Full wwPDB NMR Structure Validation Report (i)

Oct 11, 2021 – 07:38 PM EDT

PDB ID	:	2HDM
Title	:	Solution structure of V21C/V59C Lymphotactin/XCL1
Authors	:	Volkman, B.F.; Tuinstra, R.L.; Peterson, F.C.; Elgin, E.S.
Deposited on	:	2006-06-20

This is a Full wwPDB NMR Structure Validation 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 following versions of software and data (see references (i)) were used in the production of this report:

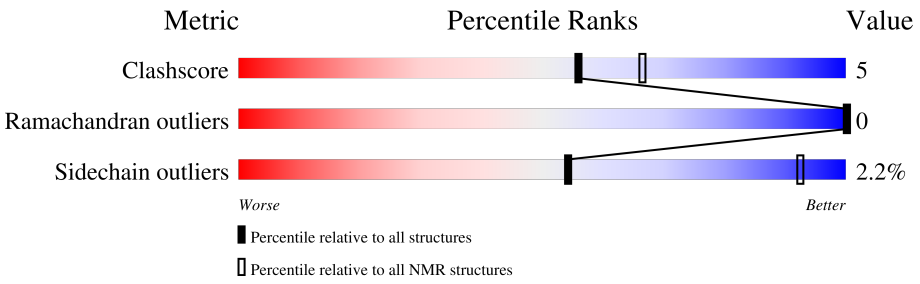
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 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	92	46% 7% 28% 20%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:14-A:30, A:35-A:65 (48)	0.28	19

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1185 atoms, of which 603 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Lymphotactin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	74	1185	354	603	113	109	6	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	CYS	VAL	engineered mutation	UNP P47992
A	59	CYS	VAL	engineered mutation	UNP P47992

4 Residue-property plots (i)

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: Lymphotactin

Chain A: 46% 7% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 R18 I24 T30 E31 G32 S33 L34 R35 F39 W55 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Lymphotactin

Chain A: 46% 7% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 R18 I24 E31 G32 S33 L34 F39 K46 W55 S62 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.2 Score per residue for model 2

• Molecule 1: Lymphotactin

Chain A: 46% 7% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 R18 I24 Y27 E31 G32 S33 L34 F39 C48 W55 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.3 Score per residue for model 3

- Molecule 1: Lymphotactin

Chain A: 43% 9% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 L14 T15 T16 Q17 R18 I24 E31 G32 S33 L34 F39 R43 W55 S62 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR

GLY

4.2.4 Score per residue for model 4

- Molecule 1: Lymphotactin

Chain A: 42% 10% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 L14 T15 T16 Q17 I24 Y27 T30 E31 G32 S33 L34 R35 F39 C59 R65 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR

GLY

4.2.5 Score per residue for model 5

- Molecule 1: Lymphotactin

Chain A: 41% 10% • 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 R18 I24 Y27 E31 G32 S33 L34 F39 K46 V47 C48 W55 D58 R61 S62 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA

VAL THR LEU THR GLY

4.2.6 Score per residue for model 6

- Molecule 1: Lymphotactin

Chain A: 42% 10% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 L14 R18 I24 K25 T26 Y27 T30 E31 G32 S33 L34 R35 F39 W55 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.7 Score per residue for model 7

- Molecule 1: Lymphotactin

Chain A: 47% 5% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 I24 T30 E31 G32 S33 L34 R35 F39 S62 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.8 Score per residue for model 8

- Molecule 1: Lymphotactin

Chain A: 45% 8% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 I24 K25 T30 E31 G32 S33 L34 R35 F39 A53 R57 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.9 Score per residue for model 9

- Molecule 1: Lymphotactin

Chain A: 38% 13% • 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 Q17 R18 I24 Y27 T30 E31 G32 S33 L34 R35 I38 F39 D50 P51 Q52 W55 S62 R65 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN

THR ALA VAL THR LEU THR GLY

4.2.10 Score per residue for model 10

- Molecule 1: Lymphotactin

Chain A: 46% 7% 28% 20%

G2 S3 E4 V5 S6 D7 K8 R9 T10 C11 V12 S13 R18 I24 K25 E31 G32 S33 L34 F39 W55 S62 K66 S67 N68 T69 R70 N71 N72 M73 I74 Q75 THR LYS PRO THR GLY THR GLN GLN SER THR ASN THR ALA VAL THR LEU THR GLY

4.2.11 Score per residue for model 11

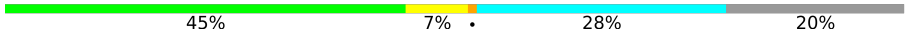
- Molecule 1: Lymphotactin

Chain A: 39% 13% 28% 20%

4.2.12 Score per residue for model 12

- Molecule 1: Lymphotactin

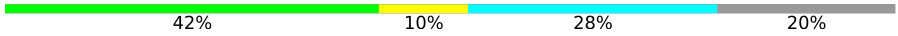
Chain A:

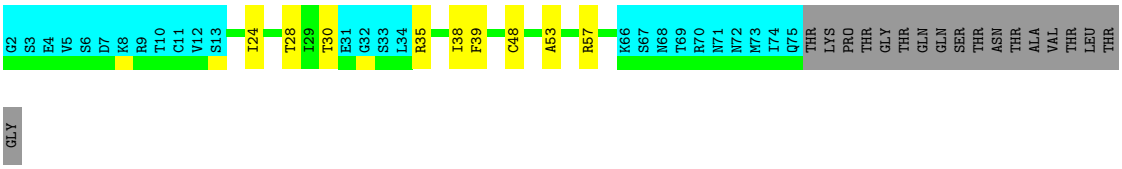


4.2.13 Score per residue for model 13

- Molecule 1: Lymphotactin

Chain A:

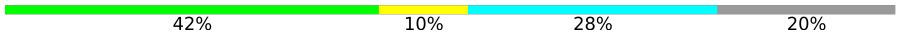


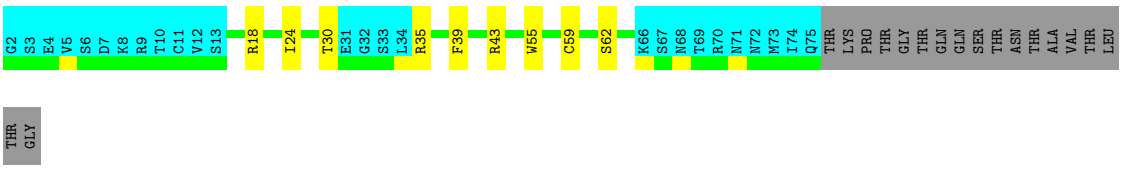


4.2.14 Score per residue for model 14

- Molecule 1: Lymphotactin

Chain A:

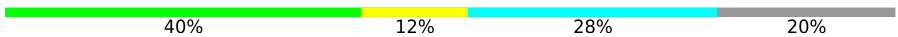




4.2.15 Score per residue for model 15

- Molecule 1: Lymphotactin

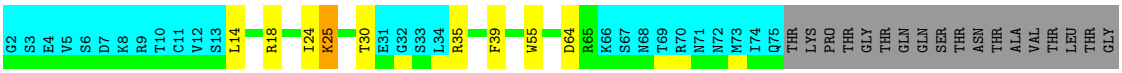
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Lymphotactin

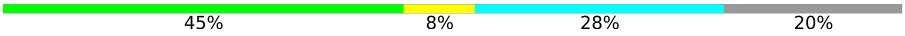
Chain A:

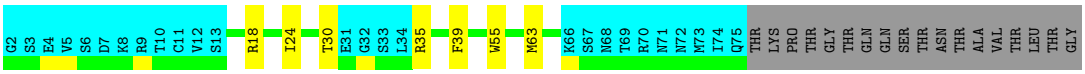


4.2.17 Score per residue for model 17

- Molecule 1: Lymphotactin

Chain A:

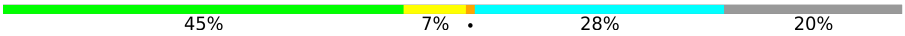


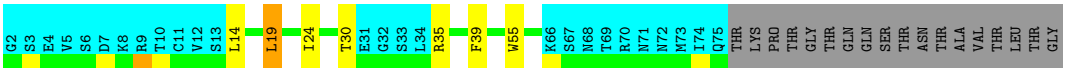


4.2.18 Score per residue for model 18

- Molecule 1: Lymphotactin

Chain A:

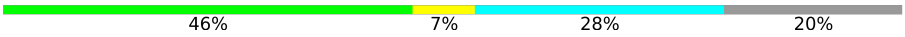


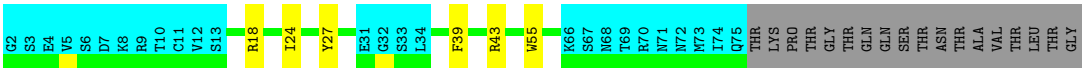


4.2.19 Score per residue for model 19 (medoid)

- Molecule 1: Lymphotactin

Chain A:

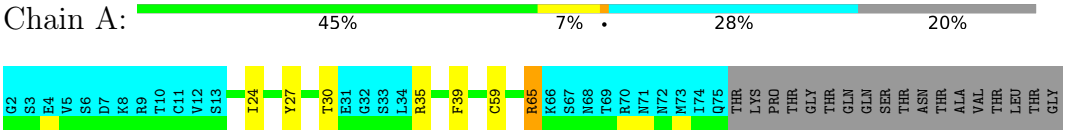




4.2.20 Score per residue for model 20

- Molecule 1: Lymphotactin

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *AUTOMATED METHODS WERE USED FOR BACKBONE CHEMICAL SHIFT ASSIGNMENT AND ITERATIVE NOE REFINEMENT.*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
CYANA	structure solution	2.1
XPLOR-NIH	refinement	2.9.3

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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	0.85±0.03	0±0/393 (0.0± 0.0%)	0.64±0.03	0±0/532 (0.0± 0.0%)
All	All	0.85	0/7860 (0.0%)	0.64	1/10640 (0.0%)

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	46	LYS	N-CA-CB	-5.05	101.50	110.60	5	1

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	387	411	411	4±1
All	All	7740	8220	8220	83

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 5.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:19:LEU:HG	1:A:55:TRP:CZ3	0.64	2.28	15	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:18:ARG:HG2	1:A:55:TRP:CD1	0.59	2.33	19	12
1:A:30:THR:O	1:A:35:ARG:HA	0.58	1.98	8	13
1:A:24:ILE:HD12	1:A:39:PHE:CD1	0.57	2.34	20	19
1:A:43:ARG:HD3	1:A:43:ARG:O	0.55	2.02	12	2
1:A:24:ILE:HD13	1:A:39:PHE:CD1	0.54	2.38	14	1
1:A:25:LYS:O	1:A:25:LYS:HE3	0.52	2.04	16	1
1:A:50:ASP:OD1	1:A:52:GLN:HG2	0.51	2.06	9	1
1:A:24:ILE:HD12	1:A:59:CYS:SG	0.51	2.46	14	1
1:A:24:ILE:HG13	1:A:59:CYS:SG	0.50	2.47	20	3
1:A:38:ILE:HG12	1:A:48:CYS:SG	0.50	2.46	13	1
1:A:27:TYR:HB3	1:A:39:PHE:CD1	0.49	2.42	11	6
1:A:27:TYR:HB3	1:A:39:PHE:CE1	0.47	2.43	2	9
1:A:53:ALA:O	1:A:57:ARG:HG3	0.46	2.11	13	2
1:A:15:THR:OG1	1:A:17:GLN:HG2	0.45	2.11	12	3
1:A:28:THR:OG1	1:A:38:ILE:HB	0.45	2.12	13	1
1:A:27:TYR:HA	1:A:38:ILE:O	0.44	2.13	9	1
1:A:57:ARG:O	1:A:61:ARG:HG2	0.44	2.12	11	1
1:A:14:LEU:HD12	1:A:14:LEU:N	0.42	2.30	6	2
1:A:20:PRO:O	1:A:23:ARG:HG2	0.42	2.15	15	1
1:A:19:LEU:HD23	1:A:55:TRP:CZ3	0.41	2.49	18	1
1:A:58:ASP:HA	1:A:61:ARG:HG3	0.40	1.93	5	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	48/92 (52%)	48±1 (99±1%)	0±1 (1±1%)	0±0 (0±0%)	100	100
All	All	960/1840 (52%)	951 (99%)	9 (1%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	44/83 (53%)	43±1 (98±2%)	1±1 (2±2%)	54 92
All	All	880/1660 (53%)	861 (98%)	19 (2%)	54 92

All 10 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	43	ARG	4
1	A	25	LYS	4
1	A	65	ARG	3
1	A	46	LYS	2
1	A	61	ARG	1
1	A	17	GLN	1
1	A	35	ARG	1
1	A	42	LYS	1
1	A	14	LEU	1
1	A	19	LEU	1

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 (i)

No chemical shift data were provided