Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2023 – 10:56 PM EDT

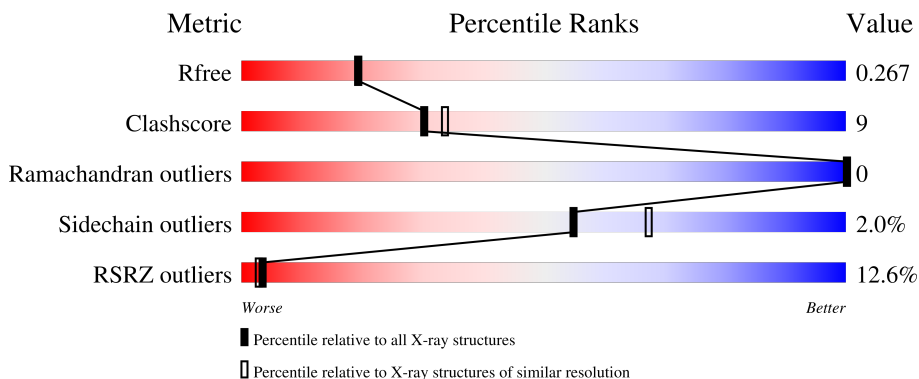|		
PDB ID	:	7K2D
Title	:	Kelch domain of human KEAP1 bound to Nrf2 linear peptide, Ac-GDEETGE-NH2
Authors	:	Muellers, S.N.; Allen, K.N.
Deposited on	:	2020-09-08
Resolution	:	2.21 Å(reported)

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 2.21 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		0.267
Clashscore		9
Ramachandran outliers		0
Sidechain outliers		2.0%
RSRZ outliers		12.6%

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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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	301	3% / 82% / 11% / 6%
1	B	301	21% / 71% / 22% / 7%
2	P	9	78% / 22%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4521 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	A	283	Total	C	N	O	S	0	0	0
			2170	1351	392	412	15			
1	B	280	Total	C	N	O	S	0	0	0
			2142	1330	389	408	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	613	SER	CYS	conflict	UNP Q14145
A	622	SER	CYS	conflict	UNP Q14145
A	624	SER	-	expression tag	UNP Q14145
B	613	SER	CYS	conflict	UNP Q14145
B	622	SER	CYS	conflict	UNP Q14145
B	624	SER	-	expression tag	UNP Q14145

- Molecule 2 is a protein called Nrf2 linear peptide, Ace-GDEETGE-NH2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	9	Total	C	N	O	0	0	1
			54	29	8	17			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	118	Total	O	0	0
			118	118		
3	B	35	Total	O	0	0
			35	35		
3	P	2	Total	O	0	0
			2	2		

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.
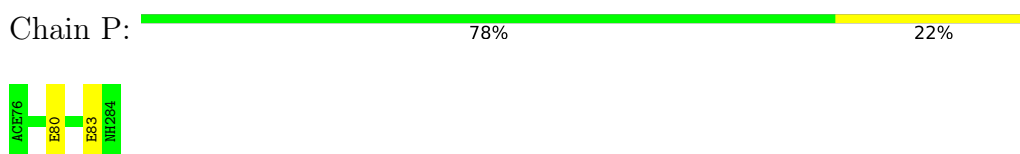
- Molecule 1: Kelch-like ECH-associated protein 1
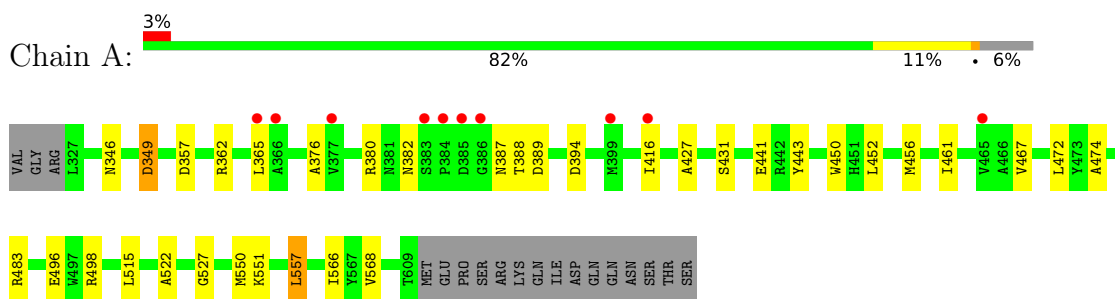
- Molecule 1: Kelch-like ECH-associated protein 1
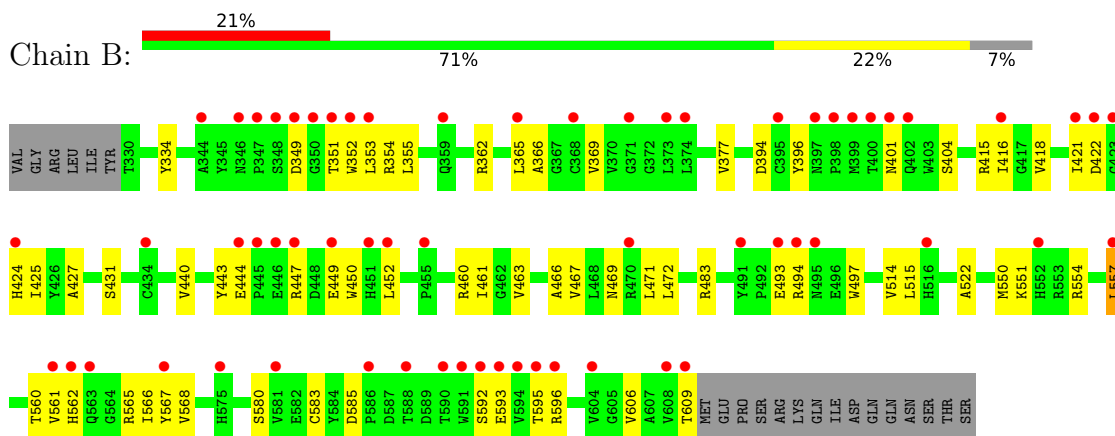
- Molecule 2: Nrf2 linear peptide, Ace-GDEETGE-NH2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	77.94Å 68.85Å 143.91Å 90.00° 91.12° 90.00°	Depositor
Resolution (Å)	27.47 – 2.21 27.47 – 2.21	Depositor EDS
% Data completeness (in resolution range)	98.9 (27.47-2.21) 98.9 (27.47-2.21)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 2.22Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.223 , 0.267 0.223 , 0.267	Depositor DCC
R_{free} test set	1261 reflections (3.32%)	wwPDB-VP
Wilson B-factor (Å^2)	32.0	Xtriage
Anisotropy	0.285	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.018 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4521	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, ACE

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	0.46	0/2223	0.67	3/3029 (0.1%)
1	B	0.37	0/2194	0.60	0/2989
2	P	0.48	0/50	0.67	0/66
All	All	0.42	0/4467	0.64	3/6084 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	498	ARG	NE-CZ-NH1	-8.25	116.18	120.30
1	A	498	ARG	NE-CZ-NH2	7.60	124.10	120.30
1	A	357	ASP	CB-CG-OD1	5.58	123.32	118.30

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	2170	0	2060	20	0
1	B	2142	0	2029	59	0
2	P	54	0	38	2	0
3	A	118	0	0	2	0
3	B	35	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	P	2	0	0	0	0
All	All	4521	0	4127	79	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 9.

All (79) 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:B:377:VAL:HG11	1:B:418:VAL:HG11	1.53	0.90
1:A:522:ALA:HB1	1:A:550:MET:HE3	1.59	0.82
1:B:415:ARG:NH2	2:P:80:GLU:OE1	2.12	0.82
1:B:444:GLU:OE1	1:B:447:ARG:HD2	1.83	0.76
1:B:561:VAL:HG12	1:B:566:ILE:HG12	1.71	0.73
1:B:425:ILE:HB	1:B:443:TYR:HB3	1.73	0.70
1:B:551:LYS:HZ2	1:B:593:GLU:H	1.44	0.66
1:A:550:MET:HE1	1:A:568:VAL:HG11	1.76	0.65
1:B:425:ILE:HG13	1:B:443:TYR:HD2	1.61	0.65
1:A:496:GLU:OE2	3:A:701:HOH:O	2.14	0.65
1:B:466:ALA:HB1	1:B:514:VAL:HG23	1.80	0.64
1:B:515:LEU:HB2	1:B:566:ILE:HD11	1.82	0.61
1:B:522:ALA:HB1	1:B:550:MET:HE3	1.81	0.61
1:A:550:MET:CE	1:A:568:VAL:HG11	2.31	0.59
1:B:444:GLU:HG3	1:B:447:ARG:HB2	1.86	0.58
1:B:424:HIS:CE1	1:B:444:GLU:HB3	2.39	0.57
1:B:562:HIS:HB3	1:B:567:TYR:CE2	2.40	0.57
1:A:515:LEU:HD22	1:A:566:ILE:HG13	1.86	0.57
1:B:349:ASP:OD1	1:B:351:THR:OG1	2.21	0.56
1:B:416:ILE:HD11	1:B:427:ALA:HB1	1.89	0.55
1:A:382:ASN:HA	1:A:387:ASN:HD22	1.72	0.55
1:B:362:ARG:HG3	1:B:365:LEU:HD13	1.89	0.55
1:B:444:GLU:OE1	1:B:447:ARG:CD	2.55	0.53
1:B:565:ARG:CG	1:B:583:CYS:SG	2.97	0.53
1:B:355:LEU:HD23	1:B:396:TYR:OH	2.08	0.53
1:B:551:LYS:HZ1	1:B:592:SER:HA	1.73	0.53
1:B:377:VAL:CG1	1:B:418:VAL:HG11	2.32	0.52
1:B:369:VAL:HG21	1:B:609:THR:HB	1.92	0.52
1:B:362:ARG:NH1	1:B:394:ASP:OD2	2.43	0.52
1:B:565:ARG:HG3	1:B:583:CYS:SG	2.50	0.52
1:B:565:ARG:NH1	1:B:585:ASP:HB2	2.25	0.51
3:A:721:HOH:O	1:B:483:ARG:HG3	2.11	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:380:ARG:NH2	1:A:387:ASN:HB3	2.26	0.51
1:B:396:TYR:HE1	1:B:401:ASN:HD22	1.59	0.51
1:A:388:THR:HG22	1:A:389:ASP:O	2.11	0.49
1:B:353:LEU:HD23	1:B:355:LEU:HD11	1.94	0.49
1:B:551:LYS:NZ	1:B:593:GLU:H	2.08	0.49
1:B:460:ARG:HB3	1:B:463:VAL:HB	1.95	0.49
1:B:365:LEU:H	1:B:365:LEU:HD23	1.78	0.48
1:B:550:MET:CE	1:B:568:VAL:HG11	2.43	0.48
1:B:551:LYS:HD2	1:B:593:GLU:HG3	1.96	0.48
1:B:443:TYR:HB2	1:B:450:TRP:CE2	2.48	0.48
1:B:469:ASN:O	1:B:471:LEU:HD22	2.13	0.48
1:B:334:TYR:CG	2:P:83:GLU:HG2	2.48	0.47
1:A:522:ALA:CB	1:A:550:MET:HE3	2.39	0.47
1:B:421:ILE:HD11	1:B:472:LEU:HB2	1.97	0.46
1:B:366:ALA:HB1	1:B:418:VAL:HG22	1.97	0.46
1:B:444:GLU:HB2	1:B:447:ARG:NH2	2.30	0.46
1:A:346:ASN:CG	1:A:349:ASP:HB2	2.36	0.46
1:B:560:THR:HB	1:B:606:VAL:HG12	1.97	0.46
1:B:467:VAL:O	1:B:514:VAL:HG21	2.16	0.45
1:A:467:VAL:HG22	1:A:472:LEU:HD12	1.99	0.45
1:A:483:ARG:HD3	1:A:527:GLY:N	2.31	0.45
1:B:595:THR:HG22	1:B:596:ARG:N	2.32	0.45
1:B:443:TYR:HB2	1:B:450:TRP:CD2	2.52	0.44
1:B:554:ARG:HG3	1:B:557:LEU:HD22	1.99	0.44
1:B:443:TYR:HA	1:B:449:GLU:O	2.17	0.44
1:A:441:GLU:HB3	1:A:452:LEU:HD23	2.00	0.44
1:B:421:ILE:HG22	1:B:422:ASP:OD2	2.17	0.44
1:B:421:ILE:O	1:B:424:HIS:HB2	2.18	0.44
1:B:425:ILE:HG13	1:B:443:TYR:CD2	2.47	0.44
1:A:365:LEU:HD12	1:A:376:ALA:HB1	2.00	0.43
1:A:416:ILE:HD11	1:A:427:ALA:HB1	2.00	0.43
1:B:493:GLU:N	1:B:493:GLU:OE2	2.52	0.43
1:B:440:VAL:HG21	1:B:497:TRP:HZ2	1.83	0.43
1:B:444:GLU:OE1	1:B:447:ARG:HB2	2.18	0.43
1:A:362:ARG:NH1	1:A:394:ASP:OD1	2.51	0.43
1:A:557:LEU:HD23	1:A:557:LEU:H	1.84	0.43
1:B:352:TRP:HE1	1:B:595:THR:CG2	2.31	0.42
1:B:352:TRP:HE1	1:B:595:THR:HG21	1.84	0.42
1:B:551:LYS:HZ2	1:B:593:GLU:N	2.13	0.42
1:A:431:SER:HB3	1:A:461:ILE:HG21	2.02	0.42
1:B:557:LEU:H	1:B:557:LEU:HD23	1.85	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:424:HIS:HE1	1:B:444:GLU:HB3	1.83	0.42
1:A:443:TYR:HB2	1:A:450:TRP:CD2	2.56	0.41
1:B:431:SER:HB3	1:B:461:ILE:HG21	2.03	0.41
1:A:456:MET:HE1	1:A:474:ALA:CB	2.51	0.40
1:B:421:ILE:HD13	1:B:421:ILE:HG21	1.89	0.40
1:B:550:MET:HE1	1:B:568:VAL:HG11	2.03	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	281/301 (93%)	271 (96%)	10 (4%)	0	100	100
1	B	278/301 (92%)	265 (95%)	13 (5%)	0	100	100
2	P	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	566/611 (93%)	542 (96%)	24 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	228/247 (92%)	225 (99%)	3 (1%)	69	80

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	225/247 (91%)	219 (97%)	6 (3%)	44	55
2	P	5/5 (100%)	5 (100%)	0	100	100
All	All	458/499 (92%)	449 (98%)	9 (2%)	55	67

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

Mol	Chain	Res	Type
1	A	349	ASP
1	A	551	LYS
1	A	557	LEU
1	B	354	ARG
1	B	404	SER
1	B	452	LEU
1	B	494	ARG
1	B	557	LEU
1	B	580	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	382	ASN
1	A	387	ASN
1	A	469	ASN
1	B	401	ASN
1	B	424	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	283/301 (94%)	0.02	10 (3%)	44	41	19, 26, 47, 76	0
1	B	280/301 (93%)	1.01	62 (22%)	0	0	23, 49, 92, 116	0
2	P	7/9 (77%)	-0.06	0	100	100	33, 34, 47, 50	0
All	All	570/611 (93%)	0.51	72 (12%)	3	3	19, 34, 83, 116	0

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	399	MET	7.0
1	B	562	HIS	6.0
1	B	351	THR	5.9
1	B	447	ARG	5.3
1	B	346	ASN	4.8
1	B	423	GLY	4.6
1	B	402	GLN	4.6
1	B	371	GLY	4.5
1	B	421	ILE	4.5
1	B	400	THR	4.4
1	B	575	HIS	4.2
1	A	384	PRO	4.2
1	B	350	GLY	4.0
1	B	398	PRO	3.9
1	B	347	PRO	3.8
1	B	444	GLU	3.7
1	B	446	GLU	3.7
1	B	373	LEU	3.7
1	B	353	LEU	3.6
1	B	422	ASP	3.6
1	B	455	PRO	3.6
1	B	561	VAL	3.5
1	B	595	THR	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	491	TYR	3.4
1	B	588	THR	3.4
1	B	594	VAL	3.4
1	A	385	ASP	3.3
1	B	563	GLN	3.3
1	B	608	VAL	3.3
1	A	386	GLY	3.2
1	A	366	ALA	3.2
1	B	452	LEU	3.2
1	B	352	TRP	3.1
1	B	395	CYS	3.0
1	B	365	LEU	3.0
1	A	399	MET	2.9
1	B	449	GLU	2.8
1	B	494	ARG	2.8
1	B	416	ILE	2.8
1	B	349	ASP	2.7
1	B	604	VAL	2.7
1	B	567	TYR	2.7
1	B	493	GLU	2.7
1	B	590	THR	2.7
1	B	348	SER	2.6
1	B	557	LEU	2.5
1	B	451	HIS	2.5
1	B	424	HIS	2.4
1	B	609	THR	2.4
1	B	359	GLN	2.4
1	B	470	ARG	2.4
1	B	596	ARG	2.4
1	B	434	CYS	2.4
1	B	397	ASN	2.3
1	A	365	LEU	2.3
1	B	344	ALA	2.3
1	B	581	VAL	2.3
1	B	516	HIS	2.3
1	B	591	TRP	2.3
1	B	586	PRO	2.2
1	B	495	ASN	2.2
1	B	552	HIS	2.2
1	B	401	ASN	2.1
1	A	383	SER	2.1
1	B	374	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	368	CYS	2.1
1	A	377	VAL	2.1
1	A	465	VAL	2.1
1	B	445	PRO	2.1
1	A	416	ILE	2.0
1	B	593	GLU	2.0
1	B	592	SER	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.