

Full wwPDB X-ray Structure Validation Report (i)

Apr 18, 2021 - 05:36 PM JST

PDB ID	:	7C4C
Title	:	The crystal structure of Trypanosoma brucei RNase D : GMP complex
Authors	:	Gao, Y.Q.; Gan, J.H.
Deposited on		
Resolution	:	2.27 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

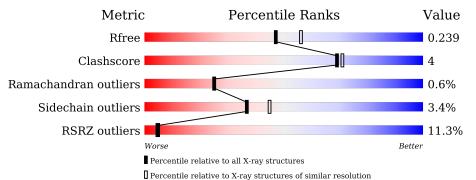
Xtriage (Phenix) EDS buster-report Percentile statistics Refmac CCP4 Ideal geometry (proteins)	:::::::::::::::::::::::::::::::::::::::	 1.8.5 (274361), CSD as541be (2020) 1.13 2.18 1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove) Engh & Huber (2001)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)	:	3

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.27 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	Δ	351	11%	10%	5%
1	Л	001	84%	10%	5%



7C4C

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2655 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 CCHC-type domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
1	А	332	Total 2524	C 1593	N 451	O 460	S 16	${ m Se} \over 4$	0	0	0

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

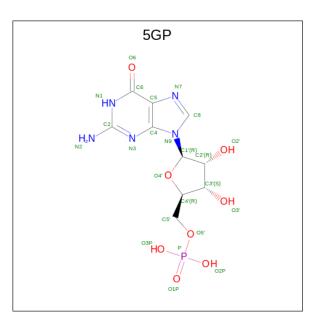
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	6	Total Zn 6 6	0	0

• Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	6	Total Mn 6 6	0	0

• Molecule 4 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula: $C_{10}H_{14}N_5O_8P$) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total C N O P 24 10 5 8 1	0	0
4	А	1	Total C N O 18 9 5 4	0	0
4	А	1	Total C N O P 24 10 5 8 1	0	0
4	A	1	Total C N O 12 6 5 1	0	0

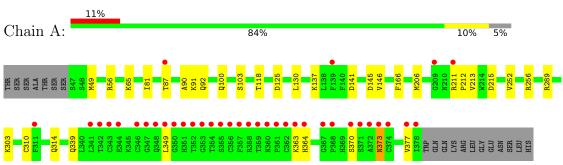
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	41	Total O 41 41	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: CCHC-type domain-containing protein



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	57.17Å 75.00Å 101.41Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 - 2.27	Depositor
Resolution (A)	30.15 - 2.26	EDS
% Data completeness	86.0 (30.00-2.27)	Depositor
(in resolution range)	86.0 (30.15-2.26)	EDS
R _{merge}	0.09	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$3.81 (at 2.26 \text{\AA})$	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
D D.	0.192 , 0.239	Depositor
R, R_{free}	0.194 , 0.239	DCC
R_{free} test set	918 reflections (5.11%)	wwPDB-VP
Wilson B-factor $(Å^2)$	26.9	Xtriage
Anisotropy	0.090	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.36 , 44.6	EDS
L-test for twinning ²	$ < L >=0.47, < L^2>=0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2655	wwPDB-VP
Average B, all atoms $(Å^2)$	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.67% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: 5GP, MN, 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).

Mal	Chain	Bond	lengths	Bond angles		
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.43	0/2573	0.62	0/3477	

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	А	2524	0	2479	19	0
2	А	6	0	0	0	0
3	А	6	0	0	0	0
4	А	78	0	37	1	0
5	А	41	0	0	2	0
All	All	2655	0	2516	19	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 4.

All (19) 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:211:ARG:HD2	1:A:212:PRO:HD2	1.65	0.77
1:A:145:ASP:OD2	5:A:501:HOH:O	2.11	0.68
1:A:211:ARG:HH21	1:A:213:VAL:H	1.40	0.66
1:A:364:HIS:CE1	1:A:377:VAL:HG11	2.31	0.65
1:A:211:ARG:NH2	1:A:215:ASP:OD1	2.28	0.65
1:A:211:ARG:HD2	1:A:212:PRO:CD	2.29	0.62
1:A:49:MSE:HA	1:A:49:MSE:HE3	1.85	0.59
1:A:211:ARG:NH2	1:A:213:VAL:H	2.05	0.54
1:A:289:ARG:HD2	1:A:289:ARG:O	2.11	0.49
1:A:90:ALA:HB1	1:A:92:GLN:OE1	2.13	0.49
1:A:65:LYS:NZ	1:A:125:ASP:OD2	2.46	0.48
1:A:363:LYS:HE2	1:A:364:HIS:NE2	2.31	0.45
1:A:303:LYS:NZ	5:A:507:HOH:O	2.48	0.45
1:A:339:GLN:HA	1:A:349:LEU:HD23	1.99	0.44
1:A:370:SER:OG	1:A:373:ASN:OD1	2.38	0.42
1:A:130:LEU:HA	1:A:137:LYS:HE2	2.00	0.42
1:A:252:VAL:O	1:A:256:ARG:HG3	2.20	0.41
1:A:310:CYS:O	1:A:314:GLN:HG3	2.20	0.41
1:A:81:ILE:O	4:A:413:5GP:H5'2	2.22	0.40

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	А	330/351~(94%)	315~(96%)	13~(4%)	2(1%)	25 25

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	141	ASP
1	А	118	THR



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	А	268/293~(92%)	259~(97%)	9~(3%)	37 45	

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	56	ARG
1	А	87	THR
1	А	91	LYS
1	А	100	GLN
1	А	103	SER
1	А	146	VAL
1	А	166	PHE
1	А	206	MSE
1	А	373	ASN

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 16 ligands modelled in this entry, 12 are monoatomic - leaving 4 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	Turne	Chain	Dec	Res Link Bond lengths				Bond angles		
10101	Type	Unam	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	5GP	А	415	-	22,26,26	0.60	0	27,40,40	1.85	5 (18%)
4	5GP	А	416	-	10,13,26	1.43	2 (20%)	10,19,40	2.85	6 (60%)
4	5GP	А	414	3	16,20,26	0.80	0	17,30,40	2.48	9 (52%)
4	5GP	А	413	3	22,26,26	0.61	0	27,40,40	1.70	4 (14%)

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
4	5GP	А	415	-	-	3/6/26/26	0/3/3/3
4	5GP	А	416	-	-	-	0/2/2/3
4	5GP	А	414	3	-	0/0/17/26	0/3/3/3
4	5GP	А	413	3	-	0/6/26/26	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	А	416	5GP	C6-C5	3.50	1.47	1.41
4	А	416	5GP	C5-C4	2.03	1.46	1.40

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
4	А	415	5GP	N3-C2-N1	-5.24	120.23	127.22
4	А	414	5GP	N3-C2-N1	-5.08	120.44	127.22
4	А	415	5GP	C2-N3-C4	4.92	120.98	115.36
4	А	413	5GP	N3-C2-N1	-4.82	120.79	127.22
4	А	416	5GP	C2-N3-C4	4.57	120.58	115.36

Continued on next page...



Mol	Chain	Res	Type	Atoms		$Observed(^{o})$	$Ideal(^{o})$
4	А	414	5GP	C2-N3-C4	4.24	120.20	115.36
4	А	416	5GP	C6-N1-C2	4.04	122.35	115.93
4	А	416	5GP	C5-C6-N1	-3.81	118.22	123.43
4	А	413	5GP	C2-N3-C4	3.66	119.54	115.36
4	А	416	5GP	C6-C5-C4	-3.61	117.35	120.80
4	А	414	5GP	C6-N1-C2	3.61	121.66	115.93
4	А	413	5GP	C6-N1-C2	3.47	121.44	115.93
4	А	415	5GP	C6-N1-C2	3.32	121.20	115.93
4	А	413	5GP	C5-C6-N1	-3.19	119.07	123.43
4	А	414	5GP	C5-C6-N1	-3.07	119.23	123.43
4	А	414	5GP	C4'-C3'-C2'	2.94	106.11	101.63
4	А	416	5GP	N3-C2-N1	-2.84	123.44	127.22
4	А	414	5GP	O3'-C3'-C2'	-2.80	105.97	111.27
4	А	415	5GP	C5-C6-N1	-2.78	119.62	123.43
4	А	416	5GP	C4-C5-N7	-2.52	106.78	109.40
4	А	415	5GP	N2-C2-N1	2.23	120.71	117.25
4	А	414	5GP	C4-C5-N7	-2.22	107.09	109.40
4	А	414	5GP	C6-C5-C4	-2.19	118.70	120.80
4	А	414	5GP	C3'-C2'-C1'	2.05	104.10	101.04

Continued from previous page...

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	А	415	5GP	C3'-C4'-C5'-O5'
4	А	415	5GP	O4'-C4'-C5'-O5'
4	А	415	5GP	C5'-O5'-P-O3P

There are no ring outliers.

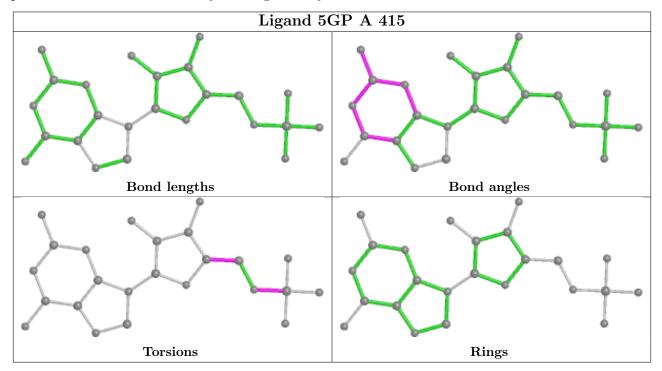
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	А	413	5GP	1	0

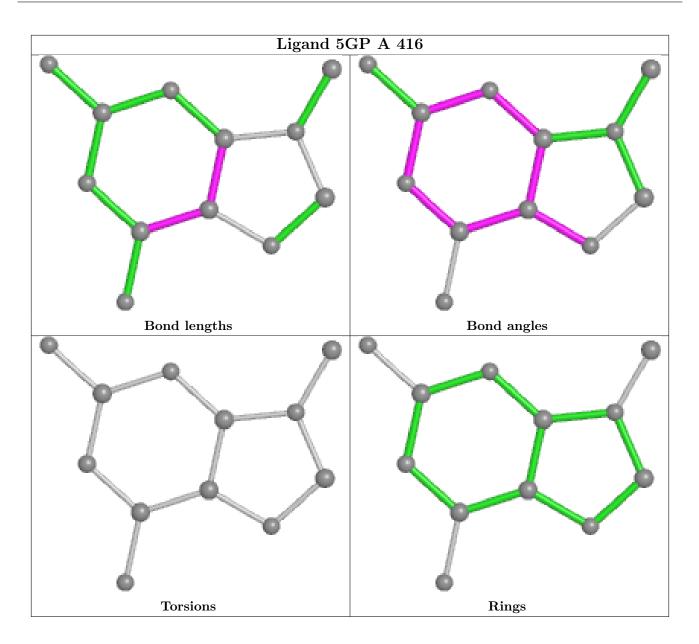
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the



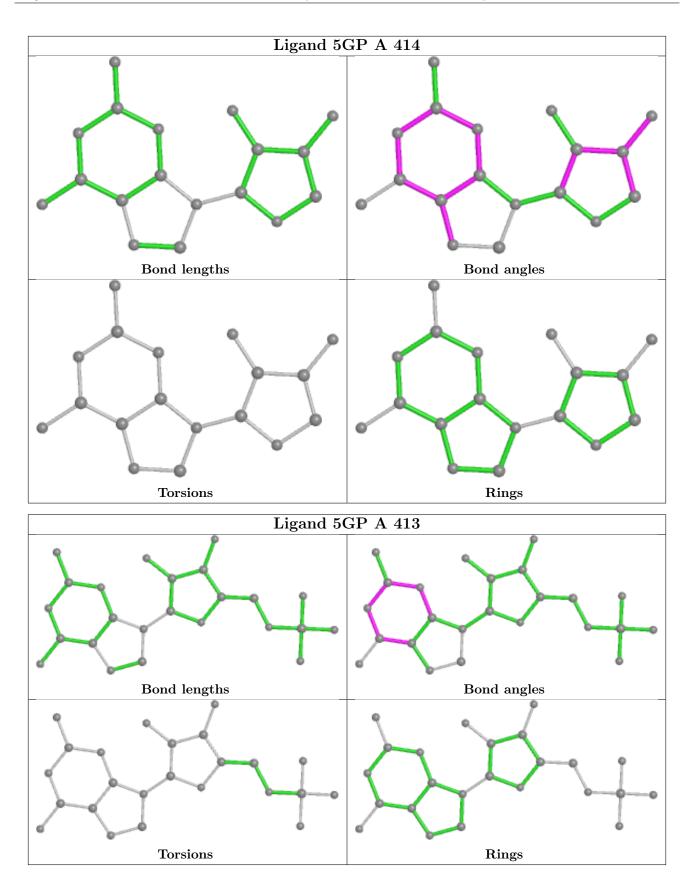
average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.













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 (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, 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	А	328/351~(93%)	0.20	37 (11%) 5 4	18, 34, 91, 146	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	347	GLY	7.3
1	А	349	LEU	6.3
1	А	87	THR	5.6
1	А	362	CYS	5.4
1	А	350	GLY	5.2
1	А	353	GLY	5.1
1	А	361	PRO	5.0
1	А	351	HIS	5.0
1	А	352	THR	4.8
1	А	357	PHE	4.5
1	А	346	CYS	4.3
1	А	348	GLN	4.3
1	А	372	ALA	4.2
1	А	360	ASN	4.0
1	А	355	SER	4.0
1	А	342	THR	4.0
1	А	378	ILE	3.8
1	А	377	VAL	3.6
1	А	364	HIS	3.5
1	А	354	THR	3.5
1	А	209	GLY	3.4
1	А	211	ARG	3.4
1	А	311	PHE	3.3
1	А	368	PRO	3.1
1	А	367	GLY	2.9
1	А	363	LYS	2.8
1	A	371	SER	2.8

Continued on next page...



Mol	Chain	Res	Type	RSRZ
1	А	358	ARG	2.8
1	А	370	SER	2.7
1	А	373	ASN	2.7
1	А	341	LEU	2.7
1	А	359	THR	2.5
1	А	344	GLU	2.4
1	А	343	CYS	2.2
1	А	139	PHE	2.2
1	А	356	CYS	2.2
1	А	374	CYS	2.1

Continued from previous page...

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 monosaccharides 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^{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	$B-factors(A^2)$	Q<0.9
4	5GP	А	415	24/24	0.69	0.34	80,92,111,114	0
2	ZN	А	406	1/1	0.82	0.06	125,125,125,125	0
3	MN	А	411	1/1	0.84	0.10	85,85,85,85	0
4	5GP	А	416	12/24	0.85	0.28	69,73,77,78	0
3	MN	А	412	1/1	0.90	0.11	120,120,120,120	0
4	5GP	А	414	18/24	0.90	0.15	58,72,79,79	0
4	5GP	А	413	24/24	0.92	0.13	$37,\!48,\!59,\!62$	0
3	MN	А	410	1/1	0.93	0.11	91,91,91,91	0
2	ZN	А	403	1/1	0.95	0.04	78,78,78,78	0
2	ZN	А	405	1/1	0.96	0.05	64,64,64,64	0
3	MN	А	409	1/1	0.97	0.06	$53,\!53,\!53,\!53$	0
3	MN	А	407	1/1	0.99	0.10	39,39,39,39	0
3	MN	А	408	1/1	0.99	0.16	41,41,41,41	0

Continued on next page...



2

ΖN

Q<0.9 0 0

0

Conti							
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(Å^2)$
2	ZN	А	402	1/1	1.00	0.07	$26,\!26,\!26,\!26$
2	ZN	А	401	1/1	1.00	0.08	34,34,34,34

1/1

404

Continued from previous page...

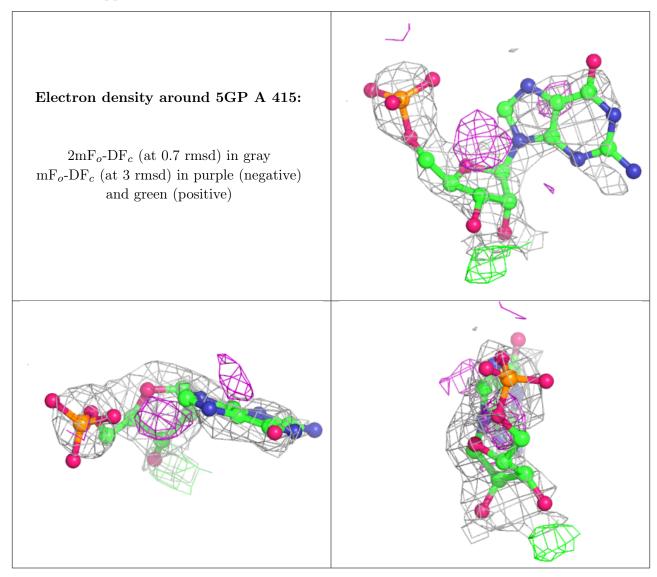
А

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.

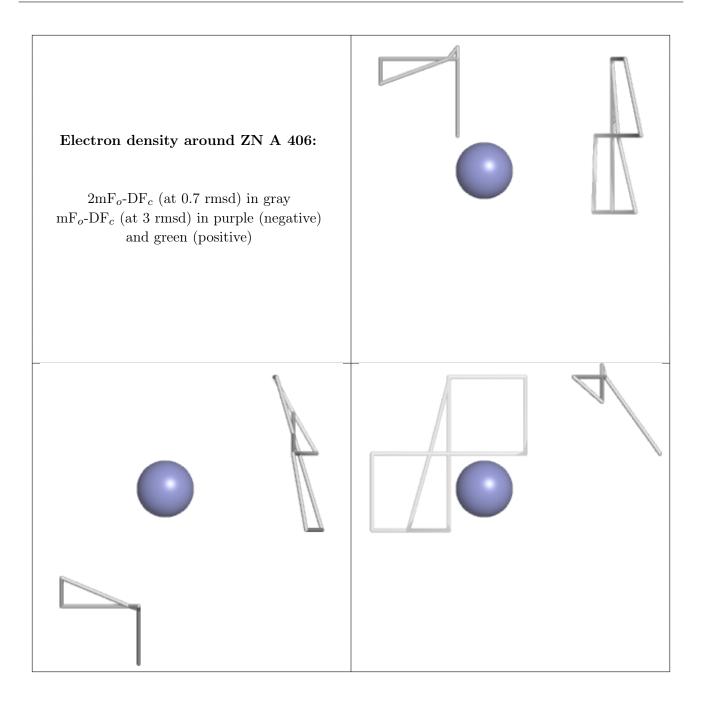
0.05

44,44,44,44

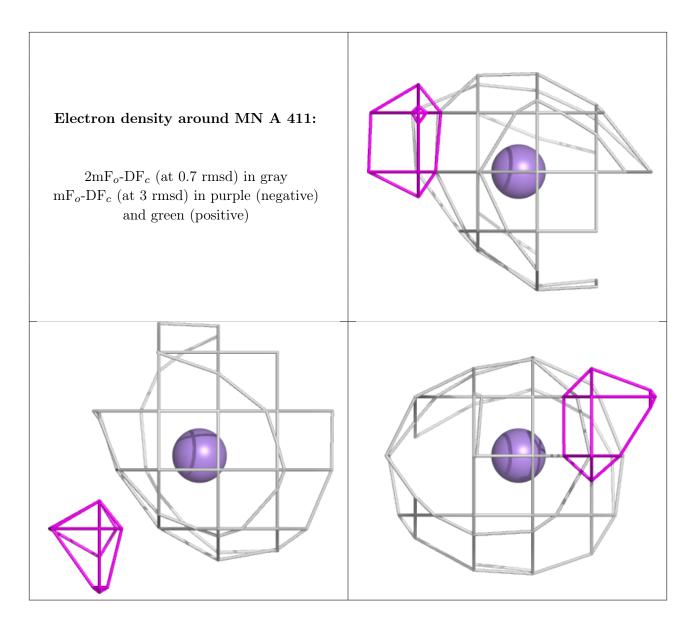
1.00



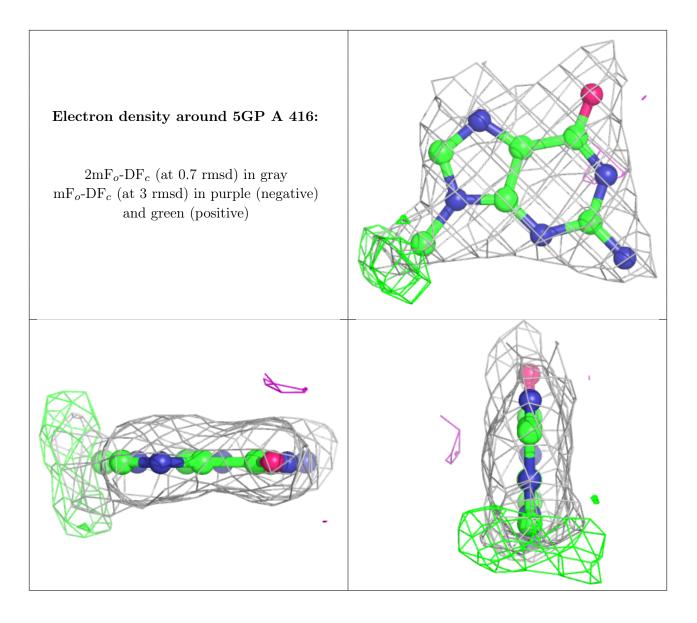




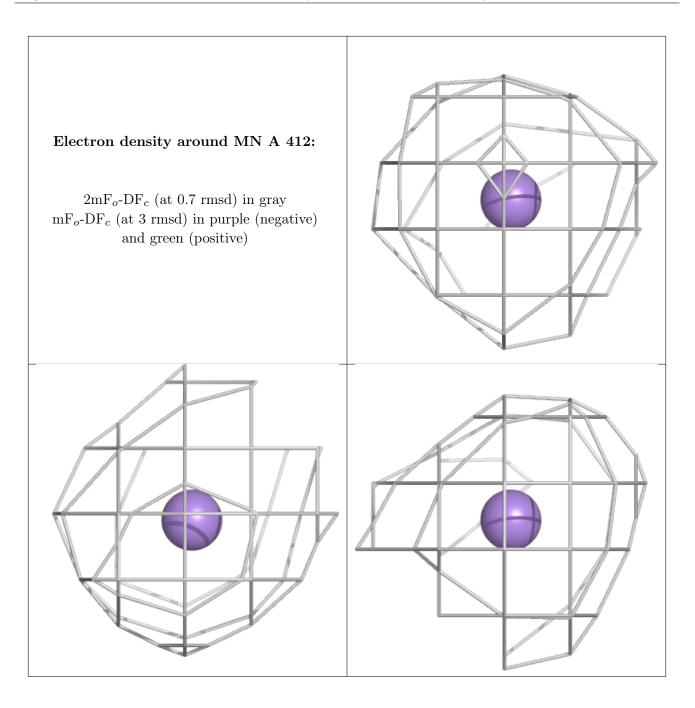




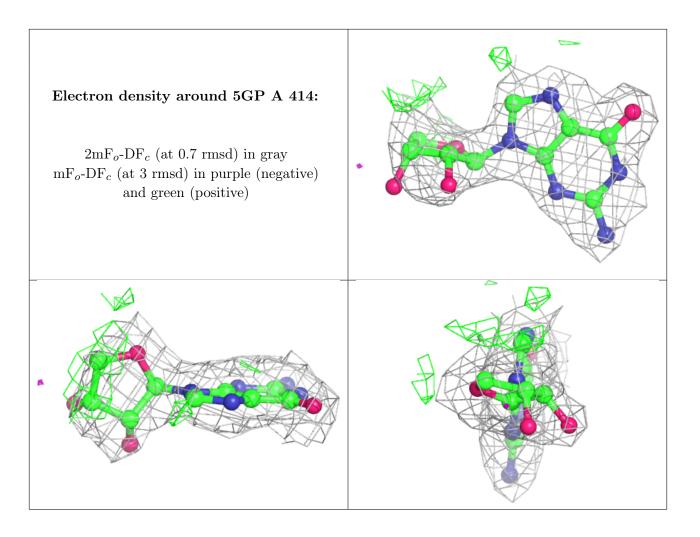




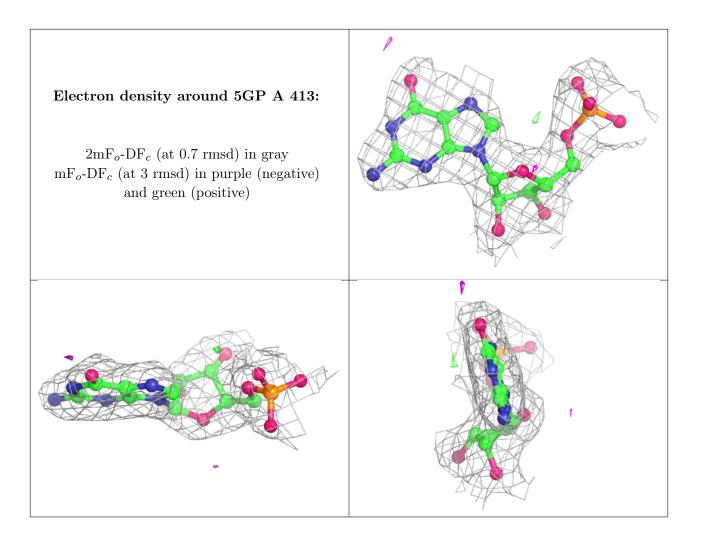




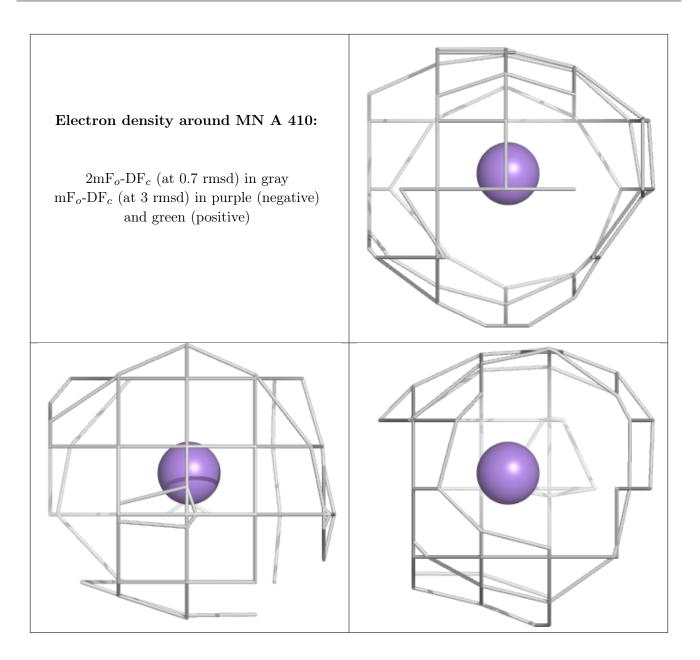




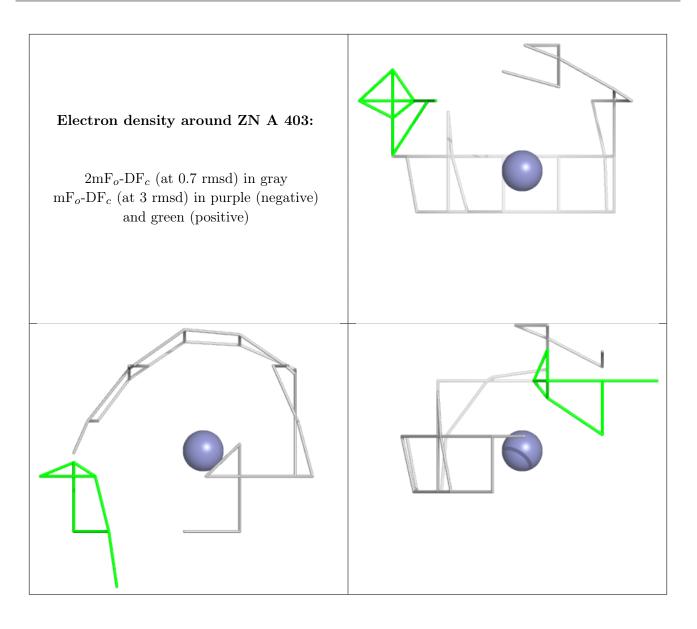




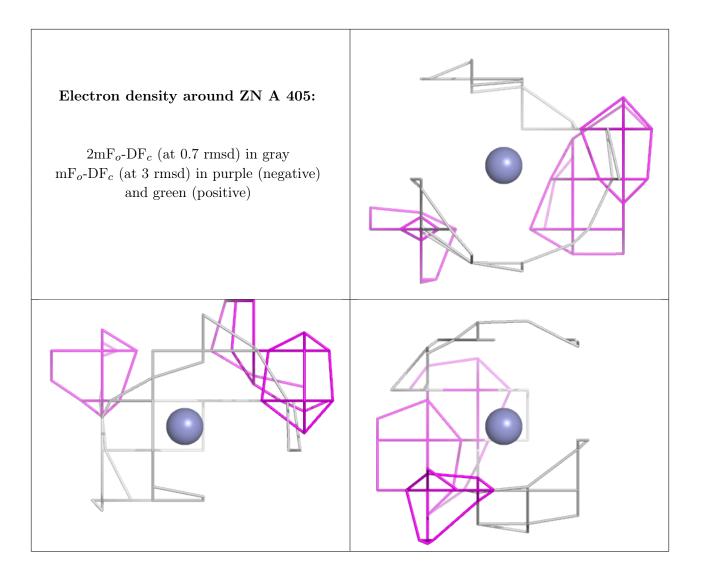




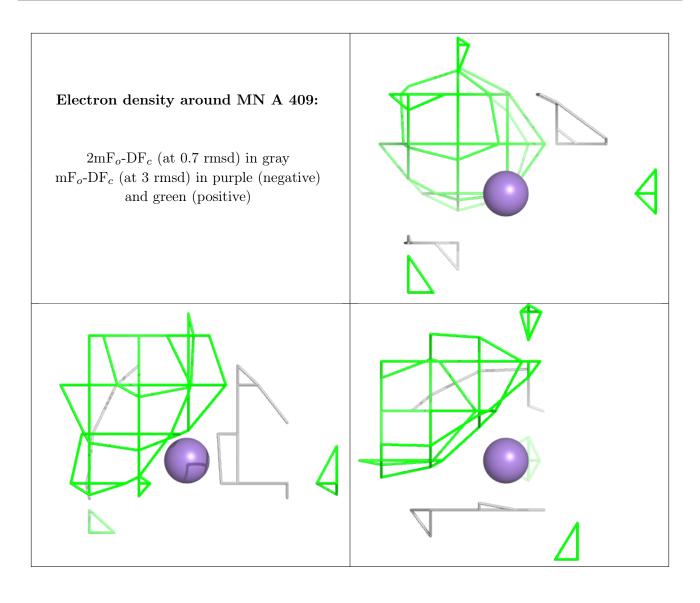




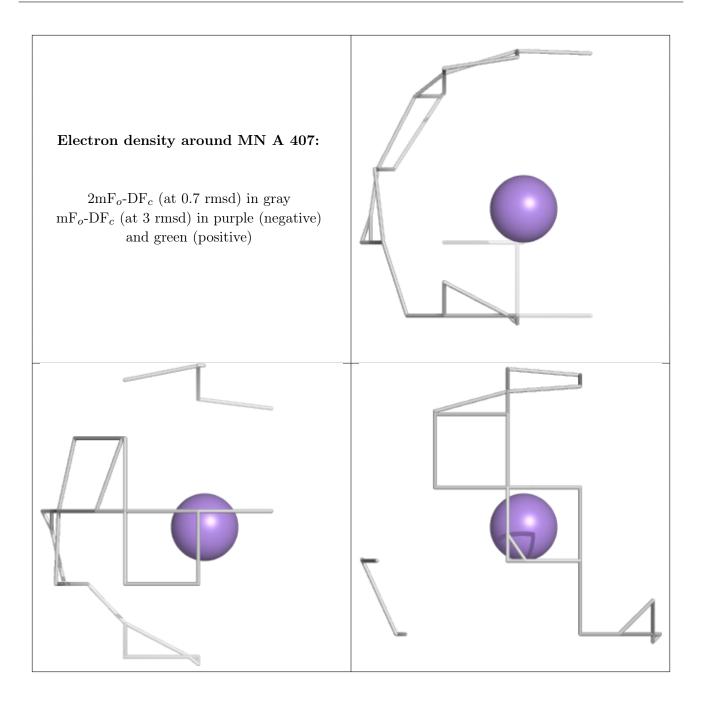




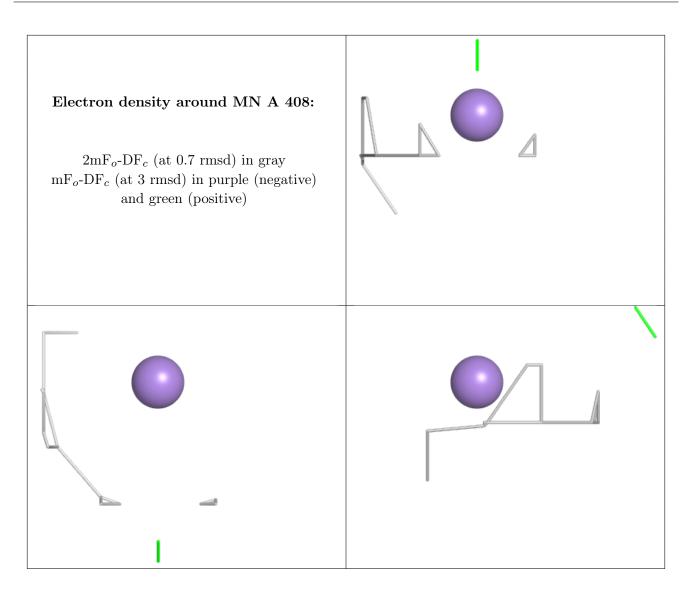




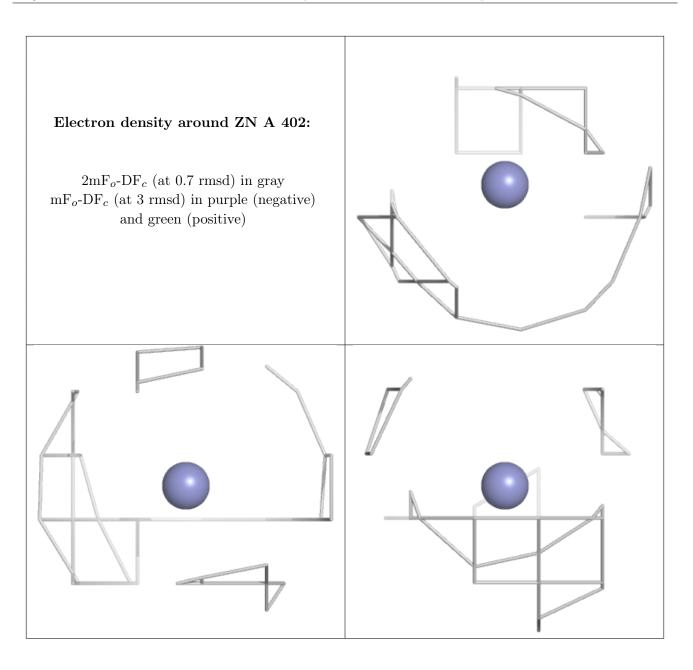




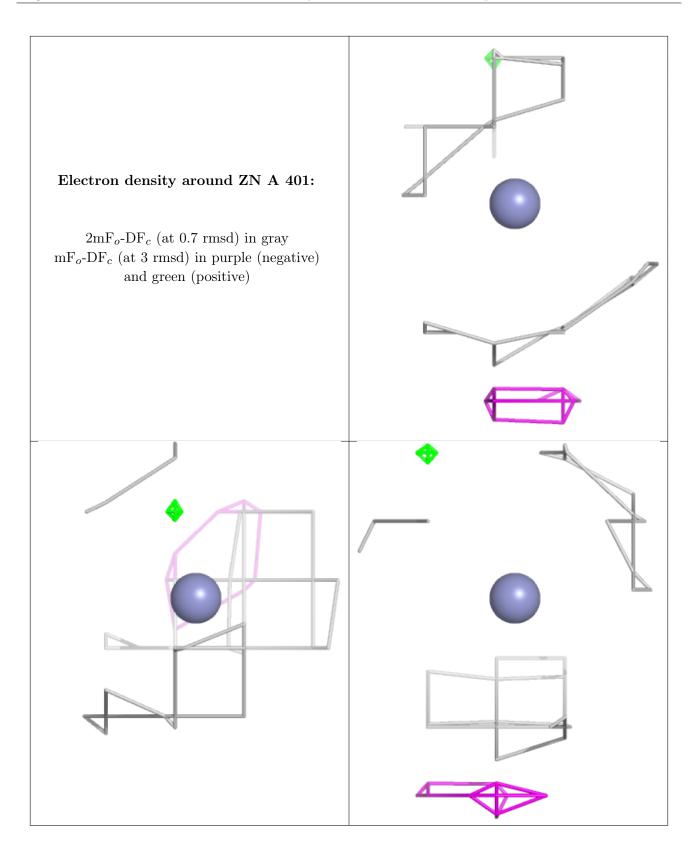




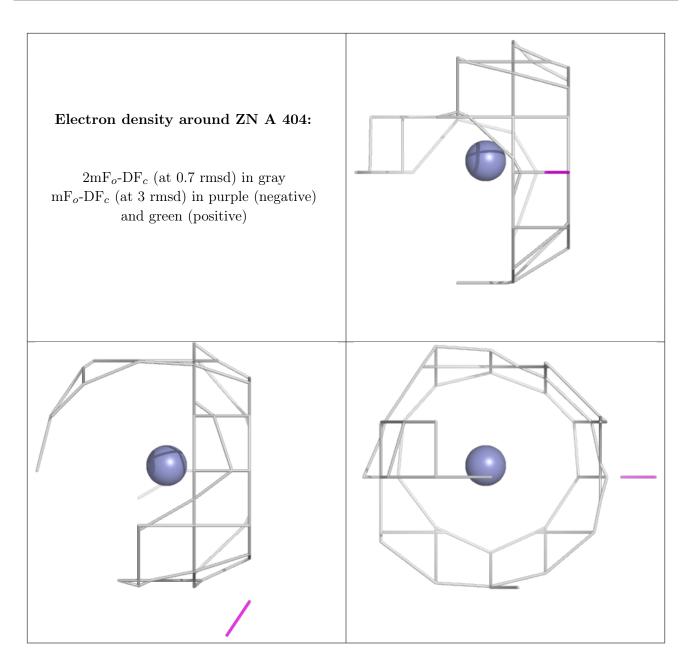












6.5 Other polymers (i)

There are no such residues in this entry.

