# W O R L D W I D E
# wwPDB
## PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2024 – 12:10 PM EST

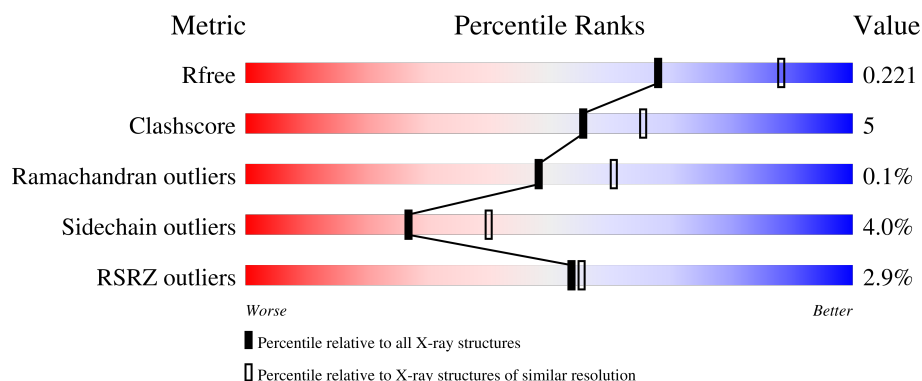|          |     |                                                                                                          |
|---------:|:---:|:---------------------------------------------------------------------------------------------------------|
| PDB ID   | :   | 3OTK                                                                                                     |
| Title    | :   | Structure and mechanisim of core 2 beta1,6-n-acetylglucosaminyltransferase: a Metal-ion independent gt-a glycosyltransferase |
| Authors  | :   | Pak, J.E.; Rini, J.M.                                                                                    |
| Deposited on | : | 2010-09-13                                                                                            |
| Resolution | : | 2.30 Å(reported)                                                                                         |

# 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 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

| Metric | Percentile Ranks | Value |
|--------|------------------|-------|
| Rfree | | 0.221 |
| Clashscore | | 5 |
| Ramachandran outliers | | 0.1% |
| Sidechain outliers | | 4.0% |
| RSRZ outliers | | 2.9% |

*Worse* — *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|--------|--------------------------|----------------------------------------------------|
| $R_{free}$ | 164625 | 5963 (2.30-2.30) |
| Clashscore | 180529 | 6698 (2.30-2.30) |
| Ramachandran outliers | 177936 | 6640 (2.30-2.30) |
| Sidechain outliers | 177891 | 6640 (2.30-2.30) |
| RSRZ outliers | 164620 | 5963 (2.30-2.30) |

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 | 391 | 2% / 81% 14% 6% |
| 1 | B | 391 | 5% / 80% 13% • 7% |
| 1 | C | 391 | 3% / 80% 11% • 8% |
| 1 | D | 391 | 2% / 83% 9% • 7% |
| 2 | E | 3 | 67% 33% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 2 | F | 3 | 67% 33% |

## 2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 12658 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 Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetyl glucosaminyltransferase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | A | 369 | Total | C | N | O | S | 0 | 2 | 0 |
| | | | 3015 | 1936 | 513 | 545 | 21 | | | |
| 1 | B | 363 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 2956 | 1897 | 501 | 537 | 21 | | | |
| 1 | C | 360 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 2934 | 1881 | 498 | 535 | 20 | | | |
| 1 | D | 364 | Total | C | N | O | S | 0 | 2 | 0 |
| | | | 2966 | 1905 | 502 | 538 | 21 | | | |

There are 4 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | 217 | SER | CYS | engineered mutation | UNP Q09324 |
| B | 217 | SER | CYS | engineered mutation | UNP Q09324 |
| C | 217 | SER | CYS | engineered mutation | UNP Q09324 |
| D | 217 | SER | CYS | engineered mutation | UNP Q09324 |

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.
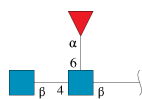


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---------|---------|-------|
| 2 | E | 3 | Total | C | N | O | 0 | 0 | 0 |
| | | | 38 | 22 | 2 | 14 | | | |
| 2 | F | 3 | Total | C | N | O | 0 | 0 | 0 |
| | | | 38 | 22 | 2 | 14 | | | |

- Molecule 3 is HEPTANE-1,2,3-TRIOL (three-letter code: HTO) (formula: $C_7H_{16}O_3$).

HTO
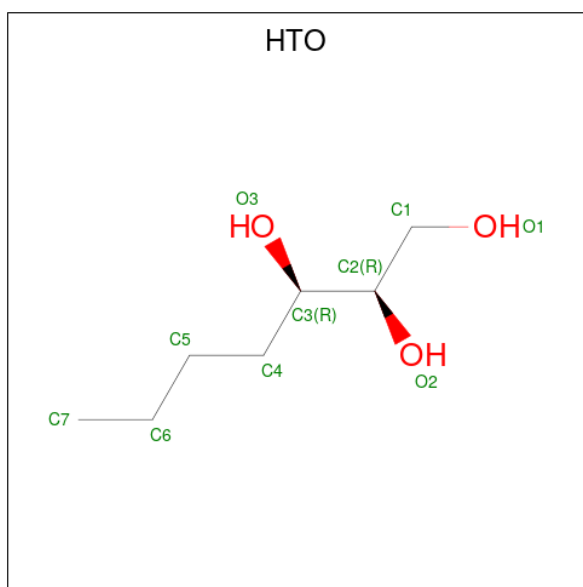


| Mol | Chain | Residues | Atoms | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|
| 3 | A | 1 | Total 10 | C 7 | O 3 | 0 | 0 |
| 3 | A | 1 | Total 10 | C 7 | O 3 | 0 | 0 |

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 4 | A | 1 | Total 1 | Na 1 | 0 | 0 |
| 4 | C | 1 | Total 1 | Na 1 | 0 | 0 |

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

**NAG**

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 5 | A | 1 | Total | C | N | O | 0 | 0 |
|   |   |   | 14 | 8 | 1 | 5 |   |   |
| 5 | B | 1 | Total | C | N | O | 0 | 0 |
|   |   |   | 14 | 8 | 1 | 5 |   |   |
| 5 | C | 1 | Total | C | N | O | 0 | 0 |
|   |   |   | 14 | 8 | 1 | 5 |   |   |
| 5 | D | 1 | Total | C | N | O | 0 | 0 |
|   |   |   | 14 | 8 | 1 | 5 |   |   |

- Molecule 6 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).

**UDP**
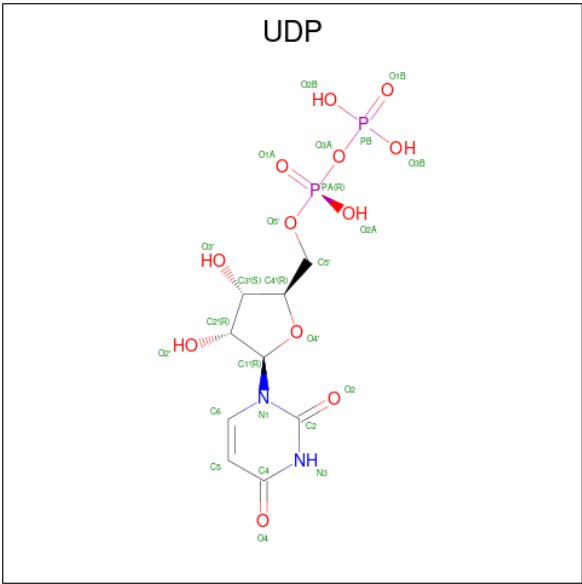
| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 6 | A | 1 | Total 25 | C 9 | N 2 | O 12 | P 2 | 0 | 0 |
| 6 | B | 1 | Total 25 | C 9 | N 2 | O 12 | P 2 | 0 | 0 |
| 6 | C | 1 | Total 25 | C 9 | N 2 | O 12 | P 2 | 0 | 0 |
| 6 | D | 1 | Total 25 | C 9 | N 2 | O 12 | P 2 | 0 | 0 |

• Molecule 7 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 7 | A | 143 | Total 143 | O 143 | 0 | 0 |
| 7 | B | 122 | Total 122 | O 122 | 0 | 0 |
| 7 | C | 120 | Total 120 | O 120 | 0 | 0 |
| 7 | D | 148 | Total 148 | O 148 | 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 summ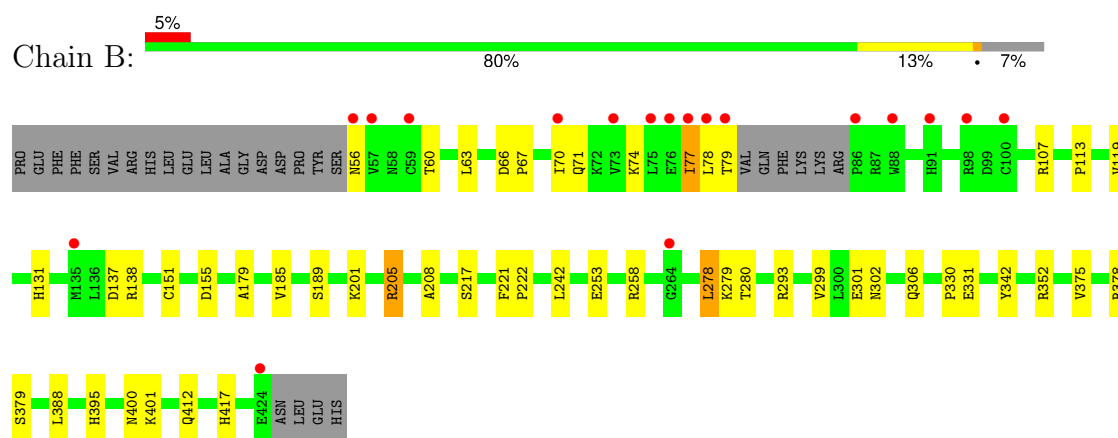arises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: 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.
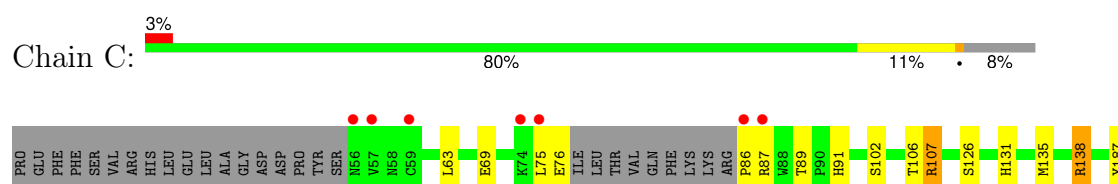
• Molecule 1: Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase



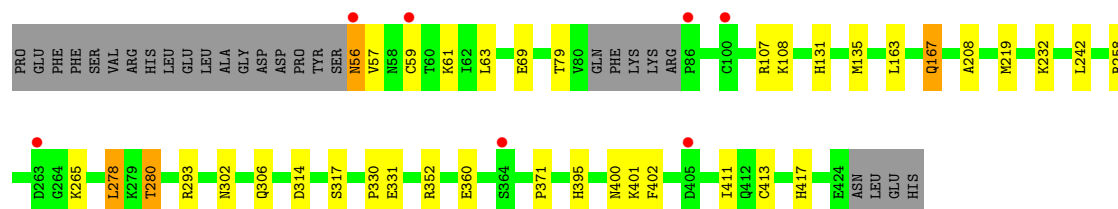• Molecule 1: Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase



• Molecule 1: Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase

D220 K232 L242 E243 T244 P248 K251 E252 K257 V262 D263 G264 V272 K273 V299 Q306 D314 S317 P318 D319 Q327 P330 E331 R352 K355 Y369 R378 S379 V380 C381 H395 H396 L397 N400 D403 M404 H417

L423 E424 ASN LEU GLU HIS

• Molecule 1: Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase

Chain D:

PRO GLU PHE PHE SER VAL ARG HIS LEU GLU LEU ALA GLY ASP ASP PRO TYR SER N56 V57 N58 C59 T60 K61 I62 L63 E69 T79 V80 GLN PHE LYS LYS ARG P86 C100 R107 K108 H131 M135 L163 Q167 A208 M219 K232 L242 R258

D263 G264 K265 L278 K279 T280 R293 N302 Q306 D314 S317 P330 E331 R352 E360 S364 P371 H395 N400 K401 F402 D405 I411 Q412 C413 H417 E424 ASN LEU GLU HIS

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

NAG1 NAG2 FUC3

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

NAG1 NAG2 FUC3

# 4   Data and refinement statistics   ⓘ

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 73.87Å    101.02Å    136.61Å<br>90.00°    93.42°      90.00° | Depositor |
| Resolution (Å) | 50.00  –  2.30<br>50.00  –  2.30 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 95.4 (50.00-2.30)<br>95.4 (50.00-2.30) | Depositor<br>EDS |
| $R_{merge}$ | 0.13 | Depositor |
| $R_{sym}$ | 0.12 | Depositor |
| $< I/\sigma(I) >$ [1] | 1.08 (at 2.20Å) | Xtriage |
| Refinement program | REFMAC 5.2.0019 | Depositor |
| R, $R_{free}$ | 0.165  ,  0.219<br>0.170  ,  0.221 | Depositor<br>DCC |
| $R_{free}$ test set | 4290 reflections (5.05%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 25.5 | Xtriage |
| Anisotropy | 0.692 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.35 , 52.2 | EDS |
| L-test for twinning[2] | $< \lvert L \rvert > = 0.48$, $< L^2 > = 0.31$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.95 | EDS |
| Total number of atoms | 12658 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 43.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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: UDP, NAG, NA, FUC, HTO

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 | 0.46 | 0/3096 | 0.56 | 0/4197 |
| 1 | B | 0.44 | 0/3032 | 0.56 | 0/4110 |
| 1 | C | 0.44 | 0/3010 | 0.57 | 0/4080 |
| 1 | D | 0.45 | 0/3048 | 0.57 | 0/4132 |
| All | All | 0.45 | 0/12186 | 0.57 | 0/16519 |

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 | 3015 | 0 | 2984 | 32 | 0 |
| 1 | B | 2956 | 0 | 2916 | 28 | 0 |
| 1 | C | 2934 | 0 | 2881 | 32 | 0 |
| 1 | D | 2966 | 0 | 2934 | 20 | 0 |
| 2 | E | 38 | 0 | 34 | 0 | 0 |
| 2 | F | 38 | 0 | 34 | 0 | 0 |
| 3 | A | 20 | 0 | 32 | 4 | 0 |
| 4 | A | 1 | 0 | 0 | 0 | 0 |
| 4 | C | 1 | 0 | 0 | 0 | 0 |

*Continued on next page...*

WORLDWIDE wwPDB PROTEIN DATA BANK

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 5 | A | 14 | 0 | 13 | 0 | 0 |
| 5 | B | 14 | 0 | 13 | 0 | 0 |
| 5 | C | 14 | 0 | 13 | 0 | 0 |
| 5 | D | 14 | 0 | 13 | 0 | 0 |
| 6 | A | 25 | 0 | 11 | 2 | 0 |
| 6 | B | 25 | 0 | 11 | 1 | 0 |
| 6 | C | 25 | 0 | 11 | 2 | 0 |
| 6 | D | 25 | 0 | 11 | 0 | 0 |
| 7 | A | 143 | 0 | 0 | 0 | 0 |
| 7 | B | 122 | 0 | 0 | 1 | 0 |
| 7 | C | 120 | 0 | 0 | 1 | 0 |
| 7 | D | 148 | 0 | 0 | 1 | 0 |
| All | All | 12658 | 0 | 11911 | 110 | 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 5.

The worst 5 of 110 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:A:89:THR:HG22 | 1:A:91:HIS:H | 1.22 | 1.04 |
| 1:A:89:THR:HG23 | 1:A:90:PRO:HD2 | 1.48 | 0.94 |
| 1:C:135:MET:HE2 | 1:C:404:MET:H | 1.43 | 0.83 |
| 1:D:302:ASN:O | 1:D:306:GLN:HG2 | 1.79 | 0.82 |
| 1:C:187:TYR:HE1 | 6:C:600:UDP:O3B | 1.62 | 0.82 |

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 | 369/391 (94%) | 359 (97%) | 9 (2%) | 1 (0%) | 37 | 47 |
| 1 | B | 360/391 (92%) | 354 (98%) | 6 (2%) | 0 | 100 | 100 |
| 1 | C | 357/391 (91%) | 344 (96%) | 13 (4%) | 0 | 100 | 100 |
| 1 | D | 362/391 (93%) | 354 (98%) | 8 (2%) | 0 | 100 | 100 |
| All | All | 1448/1564 (93%) | 1411 (97%) | 36 (2%) | 1 (0%) | 48 | 60 |

All (1) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 423 | LEU |

### 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 | 333/351 (95%) | 318 (96%) | 15 (4%) | 23 | 34 |
| 1 | B | 326/351 (93%) | 312 (96%) | 14 (4%) | 25 | 36 |
| 1 | C | 323/351 (92%) | 310 (96%) | 13 (4%) | 27 | 40 |
| 1 | D | 328/351 (93%) | 318 (97%) | 10 (3%) | 36 | 52 |
| All | All | 1310/1404 (93%) | 1258 (96%) | 52 (4%) | 27 | 40 |

5 of 52 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 395 | HIS |
| 1 | C | 299 | VAL |
| 1 | D | 317 | SER |
| 1 | C | 107 | ARG |
| 1 | C | 138 | ARG |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | D | 306 | GLN |
| 1 | D | 313 | GLN |
| 1 | D | 400 | ASN |
| 1 | D | 340 | ASN |
| 1 | B | 340 | ASN |

### 5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5   Carbohydrates ⓘ

6 monosaccharides 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 | NAG | E | 1 | 1,2 | 14,14,15 | 0.58 | 0 | 17,19,21 | 1.35 | 1 (5%) |
| 2 | NAG | E | 2 | 2 | 14,14,15 | 0.54 | 0 | 17,19,21 | 0.65 | 0 |
| 2 | FUC | E | 3 | 2 | 10,10,11 | 0.64 | 0 | 14,14,16 | 0.69 | 0 |
| 2 | NAG | F | 1 | 1,2 | 14,14,15 | 0.60 | 0 | 17,19,21 | 1.15 | 1 (5%) |
| 2 | NAG | F | 2 | 2 | 14,14,15 | 0.51 | 0 | 17,19,21 | 0.73 | 0 |
| 2 | FUC | F | 3 | 2 | 10,10,11 | 0.69 | 0 | 14,14,16 | 0.86 | 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 | NAG | E | 1 | 1,2 | - | 0/6/23/26 | 0/1/1/1 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 2 | NAG | E | 2 | 2 | - | 1/6/23/26 | 0/1/1/1 |
| 2 | FUC | E | 3 | 2 | - | - | 0/1/1/1 |
| 2 | NAG | F | 1 | 1,2 | - | 0/6/23/26 | 0/1/1/1 |
| 2 | NAG | F | 2 | 2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | FUC | F | 3 | 2 | - | - | 0/1/1/1 |

There are no bond length outliers.

All (2) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|------------|---------|
| 2 | E | 1 | NAG | C1-O5-C5 | 4.61 | 118.36 | 112.19 |
| 2 | F | 1 | NAG | C1-O5-C5 | 3.86 | 117.35 | 112.19 |

There are no chirality outliers.

All (3) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 2 | F | 2 | NAG | O5-C5-C6-O6 |
| 2 | F | 2 | NAG | C4-C5-C6-O6 |
| 2 | E | 2 | NAG | C1-C2-N2-C7 |

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

## Oligosaccharide Chain E


**Bond lengths**


**Bond angles**


**Torsions**


**Rings**

**Oligosaccharide Chain F**



Bond lengths



Bond angles



Torsions



Rings
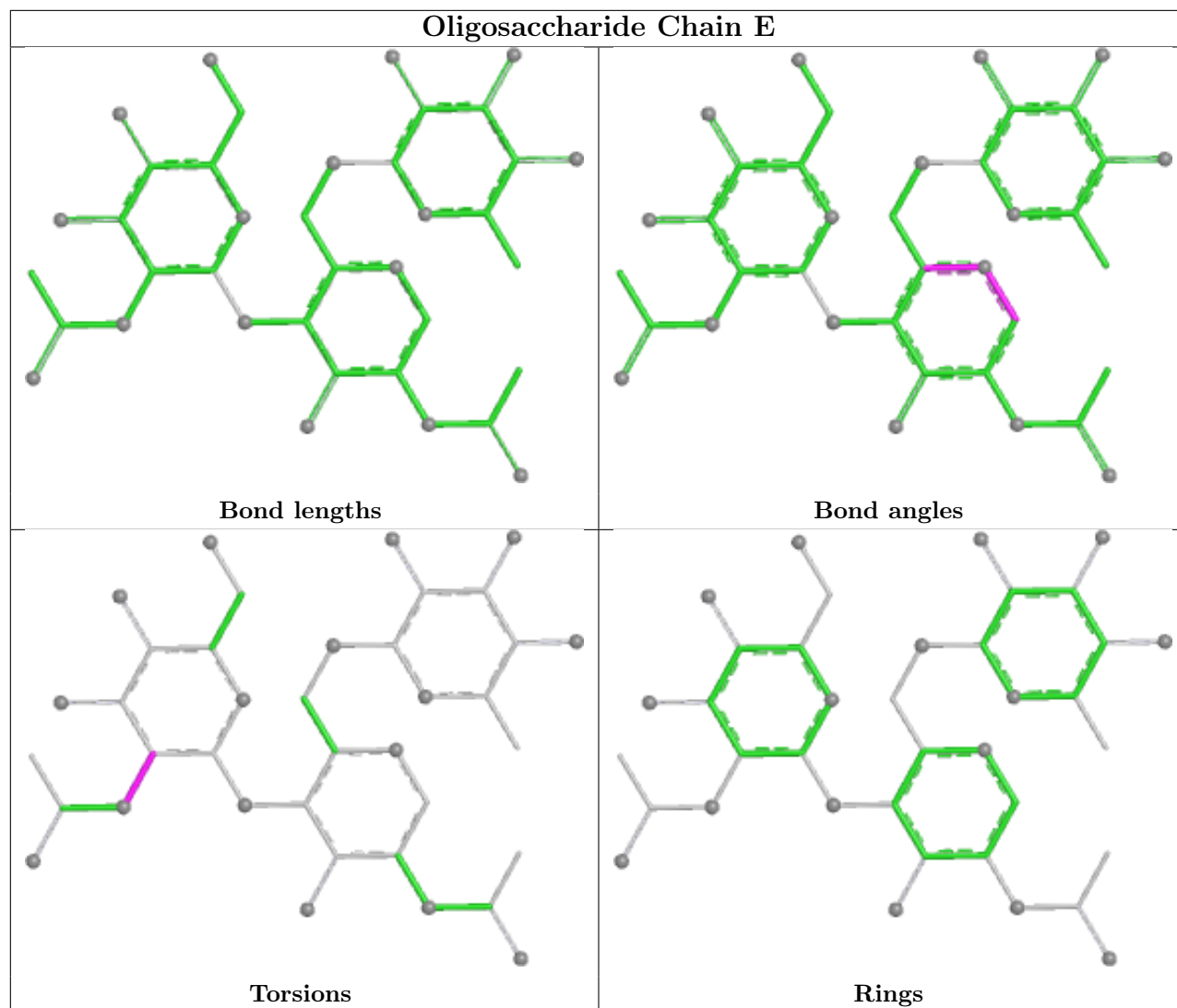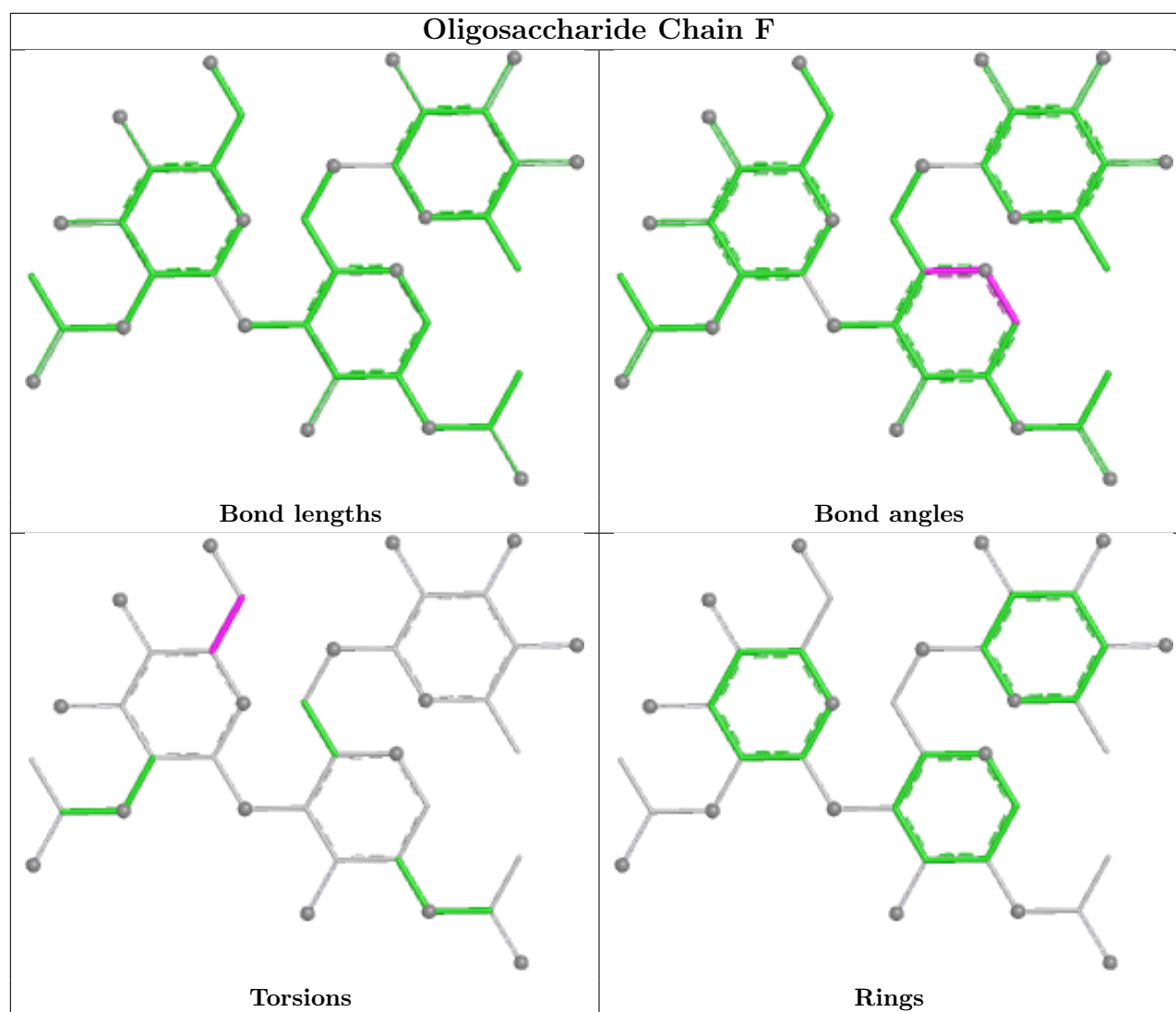
## 5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 2 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 | #$|Z| > 2$ |
| 5 | NAG | C | 593 | 1 | 14,14,15 | 0.46 | 0 | 17,19,21 | 0.98 | 1 (5%) |
| 6 | UDP | B | 599 | - | 25,26,26 | 1.03 | 2 (8%) | 38,40,40 | 1.58 | 6 (15%) |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|------------|--------|------|------------|
| | | | | | Counts | RMSZ | #\|Z\| > 2 | Counts | RMSZ | #\|Z\| > 2 |
| 6 | UDP | C | 600 | - | 25,26,26 | 0.99 | 1 (4%) | 38,40,40 | 1.48 | 4 (10%) |
| 6 | UDP | D | 601 | - | 25,26,26 | 1.01 | 1 (4%) | 38,40,40 | 1.53 | 5 (13%) |
| 5 | NAG | D | 597 | 1 | 14,14,15 | 0.52 | 0 | 17,19,21 | 0.90 | 0 |
| 6 | UDP | A | 598 | - | 25,26,26 | 0.91 | 1 (4%) | 38,40,40 | 1.42 | 5 (13%) |
| 3 | HTO | A | 584 | - | 9,9,9 | 0.47 | 0 | 10,10,10 | 0.62 | 0 |
| 3 | HTO | A | 585 | - | 9,9,9 | 0.38 | 0 | 10,10,10 | 0.77 | 0 |
| 5 | NAG | B | 592 | 1 | 14,14,15 | 0.47 | 0 | 17,19,21 | 0.79 | 0 |
| 5 | NAG | A | 588 | 1 | 14,14,15 | 0.50 | 0 | 17,19,21 | 0.92 | 1 (5%) |

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 |
|-----|------|-------|-----|------|---------|----------|-------|
| 5 | NAG | C | 593 | 1 | - | 0/6/23/26 | 0/1/1/1 |
| 6 | UDP | B | 599 | - | - | 4/16/32/32 | 0/2/2/2 |
| 6 | UDP | C | 600 | - | - | 2/16/32/32 | 0/2/2/2 |
| 6 | UDP | D | 601 | - | - | 6/16/32/32 | 0/2/2/2 |
| 5 | NAG | D | 597 | 1 | - | 2/6/23/26 | 0/1/1/1 |
| 6 | UDP | A | 598 | - | - | 2/16/32/32 | 0/2/2/2 |
| 3 | HTO | A | 584 | - | - | 7/10/10/10 | - |
| 3 | HTO | A | 585 | - | - | 1/10/10/10 | - |
| 5 | NAG | B | 592 | 1 | - | 4/6/23/26 | 0/1/1/1 |
| 5 | NAG | A | 588 | 1 | - | 0/6/23/26 | 0/1/1/1 |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|-------------|----------|
| 6 | B | 599 | UDP | PA-O3A | 2.55 | 1.62 | 1.59 |
| 6 | D | 601 | UDP | C6-C5 | 2.21 | 1.40 | 1.35 |
| 6 | B | 599 | UDP | C6-C5 | 2.06 | 1.39 | 1.35 |
| 6 | A | 598 | UDP | C6-C5 | 2.01 | 1.39 | 1.35 |
| 6 | C | 600 | UDP | PA-O3A | 2.01 | 1.61 | 1.59 |

The worst 5 of 22 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|------|----------------|-------------|
| 6 | B | 599 | UDP | C4-N3-C2 | -5.35 | 119.97 | 126.61 |

*Continued on next page...*

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|----------------|-------------|
| 6 | C | 600 | UDP | C4-N3-C2 | -5.34 | 119.98 | 126.61 |
| 6 | D | 601 | UDP | C4-N3-C2 | -5.30 | 120.04 | 126.61 |
| 6 | A | 598 | UDP | C4-N3-C2 | -4.70 | 120.77 | 126.61 |
| 6 | B | 599 | UDP | C5-C4-N3 | 3.91 | 120.28 | 114.80 |

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 3 | A | 584 | HTO | C1-C2-C3-O3 |
| 3 | A | 584 | HTO | C1-C2-C3-C4 |
| 3 | A | 584 | HTO | O2-C2-C3-O3 |
| 3 | A | 584 | HTO | O2-C2-C3-C4 |
| 5 | B | 592 | NAG | C8-C7-N2-C2 |

There are no ring outliers.

5 monomers are involved in 8 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 6 | B | 599 | UDP | 1 | 0 |
| 6 | C | 600 | UDP | 2 | 0 |
| 6 | A | 598 | UDP | 2 | 0 |
| 3 | A | 584 | HTO | 1 | 0 |
| 3 | A | 585 | HTO | 3 | 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 UDP B 599**



Bond lengths



Bond angles



Torsions



Rings

**Ligand UDP C 600**



Bond lengths



Bond angles



Torsions



Rings

## Ligand UDP D 601


Bond lengths


Bond angles


Torsions


Rings

## Ligand UDP A 598


Bond lengths


Bond angles


Torsions


Rings
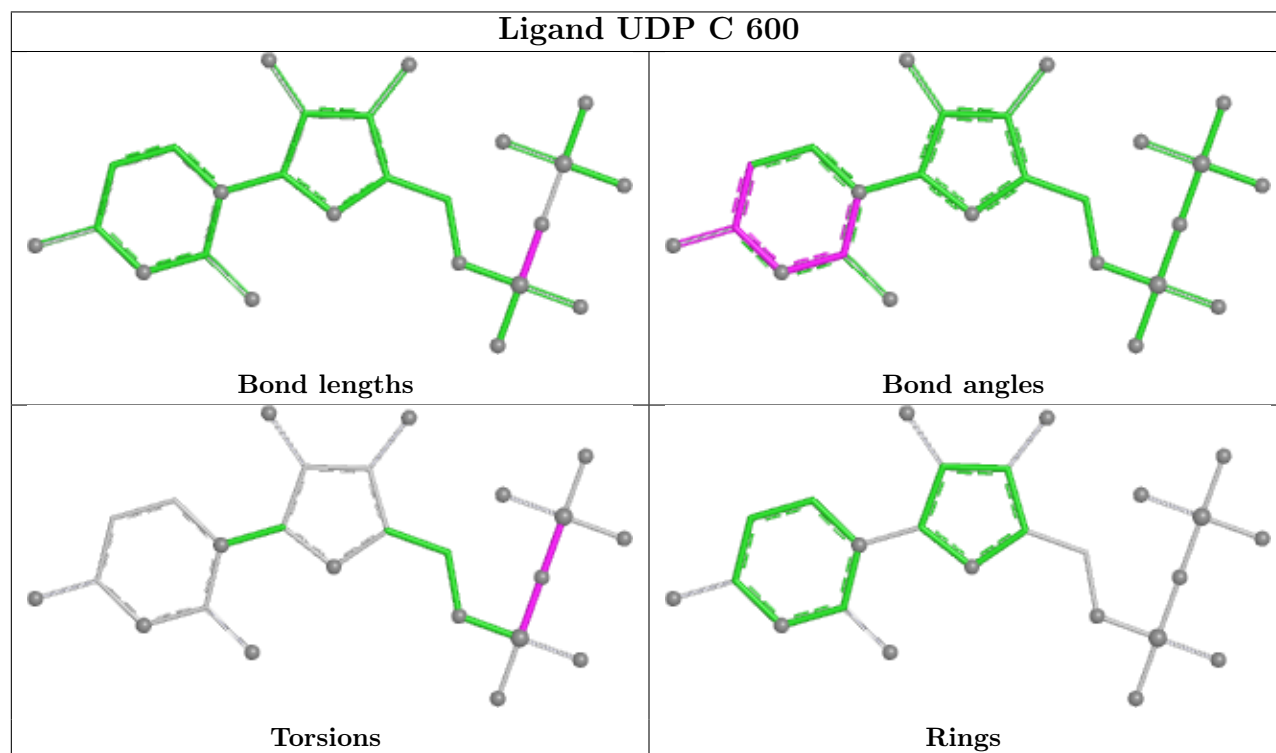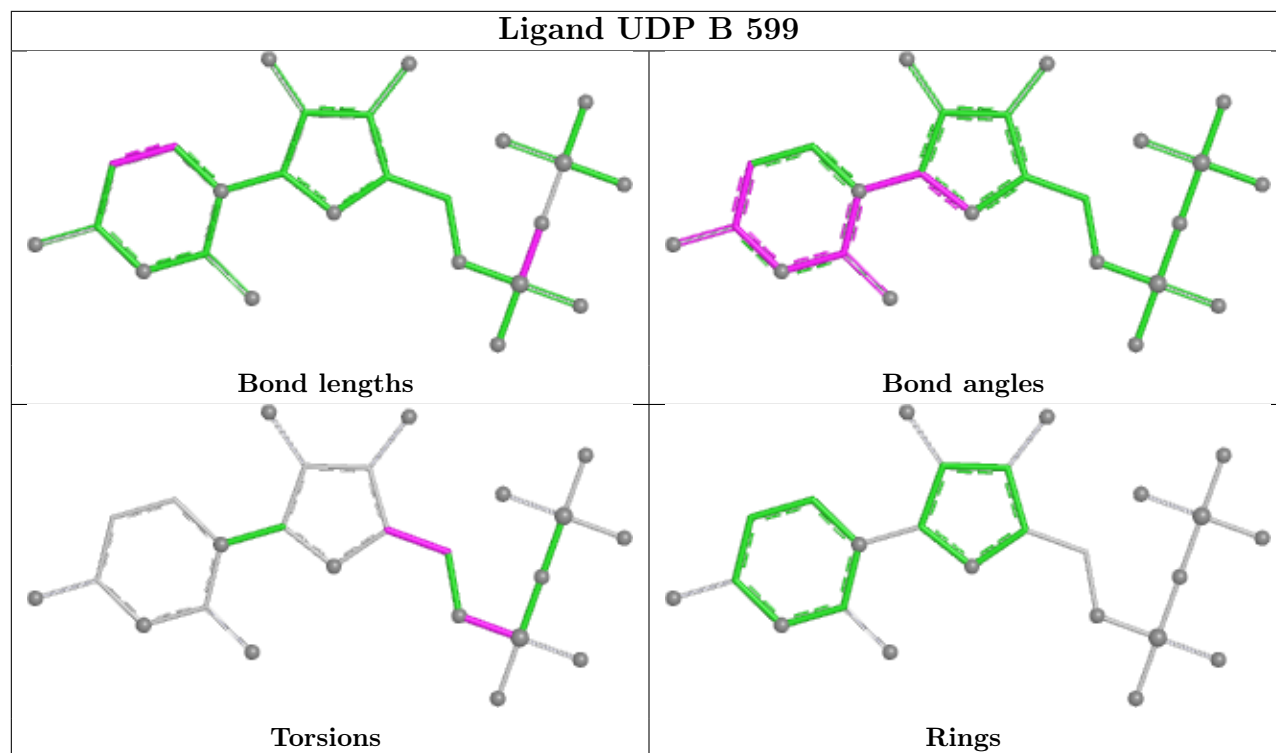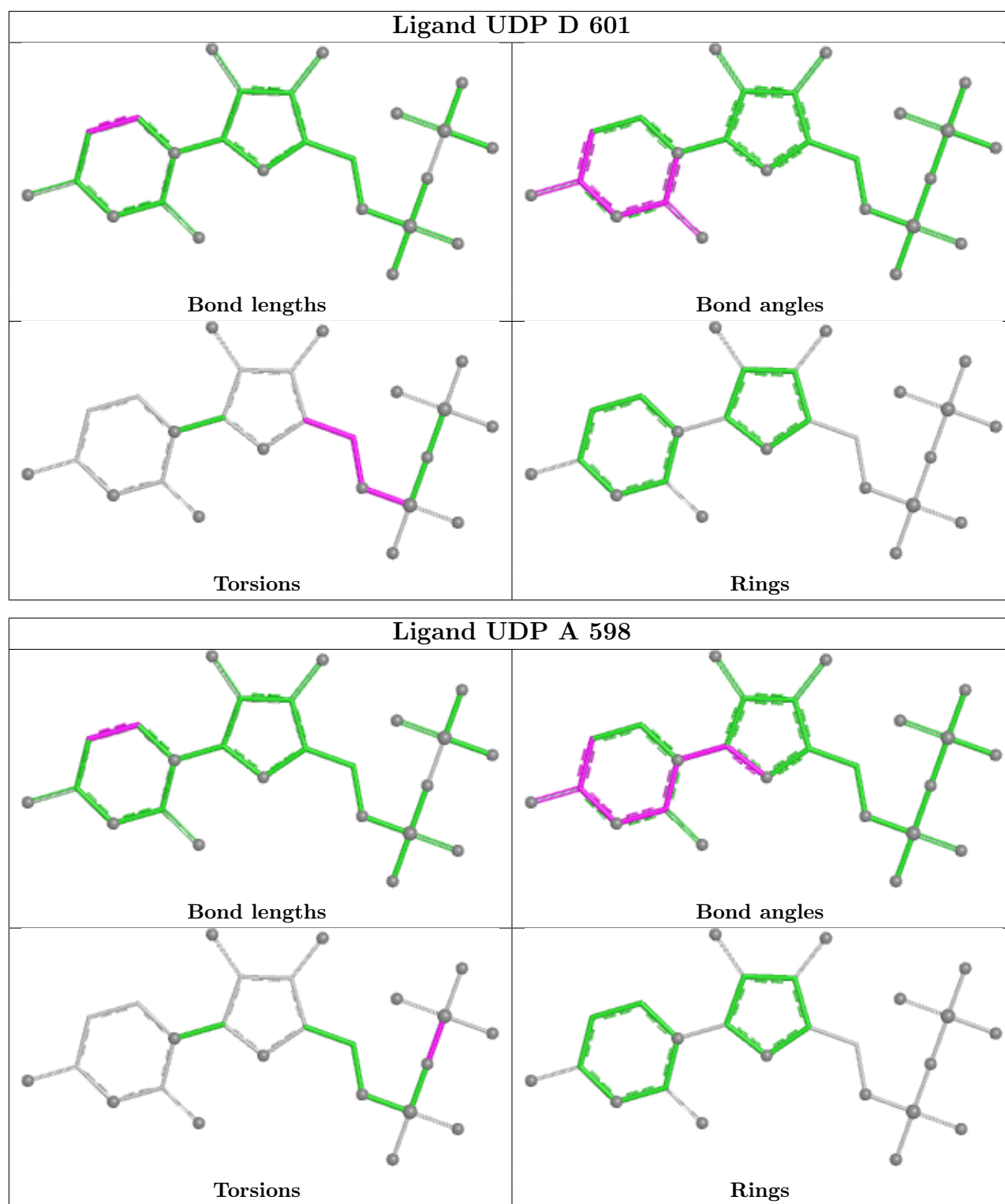
## 5.7　Other polymers ⓘ

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, $95^{th}$ 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 | 369/391 (94%) | -0.15 | 6 (1%) | 70 | 71 | 15, 38, 66, 96 | 2 (0%) |
| 1 | B | 363/391 (92%) | 0.03 | 18 (4%) | 35 | 36 | 24, 42, 76, 103 | 1 (0%) |
| 1 | C | 360/391 (92%) | -0.03 | 11 (3%) | 51 | 53 | 22, 39, 73, 95 | 1 (0%) |
| 1 | D | 364/391 (93%) | -0.09 | 7 (1%) | 66 | 67 | 22, 39, 63, 88 | 2 (0%) |
| All | All | 1456/1564 (93%) | -0.06 | 42 (2%) | 54 | 55 | 15, 39, 71, 103 | 6 (0%) |

The worst 5 of 42 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | B | 86 | PRO | 4.9 |
| 1 | D | 86 | PRO | 4.1 |
| 1 | C | 86 | PRO | 3.9 |
| 1 | C | 75 | LEU | 3.7 |
| 1 | B | 77 | ILE | 3.7 |

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

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.
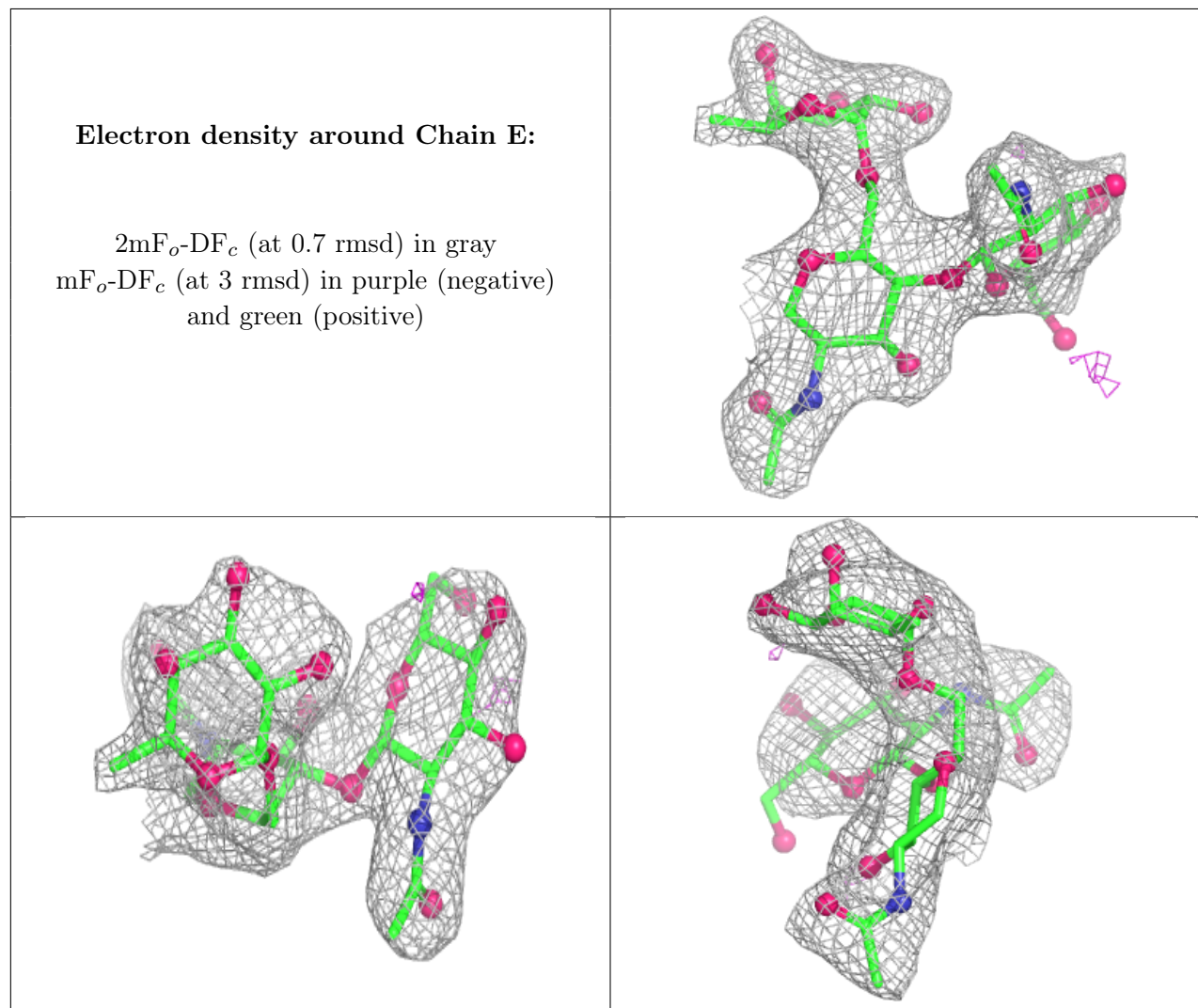
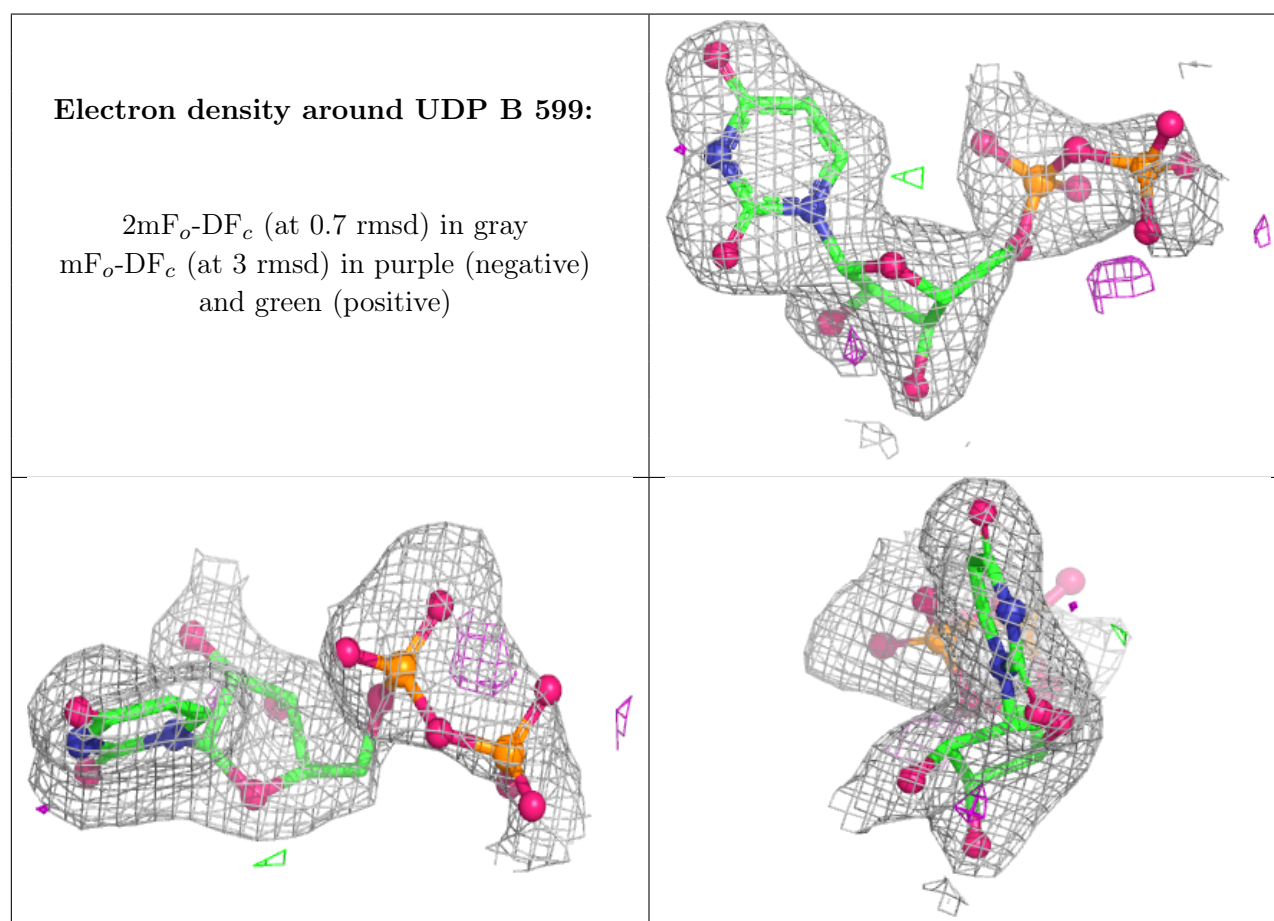| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|-----|------|-------|-----|-------|------|-----|------------------|-------|
| 2 | NAG | F | 2 | 14/15 | 0.81 | 0.16 | 51,91,120,122 | 0 |
| 2 | FUC | F | 3 | 10/11 | 0.87 | 0.12 | 60,66,79,86 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors($Å^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 2 | NAG | E | 2 | 14/15 | 0.88 | 0.13 | 48,80,98,100 | 0 |
| 2 | FUC | E | 3 | 10/11 | 0.94 | 0.10 | 45,68,78,86 | 0 |
| 2 | NAG | E | 1 | 14/15 | 0.95 | 0.08 | 38,57,67,67 | 0 |
| 2 | NAG | F | 1 | 14/15 | 0.95 | 0.09 | 37,52,65,70 | 0 |

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.
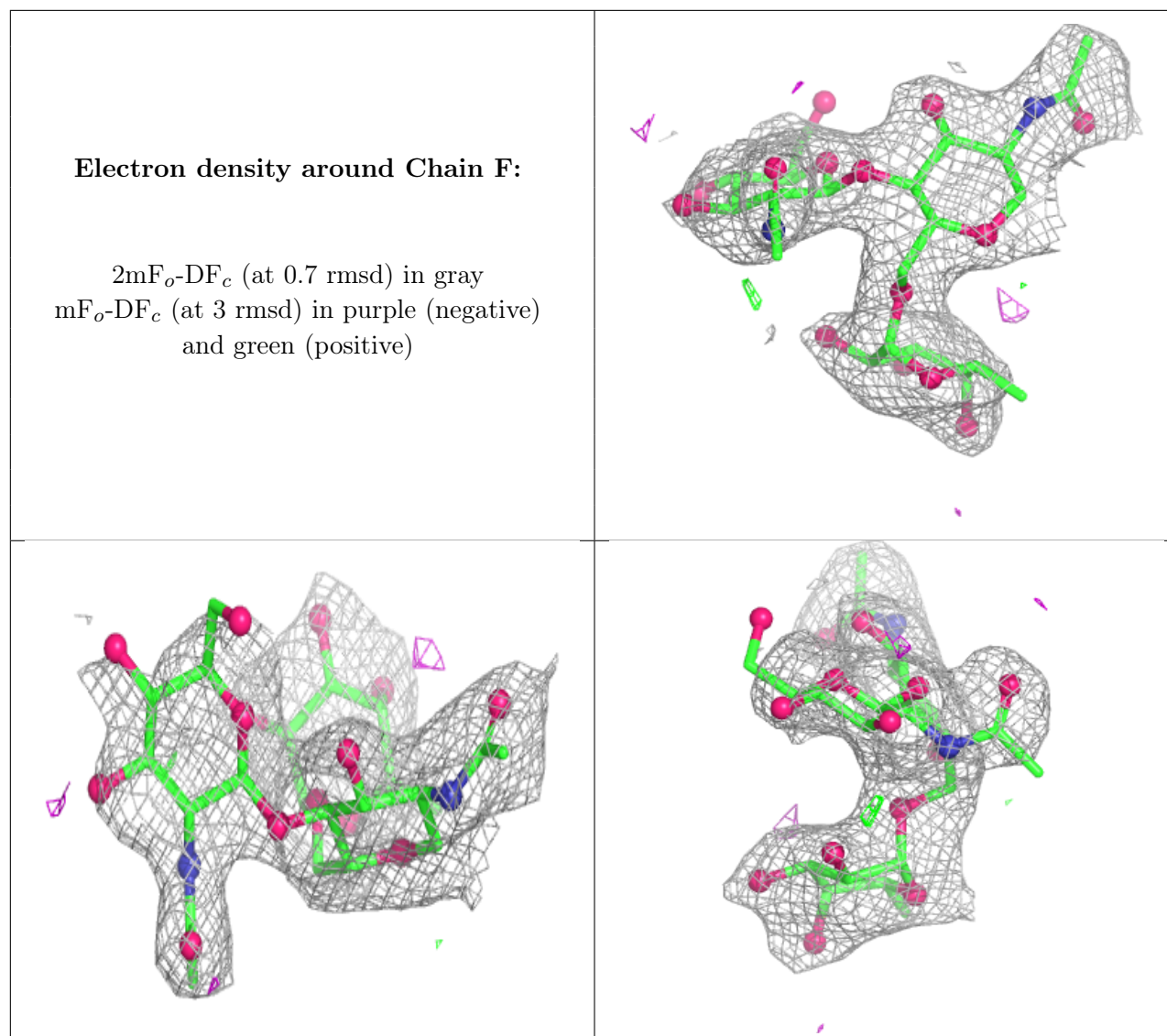


**Electron density around Chain E:**

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

Electron density around Chain F:

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

## 6.4 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, $95^{th}$ 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($\text{Å}^2$) | Q<0.9 |
|-----|------|-------|-----|-------|------|-----|--------------------------|-------|
| 5 | NAG | D | 597 | 14/15 | 0.65 | 0.18 | 82,105,118,118 | 0 |
| 5 | NAG | B | 592 | 14/15 | 0.69 | 0.18 | 94,112,124,125 | 0 |
| 5 | NAG | C | 593 | 14/15 | 0.70 | 0.21 | 75,105,122,129 | 0 |
| 3 | HTO | A | 584 | 10/10 | 0.75 | 0.25 | 47,84,110,113 | 0 |
| 5 | NAG | A | 588 | 14/15 | 0.79 | 0.17 | 73,93,107,125 | 0 |
| 3 | HTO | A | 585 | 10/10 | 0.83 | 0.19 | 42,64,75,77 | 0 |
| 6 | UDP | B | 599 | 25/25 | 0.88 | 0.12 | 33,63,149,157 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å²) | Q<0.9 |
|-----|------|-------|-----|-------|------|------|----------------|-------|
| 6 | UDP | D | 601 | 25/25 | 0.92 | 0.11 | 33,59,142,152 | 0 |
| 4 | NA | A | 586 | 1/1 | 0.94 | 0.08 | 43,43,43,43 | 0 |
| 6 | UDP | C | 600 | 25/25 | 0.95 | 0.07 | 36,53,73,90 | 0 |
| 4 | NA | C | 587 | 1/1 | 0.95 | 0.06 | 38,38,38,38 | 0 |
| 6 | UDP | A | 598 | 25/25 | 0.96 | 0.06 | 27,40,65,90 | 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.
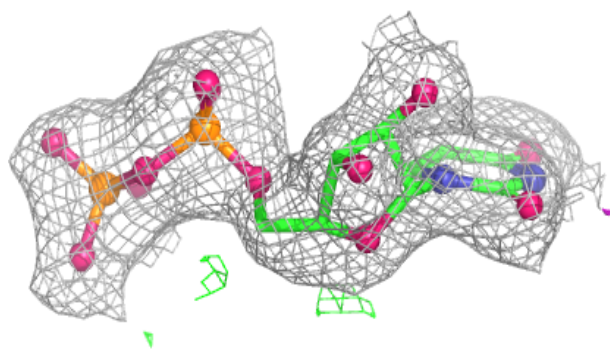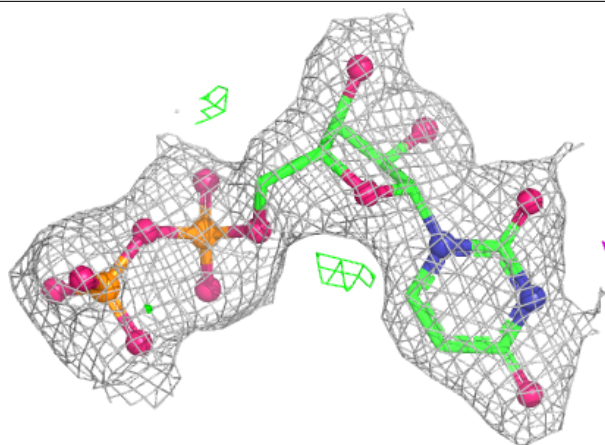


**Electron density around UDP B 599:**

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

**Electron density around UDP D 601:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
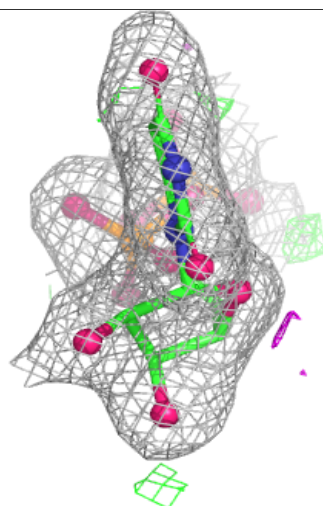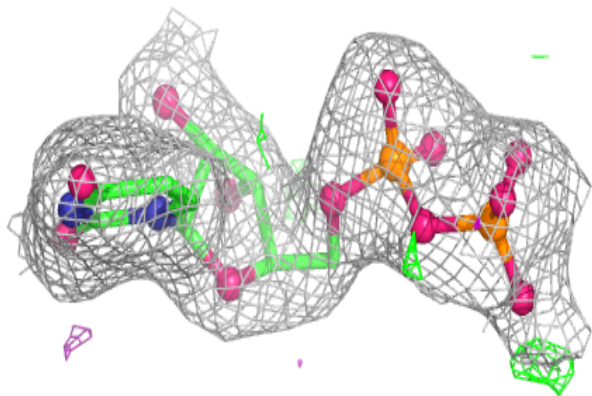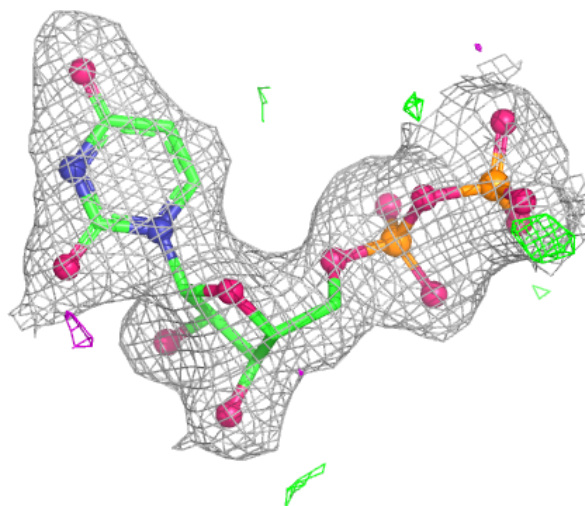and green (positive)
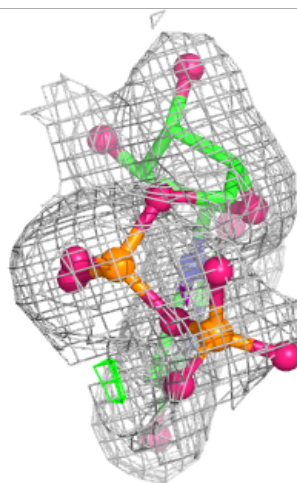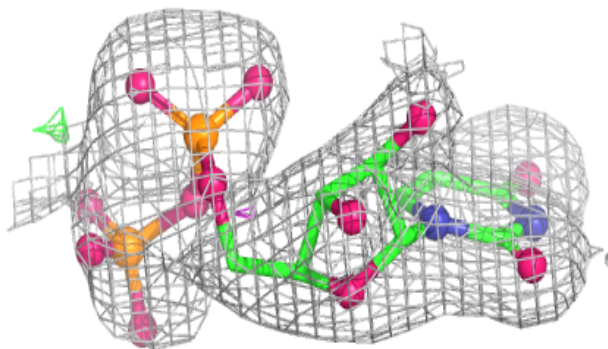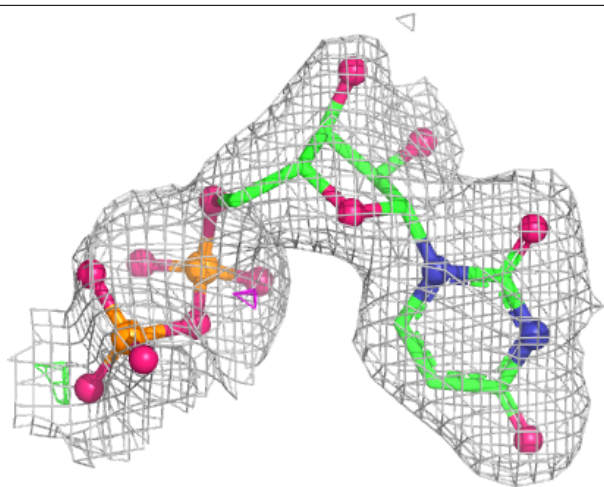
**Electron density around UDP C 600:**

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
$mF_o-DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
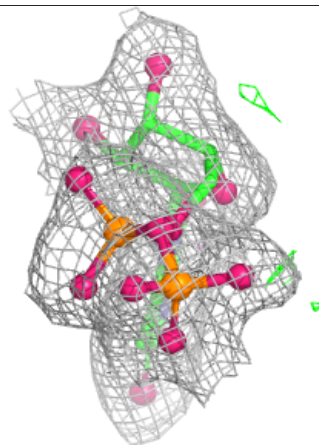
**Electron density around UDP A 598:**

$2mF_o\text{-}DF_c$ (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.