W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB NMR Structure Validation Summary Report ⓘ

Jun 28, 2021 – 01:52 PM EDT

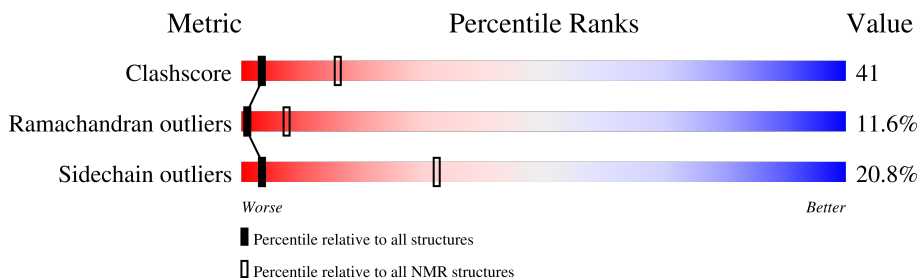|		
PDB ID	:	6XMN
Title	:	Solution NMR CXCL8-CXCR1 N-domain complex structure
Authors	:	Sepuru, K.M.; Rajarathnam, K.
Deposited on	:	2020-06-30

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 75%.

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	Percentile Ranks	Value
Clashscore		41
Ramachandran outliers		11.6%
Sidechain outliers		20.8%

Worse *Better*

▪ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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	66	39% 41% 11% 9%
2	B	29	14% · 17% 14% 52%

2   Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 14 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:7-A:66, B:14-B:27 (74)	0.32	14

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 2 single-model clusters were found.

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

3   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 1525 atoms, of which 752 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Interleukin-8.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	66	1095	345	556	95	95	4	0

- Molecule 2 is a protein called C-X-C chemokine receptor type 1.

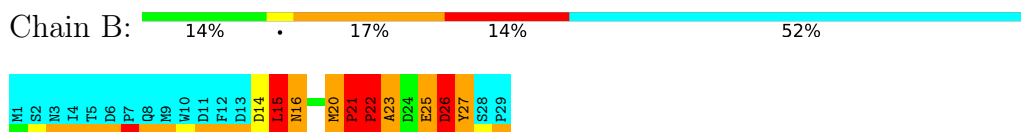| Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	29	430	146	196	33	52	3	0

4 Residue-property plots ⓘ

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.
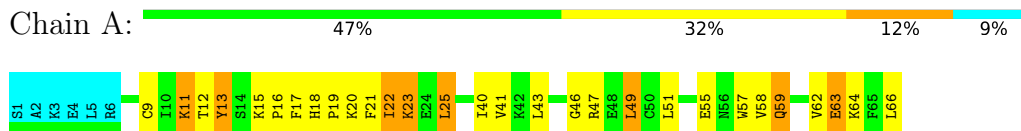
• Molecule 1: Interleukin-8

Chain A: 39%   41%   11%   9%

S1 A2 K3 E4 L5 R6 C9 I10 K11 T12 Y13 S14 K15 P16 F17 H18 P19 K20 F21 I22 K23 E24 L25 R26 H33 C34 I40 V41 K42 L43 G46 R47 E48 L49 C50 L51 E55 N56 W57 V58 Q59 R60 V61 V62 E63 K64 F65 L66

• Molecule 2: C-X-C chemokine receptor type 1

Chain B: 14%   ·   17%   14%   52%

M1 S2 N3 I4 T5 D6 P7 Q8 M9 W10 D11 F12 D13 D14 L15 N16 M20 P21 P22 A23 D24 E25 D26 Y27 S28 P29
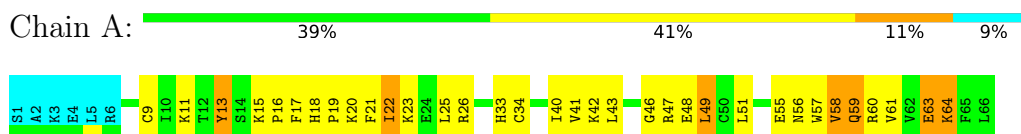
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 14. Colouring as in section 4.1 above.

• Molecule 1: Interleukin-8

Chain A: 47%   32%   12%   9%

S1 A2 K3 E4 L5 R6 C9 I10 K11 T12 Y13 S14 K15 P16 F17 H18 P19 K20 F21 I22 K23 E24 L25 I40 V41 K42 L43 G46 R47 E48 L49 C50 L51 E55 N56 W57 V58 Q59 V62 E63 K64 F65 L66

• Molecule 2: C-X-C chemokine receptor type 1

Chain B: ·   17%   14%   14%   52%

M1 S2 N3 I4 T5 D6 P7 Q8 M9 W10 D11 F12 D13 D14 L15 N16 F17 T18 G19 M20 P21 P22 A23 D24 E25 D26 Y27 S28 P29
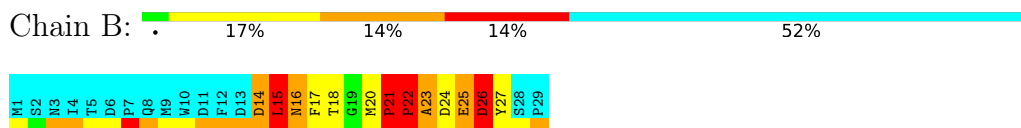
5 Refinement protocol and experimental data overview ⓘ

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

Of the 30 calculated structures, 20 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
ARIA	refinement	
ARIA	structure calculation	

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	995
Number of shifts mapped to atoms	995
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	75%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.64±0.02	0±0/502 (0.0± 0.0%)	0.70±0.03	0±0/676 (0.0± 0.0%)
2	B	0.85±0.12	0±1/114 (0.3± 0.6%)	1.66±0.09	3±1/157 (2.1± 0.9%)
All	All	0.69	7/12320 (0.1%)	0.96	65/16660 (0.4%)

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	0.1±0.2
2	B	0.0±0.0	1.2±1.0
All	All	0	25

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	27	TYR	CE2-CZ	6.43	1.47	1.38	7	4
2	B	27	TYR	CE1-CZ	-6.20	1.30	1.38	16	3

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(º)	Ideal(º)	Models	
								Worst	Total
2	B	21	PRO	N-CA-CB	-7.93	93.78	103.30	11	20
2	B	22	PRO	N-CA-CB	-7.83	93.90	103.30	8	7
2	B	22	PRO	CA-N-CD	-7.16	101.48	111.50	11	20
2	B	22	PRO	N-CD-CG	-6.02	94.17	103.20	8	7
2	B	22	PRO	CB-CA-C	5.53	125.82	112.00	2	11

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)
2	B	22	PRO	Peptide	8
2	B	21	PRO	Peptide	6
2	B	23	ALA	Peptide	5
2	B	15	LEU	Peptide	2
1	A	60	ARG	Sidechain	1

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	491	501	501	41±5
2	B	110	91	91	16±4
All	All	12020	11840	11840	974

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 41.

5 of 255 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:18:HIS:CD2	2:B:15:LEU:HG	0.90	2.01	8	14
1:A:18:HIS:CD2	1:A:19:PRO:HD2	0.83	2.08	20	18
1:A:59:GLN:O	1:A:63:GLU:HB2	0.82	1.75	5	20
1:A:13:TYR:CE2	1:A:49:LEU:HG	0.80	2.11	16	10
1:A:22:ILE:HD11	1:A:25:LEU:HD13	0.79	1.54	3	4

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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	59/66 (89%)	55±1 (94±2%)	4±1 (6±2%)	0±0 (0±0%)	54	85
2	B	14/29 (48%)	3±1 (20±4%)	3±1 (20±6%)	8±1 (60±4%)	0	0
All	All	1460/1900 (77%)	1162 (80%)	129 (9%)	169 (12%)	1	7

5 of 11 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	15	LEU	20
2	B	16	ASN	20
2	B	20	MET	20
2	B	21	PRO	20
2	B	22	PRO	20

6.3.2   Protein sidechains ⓘ

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	57/62 (92%)	47±2 (82±3%)	10±2 (18±3%)	4	37
2	B	12/27 (44%)	8±1 (67±11%)	4±1 (33±11%)	1	11
All	All	1380/1780 (78%)	1093 (79%)	287 (21%)	3	32

5 of 36 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	22	ILE	20
2	B	22	PRO	20
1	A	13	TYR	19
1	A	49	LEU	19
1	A	58	VAL	17

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 monosaccharides in this entry.

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 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 75% for the well-defined parts and 73% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_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	739
Number of shifts mapped to atoms	739
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	9

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	65	-0.10 ± 0.16	None needed (< 0.5 ppm)
$^{13}C_\beta$	61	-0.21 ± 0.15	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	60	0.21 ± 0.46	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 64%, i.e. 606 atoms were assigned a chemical shift out of a possible 946. 7 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	232/358 (65%)	116/142 (82%)	60/148 (41%)	56/68 (82%)
Sidechain	370/510 (73%)	238/302 (79%)	132/187 (71%)	0/21 (0%)

Continued on next page...

Continued from previous page...

|	**Total**	**^{1}H**	**^{13}C**	**^{15}N**
Aromatic	4/78 (5%)	4/42 (10%)	0/33 (0%)	0/3 (0%)
Overall	606/946 (64%)	358/486 (74%)	192/368 (52%)	56/92 (61%)

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	23	LYS	CE	22.66	46.00 − 37.80	-23.5
1	A	59	GLN	HB2	4.03	3.30 − 0.80	7.9
1	A	8	GLN	H	11.99	11.17 − 5.27	6.4
1	A	16	PRO	HD2	1.46	5.45 − 1.85	-6.1
1	A	16	PRO	HD3	1.46	5.52 − 1.72	-5.7
1	A	22	ILE	CG1	18.07	36.54 − 18.94	-5.5
1	A	20	LYS	HG3	-0.09	2.76 − -0.04	-5.2
1	A	18	HIS	CB	40.90	40.69 − 19.69	5.1
1	A	20	LYS	H	11.29	11.24 − 5.14	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:
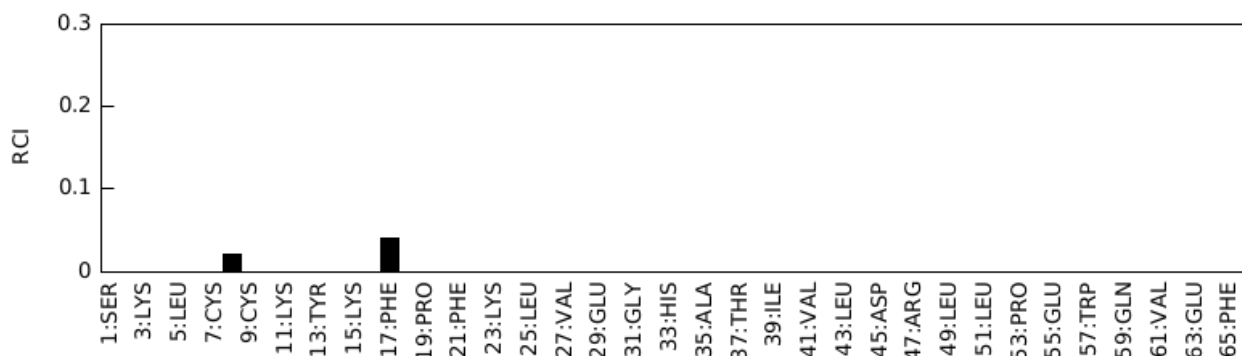
7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_2*

7.2.1 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	256
Number of shifts mapped to atoms	256
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 ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	29	-0.53 ± 0.55	None needed (imprecise)
$^{13}C_\beta$	28	0.55 ± 0.58	None needed (imprecise)
$^{13}C'$	3	—	None (insufficient data)
^{15}N	25	0.21 ± 0.64	None needed (< 0.5 ppm)

7.2.3 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 13%, i.e. 122 atoms were assigned a chemical shift out of a possible 946. 1 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

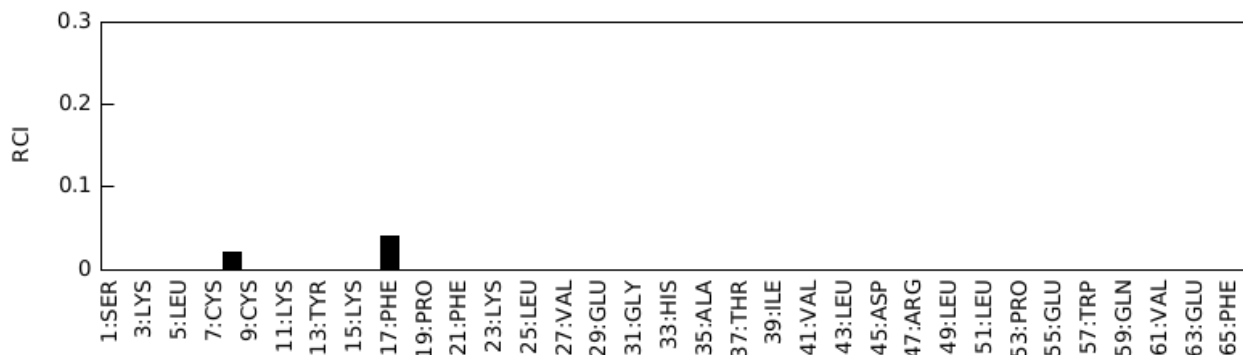|	Total	^1H	^{13}C	^{15}N
Backbone	57/358 (16%)	28/142 (20%)	16/148 (11%)	13/68 (19%)
Sidechain	65/510 (13%)	43/302 (14%)	22/187 (12%)	0/21 (0%)
Aromatic	0/78 (0%)	0/42 (0%)	0/33 (0%)	0/3 (0%)
Overall	122/946 (13%)	71/486 (15%)	38/368 (10%)	13/92 (14%)

7.2.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5   Random Coil Index (RCI) plots ⓘ

The images below report *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:

Random coil index (RCI) for chain B: