



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2023 – 06:31 pm GMT

PDB ID : 8BLS
Title : Structure of Lactobacillus salivarius (Ls) bile salt hydrolase(BSH) in complex with Glycocholate (GCA)
Authors : Karlov, D.S.; Long, S.L.; Zeng, X.; Xu, F.; Lal, K.; Cao, L.; Hayoun, K.; Lin, J.; Joyce, S.A.; Tikhonova, I.G.
Deposited on : 2022-11-10
Resolution : 2.10 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.32.1
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) : 2.32.1

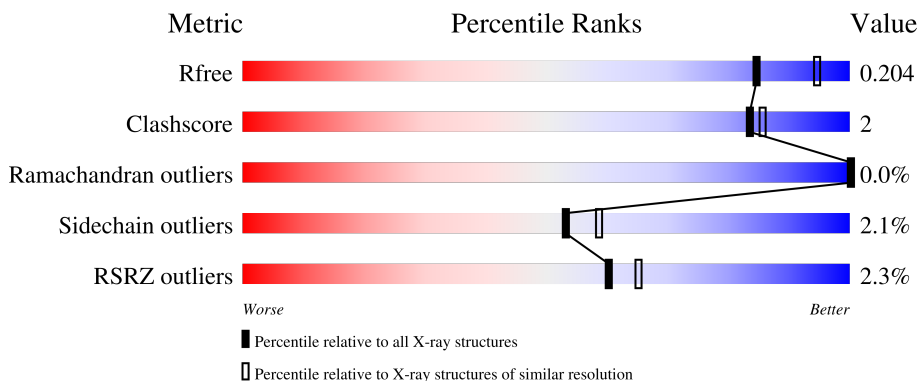
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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.



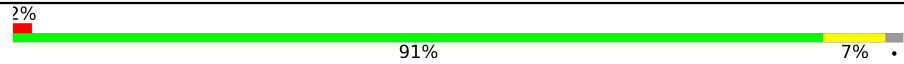
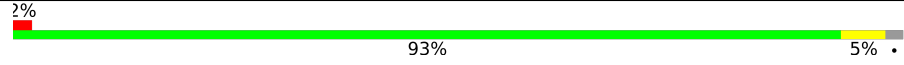
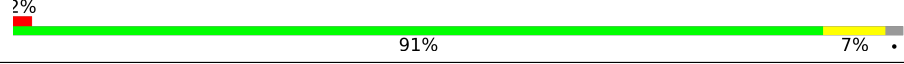
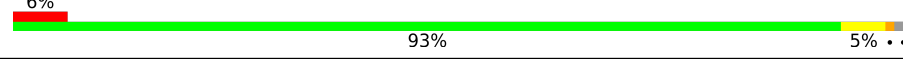
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	325	 93% 6% .
1	B	325	 88% 9% .
1	C	325	 94% 5% ..
1	D	325	 90% 7% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	325	 2% 91% 7% •
1	F	325	 2% 93% 5% •
1	G	325	 2% 91% 7% •
1	H	325	 6% 93% 5% ••

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 22113 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 Bile salt hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	321	Total 2558	C 1628	N 420	O 504	S 6	0	0	0
1	B	318	Total 2537	C 1615	N 420	O 496	S 6	0	1	0
1	C	322	Total 2545	C 1619	N 417	O 503	S 6	0	0	0
1	D	318	Total 2539	C 1616	N 417	O 500	S 6	0	1	0
1	E	320	Total 2526	C 1609	N 414	O 497	S 6	0	0	0
1	F	318	Total 2513	C 1601	N 411	O 495	S 6	0	0	0
1	G	318	Total 2527	C 1610	N 413	O 498	S 6	0	0	0
1	H	319	Total 2487	C 1587	N 407	O 487	S 6	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

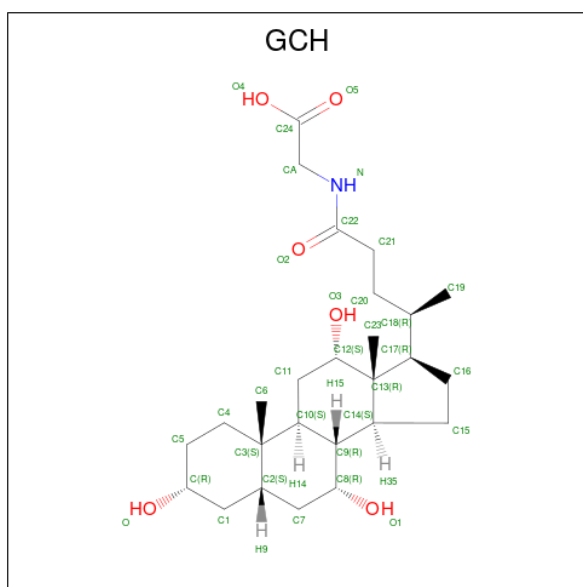
Chain	Residue	Modelled	Actual	Comment	Reference
A	325	LEU	-	expression tag	UNP J7H3P9
A	326	GLU	-	expression tag	UNP J7H3P9
B	325	LEU	-	expression tag	UNP J7H3P9
B	326	GLU	-	expression tag	UNP J7H3P9
C	325	LEU	-	expression tag	UNP J7H3P9
C	326	GLU	-	expression tag	UNP J7H3P9
D	325	LEU	-	expression tag	UNP J7H3P9
D	326	GLU	-	expression tag	UNP J7H3P9
E	325	LEU	-	expression tag	UNP J7H3P9
E	326	GLU	-	expression tag	UNP J7H3P9
F	325	LEU	-	expression tag	UNP J7H3P9
F	326	GLU	-	expression tag	UNP J7H3P9
G	325	LEU	-	expression tag	UNP J7H3P9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	326	GLU	-	expression tag	UNP J7H3P9
H	325	LEU	-	expression tag	UNP J7H3P9
H	326	GLU	-	expression tag	UNP J7H3P9

- Molecule 2 is GLYCOCHOLIC ACID (three-letter code: GCH) (formula: C₂₆H₄₃NO₆) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total	C	N	O	0	0
			33	26	1	6		
2	B	1	Total	C	N	O	0	0
			33	26	1	6		
2	C	1	Total	C	N	O	0	0
			33	26	1	6		
2	D	1	Total	C	N	O	0	0
			33	26	1	6		
2	E	1	Total	C	N	O	0	0
			33	26	1	6		
2	F	1	Total	C	N	O	0	0
			33	26	1	6		
2	G	1	Total	C	N	O	0	0
			33	26	1	6		
2	H	1	Total	C	N	O	0	0
			33	26	1	6		
2	H	1	Total	C	N	O	0	0
			33	26	1	6		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	247	Total O 248 248	0	1
3	B	198	Total O 201 201	0	3
3	C	185	Total O 186 186	0	1
3	D	225	Total O 227 227	0	2
3	E	162	Total O 163 163	0	1
3	F	196	Total O 197 197	0	1
3	G	192	Total O 194 194	0	2
3	H	167	Total O 168 168	0	1

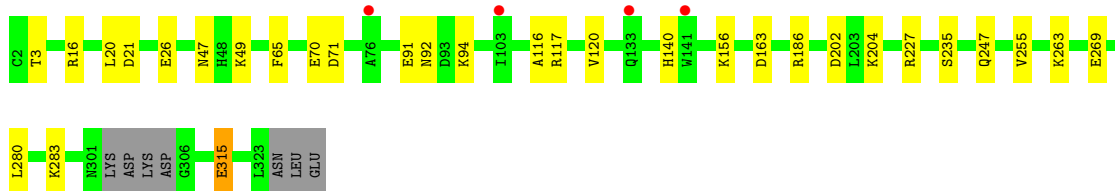
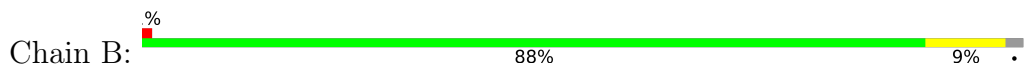
3 Residue-property plots [i](#)

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.

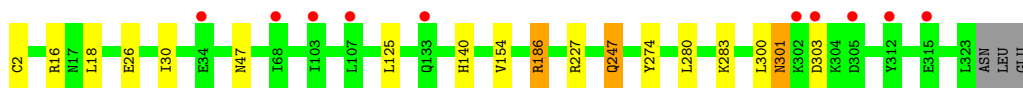
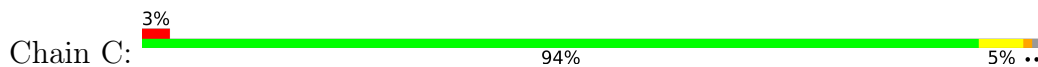
- Molecule 1: Bile salt hydrolase



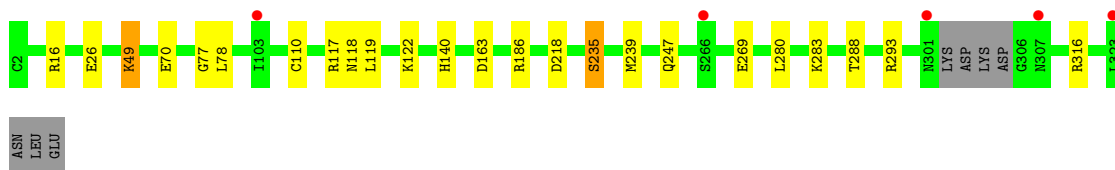
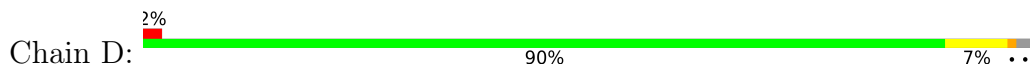
- Molecule 1: Bile salt hydrolase



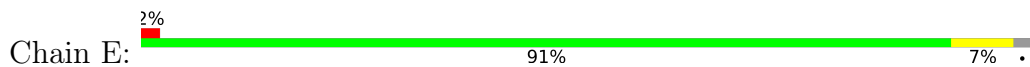
- Molecule 1: Bile salt hydrolase

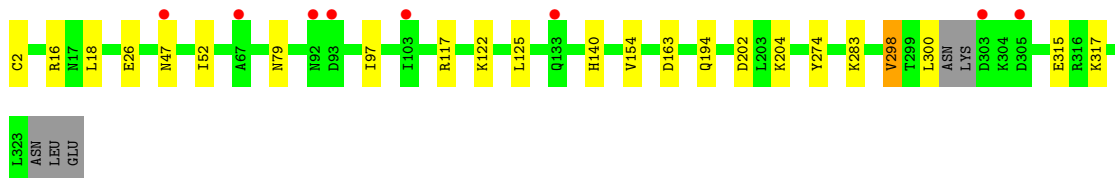


- Molecule 1: Bile salt hydrolase



- Molecule 1: Bile salt hydrolase

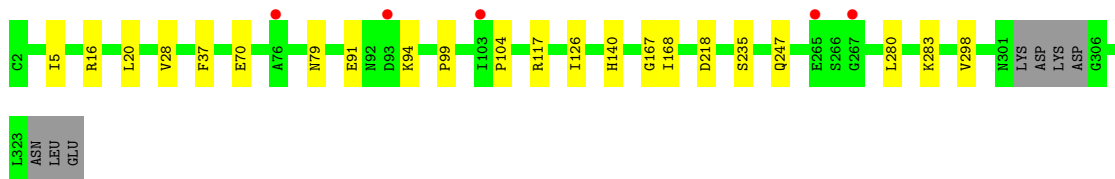
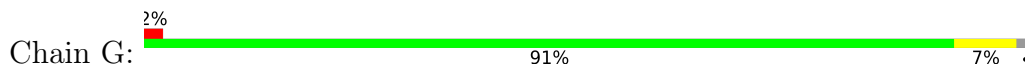




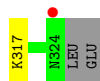
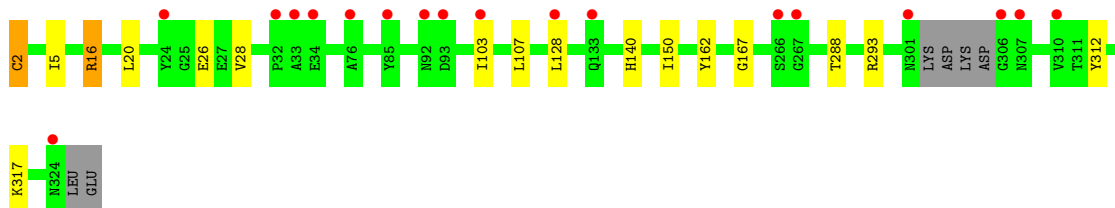
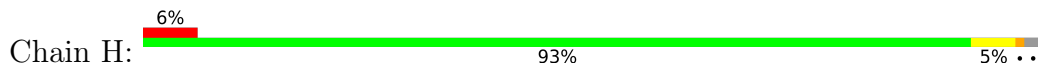
• Molecule 1: Bile salt hydrolase



• Molecule 1: Bile salt hydrolase



• Molecule 1: Bile salt hydrolase



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	84.01Å 94.09Å 166.97Å 90.00° 90.64° 90.00°	Depositor
Resolution (Å)	28.40 – 2.10 28.39 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.6 (28.40-2.10) 94.6 (28.39-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	10.19 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.147 , 0.199 0.157 , 0.204	Depositor DCC
R_{free} test set	7902 reflections (5.52%)	wwPDB-VP
Wilson B-factor (Å ²)	24.7	Xtrriage
Anisotropy	0.050	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 53.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.017 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	22113	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: OCS, GCH

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	0.75	0/2602	0.87	1/3529 (0.0%)
1	B	0.71	0/2587	0.86	1/3510 (0.0%)
1	C	0.74	0/2590	0.85	2/3518 (0.1%)
1	D	0.72	0/2583	0.85	1/3503 (0.0%)
1	E	0.72	0/2570	0.85	0/3487
1	F	0.72	0/2557	0.85	0/3471
1	G	0.73	0/2571	0.85	0/3488
1	H	0.71	0/2531	0.84	0/3442
All	All	0.73	0/20591	0.85	5/27948 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	186	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	B	227	ARG	NE-CZ-NH2	5.80	123.20	120.30
1	C	227	ARG	NE-CZ-NH2	5.60	123.10	120.30
1	A	117	ARG	NE-CZ-NH1	-5.33	117.63	120.30
1	D	293	ARG	NE-CZ-NH2	-5.07	117.77	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	2558	0	2485	8	0
1	B	2537	0	2462	18	0
1	C	2545	0	2447	8	0
1	D	2539	0	2465	14	0
1	E	2526	0	2436	10	0
1	F	2513	0	2422	6	0
1	G	2527	0	2447	12	0
1	H	2487	0	2369	8	0
2	A	33	0	42	0	0
2	B	33	0	42	2	0
2	C	33	0	42	1	0
2	D	33	0	42	1	0
2	E	33	0	42	2	0
2	F	33	0	42	4	0
2	G	33	0	42	5	0
2	H	66	0	84	3	0
3	A	248	0	0	0	0
3	B	201	0	0	3	0
3	C	186	0	0	3	0
3	D	227	0	0	4	0
3	E	163	0	0	1	0
3	F	197	0	0	1	0
3	G	194	0	0	1	0
3	H	168	0	0	2	0
All	All	22113	0	19911	97	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 2.

All (97) 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:G:20:LEU:HD13	2:G:401:GCH:H19	1.54	0.89
1:B:235:SER:HB2	1:B:247:GLN:HE21	1.41	0.84
2:G:401:GCH:H20	2:G:401:GCH:H38	1.61	0.82
1:D:70:GLU:O	1:D:283:LYS:HE2	1.80	0.81
1:A:235:SER:HB2	1:A:247:GLN:HE21	1.53	0.73
1:F:70:GLU:O	1:F:283:LYS:HE2	1.94	0.68
1:E:79:ASN:HD21	2:E:401:GCH:H28	1.64	0.62
2:F:401:GCH:H19	2:F:401:GCH:O2	2.00	0.61
1:F:280:LEU:O	1:F:283:LYS:HD3	2.03	0.58
1:B:70:GLU:O	1:B:283:LYS:HE2	2.04	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:401:GCH:H20	2:D:401:GCH:H40	1.86	0.57
3:C:566:HOH:O	1:D:269:GLU:HG3	2.04	0.57
1:D:117:ARG:NH2	1:D:163:ASP:OD1	2.39	0.56
1:G:235:SER:HB2	1:G:247:GLN:HE21	1.72	0.54
2:B:401:GCH:H20	2:B:401:GCH:H40	1.88	0.54
1:G:28:VAL:HG11	1:G:298:VAL:HG11	1.89	0.54
1:B:91:GLU:OE1	1:B:94:LYS:HG3	2.08	0.53
2:F:401:GCH:H20	2:F:401:GCH:H40	1.89	0.53
1:A:70:GLU:O	1:A:283:LYS:HE2	2.08	0.53
1:C:125:LEU:HD12	1:C:154:VAL:HG21	1.91	0.53
1:D:235:SER:HB2	1:D:247:GLN:HE21	1.75	0.52
1:D:280:LEU:O	1:D:283:LYS:HD3	2.10	0.51
1:B:49:LYS:NZ	1:B:71:ASP:OD2	2.41	0.51
2:G:401:GCH:N	3:G:502:HOH:O	2.34	0.51
1:D:118:ASN:O	1:D:122:LYS:HE3	2.12	0.50
1:E:52:ILE:HD11	1:E:298:VAL:HG22	1.93	0.50
2:E:401:GCH:H20	2:E:401:GCH:H40	1.93	0.49
2:C:401:GCH:O1	2:C:401:GCH:H10	2.12	0.49
1:B:117[B]:ARG:NH1	1:B:163:ASP:OD1	2.41	0.49
1:H:2:OCS:OD2	1:H:16:ARG:NH2	2.45	0.49
1:D:26:GLU:H	1:D:26:GLU:CD	2.16	0.49
1:B:20:LEU:HD13	2:B:401:GCH:H19	1.94	0.48
1:H:28:VAL:HB	1:H:312:TYR:HB2	1.95	0.48
1:A:2:OCS:OD2	1:A:78:LEU:HA	2.14	0.48
2:F:401:GCH:O2	2:F:401:GCH:C19	2.60	0.48
1:E:125:LEU:HD12	1:E:154:VAL:HG21	1.94	0.48
1:B:280:LEU:O	1:B:283:LYS:HD3	2.13	0.48
1:H:150:ILE:HA	1:H:162:TYR:O	2.13	0.47
1:B:315:GLU:CG	3:B:696:HOH:O	2.62	0.47
1:G:79:ASN:HB2	2:G:401:GCH:H28	1.97	0.47
1:C:26:GLU:H	1:C:26:GLU:CD	2.16	0.47
1:B:315:GLU:HG2	3:B:696:HOH:O	2.14	0.47
1:G:218:ASP:HB2	3:H:1445:HOH:O	2.14	0.47
1:A:280:LEU:O	1:A:283:LYS:HD3	2.15	0.47
1:B:70:GLU:O	1:B:283:LYS:CE	2.63	0.47
1:B:26:GLU:H	1:B:26:GLU:CD	2.18	0.46
1:B:202:ASP:OD2	1:B:204:LYS:HE3	2.16	0.46
1:B:186:ARG:HD2	3:B:544:HOH:O	2.16	0.46
1:D:110:CYS:SG	1:D:119:LEU:HD22	2.55	0.45
2:H:401:GCH:H20	2:H:401:GCH:H40	1.97	0.45
1:E:202:ASP:OD2	1:E:204:LYS:HE2	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:97:ILE:O	1:E:125:LEU:HA	2.17	0.45
1:C:186:ARG:HD2	3:C:576:HOH:O	2.16	0.45
1:A:76:ALA:O	1:A:141:TRP:HA	2.17	0.44
1:E:117:ARG:NH2	1:E:163:ASP:OD1	2.50	0.44
1:H:103:ILE:HG23	1:H:107:LEU:HD12	1.98	0.44
1:D:186:ARG:HD2	3:D:561:HOH:O	2.17	0.44
2:H:401:GCH:H20	2:H:401:GCH:H38	2.00	0.44
1:G:5:ILE:HG22	1:G:168:ILE:HD13	2.00	0.44
1:C:247:GLN:HE21	1:C:247:GLN:HA	1.82	0.44
1:C:18:LEU:HD22	1:C:274:TYR:CE2	2.53	0.44
1:E:52:ILE:CD1	1:E:298:VAL:HG22	2.48	0.44
1:G:70:GLU:O	1:G:283:LYS:HE2	2.18	0.43
2:G:401:GCH:H20	2:G:401:GCH:C23	2.39	0.43
1:G:280:LEU:O	1:G:283:LYS:HE3	2.18	0.43
1:B:3:THR:HG21	1:B:255:VAL:HG21	2.01	0.43
1:A:77:GLY:C	1:A:78:LEU:HD12	2.39	0.43
1:B:65:PHE:N	1:B:65:PHE:CD1	2.86	0.43
1:F:288:THR:HA	3:F:551:HOH:O	2.17	0.43
1:G:99:PRO:HD3	1:G:126:ILE:O	2.19	0.43
1:E:18:LEU:HD22	1:E:274:TYR:CE2	2.54	0.42
1:H:20:LEU:HD13	2:H:401:GCH:H19	2.01	0.42
1:H:26:GLU:H	1:H:26:GLU:CD	2.23	0.42
1:C:30:ILE:HD13	1:C:303:ASP:HB2	2.02	0.42
1:D:316:ARG:NH2	3:D:519:HOH:O	2.52	0.42
1:G:5:ILE:HA	1:G:167:GLY:O	2.19	0.42
1:G:37:PHE:CE1	1:G:104:PRO:HB3	2.55	0.42
1:G:91:GLU:HB2	1:G:94:LYS:HB2	2.00	0.42
1:H:288:THR:HA	3:H:1403:HOH:O	2.19	0.42
1:B:263:LYS:HE2	1:B:269:GLU:OE2	2.20	0.42
1:C:300:LEU:O	1:C:301:ASN:C	2.58	0.42
3:C:554:HOH:O	1:D:218:ASP:HB2	2.19	0.42
1:H:5:ILE:HA	1:H:167:GLY:O	2.20	0.42
1:D:288:THR:HA	3:D:621:HOH:O	2.20	0.41
1:D:77:GLY:C	1:D:78:LEU:HD12	2.41	0.41
1:F:263:LYS:HE3	1:F:269:GLU:OE2	2.21	0.41
2:F:401:GCH:H19	2:F:401:GCH:C22	2.49	0.41
1:A:78:LEU:HD12	1:A:78:LEU:N	2.36	0.41
1:B:65:PHE:N	1:B:65:PHE:HD1	2.19	0.41
1:F:235:SER:HB2	1:F:247:GLN:OE1	2.20	0.41
1:C:280:LEU:O	1:C:283:LYS:HE3	2.20	0.41
1:F:97:ILE:O	1:F:125:LEU:HA	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:26:GLU:H	1:E:26:GLU:CD	2.25	0.41
1:A:202:ASP:OD2	1:A:204:LYS:HE3	2.21	0.40
1:B:116:ALA:O	1:B:120:VAL:HG23	2.22	0.40
1:D:49:LYS:NZ	3:D:510:HOH:O	2.45	0.40
1:E:122:LYS:NZ	3:E:509:HOH:O	2.50	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	A	317/325 (98%)	311 (98%)	6 (2%)	0	100	100
1	B	316/325 (97%)	309 (98%)	7 (2%)	0	100	100
1	C	320/325 (98%)	313 (98%)	6 (2%)	1 (0%)	41	41
1	D	315/325 (97%)	307 (98%)	8 (2%)	0	100	100
1	E	316/325 (97%)	305 (96%)	11 (4%)	0	100	100
1	F	314/325 (97%)	307 (98%)	7 (2%)	0	100	100
1	G	314/325 (97%)	307 (98%)	7 (2%)	0	100	100
1	H	315/325 (97%)	307 (98%)	8 (2%)	0	100	100
All	All	2527/2600 (97%)	2466 (98%)	60 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	301	ASN

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	282/287 (98%)	275 (98%)	7 (2%)	47	52
1	B	278/287 (97%)	271 (98%)	7 (2%)	47	52
1	C	277/287 (96%)	273 (99%)	4 (1%)	67	73
1	D	280/287 (98%)	275 (98%)	5 (2%)	59	65
1	E	275/287 (96%)	266 (97%)	9 (3%)	38	40
1	F	274/287 (96%)	268 (98%)	6 (2%)	52	57
1	G	277/287 (96%)	274 (99%)	3 (1%)	73	79
1	H	265/287 (92%)	260 (98%)	5 (2%)	57	63
All	All	2208/2296 (96%)	2162 (98%)	46 (2%)	53	59

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	ARG
1	A	140	HIS
1	A	184	LYS
1	A	191	SER
1	A	232	LYS
1	A	237	LYS
1	A	326	GLU
1	B	16	ARG
1	B	21	ASP
1	B	47	ASN
1	B	92	ASN
1	B	140	HIS
1	B	156	LYS
1	B	315	GLU
1	C	16	ARG
1	C	47	ASN
1	C	140	HIS
1	C	247	GLN
1	D	16	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	49	LYS
1	D	140	HIS
1	D	235	SER
1	D	239	MET
1	E	16	ARG
1	E	47	ASN
1	E	140	HIS
1	E	194	GLN
1	E	283	LYS
1	E	298	VAL
1	E	300	LEU
1	E	315	GLU
1	E	317	LYS
1	F	16	ARG
1	F	140	HIS
1	F	148	LYS
1	F	156	LYS
1	F	204	LYS
1	F	265	GLU
1	G	16	ARG
1	G	117	ARG
1	G	140	HIS
1	H	16	ARG
1	H	128	LEU
1	H	140	HIS
1	H	293	ARG
1	H	317	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	247	GLN
1	B	247	GLN
1	B	301	ASN
1	C	10	ASN
1	C	133	GLN
1	D	133	GLN
1	E	79	ASN
1	E	133	GLN
1	F	79	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

8 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	OCS	B	2	1	7,8,9	0.95	0	6,11,13	0.86	0
1	OCS	A	2	1	7,8,9	0.70	0	6,11,13	1.13	1 (16%)
1	OCS	H	2	1	7,8,9	0.87	0	6,11,13	1.62	2 (33%)
1	OCS	G	2	1	7,8,9	0.92	0	6,11,13	1.20	0
1	OCS	E	2	1	7,8,9	1.18	1 (14%)	6,11,13	2.53	2 (33%)
1	OCS	F	2	1	7,8,9	1.44	1 (14%)	6,11,13	1.25	1 (16%)
1	OCS	C	2	1	7,8,9	0.75	0	6,11,13	1.18	1 (16%)
1	OCS	D	2	1	7,8,9	0.86	0	6,11,13	1.33	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
1	OCS	B	2	1	-	0/4/7/9	-
1	OCS	A	2	1	-	0/4/7/9	-
1	OCS	H	2	1	-	0/4/7/9	-
1	OCS	G	2	1	-	0/4/7/9	-
1	OCS	E	2	1	-	0/4/7/9	-
1	OCS	F	2	1	-	0/4/7/9	-
1	OCS	C	2	1	-	0/4/7/9	-
1	OCS	D	2	1	-	0/4/7/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	2	OCS	OD1-SG	3.15	1.54	1.45
1	E	2	OCS	OD3-SG	2.79	1.53	1.45

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	2	OCS	OD3-SG-CB	-5.50	100.40	106.94
1	H	2	OCS	OD2-SG-CB	-2.33	102.02	105.74
1	H	2	OCS	OD3-SG-OD1	2.25	121.72	113.95
1	E	2	OCS	OD2-SG-CB	2.17	109.20	105.74
1	F	2	OCS	OD2-SG-OD3	2.17	116.57	111.27
1	C	2	OCS	OD2-SG-OD3	2.15	116.54	111.27
1	A	2	OCS	OD3-SG-CB	2.07	109.41	106.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	2	OCS	1	0
1	H	2	OCS	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

9 ligands are modelled in this entry.

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	GCH	G	401	-	36,36,36	1.10	4 (11%)	56,56,56	1.39	7 (12%)
2	GCH	A	401	-	36,36,36	0.61	0	56,56,56	0.93	2 (3%)
2	GCH	B	401	-	36,36,36	0.59	0	56,56,56	0.85	1 (1%)
2	GCH	H	402	-	36,36,36	0.60	0	56,56,56	0.80	0
2	GCH	H	401	-	36,36,36	0.50	0	56,56,56	0.77	1 (1%)
2	GCH	F	401	-	36,36,36	0.65	0	56,56,56	1.09	2 (3%)
2	GCH	E	401	-	36,36,36	0.60	0	56,56,56	1.16	4 (7%)
2	GCH	D	401	-	36,36,36	0.49	0	56,56,56	0.90	2 (3%)
2	GCH	C	401	-	36,36,36	0.67	0	56,56,56	1.06	4 (7%)

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	GCH	G	401	-	-	5/14/79/79	0/4/4/4
2	GCH	A	401	-	-	5/14/79/79	0/4/4/4
2	GCH	B	401	-	-	2/14/79/79	0/4/4/4
2	GCH	H	402	-	-	0/14/79/79	0/4/4/4
2	GCH	H	401	-	-	1/14/79/79	0/4/4/4
2	GCH	F	401	-	-	5/14/79/79	0/4/4/4
2	GCH	E	401	-	-	5/14/79/79	0/4/4/4
2	GCH	D	401	-	-	3/14/79/79	0/4/4/4
2	GCH	C	401	-	-	0/14/79/79	0/4/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	401	GCH	O5-C24	2.70	1.31	1.22
2	G	401	GCH	C20-C21	2.48	1.60	1.52
2	G	401	GCH	CA-C24	2.44	1.56	1.51
2	G	401	GCH	O4-C24	-2.04	1.23	1.30

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	401	GCH	C20-C21-C22	5.11	124.45	113.04
2	F	401	GCH	C13-C17-C18	-4.30	114.36	119.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	401	GCH	C14-C13-C12	3.91	111.04	107.40
2	A	401	GCH	C13-C17-C18	3.17	123.28	119.50
2	G	401	GCH	C20-C18-C17	3.06	116.61	110.28
2	G	401	GCH	C17-C13-C14	3.03	103.15	100.09
2	E	401	GCH	C17-C13-C14	-2.74	97.33	100.09
2	G	401	GCH	C3-C10-C9	2.74	114.76	111.82
2	C	401	GCH	C2-C1-C	2.67	116.69	112.76
2	E	401	GCH	C3-C10-C9	2.52	114.52	111.82
2	C	401	GCH	C1-C2-C3	2.49	115.30	112.66
2	E	401	GCH	C16-C17-C13	-2.47	101.13	103.55
2	D	401	GCH	C13-C17-C18	-2.40	116.63	119.50
2	D	401	GCH	C11-C10-C3	-2.24	111.41	113.73
2	H	401	GCH	C10-C9-C8	2.19	114.49	111.88
2	C	401	GCH	C14-C13-C12	2.18	109.43	107.40
2	C	401	GCH	C3-C10-C9	2.10	114.08	111.82
2	G	401	GCH	C16-C17-C13	-2.10	101.50	103.55
2	G	401	GCH	C19-C18-C20	-2.09	107.08	110.36
2	A	401	GCH	C19-C18-C17	2.06	116.08	112.92
2	F	401	GCH	C16-C17-C18	2.05	115.32	112.15
2	B	401	GCH	C11-C12-C13	2.05	113.35	111.24
2	G	401	GCH	C6-C3-C4	-2.05	104.96	108.26

There are no chirality outliers.

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	401	GCH	C18-C20-C21-C22
2	E	401	GCH	C21-C22-N-CA
2	F	401	GCH	C21-C22-N-CA
2	E	401	GCH	O2-C22-N-CA
2	F	401	GCH	O2-C22-N-CA
2	G	401	GCH	C13-C17-C18-C19
2	E	401	GCH	C17-C18-C20-C21
2	G	401	GCH	C16-C17-C18-C20
2	A	401	GCH	C13-C17-C18-C20
2	E	401	GCH	C19-C18-C20-C21
2	A	401	GCH	C16-C17-C18-C20
2	A	401	GCH	C16-C17-C18-C19
2	B	401	GCH	C18-C20-C21-C22
2	D	401	GCH	C18-C20-C21-C22
2	H	401	GCH	C18-C20-C21-C22
2	G	401	GCH	C16-C17-C18-C19

Continued on next page...

Continued from previous page...

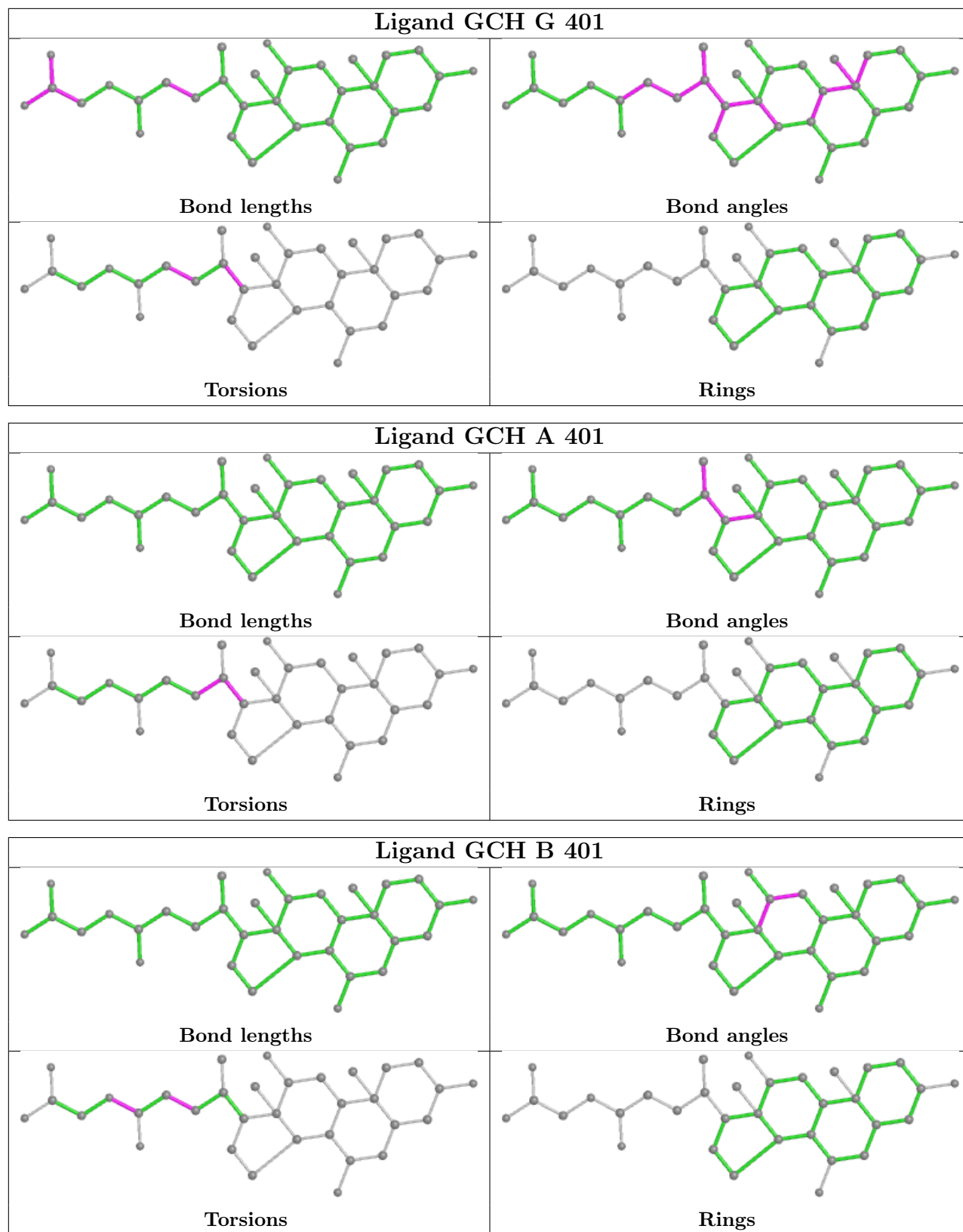
Mol	Chain	Res	Type	Atoms
2	A	401	GCH	C19-C18-C20-C21
2	G	401	GCH	C18-C20-C21-C22
2	D	401	GCH	O5-C24-CA-N
2	D	401	GCH	O4-C24-CA-N
2	F	401	GCH	O4-C24-CA-N
2	F	401	GCH	O5-C24-CA-N
2	G	401	GCH	C13-C17-C18-C20
2	A	401	GCH	C13-C17-C18-C19
2	F	401	GCH	C18-C20-C21-C22
2	B	401	GCH	O2-C22-N-CA

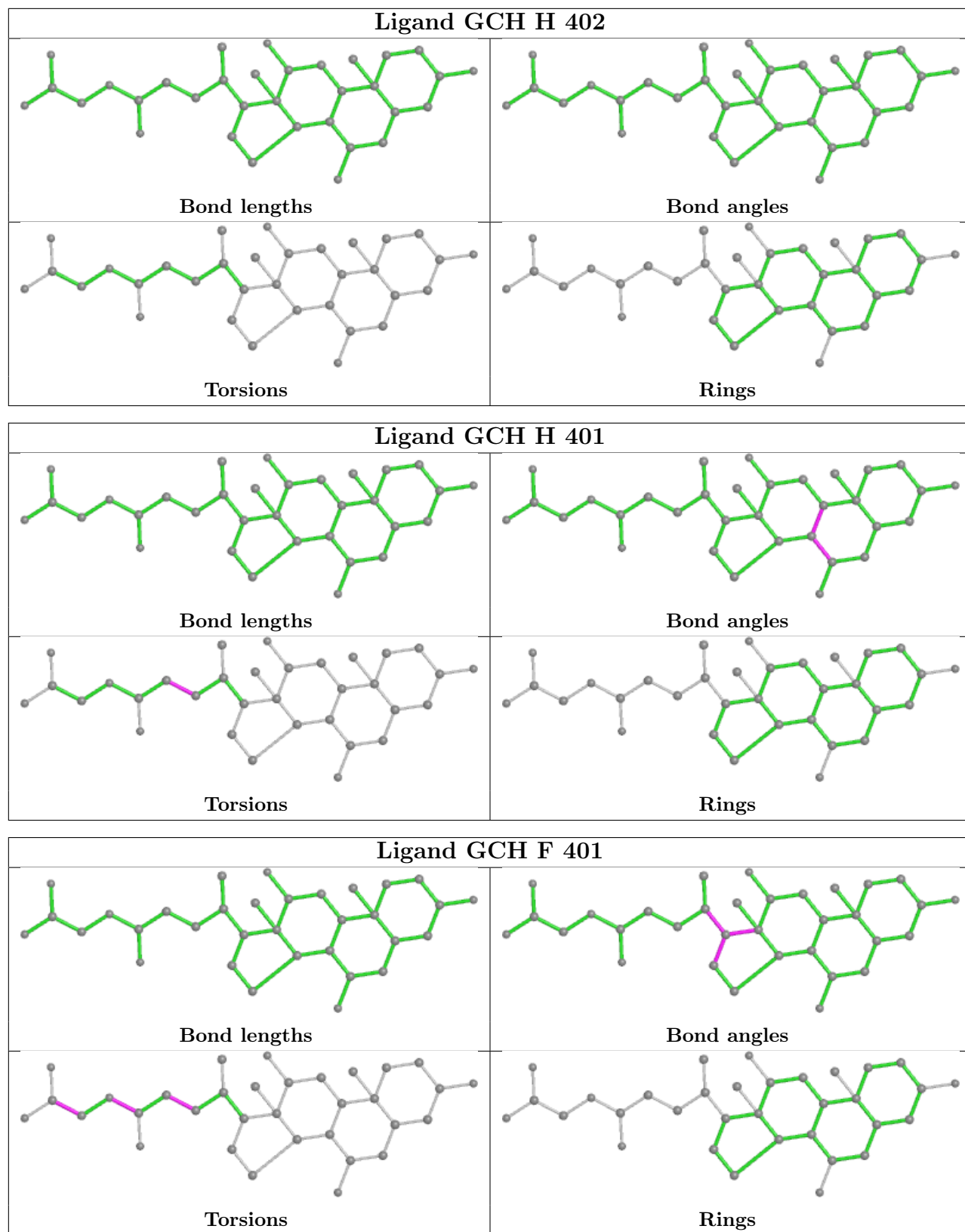
There are no ring outliers.

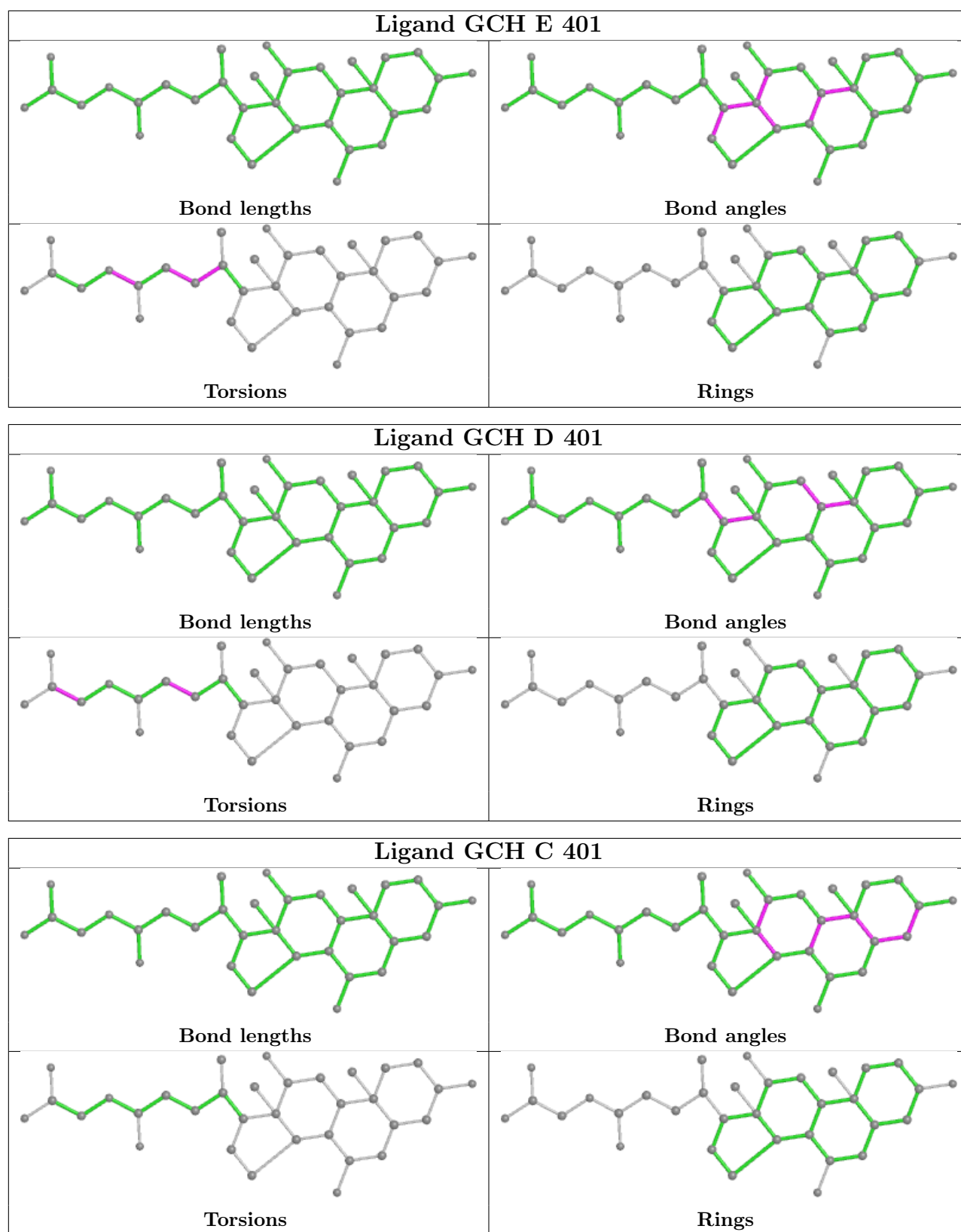
7 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	401	GCH	5	0
2	B	401	GCH	2	0
2	H	401	GCH	3	0
2	F	401	GCH	4	0
2	E	401	GCH	2	0
2	D	401	GCH	1	0
2	C	401	GCH	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 than 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

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(Å ²)	Q<0.9
1	A	320/325 (98%)	-0.36	2 (0%) 89 91	15, 21, 37, 49	0
1	B	317/325 (97%)	-0.21	4 (1%) 77 80	15, 25, 44, 53	0
1	C	321/325 (98%)	-0.16	10 (3%) 49 55	16, 26, 45, 67	0
1	D	317/325 (97%)	-0.29	5 (1%) 72 75	15, 23, 43, 60	0
1	E	319/325 (98%)	-0.13	8 (2%) 57 62	17, 27, 48, 64	0
1	F	317/325 (97%)	-0.27	6 (1%) 66 71	15, 24, 42, 61	0
1	G	317/325 (97%)	-0.22	5 (1%) 72 75	18, 28, 44, 59	0
1	H	318/325 (97%)	0.09	18 (5%) 23 29	17, 32, 53, 62	0
All	All	2546/2600 (97%)	-0.20	58 (2%) 60 65	15, 25, 46, 67	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	305	ASP	4.6
1	C	305	ASP	3.8
1	C	303	ASP	3.3
1	H	310	VAL	3.3
1	E	92	ASN	3.2
1	H	266	SER	3.1
1	H	85	TYR	3.0
1	F	47	ASN	2.9
1	G	267	GLY	2.8
1	E	103	ILE	2.8
1	B	133	GLN	2.8
1	H	324	ASN	2.7
1	B	103	ILE	2.7
1	H	301	ASN	2.7
1	F	306	GLY	2.6
1	D	323	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	107	LEU	2.6
1	H	76	ALA	2.5
1	H	32	PRO	2.4
1	F	265	GLU	2.4
1	D	103	ILE	2.4
1	E	133	GLN	2.4
1	F	301	ASN	2.4
1	H	33	ALA	2.4
1	H	103	ILE	2.3
1	F	103	ILE	2.3
1	E	303	ASP	2.3
1	C	312	TYR	2.3
1	D	301	ASN	2.3
1	H	92	ASN	2.3
1	C	103	ILE	2.3
1	H	133	GLN	2.3
1	B	141	TRP	2.3
1	F	76	ALA	2.2
1	E	67	ALA	2.2
1	H	128	LEU	2.2
1	H	93	ASP	2.2
1	A	266	SER	2.2
1	C	68	ILE	2.1
1	E	93	ASP	2.1
1	E	47	ASN	2.1
1	G	93	ASP	2.1
1	C	302	LYS	2.1
1	G	103	ILE	2.1
1	C	315	GLU	2.1
1	D	266	SER	2.1
1	H	34	GLU	2.1
1	C	133	GLN	2.1
1	H	307	ASN	2.1
1	G	76	ALA	2.1
1	H	24	TYR	2.1
1	D	307	ASN	2.0
1	H	306	GLY	2.0
1	B	76	ALA	2.0
1	G	265	GLU	2.0
1	C	34	GLU	2.0
1	A	103	ILE	2.0
1	H	267	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
1	OCS	C	2	9/10	0.96	0.12	17,22,34,41	0
1	OCS	G	2	9/10	0.97	0.10	21,25,35,38	0
1	OCS	H	2	9/10	0.97	0.09	21,28,32,38	0
1	OCS	F	2	9/10	0.98	0.09	18,24,35,37	0
1	OCS	A	2	9/10	0.98	0.09	18,23,29,32	0
1	OCS	E	2	9/10	0.98	0.09	17,25,35,36	0
1	OCS	B	2	9/10	0.99	0.07	18,26,34,35	0
1	OCS	D	2	9/10	0.99	0.10	17,23,31,35	0

6.3 Carbohydrates [i](#)

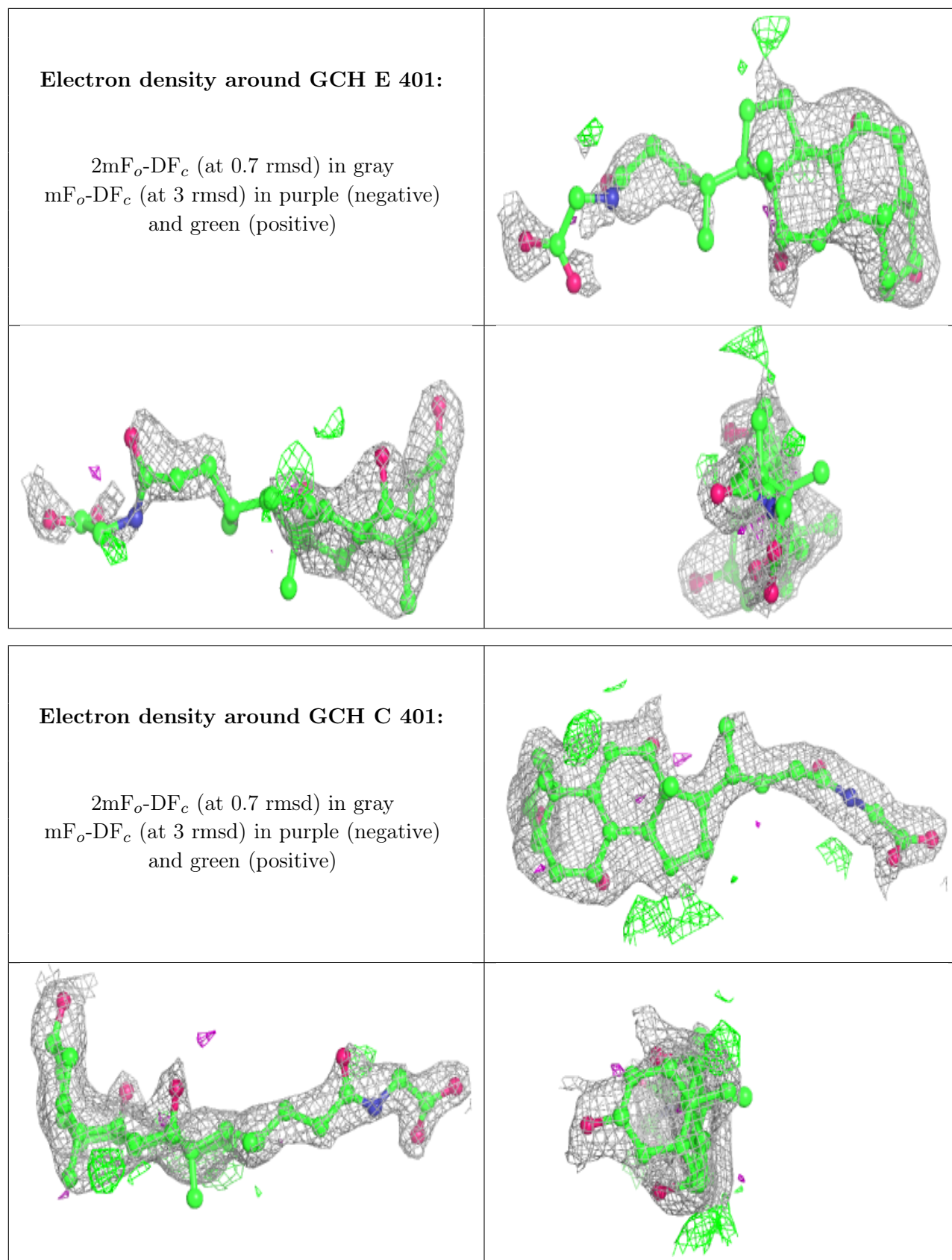
There are no monosaccharides in this entry.

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

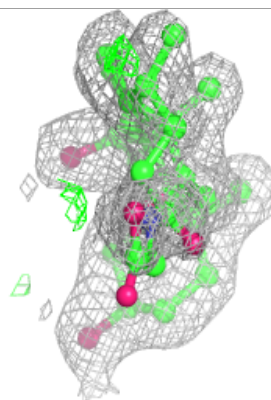
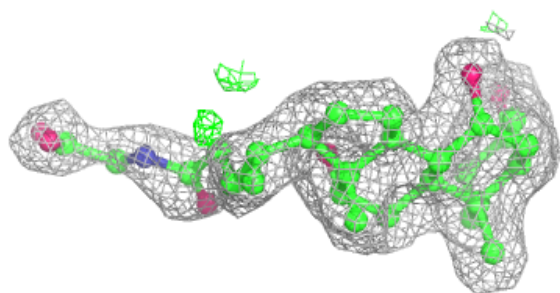
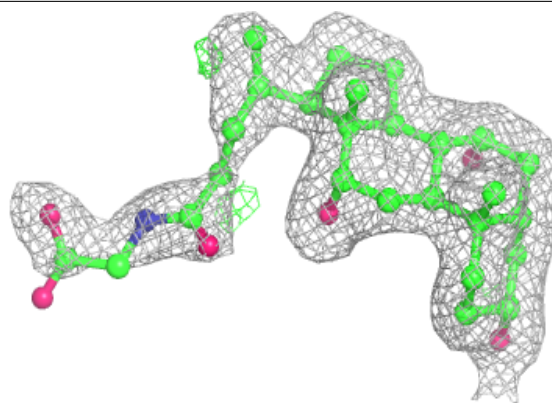
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	GCH	E	401	33/33	0.70	0.23	45,70,81,83	0
2	GCH	C	401	33/33	0.79	0.17	41,55,66,70	0
2	GCH	A	401	33/33	0.85	0.15	20,23,58,67	0
2	GCH	H	401	33/33	0.85	0.14	38,45,74,76	0
2	GCH	G	401	33/33	0.87	0.15	29,36,74,83	0
2	GCH	H	402	33/33	0.87	0.13	24,30,40,44	0
2	GCH	F	401	33/33	0.88	0.13	23,26,66,73	0
2	GCH	D	401	33/33	0.88	0.14	23,29,55,62	0
2	GCH	B	401	33/33	0.89	0.13	26,33,66,76	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.

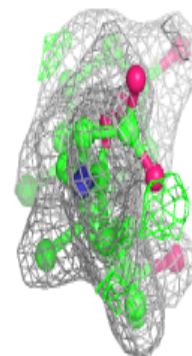
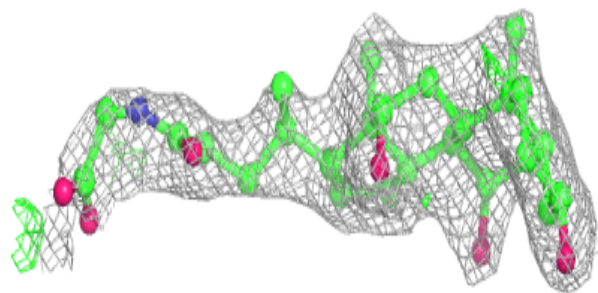
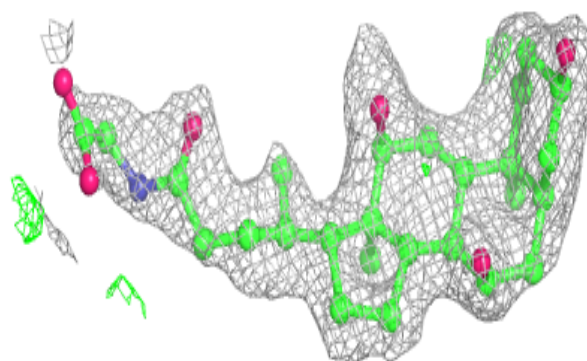


Electron density around GCH A 401:

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

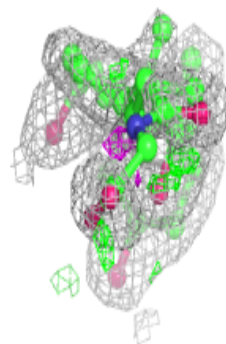
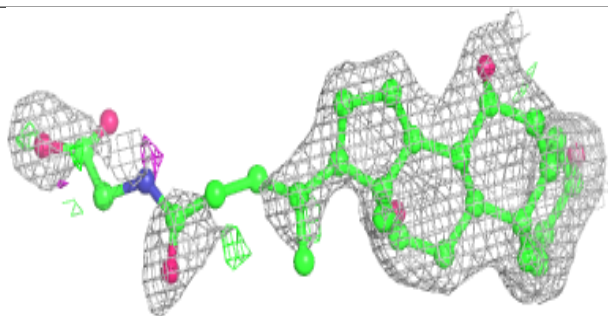
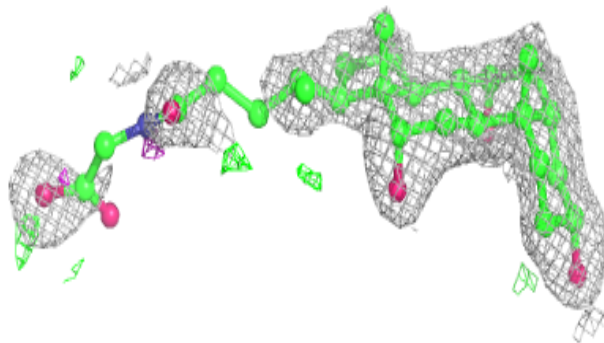
**Electron density around GCH H 401:**

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

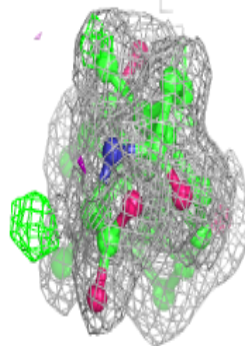
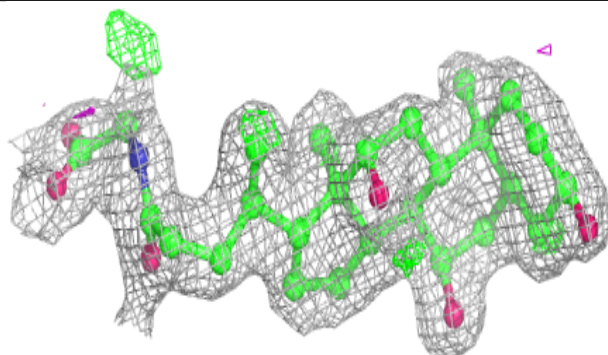
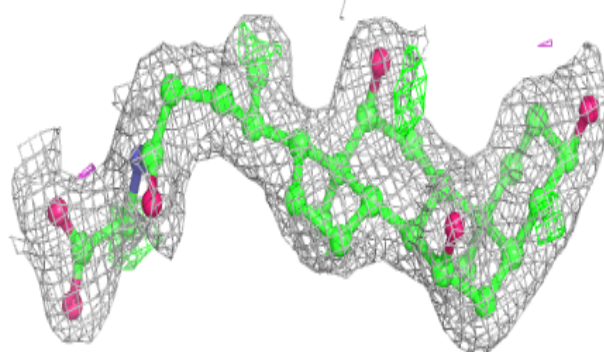


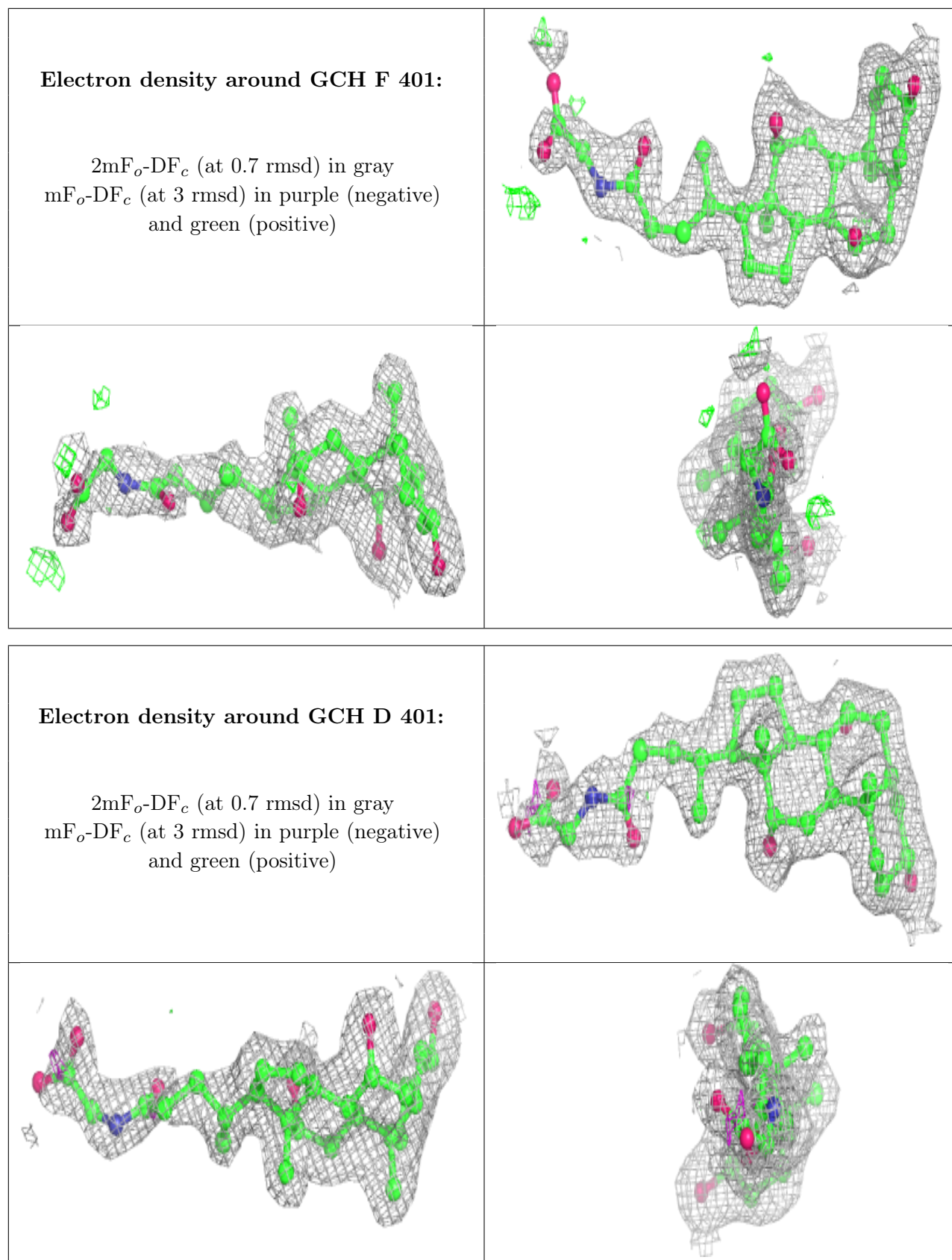
Electron density around GCH G 401:

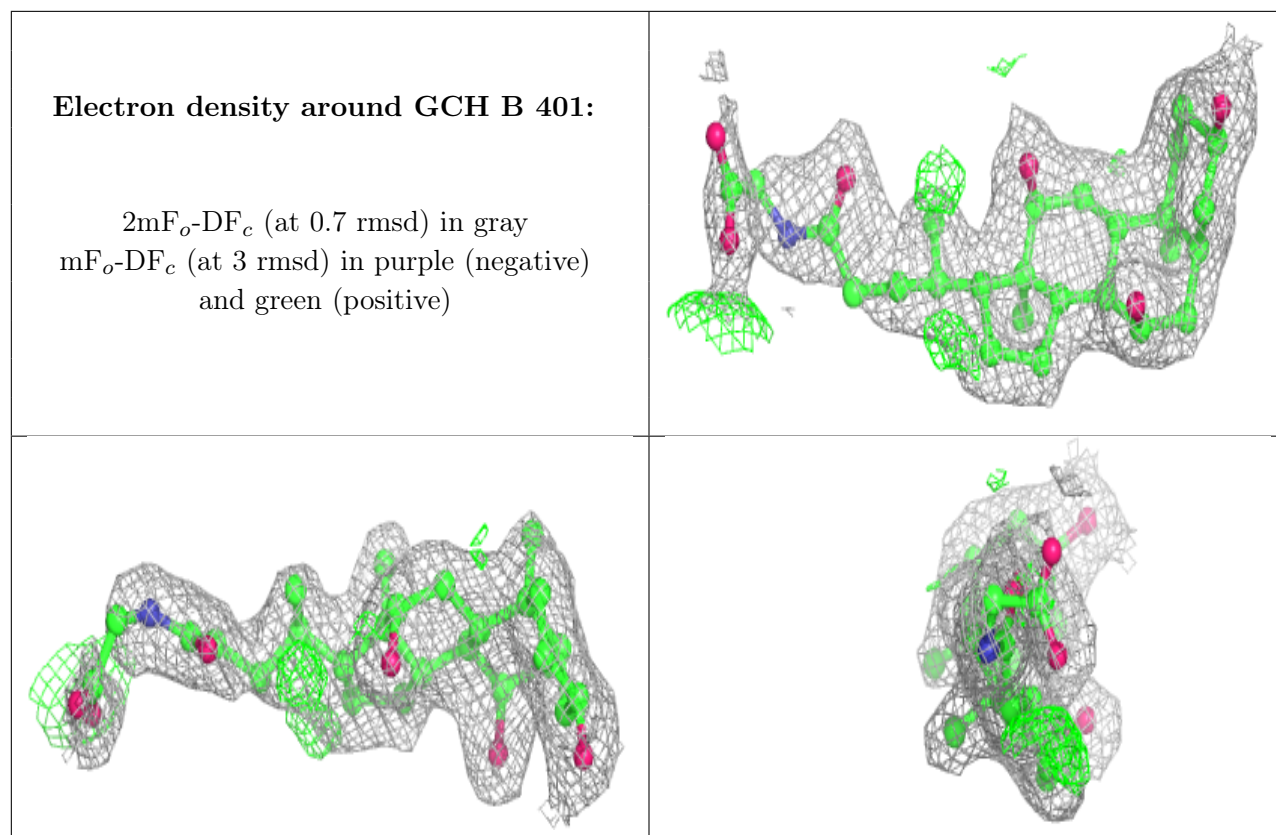
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around GCH H 402:**

$2mF_o-DF_c$ (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.