



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:30 AM GMT

PDB ID : 2XAD  
Title : Crystal structure of deacetylase-teicoplanin complex in biosynthesis pathway of teicoplanin  
Authors : Chan, H.C.; Huang, Y.T.; Lyu, S.Y.; Huang, C.J.; Li, Y.S.; Liu, Y.C.; Chou, C.C.; Tsai, M.D.; Li, T.L.  
Deposited on : 2010-03-31  
Resolution : 1.70 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

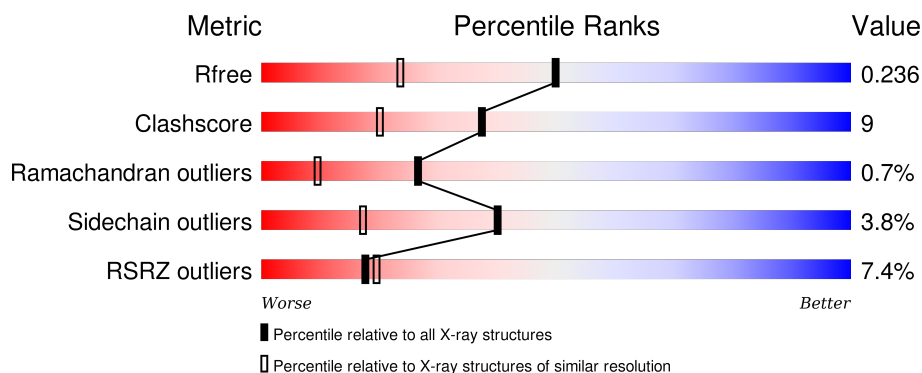
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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.




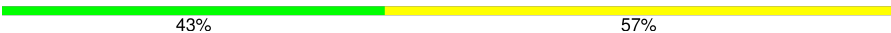

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3190 (1.70-1.70)
Clashscore	102246	3585 (1.70-1.70)
Ramachandran outliers	100387	3527 (1.70-1.70)
Sidechain outliers	100360	3527 (1.70-1.70)
RSRZ outliers	91569	3200 (1.70-1.70)

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. 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	273	<div> <div>6%</div> <div>77%</div> <div>16%</div> <div>• 5%</div> </div>
1	B	273	<div> <div>5%</div> <div>83%</div> <div>10%</div> <div>• 5%</div> </div>
1	C	273	<div> <div>12%</div> <div>72%</div> <div>19%</div> <div>• • 5%</div> </div>
1	D	273	<div> <div>5%</div> <div>81%</div> <div>13%</div> <div>• 6%</div> </div>
2	E	7	<div> <div>71%</div> <div>29%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	7	
2	G	7	
2	H	7	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	F	709	X	-	-	X
6	T55	H	711	-	-	X	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9379 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 N-ACYL GLM PEUDO-TEICOPLANIN DEACETYLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	259	Total	C	N	O	S	Se	0	0	0
			2018	1263	366	385	1	3			
1	B	258	Total	C	N	O	S	Se	0	0	1
			2006	1255	365	382	1	3			
1	C	258	Total	C	N	O	S	Se	0	0	1
			2008	1257	365	382	1	3			
1	D	257	Total	C	N	O	S	Se	0	0	0
			2006	1255	364	383	1	3			

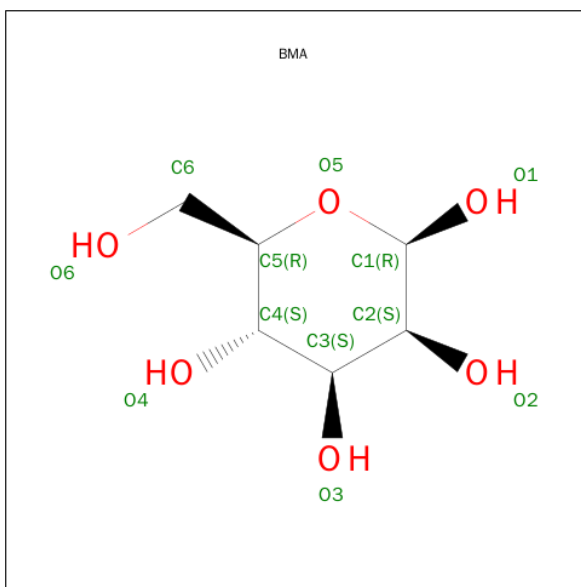
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	164	ASN	HIS	ENGINEERED MUTATION	UNP Q6ZZJ1
B	164	ASN	HIS	ENGINEERED MUTATION	UNP Q6ZZJ1
C	164	ASN	HIS	ENGINEERED MUTATION	UNP Q6ZZJ1
D	164	ASN	HIS	ENGINEERED MUTATION	UNP Q6ZZJ1

- Molecule 2 is a protein called TEICOPLANIN.

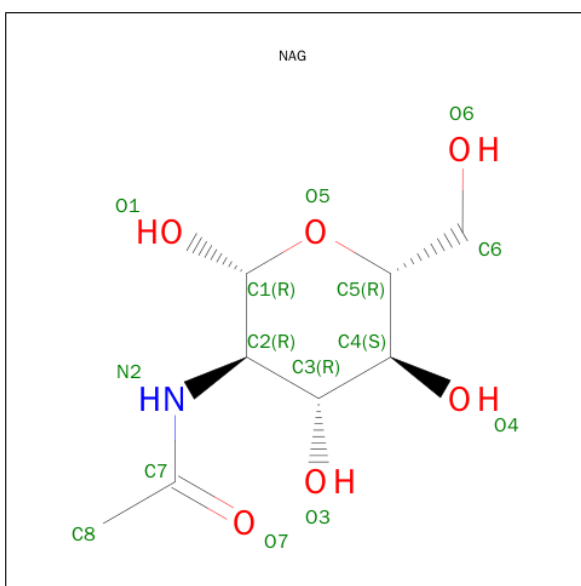
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	7	Total	C	Cl	N	O	0	0	0
			84	58	2	7	17			
2	F	7	Total	C	Cl	N	O	0	0	0
			85	58	2	7	18			
2	G	7	Total	C	Cl	N	O	0	0	0
			85	58	2	7	18			
2	H	7	Total	C	Cl	N	O	0	0	0
			85	58	2	7	18			

- Molecule 3 is SUGAR (BETA-D-MANNOSE) (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



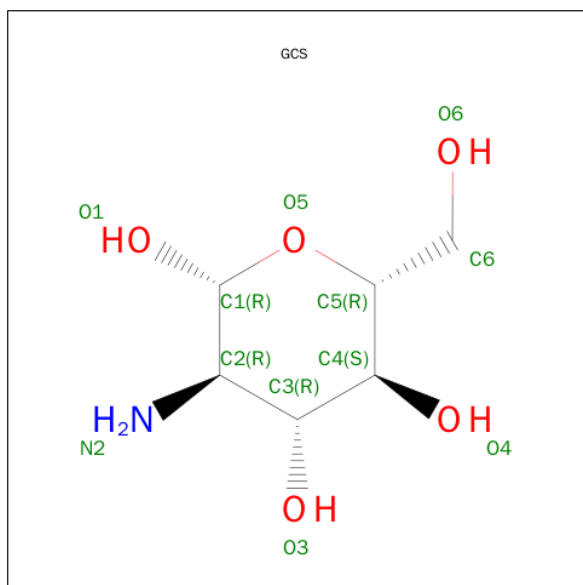
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			11	6	5		
3	F	1	Total	C	O	0	0
			11	6	5		
3	G	1	Total	C	O	0	0
			11	6	5		
3	H	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



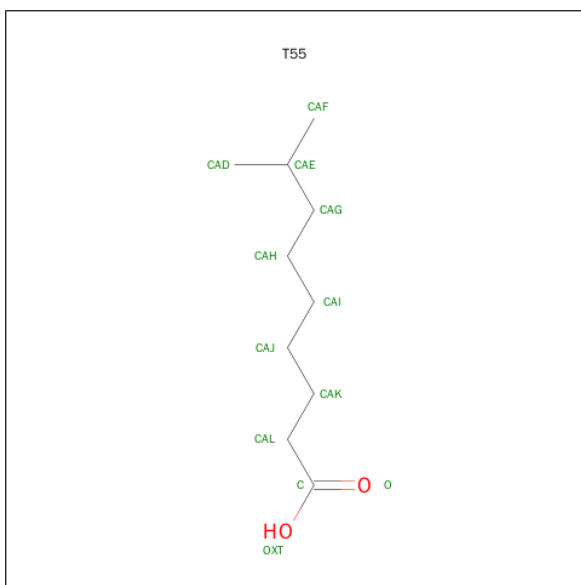
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	H	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is SUGAR (D-GLUCOSAMINE) (three-letter code: GCS) (formula:  $C_6H_{13}NO_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	E	1	Total	C	N	O	0	0
			11	6	1	4		
5	F	1	Total	C	N	O	0	0
			11	6	1	4		
5	G	1	Total	C	N	O	0	0
			11	6	1	4		
5	H	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 6 is 8-METHYLNONANOIC ACID (three-letter code: T55) (formula:  $C_{10}H_{20}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total	C	O	0	0
			11	10	1		
6	F	1	Total	C	O	0	0
			11	10	1		
6	G	1	Total	C	O	0	0
			11	10	1		
6	H	1	Total	C	O	0	0
			11	10	1		

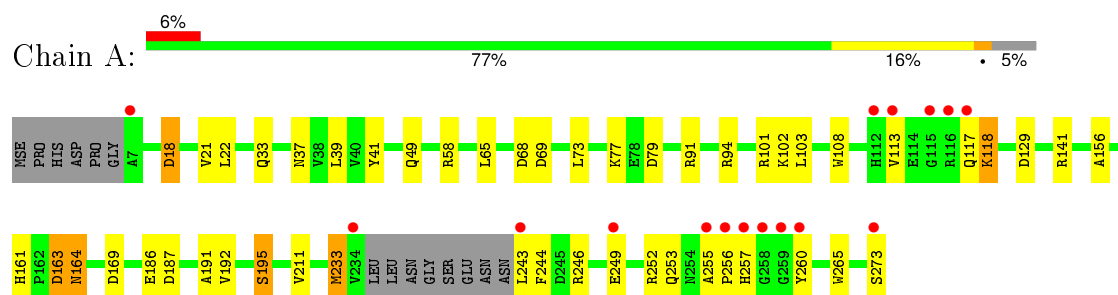
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	204	Total	O	0	0
			204	204		
7	B	200	Total	O	0	0
			200	200		
7	C	158	Total	O	0	0
			158	158		
7	D	221	Total	O	0	0
			221	221		
7	E	4	Total	O	0	0
			4	4		
7	F	13	Total	O	0	0
			13	13		
7	G	4	Total	O	0	0
			4	4		
7	H	10	Total	O	0	0
			10	10		

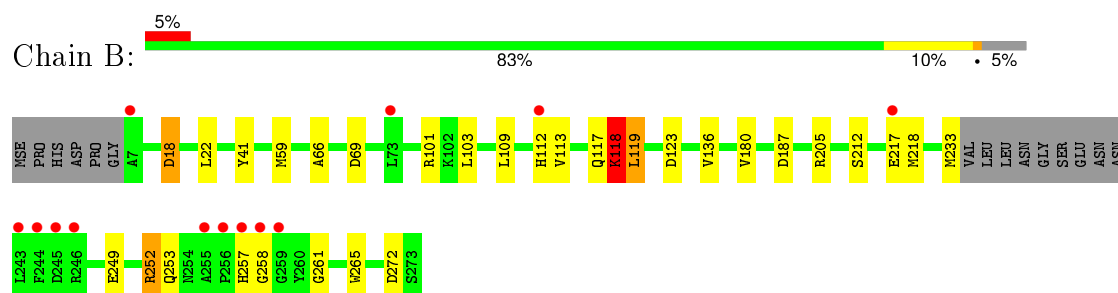
### 3 Residue-property plots [i](#)

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

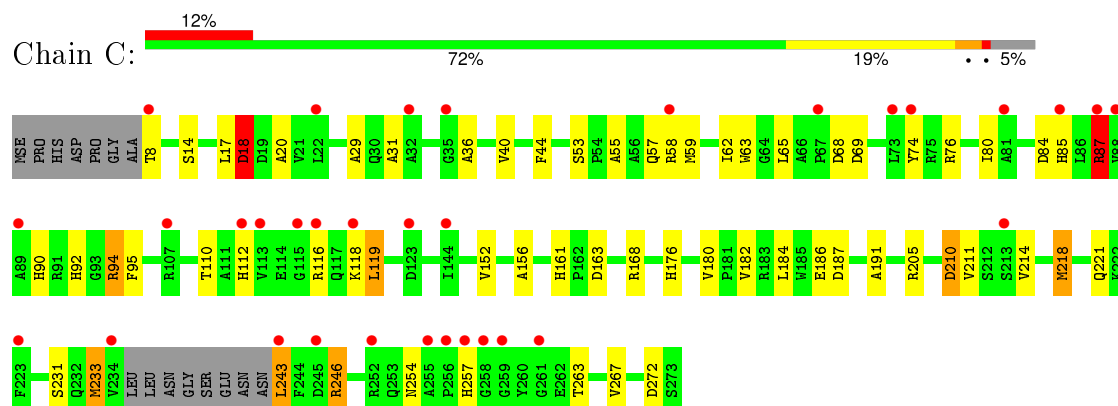
#### • Molecule 1: N-ACYL GLM PEUDO-TEICOPLANIN DEACETYLASE



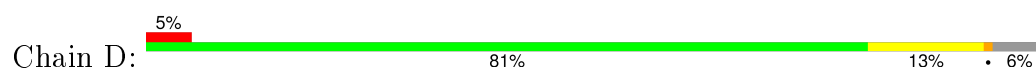
#### • Molecule 1: N-ACYL GLM PEUDO-TEICOPLANIN DEACETYLASE



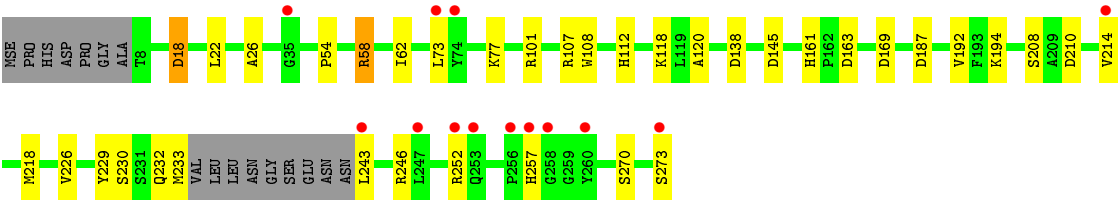
#### • Molecule 1: N-ACYL GLM PEUDO-TEICOPLANIN DEACETYLASE



#### • Molecule 1: N-ACYL GLM PEUDO-TEICOPLANIN DEACETYLASE







• Molecule 2: TEICOPLANIN



• Molecule 2: TEICOPLANIN



• Molecule 2: TEICOPLANIN



• Molecule 2: TEICOPLANIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.77Å 70.71Å 76.31Å 113.89° 108.29° 98.20°	Depositor
Resolution (Å)	20.00 – 1.70 19.94 – 1.70	Depositor EDS
% Data completeness (in resolution range)	95.8 (20.00-1.70) 77.1 (19.94-1.70)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.37 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.201 , 0.237 0.200 , 0.236	Depositor DCC
$R_{free}$ test set	5379 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.4	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 29.7	EDS
Estimated twinning fraction	0.125 for -h,-k,h+k+l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 106831 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9379	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 59.45 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.7616e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: BMA, GHP, NAG, 3MY, T55, GCS, 3FG, OMY

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	1.25	8/2064 (0.4%)	1.34	12/2805 (0.4%)
1	B	1.17	4/2052 (0.2%)	1.08	7/2789 (0.3%)
1	C	1.20	3/2054 (0.1%)	1.13	11/2792 (0.4%)
1	D	1.14	1/2052 (0.0%)	1.09	9/2788 (0.3%)
All	All	1.19	16/8222 (0.2%)	1.16	39/11174 (0.3%)

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	21	VAL	CB-CG1	7.31	1.68	1.52
1	A	265	TRP	CG-CD1	6.90	1.46	1.36
1	C	44	PHE	CE2-CZ	6.50	1.49	1.37
1	A	101	ARG	CD-NE	-6.16	1.35	1.46
1	A	41	TYR	CE1-CZ	6.12	1.46	1.38

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	101	ARG	NE-CZ-NH2	-31.52	104.54	120.30
1	A	101	ARG	NE-CZ-NH1	24.14	132.37	120.30
1	A	101	ARG	CD-NE-CZ	10.92	138.88	123.60
1	C	205	ARG	NE-CZ-NH2	-10.55	115.02	120.30
1	C	205	ARG	NE-CZ-NH1	9.81	125.20	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2018	0	1935	30	1
1	B	2006	0	1921	19	0
1	C	2008	0	1925	66	1
1	D	2006	0	1921	30	0
2	E	84	0	34	2	0
2	F	85	0	35	4	0
2	G	85	0	35	3	0
2	H	85	0	34	5	0
3	E	11	0	10	0	0
3	F	11	0	10	1	0
3	G	11	0	10	0	0
3	H	11	0	10	4	0
4	E	14	0	13	2	0
4	F	14	0	11	0	0
4	G	14	0	13	1	0
4	H	14	0	13	0	0
5	E	11	0	9	0	0
5	F	11	0	9	0	0
5	G	11	0	9	0	0
5	H	11	0	9	0	0
6	E	11	0	19	4	0
6	F	11	0	19	4	0
6	G	11	0	19	3	0
6	H	11	0	19	12	0
7	A	204	0	0	4	0
7	B	200	0	0	0	0
7	C	158	0	0	2	0
7	D	221	0	0	4	0
7	E	4	0	0	0	0
7	F	13	0	0	2	0
7	G	4	0	0	0	0
7	H	10	0	0	0	0
All	All	9379	0	8042	154	1

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 9.

The worst 5 of 154 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:C:29:ALA:CB	1:C:218:MSE:HE2	1.57	1.34
1:C:29:ALA:HB1	1:C:218:MSE:CE	1.66	1.26
1:D:233:MSE:HE2	7:D:2200:HOH:O	1.45	1.17
1:D:233:MSE:HE3	6:H:711:T55:CAF	1.76	1.15
1:C:243:LEU:HD22	6:G:711:T55:HAF1	1.32	1.11

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:195:SER:O	1:C:58:ARG:NH2[1_544]	2.15	0.05

## 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	255/273 (93%)	247 (97%)	6 (2%)	2 (1%)	24	7
1	B	254/273 (93%)	249 (98%)	3 (1%)	2 (1%)	24	7
1	C	254/273 (93%)	241 (95%)	12 (5%)	1 (0%)	39	20
1	D	253/273 (93%)	246 (97%)	5 (2%)	2 (1%)	24	7
All	All	1016/1092 (93%)	983 (97%)	26 (3%)	7 (1%)	26	9

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	ASP
1	B	18	ASP
1	B	257	HIS
1	C	18	ASP
1	D	18	ASP

### 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	210/218 (96%)	201 (96%)	9 (4%)	35	14
1	B	208/218 (95%)	200 (96%)	8 (4%)	40	17
1	C	209/218 (96%)	199 (95%)	10 (5%)	31	12
1	D	209/218 (96%)	204 (98%)	5 (2%)	57	36
All	All	836/872 (96%)	804 (96%)	32 (4%)	40	17

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	217	GLU
1	C	8	THR
1	D	187	ASP
1	B	253	GLN
1	C	69	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	122	ASN
1	C	161	HIS
1	D	122	ASN
1	C	90	HIS
1	C	117	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

28 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	GHP	E	701	2	10,11,12	1.17	1 (10%)	12,14,16	1.01	1 (8%)
2	3MY	E	702	2	12,13,14	1.44	2 (16%)	14,17,19	3.27	7 (50%)
2	3FG	E	703	2	10,11,13	1.16	1 (10%)	12,14,18	2.02	5 (41%)
2	GHP	E	704	2,5	10,11,12	2.66	4 (40%)	12,14,16	1.34	3 (25%)
2	GHP	E	705	2	10,11,12	1.82	4 (40%)	12,14,16	2.03	4 (33%)
2	OMY	E	706	2,4	13,14,15	1.80	3 (23%)	17,19,21	3.61	5 (29%)
2	3FG	E	707	3,2	9,13,13	1.28	1 (11%)	14,18,18	1.57	2 (14%)
2	GHP	F	701	2	10,11,12	1.00	0	12,14,16	1.25	1 (8%)
2	3MY	F	702	2	12,13,14	1.16	1 (8%)	14,17,19	2.93	4 (28%)
2	3FG	F	703	2	11,12,13	1.36	2 (18%)	14,16,18	2.12	2 (14%)
2	GHP	F	704	2,5	10,11,12	2.88	4 (40%)	12,14,16	2.41	5 (41%)
2	GHP	F	705	2	10,11,12	1.61	3 (30%)	12,14,16	2.13	6 (50%)
2	OMY	F	706	2,4	13,14,15	1.40	1 (7%)	17,19,21	3.47	2 (11%)
2	3FG	F	707	3,2	9,13,13	1.46	1 (11%)	14,18,18	2.50	4 (28%)
2	GHP	G	701	2	10,11,12	1.41	1 (10%)	12,14,16	1.53	3 (25%)
2	3MY	G	702	2	12,13,14	0.87	0	14,17,19	2.81	9 (64%)
2	3FG	G	703	2	11,12,13	0.69	0	14,16,18	0.71	0
2	GHP	G	704	2,5	10,11,12	2.76	3 (30%)	12,14,16	1.73	4 (33%)
2	GHP	G	705	2	10,11,12	2.75	2 (20%)	12,14,16	2.54	6 (50%)
2	OMY	G	706	2,4	13,14,15	1.27	2 (15%)	17,19,21	3.48	4 (23%)
2	3FG	G	707	3,2	9,13,13	1.39	2 (22%)	14,18,18	2.35	4 (28%)
2	GHP	H	701	2	10,11,12	1.33	2 (20%)	12,14,16	1.47	2 (16%)
2	3MY	H	702	2	12,13,14	1.30	1 (8%)	14,17,19	2.52	5 (35%)
2	3FG	H	703	2	11,12,13	1.31	1 (9%)	14,16,18	0.94	1 (7%)
2	GHP	H	704	2,5	10,11,12	1.56	3 (30%)	12,14,16	0.96	0
2	GHP	H	705	2	10,11,12	1.53	2 (20%)	12,14,16	2.42	5 (41%)
2	OMY	H	706	2,4	13,14,15	2.04	6 (46%)	17,19,21	3.33	7 (41%)
2	3FG	H	707	3,2	9,13,13	1.08	1 (11%)	14,18,18	2.71	4 (28%)

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	GHP	E	701	2	-	0/4/6/8	0/1/1/1
2	3MY	E	702	2	-	0/4/6/8	0/1/1/1
2	3FG	E	703	2	-	0/4/6/8	0/1/1/1
2	GHP	E	704	2,5	-	0/4/6/8	0/1/1/1
2	GHP	E	705	2	-	0/4/6/8	0/1/1/1
2	OMY	E	706	2,4	-	0/8/10/12	0/1/1/1
2	3FG	E	707	3,2	-	0/4/8/8	0/1/1/1
2	GHP	F	701	2	-	0/4/6/8	0/1/1/1
2	3MY	F	702	2	-	0/4/6/8	0/1/1/1
2	3FG	F	703	2	-	0/4/6/8	0/1/1/1
2	GHP	F	704	2,5	-	0/4/6/8	0/1/1/1
2	GHP	F	705	2	-	0/4/6/8	0/1/1/1
2	OMY	F	706	2,4	-	0/8/10/12	0/1/1/1
2	3FG	F	707	3,2	-	0/4/8/8	0/1/1/1
2	GHP	G	701	2	-	0/4/6/8	0/1/1/1
2	3MY	G	702	2	-	0/4/6/8	0/1/1/1
2	3FG	G	703	2	-	0/4/6/8	0/1/1/1
2	GHP	G	704	2,5	-	0/4/6/8	0/1/1/1
2	GHP	G	705	2	-	0/4/6/8	0/1/1/1
2	OMY	G	706	2,4	-	0/8/10/12	0/1/1/1
2	3FG	G	707	3,2	-	0/4/8/8	0/1/1/1
2	GHP	H	701	2	-	0/4/6/8	0/1/1/1
2	3MY	H	702	2	-	0/4/6/8	0/1/1/1
2	3FG	H	703	2	-	0/4/6/8	0/1/1/1
2	GHP	H	704	2,5	-	0/4/6/8	0/1/1/1
2	GHP	H	705	2	-	0/4/6/8	0/1/1/1
2	OMY	H	706	2,4	-	0/8/10/12	0/1/1/1
2	3FG	H	707	3,2	-	0/4/8/8	0/1/1/1

The worst 5 of 54 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	706	OMY	CZ-CE1	-3.09	1.36	1.39
2	H	706	OMY	CD1-CG	-2.18	1.35	1.39
2	E	705	GHP	C6-C1	2.01	1.42	1.39
2	H	707	3FG	OD2-CD2	2.02	1.41	1.37
2	E	703	3FG	CB-CA	2.02	1.54	1.52

The worst 5 of 105 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	706	OMY	CG-CB-CA	-13.24	93.45	111.96
2	F	706	OMY	CG-CB-CA	-11.87	95.36	111.96
2	G	706	OMY	CG-CB-CA	-11.66	95.65	111.96
2	H	706	OMY	CG-CB-CA	-9.19	99.11	111.96
2	F	702	3MY	CD1-CE1-CZ	-6.43	113.90	120.49

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

13 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	705	GHP	1	0
2	E	707	3FG	1	0
2	F	702	3MY	1	0
2	F	703	3FG	1	0
2	F	705	GHP	2	0
2	F	707	3FG	1	0
2	G	701	GHP	1	0
2	G	703	3FG	1	0
2	G	705	GHP	1	0
2	G	706	OMY	1	0
2	H	702	3MY	1	0
2	H	703	3FG	1	0
2	H	707	3FG	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

16 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	BMA	E	708	2	11,11,12	1.61	3 (27%)	14,15,17	1.77	3 (21%)
4	NAG	E	709	2	14,14,15	1.58	1 (7%)	15,19,21	1.11	1 (6%)
5	GCS	E	710	2,6	11,11,12	1.34	2 (18%)	13,15,17	2.23	4 (30%)
6	T55	E	711	5	10,10,11	0.90	1 (10%)	9,10,12	1.65	2 (22%)
3	BMA	F	708	2	11,11,12	1.51	2 (18%)	14,15,17	1.22	1 (7%)
4	NAG	F	709	2	14,14,15	3.20	6 (42%)	15,19,21	4.78	6 (40%)
5	GCS	F	710	2,6	11,11,12	1.73	1 (9%)	13,15,17	1.96	4 (30%)
6	T55	F	711	5	10,10,11	0.42	0	9,10,12	0.75	0
3	BMA	G	708	2	11,11,12	2.09	4 (36%)	14,15,17	2.55	6 (42%)
4	NAG	G	709	2	14,14,15	2.56	3 (21%)	15,19,21	2.66	3 (20%)
5	GCS	G	710	2,6	11,11,12	1.47	3 (27%)	13,15,17	1.43	3 (23%)
6	T55	G	711	5	10,10,11	0.75	0	9,10,12	1.17	1 (11%)
3	BMA	H	708	2	11,11,12	1.77	3 (27%)	14,15,17	4.32	8 (57%)
4	NAG	H	709	2	14,14,15	1.45	1 (7%)	15,19,21	2.16	6 (40%)
5	GCS	H	710	2,6	11,11,12	1.48	3 (27%)	13,15,17	1.93	4 (30%)
6	T55	H	711	5	10,10,11	1.14	1 (10%)	9,10,12	0.39	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	BMA	E	708	2	-	0/2/19/22	0/1/1/1
4	NAG	E	709	2	-	0/6/23/26	0/1/1/1
5	GCS	E	710	2,6	-	0/2/19/22	0/1/1/1
6	T55	E	711	5	-	0/7/8/9	0/0/0/0
3	BMA	F	708	2	-	0/2/19/22	0/1/1/1
4	NAG	F	709	2	1/1/5/7	0/6/23/26	0/1/1/1
5	GCS	F	710	2,6	-	0/2/19/22	0/1/1/1
6	T55	F	711	5	-	0/7/8/9	0/0/0/0
3	BMA	G	708	2	-	0/2/19/22	0/1/1/1
4	NAG	G	709	2	-	0/6/23/26	0/1/1/1
5	GCS	G	710	2,6	-	0/2/19/22	0/1/1/1
6	T55	G	711	5	-	0/7/8/9	0/0/0/0
3	BMA	H	708	2	-	0/2/19/22	0/1/1/1
4	NAG	H	709	2	-	0/6/23/26	0/1/1/1
5	GCS	H	710	2,6	-	0/2/19/22	0/1/1/1
6	T55	H	711	5	-	0/7/8/9	0/0/0/0

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	709	NAG	C4-C3	-3.57	1.43	1.52
5	H	710	GCS	C3-C2	-2.61	1.48	1.53
5	G	710	GCS	O5-C1	-2.60	1.39	1.43
3	H	708	BMA	C1-C2	-2.49	1.46	1.52
5	E	710	GCS	O5-C1	-2.42	1.39	1.43

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	708	BMA	C1-C2-C3	-12.26	95.03	109.54
3	G	708	BMA	C1-C2-C3	-5.16	103.43	109.54
4	H	709	NAG	C6-C5-C4	-4.68	101.47	113.02
3	G	708	BMA	O5-C1-C2	-4.30	103.88	110.86
5	E	710	GCS	C4-C3-C2	-4.13	105.82	111.39

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	F	709	NAG	C3

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	709	NAG	2	0
6	E	711	T55	4	0
3	F	708	BMA	1	0
6	F	711	T55	4	0
4	G	709	NAG	1	0
6	G	711	T55	3	0
3	H	708	BMA	4	0
6	H	711	T55	12	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	256/273 (93%)	0.34	16 (6%) 23 25	8, 18, 45, 61	0
1	B	255/273 (93%)	0.27	13 (5%) 32 34	9, 19, 45, 60	0
1	C	255/273 (93%)	0.74	33 (12%) 5 5	10, 26, 47, 63	0
1	D	254/273 (93%)	0.26	13 (5%) 32 34	8, 19, 37, 62	0
2	E	0/7	-	-	-	-
2	F	0/7	-	-	-	-
2	G	0/7	-	-	-	-
2	H	0/7	-	-	-	-
All	All	1020/1120 (91%)	0.41	75 (7%) 17 19	8, 20, 45, 63	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	234	VAL	9.1
1	C	243	LEU	8.3
1	C	257	HIS	6.5
1	C	259	GLY	6.3
1	A	256	PRO	5.9

### 6.2 Non-standard residues in protein, DNA, RNA chains [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	GHP	E	704	11/12	0.96	0.08	-	11,18,27,27	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	3FG	H	707	13/13	0.90	0.11	-	20,24,31,32	0
2	GHP	G	705	11/12	0.88	0.11	-	25,31,34,36	0
2	3FG	F	703	12/13	0.86	0.12	-	19,31,36,38	0
2	3FG	E	703	11/13	0.88	0.12	-	26,30,39,41	0
2	GHP	E	705	11/12	0.79	0.19	-	30,34,42,45	0
2	GHP	F	704	11/12	0.94	0.10	-	16,18,22,24	0
2	3MY	E	702	13/14	0.91	0.11	-	13,19,29,30	0
2	3MY	H	702	13/14	0.97	0.07	-	8,12,18,18	0
2	GHP	F	701	11/12	0.91	0.12	-	27,31,36,40	0
2	GHP	H	705	11/12	0.94	0.08	-	19,23,26,27	0
2	3FG	E	707	13/13	0.84	0.13	-	38,41,44,49	0
2	3MY	F	702	13/14	0.94	0.10	-	11,16,27,30	0
2	3FG	H	703	12/13	0.92	0.10	-	15,21,27,30	0
2	3MY	G	702	13/14	0.95	0.10	-	12,21,34,36	0
2	GHP	G	704	11/12	0.93	0.08	-	16,21,26,28	0
2	3FG	F	707	13/13	0.90	0.11	-	30,33,38,42	0
2	GHP	H	704	11/12	0.95	0.09	-	10,15,18,20	0
2	OMY	G	706	14/15	0.78	0.16	-	23,28,41,43	0
2	GHP	E	701	11/12	0.76	0.17	-	32,33,37,40	0
2	GHP	G	701	11/12	0.55	0.58	-	37,50,56,56	0
2	OMY	E	706	14/15	0.79	0.16	-	18,29,39,43	0
2	GHP	F	705	11/12	0.87	0.15	-	22,30,34,34	0
2	OMY	H	706	14/15	0.88	0.11	-	14,20,26,42	0
2	GHP	H	701	11/12	0.94	0.07	-	18,20,23,24	0
2	OMY	F	706	14/15	0.85	0.15	-	21,28,35,38	0
2	3FG	G	707	13/13	0.89	0.12	-	28,32,39,43	0
2	3FG	G	703	12/13	0.83	0.14	-	27,39,48,50	0

## 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
4	NAG	F	709	14/15	0.60	0.46	4.39	4,47,55,58	0
6	T55	H	711	11/12	0.95	0.15	2.44	12,31,41,41	0
6	T55	F	711	11/12	0.90	0.17	1.81	19,34,38,39	0
5	GCS	H	710	11/12	0.95	0.10	1.37	4,9,11,12	0
6	T55	E	711	11/12	0.88	0.16	1.28	15,31,40,42	0
6	T55	G	711	11/12	0.91	0.18	0.42	19,31,43,46	0
5	GCS	F	710	11/12	0.96	0.09	-0.26	9,14,17,20	0
5	GCS	G	710	11/12	0.96	0.08	-0.80	10,14,17,17	0
5	GCS	E	710	11/12	0.96	0.07	-0.99	8,10,12,12	0
4	NAG	H	709	14/15	0.66	0.40	-	39,51,56,56	0
3	BMA	G	708	11/12	0.50	0.27	-	43,47,50,50	0
3	BMA	H	708	11/12	0.69	0.28	-	36,46,53,53	0
4	NAG	E	709	14/15	0.56	0.41	-	56,64,67,67	0
3	BMA	F	708	11/12	0.58	0.31	-	51,56,59,63	0
4	NAG	G	709	14/15	0.54	0.42	-	54,60,64,65	0
3	BMA	E	708	11/12	0.43	0.37	-	57,62,64,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.