Full wwPDB X-ray Structure Validation Report ⓘ

Jun 19, 2024 – 01:41 AM EDT

PDB ID	:	4BK4
Title	:	crystal structure of the human EphA4 ectodomain
Authors	:	Seiradake, E.; Schaupp, A.; del Toro Ruiz, D.; Kaufmann, R.; Mitakidis, N.; Harlos, K.; Aricescu, A.R.; Klein, R.; Jones, E.Y.
Deposited on	:	2013-04-22
Resolution	:	3.65 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:

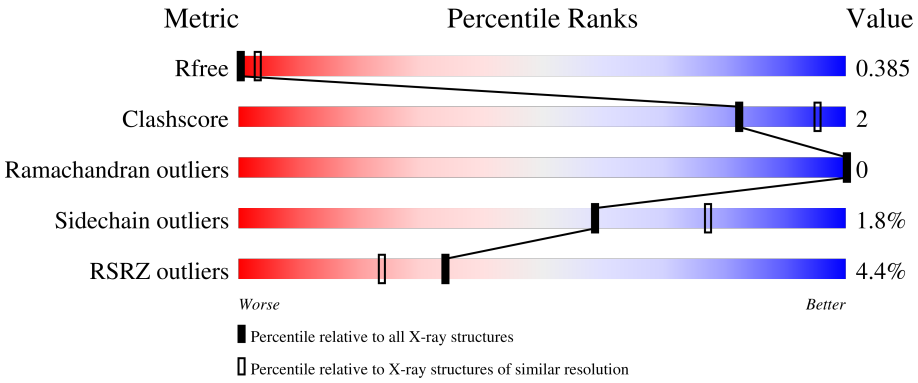
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.65 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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	568	4% 84% 7% 10%
1	B	568	2% 47% 5% 47%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6344 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 EPHRIN TYPE-A RECEPTOR 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	514	Total	C	N	O	S	0	0	0
			4002	2496	692	788	26			
1	B	300	Total	C	N	O	S	0	0	0
			2342	1460	398	462	22			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	expression tag	UNP P54764
A	-10	GLY	-	expression tag	UNP P54764
A	-9	ILE	-	expression tag	UNP P54764
A	-8	LEU	-	expression tag	UNP P54764
A	-7	PRO	-	expression tag	UNP P54764
A	-6	SER	-	expression tag	UNP P54764
A	-5	PRO	-	expression tag	UNP P54764
A	-4	GLY	-	expression tag	UNP P54764
A	-3	MET	-	expression tag	UNP P54764
A	-2	PRO	-	expression tag	UNP P54764
A	-1	ALA	-	expression tag	UNP P54764
A	0	LEU	-	expression tag	UNP P54764
A	1	LEU	-	expression tag	UNP P54764
A	2	SER	-	expression tag	UNP P54764
A	3	LEU	-	expression tag	UNP P54764
A	4	VAL	-	expression tag	UNP P54764
A	5	SER	-	expression tag	UNP P54764
A	6	LEU	-	expression tag	UNP P54764
A	7	LEU	-	expression tag	UNP P54764
A	8	SER	-	expression tag	UNP P54764
A	9	VAL	-	expression tag	UNP P54764
A	10	LEU	-	expression tag	UNP P54764
A	11	LEU	-	expression tag	UNP P54764
A	12	MET	-	expression tag	UNP P54764
A	13	GLY	-	expression tag	UNP P54764

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	CYS	-	expression tag	UNP P54764
A	15	VAL	-	expression tag	UNP P54764
A	16	ALA	-	expression tag	UNP P54764
A	17	GLU	-	expression tag	UNP P54764
A	18	THR	-	expression tag	UNP P54764
A	19	GLY	-	expression tag	UNP P54764
A	548	GLY	-	expression tag	UNP P54764
A	549	THR	-	expression tag	UNP P54764
A	550	LYS	-	expression tag	UNP P54764
A	551	HIS	-	expression tag	UNP P54764
A	552	HIS	-	expression tag	UNP P54764
A	553	HIS	-	expression tag	UNP P54764
A	554	HIS	-	expression tag	UNP P54764
A	555	HIS	-	expression tag	UNP P54764
A	556	HIS	-	expression tag	UNP P54764
B	-11	MET	-	expression tag	UNP P54764
B	-10	GLY	-	expression tag	UNP P54764
B	-9	ILE	-	expression tag	UNP P54764
B	-8	LEU	-	expression tag	UNP P54764
B	-7	PRO	-	expression tag	UNP P54764
B	-6	SER	-	expression tag	UNP P54764
B	-5	PRO	-	expression tag	UNP P54764
B	-4	GLY	-	expression tag	UNP P54764
B	-3	MET	-	expression tag	UNP P54764
B	-2	PRO	-	expression tag	UNP P54764
B	-1	ALA	-	expression tag	UNP P54764
B	0	LEU	-	expression tag	UNP P54764
B	1	LEU	-	expression tag	UNP P54764
B	2	SER	-	expression tag	UNP P54764
B	3	LEU	-	expression tag	UNP P54764
B	4	VAL	-	expression tag	UNP P54764
B	5	SER	-	expression tag	UNP P54764
B	6	LEU	-	expression tag	UNP P54764
B	7	LEU	-	expression tag	UNP P54764
B	8	SER	-	expression tag	UNP P54764
B	9	VAL	-	expression tag	UNP P54764
B	10	LEU	-	expression tag	UNP P54764
B	11	LEU	-	expression tag	UNP P54764
B	12	MET	-	expression tag	UNP P54764
B	13	GLY	-	expression tag	UNP P54764
B	14	CYS	-	expression tag	UNP P54764
B	15	VAL	-	expression tag	UNP P54764

Continued on next page...

Continued from previous page...

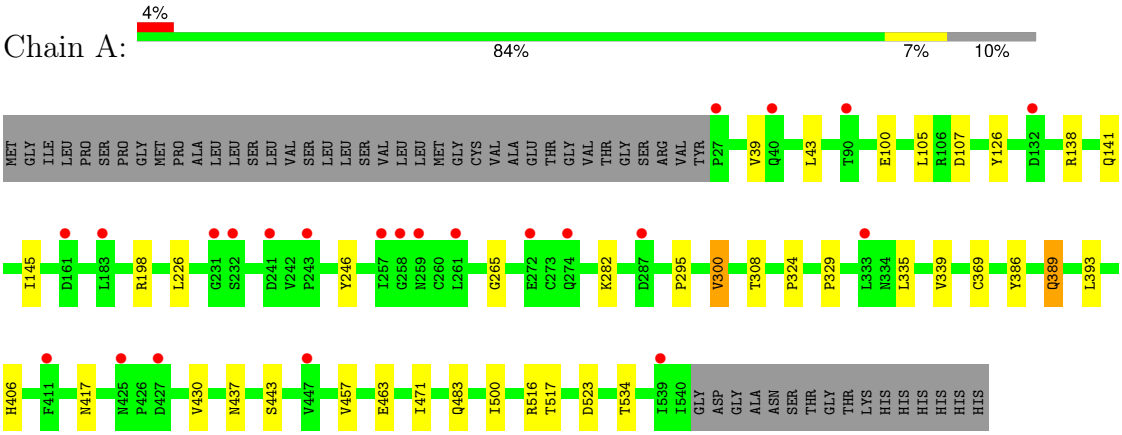
Chain	Residue	Modelled	Actual	Comment	Reference
B	16	ALA	-	expression tag	UNP P54764
B	17	GLU	-	expression tag	UNP P54764
B	18	THR	-	expression tag	UNP P54764
B	19	GLY	-	expression tag	UNP P54764
B	548	GLY	-	expression tag	UNP P54764
B	549	THR	-	expression tag	UNP P54764
B	550	LYS	-	expression tag	UNP P54764
B	551	HIS	-	expression tag	UNP P54764
B	552	HIS	-	expression tag	UNP P54764
B	553	HIS	-	expression tag	UNP P54764
B	554	HIS	-	expression tag	UNP P54764
B	555	HIS	-	expression tag	UNP P54764
B	556	HIS	-	expression tag	UNP P54764

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.

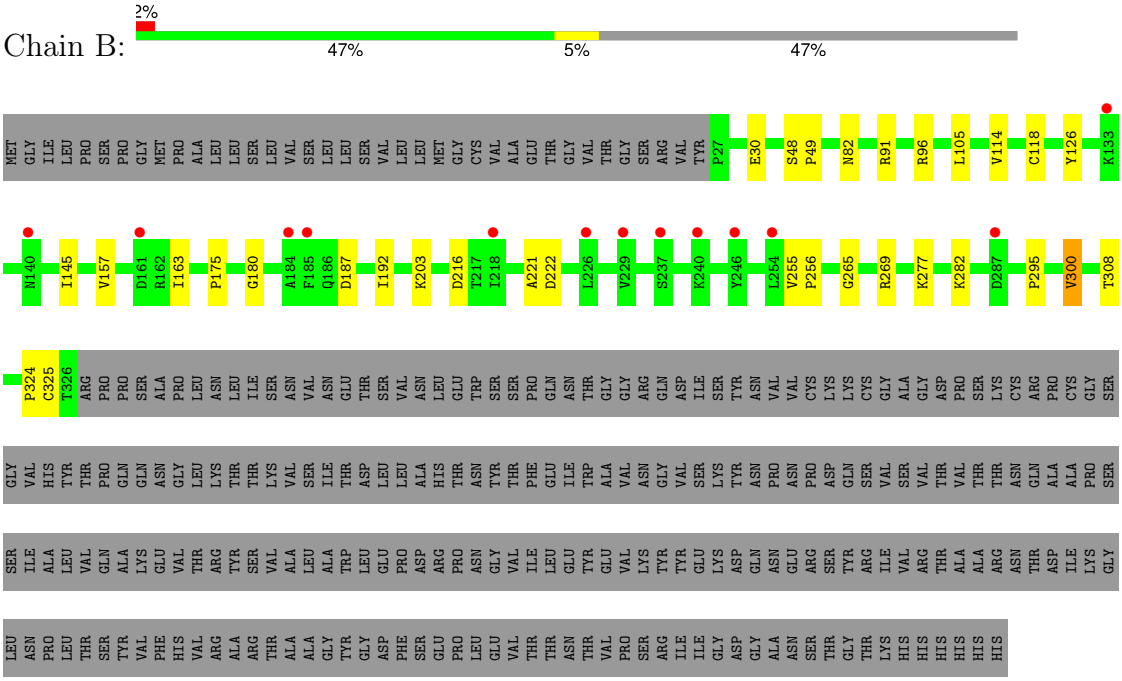
• Molecule 1: EPHRIN TYPE-A RECEPTOR 4

Chain A:



• Molecule 1: EPHRIN TYPE-A RECEPTOR 4

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	166.89Å 166.89Å 192.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.18 – 3.65 48.18 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.18-3.65) 99.9 (48.18-3.65)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.05 (at 3.67Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.349 , 0.386 0.382 , 0.385	Depositor DCC
R_{free} test set	1748 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	178.5	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 60.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.046 for -h,-k,l	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	6344	wwPDB-VP
Average B, all atoms (Å^2)	124.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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	0.35	0/4088	0.53	0/5563
1	B	0.34	0/2390	0.52	0/3240
All	All	0.35	0/6478	0.53	0/8803

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	4002	0	3865	15	0
1	B	2342	0	2236	16	0
All	All	6344	0	6101	31	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 2.

All (31) 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:126:TYR:HB3	1:A:145:ILE:HD11	1.88	0.56
1:B:126:TYR:HB3	1:B:145:ILE:HD11	1.88	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:82:ASN:HB2	1:B:187:ASP:HB3	1.89	0.54
1:A:457:VAL:H	1:A:500:ILE:HG13	1.73	0.53
1:B:91:ARG:HB2	1:B:180:GLY:H	1.74	0.52
1:A:389:GLN:HB2	1:A:393:LEU:HD13	1.92	0.52
1:A:471:ILE:HA	1:A:517:THR:HG22	1.91	0.52
1:B:105:LEU:HD12	1:B:192:ILE:HD12	1.91	0.51
1:A:443:SER:HB2	1:A:523:ASP:HB2	1.93	0.51
1:B:114:VAL:HG13	1:B:118:CYS:HB2	1.91	0.50
1:B:157:VAL:H	1:B:163:ILE:HG23	1.76	0.50
1:B:269:ARG:HH22	1:B:277:LYS:HE3	1.77	0.50
1:A:329:PRO:HG3	1:A:417:ASN:HB2	1.94	0.49
1:A:406:HIS:H	1:A:437:ASN:HB2	1.78	0.49
1:A:335:LEU:HD12	1:A:430:VAL:HG12	1.95	0.49
1:B:175:PRO:HB3	1:B:221:ALA:HB1	1.93	0.49
1:A:100:GLU:HB3	1:A:198:ARG:HB2	1.94	0.49
1:A:265:GLY:H	1:A:282:LYS:HB3	1.79	0.47
1:B:265:GLY:H	1:B:282:LYS:HB3	1.79	0.47
1:B:203:LYS:HE2	1:B:216:ASP:HB3	1.97	0.47
1:A:138:ARG:HB3	1:A:141:GLN:HB2	1.98	0.46
1:A:226:LEU:HD22	1:A:246:TYR:HB3	1.98	0.46
1:B:48:SER:HB3	1:B:49:PRO:HD3	1.98	0.44
1:A:300:VAL:HG22	1:A:308:THR:HG22	1.99	0.44
1:B:300:VAL:HG22	1:B:308:THR:HG22	1.99	0.44
1:B:30:GLU:HG3	1:B:203:LYS:H	1.83	0.44
1:B:96:ARG:HH21	1:B:222:ASP:HA	1.84	0.43
1:B:295:PRO:HD2	1:B:324:PRO:HB3	2.00	0.42
1:A:295:PRO:HD2	1:A:324:PRO:HB3	2.01	0.42
1:A:39:VAL:HG21	1:A:43:LEU:HA	2.01	0.42
1:B:255:VAL:HA	1:B:256:PRO:HD3	1.97	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	A	512/568 (90%)	483 (94%)	29 (6%)	0	100	100
1	B	298/568 (52%)	280 (94%)	18 (6%)	0	100	100
All	All	810/1136 (71%)	763 (94%)	47 (6%)	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	446/489 (91%)	435 (98%)	11 (2%)	47	69
1	B	259/489 (53%)	257 (99%)	2 (1%)	81	89
All	All	705/978 (72%)	692 (98%)	13 (2%)	59	77

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	105	LEU
1	A	107	ASP
1	A	300	VAL
1	A	339	VAL
1	A	369	CYS
1	A	386	TYR
1	A	389	GLN
1	A	463	GLU
1	A	483	GLN
1	A	516	ARG
1	A	534	THR
1	B	300	VAL
1	B	325	CYS

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

Mol	Chain	Res	Type
1	B	74	ASN
1	B	80	GLN
1	B	210	ASN
1	B	259	ASN

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)

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	514/568 (90%)	0.28	23 (4%)	33	23	77, 116, 186, 300	0
1	B	300/568 (52%)	0.31	13 (4%)	35	24	65, 115, 228, 300	0
All	All	814/1136 (71%)	0.29	36 (4%)	34	23	65, 116, 212, 300	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	237	SER	7.6
1	A	161	ASP	6.5
1	B	140	ASN	4.5
1	B	254	LEU	4.0
1	A	257	ILE	3.9
1	A	232	SER	3.8
1	A	287	ASP	3.5
1	B	240	LYS	3.0
1	A	272	GLU	3.0
1	A	258	GLY	2.9
1	A	40	GLN	2.8
1	B	246	TYR	2.8
1	B	218	ILE	2.8
1	A	132	ASP	2.8
1	A	241	ASP	2.7
1	A	427	ASP	2.7
1	A	333	LEU	2.7
1	A	411	PHE	2.7
1	B	185	PHE	2.5
1	A	231	GLY	2.5
1	A	274	GLN	2.5
1	B	184	ALA	2.4
1	A	259	ASN	2.3
1	B	226	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	229	VAL	2.3
1	A	425	ASN	2.2
1	B	161	ASP	2.2
1	A	261	LEU	2.2
1	A	243	PRO	2.2
1	A	90	THR	2.2
1	A	183	LEU	2.2
1	A	539	ILE	2.1
1	B	287	ASP	2.1
1	B	133	LYS	2.1
1	A	27	PRO	2.1
1	A	447	VAL	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.