



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:47 PM GMT

PDB ID : 4JIO
Title : Bro1 V domain and ubiquitin
Authors : Pashkova, N.; Gakhar, L.; Piper, R.C.
Deposited on : 2013-03-06
Resolution : 3.60 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

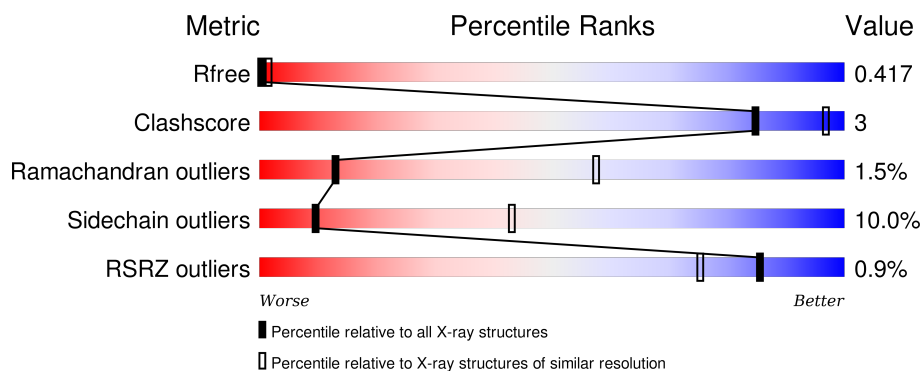
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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}	91344	1408 (3.80-3.40)
Clashscore	102246	1010 (3.74-3.46)
Ramachandran outliers	100387	1007 (3.76-3.44)
Sidechain outliers	100360	1007 (3.76-3.44)
RSRZ outliers	91569	1003 (3.78-3.42)

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 ≥ 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	344	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> % 87% • 10% </div> </div>
1	B	344	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> % 88% • • 7% </div> </div>
2	U	76	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 93% • • </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3536 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 BRO1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	310	Total	C	N	O	Se	0	0	0
			1559	934	310	310	5			
1	B	320	Total	C	N	O	Se	0	0	0
			1607	962	320	320	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	365	ILE	-	EXPRESSION TAG	UNP G0VBG6
A	366	ASP	-	EXPRESSION TAG	UNP G0VBG6
A	367	PRO	-	EXPRESSION TAG	UNP G0VBG6
A	368	PHE	-	EXPRESSION TAG	UNP G0VBG6
A	369	THR	-	EXPRESSION TAG	UNP G0VBG6
A	417	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6
A	418	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6
A	419	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6
B	365	ILE	-	EXPRESSION TAG	UNP G0VBG6
B	366	ASP	-	EXPRESSION TAG	UNP G0VBG6
B	367	PRO	-	EXPRESSION TAG	UNP G0VBG6
B	368	PHE	-	EXPRESSION TAG	UNP G0VBG6
B	369	THR	-	EXPRESSION TAG	UNP G0VBG6
B	417	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6
B	418	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6
B	419	ALA	LYS	ENGINEERED MUTATION	UNP G0VBG6

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	U	75	Total	C	N	O	0	0	0
			370	220	75	75			

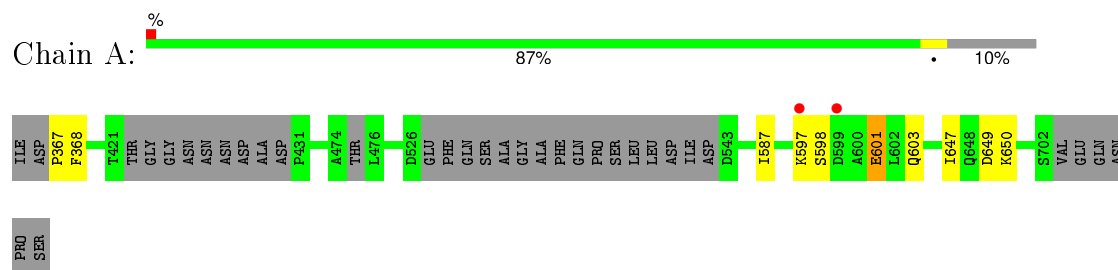
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	28	MET	ALA	ENGINEERED MUTATION	UNP P62987

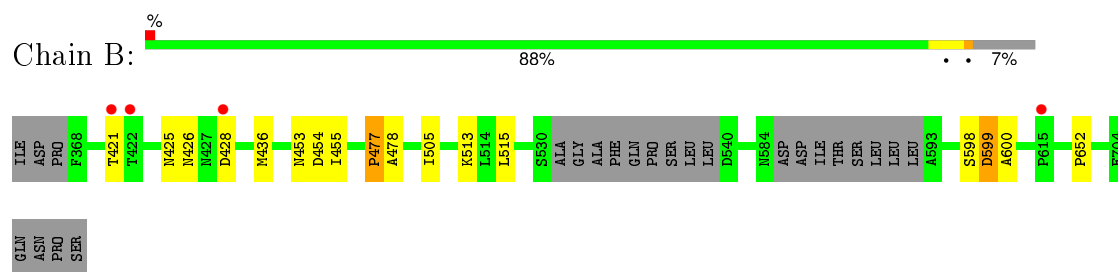
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

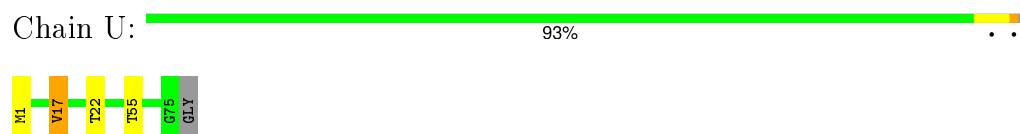
- Molecule 1: BRO1



- Molecule 1: BRO1



- Molecule 2: Ubiquitin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	63.43Å 92.32Å 229.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.49 – 3.60 34.49 – 3.60	Depositor EDS
% Data completeness (in resolution range)	98.6 (34.49-3.60) 98.5 (34.49-3.60)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.33 (at 3.56Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.366 , 0.421 0.352 , 0.417	Depositor DCC
R_{free} test set	1611 reflections (10.04%)	DCC
Wilson B-factor (Å ²)	131.3	Xtriage
Anisotropy	0.674	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 1428.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	6 of 16272 reflections (0.037%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	3536	wwPDB-VP
Average B, all atoms (Å ²)	138.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.70% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.20	0/1550	0.40	0/2146
1	B	0.20	0/1599	0.42	1/2215 (0.0%)
2	U	0.18	0/369	0.35	0/512
All	All	0.20	0/3518	0.40	1/4873 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	599	ASP	N-CA-C	5.06	124.67	111.00

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	1559	0	715	7	0
1	B	1607	0	740	6	0
2	U	370	0	157	2	0
All	All	3536	0	1612	15	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 3.

All (15) 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)
2:U:22:THR:HA	2:U:55:THR:HA	1.82	0.61
1:B:599:ASP:N	1:B:600:ALA:HA	2.18	0.58
1:A:601:GLU:HA	1:A:603:GLN:N	2.23	0.54
1:B:598:SER:C	1:B:600:ALA:HA	2.28	0.54
1:A:597:LYS:H	1:A:598:SER:HA	1.73	0.54
1:A:601:GLU:HA	1:A:603:GLN:H	1.71	0.54
1:A:649:ASP:H	1:A:650:LYS:HA	1.77	0.48
1:A:597:LYS:N	1:A:598:SER:HA	2.30	0.47
2:U:1:MET:N	2:U:17:VAL:O	2.41	0.46
1:B:477:PRO:HA	1:B:478:ALA:HA	1.58	0.46
1:A:601:GLU:CA	1:A:603:GLN:H	2.30	0.43
1:B:513:LYS:C	1:B:515:LEU:H	2.22	0.43
1:B:425:ASN:HA	1:B:426:ASN:HA	1.70	0.41
1:B:453:ASN:O	1:B:455:ILE:N	2.53	0.41
1:A:367:PRO:HA	1:A:368:PHE:HA	1.64	0.41

There are no symmetry-related clashes.

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	302/344 (88%)	277 (92%)	22 (7%)	3 (1%)	19	66
1	B	314/344 (91%)	282 (90%)	26 (8%)	6 (2%)	10	53
2	U	73/76 (96%)	70 (96%)	2 (3%)	1 (1%)	14	59
All	All	689/764 (90%)	629 (91%)	50 (7%)	10 (2%)	13	57

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	587	ILE
1	B	477	PRO
1	B	652	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	647	ILE
1	B	421	THR
1	B	505	ILE
2	U	17	VAL
1	A	601	GLU
1	B	428	ASP
1	B	454	ASP

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	5/306 (2%)	5 (100%)	0	100	100
1	B	5/306 (2%)	4 (80%)	1 (20%)	1	11
All	All	10/612 (2%)	9 (90%)	1 (10%)	9	43

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	436	MSE

Some 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 [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, 95th 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 ²)	Q<0.9
1	A	305/344 (88%)	-0.81	2 (0%) 89 81	76, 131, 233, 302	0
1	B	315/344 (91%)	-0.84	4 (1%) 79 66	39, 121, 229, 305	0
2	U	75/76 (98%)	-0.76	0 100 100	77, 144, 215, 257	0
All	All	695/764 (90%)	-0.82	6 (0%) 85 75	39, 131, 233, 305	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	421	THR	3.9
1	B	422	THR	3.5
1	A	599	ASP	2.8
1	B	615	PRO	2.2
1	A	597	LYS	2.1
1	B	428	ASP	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.