



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 4, 2026 – 10:22 PM UTC

PDB ID : 9QHL / pdb\_00009qhl  
Title : ExsF in complex with BclA N-ter peptide  
Authors : Sogues, A.; Sleutel, M.; Remaut, H.  
Deposited on : 2025-03-16  
Resolution : 1.93 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](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:

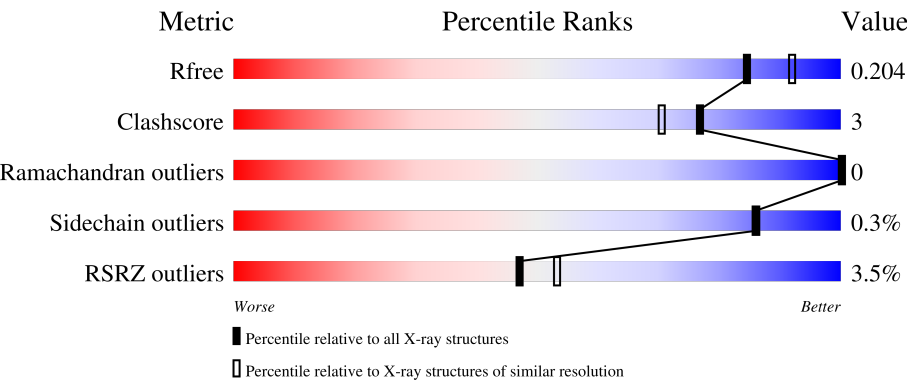
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

# 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 1.93 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.



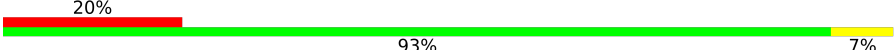
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R <sub>free</sub>	180053	1452 (1.94-1.94)
Clashscore	190562	1494 (1.94-1.94)
Ramachandran outliers	187476	1479 (1.94-1.94)
Sidechain outliers	187428	1479 (1.94-1.94)
RSRZ outliers	180081	1453 (1.94-1.94)

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	155	<div><div>2%</div><div><div></div><div>89%</div><div>6%</div><div>5%</div></div></div>
1	E	155	<div><div>3%</div><div><div></div><div>86%</div><div>10%</div><div>5%</div></div></div>
1	G	155	<div><div>2%</div><div><div></div><div>86%</div><div>8%</div><div>5%</div></div></div>
2	D	15	<div><div>13%</div><div><div></div><div>93%</div><div>7%</div></div></div>
2	F	15	<div><div>13%</div><div><div></div><div>100%</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	15	 A horizontal bar chart showing the quality of chain H. The bar is divided into three segments: a red segment on the left labeled '20%', a green segment in the middle labeled '93%', and a yellow segment on the right labeled '7%'. The segments are stacked horizontally to represent the total quality score.

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4082 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 Exosporium protein ExsF.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	148	Total	C	N	O	0	0	0
			1065	688	177	200			
1	E	148	Total	C	N	O	0	0	0
			1074	693	177	204			
1	G	147	Total	C	N	O	0	1	0
			1073	694	177	202			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	MET	-	initiating methionine	UNP A0AAW8C509
A	14	HIS	-	expression tag	UNP A0AAW8C509
A	15	HIS	-	expression tag	UNP A0AAW8C509
A	16	HIS	-	expression tag	UNP A0AAW8C509
A	17	HIS	-	expression tag	UNP A0AAW8C509
A	18	HIS	-	expression tag	UNP A0AAW8C509
A	19	HIS	-	expression tag	UNP A0AAW8C509
E	13	MET	-	initiating methionine	UNP A0AAW8C509
E	14	HIS	-	expression tag	UNP A0AAW8C509
E	15	HIS	-	expression tag	UNP A0AAW8C509
E	16	HIS	-	expression tag	UNP A0AAW8C509
E	17	HIS	-	expression tag	UNP A0AAW8C509
E	18	HIS	-	expression tag	UNP A0AAW8C509
E	19	HIS	-	expression tag	UNP A0AAW8C509
G	13	MET	-	initiating methionine	UNP A0AAW8C509
G	14	HIS	-	expression tag	UNP A0AAW8C509
G	15	HIS	-	expression tag	UNP A0AAW8C509
G	16	HIS	-	expression tag	UNP A0AAW8C509
G	17	HIS	-	expression tag	UNP A0AAW8C509
G	18	HIS	-	expression tag	UNP A0AAW8C509
G	19	HIS	-	expression tag	UNP A0AAW8C509

- Molecule 2 is a protein called Exosporium leader peptide-containing protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	15	Total	C	N	O	0	0	0
			111	76	16	19			
2	F	15	Total	C	N	O	0	0	0
			111	76	16	19			
2	H	15	Total	C	N	O	0	0	0
			111	76	16	19			

- Molecule 3 is SODIUM ION (CCD ID: NA) (formula: Na) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	2	Total	Na	0	0
			2	2		

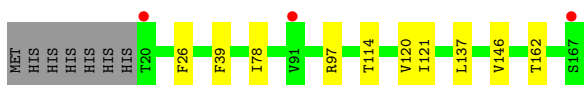
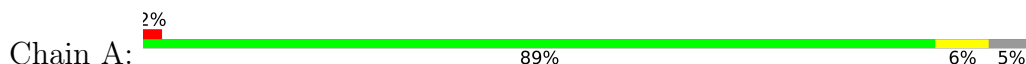
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	153	Total	O	0	0
			153	153		
4	D	18	Total	O	0	0
			18	18		
4	E	162	Total	O	0	0
			162	162		
4	F	12	Total	O	0	0
			12	12		
4	G	165	Total	O	0	0
			165	165		
4	H	25	Total	O	0	0
			25	25		

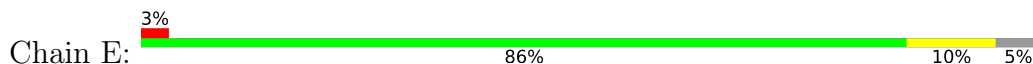
### 3 Residue-property plots [i](#)

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.

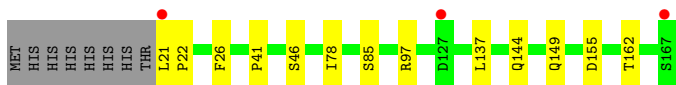
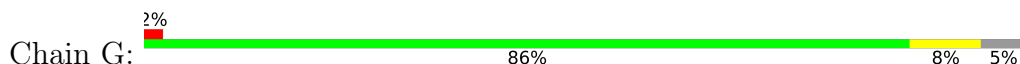
- Molecule 1: Exosporium protein ExsF



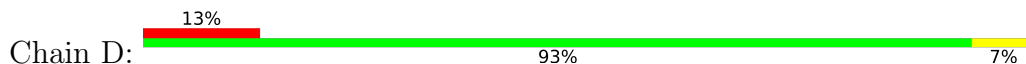
- Molecule 1: Exosporium protein ExsF



- Molecule 1: Exosporium protein ExsF



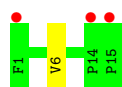
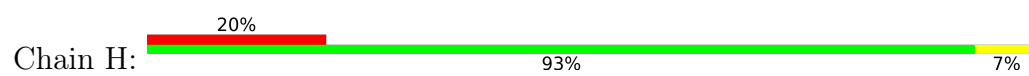
- Molecule 2: Exosporium leader peptide-containing protein



- Molecule 2: Exosporium leader peptide-containing protein



- Molecule 2: Exosporium leader peptide-containing protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.35Å 47.64Å 97.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.68 – 1.93 62.68 – 1.93	Depositor EDS
% Data completeness (in resolution range)	86.1 (62.68-1.93) 86.3 (62.68-1.93)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.08 (at 1.92Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.167 , 0.205 0.168 , 0.204	Depositor DCC
$R_{free}$ test set	1902 reflections (4.21%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.9	Xtriage
Anisotropy	0.072	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4082	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% 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: NA

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	0.31	0/1087	0.53	0/1497
1	E	0.33	0/1096	0.55	0/1509
1	G	0.32	0/1095	0.55	0/1509
2	D	0.27	0/117	0.45	0/165
2	F	0.27	0/117	0.39	0/165
2	H	0.34	0/117	0.59	0/165
All	All	0.32	0/3629	0.54	0/5010

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	1065	0	1090	6	0
1	E	1074	0	1109	9	0
1	G	1073	0	1108	8	0
2	D	111	0	115	1	0
2	F	111	0	115	0	0
2	H	111	0	115	0	0
3	G	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	153	0	0	0	0
4	D	18	0	0	0	0
4	E	162	0	0	1	3
4	F	12	0	0	0	0
4	G	165	0	0	1	2
4	H	25	0	0	0	1
All	All	4082	0	3652	21	3

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 3.

All (21) 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:G:46:SER:HB3	4:G:364:HOH:O	2.03	0.59
1:A:78:ILE:HD11	1:A:137:LEU:HD11	1.88	0.55
1:G:21:LEU:HD12	1:G:22:PRO:HD2	1.91	0.52
1:E:60:ASP:OD1	1:E:72:ARG:HD3	2.10	0.52
1:A:120:VAL:HG23	1:A:121:ILE:HG12	1.92	0.52
1:E:20:THR:N	4:E:203:HOH:O	2.45	0.49
1:E:26:PHE:CD1	1:E:162:THR:HG22	2.49	0.48
1:A:114:THR:HB	1:E:128:VAL:HG21	1.97	0.46
1:E:94:GLU:HG2	1:E:152:GLY:HA3	1.98	0.46
1:G:78:ILE:HD11	1:G:137:LEU:HD11	1.99	0.45
1:G:97:ARG:HD3	1:G:149:GLN:OE1	2.16	0.45
1:A:97:ARG:NH2	2:D:6:VAL:O	2.51	0.44
1:G:26:PHE:CD1	1:G:162:THR:HG22	2.53	0.43
1:E:97:ARG:HD3	1:E:149:GLN:OE1	2.18	0.43
1:A:26:PHE:CD1	1:A:162:THR:HG22	2.54	0.42
1:E:134:LEU:HB2	1:G:26:PHE:HB2	2.01	0.42
1:A:39:PHE:HA	1:A:146:VAL:HG13	2.03	0.41
1:E:41:PRO:HB3	1:E:144:GLN:HB3	2.03	0.41
1:G:41:PRO:HB3	1:G:144:GLN:HB3	2.03	0.40
1:G:85:SER:HB3	1:G:155:ASP:HB3	2.03	0.40
1:E:70:ILE:HG13	1:E:137:LEU:HD13	2.04	0.40

All (3) 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)
4:E:308:HOH:O	4:G:435:HOH:O[1_565]	1.96	0.24
4:E:326:HOH:O	4:H:105:HOH:O[1_565]	2.03	0.17
4:E:265:HOH:O	4:G:360:HOH:O[1_565]	2.18	0.02

## 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	146/155 (94%)	142 (97%)	4 (3%)	0	100	100
1	E	146/155 (94%)	144 (99%)	2 (1%)	0	100	100
1	G	146/155 (94%)	143 (98%)	3 (2%)	0	100	100
2	D	13/15 (87%)	13 (100%)	0	0	100	100
2	F	13/15 (87%)	13 (100%)	0	0	100	100
2	H	13/15 (87%)	13 (100%)	0	0	100	100
All	All	477/510 (94%)	468 (98%)	9 (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	Rotameric	Outliers	Percentiles	
1	A	117/128 (91%)	117 (100%)	0	100	100
1	E	121/128 (94%)	121 (100%)	0	100	100
1	G	120/128 (94%)	120 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	14/14 (100%)	14 (100%)	0	100	100
2	F	14/14 (100%)	14 (100%)	0	100	100
2	H	14/14 (100%)	13 (93%)	1 (7%)	13	3
All	All	400/426 (94%)	399 (100%)	1 (0%)	86	86

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	6	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	GLN
1	E	59	ASN
1	E	107	ASN
1	E	138	ASN
1	G	77	GLN
1	G	138	ASN
1	G	144	GLN
2	H	4	ASN

### 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 ⓘ

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 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

### 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<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	148/155 (95%)	-0.44	3 (2%) 65 72	10, 16, 26, 45	0
1	E	148/155 (95%)	-0.51	4 (2%) 56 63	9, 14, 24, 45	0
1	G	147/155 (94%)	-0.56	3 (2%) 65 72	6, 14, 26, 39	1 (0%)
2	D	15/15 (100%)	0.17	2 (13%) 7 8	13, 26, 39, 40	0
2	F	15/15 (100%)	0.38	2 (13%) 7 8	13, 27, 54, 76	0
2	H	15/15 (100%)	0.56	3 (20%) 3 3	14, 25, 45, 55	0
All	All	488/510 (95%)	-0.42	17 (3%) 47 53	6, 15, 32, 76	1 (0%)

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	15	PRO	4.5
1	A	167	SER	3.4
1	A	20	THR	3.3
1	A	91	VAL	3.3
2	H	14	PRO	3.2
2	H	15	PRO	3.2
1	E	20	THR	3.1
1	G	21	LEU	2.9
2	F	14	PRO	2.5
2	H	1	PHE	2.4
1	E	167	SER	2.4
1	E	91	VAL	2.4
2	D	1	PHE	2.2
1	G	167	SER	2.1
2	D	15	PRO	2.1
1	E	21	LEU	2.1
1	G	127	ASP	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 oligosaccharides in this entry.

## 6.4 Ligands [i](#)

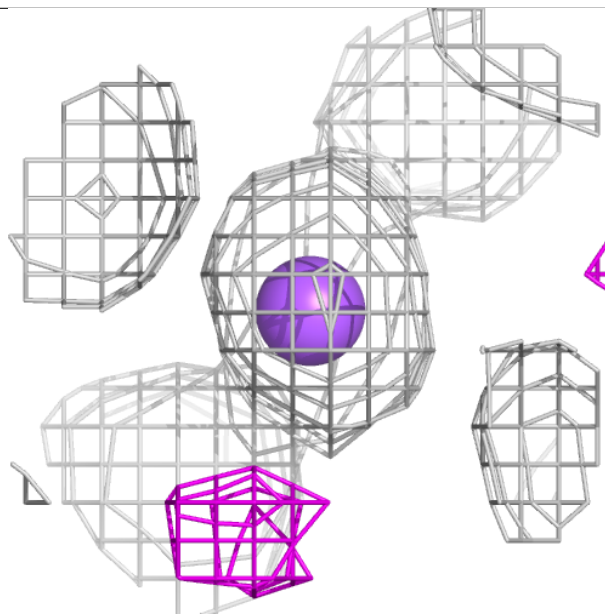
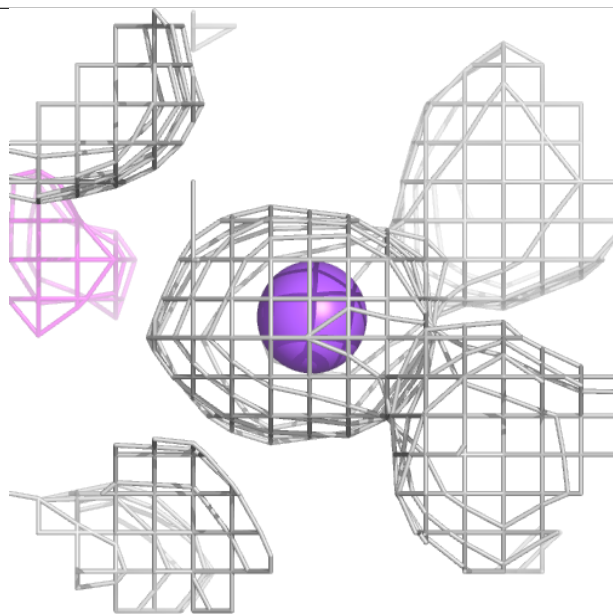
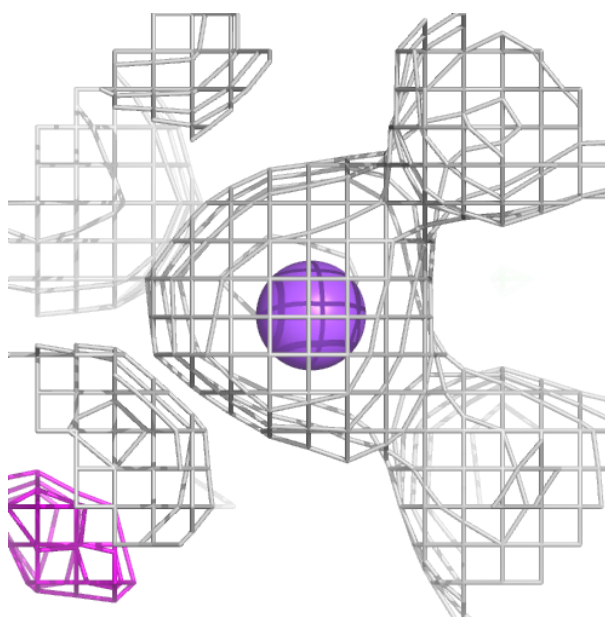
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	NA	G	201	1/1	0.93	0.11	33,33,33,33	0
3	NA	G	202	1/1	0.94	0.13	28,28,28,28	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 NA G 201:**

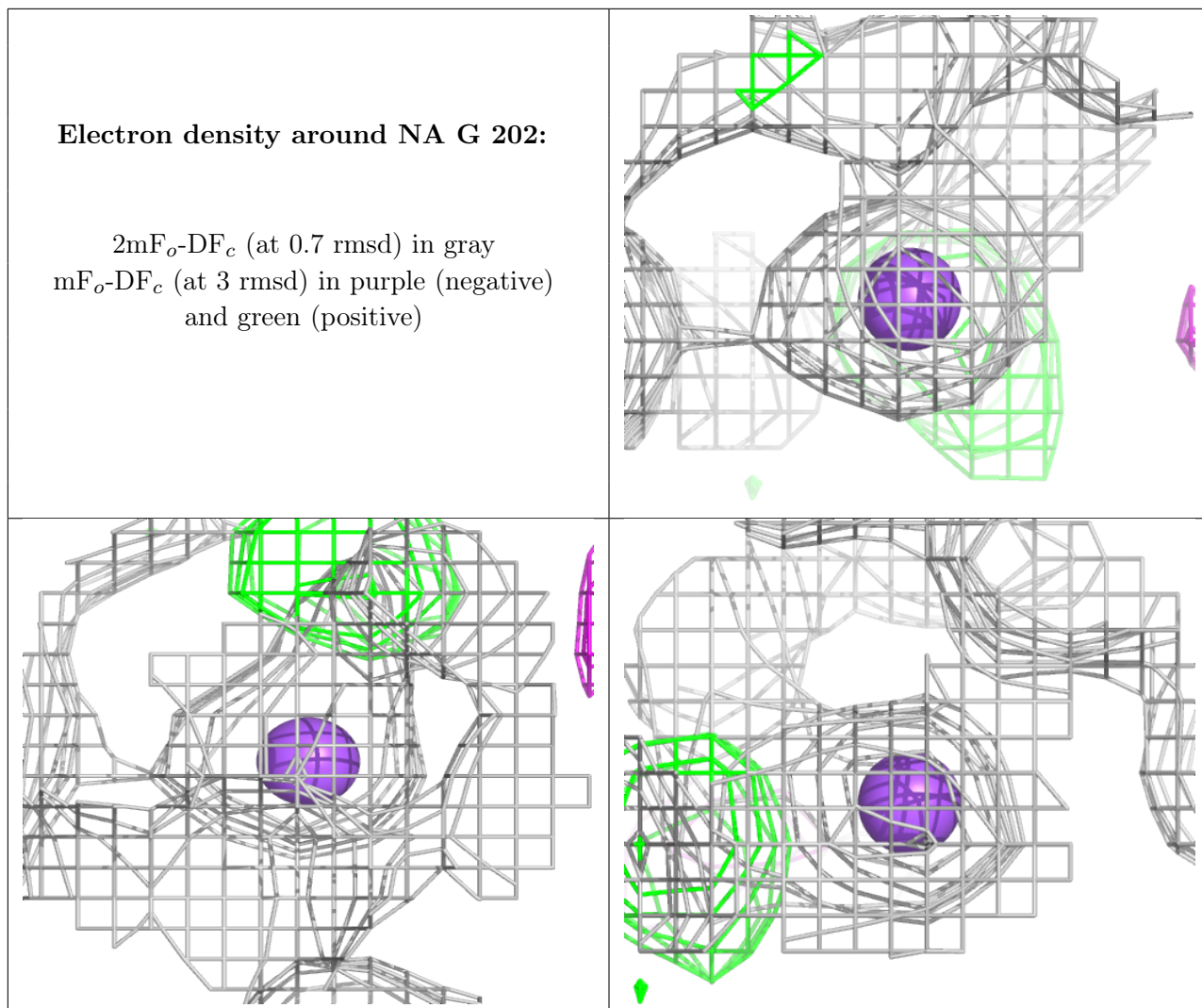
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around NA G 202:**

$2mF_o-DF_c$  (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.