wwPDB NMR Structure Validation Summary Report ⓘ

Nov 6, 2023 – 07:01 AM EST

PDB ID	:	1ILP
Title	:	CXCR-1 N-TERMINAL PEPTIDE BOUND TO INTERLEUKIN-8
Authors	:	Skelton, N.J.; Quan, C.; Lowman, H.
Deposited on	:	1998-12-16

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

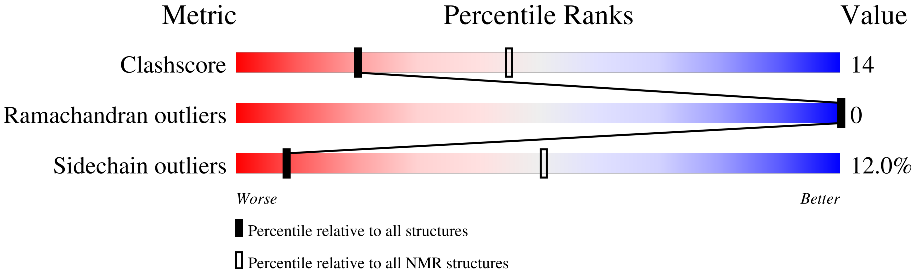
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
SOLUTION NMR

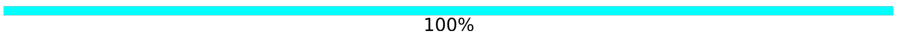
The overall completeness of chemical shifts assignment was not calculated.

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	72	
1	B	72	
2	C	19	

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:2-A:12, A:20-A:44, A:48-A:72, B:2-B:72 (132)	0.12	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 2 clusters and 1 single-model cluster was found.

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

3 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 2621 atoms, of which 1311 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Interleukin-8 (precursor).

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	71	1178	369	596	105	104	4	0
			Total	C	H	N	O	S	
1	B	71	1178	369	596	105	104	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	C	19	265	93	119	19	32	2	1

There are 7 discrepancies between the modelled and reference sequences:

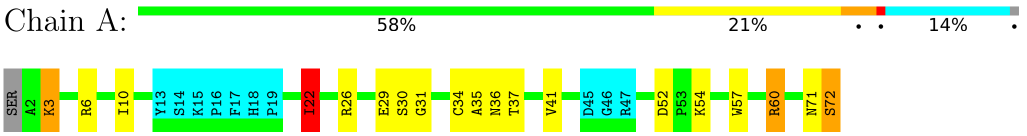
Chain	Residue	Modelled	Actual	Comment	Reference
C	0	ACE	-	acetylation	UNP P25024
C	?	-	LEU	deletion	UNP P25024
C	?	-	ASN	deletion	UNP P25024
C	?	-	PHE	deletion	UNP P25024
C	?	-	THR	deletion	UNP P25024
C	7	ACA	GLY	engineered mutation	UNP P25024
C	18	NH2	-	amidation	UNP P25024

4 Residue-property plots

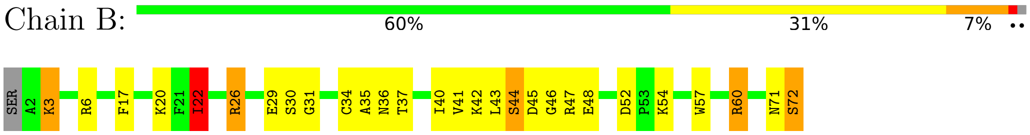
4.1 Average score per residue in the NMR ensemble

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.

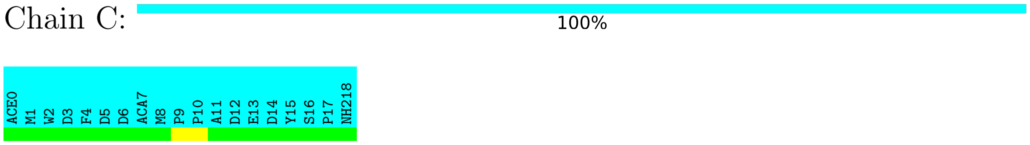
- Molecule 1: Interleukin-8 (precursor)



- Molecule 1: Interleukin-8 (precursor)



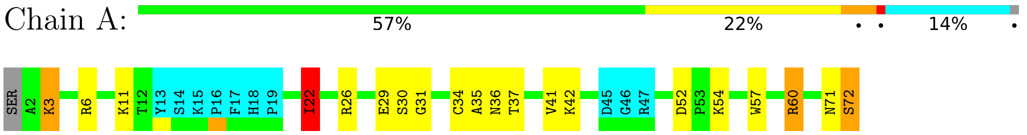
- Molecule 2: C-X-C chemokine receptor type 1



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

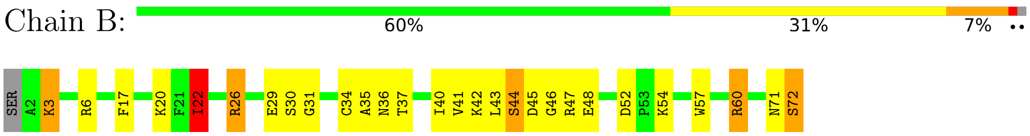
The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Interleukin-8 (precursor)



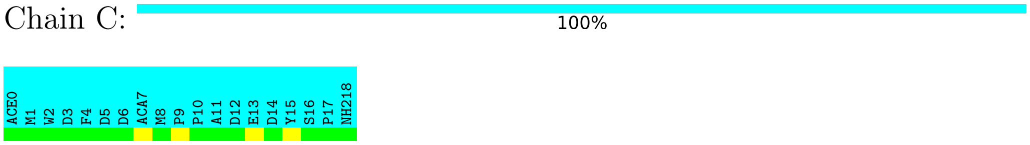
- Molecule 1: Interleukin-8 (precursor)

Chain B:



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

Chain C:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *RESTRAINED MOLECULAR DYNAMICS*.

Of the 40 calculated structures, 20 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION ENERGY*.

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

Software name	**Classification**	**Version**
Discover	refinement	
MSI DISCOVER	structure solution	DISCOVER

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, ACA, ACE

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	1.06±0.00	1±0/503 (0.2± 0.0%)	1.24±0.01	6±0/672 (0.9± 0.0%)
1	B	1.15±0.00	1±0/593 (0.2± 0.0%)	1.23±0.00	6±0/794 (0.8± 0.0%)
All	All	1.11	40/21920 (0.2%)	1.24	240/29320 (0.8%)

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	3.0±0.0
1	B	0.0±0.0	4.0±0.0
All	All	0	140

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
1	B	57	TRP	CG-CD2	-6.39	1.32	1.43	1	20
1	A	57	TRP	CG-CD2	-6.30	1.32	1.43	1	20

5 of 12 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
1	B	57	TRP	NE1-CE2-CZ2	8.15	139.37	130.40	16	20
1	A	57	TRP	NE1-CE2-CZ2	8.10	139.31	130.40	1	20
1	B	57	TRP	NE1-CE2-CD2	-6.97	100.33	107.30	1	20
1	A	57	TRP	NE1-CE2-CD2	-6.96	100.34	107.30	1	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	B	57	TRP	CG-CD1-NE1	-6.33	103.77	110.10	1	20

There are no chirality outliers.

5 of 7 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	6	ARG	Sidechain	20
1	A	26	ARG	Sidechain	20
1	A	60	ARG	Sidechain	20
1	B	6	ARG	Sidechain	20
1	B	26	ARG	Sidechain	20

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	497	519	519	11±1
1	B	582	596	596	19±0
2	C	0	0	0	0±0
All	All	21580	22300	22300	602

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 14.

5 of 33 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:17:PHE:CE1	1:B:22:ILE:HD11	0.60	2.29	1	20
1:B:3:LYS:CD	1:B:3:LYS:N	0.59	2.65	1	20
1:A:3:LYS:N	1:A:3:LYS:CD	0.58	2.65	1	20
1:B:43:LEU:HD12	1:B:43:LEU:N	0.56	2.15	1	20
1:A:22:ILE:HG13	1:A:41:VAL:HG11	0.56	1.78	16	1

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	59/72 (82%)	55±0 (93±1%)	4±0 (7±1%)	0±0 (0±0%)	100	100
1	B	69/72 (96%)	64±0 (93±0%)	5±0 (7±0%)	0±0 (0±0%)	100	100
2	C	0	-	-	-	-	
All	All	2560/3260 (79%)	2376 (93%)	184 (7%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	57/67 (85%)	50±1 (88±1%)	7±1 (12±1%)	8	51
1	B	66/67 (99%)	58±0 (88±0%)	8±0 (12±0%)	8	51
2	C	0	-	-	-	
All	All	2460/2980 (83%)	2165 (88%)	295 (12%)	8	51

5 of 19 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	3	LYS	20
1	A	22	ILE	20
1	A	37	THR	20
1	A	60	ARG	20
1	A	72	SER	20

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	ACA	C	7	2	7,7,8	0.47±0.02	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	ACA	C	7	2	6,6,8	0.61±0.09	0±0 (0±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ACA	C	7	2	-	0±0,4,5,6	-

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 ⓘ

No chemical shift data were provided