



Full wwPDB NMR Structure Validation Report ⓘ

Aug 27, 2020 – 02:06 PM BST

PDB ID : 6VEE
Title : Solution structure of the TTD and linker region of mouse UHRF1 (NP95)
Authors : Lemak, A.; Houliston, S.; Duan, S.; Arrowsmith, C.H.
Deposited on : 2019-12-31

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 ⓘ symbol.

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

Cyrange : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
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 : 2.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.13

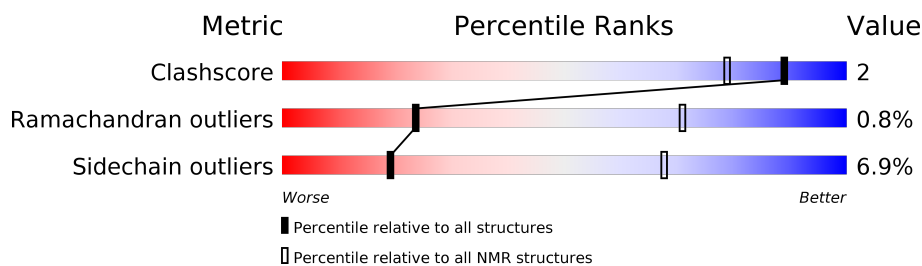
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 74%.

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	187	

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 18 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:128-A:159, A:172-A:226, A:231-A:300 (157)	0.93	18

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

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

NmrClust was unable to cluster the ensemble.

Error message: Inconsistent models

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2976 atoms, of which 1476 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called E3 ubiquitin-protein ligase UHRF1.

Mol	Chain	Residues	Atoms						Trace
1	A	184	Total	C	H	N	O	S	0
			2976	944	1476	268	281	7	

There are 3 discrepancies between the modelled and reference sequences:


Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	expression tag	UNP Q8VDF2
A	120	HIS	-	expression tag	UNP Q8VDF2
A	121	MET	-	expression tag	UNP Q8VDF2

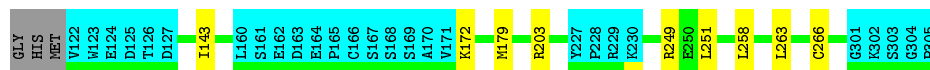
4 Residue-property plots

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: E3 ubiquitin-protein ligase UHRF1

Chain A: 



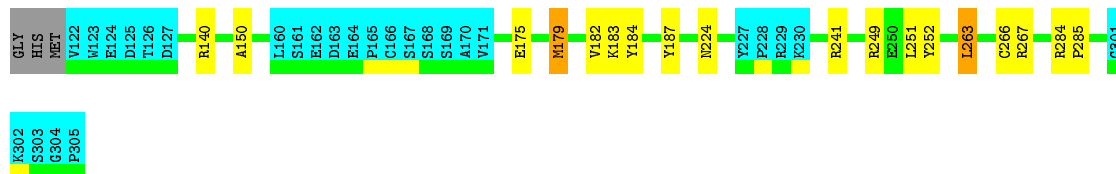
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A: 



4.2.2 Score per residue for model 2

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

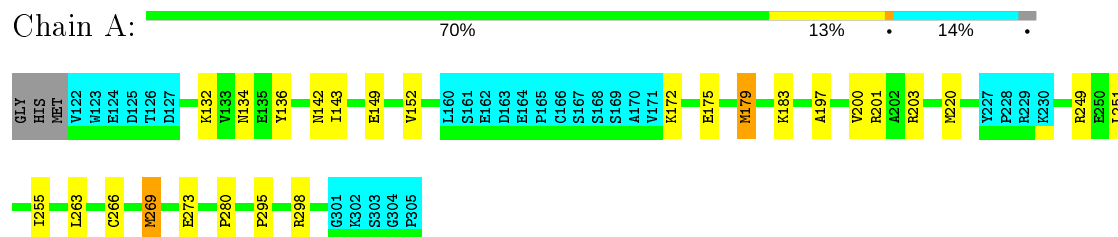
Chain A: 





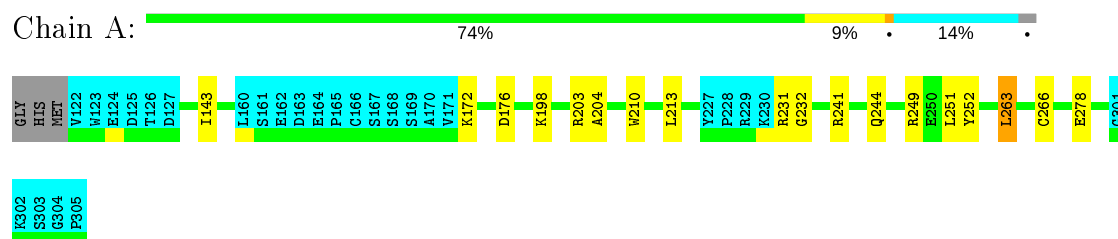
4.2.3 Score per residue for model 3

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



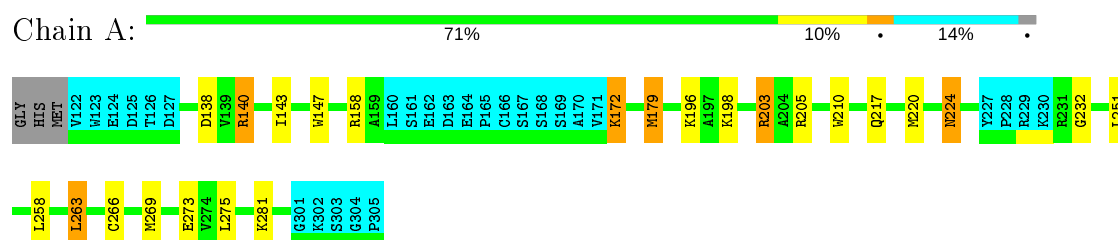
4.2.4 Score per residue for model 4

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



4.2.5 Score per residue for model 5

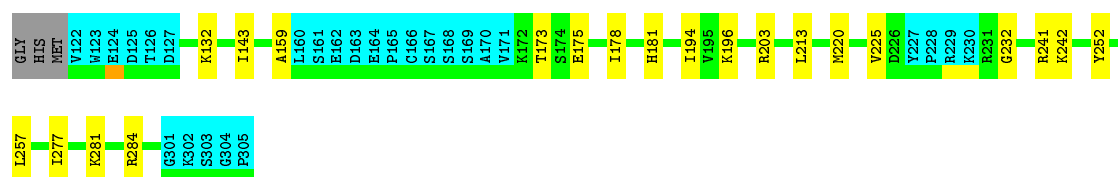
- Molecule 1: E3 ubiquitin-protein ligase UHRF1



4.2.6 Score per residue for model 6

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

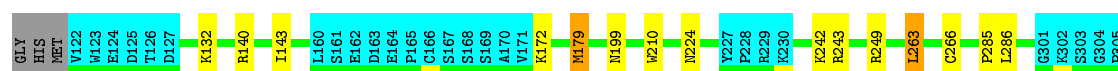




4.2.7 Score per residue for model 7

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A: 76% 7% 14%



4.2.8 Score per residue for model 8

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

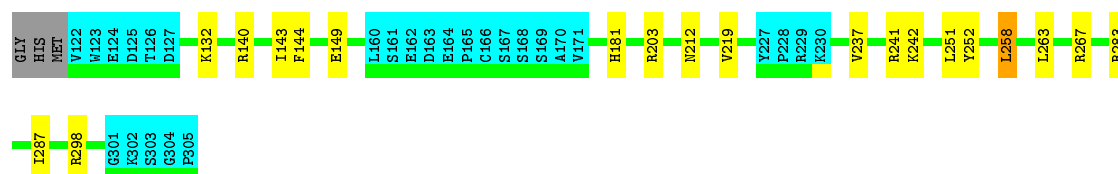
Chain A: 76% 7% 14%



4.2.9 Score per residue for model 9

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

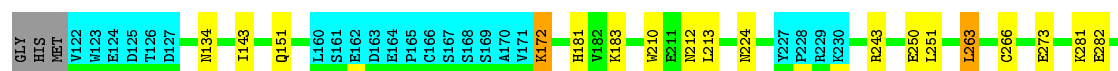
Chain A: 73% 10% 14%



4.2.10 Score per residue for model 10

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A: 74% 9% 14%



G301
K302
S303
G304
P305

4.2.11 Score per residue for model 11


- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A:  77% 6% 14%

GLY HIS MET V122 M123 E124 D125 T126 D127 I143 K157 L160 S161 E162 D163 E164 P165 C166 S167 S168 A170 V171 E175 M179 Y180 H181 Y182 K183 K196 Y227 P228 R229 R230 L251 L258 L263 C266 M269 S291 G301 K302 S303 G304 P305

4.2.12 Score per residue for model 12

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A:  76% 7% 14%

GLY HIS MET V122 M123 E124 D125 T126 D127 D138 L160 S161 E162 D163 E164 P165 C166 S167 S168 A170 V171 K172 H181 I194 R210 L213 N222 Y227 P228 R229 R230 R242 R249 E250 L251 L258 K281 E282 R283 R284 G301 K302 S303 G304 P305

4.2.13 Score per residue for model 13

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

Chain A:  73% 10% 14%

GLY HIS MET V122 M123 E124 D125 T126 D127 D138 I143 W147 K157 L160 S161 E162 D163 E164 P165 C166 S167 S168 S169 A170 V171 Y184 Y187 A202 R203 L213 Y227 P228 R229 R230 R231 D236 L239 R242 R249 R256 L257 L258 L263 C266 E282 L286 G301 K302 S303 G304 P305

C266 E282 L286 G301 K302 S303 G304 P305

4.2.14 Score per residue for model 14

- Molecule 1: E3 ubiquitin-protein ligase UHRF1

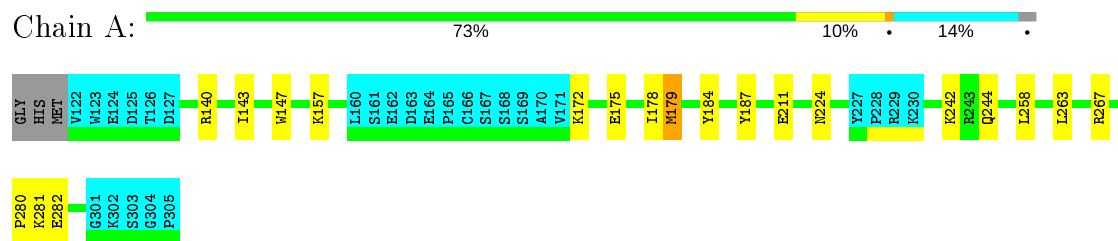
Chain A:  75% 8% 14%

GLY HIS MET V122 M123 E124 D125 T126 D127 I143 W147 K157 L160 S161 E162 D163 E164 P165 C166 S167 S168 S169 A170 V171 K172 M179 A204 N220 Y227 P228 R229 R230 R242 R249 E250 L251 L257 L258 S261 Q262 L263 M269 E282 G301 K302

S303 G304 P305

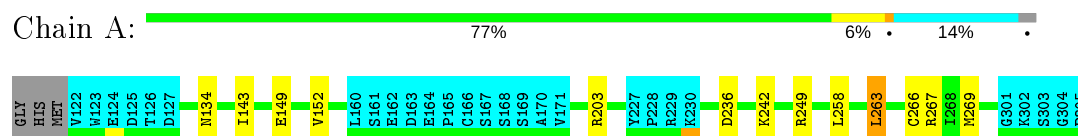
4.2.15 Score per residue for model 15

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



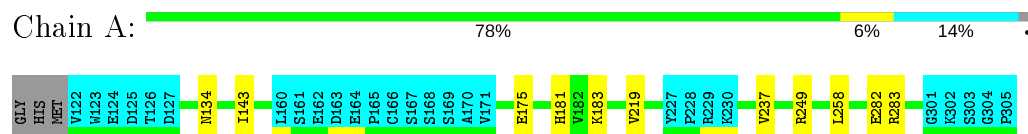
4.2.16 Score per residue for model 16

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



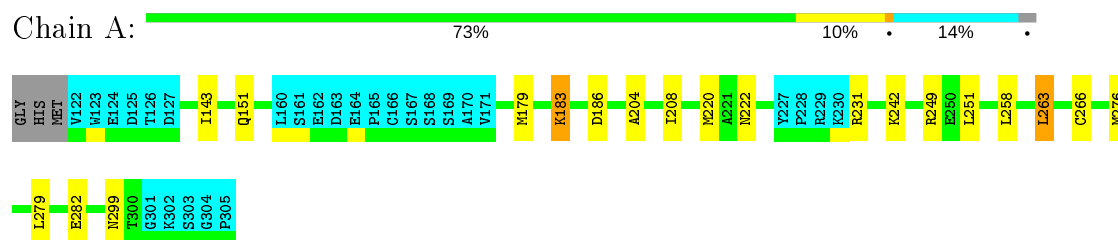
4.2.17 Score per residue for model 17

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



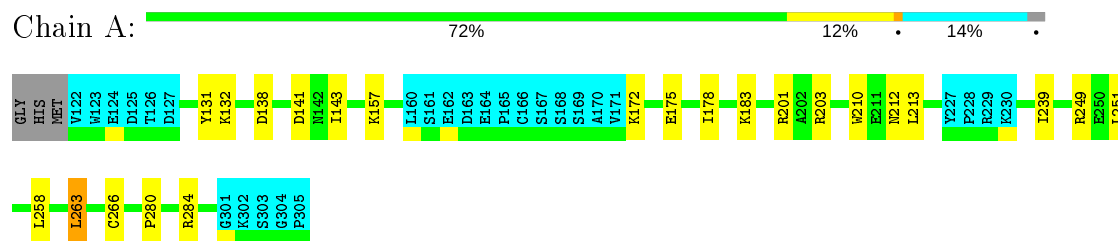
4.2.18 Score per residue for model 18 (medoid)

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



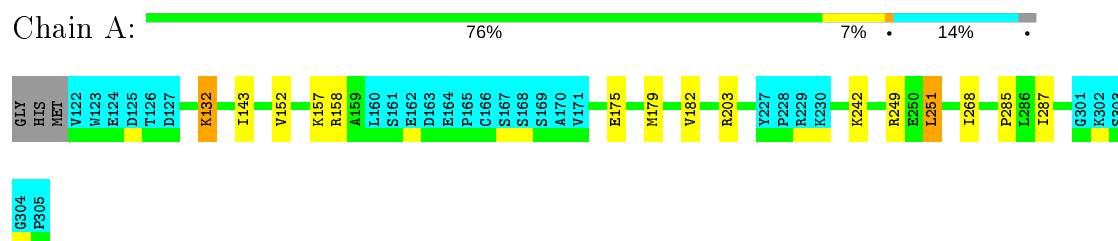
4.2.19 Score per residue for model 19

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



4.2.20 Score per residue for model 20

- Molecule 1: E3 ubiquitin-protein ligase UHRF1



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CNS	refinement	
CYANA	structure calculation	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1963
Number of shifts mapped to atoms	1963
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	74%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

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

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	1296	1290	1290	5±2
All	All	25920	25800	25800	95

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:255:ILE:HG13	1:A:266:CYS:SG	0.70	2.27	3	1
1:A:241:ARG:HB2	1:A:252:TYR:HB2	0.57	1.76	9	3
1:A:269:MET:HB3	1:A:273:GLU:HG3	0.56	1.76	5	2
1:A:204:ALA:HA	1:A:278:GLU:HB2	0.53	1.80	4	1
1:A:242:LYS:HD2	1:A:244:GLN:HE22	0.53	1.63	15	1
1:A:138:ASP:HB2	1:A:203:ARG:HA	0.53	1.80	5	2
1:A:157:LYS:HA	1:A:178:ILE:HG23	0.53	1.80	2	3
1:A:263:LEU:HD12	1:A:266:CYS:SG	0.52	2.45	4	1
1:A:172:LYS:HD3	1:A:172:LYS:H	0.51	1.66	10	1
1:A:210:TRP:HA	1:A:213:LEU:HD23	0.50	1.82	10	1
1:A:179:MET:SD	1:A:179:MET:N	0.50	2.85	15	5
1:A:210:TRP:HE1	1:A:251:LEU:HD22	0.50	1.67	5	1
1:A:149:GLU:HB3	1:A:203:ARG:NH2	0.50	2.22	2	1
1:A:263:LEU:HD12	1:A:266:CYS:HB2	0.50	1.84	10	9

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:152:VAL:HG12	1:A:182:VAL:HG12	0.49	1.84	20	1
1:A:140:ARG:HA	1:A:147:TRP:HA	0.49	1.83	15	2
1:A:222:ASN:ND2	1:A:231:ARG:HB2	0.48	2.23	18	1
1:A:213:LEU:HD21	1:A:239:ILE:HD13	0.48	1.84	2	2
1:A:147:TRP:HE1	1:A:236:ASP:HB2	0.48	1.69	13	1
1:A:220:MET:HB2	1:A:275:LEU:HB2	0.47	1.84	5	1
1:A:197:ALA:HA	1:A:200:VAL:HG12	0.47	1.87	3	1
1:A:144:PHE:HA	1:A:258:LEU:HD11	0.47	1.86	9	1
1:A:243:ARG:HB2	1:A:250:GLU:HB3	0.46	1.87	10	1
1:A:251:LEU:HD12	1:A:268:ILE:HG13	0.46	1.86	20	1
1:A:149:GLU:HB2	1:A:298:ARG:HG2	0.45	1.88	3	2
1:A:179:MET:N	1:A:179:MET:SD	0.45	2.89	18	3
1:A:251:LEU:HB3	1:A:268:ILE:HB	0.45	1.87	20	1
1:A:210:TRP:NE1	1:A:244:GLN:HB3	0.45	2.26	4	1
1:A:224:ASN:HA	1:A:232:GLY:HA3	0.45	1.89	5	1
1:A:220:MET:SD	1:A:277:ILE:HD13	0.45	2.51	6	1
1:A:219:VAL:HG23	1:A:237:VAL:HG13	0.45	1.87	9	2
1:A:279:LEU:HD22	1:A:282:GLU:H	0.45	1.71	18	1
1:A:131:TYR:CE2	1:A:280:PRO:HB3	0.45	2.46	19	1
1:A:241:ARG:HB3	1:A:252:TYR:HB2	0.44	1.89	2	1
1:A:243:ARG:HB3	1:A:250:GLU:HB2	0.44	1.89	8	1
1:A:210:TRP:NE1	1:A:251:LEU:HB2	0.44	2.28	19	1
1:A:213:LEU:HD11	1:A:239:ILE:HG21	0.44	1.89	19	1
1:A:134:ASN:HA	1:A:152:VAL:O	0.44	2.13	3	2
1:A:147:TRP:CZ2	1:A:220:MET:HB3	0.44	2.48	5	1
1:A:139:VAL:HG11	1:A:195:VAL:HG11	0.43	1.89	8	1
1:A:241:ARG:HB3	1:A:252:TYR:HB3	0.43	1.90	6	1
1:A:173:THR:HB	1:A:178:ILE:HD11	0.43	1.91	6	1
1:A:131:TYR:CZ	1:A:280:PRO:HB3	0.43	2.48	2	1
1:A:210:TRP:HE1	1:A:251:LEU:HB2	0.43	1.74	19	1
1:A:208:ILE:HD12	1:A:276:MET:HG3	0.43	1.90	18	1
1:A:151:GLN:HB3	1:A:183:LYS:HE3	0.42	1.90	18	1
1:A:204:ALA:HB3	1:A:220:MET:SD	0.42	2.54	14	1
1:A:131:TYR:OH	1:A:201:ARG:HA	0.42	2.14	2	1
1:A:149:GLU:HB2	1:A:203:ARG:NH2	0.42	2.29	16	1
1:A:172:LYS:HE2	1:A:172:LYS:N	0.42	2.30	5	1
1:A:147:TRP:CE2	1:A:220:MET:HB2	0.42	2.50	14	1
1:A:184:TYR:HB2	1:A:187:TYR:HB2	0.42	1.91	15	3
1:A:134:ASN:HA	1:A:151:GLN:NE2	0.41	2.30	10	1
1:A:179:MET:SD	1:A:196:LYS:HE3	0.41	2.55	11	1
1:A:132:LYS:HE2	1:A:287:ILE:HD11	0.41	1.93	9	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:150:ALA:HA	1:A:183:LYS:O	0.41	2.15	2	1
1:A:210:TRP:HA	1:A:213:LEU:HB2	0.41	1.92	12	1
1:A:202:ALA:HB1	1:A:286:LEU:HD12	0.41	1.92	13	1
1:A:225:VAL:HG13	1:A:257:LEU:HD12	0.41	1.93	6	1
1:A:201:ARG:NH1	1:A:220:MET:SD	0.41	2.94	3	1
1:A:211:GLU:HA	1:A:244:GLN:NE2	0.41	2.31	15	1
1:A:257:LEU:HB2	1:A:261:SER:O	0.41	2.16	14	1
1:A:204:ALA:HB2	1:A:220:MET:SD	0.40	2.56	18	1
1:A:132:LYS:HB3	1:A:287:ILE:HD11	0.40	1.93	20	1
1:A:150:ALA:HB1	1:A:182:VAL:HB	0.40	1.92	1	1
1:A:181:HIS:ND1	1:A:194:ILE:HG22	0.40	2.32	6	2
1:A:131:TYR:CZ	1:A:201:ARG:HA	0.40	2.51	19	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	157/187 (84%)	144±3 (92±2%)	12±3 (7±2%)	1±1 (1±1%)	24	71
All	All	3140/3740 (84%)	2884 (92%)	231 (7%)	25 (1%)	24	71

All 11 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	258	LEU	10
1	A	285	PRO	3
1	A	232	GLY	2
1	A	280	PRO	2
1	A	269	MET	2
1	A	283	ARG	1
1	A	295	PRO	1
1	A	205	ARG	1
1	A	281	LYS	1
1	A	174	SER	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	291	SER	1

6.3.2 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 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	141/167 (84%)	131±2 (93±1%)	10±2 (7±1%)	19 68
All	All	2820/3340 (84%)	2626 (93%)	194 (7%)	19 68

All 50 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	143	ILE	18
1	A	263	LEU	15
1	A	249	ARG	12
1	A	251	LEU	11
1	A	175	GLU	10
1	A	242	LYS	9
1	A	172	LYS	9
1	A	183	LYS	8
1	A	203	ARG	8
1	A	179	MET	6
1	A	282	GLU	6
1	A	132	LYS	6
1	A	281	LYS	5
1	A	140	ARG	5
1	A	157	LYS	5
1	A	181	HIS	5
1	A	284	ARG	4
1	A	267	ARG	4
1	A	231	ARG	3
1	A	224	ASN	3
1	A	212	ASN	3
1	A	222	ASN	2
1	A	196	LYS	2
1	A	283	ARG	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	198	LYS	2
1	A	213	LEU	2
1	A	299	ASN	2
1	A	186	ASP	2
1	A	136	TYR	2
1	A	138	ASP	2
1	A	158	ARG	2
1	A	176	ASP	1
1	A	243	ARG	1
1	A	142	ASN	1
1	A	286	LEU	1
1	A	236	ASP	1
1	A	149	GLU	1
1	A	273	GLU	1
1	A	217	GLN	1
1	A	135	GLU	1
1	A	205	ARG	1
1	A	256	ARG	1
1	A	269	MET	1
1	A	128	LEU	1
1	A	141	ASP	1
1	A	199	ASN	1
1	A	134	ASN	1
1	A	260	ASP	1
1	A	258	LEU	1
1	A	210	TRP	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates ⓘ

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

The completeness of assignment taking into account all chemical shift lists is 74% for the well-defined parts and 74% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *v1lnk_cshifts*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1963
Number of shifts mapped to atoms	1963
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	174	-0.12 ± 0.16	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	158	0.32 ± 0.12	None needed (< 0.5 ppm)
$^{13}\text{C}'$	150	0.08 ± 0.06	None needed (< 0.5 ppm)
^{15}N	150	-0.52 ± 0.47	None needed (imprecise)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 74%, i.e. 1519 atoms were assigned a chemical shift out of a possible 2062. 0 out of 29 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	688/769 (89%)	279/306 (91%)	278/314 (89%)	131/149 (88%)
Sidechain	745/1143 (65%)	456/669 (68%)	274/403 (68%)	15/71 (21%)

Continued on next page...

Continued from previous page...

	Total	^1H	^{13}C	^{15}N
Aromatic	86/150 (57%)	43/78 (55%)	43/67 (64%)	0/5 (0%)
Overall	1519/2062 (74%)	778/1053 (74%)	595/784 (76%)	146/225 (65%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 74%, i.e. 1761 atoms were assigned a chemical shift out of a possible 2366. 0 out of 32 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	793/898 (88%)	322/357 (90%)	321/368 (87%)	150/173 (87%)
Sidechain	866/1298 (67%)	533/762 (70%)	318/460 (69%)	15/76 (20%)
Aromatic	102/170 (60%)	51/88 (58%)	50/76 (66%)	1/6 (17%)
Overall	1761/2366 (74%)	906/1207 (75%)	689/904 (76%)	166/255 (65%)

7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

