Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 02:49 AM EDT

PDB ID	:	2LTW
BMRB ID	:	18499
Title	:	YAP WW1 in complex with a Smad7 derived peptide
Authors	:	Macias, M.J.; Aragon, E.; Goerner, N.; Xi, Q.; Lopes, T.; Gao, S.; Massague, J.
Deposited on	:	2012-06-04

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:

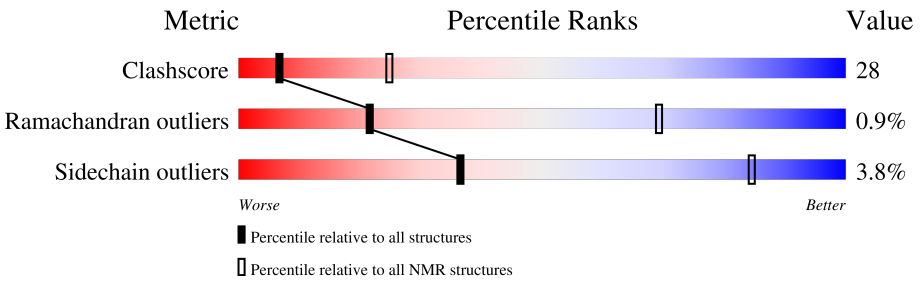
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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 is 43%.

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	36	53% 28% • 17%
2	B	14	36% 29% 36%

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 11 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:173-A:202, B:207-B:215 (39)	0.29	11

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 and 2 single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 11, 12, 13, 17, 20, 21, 24, 26, 30
2	5, 7, 8, 9, 14, 15, 16, 19, 22, 25, 27
3	4, 23, 28
4	6, 10
Single-model clusters	18; 29

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 784 atoms, of which 379 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Yorkie homolog.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	36	572	184	279	52	56	1	0

- Molecule 2 is a protein called Smad7 derived peptide.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	14	212	71	100	17	23	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	204	GLY	-	SEE REMARK 999	UNP O15105

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: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

• Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

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: Yorkie homolog

Chain A: 50% 31% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 H192 Q195 T198 W199 R203 K204 A205

• Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.2 Score per residue for model 2

- Molecule 1: Yorkie homolog

Chain A: 47% 33% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 H192 I193 D194 Q195 W199 Q200 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.3 Score per residue for model 3

- Molecule 1: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 S184 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 21% 36% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.4 Score per residue for model 4

- Molecule 1: Yorkie homolog

Chain A: 50% 31% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 H192 Q195 W199 Q200 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 29% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.5 Score per residue for model 5

- Molecule 1: Yorkie homolog

Chain A: 47% 36% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 H192 Q195 T198 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.6 Score per residue for model 6

- Molecule 1: Yorkie homolog

Chain A: 47% 36% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.7 Score per residue for model 7

- Molecule 1: Yorkie homolog

Chain A: 47% 31% 6% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 Q195 W199 Q200 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.8 Score per residue for model 8

- Molecule 1: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 36% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.9 Score per residue for model 9

- Molecule 1: Yorkie homolog

Chain A: 61% 19% • 17%

D170 V171 P172 L173 P174 E178 M179 A180 Y188 F189 L190 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 21% 43% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.10 Score per residue for model 10

- Molecule 1: Yorkie homolog

Chain A: 53% 31% 17%

D170 V171 P172 L173 P174 W177 E178 M179 T182 Q186 R187 Y188 F189 L190 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 21% 43% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.11 Score per residue for model 11 (medoid)

- Molecule 1: Yorkie homolog

Chain A: 61% 19% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 L190 Q195 Q200 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 29% 7% 36%

G204 E205 S206 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.12 Score per residue for model 12

- Molecule 1: Yorkie homolog

Chain A: 56% 25% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 S183 F189 L190 Q195 T198 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.13 Score per residue for model 13

- Molecule 1: Yorkie homolog

Chain A: 56% 22% 6% 17%

D170 V171 P172 L173 P174 W177 E178 T182 Q186 L190 N191 H192 Q195 W199 Q200 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.14 Score per residue for model 14

- Molecule 1: Yorkie homolog

Chain A: 39% 44% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 H192 Q195 T198 W199 Q200 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 43% 21% 36%

G204 E205 S206 P209 P210 Y211 P215 M216 D217

4.2.15 Score per residue for model 15

- Molecule 1: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.16 Score per residue for model 16

- Molecule 1: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

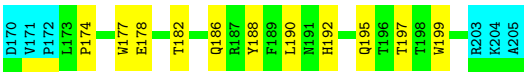
Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.17 Score per residue for model 17

- Molecule 1: Yorkie homolog

Chain A: 53% 31% 17%



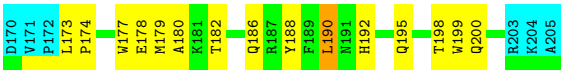
- Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

4.2.18 Score per residue for model 18

- Molecule 1: Yorkie homolog

Chain A: 42% 39% • 17%



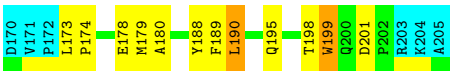
- Molecule 2: Smad7 derived peptide

Chain B: 21% 36% 7% 36%

4.2.19 Score per residue for model 19

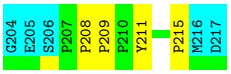
- Molecule 1: Yorkie homolog

Chain A: 50% 28% 6% 17%



- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%



4.2.20 Score per residue for model 20

- Molecule 1: Yorkie homolog

Chain A: 58% 22% • 17%

D170 V171 P172 L173 P174 W177 E178 T182 Q186 L190 Q195 T196 T197 T198 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.21 Score per residue for model 21

- Molecule 1: Yorkie homolog

Chain A: 56% 28% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 29% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.22 Score per residue for model 22

- Molecule 1: Yorkie homolog

Chain A: 50% 31% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 H192 Q195 T196 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.23 Score per residue for model 23

- Molecule 1: Yorkie homolog

Chain A: 53% 28% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Y188 F189 L190 N191 H192 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 21% 7% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.24 Score per residue for model 24

- Molecule 1: Yorkie homolog

Chain A: 47% 31% 6% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 N191 D194 Q195 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 36% 36%

G204 E205 S206 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.25 Score per residue for model 25

- Molecule 1: Yorkie homolog

Chain A: 42% 39% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 K181 T182 Q186 R187 Y188 F189 L190 Q195 T198 W199 Q200 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.26 Score per residue for model 26

- Molecule 1: Yorkie homolog

Chain A: 50% 33% 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 H192 Q195 W199 Q200 D201 P202 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 29% 36% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

4.2.27 Score per residue for model 27

- Molecule 1: Yorkie homolog

Chain A: 50% 31% • 17%

D170 V171 P172 L173 P174 W177 E178 M179 A180 Y188 F189 L190 N191 D194 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 P215 M216 D217

4.2.28 Score per residue for model 28

- Molecule 1: Yorkie homolog

Chain A: 58% 22% • 17%

D170 V171 P172 L173 P174 E178 M179 A180 K181 T182 Y188 F189 L190 Q195 W199 R203 K204 A205

- Molecule 2: Smad7 derived peptide

Chain B: 21% 43% 36%

G204 E205 S206 P207 P208 P209 P210 Y211 S212 R213 Y214 P215 M216 D217

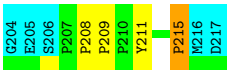
4.2.29 Score per residue for model 29

- Molecule 1: Yorkie homolog

Chain A: 50% 28% 6% 17%

- Molecule 2: Smad7 derived peptide

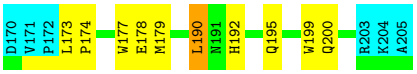
Chain B: 36% 21% 7% 36%



4.2.30 Score per residue for model 30

- Molecule 1: Yorkie homolog

Chain A: 56% 25% • 17%



- Molecule 2: Smad7 derived peptide

Chain B: 36% 29% 36%

5 Refinement protocol and experimental data overview (i)

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

Of the 300 calculated structures, 30 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
CNSSOLVE	structure solution	
CNSSOLVE	refinement	

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	2
Total number of shifts	269
Number of shifts mapped to atoms	269
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	43%

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	245	226	225	13±2
2	B	76	71	71	10±2
All	All	9630	8910	8880	513

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 28.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:178:GLU:HB3	1:A:190:LEU:HB2	0.88	1.44	29	30
2:B:209:PRO:HB3	2:B:215:PRO:HB3	0.78	1.53	29	28
1:A:178:GLU:CB	1:A:190:LEU:HB2	0.76	2.10	29	30
1:A:200:GLN:HE21	1:A:200:GLN:HA	0.76	1.41	7	1
2:B:213:ARG:HG3	2:B:214:TYR:CD1	0.75	2.17	3	3
1:A:195:GLN:HA	2:B:211:TYR:CZ	0.71	2.19	10	30
1:A:195:GLN:HA	2:B:211:TYR:CE1	0.69	2.23	17	30
2:B:209:PRO:CB	2:B:215:PRO:HB3	0.68	2.18	28	30
1:A:173:LEU:HB3	1:A:179:MET:HG3	0.67	1.65	11	9
2:B:213:ARG:HG3	2:B:214:TYR:HD1	0.67	1.50	3	2
1:A:195:GLN:HG3	2:B:211:TYR:CE2	0.63	2.29	5	26
2:B:213:ARG:HA	2:B:213:ARG:NE	0.60	2.10	24	1
1:A:195:GLN:HG3	2:B:211:TYR:CD2	0.60	2.31	17	21

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:190:LEU:HD13	2:B:211:TYR:CD1	0.60	2.32	28	28
1:A:174:PRO:HB2	1:A:177:TRP:CD1	0.60	2.31	25	22
2:B:213:ARG:NE	2:B:213:ARG:HA	0.57	2.14	8	2
1:A:173:LEU:HD22	1:A:179:MET:HG3	0.56	1.77	6	2
1:A:180:ALA:HB3	1:A:188:TYR:CE1	0.55	2.37	5	23
1:A:199:TRP:CZ2	2:B:208:PRO:HG3	0.55	2.37	28	22
2:B:209:PRO:HB3	2:B:215:PRO:CB	0.53	2.34	20	11
1:A:178:GLU:HB2	1:A:192:HIS:CD2	0.53	2.39	18	17
1:A:174:PRO:HD2	1:A:179:MET:CE	0.51	2.35	11	20
1:A:182:THR:OG1	1:A:186:GLN:HB2	0.50	2.06	5	14
2:B:208:PRO:CB	2:B:209:PRO:HD2	0.50	2.37	10	14
2:B:209:PRO:HB3	2:B:215:PRO:HA	0.50	1.84	22	2
1:A:191:ASN:CB	1:A:194:ASP:HB2	0.49	2.37	24	3
1:A:195:GLN:HG2	1:A:195:GLN:O	0.49	2.07	17	6
1:A:173:LEU:HA	1:A:179:MET:HE2	0.48	1.85	25	3
1:A:188:TYR:CD1	2:B:209:PRO:HG3	0.48	2.43	29	6
1:A:200:GLN:H	1:A:200:GLN:NE2	0.48	2.05	13	1
1:A:189:PHE:O	1:A:198:THR:HB	0.48	2.08	14	6
1:A:174:PRO:HD2	1:A:179:MET:HE2	0.48	1.86	7	7
1:A:188:TYR:CE1	2:B:209:PRO:HG3	0.47	2.43	1	6
2:B:209:PRO:HB3	2:B:215:PRO:CA	0.47	2.38	22	2
2:B:209:PRO:CG	2:B:215:PRO:HB3	0.47	2.40	22	7
1:A:191:ASN:HB2	1:A:194:ASP:HB2	0.47	1.85	24	2
1:A:199:TRP:CH2	2:B:208:PRO:HB3	0.46	2.45	15	7
1:A:177:TRP:HA	1:A:190:LEU:O	0.46	2.10	22	9
1:A:190:LEU:HD11	2:B:211:TYR:HA	0.46	1.88	14	2
2:B:213:ARG:C	2:B:214:TYR:HD1	0.45	2.15	11	4
1:A:197:THR:HG21	2:B:208:PRO:HB2	0.44	1.89	17	2
1:A:188:TYR:CE1	2:B:209:PRO:HG2	0.44	2.48	18	1
1:A:173:LEU:HD22	1:A:179:MET:HB3	0.44	1.88	7	1
2:B:213:ARG:HG3	2:B:214:TYR:CE1	0.43	2.48	3	1
1:A:178:GLU:HB2	1:A:192:HIS:HD2	0.43	1.72	16	1
1:A:198:THR:HG22	1:A:200:GLN:O	0.43	2.14	18	2
1:A:199:TRP:CE2	2:B:208:PRO:HG3	0.42	2.50	4	3
1:A:177:TRP:CH2	1:A:202:PRO:HB3	0.42	2.50	14	1
1:A:178:GLU:O	1:A:189:PHE:HA	0.42	2.15	29	2
2:B:211:TYR:O	2:B:211:TYR:CG	0.42	2.72	27	7
1:A:186:GLN:HB3	1:A:199:TRP:CZ3	0.41	2.51	18	1
1:A:195:GLN:O	1:A:195:GLN:HG2	0.41	2.15	3	1
1:A:200:GLN:HE21	1:A:200:GLN:CA	0.41	2.20	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	30/36 (83%)	30±0 (99±1%)	0±0 (1±1%)	0±0 (0±0%)	100 100
2	B	9/14 (64%)	6±1 (61±6%)	3±0 (34±5%)	0±0 (4±5%)	5 31
All	All	1170/1500 (78%)	1058 (90%)	101 (9%)	11 (1%)	21 69

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
2	B	215	PRO	11

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	26/31 (84%)	25±1 (95±3%)	1±1 (5±3%)	31 79
2	B	9/13 (69%)	9±0 (99±4%)	0±0 (1±4%)	66 95
All	All	1050/1320 (80%)	1010 (96%)	40 (4%)	36 84

All 11 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	190	LEU	17
1	A	199	TRP	6
1	A	200	GLN	4
1	A	182	THR	3
1	A	191	ASN	2
2	B	208	PRO	2
2	B	213	ARG	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	184	SER	1
1	A	183	SER	1
1	A	192	HIS	1
1	A	196	THR	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)

The completeness of assignment taking into account all chemical shift lists is 43% for the well-defined parts and 40% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *YAPWW1*

7.1.1 Bookkeeping (i)

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	229
Number of shifts mapped to atoms	229
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 36%, i.e. 190 atoms were assigned a chemical shift out of a possible 527. 0 out of 2 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}**H**	^{13}**C**	^{15}**N**
Backbone	58/183 (32%)	58/73 (79%)	0/78 (0%)	0/32 (0%)
Sidechain	113/275 (41%)	113/178 (63%)	0/86 (0%)	0/11 (0%)
Aromatic	19/69 (28%)	19/33 (58%)	0/32 (0%)	0/4 (0%)
Overall	190/527 (36%)	190/284 (67%)	0/196 (0%)	0/47 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 34%, i.e. 227 atoms were assigned a chemical shift out of a possible 663. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	68/237 (29%)	68/95 (72%)	0/100 (0%)	0/42 (0%)
Sidechain	140/357 (39%)	140/230 (61%)	0/112 (0%)	0/15 (0%)
Aromatic	19/69 (28%)	19/33 (58%)	0/32 (0%)	0/4 (0%)
Overall	227/663 (34%)	227/358 (63%)	0/244 (0%)	0/61 (0%)

7.1.4 Statistically unusual chemical shifts (i)

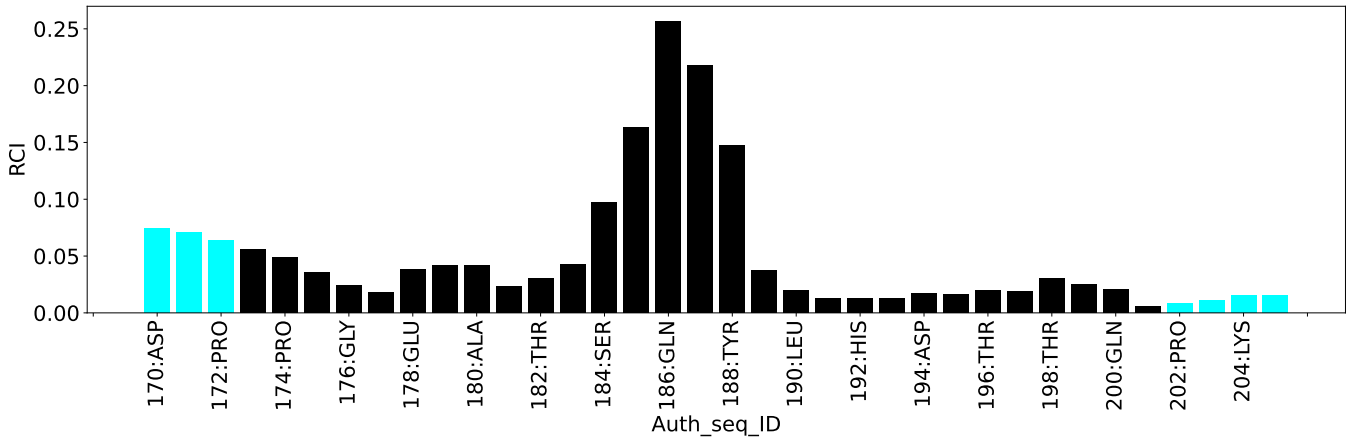
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	182	THR	HG1	4.76	0.08 – 2.19	17.2
1	A	191	ASN	HB3	0.02	1.12 – 4.38	-8.4
1	A	201	ASP	HA	2.36	3.04 – 6.12	-7.2
1	A	202	PRO	HB3	0.12	0.25 – 3.76	-5.4
1	A	202	PRO	HB2	0.34	0.37 – 3.78	-5.1

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *SMAD7*

7.2.1 Bookkeeping (i)

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	40
Number of shifts mapped to atoms	40
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

7.2.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 7%, i.e. 35 atoms were assigned a chemical shift out of a possible 527. 0 out of 2 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	9/183 (5%)	9/73 (12%)	0/78 (0%)	0/32 (0%)
Sidechain	18/275 (7%)	18/178 (10%)	0/86 (0%)	0/11 (0%)
Aromatic	8/69 (12%)	8/33 (24%)	0/32 (0%)	0/4 (0%)
Overall	35/527 (7%)	35/284 (12%)	0/196 (0%)	0/47 (0%)

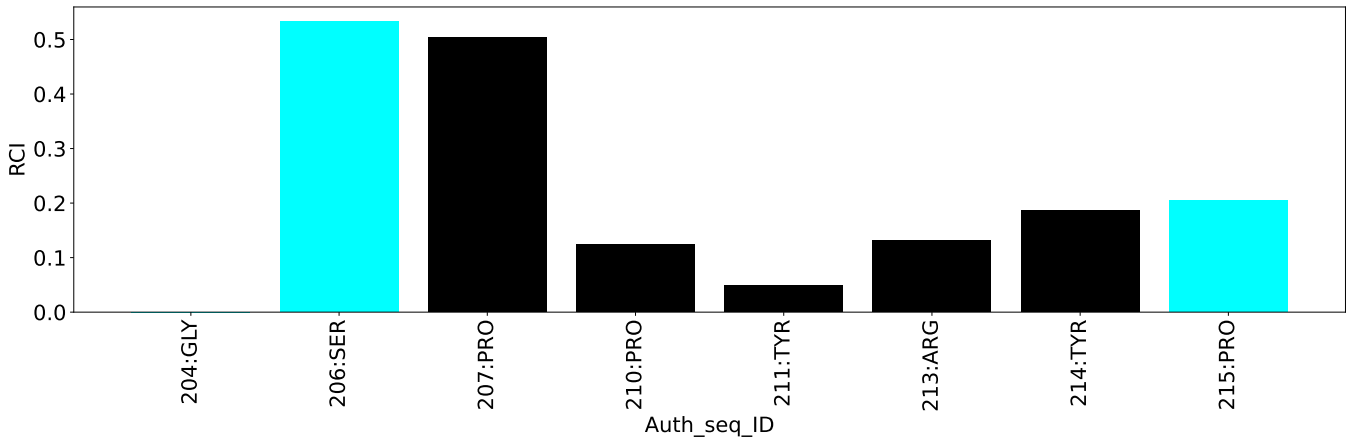
The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 6%, i.e. 39 atoms were assigned a chemical shift out of a possible 663. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	12/237 (5%)	12/95 (13%)	0/100 (0%)	0/42 (0%)
Sidechain	19/357 (5%)	19/230 (8%)	0/112 (0%)	0/15 (0%)
Aromatic	8/69 (12%)	8/33 (24%)	0/32 (0%)	0/4 (0%)
Overall	39/663 (6%)	39/358 (11%)	0/244 (0%)	0/61 (0%)

7.2.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.2.5 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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain B: