	1.76	2/994 (0.2%)	0.72	2/1355 (0.1%)
All	All	1.09	17/5797 (0.3%)	0.76	9/7892 (0.1%)

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
1	B	0	1
1	C	0	2
All	All	0	4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	6	GLU	CD-OE1	46.54	1.76	1.25
1	B	6	GLU	CD-OE2	26.45	1.54	1.25
1	F	6	GLU	CD-OE2	25.38	1.53	1.25
1	B	6	GLU	CD-OE1	-22.09	1.01	1.25
1	A	86	ASN	CG-ND2	16.49	1.74	1.32

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	86	ASN	CA-CB-CG	-8.38	94.96	113.40
1	F	6	GLU	OE1-CD-OE2	-8.05	113.64	123.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	90	GLN	CG-CD-OE1	7.19	135.99	121.60
1	B	6	GLU	CG-CD-OE1	7.04	132.38	118.30
1	C	43	PHE	CZ-CE2-CD2	-7.03	111.67	120.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	86	ASN	Sidechain
1	B	6	GLU	Sidechain
1	C	16	GLN	Sidechain
1	C	6	GLU	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	902	0	876	9	0
1	B	906	0	882	9	0
1	C	931	0	913	11	0
1	D	927	0	916	12	0
1	E	910	0	887	4	0
1	F	934	0	927	11	0
2	A	155	0	0	1	0
2	B	140	0	0	1	0
2	C	136	0	0	0	0
2	D	137	0	0	2	0
2	E	136	0	0	3	0
2	F	147	0	0	0	0
All	All	6361	0	5401	55	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 55 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:52:VAL:HG22	1:A:65[B]:THR:HG22	1.46	0.97
1:C:52:VAL:HG22	1:C:65[B]:THR:HG22	1.47	0.96
1:D:33[B]:THR:HG23	2:D:235:HOH:O	1.70	0.90
1:D:14[B]:VAL:HG23	1:D:120:THR:HG22	1.54	0.89
1:F:68[A]:LEU:CD1	1:F:73[A]:SER:HB3	2.08	0.83

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	118/122 (97%)	116 (98%)	2 (2%)	0	100	100
1	B	119/122 (98%)	118 (99%)	1 (1%)	0	100	100
1	C	121/122 (99%)	119 (98%)	2 (2%)	0	100	100
1	D	121/122 (99%)	119 (98%)	1 (1%)	1 (1%)	16	3
1	E	120/122 (98%)	120 (100%)	0	0	100	100
1	F	128/122 (105%)	126 (98%)	2 (2%)	0	100	100
All	All	727/732 (99%)	718 (99%)	8 (1%)	1 (0%)	48	15

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	44	ALA

### 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	104/100 (104%)	101 (97%)	3 (3%)	37 5
1	B	104/100 (104%)	100 (96%)	4 (4%)	28 2
1	C	108/100 (108%)	104 (96%)	4 (4%)	29 3
1	D	109/100 (109%)	105 (96%)	4 (4%)	29 3
1	E	105/100 (105%)	102 (97%)	3 (3%)	37 5
1	F	110/100 (110%)	107 (97%)	3 (3%)	40 6
All	All	640/600 (107%)	619 (97%)	21 (3%)	33 4

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	112	LYS
1	E	57	PHE
1	F	85	ILE
1	F	27	ASN
1	E	10[B]	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	5	GLN
1	F	27	ASN
1	D	24	ASN
1	F	98	GLN
1	E	5	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 [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	115/122 (94%)	0.59	9 (7%) 20 23	6, 12, 18, 28	9 (7%)
1	B	116/122 (95%)	0.56	13 (11%) 11 13	5, 12, 19, 29	10 (8%)
1	C	117/122 (95%)	0.67	13 (11%) 12 13	6, 13, 21, 27	13 (11%)
1	D	116/122 (95%)	0.50	12 (10%) 13 15	6, 12, 21, 29	13 (11%)
1	E	116/122 (95%)	0.66	8 (6%) 24 27	7, 13, 19, 25	10 (8%)
1	F	117/122 (95%)	0.90	14 (11%) 10 11	7, 14, 24, 29	16 (13%)
All	All	697/732 (95%)	0.65	69 (9%) 14 16	5, 13, 21, 29	71 (10%)

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	88	GLY	7.3
1	E	41	ALA	7.1
1	E	43	PHE	6.6
1	D	41	ALA	6.3
1	F	42	GLY	5.9

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.