Full wwPDB NMR Structure Validation Report (i)

Mar 1, 2022 – 05:03 PM EST

PDB ID	:	2FTU
Title	:	solution structure of domain 3 of RAP
Authors	:	Lee, D.; Walsh, J.D.; Wang, Y.-X.
Deposited on	:	2006-01-24

This is a Full wwPDB NMR 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

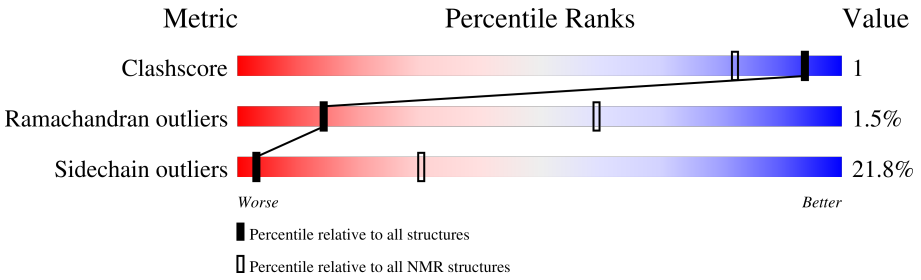
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	118	69% 14% 18%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:17-A:113 (97)	0.41	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 9, 11, 12, 13, 14, 15, 16, 18, 19, 20
2	7, 8, 10
Single-model clusters	17

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1934 atoms, of which 956 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Alpha-2-macroglobulin receptor-associated protein, domain 3.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	118	1934	603	956	188	187	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 69% 14% 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 L21 L24 N29 L30 F38 R39 E46 A47 K48 R77 L87 L88 R91 L95 G96 Y97 L103 R109 R114 H115 N116 E117 L118

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 61% 20% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 D20 L21 L24 A28 F38 R39 L42 E46 E50 Y55 L59 E60 E64 R67 R80 S81 R82 E83 K84 L87 L88 E89 K93 H102 L103 A113 R114

H115 N116 E117 L118

4.2.2 Score per residue for model 2

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 63% 18% •• 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 V18 I19 L24 N29 F38 R39 E40 E46 L66 R67 S71 V72 G73 R77 R80 E83 L87 I88 E89 G90 R91 L95 K100 L103 Q104 D105 R114 H115 N116

4.2.3 Score per residue for model 3

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A:

4.2.4 Score per residue for model 4

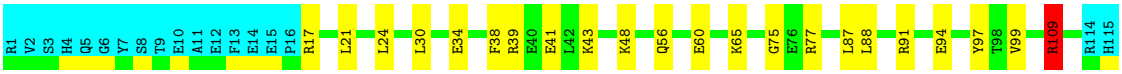
• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A:

4.2.7 Score per residue for model 7

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 63% 17% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 D20 L24 N29 F38 R39 E40 K43 E46 A47 K48 H54 Y55 Q56 A62 H63 E64 K65 E83 K84 H85 A86 L87 L95 G96 Y97 R109 A113 R114 H115 N116 E117

L118

4.2.8 Score per residue for model 8

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 61% 20% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 L21 L24 N29 L30 K33 E34 L35 F38 R39 K43 E46 I49 E50 K51 H52 K57 K65 L66 R67 H68 R80 S81 L87 R91 L95 K101 S111

R114 H115 N116 E117 L118

4.2.9 Score per residue for model 9

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 61% 19% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 D20 L24 N29 L30 F38 R39 K48 I49 E50 L59 E60 L66 R67 R77 R82 E89 G90 R91 T92 K93 E94 L95 G96 Y97 K100 K101 H102 L103 A113 R114 H115

N116 E117 L118

4.2.10 Score per residue for model 10

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

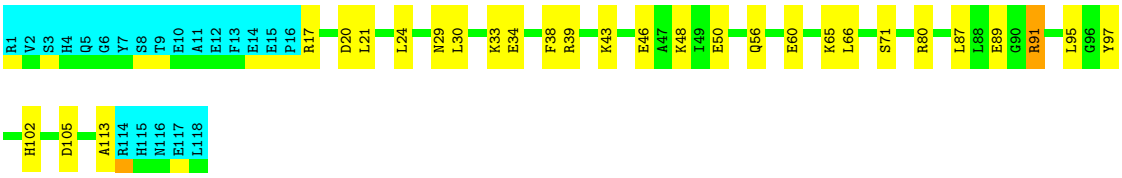
Chain A: 64% 15% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 V18 I19 L24 A28 E34 F38 R39 E46 I49 H52 L66 L87 L88 R91 L95 G96 Y97 L103 S107 R112 A113 R114 H115 N116 E117 L118

4.2.11 Score per residue for model 11

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

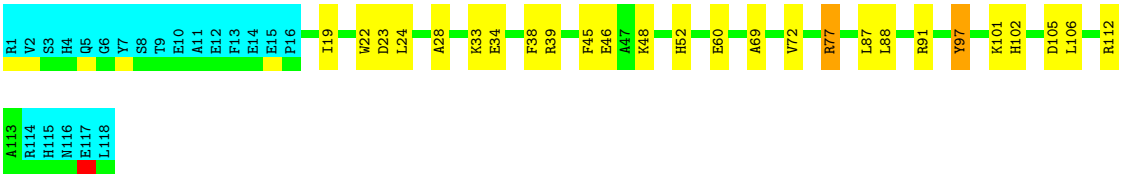
Chain A: 58% 23% • 18%



4.2.12 Score per residue for model 12

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

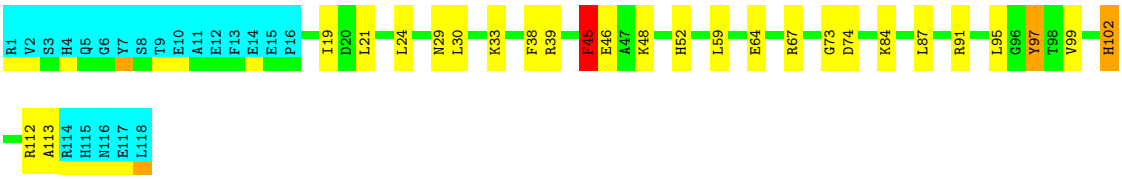
Chain A: 60% 20% • 18%



4.2.13 Score per residue for model 13

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

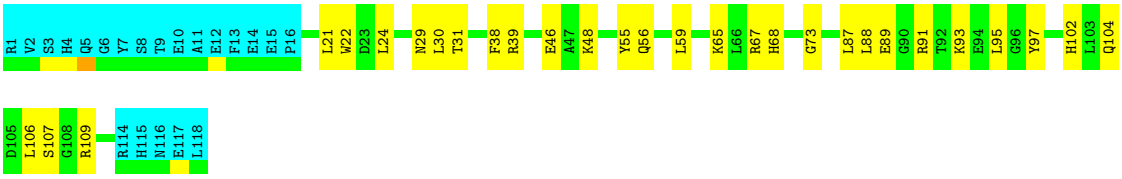
Chain A: 60% 19% • • 18%



4.2.14 Score per residue for model 14

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

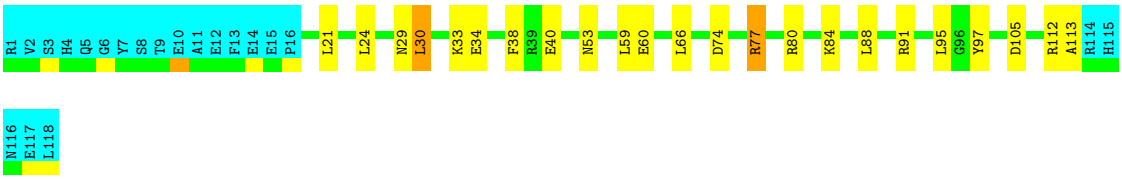
Chain A: 58% 25% 18%



4.2.15 Score per residue for model 15

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 63% 18% • 18%



4.2.16 Score per residue for model 16

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 69% 14% 18%

4.2.17 Score per residue for model 17

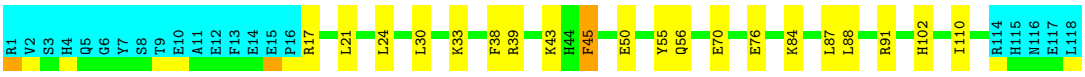
• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 63% 16% • 18%

4.2.18 Score per residue for model 18

• Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 65% 16% • 18%



4.2.19 Score per residue for model 19

- Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 53% 25% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 V18 I19 D20 L21 W22 D23 L24 L30 E34 F38 R39 E40 E41 L42 K43 H44 F45 E46 A47 K48 H52 K65 H68 S71 D74 R77 R80 K84 L87 R91

Y97 K101 H102 L103 Q104 D105 L106 S107 G108 R109 A113 R114 H115 N116 E117 L118

4.2.20 Score per residue for model 20

- Molecule 1: Alpha-2-macroglobulin receptor-associated protein, domain 3

Chain A: 58% 21% • 18%

R1 V2 S3 H4 Q5 G6 Y7 S8 T9 E10 A11 E12 F13 E14 E15 P16 R17 V18 L21 W22 D23 L24 A28 M29 L30 F38 R39 L42 K43 K48 H52 H63 L66 R77 V78 S79 R80 K84 L87 L88 R91 Y97 H102 L103 S107

S111 R112 A113 R114 H115 N116 E117 L118

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics*.

Of the ? calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	refinement	2.1.4
CYANA	structure solution	2.1.4

No chemical shift data was provided.

6 Model quality (i)

6.1 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 (average) 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	0.55±0.01	0±0/814 (0.0± 0.0%)	1.07±0.03	1±1/1088 (0.1± 0.1%)
All	All	0.55	0/16280 (0.0%)	1.07	17/21760 (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	Planarity
1	A	0.0±0.0	2.0±1.0
All	All	0	40

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	67	ARG	NE-CZ-NH2	-6.73	116.94	120.30	9	1
1	A	109	ARG	NE-CZ-NH2	-6.71	116.94	120.30	17	1
1	A	55	TYR	CB-CG-CD2	-6.68	116.99	121.00	3	3
1	A	82	ARG	NE-CZ-NH2	-6.58	117.01	120.30	4	2
1	A	91	ARG	NE-CZ-NH1	6.39	123.49	120.30	11	1
1	A	39	ARG	NE-CZ-NH2	-5.84	117.38	120.30	2	1
1	A	77	ARG	NE-CZ-NH2	-5.80	117.40	120.30	12	1
1	A	80	ARG	NE-CZ-NH2	-5.77	117.41	120.30	4	2
1	A	106	LEU	CB-CG-CD1	5.69	120.68	111.00	12	1
1	A	99	VAL	CA-CB-CG1	5.44	119.06	110.90	5	1
1	A	91	ARG	NE-CZ-NH2	-5.42	117.59	120.30	19	1
1	A	80	ARG	NE-CZ-NH1	5.21	122.90	120.30	20	1
1	A	45	PHE	CB-CG-CD2	-5.11	117.22	120.80	18	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	77	ARG	Sidechain	6
1	A	97	TYR	Sidechain	6
1	A	109	ARG	Sidechain	5
1	A	112	ARG	Peptide,Sidechain	5
1	A	91	ARG	Sidechain	4
1	A	80	ARG	Sidechain	3
1	A	39	ARG	Sidechain	3
1	A	55	TYR	Sidechain	2
1	A	67	ARG	Sidechain	2
1	A	17	ARG	Sidechain	2
1	A	82	ARG	Sidechain	1
1	A	45	PHE	Sidechain	1

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	801	802	802	2±2
All	All	16020	16040	16040	40

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 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:52:HIS:HB2	1:A:95:LEU:HD12	0.68	1.66	10	1
1:A:22:TRP:CH2	1:A:103:LEU:HD23	0.65	2.25	3	2
1:A:22:TRP:CZ2	1:A:106:LEU:HD12	0.59	2.32	19	1
1:A:20:ASP:O	1:A:24:LEU:HD22	0.59	1.97	6	3
1:A:62:ALA:HB1	1:A:85:HIS:CE1	0.58	2.33	4	2
1:A:69:ALA:O	1:A:72:VAL:HG22	0.55	2.02	12	2
1:A:19:ILE:HD12	1:A:45:PHE:CZ	0.54	2.37	13	1
1:A:18:VAL:HG21	1:A:103:LEU:HD13	0.54	1.79	10	1
1:A:29:ASN:H	1:A:30:LEU:HD12	0.53	1.64	14	2
1:A:103:LEU:HD12	1:A:103:LEU:C	0.53	2.24	10	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:102:HIS:CD2	1:A:106:LEU:HD13	0.51	2.40	14	1
1:A:22:TRP:CE3	1:A:42:LEU:HD11	0.51	2.41	19	1
1:A:38:PHE:CE2	1:A:42:LEU:HD12	0.50	2.40	20	1
1:A:48:LYS:HG3	1:A:95:LEU:HD21	0.49	1.84	7	1
1:A:69:ALA:HA	1:A:72:VAL:HG13	0.48	1.85	17	1
1:A:52:HIS:CB	1:A:95:LEU:HD12	0.48	2.39	10	1
1:A:30:LEU:HD22	1:A:30:LEU:N	0.46	2.24	19	1
1:A:22:TRP:CH2	1:A:103:LEU:HD22	0.46	2.46	20	1
1:A:22:TRP:CH2	1:A:103:LEU:HD12	0.45	2.46	17	1
1:A:22:TRP:CZ2	1:A:103:LEU:HD23	0.45	2.47	4	1
1:A:30:LEU:HD22	1:A:30:LEU:H	0.45	1.71	19	1
1:A:45:PHE:CE2	1:A:106:LEU:HD11	0.45	2.47	19	1
1:A:18:VAL:HG11	1:A:103:LEU:HD13	0.44	1.89	19	1
1:A:18:VAL:HG21	1:A:103:LEU:HD11	0.44	1.87	2	1
1:A:22:TRP:CH2	1:A:106:LEU:HD12	0.44	2.48	19	1
1:A:52:HIS:CD2	1:A:95:LEU:HB3	0.43	2.47	8	1
1:A:30:LEU:HD12	1:A:30:LEU:N	0.43	2.28	14	1
1:A:20:ASP:O	1:A:24:LEU:HD12	0.43	2.13	1	1
1:A:72:VAL:HG23	1:A:74:ASP:H	0.43	1.74	17	1
1:A:35:LEU:C	1:A:35:LEU:HD13	0.41	2.36	8	1
1:A:18:VAL:HG21	1:A:103:LEU:HD22	0.41	1.92	19	1
1:A:74:ASP:CG	1:A:77:ARG:HH12	0.41	2.19	15	1
1:A:99:VAL:HA	1:A:102:HIS:CE1	0.40	2.51	13	1
1:A:30:LEU:HD22	1:A:113:ALA:CB	0.40	2.46	15	1

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	97/118 (82%)	86±2 (89±2%)	9±2 (10±2%)	2±1 (2±1%)	14	59
All	All	1940/2360 (82%)	1724 (89%)	186 (10%)	30 (2%)	14	59

All 9 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	113	ALA	8
1	A	29	ASN	8
1	A	28	ALA	4
1	A	73	GLY	3
1	A	17	ARG	3
1	A	75	GLY	1
1	A	30	LEU	1
1	A	74	ASP	1
1	A	18	VAL	1

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	84/103 (82%)	66±3 (78±4%)	18±3 (22±4%)	3 30
All	All	1680/2060 (82%)	1314 (78%)	366 (22%)	3 30

All 70 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	24	LEU	20
1	A	38	PHE	20
1	A	91	ARG	16
1	A	87	LEU	15
1	A	39	ARG	14
1	A	21	LEU	13
1	A	95	LEU	13
1	A	46	GLU	12
1	A	88	LEU	12
1	A	48	LYS	11
1	A	97	TYR	10
1	A	84	LYS	9
1	A	65	LYS	9
1	A	89	GLU	8
1	A	66	LEU	8
1	A	105	ASP	8
1	A	34	GLU	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	60	GLU	7
1	A	102	HIS	7
1	A	30	LEU	7
1	A	43	LYS	7
1	A	50	GLU	6
1	A	56	GLN	6
1	A	33	LYS	6
1	A	17	ARG	5
1	A	59	LEU	5
1	A	67	ARG	5
1	A	103	LEU	5
1	A	71	SER	5
1	A	52	HIS	5
1	A	19	ILE	4
1	A	40	GLU	4
1	A	80	ARG	4
1	A	109	ARG	4
1	A	107	SER	4
1	A	64	GLU	3
1	A	93	LYS	3
1	A	104	GLN	3
1	A	23	ASP	3
1	A	63	HIS	3
1	A	68	HIS	3
1	A	101	LYS	3
1	A	45	PHE	3
1	A	77	ARG	3
1	A	83	GLU	2
1	A	100	LYS	2
1	A	18	VAL	2
1	A	35	LEU	2
1	A	54	HIS	2
1	A	79	SER	2
1	A	76	GLU	2
1	A	111	SER	2
1	A	22	TRP	2
1	A	112	ARG	2
1	A	106	LEU	2
1	A	42	LEU	1
1	A	29	ASN	1
1	A	41	GLU	1
1	A	94	GLU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	57	LYS	1
1	A	81	SER	1
1	A	82	ARG	1
1	A	20	ASP	1
1	A	31	THR	1
1	A	53	ASN	1
1	A	32	ASP	1
1	A	51	LYS	1
1	A	70	GLU	1
1	A	110	ILE	1
1	A	74	ASP	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided