# WORLDWIDE wwPDB PROTEIN DATA BANK

# wwPDB EM Validation Summary Report (i)

Dec 9, 2024 – 08:18 PM JST

|  |  |  |
|---|---|---|
| PDB ID | : | 9JJE |
| EMDB ID | : | EMD-61524 |
| Title | : | Nematostella vectensis TRPM2 tetramer in complex with ADPRP/Ca2+ |
| Authors | : | Jiang, Y.; Zhang, Z.; Toth, B.; Szollosi, A.; Csanady, L. |
| Deposited on | : | 2024-09-13 |
| Resolution | : | 2.52 Å(reported) |

This is a wwPDB EM Validation Summary 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/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

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

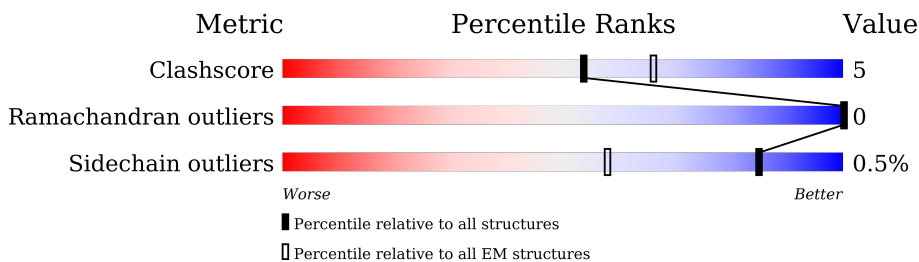|  |  |  |
|---|---|---|
| EMDB validation analysis | : | FAILED |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| MolProbity | : | 4.02b-467 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20231227.v01 (using entries in the PDB archive December 27th 2023) |
| MapQ | : | FAILED |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.40 |

# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.52 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 |
|---|---|---|
| Clashscore | | 5 |
| Ramachandran outliers | | 0 |
| Sidechain outliers | | 0.5% |

*Worse*      *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures

| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|---|---|---|
| Clashscore | 210492 | 15764 |
| Ramachandran outliers | 207382 | 16835 |
| Sidechain outliers | 206894 | 16415 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 1560 | 52%   9%   40% |
| 1 | B | 1560 | 51%   9%   40% |
| 1 | C | 1560 | 52%   9%   40% |
| 1 | D | 1560 | 52%   8%   40% |
| 2 | E | 2 | 100% |
| 2 | F | 2 | 100% |
| 2 | G | 2 | 100% |
| 2 | H | 2 | 100% |

# 2   Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 30472 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Transient receptor potential cation channel subfamily M member-like 2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|-------|
| 1 | A | 941 | Total | C | N | O | S | 0 | 0 |
| | | | 7549 | 4906 | 1243 | 1359 | 41 | | |
| 1 | B | 941 | Total | C | N | O | S | 0 | 0 |
| | | | 7549 | 4906 | 1243 | 1359 | 41 | | |
| 1 | C | 941 | Total | C | N | O | S | 0 | 0 |
| | | | 7549 | 4906 | 1243 | 1359 | 41 | | |
| 1 | D | 941 | Total | C | N | O | S | 0 | 0 |
| | | | 7549 | 4906 | 1243 | 1359 | 41 | | |

There are 36 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | 1552 | SER | - | expression tag | UNP A7T1N0 |
| A | 1553 | ASN | - | expression tag | UNP A7T1N0 |
| A | 1554 | SER | - | expression tag | UNP A7T1N0 |
| A | 1555 | LEU | - | expression tag | UNP A7T1N0 |
| A | 1556 | GLU | - | expression tag | UNP A7T1N0 |
| A | 1557 | VAL | - | expression tag | UNP A7T1N0 |
| A | 1558 | LEU | - | expression tag | UNP A7T1N0 |
| A | 1559 | PHE | - | expression tag | UNP A7T1N0 |
| A | 1560 | GLN | - | expression tag | UNP A7T1N0 |
| B | 1552 | SER | - | expression tag | UNP A7T1N0 |
| B | 1553 | ASN | - | expression tag | UNP A7T1N0 |
| B | 1554 | SER | - | expression tag | UNP A7T1N0 |
| B | 1555 | LEU | - | expression tag | UNP A7T1N0 |
| B | 1556 | GLU | - | expression tag | UNP A7T1N0 |
| B | 1557 | VAL | - | expression tag | UNP A7T1N0 |
| B | 1558 | LEU | - | expression tag | UNP A7T1N0 |
| B | 1559 | PHE | - | expression tag | UNP A7T1N0 |
| B | 1560 | GLN | - | expression tag | UNP A7T1N0 |
| C | 1552 | SER | - | expression tag | UNP A7T1N0 |
| C | 1553 | ASN | - | expression tag | UNP A7T1N0 |
| C | 1554 | SER | - | expression tag | UNP A7T1N0 |

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

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| C | 1555 | LEU | - | expression tag | UNP A7T1N0 |
| C | 1556 | GLU | - | expression tag | UNP A7T1N0 |
| C | 1557 | VAL | - | expression tag | UNP A7T1N0 |
| C | 1558 | LEU | - | expression tag | UNP A7T1N0 |
| C | 1559 | PHE | - | expression tag | UNP A7T1N0 |
| C | 1560 | GLN | - | expression tag | UNP A7T1N0 |
| D | 1552 | SER | - | expression tag | UNP A7T1N0 |
| D | 1553 | ASN | - | expression tag | UNP A7T1N0 |
| D | 1554 | SER | - | expression tag | UNP A7T1N0 |
| D | 1555 | LEU | - | expression tag | UNP A7T1N0 |
| D | 1556 | GLU | - | expression tag | UNP A7T1N0 |
| D | 1557 | VAL | - | expression tag | UNP A7T1N0 |
| D | 1558 | LEU | - | expression tag | UNP A7T1N0 |
| D | 1559 | PHE | - | expression tag | UNP A7T1N0 |
| D | 1560 | GLN | - | expression tag | UNP A7T1N0 |

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---------|-------|
| 2 | E | 2 | Total 28 | C 16 | N 2 | O 10 | 0 | 0 |
| 2 | F | 2 | Total 28 | C 16 | N 2 | O 10 | 0 | 0 |
| 2 | G | 2 | Total 28 | C 16 | N 2 | O 10 | 0 | 0 |
| 2 | H | 2 | Total 28 | C 16 | N 2 | O 10 | 0 | 0 |

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

| Mol | Chain | Residues | Atoms | | AltConf |
|-----|-------|----------|-------|---|---------|
| 3 | A | 1 | Total 1 | Ca 1 | 0 |
| 3 | B | 1 | Total 1 | Ca 1 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 3 | C | 1 | Total | Ca | 0 |
| | | | 1 | 1 | |
| 3 | D | 1 | Total | Ca | 0 |
| | | | 1 | 1 | |

- Molecule 4 is [(2R,3R,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-3-HYDROXY-4-(PHOSPH ONOOXY)TETRAHYDROFURAN-2-YL]METHYL [(2R,3S,4R,5R)-3,4,5-TRIHYDROX YTETRAHYDROFURAN-2-YL]METHYL DIHYDROGEN DIPHOSPHATE (three-letter code: A2R) (formula: $C_{15}H_{24}N_5O_{17}P_3$) (labeled as "Ligand of Interest" by depositor).



| Mol | Chain | Residues | Atoms | | | | | AltConf |
|---|---|---|---|---|---|---|---|---|
| 4 | A | 1 | Total | C | N | O | P | 0 |
| | | | 40 | 15 | 5 | 17 | 3 | |
| 4 | B | 1 | Total | C | N | O | P | 0 |
| | | | 40 | 15 | 5 | 17 | 3 | |
| 4 | C | 1 | Total | C | N | O | P | 0 |
| | | | 40 | 15 | 5 | 17 | 3 | |
| 4 | D | 1 | Total | C | N | O | P | 0 |
| | | | 40 | 15 | 5 | 17 | 3 | |

# 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. 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. 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: Transient receptor potential cation channel subfamily M member-like 2

GLY VAL MET LEU GLN GLY GLY LYS LYS VAL VAL LEU GLU PHE VAL ALA ILE ILE ARG LYS ASP ASN ASN GLN TRP ALA ILE PRO GLY MET VAL GLU PRO GLY GLN LEU VAL THR ALA LEU LYS ALA GLU GLY PHE GLU GLU ALA MET ALA LYS LEU ASN VAL SER GLU GLU LYS

GLU ARG ILE ALA LYS GLN ILE GLU ARG LEU PHE GLN GLN GLN GLY ILE GLU TYR LYS GLY TYR VAL ASP ASP PRO ARG ASN THR ASP ASN ALA TRP MET GLU THR VAL ALA VAL ASN PHE HIS ASP ASP LYS GLY LEU GLU PHE GLY ASP ILE THR LEU ALA GLY ASP ASP ALA ALA

ALA VAL ARG GLN ARG SER GLY ILE PRO TYR ALA HIS VAL ILE LEU SER VAL LYS MET ARG ASP ALA ALA PHE SER SER LEU GLU VAL LEU PHE GLN

• Molecule 1: Transient receptor potential cation channel subfamily M member-like 2

Chain B:


51% 9% 40%

MET GLY LYS SER ASP SER PHE THR PRO LEU TYR ASP GLY ASP SER HIS VAL LEU ASN PHE GLY SER ASN GLN LEU SER GLN SER LYS LYS SER TRP ILE ALA ARG ASN PHE SER ARG ARG GLU CYS ILE ARG PHE VAL PRO LYS SER HIS VAL SER ARG CYS LYS CYS

GLY ARG PRO GLU ARG GLU HIS SER GLN GLN ALA LEU GLU LEU SER GLY GLN GLY GLU GLU TRP ASN VAL ALA SER CYS THR THR LYS HIS PRO THR ASN A94 E97 P110 M114 T118 V139 V142 T143 P153 R154 L155 R156 T176 G177 G178 T179 N180 T181

M184 E193 Q194 A210 D216 LYS GLN SER ASP LEU ILE SER GLU LYS ASN GLY LYS TYR PRO ALA LEU TYR SER MET GLU PRO THR PRO GLY HIS GLN G243 H250 V256 Y264 G265 I268 G269 M270 R271 I278 V281 LYS THR ASP SER ARG SER GLU ALA GLY

SER I292 V296 V297 L298 L299 E302 T307 A321 V322 G328 R329 V346 I351 A364 V367 E368 V369 Y395 S396 L397 E400 L407 A408 I409 L429 R433 L461 K465 Q474 L477 S478 M479 R480 E481 F482 L483 S484 L485 D486 P496

G497 T500 I501 K502 P503 L504 R512 D520 V525 L529 E535 K540 F545 GLY GLU LEU ALA SER TYR ALA GLY LEU VAL LEU LYS ASN ARG LYS SER LYS ASP LEU LEU ALA ASN ILE ASN ARG ILE ASP P575 L576 P577 T578 P579 Y580 L581 L589

R592 R593 R623 I630 D636 D639 R642 E660 T661 L662 S663 Q664 T665 E681 V684 V718 C721 L722 I728 P736 ASP GLU GLN HIS GLU ARG GLN ALA ALA GLU ARG GLU HIS LYS GLN SER LEU ASN GLN LYS SER LYS VAL LYS SER HIS LYS

GLU LYS ASN ASP ALA PRO VAL VAL PRO VAL TYR ARG SER LYS GLU GLU LYS ALA VAL SER ASN ASP GLU GLU ARG VAL GLY THR GLU GLU ASP ASP PHE GLN LEU ASP TYR ILE PRO GLY ILE ARG GLU HIS LYS SER ASP SER MET GLU VAL ILE MET ARG ASN LYS LYS LEU

GLY F826 K840 N844 Y848 Y855 L860 D867 K870 L882 E892 R895 K900 Y912 D915 T916 W917 N918 F919 A931 L934 R941 I942 I949 F957 L962 N968 L975 T996 R1016 N1017 N1018 V1021 T1024 S1025

Q1028 Y1029 P1030 E1037 L1038 F1039 E1042 F1051 D1057 N1074 I1075 R1114 A1118 P1119 P1120 L1121 V1122 H1126 I1131 ARG TRP VAL TRP ARG MET CYS ARG CYS GLY HIS PRO PRO ARG GLY SER THR MET K1150 Q1167 L1206 ASP SER MET SER GLY THR ARG ALA

HIS ALA LEU THR ASP GLY ASN GLY LEU GLU GLY GLY HIS ASP SER GLU GLY ARG LEU ALA ARG MET GLU VAL GLU LEU SER SER ASN SER GLU LEU GLN LYS ILE LEU ALA LEU GLN GLN GLN PRO PRO VAL LYS GLY GLN ALA ALA VAL PRO ILE GLN LEU THR LEU LEU HIS

TYR LYS ALA ARG SER SER PRO TYR PRO GLY SER THR ALA LYS ARG PHE ALA VAL GLN ASP ASN MET VAL ASP TRP GLN VAL PRO PHE ASP TYR LYS PRO VAL ASN THR TYR ALA PRO VAL VAL LEU ALA ASN PRO VAL TRP ALA ASP LEU MET ALA MET SER PRO ARG PRO

GLU LEU PRO ASN GLN MET ASP HIS THR CYS ASN VAL ASN ARG SER TYR ASN GLY THR TYR VAL LYS ASP GLY LEU PRO LEU ASN PRO MET GLY ARG THR GLY MET GLN GLY ARG GLY LEU LEU GLY ARG PHE GLY PRO ASN HIS ALA ALA ASP PRO VAL VAL THR ARG TRP

LYS ARG THR SER ALA GLY VAL GLY VAL MET LEU GLN GLY GLY LYS VAL VAL LEU PHE GLN GLY GLN ILE TYR LYS TYR VAL ALA ILE GLN ILE ARG LYS ASP ASN GLN THR TYR VAL ALA LEU LEU LYS ALA GLU GLU PHE GLY GLU ALA MET ALA LYS LEU ASN VAL

SER GLN GLU GLU LYS GLU ARG ILE ALA LYS GLN ILE GLU ARG LEU PHE ALA ILE LYS ALA LYS GLN ILE GLY TYR VAL ASP PRO ARG ASN THR ASN ALA TRP MET GLU THR VAL ALA VAL ASN PHE HIS ASP ASP LYS GLY LEU GLU PHE GLY ASP ILE THR LEU GLN ALA

GLY ASP ASP ALA ALA ALA VAL ARG TRP GLN ARG VAL SER GLY ASN ILE PRO LEU TYR ALA SER HIS SER ILE LEU GLU LYS VAL ALA LYS MET ARG ASP ALA ALA PHE SER ASN SER LEU GLU VAL LEU PHE GLN

• Molecule 1: Transient receptor potential cation channel subfamily M member-like 2

Chain C:

52%                    9%                    40%

MET GLY LYS ASP SER PHE THR PRO LEU TYR ASP GLY GLY ASP SER SER HIS VAL LEU ASN PHE GLY SER ASN GLN LEU SER GLN LYS LYS SER TRP ILE ALA ARG ASN PHE SER ARG ARG GLU CYS ILE ARG PHE VAL PRO LYS SER HIS ASP VAL SER ARG CYS LYS CYS

GLY ARG PRO GLU ARG HIS SER GLN GLN ALA LEU GLU LEU SER GLY SER GLN GLY SER GLU GLU TRP ASN VAL ALA SER CYS THR THR LYS HIS PRO THR ASN A94 E97 P110 V139 V142 T143 P153 R154 L155 R156 T176 G177 G178 T179 N180 T181 M184 E193 Q194

A210 D216 LYS GLN SER ASP LEU ILE SER GLU LYS GLY LYS GLN GLY LYS TYR PRO ALA LEU TYR SER MET GLU PRO THR PRO GLY HIS GLN G243 H250 V256 Y264 G265 I268 G269 M270 R271 I278 V281 THR ASP SER ARG SER GLU ALA GLY SER I292 V296

V297 L298 L299 E302 T307 A321 V322 G328 R329 V346 I351 V369 Y395 S396 L397 E400 L407 A408 I409 L429 R433 I461 K465 Q474 L477 S478 M479 R480 E481 F482 L483 S484 L485 D486 P496 G497 T500 I501 K502 R512

D520 V525 L529 E535 K640 F545 GLY GLU LEU ALA SER TYR ALA GLU GLY LEU VAL LEU LYS ASN ARG LYS SER ASP LYS ASP LEU LEU ALA ASN ILE ARG ILE ASP P575 L576 P577 T578 P579 Y580 L581 L589 R592 R593 I630 D636 D639 R642

E660 T661 L662 S663 Q664 T665 E681 V684 V718 L719 V720 C721 L722 I728 P736 ASP GLU GLN HIS GLY ARG GLN ALA ALA GLU GLU HIS GLN LYS SER LEU ASN GLN LYS SER SER LYS SER GLU GLU LYS GLU ASN ASP PRO VAL VAL PRO VAL TYR

ARG SER LYS GLU GLU LYS ALA VAL SER ASN ASP GLU GLU ALA ARG VAL GLY THR GLU ASN GLU GLU ASP PHE GLN LEU GLU ASP ASP TYR ILE GLU PRO GLN ILE LEU ARG GLU ASP SER MET GLU VAL ILE MET ARG ASN LYS LYS SER LYS GLU LEU GLY F826 K840 N844 V845 V846 G847 Y848

Y855 D867 K870 L882 E892 R895 K900 Y912 D915 T916 W917 N918 F919 A931 L934 R941 I942 I949 F957 L962 N968 L975 T996 I1001 R1016 N1017 N1018 V1021 T1024 S1025 Q1028 Y1029 P1030 L1038 F1039 E1042

F1051 D1057 N1074 F1085 D1097 K1101 R1114 A1118 L1121 V1122 H1126 I1131 ARG VAL TRP ARG MET CYS ARG CYS GLY HIS PRO PRO ARG GLY SER THR MET K1150 Q1167 L1206 ASP SER MET SER GLY THR ARG HIS ALA LEU THR ASP GLY

ASN GLY LEU GLU GLY GLY HIS ASP SER GLU GLY ARG LEU ALA ARG MET GLU VAL VAL GLU LEU ASN SER ASN ASN GLU SER LEU GLN PRO PHE PRO ASP TYR LYS PRO VAL ASN ILE THR ALA LEU LEU GLN GLN PRO VAL LYS GLY LYS GLN ALA LEU VAL PRO ILE LEU THR LEU HIS ALA LYS ALA ARG SER SER

PRO TYR PRO GLY SER THR ALA LYS ARG PHE ALA VAL GLN ASP ASN MET VAL ASP TRP GLN VAL PHE PRO ASP TYR LYS PRO VAL ASN ILE THR ALA PRO VAL VAL LEU ALA ASN PRO VAL TRP ALA ASP LYS ASP LEU MET ALA MET ALA LYS LEU ASN ASN VAL SER GLN GLU GLU LEU PRO GLU LEU TYR ASN GLN

MET ASP HIS THR CYS ASN VAL ASN ARG VAL SER TYR ASN GLY THR VAL LYS ASP GLY LEU PRO LEU ASN PRO MET GLY ARG THR GLY MET GLN ARG GLY LEU LEU ALA ASN PRO PHE GLY PRO ASN HIS ALA ALA ASP PRO VAL THR ARG TRP LYS ARG THR SER ALA GLY

VAL MET LEU GLN GLY GLY LYS LYS VAL LEU GLU PHE VAL GLN ALA ILE GLN ARG LYS ASN ASN GLN TRP ALA ILE PRO GLY GLY MET VAL GLU PRO GLY GLN LEU VAL THR GLN ALA LEU LYS ALA GLU LEU GLY PHE GLY GLU GLU GLY MET ALA LYS LEU ASN VAL SER GLN GLU GLU LYS GLU

ARG ILE ALA LYS GLN ILE GLU ARG LEU PHE GLN GLN GLY GLN GLU GLY ILE LYS TYR LYS VAL VAL ASP PRO ARG ASP ASN THR ASP ASN ALA TRP MET GLU THR VAL ALA VAL ASN PHE HIS ASP ASP LYS GLY LEU PHE GLY ASP ILE THR LEU GLN ALA GLY ASP ASP ALA ALA ALA

VAL ARG TRP GLN ARG VAL SER GLY ASN ILE PRO LEU TYR ALA SER HIS SER ILE LEU GLU LYS VAL ALA LYS MET ARG ASP ALA ALA PHE SER ASN SER LEU GLU VAL LEU PHE GLN

• Molecule 1: Transient receptor potential cation channel subfamily M member-like 2

Chain D:

52%      8%      40%



• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

100%

NAG1
NAG2

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc opyranose

Chain F: 100%

NAG1
NAG2

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc opyranose

Chain G: 100%

NAG1
NAG2

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc opyranose

Chain H: 100%

NAG1
NAG2

# 4 Experimental information ⓘ

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 377471 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | NONE | Depositor |
| Microscope | FEI TITAN KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose $(e^-/\text{Å}^2)$ | 60 | Depositor |
| Minimum defocus (nm) | 1000 | Depositor |
| Maximum defocus (nm) | 2200 | Depositor |
| Magnification | Not provided | |
| Image detector | GATAN K3 (6k x 4k) | Depositor |

# 5   Model quality ⓘ

## 5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CA, A2R

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 | 0.30 | 0/7715 | 0.53 | 0/10457 |
| 1 | B | 0.30 | 0/7715 | 0.53 | 0/10457 |
| 1 | C | 0.30 | 0/7715 | 0.53 | 0/10457 |
| 1 | D | 0.30 | 0/7715 | 0.53 | 0/10457 |
| All | All | 0.30 | 0/30860 | 0.53 | 0/41828 |

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 | 7549 | 0 | 7629 | 80 | 0 |
| 1 | B | 7549 | 0 | 7629 | 82 | 0 |
| 1 | C | 7549 | 0 | 7629 | 80 | 0 |
| 1 | D | 7549 | 0 | 7629 | 80 | 0 |
| 2 | E | 28 | 0 | 25 | 0 | 0 |
| 2 | F | 28 | 0 | 25 | 0 | 0 |
| 2 | G | 28 | 0 | 25 | 0 | 0 |
| 2 | H | 28 | 0 | 25 | 0 | 0 |
| 3 | A | 1 | 0 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 3 | B | 1 | 0 | 0 | 0 | 0 |
| 3 | C | 1 | 0 | 0 | 0 | 0 |
| 3 | D | 1 | 0 | 0 | 0 | 0 |
| 4 | A | 40 | 0 | 20 | 2 | 0 |
| 4 | B | 40 | 0 | 20 | 2 | 0 |
| 4 | C | 40 | 0 | 20 | 2 | 0 |
| 4 | D | 40 | 0 | 20 | 3 | 0 |
| All | All | 30472 | 0 | 30696 | 304 | 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 304 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:D:840:LYS:HD3 | 1:D:1114:ARG:HB2 | 1.74 | 0.69 |
| 1:C:840:LYS:HD3 | 1:C:1114:ARG:HB2 | 1.74 | 0.69 |
| 1:B:840:LYS:HD3 | 1:B:1114:ARG:HB2 | 1.74 | 0.69 |
| 1:A:840:LYS:HD3 | 1:A:1114:ARG:HB2 | 1.74 | 0.68 |
| 1:C:1018:ASN:HB3 | 1:C:1021:VAL:HG22 | 1.82 | 0.62 |

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 PDB entries followed by that with respect to all EM entries.

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 | 929/1560 (60%) | 887 (96%) | 42 (4%) | 0 | 100 | 100 |
| 1 | B | 929/1560 (60%) | 887 (96%) | 42 (4%) | 0 | 100 | 100 |
| 1 | C | 929/1560 (60%) | 888 (96%) | 41 (4%) | 0 | 100 | 100 |
| 1 | D | 929/1560 (60%) | 887 (96%) | 42 (4%) | 0 | 100 | 100 |

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|-----|-------|----------|----------|---------|----------|-------------|
| All | All | 3716/6240 (60%) | 3549 (96%) | 167 (4%) | 0 | 100 100 |

There are no Ramachandran outliers to report.

### 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 PDB entries followed by that with respect to all EM entries.

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 | 818/1346 (61%) | 814 (100%) | 4 (0%) | 86 94 |
| 1 | B | 818/1346 (61%) | 814 (100%) | 4 (0%) | 86 94 |
| 1 | C | 818/1346 (61%) | 814 (100%) | 4 (0%) | 86 94 |
| 1 | D | 818/1346 (61%) | 814 (100%) | 4 (0%) | 86 94 |
| All | All | 3272/5384 (61%) | 3256 (100%) | 16 (0%) | 85 94 |

5 of 16 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | D | 1018 | ASN |
| 1 | D | 968 | ASN |
| 1 | C | 844 | ASN |
| 1 | D | 844 | ASN |
| 1 | B | 1074 | ASN |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | C | 403 | GLN |
| 1 | D | 1160 | GLN |
| 1 | C | 968 | ASN |
| 1 | D | 968 | ASN |
| 1 | C | 664 | GLN |

### 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 ⓘ

8 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 | 0.32 | 0 | 17,19,21 | 0.49 | 0 |
| 2 | NAG | E | 2 | 2 | 14,14,15 | 0.29 | 0 | 17,19,21 | 0.52 | 0 |
| 2 | NAG | F | 1 | 1,2 | 14,14,15 | 0.32 | 0 | 17,19,21 | 0.49 | 0 |
| 2 | NAG | F | 2 | 2 | 14,14,15 | 0.29 | 0 | 17,19,21 | 0.52 | 0 |
| 2 | NAG | G | 1 | 1,2 | 14,14,15 | 0.32 | 0 | 17,19,21 | 0.49 | 0 |
| 2 | NAG | G | 2 | 2 | 14,14,15 | 0.28 | 0 | 17,19,21 | 0.51 | 0 |
| 2 | NAG | H | 1 | 1,2 | 14,14,15 | 0.32 | 0 | 17,19,21 | 0.49 | 0 |
| 2 | NAG | H | 2 | 2 | 14,14,15 | 0.29 | 0 | 17,19,21 | 0.52 | 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 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | E | 2 | 2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | F | 1 | 1,2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | F | 2 | 2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | G | 1 | 1,2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | G | 2 | 2 | - | 2/6/23/26 | 0/1/1/1 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 2 | NAG | H | 1 | 1,2 | - | 2/6/23/26 | 0/1/1/1 |
| 2 | NAG | H | 2 | 2 | - | 2/6/23/26 | 0/1/1/1 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 2 | E | 2 | NAG | O5-C5-C6-O6 |
| 2 | F | 2 | NAG | O5-C5-C6-O6 |
| 2 | G | 2 | NAG | O5-C5-C6-O6 |
| 2 | H | 2 | NAG | O5-C5-C6-O6 |
| 2 | E | 2 | NAG | C4-C5-C6-O6 |

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

## Oligosaccharide Chain G



Bond lengths



Bond angles



Torsions



Rings

**Oligosaccharide Chain H**

Bond lengths

Bond angles

Torsions

Rings
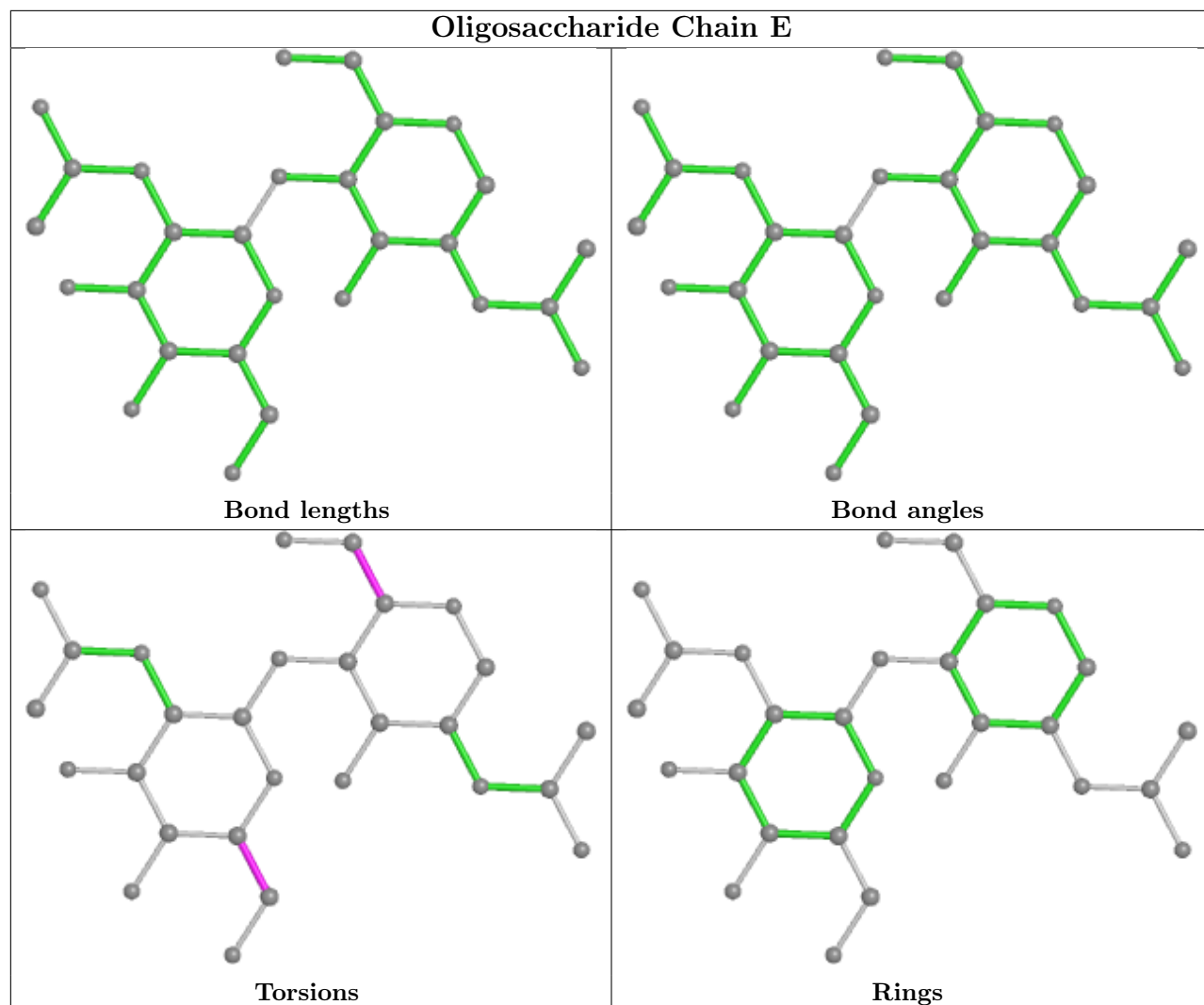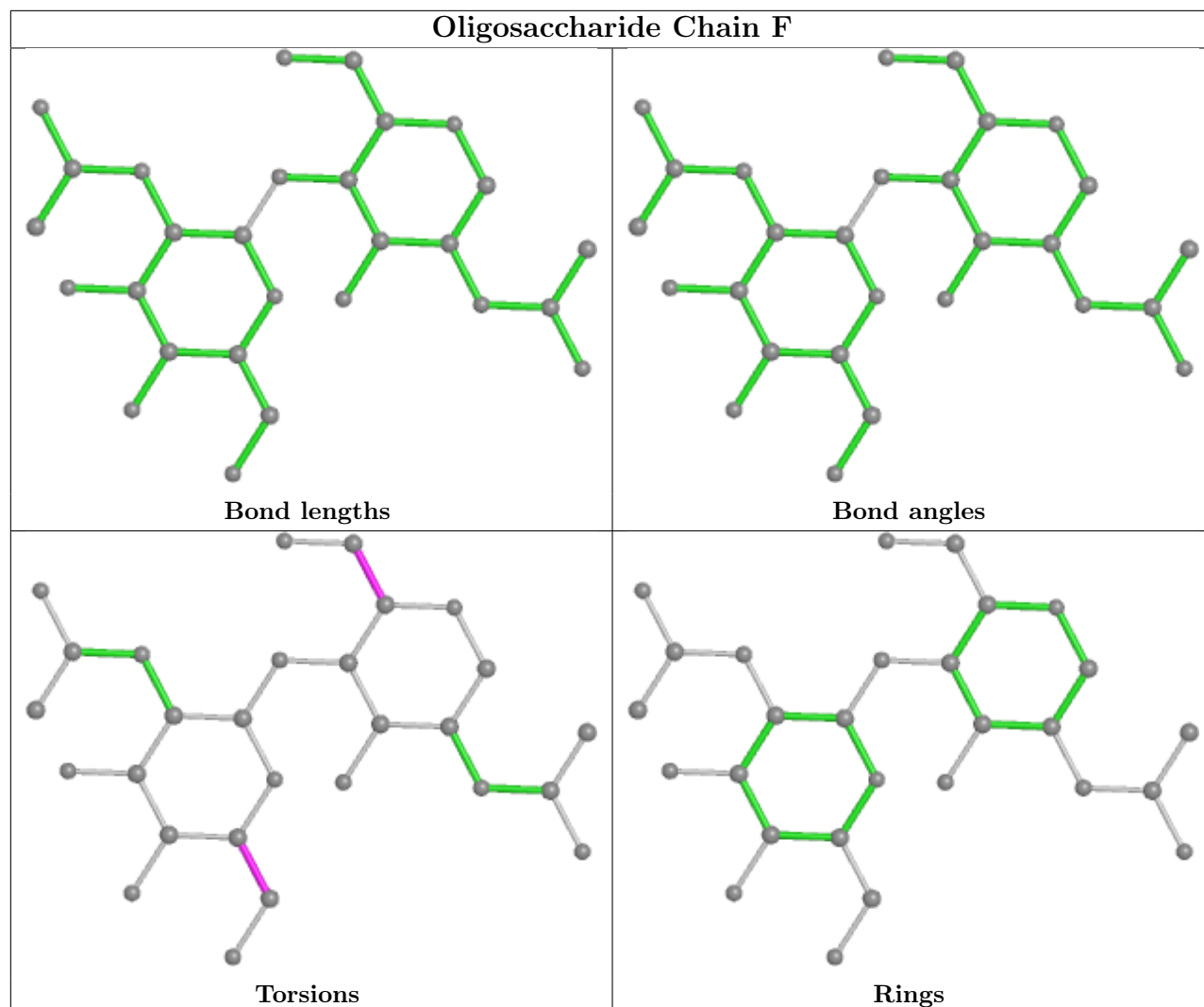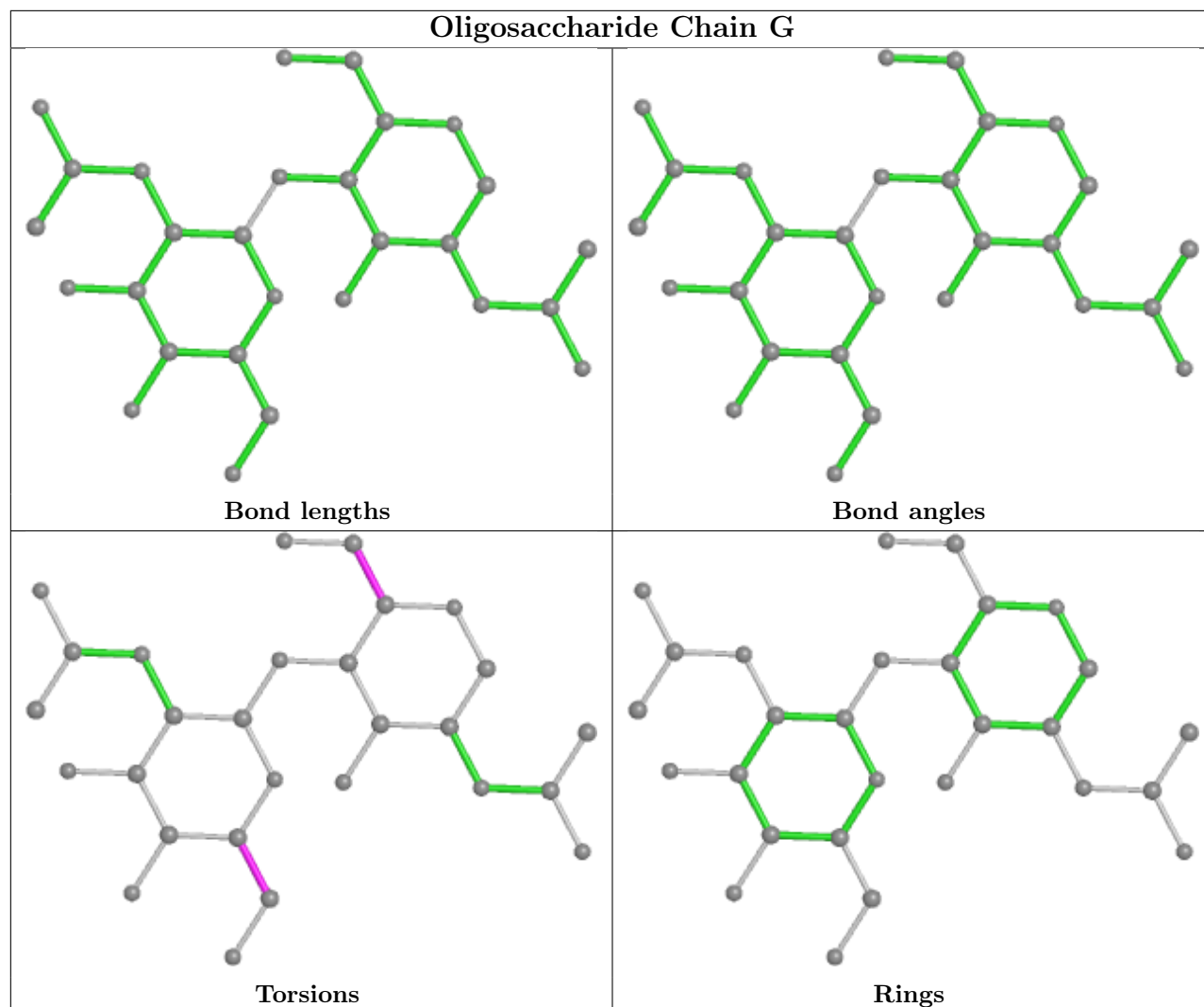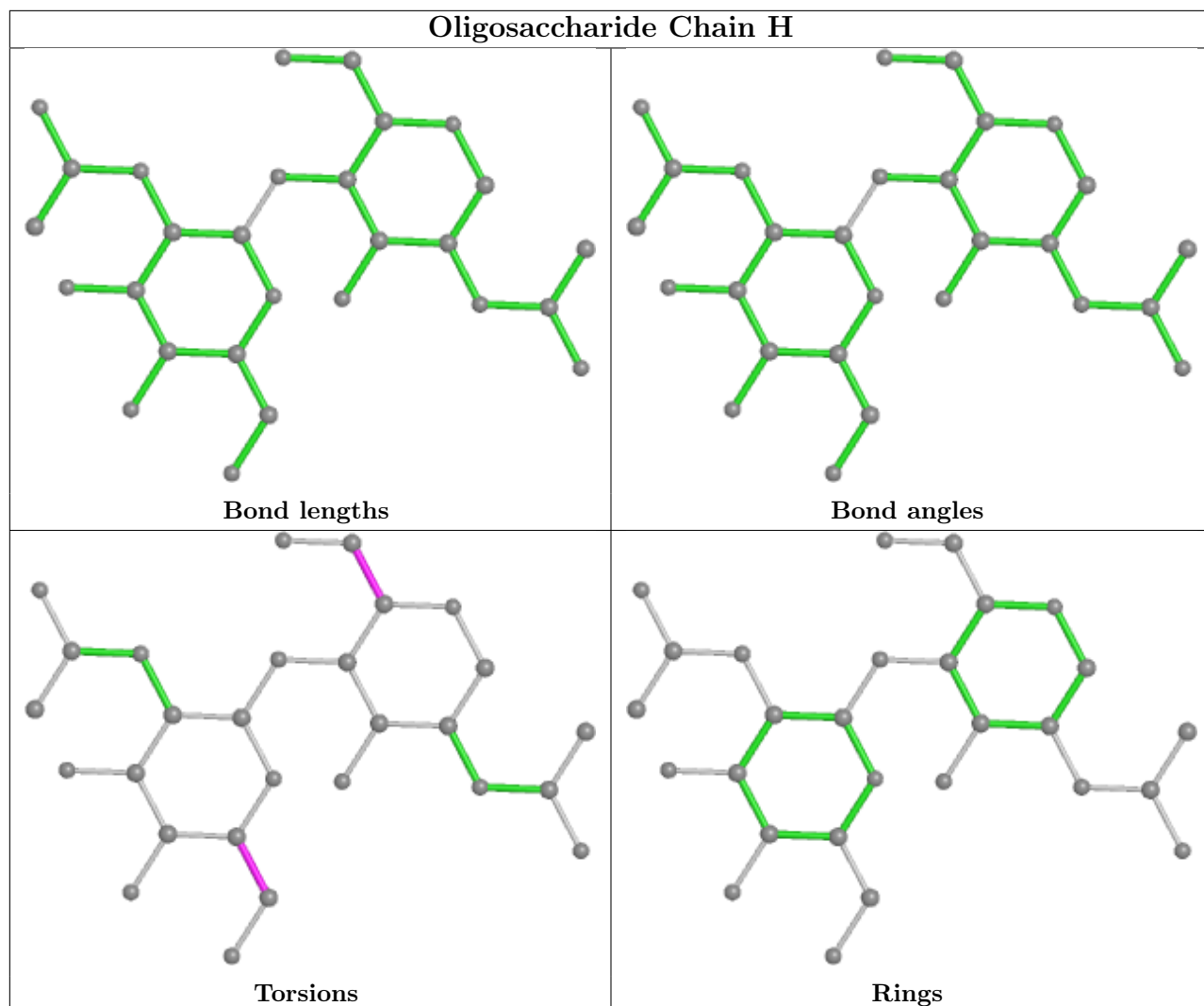
## 5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|-----------|--------|------|-----------|
| | | | | | Counts | RMSZ | $\#|Z| > 2$ | Counts | RMSZ | $\#|Z| > 2$ |
| 4 | A2R | D | 4001 | - | 37,43,43 | 3.72 | 18 (48%) | 46,67,67 | 1.31 | 3 (6%) |
| 4 | A2R | A | 4001 | - | 37,43,43 | 3.72 | 17 (45%) | 46,67,67 | 1.31 | 3 (6%) |
| 4 | A2R | C | 4001 | - | 37,43,43 | 3.72 | 17 (45%) | 46,67,67 | 1.31 | 3 (6%) |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|----------|--------|------|----------|
|     |      |       |     |      | Counts | RMSZ | #\|Z\| > 2 | Counts | RMSZ | #\|Z\| > 2 |
| 4 | A2R | B | 4001 | - | 37,43,43 | 3.72 | 18 (48%) | 46,67,67 | 1.31 | 3 (6%) |

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

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 4 | A2R | D | 4001 | - | - | 10/23/59/59 | 0/4/4/4 |
| 4 | A2R | A | 4001 | - | - | 10/23/59/59 | 0/4/4/4 |
| 4 | A2R | C | 4001 | - | - | 10/23/59/59 | 0/4/4/4 |
| 4 | A2R | B | 4001 | - | - | 10/23/59/59 | 0/4/4/4 |

The worst 5 of 70 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 4 | A | 4001 | A2R | C2-N3 | 9.83 | 1.48 | 1.32 |
| 4 | B | 4001 | A2R | C2-N3 | 9.83 | 1.48 | 1.32 |
| 4 | D | 4001 | A2R | C2-N3 | 9.83 | 1.48 | 1.32 |
| 4 | C | 4001 | A2R | C2-N3 | 9.80 | 1.47 | 1.32 |
| 4 | B | 4001 | A2R | O14'-C11' | -9.34 | 1.31 | 1.43 |

The worst 5 of 12 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|------------|---------|
| 4 | A | 4001 | A2R | N3-C2-N1 | -5.84 | 119.55 | 128.68 |
| 4 | B | 4001 | A2R | N3-C2-N1 | -5.84 | 119.55 | 128.68 |
| 4 | D | 4001 | A2R | N3-C2-N1 | -5.84 | 119.56 | 128.68 |
| 4 | C | 4001 | A2R | N3-C2-N1 | -5.81 | 119.60 | 128.68 |
| 4 | C | 4001 | A2R | C11'-C12'-C13' | 3.33 | 106.47 | 102.30 |

There are no chirality outliers.
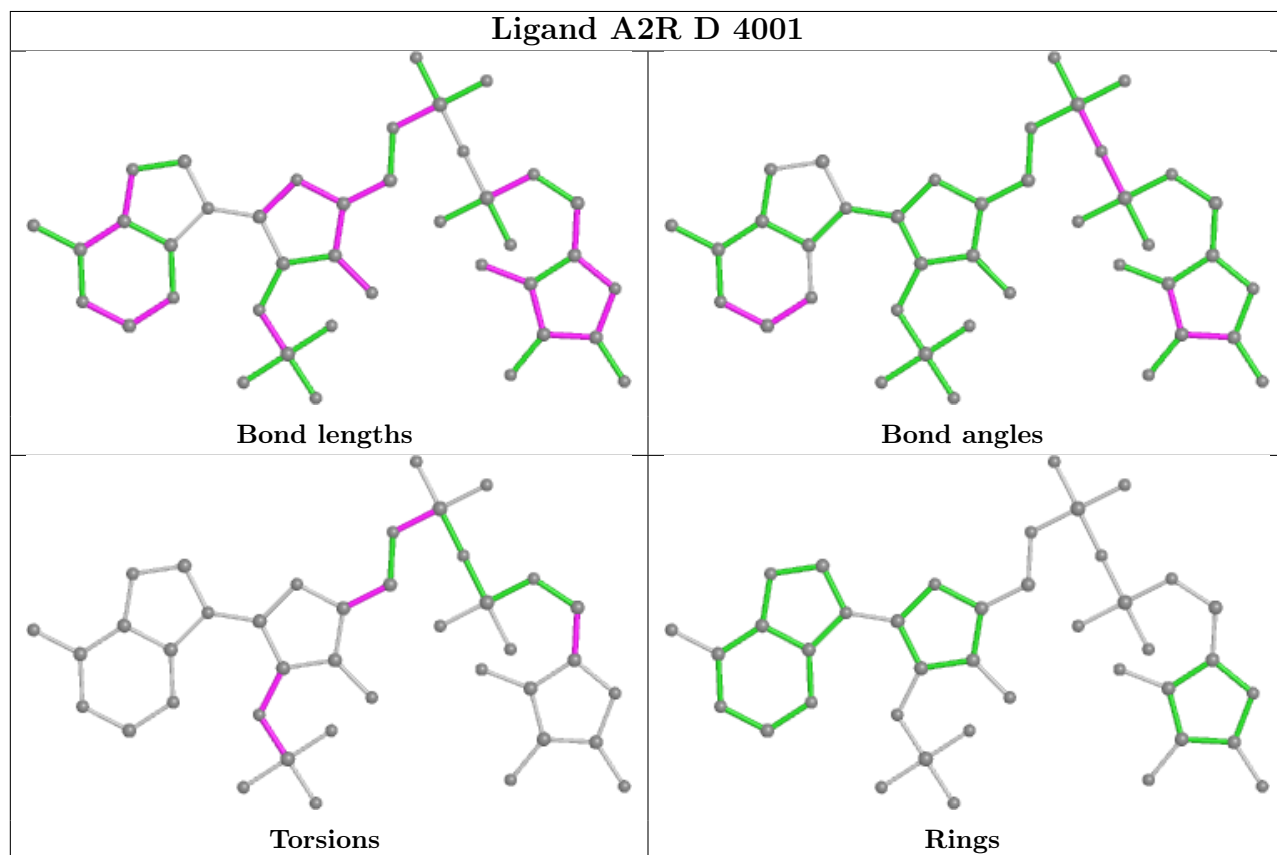
5 of 40 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 4 | A | 4001 | A2R | C2'-O2'-P2'-OP2 |
| 4 | A | 4001 | A2R | O4'-C4'-C5'-O5' |
| 4 | A | 4001 | A2R | C3'-C4'-C5'-O5' |
| 4 | A | 4001 | A2R | C5'-O5'-P-O1 |
| 4 | A | 4001 | A2R | O14'-C14'-C15'-O15' |

There are no ring outliers.

4 monomers are involved in 9 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 4 | D | 4001 | A2R | 3 | 0 |
| 4 | A | 4001 | A2R | 2 | 0 |
| 4 | C | 4001 | A2R | 2 | 0 |
| 4 | B | 4001 | A2R | 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 A2R D 4001

Bond lengths      Bond angles

Torsions      Rings

## Ligand A2R A 4001



Bond lengths



Bond angles



Torsions



Rings

## Ligand A2R C 4001



Bond lengths



Bond angles



Torsions



Rings

**Ligand A2R B 4001**



Bond lengths

Bond angles

Torsions

Rings
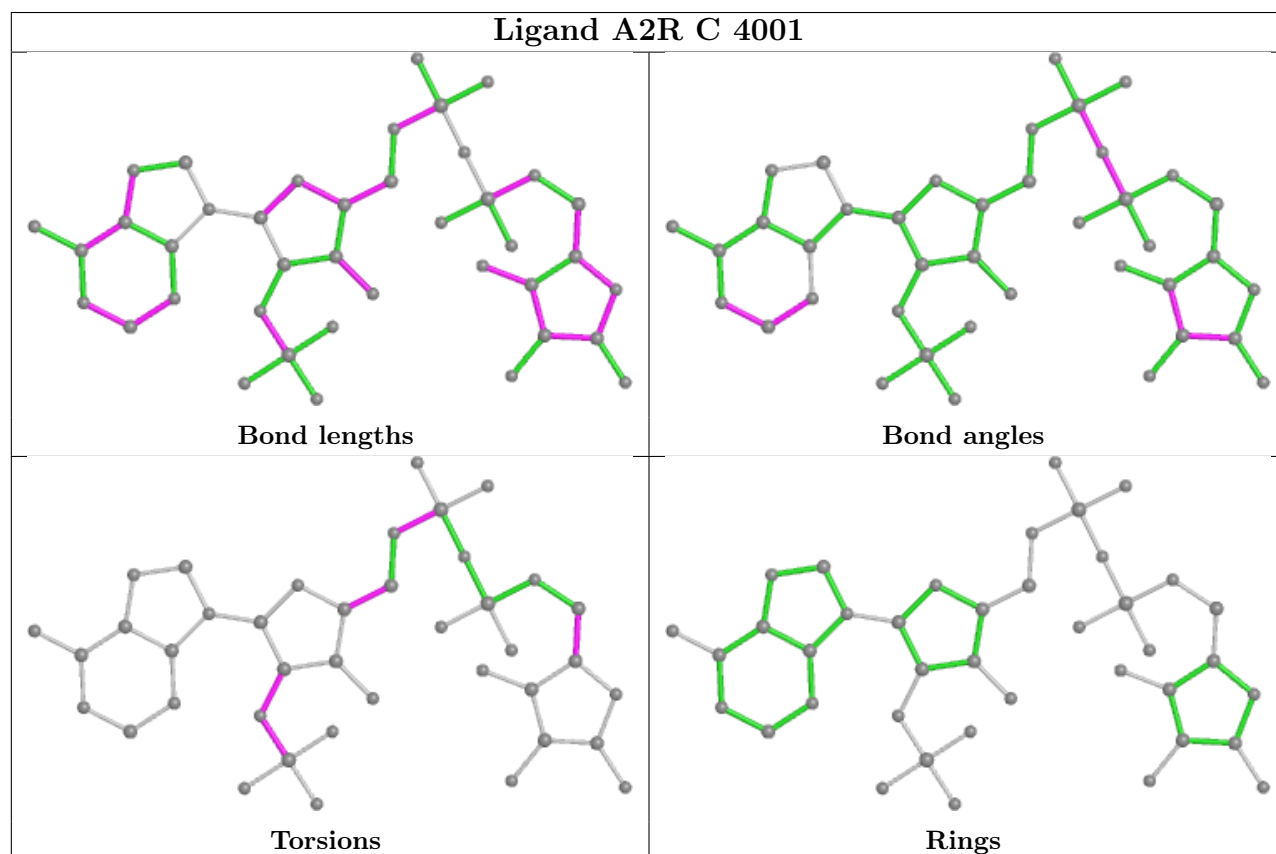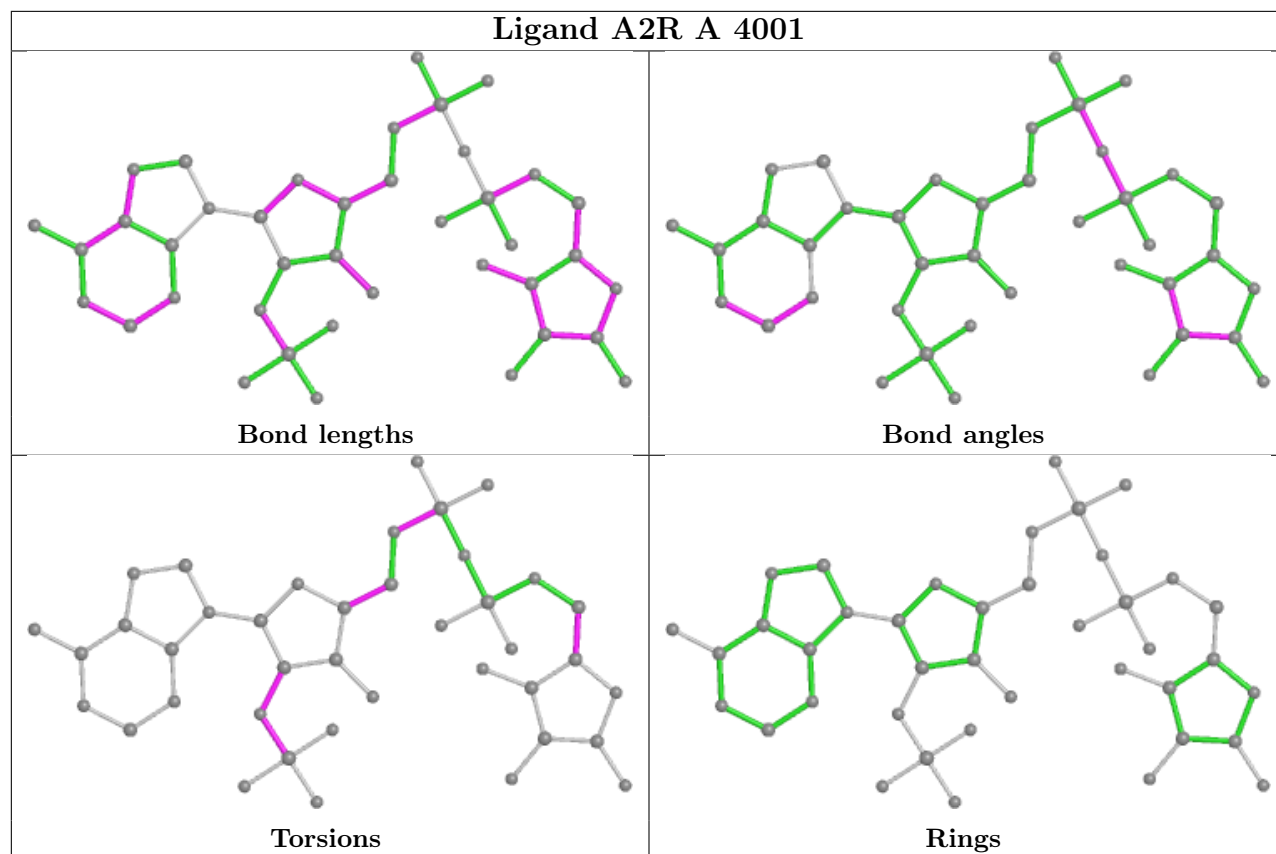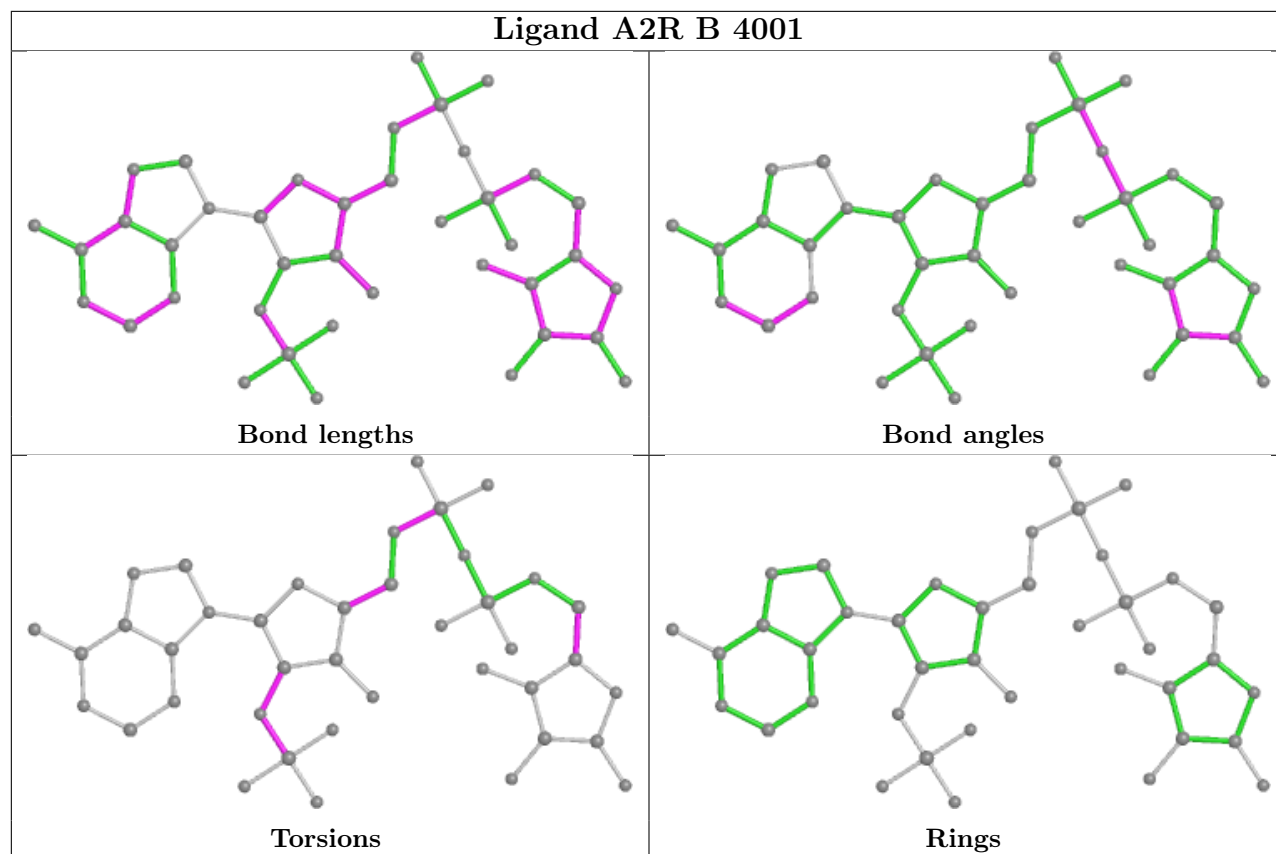
## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.