Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 10:30 am BST

PDB ID	:	1XDT
Title	:	COMPLEX OF DIPHTHERIA TOXIN AND HEPARIN-BINDING EPIDER-MAL GROWTH FACTOR
Authors	:	Louie, G.V.; Yang, W.; Bowman, M.E.; Choe, S.
Deposited on	:	1997-11-18
Resolution	:	2.65 Å(reported)

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	:	2.11
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)	:	2.11

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 2.65 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
Clashscore		10
Ramachandran outliers		1.1%
Sidechain outliers		5.1%
RSRZ outliers		3.6%

Worse　　*Better*

█ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
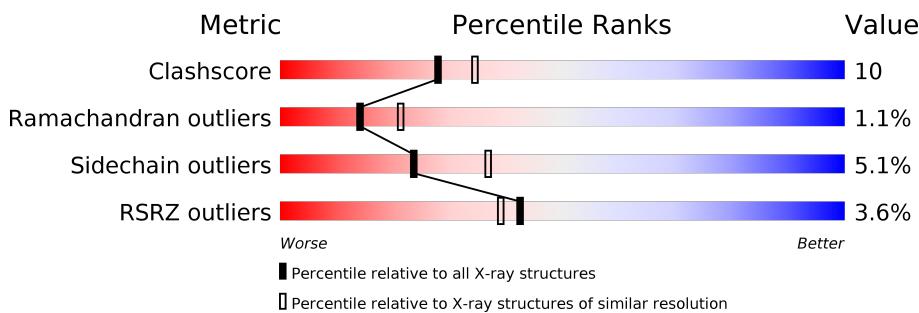
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 >=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	T	535	3% 72% 22% · ·
2	R	79	3% 43% 6% · 48%
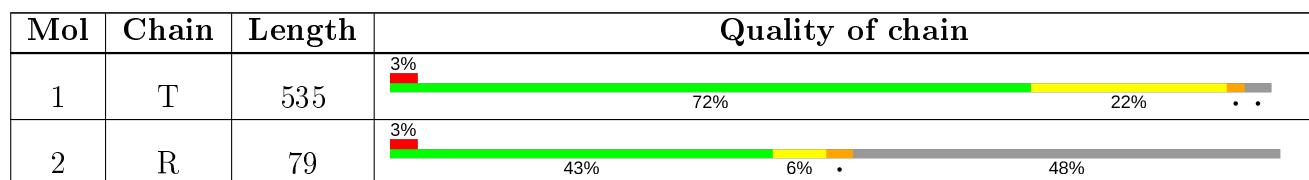
2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4354 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 DIPHTHERIA TOXIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	518	Total	C	N	O	S	1	0	0
			3983	2510	676	785	12			

- Molecule 2 is a protein called HEPARIN-BINDING EPIDERMAL GROWTH FACTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	R	41	Total	C	N	O	S	0	0	0
			323	203	59	55	6			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	T	45	Total	O	0	0
			45	45		
3	R	3	Total	O	0	0
			3	3		

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 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: DIPHTHERIA TOXIN

Chain T:
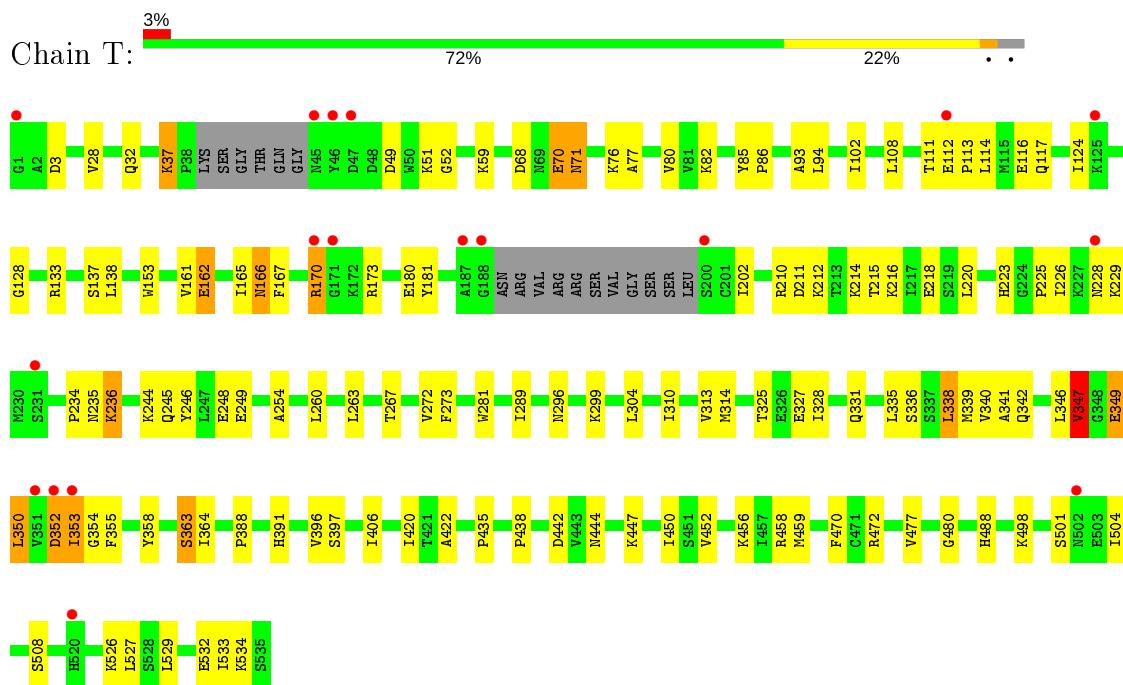
- Molecule 2: HEPARIN-BINDING EPIDERMAL GROWTH FACTOR

Chain R:
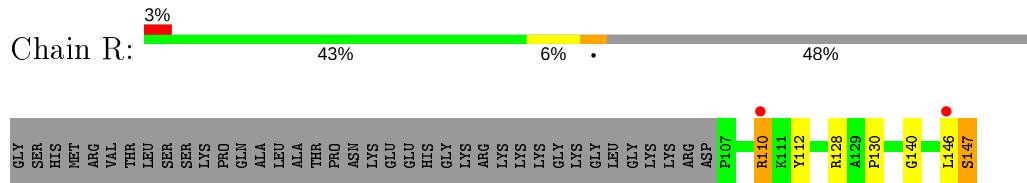
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	89.40Å 103.70Å 127.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.65 20.30 – 2.65	Depositor EDS
% Data completeness (in resolution range)	88.5 (25.00-2.65) 88.4 (20.30-2.65)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	0.62 (at 2.67Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.172 , 0.290 0.183 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	35.0	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 65.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4354	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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 (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	T	0.47	1/4063 (0.0%)	0.68	1/5503 (0.0%)
2	R	0.87	1/333 (0.3%)	0.82	1/444 (0.2%)
All	All	0.51	2/4396 (0.0%)	0.70	2/5947 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	R	128	ARG	NE-CZ	13.00	1.50	1.33
1	T	70	GLU	CB-CG	-5.03	1.42	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	R	128	ARG	CD-NE-CZ	-9.10	110.87	123.60
1	T	352	ASP	N-CA-C	5.57	126.03	111.00

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	T	3983	0	3912	80	0
2	R	323	0	298	5	0
3	R	3	0	0	0	0
3	T	45	0	0	0	0
All	All	4354	0	4210	84	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 10.

All (84) 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)
2:R:146:LEU:O	2:R:147:SER:HB2	1.88	0.72
1:T:170:ARG:HH11	1:T:170:ARG:HB2	1.55	0.70
1:T:170:ARG:NH1	1:T:170:ARG:HB2	2.06	0.69
2:R:110:ARG:HD2	2:R:110:ARG:H	1.58	0.69
1:T:304:LEU:HD12	1:T:363:SER:HB2	1.77	0.66
1:T:234:PRO:O	1:T:236:LYS:HD3	1.96	0.66
1:T:335:LEU:O	1:T:339:MET:HG3	1.97	0.64
1:T:68:ASP:HB3	1:T:70:GLU:CG	2.29	0.62
1:T:458:ARG:HH11	1:T:459:MET:H	1.49	0.60
1:T:325:THR:OG1	1:T:328:ILE:HD13	2.02	0.60
1:T:215:THR:O	1:T:218:GLU:HG2	2.01	0.60
1:T:94:LEU:HB2	1:T:138:LEU:HD21	1.83	0.60
1:T:458:ARG:NH1	1:T:459:MET:H	2.00	0.59
1:T:327:GLU:O	1:T:331:GLN:HG3	2.03	0.58
1:T:49:ASP:HB3	1:T:114:LEU:HB3	1.85	0.58
1:T:37:LYS:H	1:T:37:LYS:HD3	1.69	0.58
1:T:289:ILE:HG12	1:T:355:PHE:CZ	2.41	0.56
1:T:341:ALA:HB1	1:T:353:ILE:HG23	1.88	0.55
1:T:37:LYS:HD2	1:T:153:TRP:HB3	1.89	0.55
1:T:327:GLU:HG3	1:T:331:GLN:HE21	1.71	0.55
1:T:112:GLU:HB2	1:T:117:GLN:NE2	2.22	0.54
1:T:94:LEU:HB2	1:T:138:LEU:CD2	2.38	0.54
1:T:444:ASN:ND2	1:T:447:LYS:HG3	2.22	0.54
1:T:70:GLU:CG	1:T:71:ASN:N	2.71	0.54
1:T:532:GLU:OE2	1:T:534:LYS:HE2	2.08	0.53
1:T:388:PRO:HA	1:T:397:SER:HB3	1.90	0.53
1:T:226:ILE:HD11	1:T:254:ALA:HA	1.89	0.53
1:T:406:ILE:HB	1:T:533:ILE:HD13	1.91	0.53
1:T:338:LEU:O	1:T:342:GLN:HG2	2.08	0.52
1:T:396:VAL:HG12	1:T:420:ILE:HG23	1.91	0.52
1:T:113:PRO:HG2	1:T:116:GLU:HB2	1.92	0.52
2:R:110:ARG:N	2:R:110:ARG:HD2	2.25	0.52
1:T:391:HIS:HB3	1:T:527:LEU:HD23	1.90	0.51
1:T:296:ASN:HB3	1:T:299:LYS:HB2	1.91	0.51
1:T:260:LEU:HA	1:T:331:GLN:OE1	2.09	0.51
1:T:406:ILE:O	1:T:533:ILE:HA	2.10	0.51
1:T:245:GLN:O	1:T:249:GLU:HG3	2.10	0.51

<div align="center">Continued on next page...</div>

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:T:113:PRO:O	1:T:117:GLN:HG3	2.11	0.51
1:T:349:GLU:H	1:T:350:LEU:HD23	1.76	0.50
1:T:124:ILE:O	1:T:128:GLY:HA3	2.12	0.50
1:T:310:ILE:HD13	1:T:314:MET:HE3	1.94	0.49
1:T:28:VAL:HG11	1:T:166:ASN:OD1	2.12	0.49
1:T:173:ARG:HD2	1:T:210:ARG:HD3	1.93	0.49
1:T:272:VAL:HG12	1:T:273:PHE:CD1	2.48	0.48
1:T:391:HIS:HB3	1:T:527:LEU:CD2	2.44	0.48
1:T:229:LYS:HE3	1:T:246:TYR:CE1	2.49	0.48
1:T:70:GLU:CG	1:T:71:ASN:H	2.27	0.48
1:T:52:GLY:HA2	1:T:153:TRP:CD1	2.49	0.47
1:T:470:PHE:CE2	1:T:472:ARG:HB2	2.48	0.47
1:T:526:LYS:HE2	2:R:140:GLY:O	2.14	0.47
1:T:216:LYS:HE3	1:T:331:GLN:HG2	1.97	0.47
1:T:435:PRO:HG2	1:T:508:SER:O	2.15	0.47
1:T:202:ILE:N	1:T:202:ILE:HD12	2.31	0.46
1:T:452:VAL:HG21	1:T:477:VAL:HG13	1.97	0.46
1:T:529:LEU:N	1:T:529:LEU:HD22	2.30	0.46
1:T:223:HIS:CD2	1:T:225:PRO:HD2	2.51	0.46
1:T:225:PRO:HA	1:T:228:ASN:HB3	1.98	0.46
1:T:346:LEU:O	1:T:347:VAL:HG13	2.16	0.45
1:T:32:GLN:HA	1:T:161:VAL:CG1	2.47	0.45
1:T:281:TRP:CZ2	1:T:364:ILE:HD11	2.51	0.45
1:T:244:LYS:O	1:T:248:GLU:HG3	2.16	0.45
1:T:336:SER:O	1:T:340:VAL:HG23	2.16	0.45
1:T:49:ASP:OD1	1:T:113:PRO:HA	2.17	0.45
1:T:94:LEU:HD13	1:T:102:ILE:HD13	1.99	0.45
1:T:112:GLU:HB2	1:T:117:GLN:HE21	1.82	0.45
1:T:212:LYS:O	1:T:216:LYS:HG3	2.17	0.44
1:T:214:LYS:HE2	1:T:214:LYS:HB3	1.87	0.44
1:T:202:ILE:H	1:T:202:ILE:HD12	1.82	0.44
1:T:59:LYS:HE2	1:T:181:TYR:OH	2.18	0.44
1:T:77:ALA:HB1	1:T:167:PHE:CD2	2.53	0.44
1:T:263:LEU:O	1:T:267:THR:HB	2.17	0.44
1:T:438:PRO:HA	1:T:442:ASP:OD1	2.18	0.44
1:T:260:LEU:HD22	1:T:331:GLN:HB3	2.00	0.43
1:T:93:ALA:HA	1:T:137:SER:O	2.19	0.43
1:T:447:LYS:O	1:T:488:HIS:HD2	2.01	0.43
1:T:216:LYS:O	1:T:220:LEU:HG	2.19	0.42
1:T:82:LYS:O	1:T:162:GLU:HG2	2.18	0.42
1:T:80:VAL:HB	1:T:165:ILE:HG22	2.00	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:T:422:ALA:HB3	1:T:480:GLY:HA2	2.01	0.42
1:T:281:TRP:HZ2	1:T:364:ILE:HD11	1.85	0.41
2:R:112:TYR:CD2	2:R:130:PRO:HB3	2.55	0.41
1:T:85:TYR:HA	1:T:86:PRO:HD3	1.93	0.41
1:T:501:SER:HA	1:T:504:ILE:HD12	2.01	0.41
1:T:108:LEU:HB3	1:T:117:GLN:OE1	2.22	0.40

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	T	512/535 (96%)	476 (93%)	31 (6%)	5 (1%)	15	23
2	R	39/79 (49%)	34 (87%)	4 (10%)	1 (3%)	5	7
All	All	551/614 (90%)	510 (93%)	35 (6%)	6 (1%)	14	21

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	T	352	ASP
1	T	235	ASN
1	T	347	VAL
1	T	349	GLU
2	R	110	ARG
1	T	354	GLY

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	T	437/452 (97%)	414 (95%)	23 (5%)	22	35
2	R	35/67 (52%)	34 (97%)	1 (3%)	42	60
All	All	472/519 (91%)	448 (95%)	24 (5%)	24	37

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	T	3	ASP
1	T	37	LYS
1	T	51	LYS
1	T	71	ASN
1	T	76	LYS
1	T	111	THR
1	T	133	ARG
1	T	162	GLU
1	T	166	ASN
1	T	170	ARG
1	T	180	GLU
1	T	211	ASP
1	T	236	LYS
1	T	313	VAL
1	T	338	LEU
1	T	347	VAL
1	T	350	LEU
1	T	353	ILE
1	T	358	TYR
1	T	363	SER
1	T	450	ILE
1	T	456	LYS
1	T	498	LYS
2	R	147	SER

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

Mol	Chain	Res	Type
1	T	45	ASN
1	T	71	ASN
1	T	166	ASN
1	T	387	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 carbohydrates 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 (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, 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	T	518/535 (96%)	-0.25	18 (3%)	44	40	7, 34, 75, 100	1 (0%)
2	R	41/79 (51%)	-0.08	2 (4%)	29	26	17, 35, 72, 86	0
All	All	559/614 (91%)	-0.24	20 (3%)	42	39	7, 34, 75, 100	1 (0%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	T	45	ASN	7.9
1	T	353	ILE	6.3
1	T	46	TYR	5.1
1	T	47	ASP	4.4
1	T	352	ASP	4.4
1	T	351	VAL	4.3
1	T	188	GLY	3.8
1	T	187	ALA	3.7
1	T	171	GLY	3.1
2	R	110	ARG	2.8
1	T	112	GLU	2.5
1	T	231	SER	2.3
1	T	520	HIS	2.3
2	R	146	LEU	2.3
1	T	200	SER	2.2
1	T	502	ASN	2.2
1	T	125	LYS	2.1
1	T	228	ASN	2.1
1	T	170	ARG	2.1
1	T	1	GLY	2.1

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.