Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2024 – 07:14 pm BST

PDB ID	:	8RYM
Title	:	Structure of S2 TCR in complex with HLA-A*03:01 bound to ELFSYLIEK peptide
Authors	:	Karuppiah, V.
Deposited on	:	2024-02-09
Resolution	:	2.34 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

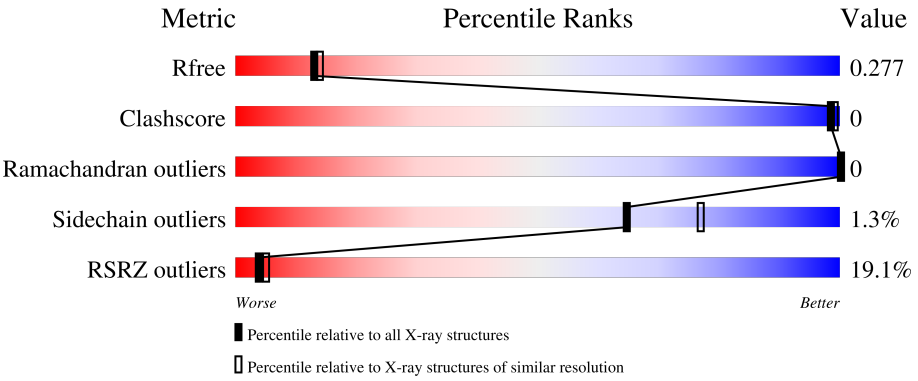
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.34 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}	164625	2747 (2.36-2.32)
Clashscore	180529	2936 (2.36-2.32)
Ramachandran outliers	177936	2912 (2.36-2.32)
Sidechain outliers	177891	2912 (2.36-2.32)
RSRZ outliers	164620	2747 (2.36-2.32)

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	7% 98%
2	B	100	% 94%
3	C	9	89% 11%
4	D	198	25% 86% 10%
5	E	245	34% 96%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6430 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 HLA class I histocompatibility antigen, A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2244	1396	405	434	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	276	PRO	-	expression tag	UNP P04439

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	98	Total	C	N	O	S	0	0	0
			821	522	139	157	3			

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 ELFSYLIEK peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			81	55	10	16			

- Molecule 4 is a protein called TCR alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	178	Total	C	N	O	S	0	0	0
			1349	834	233	274	8			

- Molecule 5 is a protein called TCR beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	238	Total	C	N	O	S	0	0	0
			1875	1184	323	359	9			

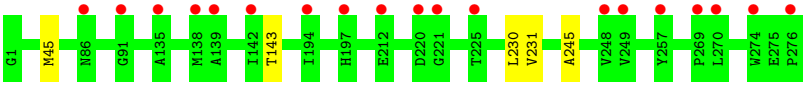
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	24	Total	O	0	0
			24	24		
6	B	22	Total	O	0	0
			22	22		
6	C	2	Total	O	0	0
			2	2		
6	D	5	Total	O	0	0
			5	5		
6	E	7	Total	O	0	0
			7	7		

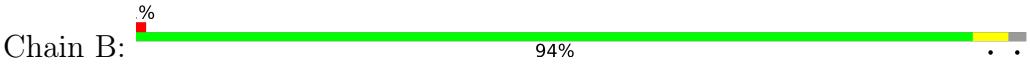
3 Residue-property plots (i)

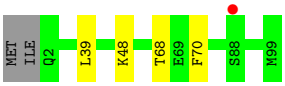
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: HLA class I histocompatibility antigen, A alpha chain

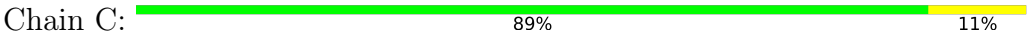


• Molecule 2: Beta-2-microglobulin

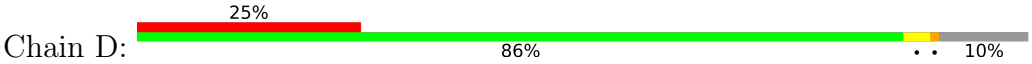


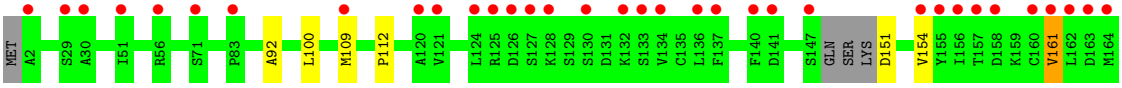


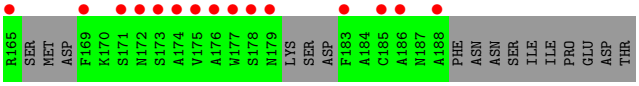
• Molecule 3: ELFSYLIEK peptide



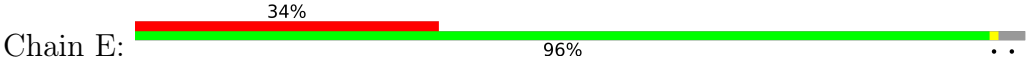
• Molecule 4: TCR alpha

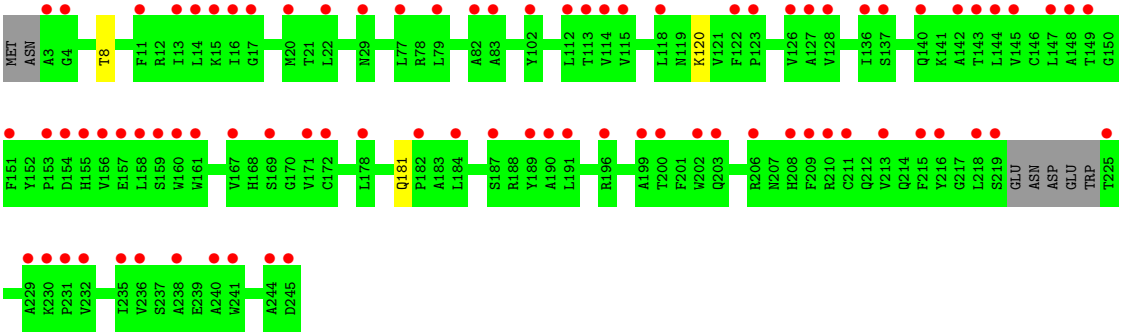






• Molecule 5: TCR beta




MET ASN A3 G4 T8 F11 R12 I13 L14 K15 I16 G17 M20 T21 L22 N29 L77 R78 L79 A82 A83 Y102 L112 T113 V114 V115 L118 N119 K120 V121 F122 P123 V126 A127 V128 I136 S137 Q140 K141 A142 T143 L144 V145 C146 L147 A148 T149 G150
F151 Y152 P153 D154 H155 V156 E157 L158 S159 W160 W161 V167 H168 S169 G170 V171 C172 L178 Q181 P182 A183 L184 S187 R188 Y189 A190 L191 H195 A199 T200 F201 W202 Q203 R206 N207 H208 P209 R210 C211 Q212 V213 Q214 P215 Y216 G217 L218 S219 GLU ASN ASP GLU TRP T225
A229 K230 P231 V232 I235 V236 S237 A238 E239 A240 W241 A244 D245

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.66Å 49.20Å 124.67Å 90.00° 92.93° 90.00°	Depositor
Resolution (Å)	41.50 – 2.34 41.50 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.8 (41.50-2.34) 99.8 (41.50-2.34)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.76 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.243 , 0.272 0.252 , 0.277	Depositor DCC
R_{free} test set	2145 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	54.9	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 36.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6430	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% 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

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	0.64	0/2305	0.70	0/3130
2	B	0.62	0/844	0.69	0/1141
3	C	0.64	0/82	0.66	0/107
4	D	0.67	0/1369	0.72	0/1854
5	E	0.65	0/1925	0.71	0/2620
All	All	0.65	0/6525	0.71	0/8852

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	2244	0	2093	2	0
2	B	821	0	780	1	0
3	C	81	0	83	1	0
4	D	1349	0	1299	3	0
5	E	1875	0	1792	0	0
6	A	24	0	0	0	0
6	B	22	0	0	0	0
6	C	2	0	0	0	0
6	D	5	0	0	0	0
6	E	7	0	0	0	0
All	All	6430	0	6047	6	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 0.

All (6) 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)
4:D:112:PRO:HG3	4:D:161:VAL:HG23	1.94	0.49
4:D:151:ASP:HB3	4:D:154:VAL:HG12	1.97	0.47
1:A:230:LEU:HD23	1:A:245:ALA:HB2	1.97	0.46
1:A:143:THR:HG23	3:C:9:LYS:HA	1.99	0.43
2:B:39:LEU:HD13	2:B:68:THR:HG22	2.00	0.43
4:D:92:ALA:HB1	4:D:100:LEU:HD11	2.01	0.43

There are no symmetry-related clashes.

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	274/276 (99%)	266 (97%)	8 (3%)	0	100	100
2	B	96/100 (96%)	93 (97%)	3 (3%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
4	D	170/198 (86%)	163 (96%)	7 (4%)	0	100	100
5	E	234/245 (96%)	225 (96%)	9 (4%)	0	100	100
All	All	781/828 (94%)	754 (96%)	27 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	233/233 (100%)	231 (99%)	2 (1%)	75 85
2	B	93/95 (98%)	91 (98%)	2 (2%)	47 58
3	C	9/9 (100%)	9 (100%)	0	100 100
4	D	152/172 (88%)	150 (99%)	2 (1%)	65 77
5	E	204/211 (97%)	201 (98%)	3 (2%)	60 72
All	All	691/720 (96%)	682 (99%)	9 (1%)	65 77

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	MET
1	A	231	VAL
2	B	48	LYS
2	B	70	PHE
4	D	109	MET
4	D	161	VAL
5	E	8	THR
5	E	120	LYS
5	E	181	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	86	ASN
1	A	218	GLN
4	D	35	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

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, 95^{th} 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%)	0.73	19 (6%) 24 30	41, 61, 100, 115	0
2	B	98/100 (98%)	0.29	1 (1%) 79 83	41, 54, 72, 76	0
3	C	9/9 (100%)	0.76	0 100 100	55, 58, 64, 68	0
4	D	178/198 (89%)	1.49	49 (27%) 2 2	50, 82, 121, 137	0
5	E	238/245 (97%)	1.67	84 (35%) 1 1	53, 85, 120, 154	0
All	All	799/828 (96%)	1.13	153 (19%) 4 5	41, 70, 114, 154	0

All (153) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	188	ALA	5.2
4	D	169	PHE	5.1
4	D	183	PHE	5.0
4	D	161	VAL	4.5
5	E	225	THR	4.5
4	D	165	ARG	4.4
5	E	16	ILE	4.4
4	D	137	PHE	4.4
5	E	147	LEU	4.3
5	E	158	LEU	4.2
5	E	191	LEU	4.0
4	D	156	ILE	4.0
5	E	160	TRP	3.9
5	E	235	ILE	3.8
1	A	197	HIS	3.8
1	A	248	VAL	3.8
5	E	82	ALA	3.7
1	A	276	PRO	3.7
5	E	208	HIS	3.6
4	D	186	ALA	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	249	VAL	3.6
5	E	128	VAL	3.6
4	D	30	ALA	3.6
5	E	199	ALA	3.6
5	E	218	LEU	3.5
5	E	3	ALA	3.5
5	E	102	TYR	3.4
5	E	184	LEU	3.4
5	E	238	ALA	3.4
4	D	136	LEU	3.3
4	D	177	TRP	3.3
5	E	148	ALA	3.3
4	D	175	VAL	3.3
4	D	160	CYS	3.3
5	E	244	ALA	3.3
5	E	112	LEU	3.2
5	E	219	SER	3.2
1	A	91	GLY	3.2
4	D	164	MET	3.1
5	E	11	PHE	3.1
5	E	118	LEU	3.1
5	E	178	LEU	3.1
5	E	155	HIS	3.1
4	D	154	VAL	3.1
5	E	240	ALA	3.1
5	E	206	ARG	3.1
5	E	13	ILE	3.0
5	E	126	VAL	3.0
4	D	147	SER	3.0
5	E	245	ASP	3.0
5	E	200	THR	2.9
4	D	130	SER	2.9
4	D	173	SER	2.9
4	D	157	THR	2.8
5	E	22	LEU	2.8
1	A	269	PRO	2.8
1	A	86	ASN	2.8
4	D	155	TYR	2.8
5	E	190	ALA	2.8
5	E	113	THR	2.8
4	D	176	ALA	2.8
5	E	241	TRP	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	132	LYS	2.7
5	E	123	PRO	2.7
5	E	145	VAL	2.7
5	E	213	VAL	2.7
5	E	15	LYS	2.7
5	E	122	PHE	2.7
1	A	139	ALA	2.7
5	E	127	ALA	2.7
1	A	221	GLY	2.7
1	A	212	GLU	2.7
5	E	136	ILE	2.7
4	D	133	SER	2.7
5	E	209	PHE	2.7
5	E	232	VAL	2.7
5	E	169	SER	2.7
1	A	194	ILE	2.6
4	D	2	ALA	2.6
5	E	189	TYR	2.6
5	E	115	VAL	2.6
5	E	17	GLY	2.6
1	A	138	MET	2.6
4	D	174	ALA	2.6
5	E	171	VAL	2.6
4	D	83	PRO	2.6
5	E	137	SER	2.6
5	E	172	CYS	2.6
5	E	203	GLN	2.6
5	E	114	VAL	2.5
4	D	172	ASN	2.5
5	E	187	SER	2.5
5	E	83	ALA	2.5
5	E	156	VAL	2.5
5	E	215	PHE	2.5
5	E	154	ASP	2.4
4	D	128	LYS	2.4
5	E	144	LEU	2.4
5	E	210	ARG	2.4
4	D	185	CYS	2.4
5	E	140	GLN	2.4
5	E	14	LEU	2.4
4	D	140	PHE	2.4
4	D	179	ASN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	E	167	VAL	2.4
5	E	236	VAL	2.4
4	D	162	LEU	2.4
5	E	229	ALA	2.4
1	A	274	TRP	2.3
4	D	171	SER	2.3
4	D	120	ALA	2.3
5	E	149	THR	2.3
5	E	231	PRO	2.3
5	E	161	TRP	2.3
4	D	158	ASP	2.3
5	E	211	CYS	2.3
5	E	196	ARG	2.3
4	D	126	ASP	2.3
1	A	257	TYR	2.3
5	E	153	PRO	2.3
5	E	77	LEU	2.2
1	A	220	ASP	2.2
5	E	4	GLY	2.2
4	D	124	LEU	2.2
4	D	134	VAL	2.2
5	E	142	ALA	2.2
5	E	143	THR	2.2
4	D	109	MET	2.2
5	E	216	TYR	2.2
4	D	121	VAL	2.2
5	E	151	PHE	2.2
1	A	142	ILE	2.1
4	D	56	ARG	2.1
4	D	125	ARG	2.1
5	E	29	ASN	2.1
5	E	157	GLU	2.1
4	D	163	ASP	2.1
4	D	71	SER	2.1
4	D	141	ASP	2.1
5	E	182	PRO	2.1
1	A	225	THR	2.1
4	D	29	SER	2.0
4	D	178	SER	2.0
4	D	51	ILE	2.0
5	E	79	LEU	2.0
1	A	135	ALA	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	E	202	TRP	2.0
2	B	88	SER	2.0
4	D	127	SER	2.0
5	E	159	SER	2.0
5	E	230	LYS	2.0
1	A	270	LEU	2.0
5	E	20	MET	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.