W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 30, 2023 – 07:17 pm BST

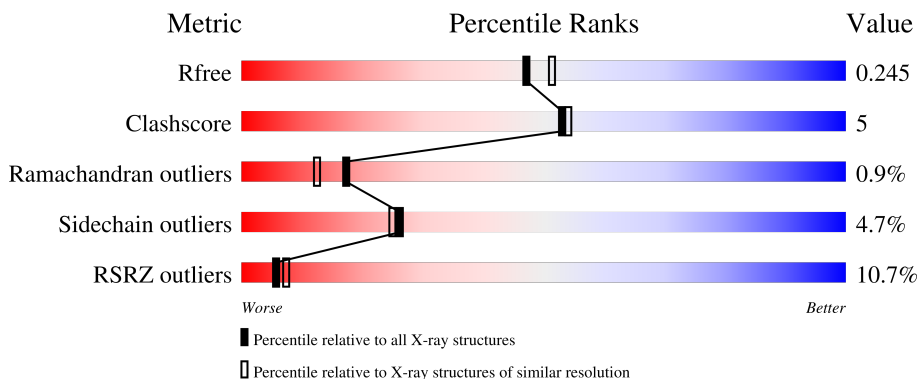|		
PDB ID	:	7Q9A
Title	:	MHC Class I A02 Allele presenting LLLGIGILVL, in complex with Mel5 TCR
Authors	:	Rizkallah, P.J.; Sewell, A.K.; Wall, A.; Fuller, A.
Deposited on	:	2021-11-12
Resolution	:	2.10 Å(reported)

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 2.10 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	Percentile Ranks	Value
Rfree		0.245
Clashscore		5
Ramachandran outliers		0.9%
Sidechain outliers		4.7%
RSRZ outliers		10.7%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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 <=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	20% / 81% / 18% ·
2	B	100	13% / 84% / 14% ·
3	C	10	80% / 20%
4	D	199	7% / 83% / 15% ·
5	E	244	3% / 84% / 16%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 6911 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total 2265	C 1414	N 414	O 428	S 9	0	1	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total 837	C 533	N 141	O 159	S 4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called LEU-LEU-LEU-GLY-ILE-GLY-ILE-LEU-VAL-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total 72	C 51	N 10	O 11	0	0	0

- Molecule 4 is a protein called Human Mel5 T Cell Receptor, Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	199	Total 1549	C 963	N 255	O 323	S 8	0	1	0

- Molecule 5 is a protein called Human Mel5 T Cell Receptor, Beta Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	244	Total 1936	C 1227	N 333	O 371	S 5	0	1	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 6	C 3	O 3	0	0
6	A	1	Total 6	C 3	O 3	0	0
6	D	1	Total 6	C 3	O 3	0	0
6	E	1	Total 6	C 3	O 3	0	0

- Molecule 7 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	84	Total 84	O 84	0	0
7	B	23	Total 23	O 23	0	0
7	C	3	Total 3	O 3	0	0
7	D	51	Total 51	O 51	0	0
7	E	67	Total 67	O 67	0	0

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.
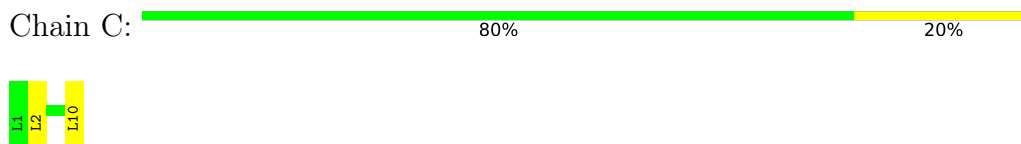
- Molecule 1: MHC class I antigen
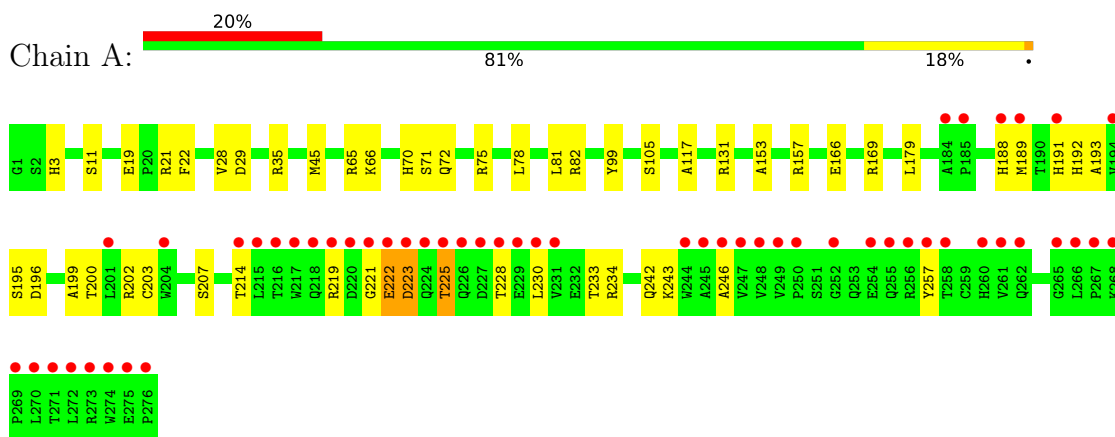
- Molecule 2: Beta-2-microglobulin
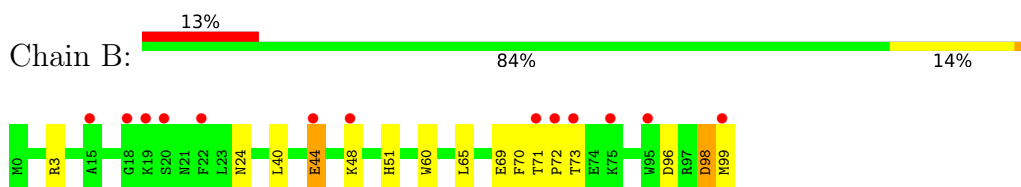
- Molecule 3: LEU-LEU-LEU-GLY-ILE-GLY-ILE-LEU-VAL-LEU

- Molecule 4: Human Mel5 T Cell Receptor, Alpha Chain

C182 A183 N187 N188 S189 I190 I191 P192 E193 D194 P200
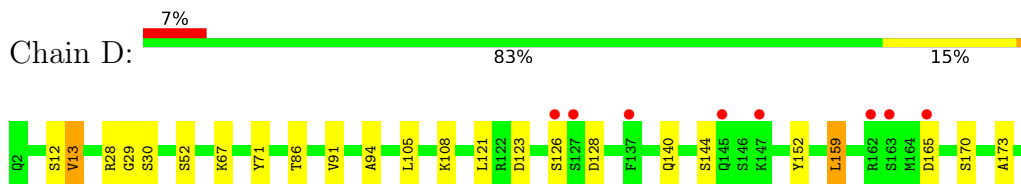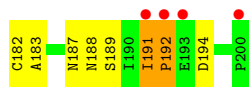
• Molecule 5: Human Mel5 T Cell Receptor, Beta Chain

Chain E:

3%

84% 16%

S1 Q2 T3 I4 H5 Q6 W7 L11 V25 N30 P31 N32 L33 Q38 R42 G43 L44 V60 P61 R72 Q73 F74 L83 W93 E102 T112 E115 E129 E132 I135 Q139 D153 H154 V155 S168 C171 L177 L183

Y188 R193 L194 R195 A198 Q213 E219 N220 T224 Q225 D226 K229 D244
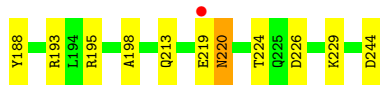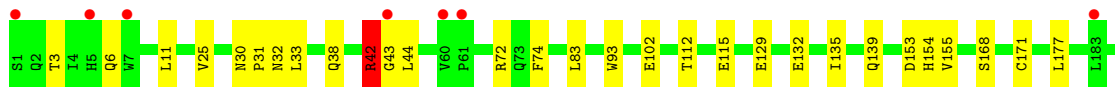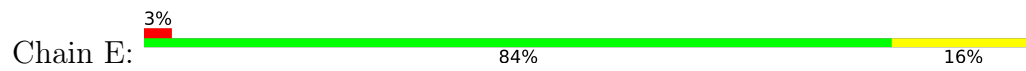
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α, β, γ	122.02Å 122.02Å 82.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	61.01 – 2.10 61.01 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (61.01-2.10) 99.8 (61.01-2.10)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.00 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.204 , 0.243 0.211 , 0.245	Depositor DCC
R_{free} test set	3455 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	53.1	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 38.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6911	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: GOL

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	0.73	0/2331	0.92	4/3163 (0.1%)
2	B	0.73	0/860	0.93	2/1162 (0.2%)
3	C	0.77	0/71	0.84	0/94
4	D	0.74	0/1582	0.97	0/2146
5	E	0.72	1/1990 (0.1%)	0.92	2/2713 (0.1%)
All	All	0.73	1/6834 (0.0%)	0.93	8/9278 (0.1%)

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
4	D	0	1
5	E	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	102	GLU	CD-OE2	5.07	1.31	1.25

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	3	ARG	NE-CZ-NH2	-6.94	116.83	120.30
2	B	3	ARG	NE-CZ-NH1	6.66	123.63	120.30
5	E	6	GLN	CB-CA-C	6.40	123.20	110.40
5	E	42	ARG	CG-CD-NE	-6.07	99.06	111.80
1	A	157	ARG	NE-CZ-NH1	-6.03	117.29	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	191	ILE	Peptide
5	E	43	GLY	Peptide

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	2265	0	2115	27	0
2	B	837	0	803	8	0
3	C	72	0	94	2	0
4	D	1549	0	1452	25	0
5	E	1936	0	1851	18	0
6	A	12	0	16	0	0
6	D	6	0	8	1	0
6	E	6	0	8	0	0
7	A	84	0	0	2	0
7	B	23	0	0	0	0
7	C	3	0	0	0	0
7	D	51	0	0	1	0
7	E	67	0	0	0	0
All	All	6911	0	6347	70	0

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 5.

The worst 5 of 70 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:166:GLU:OE2	4:D:28:ARG:NH1	2.19	0.76
1:A:193:ALA:HB1	1:A:199:ALA:HA	1.76	0.66
1:A:72:GLN:NE2	7:A:401:HOH:O	2.29	0.65
4:D:86:THR:HG21	5:E:42:ARG:HG2	1.78	0.64
4:D:123:ASP:HB3	4:D:126:SER:O	1.98	0.64

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	275/276 (100%)	255 (93%)	17 (6%)	3 (1%)	14	9
2	B	98/100 (98%)	95 (97%)	3 (3%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
4	D	198/199 (100%)	184 (93%)	10 (5%)	4 (2%)	7	3
5	E	243/244 (100%)	231 (95%)	12 (5%)	0	100	100
All	All	822/829 (99%)	773 (94%)	42 (5%)	7 (1%)	17	12

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	192	HIS
1	A	222	GLU
4	D	192	PRO
1	A	223	ASP
4	D	128	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	233/232 (100%)	221 (95%)	12 (5%)	23	21
2	B	95/95 (100%)	89 (94%)	6 (6%)	18	15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	8/8 (100%)	8 (100%)	0	100	100
4	D	178/177 (101%)	173 (97%)	5 (3%)	43	47
5	E	212/211 (100%)	201 (95%)	11 (5%)	23	21
All	All	726/723 (100%)	692 (95%)	34 (5%)	26	25

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	171	CYS
5	E	193	ARG
5	E	226	ASP
2	B	44	GLU
1	A	230	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	176	ASN
5	E	38	GLN
5	E	220	ASN
5	E	207	HIS
1	A	188	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 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	$\#|Z| > 2$
6	GOL	D	301	-	5,5,5	0.19	0	5,5,5	0.46	0
6	GOL	A	301	-	5,5,5	0.10	0	5,5,5	0.29	0
6	GOL	E	301	-	5,5,5	0.11	0	5,5,5	0.32	0
6	GOL	A	302	-	5,5,5	0.08	0	5,5,5	0.26	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
6	GOL	D	301	-	-	2/4/4/4	-
6	GOL	A	301	-	-	2/4/4/4	-
6	GOL	E	301	-	-	2/4/4/4	-
6	GOL	A	302	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	301	GOL	O1-C1-C2-C3
6	A	302	GOL	O1-C1-C2-O2
6	A	302	GOL	O1-C1-C2-C3
6	D	301	GOL	C1-C2-C3-O3
6	D	301	GOL	O2-C2-C3-O3

There are no ring outliers.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	301	GOL	1	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 ⓘ

6.1   Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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^2)	Q<0.9
1	A	276/276 (100%)	1.55	54 (19%)	1	1	36, 60, 133, 149	0
2	B	100/100 (100%)	0.96	13 (13%)	3	4	37, 68, 117, 131	0
3	C	10/10 (100%)	0.99	0	100	100	37, 41, 46, 46	0
4	D	199/199 (100%)	0.86	14 (7%)	16	20	38, 59, 117, 131	0
5	E	244/244 (100%)	0.62	8 (3%)	46	53	38, 57, 100, 117	0
All	All	829/829 (100%)	1.03	89 (10%)	6	7	36, 59, 120, 149	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	228	THR	16.1
1	A	273	ARG	10.8
1	A	260	HIS	10.1
1	A	227	ASP	10.1
1	A	230	LEU	9.6

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

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

6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95th 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(Å^2)	Q<0.9
6	GOL	D	301	6/6	0.79	0.36	77,79,83,88	0
6	GOL	E	301	6/6	0.81	0.19	61,75,76,76	0
6	GOL	A	301	6/6	0.84	0.14	83,86,90,96	0
6	GOL	A	302	6/6	0.87	0.15	75,76,79,80	0

6.5 Other polymers ⓘ

There are no such residues in this entry.