wwPDB EM Validation Summary Report ⓘ

Dec 18, 2022 – 06:57 am GMT

PDB ID	:	7A5O
EMDB ID	:	EMD-10517
Title	:	Human MUC2 AAs 21-1397
Authors	:	Javitt, G.; Khmelnitsky, L.; Albert, L.; Elad, N.; Ilani, T.; Diskin, R.; Fass, D.
Deposited on	:	2020-08-21
Resolution	:	2.95 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 ⓘ 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 ⓘ) were used in the production of this report:

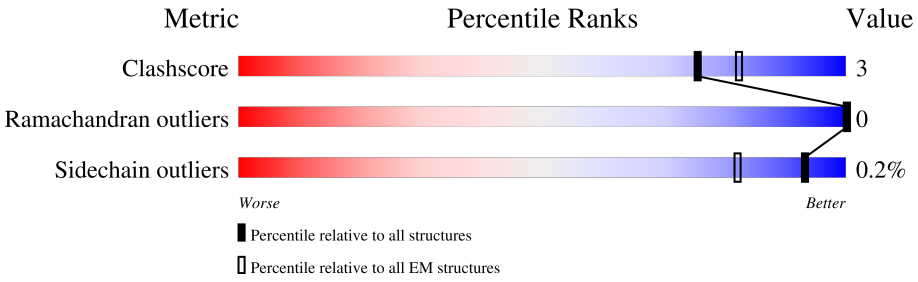
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.95 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.



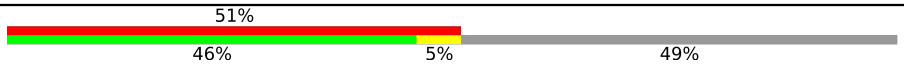
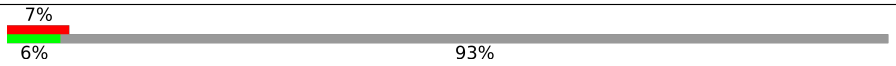
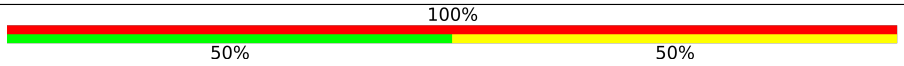
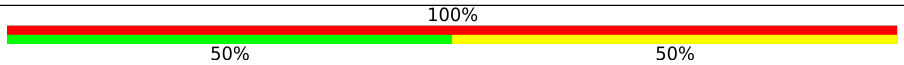
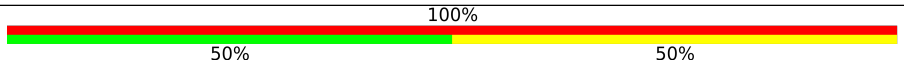
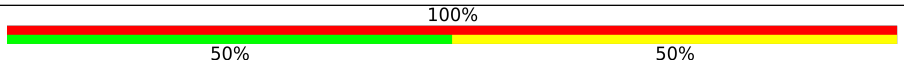
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1383	41% 80% 7% 13%
1	B	1383	5% 46% 5% 49%
1	C	1383	50% 73% 7% 19%
1	D	1383	7% 33% 64%
1	E	1383	81% 80% 8% 13%
1	F	1383	6% 93%
1	G	1383	81% 73% 7% 19%
1	H	1383	36% 33% 64%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	1383	
1	J	1383	
2	K	2	
2	L	2	
2	M	2	
2	N	2	
2	O	2	
2	P	2	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 55254 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 Mucin-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1205	Total	C	N	O	S	4	0
			9125	5668	1579	1760	118		
1	B	708	Total	C	N	O	S	3	0
			5276	3278	918	1013	67		
1	C	1115	Total	C	N	O	S	4	0
			8429	5239	1459	1622	109		
1	D	497	Total	C	N	O	S	1	0
			3849	2390	661	747	51		
1	E	1205	Total	C	N	O	S	4	0
			9125	5668	1579	1760	118		
1	F	90	Total	C	N	O	S	0	0
			696	429	120	138	9		
1	G	1115	Total	C	N	O	S	4	0
			8429	5239	1459	1622	109		
1	H	497	Total	C	N	O	S	1	0
			3849	2390	661	747	51		
1	I	708	Total	C	N	O	S	3	0
			5276	3278	918	1013	67		
1	J	90	Total	C	N	O	S	0	0
			696	429	120	138	9		

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1325	THR	PRO	conflict	UNP Q02817
A	1398	HIS	-	expression tag	UNP Q02817
A	1399	HIS	-	expression tag	UNP Q02817
A	1400	HIS	-	expression tag	UNP Q02817
A	1401	HIS	-	expression tag	UNP Q02817
A	1402	HIS	-	expression tag	UNP Q02817
A	1403	HIS	-	expression tag	UNP Q02817
B	1325	THR	PRO	conflict	UNP Q02817
B	1398	HIS	-	expression tag	UNP Q02817
B	1399	HIS	-	expression tag	UNP Q02817

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1400	HIS	-	expression tag	UNP Q02817
B	1401	HIS	-	expression tag	UNP Q02817
B	1402	HIS	-	expression tag	UNP Q02817
B	1403	HIS	-	expression tag	UNP Q02817
C	1325	THR	PRO	conflict	UNP Q02817
C	1398	HIS	-	expression tag	UNP Q02817
C	1399	HIS	-	expression tag	UNP Q02817
C	1400	HIS	-	expression tag	UNP Q02817
C	1401	HIS	-	expression tag	UNP Q02817
C	1402	HIS	-	expression tag	UNP Q02817
C	1403	HIS	-	expression tag	UNP Q02817
D	1325	THR	PRO	conflict	UNP Q02817
D	1398	HIS	-	expression tag	UNP Q02817
D	1399	HIS	-	expression tag	UNP Q02817
D	1400	HIS	-	expression tag	UNP Q02817
D	1401	HIS	-	expression tag	UNP Q02817
D	1402	HIS	-	expression tag	UNP Q02817
D	1403	HIS	-	expression tag	UNP Q02817
E	1325	THR	PRO	conflict	UNP Q02817
E	1398	HIS	-	expression tag	UNP Q02817
E	1399	HIS	-	expression tag	UNP Q02817
E	1400	HIS	-	expression tag	UNP Q02817
E	1401	HIS	-	expression tag	UNP Q02817
E	1402	HIS	-	expression tag	UNP Q02817
E	1403	HIS	-	expression tag	UNP Q02817
F	1325	THR	PRO	conflict	UNP Q02817
F	1398	HIS	-	expression tag	UNP Q02817
F	1399	HIS	-	expression tag	UNP Q02817
F	1400	HIS	-	expression tag	UNP Q02817
F	1401	HIS	-	expression tag	UNP Q02817
F	1402	HIS	-	expression tag	UNP Q02817
F	1403	HIS	-	expression tag	UNP Q02817
G	1325	THR	PRO	conflict	UNP Q02817
G	1398	HIS	-	expression tag	UNP Q02817
G	1399	HIS	-	expression tag	UNP Q02817
G	1400	HIS	-	expression tag	UNP Q02817
G	1401	HIS	-	expression tag	UNP Q02817
G	1402	HIS	-	expression tag	UNP Q02817
G	1403	HIS	-	expression tag	UNP Q02817
H	1325	THR	PRO	conflict	UNP Q02817
H	1398	HIS	-	expression tag	UNP Q02817
H	1399	HIS	-	expression tag	UNP Q02817

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	1400	HIS	-	expression tag	UNP Q02817
H	1401	HIS	-	expression tag	UNP Q02817
H	1402	HIS	-	expression tag	UNP Q02817
H	1403	HIS	-	expression tag	UNP Q02817
I	1325	THR	PRO	conflict	UNP Q02817
I	1398	HIS	-	expression tag	UNP Q02817
I	1399	HIS	-	expression tag	UNP Q02817
I	1400	HIS	-	expression tag	UNP Q02817
I	1401	HIS	-	expression tag	UNP Q02817
I	1402	HIS	-	expression tag	UNP Q02817
I	1403	HIS	-	expression tag	UNP Q02817
J	1325	THR	PRO	conflict	UNP Q02817
J	1398	HIS	-	expression tag	UNP Q02817
J	1399	HIS	-	expression tag	UNP Q02817
J	1400	HIS	-	expression tag	UNP Q02817
J	1401	HIS	-	expression tag	UNP Q02817
J	1402	HIS	-	expression tag	UNP Q02817
J	1403	HIS	-	expression tag	UNP Q02817

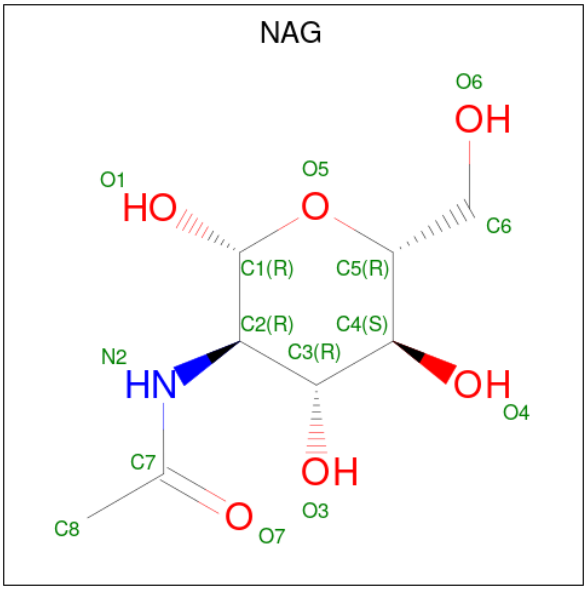
- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	K	2	Total	C	N	O	0	0
			28	16	2	10		
2	L	2	Total	C	N	O	0	0
			28	16	2	10		
2	M	2	Total	C	N	O	0	0
			28	16	2	10		
2	N	2	Total	C	N	O	0	0
			28	16	2	10		
2	O	2	Total	C	N	O	0	0
			28	16	2	10		
2	P	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	AltConf
3	A	5	Total Ca 5 5	0
3	B	2	Total Ca 2 2	0
3	C	3	Total Ca 3 3	0
3	D	3	Total Ca 3 3	0
3	E	5	Total Ca 5 5	0
3	F	2	Total Ca 2 2	0
3	G	3	Total Ca 3 3	0
3	H	3	Total Ca 3 3	0
3	I	2	Total Ca 2 2	0
3	J	2	Total Ca 2 2	0

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



Mol	Chain	Residues	Atoms	AltConf
4	A	1	Total C N O 28 16 2 10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf
4	A	1	Total C N O 28 16 2 10	0
4	B	1	Total C N O 28 16 2 10	0
4	B	1	Total C N O 28 16 2 10	0
4	C	1	Total C N O 28 16 2 10	0
4	C	1	Total C N O 28 16 2 10	0
4	E	1	Total C N O 28 16 2 10	0
4	E	1	Total C N O 28 16 2 10	0
4	G	1	Total C N O 28 16 2 10	0
4	G	1	Total C N O 28 16 2 10	0
4	I	1	Total C N O 28 16 2 10	0
4	I	1	Total C N O 28 16 2 10	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	AltConf
5	A	23	Total O 23 23	0
5	B	10	Total O 10 10	0
5	C	22	Total O 22 22	0
5	D	13	Total O 13 13	0
5	E	23	Total O 23 23	0
5	F	1	Total O 1 1	0
5	G	22	Total O 22 22	0
5	H	13	Total O 13 13	0

Continued on next page...

Continued from previous page...

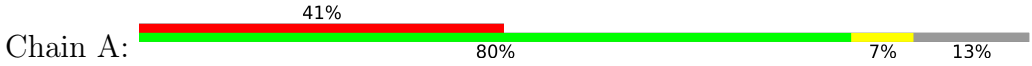
Mol	Chain	Residues	Atoms	AltConf
5	I	10	Total O 10 10	0
5	J	1	Total O 1 1	0

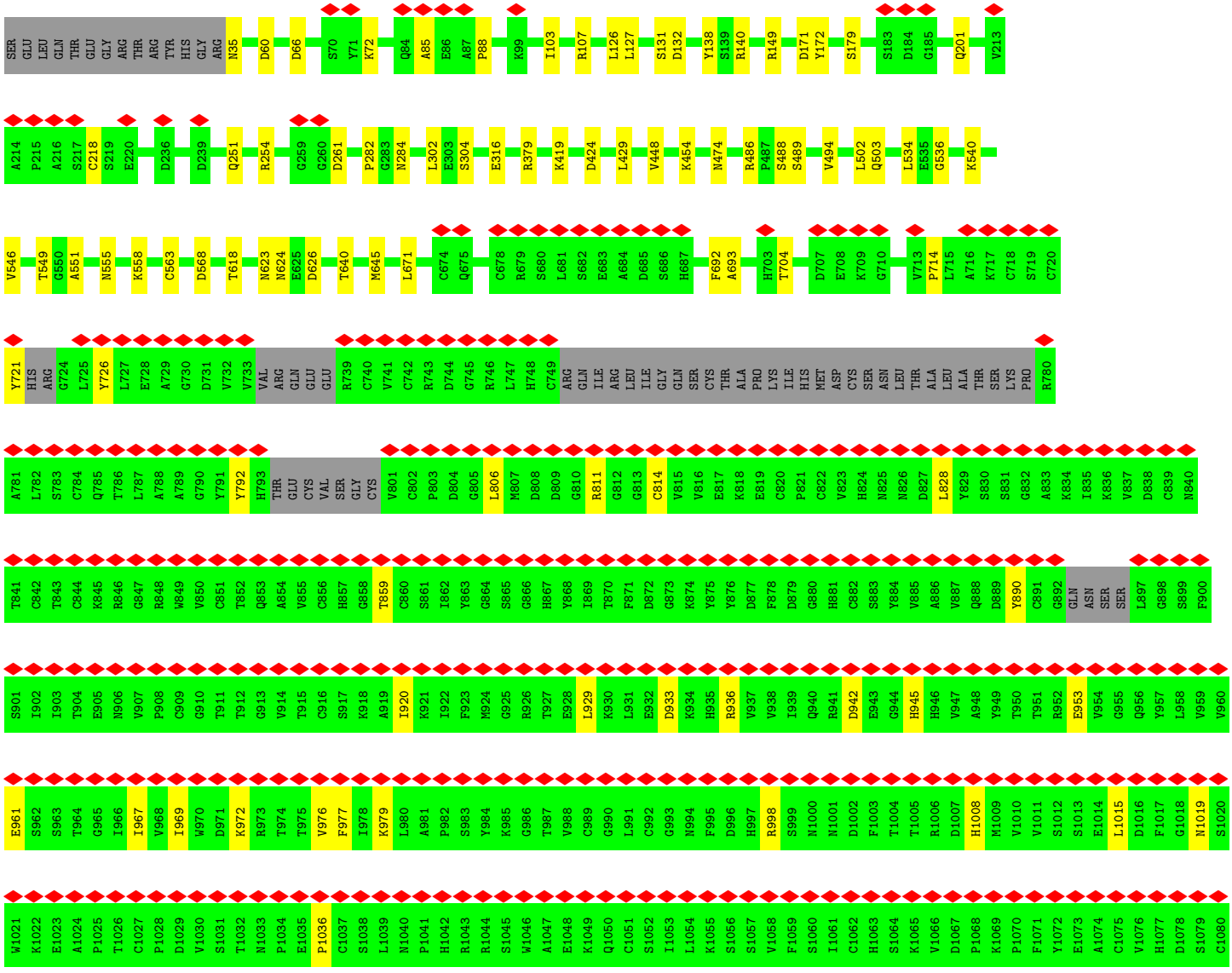
3 Residue-property plots (i)

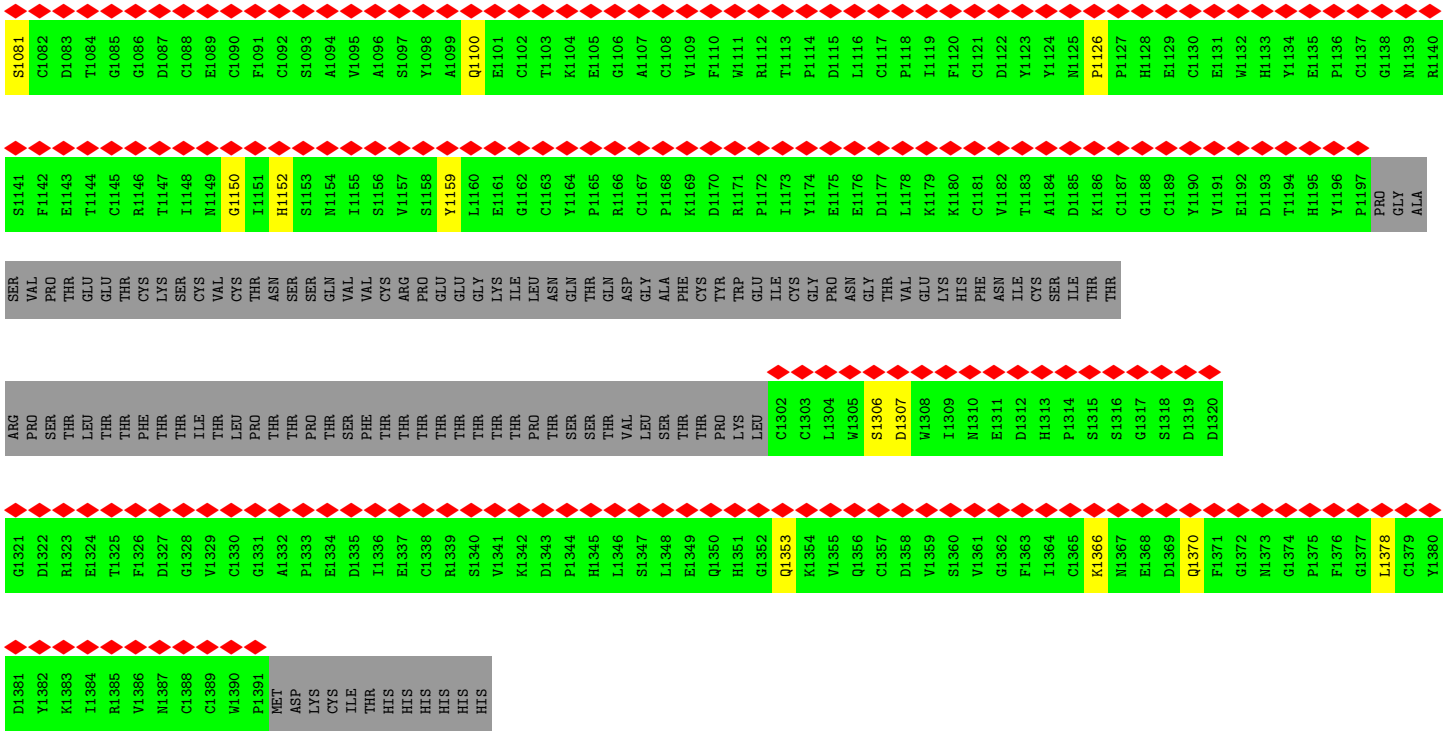
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: Mucin-2

Chain A:

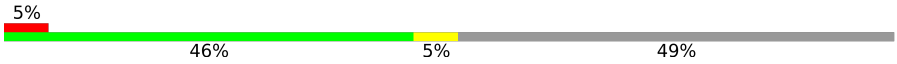


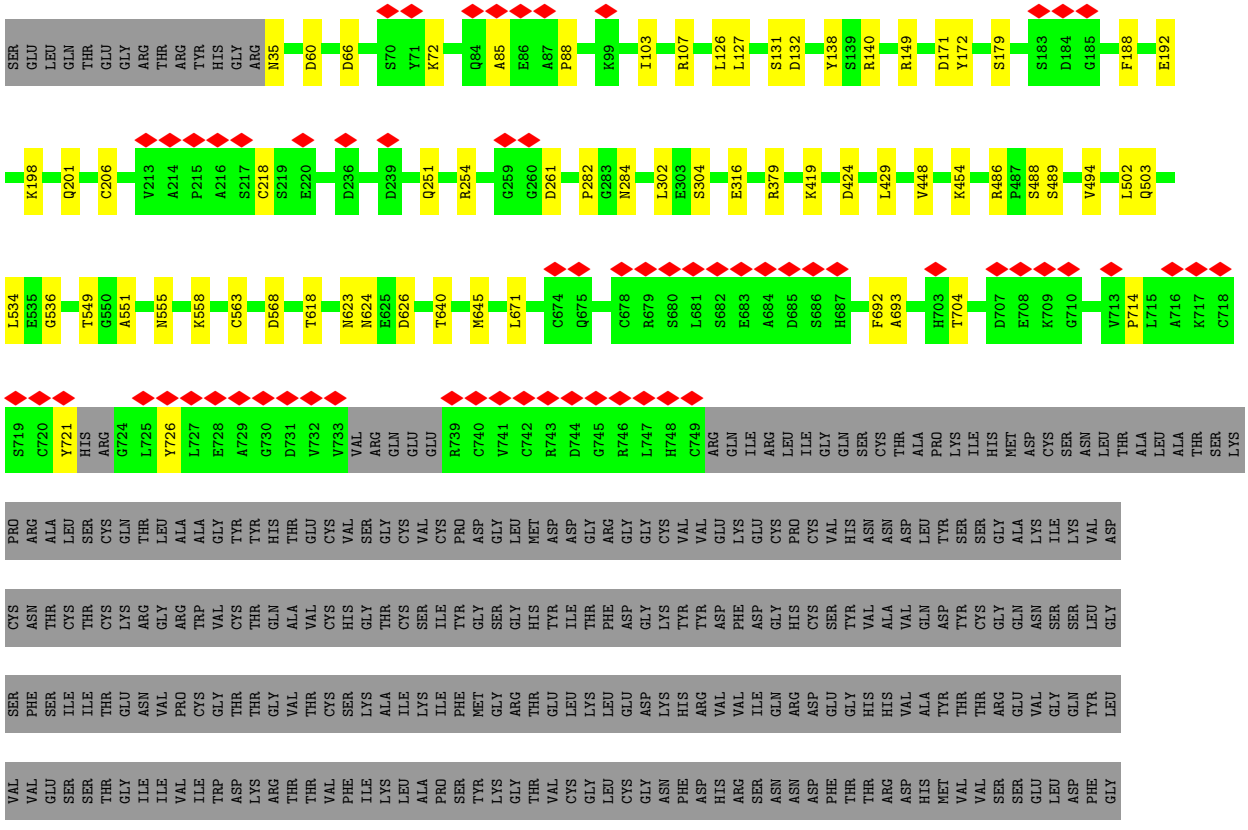


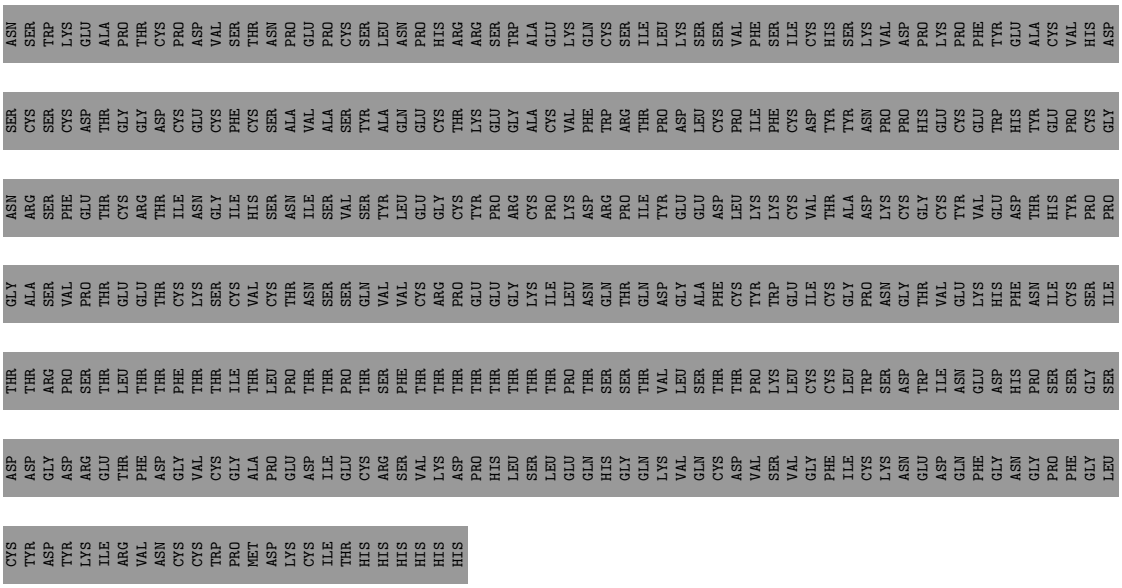


• Molecule 1: Mucin-2

Chain B:

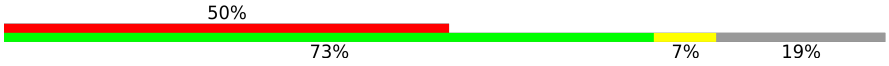


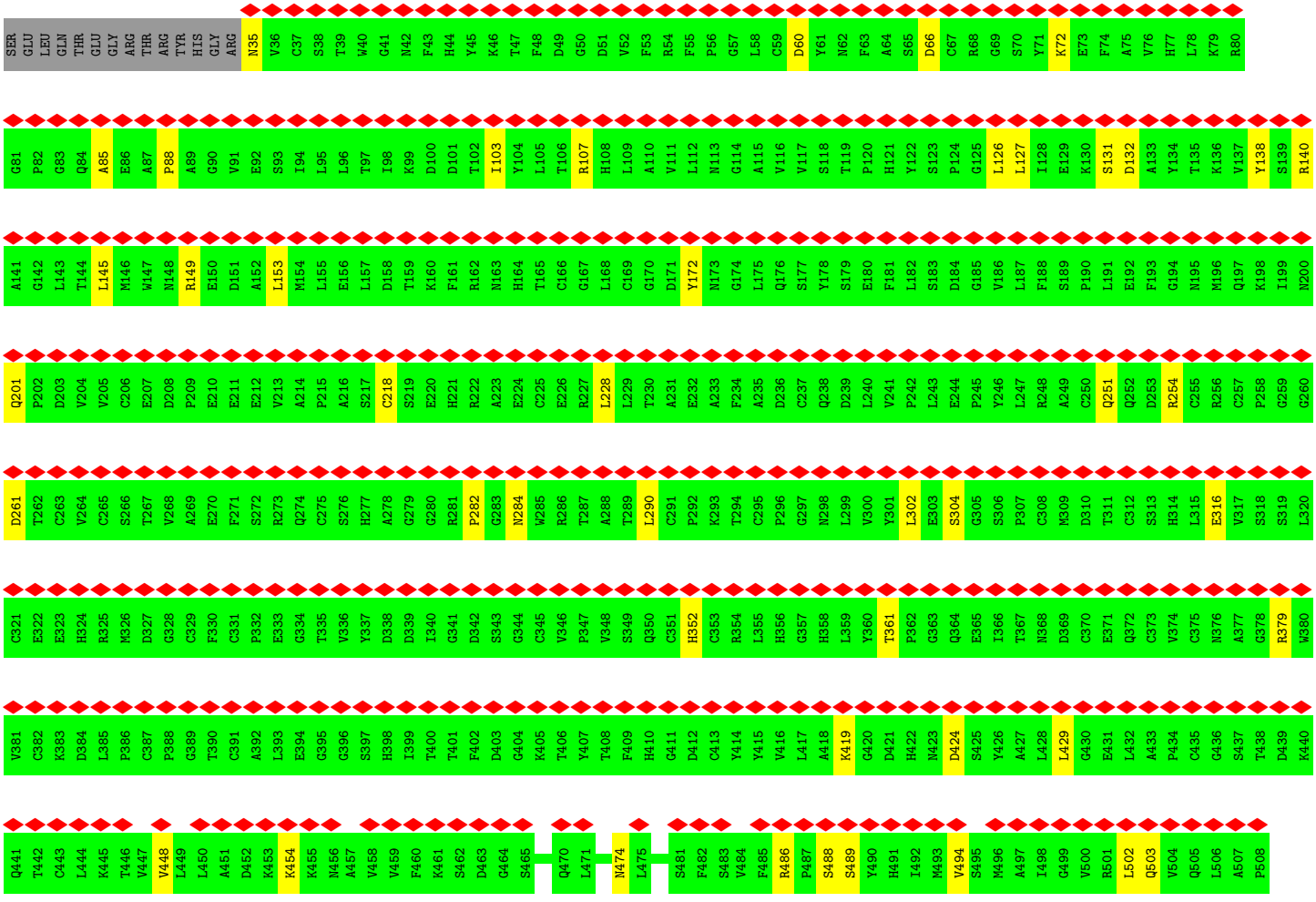


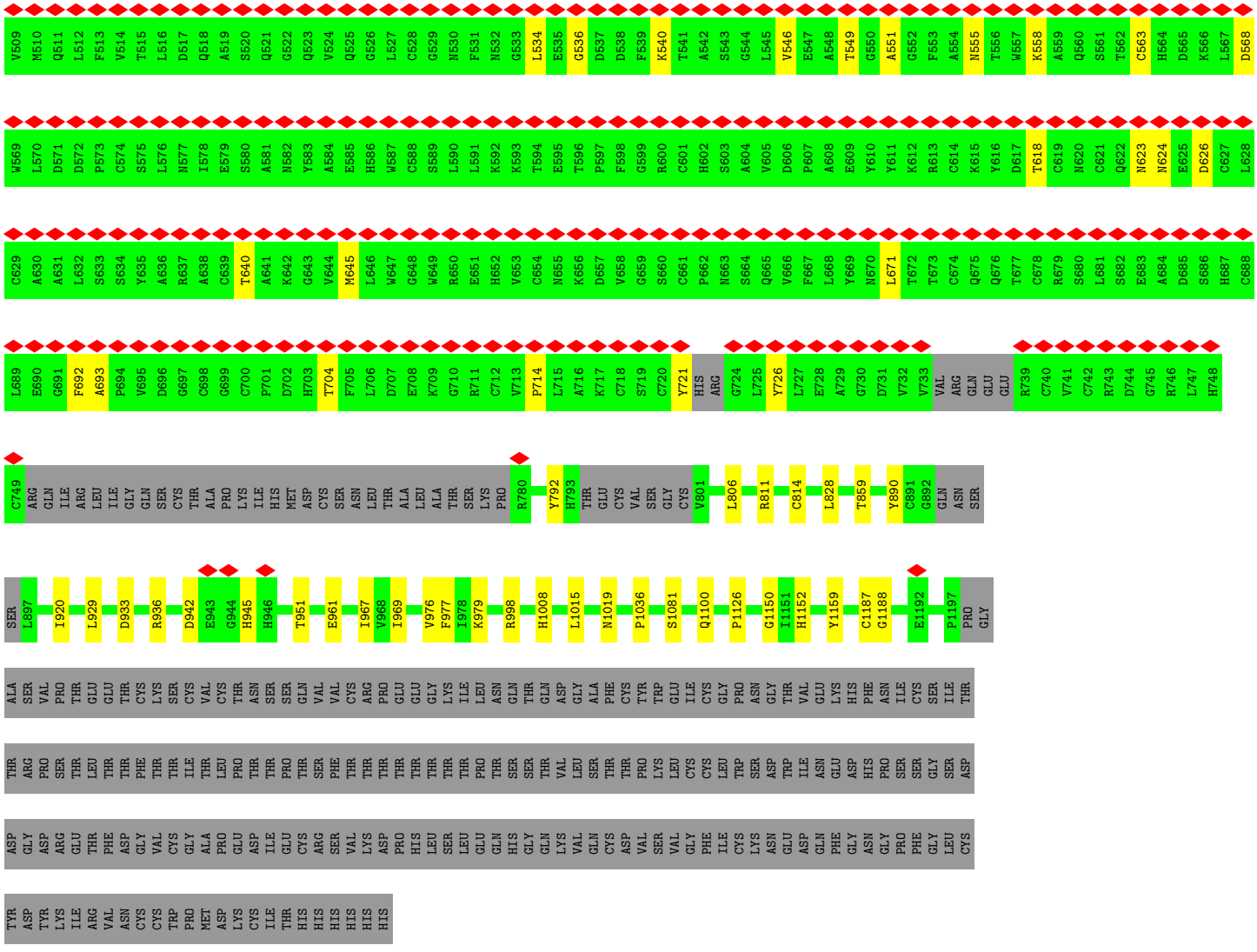


• Molecule 1: Mucin-2

Chain C:

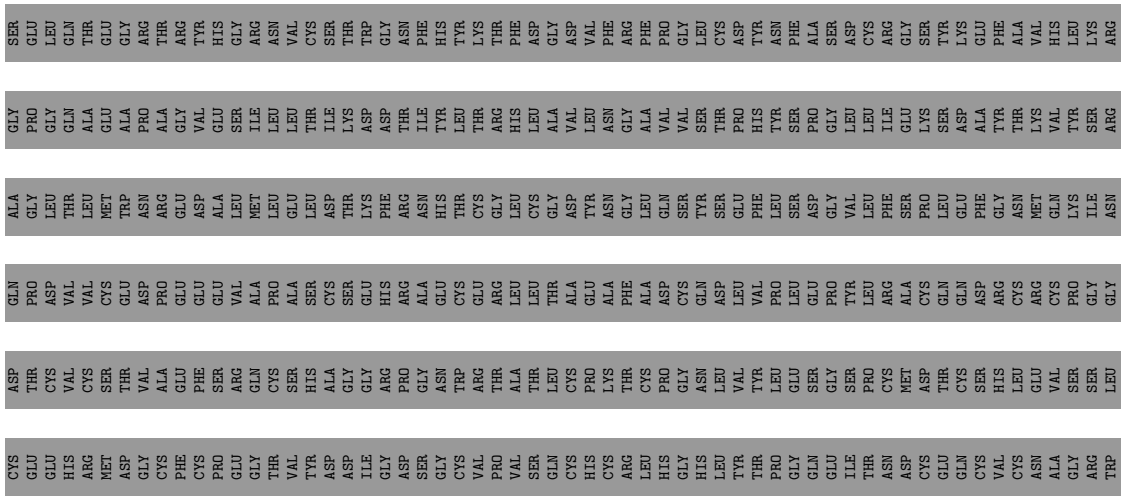


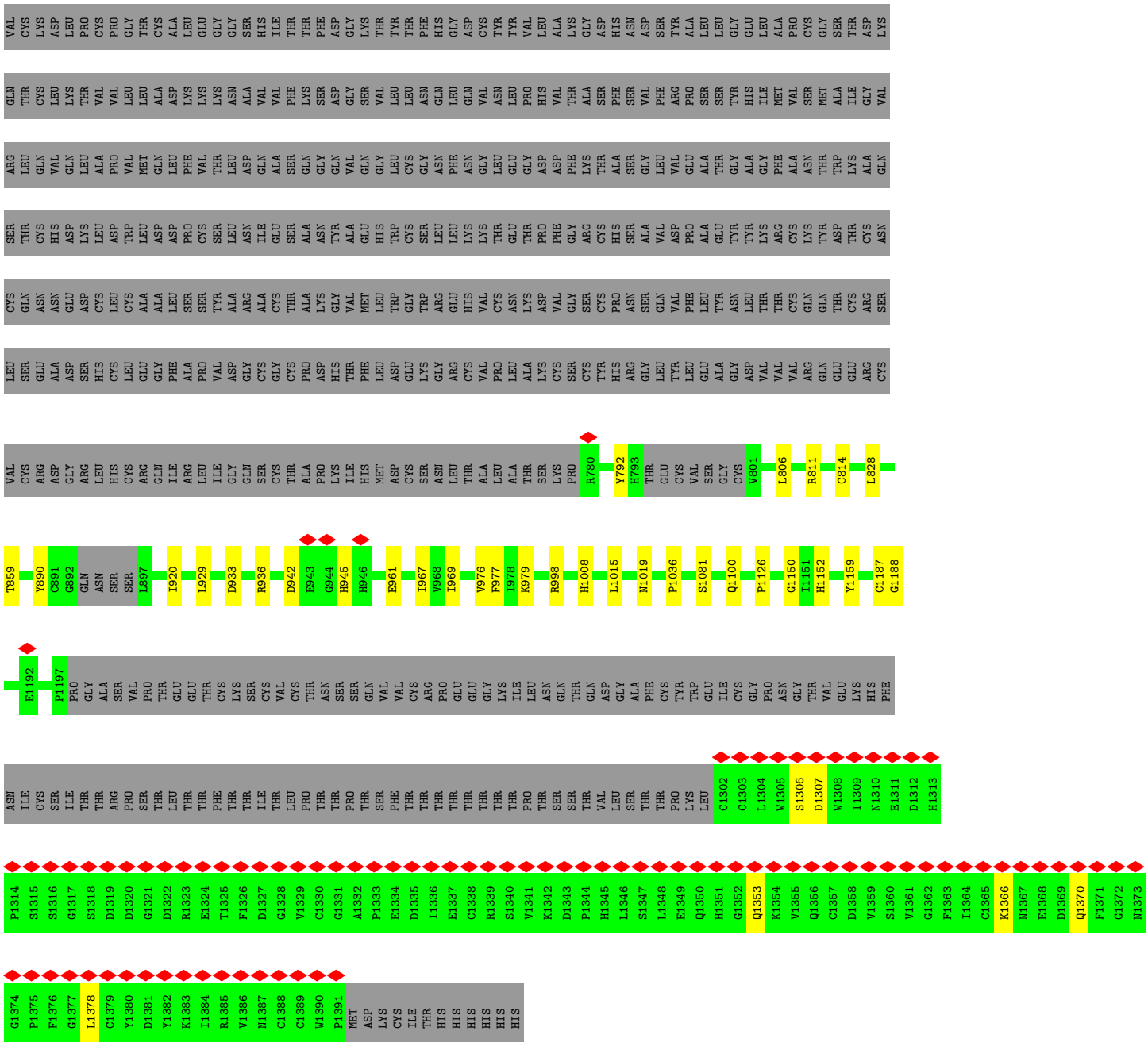




- Molecule 1: Mucin-2

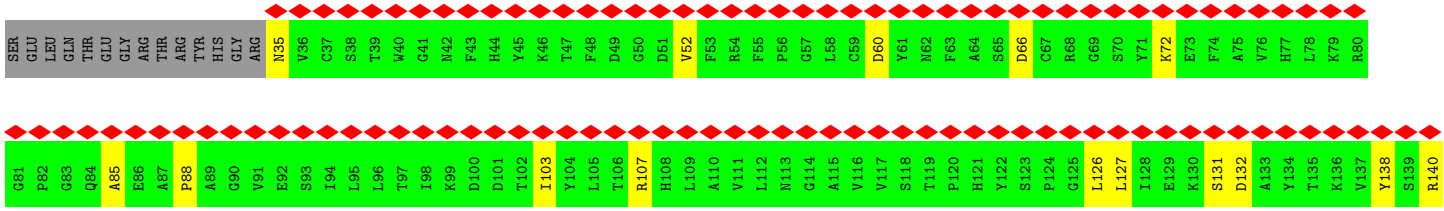
Chain D:

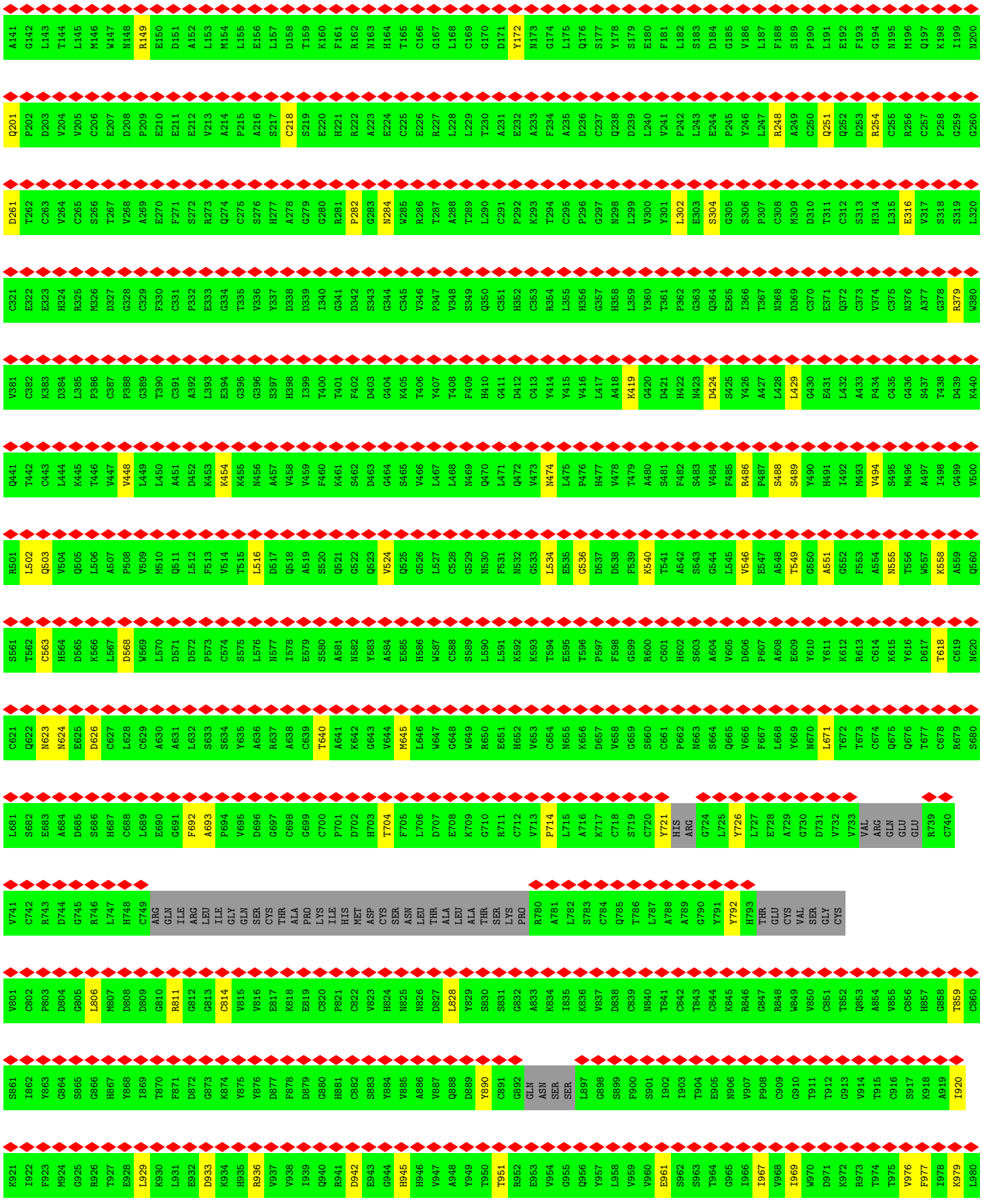


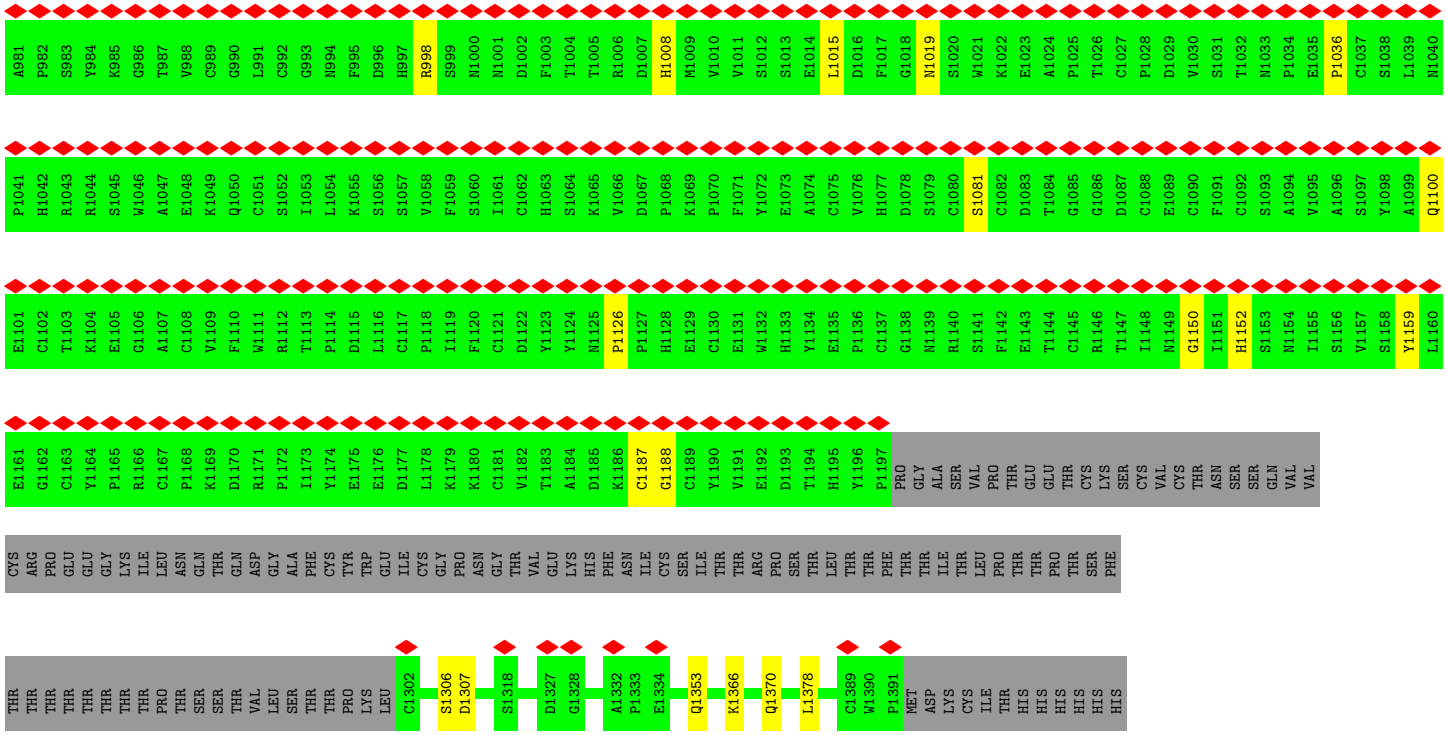


• Molecule 1: Mucin-2

Chain E:

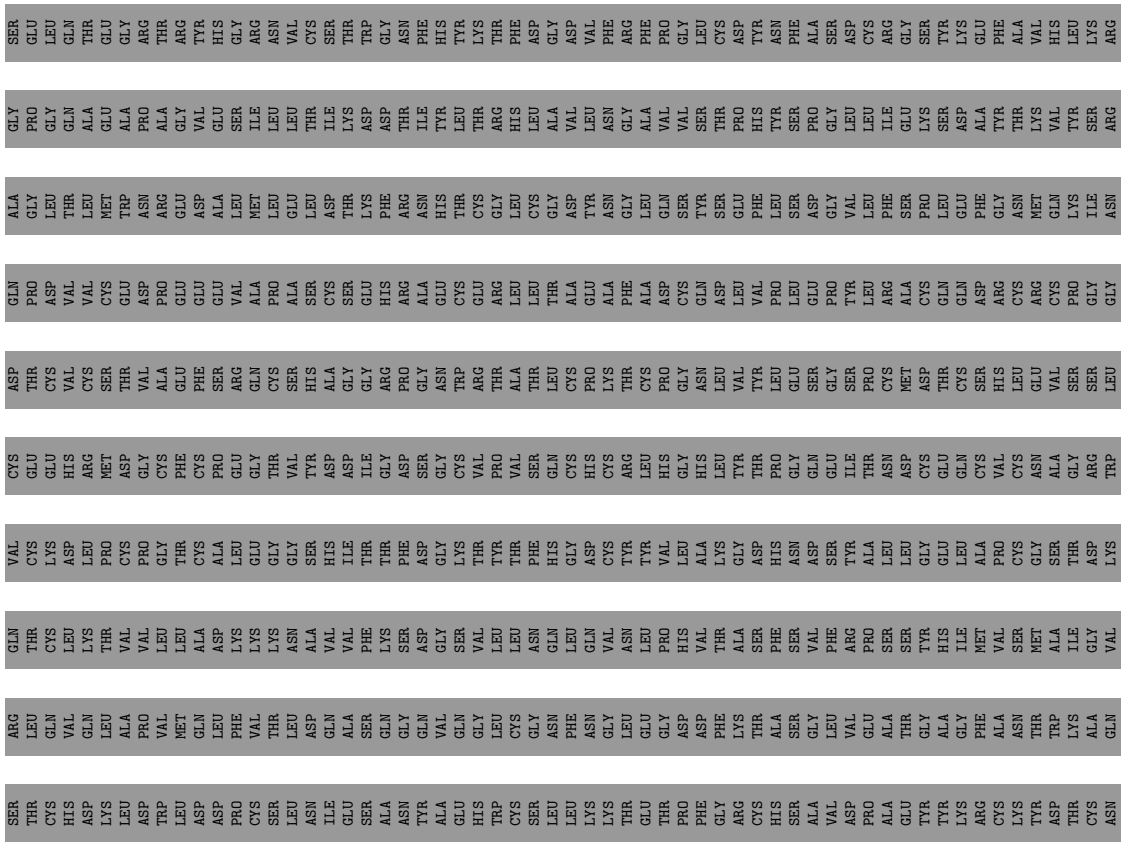


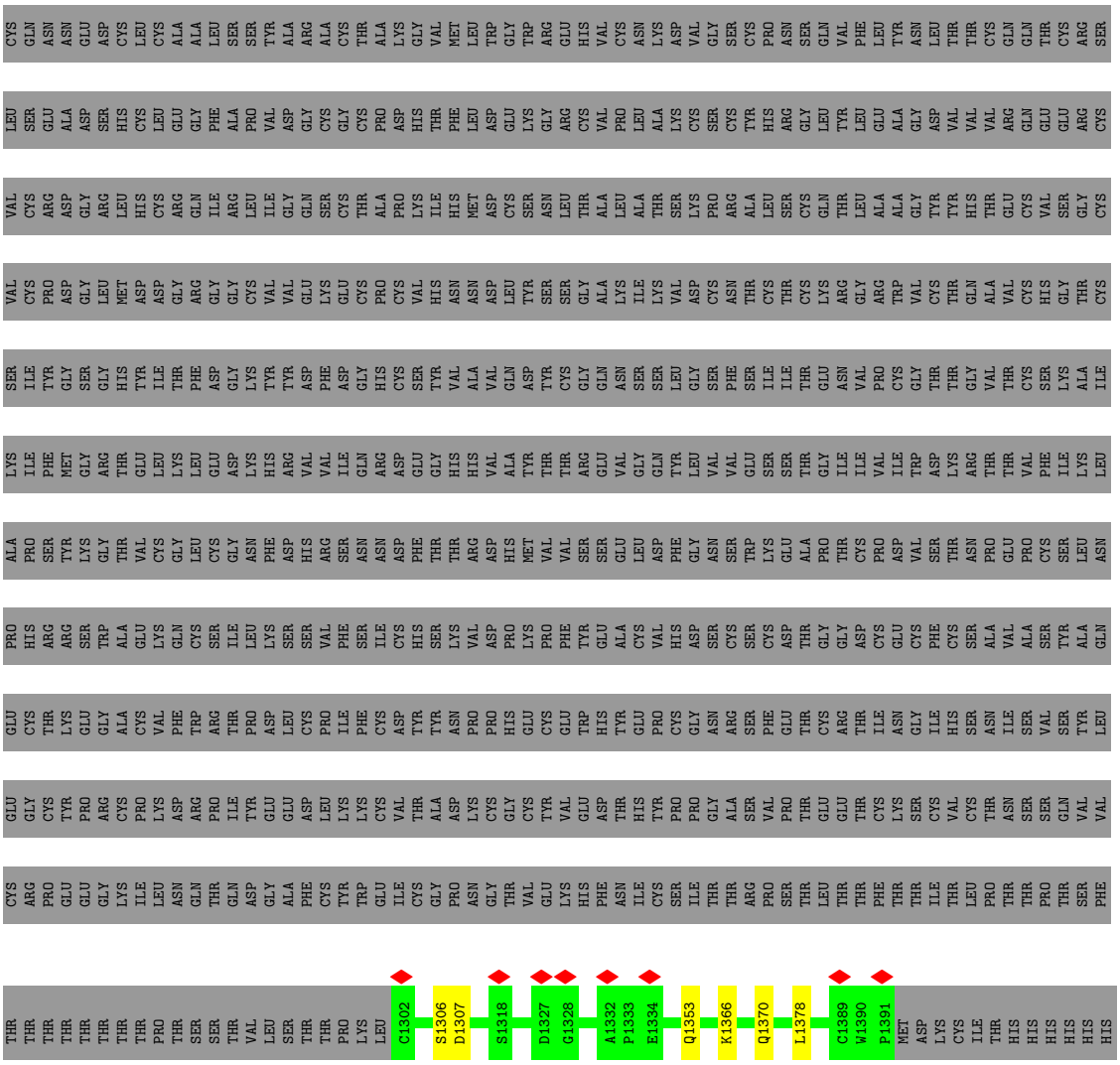

A141 G142 L143 T144 L145 M146 W147 N148 R149 E150 D151 A152 L153 M154 L155 E156 L157 D158 T159 K160 F161 R162 N163 H164 T165 C166 G167 L168 C169 G170 D171 Y172 N173 G174 L175 Q176 S177 Y178 S179 E180 F181 L182 S183 D184 G185 V186 L187 F188 S189 P190 L191 E192 F193 G194 N195 M196 Q197 K198 I199 N200
Q201 P202 D203 V204 V205 C206 E207 D208 P209 E210 E211 E212 V213 A214 P215 A216 S217 C218 S219 E220 H221 R222 A223 E224 C225 E226 R227 L228 L229 T230 A231 E232 A233 F234 A235 D236 C237 Q238 D239 L240 V241 P242 L243 E244 P245 Y246 L247 R248 A249 C250 Q251 Q252 D253 R254 C255 R256 C257 P258 G259 G260
D261 T262 C263 V264 C265 S266 T267 V268 A269 E270 F271 S272 R273 Q274 C275 S276 H277 A278 G279 G280 R281 P282 G283 N284 W285 R286 T287 A288 T289 L290 C291 P292 K293 T294 C295 P296 G297 N298 L299 V300 Y301 L302 E303 S304 G305 S306 P307 C308 M309 D310 T311 C312 S313 H314 L315 E316 V317 S318 S319 L320
C321 E322 E323 H324 R325 M326 D327 G328 C329 F330 C331 P332 E333 G334 T335 V336 Y337 D338 D339 I340 G341 D342 S343 G344 C345 V346 P347 V348 S349 Q350 C351 H352 C353 R354 L355 H356 G357 H358 L359 Y360 T361 P362 G363 Q364 E365 I366 T367 M368 D369 C370 E371 Q372 C373 V374 C375 N376 A377 G378 R379 W380
V381 C382 K383 D384 L385 P386 C387 P388 G389 T390 C391 A392 L393 E394 G395 G396 S397 H398 I399 T400 T401 F402 D403 G404 K405 T406 Y407 T408 F409 H410 G411 D412 C413 Y414 Y415 V416 L417 A418 K419 G420 D421 H422 N423 D424 S425 Y426 A427 L428 L429 G430 E431 L432 A433 P434 C435 G436 S437 T438 D439 K440
Q441 T442 C443 L444 K445 T446 V447 V448 L449 L450 A451 D452 K453 K454 K455 N456 A457 V458 V459 F460 K461 S462 D463 G464 S465 V466 L467 L468 N469 Q470 L471 Q472 V473 N474 L475 P476 H477 V478 T479 A480 S481 F482 S483 V484 F485 R486 P487 S488 S489 Y490 H491 I492 M493 V494 S495 M496 A497 I498 G499 V500
R501 L502 Q503 V504 Q505 L506 A507 P508 V509 H510 Q511 L512 F513 V514 T515 L516 D517 Q518 A519 S520 Q521 G522 Q523 V524 Q525 G526 L527 C528 G529 N530 F531 N532 G533 L534 E535 G536 D537 D538 F539 K540 T541 A542 S543 G544 L545 V546 E547 A548 T549 G550 A551 G552 F553 A554 N555 T556 W557 K558 A559 Q560
S561 T562 C563 H564 D565 K566 L567 D568 W569 L570 D571 D572 P573 C574 S575 L576 N577 I578 E579 S580 A581 N582 Y583 A584 E585 H586 W587 C588 S589 L590 L591 K592 K593 T594 E595 T596 P597 F598 G599 R600 C601 H602 S603 A604 V605 D606 P607 A608 E609 Y610 Y611 K612 R613 C614 K615 Y616 D617 T618 C619 N620
C621 Q622 N623 N624 E625 D626 C627 L628 C629 A630 A631 L632 S633 S634 Y635 A636 R637 A638 C639 T640 A641 K642 G643 V644 M645 L646 W647 G648 W649 R650 E651 H652 V653 C654 N655 K656 D657 V658 G659 S660 C661 P662 N663 S664 Q665 V666 F667 L668 Y669 N670 L671 T672 T673 C674 Q675 Q676 T677 C678 R679 S680
L681 S682 E683 A684 D685 S686 H687 C688 L689 E690 G691 F692 A693 P694 V695 D696 G697 C698 G699 C700 P701 D702 H703 T704 F705 L706 D707 E708 K709 G710 R711 C712 V713 P714 L715 A716 K717 C718 S719 C720 Y721 HIS ARG G724 L725 Y726 L727 E728 A729 G730 D731 V732 V733 VAL ARG GLN GLU GLU R739 C740
V741 C742 R743 D744 G745 R746 L747 H748 C749 ARG GLN ILE ARG LEU ILE GLY GLN SER CYS THR ALA PRO LYS ILE HIS MET ASP CYS SER ASN LEU THR ALA LEU ALA THR SER LYS PRO R780 A781 L782 S783 C784 Q785 T786 L787 A788 A789 G790 Y791 Y792 H793 THR GLU CYS VAL SER GLY CYS
V801 C802 P803 D804 G805 L806 M807 D808 D809 G810 R811 G812 G813 C814 V815 V816 E817 K818 E819 C820 P821 C822 V823 H824 N825 N826 D827 L828 Y829 S830 S831 G832 A833 K834 I835 K836 V837 D838 C839 N840 T841 C842 T843 C844 K845 R846 G847 R848 W849 V850 C851 T852 Q853 A854 V855 C856 H857 G858 T859 C860
S861 I862 Y863 G864 S865 G866 H867 Y868 I869 T870 F871 D872 G873 K874 Y875 Y876 D877 F878 D879 G880 H881 C882 S883 Y884 V885 A886 V887 Q888 D889 Y890 C891 G892 GLN ASN SER SER L897 G898 S899 F900 S901 I902 I903 T904 E905 N906 V907 P908 C909 G910 T911 T912 G913 V914 T915 C916 S917 K918 A919 I920
K921 I922 F923 M924 G925 R926 T927 E928 L929 K930 L931 E932 D933 K934 H935 R936 V937 V938 I939 Q940 R941 D942 E943 G944 H945 H946 V947 A948 Y949 T950 T951 R952 E953 V954 G955 Q956 Y957 L958 V959 V960 E961 S962 S963 T964 G965 I966 I967 V968 I969 W970 D971 K972 R973 T974 T975 V976 F977 I978 K979 L980



- Molecule 1: Mucin-2

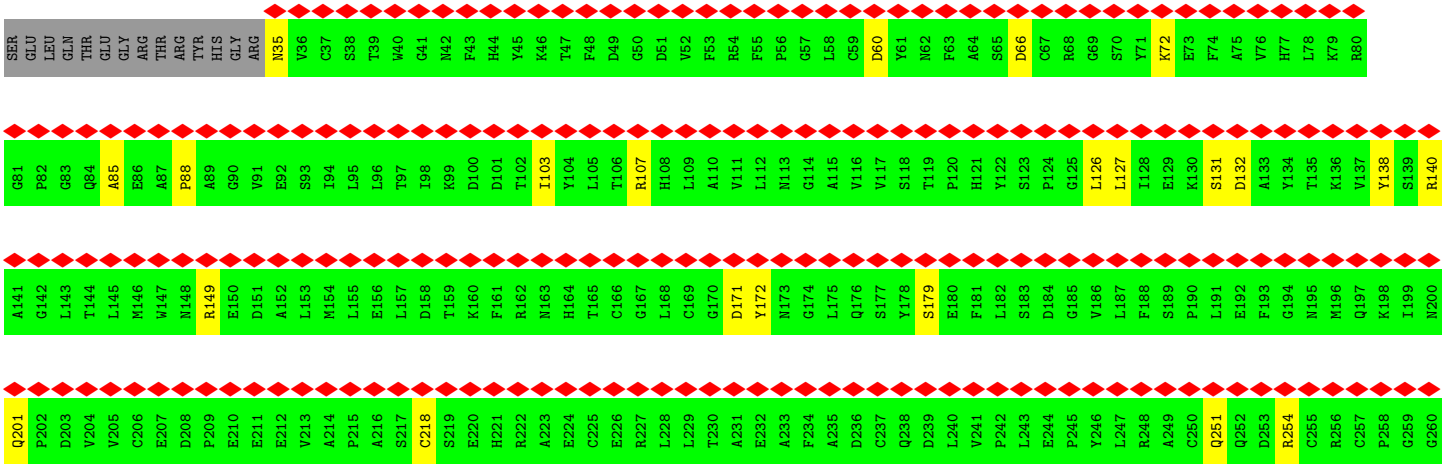
Chain F: 6% 93%

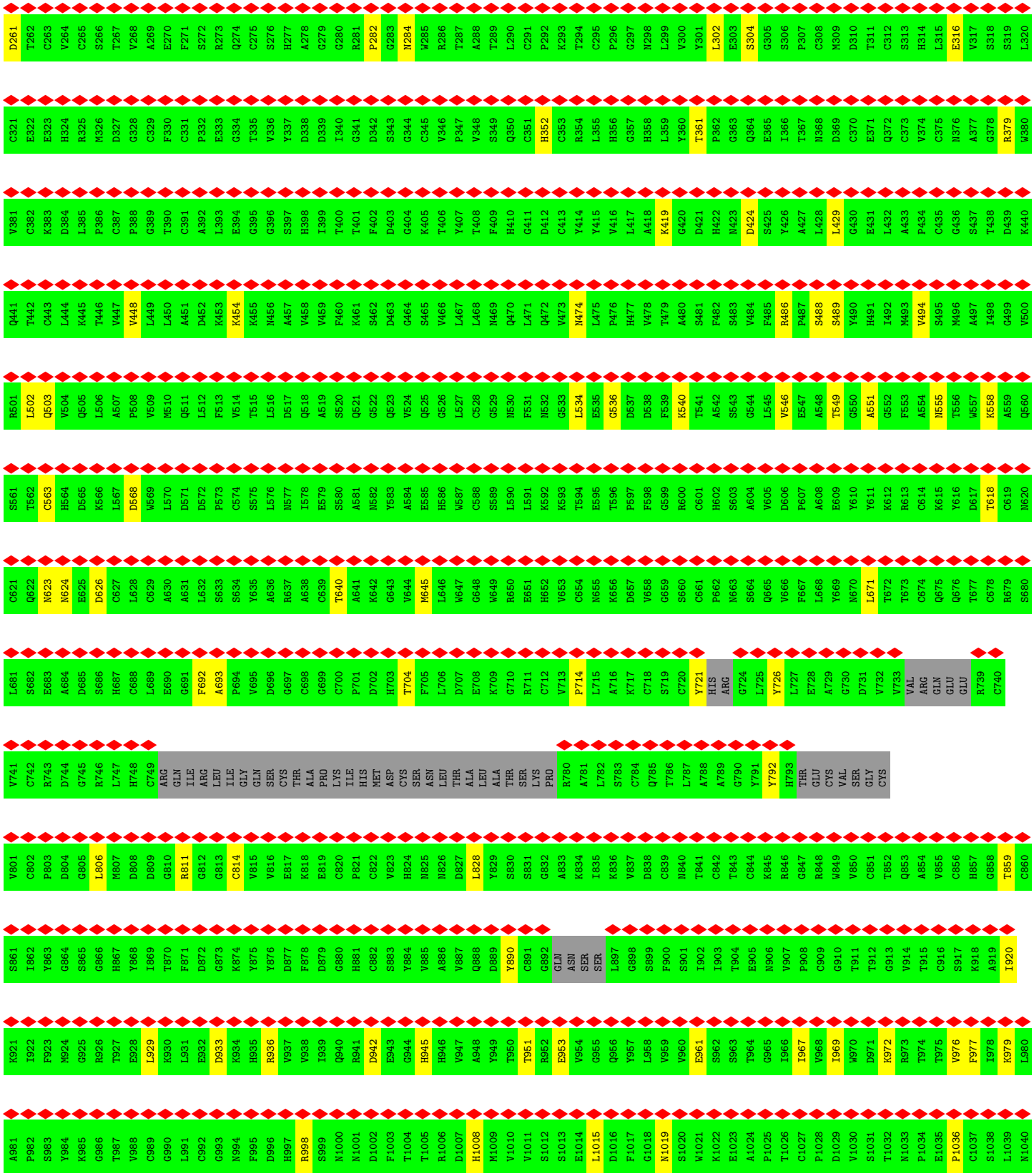


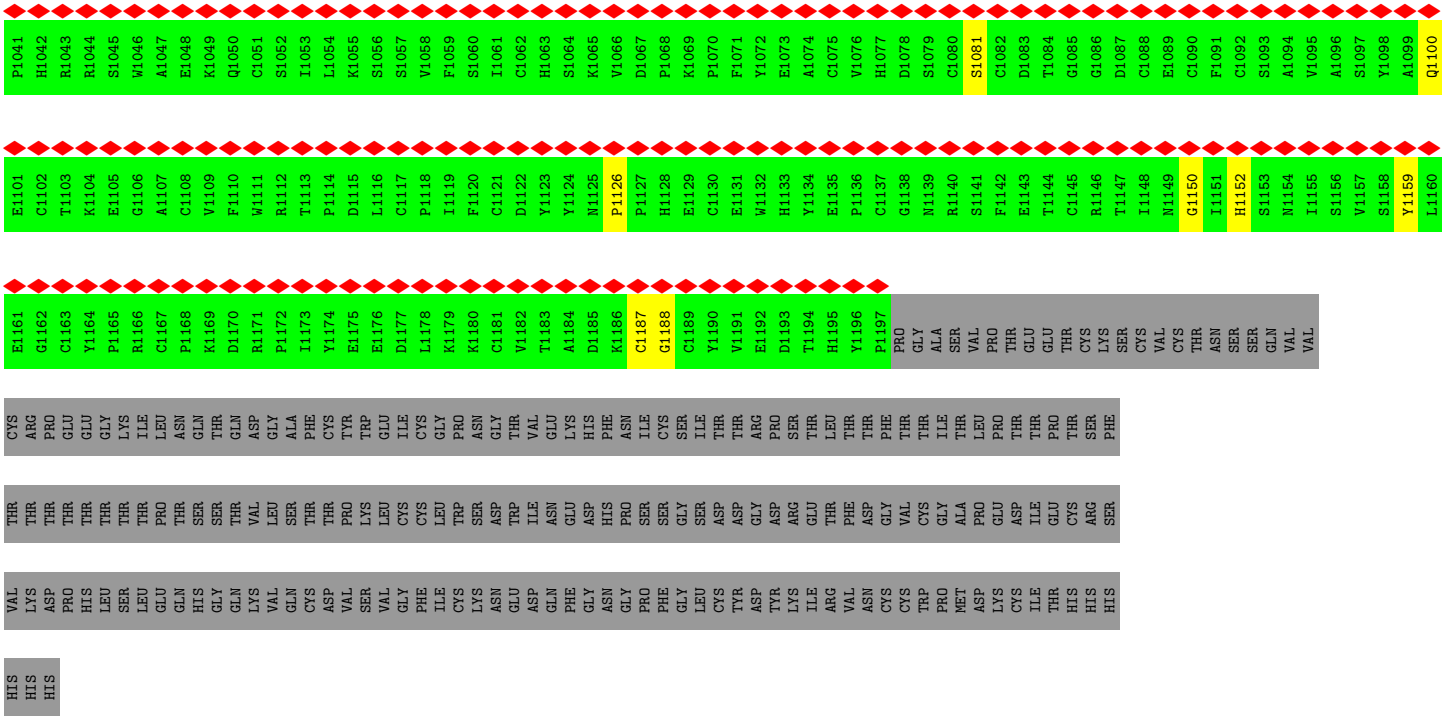


• Molecule 1: Mucin-2

Chain G:

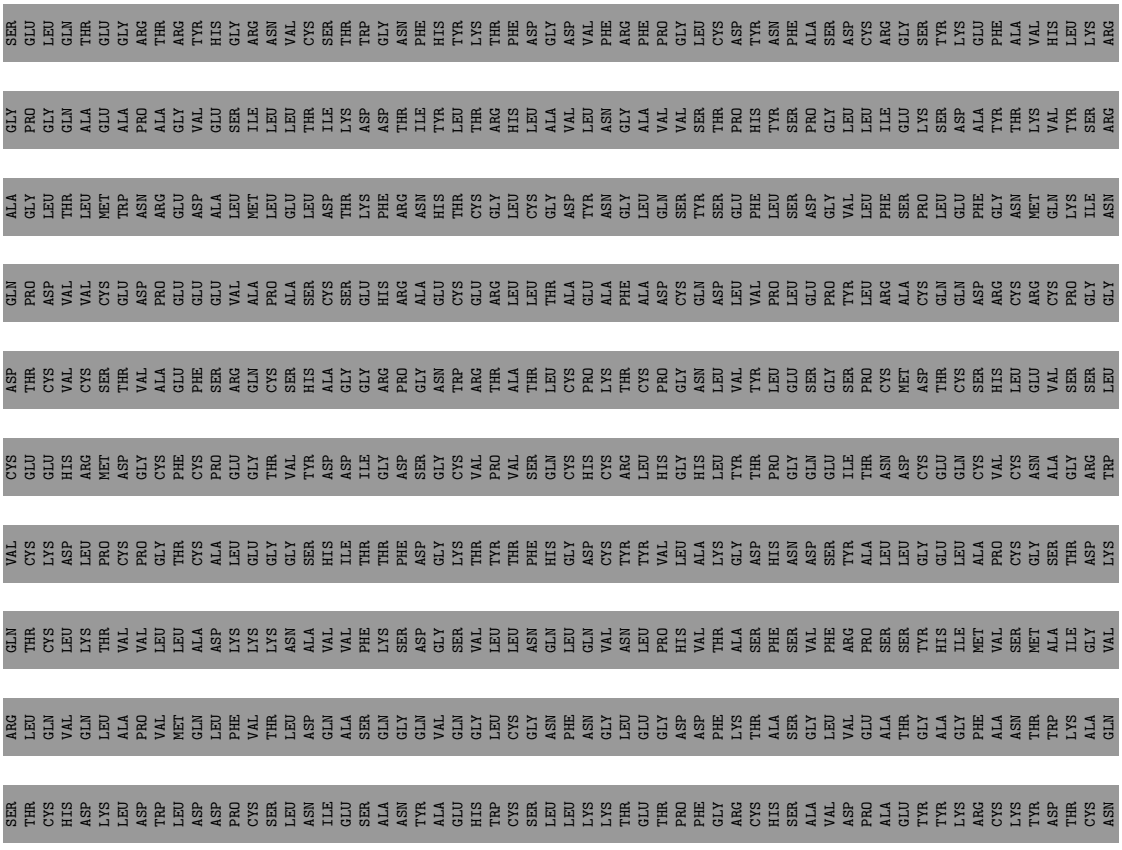


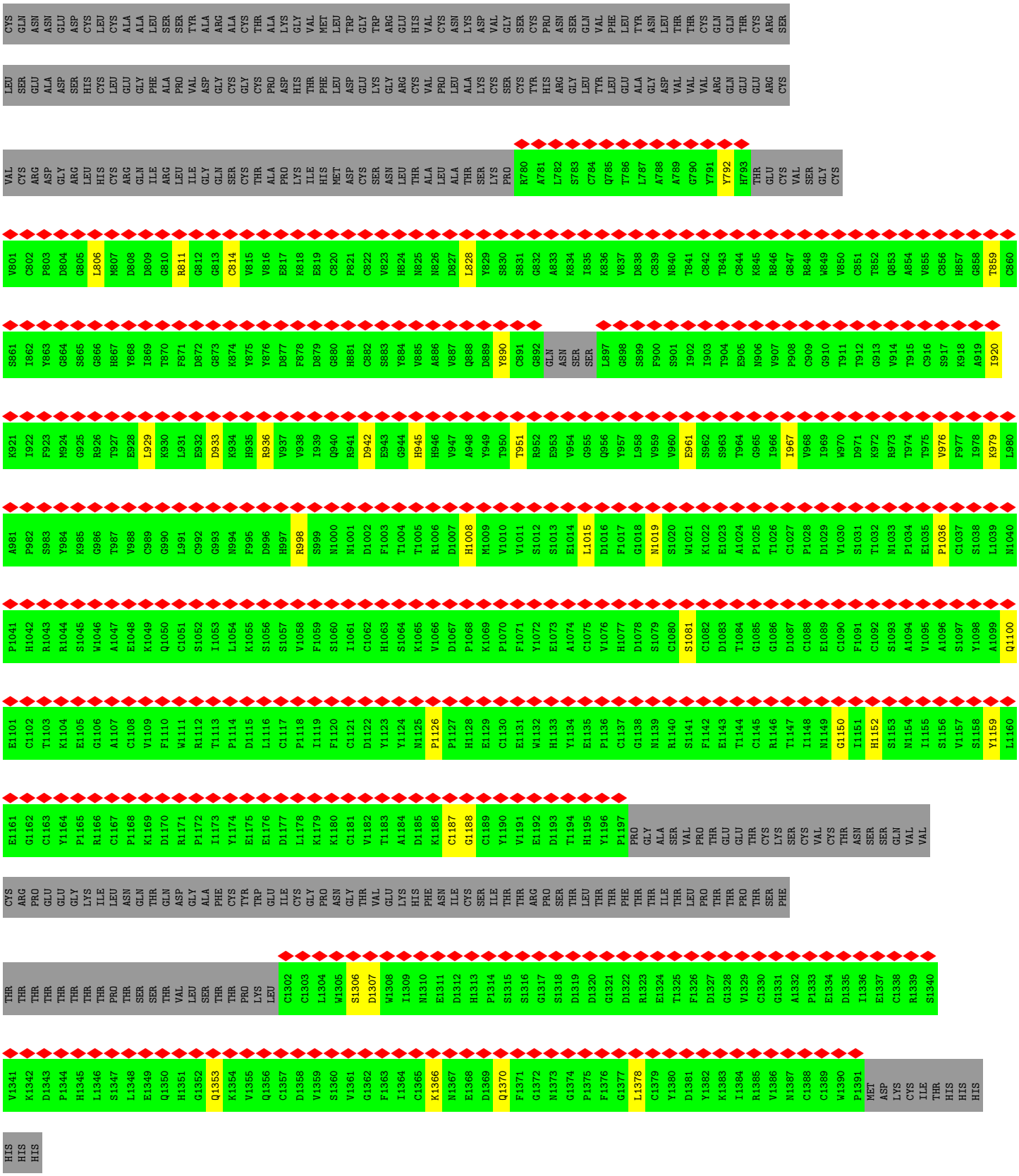

D261 T262 C263 V264 C265 S266 T267 V268 A269 E270 F271 S272 R273 Q274 C275 S276 H277 A278 G279 G280 R281 P282 G283 N284 W285 R286 T287 A288 T289 L290 C291 P292 K293 T294 C295 P296 G297 N298 L299 V300 Y301 L302 E303 S304 G305 S306 P307 C308 M309 D310 T311 C312 S313 H314 L315 E316 V317 S318 S319 L320
C321 E322 E323 H324 R325 M326 D327 G328 C329 F330 C331 P332 E333 G334 T335 V336 Y337 D338 D339 I340 G341 D342 S343 G344 C345 V346 P347 V348 S349 Q350 C351 H352 C353 R354 L355 H356 G357 H358 L359 Y360 T361 P362 G363 Q364 E365 I366 T367 N368 D369 C370 E371 Q372 C373 V374 C375 N376 A377 G378 R379 W380
V381 C382 K383 D384 L385 P386 C387 P388 G389 T390 C391 A392 L393 E394 G395 G396 S397 H398 I399 T400 T401 F402 D403 G404 K405 T406 Y407 T408 F409 H410 G411 D412 C413 Y414 Y415 V416 L417 A418 K419 G420 D421 H422 N423 D424 S425 Y426 A427 L428 L429 G430 E431 L432 A433 P434 C435 G436 S437 T438 D439 K440
Q441 T442 C443 L444 K445 T446 V447 V448 L449 L450 A451 D452 K453 K454 K455 N456 A457 V458 V459 F460 K461 S462 D463 G464 S465 V466 L467 L468 N469 Q470 G471 Q472 V473 N474 L475 P476 H477 V478 T479 A480 S481 F482 S483 V484 F485 R486 P487 S488 S489 Y490 H491 I492 M493 V494 S495 M496 A497 I498 G499 V500
R501 L502 Q503 V504 Q505 L506 A507 P508 V509 M510 Q511 L512 F513 V514 T515 L516 D517 Q518 A519 S520 Q521 G522 Q523 V524 Q525 G526 L527 C528 G529 N530 F531 N532 G533 L534 E535 G536 D537 D538 F539 K540 T541 A542 S543 G544 L545 V546 E547 A548 T549 G550 A551 G552 F553 A554 N555 T556 W557 K558 A559 Q560
S561 T562 C563 H564 D565 K566 L567 D568 W569 L570 D571 D572 P573 C574 S575 L576 N577 I578 E579 S580 A581 N582 Y583 A584 E585 H586 W587 C588 S589 L590 L591 K592 K593 T594 E595 T596 P597 F598 G599 R600 C601 H602 S603 A604 V605 D606 P607 A608 E609 Y610 Y611 K612 R613 C614 K615 Y616 D617 T618 C619 N620
C621 Q622 N623 N624 E625 D626 C627 L628 C629 A630 A631 L632 S633 S634 Y635 A636 R637 A638 C639 T640 A641 K642 G643 V644 M645 L646 W647 G648 W649 R650 E651 H652 V653 C654 N655 K656 D657 V658 G659 S660 C661 P662 N663 S664 Q665 V666 F667 L668 Y669 N670 L671 T672 T673 C674 Q675 Q676 T677 C678 R679 S680
L681 S682 E683 A684 D685 S686 H687 C688 L689 E690 G691 F692 A693 P694 V695 D696 G697 C698 G699 C700 P701 D702 H703 T704 F705 L706 D707 E708 K709 G710 R711 C712 V713 P714 L715 A716 K717 C718 S719 C720 Y721 HIS ARG G724 L725 Y726 L727 E728 A729 G730 D731 V732 V733 VAL ARG GLN GLU GLU R739 C740
V741 C742 R743 D744 G745 R746 L747 H748 C749 ARG GLN ILE ARG LEU ILE GLY GLN SER CYS THR ALA PRO LYS ILE HIS MET ASP CYS SER ASN LEU THR ALA LEU ALA THR SER LYS PRO R780 A781 L782 S783 C784 Q785 T786 L787 A788 A789 G790 Y791 Y792 H793 THR GLU CYS VAL SER GLY CYS
V801 C802 P803 D804 G805 L806 M807 D808 D809 G810 R811 G812 G813 C814 V815 V816 E817 K818 E819 C820 P821 C822 V823 H824 N825 N826 D827 L828 Y829 S830 S831 G832 A833 K834 I835 L782 K836 V837 D838 C839 N840 T841 C842 T843 C844 K845 R846 G847 R848 W849 V850 C851 T852 Q853 A854 V855 C856 H857 G858 T859 C860
S861 I862 Y863 G864 S865 G866 H867 Y868 I869 T870 F871 D872 G873 K874 Y875 Y876 D877 F878 D879 G880 H881 C882 S883 Y884 V885 A886 V887 Q888 D889 Y890 C891 G892 GLN ASN SER SER L897 G898 S899 F900 S901 I902 I903 T904 E905 N906 V907 P908 C909 G910 T911 T912 G913 V914 T915 C916 S917 K918 A919 I920
K921 I922 F923 M924 G925 R926 T927 E928 L929 K930 L931 E932 D933 K934 H935 R936 V937 V938 I939 Q940 R941 D942 E943 G944 H945 H946 V947 A948 Y949 T950 T951 R952 E953 V954 G955 Q956 Y957 L958 V959 V960 E961 S962 S963 T964 G965 I966 I967 V968 I969 W970 D971 K972 R973 T974 T975 V976 F977 I978 K979 L980
A981 P982 S983 Y984 K985 G986 T987 V988 C989 G990 L991 C992 G993 N994 F995 D996 H997 R998 S999 N1000 N1001 D1002 F1003 T1004 T1005 R1006 D1007 H1008 M1009 V1010 V1011 S1012 S1013 E1014 L1015 D1016 F1017 G1018 N1019 S1020 W1021 K1022 E1023 A1024 P1025 T1026 C1027 P1028 D1029 V1030 S1031 T1032 N1033 P1034 E1035 P1036 C1037 S1038 L1039 N1040



- Molecule 1: Mucin-2

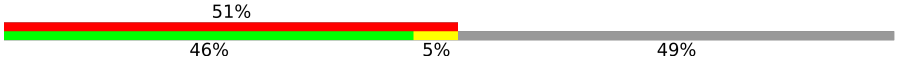
Chain H:

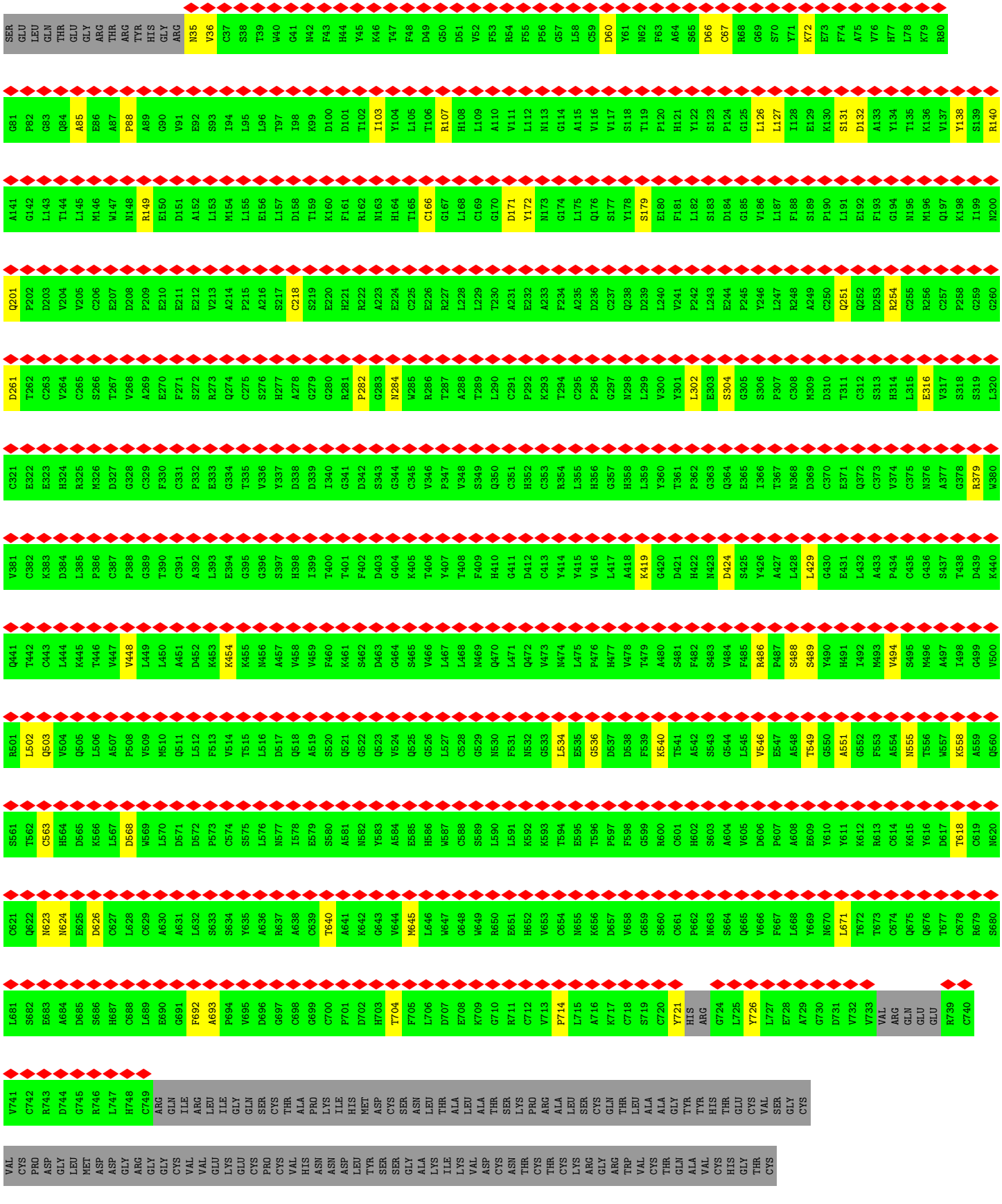




- Molecule 1: Mucin-2

Chain I:




SER GLU LEU GLN THR GLU GLY ARG THR ARG TYR HIS GLY ARG N35 V36 C37 S38 T39 W40 G41 N42 F43 H44 Y45 K46 T47 F48 D49 G50 D51 V52 F53 R54 F55 P56 G57 L58 C59 D60 Y61 N62 F63 A64 S65 D66 C67 R68 G69 S70 Y71 K72 E73 F74 A75 V76 H77 L78 K79 R80
G81 P82 G83 Q84 A85 E86 A87 P88 A89 G90 V91 E92 S93 I94 L95 L96 T97 I98 K99 D100 D101 T102 I103 Y104 L105 T106 R107 H108 L109 A110 V111 L112 N113 G114 A115 V116 V117 S118 T119 P120 H121 Y122 S123 P124 G125 L126 L127 I128 E129 K130 S131 D132 A133 Y134 T135 K136 V137 Y138 S139 R140
A141 G142 L143 T144 L145 M146 W147 N148 R149 E150 D151 A152 L153 M154 L155 E156 L157 D158 T159 K160 F161 R162 N163 H164 T165 C166 G167 L168 C169 G170 D171 Y172 N173 G174 L175 Q176 S177 Y178 S179 E180 F181 L182 S183 D184 G185 V186 L187 F188 S189 P190 L191 E192 F193 G194 N195 M196 Q197 K198 I199 N200
Q201 P202 D203 V204 V205 C206 E207 D208 P209 E210 E211 E212 V213 A214 P215 A216 S217 C218 S219 E220 H221 R222 A223 E224 C225 E226 R227 L228 L229 T230 A231 E232 A233 F234 A235 D236 C237 Q238 D239 L240 V241 P242 L243 E244 P245 Y246 L247 R248 A249 C250 Q251 Q252 D253 R254 C255 R256 C257 P258 G259 G260
D261 T262 C263 V264 C265 S266 T267 V268 A269 E270 F271 S272 R273 Q274 C275 S276 H277 A278 G279 G280 R281 P282 G283 N284 W285 R286 T287 A288 T289 L290 C291 P292 K293 T294 C295 P296 G297 N298 L299 V300 Y301 L302 E303 S304 G305 S306 P307 C308 M309 D310 T311 C312 S313 H314 L315 E316 V317 S318 S319 L320
C321 E322 E323 H324 R325 M326 D327 G328 C329 F330 C331 P332 E333 G334 T335 V336 Y337 D338 D339 I340 G341 D342 S343 G344 C345 V346 P347 V348 S349 Q350 C351 H352 C353 R354 L355 H356 G357 H358 L359 Y360 T361 P362 G363 Q364 E365 I366 T367 N368 D369 C370 E371 Q372 C373 V374 C375 N376 A377 G378 R379 W380
V381 C382 K383 D384 L385 P386 C387 P388 G389 T390 C391 A392 L393 E394 G395 G396 S397 H398 I399 T400 T401 F402 D403 G404 K405 T406 Y407 T408 F409 H410 G411 D412 C413 Y414 Y415 V416 L417 A418 K419 G420 D421 H422 N423 D424 S425 Y426 A427 L428 L429 G430 E431 L432 A433 P434 C435 G436 S437 T438 D439 K440
Q441 T442 C443 L444 K445 T446 V447 V448 L449 L450 A451 D452 K453 K454 K455 N456 A457 V458 V459 F460 K461 S462 D463 G464 S465 V466 L467 L468 N469 Q470 L471 Q472 V473 N474 L475 P476 H477 V478 T479 A480 S481 F482 S483 V484 F485 R486 P487 S488 S489 Y490 H491 I492 M493 V494 S495 M496 A497 I498 G499 V500
R501 L502 Q503 V504 Q505 L506 A507 P508 V509 M510 Q511 L512 F513 V514 T515 L516 D517 Q518 A519 S520 Q521 G522 Q523 V524 Q525 G526 L527 C528 G529 N530 F531 N532 G533 L534 E535 G536 D537 D538 F539 K540 T541 A542 S543 G544 L545 V546 E547 A548 T549 G550 A551 G552 F553 A554 N555 T556 W557 K558 A559 Q560
S561 T562 C563 H564 D565 K566 L567 D568 W569 L570 D571 D572 P573 C574 S575 L576 N577 I578 E579 S580 A581 N582 Y583 A584 E585 H586 W587 C588 S589 L590 L591 K592 K593 T594 E595 T596 P597 F598 G599 R600 C601 H602 S603 A604 V605 D606 P607 A608 E609 Y610 Y611 K612 R613 C614 K615 Y616 D617 T618 C619 N620
C621 Q622 N623 N624 E625 D626 C627 L628 C629 A630 A631 L632 S633 S634 Y635 A636 R637 A638 C639 T640 A641 K642 G643 V644 M645 L646 W647 G648 W649 R650 E651 H652 V653 C654 N655 K656 D657 V658 G659 S660 C661 P662 N663 S664 Q665 V666 F667 L668 Y669 N670 L671 T672 T673 C674 Q675 Q676 T677 C678 R679 S680
L681 S682 E683 A684 D685 S686 H687 C688 L689 E690 G691 F692 A693 P694 V695 D696 G697 C698 G699 C700 P701 D702 H703 T704 F705 L706 D707 E708 K709 G710 R711 C712 V713 P714 L715 A716 K717 C718 S719 C720 Y721 HIS ARG G724 L725 Y726 L727 E728 A729 G730 D731 V732 V733 VAL ARG GLN GLU GLU R739 C740
V741 C742 R743 D744 G745 R746 L747 H748 C749 ARG GLN ILE ARG LEU ILE GLY GLN SER CYS THR ALA PRO LYS ILE HIS MET ASP CYS SER ASN LEU THR ALA LEU ALA THR SER LYS PRO ARG ALA LEU SER CYS GLN THR LEU ALA ALA GLY TYR TYR HIS THR GLU CYS VAL SER GLY CYS
VAL CYS PRO ASP GLY LEU MET ASP ASP GLY ARG GLY GLY CYS VAL VAL GLU LYS GLU CYS PRO CYS VAL HIS ASN ASN ASP LEU TYR SER SER GLY ALA LYS ILE LYS VAL ASP CYS ASN THR CYS THR CYS LYS ARG GLY ARG TRP VAL CYS THR GLN ALA VAL CYS HIS GLY THR CYS

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

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

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

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

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

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

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

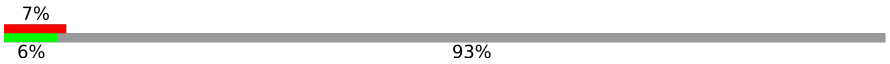
THR THR THR THR THR THR THR PRO THR SER SER THR VAL LEU SER THR THR PRO LYS LEU CYS CYS LEU TRP SER ASP TRP ILE ASN GLU ASP HIS PRO SER SER GLY SER ASP ASP GLY ASP ARG GLU THR PHE ASP GLY VAL CYS GLY VAL PHE THR THR ILE LEU PRO THR PRO THR SER PHE

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

HIS HIS HIS

- Molecule 1: Mucin-2

Chain J:



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

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

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

GLN PRO ASP VAL VAL CYS GLU ASP PRO GLU GLU GLU VAL ALA PRO ALA SER CYS SER GLU HIS ARG ALA GLU CYS GLU ARG LEU LEU THR ALA GLU GLU ALA PHE ALA ASP CYS GLN ASP LEU VAL PRO LEU GLU GLU PRO TYR LEU LEU ARG ALA CYS GLN GLN ASP ARG CYS ARG CYS PRO GLY GLY GLY

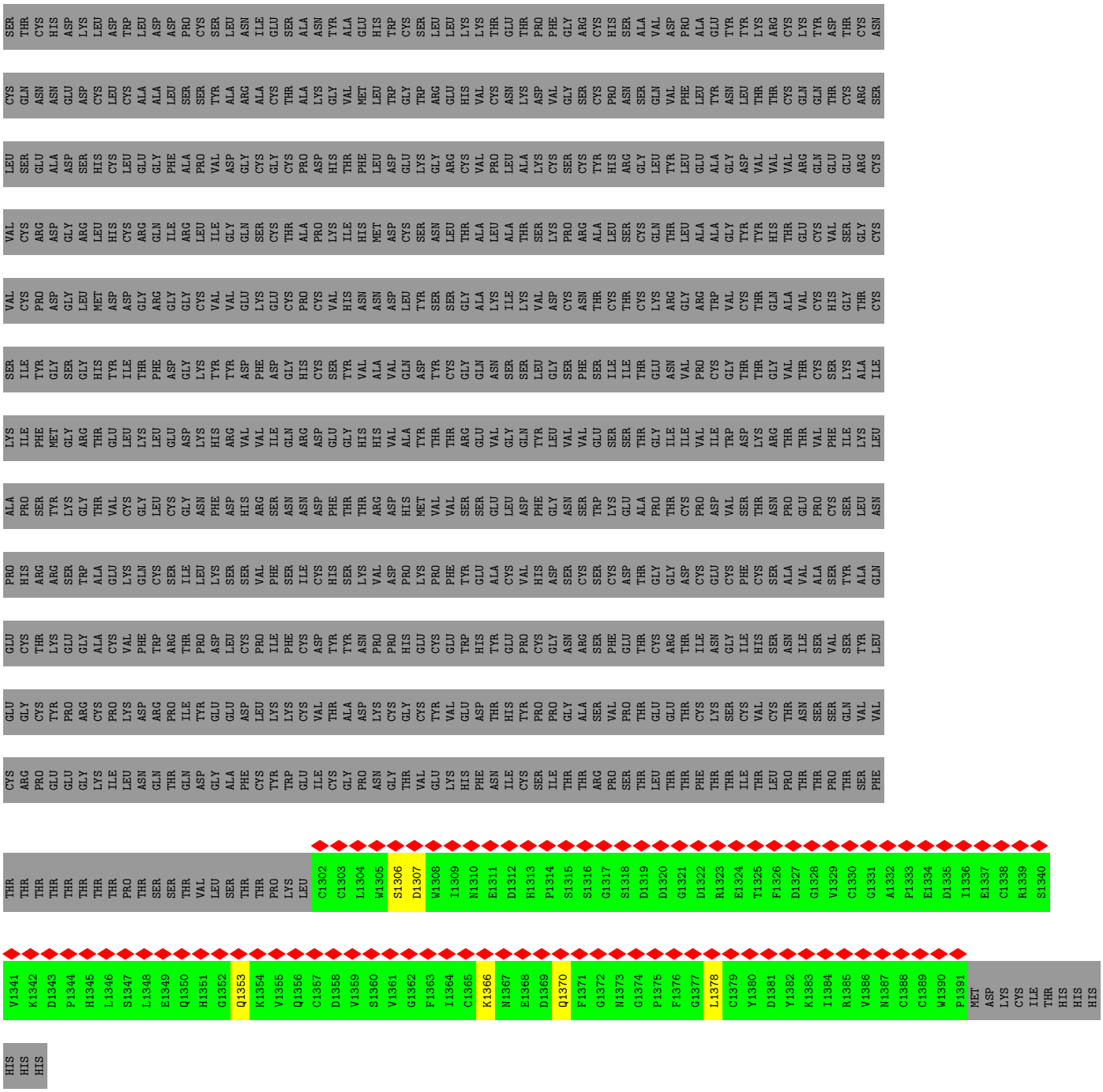
ASP THR CYS VAL CYS SER THR VAL ALA GLU PHE SER ARG GLN CYS SER HIS ALA GLY GLY ARG PRO GLY ASN TRP ARG THR ALA THR LEU CYS PRO LYS THR CYS PRO GLY ASN LEU VAL TYR LEU GLU SER GLY SER PRO CYS MET ASP THR CYS SER ASP LEU LEU GLU VAL GLY

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

VAL CYS LYS ASP LEU PRO THR CYS ALA LEU GLU GLY THR PHE HIS GLY ASP CYS TYR VAL LEU ALA LYS GLY ASP HIS ASN ASP SER TYR ALA LEU LEU GLY GLU LEU ALA PRO CYS GLY SER ASP LYS

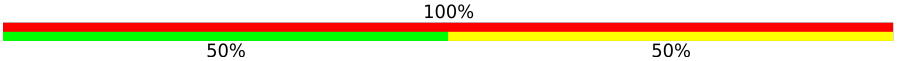
GLN THR CYS LEU LYS THR VAL VAL LEU LEU ALA ASP LYS LYS LYS ASN ALA VAL PHE LYS SER ASP GLY GLN ASN LEU GLN VAL ASN LEU PRO HIS VAL THR ALA SER PHE SER VAL PHE ARG PRO SER SER TYR HIS ILE MET VAL SER MET ALA ILE GLY VAL

ARG LEU GLN VAL GLN LEU PHE VAL THR LEU ASP GLN ALA SER GLN GLY VAL GLN GLY LEU CYS GLY ASN PHE ASN GLY LEU GLU GLY ASP ASP PHE LYS THR ALA SER GLY LEU VAL GLU ALA THR GLY ALA ASN THR TRP LYS ALA GLN



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

Chain K:



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

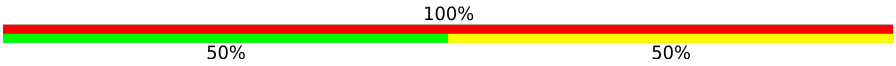
Chain L:

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

Chain M:

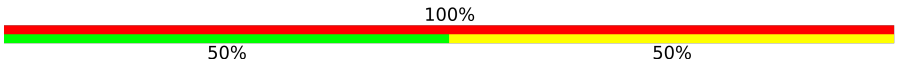
• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain N:



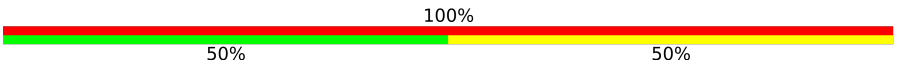
• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain O:



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

Chain P:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	178136	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	48	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	2.714	Depositor
Minimum map value	-1.867	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.037	Depositor
Recommended contour level	0.17	Depositor
Map size (Å)	340.0, 340.0, 340.0	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.85, 0.85, 0.85	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, 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	0.39	0/9358	0.61	0/12718
1	B	0.35	0/5406	0.61	0/7355
1	C	0.40	0/8643	0.62	0/11748
1	D	0.44	0/3952	0.62	0/5363
1	E	0.39	0/9358	0.62	0/12718
1	F	0.31	0/715	0.56	0/970
1	G	0.40	0/8643	0.62	0/11748
1	H	0.44	0/3952	0.62	0/5363
1	I	0.35	0/5406	0.61	0/7355
1	J	0.31	0/715	0.56	0/970
All	All	0.39	0/56148	0.61	0/76308

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	2
1	B	0	1
1	C	0	2
1	D	0	1
1	E	0	2
1	G	0	2
1	H	0	1
1	I	0	1
All	All	0	12

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1126	PRO	Peptide
1	A	424	ASP	Peptide
1	B	424	ASP	Peptide
1	C	1126	PRO	Peptide
1	C	424	ASP	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	A	9125	0	8366	61	0
1	B	5276	0	4839	35	0
1	C	8429	0	7755	61	0
1	D	3849	0	3527	20	0
1	E	9125	0	8366	63	0
1	F	696	0	611	3	0
1	G	8429	0	7755	61	0
1	H	3849	0	3527	20	0
1	I	5276	0	4839	35	0
1	J	696	0	611	3	0
2	K	28	0	25	0	0
2	L	28	0	25	0	0
2	M	28	0	25	0	0
2	N	28	0	25	0	0
2	O	28	0	25	0	0
2	P	28	0	25	0	0
3	A	5	0	0	0	0
3	B	2	0	0	0	0
3	C	3	0	0	0	0
3	D	3	0	0	0	0
3	E	5	0	0	0	0
3	F	2	0	0	0	0
3	G	3	0	0	0	0
3	H	3	0	0	0	0
3	I	2	0	0	0	0
3	J	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	28	0	26	0	0
4	B	28	0	26	0	0
4	C	28	0	26	0	0
4	E	28	0	26	0	0
4	G	28	0	26	0	0
4	I	28	0	26	0	0
5	A	23	0	0	0	0
5	B	10	0	0	0	0
5	C	22	0	0	0	0
5	D	13	0	0	0	0
5	E	23	0	0	0	0
5	F	1	0	0	0	0
5	G	22	0	0	0	0
5	H	13	0	0	0	0
5	I	10	0	0	0	0
5	J	1	0	0	0	0
All	All	55254	0	50502	322	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 3.

The worst 5 of 322 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:35:ASN:N	1:C:172:TYR:HH	1.66	0.93
1:G:35:ASN:N	1:G:172:TYR:HH	1.66	0.92
1:E:474:ASN:ND2	1:G:474:ASN:ND2	2.18	0.91
1:A:474:ASN:ND2	1:C:474:ASN:ND2	2.18	0.91
1:A:35:ASN:N	1:A:172:TYR:HH	1.71	0.89

There are no symmetry-related clashes.

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 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	1195/1383 (86%)	1061 (89%)	134 (11%)	0	100	100
1	B	705/1383 (51%)	611 (87%)	94 (13%)	0	100	100
1	C	1107/1383 (80%)	971 (88%)	136 (12%)	0	100	100
1	D	490/1383 (35%)	450 (92%)	40 (8%)	0	100	100
1	E	1195/1383 (86%)	1062 (89%)	133 (11%)	0	100	100
1	F	88/1383 (6%)	87 (99%)	1 (1%)	0	100	100
1	G	1107/1383 (80%)	974 (88%)	133 (12%)	0	100	100
1	H	490/1383 (35%)	450 (92%)	40 (8%)	0	100	100
1	I	705/1383 (51%)	612 (87%)	93 (13%)	0	100	100
1	J	88/1383 (6%)	87 (99%)	1 (1%)	0	100	100
All	All	7170/13830 (52%)	6365 (89%)	805 (11%)	0	100	100

There are no Ramachandran outliers to report.

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 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	996/1204 (83%)	994 (100%)	2 (0%)	93	98
1	B	562/1204 (47%)	561 (100%)	1 (0%)	93	98
1	C	918/1204 (76%)	916 (100%)	2 (0%)	93	98
1	D	434/1204 (36%)	433 (100%)	1 (0%)	93	98
1	E	996/1204 (83%)	994 (100%)	2 (0%)	93	98
1	F	78/1204 (6%)	78 (100%)	0	100	100
1	G	918/1204 (76%)	916 (100%)	2 (0%)	93	98
1	H	434/1204 (36%)	433 (100%)	1 (0%)	93	98
1	I	562/1204 (47%)	561 (100%)	1 (0%)	93	98

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	J	78/1204 (6%)	78 (100%)	0	100 100
All	All	5976/12040 (50%)	5964 (100%)	12 (0%)	93 98

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	998	ARG
1	G	419	LYS
1	I	419	LYS
1	G	998	ARG
1	C	419	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	352	HIS
1	F	1370	GLN
1	I	503	GLN
1	E	474	ASN
1	E	665	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)

12 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	K	1	2,1	14,14,15	0.31	0	17,19,21	0.46	0
2	NAG	K	2	2	14,14,15	0.28	0	17,19,21	0.59	1 (5%)
2	NAG	L	1	2,1	14,14,15	0.31	0	17,19,21	0.46	0
2	NAG	L	2	2	14,14,15	0.27	0	17,19,21	0.58	1 (5%)
2	NAG	M	1	2,1	14,14,15	0.32	0	17,19,21	0.48	0
2	NAG	M	2	2	14,14,15	0.29	0	17,19,21	0.59	1 (5%)
2	NAG	N	1	2,1	14,14,15	0.31	0	17,19,21	0.47	0
2	NAG	N	2	2	14,14,15	0.28	0	17,19,21	0.58	1 (5%)
2	NAG	O	1	2,1	14,14,15	0.30	0	17,19,21	0.46	0
2	NAG	O	2	2	14,14,15	0.27	0	17,19,21	0.58	1 (5%)
2	NAG	P	1	2,1	14,14,15	0.33	0	17,19,21	0.47	0
2	NAG	P	2	2	14,14,15	0.28	0	17,19,21	0.59	1 (5%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	K	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	K	2	2	-	2/6/23/26	0/1/1/1
2	NAG	L	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	L	2	2	-	2/6/23/26	0/1/1/1
2	NAG	M	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	M	2	2	-	2/6/23/26	0/1/1/1
2	NAG	N	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	N	2	2	-	2/6/23/26	0/1/1/1
2	NAG	O	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	O	2	2	-	2/6/23/26	0/1/1/1
2	NAG	P	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	P	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	P	2	NAG	C1-O5-C5	2.04	114.95	112.19
2	K	2	NAG	C1-O5-C5	2.03	114.95	112.19
2	L	2	NAG	C1-O5-C5	2.03	114.94	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	M	2	NAG	C1-O5-C5	2.02	114.92	112.19
2	O	2	NAG	C1-O5-C5	2.01	114.92	112.19

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

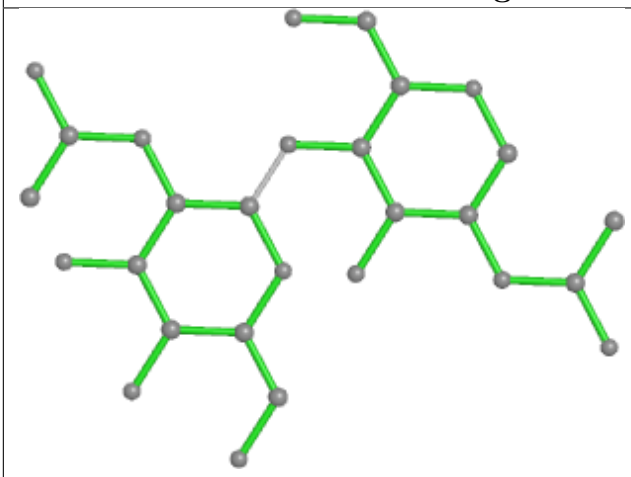
Mol	Chain	Res	Type	Atoms
2	K	2	NAG	O5-C5-C6-O6
2	L	2	NAG	O5-C5-C6-O6
2	M	2	NAG	O5-C5-C6-O6
2	N	2	NAG	O5-C5-C6-O6
2	O	2	NAG	O5-C5-C6-O6

There are no ring outliers.

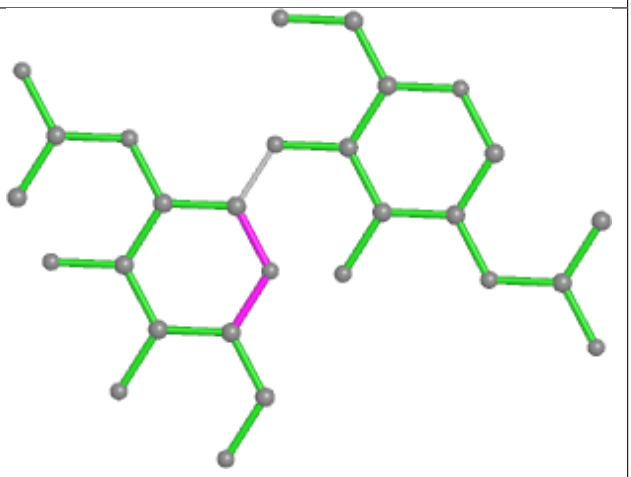
No monomer is involved in short contacts.

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

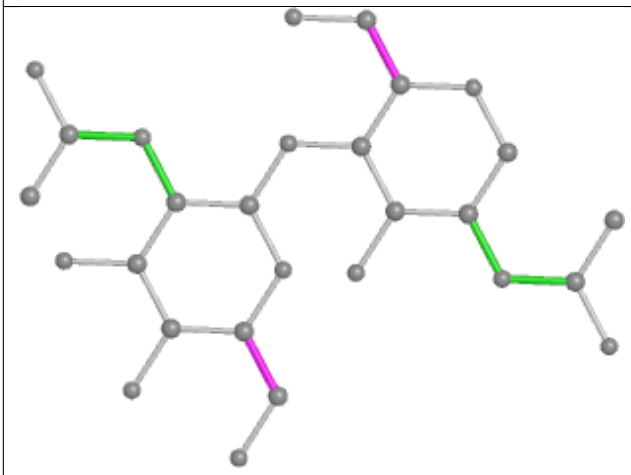
Oligosaccharide Chain K



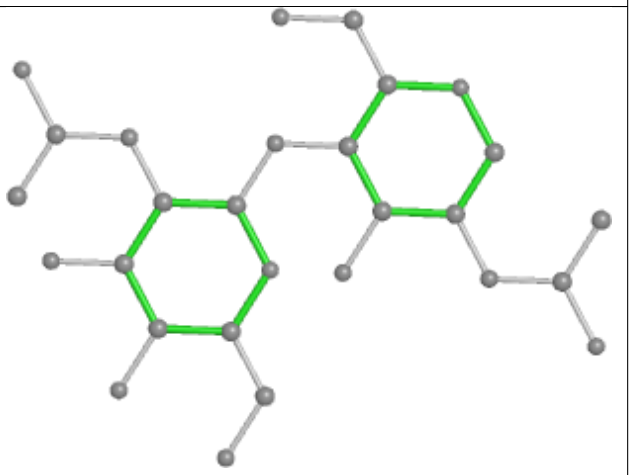
Bond lengths



Bond angles

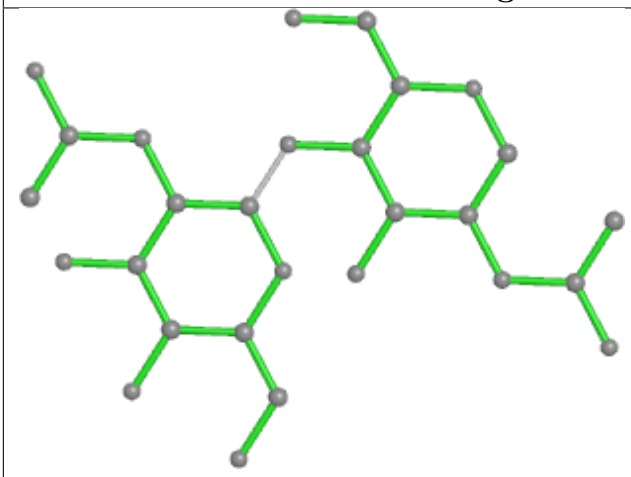


Torsions

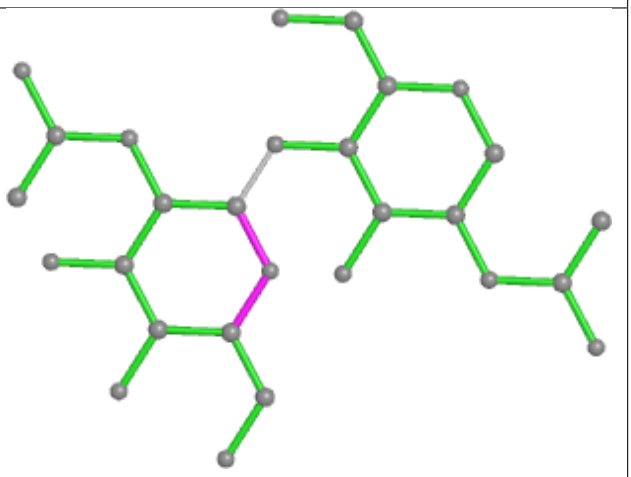


Rings

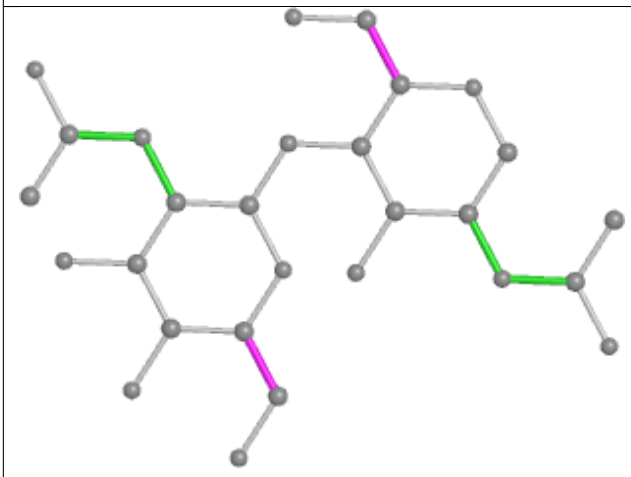
Oligosaccharide Chain L



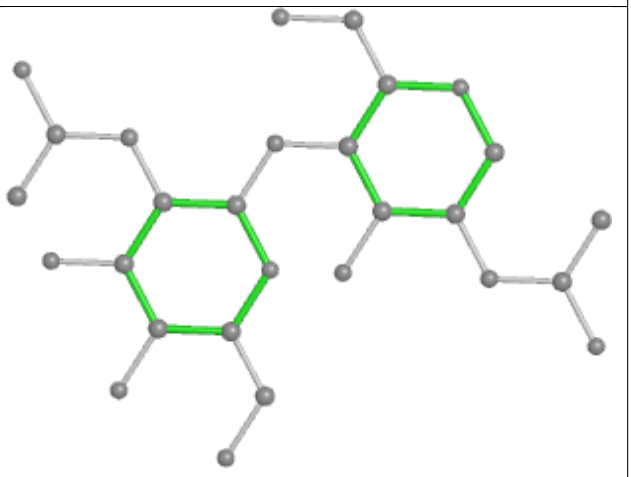
Bond lengths



Bond angles

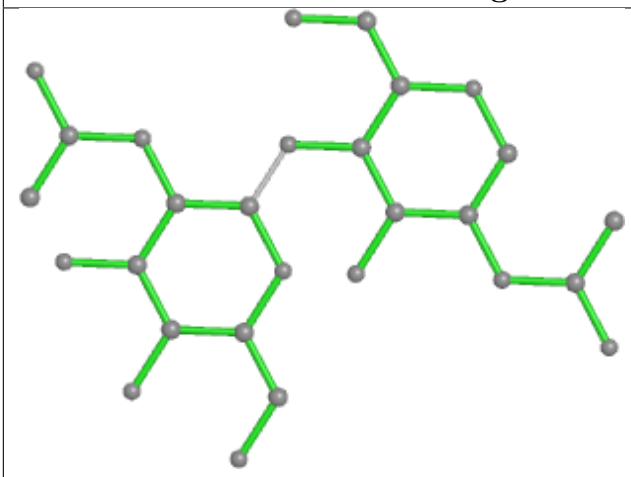


Torsions

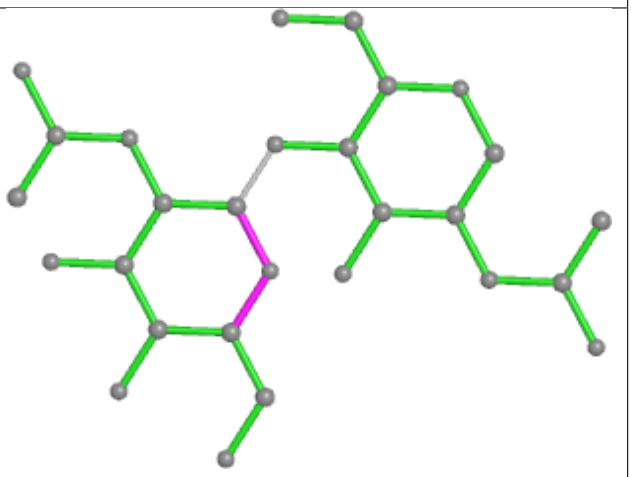


Rings

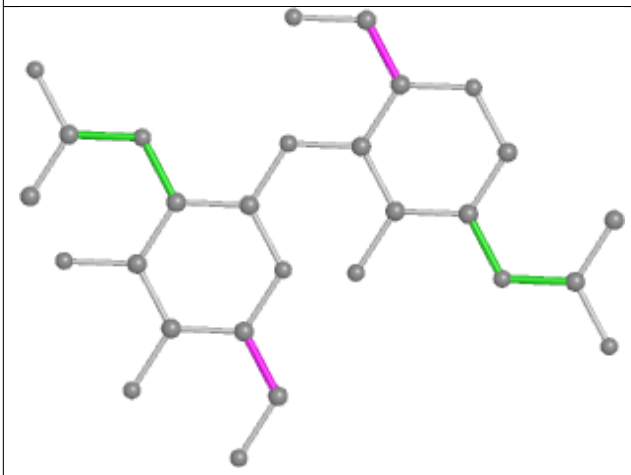
Oligosaccharide Chain M



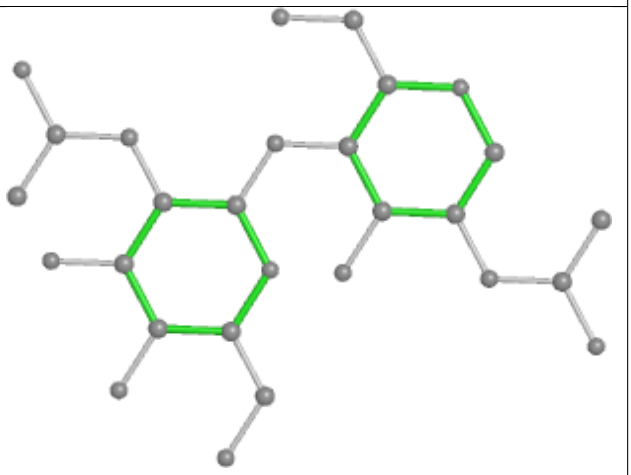
Bond lengths



Bond angles

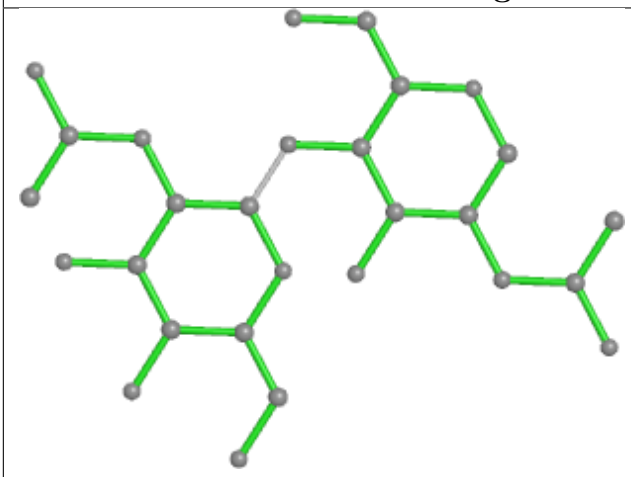


Torsions

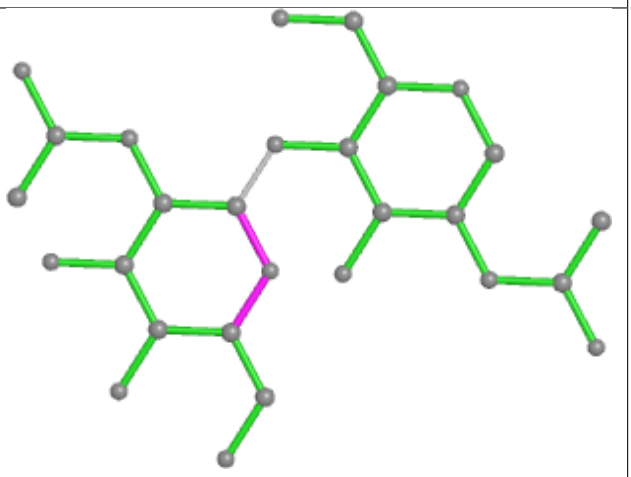


Rings

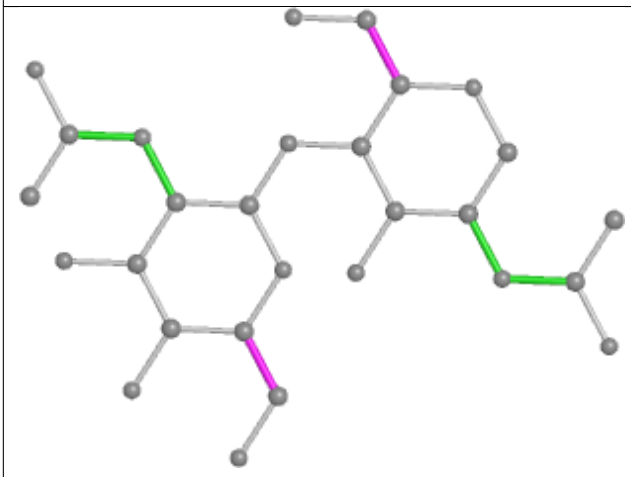
Oligosaccharide Chain N



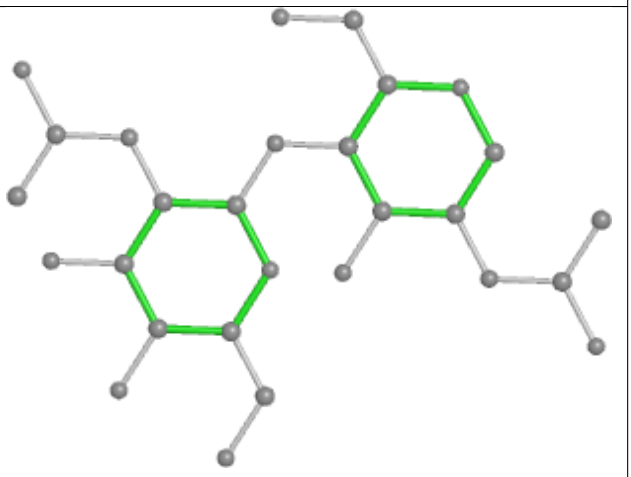
Bond lengths



Bond angles

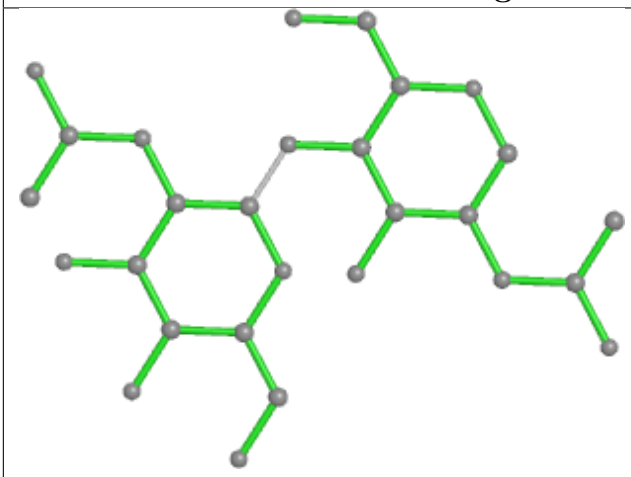


Torsions

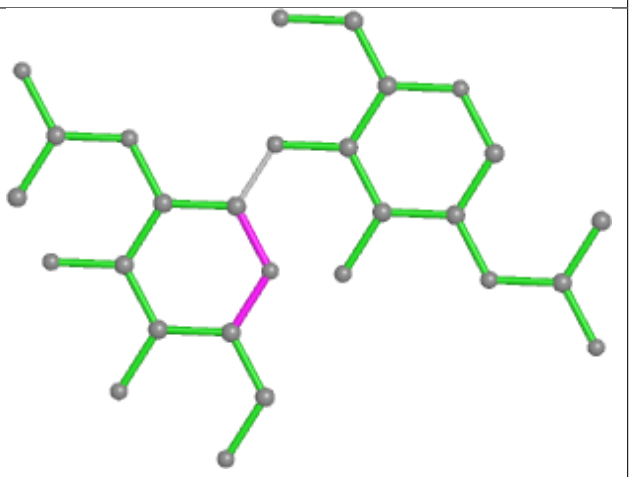


Rings

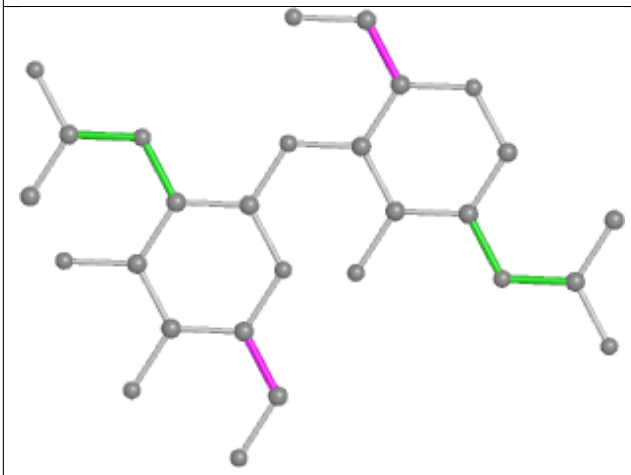
Oligosaccharide Chain O



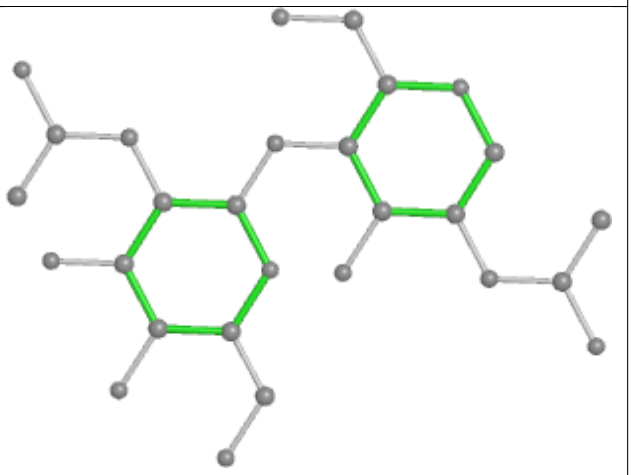
Bond lengths



Bond angles

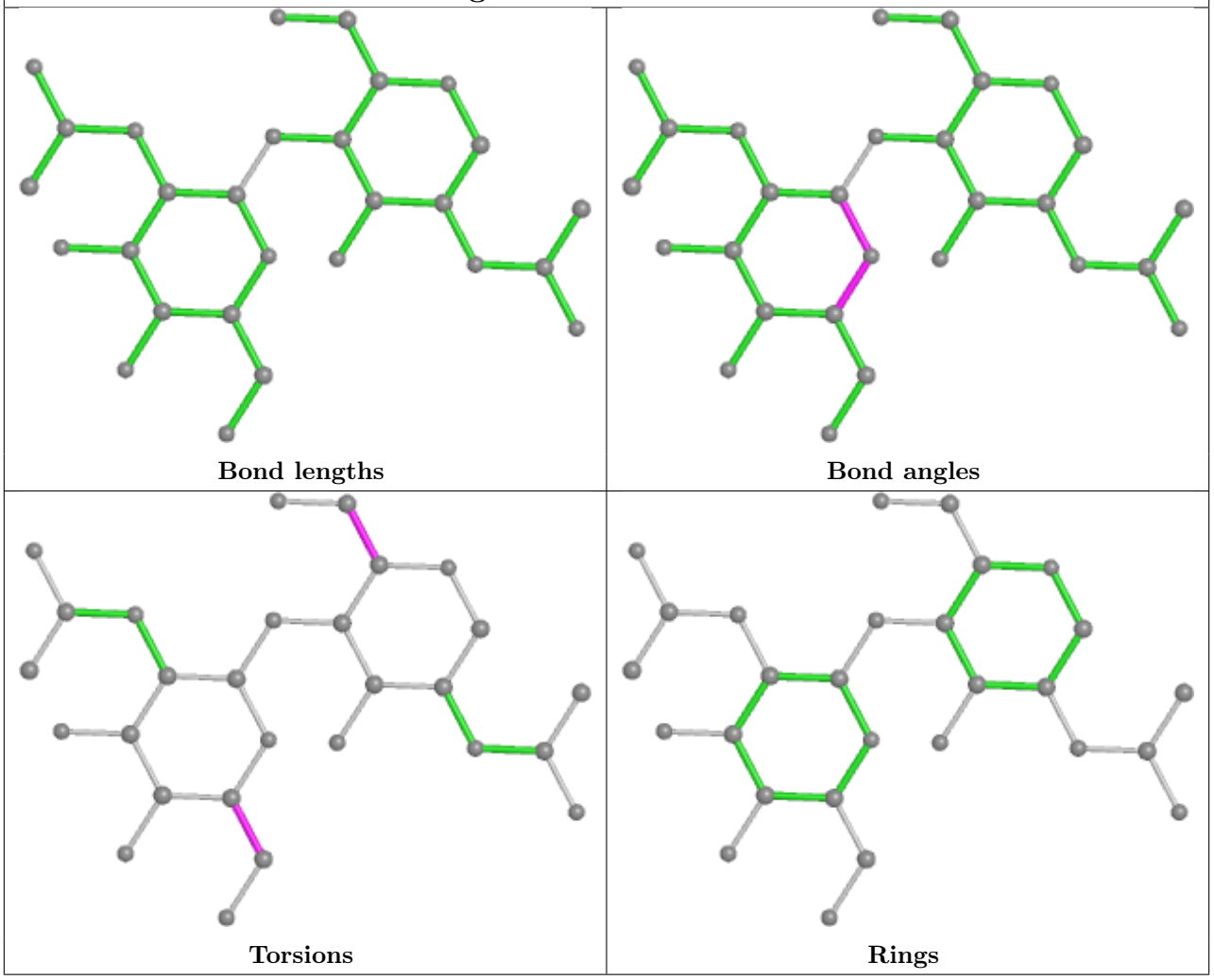


Torsions



Rings

Oligosaccharide Chain P



5.6 Ligand geometry (i)

Of 42 ligands modelled in this entry, 30 are monoatomic - leaving 12 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	NAG	I	2004	1	14,14,15	0.33	0	17,19,21	0.49	0
4	NAG	G	2004	1	14,14,15	1.33	2 (14%)	17,19,21	1.27	1 (5%)
4	NAG	B	2004	1	14,14,15	0.33	0	17,19,21	0.49	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	E	2006	1	14,14,15	1.35	2 (14%)	17,19,21	1.26	1 (5%)
4	NAG	B	2003	1	14,14,15	1.33	2 (14%)	17,19,21	1.26	1 (5%)
4	NAG	A	1507	1	14,14,15	0.34	0	17,19,21	0.49	0
4	NAG	C	2004	1	14,14,15	1.34	2 (14%)	17,19,21	1.27	1 (5%)
4	NAG	E	2007	1	14,14,15	0.34	0	17,19,21	0.49	0
4	NAG	C	2005	1	14,14,15	0.35	0	17,19,21	0.49	0
4	NAG	A	1506	1	14,14,15	1.34	2 (14%)	17,19,21	1.27	1 (5%)
4	NAG	G	2005	1	14,14,15	0.33	0	17,19,21	0.50	0
4	NAG	I	2003	1	14,14,15	1.34	2 (14%)	17,19,21	1.26	1 (5%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	I	2004	1	-	4/6/23/26	0/1/1/1
4	NAG	G	2004	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2004	1	-	4/6/23/26	0/1/1/1
4	NAG	E	2006	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2003	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1507	1	-	4/6/23/26	0/1/1/1
4	NAG	C	2004	1	-	2/6/23/26	0/1/1/1
4	NAG	E	2007	1	-	4/6/23/26	0/1/1/1
4	NAG	C	2005	1	-	4/6/23/26	0/1/1/1
4	NAG	A	1506	1	-	2/6/23/26	0/1/1/1
4	NAG	G	2005	1	-	4/6/23/26	0/1/1/1
4	NAG	I	2003	1	-	2/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	I	2003	NAG	O5-C1	4.44	1.50	1.43
4	A	1506	NAG	O5-C1	4.43	1.50	1.43
4	E	2006	NAG	O5-C1	4.43	1.50	1.43
4	B	2003	NAG	O5-C1	4.42	1.50	1.43
4	C	2004	NAG	O5-C1	4.40	1.50	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	2004	NAG	C1-O5-C5	5.04	119.02	112.19
4	A	1506	NAG	C1-O5-C5	5.03	119.00	112.19
4	G	2004	NAG	C1-O5-C5	5.03	119.00	112.19
4	E	2006	NAG	C1-O5-C5	5.00	118.97	112.19
4	B	2003	NAG	C1-O5-C5	5.00	118.97	112.19

There are no chirality outliers.

5 of 36 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1507	NAG	O5-C5-C6-O6
4	B	2004	NAG	O5-C5-C6-O6
4	C	2005	NAG	O5-C5-C6-O6
4	E	2007	NAG	O5-C5-C6-O6
4	G	2005	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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 Map visualisation

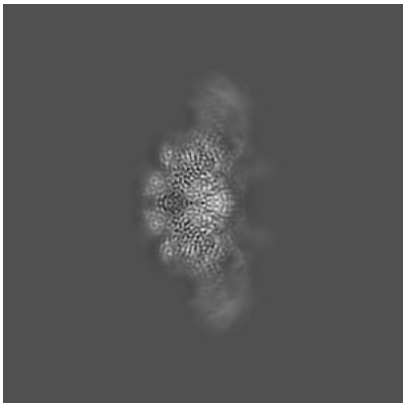
This section contains visualisations of the EMDB entry EMD-10517. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

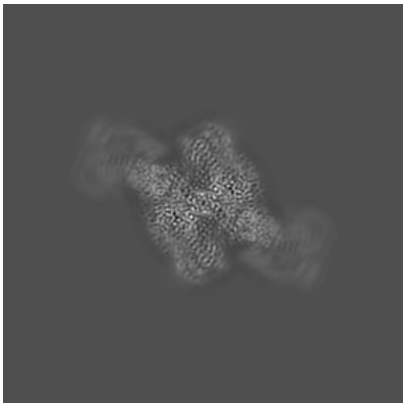
6.1 Orthogonal projections

6.1.1 Primary map

X



Y

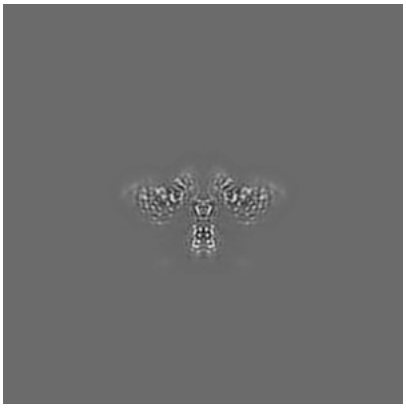


Z

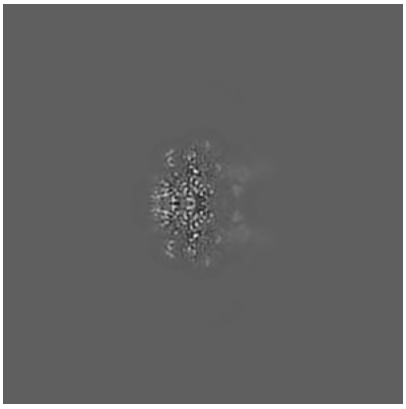
The images above show the map projected in three orthogonal directions.

6.2 Central slices

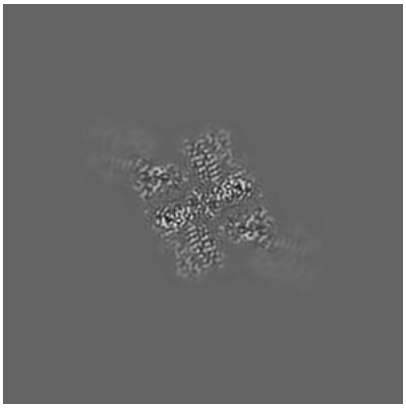
6.2.1 Primary map



X Index: 200



Y Index: 200

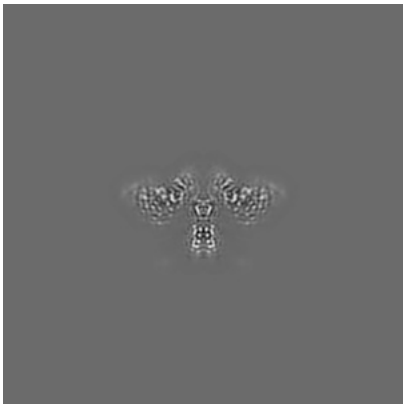


Z Index: 200

The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map



X Index: 200

Y Index: 212

Z Index: 189

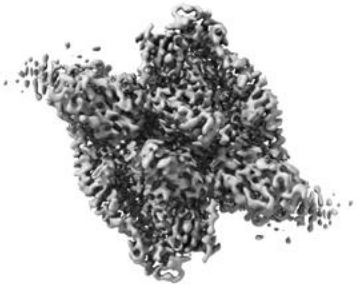
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.17. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

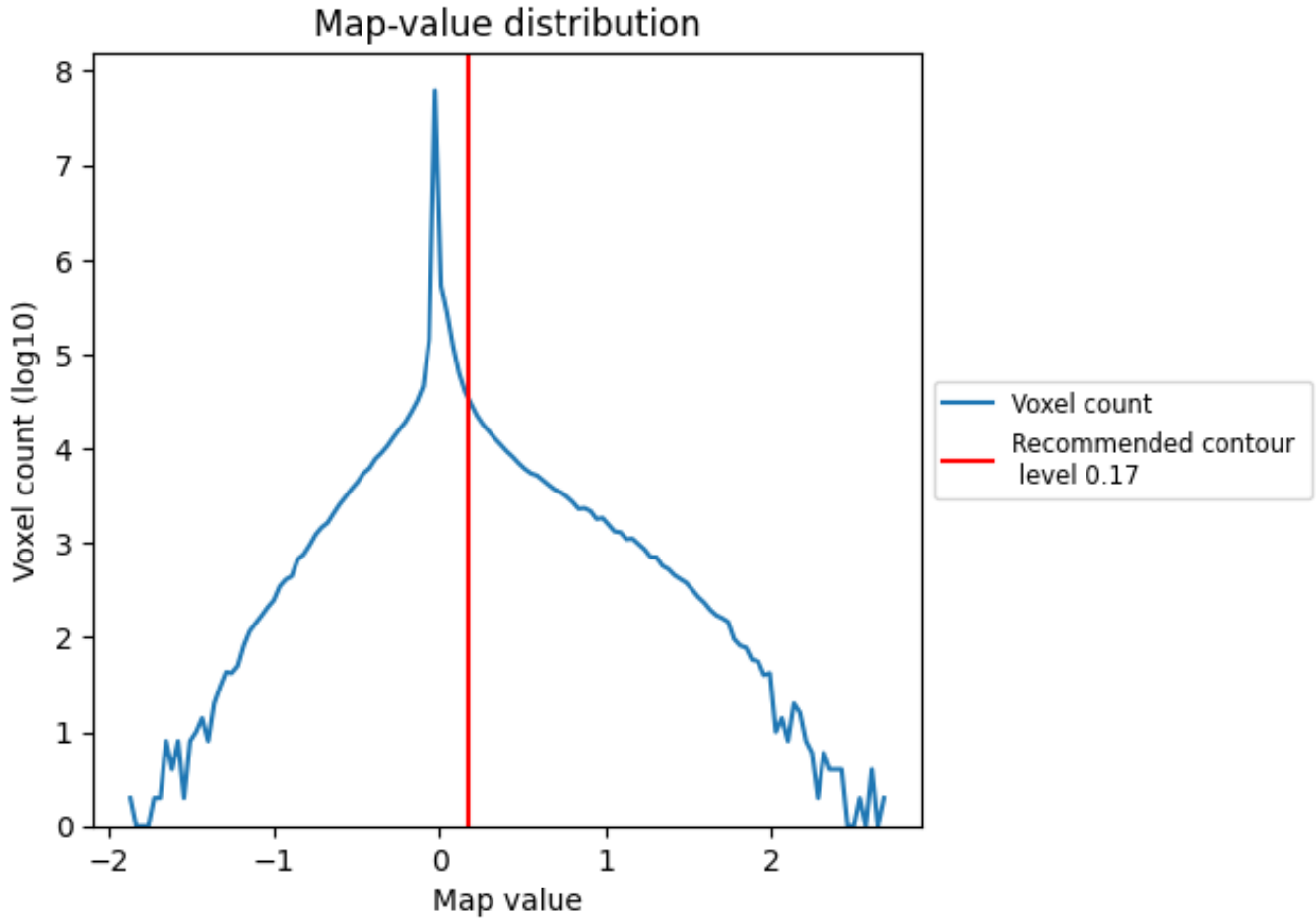
6.5 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

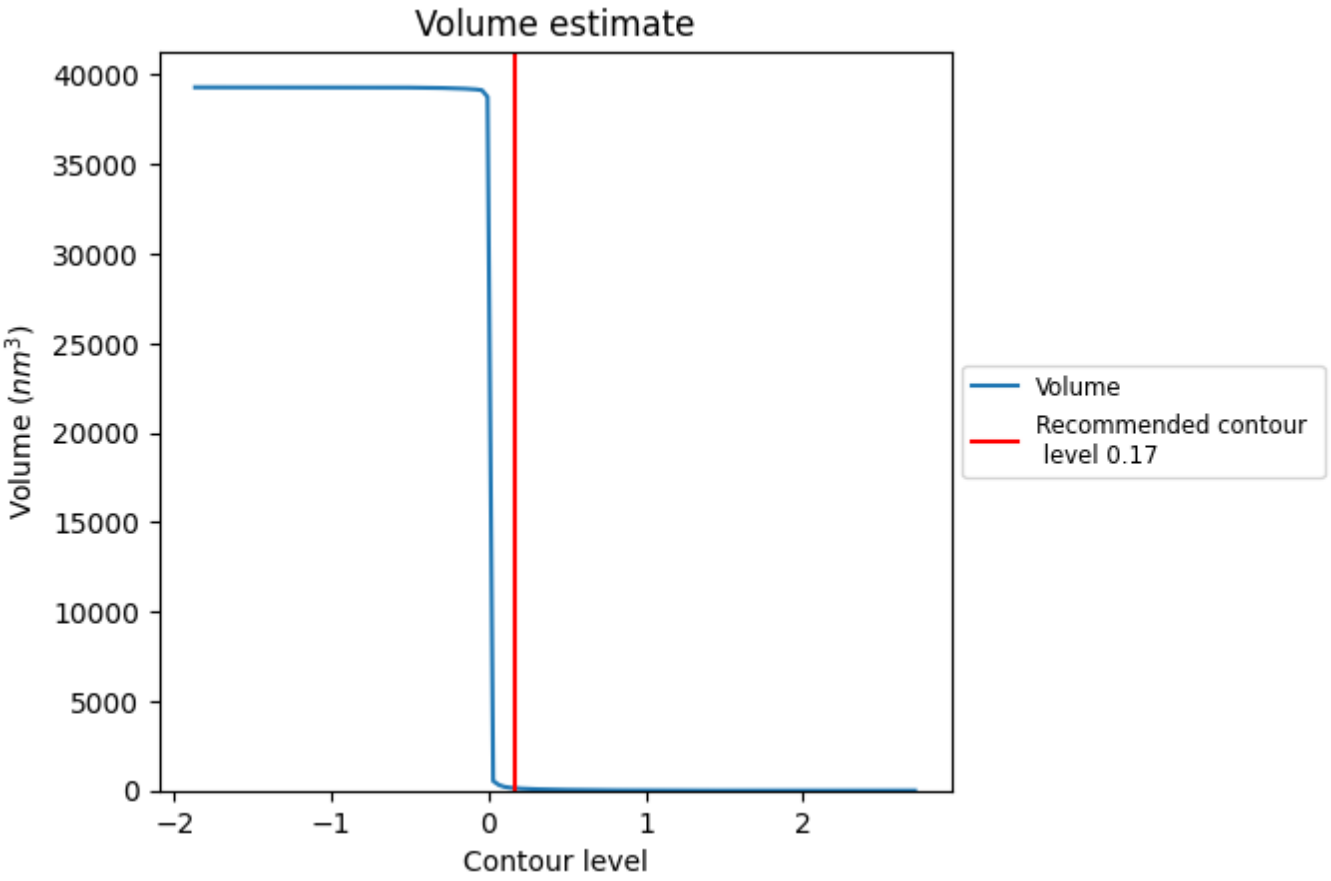
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

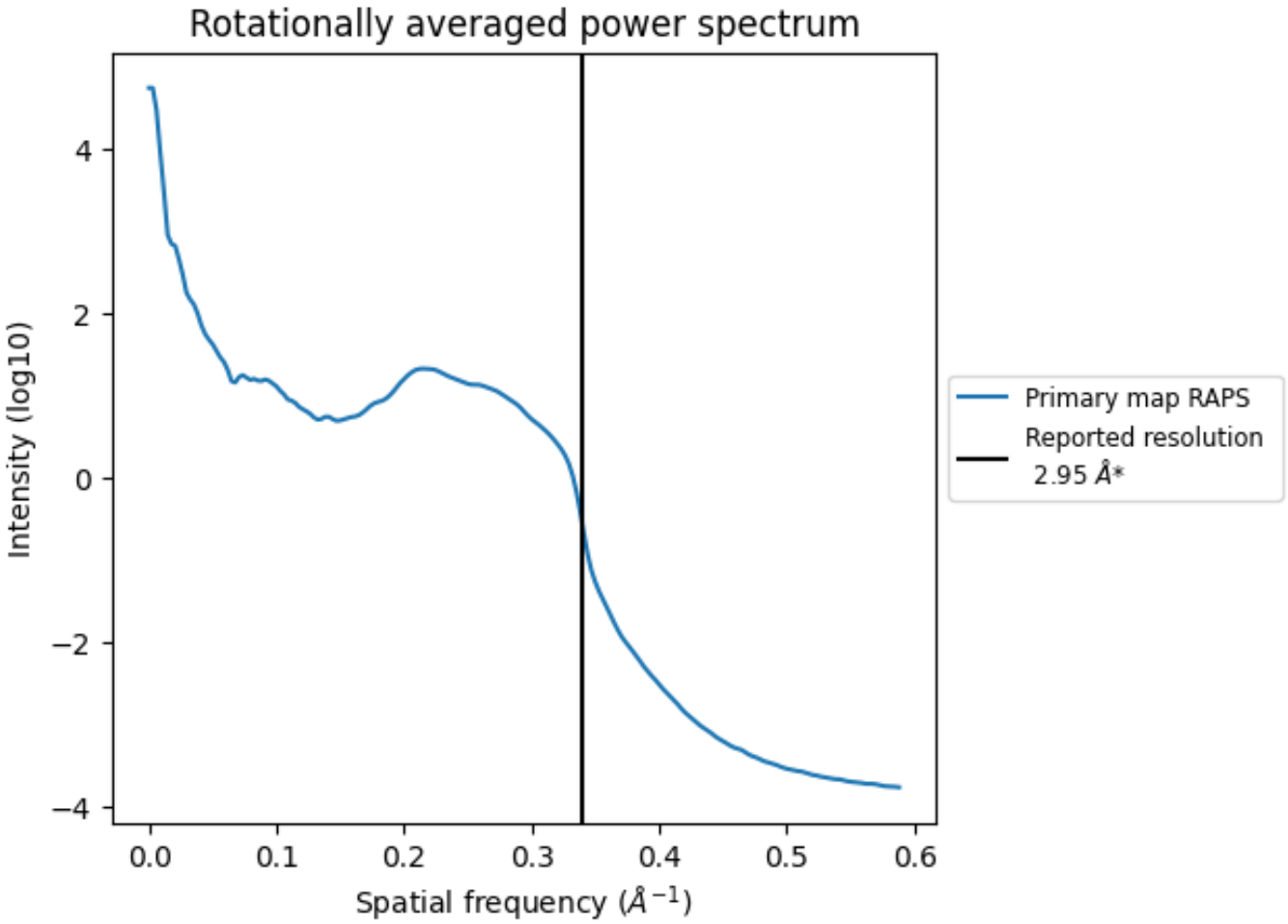
7.2 Volume estimate (i)



The volume at the recommended contour level is 133 nm^3; this corresponds to an approximate mass of 120 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)

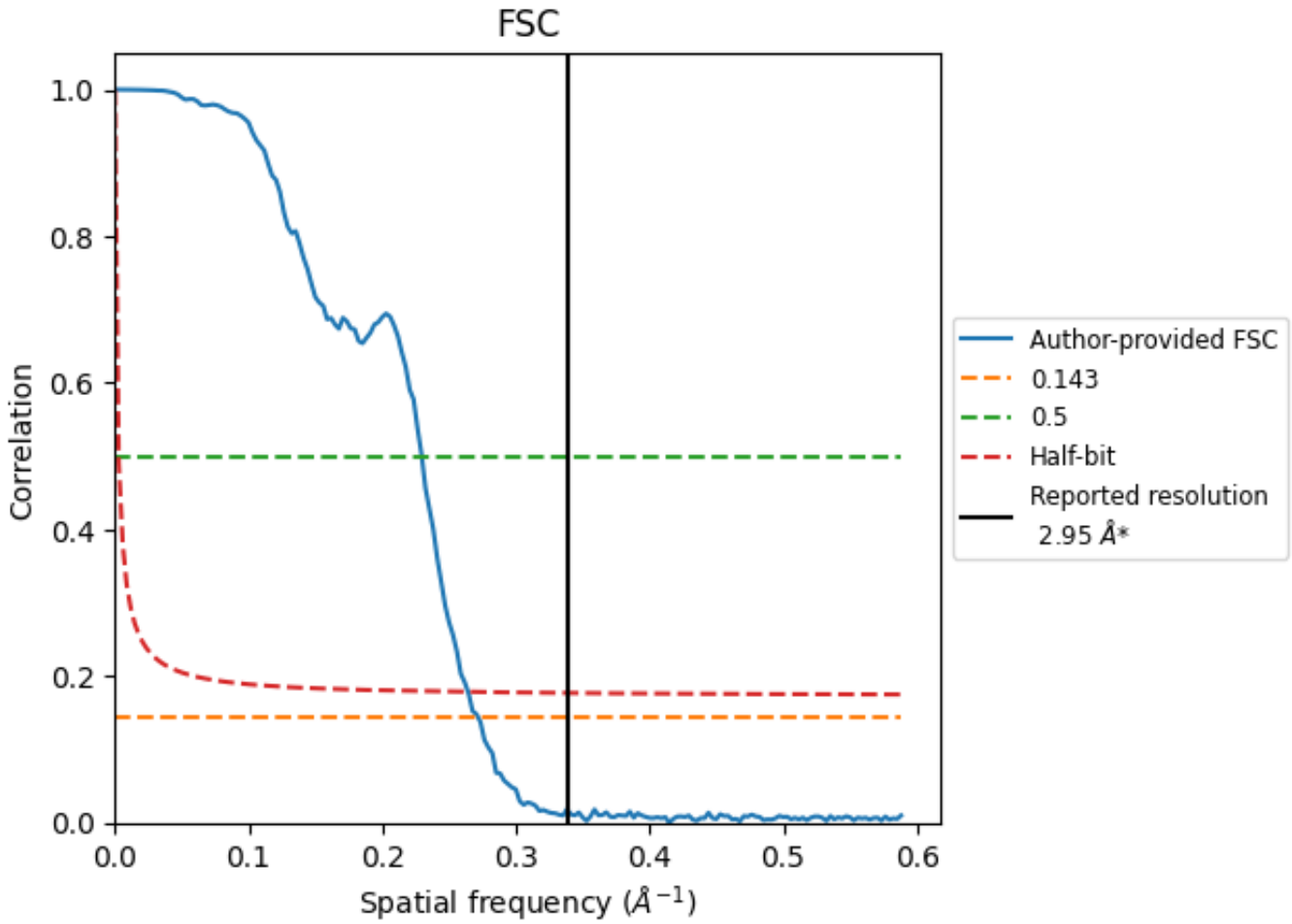


*Reported resolution corresponds to spatial frequency of 0.339 Å^{-1}

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.339 Å^{-1}

8.2 Resolution estimates (i)

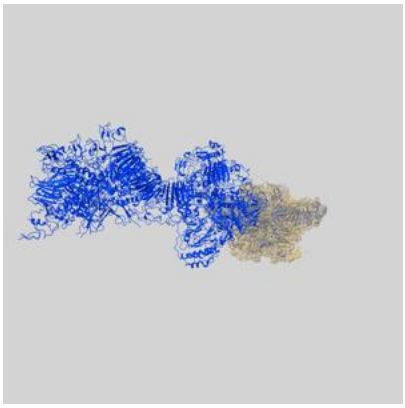
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.95	-	-
Author-provided FSC curve	3.68	4.36	3.79
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 3.68 differs from the reported value 2.95 by more than 10 %

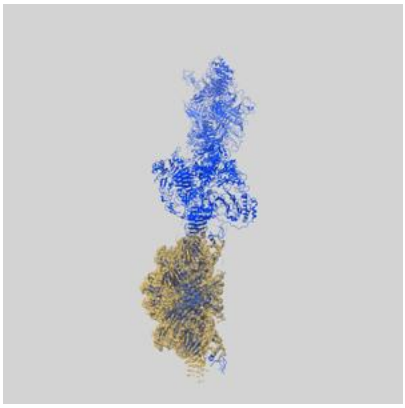
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-10517 and PDB model 7A5O. Per-residue inclusion information can be found in section 3 on page 10.

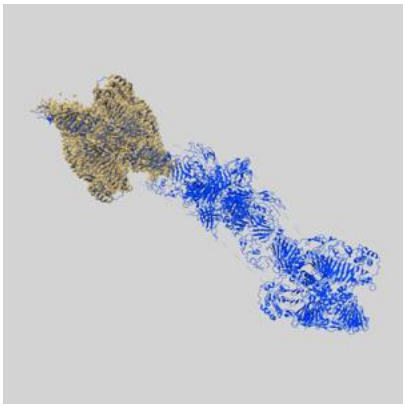
9.1 Map-model overlay



X



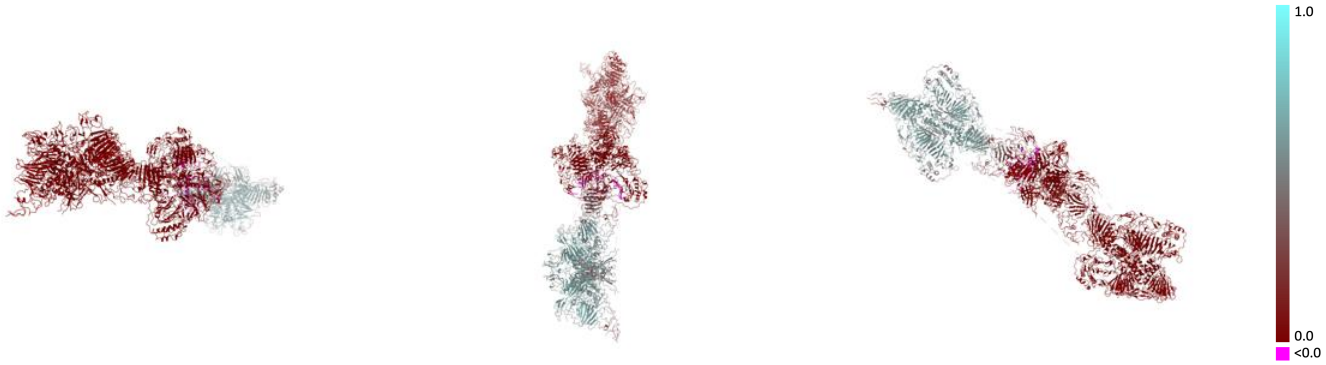
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.17 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

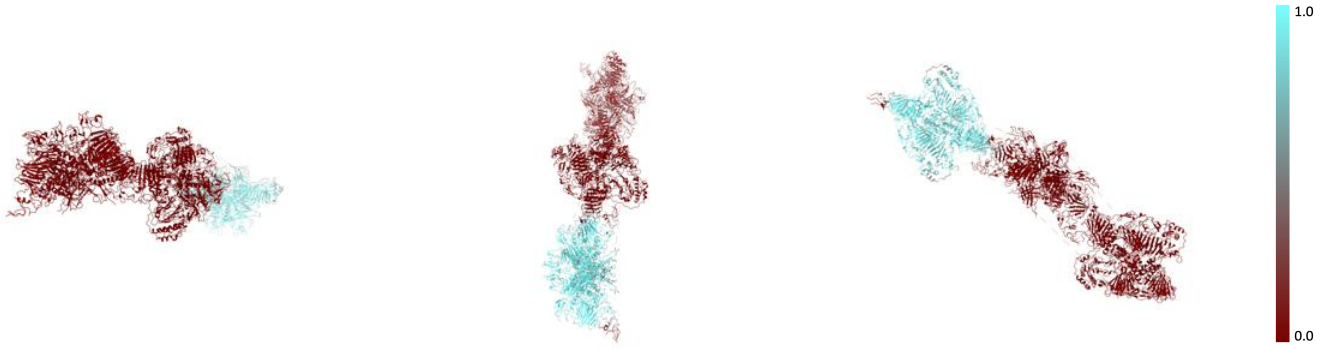
9.2 Q-score mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

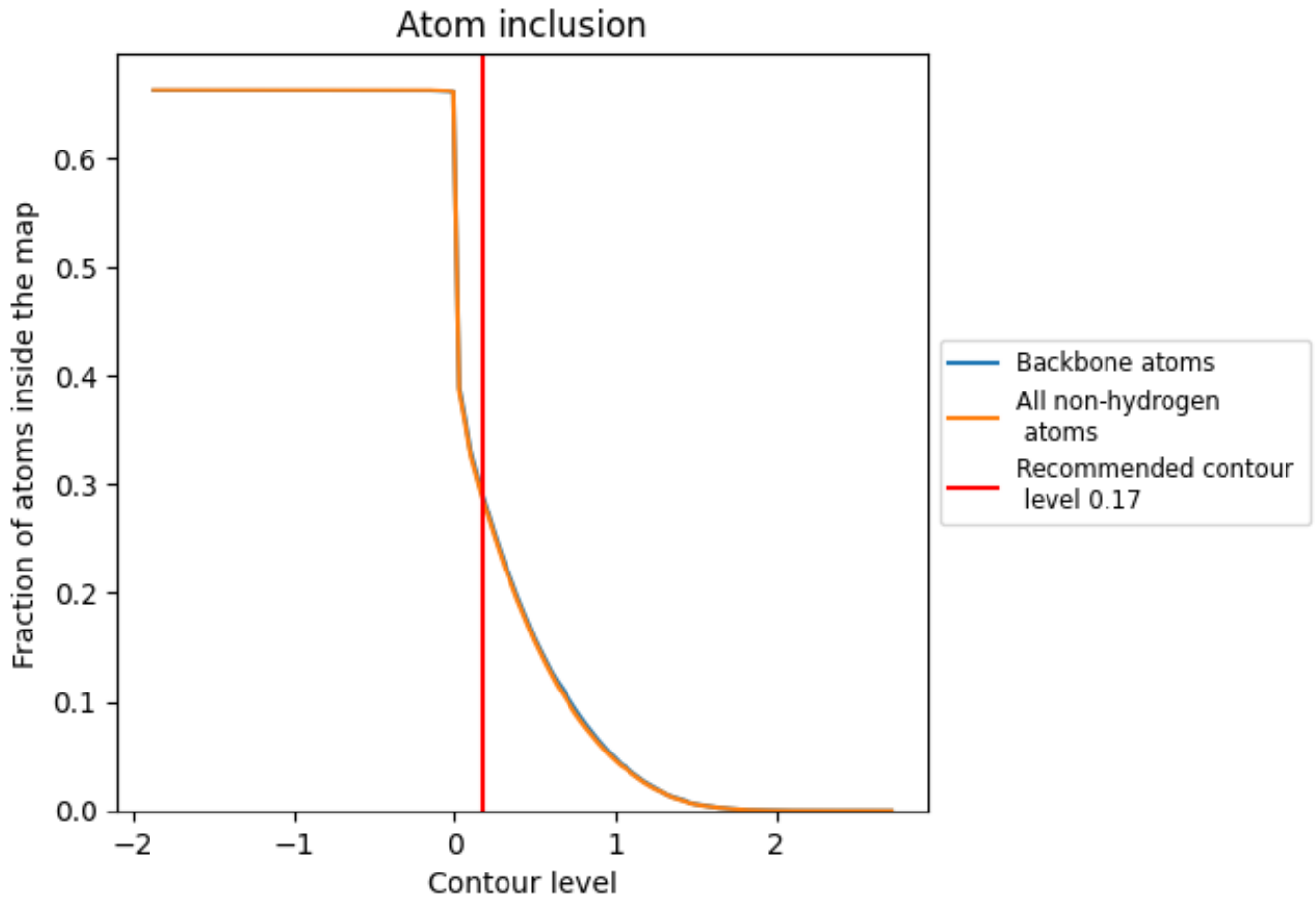
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.17).

9.4 Atom inclusion (i)



At the recommended contour level, 29% of all backbone atoms, 29% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.17) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.2884	0.1950
A	0.4809	0.3130
B	0.8299	0.5300
C	0.3596	0.2920
D	0.7620	0.4810
E	0.0604	0.0390
F	0.7889	0.5130
G	0.0000	-0.0020
H	0.0000	0.0010
I	0.0000	0.0000
J	0.0000	0.0000
K	0.0000	0.0000
L	0.8214	0.5480
M	0.8214	0.5450
N	0.0000	0.0000
O	0.0000	0.0000
P	0.0000	0.0000