WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Jun 10, 2024 – 05:09 PM EDT

PDB ID	:	8UKV
Title	:	Crystal structure of nanobody/VHH domain of 34E5 in complex with the extracellular region of the epidermal growth factor variant III (EGFRvIII)
Authors	:	Stayrook, S.E.; Ferguson, K.M.; Bagchi, A.
Deposited on	:	2023-10-15
Resolution	:	2.94 Å(reported)

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

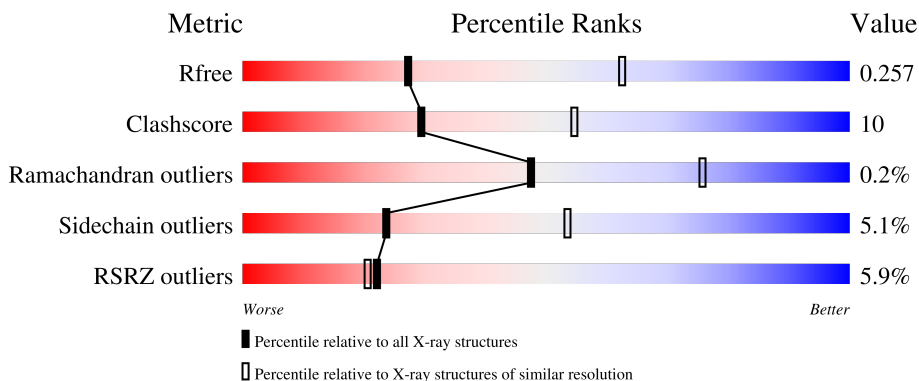|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.94 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		0.257
Clashscore		10
Ramachandran outliers		0.2%
Sidechain outliers		5.1%
RSRZ outliers		5.9%

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$_{free}$	130704	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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	357	11% 67% / 20% · 10%
1	B	357	2% 73% / 14% · 13%
2	C	141	% 60% / 23% · 16%
2	D	141	3% 52% / 30% · 16%
3	E	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	2	100%
3	G	2	100%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6739 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 Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total 2446	C 1511	N 438	O 467	S 30	0	0	0
1	B	310	Total 2352	C 1453	N 424	O 447	S 28	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	273	GLY	VAL	engineered mutation	UNP P00533
A	619	HIS	-	expression tag	UNP P00533
A	620	HIS	-	expression tag	UNP P00533
A	621	HIS	-	expression tag	UNP P00533
A	622	HIS	-	expression tag	UNP P00533
A	623	HIS	-	expression tag	UNP P00533
A	624	HIS	-	expression tag	UNP P00533
B	273	GLY	VAL	engineered mutation	UNP P00533
B	619	HIS	-	expression tag	UNP P00533
B	620	HIS	-	expression tag	UNP P00533
B	621	HIS	-	expression tag	UNP P00533
B	622	HIS	-	expression tag	UNP P00533
B	623	HIS	-	expression tag	UNP P00533
B	624	HIS	-	expression tag	UNP P00533

- Molecule 2 is a protein called Nanobody/VHH domain 34E5.

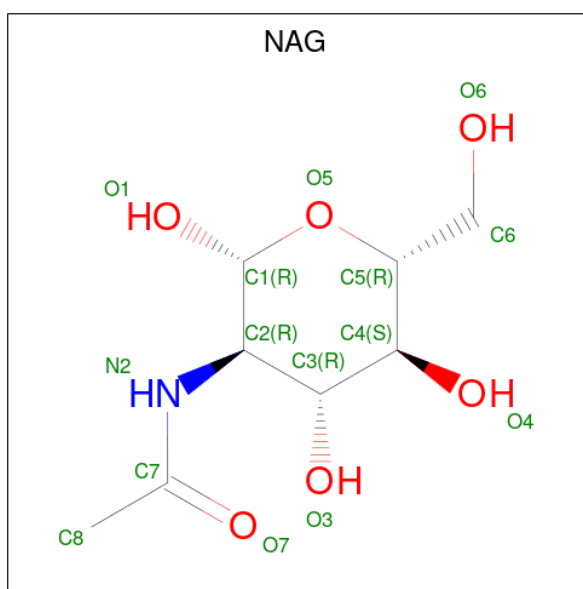| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	119	Total 924	C 574	N 163	O 182	S 5	0	0	0
2	D	118	Total 919	C 571	N 162	O 181	S 5	0	0	0

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a

cetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

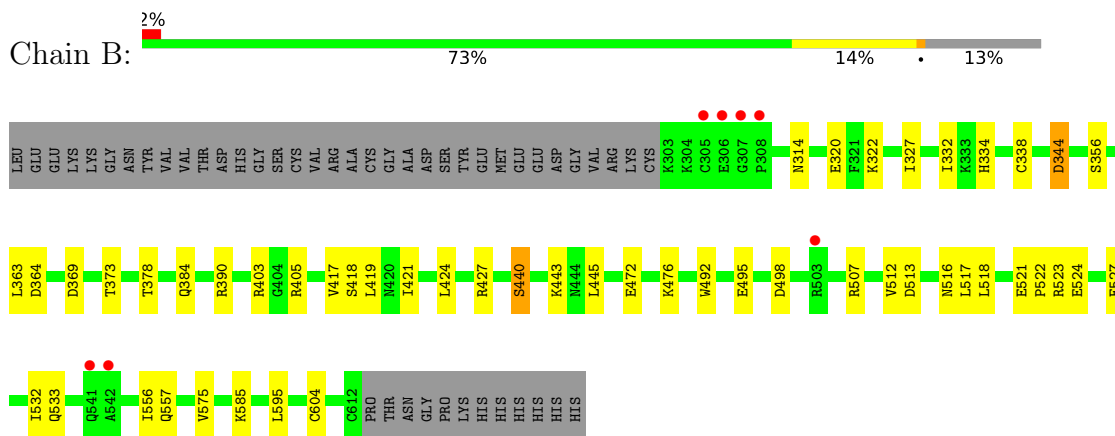| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

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: Epidermal growth factor receptor

Chain A:

• Molecule 1: Epidermal growth factor receptor

Chain B:

• Molecule 2: Nanobody/VHH domain 34E5

Chain C:

E1 V2 Q3 L4 S7 L11 S17 A24 N32 V33 R38 P41 G42 N43 Q44 R45 A49 T52 T53 D54 N58 Y59 A60 D61 S62 V63 I69 Y77 M80 L83 K84 P85 E86 C93 H94 L95 K96 N97 I98 R99 D100 Y101 Q113 V114

T115 A119 ALA ALA GLU GLN LYS LEU ILE SER GLU GLU ASP LEU ASN GLY LEU GLU HIS HIS HIS HIS HIS

• Molecule 2: Nanobody/VHH domain 34E5

Chain D:

3%

52% 30% • 16%

E1 V2 E6 S7 G8 L11 V12 L20 S25 I28 F29 S30 F31 N32 V33 W36 P41 G42 N43 Q44 R45 T52 T53 D61 R66 F67 T68 I69 S70 R71 V76 Y77 L78 Q79 M80 L83 K84 P85 E86 D87 V90 Y91 Y92 C93 H94 L95 K96

N97 I98 R99 D100 Y101 R102 N103 T112 Q113 V114 T115 V116 S117 S118 ALA ALA ALA GLU GLN LYS LEU ILE SER GLU GLU ASP LEU ASN GLY LEU GLU HIS HIS HIS HIS HIS

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

100%

NAG1 NAG2

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

100%

NAG1 NAG2

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

100%

NAG1 NAG2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	66.29Å　　80.36Å　　91.84Å 106.92°　　103.50°　　104.56°	Depositor
Resolution (Å)	46.55　–　2.94 46.58　–　2.94	Depositor EDS
% Data completeness (in resolution range)	98.2 (46.55-2.94) 98.2 (46.58-2.94)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.92 (at 2.96Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.215　,　0.252 0.227　,　0.257	Depositor DCC
R_{free} test set	1682 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	76.2	Xtriage
Anisotropy	0.504	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 38.6	EDS
L-test for twinning[2]	$<\|L\|>\ =\ 0.47,\ <L^2>\ =\ 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6739	wwPDB-VP
Average B, all atoms (Å^2)	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.66% 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 (i)

5.1 Standard geometry (i)

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

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	0.33	0/2496	0.60	0/3377
1	B	0.31	0/2401	0.59	0/3253
2	C	0.36	0/942	0.64	0/1276
2	D	0.34	0/937	0.66	0/1269
All	All	0.33	0/6776	0.61	0/9175

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	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	578	GLU	Peptide
1	B	523	ARG	Sidechain

5.2 Too-close contacts (i)

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	2446	0	2358	52	0
1	B	2352	0	2262	30	0
2	C	924	0	889	20	0
2	D	919	0	886	27	0
3	E	28	0	25	0	0
3	F	28	0	25	0	0
3	G	28	0	25	0	0
4	A	14	0	13	1	0
All	All	6739	0	6483	126	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 10.

All (126) 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)
2:D:99:ARG:HD2	2:D:103:ASN:HB3	1.28	1.08
1:A:484:ALA:O	1:A:503:ARG:NH1	2.06	0.88
1:B:521:GLU:HB3	1:B:522:PRO:HD3	1.64	0.79
2:D:28:ILE:HA	2:D:97:ASN:HD21	1.47	0.78
1:A:521:GLU:HB3	1:A:522:PRO:HD3	1.67	0.77
2:D:66:ARG:NH2	2:D:87:ASP:OD2	2.18	0.77
1:B:522:PRO:HB2	1:B:533:GLN:HG2	1.69	0.73
1:B:320:GLU:OE1	1:B:334:HIS:ND1	2.23	0.69
1:A:568:VAL:HG12	1:A:570:THR:H	1.58	0.69
1:B:338:CYS:O	1:B:373:THR:OG1	2.14	0.66
1:A:436:ASP:OD1	1:A:463:LYS:N	2.25	0.66
1:A:478:THR:HG23	1:A:480:GLN:HG3	1.76	0.65
1:A:512:VAL:HG12	1:A:513:ASP:H	1.61	0.65
2:C:2:VAL:HG11	2:C:95:LEU:HD13	1.78	0.65
2:D:2:VAL:HG11	2:D:95:LEU:HD13	1.78	0.64
1:A:576:MET:HG3	1:A:580:ASN:HB3	1.80	0.63
2:D:11:LEU:HD22	2:D:118:SER:HB3	1.80	0.63
1:A:407:LYS:NZ	1:A:434:ASP:OD2	2.30	0.62
2:D:8:GLY:O	2:D:112:THR:HG21	2.01	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:397:GLU:OE1	1:A:427:ARG:NH2	2.35	0.60
1:A:572:PRO:HD2	1:A:585:LYS:HD3	1.83	0.60
2:D:11:LEU:HD23	2:D:115:THR:HB	1.85	0.59
1:A:472:GLU:H	1:A:472:GLU:CD	2.06	0.58
1:B:585:LYS:HE2	1:B:595:LEU:HD23	1.84	0.58
2:C:113:GLN:HE21	2:C:113:GLN:HA	1.68	0.57
2:C:52:THR:HG23	2:C:54:ASP:H	1.68	0.57
2:D:6:GLU:HB3	2:D:93:CYS:SG	2.44	0.57
1:B:421:ILE:HG13	1:B:445:LEU:HD13	1.88	0.56
1:A:519:GLU:HA	1:A:523:ARG:HH21	1.70	0.56
1:A:384:GLN:HG2	1:A:417:VAL:HG23	1.87	0.56
2:D:33:VAL:HG22	2:D:52:THR:HG22	1.88	0.55
1:A:337:ASN:HD22	4:A:701:NAG:H83	1.71	0.55
1:A:584:TRP:CE3	1:A:603:GLY:HA2	2.42	0.55
1:A:519:GLU:N	1:A:519:GLU:OE2	2.40	0.54
1:B:344:ASP:N	1:B:344:ASP:OD1	2.40	0.53
1:A:457:PHE:CD2	1:A:462:GLN:HB3	2.44	0.53
1:A:487:SER:HB2	1:A:488:PRO:HD2	1.91	0.52
1:B:527:GLU:HB3	1:B:532:ILE:HD13	1.92	0.52
2:D:67:PHE:CE1	2:D:80:MET:HG2	2.45	0.52
1:A:515:CYS:SG	1:A:526:VAL:HG12	2.50	0.52
1:B:403:ARG:HD3	1:B:405:ARG:NH2	2.25	0.52
1:B:513:ASP:OD1	1:B:513:ASP:N	2.36	0.51
2:C:32:ASN:O	2:C:53:THR:HG23	2.11	0.51
1:A:405:ARG:NH1	2:D:41:PRO:HB3	2.26	0.51
2:C:52:THR:HG23	2:C:54:ASP:N	2.26	0.49
1:B:378:THR:HA	1:B:403:ARG:HB2	1.93	0.49
1:A:444:ASN:N	1:A:444:ASN:OD1	2.46	0.49
1:B:557:GLN:HB2	2:D:102:ARG:HA	1.94	0.49
2:C:33:VAL:HG13	2:C:96:LYS:HB3	1.95	0.49
2:D:69:ILE:HD12	2:D:77:TYR:O	2.13	0.49
1:B:517:LEU:HB2	1:B:518:LEU:HD22	1.95	0.48
1:A:537:GLU:OE1	1:A:569:LYS:HG3	2.13	0.48
1:A:518:LEU:HB2	1:A:519:GLU:OE2	2.13	0.48
2:D:85:PRO:HA	2:D:116:VAL:CG2	2.44	0.48
1:B:527:GLU:HB3	1:B:532:ILE:CD1	2.44	0.48
1:B:556:ILE:HD12	1:B:557:GLN:N	2.29	0.47
1:A:585:LYS:HG2	1:A:595:LEU:HA	1.94	0.47
2:C:80:MET:HB3	2:C:83:LEU:HD21	1.95	0.47
1:B:369:ASP:OD1	1:B:390:ARG:NH2	2.45	0.47
1:B:427:ARG:HA	1:B:492:TRP:CD1	2.49	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:29:PHE:N	2:D:29:PHE:CD1	2.82	0.47
2:D:30:SER:O	2:D:53:THR:HG21	2.15	0.47
1:B:363:LEU:HD12	1:B:364:ASP:N	2.29	0.47
1:B:417:VAL:HA	1:B:440:SER:O	2.15	0.47
2:D:67:PHE:CZ	2:D:80:MET:HG2	2.50	0.47
1:A:578:GLU:HA	1:A:580:ASN:ND2	2.30	0.46
2:C:43:ASN:HD22	2:C:43:ASN:N	2.13	0.46
2:C:60:ALA:HB3	2:C:63:VAL:HG22	1.97	0.46
1:A:317:GLY:O	1:A:318:ILE:HD12	2.16	0.46
1:A:576:MET:CG	1:A:580:ASN:HB3	2.44	0.46
1:B:472:GLU:O	1:B:476:LYS:HG2	2.16	0.46
1:B:314:ASN:HD21	2:C:11:LEU:H	1.62	0.46
1:A:568:VAL:HG12	1:A:569:LYS:N	2.31	0.46
1:B:424:LEU:HD23	1:B:424:LEU:HA	1.81	0.46
2:C:33:VAL:CG1	2:C:96:LYS:HB3	2.45	0.46
1:A:364:ASP:OD1	1:A:365:PRO:HD2	2.15	0.45
2:C:11:LEU:HA	2:C:115:THR:O	2.16	0.45
1:A:422:THR:HA	1:A:444:ASN:O	2.16	0.45
2:D:12:VAL:O	2:D:116:VAL:HA	2.16	0.45
1:A:523:ARG:HG2	1:A:523:ARG:HH11	1.81	0.45
1:A:474:SER:O	1:A:478:THR:HG22	2.17	0.45
1:A:584:TRP:CD2	1:A:603:GLY:HA2	2.52	0.45
1:A:295:GLU:HA	1:A:300:ARG:HA	1.99	0.44
1:A:515:CYS:O	1:A:524:GLU:HG3	2.16	0.44
2:C:41:PRO:HA	2:C:42:GLY:HA2	1.72	0.44
1:A:528:ASN:O	1:A:530:GLU:HG3	2.17	0.44
2:D:36:TRP:CZ3	2:D:76:VAL:HG23	2.53	0.44
1:A:588:ASP:OD1	1:A:590:GLY:N	2.51	0.44
2:D:43:ASN:N	2:D:43:ASN:OD1	2.50	0.44
1:A:376:GLU:OE1	1:A:403:ARG:NE	2.38	0.44
1:B:384:GLN:HG2	1:B:417:VAL:HG22	2.00	0.43
2:C:61:ASP:N	2:C:61:ASP:OD1	2.49	0.43
1:A:519:GLU:HA	1:A:523:ARG:NH2	2.33	0.43
1:B:327:ILE:CG2	1:B:332:ILE:HD12	2.48	0.43
2:C:4:LEU:HD21	2:C:95:LEU:HB2	2.00	0.43
1:A:597:HIS:ND1	1:A:598:PRO:HD2	2.34	0.43
1:B:516:ASN:HD22	1:B:522:PRO:HD2	1.83	0.43
2:D:99:ARG:HH11	2:D:103:ASN:C	2.22	0.43
1:A:588:ASP:OD2	1:A:594:HIS:HE1	2.01	0.43
1:A:523:ARG:HD2	1:A:546:THR:OG1	2.19	0.43
1:A:365:PRO:HB3	1:A:387:PRO:CG	2.49	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:495:GLU:OE2	1:A:497:ARG:NH1	2.52	0.42
1:A:558:CYS:HB3	1:A:567:CYS:HB3	1.87	0.42
2:C:84:LYS:HB3	2:C:86:GLU:OE2	2.19	0.42
2:D:90:VAL:HG12	2:D:92:TYR:CE2	2.54	0.42
2:C:93:CYS:SG	2:C:94:HIS:N	2.92	0.42
1:A:413:SER:HB3	1:A:435:GLY:HA3	2.02	0.42
1:A:516:ASN:O	1:A:523:ARG:HB2	2.19	0.42
1:B:507:ARG:NH1	1:B:524:GLU:OE2	2.52	0.42
1:B:495:GLU:HB2	1:B:498:ASP:OD2	2.19	0.42
1:B:585:LYS:HE2	1:B:595:LEU:CD2	2.50	0.42
2:D:32:ASN:OD1	2:D:71:ARG:NH2	2.37	0.42
2:D:61:ASP:OD1	2:D:61:ASP:N	2.53	0.42
1:A:512:VAL:HG12	1:A:513:ASP:OD1	2.20	0.42
1:A:524:GLU:HB3	1:A:533:GLN:HA	2.02	0.42
1:A:344:ASP:N	1:A:344:ASP:OD1	2.53	0.41
1:A:358:THR:OG1	1:A:360:THR:HG23	2.20	0.41
1:B:512:VAL:HG12	1:B:513:ASP:N	2.35	0.41
2:D:69:ILE:HD12	2:D:69:ILE:HA	1.89	0.41
1:B:418:SER:C	1:B:419:LEU:HD12	2.41	0.41
2:D:20:LEU:HB2	2:D:78:LEU:HB3	2.01	0.41
1:A:421:ILE:HG13	1:A:445:LEU:HD13	2.02	0.41
2:C:4:LEU:HD23	2:C:24:ALA:HB2	2.02	0.41
2:C:69:ILE:HD12	2:C:77:TYR:O	2.20	0.41
2:C:49:ALA:HA	2:C:58:ASN:O	2.21	0.40
2:D:12:VAL:CG1	2:D:116:VAL:HG12	2.52	0.40

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	319/357 (89%)	288 (90%)	30 (9%)	1 (0%)	41	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	308/357 (86%)	274 (89%)	33 (11%)	1 (0%)	41	69
2	C	117/141 (83%)	111 (95%)	6 (5%)	0	100	100
2	D	116/141 (82%)	108 (93%)	8 (7%)	0	100	100
All	All	860/996 (86%)	781 (91%)	77 (9%)	2 (0%)	47	76

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	443	LYS
1	A	522	PRO

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	276/307 (90%)	259 (94%)	17 (6%)	18	45
1	B	265/307 (86%)	259 (98%)	6 (2%)	50	78
2	C	100/119 (84%)	92 (92%)	8 (8%)	12	32
2	D	100/119 (84%)	93 (93%)	7 (7%)	15	39
All	All	741/852 (87%)	703 (95%)	38 (5%)	24	53

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	301	LYS
1	A	304	LYS
1	A	306	GLU
1	A	309	CYS
1	A	334	HIS
1	A	344	ASP
1	A	394	HIS
1	A	503	ARG
1	A	515	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	523	ARG
1	A	537	GLU
1	A	548	THR
1	A	576	MET
1	A	578	GLU
1	A	580	ASN
1	A	596	CYS
1	A	597	HIS
1	B	322	LYS
1	B	344	ASP
1	B	356	SER
1	B	440	SER
1	B	575	VAL
1	B	604	CYS
2	C	3	GLN
2	C	7	SER
2	C	17	SER
2	C	38	ARG
2	C	45	ARG
2	C	61	ASP
2	C	101	TYR
2	C	113	GLN
2	D	7	SER
2	D	25	SER
2	D	43	ASN
2	D	45	ARG
2	D	93	CYS
2	D	101	TYR
2	D	113	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	566	HIS
1	A	580	ASN
1	A	594	HIS
1	B	346	HIS
1	B	516	ASN
2	C	43	ASN
2	C	113	GLN
2	D	39	GLN
2	D	97	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	113	GLN

5.3.3   RNA (i)

There are no RNA molecules in this entry.

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

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

5.5   Carbohydrates (i)

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	$\#|Z| > 2$
3	NAG	E	1	3,1	14,14,15	0.49	0	17,19,21	0.63	0
3	NAG	E	2	3	14,14,15	0.20	0	17,19,21	0.53	0
3	NAG	F	1	3,1	14,14,15	0.54	0	17,19,21	0.62	0
3	NAG	F	2	3	14,14,15	0.27	0	17,19,21	0.47	0
3	NAG	G	1	3,1	14,14,15	0.35	0	17,19,21	0.59	0
3	NAG	G	2	3	14,14,15	0.20	0	17,19,21	0.44	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
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	G	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

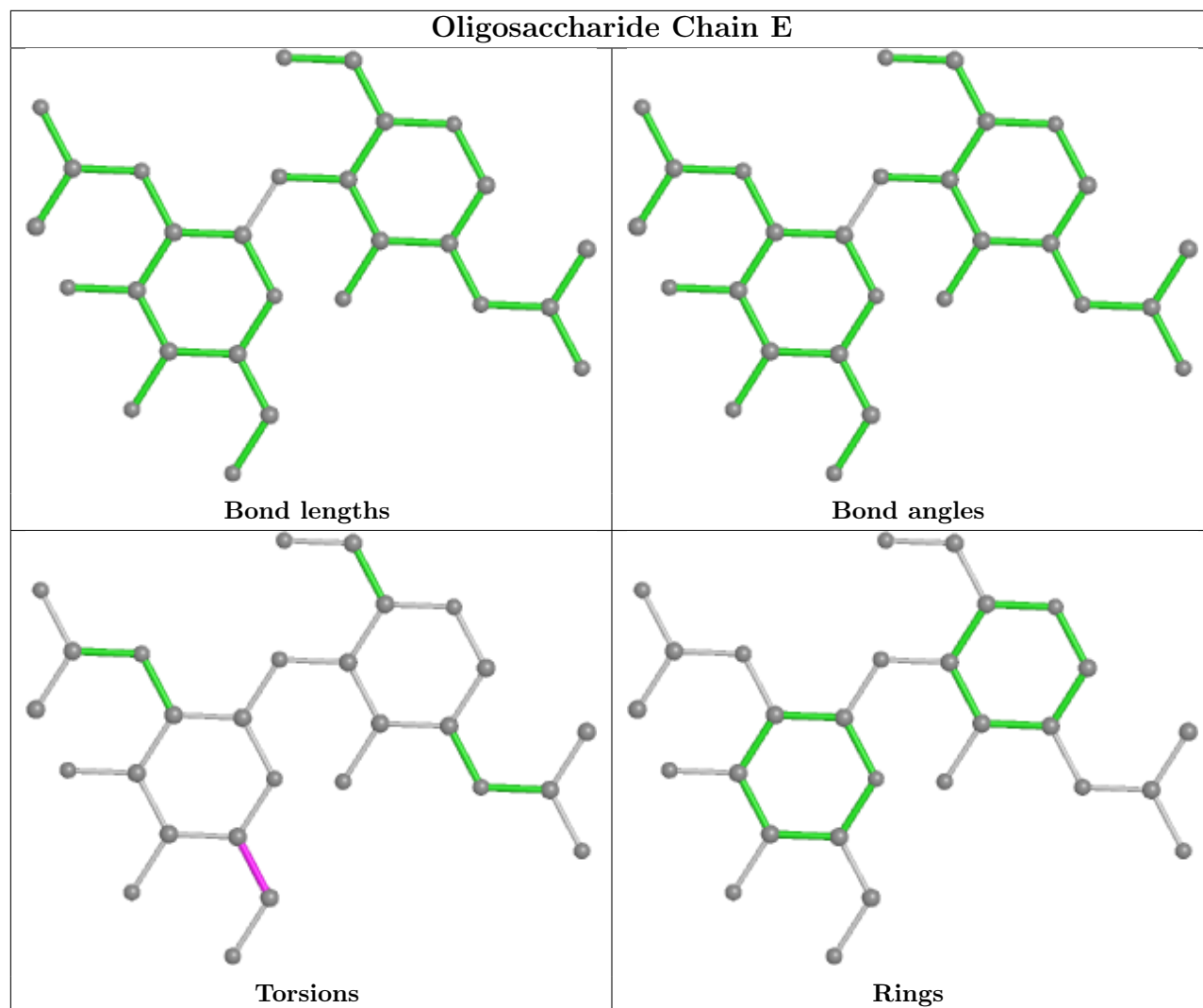There are no bond angle outliers.

There are no chirality outliers.

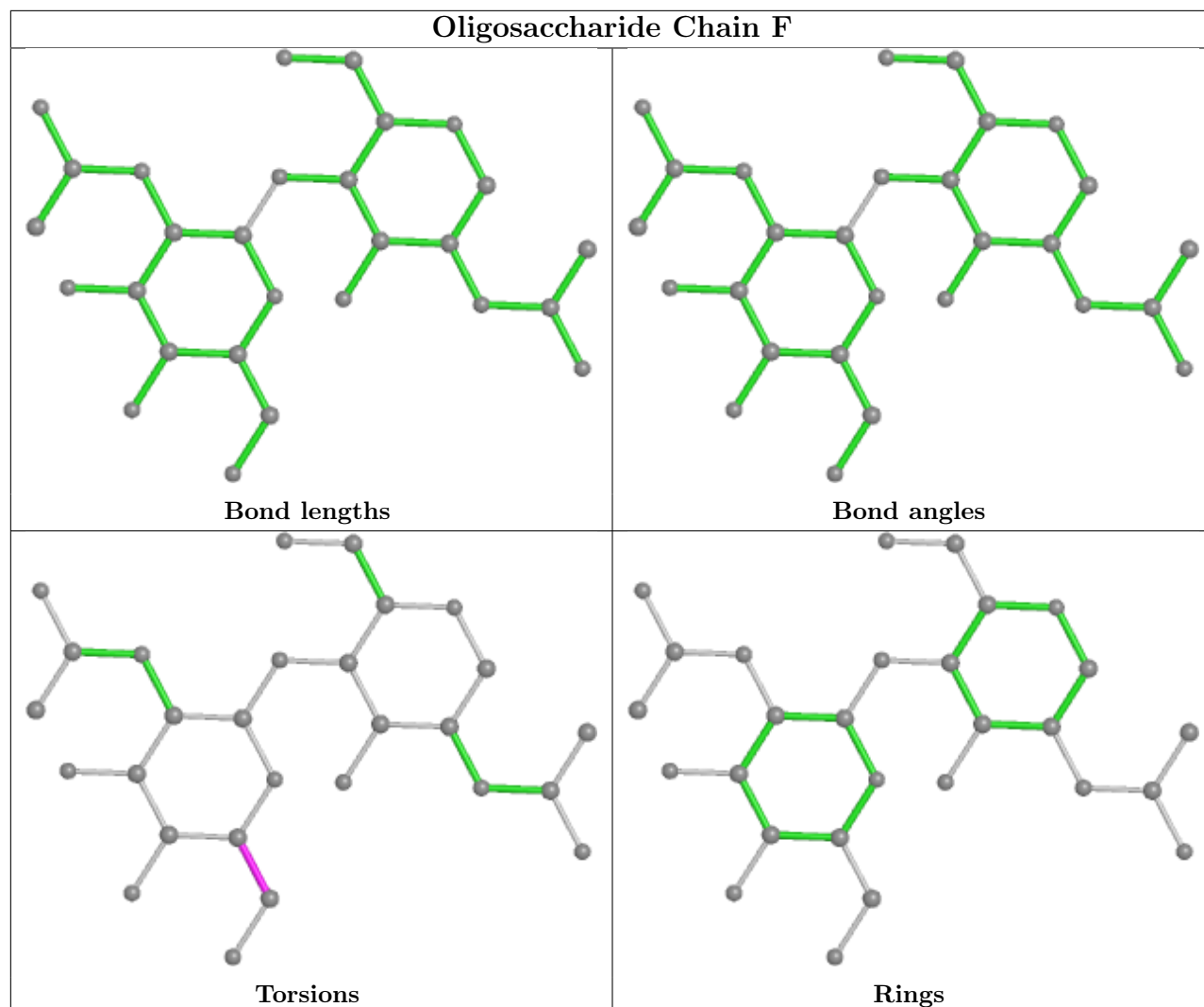All (10) torsion outliers are listed below:

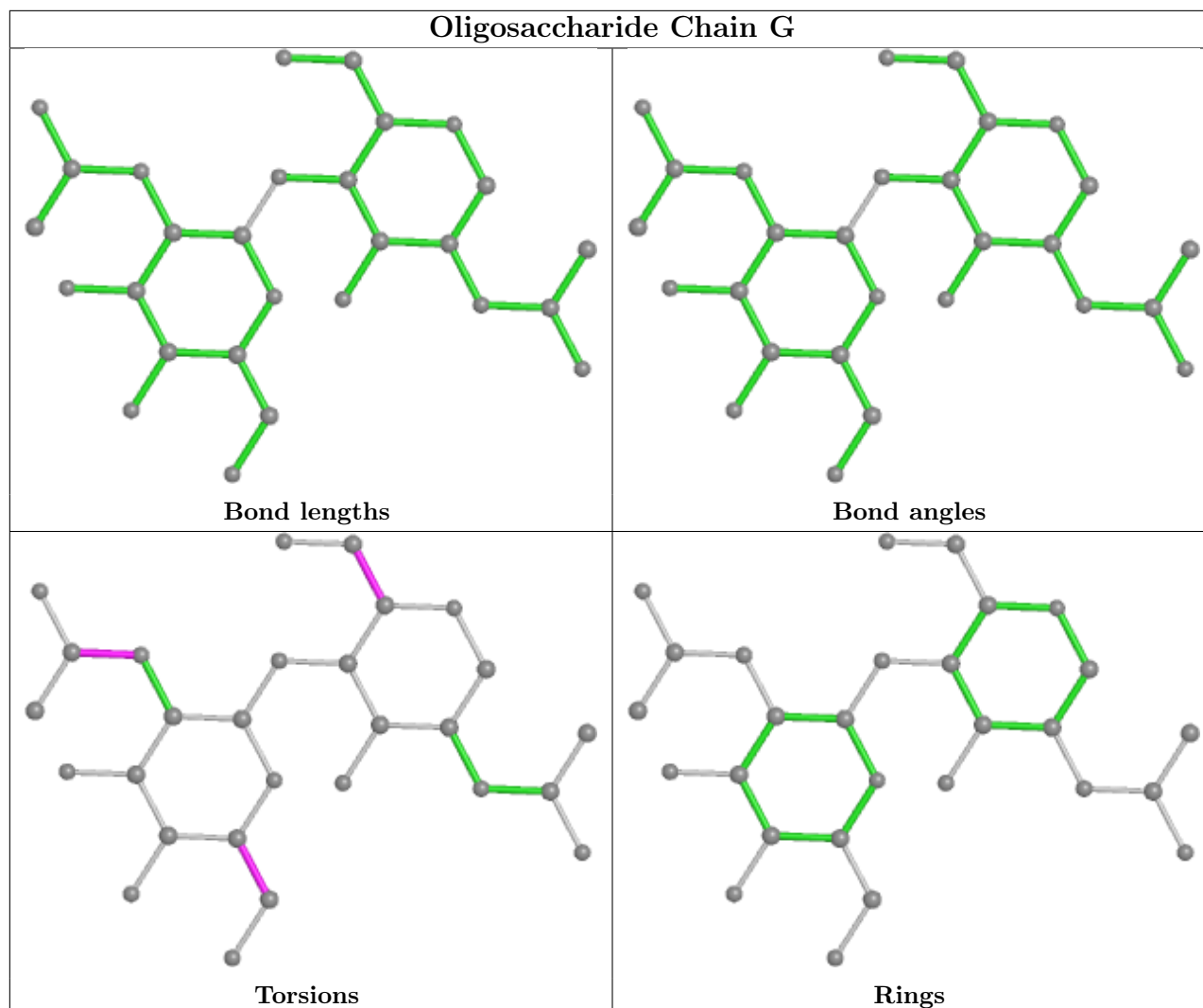| Mol	Chain	Res	Type	Atoms
3	G	2	NAG	O5-C5-C6-O6
3	G	1	NAG	O5-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6
3	G	2	NAG	C8-C7-N2-C2
3	G	2	NAG	O7-C7-N2-C2
3	E	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6
3	F	2	NAG	O5-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6

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

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).

5.6 Ligand geometry (i)

1 ligand is 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	$\#|Z| > 2$
4	NAG	A	701	1	14,14,15	0.29	0	17,19,21	0.57	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
4	NAG	A	701	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	701	NAG	O5-C5-C6-O6
4	A	701	NAG	C8-C7-N2-C2
4	A	701	NAG	O7-C7-N2-C2
4	A	701	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	701	NAG	1	0

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	321/357 (89%)	0.59	38 (11%)	4	3	53, 95, 180, 219	0
1	B	310/357 (86%)	0.18	7 (2%)	60	61	47, 82, 137, 190	0
2	C	119/141 (84%)	0.36	2 (1%)	70	71	57, 79, 117, 176	0
2	D	118/141 (83%)	0.29	4 (3%)	45	43	55, 86, 116, 155	0
All	All	868/996 (87%)	0.37	51 (5%)	22	20	47, 87, 161, 219	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	609	LEU	9.6
2	D	1	GLU	6.8
1	A	612	CYS	6.7
1	A	292	TYR	6.3
1	A	294	MET	6.0
1	A	307	GLY	5.5
1	A	305	CYS	5.4
1	B	305	CYS	5.2
2	C	100	ASP	5.1
1	A	517	LEU	4.9
1	A	611	GLY	4.8
1	A	308	PRO	4.8
1	A	579	ASN	4.6
1	A	300	ARG	4.1
1	A	572	PRO	4.0
1	A	582	LEU	3.8
1	A	293	GLU	3.7
1	A	583	VAL	3.6
1	B	307	GLY	3.5
1	A	295	GLU	3.5
1	A	515	CYS	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	301	LYS	3.3
1	B	308	PRO	3.2
1	B	541	GLN	3.2
1	A	302	CYS	3.2
1	B	503	ARG	3.2
1	A	608	GLY	3.2
1	B	306	GLU	3.2
1	A	581	THR	3.1
1	A	299	VAL	3.0
1	A	518	LEU	2.9
1	A	600	CYS	2.8
1	A	605	THR	2.8
2	D	83	LEU	2.8
1	A	571	CYS	2.8
1	A	514	LYS	2.7
1	A	557	GLN	2.6
1	A	602	TYR	2.6
1	A	584	TRP	2.6
2	D	80	MET	2.5
2	C	98	ILE	2.5
1	A	560	HIS	2.4
1	A	327	ILE	2.3
1	A	606	GLY	2.2
1	A	578	GLU	2.2
1	B	542	ALA	2.2
1	A	519	GLU	2.1
1	A	521	GLU	2.1
1	A	539	LEU	2.1
2	D	118	SER	2.0
1	A	601	THR	2.0

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, 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.

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	F	2	14/15	0.88	0.21	123,127,129,130	0
3	NAG	G	1	14/15	0.91	0.20	117,117,124,127	0
3	NAG	E	2	14/15	0.94	0.16	121,127,129,130	0
3	NAG	G	2	14/15	0.94	0.32	129,131,132,132	0
3	NAG	F	1	14/15	0.95	0.19	117,117,117,123	0
3	NAG	E	1	14/15	0.97	0.23	117,117,117,123	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.
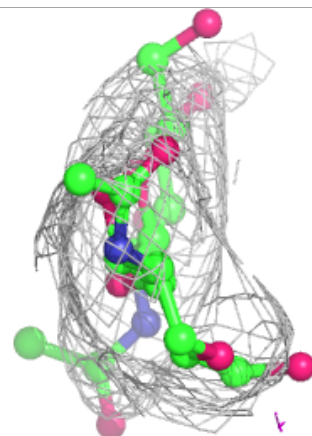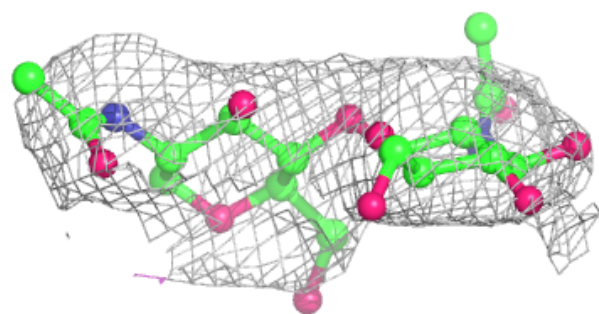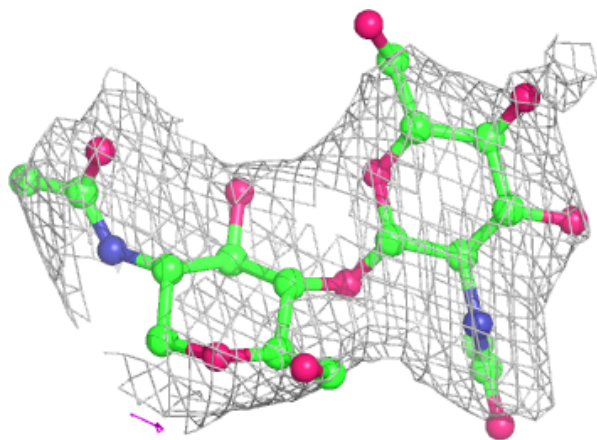

Electron density around Chain E:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
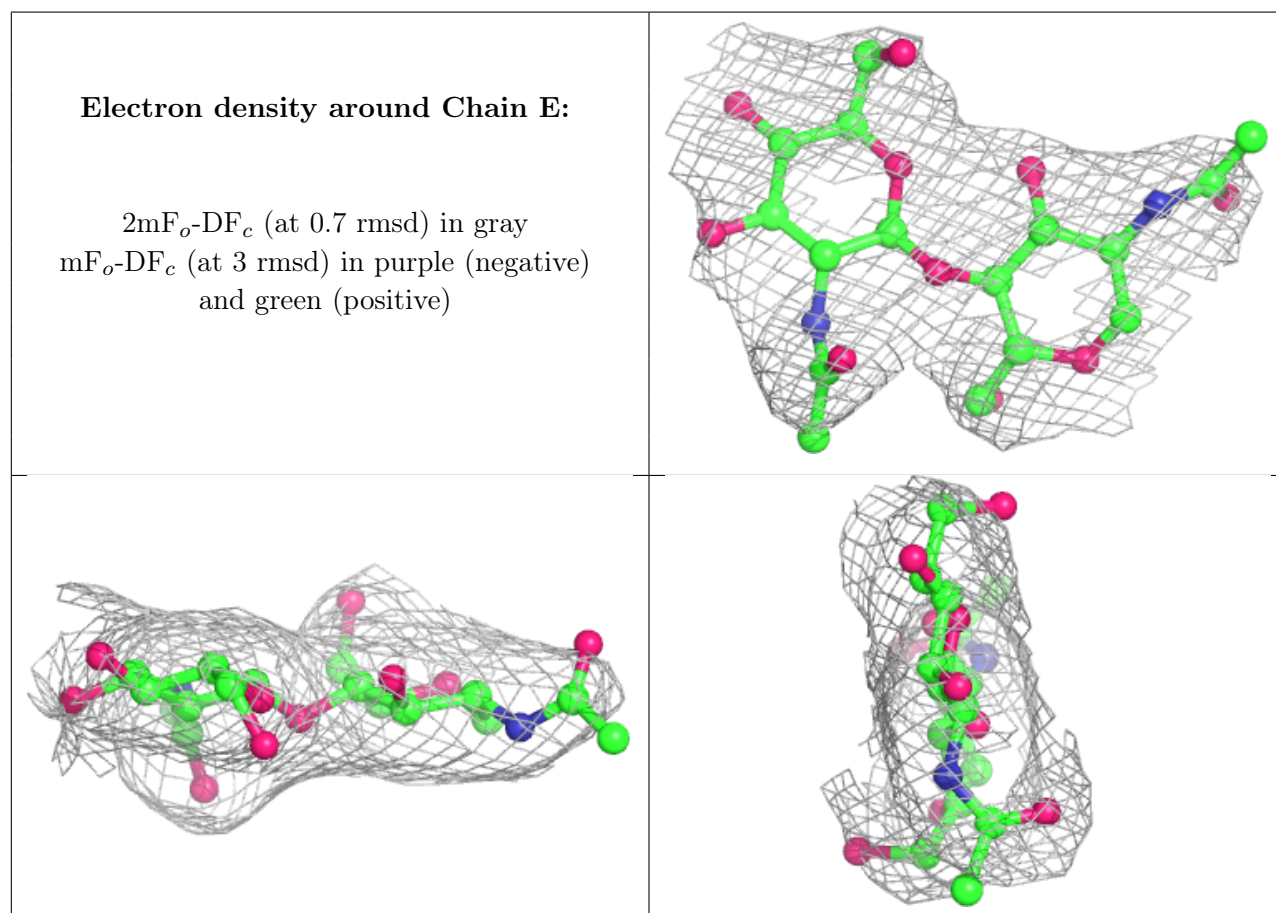
Electron density around Chain F:

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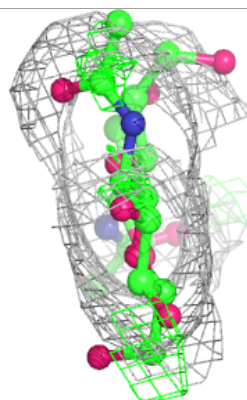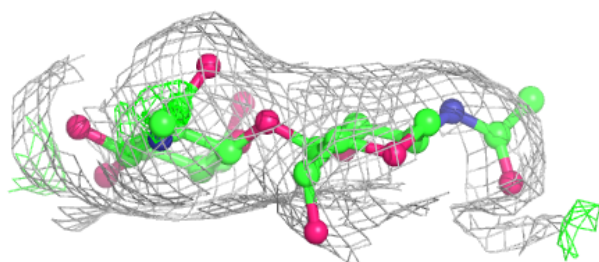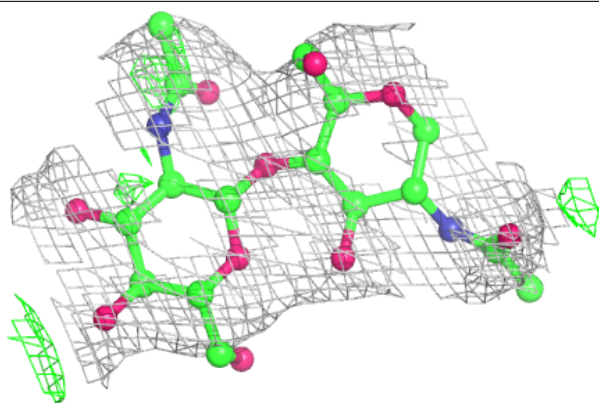
Electron density around Chain G:

$2mF_o\text{-}DF_c$ (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, 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
4	NAG	A	701	14/15	0.84	0.19	120,124,129,129	0

6.5 Other polymers ⓘ

There are no such residues in this entry.