



wwPDB X-ray Structure Validation Summary Report

Nov 2, 2015 – 07:20 AM EST

PDB ID : 5DFD
Title : Crystal structure of BRD2(BD2) W370F mutant with ligand 28 bound
Authors : Tallant, C.; Baud, M.; Lin-Shiao, E.; Chirgadze, D.Y.; Ciulli, A.
Deposited on : 2015-08-26
Resolution : 1.50 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
A user guide is available at
<http://wwpdb.org/validation/ValidationPDFNotes.html>

The following versions of software and data (see [references](#)) were used in the production of this report:

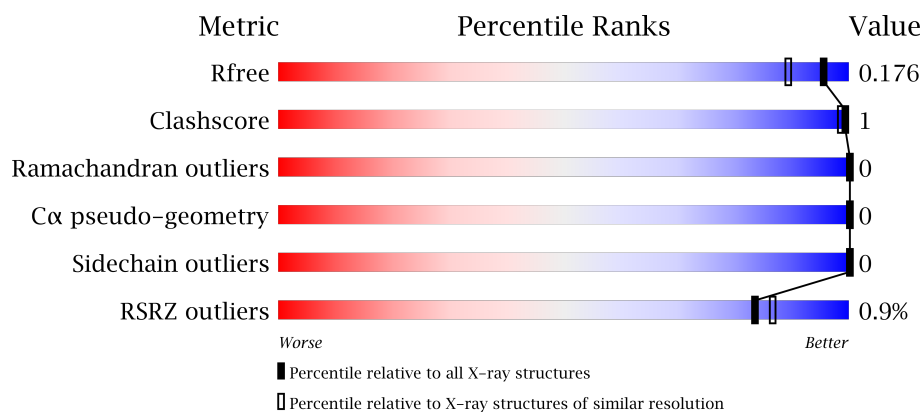
MolProbity : 4.02b-467
Mogul : 1.17 November 2013 2014 2015
Xtriage (Phenix) : 1.9-1692
EDS : rb-20025774
Percentile statistics : 25063
Refmac : 5.8.0122
CCP4 : 6.1.3
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20025774

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 1.50 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}	83175	1879 (1.50-1.50)
Clashscore	93826	2076 (1.50-1.50)
Ramachandran outliers	92073	2023 (1.50-1.50)
Cα geometry	92159	2022 (1.50-1.50)
Sidechain outliers	92050	2021 (1.50-1.50)
RSRZ outliers	83181	1880 (1.50-1.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	114	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	59E	A	501	-	-	-	X
3	GOL	A	502	-	-	-	X

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 1018 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.

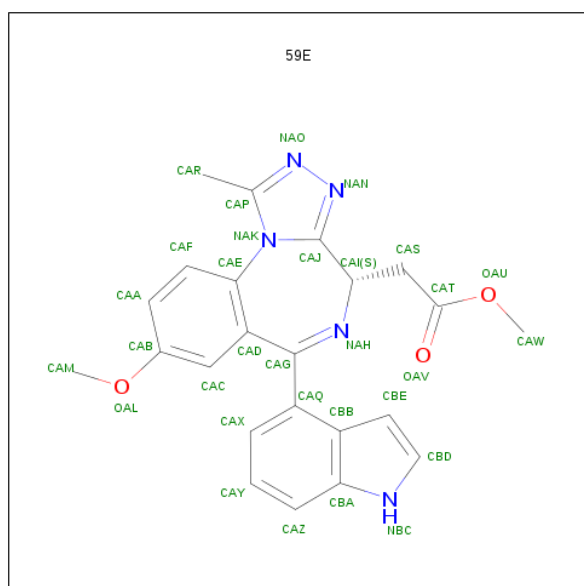
- Molecule 1 is a protein called Bromodomain-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	108	Total	C	N	O	S	0	1	0
			895	572	156	160	7			

There are 3 discrepancies between the modelled and reference sequences:

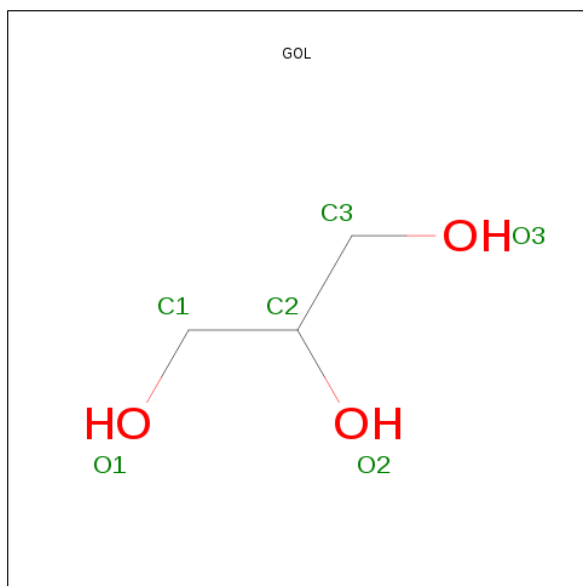
Chain	Residue	Modelled	Actual	Comment	Reference
A	342	SER	-	expression tag	UNP P25440
A	343	MET	-	expression tag	UNP P25440
A	370	PHE	TRP	engineered mutation	UNP P25440

- Molecule 2 is methyl [(4S)-6-(1H-indol-4-yl)-8-methoxy-1-methyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepin-4-yl]acetate (three-letter code: 59E) (formula: C₂₃H₂₁N₅O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			31	23	5	3		

- 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 water.

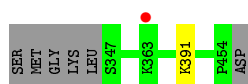
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	86	Total	O	0	0
			86	86		

3 Residue-property plots

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: Bromodomain-containing protein 2

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2 ₁	Depositor
Cell constants a, b, c, α , β , γ	32.12Å 52.44Å 71.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.67 – 1.50 29.67 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.67-1.50) 100.0 (29.67-1.50)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.01	Depositor
$\langle I/\sigma(I) \rangle$ ¹	12.54 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.160 , 0.175 0.160 , 0.176	Depositor DCC
R_{free} test set	1025 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	12.5	Xtriage
Anisotropy	0.146	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 38.9	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	0 of 20134 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	1018	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.73% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 59E

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	0.46	0/923	0.65	0/1242

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	895	0	873	1	0
2	A	31	0	0	0	0
3	A	6	0	8	0	0
4	A	86	0	0	0	1
All	All	1018	0	881	1	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (1) close contacts within the same asymmetric unit are listed below. Note that the hydrogen positions were calculated by MolProbity and these may be different from any deposited coordinates. The reason for using them is that the clashscore method is a tool to assess close contacts between **heavy atoms**, and specifically not a validation tool for hydrogen positions. Therefore,

the algorithm must be used consistently for every structure.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:391:LYS:HA	1:A:391:LYS:HE2	2.01	0.41

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	Distance(Å)	Clash(Å)
4:A:601:HOH:O	4:A:635:HOH:O[2_545]	0.92	1.28

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	107/114 (94%)	107 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

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	96/100 (96%)	96 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 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	59E	A	501	-	28,35,35	1.95	6 (21%)	35,51,51	1.99	9 (25%)
3	GOL	A	502	-	5,5,5	0.41	0	5,5,5	0.41	0

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	59E	A	501	-	-	0/12/28/28	0/4/5/5
3	GOL	A	502	-	-	0/4/4/4	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	59E	CAD-CAG	-4.45	1.42	1.49
2	A	501	59E	CAZ-CBA	-3.93	1.34	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	59E	CAQ-CBB	-3.03	1.37	1.43
2	A	501	59E	CBB-CBA	-2.50	1.35	1.42
2	A	501	59E	CAI-NAH	3.30	1.50	1.47

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	59E	CAF-CAE-CAD	-3.95	119.76	122.83
2	A	501	59E	CAF-CAE-NAK	-3.64	115.67	119.49
2	A	501	59E	CAD-CAG-NAH	-3.01	120.28	124.46
2	A	501	59E	CAC-CAD-CAG	-2.46	115.06	118.87
2	A	501	59E	NAN-CAJ-NAK	-2.05	109.49	111.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.

5.9 Trace atom model geometry

5.9.1 Abnormal distances between consecutive trace atoms ⓘ

The following table provides a summary of the distances observed between consecutive C α atoms (in protein chains) or P atoms (in RNA or DNA chains).

Mol	Chain	Analysed	Outliers
1	A	105/114 (92.1%)	0/105 (0.0%)

5.9.2 C α pseudo-geometry ⓘ

There are no C α pseudo-geometry outliers to report.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	108/114 (94%)	0.02	1 (0%) 83 86	8, 12, 23, 27	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	363	LYS	2.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	GOL	A	502	6/6	0.75	0.30	6.46	25,30,32,35	0
2	59E	A	501	31/31	0.89	0.18	2.85	13,16,21,24	0

6.5 Other polymers

There are no such residues in this entry.