



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 31, 2019 – 07:57 PM EDT

PDB ID : 6K5A  
Title : Crystal structure of the E148D/R147A/F317A mutant in presence of 200 mM NaBr  
Authors : Park, K.; Lim, H.H.  
Deposited on : 2019-05-28  
Resolution : 3.16 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](mailto: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  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

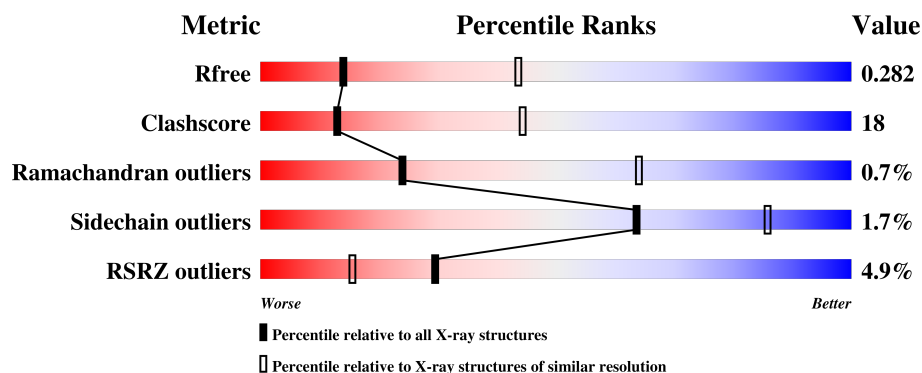
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.16 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))
$R_{free}$	111664	1389 (3.20-3.12)
Clashscore	122126	1522 (3.20-3.12)
Ramachandran outliers	120053	1493 (3.20-3.12)
Sidechain outliers	120020	1492 (3.20-3.12)
RSRZ outliers	108989	1344 (3.20-3.12)

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  $\geq 3$ , 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	473	<div> <div>2%</div> <div> <div></div> <div>60%</div> <div>33%</div> <div>7%</div> </div> </div>
1	B	473	<div> <div>5%</div> <div> <div></div> <div>55%</div> <div>36%</div> <div>7%</div> </div> </div>
2	C	222	<div> <div>5%</div> <div> <div></div> <div>74%</div> <div>24%</div> <div>.</div> </div> </div>
2	E	222	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>30%</div> <div>.</div> </div> </div>
3	D	211	<div> <div>9%</div> <div> <div></div> <div>61%</div> <div>38%</div> <div>.</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	211	

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
4	BR	A	501	-	-	X	-
4	BR	B	501	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13175 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 H(+)/Cl(-) exchange transporter ClcA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	440	Total	C	N	O	S	0	0	0
			3288	2162	552	554	20			
1	B	440	Total	C	N	O	S	0	0	0
			3288	2162	552	554	20			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	ALA	ARG	engineered mutation	UNP E9TIA0
A	148	ASP	GLU	engineered mutation	UNP E9TIA0
A	317	ALA	PHE	engineered mutation	UNP E9TIA0
B	147	ALA	ARG	engineered mutation	UNP E9TIA0
B	148	ASP	GLU	engineered mutation	UNP E9TIA0
B	317	ALA	PHE	engineered mutation	UNP E9TIA0

- Molecule 2 is a protein called Fab fragment, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	222	Total	C	N	O	S	0	0	0
			1681	1082	275	318	6			
2	E	221	Total	C	N	O	S	0	0	0
			1672	1077	274	315	6			

- Molecule 3 is a protein called Fab fragment, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			
3	F	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			

- Molecule 4 is BROMIDE ION (three-letter code: BR) (formula: Br) (labeled as "Ligand of

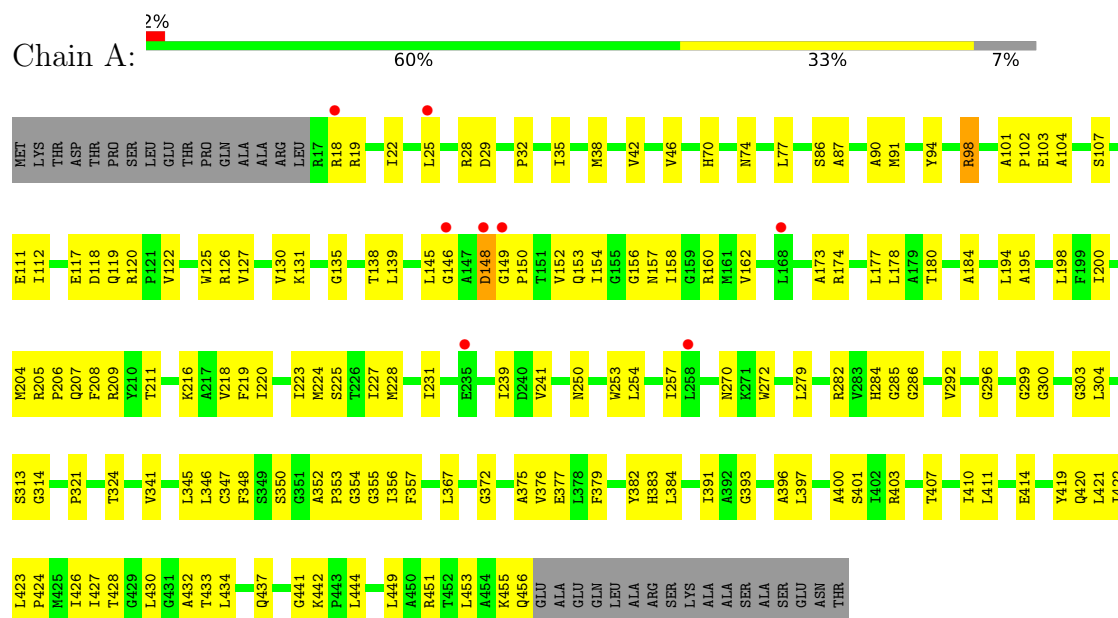
Interest" by author).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total 2	Br 2	0	0
4	A	2	Total 2	Br 2	0	0

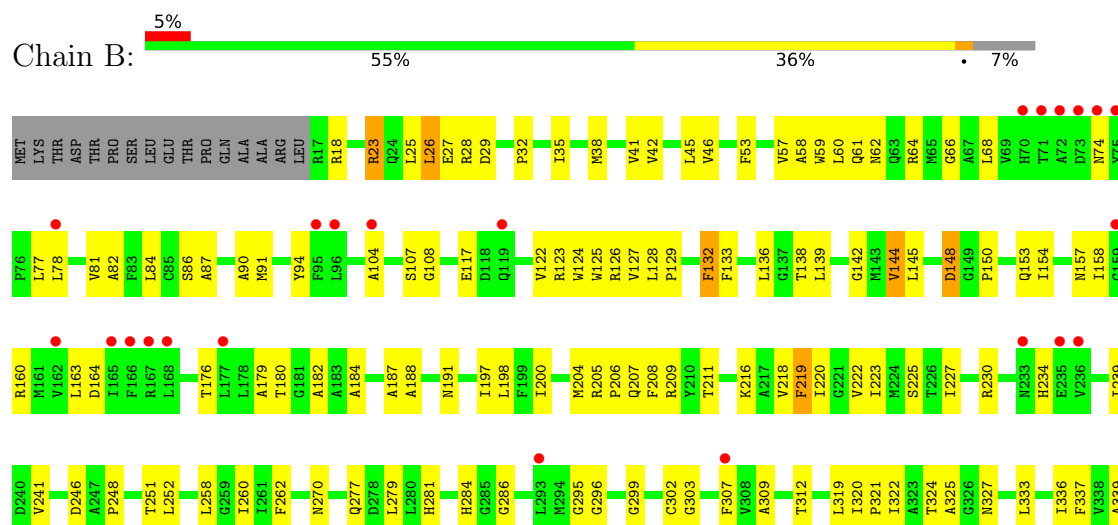
### 3 Residue-property plots

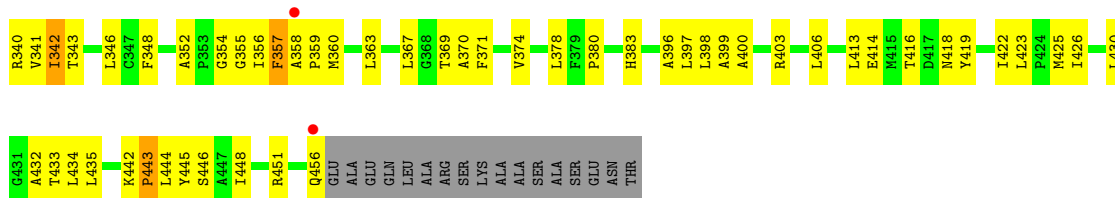
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: H(+)/Cl(-) exchange transporter ClcA

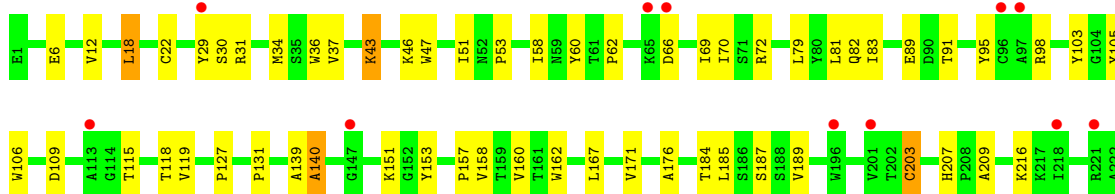
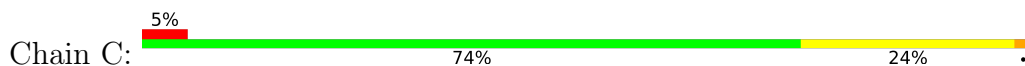


#### • Molecule 1: H(+)/Cl(-) exchange transporter ClcA

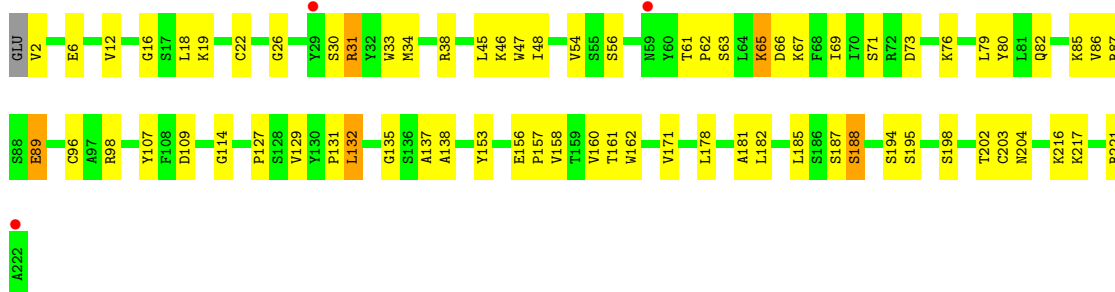




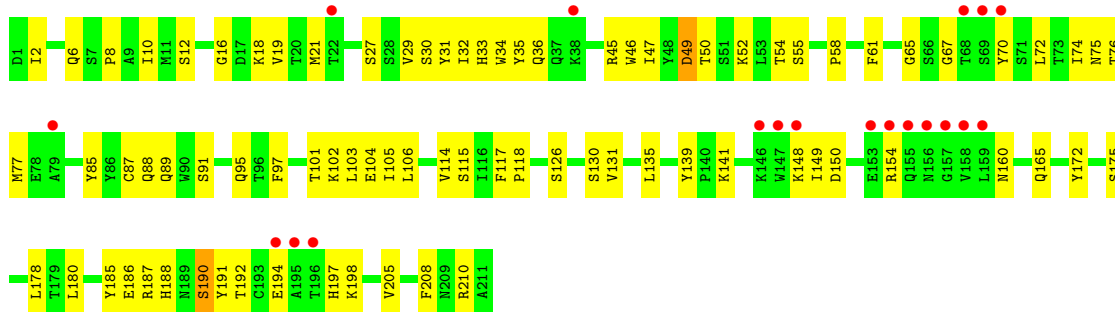
- Molecule 2: Fab fragment, heavy chain



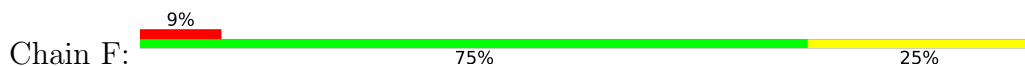
- Molecule 2: Fab fragment, heavy chain

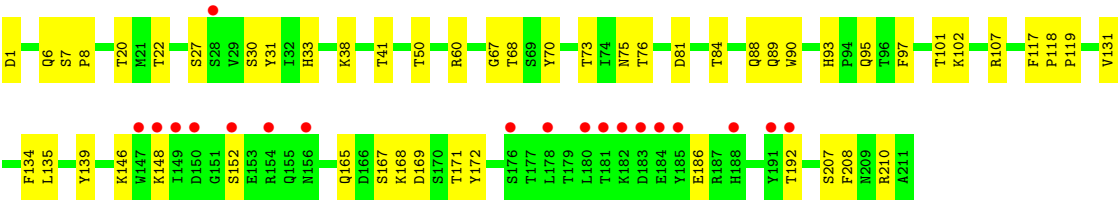


- Molecule 3: Fab fragment, light chain



- Molecule 3: Fab fragment, light chain







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	232.18Å 101.34Å 172.07Å 90.00° 132.10° 90.00°	Depositor
Resolution (Å)	39.69 – 3.16 44.79 – 3.16	Depositor EDS
% Data completeness (in resolution range)	98.9 (39.69-3.16) 98.9 (44.79-3.16)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.29 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.222 , 0.276 0.230 , 0.282	Depositor DCC
$R_{free}$ test set	2521 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	108.5	Xtriage
Anisotropy	0.534	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 64.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.010 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	13175	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.47	0/3359	0.70	3/4561 (0.1%)
1	B	0.47	0/3359	0.69	1/4561 (0.0%)
2	C	0.54	0/1730	0.79	2/2367 (0.1%)
2	E	0.57	0/1721	0.74	3/2355 (0.1%)
3	D	0.53	0/1660	0.74	0/2257
3	F	0.51	0/1660	0.73	0/2257
All	All	0.50	0/13489	0.72	9/18358 (0.0%)

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
3	D	0	1
3	F	0	1
All	All	0	2

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	140	ALA	CB-CA-C	7.25	120.97	110.10
2	E	89	GLU	OE1-CD-OE2	-5.76	116.39	123.30
2	E	65	LYS	CB-CA-C	-5.66	99.08	110.40
2	C	140	ALA	N-CA-C	-5.54	96.03	111.00
1	A	107	SER	N-CA-C	5.49	125.82	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	49	ASP	Mainchain
3	F	6	GLN	Peptide

## 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	3288	0	3445	137	0
1	B	3288	0	3445	182	0
2	C	1681	0	1663	42	0
2	E	1672	0	1654	44	0
3	D	1621	0	1546	71	0
3	F	1621	0	1546	40	0
4	A	2	0	0	4	0
4	B	2	0	0	7	0
All	All	13175	0	13299	465	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 18.

The worst 5 of 465 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:B:356:ILE:N	4:B:501:BR:BR	2.40	1.09
1:B:445:TYR:OH	4:B:501:BR:BR	2.31	1.02
1:A:200:ILE:HA	1:A:204:MET:HB2	1.49	0.91
1:A:219:PHE:HB3	1:B:430:LEU:HD21	1.53	0.90
3:D:95:GLN:N	3:D:95:GLN:OE1	2.06	0.88

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	
1	A	438/473 (93%)	398 (91%)	39 (9%)	1 (0%)	49	83
1	B	438/473 (93%)	400 (91%)	34 (8%)	4 (1%)	19	58
2	C	220/222 (99%)	194 (88%)	24 (11%)	2 (1%)	19	58
2	E	219/222 (99%)	197 (90%)	19 (9%)	3 (1%)	12	48
3	D	209/211 (99%)	191 (91%)	17 (8%)	1 (0%)	31	70
3	F	209/211 (99%)	197 (94%)	11 (5%)	1 (0%)	31	70
All	All	1733/1812 (96%)	1577 (91%)	144 (8%)	12 (1%)	24	64

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	26	LEU
1	B	325	ALA
2	E	62	PRO
1	A	285	GLY
2	C	157	PRO

### 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	330/356 (93%)	324 (98%)	6 (2%)	62	85
1	B	330/356 (93%)	324 (98%)	6 (2%)	62	85
2	C	182/182 (100%)	178 (98%)	4 (2%)	55	81
2	E	181/182 (100%)	177 (98%)	4 (2%)	55	81
3	D	185/185 (100%)	182 (98%)	3 (2%)	65	87
3	F	185/185 (100%)	184 (100%)	1 (0%)	90	96
All	All	1393/1446 (96%)	1369 (98%)	24 (2%)	63	86

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	246	ASP
2	C	43	LYS
2	E	188	SER
1	B	357	PHE
2	C	18	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	284	HIS
3	F	95	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	440/473 (93%)	0.05	8 (1%) 68 53	81, 111, 139, 157	0
1	B	440/473 (93%)	0.29	25 (5%) 24 12	90, 117, 149, 173	0
2	C	222/222 (100%)	0.12	11 (4%) 29 15	75, 104, 140, 171	0
2	E	221/222 (99%)	-0.19	3 (1%) 75 62	79, 102, 134, 166	0
3	D	211/211 (100%)	0.20	19 (9%) 9 5	92, 115, 134, 151	0
3	F	211/211 (100%)	0.41	19 (9%) 9 5	76, 96, 149, 160	0
All	All	1745/1812 (96%)	0.15	85 (4%) 29 15	75, 110, 143, 173	0

The worst 5 of 85 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	73	ASP	6.0
1	B	119	GLN	5.9
3	D	156	ASN	5.9
3	D	155	GLN	5.5
3	D	79	ALA	5.3

### 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

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<sup>th</sup> 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.

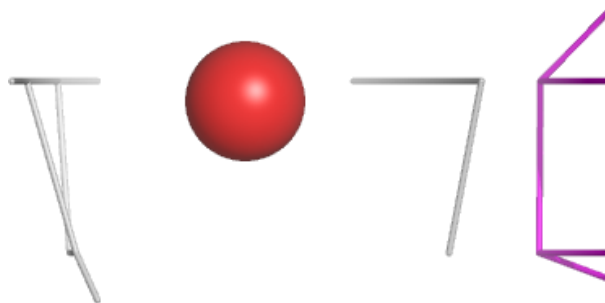
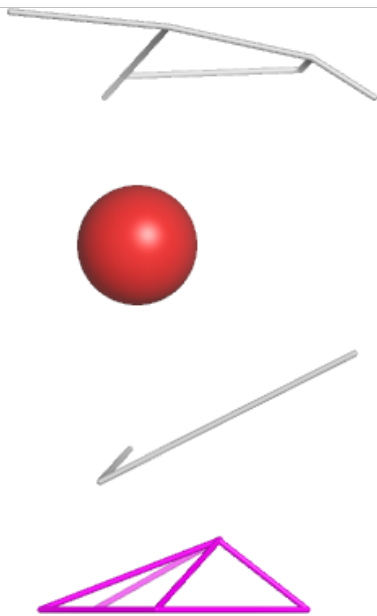
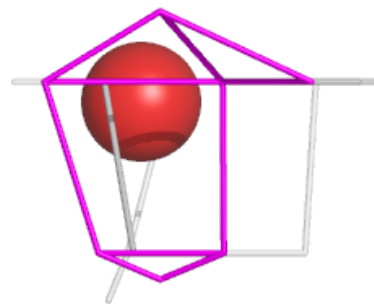
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	BR	B	501	1/1	0.82	0.15	203,203,203,203	0
4	BR	B	502	1/1	0.86	0.08	150,150,150,150	0
4	BR	A	501	1/1	0.89	0.24	196,196,196,196	0
4	BR	A	502	1/1	0.93	0.15	160,160,160,160	0

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.



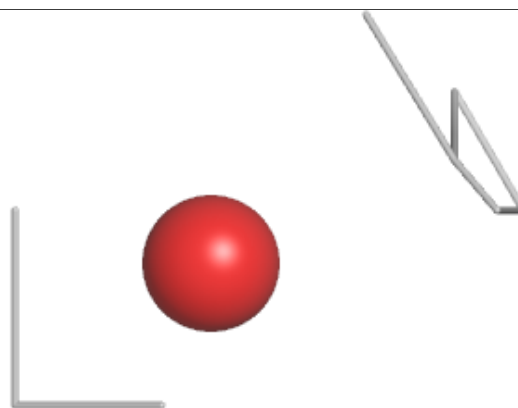
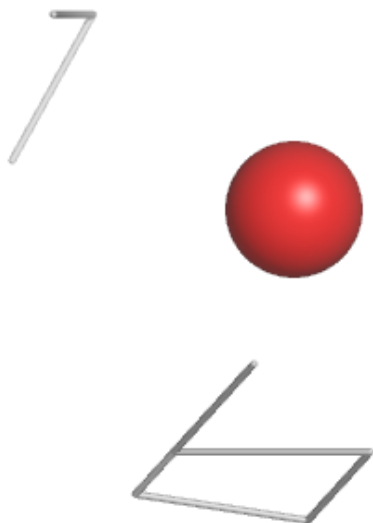
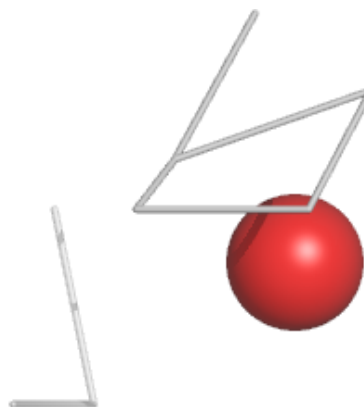
**Electron density around BR B 501:**

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



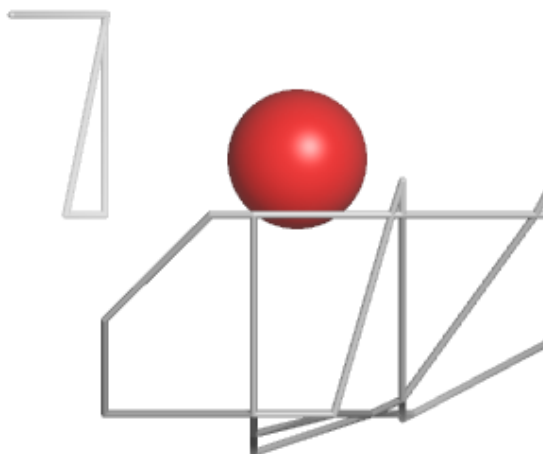
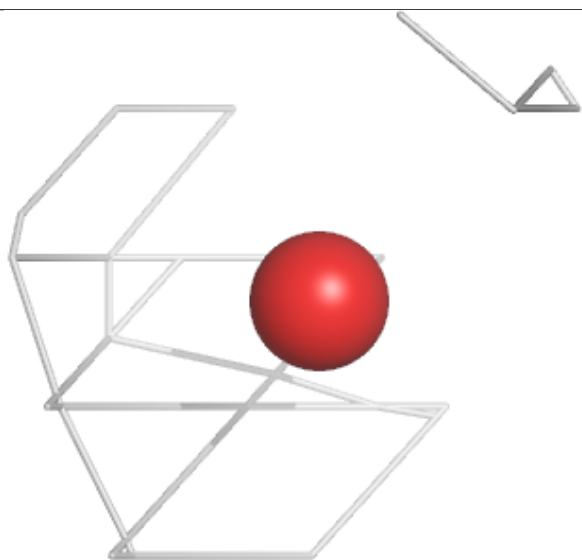
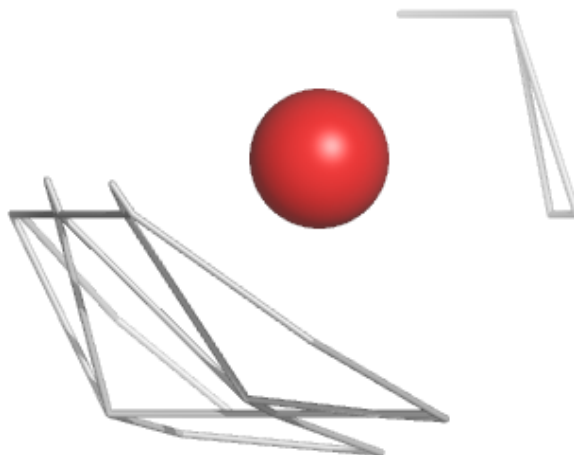
**Electron density around BR B 502:**

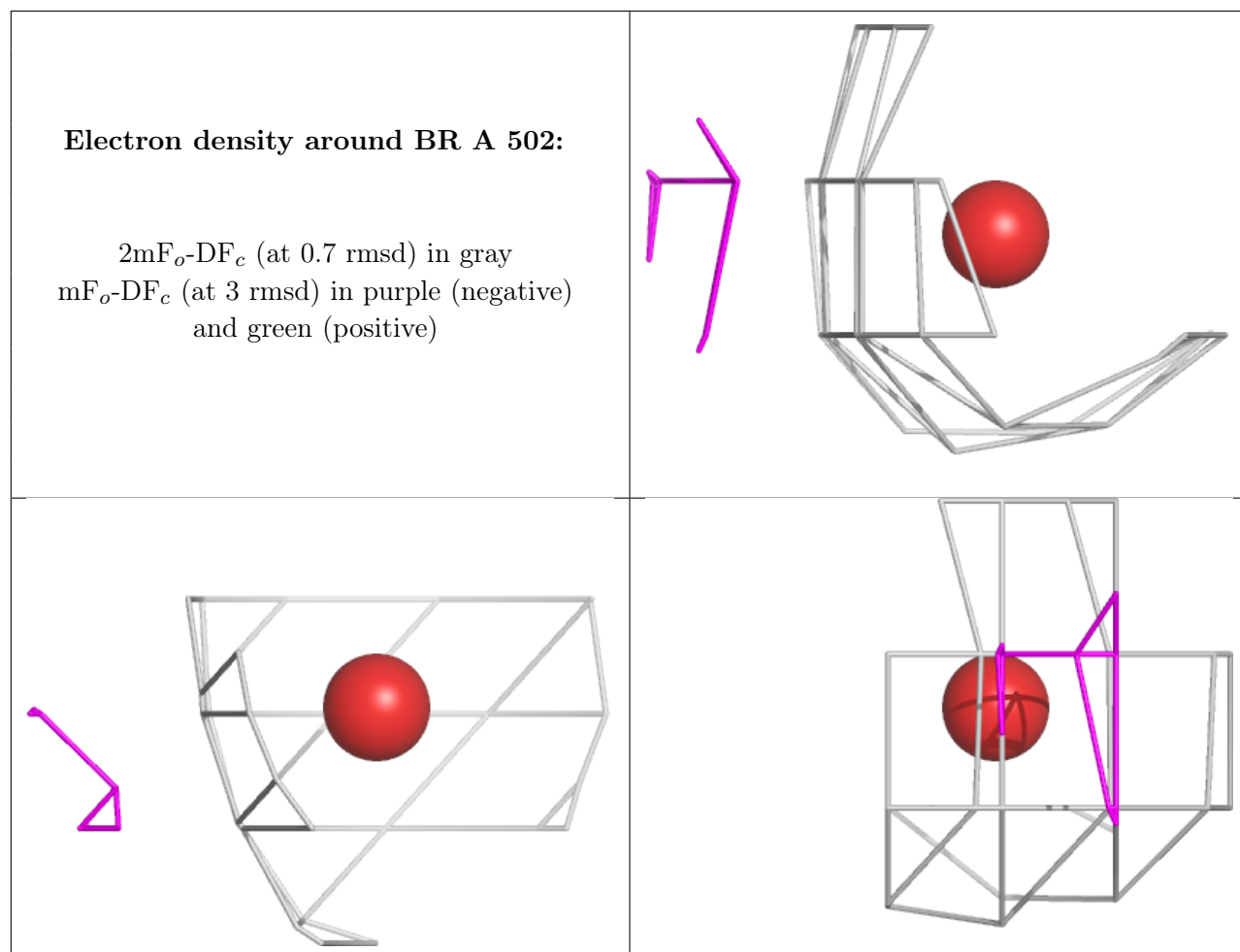
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around BR A 501:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.