



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 05:32 AM UTC

PDB ID : 5Y8V / pdb_00005y8v
Title : Crystal structure of GAS41
Authors : Wang, Y.; Hao, Q.
Deposited on : 2017-08-21
Resolution : 2.61 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 ⓘ 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

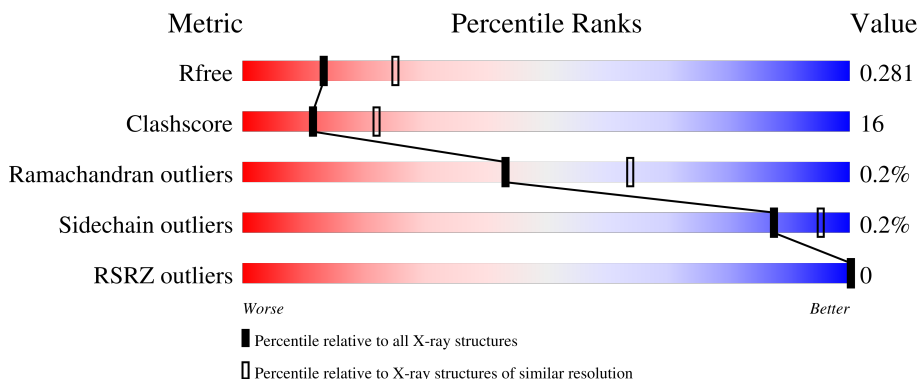
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.61 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}	180053	4951 (2.64-2.60)
Clashscore	190562	5303 (2.64-2.60)
Ramachandran outliers	187476	5217 (2.64-2.60)
Sidechain outliers	187428	5217 (2.64-2.60)
RSRZ outliers	180081	4950 (2.64-2.60)

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 $\leq 5\%$. 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	140	
1	B	140	
1	C	140	
1	D	140	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4543 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 YEATS domain-containing protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	134	Total	C	N	O	S	0	0	0
			1120	732	182	201	5			
1	A	134	Total	C	N	O	S	0	0	0
			1120	732	182	201	5			
1	C	134	Total	C	N	O	S	0	0	0
			1120	732	182	201	5			
1	D	130	Total	C	N	O	S	0	0	0
			1091	714	177	196	4			

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		

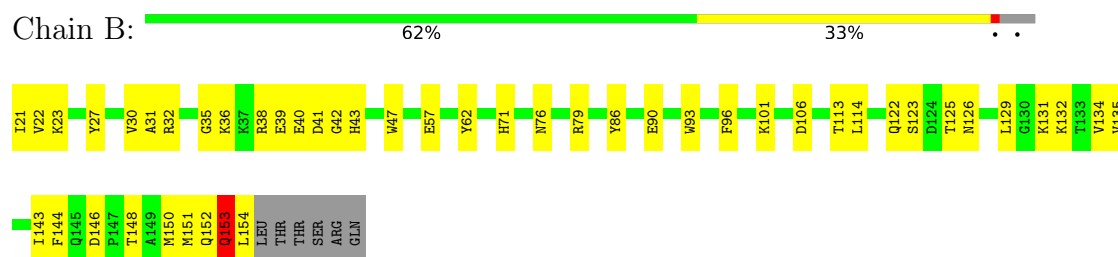
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	17	Total	O	0	0
			17	17		
3	A	12	Total	O	0	0
			12	12		
3	C	25	Total	O	0	0
			25	25		
3	D	34	Total	O	0	0
			34	34		

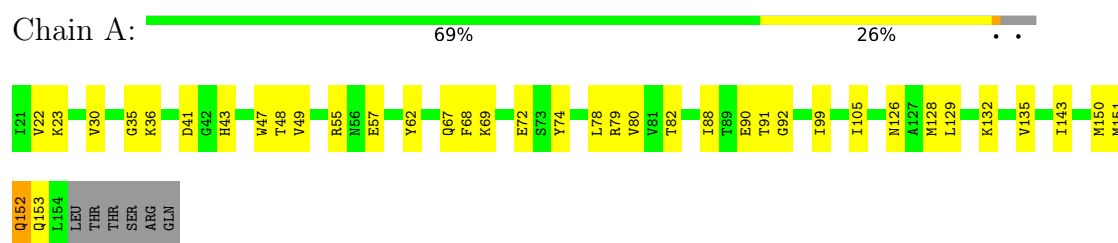
3 Residue-property plots [i](#)

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.

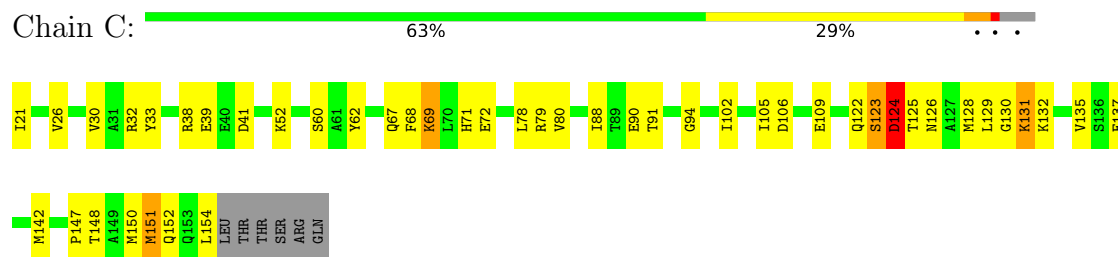
- Molecule 1: YEATS domain-containing protein 4



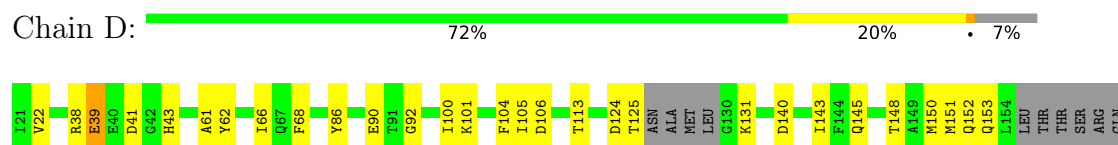
- Molecule 1: YEATS domain-containing protein 4



- Molecule 1: YEATS domain-containing protein 4



- Molecule 1: YEATS domain-containing protein 4



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	55.47Å 57.21Å 205.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.50 – 2.61 31.50 – 2.61	Depositor EDS
% Data completeness (in resolution range)	92.6 (31.50-2.61) 87.8 (31.50-2.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.237 , 0.280 0.239 , 0.281	Depositor DCC
R_{free} test set	858 reflections (4.13%)	wwPDB-VP
Wilson B-factor (Å ²)	45.2	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 29.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.186 for k,h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4543	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.60% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG

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	0.57	0/1152	0.87	3/1559 (0.2%)
1	B	0.58	0/1152	0.82	4/1559 (0.3%)
1	C	0.56	0/1152	0.90	4/1559 (0.3%)
1	D	0.61	0/1122	0.90	3/1517 (0.2%)
All	All	0.58	0/4578	0.87	14/6194 (0.2%)

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
1	C	0	4

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	131	LYS	CA-CB-CG	7.19	128.47	114.10
1	A	152	GLN	CA-C-N	6.30	133.04	121.70
1	A	152	GLN	C-N-CA	6.30	133.04	121.70
1	C	151	MET	CA-C-N	-6.14	113.28	122.24
1	C	151	MET	C-N-CA	-6.14	113.28	122.24

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	123	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	124	ASP	Peptide
1	C	152	GLN	Peptide
1	C	39	GLU	Sidechain

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	1120	0	1111	36	0
1	B	1120	0	1111	46	0
1	C	1120	0	1111	48	0
1	D	1091	0	1079	23	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	12	0	0	1	0
3	B	17	0	0	2	0
3	C	25	0	0	7	0
3	D	34	0	0	4	0
All	All	4543	0	4412	142	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 16.

The worst 5 of 142 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:151:MET:O	1:B:154:LEU:HD13	1.12	1.26
1:D:145:GLN:NE2	3:D:301:HOH:O	1.86	1.07
1:B:151:MET:O	1:B:154:LEU:CD1	2.07	1.02
1:B:86:TYR:OH	3:B:301:HOH:O	1.82	0.96
1:B:40:GLU:HG3	1:A:79:ARG:HB3	1.48	0.94

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	132/140 (94%)	125 (95%)	7 (5%)	0	100	100
1	B	132/140 (94%)	129 (98%)	3 (2%)	0	100	100
1	C	132/140 (94%)	123 (93%)	8 (6%)	1 (1%)	16	31
1	D	126/140 (90%)	119 (94%)	7 (6%)	0	100	100
All	All	522/560 (93%)	496 (95%)	25 (5%)	1 (0%)	43	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	124	ASP

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	123/129 (95%)	123 (100%)	0	100	100
1	B	123/129 (95%)	122 (99%)	1 (1%)	73	87
1	C	123/129 (95%)	123 (100%)	0	100	100
1	D	120/129 (93%)	120 (100%)	0	100	100
All	All	489/516 (95%)	488 (100%)	1 (0%)	87	95

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

Mol	Chain	Res	Type
1	B	153	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	153	GLN
1	C	126	ASN
1	D	46	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 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 ⓘ

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, 95th 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	A	134/140 (95%)	-1.43	0 100 100	34, 41, 63, 89	0
1	B	134/140 (95%)	-1.52	0 100 100	31, 39, 66, 79	0
1	C	134/140 (95%)	-1.40	0 100 100	36, 43, 85, 109	0
1	D	130/140 (92%)	-1.49	0 100 100	31, 38, 64, 82	0
All	All	532/560 (95%)	-1.46	0 100 100	31, 41, 68, 109	0

There are no RSRZ outliers to report.

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](#)

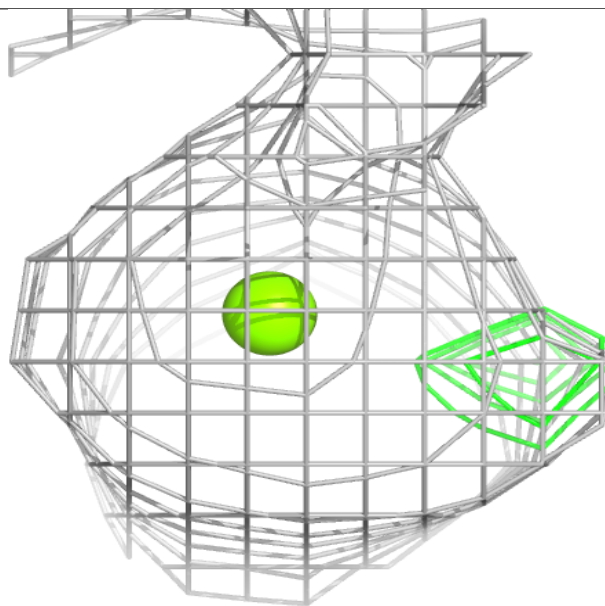
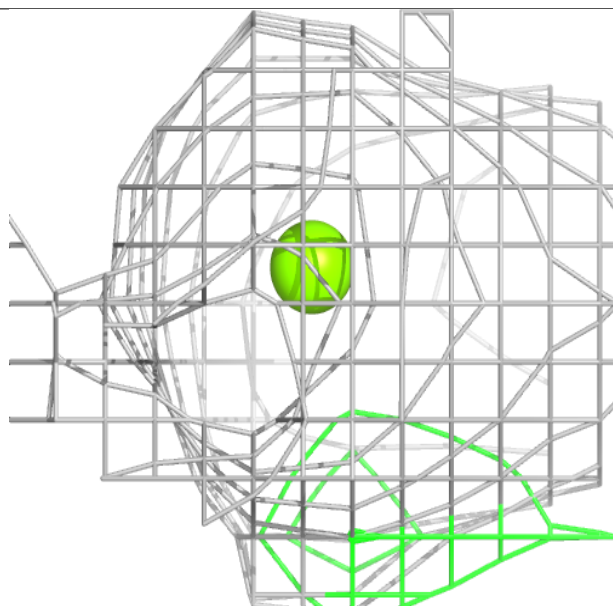
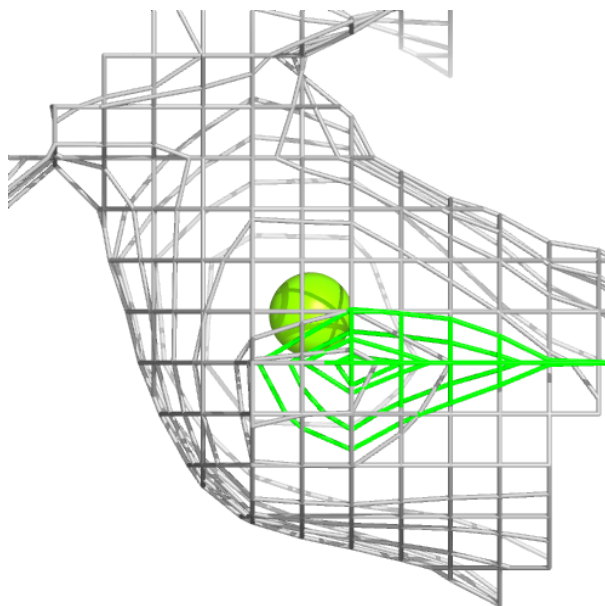
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists 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	MG	D	201	1/1	0.98	0.07	52,52,52,52	0
2	MG	C	201	1/1	0.99	0.05	60,60,60,60	0
2	MG	B	201	1/1	0.99	0.12	40,40,40,40	0
2	MG	A	201	1/1	1.00	0.11	52,52,52,52	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.

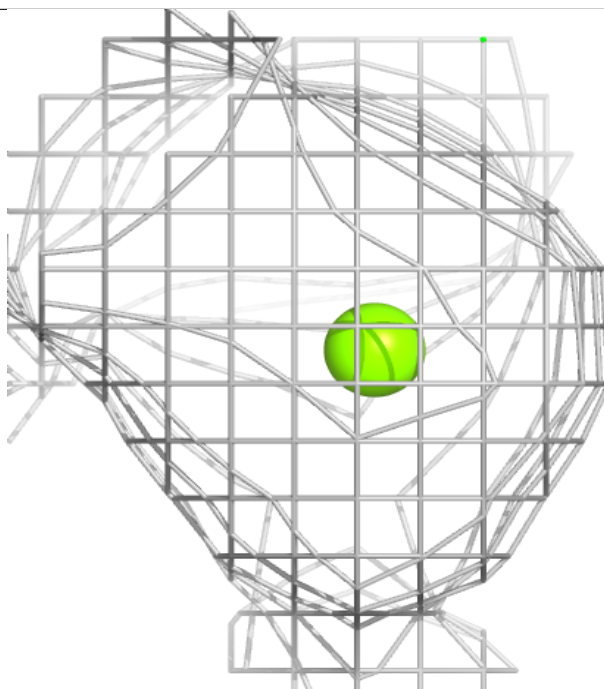
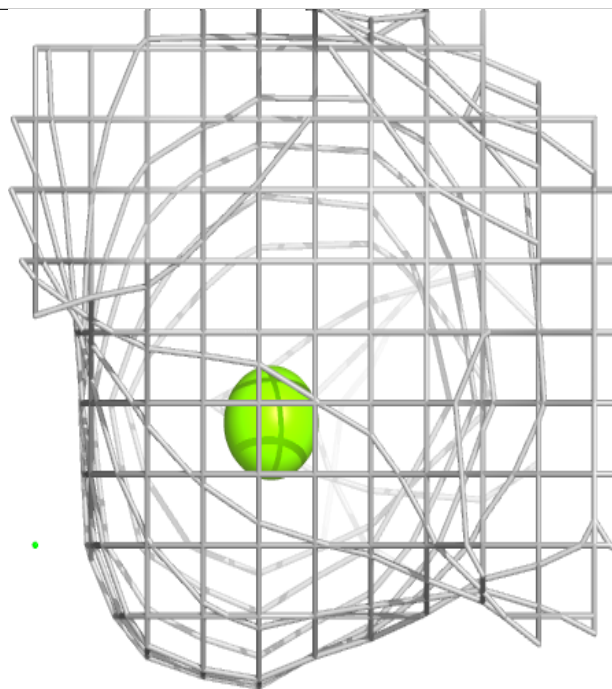
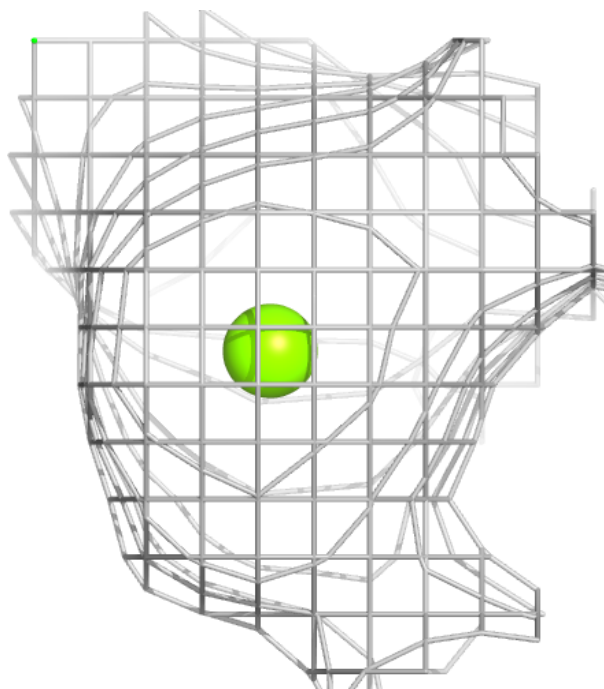
Electron density around MG D 201:

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



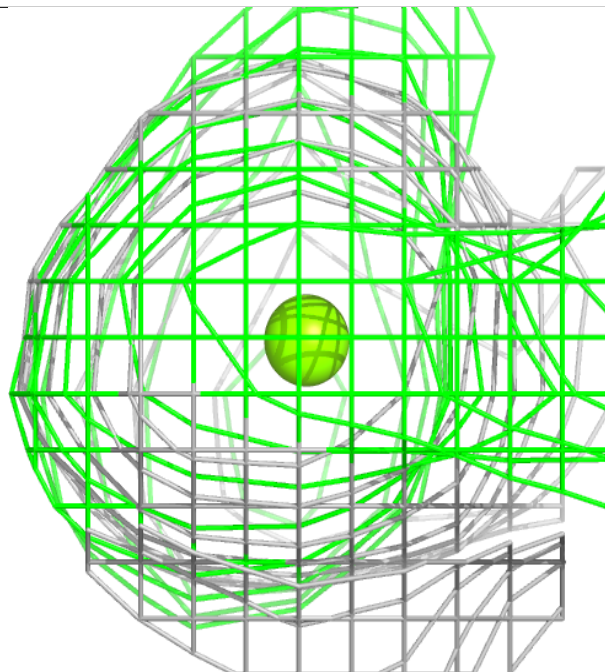
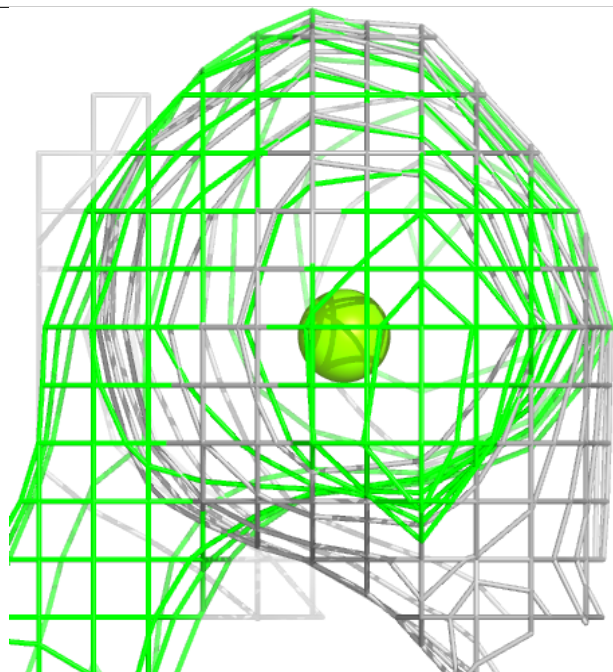
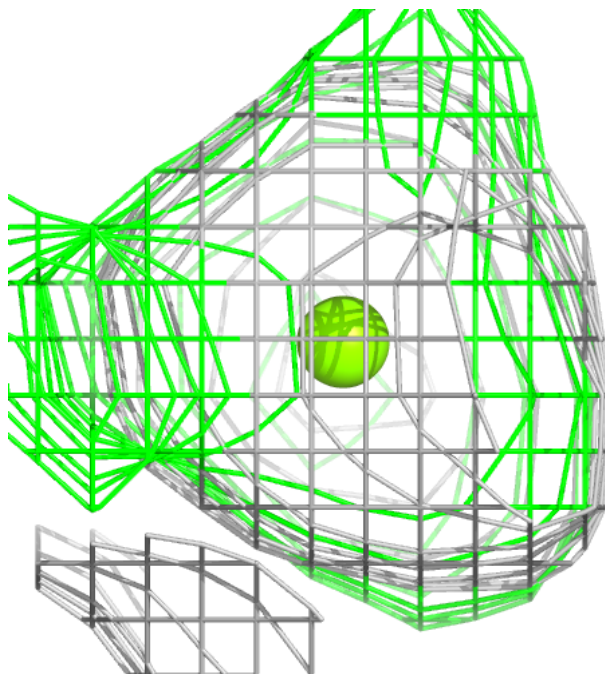
Electron density around MG C 201:

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



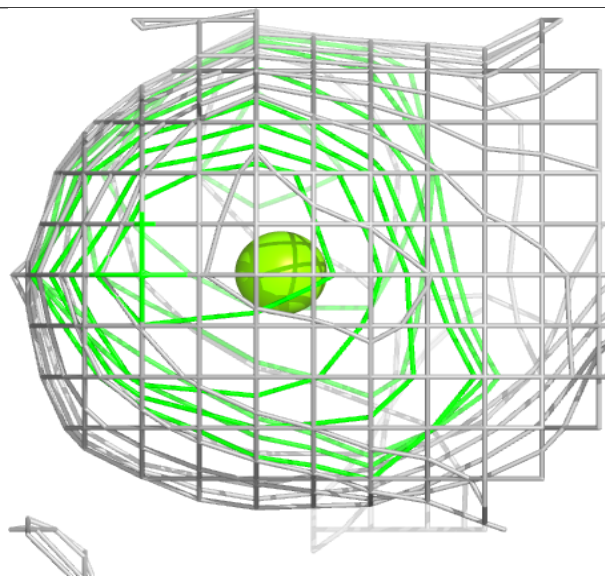
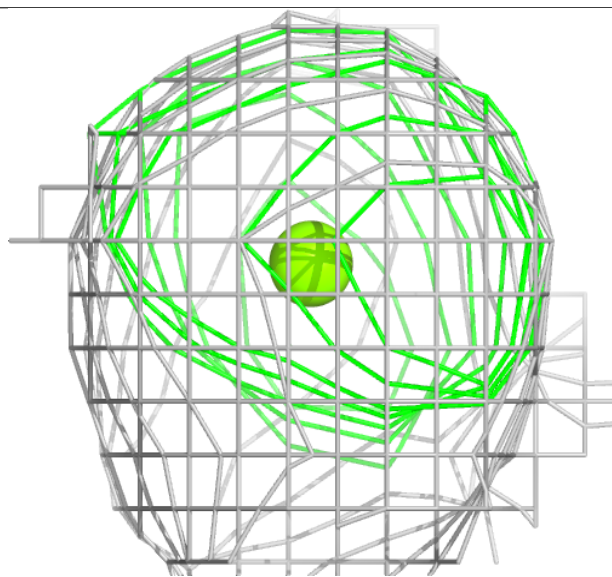
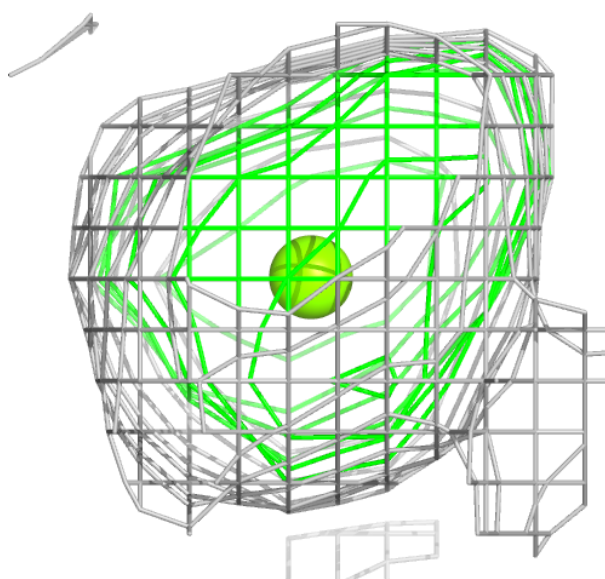
Electron density around MG B 201:

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



Electron density around MG A 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.