

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

Jan 30, 2024 – 03:48 pm GMT

PDB ID : 8Q34

Title : Crystal structure of the first bromodomain of human BRD4 in complex with

the ligand ZZ001229a

Authors: MacLean, E.M.; Gao, Q.; Williams, E.; Balcomb, B.H.; von Delft, F.; Bajusz,

D.; Keeley, A.; Abranyi-Balogh, P.; Koekemoer, L.; Keseru, G.M.

Deposited on : 2023-08-03

Resolution : 1.48 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 types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

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

MolProbity: 4.02b-467

Mogul : 1.8.4, CSD as541be (2020)

Xtriage (Phenix) : 1.13

EDS : 2.36

buster-report : 1.1.7 (2018)

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Refmac : 5.8.0158

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

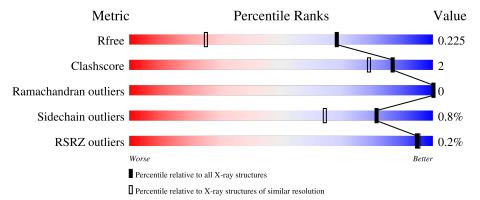
Validation Pipeline (wwPDB-VP) : 2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY\ DIFFRACTION$ 

The reported resolution of this entry is 1.48 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	Similar resolution
Metric	$(\#  ext{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$
$R_{free}$	130704	4690 (1.50-1.46)
Clashscore	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

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% 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	130	93%	5% •
1	В	130	93%	5% •
1	С	130	90%	7% •
1	D	130	92%	5% • •



# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9424 atoms, of which 4258 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.

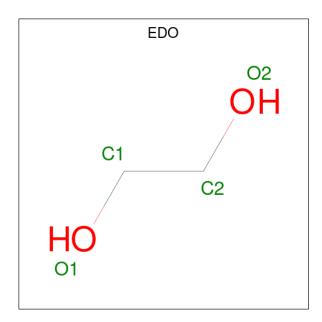
Mol	Chain	Residues		Atoms				ZeroOcc	AltConf	Trace	
1	Λ	127	Total	С	Н	N	О	S	0	2	0
1	A	121	2133	697	1056	177	196	7	0		0
1	В	127	Total	С	Н	N	О	S	0	2	0
1	Ъ	121	2136	698	1058	177	196	7	0		0
1	С	126	Total	С	Н	N	О	S	0	1	0
1		120	2093	687	1034	173	192	7	0	1	U
1	D	126	Total	С	Н	N	О	S	0	1	0
1	ש	120	2090	685	1034	173	191	7	U	1	

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	SER	-	expression tag	UNP O60885
A	43	MET	-	expression tag	UNP O60885
В	42	SER	-	expression tag	UNP O60885
В	43	MET	-	expression tag	UNP O60885
С	42	SER	-	expression tag	UNP O60885
С	43	MET	-	expression tag	UNP O60885
D	42	SER	-	expression tag	UNP O60885
D	43	MET	-	expression tag	UNP O60885

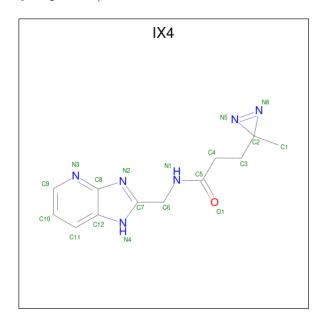
• Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total		Н	О	0	0
	11	1	9	2	5	2	O O	Ŭ.
2	В	1	Total	$\mathbf{C}$	Η	Ο	0	0
2	Б	1	9	2	5	2	0	U
2	C	1	Total	С	Н	О	0	0
2		1	9	2	5	2	0	U
2	D	1	Total	С	Н	О	0	0
2	ע	1	9	2	5	2	0	U

• Molecule 3 is  $\{N\}$ -(1  $\{H\}$ -imidazo[4,5-b]pyridin-2-ylmethyl)-3-(3-methyl-1,2-diazirin-3-yl)p ropanamide (three-letter code: IX4) (formula:  $C_{12}H_{14}N_6O$ ) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	Λ	1	Total	С	Н	N	О	0	0
3	A	1	33	12	14	6	1	U	0
3	В	1	Total	С	Н	N	О	0	0
3	Б	1	33	12	14	6	1	U	0
3	С	1	Total	С	Н	N	О	0	0
3		1	33	12	14	6	1	U	0
3	D	1	Total	С	Н	N	О	0	0
3	ט	1	33	12	14	6	1	U	0

### • Molecule 4 is water.

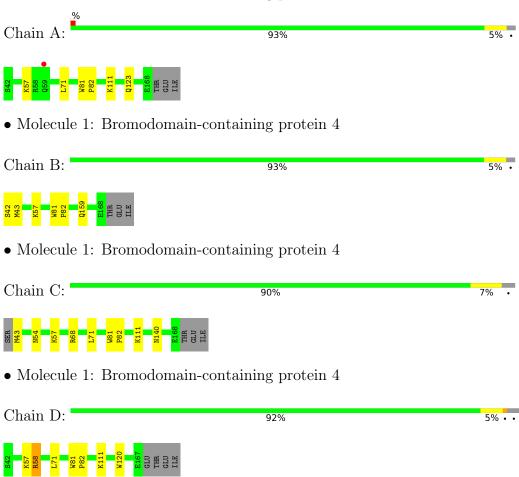
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	203	Total O 203 203	0	0
4	В	205	Total O 205 205	0	0
4	С	199	Total O 199 199	0	0
4	D	197	Total O 197 197	0	0



# 3 Residue-property plots (i)

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 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 4





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants	37.24Å 44.09Å 78.48Å	Donogiton
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.03^{\circ}$ $90.00^{\circ}$ $90.03^{\circ}$	Depositor
Resolution (Å)	78.48 - 1.48	Depositor
Resolution (A)	78.48 - 1.48	EDS
% Data completeness	90.8 (78.48-1.48)	Depositor
(in resolution range)	90.7 (78.48-1.48)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.15 (at 1.48Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
$R, R_{free}$	0.184 , $0.216$	Depositor
it, it free	0.193 , $0.225$	DCC
$R_{free}$ test set	3632 reflections $(4.81%)$	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.9	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.34, 33.1	EDS
L-test for twinning <sup>2</sup>	$< L > = 0.52, < L^2> = 0.36$	Xtriage
	0.467 for h,-k,-l	
Estimated twinning fraction	0.469  for -h,k,-l	Xtriage
	0.467  for -h,-k,l	
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	9424	wwPDB-VP
Average B, all atoms $(\mathring{A}^2)$	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 10.07% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $< L^2>$  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, IX4

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	Mol Chain		lengths	Bond angles		
MIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.40	0/1107	0.67	0/1506	
1	В	0.38	0/1108	0.67	0/1507	
1	С	0.37	0/1089	0.68	1/1483 (0.1%)	
1	D	0.37	0/1086	0.73	$2/1479 \ (0.1\%)$	
All	All	0.38	0/4390	0.69	3/5975 (0.1%)	

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	С	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	D	58	ARG	NE-CZ-NH2	-9.81	115.39	120.30
1	D	58	ARG	NE-CZ-NH1	6.37	123.49	120.30
1	С	57	LYS	CB-CA-C	-6.00	98.41	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	С	68	ARG	Sidechain



### 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	1077	1056	1071	4	0
1	В	1078	1058	1073	4	0
1	С	1059	1034	1050	4	0
1	D	1056	1034	1049	6	0
2	A	4	5	6	0	0
2	В	4	5	6	0	0
2	С	4	5	6	0	0
2	D	4	5	6	0	0
3	A	19	14	0	0	0
3	В	19	14	0	0	0
3	С	19	14	0	1	0
3	D	19	14	0	0	0
4	A	203	0	0	2	1
4	В	205	0	0	1	0
4	С	199	0	0	3	0
4	D	197	0	0	0	1
All	All	5166	4258	4267	18	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 2.

The worst 5 of 18 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{aligned}  ext{Interatomic} \  ext{distance} & ( ext{Å}) \end{aligned}$	$egin{aligned}  ext{Clash} \  ext{overlap } ( ext{Å}) \end{aligned}$	
1:D:58:ARG:HD2	1:D:120:TRP:CD2	2.36	0.61	
1:A:123:GLN:NE2	4:A:803:HOH:O	2.39	0.55	
1:D:58:ARG:HD3	1:D:120:TRP:O	2.07	0.54	
1:B:159[B]:GLN:HG3	4:B:858:HOH:O	2.08	0.53	
1:A:71[A]:LEU:HD21	1:A:111:LYS:HG3	1.93	0.50	

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	$egin{aligned} &  ext{Interatomic} \ &  ext{distance} \ &  ext{(Å)} \end{aligned}$	$egin{aligned} \operatorname{Clash} \ \operatorname{overlap}\ (\begin{subarray}{c} \begin{subarray}{c} \begi$
4:A:885:HOH:O	4:D:910:HOH:O[1_646]	1.66	0.54

### 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	Percen	ntiles
1	A	127/130 (98%)	126 (99%)	1 (1%)	0	100	100
1	В	127/130 (98%)	126 (99%)	1 (1%)	0	100	100
1	С	125/130~(96%)	124 (99%)	1 (1%)	0	100	100
1	D	125/130~(96%)	124 (99%)	1 (1%)	0	100	100
All	All	504/520 (97%)	500 (99%)	4 (1%)	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	122/123 (99%)	121 (99%)	1 (1%)	81 64
1	В	122/123 (99%)	121 (99%)	1 (1%)	81 64
1	С	119/123 (97%)	118 (99%)	1 (1%)	81 64
1	D	119/123 (97%)	118 (99%)	1 (1%)	81 64
All	All	482/492 (98%)	478 (99%)	4 (1%)	81 64

All (4) residues with a non-rotameric sidechain are listed below:



Mol	Chain	$\operatorname{Res}$	Type
1	A	57	LYS
1	В	57	LYS
1	С	43	MET
1	D	57	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	123	GLN
1	В	117	ASN
1	С	116	ASN
1	С	117	ASN
1	D	59	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 monosaccharides in this entry.

### 5.6 Ligand geometry (i)

8 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	Trons	Chain	Dag	Bond lengths		Bond angles				
IVIOI	Type	Chain	nes	Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	IX4	С	702	-	15,21,21	0.84	0	14,30,30	1.55	3 (21%)



Mol	Tuna	Chain	Dog	Link	Bond lengths			Bond angles			
MIOI	Type	Chain	Res	nes	nes   Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	EDO	С	701	-	3,3,3	0.15	0	2,2,2	0.16	0	
3	IX4	D	702	-	15,21,21	0.80	0	14,30,30	1.35	2 (14%)	
2	EDO	A	701	-	3,3,3	0.17	0	2,2,2	0.19	0	
2	EDO	D	701	-	3,3,3	0.10	0	2,2,2	0.16	0	
3	IX4	В	702	-	15,21,21	0.90	0	14,30,30	1.64	4 (28%)	
3	IX4	A	702	-	15,21,21	0.83	0	14,30,30	1.43	2 (14%)	
2	EDO	В	701	-	3,3,3	0.13	0	2,2,2	0.05	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
3	IX4	С	702	-	-	0/6/15/15	0/3/3/3
2	EDO	С	701	_	-	0/1/1/1	_
3	IX4	D	702	-	-	0/6/15/15	0/3/3/3
2	EDO	A	701	-	-	0/1/1/1	-
2	EDO	D	701	-	-	0/1/1/1	-
3	IX4	В	702	-	-	0/6/15/15	0/3/3/3
3	IX4	A	702	-	-	0/6/15/15	0/3/3/3
2	EDO	В	701	-	-	0/1/1/1	-

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
3	В	702	IX4	C8-C12-N4	-3.64	104.76	109.19
3	С	702	IX4	C8-C12-N4	-3.58	104.83	109.19
3	A	702	IX4	C8-C12-N4	-3.09	105.43	109.19
3	D	702	IX4	C8-C12-N4	-2.93	105.62	109.19
3	С	702	IX4	C10-C11-C12	-2.59	116.36	120.08

There are no chirality outliers.

There are no torsion outliers.

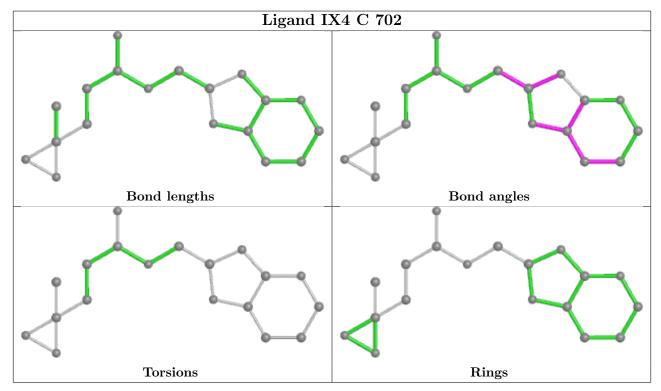
There are no ring outliers.

1 monomer is involved in 1 short contact:

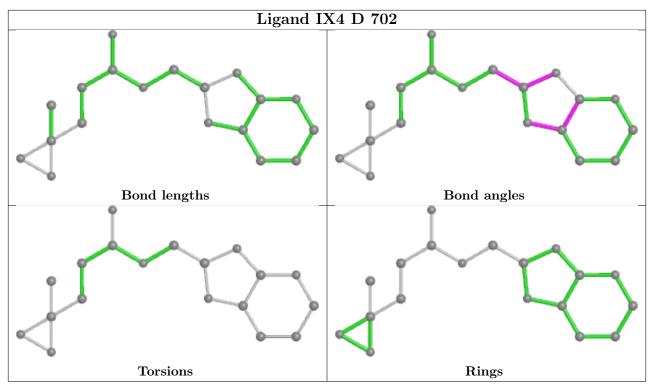
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	С	702	IX4	1	0

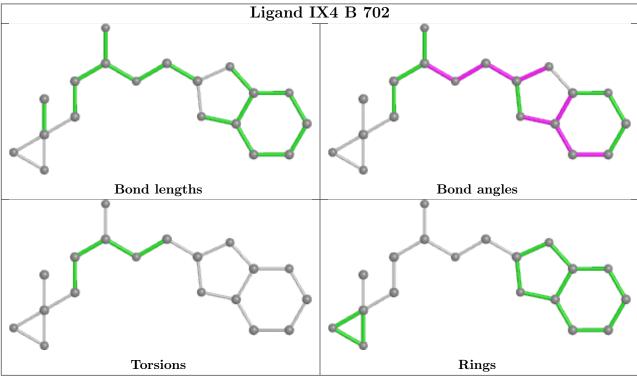


The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

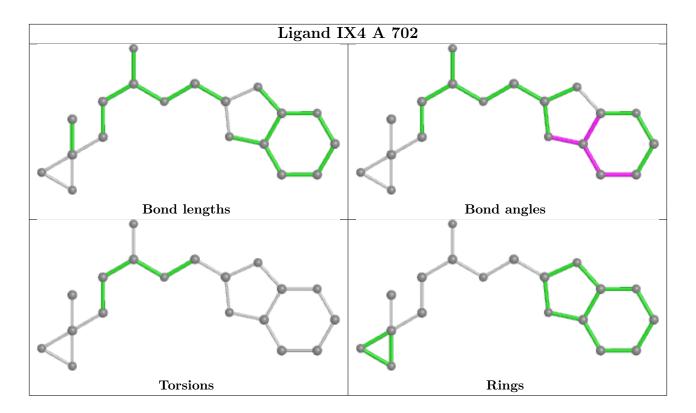












# 5.7 Other polymers (i)

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)

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,  $95^{th}$  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>	# RSRZ > 2	$\mathbf{OWAB}(\mathbf{\mathring{A}}^2)$	Q < 0.9
1	A	127/130 (97%)	-0.56	1 (0%) 86 88	10, 14, 28, 50	0
1	В	127/130 (97%)	-0.56	0 100 100	10, 13, 25, 47	0
1	С	126/130 (96%)	-0.58	0 100 100	10, 13, 24, 50	0
1	D	126/130 (96%)	-0.58	0 100 100	10, 13, 25, 37	0
All	All	506/520 (97%)	-0.57	1 (0%) 95 95	10, 14, 26, 50	0

### All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	59	GLN	2.1

# 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 monosaccharides in this entry.

### 6.4 Ligands (i)

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. The B-factors column lists the minimum, median,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	EDO	A	701	4/4	0.89	0.13	23,26,27,27	0

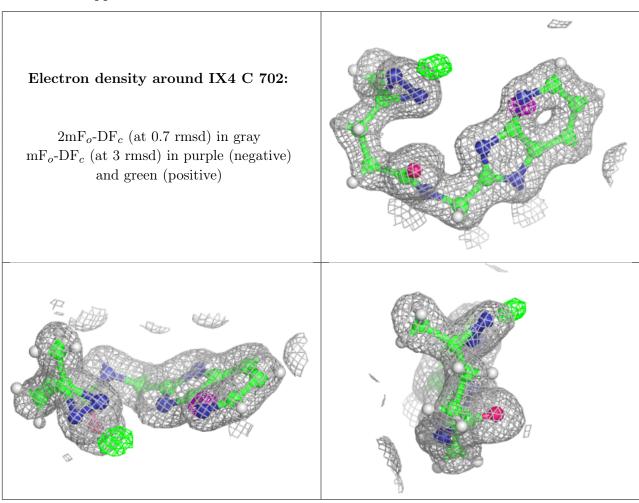
Continued on next page...



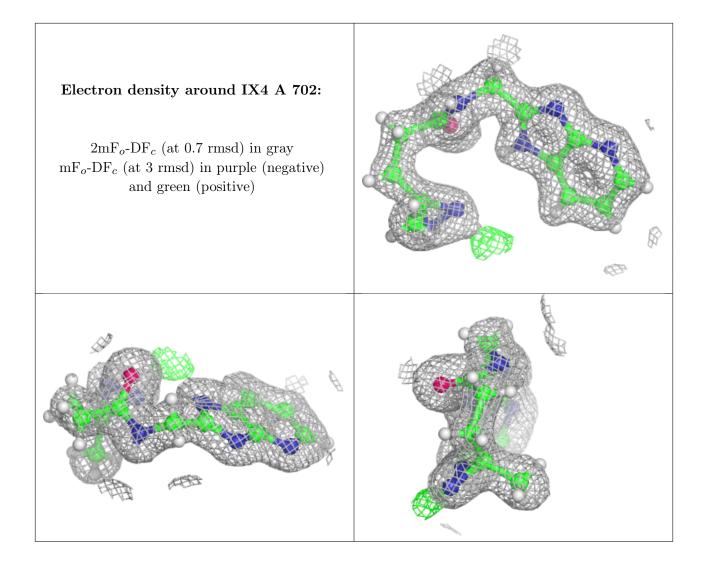
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	EDO	В	701	4/4	0.89	0.14	21,25,26,27	0
2	EDO	С	701	4/4	0.90	0.17	21,24,25,26	0
3	IX4	С	702	19/19	0.94	0.08	12,14,20,22	0
3	IX4	A	702	19/19	0.96	0.07	12,15,20,22	0
3	IX4	В	702	19/19	0.96	0.07	12,14,20,23	0
2	EDO	D	701	4/4	0.96	0.08	22,24,25,26	0
3	IX4	D	702	19/19	0.96	0.07	13,14,19,22	0

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.



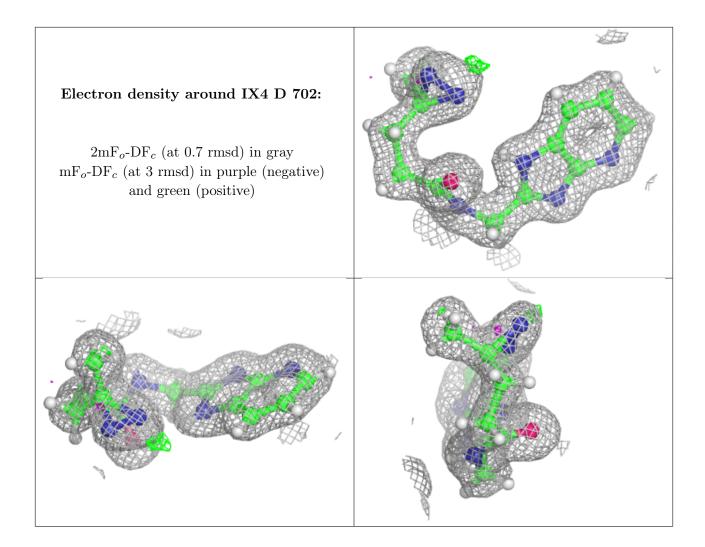






# Electron density around IX4 B 702: 2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative) and green (positive)





# 6.5 Other polymers (i)

There are no such residues in this entry.

