



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2024 – 12:16 PM JST

PDB ID : 5ZJK  
Title : Structure of myroilysin  
Authors : Li, W.D.; Ran, T.T.; Xu, D.Q.; Wang, W.W.  
Deposited on : 2018-03-20  
Resolution : 2.60 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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

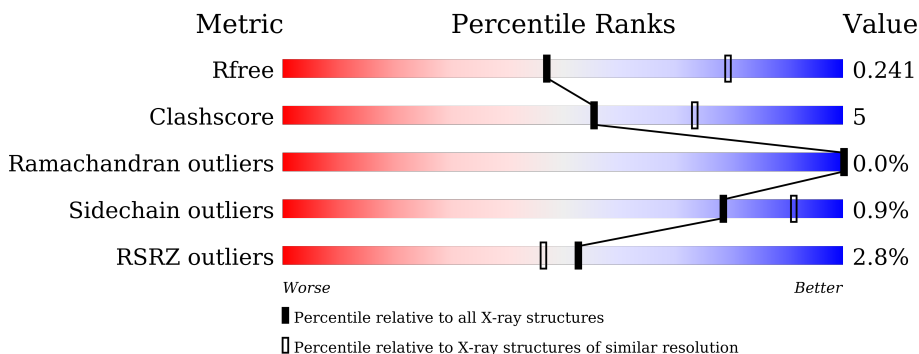
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.






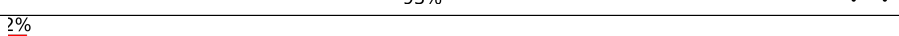
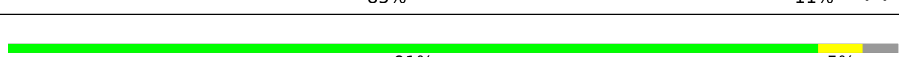



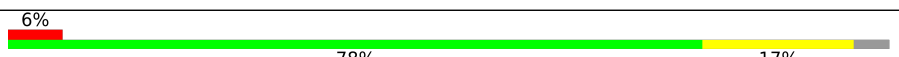


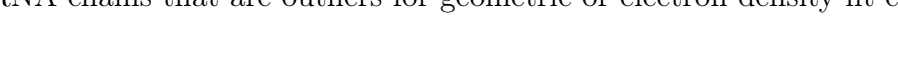
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	213	91% 5% .
1	B	213	92% 5% .
1	C	213	91% 5% .
1	D	213	91% 5% .
1	E	213	% 90% 7% .
1	F	213	% 91% 5% .

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	213	
1	H	213	
1	I	213	
1	J	213	
1	K	213	
1	L	213	
1	M	213	
1	N	213	
1	O	213	
1	P	213	
1	Q	213	
1	R	213	

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
3	PO4	E	301	-	X	-	-
3	PO4	M	301	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 30072 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 Myroilysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	B	206	Total 1637	C 1042	N 271	O 321	S 3	0	0	0
1	C	205	Total 1629	C 1036	N 270	O 320	S 3	0	0	0
1	D	205	Total 1629	C 1036	N 270	O 320	S 3	0	0	0
1	E	205	Total 1629	C 1036	N 270	O 320	S 3	0	0	0
1	F	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	G	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	H	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	I	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	J	207	Total 1646	C 1047	N 272	O 324	S 3	0	0	0
1	K	206	Total 1637	C 1042	N 271	O 321	S 3	0	0	0
1	L	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	M	205	Total 1629	C 1036	N 270	O 320	S 3	0	0	0
1	N	204	Total 1601	C 1019	N 262	O 317	S 3	0	0	0
1	O	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0
1	P	204	Total 1621	C 1032	N 268	O 318	S 3	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	Q	204	1621	1032	268	318	3	0	0	0
1	R	204	1621	1032	268	318	3	0	0	0

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	243	LEU	-	expression tag	UNP A0A0P0DZ84
A	244	GLU	-	expression tag	UNP A0A0P0DZ84
A	245	HIS	-	expression tag	UNP A0A0P0DZ84
A	246	HIS	-	expression tag	UNP A0A0P0DZ84
A	247	HIS	-	expression tag	UNP A0A0P0DZ84
A	248	HIS	-	expression tag	UNP A0A0P0DZ84
A	249	HIS	-	expression tag	UNP A0A0P0DZ84
A	250	HIS	-	expression tag	UNP A0A0P0DZ84
B	243	LEU	-	expression tag	UNP A0A0P0DZ84
B	244	GLU	-	expression tag	UNP A0A0P0DZ84
B	245	HIS	-	expression tag	UNP A0A0P0DZ84
B	246	HIS	-	expression tag	UNP A0A0P0DZ84
B	247	HIS	-	expression tag	UNP A0A0P0DZ84
B	248	HIS	-	expression tag	UNP A0A0P0DZ84
B	249	HIS	-	expression tag	UNP A0A0P0DZ84
B	250	HIS	-	expression tag	UNP A0A0P0DZ84
C	243	LEU	-	expression tag	UNP A0A0P0DZ84
C	244	GLU	-	expression tag	UNP A0A0P0DZ84
C	245	HIS	-	expression tag	UNP A0A0P0DZ84
C	246	HIS	-	expression tag	UNP A0A0P0DZ84
C	247	HIS	-	expression tag	UNP A0A0P0DZ84
C	248	HIS	-	expression tag	UNP A0A0P0DZ84
C	249	HIS	-	expression tag	UNP A0A0P0DZ84
C	250	HIS	-	expression tag	UNP A0A0P0DZ84
D	243	LEU	-	expression tag	UNP A0A0P0DZ84
D	244	GLU	-	expression tag	UNP A0A0P0DZ84
D	245	HIS	-	expression tag	UNP A0A0P0DZ84
D	246	HIS	-	expression tag	UNP A0A0P0DZ84
D	247	HIS	-	expression tag	UNP A0A0P0DZ84
D	248	HIS	-	expression tag	UNP A0A0P0DZ84
D	249	HIS	-	expression tag	UNP A0A0P0DZ84
D	250	HIS	-	expression tag	UNP A0A0P0DZ84
E	243	LEU	-	expression tag	UNP A0A0P0DZ84
E	244	GLU	-	expression tag	UNP A0A0P0DZ84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	245	HIS	-	expression tag	UNP A0A0P0DZ84
E	246	HIS	-	expression tag	UNP A0A0P0DZ84
E	247	HIS	-	expression tag	UNP A0A0P0DZ84
E	248	HIS	-	expression tag	UNP A0A0P0DZ84
E	249	HIS	-	expression tag	UNP A0A0P0DZ84
E	250	HIS	-	expression tag	UNP A0A0P0DZ84
F	243	LEU	-	expression tag	UNP A0A0P0DZ84
F	244	GLU	-	expression tag	UNP A0A0P0DZ84
F	245	HIS	-	expression tag	UNP A0A0P0DZ84
F	246	HIS	-	expression tag	UNP A0A0P0DZ84
F	247	HIS	-	expression tag	UNP A0A0P0DZ84
F	248	HIS	-	expression tag	UNP A0A0P0DZ84
F	249	HIS	-	expression tag	UNP A0A0P0DZ84
F	250	HIS	-	expression tag	UNP A0A0P0DZ84
G	243	LEU	-	expression tag	UNP A0A0P0DZ84
G	244	GLU	-	expression tag	UNP A0A0P0DZ84
G	245	HIS	-	expression tag	UNP A0A0P0DZ84
G	246	HIS	-	expression tag	UNP A0A0P0DZ84
G	247	HIS	-	expression tag	UNP A0A0P0DZ84
G	248	HIS	-	expression tag	UNP A0A0P0DZ84
G	249	HIS	-	expression tag	UNP A0A0P0DZ84
G	250	HIS	-	expression tag	UNP A0A0P0DZ84
H	243	LEU	-	expression tag	UNP A0A0P0DZ84
H	244	GLU	-	expression tag	UNP A0A0P0DZ84
H	245	HIS	-	expression tag	UNP A0A0P0DZ84
H	246	HIS	-	expression tag	UNP A0A0P0DZ84
H	247	HIS	-	expression tag	UNP A0A0P0DZ84
H	248	HIS	-	expression tag	UNP A0A0P0DZ84
H	249	HIS	-	expression tag	UNP A0A0P0DZ84
H	250	HIS	-	expression tag	UNP A0A0P0DZ84
I	243	LEU	-	expression tag	UNP A0A0P0DZ84
I	244	GLU	-	expression tag	UNP A0A0P0DZ84
I	245	HIS	-	expression tag	UNP A0A0P0DZ84
I	246	HIS	-	expression tag	UNP A0A0P0DZ84
I	247	HIS	-	expression tag	UNP A0A0P0DZ84
I	248	HIS	-	expression tag	UNP A0A0P0DZ84
I	249	HIS	-	expression tag	UNP A0A0P0DZ84
I	250	HIS	-	expression tag	UNP A0A0P0DZ84
J	243	LEU	-	expression tag	UNP A0A0P0DZ84
J	244	GLU	-	expression tag	UNP A0A0P0DZ84
J	245	HIS	-	expression tag	UNP A0A0P0DZ84
J	246	HIS	-	expression tag	UNP A0A0P0DZ84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
J	247	HIS	-	expression tag	UNP A0A0P0DZ84
J	248	HIS	-	expression tag	UNP A0A0P0DZ84
J	249	HIS	-	expression tag	UNP A0A0P0DZ84
J	250	HIS	-	expression tag	UNP A0A0P0DZ84
K	243	LEU	-	expression tag	UNP A0A0P0DZ84
K	244	GLU	-	expression tag	UNP A0A0P0DZ84
K	245	HIS	-	expression tag	UNP A0A0P0DZ84
K	246	HIS	-	expression tag	UNP A0A0P0DZ84
K	247	HIS	-	expression tag	UNP A0A0P0DZ84
K	248	HIS	-	expression tag	UNP A0A0P0DZ84
K	249	HIS	-	expression tag	UNP A0A0P0DZ84
K	250	HIS	-	expression tag	UNP A0A0P0DZ84
L	243	LEU	-	expression tag	UNP A0A0P0DZ84
L	244	GLU	-	expression tag	UNP A0A0P0DZ84
L	245	HIS	-	expression tag	UNP A0A0P0DZ84
L	246	HIS	-	expression tag	UNP A0A0P0DZ84
L	247	HIS	-	expression tag	UNP A0A0P0DZ84
L	248	HIS	-	expression tag	UNP A0A0P0DZ84
L	249	HIS	-	expression tag	UNP A0A0P0DZ84
L	250	HIS	-	expression tag	UNP A0A0P0DZ84
M	243	LEU	-	expression tag	UNP A0A0P0DZ84
M	244	GLU	-	expression tag	UNP A0A0P0DZ84
M	245	HIS	-	expression tag	UNP A0A0P0DZ84
M	246	HIS	-	expression tag	UNP A0A0P0DZ84
M	247	HIS	-	expression tag	UNP A0A0P0DZ84
M	248	HIS	-	expression tag	UNP A0A0P0DZ84
M	249	HIS	-	expression tag	UNP A0A0P0DZ84
M	250	HIS	-	expression tag	UNP A0A0P0DZ84
N	243	LEU	-	expression tag	UNP A0A0P0DZ84
N	244	GLU	-	expression tag	UNP A0A0P0DZ84
N	245	HIS	-	expression tag	UNP A0A0P0DZ84
N	246	HIS	-	expression tag	UNP A0A0P0DZ84
N	247	HIS	-	expression tag	UNP A0A0P0DZ84
N	248	HIS	-	expression tag	UNP A0A0P0DZ84
N	249	HIS	-	expression tag	UNP A0A0P0DZ84
N	250	HIS	-	expression tag	UNP A0A0P0DZ84
O	243	LEU	-	expression tag	UNP A0A0P0DZ84
O	244	GLU	-	expression tag	UNP A0A0P0DZ84
O	245	HIS	-	expression tag	UNP A0A0P0DZ84
O	246	HIS	-	expression tag	UNP A0A0P0DZ84
O	247	HIS	-	expression tag	UNP A0A0P0DZ84
O	248	HIS	-	expression tag	UNP A0A0P0DZ84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
O	249	HIS	-	expression tag	UNP A0A0P0DZ84
O	250	HIS	-	expression tag	UNP A0A0P0DZ84
P	243	LEU	-	expression tag	UNP A0A0P0DZ84
P	244	GLU	-	expression tag	UNP A0A0P0DZ84
P	245	HIS	-	expression tag	UNP A0A0P0DZ84
P	246	HIS	-	expression tag	UNP A0A0P0DZ84
P	247	HIS	-	expression tag	UNP A0A0P0DZ84
P	248	HIS	-	expression tag	UNP A0A0P0DZ84
P	249	HIS	-	expression tag	UNP A0A0P0DZ84
P	250	HIS	-	expression tag	UNP A0A0P0DZ84
Q	243	LEU	-	expression tag	UNP A0A0P0DZ84
Q	244	GLU	-	expression tag	UNP A0A0P0DZ84
Q	245	HIS	-	expression tag	UNP A0A0P0DZ84
Q	246	HIS	-	expression tag	UNP A0A0P0DZ84
Q	247	HIS	-	expression tag	UNP A0A0P0DZ84
Q	248	HIS	-	expression tag	UNP A0A0P0DZ84
Q	249	HIS	-	expression tag	UNP A0A0P0DZ84
Q	250	HIS	-	expression tag	UNP A0A0P0DZ84
R	243	LEU	-	expression tag	UNP A0A0P0DZ84
R	244	GLU	-	expression tag	UNP A0A0P0DZ84
R	245	HIS	-	expression tag	UNP A0A0P0DZ84
R	246	HIS	-	expression tag	UNP A0A0P0DZ84
R	247	HIS	-	expression tag	UNP A0A0P0DZ84
R	248	HIS	-	expression tag	UNP A0A0P0DZ84
R	249	HIS	-	expression tag	UNP A0A0P0DZ84
R	250	HIS	-	expression tag	UNP A0A0P0DZ84

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0

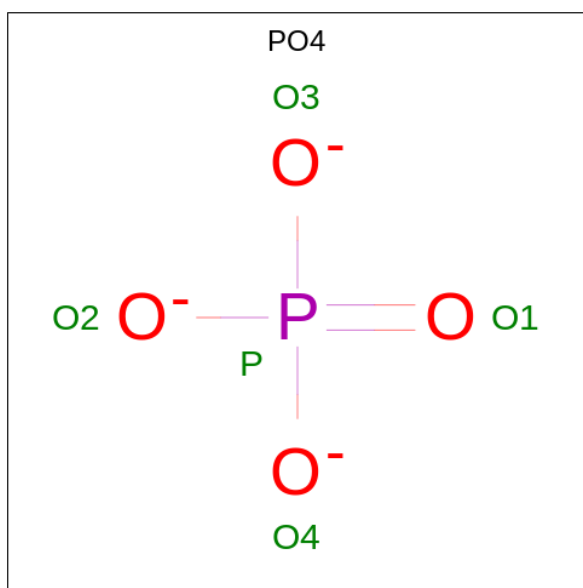
*Continued on next page...*



Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Zn 1 1	0	0
2	H	1	Total Zn 1 1	0	0
2	I	1	Total Zn 1 1	0	0
2	J	1	Total Zn 1 1	0	0
2	K	1	Total Zn 1 1	0	0
2	L	1	Total Zn 1 1	0	0
2	M	1	Total Zn 1 1	0	0
2	N	1	Total Zn 1 1	0	0
2	O	1	Total Zn 1 1	0	0
2	P	1	Total Zn 1 1	0	0
2	Q	1	Total Zn 1 1	0	0
2	R	1	Total Zn 1 1	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		
3	E	1	Total	O	P	0	0
			5	4	1		
3	F	1	Total	O	P	0	0
			5	4	1		
3	G	1	Total	O	P	0	0
			5	4	1		
3	H	1	Total	O	P	0	0
			5	4	1		
3	I	1	Total	O	P	0	0
			5	4	1		
3	J	1	Total	O	P	0	0
			5	4	1		
3	K	1	Total	O	P	0	0
			5	4	1		
3	M	1	Total	O	P	0	0
			5	4	1		
3	P	1	Total	O	P	0	0
			5	4	1		
3	Q	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	75	Total	O	0	0
			75	75		
4	B	61	Total	O	0	0
			61	61		
4	C	74	Total	O	0	0
			74	74		
4	D	80	Total	O	0	0
			80	80		
4	E	54	Total	O	0	0
			54	54		
4	F	50	Total	O	0	0
			50	50		
4	G	33	Total	O	0	0
			33	33		
4	H	39	Total	O	0	0
			39	39		

*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>
4	I	32	Total 32	O 32	0	0
4	J	81	Total 81	O 81	0	0
4	K	18	Total 18	O 18	0	0
4	L	52	Total 52	O 52	0	0
4	M	5	Total 5	O 5	0	0
4	N	2	Total 2	O 2	0	0
4	O	28	Total 28	O 28	0	0
4	P	29	Total 29	O 29	0	0
4	Q	32	Total 32	O 32	0	0
4	R	2	Total 2	O 2	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.

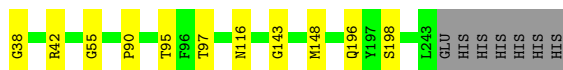
- Molecule 1: Myroilysin

Chain A:  91% 5% .



- Molecule 1: Myroilysin

Chain B:  92% 5% .



- Molecule 1: Myroilysin

Chain C:  91% 5% .




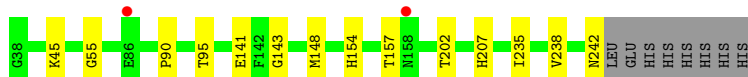
- Molecule 1: Myroilysin

Chain D:  91% 5% .



- Molecule 1: Myroilysin

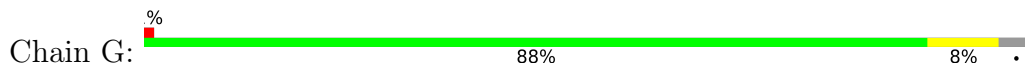
Chain E:  90% 7% .



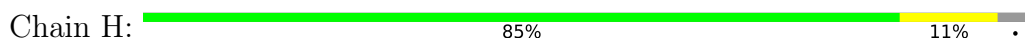
- Molecule 1: Myroilysin



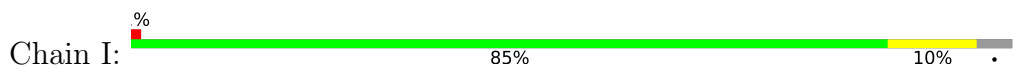
● Molecule 1: Myroilysin



● Molecule 1: Myroilysin



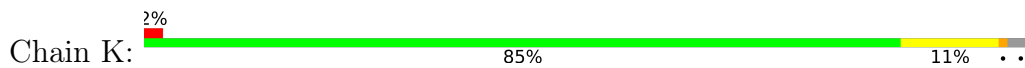
● Molecule 1: Myroilysin



● Molecule 1: Myroilysin



● Molecule 1: Myroilysin



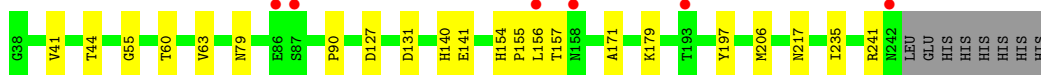
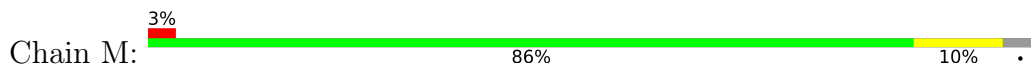
HIS

● Molecule 1: Myroilysin

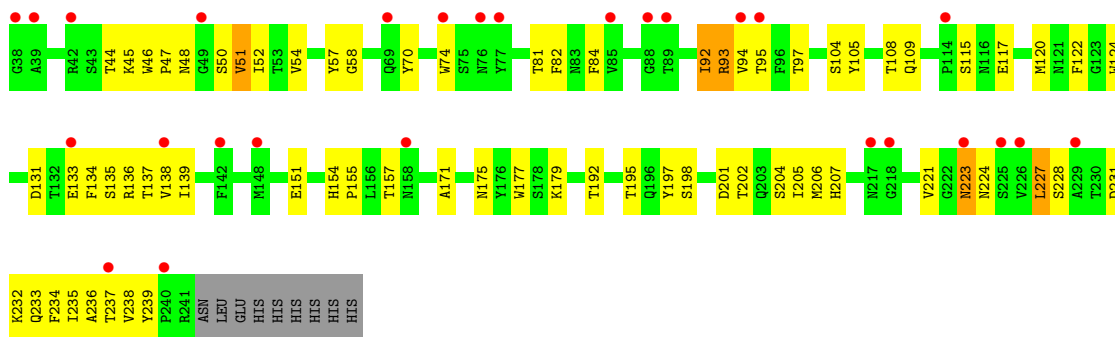




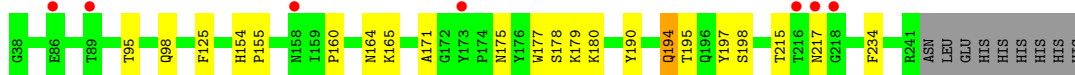
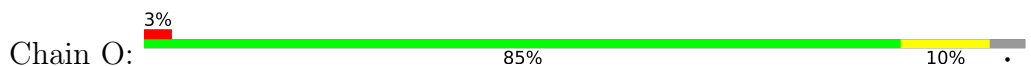
- Molecule 1: Myroilysin



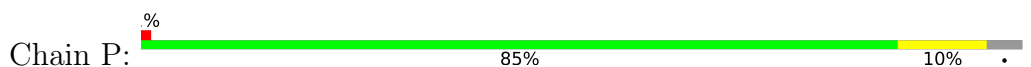
- Molecule 1: Myroilysin



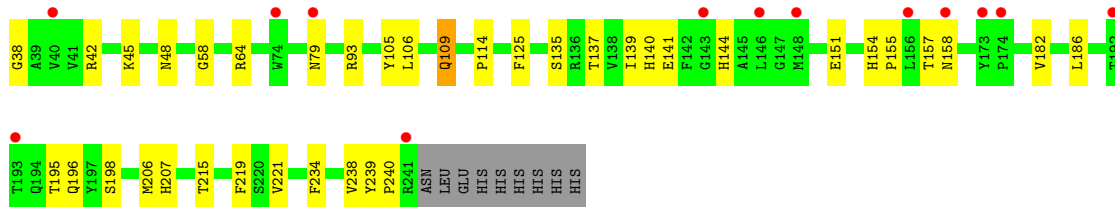
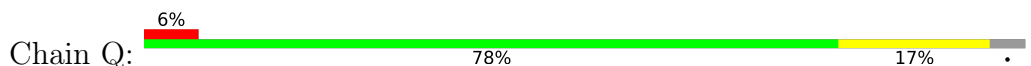
- Molecule 1: Myroilysin



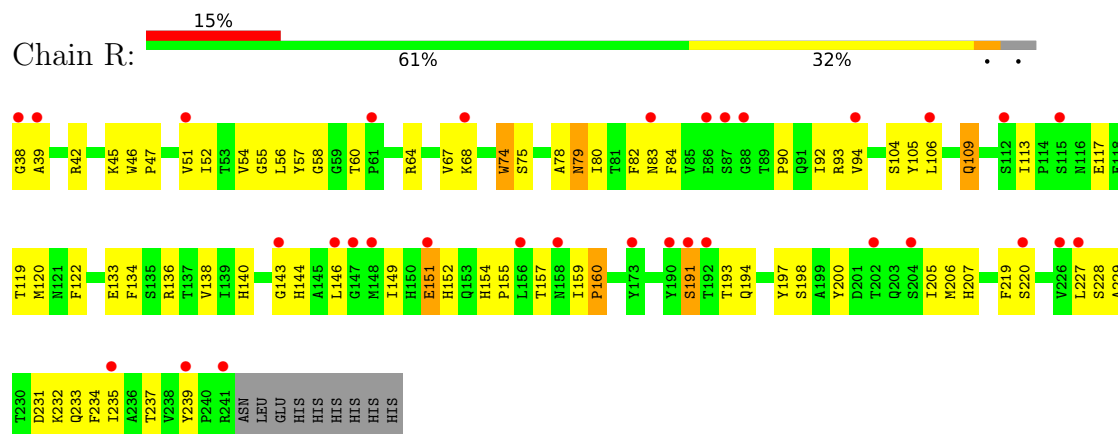
- Molecule 1: Myroilysin



- Molecule 1: Myroilysin



- Molecule 1: Myroilysin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.18Å 115.90Å 163.35Å 90.00° 93.71° 90.00°	Depositor
Resolution (Å)	19.99 – 2.60 19.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (19.99-2.60) 99.7 (19.99-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.31 (at 2.59Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.187 , 0.254 0.206 , 0.241	Depositor DCC
$R_{free}$ test set	7188 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.1	Xtrriage
Anisotropy	0.447	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 32.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	30072	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: ZN, PO4

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	0.42	0/1670	0.43	0/2280
1	B	0.41	0/1686	0.43	0/2302
1	C	0.41	0/1678	0.42	0/2291
1	D	0.57	0/1678	0.44	0/2291
1	E	0.38	0/1678	0.42	0/2291
1	F	0.35	0/1670	0.40	0/2280
1	G	0.37	0/1670	0.42	0/2280
1	H	0.37	0/1670	0.40	0/2280
1	I	0.45	0/1670	0.43	0/2280
1	J	0.39	0/1695	0.41	0/2314
1	K	0.36	0/1686	0.42	0/2302
1	L	0.41	0/1670	0.42	0/2280
1	M	0.49	0/1678	0.45	0/2291
1	N	0.46	0/1650	0.55	0/2257
1	O	0.36	0/1670	0.41	0/2280
1	P	0.35	0/1670	0.42	0/2280
1	Q	0.37	0/1670	0.48	0/2280
1	R	0.50	0/1670	0.63	0/2280
All	All	0.42	0/30129	0.45	0/41139

There are no bond length outliers.

There are no bond angle outliers.

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	1621	0	1513	9	0
1	B	1637	0	1530	7	0
1	C	1629	0	1519	8	0
1	D	1629	0	1519	6	0
1	E	1629	0	1519	10	0
1	F	1621	0	1513	6	0
1	G	1621	0	1513	11	0
1	H	1621	0	1513	15	0
1	I	1621	0	1513	16	0
1	J	1646	0	1536	5	0
1	K	1637	0	1530	17	0
1	L	1621	0	1513	6	0
1	M	1629	0	1519	15	0
1	N	1601	0	1466	62	0
1	O	1621	0	1513	16	0
1	P	1621	0	1513	13	0
1	Q	1621	0	1512	22	0
1	R	1621	0	1512	59	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
2	M	1	0	0	0	0
2	N	1	0	0	0	0
2	O	1	0	0	0	0
2	P	1	0	0	0	0
2	Q	1	0	0	0	0
2	R	1	0	0	0	0
3	B	5	0	0	0	0
3	D	5	0	0	0	0
3	E	5	0	0	0	0
3	F	5	0	0	0	0
3	G	5	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	5	0	0	0	0
3	I	5	0	0	0	0
3	J	5	0	0	0	0
3	K	5	0	0	0	0
3	M	5	0	0	0	0
3	P	5	0	0	0	0
3	Q	5	0	0	0	0
4	A	75	0	0	1	0
4	B	61	0	0	1	0
4	C	74	0	0	0	0
4	D	80	0	0	0	0
4	E	54	0	0	1	0
4	F	50	0	0	1	0
4	G	33	0	0	0	0
4	H	39	0	0	1	0
4	I	32	0	0	1	0
4	J	81	0	0	1	0
4	K	18	0	0	0	0
4	L	52	0	0	0	0
4	M	5	0	0	0	0
4	N	2	0	0	0	0
4	O	28	0	0	0	0
4	P	29	0	0	0	0
4	Q	32	0	0	0	0
4	R	2	0	0	0	0
All	All	30072	0	27266	286	0

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 5.

The worst 5 of 286 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:Q:154:HIS:O	1:Q:157:THR:HG22	1.64	0.97
1:N:54:VAL:HG12	1:N:92:ILE:HB	1.57	0.85
1:N:223:ASN:C	1:N:223:ASN:HD22	1.82	0.83
1:R:39:ALA:HB3	1:R:149:ILE:HG12	1.60	0.83
1:N:57:TYR:CE1	1:N:93:ARG:HG2	2.13	0.83

There are no symmetry-related clashes.

## 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	202/213 (95%)	196 (97%)	6 (3%)	0	100	100
1	B	204/213 (96%)	200 (98%)	4 (2%)	0	100	100
1	C	203/213 (95%)	194 (96%)	9 (4%)	0	100	100
1	D	203/213 (95%)	195 (96%)	8 (4%)	0	100	100
1	E	203/213 (95%)	196 (97%)	7 (3%)	0	100	100
1	F	202/213 (95%)	195 (96%)	7 (4%)	0	100	100
1	G	202/213 (95%)	192 (95%)	10 (5%)	0	100	100
1	H	202/213 (95%)	195 (96%)	7 (4%)	0	100	100
1	I	202/213 (95%)	195 (96%)	7 (4%)	0	100	100
1	J	205/213 (96%)	203 (99%)	2 (1%)	0	100	100
1	K	204/213 (96%)	197 (97%)	7 (3%)	0	100	100
1	L	202/213 (95%)	195 (96%)	7 (4%)	0	100	100
1	M	203/213 (95%)	197 (97%)	6 (3%)	0	100	100
1	N	202/213 (95%)	186 (92%)	16 (8%)	0	100	100
1	O	202/213 (95%)	196 (97%)	6 (3%)	0	100	100
1	P	202/213 (95%)	196 (97%)	6 (3%)	0	100	100
1	Q	202/213 (95%)	193 (96%)	9 (4%)	0	100	100
1	R	202/213 (95%)	187 (93%)	14 (7%)	1 (0%)	29	52
All	All	3647/3834 (95%)	3508 (96%)	138 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	R	160	PRO

### 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	176/185 (95%)	176 (100%)	0	100	100
1	B	178/185 (96%)	178 (100%)	0	100	100
1	C	177/185 (96%)	177 (100%)	0	100	100
1	D	177/185 (96%)	177 (100%)	0	100	100
1	E	177/185 (96%)	176 (99%)	1 (1%)	86	95
1	F	176/185 (95%)	176 (100%)	0	100	100
1	G	176/185 (95%)	176 (100%)	0	100	100
1	H	176/185 (95%)	176 (100%)	0	100	100
1	I	176/185 (95%)	173 (98%)	3 (2%)	60	81
1	J	179/185 (97%)	179 (100%)	0	100	100
1	K	178/185 (96%)	176 (99%)	2 (1%)	73	88
1	L	176/185 (95%)	175 (99%)	1 (1%)	86	95
1	M	177/185 (96%)	174 (98%)	3 (2%)	60	81
1	N	171/185 (92%)	165 (96%)	6 (4%)	36	62
1	O	176/185 (95%)	174 (99%)	2 (1%)	73	88
1	P	176/185 (95%)	175 (99%)	1 (1%)	86	95
1	Q	176/185 (95%)	173 (98%)	3 (2%)	60	81
1	R	176/185 (95%)	169 (96%)	7 (4%)	31	57
All	All	3174/3330 (95%)	3145 (99%)	29 (1%)	78	91

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	N	223	ASN
1	R	220	SER
1	O	194	GLN
1	R	109	GLN
1	O	165	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	R	79	ASN
1	Q	158	ASN
1	N	223	ASN
1	K	242	ASN
1	O	194	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](#)

Of 30 ligands modelled in this entry, 18 are monoatomic - leaving 12 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	$\# Z  > 2$
3	PO4	E	301	-	4,4,4	2.67	4 (100%)	6,6,6	1.20	1 (16%)
3	PO4	G	301	-	4,4,4	0.93	0	6,6,6	0.41	0
3	PO4	P	301	-	4,4,4	0.94	0	6,6,6	0.43	0
3	PO4	H	301	-	4,4,4	0.91	0	6,6,6	0.44	0
3	PO4	K	301	-	4,4,4	0.91	0	6,6,6	0.46	0
3	PO4	Q	301	-	4,4,4	0.92	0	6,6,6	0.44	0
3	PO4	B	301	-	4,4,4	0.92	0	6,6,6	0.43	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	PO4	I	301	-	4,4,4	0.92	0	6,6,6	0.44	0
3	PO4	M	301	-	4,4,4	0.92	0	6,6,6	0.43	0
3	PO4	D	301	-	4,4,4	0.94	0	6,6,6	0.45	0
3	PO4	J	301	-	4,4,4	0.93	0	6,6,6	0.43	0
3	PO4	F	301	-	4,4,4	0.94	0	6,6,6	0.43	0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	301	PO4	P-O4	-3.16	1.45	1.54
3	E	301	PO4	P-O3	-2.60	1.46	1.54
3	E	301	PO4	P-O2	-2.53	1.47	1.54
3	E	301	PO4	P-O1	-2.29	1.45	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	301	PO4	O4-P-O1	-2.35	102.31	110.89

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	204/213 (95%)	-0.52	0 <b>100</b> <b>100</b>	12, 18, 27, 34	0
1	B	206/213 (96%)	-0.50	0 <b>100</b> <b>100</b>	14, 20, 30, 49	0
1	C	205/213 (96%)	-0.32	1 (0%) 91 89	13, 18, 34, 55	0
1	D	205/213 (96%)	-0.44	1 (0%) 91 89	15, 21, 30, 63	0
1	E	205/213 (96%)	-0.36	2 (0%) 82 80	19, 24, 35, 54	0
1	F	204/213 (95%)	-0.31	2 (0%) 82 80	21, 28, 38, 55	0
1	G	204/213 (95%)	-0.15	2 (0%) 82 80	24, 31, 42, 63	0
1	H	204/213 (95%)	-0.24	0 <b>100</b> <b>100</b>	23, 31, 42, 57	0
1	I	204/213 (95%)	-0.14	2 (0%) 82 80	22, 33, 45, 59	0
1	J	207/213 (97%)	-0.49	0 <b>100</b> <b>100</b>	15, 19, 30, 44	0
1	K	206/213 (96%)	-0.02	5 (2%) 59 53	30, 37, 56, 66	0
1	L	204/213 (95%)	-0.43	0 <b>100</b> <b>100</b>	14, 24, 35, 47	0
1	M	205/213 (96%)	0.07	6 (2%) 51 45	30, 39, 54, 71	0
1	N	204/213 (95%)	0.88	27 (13%) <b>3</b> <b>2</b>	40, 77, 106, 122	0
1	O	204/213 (95%)	-0.01	7 (3%) 45 38	24, 37, 61, 72	0
1	P	204/213 (95%)	-0.03	3 (1%) 73 70	20, 35, 51, 65	0
1	Q	204/213 (95%)	0.38	13 (6%) <b>19</b> <b>14</b>	19, 45, 65, 99	0
1	R	204/213 (95%)	0.93	32 (15%) <b>2</b> <b>1</b>	53, 80, 107, 120	0
All	All	3683/3834 (96%)	-0.10	103 (2%) 53 46	12, 30, 80, 122	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	38	GLY	6.2
1	D	242	ASN	4.6
1	N	225	SER	4.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	O	217	ASN	4.4
1	R	158	ASN	4.3

## 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
3	PO4	D	301	5/5	0.70	0.39	38,44,64,77	0
3	PO4	M	301	5/5	0.78	0.47	59,62,71,90	0
3	PO4	K	301	5/5	0.91	0.29	49,51,60,71	0
3	PO4	I	301	5/5	0.91	0.28	44,47,59,81	0
3	PO4	P	301	5/5	0.92	0.35	40,42,60,61	0
3	PO4	H	301	5/5	0.93	0.35	48,55,59,73	0
2	ZN	R	301	1/1	0.94	0.04	77,77,77,77	0
3	PO4	F	301	5/5	0.94	0.30	38,50,55,58	0
3	PO4	G	301	5/5	0.94	0.38	43,50,66,71	0
3	PO4	B	301	5/5	0.94	0.36	37,37,46,63	0
3	PO4	Q	301	5/5	0.94	0.32	44,50,66,78	0
3	PO4	E	301	5/5	0.96	0.35	43,45,59,60	0
3	PO4	J	301	5/5	0.96	0.21	29,30,43,48	0
2	ZN	N	301	1/1	0.96	0.09	70,70,70,70	0
2	ZN	H	302	1/1	0.97	0.06	24,24,24,24	0
2	ZN	M	302	1/1	0.97	0.05	31,31,31,31	0
2	ZN	P	302	1/1	0.98	0.06	31,31,31,31	0
2	ZN	L	301	1/1	0.98	0.18	38,38,38,38	0
2	ZN	Q	302	1/1	0.99	0.04	40,40,40,40	0
2	ZN	J	302	1/1	0.99	0.13	30,30,30,30	0
2	ZN	F	302	1/1	0.99	0.16	32,32,32,32	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ZN	G	302	1/1	0.99	0.09	28,28,28,28	0
2	ZN	A	301	1/1	0.99	0.16	29,29,29,29	0
2	ZN	O	301	1/1	0.99	0.04	31,31,31,31	0
2	ZN	I	302	1/1	0.99	0.02	25,25,25,25	0
2	ZN	K	302	1/1	1.00	0.03	31,31,31,31	0
2	ZN	C	301	1/1	1.00	0.22	43,43,43,43	0
2	ZN	D	302	1/1	1.00	0.15	30,30,30,30	0
2	ZN	E	302	1/1	1.00	0.12	31,31,31,31	0
2	ZN	B	302	1/1	1.00	0.13	28,28,28,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.