

# wwPDB X-ray Structure Validation Summary Report (i)

#### May 16, 2020 - 09:07 am BST

PDB ID	:	6FNX
Title	:	FIRST DOMAIN OF HUMAN BROMODOMAIN BRD4 IN COMPLEX
		WITH INHIBITOR F1
Authors	:	Raux, B.; Betzi, S.
Deposited on		
$\operatorname{Resolution}$	:	1.19 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 https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

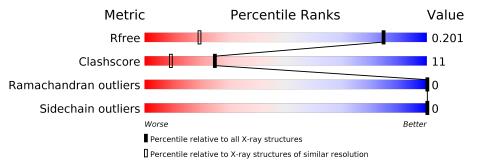
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.11
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
$\operatorname{Refmac}$	:	5.8.0158
$\operatorname{CCP4}$	:	$7.0.044 (\mathrm{Gargrove})$
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

# 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 1.19 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R <sub>free</sub>	130704	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)

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 >=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%

Mol	Chain	Length	Quality of chain					
1	А	127	83%	16%	•			

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
3	EDO	А	204	-	Х	-	-



# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1439 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.

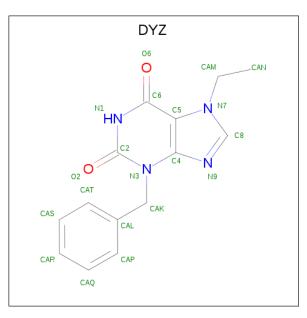
• Molecule 1 is a protein called Bromodomain-containing protein 4.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	127	Total 1127	C 728	N 186	O 206	S 7	0	7	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	42	SER	-	expression tag	UNP 060885
А	43	MET	-	expression tag	UNP 060885

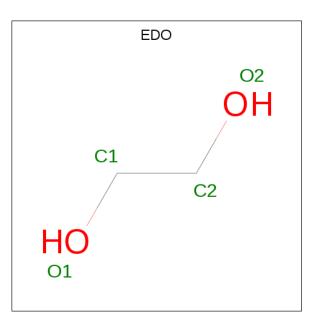
• Molecule 2 is 7-ethyl-3-(phenylmethyl) purine-2,6-dione (three-letter code: DYZ) (formula:  $\rm C_{14}H_{14}N_4O_2).$ 



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf		
2	А	1	Total 20	C 14	N 4	O 2	0	0

• Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	$\begin{array}{ccc} \text{Total}  \text{C}  \text{O} \\ 4  2  2 \end{array}$	0	0
3	А	1	$\begin{array}{ccc} \text{Total}  \text{C}  \text{O} \\ 4  2  2 \end{array}$	0	0
3	А	1	$\begin{array}{ccc} \text{Total}  \text{C}  \text{O} \\ 4  2  2 \end{array}$	0	0
3	А	1	$\begin{array}{ccc} \text{Total}  \text{C}  \text{O} \\ 4  2  2 \end{array}$	0	0

• Molecule 4 is water.

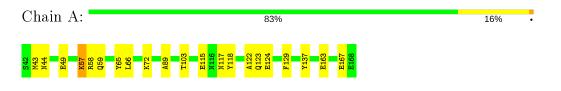
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	276	Total         O           276         276	0	0



# 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 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.

• Molecule 1: Bromodomain-containing protein 4





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	37.86Å $44.17$ Å $78.16$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	39.08 - 1.19	Depositor
	39.08 - 1.19	EDS
% Data completeness	98.5 (39.08-1.19)	Depositor
(in resolution range)	90.7(39.08-1.19)	EDS
R <sub>merge</sub>	0.10	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$0.44 (at 1.19 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0189	Depositor
B B.	0.232 , $0.276$	Depositor
$R, R_{free}$	0.160 , $0.201$	DCC
$R_{free}$ test set	2092 reflections $(5.00%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	7.8	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.28 , $40.3$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.48, \langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1439	wwPDB-VP
Average B, all atoms $(Å^2)$	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 8.88% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: EDO, DYZ

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).

Mal	Chain	Bo	nd lengths	Bond angles		
Mol		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	1.63	7/1160~(0.6%)	1.36	5/1573~(0.3%)	

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
1	А	115	GLU	CD-OE2	8.52	1.35	1.25
1	А	57	LYS	CE-NZ	-7.80	1.29	1.49
1	А	124	GLU	CD-OE1	-6.52	1.18	1.25
1	А	129	PHE	CG-CD2	5.88	1.47	1.38
1	А	137	TYR	CD2-CE2	-5.75	1.30	1.39

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	58	ARG	NE-CZ-NH1	9.18	124.89	120.30
1	А	57	LYS	CD-CE-NZ	-7.43	94.62	111.70
1	А	58	ARG	NE-CZ-NH2	-6.24	117.18	120.30
1	А	115	GLU	OE1-CD-OE2	5.36	129.73	123.30
1	А	65	TYR	CB-CG-CD2	-5.21	117.87	121.00

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	А	1127	0	1127	21	0
2	А	20	0	0	0	0
3	А	16	0	21	6	0
4	А	276	0	0	12	0
All	All	1439	0	1148	25	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 11.

The worst 5 of 25 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:123[B]:GLN:OE1	4:A:301:HOH:O	1.61	1.18
1:A:123[A]:GLN:OE1	4:A:302:HOH:O	1.89	0.91
1:A:59:GLN:OE1	1:A:117[B]:ASN:OD1	1.93	0.86
1:A:44:ASN:HD21	1:A:89:ALA:H	1.20	0.84
3:A:204:EDO:H22	4:A:479:HOH:O	1.82	0.80

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	5
1	А	133/127~(105%)	130~(98%)	3~(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	Analysed Rotameric		Percentiles	
1	А	128/120~(107%)	128~(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. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	44	ASN
1	А	54	ASN
1	А	78	GLN
1	А	130	ASN

#### 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)

5 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	Tune	Chain	Res	Link	Bo	ond leng	$\mathbf{ths}$	Bond angles		
	Type	Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	DYZ	А	201	-	16,22,22	2.92	5 (31%)	17,31,31	<mark>3.90</mark>	10 (58%)
3	EDO	А	204	-	$^{3,3,3}$	1.60	1 (33%)	2,2,2	2.70	1(50%)
3	EDO	А	202	-	$^{3,3,3}$	1.74	1 (33%)	2,2,2	0.83	0
3	EDO	А	205	3	$3,\!3,\!3$	0.36	0	2,2,2	0.96	0
3	EDO	А	203	3	$^{3,3,3}$	0.86	0	2,2,2	0.21	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	DYZ	А	201	-	-	0/6/6/6	0/3/3/3
3	EDO	А	204	-	-	1/1/1/1	-
3	EDO	А	202	-	-	1/1/1/1	-
3	EDO	А	205	3	-	1/1/1/1	-
3	EDO	А	203	3	-	0/1/1/1	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms		$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
2	А	201	DYZ	C6-N1	6.86	1.45	1.33
2	А	201	DYZ	C5-C4	-6.11	1.31	1.39
2	А	201	DYZ	CAM-N7	-4.76	1.45	1.49
2	А	201	DYZ	C6-C5	-2.94	1.36	1.41
3	А	202	EDO	O1-C1	2.84	1.56	1.42

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
2	А	201	DYZ	C6-N1-C2	9.03	122.77	115.14
2	А	201	DYZ	C5-C6-N1	-7.79	112.78	123.43
2	А	201	DYZ	C6-C5-C4	5.93	126.45	120.80
2	А	201	DYZ	CAK-N3-C2	4.91	123.67	117.92
2	А	201	DYZ	CAR-CAS-CAT	4.18	126.56	120.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	Atoms
3	А	205	EDO	O1-C1-C2-O2
		2		,



Continued from previous page...

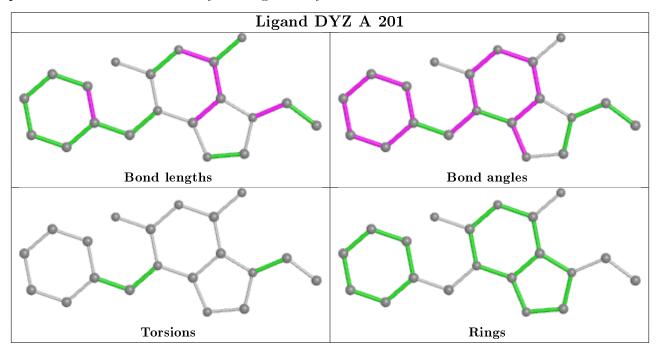
Mol	Chain	Res	Type	Atoms
3	А	202	EDO	O1-C1-C2-O2
3	А	204	EDO	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	А	204	EDO	3	0
3	А	205	EDO	3	0
3	А	203	EDO	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates (i)

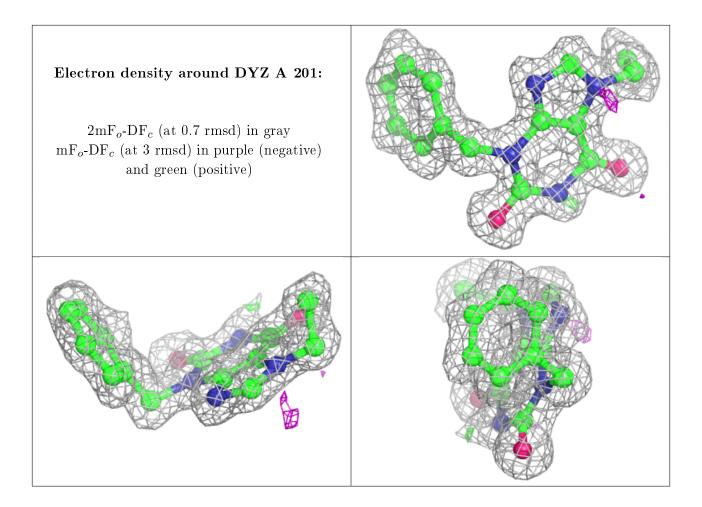
Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





### 6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

