Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 05:19 PM EDT

PDB ID	:	2KP5
Title	:	NMR structure of Hahellin, a beta-gamma crystallin
Authors	:	Srivastava, A.K.; Sharma, Y.; Chary, K.V.
Deposited on	:	2009-10-07

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

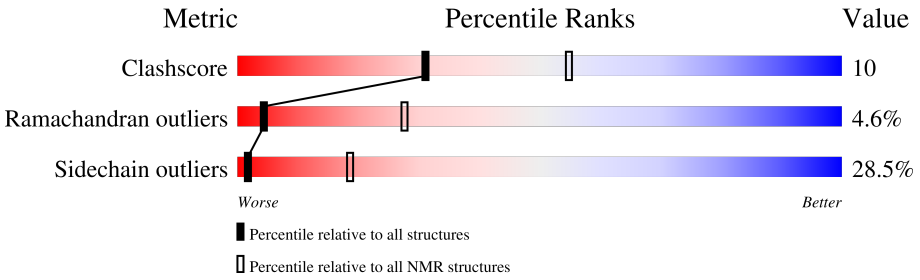
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	92	50% 41% • 7% •

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:3-A:87 (85)	0.78	2

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 3 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1396 atoms, of which 684 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	91	1396	439	684	125	147	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q2SHN6

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: Putative uncharacterized protein

Chain A: 50% 41% 7%

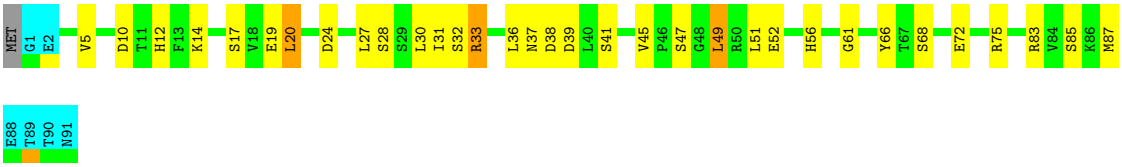
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Putative uncharacterized protein

Chain A: 57% 33% 7%



4.2.2 Score per residue for model 2 (medoid)

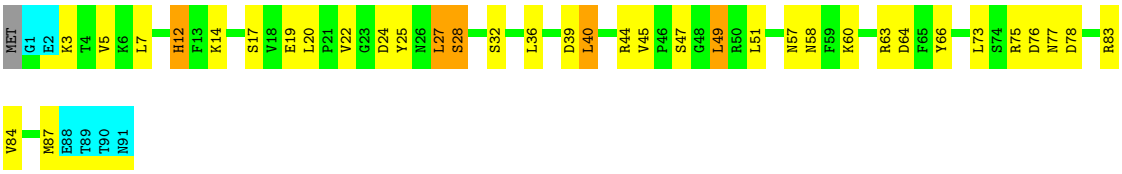
• Molecule 1: Putative uncharacterized protein

Chain A: 55% 36% 7%

4.2.3 Score per residue for model 3

- Molecule 1: Putative uncharacterized protein

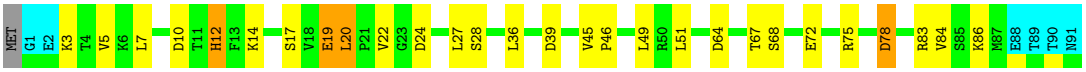
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Putative uncharacterized protein

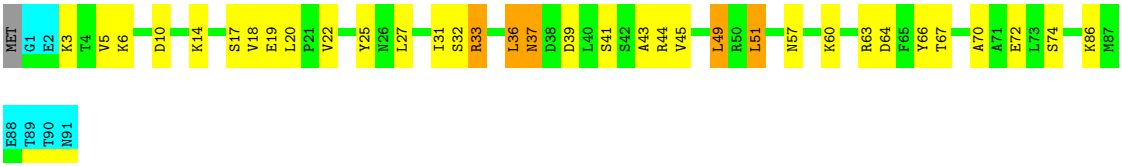
Chain A:



4.2.5 Score per residue for model 5

- Molecule 1: Putative uncharacterized protein

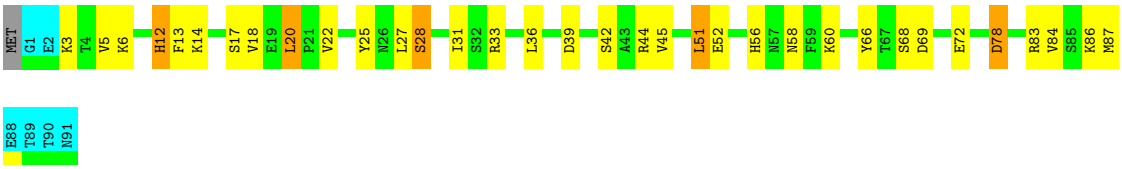
Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: Putative uncharacterized protein

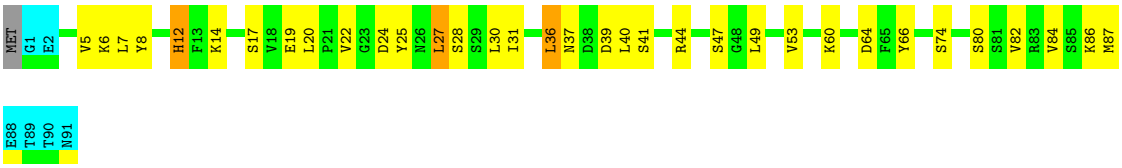
Chain A:



4.2.7 Score per residue for model 7

- Molecule 1: Putative uncharacterized protein

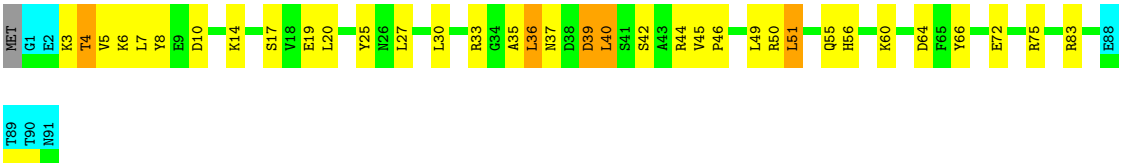
Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Putative uncharacterized protein

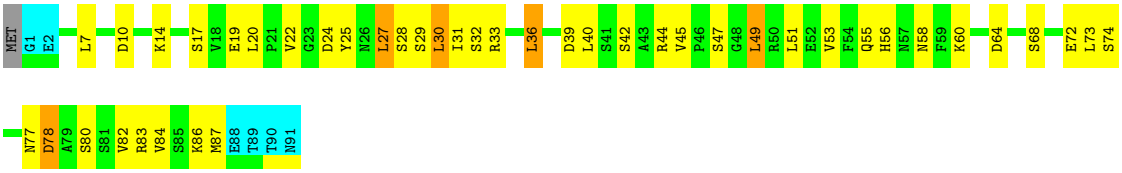
Chain A:



4.2.9 Score per residue for model 9

- Molecule 1: Putative uncharacterized protein

Chain A:



4.2.10 Score per residue for model 10

- Molecule 1: Putative uncharacterized protein

Chain A: 51% 32% 10% 7%

MET G1 E2 K3 T4 V5 K6 L7 Y8 H12 F13 K14 S17 V18 E19 L20 P21 V22 G23 D24 L27 L30 I31 S32 R33 L36 N37 D38 D39 L40 R44 L49 R50 L51 H56 K60 Y66 T67 S68 E72 L73 S74 R75 D76 N77 D78 R83 V84 S85 K86 M87 E88 T89 T90 N91

4.2.11 Score per residue for model 11

- Molecule 1: Putative uncharacterized protein

Chain A: 62% 24% 7% 7%

MET G1 E2 K3 T4 V5 K6 L7 H12 F13 K14 S17 V18 E19 L20 D24 Y25 N26 L27 S28 L36 N37 D38 D39 L40 S41 V45 L51 K60 Y66 E72 R75 D78 R83 K86 M87 E88 T89 T90 N91

4.2.12 Score per residue for model 12

- Molecule 1: Putative uncharacterized protein

Chain A: 49% 35% 9% 7%

MET G1 E2 K3 T4 V5 K6 L7 Y8 E9 D10 T11 H12 F13 K14 S17 V18 E19 L20 P21 V22 G23 D24 L27 S28 S29 L30 I31 S32 A35 L36 N37 D38 D39 L40 R44 V45 R50 L51 H56 N57 N58 D64 F65 Y66 T67 E72 R75 D76 N77 D78 V84 S85 K86 M87 E88 T89 T90 N91

4.2.13 Score per residue for model 13

- Molecule 1: Putative uncharacterized protein

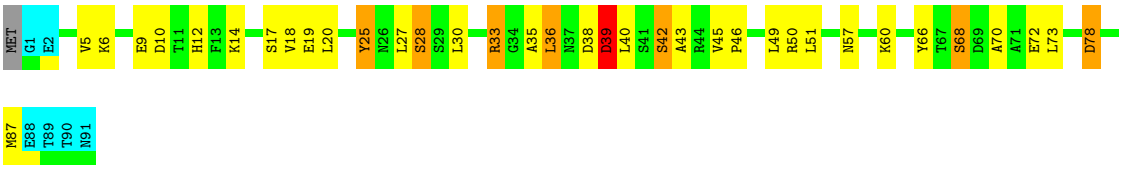
Chain A: 53% 32% 8% 7%

MET G1 E2 K3 T4 V5 E9 H12 Y16 S17 L20 P21 V22 G23 D24 Y25 N26 L27 S28 S29 L30 I31 S32 R33 G34 A35 L36 N37 D38 D39 S42 V45 P46 S47 G48 L49 R50 K60 D64 T67 S68 E72 L73 S74 R75 D76 S85 E88 T89 T90 N91

4.2.14 Score per residue for model 14

- Molecule 1: Putative uncharacterized protein

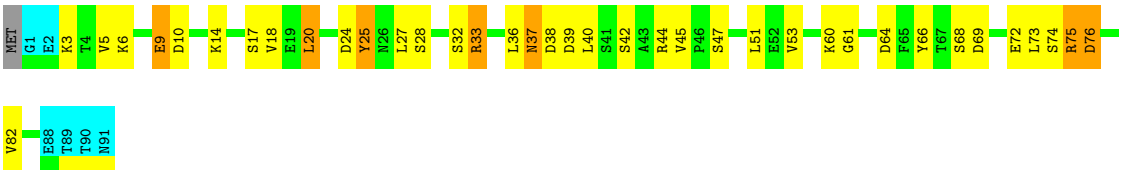
Chain A: 53% 30% 8% 7%



4.2.15 Score per residue for model 15

- Molecule 1: Putative uncharacterized protein

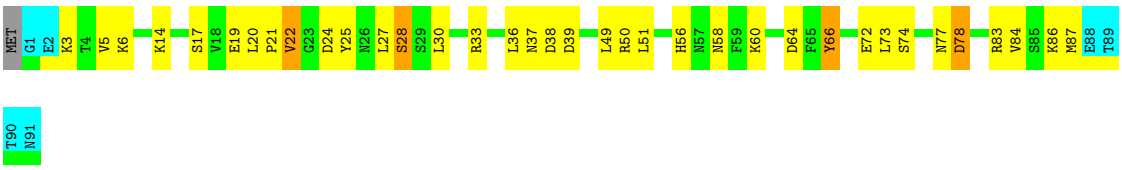
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Putative uncharacterized protein

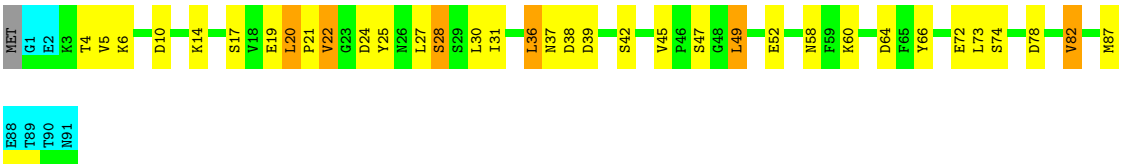
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Putative uncharacterized protein

Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Putative uncharacterized protein

Chain A: 46% 41% 5% 7% •

MET G1 E2 K3 T4 V5 K6 L7 Y8 E9 D10 K14 S17 V18 E19 L20 P21 V22 G23 D24 L27 S28 S29 L30 I31 S32 R33 L36 N37 D38 D39 L40 S41 S42 A43 R44 V45 P46 S47 L51 H56 N57 N58 F59 K60 Y66 T67 S68 E72 L73 D78

A79 R83 V84 S85 K86 M87 E88 T89 T90 N91

4.2.19 Score per residue for model 19

- Molecule 1: Putative uncharacterized protein

Chain A: 52% 36% • 7% •

MET G1 E2 K3 T4 V5 K6 L7 Y8 E9 D10 K14 S17 L20 P21 V22 G23 D24 Y25 N26 L27 S28 I31 S32 R33 D38 D39 L40 R44 V45 P46 S47 G48 L49 R50 L51 N58 F59 K60 D64 F65 Y66 L73 S74 R75 D78 A79 R83 V84

M87 E88 T89 T90 N91

4.2.20 Score per residue for model 20

- Molecule 1: Putative uncharacterized protein

Chain A: 53% 33% 7% 7% •

MET G1 E2 K3 K6 L7 D10 K14 S17 V18 E19 L20 D24 L27 S28 S29 L30 I31 L36 N37 D38 D39 S42 A43 R44 S47 G48 L49 R50 L51 H56 K60 R63 D64 F65 Y66 E72 L73 S74 R75 D76 N77 D78 V84 S85 K86

M87 E88 T89 T90 N91

5 Refinement protocol and experimental data overview (i)

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	refinement	

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	668	649	649	13±2
All	All	13360	12980	12980	256

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 10.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:5:VAL:HG13	1:A:49:LEU:HD12	0.98	1.33	1	1
1:A:5:VAL:HG23	1:A:45:VAL:HG12	0.81	1.52	2	8
1:A:51:LEU:HD23	1:A:84:VAL:HG12	0.80	1.52	3	3
1:A:20:LEU:HD12	1:A:84:VAL:HG11	0.79	1.51	20	5
1:A:5:VAL:HG13	1:A:49:LEU:HD13	0.79	1.53	10	2
1:A:49:LEU:HD21	1:A:51:LEU:HD23	0.77	1.56	2	1
1:A:45:VAL:HG11	1:A:51:LEU:HD11	0.76	1.58	12	1
1:A:45:VAL:HG11	1:A:51:LEU:HD21	0.75	1.58	5	4
1:A:49:LEU:HD21	1:A:84:VAL:HG23	0.66	1.68	16	1
1:A:7:LEU:HD11	1:A:20:LEU:HD11	0.66	1.66	3	3
1:A:49:LEU:HD13	1:A:85:SER:O	0.64	1.92	1	1
1:A:7:LEU:HD22	1:A:40:LEU:HD23	0.64	1.69	11	1
1:A:36:LEU:HD12	1:A:39:ASP:HB2	0.64	1.69	11	1
1:A:45:VAL:HG11	1:A:51:LEU:CD1	0.64	2.23	12	3

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:19:GLU:O	1:A:20:LEU:HD22	0.64	1.92	16	2
1:A:73:LEU:HD13	1:A:74:SER:N	0.63	2.08	15	6
1:A:5:VAL:HG12	1:A:46:PRO:HD2	0.63	1.69	4	1
1:A:20:LEU:HD22	1:A:33:ARG:CD	0.63	2.23	1	1
1:A:31:ILE:HD12	1:A:36:LEU:HD22	0.63	1.71	18	1
1:A:73:LEU:HD12	1:A:76:ASP:O	0.62	1.94	15	1
1:A:19:GLU:O	1:A:20:LEU:HD23	0.62	1.94	9	4
1:A:7:LEU:HD23	1:A:43:ALA:HB2	0.62	1.70	18	1
1:A:49:LEU:CD2	1:A:51:LEU:HD23	0.61	2.25	2	1
1:A:7:LEU:HD23	1:A:40:LEU:CD2	0.61	2.26	9	1
1:A:20:LEU:HD22	1:A:25:TYR:CZ	0.61	2.30	9	1
1:A:20:LEU:CB	1:A:84:VAL:HG11	0.61	2.25	9	1
1:A:30:LEU:HD13	1:A:37:ASN:HA	0.61	1.73	1	2
1:A:53:VAL:HA	1:A:82:VAL:HG13	0.61	1.72	7	2
1:A:45:VAL:HB	1:A:49:LEU:HD22	0.61	1.73	2	1
1:A:5:VAL:CG2	1:A:45:VAL:HG12	0.60	2.24	12	8
1:A:27:LEU:HA	1:A:30:LEU:HD12	0.60	1.73	17	4
1:A:7:LEU:HD22	1:A:40:LEU:HG	0.60	1.73	18	1
1:A:46:PRO:HD2	1:A:49:LEU:HD13	0.60	1.74	2	3
1:A:36:LEU:HD22	1:A:39:ASP:CG	0.60	2.16	8	1
1:A:20:LEU:HD22	1:A:33:ARG:HD3	0.59	1.74	1	4
1:A:27:LEU:O	1:A:31:ILE:HD12	0.59	1.98	17	10
1:A:20:LEU:O	1:A:84:VAL:HG21	0.58	1.98	12	8
1:A:45:VAL:HG11	1:A:51:LEU:CD2	0.58	2.28	5	2
1:A:45:VAL:HG11	1:A:51:LEU:HD13	0.58	1.74	3	3
1:A:5:VAL:HG13	1:A:45:VAL:CG1	0.57	2.28	18	1
1:A:7:LEU:HD23	1:A:40:LEU:HG	0.57	1.76	10	2
1:A:30:LEU:HD22	1:A:36:LEU:O	0.57	1.99	12	2
1:A:7:LEU:HD22	1:A:40:LEU:HD21	0.57	1.74	7	1
1:A:49:LEU:HD23	1:A:50:ARG:N	0.57	2.14	19	5
1:A:20:LEU:HB2	1:A:84:VAL:HG11	0.57	1.77	9	1
1:A:20:LEU:HD11	1:A:25:TYR:CE1	0.57	2.34	15	1
1:A:7:LEU:HD23	1:A:40:LEU:HD21	0.57	1.76	9	1
1:A:51:LEU:HD23	1:A:84:VAL:CG1	0.56	2.27	3	3
1:A:36:LEU:HD12	1:A:37:ASN:O	0.56	2.00	15	1
1:A:19:GLU:C	1:A:20:LEU:HD23	0.56	2.21	17	8
1:A:45:VAL:HG11	1:A:51:LEU:HD23	0.56	1.78	14	1
1:A:18:VAL:HG12	1:A:20:LEU:CD2	0.55	2.31	6	1
1:A:49:LEU:O	1:A:67:THR:HG23	0.55	2.01	4	3
1:A:27:LEU:HD12	1:A:37:ASN:CG	0.55	2.22	17	1
1:A:31:ILE:HD11	1:A:37:ASN:ND2	0.55	2.16	20	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:45:VAL:HG21	1:A:51:LEU:HD23	0.55	1.79	11	2
1:A:36:LEU:O	1:A:36:LEU:HD13	0.54	2.02	11	1
1:A:36:LEU:HD13	1:A:39:ASP:HB3	0.54	1.78	14	1
1:A:7:LEU:HD23	1:A:43:ALA:HB1	0.54	1.78	2	1
1:A:45:VAL:HG23	1:A:49:LEU:HB3	0.54	1.80	1	1
1:A:7:LEU:HB3	1:A:40:LEU:HD21	0.54	1.79	8	1
1:A:49:LEU:HD23	1:A:85:SER:O	0.54	2.02	10	1
1:A:31:ILE:HD11	1:A:37:ASN:HD21	0.53	1.64	20	1
1:A:36:LEU:HD13	1:A:37:ASN:N	0.53	2.18	5	3
1:A:36:LEU:HD12	1:A:39:ASP:CB	0.53	2.32	11	1
1:A:5:VAL:HG21	1:A:51:LEU:HD22	0.52	1.81	11	1
1:A:5:VAL:HG11	1:A:51:LEU:HD23	0.52	1.80	10	2
1:A:27:LEU:HD23	1:A:28:SER:N	0.52	2.19	19	11
1:A:31:ILE:HG12	1:A:36:LEU:HD22	0.52	1.81	9	1
1:A:7:LEU:HG	1:A:43:ALA:HB2	0.52	1.81	20	1
1:A:52:GLU:O	1:A:82:VAL:HG12	0.52	2.05	17	1
1:A:20:LEU:HD12	1:A:84:VAL:CG1	0.51	2.32	20	2
1:A:45:VAL:HG22	1:A:68:SER:O	0.51	2.06	14	1
1:A:21:PRO:O	1:A:22:VAL:C	0.51	2.49	16	5
1:A:25:TYR:N	1:A:25:TYR:CD1	0.51	2.79	14	8
1:A:45:VAL:HG23	1:A:49:LEU:HD22	0.50	1.82	19	1
1:A:5:VAL:CG1	1:A:49:LEU:HD13	0.50	2.31	10	2
1:A:73:LEU:HD12	1:A:78:ASP:OD1	0.50	2.07	19	1
1:A:27:LEU:HD12	1:A:37:ASN:OD1	0.50	2.06	17	1
1:A:7:LEU:HD23	1:A:40:LEU:CD1	0.50	2.37	12	1
1:A:51:LEU:HG	1:A:84:VAL:HG23	0.49	1.82	9	1
1:A:18:VAL:HG11	1:A:33:ARG:CB	0.49	2.38	14	1
1:A:35:ALA:O	1:A:36:LEU:HD22	0.49	2.08	13	1
1:A:40:LEU:CB	1:A:73:LEU:HD11	0.49	2.38	14	1
1:A:7:LEU:HD12	1:A:20:LEU:CD1	0.49	2.38	18	1
1:A:7:LEU:HD13	1:A:20:LEU:HD21	0.48	1.85	20	1
1:A:36:LEU:HD22	1:A:39:ASP:CB	0.48	2.38	8	1
1:A:5:VAL:HB	1:A:51:LEU:HD21	0.48	1.85	3	1
1:A:49:LEU:HD22	1:A:51:LEU:HD11	0.48	1.85	3	1
1:A:36:LEU:HD12	1:A:40:LEU:HD12	0.48	1.85	8	1
1:A:18:VAL:HG13	1:A:33:ARG:HG2	0.47	1.86	5	1
1:A:30:LEU:HD13	1:A:36:LEU:C	0.47	2.30	13	1
1:A:27:LEU:HD12	1:A:37:ASN:ND2	0.47	2.25	11	2
1:A:5:VAL:HG13	1:A:45:VAL:HG12	0.47	1.87	18	1
1:A:49:LEU:HD21	1:A:84:VAL:CG2	0.47	2.39	9	1
1:A:30:LEU:O	1:A:35:ALA:HB3	0.47	2.09	8	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:8:TYR:O	1:A:40:LEU:HD13	0.47	2.09	12	1
1:A:9:GLU:HA	1:A:40:LEU:HD13	0.47	1.87	14	2
1:A:30:LEU:HD13	1:A:36:LEU:O	0.46	2.10	13	1
1:A:5:VAL:HG12	1:A:45:VAL:HB	0.46	1.85	13	1
1:A:4:THR:HG22	1:A:21:PRO:HA	0.46	1.87	19	3
1:A:40:LEU:O	1:A:73:LEU:HD12	0.46	2.10	3	1
1:A:53:VAL:HA	1:A:82:VAL:HG12	0.46	1.86	9	1
1:A:22:VAL:HA	1:A:84:VAL:HG23	0.46	1.88	19	2
1:A:7:LEU:HD12	1:A:20:LEU:HD12	0.45	1.88	18	1
1:A:3:LYS:O	1:A:4:THR:HG23	0.45	2.12	8	2
1:A:36:LEU:HD22	1:A:39:ASP:HB3	0.45	1.89	8	1
1:A:20:LEU:HD12	1:A:25:TYR:CD1	0.45	2.46	7	1
1:A:43:ALA:O	1:A:70:ALA:HB3	0.44	2.12	14	2
1:A:18:VAL:HG12	1:A:20:LEU:HD23	0.44	1.90	15	1
1:A:25:TYR:CD1	1:A:25:TYR:N	0.44	2.86	16	4
1:A:36:LEU:HD12	1:A:38:ASP:OD1	0.44	2.13	18	2
1:A:8:TYR:CD2	1:A:10:ASP:O	0.43	2.71	18	1
1:A:78:ASP:O	1:A:79:ALA:HB2	0.43	2.14	19	1
1:A:5:VAL:HG12	1:A:20:LEU:O	0.43	2.14	16	1
1:A:45:VAL:CB	1:A:49:LEU:HD22	0.42	2.43	2	1
1:A:41:SER:C	1:A:73:LEU:HD23	0.42	2.35	18	1
1:A:5:VAL:HG13	1:A:20:LEU:HG	0.42	1.89	2	1
1:A:7:LEU:HD23	1:A:43:ALA:CB	0.42	2.43	2	1
1:A:30:LEU:HD13	1:A:37:ASN:HB3	0.42	1.91	12	1
1:A:20:LEU:HD12	1:A:25:TYR:CE1	0.42	2.49	19	1
1:A:40:LEU:HB3	1:A:73:LEU:HD23	0.42	1.91	15	1
1:A:37:ASN:O	1:A:79:ALA:HB3	0.42	2.15	18	1
1:A:7:LEU:HD13	1:A:40:LEU:HD21	0.41	1.92	2	1
1:A:5:VAL:HG12	1:A:45:VAL:HG23	0.41	1.92	17	1
1:A:5:VAL:CG1	1:A:49:LEU:HD22	0.41	2.45	7	1
1:A:30:LEU:HB3	1:A:35:ALA:HB3	0.41	1.92	14	1
1:A:7:LEU:CD1	1:A:20:LEU:HD11	0.41	2.43	4	1
1:A:18:VAL:HG21	1:A:35:ALA:HB2	0.41	1.91	12	1
1:A:7:LEU:HD23	1:A:40:LEU:CG	0.41	2.46	12	1
1:A:36:LEU:HD13	1:A:37:ASN:H	0.41	1.74	7	1
1:A:30:LEU:HD12	1:A:80:SER:O	0.40	2.17	9	1
1:A:22:VAL:HG23	1:A:85:SER:OG	0.40	2.16	13	1
1:A:51:LEU:HD11	1:A:66:TYR:CB	0.40	2.46	16	1
1:A:19:GLU:C	1:A:20:LEU:HD22	0.40	2.37	7	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	85/92 (92%)	69±3 (81±3%)	12±3 (14±3%)	4±1 (5±2%)	4 27
All	All	1700/1840 (92%)	1380 (81%)	241 (14%)	79 (5%)	4 27

All 19 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	39	ASP	18
1	A	78	ASP	12
1	A	22	VAL	9
1	A	12	HIS	8
1	A	77	ASN	5
1	A	74	SER	4
1	A	13	PHE	3
1	A	42	SER	3
1	A	38	ASP	3
1	A	61	GLY	2
1	A	69	ASP	2
1	A	57	ASN	2
1	A	67	THR	2
1	A	79	ALA	1
1	A	58	ASN	1
1	A	80	SER	1
1	A	37	ASN	1
1	A	75	ARG	1
1	A	40	LEU	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	75/81 (93%)	54±3 (71±4%)	21±3 (29±4%)	2 18
All	All	1500/1620 (93%)	1072 (71%)	428 (29%)	2 18

All 55 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	17	SER	20
1	A	14	LYS	19
1	A	66	TYR	17
1	A	72	GLU	17
1	A	60	LYS	17
1	A	24	ASP	16
1	A	36	LEU	15
1	A	20	LEU	14
1	A	64	ASP	14
1	A	10	ASP	13
1	A	87	MET	13
1	A	3	LYS	13
1	A	28	SER	12
1	A	75	ARG	11
1	A	83	ARG	11
1	A	27	LEU	11
1	A	44	ARG	11
1	A	86	LYS	11
1	A	6	LYS	11
1	A	12	HIS	10
1	A	33	ARG	10
1	A	47	SER	10
1	A	78	ASP	10
1	A	32	SER	9
1	A	56	HIS	9
1	A	68	SER	9
1	A	42	SER	8
1	A	51	LEU	8
1	A	58	ASN	7
1	A	49	LEU	6
1	A	76	ASP	6
1	A	19	GLU	6
1	A	41	SER	5
1	A	37	ASN	4
1	A	8	TYR	4
1	A	30	LEU	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	50	ARG	4
1	A	38	ASP	3
1	A	55	GLN	3
1	A	40	LEU	3
1	A	63	ARG	3
1	A	52	GLU	2
1	A	4	THR	2
1	A	29	SER	2
1	A	39	ASP	2
1	A	9	GLU	2
1	A	73	LEU	2
1	A	25	TYR	2
1	A	74	SER	1
1	A	13	PHE	1
1	A	7	LEU	1
1	A	16	TYR	1
1	A	57	ASN	1
1	A	69	ASP	1
1	A	82	VAL	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