wwPDB X-ray Structure Validation Summary Report (i)

Feb 22, 2024 – 01:33 PM EST

PDB ID	:	4V40
Title	:	BETA-GALACTOSIDASE
Authors	:	Jacobson, R.H.; Zhang, X.; Dubose, R.F.; Matthews, B.W.
Deposited on	:	1994-07-18
Resolution	:	2.50 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

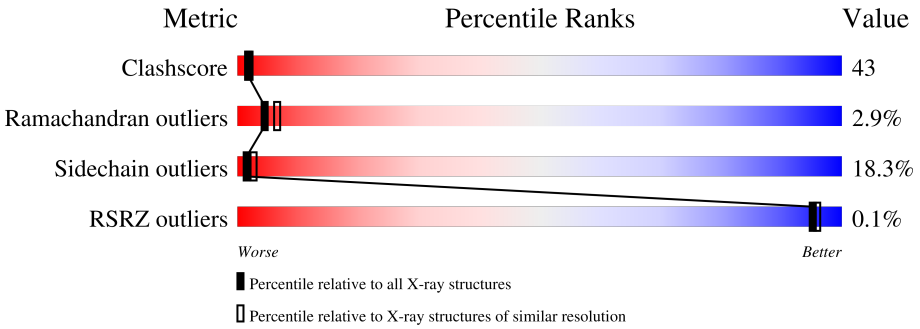
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1023	39% 45% 14% ·
1	B	1023	38% 43% 16% ·
1	C	1023	44% 41% 12% ·
1	D	1023	38% 46% 14% ·
1	E	1023	26% 49% 21% ·
1	F	1023	37% 46% 13% ·
1	G	1023	35% 44% 18% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	1023	26% 50% 20%
1	I	1023	33% 49% 15%
1	J	1023	39% 46% 13%
1	K	1023	30% 50% 16%
1	L	1023	29% 49% 19%
1	M	1023	19% 54% 23%
1	N	1023	34% 46% 17%
1	O	1023	32% 49% 16%
1	P	1023	% 19% 52% 25%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 132654 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called BETA-GALACTOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	B	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	C	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	D	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	E	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	F	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	G	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	H	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	I	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	J	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	K	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	L	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	M	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	N	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	O	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			
1	P	1021	Total	C	N	O	S	0	0	0
			8198	5185	1451	1524	38			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Mg 2 2	0	0
2	B	2	Total Mg 2 2	0	0
2	C	2	Total Mg 2 2	0	0
2	D	2	Total Mg 2 2	0	0
2	E	2	Total Mg 2 2	0	0
2	F	2	Total Mg 2 2	0	0
2	G	2	Total Mg 2 2	0	0
2	H	2	Total Mg 2 2	0	0
2	I	2	Total Mg 2 2	0	0
2	J	2	Total Mg 2 2	0	0
2	K	2	Total Mg 2 2	0	0
2	L	2	Total Mg 2 2	0	0
2	M	1	Total Mg 1 1	0	0
2	N	2	Total Mg 2 2	0	0
2	O	2	Total Mg 2 2	0	0
2	P	2	Total Mg 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	88	Total O 88 88	0	0
3	B	96	Total O 96 96	0	0
3	C	91	Total O 91 91	0	0

Continued on next page...

Continued from previous page...

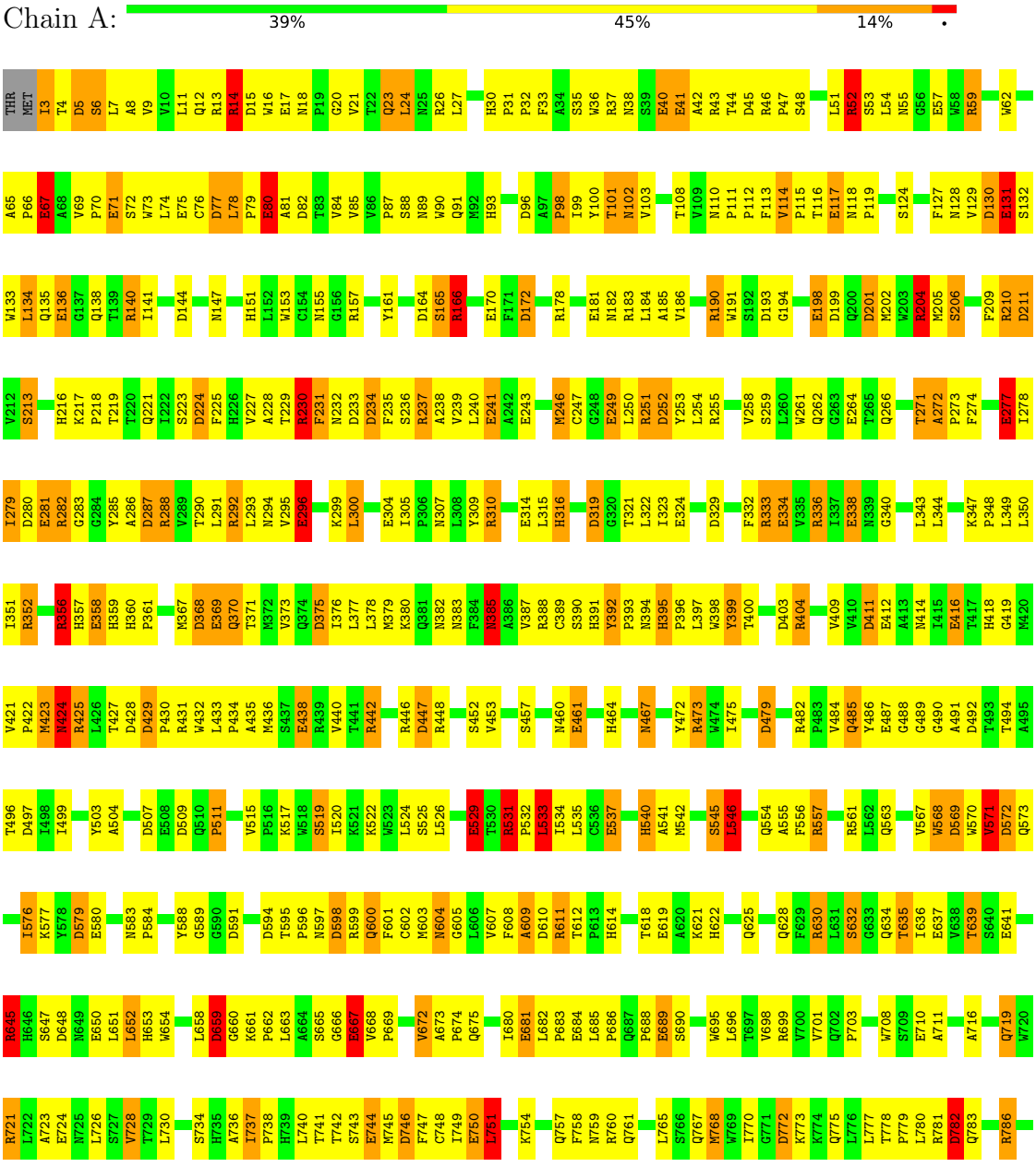
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	97	Total O 97 97	0	0
3	E	94	Total O 94 94	0	0
3	F	91	Total O 91 91	0	0
3	G	95	Total O 95 95	0	0
3	H	92	Total O 92 92	0	0
3	I	90	Total O 90 90	0	0
3	J	97	Total O 97 97	0	0
3	K	87	Total O 87 87	0	0
3	L	84	Total O 84 84	0	0
3	M	79	Total O 79 79	0	0
3	N	94	Total O 94 94	0	0
3	O	95	Total O 95 95	0	0
3	P	85	Total O 85 85	0	0

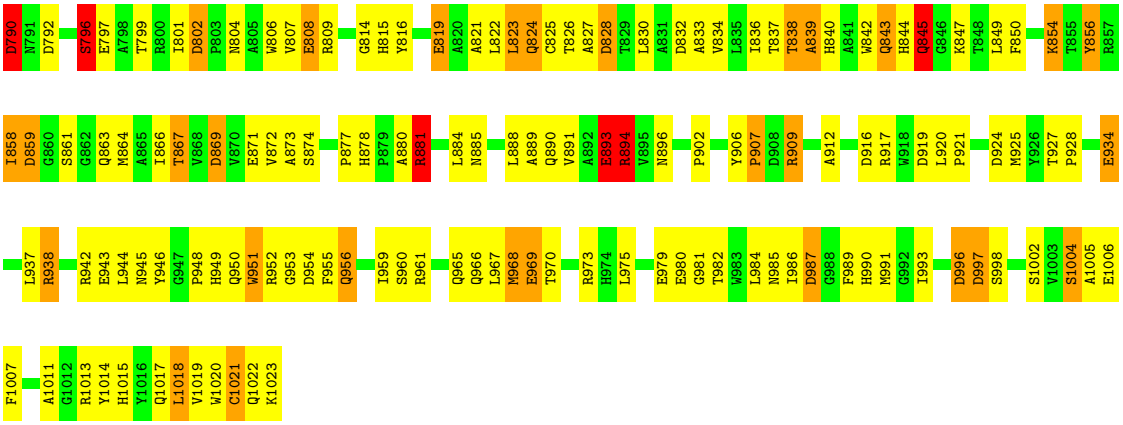
3 Residue-property plots (i)

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: BETA-GALACTOSIDASE

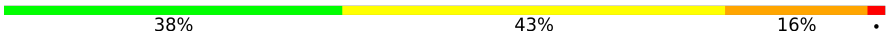
Chain A:

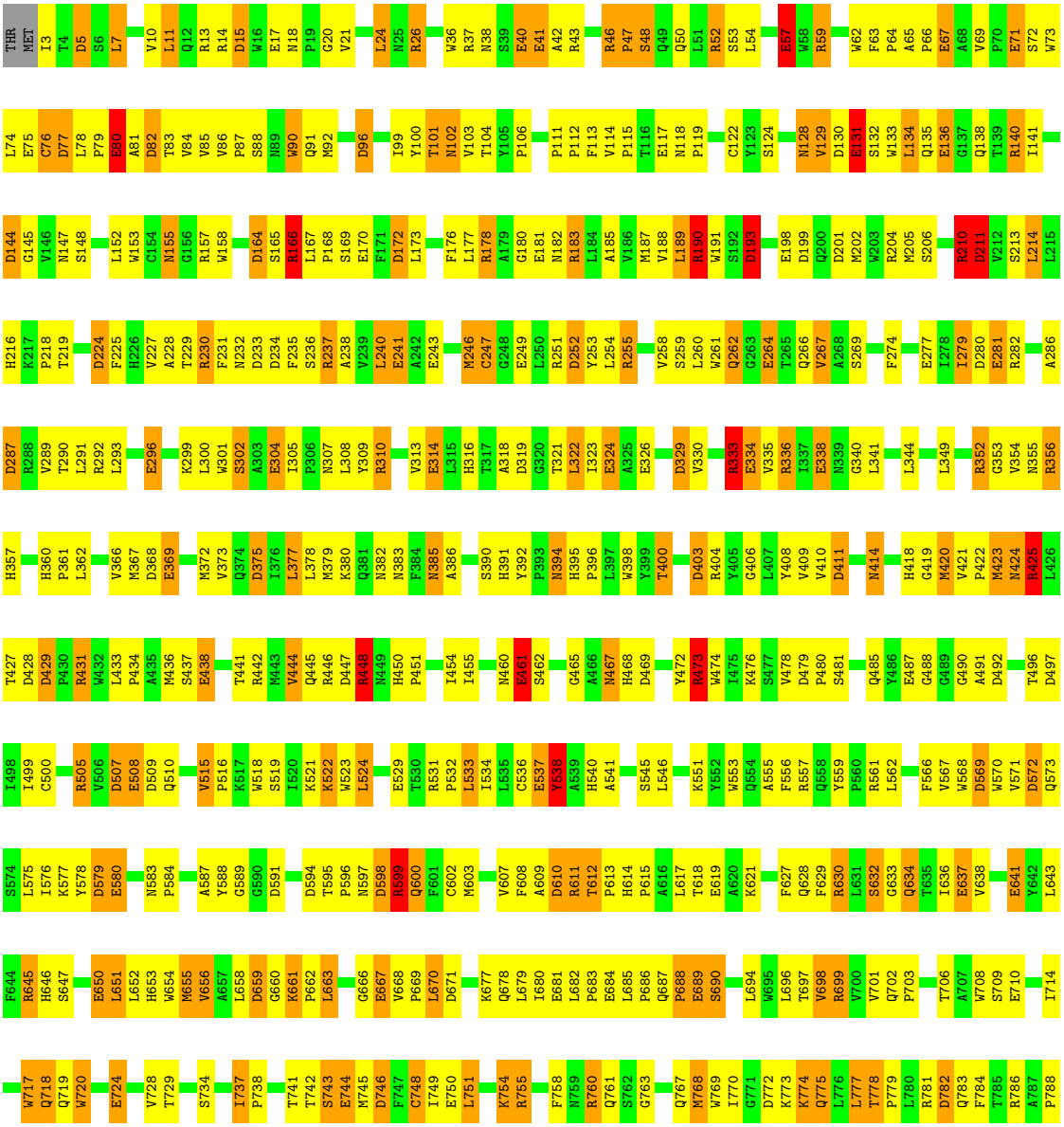


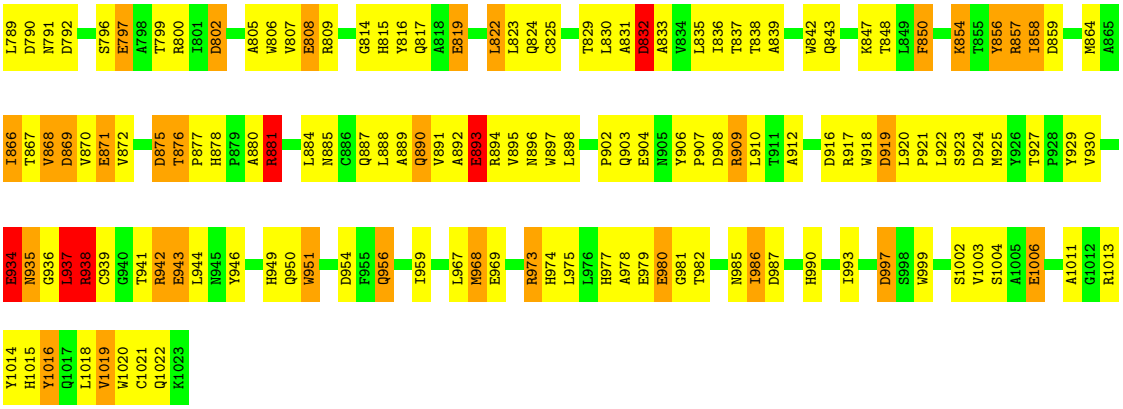


• Molecule 1: BETA-GALACTOSIDASE

Chain B:

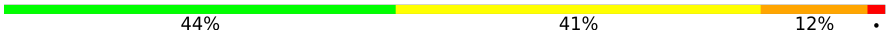


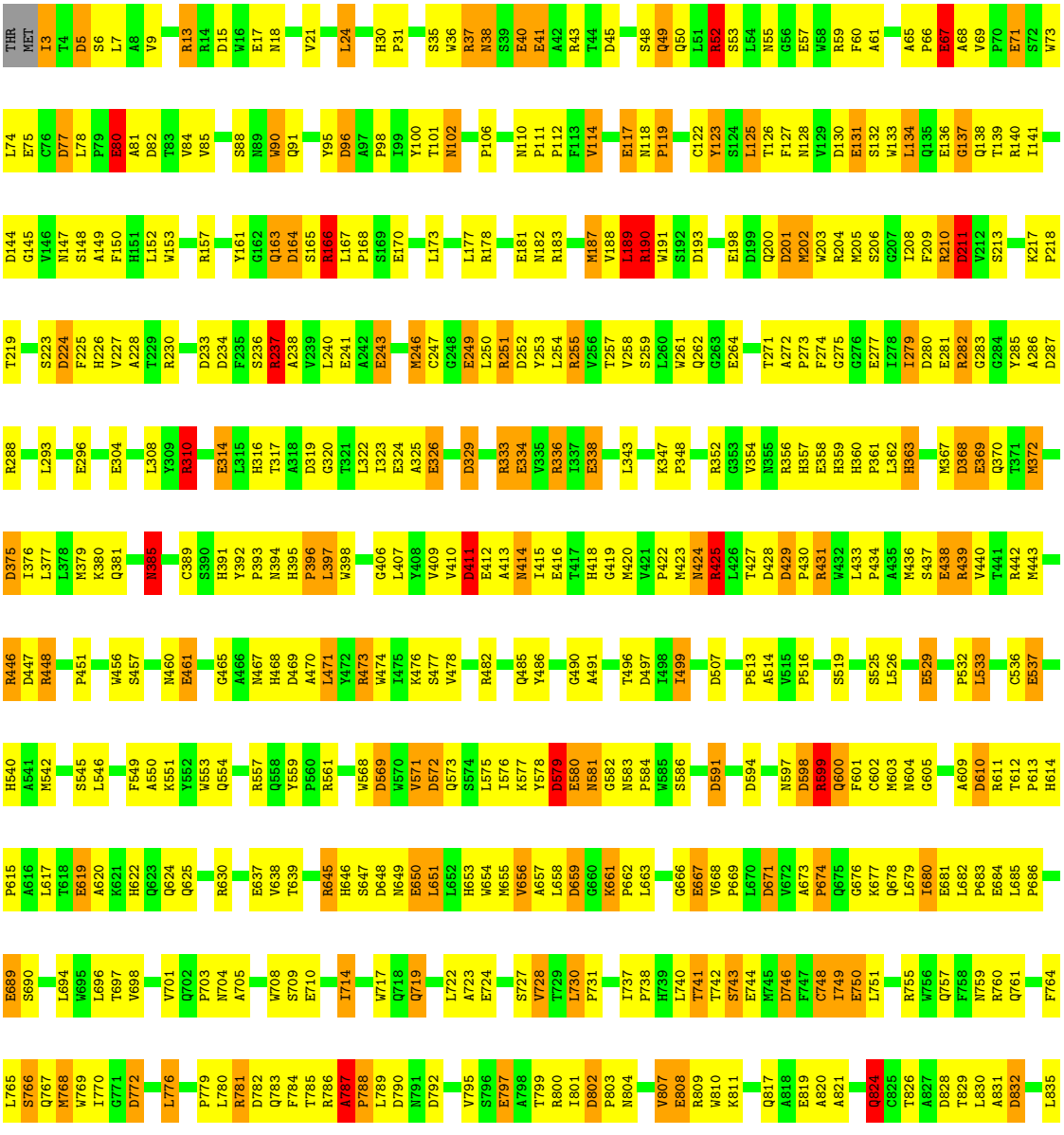


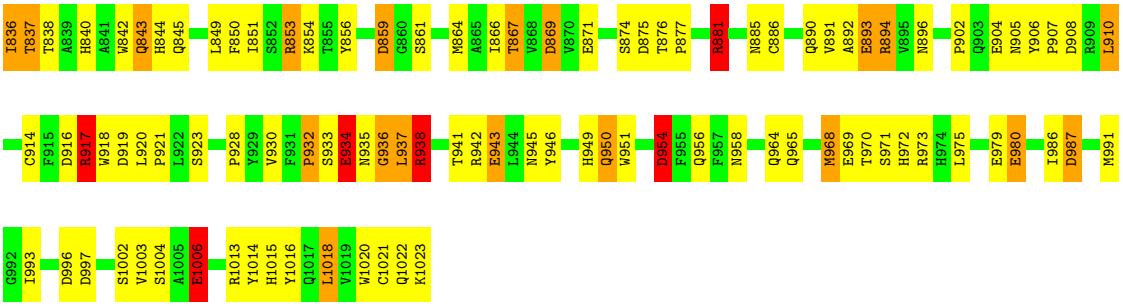


• Molecule 1: BETA-GALACTOSIDASE

Chain C:

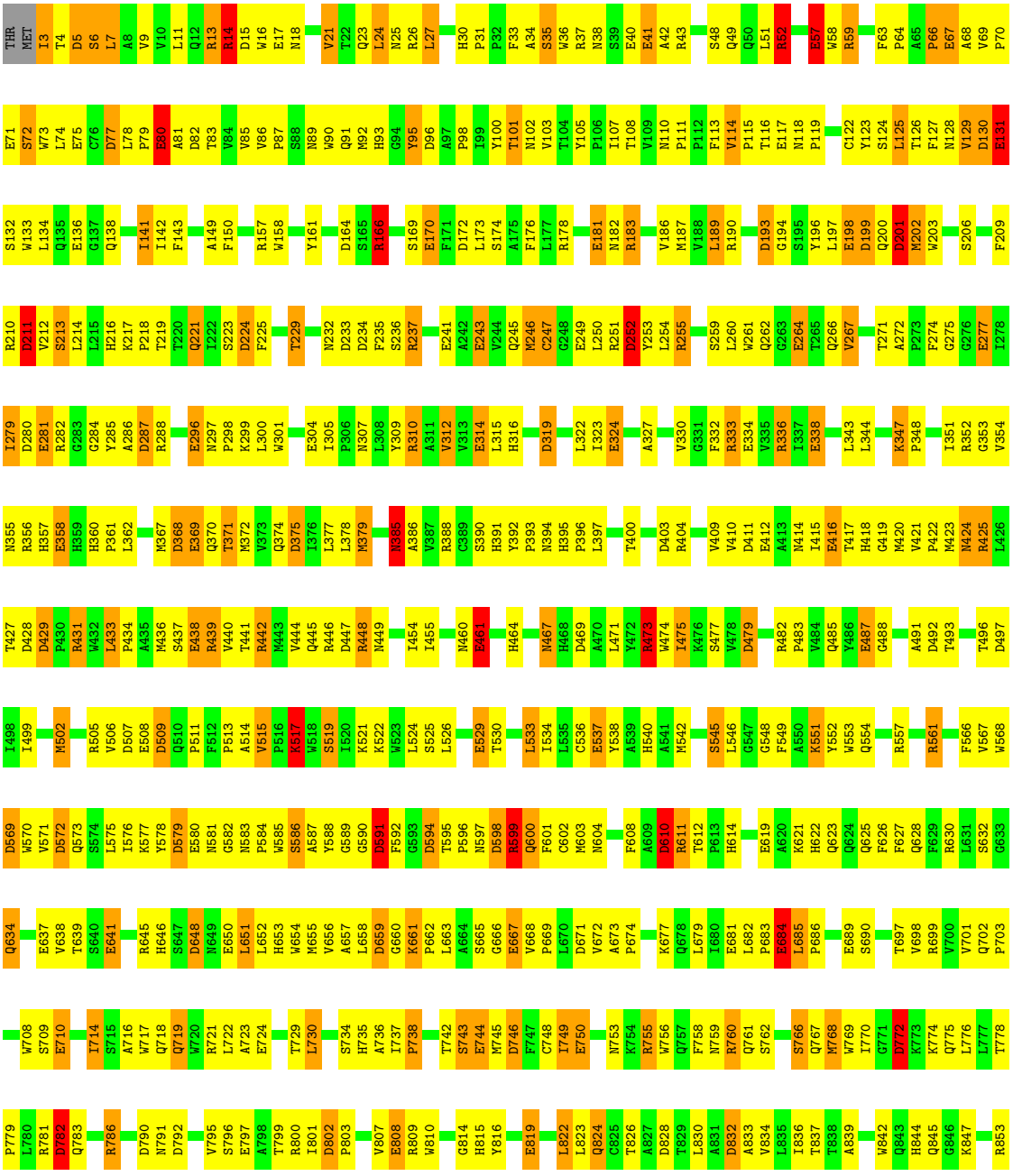


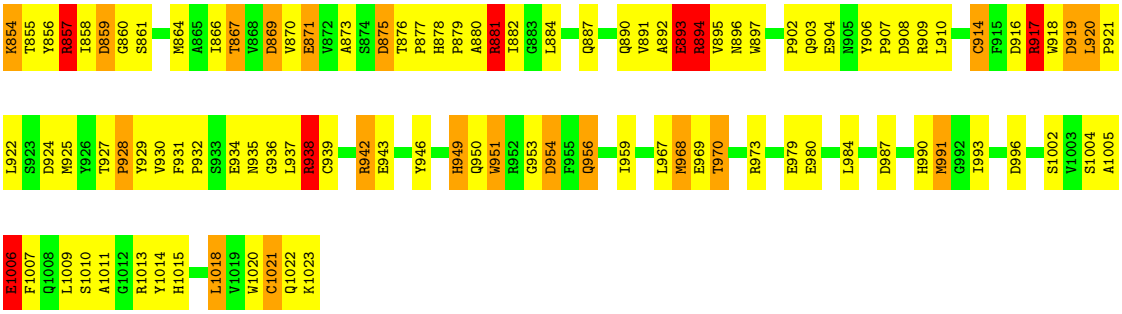


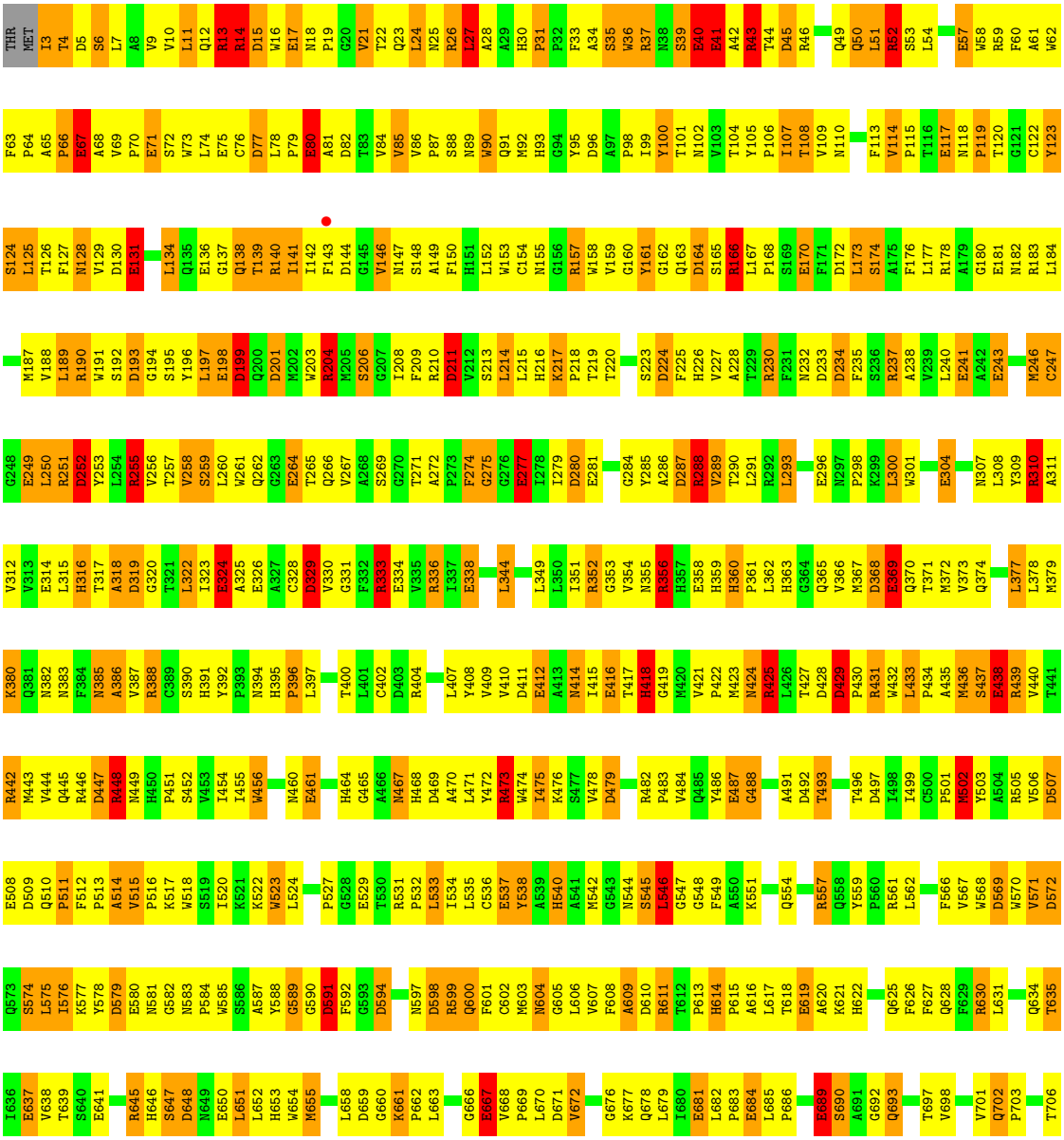


• Molecule 1: BETA-GALACTOSIDASE

Chain D:

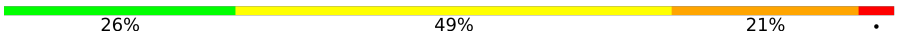


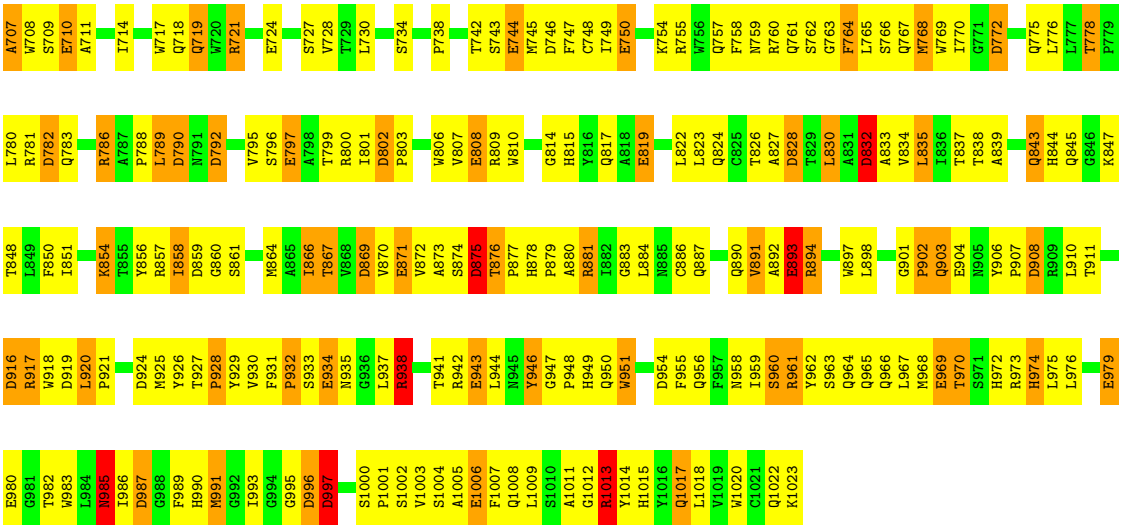




• Molecule 1: BETA-GALACTOSIDASE

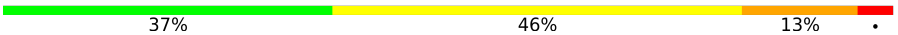
Chain E:

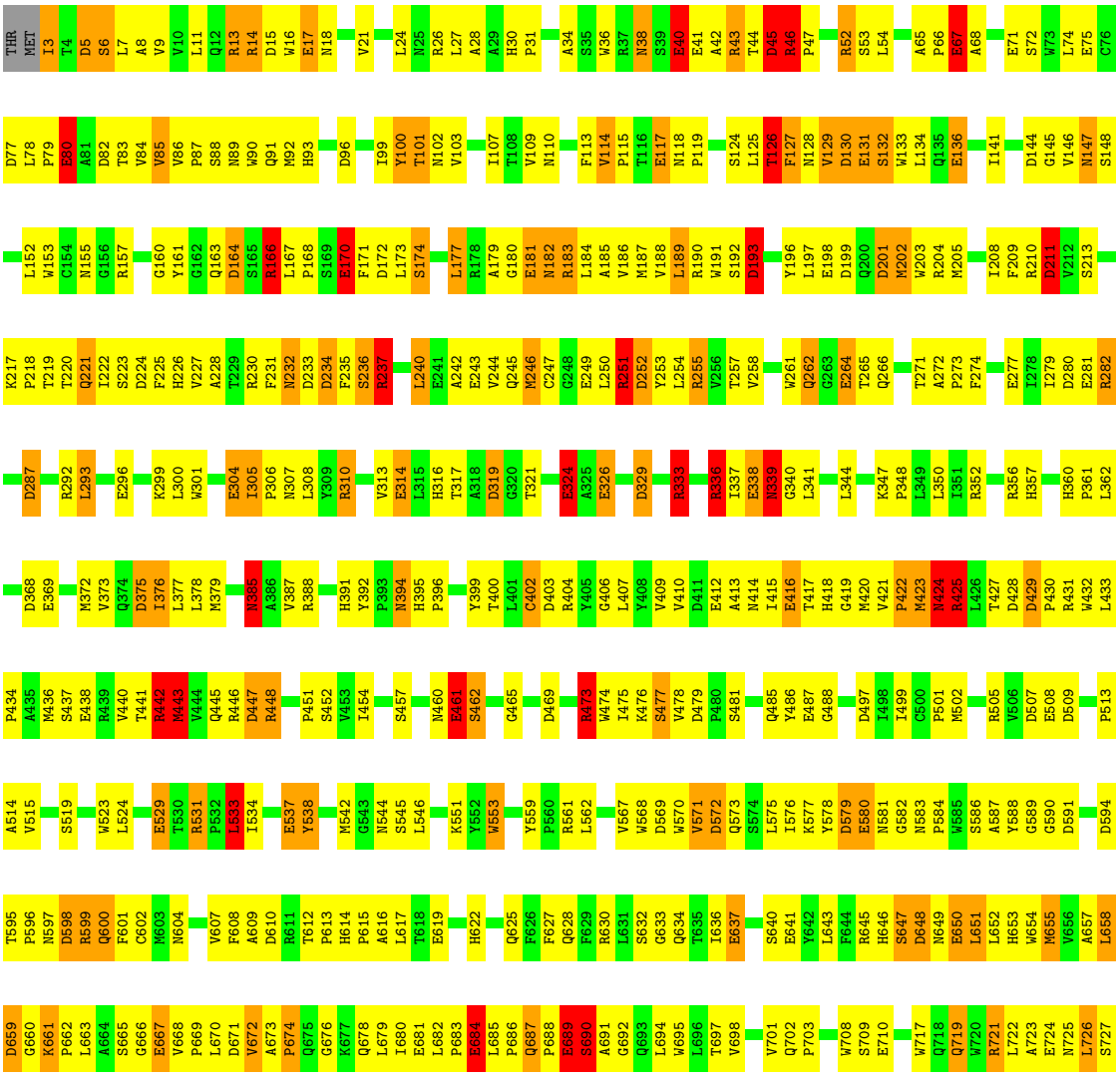


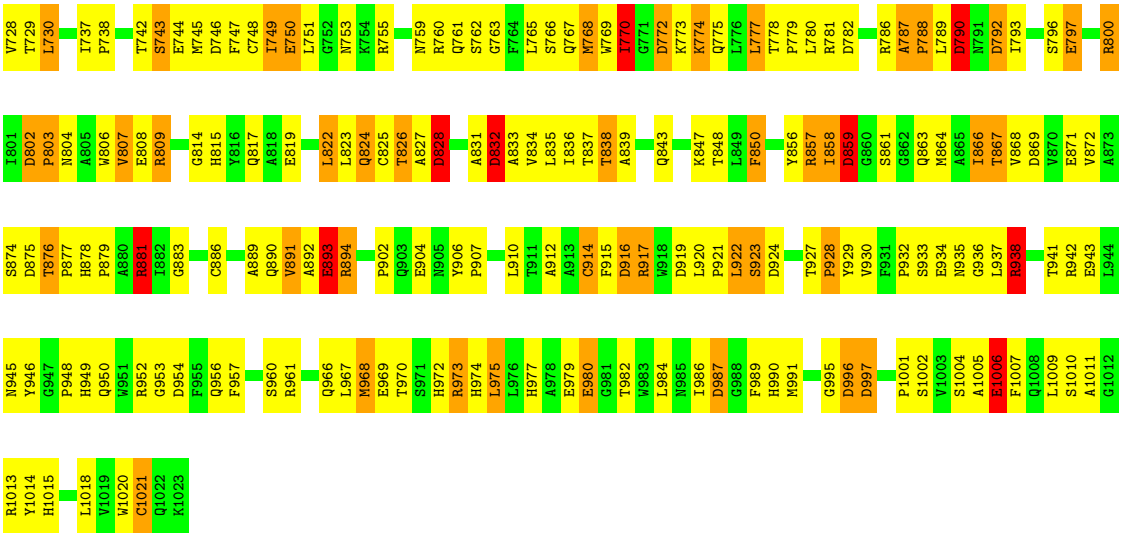


• Molecule 1: BETA-GALACTOSIDASE

Chain F:

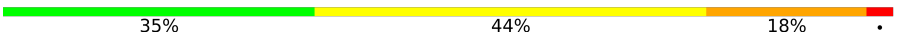


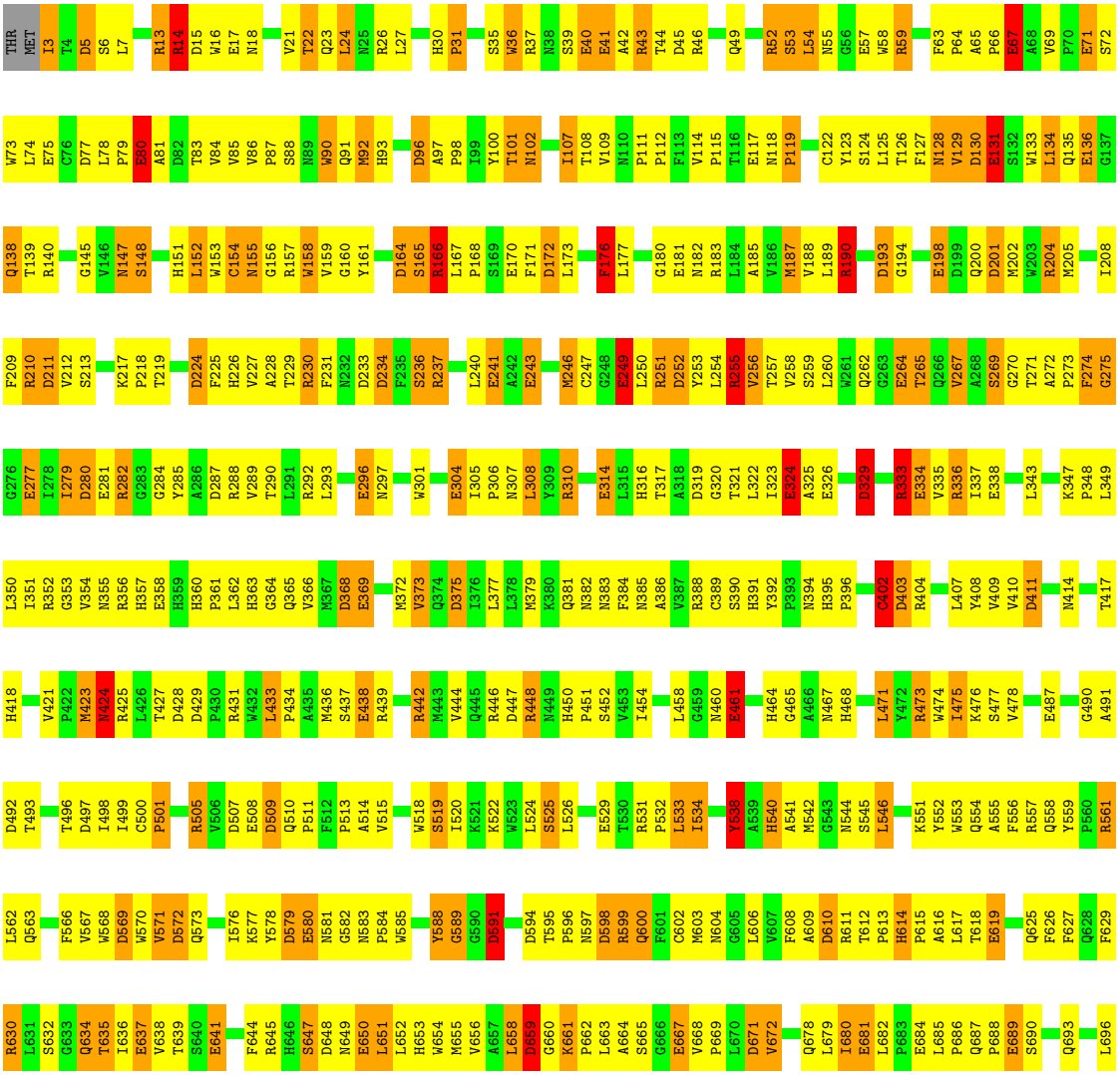


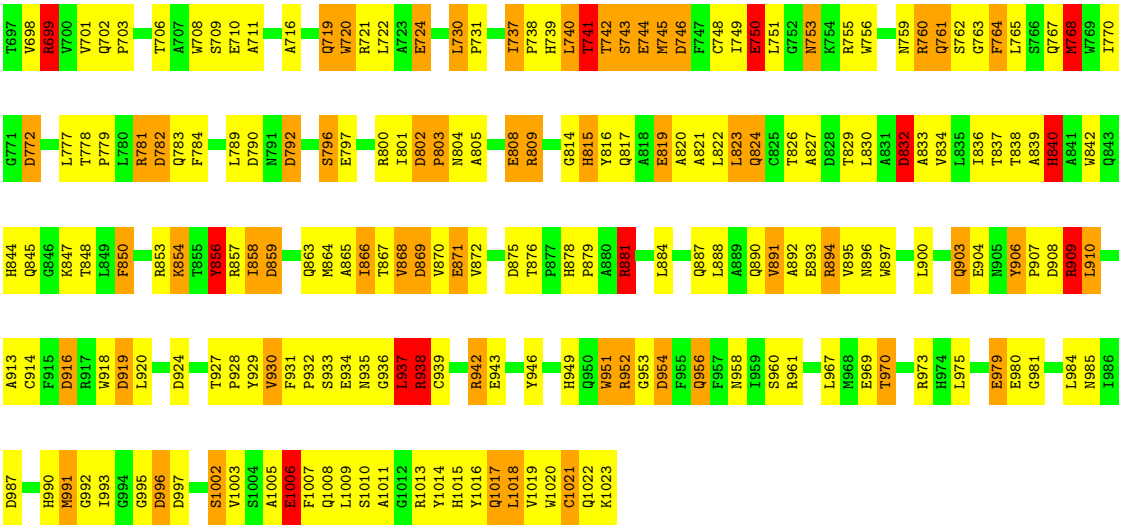


• Molecule 1: BETA-GALACTOSIDASE

Chain G:

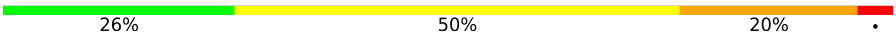


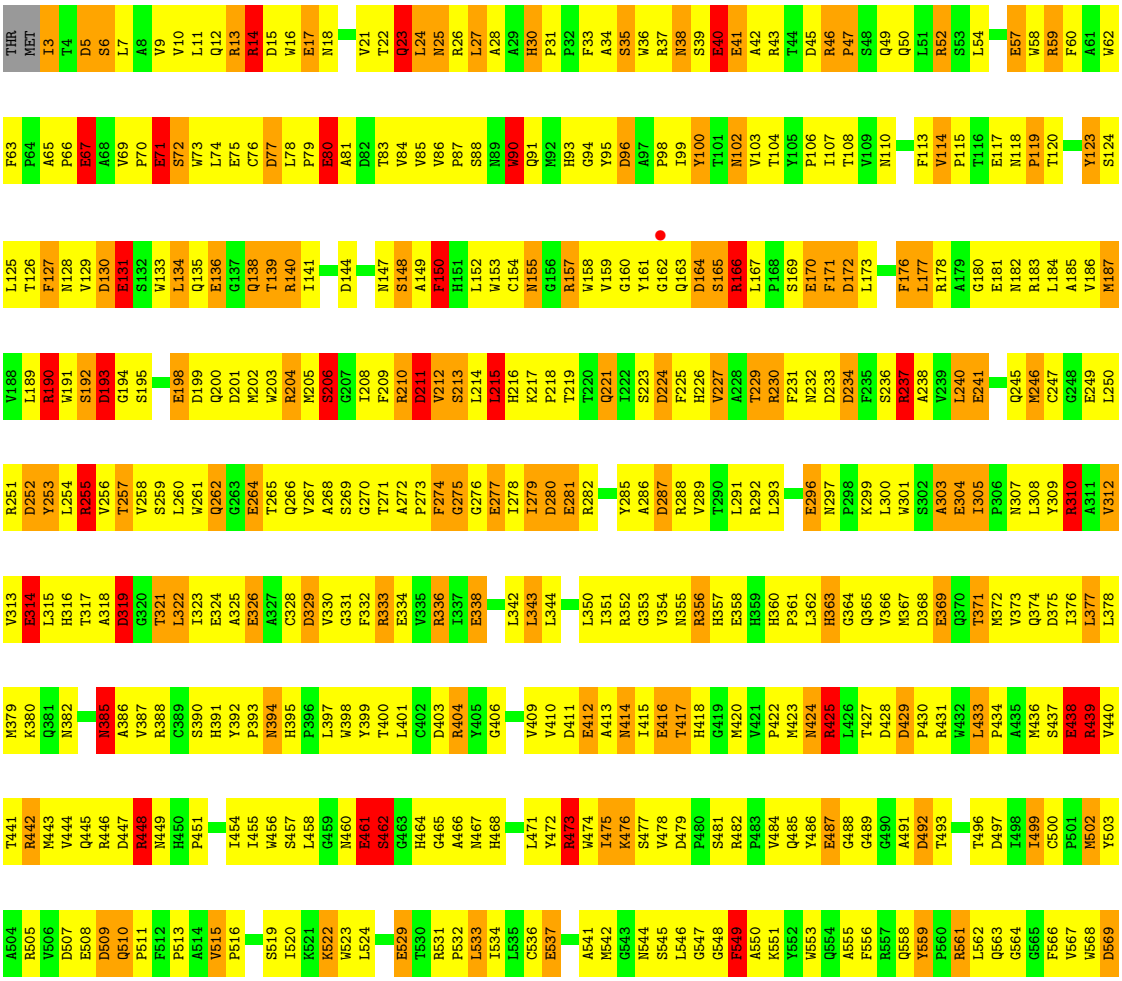


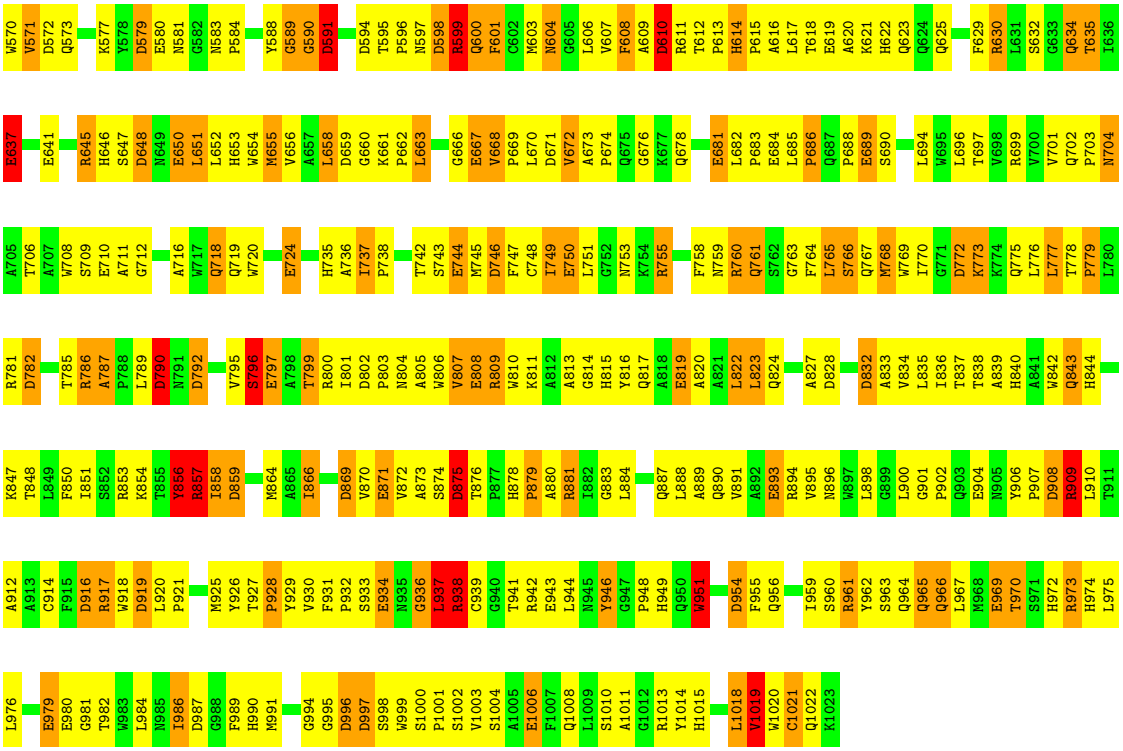


• Molecule 1: BETA-GALACTOSIDASE

Chain H:

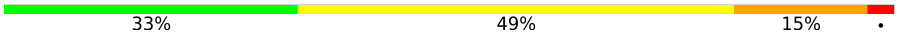


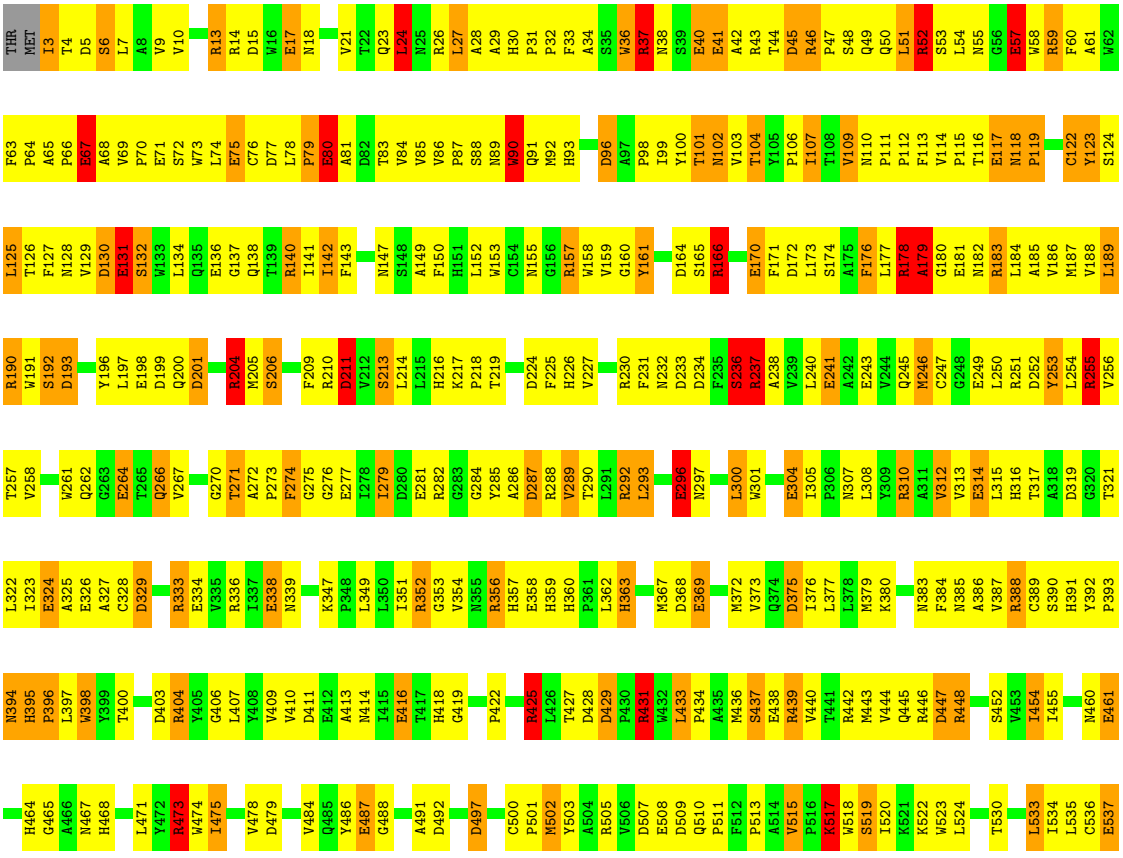


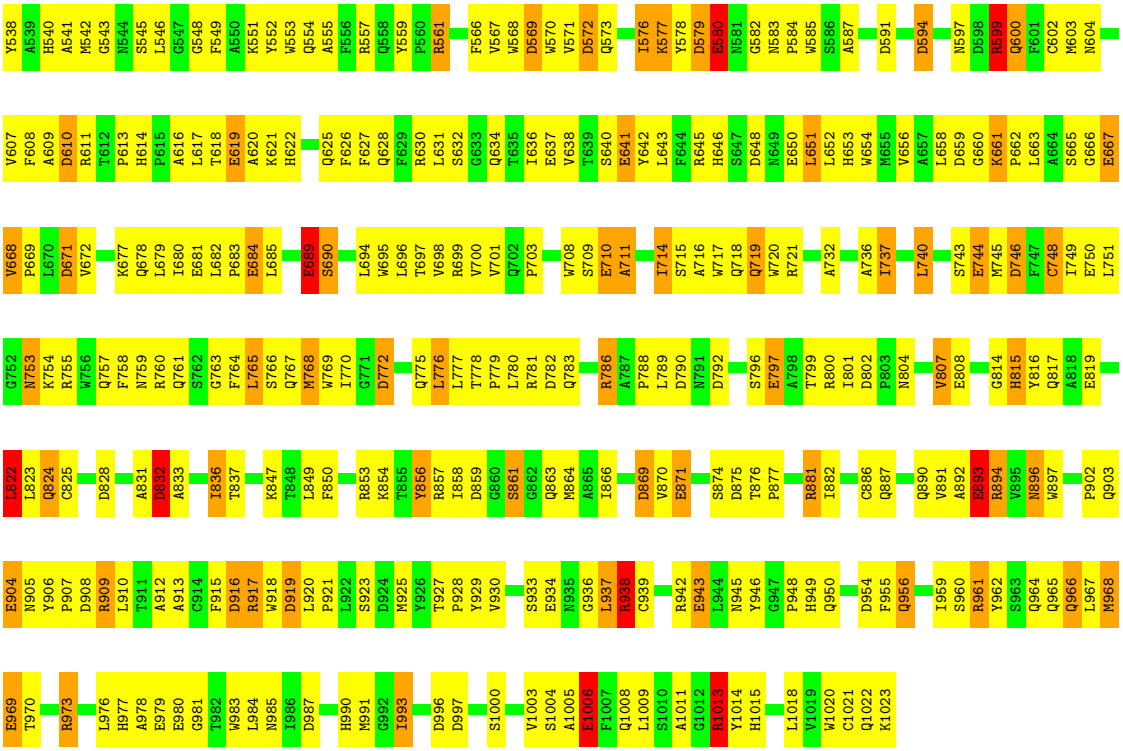


• Molecule 1: BETA-GALACTOSIDASE

Chain I:

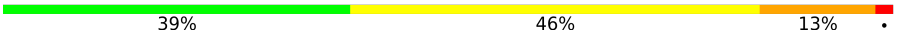


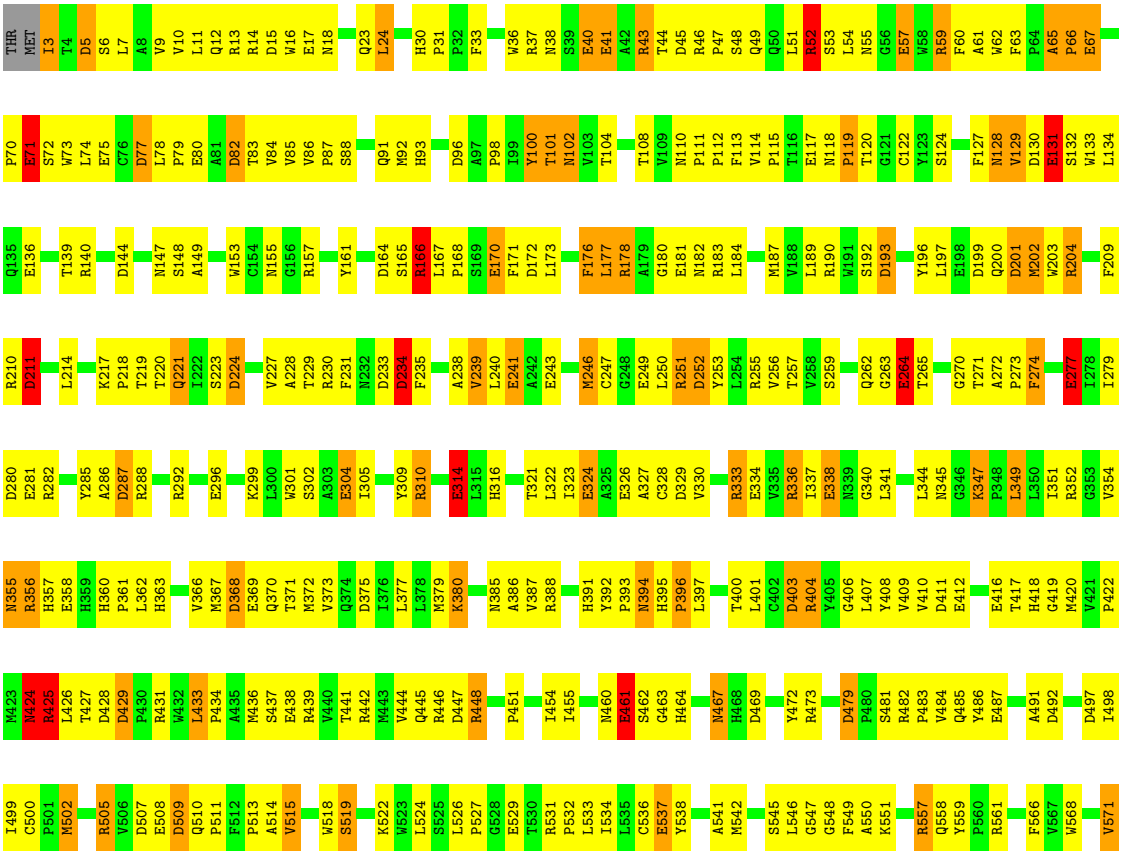


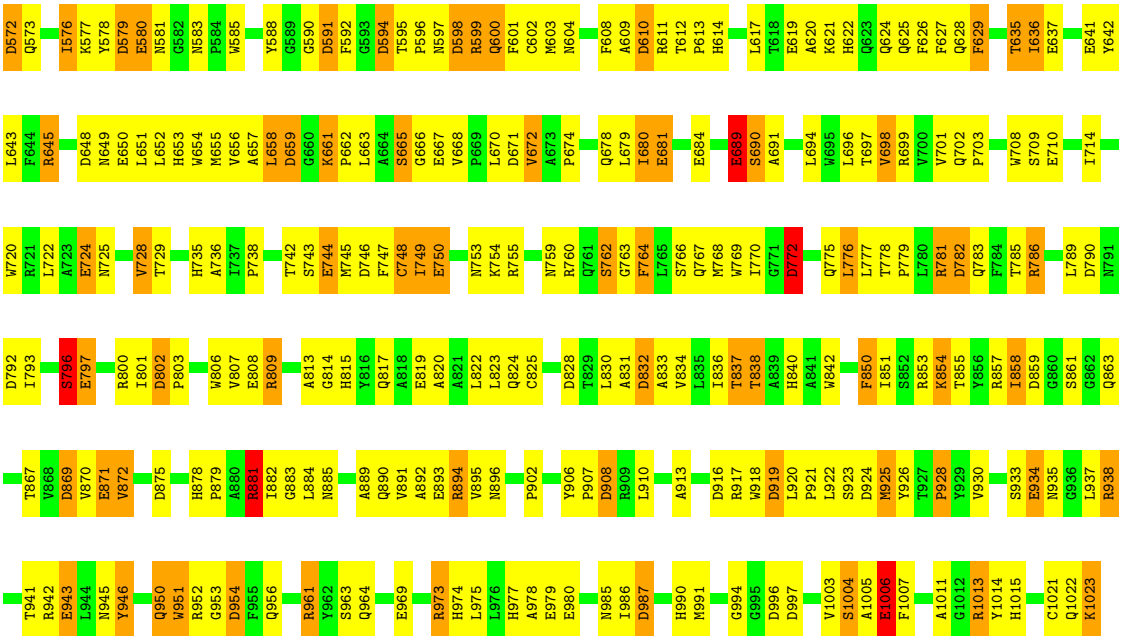


• Molecule 1: BETA-GALACTOSIDASE

Chain J:

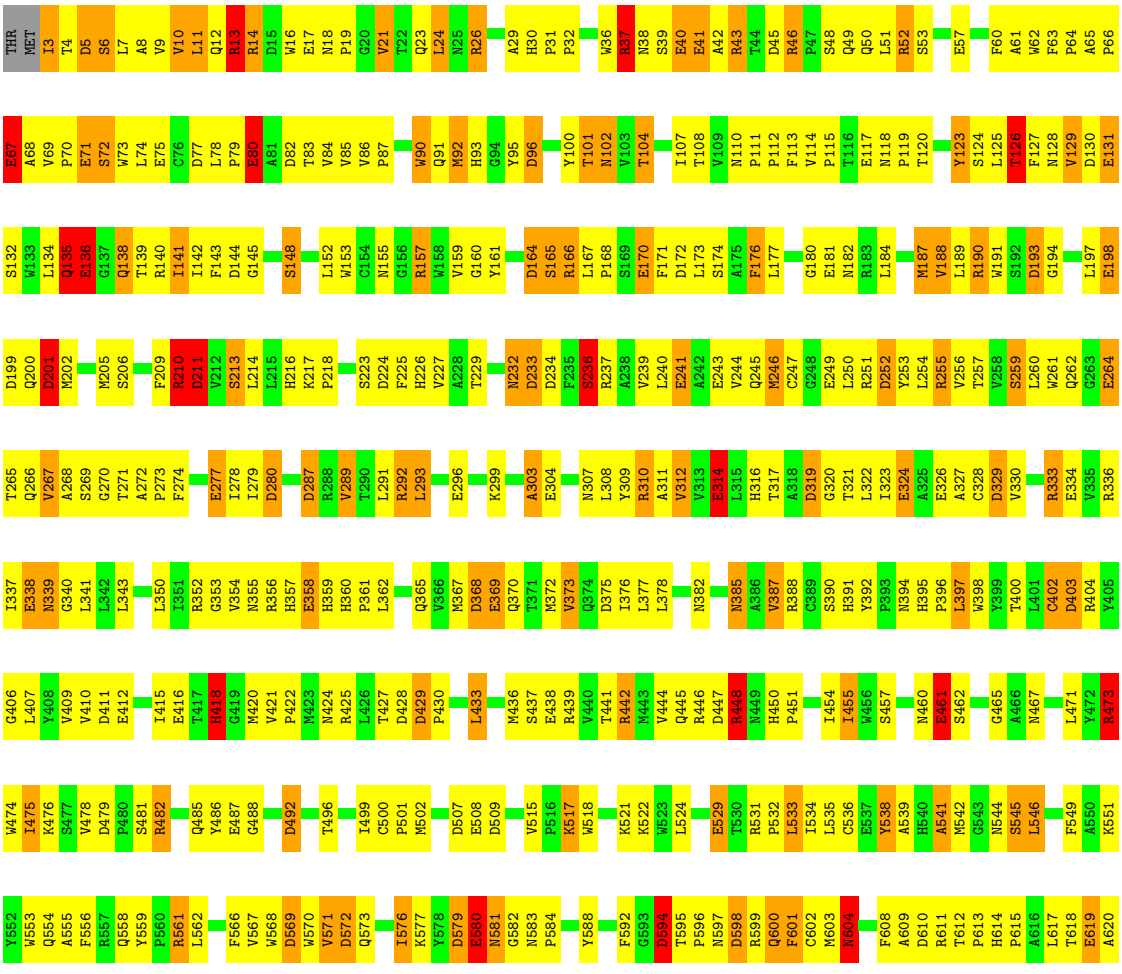


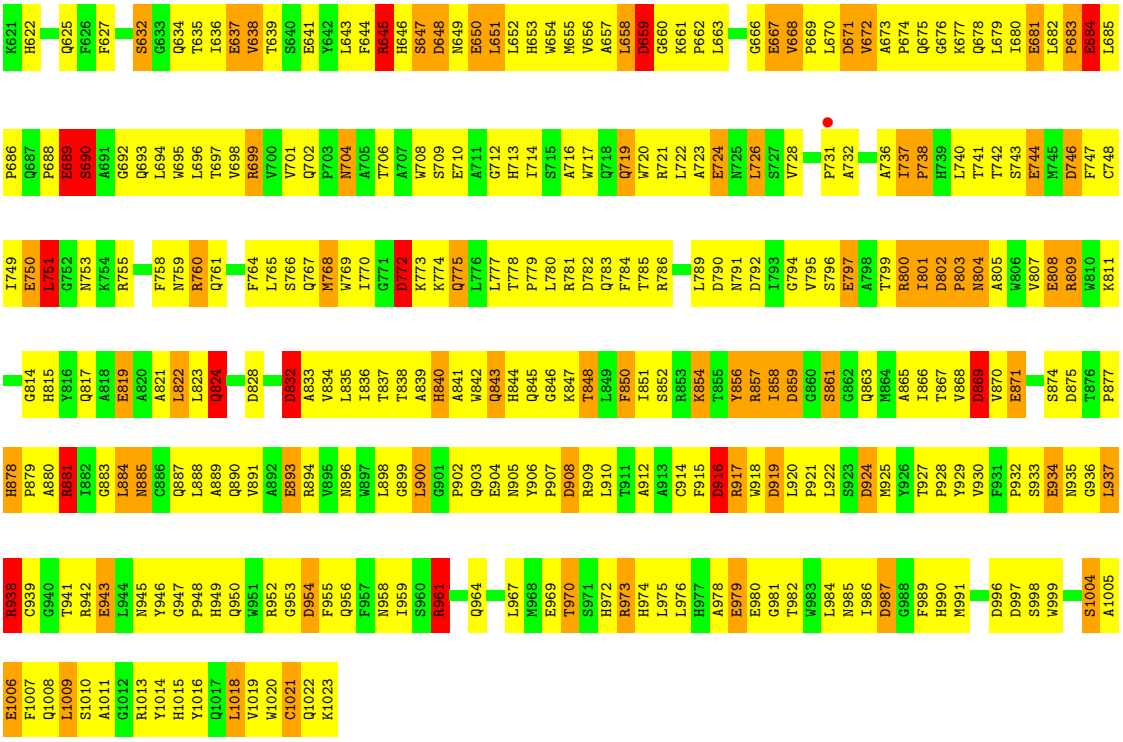




• Molecule 1: BETA-GALACTOSIDASE

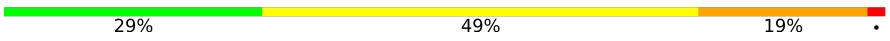
Chain K:

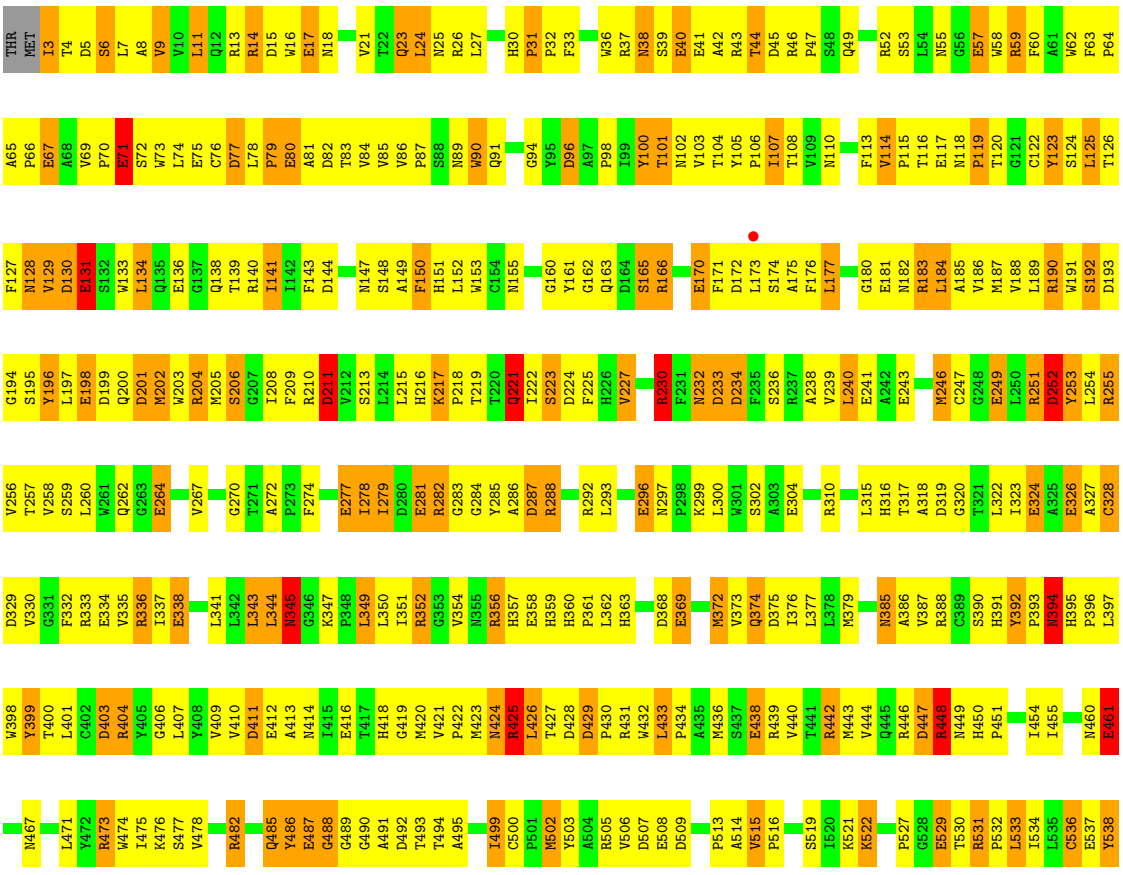


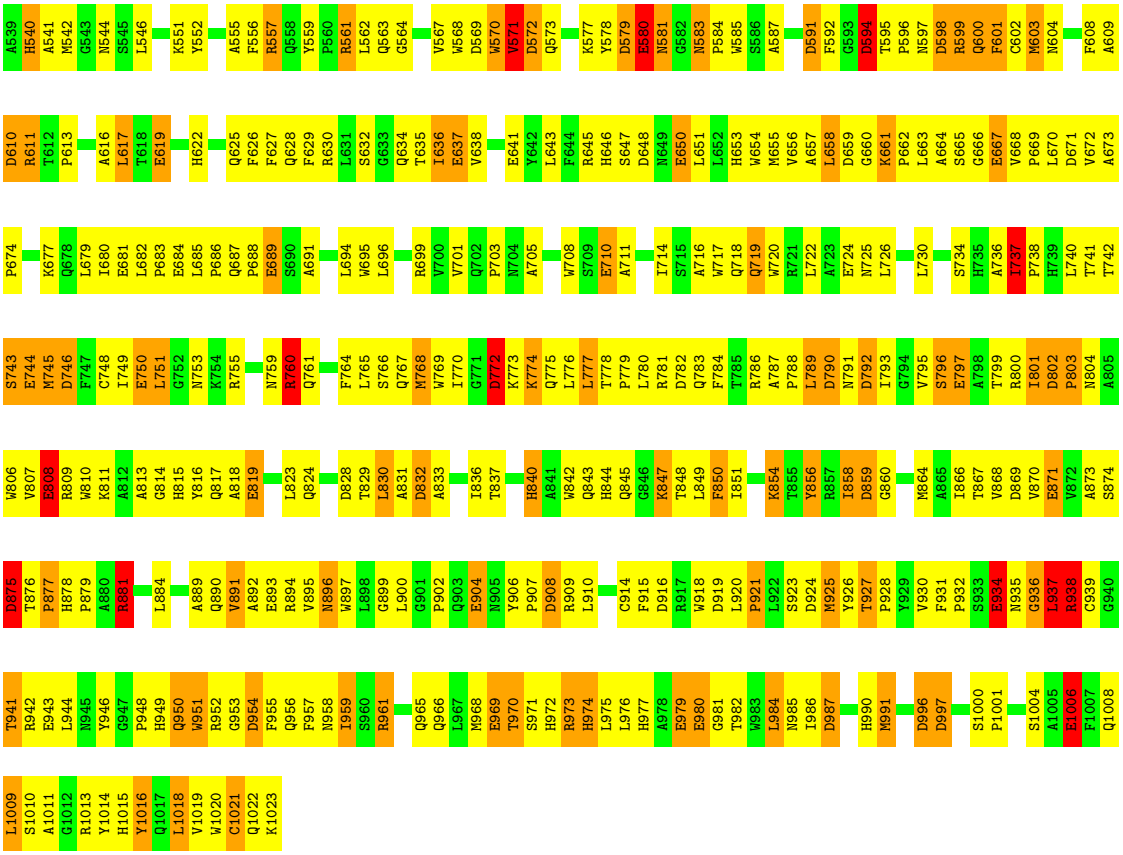


• Molecule 1: BETA-GALACTOSIDASE

Chain L:

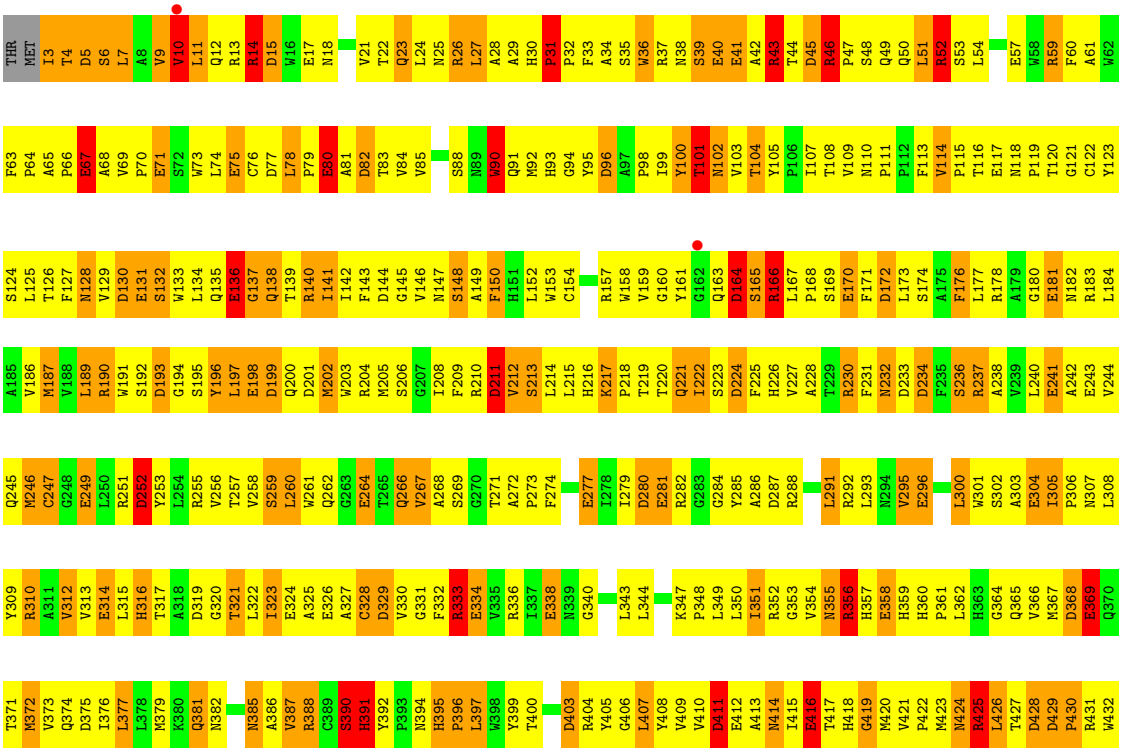


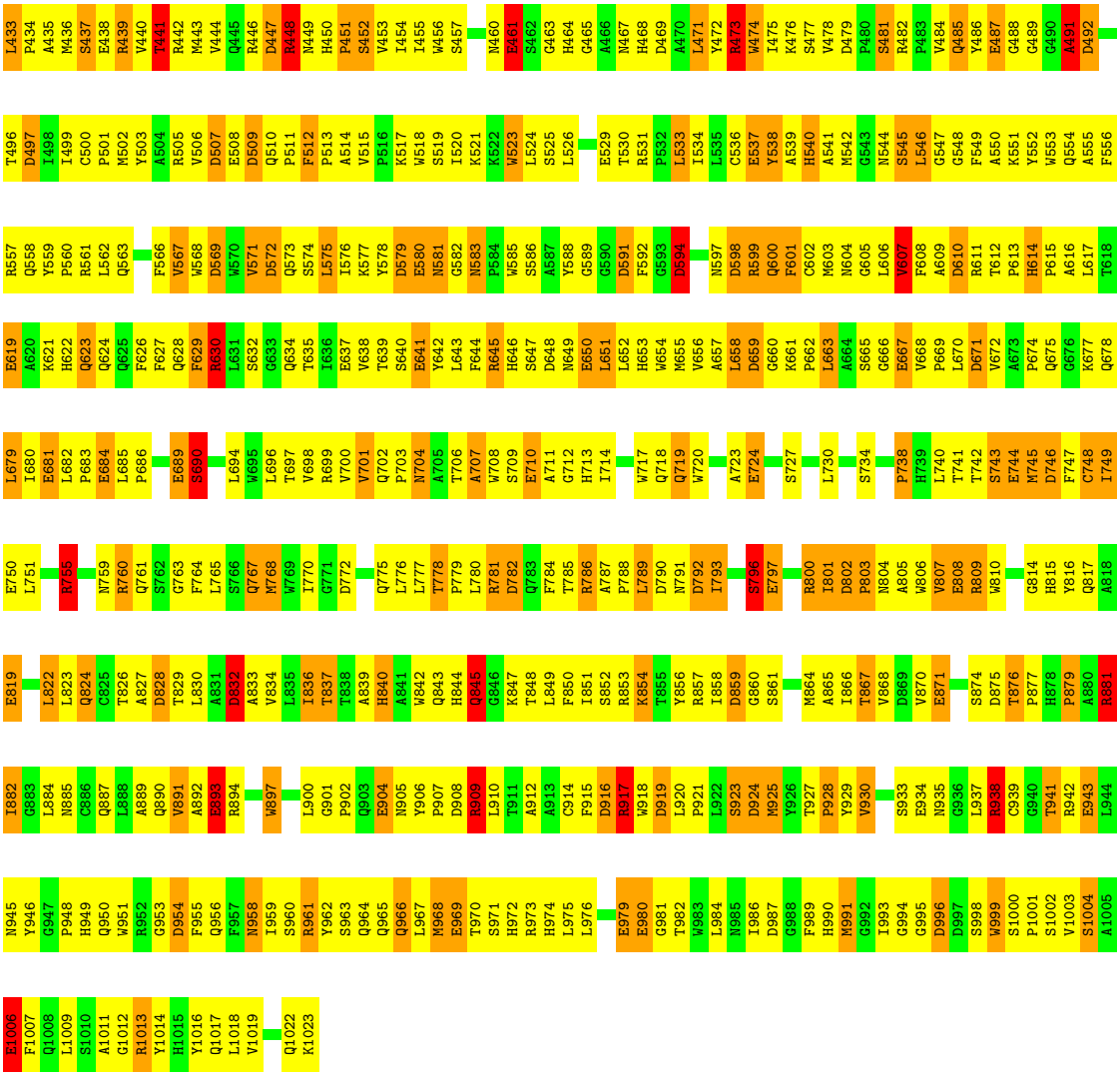




• Molecule 1: BETA-GALACTOSIDASE

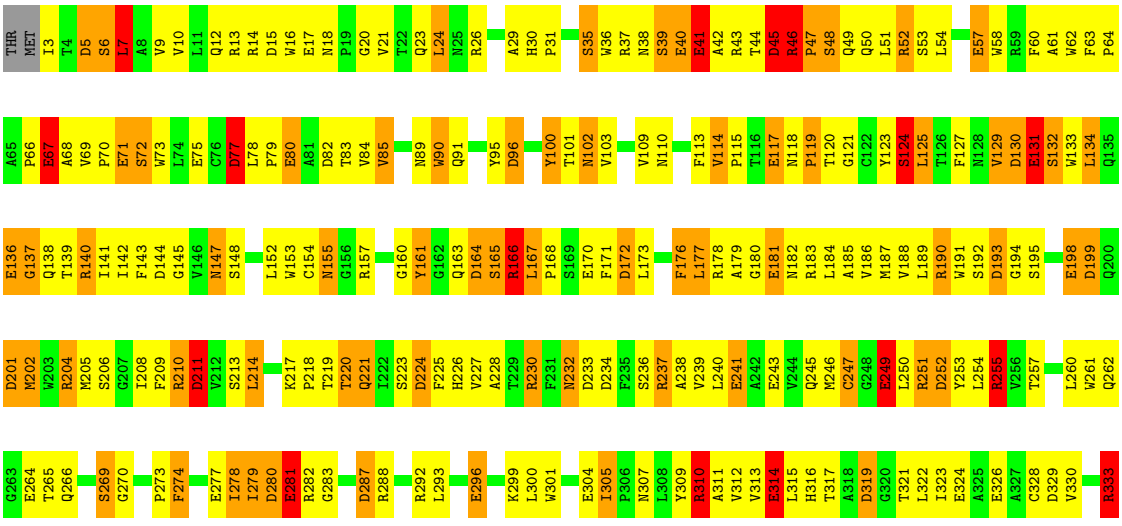
Chain M:

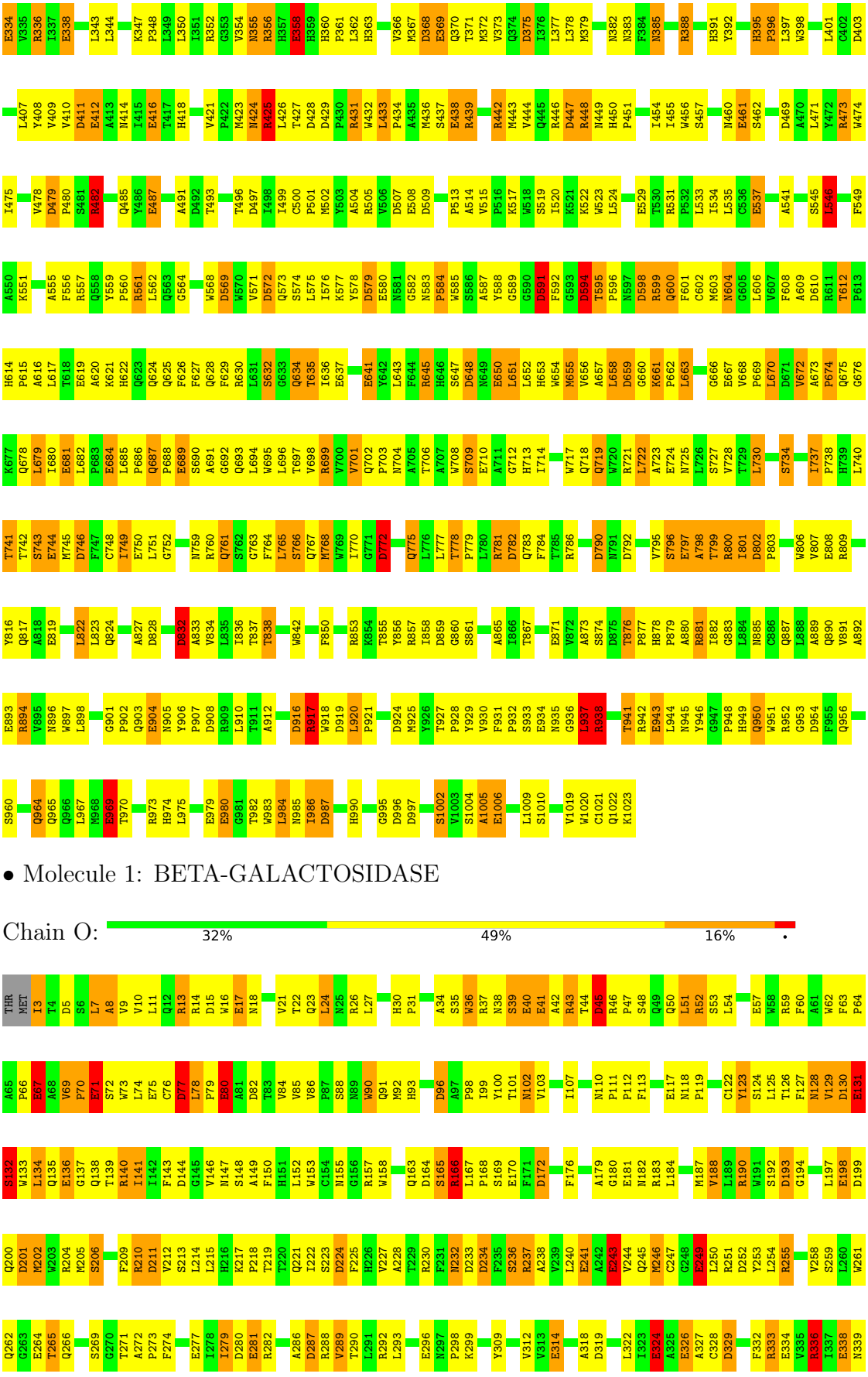




• Molecule 1: BETA-GALACTOSIDASE

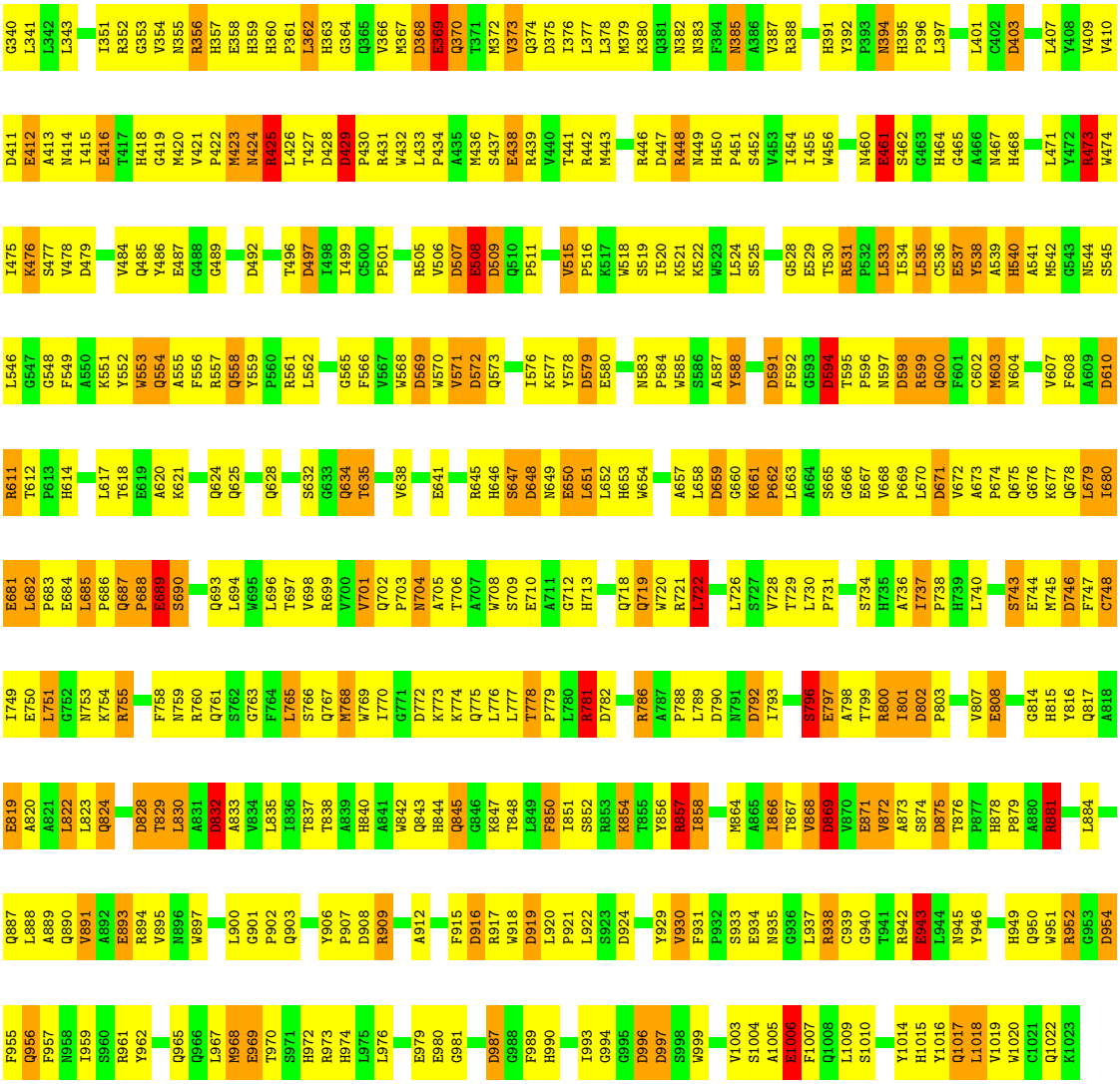
Chain N:





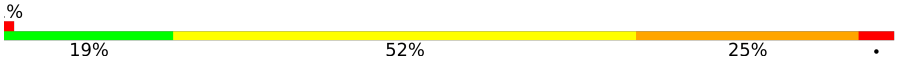
• Molecule 1: BETA-GALACTOSIDASE

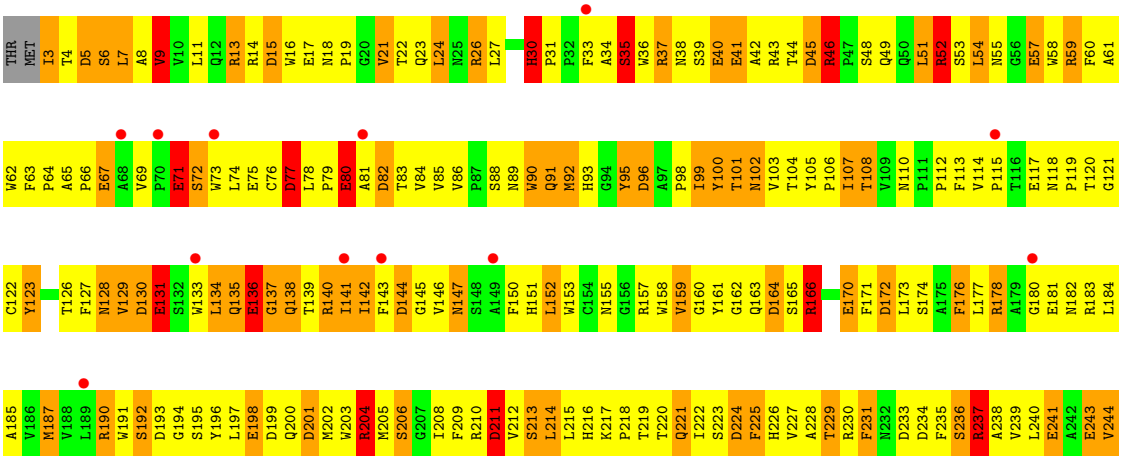
Chain O:

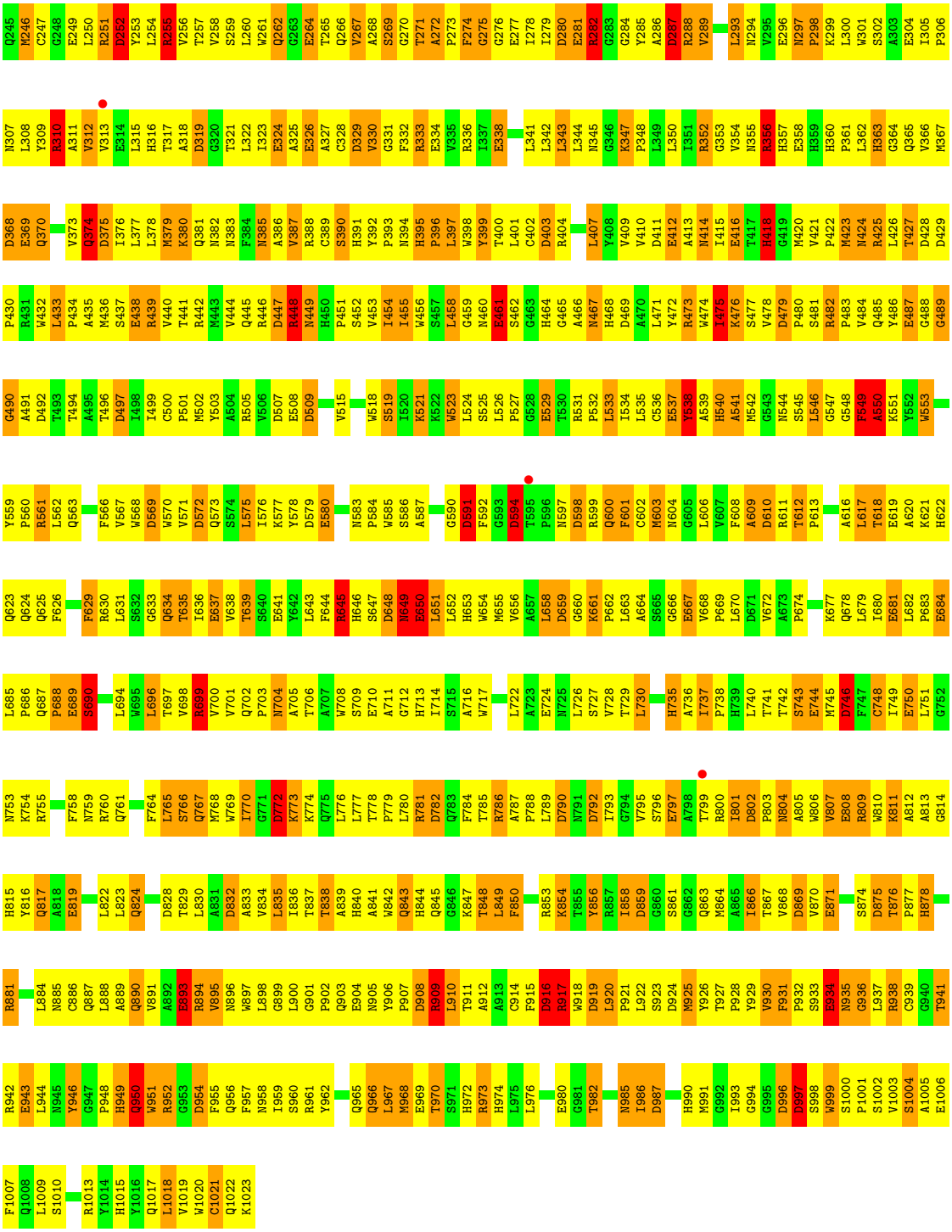


• Molecule 1: BETA-GALACTOSIDASE

Chain P:






Q245 M246 C247 G248 E249 L250 R251 D252 Y253 L254 R255 V256 T257 V258 S259 L260 W261 Q262 G263 E264 T265 Q266 V267 A268 S269 G270 T271 A272 P273 F274 G275 G276 E277 I278 I279 D280 E281 R282 G283 G284 Y285 A286 D287 R288 V289 L293 N294 V295 E296 N297 P298 K299 L300 W301 S302 A303 E304 I305 P306
N307 L308 Y309 R310 A311 V312 V313 E314 L315 H316 T317 A318 D319 G320 T321 L322 I323 E324 A325 E326 A327 C328 D329 V330 G331 F332 R333 E334 V335 R336 I337 E338 L341 L342 L343 L344 N345 G346 K347 P348 L349 L350 I351 R352 G353 V354 N355 R356 H357 E358 H359 H360 P361 L362 H363 G364 Q365 V366 M367
D368 E369 Q370 V373 Q374 D375 I376 L377 L378 M379 K380 Q381 N382 N383 F384 N385 A386 V387 R388 C389 S390 D391 Y392 P393 N394 H395 P396 L397 W398 Y399 T400 L401 C402 D403 R404 L407 Y408 V409 V410 D411 E412 A413 N414 I415 E416 T417 H418 G419 M420 V421 P422 M423 N424 R425 L426 T427 D428 D429
P430 R431 W432 L433 P434 A435 M436 S437 E438 R439 V440 T441 R442 M443 V444 Q445 R446 D447 R448 N449 H450 P451 S452 V453 I454 I455 W456 S457 L458 G459 N460 E461 S462 G463 H464 G465 A466 N467 H468 D469 A470 L471 Y472 R473 W474 I475 K476 S477 V478 D479 P480 S481 R482 P483 V484 Q485 Y486 E487 G488 G489
G490 A491 D492 T493 T494 A495 T496 D497 T498 I499 C500 P501 M502 Y503 A504 R505 V506 D507 E508 D509 V515 W518 S519 I520 K521 K522 W523 L524 S525 L526 P527 G528 E529 T530 R531 P532 L533 I534 L535 C536 E537 Y538 A539 H540 A541 M542 G543 N544 S545 L546 G547 G548 F549 A550 K551 Y552 W553
Y559 P560 R561 L562 Q563 F566 V567 W568 D569 W570 V571 D572 Q573 S574 L575 I576 K577 Y578 D579 E580 N583 P584 W585 S586 A587 G590 D591 F592 G593 D594 T595 P596 N597 D598 R599 Q600 F601 C602 W603 N604 G605 L606 V607 F608 A609 D610 R611 T612 P613 A616 L617 T618 E619 A620 K621 H622
Q623 Q624 Q625 F626 F629 R630 L631 S632 G633 Q634 T635 I636 E637 V638 T639 S640 E641 Y642 L643 F644 R645 H646 S647 D648 N649 E650 L651 L652 H653 W654 M655 V656 A657 L658 D659 G660 K661 P662 L663 A664 S665 G666 E667 V668 P669 L670 D671 V672 A673 P674 K677 Q678 L679 I680 E681 L682 P683 E684
L685 P686 Q687 P688 E689 E690 L694 W695 L696 T697 V698 R699 V700 V701 Q702 P703 N704 A705 T706 A707 W708 S709 E710 A711 G712 H713 I714 S715 A716 W717 L722 A723 E724 N725 L726 S727 V728 T729 L730 H735 A736 I737 P738 H739 L740 T741 T742 S743 E744 M745 D746 F747 C748 I749 E750 L751 G752
N753 K754 R755 F758 N759 R760 Q761 F764 L765 S766 Q767 M768 W769 I770 G771 D772 K773 K774 Q775 L776 L777 T778 P779 L780 R781 D782 Q783 F784 T785 R786 A787 P788 L789 D790 N791 D792 I793 G794 V795 S796 E797 A798 T799 R800 I801 D802 P803 N804 A805 W806 V807 E808 R809 W810 K811 A812 A813 G814
H815 Y816 Q817 A818 E819 L822 L823 Q824 D828 T829 L830 A831 D832 A833 V834 L835 I836 T837 T838 A839 H840 A841 W842 Q843 H844 Q845 G846 K847 T848 L849 F850 R853 K854 T855 Y856 R857 I858 D859 G860 S861 G862 Q863 M864 A865 I866 T867 V868 D869 V870 E871 S874 D875 T876 P877 H878
R881 L884 N885 C886 Q887 L888 A889 Q890 V891 A892 E893 R894 V895 N896 W897 L898 G899 L900 G901 P902 Q903 E904 N905 Y906 P907 D908 R909 L910 T911 A912 A913 C914 F915 D916 R917 W918 D919 L920 P921 L922 S923 D924 M925 Y926 T927 P928 Y929 V930 F931 P932 S933 E934 N935 G936 L937 R938 C939 G940 T941
R942 E943 L944 N945 Y946 G947 P948 H949 Q950 W951 R952 G953 D954 F955 Q956 F957 N958 I959 S960 R961 Y962 Q965 Q966 L967 M968 E969 T970 S971 H972 R973 H974 L975 L976 E980 G981 T982 N985 I986 D987 H990 M991 G992 I993 G994 G995 D996 D997 S998 W999 S1000 P1001 S1002 V1003 S1004 A1005 E1006
F1007 Q1008 L1009 S1010 R1013 Y1014 H1015 Y1016 Q1017 L1018 V1019 W1020 C1021 Q1022 K1023

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	107.90Å 207.50Å 509.90Å 90.00° 94.70° 90.00°	Depositor
Resolution (Å)	8.00 – 2.50 92.62 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-2.50) 39.3 (92.62-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.00Å)	Xtriage
Refinement program	TNT 5D, TNT V. 5-D	Depositor
R, R_{free}	0.174 , (Not available) 0.169 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	27.6	Xtriage
Anisotropy	0.197	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 89.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.009 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	132654	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.98% 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

5.1 Standard geometry

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

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	1.16	51/8440 (0.6%)	1.52	139/11516 (1.2%)
1	B	1.17	54/8440 (0.6%)	1.51	130/11516 (1.1%)
1	C	1.18	56/8440 (0.7%)	1.50	132/11516 (1.1%)
1	D	1.16	55/8440 (0.7%)	1.52	148/11516 (1.3%)
1	E	1.16	55/8440 (0.7%)	1.56	145/11516 (1.3%)
1	F	1.18	45/8440 (0.5%)	1.53	144/11516 (1.3%)
1	G	1.16	58/8440 (0.7%)	1.51	151/11516 (1.3%)
1	H	1.16	56/8440 (0.7%)	1.57	150/11516 (1.3%)
1	I	1.13	53/8440 (0.6%)	1.52	140/11516 (1.2%)
1	J	1.12	48/8440 (0.6%)	1.48	134/11516 (1.2%)
1	K	1.09	53/8440 (0.6%)	1.45	115/11516 (1.0%)
1	L	1.13	53/8440 (0.6%)	1.54	134/11516 (1.2%)
1	M	1.16	55/8440 (0.7%)	1.58	142/11516 (1.2%)
1	N	1.14	51/8440 (0.6%)	1.49	127/11516 (1.1%)
1	O	1.11	54/8440 (0.6%)	1.49	141/11516 (1.2%)
1	P	1.17	57/8440 (0.7%)	1.60	151/11516 (1.3%)
All	All	1.15	854/135040 (0.6%)	1.52	2223/184256 (1.2%)

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	1	0
1	B	1	0
1	D	2	1
1	E	1	0
1	F	2	0
1	G	2	0
1	H	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	I	1	0
1	J	1	0
1	L	1	0
1	M	2	0
1	P	2	0
All	All	17	1

The worst 5 of 854 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	358	GLU	CD-OE2	11.52	1.38	1.25
1	F	75	GLU	CD-OE1	10.30	1.36	1.25
1	K	358	GLU	CD-OE2	9.34	1.35	1.25
1	B	650	GLU	CD-OE1	9.31	1.35	1.25
1	F	326	GLU	CD-OE2	9.01	1.35	1.25

The worst 5 of 2223 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	881	ARG	NE-CZ-NH2	-14.45	113.08	120.30
1	B	166	ARG	NE-CZ-NH2	-14.23	113.19	120.30
1	N	561	ARG	NE-CZ-NH2	-14.06	113.27	120.30
1	L	425	ARG	NE-CZ-NH2	13.74	127.17	120.30
1	L	997	ASP	CB-CG-OD2	-13.64	106.02	118.30

5 of 17 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	90	TRP	CA
1	B	718	GLN	CA
1	D	95	TYR	CA
1	D	914	CYS	CA
1	E	118	ASN	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	473	ARG	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8198	0	7796	540	0
1	B	8198	0	7796	546	0
1	C	8198	0	7796	446	0
1	D	8198	0	7796	551	0
1	E	8198	0	7795	892	0
1	F	8198	0	7796	579	0
1	G	8198	0	7796	641	0
1	H	8198	0	7796	882	0
1	I	8198	0	7796	618	0
1	J	8198	0	7795	507	0
1	K	8198	0	7796	781	0
1	L	8198	0	7796	792	0
1	M	8198	0	7796	1078	0
1	N	8198	0	7795	630	0
1	O	8198	0	7796	659	0
1	P	8198	0	7796	1151	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
2	G	2	0	0	0	0
2	H	2	0	0	0	0
2	I	2	0	0	0	0
2	J	2	0	0	0	0
2	K	2	0	0	0	0
2	L	2	0	0	0	0
2	M	1	0	0	0	0
2	N	2	0	0	0	0
2	O	2	0	0	0	0
2	P	2	0	0	0	0
3	A	88	0	0	7	0
3	B	96	0	0	14	0
3	C	91	0	0	9	0
3	D	97	0	0	13	0
3	E	94	0	0	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	91	0	0	9	0
3	G	95	0	0	13	0
3	H	92	0	0	18	0
3	I	90	0	0	15	0
3	J	97	0	0	9	0
3	K	87	0	0	9	0
3	L	84	0	0	12	0
3	M	79	0	0	17	0
3	N	94	0	0	19	0
3	O	95	0	0	12	0
3	P	85	0	0	21	0
All	All	132654	0	124733	11096	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 43.

The worst 5 of 11096 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:E:7:LEU:HD13	1:E:74:LEU:HD11	1.23	1.17
1:C:427:THR:HA	1:C:436:MET:HE1	1.21	1.16
1:D:572:ASP:HB3	1:D:603:MET:HG2	1.25	1.16
1:A:770:ILE:HD11	1:A:1022:GLN:HG2	1.18	1.15
1:I:316:HIS:HA	1:I:323:ILE:HD13	1.24	1.13

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 X-ray entries followed by that with respect to entries of similar resolution.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	1019/1023 (100%)	909 (89%)	92 (9%)	18 (2%)	8 14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	1019/1023 (100%)	914 (90%)	85 (8%)	20 (2%)	7 12
1	C	1019/1023 (100%)	919 (90%)	81 (8%)	19 (2%)	8 13
1	D	1019/1023 (100%)	910 (89%)	98 (10%)	11 (1%)	14 26
1	E	1019/1023 (100%)	842 (83%)	135 (13%)	42 (4%)	3 3
1	F	1019/1023 (100%)	892 (88%)	101 (10%)	26 (3%)	5 8
1	G	1019/1023 (100%)	893 (88%)	101 (10%)	25 (2%)	5 8
1	H	1019/1023 (100%)	845 (83%)	140 (14%)	34 (3%)	4 5
1	I	1019/1023 (100%)	884 (87%)	111 (11%)	24 (2%)	6 9
1	J	1019/1023 (100%)	887 (87%)	118 (12%)	14 (1%)	11 20
1	K	1019/1023 (100%)	855 (84%)	131 (13%)	33 (3%)	4 5
1	L	1019/1023 (100%)	838 (82%)	146 (14%)	35 (3%)	3 5
1	M	1019/1023 (100%)	836 (82%)	127 (12%)	56 (6%)	2 2
1	N	1019/1023 (100%)	875 (86%)	117 (12%)	27 (3%)	5 8
1	O	1019/1023 (100%)	889 (87%)	102 (10%)	28 (3%)	5 7
1	P	1019/1023 (100%)	797 (78%)	164 (16%)	58 (6%)	1 1
All	All	16304/16368 (100%)	13985 (86%)	1849 (11%)	470 (3%)	4 6

5 of 470 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ASP
1	A	277	GLU
1	A	389	CYS
1	A	541	ALA
1	A	659	ASP

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	873/875 (100%)	723 (83%)	150 (17%)	2 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	873/875 (100%)	709 (81%)	164 (19%)	1	2
1	C	873/875 (100%)	754 (86%)	119 (14%)	3	7
1	D	873/875 (100%)	729 (84%)	144 (16%)	2	4
1	E	873/875 (100%)	686 (79%)	187 (21%)	1	2
1	F	873/875 (100%)	735 (84%)	138 (16%)	2	4
1	G	873/875 (100%)	717 (82%)	156 (18%)	2	3
1	H	873/875 (100%)	693 (79%)	180 (21%)	1	2
1	I	873/875 (100%)	716 (82%)	157 (18%)	1	3
1	J	873/875 (100%)	755 (86%)	118 (14%)	4	7
1	K	873/875 (100%)	722 (83%)	151 (17%)	2	3
1	L	873/875 (100%)	704 (81%)	169 (19%)	1	2
1	M	873/875 (100%)	677 (78%)	196 (22%)	1	1
1	N	873/875 (100%)	717 (82%)	156 (18%)	2	3
1	O	873/875 (100%)	715 (82%)	158 (18%)	1	3
1	P	873/875 (100%)	665 (76%)	208 (24%)	0	1
All	All	13968/14000 (100%)	11417 (82%)	2551 (18%)	1	3

5 of 2551 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	75	GLU
1	O	704	ASN
1	M	321	THR
1	M	67	GLU
1	N	279	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 370 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	316	HIS
1	M	485	GLN
1	K	604	ASN
1	L	624	GLN
1	M	949	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 31 ligands modelled in this entry, 31 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	1021/1023 (99%)	-1.08	0 100 100	2, 24, 54, 79	0
1	B	1021/1023 (99%)	-1.10	0 100 100	2, 23, 54, 81	0
1	C	1021/1023 (99%)	-1.06	0 100 100	4, 21, 52, 80	0
1	D	1021/1023 (99%)	-1.10	0 100 100	4, 25, 56, 81	0
1	E	1021/1023 (99%)	-0.94	1 (0%) 95 96	8, 34, 61, 86	0
1	F	1021/1023 (99%)	-1.03	0 100 100	2, 23, 55, 78	0
1	G	1021/1023 (99%)	-1.07	0 100 100	3, 27, 58, 82	0
1	H	1021/1023 (99%)	-0.92	1 (0%) 95 96	6, 33, 61, 89	0
1	I	1021/1023 (99%)	-1.02	0 100 100	4, 30, 59, 80	0
1	J	1021/1023 (99%)	-1.02	0 100 100	8, 28, 56, 85	0
1	K	1021/1023 (99%)	-0.89	1 (0%) 95 96	10, 35, 64, 92	0
1	L	1021/1023 (99%)	-0.86	1 (0%) 95 96	6, 35, 63, 84	0
1	M	1021/1023 (99%)	-0.80	2 (0%) 95 95	12, 39, 65, 80	0
1	N	1021/1023 (99%)	-0.97	0 100 100	9, 30, 59, 89	0
1	O	1021/1023 (99%)	-1.00	0 100 100	11, 31, 60, 81	0
1	P	1021/1023 (99%)	-0.47	15 (1%) 73 75	14, 43, 69, 89	0
All	All	16336/16368 (99%)	-0.96	21 (0%) 95 96	2, 30, 60, 92	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	81	ALA	4.1
1	P	313	VAL	3.4
1	P	143	PHE	3.3
1	P	70	PRO	3.1
1	P	141	ILE	3.0

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)

There are no monosaccharides in this entry.

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, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	F	1101	1/1	0.78	0.19	35,35,35,35	0
2	MG	K	1101	1/1	0.89	0.08	34,34,34,34	0
2	MG	G	1101	1/1	0.91	0.17	32,32,32,32	0
2	MG	O	1101	1/1	0.92	0.12	40,40,40,40	0
2	MG	C	1102	1/1	0.93	0.08	28,28,28,28	0
2	MG	P	1101	1/1	0.93	0.12	49,49,49,49	0
2	MG	J	1101	1/1	0.94	0.14	34,34,34,34	0
2	MG	M	1101	1/1	0.94	0.14	56,56,56,56	0
2	MG	L	1102	1/1	0.95	0.04	28,28,28,28	0
2	MG	D	1102	1/1	0.95	0.13	42,42,42,42	0
2	MG	E	1102	1/1	0.95	0.08	32,32,32,32	0
2	MG	A	1102	1/1	0.95	0.08	37,37,37,37	0
2	MG	H	1101	1/1	0.96	0.15	27,27,27,27	0
2	MG	E	1101	1/1	0.96	0.15	39,39,39,39	0
2	MG	J	1102	1/1	0.96	0.05	29,29,29,29	0
2	MG	B	1101	1/1	0.96	0.15	25,25,25,25	0
2	MG	N	1101	1/1	0.97	0.14	32,32,32,32	0
2	MG	D	1101	1/1	0.97	0.12	28,28,28,28	0
2	MG	C	1101	1/1	0.97	0.18	23,23,23,23	0
2	MG	A	1101	1/1	0.98	0.15	37,37,37,37	0
2	MG	K	1102	1/1	0.98	0.05	25,25,25,25	0
2	MG	N	1102	1/1	0.98	0.11	26,26,26,26	0
2	MG	L	1101	1/1	0.98	0.14	31,31,31,31	0
2	MG	I	1102	1/1	0.98	0.08	33,33,33,33	0
2	MG	F	1102	1/1	0.99	0.07	26,26,26,26	0
2	MG	H	1102	1/1	0.99	0.06	22,22,22,22	0
2	MG	I	1101	1/1	0.99	0.13	33,33,33,33	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	B	1102	1/1	0.99	0.05	23,23,23,23	0
2	MG	O	1102	1/1	0.99	0.09	15,15,15,15	0
2	MG	G	1102	1/1	0.99	0.03	18,18,18,18	0
2	MG	P	1102	1/1	0.99	0.05	26,26,26,26	0

6.5 Other polymers (i)

There are no such residues in this entry.