



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:28 PM GMT

PDB ID : 4RYF  
Title : ClpP1/2 heterocomplex from *Listeria monocytogenes*  
Authors : Dahmen, M.; Vielberg, M.-T.; Groll, M.; Sieber, S.A.  
Deposited on : 2014-12-15  
Resolution : 2.80 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

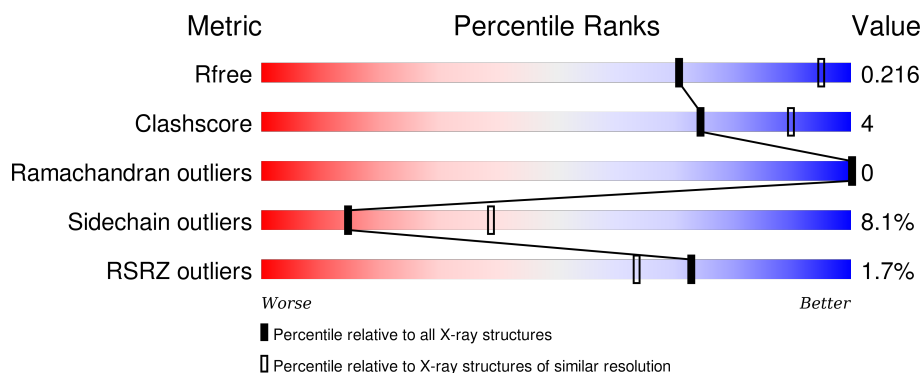
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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. 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  $\leq 5\%$ . 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	201	<div> <div>2%</div> <div>72%</div> <div>15%</div> <div>•</div> <div>11%</div> </div>
1	B	201	<div> <div>2%</div> <div>72%</div> <div>15%</div> <div>•</div> <div>11%</div> </div>
1	C	201	<div> <div>2%</div> <div>73%</div> <div>13%</div> <div>•</div> <div>11%</div> </div>
1	D	201	<div> <div>2%</div> <div>72%</div> <div>15%</div> <div>•</div> <div>11%</div> </div>
1	E	201	<div> <div>2%</div> <div>72%</div> <div>15%</div> <div>•</div> <div>11%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	201	
1	G	201	
2	H	204	
2	I	204	
2	J	204	
2	K	204	
2	L	204	
2	M	204	
2	N	204	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MLI	H	301	-	-	-	X
4	MLI	K	301	-	-	-	X
4	MLI	L	301	-	-	-	X
4	MLI	M	301	-	-	-	X
4	MLI	N	301	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 19892 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 ATP-dependent Clp protease proteolytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	179	Total	C	N	O	S	0	0	0
			1412	887	243	279	3			
1	B	178	Total	C	N	O	S	0	0	0
			1408	885	242	278	3			
1	C	178	Total	C	N	O	S	0	0	0
			1408	885	242	278	3			
1	D	179	Total	C	N	O	S	0	0	0
			1412	887	243	279	3			
1	E	178	Total	C	N	O	S	0	0	0
			1408	885	242	278	3			
1	F	178	Total	C	N	O	S	0	0	0
			1408	885	242	278	3			
1	G	178	Total	C	N	O	S	0	0	0
			1408	885	242	278	3			

There are 77 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
A	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
A	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
A	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
A	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
A	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
A	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
A	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
A	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
A	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
A	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
B	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
B	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
B	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
B	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
B	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
B	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
B	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
B	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
B	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
B	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
C	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
C	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
C	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
C	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
C	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
C	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
C	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
C	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
C	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
C	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
C	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
D	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
D	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
D	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
D	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
D	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
D	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
D	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
D	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
D	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
D	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
D	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
E	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
E	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
E	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
E	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
E	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
E	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
E	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
E	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
E	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
E	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
E	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
F	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
F	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
F	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
F	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
F	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
F	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
F	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
F	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
F	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
F	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1
G	198	MET	-	EXPRESSION TAG	UNP Q8Y7Y1
G	199	ALA	-	EXPRESSION TAG	UNP Q8Y7Y1
G	200	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
G	201	TRP	-	EXPRESSION TAG	UNP Q8Y7Y1
G	202	SER	-	EXPRESSION TAG	UNP Q8Y7Y1
G	203	HIS	-	EXPRESSION TAG	UNP Q8Y7Y1
G	204	PRO	-	EXPRESSION TAG	UNP Q8Y7Y1
G	205	GLN	-	EXPRESSION TAG	UNP Q8Y7Y1
G	206	PHE	-	EXPRESSION TAG	UNP Q8Y7Y1
G	207	GLU	-	EXPRESSION TAG	UNP Q8Y7Y1
G	208	LYS	-	EXPRESSION TAG	UNP Q8Y7Y1

- Molecule 2 is a protein called ATP-dependent Clp protease proteolytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	183	Total	C	N	O	S	0	0	0
			1398	884	234	270	10			
2	I	183	Total	C	N	O	S	0	0	0
			1398	884	234	270	10			
2	J	184	Total	C	N	O	S	0	0	0
			1404	887	235	272	10			
2	K	184	Total	C	N	O	S	0	0	0
			1404	887	235	272	10			
2	L	184	Total	C	N	O	S	0	0	0
			1404	887	235	272	10			
2	M	183	Total	C	N	O	S	0	0	0
			1398	884	234	270	10			
2	N	184	Total	C	N	O	S	0	0	0
			1404	887	235	272	10			

There are 42 discrepancies between the modelled and reference sequences:

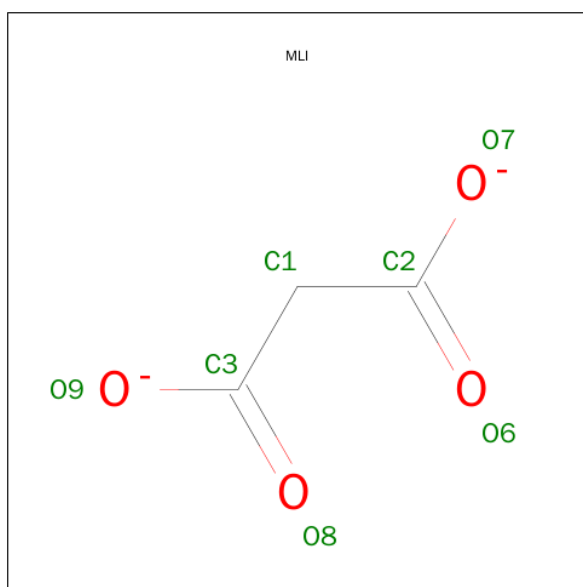
Chain	Residue	Modelled	Actual	Comment	Reference
H	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
H	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
H	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
H	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
H	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
H	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
I	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
J	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
K	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
L	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
M	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	199	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	200	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	201	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	202	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	203	HIS	-	EXPRESSION TAG	UNP Q9RQI6
N	204	HIS	-	EXPRESSION TAG	UNP Q9RQI6

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Na 1 1	0	0
3	J	1	Total Na 1 1	0	0
3	D	1	Total Na 1 1	0	0
3	K	1	Total Na 1 1	0	0
3	E	1	Total Na 1 1	0	0
3	H	1	Total Na 1 1	0	0
3	B	1	Total Na 1 1	0	0
3	I	1	Total Na 1 1	0	0
3	C	1	Total Na 1 1	0	0
3	A	1	Total Na 1 1	0	0
3	N	1	Total Na 1 1	0	0
3	L	1	Total Na 1 1	0	0
3	F	1	Total Na 1 1	0	0
3	M	1	Total Na 1 1	0	0

- Molecule 4 is MALONATE ION (three-letter code: MLI) (formula: C<sub>3</sub>H<sub>2</sub>O<sub>4</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	H	1	Total	C	O	0	0
			7	3	4		
4	I	1	Total	C	O	0	0
			7	3	4		
4	J	1	Total	C	O	0	0
			7	3	4		
4	K	1	Total	C	O	0	0
			7	3	4		
4	L	1	Total	C	O	0	0
			7	3	4		
4	M	1	Total	C	O	0	0
			7	3	4		
4	N	1	Total	C	O	0	0
			7	3	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	13	Total	O	0	0
			13	13		
5	B	7	Total	O	0	0
			7	7		
5	C	7	Total	O	0	0
			7	7		
5	D	9	Total	O	0	0
			9	9		
5	E	11	Total	O	0	0
			11	11		

*Continued on next page...*

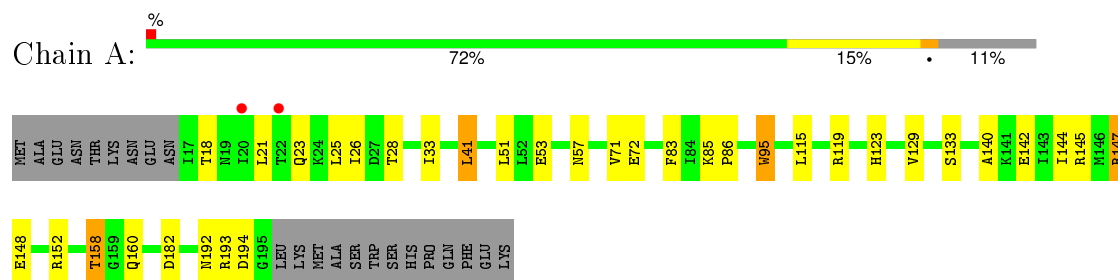
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	17	Total 17	O 17	0	0
5	G	23	Total 23	O 23	0	0
5	H	3	Total 3	O 3	0	0
5	I	9	Total 9	O 9	0	0
5	J	16	Total 16	O 16	0	0
5	K	19	Total 19	O 19	0	0
5	L	10	Total 10	O 10	0	0
5	M	8	Total 8	O 8	0	0
5	N	3	Total 3	O 3	0	0

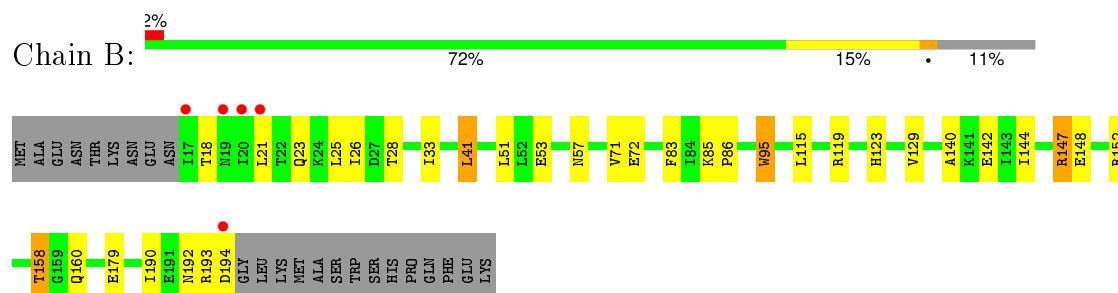
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

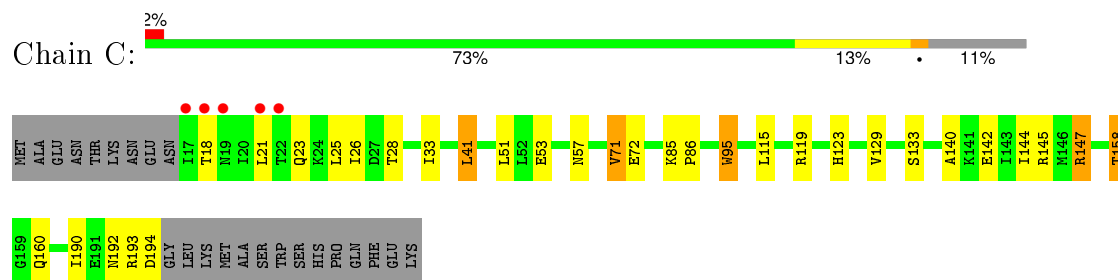
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



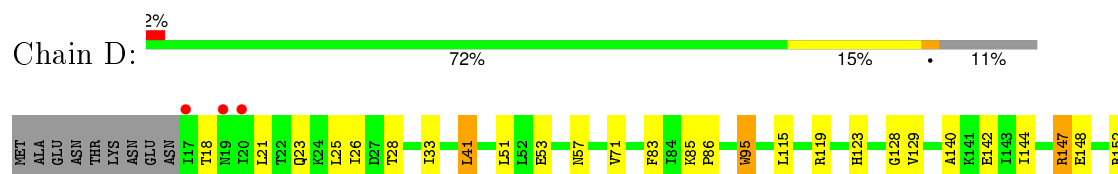
- Molecule 1: ATP-dependent Clp protease proteolytic subunit

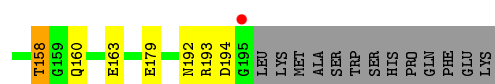


- Molecule 1: ATP-dependent Clp protease proteolytic subunit

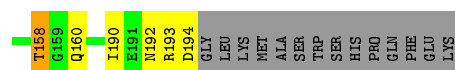
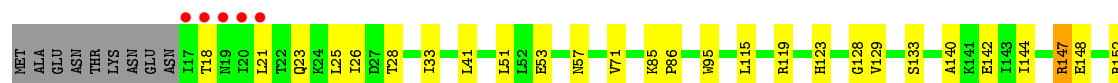


- Molecule 1: ATP-dependent Clp protease proteolytic subunit

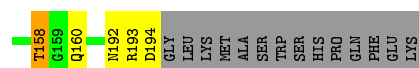
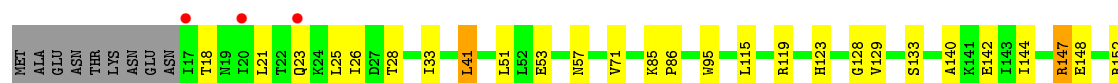




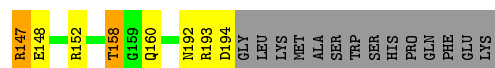
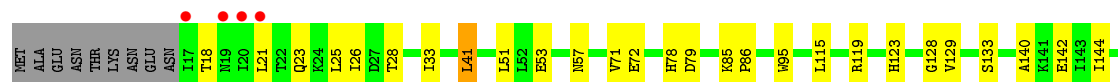
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



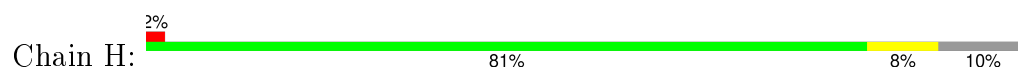
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



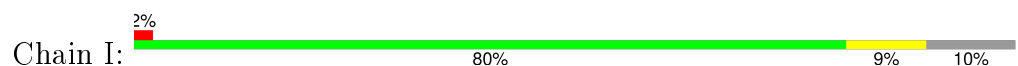
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



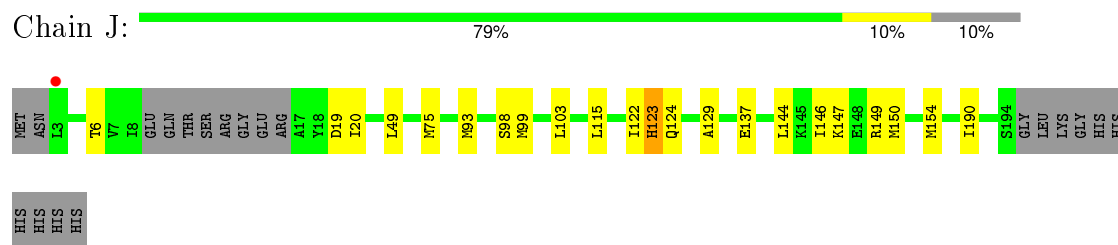
- Molecule 2: ATP-dependent Clp protease proteolytic subunit



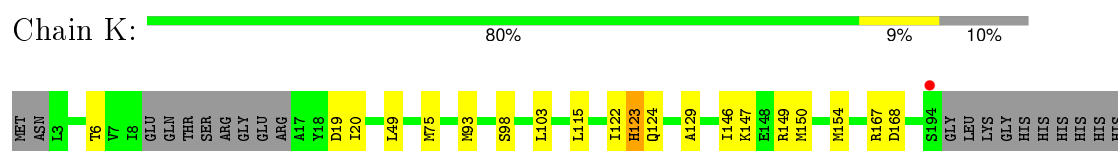
- Molecule 2: ATP-dependent Clp protease proteolytic subunit



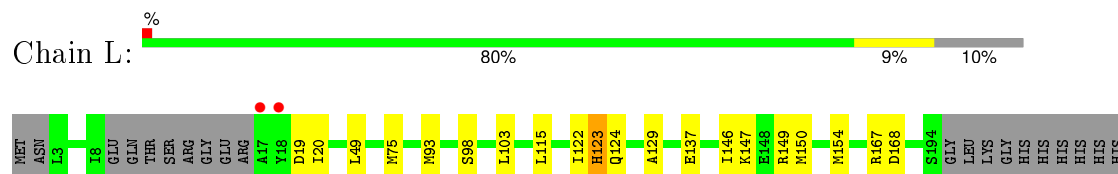
- Molecule 2: ATP-dependent Clp protease proteolytic subunit



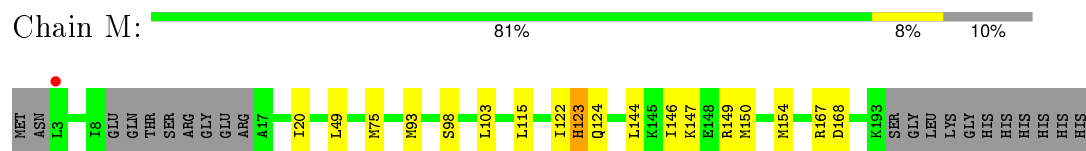
- Molecule 2: ATP-dependent Clp protease proteolytic subunit



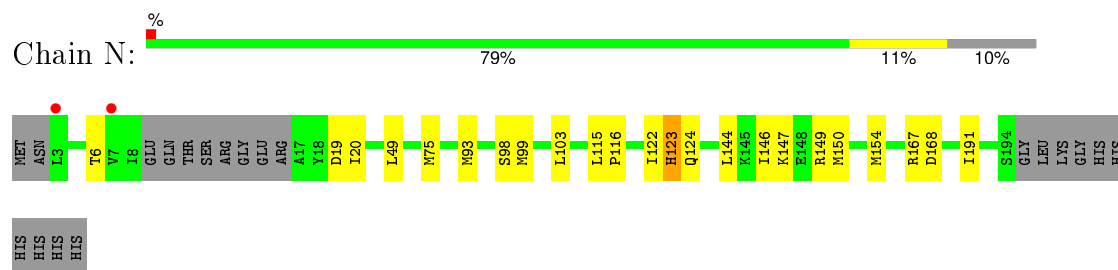
- Molecule 2: ATP-dependent Clp protease proteolytic subunit



- Molecule 2: ATP-dependent Clp protease proteolytic subunit



- Molecule 2: ATP-dependent Clp protease proteolytic subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.23Å 127.15Å 265.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 49.39 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (30.00-2.80) 96.2 (49.39-2.80)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.16 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.180 , 0.216 0.179 , 0.216	Depositor DCC
$R_{free}$ test set	3934 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.0	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 31.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 78735 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	19892	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.62% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: NA, MLI

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.28	0/1431	0.50	0/1936
1	B	0.28	0/1427	0.50	0/1931
1	C	0.28	0/1427	0.50	0/1931
1	D	0.28	0/1431	0.51	0/1936
1	E	0.28	0/1427	0.51	0/1931
1	F	0.28	0/1427	0.51	0/1931
1	G	0.28	0/1427	0.50	0/1931
2	H	0.26	0/1416	0.46	0/1910
2	I	0.27	0/1416	0.47	0/1910
2	J	0.27	0/1422	0.48	0/1918
2	K	0.27	0/1422	0.47	0/1918
2	L	0.26	0/1422	0.47	0/1918
2	M	0.26	0/1416	0.47	0/1910
2	N	0.26	0/1422	0.47	0/1918
All	All	0.27	0/19933	0.49	0/26929

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
2	H	0	1
2	I	0	1
2	J	0	1
2	K	0	1
2	L	0	1
2	M	0	1
2	N	0	1
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	122	ILE	Peptide
2	I	122	ILE	Peptide
2	J	122	ILE	Peptide
2	K	122	ILE	Peptide
2	L	122	ILE	Peptide

## 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1412	0	1418	21	1
1	B	1408	0	1415	20	0
1	C	1408	0	1415	20	0
1	D	1412	0	1418	21	1
1	E	1408	0	1415	19	0
1	F	1408	0	1415	18	0
1	G	1408	0	1415	23	0
2	H	1398	0	1413	7	0
2	I	1398	0	1413	8	0
2	J	1404	0	1418	9	0
2	K	1404	0	1418	8	0
2	L	1404	0	1418	7	0
2	M	1398	0	1413	6	0
2	N	1404	0	1418	8	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
3	M	1	0	0	0	0
3	N	1	0	0	0	0
4	H	7	0	2	0	0
4	I	7	0	2	0	0
4	J	7	0	2	0	0
4	K	7	0	2	0	0
4	L	7	0	2	0	0
4	M	7	0	2	0	0
4	N	7	0	2	0	0
5	A	13	0	0	0	0
5	B	7	0	0	0	0
5	C	7	0	0	1	0
5	D	9	0	0	0	0
5	E	11	0	0	0	0
5	F	17	0	0	0	0
5	G	23	0	0	2	0
5	H	3	0	0	0	0
5	I	9	0	0	0	0
5	J	16	0	0	0	0
5	K	19	0	0	0	0
5	L	10	0	0	0	0
5	M	8	0	0	0	0
5	N	3	0	0	0	0
All	All	19892	0	19836	159	1

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 4.

The worst 5 of 159 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:B:147:ARG:HH21	1:B:147:ARG:HG2	1.22	1.05
1:A:147:ARG:HH21	1:A:147:ARG:HG2	1.22	1.03
1:F:147:ARG:HG2	1:F:147:ARG:HH21	1.22	1.01
1:E:147:ARG:HH21	1:E:147:ARG:HG2	1.22	1.01
1:G:147:ARG:HG2	1:G:147:ARG:HH21	1.22	1.01

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:182:ASP:O	1:D:163:GLU:OE1[1_655]	2.16	0.04

## 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	177/201 (88%)	171 (97%)	6 (3%)	0	100	100
1	B	176/201 (88%)	171 (97%)	5 (3%)	0	100	100
1	C	176/201 (88%)	171 (97%)	5 (3%)	0	100	100
1	D	177/201 (88%)	171 (97%)	6 (3%)	0	100	100
1	E	176/201 (88%)	171 (97%)	5 (3%)	0	100	100
1	F	176/201 (88%)	171 (97%)	5 (3%)	0	100	100
1	G	176/201 (88%)	171 (97%)	5 (3%)	0	100	100
2	H	179/204 (88%)	175 (98%)	4 (2%)	0	100	100
2	I	179/204 (88%)	174 (97%)	5 (3%)	0	100	100
2	J	180/204 (88%)	175 (97%)	5 (3%)	0	100	100
2	K	180/204 (88%)	177 (98%)	3 (2%)	0	100	100
2	L	180/204 (88%)	176 (98%)	4 (2%)	0	100	100
2	M	179/204 (88%)	175 (98%)	4 (2%)	0	100	100
2	N	180/204 (88%)	176 (98%)	4 (2%)	0	100	100
All	All	2491/2835 (88%)	2425 (97%)	66 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	155/175 (89%)	138 (89%)	17 (11%)	8	23
1	B	155/175 (89%)	138 (89%)	17 (11%)	8	23
1	C	155/175 (89%)	138 (89%)	17 (11%)	8	23
1	D	155/175 (89%)	138 (89%)	17 (11%)	8	23
1	E	155/175 (89%)	139 (90%)	16 (10%)	9	26
1	F	155/175 (89%)	138 (89%)	17 (11%)	8	23
1	G	155/175 (89%)	138 (89%)	17 (11%)	8	23
2	H	148/166 (89%)	141 (95%)	7 (5%)	32	67
2	I	148/166 (89%)	140 (95%)	8 (5%)	27	60
2	J	149/166 (90%)	140 (94%)	9 (6%)	24	56
2	K	149/166 (90%)	142 (95%)	7 (5%)	32	67
2	L	149/166 (90%)	141 (95%)	8 (5%)	27	60
2	M	148/166 (89%)	140 (95%)	8 (5%)	27	60
2	N	149/166 (90%)	141 (95%)	8 (5%)	27	60
All	All	2125/2387 (89%)	1952 (92%)	173 (8%)	15	39

5 of 173 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	95	TRP
1	F	115	LEU
2	M	75	MET
1	E	123	HIS
1	F	25	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	82	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	I	82	ASN
2	M	160	GLN
2	H	160	GLN
2	H	192	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 14 are monoatomic - leaving 7 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	MLI	H	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	I	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	J	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	K	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	L	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	M	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	MLI	N	301	-	0,6,6	0.00	-	0,7,7	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical

component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MLI	H	301	-	-	0/0/4/4	0/0/0/0
4	MLI	I	301	-	-	0/0/4/4	0/0/0/0
4	MLI	J	301	-	-	0/0/4/4	0/0/0/0
4	MLI	K	301	-	-	0/0/4/4	0/0/0/0
4	MLI	L	301	-	-	0/0/4/4	0/0/0/0
4	MLI	M	301	-	-	0/0/4/4	0/0/0/0
4	MLI	N	301	-	-	0/0/4/4	0/0/0/0

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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 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	A	179/201 (89%)	-0.20	2 (1%) 82 74	28, 42, 94, 123	0
1	B	178/201 (88%)	-0.13	5 (2%) 56 44	32, 47, 87, 138	0
1	C	178/201 (88%)	-0.16	5 (2%) 56 44	38, 49, 97, 146	0
1	D	179/201 (89%)	-0.13	4 (2%) 65 54	32, 46, 91, 126	0
1	E	178/201 (88%)	-0.20	5 (2%) 56 44	27, 38, 93, 140	0
1	F	178/201 (88%)	-0.25	3 (1%) 73 63	24, 34, 79, 136	0
1	G	178/201 (88%)	-0.18	4 (2%) 65 54	24, 34, 82, 148	0
2	H	183/204 (89%)	-0.05	4 (2%) 65 54	31, 51, 82, 127	0
2	I	183/204 (89%)	-0.23	4 (2%) 65 54	23, 39, 75, 124	0
2	J	184/204 (90%)	-0.26	1 (0%) 91 88	23, 33, 74, 120	0
2	K	184/204 (90%)	-0.26	1 (0%) 91 88	23, 33, 74, 114	0
2	L	184/204 (90%)	-0.21	2 (1%) 82 74	27, 37, 79, 119	0
2	M	183/204 (89%)	-0.26	1 (0%) 91 88	30, 43, 78, 125	0
2	N	184/204 (90%)	-0.13	2 (1%) 82 74	35, 52, 85, 110	0
All	All	2533/2835 (89%)	-0.19	43 (1%) 73 63	23, 42, 88, 148	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	17	ILE	8.7
2	H	17	ALA	6.8
1	B	20	ILE	5.5
1	G	20	ILE	5.5
2	J	3	LEU	5.1

## 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 carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	MLI	N	301	7/7	0.90	0.35	6.68	77,81,84,87	0
4	MLI	M	301	7/7	0.91	0.29	6.10	77,82,89,91	0
4	MLI	K	301	7/7	0.88	0.35	5.25	62,68,78,80	0
4	MLI	H	301	7/7	0.90	0.35	4.79	65,75,83,85	0
4	MLI	L	301	7/7	0.89	0.25	2.76	72,74,82,87	0
4	MLI	I	301	7/7	0.88	0.23	1.37	61,66,72,79	0
4	MLI	J	301	7/7	0.90	0.20	0.77	48,56,69,81	0
3	NA	F	301	1/1	0.73	0.18	0.22	57,57,57,57	0
3	NA	B	301	1/1	0.92	0.16	-0.54	51,51,51,51	0
3	NA	C	301	1/1	0.88	0.16	-0.57	46,46,46,46	0
3	NA	M	302	1/1	0.97	0.13	-0.93	44,44,44,44	0
3	NA	I	302	1/1	0.97	0.11	-1.26	25,25,25,25	0
3	NA	J	302	1/1	0.97	0.10	-1.59	24,24,24,24	0
3	NA	D	301	1/1	0.94	0.09	-1.69	46,46,46,46	0
3	NA	G	301	1/1	0.93	0.15	-1.81	44,44,44,44	0
3	NA	A	301	1/1	0.92	0.09	-1.85	51,51,51,51	0
3	NA	K	302	1/1	0.96	0.08	-2.64	26,26,26,26	0
3	NA	H	302	1/1	0.96	0.08	-3.16	38,38,38,38	0
3	NA	E	301	1/1	0.97	0.07	-3.32	34,34,34,34	0
3	NA	N	302	1/1	0.96	0.06	-3.38	37,37,37,37	0
3	NA	L	302	1/1	0.96	0.06	-3.40	29,29,29,29	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.