Full wwPDB NMR Structure Validation Report (i)

Feb 26, 2022 – 04:47 PM EST

PDB ID	:	2CRZ
Title	:	Solution structure of the fifth FNIII domain of human fibronectin type III domain containing protein 3a
Authors	:	Kurosaki, C.; Hayashi, F.; Yoshida, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-05-20

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:

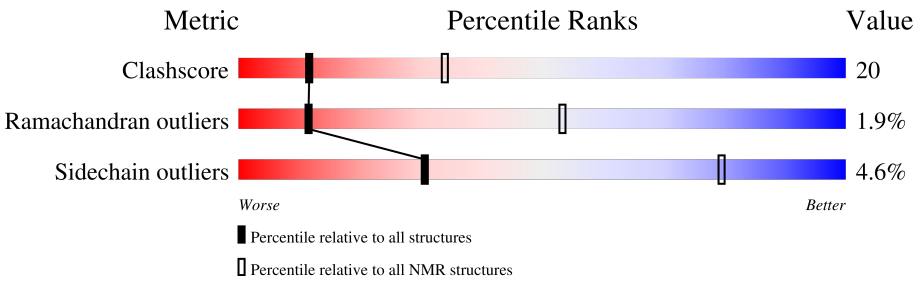
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	110	61% 25% • 12%

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:8-A:104 (97)	0.12	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 4 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1598 atoms, of which 792 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Fibronectin type-III domain containing protein 3a.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	110	1598	502	792	139	159	6	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q9Y2H6
A	2	SER	-	cloning artifact	UNP Q9Y2H6
A	3	SER	-	cloning artifact	UNP Q9Y2H6
A	4	GLY	-	cloning artifact	UNP Q9Y2H6
A	5	SER	-	cloning artifact	UNP Q9Y2H6
A	6	SER	-	cloning artifact	UNP Q9Y2H6
A	7	GLY	-	cloning artifact	UNP Q9Y2H6
A	105	SER	-	cloning artifact	UNP Q9Y2H6
A	106	GLY	-	cloning artifact	UNP Q9Y2H6
A	107	PRO	-	cloning artifact	UNP Q9Y2H6
A	108	SER	-	cloning artifact	UNP Q9Y2H6
A	109	SER	-	cloning artifact	UNP Q9Y2H6
A	110	GLY	-	cloning artifact	UNP Q9Y2H6

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: Fibronectin type-III domain containing protein 3a

Chain A: 61% 25% • 12%

G1 S2 S3 G4 S5 S6 G7 C12 L17 Q18 G19 R20 P21 I26 W30 D36 S39 P40 I41 Y44 M48 S49 P50 I51 E52 K53 R57 Y60 C67 T77 Y78 S79 F80 R83 A84 A85 F90 F93 S94 D98 T101 S105

G106 P107 S108 S109 G110

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: Fibronectin type-III domain containing protein 3a

Chain A: 58% 28% •• 12%

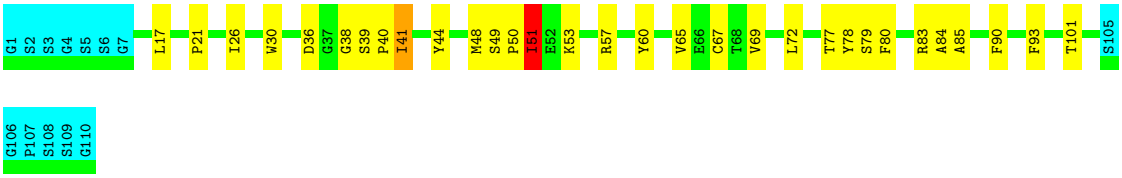
G1 S2 S3 G4 S5 S6 G7 L13 L17 Q18 G19 R20 P21 I26 Q27 D36 G37 G38 I41 S42 C43 Y44 E47 M48 S49 P50 I51 E52 K53 R57 Y60 C67 T77 Y78 S79 F80 R81 L82 R83 A84 A85 F90 F93 K96 T101

S105 G106 P107 S108 S109 G110

4.2.2 Score per residue for model 2

• Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 60% 26% •• 12%



4.2.3 Score per residue for model 3

• Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 65% 22% • 12%

4.2.4 Score per residue for model 4

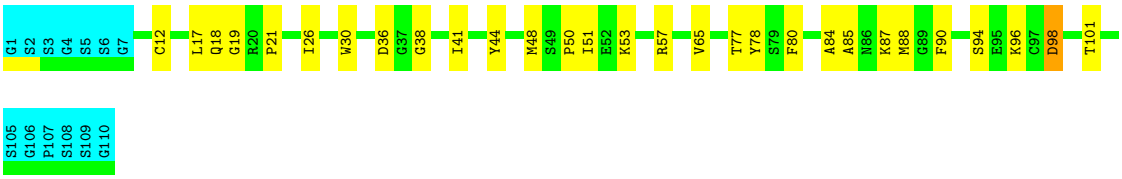
• Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 65% 22% • 12%

4.2.5 Score per residue for model 5

• Molecule 1: Fibronectin type-III domain containing protein 3a

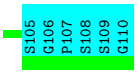
Chain A: 62% 25% • 12%



4.2.6 Score per residue for model 6

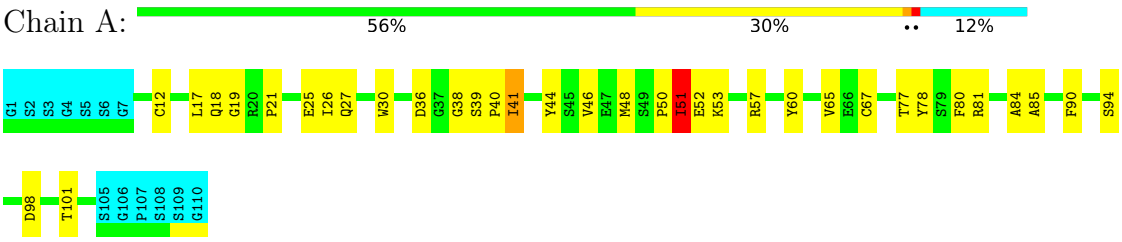
• Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 59% 26% • 12%



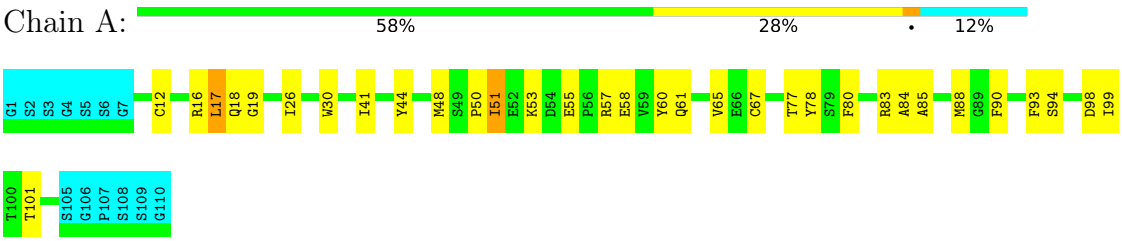
4.2.7 Score per residue for model 7

• Molecule 1: Fibronectin type-III domain containing protein 3a



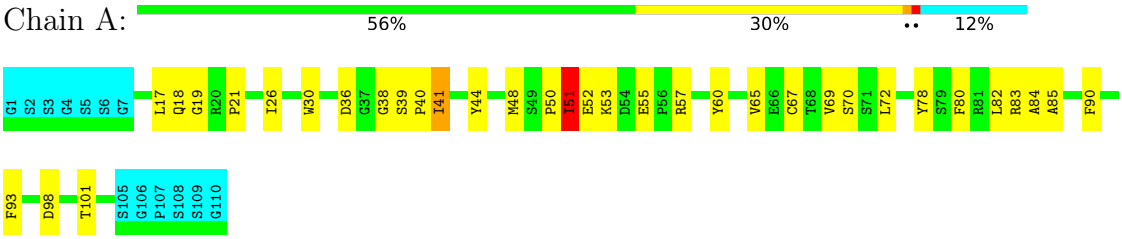
4.2.8 Score per residue for model 8

• Molecule 1: Fibronectin type-III domain containing protein 3a



4.2.9 Score per residue for model 9

• Molecule 1: Fibronectin type-III domain containing protein 3a



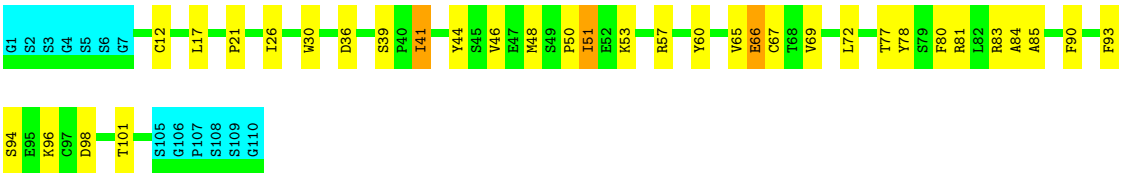
4.2.10 Score per residue for model 10

• Molecule 1: Fibronectin type-III domain containing protein 3a

4.2.11 Score per residue for model 11

- Molecule 1: Fibronectin type-III domain containing protein 3a

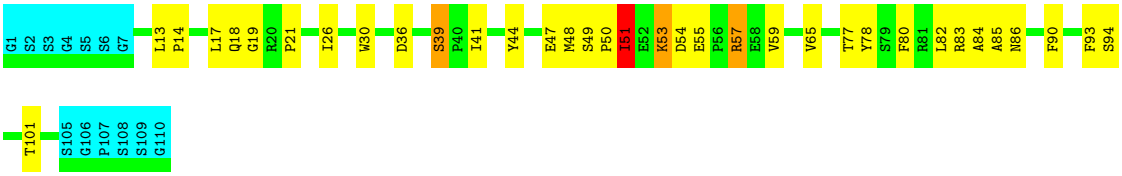
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Fibronectin type-III domain containing protein 3a

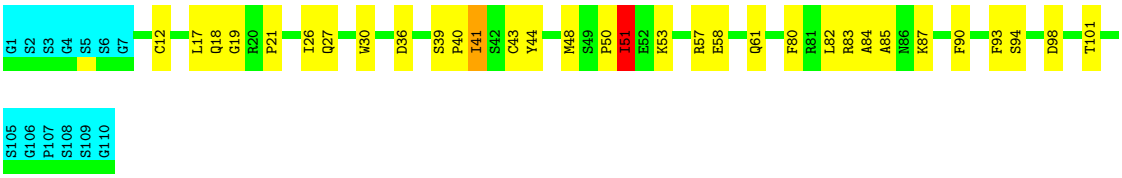
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Fibronectin type-III domain containing protein 3a

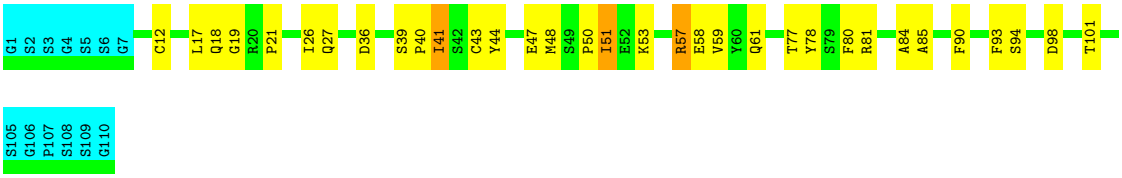
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Fibronectin type-III domain containing protein 3a

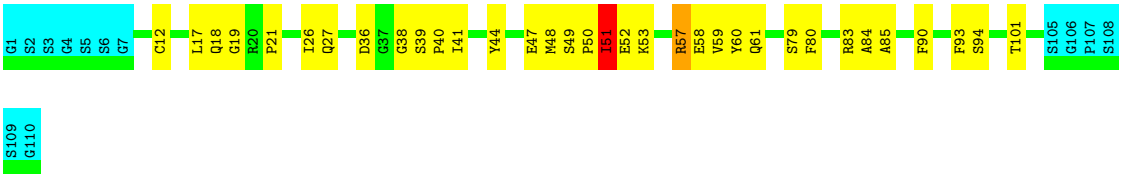
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Fibronectin type-III domain containing protein 3a

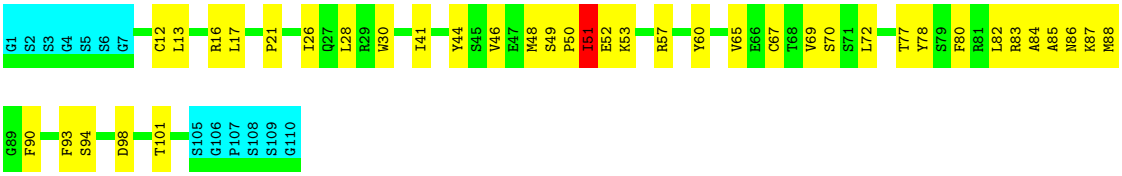
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Fibronectin type-III domain containing protein 3a

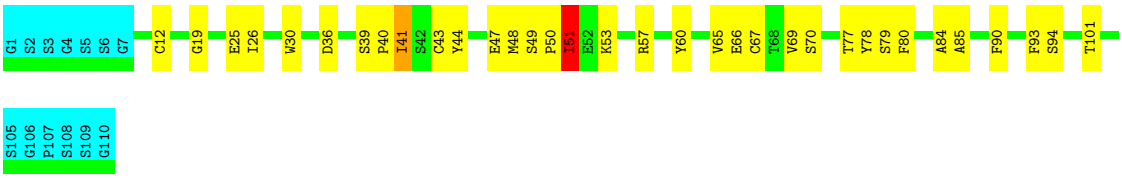
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 62% 25% • 12%

G1 S2 S3 G4 S5 S6 G7 C12 L17 Q18 G19 R20 P21 E25 I26 Q27 D36 S39 P40 I41 S42 C43 Y44 M48 S49 P50 I51 E52 K53 R57 T77 Y78 S79 F80 A84 A85 F90 S94 E95 K96 C97 D98 T101 S105 G106 P107 S108 S109 G110

4.2.19 Score per residue for model 19

- Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 49% 36% • 12%

G1 S2 S3 G4 S5 S6 G7 C12 L17 Q18 G19 R20 P21 K22 E25 I26 Q27 W30 L34 V35 D36 G37 G38 S39 P40 I41 Y44 E47 M48 S49 P50 I51 E52 K53 R57 Y60 V65 E66 C67 S70 T77 Y78 S79 F80 R81 L82 R83 A84 A85 N86 K87 M88 G89 F90 F93 S94 E95 K96 C97 D98 T101 S105 G106 P107 S108 S109 G110

4.2.20 Score per residue for model 20

- Molecule 1: Fibronectin type-III domain containing protein 3a

Chain A: 58% 27% • 12%

G1 S2 S3 G4 S5 S6 G7 C12 L17 Q18 G19 R20 P21 I26 Q27 D36 G37 G38 S39 P40 I41 Y44 E47 M48 S49 P50 I51 E52 K53 D54 R57 Y60 C67 T77 Y78 S79 F80 R81 L82 A85 N86 K87 M88 G89 F90 F93 S94 D98 T101 S105 G106 P107 S108 S109 G110

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function, structures with the lowest energy, structures with the least restraint violations*.

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

Software name	Classification	Version
CYANA	structure solution	2.0.17
CYANA	refinement	2.0.17

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	737	735	735	29±5
All	All	14740	14700	14700	584

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 20.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:26:ILE:HD11	1:A:101:THR:HG22	1.00	1.32	19	20
1:A:26:ILE:HD11	1:A:101:THR:CG2	0.92	1.93	9	20
1:A:41:ILE:HG21	1:A:44:TYR:CZ	0.80	2.12	6	20
1:A:17:LEU:HD21	1:A:21:PRO:HD3	0.80	1.54	11	16
1:A:85:ALA:HB2	1:A:90:PHE:CD1	0.72	2.19	17	11
1:A:41:ILE:HD12	1:A:84:ALA:HB1	0.71	1.60	9	16
1:A:26:ILE:CD1	1:A:101:THR:HG22	0.67	2.17	19	16
1:A:60:TYR:CD1	1:A:67:CYS:CB	0.64	2.81	8	8
1:A:85:ALA:HB2	1:A:90:PHE:CG	0.61	2.30	19	10
1:A:85:ALA:HB2	1:A:90:PHE:CD2	0.61	2.30	11	9
1:A:48:MET:SD	1:A:80:PHE:CE2	0.60	2.94	9	19
1:A:77:THR:O	1:A:78:TYR:CD1	0.58	2.56	17	15
1:A:17:LEU:HD12	1:A:27:GLN:O	0.58	1.99	7	8
1:A:44:TYR:CE1	1:A:84:ALA:HB2	0.57	2.34	10	12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:43:CYS:SG	1:A:85:ALA:HB3	0.57	2.38	17	1
1:A:69:VAL:HG21	1:A:80:PHE:HE2	0.56	1.60	2	2
1:A:26:ILE:HD11	1:A:101:THR:HG23	0.56	1.77	18	7
1:A:85:ALA:HB2	1:A:90:PHE:CE1	0.56	2.35	17	4
1:A:30:TRP:CD2	1:A:65:VAL:O	0.56	2.59	11	7
1:A:48:MET:CG	1:A:80:PHE:CE2	0.55	2.90	15	19
1:A:26:ILE:HG13	1:A:72:LEU:HD12	0.55	1.78	16	4
1:A:60:TYR:CD2	1:A:67:CYS:CB	0.55	2.89	1	1
1:A:17:LEU:HD13	1:A:26:ILE:CG2	0.54	2.33	18	5
1:A:60:TYR:CD1	1:A:67:CYS:HB2	0.54	2.38	11	7
1:A:60:TYR:CE1	1:A:67:CYS:CB	0.53	2.92	19	4
1:A:41:ILE:HG21	1:A:44:TYR:CE1	0.52	2.40	7	3
1:A:48:MET:HG3	1:A:80:PHE:CE2	0.52	2.40	12	18
1:A:41:ILE:HD12	1:A:84:ALA:CB	0.52	2.34	14	2
1:A:87:LYS:HG3	1:A:88:MET:HE3	0.51	1.81	5	1
1:A:60:TYR:CE1	1:A:67:CYS:HB3	0.51	2.41	8	3
1:A:44:TYR:CD1	1:A:84:ALA:HB2	0.51	2.40	16	1
1:A:69:VAL:HG21	1:A:80:PHE:CE2	0.51	2.41	9	2
1:A:53:LYS:O	1:A:55:GLU:N	0.50	2.44	12	1
1:A:66:GLU:N	1:A:66:GLU:OE1	0.50	2.45	11	1
1:A:41:ILE:HG21	1:A:44:TYR:CE2	0.50	2.42	6	5
1:A:30:TRP:CH2	1:A:46:VAL:HG21	0.50	2.42	16	1
1:A:30:TRP:O	1:A:65:VAL:CG2	0.49	2.60	9	6
1:A:50:PRO:O	1:A:51:ILE:C	0.49	2.50	9	18
1:A:50:PRO:O	1:A:53:LYS:N	0.49	2.45	15	18
1:A:12:CYS:O	1:A:94:SER:CB	0.49	2.60	17	14
1:A:58:GLU:OE1	1:A:61:GLN:NE2	0.49	2.45	14	3
1:A:47:GLU:HB2	1:A:93:PHE:CZ	0.49	2.42	12	8
1:A:48:MET:SD	1:A:80:PHE:CZ	0.49	3.06	3	6
1:A:86:ASN:O	1:A:88:MET:N	0.49	2.45	10	6
1:A:60:TYR:CD1	1:A:61:GLN:N	0.48	2.80	15	1
1:A:43:CYS:O	1:A:85:ALA:N	0.48	2.46	14	1
1:A:36:ASP:OD1	1:A:38:GLY:N	0.48	2.46	6	10
1:A:50:PRO:O	1:A:52:GLU:N	0.48	2.47	15	4
1:A:60:TYR:CG	1:A:67:CYS:CB	0.48	2.97	2	2
1:A:69:VAL:HG12	1:A:70:SER:N	0.48	2.23	16	3
1:A:60:TYR:CD2	1:A:67:CYS:HB3	0.47	2.44	2	2
1:A:81:ARG:HH11	1:A:96:LYS:N	0.47	2.07	19	1
1:A:13:LEU:O	1:A:82:LEU:HD23	0.47	2.08	16	2
1:A:69:VAL:HG12	1:A:72:LEU:HD21	0.47	1.86	3	2
1:A:77:THR:HG22	1:A:78:TYR:N	0.47	2.24	1	10

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:83:ARG:HG2	1:A:93:PHE:CE1	0.46	2.45	9	2
1:A:83:ARG:HD3	1:A:93:PHE:CE2	0.46	2.45	11	7
1:A:36:ASP:OD1	1:A:36:ASP:C	0.46	2.54	3	15
1:A:48:MET:HB2	1:A:80:PHE:CD2	0.46	2.45	2	6
1:A:60:TYR:CE2	1:A:67:CYS:HB3	0.46	2.45	1	1
1:A:36:ASP:CG	1:A:39:SER:O	0.46	2.55	13	8
1:A:40:PRO:O	1:A:41:ILE:C	0.46	2.54	10	12
1:A:46:VAL:HG13	1:A:81:ARG:O	0.46	2.10	7	2
1:A:25:GLU:O	1:A:26:ILE:CG1	0.45	2.64	19	3
1:A:17:LEU:HD21	1:A:21:PRO:CD	0.45	2.41	3	3
1:A:50:PRO:O	1:A:53:LYS:O	0.45	2.34	15	14
1:A:36:ASP:OD1	1:A:39:SER:N	0.45	2.49	10	2
1:A:39:SER:OG	1:A:87:LYS:CE	0.45	2.65	13	1
1:A:18:GLN:O	1:A:19:GLY:C	0.45	2.55	12	11
1:A:83:ARG:HB3	1:A:93:PHE:CD1	0.45	2.47	3	7
1:A:17:LEU:HD23	1:A:18:GLN:N	0.45	2.26	8	1
1:A:16:ARG:O	1:A:17:LEU:C	0.45	2.56	8	2
1:A:77:THR:C	1:A:78:TYR:CD1	0.44	2.91	17	4
1:A:82:LEU:O	1:A:94:SER:N	0.44	2.50	12	4
1:A:81:ARG:NH1	1:A:94:SER:O	0.44	2.50	7	2
1:A:83:ARG:HD3	1:A:93:PHE:CZ	0.44	2.48	11	2
1:A:39:SER:OG	1:A:86:ASN:ND2	0.44	2.51	20	2
1:A:85:ALA:HB2	1:A:90:PHE:CE2	0.44	2.48	14	2
1:A:30:TRP:CE2	1:A:65:VAL:HA	0.44	2.48	7	3
1:A:50:PRO:HB3	1:A:78:TYR:CE2	0.44	2.47	18	5
1:A:98:ASP:CG	1:A:98:ASP:O	0.44	2.56	20	1
1:A:83:ARG:CD	1:A:93:PHE:CZ	0.43	3.01	9	2
1:A:83:ARG:NE	1:A:93:PHE:CZ	0.43	2.86	2	1
1:A:60:TYR:CE1	1:A:67:CYS:HB2	0.43	2.48	19	1
1:A:51:ILE:HG21	1:A:77:THR:HB	0.42	1.90	12	2
1:A:57:ARG:O	1:A:59:VAL:HG13	0.42	2.13	12	1
1:A:86:ASN:C	1:A:88:MET:N	0.42	2.73	16	5
1:A:57:ARG:HD2	1:A:59:VAL:CG1	0.42	2.44	15	2
1:A:49:SER:O	1:A:79:SER:N	0.42	2.53	15	4
1:A:48:MET:HG3	1:A:80:PHE:CD2	0.42	2.50	4	3
1:A:47:GLU:CB	1:A:93:PHE:CZ	0.42	3.03	15	1
1:A:77:THR:CG2	1:A:78:TYR:N	0.42	2.82	1	2
1:A:86:ASN:O	1:A:87:LYS:C	0.42	2.57	19	1
1:A:69:VAL:CG1	1:A:72:LEU:HD21	0.42	2.44	3	1
1:A:81:ARG:NE	1:A:93:PHE:HB3	0.41	2.29	14	1
1:A:30:TRP:CD1	1:A:82:LEU:HD21	0.41	2.50	13	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:60:TYR:CG	1:A:67:CYS:HB3	0.41	2.50	16	2
1:A:41:ILE:HG22	1:A:42:SER:N	0.41	2.31	1	1
1:A:25:GLU:HG2	1:A:70:SER:CB	0.41	2.46	17	3
1:A:60:TYR:CD1	1:A:60:TYR:C	0.41	2.93	1	1
1:A:98:ASP:O	1:A:98:ASP:CG	0.41	2.59	5	1
1:A:88:MET:HA	1:A:88:MET:CE	0.41	2.46	8	1
1:A:58:GLU:OE2	1:A:61:GLN:CB	0.41	2.69	15	1
1:A:65:VAL:CG2	1:A:66:GLU:OE1	0.41	2.68	17	1
1:A:26:ILE:HG21	1:A:99:ILE:HG21	0.41	1.92	6	2
1:A:48:MET:HG2	1:A:49:SER:N	0.40	2.30	12	3
1:A:13:LEU:CD2	1:A:14:PRO:HD2	0.40	2.46	12	1
1:A:30:TRP:CD1	1:A:82:LEU:CD2	0.40	3.04	9	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	97/110 (88%)	85±2 (88±3%)	10±3 (11±3%)	2±1 (2±1%)	11	53
All	All	1940/2200 (88%)	1699 (88%)	204 (11%)	37 (2%)	11	53

All 7 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	41	ILE	12
1	A	51	ILE	11
1	A	87	LYS	5
1	A	19	GLY	4
1	A	17	LEU	2
1	A	96	LYS	2
1	A	54	ASP	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	83/91 (91%)	79±1 (95±1%)	4±1 (5±1%)	31	79
All	All	1660/1820 (91%)	1583 (95%)	77 (5%)	31	79

All 14 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	51	ILE	20
1	A	57	ARG	20
1	A	98	ASP	14
1	A	39	SER	7
1	A	55	GLU	3
1	A	43	CYS	3
1	A	96	LYS	2
1	A	53	LYS	2
1	A	66	GLU	1
1	A	16	ARG	1
1	A	28	LEU	1
1	A	22	LYS	1
1	A	34	LEU	1
1	A	54	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