WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Nov 6, 2023 – 09:05 am GMT

PDB ID	:	7ZVE
Title	:	K403 acetylated glucose-6-phosphate dehydrogenase (G6PD)
Authors	:	Wu, F.; Muskat, N.H.; Shahar, A.; Arbely, E.
Deposited on	:	2022-05-15
Resolution	:	2.28 Å(reported)

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	:	2.36
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 2.28 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.
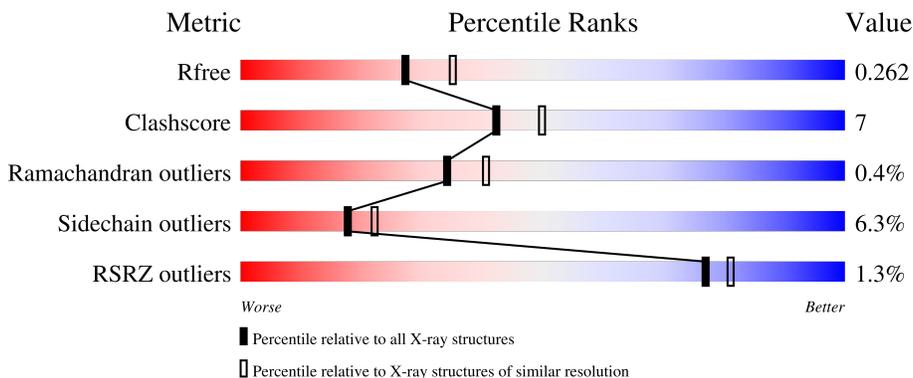
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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

Mol	Chain	Length	Quality of chain
1	A	501	2% / 79% / 16% / . .
2	B	498	.% / 77% / 18% / 5% .
3	C	497	.% / 79% / 15% / . .
4	D	499	.% / 79% / 16% / . .
4	H	499	.% / 80% / 15% / 5% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	500	
6	F	497	
7	G	496	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ALY	B	403	-	-	X	-
8	GOL	B	602	-	-	X	-
8	GOL	E	601	-	-	X	-

2 Entity composition (i)

There are 10 unique types of molecules in this entry. The entry contains 32772 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 Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	501	Total	C	N	O	S	0	0	0
			4053	2582	706	744	21			

- Molecule 2 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	498	Total	C	N	O	S	0	0	0
			4036	2572	703	740	21			

- Molecule 3 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	497	Total	C	N	O	S	0	1	0
			4039	2574	705	739	21			

- Molecule 4 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	499	Total	C	N	O	S	0	1	0
			4051	2582	707	741	21			
4	H	499	Total	C	N	O	S	0	0	0
			4043	2577	704	741	21			

- Molecule 5 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	500	Total	C	N	O	S	0	0	0
			4052	2582	705	744	21			

- Molecule 6 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	497	Total 4028	C 2567	N 702	O 738	S 21	0	0	0

- Molecule 7 is a protein called Glucose-6-phosphate 1-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	496	Total 4023	C 2563	N 701	O 738	S 21	0	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total 6	C 3	O 3	0	0
8	B	1	Total 6	C 3	O 3	0	0
8	C	1	Total 6	C 3	O 3	0	0
8	D	1	Total 6	C 3	O 3	0	0
8	E	1	Total 6	C 3	O 3	0	0
8	F	1	Total 6	C 3	O 3	0	0
8	G	1	Total 6	C 3	O 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	H	1	Total 6	C 3	O 3	0	0

- Molecule 9 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	J	1	Total 1	Cu 1	0	0

- Molecule 10 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	I	398	Total 398	O 398	0	0

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.
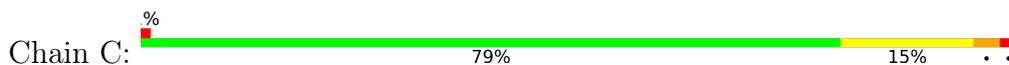
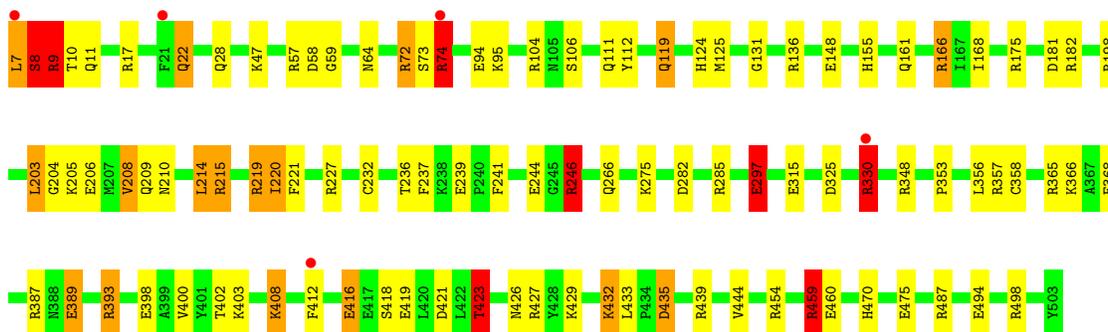- Molecule 1: Glucose-6-phosphate 1-dehydrogenase
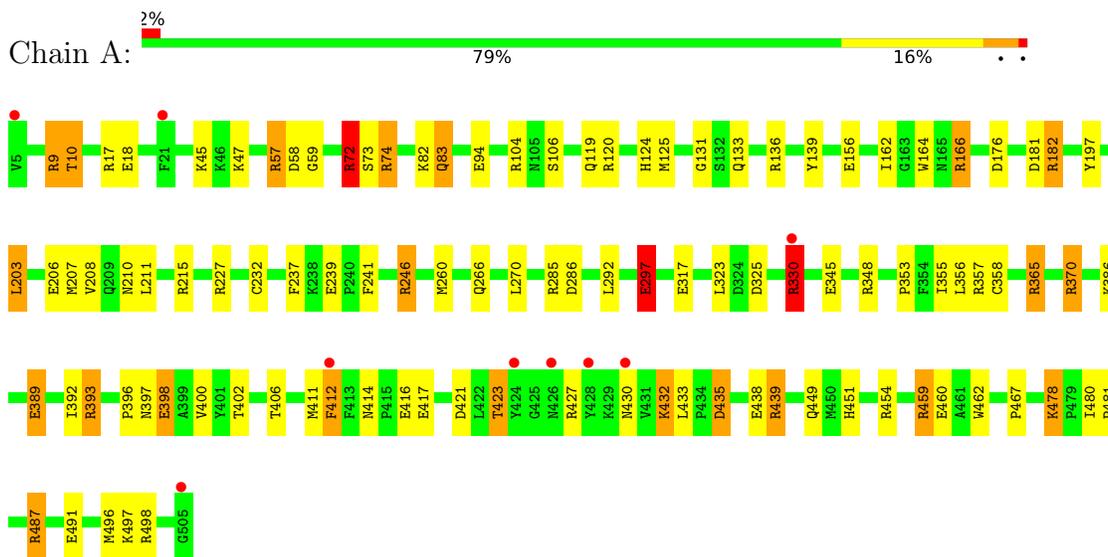
Chain A:

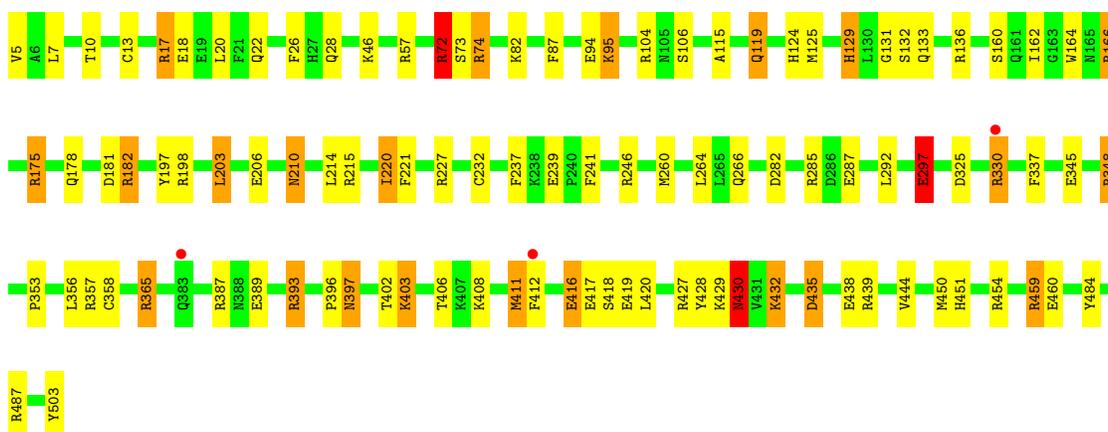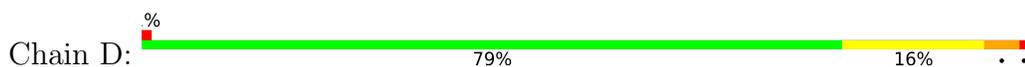- Molecule 2: Glucose-6-phosphate 1-dehydrogenase
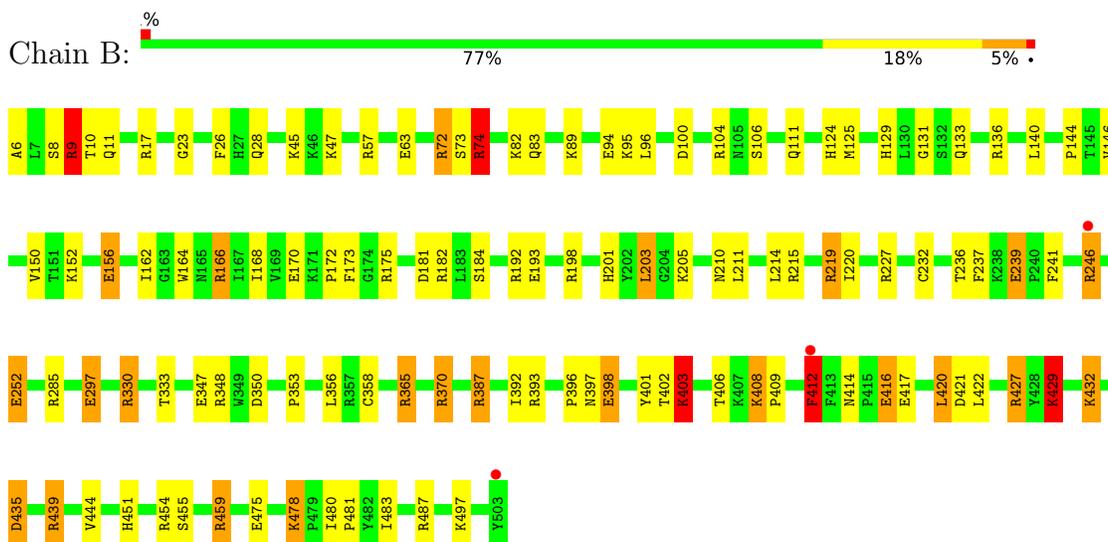
Chain B:

- Molecule 3: Glucose-6-phosphate 1-dehydrogenase
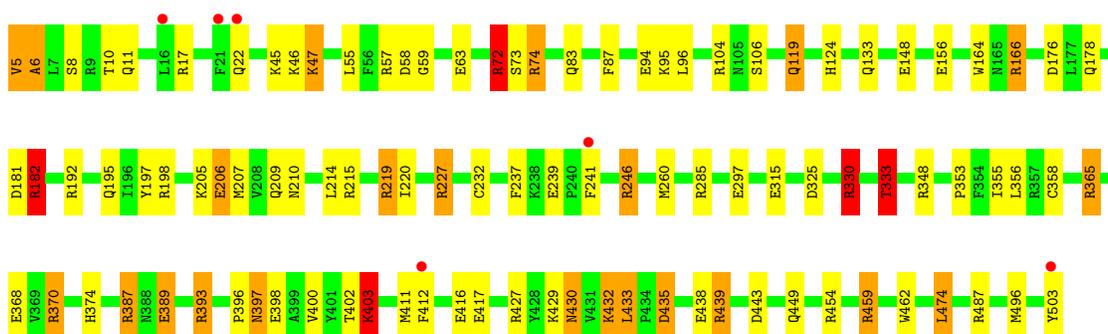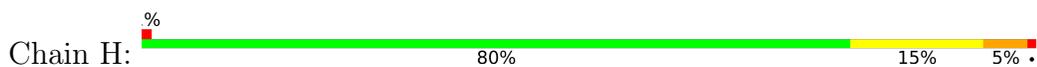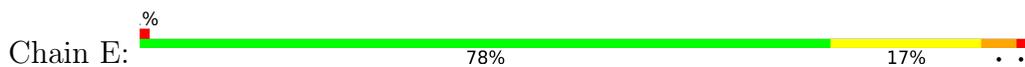
Chain C:

• Molecule 4: Glucose-6-phosphate 1-dehydrogenase
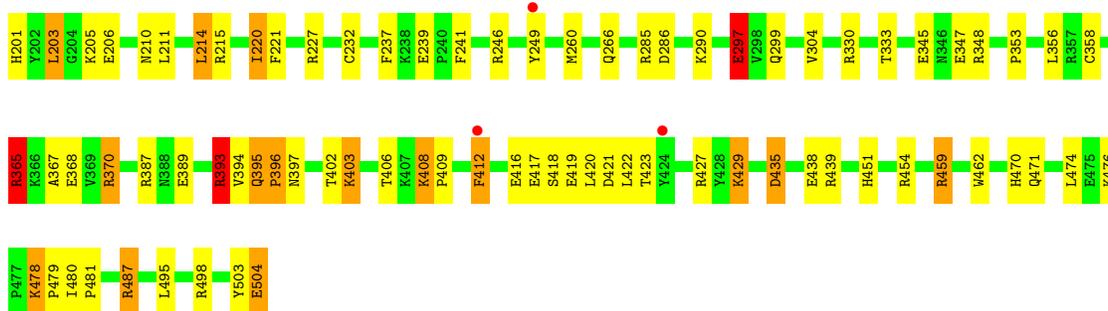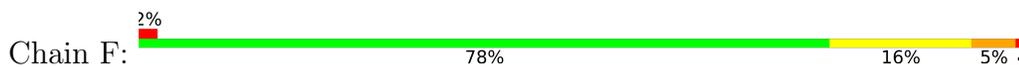
Chain D:

• Molecule 4: Glucose-6-phosphate 1-dehydrogenase

Chain H:

• Molecule 5: Glucose-6-phosphate 1-dehydrogenase

Chain E:

H201 Y202 L203 G204 K205 E206 N210 L211 L214 R215 I220 F221 R227 C232 F237 K238 E239 P240 F241 R246 Y249 M260 Q266 R285 D286 K290 E297 V298 Q299 V304 R330 T333 E345 N346 E347 R348 P353 L356 R357 C358

R365 K366 A367 E368 V369 R370 R387 N388 E389 R394 V394 Q395 P396 N397 T402 K403 T406 K407 K408 P409 F412 E416 E417 S418 E419 L420 D421 L422 T423 Y424 R427 Y428 K429 D435 E438 R439 H451 R454 R459 W462 H470 Q471 L474 E475 K476

P477 K478 P479 I480 P481 R487 L495 R498 Y503 E504

• Molecule 6: Glucose-6-phosphate 1-dehydrogenase

Chain F:

2%
78% 16% 5% •

L7 S8 R9 T10 Q11 V12 C13 G14 I15 L16 R17 E18 F21 Q22 F26 H27 Q28 K47 D58 R72 S73 R74 E85 K95 R104 N105 S106 Q119 R120 H124 R136 V146 V150 R166 I167 I168 V169 E170 R175 D176 D181 R182 E193

R198 H201 Y202 L203 E206 N210 L211 L214 R215 F216 R219 I220 F221 R227 C232 T236 F237 K238 E239 P240 F241 R246 Q266 E274 R285 E297 L323 D324 D325 R330 E347 R348 P353 L356 R357 C358 R365 R370

D379 Q383 R387 N388 E389 R393 V394 Q395 P396 N397 T402 K403 M404 M405 T406 K407 K408 F412 E416 D421 L422 T423 Y424 G425 N426 R427 Y428 K429 N430 V431 P434 D435 E438 R439 L440 I441 L442 D443 V444 H451 R454 S455 R459 W462 Q471

E475 K478 P479 I480 P481 R487 Y503

• Molecule 7: Glucose-6-phosphate 1-dehydrogenase

Chain G:

2%
77% 17% 5% •

S8 R9 T10 Q11 V12 C13 R17 E18 E19 L20 F21 Q22 F26 H27 Q28 K47 R57 D58 R72 S73 R74 D79 K82 Q83 P92 K95 L96 K97 R104 N105 S106 Y112 Q119 H124 Q133 R136 E148 I162 G163 W164 N165 R166

R175 D176 D181 R182 E193 Y197 R198 I199 D200 L203 G204 K205 Q209 N210 L214 R215 R219 I220 F221 R227 C232 F237 K238 E239 P240 F241 G242 T243 R246 Y249 M260 H263 Q266 R285 D286 L292 E297 E315 L323 D324

D325 R330 E347 R348 P353 L356 R357 C358 R365 E368 V369 R370 R387 N388 R393 V394 Q395 P396 N397 T402 K403 M404 K408 M411 F412 E416 E417 S418 T423 Y424 R427 Y428 K429 K432 D435 E438 R439 Q449 M450 H451 R454

R459 E460 A461 W462 E473 K476 P477 K478 I483 R487 K497 F501 Q502 Y503

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	64.36Å 122.61Å 161.31Å 77.23° 80.99° 77.40°	Depositor
Resolution (Å)	50.01 – 2.28 117.39 – 2.28	Depositor EDS
% Data completeness (in resolution range)	96.9 (50.01-2.28) 96.9 (117.39-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.211 , 0.261 0.219 , 0.262	Depositor DCC
R_{free} test set	10333 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	43.7	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 34.6	EDS
L-test for twinning[2]	$< |L| > = 0.52, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	32772	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.94% 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: GOL, ALY, CU

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	0.76	6/4147 (0.1%)	1.11	27/5606 (0.5%)
2	B	0.75	8/4117 (0.2%)	1.11	23/5565 (0.4%)
3	C	0.77	10/4123 (0.2%)	1.14	33/5572 (0.6%)
4	D	0.76	10/4135 (0.2%)	1.10	27/5589 (0.5%)
4	H	0.75	8/4124 (0.2%)	1.11	26/5575 (0.5%)
5	E	0.73	8/4133 (0.2%)	1.11	19/5587 (0.3%)
6	F	0.81	11/4122 (0.3%)	1.18	38/5572 (0.7%)
7	G	0.71	5/4104 (0.1%)	1.12	31/5547 (0.6%)
All	All	0.75	66/33005 (0.2%)	1.12	224/44613 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	17
2	B	0	16
3	C	0	18
4	D	0	18
4	H	0	16
5	E	0	18
6	F	0	17
7	G	0	19
All	All	0	139

The worst 5 of 66 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	193	GLU	CD-OE2	13.37	1.40	1.25
6	F	206	GLU	CD-OE1	12.37	1.39	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	274	GLU	CD-OE1	11.70	1.38	1.25
3	C	416	GLU	CD-OE2	10.03	1.36	1.25
4	H	148	GLU	CD-OE1	9.99	1.36	1.25

The worst 5 of 224 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	348	ARG	NE-CZ-NH1	12.78	126.69	120.30
3	C	330	ARG	NE-CZ-NH2	-12.08	114.26	120.30
3	C	387	ARG	NE-CZ-NH2	-11.88	114.36	120.30
5	E	387	ARG	NE-CZ-NH2	-11.64	114.48	120.30
2	B	215	ARG	NE-CZ-NH2	-11.19	114.71	120.30

There are no chirality outliers.

5 of 139 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	ARG	Sidechain
1	A	120	ARG	Sidechain
1	A	57	ARG	Sidechain
1	A	72	ARG	Sidechain
1	A	74	ARG	Sidechain

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	4053	0	3996	47	0
2	B	4036	0	3979	77	0
3	C	4039	0	3987	54	0
4	D	4051	0	4001	43	0
4	H	4043	0	3988	56	0
5	E	4052	0	3994	60	0
6	F	4028	0	3973	65	0
7	G	4023	0	3963	52	0
8	B	12	0	16	6	0
8	C	6	0	8	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	D	6	0	8	1	0
8	E	6	0	8	4	0
8	F	6	0	8	2	0
8	G	6	0	8	0	0
8	H	6	0	8	2	0
9	J	1	0	0	0	0
10	I	398	0	0	29	0
All	All	32772	0	31945	444	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 7.

The worst 5 of 444 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)
3:C:7:LEU:HD13	3:C:11:GLN:HE21	0.93	1.09
5:E:408:LYS:CE	5:E:409:PRO:O	2.01	1.08
3:C:421:ASP:OD1	3:C:423:THR:HG22	1.54	1.06
1:A:421:ASP:OD1	1:A:423:THR:HG22	1.57	1.04
3:C:7:LEU:HD13	3:C:11:GLN:NE2	1.74	1.02

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	499/501 (100%)	484 (97%)	14 (3%)	1 (0%)	47	57
2	B	495/498 (99%)	483 (98%)	10 (2%)	2 (0%)	34	40
3	C	495/497 (100%)	483 (98%)	10 (2%)	2 (0%)	34	40
4	D	497/499 (100%)	486 (98%)	9 (2%)	2 (0%)	34	40

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	H	496/499 (99%)	482 (97%)	11 (2%)	3 (1%)	25	29
5	E	497/500 (99%)	485 (98%)	10 (2%)	2 (0%)	34	40
6	F	495/497 (100%)	477 (96%)	16 (3%)	2 (0%)	34	40
7	G	493/496 (99%)	481 (98%)	10 (2%)	2 (0%)	34	40
All	All	3967/3987 (100%)	3861 (97%)	90 (2%)	16 (0%)	34	40

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	9	ARG
4	D	430	ASN
7	G	429	LYS
4	H	430	ASN
5	E	396	PRO

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	437/437 (100%)	415 (95%)	22 (5%)	24	32
2	B	434/434 (100%)	402 (93%)	32 (7%)	13	16
3	C	435/434 (100%)	408 (94%)	27 (6%)	18	23
4	D	436/435 (100%)	411 (94%)	25 (6%)	20	26
4	H	435/435 (100%)	410 (94%)	25 (6%)	20	26
5	E	436/436 (100%)	407 (93%)	29 (7%)	16	19
6	F	435/435 (100%)	405 (93%)	30 (7%)	15	18
7	G	433/433 (100%)	403 (93%)	30 (7%)	15	18
All	All	3481/3479 (100%)	3261 (94%)	220 (6%)	18	22

5 of 220 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	203	LEU
6	F	74	ARG
4	H	474	LEU
4	H	72	ARG
5	E	249	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 70 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	388	ASN
4	H	11	GLN
4	H	266	GLN
3	C	161	GLN
3	C	124	HIS

5.3.3   RNA   ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains   ⓘ

6 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	$\#|Z| > 2$
3	ALY	C	403	3	10,11,12	0.89	0	7,12,14	1.06	0
7	ALY	G	403	7	10,11,12	1.15	1 (10%)	7,12,14	1.40	0
5	ALY	E	403	5	10,11,12	0.73	0	7,12,14	1.14	1 (14%)
4	ALY	D	403	4	10,11,12	0.45	0	7,12,14	0.89	0
2	ALY	B	403	2	10,11,12	0.87	1 (10%)	7,12,14	1.73	2 (28%)
4	ALY	H	403	4	10,11,12	0.98	1 (10%)	7,12,14	1.20	1 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ALY	C	403	3	-	3/9/10/12	-
7	ALY	G	403	7	-	5/9/10/12	-
5	ALY	E	403	5	-	3/9/10/12	-
4	ALY	D	403	4	-	3/9/10/12	-
2	ALY	B	403	2	-	4/9/10/12	-
4	ALY	H	403	4	-	2/9/10/12	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	403	ALY	O-C	2.04	1.28	1.19
4	H	403	ALY	O-C	2.04	1.28	1.19
2	B	403	ALY	O-C	2.03	1.28	1.19

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(º)	Ideal(º)
2	B	403	ALY	CE-NZ-CH	2.76	126.81	122.56
5	E	403	ALY	CE-NZ-CH	2.69	126.69	122.56
2	B	403	ALY	OH-CH-CH3	-2.52	117.38	122.06
4	H	403	ALY	CE-NZ-CH	2.48	126.37	122.56

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

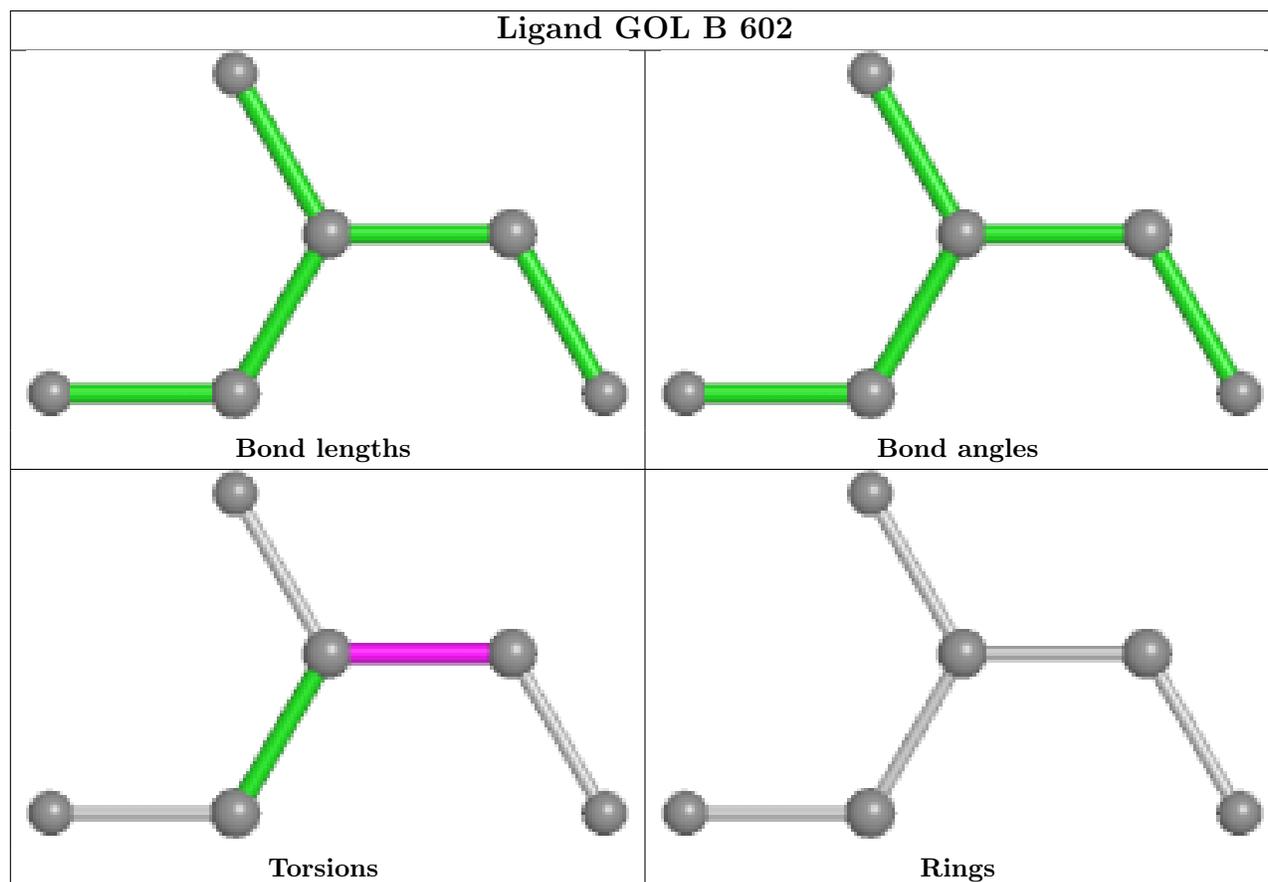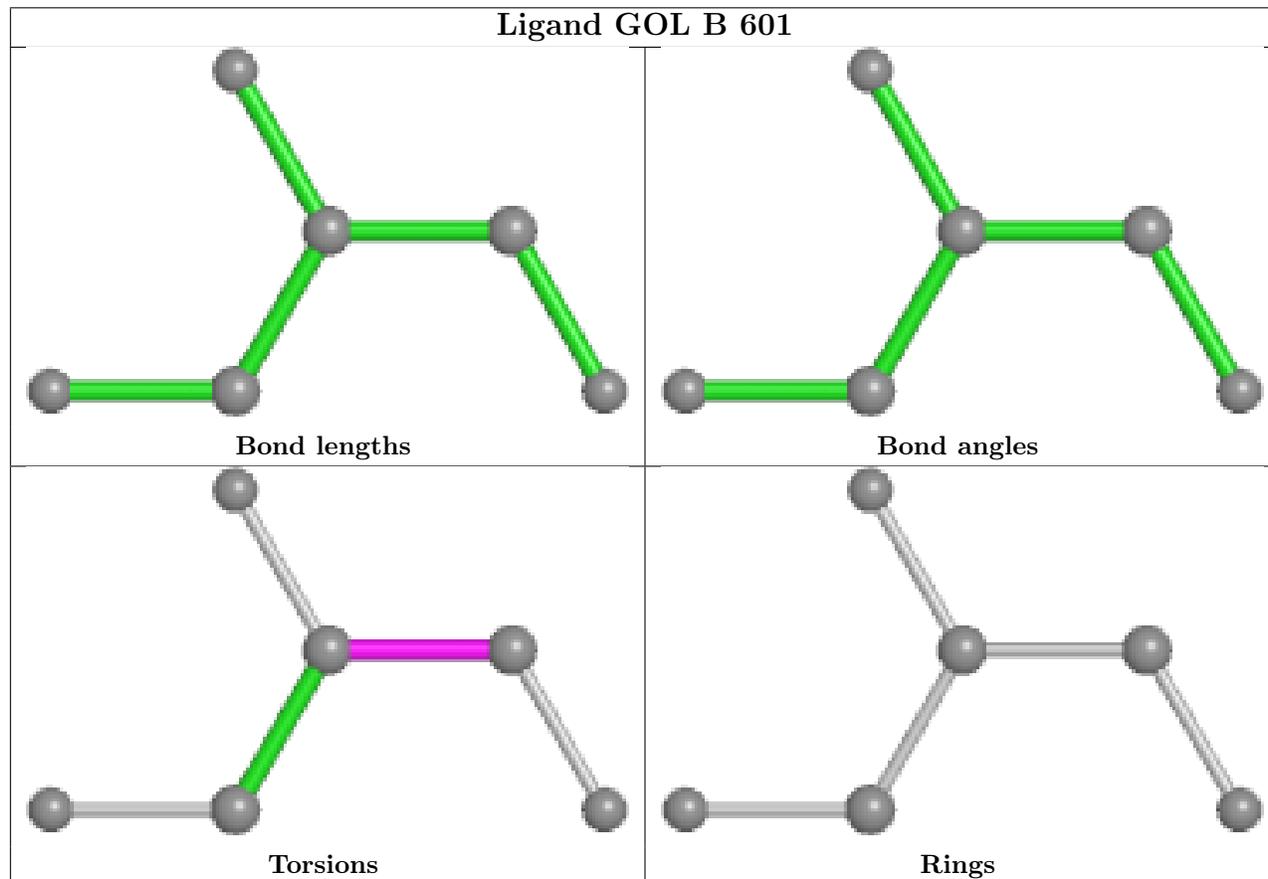| Mol	Chain	Res	Type	Atoms
3	C	403	ALY	C-CA-CB-CG
4	D	403	ALY	O-C-CA-CB
7	G	403	ALY	C-CA-CB-CG
2	B	403	ALY	CD-CE-NZ-CH
7	G	403	ALY	CH3-CH-NZ-CE

There are no ring outliers.

5 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	403	ALY	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	403	ALY	4	0
4	D	403	ALY	1	0
2	B	403	ALY	15	0
4	H	403	ALY	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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).

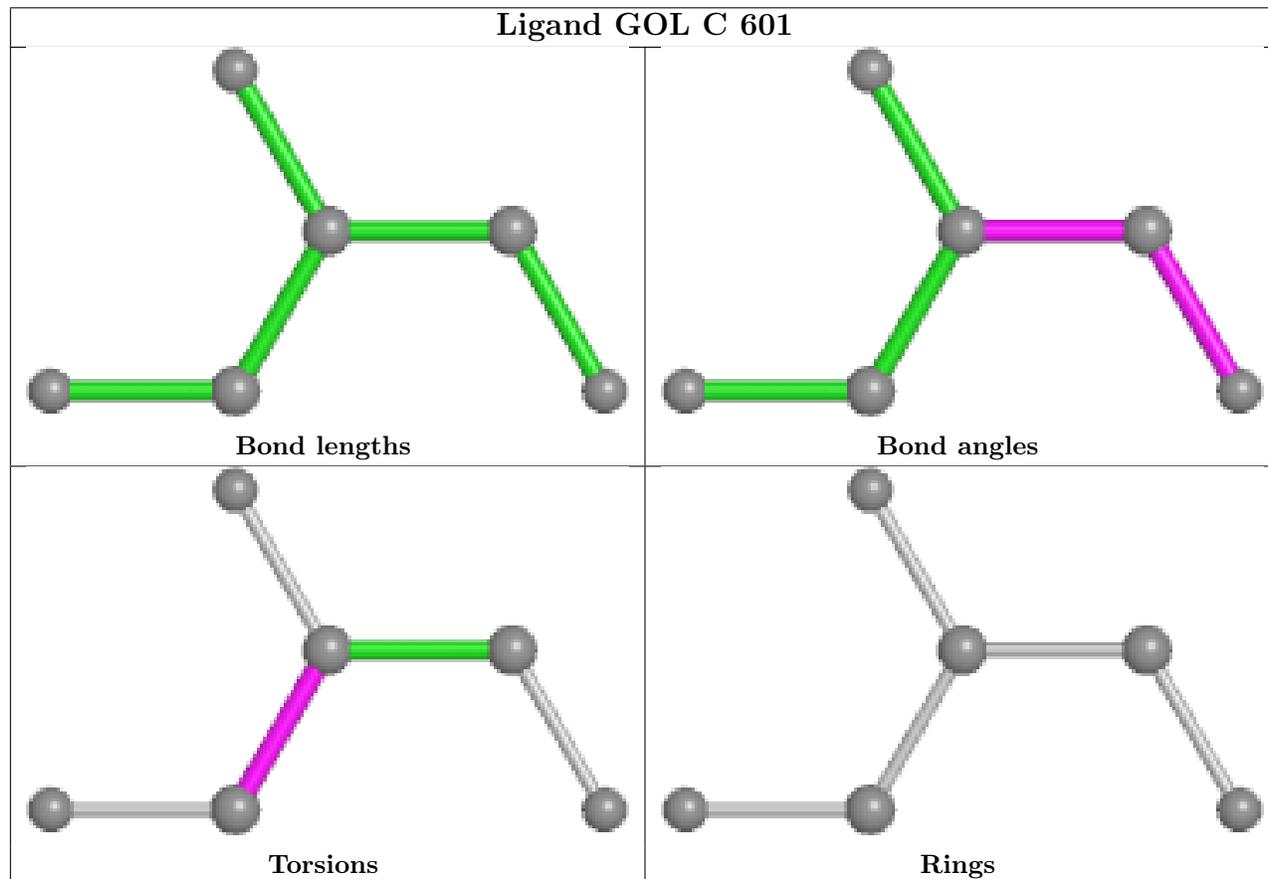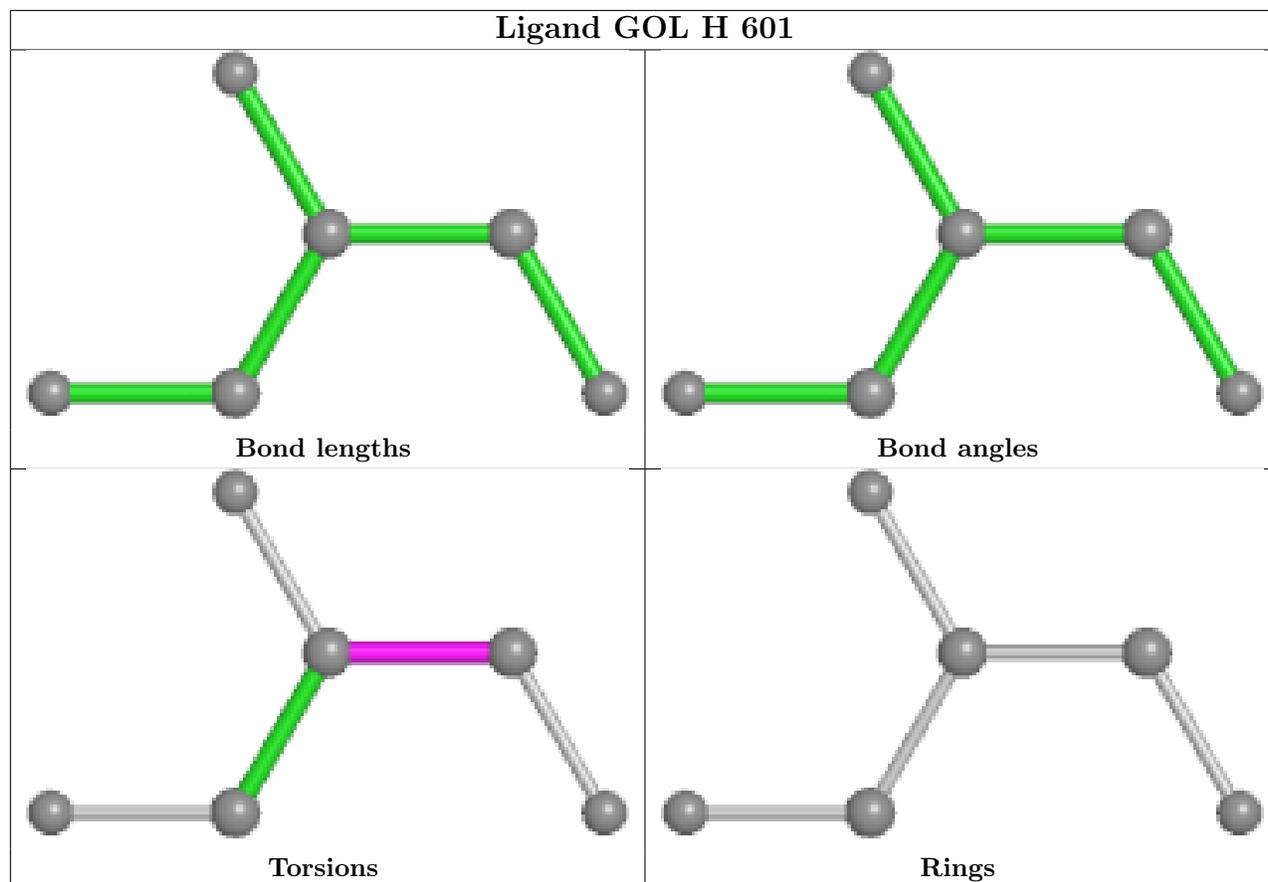| Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
8	GOL	B	602	-	5,5,5	0.21	0	5,5,5	0.42	0
8	GOL	D	601	-	5,5,5	0.62	0	5,5,5	1.21	0
8	GOL	B	601	-	5,5,5	0.42	0	5,5,5	1.04	0
8	GOL	H	601	-	5,5,5	0.29	0	5,5,5	0.68	0
8	GOL	C	601	-	5,5,5	0.94	0	5,5,5	1.28	1 (20%)
8	GOL	E	601	-	5,5,5	0.30	0	5,5,5	0.74	0
8	GOL	F	601	-	5,5,5	0.39	0	5,5,5	1.60	1 (20%)
8	GOL	G	601	-	5,5,5	0.66	0	5,5,5	1.36	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
8	GOL	B	602	-	-	2/4/4/4	-
8	GOL	D	601	-	-	4/4/4/4	-
8	GOL	B	601	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	GOL	H	601	-	-	2/4/4/4	-
8	GOL	C	601	-	-	2/4/4/4	-
8	GOL	E	601	-	-	1/4/4/4	-
8	GOL	F	601	-	-	4/4/4/4	-
8	GOL	G	601	-	-	4/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	F	601	GOL	O2-C2-C3	-2.35	98.78	109.12
8	C	601	GOL	O1-C1-C2	2.02	119.89	110.20

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	602	GOL	O1-C1-C2-C3
8	D	601	GOL	O1-C1-C2-C3
8	D	601	GOL	C1-C2-C3-O3
8	F	601	GOL	O1-C1-C2-C3
8	G	601	GOL	C1-C2-C3-O3

There are no ring outliers.

7 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	602	GOL	4	0
8	D	601	GOL	1	0
8	B	601	GOL	2	0
8	H	601	GOL	2	0
8	C	601	GOL	1	0
8	E	601	GOL	4	0
8	F	601	GOL	2	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.
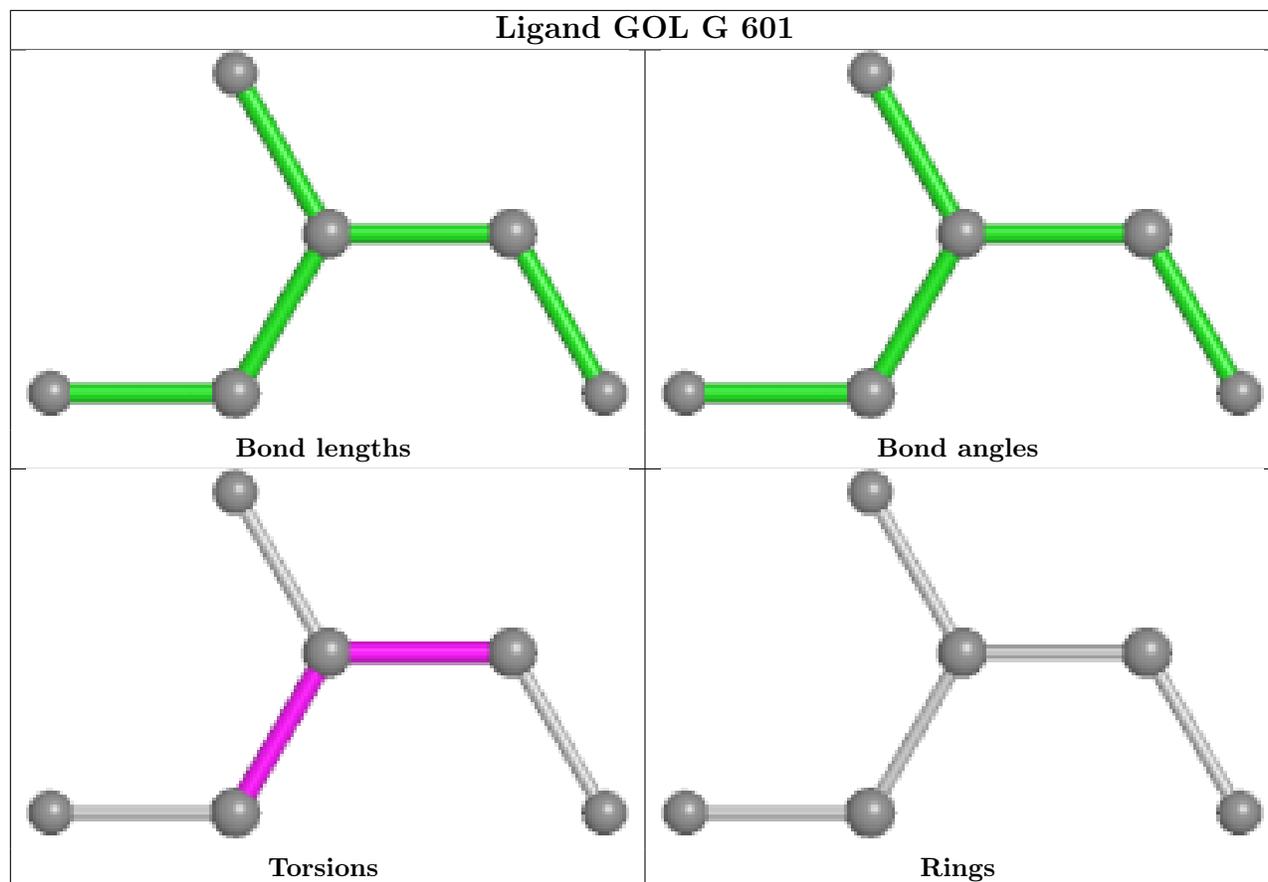
Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD 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 GOL B 602

Ligand GOL D 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL B 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL H 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL C 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL E 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL F 601

Bond lengths

Bond angles

Torsions

Rings

Ligand GOL G 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(Å^2)	Q<0.9
1	A	501/501 (100%)	0.19	9 (1%)	68	74	35, 51, 87, 129	0
2	B	497/498 (99%)	0.19	3 (0%)	89	91	35, 50, 79, 127	0
3	C	496/497 (99%)	0.20	5 (1%)	82	86	34, 51, 85, 112	0
4	D	498/499 (99%)	0.15	3 (0%)	89	91	33, 52, 79, 104	0
4	H	498/499 (99%)	0.24	6 (1%)	79	82	36, 56, 91, 114	0
5	E	499/500 (99%)	0.22	4 (0%)	86	89	34, 58, 93, 126	0
6	F	497/497 (100%)	0.26	11 (2%)	62	68	36, 52, 92, 157	0
7	G	495/496 (99%)	0.34	10 (2%)	65	70	38, 65, 100, 128	0
All	All	3981/3987 (99%)	0.22	51 (1%)	77	81	33, 54, 90, 157	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	412	PHE	6.9
5	E	249	TYR	5.3
4	H	412	PHE	5.0
4	D	412	PHE	5.0
3	C	7	LEU	4.4

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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
7	ALY	G	403	12/13	0.95	0.16	41,63,82,82	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
4	ALY	H	403	12/13	0.95	0.16	34,60,73,77	0
2	ALY	B	403	12/13	0.96	0.20	44,68,86,86	0
5	ALY	E	403	12/13	0.96	0.18	45,61,74,76	0
4	ALY	D	403	12/13	0.97	0.19	40,60,73,76	0
3	ALY	C	403	12/13	0.97	0.17	35,58,73,76	0

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(Å²)	Q<0.9
9	CU	J	1	1/1	0.64	0.16	124,124,124,124	0
8	GOL	D	601	6/6	0.77	0.25	52,68,71,80	0
8	GOL	B	602	6/6	0.78	0.23	52,76,81,84	0
8	GOL	C	601	6/6	0.90	0.31	41,45,66,70	0
8	GOL	H	601	6/6	0.92	0.19	56,68,71,72	0
8	GOL	B	601	6/6	0.93	0.26	45,59,72,76	0
8	GOL	E	601	6/6	0.94	0.24	61,74,88,94	0
8	GOL	F	601	6/6	0.94	0.26	56,62,68,75	0
8	GOL	G	601	6/6	0.96	0.29	44,48,53,60	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.

WORLDWIDE PDB PROTEIN DATA BANK

Electron density around CU J 1:

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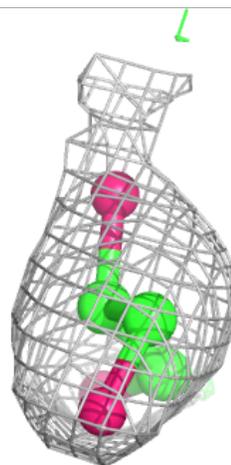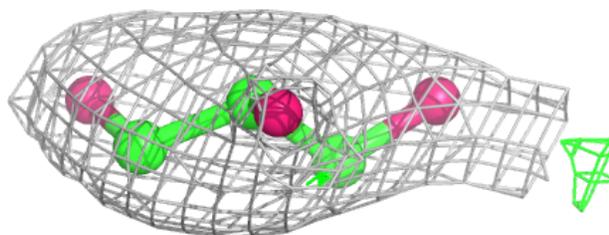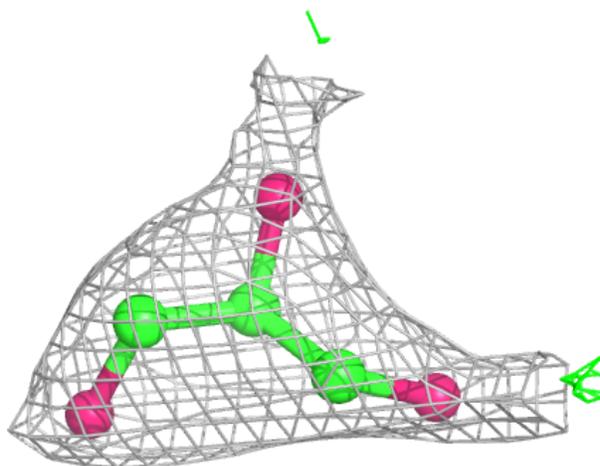$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around GOL D 601:

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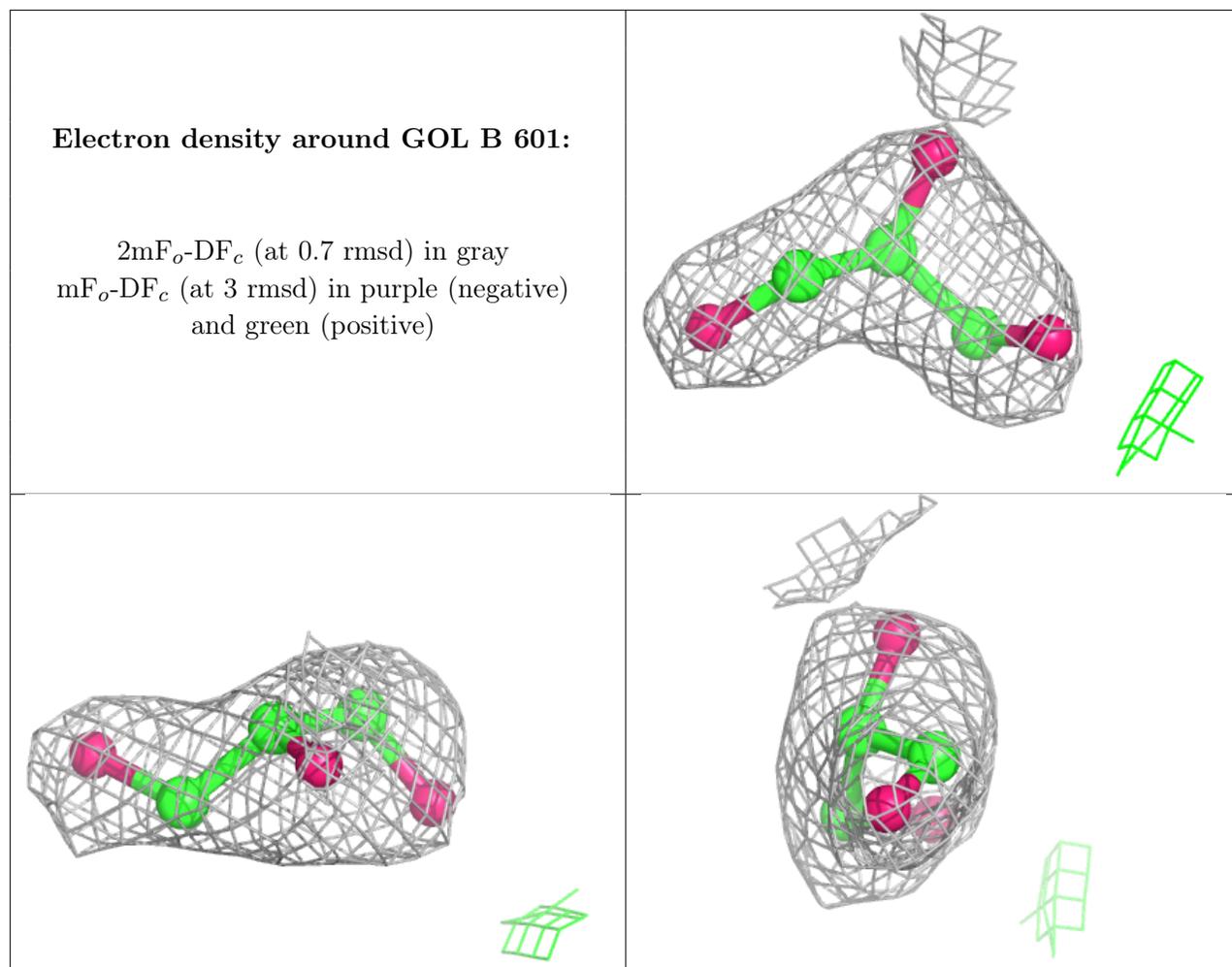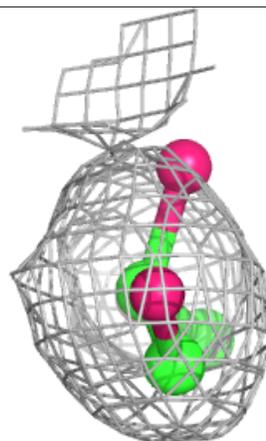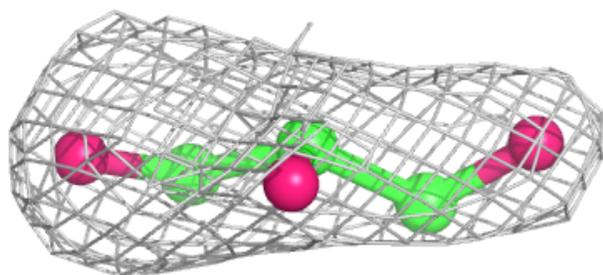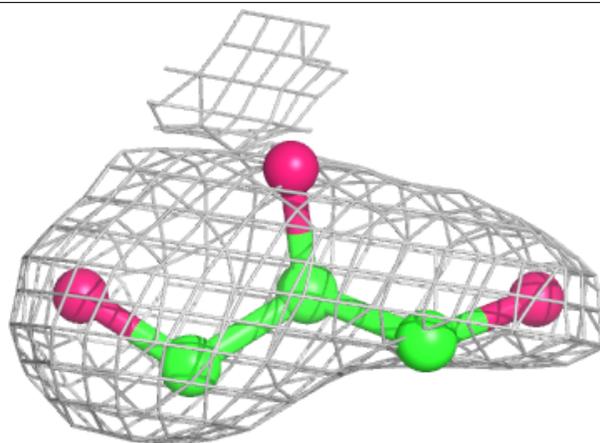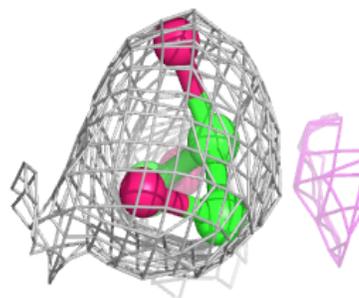mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around GOL B 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 GOL C 601:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
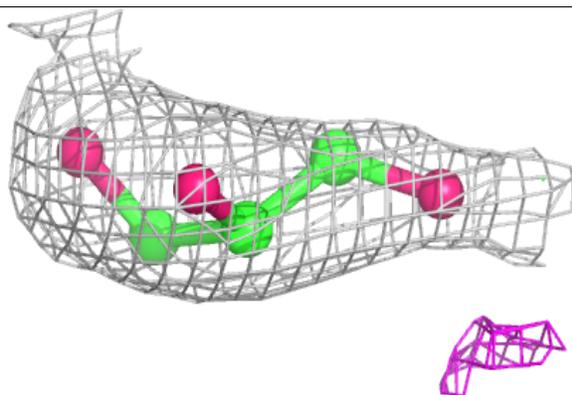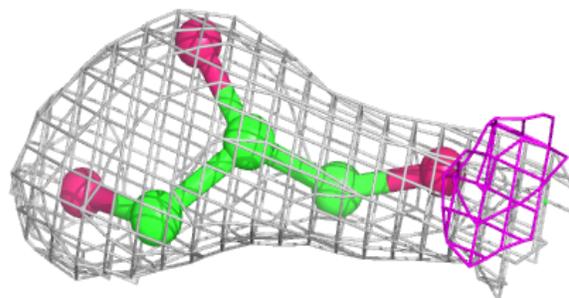and green (positive)
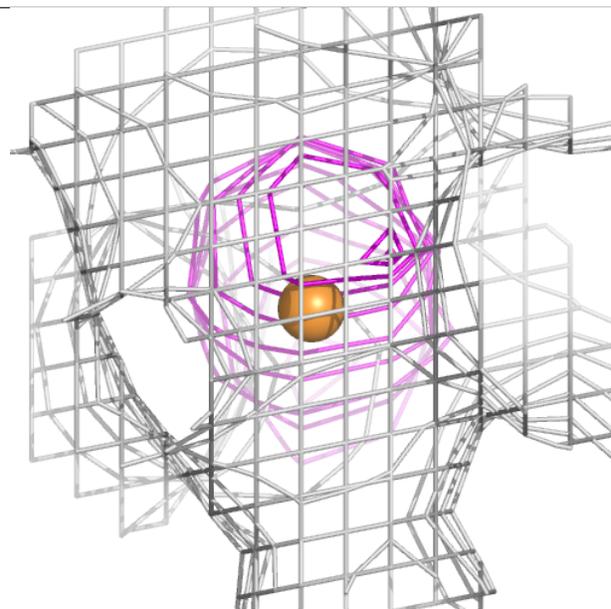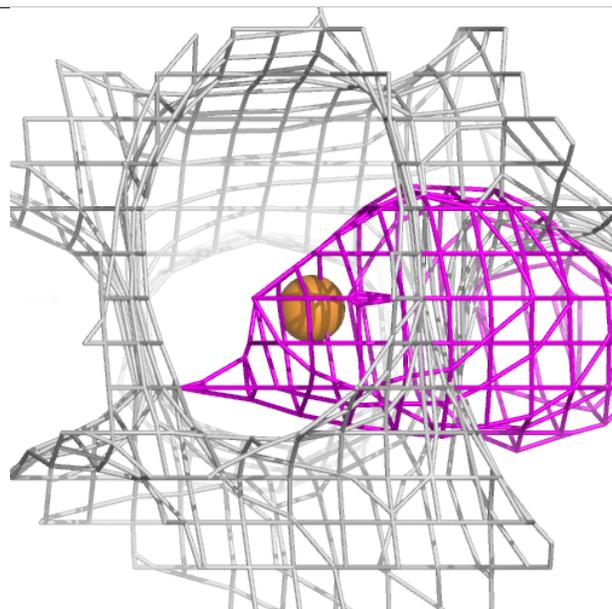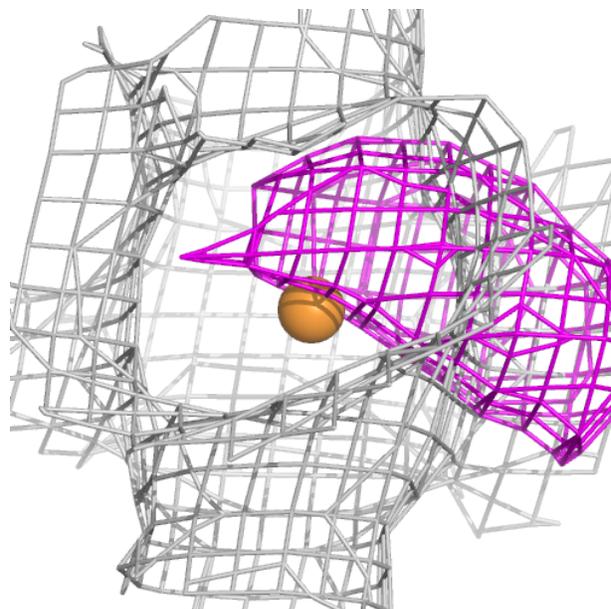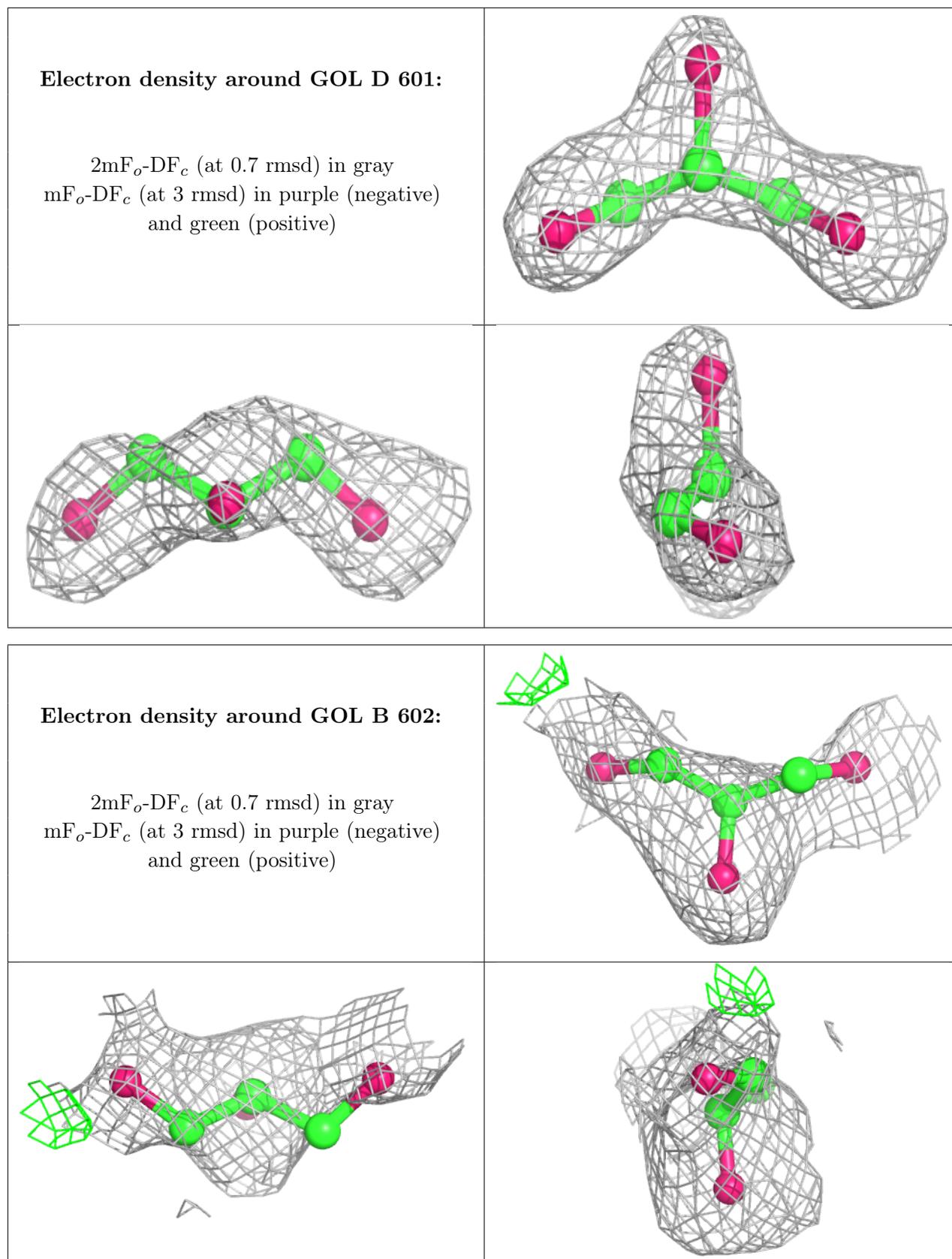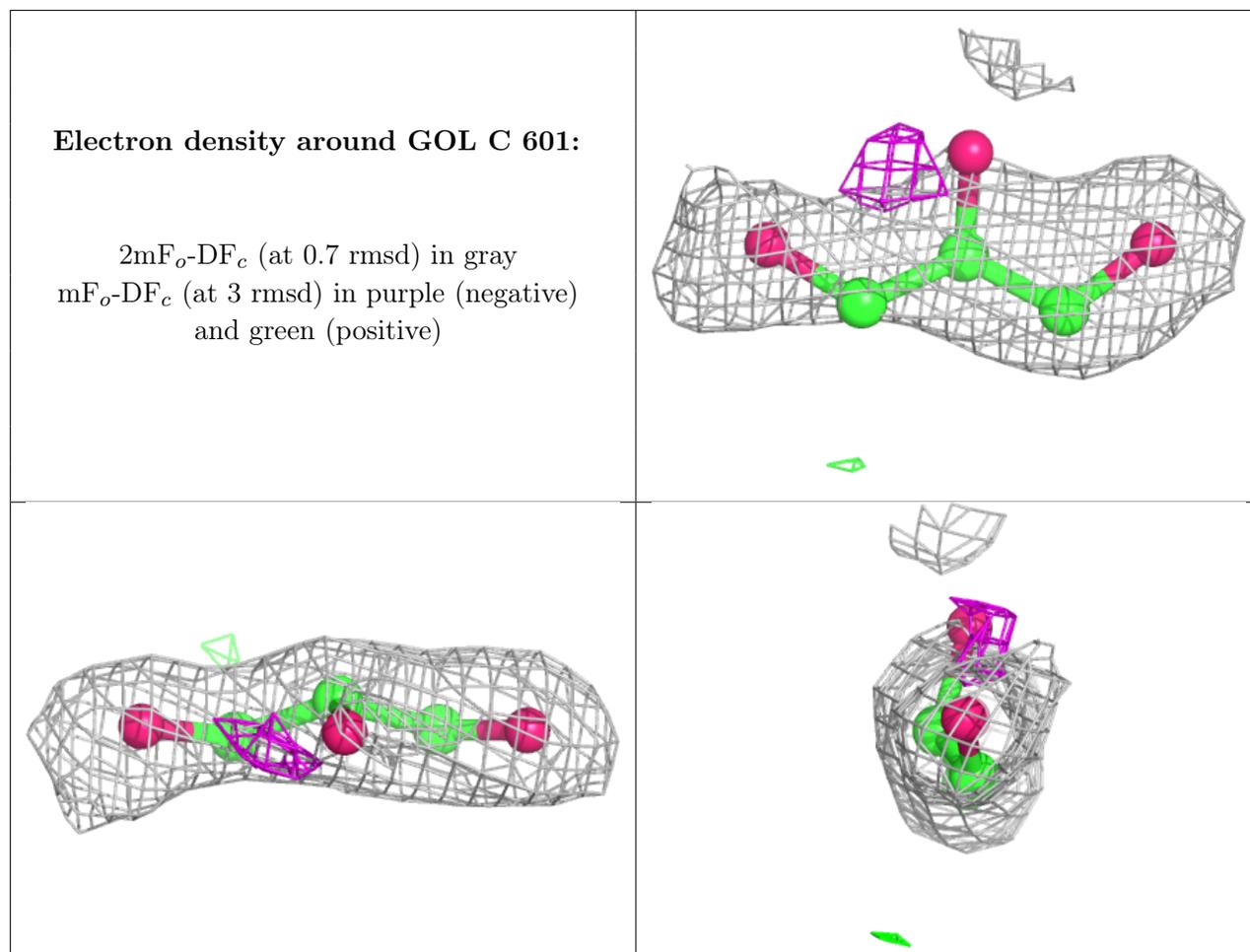
Electron density around GOL H 601:

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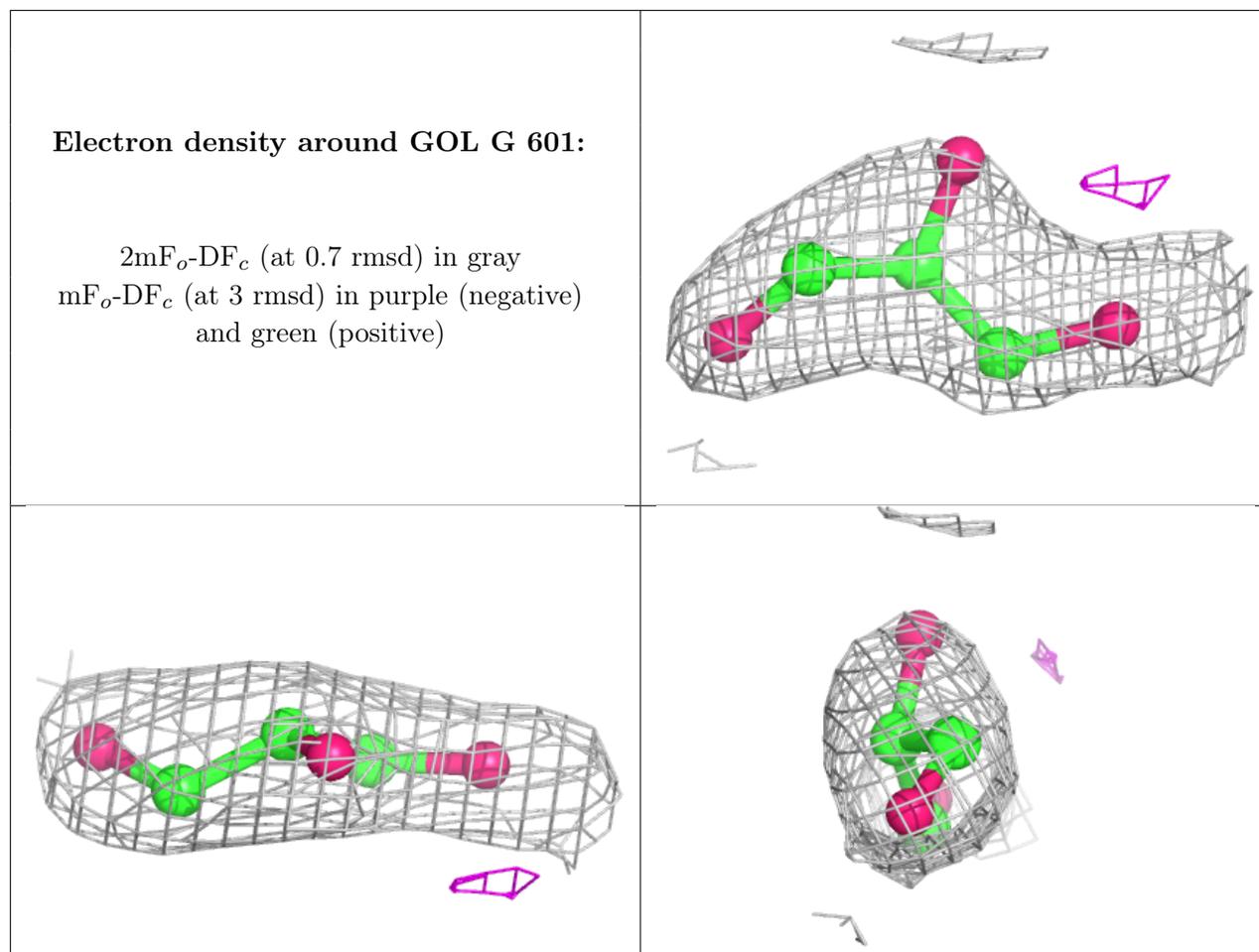mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around GOL 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 GOL E 601:

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

Electron density around GOL F 601:

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

Electron density around GOL G 601:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.