



wwPDB NMR Structure Validation Summary Report ⓘ

Feb 12, 2017 – 10:31 pm GMT

PDB ID : 2JXY
Title : Solution structure of the hemopexin-like domain of MMP12
Authors : Bertini, I.; Calderone, V.; Fragai, M.; Jaiswal, R.; Luchinat, C.; Melikian, M.
Deposited on : 2007-12-01

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : trunk28760
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

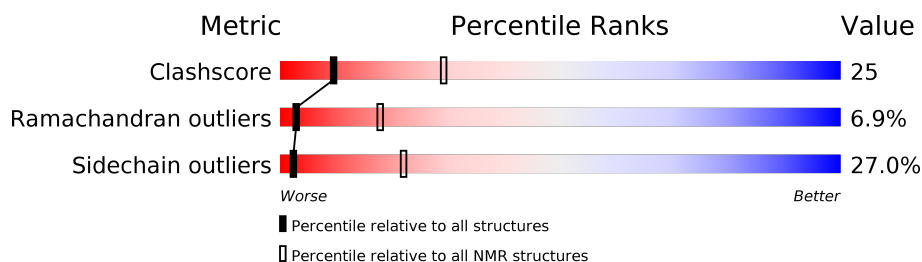
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 79%.

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	194	

2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:281-A:470 (190)	0.70	1

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 4 single-model clusters were found.

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

3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3243 atoms, of which 1595 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Macrophage metalloelastase.

Mol	Chain	Residues	Atoms						Trace
1	A	194	Total	C	H	N	O	S	0
			3242	1089	1595	267	286	5	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	MET	-	EXPRESSION TAG	UNP P39900

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

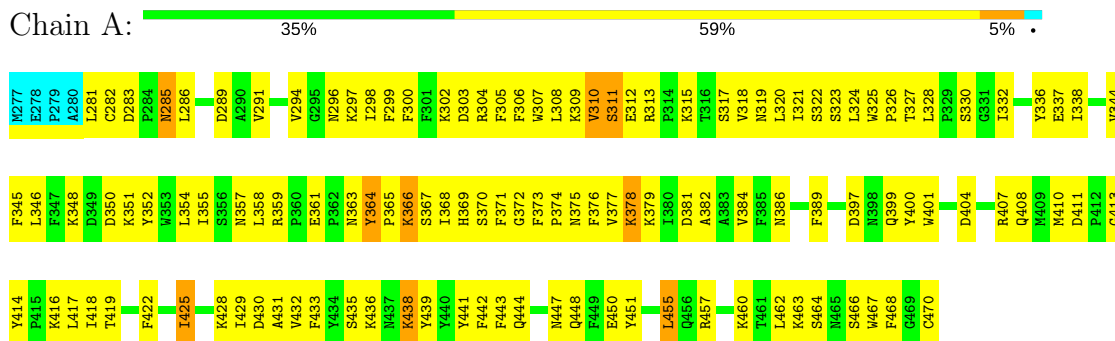
Mol	Chain	Residues	Atoms	
2	A	1	Total	Ca
			1	1

4 Residue-property plots [i](#)

4.1 Average score per residue in the NMR ensemble

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.

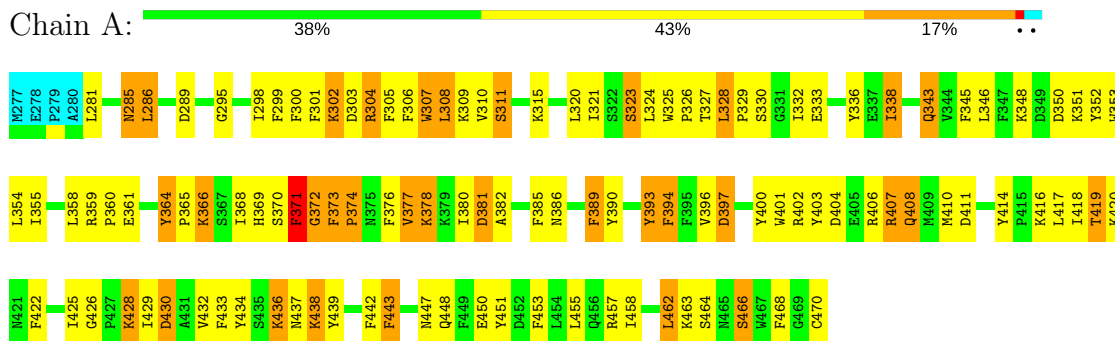
- Molecule 1: Macrophage metalloelastase



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Macrophage metalloelastase



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing*.

Of the 1600 calculated structures, 20 were deposited, based on the following criterion: *lowest target function*.

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

Software name	Classification	Version
DYANA	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)	BMRB entry 7414, BMRB entry 15578
Number of chemical shift lists	2
Total number of shifts	5866
Number of shifts mapped to atoms	4253
Number of unparsed shifts	0
Number of shifts with mapping errors	1613
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	79%

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:
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	1618	1568	1565	80±15
All	All	32380	31360	31324	1602

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 25.

5 of 1014 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:281:LEU:HD11	1:A:307:TRP:CZ3	1.00	1.91	17	1
1:A:328:LEU:HD21	1:A:355:ILE:HD11	0.94	1.40	10	3
1:A:346:LEU:HD23	1:A:355:ILE:HD13	0.92	1.37	2	2
1:A:291:VAL:HG21	1:A:442:PHE:CZ	0.91	2.01	15	1
1:A:298:ILE:HD13	1:A:300:PHE:CZ	0.90	2.01	13	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	189/194 (97%)	138±5 (73±3%)	38±5 (20±3%)	13±3 (7±1%)	3	18
All	All	3780/3880 (97%)	2753 (73%)	767 (20%)	260 (7%)	3	18

5 of 48 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	326	PRO	20
1	A	378	LYS	16
1	A	413	GLY	15
1	A	455	LEU	14
1	A	303	ASP	13

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	175/178 (98%)	128±5 (73±3%)	47±5 (27±3%)	2	22
All	All	3500/3560 (98%)	2556 (73%)	944 (27%)	2	22

5 of 139 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	304	ARG	18
1	A	410	MET	18
1	A	364	TYR	17
1	A	407	ARG	16
1	A	378	LYS	16

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.

6.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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 79% for the well-defined parts and 78% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 7414

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	2173
Number of shifts mapped to atoms	2173
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	31

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	192	2.92 ± 0.14	Should be applied
$^{13}\text{C}_\beta$	184	3.19 ± 0.17	Should be applied
$^{13}\text{C}'$	176	3.15 ± 0.12	Should be applied
^{15}N	179	0.21 ± 0.22	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 77%, i.e. 1972 atoms were assigned a chemical shift out of a possible 2546. 22 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	909/926 (98%)	367/368 (100%)	364/380 (96%)	178/178 (100%)
Sidechain	992/1250 (79%)	596/746 (80%)	380/438 (87%)	16/66 (24%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	71/370 (19%)	66/196 (34%)	0/167 (0%)	5/7 (71%)
Overall	1972/2546 (77%)	1029/1310 (79%)	744/985 (76%)	199/251 (79%)

7.1.4 Statistically unusual chemical shifts ⓘ

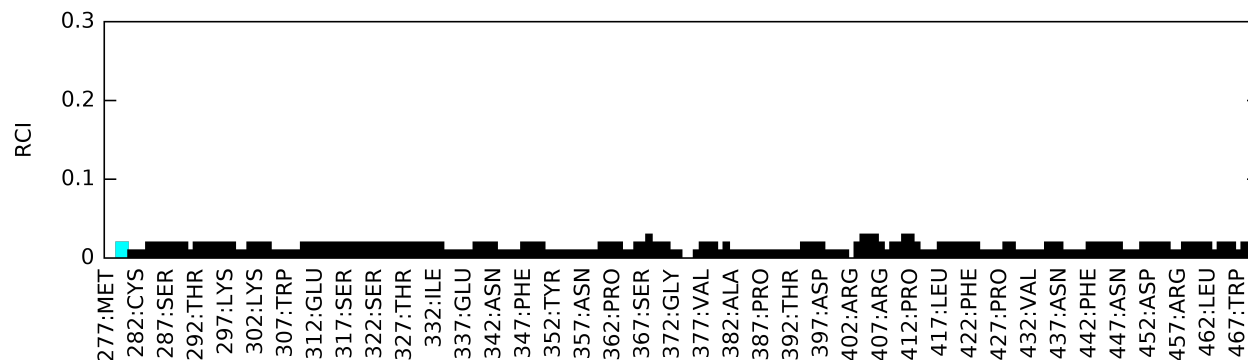
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	359	ARG	CD	23.47	47.57 – 38.77	-22.4
1	A	359	ARG	CG	40.33	33.23 – 21.23	10.9
1	A	392	THR	HG23	-1.21	2.29 – -0.01	-10.2
1	A	392	THR	HG21	-1.21	2.29 – -0.01	-10.2
1	A	392	THR	HG22	-1.21	2.29 – -0.01	-10.2
1	A	369	HIS	HD2	10.35	9.28 – 4.78	7.4
1	A	450	GLU	HG3	0.75	3.31 – 1.21	-7.2
1	A	415	PRO	HB3	-0.47	3.81 – 0.21	-6.9
1	A	410	MET	CG	23.73	38.33 – 25.73	-6.6
1	A	365	PRO	CG	20.18	32.66 – 21.76	-6.4
1	A	365	PRO	HB3	-0.30	3.81 – 0.21	-6.4
1	A	359	ARG	HD2	1.65	4.27 – 1.97	-6.4
1	A	298	ILE	HG21	-0.94	2.13 – -0.57	-6.4
1	A	298	ILE	HG23	-0.94	2.13 – -0.57	-6.4
1	A	298	ILE	HG22	-0.94	2.13 – -0.57	-6.4
1	A	416	LYS	HE3	1.70	3.86 – 1.96	-6.3
1	A	297	LYS	CD	21.56	34.86 – 23.06	-6.3
1	A	359	ARG	HD3	1.54	4.36 – 1.86	-6.3
1	A	359	ARG	HG2	3.19	2.92 – 0.22	6.0
1	A	402	ARG	CG	20.12	33.23 – 21.23	-5.9
1	A	425	ILE	HG23	-0.78	2.13 – -0.57	-5.8
1	A	425	ILE	HG22	-0.78	2.13 – -0.57	-5.8
1	A	425	ILE	HG21	-0.78	2.13 – -0.57	-5.8
1	A	319	ASN	HB3	0.88	4.41 – 1.11	-5.7
1	A	415	PRO	CG	21.23	32.66 – 21.76	-5.5
1	A	359	ARG	HG3	3.13	3.00 – 0.10	5.5
1	A	427	PRO	CG	21.31	32.66 – 21.76	-5.4
1	A	448	GLN	HG3	0.82	3.75 – 0.85	-5.1
1	A	380	ILE	HD13	-0.79	2.13 – -0.77	-5.1
1	A	380	ILE	HD12	-0.79	2.13 – -0.77	-5.1
1	A	380	ILE	HD11	-0.79	2.13 – -0.77	-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:



7.2 Chemical shift list 2

File name: BMRB entry 15578

Chemical shift list name: *assigned_chem_shift_list_1*

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	3693
Number of shifts mapped to atoms	2080
Number of unparsed shifts	0
Number of shifts with mapping errors	1613
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	19

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

- Residue not found in structure. First 5 (of 1613) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	152	ARG	CD	40.569	0.3	1
A	126	SER	HB3	3.839	0.02	2
A	68	HIS	HB2	2.02	0.02	2
A	126	SER	C	170.638	0.3	1
A	109	PHE	H	9.128	0.02	1

7.2.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	344	2.59 ± 0.09	Should be applied
$^{13}\text{C}_\beta$	313	2.79 ± 0.18	Should be applied
$^{13}\text{C}'$	338	3.09 ± 0.07	Should be applied
^{15}N	342	0.22 ± 0.41	None needed (< 0.5 ppm)

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 74%, i.e. 1874 atoms were assigned a chemical shift out of a possible 2546. 22 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	909/926 (98%)	362/368 (98%)	371/380 (98%)	176/178 (99%)
Sidechain	955/1250 (76%)	569/746 (76%)	370/438 (84%)	16/66 (24%)
Aromatic	10/370 (3%)	5/196 (3%)	0/167 (0%)	5/7 (71%)
Overall	1874/2546 (74%)	936/1310 (71%)	741/985 (75%)	197/251 (78%)

7.2.4 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	392	THR	HG22	-1.19	2.29 – -0.01	-10.1
1	A	392	THR	HG23	-1.19	2.29 – -0.01	-10.1
1	A	392	THR	HG21	-1.19	2.29 – -0.01	-10.1
1	A	450	GLU	HG3	0.81	3.31 – 1.21	-6.9
1	A	298	ILE	HG21	-0.94	2.13 – -0.57	-6.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	298	ILE	HG23	-0.94	2.13 – -0.57	-6.4
1	A	298	ILE	HG22	-0.94	2.13 – -0.57	-6.4
1	A	410	MET	CG	24.07	38.33 – 25.73	-6.3
1	A	134	PRO	CB	24.69	37.79 – 25.89	-6.0
1	A	128	PRO	HA	2.46	6.05 – 2.75	-5.9
1	A	425	ILE	HG21	-0.79	2.13 – -0.57	-5.8
1	A	425	ILE	HG23	-0.79	2.13 – -0.57	-5.8
1	A	425	ILE	HG22	-0.79	2.13 – -0.57	-5.8
1	A	106	THR	HB	2.35	5.82 – 2.52	-5.5
1	A	352	TYR	HB3	0.91	4.75 – 0.95	-5.1
1	A	448	GLN	HG3	0.83	3.75 – 0.85	-5.1
1	A	380	ILE	HD13	-0.79	2.13 – -0.77	-5.1
1	A	380	ILE	HD12	-0.79	2.13 – -0.77	-5.1
1	A	380	ILE	HD11	-0.79	2.13 – -0.77	-5.1

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.

Random coil index (RCI) for chain A:

