



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 22, 2021 – 12:20 PM JST

PDB ID : 7EK8  
Title : Crystal structure of apo streptavidin at ambient temperature  
Authors : DeMirici, H.; Ertem, F.B.; Destan, E.; Ayan, E.  
Deposited on : 2021-04-04  
Resolution : 1.70 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.23.1  
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) : 2.23.1

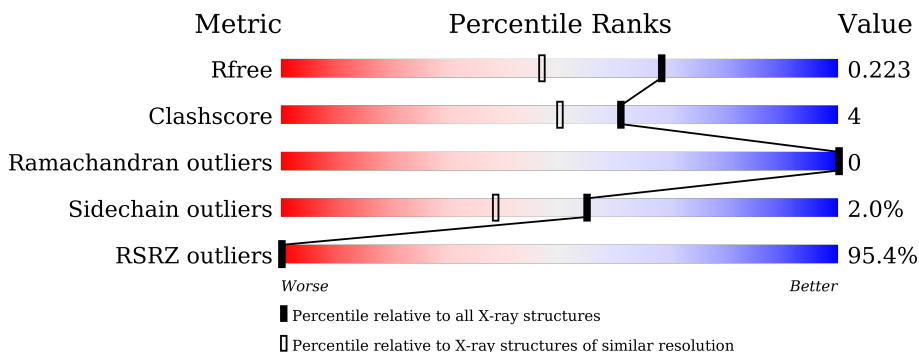
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	123	<div style="display: flex; justify-content: space-between;"> <div style="width: 93%; height: 10px; background-color: red;"></div> <div style="width: 89%; height: 10px; background-color: green;"></div> <div style="width: 10%; height: 10px; background-color: yellow;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">93% 89% 10% .</p>
1	B	123	<div style="display: flex; justify-content: space-between;"> <div style="width: 93%; height: 10px; background-color: red;"></div> <div style="width: 85%; height: 10px; background-color: green;"></div> <div style="width: 13%; height: 10px; background-color: yellow;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">93% 85% 13% ..</p>
1	C	123	<div style="display: flex; justify-content: space-between;"> <div style="width: 94%; height: 10px; background-color: red;"></div> <div style="width: 92%; height: 10px; background-color: green;"></div> <div style="width: 5%; height: 10px; background-color: yellow;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">94% 92% 5% ..</p>
1	D	123	<div style="display: flex; justify-content: space-between;"> <div style="width: 94%; height: 10px; background-color: red;"></div> <div style="width: 91%; height: 10px; background-color: green;"></div> <div style="width: 7%; height: 10px; background-color: yellow;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">94% 91% 7% .</p>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4166 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 Streptavidin.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace	
			Total	C	N				O
1	A	121	942	594	159	189	0	8	0
1	B	121	937	588	163	186	0	8	0
1	C	120	950	597	161	192	0	10	0
1	D	120	940	591	160	189	0	10	0

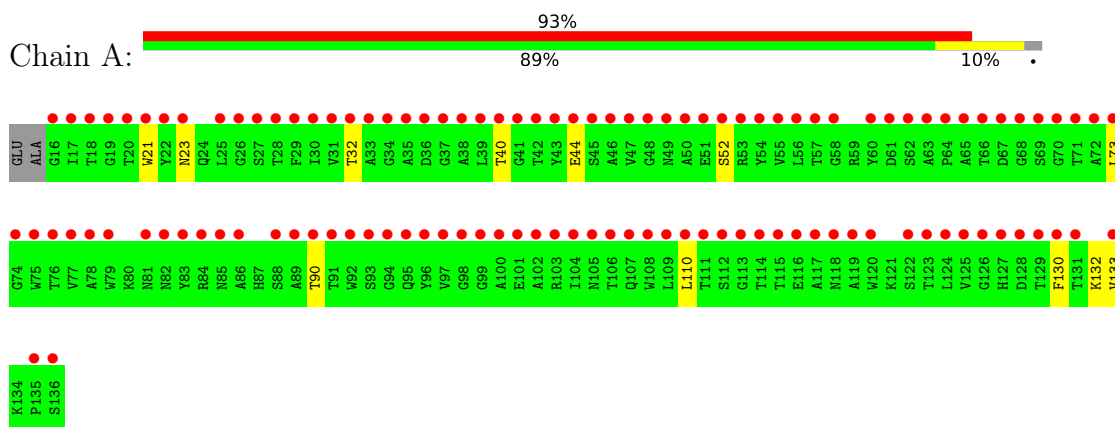
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	100	101	101	0	1
2	B	94	95	95	0	1
2	C	111	115	115	0	4
2	D	85	86	86	0	1

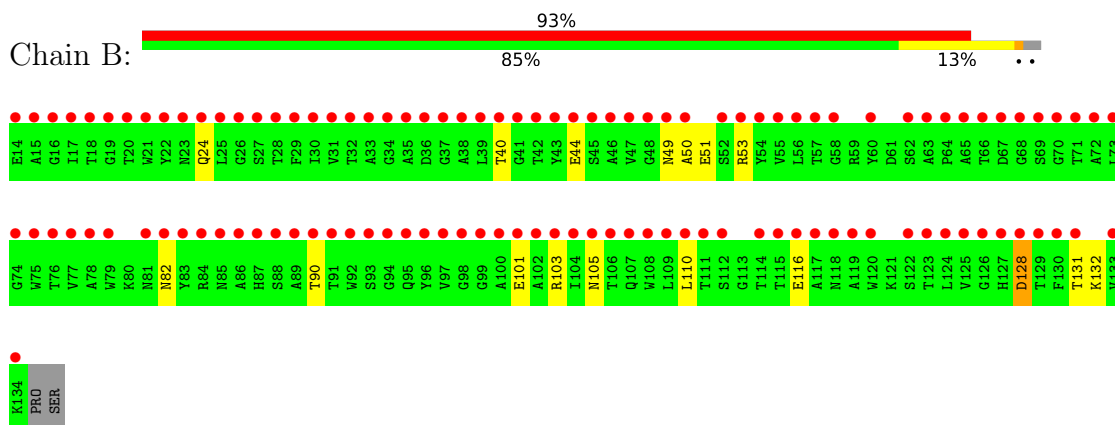
### 3 Residue-property plots [i](#)

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.

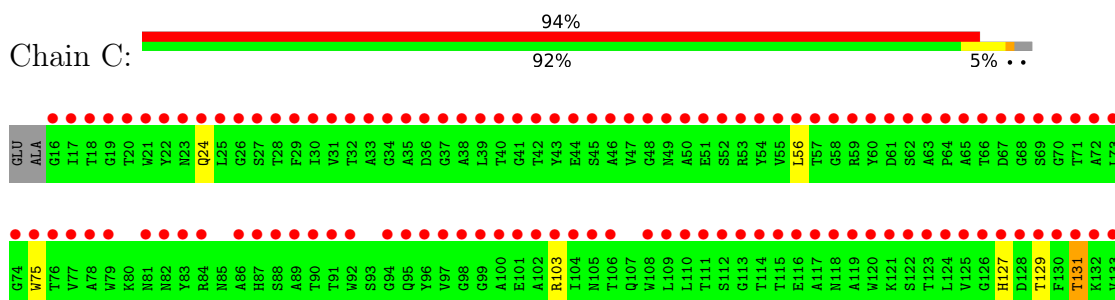
- Molecule 1: Streptavidin



- Molecule 1: Streptavidin

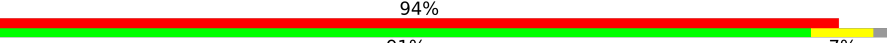


- Molecule 1: Streptavidin



K134  
P135  
SER

● Molecule 1: Streptavidin

Chain D:  94%  
91% 7%

GLU  
ALA  
G16  
I17  
T18  
G19  
T20  
W21  
Y22  
N23  
Q24  
L25  
G26  
S27  
T28  
F29  
I30  
V31  
T32  
A33  
G34  
A35  
D36  
G37  
A38  
L39  
T40  
E41  
T42  
Y43  
E44  
S45  
A46  
V47  
G48  
N49  
A50  
E51  
S52  
R53  
Y54  
V55  
L56  
T57  
G58  
R59  
Y60  
D61  
S62  
A63  
P64  
A65  
T66  
D67  
G68  
S69  
G70  
T71  
A72  
L73

G74  
W75  
T76  
V77  
A78  
W79  
K80  
N81  
N82  
Y83  
R84  
N85  
A86  
H87  
S88  
A89  
T90  
T91  
R92  
S93  
G94  
D95  
Y96  
Y97  
G98  
G99  
A100  
E101  
A102  
R103  
I104  
H105  
T106  
Q107  
W108  
L109  
L110  
T111  
S112  
G113  
T114  
T115  
E116  
A117  
M118  
A119  
W120  
K121  
S122  
T123  
L124  
V125  
G126  
H127  
D128  
T129  
F130  
T131  
K132  
V133

K134  
P135  
SER

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.40Å 87.70Å 58.90Å 90.00° 98.90° 90.00°	Depositor
Resolution (Å)	48.49 – 1.70 48.49 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.49-1.70) 100.0 (48.49-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 1.70Å)	Xtrriage
Refinement program	PHENIX dev_3318	Depositor
R, $R_{free}$	0.190 , 0.224 0.190 , 0.223	Depositor DCC
$R_{free}$ test set	1932 reflections (3.70%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.9	Xtrriage
Anisotropy	0.560	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 87.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	4166	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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	0.71	2/982 (0.2%)	0.75	0/1346
1	B	0.53	0/975	0.69	0/1334
1	C	0.56	0/994	0.68	0/1362
1	D	0.57	0/984	0.64	0/1349
All	All	0.60	2/3935 (0.1%)	0.69	0/5391

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	52[A]	SER	C-O	8.61	1.39	1.23
1	A	52[B]	SER	C-O	8.61	1.39	1.23

There are no bond angle outliers.

There are no chirality outliers.

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	A	942	0	890	6	0
1	B	937	0	891	17	0
1	C	950	0	903	5	0
1	D	940	0	896	5	0
2	A	101	0	0	3	1
2	B	95	0	0	13	0
2	C	115	0	0	3	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	86	0	0	2	1
All	All	4166	0	3580	33	3

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 4.

The worst 5 of 33 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:B:128:ASP:OD1	2:B:201:HOH:O	1.94	0.85
1:B:116:GLU:OE1	2:B:202:HOH:O	2.01	0.77
1:B:103[B]:ARG:NH2	2:B:204:HOH:O	2.13	0.76
1:D:36[A]:ASP:OD2	2:D:201:HOH:O	2.08	0.71
1:C:129[A]:THR:OG1	2:C:201[A]:HOH:O	2.11	0.68

All (3) 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)
2:C:244:HOH:O	2:C:282:HOH:O[1_655]	2.09	0.11
2:C:293:HOH:O	2:C:299:HOH:O[1_655]	2.13	0.07
2:A:292:HOH:O	2:D:209:HOH:O[2_846]	2.17	0.03

## 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	126/123 (102%)	124 (98%)	2 (2%)	0	100	100
1	B	126/123 (102%)	125 (99%)	1 (1%)	0	100	100
1	C	128/123 (104%)	125 (98%)	3 (2%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	128/123 (104%)	127 (99%)	1 (1%)	0	100	100
All	All	508/492 (103%)	501 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	96/91 (106%)	94 (98%)	2 (2%)	53	36
1	B	94/91 (103%)	91 (97%)	3 (3%)	39	20
1	C	98/91 (108%)	96 (98%)	2 (2%)	55	38
1	D	96/91 (106%)	94 (98%)	2 (2%)	53	36
All	All	384/364 (106%)	375 (98%)	9 (2%)	55	33

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	107[A]	GLN
1	D	107[B]	GLN
1	B	101	GLU
1	B	128	ASP
1	C	131[A]	THR

Sometimes 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 monosaccharides 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	121/123 (98%)	3.64	114 (94%)	0   0	6, 14, 31, 41	2 (1%)
1	B	121/123 (98%)	3.85	114 (94%)	0   0	5, 14, 38, 65	1 (0%)
1	C	120/123 (97%)	4.17	116 (96%)	0   0	5, 12, 60, 90	1 (0%)
1	D	120/123 (97%)	4.39	116 (96%)	0   0	6, 16, 56, 101	1 (0%)
All	All	482/492 (97%)	4.01	460 (95%)	0   0	5, 14, 42, 101	5 (1%)

The worst 5 of 460 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	46	ALA	19.7
1	D	47	VAL	18.9
1	C	47	VAL	17.5
1	D	45	SER	16.4
1	C	48	GLY	15.3

### 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 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.