wwPDB X-ray Structure Validation Summary Report (i)

Mar 5, 2024 – 01:23 AM EST

PDB ID	:	8GDW
Title	:	Crystal structure of Domain Related to Iron (DRI) from cyanobacteria
Authors	:	Kumaran, D.; Grosjean, N.; Blaby, E.C.
Deposited on	:	2023-03-06
Resolution	:	2.35 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

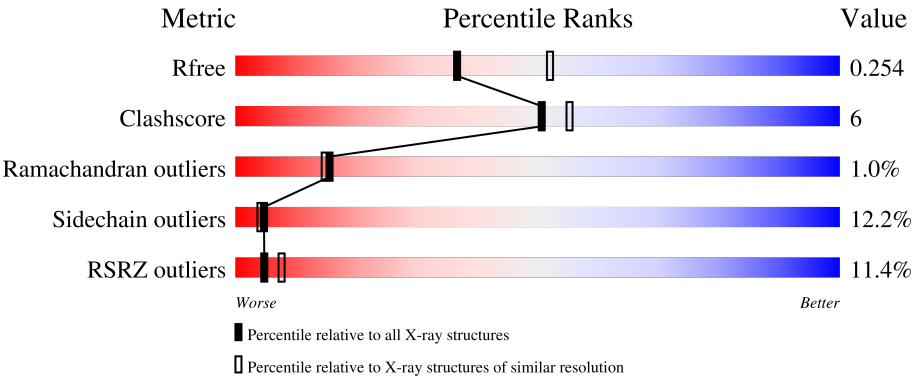
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.35 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	103	11% 78% 16% 5%
1	BBB	103	12% 83% 13%
1	CCC	103	15% 76% 20%
1	DDD	103	8% 74% 23%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3134 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 Ssr1698 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	101	Total	C	N	O	S	0	0	0
			774	478	133	157	6			
1	BBB	101	Total	C	N	O	S	0	0	0
			774	478	133	157	6			
1	CCC	101	Total	C	N	O	S	0	0	0
			774	478	133	157	6			
1	DDD	101	Total	C	N	O	S	0	0	0
			774	478	133	157	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	97	ALA	-	expression tag	UNP P73129
AAA	98	GLU	-	expression tag	UNP P73129
AAA	99	ASN	-	expression tag	UNP P73129
AAA	100	LEU	-	expression tag	UNP P73129
AAA	101	TYR	-	expression tag	UNP P73129
AAA	102	PHE	-	expression tag	UNP P73129
AAA	103	GLN	-	expression tag	UNP P73129
BBB	97	ALA	-	expression tag	UNP P73129
BBB	98	GLU	-	expression tag	UNP P73129
BBB	99	ASN	-	expression tag	UNP P73129
BBB	100	LEU	-	expression tag	UNP P73129
BBB	101	TYR	-	expression tag	UNP P73129
BBB	102	PHE	-	expression tag	UNP P73129
BBB	103	GLN	-	expression tag	UNP P73129
CCC	97	ALA	-	expression tag	UNP P73129
CCC	98	GLU	-	expression tag	UNP P73129
CCC	99	ASN	-	expression tag	UNP P73129
CCC	100	LEU	-	expression tag	UNP P73129
CCC	101	TYR	-	expression tag	UNP P73129
CCC	102	PHE	-	expression tag	UNP P73129
CCC	103	GLN	-	expression tag	UNP P73129

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	97	ALA	-	expression tag	UNP P73129
DDD	98	GLU	-	expression tag	UNP P73129
DDD	99	ASN	-	expression tag	UNP P73129
DDD	100	LEU	-	expression tag	UNP P73129
DDD	101	TYR	-	expression tag	UNP P73129
DDD	102	PHE	-	expression tag	UNP P73129
DDD	103	GLN	-	expression tag	UNP P73129

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total Zn 1 1	0	0
2	BBB	1	Total Zn 1 1	0	0
2	CCC	1	Total Zn 1 1	0	0
2	DDD	1	Total Zn 1 1	0	0

- Molecule 3 is water.

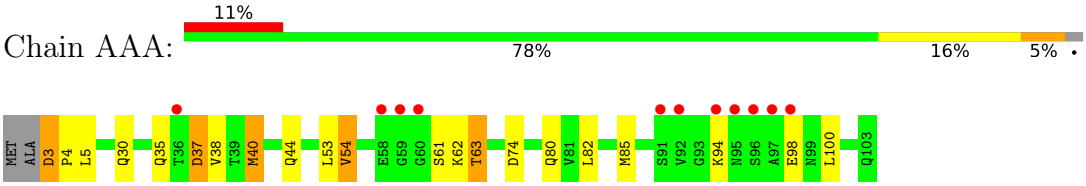
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	13	Total O 13 13	0	0
3	BBB	5	Total O 5 5	0	0
3	CCC	3	Total O 3 3	0	0
3	DDD	13	Total O 13 13	0	0

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.

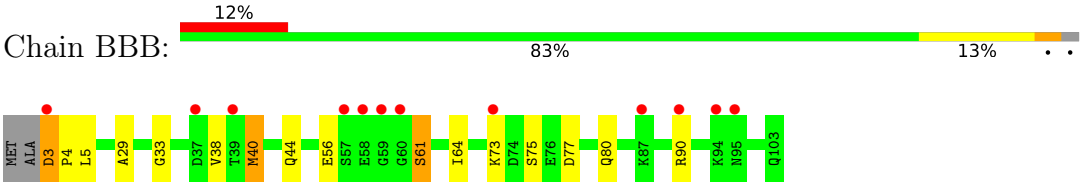
• Molecule 1: Ssr1698 protein

Chain AAA:



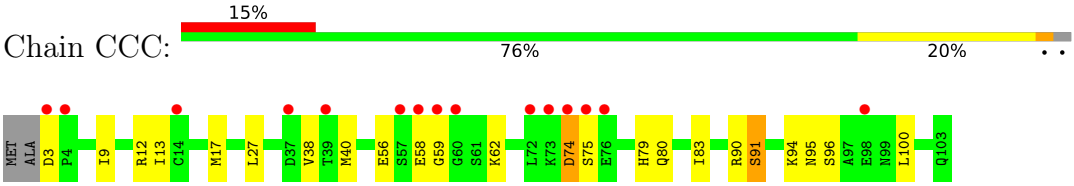
• Molecule 1: Ssr1698 protein

Chain BBB:



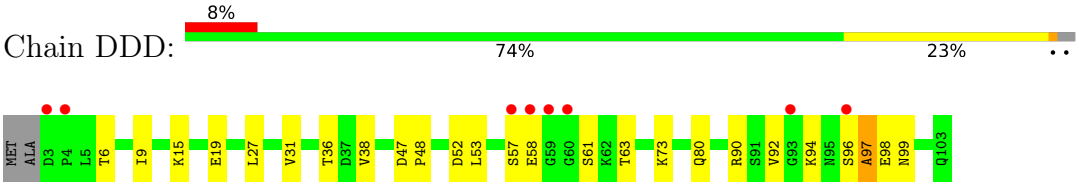
• Molecule 1: Ssr1698 protein

Chain CCC:



• Molecule 1: Ssr1698 protein

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	51.72Å 87.23Å 57.24Å 90.00° 94.98° 90.00°	Depositor
Resolution (Å)	29.50 – 2.35 29.48 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.0 (29.50-2.35) 99.0 (29.48-2.35)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 2.36Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.197 , 0.247 0.202 , 0.254	Depositor DCC
R_{free} test set	1027 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	56.6	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 39.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3134	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.01% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	AAA	0.67	0/785	0.83	0/1060
1	BBB	0.68	0/785	0.87	0/1060
1	CCC	0.67	0/785	0.85	0/1060
1	DDD	0.68	0/785	0.84	0/1060
All	All	0.67	0/3140	0.85	0/4240

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

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	AAA	774	0	753	13	0
1	BBB	774	0	753	5	0
1	CCC	774	0	753	8	0
1	DDD	774	0	753	12	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
2	CCC	1	0	0	0	0
2	DDD	1	0	0	0	0
3	AAA	13	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	BBB	5	0	0	0	0
3	CCC	3	0	0	1	0
3	DDD	13	0	0	0	0
All	All	3134	0	3012	35	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 6.

The worst 5 of 35 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:AAA:3:ASP:HB3	1:AAA:4:PRO:HD3	1.62	0.79
1:DDD:15:LYS:HE2	1:DDD:19:GLU:OE2	2.01	0.61
1:AAA:40:MET:HE1	1:BBB:40:MET:SD	2.48	0.53
1:DDD:27:LEU:HD21	1:DDD:90:ARG:HG3	1.90	0.53
1:AAA:37:ASP:OD1	1:AAA:37:ASP:N	2.28	0.53

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	AAA	99/103 (96%)	94 (95%)	5 (5%)	0	100 100
1	BBB	99/103 (96%)	93 (94%)	5 (5%)	1 (1%)	15 15
1	CCC	99/103 (96%)	91 (92%)	7 (7%)	1 (1%)	15 15
1	DDD	99/103 (96%)	92 (93%)	5 (5%)	2 (2%)	7 5
All	All	396/412 (96%)	370 (93%)	22 (6%)	4 (1%)	15 15

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	DDD	97	ALA
1	BBB	75	SER
1	DDD	98	GLU
1	CCC	59	GLY

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	AAA	84/85 (99%)	72 (86%)	12 (14%)	3	3
1	BBB	84/85 (99%)	74 (88%)	10 (12%)	5	4
1	CCC	84/85 (99%)	73 (87%)	11 (13%)	4	4
1	DDD	84/85 (99%)	76 (90%)	8 (10%)	8	7
All	All	336/340 (99%)	295 (88%)	41 (12%)	5	4

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	62	LYS
1	DDD	57	SER
1	CCC	74	ASP
1	CCC	100	LEU
1	DDD	61	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides 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

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	AAA	101/103 (98%)	0.52	11 (10%)	5	9	44, 57, 112, 152	0
1	BBB	101/103 (98%)	0.36	12 (11%)	4	7	46, 61, 100, 118	0
1	CCC	101/103 (98%)	0.54	15 (14%)	2	3	47, 65, 104, 174	0
1	DDD	101/103 (98%)	0.36	8 (7%)	12	19	44, 57, 107, 136	0
All	All	404/412 (98%)	0.45	46 (11%)	5	7	44, 61, 108, 174	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	95	ASN	10.8
1	CCC	74	ASP	6.5
1	DDD	59	GLY	5.1
1	AAA	59	GLY	5.0
1	BBB	90	ARG	4.6

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

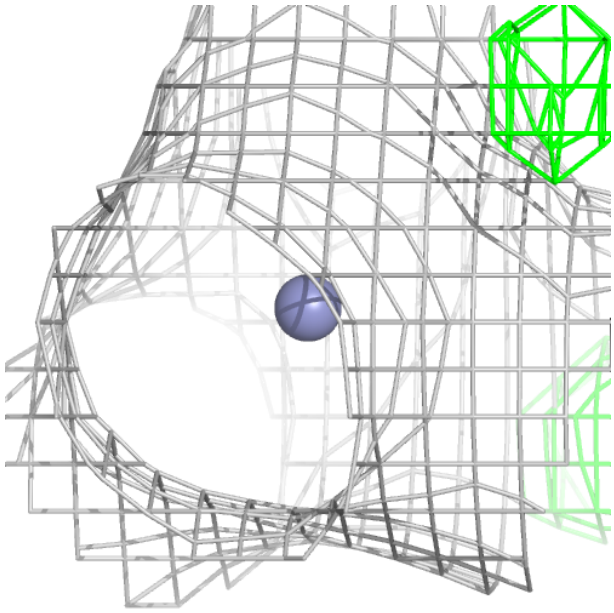
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

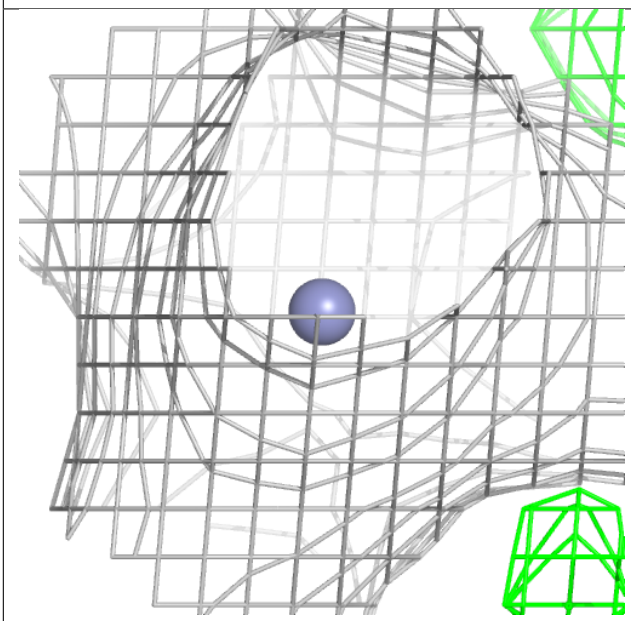
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	ZN	CCC	201	1/1	0.98	0.09	66,66,66,66	0
2	ZN	BBB	201	1/1	0.99	0.10	56,56,56,56	0
2	ZN	AAA	201	1/1	1.00	0.11	48,48,48,48	0
2	ZN	DDD	201	1/1	1.00	0.11	51,51,51,51	0

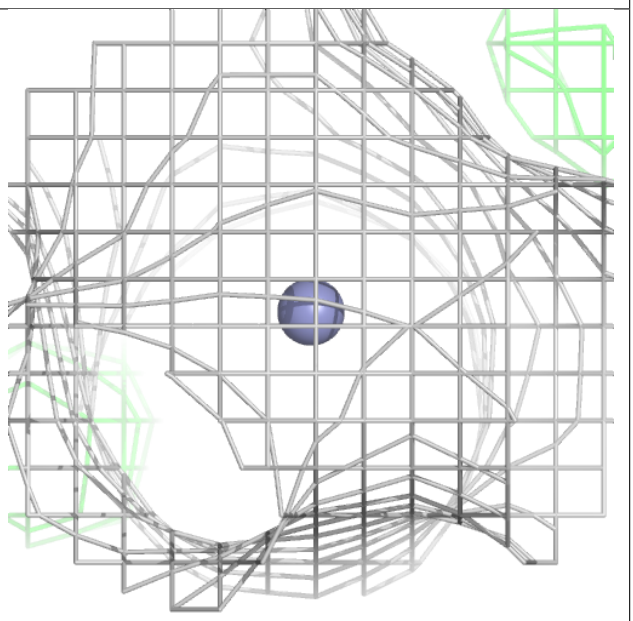
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around ZN CCC 201:

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

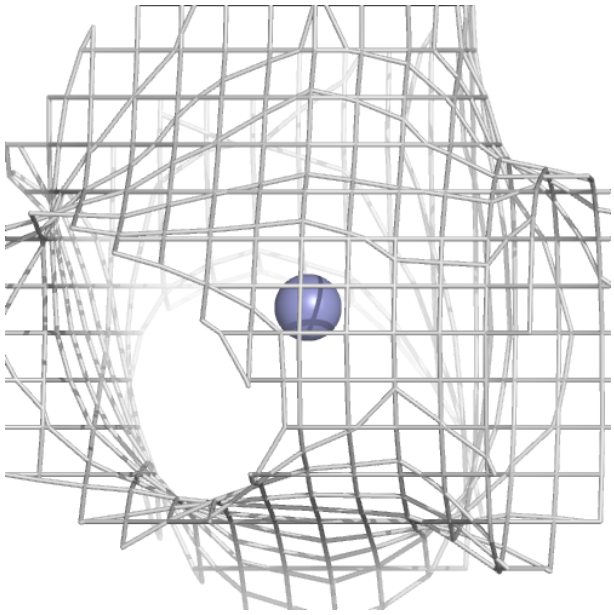


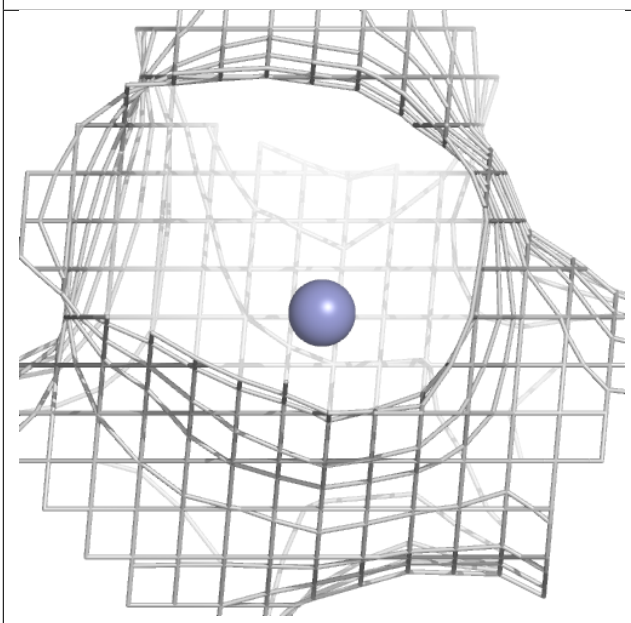


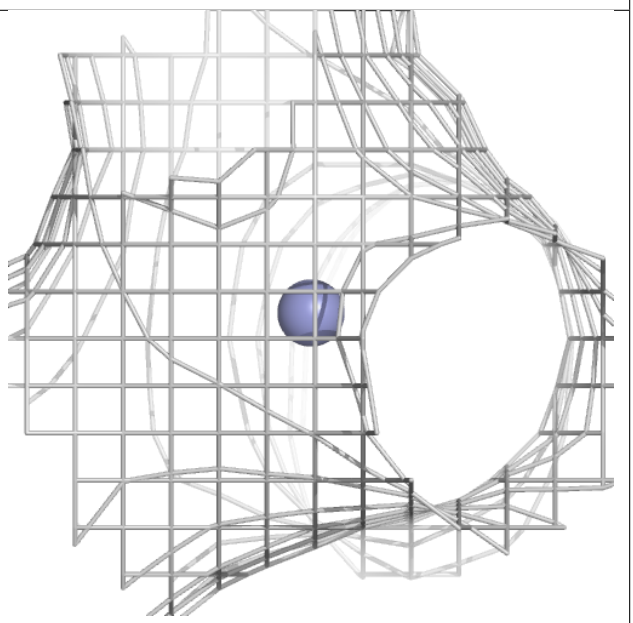


Electron density around ZN BBB 201:

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

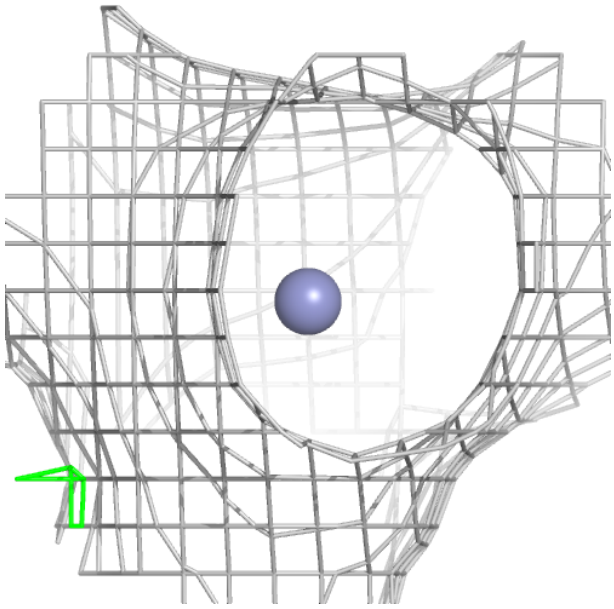


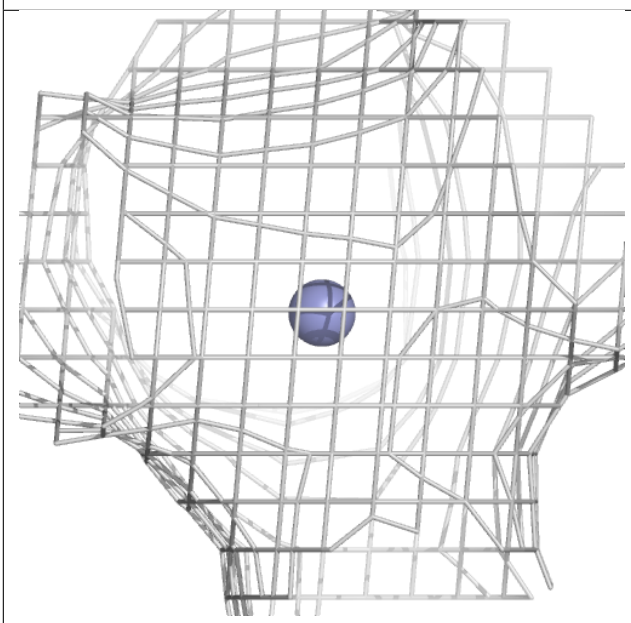


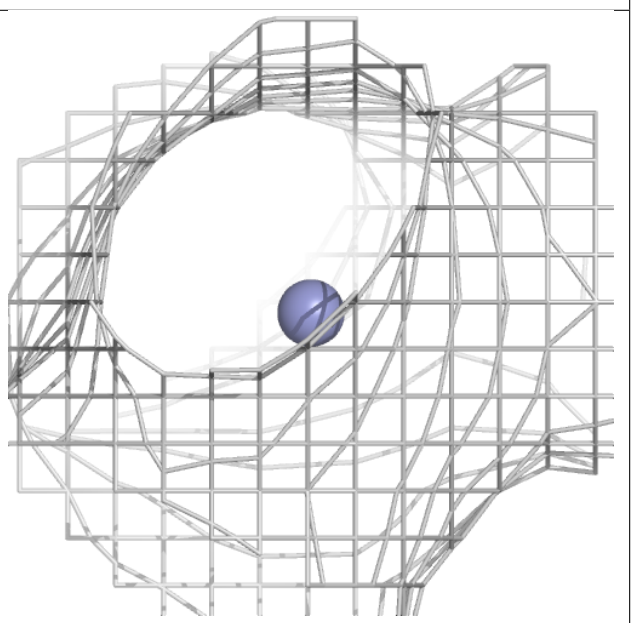


Electron density around ZN AAA 201:

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

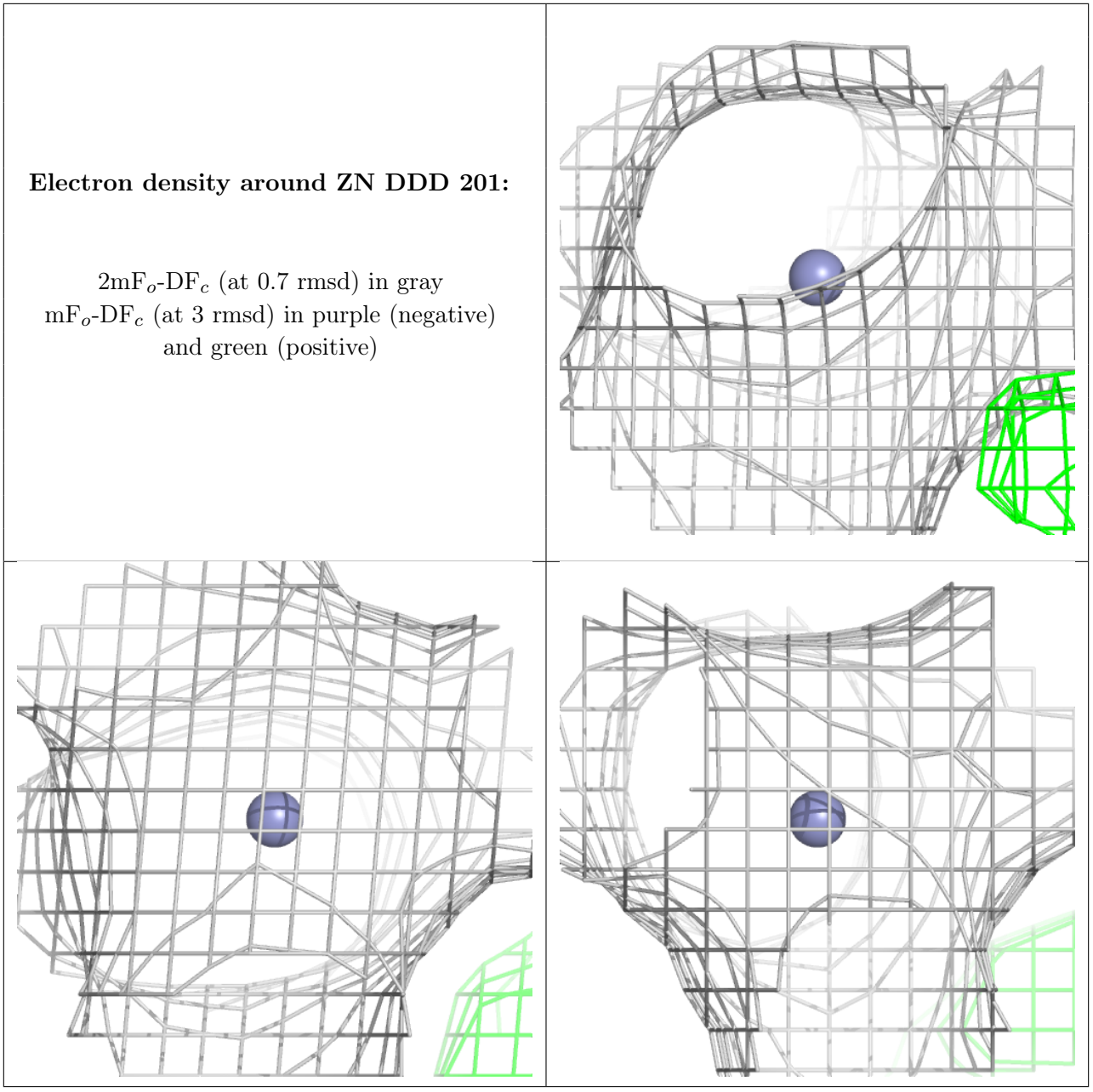






Electron density around ZN DDD 201:

$2mF_o$-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.