



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2025 – 04:26 PM JST

PDB ID : 9J71 / pdb\_00009j71  
Title : Crystal structure of Keap1\_compound\_7  
Authors : Xu, K.  
Deposited on : 2024-08-17  
Resolution : 2.99 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0rc1  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.44

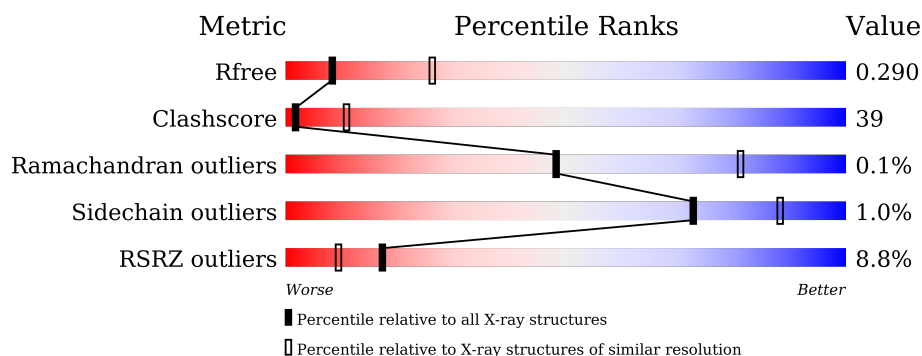
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.99 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))
$R_{free}$	164625	2511 (3.00-3.00)
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)
RSRZ outliers	164620	2523 (3.00-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	288	<div> <div>3%</div> <div> <div>50%</div> <div>48%</div> </div> <div>..</div> </div>
1	B	288	<div> <div>8%</div> <div> <div>47%</div> <div>52%</div> </div> <div>.</div> </div>
1	C	288	<div> <div>22%</div> <div> <div>37%</div> <div>60%</div> </div> <div>.</div> </div>
1	X	288	<div> <div>2%</div> <div> <div>50%</div> <div>48%</div> </div> <div>..</div> </div>

## 2 Entry composition [i](#)

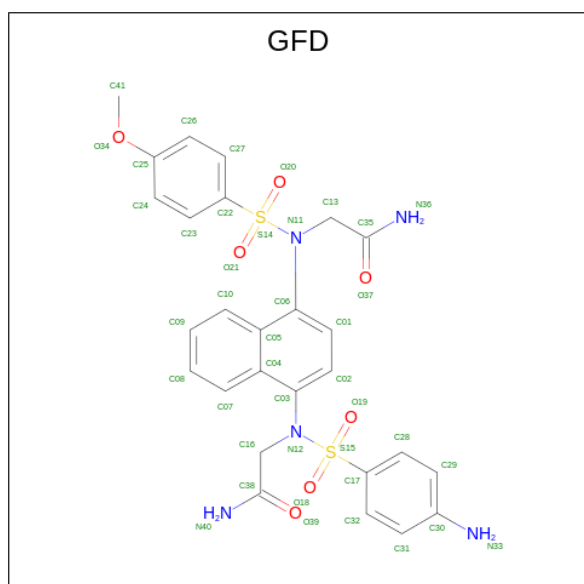
There are 3 unique types of molecules in this entry. The entry contains 9199 atoms, of which 176 are hydrogens and 0 are deuteriums.

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 Kelch-like ECH-associated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	0	0	0
			2201	1368	398	420	15			
1	B	285	Total	C	N	O	S	0	0	0
			2194	1363	397	419	15			
1	C	280	Total	C	N	O	S	0	1	0
			2166	1348	391	412	15			
1	X	286	Total	C	N	O	S	0	0	0
			2201	1368	398	420	15			

- Molecule 2 is 2-[(4-aminophenyl)sulfonyl]-[4-[(2-azanyl-2-oxidanylidene-ethyl)-(4-methoxyphenyl)sulfonyl-amino]naphthalen-1-yl]amino]ethanamide (CCD ID: GFD) (formula: C<sub>27</sub>H<sub>27</sub>N<sub>5</sub>O<sub>7</sub>S<sub>2</sub>).



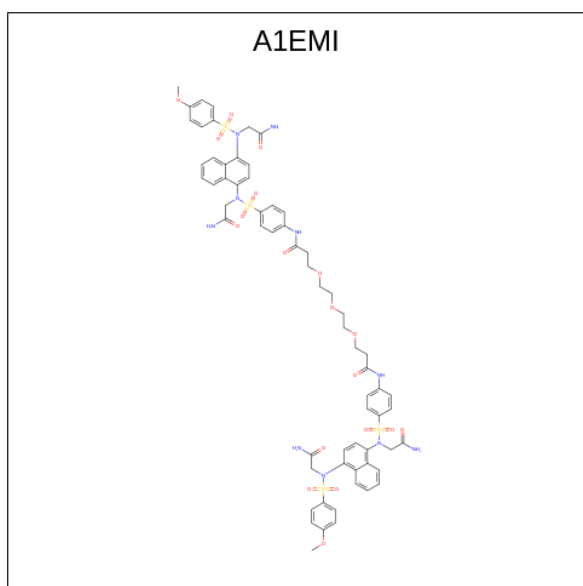
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	1
			136	54	54	10	14	4		

*Continued on next page...*

Continued from previous page...

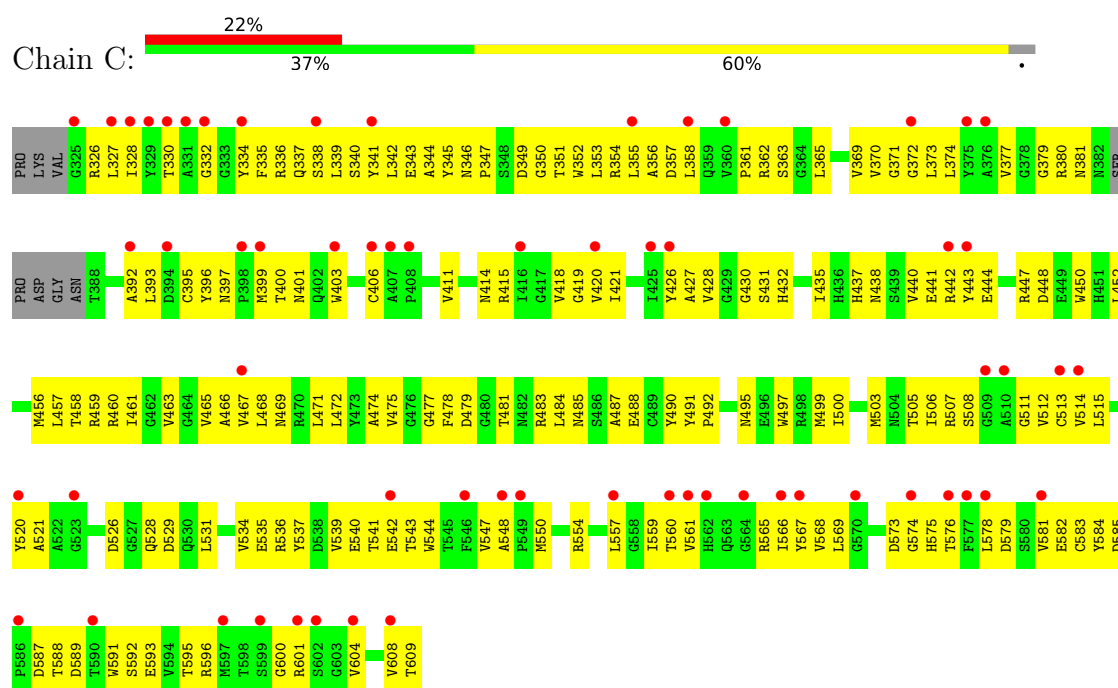
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	H	N	O	S	
			136	54	54	10	14	4	
								0	1

- Molecule 3 is {N}-[4-[(2-azanyl-2-oxidanylidene-ethyl)-[4-[(2-azanyl-2-oxidanylidene-ethyl)-(4-methoxyphenyl)sulfonyl-amino]naphthalen-1-yl]sulfamoyl]phenyl]-3-[2-[2-[3-[[4-[(2-azanyl-2-oxidanylidene-ethyl)-[4-[(2-azanyl-2-oxidanylidene-ethyl)-(4-methoxyphenyl)sulfonyl-amino]naphthalen-1-yl]sulfamoyl]phenyl]amino]-3-oxidanylidene-propoxy]ethoxy]ethoxy]prop anamide (CCD ID: A1EMI) (formula: C<sub>64</sub>H<sub>68</sub>N<sub>10</sub>O<sub>19</sub>S<sub>4</sub>) (labeled as "Ligand of Interest" by depositor).

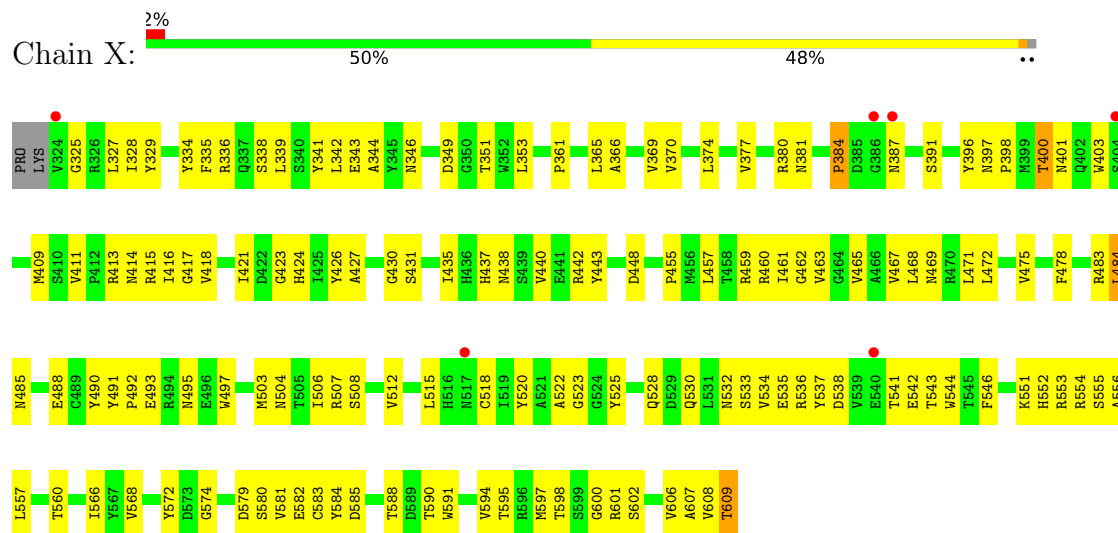


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	X	1	Total	C	H	N	O	S	
			165	64	68	10	19	4	
								0	0





• Molecule 1: Kelch-like ECH-associated protein 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.47Å 73.87Å 123.08Å 90.00° 100.55° 90.00°	Depositor
Resolution (Å)	44.45 – 2.99 44.45 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.4 (44.45-2.99) 85.2 (44.45-2.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.57 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, $R_{free}$	0.253 , 0.291 0.253 , 0.290	Depositor DCC
$R_{free}$ test set	37220 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.3	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	9199	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GFD, A1EMI

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	0.20	0/2254	0.47	0/3068
1	B	0.17	0/2247	0.43	0/3058
1	C	0.18	0/2220	0.46	0/3019
1	X	0.20	0/2254	0.44	0/3068
All	All	0.19	0/8975	0.45	0/12213

There are no bond length outliers.

There are no bond angle outliers.

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	A	2201	0	2091	169	0
1	B	2194	0	2083	164	1
1	C	2166	0	2058	207	1
1	X	2201	0	2091	153	0
2	A	82	54	0	11	0
2	C	82	54	0	4	0
3	X	97	68	0	12	0
All	All	9023	176	8323	683	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 39.

The worst 5 of 683 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:X:555:SER:OG	3:X:701:A1EMI:O18	1.54	1.21
1:C:585:ASP:OD2	1:C:588:THR:HG23	1.46	1.12
1:C:515:LEU:HD13	1:C:561:VAL:HG21	1.33	1.07
1:A:579:ASP:HB3	1:A:600:GLY:HA2	1.35	1.07
1:A:503:MET:HE1	1:A:507:ARG:HD2	1.35	1.06

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	Interatomic distance (Å)	Clash overlap (Å)
1:B:404:SER:OG	1:C:540:GLU:OE2[1_455]	1.75	0.45

## 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	284/288 (99%)	273 (96%)	11 (4%)	0	100	100
1	B	283/288 (98%)	273 (96%)	10 (4%)	0	100	100
1	C	277/288 (96%)	269 (97%)	8 (3%)	0	100	100
1	X	284/288 (99%)	272 (96%)	11 (4%)	1 (0%)	30	66
All	All	1128/1152 (98%)	1087 (96%)	40 (4%)	1 (0%)	48	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	384	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	232/234 (99%)	226 (97%)	6 (3%)	41	72
1	B	231/234 (99%)	231 (100%)	0	100	100
1	C	228/234 (97%)	228 (100%)	0	100	100
1	X	232/234 (99%)	229 (99%)	3 (1%)	65	85
All	All	923/936 (99%)	914 (99%)	9 (1%)	73	88

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	484	LEU
1	X	609	THR
1	A	568	VAL
1	A	604	VAL
1	A	609	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	381	ASN
1	C	517	ASN
1	X	575	HIS
1	X	517	ASN
1	X	552	HIS

### 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 [i](#)

There are no oligosaccharides 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GFD	A	701[A]	-	44,44,44	1.06	2 (4%)	59,65,65	1.27	6 (10%)
2	GFD	C	701[B]	-	44,44,44	1.23	4 (9%)	59,65,65	1.25	8 (13%)
2	GFD	C	701[A]	-	44,44,44	1.13	3 (6%)	59,65,65	1.41	11 (18%)
2	GFD	A	701[B]	-	44,44,44	1.12	2 (4%)	59,65,65	1.25	6 (10%)
3	A1EMI	X	701	-	104,104,104	2.85	46 (44%)	135,149,149	2.85	46 (34%)

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	GFD	A	701[A]	-	-	13/42/42/42	0/4/4/4
2	GFD	C	701[B]	-	-	15/42/42/42	0/4/4/4
2	GFD	C	701[A]	-	-	9/42/42/42	0/4/4/4
2	GFD	A	701[B]	-	-	11/42/42/42	0/4/4/4
3	A1EMI	X	701	-	-	30/106/106/106	0/8/8/8

The worst 5 of 57 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	X	701	A1EMI	S14-N11	8.91	1.77	1.65
3	X	701	A1EMI	C16-S13	7.87	1.87	1.76
3	X	701	A1EMI	C61-S64	7.54	1.87	1.76

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	X	701	A1EMI	S64-N65	7.03	1.75	1.65
3	X	701	A1EMI	C68-N65	6.90	1.50	1.44

The worst 5 of 77 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	X	701	A1EMI	O67-S64-O66	-11.46	100.95	119.52
3	X	701	A1EMI	O82-S79-O81	-11.37	101.09	119.52
3	X	701	A1EMI	O19-S14-O20	-10.62	102.31	119.52
3	X	701	A1EMI	C50-C51-N52	7.19	127.21	115.86
3	X	701	A1EMI	O17-S13-O18	-6.75	108.58	119.52

There are no chirality outliers.

5 of 78 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701[B]	GFD	C38-C16-N12-S15
2	C	701[A]	GFD	C03-N12-S15-C17
2	C	701[A]	GFD	C35-C13-N11-S14
2	C	701[B]	GFD	C38-C16-N12-C03
2	C	701[B]	GFD	C38-C16-N12-S15

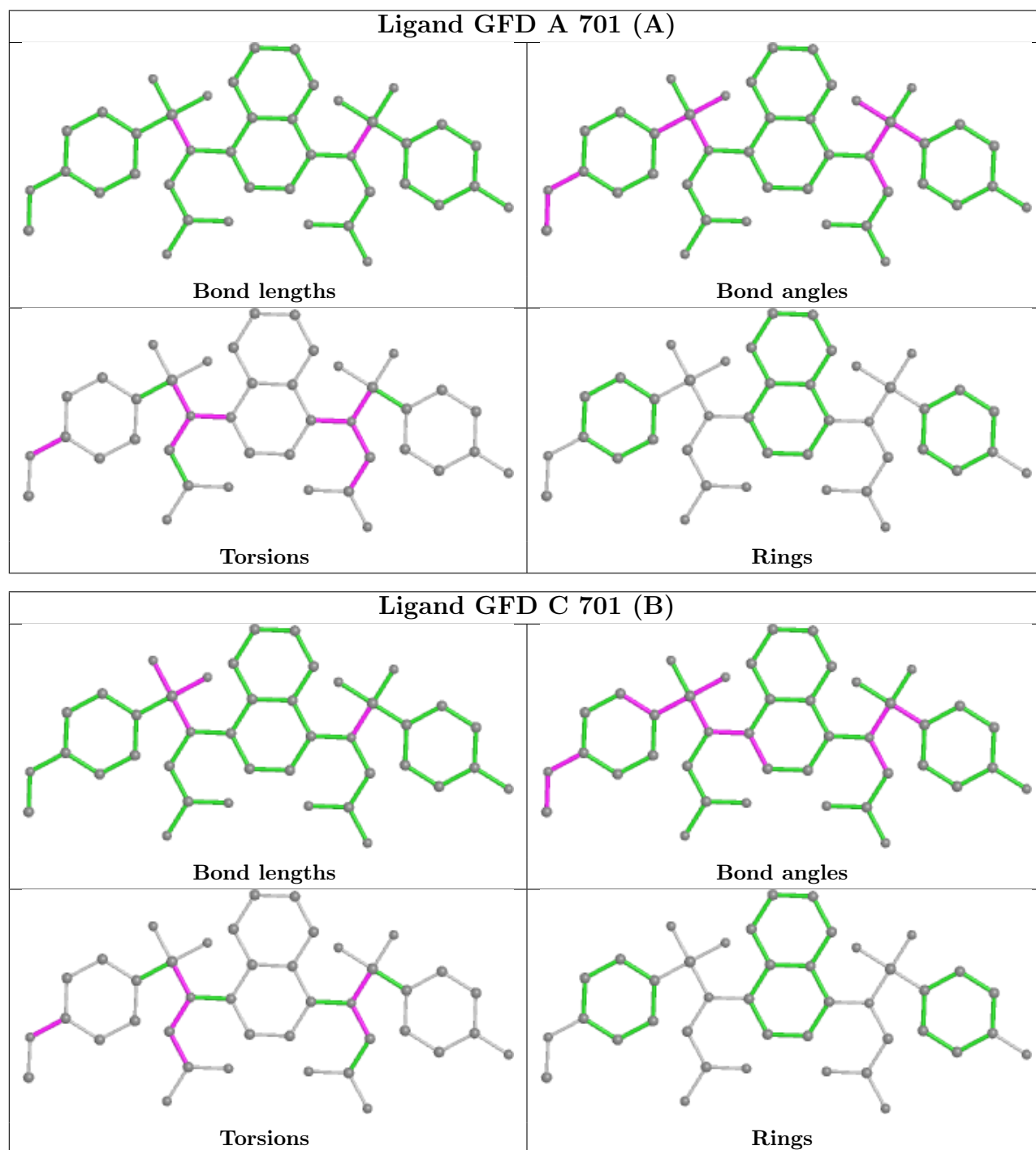
There are no ring outliers.

5 monomers are involved in 27 short contacts:

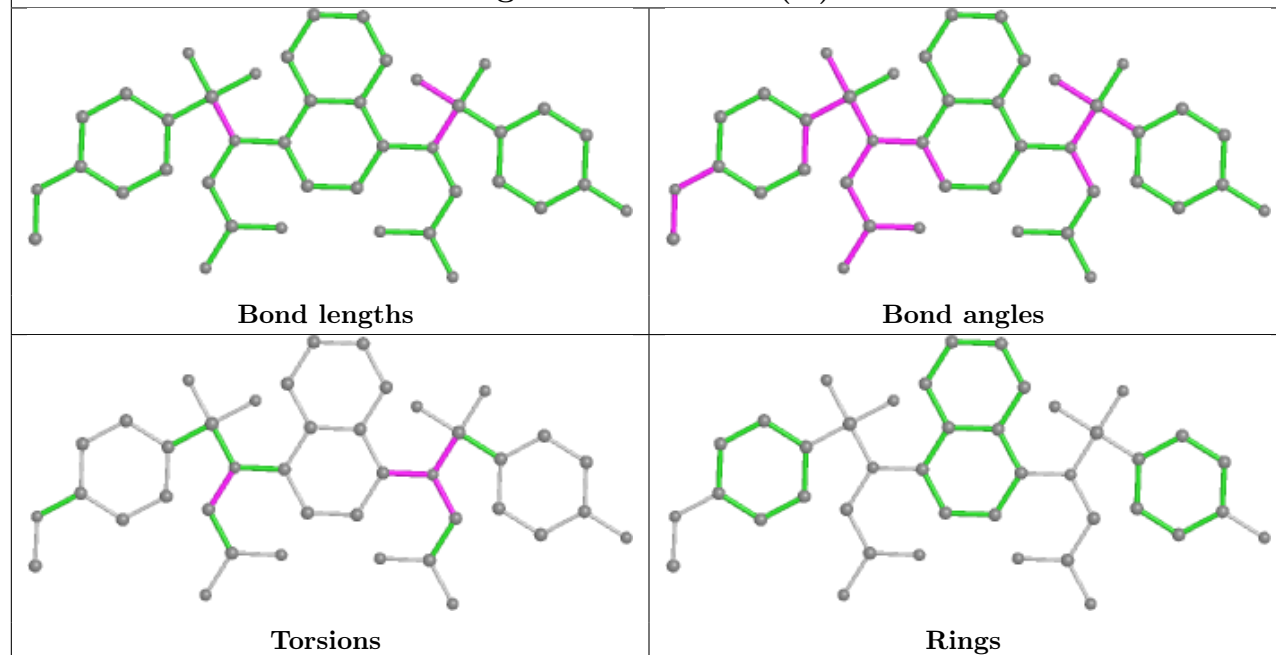
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701[A]	GFD	8	0
2	C	701[B]	GFD	2	0
2	C	701[A]	GFD	2	0
2	A	701[B]	GFD	3	0
3	X	701	A1EMI	12	0

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 than 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.

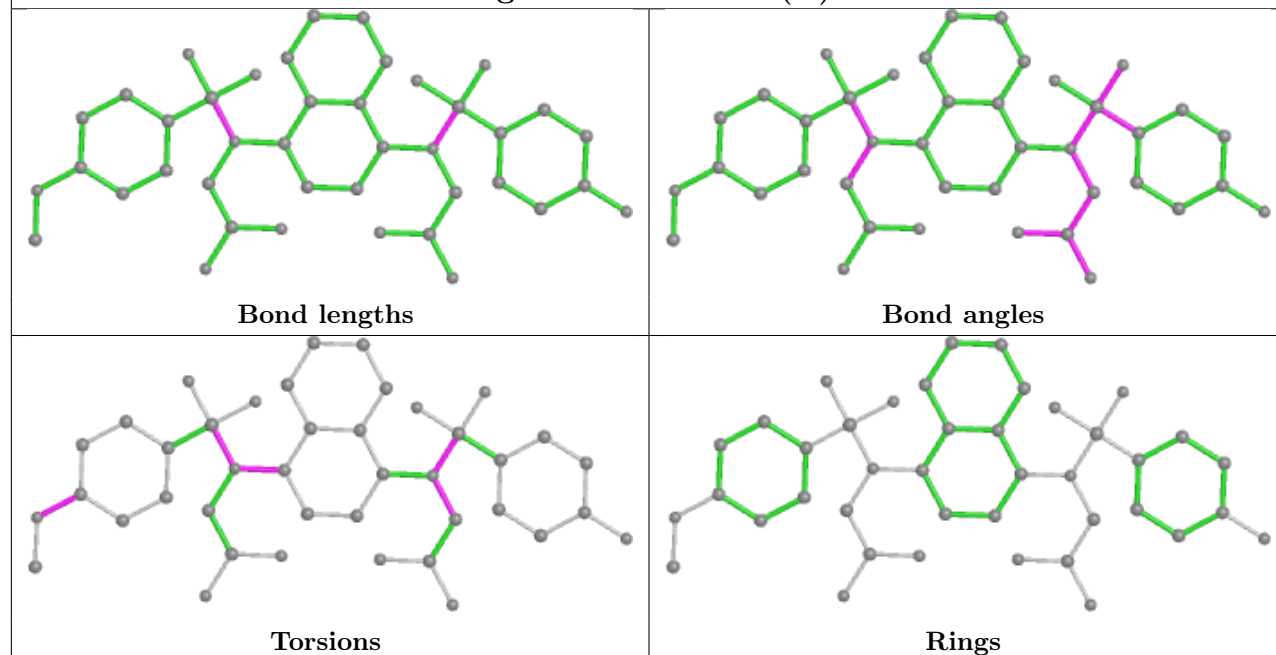
The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

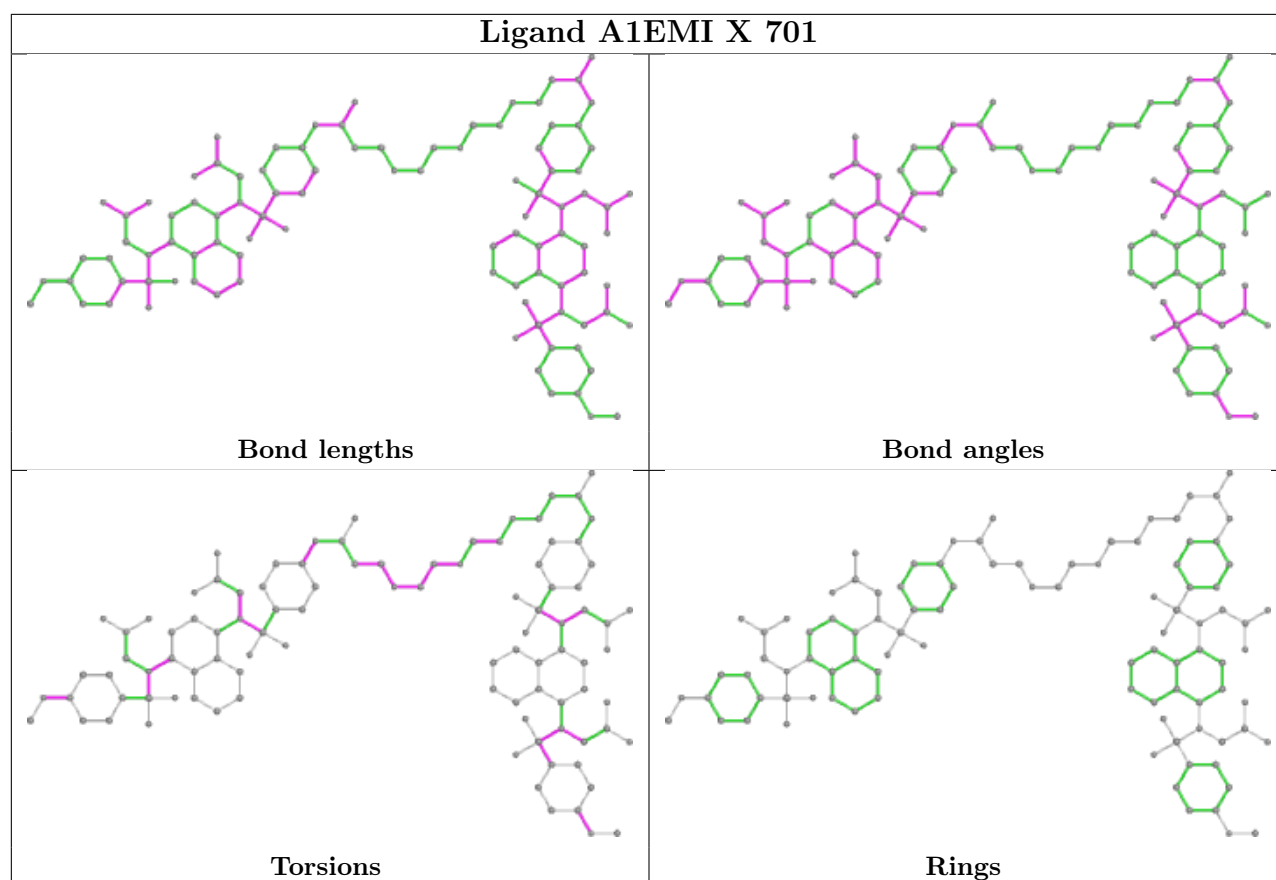


## Ligand GFD C 701 (A)



## Ligand GFD A 701 (B)





## 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	286/288 (99%)	0.56	8 (2%) 55 33	50, 68, 88, 108	0
1	B	285/288 (98%)	0.88	24 (8%) 18 10	51, 79, 101, 110	0
1	C	280/288 (97%)	1.37	62 (22%) 3 2	72, 112, 135, 148	1 (0%)
1	X	286/288 (99%)	0.53	6 (2%) 63 41	45, 63, 85, 129	0
All	All	1137/1152 (98%)	0.83	100 (8%) 17 9	45, 75, 126, 148	1 (0%)

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	332	GLY	4.2
1	B	380	ARG	3.7
1	C	394	ASP	3.6
1	X	540	GLU	3.6
1	C	329	TYR	3.5

### 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 oligosaccharides 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<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.

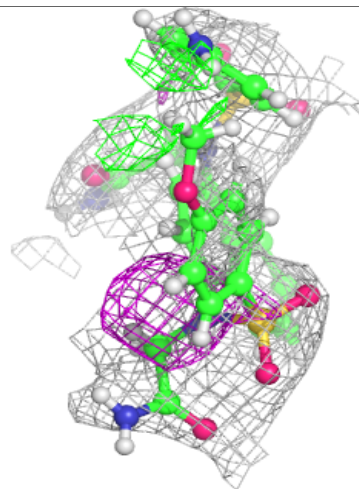
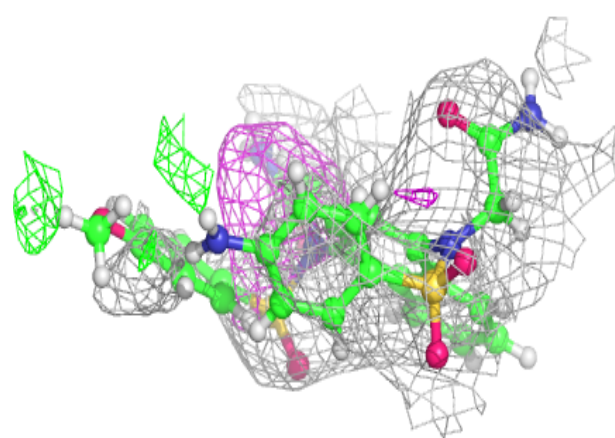
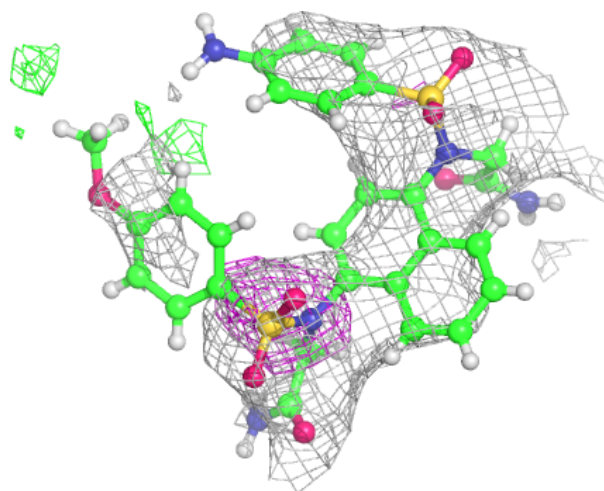


Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GFD	C	701[A]	41/41	0.74	0.21	114,128,154,156	68
2	GFD	C	701[B]	41/41	0.74	0.21	113,127,153,154	68
2	GFD	A	701[A]	41/41	0.82	0.23	78,86,105,110	68
2	GFD	A	701[B]	41/41	0.82	0.23	77,88,107,110	68
3	A1EMI	X	701	97/97	0.82	0.19	69,95,118,133	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.

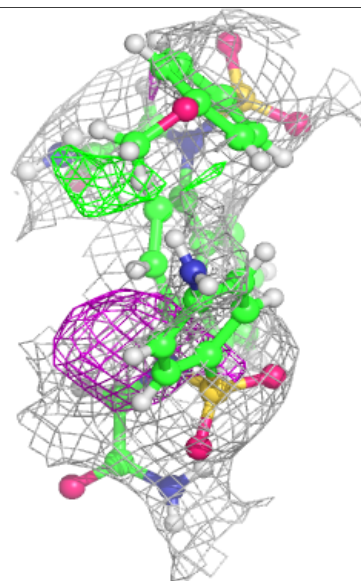
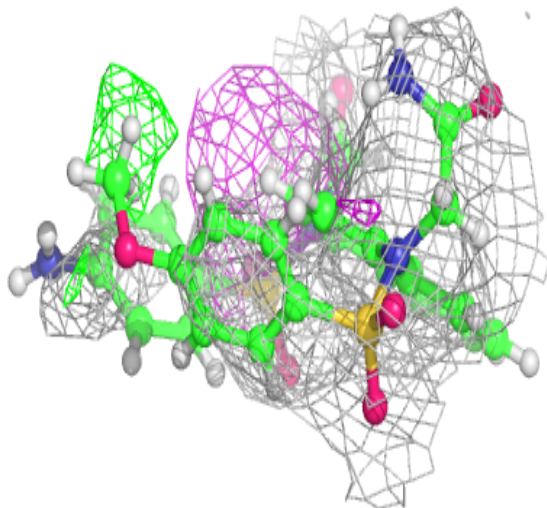
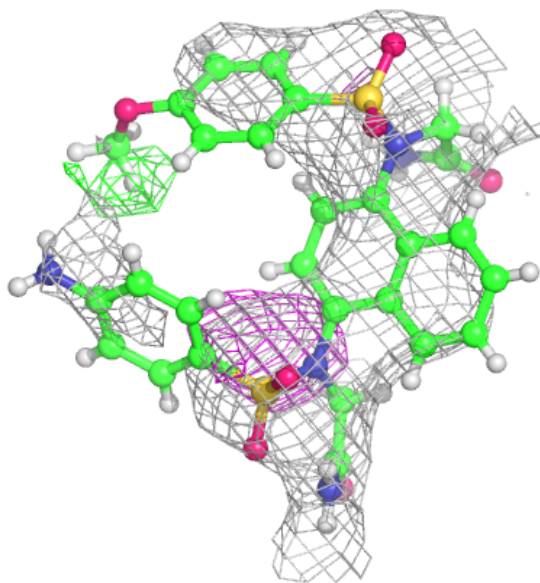
**Electron density around GFD C 701 (A):**

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)



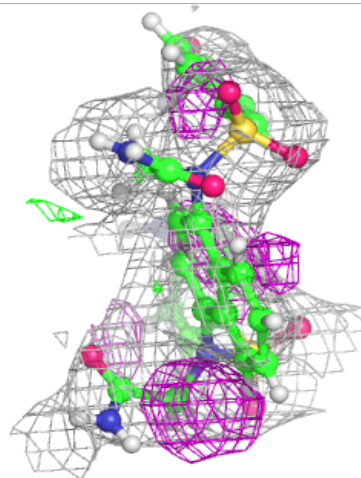
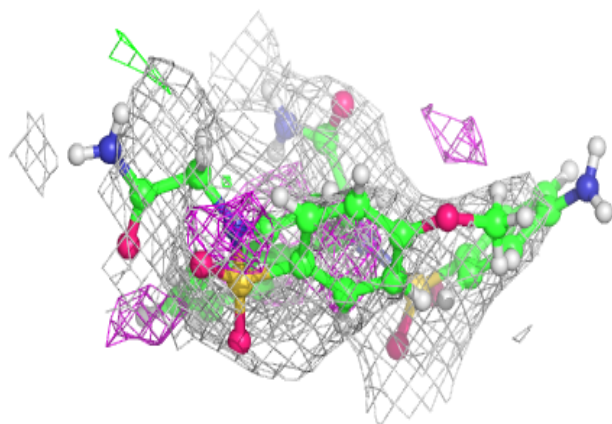
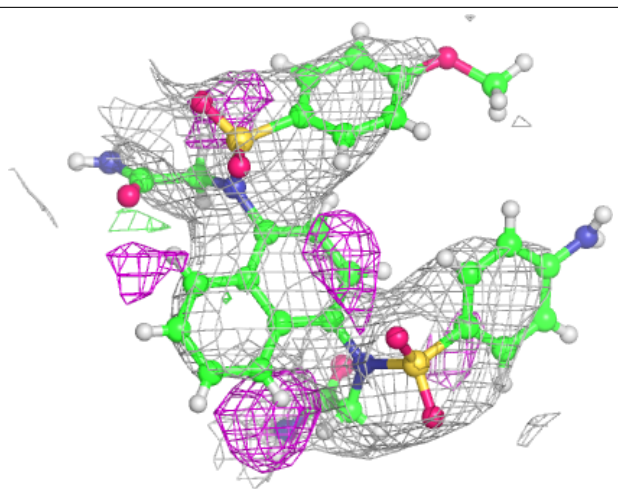
**Electron density around GFD C 701 (B):**

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



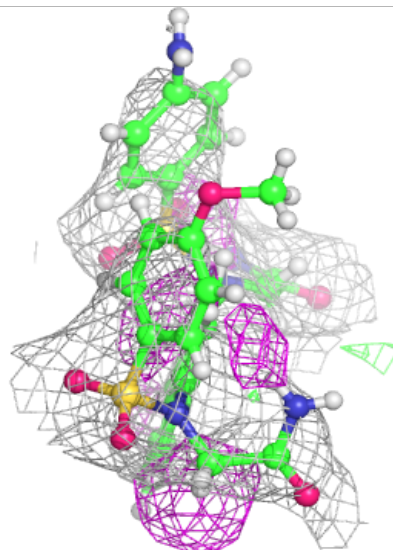
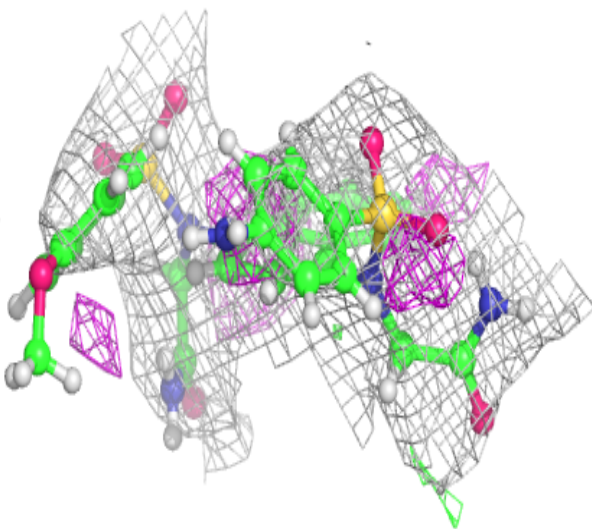
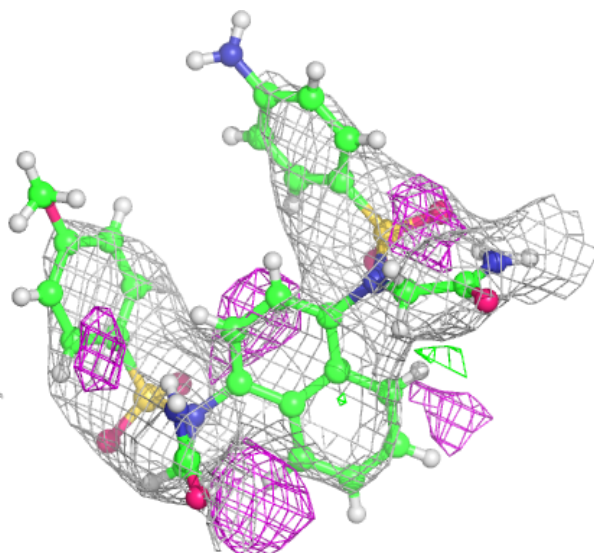
**Electron density around GFD A 701 (A):**

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



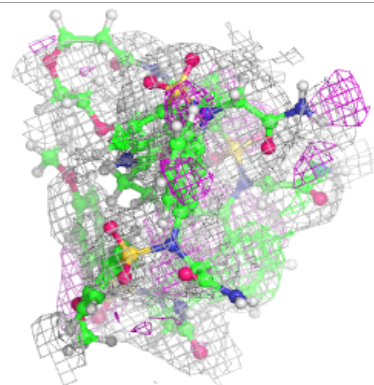
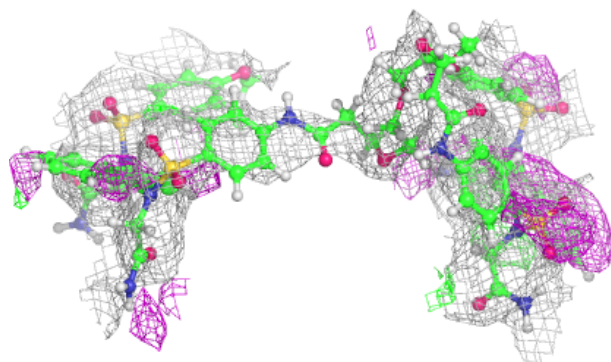
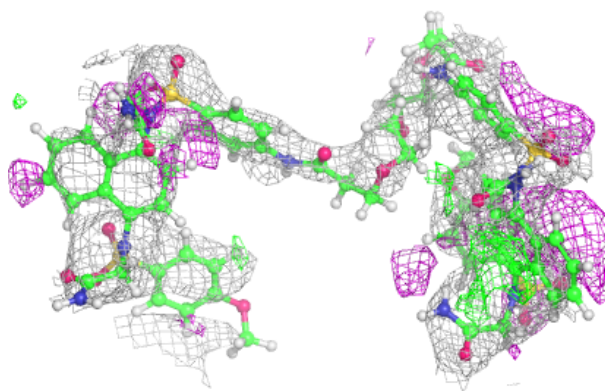
**Electron density around GFD A 701 (B):**

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



**Electron density around A1EMI X 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 [i](#)

There are no such residues in this entry.