wwPDB NMR Structure Validation Summary Report

Jun 12, 2024 – 02:33 AM EDT

PDB ID	:	2ARW
BMRB ID	:	5940
Title	:	The solution structure of the membrane proximal cytokine receptor domain of the human interleukin-6 receptor
Authors	:	Hecht, O.; Dingley, A.J.; Schwantner, A.; Ozbek, S.; Rose-John, S.; Grotzinger, J.
Deposited on	:	2005-08-22

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
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

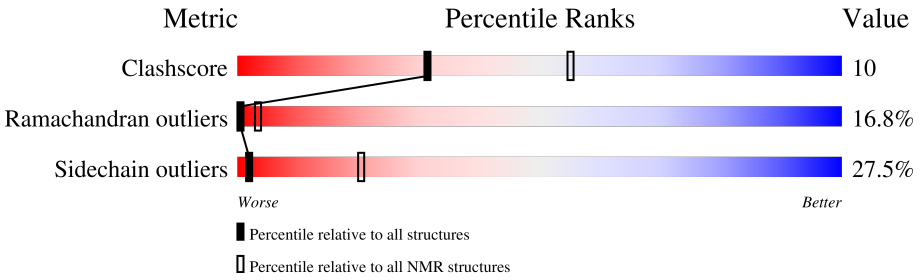
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 48%.

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	126	37% 29% 12% 5% 18%

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 is only 1 type of molecule in this entry. The entry contains 1687 atoms, of which 828 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Interleukin-6 receptor alpha chain.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	103	1687	546	828	158	151	4	0

There is a discrepancy between the modelled and reference sequences:

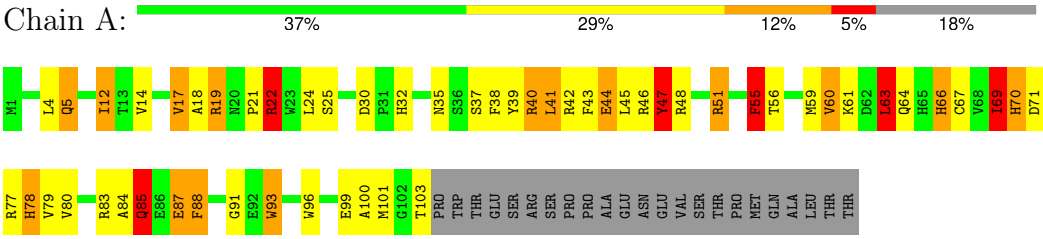
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P08887

4 Residue-property plots (i)

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-6 receptor alpha chain

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamicsu*.

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
DYANA	structure solution	1.5
DYANA	refinement	1.5

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	1
Total number of shifts	856
Number of shifts mapped to atoms	696
Number of unparsed shifts	0
Number of shifts with mapping errors	160
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	48%

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	0.94	0/889 (0.0%)	1.30	7/1212 (0.6%)
All	All	0.94	0/889 (0.0%)	1.30	7/1212 (0.6%)

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	8
All	All	0	8

There are no bond-length outliers.

5 of 7 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	47	TYR	CB-CG-CD2	-6.66	117.01	121.00
1	A	39	TYR	CB-CG-CD2	-5.88	117.47	121.00
1	A	22	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	A	48	ARG	NE-CZ-NH1	5.67	123.13	120.30
1	A	55	PHE	CB-CG-CD2	-5.10	117.23	120.80

There are no chirality outliers.

5 of 8 planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	18	ALA	Mainchain
1	A	19	ARG	Sidechain
1	A	47	TYR	Sidechain
1	A	55	PHE	Sidechain
1	A	56	THR	Mainchain

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	859	828	814	16
All	All	859	828	814	16

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.

5 of 16 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:44:GLU:HB2	1:A:83:ARG:HB2	0.55	1.77
1:A:60:VAL:HG23	1:A:67:CYS:HB2	0.54	1.79
1:A:17:VAL:HG22	1:A:25:SER:OG	0.50	2.07
1:A:69:ILE:HG21	1:A:72:ALA:HB2	0.48	1.85
1:A:79:VAL:HG23	1:A:99:GLU:HB2	0.46	1.87

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	101/126 (80%)	58 (57%)	26 (26%)	17 (17%)	0 3
All	All	101/126 (80%)	58 (57%)	26 (26%)	17 (17%)	0 3

5 of 17 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	5	GLN
1	A	12	ILE
1	A	17	VAL
1	A	30	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	37	SER

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	91/112 (81%)	66 (73%)	25 (27%)	2 20
All	All	91/112 (81%)	66 (73%)	25 (27%)	2 20

5 of 25 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	4	LEU
1	A	5	GLN
1	A	12	ILE
1	A	19	ARG
1	A	22	ARG

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 monosaccharides 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

The completeness of assignment taking into account all chemical shift lists is 48% for the well-defined parts and 48% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_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	856
Number of shifts mapped to atoms	696
Number of unparsed shifts	0
Number of shifts with mapping errors	160
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- No matching atom found in the structure. First 5 (of 160) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	104	PRO	CA	59.218	.	1
1	A	104	PRO	HA	3.972	.	1
1	A	104	PRO	CB	28.101	.	1
1	A	104	PRO	HB2	2.261	.	2
1	A	104	PRO	C	178.15	.	1
1	A	105	TRP	N	121.237	.	1
1	A	105	TRP	H	8.256	.	1
1	A	105	TRP	CA	59.117	.	1
1	A	105	TRP	HA	4.141	.	1
1	A	105	TRP	CB	29.822	.	1
1	A	105	TRP	HB3	3.085	.	2
1	A	105	TRP	HB2	2.825	.	2
1	A	105	TRP	C	175.818	.	1
1	A	106	THR	N	118.365	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	106	THR	H	7.12	.	1
1	A	106	THR	CA	59.554	.	1
1	A	106	THR	HA	3.968	.	1
1	A	106	THR	CB	71.36	.	1
1	A	106	THR	HB	3.791	.	1
1	A	106	THR	CG2	21.217	.	1
1	A	106	THR	C	172.186	.	1
1	A	107	GLU	N	121.991	.	1
1	A	107	GLU	H	7.933	.	1
1	A	107	GLU	CA	56.494	.	1
1	A	107	GLU	HA	3.572	.	1
1	A	107	GLU	CB	29.822	.	1
1	A	107	GLU	HB2	1.792	.	2
1	A	107	GLU	CG	36.067	.	1
1	A	107	GLU	C	176.541	.	1
1	A	108	SER	N	117.559	.	1
1	A	108	SER	H	8.268	.	1
1	A	108	SER	CA	58.243	.	1
1	A	108	SER	HA	4.344	.	1
1	A	108	SER	CB	63.489	.	1
1	A	108	SER	HB2	3.78	.	2
1	A	108	SER	C	174.292	.	1
1	A	109	ARG	N	123.469	.	1
1	A	109	ARG	H	8.375	.	1
1	A	109	ARG	CA	56.056	.	1
1	A	109	ARG	HA	4.309	.	1
1	A	109	ARG	CB	31.134	.	1
1	A	109	ARG	HB3	1.791	.	2
1	A	109	ARG	HB2	1.706	.	2
1	A	109	ARG	CG	27.016	.	1
1	A	109	ARG	CD	43.219	.	1
1	A	109	ARG	C	175.227	.	1
1	A	110	SER	N	116.669	.	1
1	A	110	SER	H	8.191	.	1
1	A	110	SER	CA	54.745	.	1
1	A	110	SER	CB	64.364	.	1
1	A	112	PRO	CA	63.074	.	1
1	A	112	PRO	HA	4.397	.	1
1	A	112	PRO	CB	31.921	.	1
1	A	112	PRO	HB3	2.314	.	2
1	A	112	PRO	HB2	1.924	.	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	112	PRO	CG	27.284	.	1
1	A	112	PRO	CD	50.408	.	1
1	A	112	PRO	C	176.869	.	1
1	A	113	ALA	N	124.019	.	1
1	A	113	ALA	H	8.45	.	1
1	A	113	ALA	CA	52.121	.	1
1	A	113	ALA	HA	4.316	.	1
1	A	113	ALA	CB	19.329	.	1
1	A	113	ALA	HB1	1.409	.	1
1	A	113	ALA	HB2	1.409	.	1
1	A	113	ALA	HB3	1.409	.	1
1	A	114	GLU	N	119.962	.	1
1	A	114	GLU	H	8.399	.	1
1	A	114	GLU	CA	56.056	.	1
1	A	114	GLU	HA	4.3	.	1
1	A	114	GLU	CB	30.259	.	1
1	A	114	GLU	HB3	2.096	.	2
1	A	114	GLU	HB2	1.923	.	2
1	A	114	GLU	CG	35.536	.	1
1	A	115	ASN	N	124.027	.	1
1	A	115	ASN	H	8.038	.	1
1	A	115	ASN	CA	54.745	.	1
1	A	115	ASN	HA	4.363	.	1
1	A	115	ASN	CB	40.316	.	1
1	A	115	ASN	HB3	2.974	.	2
1	A	116	GLU	N	121.351	.	1
1	A	116	GLU	H	8.851	.	1
1	A	116	GLU	CA	56.056	.	1
1	A	116	GLU	HA	4.436	.	1
1	A	116	GLU	CB	30.259	.	1
1	A	116	GLU	HB3	2.092	.	2
1	A	116	GLU	HB2	1.919	.	2
1	A	116	GLU	CG	35.458	.	1
1	A	116	GLU	C	176.502	.	1
1	A	117	VAL	N	121.511	.	1
1	A	117	VAL	H	8.362	.	1
1	A	117	VAL	CA	62.178	.	1
1	A	117	VAL	HA	4.218	.	1
1	A	117	VAL	CB	32.883	.	1
1	A	117	VAL	HB	2.092	.	1
1	A	117	VAL	CG1	20.619	.	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	117	VAL	C	176.089	.	1
1	A	118	SER	N	119.98	.	1
1	A	118	SER	H	8.565	.	1
1	A	118	SER	CA	58.243	.	1
1	A	118	SER	HA	4.533	.	1
1	A	118	SER	CB	63.927	.	1
1	A	118	SER	HB3	3.664	.	2
1	A	118	SER	HB2	3.839	.	2
1	A	118	SER	C	176.332	.	1
1	A	119	THR	N	118.516	.	1
1	A	119	THR	H	8.287	.	1
1	A	119	THR	CA	59.554	.	1
1	A	119	THR	CB	69.611	.	1
1	A	120	PRO	CA	63.14	.	1
1	A	120	PRO	HA	4.4	.	1
1	A	120	PRO	CB	32.322	.	1
1	A	120	PRO	HB3	2.317	.	2
1	A	120	PRO	HB2	1.883	.	2
1	A	120	PRO	CG	27.356	.	1
1	A	120	PRO	CD	51.041	.	1
1	A	120	PRO	C	177.986	.	1
1	A	121	MET	N	120.659	.	1
1	A	121	MET	H	8.491	.	1
1	A	121	MET	CA	55.619	.	1
1	A	121	MET	HA	4.444	.	1
1	A	121	MET	CB	32.883	.	1
1	A	121	MET	HB3	2.013	.	2
1	A	121	MET	HB2	2.1	.	2
1	A	121	MET	CG	31.956	.	1
1	A	122	GLN	N	125.9	.	1
1	A	122	GLN	H	7.993	.	1
1	A	122	GLN	CA	57.368	.	1
1	A	122	GLN	HA	4.047	.	1
1	A	122	GLN	CB	30.259	.	1
1	A	122	GLN	HB2	2.137	.	2
1	A	122	GLN	CG	32.783	.	1
1	A	122	GLN	C	172.745	.	1
1	A	123	ALA	N	127.085	.	1
1	A	123	ALA	H	8.799	.	1
1	A	123	ALA	CA	52.347	.	1
1	A	123	ALA	HA	4.397	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	123	ALA	CB	19.329	.	1
1	A	123	ALA	HB1	1.402	.	1
1	A	123	ALA	HB2	1.402	.	1
1	A	123	ALA	HB3	1.402	.	1
1	A	124	LEU	N	122.148	.	1
1	A	124	LEU	H	8.518	.	1
1	A	124	LEU	CA	54.444	.	1
1	A	124	LEU	HA	4.147	.	1
1	A	124	LEU	CB	42.597	.	1
1	A	124	LEU	HB2	1.76	.	2
1	A	124	LEU	CG	26.632	.	1
1	A	124	LEU	CD1	23.872	.	2
1	A	125	THR	N	118.072	.	1
1	A	125	THR	H	8.698	.	1
1	A	125	THR	CA	61.971	.	1
1	A	125	THR	HA	4.49	.	1
1	A	125	THR	CB	69.681	.	1
1	A	125	THR	HB	4.23	.	1
1	A	125	THR	CG2	21.492	.	1
1	A	126	THR	N	121.944	.	1
1	A	126	THR	H	7.978	.	1
1	A	126	THR	CA	63.052	.	1
1	A	126	THR	CB	70.485	.	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	112	0.00 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	105	0.16 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	84	-0.07 ± 0.11	None needed (< 0.5 ppm)
^{15}N	103	0.73 ± 0.32	Should be applied

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 48%, i.e. 696 atoms were assigned a chemical shift out of a possible 1459. 0 out of 14 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	415/508 (82%)	169/205 (82%)	162/206 (79%)	84/97 (87%)
Sidechain	281/769 (37%)	122/498 (24%)	159/233 (68%)	0/38 (0%)
Aromatic	0/182 (0%)	0/90 (0%)	0/75 (0%)	0/17 (0%)
Overall	696/1459 (48%)	291/793 (37%)	321/514 (62%)	84/152 (55%)

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: