



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 11:51 am GMT

PDB ID : 3MQB  
Title : Crystal Structure of Ectodomain of BST-2/Tetherin/CD317 (C2)  
Authors : Xiong, Y.; Yang, H.; Wang, J.; Meng, W.  
Deposited on : 2010-04-27  
Resolution : 3.20 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](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 : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

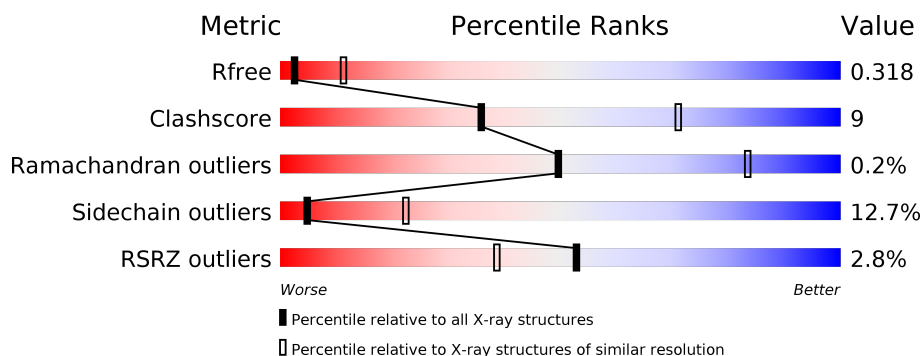
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
$R_{free}$	111664	1121 (3.22-3.18)
Clashscore	122126	1091 (3.20-3.20)
Ramachandran outliers	120053	1074 (3.20-3.20)
Sidechain outliers	120020	1073 (3.20-3.20)
RSRZ outliers	108989	1083 (3.22-3.18)

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.

Mol	Chain	Length	Quality of chain
1	A	121	<div> <div>3%</div> <div> <div></div> <div>67%</div> <div>17%</div> <div>5%</div> <div>10%</div> </div> </div>
1	B	121	<div> <div>2%</div> <div> <div></div> <div>70%</div> <div>17%</div> <div>12%</div> </div> </div>
1	E	121	<div> <div>2%</div> <div> <div></div> <div>67%</div> <div>16%</div> <div>5%</div> <div>12%</div> </div> </div>
1	F	121	<div> <div>2%</div> <div> <div></div> <div>66%</div> <div>15%</div> <div>16%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3305 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 Bone marrow stromal antigen 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	109	Total	C	N	O	S	0	0	0
			852	516	157	173	6			
1	B	106	Total	C	N	O	S	0	0	0
			831	505	153	167	6			
1	E	106	Total	C	N	O	S	0	0	0
			831	505	153	167	6			
1	F	102	Total	C	N	O	S	0	0	0
			791	476	148	161	6			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	41	ALA	-	EXPRESSION TAG	UNP Q10589
A	42	GLY	-	EXPRESSION TAG	UNP Q10589
A	43	PHE	-	EXPRESSION TAG	UNP Q10589
A	44	SER	-	EXPRESSION TAG	UNP Q10589
A	45	MET	-	EXPRESSION TAG	UNP Q10589
A	46	ASP	-	EXPRESSION TAG	UNP Q10589
B	41	ALA	-	EXPRESSION TAG	UNP Q10589
B	42	GLY	-	EXPRESSION TAG	UNP Q10589
B	43	PHE	-	EXPRESSION TAG	UNP Q10589
B	44	SER	-	EXPRESSION TAG	UNP Q10589
B	45	MET	-	EXPRESSION TAG	UNP Q10589
B	46	ASP	-	EXPRESSION TAG	UNP Q10589
E	41	ALA	-	EXPRESSION TAG	UNP Q10589
E	42	GLY	-	EXPRESSION TAG	UNP Q10589
E	43	PHE	-	EXPRESSION TAG	UNP Q10589
E	44	SER	-	EXPRESSION TAG	UNP Q10589
E	45	MET	-	EXPRESSION TAG	UNP Q10589
E	46	ASP	-	EXPRESSION TAG	UNP Q10589
F	41	ALA	-	EXPRESSION TAG	UNP Q10589
F	42	GLY	-	EXPRESSION TAG	UNP Q10589
F	43	PHE	-	EXPRESSION TAG	UNP Q10589

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	44	SER	-	EXPRESSION TAG	UNP Q10589
F	45	MET	-	EXPRESSION TAG	UNP Q10589
F	46	ASP	-	EXPRESSION TAG	UNP Q10589



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	181.65Å 28.59Å 94.51Å 90.00° 111.43° 90.00°	Depositor
Resolution (Å)	20.00 – 3.20 19.91 – 3.20	Depositor EDS
% Data completeness (in resolution range)	87.0 (20.00-3.20) 87.0 (19.91-3.20)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 3.22Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.255 , 0.315 0.257 , 0.318	Depositor DCC
$R_{free}$ test set	338 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.7	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 58.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3305	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.06% 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	0.63	0/858	0.65	0/1152
1	B	0.60	0/837	0.60	0/1124
1	E	0.63	0/837	0.66	0/1124
1	F	0.59	0/794	0.66	1/1065 (0.1%)
All	All	0.61	0/3326	0.64	1/4465 (0.0%)

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

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	F	58	ARG	NE-CZ-NH1	5.17	122.89	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	151	LYS	Peptide

### 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	852	0	843	21	1
1	B	831	0	825	24	0
1	E	831	0	825	21	0
1	F	791	0	787	27	0
All	All	3305	0	3280	57	1

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 9.

All (57) 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:156:SER:HB2	1:A:158:GLN:HE21	1.13	1.10
1:A:84:VAL:HG11	1:B:84:VAL:HG11	1.22	1.07
1:A:134:VAL:HG23	1:B:134:VAL:HG23	1.35	1.03
1:A:150:ASP:O	1:A:152:LYS:N	2.00	0.95
1:A:84:VAL:CG1	1:B:84:VAL:HG11	2.04	0.87
1:A:134:VAL:HG23	1:B:134:VAL:CG2	2.06	0.85
1:E:84:VAL:HG11	1:F:84:VAL:HG11	1.58	0.85
1:A:84:VAL:HG11	1:B:84:VAL:CG1	2.08	0.84
1:A:156:SER:HB2	1:A:158:GLN:NE2	1.97	0.78
1:A:136:ARG:HD3	1:A:137:LEU:HD13	1.67	0.77
1:B:126:LYS:HZ2	1:F:146:VAL:CG1	1.98	0.77
1:E:147:ARG:HD2	1:F:148:ILE:CD1	2.16	0.76
1:B:126:LYS:NZ	1:F:146:VAL:HG11	2.01	0.76
1:E:147:ARG:HD2	1:F:148:ILE:HD13	1.70	0.73
1:E:144:LEU:HD22	1:F:148:ILE:HD12	1.75	0.68
1:E:63:CYS:O	1:E:67:THR:HG23	1.96	0.65
1:E:141:ASN:ND2	1:F:141:ASN:OD1	2.30	0.64
1:B:126:LYS:HZ2	1:F:146:VAL:HG11	1.61	0.63
1:B:147:ARG:O	1:B:151:LYS:HB2	2.00	0.62
1:E:150:ASP:HA	1:E:153:TYR:HB3	1.81	0.61
1:A:151:LYS:HA	1:A:154:TYR:HD2	1.66	0.59
1:A:84:VAL:CG1	1:B:84:VAL:CG1	2.76	0.58
1:E:137:LEU:HD21	1:F:138:ARG:HG3	1.84	0.58
1:B:112:LYS:NZ	1:F:135:GLU:OE2	2.29	0.58
1:B:126:LYS:HZ2	1:F:146:VAL:CG2	2.17	0.58
1:B:126:LYS:NZ	1:F:146:VAL:HG21	2.19	0.57
1:F:136:ARG:HD2	1:F:137:LEU:HD13	1.85	0.57
1:A:151:LYS:HA	1:A:154:TYR:CD2	2.41	0.55

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:63:CYS:O	1:B:67:THR:HG23	2.09	0.52
1:E:154:TYR:HB3	1:E:155:PRO:HD3	1.92	0.51
1:A:133:GLU:OE1	1:A:136:ARG:NH1	2.43	0.51
1:A:141:ASN:O	1:A:145:SER:OG	2.25	0.51
1:E:95:VAL:HG23	1:F:95:VAL:HG23	1.93	0.51
1:E:134:VAL:HG23	1:F:134:VAL:HG23	1.91	0.51
1:E:136:ARG:HD3	1:E:137:LEU:HD13	1.95	0.48
1:B:112:LYS:NZ	1:F:139:ARG:HH22	2.11	0.48
1:A:134:VAL:CG2	1:B:134:VAL:HG23	2.25	0.48
1:B:112:LYS:HZ2	1:F:139:ARG:HH22	1.60	0.47
1:B:126:LYS:HZ3	1:F:146:VAL:HG11	1.76	0.47
1:A:141:ASN:OD1	1:B:141:ASN:ND2	2.40	0.46
1:E:147:ARG:CD	1:F:148:ILE:HD13	2.44	0.46
1:E:95:VAL:CG2	1:F:95:VAL:HG23	2.46	0.46
1:E:144:LEU:HD22	1:F:148:ILE:CD1	2.45	0.46
1:A:138:ARG:HG3	1:B:137:LEU:HD21	1.97	0.45
1:E:95:VAL:HG23	1:F:95:VAL:CG2	2.46	0.45
1:E:151:LYS:HZ2	1:F:151:LYS:C	2.20	0.45
1:E:148:ILE:HG13	1:F:148:ILE:HG13	2.00	0.43
1:B:126:LYS:HZ1	1:F:146:VAL:HG21	1.81	0.43
1:E:84:VAL:HG11	1:F:84:VAL:CG1	2.38	0.43
1:E:61:MET:O	1:E:64:ARG:HB2	2.19	0.42
1:A:110:GLN:O	1:A:114:GLU:HB2	2.19	0.42
1:A:135:GLU:HG3	1:E:151:LYS:HZ1	1.85	0.41
1:F:83:ASP:O	1:F:87:GLN:HG3	2.21	0.41
1:A:148:ILE:HD11	1:B:148:ILE:HG13	2.03	0.41
1:A:125:HIS:O	1:A:128:GLN:HB2	2.20	0.41
1:B:82:GLN:O	1:B:85:GLU:HB3	2.22	0.40
1:B:137:LEU:HD12	1:B:137:LEU:HA	1.85	0.40

All (1) 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)
1:A:64:ARG:NH2	1:A:75:THR:OG1[2_656]	2.16	0.04

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	107/121 (88%)	105 (98%)	1 (1%)	1 (1%)	19	60
1	B	104/121 (86%)	102 (98%)	2 (2%)	0	100	100
1	E	104/121 (86%)	103 (99%)	1 (1%)	0	100	100
1	F	100/121 (83%)	98 (98%)	2 (2%)	0	100	100
All	All	415/484 (86%)	408 (98%)	6 (1%)	1 (0%)	49	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	151	LYS

### 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	A	92/101 (91%)	80 (87%)	12 (13%)	4	21
1	B	89/101 (88%)	80 (90%)	9 (10%)	8	32
1	E	89/101 (88%)	76 (85%)	13 (15%)	3	16
1	F	85/101 (84%)	74 (87%)	11 (13%)	5	21
All	All	355/404 (88%)	310 (87%)	45 (13%)	5	22

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	LEU
1	A	58	ARG
1	A	76	GLU
1	A	83	ASP
1	A	123	LEU
1	A	128	GLN
1	A	135	GLU
1	A	136	ARG
1	A	137	LEU
1	A	140	GLU
1	A	145	SER
1	A	152	LYS
1	B	50	SER
1	B	57	LEU
1	B	76	GLU
1	B	123	LEU
1	B	128	GLN
1	B	135	GLU
1	B	137	LEU
1	B	145	SER
1	B	152	LYS
1	E	57	LEU
1	E	58	ARG
1	E	64	ARG
1	E	76	GLU
1	E	123	LEU
1	E	128	GLN
1	E	135	GLU
1	E	136	ARG
1	E	137	LEU
1	E	141	ASN
1	E	145	SER
1	E	150	ASP
1	E	153	TYR
1	F	57	LEU
1	F	58	ARG
1	F	64	ARG
1	F	76	GLU
1	F	123	LEU
1	F	128	GLN
1	F	135	GLU
1	F	136	ARG
1	F	137	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	145	SER
1	F	150	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	110	GLN
1	A	158	GLN
1	B	110	GLN
1	E	110	GLN
1	E	128	GLN
1	F	110	GLN

### 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](#)

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	109/121 (90%)	0.14	4 (3%)	41	27	34, 65, 107, 164	0
1	B	106/121 (87%)	0.05	3 (2%)	53	39	38, 70, 131, 190	0
1	E	106/121 (87%)	0.06	3 (2%)	53	39	30, 54, 125, 216	0
1	F	102/121 (84%)	-0.06	2 (1%)	65	51	32, 58, 96, 157	0
All	All	423/484 (87%)	0.05	12 (2%)	53	39	30, 62, 122, 216	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	50	SER	4.8
1	E	154	TYR	4.3
1	A	51	GLU	3.9
1	B	155	PRO	3.6
1	B	153	TYR	3.4
1	B	154	TYR	3.1
1	F	51	GLU	2.9
1	A	87	GLN	2.6
1	E	155	PRO	2.6
1	A	157	SER	2.2
1	F	150	ASP	2.0
1	E	51	GLU	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.