Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 01:48 PM GMT

PDB ID	:	3UVU
Title	:	Structural basis of nuclear import of Flap endonuclease 1 (FEN1)
Authors	:	Barros, A.C.; Takeda, A.A.S.; Fontes, M.R.M.
Deposited on	:	2011-11-30
Resolution	:	2.38 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

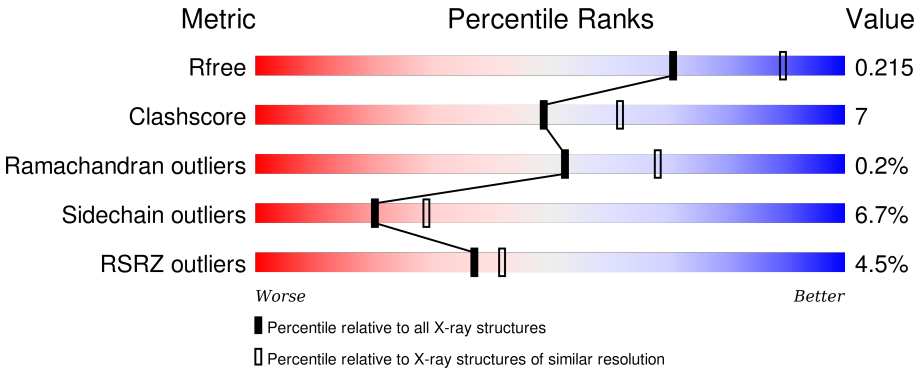
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.38 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}	91344	4019 (2.40-2.36)
Clashscore	102246	4595 (2.40-2.36)
Ramachandran outliers	100387	4520 (2.40-2.36)
Sidechain outliers	100360	4522 (2.40-2.36)
RSRZ outliers	91569	4034 (2.40-2.36)

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. 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	B	19	63% 26% 5% 5%
2	A	510	4% 68% 13% • 17%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3477 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 Flap endonuclease 1 (Fen1) peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	B	19	146	90	29	27	0	0	0

- Molecule 2 is a protein called Importin subunit alpha-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	425	3174	2023	541	600	10	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	EXPRESSION TAG	UNP P52293
A	21	HIS	-	EXPRESSION TAG	UNP P52293
A	22	HIS	-	EXPRESSION TAG	UNP P52293
A	23	HIS	-	EXPRESSION TAG	UNP P52293
A	24	HIS	-	EXPRESSION TAG	UNP P52293
A	25	HIS	-	EXPRESSION TAG	UNP P52293
A	26	HIS	-	EXPRESSION TAG	UNP P52293
A	27	SER	-	EXPRESSION TAG	UNP P52293
A	28	SER	-	EXPRESSION TAG	UNP P52293
A	29	GLY	-	EXPRESSION TAG	UNP P52293
A	30	LEU	-	EXPRESSION TAG	UNP P52293
A	31	VAL	-	EXPRESSION TAG	UNP P52293
A	32	PRO	-	EXPRESSION TAG	UNP P52293
A	33	ARG	-	EXPRESSION TAG	UNP P52293
A	34	GLY	-	EXPRESSION TAG	UNP P52293
A	35	SER	-	EXPRESSION TAG	UNP P52293
A	36	GLY	-	EXPRESSION TAG	UNP P52293
A	37	MET	-	EXPRESSION TAG	UNP P52293
A	38	LYS	-	EXPRESSION TAG	UNP P52293
A	39	GLU	-	EXPRESSION TAG	UNP P52293

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	THR	-	EXPRESSION TAG	UNP P52293
A	41	ALA	-	EXPRESSION TAG	UNP P52293
A	42	ALA	-	EXPRESSION TAG	UNP P52293
A	43	ALA	-	EXPRESSION TAG	UNP P52293
A	44	LYS	-	EXPRESSION TAG	UNP P52293
A	45	PHE	-	EXPRESSION TAG	UNP P52293
A	46	GLU	-	EXPRESSION TAG	UNP P52293
A	47	ARG	-	EXPRESSION TAG	UNP P52293
A	48	GLN	-	EXPRESSION TAG	UNP P52293
A	49	HIS	-	EXPRESSION TAG	UNP P52293
A	50	MET	-	EXPRESSION TAG	UNP P52293
A	51	ASP	-	EXPRESSION TAG	UNP P52293
A	52	SER	-	EXPRESSION TAG	UNP P52293
A	53	PRO	-	EXPRESSION TAG	UNP P52293
A	54	ASP	-	EXPRESSION TAG	UNP P52293
A	55	LEU	-	EXPRESSION TAG	UNP P52293
A	56	GLY	-	EXPRESSION TAG	UNP P52293
A	57	THR	-	EXPRESSION TAG	UNP P52293
A	58	ASP	-	EXPRESSION TAG	UNP P52293
A	59	ASP	-	EXPRESSION TAG	UNP P52293
A	60	ASP	-	EXPRESSION TAG	UNP P52293
A	61	ASP	-	EXPRESSION TAG	UNP P52293
A	62	LYS	-	EXPRESSION TAG	UNP P52293
A	63	ALA	-	EXPRESSION TAG	UNP P52293
A	64	MET	-	EXPRESSION TAG	UNP P52293
A	65	ALA	-	EXPRESSION TAG	UNP P52293
A	66	ASP	-	EXPRESSION TAG	UNP P52293
A	67	ILE	-	EXPRESSION TAG	UNP P52293
A	68	GLY	-	EXPRESSION TAG	UNP P52293
A	69	SER	-	EXPRESSION TAG	UNP P52293

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	11	Total O 11 11	0	0
3	A	146	Total O 146 146	0	0

3 Residue-property plots (i)

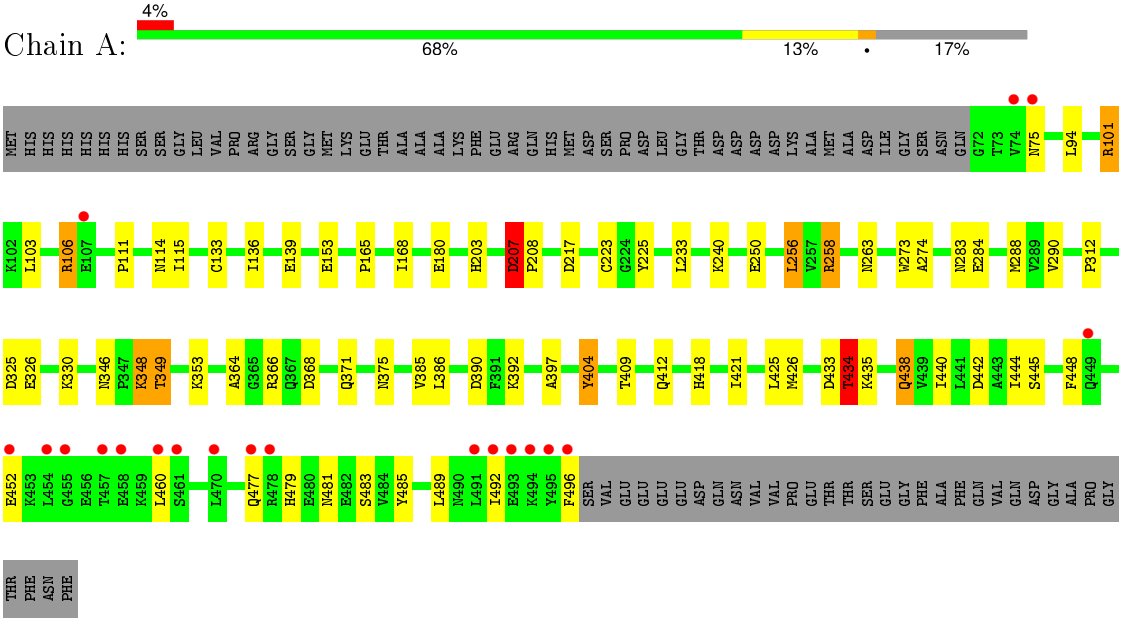
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Flap endonuclease 1 (Fen1) peptide

Chain B:

- Molecule 2: Importin subunit alpha-2

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.87Å 89.56Å 100.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.71 – 2.38 36.70 – 2.38	Depositor EDS
% Data completeness (in resolution range)	97.0 (36.71-2.38) 97.1 (36.70-2.38)	Depositor EDS
R_{merge}	10.50	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.19 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.168 , 0.215 0.169 , 0.215	Depositor DCC
R_{free} test set	1443 reflections (5.39%)	DCC
Wilson B-factor (Å^2)	45.5	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 40.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	1 of 28235 reflections (0.004%)	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3477	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	B	1.31	1/147 (0.7%)	1.39	3/191 (1.6%)
2	A	1.06	5/3231 (0.2%)	0.91	8/4414 (0.2%)
All	All	1.07	6/3378 (0.2%)	0.94	11/4605 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	326	GLU	CG-CD	6.00	1.60	1.51
2	A	153	GLU	CG-CD	5.93	1.60	1.51
2	A	250	GLU	CG-CD	5.58	1.60	1.51
2	A	284	GLU	CG-CD	5.51	1.60	1.51
1	B	359	GLU	CG-CD	5.42	1.60	1.51
2	A	404	TYR	CE2-CZ	5.35	1.45	1.38

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	258	ARG	NE-CZ-NH2	-8.87	115.86	120.30
2	A	258	ARG	NE-CZ-NH1	8.66	124.63	120.30
2	A	217	ASP	CB-CG-OD1	7.97	125.47	118.30
1	B	356	LYS	C-N-CA	-7.27	103.53	121.70
2	A	233	LEU	CB-CG-CD1	-6.61	99.76	111.00
2	A	207	ASP	CB-CG-OD1	-6.32	112.61	118.30
2	A	256	LEU	CA-CB-CG	6.11	129.35	115.30
1	B	355	ARG	NE-CZ-NH1	5.87	123.23	120.30
2	A	325	ASP	CB-CG-OD1	5.75	123.48	118.30
2	A	217	ASP	CB-CG-OD2	-5.57	113.29	118.30
1	B	355	ARG	NE-CZ-NH2	-5.05	117.77	120.30

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	B	146	0	166	4	0
2	A	3174	0	3210	47	0
3	A	146	0	0	5	0
3	B	11	0	0	0	0
All	All	3477	0	3376	48	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 7.

All (48) 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:A:349:THR:HG23	3:A:602:HOH:O	1.83	0.77
2:A:371:GLN:HE21	2:A:375:ASN:HD21	1.35	0.74
2:A:346:ASN:HD22	2:A:348:LYS:H	1.36	0.73
2:A:133:CYS:HB3	2:A:136:ILE:HG22	1.72	0.71
2:A:348:LYS:HG3	3:A:684:HOH:O	1.90	0.71
2:A:426:MET:HE1	2:A:444:ILE:HD12	1.82	0.60
2:A:485:TYR:CB	3:A:692:HOH:O	2.49	0.60
2:A:273:TRP:CD2	2:A:312:PRO:HB3	2.37	0.60
1:B:353:ALA:HB3	2:A:364:ALA:HB1	1.85	0.58
2:A:425:LEU:HG	2:A:440:ILE:HG23	1.85	0.57
2:A:435:LYS:HE3	2:A:438:GLN:HE21	1.68	0.57
2:A:448:PHE:CD1	2:A:460:LEU:HD23	2.41	0.56
2:A:386:LEU:HD22	2:A:421:ILE:HD11	1.87	0.55
2:A:290:VAL:CG1	2:A:330:LYS:HD3	2.38	0.54
2:A:346:ASN:ND2	2:A:348:LYS:H	2.07	0.52
2:A:371:GLN:HE21	2:A:375:ASN:ND2	2.04	0.51
2:A:180:GLU:HB2	2:A:225:TYR:CD1	2.45	0.51
2:A:207:ASP:HB2	2:A:208:PRO:HD3	1.93	0.51
2:A:418:HIS:HD2	3:A:697:HOH:O	1.95	0.50
2:A:101:ARG:CD	2:A:139:GLU:OE1	2.59	0.50
2:A:477:GLN:HE21	2:A:489:LEU:HA	1.77	0.50
2:A:385:VAL:HG12	2:A:397:ALA:HB2	1.93	0.50
2:A:409:THR:H	2:A:412:GLN:NE2	2.12	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:479:HIS:HD2	2:A:481:ASN:HB3	1.79	0.47
1:B:356:LYS:NZ	2:A:283:ASN:ND2	2.62	0.47
2:A:366:ARG:HD2	2:A:368:ASP:OD1	2.15	0.47
2:A:426:MET:CE	2:A:444:ILE:HD12	2.44	0.46
2:A:101:ARG:HD3	2:A:139:GLU:OE1	2.15	0.46
2:A:435:LYS:HD2	2:A:435:LYS:HA	1.75	0.45
2:A:165:PRO:HD3	2:A:203:HIS:CD2	2.52	0.45
2:A:101:ARG:HD2	2:A:139:GLU:OE1	2.17	0.45
2:A:426:MET:HE1	2:A:444:ILE:CD1	2.47	0.44
2:A:168:ILE:HD12	2:A:203:HIS:HB3	1.99	0.44
2:A:168:ILE:CD1	2:A:203:HIS:HB3	2.48	0.44
1:B:355:ARG:HD3	3:A:619:HOH:O	2.18	0.44
2:A:492:ILE:HG23	2:A:496:PHE:HB2	2.00	0.44
2:A:290:VAL:HG13	2:A:330:LYS:HD3	2.00	0.43
2:A:103:LEU:O	2:A:106:ARG:HG2	2.19	0.43
2:A:433:ASP:O	2:A:434:THR:C	2.56	0.42
1:B:356:LYS:HZ1	2:A:283:ASN:ND2	2.17	0.42
2:A:426:MET:CE	2:A:444:ILE:CD1	2.98	0.42
2:A:256:LEU:HD11	2:A:274:ALA:HB3	2.01	0.42
2:A:111:PRO:HB2	2:A:114:ASN:HB2	2.02	0.41
2:A:94:LEU:HD12	2:A:94:LEU:O	2.21	0.41
2:A:263:ASN:C	2:A:263:ASN:OD1	2.59	0.41
2:A:386:LEU:HD21	2:A:425:LEU:HD13	2.04	0.40
2:A:448:PHE:CE1	2:A:460:LEU:HD23	2.56	0.40
2:A:111:PRO:O	2:A:115:ILE:HG12	2.21	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	B	17/19 (90%)	16 (94%)	1 (6%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	A	423/510 (83%)	412 (97%)	10 (2%)	1 (0%)	52 68
All	All	440/529 (83%)	428 (97%)	11 (2%)	1 (0%)	52 68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	434	THR

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	B	16/16 (100%)	12 (75%)	4 (25%)	1 0
2	A	341/426 (80%)	321 (94%)	20 (6%)	24 36
All	All	357/442 (81%)	333 (93%)	24 (7%)	20 29

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

Mol	Chain	Res	Type
1	B	356	LYS
1	B	357	GLU
1	B	363	SER
1	B	370	THR
2	A	75	ASN
2	A	101	ARG
2	A	106	ARG
2	A	207	ASP
2	A	223	CYS
2	A	240	LYS
2	A	258	ARG
2	A	288	MET
2	A	348	LYS
2	A	349	THR
2	A	353	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	390	ASP
2	A	392	LYS
2	A	404	TYR
2	A	434	THR
2	A	438	GLN
2	A	442	ASP
2	A	445	SER
2	A	452	GLU
2	A	483	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	86	ASN
2	A	261	HIS
2	A	283	ASN
2	A	346	ASN
2	A	350	ASN
2	A	352	GLN
2	A	375	ASN
2	A	412	GLN
2	A	438	GLN
2	A	446	ASN
2	A	479	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 carbohydrates 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, 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	B	19/19 (100%)	-0.19	0 100 100	39, 48, 62, 67	0
2	A	425/510 (83%)	0.11	20 (4%) 35 40	32, 45, 84, 94	0
All	All	444/529 (83%)	0.09	20 (4%) 37 42	32, 45, 82, 94	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	107	GLU	4.5
2	A	496	PHE	4.5
2	A	478	ARG	4.0
2	A	74	VAL	3.9
2	A	455	GLY	3.6
2	A	494	LYS	3.2
2	A	491	LEU	3.0
2	A	452	GLU	3.0
2	A	457	THR	2.8
2	A	495	TYR	2.7
2	A	454	LEU	2.6
2	A	75	ASN	2.4
2	A	493	GLU	2.4
2	A	477	GLN	2.2
2	A	470	LEU	2.2
2	A	492	ILE	2.2
2	A	461	SER	2.1
2	A	449	GLN	2.1
2	A	460	LEU	2.1
2	A	458	GLU	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 carbohydrates 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.