



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2018 – 09:44 PM EST

PDB ID : 5CGE  
Title : Structure of Hydroxyethylthiazole Kinase ThiM from *Staphylococcus aureus* in complex with substrate analog 2-(2-methyl-1H-imidazole-1-yl)ethanol  
Authors : Kuenz, M.; Drebes, J.; Windshuegel, B.; Cang, H.; Wrenger, C.; Betzel, C.  
Deposited on : 2015-07-09  
Resolution : 1.62 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
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) : rb-20030736

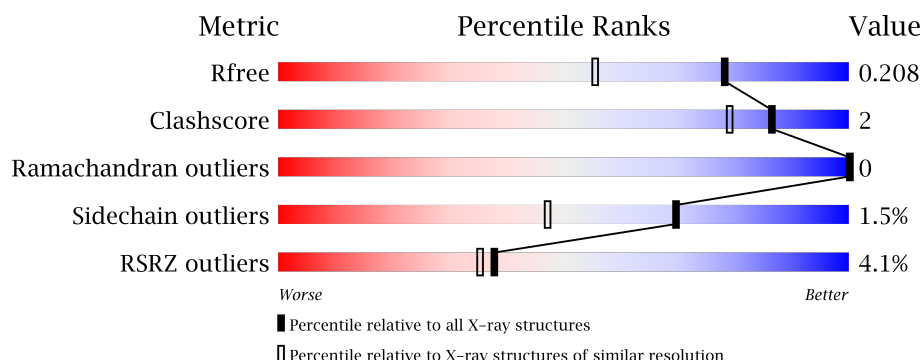
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.62 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	3539 (1.64-1.60)
Clashscore	112137	3855 (1.64-1.60)
Ramachandran outliers	110173	3764 (1.64-1.60)
Sidechain outliers	110143	3763 (1.64-1.60)
RSRZ outliers	101464	3562 (1.64-1.60)

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 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	277	<div> <div>3%</div> <div>82% 8% • 8%</div> </div>
1	B	277	<div> <div>4%</div> <div>85% 9% • 5%</div> </div>
1	C	277	<div> <div>%</div> <div>83% 6% • 10%</div> </div>
1	D	277	<div> <div>6%</div> <div>81% 7% • 12%</div> </div>
1	E	277	<div> <div>5%</div> <div>89% 7% •</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	277	<div><div></div><div>4%</div><div>83%</div><div>8%</div><div>8%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12018 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 Hydroxyethylthiazole kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	27	0	0
			1901	1211	312	372	6			
1	B	262	Total	C	N	O	S	31	0	0
			1954	1242	320	385	7			
1	C	250	Total	C	N	O	S	21	0	0
			1878	1198	307	367	6			
1	E	267	Total	C	N	O	S	29	0	0
			2006	1276	327	396	7			
1	D	245	Total	C	N	O	S	33	0	0
			1817	1164	299	348	6			
1	F	255	Total	C	N	O	S	19	0	0
			1897	1209	311	370	7			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	264	GLU	-	expression tag	UNP Q6GEY3
A	265	ASN	-	expression tag	UNP Q6GEY3
A	266	LEU	-	expression tag	UNP Q6GEY3
A	267	TYR	-	expression tag	UNP Q6GEY3
A	268	PHE	-	expression tag	UNP Q6GEY3
A	269	GLN	-	expression tag	UNP Q6GEY3
A	270	SER	-	expression tag	UNP Q6GEY3
A	271	GLY	-	expression tag	UNP Q6GEY3
A	272	HIS	-	expression tag	UNP Q6GEY3
A	273	HIS	-	expression tag	UNP Q6GEY3
A	274	HIS	-	expression tag	UNP Q6GEY3
A	275	HIS	-	expression tag	UNP Q6GEY3
A	276	HIS	-	expression tag	UNP Q6GEY3
A	277	HIS	-	expression tag	UNP Q6GEY3
B	264	GLU	-	expression tag	UNP Q6GEY3
B	265	ASN	-	expression tag	UNP Q6GEY3
B	266	LEU	-	expression tag	UNP Q6GEY3

*Continued on next page...*

*Continued from previous page...*

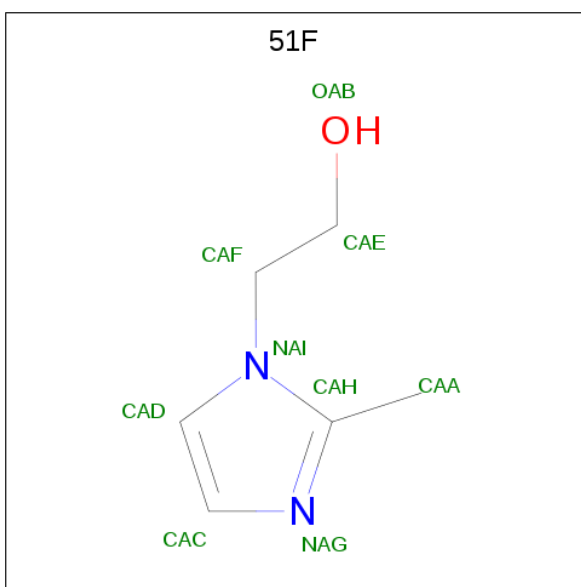
Chain	Residue	Modelled	Actual	Comment	Reference
B	267	TYR	-	expression tag	UNP Q6GEY3
B	268	PHE	-	expression tag	UNP Q6GEY3
B	269	GLN	-	expression tag	UNP Q6GEY3
B	270	SER	-	expression tag	UNP Q6GEY3
B	271	GLY	-	expression tag	UNP Q6GEY3
B	272	HIS	-	expression tag	UNP Q6GEY3
B	273	HIS	-	expression tag	UNP Q6GEY3
B	274	HIS	-	expression tag	UNP Q6GEY3
B	275	HIS	-	expression tag	UNP Q6GEY3
B	276	HIS	-	expression tag	UNP Q6GEY3
B	277	HIS	-	expression tag	UNP Q6GEY3
C	264	GLU	-	expression tag	UNP Q6GEY3
C	265	ASN	-	expression tag	UNP Q6GEY3
C	266	LEU	-	expression tag	UNP Q6GEY3
C	267	TYR	-	expression tag	UNP Q6GEY3
C	268	PHE	-	expression tag	UNP Q6GEY3
C	269	GLN	-	expression tag	UNP Q6GEY3
C	270	SER	-	expression tag	UNP Q6GEY3
C	271	GLY	-	expression tag	UNP Q6GEY3
C	272	HIS	-	expression tag	UNP Q6GEY3
C	273	HIS	-	expression tag	UNP Q6GEY3
C	274	HIS	-	expression tag	UNP Q6GEY3
C	275	HIS	-	expression tag	UNP Q6GEY3
C	276	HIS	-	expression tag	UNP Q6GEY3
C	277	HIS	-	expression tag	UNP Q6GEY3
E	264	GLU	-	expression tag	UNP Q6GEY3
E	265	ASN	-	expression tag	UNP Q6GEY3
E	266	LEU	-	expression tag	UNP Q6GEY3
E	267	TYR	-	expression tag	UNP Q6GEY3
E	268	PHE	-	expression tag	UNP Q6GEY3
E	269	GLN	-	expression tag	UNP Q6GEY3
E	270	SER	-	expression tag	UNP Q6GEY3
E	271	GLY	-	expression tag	UNP Q6GEY3
E	272	HIS	-	expression tag	UNP Q6GEY3
E	273	HIS	-	expression tag	UNP Q6GEY3
E	274	HIS	-	expression tag	UNP Q6GEY3
E	275	HIS	-	expression tag	UNP Q6GEY3
E	276	HIS	-	expression tag	UNP Q6GEY3
E	277	HIS	-	expression tag	UNP Q6GEY3
D	264	GLU	-	expression tag	UNP Q6GEY3
D	265	ASN	-	expression tag	UNP Q6GEY3
D	266	LEU	-	expression tag	UNP Q6GEY3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	267	TYR	-	expression tag	UNP Q6GEY3
D	268	PHE	-	expression tag	UNP Q6GEY3
D	269	GLN	-	expression tag	UNP Q6GEY3
D	270	SER	-	expression tag	UNP Q6GEY3
D	271	GLY	-	expression tag	UNP Q6GEY3
D	272	HIS	-	expression tag	UNP Q6GEY3
D	273	HIS	-	expression tag	UNP Q6GEY3
D	274	HIS	-	expression tag	UNP Q6GEY3
D	275	HIS	-	expression tag	UNP Q6GEY3
D	276	HIS	-	expression tag	UNP Q6GEY3
D	277	HIS	-	expression tag	UNP Q6GEY3
F	264	GLU	-	expression tag	UNP Q6GEY3
F	265	ASN	-	expression tag	UNP Q6GEY3
F	266	LEU	-	expression tag	UNP Q6GEY3
F	267	TYR	-	expression tag	UNP Q6GEY3
F	268	PHE	-	expression tag	UNP Q6GEY3
F	269	GLN	-	expression tag	UNP Q6GEY3
F	270	SER	-	expression tag	UNP Q6GEY3
F	271	GLY	-	expression tag	UNP Q6GEY3
F	272	HIS	-	expression tag	UNP Q6GEY3
F	273	HIS	-	expression tag	UNP Q6GEY3
F	274	HIS	-	expression tag	UNP Q6GEY3
F	275	HIS	-	expression tag	UNP Q6GEY3
F	276	HIS	-	expression tag	UNP Q6GEY3
F	277	HIS	-	expression tag	UNP Q6GEY3

- Molecule 2 is 2-(2-methyl-1H-imidazol-1-yl)ethanol (three-letter code: 51F) (formula: C<sub>6</sub>H<sub>10</sub>N<sub>2</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			9	6	2	1		
2	B	1	Total	C	N	O	0	0
			9	6	2	1		
2	C	1	Total	C	N	O	0	0
			9	6	2	1		
2	E	1	Total	C	N	O	0	0
			9	6	2	1		
2	D	1	Total	C	N	O	0	0
			9	6	2	1		
2	D	1	Total	C	N	O	0	0
			9	6	2	1		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	C	2	Total	Mg	0	0
			2	2		
3	E	1	Total	Mg	0	0
			1	1		

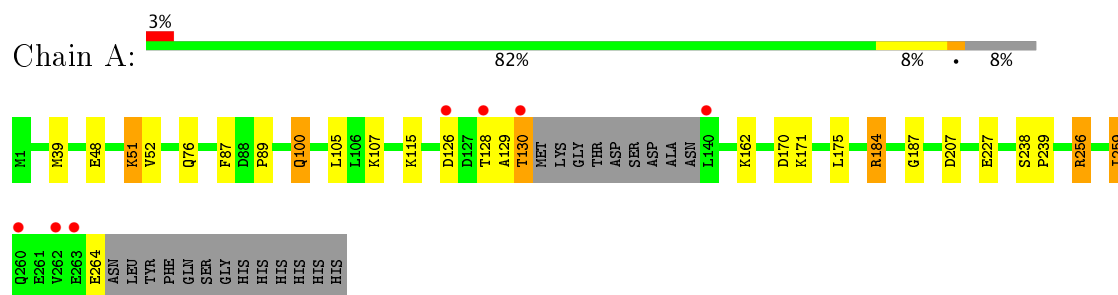
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	105	Total 105	O 105	0	0
4	B	129	Total 129	O 129	0	0
4	C	145	Total 145	O 145	0	0
4	E	58	Total 58	O 58	0	0
4	D	30	Total 30	O 30	0	0
4	F	39	Total 39	O 39	0	0

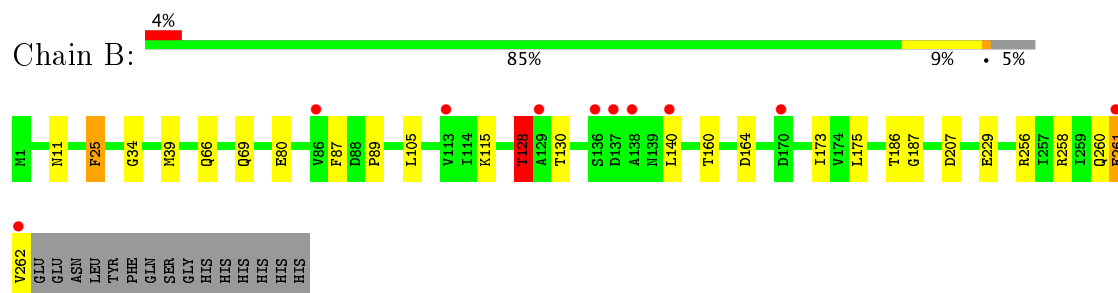
### 3 Residue-property plots [i](#)

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.

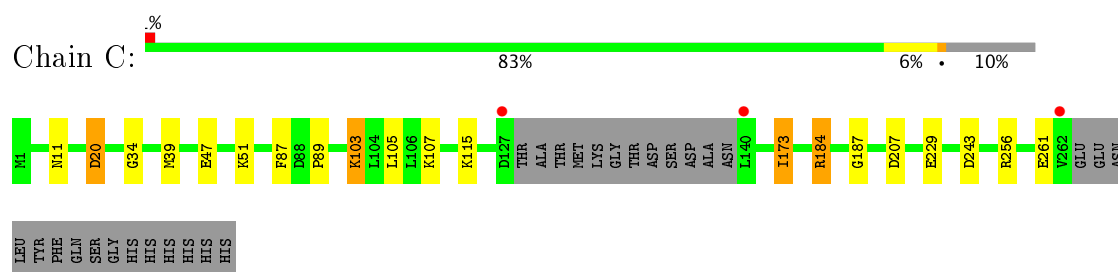
#### • Molecule 1: Hydroxyethylthiazole kinase



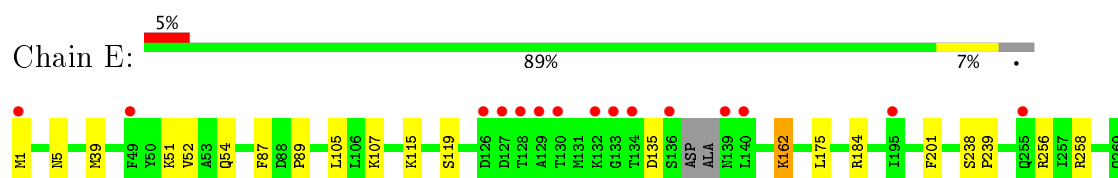
#### • Molecule 1: Hydroxyethylthiazole kinase



#### • Molecule 1: Hydroxyethylthiazole kinase

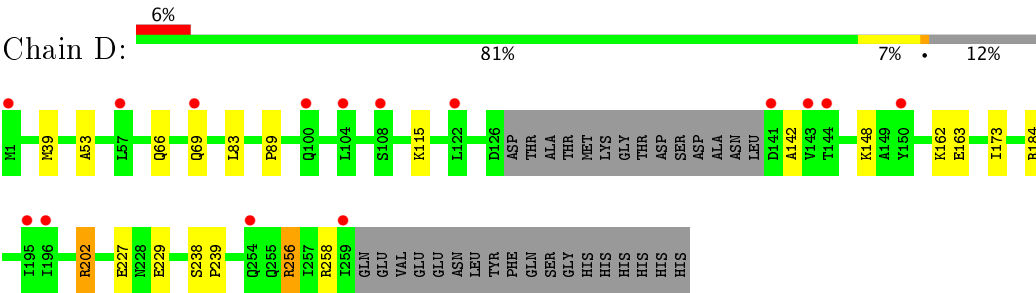


#### • Molecule 1: Hydroxyethylthiazole kinase

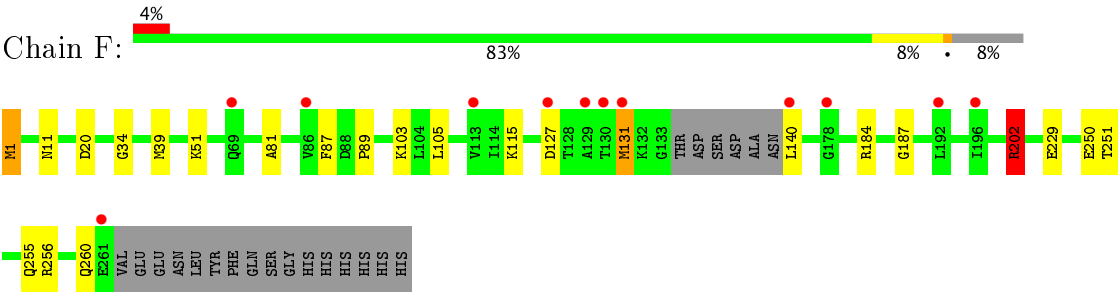


SER  
GLY  
HIS  
HIS  
HIS  
HIS  
HIS

• Molecule 1: Hydroxyethylthiazole kinase



• Molecule 1: Hydroxyethylthiazole kinase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.34Å 62.48Å 109.17Å 92.64° 92.05° 101.46°	Depositor
Resolution (Å)	30.00 – 1.62 29.86 – 1.62	Depositor EDS
% Data completeness (in resolution range)	93.2 (30.00-1.62) 91.7 (29.86-1.62)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.00 (at 1.62Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.182 , 0.203 0.191 , 0.208	Depositor DCC
$R_{free}$ test set	9508 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.5	Xtriage
Anisotropy	0.562	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.025 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12018	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% 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: MG, 51F

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	1.03	4/1925 (0.2%)	1.17	14/2620 (0.5%)
1	B	1.22	3/1979 (0.2%)	1.06	9/2693 (0.3%)
1	C	1.29	6/1902 (0.3%)	1.18	11/2587 (0.4%)
1	D	1.09	7/1841 (0.4%)	0.89	4/2506 (0.2%)
1	E	0.80	2/2032 (0.1%)	1.01	8/2764 (0.3%)
1	F	0.81	2/1921 (0.1%)	0.89	9/2613 (0.3%)
All	All	1.06	24/11600 (0.2%)	1.04	55/15783 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
All	All	0	6

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	261	GLU	CG-CD	-30.14	1.06	1.51
1	D	184	ARG	CD-NE	-26.40	1.01	1.46
1	C	51	LYS	CE-NZ	-22.25	0.93	1.49
1	F	140	LEU	CB-CG	-17.37	1.02	1.52
1	C	47	GLU	CD-OE1	-17.07	1.06	1.25
1	D	148	LYS	CG-CD	14.62	2.02	1.52
1	C	184	ARG	CZ-NH2	-14.45	1.14	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	107	LYS	CG-CD	12.99	1.96	1.52
1	C	47	GLU	CD-OE2	12.63	1.39	1.25
1	E	51	LYS	CD-CE	-11.87	1.21	1.51
1	D	163	GLU	CG-CD	-11.55	1.34	1.51
1	A	184	ARG	CZ-NH2	11.46	1.48	1.33
1	B	66	GLN	CB-CG	-10.57	1.24	1.52
1	A	184	ARG	CD-NE	-10.53	1.28	1.46
1	F	229	GLU	CB-CG	-9.73	1.33	1.52
1	D	142	ALA	CA-CB	-9.63	1.32	1.52
1	D	229	GLU	CB-CG	-8.06	1.36	1.52
1	D	66	GLN	CB-CG	-7.61	1.32	1.52
1	D	162	LYS	CA-CB	-7.40	1.37	1.53
1	E	52	VAL	CA-CB	-6.61	1.40	1.54
1	C	229	GLU	CG-CD	6.37	1.61	1.51
1	B	229	GLU	CG-CD	5.74	1.60	1.51
1	A	162	LYS	CG-CD	-5.73	1.32	1.52
1	C	103	LYS	CD-CE	-5.62	1.37	1.51

All (55) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	184	ARG	NE-CZ-NH2	-22.42	109.09	120.30
1	A	184	ARG	NE-CZ-NH1	20.78	130.69	120.30
1	C	51	LYS	CD-CE-NZ	19.72	157.05	111.70
1	E	258	ARG	NE-CZ-NH1	19.37	129.99	120.30
1	E	258	ARG	NE-CZ-NH2	-17.98	111.31	120.30
1	D	184	ARG	CG-CD-NE	17.33	148.20	111.80
1	A	184	ARG	NE-CZ-NH2	-13.95	113.33	120.30
1	E	51	LYS	CG-CD-CE	12.83	150.40	111.90
1	A	107	LYS	CB-CG-CD	-12.49	79.13	111.60
1	A	256	ARG	NE-CZ-NH2	-12.14	114.23	120.30
1	A	184	ARG	NH1-CZ-NH2	-12.10	106.09	119.40
1	F	140	LEU	CA-CB-CG	12.05	143.03	115.30
1	B	256	ARG	NE-CZ-NH1	11.08	125.84	120.30
1	A	256	ARG	NE-CZ-NH1	11.02	125.81	120.30
1	C	20	ASP	CB-CG-OD2	-10.75	108.62	118.30
1	F	140	LEU	CB-CG-CD2	10.66	129.12	111.00
1	B	261	GLU	CB-CG-CD	10.38	142.22	114.20
1	D	184	ARG	CD-NE-CZ	10.36	138.10	123.60
1	B	256	ARG	NE-CZ-NH2	-10.07	115.26	120.30
1	C	256	ARG	NE-CZ-NH2	-9.96	115.32	120.30
1	E	51	LYS	CD-CE-NZ	9.61	133.79	111.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	80	GLU	OE1-CD-OE2	9.47	134.67	123.30
1	D	256	ARG	NE-CZ-NH2	-8.72	115.94	120.30
1	D	256	ARG	NE-CZ-NH1	8.51	124.55	120.30
1	F	256	ARG	NE-CZ-NH2	-8.38	116.11	120.30
1	A	51	LYS	CG-CD-CE	8.24	136.62	111.90
1	A	259	ILE	CB-CA-C	-8.08	95.44	111.60
1	B	260	GLN	CA-CB-CG	7.92	130.83	113.40
1	C	256	ARG	NE-CZ-NH1	7.87	124.23	120.30
1	A	162	LYS	CB-CG-CD	7.76	131.78	111.60
1	C	184	ARG	NE-CZ-NH1	7.64	124.12	120.30
1	B	207	ASP	CB-CG-OD1	7.51	125.06	118.30
1	F	131	MET	N-CA-CB	-7.30	97.47	110.60
1	C	207	ASP	CB-CG-OD1	7.22	124.80	118.30
1	A	107	LYS	CG-CD-CE	6.89	132.56	111.90
1	F	256	ARG	NE-CZ-NH1	6.85	123.72	120.30
1	A	100	GLN	OE1-CD-NE2	-6.62	106.68	121.90
1	E	256	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	B	128	THR	N-CA-CB	-6.51	97.94	110.30
1	A	51	LYS	CB-CG-CD	6.30	127.99	111.60
1	B	186	THR	C-N-CA	-5.62	110.50	122.30
1	C	47	GLU	CG-CD-OE2	-5.50	107.31	118.30
1	B	25	PHE	CB-CG-CD2	-5.50	116.95	120.80
1	E	256	ARG	NE-CZ-NH1	5.44	123.02	120.30
1	E	184	ARG	CB-CG-CD	-5.44	97.47	111.60
1	E	162	LYS	CG-CD-CE	5.41	128.13	111.90
1	A	207	ASP	CB-CG-OD1	5.41	123.16	118.30
1	F	184	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	C	243	ASP	CB-CG-OD1	5.17	122.95	118.30
1	C	173	ILE	CG1-CB-CG2	5.15	122.73	111.40
1	F	20	ASP	CB-CG-OD1	5.14	122.93	118.30
1	A	171	LYS	CB-CG-CD	-5.12	98.29	111.60
1	F	202	ARG	CG-CD-NE	-5.11	101.07	111.80
1	F	1	MET	CG-SD-CE	5.09	108.34	100.20
1	C	47	GLU	CG-CD-OE1	5.02	128.34	118.30

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	GLN	Sidechain
1	A	184	ARG	Sidechain
1	C	20	ASP	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	D	258	ARG	Peptide
1	E	175	LEU	Mainchain
1	F	260	GLN	Peptide

## 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	1901	0	1933	14	0
1	B	1954	0	1991	13	0
1	C	1878	0	1921	8	0
1	D	1817	0	1858	7	0
1	E	2006	0	2029	6	0
1	F	1897	0	1930	9	0
2	A	9	0	0	0	0
2	B	9	0	0	0	0
2	C	9	0	0	0	0
2	D	18	0	0	0	0
2	E	9	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	E	1	0	0	0	0
4	A	105	0	0	4	0
4	B	129	0	0	2	0
4	C	145	0	0	1	0
4	D	30	0	0	1	0
4	E	58	0	0	0	0
4	F	39	0	0	1	0
All	All	12018	0	11662	57	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 2.

All (57) 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:D:53:ALA:O	1:D:202:ARG:NH1	1.88	1.05
1:A:76:GLN:HG3	4:A:462:HOH:O	1.70	0.91
1:A:128:THR:O	1:A:130:THR:HB	1.86	0.75
4:A:474:HOH:O	1:C:184:ARG:HD2	1.90	0.71
1:D:83:LEU:HD13	1:D:202:ARG:NH1	2.13	0.64
1:C:173:ILE:HD13	1:C:261:GLU:HA	1.82	0.61
1:A:175:LEU:CD2	1:A:259:ILE:HD12	2.31	0.60
1:D:83:LEU:HD13	1:D:202:ARG:HH12	1.66	0.60
1:C:187:GLY:HA2	4:C:501:HOH:O	2.04	0.58
1:A:187:GLY:HA2	4:A:484:HOH:O	2.06	0.55
1:F:103:LYS:HE3	1:F:127:ASP:OD2	2.06	0.55
1:A:48:GLU:O	1:A:51:LYS:HG2	2.07	0.54
1:F:81:ALA:O	1:F:202:ARG:NH2	2.40	0.52
1:A:51:LYS:HG3	1:A:52:VAL:HG13	1.90	0.52
1:D:227:GLU:OE1	1:D:256:ARG:NH2	2.43	0.51
1:E:54:GLN:HE21	1:E:201:PHE:HA	1.75	0.51
1:B:128:THR:HG22	1:B:130:THR:H	1.74	0.51
1:B:87:PHE:CE1	1:B:105:LEU:HD23	2.47	0.50
1:B:261:GLU:HG3	1:B:262:VAL:HG23	1.95	0.48
1:E:119:SER:N	1:E:135:ASP:OD1	2.37	0.48
1:A:175:LEU:HD23	1:A:259:ILE:HD12	1.95	0.48
1:B:128:THR:CG2	1:B:130:THR:OG1	2.63	0.47
1:F:103:LYS:CE	1:F:127:ASP:OD2	2.62	0.47
1:D:187:GLY:HA2	4:D:421:HOH:O	2.14	0.46
1:F:187:GLY:HA2	4:F:335:HOH:O	2.16	0.46
1:B:25:PHE:HD2	1:B:187:GLY:HA3	1.81	0.46
1:A:87:PHE:CE1	1:A:105:LEU:HD23	2.51	0.46
1:F:87:PHE:CE1	1:F:105:LEU:HD23	2.53	0.44
1:A:259:ILE:HG22	1:A:259:ILE:O	2.09	0.44
1:F:251:THR:HG22	1:F:255:GLN:NE2	2.32	0.44
1:B:258:ARG:NH1	4:B:407:HOH:O	2.49	0.44
1:D:89:PRO:HD2	1:D:115:LYS:O	2.18	0.44
1:E:1:MET:CE	1:E:5:ASN:OD1	2.66	0.43
1:C:89:PRO:HD2	1:C:115:LYS:O	2.17	0.43
1:A:89:PRO:HD2	1:A:115:LYS:O	2.18	0.43
1:B:89:PRO:HD2	1:B:115:LYS:O	2.19	0.43
1:F:89:PRO:HD2	1:F:115:LYS:O	2.19	0.43
1:F:1:MET:HG3	1:F:250:GLU:HA	2.00	0.43
1:B:128:THR:HG23	1:B:130:THR:HG23	2.01	0.43
1:C:103:LYS:O	1:C:107:LYS:HG2	2.19	0.43
1:A:227:GLU:OE1	1:A:256:ARG:NH2	2.53	0.42
1:E:238:SER:HB3	1:E:239:PRO:HD3	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:129:ALA:O	1:A:130:THR:HG22	2.19	0.42
1:B:160:THR:HA	1:B:164:ASP:OD1	2.19	0.42
1:E:89:PRO:HD2	1:E:115:LYS:O	2.19	0.42
1:B:173:ILE:CD1	1:B:175:LEU:HG	2.50	0.42
1:B:69:GLN:HG2	4:B:423:HOH:O	2.19	0.42
4:A:474:HOH:O	1:C:184:ARG:CD	2.58	0.42
1:D:238:SER:HB3	1:D:239:PRO:HD3	2.02	0.41
1:B:11:ASN:HA	1:B:34:GLY:O	2.20	0.41
1:C:87:PHE:CE1	1:C:105:LEU:HD23	2.55	0.41
1:A:238:SER:HB3	1:A:239:PRO:HD3	2.01	0.41
1:A:170:ASP:O	1:A:264:GLU:HB3	2.20	0.41
1:B:261:GLU:HG3	1:B:262:VAL:N	2.36	0.41
1:F:11:ASN:HA	1:F:34:GLY:O	2.21	0.41
1:C:11:ASN:HA	1:C:34:GLY:O	2.21	0.41
1:E:87:PHE:CE1	1:E:105:LEU:HD23	2.56	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	251/277 (91%)	249 (99%)	2 (1%)	0	100	100
1	B	260/277 (94%)	259 (100%)	1 (0%)	0	100	100
1	C	246/277 (89%)	242 (98%)	4 (2%)	0	100	100
1	D	241/277 (87%)	240 (100%)	1 (0%)	0	100	100
1	E	263/277 (95%)	261 (99%)	2 (1%)	0	100	100
1	F	251/277 (91%)	249 (99%)	2 (1%)	0	100	100
All	All	1512/1662 (91%)	1500 (99%)	12 (1%)	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	200/223 (90%)	197 (98%)	3 (2%)	70	47
1	B	207/223 (93%)	204 (99%)	3 (1%)	71	50
1	C	200/223 (90%)	199 (100%)	1 (0%)	91	83
1	D	189/223 (85%)	185 (98%)	4 (2%)	59	31
1	E	213/223 (96%)	210 (99%)	3 (1%)	71	50
1	F	199/223 (89%)	195 (98%)	4 (2%)	60	33
All	All	1208/1338 (90%)	1190 (98%)	18 (2%)	70	47

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	MET
1	A	126	ASP
1	A	130	THR
1	B	39	MET
1	B	128	THR
1	B	140	LEU
1	C	39	MET
1	E	39	MET
1	E	107	LYS
1	E	162	LYS
1	D	39	MET
1	D	69	GLN
1	D	173	ILE
1	D	202	ARG
1	F	39	MET
1	F	51	LYS
1	F	131	MET
1	F	202	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	6	ASN
1	E	11	ASN
1	E	54	GLN
1	D	67	ASN
1	F	67	ASN
1	F	255	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 5 are monoatomic - leaving 6 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	51F	A	301	-	8,9,9	0.80	0	6,11,11	3.95	1 (16%)
2	51F	B	301	-	8,9,9	1.40	2 (25%)	6,11,11	2.37	2 (33%)
2	51F	C	301	-	8,9,9	2.79	4 (50%)	6,11,11	4.16	4 (66%)
2	51F	D	301	-	8,9,9	1.26	1 (12%)	6,11,11	1.11	0
2	51F	D	302	-	8,9,9	1.12	1 (12%)	6,11,11	0.60	0
2	51F	E	301	-	8,9,9	0.67	0	6,11,11	1.45	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	51F	A	301	-	-	0/3/3/3	0/1/1/1
2	51F	B	301	-	-	0/3/3/3	0/1/1/1
2	51F	C	301	-	-	0/3/3/3	0/1/1/1
2	51F	D	301	-	-	0/3/3/3	0/1/1/1
2	51F	D	302	-	-	0/3/3/3	0/1/1/1
2	51F	E	301	-	-	0/3/3/3	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	51F	CAD-NAI	-4.49	1.32	1.38
2	D	301	51F	CAD-NAI	-3.15	1.34	1.38
2	D	302	51F	CAD-NAI	-3.06	1.34	1.38
2	B	301	51F	CAD-NAI	-2.31	1.35	1.38
2	C	301	51F	CAD-CAC	2.02	1.44	1.36
2	C	301	51F	CAH-NAG	2.15	1.40	1.35
2	B	301	51F	CAF-NAI	2.51	1.51	1.48
2	C	301	51F	CAF-NAI	5.40	1.56	1.48

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	51F	CAD-NAI-CAH	-9.47	100.00	108.64
2	B	301	51F	CAD-NAI-CAH	-4.77	104.29	108.64
2	C	301	51F	CAD-CAC-NAG	-2.75	102.42	109.38
2	C	301	51F	CAF-NAI-CAD	-2.26	118.94	123.55
2	B	301	51F	OAB-CAE-CAF	-2.19	102.50	110.59
2	C	301	51F	CAA-CAH-NAG	4.51	134.41	121.61
2	C	301	51F	CAD-NAI-CAH	8.39	116.29	108.64

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	255/277 (92%)	-0.13	7 (2%)	55 53	15, 20, 42, 71	14 (5%)
1	B	262/277 (94%)	-0.16	10 (3%)	41 38	14, 18, 34, 69	16 (6%)
1	C	250/277 (90%)	-0.26	3 (1%)	79 79	14, 18, 31, 71	19 (7%)
1	D	245/277 (88%)	0.52	16 (6%)	20 18	24, 43, 67, 85	13 (5%)
1	E	267/277 (96%)	0.12	15 (5%)	25 23	18, 30, 54, 67	10 (3%)
1	F	255/277 (92%)	0.29	12 (4%)	32 30	22, 40, 62, 76	8 (3%)
All	All	1534/1662 (92%)	0.06	63 (4%)	38 35	14, 27, 58, 85	80 (5%)

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	259	ILE	6.0
1	A	130	THR	4.7
1	E	130	THR	4.7
1	F	127	ASP	4.7
1	E	128	THR	4.5
1	E	134	THR	4.5
1	B	138	ALA	4.3
1	B	136	SER	4.2
1	B	262	VAL	4.1
1	C	140	LEU	3.8
1	A	260	GLN	3.8
1	F	129	ALA	3.7
1	A	262	VAL	3.4
1	B	137	ASP	3.3
1	F	140	LEU	3.3
1	E	1	MET	3.2
1	F	130	THR	3.2
1	D	69	GLN	3.2
1	D	196	ILE	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	133	GLY	3.0
1	E	129	ALA	3.0
1	E	126	ASP	2.9
1	E	127	ASP	2.9
1	B	140	LEU	2.9
1	A	126	ASP	2.9
1	C	127	ASP	2.8
1	F	196	ILE	2.7
1	F	69	GLN	2.7
1	F	192	LEU	2.7
1	A	128	THR	2.7
1	D	254	GLN	2.6
1	E	136	SER	2.6
1	E	140	LEU	2.6
1	F	261	GLU	2.5
1	D	192	LEU	2.4
1	D	141	ASP	2.4
1	D	143	VAL	2.4
1	D	1	MET	2.3
1	F	86	VAL	2.3
1	B	129	ALA	2.3
1	E	49	PHE	2.3
1	B	113	VAL	2.3
1	B	86	VAL	2.2
1	A	140	LEU	2.2
1	D	150	TYR	2.2
1	F	131	MET	2.2
1	D	195	ILE	2.2
1	D	144	THR	2.2
1	D	100	GLN	2.2
1	E	139	ASN	2.1
1	F	113	VAL	2.1
1	B	261	GLU	2.1
1	E	132	LYS	2.1
1	B	170	ASP	2.1
1	E	195	ILE	2.1
1	A	263	GLU	2.1
1	D	122	LEU	2.1
1	D	108	SER	2.1
1	E	255	GLN	2.0
1	F	178	GLY	2.0
1	C	262	VAL	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	57	LEU	2.0
1	D	104	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. 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	51F	E	301	9/9	0.90	0.11	0.81	26,30,38,50	0
3	MG	C	302	1/1	0.95	0.09	0.51	45,45,45,45	0
2	51F	B	301	9/9	0.95	0.08	0.35	14,16,24,34	0
2	51F	A	301	9/9	0.97	0.07	-0.07	11,13,25,34	0
2	51F	C	301	9/9	0.95	0.09	-0.08	11,13,17,29	0
3	MG	E	302	1/1	0.92	0.09	-0.46	46,46,46,46	0
2	51F	D	302	9/9	0.95	0.10	-0.74	24,30,42,48	0
3	MG	B	302	1/1	0.94	0.06	-1.28	45,45,45,45	0
2	51F	D	301	9/9	0.97	0.06	-1.47	23,25,35,50	0
3	MG	C	303	1/1	0.99	0.06	-1.82	27,27,27,27	0
3	MG	A	302	1/1	0.96	0.06	-	52,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.