Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 10:12 pm BST

PDB ID	:	2JT5
Title	:	solution structure of matrix metalloproteinase 3 (MMP-3) in the presence of n-hydroxy-2-[n-(2-hydroxyethyl)biphenyl-4-sulfonamide] hydroxamic acid (MLC88)
Authors	:	Alcaraz, L.A.; Banci, L.; Bertini, I.; Cantini, F.; Donaire, A.; Gonnelli, L.
Deposited on	:	2007-07-20

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 (1)) were used in the production of this report:

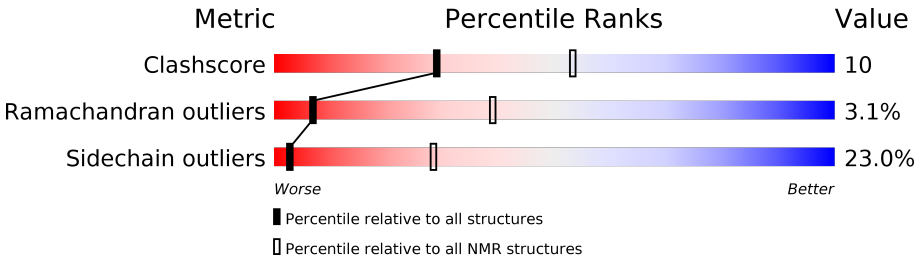
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 15%.

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	161	61% 33% 6%

2 Ensemble composition and analysis

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2536 atoms, of which 1230 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Stromelysin-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	161	2491	818	1213	213	245	2	0

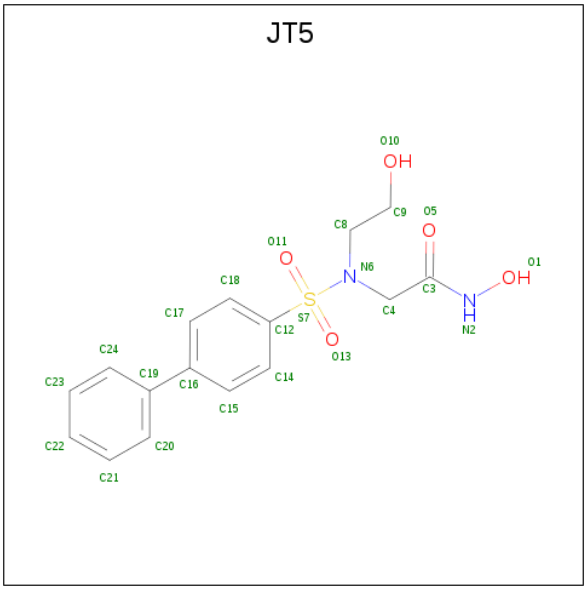
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	2	2	2

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
			Total	Ca
3	A	2	2	2

- Molecule 4 is N 2 -(biphenyl-4-ylsulfonyl)-N-hydroxy-N 2 -(2-hydroxyethyl)glycinamide (three-letter code: JT5) (formula: $C_{16}H_{18}N_2O_5S$).



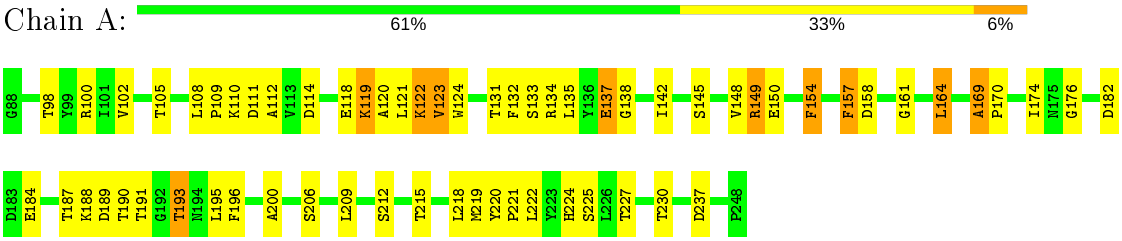
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
4	A	1	41	16	17	2	5	1

4 Residue-property plots (i)

These plots are provided for all protein, RNA and DNA 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: Stromelysin-1

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 1 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
X-PLOR NIH	refinement	
AutoDock	refinement	3.0

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)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	300
Number of shifts mapped to atoms	298
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	15%

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

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: JT5, ZN, CA

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	1278	1213	1207	25
4	A	24	17	17	2
All	All	1306	1230	1224	26

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 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:149:ARG:HA	1:A:157:PHE:HB3	0.78	1.55
1:A:120:ALA:HB1	1:A:200:ALA:HA	0.66	1.68
1:A:142:ILE:HA	1:A:176:GLY:O	0.59	1.96
1:A:148:VAL:HG22	1:A:149:ARG:H	0.59	1.56
1:A:112:ALA:HA	1:A:188:LYS:CE	0.58	2.28
1:A:200:ALA:HB1	1:A:218:LEU:HD21	0.56	1.77
1:A:176:GLY:HA2	1:A:206:SER:HB3	0.55	1.77
1:A:169:ALA:HB3	1:A:170:PRO:HD3	0.54	1.79
1:A:158:ASP:OD2	1:A:161:GLY:HA3	0.54	2.03
1:A:220:TYR:CD1	1:A:221:PRO:HD2	0.54	2.37
1:A:98:THR:HG22	1:A:133:SER:HB2	0.52	1.82
1:A:112:ALA:HA	1:A:188:LYS:HE3	0.50	1.83

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:108:LEU:HB3	1:A:109:PRO:HD2	0.50	1.84
1:A:212:SER:HB3	1:A:237:ASP:OD2	0.48	2.08
1:A:184:GLU:HG2	1:A:193:THR:OG1	0.47	2.09
1:A:220:TYR:CE2	1:A:222:LEU:HB2	0.46	2.45
1:A:209:LEU:HD23	1:A:219:MET:HG3	0.45	1.89
1:A:148:VAL:HG13	1:A:150:GLU:HG2	0.44	1.88
4:A:1:JT5:H33	4:A:1:JT5:H30	0.44	1.89
1:A:124:TRP:CZ2	1:A:218:LEU:HD23	0.43	2.49
1:A:119:LYS:O	1:A:123:VAL:HB	0.43	2.13
1:A:121:LEU:HD11	1:A:132:PHE:CB	0.43	2.44
1:A:164:LEU:HD12	4:A:1:JT5:H36	0.42	1.92
1:A:119:LYS:HB3	1:A:196:PHE:CE1	0.41	2.51
1:A:122:LYS:HB2	1:A:122:LYS:HZ3	0.41	1.76
1:A:188:LYS:HG3	1:A:188:LYS:O	0.40	2.16

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	159/161 (99%)	127 (80%)	27 (17%)	5 (3%)	7 39
All	All	159/161 (99%)	127 (80%)	27 (17%)	5 (3%)	7 39

All 5 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	137	GLU
1	A	138	GLY
1	A	154	PHE
1	A	169	ALA
1	A	174	ILE

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	135/135 (100%)	104 (77%)	31 (23%)	3 28
All	All	135/135 (100%)	104 (77%)	31 (23%)	3 28

All 31 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	227	THR
1	A	195	LEU
1	A	114	ASP
1	A	190	THR
1	A	157	PHE
1	A	123	VAL
1	A	164	LEU
1	A	119	LYS
1	A	149	ARG
1	A	110	LYS
1	A	135	LEU
1	A	134	ARG
1	A	187	THR
1	A	182	ASP
1	A	191	THR
1	A	105	THR
1	A	131	THR
1	A	224	HIS
1	A	100	ARG
1	A	122	LYS
1	A	111	ASP
1	A	118	GLU
1	A	154	PHE
1	A	145	SER
1	A	225	SER
1	A	137	GLU
1	A	193	THR
1	A	230	THR
1	A	215	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	102	VAL
1	A	189	ASP

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 carbohydrates in this entry.

LIGAND-GEOMETRY INFOmissingINFO

6.6 Other polymers

There are no such molecules in this entry.

6.7 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 15% for the well-defined parts and 15% for the entire structure.

7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

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	300
Number of shifts mapped to atoms	298
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. All 2 occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	67	ASP	H	7.76	-1.0	1
A	67	ASP	N	115.788	-1.0	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	0	—	None (insufficient data)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	150	0.16 ± 0.34	None needed (< 0.5 ppm)

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 15%, i.e. 298 atoms were assigned a chemical shift out of a possible 1941. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	298/785 (38%)	149/312 (48%)	0/322 (0%)	149/151 (99%)
Sidechain	0/926 (0%)	0/541 (0%)	0/352 (0%)	0/33 (0%)
Aromatic	0/230 (0%)	0/122 (0%)	0/95 (0%)	0/13 (0%)
Overall	298/1941 (15%)	149/975 (15%)	0/769 (0%)	149/197 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 15%, i.e. 298 atoms were assigned a chemical shift out of a possible 1941. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	298/785 (38%)	149/312 (48%)	0/322 (0%)	149/151 (99%)
Sidechain	0/926 (0%)	0/541 (0%)	0/352 (0%)	0/33 (0%)
Aromatic	0/230 (0%)	0/122 (0%)	0/95 (0%)	0/13 (0%)
Overall	298/1941 (15%)	149/975 (15%)	0/769 (0%)	149/197 (76%)

7.1.4 Statistically unusual chemical shifts

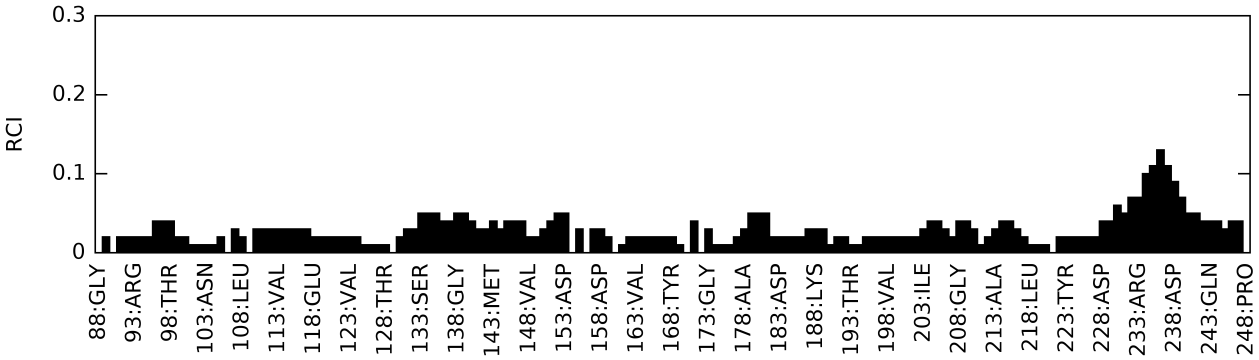
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	218	LEU	H	11.51	11.47 – 4.97	5.1

7.1.5 Random Coil Index (RCI) plots

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:


RCI
0.3
0.2
0.1
0
88:GLY
93:ARG
98:THR
103:ASN
108:LEU
113:VAL
118:GLU
123:VAL
128:THR
133:SER
138:GLY
143:MET
148:VAL
153:ASP
158:ASP
163:VAL
168:TYR
173:GLY
178:ALA
183:ASP
188:LYS
193:THR
198:VAL
203:ILE
208:GLY
213:ALA
218:LEU
223:TYR
228:ASP
233:ARG
238:ASP
243:GLN
248:PRO