



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:47 PM GMT

PDB ID : 4MRN  
Title : Structure of a bacterial Atm1-family ABC transporter  
Authors : Lee, J.Y.; Yang, J.G.; Zhitnitsky, D.; Lewinson, O.; Rees, D.C.  
Deposited on : 2013-09-17  
Resolution : 2.50 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

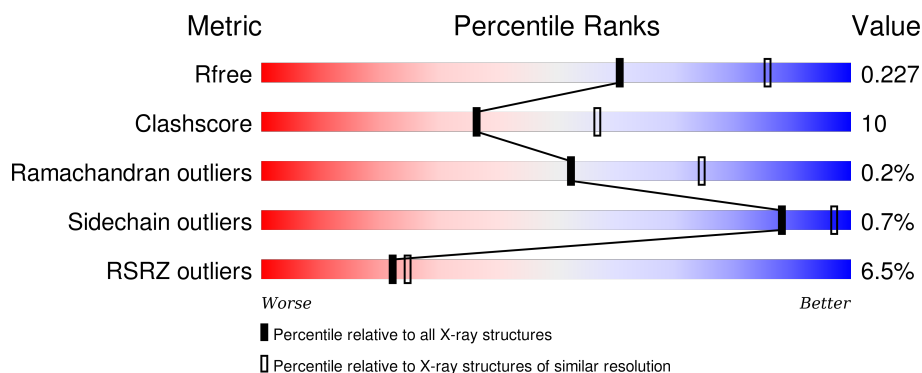
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	614	<div> <div>7%</div> <div>78%</div> <div>20%</div> <div>.</div> </div>
1	B	614	<div> <div>5%</div> <div>74%</div> <div>23%</div> <div>..</div> </div>

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	LDA	A	701	-	-	-	X
2	LDA	A	702	-	-	-	X
2	LDA	B	701	-	-	-	X
2	LDA	B	702	-	-	-	X
3	PO4	A	704	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9686 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.

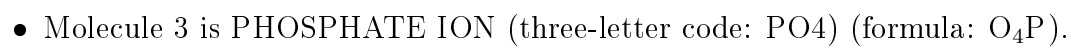
- Molecule 1 is a protein called ABC transporter related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	600	Total	C	N	O	Se	0	0	0
			4658	2970	828	848	12			
1	B	598	Total	C	N	O	Se	0	0	0
			4645	2963	825	845	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	609	HIS	-	EXPRESSION TAG	UNP Q2G506
A	610	HIS	-	EXPRESSION TAG	UNP Q2G506
A	611	HIS	-	EXPRESSION TAG	UNP Q2G506
A	612	HIS	-	EXPRESSION TAG	UNP Q2G506
A	613	HIS	-	EXPRESSION TAG	UNP Q2G506
A	614	HIS	-	EXPRESSION TAG	UNP Q2G506
B	609	HIS	-	EXPRESSION TAG	UNP Q2G506
B	610	HIS	-	EXPRESSION TAG	UNP Q2G506
B	611	HIS	-	EXPRESSION TAG	UNP Q2G506
B	612	HIS	-	EXPRESSION TAG	UNP Q2G506
B	613	HIS	-	EXPRESSION TAG	UNP Q2G506
B	614	HIS	-	EXPRESSION TAG	UNP Q2G506

- Molecule 2 is LAURYL DIMETHYLAMINE-N-OXIDE (three-letter code: LDA) (formula:  $C_{14}H_{31}NO$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	141	Total	O	0	0
			141	141		
4	B	148	Total	O	0	0
			148	148		

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 ( $\text{RSRZ} > 2$ ). 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.

Chain A:

Chain B:

Token	Color
R101	Green
M102	Green
E106	Green
R123	Green
L127	Green
A122	Green
R223	Green
A224	Green
V225	Green
D226	Green
S227	Green
L228	Green
K235	Green
V236	Green
R372	Green
D373	Green
R374	Green
S381	Green
A398	Green
S401	Green
L407	Green
F410	Green
W414	Green
I418	Green
R434	Green
V445	Green
L446	Green
D449	Green
E493	Green
R494	Green
L498	Green
L512	Green
V513	Green
K514	Green
L520	Green
F521	Green
E522	Green
E523	Green
A524	Green
T525	Green
S526	Green
A527	Green
L528	Green
R102	Green
E106	Green
R123	Green
L127	Green
A122	Green
R223	Green
A224	Green
V225	Green
D226	Green
S227	Green
L228	Green
K235	Green
V236	Green
R372	Green
D373	Green
R374	Green
S381	Green
A398	Green
S401	Green
L407	Green
F410	Green
W414	Green
I418	Green
R434	Green
V445	Green
L446	Green
D449	Green
E493	Green
R494	Green
L498	Green
L512	Green
V513	Green
K514	Green
L520	Green
F521	Green
E522	Green
E523	Green
A524	Green
T525	Green
S526	Green
A527	Green
L528	Green
G147	Green
I151	Green
D152	Green
T153	Green
L158	Green
L159	Green
F160	Green
A163	Green
P164	Green
T165	Green
E168	Green
L169	Green
T170	Green
A171	Green
V172	Green
I173	Green
V174	Green
I175	Green
F176	Green
W177	Green
L178	Green
G181	Green
L182	Green
T188	Green
A191	Green
W197	Green
T198	Green
T199	Green
R200	Green
T201	Green
I202	Green
W205	Green
R206	Green
R210	Green
P25	Green
P29	Green
A33	Green
R37	Green
R38	Green
V39	Green
I43	Green
K50	Green
L56	Green
P57	Green
Y60	Green
K61	Green
D65	Green
T68	Green
L69	Green
G70	Green
G71	Green
G72	Green
A73	Green
Q74	Green
P75	Green
L76	Green
A77	Green
T78	Green
V79	Green
A80	Green
L81	Green
N99	Green
L100	Green
D11	Red
A12	Red
R13	Red
H14	Red
L11	Red
R12	Red
L12	Red
R13	Red
H14	Red
L15	Red
R16	Red
L17	Red
A18	Red
T19	Red
L20	Red
K21	Red
R22	Red
P25	Red
P29	Red
A33	Red
R37	Red
R38	Red
V39	Red
I43	Red
K50	Red
L56	Red
P57	Red
Y60	Red
K61	Red
D65	Red
T68	Red
L69	Red
G70	Red
G71	Red
G72	Red
A73	Red
Q74	Red
P75	Red
L76	Red
A77	Red
T78	Red
V79	Red
A80	Red
L81	Red
N99	Red
L100	Red
D11	Grey
A12	Grey
R13	Grey
H14	Grey
L15	Grey
R16	Grey
L17	Grey
A18	Grey
T19	Grey
L20	Grey
K21	Grey
R22	Grey
P25	Grey
P29	Grey
A33	Grey
R37	Grey
R38	Grey
V39	Grey
I43	Grey
K50	





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	320.77Å 95.50Å 79.67Å 90.00° 101.84° 90.00°	Depositor
Resolution (Å)	39.80 – 2.50 39.80 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (39.80-2.50) 84.4 (39.80-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.200 , 0.227 0.197 , 0.227	Depositor DCC
$R_{free}$ test set	3433 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	60.2	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 49.8	EDS
Estimated twinning fraction	0.019 for -h-2*k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 80406 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9686	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, LDA

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	0.43	0/4732	0.60	0/6408
1	B	0.44	0/4719	0.64	0/6389
All	All	0.44	0/9451	0.62	0/12797

There are no bond length outliers.

There are no bond angle outliers.

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	4658	0	4733	90	1
1	B	4645	0	4723	106	1
2	A	32	0	62	4	0
2	B	32	0	62	5	0
3	A	15	0	0	2	0
3	B	15	0	0	2	0
4	A	141	0	0	13	1
4	B	148	0	0	6	1
All	All	9686	0	9580	181	2

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 10.

The worst 5 of 181 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:147:GLY:HA3	2:A:701:LDA:H51	1.60	0.83
1:B:132:HIS:HB2	4:B:913:HOH:O	1.82	0.80
1:A:581:ASP:OD2	4:A:871:HOH:O	2.01	0.79
3:A:704:PO4:O2	4:A:932:HOH:O	2.01	0.78
1:B:61:LYS:NZ	1:B:300:ASP:OD1	2.18	0.76

All (2) 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:591:GLU:OE1	1:B:13:ARG:NH2[4_547]	2.13	0.07
4:A:928:HOH:O	4:B:927:HOH:O[4_547]	2.17	0.03

## 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	598/614 (97%)	589 (98%)	9 (2%)	0	100	100
1	B	596/614 (97%)	583 (98%)	11 (2%)	2 (0%)	46	68
All	All	1194/1228 (97%)	1172 (98%)	20 (2%)	2 (0%)	52	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	74	GLN
1	B	131	PHE

### 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	478/478 (100%)	476 (100%)	2 (0%)	93	98
1	B	477/478 (100%)	472 (99%)	5 (1%)	82	95
All	All	955/956 (100%)	948 (99%)	7 (1%)	88	97

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	130	ARG
1	B	445	VAL
1	B	215	ARG
1	A	307	TYR
1	B	307	TYR

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](#)

10 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	LDA	A	701	-	15,15,15	2.79	3 (20%)	16,17,17	0.73	0
2	LDA	A	702	-	15,15,15	3.09	3 (20%)	16,17,17	1.21	2 (12%)
3	PO4	A	703	-	4,4,4	0.45	0	6,6,6	0.28	0
3	PO4	A	704	-	4,4,4	0.55	0	6,6,6	0.28	0
3	PO4	A	705	-	4,4,4	0.50	0	6,6,6	0.27	0
2	LDA	B	701	-	15,15,15	2.92	3 (20%)	16,17,17	0.83	0
2	LDA	B	702	-	15,15,15	2.82	3 (20%)	16,17,17	0.80	0
3	PO4	B	703	-	4,4,4	0.47	0	6,6,6	0.27	0
3	PO4	B	704	-	4,4,4	0.43	0	6,6,6	0.27	0
3	PO4	B	705	-	4,4,4	0.47	0	6,6,6	0.28	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	LDA	A	701	-	-	0/13/13/13	0/0/0/0
2	LDA	A	702	-	-	0/13/13/13	0/0/0/0
3	PO4	A	703	-	-	0/0/0/0	0/0/0/0
3	PO4	A	704	-	-	0/0/0/0	0/0/0/0
3	PO4	A	705	-	-	0/0/0/0	0/0/0/0
2	LDA	B	701	-	-	0/13/13/13	0/0/0/0
2	LDA	B	702	-	-	0/13/13/13	0/0/0/0
3	PO4	B	703	-	-	0/0/0/0	0/0/0/0
3	PO4	B	704	-	-	0/0/0/0	0/0/0/0
3	PO4	B	705	-	-	0/0/0/0	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	702	LDA	O1-N1	-10.69	1.29	1.39
2	B	701	LDA	O1-N1	-10.22	1.29	1.39
2	B	702	LDA	O1-N1	-9.85	1.30	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	LDA	O1-N1	-9.69	1.30	1.39
2	A	702	LDA	CM1-N1	-4.15	1.43	1.49

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	702	LDA	O1-N1-C1	-2.79	107.13	110.27
2	A	702	LDA	CM2-N1-CM1	-2.27	106.27	108.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	LDA	4	0
3	A	704	PO4	2	0
2	B	701	LDA	1	0
2	B	702	LDA	4	0
3	B	703	PO4	1	0
3	B	705	PO4	1	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	588/614 (95%)	0.38	45 (7%)	16 18	31, 65, 107, 149	0
1	B	586/614 (95%)	0.50	31 (5%)	30 34	24, 59, 106, 189	0
All	All	1174/1228 (95%)	0.44	76 (6%)	22 25	24, 62, 108, 189	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	72	GLY	35.3
1	B	71	GLY	17.6
1	B	73	ALA	17.3
1	B	70	GLY	12.3
1	B	74	GLN	9.4

### 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( $\text{\AA}^2$ )	Q<0.9
2	LDA	A	701	16/16	0.82	0.66	15.63	50,65,110,118	0
2	LDA	B	701	16/16	0.94	0.66	10.18	57,75,87,94	0
2	LDA	A	702	16/16	0.95	0.47	6.94	49,59,75,79	0
2	LDA	B	702	16/16	0.89	0.36	4.81	46,61,99,104	0
3	PO4	B	703	5/5	0.85	0.19	1.00	116,118,119,123	0
3	PO4	A	704	5/5	0.94	0.13	-0.58	83,87,88,92	5
3	PO4	B	704	5/5	0.83	0.11	-	122,127,128,129	0
3	PO4	A	705	5/5	0.80	0.37	-	149,151,153,156	0
3	PO4	B	705	5/5	0.68	0.50	-	141,141,142,145	0
3	PO4	A	703	5/5	0.72	0.12	-	157,158,158,159	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.