



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:29 PM GMT

PDB ID : 4OSD  
Title : Dimer of a C-terminal fragment of phage T4 gp5 beta-helix  
Authors : Buth, S.A.; Leiman, P.G.; Shneider, M.M.  
Deposited on : 2014-02-12  
Resolution : 1.96 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

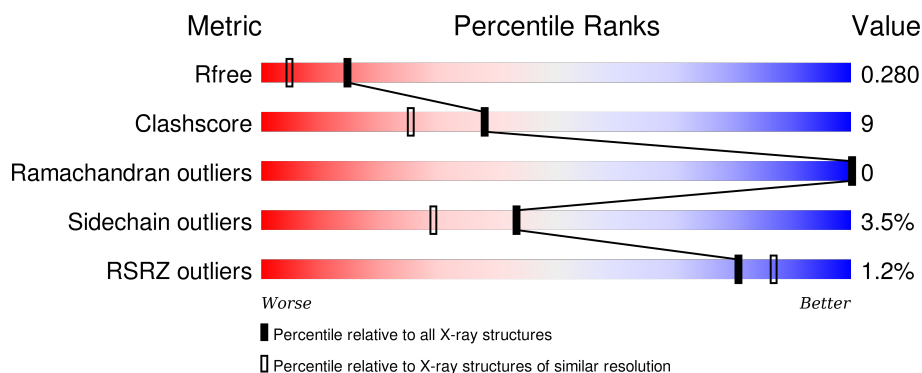
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.96 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	95	 80% 18% •
1	B	95	 78% 18% ••
1	C	95	 79% 19% •
1	D	95	 79% 19% •
1	E	95	 76% 21% ••

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	95	
1	G	95	
1	H	95	
1	I	95	
1	J	95	
1	K	95	
1	L	95	
1	M	95	
1	N	95	
1	O	95	
1	P	95	
1	Q	95	
1	R	95	

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
2	ELA	A	601	-	-	-	X
2	ELA	A	602	-	-	-	X
2	ELA	F	603	-	-	-	X
2	ELA	G	602	-	-	-	X
2	ELA	J	601	-	-	-	X
2	ELA	M	601	-	-	-	X
2	ELA	P	602	-	-	-	X
4	STE	C	601	-	-	-	X
4	STE	F	602	-	-	-	X
4	STE	G	601	-	-	-	X
4	STE	K	601	-	-	-	X
4	STE	O	602	-	-	-	X
4	STE	P	601	-	-	-	X
5	PLM	D	601	-	-	-	X
5	PLM	H	602	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PLM	K	602	-	-	-	X
5	PLM	N	601	-	-	-	X
5	PLM	Q	602	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13209 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 Tail-associated lysozyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	B	93	Total	C	N	O	S	0	1	0
			684	417	114	151	2			
1	C	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	D	93	Total	C	N	O	S	0	1	0
			684	417	114	151	2			
1	E	93	Total	C	N	O	S	0	1	0
			685	417	115	151	2			
1	F	93	Total	C	N	O	S	0	1	0
			682	414	114	152	2			
1	G	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	H	93	Total	C	N	O	S	0	1	0
			685	417	115	151	2			
1	I	92	Total	C	N	O	S	0	0	0
			670	408	112	148	2			
1	J	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	K	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	L	92	Total	C	N	O	S	0	1	0
			679	413	113	151	2			
1	M	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	N	93	Total	C	N	O	S	0	1	0
			682	414	114	152	2			
1	O	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	P	93	Total	C	N	O	S	0	1	0
			682	414	114	152	2			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			
1	R	93	Total	C	N	O	S	0	0	0
			676	411	113	150	2			

There are 54 discrepancies between the modelled and reference sequences:

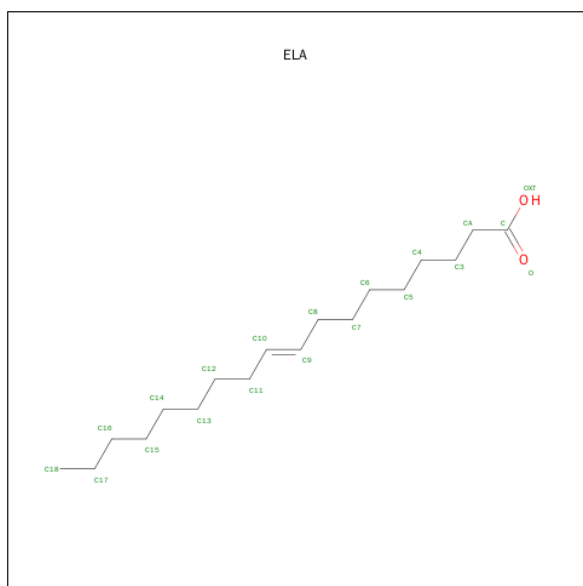
Chain	Residue	Modelled	Actual	Comment	Reference
A	481	SER	-	EXPRESSION TAG	UNP P16009
A	482	GLY	-	EXPRESSION TAG	UNP P16009
A	483	SER	-	EXPRESSION TAG	UNP P16009
B	481	SER	-	EXPRESSION TAG	UNP P16009
B	482	GLY	-	EXPRESSION TAG	UNP P16009
B	483	SER	-	EXPRESSION TAG	UNP P16009
C	481	SER	-	EXPRESSION TAG	UNP P16009
C	482	GLY	-	EXPRESSION TAG	UNP P16009
C	483	SER	-	EXPRESSION TAG	UNP P16009
D	481	SER	-	EXPRESSION TAG	UNP P16009
D	482	GLY	-	EXPRESSION TAG	UNP P16009
D	483	SER	-	EXPRESSION TAG	UNP P16009
E	481	SER	-	EXPRESSION TAG	UNP P16009
E	482	GLY	-	EXPRESSION TAG	UNP P16009
E	483	SER	-	EXPRESSION TAG	UNP P16009
F	481	SER	-	EXPRESSION TAG	UNP P16009
F	482	GLY	-	EXPRESSION TAG	UNP P16009
F	483	SER	-	EXPRESSION TAG	UNP P16009
G	481	SER	-	EXPRESSION TAG	UNP P16009
G	482	GLY	-	EXPRESSION TAG	UNP P16009
G	483	SER	-	EXPRESSION TAG	UNP P16009
H	481	SER	-	EXPRESSION TAG	UNP P16009
H	482	GLY	-	EXPRESSION TAG	UNP P16009
H	483	SER	-	EXPRESSION TAG	UNP P16009
I	481	SER	-	EXPRESSION TAG	UNP P16009
I	482	GLY	-	EXPRESSION TAG	UNP P16009
I	483	SER	-	EXPRESSION TAG	UNP P16009
J	481	SER	-	EXPRESSION TAG	UNP P16009
J	482	GLY	-	EXPRESSION TAG	UNP P16009
J	483	SER	-	EXPRESSION TAG	UNP P16009
K	481	SER	-	EXPRESSION TAG	UNP P16009
K	482	GLY	-	EXPRESSION TAG	UNP P16009
K	483	SER	-	EXPRESSION TAG	UNP P16009
L	481	SER	-	EXPRESSION TAG	UNP P16009

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	482	GLY	-	EXPRESSION TAG	UNP P16009
L	483	SER	-	EXPRESSION TAG	UNP P16009
M	481	SER	-	EXPRESSION TAG	UNP P16009
M	482	GLY	-	EXPRESSION TAG	UNP P16009
M	483	SER	-	EXPRESSION TAG	UNP P16009
N	481	SER	-	EXPRESSION TAG	UNP P16009
N	482	GLY	-	EXPRESSION TAG	UNP P16009
N	483	SER	-	EXPRESSION TAG	UNP P16009
O	481	SER	-	EXPRESSION TAG	UNP P16009
O	482	GLY	-	EXPRESSION TAG	UNP P16009
O	483	SER	-	EXPRESSION TAG	UNP P16009
P	481	SER	-	EXPRESSION TAG	UNP P16009
P	482	GLY	-	EXPRESSION TAG	UNP P16009
P	483	SER	-	EXPRESSION TAG	UNP P16009
Q	481	SER	-	EXPRESSION TAG	UNP P16009
Q	482	GLY	-	EXPRESSION TAG	UNP P16009
Q	483	SER	-	EXPRESSION TAG	UNP P16009
R	481	SER	-	EXPRESSION TAG	UNP P16009
R	482	GLY	-	EXPRESSION TAG	UNP P16009
R	483	SER	-	EXPRESSION TAG	UNP P16009

- Molecule 2 is 9-OCTADECENOIC ACID (three-letter code: ELA) (formula:  $C_{18}H_{34}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			20	18	2		

Continued on next page...

*Continued from previous page...*

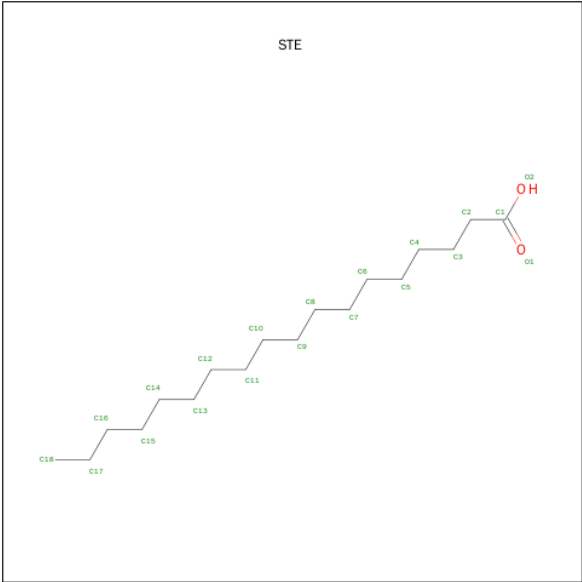
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			18	16	2		
2	F	1	Total	C	O	0	0
			20	18	2		
2	G	1	Total	C	O	0	0
			20	18	2		
2	J	1	Total	C	O	0	0
			20	18	2		
2	M	1	Total	C	O	0	0
			20	18	2		
2	P	1	Total	C	O	0	0
			20	18	2		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	Q	1	Total	Mg	0	0
			1	1		
3	H	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	O	1	Total	Mg	0	0
			1	1		
3	L	1	Total	Mg	0	0
			1	1		
3	F	1	Total	Mg	0	0
			1	1		

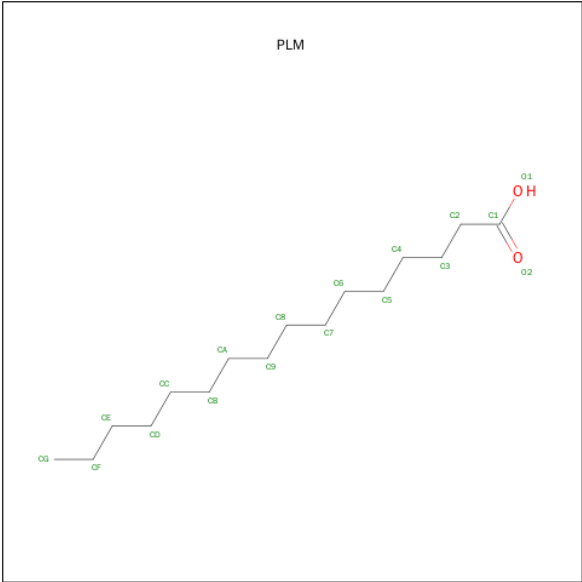
- Molecule 4 is STEARIC ACID (three-letter code: STE) (formula: C<sub>18</sub>H<sub>36</sub>O<sub>2</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			20	18	2		
4	F	1	Total	C	O	0	0
			20	18	2		
4	G	1	Total	C	O	0	0
			20	18	2		
4	K	1	Total	C	O	0	0
			20	18	2		
4	O	1	Total	C	O	0	0
			20	18	2		
4	P	1	Total	C	O	0	0
			20	18	2		

- Molecule 5 is PALMITIC ACID (three-letter code: PLM) (formula: C<sub>16</sub>H<sub>32</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	C	O	0	0
			18	16	2		
5	H	1	Total	C	O	0	0
			18	16	2		
5	K	1	Total	C	O	0	0
			18	16	2		
5	N	1	Total	C	O	0	0
			18	16	2		
5	Q	1	Total	C	O	0	0
			18	16	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	36	Total	O	0	0
			36	36		
6	B	38	Total	O	0	0
			38	38		
6	C	41	Total	O	0	0
			41	41		
6	D	38	Total	O	0	0
			38	38		
6	E	36	Total	O	0	0
			36	36		
6	F	42	Total	O	0	0
			42	42		
6	G	28	Total	O	0	0
			28	28		

Continued on next page...

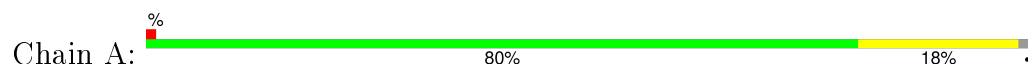
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	21	Total 21	O 21	0	0
6	I	20	Total 20	O 20	0	0
6	J	46	Total 46	O 46	0	0
6	K	32	Total 32	O 32	0	0
6	L	37	Total 37	O 37	0	0
6	M	21	Total 21	O 21	0	0
6	N	22	Total 22	O 22	0	0
6	O	34	Total 34	O 34	0	0
6	P	45	Total 45	O 45	0	0
6	Q	44	Total 44	O 44	0	0
6	R	57	Total 57	O 57	0	0

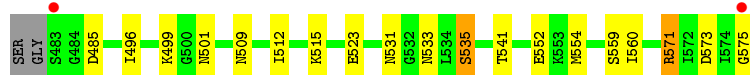
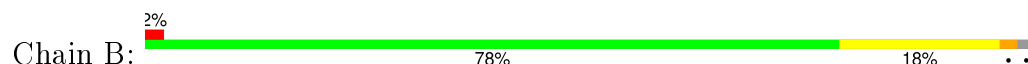
### 3 Residue-property plots [i](#)

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 ( $RSRZ > 2$ ). 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.

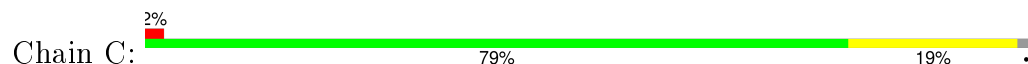
- Molecule 1: Tail-associated lysozyme



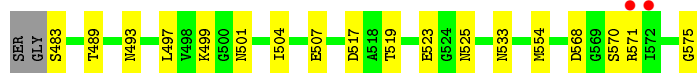
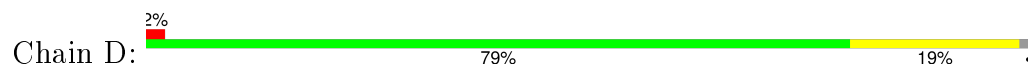
- Molecule 1: Tail-associated lysozyme



- Molecule 1: Tail-associated lysozyme



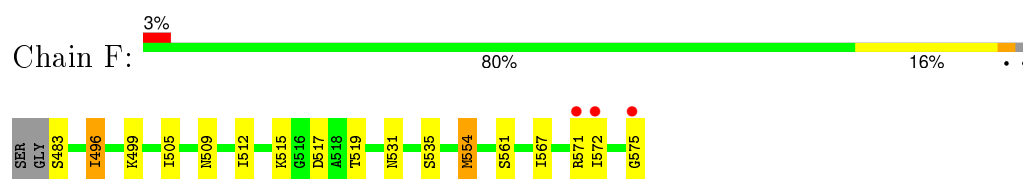
- Molecule 1: Tail-associated lysozyme



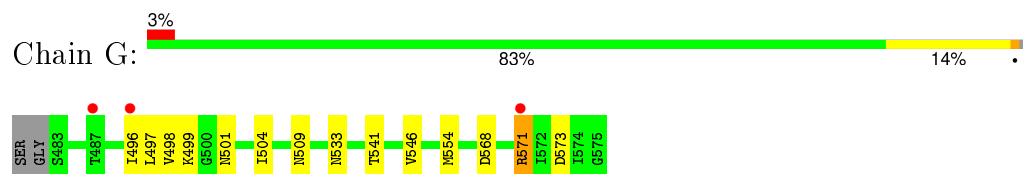
- Molecule 1: Tail-associated lysozyme



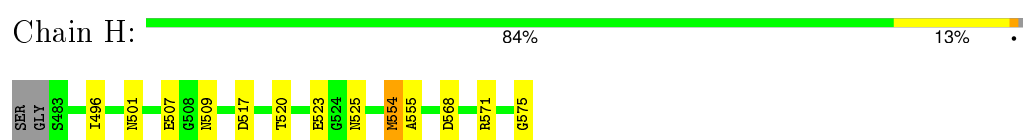
- Molecule 1: Tail-associated lysozyme



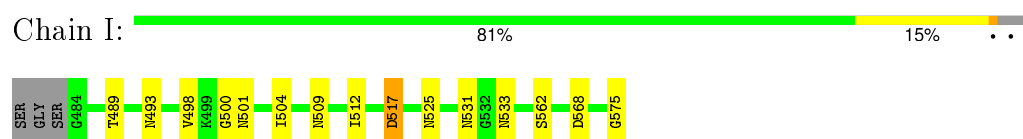
- Molecule 1: Tail-associated lysozyme



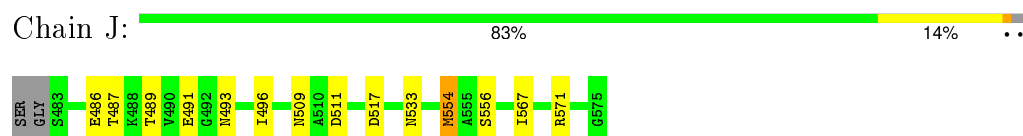
- Molecule 1: Tail-associated lysozyme



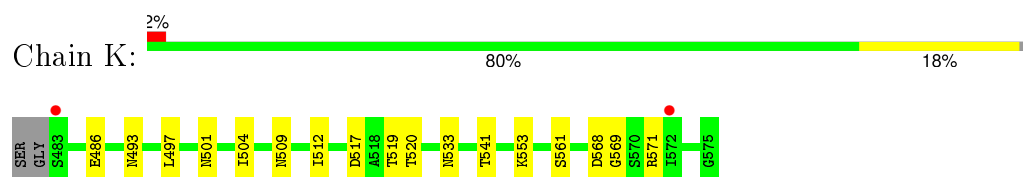
- Molecule 1: Tail-associated lysozyme



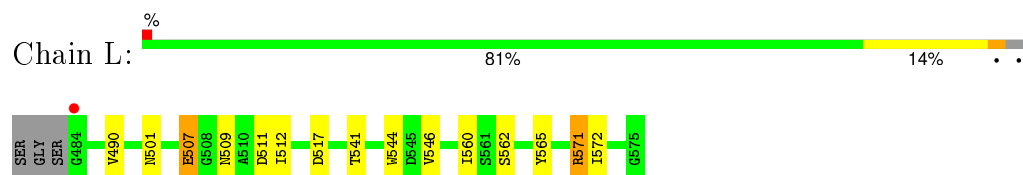
- Molecule 1: Tail-associated lysozyme



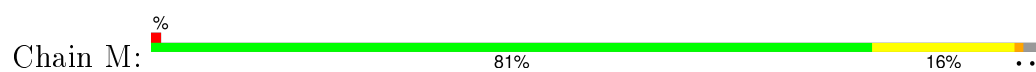
- Molecule 1: Tail-associated lysozyme



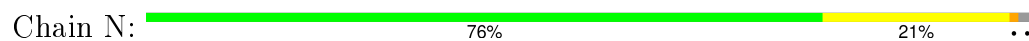
- Molecule 1: Tail-associated lysozyme



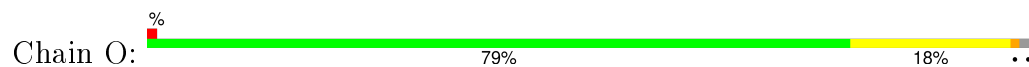
- Molecule 1: Tail-associated lysozyme



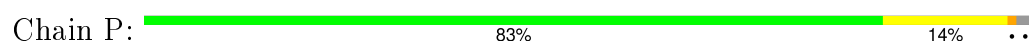
- Molecule 1: Tail-associated lysozyme



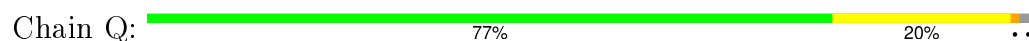
- Molecule 1: Tail-associated lysozyme



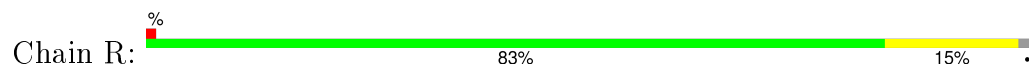
- Molecule 1: Tail-associated lysozyme



- Molecule 1: Tail-associated lysozyme



- Molecule 1: Tail-associated lysozyme



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.29 Å   73.79 Å   111.60 Å 90.00°   113.39°   90.00°	Depositor
Resolution (Å)	59.87 – 1.96 59.87 – 1.96	Depositor EDS
% Data completeness (in resolution range)	99.1 (59.87-1.96) 99.1 (59.87-1.96)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.02 (at 1.97 Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.219 , 0.279 0.225 , 0.280	Depositor DCC
$R_{free}$ test set	5818 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.9	Xtriage
Anisotropy	0.214	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 32.0	EDS
Estimated twinning fraction	0.022 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	33 of 116355 reflections (0.028%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13209	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.87 % 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 1.2412e-04. 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, 0.375 respectively for untwinned datasets, and 0.333, 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: STE, MG, ELA, PLM

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	0.98	0/682	0.97	1/925 (0.1%)
1	B	1.00	1/690 (0.1%)	0.87	0/936
1	C	1.05	0/682	0.99	0/925
1	D	0.99	0/690	0.98	1/936 (0.1%)
1	E	0.99	1/691 (0.1%)	0.89	0/936
1	F	1.03	1/688 (0.1%)	0.93	0/933
1	G	0.84	0/682	0.88	0/925
1	H	0.85	0/691	0.88	0/936
1	I	0.88	0/676	0.87	1/917 (0.1%)
1	J	1.00	0/682	0.98	0/925
1	K	1.05	0/682	0.98	1/925 (0.1%)
1	L	1.05	1/685 (0.1%)	0.97	1/929 (0.1%)
1	M	0.93	0/682	0.88	0/925
1	N	0.87	0/688	0.88	1/933 (0.1%)
1	O	0.89	0/682	0.90	0/925
1	P	1.04	0/688	1.01	2/933 (0.2%)
1	Q	1.09	0/682	0.97	0/925
1	R	1.09	0/682	0.98	0/925
All	All	0.98	4/12325 (0.0%)	0.94	8/16714 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	544	TRP	CB-CG	5.80	1.60	1.50
1	B	535	SER	CB-OG	-5.52	1.35	1.42
1	F	535	SER	CB-OG	-5.45	1.35	1.42
1	E	535	SER	CB-OG	-5.11	1.35	1.42

The worst 5 of 8 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	549	ASP	CB-CG-OD1	6.10	123.79	118.30
1	K	568	ASP	CB-CG-OD1	6.00	123.70	118.30
1	N	545	ASP	CB-CG-OD1	5.98	123.68	118.30
1	P	511	ASP	CB-CG-OD1	5.70	123.43	118.30
1	D	568	ASP	CB-CG-OD1	5.63	123.37	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-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 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	676	0	651	22	0
1	B	684	0	661	24	0
1	C	676	0	651	25	1
1	D	684	0	661	22	0
1	E	685	0	663	28	0
1	F	682	0	655	22	0
1	G	676	0	651	20	0
1	H	685	0	663	17	0
1	I	670	0	646	20	0
1	J	676	0	651	18	0
1	K	676	0	651	30	0
1	L	679	0	651	23	0
1	M	676	0	651	19	0
1	N	682	0	655	18	0
1	O	676	0	651	17	0
1	P	682	0	655	17	0
1	Q	676	0	651	28	0
1	R	676	0	651	20	0
2	A	38	0	59	7	0
2	F	20	0	33	0	0
2	G	20	0	33	1	0
2	J	20	0	33	1	0
2	M	20	0	33	0	0
2	P	20	0	33	2	0
3	B	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	F	1	0	0	0	0
3	H	1	0	0	0	0
3	L	1	0	0	0	0
3	O	1	0	0	0	0
3	Q	1	0	0	0	0
4	C	20	0	35	2	0
4	F	20	0	35	2	0
4	G	20	0	35	1	0
4	K	20	0	35	3	0
4	O	20	0	35	1	0
4	P	20	0	35	1	0
5	D	18	0	31	4	0
5	H	18	0	31	2	0
5	K	18	0	31	3	0
5	N	18	0	31	2	0
5	Q	18	0	31	5	0
6	A	36	0	0	3	0
6	B	38	0	0	1	0
6	C	41	0	0	0	0
6	D	38	0	0	1	0
6	E	36	0	0	3	0
6	F	42	0	0	2	0
6	G	28	0	0	0	0
6	H	21	0	0	0	0
6	I	20	0	0	0	0
6	J	46	0	0	1	0
6	K	32	0	0	1	0
6	L	37	0	0	2	0
6	M	21	0	0	0	0
6	N	22	0	0	0	0
6	O	34	0	0	2	1
6	P	45	0	0	4	0
6	Q	44	0	0	0	0
6	R	57	0	0	1	0
All	All	13209	0	12358	229	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 9.

The worst 5 of 229 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:J:511:ASP:OD1	1:K:519:THR:HB	1.64	0.96
1:K:512:ILE:HD11	5:K:602:PLM:H31	1.47	0.94
1:P:493:ASN:HD22	1:Q:501:ASN:H	1.22	0.85
1:G:499:LYS:HA	1:H:507:GLU:HG3	1.59	0.84
1:M:499:LYS:HA	1:N:507:GLU:HG3	1.60	0.83

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)
1:C:483:SER:N	6:O:706:HOH:O[2_556]	1.99	0.21

## 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	91/95 (96%)	91 (100%)	0	0	100	100
1	B	92/95 (97%)	92 (100%)	0	0	100	100
1	C	91/95 (96%)	91 (100%)	0	0	100	100
1	D	92/95 (97%)	92 (100%)	0	0	100	100
1	E	92/95 (97%)	92 (100%)	0	0	100	100
1	F	92/95 (97%)	92 (100%)	0	0	100	100
1	G	91/95 (96%)	91 (100%)	0	0	100	100
1	H	92/95 (97%)	90 (98%)	2 (2%)	0	100	100
1	I	90/95 (95%)	90 (100%)	0	0	100	100
1	J	91/95 (96%)	90 (99%)	1 (1%)	0	100	100
1	K	91/95 (96%)	90 (99%)	1 (1%)	0	100	100
1	L	91/95 (96%)	90 (99%)	1 (1%)	0	100	100
1	M	91/95 (96%)	91 (100%)	0	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	92/95 (97%)	92 (100%)	0	0	100	100
1	O	91/95 (96%)	91 (100%)	0	0	100	100
1	P	92/95 (97%)	92 (100%)	0	0	100	100
1	Q	91/95 (96%)	91 (100%)	0	0	100	100
1	R	91/95 (96%)	90 (99%)	1 (1%)	0	100	100
All	All	1644/1710 (96%)	1638 (100%)	6 (0%)	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	75/76 (99%)	73 (97%)	2 (3%)	52	41
1	B	76/76 (100%)	73 (96%)	3 (4%)	39	24
1	C	75/76 (99%)	72 (96%)	3 (4%)	38	23
1	D	76/76 (100%)	75 (99%)	1 (1%)	76	72
1	E	76/76 (100%)	74 (97%)	2 (3%)	54	43
1	F	76/76 (100%)	73 (96%)	3 (4%)	39	24
1	G	75/76 (99%)	71 (95%)	4 (5%)	28	13
1	H	76/76 (100%)	73 (96%)	3 (4%)	39	24
1	I	74/76 (97%)	74 (100%)	0	100	100
1	J	75/76 (99%)	73 (97%)	2 (3%)	52	41
1	K	75/76 (99%)	75 (100%)	0	100	100
1	L	75/76 (99%)	71 (95%)	4 (5%)	28	13
1	M	75/76 (99%)	71 (95%)	4 (5%)	28	13
1	N	76/76 (100%)	71 (93%)	5 (7%)	21	7
1	O	75/76 (99%)	71 (95%)	4 (5%)	28	13
1	P	76/76 (100%)	74 (97%)	2 (3%)	54	43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	Q	75/76 (99%)	72 (96%)	3 (4%)	38	23
1	R	75/76 (99%)	72 (96%)	3 (4%)	38	23
All	All	1356/1368 (99%)	1308 (96%)	48 (4%)	43	29

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	496	ILE
1	L	562	SER
1	Q	570	SER
1	J	554	MET
1	L	507[B]	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	525	ASN
1	K	501	ASN
1	Q	509	ASN
1	I	533	ASN
1	J	509	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 24 ligands modelled in this entry, 6 are monoatomic - leaving 18 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	ELA	A	601	-	16,19,19	1.27	1 (6%)	16,19,19	1.30	2 (12%)
2	ELA	A	602	-	14,17,19	1.32	1 (7%)	14,17,19	0.96	1 (7%)
4	STE	C	601	-	16,19,19	0.46	0	16,19,19	0.39	0
5	PLM	D	601	-	14,17,17	0.44	0	14,17,17	1.45	1 (7%)
4	STE	F	602	-	16,19,19	0.47	0	16,19,19	0.39	0
2	ELA	F	603	-	16,19,19	1.29	1 (6%)	16,19,19	1.30	3 (18%)
4	STE	G	601	-	16,19,19	0.57	0	16,19,19	0.32	0
2	ELA	G	602	-	16,19,19	1.07	1 (6%)	16,19,19	1.34	1 (6%)
5	PLM	H	602	-	14,17,17	0.49	0	14,17,17	1.17	0
2	ELA	J	601	-	16,19,19	1.13	1 (6%)	16,19,19	1.20	1 (6%)
4	STE	K	601	-	16,19,19	0.37	0	16,19,19	0.47	0
5	PLM	K	602	-	14,17,17	0.57	0	14,17,17	1.27	0
2	ELA	M	601	-	16,19,19	1.05	1 (6%)	16,19,19	1.20	2 (12%)
5	PLM	N	601	-	14,17,17	0.50	0	14,17,17	1.23	1 (7%)
4	STE	O	602	-	16,19,19	0.45	0	16,19,19	0.34	0
4	STE	P	601	-	16,19,19	0.53	0	16,19,19	0.35	0
2	ELA	P	602	-	16,19,19	1.24	1 (6%)	16,19,19	1.44	2 (12%)
5	PLM	Q	602	-	14,17,17	0.65	0	14,17,17	1.19	1 (7%)

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	ELA	A	601	-	-	0/15/17/17	0/0/0/0
2	ELA	A	602	-	-	0/13/15/17	0/0/0/0
4	STE	C	601	-	-	0/15/17/17	0/0/0/0
5	PLM	D	601	-	-	0/13/15/15	0/0/0/0
4	STE	F	602	-	-	0/15/17/17	0/0/0/0
2	ELA	F	603	-	-	0/15/17/17	0/0/0/0
4	STE	G	601	-	-	0/15/17/17	0/0/0/0
2	ELA	G	602	-	-	0/15/17/17	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PLM	H	602	-	-	0/13/15/15	0/0/0/0
2	ELA	J	601	-	-	0/15/17/17	0/0/0/0
4	STE	K	601	-	-	0/15/17/17	0/0/0/0
5	PLM	K	602	-	-	0/13/15/15	0/0/0/0
2	ELA	M	601	-	-	0/15/17/17	0/0/0/0
5	PLM	N	601	-	-	0/13/15/15	0/0/0/0
4	STE	O	602	-	-	0/15/17/17	0/0/0/0
4	STE	P	601	-	-	0/15/17/17	0/0/0/0
2	ELA	P	602	-	-	0/15/17/17	0/0/0/0
5	PLM	Q	602	-	-	0/13/15/15	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	601	ELA	C10-C9	3.86	1.54	1.31
2	G	602	ELA	C10-C9	3.90	1.54	1.31
2	M	601	ELA	C10-C9	3.90	1.54	1.31
2	P	602	ELA	C10-C9	4.13	1.55	1.31
2	A	602	ELA	C10-C9	4.29	1.56	1.31

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	601	ELA	C12-C11-C10	-2.58	98.89	112.45
2	P	602	ELA	C6-C7-C8	-2.46	104.28	113.86
2	F	603	ELA	C13-C12-C11	-2.41	104.46	113.86
5	Q	602	PLM	CA-C9-C8	-2.37	102.31	114.53
2	F	603	ELA	C4-C3-CA	-2.33	104.04	113.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	ELA	6	0
2	A	602	ELA	1	0
4	C	601	STE	2	0
5	D	601	PLM	4	0
4	F	602	STE	2	0
4	G	601	STE	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	602	ELA	1	0
5	H	602	PLM	2	0
2	J	601	ELA	1	0
4	K	601	STE	3	0
5	K	602	PLM	3	0
5	N	601	PLM	2	0
4	O	602	STE	1	0
4	P	601	STE	1	0
2	P	602	ELA	2	0
5	Q	602	PLM	5	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

### 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	93/95 (97%)	-0.15	1 (1%) 82 88	20, 30, 48, 76	0
1	B	93/95 (97%)	-0.14	2 (2%) 65 74	19, 30, 53, 75	0
1	C	93/95 (97%)	-0.12	2 (2%) 65 74	17, 30, 46, 71	1 (1%)
1	D	93/95 (97%)	-0.14	2 (2%) 65 74	22, 30, 49, 79	0
1	E	93/95 (97%)	-0.19	1 (1%) 82 88	21, 30, 49, 72	0
1	F	93/95 (97%)	-0.09	3 (3%) 51 61	21, 30, 45, 75	1 (1%)
1	G	93/95 (97%)	-0.03	3 (3%) 51 61	24, 37, 59, 69	0
1	H	93/95 (97%)	-0.11	0 100 100	25, 35, 57, 83	0
1	I	92/95 (96%)	-0.05	0 100 100	25, 36, 60, 69	1 (1%)
1	J	93/95 (97%)	-0.16	0 100 100	21, 30, 44, 51	0
1	K	93/95 (97%)	-0.14	2 (2%) 65 74	21, 30, 49, 75	0
1	L	92/95 (96%)	-0.12	1 (1%) 82 88	22, 30, 48, 61	1 (1%)
1	M	93/95 (97%)	-0.14	1 (1%) 82 88	24, 35, 59, 72	0
1	N	93/95 (97%)	-0.12	0 100 100	25, 36, 60, 90	0
1	O	93/95 (97%)	-0.06	1 (1%) 82 88	25, 35, 62, 76	1 (1%)
1	P	93/95 (97%)	-0.19	0 100 100	19, 30, 43, 51	0
1	Q	93/95 (97%)	-0.13	0 100 100	20, 29, 45, 69	0
1	R	93/95 (97%)	-0.15	1 (1%) 82 88	19, 30, 46, 58	1 (1%)
All	All	1672/1710 (97%)	-0.13	20 (1%) 81 87	17, 32, 54, 90	6 (0%)

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	483	SER	5.4
1	K	483	SER	3.4
1	E	575	GLY	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	575	GLY	2.7
1	D	571	ARG	2.7

## 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
2	ELA	P	602	20/20	0.80	0.30	12.26	76,80,83,85	0
4	STE	O	602	20/20	0.52	0.34	12.04	75,88,97,98	0
2	ELA	G	602	20/20	0.65	0.47	10.82	79,101,108,110	0
5	PLM	N	601	18/18	0.61	0.40	9.64	60,102,110,111	0
2	ELA	M	601	20/20	0.73	0.38	8.54	73,83,87,88	0
2	ELA	J	601	20/20	0.79	0.25	7.17	60,68,73,74	0
5	PLM	Q	602	18/18	0.83	0.21	6.77	59,69,83,86	0
4	STE	G	601	20/20	0.75	0.33	6.62	66,78,91,94	0
4	STE	C	601	20/20	0.75	0.28	6.56	52,58,81,81	0
2	ELA	F	603	20/20	0.85	0.25	5.83	62,70,80,83	0
4	STE	F	602	20/20	0.81	0.24	5.66	54,62,73,80	0
2	ELA	A	601	20/20	0.80	0.26	5.43	53,67,89,98	0
4	STE	K	601	20/20	0.81	0.21	5.27	54,63,69,72	0
5	PLM	D	601	18/18	0.83	0.19	4.43	51,62,71,73	0
5	PLM	K	602	18/18	0.78	0.20	4.28	62,68,76,77	0
4	STE	P	601	20/20	0.79	0.17	3.64	63,71,77,78	0
2	ELA	A	602	18/20	0.83	0.22	3.49	56,63,76,76	0
5	PLM	H	602	18/18	0.75	0.20	2.62	72,85,90,90	0
3	MG	F	601	1/1	0.95	0.05	-2.54	40,40,40,40	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	H	601	1/1	0.90	0.05	-2.65	46,46,46,46	0
3	MG	B	601	1/1	0.98	0.03	-4.15	41,41,41,41	0
3	MG	Q	601	1/1	0.99	0.04	-5.16	43,43,43,43	0
3	MG	O	601	1/1	0.98	0.03	-5.73	40,40,40,40	0
3	MG	L	601	1/1	0.92	0.03	-8.40	35,35,35,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.