WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 22, 2024 – 06:54 PM JST

PDB ID	:	8I3Y
Title	:	Crystal structure of ASCT from Trypanosoma brucei in complex with Succinyl-CoA.
Authors	:	Mochizuki, K.; Inaoka, D.K.; Fukuda, K.; Kurasawa, H.; Iyoda, K.; Nakai, U.; Harada, S.; Balogun, E.O.; Mazet, M.; Millerioux, Y.; Bringaud, F.; Boshart, M.; Hirayama, K.; Kita, K.; Shiba, T.
Deposited on	:	2023-01-18
Resolution	:	2.40 Å(reported)

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

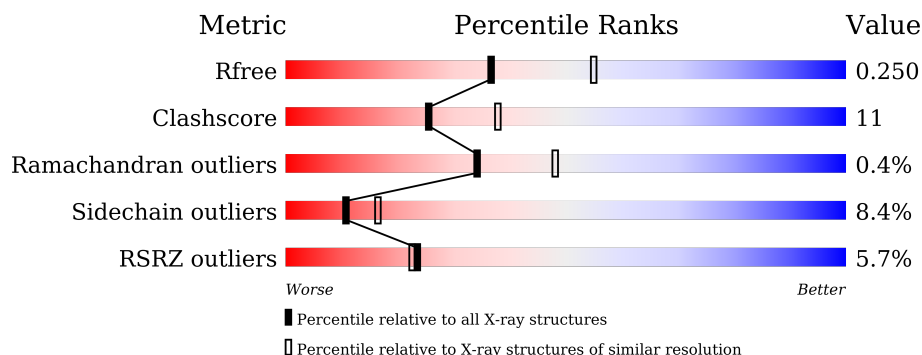|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	Percentile Ranks	Value
Rfree		0.250
Clashscore		11
Ramachandran outliers		0.4%
Sidechain outliers		8.4%
RSRZ outliers		5.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	A	483	13% 67% 22% 5% · 5%
1	B	483	86% 9% · ·
1	C	483	8% 80% 13% · ·
1	D	483	% 85% 11% · ·

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 14225 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 Succinyl-CoA:3-ketoacid-coenzyme A transferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	459	Total	C	N	O	S	0	0	0
			3454	2173	603	655	23			
1	B	470	Total	C	N	O	S	0	0	0
			3539	2222	621	673	23			
1	C	464	Total	C	N	O	S	0	0	0
			3493	2197	612	661	23			
1	D	471	Total	C	N	O	S	0	0	0
			3544	2225	622	674	23			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	419	ASN	SER	conflict	UNP Q386P1
B	419	ASN	SER	conflict	UNP Q386P1
C	419	ASN	SER	conflict	UNP Q386P1
D	419	ASN	SER	conflict	UNP Q386P1

- Molecule 2 is SUCCINYL-COENZYME A (three-letter code: SCA) (formula: $C_{25}H_{40}N_7O_{19}P_3S$) (labeled as "Ligand of Interest" by depositor).
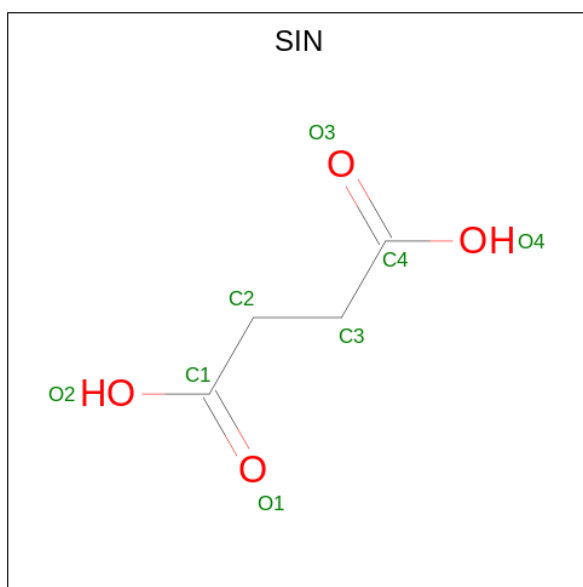
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
2	A	1	55	25	7	19	3	1	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 4 is SUCCINIC ACID (three-letter code: SIN) (formula: $C_4H_6O_4$) (labeled as "Ligand of Interest" by depositor).

SIN

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 8	C 4	O 4	0	0
4	D	1	Total 8	C 4	O 4	0	0

• Molecule 5 is water.

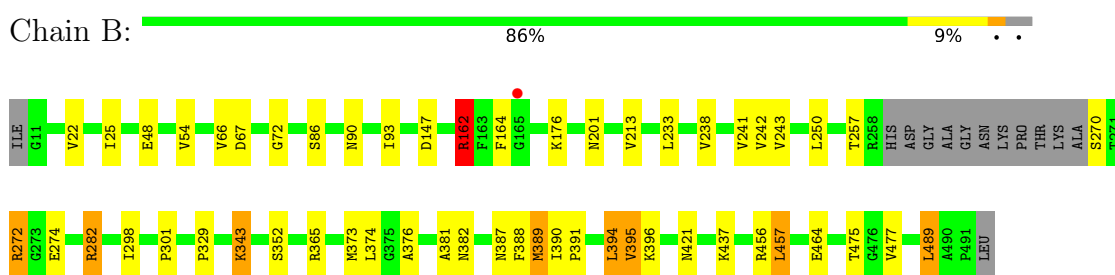| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	28	Total 28	O 28	0	0
5	B	28	Total 28	O 28	0	0
5	C	24	Total 24	O 24	0	0
5	D	40	Total 40	O 40	0	0

WORLD WIDE
PDB
PROTEIN DATA BANK

3 Residue-property plots ⓘ

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.
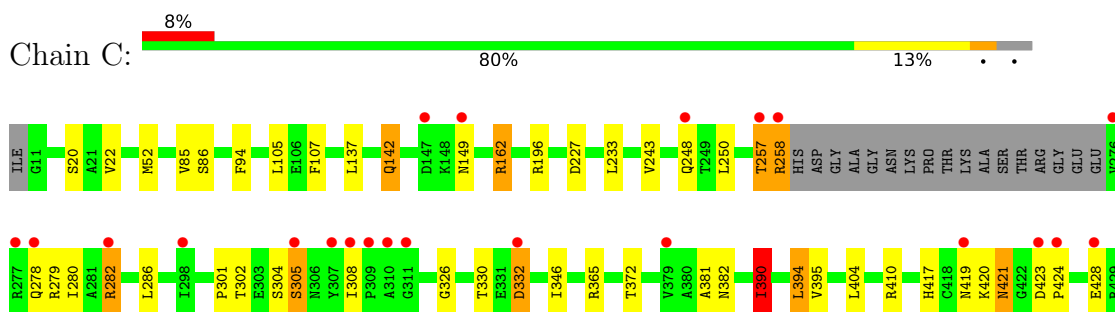
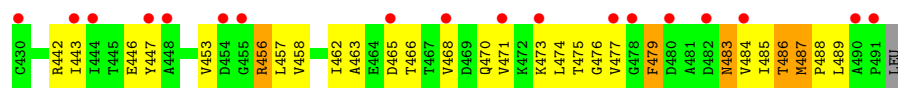• Molecule 1: Succinyl-CoA:3-ketoacid-coenzyme A transferase

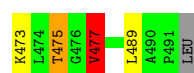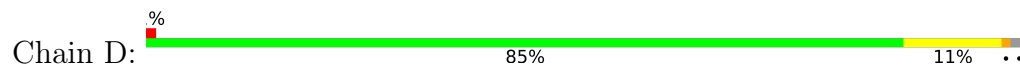• Molecule 1: Succinyl-CoA:3-ketoacid-coenzyme A transferase

• Molecule 1: Succinyl-CoA:3-ketoacid-coenzyme A transferase

C430 R442 I443 I444 T445 E446 Y447 A448 V453 D454 G455 R456 L457 V458 I462 A463 E464 D465 T466 T467 V468 D469 Q470 V471 K472 K473 L474 T475 G476 V477 G478 F479 D480 A481 D482 N483 V484 I485 T486 M487 P488 L489 A490 P491 LEU

• Molecule 1: Succinyl-CoA:3-ketoacid-coenzyme A transferase

Chain D:

85%　　　　11%

ILE G11 V15 L18 S19 I25 M52 G53 V54 V66 K82 V85 N90 N91 K92 P109 V136 L137 Q142 V152 R162 F163 F164 S180 R196 V213 L233 V238 V241 V242 V243 Q248 T249 L250 T255 R258 HIS ASP

GLY ALA GLY ASN LYS PRO THR LYS A269 V276 E287 V294 L316 P329 V334 D337 N340 K343 Q344 R365 A376 L385 I390 P391 L394 V405 S406 C407 G408 T409 R410 K420 E428 K437 R442 T466 Q470 V471 K472

K473 L474 T475 G476 V477 L489 A490 P491 LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	59.80Å 163.65Å 186.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 2.40 19.98 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (19.98-2.40) 99.9 (19.98-2.40)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.91 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.192 , 0.248 0.199 , 0.250	Depositor DCC
R_{free} test set	3585 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	37.8	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 38.2	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14225	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.24% 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: SIN, CA, SCA

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	0.72	0/3510	0.97	12/4749 (0.3%)
1	B	0.67	0/3596	0.90	6/4865 (0.1%)
1	C	0.71	0/3550	0.90	8/4804 (0.2%)
1	D	0.63	0/3601	0.87	6/4872 (0.1%)
All	All	0.68	0/14257	0.91	32/19290 (0.2%)

There are no bond length outliers.

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	162	ARG	NE-CZ-NH1	8.71	124.66	120.30
1	D	365	ARG	NE-CZ-NH2	-8.45	116.08	120.30
1	B	162	ARG	NE-CZ-NH1	8.39	124.50	120.30
1	A	365	ARG	NE-CZ-NH2	-8.01	116.29	120.30
1	C	162	ARG	NE-CZ-NH1	7.38	123.99	120.30

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	3454	0	3470	182	0
1	B	3539	0	3552	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3493	0	3512	58	0
1	D	3544	0	3557	22	0
2	A	55	0	35	14	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	B	8	0	4	2	0
4	D	8	0	4	1	0
5	A	28	0	0	0	0
5	B	28	0	0	0	0
5	C	24	0	0	0	0
5	D	40	0	0	0	0
All	All	14225	0	14134	298	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 11.

The worst 5 of 298 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:442:ARG:HG2	1:A:443:ILE:CD1	1.64	1.26
1:A:443:ILE:CG1	1:A:449:VAL:H	1.49	1.25
1:A:442:ARG:CG	1:A:443:ILE:HD12	1.70	1.22
1:A:471:VAL:O	1:A:474:LEU:HD12	1.45	1.14
1:A:443:ILE:HG13	1:A:449:VAL:HA	1.14	1.13

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	455/483 (94%)	432 (95%)	18 (4%)	5 (1%)	14	20
1	B	466/483 (96%)	457 (98%)	8 (2%)	1 (0%)	47	62
1	C	460/483 (95%)	448 (97%)	11 (2%)	1 (0%)	47	62
1	D	467/483 (97%)	456 (98%)	11 (2%)	0	100	100
All	All	1848/1932 (96%)	1793 (97%)	48 (3%)	7 (0%)	34	48

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	280	ILE
1	C	465	ASP
1	A	488	PRO
1	A	445	THR
1	A	449	VAL

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	369/387 (95%)	326 (88%)	43 (12%)	5	7
1	B	378/387 (98%)	351 (93%)	27 (7%)	14	23
1	C	373/387 (96%)	344 (92%)	29 (8%)	12	19
1	D	378/387 (98%)	351 (93%)	27 (7%)	14	23
All	All	1498/1548 (97%)	1372 (92%)	126 (8%)	11	16

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	382	ASN
1	D	233	LEU
1	C	85	VAL
1	D	213	VAL
1	D	420	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	208	GLN
1	D	344	GLN
1	B	97	GLN
1	B	47	GLN
1	D	417	HIS

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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
4	SIN	B	501	-	7,7,7	1.20	0	8,8,8	1.45	2 (25%)
4	SIN	D	501	-	7,7,7	1.15	0	8,8,8	1.29	1 (12%)
2	SCA	A	601	-	49,57,57	1.20	3 (6%)	61,84,84	2.06	12 (19%)

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	SIN	B	501	-	-	0/5/5/5	-
4	SIN	D	501	-	-	2/5/5/5	-
2	SCA	A	601	-	-	12/52/72/72	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	SCA	CS1-S	-2.77	1.69	1.76
2	A	601	SCA	OS4-CS4	-2.36	1.22	1.30
2	A	601	SCA	C4-N3	-2.02	1.32	1.35

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	SCA	CS2-CS1-S	7.97	122.73	113.46
2	A	601	SCA	OS1-CS1-CS2	-7.88	114.68	123.99
2	A	601	SCA	CP1-CP2-NP1	-4.37	103.24	112.42
2	A	601	SCA	CS3-CS2-CS1	-3.58	107.03	112.41
2	A	601	SCA	C4-C5-N7	-3.12	106.15	109.40

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	SCA	CPB-O7-P2-O6
2	A	601	SCA	CPB-O7-P2-O21
2	A	601	SCA	CPB-O7-P2-O22
2	A	601	SCA	OP2-CP6-CP7-CPA
2	A	601	SCA	NP2-CP6-CP7-CPA

There are no ring outliers.

3 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	501	SIN	2	0
4	D	501	SIN	1	0
2	A	601	SCA	14	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.

Ligand SIN B 501

Ligand SIN D 501

Bond lengths

Bond angles

Torsions

Rings

Ligand SCA A 601

Bond lengths

Bond angles

Torsions

Rings
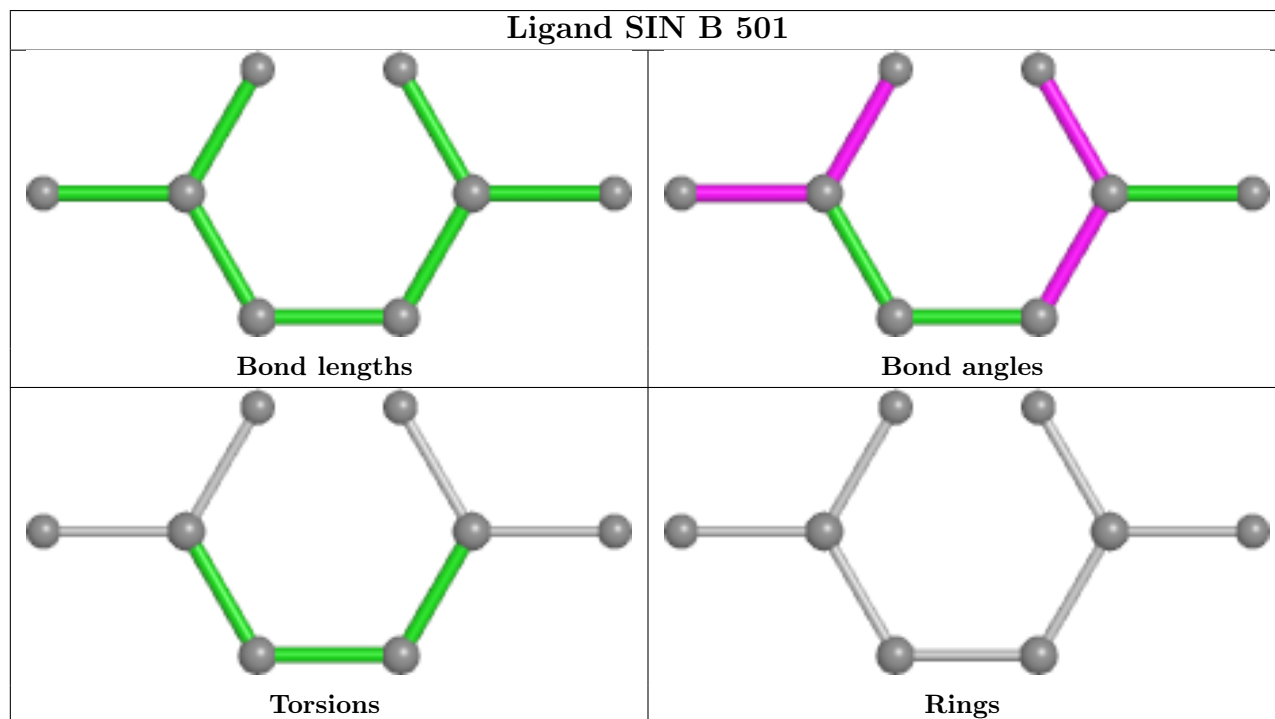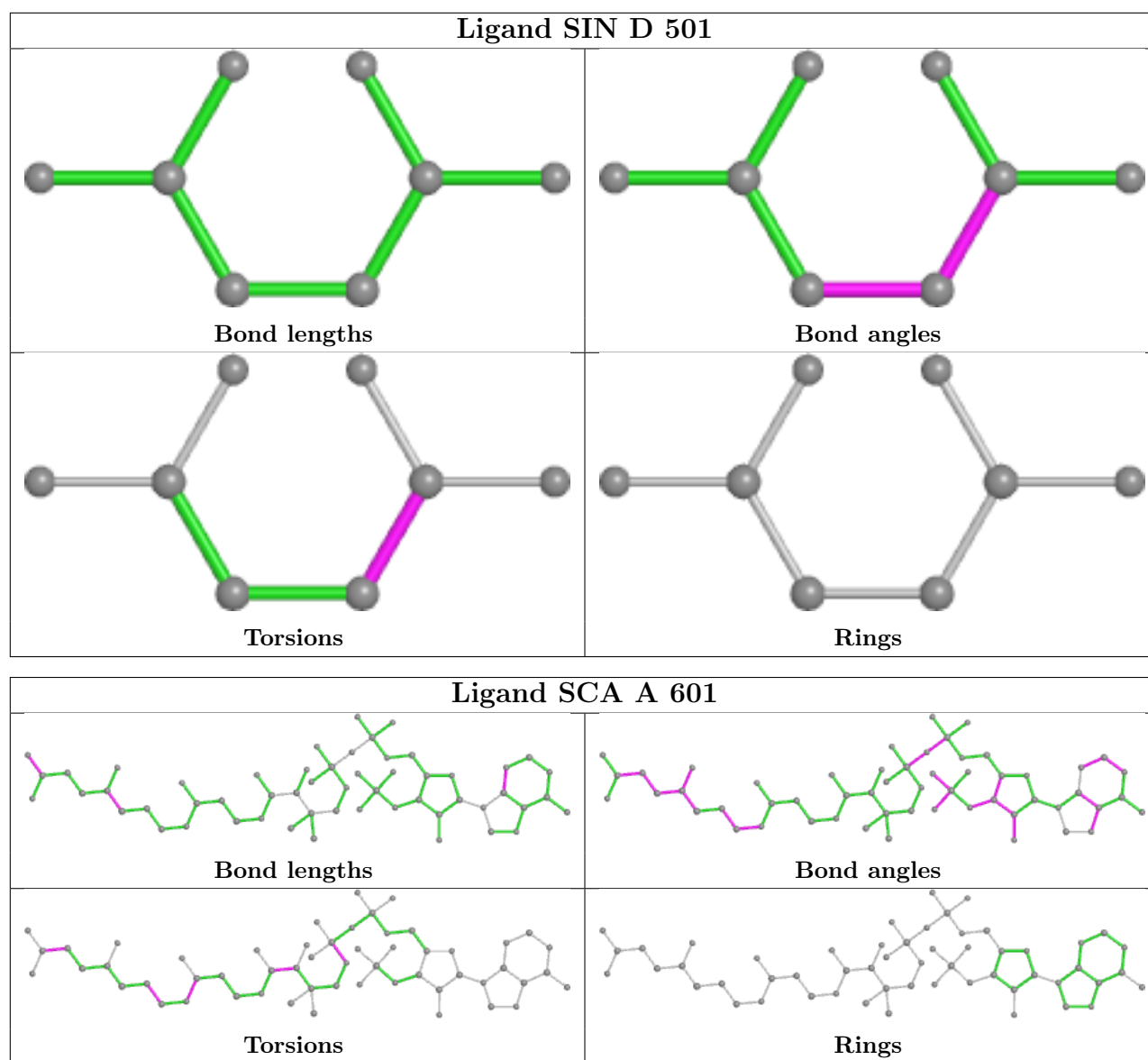
5.7   Other polymers ⓘ

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, 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(Å²)	Q<0.9
1	A	459/483 (95%)	0.33	62 (13%)	3	2	23, 39, 100, 153	1 (0%)
1	B	470/483 (97%)	-0.32	1 (0%)	95	94	23, 35, 56, 91	1 (0%)
1	C	464/483 (96%)	0.16	40 (8%)	10	9	24, 45, 92, 146	1 (0%)
1	D	471/483 (97%)	-0.29	3 (0%)	89	88	23, 37, 58, 110	1 (0%)
All	All	1864/1932 (96%)	-0.03	106 (5%)	23	22	23, 38, 86, 153	4 (0%)

The worst 5 of 106 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	486	THR	11.6
1	A	458	VAL	6.7
1	A	464	GLU	6.5
1	C	147	ASP	6.2
1	A	483	ASN	5.3

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, 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
3	CA	D	502	1/1	0.72	0.17	84,84,84,84	0
4	SIN	D	501	8/8	0.78	0.21	60,65,69,75	0
3	CA	B	502	1/1	0.80	0.09	70,70,70,70	0
2	SCA	A	601	55/55	0.84	0.23	65,79,98,108	0
3	CA	C	701	1/1	0.85	0.08	69,69,69,69	0
4	SIN	B	501	8/8	0.87	0.25	55,58,63,65	0
3	CA	A	602	1/1	0.94	0.06	61,61,61,61	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 CA D 502:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
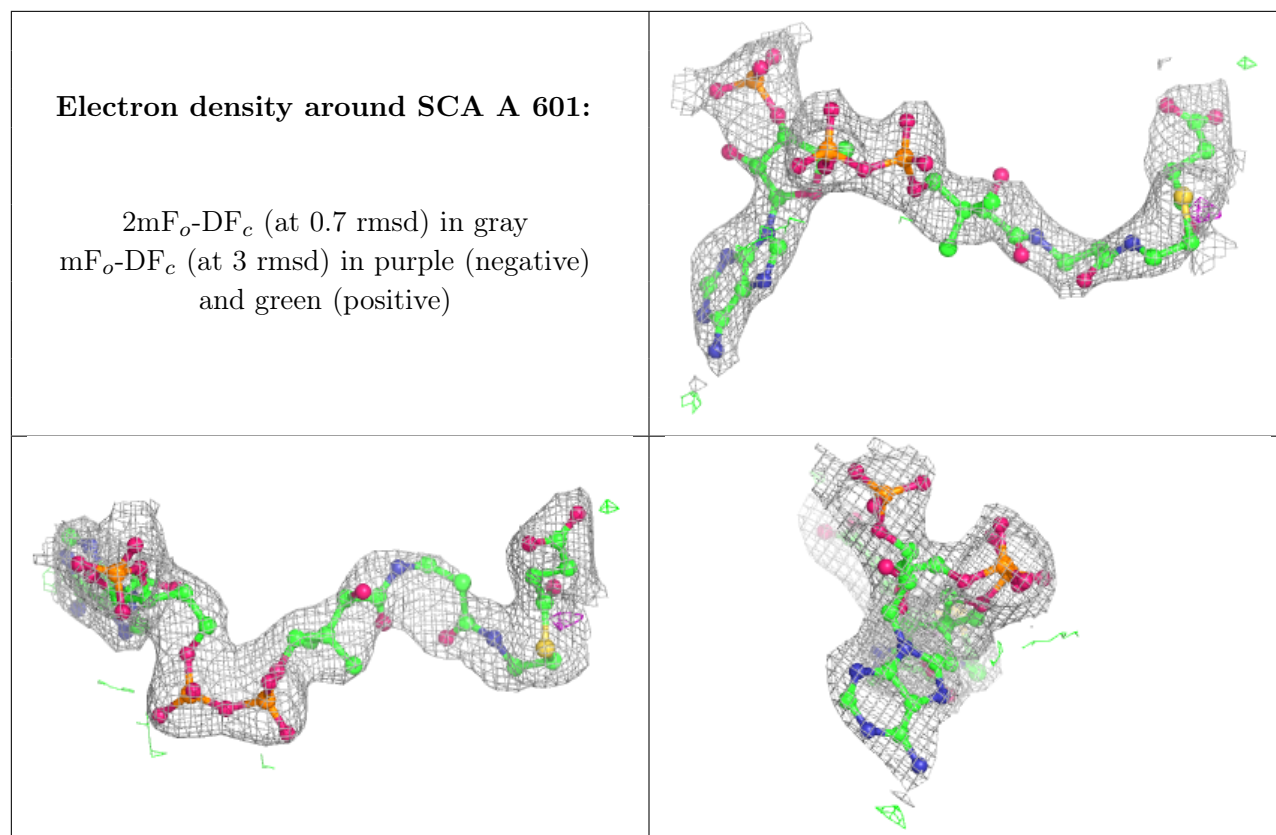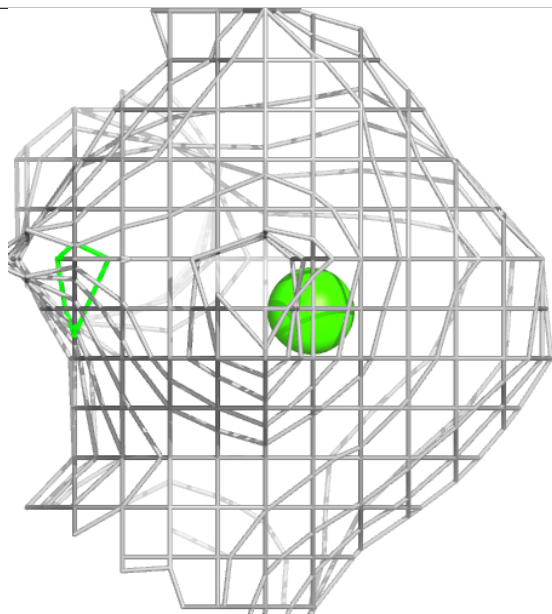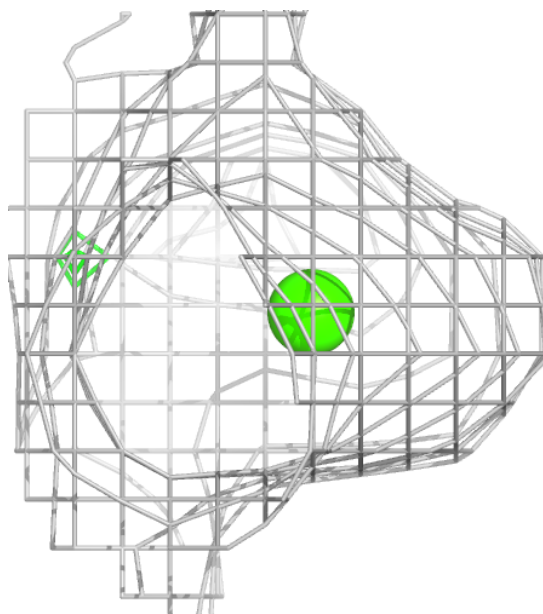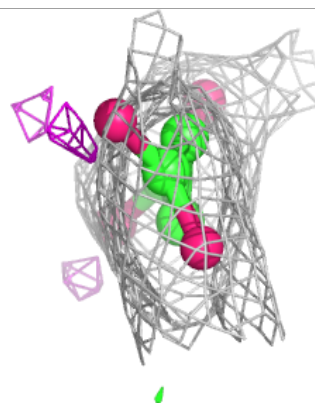and green (positive)
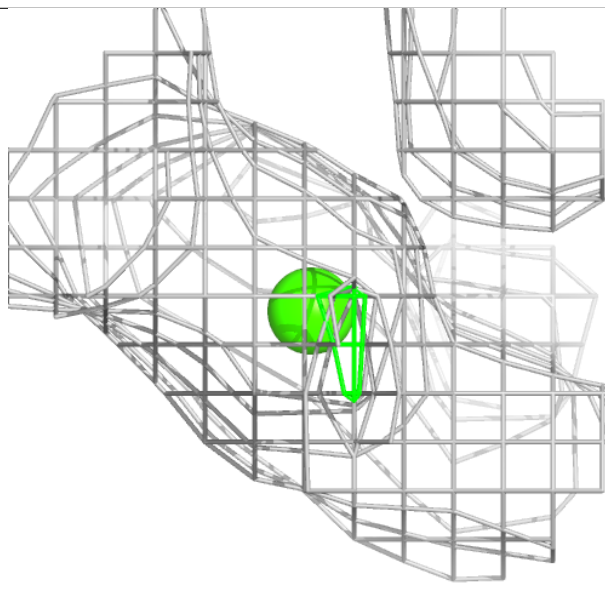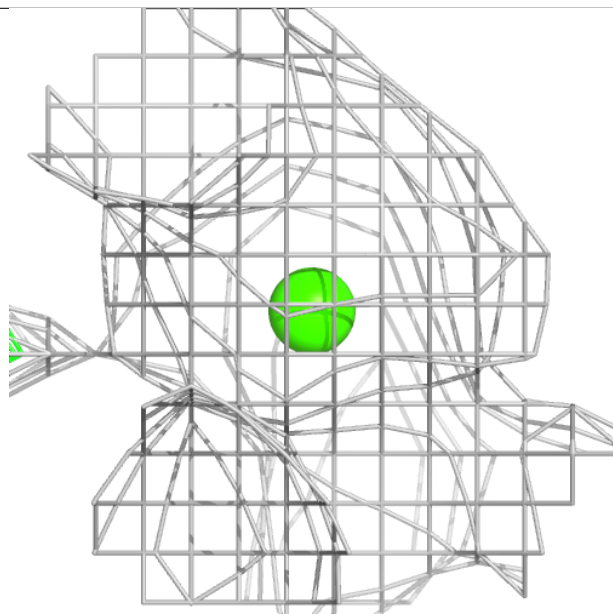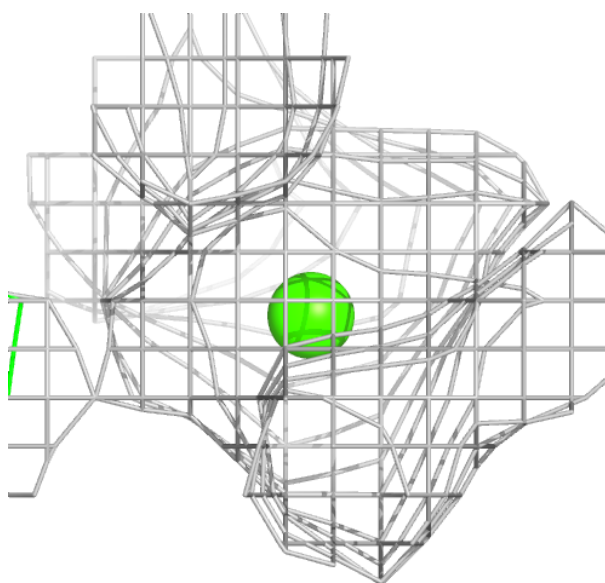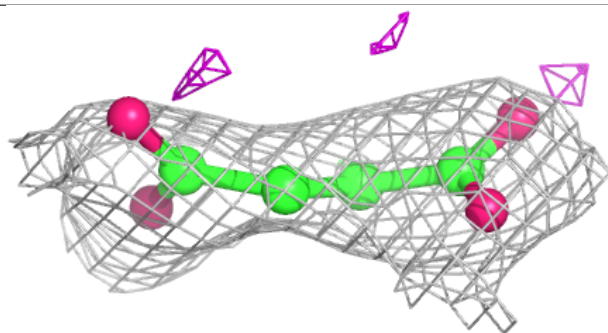
Electron density around SIN D 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
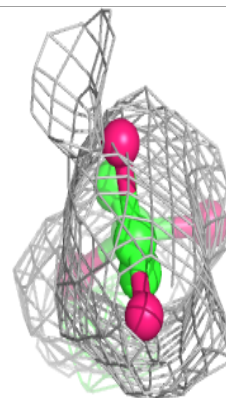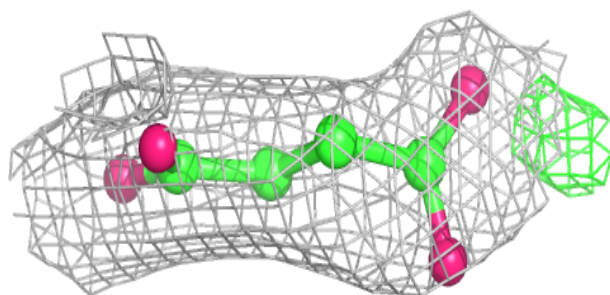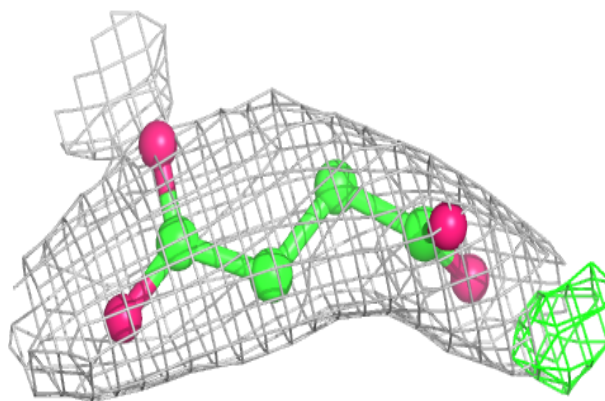
Electron density around CA B 502:

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

Electron density around SCA A 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
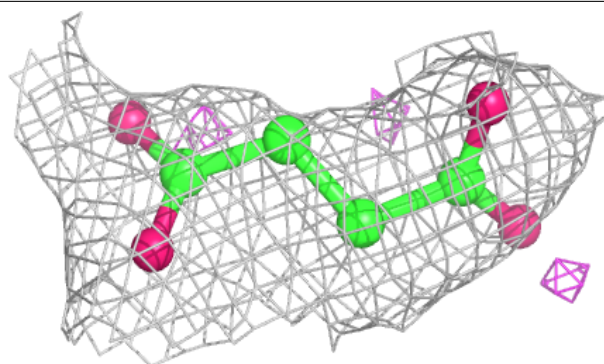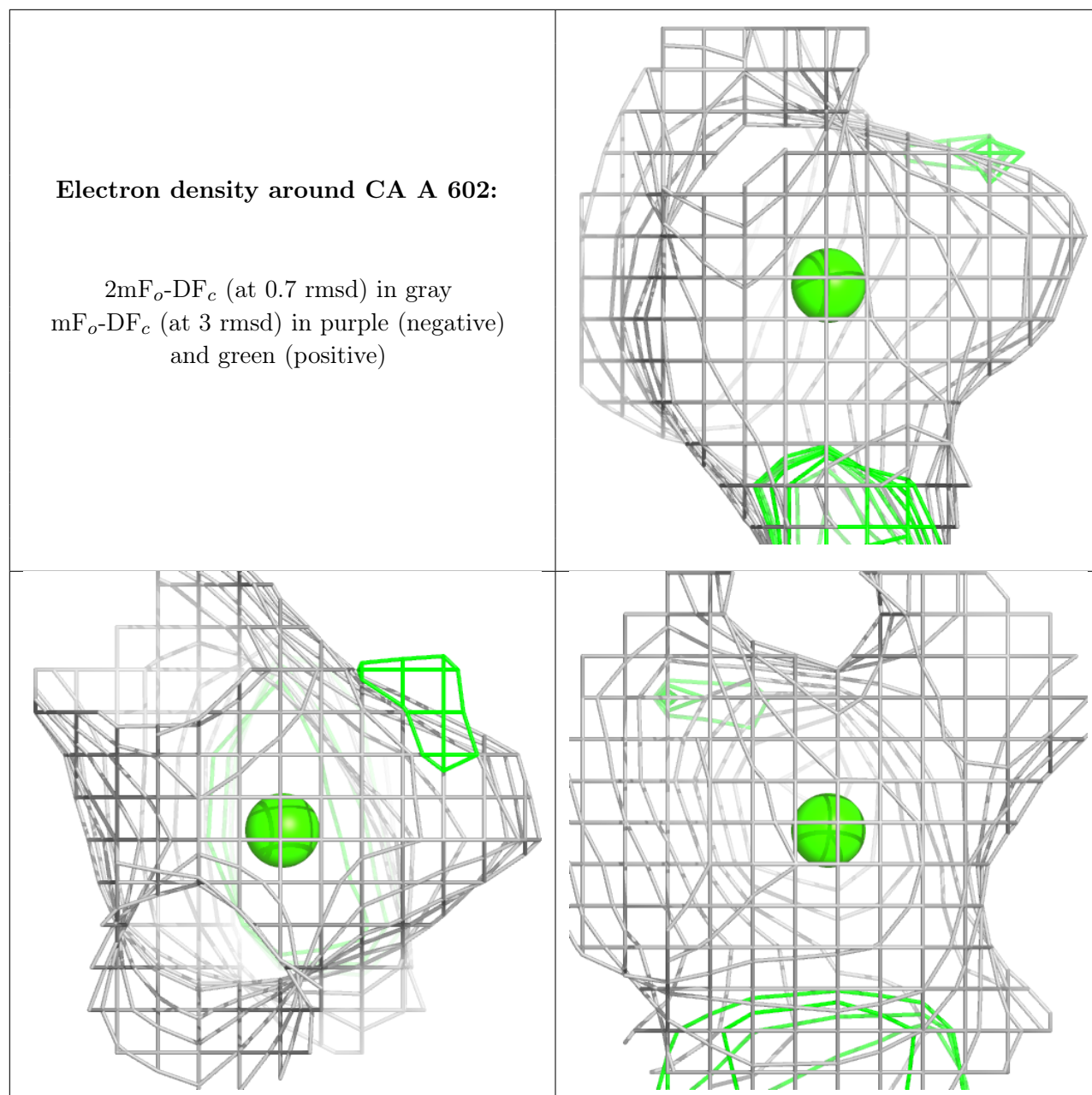and green (positive)
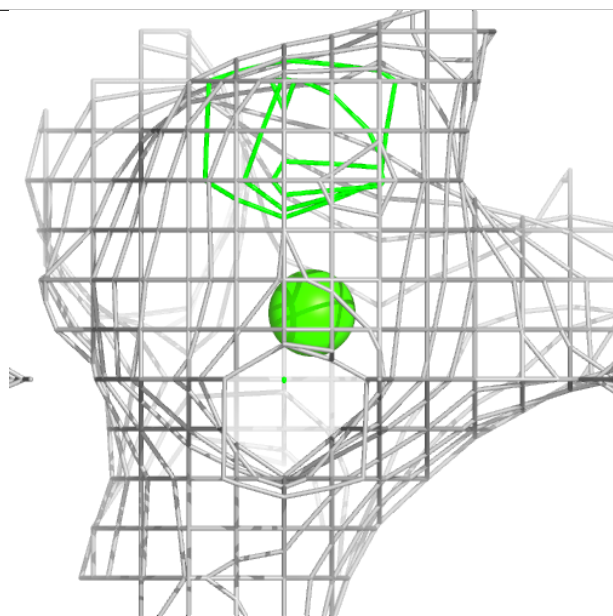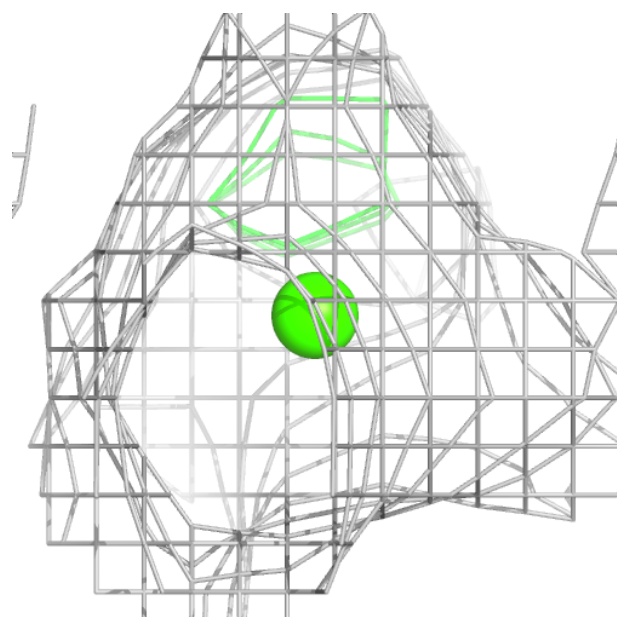
Electron density around CA C 701:

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

Electron density around SIN B 501:

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

Electron density around CA A 602:

$2mF_o$-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.