Full wwPDB X-ray Structure Validation Report (i)

Oct 10, 2023 – 07:25 PM EDT

|		
PDB ID	:	7K2M
Title	:	Kelch domain of human KEAP1 bound to Nrf2 cyclic peptide, c[GEPETGE]
Authors	:	Muellers, S.N.; Allen, K.N.
Deposited on	:	2020-09-08
Resolution	:	2.02 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 2.02 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		0.261
Clashscore		5
Ramachandran outliers		0.3%
Sidechain outliers		1.6%
RSRZ outliers		22.9%

Worse   *Better*

∎ Percentile relative to all X-ray structures

⟦ Percentile relative to X-ray structures of similar resolution
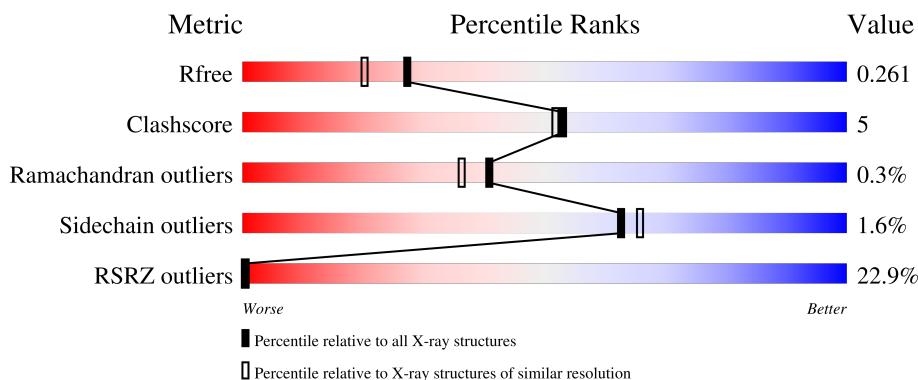
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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	289	37% — 84% / 13% ·
1	B	289	9% — 91% / 9% ·
2	P	7	86% / 14%
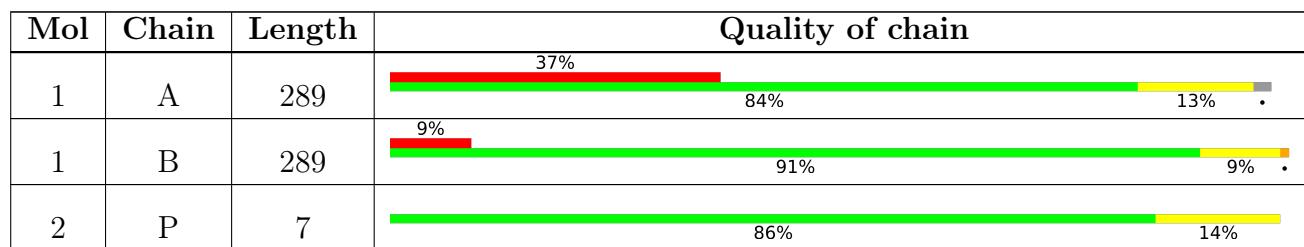
2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4511 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 Kelch-like ECH-associated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	289	Total	C	N	O	S	0	1	0
			2225	1383	404	421	17			
1	A	284	Total	C	N	O	S	0	1	0
			2101	1313	376	397	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	540	ALA	GLU	conflict	UNP Q14145
B	542	ALA	GLU	conflict	UNP Q14145
B	613	SER	CYS	conflict	UNP Q14145
A	540	ALA	GLU	conflict	UNP Q14145
A	542	ALA	GLU	conflict	UNP Q14145
A	613	SER	CYS	conflict	UNP Q14145

- Molecule 2 is a protein called Nrf2 cyclic peptide,c[GEPETGE].

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	0	0	0
			49	28	7	14			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	110	Total	O	0	0
			110	110		
3	A	25	Total	O	0	0
			25	25		
3	P	1	Total	O	0	0
			1	1		

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: Kelch-like ECH-associated protein 1
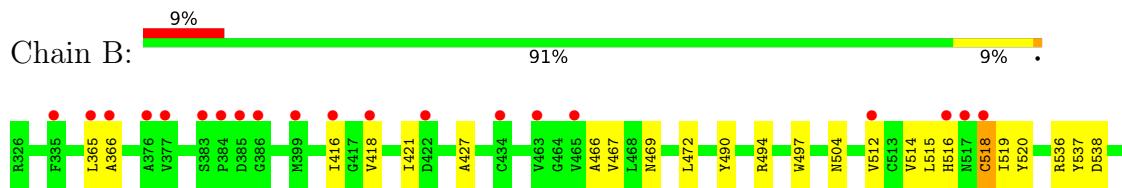
• Molecule 1: Kelch-like ECH-associated protein 1
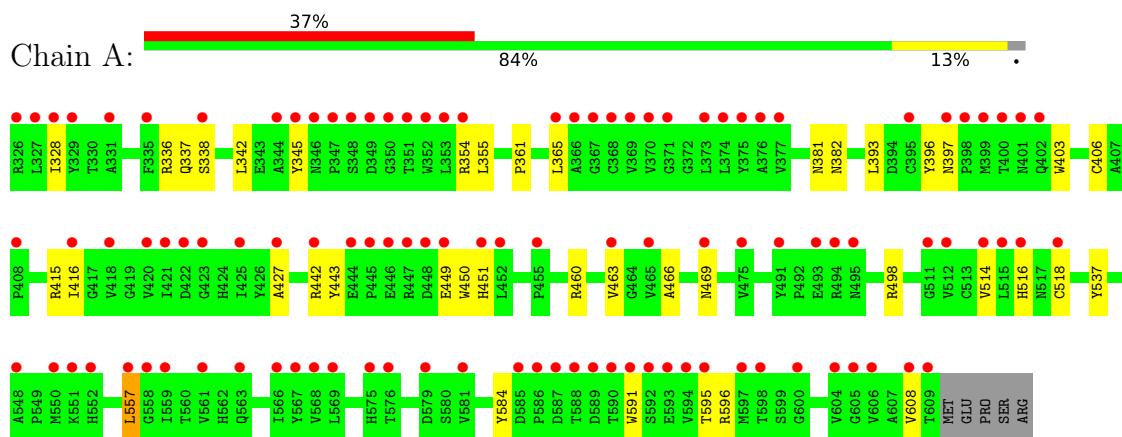
• Molecule 2: Nrf2 cyclic peptide,c[GEPETGE]
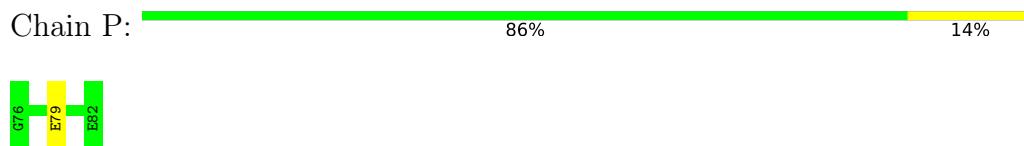
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.76Å 68.86Å 77.39Å 90.00° 117.90° 90.00°	Depositor
Resolution (Å)	29.55 – 2.02 29.55 – 2.02	Depositor EDS
% Data completeness (in resolution range)	97.9 (29.55-2.02) 97.9 (29.55-2.02)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.237 , 0.261 0.237 , 0.261	Depositor DCC
R_{free} test set	1995 reflections (4.10%)	wwPDB-VP
Wilson B-factor (Å^2)	33.9	Xtriage
Anisotropy	0.476	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 44.1	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.011 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4511	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.89% 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	A	0.27	0/2153	0.53	0/2941
1	B	0.29	0/2282	0.56	0/3107
2	P	0.26	0/49	0.52	0/65
All	All	0.28	0/4484	0.55	0/6113

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 (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	A	2101	0	1960	27	0
1	B	2225	0	2124	21	0
2	P	49	0	37	1	0
3	A	25	0	0	0	0
3	B	110	0	0	0	0
3	P	1	0	0	0	0
All	All	4511	0	4121	46	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.

All (46) 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:466:ALA:HB1	1:A:514:VAL:HG23	1.51	0.92
1:B:515:LEU:HD22	1:B:566:ILE:HG13	1.58	0.83
1:B:518:CYS:SG	1:B:537:TYR:O	2.38	0.81
1:B:466:ALA:HB1	1:B:514:VAL:HG23	1.71	0.73
1:A:415:ARG:NH2	2:P:79:GLU:OE2	2.22	0.71
1:A:514:VAL:HG12	1:A:514:VAL:O	2.00	0.61
1:A:416:ILE:HD11	1:A:427:ALA:HB1	1.82	0.60
1:B:515:LEU:HB3	1:B:520:TYR:CE1	2.37	0.59
1:B:518:CYS:HG	1:B:538:ASP:HA	1.69	0.57
1:A:393:LEU:HD13	1:A:450:TRP:HZ2	1.70	0.57
1:A:338:SER:OG	1:A:381:ASN:HA	2.05	0.56
1:A:337:GLN:HA	1:A:382:ASN:O	2.05	0.55
1:A:338:SER:HB2	1:A:381:ASN:OD1	2.07	0.55
1:B:515:LEU:HD22	1:B:566:ILE:CG1	2.33	0.54
1:A:518:CYS:HB3	1:A:537:TYR:O	2.08	0.53
1:A:355:LEU:HD23	1:A:396:TYR:OH	2.09	0.53
1:B:518:CYS:SG	1:B:538:ASP:HA	2.49	0.53
1:A:328:ILE:HG12	1:A:608:VAL:HG22	1.90	0.53
1:B:518:CYS:HG	1:B:537:TYR:C	2.11	0.52
1:A:443:TYR:HB2	1:A:450:TRP:CE2	2.46	0.51
1:A:557:LEU:HD23	1:A:557:LEU:H	1.78	0.48
1:A:595:THR:HG22	1:A:596:ARG:H	1.79	0.48
1:B:416:ILE:HD11	1:B:427:ALA:HB1	1.96	0.47
1:B:557:LEU:H	1:B:557:LEU:HD23	1.80	0.47
1:B:494:ARG:NH1	1:A:518:CYS:SG	2.88	0.46
1:A:393:LEU:HD23	1:A:406:CYS:HB2	1.98	0.46
1:B:519:ILE:O	1:B:536:ARG:HA	2.16	0.45
1:A:342:LEU:HD22	1:A:403:TRP:CZ2	2.52	0.45
1:A:460:ARG:HB3	1:A:463:VAL:HB	1.99	0.44
1:B:365:LEU:H	1:B:365:LEU:HD23	1.82	0.44
1:A:442:ARG:NH2	1:A:451:HIS:HB2	2.33	0.44
1:B:467:VAL:O	1:B:514:VAL:HG21	2.19	0.43
1:A:345:TYR:CE1	1:A:595:THR:HG21	2.53	0.43
1:A:338:SER:O	1:A:361:PRO:HB3	2.19	0.43
1:B:512:VAL:HA	1:B:520:TYR:O	2.19	0.42
1:A:443:TYR:HB2	1:A:450:TRP:CD2	2.55	0.42
1:B:504:ASN:CB	1:B:546:PHE:CE2	3.02	0.42
1:B:366:ALA:HB1	1:B:418:VAL:HG22	2.01	0.42
1:A:365:LEU:HD23	1:A:365:LEU:H	1.84	0.42
1:A:443:TYR:HA	1:A:449:GLU:O	2.19	0.42
1:A:584:TYR:HB2	1:A:591:TRP:CZ2	2.54	0.42
1:B:504:ASN:HB2	1:B:546:PHE:CZ	2.55	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:490:TYR:HB2	1:B:497:TRP:CE2	2.56	0.41
1:B:469:ASN:OD1	1:A:469:ASN:ND2	2.54	0.41
1:B:421:ILE:HD11	1:B:472:LEU:HB2	2.03	0.41
1:A:584:TYR:HB2	1:A:591:TRP:CH2	2.55	0.40

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	283/289 (98%)	271 (96%)	11 (4%)	1 (0%)	34	28
1	B	288/289 (100%)	276 (96%)	11 (4%)	1 (0%)	41	36
2	P	5/7 (71%)	5 (100%)	0	0	100	100
All	All	576/585 (98%)	552 (96%)	22 (4%)	2 (0%)	41	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	516	HIS
1	A	516	HIS

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	210/234 (90%)	205 (98%)	5 (2%)	49	49
1	B	235/234 (100%)	233 (99%)	2 (1%)	78	82
2	P	5/5 (100%)	5 (100%)	0	100	100
All	All	450/473 (95%)	443 (98%)	7 (2%)	62	66

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	518	CYS
1	B	557	LEU
1	A	336	ARG
1	A	354	ARG
1	A	397	ASN
1	A	498	ARG
1	A	557	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	397	ASN
1	A	424	HIS
1	A	451	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 ⓘ

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(Å^2)	Q<0.9
1	A	284/289 (98%)	1.70	106 (37%)	0	0	25, 57, 103, 126	0
1	B	289/289 (100%)	0.55	27 (9%)	8	8	24, 31, 51, 74	0
2	P	7/7 (100%)	0.22	0	100	100	37, 39, 40, 43	0
All	All	580/585 (99%)	1.11	133 (22%)	0	0	24, 38, 95, 126	0

All (133) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	399	MET	8.6
1	A	518	CYS	7.0
1	A	353	LEU	6.9
1	A	351	THR	6.9
1	A	515	LEU	5.7
1	A	328	ILE	5.6
1	A	349	ASP	5.4
1	A	422	ASP	5.3
1	A	423	GLY	5.3
1	A	347	PRO	5.2
1	A	346	ASN	5.0
1	A	590	THR	5.0
1	A	395	CYS	4.9
1	A	446	GLU	4.9
1	B	614	ARG	4.8
1	A	592	SER	4.8
1	A	421	ILE	4.8
1	B	366	ALA	4.6
1	A	516	HIS	4.6
1	A	591	TRP	4.6
1	A	400	THR	4.5
1	A	608	VAL	4.5
1	A	581	VAL	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	385	ASP	4.4
1	A	371	GLY	4.3
1	A	416	ILE	4.2
1	B	613	SER	4.2
1	A	594	VAL	4.2
1	A	587	ASP	4.1
1	B	384	PRO	4.1
1	A	370	VAL	4.1
1	A	368	CYS	4.1
1	A	326	ARG	4.1
1	A	604	VAL	4.0
1	A	397	ASN	4.0
1	A	491	TYR	4.0
1	A	352	TRP	3.9
1	A	348	SER	3.8
1	A	449	GLU	3.8
1	A	350	GLY	3.7
1	A	575	HIS	3.7
1	A	561	VAL	3.7
1	B	518	CYS	3.6
1	A	559	ILE	3.6
1	A	344	ALA	3.6
1	A	448	ASP	3.6
1	B	365	LEU	3.6
1	A	598	THR	3.6
1	A	588	THR	3.5
1	A	595	THR	3.5
1	A	586	PRO	3.5
1	A	568	VAL	3.5
1	A	338	SER	3.4
1	B	517	ASN	3.4
1	A	442	ARG	3.3
1	A	557	LEU	3.3
1	B	399	MET	3.3
1	A	418	VAL	3.2
1	A	444	GLU	3.2
1	A	597	MET	3.2
1	A	494	ARG	3.2
1	A	563	GLN	3.2
1	B	516	HIS	3.2
1	A	365	LEU	3.2
1	A	366	ALA	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	374	LEU	3.1
1	B	335	PHE	3.1
1	A	377	VAL	3.1
1	B	512	VAL	3.1
1	A	463	VAL	3.1
1	B	416	ILE	3.0
1	A	609	THR	3.0
1	A	585	ASP	3.0
1	A	576	THR	3.0
1	A	552	HIS	3.0
1	B	605	GLY	3.0
1	A	327	LEU	3.0
1	A	354	ARG	2.9
1	A	402	GLN	2.9
1	A	606	VAL	2.9
1	A	335	PHE	2.9
1	A	401	ASN	2.9
1	A	452	LEU	2.9
1	A	373	LEU	2.8
1	B	606	VAL	2.8
1	A	465	VAL	2.7
1	B	386	GLY	2.7
1	B	465	VAL	2.7
1	A	331	ALA	2.7
1	A	569	LEU	2.7
1	A	451	HIS	2.6
1	A	512	VAL	2.6
1	A	548	ALA	2.6
1	A	511	GLY	2.6
1	A	605	GLY	2.6
1	A	593	GLU	2.6
1	B	569	LEU	2.6
1	B	604	VAL	2.5
1	A	425	ILE	2.5
1	A	345	TYR	2.5
1	A	558	GLY	2.5
1	A	445	PRO	2.5
1	B	422	ASP	2.5
1	A	567	TYR	2.4
1	A	469	ASN	2.4
1	A	551	LYS	2.4
1	A	495	ASN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	383	SER	2.4
1	B	376	ALA	2.4
1	A	408	PRO	2.3
1	A	369	VAL	2.3
1	A	420	VAL	2.3
1	A	514	VAL	2.3
1	A	375	TYR	2.3
1	A	376	ALA	2.3
1	A	455	PRO	2.3
1	A	367	GLY	2.2
1	B	559	ILE	2.2
1	A	475	VAL	2.2
1	A	550	MET	2.2
1	A	566	ILE	2.2
1	A	447	ARG	2.2
1	B	377	VAL	2.2
1	A	589	ASP	2.2
1	A	600	GLY	2.2
1	A	427	ALA	2.1
1	A	398	PRO	2.1
1	A	329	TYR	2.1
1	A	493	GLU	2.0
1	A	579	ASP	2.0
1	B	418	VAL	2.0
1	B	463	VAL	2.0
1	B	434[A]	CYS	2.0

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.