Full wwPDB X-ray Structure Validation Report (i)

Jul 26, 2022 – 04:14 pm BST

PDB ID	:	7PTH
Title	:	C54S mutant of choline-sulfatase from E. meliloti CECT4857 bound to choline
Authors	:	Gavira, J.A.; Martinez-Rodriguez, S.
Deposited on	:	2021-09-27
Resolution	:	1.85 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 types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

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

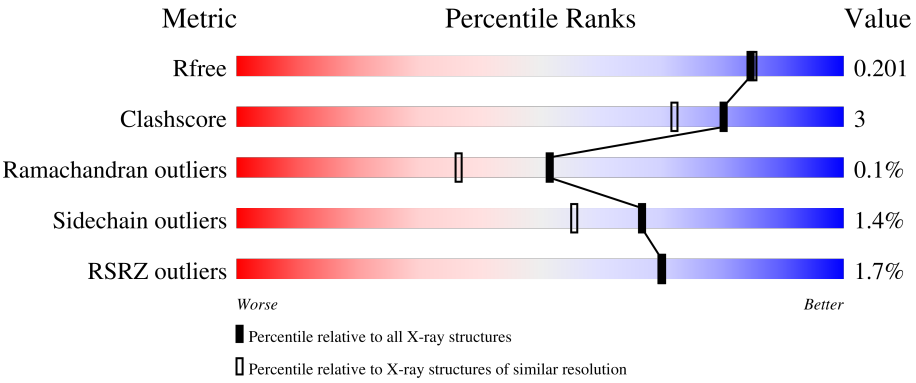
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	520	2% 94% 6%
1	B	520	% 90% 8%
1	C	520	2% 92% 7%
1	D	520	% 91% 7%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 35148 atoms, of which 16034 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 Choline sulfatase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	517	Total	C	H	N	O	S	0	22	0
			8394	2726	4061	766	820	21			
1	B	512	Total	C	H	N	O	S	0	12	0
			8139	2645	3954	732	786	22			
1	C	517	Total	C	H	N	O	S	0	8	0
			8185	2662	3969	746	787	21			
1	D	512	Total	C	H	N	O	S	0	10	0
			8138	2638	3960	729	790	21			

There are 36 discrepancies between the modelled and reference sequences:

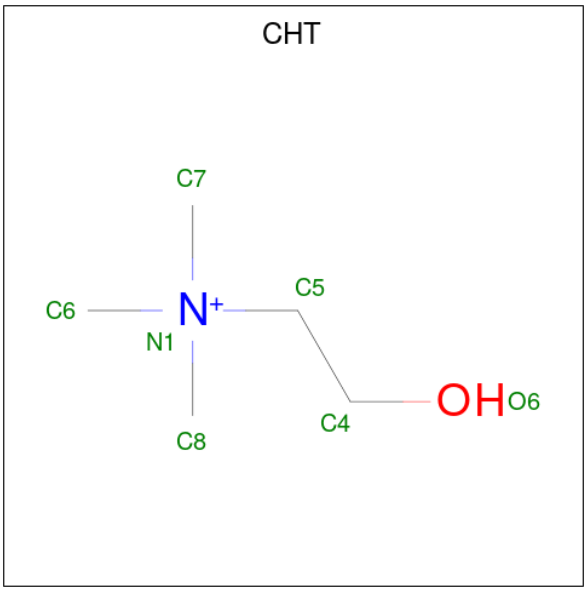
Chain	Residue	Modelled	Actual	Comment	Reference
A	54	SER	CYS	engineered mutation	UNP A0A410NSD4
A	513	GLY	-	expression tag	UNP A0A410NSD4
A	514	SER	-	expression tag	UNP A0A410NSD4
A	515	HIS	-	expression tag	UNP A0A410NSD4
A	516	HIS	-	expression tag	UNP A0A410NSD4
A	517	HIS	-	expression tag	UNP A0A410NSD4
A	518	HIS	-	expression tag	UNP A0A410NSD4
A	519	HIS	-	expression tag	UNP A0A410NSD4
A	520	HIS	-	expression tag	UNP A0A410NSD4
B	54	SER	CYS	engineered mutation	UNP A0A410NSD4
B	513	GLY	-	expression tag	UNP A0A410NSD4
B	514	SER	-	expression tag	UNP A0A410NSD4
B	515	HIS	-	expression tag	UNP A0A410NSD4
B	516	HIS	-	expression tag	UNP A0A410NSD4
B	517	HIS	-	expression tag	UNP A0A410NSD4
B	518	HIS	-	expression tag	UNP A0A410NSD4
B	519	HIS	-	expression tag	UNP A0A410NSD4
B	520	HIS	-	expression tag	UNP A0A410NSD4
C	54	SER	CYS	engineered mutation	UNP A0A410NSD4
C	513	GLY	-	expression tag	UNP A0A410NSD4
C	514	SER	-	expression tag	UNP A0A410NSD4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	515	HIS	-	expression tag	UNP A0A410NSD4
C	516	HIS	-	expression tag	UNP A0A410NSD4
C	517	HIS	-	expression tag	UNP A0A410NSD4
C	518	HIS	-	expression tag	UNP A0A410NSD4
C	519	HIS	-	expression tag	UNP A0A410NSD4
C	520	HIS	-	expression tag	UNP A0A410NSD4
D	54	SER	CYS	engineered mutation	UNP A0A410NSD4
D	513	GLY	-	expression tag	UNP A0A410NSD4
D	514	SER	-	expression tag	UNP A0A410NSD4
D	515	HIS	-	expression tag	UNP A0A410NSD4
D	516	HIS	-	expression tag	UNP A0A410NSD4
D	517	HIS	-	expression tag	UNP A0A410NSD4
D	518	HIS	-	expression tag	UNP A0A410NSD4
D	519	HIS	-	expression tag	UNP A0A410NSD4
D	520	HIS	-	expression tag	UNP A0A410NSD4

- Molecule 2 is CHOLINE ION (three-letter code: CHT) (formula: $C_5H_{14}NO$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			21	5	14	1	1		
2	B	1	Total	C	H	N	O	0	0
			21	5	14	1	1		
2	C	1	Total	C	H	N	O	0	0
			21	5	14	1	1		

Continued on next page...

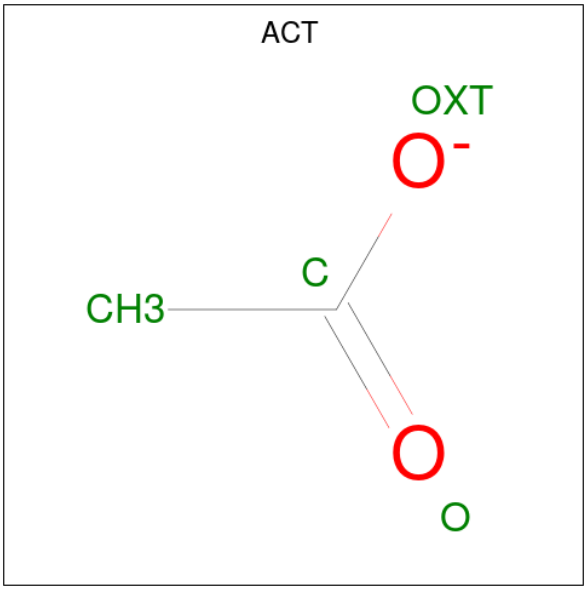
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	D	1	Total	C	H	N	O	0	0
			21	5	14	1	1		

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

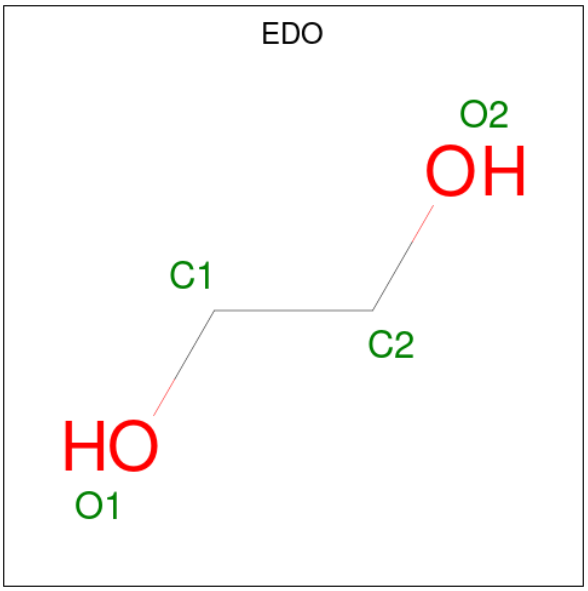
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



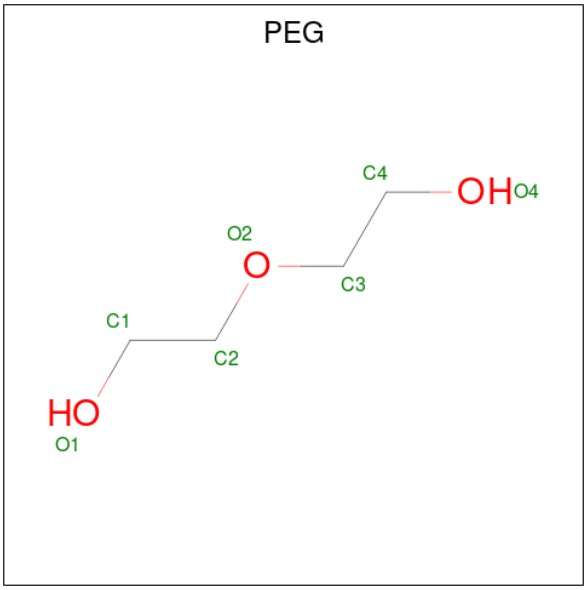
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	O	0	0
			7	2	3	2		
4	B	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	H	O	0	1
			20	4	12	4		
5	D	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 7 is water.

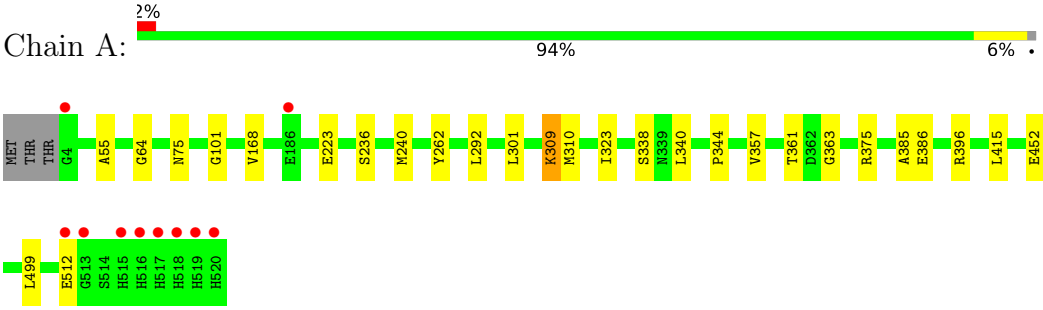
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	529	Total O 529 529	0	0
7	B	520	Total O 520 520	0	0
7	C	582	Total O 582 582	0	0
7	D	512	Total O 512 512	0	0

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.

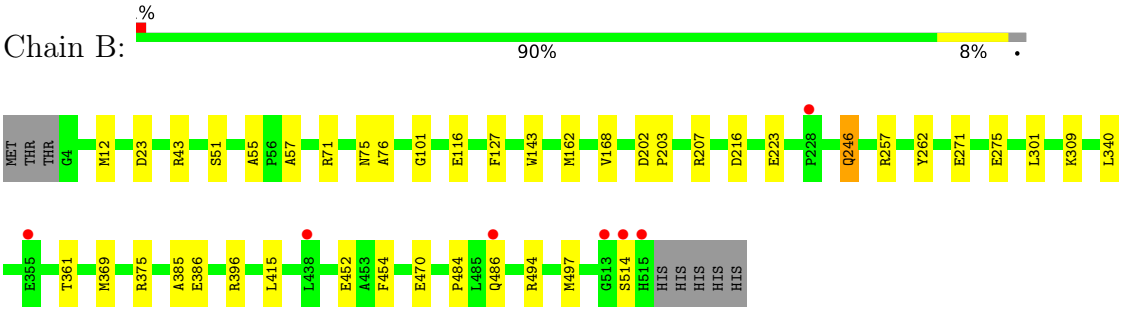
• Molecule 1: Choline sulfatase

Chain A:



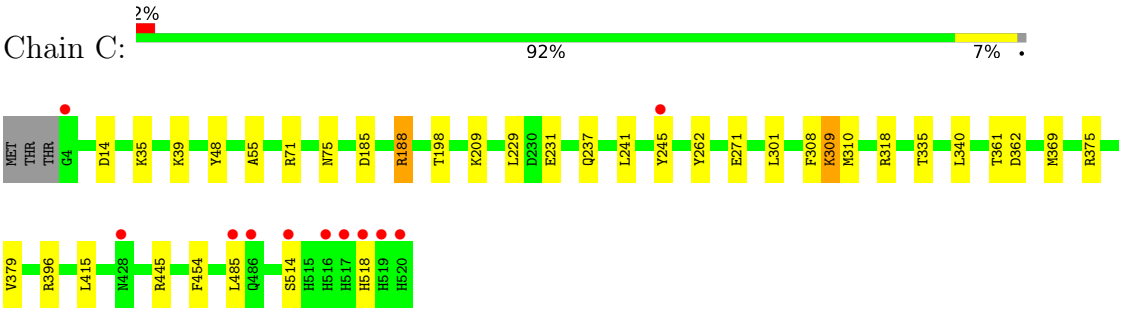
• Molecule 1: Choline sulfatase

Chain B:



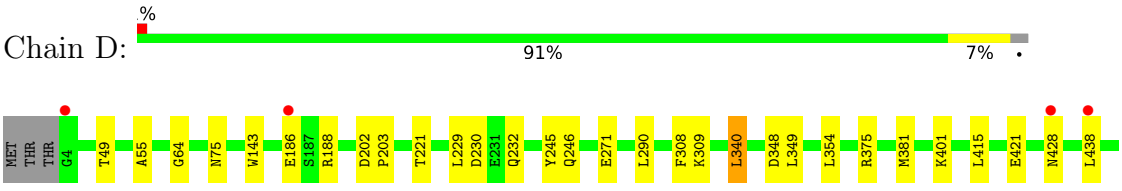
• Molecule 1: Choline sulfatase

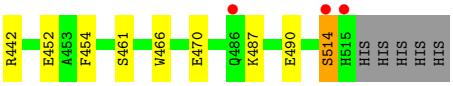
Chain C:



• Molecule 1: Choline sulfatase

Chain D:




R442
E452
A453
F454
S461
W466
E470
Q486
K487
E490
S514
H515
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.77Å 103.97Å 108.97Å 90.00° 104.04° 90.00°	Depositor
Resolution (Å)	45.49 – 1.85 45.49 – 1.85	Depositor EDS
% Data completeness (in resolution range)	93.2 (45.49-1.85) 93.3 (45.49-1.85)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.86Å)	Xtriage
Refinement program	PHENIX 1.19-4092, REFMAC 5	Depositor
R, R_{free}	0.163 , 0.201 0.164 , 0.201	Depositor DCC
R_{free} test set	8020 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	22.0	Xtriage
Anisotropy	0.614	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	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	0.97	EDS
Total number of atoms	35148	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.7919e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: PEG, CA, CHT, ACT, EDO

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	0.41	0/4475	0.63	0/6089
1	B	0.38	0/4338	0.62	0/5901
1	C	0.41	0/4347	0.62	0/5913
1	D	0.40	0/4319	0.62	0/5875
All	All	0.40	0/17479	0.62	0/23778

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	4333	4061	4088	17	0
1	B	4185	3954	4009	26	0
1	C	4216	3969	4025	22	0
1	D	4178	3960	3993	23	0
2	A	7	14	14	1	0
2	B	7	14	14	1	0
2	C	7	14	14	1	0
2	D	7	14	14	1	0
3	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
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	B	8	6	6	0	0
5	C	8	12	12	0	0
5	D	4	6	6	1	0
6	C	7	10	10	0	0
7	A	529	0	0	3	2
7	B	520	0	0	10	1
7	C	582	0	0	8	3
7	D	512	0	0	9	0
All	All	19114	16034	16205	90	3

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 3.

All (90) 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:D:452[A]:GLU:OE2	7:D:701:HOH:O	1.93	0.86
1:B:484:PRO:O	7:B:701:HOH:O	1.93	0.85
1:D:230:ASP:OD2	7:D:703:HOH:O	1.95	0.85
1:D:490:GLU:OE2	7:D:702:HOH:O	1.94	0.85
1:A:452[B]:GLU:OE2	7:A:701:HOH:O	1.96	0.84
1:D:186[A]:GLU:OE2	7:D:704:HOH:O	1.98	0.80
1:B:486:GLN:NE2	7:B:707:HOH:O	2.15	0.79
1:B:452[A]:GLU:OE2	7:B:702:HOH:O	1.99	0.79
1:D:271:GLU:OE2	7:D:705:HOH:O	2.01	0.77
1:B:257:ARG:HD2	1:D:514:SER:HA	1.66	0.77
1:C:362:ASP:OD2	7:C:701:HOH:O	2.04	0.76
1:A:223[A]:GLU:OE2	7:A:702:HOH:O	2.03	0.75
1:B:223:GLU:OE1	7:B:704:HOH:O	2.10	0.69
1:B:470:GLU:OE2	7:B:705:HOH:O	2.11	0.68
1:B:162:MET:SD	1:B:207[A]:ARG:HG3	2.36	0.65
1:B:116:GLU:OE2	7:B:706:HOH:O	2.15	0.64
1:A:361[B]:THR:HG22	1:A:363:GLY:H	1.63	0.64
1:A:344:PRO:HD3	1:A:361[B]:THR:HG21	1.80	0.63
1:C:231:GLU:OE2	7:C:702:HOH:O	2.16	0.61
1:B:216:ASP:OD1	1:B:216:ASP:N	2.32	0.60
1:A:396:ARG:CZ	1:A:415:LEU:HD12	2.33	0.58
1:C:71:ARG:NH2	7:C:708:HOH:O	2.35	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:396:ARG:CZ	1:B:415:LEU:HD12	2.35	0.55
1:B:262:TYR:CE2	1:B:301:LEU:HD11	2.41	0.55
1:C:271:GLU:OE2	7:C:703:HOH:O	2.18	0.55
1:C:55:ALA:HB3	1:C:75:ASN:HA	1.90	0.53
1:A:385:ALA:HB1	1:A:386[A]:GLU:OE2	2.11	0.51
1:D:55:ALA:HB3	1:D:75:ASN:HA	1.93	0.51
1:B:494:ARG:O	1:B:497[A]:MET:HG2	2.10	0.50
1:A:236:SER:O	1:A:240:MET:HG3	2.12	0.50
1:A:55:ALA:HB3	1:A:75:ASN:HA	1.93	0.49
1:B:361[B]:THR:HB	7:B:802:HOH:O	2.11	0.49
1:B:514:SER:HB3	7:B:864:HOH:O	2.12	0.48
1:C:514:SER:HB3	1:C:518:HIS:HA	1.95	0.48
1:A:262:TYR:CE2	1:A:301:LEU:HD11	2.47	0.48
1:B:12:MET:HG2	7:B:772:HOH:O	2.13	0.48
1:D:401:LYS:HE2	1:D:421:GLU:OE2	2.14	0.48
1:C:514:SER:CB	1:C:518:HIS:HA	2.44	0.48
1:C:185:ASP:O	1:C:188:ARG:HD3	2.14	0.48
1:D:348:ASP:O	7:D:706:HOH:O	2.20	0.47
1:A:309:LYS:O	1:A:310:MET:HB2	2.14	0.47
1:A:512:GLU:OE2	7:A:703:HOH:O	2.20	0.47
1:C:237:GLN:O	1:C:241:LEU:HG	2.15	0.47
1:C:361:THR:HB	7:C:844:HOH:O	2.15	0.47
1:D:229:LEU:HD22	7:D:763:HOH:O	2.15	0.47
1:C:48:TYR:CE2	1:C:318:ARG:HG3	2.50	0.47
1:B:55:ALA:HB3	1:B:75:ASN:HA	1.96	0.46
1:A:101:GLY:HA2	1:A:168:VAL:HG22	1.98	0.45
1:D:232:GLN:OE1	7:D:707:HOH:O	2.21	0.45
1:D:438:LEU:HD21	1:D:442:ARG:NH2	2.32	0.44
1:B:101:GLY:HA2	1:B:168:VAL:HG22	1.98	0.44
1:A:493:MET:HG2	1:A:499:LEU:HD13	2.00	0.44
1:A:357[A]:VAL:O	1:A:361[A]:THR:HG22	2.18	0.44
1:A:64:GLY:HA3	1:A:361[B]:THR:HG23	1.99	0.43
1:B:23:ASP:OD1	1:B:43:ARG:NE	2.46	0.43
1:D:466:TRP:O	1:D:470:GLU:HG3	2.19	0.43
1:C:445[B]:ARG:HD3	7:C:805:HOH:O	2.18	0.43
1:B:202:ASP:HB2	1:B:203:PRO:HA	2.00	0.43
1:C:396:ARG:CZ	1:C:415:LEU:HD12	2.48	0.43
1:C:335:THR:HB	1:C:369:MET:CE	2.49	0.43
1:C:379:VAL:HB	1:C:396:ARG:HB3	2.01	0.43
1:C:209:LYS:NZ	7:C:729:HOH:O	2.49	0.42
1:C:262:TYR:CE2	1:C:301:LEU:HD11	2.53	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:14:ASP:HB3	1:C:198:THR:O	2.19	0.42
1:D:188:ARG:NH1	1:D:188:ARG:HG2	2.33	0.42
1:C:229:LEU:HD22	1:C:245:TYR:OH	2.19	0.42
1:D:64:GLY:C	1:D:340:LEU:HD22	2.40	0.42
1:C:309:LYS:O	1:C:310:MET:HB2	2.18	0.42
1:C:340:LEU:HD21	1:C:362:ASP:HB2	2.02	0.42
1:D:461:SER:HB2	5:D:602:EDO:H21	2.02	0.42
2:A:601:CHT:O6	2:A:601:CHT:C7	2.68	0.42
1:B:385:ALA:HB1	1:B:386:GLU:OE2	2.20	0.41
1:D:245:TYR:CD1	1:D:246:GLN:OE1	2.73	0.41
1:D:354:LEU:HD23	7:D:802:HOH:O	2.20	0.41
1:B:246:GLN:H	1:B:246:GLN:CD	2.22	0.41
1:D:49:THR:O	1:D:381:MET:HG2	2.21	0.41
1:B:51:SER:O	1:B:57:ALA:HB2	2.21	0.41
1:D:188:ARG:HG2	1:D:188:ARG:HH11	1.84	0.41
1:A:292[B]:LEU:HD22	1:A:323:ILE:HG23	2.03	0.41
1:B:271:GLU:O	1:B:275:GLU:HG3	2.19	0.41
1:D:202:ASP:HA	1:D:203:PRO:C	2.42	0.41
1:D:290:LEU:HD13	1:D:349:LEU:HD13	2.02	0.41
2:C:601:CHT:HC41	7:C:743:HOH:O	2.20	0.40
1:B:76:ALA:HB1	1:B:127:PHE:HA	2.04	0.40
1:C:35:LYS:O	1:C:39[A]:LYS:HG2	2.21	0.40
1:D:143:TRP:O	2:D:601:CHT:H71	2.21	0.40
1:B:143:TRP:O	2:B:601:CHT:H61	2.21	0.40

All (3) 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)
7:A:1160:HOH:O	7:C:1119:HOH:O[2_554]	2.11	0.09
7:B:873:HOH:O	7:C:1164:HOH:O[1_455]	2.11	0.09
7:A:1144:HOH:O	7:C:1214:HOH:O[2_555]	2.14	0.06

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	537/520 (103%)	523 (97%)	14 (3%)	0	100	100
1	B	522/520 (100%)	504 (97%)	18 (3%)	0	100	100
1	C	522/520 (100%)	505 (97%)	16 (3%)	1 (0%)	47	33
1	D	520/520 (100%)	504 (97%)	15 (3%)	1 (0%)	47	33
All	All	2101/2080 (101%)	2036 (97%)	63 (3%)	2 (0%)	51	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	308	PHE
1	C	308	PHE

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	460/443 (104%)	456 (99%)	4 (1%)	78	72
1	B	446/443 (101%)	440 (99%)	6 (1%)	69	58
1	C	446/443 (101%)	440 (99%)	6 (1%)	69	58
1	D	444/443 (100%)	435 (98%)	9 (2%)	55	40
All	All	1796/1772 (101%)	1771 (99%)	25 (1%)	67	55

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	309	LYS
1	A	338	SER
1	A	375	ARG
1	A	454	PHE
1	B	71	ARG
1	B	246	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	309	LYS
1	B	369	MET
1	B	375	ARG
1	B	454	PHE
1	C	188	ARG
1	C	309	LYS
1	C	375	ARG
1	C	454	PHE
1	C	485[A]	LEU
1	C	485[B]	LEU
1	D	221	THR
1	D	309	LYS
1	D	340	LEU
1	D	375	ARG
1	D	415	LEU
1	D	428	ASN
1	D	454	PHE
1	D	487	LYS
1	D	514	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	371	ASN
1	C	219	HIS

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 14 ligands modelled in this entry, 4 are monoatomic - leaving 10 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	$\#\|Z\| > 2$
2	CHT	A	601	-	6,6,6	1.26	0	8,8,8	0.32	0
4	ACT	B	603	-	3,3,3	1.20	0	3,3,3	1.44	0
5	EDO	C	602[B]	-	3,3,3	0.47	0	2,2,2	0.35	0
5	EDO	D	602	-	3,3,3	0.62	0	2,2,2	0.21	0
2	CHT	C	601	-	6,6,6	1.15	0	8,8,8	0.34	0
4	ACT	B	602	-	3,3,3	1.16	0	3,3,3	1.34	0
6	PEG	C	603	-	6,6,6	0.17	0	5,5,5	0.09	0
2	CHT	D	601	-	6,6,6	1.17	0	8,8,8	0.36	0
5	EDO	C	602[A]	-	3,3,3	0.44	0	2,2,2	0.56	0
2	CHT	B	601	-	6,6,6	1.17	0	8,8,8	0.36	0

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
2	CHT	A	601	-	-	0/4/4/4	-
5	EDO	C	602[B]	-	-	0/1/1/1	-
5	EDO	D	602	-	-	1/1/1/1	-
2	CHT	C	601	-	-	0/4/4/4	-
6	PEG	C	603	-	-	1/4/4/4	-
2	CHT	D	601	-	-	0/4/4/4	-
5	EDO	C	602[A]	-	-	0/1/1/1	-
2	CHT	B	601	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	602	EDO	O1-C1-C2-O2
6	C	603	PEG	C1-C2-O2-C3

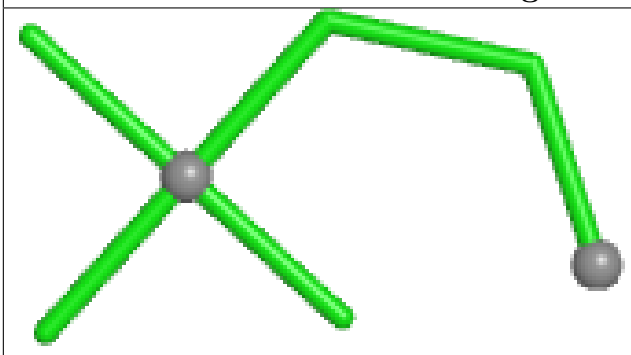
There are no ring outliers.

5 monomers are involved in 5 short contacts:

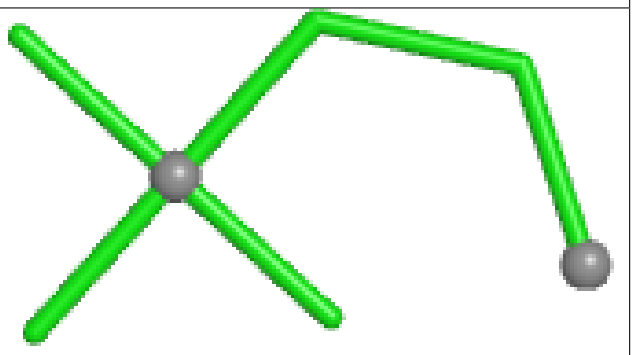
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	CHT	1	0
5	D	602	EDO	1	0
2	C	601	CHT	1	0
2	D	601	CHT	1	0
2	B	601	CHT	1	0

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.

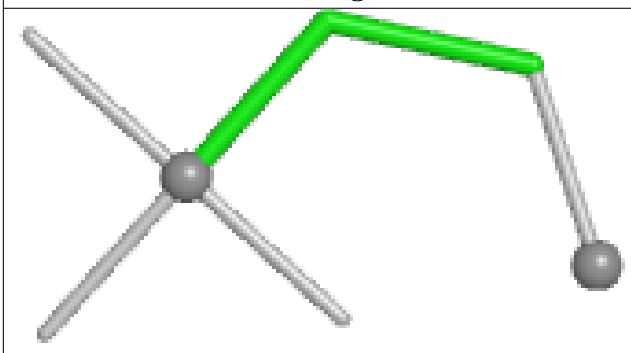
Ligand CHT A 601



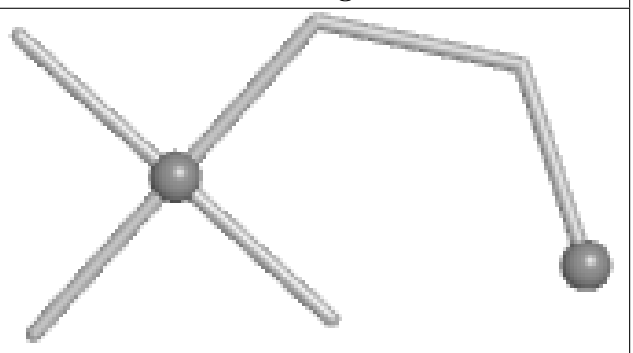
Bond lengths



Bond angles

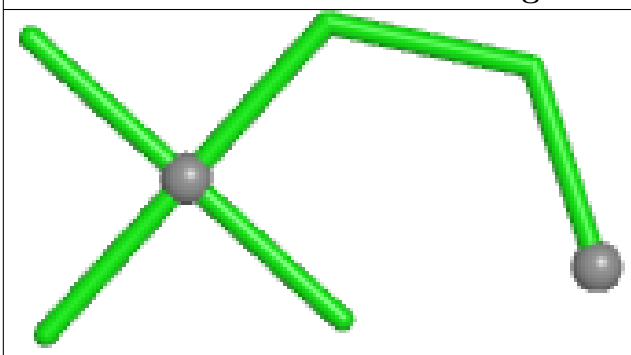


Torsions

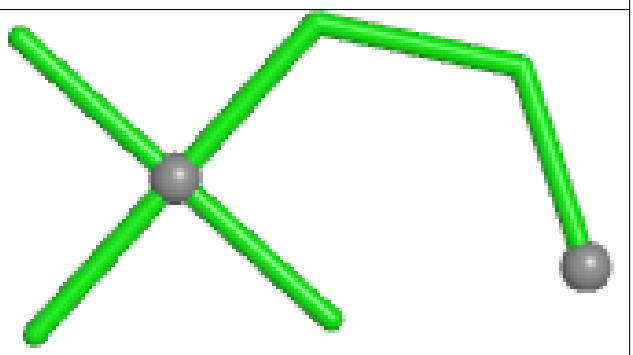


Rings

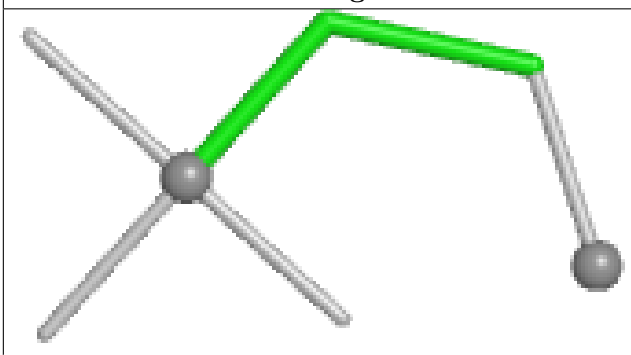
Ligand CHT C 601



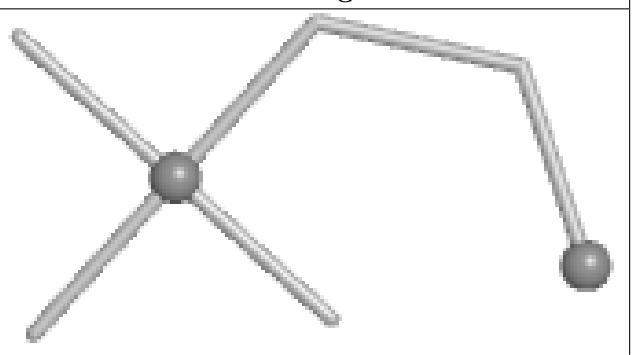
Bond lengths



Bond angles

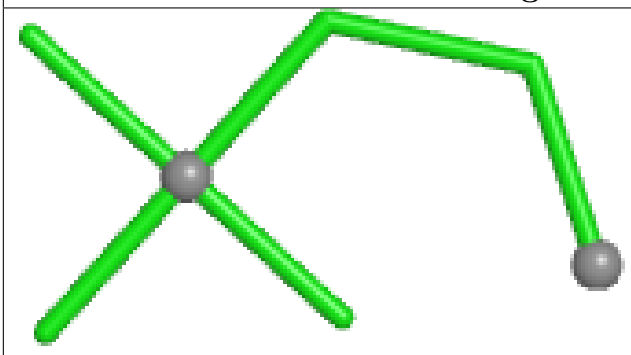


Torsions

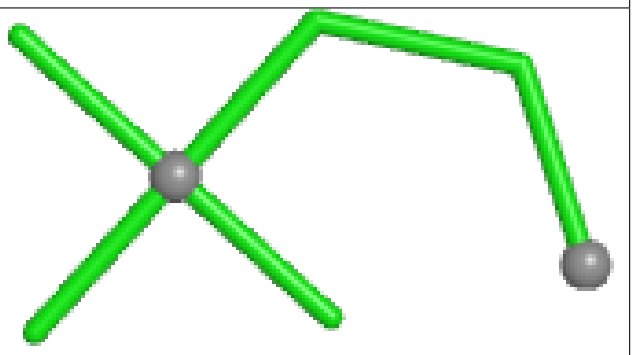


Rings

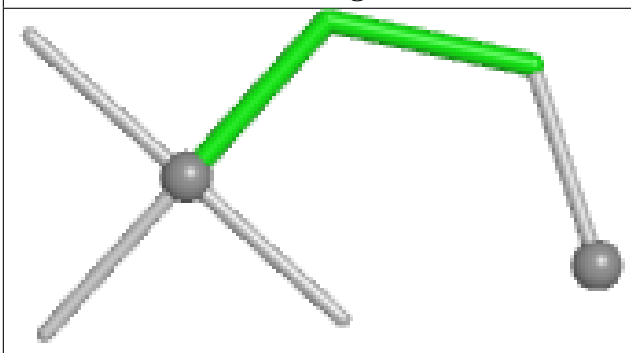
Ligand CHT D 601



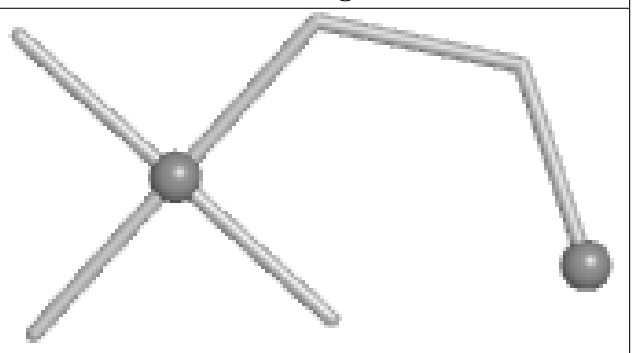
Bond lengths



Bond angles

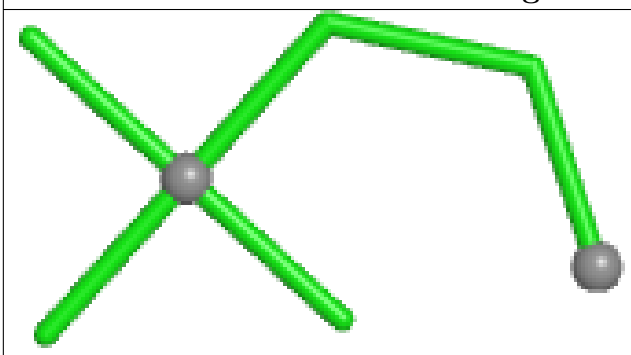


Torsions

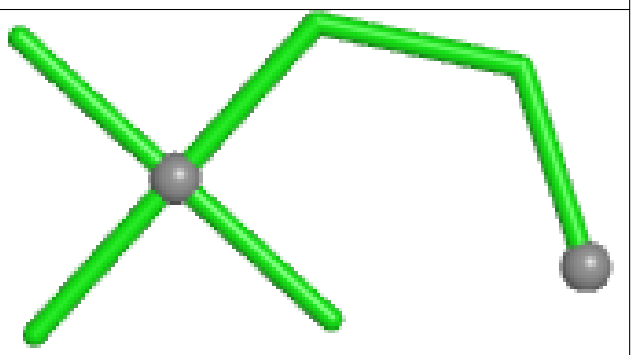


Rings

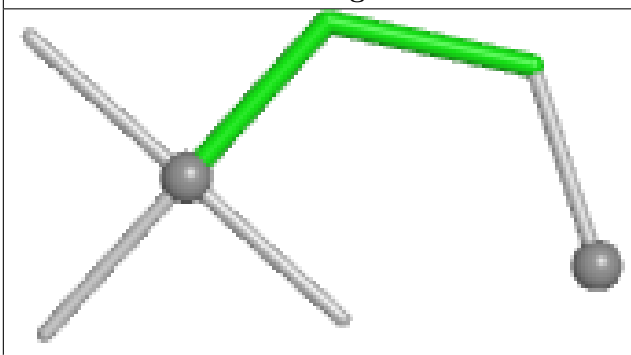
Ligand CHT B 601



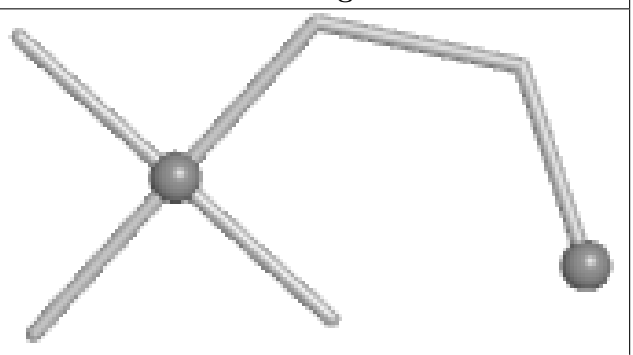
Bond lengths



Bond angles



Torsions



Rings

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	A	517/520 (99%)	-0.15	10 (1%)	66	66	18, 25, 41, 72	0
1	B	512/520 (98%)	-0.11	7 (1%)	75	76	20, 28, 44, 71	0
1	C	517/520 (99%)	-0.19	11 (2%)	63	63	18, 25, 41, 77	0
1	D	512/520 (98%)	-0.31	7 (1%)	75	76	19, 25, 42, 72	0
All	All	2058/2080 (98%)	-0.19	35 (1%)	70	70	18, 26, 42, 77	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	514	SER	6.1
1	D	515	HIS	5.7
1	B	514	SER	5.6
1	A	519	HIS	5.2
1	A	516	HIS	4.9
1	C	518	HIS	4.7
1	D	486	GLN	4.3
1	C	485[A]	LEU	3.9
1	B	513	GLY	3.9
1	A	520	HIS	3.8
1	C	516	HIS	3.6
1	C	520	HIS	3.4
1	A	515	HIS	3.4
1	C	517	HIS	3.4
1	A	517	HIS	3.3
1	C	486[A]	GLN	3.3
1	C	514	SER	3.2
1	D	186[A]	GLU	3.2
1	C	519	HIS	3.2
1	D	428	ASN	3.1
1	B	355[A]	GLU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	513	GLY	2.9
1	A	518	HIS	2.9
1	B	486	GLN	2.8
1	A	4	GLY	2.7
1	D	438	LEU	2.6
1	C	4	GLY	2.6
1	A	186[A]	GLU	2.6
1	D	4	GLY	2.6
1	B	515	HIS	2.4
1	B	228	PRO	2.3
1	B	438	LEU	2.3
1	C	245	TYR	2.2
1	C	428	ASN	2.1
1	A	512	GLU	2.0

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(Å^2)	Q<0.9
4	ACT	B	603	4/4	0.63	0.35	63,71,99,99	0
5	EDO	D	602	4/4	0.70	0.21	31,44,52,52	0
4	ACT	B	602	4/4	0.84	0.12	38,39,47,47	0
6	PEG	C	603	7/7	0.85	0.32	38,46,65,72	0
2	CHT	C	601	7/7	0.86	0.13	27,32,38,38	0
2	CHT	D	601	7/7	0.92	0.09	27,32,39,39	0
5	EDO	C	602[B]	4/4	0.93	0.34	32,38,42,42	10
2	CHT	A	601	7/7	0.93	0.11	26,31,36,37	0
5	EDO	C	602[A]	4/4	0.93	0.34	34,41,42,46	10

Continued on next page...

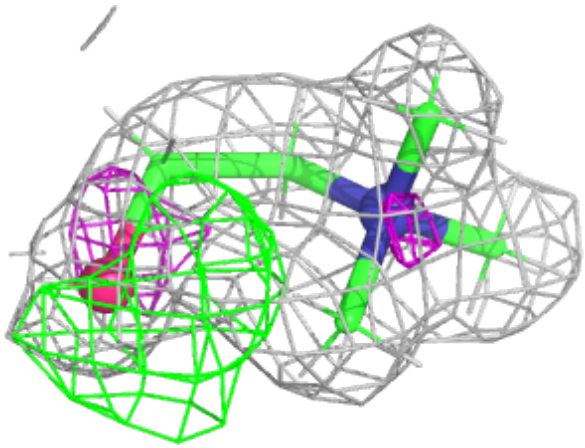
Continued from previous page...

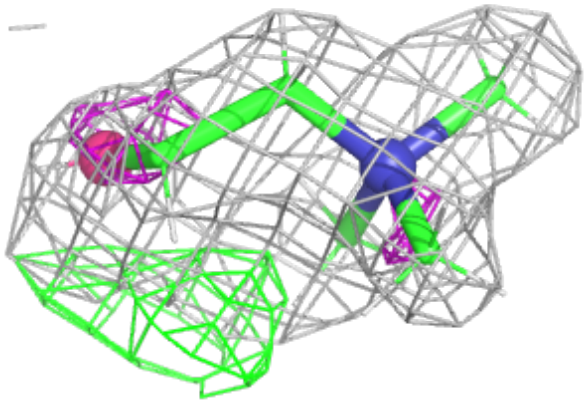
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CHT	B	601	7/7	0.94	0.11	31,38,44,46	0
3	CA	B	604	1/1	0.98	0.04	33,33,33,33	0
3	CA	C	604	1/1	0.98	0.04	33,33,33,33	0
3	CA	A	602	1/1	0.99	0.05	31,31,31,31	0
3	CA	D	603	1/1	0.99	0.05	33,33,33,33	0

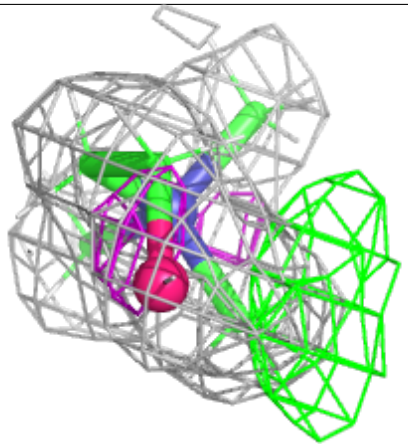
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CHT C 601:

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

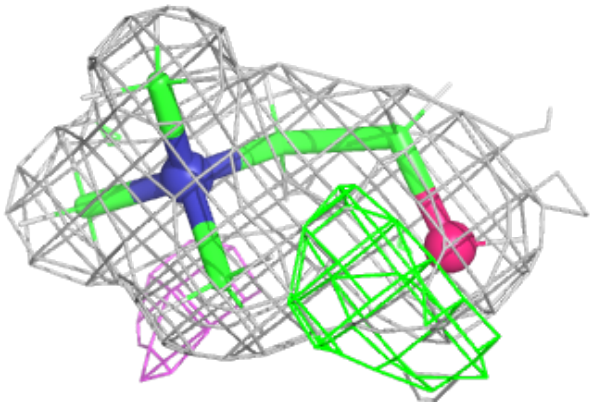


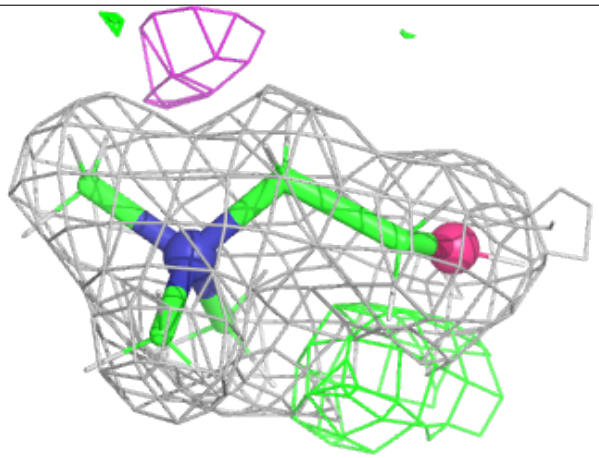


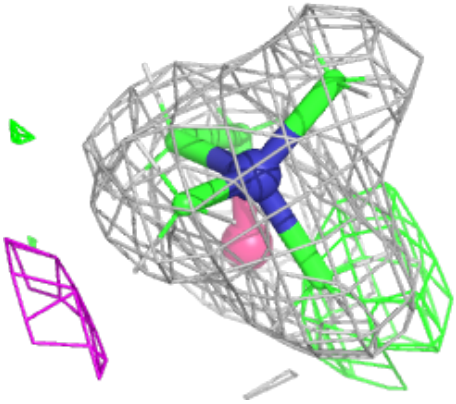


Electron density around CHT D 601:

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

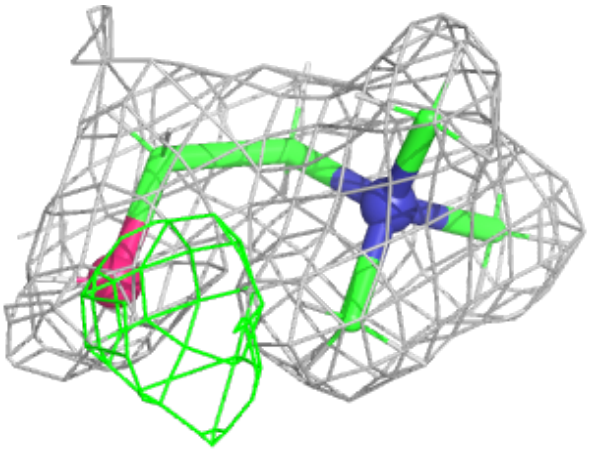


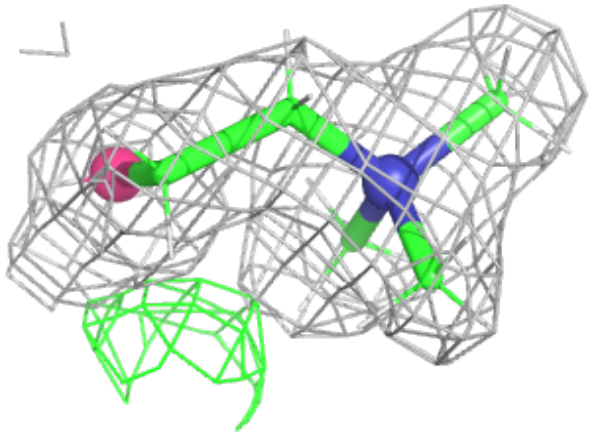


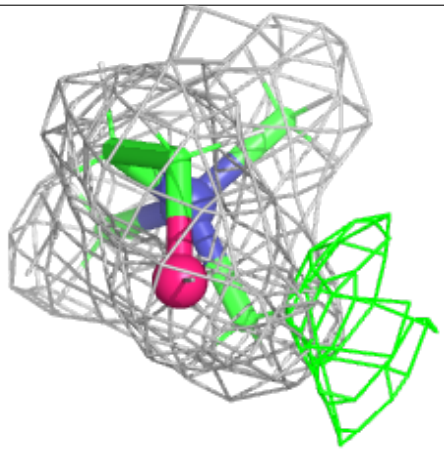


Electron density around CHT A 601:

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

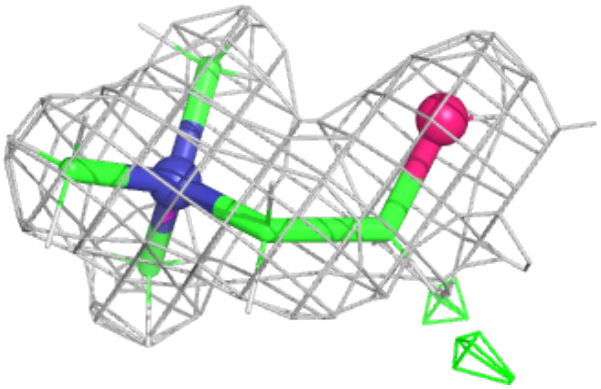


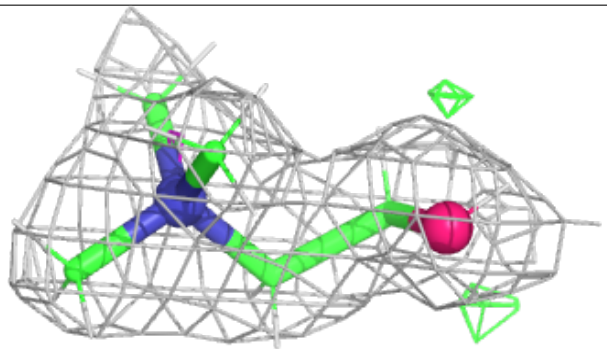


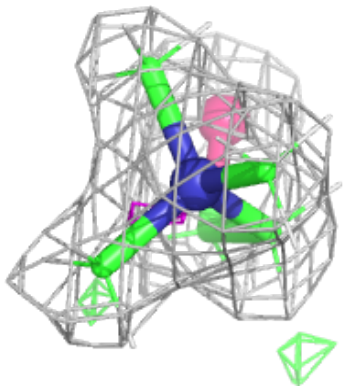


Electron density around CHT B 601:

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

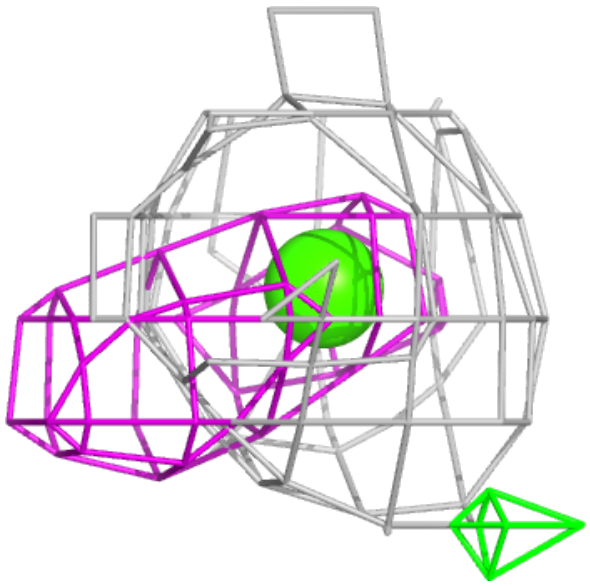


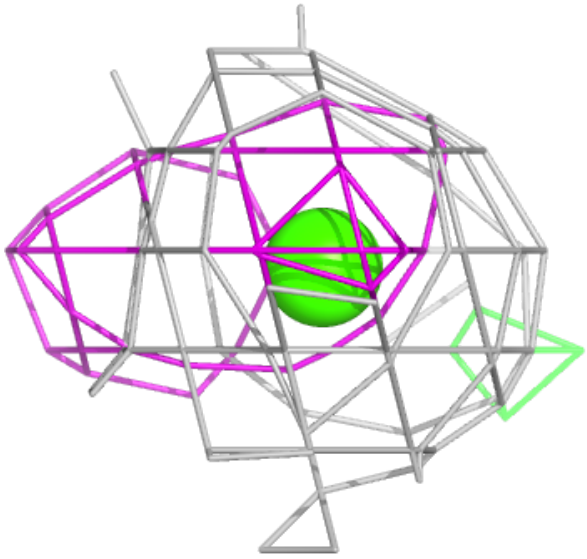


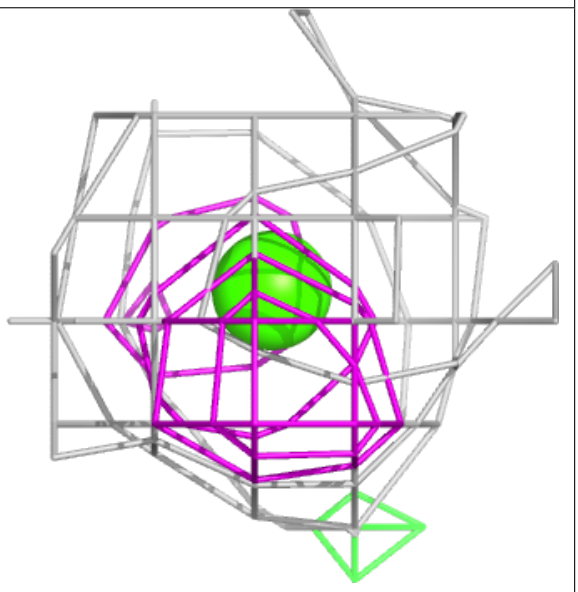


Electron density around CA B 604:

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

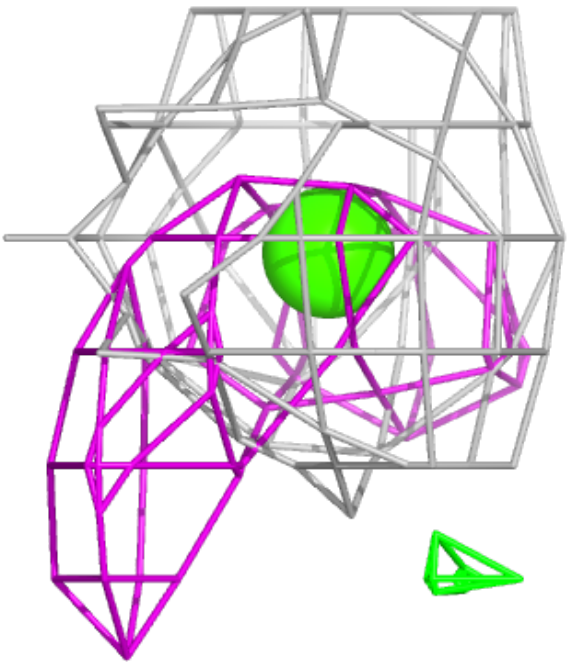


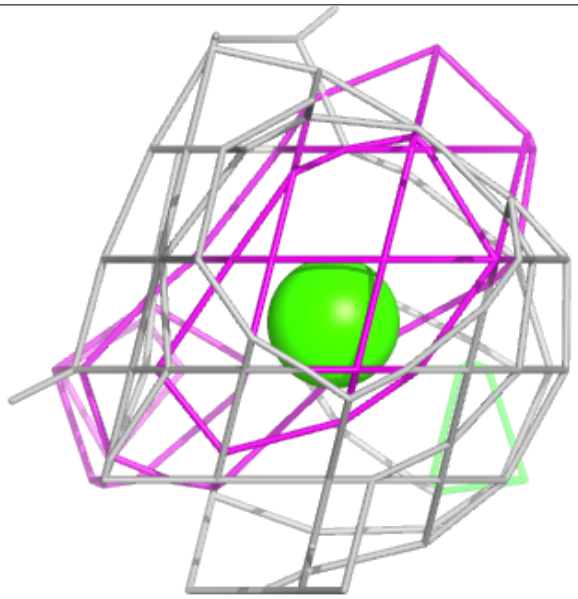


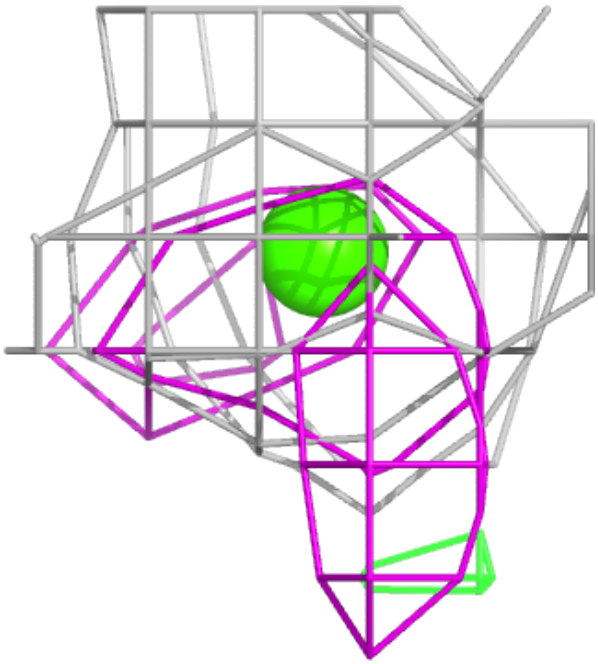


Electron density around CA C 604:

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

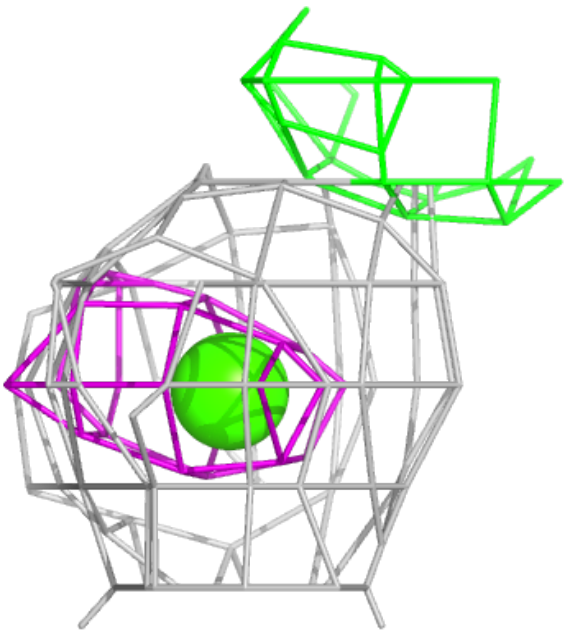


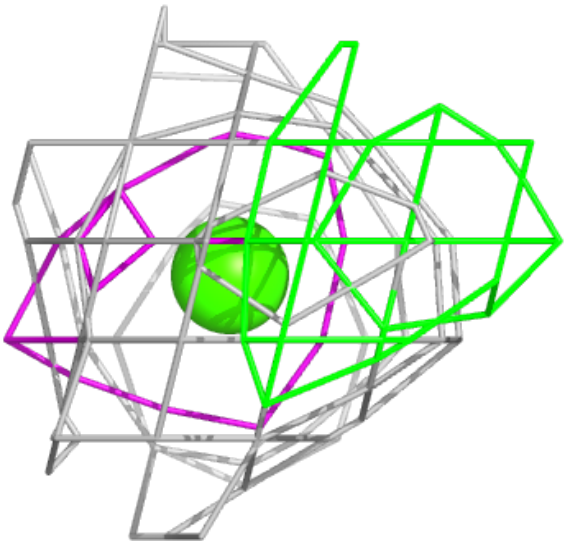


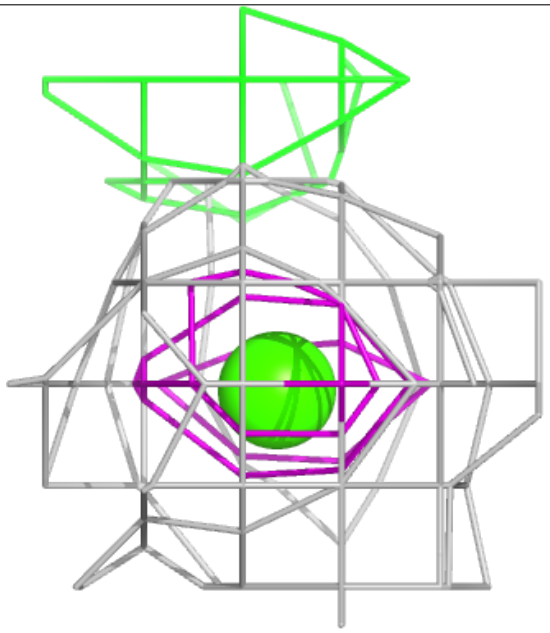


Electron density around CA A 602:

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

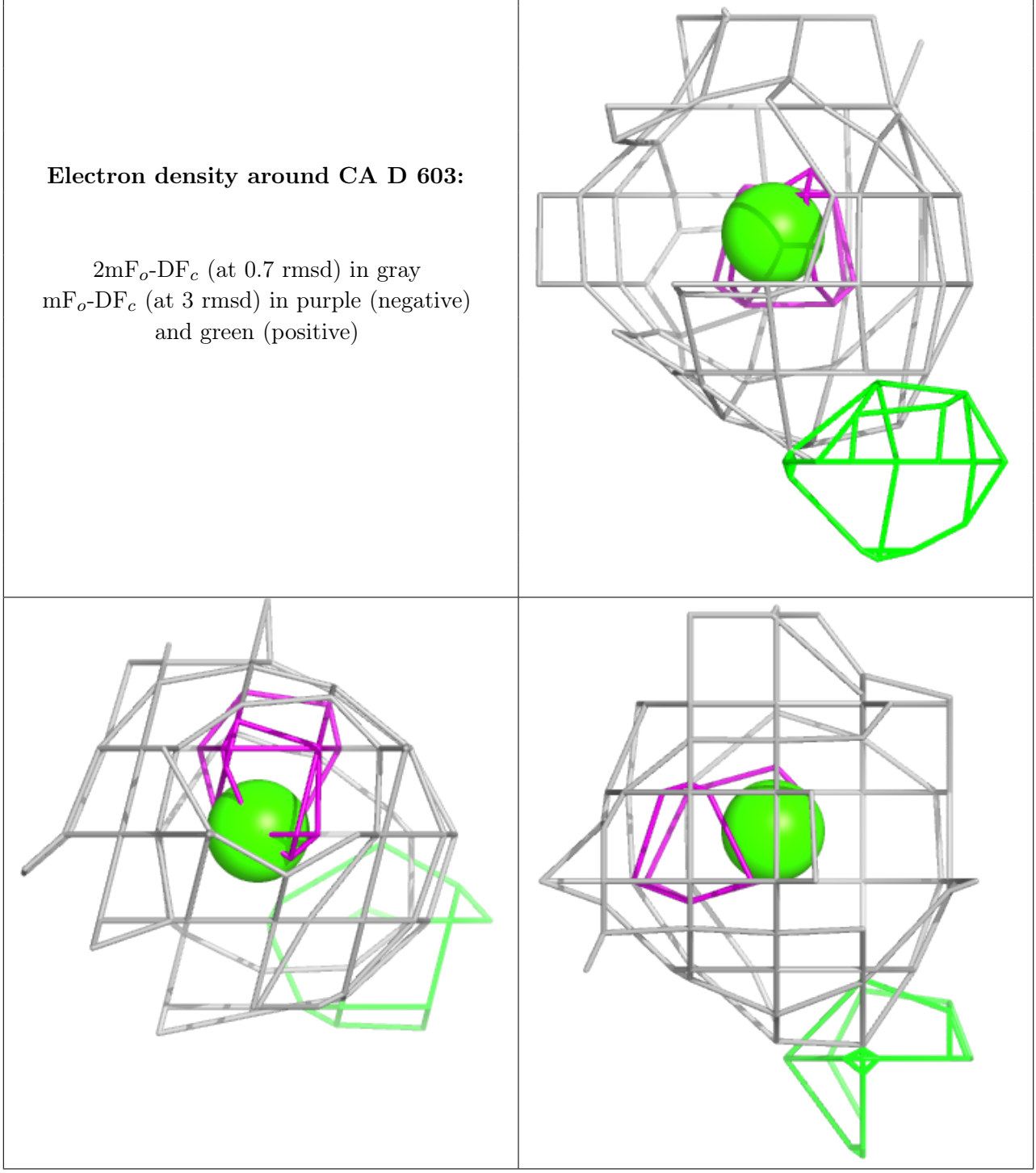






Electron density around CA D 603:

$2mF_o$-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.