Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 10:13 pm BST

PDB ID	:	2JT6
Title	:	Solution structure of matrix metalloproteinase 3 (MMP-3) in the presence of 3 -4'-cyanobyphenyl-4-yloxy)-n-hdydroxypropionamide (MMP-3 inhibitor VII)
Authors	:	Alcaraz, L.A.; Banci, L.; Bertini, I.; Cantini, F.; Donaire, A.; Gonnelli, L.
Deposited on	:	2007-07-23

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:

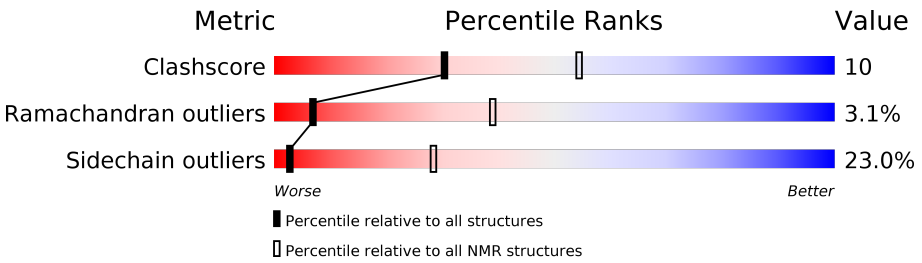
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
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 (i)

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	63% 31% 6% •

2 Ensemble composition and analysis (i)

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

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2529 atoms, of which 1226 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

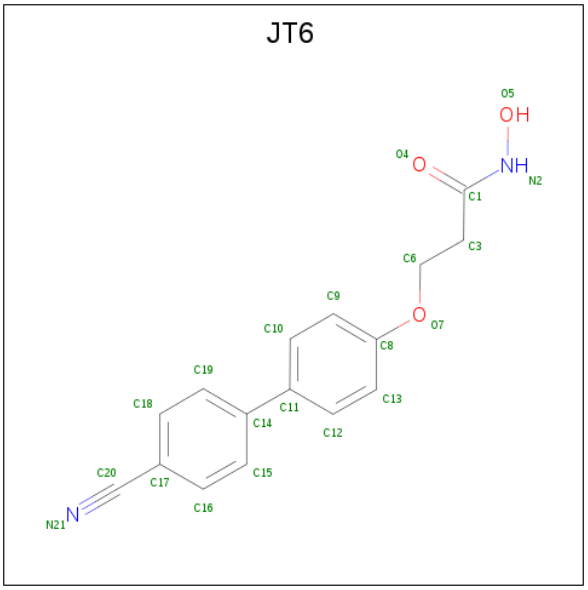
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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 3-[(4'-cyanobiphenyl-4-yl)oxy]-N-hydroxypropanamide (three-letter code: JT6) (formula: $C_{16}H_{14}N_2O_3$).



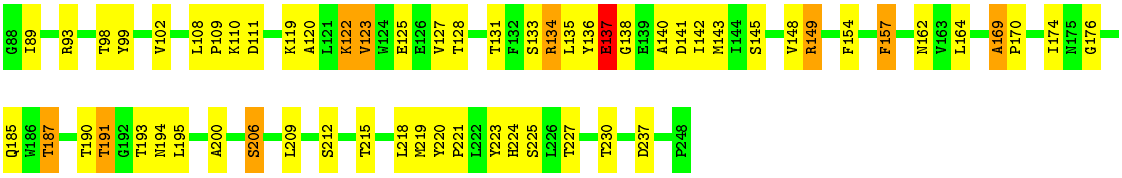
Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
4	A	1	34	16	13	2	3

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: 63% 31% 6% •



5 Refinement protocol and experimental data overview (i)

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 (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, JT6, 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 (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	1278	1213	1207	25
All	All	1303	1226	1220	25

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	0.69	1.62
1:A:120:ALA:HB1	1:A:200:ALA:HA	0.67	1.66
1:A:135:LEU:HD13	1:A:137:GLU:HG2	0.65	1.68
1:A:200:ALA:HB1	1:A:218:LEU:HD21	0.59	1.73
1:A:169:ALA:HB3	1:A:170:PRO:HD3	0.55	1.79
1:A:148:VAL:HG22	1:A:149:ARG:H	0.51	1.64
1:A:135:LEU:HB2	1:A:137:GLU:OE2	0.51	2.05
1:A:176:GLY:HA2	1:A:206:SER:CB	0.51	2.35
1:A:135:LEU:HD11	1:A:140:ALA:HB2	0.51	1.82
1:A:209:LEU:HD23	1:A:219:MET:HG3	0.49	1.82
1:A:99:TYR:CZ	1:A:134:ARG:HB2	0.47	2.44
1:A:142:ILE:HA	1:A:176:GLY:O	0.47	2.08
1:A:176:GLY:HA2	1:A:206:SER:HB2	0.46	1.87

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	0.46	1.88
1:A:164:LEU:HD22	1:A:193:THR:HG23	0.45	1.88
1:A:187:THR:HG21	1:A:191:THR:HB	0.45	1.89
1:A:187:THR:O	1:A:194:ASN:HA	0.44	2.12
1:A:220:TYR:CD1	1:A:221:PRO:HD2	0.44	2.47
1:A:122:LYS:HA	1:A:125:GLU:HB3	0.42	1.92
1:A:212:SER:HB3	1:A:237:ASP:OD2	0.42	2.15
1:A:89:ILE:HG13	1:A:93:ARG:NH1	0.42	2.30
1:A:119:LYS:O	1:A:123:VAL:HB	0.41	2.16
1:A:149:ARG:CA	1:A:157:PHE:HB3	0.41	2.40
1:A:212:SER:OG	1:A:219:MET:HB2	0.41	2.16
1:A:137:GLU:H	1:A:137:GLU:CD	0.40	2.19

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%)	125 (79%)	29 (18%)	5 (3%)	7 39
All	All	159/161 (99%)	125 (79%)	29 (18%)	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	174	ILE
1	A	137	GLU
1	A	169	ALA
1	A	138	GLY
1	A	154	PHE

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	187	THR
1	A	122	LYS
1	A	190	THR
1	A	137	GLU
1	A	131	THR
1	A	206	SER
1	A	223	TYR
1	A	191	THR
1	A	123	VAL
1	A	133	SER
1	A	230	THR
1	A	136	TYR
1	A	162	ASN
1	A	98	THR
1	A	149	ARG
1	A	110	LYS
1	A	195	LEU
1	A	128	THR
1	A	127	VAL
1	A	145	SER
1	A	215	THR
1	A	227	THR
1	A	225	SER
1	A	224	HIS
1	A	141	ASP
1	A	134	ARG
1	A	102	VAL
1	A	185	GLN
1	A	143	MET
1	A	157	PHE
1	A	111	ASP

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 carbohydrates in this entry.

LIGAND-GEOMETRY INFOmissingINFO

6.6 Other polymers (i)

There are no such molecules in this entry.

6.7 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 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 (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	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	7.78	-1.0	1
A	67	ASP	N	115.56	-1.0	1

7.1.2 Chemical shift referencing (i)

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	0.17 ± 0.35	None needed (< 0.5 ppm)

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 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 (i)

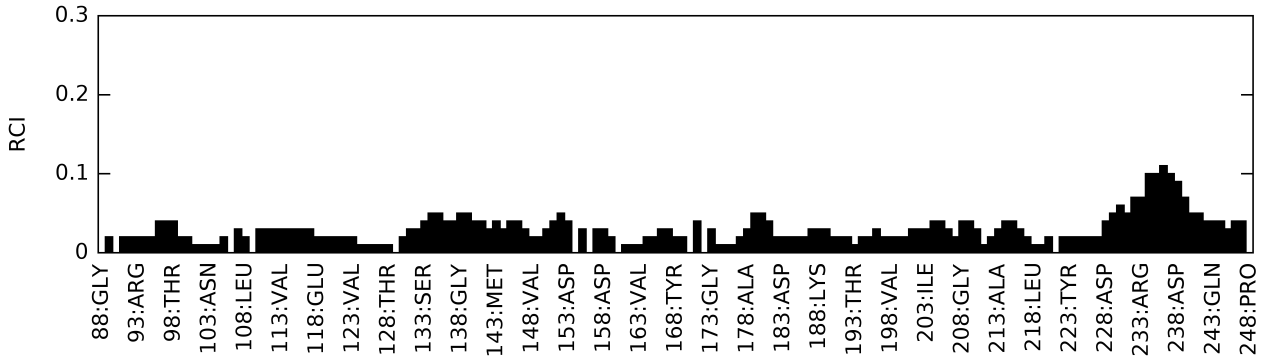
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

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 (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:


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