



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 21, 2025 – 11:19 am BST

PDB ID : 9G5J / pdb\_00009g5j  
Title : Structure of the PRO-PRO endopeptidase (PPEP-3) E153A Y189F in complex with substrate peptide Ac-EPLPPPP-NH2 from *Geobacillus thermodenitrificans*  
Authors : Claushuis, B.; Wojtalla, F.; van Leeuwen, H.; Corver, J.; Baumann, U.; Hensbergen, P.  
Deposited on : 2024-07-17  
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](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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
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

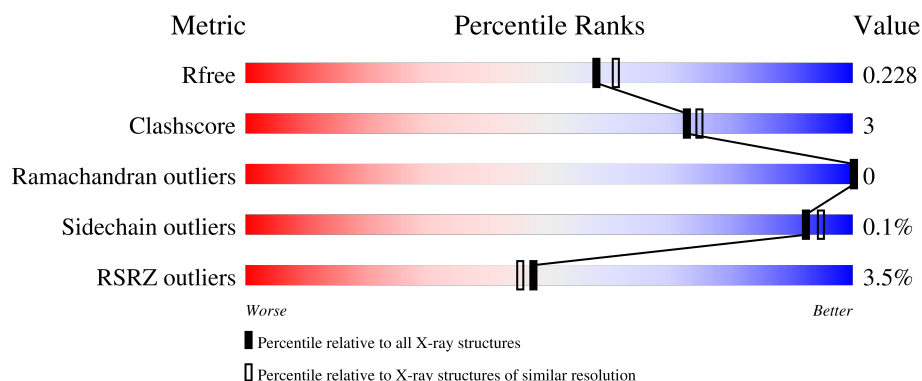
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (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 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	230	<div> <div>2%</div> <div> <div></div> <div>83%</div> <div>7%</div> <div>10%</div> </div> </div>
1	B	230	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>7%</div> <div>10%</div> </div> </div>
1	C	230	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>•</div> <div>10%</div> </div> </div>
1	D	230	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div>6%</div> <div>7%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	9	<div><div></div><div>78%</div><div>22%</div></div>
2	F	9	<div><div>22%</div><div>78%</div><div>11%</div><div>11%</div></div>
2	G	9	<div><div>22%</div><div>56%</div><div>33%</div><div>11%</div></div>
2	H	9	<div><div>11%</div><div>56%</div><div>44%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 14254 atoms, of which 6769 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 ATLF-like domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	208	Total	C	H	N	O	S	0	1	0
			3329	1083	1636	293	316	1			
1	B	207	Total	C	H	N	O	S	0	0	0
			3252	1068	1587	286	310	1			
1	C	207	Total	C	H	N	O	S	0	3	0
			3316	1084	1630	296	305	1			
1	D	215	Total	C	H	N	O	S	0	0	0
			3397	1110	1666	301	318	2			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	MET	-	initiating methionine	UNP A4IN Y2
A	7	GLY	-	expression tag	UNP A4IN Y2
A	8	SER	-	expression tag	UNP A4IN Y2
A	9	SER	-	expression tag	UNP A4IN Y2
A	10	HIS	-	expression tag	UNP A4IN Y2
A	11	HIS	-	expression tag	UNP A4IN Y2
A	12	HIS	-	expression tag	UNP A4IN Y2
A	13	HIS	-	expression tag	UNP A4IN Y2
A	14	HIS	-	expression tag	UNP A4IN Y2
A	15	HIS	-	expression tag	UNP A4IN Y2
A	16	SER	-	expression tag	UNP A4IN Y2
A	17	SER	-	expression tag	UNP A4IN Y2
A	18	GLY	-	expression tag	UNP A4IN Y2
A	19	LEU	-	expression tag	UNP A4IN Y2
A	20	VAL	-	expression tag	UNP A4IN Y2
A	21	PRO	-	expression tag	UNP A4IN Y2
A	22	ARG	-	expression tag	UNP A4IN Y2
A	23	GLY	-	expression tag	UNP A4IN Y2
A	24	SER	-	expression tag	UNP A4IN Y2
A	25	HIS	-	expression tag	UNP A4IN Y2
A	26	MET	-	expression tag	UNP A4IN Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	154	ALA	GLU	engineered mutation	UNP A4IN Y2
A	190	PHE	TYR	engineered mutation	UNP A4IN Y2
B	6	MET	-	initiating methionine	UNP A4IN Y2
B	7	GLY	-	expression tag	UNP A4IN Y2
B	8	SER	-	expression tag	UNP A4IN Y2
B	9	SER	-	expression tag	UNP A4IN Y2
B	10	HIS	-	expression tag	UNP A4IN Y2
B	11	HIS	-	expression tag	UNP A4IN Y2
B	12	HIS	-	expression tag	UNP A4IN Y2
B	13	HIS	-	expression tag	UNP A4IN Y2
B	14	HIS	-	expression tag	UNP A4IN Y2
B	15	HIS	-	expression tag	UNP A4IN Y2
B	16	SER	-	expression tag	UNP A4IN Y2
B	17	SER	-	expression tag	UNP A4IN Y2
B	18	GLY	-	expression tag	UNP A4IN Y2
B	19	LEU	-	expression tag	UNP A4IN Y2
B	20	VAL	-	expression tag	UNP A4IN Y2
B	21	PRO	-	expression tag	UNP A4IN Y2
B	22	ARG	-	expression tag	UNP A4IN Y2
B	23	GLY	-	expression tag	UNP A4IN Y2
B	24	SER	-	expression tag	UNP A4IN Y2
B	25	HIS	-	expression tag	UNP A4IN Y2
B	26	MET	-	expression tag	UNP A4IN Y2
B	154	ALA	GLU	engineered mutation	UNP A4IN Y2
B	190	PHE	TYR	engineered mutation	UNP A4IN Y2
C	6	MET	-	initiating methionine	UNP A4IN Y2
C	7	GLY	-	expression tag	UNP A4IN Y2
C	8	SER	-	expression tag	UNP A4IN Y2
C	9	SER	-	expression tag	UNP A4IN Y2
C	10	HIS	-	expression tag	UNP A4IN Y2
C	11	HIS	-	expression tag	UNP A4IN Y2
C	12	HIS	-	expression tag	UNP A4IN Y2
C	13	HIS	-	expression tag	UNP A4IN Y2
C	14	HIS	-	expression tag	UNP A4IN Y2
C	15	HIS	-	expression tag	UNP A4IN Y2
C	16	SER	-	expression tag	UNP A4IN Y2
C	17	SER	-	expression tag	UNP A4IN Y2
C	18	GLY	-	expression tag	UNP A4IN Y2
C	19	LEU	-	expression tag	UNP A4IN Y2
C	20	VAL	-	expression tag	UNP A4IN Y2
C	21	PRO	-	expression tag	UNP A4IN Y2
C	22	ARG	-	expression tag	UNP A4IN Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	23	GLY	-	expression tag	UNP A4IN2
C	24	SER	-	expression tag	UNP A4IN2
C	25	HIS	-	expression tag	UNP A4IN2
C	26	MET	-	expression tag	UNP A4IN2
C	154	ALA	GLU	engineered mutation	UNP A4IN2
C	190	PHE	TYR	engineered mutation	UNP A4IN2
D	6	MET	-	initiating methionine	UNP A4IN2
D	7	GLY	-	expression tag	UNP A4IN2
D	8	SER	-	expression tag	UNP A4IN2
D	9	SER	-	expression tag	UNP A4IN2
D	10	HIS	-	expression tag	UNP A4IN2
D	11	HIS	-	expression tag	UNP A4IN2
D	12	HIS	-	expression tag	UNP A4IN2
D	13	HIS	-	expression tag	UNP A4IN2
D	14	HIS	-	expression tag	UNP A4IN2
D	15	HIS	-	expression tag	UNP A4IN2
D	16	SER	-	expression tag	UNP A4IN2
D	17	SER	-	expression tag	UNP A4IN2
D	18	GLY	-	expression tag	UNP A4IN2
D	19	LEU	-	expression tag	UNP A4IN2
D	20	VAL	-	expression tag	UNP A4IN2
D	21	PRO	-	expression tag	UNP A4IN2
D	22	ARG	-	expression tag	UNP A4IN2
D	23	GLY	-	expression tag	UNP A4IN2
D	24	SER	-	expression tag	UNP A4IN2
D	25	HIS	-	expression tag	UNP A4IN2
D	26	MET	-	expression tag	UNP A4IN2
D	154	ALA	GLU	engineered mutation	UNP A4IN2
D	190	PHE	TYR	engineered mutation	UNP A4IN2

- Molecule 2 is a protein called ACE-GLU-PRO-LEU-PRO-PRO-PRO-NH<sub>2</sub>.

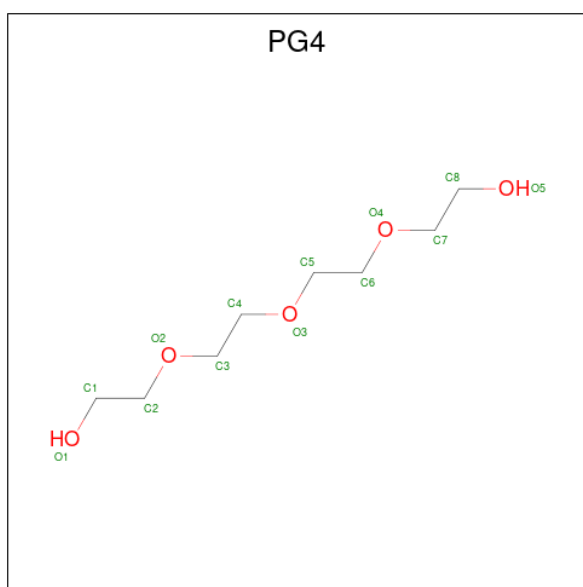
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	9	Total	C	H	N	O	0	0	1
			113	38	57	8	10			
2	F	8	Total	C	H	N	O	0	0	1
			99	34	50	8	7			
2	G	8	Total	C	H	N	O	0	0	1
			107	36	54	8	9			
2	H	9	Total	C	H	N	O	0	0	1
			113	38	57	8	10			

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by

depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		
3	D	1	Total	Zn	0	0
			1	1		

- Molecule 4 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			31	8	18	5		

- Molecule 5 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 6 is water.

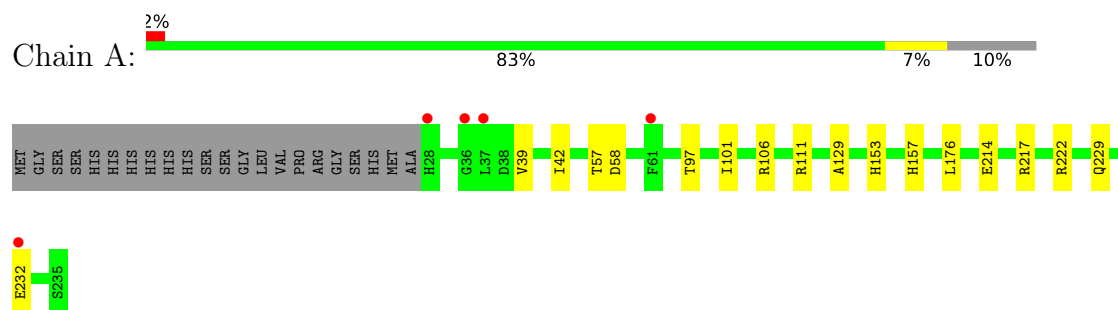
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	135	Total	O	0	0
			135	135		
6	B	111	Total	O	0	0
			111	111		
6	C	115	Total	O	0	0
			115	115		
6	D	93	Total	O	0	0
			93	93		
6	E	5	Total	O	0	0
			5	5		
6	F	3	Total	O	0	0
			3	3		
6	G	3	Total	O	0	0
			3	3		
6	H	4	Total	O	0	0
			4	4		



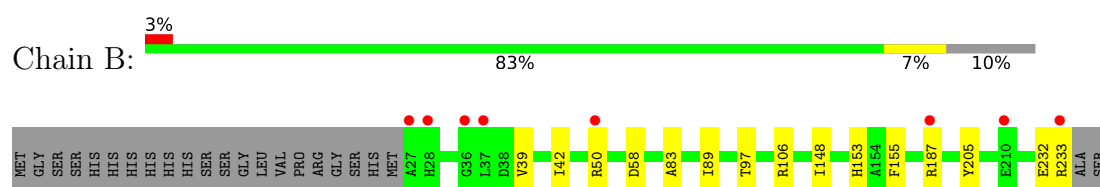
### 3 Residue-property plots [i](#)

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.

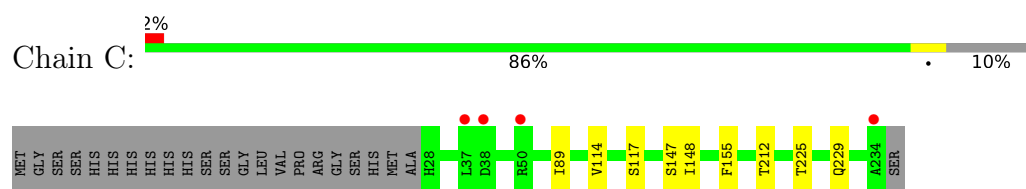
- Molecule 1: ATLF-like domain-containing protein



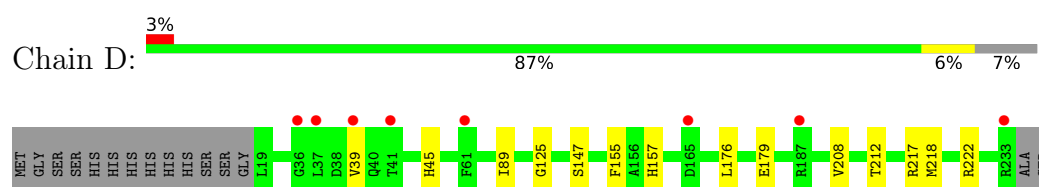
- Molecule 1: ATLF-like domain-containing protein



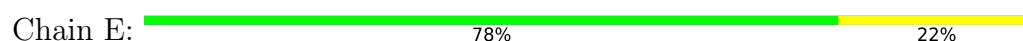
- Molecule 1: ATLF-like domain-containing protein



- Molecule 1: ATLF-like domain-containing protein

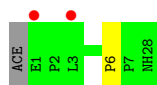
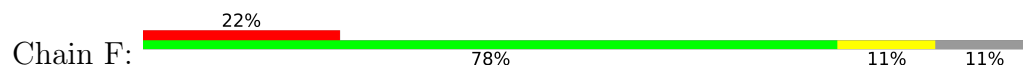


- Molecule 2: ACE-GLU-PRO-LEU-PRO-PRO-PRO-NH2

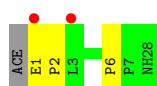




- Molecule 2: ACE-GLU-PRO-LEU-PRO-PRO-PRO-PRO-NH2



- Molecule 2: ACE-GLU-PRO-LEU-PRO-PRO-PRO-PRO-NH2



- Molecule 2: ACE-GLU-PRO-LEU-PRO-PRO-PRO-PRO-NH2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.43Å 95.41Å 126.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.71 – 2.00 47.71 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.71-2.00) 99.0 (47.71-2.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.00Å)	Xtriage
Refinement program	PHENIX (2.0rc1_5617: ???)	Depositor
R, $R_{free}$	0.190 , 0.228 0.191 , 0.228	Depositor DCC
$R_{free}$ test set	58252 reflections (4.32%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.4	Xtriage
Anisotropy	0.291	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 51.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14254	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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: ACE, PGE, PG4, NH2, ZN

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	0.15	0/1742	0.33	0/2368
1	B	0.14	0/1712	0.32	0/2332
1	C	0.13	0/1743	0.32	0/2373
1	D	0.15	0/1780	0.34	0/2422
2	E	0.48	0/57	0.73	0/82
2	F	0.56	0/52	0.73	0/75
2	G	0.55	0/56	0.75	0/80
2	H	0.52	0/57	0.75	0/82
All	All	0.17	0/7199	0.35	0/9814

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	1693	1636	1635	13	0
1	B	1665	1587	1587	11	0
1	C	1686	1630	1632	5	0
1	D	1731	1666	1666	13	0
2	E	56	57	55	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	49	50	50	1	0
2	G	53	54	54	3	0
2	H	56	57	55	4	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	13	18	18	0	0
5	D	10	14	14	0	0
6	A	135	0	0	0	0
6	B	111	0	0	1	0
6	C	115	0	0	0	0
6	D	93	0	0	0	0
6	E	5	0	0	0	0
6	F	3	0	0	0	0
6	G	3	0	0	0	0
6	H	4	0	0	0	0
All	All	7485	6769	6766	42	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 3.

All (42) 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:176:LEU:HD22	1:A:222:ARG:HB3	1.65	0.78
1:D:176:LEU:HD22	1:D:222:ARG:HB3	1.69	0.74
1:D:179:GLU:OE1	1:D:222:ARG:NH1	2.25	0.69
1:D:176:LEU:CD2	1:D:222:ARG:HB3	2.28	0.64
1:C:148:ILE:HD11	1:C:212:THR:HG21	1.82	0.62
1:D:157:HIS:ND1	2:H:3:LEU:HD13	2.15	0.60
1:B:50:ARG:HD3	1:B:83:ALA:HB1	1.83	0.59
1:A:176:LEU:CD2	1:A:222:ARG:HB3	2.34	0.55
1:A:111:ARG:NH2	2:E:3:LEU:HD22	2.24	0.53
1:B:39:VAL:HG12	1:B:42:ILE:HD12	1.92	0.52
1:B:50:ARG:CD	1:B:83:ALA:HB1	2.40	0.50
1:A:101:ILE:HD11	1:B:58:ASP:OD1	2.12	0.50
1:B:187:ARG:HH22	1:D:212:THR:CG2	2.25	0.49
1:A:39:VAL:HB	1:A:42:ILE:HD12	1.94	0.49
1:D:176:LEU:HD22	1:D:222:ARG:CB	2.41	0.49
2:H:1:GLU:OE1	2:H:3:LEU:HG	2.14	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:89:ILE:HD11	1:C:155:PHE:CD1	2.48	0.47
1:D:39:VAL:HG23	1:D:45:HIS:HB2	1.97	0.47
1:C:114[A]:VAL:HG23	1:C:117:SER:HB3	1.97	0.47
1:A:97:THR:HB	1:A:106:ARG:HG2	1.98	0.46
1:A:157:HIS:ND1	2:E:3:LEU:HD13	2.31	0.46
1:B:153:HIS:CE1	2:F:6:PRO:HD3	2.50	0.46
1:D:147:SER:N	2:H:6:PRO:HG2	2.31	0.45
1:D:179:GLU:OE2	1:D:222:ARG:HD2	2.16	0.45
1:A:129:ALA:HB3	6:B:809:HOH:O	2.17	0.45
1:A:176:LEU:HD22	1:A:222:ARG:CB	2.41	0.45
1:A:153:HIS:CE1	2:E:6:PRO:HD3	2.52	0.44
1:B:97:THR:HB	1:B:106:ARG:HG2	1.99	0.44
1:A:57:THR:O	1:A:58:ASP:C	2.61	0.44
1:A:214:GLU:OE1	1:A:217:ARG:NH1	2.52	0.43
1:D:89:ILE:HD11	1:D:155:PHE:CD1	2.53	0.42
1:B:187:ARG:NH2	1:D:212:THR:CG2	2.82	0.42
1:B:232:GLU:O	1:B:233:ARG:HB2	2.19	0.42
1:B:89:ILE:HD11	1:B:155:PHE:CD1	2.54	0.42
1:B:148:ILE:HD11	1:B:205:TYR:HA	2.00	0.42
1:D:125:GLY:O	2:H:5:PRO:HD3	2.19	0.42
1:C:225:THR:O	1:C:229:GLN:HG3	2.20	0.42
2:G:1:GLU:HB3	2:G:2:PRO:HD3	2.02	0.42
1:D:217:ARG:C	1:D:218:MET:HE2	2.46	0.41
1:C:147:SER:HB2	2:G:6:PRO:HD2	2.02	0.41
1:A:229:GLN:HA	1:A:232:GLU:OE1	2.21	0.41
2:G:1:GLU:N	2:G:2:PRO:HD2	2.37	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	207/230 (90%)	203 (98%)	4 (2%)	0	100	100
1	B	205/230 (89%)	203 (99%)	2 (1%)	0	100	100
1	C	208/230 (90%)	206 (99%)	2 (1%)	0	100	100
1	D	213/230 (93%)	208 (98%)	5 (2%)	0	100	100
2	E	7/9 (78%)	7 (100%)	0	0	100	100
2	F	6/9 (67%)	6 (100%)	0	0	100	100
2	G	6/9 (67%)	6 (100%)	0	0	100	100
2	H	7/9 (78%)	7 (100%)	0	0	100	100
All	All	859/956 (90%)	846 (98%)	13 (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	179/197 (91%)	179 (100%)	0	100	100
1	B	173/197 (88%)	173 (100%)	0	100	100
1	C	176/197 (89%)	176 (100%)	0	100	100
1	D	181/197 (92%)	180 (99%)	1 (1%)	84	88
2	E	7/7 (100%)	7 (100%)	0	100	100
2	F	6/7 (86%)	6 (100%)	0	100	100
2	G	7/7 (100%)	7 (100%)	0	100	100
2	H	7/7 (100%)	7 (100%)	0	100	100
All	All	736/816 (90%)	735 (100%)	1 (0%)	92	95

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

Mol	Chain	Res	Type
1	D	208	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	45	HIS
1	C	81	GLN
1	C	85	HIS
1	D	81	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

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	$\# Z  > 2$
5	PGE	D	702	-	9,9,9	0.27	0	8,8,8	0.48	0
4	PG4	A	702	-	12,12,12	0.28	0	11,11,11	0.19	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
5	PGE	D	702	-	-	3/7/7/7	-
4	PG4	A	702	-	-	5/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	702	PGE	O2-C3-C4-O3
4	A	702	PG4	O3-C5-C6-O4
4	A	702	PG4	O1-C1-C2-O2
5	D	702	PGE	C4-C3-O2-C2
4	A	702	PG4	C1-C2-O2-C3
5	D	702	PGE	C1-C2-O2-C3
4	A	702	PG4	O2-C3-C4-O3
4	A	702	PG4	C6-C5-O3-C4

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	208/230 (90%)	-0.21	5 (2%) 59 58	17, 30, 56, 72	1 (0%)
1	B	207/230 (90%)	-0.16	8 (3%) 44 42	21, 31, 53, 93	0
1	C	207/230 (90%)	-0.15	4 (1%) 66 65	17, 33, 60, 88	3 (1%)
1	D	215/230 (93%)	-0.03	8 (3%) 45 43	24, 35, 63, 81	0
2	E	7/9 (77%)	-0.32	0 100 100	25, 26, 32, 40	0
2	F	7/9 (77%)	0.97	2 (28%) 1 1	28, 33, 69, 80	0
2	G	7/9 (77%)	0.87	2 (28%) 1 1	27, 32, 62, 101	0
2	H	7/9 (77%)	0.28	1 (14%) 7 6	28, 33, 41, 60	0
All	All	865/956 (90%)	-0.12	30 (3%) 47 45	17, 32, 60, 101	4 (0%)

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	27	ALA	4.8
1	B	28	HIS	4.6
1	A	61	PHE	3.6
2	F	3	LEU	3.4
1	D	37	LEU	3.4
1	A	37	LEU	3.3
1	D	36	GLY	3.1
1	D	61	PHE	3.0
2	G	1	GLU	3.0
1	A	36	GLY	3.0
2	H	1	GLU	2.9
1	B	37	LEU	2.8
1	A	232	GLU	2.8
1	B	50	ARG	2.7
2	F	1	GLU	2.7
1	B	233	ARG	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	50	ARG	2.6
1	D	233	ARG	2.5
1	D	187	ARG	2.5
1	C	38	ASP	2.4
1	C	234	ALA	2.4
1	A	28	HIS	2.3
1	B	187	ARG	2.2
2	G	3	LEU	2.2
1	D	41	THR	2.1
1	C	37	LEU	2.1
1	B	210	GLU	2.1
1	D	39	VAL	2.0
1	B	36	GLY	2.0
1	D	165	ASP	2.0

## 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.

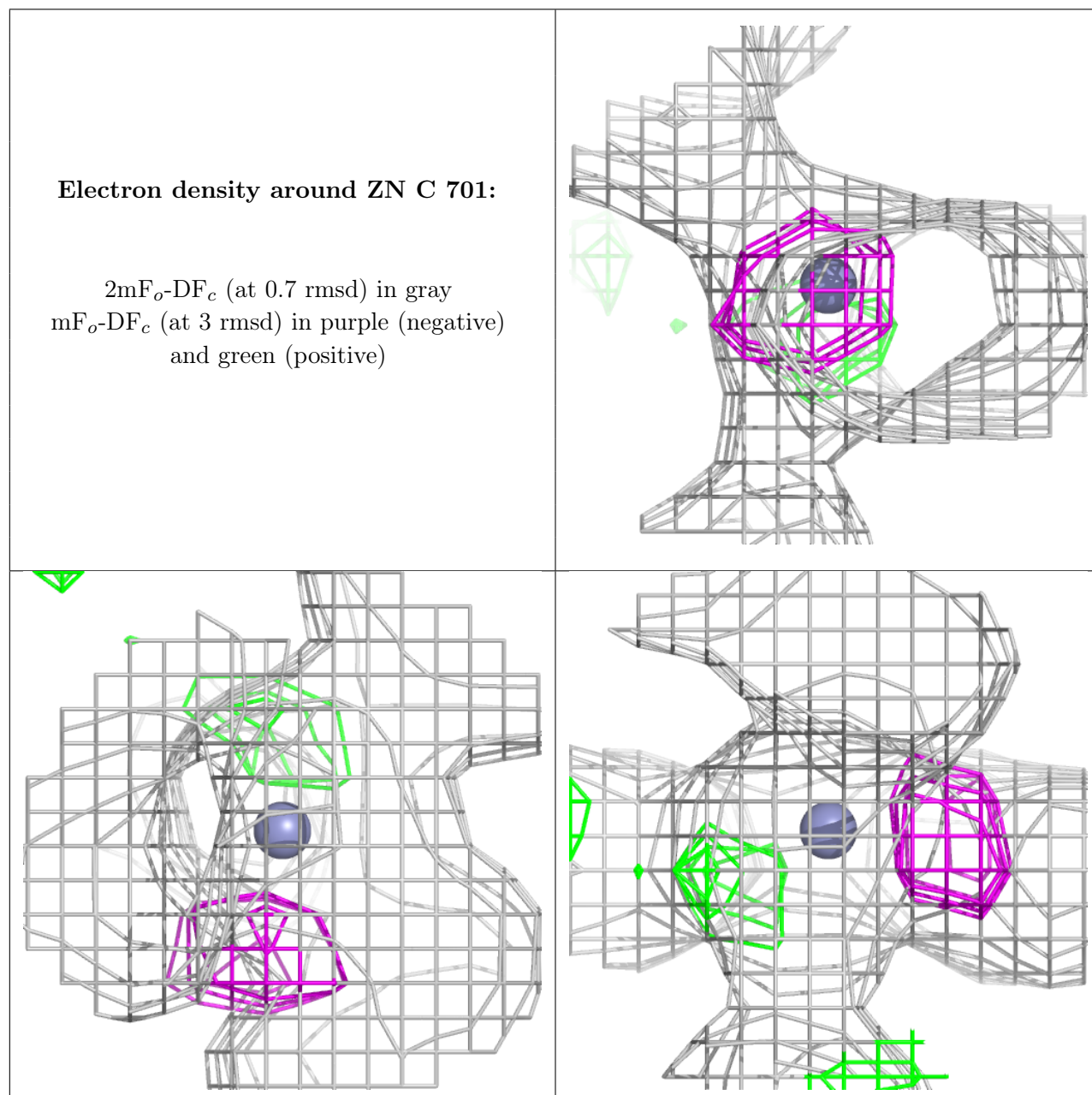
## 6.4 Ligands [i](#)

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	ZN	C	701	1/1	0.81	0.11	25,25,25,25	1
4	PG4	A	702	13/13	0.81	0.15	42,53,74,91	0
5	PGE	D	702	10/10	0.87	0.10	40,54,67,69	0
3	ZN	A	701	1/1	0.88	0.08	18,18,18,18	1
3	ZN	D	701	1/1	0.90	0.07	26,26,26,26	1
3	ZN	B	701	1/1	0.91	0.07	23,23,23,23	1

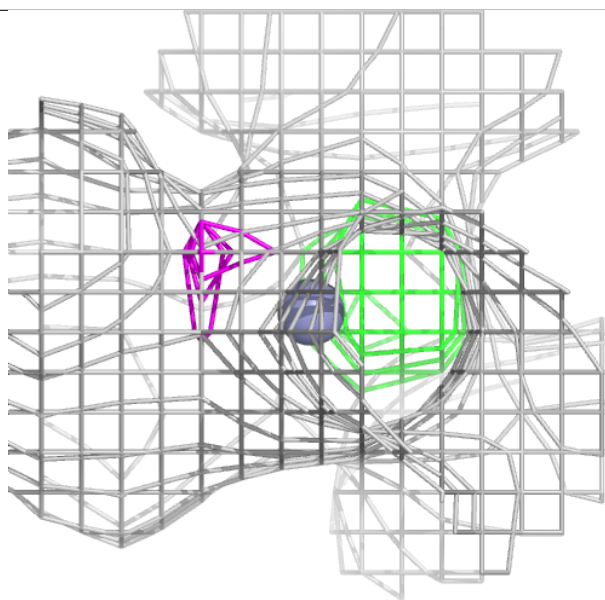
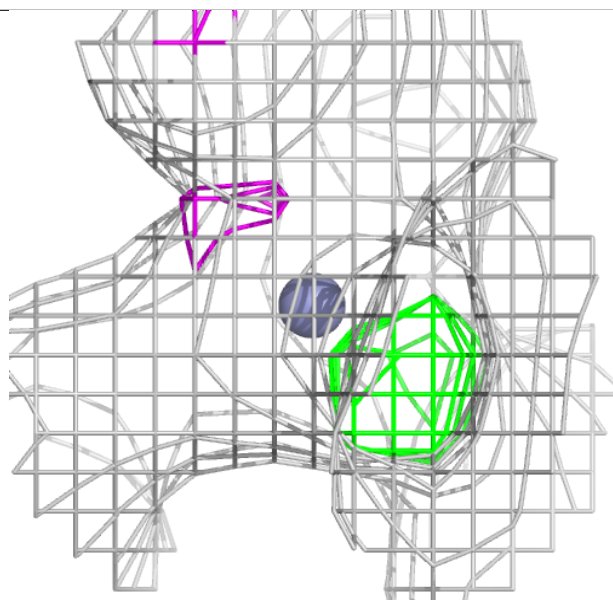
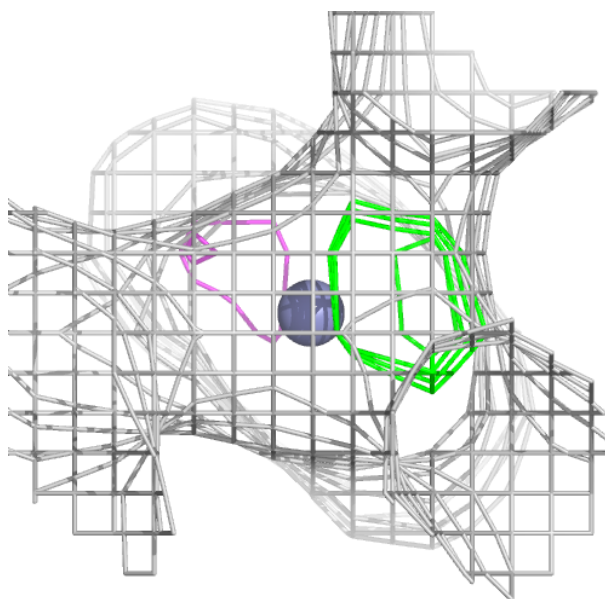
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



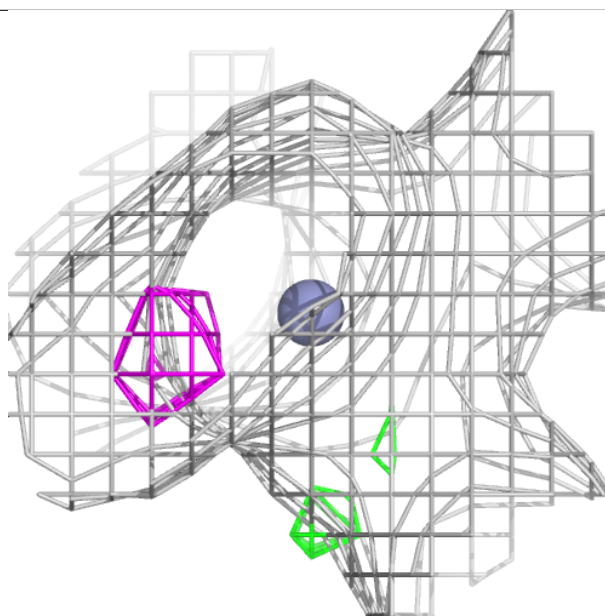
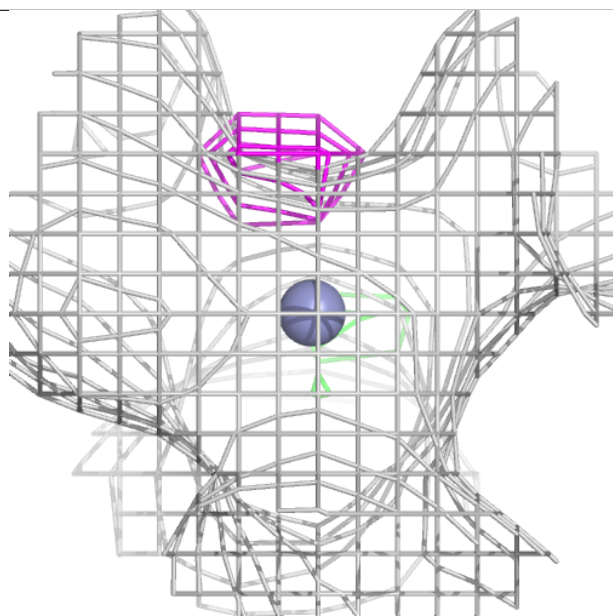
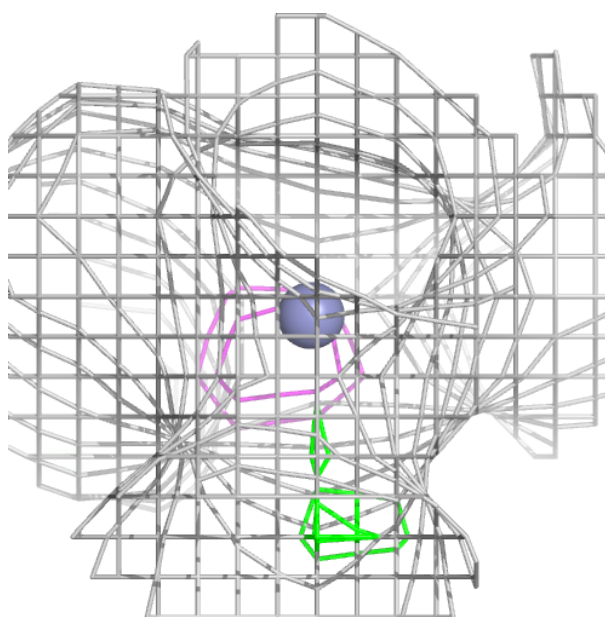
**Electron density around ZN A 701:**

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



**Electron density around ZN D 701:**

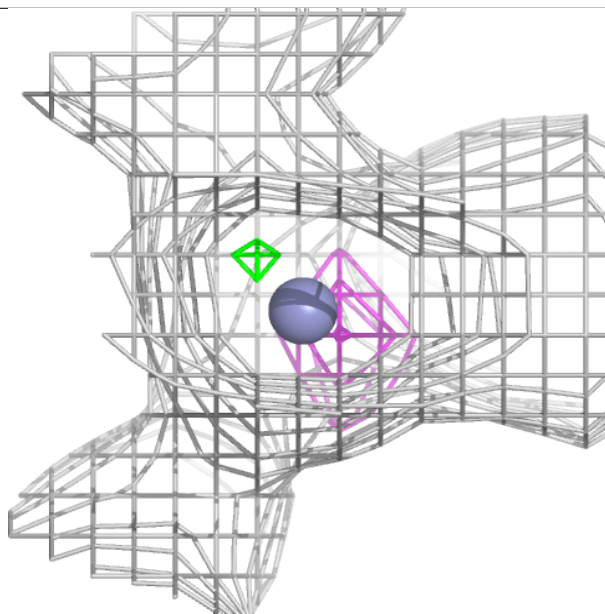
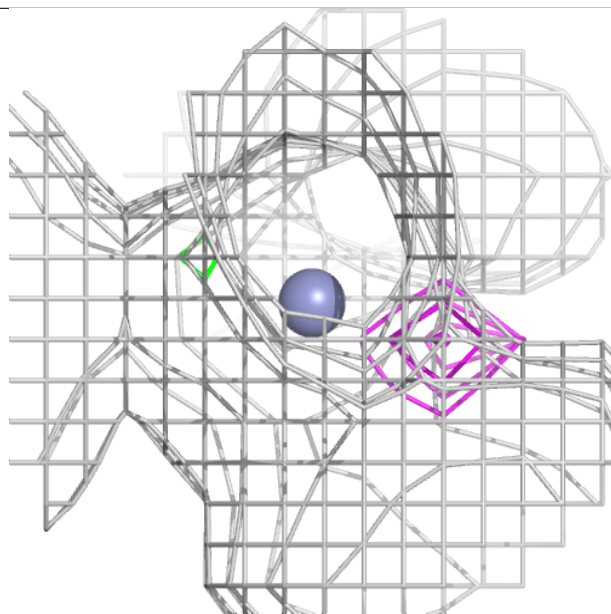
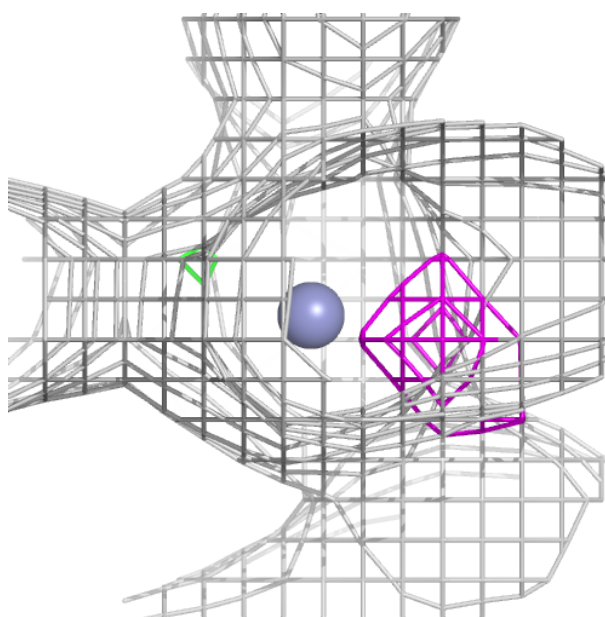
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around ZN 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)



## 6.5 Other polymers ⓘ

There are no such residues in this entry.