



# wwPDB X-ray Structure Validation Summary Report

Feb 19, 2016 – 11:15 PM GMT

PDB ID : 4ZFJ  
Title : Ergothioneine-biosynthetic Ntn hydrolase EgtC, apo form  
Authors : Vit, A.; Seebeck, F.P.; Blankenfeldt, W.  
Deposited on : 2015-04-21  
Resolution : 1.75 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

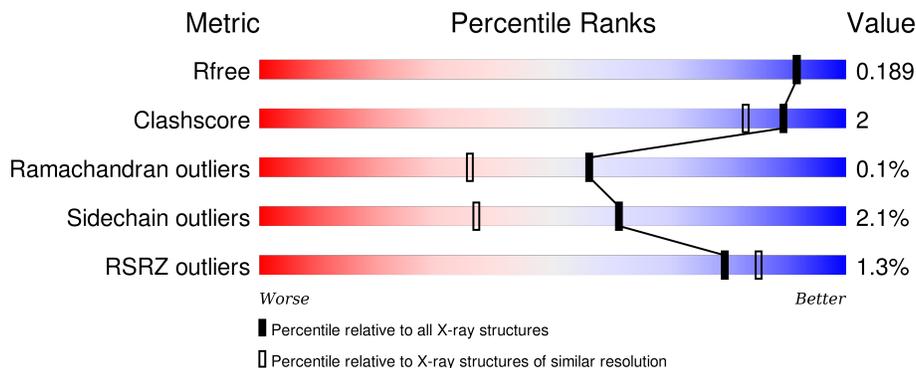
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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.



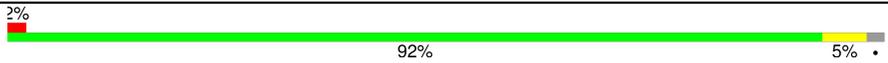
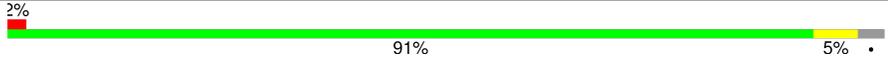
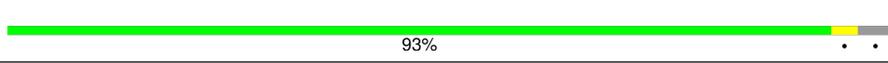
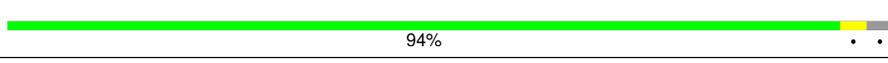
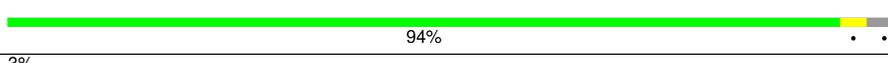
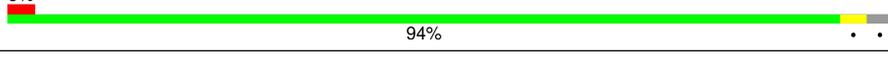
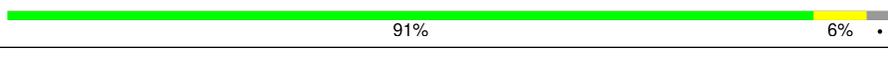
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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	235	
1	B	235	
1	C	235	
1	D	235	
1	E	235	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	235	 2% 92% 5%
1	G	235	 2% 91% 5%
1	H	235	 93%
1	I	235	 94%
1	J	235	 94%
1	K	235	 3% 94%
1	L	235	 91% 6%

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
2	MES	A	302	-	-	-	X
3	XPE	B	302	-	-	-	X
3	XPE	C	301	-	-	-	X
3	XPE	D	302	-	-	-	X
3	XPE	D	303	-	-	-	X
3	XPE	E	302	-	-	-	X
3	XPE	G	302	-	-	-	X
3	XPE	H	302	-	-	-	X
3	XPE	J	302	-	-	-	X
3	XPE	L	302	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 43942 atoms, of which 20573 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 Amidohydrolase EgtC.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se			
1	A	229	3399	1075	1683	312	325	2	2	0	4	0
1	B	228	3391	1071	1683	311	322	2	2	0	2	0
1	C	229	3438	1090	1696	316	332	2	2	0	14	0
1	D	227	3356	1063	1665	309	315	2	2	0	0	0
1	E	228	3379	1069	1671	311	324	2	2	0	4	0
1	F	230	3382	1074	1671	314	319	2	2	0	0	0
1	G	227	3442	1089	1703	320	326	2	2	0	10	0
1	H	226	3338	1058	1653	307	316	2	2	0	0	0
1	I	228	3373	1068	1672	309	320	2	2	0	2	0
1	J	226	3312	1054	1635	303	316	2	2	0	1	0
1	K	227	3358	1066	1663	309	316	2	2	0	1	0
1	L	227	3359	1065	1663	309	318	2	2	0	1	0

There are 168 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	53	ASP	GLU	engineered mutation	UNP A0R5M9
A	84	VAL	LEU	engineered mutation	UNP A0R5M9
A	95	SER	PRO	engineered mutation	UNP A0R5M9
A	118	ALA	ASP	engineered mutation	UNP A0R5M9
A	137	LEU	VAL	engineered mutation	UNP A0R5M9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	188	ARG	HIS	engineered mutation	UNP A0R5M9
A	228	SER	-	expression tag	UNP A0R5M9
A	229	SER	-	expression tag	UNP A0R5M9
A	230	HIS	-	expression tag	UNP A0R5M9
A	231	HIS	-	expression tag	UNP A0R5M9
A	232	HIS	-	expression tag	UNP A0R5M9
A	233	HIS	-	expression tag	UNP A0R5M9
A	234	HIS	-	expression tag	UNP A0R5M9
A	235	HIS	-	expression tag	UNP A0R5M9
B	53	ASP	GLU	engineered mutation	UNP A0R5M9
B	84	VAL	LEU	engineered mutation	UNP A0R5M9
B	95	SER	PRO	engineered mutation	UNP A0R5M9
B	118	ALA	ASP	engineered mutation	UNP A0R5M9
B	137	LEU	VAL	engineered mutation	UNP A0R5M9
B	188	ARG	HIS	engineered mutation	UNP A0R5M9
B	228	SER	-	expression tag	UNP A0R5M9
B	229	SER	-	expression tag	UNP A0R5M9
B	230	HIS	-	expression tag	UNP A0R5M9
B	231	HIS	-	expression tag	UNP A0R5M9
B	232	HIS	-	expression tag	UNP A0R5M9
B	233	HIS	-	expression tag	UNP A0R5M9
B	234	HIS	-	expression tag	UNP A0R5M9
B	235	HIS	-	expression tag	UNP A0R5M9
C	53	ASP	GLU	engineered mutation	UNP A0R5M9
C	84	VAL	LEU	engineered mutation	UNP A0R5M9
C	95	SER	PRO	engineered mutation	UNP A0R5M9
C	118	ALA	ASP	engineered mutation	UNP A0R5M9
C	137	LEU	VAL	engineered mutation	UNP A0R5M9
C	188	ARG	HIS	engineered mutation	UNP A0R5M9
C	228	SER	-	expression tag	UNP A0R5M9
C	229	SER	-	expression tag	UNP A0R5M9
C	230	HIS	-	expression tag	UNP A0R5M9
C	231	HIS	-	expression tag	UNP A0R5M9
C	232	HIS	-	expression tag	UNP A0R5M9
C	233	HIS	-	expression tag	UNP A0R5M9
C	234	HIS	-	expression tag	UNP A0R5M9
C	235	HIS	-	expression tag	UNP A0R5M9
D	53	ASP	GLU	engineered mutation	UNP A0R5M9
D	84	VAL	LEU	engineered mutation	UNP A0R5M9
D	95	SER	PRO	engineered mutation	UNP A0R5M9
D	118	ALA	ASP	engineered mutation	UNP A0R5M9
D	137	LEU	VAL	engineered mutation	UNP A0R5M9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	188	ARG	HIS	engineered mutation	UNP A0R5M9
D	228	SER	-	expression tag	UNP A0R5M9
D	229	SER	-	expression tag	UNP A0R5M9
D	230	HIS	-	expression tag	UNP A0R5M9
D	231	HIS	-	expression tag	UNP A0R5M9
D	232	HIS	-	expression tag	UNP A0R5M9
D	233	HIS	-	expression tag	UNP A0R5M9
D	234	HIS	-	expression tag	UNP A0R5M9
D	235	HIS	-	expression tag	UNP A0R5M9
E	53	ASP	GLU	engineered mutation	UNP A0R5M9
E	84	VAL	LEU	engineered mutation	UNP A0R5M9
E	95	SER	PRO	engineered mutation	UNP A0R5M9
E	118	ALA	ASP	engineered mutation	UNP A0R5M9
E	137	LEU	VAL	engineered mutation	UNP A0R5M9
E	188	ARG	HIS	engineered mutation	UNP A0R5M9
E	228	SER	-	expression tag	UNP A0R5M9
E	229	SER	-	expression tag	UNP A0R5M9
E	230	HIS	-	expression tag	UNP A0R5M9
E	231	HIS	-	expression tag	UNP A0R5M9
E	232	HIS	-	expression tag	UNP A0R5M9
E	233	HIS	-	expression tag	UNP A0R5M9
E	234	HIS	-	expression tag	UNP A0R5M9
E	235	HIS	-	expression tag	UNP A0R5M9
F	53	ASP	GLU	engineered mutation	UNP A0R5M9
F	84	VAL	LEU	engineered mutation	UNP A0R5M9
F	95	SER	PRO	engineered mutation	UNP A0R5M9
F	118	ALA	ASP	engineered mutation	UNP A0R5M9
F	137	LEU	VAL	engineered mutation	UNP A0R5M9
F	188	ARG	HIS	engineered mutation	UNP A0R5M9
F	228	SER	-	expression tag	UNP A0R5M9
F	229	SER	-	expression tag	UNP A0R5M9
F	230	HIS	-	expression tag	UNP A0R5M9
F	231	HIS	-	expression tag	UNP A0R5M9
F	232	HIS	-	expression tag	UNP A0R5M9
F	233	HIS	-	expression tag	UNP A0R5M9
F	234	HIS	-	expression tag	UNP A0R5M9
F	235	HIS	-	expression tag	UNP A0R5M9
G	53	ASP	GLU	engineered mutation	UNP A0R5M9
G	84	VAL	LEU	engineered mutation	UNP A0R5M9
G	95	SER	PRO	engineered mutation	UNP A0R5M9
G	118	ALA	ASP	engineered mutation	UNP A0R5M9
G	137	LEU	VAL	engineered mutation	UNP A0R5M9

*Continued on next page...*

*Continued from previous page...*

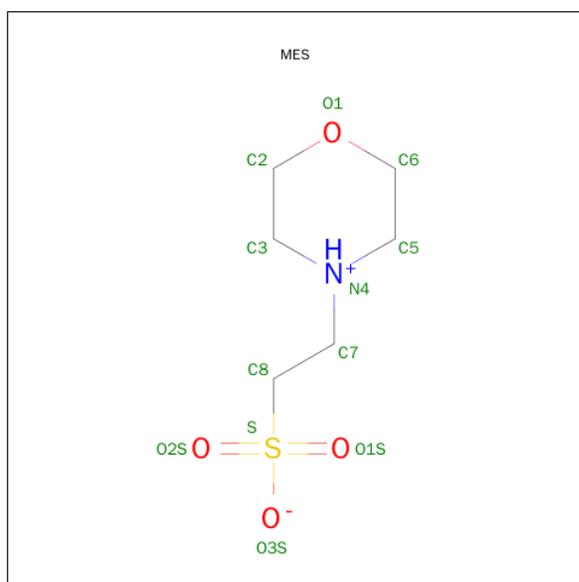
Chain	Residue	Modelled	Actual	Comment	Reference
G	188	ARG	HIS	engineered mutation	UNP A0R5M9
G	228	SER	-	expression tag	UNP A0R5M9
G	229	SER	-	expression tag	UNP A0R5M9
G	230	HIS	-	expression tag	UNP A0R5M9
G	231	HIS	-	expression tag	UNP A0R5M9
G	232	HIS	-	expression tag	UNP A0R5M9
G	233	HIS	-	expression tag	UNP A0R5M9
G	234	HIS	-	expression tag	UNP A0R5M9
G	235	HIS	-	expression tag	UNP A0R5M9
H	53	ASP	GLU	engineered mutation	UNP A0R5M9
H	84	VAL	LEU	engineered mutation	UNP A0R5M9
H	95	SER	PRO	engineered mutation	UNP A0R5M9
H	118	ALA	ASP	engineered mutation	UNP A0R5M9
H	137	LEU	VAL	engineered mutation	UNP A0R5M9
H	188	ARG	HIS	engineered mutation	UNP A0R5M9
H	228	SER	-	expression tag	UNP A0R5M9
H	229	SER	-	expression tag	UNP A0R5M9
H	230	HIS	-	expression tag	UNP A0R5M9
H	231	HIS	-	expression tag	UNP A0R5M9
H	232	HIS	-	expression tag	UNP A0R5M9
H	233	HIS	-	expression tag	UNP A0R5M9
H	234	HIS	-	expression tag	UNP A0R5M9
H	235	HIS	-	expression tag	UNP A0R5M9
I	53	ASP	GLU	engineered mutation	UNP A0R5M9
I	84	VAL	LEU	engineered mutation	UNP A0R5M9
I	95	SER	PRO	engineered mutation	UNP A0R5M9
I	118	ALA	ASP	engineered mutation	UNP A0R5M9
I	137	LEU	VAL	engineered mutation	UNP A0R5M9
I	188	ARG	HIS	engineered mutation	UNP A0R5M9
I	228	SER	-	expression tag	UNP A0R5M9
I	229	SER	-	expression tag	UNP A0R5M9
I	230	HIS	-	expression tag	UNP A0R5M9
I	231	HIS	-	expression tag	UNP A0R5M9
I	232	HIS	-	expression tag	UNP A0R5M9
I	233	HIS	-	expression tag	UNP A0R5M9
I	234	HIS	-	expression tag	UNP A0R5M9
I	235	HIS	-	expression tag	UNP A0R5M9
J	53	ASP	GLU	engineered mutation	UNP A0R5M9
J	84	VAL	LEU	engineered mutation	UNP A0R5M9
J	95	SER	PRO	engineered mutation	UNP A0R5M9
J	118	ALA	ASP	engineered mutation	UNP A0R5M9
J	137	LEU	VAL	engineered mutation	UNP A0R5M9

*Continued on next page...*

*Continued from previous page...*

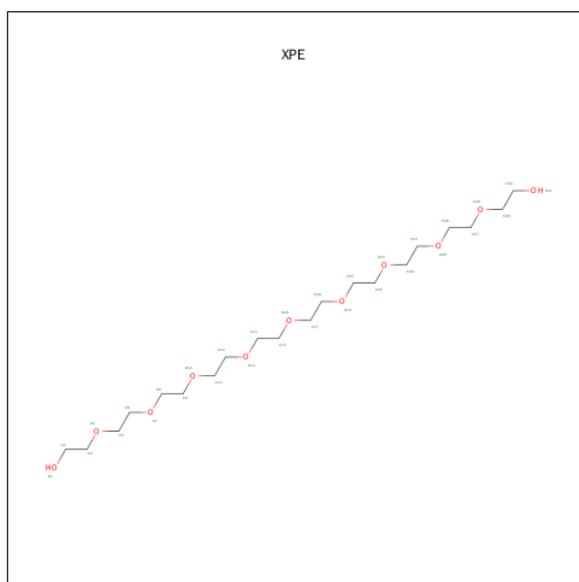
Chain	Residue	Modelled	Actual	Comment	Reference
J	188	ARG	HIS	engineered mutation	UNP A0R5M9
J	228	SER	-	expression tag	UNP A0R5M9
J	229	SER	-	expression tag	UNP A0R5M9
J	230	HIS	-	expression tag	UNP A0R5M9
J	231	HIS	-	expression tag	UNP A0R5M9
J	232	HIS	-	expression tag	UNP A0R5M9
J	233	HIS	-	expression tag	UNP A0R5M9
J	234	HIS	-	expression tag	UNP A0R5M9
J	235	HIS	-	expression tag	UNP A0R5M9
K	53	ASP	GLU	engineered mutation	UNP A0R5M9
K	84	VAL	LEU	engineered mutation	UNP A0R5M9
K	95	SER	PRO	engineered mutation	UNP A0R5M9
K	118	ALA	ASP	engineered mutation	UNP A0R5M9
K	137	LEU	VAL	engineered mutation	UNP A0R5M9
K	188	ARG	HIS	engineered mutation	UNP A0R5M9
K	228	SER	-	expression tag	UNP A0R5M9
K	229	SER	-	expression tag	UNP A0R5M9
K	230	HIS	-	expression tag	UNP A0R5M9
K	231	HIS	-	expression tag	UNP A0R5M9
K	232	HIS	-	expression tag	UNP A0R5M9
K	233	HIS	-	expression tag	UNP A0R5M9
K	234	HIS	-	expression tag	UNP A0R5M9
K	235	HIS	-	expression tag	UNP A0R5M9
L	53	ASP	GLU	engineered mutation	UNP A0R5M9
L	84	VAL	LEU	engineered mutation	UNP A0R5M9
L	95	SER	PRO	engineered mutation	UNP A0R5M9
L	118	ALA	ASP	engineered mutation	UNP A0R5M9
L	137	LEU	VAL	engineered mutation	UNP A0R5M9
L	188	ARG	HIS	engineered mutation	UNP A0R5M9
L	228	SER	-	expression tag	UNP A0R5M9
L	229	SER	-	expression tag	UNP A0R5M9
L	230	HIS	-	expression tag	UNP A0R5M9
L	231	HIS	-	expression tag	UNP A0R5M9
L	232	HIS	-	expression tag	UNP A0R5M9
L	233	HIS	-	expression tag	UNP A0R5M9
L	234	HIS	-	expression tag	UNP A0R5M9
L	235	HIS	-	expression tag	UNP A0R5M9

- Molecule 2 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			S
2	A	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	A	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	B	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	D	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	E	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	F	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	G	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	H	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	I	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	J	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	K	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
2	L	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		

- Molecule 3 is 3,6,9,12,15,18,21,24,27-NONAOXANONACOSANE-1,29-DIOL (three-letter code: XPE) (formula: C<sub>20</sub>H<sub>42</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	H	O	0	0
			63	18	36	9		
3	C	1	Total	C	H	O	0	0
			55	16	31	8		
3	D	1	Total	C	H	O	0	0
			60	17	34	9		
3	D	1	Total	C	H	O	0	0
			55	16	31	8		
3	E	1	Total	C	H	O	0	0
			55	16	31	8		
3	G	1	Total	C	H	O	0	0
			63	18	36	9		
3	H	1	Total	C	H	O	0	0
			55	16	31	8		
3	I	1	Total	C	H	O	0	0
			63	18	36	9		
3	J	1	Total	C	H	O	0	0
			47	14	26	7		
3	K	1	Total	C	H	O	0	0
			55	16	31	8		
3	L	1	Total	C	H	O	0	0
			63	18	36	9		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	230	Total	O	0	0
			230	230		

*Continued on next page...*

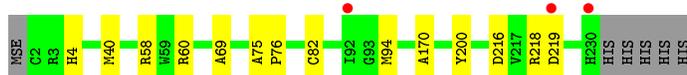
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	252	Total 252	O 252	0	0
4	C	199	Total 199	O 199	0	0
4	D	197	Total 197	O 197	0	0
4	E	189	Total 189	O 189	0	0
4	F	193	Total 193	O 193	0	0
4	G	208	Total 208	O 208	0	0
4	H	231	Total 231	O 231	0	0
4	I	211	Total 211	O 211	0	0
4	J	215	Total 215	O 215	0	0
4	K	181	Total 181	O 181	0	0
4	L	175	Total 175	O 175	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 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.

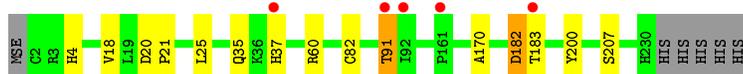
- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



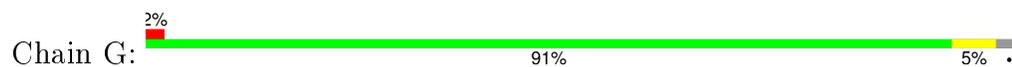
- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.98Å 69.51Å 160.38Å 90.00° 94.84° 90.00°	Depositor
Resolution (Å)	47.64 – 1.75 47.64 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.64-1.75) 99.4 (47.64-1.75)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 1.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.149 , 0.180 0.163 , 0.189	Depositor DCC
$R_{free}$ test set	13873 reflections (5.14%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.6	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 50.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	2 of 283637 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	43942	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	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 44.94 % 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 1.4203e-04. 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: MES, XPE

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	0.34	0/1776	0.55	0/2425
1	B	0.36	0/1757	0.58	1/2399 (0.0%)
1	C	0.34	0/1819	0.56	0/2484
1	D	0.32	0/1729	0.54	0/2363
1	E	0.33	0/1762	0.53	0/2408
1	F	0.32	0/1750	0.54	0/2391
1	G	0.32	0/1808	0.53	0/2466
1	H	0.32	0/1723	0.52	0/2355
1	I	0.35	0/1749	0.56	0/2389
1	J	0.31	0/1718	0.53	0/2350
1	K	0.32	0/1737	0.52	0/2375
1	L	0.31	0/1737	0.53	0/2375
All	All	0.33	0/21065	0.54	1/28780 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	219	ASP	C-N-CA	5.07	134.38	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	182	ASP	Peptide
1	C	183[B]	THR	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	1716	1683	1654	14	0
1	B	1708	1683	1669	16	0
1	C	1742	1696	1631	6	0
1	D	1691	1665	1665	10	0
1	E	1708	1671	1654	10	0
1	F	1711	1671	1669	10	0
1	G	1739	1703	1669	10	0
1	H	1685	1653	1651	5	0
1	I	1701	1672	1658	7	0
1	J	1677	1635	1637	5	0
1	K	1695	1663	1661	5	0
1	L	1696	1663	1664	11	0
2	A	24	26	25	0	0
2	B	12	13	13	0	0
2	D	12	13	13	0	0
2	E	12	13	13	0	0
2	F	12	13	13	0	0
2	G	12	13	13	0	0
2	H	12	13	13	0	0
2	I	12	13	13	0	0
2	J	12	13	12	0	0
2	K	12	13	13	0	0
2	L	12	13	13	0	0
3	B	27	36	32	5	0
3	C	24	31	28	4	0
3	D	50	65	60	11	0
3	E	24	31	28	4	0
3	G	27	36	32	3	0
3	H	24	31	28	1	0
3	I	27	36	32	2	0
3	J	21	26	24	0	0
3	K	24	31	28	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	L	27	36	32	3	0
4	A	230	0	0	1	1
4	B	252	0	0	0	2
4	C	199	0	0	0	0
4	D	197	0	0	0	1
4	E	189	0	0	2	0
4	F	193	0	0	2	0
4	G	208	0	0	0	0
4	H	231	0	0	2	1
4	I	211	0	0	1	1
4	J	215	0	0	0	0
4	K	181	0	0	1	0
4	L	175	0	0	0	0
All	All	23369	20573	20360	99	3

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.

The worst 5 of 99 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:E:119:ARG:NH1	1:E:131:THR:O	2.04	0.91
1:F:40:MSE:HE1	1:L:40:MSE:SE	2.37	0.75
1:B:219:ASP:N	1:B:220:ALA:HB3	2.06	0.71
1:F:188:ARG:NH2	4:F:402:HOH:O	2.24	0.70
1:L:40:MSE:HE2	1:L:42:ALA:HB2	1.75	0.68

All (3) 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)
4:B:633:HOH:O	4:D:592:HOH:O[1_565]	2.15	0.05
4:A:555:HOH:O	4:B:419:HOH:O[1_545]	2.18	0.02
4:H:584:HOH:O	4:I:570:HOH:O[1_445]	2.18	0.02

## 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	231/235 (98%)	223 (96%)	8 (4%)	0	100	100
1	B	228/235 (97%)	219 (96%)	8 (4%)	1 (0%)	39	19
1	C	240/235 (102%)	230 (96%)	8 (3%)	2 (1%)	24	8
1	D	225/235 (96%)	219 (97%)	6 (3%)	0	100	100
1	E	230/235 (98%)	220 (96%)	8 (4%)	2 (1%)	21	6
1	F	228/235 (97%)	221 (97%)	7 (3%)	0	100	100
1	G	235/235 (100%)	227 (97%)	8 (3%)	0	100	100
1	H	224/235 (95%)	217 (97%)	7 (3%)	0	100	100
1	I	228/235 (97%)	222 (97%)	6 (3%)	0	100	100
1	J	225/235 (96%)	219 (97%)	6 (3%)	0	100	100
1	K	226/235 (96%)	220 (97%)	6 (3%)	0	100	100
1	L	226/235 (96%)	220 (97%)	6 (3%)	0	100	100
All	All	2746/2820 (97%)	2657 (97%)	84 (3%)	5 (0%)	56	32

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	37[A]	HIS
1	C	37[B]	HIS
1	E	92[A]	ILE
1	E	92[B]	ILE
1	B	219	ASP

### 5.3.2 Protein sidechains [i](#)

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	180/181 (99%)	176 (98%)	4 (2%)	60	35
1	B	179/181 (99%)	175 (98%)	4 (2%)	60	35
1	C	181/181 (100%)	173 (96%)	8 (4%)	35	11
1	D	174/181 (96%)	170 (98%)	4 (2%)	58	33
1	E	178/181 (98%)	173 (97%)	5 (3%)	51	25
1	F	174/181 (96%)	171 (98%)	3 (2%)	68	49
1	G	182/181 (101%)	179 (98%)	3 (2%)	70	52
1	H	173/181 (96%)	170 (98%)	3 (2%)	68	49
1	I	176/181 (97%)	173 (98%)	3 (2%)	68	49
1	J	172/181 (95%)	170 (99%)	2 (1%)	78	62
1	K	174/181 (96%)	172 (99%)	2 (1%)	80	66
1	L	175/181 (97%)	171 (98%)	4 (2%)	58	33
All	All	2118/2172 (98%)	2073 (98%)	45 (2%)	61	37

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	4	HIS
1	F	4	HIS
1	L	4	HIS
1	E	198	GLU
1	F	40	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

23 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
2	MES	A	301	-	12,12,12	1.98	1 (8%)	16,16,16	1.12	1 (6%)
2	MES	A	302	-	12,12,12	2.07	1 (8%)	16,16,16	1.43	1 (6%)
2	MES	B	301	-	12,12,12	2.07	1 (8%)	16,16,16	1.54	3 (18%)
3	XPE	B	302	-	26,26,30	0.71	0	25,25,29	0.64	0
3	XPE	C	301	-	23,23,30	0.65	0	22,22,29	0.34	0
2	MES	D	301	-	12,12,12	1.99	1 (8%)	16,16,16	1.21	0
3	XPE	D	302	-	25,25,30	0.67	0	24,24,29	0.59	0
3	XPE	D	303	-	23,23,30	0.73	0	22,22,29	0.55	0
2	MES	E	301	-	12,12,12	1.96	1 (8%)	16,16,16	1.47	3 (18%)
3	XPE	E	302	-	23,23,30	0.69	0	22,22,29	0.60	0
2	MES	F	301	-	12,12,12	2.01	1 (8%)	16,16,16	1.59	3 (18%)
2	MES	G	301	-	12,12,12	2.00	1 (8%)	16,16,16	1.21	0
3	XPE	G	302	-	26,26,30	0.76	0	25,25,29	0.62	1 (4%)
2	MES	H	301	-	12,12,12	1.96	1 (8%)	16,16,16	1.38	2 (12%)
3	XPE	H	302	-	23,23,30	0.77	0	22,22,29	0.57	0
2	MES	I	301	-	12,12,12	2.00	1 (8%)	16,16,16	1.23	1 (6%)
3	XPE	I	302	-	26,26,30	0.63	0	25,25,29	0.46	0
2	MES	J	301	-	12,12,12	1.93	1 (8%)	16,16,16	1.90	4 (25%)
3	XPE	J	302	-	20,20,30	0.72	0	19,19,29	0.31	0
2	MES	K	301	-	12,12,12	1.95	1 (8%)	16,16,16	1.39	1 (6%)
3	XPE	K	302	-	23,23,30	0.67	0	22,22,29	0.26	0
2	MES	L	301	-	12,12,12	2.06	1 (8%)	16,16,16	1.20	1 (6%)
3	XPE	L	302	-	26,26,30	0.67	0	25,25,29	0.38	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	MES	A	301	-	-	0/6/14/14	0/1/1/1
2	MES	A	302	-	-	0/6/14/14	0/1/1/1
2	MES	B	301	-	-	0/6/14/14	0/1/1/1
3	XPE	B	302	-	-	0/24/24/28	0/0/0/0
3	XPE	C	301	-	-	0/21/21/28	0/0/0/0
2	MES	D	301	-	-	0/6/14/14	0/1/1/1
3	XPE	D	302	-	-	0/23/23/28	0/0/0/0
3	XPE	D	303	-	-	0/21/21/28	0/0/0/0
2	MES	E	301	-	-	0/6/14/14	0/1/1/1
3	XPE	E	302	-	-	0/21/21/28	0/0/0/0
2	MES	F	301	-	-	0/6/14/14	0/1/1/1
2	MES	G	301	-	-	0/6/14/14	0/1/1/1
3	XPE	G	302	-	-	0/24/24/28	0/0/0/0
2	MES	H	301	-	-	0/6/14/14	0/1/1/1
3	XPE	H	302	-	-	0/21/21/28	0/0/0/0
2	MES	I	301	-	-	0/6/14/14	0/1/1/1
3	XPE	I	302	-	-	0/24/24/28	0/0/0/0
2	MES	J	301	-	-	0/6/14/14	0/1/1/1
3	XPE	J	302	-	-	0/18/18/28	0/0/0/0
2	MES	K	301	-	-	0/6/14/14	0/1/1/1
3	XPE	K	302	-	-	0/21/21/28	0/0/0/0
2	MES	L	301	-	-	0/6/14/14	0/1/1/1
3	XPE	L	302	-	-	0/24/24/28	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	302	MES	C8-S	-6.29	1.66	1.78
2	L	301	MES	C8-S	-6.25	1.66	1.78
2	B	301	MES	C8-S	-6.14	1.66	1.78
2	G	301	MES	C8-S	-6.08	1.67	1.78
2	F	301	MES	C8-S	-6.05	1.67	1.78

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	301	MES	O1-C2-C3	-3.50	107.86	111.42
2	F	301	MES	O1-C2-C3	-2.88	108.49	111.42
2	J	301	MES	C6-C5-N4	-2.36	107.06	109.97

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	301	MES	O1-C6-C5	-2.32	109.06	111.42
2	E	301	MES	O2S-S-O1S	-2.00	108.42	112.51

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 33 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	302	XPE	5	0
3	C	301	XPE	4	0
3	D	302	XPE	7	0
3	D	303	XPE	4	0
3	E	302	XPE	4	0
3	G	302	XPE	3	0
3	H	302	XPE	1	0
3	I	302	XPE	2	0
3	L	302	XPE	3	0

## 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 [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	227/235 (96%)	-0.26	3 (1%) 79 85	10, 16, 35, 76	0
1	B	226/235 (96%)	-0.34	3 (1%) 79 85	11, 17, 36, 72	0
1	C	227/235 (96%)	-0.26	5 (2%) 65 72	11, 18, 38, 67	0
1	D	225/235 (95%)	-0.23	3 (1%) 79 85	11, 23, 43, 60	0
1	E	226/235 (96%)	-0.15	4 (1%) 71 78	13, 21, 42, 75	0
1	F	228/235 (97%)	-0.26	4 (1%) 71 78	12, 22, 42, 74	0
1	G	225/235 (95%)	-0.27	4 (1%) 71 78	12, 21, 41, 55	0
1	H	224/235 (95%)	-0.36	1 (0%) 93 94	12, 19, 37, 50	0
1	I	226/235 (96%)	-0.24	1 (0%) 93 94	11, 19, 41, 59	0
1	J	224/235 (95%)	-0.26	0 100 100	13, 21, 36, 51	0
1	K	225/235 (95%)	-0.12	6 (2%) 58 64	12, 21, 45, 66	0
1	L	225/235 (95%)	-0.19	1 (0%) 93 94	14, 24, 44, 68	0
All	All	2708/2820 (96%)	-0.24	35 (1%) 79 85	10, 20, 41, 76	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	92[A]	ILE	8.3
1	C	183[A]	THR	6.1
1	B	220	ALA	4.9
1	F	230	HIS	3.8
1	B	219	ASP	3.6

### 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
3	XPE	D	303	24/31	0.58	0.32	12.40	70,84,96,96	0
3	XPE	J	302	21/31	0.66	0.40	5.89	69,87,95,96	0
3	XPE	G	302	27/31	0.68	0.22	5.49	42,55,66,70	0
3	XPE	E	302	24/31	0.84	0.19	3.94	37,52,79,80	0
3	XPE	C	301	24/31	0.81	0.20	3.82	36,47,63,67	0
3	XPE	H	302	24/31	0.80	0.21	3.63	33,46,72,73	0
3	XPE	D	302	26/31	0.90	0.15	2.64	21,35,55,58	0
3	XPE	L	302	27/31	0.88	0.14	2.54	29,41,61,62	0
3	XPE	B	302	27/31	0.83	0.16	2.40	24,38,58,61	0
2	MES	A	302	12/12	0.67	0.23	2.17	44,60,90,93	0
2	MES	F	301	12/12	0.81	0.17	1.80	48,60,86,92	0
2	MES	H	301	12/12	0.93	0.14	1.70	36,44,53,57	0
3	XPE	I	302	27/31	0.89	0.12	1.30	23,34,53,64	0
2	MES	K	301	12/12	0.87	0.17	1.22	53,64,79,83	0
2	MES	E	301	12/12	0.89	0.14	1.21	42,52,65,67	0
2	MES	D	301	12/12	0.82	0.13	1.09	43,52,75,77	0
2	MES	B	301	12/12	0.96	0.10	0.92	32,40,49,50	0
2	MES	J	301	12/12	0.94	0.12	0.70	30,40,45,52	0
2	MES	L	301	12/12	0.96	0.11	0.62	32,40,47,49	0
2	MES	G	301	12/12	0.84	0.16	0.52	39,52,80,83	0
2	MES	I	301	12/12	0.96	0.10	0.46	34,43,51,56	0
3	XPE	K	302	24/31	0.93	0.10	0.10	29,39,54,54	0
2	MES	A	301	12/12	0.97	0.08	-0.40	25,31,36,38	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.