WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 17, 2021 – 03:21 AM EDT

PDB ID	:	1LUQ
Title	:	Full Matrix Error Analysis of Streptavidin
Authors	:	Merritt, E.A.; Le Trong, I.
Deposited on	:	2002-05-23
Resolution	:	0.96 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
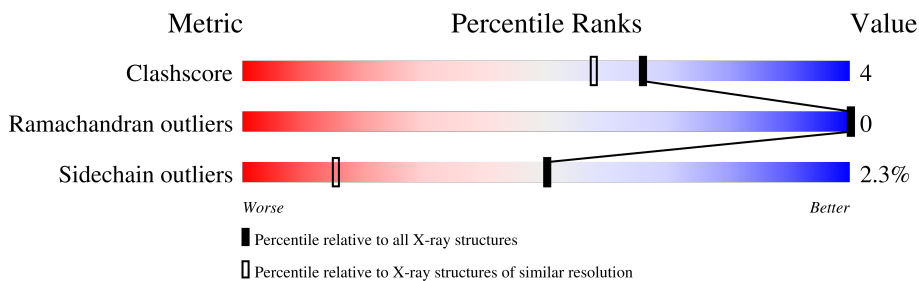with specific help available everywhere you see the (i) symbol.

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 0.96 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	Percentile Ranks	Value
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		2.3%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1321 (1.06-0.86)
Ramachandran outliers	138981	1233 (1.06-0.86)
Sidechain outliers	138945	1235 (1.06-0.86)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	127	82% 11% · 6%
1	B	127	87% 6% · 6%

2   Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2236 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.
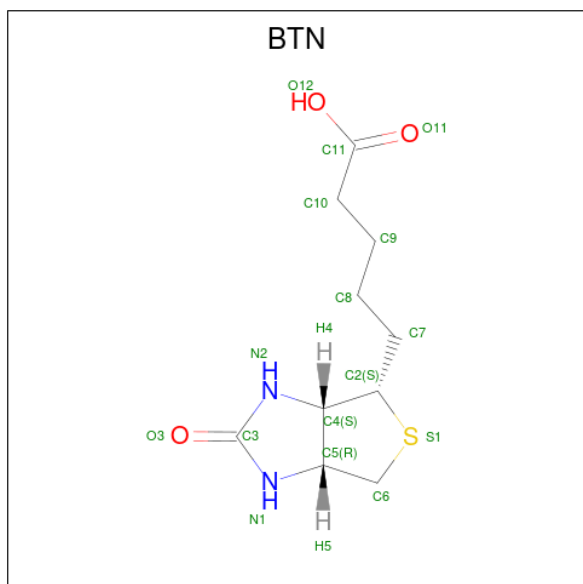
- Molecule 1 is a protein called Streptavidin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	119	Total	C	N	O	0	8	0
			954	605	160	189			
1	B	120	Total	C	N	O	0	6	0
			934	583	163	188			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1045	ALA	SER	engineered mutation	GB 4376079
B	2045	ALA	SER	engineered mutation	GB 4376079

- Molecule 2 is BIOTIN (three-letter code: BTN) (formula: $C_{10}H_{16}N_2O_3S$).

BTN

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			16	10	2	3	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	S	0	0
			16	10	2	3	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).
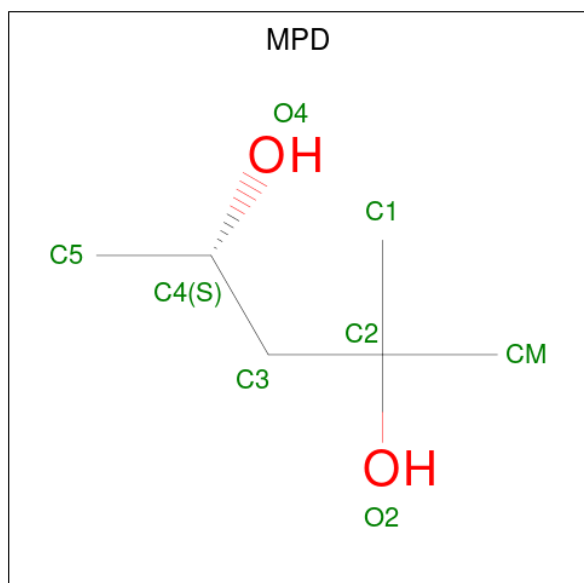
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total C O 8 6 2			0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	145	Total O 145 145		0	0
5	B	157	Total O 157 157		0	0

3 Residue-property plots ⓘ

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. 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. 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: Streptavidin

Chain A:
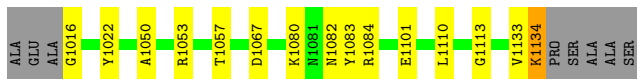
• Molecule 1: Streptavidin

Chain B:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	46.03Å 93.54Å 104.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.00 – 0.96	Depositor
% Data completeness (in resolution range)	96.4 (14.00-0.96)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.139 , 0.165	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2236	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MPD, BTN

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	0.92	0/987	1.56	24/1352 (1.8%)
1	B	0.91	0/968	1.68	15/1325 (1.1%)
All	All	0.91	0/1955	1.62	39/2677 (1.5%)

There are no bond length outliers.

All (39) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2103	ARG	NE-CZ-NH2	32.08	136.34	120.30
1	A	1053	ARG	CD-NE-CZ	19.40	150.76	123.60
1	A	1084	ARG	NE-CZ-NH1	-14.36	113.12	120.30
1	A	1084	ARG	NE-CZ-NH2	10.88	125.74	120.30
1	B	2103	ARG	NH1-CZ-NH2	-10.74	107.59	119.40
1	B	2084	ARG	NE-CZ-NH1	-9.66	115.47	120.30
1	A	1083[A]	TYR	CB-CG-CD2	8.93	126.36	121.00
1	A	1083[B]	TYR	CB-CG-CD2	8.93	126.36	121.00
1	B	2073[A]	LEU	CB-CG-CD1	8.84	126.03	111.00
1	B	2073[B]	LEU	CB-CG-CD1	8.84	126.03	111.00
1	B	2103	ARG	NE-CZ-NH1	-8.50	116.05	120.30
1	B	2053[A]	ARG	NE-CZ-NH1	7.97	124.29	120.30
1	B	2053[B]	ARG	NE-CZ-NH1	7.97	124.29	120.30
1	B	2103	ARG	CD-NE-CZ	7.97	134.76	123.60
1	B	2060	TYR	CB-CG-CD1	7.92	125.75	121.00
1	A	1133	VAL	CA-CB-CG2	7.52	122.19	110.90
1	A	1067	ASP	O-C-N	6.86	134.86	123.20
1	A	1101	GLU	OE1-CD-OE2	-6.59	115.40	123.30
1	A	1022[A]	TYR	CA-CB-CG	6.24	125.25	113.40
1	A	1022[B]	TYR	CA-CB-CG	6.24	125.25	113.40
1	B	2110	LEU	CB-CG-CD1	6.19	121.52	111.00
1	A	1133	VAL	CA-CB-CG1	-6.13	101.71	110.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1134	LYS	CA-C-O	6.05	132.81	120.10
1	B	2051	GLU	OE1-CD-OE2	5.97	130.47	123.30
1	B	2053[A]	ARG	NH1-CZ-NH2	-5.89	112.92	119.40
1	B	2053[B]	ARG	NH1-CZ-NH2	-5.89	112.92	119.40
1	A	1113	GLY	C-N-CA	5.71	135.97	121.70
1	A	1110[A]	LEU	CB-CG-CD1	5.57	120.47	111.00
1	A	1110[B]	LEU	CB-CG-CD1	5.57	120.47	111.00
1	A	1067	ASP	CA-C-O	-5.46	108.62	120.10
1	A	1113	GLY	O-C-N	-5.45	113.97	122.70
1	A	1133	VAL	CB-CA-C	5.30	121.47	111.40
1	B	2084	ARG	NH1-CZ-NH2	5.24	125.17	119.40
1	A	1101	GLU	CA-CB-CG	5.24	124.93	113.40
1	A	1050	ALA	C-N-CA	5.12	134.49	121.70
1	A	1083[A]	TYR	CB-CG-CD1	-5.06	117.96	121.00
1	A	1083[B]	TYR	CB-CG-CD1	-5.06	117.96	121.00
1	A	1057[A]	THR	CA-CB-CG2	-5.02	105.37	112.40
1	A	1057[B]	THR	CA-CB-CG2	-5.02	105.37	112.40

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	954	0	12	3	0
1	B	934	0	0	1	0
2	A	16	0	15	0	0
2	B	16	0	15	0	0
3	A	6	0	7	2	0
4	A	8	0	13	4	0
5	A	145	0	0	4	0
5	B	157	0	0	1	0
All	All	2236	0	62	8	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 4.

All (8) 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:1080:LYS:NZ	3:A:3001:GOL:H12	1.98	0.78
4:A:4001:MPD:H13	5:A:6131:HOH:O	1.93	0.67
4:A:4001:MPD:H32	5:A:6076:HOH:O	1.94	0.67
4:A:4001:MPD:H11	5:A:6161:HOH:O	2.11	0.49
1:B:2103:ARG:NH2	5:B:6231:HOH:O	2.48	0.46
1:A:1082:ASN:ND2	3:A:3001:GOL:O1	2.49	0.46
4:A:4001:MPD:O2	4:A:4001:MPD:H52	2.16	0.45
1:A:1016:GLY:N	5:A:6220:HOH:O	2.53	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	125/127 (98%)	121 (97%)	4 (3%)	0	100	100
1	B	124/127 (98%)	122 (98%)	2 (2%)	0	100	100
All	All	249/254 (98%)	243 (98%)	6 (2%)	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	95/91 (104%)	94 (99%)	1 (1%)	73	41
1	B	94/91 (103%)	89 (95%)	5 (5%)	22	2
All	All	189/182 (104%)	183 (97%)	6 (3%)	50	8

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

Mol	Chain	Res	Type
1	A	1134	LYS
1	B	2023	ASN
1	B	2073[A]	LEU
1	B	2073[B]	LEU
1	B	2116[A]	GLU
1	B	2116[B]	GLU

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

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	BTN	B	5200	-	14,17,17	0.68	0	19,23,23	0.76	0
3	GOL	A	3001	-	5,5,5	0.90	0	5,5,5	2.88	3 (60%)
4	MPD	A	4001	-	7,7,7	0.79	0	9,10,10	1.68	3 (33%)
2	BTN	A	5100	-	14,17,17	0.93	1 (7%)	19,23,23	0.99	1 (5%)

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	BTN	B	5200	-	-	0/5/28/28	0/2/2/2
3	GOL	A	3001	-	-	2/4/4/4	-
4	MPD	A	4001	-	-	4/5/5/5	-
2	BTN	A	5100	-	-	0/5/28/28	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	5100	BTN	C2-S1	-2.36	1.78	1.82

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	3001	GOL	O1-C1-C2	5.46	136.37	110.20
4	A	4001	MPD	O4-C4-C3	3.04	123.65	111.36
3	A	3001	GOL	O2-C2-C1	2.57	120.43	109.12
4	A	4001	MPD	CM-C2-C1	2.29	115.34	110.57
3	A	3001	GOL	C3-C2-C1	2.23	120.39	111.70
4	A	4001	MPD	C1-C2-C3	-2.18	99.79	109.96
2	A	5100	BTN	N2-C3-N1	2.15	110.78	108.76

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	3001	GOL	C1-C2-C3-O3
4	A	4001	MPD	C2-C3-C4-O4
3	A	3001	GOL	O2-C2-C3-O3
4	A	4001	MPD	O2-C2-C3-C4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	A	4001	MPD	C2-C3-C4-C5
4	A	4001	MPD	C1-C2-C3-C4

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	3001	GOL	2	0
4	A	4001	MPD	4	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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.