wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 08:39 PM BST

PDB ID	:	2FIN
Title	:	Solution Structure of the complex between poxvirus-encoded CC chemokine inhibitor vCCI and human MIP-1beta, ensemble structure
Authors	:	Zhang, L.
Deposited on	:	2005-12-29

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 (i) symbol.

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

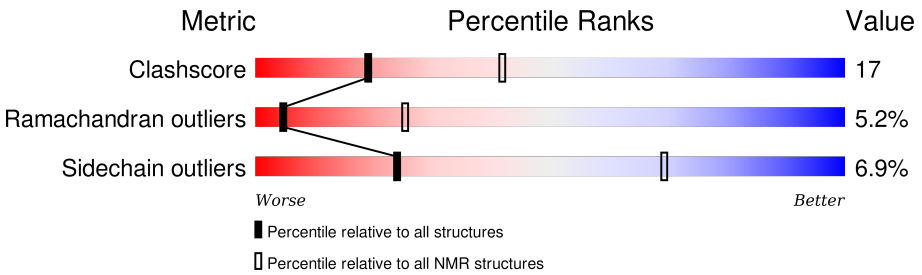
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 27%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	242	61% 26% • 12%
2	B	69	58% 25% • 13%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 15 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:8-A:54, A:77-A:242, B:10-B:69 (273)	0.31	15

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 1 single-model cluster was found.

Cluster number	Models
1	3, 5, 7, 8, 12, 15
2	2, 4, 13
3	1, 6, 9
4	10, 14
Single-model clusters	11

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4601 atoms, of which 2247 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called rabbitpox encoded CC chemokine inhibitor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	242	3575	1104	1755	298	405	13	0

- Molecule 2 is a protein called Small inducible cytokine A4.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	69	1026	337	492	83	109	5	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	45	ALA	LYS	ENGINEERED	UNP P13236
B	46	ALA	ARG	ENGINEERED	UNP P13236
B	48	ALA	LYS	ENGINEERED	UNP P13236

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: rabbitpox encoded CC chemokine inhibitor

Chain A: 61% 26% • 12%

M1 P2 A3 S4 L5 Q6 Q7 E18 E19 N20 K21 H22 H23 M24 P39 T40 N41 D42 K43 I44 C45 Q46 T52 E53 S54 E55 S56 D57 P58 D59 P60 E61 V62 E63 S64 E65 D66 D67 S68 T69 S70 V71 E72 D73 V74 D75 P76 Y80 Y81 S82 I83 R89 M90 N91 F92

G93 F94 P98 N113 A114 R115 S120 P121 I130 T131 H132 M138 E143 R149 E153 E154 K155 D158 I159 K160 P163 S170 H171 V174 E177 D178 T183 K188 C189 V190 L193 V197 R198 M202 S206 S207 E208 L209 K212 D213

G214 V218 D228 L232 I233 L238 C241 V242

- Molecule 2: Small inducible cytokine A4

Chain B: 58% 25% • 13%

A1 P2 M3 G4 S5 D6 P7 P8 T9 A10 C11 C12 F13 S14 Y15 R18 K19 L20 P21 R22 N23 F24 Y29 E30 T31 S32 S33 L34 C35 P38 T44 A45 A46 S47 A48 C51 N69

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

The representative model is number 15. Colouring as in section 4.1 above.

- Molecule 1: rabbitpox encoded CC chemokine inhibitor

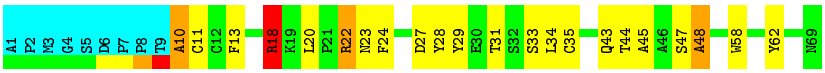
Chain A: 60% 24% 5% 12%

M1 P2 A3 S4 L5 Q6 Q7 E18 K21 H22 H23 M24 I29 I30 Q37 I44 C45 Q46 S47 V48 T49 E50 E53 S54 E55 S56 D57 P58 D59 P60 E61 V62 E63 S64 E65 D66 D67 S68 T69 S70 V71 E72 D73 V74 D75 P76 Y80 Y81 S82 R89 M90 N91

Q99 S106 N113 A114 R115 S120 P121 P128 A129 I130 T131 H132 I139 K140 D141 C142 E143 R149 E153 E154 K155 D158 I159 K160 S167 N168 I169 V174 S175 Y176 E177 D178 I179 I180 T183 I184 V185 D186 T187 K188 C189 L193 V197 R198 L209

- Molecule 2: Small inducible cytokine A4

Chain B:



5 Refinement protocol and experimental data overview (i)

Of the 100 calculated structures, 15 were deposited, based on the following criterion: *structures with the least restraint violations,structures with the lowest energy.*

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

Software name	Classification	Version
Dynamo	structure solution	3.1
Dynamo	refinement	3.1

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 6809
Number of chemical shift lists	1
Total number of shifts	973
Number of shifts mapped to atoms	973
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	27%

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)

There are no covalent bond-length or bond-angle outliers.

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	3.9±0.6
2	B	0.0±0.0	1.7±0.5
All	All	0	84

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 8 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	149	ARG	Sidechain	15
1	A	89	ARG	Sidechain	15
2	B	18	ARG	Sidechain	14
1	A	198	ARG	Sidechain	13
2	B	22	ARG	Sidechain	11

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	1602	1571	1571	54±10
2	B	475	436	436	17±2
All	All	31155	30105	30105	1026

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 17.

5 of 644 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:139:ILE:HD13	1:A:140:LYS:N	0.85	1.86	4	2
1:A:116:LEU:HD22	1:A:116:LEU:N	0.79	1.92	11	2
2:B:68:LEU:HD23	2:B:69:ASN:N	0.75	1.96	2	1
1:A:116:LEU:HD12	1:A:116:LEU:N	0.71	1.99	3	1
1:A:125:LYS:HZ1	1:A:127:SER:CB	0.70	1.99	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	212/242 (88%)	168±4 (79±2%)	34±3 (16±1%)	10±2 (5±1%)	5	28
2	B	59/69 (86%)	41±2 (70±3%)	14±3 (23±4%)	4±1 (7±2%)	3	18
All	All	4065/4665 (87%)	3142 (77%)	711 (17%)	212 (5%)	5	25

5 of 49 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	48	ALA	15
2	B	31	THR	15
1	A	163	PRO	13
1	A	207	SER	11
1	A	39	PRO	10

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	191/219 (87%)	176±4 (92±2%)	15±4 (8±2%)	21 66
2	B	53/60 (88%)	51±1 (96±2%)	2±1 (4±2%)	40 83
All	All	3660/4185 (87%)	3407 (93%)	253 (7%)	24 69

5 of 109 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	80	TYR	11
1	A	232	LEU	8
2	B	18	ARG	8
1	A	202	MET	6
1	A	24	MET	6

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)

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 27% for the well-defined parts and 28% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 6809

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	973
Number of shifts mapped to atoms	973
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 (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	212	0.32 ± 0.08	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	200	0.67 ± 0.11	Should be applied
$^{13}C'$	171	0.18 ± 0.14	None needed (< 0.5 ppm)
^{15}N	195	-0.23 ± 0.32	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 27%, i.e. 856 atoms were assigned a chemical shift out of a possible 3143. 0 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	682/1347 (51%)	172/537 (32%)	338/546 (62%)	172/264 (65%)
Sidechain	174/1614 (11%)	0/944 (0%)	174/613 (28%)	0/57 (0%)

Continued on next page...

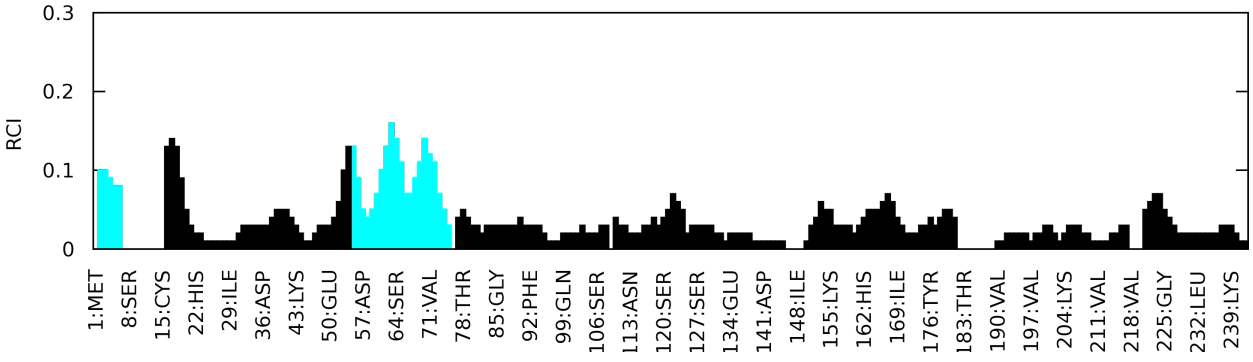
Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/182 (0%)	0/97 (0%)	0/79 (0%)	0/6 (0%)
Overall	856/3143 (27%)	172/1578 (11%)	512/1238 (41%)	172/327 (53%)

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: