WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 05:23 AM EDT

PDB ID	:	6BPC
Title	:	Plasmodium vivax reticulocyte binding protein 2b (PvRBP2b) bound to monoclonal antibody 4F7
Authors	:	Gruszczyk, J.; Chan, L.J.; Tham, W.H.
Deposited on	:	2017-11-22
Resolution	:	2.66 Å(reported)

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

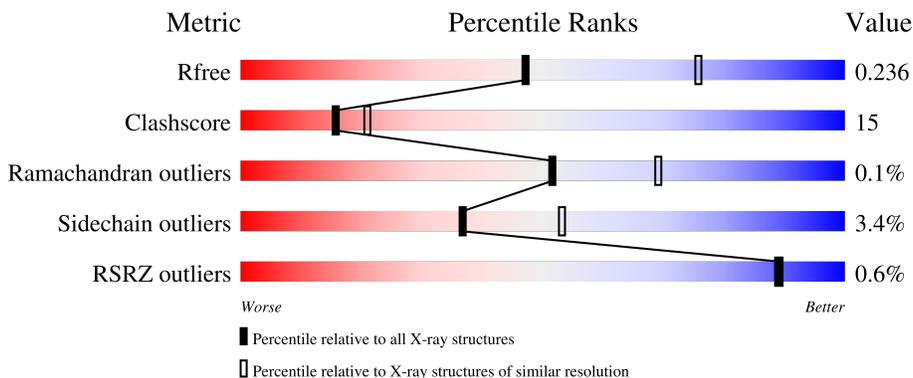|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.66 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		0.236
Clashscore		15
Ramachandran outliers		0.1%
Sidechain outliers		3.4%
RSRZ outliers		0.6%

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))
R$_{free}$	130704	1332 (2.68-2.64)
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 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	307	2% 70% 24% · ·
1	D	307	.% 53% 38% · 7%
2	B	254	53% 30% · 16%
2	E	254	61% 21% · 17%
3	C	237	67% 23% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	237	64% 25% 11%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 12045 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 Reticulocyte binding protein 2, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	0	0
			2494	1599	426	460	9			
1	D	286	Total	C	N	O	S	0	0	0
			2410	1546	411	444	9			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	164	GLY	-	expression tag	UNP A5K736
A	165	ALA	-	expression tag	UNP A5K736
A	166	MET	-	expression tag	UNP A5K736
A	167	GLY	-	expression tag	UNP A5K736
A	168	SER	-	expression tag	UNP A5K736
D	164	GLY	-	expression tag	UNP A5K736
D	165	ALA	-	expression tag	UNP A5K736
D	166	MET	-	expression tag	UNP A5K736
D	167	GLY	-	expression tag	UNP A5K736
D	168	SER	-	expression tag	UNP A5K736

- Molecule 2 is a protein called Monoclonal antibody 4F7 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	213	Total	C	N	O	S	0	0	0
			1620	1029	260	324	7			
2	E	212	Total	C	N	O	S	0	0	0
			1613	1024	259	323	7			

- Molecule 3 is a protein called Monoclonal antibody 4F7 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	212	Total	C	N	O	S	0	0	0
			1646	1024	278	337	7			

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	F	211	1642	1022	277	336	7	0	0	0

- Molecule 4 is water.

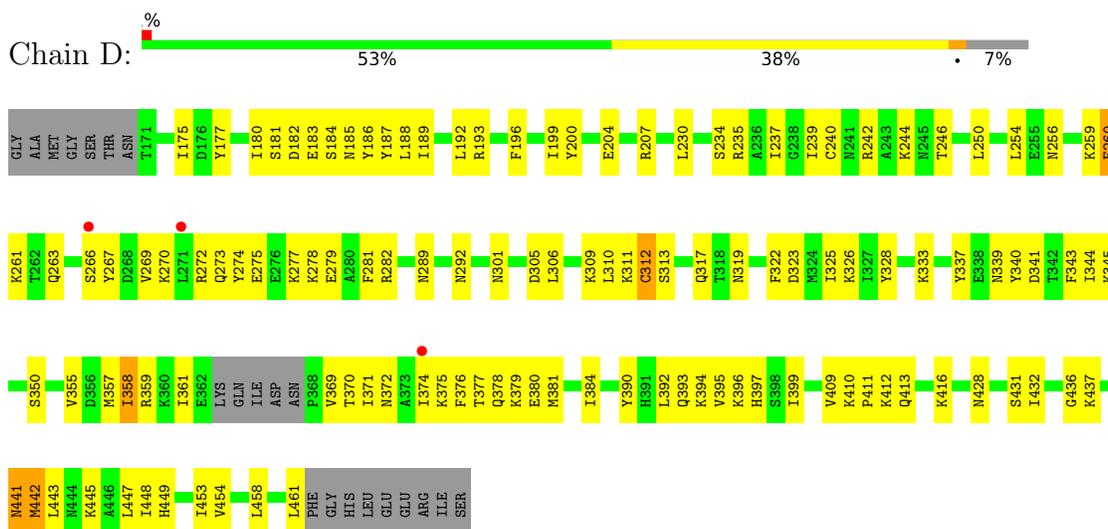| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	90	Total	O	0	0
			90	90		
4	B	98	Total	O	0	0
			98	98		
4	C	130	Total	O	0	0
			130	130		
4	D	62	Total	O	0	0
			62	62		
4	E	109	Total	O	0	0
			109	109		
4	F	131	Total	O	0	0
			131	131		

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.
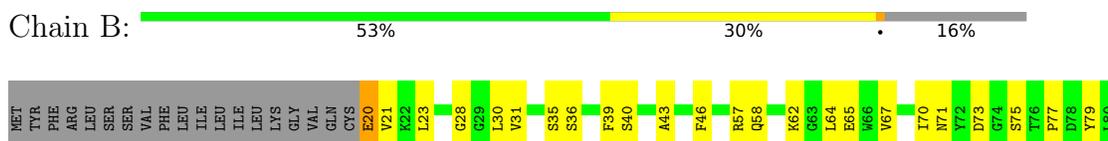
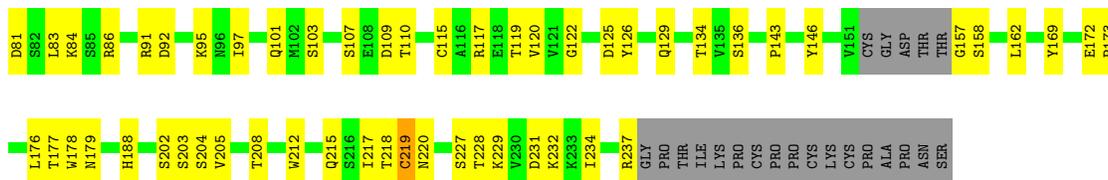• Molecule 1: Reticulocyte binding protein 2, putative
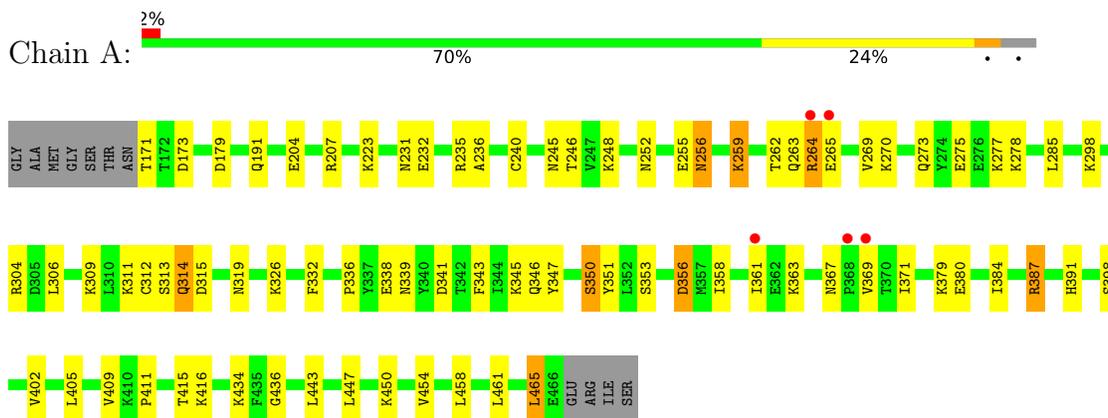
Chain A:

• Molecule 1: Reticulocyte binding protein 2, putative

Chain D:

• Molecule 2: Monoclonal antibody 4F7 Fab heavy chain

Chain B:

• Molecule 2: Monoclonal antibody 4F7 Fab heavy chain

Chain E: 61% 21% · 17%

• Molecule 3: Monoclonal antibody 4F7 Fab light chain

Chain C: 67% 23% 11%

• Molecule 3: Monoclonal antibody 4F7 Fab light chain

Chain F: 64% 25% 11%
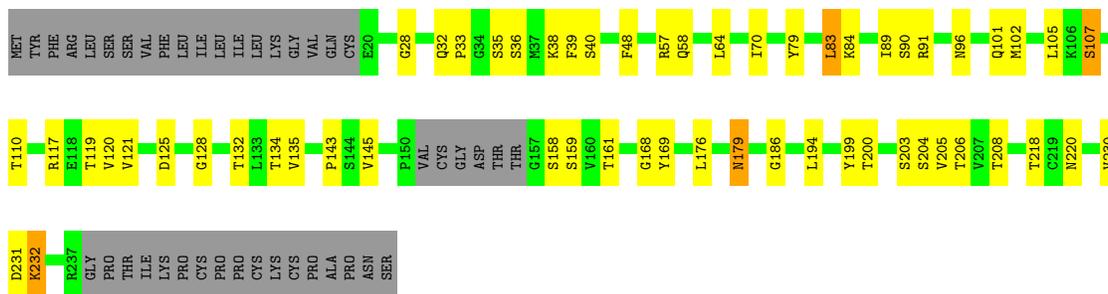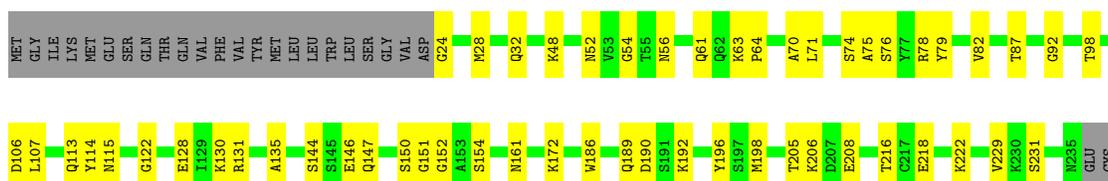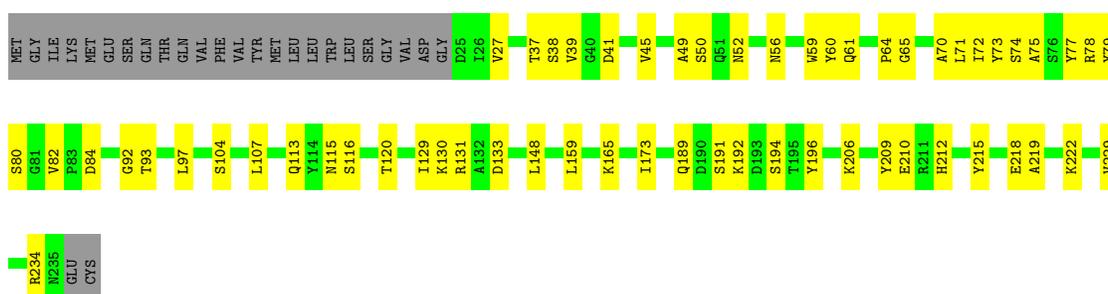
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	70.09Å 70.31Å 140.36Å 104.18° 100.34° 98.41°	Depositor
Resolution (Å)	66.31 – 2.66 66.31 – 2.66	Depositor EDS
% Data completeness (in resolution range)	94.3 (66.31-2.66) 94.3 (66.31-2.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 2.65Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.197 , 0.215 0.212 , 0.236	Depositor DCC
R_{free} test set	934 reflections (1.38%)	wwPDB-VP
Wilson B-factor (Å^2)	47.3	Xtriage
Anisotropy	0.202	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 42.4	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.053 for k,h,-h-k-l 0.107 for -k,-h,-l 0.046 for -h,-k,h+k+l	Xtriage
Reported twinning fraction	0.120 for -k,-h,-l	Depositor
Outliers	0 of 67840 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12045	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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	0.26	0/2543	0.40	0/3411
1	D	0.28	0/2456	0.44	0/3292
2	B	0.28	0/1660	0.50	0/2265
2	E	0.28	0/1653	0.49	0/2255
3	C	0.27	0/1683	0.48	0/2284
3	F	0.27	0/1679	0.47	0/2279
All	All	0.27	0/11674	0.46	0/15786

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	2494	0	2504	69	0
1	D	2410	0	2426	88	0
2	B	1620	0	1577	60	0
2	E	1613	0	1568	42	0
3	C	1646	0	1562	39	0
3	F	1642	0	1559	46	0
4	A	90	0	0	15	0
4	B	98	0	0	10	0
4	C	130	0	0	12	1
4	D	62	0	0	13	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	109	0	0	12	1
4	F	131	0	0	11	0
All	All	12045	0	11196	333	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 15.

The worst 5 of 333 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:265:GLU:OE2	1:A:270:LYS:CE	1.93	1.16
1:A:265:GLU:OE2	1:A:270:LYS:HE2	0.98	1.13
2:B:176:LEU:HD13	2:B:203:SER:HB2	1.42	1.00
2:B:101:GLN:OE1	4:B:301:HOH:O	1.87	0.92
1:A:264:ARG:HG2	1:A:264:ARG:HH11	1.33	0.90

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)
4:C:419:HOH:O	4:E:401:HOH:O[1_565]	2.13	0.07

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	294/307 (96%)	283 (96%)	11 (4%)	0	100	100
1	D	282/307 (92%)	269 (95%)	12 (4%)	1 (0%)	34	48
2	B	209/254 (82%)	207 (99%)	2 (1%)	0	100	100
2	E	208/254 (82%)	204 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	210/237 (89%)	202 (96%)	7 (3%)	1 (0%)	29	43
3	F	209/237 (88%)	202 (97%)	7 (3%)	0	100	100
All	All	1412/1596 (88%)	1367 (97%)	43 (3%)	2 (0%)	51	69

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	75	ALA
1	D	358	ILE

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	278/286 (97%)	267 (96%)	11 (4%)	31	47
1	D	269/286 (94%)	255 (95%)	14 (5%)	23	36
2	B	186/223 (83%)	177 (95%)	9 (5%)	25	39
2	E	185/223 (83%)	178 (96%)	7 (4%)	33	49
3	C	187/210 (89%)	184 (98%)	3 (2%)	62	78
3	F	187/210 (89%)	187 (100%)	0	100	100
All	All	1292/1438 (90%)	1248 (97%)	44 (3%)	37	53

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	263	GLN
1	D	442	MET
1	D	312	CYS
1	D	379	LYS
2	E	40	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	61	GLN
2	E	179	ASN
3	C	179	GLN
2	E	129	GLN
3	C	147	GLN

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 ⓘ

There are no ligands in this entry.

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(Å²)	Q<0.9
1	A	296/307 (96%)	0.00	5 (1%)	70	67	25, 54, 86, 109	0
1	D	286/307 (93%)	0.16	3 (1%)	82	81	33, 66, 92, 109	0
2	B	213/254 (83%)	-0.14	0	100	100	20, 38, 61, 94	0
2	E	212/254 (83%)	-0.14	0	100	100	22, 43, 68, 81	0
3	C	212/237 (89%)	-0.12	0	100	100	21, 36, 55, 81	0
3	F	211/237 (89%)	-0.11	0	100	100	22, 36, 57, 70	0
All	All	1430/1596 (89%)	-0.04	8 (0%)	89	89	20, 46, 82, 109	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	374	ILE	4.3
1	D	271	LEU	3.4
1	A	369	VAL	3.2
1	A	361	ILE	2.6
1	A	264	ARG	2.2

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.