



wwPDB X-ray Structure Validation Summary Report

Jun 16, 2014 – 08:15 PM BST

PDB ID : 4V41
Title : E. COLI (LAC Z) BETA-GALACTOSIDASE (NCS CONSTRAINED MONOMER-MONOCLINIC)
Authors : Juers, D.H.; Jacobson, R.H.; Wigley, D.; Zhang, X.J.; Huber, R.E.; Tronrud, D.E.; Matthews, B.W.
Deposited on : 2000-06-07
Resolution : 2.50 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

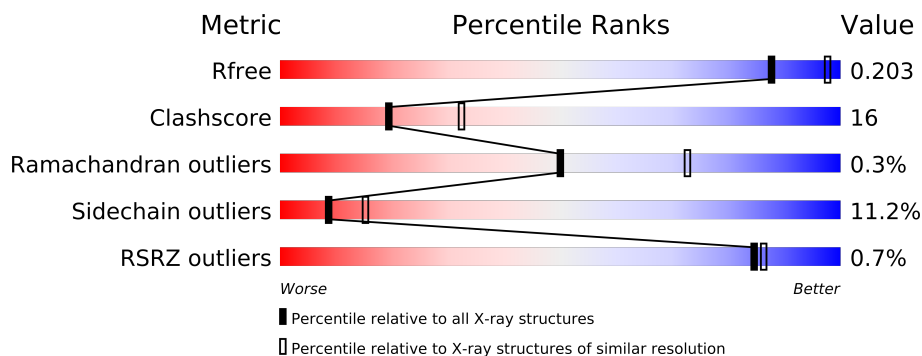
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.16 November 2013
Xtriage (Phenix) : dev-1323
EDS : stable23397
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23397

1 Overall quality at a glance

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))
R_{free}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1023	
1	B	1023	
1	C	1023	
1	D	1023	
1	E	1023	
1	F	1023	
1	G	1023	
1	H	1023	
1	I	1023	
1	J	1023	
1	K	1023	
1	L	1023	
1	M	1023	
1	N	1023	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	O	1023	
1	P	1023	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 138704 atoms, of which 0 are hydrogen and 0 are deuterium.

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	5	0
			8232	5201	1462	1528	41			
1	B	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	C	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	D	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	E	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	F	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	G	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	H	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	I	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	J	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	K	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	L	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	M	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	N	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	O	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			
1	P	1021	Total	C	N	O	S	0	5	0
			8232	5201	1462	1528	41			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
A	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
A	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
B	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
B	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
B	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
C	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
C	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
C	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
D	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
D	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
D	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
E	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
E	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
E	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
F	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
F	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
F	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
G	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
G	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
G	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
H	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
H	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
H	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
I	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
I	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
I	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
K	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
K	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
K	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
L	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
L	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
L	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
M	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
M	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
M	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
N	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
N	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
N	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
O	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
O	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
P	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
P	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
P	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722

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

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

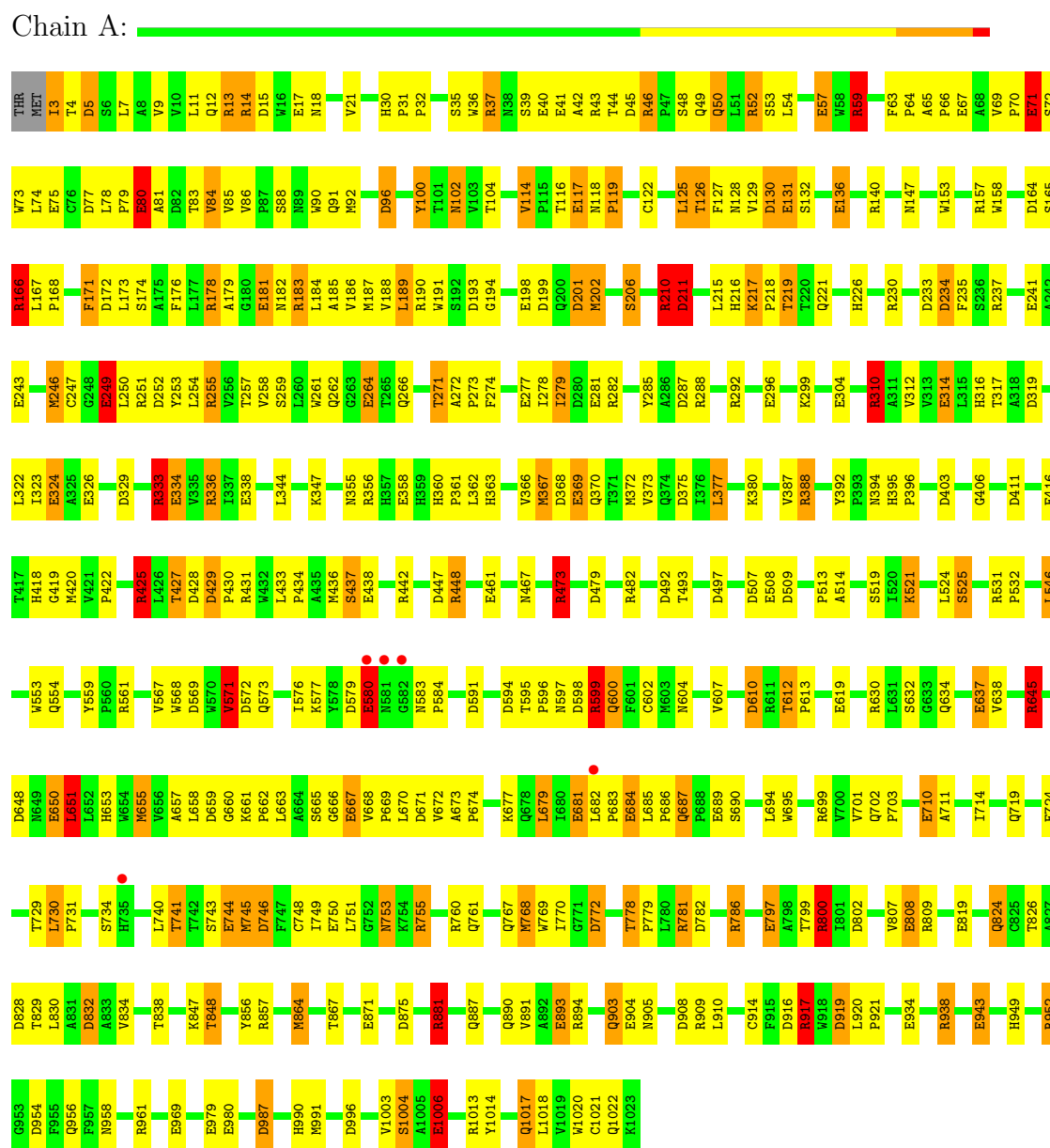
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	434	Total O 434 434	0	0
3	B	436	Total O 436 436	0	0
3	C	433	Total O 433 433	0	0
3	D	437	Total O 437 437	0	0
3	E	435	Total O 435 435	0	0
3	F	436	Total O 436 436	0	0
3	G	434	Total O 434 434	0	0
3	H	435	Total O 435 435	0	0
3	I	434	Total O 434 434	0	0
3	J	436	Total O 436 436	0	0
3	K	435	Total O 435 435	0	0
3	L	435	Total O 435 435	0	0
3	M	434	Total O 434 434	0	0
3	N	436	Total O 436 436	0	0
3	O	433	Total O 433 433	0	0
3	P	437	Total O 437 437	0	0

3 Residue-property plots

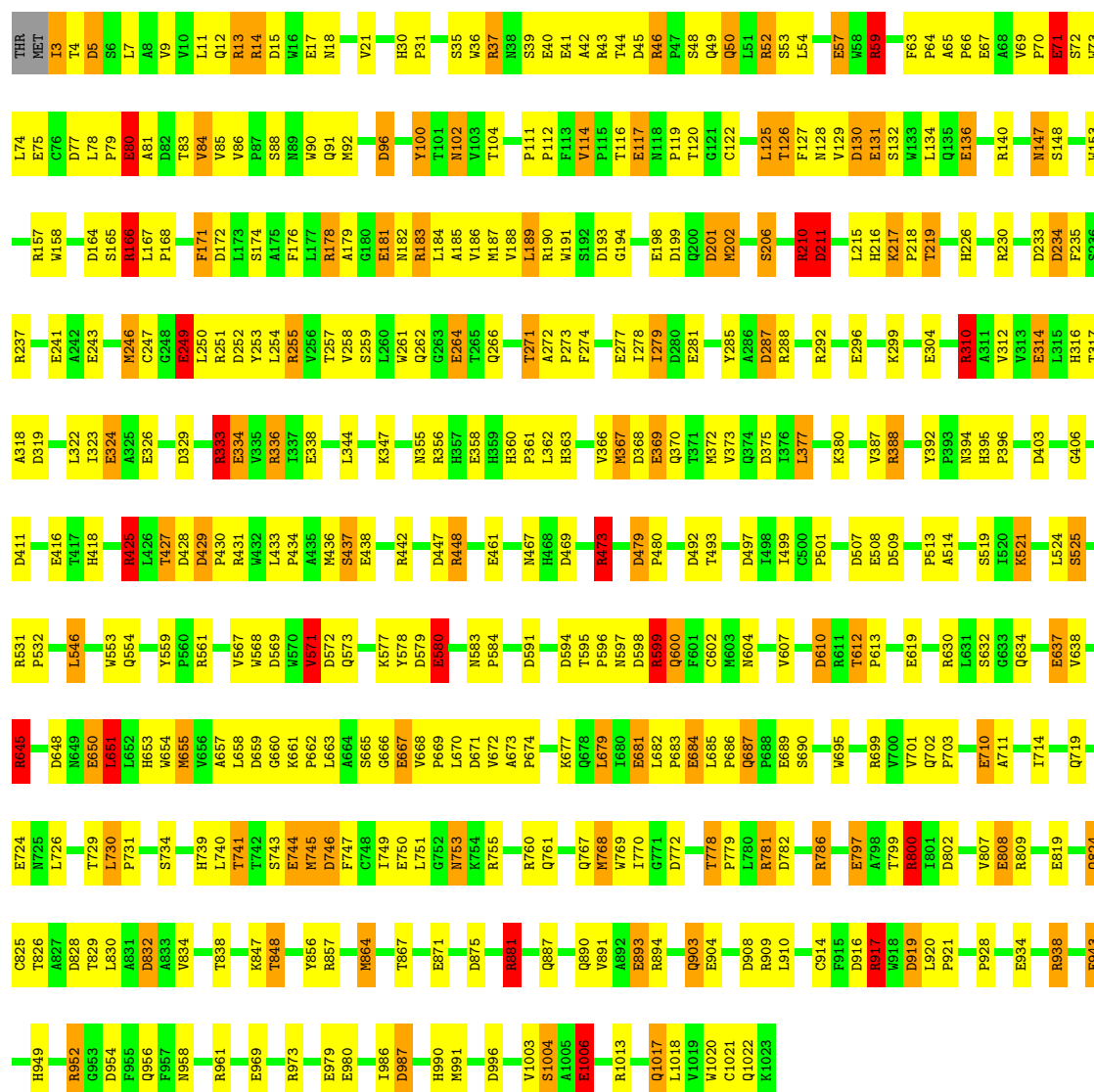
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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



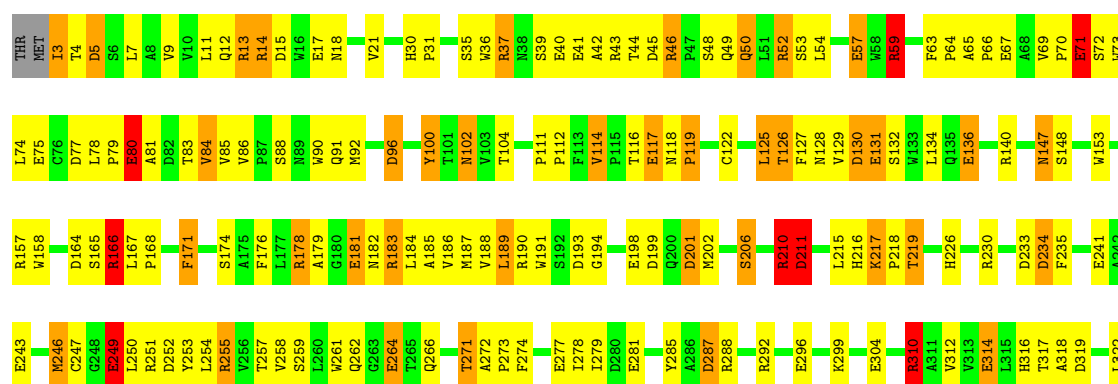
• Molecule 1: BETA-GALACTOSIDASE

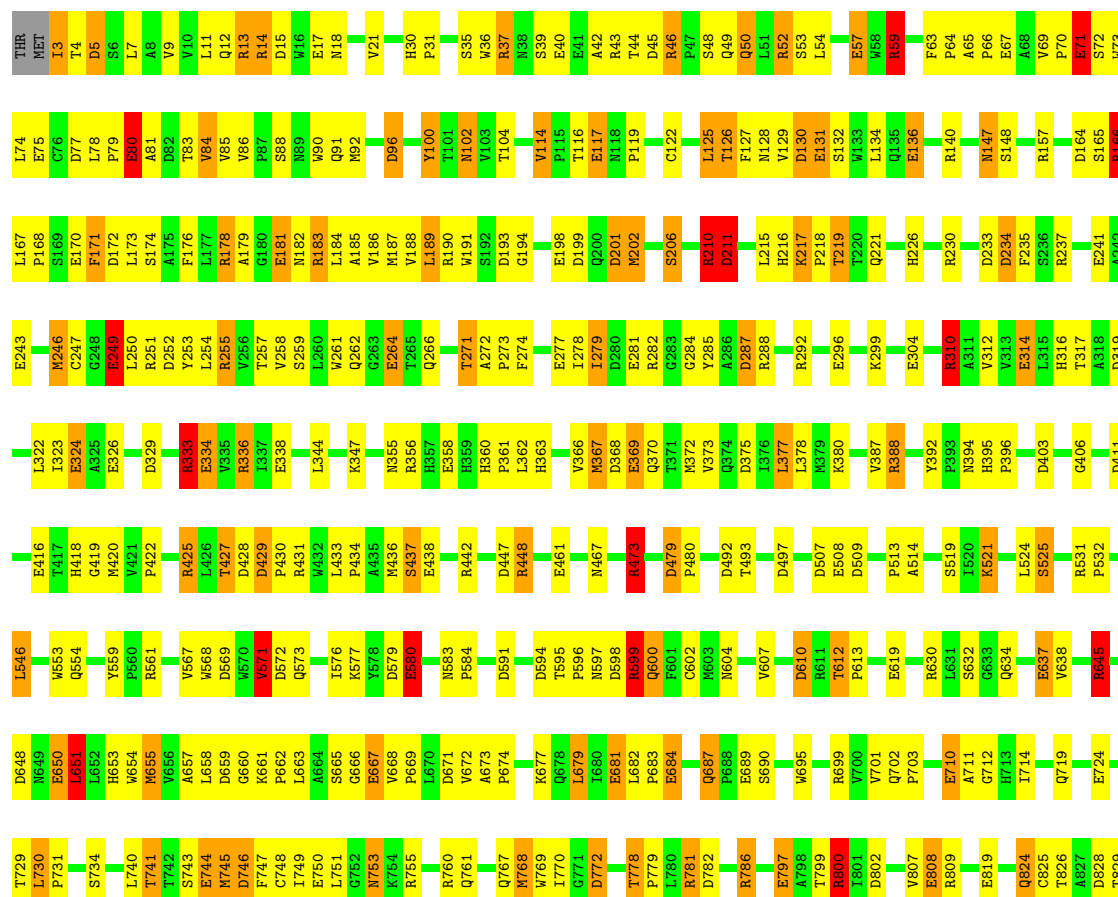
Chain B:

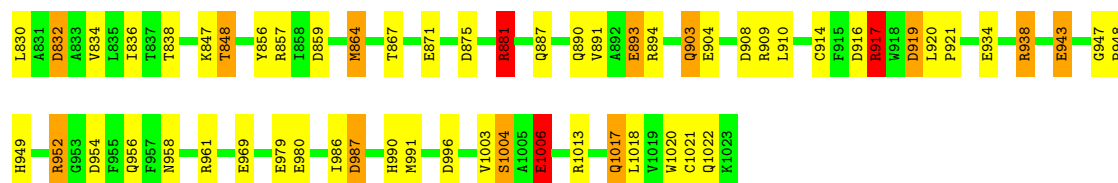


- Molecule 1: BETA-GALACTOSIDASE

Chain C:

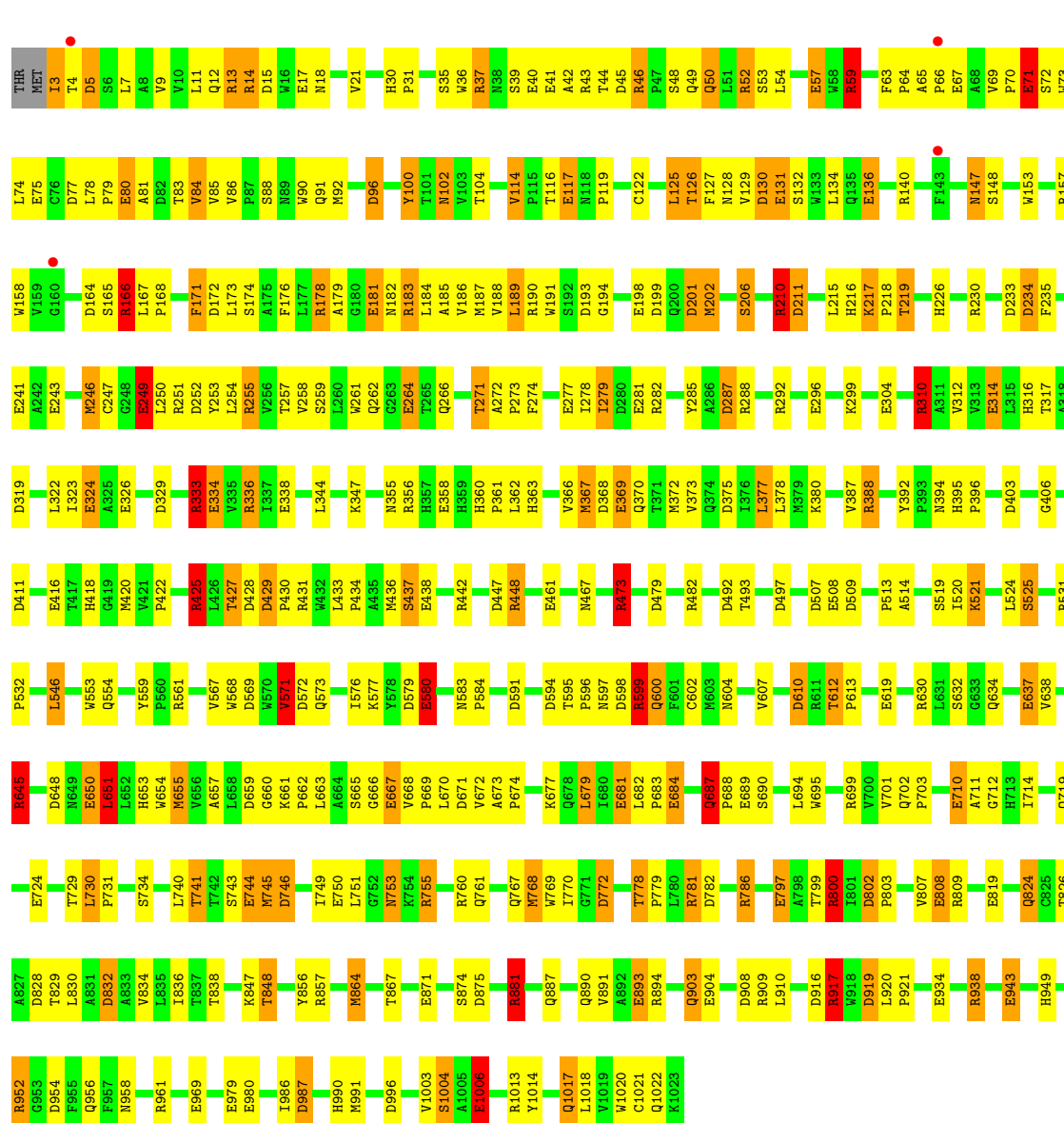






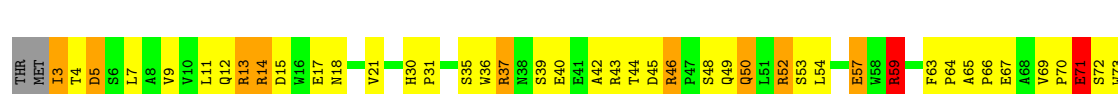
• Molecule 1: BETA-GALACTOSIDASE

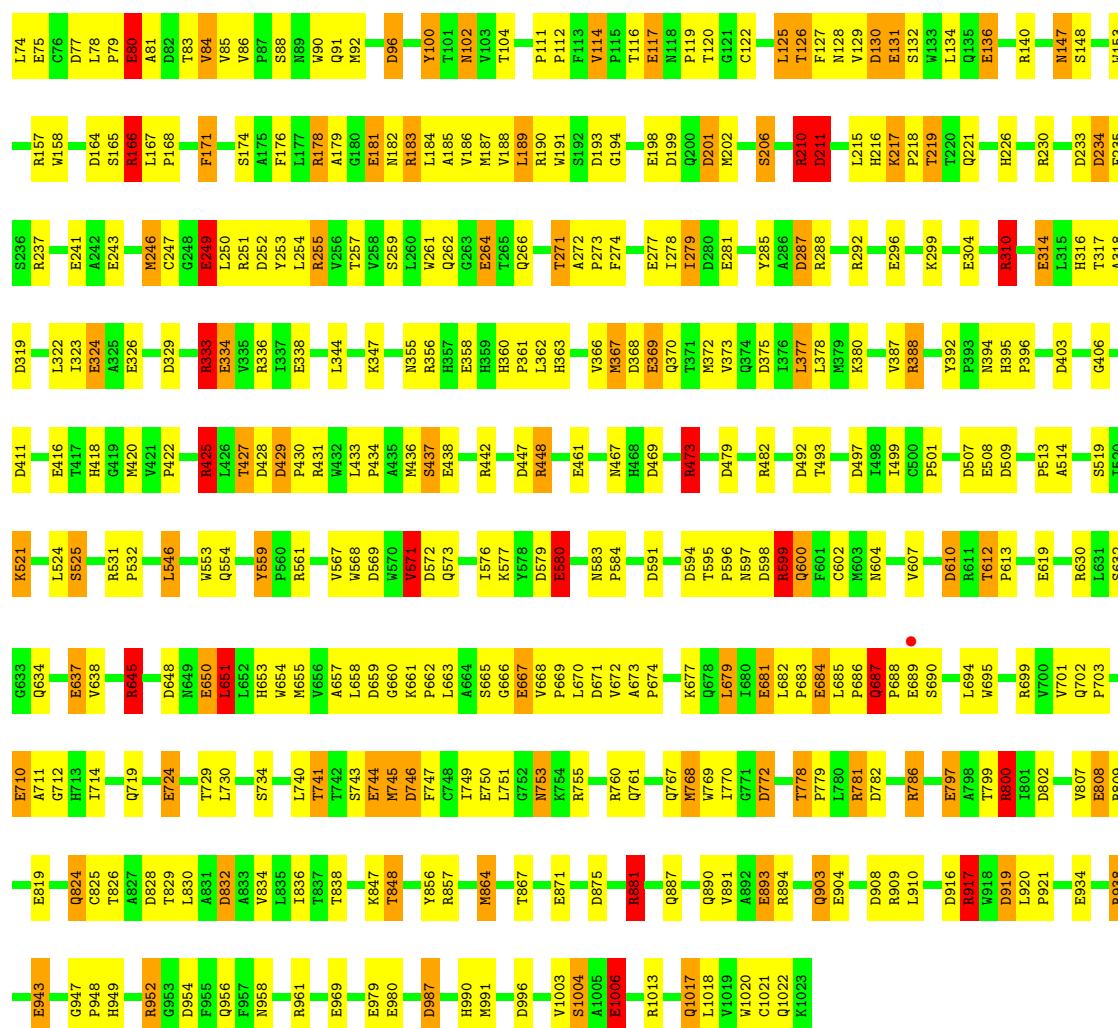
Chain E:



• Molecule 1: BETA-GALACTOSIDASE

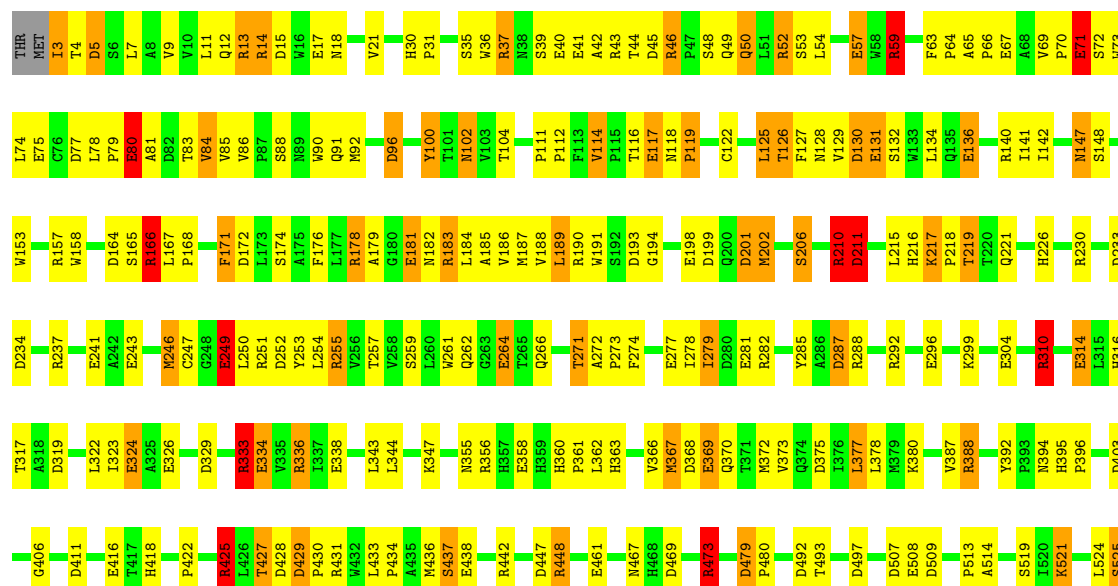
Chain F:

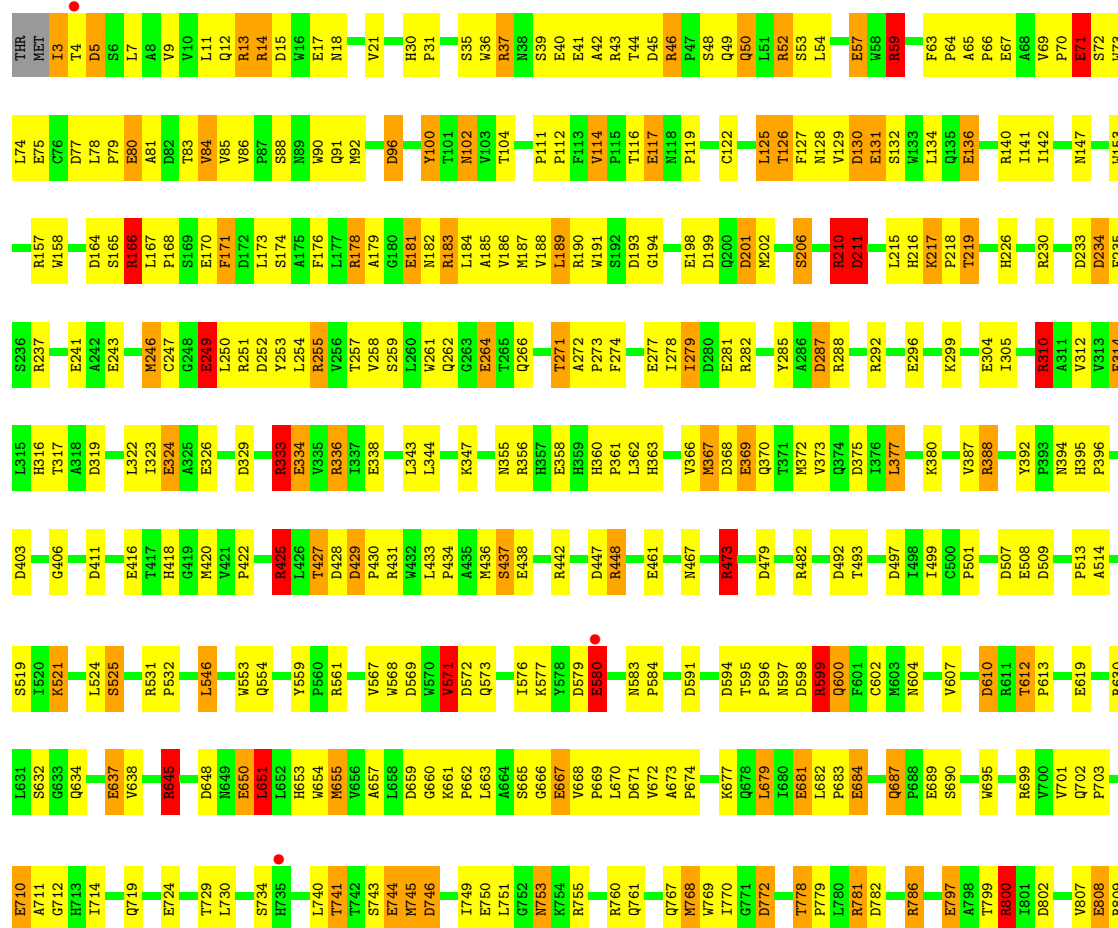




• Molecule 1: BETA-GALACTOSIDASE

Chain G:

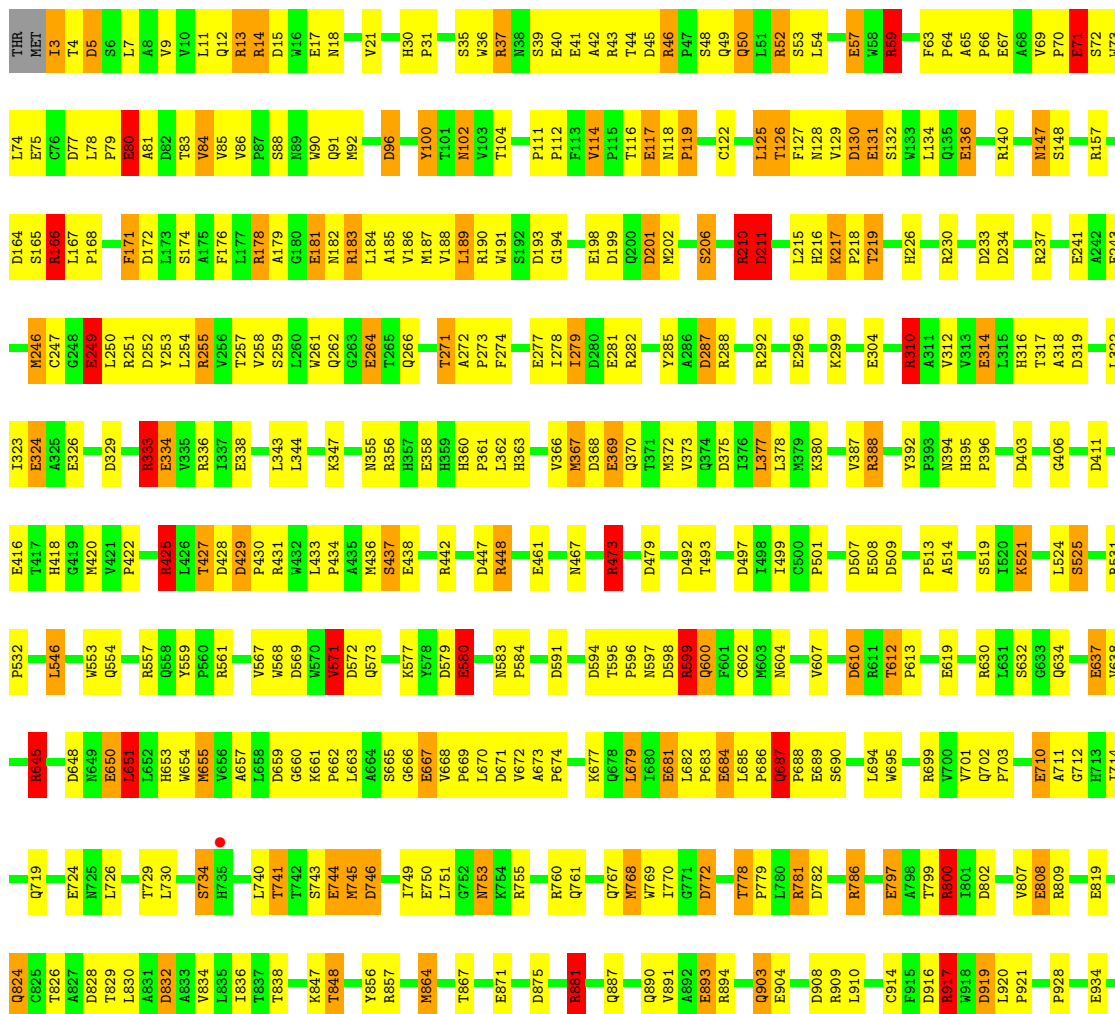






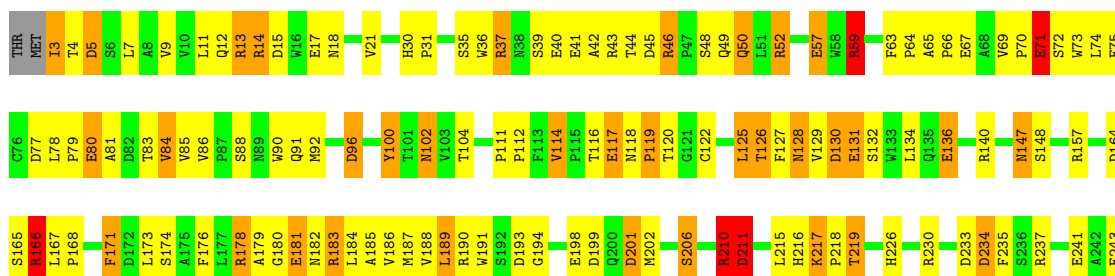
• Molecule 1: BETA-GALACTOSIDASE

Chain I:

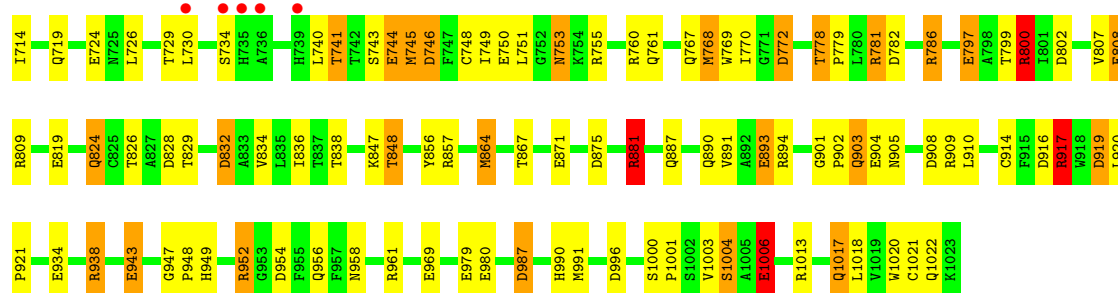


• Molecule 1: BETA-GALACTOSIDASE

Chain J:

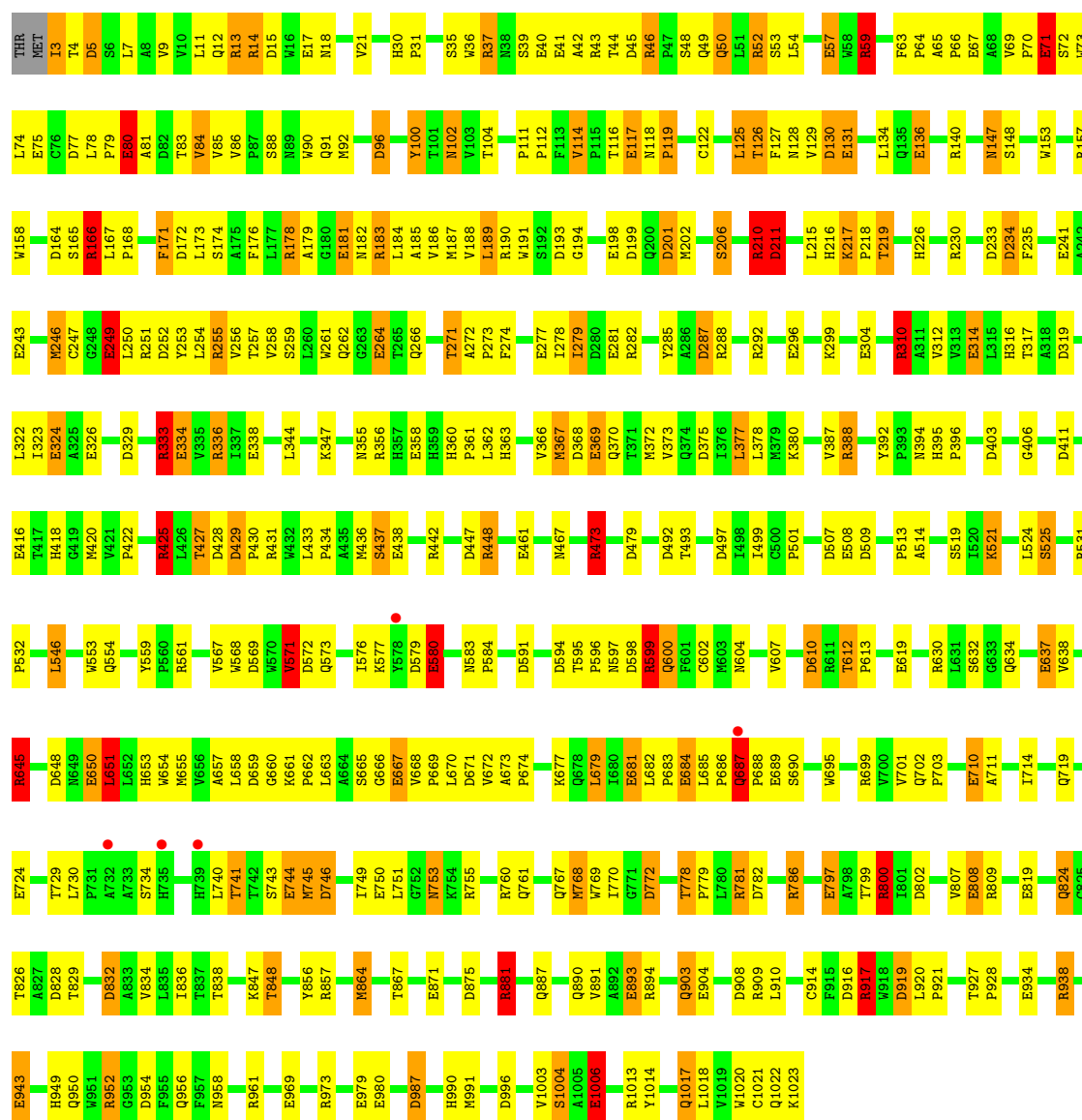






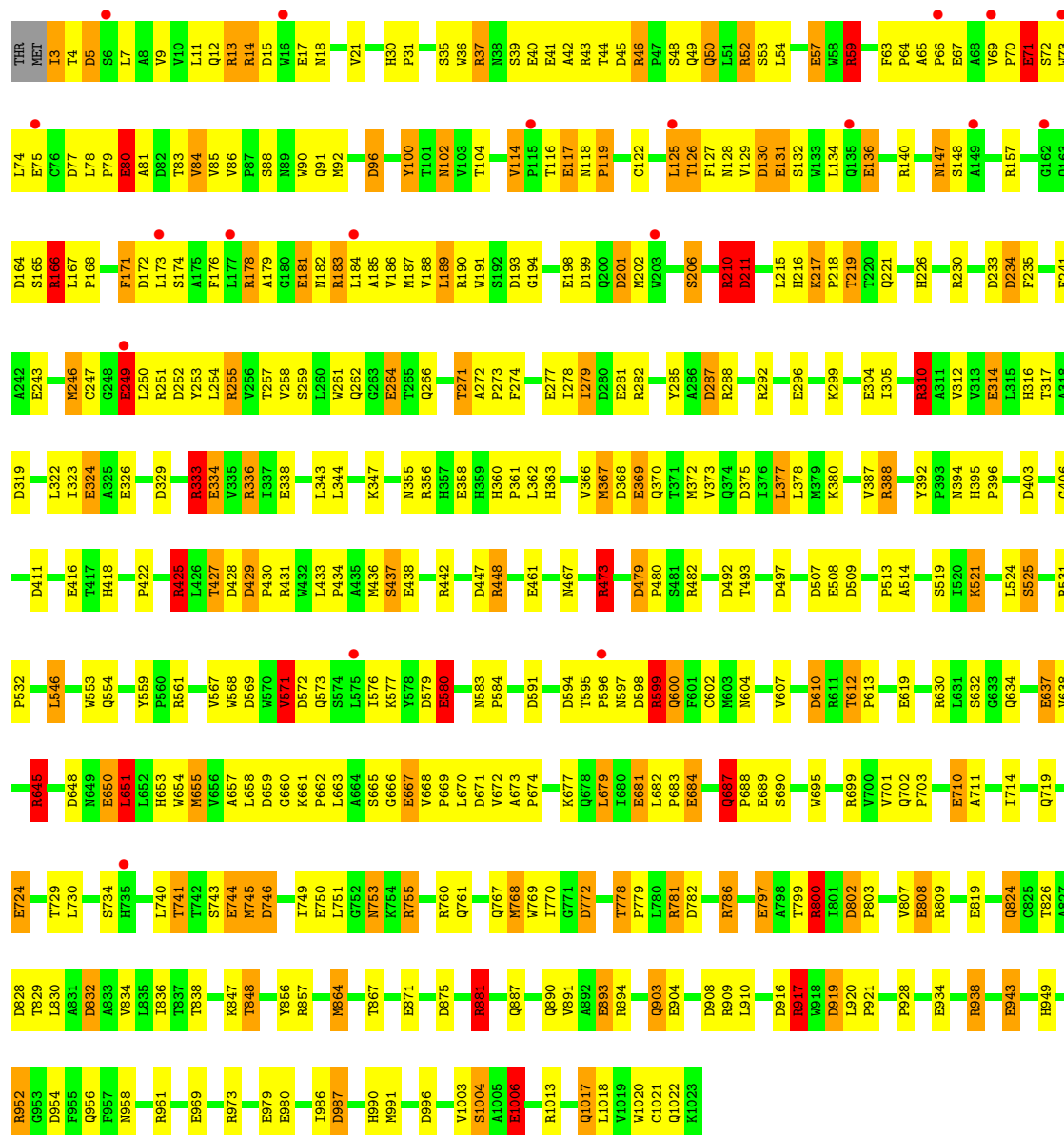
• Molecule 1: BETA-GALACTOSIDASE

Chain L:



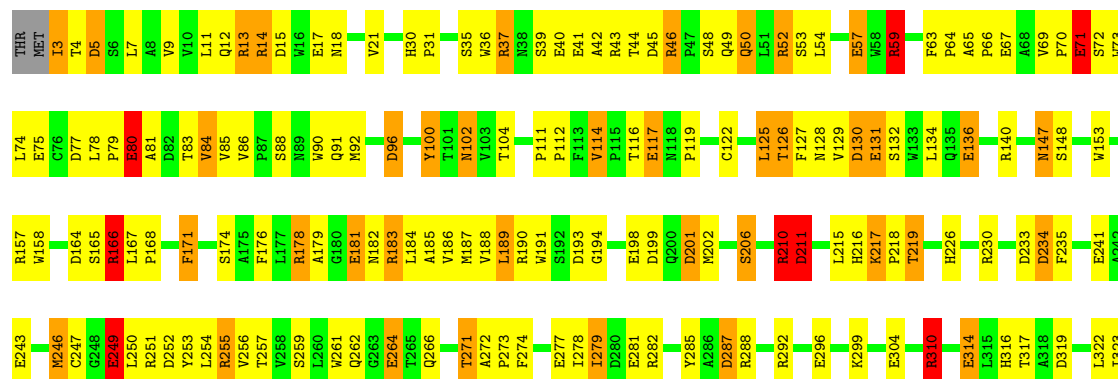
• Molecule 1: BETA-GALACTOSIDASE

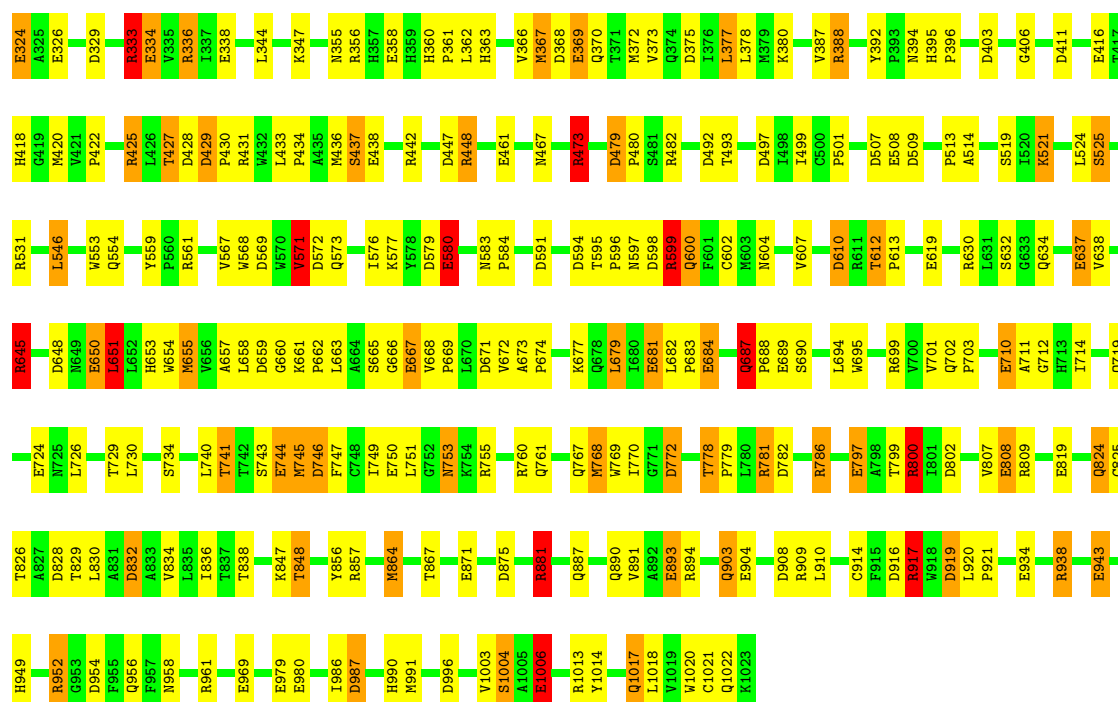
Chain M:

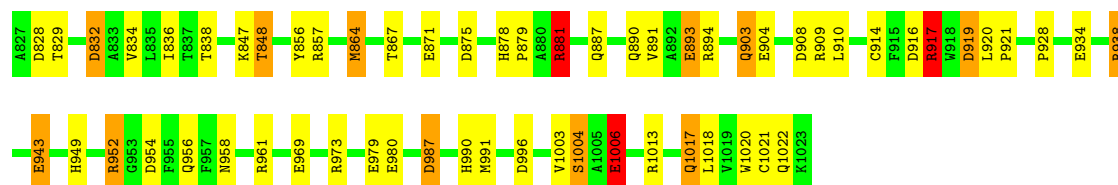


• Molecule 1: BETA-GALACTOSIDASE

Chain N:

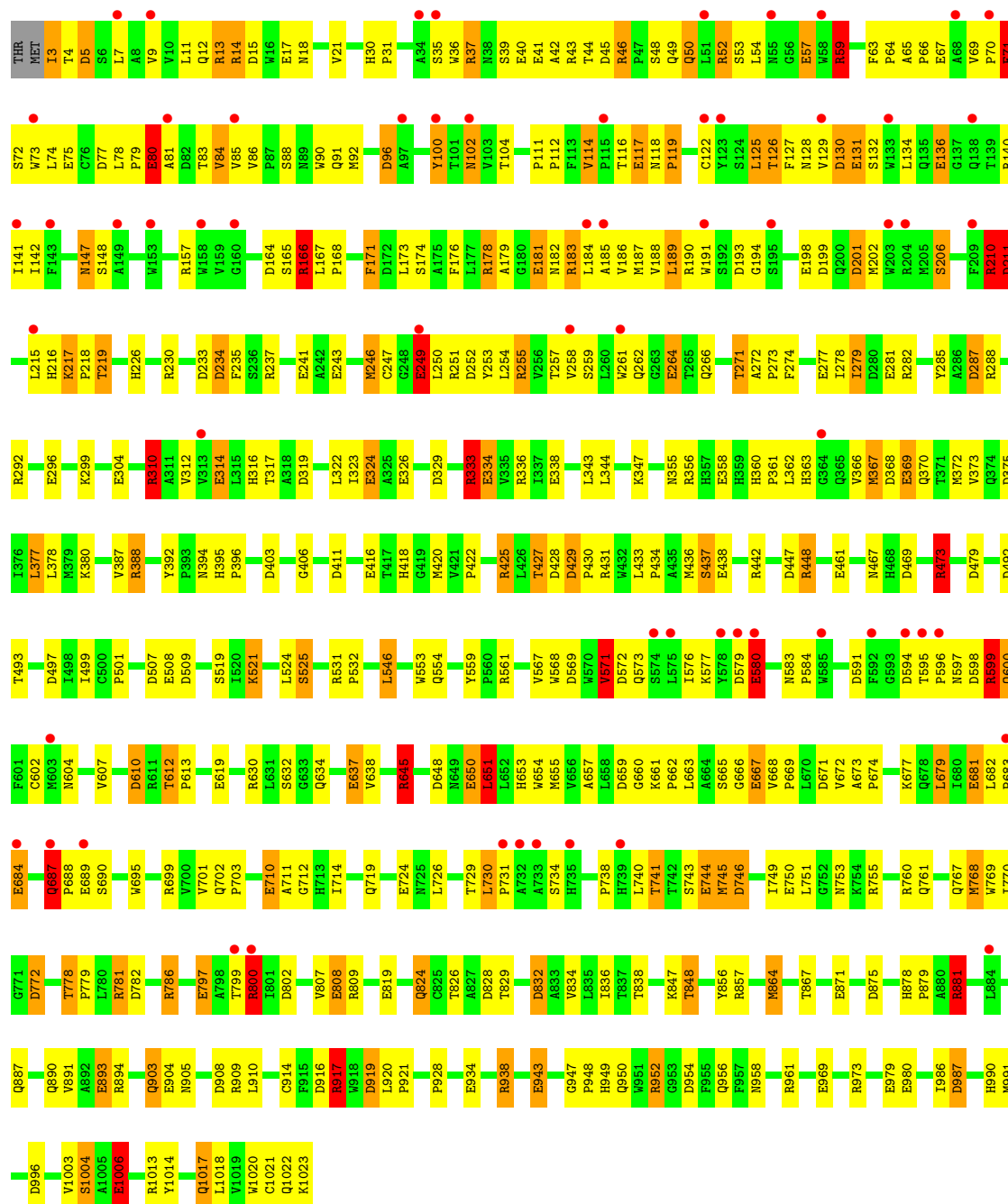






● Molecule 1: BETA-GALACTOSIDASE

Chain P: 



4 Data and refinement statistics

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)	(Not available) – 2.50 92.62 – 2.00	Depositor EDS
% Data completeness (in resolution range)	71.0 ((Not available)-2.50) 39.3 (92.62-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.07 (at 2.00Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.199 , 0.207 0.191 , 0.203	Depositor DCC
R_{free} test set	1680 reflections (0.29%)	DCC
Wilson B-factor (Å ²)	27.6	Xtriage
Anisotropy	0.197	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 79.9	EDS
Estimated twinning fraction	0.009 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 590207 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	138704	wwPDB-VP
Average B, all atoms (Å ²)	35.0	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.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CME, 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	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	B	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	C	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	D	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	E	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	F	1.23	55/8472 (0.6%)	1.69	186/11553 (1.6%)
1	G	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	H	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	I	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	J	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	K	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	L	1.23	56/8472 (0.7%)	1.69	188/11553 (1.6%)
1	M	1.23	56/8472 (0.7%)	1.69	189/11553 (1.6%)
1	N	1.23	56/8472 (0.7%)	1.69	187/11553 (1.6%)
1	O	1.23	57/8472 (0.7%)	1.69	188/11553 (1.6%)
1	P	1.23	56/8472 (0.7%)	1.69	187/11553 (1.6%)
All	All	1.23	896/135552 (0.7%)	1.69	3010/184848 (1.6%)

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

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	H	2	0
1	I	2	0
1	J	2	0
1	K	2	0
1	L	2	0
1	M	2	0
1	N	2	0
1	O	2	0
1	P	2	0
All	All	32	0

The worst 5 of 896 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	249	GLU	CD-OE2	12.04	1.38	1.25
1	O	249	GLU	CD-OE2	12.03	1.38	1.25
1	H	249	GLU	CD-OE2	12.01	1.38	1.25
1	J	249	GLU	CD-OE2	12.01	1.38	1.25
1	P	249	GLU	CD-OE2	12.01	1.38	1.25

The worst 5 of 3010 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	210	ARG	NE-CZ-NH1	18.08	129.34	120.30
1	L	210	ARG	NE-CZ-NH1	18.07	129.34	120.30
1	J	210	ARG	NE-CZ-NH1	18.07	129.33	120.30
1	M	210	ARG	NE-CZ-NH1	18.03	129.31	120.30
1	G	210	ARG	NE-CZ-NH1	18.01	129.30	120.30

5 of 32 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	166	ARG	CA
1	A	249	GLU	CA
1	B	166	ARG	CA
1	B	249	GLU	CA
1	C	166	ARG	CA

There are no planarity outliers.

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8232	0	7817	270	3
1	B	8232	0	7817	267	4
1	C	8232	0	7817	262	1
1	D	8232	0	7817	275	0
1	E	8232	0	7817	266	0
1	F	8232	0	7817	269	0
1	G	8232	0	7817	267	0
1	H	8232	0	7817	261	0
1	I	8232	0	7817	269	1
1	J	8232	0	7817	274	0
1	K	8232	0	7817	277	0
1	L	8232	0	7817	264	0
1	M	8232	0	7817	271	0
1	N	8232	0	7817	269	0
1	O	8232	0	7817	271	0
1	P	8232	0	7817	269	1
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	2	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	434	0	0	13	0
3	B	436	0	0	13	0
3	C	433	0	0	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	437	0	0	13	0
3	E	435	0	0	13	0
3	F	436	0	0	13	0
3	G	434	0	0	13	0
3	H	435	0	0	13	0
3	I	434	0	0	13	0
3	J	436	0	0	13	0
3	K	435	0	0	13	0
3	L	435	0	0	13	0
3	M	434	0	0	13	0
3	N	436	0	0	13	0
3	O	433	0	0	13	0
3	P	437	0	0	13	0
All	All	138704	0	125072	4201	5

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 16.

The worst 5 of 4201 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:L:427:THR:HA	1:L:436:MET:HE1	1.43	1.01
1:E:427:THR:HA	1:E:436:MET:HE1	1.41	1.00
1:M:427:THR:HA	1:M:436:MET:HE1	1.43	1.00
1:J:427:THR:HA	1:J:436:MET:HE1	1.44	0.99
1:C:427:THR:HA	1:C:436:MET:HE1	1.45	0.98

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:580:GLU:O	1:B:578:TYR:CB[2_555]	1.77	0.43
1:A:580:GLU:O	1:B:578:TYR:CG[2_555]	1.85	0.35
1:A:580:GLU:O	1:B:578:TYR:CD1[2_555]	2.10	0.10
1:B:739:HIS:NE2	1:P:738:PRO:O[1_354]	2.10	0.10
1:C:739:HIS:ND1	1:I:734:SER:O[1_655]	2.15	0.05

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	B	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	C	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	D	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	E	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	F	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	G	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	H	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	I	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	J	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	K	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	L	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	M	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	N	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	O	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
1	P	1021/1023 (100%)	976 (96%)	42 (4%)	3 (0%)	50	73
All	All	16336/16368 (100%)	15616 (96%)	672 (4%)	48 (0%)	50	73

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	174	SER
1	B	174	SER
1	C	174	SER
1	D	174	SER
1	E	174	SER

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	B	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	C	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	D	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	E	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	F	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	G	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	H	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	I	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	J	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	K	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	L	875/872 (100%)	777 (89%)	98 (11%)	9	15
1	M	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	N	875/872 (100%)	777 (89%)	98 (11%)	9	15
1	O	875/872 (100%)	776 (89%)	99 (11%)	9	15
1	P	875/872 (100%)	776 (89%)	99 (11%)	9	15
All	All	14000/13952 (100%)	12418 (89%)	1582 (11%)	9	15

5 of 1582 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	166	ARG
1	I	917	ARG
1	O	753	ASN
1	H	344	LEU
1	I	72	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 234 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	316	HIS
1	I	1017	GLN
1	O	824	GLN
1	H	604	ASN
1	I	226	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

48 non-standard protein/DNA/RNA residues 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$
1	CME	A	1021	1	9,9,10	5.91	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	A	748	1	9,9,10	5.12	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	A	914	1	9,9,10	6.30	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	B	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	B	748	1	9,9,10	5.12	2 (22%)	7,9,11	1.45	1 (14%)
1	CME	B	914	1	9,9,10	6.31	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	C	1021	1	9,9,10	5.91	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	C	748	1	9,9,10	5.13	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	C	914	1	9,9,10	6.29	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	D	1021	1	9,9,10	5.89	2 (22%)	7,9,11	4.01	1 (14%)
1	CME	D	748	1	9,9,10	5.12	2 (22%)	7,9,11	1.43	1 (14%)
1	CME	D	914	1	9,9,10	6.30	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	E	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	E	748	1	9,9,10	5.08	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	E	914	1	9,9,10	6.29	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	F	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.00	1 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	CME	F	748	1	9,9,10	5.15	2 (22%)	7,9,11	1.43	1 (14%)
1	CME	F	914	1	9,9,10	6.31	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	G	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	G	748	1	9,9,10	5.11	2 (22%)	7,9,11	1.43	1 (14%)
1	CME	G	914	1	9,9,10	6.30	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	H	1021	1	9,9,10	5.93	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	H	748	1	9,9,10	5.13	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	H	914	1	9,9,10	6.35	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	I	1021	1	9,9,10	5.94	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	I	748	1	9,9,10	5.12	2 (22%)	7,9,11	1.46	1 (14%)
1	CME	I	914	1	9,9,10	6.30	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	J	1021	1	9,9,10	5.91	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	J	748	1	9,9,10	5.11	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	J	914	1	9,9,10	6.27	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	K	1021	1	9,9,10	5.89	2 (22%)	7,9,11	4.01	1 (14%)
1	CME	K	748	1	9,9,10	5.14	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	K	914	1	9,9,10	6.28	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	L	1021	1	9,9,10	5.91	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	L	748	1	9,9,10	5.13	2 (22%)	7,9,11	1.45	1 (14%)
1	CME	L	914	1	9,9,10	6.33	2 (22%)	7,9,11	2.21	1 (14%)
1	CME	M	1021	1	9,9,10	5.91	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	M	748	1	9,9,10	5.09	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	M	914	1	9,9,10	6.31	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	N	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.00	1 (14%)
1	CME	N	748	1	9,9,10	5.14	2 (22%)	7,9,11	1.44	1 (14%)
1	CME	N	914	1	9,9,10	6.32	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	O	1021	1	9,9,10	5.93	2 (22%)	7,9,11	3.99	1 (14%)
1	CME	O	748	1	9,9,10	5.11	2 (22%)	7,9,11	1.43	1 (14%)
1	CME	O	914	1	9,9,10	6.33	2 (22%)	7,9,11	2.20	1 (14%)
1	CME	P	1021	1	9,9,10	5.92	2 (22%)	7,9,11	4.01	1 (14%)
1	CME	P	748	1	9,9,10	5.13	2 (22%)	7,9,11	1.45	1 (14%)
1	CME	P	914	1	9,9,10	6.27	2 (22%)	7,9,11	2.20	1 (14%)

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
1	CME	A	1021	1	-	0/6/8/10	0/0/0/0
1	CME	A	748	1	-	0/6/8/10	0/0/0/0
1	CME	A	914	1	-	0/6/8/10	0/0/0/0
1	CME	B	1021	1	-	0/6/8/10	0/0/0/0
1	CME	B	748	1	-	0/6/8/10	0/0/0/0
1	CME	B	914	1	-	0/6/8/10	0/0/0/0
1	CME	C	1021	1	-	0/6/8/10	0/0/0/0
1	CME	C	748	1	-	0/6/8/10	0/0/0/0
1	CME	C	914	1	-	0/6/8/10	0/0/0/0
1	CME	D	1021	1	-	0/6/8/10	0/0/0/0
1	CME	D	748	1	-	0/6/8/10	0/0/0/0
1	CME	D	914	1	-	0/6/8/10	0/0/0/0
1	CME	E	1021	1	-	0/6/8/10	0/0/0/0
1	CME	E	748	1	-	0/6/8/10	0/0/0/0
1	CME	E	914	1	-	0/6/8/10	0/0/0/0
1	CME	F	1021	1	-	0/6/8/10	0/0/0/0
1	CME	F	748	1	-	0/6/8/10	0/0/0/0
1	CME	F	914	1	-	0/6/8/10	0/0/0/0
1	CME	G	1021	1	-	0/6/8/10	0/0/0/0
1	CME	G	748	1	-	0/6/8/10	0/0/0/0
1	CME	G	914	1	-	0/6/8/10	0/0/0/0
1	CME	H	1021	1	-	0/6/8/10	0/0/0/0
1	CME	H	748	1	-	0/6/8/10	0/0/0/0
1	CME	H	914	1	-	0/6/8/10	0/0/0/0
1	CME	I	1021	1	-	0/6/8/10	0/0/0/0
1	CME	I	748	1	-	0/6/8/10	0/0/0/0
1	CME	I	914	1	-	0/6/8/10	0/0/0/0
1	CME	J	1021	1	-	0/6/8/10	0/0/0/0
1	CME	J	748	1	-	0/6/8/10	0/0/0/0
1	CME	J	914	1	-	0/6/8/10	0/0/0/0
1	CME	K	1021	1	-	0/6/8/10	0/0/0/0
1	CME	K	748	1	-	0/6/8/10	0/0/0/0
1	CME	K	914	1	-	0/6/8/10	0/0/0/0
1	CME	L	1021	1	-	0/6/8/10	0/0/0/0
1	CME	L	748	1	-	0/6/8/10	0/0/0/0
1	CME	L	914	1	-	0/6/8/10	0/0/0/0
1	CME	M	1021	1	-	0/6/8/10	0/0/0/0
1	CME	M	748	1	-	0/6/8/10	0/0/0/0
1	CME	M	914	1	-	0/6/8/10	0/0/0/0
1	CME	N	1021	1	-	0/6/8/10	0/0/0/0
1	CME	N	748	1	-	0/6/8/10	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CME	N	914	1	-	0/6/8/10	0/0/0/0
1	CME	O	1021	1	-	0/6/8/10	0/0/0/0
1	CME	O	748	1	-	0/6/8/10	0/0/0/0
1	CME	O	914	1	-	0/6/8/10	0/0/0/0
1	CME	P	1021	1	-	0/6/8/10	0/0/0/0
1	CME	P	748	1	-	0/6/8/10	0/0/0/0
1	CME	P	914	1	-	0/6/8/10	0/0/0/0

The worst 5 of 96 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	914	CME	O-C	18.81	1.24	1.11
1	L	914	CME	O-C	18.76	1.24	1.11
1	O	914	CME	O-C	18.74	1.24	1.11
1	N	914	CME	O-C	18.71	1.24	1.11
1	F	914	CME	O-C	18.70	1.24	1.11

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1021	CME	CB-SG-SD	-10.43	83.06	103.90
1	K	1021	CME	CB-SG-SD	-10.43	83.07	103.90
1	P	1021	CME	CB-SG-SD	-10.42	83.08	103.90
1	N	1021	CME	CB-SG-SD	-10.42	83.09	103.90
1	A	1021	CME	CB-SG-SD	-10.41	83.11	103.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 32 ligands modelled in this entry, 32 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.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å ²)	Q < 0.9
1	A	1021/1023 (99%)	-0.73	5 (0%) 88 90	3, 21, 65, 100	1 (0%)
1	B	1021/1023 (99%)	-0.77	0 100 100	3, 21, 65, 100	1 (0%)
1	C	1021/1023 (99%)	-0.68	1 (0%) 93 95	2, 19, 63, 98	1 (0%)
1	D	1021/1023 (99%)	-0.69	0 100 100	5, 24, 66, 100	1 (0%)
1	E	1021/1023 (99%)	-0.52	4 (0%) 90 92	15, 33, 74, 100	1 (0%)
1	F	1021/1023 (99%)	-0.75	1 (0%) 93 95	3, 21, 66, 100	1 (0%)
1	G	1021/1023 (99%)	-0.76	1 (0%) 93 95	7, 25, 68, 100	1 (0%)
1	H	1021/1023 (99%)	-0.52	3 (0%) 91 93	14, 33, 73, 100	1 (0%)
1	I	1021/1023 (99%)	-0.68	1 (0%) 93 95	10, 28, 70, 100	1 (0%)
1	J	1021/1023 (99%)	-0.70	1 (0%) 93 95	8, 26, 69, 100	1 (0%)
1	K	1021/1023 (99%)	-0.60	5 (0%) 88 90	17, 35, 75, 100	1 (0%)
1	L	1021/1023 (99%)	-0.52	5 (0%) 88 90	16, 34, 74, 100	1 (0%)
1	M	1021/1023 (99%)	-0.17	19 (1%) 64 66	22, 40, 78, 100	1 (0%)
1	N	1021/1023 (99%)	-0.64	0 100 100	11, 29, 71, 100	1 (0%)
1	O	1021/1023 (99%)	-0.69	2 (0%) 93 94	12, 30, 72, 100	1 (0%)
1	P	1021/1023 (99%)	0.31	63 (6%) 20 20	29, 47, 83, 100	1 (0%)
All	All	16336/16368 (99%)	-0.57	111 (0%) 84 86	2, 30, 72, 100	16 (0%)

The worst 5 of 111 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	313	VAL	5.4
1	P	732	ALA	5.1
1	P	70	PRO	5.1
1	K	735	HIS	4.9
1	L	735	HIS	4.9

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å ²)	Q<0.9
1	CME	F	914	10/11	0.12	1.42	8,14,100,100	0
1	CME	H	1021	10/11	0.13	0.84	13,36,100,100	0
1	CME	D	1021	10/11	0.13	0.57	4,27,100,100	0
1	CME	D	748	10/11	0.16	0.54	21,30,100,100	0
1	CME	L	748	10/11	0.16	0.22	31,41,100,100	0
1	CME	I	1021	10/11	0.14	0.19	9,32,100,100	0
1	CME	O	1021	10/11	0.12	0.17	11,34,100,100	0
1	CME	H	748	10/11	0.13	0.16	29,39,100,100	0
1	CME	C	914	10/11	0.10	0.14	6,12,97,97	0
1	CME	C	1021	10/11	0.11	0.12	1,23,97,97	0
1	CME	K	1021	10/11	0.13	0.07	16,39,100,100	0
1	CME	E	914	10/11	0.12	0.02	20,26,100,100	0
1	CME	H	914	10/11	0.11	-0.07	19,25,100,100	0
1	CME	F	1021	10/11	0.10	-0.10	2,25,100,100	0
1	CME	M	1021	10/11	0.14	-0.11	21,44,100,100	0
1	CME	A	1021	10/11	0.10	-0.11	2,25,100,100	0
1	CME	E	1021	10/11	0.12	-0.14	14,37,100,100	0
1	CME	I	914	10/11	0.11	-0.17	15,21,100,100	0
1	CME	M	914	10/11	0.13	-0.19	27,33,100,100	0
1	CME	K	914	10/11	0.08	-0.25	22,28,100,100	0
1	CME	N	1021	10/11	0.11	-0.31	9,33,100,100	0
1	CME	L	914	10/11	0.09	-0.32	21,27,100,100	0
1	CME	J	914	10/11	0.09	-0.41	13,19,100,100	0
1	CME	G	1021	10/11	0.08	-0.44	6,29,100,100	0
1	CME	C	748	10/11	0.12	-0.46	16,26,97,97	0
1	CME	J	1021	10/11	0.10	-0.48	7,30,100,100	0
1	CME	D	914	10/11	0.10	-0.49	10,16,100,100	0
1	CME	B	914	10/11	0.08	-0.59	8,14,99,99	0
1	CME	B	1021	10/11	0.08	-0.60	2,25,99,99	0
1	CME	O	748	10/11	0.12	-0.61	27,37,100,100	0
1	CME	K	748	10/11	0.14	-0.66	32,42,100,100	0
1	CME	P	1021	10/11	0.13	-0.73	28,51,100,100	0
1	CME	P	748	10/11	0.11	-0.75	44,54,100,100	0
1	CME	N	914	10/11	0.08	-0.77	16,22,100,100	0
1	CME	E	748	10/11	0.10	-0.90	30,40,100,100	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
1	CME	O	914	10/11	0.08	-0.96	17,23,100,100	0
1	CME	P	914	10/11	0.11	-0.98	34,40,100,100	0
1	CME	A	748	10/11	0.09	-0.99	18,28,100,100	0
1	CME	G	914	10/11	0.07	-1.00	12,18,100,100	0
1	CME	M	748	10/11	0.11	-1.00	37,47,100,100	0
1	CME	A	914	10/11	0.10	-1.01	8,14,100,100	0
1	CME	B	748	10/11	0.08	-1.08	18,28,99,99	0
1	CME	I	748	10/11	0.10	-1.10	25,35,100,100	0
1	CME	G	748	10/11	0.09	-1.15	22,32,100,100	0
1	CME	F	748	10/11	0.09	-1.15	18,28,100,100	0
1	CME	L	1021	10/11	0.09	-1.17	15,38,100,100	0
1	CME	J	748	10/11	0.08	-1.57	23,33,100,100	0
1	CME	N	748	10/11	0.07	-1.78	26,35,100,100	0

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	MG	A	3001	1/1	0.10	0.94	18,18,18,18	0
2	MG	J	3002	1/1	0.12	0.83	23,23,23,23	0
2	MG	N	3002	1/1	0.14	0.51	26,26,26,26	0
2	MG	A	3002	1/1	0.11	0.39	18,18,18,18	0
2	MG	D	3002	1/1	0.10	0.35	21,21,21,21	0
2	MG	F	3001	1/1	0.10	0.35	18,18,18,18	0
2	MG	H	3002	1/1	0.12	-0.22	30,30,30,30	0
2	MG	O	3002	1/1	0.10	-0.56	27,27,27,27	0
2	MG	M	3002	1/1	0.15	-0.61	37,37,37,37	0
2	MG	I	3002	1/1	0.09	-0.68	25,25,25,25	0
2	MG	E	3002	1/1	0.10	-0.99	30,30,30,30	0
2	MG	P	3001	1/1	0.12	-1.00	44,44,44,44	0
2	MG	C	3002	1/1	0.09	-1.11	16,16,16,16	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	MG	M	3001	1/1	0.09	-1.29	37,37,37,37	0
2	MG	K	3001	1/1	0.07	-1.62	32,32,32,32	0
2	MG	B	3001	1/1	0.08	-1.70	18,18,18,18	0
2	MG	K	3002	1/1	0.07	-1.83	32,32,32,32	0
2	MG	F	3002	1/1	0.06	-1.89	19,19,19,19	0
2	MG	J	3001	1/1	0.05	-2.08	23,23,23,23	0
2	MG	I	3001	1/1	0.04	-2.47	25,25,25,25	0
2	MG	L	3001	1/1	0.05	-2.56	31,31,31,31	0
2	MG	L	3002	1/1	0.06	-2.68	31,31,31,31	0
2	MG	P	3002	1/1	0.08	-2.71	44,44,44,44	0
2	MG	D	3001	1/1	0.04	-3.18	20,20,20,20	0
2	MG	G	3001	1/1	0.05	-3.46	22,22,22,22	0
2	MG	H	3001	1/1	0.04	-3.54	29,29,29,29	0
2	MG	E	3001	1/1	0.03	-3.59	30,30,30,30	0
2	MG	B	3002	1/1	0.05	-3.83	18,18,18,18	0
2	MG	O	3001	1/1	0.05	-4.74	27,27,27,27	0
2	MG	G	3002	1/1	0.04	-4.89	22,22,22,22	0
2	MG	C	3001	1/1	0.04	-6.07	16,16,16,16	0
2	MG	N	3001	1/1	0.04	-7.97	26,26,26,26	0

6.5 Other polymers ⓘ

There are no such residues in this entry.