Understanding the Similarities and Differences in Various Interleukins by Structure and Sequence Mapping

Sukhpal

Ms14094

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali

August 2020

Certification of Examination

This to certify that the dissertation titled **"Understanding the Similarities and Differences in Various Interleukins by Structure and Sequence Mapping"** submitted by **Mr. Sukhpal** (Reg. No. MS14094) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Sugumar Venkataramani Dr. Sabyasachi Rakshit Dr. Monika Sharma

(Supervisor)

Dated: August 10, 2020

Declaration

The work presented in the dissertation has been carried by me under the guidance of Dr. Monika Sharma at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, or diploma, or a fellowship to any other University or Institute. Whenever contribution of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in bibliography.

Sukhpal

(Candidate)

Dated: August 10, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge

Dr. Monika Sharma

(Supervisor)

Acknowledgement

First and foremost I thank God for making me able and constantly push me through the ups and downs.

I express my gratitude to Dr. Monika Sharma for giving me opportunity to work with her. This thesis would have not been possible without her immense guidance. Only with her constant support I have been able to complete my thesis. I am really grateful to have her as my supervisor.

I thank my committee members, Dr. Sugumar Vekataramani and Dr. Sabyasachi Rakshit for investing their precious time in me.

I thank Mr. Ravi Kumar, Mr. Naman Kumar Bharti, Mr. Prateek Meena, Mr. Milind Kale, Mr. Asish , Mr. Ravineet Yadav and Mr. Mohit Sippy for helping me throughout my project.

I thank my Parents, Mr. Krishan Lal and Mrs. Kiran Bala for their never ending support throughout my life.

List of Figures

3.1.3.5: RMSD= 12.043 Å

3.1.3.6: RMSD= 12.168 Å

3.1.3.7: RMSD= 12.916 Å

3.1.3.8: RMSD= 14.288 Å

3.1.3.9: RMSD= 14.327 Å

3.1.3.10: RMSD= 14.339 Å

3.1.3.11:RMSD= 15.859 Å

3.1.3.12:RMSD= 17.617 Å

List of Tables

Notations

IL- Interleukin

Th2- Type-2 helper cell

DCs- Dendritic Cells

IFN- Interferon

ILR- Interleukin receptor

ILRa- Interleukin receptor antagonist

CD4- Cluster of Differentiation 4

RMSD- Root Mean Square Deviation

γc- gamma c

PyMOL- Python Molecule

MATLAB- Matrix Laboratory

COVID-19- Coronavirus Disease 2019

NKT- Natural Killer T

PDB- Protein Data Bank

FASTA- Fast All

SW- Smith-Waterman

Treg- T regulatory

Abstract

Cytokines are small proteins with low molecular weights having a complex regulatory influence on inflammation and immune responses. It has been reported in previous experimental studies that development of immune and inflammatory responses involve hemaetopoetic cells, lymphoid cell and various pro-inflammatory and anti-inflammatory cells, and cytokines mediate the complex interactions of these cells. Interleukins are a type of cytokine that play essential roles in the activation and differentiation of immune cells, as well as proliferation, maturation, migration, and adhesion. Many Interleukins are observed to have similar signalling pathways but exert different functions or having different origin and signalling but are observed to show common functions. Several Interleukins are observed to show receptor pleiotropy i.e. same receptor complex is shared among more than one Interleukin. Here we have analysed the structural aspects and sequence homology possessed by various Interleukins, so as to get some idea about the cause of the similarity or difference in their respective functions and the Pleiotropic behaviour of the Interleukins. The reported experimental structures of interleukins were downloaded from PBD website and their pairwise sequence and structural alignments were carried out.

Contents

Chapter 1

Introduction

1.1 Cytokines

Cytokines are small proteins that are important in cell signaling and work as immunomodulating agents(1). The term cytokine is derived from a combination of two Greek words, 'Cyto' meaning cell and 'Kinos' meaning movement. Cytokines include proteins and peptides that act in nano to picomolar concentrations as humoral regulators and modulators for functional activities of cells and tissues. Cytokines act as mediators in the immune system of a body. Cytokines are basically regulatory or signalling molecules which play an important role in regulation of various immune responses including the intensity and duration of the regulated responses by the immune cells(2). Cytokines are produced by various kinds of cells, comprising some immune cells like macrophages, lymphocytes and mast cells and also endothelial cells, fibroblasts and stromal cells. Cytokines act through receptors and modulate balance between humoral and cell based immune responses. Cytokines act on a large number of target cells than hormones(2).

Cytokines play a very different role than hormones in the body. The major feature distinguishing cytokines from hormones is the fact that cytokines are not produced by specialized cells organized in specialized glands. Cytokines are secreted proteins which means that their expression sites does not predict where they exert their biological function(1). The immunomodulating effects of cytokines are also systematic and not local.

Cytokines are peptides and are unable to cross the lipid bilayer of the cells. They bind to the receptors present on the surface of the cells and pass on the desired signal to the cell. Chemokines, interleukins, interferons, lymphokines and tumour necrosis factor are various kinds of cytokines named after their specific functions(3). There are three modes of action of cytokines, namely- 1) Autocrine- The cytokines bind to the same cell by which it is produced. 2) Paracrine- The cytokines bind to one or more nearby cells. 3) Endocrine- The cytokines generated bind to one or more distant cells; based on different types of responses. Although Cytokines are crucial for countering infections and in other immune responses but cytokines may cause some adverse effects in some cases. Over secretion of cytokines triggers a cytokine storm syndrome. A severe cytokine storm could be life threatening and can lead to multiple organ failure. Some deaths of COVID-19 are also attributed to cytokine release storms.

1.2 Pleiotropy and Redundancy in Cytokines

Pleiotropy is the ability of different cell types to secrete same cytokine or of single cytokine to act on different cell types whereas **Redundancy** is the ability of multiple cytokines to exert similar functions.There are many different mechanisms which can explain the pleiotropic and the overlapping behaviour of different cytokines. Pleiotropic actions can be explained by a cytokine having an ability to activate multiple signalling pathways in which the different signalling pathways result in different functions(4). Cytokines exerting overlapping actions have similar motifs in their receptor complexes which mediate coupling to the same pathways. Cytokine pleiotropy and redundancy can be somewhat clarified, separately, by the ability of specific cytokines to signal by more than one sort of receptor complex and by the sharing of an individual receptor motif by more than one cytokine.

1.3 Interleukins

Interleukins are a group of cytokines that regulate and mediate the communication between the immune cells. The key function of interleukins in to modulate growth, activation and differentiation during the immune responses. Interleukins help in the development and differentiation of T, B, and hematopoietic cells and play major role in both innate and adaptive immunity(6). 'Inter' meaning means of communication and 'leukin' abbreviating leukocytes, interleukins were first observed in leukocytes, later found to be produced by a wide variety of cells. The interleukins also play essential roles in proliferation, maturation, migration, and adhesion of the cells(5). There are more than 60 cytokines that are listed as Interleukins from the initial discovery of ILs. Human genome itself consists of more than 50 ILs. Majority of interleukins are produced by CD4 T lymphocytes, and macrophages, monocytes and endothelial cells. Some interleukins have been divided into various groups or families considering their similarities in ligand chains and receptor complexes. Interleukin-1 family consists of 11 members including 7 proinflammatory agonists and 4 putative antagonists. Common gamma chain family consists of 6 members, and are grouped together due to their ability to bind to γc receptor(CD132). IL10 subfamily consists of 8 members and are characterised due to their receptor sharing among themselves. IL12 family and IL17 family both have 4 members and share receptor and ligand chain among their family.

1.3.1 IL1 Family

Interleukin-1 family consists of 11 Interleukins, 7 proinflammatory agonists (IL-1α, IL-1β, IL-18, IL-33, IL-36α, IL-36β, and IL-36γ) and 4 putative antagonists (IL-1Ra, IL-36Ra, IL-37, and IL-38) which exert anti inflammatory actions(3).

Interleukin-1 was first discovered as fever inducing protein and was called human leukocytic pyrogen. It comprises of two major proteins IL-1α and IL-1β. Initially they were thought to have similar biological properties, but later on fundamental differences in their localisation, maturation and secretion were observed. IL-1α and IL-1β exert similar effects by binding to the IL-1 type I receptor. They can also bind to the IL-1 type II receptor, which acts as a decoy receptor and is not involved in signal transduction(3).

IL18 is expressed by a range of inflammatory cells. IL18 combining with IL12 induces high levels of production by T cells IFN-γ production by T cells. IL18 is majorly secreted by macrophages, DC's and epithelial cells and is associated to the diseases: psoriasis, crohn disease, bacterial and viral infections and more. IL33 is a member of IL1 family, its receptor ST2 and induces type 2 responses in T cells and ILC's. It is secreted by Necrotic cells, nuocytes, and fibroblasts and majorly targets cells basophils, eosinophils and mast cells. It has a β-Trefoil fold structure and is associated to asthma and autoimmune and cardiovascular diseases.IL36 is another proinflammatory member of IL1 family and act as a mediator for innate and adaptive immune responses.(7) IL36Ra inhibits IL36 and is associated with psoriasis and asthma.

IL37 is found in monocytes and tonsil plasma cells and target DCs.it inhibits the IL18 activity and is associated with rheumatoid arthritis and atopic dermatitis. IL38 is similar to IL1Ra and bind to the IL1R type 1 receptor. IL-38 is highly homologous to IL-36Ra and IL-1Ra, suggesting that it might act as an IL-1 family antagonist. IL38 is associated to Systemic lupus erythematosus disease(3).

1.3.2 Common Gamma Chain Cytokine Family

Common gamma chain family includes IL2, IL4, IL7, IL9, IL15 and IL21 and are grouped together due to their binding to the γc receptor. They act mainly as growth and proliferation factors for progenitors and mature cells and also have roles in lineage-specific cell differentiation(3).

IL2 is a monomer of around 15.5kDa and binds to IL2R. It is produced by CD4 + and CD8 + activated T cells, DCs, NK and NKT cells, mast cells, and ILCs and functions for proliferation of effector T and B cells, development of Treg cells, differentiation and proliferation of NK cells; **IL4** is also a monomer weighing 15kDa and binds to IL-4R type I, IL-4R type II receptors. It is produced by TH2 cells, basophils, eosinophils, mast cells, NKT cells and γ/δ T cells; **IL7** binds to IL7R, its receptor complex consists of IL-7Rα (CD127) and γc (CD132) chains. Major function of IL7 is Proliferation of pre-B and pro-B cells in mice; **IL9**: TH2 cells and ILC2s are the main sources of IL-9 production. It functions as T and mast cell growth factor, inhibition of TH1- cytokines and proliferation of CD8 + T cells and mast cells; **IL15** and IL2 share structural homology and has ability to induce T-cells, NK cells and ILC proliferation like IL2(10). It is associated to Rheumatoid arthritis, psoriasis, diabetes mellitus diseases; **IL21** is four helix bundle monomer induced by T cells and NKT cells and is acts in proliferation, differentiation and survival of B cells.

1.3.3 IL10 Family

IL10 family comprises of IL10, IL19, IL20, IL22, IL24, IL26, IL28, and IL29. Their functions are characterised by their binding to the receptors IL-10R2, IL-20R1, IL-20R2, IL-22R1, and IL-28R1 and distinct expression of these receptors in the immune system(3).

IL10 is a homodimer and has an immunosuppressive effect through APCs or direct effect to T cells. It is induced mainly by T cells , B cells and monocytes. It binds to IL-10R1/IL-10R2 complex; IL19 and IL20 both have common receptor complex IL-20R1/IL-20R2, IL20 also binds to IL-22R1/IL-20R2 complex. Both IL19 an IL20 are mainly produced by Monocytes, keratinocytes, endothelial and epithelial cells. They have different functions as IL19 helps in induction of TH2 cytokines, and IL20 plays a role in skin biology; IL22 is a monomer and consists of Six anti-parallel α-helices. It is expressed by activated T cells and NK cells and plays role in Pathogen defence, wound healing and tissue reorganization; IL24 binds to complexes comprising IL-22R1 and IL-20R2 or IL-20R1 and IL-20R2(25). Generated by melanocytes and T cells it acts on Cancer cells and plays major role in tumour suppression; IL26 expression is limited to memory T cells, NK cells and TH17 cells. Its receptor consists of IL10R2 chain and IL20R chain. It functions in activating and regulating epithelial cells; IL-28A/B and IL29 (IFN-λ) has a receptor complex comprising of IL-28R1 and IL-10R2 chains. Their basic function is the downregulation of TH2 cells and upregulation of TH1 responses.

1.3.4 IL12 Family

IL-12, IL-23, IL-27, and IL-35 share receptor and ligand chains. They have different functions because of their expression on different cell types and combinations of different receptor chains(3).

IL12 is a heterodimer of two chains(p35 and p40) produced by monocytes and macrophages. It basically functions in development and maintenance of TH1 cells and activation of NK cells; IL23 includes the IL12p40 subunit in its sequence along with a distinct IL23p19 subunit. Produced by phagocytic cells and macrophages, it simulates the production of proinflammatory IL17 and enhancement of T cell proliferation; IL27 consists of an IL12p35 like subunit and an IL12p40 like subunit. Produced by activated DCs, macrophages, epithelial cells, it promotes TH1 cell differentiation and TH17 cell inhibition; IL35 is stimulated by Treg cells, monocytes, vascular endothelial cells and works in reducing of effector T-cell proliferation and increasing IL-10 production and Treg cell proliferation.

1.3.5 Other Interleukins

IL3 and IL5 share a common receptor subunit β chain(CD131) resulting in their similar functions.Il3 is sourced to T cells, macrophages, NK cells, mast cells, eosinophils, stromal cells and act as a hematopoietic growth factor IL-3 and TNF-α promote proliferation of CD34+ progenitor cells(18).
IL6 family includes IL6, leukemia inhibitor factor, ciliary neurotrophic factor, and oncostatin M. IL-6 is a pleiotropic cytokine involved in regulation of immune responses, acute-phase responses, hematopoiesis, and inflammation. IL6 Induces acute-phase proteins

in hepatocytes and promotes T-cell and B-cell production, differentiation, activation, and survival.
IL11 is a monomer and bind to a heterodimer receptor comprising IL-11Rα and gp130 subunits. It is expressed in bone marrow cells, fibroblasts, epithelial cells and endothelial cells and works as a growth factor for myeoloid, erythroid, megakaryocyte progenitors and plasmacytoma cells(23).
IL14 bind to its receptor IL4R and has main function of proliferation to activated B cells. It is observed to be expressed especially on germinal B-cells and human tonsil B-cells.

1.4 Software Used

1.4.4 PyMOL

PyMOL is a open-source software available for model visualization and used in structural biology. The '*Py'* part of the software name refers to the Python programming language. PyMOL is also used for editing pdb files i.e. adding or removing fragments and residues, morphing, sculpting and aligning various pdb files(26).

Here we used PyMOL version 2.4.0 for our analysis.

1.4.5 MATLAB

MATLAB is a high-performance language for technical computing developed by mathworks. It includes computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. It is used in modelling, simulation, alignment, analysis, exploration and visualisation.

We used MATLAB v9.4 for carrying out sequence homology.

Chapter 2

Methodology

2.1 Protein Data Bank(pdb) File

The Protein Data Bank file provides all the detailed information about the structure of the molecule associated with it. Various experimental techniques like NMR spectroscopy and X-ray diffraction etc, are used to determine the various aspects of the structure. A large set of molecules with their associated pdb ID can be found on the server https://www.rcsb.org/. A pdb file is nothing but a text file which has information about every atom of the molecule, its coordinates, what residue it is present in, what heterogeneous atoms are present, which atoms are missing etc. A pdb file can easily be accessed from the pdb server to work with. We took pdb files of all observed Interleukins from https://www.rcsb.org/ listed below

Table 2.1.1 PDB IDs for Interleukins observed

INTERLEUKIN	PDB ID
IL1α	2ILA
IL1β	91LB
IL2	1M47
IL3	5UWC
IL4	2INT
IL5	3VA2
IL6	1ALU
IL7	3DI3
IL10	1INR
IL11	6O4O
IL12	6UIB
IL13	3BPO
IL15	2Z3Q

IL17A	4HR9
IL18	3WO2
IL19	1N1F
IL20	4DOH
IL21	3TGX
IL22	1M4R
IL23	4GRW
IL24	6GG1
IL29	3OG4
IL33	4KC3
IL37	5HN1
IL38	5BOW

```
ATOM    989  CB  LEU A 132      -4.985  37.258   4.296  1.00 28.00           C
ATOM    990  CG  LEU A 132      -3.461  37.383   4.327  1.00 28.51           C
ATOM    991  CD1 LEU A 132      -2.811  36.012   4.218  1.00 27.96           C
ATOM    992  CD2 LEU A 132      -2.966  38.308   3.221  1.00 29.16           C
ATOM    993  N   THR A 133      -7.971  38.313   5.747  1.00 32.21           N
ATOM    994  CA  THR A 133      -9.396  37.992   5.800  1.00 36.42           C
ATOM    995  C   THR A 133     -10.234  38.798   4.807  1.00 37.78           C
ATOM    996  O   THR A 133      -9.859  39.899   4.405  1.00 38.67           O
ATOM    997  CB  THR A 133      -9.940  38.181   7.226  1.00 39.03           C
ATOM    998  OG1 THR A 133      -9.449  37.131   8.068  1.00 40.35           O
ATOM    999  CG2 THR A 133     -11.446  37.970   7.247  1.00 42.05           C
TER    1000       THR A 133
HETATM 1001  S   SO4 A 134      12.936  43.311   8.182  1.00 61.45           S
HETATM 1002  O1  SO4 A 134      11.669  43.174   7.469  1.00 61.47           O
HETATM 1003  O2  SO4 A 134      12.800  44.341   9.207  1.00 61.54           O
HETATM 1004  O3  SO4 A 134      13.286  42.044   8.818  1.00 60.87           O
HETATM 1005  O4  SO4 A 134      13.983  43.693   7.239  1.00 61.20           O
HETATM 1006  S   SO4 A 135       3.975   9.982  26.637  1.00 62.08           S
HETATM 1007  O1  SO4 A 135       4.829  10.167  25.468  1.00 61.20           O
HETATM 1008  O2  SO4 A 135       2.615  10.409  26.316  1.00 61.74           O
HETATM 1009  O3  SO4 A 135       3.962   8.571  27.010  1.00 61.87           O
HETATM 1010  O4  SO4 A 135       4.491  10.775  27.751  1.00 61.04           O
HETATM 1011  O   HOH A 136      12.440  30.585  14.920  1.00 10.31           O
HETATM 1012  O   HOH A 137      -5.723  23.172  16.513  1.00 11.61           O
HETATM 1013  O   HOH A 138       8.131  27.182  28.214  1.00 18.45           O
HETATM 1014  O   HOH A 139      13.491  21.124  32.927  1.00 18.30           O
HETATM 1015  O   HOH A 140      16.545  33.206  19.929  1.00 15.18           O
HETATM 1016  O   HOH A 141      12.396  22.925  24.847  1.00 11.57           O
HETATM 1017  O   HOH A 142      -6.161  34.301  13.178  1.00 22.23           O
HETATM 1018  O   HOH A 143      -4.229  21.054  15.115  1.00 18.56           O
```

Figure 2.1.1 pdb file 1M47 showing atoms, residues and xyz-coordinates of IL2.

2.2 FASTA file

FASTA format is a text based format to represent either nucleotide sequences or amino acid peptide chains, in which base pairs of amino acids are represented by single letter codes. A sequence in FASTA format starts with a single line description, followed by the sequence data lines. The description line is distinguished from the sequence data by a greater than(>) sign in the first column.
We took the FASTA format files of the interleukins from https://www.rcsb.org/
Below mentioned is a FASTA format of Interleukin-7 from pdb file-3DI3.

>3DI3_1|Chain A|Interleukin-7|Homo sapiens (9606)
MGDCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSNCLNNEFNFFKRHICDANKEG
MFLFRAARKLRQFLKMNSTGDFDLHLLKVSEGTTILLNCTGQVKGRKPAALGAA
QPTKSLEENKSLKEQKKLNDLCFLKRLLQEIKTCWNKILMGTKEH

The accepted amino acid codes are listed below.

Table 2.2.1 Amino acids and their codes

Amino Acid	3 Letter Code	One Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

2.3 Structural Homology

Based on the previous knowledge available about various Interleukins and the structures available in the public domain we aligned the different Interleukins to each other so as to look into their similarities and differences in their structures, sequences and the function they possess in the body.

The Interleukins were aligned using the "align" plugin in PyMOL. PyMOL is a open-source software available for model visualization and used in structural biology(26). "align" command performs a sequence alignment followed by a structural superposition, and then carries out zero or more cycles of refinement in order to reject outliers, the program then returns the RMSD value for all the aligned atoms. If one wants to perform an sequence-independent structure based alignment then they can use "super" command of pymol. It aligns two selections on the basis of their structures(2). "Super" is more preferred than "align" for proteins with low sequence similarity.

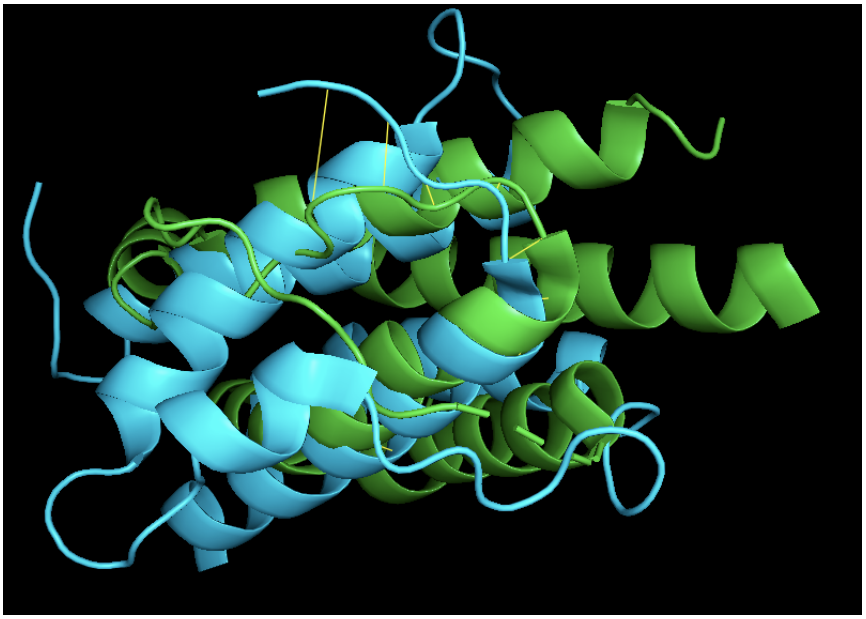

Figure 2.3.1 IL2(green) and IL7(cyan) aligned

The above image is of Interleukin-2(green) and Interleukin-7(blue) aligned using the "align" tool. The yellow lines show the residues which are aligned together in the proteins. Here we see few yellow lines indicating a better alignment of the two Interleukins.

On the other hand in the figure 2.3.2(below) Interleukin-4 and Interleukin-6 are aligned. Clearly we see a lot of yellow lines in this figure indicating the poor alignment of the residues of the two proteins.

Figure 1.3.2 IL4(cyan) and IL6(red) aligned

2.4 Sequence Homology

Various Interleukins are observed to show common receptors complexes, pathways of action, inflammatory actions and similar functions. We here have examined the sequence homology of various interleukins and their receptor complexes so as to figure out the similarities in their signalling and functions.

We here used the "swalign" command in MATLAB's bio-informatics toolbox to compute the sequence homology, percent identity, similarity and percent similarity of the Interleukins pairs. "swalign" uses Smith-Waterman algorithm to perform a local alignment of the sequence of the DNA or protein.

Syntax

Score = swalign(*Seq1*, *Seq2*)
[*Score, Alignment*] = swalign(*Seq1*, *Seq2*)
[*Score, Alignment, Start*] = swalign(*Seq1*, *Seq2*)
... = swalign(*Seq1*,*Seq2*, ...'Alphabet', *AlphabetValue*)
... = swalign(*Seq1*,*Seq2*, ...'ScoringMatrix', *ScoringMatrixValue*, ...)
... = swalign(*Seq1*,*Seq2*, ...'Scale', *ScaleValue*, ...)
... = swalign(*Seq1*,*Seq2*, ...'GapOpen', *GapOpenValue*, ...)
... = swalign(*Seq1*,*Seq2*, ...'ExtendGap', *ExtendGapValue*, ...)
... = swalign(*Seq1*,*Seq2*, ...'Showscore', *ShowscoreValue*, ...)

Chapter 3

Results and Discussion

3.1 Structural Analysis

Table 3.1.1 RMSD values of Interleukin pairs aligned

Sr. No.	INTERLEUKINS	RMSD
1	IL1β- IL38	0.816
2	IL37-IL38	1.026
3	IL1β- IL37	1.219
4	IL1α- IL18	1.542
5	IL18-IL38	2.480
6	IL1α-IL37	2.588
7	IL1α-IL38	2.629
8	IL19-IL22	2.891
9	IL1β- IL18	2.902
10	IL2-IL7	3.213
11	IL15-IL21	3.605
12	IL4-IL21	3.636
13	IL3-IL5	4.129
14	IL2-IL6	5.339
15	IL1α-IL6	5.898
16	IL1β-IL5	6.440
17	IL17A-IL23	6.529
18	IL22-IL24	6.630
19	IL1α-IL5	6.638
20	IL4-IL15	7.277
21	IL3-IL4	7.358
22	IL3-IL6	7.607
23	IL2-IL21	7.706
24	IL7-IL15	9.598
25	IL17A-IL33	9.786
26	IL10-IL19	9.810
27	IL2-IL3	10.592
28	IL4-IL5	10.649
29	IL2-IL5	10.758
30	IL4-IL6	11.088
31	IL4-IL7	12.043

32	IL2-IL15	12.168
33	IL5-IL6	12.916
34	IL1β-IL3	14.288
35	IL18-IL37	14.327
36	IL7-IL21	14.390
37	IL10-IL24	15.859
38	IL10-IL22	17.617

3.1.1 Aligned structures resulting RMSD value less than 5

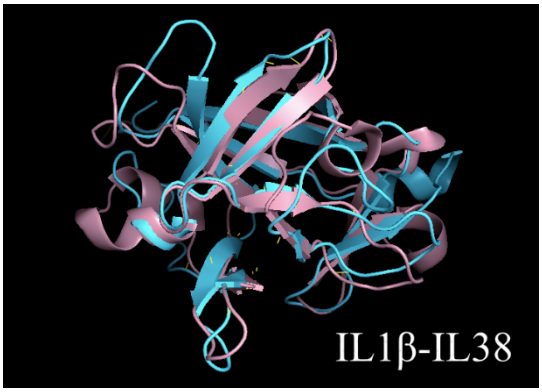

Figure 3.1.1.1 IL1β(cyan)-IL38(pink); RMSD=0.816 Å

Figure 3.1.1.2 IL37(red)-IL38(pink); RMSD= 1.026 Å

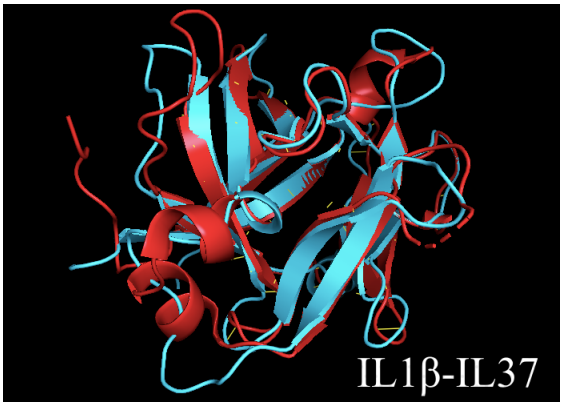

Figure 3.1.1.3 IL1β(cyan)-IL37(red); RMSD=1.219 Å

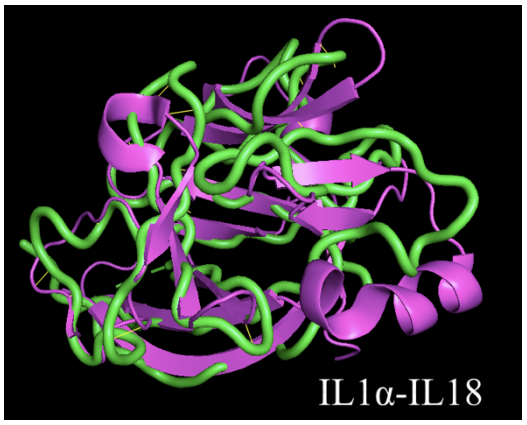

Figure 3.1.1.4 IL1(green)-IL18(magenta) ; RMSD=1.542 Å

Figure 3.1.1.5 IL18(magenta)-IL38(pink); RMSD= 2.480 Å

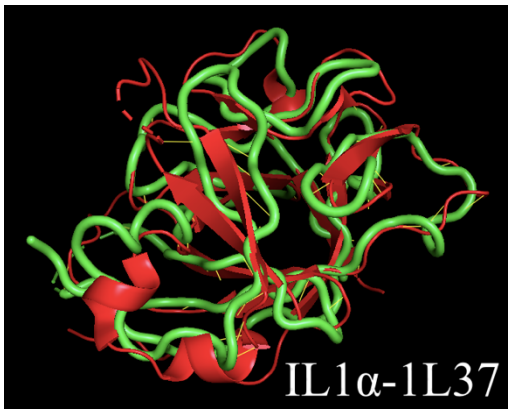

Figure 3.1.1.6 IL1(green)-IL37(red); RMSD= 2.588 Å

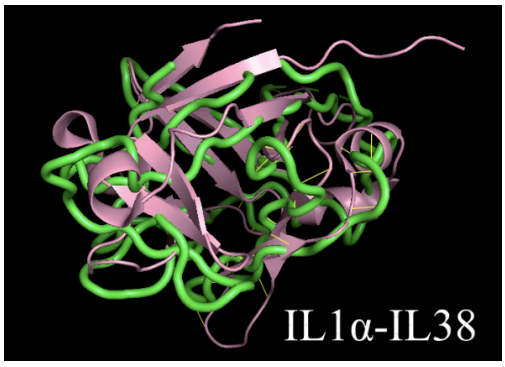

Figure 3.1.1.7IL1(green)-IL38(pink); RMSD= 2.629 Å

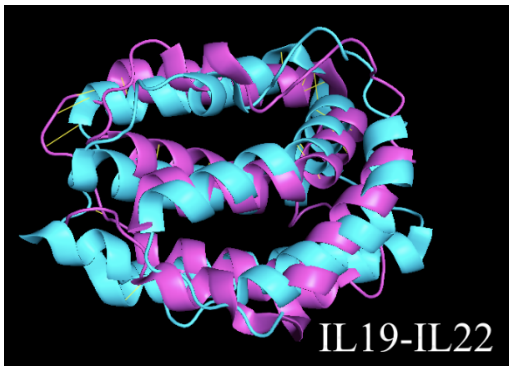

Figure 3.1.1.8 IL19(cyan)-IL22(magenta); RMSD= 2.891 Å

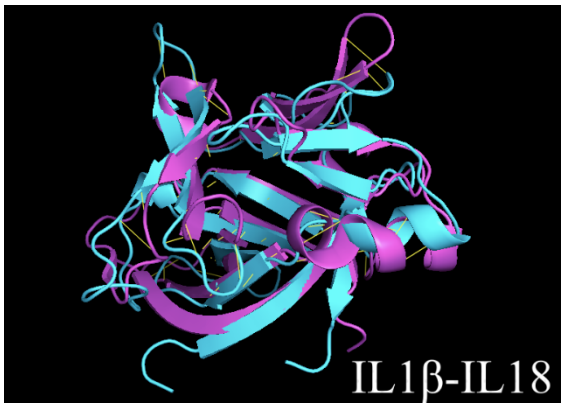

Figure 3.1.1.9 IL1β(cyan)-IL18(magenta); RMSD= 2.902 Å

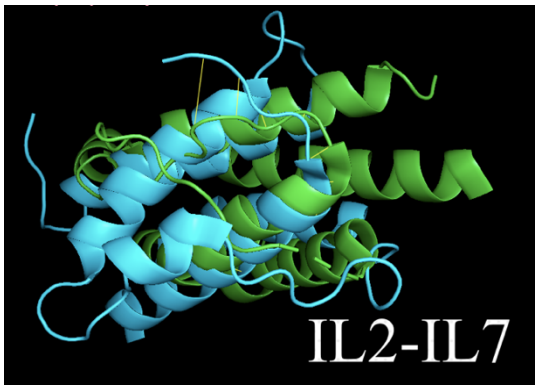

Figure 3.1.1.10 IL2(green)-IL7(cyan); RMSD= 3.213 Å

Figure 3.1.1.11 IL15(cyan)-IL21(grey): RMSD= 3.605 Å

Figure 3.1.1.12 IL4(magenta)-IL21(grey); RMSD=3.636 Å

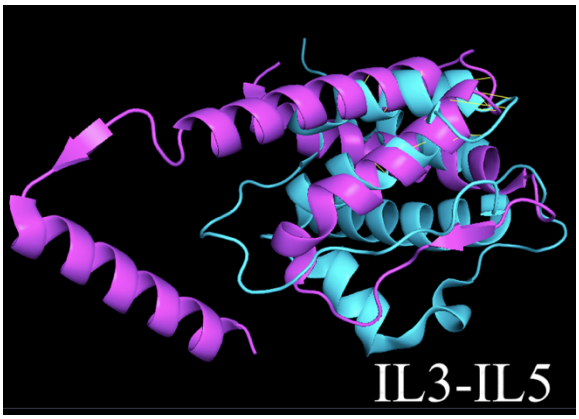

Figure 3.1.1.13 IL3(cyan)-IL5(magenta); RMSD= 4.129 Å

The above images illustrate the complexes of various Interleukins pairs aligned together which result in an root mean square deviation(RMSD) value ranging from 0.816 to 4.129. In the initial figures i.e. from 3.1.1.1 to 3.1.1.9, pretty good overlap is observed having RMSD values less than 3. Except the IL19-IL22 map(figure 3.1.1.8) all other interleukins aligned are of IL-1 family. IL1 family comprises of 11 members, including 7 pro-inflamatory agonists (IL-1α, IL-1β, IL-18, IL-33α, IL-36α, IL-36β, and IL-36γ) and 4 defined or putative antagonists (IL-1Ra, IL-36Ra, IL-37 and IL-38). Structures of various members of the IL1 family are observed to be identical showing a good overlap. IL1β an IL38 are observed to have the lowest RMSD value indicating the most similar structure of all the ILs observed, also contributing to sharing of common receptor (IL1R1) by both Interleukins. On the other hand IL19 and IL22, both belong to the IL10 family, their structural overlap explains their common disease association in psoriasis and cancer but even after such similarity in structure they share different receptors resulting in major differences in their pathways.

Further from figure 3.1.1.10 to 3.1.1.12 all Interleukin pairs aligned belong to the common gamma chain family, as they all share gamma chain in their receptor complexes. This data contribute to the binding of these ILs to the common gamma chain.
IL3 and IL5 also show an average overlap with RMSD 4.1. Presently, their only common association is in asthma, but observing the similarity in their structures, they could have various more common associations which can be looked at in future.

3.1.2 Structures resulting RMSD value ranging 5-10

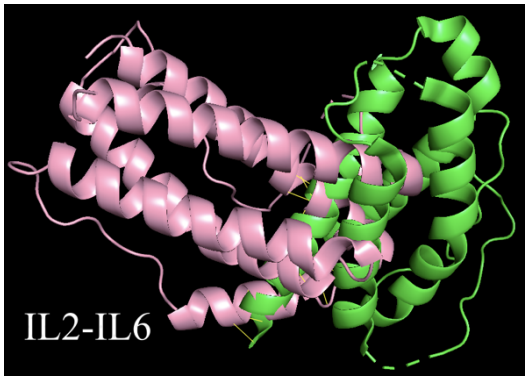

Figure 3.1.2.1 IL2(green)-IL6(pink); RMSD= 5.339 Å

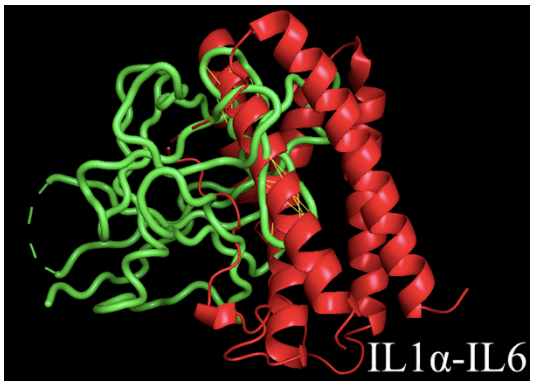

Figure 3.2.1.2 IL1(green)-IL6(red); RMSD= 5.898 Å

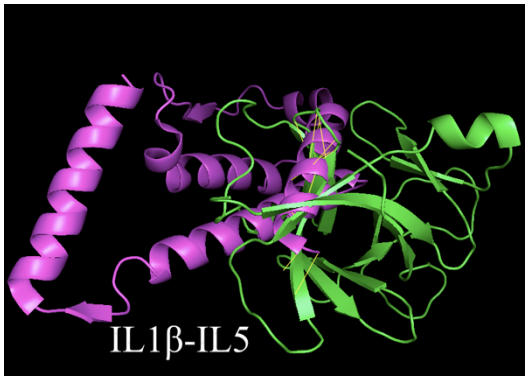

Figure 2.1.2.3 IL1β(green)-IL5(magenta);

RMSD= 6.440 Å

Figure 3.1.2.4 IL17A(cyan)-IL23(magenta);

RMSD= 6.529 Å

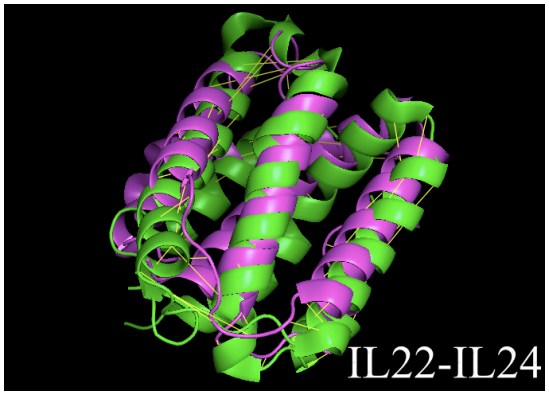

Figure 3.1.2.5 IL22(magenta)-IL24(green);

RMSD= 6.630 Å

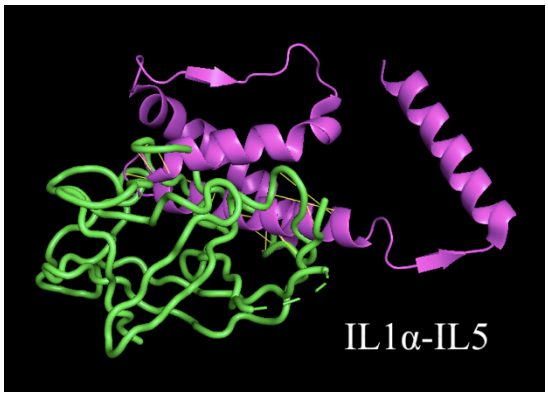

Figure 3.1.2.6 IL1(green)-IL5(magenta); RMSD= 6.638 Å

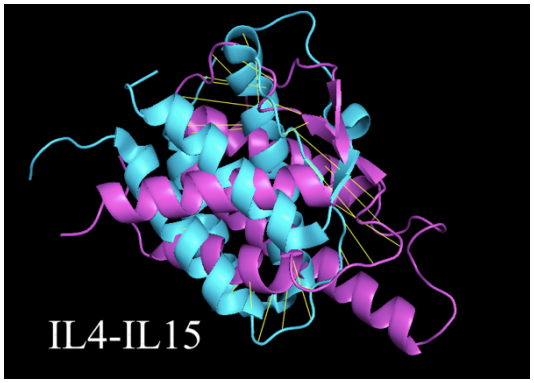

Figure 3.1.2.7 IL4(magenta)-IL15(cyan); RMSD= 7.277 Å

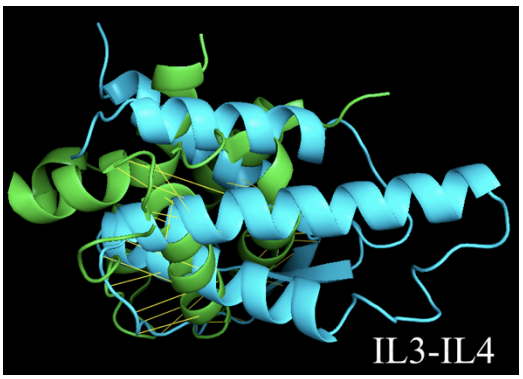

Figure 3.1.2.8 IL3(green)-IL4(cyan); RMSD=7.358 Å

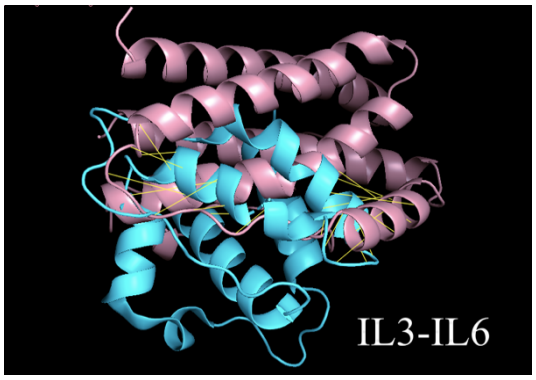

Figure 3.1.2.9 IL3(cyan)-IL6(pink); RMSD= 7.607 Å

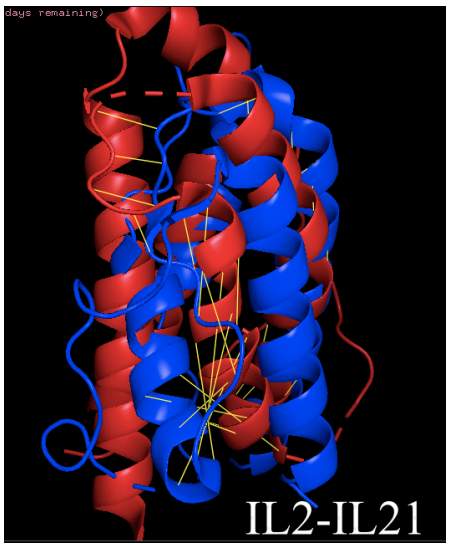

Figure 3.1.2.10 IL2(red)-IL21(blue); RMSD= 7.706 Å

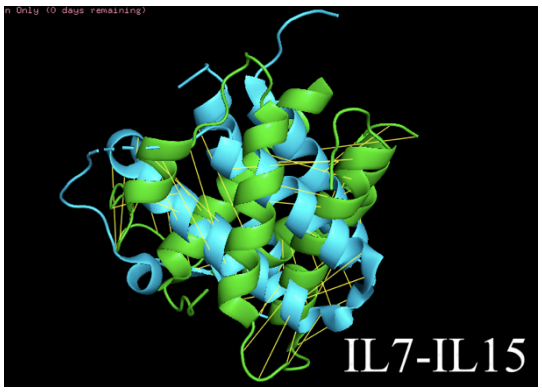

Figure 3.2.1.11 IL7(green)-IL15(cyan); RMSD= 9.598 Å

Figure 3.1.2.12 IL17a(cyan)-IL33(pink); RMSD= 9.786 Å

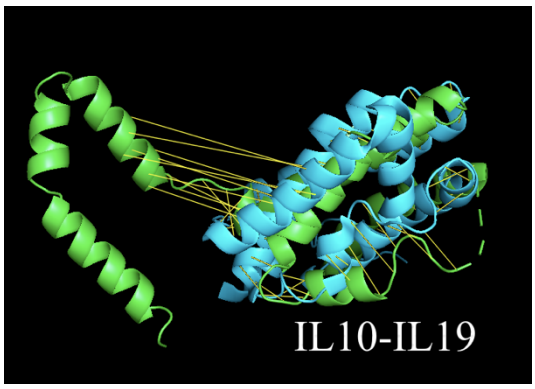

Figure 3.1.2.13 IL10(green)-IL19(cyan); RMSD= 9.810 Å

In all the above images from 3.1.2.1 to 3.1.2.13, Interleukin pairs are observed with a poor alignment with each other. Long yellow lines visible here show that the residues which are aligned are far from each other and do not align well. This confirms herewith to the previous information available about these interleukins and proclaim the different association and function of these interleukins as these Interleukin pairs have very less similar disease association and therapeutic applications(4).

3.1.3 RMSD Value greater than 10

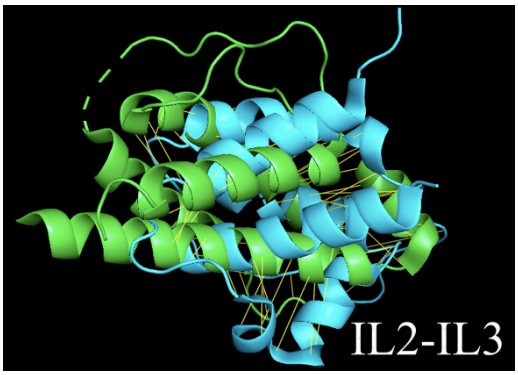

Figure 3.1.3.1IL2(green)-IL3(cyan); RMSD= 10.592 Å

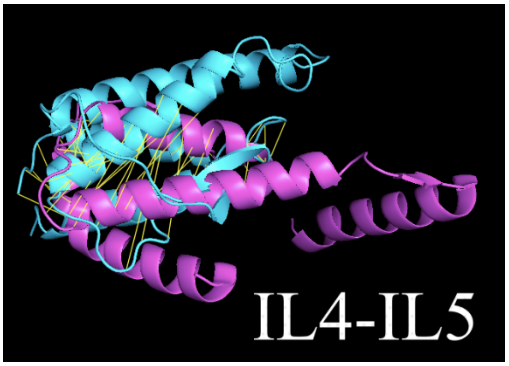

Figure 3.1.3.2 IL4(cyan)-IL5(magenta); RMSD= 10.649 Å

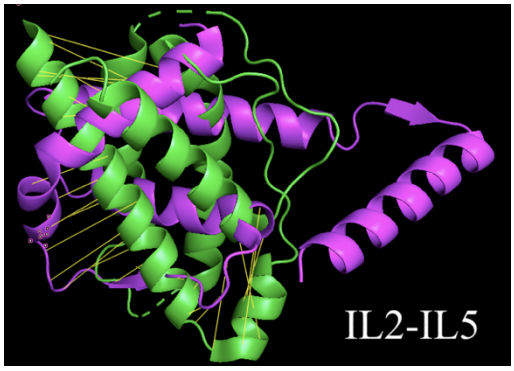

Figure 3.1.3.3 IL2(green)-IL5(magenta); RMSD= 10.758 Å

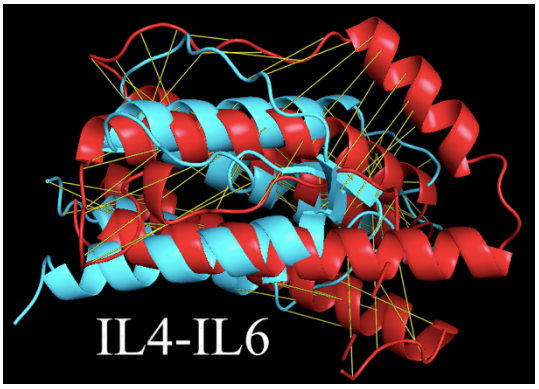

Figure 3.1.3.4 IL4(cyan)-IL6(red); RMSD= 11.088 Å

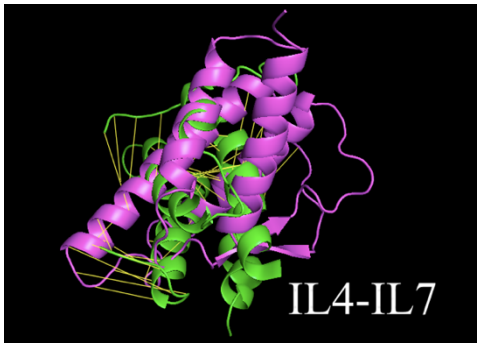

Figure 3.1.3.5 IL4(magenta)-IL7(green); RMSD= 12.043 Å

Figure 3.1.3.6 IL2(green)-IL15(magenta); RMSD= 12.168 Å

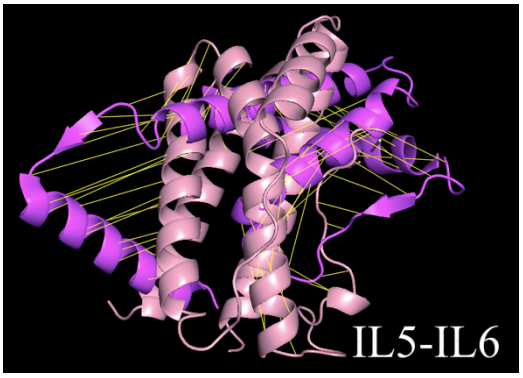

Figure 3.1.3.7IL5(magenta)-IL6(pink); RMSD= 12.916

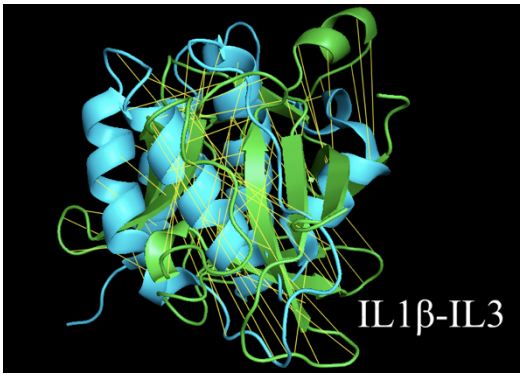

Figure 3.1.3.8 IL1β(green)-IL3(cyan); RMSD= 14.288

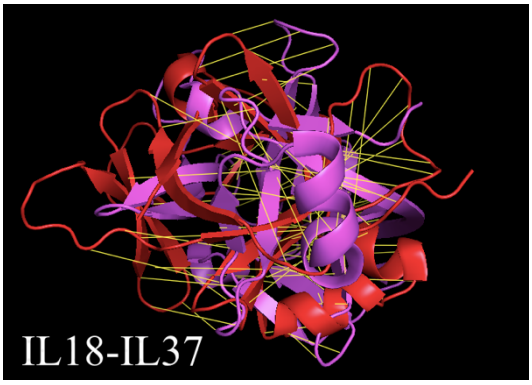

Figure 3.1.3.9 IL18(magenta)-IL37(red); RMSD= 14.327 Å

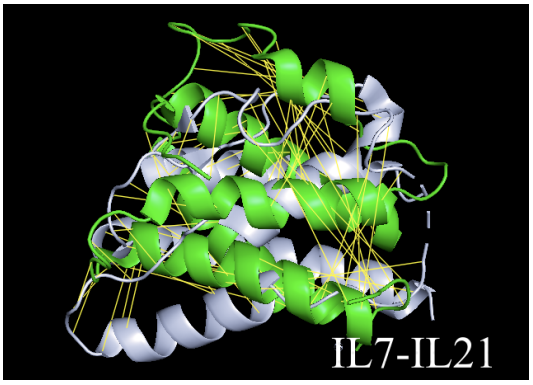

Figure 3.1.3.10 IL7(green)-IL21(grey); RMSD= 14.339 Å

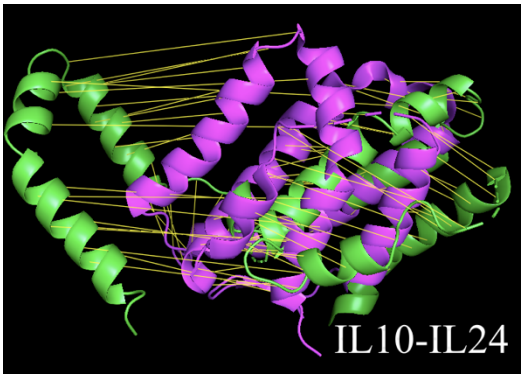

Figure 3.1.3.11 IL10(green)IL24(magenta);

-RMSD= 15.859 Å

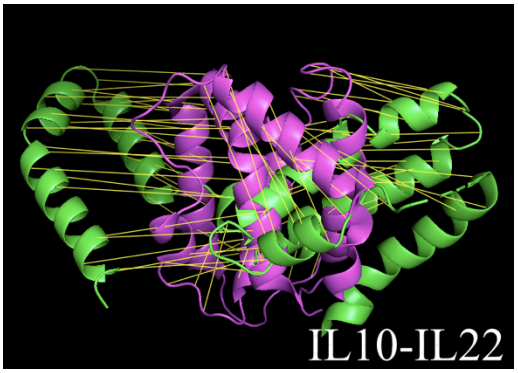

Figure 3.1.3.12 IL10(green)-IL22(light magenta);

RMSD= 17.617 Å

Analysis of the above mentioned IL pairs show a very poor alignment as can be seen from the even longer yellow lines, revealing very distant and poor residue alignment. The IL4,

IL7 and IL21(Figures 3.1.3.5 and 3.1.3.10) all belong to common gamma chain family and yet are observed to show a great difference in their structures. Similarly, IL10, IL22 and IL24 belong to IL10 family and like the previous three these also show a great structural disparity(Figure 3.1.3.11 and 3.1.3.12). This data verifies the different functional aspects executed by these interleukins by showing the difference in their structural mapping.

3.2 Sequence Homology

Sequence homology of Interleukin pairs were calculated using MATLAB. The detailed data is shown below

Table 3.2.1 Common Gamma Chain Interleukins Pairwise Alignment

Pairwise Alignment	Identity	Total Sequence after alignment using SW	Percent Identity	Similarity	Percent Similarity
IL21 - IL15	24	111	21.621	67	60.361
IL21-IL9	6	27	22.222	16	59.259
IL21-IL7	7	32	21.875	22	68.750
IL21-IL4	7	26	26.923	21	80.769
IL21-IL2	26	123	21.138	64	52.032
IL15-IL9	6	15	40.000	9	60.000
IL15-IL7	18	57	31.578	33	57.894
IL15-IL4	6	22	27.272	14	63.636
IL15-IL2	24	83	28.915	45	54.216
IL9-IL7	9	34	26.470	23	67.647
IL9-IL4	12	38	31.578	25	65.789
IL9-IL2	9	42	21.428	23	54.761
IL7-IL4	16	54	29.629	29	53.703
IL7-IL2	18	91	19.780	53	58.241
IL4-IL2	13	49	26.530	27	55.102

We observe the percent similarity of the interleukins of common gamma family ranging from around 50- 70%. Least similarity is shown by IL2-IL21 pair i.e. 52.032%, and various other IL pairs show similarity going up to the second highest observed to be 68.750 of IL7-IL21 complex. A surprisingly high sequence similarity is observed in the pair IL4-IL21 i.e. 80% which is the highest among the whole family. We had observed earlier, IL4-IL21 pair also had comparatively good structural homology with RMSD of 3.636(section 3.1.1). This shows that there might be some more aspects which can be looked into, that are not yet being observed or researched.

Table 3.2.2 IL1 Family Interleukins Pairwise Alignment

Pairwise Alignment	**Identity**	**Total Sequence after alignment using SW**	**Percent Identity**	**Similarity**	**Percent Similarity**
IL1A-IL18	33	173	19.075	66	38.150
IL1A-IL33	24	217	11.059	44	20.276
IL1A-IL37	36	187	19.251	67	35.828
IL1A-IL38	35	168	20.834	59	35.119
IL1B-IL18	31	171	18.128	66	38.596
IL1B-IL33	31	183	16.939	58	31.693
IL1B-IL37	43	182	23.626	74	40.659
IL1B-IL38	36	175	20.571	64	36.571
IL18-IL33	32	175	18.285	60	34.285
IL18-IL37	38	206	18.446	71	34.466
IL18-IL38	29	170	17.058	65	38.235
IL33-IL37	16	280	5.714	21	7.500
IL33-IL38	15	249	6.024	33	13.253
IL37-IL38	46	188	24.468	79	42.021

IL1 Family are observed to share their several receptor complexes among themselves and also they have very common signalling and functional homology, yet they are not observed to have a good sequence homology. Their similarity in the IL pairs range from as low as

7.5% in IL33-IL37 alignment to the highest of 42.021% in IL37-IL38 pair. Even after having a low sequence homology they have share a lot in common, probably because of the comparatively good structural homology.

Table 3.2.3 Pairwise Alignment of IL10 Family Interleukins

Pairwise Alignment	Identity	Total Sequence after alignment using SW	Percent Identity	Similarity	Percent Similarity
IL10-IL19	29	183	15.846	67	36.612
IL10-IL20	41	166	24.698	73	43.975
IL10-IL22	39	170	22.941	67	39.411
IL10-IL24	36	172	20.930	73	42.441
IL10-IL29	38	220	17.272	66	30.000
IL19-IL20	64	159	40.251	97	61.006
IL19-IL22	30	182	16.483	58	31.868
IL19-IL24	48	173	27.745	82	47.398
IL19-IL29	6	304	1.973	17	5.592
IL20-IL22	37	143	25.874	87	60.839
IL20-IL24	63	151	41.721	113	74.834
IL20-IL29	10	41	24.390	26	63.414
IL22-IL24	35	133	26.315	84	63.157
IL22-IL29	15	86	17.441	48	55.813
IL24-IL29	11	27	40.740	19	70.370

IL10 Family is observed to have a large range of sequence homology range from lowest of 5.592% in IL19-IL21 pair and the highest to 74.834% in IL20-IL24 aligned pair. IL20-IL24 share 113 similar residues out of 151, which is a pretty good alignment for consideration for further detailed research of the two. IL24-IL29, IL22-Il24, IL20-IL29, IL19-Il20 also show significantly good homology with similarity greater than 60%. These complexes are observed to share several receptors among themselves. Our analysis shows a positive result to the previous observations.

3.3 Correlation between align RMSD and sequence homology

The Rmsd value of the structures aligned, where in some Interleukin pairs sync with the sequence homology, in others it seems to be not so convincing. From the above data we can see that the IL1 family has the best structural alignment but the sequence similarity is pretty low in the same. Similar trend is observed in IL10 family, sequence homology in IL 10 family is seen to be decent but the align rmsd is very high(highest: 17.617 in IL10-IL22 map) for most of the members here.
However, common gamma chain cytokines are observed to have a decent both structural and sequence homology and their align rmsd values and percent similarity are observed to be homologous.

3.4 Pleiotropy

Cytokine Pleiotropy is the ability of a cytokine to exert multiple different responses on different cells types. Several cytokines are observed to have common receptor complexes shared among themselves. Same Interleukin binding to two or more different receptor complexes, exert different functions through each complex. The reason for this phenomenon could be any, structural similarity, sequence similarity or energy degeneracy. The exact reason for these pleiotropic complexes binding is yet not clear.

From the above data we observe that, the interleukins possessing pleiotropy have a large amount of sequence and structural similarity. The interleukins shared by the gamma c receptor show sequence similarity ranging from 52 to 80 percent, meaning that the interleukins have more than half of the sequence similar among themselves. IL10 family is also observed to have an average of 48 percent sequence similarity. IL1 family also shows an average sequence similarity of about 31 percent. IL1 family shows a comparative low sequence homology but have a high structural overlap. Most of the IL1 family map pairs show a significant low rmsd values evidencing their common functions.

Pleiotropic cytokines not evident from the above data could have the energy degeneracy, One can look further into the energetics to have more understanding about the pleiotropic nature of the cytokines.

3.4.1 Type I and Type II receptor complex

Some cytokines exert their responses binding to two different type of receptor, Type I and Type II. Type II receptors lack the common WSXWS box observed in the Type I cytokine binding. Type I and Type II substrate pleiotropy is observed in several interleukins like, IL1, IL4, IL13, IL7, IL12, IL24 etc. The overall sequence homology or structural alignment of these interleukins is decent only and not extraordinarily high and does not signify this behaviour. The probable reason for this could be the specific sequence or a specific structural part of the Interleukin that has high affinity towards the receptor complex, which gives the ability to the ILs to function through various receptor complexes.

Chapter 4

Conclusion

Sturcture alignment of IL1 family was observed to be very good resulting in low RMSD values and a good sequence alignment was observed in the common gamma chain ILs. Various ILs confirm by their structural and sequence mapping, the similarity observed in their functions and disease association. Structural overlap of IL1β and IL38 with RMSD of 0.816 was very good contributing to their sharing of receptor but have a poor sequence homology which shows their association to different functions.

IL4 and IL21 were observed to have the highest sequence homology(80%), they both have common production source i.e. TH2 cells. Where IL21 plays major role in B cell proliferation , differentiation and survival, IL4 also contributes to survival of B cells but not primarily. Structural overlap of IL4 and IL21 was also comparatively observed to be decently better. IL4 and IL21 may have more common function which may be observed in the future.

Why Interleukins show similarities or differences in their receptors, signalling or functions can be upto some extent be answered from the observed data. Interleukins showing more than one kind of responses and those that bind two more than one receptor complexes could have both structural or sequence homology or could have either one of them to exert such responses as evident through the given results. Further, the interleukins not showing any structural or sequence homology exert this behaviour probably due to the similar energetics.

Bibliography

1. Paul WE, Seder RA. Lymphocyte responses and cytokines. Cell. 1994;76:241–251.

2. Ivashkiv LB. Cytokines and STATs: How can signals achieve specificity? Immunity. 1995;3:1–4.

3. Cezmi A. Akdis, Interleukins (from IL-1 to IL-38), interferons, transforming growth factor β, and TNF-α: Receptors, functions, and roles in diseases, Journal of Allergy and Clinical Immunology, Volume 138, Issue 4, 2016, Pages 984-1010;

4. Glial proinflammatory cytokines mediate exaggerated pain states: implications for clinical pain,Watkins LR, Milligan ED, Maier SF,Adv Exp Med Biol. 2003; 521():1-21.

5. Nicola NA. Cytokine pleiotropy and redundancy: a view from the receptor. *Stem Cells*. 1994;12 Suppl 1:3-14.

6. Brocker C, Thompson D, Matsumoto A, Nebert DW, Vasiliou V (Oct 2010). "Evolutionary divergence and functions of the human interleukin (IL) gene family". Human Genomics. **5** (1): 30–55.

7. Dinarello CA. Interleukin-1. Cytokine Growth Factor Rev. 1997;8:253–265.

8. Dinarello CA, Novick D, Puren AJ. et al. Overview of interleukin-18: More than an interferon-γamma inducing factor. J Leukoc Biol. 1998;63:658–664.

9. Dinarello C, Arend W, Sims J, Smith D, Blumberg H, O'Neill L, et al. IL-1 fam- ily nomenclature. Nat Immunol 2010;11:973.

10. Akdis M, Burgler S, Crameri R, Eiwegger T, Fujita H, Gomez E, et al. Interleu- kins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in dis- eases. J Allergy Clin Immunol 2011;127:701-21, e1-70.

11. Akdis M, Akdis CA. Mechanisms of allergen-specific immunotherapy: multiple suppressor factors at work in immune tolerance to allergens. J Allergy Clin Immunol 2014;133:621-31.

12. Suzuki H, Zhou YW, Kato M. et al. Normal regulatory alpha/beta T cells effectively eliminate abnormally activated T cells lacking the interleukin 2 receptor beta in vivo. J Exp Med. 1999;190:1561–1572.

13. Paul WE. Interleukin-4: A prototypic immunoregulatory lymphokine. Blood. 1991;77:1859–1870.

14. Maki K, Sunaga S, Ikuta K. The V-J recombination of T cell receptor-gamma genes is blocked in interleukin-7 receptor-deficient mice. J Exp Med. 1996;184:2423–2427

15. Renauld JC, Goethals A, Houssiau F. et al. Human P40/IL-9. Expression in activated CD4+ T cells, genomic organization, and comparison with the mouse gene. J Immunol. 1990;144:4235–4241.

16. Grabstein KH, Eisenman J, Shanebeck K. et al. Cloning of a T cell growth factor that interacts with the β chain of the interleukin-2 receptor. Science. 1994;264:965–968.

17. Lodolce JP, Boone DL, Chai S. et al. IL-15 receptor maintains lymphoid homeostasis by supporting lymphocyte homing and proliferation. Immunity. 1998;9:669–676.

18. Miyajima A, Mui AL, Ogorochi T. et al. Receptors for granulocyte-macrophage colony-stimulating factor, interleukin-3, and interleukin-5. Blood. 1993;82:1960–1974.

19. Lantz CS, Boesiger J, Song CH. et al. Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. Nature. 1998;392:90–93.

20. Campbell HD, Tucker WQ, Hort Y. et al. Molecular cloning, nucleotide sequence, and expression of the gene encoding human eosinophil differentiation factor (interleukin 5). Proc Natl Acad Sci U S A. 1987;84:6629–6633.

21. Kopf M, Herren S, Wiles MV. et al. Interleukin 6 influences germinal center development and antibody production via a contribution of C3 complement component. J Exp Med. 1998;188:1895–1906.

22. Du XX, Williams DA. Interleukin-11: A multifunctional growth factor derived from the hematopoietic microenvironment. Blood. 1994;83:2023–2030.

23. Du X, Williams DA. Interleukin-11: Review of molecular, cell biology, and clinical use. Blood. 1997;89:3897–3908.

24. Gately MK, Renzetti LM, Magram J. et al. The interleukin-12/interleukin-12-receptor system: Role in normal and pathologic immune responses. Annu Rev Immunol. 1998;16:495–521.

25. Moore KW, O'Garra A, de Waal Malefyt R. et al. Interleukin-10. Annu Rev Immunol. 1993;11:165–190.

26. DeLano WL. The PyMOL Molecular Graphics System. Palo Alto, CA, USA: DeLano Scientific; 2002.

27. Barrett, G. C., and Elmore, D. T. (1998). *Amino Acids and Peptides.* New York: Cambridge University Press.