



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 07:05 AM GMT

PDB ID : 1JYZ  
Title : E. COLI (lacZ) BETA-GALACTOSIDASE IN COMPLEX WITH 2-F-LACTOSE. Chains I-P, see REMARK 400.  
Authors : Juers, D.H.; McCarter, J.D.; Withers, S.G.; Matthews, B.W.  
Deposited on : 2001-09-13  
Resolution : 2.70 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

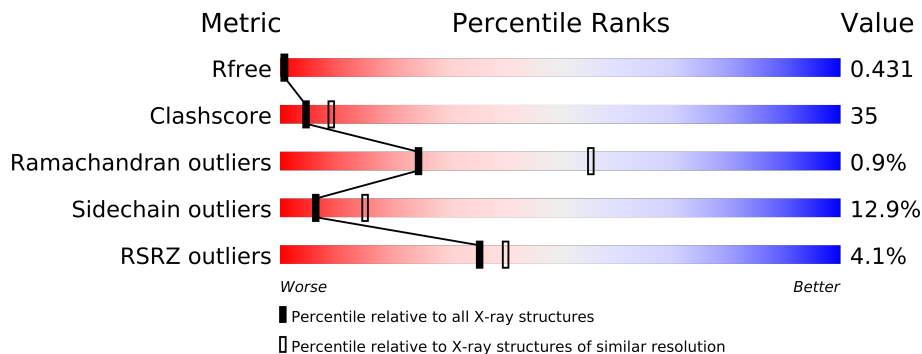
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 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	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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  $\geq 3$ , 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	I	1023	
1	J	1023	
1	K	1023	
1	L	1023	
1	M	1023	
1	N	1023	
1	O	1023	
1	P	1023	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	2FL	J	2001	-	X

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
2	2FL	K	2001	-	X
2	2FL	L	2001	-	X
2	2FL	M	2001	-	X
2	2FL	N	2001	-	X
2	2FL	O	2001	-	X
2	2FL	P	2001	-	X
3	MG	N	3001	-	X
3	MG	O	3001	-	X
4	NA	J	3102	-	X
4	NA	K	3102	-	X
4	NA	N	3101	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 67264 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	I	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	J	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	K	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	L	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	M	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	N	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	O	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			
1	P	1021	Total	C	N	O	S	0	2	0
			8219	5196	1454	1528	41			

There are 24 discrepancies between the modelled and reference sequences:

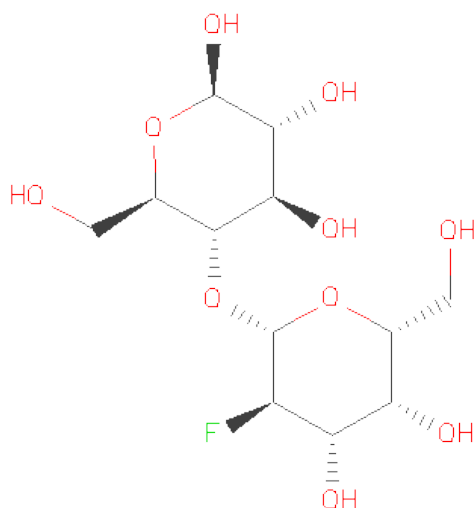
Chain	Residue	Modelled	Actual	Comment	Reference
I	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
I	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
I	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	748	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	914	CME	CYS	MODIFIED RESIDUE	UNP P00722
J	1021	CME	CYS	MODIFIED RESIDUE	UNP P00722
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

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
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
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 2-FLUORO-2-DEOXY-LACTOSE (three-letter code: 2FL) (formula:  $C_{12}H_{21}FO_{10}$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	I	1	Total	C	F	O	0	0
			23	12	1	10		
2	J	1	Total	C	F	O	0	0
			23	12	1	10		
2	K	1	Total	C	F	O	0	0
			23	12	1	10		
2	L	1	Total	C	F	O	0	0
			23	12	1	10		
2	M	1	Total	C	F	O	0	0
			23	12	1	10		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	N	1	Total	C	F	O	0	0
			23	12	1	10		
2	O	1	Total	C	F	O	0	0
			23	12	1	10		
2	P	1	Total	C	F	O	0	0
			23	12	1	10		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	2	Total	Mg	0	0
			2	2		
3	J	2	Total	Mg	0	0
			2	2		
3	K	2	Total	Mg	0	0
			2	2		
3	I	2	Total	Mg	0	0
			2	2		
3	N	2	Total	Mg	0	0
			2	2		
3	O	2	Total	Mg	0	0
			2	2		
3	L	2	Total	Mg	0	0
			2	2		
3	M	2	Total	Mg	0	0
			2	2		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	P	2	Total	Na	0	0
			2	2		
4	J	2	Total	Na	0	0
			2	2		
4	K	2	Total	Na	0	0
			2	2		
4	I	2	Total	Na	0	0
			2	2		
4	N	2	Total	Na	0	0
			2	2		
4	O	2	Total	Na	0	0
			2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	2	Total 2	Na 2	0	0
4	M	2	Total 2	Na 2	0	0

- Molecule 5 is water.

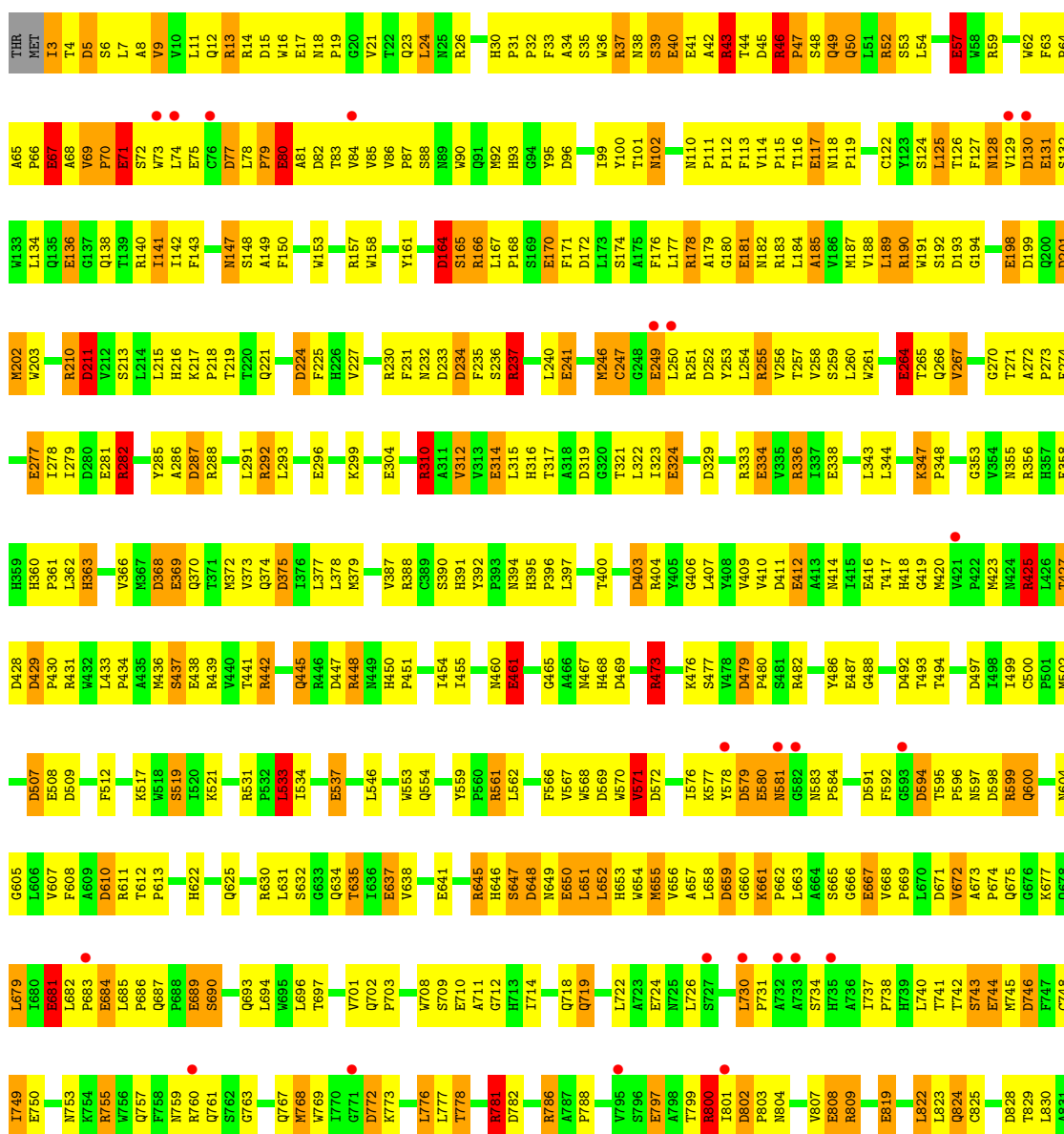
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	I	162	Total 162	O 162	0	0
5	J	162	Total 162	O 162	0	0
5	K	162	Total 162	O 162	0	0
5	L	162	Total 162	O 162	0	0
5	M	161	Total 161	O 161	0	0
5	N	163	Total 163	O 163	0	0
5	O	161	Total 161	O 161	0	0
5	P	163	Total 163	O 163	0	0

### 3 Residue-property plots

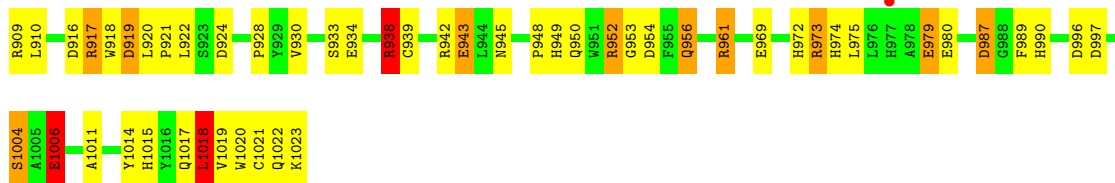
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Chain I: 

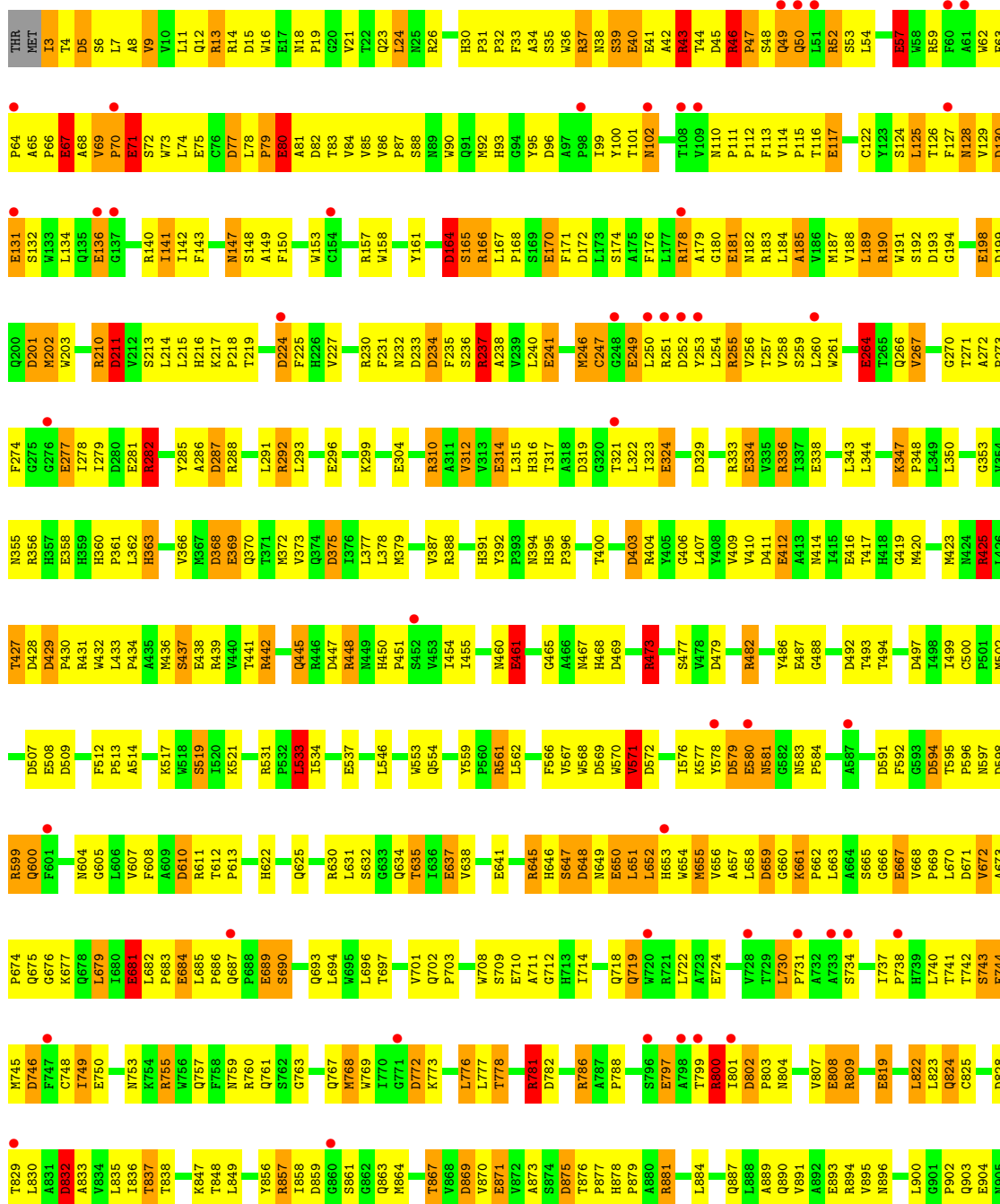


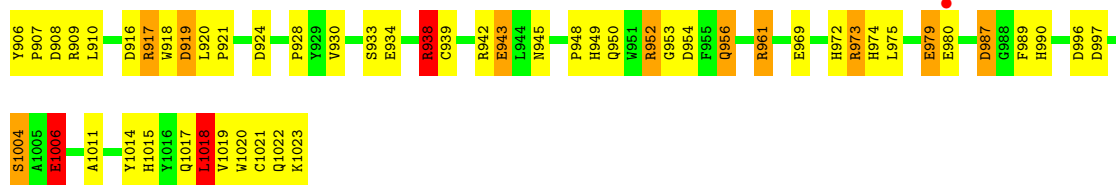




● Molecule 1: Beta-Galactosidase

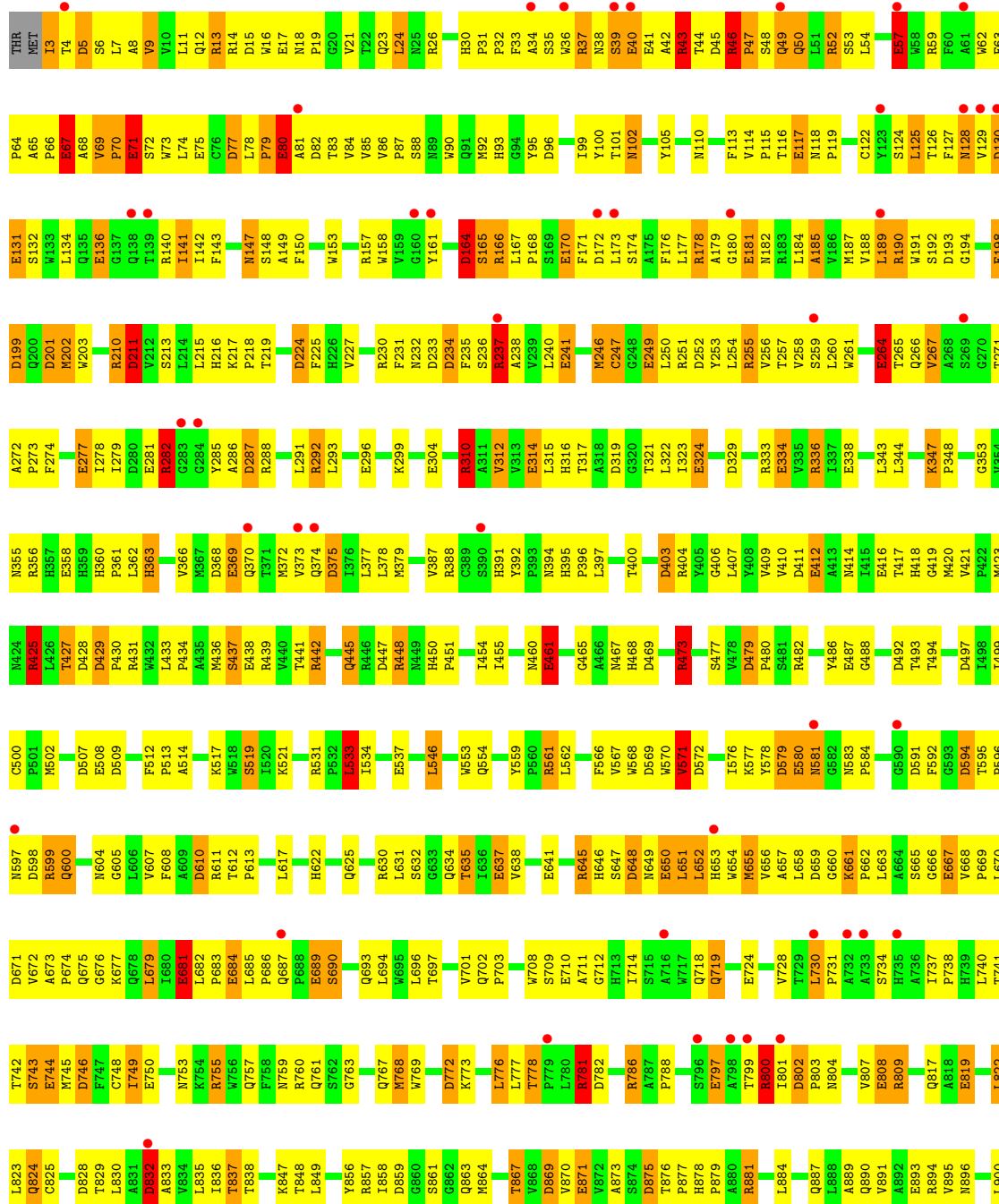
Chain K:





• Molecule 1: Beta-Galactosidase

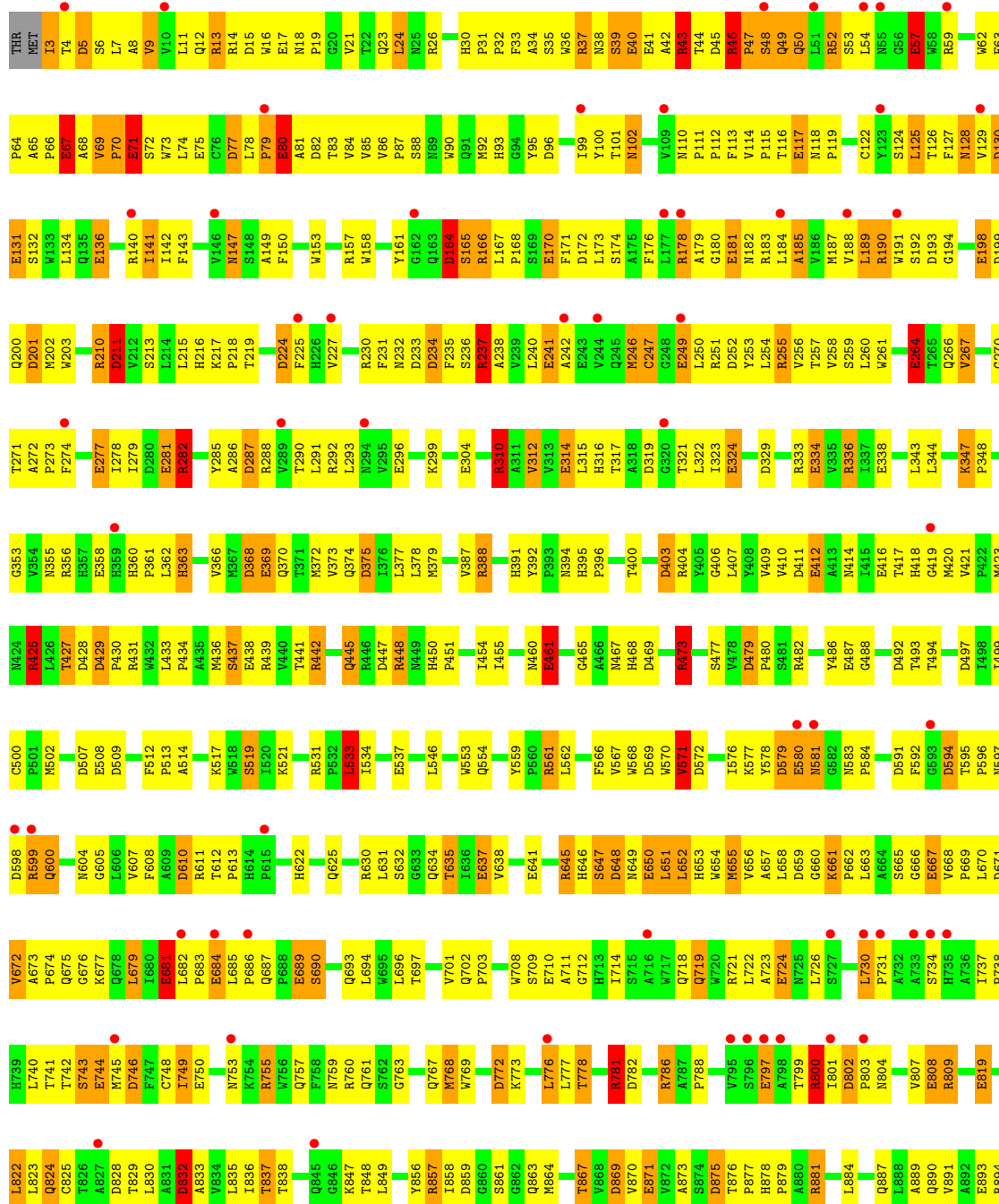
Chain L:

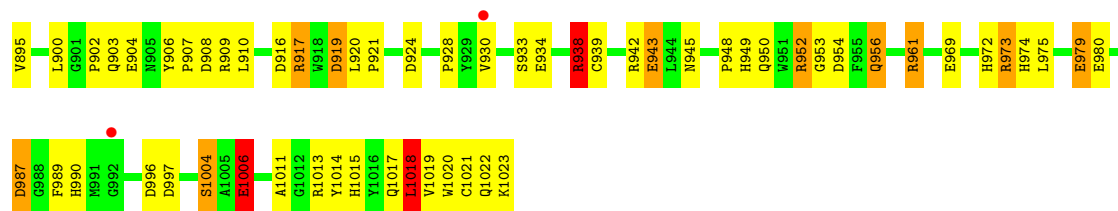




● Molecule 1: Beta-Galactosidase

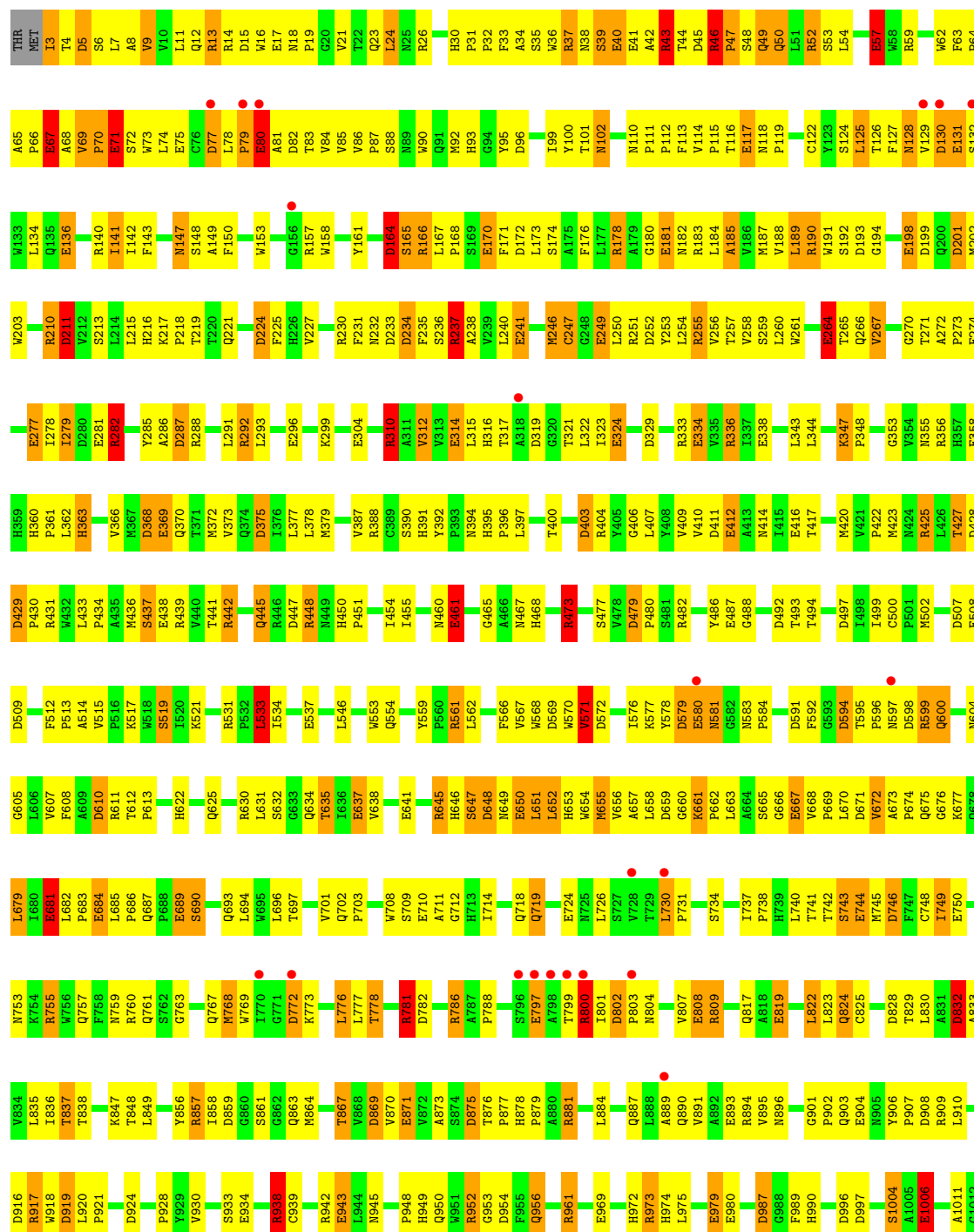
Chain M:





### • Molecule 1: Beta-Galactosidase

Chain N:

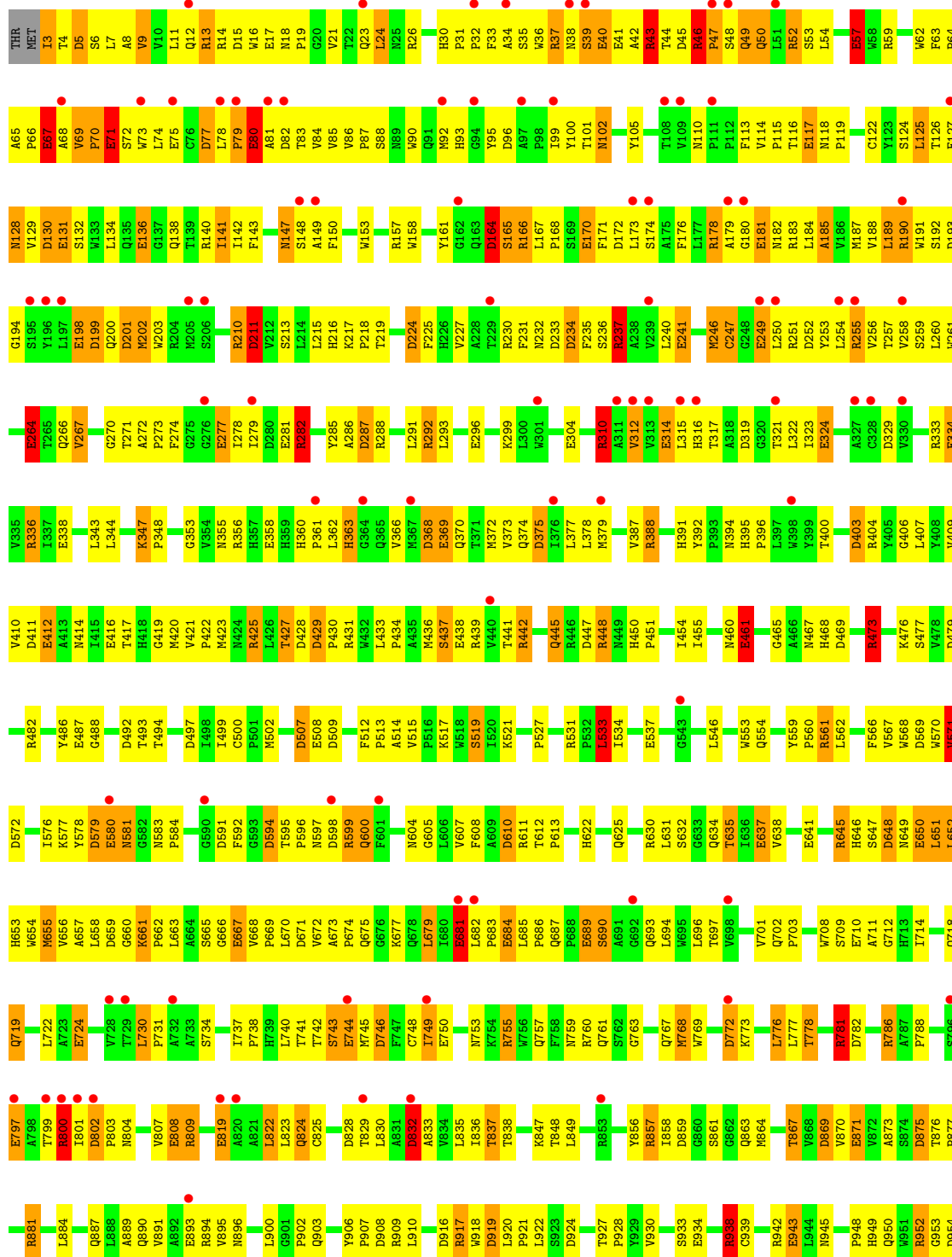






• Molecule 1: Beta-Galactosidase

Chain P:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.60Å 207.30Å 510.30Å 90.00° 95.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 68.54 – 2.50	Depositor EDS
% Data completeness (in resolution range)	66.0 (20.00-2.70) 66.8 (68.54-2.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 2.51Å)	Xtriage
Refinement program	TNT V. 5-E	Depositor
R, $R_{free}$	0.234 , (Not available) 0.446 , 0.431	Depositor DCC
$R_{free}$ test set	2111 reflections (0.48%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.0	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 54.3	EDS
Estimated twinning fraction	0.010 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 523624 reflections (0.000%)	Xtriage
$F_o, F_c$ correlation	0.58	EDS
Total number of atoms	67264	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.25% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities 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: NA, CME, MG, 2FL

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	I	1.11	56/8439 (0.7%)	1.47	141/11510 (1.2%)
1	J	1.11	56/8439 (0.7%)	1.47	139/11510 (1.2%)
1	K	1.11	55/8439 (0.7%)	1.47	139/11510 (1.2%)
1	L	1.11	56/8439 (0.7%)	1.47	139/11510 (1.2%)
1	M	1.11	56/8439 (0.7%)	1.47	141/11510 (1.2%)
1	N	1.11	56/8439 (0.7%)	1.47	141/11510 (1.2%)
1	O	1.11	56/8439 (0.7%)	1.47	140/11510 (1.2%)
1	P	1.11	55/8439 (0.7%)	1.47	143/11510 (1.2%)
All	All	1.11	446/67512 (0.7%)	1.47	1123/92080 (1.2%)

The worst 5 of 446 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	710	GLU	CD-OE2	7.97	1.34	1.25
1	N	710	GLU	CD-OE2	7.97	1.34	1.25
1	P	710	GLU	CD-OE2	7.94	1.34	1.25
1	K	710	GLU	CD-OE2	7.92	1.34	1.25
1	O	819	GLU	CD-OE2	7.91	1.34	1.25

The worst 5 of 1123 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	561	ARG	NE-CZ-NH1	10.76	125.68	120.30
1	M	561	ARG	NE-CZ-NH1	10.72	125.66	120.30
1	I	561	ARG	NE-CZ-NH1	10.71	125.65	120.30
1	N	561	ARG	NE-CZ-NH1	10.65	125.62	120.30
1	K	561	ARG	NE-CZ-NH1	10.61	125.61	120.30

There are no chirality outliers.

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	I	8219	0	7811	573	0
1	J	8219	0	7811	566	0
1	K	8219	0	7811	561	0
1	L	8219	0	7811	574	0
1	M	8219	0	7811	566	0
1	N	8219	0	7811	566	0
1	O	8219	0	7811	575	0
1	P	8219	0	7811	576	0
2	I	23	0	21	0	0
2	J	23	0	21	0	0
2	K	23	0	21	0	0
2	L	23	0	21	0	0
2	M	23	0	21	0	0
2	N	23	0	21	0	0
2	O	23	0	21	0	0
2	P	23	0	21	0	0
3	I	2	0	0	0	0
3	J	2	0	0	0	0
3	K	2	0	0	0	0
3	L	2	0	0	0	0
3	M	2	0	0	0	0
3	N	2	0	0	0	0
3	O	2	0	0	0	0
3	P	2	0	0	0	0
4	I	2	0	0	0	0
4	J	2	0	0	0	0
4	K	2	0	0	0	0
4	L	2	0	0	0	0
4	M	2	0	0	0	0
4	N	2	0	0	0	0
4	O	2	0	0	0	0
4	P	2	0	0	0	0
5	I	162	0	0	6	0
5	J	162	0	0	6	0
5	K	162	0	0	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	L	162	0	0	6	0
5	M	161	0	0	6	0
5	N	163	0	0	6	0
5	O	161	0	0	6	0
5	P	163	0	0	6	0
All	All	67264	0	62656	4491	0

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 35.

The worst 5 of 4491 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:J:43:ARG:HH11	1:J:43:ARG:HG2	1.20	1.07
1:I:43:ARG:HH11	1:I:43:ARG:HG2	1.20	1.07
1:M:43:ARG:HH11	1:M:43:ARG:HG2	1.20	1.06
1:P:427:THR:HA	1:P:436:MET:HE1	1.39	1.04
1:P:43:ARG:HH11	1:P:43:ARG:HG2	1.20	1.02

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	I	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	J	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	K	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	L	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	M	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	N	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	O	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55
1	P	1018/1023 (100%)	956 (94%)	53 (5%)	9 (1%)	25	55

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	8144/8184 (100%)	7648 (94%)	424 (5%)	72 (1%)	25	55

5 of 72 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	174	SER
1	J	174	SER
1	K	174	SER
1	L	174	SER
1	M	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	I	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	J	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	K	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	L	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	M	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	N	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	O	872/872 (100%)	759 (87%)	113 (13%)	6	15
1	P	872/872 (100%)	759 (87%)	113 (13%)	6	15
All	All	6976/6976 (100%)	6072 (87%)	904 (13%)	6	15

5 of 904 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	600	GLN
1	M	445	GLN
1	P	299	LYS
1	L	690	SER
1	M	43	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 159 such

sidechains are listed below:

Mol	Chain	Res	Type
1	L	624	GLN
1	M	467	ASN
1	P	363	HIS
1	L	761	GLN
1	M	49	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

24 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	I	1021	1	9,9,10	4.84	1 (11%)	7,9,11	1.30	0
1	CME	I	748	1	9,9,10	5.78	1 (11%)	7,9,11	2.67	4 (57%)
1	CME	I	914	1	9,9,10	6.34	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	J	1021	1	9,9,10	4.88	1 (11%)	7,9,11	1.30	0
1	CME	J	748	1	9,9,10	5.77	1 (11%)	7,9,11	2.67	4 (57%)
1	CME	J	914	1	9,9,10	6.38	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	K	1021	1	9,9,10	4.87	1 (11%)	7,9,11	1.31	0
1	CME	K	748	1	9,9,10	5.79	1 (11%)	7,9,11	2.67	4 (57%)
1	CME	K	914	1	9,9,10	6.41	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	L	1021	1	9,9,10	4.88	1 (11%)	7,9,11	1.31	0
1	CME	L	748	1	9,9,10	5.77	1 (11%)	7,9,11	2.66	4 (57%)
1	CME	L	914	1	9,9,10	6.36	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	M	1021	1	9,9,10	4.86	1 (11%)	7,9,11	1.30	0
1	CME	M	748	1	9,9,10	5.79	1 (11%)	7,9,11	2.67	4 (57%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	CME	M	914	1	9,9,10	6.36	1 (11%)	7,9,11	1.96	2 (28%)
1	CME	N	1021	1	9,9,10	4.85	1 (11%)	7,9,11	1.30	0
1	CME	N	748	1	9,9,10	5.75	1 (11%)	7,9,11	2.68	4 (57%)
1	CME	N	914	1	9,9,10	6.40	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	O	1021	1	9,9,10	4.90	1 (11%)	7,9,11	1.31	0
1	CME	O	748	1	9,9,10	5.79	1 (11%)	7,9,11	2.68	4 (57%)
1	CME	O	914	1	9,9,10	6.36	1 (11%)	7,9,11	1.95	2 (28%)
1	CME	P	1021	1	9,9,10	4.88	1 (11%)	7,9,11	1.30	0
1	CME	P	748	1	9,9,10	5.81	1 (11%)	7,9,11	2.67	4 (57%)
1	CME	P	914	1	9,9,10	6.37	1 (11%)	7,9,11	1.95	2 (28%)

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	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
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

*Continued on next page...*

*Continued from previous page...*

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

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	914	CME	O-C	19.16	1.24	1.11
1	N	914	CME	O-C	19.13	1.24	1.11
1	J	914	CME	O-C	19.08	1.24	1.11
1	P	914	CME	O-C	19.05	1.24	1.11
1	L	914	CME	O-C	19.03	1.24	1.11

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	N	748	CME	CB-CA-N	4.88	118.54	110.27
1	O	748	CME	CB-CA-N	4.87	118.53	110.27
1	P	748	CME	CB-CA-N	4.85	118.49	110.27
1	I	748	CME	CB-CA-N	4.84	118.48	110.27
1	K	748	CME	CB-CA-N	4.84	118.47	110.27

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 40 ligands modelled in this entry, 32 are monoatomic - leaving 8 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
2	2FL	I	2001	4	24,24,24	0.56	0	35,35,35	1.43	3 (8%)
2	2FL	J	2001	4	24,24,24	0.57	0	35,35,35	1.43	3 (8%)
2	2FL	K	2001	4	24,24,24	0.56	0	35,35,35	1.43	3 (8%)
2	2FL	L	2001	4	24,24,24	0.56	0	35,35,35	1.43	3 (8%)
2	2FL	M	2001	4	24,24,24	0.56	0	35,35,35	1.43	3 (8%)
2	2FL	N	2001	4	24,24,24	0.57	0	35,35,35	1.43	3 (8%)
2	2FL	O	2001	4	24,24,24	0.56	0	35,35,35	1.43	3 (8%)
2	2FL	P	2001	4	24,24,24	0.57	0	35,35,35	1.43	3 (8%)

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	2FL	I	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	J	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	K	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	L	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	M	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	N	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	O	2001	4	-	0/8/48/48	0/2/2/2
2	2FL	P	2001	4	-	0/8/48/48	0/2/2/2

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	N	2001	2FL	F2-C2-C3	6.08	113.39	108.55
2	M	2001	2FL	F2-C2-C3	6.07	113.38	108.55
2	L	2001	2FL	F2-C2-C3	6.07	113.38	108.55
2	J	2001	2FL	F2-C2-C3	6.07	113.38	108.55
2	K	2001	2FL	F2-C2-C3	6.06	113.37	108.55

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	I	1021/1023 (99%)	-0.00	23 (2%) 57 64	11, 32, 70, 100	0
1	J	1021/1023 (99%)	-0.09	18 (1%) 65 71	10, 30, 69, 100	0
1	K	1021/1023 (99%)	0.31	48 (4%) 30 34	24, 45, 81, 100	0
1	L	1021/1023 (99%)	0.36	47 (4%) 31 35	25, 45, 81, 100	0
1	M	1021/1023 (99%)	0.45	60 (5%) 22 23	27, 48, 83, 100	0
1	N	1021/1023 (99%)	0.05	22 (2%) 59 65	16, 36, 74, 100	0
1	O	1021/1023 (99%)	0.22	27 (2%) 53 59	21, 42, 78, 100	0
1	P	1021/1023 (99%)	0.76	91 (8%) 10 10	39, 60, 91, 100	0
All	All	8168/8184 (99%)	0.26	336 (4%) 35 40	10, 43, 80, 100	0

The worst 5 of 336 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	801	ILE	9.6
1	O	796	SER	7.8
1	P	797	GLU	7.0
1	O	1023	LYS	6.9
1	N	799	THR	6.7

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
1	CME	O	914	10/11	0.20	1.89	36,44,80,86	0
1	CME	M	748	10/11	0.26	1.45	56,68,100,100	0
1	CME	L	748	10/11	0.26	1.00	53,65,99,100	0
1	CME	K	1021	10/11	0.27	0.96	57,73,100,100	0
1	CME	I	1021	10/11	0.20	0.57	44,60,94,100	0
1	CME	K	748	10/11	0.24	0.53	53,65,98,100	0
1	CME	N	914	10/11	0.20	0.48	31,38,75,81	0
1	CME	L	1021	10/11	0.24	0.33	57,73,100,100	0
1	CME	N	748	10/11	0.20	0.23	44,56,90,100	0
1	CME	J	1021	10/11	0.22	0.19	42,58,93,100	0
1	CME	N	1021	10/11	0.24	0.17	48,64,99,100	0
1	CME	I	914	10/11	0.18	0.11	26,34,70,76	0
1	CME	J	914	10/11	0.16	0.05	25,32,69,75	0
1	CME	M	1021	10/11	0.23	-0.11	60,75,100,100	0
1	CME	L	914	10/11	0.19	-0.15	40,47,84,89	0
1	CME	J	748	10/11	0.18	-0.20	38,50,84,95	0
1	CME	I	748	10/11	0.17	-0.27	40,52,85,97	0
1	CME	P	748	10/11	0.21	-0.32	68,80,100,100	0
1	CME	P	1021	10/11	0.23	-0.46	72,87,100,100	0
1	CME	O	1021	10/11	0.17	-0.52	54,69,100,100	0
1	CME	M	914	10/11	0.17	-0.59	42,50,86,92	0
1	CME	K	914	10/11	0.13	-0.79	40,47,84,89	0
1	CME	P	914	10/11	0.16	-0.98	54,62,98,100	0
1	CME	O	748	10/11	0.17	-1.14	50,62,95,100	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	2FL	M	2001	23/23	0.63	8.39	65,77,93,100	0
2	2FL	N	2001	23/23	0.42	6.03	54,65,82,100	0
2	2FL	O	2001	23/23	0.46	5.15	59,70,87,100	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NA	N	3101	1/1	0.39	4.60	44,44,44,44	0
4	NA	K	3102	1/1	0.29	4.25	40,40,40,40	0
2	2FL	L	2001	23/23	0.41	3.99	63,74,91,100	0
2	2FL	J	2001	23/23	0.39	3.39	48,59,76,99	0
2	2FL	K	2001	23/23	0.40	2.96	63,74,91,100	0
3	MG	N	3001	1/1	0.28	2.77	37,37,37,37	0
2	2FL	P	2001	23/23	0.38	2.51	77,89,100,100	0
3	MG	O	3001	1/1	0.29	2.40	42,42,42,42	0
4	NA	J	3102	1/1	0.18	2.15	25,25,25,25	0
4	NA	K	3101	1/1	0.30	1.63	53,53,53,53	0
4	NA	O	3101	1/1	0.26	1.22	49,49,49,49	0
4	NA	I	3102	1/1	0.16	1.19	27,27,27,27	0
3	MG	P	3001	1/1	0.28	1.08	60,60,60,60	0
2	2FL	I	2001	23/23	0.26	0.92	49,61,78,100	0
4	NA	I	3101	1/1	0.23	0.63	39,39,39,39	0
4	NA	M	3101	1/1	0.20	0.51	55,55,55,55	0
3	MG	O	3002	1/1	0.21	0.35	35,35,35,35	0
3	MG	I	3001	1/1	0.17	0.30	32,32,32,32	0
4	NA	P	3101	1/1	0.25	0.10	67,67,67,67	0
4	NA	L	3101	1/1	0.20	0.08	53,53,53,53	0
4	NA	P	3102	1/1	0.21	0.05	55,55,55,55	0
4	NA	O	3102	1/1	0.18	-0.20	37,37,37,37	0
4	NA	M	3102	1/1	0.19	-0.22	43,43,43,43	0
4	NA	N	3102	1/1	0.15	-0.55	31,31,31,31	0
3	MG	N	3002	1/1	0.17	-0.63	30,30,30,30	0
3	MG	K	3001	1/1	0.16	-0.77	45,45,45,45	0
3	MG	K	3002	1/1	0.16	-0.93	39,39,39,39	0
3	MG	L	3002	1/1	0.15	-1.00	39,39,39,39	0
3	MG	J	3001	1/1	0.06	-1.85	31,31,31,31	0
4	NA	J	3101	1/1	0.10	-1.91	38,38,38,38	0
3	MG	L	3001	1/1	0.10	-2.03	46,46,46,46	0
3	MG	M	3001	1/1	0.13	-2.36	48,48,48,48	0
3	MG	I	3002	1/1	0.06	-2.76	26,26,26,26	0
3	MG	P	3002	1/1	0.10	-3.48	53,53,53,53	0
3	MG	J	3002	1/1	0.08	-3.54	24,24,24,24	0
3	MG	M	3002	1/1	0.10	-3.73	41,41,41,41	0
4	NA	L	3102	1/1	0.10	-3.74	40,40,40,40	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.