Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2020 – 03:42 am BST

PDB ID	:	6I99
Title	:	Bone Marrow Tyrosine Kinase in Chromosome X in complex with a newly designed covalent inhibitor JS24
Authors	:	Sousa, B.B.; Matias, P.M.; Marques, M.C.; Seixas, J.D.; Bernardes, G.J.L.
Deposited on	:	2018-11-22
Resolution	:	2.00 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
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:

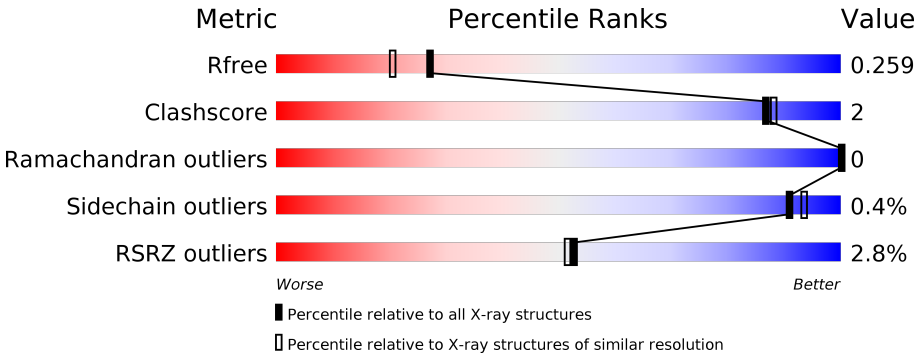
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

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 2.00 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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% 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	278	2% 90% 5% ·
1	B	278	4% 87% 6% 6%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8804 atoms, of which 4294 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 Cytoplasmic tyrosine-protein kinase BMX.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	266	Total	C	H	N	O	S	0	0	0
			4318	1410	2142	353	396	17			
1	B	261	Total	C	H	N	O	S	0	0	0
			4234	1379	2102	347	389	17			

There are 34 discrepancies between the modelled and reference sequences:

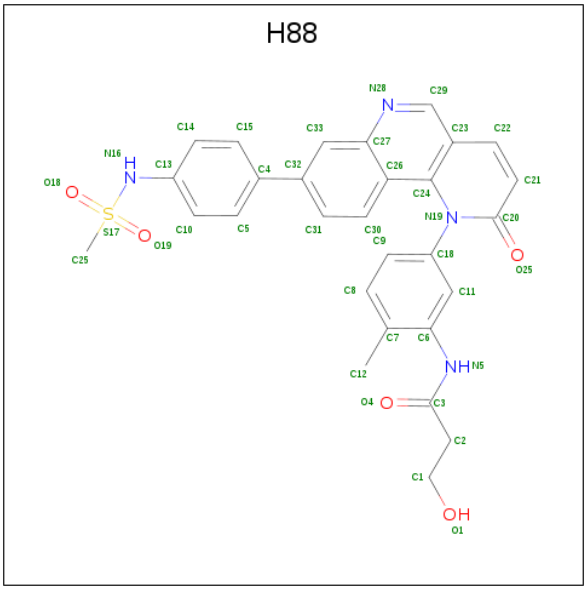
Chain	Residue	Modelled	Actual	Comment	Reference
A	398	HIS	-	expression tag	UNP P51813
A	399	HIS	-	expression tag	UNP P51813
A	400	HIS	-	expression tag	UNP P51813
A	401	HIS	-	expression tag	UNP P51813
A	402	HIS	-	expression tag	UNP P51813
A	403	HIS	-	expression tag	UNP P51813
A	404	GLU	-	expression tag	UNP P51813
A	405	ASN	-	expression tag	UNP P51813
A	406	LEU	-	expression tag	UNP P51813
A	407	TYR	-	expression tag	UNP P51813
A	408	PHE	-	expression tag	UNP P51813
A	409	GLN	-	expression tag	UNP P51813
A	410	GLY	-	expression tag	UNP P51813
A	432	LYS	GLN	engineered mutation	UNP P51813
A	611	MET	GLN	engineered mutation	UNP P51813
A	617	THR	ASP	engineered mutation	UNP P51813
A	620	GLU	GLN	engineered mutation	UNP P51813
B	398	HIS	-	expression tag	UNP P51813
B	399	HIS	-	expression tag	UNP P51813
B	400	HIS	-	expression tag	UNP P51813
B	401	HIS	-	expression tag	UNP P51813
B	402	HIS	-	expression tag	UNP P51813
B	403	HIS	-	expression tag	UNP P51813
B	404	GLU	-	expression tag	UNP P51813
B	405	ASN	-	expression tag	UNP P51813

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	406	LEU	-	expression tag	UNP P51813
B	407	TYR	-	expression tag	UNP P51813
B	408	PHE	-	expression tag	UNP P51813
B	409	GLN	-	expression tag	UNP P51813
B	410	GLY	-	expression tag	UNP P51813
B	432	LYS	GLN	engineered mutation	UNP P51813
B	611	MET	GLN	engineered mutation	UNP P51813
B	617	THR	ASP	engineered mutation	UNP P51813
B	620	GLU	GLN	engineered mutation	UNP P51813

- Molecule 2 is {N}-[2-methyl-5-[8-[4-(methylsulfonylamino)phenyl]-2-oxidanylidene-benzo[h][1,6]naphthyridin-1-yl]phenyl]-3-oxidanyl-propanamide (three-letter code: H88) (formula: $C_{29}H_{26}N_4O_5S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			63	29	25	4	4	1		
2	B	1	Total	C	H	N	O	S	0	0
			63	29	25	4	4	1		

- Molecule 3 is water.

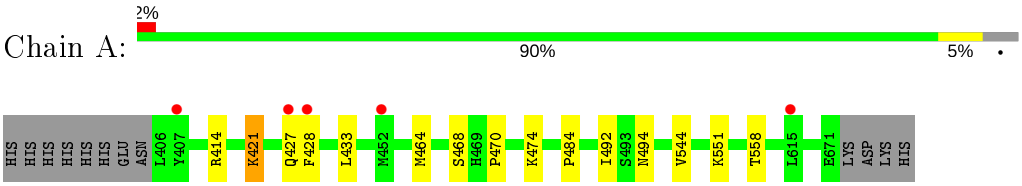
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	73	Total	O	0	0
			73	73		
3	B	53	Total	O	0	0
			53	53		

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 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.

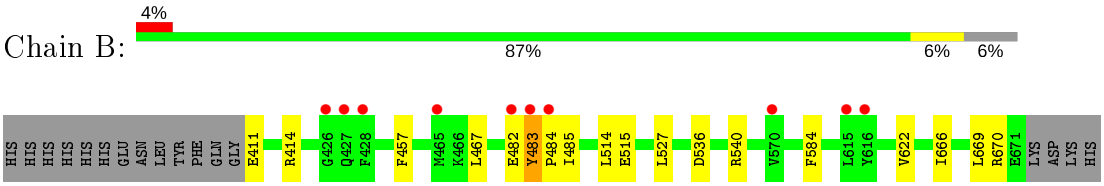
- Molecule 1: Cytoplasmic tyrosine-protein kinase BMX

Chain A:



- Molecule 1: Cytoplasmic tyrosine-protein kinase BMX

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	66.71Å 63.29Å 74.96Å 90.00° 104.55° 90.00°	Depositor
Resolution (Å)	72.56 – 2.00 72.56 – 2.00	Depositor EDS
% Data completeness (in resolution range)	57.3 (72.56-2.00) 57.3 (72.56-2.00)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.232 , 0.259 0.232 , 0.259	Depositor DCC
R_{free} test set	1185 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	23.6	Xtriage
Anisotropy	0.179	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 46.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8804	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.62 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.4930e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: H88

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	0.24	0/2234	0.41	0/3014
1	B	0.25	0/2188	0.41	0/2952
All	All	0.25	0/4422	0.41	0/5966

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	482	GLU	Peptide

5.2 Too-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 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	2176	2142	2142	8	0
1	B	2132	2102	2102	10	0
2	A	38	25	0	0	0
2	B	38	25	0	0	0
3	A	73	0	0	0	0
3	B	53	0	0	1	0
All	All	4510	4294	4244	18	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 2.

All (18) 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:468:SER:OG	1:A:474:LYS:NZ	2.14	0.80
1:B:414:ARG:NH2	1:B:484:PRO:O	2.21	0.74
1:A:414:ARG:NH2	1:A:484:PRO:O	2.34	0.61
1:B:411:GLU:N	3:B:803:HOH:O	2.37	0.57
1:A:470:PRO:O	1:A:551:LYS:NZ	2.35	0.56
1:A:421:LYS:O	1:A:433:LEU:N	2.39	0.55
1:B:536:ASP:OD1	1:B:540:ARG:NH2	2.43	0.47
1:B:515:GLU:OE2	1:B:670:ARG:NH2	2.45	0.47
1:B:514:LEU:HD11	1:B:666:ILE:HG23	1.96	0.46
1:B:514:LEU:HD21	1:B:669:LEU:HB2	1.98	0.46
1:B:457:PHE:CE1	1:B:485:ILE:HG21	2.53	0.43
1:A:492:ILE:HG22	1:A:544:VAL:O	2.19	0.43
1:B:467:LEU:HD23	1:B:527:LEU:CD2	2.48	0.43
1:A:427:GLN:NE2	1:A:428:PHE:CE2	2.88	0.41
1:A:494:ASN:HB2	1:A:544:VAL:O	2.20	0.41
1:B:584:PHE:CE1	1:B:622:VAL:HG13	2.55	0.41
1:A:464:MET:SD	1:A:558:THR:HG21	2.60	0.41
1:B:483:TYR:HB2	1:B:484:PRO:HD3	2.03	0.40

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	264/278 (95%)	260 (98%)	4 (2%)	0	100	100
1	B	259/278 (93%)	254 (98%)	5 (2%)	0	100	100
All	All	523/556 (94%)	514 (98%)	9 (2%)	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	241/253 (95%)	240 (100%)	1 (0%)	91	93
1	B	237/253 (94%)	236 (100%)	1 (0%)	91	93
All	All	478/506 (94%)	476 (100%)	2 (0%)	91	93

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	421	LYS
1	B	483	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	427	GLN

5.3.3 RNA

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)

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	H88	B	701	1	41,42,43	1.55	8 (19%)	56,62,63	1.89	9 (16%)
2	H88	A	701	1	41,42,43	1.54	8 (19%)	56,62,63	1.86	9 (16%)

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	H88	B	701	1	-	2/19/19/20	0/5/5/5
2	H88	A	701	1	-	4/19/19/20	0/5/5/5

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	H88	C13-N16	-4.51	1.35	1.43
2	A	701	H88	C13-N16	-4.37	1.36	1.43
2	B	701	H88	C25-S17	-3.71	1.66	1.75
2	A	701	H88	C25-S17	-3.69	1.66	1.75
2	A	701	H88	C24-N19	3.69	1.44	1.38
2	B	701	H88	C24-N19	3.58	1.44	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	H88	C18-N19	-3.35	1.41	1.45
2	A	701	H88	C6-N5	-3.29	1.35	1.41
2	A	701	H88	C18-N19	-3.29	1.41	1.45
2	B	701	H88	C6-N5	-3.25	1.35	1.41
2	A	701	H88	C29-N28	2.31	1.33	1.30
2	B	701	H88	C29-N28	2.31	1.33	1.30
2	B	701	H88	S17-N16	-2.28	1.60	1.63
2	B	701	H88	C20-N19	2.16	1.41	1.37
2	A	701	H88	S17-N16	-2.16	1.60	1.63
2	A	701	H88	C20-N19	2.05	1.41	1.37

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	701	H88	O19-S17-O18	-6.79	109.09	118.85
2	A	701	H88	O19-S17-O18	-6.60	109.36	118.85
2	B	701	H88	C18-N19-C24	6.20	125.77	118.76
2	A	701	H88	C18-N19-C24	6.01	125.56	118.76
2	B	701	H88	C20-N19-C18	-4.69	113.77	118.69
2	A	701	H88	C20-N19-C18	-4.45	114.02	118.69
2	B	701	H88	C22-C23-C29	-3.58	116.21	122.63
2	A	701	H88	C22-C23-C29	-3.57	116.23	122.63
2	A	701	H88	C23-C29-N28	-3.17	120.74	125.05
2	B	701	H88	C23-C29-N28	-3.17	120.75	125.05
2	A	701	H88	C22-C23-C24	2.78	122.46	117.94
2	B	701	H88	C22-C23-C24	2.78	122.45	117.94
2	B	701	H88	C8-C7-C6	2.77	120.05	117.44
2	A	701	H88	C8-C7-C6	2.76	120.04	117.44
2	B	701	H88	C25-S17-N16	2.69	109.70	106.63
2	A	701	H88	C25-S17-N16	2.53	109.52	106.63
2	A	701	H88	C29-N28-C27	2.49	121.12	117.69
2	B	701	H88	C29-N28-C27	2.47	121.09	117.69

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	H88	C11-C6-N5-C3
2	A	701	H88	C10-C13-N16-S17
2	A	701	H88	C7-C6-N5-C3
2	A	701	H88	C14-C13-N16-S17
2	B	701	H88	C11-C6-N5-C3

Continued on next page...

Continued from previous page...

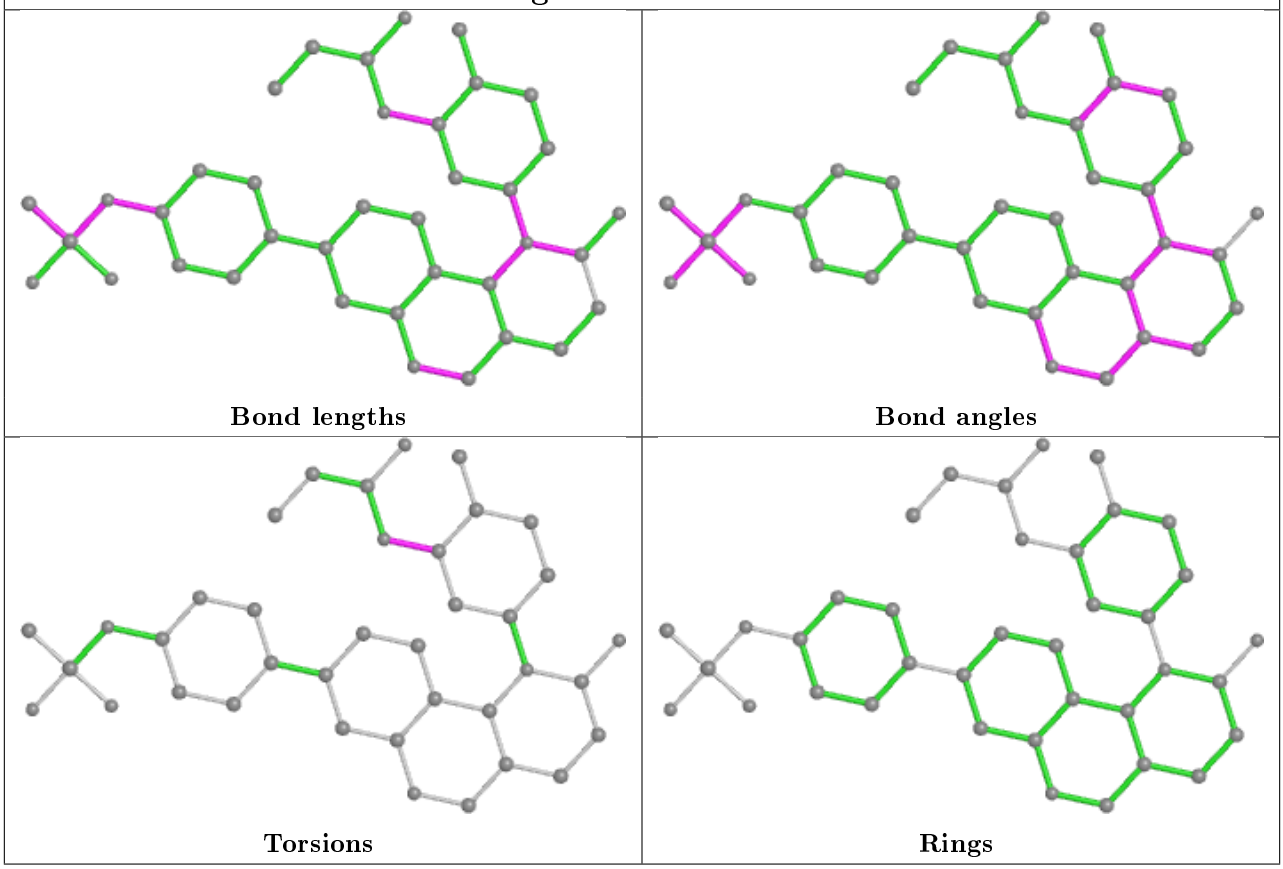
Mol	Chain	Res	Type	Atoms
2	B	701	H88	C7-C6-N5-C3

There are no ring outliers.

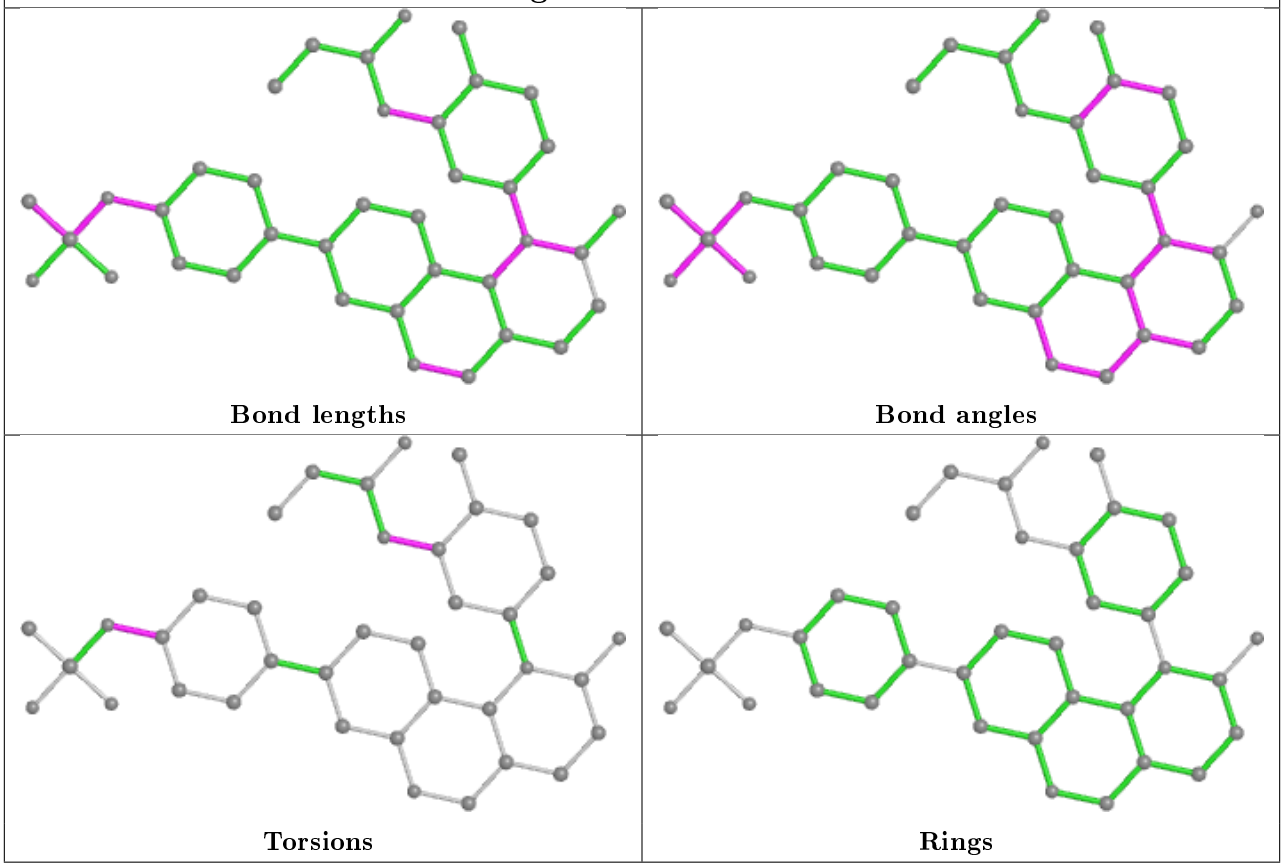
No monomer is involved in short contacts.

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.

Ligand H88 B 701



Ligand H88 A 701



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.

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, 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>2	OWAB(Å^2)	Q<0.9
1	A	266/278 (95%)	0.24	5 (1%) 66 65	15, 31, 54, 109	0
1	B	261/278 (93%)	0.32	10 (3%) 40 39	15, 31, 52, 122	0
All	All	527/556 (94%)	0.27	15 (2%) 53 51	15, 31, 54, 122	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	483	TYR	10.0
1	A	427	GLN	6.7
1	A	407	TYR	3.9
1	B	615	LEU	3.8
1	A	615	LEU	3.4
1	A	452	MET	3.1
1	B	570	VAL	2.8
1	B	465	MET	2.7
1	B	427	GLN	2.4
1	B	484	PRO	2.4
1	B	482	GLU	2.4
1	B	616	TYR	2.3
1	B	428	PHE	2.2
1	A	428	PHE	2.1
1	B	426	GLY	2.0

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 (i)

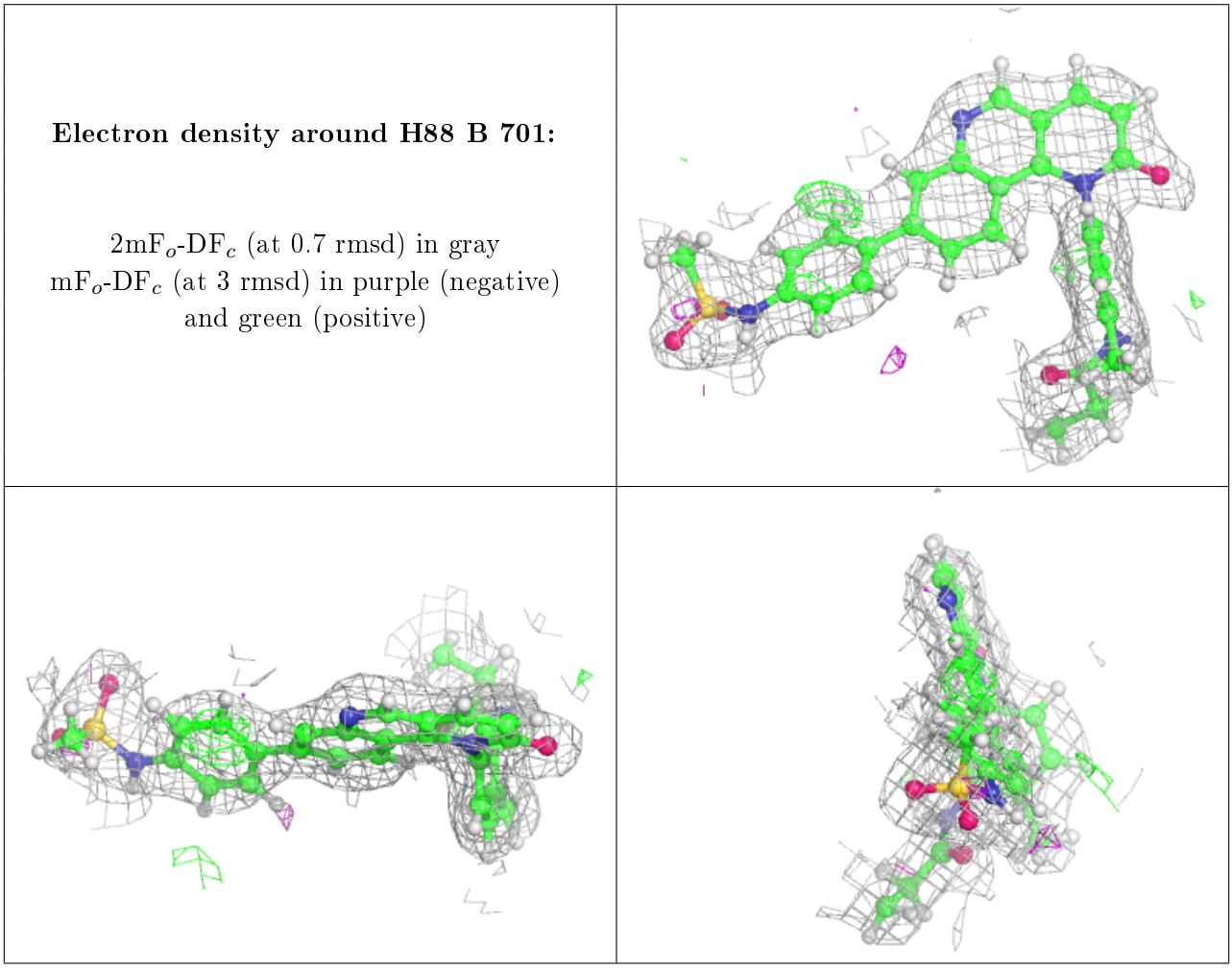
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	H88	B	701	38/39	0.92	0.13	11,22,39,44	0
2	H88	A	701	38/39	0.94	0.12	15,22,38,43	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.

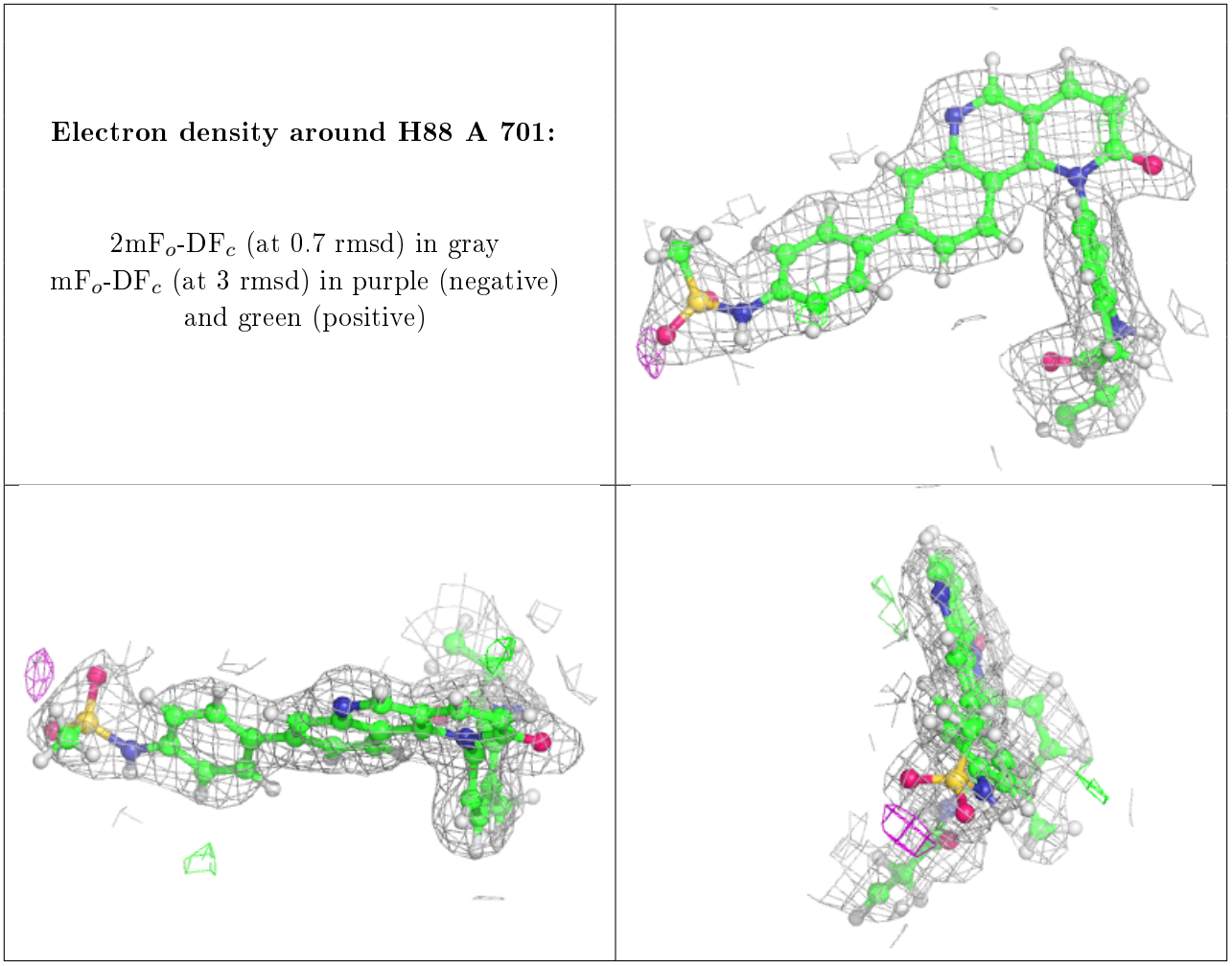
Electron density around H88 B 701:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around H88 A 701:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers

There are no such residues in this entry.