



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 21, 2024 – 06:24 am BST

PDB ID : 5G1S  
Title : Open conformation of Francisella tularensis ClpP at 1.7 Å  
Authors : Diaz-Saez, L.; Pankov, G.; Hunter, W.N.  
Deposited on : 2016-03-30  
Resolution : 1.70 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36.2

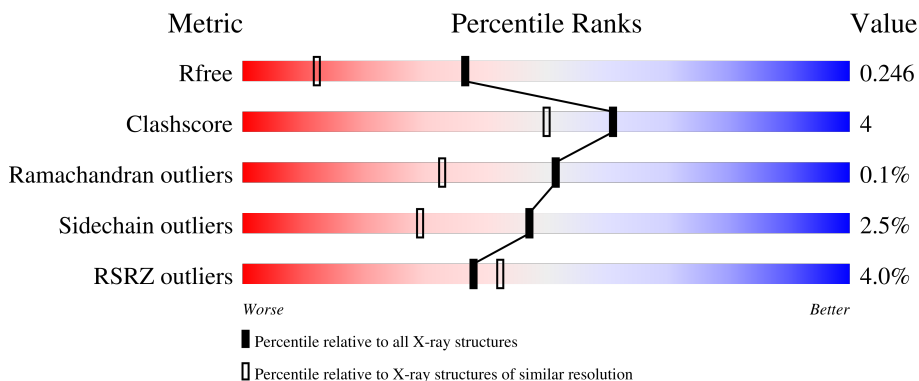
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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))
$R_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . 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	 2% 87% 6% • 6%
1	B	201	 % 87% • 9%
1	C	201	 4% 88% 5% 6%
1	D	201	 4% 84% 9% 7%
1	E	201	 5% 83% 9% • 6%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	201	
1	G	201	
1	H	201	
1	I	201	
1	J	201	
1	K	201	
1	L	201	
1	M	201	
1	N	201	
1	O	201	
1	P	201	
1	Q	201	
1	R	201	
1	S	201	
1	T	201	
1	U	201	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 37499 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.

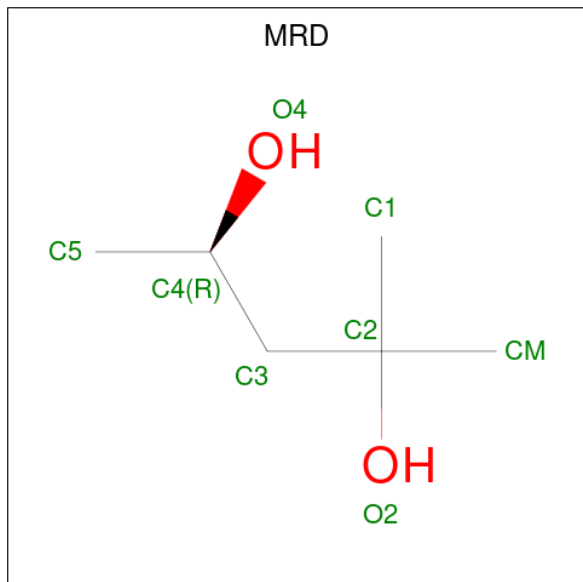
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	188	Total 1586	C 1015	N 266	O 293	S 12	4	21	0
1	B	183	Total 1510	C 965	N 257	O 276	S 12	4	16	0
1	C	189	Total 1577	C 1010	N 269	O 286	S 12	7	19	0
1	D	187	Total 1568	C 1004	N 267	O 285	S 12	7	21	0
1	E	189	Total 1575	C 1010	N 265	O 289	S 11	4	20	0
1	F	191	Total 1631	C 1040	N 283	O 297	S 11	4	23	0
1	G	193	Total 1628	C 1036	N 280	O 300	S 12	1	21	0
1	H	190	Total 1603	C 1024	N 275	O 292	S 12	5	22	0
1	I	184	Total 1552	C 997	N 264	O 279	S 12	7	22	0
1	J	188	Total 1570	C 1007	N 261	O 289	S 13	5	21	0
1	K	193	Total 1598	C 1017	N 273	O 297	S 11	10	19	0
1	L	188	Total 1545	C 988	N 264	O 283	S 10	5	17	0
1	M	192	Total 1613	C 1030	N 277	O 294	S 12	2	21	0
1	N	193	Total 1608	C 1024	N 277	O 295	S 12	4	22	0
1	O	189	Total 1593	C 1014	N 272	O 295	S 12	3	21	0
1	P	181	Total 1501	C 959	N 254	O 276	S 12	2	19	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	188	Total	C	N	O	S	10	15	0
			1544	986	264	282	12			
1	R	186	Total	C	N	O	S	7	18	0
			1543	983	263	286	11			
1	S	190	Total	C	N	O	S	4	19	0
			1573	1010	264	287	12			
1	T	191	Total	C	N	O	S	1	20	0
			1598	1017	276	293	12			
1	U	193	Total	C	N	O	S	0	24	0
			1631	1039	278	302	12			

- Molecule 2 is (4R)-2-METHYLPENTANE-2,4-DIOL (three-letter code: MRD) (formula:  $C_6H_{14}O_2$ ).



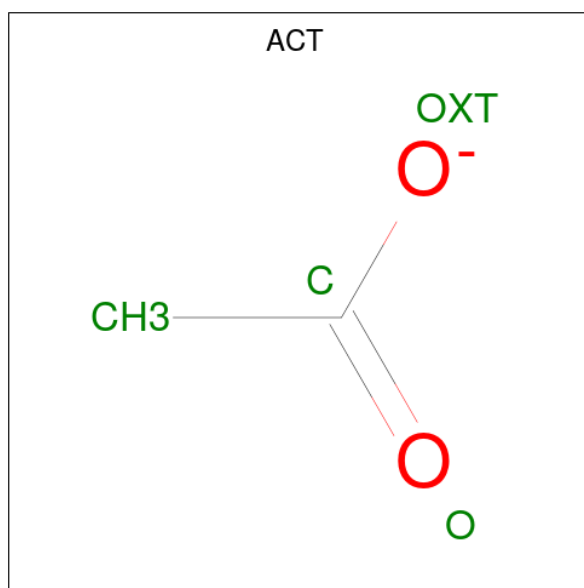
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			8	6	2		
2	C	1	Total	C	O	0	0
			8	6	2		
2	C	1	Total	C	O	0	0
			8	6	2		
2	D	1	Total	C	O	0	0
			8	6	2		
2	G	1	Total	C	O	0	0
			8	6	2		
2	I	1	Total	C	O	3	0
			8	6	2		

Continued on next page...

Continued from previous page...

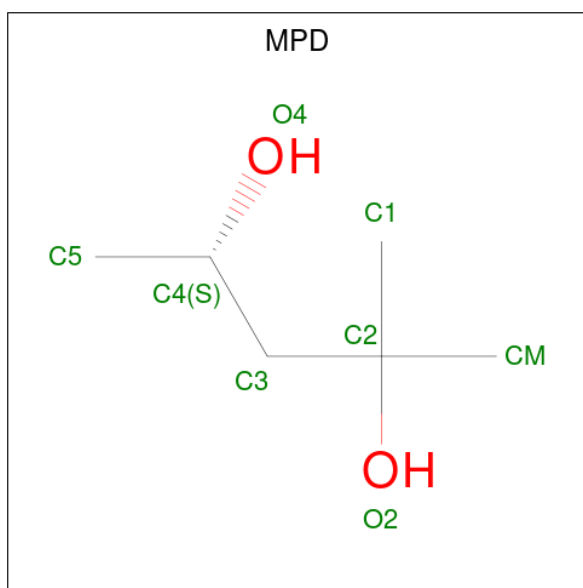
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	K	1	Total C O 8 6 2	0	0
2	K	1	Total C O 8 6 2	0	0
2	K	1	Total C O 8 6 2	0	0
2	L	1	Total C O 8 6 2	0	0
2	N	1	Total C O 8 6 2	0	0
2	O	1	Total C O 8 6 2	0	0
2	P	1	Total C O 8 6 2	0	0
2	Q	1	Total C O 8 6 2	0	0
2	S	1	Total C O 8 6 2	0	0
2	T	1	Total C O 8 6 2	0	0
2	T	1	Total C O 8 6 2	0	0
2	U	1	Total C O 8 6 2	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	E	1	Total C O 4 2 2	0	0
3	F	1	Total C O 4 2 2	0	0
3	G	1	Total C O 4 2 2	0	0
3	H	1	Total C O 4 2 2	0	0
3	I	1	Total C O 4 2 2	0	0
3	J	1	Total C O 4 2 2	0	0
3	K	1	Total C O 4 2 2	0	0
3	L	1	Total C O 4 2 2	0	0
3	M	1	Total C O 4 2 2	0	0
3	N	1	Total C O 4 2 2	0	0
3	O	1	Total C O 4 2 2	0	0
3	P	1	Total C O 4 2 2	0	0
3	Q	1	Total C O 4 2 2	0	0
3	R	1	Total C O 4 2 2	0	0
3	S	1	Total C O 4 2 2	0	0
3	T	1	Total C O 4 2 2	0	0
3	U	1	Total C O 4 2 2	0	0

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:

C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 8 6 2	0	0
4	C	1	Total C O 8 6 2	0	0
4	E	1	Total C O 8 6 2	0	0
4	F	1	Total C O 8 6 2	0	0
4	F	1	Total C O 8 6 2	0	0
4	G	1	Total C O 8 6 2	0	0
4	I	1	Total C O 8 6 2	0	0
4	J	1	Total C O 8 6 2	0	0
4	L	1	Total C O 8 6 2	0	0
4	M	1	Total C O 8 6 2	0	0
4	N	1	Total C O 8 6 2	0	0
4	S	1	Total C O 8 6 2	0	0
4	T	1	Total C O 8 6 2	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	U	1	Total	C	O	0	0
			8	6	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	256	Total	O	0	0
			256	256		
5	B	218	Total	O	0	0
			218	218		
5	C	209	Total	O	0	0
			209	209		
5	D	168	Total	O	0	0
			168	168		
5	E	185	Total	O	0	0
			185	185		
5	F	214	Total	O	0	0
			214	214		
5	G	189	Total	O	0	0
			189	189		
5	H	252	Total	O	0	0
			252	252		
5	I	167	Total	O	0	0
			167	167		
5	J	176	Total	O	0	0
			176	176		
5	K	172	Total	O	0	0
			172	172		
5	L	177	Total	O	0	0
			177	177		
5	M	206	Total	O	0	0
			206	206		
5	N	195	Total	O	0	0
			195	195		
5	O	214	Total	O	0	0
			214	214		
5	P	184	Total	O	0	0
			184	184		
5	Q	157	Total	O	0	0
			157	157		
5	R	135	Total	O	0	0
			135	135		

*Continued on next page...*

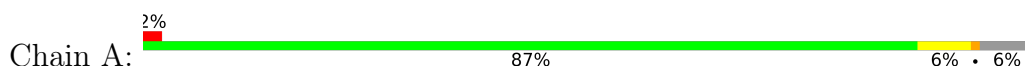
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	S	150	Total 150	O 150	0	0
5	T	205	Total 205	O 205	0	0
5	U	183	Total 183	O 183	0	0

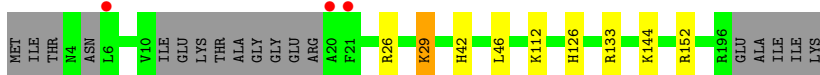
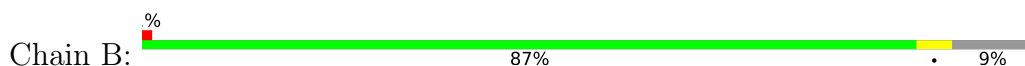
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

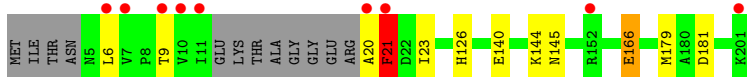
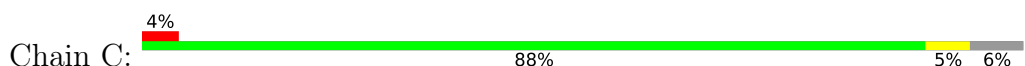
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



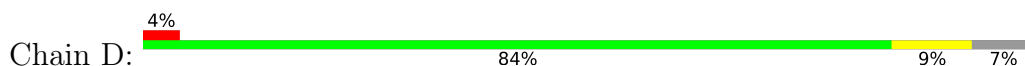
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



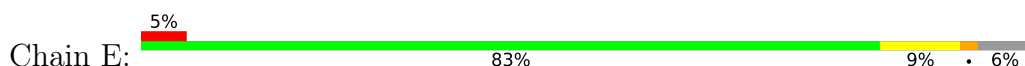
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



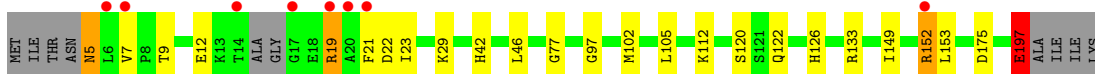
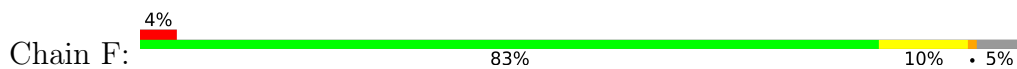
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



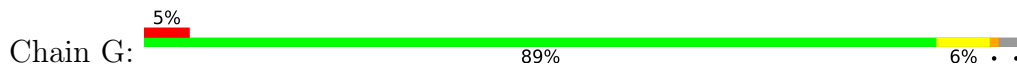
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



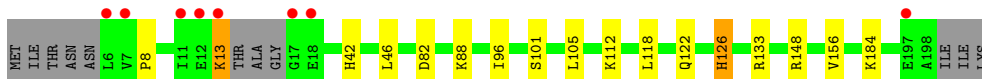
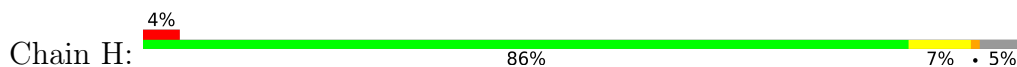
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



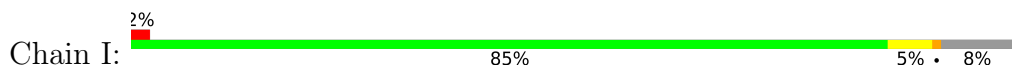
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



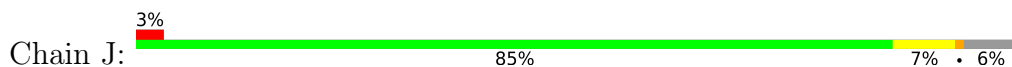
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



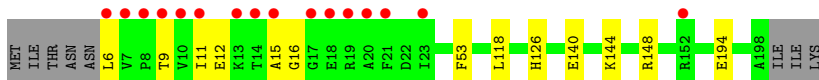
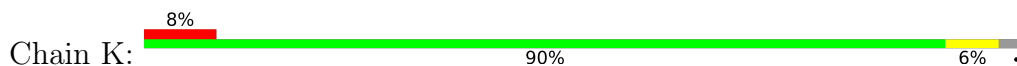
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



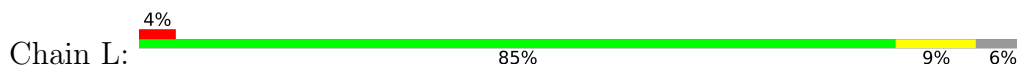
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



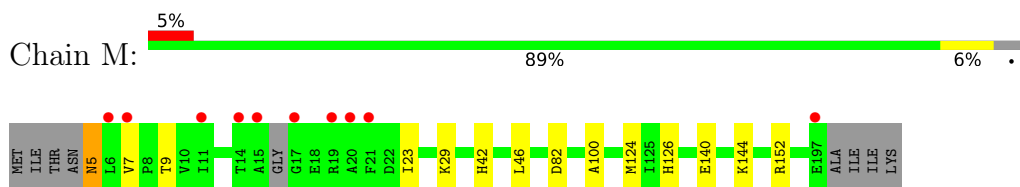
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



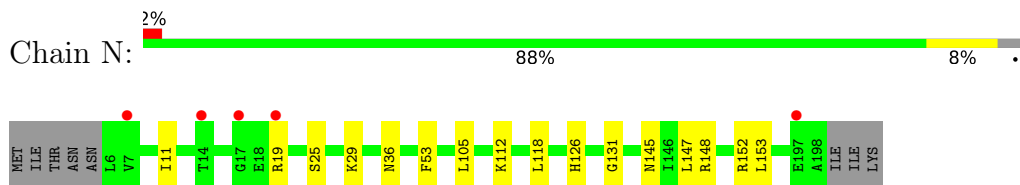
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



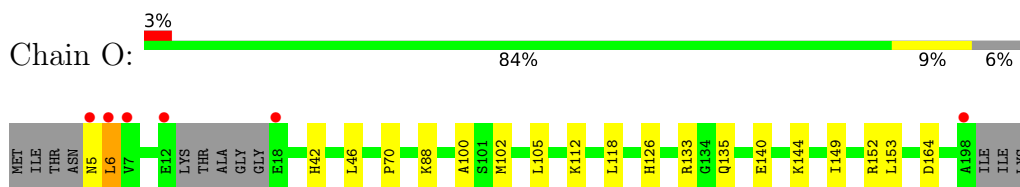
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



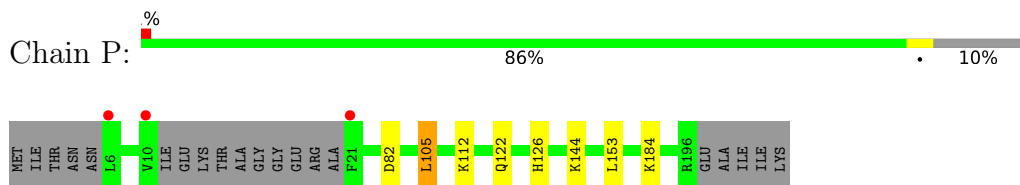
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



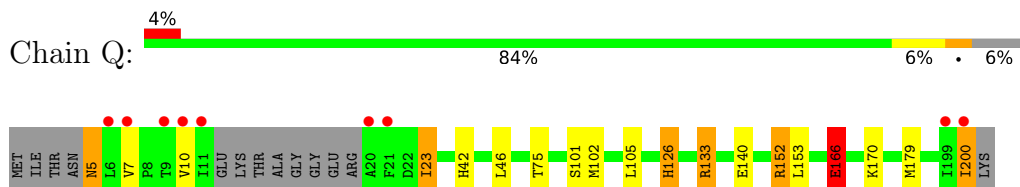
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



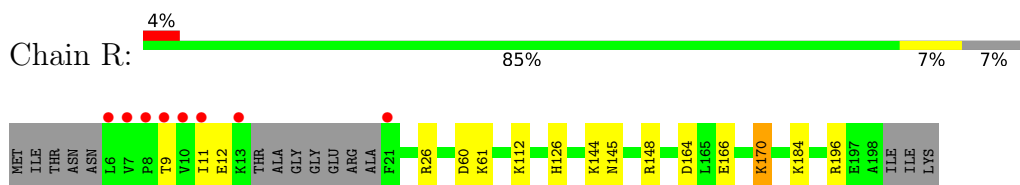
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



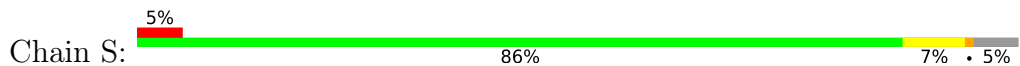
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



- Molecule 1: ATP-dependent Clp protease proteolytic subunit

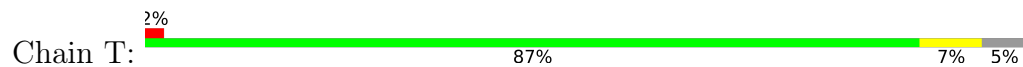


- Molecule 1: ATP-dependent Clp protease proteolytic subunit

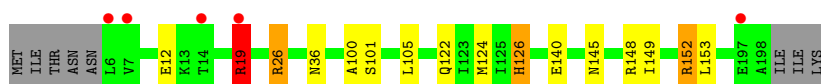
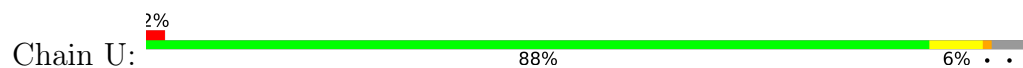




- Molecule 1: ATP-dependent Clp protease proteolytic subunit



- Molecule 1: ATP-dependent Clp protease proteolytic subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.23Å 127.92Å 353.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	176.92 – 1.70 48.80 – 1.70	Depositor EDS
% Data completeness (in resolution range)	94.7 (176.92-1.70) 94.7 (48.80-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.78 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.204 , 0.242 0.212 , 0.246	Depositor DCC
$R_{free}$ test set	22958 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.7	Xtrriage
Anisotropy	0.486	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 43.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	37499	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 83.80 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9744e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: MRD, MPD, ACT

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	0.53	2/1691 (0.1%)	0.74	2/2272 (0.1%)
1	B	0.91	3/1598 (0.2%)	0.70	1/2146 (0.0%)
1	C	0.64	1/1665 (0.1%)	0.66	1/2238 (0.0%)
1	D	0.59	3/1669 (0.2%)	0.65	1/2241 (0.0%)
1	E	0.77	3/1674 (0.2%)	0.66	1/2250 (0.0%)
1	F	0.49	1/1744 (0.1%)	0.80	3/2340 (0.1%)
1	G	0.44	0/1741	0.71	1/2338 (0.0%)
1	H	0.86	2/1714 (0.1%)	0.75	3/2299 (0.1%)
1	I	0.89	4/1654 (0.2%)	0.80	4/2222 (0.2%)
1	J	0.49	1/1667 (0.1%)	0.70	3/2241 (0.1%)
1	K	0.60	2/1693 (0.1%)	0.70	2/2276 (0.1%)
1	L	0.91	4/1630 (0.2%)	0.77	4/2190 (0.2%)
1	M	0.68	1/1720 (0.1%)	0.67	0/2309
1	N	0.60	2/1719 (0.1%)	0.72	2/2310 (0.1%)
1	O	0.55	1/1687 (0.1%)	0.71	0/2266
1	P	0.76	2/1594 (0.1%)	0.79	2/2143 (0.1%)
1	Q	1.48	4/1625 (0.2%)	1.42	9/2185 (0.4%)
1	R	0.60	3/1631 (0.2%)	0.71	3/2192 (0.1%)
1	S	0.52	1/1667 (0.1%)	0.65	2/2241 (0.1%)
1	T	0.46	1/1699 (0.1%)	0.71	1/2280 (0.0%)
1	U	0.43	0/1751	0.66	1/2353 (0.0%)
All	All	0.71	41/35233 (0.1%)	0.76	46/47332 (0.1%)

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	F	0	1



The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Q	200	ILE	CB-CG1	44.41	2.78	1.54
1	L	112	LYS	CD-CE	-27.58	0.82	1.51
1	Q	166	GLU	CG-CD	27.43	1.93	1.51
1	H	13	LYS	CB-CG	27.23	2.26	1.52
1	B	112	LYS	CE-NZ	25.77	2.13	1.49

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	Q	200	ILE	CA-CB-CG1	-47.50	20.75	111.00
1	Q	200	ILE	CG1-CB-CG2	-23.79	59.07	111.40
1	Q	200	ILE	CA-CB-CG2	-17.77	75.36	110.90
1	L	112	LYS	CG-CD-CE	16.74	162.13	111.90
1	P	184	LYS	CD-CE-NZ	15.38	147.07	111.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	197	GLU	Sidechain

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1586	0	1632	17	0
1	B	1510	0	1538	21	0
1	C	1577	0	1628	13	0
1	D	1568	0	1628	16	0
1	E	1575	0	1625	17	0
1	F	1631	0	1685	34	0
1	G	1628	0	1671	17	0
1	H	1603	0	1655	21	0
1	I	1552	0	1610	13	0
1	J	1570	0	1616	14	0
1	K	1598	0	1639	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	L	1545	0	1585	20	0
1	M	1613	0	1659	16	0
1	N	1608	0	1650	25	0
1	O	1593	0	1630	18	0
1	P	1501	0	1534	8	0
1	Q	1544	0	1588	29	0
1	R	1543	0	1571	17	0
1	S	1573	0	1625	16	0
1	T	1598	0	1635	13	0
1	U	1631	0	1662	22	0
2	A	8	0	14	0	0
2	C	16	0	28	0	0
2	D	8	0	14	0	0
2	G	8	0	14	0	0
2	I	8	0	14	0	0
2	K	24	0	42	1	0
2	L	8	0	14	3	0
2	N	8	0	14	0	0
2	O	8	0	14	0	0
2	P	8	0	14	0	0
2	Q	8	0	14	0	0
2	S	8	0	14	0	0
2	T	16	0	28	0	0
2	U	8	0	14	0	0
3	A	4	0	3	0	0
3	B	4	0	3	0	0
3	C	4	0	3	0	0
3	D	4	0	3	0	0
3	E	4	0	3	0	0
3	F	4	0	3	0	0
3	G	4	0	3	0	0
3	H	4	0	3	0	0
3	I	4	0	3	0	0
3	J	4	0	3	0	0
3	K	4	0	3	0	0
3	L	4	0	3	0	0
3	M	4	0	3	0	0
3	N	4	0	3	0	0
3	O	4	0	3	0	0
3	P	4	0	3	0	0
3	Q	4	0	3	0	0
3	R	4	0	3	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	S	4	0	3	0	0
3	T	4	0	3	0	0
3	U	4	0	3	0	0
4	B	8	0	14	0	0
4	C	8	0	14	0	0
4	E	8	0	14	0	0
4	F	16	0	28	0	0
4	G	8	0	14	0	0
4	I	8	0	14	0	0
4	J	8	0	14	0	0
4	L	8	0	14	0	0
4	M	8	0	14	0	0
4	N	8	0	14	1	0
4	S	8	0	14	1	0
4	T	8	0	14	0	0
4	U	8	0	14	0	0
5	A	256	0	0	2	0
5	B	218	0	0	0	0
5	C	209	0	0	1	0
5	D	168	0	0	0	1
5	E	185	0	0	2	0
5	F	214	0	0	3	0
5	G	189	0	0	2	0
5	H	252	0	0	3	0
5	I	167	0	0	0	0
5	J	176	0	0	2	0
5	K	172	0	0	0	0
5	L	177	0	0	4	0
5	M	206	0	0	0	0
5	N	195	0	0	0	0
5	O	214	0	0	1	0
5	P	184	0	0	0	0
5	Q	157	0	0	1	0
5	R	135	0	0	1	0
5	S	150	0	0	1	0
5	T	205	0	0	2	0
5	U	183	0	0	2	0
All	All	37499	0	34577	306	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 306 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:A:179[B]:MET:CE	1:B:152[B]:ARG:NH2	1.80	1.44
1:A:179[B]:MET:CE	1:B:152[B]:ARG:HH22	1.32	1.39
1:N:145[B]:ASN:ND2	1:N:148[B]:ARG:NH2	1.69	1.38
1:N:145[B]:ASN:ND2	1:N:148[B]:ARG:HH22	1.21	1.37
1:A:179[B]:MET:HE3	1:B:152[B]:ARG:NH2	1.29	1.36

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
5:D:2101:HOH:O	5:D:2101:HOH:O[2_565]	1.29	0.91

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	207/201 (103%)	204 (99%)	3 (1%)	0	100	100
1	B	195/201 (97%)	191 (98%)	4 (2%)	0	100	100
1	C	204/201 (102%)	198 (97%)	4 (2%)	2 (1%)	15	4
1	D	205/201 (102%)	201 (98%)	4 (2%)	0	100	100
1	E	206/201 (102%)	202 (98%)	4 (2%)	0	100	100
1	F	213/201 (106%)	210 (99%)	3 (1%)	0	100	100
1	G	216/201 (108%)	212 (98%)	4 (2%)	0	100	100
1	H	210/201 (104%)	206 (98%)	4 (2%)	0	100	100
1	I	202/201 (100%)	198 (98%)	4 (2%)	0	100	100
1	J	205/201 (102%)	201 (98%)	3 (2%)	1 (0%)	29	13
1	K	211/201 (105%)	207 (98%)	4 (2%)	0	100	100
1	L	199/201 (99%)	194 (98%)	5 (2%)	0	100	100
1	M	211/201 (105%)	207 (98%)	4 (2%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	214/201 (106%)	210 (98%)	4 (2%)	0	100	100
1	O	207/201 (103%)	203 (98%)	4 (2%)	0	100	100
1	P	196/201 (98%)	190 (97%)	6 (3%)	0	100	100
1	Q	200/201 (100%)	195 (98%)	5 (2%)	0	100	100
1	R	200/201 (100%)	196 (98%)	4 (2%)	0	100	100
1	S	206/201 (102%)	202 (98%)	4 (2%)	0	100	100
1	T	208/201 (104%)	204 (98%)	4 (2%)	0	100	100
1	U	217/201 (108%)	213 (98%)	4 (2%)	0	100	100
All	All	4332/4221 (103%)	4244 (98%)	85 (2%)	3 (0%)	51	33

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	21[A]	PHE
1	C	21[B]	PHE
1	J	97	GLY

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	181/168 (108%)	176 (97%)	5 (3%)	43	25
1	B	171/168 (102%)	170 (99%)	1 (1%)	86	80
1	C	178/168 (106%)	173 (97%)	5 (3%)	43	25
1	D	179/168 (106%)	175 (98%)	4 (2%)	52	34
1	E	180/168 (107%)	171 (95%)	9 (5%)	24	8
1	F	187/168 (111%)	177 (95%)	10 (5%)	22	7
1	G	185/168 (110%)	180 (97%)	5 (3%)	44	26
1	H	183/168 (109%)	177 (97%)	6 (3%)	38	19
1	I	177/168 (105%)	172 (97%)	5 (3%)	43	25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	180/168 (107%)	173 (96%)	7 (4%)	32	13
1	K	180/168 (107%)	179 (99%)	1 (1%)	86	80
1	L	175/168 (104%)	170 (97%)	5 (3%)	42	23
1	M	184/168 (110%)	182 (99%)	2 (1%)	73	63
1	N	183/168 (109%)	180 (98%)	3 (2%)	62	48
1	O	181/168 (108%)	174 (96%)	7 (4%)	32	13
1	P	172/168 (102%)	170 (99%)	2 (1%)	71	59
1	Q	174/168 (104%)	162 (93%)	12 (7%)	15	4
1	R	175/168 (104%)	168 (96%)	7 (4%)	31	13
1	S	178/168 (106%)	169 (95%)	9 (5%)	24	8
1	T	182/168 (108%)	175 (96%)	7 (4%)	33	14
1	U	186/168 (111%)	179 (96%)	7 (4%)	33	14
All	All	3771/3528 (107%)	3652 (97%)	119 (3%)	47	20

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	126	HIS
1	T	133[B]	ARG
1	O	112[A]	LYS
1	T	133[A]	ARG
1	U	152[C]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	Q	5	ASN
1	T	5	ASN
1	S	154	ASN
1	J	135	GLN
1	O	135	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

53 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	MRD	D	801	-	7,7,7	0.30	0	9,10,10	0.67	0
3	ACT	B	801	-	3,3,3	0.84	0	3,3,3	0.75	0
3	ACT	T	803	-	3,3,3	0.59	0	3,3,3	1.07	0
3	ACT	R	800	-	3,3,3	0.71	0	3,3,3	0.84	0
2	MRD	K	801	-	7,7,7	0.28	0	9,10,10	0.32	0
2	MRD	C	802	-	7,7,7	0.29	0	9,10,10	0.48	0
2	MRD	T	801	-	7,7,7	0.24	0	9,10,10	0.48	0
2	MRD	G	800	-	7,7,7	0.35	0	9,10,10	0.34	0
2	MRD	T	802	-	7,7,7	0.27	0	9,10,10	0.28	0
2	MRD	U	800	-	7,7,7	0.32	0	9,10,10	0.22	0
2	MRD	P	800	-	7,7,7	0.38	0	9,10,10	0.41	0
4	MPD	U	801	-	7,7,7	0.24	0	9,10,10	0.28	0
3	ACT	Q	801	-	3,3,3	0.83	0	3,3,3	0.66	0
3	ACT	U	802	-	3,3,3	0.98	0	3,3,3	0.47	0
3	ACT	J	801	-	3,3,3	0.70	0	3,3,3	0.76	0
4	MPD	N	800	-	7,7,7	0.28	0	9,10,10	0.58	0
3	ACT	O	801	-	3,3,3	0.89	0	3,3,3	0.55	0
3	ACT	D	800	-	3,3,3	0.83	0	3,3,3	0.47	0
3	ACT	H	800	-	3,3,3	0.77	0	3,3,3	0.79	0
2	MRD	I	801	-	7,7,7	0.31	0	9,10,10	0.47	0
3	ACT	E	801	-	3,3,3	0.86	0	3,3,3	0.70	0
2	MRD	K	802	-	7,7,7	0.29	0	9,10,10	0.27	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MPD	S	800	-	7,7,7	0.34	0	9,10,10	0.36	0
2	MRD	C	801	-	7,7,7	0.32	0	9,10,10	0.53	0
4	MPD	F	801	-	7,7,7	0.26	0	9,10,10	0.36	0
4	MPD	M	800	-	7,7,7	0.26	0	9,10,10	0.57	0
3	ACT	F	802	-	3,3,3	0.86	0	3,3,3	0.76	0
4	MPD	B	800	-	7,7,7	0.26	0	9,10,10	0.62	0
3	ACT	K	803	-	3,3,3	0.80	0	3,3,3	0.74	0
4	MPD	C	800	-	7,7,7	0.33	0	9,10,10	0.36	0
2	MRD	O	800	-	7,7,7	0.35	0	9,10,10	0.51	0
4	MPD	T	800	-	7,7,7	0.38	0	9,10,10	0.54	0
4	MPD	I	800	-	7,7,7	0.25	0	9,10,10	0.46	0
3	ACT	G	802	-	3,3,3	0.91	0	3,3,3	0.51	0
2	MRD	N	801	-	7,7,7	0.28	0	9,10,10	0.39	0
3	ACT	I	802	-	3,3,3	0.79	0	3,3,3	0.83	0
3	ACT	L	802	-	3,3,3	0.70	0	3,3,3	0.92	0
3	ACT	C	803	-	3,3,3	0.84	0	3,3,3	0.74	0
2	MRD	Q	800	-	7,7,7	0.35	0	9,10,10	0.43	0
4	MPD	G	801	-	7,7,7	0.32	0	9,10,10	0.34	0
4	MPD	L	801	-	7,7,7	0.27	0	9,10,10	0.46	0
3	ACT	P	801	-	3,3,3	0.74	0	3,3,3	0.70	0
3	ACT	A	801	-	3,3,3	0.73	0	3,3,3	0.97	0
4	MPD	E	800	-	7,7,7	0.28	0	9,10,10	0.27	0
2	MRD	K	800	-	7,7,7	0.33	0	9,10,10	0.58	0
3	ACT	N	802	-	3,3,3	0.83	0	3,3,3	0.53	0
4	MPD	J	800	-	7,7,7	0.27	0	9,10,10	0.36	0
3	ACT	S	802	-	3,3,3	0.75	0	3,3,3	0.78	0
4	MPD	F	800	-	7,7,7	0.30	0	9,10,10	0.37	0
2	MRD	A	800	-	7,7,7	0.29	0	9,10,10	0.25	0
3	ACT	M	801	-	3,3,3	0.90	0	3,3,3	0.61	0
2	MRD	L	800	-	7,7,7	0.33	0	9,10,10	0.33	0
2	MRD	S	801	-	7,7,7	0.25	0	9,10,10	0.35	0

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	MRD	D	801	-	-	3/5/5/5	-
2	MRD	K	801	-	-	1/5/5/5	-
2	MRD	C	802	-	-	0/5/5/5	-
2	MRD	T	801	-	-	0/5/5/5	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MRD	G	800	-	-	2/5/5/5	-
2	MRD	T	802	-	-	2/5/5/5	-
2	MRD	U	800	-	-	0/5/5/5	-
2	MRD	P	800	-	-	2/5/5/5	-
4	MPD	U	801	-	-	1/5/5/5	-
4	MPD	N	800	-	-	1/5/5/5	-
2	MRD	I	801	-	-	4/5/5/5	-
2	MRD	K	802	-	-	0/5/5/5	-
4	MPD	S	800	-	-	0/5/5/5	-
2	MRD	C	801	-	-	4/5/5/5	-
4	MPD	F	801	-	-	2/5/5/5	-
4	MPD	M	800	-	-	2/5/5/5	-
4	MPD	B	800	-	-	0/5/5/5	-
4	MPD	C	800	-	-	0/5/5/5	-
2	MRD	O	800	-	-	4/5/5/5	-
4	MPD	T	800	-	-	1/5/5/5	-
4	MPD	I	800	-	-	2/5/5/5	-
2	MRD	N	801	-	-	1/5/5/5	-
2	MRD	Q	800	-	-	0/5/5/5	-
4	MPD	G	801	-	-	0/5/5/5	-
4	MPD	L	801	-	-	0/5/5/5	-
4	MPD	E	800	-	-	2/5/5/5	-
2	MRD	K	800	-	-	1/5/5/5	-
4	MPD	J	800	-	-	2/5/5/5	-
4	MPD	F	800	-	-	2/5/5/5	-
2	MRD	A	800	-	-	0/5/5/5	-
2	MRD	L	800	-	-	2/5/5/5	-
2	MRD	S	801	-	-	0/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 41 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	801	MRD	C2-C3-C4-O4
2	T	802	MRD	C2-C3-C4-O4
4	E	800	MPD	C2-C3-C4-O4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	E	800	MPD	C2-C3-C4-C5
4	F	800	MPD	C2-C3-C4-O4

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	N	800	MPD	1	0
4	S	800	MPD	1	0
2	K	800	MRD	1	0
2	L	800	MRD	3	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	188/201 (93%)	-0.07	4 (2%) 63 67	7, 10, 25, 60	5 (2%)
1	B	183/201 (91%)	-0.12	3 (1%) 72 76	8, 12, 26, 38	6 (3%)
1	C	189/201 (94%)	0.07	9 (4%) 30 34	10, 14, 26, 56	8 (4%)
1	D	187/201 (93%)	0.11	8 (4%) 35 39	10, 14, 32, 58	8 (4%)
1	E	189/201 (94%)	-0.02	10 (5%) 26 29	9, 13, 32, 55	7 (3%)
1	F	191/201 (95%)	-0.08	8 (4%) 36 40	7, 10, 28, 50	5 (2%)
1	G	193/201 (96%)	0.13	10 (5%) 27 30	7, 10, 28, 44	8 (4%)
1	H	190/201 (94%)	-0.10	8 (4%) 36 40	6, 9, 28, 53	6 (3%)
1	I	184/201 (91%)	0.04	5 (2%) 54 58	6, 13, 30, 58	6 (3%)
1	J	188/201 (93%)	0.26	7 (3%) 41 46	11, 16, 28, 45	5 (2%)
1	K	193/201 (96%)	0.49	16 (8%) 11 13	11, 15, 38, 49	9 (4%)
1	L	188/201 (93%)	0.14	8 (4%) 35 39	11, 15, 30, 58	8 (4%)
1	M	192/201 (95%)	0.03	10 (5%) 27 30	8, 12, 28, 48	6 (3%)
1	N	193/201 (96%)	0.06	5 (2%) 56 60	7, 10, 27, 45	8 (4%)
1	O	189/201 (94%)	-0.08	6 (3%) 47 52	5, 9, 27, 58	6 (3%)
1	P	181/201 (90%)	-0.06	3 (1%) 70 74	8, 12, 23, 42	5 (2%)
1	Q	188/201 (93%)	0.21	9 (4%) 30 34	10, 16, 39, 72	7 (3%)
1	R	186/201 (92%)	0.31	8 (4%) 35 39	12, 17, 32, 49	8 (4%)
1	S	190/201 (94%)	0.05	10 (5%) 26 29	9, 14, 31, 66	8 (4%)
1	T	191/201 (95%)	-0.20	5 (2%) 56 60	6, 8, 24, 39	7 (3%)
1	U	193/201 (96%)	0.00	5 (2%) 56 60	5, 8, 26, 43	5 (2%)
All	All	3966/4221 (93%)	0.06	157 (3%) 38 42	5, 13, 31, 72	141 (3%)

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Q	11	ILE	8.9
1	K	17	GLY	8.7
1	I	20	ALA	8.4
1	S	200	ILE	7.6
1	K	14	THR	7.4

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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	B-factors(Å <sup>2</sup> )	Q<0.9
4	MPD	J	800	8/8	0.55	0.23	48,49,49,49	0
4	MPD	S	800	8/8	0.61	0.39	43,45,46,47	0
2	MRD	K	800	8/8	0.65	0.24	47,48,49,49	0
4	MPD	M	800	8/8	0.66	0.18	45,45,47,47	0
2	MRD	O	800	8/8	0.66	0.26	43,44,45,45	0
2	MRD	U	800	8/8	0.69	0.28	46,48,49,49	0
2	MRD	A	800	8/8	0.70	0.18	48,49,50,50	0
2	MRD	D	801	8/8	0.71	0.18	45,47,47,48	0
2	MRD	Q	800	8/8	0.71	0.22	41,43,44,44	0
3	ACT	E	801	4/4	0.72	0.24	22,24,25,26	0
4	MPD	T	800	8/8	0.72	0.33	49,51,55,56	0
2	MRD	T	801	8/8	0.73	0.18	39,40,41,41	0
4	MPD	G	801	8/8	0.73	0.19	36,37,38,38	0
2	MRD	L	800	8/8	0.73	0.34	42,44,45,46	0
4	MPD	L	801	8/8	0.74	0.26	37,39,40,40	0
2	MRD	C	802	8/8	0.74	0.19	45,48,50,51	0
2	MRD	T	802	8/8	0.75	0.22	42,45,48,49	0
4	MPD	N	800	8/8	0.75	0.20	40,41,42,43	0
4	MPD	F	801	8/8	0.76	0.17	42,45,47,48	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MPD	U	801	8/8	0.76	0.21	38,41,44,44	0
2	MRD	S	801	8/8	0.78	0.21	34,35,36,37	0
4	MPD	I	800	8/8	0.78	0.13	36,38,38,39	0
2	MRD	I	801	8/8	0.78	0.21	48,49,52,54	3
4	MPD	E	800	8/8	0.79	0.18	47,48,49,50	0
4	MPD	F	800	8/8	0.79	0.24	45,47,49,49	0
2	MRD	P	800	8/8	0.79	0.15	28,29,30,33	0
3	ACT	J	801	4/4	0.79	0.17	33,35,36,38	0
3	ACT	K	803	4/4	0.79	0.16	29,31,31,31	0
4	MPD	C	800	8/8	0.79	0.19	38,41,43,43	0
3	ACT	L	802	4/4	0.80	0.16	32,34,34,34	0
3	ACT	O	801	4/4	0.80	0.14	22,24,25,25	0
2	MRD	K	802	8/8	0.80	0.21	39,41,42,43	0
2	MRD	C	801	8/8	0.81	0.21	42,45,45,45	0
3	ACT	N	802	4/4	0.81	0.14	19,20,21,21	0
3	ACT	T	803	4/4	0.82	0.16	17,18,19,19	0
4	MPD	B	800	8/8	0.82	0.12	25,25,26,27	0
3	ACT	F	802	4/4	0.82	0.16	18,22,22,22	0
3	ACT	S	802	4/4	0.83	0.18	26,28,28,28	0
3	ACT	D	800	4/4	0.84	0.13	27,28,28,29	0
3	ACT	M	801	4/4	0.84	0.19	22,26,26,26	0
3	ACT	U	802	4/4	0.85	0.12	17,19,20,20	0
2	MRD	K	801	8/8	0.85	0.19	42,44,45,45	0
2	MRD	N	801	8/8	0.85	0.14	43,45,46,46	0
3	ACT	I	802	4/4	0.87	0.15	23,24,24,26	0
3	ACT	C	803	4/4	0.87	0.11	24,27,27,29	0
2	MRD	G	800	8/8	0.87	0.17	40,41,42,43	0
3	ACT	G	802	4/4	0.87	0.11	21,23,24,25	0
3	ACT	H	800	4/4	0.89	0.14	23,25,25,26	0
3	ACT	A	801	4/4	0.90	0.11	23,24,25,25	0
3	ACT	P	801	4/4	0.92	0.20	20,22,23,23	0
3	ACT	Q	801	4/4	0.93	0.15	25,27,27,27	0
3	ACT	R	800	4/4	0.94	0.09	25,26,27,28	0
3	ACT	B	801	4/4	0.94	0.16	21,24,24,25	0

## 6.5 Other polymers

There are no such residues in this entry.