Full wwPDB X-ray Structure Validation Report ⓘ

Dec 3, 2023 – 01:29 pm GMT

PDB ID	:	1V1H
Title	:	Adenovirus fibre shaft sequence N-terminally fused to the bacteriophage T4 fibritin foldon trimerisation motif with a short linker
Authors	:	Papanikolopoulou, K.; Teixeira, S.; Belrhali, H.; Forsyth, V.T.; Mitraki, A.; van Raaij, M.J.
Deposited on	:	2004-04-16
Resolution	:	1.90 Å(reported)

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

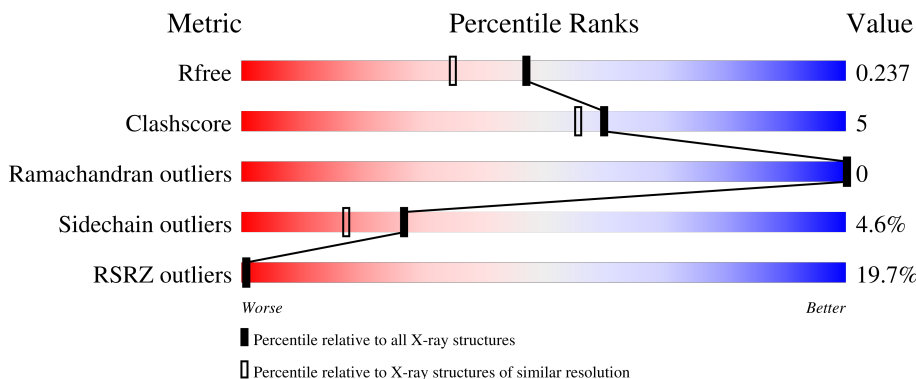|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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.
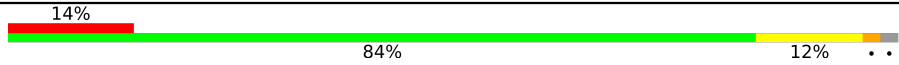
Metric	Percentile Ranks	Value
Rfree		0.237
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		4.6%
RSRZ outliers		19.7%

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	103	26% / 82% / 17% •
1	B	103	25% / 85% / 12% • •
1	C	103	28% / 90% / 8% •
1	D	103	11% / 85% / 12% • •
1	E	103	13% / 89% / 10% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	103	

14%

84% 12%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5033 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 FIBRITIN, FIBER PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	103	Total	C	N	O	S	0	2	0
			766	481	126	158	1			
1	B	101	Total	C	N	O	S	0	1	0
			755	474	124	156	1			
1	C	101	Total	C	N	O	S	0	3	0
			761	480	125	155	1			
1	D	101	Total	C	N	O	S	0	1	0
			754	475	124	154	1			
1	E	103	Total	C	N	O	S	0	4	0
			778	489	128	160	1			
1	F	101	Total	C	N	O	S	0	7	0
			777	491	126	159	1			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	401	GLY	-	linker	UNP P10104
B	402	SER	-	linker	UNP P10104
A	478	LEU	PHE	conflict	UNP P10104
B	401	GLY	-	linker	UNP P10104
B	402	SER	-	linker	UNP P10104
B	478	LEU	PHE	conflict	UNP P10104
C	401	GLY	-	linker	UNP P10104
C	402	SER	-	linker	UNP P10104
C	478	LEU	PHE	conflict	UNP P10104
D	401	GLY	-	linker	UNP P10104
D	402	SER	-	linker	UNP P10104
D	478	LEU	PHE	conflict	UNP P10104
E	401	GLY	-	linker	UNP P10104
E	402	SER	-	linker	UNP P10104
E	478	LEU	PHE	conflict	UNP P10104
F	401	GLY	-	linker	UNP P10104
F	402	SER	-	linker	UNP P10104

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	478	LEU	PHE	conflict	UNP P10104

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	77	Total 77	O 77	0	0
2	B	72	Total 72	O 72	0	0
2	C	73	Total 73	O 73	0	0
2	D	71	Total 71	O 71	0	0
2	E	74	Total 74	O 74	0	0
2	F	75	Total 75	O 75	0	0

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: FIBRITIN, FIBER PROTEIN
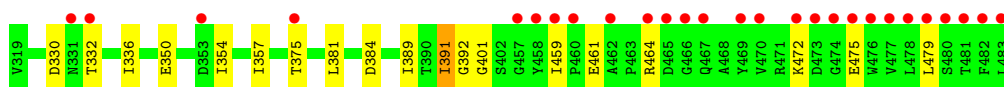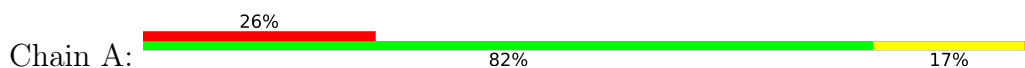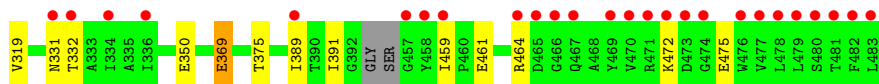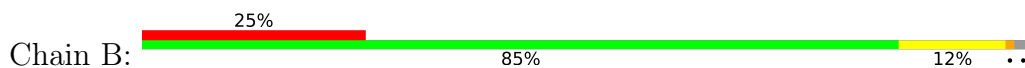
Chain A:

• Molecule 1: FIBRITIN, FIBER PROTEIN

Chain B:

• Molecule 1: FIBRITIN, FIBER PROTEIN

Chain C:

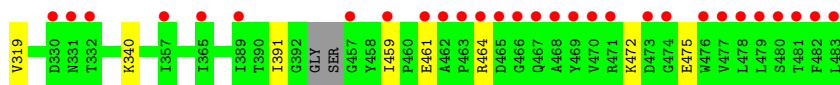• Molecule 1: FIBRITIN, FIBER PROTEIN
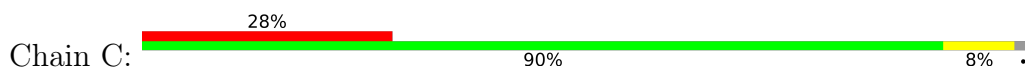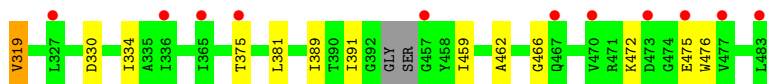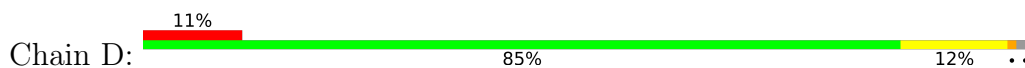
Chain D:

• Molecule 1: FIBRITIN, FIBER PROTEIN

Chain E:

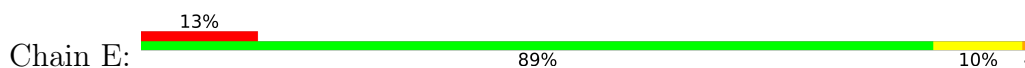• Molecule 1: FIBRITIN, FIBER PROTEIN

Chain F:

14%

84%

12%
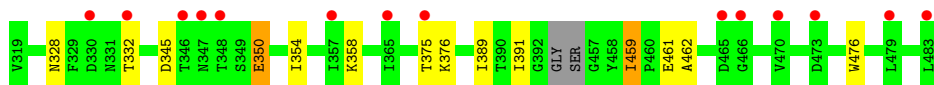
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	77.77Å 183.33Å 58.97Å	Depositor
a, b, c, α, β, γ	90.00° 129.29° 90.00°	
Resolution (Å)	20.00 – 1.90	Depositor
	19.98 – 1.90	EDS
% Data completeness	89.0 (20.00-1.90)	Depositor
(in resolution range)	88.9 (19.98-1.90)	EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	4.85 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.182 , 0.240	Depositor
	0.184 , 0.237	DCC
R_{free} test set	1625 reflections (3.65%)	wwPDB-VP
Wilson B-factor (Å^2)	22.3	Xtriage
Anisotropy	0.463	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 58.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5033	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.69% 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	0.75	0/786	0.87	2/1061 (0.2%)
1	B	0.72	0/770	0.84	0/1038
1	C	0.75	0/784	0.84	0/1056
1	D	0.75	0/769	0.85	1/1037 (0.1%)
1	E	0.73	0/806	0.84	2/1087 (0.2%)
1	F	0.71	0/816	0.81	0/1100
All	All	0.73	0/4731	0.84	5/6379 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	345	ASP	CB-CG-OD2	5.46	123.21	118.30
1	A	384	ASP	CB-CG-OD2	5.44	123.20	118.30
1	D	330	ASP	CB-CG-OD2	5.25	123.03	118.30
1	E	465	ASP	CB-CG-OD2	5.06	122.85	118.30
1	A	330	ASP	CB-CG-OD2	5.06	122.85	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	391	ILE	Peptide
1	D	391	ILE	Peptide

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	766	0	754	12	0
1	B	755	0	739	8	0
1	C	761	0	754	7	0
1	D	754	0	744	7	0
1	E	778	0	766	8	0
1	F	777	0	768	13	0
2	A	77	0	0	2	0
2	B	72	0	0	1	0
2	C	73	0	0	2	0
2	D	71	0	0	0	0
2	E	74	0	0	1	0
2	F	75	0	0	1	0
All	All	5033	0	4525	42	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 (42) 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:C:340[B]:LYS:NZ	2:C:2018:HOH:O	1.81	1.12
1:C:340[A]:LYS:NZ	2:C:2019:HOH:O	1.97	0.90
1:D:459:ILE:HD13	1:F:476:TRP:HZ2	1.49	0.78
1:B:461:GLU:HA	1:C:459:ILE:HD11	1.67	0.77
1:D:319:VAL:HG21	1:D:334:ILE:HD11	1.69	0.74
1:B:331:ASN:O	1:B:332:THR:HB	1.94	0.68
1:B:350:GLU:HG3	2:B:2037:HOH:O	1.95	0.67
1:A:392:GLY:HA3	2:A:2062:HOH:O	1.97	0.64
1:A:350:GLU:HG2	1:A:354:ILE:HD12	1.80	0.63
1:D:375:THR:HG21	1:D:389:ILE:HD11	1.82	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:375:THR:HG21	1:A:389:ILE:HD11	1.86	0.58
1:A:459:ILE:HD11	1:C:461:GLU:HA	1.88	0.56
1:D:466:GLY:HA2	1:E:473:ASP:OD2	2.06	0.56
1:D:459:ILE:HD13	1:F:476:TRP:CZ2	2.37	0.54
1:B:375:THR:HG21	1:B:389:ILE:HD11	1.91	0.53
1:A:391:ILE:HG22	1:C:391:ILE:HD11	1.89	0.53
1:D:381:LEU:HD12	1:F:391[A]:ILE:HG13	1.90	0.52
1:E:402:SER:HB2	1:E:458:TYR:CD2	2.44	0.52
1:B:461:GLU:HA	1:C:459:ILE:CD1	2.38	0.51
1:A:479:LEU:HB3	2:A:2074:HOH:O	2.11	0.51
1:F:328[B]:ASN:ND2	2:F:2007:HOH:O	2.45	0.49
1:E:389:ILE:HD12	1:F:375[B]:THR:CG2	2.43	0.48
1:F:345:ASP:HB2	1:F:358:LYS:HD2	1.96	0.47
1:A:391:ILE:CG2	1:C:391:ILE:HD11	2.46	0.45
1:A:381:LEU:HD23	1:A:381:LEU:HA	1.78	0.45
1:A:350:GLU:CG	1:A:354:ILE:HD12	2.47	0.45
1:F:350:GLU:H	1:F:350:GLU:HG2	1.65	0.44
1:E:459:ILE:HG13	1:F:459:ILE:HD11	2.00	0.43
1:F:461[A]:GLU:HG3	1:F:462:ALA:O	2.18	0.43
2:E:2067:HOH:O	1:F:376:LYS:HE2	2.18	0.43
1:A:461:GLU:HA	1:B:459:ILE:HD11	2.01	0.42
1:E:462:ALA:HA	1:E:476:TRP:CD1	2.54	0.42
1:E:389:ILE:HD12	1:F:375[B]:THR:HG22	2.01	0.42
1:F:350:GLU:HB2	1:F:354:ILE:HB	2.02	0.42
1:D:462:ALA:HA	1:D:476:TRP:CD1	2.55	0.41
1:F:375[A]:THR:HG21	1:F:389:ILE:HD11	2.02	0.41
1:A:392:GLY:HA3	1:A:401:GLY:HA3	1.61	0.41
1:A:336:ILE:HG21	1:A:357:ILE:HD12	2.02	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	103/103 (100%)	99 (96%)	4 (4%)	0	100	100
1	B	98/103 (95%)	94 (96%)	4 (4%)	0	100	100
1	C	100/103 (97%)	97 (97%)	3 (3%)	0	100	100
1	D	98/103 (95%)	95 (97%)	3 (3%)	0	100	100
1	E	105/103 (102%)	101 (96%)	4 (4%)	0	100	100
1	F	104/103 (101%)	98 (94%)	6 (6%)	0	100	100
All	All	608/618 (98%)	584 (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	83/81 (102%)	79 (95%)	4 (5%)	25	16
1	B	81/81 (100%)	74 (91%)	7 (9%)	10	4
1	C	83/81 (102%)	79 (95%)	4 (5%)	25	16
1	D	81/81 (100%)	78 (96%)	3 (4%)	34	25
1	E	85/81 (105%)	82 (96%)	3 (4%)	36	27
1	F	87/81 (107%)	84 (97%)	3 (3%)	37	28
All	All	500/486 (103%)	476 (95%)	24 (5%)	27	16

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

Mol	Chain	Res	Type
1	A	332	THR
1	A	464	ARG
1	A	472	LYS
1	A	475	GLU
1	B	319	VAL
1	B	369[A]	GLU
1	B	369[B]	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	391	ILE
1	B	464	ARG
1	B	472	LYS
1	B	475	GLU
1	C	319	VAL
1	C	464	ARG
1	C	472	LYS
1	C	475	GLU
1	D	319	VAL
1	D	472	LYS
1	D	475	GLU
1	E	369[A]	GLU
1	E	369[B]	GLU
1	E	464	ARG
1	F	332	THR
1	F	350	GLU
1	F	459	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	331	ASN

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	103/103 (100%)	1.39	27 (26%)	0	0	14, 28, 85, 90	0
1	B	101/103 (98%)	1.13	26 (25%)	0	0	14, 28, 85, 90	0
1	C	101/103 (98%)	1.12	29 (28%)	0	0	13, 28, 85, 90	0
1	D	101/103 (98%)	0.53	11 (10%)	5	6	15, 29, 59, 68	0
1	E	103/103 (100%)	0.54	13 (12%)	3	4	16, 29, 60, 68	0
1	F	101/103 (98%)	0.65	14 (13%)	2	3	17, 30, 60, 68	0
All	All	610/618 (98%)	0.89	120 (19%)	1	1	13, 29, 82, 90	0

All (120) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	479	LEU	8.3
1	A	470	VAL	8.2
1	A	466	GLY	8.1
1	C	481	THR	7.9
1	A	477	VAL	7.3
1	C	463	PRO	7.2
1	A	473	ASP	7.1
1	B	483	LEU	6.9
1	B	464	ARG	6.8
1	A	483	LEU	6.7
1	C	483	LEU	6.7
1	B	470	VAL	6.6
1	C	470	VAL	6.5
1	A	474	GLY	6.2
1	B	481	THR	6.1
1	A	467	GLN	6.1
1	B	473	ASP	6.1
1	A	469	TYR	6.0
1	A	465	ASP	5.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	477	VAL	5.5
1	B	466	GLY	5.5
1	A	478	LEU	5.3
1	C	467	GLN	5.2
1	B	480	SER	5.1
1	A	481	THR	5.1
1	E	483	LEU	4.7
1	A	476	TRP	4.7
1	B	469	TYR	4.6
1	C	468	ALA	4.5
1	B	482	PHE	4.5
1	C	465	ASP	4.5
1	C	476	TRP	4.4
1	C	459	ILE	4.4
1	A	482	PHE	4.4
1	B	479	LEU	4.3
1	B	472	LYS	4.3
1	D	483	LEU	4.3
1	C	469	TYR	4.2
1	A	457	GLY	4.2
1	B	476	TRP	4.1
1	B	331	ASN	4.1
1	F	470	VAL	4.1
1	B	477	VAL	4.0
1	A	480	SER	3.9
1	D	470	VAL	3.9
1	E	331	ASN	3.8
1	C	480	SER	3.8
1	C	462	ALA	3.7
1	B	458	TYR	3.7
1	B	465	ASP	3.6
1	A	331	ASN	3.6
1	B	474	GLY	3.6
1	A	332	THR	3.6
1	B	467	GLN	3.5
1	A	475	GLU	3.5
1	F	483	LEU	3.5
1	E	466	GLY	3.5
1	B	457	GLY	3.4
1	F	357	ILE	3.4
1	A	459	ILE	3.3
1	B	459	ILE	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	458	TYR	3.2
1	E	332	THR	3.2
1	C	466	GLY	3.2
1	E	481	THR	3.2
1	C	478	LEU	3.1
1	E	473	ASP	3.1
1	B	336	ILE	3.0
1	F	348	THR	3.0
1	C	457	GLY	3.0
1	D	457	GLY	3.0
1	F	473	ASP	2.9
1	C	482	PHE	2.9
1	B	471	ARG	2.9
1	F	346	THR	2.9
1	F	479	LEU	2.9
1	C	464	ARG	2.9
1	E	464	ARG	2.9
1	C	474	GLY	2.8
1	C	473	ASP	2.8
1	F	465	ASP	2.8
1	C	331	ASN	2.8
1	D	327	LEU	2.7
1	B	334	ILE	2.7
1	F	332	THR	2.7
1	A	464	ARG	2.6
1	B	478	LEU	2.6
1	F	375[A]	THR	2.6
1	C	461	GLU	2.6
1	C	357	ILE	2.6
1	C	330	ASP	2.6
1	A	472	LYS	2.5
1	D	477	VAL	2.5
1	D	467	GLN	2.5
1	E	336	ILE	2.5
1	A	462	ALA	2.5
1	A	353	ASP	2.4
1	D	375	THR	2.4
1	C	479	LEU	2.3
1	D	365	ILE	2.3
1	F	347	ASN	2.3
1	E	467	GLN	2.2
1	D	336	ILE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	332	THR	2.2
1	E	480	SER	2.2
1	D	475	GLU	2.2
1	F	466	GLY	2.2
1	C	332	THR	2.1
1	D	473	ASP	2.1
1	C	365[A]	ILE	2.1
1	F	330	ASP	2.1
1	C	389	ILE	2.1
1	F	365[A]	ILE	2.1
1	A	460	PRO	2.1
1	E	470	VAL	2.1
1	B	389	ILE	2.1
1	A	375	THR	2.0
1	C	471	ARG	2.0
1	E	357	ILE	2.0
1	E	365[A]	ILE	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.