Full wwPDB X-ray Structure Validation Report

May 29, 2020 – 09:50 am BST

PDB ID	:	6CYP
Title	:	Crystal structure analysis of the H75I mutant of Superoxide Dismutase from Trichoderma reesei
Authors	:	Mendoza Rengifo, E.; Ferreira Jr., J.R.S.; Garratt, R.C.
Deposited on	:	2018-04-06
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the symbol.

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

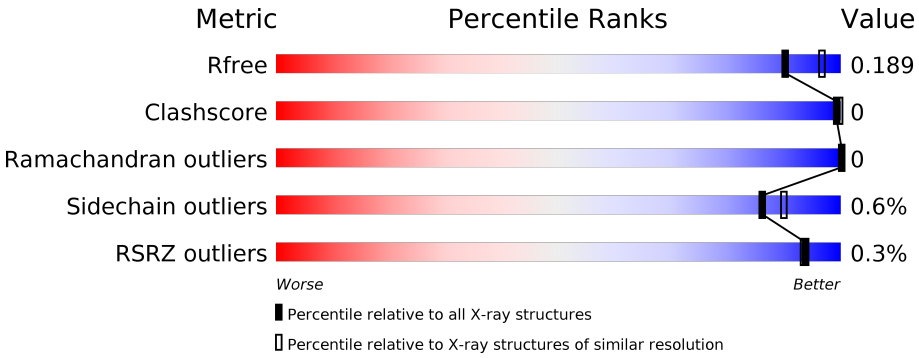
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.00 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 on 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	A	211	100%
1	B	211	100%
1	C	211	98%
1	D	211	100%
1	E	211	99%
1	F	211	100%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20544 atoms, of which 9589 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 Superoxide dismutase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	211	Total	C	H	N	O	S	0	0	0
			3231	1050	1599	272	306	4			
1	B	211	Total	C	H	N	O	S	0	0	0
			3231	1050	1599	272	306	4			
1	C	211	Total	C	H	N	O	S	0	0	0
			3231	1050	1599	272	306	4			
1	D	211	Total	C	H	N	O	S	0	0	0
			3227	1050	1595	272	306	4			
1	E	211	Total	C	H	N	O	S	0	0	0
			3231	1050	1599	272	306	4			
1	F	211	Total	C	H	N	O	S	0	0	0
			3230	1050	1598	272	306	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	75	ILE	HIS	engineered mutation	UNP G0RQS7
B	75	ILE	HIS	engineered mutation	UNP G0RQS7
C	75	ILE	HIS	engineered mutation	UNP G0RQS7
D	75	ILE	HIS	engineered mutation	UNP G0RQS7
E	75	ILE	HIS	engineered mutation	UNP G0RQS7
F	75	ILE	HIS	engineered mutation	UNP G0RQS7

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Fe	0	0
			1	1		
2	E	1	Total	Fe	0	0
			1	1		
2	B	1	Total	Fe	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total Fe 1 1	0	0
2	A	1	Total Fe 1 1	0	0
2	F	1	Total Fe 1 1	0	0

- Molecule 3 is water.

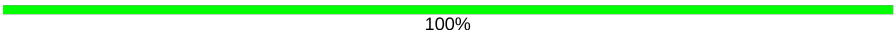
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	182	Total O 182 182	0	0
3	B	209	Total O 209 209	0	0
3	C	222	Total O 222 222	0	0
3	D	175	Total O 175 175	0	0
3	E	200	Total O 200 200	0	0
3	F	169	Total O 169 169	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

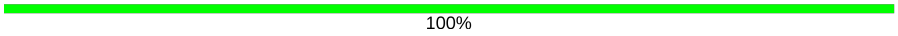
- Molecule 1: Superoxide dismutase

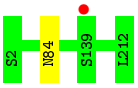
Chain A:



- Molecule 1: Superoxide dismutase

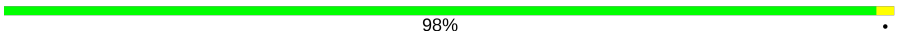
Chain B:





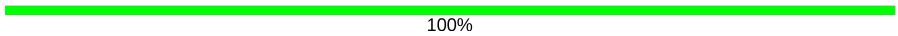
- Molecule 1: Superoxide dismutase

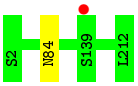
Chain C:



- Molecule 1: Superoxide dismutase

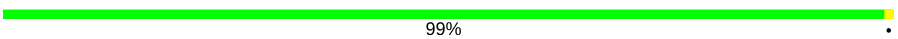
Chain D:

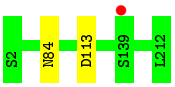




- Molecule 1: Superoxide dismutase

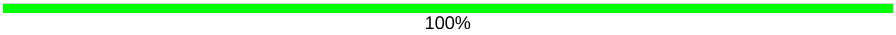
Chain E:

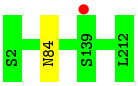




- Molecule 1: Superoxide dismutase

Chain F:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.43Å 75.13Å 241.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.09 – 2.00 54.95 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (47.09-2.00) 99.8 (54.95-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	4.83 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.159 , 0.189 0.159 , 0.189	Depositor DCC
R_{free} test set	4649 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	22.5	Xtriage
Anisotropy	0.309	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 48.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	20544	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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.*

[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: FE

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	0.26	0/1676	0.44	0/2281
1	B	0.26	0/1676	0.44	0/2281
1	C	0.27	0/1676	0.45	0/2281
1	D	0.27	0/1676	0.43	0/2281
1	E	0.26	0/1676	0.42	0/2281
1	F	0.26	0/1676	0.43	0/2281
All	All	0.26	0/10056	0.43	0/13686

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	A	1632	1599	1599	0	0
1	B	1632	1599	1599	0	0
1	C	1632	1599	1599	1	1
1	D	1632	1595	1599	0	0
1	E	1632	1599	1599	0	1
1	F	1632	1598	1599	0	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	182	0	0	0	1
3	B	209	0	0	0	0
3	C	222	0	0	0	1
3	D	175	0	0	0	0
3	E	200	0	0	0	0
3	F	169	0	0	0	0
All	All	10955	9589	9594	1	2

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 0.

All (1) 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:C:206:LYS:HB3	1:C:209:LYS:HE2	2.02	0.42

All (2) 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)
3:A:511:HOH:O	3:C:548:HOH:O[2_567]	2.14	0.06
1:C:114:LYS:HZ1	1:E:113:ASP:OD2[2_477]	1.58	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	209/211 (99%)	203 (97%)	6 (3%)	0	100	100
1	B	209/211 (99%)	201 (96%)	8 (4%)	0	100	100
1	C	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
1	D	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
1	E	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
1	F	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
All	All	1254/1266 (99%)	1212 (97%)	42 (3%)	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	169/169 (100%)	168 (99%)	1 (1%)	86	90
1	B	169/169 (100%)	168 (99%)	1 (1%)	86	90
1	C	169/169 (100%)	168 (99%)	1 (1%)	86	90
1	D	169/169 (100%)	168 (99%)	1 (1%)	86	90
1	E	169/169 (100%)	168 (99%)	1 (1%)	86	90
1	F	169/169 (100%)	168 (99%)	1 (1%)	86	90
All	All	1014/1014 (100%)	1008 (99%)	6 (1%)	86	90

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	84	ASN
1	B	84	ASN
1	C	84	ASN
1	D	84	ASN
1	E	84	ASN
1	F	84	ASN

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

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 carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 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

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

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	A	211/211 (100%)	-0.33	0 100 100	15, 27, 41, 51	0
1	B	211/211 (100%)	-0.30	1 (0%) 91 90	12, 22, 43, 73	0
1	C	211/211 (100%)	-0.34	0 100 100	12, 21, 39, 61	0
1	D	211/211 (100%)	-0.20	1 (0%) 91 90	14, 29, 50, 66	0
1	E	211/211 (100%)	-0.37	1 (0%) 91 90	15, 25, 38, 53	0
1	F	211/211 (100%)	-0.30	1 (0%) 91 90	17, 28, 42, 68	0
All	All	1266/1266 (100%)	-0.31	4 (0%) 94 93	12, 26, 43, 73	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	139	SER	5.5
1	D	139	SER	2.2
1	E	139	SER	2.1
1	F	139	SER	2.0

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 carbohydrates 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,

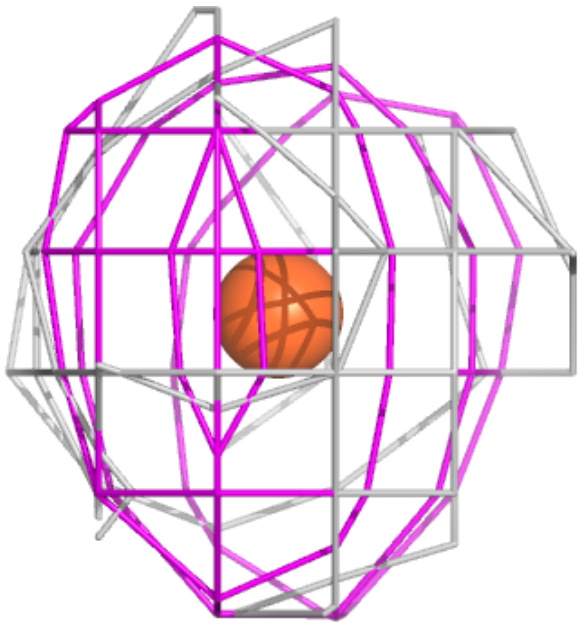
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

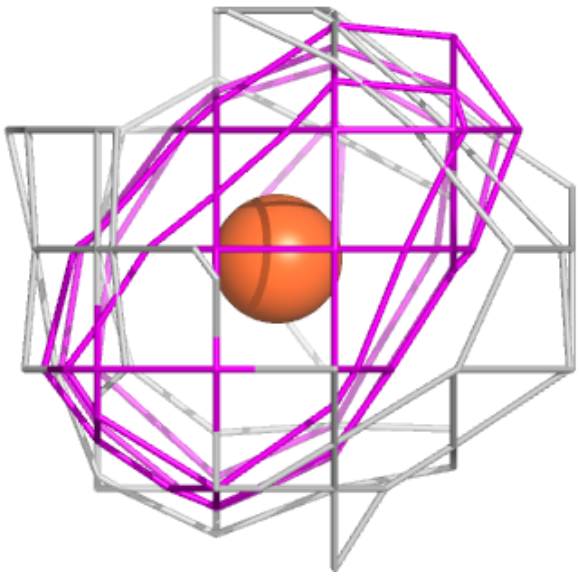
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	FE	F	301	1/1	0.97	0.11	29,29,29,29	0
2	FE	A	301	1/1	0.98	0.04	29,29,29,29	0
2	FE	C	301	1/1	0.99	0.04	24,24,24,24	0
2	FE	E	301	1/1	0.99	0.03	27,27,27,27	0
2	FE	D	301	1/1	0.99	0.09	27,27,27,27	0
2	FE	B	301	1/1	1.00	0.04	27,27,27,27	0

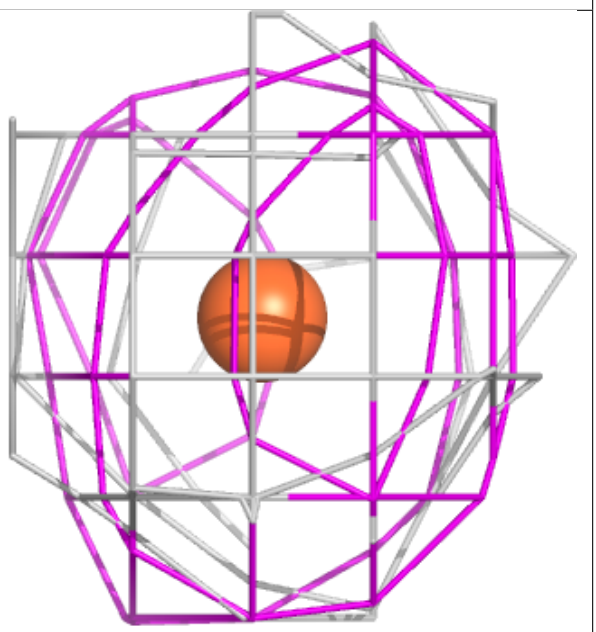
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around FE F 301:

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

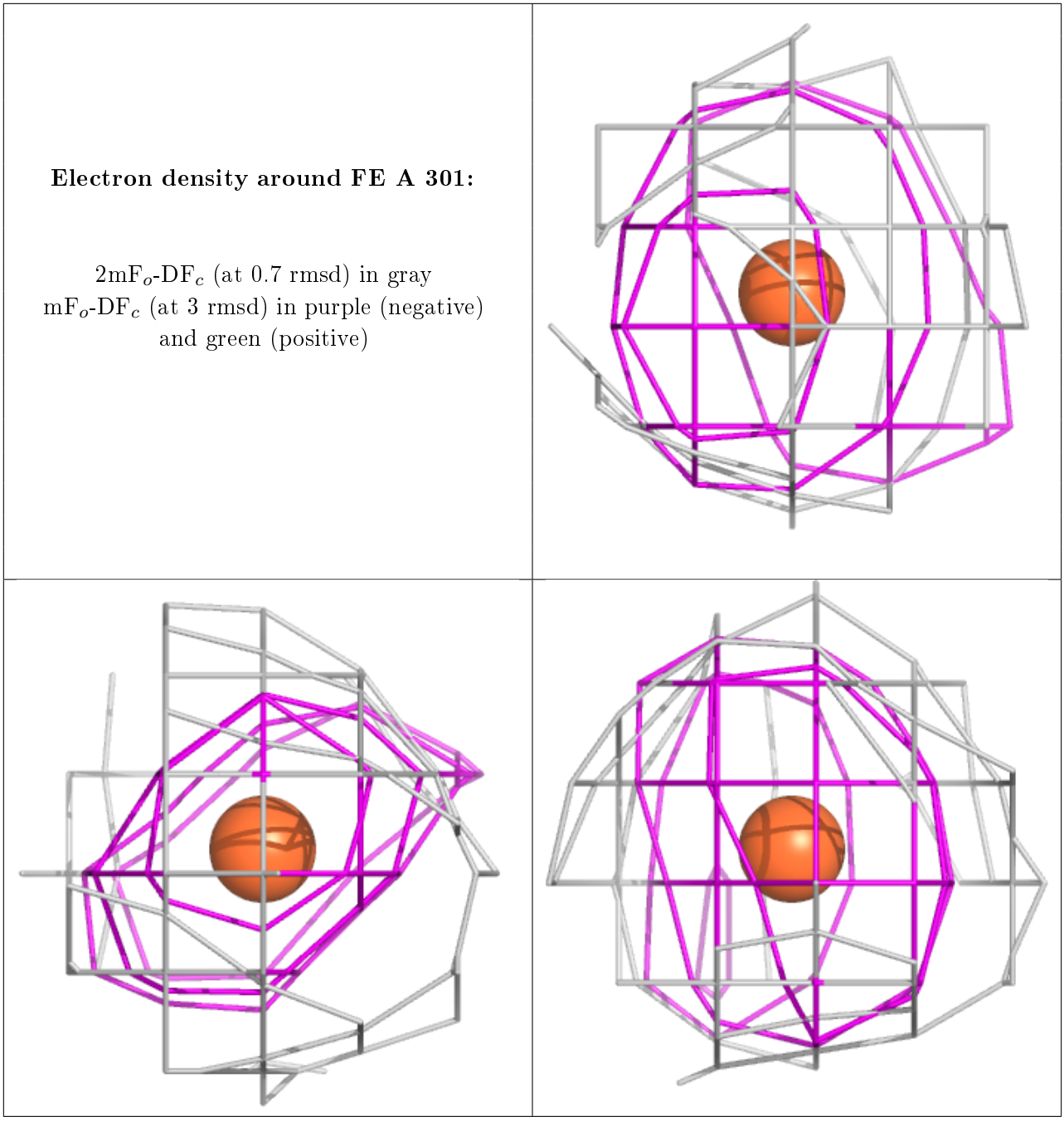






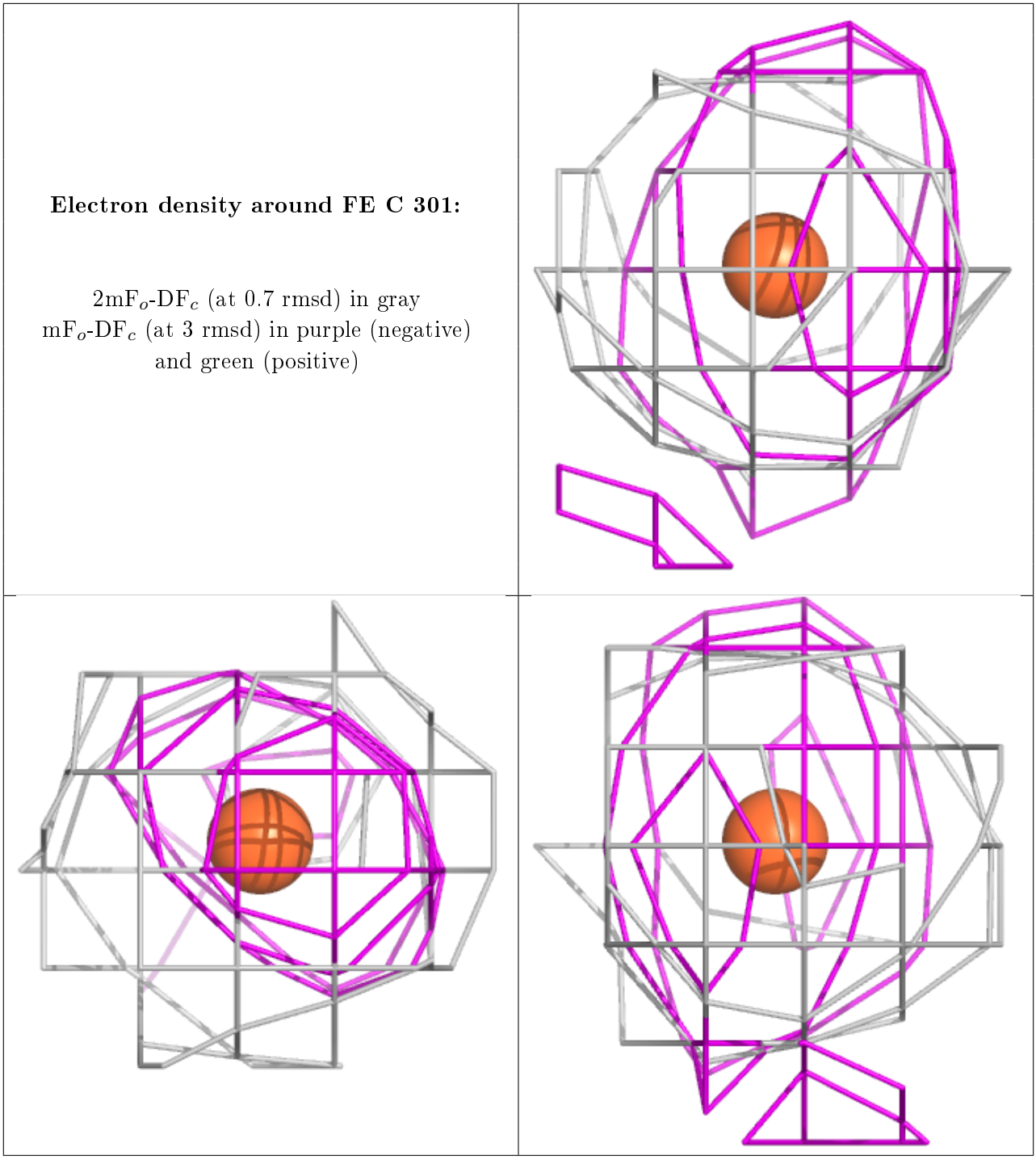
Electron density around FE A 301:

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



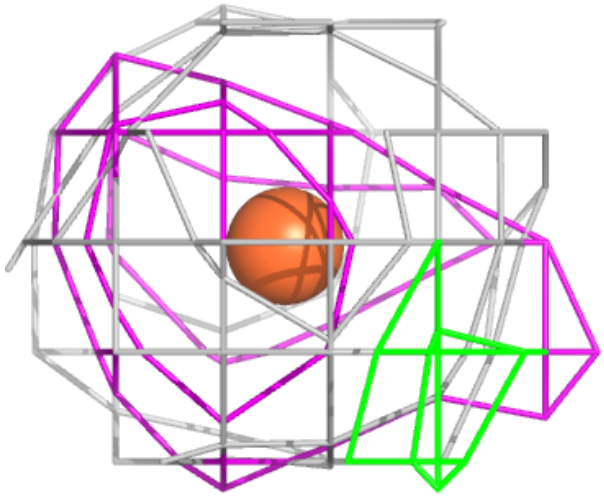
Electron density around FE C 301:

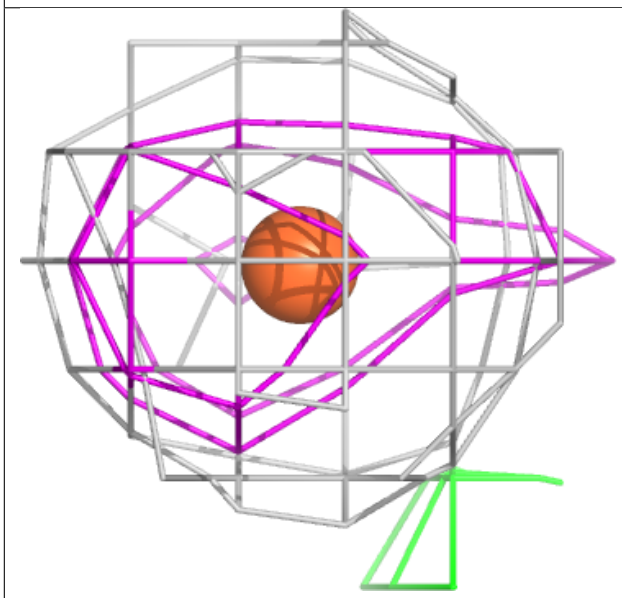
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

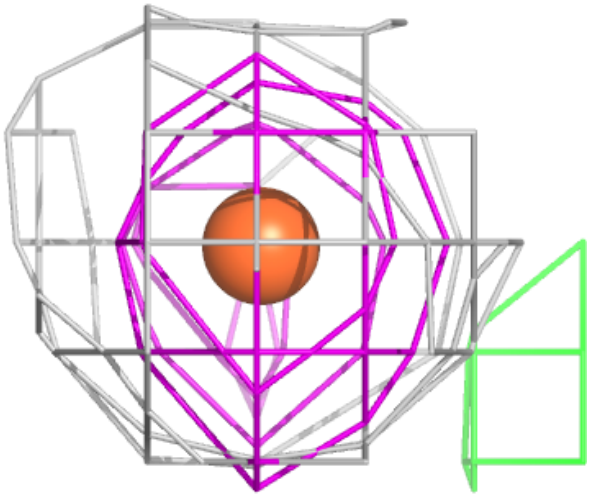


Electron density around FE E 301:

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

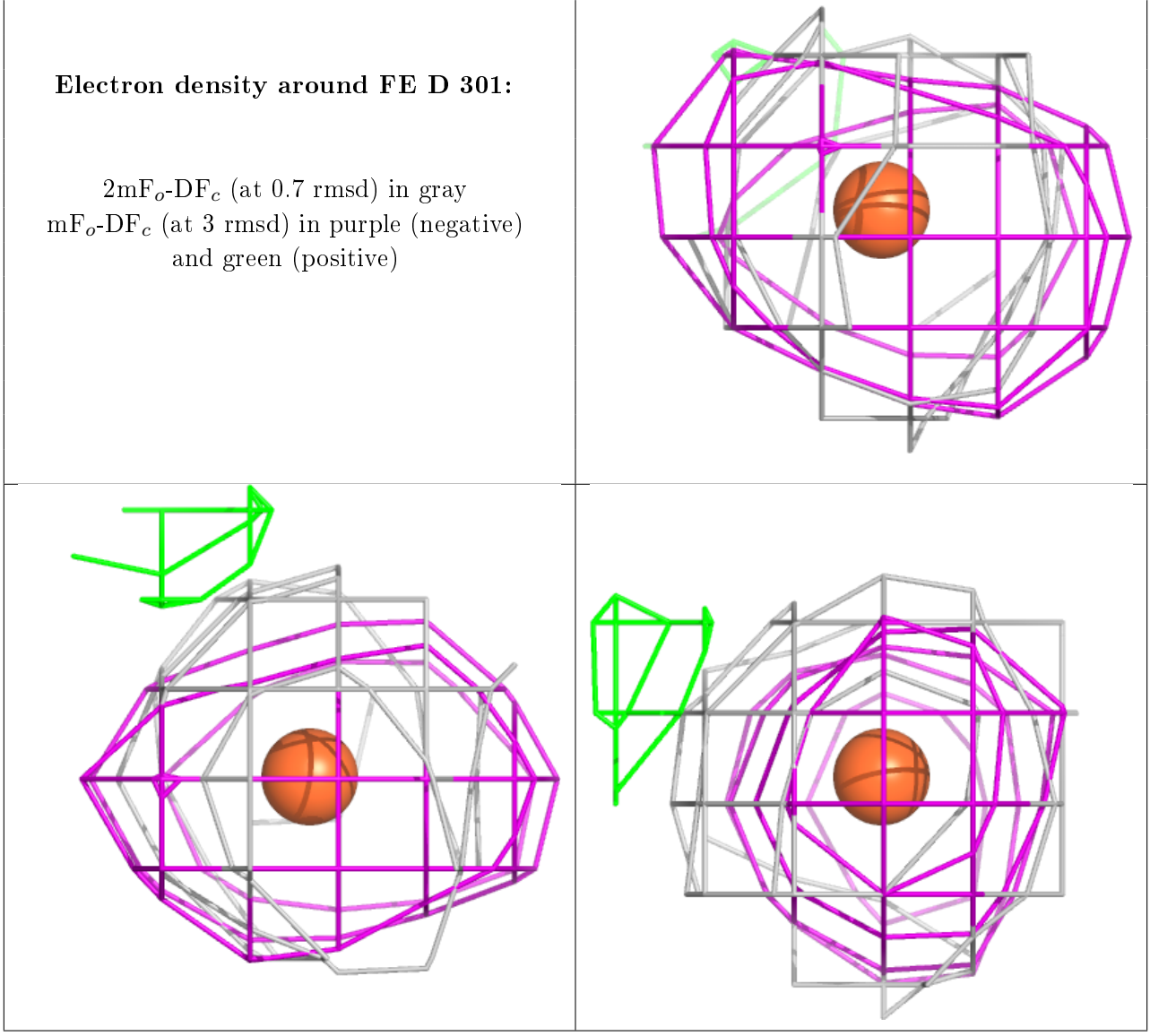






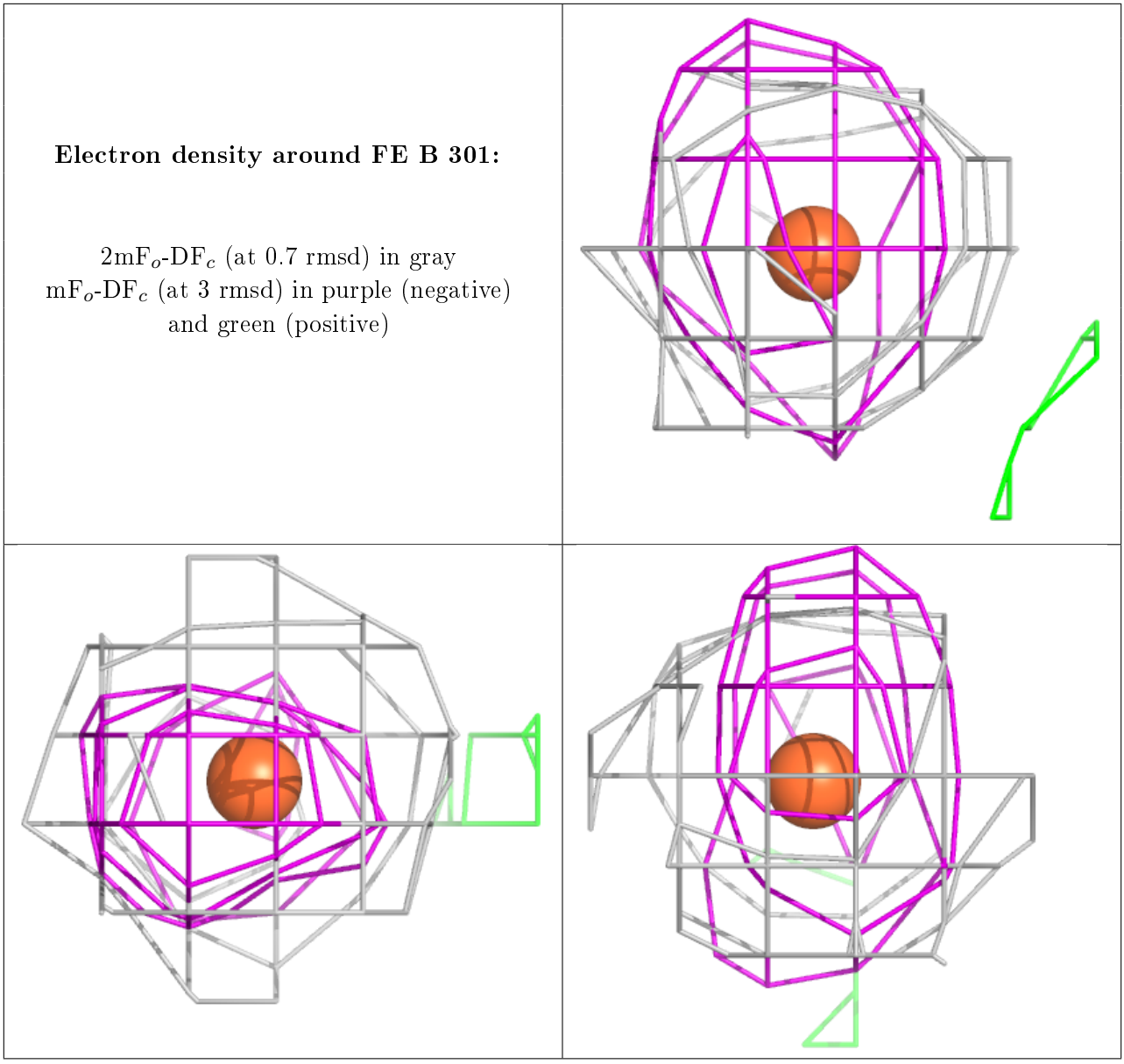
Electron density around FE D 301:

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



Electron density around FE B 301:

$2mF_o$-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.