



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 06:45 pm GMT

PDB ID : 3B0X  
Title : K263A mutant of PolX from *Thermus thermophilus* HB8 complexed with Ca-dGTP  
Authors : Nakane, S.; Nakagawa, N.; Masui, R.; Kuramitsu, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2011-06-17  
Resolution : 1.36 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

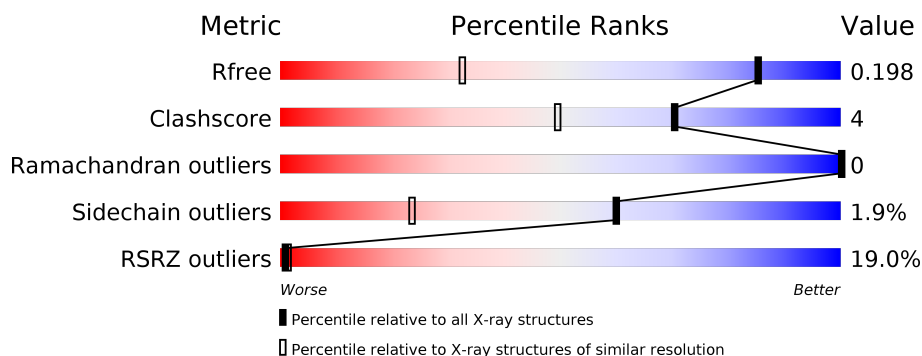
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.36 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))
$R_{free}$	111664	1197 (1.38-1.34)
Clashscore	122126	1232 (1.38-1.34)
Ramachandran outliers	120053	1215 (1.38-1.34)
Sidechain outliers	120020	1215 (1.38-1.34)
RSRZ outliers	108989	1177 (1.38-1.34)

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	575	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5125 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 polymerase beta family (X family).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	575	Total	C	N	O	S	0	4	0
			4536	2875	813	838	10			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	263	ALA	LYS	ENGINEERED MUTATION	UNP Q5SJ64

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	Ca	0	0
			4	4		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

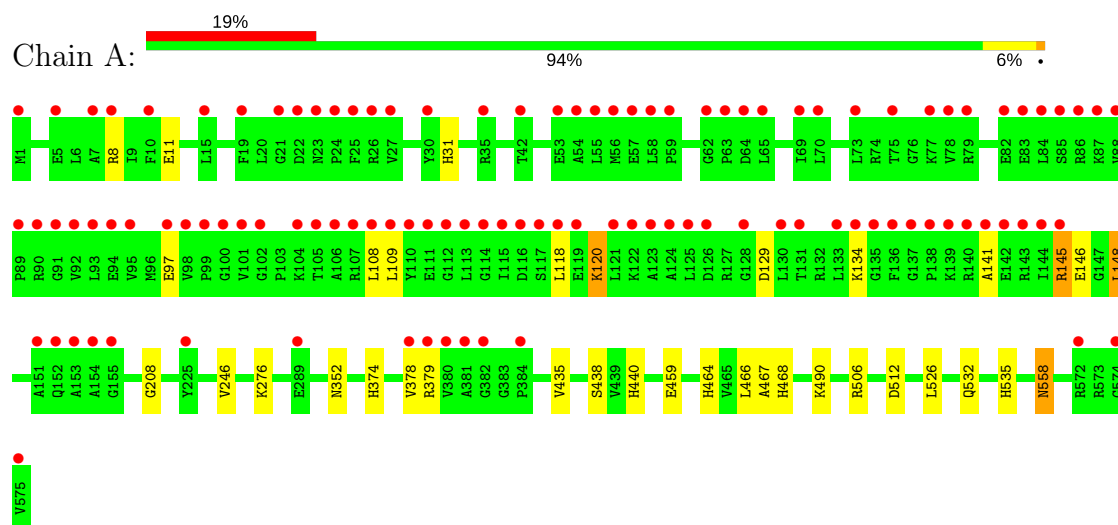
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	551	Total O 551 551	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 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 polymerase beta family (X family)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.09Å 53.20Å 84.86Å 90.00° 107.65° 90.00°	Depositor
Resolution (Å)	50.00 – 1.36 17.31 – 1.36	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-1.36) 99.4 (17.31-1.36)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.80 (at 1.36Å)	Xtriage
Refinement program	REFMAC refmac_5.5.0110	Depositor
R, $R_{free}$	0.171 , 0.198 0.171 , 0.198	Depositor DCC
$R_{free}$ test set	6294 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.4	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.46 , 57.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5125	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.25% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CA, DGT, CL

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	0.50	0/4637	0.70	2/6269 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	506	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	A	512	ASP	CB-CG-OD1	5.36	123.12	118.30

There are no chirality outliers.

There are no planarity outliers.

### 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	4536	0	4610	36	0
2	A	2	0	0	0	0
3	A	4	0	0	0	0
4	A	1	0	0	0	0
5	A	31	0	12	0	0
6	A	551	0	0	9	0
All	All	5125	0	4622	36	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 4.

All (36) 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:145:ARG:HH11	1:A:145:ARG:HG2	1.21	1.02
1:A:435:VAL:H	1:A:464:HIS:HD2	1.13	0.95
1:A:374:HIS:CE1	1:A:378:VAL:HG11	2.04	0.92
1:A:374:HIS:ND1	1:A:378:VAL:HG11	1.84	0.92
1:A:141:ALA:HA	6:A:997:HOH:O	1.74	0.87
1:A:526:LEU:H	1:A:558:ASN:HD21	1.19	0.84
1:A:129:ASP:HB3	6:A:797:HOH:O	1.88	0.73
1:A:435:VAL:H	1:A:464:HIS:CD2	2.05	0.70
1:A:466[B]:LEU:CD1	6:A:758:HOH:O	2.42	0.68
1:A:352:ASN:HD21	1:A:532:GLN:HA	1.61	0.64
1:A:558:ASN:H	1:A:558:ASN:HD22	1.47	0.62
1:A:145:ARG:NH1	1:A:145:ARG:HG2	2.01	0.61
1:A:145:ARG:CG	1:A:145:ARG:HH11	2.03	0.61
1:A:435:VAL:N	1:A:464:HIS:HD2	1.93	0.60
1:A:526:LEU:H	1:A:558:ASN:ND2	1.97	0.58
1:A:374:HIS:HB2	1:A:379:ARG:HD3	1.86	0.57
1:A:438:SER:OG	1:A:440:HIS:CE1	2.60	0.54
1:A:11:GLU:OE1	1:A:31:HIS:HD2	1.92	0.52
1:A:352:ASN:ND2	1:A:532:GLN:HA	2.24	0.52
1:A:378:VAL:HG13	1:A:440:HIS:CD2	2.46	0.49
1:A:8:ARG:NH2	1:A:97:GLU:OE1	2.39	0.49
1:A:459:GLU:HG2	1:A:490:LYS:NZ	2.29	0.48
1:A:558:ASN:H	1:A:558:ASN:ND2	2.14	0.46
1:A:532:GLN:NE2	6:A:731:HOH:O	2.49	0.45
1:A:129:ASP:C	6:A:797:HOH:O	2.55	0.45
1:A:120:LYS:O	1:A:120:LYS:HD2	2.18	0.44
1:A:374:HIS:ND1	1:A:378:VAL:CG1	2.70	0.43
1:A:466[B]:LEU:HD12	6:A:758:HOH:O	2.12	0.43
1:A:118:LEU:HB3	1:A:148:LEU:HD11	2.00	0.43
1:A:467:ALA:O	1:A:468:HIS:C	2.56	0.43
1:A:532:GLN:HG2	1:A:535:HIS:CE1	2.54	0.42
1:A:466[B]:LEU:HG	6:A:758:HOH:O	2.17	0.42
1:A:208:GLY:HA2	1:A:246:VAL:HG11	2.01	0.42
1:A:379:ARG:NH1	6:A:885:HOH:O	2.51	0.42
1:A:129:ASP:CB	6:A:797:HOH:O	2.58	0.42
1:A:145:ARG:NH1	1:A:145:ARG:CG	2.70	0.40

There are no symmetry-related clashes.



## 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	577/575 (100%)	564 (98%)	13 (2%)	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	469/465 (101%)	460 (98%)	9 (2%)	60	23

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	LEU
1	A	109	LEU
1	A	120	LYS
1	A	134	LYS
1	A	145	ARG
1	A	146	GLU
1	A	148	LEU
1	A	276	LYS
1	A	558	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	HIS
1	A	152	GLN
1	A	274	GLN
1	A	317	GLN
1	A	352	ASN
1	A	440	HIS
1	A	464	HIS
1	A	532	GLN
1	A	558	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 8 ligands modelled in this entry, 7 are monoatomic - leaving 1 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	$\# Z  > 2$
5	DGT	A	583	3	27,33,33	1.11	2 (7%)	31,52,52	1.74	4 (12%)

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
5	DGT	A	583	3	-	0/18/34/34	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	583	DGT	C5-C4	2.98	1.47	1.40
5	A	583	DGT	C6-C5	3.42	1.47	1.41

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	583	DGT	C5-C6-N1	-4.74	116.72	123.47
5	A	583	DGT	C6-C5-C4	-4.37	116.56	120.85
5	A	583	DGT	N3-C2-N1	-2.06	124.39	127.41
5	A	583	DGT	C6-N1-C2	5.04	123.32	116.06

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	575/575 (100%)	0.90	109 (18%) ⓘ ⓘ	8, 14, 42, 56	1 (0%)

All (109) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	135	GLY	10.7
1	A	25	PHE	9.7
1	A	154	ALA	7.9
1	A	89	PRO	7.5
1	A	381	ALA	7.4
1	A	108	LEU	6.8
1	A	136	PHE	6.7
1	A	1	MET	6.5
1	A	110	TYR	6.5
1	A	133	LEU	6.5
1	A	91	GLY	6.4
1	A	134	LYS	6.1
1	A	113	LEU	6.1
1	A	138	PRO	6.1
1	A	380	VAL	6.0
1	A	101	VAL	5.9
1	A	382	GLY	5.7
1	A	139	LYS	5.7
1	A	125	LEU	5.7
1	A	111	GLU	5.5
1	A	572	ARG	5.4
1	A	79	ARG	5.2
1	A	23	ASN	5.0
1	A	155	GLY	5.0
1	A	142	GLU	4.9
1	A	137	GLY	4.9
1	A	86	ARG	4.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	107	ARG	4.8
1	A	90	ARG	4.8
1	A	92	VAL	4.8
1	A	105	THR	4.8
1	A	379	ARG	4.8
1	A	27	VAL	4.7
1	A	124	ALA	4.6
1	A	104	LYS	4.6
1	A	98	VAL	4.5
1	A	63	PRO	4.5
1	A	112	GLY	4.4
1	A	378	VAL	4.3
1	A	143	ARG	4.1
1	A	119	GLU	4.0
1	A	64	ASP	4.0
1	A	100	GLY	4.0
1	A	575	VAL	3.9
1	A	122	LYS	3.9
1	A	116	ASP	3.8
1	A	82	GLU	3.7
1	A	73	LEU	3.7
1	A	131	THR	3.6
1	A	69	ILE	3.6
1	A	109	LEU	3.5
1	A	53	GLU	3.5
1	A	118	LEU	3.5
1	A	106	ALA	3.4
1	A	58	LEU	3.4
1	A	26	ARG	3.3
1	A	88	VAL	3.2
1	A	123	ALA	3.2
1	A	59	PRO	3.2
1	A	21	GLY	3.2
1	A	126	ASP	3.2
1	A	55	LEU	3.1
1	A	78	VAL	3.1
1	A	151	ALA	3.1
1	A	77	LYS	3.0
1	A	115	ILE	2.9
1	A	87	LYS	2.9
1	A	8	ARG	2.9
1	A	93	LEU	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	144	ILE	2.9
1	A	574	GLY	2.8
1	A	117	SER	2.8
1	A	121	LEU	2.8
1	A	56	MET	2.7
1	A	99	PRO	2.7
1	A	140	ARG	2.7
1	A	54	ALA	2.6
1	A	57	GLU	2.6
1	A	128	GLY	2.5
1	A	153	ALA	2.5
1	A	42	THR	2.5
1	A	65	LEU	2.5
1	A	85	SER	2.5
1	A	102	GLY	2.4
1	A	15	LEU	2.4
1	A	75	THR	2.4
1	A	83	GLU	2.4
1	A	130	LEU	2.3
1	A	384	PRO	2.3
1	A	141	ALA	2.3
1	A	95	VAL	2.3
1	A	84	LEU	2.3
1	A	289	GLU	2.2
1	A	94	GLU	2.2
1	A	145	ARG	2.2
1	A	62	GLY	2.2
1	A	10	PHE	2.2
1	A	24	PRO	2.2
1	A	30	TYR	2.1
1	A	19	PHE	2.1
1	A	7	ALA	2.1
1	A	114	GLY	2.1
1	A	152	GLN	2.1
1	A	5	GLU	2.1
1	A	97	GLU	2.1
1	A	22	ASP	2.0
1	A	35	ARG	2.0
1	A	225	TYR	2.0
1	A	70	LEU	2.0

## 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. 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( $\text{\AA}^2$ )	Q<0.9
3	CA	A	581	1/1	0.95	0.20	39,39,39,39	0
5	DGT	A	583	31/31	0.97	0.08	9,12,22,26	0
2	ZN	A	577	1/1	0.98	0.17	18,18,18,18	1
2	ZN	A	576	1/1	0.98	0.14	22,22,22,22	1
3	CA	A	578	1/1	1.00	0.02	10,10,10,10	0
4	CL	A	582	1/1	1.00	0.04	21,21,21,21	0
3	CA	A	579	1/1	1.00	0.02	11,11,11,11	0
3	CA	A	580	1/1	1.00	0.02	13,13,13,13	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.