



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 02:58 PM GMT

PDB ID : 2X4Y  
Title : MOLECULAR BASIS OF HISTONE H3K36ME3 RECOGNITION BY THE  
PWWP DOMAIN OF BRPF1.  
Authors : Vezzoli, A.; Bonadies, N.; Allen, M.D.; Freund, S.M.V.; Santiveri, C.M.; Kvinn-  
laug, B.; Huntly, B.J.P.; Gottgens, B.; Bycroft, M.  
Deposited on : 2010-02-02  
Resolution : 1.70 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

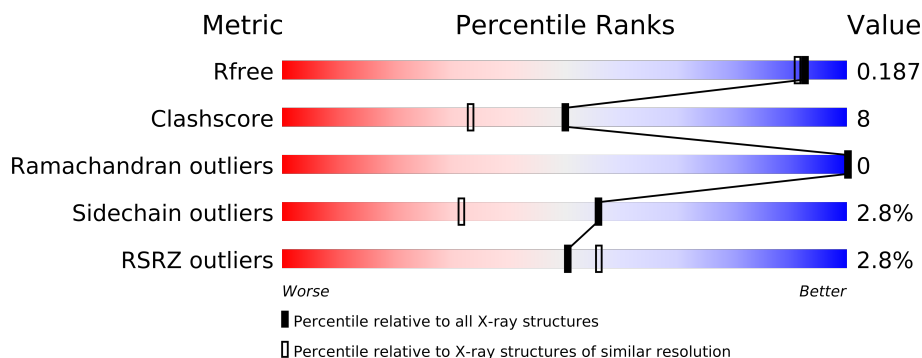
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.70 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.




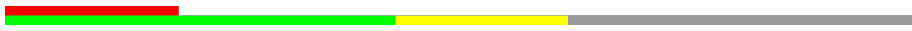
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	132	
1	C	132	
1	E	132	
1	G	132	
1	I	132	
1	K	132	
1	M	132	
1	O	132	
2	B	21	
2	D	21	
2	F	21	
2	H	21	
2	J	21	
2	L	21	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	N	21	
2	P	21	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10140 atoms, of which 0 are hydrogen and 0 are deuterium.

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 PEREGRIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	127	Total	C	N	O	S	0	3	0
			1059	679	193	182	5			
1	C	125	Total	C	N	O	S	0	0	0
			1023	658	185	175	5			
1	E	126	Total	C	N	O	S	0	0	0
			1033	663	186	179	5			
1	G	124	Total	C	N	O	S	0	0	0
			1019	656	184	174	5			
1	I	125	Total	C	N	O	S	0	0	0
			1023	658	185	175	5			
1	K	126	Total	C	N	O	S	0	3	0
			1050	674	188	183	5			
1	M	126	Total	C	N	O	S	0	0	0
			1033	663	186	179	5			
1	O	123	Total	C	N	O	S	0	0	0
			1010	651	182	172	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1074	GLY	-	EXPRESSION TAG	UNP P55201
A	1075	GLY	-	EXPRESSION TAG	UNP P55201
C	1074	GLY	-	EXPRESSION TAG	UNP P55201
C	1075	GLY	-	EXPRESSION TAG	UNP P55201
E	1074	GLY	-	EXPRESSION TAG	UNP P55201
E	1075	GLY	-	EXPRESSION TAG	UNP P55201
G	1074	GLY	-	EXPRESSION TAG	UNP P55201
G	1075	GLY	-	EXPRESSION TAG	UNP P55201
I	1074	GLY	-	EXPRESSION TAG	UNP P55201
I	1075	GLY	-	EXPRESSION TAG	UNP P55201
K	1074	GLY	-	EXPRESSION TAG	UNP P55201
K	1075	GLY	-	EXPRESSION TAG	UNP P55201
M	1074	GLY	-	EXPRESSION TAG	UNP P55201

*Continued on next page...*

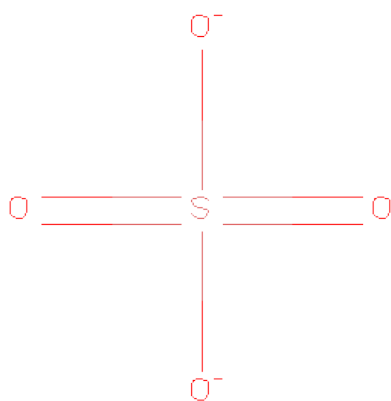
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
M	1075	GLY	-	EXPRESSION TAG	UNP P55201
O	1074	GLY	-	EXPRESSION TAG	UNP P55201
O	1075	GLY	-	EXPRESSION TAG	UNP P55201

- Molecule 2 is a protein called HISTONE H3.2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	18	Total	C	N	O	0	0	0
			136	86	29	21			
2	D	14	Total	C	N	O	0	0	0
			103	65	22	16			
2	F	15	Total	C	N	O	0	0	0
			108	68	23	17			
2	H	11	Total	C	N	O	0	0	0
			73	47	13	13			
2	J	12	Total	C	N	O	0	0	0
			83	53	16	14			
2	L	11	Total	C	N	O	0	0	0
			77	50	15	12			
2	N	14	Total	C	N	O	0	0	0
			106	68	21	17			
2	P	13	Total	C	N	O	0	0	0
			91	57	19	15			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	E	1	Total O S 5 4 1	0	0
3	G	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	154	Total O 154 154	0	0
4	B	24	Total O 24 24	0	0
4	C	113	Total O 113 113	0	0
4	D	14	Total O 14 14	0	0
4	E	166	Total O 166 166	0	0
4	F	19	Total O 19 19	0	0
4	G	140	Total O 140 140	0	0
4	H	9	Total O 9 9	0	0
4	I	119	Total O 119 119	0	0
4	J	12	Total O 12 12	0	0
4	K	123	Total O 123 123	0	0
4	L	6	Total O 6 6	0	0
4	M	108	Total O 108 108	0	0

*Continued on next page...*

*Continued from previous page...*

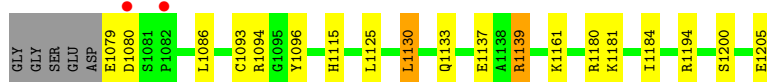
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	N	6	Total 6	O 6	0	0
4	O	62	Total 62	O 62	0	0
4	P	3	Total 3	O 3	0	0

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

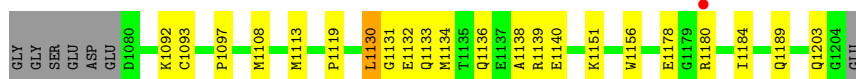
#### • Molecule 1: PEREGRIN

Chain A: 



#### • Molecule 1: PEREGRIN

Chain C: 



#### • Molecule 1: PEREGRIN

Chain E: 



#### • Molecule 1: PEREGRIN

Chain G: 



#### • Molecule 1: PEREGRIN

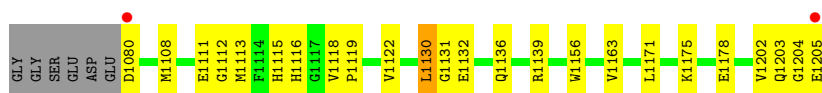
Chain I: 



#### • Molecule 1: PEREGRIN

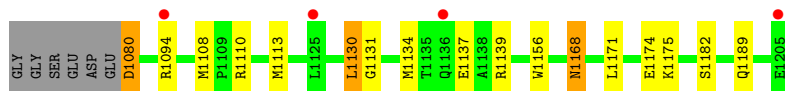
Chain K: 





- Molecule 1: PEREGRIN

Chain M:



- Molecule 1: PEREGRIN

Chain O:



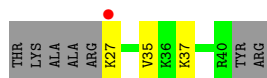
- Molecule 2: HISTONE H3.2

Chain B:



- Molecule 2: HISTONE H3.2

Chain D:



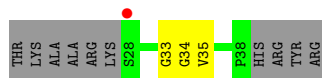
- Molecule 2: HISTONE H3.2

Chain F:



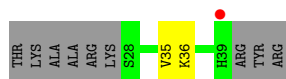
- Molecule 2: HISTONE H3.2

Chain H:



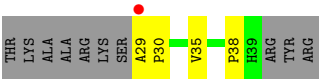
- Molecule 2: HISTONE H3.2

Chain J:



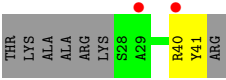
- Molecule 2: HISTONE H3.2

Chain L:



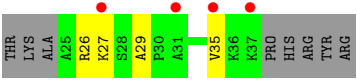
● Molecule 2: HISTONE H3.2

Chain N:



● Molecule 2: HISTONE H3.2

Chain P:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	35.76Å 97.91Å 102.52Å 86.80° 80.96° 87.24°	Depositor
Resolution (Å)	24.77 – 1.70 48.84 – 1.70	Depositor EDS
% Data completeness (in resolution range)	88.0 (24.77-1.70) 94.2 (48.84-1.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 1.70Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.190 , 0.227 0.186 , 0.187	Depositor DCC
$R_{free}$ test set	7136 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.1	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 47.3	EDS
Estimated twinning fraction	0.016 for -h,-k,-h+l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 141891 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10140	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 5.04% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: M3L, SO4

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	0.92	1/1099 (0.1%)	0.91	0/1485
1	C	0.72	0/1051	0.81	0/1423
1	E	0.79	0/1061	0.84	1/1435 (0.1%)
1	G	0.74	0/1047	0.80	0/1418
1	I	0.71	1/1051 (0.1%)	0.76	0/1423
1	K	0.72	0/1086	0.78	0/1469
1	M	0.70	0/1061	0.75	0/1435
1	O	0.58	0/1038	0.70	0/1406
2	B	0.79	0/127	0.96	0/170
2	D	0.63	0/93	0.87	0/124
2	F	0.74	0/98	0.79	0/131
2	H	0.67	0/62	0.85	0/84
2	J	0.57	0/73	0.62	0/99
2	L	0.50	0/67	0.63	0/91
2	N	0.50	0/97	0.56	0/131
2	P	0.52	0/79	0.70	0/104
All	All	0.73	2/9190 (0.0%)	0.79	1/12428 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1093	CYS	CB-SG	-11.65	1.62	1.82
1	I	1093	CYS	CB-SG	-5.79	1.72	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	1172	ASP	CB-CG-OD1	-5.51	113.34	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1059	0	1080	16	0
1	C	1023	0	1045	22	0
1	E	1033	0	1051	11	0
1	G	1019	0	1042	11	0
1	I	1023	0	1045	11	0
1	K	1050	0	1065	20	0
1	M	1033	0	1051	15	0
1	O	1010	0	1034	18	0
2	B	136	0	147	4	0
2	D	103	0	115	4	0
2	F	108	0	120	5	0
2	H	73	0	82	4	0
2	J	83	0	89	3	0
2	L	77	0	84	4	0
2	N	106	0	111	1	0
2	P	91	0	106	4	0
3	A	15	0	0	2	0
3	C	5	0	0	0	0
3	D	5	0	0	1	0
3	E	5	0	0	1	0
3	G	5	0	0	0	0
4	A	154	0	0	7	0
4	B	24	0	0	1	0
4	C	113	0	0	7	0
4	D	14	0	0	2	0
4	E	166	0	0	3	0
4	F	19	0	0	4	0
4	G	140	0	0	3	0
4	H	9	0	0	3	0
4	I	119	0	0	4	0
4	J	12	0	0	0	0
4	K	123	0	0	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	6	0	0	0	0
4	M	108	0	0	6	0
4	N	6	0	0	0	0
4	O	62	0	0	2	0
4	P	3	0	0	0	0
All	All	10140	0	9267	138	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

All (138) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:K:1130:LEU:HD12	2:L:35:VAL:HG21	1.41	1.00
1:C:1184:ILE:HD11	3:D:1041:SO4:O3	1.64	0.97
1:K:1115:HIS:O	1:K:1118:VAL:HG12	1.70	0.91
1:C:1203:GLN:HB3	4:C:2107:HOH:O	1.70	0.90
1:G:1130:LEU:HD12	2:H:35:VAL:HG21	1.55	0.87
1:O:1090:TRP:HB2	1:O:1163:VAL:HG13	1.58	0.83
1:A:1133:GLN:HB3	4:A:2050:HOH:O	1.79	0.81
1:K:1178[B]:GLU:HG2	4:K:2078:HOH:O	1.78	0.81
1:O:1163:VAL:HG11	1:O:1175:LYS:HE2	1.65	0.77
1:C:1133:GLN:HG3	4:C:2035:HOH:O	1.84	0.76
1:E:1080:ASP:N	1:E:1080:ASP:OD2	2.17	0.74
1:I:1130:LEU:HD12	2:J:35:VAL:HG21	1.68	0.73
1:G:1090:TRP:HB2	1:G:1163:VAL:HG13	1.69	0.72
1:C:1132:GLU:OE1	1:E:1107:LYS:HE2	1.90	0.70
1:K:1080:ASP:N	4:K:2002:HOH:O	2.27	0.67
1:I:1108:MET:HE3	1:I:1113:MET:HG2	1.75	0.67
1:K:1116:HIS:HD2	1:K:1202[A]:VAL:HG13	1.58	0.67
2:D:37:LYS:HE2	4:D:2009:HOH:O	1.94	0.67
1:A:1125:LEU:HD11	3:A:2206:SO4:O1	1.96	0.66
2:H:33:GLY:HA3	4:H:2006:HOH:O	1.96	0.66
1:A:1133:GLN:O	1:A:1137:GLU:HG2	1.96	0.65
2:B:40:ARG:HD2	4:B:2018:HOH:O	1.97	0.65
1:C:1136:GLN:HG3	4:C:2036:HOH:O	1.97	0.64
4:K:2017:HOH:O	2:L:38:PRO:HD2	1.98	0.64
2:H:33:GLY:CA	4:H:2006:HOH:O	2.45	0.63
1:C:1130:LEU:HD12	2:D:35:VAL:HG21	1.80	0.62
2:F:34:GLY:N	4:F:2013:HOH:O	1.97	0.62
1:M:1094:ARG:HG3	4:M:2017:HOH:O	2.00	0.62
2:F:25:ALA:N	4:F:2001:HOH:O	2.33	0.61
1:O:1097:PRO:HB2	4:O:2008:HOH:O	2.01	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:O:1080:ASP:N	1:O:1080:ASP:OD2	2.34	0.61
1:K:1116:HIS:CD2	1:K:1202[A]:VAL:HG13	2.36	0.60
1:M:1108:MET:CE	1:M:1113:MET:HG2	2.30	0.60
1:K:1115:HIS:CE1	1:K:1202[A]:VAL:HG11	2.36	0.60
1:K:1111:GLU:H	1:K:1111:GLU:CD	2.05	0.60
1:M:1168:ASN:ND2	1:M:1171:LEU:H	2.00	0.59
1:C:1189:GLN:NE2	4:C:2090:HOH:O	2.34	0.59
2:H:34:GLY:N	4:H:2006:HOH:O	2.08	0.59
1:A:1139:ARG:CZ	4:A:2052:HOH:O	2.50	0.59
1:I:1136:GLN:HG3	4:I:2040:HOH:O	2.03	0.59
1:G:1090:TRP:HB2	1:G:1163:VAL:CG1	2.31	0.59
1:G:1137:GLU:OE2	4:G:2052:HOH:O	2.17	0.58
1:C:1130:LEU:HD22	1:C:1134:MET:HG3	1.85	0.58
2:D:37:LYS:HG3	4:D:2010:HOH:O	2.03	0.58
1:A:1130:LEU:HD12	2:B:35:VAL:HG21	1.83	0.58
1:M:1175:LYS:NZ	4:M:2076:HOH:O	2.37	0.58
1:A:1086:LEU:HD13	1:A:1115:HIS:CD2	2.38	0.57
1:O:1097:PRO:HG3	1:O:1179:GLY:HA2	1.88	0.56
1:M:1189:GLN:NE2	4:M:2093:HOH:O	2.39	0.56
1:O:1119:PRO:HB2	2:P:29:ALA:HA	1.87	0.56
1:O:1130:LEU:HD12	2:P:35:VAL:HG21	1.88	0.56
1:E:1137:GLU:HG3	4:E:2049:HOH:O	2.05	0.56
1:O:1090:TRP:HB2	1:O:1163:VAL:CG1	2.32	0.55
1:O:1168:ASN:C	1:O:1168:ASN:HD22	2.10	0.55
1:O:1108:MET:HE3	1:O:1113:MET:HG2	1.89	0.55
1:G:1152:ARG:HD3	4:G:2069:HOH:O	2.06	0.54
1:M:1130:LEU:HD22	1:M:1130:LEU:O	2.06	0.54
1:O:1092:LYS:NZ	1:O:1093:CYS:O	2.40	0.54
1:K:1115:HIS:O	1:K:1118:VAL:CG1	2.51	0.53
1:K:1108:MET:HE3	1:K:1113:MET:HG2	1.90	0.52
1:I:1096:TYR:CD1	2:J:36:M3L:HM33	2.44	0.52
1:A:1079:GLU:O	1:A:1079:GLU:HG2	2.08	0.52
1:O:1131:GLY:HA3	1:O:1156:TRP:CE2	2.44	0.52
1:I:1189:GLN:NE2	4:I:2103:HOH:O	2.42	0.52
1:C:1133:GLN:HB3	4:C:2041:HOH:O	2.10	0.52
1:E:1137:GLU:CG	4:E:2049:HOH:O	2.57	0.51
2:F:33:GLY:HA3	4:F:2013:HOH:O	2.10	0.51
1:I:1108:MET:HE3	1:I:1113:MET:CG	2.41	0.51
1:I:1160:THR:HG23	4:I:2047:HOH:O	2.10	0.51
1:C:1108:MET:CE	1:C:1113:MET:HG2	2.41	0.50
1:C:1119:PRO:HD3	2:D:27:LYS:HD3	1.94	0.50
1:I:1181:LYS:HE3	4:I:2092:HOH:O	2.10	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:M:1168:ASN:HD22	1:M:1168:ASN:C	2.15	0.50
1:C:1130:LEU:HD22	1:C:1134:MET:CG	2.43	0.49
1:E:1108:MET:HE3	1:E:1113:MET:HG2	1.94	0.48
1:M:1080:ASP:OD2	1:M:1080:ASP:N	2.46	0.48
1:A:1139:ARG:NE	4:A:2052:HOH:O	2.46	0.48
1:G:1140:GLU:OE2	4:G:2054:HOH:O	2.20	0.48
2:N:40:ARG:HD2	2:N:41:TYR:CE2	2.49	0.48
1:A:1200:SER:OG	1:A:1205:GLU:HB2	2.13	0.48
2:F:33:GLY:CA	4:F:2013:HOH:O	2.61	0.48
1:C:1092:LYS:NZ	1:C:1093:CYS:O	2.47	0.48
1:K:1203:GLN:NE2	4:K:2122:HOH:O	2.47	0.47
1:K:1163:VAL:HG11	1:K:1171:LEU:HD11	1.96	0.47
1:A:1180:ARG:HG3	1:A:1181:LYS:N	2.29	0.47
1:M:1175:LYS:CE	4:M:2076:HOH:O	2.62	0.47
1:K:1132:GLU:O	1:K:1136:GLN:HG2	2.15	0.47
1:K:1204:GLY:O	1:K:1205:GLU:HB2	2.15	0.47
2:B:27:LYS:HG3	2:P:27:LYS:HG2	1.96	0.47
1:M:1130:LEU:HD22	1:M:1134:MET:HG3	1.98	0.46
1:K:1118:VAL:HG11	1:K:1202[B]:VAL:CG1	2.45	0.46
1:I:1108:MET:CE	1:I:1113:MET:HG2	2.43	0.46
1:E:1171:LEU:C	1:E:1171:LEU:HD23	2.36	0.46
1:A:1096:TYR:CD1	2:B:36:M3L:HM33	2.51	0.46
1:O:1102:LEU:HD23	1:O:1104:ILE:HD11	1.98	0.46
1:M:1175:LYS:HE3	4:M:2076:HOH:O	2.15	0.46
1:A:1184:ILE:HG23	4:A:2111:HOH:O	2.17	0.45
1:K:1131:GLY:HA3	1:K:1156:TRP:CE2	2.52	0.44
1:O:1081:SER:O	1:O:1159:ARG:NE	2.45	0.44
1:C:1108:MET:HE3	1:C:1113:MET:CG	2.47	0.44
1:M:1110:ARG:HB3	4:M:2029:HOH:O	2.18	0.44
1:I:1130:LEU:HD12	2:J:35:VAL:CG2	2.44	0.44
1:C:1138:ALA:HB1	1:C:1140:GLU:HG2	2.00	0.44
1:A:1094[A]:ARG:HD2	4:A:2021:HOH:O	2.17	0.44
1:C:1130:LEU:O	1:C:1134:MET:HG3	2.19	0.43
1:I:1080:ASP:OD1	1:I:1080:ASP:N	2.50	0.43
1:C:1097:PRO:HB2	4:C:2018:HOH:O	2.18	0.43
1:O:1144:LEU:HD12	1:O:1155:GLN:O	2.18	0.43
1:G:1171:LEU:HD23	1:G:1171:LEU:C	2.38	0.43
2:F:37:LYS:HA	2:F:38:PRO:HD3	1.88	0.43
1:K:1119:PRO:HB2	2:L:29:ALA:HA	2.01	0.42
1:C:1097:PRO:HA	1:C:1178:GLU:OE2	2.19	0.42
3:E:2206:SO4:O4	4:E:2165:HOH:O	2.19	0.42
1:K:1112:GLY:HA2	1:K:1122:VAL:CG2	2.50	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:G:1094:ARG:H	1:G:1094:ARG:HG2	1.61	0.42
1:O:1110:ARG:O	2:P:26:ARG:HD2	2.20	0.42
1:C:1108:MET:HE2	1:C:1113:MET:HG2	2.00	0.42
1:G:1108:MET:HE3	1:G:1113:MET:HG2	2.02	0.42
1:C:1133:GLN:CB	4:C:2041:HOH:O	2.68	0.42
1:G:1163:VAL:HG21	1:G:1171:LEU:HD11	2.01	0.42
2:L:29:ALA:HA	2:L:30:PRO:HD3	1.92	0.42
1:A:1094[A]:ARG:NH1	4:A:2019:HOH:O	2.48	0.42
1:A:1194:ARG:HD3	4:A:2059:HOH:O	2.18	0.42
1:E:1080:ASP:HB2	1:E:1081:SER:H	1.41	0.41
1:E:1190:ILE:HG23	1:M:1174:GLU:HG3	2.02	0.41
1:C:1151:LYS:HB2	1:C:1151:LYS:HE3	1.92	0.41
1:O:1189:GLN:NE2	4:O:2055:HOH:O	2.54	0.41
1:E:1108:MET:CE	1:E:1113:MET:HG2	2.50	0.41
1:G:1120:ILE:HA	1:G:1121:PRO:HD3	1.95	0.41
1:K:1175:LYS:O	1:K:1178[B]:GLU:HG3	2.21	0.41
1:M:1168:ASN:HD22	1:M:1171:LEU:H	1.68	0.41
1:O:1199:ARG:O	1:O:1202:VAL:HB	2.20	0.41
1:C:1131:GLY:HA3	1:C:1156:TRP:CE2	2.56	0.41
1:A:1161:LYS:NZ	3:A:2208:SO4:O2	2.52	0.40
1:M:1131:GLY:HA3	1:M:1156:TRP:CE2	2.57	0.40
1:E:1130:LEU:HD22	1:E:1134:MET:HG3	2.04	0.40
1:E:1151:LYS:HD3	1:E:1151:LYS:HA	1.95	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	128/132 (97%)	126 (98%)	2 (2%)	0	100	100
1	C	123/132 (93%)	121 (98%)	2 (2%)	0	100	100
1	E	124/132 (94%)	121 (98%)	3 (2%)	0	100	100
1	G	122/132 (92%)	120 (98%)	2 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	I	123/132 (93%)	121 (98%)	2 (2%)	0	100	100
1	K	127/132 (96%)	125 (98%)	2 (2%)	0	100	100
1	M	124/132 (94%)	121 (98%)	3 (2%)	0	100	100
1	O	121/132 (92%)	115 (95%)	6 (5%)	0	100	100
2	B	15/21 (71%)	15 (100%)	0	0	100	100
2	D	11/21 (52%)	11 (100%)	0	0	100	100
2	F	12/21 (57%)	12 (100%)	0	0	100	100
2	H	8/21 (38%)	8 (100%)	0	0	100	100
2	J	9/21 (43%)	9 (100%)	0	0	100	100
2	L	8/21 (38%)	8 (100%)	0	0	100	100
2	N	11/21 (52%)	11 (100%)	0	0	100	100
2	P	10/21 (48%)	10 (100%)	0	0	100	100
All	All	1076/1224 (88%)	1054 (98%)	22 (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	117/117 (100%)	114 (97%)	3 (3%)	59	35
1	C	112/117 (96%)	109 (97%)	3 (3%)	57	33
1	E	113/117 (97%)	110 (97%)	3 (3%)	57	33
1	G	112/117 (96%)	108 (96%)	4 (4%)	47	22
1	I	112/117 (96%)	109 (97%)	3 (3%)	57	33
1	K	116/117 (99%)	114 (98%)	2 (2%)	73	55
1	M	113/117 (97%)	107 (95%)	6 (5%)	32	11
1	O	111/117 (95%)	108 (97%)	3 (3%)	57	33
2	B	11/14 (79%)	11 (100%)	0	100	100
2	D	9/14 (64%)	9 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	9/14 (64%)	9 (100%)	0	100	100
2	H	6/14 (43%)	6 (100%)	0	100	100
2	J	7/14 (50%)	7 (100%)	0	100	100
2	L	6/14 (43%)	6 (100%)	0	100	100
2	N	9/14 (64%)	9 (100%)	0	100	100
2	P	7/14 (50%)	7 (100%)	0	100	100
All	All	970/1048 (93%)	943 (97%)	27 (3%)	56	32

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1080	ASP
1	A	1130	LEU
1	A	1139	ARG
1	C	1130	LEU
1	C	1139	ARG
1	C	1180	ARG
1	E	1080	ASP
1	E	1111	GLU
1	E	1130	LEU
1	G	1094	ARG
1	G	1130	LEU
1	G	1139	ARG
1	G	1169	GLN
1	I	1080	ASP
1	I	1130	LEU
1	I	1180	ARG
1	K	1130	LEU
1	K	1139	ARG
1	M	1080	ASP
1	M	1130	LEU
1	M	1137	GLU
1	M	1139	ARG
1	M	1168	ASN
1	M	1182	SER
1	O	1080	ASP
1	O	1130	LEU
1	O	1168	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (31) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	1133	GLN
1	A	1169	GLN
1	A	1189	GLN
1	A	1203	GLN
2	B	39	HIS
1	C	1116	HIS
1	C	1203	GLN
1	E	1133	GLN
1	E	1189	GLN
1	E	1197	GLN
1	E	1203	GLN
1	G	1116	HIS
1	G	1133	GLN
1	G	1203	GLN
1	I	1189	GLN
1	I	1203	GLN
1	K	1116	HIS
1	K	1136	GLN
1	K	1203	GLN
1	M	1115	HIS
1	M	1116	HIS
1	M	1133	GLN
1	M	1168	ASN
1	M	1189	GLN
1	M	1203	GLN
1	O	1115	HIS
1	O	1133	GLN
1	O	1136	GLN
1	O	1168	ASN
1	O	1169	GLN
1	O	1197	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

8 non-standard protein/DNA/RNA residues are modelled in this entry.

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	M3L	B	36	2	11,11,12	2.79	2 (18%)	12,14,16	1.36	1 (8%)
2	M3L	D	36	2	11,11,12	4.16	2 (18%)	12,14,16	0.73	0
2	M3L	F	36	2	11,11,12	3.64	4 (36%)	12,14,16	1.03	1 (8%)
2	M3L	H	36	2	11,11,12	4.49	2 (18%)	12,14,16	1.25	1 (8%)
2	M3L	J	36	2	11,11,12	4.08	2 (18%)	12,14,16	0.98	1 (8%)
2	M3L	L	36	2	11,11,12	4.43	3 (27%)	12,14,16	0.83	1 (8%)
2	M3L	N	36	2	11,11,12	3.46	2 (18%)	12,14,16	0.65	0
2	M3L	P	36	2	11,11,12	3.55	2 (18%)	12,14,16	1.04	2 (16%)

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	M3L	B	36	2	-	0/8/10/12	0/0/0/0
2	M3L	D	36	2	-	0/8/10/12	0/0/0/0
2	M3L	F	36	2	-	0/8/10/12	0/0/0/0
2	M3L	H	36	2	-	0/8/10/12	0/0/0/0
2	M3L	J	36	2	-	0/8/10/12	0/0/0/0
2	M3L	L	36	2	-	0/8/10/12	0/0/0/0
2	M3L	N	36	2	-	0/8/10/12	0/0/0/0
2	M3L	P	36	2	-	0/8/10/12	0/0/0/0

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	36	M3L	O-C	14.40	1.21	1.11
2	L	36	M3L	O-C	14.03	1.21	1.11
2	D	36	M3L	O-C	13.14	1.20	1.11
2	J	36	M3L	O-C	12.92	1.20	1.11
2	F	36	M3L	O-C	11.06	1.18	1.11
2	P	36	M3L	O-C	11.03	1.18	1.11
2	N	36	M3L	O-C	10.44	1.18	1.11

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	36	M3L	O-C	8.15	1.16	1.11
2	N	36	M3L	CA-C	3.39	1.54	1.48
2	B	36	M3L	CA-C	3.18	1.54	1.48
2	F	36	M3L	CA-C	3.10	1.54	1.48
2	P	36	M3L	CA-C	2.80	1.53	1.48
2	J	36	M3L	CA-C	2.70	1.53	1.48
2	L	36	M3L	CA-C	2.63	1.53	1.48
2	F	36	M3L	CM2-NZ	-2.33	1.43	1.50
2	H	36	M3L	CA-C	2.24	1.52	1.48
2	L	36	M3L	CE-NZ	-2.16	1.44	1.51
2	D	36	M3L	CA-C	2.08	1.52	1.48
2	F	36	M3L	CM1-NZ	-2.02	1.44	1.50

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	36	M3L	C-CA-N	-3.48	110.35	113.83
2	B	36	M3L	C-CA-N	-3.45	110.38	113.83
2	J	36	M3L	C-CA-N	-2.48	111.35	113.83
2	F	36	M3L	C-CA-N	-2.46	111.37	113.83
2	L	36	M3L	CM3-NZ-CM1	2.14	114.69	109.00
2	P	36	M3L	CG-CB-CA	2.10	120.45	113.13
2	P	36	M3L	CE-CD-CG	2.10	118.69	110.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

7 ligands are modelled in this entry.

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	SO4	A	2206	-	4,4,4	0.53	0	6,6,6	0.63	0
3	SO4	A	2207	-	4,4,4	0.31	0	6,6,6	0.55	0
3	SO4	A	2208	-	4,4,4	0.51	0	6,6,6	0.63	0
3	SO4	C	2205	-	4,4,4	0.23	0	6,6,6	0.19	0
3	SO4	D	1041	-	4,4,4	0.40	0	6,6,6	0.32	0
3	SO4	E	2206	-	4,4,4	0.26	0	6,6,6	0.53	0
3	SO4	G	2204	-	4,4,4	0.47	0	6,6,6	0.16	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
3	SO4	A	2206	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2207	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2208	-	-	0/0/0/0	0/0/0/0
3	SO4	C	2205	-	-	0/0/0/0	0/0/0/0
3	SO4	D	1041	-	-	0/0/0/0	0/0/0/0
3	SO4	E	2206	-	-	0/0/0/0	0/0/0/0
3	SO4	G	2204	-	-	0/0/0/0	0/0/0/0

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.

## 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	127/132 (96%)	-0.02	2 (1%) 68 74	12, 21, 40, 57	0
1	C	125/132 (94%)	-0.12	1 (0%) 83 88	15, 28, 43, 55	0
1	E	126/132 (95%)	-0.08	3 (2%) 56 61	13, 23, 41, 59	0
1	G	124/132 (93%)	-0.13	0 100 100	15, 25, 43, 50	0
1	I	125/132 (94%)	-0.14	2 (1%) 68 74	17, 28, 46, 53	0
1	K	126/132 (95%)	-0.09	2 (1%) 68 74	15, 26, 44, 56	0
1	M	126/132 (95%)	0.02	4 (3%) 45 50	16, 30, 49, 59	0
1	O	123/132 (93%)	0.22	5 (4%) 35 40	23, 35, 56, 64	0
2	B	18/21 (85%)	0.16	1 (5%) 24 27	20, 28, 40, 46	0
2	D	14/21 (66%)	0.57	1 (7%) 16 19	30, 36, 48, 52	0
2	F	15/21 (71%)	0.37	1 (6%) 17 21	22, 32, 41, 45	0
2	H	11/21 (52%)	0.43	1 (9%) 9 12	27, 32, 46, 47	0
2	J	12/21 (57%)	0.37	1 (8%) 11 14	28, 33, 49, 56	0
2	L	11/21 (52%)	0.52	1 (9%) 9 12	28, 36, 42, 45	0
2	N	14/21 (66%)	0.53	2 (14%) 3 5	29, 39, 46, 52	0
2	P	13/21 (61%)	1.26	4 (30%) 1 2	39, 43, 49, 60	0
All	All	1110/1224 (90%)	0.01	31 (2%) 50 56	12, 28, 48, 64	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	1082	PRO	5.4
1	M	1205	GLU	4.1
1	E	1205	GLU	4.0
1	O	1095	GLY	3.9
2	P	37	LYS	3.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	P	31	ALA	3.5
1	E	1204	GLY	3.5
1	K	1205	GLU	3.5
2	P	35	VAL	3.0
1	C	1180	ARG	3.0
2	B	24	ALA	3.0
2	D	27	LYS	2.9
1	M	1136	GLN	2.8
2	J	39	HIS	2.8
1	O	1096	TYR	2.7
1	I	1204	GLY	2.7
1	M	1125	LEU	2.6
1	M	1094	ARG	2.4
1	O	1094	ARG	2.4
2	F	25	ALA	2.4
2	N	40	ARG	2.4
2	L	29	ALA	2.4
1	I	1081	SER	2.3
2	N	29	ALA	2.3
2	H	28	SER	2.3
1	K	1080	ASP	2.3
2	P	27	LYS	2.2
1	E	1082	PRO	2.2
1	A	1080	ASP	2.1
1	O	1080	ASP	2.1
1	A	1082	PRO	2.0

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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	M3L	N	36	12/13	0.10	1.24	26,28,32,33	0
2	M3L	L	36	12/13	0.09	0.24	21,26,33,34	0
2	M3L	B	36	12/13	0.09	-0.01	15,20,23,25	0
2	M3L	P	36	12/13	0.14	-0.17	37,41,52,53	0
2	M3L	D	36	12/13	0.09	-0.18	24,29,33,34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	M3L	F	36	12/13	0.09	-0.23	18,21,24,28	0
2	M3L	J	36	12/13	0.07	-0.62	24,25,32,32	0
2	M3L	H	36	12/13	0.07	-0.66	20,26,32,32	0

### 6.3 Carbohydrates

There are no carbohydrates in this entry.

### 6.4 Ligands

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	2208	5/5	0.10	0.23	29,32,33,35	5
3	SO4	A	2207	5/5	0.09	-0.19	28,29,31,31	5
3	SO4	D	1041	5/5	0.12	-0.30	44,46,48,51	5
3	SO4	E	2206	5/5	0.08	-0.81	30,31,36,37	4
3	SO4	A	2206	5/5	0.09	-1.06	29,34,36,37	0
3	SO4	G	2204	5/5	0.07	-1.14	28,33,36,38	4
3	SO4	C	2205	5/5	0.07	-1.47	30,33,36,38	5

### 6.5 Other polymers

There are no such residues in this entry.