



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:36 am GMT

PDB ID : 4S20  
Title : Structural basis for transcription reactivation by RapA  
Authors : Liu, B.; Zuo, Y.; Steitz, T.A.  
Deposited on : 2015-01-16  
Resolution : 4.70 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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28683
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

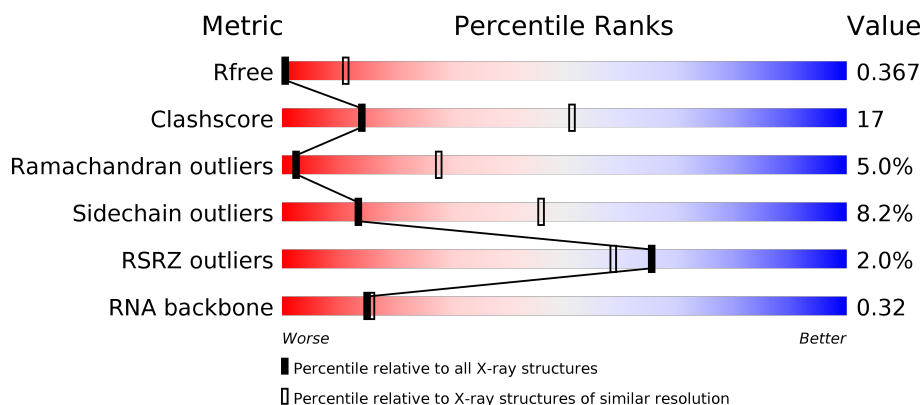
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.70 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))
$R_{free}$	100719	1010 (5.70-3.66)
Clashscore	112137	1059 (5.70-3.70)
Ramachandran outliers	110173	1004 (5.70-3.68)
Sidechain outliers	110143	1026 (5.70-3.66)
RSRZ outliers	101464	1019 (5.70-3.66)
RNA backbone	2435	1040 (6.46-2.90)

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	329	<div> <div></div> <div>62% 7% 31%</div> </div>
1	B	329	<div> <div>4%</div> <div>62% 7% 31%</div> </div>
1	F	329	<div> <div></div> <div>62% 6% 31%</div> </div>
1	G	329	<div> <div></div> <div>62% 7% 31%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	C	1342	
2	H	1342	
3	D	1416	
3	I	1416	
4	E	90	
4	J	90	
5	K	974	
5	L	974	
6	M	15	
6	O	15	
7	N	9	
7	P	9	

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 61154 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 DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	B	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	F	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	G	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1179	Total	C	N	O	S	0	0	0
			9285	5828	1620	1798	39			
2	H	1173	Total	C	N	O	S	0	0	0
			9235	5799	1609	1788	39			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1149	Total	C	N	O	S	0	0	0
			8990	5649	1614	1681	46			
3	I	1160	Total	C	N	O	S	0	0	0
			9068	5701	1626	1695	46			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1408	LEU	-	EXPRESSION TAG	UNP K0BCS5
D	1409	GLU	-	EXPRESSION TAG	UNP K0BCS5
D	1410	VAL	-	EXPRESSION TAG	UNP K0BCS5
D	1411	HIS	-	EXPRESSION TAG	UNP K0BCS5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	1412	HIS	-	EXPRESSION TAG	UNP K0BCS5
D	1413	HIS	-	EXPRESSION TAG	UNP K0BCS5
D	1414	HIS	-	EXPRESSION TAG	UNP K0BCS5
D	1415	HIS	-	EXPRESSION TAG	UNP K0BCS5
D	1416	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1408	LEU	-	EXPRESSION TAG	UNP K0BCS5
I	1409	GLU	-	EXPRESSION TAG	UNP K0BCS5
I	1410	VAL	-	EXPRESSION TAG	UNP K0BCS5
I	1411	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1412	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1413	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1414	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1415	HIS	-	EXPRESSION TAG	UNP K0BCS5
I	1416	HIS	-	EXPRESSION TAG	UNP K0BCS5

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	75	Total	C	N	O	S	0	0	0
			600	365	114	120	1			
4	J	75	Total	C	N	O	S	0	0	0
			600	365	114	120	1			

- Molecule 5 is a protein called RNA polymerase-associated protein RapA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	K	961	Total	C	N	O	S	0	0	0
			7665	4797	1370	1468	30			
5	L	961	Total	C	N	O	S	0	0	0
			7665	4797	1370	1468	30			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-5	HIS	-	EXPRESSION TAG	UNP P60240
K	-4	HIS	-	EXPRESSION TAG	UNP P60240
K	-3	HIS	-	EXPRESSION TAG	UNP P60240
K	-2	HIS	-	EXPRESSION TAG	UNP P60240
K	-1	HIS	-	EXPRESSION TAG	UNP P60240
K	0	HIS	-	EXPRESSION TAG	UNP P60240
K	350	CYS	ARG	ENGINEERED MUTATION	UNP P60240
L	-5	HIS	-	EXPRESSION TAG	UNP P60240

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
L	-4	HIS	-	EXPRESSION TAG	UNP P60240
L	-3	HIS	-	EXPRESSION TAG	UNP P60240
L	-2	HIS	-	EXPRESSION TAG	UNP P60240
L	-1	HIS	-	EXPRESSION TAG	UNP P60240
L	0	HIS	-	EXPRESSION TAG	UNP P60240
L	350	CYS	ARG	ENGINEERED MUTATION	UNP P60240

- Molecule 6 is a DNA chain called 5'-D(P\*AP\*CP\*GP\*AP\*CP\*TP\*GP\*AP\*GP\*CP\*CP\*GP\*AP\*TP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	M	15	Total	C	N	O	P	0	0	0
			311	146	61	89	15			
6	O	15	Total	C	N	O	P	0	0	0
			311	146	61	89	15			

- Molecule 7 is a RNA chain called 5'-R(P\*AP\*UP\*CP\*GP\*GP\*CP\*UP\*CP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	N	9	Total	C	N	O	P	0	0	0
			191	85	33	64	9			
7	P	9	Total	C	N	O	P	0	0	0
			191	85	33	64	9			

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	2	Total	Zn	0	0
			2	2		
8	D	2	Total	Zn	0	0
			2	2		

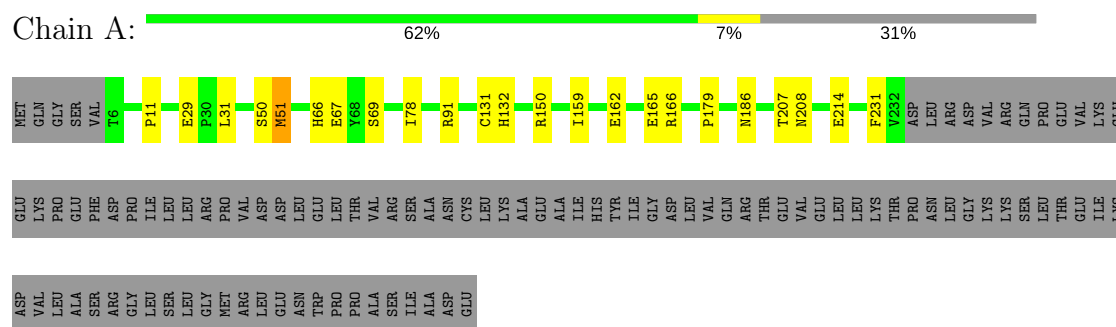
- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	I	1	Total	Mg	0	0
			1	1		
9	D	1	Total	Mg	0	0
			1	1		

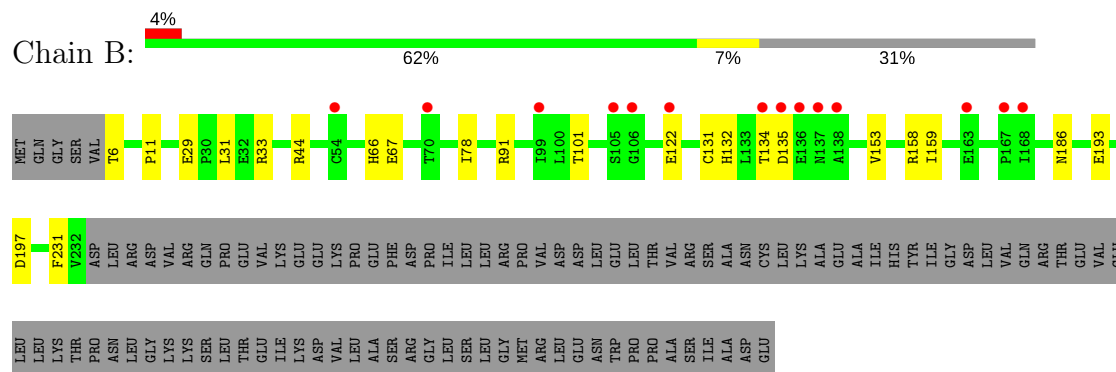
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

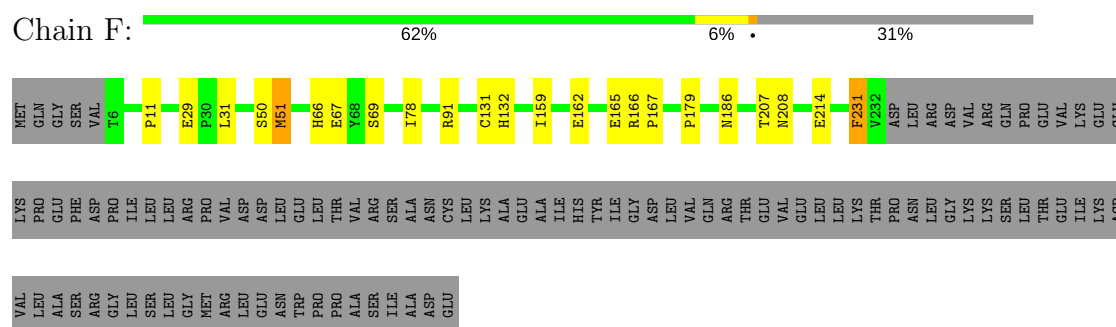
- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha

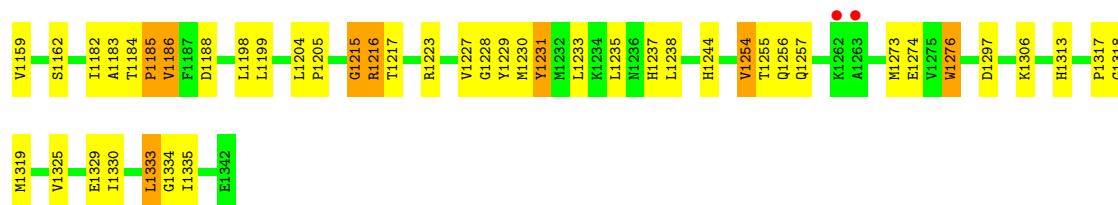
Category	Percentage
Very bad	62%
Bad	7%
Good	31%



Category	Percentage
Very good	63%
Good	22%
Not good	12%
Very bad	3%

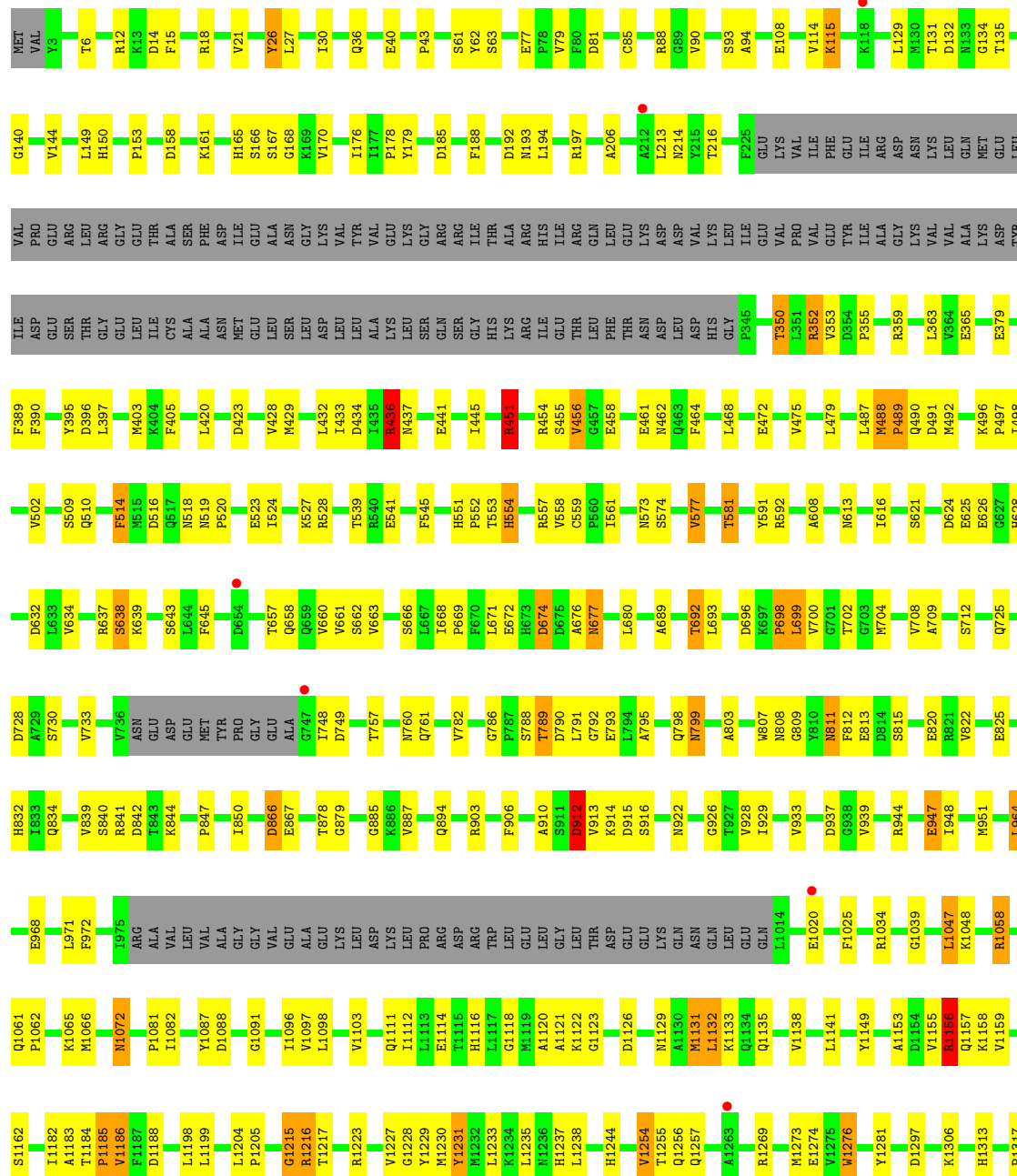




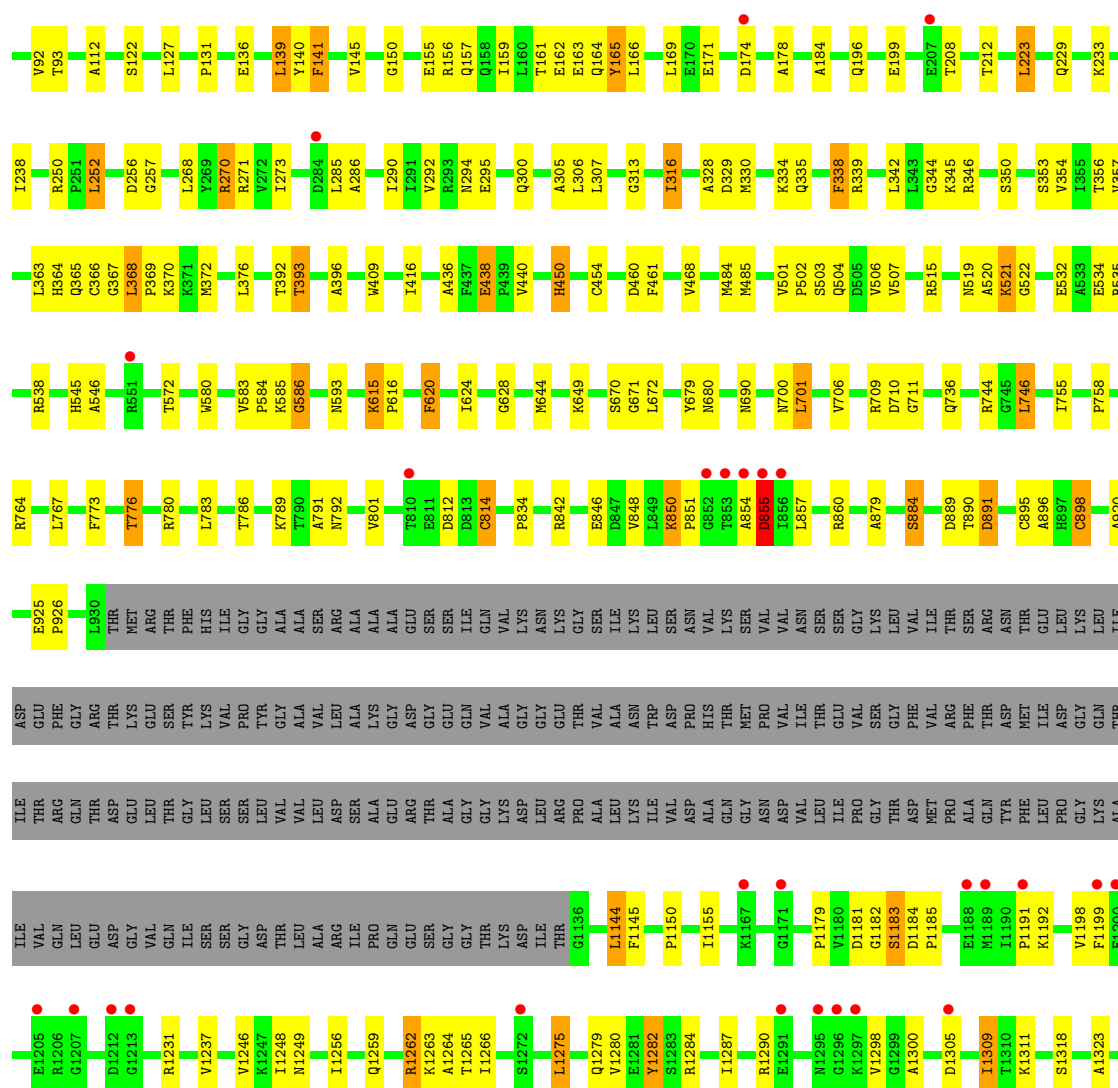


● Molecule 2: DNA-directed RNA polymerase subunit beta

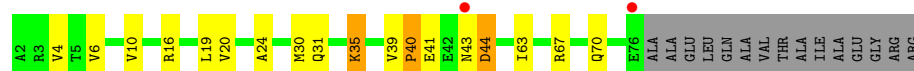
Chain H: 62% 22% 13%



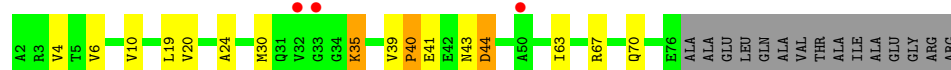




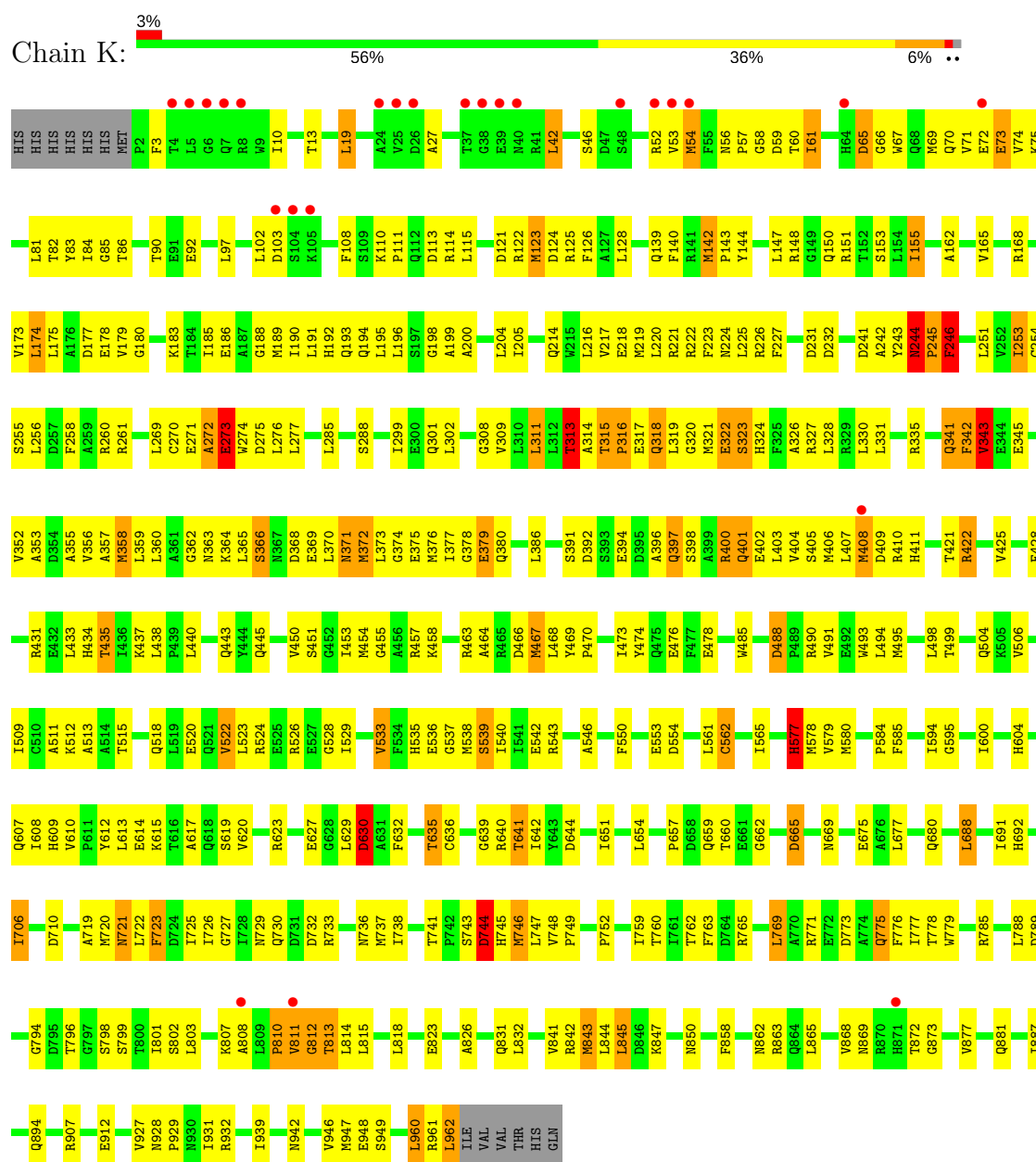
• Molecule 4: DNA-directed RNA polymerase subunit omega

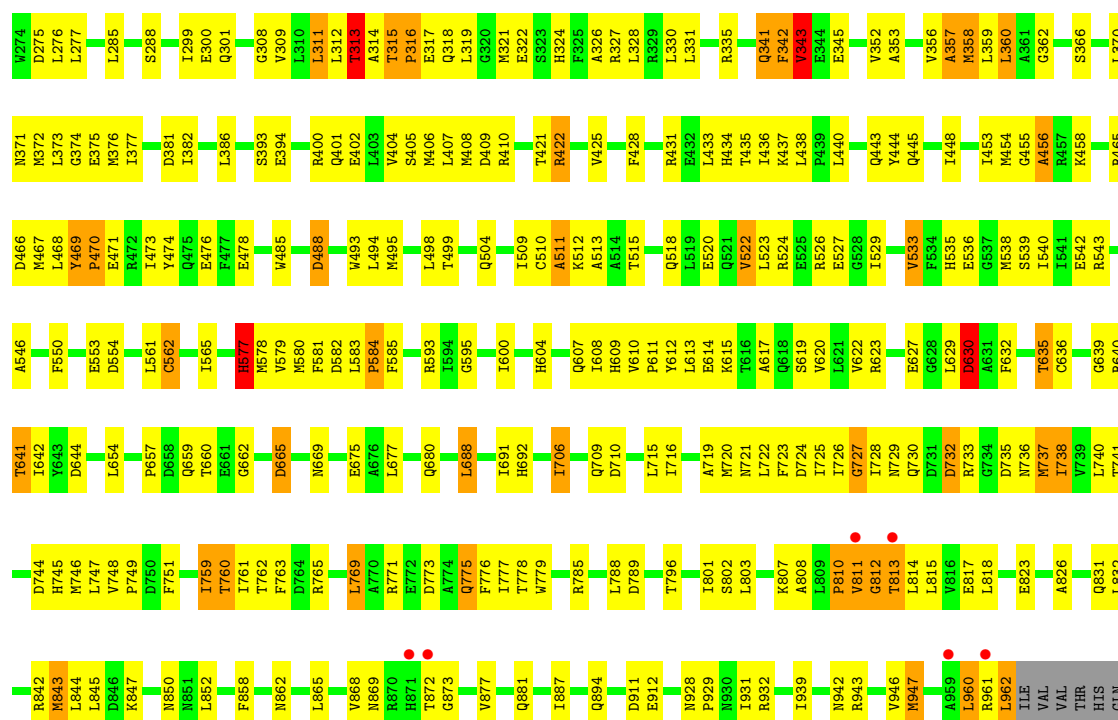


• Molecule 4: DNA-directed RNA polymerase subunit omega

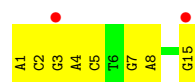


• Molecule 5: RNA polymerase-associated protein RapA

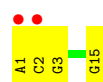
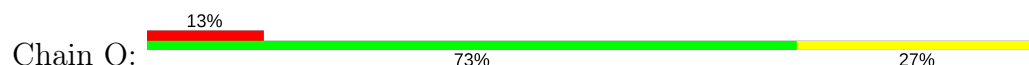




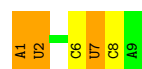
- Molecule 6: 5'-D(P\*AP\*CP\*GP\*AP\*CP\*TP\*GP\*AP\*GP\*CP\*CP\*GP\*AP\*TP\*G)-3'



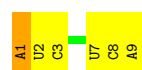
- Molecule 6: 5'-D(P\*AP\*CP\*GP\*AP\*CP\*TP\*GP\*AP\*GP\*CP\*CP\*GP\*AP\*TP\*G)-3'



- Molecule 7: 5'-R(P\*AP\*UP\*CP\*GP\*GP\*CP\*UP\*CP\*A)-3'



- Molecule 7: 5'-R(P\*AP\*UP\*CP\*GP\*GP\*CP\*UP\*CP\*A)-3'



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	336.09Å 158.93Å 255.01Å 90.00° 101.28° 90.00°	Depositor
Resolution (Å)	20.00 – 4.70 20.00 – 4.70	Depositor EDS
% Data completeness (in resolution range)	93.2 (20.00-4.70) 94.5 (20.00-4.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.21 (at 4.74Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.273 , 0.369 0.288 , 0.367	Depositor DCC
$R_{free}$ test set	3240 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	196.2	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.19 , 247.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	61154	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	292.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% 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, MG

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	0.33	0/1781	0.54	0/2414
1	B	0.32	0/1781	0.55	0/2414
1	F	0.33	0/1781	0.54	0/2414
1	G	0.32	0/1781	0.55	0/2414
2	C	0.36	0/9435	0.60	1/12729 (0.0%)
2	H	0.35	0/9385	0.60	1/12662 (0.0%)
3	D	0.35	0/9128	0.58	1/12313 (0.0%)
3	I	0.35	0/9208	0.58	3/12426 (0.0%)
4	E	0.36	0/602	0.52	0/810
4	J	0.35	0/602	0.52	0/810
5	K	0.38	0/7808	0.65	2/10576 (0.0%)
5	L	0.38	0/7808	0.66	2/10576 (0.0%)
6	M	0.60	1/349 (0.3%)	0.70	0/535
6	O	0.61	1/349 (0.3%)	0.72	0/535
7	N	0.73	1/212 (0.5%)	0.63	0/326
7	P	0.72	1/212 (0.5%)	0.67	0/326
All	All	0.36	4/62222 (0.0%)	0.60	10/84280 (0.0%)

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
5	L	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	N	1	A	OP3-P	-10.19	1.49	1.61
6	O	1	DA	OP3-P	-10.18	1.49	1.61

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	P	1	A	OP3-P	-10.14	1.49	1.61
6	M	1	DA	OP3-P	-10.10	1.49	1.61

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	814	CYS	N-CA-C	-5.33	96.60	111.00
5	L	577	HIS	C-N-CA	-5.32	108.39	121.70
5	K	577	HIS	N-CA-C	-5.27	96.77	111.00
2	C	436	ARG	NE-CZ-NH1	5.24	122.92	120.30
3	I	850	LYS	C-N-CD	5.22	139.37	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	L	148	ARG	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	1759	0	1785	10	0
1	B	1759	0	1785	11	0
1	F	1759	0	1785	11	0
1	G	1759	0	1785	11	0
2	C	9285	0	9291	156	0
2	H	9235	0	9242	161	0
3	D	8990	0	9174	155	0
3	I	9068	0	9265	152	0
4	E	600	0	607	10	0
4	J	600	0	607	6	0
5	K	7665	0	7524	735	0
5	L	7665	0	7524	786	0
6	M	311	0	168	19	0
6	O	311	0	168	5	0
7	N	191	0	98	11	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	P	191	0	98	8	0
8	D	2	0	0	1	0
8	I	2	0	0	0	0
9	D	1	0	0	0	0
9	I	1	0	0	0	0
All	All	61154	0	60906	2127	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 17.

The worst 5 of 2127 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)
5:L:720:MET:HE3	5:L:723:PHE:CE2	1.20	1.64
5:K:69:MET:CG	5:K:97:LEU:HD22	1.28	1.61
5:L:184:THR:HG22	5:L:312:LEU:CD2	1.17	1.61
5:L:219:MET:HG2	5:L:227:PHE:CE1	1.37	1.60
5:L:468:LEU:HA	5:L:617:ALA:CB	1.24	1.59

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	225/329 (68%)	202 (90%)	19 (8%)	4 (2%)	10	50
1	B	225/329 (68%)	205 (91%)	20 (9%)	0	100	100
1	F	225/329 (68%)	202 (90%)	19 (8%)	4 (2%)	10	50
1	G	225/329 (68%)	205 (91%)	19 (8%)	1 (0%)	38	77
2	C	1171/1342 (87%)	894 (76%)	211 (18%)	66 (6%)	2	25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	1165/1342 (87%)	889 (76%)	206 (18%)	70 (6%)	2	23
3	D	1143/1416 (81%)	867 (76%)	235 (21%)	41 (4%)	4	35
3	I	1156/1416 (82%)	869 (75%)	240 (21%)	47 (4%)	3	31
4	E	73/90 (81%)	59 (81%)	8 (11%)	6 (8%)	1	16
4	J	73/90 (81%)	59 (81%)	8 (11%)	6 (8%)	1	16
5	K	959/974 (98%)	689 (72%)	201 (21%)	69 (7%)	1	20
5	L	959/974 (98%)	694 (72%)	200 (21%)	65 (7%)	1	21
All	All	7599/8960 (85%)	5834 (77%)	1386 (18%)	379 (5%)	2	28

5 of 379 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	51	MET
2	C	43	PRO
2	C	63	SER
2	C	115	LYS
2	C	216	THR

### 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	195/286 (68%)	191 (98%)	4 (2%)	59	81
1	B	195/286 (68%)	188 (96%)	7 (4%)	40	69
1	F	195/286 (68%)	191 (98%)	4 (2%)	59	81
1	G	195/286 (68%)	188 (96%)	7 (4%)	40	69
2	C	1016/1157 (88%)	937 (92%)	79 (8%)	15	48
2	H	1011/1157 (87%)	934 (92%)	77 (8%)	15	49
3	D	962/1177 (82%)	879 (91%)	83 (9%)	12	44
3	I	971/1177 (82%)	889 (92%)	82 (8%)	13	45
4	E	65/74 (88%)	62 (95%)	3 (5%)	31	64

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	J	65/74 (88%)	62 (95%)	3 (5%)	31	64
5	K	822/835 (98%)	724 (88%)	98 (12%)	6	29
5	L	822/835 (98%)	732 (89%)	90 (11%)	7	33
All	All	6514/7630 (85%)	5977 (92%)	537 (8%)	13	46

5 of 537 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	947	GLU
3	I	416	ILE
5	L	518	GLN
2	H	1072	ASN
3	I	88	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
3	I	419	HIS
5	K	150	GLN
5	L	729	ASN
3	I	875	ASN
5	K	239	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
7	N	8/9 (88%)	3 (37%)	0
7	P	8/9 (88%)	1 (12%)	0
All	All	16/18 (88%)	4 (25%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
7	N	2	U
7	N	6	C
7	N	7	U
7	P	7	U

There are no RNA pucker outliers to report.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	227/329 (68%)	-0.50	0 100 100	178, 268, 367, 400	0
1	B	227/329 (68%)	-0.05	14 (6%) 21 19	237, 380, 491, 539	0
1	F	227/329 (68%)	-0.49	0 100 100	192, 268, 357, 401	0
1	G	227/329 (68%)	-0.20	1 (0%) 92 88	256, 379, 484, 537	0
2	C	1179/1342 (87%)	-0.58	15 (1%) 77 70	126, 241, 378, 510	0
2	H	1173/1342 (87%)	-0.57	6 (0%) 90 86	143, 251, 375, 532	0
3	D	1149/1416 (81%)	-0.40	29 (2%) 58 50	181, 301, 441, 641	0
3	I	1160/1416 (81%)	-0.38	32 (2%) 53 46	175, 282, 433, 662	0
4	E	75/90 (83%)	0.14	2 (2%) 55 47	285, 343, 385, 397	0
4	J	75/90 (83%)	0.10	3 (4%) 39 32	329, 349, 365, 376	0
5	K	961/974 (98%)	-0.35	25 (2%) 56 49	178, 278, 461, 664	0
5	L	961/974 (98%)	-0.35	26 (2%) 55 47	169, 269, 469, 638	0
6	M	15/15 (100%)	0.51	2 (13%) 4 6	321, 353, 461, 465	0
6	O	15/15 (100%)	0.45	2 (13%) 4 6	331, 351, 543, 550	0
7	N	9/9 (100%)	-0.01	0 100 100	258, 268, 381, 391	0
7	P	9/9 (100%)	-0.04	0 100 100	236, 245, 375, 395	0
All	All	7689/9008 (85%)	-0.41	157 (2%) 65 58	126, 279, 441, 664	0

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	L	65	ASP	15.6
5	K	6	GLY	8.5
5	L	64	HIS	8.1
5	K	5	LEU	6.4
5	K	25	VAL	6.2

## 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( $\text{\AA}^2$ )	Q<0.9
8	ZN	D	1502	1/1	0.98	0.20	-0.01	84,84,84,84	0
8	ZN	D	1501	1/1	0.90	0.17	-0.50	106,106,106,106	0
8	ZN	I	1501	1/1	0.76	0.13	-0.89	131,131,131,131	0
8	ZN	I	1502	1/1	0.96	0.10	-1.48	117,117,117,117	0
9	MG	I	1503	1/1	0.86	0.11	-	52,52,52,52	0
9	MG	D	1503	1/1	0.80	0.24	-	10,10,10,10	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.