wwPDB X-ray Structure Validation Summary Report (i)

Apr 5, 2022 – 01:12 pm BST

PDB ID	:	7AKK
Title	:	Structure of a complement factor-receptor complex
Authors	:	Fernandez, F.J.; Santos-Lopez, J.; Martinez-Barricarte, R.; Querol-Garcia, J.; Navas-Yuste, S.; Savko, M.; Shepard, W.E.; Rodriguez de Cordoba, S.; Vega, M.C.
Deposited on	:	2020-10-01
Resolution	:	3.40 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

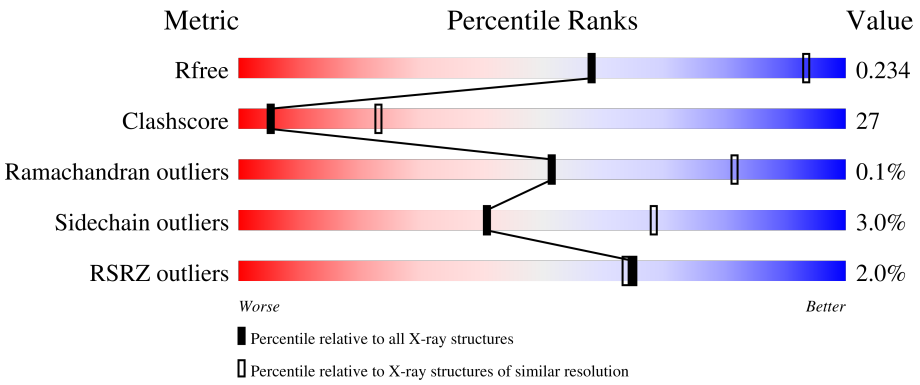
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	B	645	2%, 53%, 46%
1	C	645	2%, 49%, 49%
2	A	898	3%, 47%, 44%, 6%
2	E	898	2%, 27%, 28%, 44%
3	D	195	50%, 43%, 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	195	% 57% 34% 5%
4	I	3	100%
5	J	4	50% 50%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 24012 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 Complement C3 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	645	5024	3198	851	960	15	0	0	0
1	C	645	5024	3198	851	960	15	0	0	0

- Molecule 2 is a protein called Complement C3b alpha' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	842	6736	4276	1128	1295	37	0	0	0
2	E	504	4084	2585	684	789	26	0	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ARG	deletion	UNP P01024
A	?	-	SER	deletion	UNP P01024
A	?	-	SER	deletion	UNP P01024
A	?	-	LYS	deletion	UNP P01024
A	?	-	ILE	deletion	UNP P01024
A	?	-	THR	deletion	UNP P01024
A	?	-	HIS	deletion	UNP P01024
A	?	-	ARG	deletion	UNP P01024
A	?	-	ILE	deletion	UNP P01024
A	?	-	HIS	deletion	UNP P01024
A	?	-	TRP	deletion	UNP P01024
A	?	-	GLU	deletion	UNP P01024
A	?	-	SER	deletion	UNP P01024
A	?	-	ALA	deletion	UNP P01024
A	?	-	SER	deletion	UNP P01024
A	?	-	LEU	deletion	UNP P01024

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LEU	deletion	UNP P01024
E	?	-	ARG	deletion	UNP P01024
E	?	-	SER	deletion	UNP P01024
E	?	-	SER	deletion	UNP P01024
E	?	-	LYS	deletion	UNP P01024
E	?	-	ILE	deletion	UNP P01024
E	?	-	THR	deletion	UNP P01024
E	?	-	HIS	deletion	UNP P01024
E	?	-	ARG	deletion	UNP P01024
E	?	-	ILE	deletion	UNP P01024
E	?	-	HIS	deletion	UNP P01024
E	?	-	TRP	deletion	UNP P01024
E	?	-	GLU	deletion	UNP P01024
E	?	-	SER	deletion	UNP P01024
E	?	-	ALA	deletion	UNP P01024
E	?	-	SER	deletion	UNP P01024
E	?	-	LEU	deletion	UNP P01024
E	?	-	LEU	deletion	UNP P01024

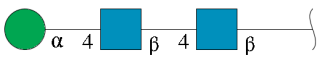
- Molecule 3 is a protein called Integrin alpha-M.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	185	1499	951	267	278	3	0	0	0
			Total	C	N	O	S			
3	H	185	1499	951	267	278	3	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	128	SER	CYS	engineered mutation	UNP P11215
D	316	GLY	ILE	engineered mutation	UNP P11215
H	128	SER	CYS	engineered mutation	UNP P11215
H	316	GLY	ILE	engineered mutation	UNP P11215

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

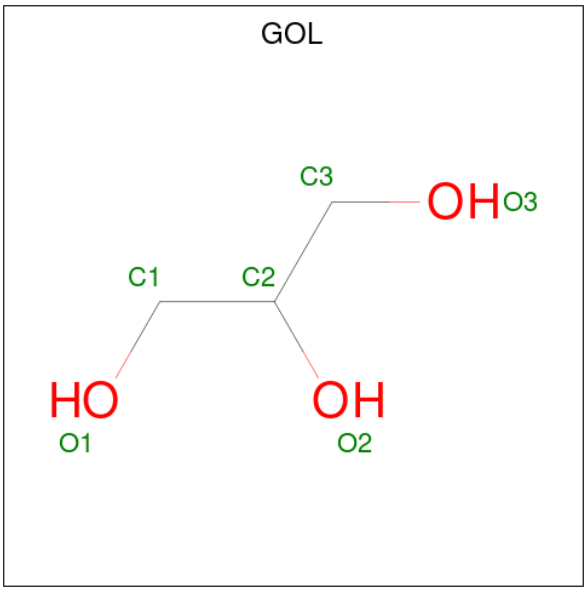


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	I	3	39	22	2	15	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-6)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	J	4	50	28	2	20	0	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	B	1	6	3	3	0	0
6	C	1	6	3	3	0	0
6	E	1	6	3	3	0	0

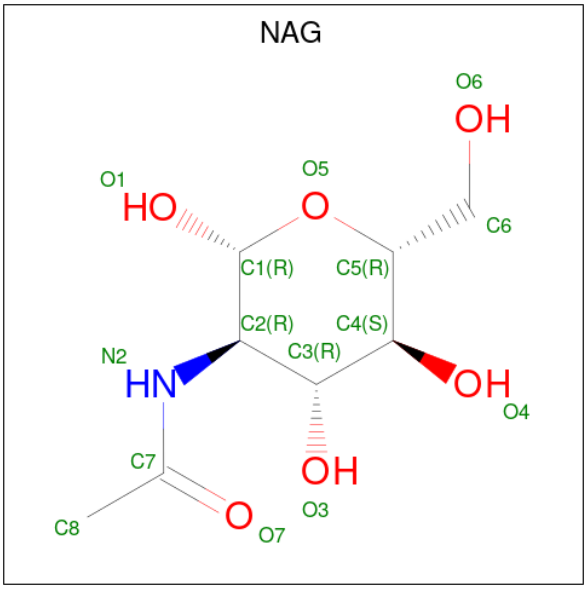
- Molecule 7 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total K 1 1	0	0
7	C	1	Total K 1 1	0	0

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	1	Total Mg 1 1	0	0
8	H	1	Total Mg 1 1	0	0

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	E	1	Total C N O 14 8 1 5	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	B	6	Total O 6 6	0	0
10	A	7	Total O 7 7	0	0

Continued on next page...

Continued from previous page...

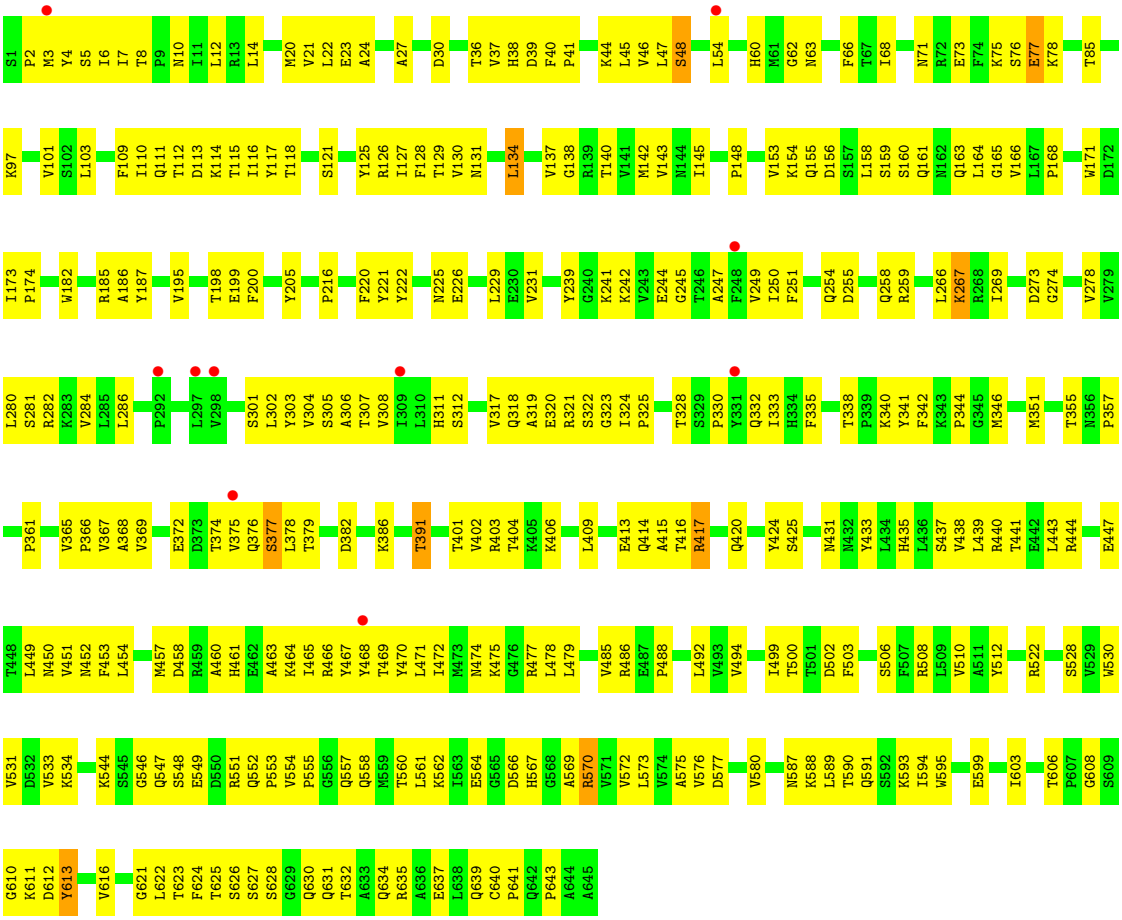
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
10	D	2	2	2	0	0
			Total	O		
10	C	4	4	4	0	0
			Total	O		
10	E	2	2	2	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

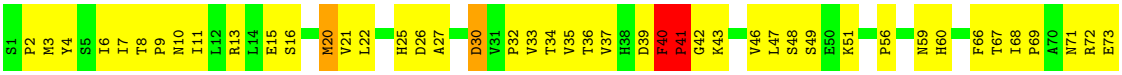
• Molecule 1: Complement C3 beta chain

Chain B: 2% 53% 46%



• Molecule 1: Complement C3 beta chain

Chain C: 2% 49% 49%



F74 K75 G79 R80 N81 V86 Q87 A88 T89 F90 G91 T92 Q93 K97 V98 V99 L100 V101 S102 L103 Q104 S105 G106 Y107 L108 F109 D113 K114 T115 I116 Y117 L124 Y125 R126 I127 F128 T129 V130 N131 H132 L135 P136 V137 G138 R139 T140 V141 M142 V143 N144 I145 E146 N147 P148

E149 P152 V153 K154 Q155 D156 S157 L158 S159 S160 Q161 N162 Q163 L164 G165 V166 L167 P168 M179 G180 Q181 W182 K183 I184 R185 Y188 Q194 V195 F196 S197 T198 E199 V202 K203 E204 Y205 V206 L207 P208 S209 F210 F220 Y224 N225 E226 V231 F237 L238 Y239 G240 K241

K242 V243 E244 A247 F248 V249 I250 F251 G252 R259 K267 R268 I269 P270 I271 G274 V278 R282 D287 P292 R293 A294 E295 D296 L297 V298 S301 L302 Y303 V304 S305 A306 T307 V308 H311 D315 M316 V317 E320 R321 I324 P325 I326 V327 T328 S329

Q332 I333 H334 F335 T336 K337 T338 P339 K343 M346 M351 V352 F353 V354 T355 D358 G359 S360 P361 A362 V365 P366 V367 A368 V369 Q370 T374 V375 Q376 S377 L378 T379 Q380 G383 K386 L387 S388 I389 N390 T391 Q395 L398 T401 V402 R403 T404 K405 K406

S410 E411 A412 E413 Q414 A415 T416 R417 Q420 P423 Y424 S425 T426 N431 N432 Y433 L434 S437 V438 L439 R440 T441 E442 L443 R444 L449 N450 V451 N452 F453 L454 L455 R456 M457 D458 R459 K464 I465 R466 Y467 Y468 T469 Y470 L471 I472 M473 N474 R477 L478 L479 K480 P488

L492 V493 V494 L495 P496 T500 T501 D502 F503 I504 P505 S506 F507 R508 L509 V510 A511 Y512 Y513 T514 L515 R522 E523 V524 V525 A526 D527 S528 D532 D535 V538 G539 S540 L541 V542 V543 K544 Q547 D550 P553 G556 Q557 Q558 M559 T560 L561 K562 I563 E564 G565

R570 V571 V572 L573 D577 K578 G579 V580 F581 V582 L583 K588 L589 T590 Q591 S592 K593 I594 W595 D596 V597 V598 E599 I603 G604 G605 T606 P607 G608 S609 G610 Y613 V616 D619 L622 T623 F624 T625 S626 S627 S628 G629 Q630 Q631 T632 A633 Q634 R635 L638 Q639 C640

P641 Q642 A645

- Molecule 2: Complement C3b alpha' chain

Chain A: 3% 47% 44% • 6%

SER ASN LEU D649 E650 D651 I652 E655 E656 R661 P665 W668 L669 W670 N671 V672 E673 E677 K680 N681 G682 I683 S684 T685 K686 L687 M688 M689 I690 F691 L692 K693 D694 S695 I696 T697 T698 W699 E700 T701 L702 A703 V704 S705 M706 S707 K710 G711 V714 P717 F718

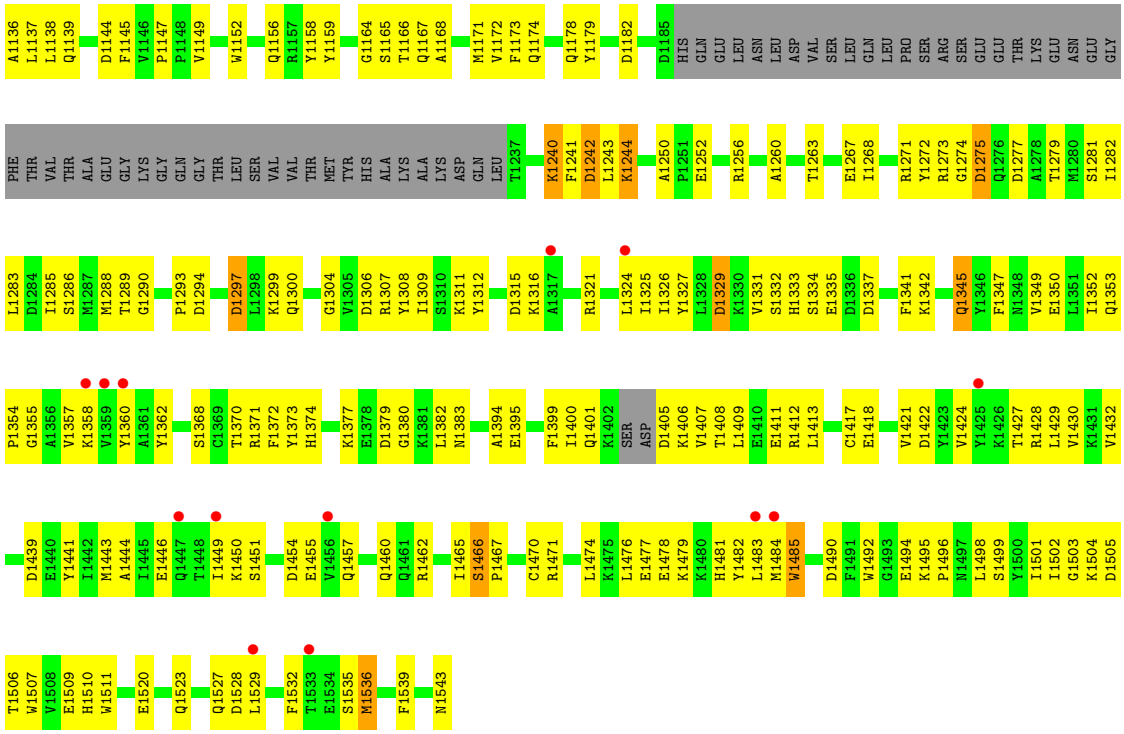
E719 V720 T721 V722 M723 F727 I728 D729 L730 R731 Y734 S735 V736 V737 R738 N739 E740 Q741 V742 E743 I744 Y749 N750 Y751 R752 Q753 E756 L757 K758 V759 R760 V761 E762 L763 L764 A768 F769 C770 S771 L772 A773 T774 T775 K776 R777 R778 Q781 T782 I785 P786 P787 K788 S789

S790 L791 S792 V793 V796 I797 V798 P799 L800 K801 T802 G803 L804 Q805 K810 A811 A812 V813 Y814 H815 H816 F817 I818 S819 D820 G821 V822 K827 V828 V829 P830 E831 G832 I833 R834 M835 N836 K837 V841 L844 Q855 T882 P883 V884 A885 Q886 A894 E895 R896 L897 L900

S905 G906 C907 G908 E909 Q910 N911 M912 I913 G914 M915 T916 P917 T918 V919 I920 D926 Q930 W931 F934 G935 L936 E937 K938 R939 Q940 G941 A942 L943 E944 L945 I946 K947 K948 G949 Y950 T951 Q952 Q953 L954 A955 F956 R957 Q958 P959 S960 F963 A964 A965 F966 V967 K968 R969 A970 P971

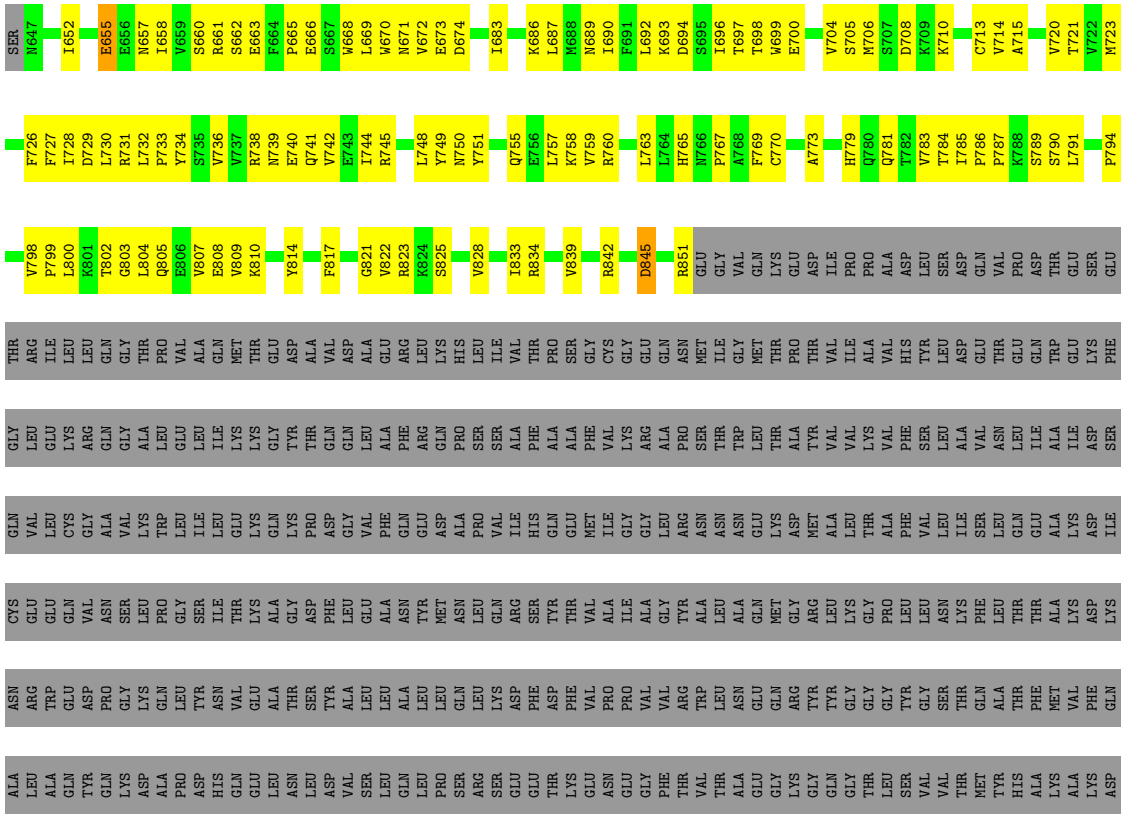
S972 T973 W974 L975 T976 A977 Y978 V979 V982 L989 I990 V996 L997 C998 V1001 K1002 W1003 K1008 Q1009 K1010 P1011 D1012 G1013 V1014 F1015 D1018 A1019 P1020 V1021 I1022 H1023 Q1024 E1025 M1026 I1027 L1030 N1033 N1034 D1037 V1044 L1045 I1046 S1047 A1051 I1054 V1059 N1060

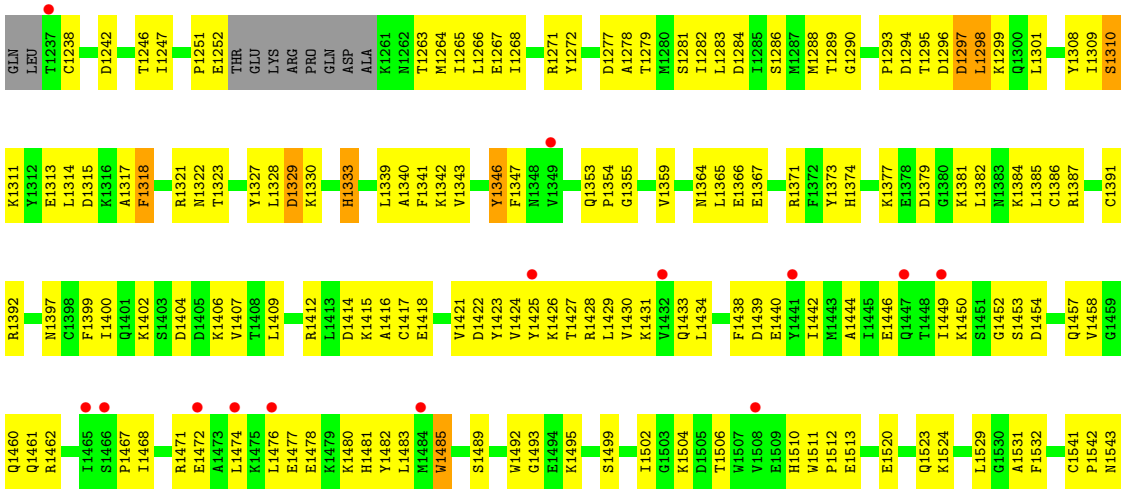
S1061 L1062 I1066 G1070 D1071 F1072 L1073 E1074 Y1077 M1078 N1079 L1080 S1083 V1086 G1090 L1093 M1096 G1097 R1098 L1099 K1100 G1101 P1102 L1103 L1104 N1105 K1106 F1107 L1108 T1109 T1110 A1111 K1112 D1113 K1114 N1115 R1116 W1117 E1118 D1119 P1120 G1121 L1124 Y1125 E1128 Y1132 A1133 L1134 L1135



• Molecule 2: Complement C3b alpha' chain

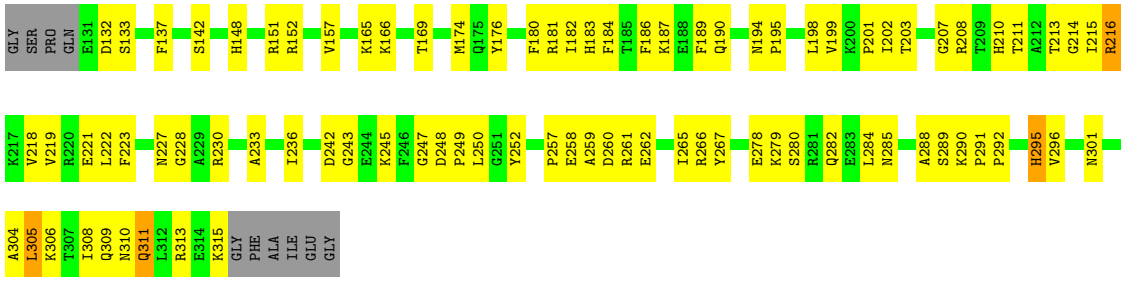
Chain E:





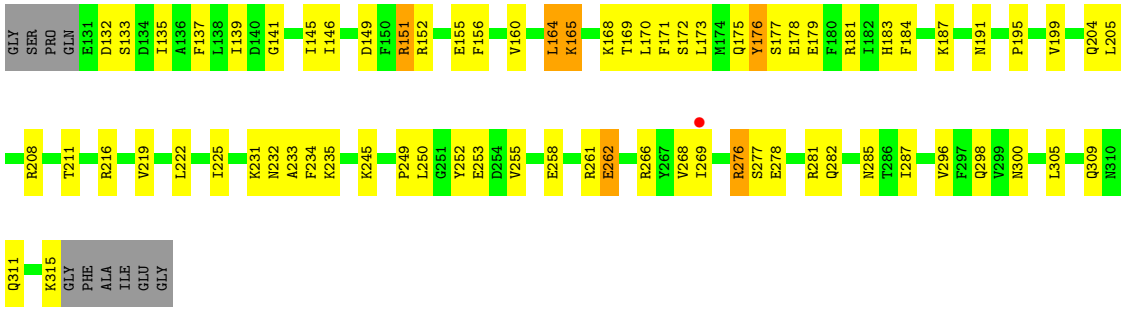
• Molecule 3: Integrin alpha-M

Chain D:



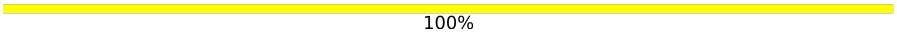
• Molecule 3: Integrin alpha-M

Chain H:



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

Chain I:



• Molecule 5: beta-D-mannopyranose-(1-6)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	111.29Å 150.73Å 111.30Å 90.00° 92.80° 90.00°	Depositor
Resolution (Å)	53.77 – 3.40 53.77 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.3 (53.77-3.40) 99.3 (53.77-3.40)	Depositor EDS
R_{merge}	0.34	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.38 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.192 , 0.229 0.192 , 0.234	Depositor DCC
R_{free} test set	2520 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	105.3	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.37$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.109 for l,k,-h 0.117 for h,-k,-l 0.326 for l,-k,h	Xtriage
Reported twinning fraction	0.500 for l,-k,h	Depositor
Outliers	0 of 50383 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	24012	wwPDB-VP
Average B, all atoms (Å^2)	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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: GOL, MG, MAN, K, NAG, BMA

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	B	0.63	0/5126	0.82	4/6966 (0.1%)
1	C	0.61	1/5126 (0.0%)	0.83	2/6966 (0.0%)
2	A	0.56	1/6871 (0.0%)	0.73	5/9303 (0.1%)
2	E	0.61	0/4164	0.77	1/5629 (0.0%)
3	D	0.55	0/1527	0.70	0/2054
3	H	0.67	2/1527 (0.1%)	0.80	1/2054 (0.0%)
All	All	0.60	4/24341 (0.0%)	0.78	13/32972 (0.0%)

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	B	0	1
1	C	0	1
2	A	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	640	CYS	CB-SG	6.77	1.93	1.82
3	H	262	GLU	CG-CD	5.90	1.60	1.51
2	A	749	TYR	CD1-CE1	-5.24	1.31	1.39
3	H	262	GLU	CB-CG	5.21	1.62	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1119	ASP	CB-CG-OD2	-9.79	109.49	118.30
1	B	391	THR	C-N-CA	-7.50	102.95	121.70
1	C	41	PRO	CA-N-CD	-6.88	101.87	111.50
2	E	713	CYS	CA-CB-SG	6.30	125.35	114.00
2	A	1119	ASP	CB-CG-OD1	6.14	123.82	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	817	PHE	Peptide
1	B	640	CYS	Peptide
1	C	40	PHE	Peptide

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	B	5024	0	5084	275	0
1	C	5024	0	5084	332	1
2	A	6736	0	6668	365	2
2	E	4084	0	4022	250	0
3	D	1499	0	1509	66	1
3	H	1499	0	1511	68	2
4	I	39	0	33	2	0
5	J	50	0	43	1	0
6	B	6	0	8	0	0
6	C	6	0	8	0	0
6	E	6	0	8	1	0
7	A	1	0	0	0	0
7	C	1	0	0	0	0
8	D	1	0	0	0	0
8	H	1	0	0	0	0
9	E	14	0	13	0	0
10	A	7	0	0	0	0
10	B	6	0	0	0	0
10	C	4	0	0	0	0
10	D	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	E	2	0	0	0	0
All	All	24012	0	23991	1272	3

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 27.

The worst 5 of 1272 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:C:370:GLN:HG3	1:C:401:THR:HB	1.14	1.12
1:C:369:VAL:H	1:C:375:VAL:CG2	1.65	1.09
1:C:369:VAL:H	1:C:375:VAL:HG21	0.94	1.07
1:C:369:VAL:HG22	1:C:375:VAL:HG11	1.34	1.07
1:B:155:GLN:HE21	2:E:804:LEU:HD13	1.13	1.06

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:756:GLU:OE2	3:H:300:ASN:ND2[2_747]	1.93	0.27
2:A:1428:ARG:NH2	3:H:191:ASN:O[1_455]	2.07	0.13
3:D:249:PRO:O	1:C:71:ASN:ND2[2_657]	2.08	0.12

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	643/645 (100%)	609 (95%)	34 (5%)	0	100 100
1	C	643/645 (100%)	617 (96%)	25 (4%)	1 (0%)	47 78
2	A	836/898 (93%)	743 (89%)	92 (11%)	1 (0%)	51 82

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	498/898 (56%)	441 (89%)	57 (11%)	0	100	100
3	D	183/195 (94%)	169 (92%)	14 (8%)	0	100	100
3	H	183/195 (94%)	173 (94%)	9 (5%)	1 (0%)	29	61
All	All	2986/3476 (86%)	2752 (92%)	231 (8%)	3 (0%)	51	82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	41	PRO
2	A	656	GLU
3	H	177	SER

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	B	567/567 (100%)	551 (97%)	16 (3%)	43	70
1	C	567/567 (100%)	555 (98%)	12 (2%)	53	76
2	A	745/794 (94%)	720 (97%)	25 (3%)	37	65
2	E	463/794 (58%)	448 (97%)	15 (3%)	39	67
3	D	166/172 (96%)	159 (96%)	7 (4%)	30	59
3	H	166/172 (96%)	162 (98%)	4 (2%)	49	74
All	All	2674/3066 (87%)	2595 (97%)	79 (3%)	41	68

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	440	ARG
2	E	1346	TYR
1	C	583	LEU
2	E	1297	ASP
3	H	151	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	642	GLN
3	H	191	ASN
2	E	689	ASN
2	E	1383	ASN
2	A	1276	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)

7 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	$\#\|Z\| > 2$
4	NAG	I	1	1,4	14,14,15	0.26	0	17,19,21	0.48	0
4	NAG	I	2	4	14,14,15	0.75	0	17,19,21	1.12	2 (11%)
4	MAN	I	3	4	11,11,12	2.27	4 (36%)	15,15,17	1.58	4 (26%)
5	NAG	J	1	5,1	14,14,15	0.94	1 (7%)	17,19,21	0.57	0
5	NAG	J	2	5	14,14,15	0.87	1 (7%)	17,19,21	0.80	0
5	MAN	J	3	5	11,11,12	2.16	4 (36%)	15,15,17	1.87	2 (13%)
5	BMA	J	4	5	11,11,12	1.85	4 (36%)	15,15,17	1.13	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	I	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	I	2	4	-	2/6/23/26	0/1/1/1
4	MAN	I	3	4	-	1/2/19/22	0/1/1/1
5	NAG	J	1	5,1	-	2/6/23/26	0/1/1/1
5	NAG	J	2	5	-	2/6/23/26	0/1/1/1
5	MAN	J	3	5	-	1/2/19/22	0/1/1/1
5	BMA	J	4	5	-	0/2/19/22	0/1/1/1

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	I	3	MAN	O5-C1	5.53	1.52	1.43
5	J	3	MAN	C2-C3	4.80	1.59	1.52
5	J	4	BMA	C4-C3	3.18	1.60	1.52
5	J	1	NAG	O5-C1	2.98	1.48	1.43
5	J	4	BMA	O5-C1	2.89	1.48	1.43

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	J	3	MAN	C1-O5-C5	5.18	119.21	112.19
4	I	3	MAN	C1-C2-C3	3.22	113.63	109.67
5	J	3	MAN	C1-C2-C3	3.16	113.55	109.67
4	I	3	MAN	C1-O5-C5	2.72	115.87	112.19
4	I	3	MAN	O5-C1-C2	-2.55	106.84	110.77

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	J	1	NAG	O5-C5-C6-O6
4	I	1	NAG	O5-C5-C6-O6
4	I	2	NAG	O5-C5-C6-O6
5	J	1	NAG	C4-C5-C6-O6
4	I	1	NAG	C4-C5-C6-O6

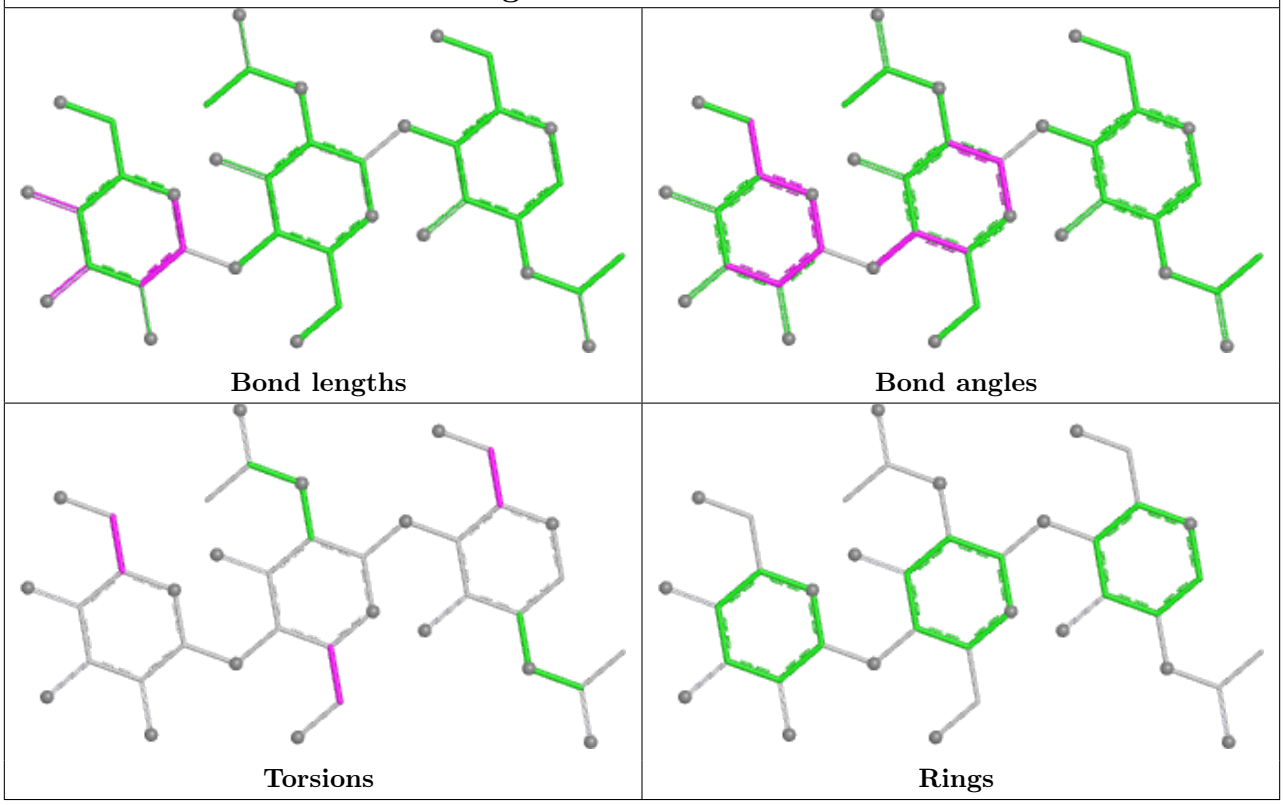
There are no ring outliers.

3 monomers are involved in 3 short contacts:

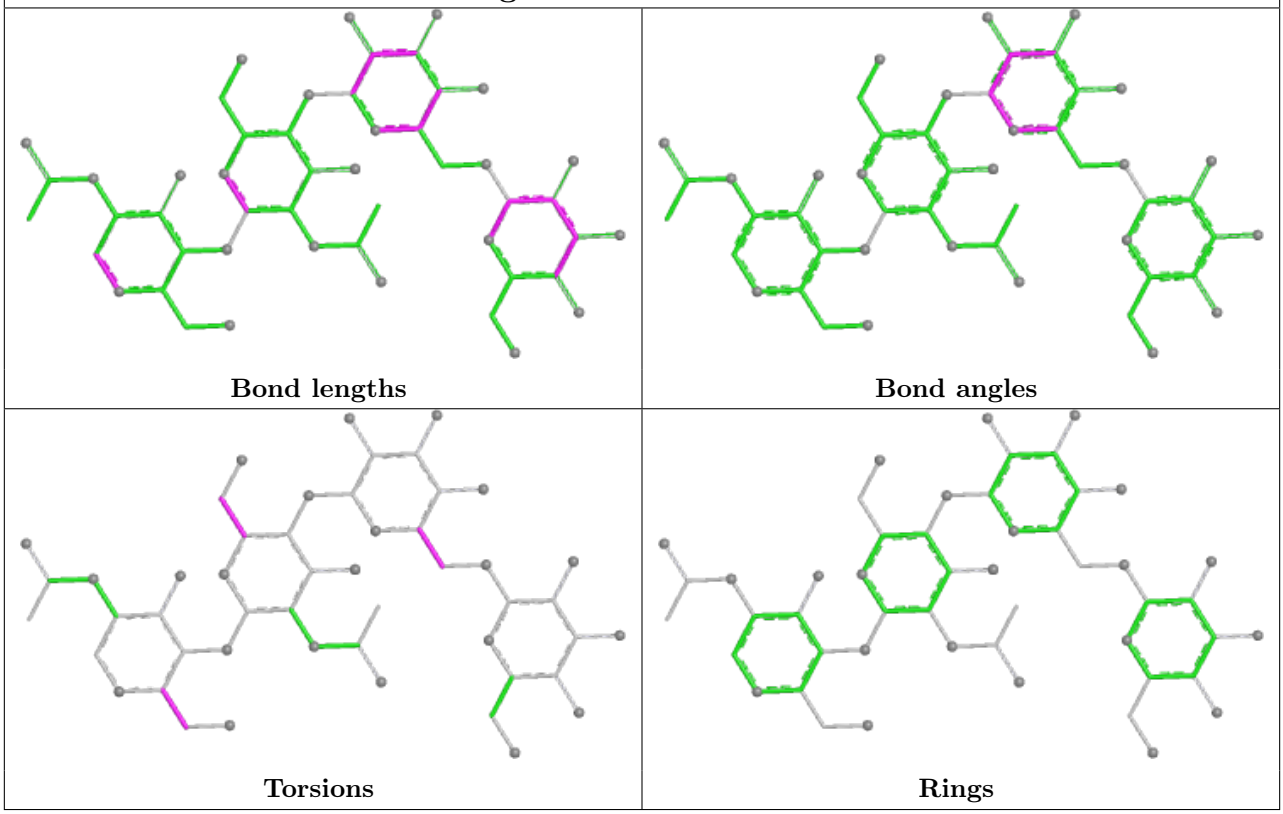
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	4	BMA	1	0
5	J	2	NAG	1	0
4	I	1	NAG	2	0

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

Oligosaccharide Chain I



Oligosaccharide Chain J



5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
6	GOL	C	1601	-	5,5,5	0.90	0	5,5,5	0.99	0
9	NAG	E	1601	2	14,14,15	0.38	0	17,19,21	0.79	1 (5%)
6	GOL	E	1602	-	5,5,5	1.11	0	5,5,5	1.00	0
6	GOL	B	1601	-	5,5,5	1.66	2 (40%)	5,5,5	0.58	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
6	GOL	C	1601	-	-	0/4/4/4	-
9	NAG	E	1601	2	-	0/6/23/26	0/1/1/1
6	GOL	E	1602	-	-	2/4/4/4	-
6	GOL	B	1601	-	-	3/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1601	GOL	C3-C2	2.93	1.63	1.51
6	B	1601	GOL	C1-C2	2.08	1.60	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	E	1601	NAG	C1-O5-C5	2.19	115.16	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	1601	GOL	C1-C2-C3-O3
6	E	1602	GOL	O1-C1-C2-C3
6	E	1602	GOL	O1-C1-C2-O2
6	B	1601	GOL	O2-C2-C3-O3
6	B	1601	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	1602	GOL	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	B	645/645 (100%)	-0.09	10 (1%)	72	70	47, 108, 152, 185	0
1	C	645/645 (100%)	-0.09	10 (1%)	72	70	57, 104, 150, 179	0
2	A	842/898 (93%)	-0.15	25 (2%)	50	49	55, 127, 164, 192	0
2	E	504/898 (56%)	-0.16	14 (2%)	53	51	58, 111, 170, 204	0
3	D	185/195 (94%)	-0.18	0	100	100	78, 108, 133, 155	0
3	H	185/195 (94%)	-0.26	1 (0%)	91	90	68, 93, 123, 145	0
All	All	3006/3476 (86%)	-0.14	60 (1%)	65	64	47, 111, 159, 204	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	388	SER	6.9
2	A	1317	ALA	5.3
2	E	1447	GLN	5.2
2	A	1447	GLN	5.1
2	A	1037	ASP	4.9

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

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

6.3 Carbohydrates (i)

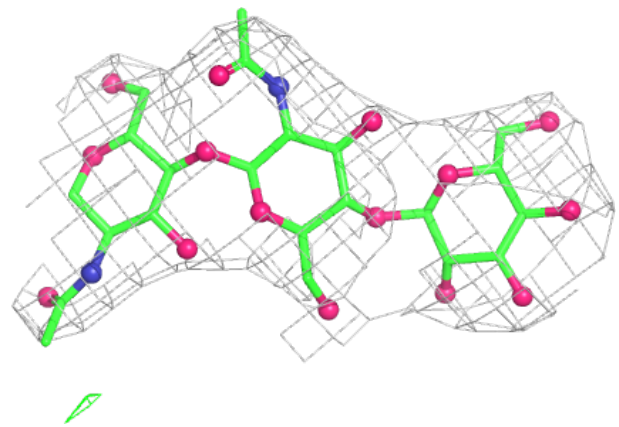
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

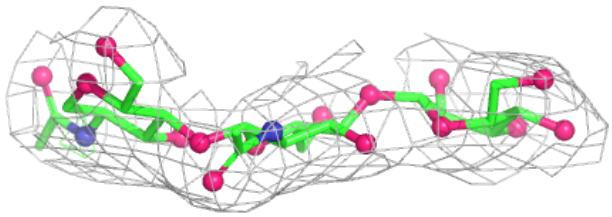
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	BMA	J	4	11/12	0.91	0.23	119,126,140,143	0
5	MAN	J	3	11/12	0.92	0.16	135,144,147,150	0
4	MAN	I	3	11/12	0.94	0.11	120,126,137,141	0
4	NAG	I	1	14/15	0.95	0.14	123,128,134,135	0
4	NAG	I	2	14/15	0.96	0.08	138,141,144,145	0
5	NAG	J	1	14/15	0.96	0.20	98,107,131,132	0
5	NAG	J	2	14/15	0.97	0.17	106,118,132,140	0

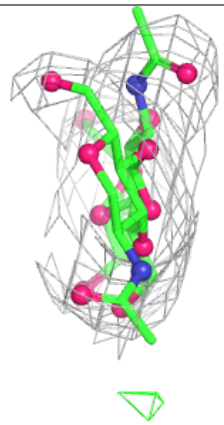
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain I:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

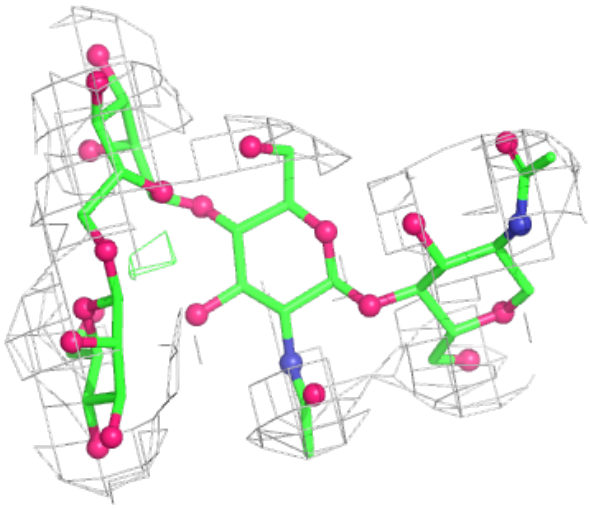


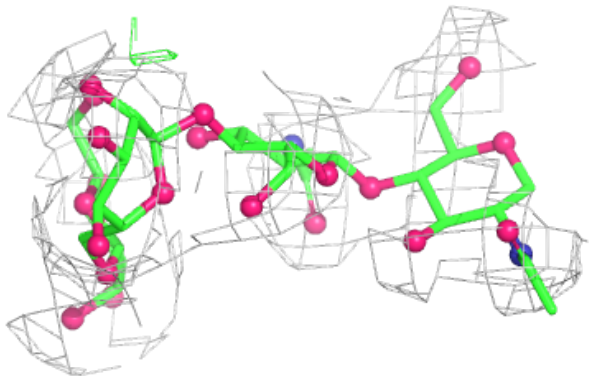


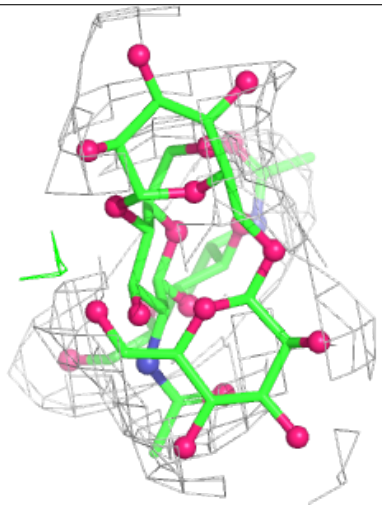


Electron density around Chain J:

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







6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	NAG	E	1601	14/15	0.74	0.24	157,161,166,168	0
6	GOL	C	1601	6/6	0.90	0.19	81,91,93,93	0
6	GOL	E	1602	6/6	0.95	0.20	69,73,75,76	0
6	GOL	B	1601	6/6	0.96	0.23	67,70,71,73	0
7	K	C	1602	1/1	0.98	0.14	71,71,71,71	0
8	MG	D	2101	1/1	0.99	0.11	117,117,117,117	0
8	MG	H	2001	1/1	0.99	0.15	82,82,82,82	0
7	K	A	1601	1/1	0.99	0.22	71,71,71,71	0

6.5 Other polymers (i)

There are no such residues in this entry.