



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2022 – 02:04 PM EST

PDB ID : 6VQZ  
Title : HLA-B\*27:05 presenting an HIV-1 6mer peptide  
Authors : Pymm, P.; Tenzer, S.; Wee, E.; Weimershaus, M.; Burgevin, A.; Kollnberger, S.; Gerstoft, J.; Josephs, T.M.; Ladell, K.; McLaren, J.E.; Appay, V.; Price, D.A.; Fugger, L.; Bell, J.I.; Hansjorg, S.; Van Endert, P.; Harkiolaki, M.; Iversen, A.K.N.  
Deposited on : 2020-02-06  
Resolution : 2.25 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.27  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

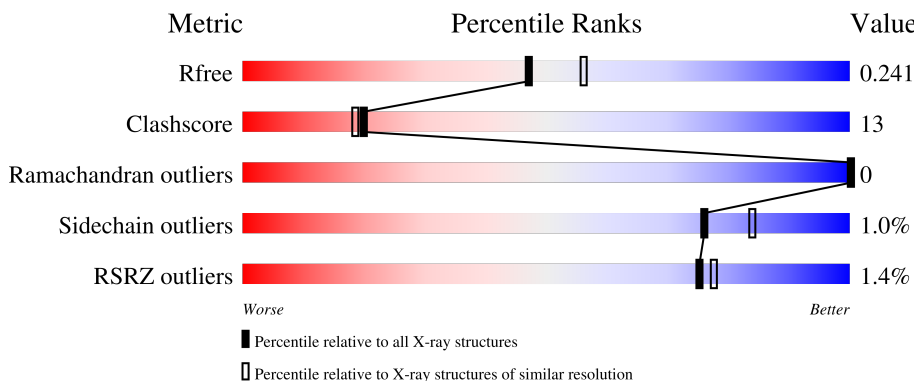
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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))
$R_{free}$	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	276	<div> <div>79%</div> <div>21%</div> </div>
1	C	276	<div> <div>3%</div> <div>68%</div> <div>30%</div> <div>.</div> </div>
2	B	99	<div> <div>65%</div> <div>35%</div> </div>
2	D	99	<div> <div>2%</div> <div>71%</div> <div>28%</div> <div>.</div> </div>
3	F	6	<div> <div>83%</div> <div>17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	6	 100%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6694 atoms, of which 169 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	1	5	0
			2290	1423	422	437	8			
1	C	275	Total	C	N	O	S	0	5	0
			2289	1422	417	442	8			

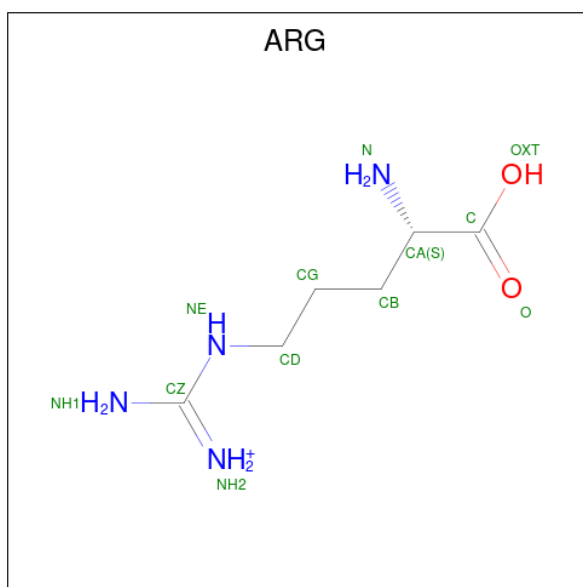
- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	3	0
			856	544	144	165	3			
2	D	99	Total	C	N	O	S	0	1	0
			837	533	143	158	3			

- Molecule 3 is a protein called 6-mer peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	6	Total	C	H	N	O	0	0	0
			129	41	71	11	6			
3	F	6	Total	C	H	N	O	0	0	0
			129	41	71	11	6			

- Molecule 4 is ARGinine (three-letter code: ARG) (formula: C<sub>6</sub>H<sub>15</sub>N<sub>4</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			23	6	12	4	1		
4	C	1	Total	C	H	N	O	0	0
			26	6	15	4	1		

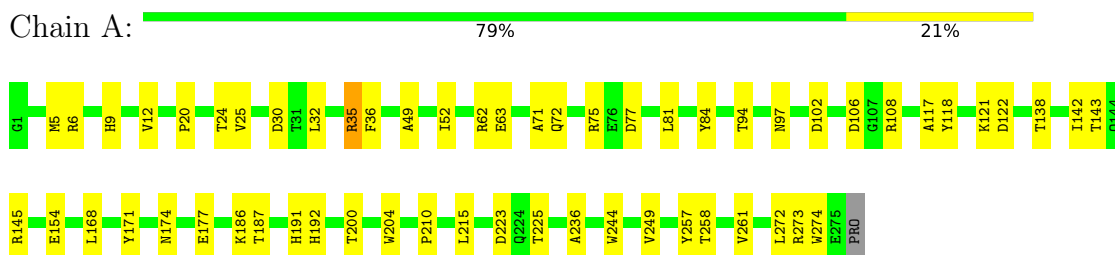
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	51	Total	O	0	0
			51	51		
5	B	11	Total	O	0	0
			11	11		
5	C	38	Total	O	0	0
			38	38		
5	D	14	Total	O	0	0
			14	14		
5	F	1	Total	O	0	0
			1	1		

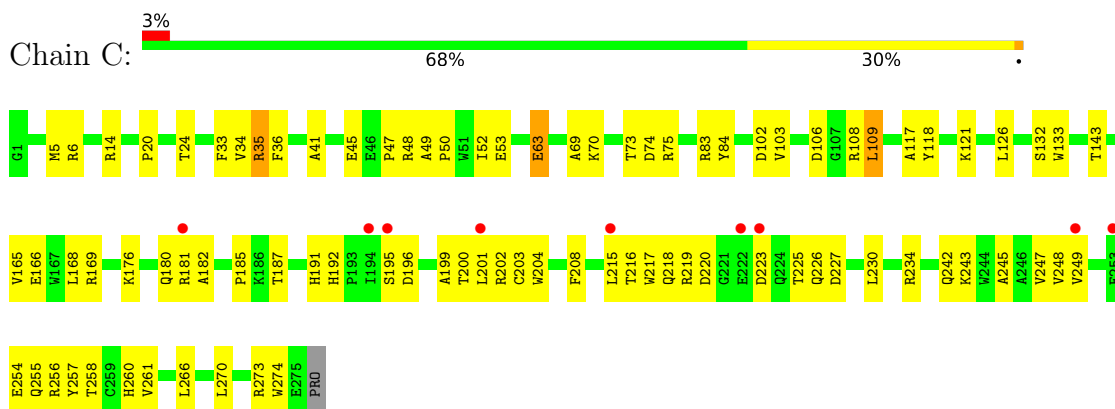
### 3 Residue-property plots

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

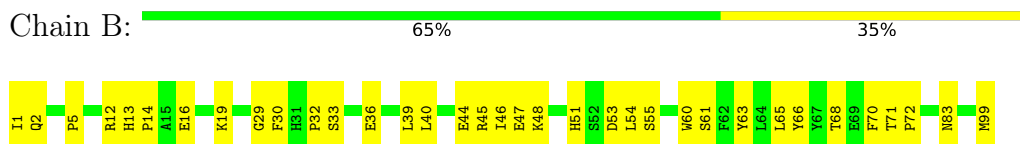
#### • Molecule 1: MHC class I antigen



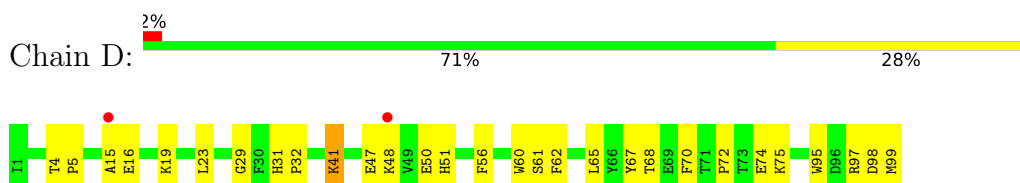
#### • Molecule 1: MHC class I antigen



#### • Molecule 2: Beta-2-microglobulin



#### • Molecule 2: Beta-2-microglobulin




- Molecule 3: 6-mer peptide

Chain G:  100%

There are no outlier residues recorded for this chain.

- Molecule 3: 6-mer peptide

Chain F:  83% 17%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.49Å 130.52Å 90.17Å 90.00° 104.41° 90.00°	Depositor
Resolution (Å)	87.33 – 2.25 87.33 – 2.25	Depositor EDS
% Data completeness (in resolution range)	97.9 (87.33-2.25) 96.9 (87.33-2.25)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.43 (at 2.25Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.201 , 0.244 0.201 , 0.241	Depositor DCC
$R_{free}$ test set	2259 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtriage
Anisotropy	0.570	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 15.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.427 for h,-k,-h-l	Xtriage
Reported twinning fraction	0.440 for h,-k,-h-l	Depositor
Outliers	0 of 47259 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6694	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.46% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.28	0/2357	0.48	0/3201
1	C	0.27	0/2353	0.45	0/3197
2	B	0.28	0/879	0.44	0/1187
2	D	0.30	0/863	0.43	0/1166
3	F	0.39	0/59	0.48	0/78
3	G	0.38	0/59	0.57	0/78
All	All	0.28	0/6570	0.46	0/8907

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 [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	2290	0	2160	50	1
1	C	2289	0	2145	70	1
2	B	856	0	816	28	0
2	D	837	0	807	23	0
3	F	58	71	71	3	0
3	G	58	71	71	0	0
4	A	11	12	12	3	0
4	C	11	15	12	2	0
5	A	51	0	0	2	0
5	B	11	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	38	0	0	1	0
5	D	14	0	0	0	0
5	F	1	0	0	0	0
All	All	6525	169	6094	158	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 13.

The worst 5 of 158 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:106:ASP:OD2	1:A:108:ARG:NH1	1.93	1.02
1:C:258:THR:HG22	1:C:273:ARG:HD3	1.49	0.95
1:C:45:GLU:OE2	1:C:63:GLU:HG2	1.68	0.94
1:C:200:THR:HG22	1:C:248:VAL:HG12	1.50	0.91
1:A:223:ASP:OD2	1:A:225:THR:HG23	1.70	0.91

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:154:GLU:OE1	1:C:83:ARG:NH1[1_454]	2.05	0.15

## 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	278/276 (101%)	272 (98%)	6 (2%)	0	100	100
1	C	278/276 (101%)	273 (98%)	5 (2%)	0	100	100
2	B	100/99 (101%)	100 (100%)	0	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	98/99 (99%)	97 (99%)	1 (1%)	0	100	100
3	F	4/6 (67%)	4 (100%)	0	0	100	100
3	G	4/6 (67%)	4 (100%)	0	0	100	100
All	All	762/762 (100%)	750 (98%)	12 (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	239/235 (102%)	238 (100%)	1 (0%)	91	94
1	C	239/235 (102%)	236 (99%)	3 (1%)	69	79
2	B	97/94 (103%)	96 (99%)	1 (1%)	76	84
2	D	95/94 (101%)	93 (98%)	2 (2%)	53	62
3	F	6/6 (100%)	6 (100%)	0	100	100
3	G	6/6 (100%)	6 (100%)	0	100	100
All	All	682/670 (102%)	675 (99%)	7 (1%)	76	84

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

Mol	Chain	Res	Type
1	C	63	GLU
1	C	109	LEU
2	D	70	PHE
2	D	41	LYS
1	C	35	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	GLN
1	A	97	ASN
1	C	218	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 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	$\# Z  > 2$
4	ARG	A	301	-	9,10,11	0.43	0	5,11,13	0.60	0
4	ARG	C	301	-	9,10,11	0.39	0	5,11,13	0.25	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
4	ARG	A	301	-	-	0/8/9/11	-
4	ARG	C	301	-	-	0/8/9/11	-

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.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	301	ARG	3	0
4	C	301	ARG	2	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	275/276 (99%)	-0.37	0 <span>100</span> <span>100</span>	17, 28, 55, 103	4 (1%)
1	C	275/276 (99%)	0.01	9 (3%) <span>46</span> <span>48</span>	19, 34, 114, 166	5 (1%)
2	B	99/99 (100%)	-0.35	0 <span>100</span> <span>100</span>	17, 30, 50, 69	0
2	D	99/99 (100%)	-0.10	2 (2%) <span>65</span> <span>68</span>	22, 39, 78, 96	1 (1%)
3	F	6/6 (100%)	-0.65	0 <span>100</span> <span>100</span>	22, 28, 34, 35	0
3	G	6/6 (100%)	-0.58	0 <span>100</span> <span>100</span>	22, 24, 32, 34	0
All	All	760/762 (99%)	-0.20	11 (1%) <span>75</span> <span>77</span>	17, 31, 83, 166	10 (1%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	223	ASP	8.7
1	C	222	GLU	4.9
1	C	249	VAL	4.8
1	C	194	ILE	4.6
1	C	195	SER	4.4

### 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 monosaccharides 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ARG	C	301	11/12	0.89	0.18	24,33,41,54	0
4	ARG	A	301	11/12	0.93	0.16	23,33,43,43	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.