INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
Identification, Genomic Structure, and Functional Studies of the Human Novel CC Chemokine MIP-4, and the Cloning of the Murine Homologue

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Ping Guan, MS

The Ohio State University
1998

Dissertation Committee:
Dr. Arthur H.M. Burghes (adviser)
Dr. Paul A. Fuerst
Dr. Michael C. Ostrowski
Dr. C. Yung Yu

Approved by

Dr. Arthur H.M. Burghes
Adviser

Department of Molecular Genetics
Abstract

The chemokines are a group of chemotactic molecules that appear to regulate the directed movement of white blood cells *in vitro* and *in vivo* and may therefore play important roles in inflammation and immunity. The genes encoding the chemokines are clustered in close physical proximity to each other. A large cluster of human CC chemokine genes reside on chromosome 17.

A positional cloning approach was used to identify novel chemokine genes within this cluster. A YAC contig was constructed, which encompasses the MIP-1α gene region and exon trapping and sequence analysis was used to isolate novel chemokine genes. Using this approach, a gene encoding a chemokine, named MIP-4, was isolated. The MIP-4 gene is closely related to MIP-1α with 49.5% identity at the nucleotide level and 59.6% at the predicted amino acid level. The MIP-4 gene consists of 3 exons spreading over 7.1 kb and is separated from the MIP-1α gene by 16 kb. The MIP-4 gene encodes an 800 bp mRNA that is expressed in lung and macrophages but not in brain or muscle. The mRNA encodes an 89 amino acid protein and includes a signal peptide of 21 amino acids.
Recombinant or synthetic MIP-4 induced calcium mobilization in naïve and activated T lymphocyte subpopulations \textit{in vitro}. Injection of synthetic MIP-4 into the peritoneal cavity of mice led to the accumulation of both CD4+ and CD8+ T lymphocytes, but not monocytes or granulocytes. These observations provide new information concerning the arrangement of the CC chemokine gene cluster on human chromosome 17 and indicate that the MIP-4 gene product is chemotactic \textit{in vivo} for both CD4+ and CD8+ T lymphocytes, and may therefore be implicated in both humoral and cell mediated immunity. Under the stimulation of cytokines IL-12 and IL-15, human natural killer cells express MIP-4. However, the incubation of MIP-4 protein with PBL depleted CD8+ cells did not inhibit the infection rate of either M-tropic or T-tropic HIV-1 virus. Instead, it appeared to enhance the HIV uptake. This might suggest that MIP-4 upregulate the expression of the chemokine receptor in the target cells.
To my family
Acknowledgments

I would like to acknowledge many people for their help and support. I would like to acknowledge my adviser, Dr. Arthur Burghes for his patience, guidance and encouragement throughout my graduate studies. I would like to thank the members of my graduate committee Dr. Paul Fuerst, Dr. Mike Ostrowski and Dr. Yung Yu for their helpful suggestions and support.

I would also like to acknowledge Dr. Kudeep Neote and his staff at Pfizer Central Research for performing MIP-4 chemokine assay and Dr. Shaun McColl at University of Adelaide, Australia for MIP-4 animal studies presented in this thesis. I owe a great deal of gratitude to Dr. Mike Caligiuri and Todd Fehninger at OSU Comprehensive Cancer Center for their help in the NK cell studies and Dr. Alessandra Oliva at NIH for the HIV infection assay.

I want to thank all the past and present members of the Burghes lab who helped me greatly in the course of my graduate studies. In particular, I want to thank Dr. John Carpten and Dr. Christine DiDonato for their help in the early phase of my project and Dr. Jegatheesan Seharaseyon, Dr. Dan Coover, Dr. Alissa Winnard, Dr. Thanh Le, Dr. Umrao Monani, Susan Burghes and Song Han who shared me with their technical expertise and support. Especially, I want to thank Dr. Patricia McAndrew for her tremendous help and comments on the writing of my thesis.

Last, I want to thank my parents for their unconditional love and constant support that has brought me to this far.
Vita

April 1, 1965 ............................................................... Born, Sichuan, China

1986.............................................................................. B.S., East China Normal University, Shanghai, China

1990............................................................................... M.S., the Chinese Academy of Sciences, Beijing, China

1990—1992................................................................... Instructor, Beijing Medical University, Beijing, China

1992—present................................................................ Graduate teaching and research associate, the Ohio State University, Columbus, OH

Publications


Field of study

Major Field: Molecular Genetics
Table of Contents

Abstract ......................................................................................................................... ii

Dedication ...................................................................................................................... iv

Acknowledgments ........................................................................................................ v

Vita ......................................................................................................................................... vi

List of Figures ............................................................................................................... xi

List of Tables .................................................................................................................. xiii

List of Abbreviations ..................................................................................................... xiv

Chapter Page

1 Introduction .................................................................................................................. 1
  1.1 Purpose, significance and organization of this thesis ....................................... 1
  1.2 Summary ............................................................................................................. 3

2 Introduction to Chemokines ..................................................................................... 7
  2.1 Background ......................................................................................................... 7
  2.2 Biological properties of chemokines ................................................................. 8
  2.3 CXC chemokine (α chemokine family) .............................................................. 8
  2.4 CC chemokine (β chemokine family) ............................................................... 9
  2.5 C chemokine (γ chemokine family) ................................................................. 10
  2.6 CXXXC chemokine (δ chemokine family) ........................................................ 10
  2.7 Three-dimensional structure of chemokines ................................................... 11
  2.8 Chemokine receptors ....................................................................................... 11
  2.9 Chemokine ligand-receptor interaction ......................................................... 12
  2.10 Chemokine receptors and signal transduction ............................................... 13
  2.11 Genomic structures of chemokine genes ....................................................... 14
  2.12 Chromosomal locations of chemokine genes ............................................... 14
  2.13 Regulation of chemokine gene expression .................................................... 16
  2.14 Chemokines in leukocyte trafficking .............................................................. 16
  2.15 Chemokines in inflammation ........................................................................ 17
  2.16 Chemokines and infectious diseases ............................................................... 18
List of Figures

Figures

Page
2.1 Amino acid sequences of human chemokines .................................................. 23
2.2 Two-loop structure of human chemokine proteins .......................................... 24
2.3 Structural similarity diagram of human chemokine proteins ......................... 25
2.4 Three-dimensional structures of human chemokine dimers ............................ 26
2.5 Schematic diagram of human chemokine receptor ......................................... 27
2.6 Signaling pathway of chemokines .................................................................. 28
2.7 Genomic organizations and chromosomal localization of human Chemokines ....................................................................................................... 29
2.8 Cluster regions of human CC and CXC chemokine genes .............................. 30
4.1 Determination of YAC sizes by Pulse Field Gel Electrophoresis (PFGE) ........ 55
4.2 Schematic representation of Vectorette PCR ................................................... 56
4.3 Assembly of a YAC contig ............................................................................... 57
4.4 Southern analysis of the YAC contig ............................................................... 58
4.5 Comparison of the contig map with the contig published by Hieshima et al (Hieshima et A., 1997) ............................................................ 59
5.1 Map of SuperCos I cosmid vector ................................................................... 82
5.2 Vector map of pSPL3B ...................................................................................... 83
5.3 Schematic representation of exon trapping procedure .................................... 84
5.4 Map of the 3'-exon trapping vector pTAG4 .................................................... 85
5.5 Nucleotide sequences of THC-14098_1 and THC-79282 ................................ 86
5.6 RT-PCR amplification of cDNA from different cell lines ............................... 87
5.7 Expression profile of THC-14098_1 in human tissues .................................... 88
5.8 Expression profile of THC-79282 ................................................................. 89
5.9 Alignment of human CC chemokine sequences at amino acid level ............. 90
5.10 Genomic sequence of the human MIP-4 gene ................................................ 91
5.11 Southern analysis of cosmid DNA ................................................................. 92
5.12 Genomic structure of the human MIP-4 gene ................................................ 93
5.13 The chromosomal location of the MIP-4 gene related to the MIP-1α Gene ................................................................................................................ 94
6.1 Map of the prokaryotic expression vector pRSET ........................................... 112
6.2 SDS-PAGE (12%) of the MIP-4 fusion protein generated from pRSET_C ................................................................. 113
6.3 Map of the prokaryotic expression vector pET-32c ......................................... 114
6.4 SDS-PAGE (18%) of the MIP-4 fusion protein generated from pET-32c .......... 115
6.5 MIP-4 mobilizes intracellular calcium in fresh human T lymphocytes .......... 116
6.6 MIP-4 mobilizes intracellular calcium in activated human T lymphocytes ................................................................. 117
6.7 MIP-4 has no effects on the total number of leukocytes \textit{in vivo} .................. 118
6.8 MIP-4 attracts both CD4+ and CD8+ T lymphocytes \textit{in vivo} .................. 119
6.9 Expression of MIP-4 in NK cells under cytokine stimulation ......................... 120
7.1 Southern hybridization of BAC 114K21 DNA with mouse MIP-1\textalpha\ cDNA probe and human MIP-4 cDNA probe .................................................. 127
List of Tables

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>54</td>
</tr>
<tr>
<td>5.1</td>
<td>81</td>
</tr>
</tbody>
</table>
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>CAM</td>
<td>Cellular adhesion molecule</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CCR</td>
<td>CC chemokine receptor</td>
</tr>
<tr>
<td>CD</td>
<td>Clusters of differentiation</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CEPH</td>
<td>The Centre d’Etude du Polymorphisme Humain</td>
</tr>
<tr>
<td>cfu</td>
<td>colony forming unit</td>
</tr>
<tr>
<td>CIAP</td>
<td>Calf intestinal alkaline phosphatase</td>
</tr>
<tr>
<td>CXCR</td>
<td>CXC chemokine receptor</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethyl pyrocarbonate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diamine trichloroacetic acid</td>
</tr>
<tr>
<td>EK</td>
<td>Enterokinase</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tag</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>GAG</td>
<td>Glycosaminoglycan</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukine</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-beta-D-1-thiogalact pyranoside</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>μl</td>
<td>Microliter</td>
</tr>
<tr>
<td>ml</td>
<td>Milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>Milimolar</td>
</tr>
<tr>
<td>MIP</td>
<td>Macrophage inflammatory protein</td>
</tr>
<tr>
<td>MCP</td>
<td>Monocyte chemotactic protein</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple cloning site</td>
</tr>
<tr>
<td>mw</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>NK</td>
<td>Natural killer cell</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pd(N)₆</td>
<td>Random hexamers</td>
</tr>
<tr>
<td>PFGE</td>
<td>Pulse field gel electrophoresis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pfu</td>
<td>Plaque forming unit</td>
</tr>
<tr>
<td>PHA</td>
<td>Phytohaemagglutinin</td>
</tr>
<tr>
<td>PMBC</td>
<td>Peripheral blood mononuclear cell</td>
</tr>
<tr>
<td>RANTES</td>
<td>Regulated on activation, normal T cell expressed and secreted</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TIGR</td>
<td>The Institute of Genomic Research</td>
</tr>
<tr>
<td>UDG</td>
<td>Uracil DNA glycosylase</td>
</tr>
<tr>
<td>YAC</td>
<td>Yeast artificial chromosome</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Purpose, significance and organization of this thesis

Chemokines are the largest subfamily of the cytokine superfamily and these molecules play a key role in directing white blood cell traffic during immune responses. They are small secreted proteins which function as chemoattractants and activators of various cell populations in the immune system (Baggiolini et al., 1997; Schall et al., 1994). Members of the chemokine family have a high degree of amino acid homology and share a conserved four-cysteine residue motif. Chemokines are further subdivided into four groups based on their structures: CXC, CC, C, and CXXXC chemokines (also called α, β, γ, δchemokine) (Oppenheim et al., 1991; Pan et al., 1997). Recent investigations have demonstrated that the abnormal expression of some chemokines is associated with inflammatory diseases such as psoriasis, asthma, rheumatoid arthritis and respiratory/lung disorders (Strieter et al., 1996). Chemokines also promote wound healing, inhibit tumor cell growth, and are involved in the regulation of viral infection. CC chemokines produced by CD8+ T lymphocytes, such as MIP-1α, MIP-1β and RANTES, have been identified as major HIV-suppressive factors (Cocchi et al., 1995). Several chemokine receptors, CCR5, -2b, -3 are found to be cofactors for HIV infection (Dragic et al., 1996; Feng et al., 1996). The identification of new members in this gene
family yields further knowledge about the important functions of chemokines, contributes to our understanding of the pathogenesis of several diseases, allows the development of therapeutic interventions.

My overall goal was to identify novel chemokine genes. The project focused on CC chemokines because of their functional significance. It has been postulated that only half of the human CC chemokine genes have been cloned to date, and all except four of the twenty-four cloned genes have been mapped in a cluster on chromosome 17 within the region: 17q11.2 (Opdenakker et al., 1994; Naruse et al., 1996). This created a unique opportunity for me to search a known region of the human chromosome 17 for new members of the CC chemokine family using a positional cloning approach.

A new chemokine gene MIP-4 has been identified and its recombinant protein is generated. The protein has shown to have chemoattractant properties for naïve T cell. This would indicate that I have identified a molecule that is critical in the initial stages of immune response.

The results of this study will have three important outcomes. First, it lays the foundation for further characterization of the novel chemokine, MIP-4, and studies of its receptor. These studies will elucidate the function of MIP-4 and provide evidence that it is an important immunoregulator involved in the pathogenesis of HIV. Second, an animal models can be generated by making transgenic/knock-out mice which will serve as an excellent model system for studying the pathogenesis of the diseases associated with MIP-4. Third, by administering the MIP-4 protein directly to mouse models correlated to human diseases, we would be able to explore its therapeutic potential for drug development.
This thesis includes eight chapters. Chapter one summarizes the overall structure of this thesis. Chapter two contains an introduction to chemokines, including genes and their expression patterns, protein structures and functions, and their relationship to human diseases. Chapter three is a literature review of the cloning of chemokine genes. Chapter four explains the construction and characterization of a YAC contig encompassing the CC chemokine cluster region on chromosome 17. Chapter five describes the isolation of chemokine candidate genes from the cluster region through exon trapping and homology analysis. Chapter six summarizes the generation of the MIP-4 recombinant protein and functional in vitro and in vivo assays of this protein, and discusses the studies of MIP-4 expression in natural killer (NK) cells and of its inhibitory effects on HIV infection. Chapter seven describes the cloning of the murine homologue of the human MIP-4 gene. Chapter eight states the conclusions of this project.

1.2 Summary

The goal of my project was to identify novel CC chemokine genes and to study their biological functions. This goal was attained, and the results are outlined as follows:

A YAC (Yeast Artificial Chromosome) contig was first constructed which spanned the cluster region where human CC chemokine genes MIP-1α and MIP-1β reside. MIP-1α gene primer sets were used to isolate YAC clones from the human CEPH mega YAC library by PCR-based screening. Each selected clone was analyzed by vectorette PCR and chromosomal mapping to determine whether it was chimeric. If
both left and right ends of a YAC clone map to the same chromosome, the clone is non-chimerical. Vectorette PCR was used to amplify YAC end fragments using a vectorette primer and the YAC end primer. The PCR products were then sequenced. Chromosomal mapping analyses were performed using the ATCC human-rodent somatic hybrid panel I. Two initially identified non-chimeric clones were used to obtain three additional clones from the YAC library, then aligned to assemble a 1.8 Mb YAC contig which contained both human MIP-1α and -1β genes.

Two candidate chemokine genes were isolated from this region. YAC clones from the contig were used to construct cosmid libraries, which were further screened with human genomic DNA probes. These cosmid clones were then screened with oligonucleotides, which corresponded to conserved chemokine regions. The positive clones were pooled together for exon trapping. Briefly, cosmid DNA was subcloned into the multiple subcloning site (MCS) of an exon-trapping vector. The MCS was flanked by a splicing acceptor and a splicing donor site. If the cloned fragment contains a complete exon in the correct orientation, then a splicing event can occur between the vector and the cloned fragment. Spliced products were then amplified by reverse transcription followed by the polymerase chain reaction (RT-PCR).

Two of the exons sequenced showed extensive homology to the known chemokine genes and to several EST clones in a few databases. Thus, the two exons identified by these methods corresponded to new putative chemokine genes. One was soon identified as HCCI (Schulz-Knappe et al., 1996) and the other, named MIP-4, was present in The Institute of Genomic Research (TIGR) human cDNA database. The predicted protein sequence of the MIP-4 gene consists of 89 amino acid residues and is
closely related to human MIP-1α with 59.6% homology. The full length cDNA of MIP-4 was cloned from a 72-hour culture of monocytes by RT-PCR using primers based on the sequence information from TIGR database. Northern analysis showed that the gene is expressed in lung and lymph nodes. Long range PCR demonstrated that the MIP-4 gene has three exons. Gene mapping studies indicated that the gene is located about 16kb downstream of the MIP-1α gene on chromosome 17, in the opposite orientation.

Recombinant MIP-4 protein was generated and used to study its activity in vitro and in vivo. The MIP-4 cDNA fragment encoding the predicted mature protein was subcloned into a bacterial expression vector pET32 to produce the fusion protein with a 6XHis tag at the N-terminus. The protein was purified using the TALON affinity resin and its tag was cleaved with enterokinase. The purified protein activated both naive and activated T cells in Ca++ flux assays and attracted CD4+ and CD8+ T cells in animal injection studies. This suggested that MIP-4 might play a critical role in the initiation of an immune response.

Finally, MIP-4 gene expression in natural killer cells (NK cells) was studied by RT-PCR. MIP-4 protein's inhibitory effects on HIV infection is being analyzed using synthetic MIP-4 protein in in vitro HIV infection assays. Under the combined stimulation of cytokines IL-12 and IL-15, the NK cells started to express MIP-4 message four hours after incubation. The preliminary results from the in vitro assay of HIV infection have shown that MIP-4 did not inhibit the infection rate of either M-tropic or T-tropic HIV-1.
Based on the close physical distance between the human MIP-4 gene to the MIP-1α gene, I also demonstrated that mouse contains a potential homologue of the human MIP-4 gene by using low stringency hybridization. The positive fragments have been cloned into pBluescript and are being analyzed through sequencing.
Chapter 2

Introduction to Chemokines

2.1 Background

The first human CXC chemokine protein, platelet factor 4, was identified through protein purification followed by amino acid sequencing in 1977 (Deuel et al., 1977). IL-8 was the first cloned chemokine gene (Matsushima et al., 1988). Since then, more than 40 human and mouse chemokines have been added to the family. Chemokines were originally characterized as low molecular weight proteins which had similar primary sequences, and four conserved cysteine residues (Lodi et al., 1994). The chemokine superfamily has been further divided into four groups based on the positions of these cysteine residues: CXC chemokines (X refers to any amino acid separating the first two cysteines), CC chemokines (there is no intervening amino acid between the first two cysteines), C chemokines (there is only one cysteine present in the both pairs of the cysteines) and CXXXC chemokines (there are three intervening amino acids between the first pair of cysteines) (Figure 2.1). These subfamilies are also referred to as the $\alpha$, $\beta$, $\gamma$ and $\delta$ chemokines, respectively.
2.2 Biochemical properties of chemokines

Chemokines are the largest group in the cytokine family. They are inducible secreted proteins, which function as chemoattractants and activators for many different cell types, especially those from the immune system (Schall 1991; Oppenheim et al., 1991). Most of the chemokines in their monomeric forms have a molecular weight ranging from 7 kDa to 15 kDa and are active at concentrations ranging from $10^{-8}$ to $10^{-10}$ M (Oppenheim et al., 1991). However, many chemokines also form dimers, trimers and tetramers upon secretion. The four cysteine residues form two disulfide bonds, which fold the chemokines into loop structures (Matsushima et al., 1992) (Figure 2.2). Chemokines are basic proteins with glycosaminoglycan (GAG)-binding sites. Their carboxyl termini bind heparin. Studies have suggested that chemokines are preferably immobilized by binding to proteoglycans on cell surfaces at inflammation sites or extracellular matrix protein, which recruits lymphocytes to the sites and leads to activation of host defense mechanisms (Rot, A., 1993).

2.3 CXC chemokine (α chemokine family)

Most of the CXC chemokine family members were first identified by cDNA cloning, so the functions of these proteins were initially unknown. They all share a typical amino acid structure. The precursor protein contains a hydrophobic leader peptide of about 20 to 30 amino acid long, which is cleaved after the protein is secreted, leaving the mature protein that ranges from 60 to 80 amino acids. The four-cysteine-motif is conserved in all members, with one intervening amino acid positioned between the first pair of cysteines. They can be further divided into two branches; those
containing the sequence of Glutamic acid—Leucine—Arginine (E-L-R) near the N-terminus (preceding the CXC motif) and those which do not contain the E-L-R motif in their amino acid sequences. They act on distinguished sub-populations of leukocytes. The E-L-R CXC chemokines are chemoattractants to neutrophils, while those without the E-L-R motif act on lymphocytes (Loetscher et al., 1996).

2.4 CC chemokine (β chemokine family)

Thus far, more than twenty human CC chemokine genes have been cloned. Many of them, including the MCP/RANTES and the MIP groups, have been localized and assigned to a cluster region on human chromosome 17q11.2-q12. All CC chemokines have a similar primary structure and contain a conserved four-cysteine-motif. They have a signal peptide of 20 to 24 amino acids, which is cleaved to produce a mature protein of 68 to 76 amino acids with predicted molecular masses of 7-12 kDa. The intraspecies sequence of CC chemokines varies from 28-70% homology at the amino acid level, while interspecies sequence homology varies from 25-55%. Lower similarity exits between CC and CXC chemokines with a homology from 20-40% at the amino acid level (Figure 2.3). The biological roles of many CC chemokines have been identified. They have been shown to be potent monocyte/macrophage but not neutrophil chemoattractants. They also act on basophils, eosinophils, and some T lymphocytes. In addition, they can stimulate target cells to release enzymes, induce intracellular Ca++ movement and inhibit tumor cell growth.
2.5 C chemokine (γ chemokine family)

Lymphotactin (Ltn) and activation-induced, T cell-derived and chemokine-related molecule (ATAC), were both identified in 1995, and are members of the new chemokine family, called “C” chemokines (Kelner et al., 1994; Muller et al., 1995). They contain only two of the conserved four cysteines present in other chemokines. C chemokines share a great deal of sequence similarity with both CXC and CC chemokines, particularly the C-terminal domains of CC chemokines. They are expressed by activated CD8+ T cells (Muller et al., 1995) and natural killer (NK) cells (Hedrick et al., 1997). Lymphotactin is a potent chemoattractant for both lymphocytes and NK cells but not for monocytes or neutrophils (Hedrick et al., 1997).

2.6 CXXXC chemokine (δ chemokine family)

Only one member of the CXXXC (or CX3C) chemokine family has been identified so far. It has a unique cysteine pattern of three intervening amino acids between the first pair of cysteines, and two isoforms have been isolated. One isoform is a membrane-bound protein (Neurotactin) (Pan et al., 1997) and the other isoform is a glycoprotein (Fracttalkine) (Bazan et al., 1997). The membrane-anchored isoform is a potent chemoattractant to neutrophils both in vitro and in vivo, and is expressed in brain tissue. The soluble glycoprotein is active on both monocytes and T cells but not neutrophils, and induces adhesion of these cells to the endothelium. Although the biological significance of CXXXC chemokines is not clear, it has been suggested that they are involved in the inflammation process.
2.7 Three-dimensional structure of chemokines

The first chemokines for which the three-dimensional structure was determined include platelet factor 4 (PF-4) and interleukin 8 (IL-8) (St. Charles et al., 1989; Chore et al., 1990). Since then, the structures of both CXC and CC chemokine members have been determined (Chore et al., 1995). Although their monomeric structures are very similar, the dimer structures of CC and CXC chemokines are entirely distinct (Figure 2.4). They all contain an NH\textsubscript{2}-terminal loop, three antiparallel \(\beta\)-strands connected by loops, and a COOH-terminal \(\alpha\)-helix. However, the three-dimensional structures of MIP-1\(\beta\) and RANTES consist of dimers formed by the interaction of the NH\textsubscript{2}-terminal regions of the monomers, which results in an elongated, cylindrical shape (Lodi et al., 1994; Skelton et al., 1995). The IL-8 dimer is globular in shape and the interface comprises an antiparallel \(\beta\)-sheet (Clore et al., 1995). The dramatic differences of receptor binding sites between these chemokines ensure no receptor cross-binding between the two groups.

2.8 Chemokine receptors

Chemokines induce cell migration and activation by binding to receptors on their target cells. The chemokine receptors are a group of specific G-protein-coupled proteins with seven transmembrane domains (Figure 2.5). The expanding array of chemokine members is reflected in the growing number of chemokine receptors that have been identified (Luster, 1998). Thus far, four human CXC chemokine receptors (CXCR1, CXCR2, CXCR3 and CXCR4), ten human CC chemokine receptors (CCR1, CCR2, CCR3, CCR4, CCR5, CCR6, CCR7, CCR8, CCR9 and CCR10) and one human
CX3CR chemokine receptor (CX3CR1) have been identified. No receptor has been identified for any C chemokine yet. The chemokine receptors share a great deal of similarity between each other. The seven transmembrane domains are most conserved while the variations in extracellular loops determine the chemokine binding specificity and the variations in intracellular loops decide the downstream signal pathway.

2.9 Chemokine ligand-receptor interaction

Many chemokine receptors can have more than one chemokine ligands, however CCRs bind only CC chemokines and CXCRs bind only CXC chemokines. Even more specifically, within the CXCR group, CXCR2 tends to bind to all E-L-R CXC chemokines while CXCR3 binds to non-E-L-R CXC chemokines. Chemokine receptors initiate unique and specific cellular activities only in response to their specific chemokines by relying on the dynamic relationship involving four components: the "presenting" molecule, the chemokine, the chemokine receptor and the G-protein complex linked to the receptor. The expression profiles of chemokine receptors are highly temporally and spatially regulated and cell state-dependent.

Chemokine receptors are expressed on different subsets of leukocytes. Some receptors are widely expressed in different cell types, whereas others are highly restricted to a certain cell type. For example, CXCR1 is predominantly expressed by neutrophils but CCR2 can be found on monocytes, basophils, T lymphocytes, NK cells and dendritic cells. The same chemokine receptors can be constitutively expressed on one type of cells but are inducible on another type of cells. The constitutive expression
can be down-regulated. CCR1 and CCR2 are expressed constitutively on monocytes but on lymphocytes only after induction with IL-2 (Loetscher et al., 1996).

In addition, the expression of chemokine receptors can be regulated by cell activation as well as differentiation. Some non-hematopoietic cells, such as epithelial cell, endothelial cells and neurons, can also express chemokine receptors, which suggests that chemokines play other roles in addition to leukocyte recruitment. There are two additional types of chemokine receptors that are not involved in signaling. One is the erythrocyte chemokine receptor (also called DARC) (Horuk et al., 1993). Both CXC and CC chemokines can bind to this receptor but don't induce Ca++ influx in the target cells, erythocytes and endothelial cells. The other type of receptor is a viral encoded chemokine receptor, whose biological relevance is unclear (Ahuja et al., 1993).

2.10 Chemokine receptors and signal transduction

Similar to many seven transmembrane domain receptors, the chemokine receptors are functionally linked to phospholipases via heterotrimeric G proteins, with downstream generation of inositol trisphosphate, which results in intracellular Ca++ flux, Ca++ channel opening and protein kinase C activation (Bokch et al., 1995) (Figure 2.6). The leukocyte movement in response to chemokines are blocked by pre-treating the cells with Bordetella pertussis toxin (Ptx), a Gi-type G-protein inhibitor by ADP-ribosylation, suggesting that chemokine receptors are linked to G proteins of Gi class. The chemokine receptor signaling pathway in leukocytes also includes activation of small GTP-binding proteins of the Ras, Rac and Rho families, which are involved in cell motility. Chemokine receptors can signal their mitogenic stimuli and regulate cell
differentiation and proliferation by interacting with Ras/Raf/MAP kinase pathway via a tyrosine kinase connection (Bourne, 1995). For example, in T cells, RANTES stimulates chemotaxis at low concentrations. At high concentrations, it induces a variety of cellular responses, including cytokine productions and cell proliferation, normally associated with T-cell activation (Bacon et al., 1995).

2.11 Genomic structures of chemokine genes

All chemokine genes have similar genomic structures. Typically, a CC chemokine gene consists of two introns and three exons while a CXC chemokine gene consists of three intron and four exons (Figure 2.7). The gene structures for human MIPs, MCPs and RANTES, as well as some of their mouse homologs have been characterized. All exhibit a three exon/two intron structure, with Exon1 containing the 5’ untranslated region and signal peptide coding sequence, Exon 2 encoding the N-terminal portion of the mature peptide and Exon 3 for C-terminal region and 3’-untranslated region (Blum et al., 1990; Wilson et al., 1990). The intron-exon junctions among these genes are also highly conserved. There is one additional exon which encodes a few C-terminal amino acids and the 3’ untranslated region present in the CXC chemokine genes. But the splice junctions between CC and CXC chemokine genes are conserved (Miller et al., 1990).

2.12 Chromosomal locations of chemokine genes

The structural differences among chemokines are also correlated with the chromosomal locations of the chemokine genes. It has been shown that chemokine
genes are clustered in particular regions. The human CC and CXC chemokine genes have been mapped to separate regions, mainly on chromosome 17q11.2-q12 and chromosome 4q12-21, respectively (Oppenheim et al., 1991; Opdenakker et al., 1994) (Figure 2.8). More than twenty human CC chemokine genes have been identified and all except four have been mapped to the long arm of chromosome 17 within one region 17q11.2-q12. The only CC chemokines outside these regions are TARC, which has been mapped to chromosome 16q13 (Nomiyama et al., 1997) MIP-3α (also called LARC, exodus), which has been mapped to chromosome 2q33-q37 (Hieshima et al., 1997). MIP-3β and 6Ckine (exodus-2), which have been mapped to chromosome 9 (Ross et al., 1997; Hromas et al., 1997). A YAC contig of chromosome 17q11.2 that contains the most recently identified CC chemokines was constructed (Narruse et al., 1996). Therefore, the specific location of a particular chemokine can be easily determined using sequence-tagged site (STS) markers within the contig by PCR amplifications. SDF-1 was localized on chromosome 10q11.1 by in situ fluorescent hybridization (Shiroozu et al., 1995). This is the only CXC chemokine that has been mapped outside of the 4q12-21 cluster region. The C chemokine gene, lymphotactin, was localized on chromosome 1 through human/hamster hybrid mapping (Kennedy et al., 1995). The corresponding mouse CC chemokine genes are clustered on chromosome 11 and CXC chemokines are clustered on chromosome 5, which are the syntenic regions to human chromosome 17 and chromosome 4, respectively. The mouse SDF-1 gene was localized to chromosome 6 and the lymphotactin gene was localized to mouse chromosome 1 by interspecies backcross analysis (Kelner et al., 1994).
2.13 Regulation of chemokine gene expression

The host response to invasive stimuli is multi-factorial. The recruitment and activation of inflammatory cells to these sites are important steps involved in the process. Chemokines play central roles in the initiation of these steps by serving as mediators of the cellular movement. The expression of most chemokines is inducible, while a few of them can be expressed constitutively. Promoter analysis revealed that there are various cis regulatory elements present in the chemokine promoter/enhancer regions, including SP1 site, NFκB site, AP-1 and AP-3 sites, a GRE element, the sites for interferon regulatory factor-1, hepatocyte nuclear factor-1 and an octamer-binding motif (Mukaida et al., 1989).

2.14 Chemokines in leukocyte trafficking

Chemokines are considered as the directing factors for the movement of leukocytes in development, homeostasis and inflammation. Several chemokine characteristics enable them to function in the selective regulation of leukocytes (Taub, 1996). First, chemokines are transcribed and produced rapidly upon stimulation with a variety of stimuli. Second, the chemokine family consists of a large number of members, which have distinct and overlapping target cells. Third, the localization of the chemokines on the surface of endothelium via proteoglycans and extracellular matrix molecules provide a more efficient presentation to the target cells. Fourth, chemokines can induce a rapid increase of leukocyte adhesion to cellular adhesion molecules (CAMs). Finally, chemokines can form a concentration gradient, which can attract leukocyte to the sites. The concentration, distribution and availability of
Chemokines have been shown to be important factors in promoting cellular adhesion and emigration. The role of chemokines in regulating lymphocyte movement has been determined in studies of mice deficient in a particular chemokine. Among the five major classes of blood leukocytes, the CXC chemokines primarily direct neutrophil movement while CC chemokines are capable of inducing a large variety of cell movement, including monocytes, lymphocytes (including NK cells), basophils and eosinophils. The C chemokine acts only upon lymphocytes (Baggiolini et al., 1997).

2.15 Chemokines in inflammation

The abnormal expression of chemokines has been detected in a wide variety of inflammatory diseases (Baggiolini et al., 1994). The inflammation in tissues is likely due to the chemokine secretion that causes the accumulation and activation of leukocytes in the inflamed tissues. Chemokines have been detected during inflammation in most organs, including skin, brain, joints, lung, blood vessels, kidneys, and gastrointestinal tract and in many cell types of these organs. The major stimuli for chemokine production in these cells are early pro-inflammatory cytokines, such as interleukin-1 and tumor necrosis factor α, bacterial products such as lipoploysaccharide (LPS), and viral infection (Baggiolini et al., 1995). The subgroups of chemokines expressed in a tissue control the type of inflammatory infiltrates that are specific to a particular disease. In many acute disease states, such as bacterial infection, there is a massive infiltration of neutrophils into the tissue, while in chronic diseases, lymphocytes and macrophages infiltrate into tissues.
2.16 Chemokines and infectious diseases

Two types of chemokine receptors serve as co-receptors for human pathogenic organisms. *Plasmodium vivax* binds to the DARC receptor on erythrocytes that cause malaria (Horuk et al., 1993). Different HIV strains bind to two different groups of chemokine receptors for entry into specific target cells: M-tropic HIV isolates bind CCR5 as well as CCR2b, CCR3 (a group of CC chemokine receptors) for infection of macrophages and activated T-cells (Doranz et al., 1996). CXC4 (a CXC chemokine receptor) is the co-receptor for T-tropic HIV strains that infect T-cell lines (Feng et al., 1996). MIP-1α and −1β, and RANTES, ligands for CCR5, block the entry of M-tropic HIV into cells (Cocchi et al., 1995). SDF-1, a ligand for CXC4, inhibits T-tropic HIV infection (Oberlin et al., 1996). The natural killer (NK) cells stimulated with cytokines IL-12 and IL-15 have been shown to express all three chemokines. Even after depletion of the chemokines by specific polyclonal antibodies, the NK cell supernatant still retains some HIV inhibitory effects (Oliva et al., 1998; Fehniger et al., 1998). This suggests that there is as yet unknown HIV inhibition factor(s). Although the molecular mechanism of the fusion event has not been determined, an association between HIV glycoprotein 120, CD4, and a chemokine receptor must occur before HIV enters into cells (Wu et al., 1996). Persons who have homozygous mutations in CCR5 gene are naturally resist HIV infection (Samson et al., 1996).

2.17 Chemokines and inflammation in the central nervous system

The chemoattractant properties of chemokines and their association with inflammatory states suggested a potential role in the immunopathologic diseases such as
experimental autoimmune encephalomyelitis (EAE) and multiple sclerosis (MS) in the central nervous system (CNS). The striking burst of IP-10 and MCP-1 expression correlates with EAE onset (Chaudhuri et al., 1993). MIP-1α has been found (by ELISA) to be elevated in the cerebrospinal fluid (CSF) of MS patients in relapse, and there is a direct correlation between MIP-1α levels and CSF-cell counts, consistent with the possibility that this chemokine is expressed by leukocytes during chronic inflammation in the human CNS (Miyagishi et al., 1995).

2.18 Chemokines in angiogenesis, tumor growth and stem-cell proliferation

Some chemokines are also found to be involved in the regulation of angiogenesis and tumor growth. Platelet factor 4 (PF-4) and IP-10 inhibit neovascularization, tumor growth and metastasis (Sharpe et al., 1990; Luster et al., 1993). It has been suggested that these chemokines may bind to heparin sulfate sites on endothelial cells and prevent other growth factors, such as bFGF, TGF-α from binding to these sites, which would have inhibitory effects on angiogenesis. MIP-1α is known to be a potent factor that inhibits stem-cell proliferation (Graham et al., 1990). MCP-1, -2 and -3 have been isolated from human tumor cell lines, but the roles of these chemokines in tumor development are still unclear (Van Damme et al., 1992).

2.19 Chemokines in animal studies

Animal models of human diseases have been widely used to study the relationship between chemokine expression and pathophysiologic processes of the diseases. Effective antimicrobial host defense require the generation of a vigorous
inflammatory response that involves the recruitment of neutrophils and mononuclear phagocytes, leading to rapid clearance of the inciting organism. Early bacterial or fungal clearance is mediated by a dual phagocytic system involving both neutrophils and mononuclear phagocytes. The recruitment and activation of these leukocyte populations is partially dependent on the establishment of chemotactic gradient, in which chemokines play a major role. A mouse model of *Crypyococcus neoformans* pneumonia was used to study the possible roles of chemokines in the recruitment of mononuclear cells during the development of the disease (Huffnagle et al., 1995). The infiltration of leukocytes into the lung during the pathogenesis of *C. neoformans* pneumonia correlated with the temporal production of MCP-1 in the lung. The depletion of MCP-1 using neutralizing anti-MCP-1 antibodies had profound effects on leukocyte recruitment, with significant attenuation of mononuclear phagocytes and CD4+ T cells.

A mouse model of allergic airway inflammation was used to determine if CC chemokines were potent eosinophil chemoattractants (Lukacs et al., 1996). The temporal expression of both MIP-1α and RANTES is coincident with the recruitment of eosinophils into the airway. Depletion of both chemokines with neutralizing antibodies reduced the eosinophil infiltration into the lungs by more than 50%. These results establish a role for both MIP-1α and RANTES in eosinophil recruitment during the pathogenesis of allergic airway inflammation and suggests that targeting and attenuating the biologic effect of these two CC chemokines may impact the subsequent pathogenesis of this disorder. Direct administrations of chemokine proteins to animals
have revealed a critical role for chemokines in the inflammation process (Meurer et al., 1993; Alam et al., 1994).

2.20 Transgenic/knockout mouse models for chemokine functional studies

The ability to induce chemotaxis is the signature biological characteristic of chemokines. Chemotactic assays demonstrate specific groups of leukocytes migrating toward the gradients of these chemokines in vitro. However, do chemokines regulate leukocyte trafficking in vivo? Do all chemokines have this property? How is the chemokine expression regulated? Why and how do chemokines establish gradients and how eventually are the gradients destroyed? To address these questions regarding the chemokine functions in vivo, genetically engineered mutant animals are developed. Two approaches, gain- and loss-of-function of the gene of interest, were widely employed in generating mouse models. Knockout mouse models have been generated for MIP-1α and SDF-1 genes (Cook et al., 1995; Nagasawa et al., 1996). These studies demonstrated that MIP-1α is required for inflammatory response for viral infection while the homozygous animals showed otherwise normal phenotype. However, missing of SDF-1 gene is embryonic lethal and severely reduced B-cell lymphopoiesis and bone marrow myelopoiesis were found in the mutant embryos. Mice lacking IL-8 receptor gene showed expansion of B cells and neutrophils (Cacalano et al., 1994). A KC transgenic model, which is a CXC chemokine gene under the hGH promoter control showed remarkable infiltration of the thymic cortex by polymorphonuclear cells histologically (Lira et al., 1994). But, their organs and blood cells did not shown any abnormality. The studies have made it clear that a single chemokine can induce
recruitment of a particular class of cells when expressed in a tissue-specific fashion but abolishing signaling through knocking out one kind of receptor can affect functions other than leukocyte recruitment.
Figure 2.1 Amino acid sequences of human chemokines. The sequences are aligned according to their cysteines (framed areas). MIP-1α, MIP-1β, RANTES, MCP-1 and MIP-4 are from CC chemokine subfamily. IL-8, PF-4, SDF-1 and IP-10 are CXC chemokines. Lymphotactin is a C chemokine and Neurotactin is a CXXXC chemokine.
Figure 2.2 Two-loop structure of human chemokine proteins. The disulfate bonds are showed as filled bars. Letters N and C represent the N termini and the C termini, respectively.
Figure 2.3 Structure similarity diagram of human chemokine proteins. The similarity scores of the proteins sequences were determined by Clustal Align method from DNASTAR package. Distance to the branching points indicates the percentage of sequence divergence. The input sequences include C, CC, CXC and CXXXC chemokines retrieved from GeneBank.
Figure 2.4 Three-dimensional structures of human chemokine dimers. The images were adapted from SWISS-3D IMAGE Database (http://expasy.hcuge.ch/cgi-bin/sw3d-search-de). The spirals indicate α helix, the strips represent β-strands while the random lines are for loop structures.
Figure 2.5 Schematic diagram of human chemokine receptor. Each small circle represents an amino acid. The filled circles indicate conserved amino acids in different chemokine receptors while open circles indicate variations. The seven transmembrane domains and the intracellular loops are most conserved.
Figure 2.6 Signaling pathway of chemokines. Arrows denote the pathways of the signals generated by chemokine and chemokine receptor interactions.
Figure 2.7 Genomic organizations and chromosomal localization of human chemokines. Exons are denoted by boxes and introns by lines. The coding regions for signal peptides and the mature proteins are hatched and filled, respectively. The 5' and 3' UTR are open boxes.
Figure 2.8 Cluster regions of human CC and CXC chemokine genes. Chromosomes are represented with G-banding patterns. The cluster regions for CC and CXC chemokine genes are indicated.
3.1 Introduction

New chemokines were initially identified through their chemotactic effects on leukocytes, and they were mainly isolated from blood leukocytes or cell lines. Human IL-8 was first cloned from LPS-stimulated human monocyte mRNA (Matsushima et al., 1988), which encodes a 99 amino acid peptide. It has a signal peptide of 20 amino acid and the mature protein is identical to the protein originally isolated from the culture media of stimulated human blood monocytes (Yoshimura et al., 1987). The first CC chemokine, LD78, was identified by differential hybridization from human tonsillar lymphocytes (Obaru et al., 1986), and a closely related chemokine Act-2 was later identified (Lipes et al., 1988). MIP-1α and MIP-1β were purified from the culture medium of murine macrophages stimulated with LPS (Wolpe et al., 1988) and subsequently cloned (Sherry et al., 1988). Since the murine MIP-1α and MIP-1β are considered to be the homologues of the human LD78 and Act-2 genes, with an amino acid identity of more than 70%, the two human gene were renamed as human MIP-1α and MIP-1β, respectively. MCP-1 is the best characterized CC chemokine. The protein was purified from the culture media of fresh blood monocytes and stable monocyte cell lines (Yoshimura et al., 1989; Matsushima et al., 1989). The MCP-1 cDNA was cloned by different groups from different sources (Furutani et al., 1989; Sica et al., 1990).
More recently, approaches based on the selective cloning of secreted proteins by signal sequence trapping or the use of public/private EST databases through bioinformatic software have allowed the rapid identification of novel chemokines based on their sequence and structural homologies. To date, more than 40 human chemokine genes have been identified (Baggiolini et al., 1997). The various approaches used to clone chemokine genes are described in more details in the following paragraphs. All the cloned human chemokines genes are listed in Table 3.1.

3.2 Identification of chemokine genes according to their protein sequences

Many chemokine proteins were first purified to homogeneity using advanced protein purification and sequencing techniques, and then their amino acid sequences were determined. The first chemokine protein to be characterized was human platelet factor 4 (PF-4), whose amino acid sequence was reported in 1977 (Deuel et al., 1977). Its cDNA was cloned from a human erythroleukemic (HEL) cell-derived λgt 11 expression library using commercial human PF-4 antibody prepared in goats (Poncz et al., 1987).

IL-8 cDNA was cloned by screening a LPS-stimulated human monocyte library with degenerate oligos based on a unique protein sequence in IL-8 (Matsushima et al., 1988). Human MCP-1 cDNA was cloned by screening a glioma cell cDNA library using one pair of degenerate oligos deduced from the amino acid sequence of the N-terminus of the MCP-1 protein (Yoshimura et al., 1989). Mouse MIP-2 cDNA was isolated using degenerate oligos which corresponded to the amino acid residues 9 to 14 of the N-terminal sequence of MIP-2 (Tekamp-Olson et al., 1990).

Both the human MCP-2 cDNA and gene were identified by degenerate PCR amplifications using primers derived from the conserved protein sequences (Coillie et al., 1997a; Coillie et al., 1997b). One set of degenerate primers was deduced from the
CC chemokine consensus sequences and used for PCR amplifications of both cDNA and the genomic region.

3.3 Differential expression cloning of chemokine genes

Chinese hamster GRO-α, a CXC chemokine gene, was cloned through subtractive hybridization of cDNAs from the tumorigenic CHEF/16 cells and the non-tumorigenic CHEF/18 cells (Anisowicz et al., 1987). The cDNAs from CHEF/16 cells were hybridized with cDNAs from CHEF/18 cells and the cDNAs that did not hybridize were used as probes to screen the CHEF/16 cDNA library. Three human GRO genes were cloned using a similar strategy (Haskill et al., 1990). Mouse Mig, a CXC chemokine gene, was cloned by differential screening (Farber, 1990). The cDNA library generated from the stimulated mouse RAW 264.7 cell line (a mouse monocyte/macrophage cell line) was screened with a cDNA probe from stimulated RAW 264.7 cells and a cDNA probe from unstimulated RAW 264.7 cells. The clones present in the stimulated but not in the unstimulated RAW 264.7 cells were isolated and sequenced.

3.4 Low stringency hybridization

Many mouse and human chemokine genes were cloned by low stringency hybridization using the known chemokine cDNAs as probes. This was possible because of the high degree of sequence conservation among intra- and inter-species chemokine genes. Human Mig, SDF-1 and lymphotactin were identified using corresponding mouse cDNA probes (Farber, 1993; Kennedy et al., 1995; Shirozu et al., 1995). Both human MCP-5 and mouse MCP-5 genes were isolated using human MCP-4 cDNA as the probe (Sarafi et al., 1997) and human MCP-4 was isolated using a genomic fragment of from the human eotaxin gene (Garcia-Zepeda et al., 1996).
3.5 Random sequencing of cDNA clones

It is possible to identify novel genes random sequencing of clones from a particular library because of the rapidly improvement of DNA sequencing techniques and automation. During the analysis of cytokine expression profiles of freshly-sorted and activated pro-T cell cDNA libraries, a clone which represented a new branch of chemokine, the C chemokine, was identified and named Lymphotactin (Kelner et al., 1994). Two novel CC chemokines, DC-CK-1 and TECK, were cloned through randomly sequencing cDNA clones from lung dendritic cells and thymus cells, respectively (Adema et al., 1997; Vicari et al., 1997). A new chemokine, representing another branch of the chemokine family, CXXXC chemokine, neurotactin, was identified by studying a mouse choroid plexus cDNA library through automated high-throughput single-pass sequencing and computer-aided analysis (Pan et al., 1997).

3.6 Cloning chemokine gene through signal sequence trapping

Since chemokines are secreted proteins with a hydrophobic signal peptide that directs the secretion of the precursor protein, their cDNAs can be cloned through signal sequence trapping systems. The first chemokines to be identified by this strategy were mouse storm cell-derived factor-1 (SDF-1_ and SDF-1_) from a bone marrow stormal cell line, ST-2 (Nishikawa et al., 1988). Basically, the trapping vector was engineered so that a reporter gene product can be monitored on host cell plasma membrane or in the medium. When the cDNA inserts containing the signal peptide coding sequences were cloned in-frame with the correct orientation in the trapping vectors, the reporter gene products can be detected either by immunological screening (Tashiro et al., 1993), or by genetic selection (Jacobs et al., 1997).
3.7 Cloning chemokine genes through database searching

Recently, the methodology advances in two areas have dramatically improved the speed of the identification of novel chemokine genes. First, the human EST (expressed sequence tag) projects have generated a huge database which is a collection of identification sequences or tags for the majority of expressed human genes. These sequences are usually correspond to fragments which are approximately 250 bp to 500 bp in length. Chemokines are small peptides that are encoded by cDNA fragments around 300bp in length. Therefore, the entire chemokine cDNA is often found in a single EST clone. There are presently over 600,000 EST sequences available on public databases, which is an enormous resource for the identification of novel chemokine genes. Second, the rapid development of the WorldWideWeb, which provides communication between various research sites, allows researchers to search the shared databases from long distances without having the traditional complications in maintaining the databases themselves. It also allows access to a number of automated searching tools, and new sequences can be identified by servers and automatically sent to an individual researcher promptly. This approach has lead to the rapid identification of more than 10 new chemokine genes, including one CXC chemokine leukocyte adhesion inhibitor-1/T39765, eight CC chemokines PARC/R83915 (DC-CK-1, MIP-4, AMAC-1), MIP-5/R91733, HCC-1/Z49270, T58847, D31065, MCP-4/T64262, W17274 and W05519, and one CXXXC chemokine Neurotactin/Z44443 (Wells et al., 1997). Most recently, a novel CXC chemokine was identified, B-lymphocytes chemoattranct (BLC), by in situ hybridization of mouse tissue with antisense transcripts of ESTs that showed homology to chemokines (Gunn et al., 1998). It is the first chemokine to be identified that is selective towards B cells.
Table 3.1 A list of the cloned human chemokine genes. Multiple names for the same gene are in ( ) next to the GenBank nomenclature for that gene. This table was assembled from the chemokine database at: http://cytokine.medic.kumamoto-u.ac.jp/CFC/CK/Chemokine.html
<table>
<thead>
<tr>
<th>Chemokine Subfamily</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>C chemokine</td>
<td>Lymphotactin (SCM-1α)</td>
</tr>
<tr>
<td></td>
<td>SCM-1β</td>
</tr>
<tr>
<td>CC chemokine</td>
<td>MIP-1a (LD78a, LD78b, SCI, AT464.1, AT464.2)</td>
</tr>
<tr>
<td></td>
<td>MIP-1b (Act-2, HC21, AT744.1, AT744.2)</td>
</tr>
<tr>
<td></td>
<td>MIP-1g (MRP-2, CCF18)</td>
</tr>
<tr>
<td></td>
<td>MIP-3 (CKb8, MPIF-1)</td>
</tr>
<tr>
<td></td>
<td>MIP-3a (exodus-1, LARC)</td>
</tr>
<tr>
<td></td>
<td>MIP-3b (exodus-3, ELC, CKb11)</td>
</tr>
<tr>
<td></td>
<td>MIP-4 (DC-CK-1, PARC, AMAC-1)</td>
</tr>
<tr>
<td></td>
<td>MIP-5 (MIP-1d, HCC-2, Lkn-1)</td>
</tr>
<tr>
<td></td>
<td>HCC-1 (HCC-3)</td>
</tr>
<tr>
<td></td>
<td>HCC-4 (LEC)</td>
</tr>
<tr>
<td></td>
<td>MCP-1 (JE, MCAF)</td>
</tr>
<tr>
<td></td>
<td>MCP-2 (HC14)</td>
</tr>
<tr>
<td></td>
<td>MCP-3 (MARC, FIC, NC28)</td>
</tr>
<tr>
<td></td>
<td>MCP-4 (CKb10)</td>
</tr>
<tr>
<td></td>
<td>MCP-5</td>
</tr>
<tr>
<td></td>
<td>RANTES</td>
</tr>
<tr>
<td></td>
<td>MRP-1 (C10)</td>
</tr>
<tr>
<td></td>
<td>I-309 (TCA3)</td>
</tr>
<tr>
<td></td>
<td>6Ckine (SLC, exodus-2, TCA4)</td>
</tr>
<tr>
<td></td>
<td>eotaxin</td>
</tr>
<tr>
<td></td>
<td>Clone 391</td>
</tr>
<tr>
<td></td>
<td>STCP-1 (MDC)</td>
</tr>
<tr>
<td></td>
<td>TARC</td>
</tr>
<tr>
<td></td>
<td>TECK</td>
</tr>
<tr>
<td>CXC chemokine</td>
<td>PF-4</td>
</tr>
<tr>
<td></td>
<td>GRO-a (MGSA-a)</td>
</tr>
<tr>
<td></td>
<td>GRO-b (MGSA-b, MIP-2a)</td>
</tr>
<tr>
<td></td>
<td>GRO-g (MIP-2b)</td>
</tr>
<tr>
<td></td>
<td>ENA-78 (LIX, AMCF-II)</td>
</tr>
<tr>
<td></td>
<td>GCP-2</td>
</tr>
<tr>
<td></td>
<td>IL-8 (NAP-1, MDCF, AMCF-1)</td>
</tr>
<tr>
<td></td>
<td>NAP-2 (PBP, b-GF1, bGF-2, CTAPIII)</td>
</tr>
<tr>
<td></td>
<td>NAP-4</td>
</tr>
<tr>
<td></td>
<td>BLC (BCA-1)</td>
</tr>
<tr>
<td></td>
<td>Mig</td>
</tr>
<tr>
<td></td>
<td>IP-10 (crg-2, mob-1)</td>
</tr>
<tr>
<td></td>
<td>SDF-1 (PBSF)</td>
</tr>
<tr>
<td></td>
<td>I-TAC (H174)</td>
</tr>
<tr>
<td>CXXXC chemokine</td>
<td>Fractalkine (Neurotactin)</td>
</tr>
</tbody>
</table>

Table 3.1
Chapter 4

Construction and Characterization of a YAC contig encompassing the CC Chemokine Cluster Region on Chromosome 17

4.1 Introduction

Chemokine genes are located in clusters in both humans and mice. Their structural as well as functional characteristics are correlated with the chromosomal locations of these genes. The CC chemokines, which preferentially activate monocytes, basophils, eosinophils, and T cells, map to human chromosome 17q11-q12 (Opdenakker et al., 1994) while the CXC chemokines which mainly attract neutrophils map to human chromosome 4q12-q21 (Damme, 1994).

The mouse CC chemokine genes are clustered on chromosome 11 and CXC chemokines on chromosome 5, which are the syntenic regions of human chromosome 17 and chromosome 4, respectively. Positional cloning is a traditional approach to identify a gene based on its chromosomal map position without knowledge of its protein product. Basically, it involves four steps: mapping the gene(s) of interests to a chromosomal region, constructing genetic/physical maps of the candidate region, searching the region for conserved/homologous or transcribed sequences, then identifying and characterizing the candidate gene(s). If the gene of interest is associated
with a disease phenotype, then it is essential to identify mutation(s) within the gene, which is present in affected persons but not in the normal population. Positional cloning is a powerful tool which has led to the cloning of more than fifty genes for human diseases, including retinoblastoma (Friend et al., 1986), Duchenne muscular dystrophy (Ray et al., 1985; Monaco et al., 1986; Burghes et al., 1987), and cystic fibrosis (Rommens et al., 1989), Huntington’s disease (Huntington’s disease collaborative research group HDCRG, 1993) and breast cancer (Miki et al., 1994), and it is currently being used to clone genes responsible for early onset Alzheimer’s disease (Campion et al., 1995), stroke (Joutel et al., 1997) and prostate cancer (Chekmareva et al., 1997). Narrowing down the location of a gene on a particular chromosomal subregion greatly facilitates the search for candidate genes.

Positional cloning is a very useful tool for CC chemokine gene cloning. Conventional approaches, such as subtractive screening, differential hybridization, or protein purification, all depend on the expression of the genes of interest in special tissues or cells. Positional cloning does not require previous knowledge about the expression profile or function of the gene. Combined with efficient gene trapping techniques and powerful database analysis, positional cloning could lead to isolation of multiple genes within the gene cluster. In addition, a thorough search of the cluster region encompassed by the YAC contig could lead to clone all novel genes in a cluster. Positional cloning utilizes genomic DNA as the direct cloning source, which allows easy access to the genomic structure and promoter sequence of newly cloned genes. The promoter sequence often provides important information on the expression pattern and
regulation of chemokine genes, which is the key to understand the roles that the individual chemokines play in the immune responses.

It has been shown that chemokines play a key role in directing lymphocyte traffic during immune responses. Recent investigations revealed that the abnormal expression of some chemokines is associated with inflammatory diseases such as psoriasis, asthma, rheumatoid arthritis and respiratory/lung disorders. Chemokines also promote wound healing, inhibit tumor cell growth, and are involved in the regulation of viral infection. CC chemokines produced by CD8+ T lymphocytes, such as MIP-1α, MIP-1β and RANTES, have been identified as the major HIV-suppressive factors. Several chemokine receptors, CCR-1, -2, -3, -5, have been found to be the cofactors for HIV infection. The cloning of every new member from this gene family provides further knowledge of the important functions of chemokines and could potentially unravel the pathogenesis of a variety of diseases, which would eventually lead to the discovery of therapeutic interventions.

My overall goal was to identify novel chemokine genes because of their functional significance. It has been postulated that only half of the human CC chemokine genes have been cloned to date. All CC chemokines except for four have been mapped to chromosome 17 within one clustered regions: 17q11.2-q12. This created a unique opportunity for me to search a known region of the human chromosome 17 for new members of the CC chemokine family using a positional cloning approach.
In this chapter, I describe the construction and further characterization of a YAC (Yeast Artificial Chromosome) contig. This contig is about 1.8 Mb long and spans the cluster region where human CC chemokine genes MIP-1α and MIP-1β reside.

4.2 Materials and Methods

4.2.1 Screening the CEPH mega-YAC library

The human CEPH mega-YAC DNA pools (Research Genetics, Huntsville, Al) were screened by the polymerase chain reaction (PCR) amplification using primer sets that amplify the human MIP-1α gene. Two MIP-1α primer sets were synthesized: 5'-TGCAGCTCTGTGGGAAGGAATGG-3'; 5'-TTTATAGGCCAGCCCTGGCGGATG-3' derived from the 5' untranslated region and exon 1, and 5'-TTCTCCACAGTTGCTGCTGACACG-3'; 5'-TTATCATGACACCCTGGGCTTGAGCA-3' derived from exon 2. PCR was performed in either 25μl or 50μl reactions containing 20ng of DNA sample, 50nM of each primer, 1X Taq polymerase buffer (10mM Tris-HCL, 50mM KCL, pH 9.0, 200μM dNTPs, 2.5mM MgCl₂) and 2.5 units of Taq polymerase (GIBCO BRL, Gaithersburg, MD). The amplification was started with a 3 min denaturization at 94°C, followed by 35 cycles of amplification at 94°C for 30 sec, 60°C for 30 sec, 72°C for 45 sec, with a final extension at 72°C for 7 min in a Perkin Elmer Cetus DNA Thermal Cycler (Perkin Elmer; Norwalk, CT). The PCR products were verified by running 10μl of each PCR reaction on a 1.5% agarose gel and checking for the presence/absence of the expected size band. The screening started with the DNA block pools from the YAC library, in which each sample contains DNA isolated from YAC clones collected from eight 96-well-microtiter plates. The second round of PCR
amplification screened the DNA plate pools, in which each sample contains DNA collected from either every column or every row of the eight plates. The PCR results provided the plate address, the column address and the row address of all positive clones, which were then ordered from Research Genetics (Huntsville, Al).

4.2.2 Preparation of chromosomal DNA from yeast

YAC chromosomal DNA for pulse field gel electrophoresis (PFGE) analysis was prepared according to a previously established protocol (Carpten et al., 1994). Briefly, each YAC clone was grown in 5ml of yeast minimal media lacking uracil and tryptophan (0.67% w/v yeast nitrogen base without amino acids, 2% w/v glucose, appropriate amino acid supplements) at 30°C with vigorous shaking for two to three days until O.D₆₀₀ reached 0.8 to 1.0. The cell pellet was washed once with 50mM EDTA (pH 8.0) and then resuspended in 100ul of EDTA (50mM, pH 8.0) with 200ng/ml of Zymolase 70T (ICN, Irvine, CA). One hundred microliters of 2% of a low-melt agarose (FMC/MCD, Rockland, MD) in 50mM EDTA (pH 8.0, cooled to 55°C) was added to the cell suspension and mixed well. A 100ul aliquot of each sample was immediately transferred into a plug mold (Bio-Rad, Hercules, CA) and incubated at -20°C for 5 minutes before being removed from the mold. The agarose plugs from the same cell culture were equilibrated in 40ml of spheroplasting buffer (0.5mM EDTA, pH8.0, 7.5% v/v β-mercaptoethanol) in a sealed petri dish at 37°C overnight and then replaced with 10 ml of ESP lysis buffer (0.5M EDTA pH8.0, 1%w/v sarkosyl and 0.5mg/ml proteinase K) for two days at 50°C. Plugs were stored at 4°C in 0.5M EDTA (pH8.0) for up to one month.
4.2.3 **Determination of YAC sizes by PFGE**

YAC chromosomal DNAs were separated by pulse field gel electrophoresis (PFGE) (Carle et al., 1986). The YAC plugs containing the high molecular weight DNA isolated from individual YACs were loaded (sealed with low-melt agarose after loading) in 1% agarose gel (prepared in 0.5% TBE without ethidium bromide). The plugs were electrophoresed for 48 hours at 200V with ramped switch times of 60 seconds to 90 seconds (optimal for separation range from 400Kb to 2000Kb) in 0.5XTBE at 14°C in a CHEF-DR™ II Electrophoresis apparatus (Bio-Rad, Hercules, CA). The gel was stained with ethidium bromide, destained in deionized water, and photographed. The resolved yeast chromosomes range from 250kb to 2200kb and served as molecular weight markers.

The DNA in the gel was transferred to Hybond-N+ nylon filters (Amersham, Clearbrook, IL) by a standard protocol (Sambrook et al., 1982) and baked at 80°C for 2 hours, and hybridized to human genomic DNA probes to detect human inserts. The high molecular weight DNA was nicked by incubating the gel in 0.25M HCL with gentle shaking for 5 minutes. The gel was incubated in denaturization buffer (5M NaCl, 0.5MnaOH) for 30 minutes at room temperature, and transferred to neutralization buffer (1M Tris-HCL, pH7.4, 1.5M NaCl) for 30 minutes at room temperature.

The hybridization probes were labeled with $\alpha$-$^{32}$P [dCTP] by the random primer method (Feinberg et al., 1984). One hundred nanograms of human genomic DNA in 20ul TE was denatured by boiling for 3 minutes, then 2.5 units of the random primer (pdN6) was immediately added to the DNA sample. The mixture was cooled down to 37°C to allow the priming of the DNA template. The followings were added in order:
20ul of 2.5X RPLB (0.5M HEPES, pH6.6, 12.5mM MgCl2, 25mM β-mercaptoethanol, 125mMTris-HCL, pH8.0, 50uM of each dATP, dGTP and dTTP), 2ul of 100XBSA (New England BioLabs, MA), 5ul of a-32P [dCTP] (Amersham, Clearbrook, IL) and 2.5unit of Klenow DNA polymerase (USB, Cleveland, OH). The sample was mixed briefly and incubated at 37°C for 1 to 2 hours. The reaction was terminated by adding 50ul of TE-0.1%SDS. Unincorporated radioactive label was removed by running the sample through a Sephadex G50 column. The incorporation rate of the labeling was measured by scintillation counting.

The blot was hybridized with 32P-labeled human genomic DNA probes. Prior to hybridization, the blot was incubated at 42°C in 50ml of prehybridization buffer (6XSSPE, 40% formamide, 2%SDS, 0.5% dry milk) for 2 hours. The blot was then hybridized in 20ml of hybridization buffer (prehybridization buffer plus 10% dextran sulfate) with 1X10^7 cpm 32P labeled probe. The hybridization was carried out overnight at 42°C. The filter was washed once in 2XSSC-0.1%SDS for 10 minutes at room temperature, followed by 15 minute washes in 2XSSC-0.1%SDS, in 1XSSC-0.1%SDS, in 0.5XSSC-0.1%SDS and in 0.2XSSC-0.1%SDS at 68°C with monitoring. The filter was autoradiographed for two days at -70°C.

4.2.4 Preparation of DNA from yeast for PCR

Total low molecular weight yeast DNA was prepared from YAC clones (Carpten, et al., 1994). Yeast culture from a single YAC clone was prepared as previously described. Cells were centrifuged at 1000xg for 5 minutes and washed once with 50mM EDTA (pH8.0). The cell pellets were resuspended in 2ml of SPEM buffer
(1M sorbitol, 100mM sodium phosphate buffer pH 7.4, 60mM EDTA pH 8.0, 90mM β-mercaptoethanol freshly added) and 0.5mg/ml Zymolase (ICN, Irvine, CA) and incubated at 37°C for 1 to 2 hours until 90% of yeast cells formed spheroplasts. The spheroplasts were harvested at 600xg for 5 minutes, resuspended in 700ul of yeast lysis buffer and incubated at 68°C for 15 minutes. The lysate was extracted twice with 1 volume of ice-cold isopropanol and the DNA pellet was collected by centrifugation in a microfuge at full speed at 4°C for 5 minutes. The DNA was washed once with 70% ethanol and dried in a speed vacuum. The DNA was then resuspended in 200-300ul of TE (10mM Tris-HCL, pH8.0, 1mM EDTA pH8.0) with 20mg/ml RNase (Amersham, Cleveland, OH) and incubated at 68°C for 15 minutes. DNA was precipitated with 1 volume of isopropanol and 1/10 volume of 3M sodium acetate (pH 5.2) and centrifuged at 16000Xg for 5 minutes. DNA pellet was washed twice with 70% ethanol, briefly dried and resuspended in 50ul of TE (pH 8.0).

4.2.5 Isolation of YAC chromosome ends by Vectorette PCR

The left and right end sequences of each YAC clone were cloned and mapped to a chromosomal location to determine if any were chimeric. Vectorette PCR (also called bubble-PCR) was used to clone the YAC ends (Riley et al., 1990). Briefly, prepared YAC DNA plugs, each of which was cut into three equal slices, were equilibrated in 50 ml of TE buffer with gentle agitation for 2 hours at 4°C. The TE buffer was changed once, and then the plugs were equilibrated overnight at 4°C. One third of each plug, or one slice, was used for each digestion with a blund-end enzyme. The slices were equilibrated in 500ul of the appropriate restriction buffers at 4°C for 2
hours, and then digested with 20U of either EcoRV, HincII, or RsaI in 100ul fresh restriction buffer for 2 hours at 37°C. The slices were equilibrated in 1X ligation buffer (50mM Tris-HCL pH 7.6, 10mM MgCl₂, 1mM DTT) for 2 hours at 4°C. The buffer was replaced with 100ul fresh ligation buffer with 2 pmole blunt-ended annealed vectorette adapters (top strand 5'-AAGGAGAGGACGCTGTCTGTCGAAGGTA AGGAACGGACGAGAGAAGGGAGG-3', bottom strand 5'-CTCTCCCTTCTCGAA TCGTAAACCCTCTACGAGAATCGCTGCTCTCTCTCCTT-3'). The tubes were heated to 65°C for 15 minutes, cooled down to 37°C, and ATP was added to a final concentration of 1mM. Five units of T4 ligase (New England BioLabs, Beverly, MA) was added to the ligation mix, which was incubated for 1 hour at 37°C and diluted to 500ul with sterile water. Vectorette libraries were stored in aliquots at -20°C.

Five microliters of each Vectorette library was amplified with primers 1089 (YAC left-arm primer 5'-CACCCGTTCTCGGAGCACTGTCCGACCGC-3') and 224 (universal vectorette primer 5'-CGAATCGTAACCGTGTTACGAGAATCGCTG-3') or primers 1091 (YAC right-arm primer 5'-ATATAGGCGCCAGCAACCGCACC TGTTGGCG-3') and 224 in a 50ul 1X PCR buffer. The amplification was initiated by 3 minutes of denaturization at 94 °C, followed by 35 cycles of amplification at 94°C for 1 min, 60°C for 1 min, 72°C for 3 min, with a final extension at 72°C for 7 min. Ten microliters of each PCR product was run out in a 1.5% agarose gel.

Vectorette PCR products were purified using the Wizard PCR purification kit (Promega, Madison,WI) and directly ligated into the TA vector according to the manufacturer’s protocol (Invitrogen, CA). White colonies were grown up in 5ml of LB media (with 50μg/ml of amp) and plasmid DNA was isolated for sequencing analysis.
4.2.6 DNA sequencing

Plasmid DNA was purified using Wizard mini-prep system (Promega, Madison, WI) and then sequenced using vector-specific primers T7 promoter and M13 reverse primer. Sequencing reactions were carried out with the Sequenase Version 2.0 sequencing kit according to the manufacturer’s protocol (Amersham, Cleveland, OH).

4.2.7 Mapping YAC ends

PCR primer sets developed from the YAC end sequences were used in chromosomal mapping. The human-rodent hybrid mapping panel I (ATCC, Manassas, VA), which contains a different set of human chromosomes in each hybrid cell line, was amplified with the primer sets developed from each YAC (Table 2). One hundred nanograms of DNA from each hybrid was used in each 50ul PCR reaction with 1X Taq polymerase buffer and 2.5 units Taq (GIBCO BRL, Gaithburg, MD), and 50nM of each primer. The PCR cycling conditions consisted of 3 min of denaturization at 94 °C, followed by 35 cycles of amplification at 94°C for 30 sec, 50°C to 55°C for 30 sec, 72°C for 30 sec, with a final extension at 72°C for 7 min in a Perkin Elmer Cetus DNA Thermal Cycler (Perkin Elmer; Norwalk, CT). The PCR products were verified by running 10ul of each PCR reaction on a 1.5% agarose gel to evaluate the presence/absence of the expected bands from each hybrid cell line. Both positive and negative amplifications were used to map a YAC end to a particular chromosome.
4.2.8 Assembly of a YAC contig

The primer sets derived from non-chimeric YAC clones were used to assemble a YAC contig containing the human MIP-1α gene. The primer sets from one YAC clone were used to amplify other YACs. The primer set, which amplified the expected products, linked two YACs together. The primer set, which generated negative results, identified non-overlapping clones. The non-overlapped primer set from the linked clones was used to expand the YAC contig in both directions as previously described. The additional clones isolated were further characterized as described and arranged into the contig.

4.2.9 Analysis of the YAC contig for the presence of the known CC chemokine genes

To check for the presence of the known CC chemokine genes in the assembled YAC contig, a series of Southern hybridization was performed by using probes from known human CC chemokine genes. The probes included MIP-1α, MIP-β, RANTES, I-309, MCP-1, MCP-2 and MCP-3 gene fragments. DNA isolated from individual YAC clones was digested with EcoRI, separated in a 1.0% agarose gel, and transferred to Hybond-N+ filters. The filters were hybridized to each probe and scored for the presence or absence of a particular chemokine gene.

4.3 Results

MIP-1α was the first cloned human CC chemokine gene. Since MIP-1α maps to 17q11.2-12 (Nakao et al, 1990), it was used as a marker to construct the YAC contig.
A PCR screening strategy was used to isolate YAC clones that contained the MIP-1α gene. Three sets of PCR primers were designed which flanked the different regions of MIP-1α gene. Among these, primers amplifying exon 2 gave a single band with the expected size of 117 bp in trial PCR reactions with human genomic DNA. The exon 2 primer set was used to screen the DNA pools of the human CEPH mega-YAC library by PCR (Research Genetics Inc., AL). Five positive clones, 737-d-3, 756-d-6, 770-b-11, 801-g-4, and 803-a-2, were isolated.

Since the human DNA inserts in mega-YAC clones are quite large (300Kb to 2000Kb) and rich in repetitive sequences, they are unstable and prone to deletions (Ling et al., 1993). In some circumstances, one yeast colony could harbor more than one YAC chromosome, which could complicate further analysis if it is not detected. Pulse field gel electrophoresis (PFGE) (Carle et al., 1986) was used to separate huge YAC DNA fragments. The size(s) and the numbers of YAC chromosomes within one YAC clone were determined by Southern hybridization. The results demonstrated that only one YAC chromosome was present in each of the five isolated YAC clones, with sizes of 870Kb (737-d-3), 830Kb (756-d-6), 1400Kb (770-b-11), 1420Kb (801-g-4) and 1520Kb (803-a-2) (Figure 4.1).

Most total-human-genomic YAC libraries contain a high proportion (typically 20% to 40%) of chimeric clones, or clones carrying two or more non-contiguous genomic DNA segments (Green et al., 1991). To determine if any of the YAC clones were chimeric, the left and right end sequences were cloned and map to a chromosomal location. Vectorette PCR (also called bulb-PCR) was used to clone the YAC ends of the five YAC clones (Riley et al., 1990) (Figure 4.2). The first round of
PCR amplification had to be primed by a YAC end primer because the vectorette primer has an internal sequence which is identical, but not complementary to the vectorette adapter. This ensures that only the fragments adjacent to YAC arms can be amplified. Three blunt-end enzymes, RsaI, EcoRV and HincII, were used separately in partial digestions of each YAC clone to produce end fragments with sizes suitable for standard PCR amplification. YAC left- and right-end primers were used to amplify each end fragment. Vectorette PCR products were directly cloned into TA cloning vectors and sequenced. PCR primer sets were designed based on the YAC end sequences, and were used in chromosomal mapping studies to determine the chimeric status of each YAC clone. DNA samples from the human-rodent somatic hybrid mapping panel I, in which each hybrid contains a different set of human chromosomes, were amplified by primer sets from each YAC clone. Analysis of these results led to the chromosomal assignment of each YAC end (Table 4.1). If both ends mapped to the same chromosome, the YAC was considered non-chimeric. Among the five clones isolated, two (737-d-3 and 803-a-2) were found to be non-chimeric, and both of their ends mapped to chromosome 17. YAC clone 801-g-4 was chimeric, with a right end which mapped to Chromosome 17 and a left end which mapped to chromosome 1 (Table 4.2). The right ends of the remaining clones 756-d-6 and 770-b-11 were mapped to chromosome 17, but their left ends were not attained through vectorette PCR.

The two non-chimeric clones, 737-d-3 and 803-a-2, served as the core to assemble the YAC contig spanning the huMIP-1α region. The non-overlapping PCR primer set from each clone was used to expand the YAC contig in both directions. The same PCR screening strategy was used as described above to isolate three additional
clones. Clones 756-d-2 and 786-h-7 extended from the 737-d-3 side, while clone 784-f-11 extended from the 803-a-2 side. The whole contig spanned approximately 1.8Mb (Figure 4.3).

To check for the presence of known CC chemokine genes in the assembled YAC contig, a series of Southern hybridization were performed using the probes from the known human CC chemokine genes (Figure 4.4). The results showed that the 1.8Mb contig spans the MIP-1α and MIP-1β region, but not the MCP-1, MCP-2, MCP-3 region. This result indicated that the contig covers the cluster region on 17q11.2-q12 where both MIP-1α and MIP-1β reside.

4.4 Discussion

More than 20 human CC chemokine genes have been identified thus far. They include MIP-1α, -1β, -1γ (MRP-2), MIP-3α (LARC, exodus-1), -3β, C10(MRP-1), RANTES, I-309, Eotaxin, MCP-1, -2, -3, -4, -5, DC-CK-1 (PRAC, MIP-4, AMAC), TARC, TECK, CKb6 (Eotaxin-2), 6Ckine (exodus-2), MDC, HCC-1, -2 (MIP-5), HCC-3, HCC-4. Most of these CC chemokine genes have been mapped to chromosome 17q11.2-q12 region (Opdenakker, et al. 1994). A few CC chemokines reside outside of these clusters, including TARC, which has been mapped to chromosome 16q13 (Nomiyama et al., 1997), MIP-3γ (also called LARC, exodus), which has been mapped to chromosome 2q33-q37 (Hieshima et al., 1997), MIP-3β and 6Ckine (exodus-2), which maps to chromosome 9 (Rossi et al., 1997; Hromas et al., 1997). Analysis of the human EST databases and cytokine websites indicated that there may be unidentified CC chemokine genes in the cluster regions. Recently, a YAC contig of chromosome...
17q11.2 which contains newly identified CC chemokines was constructed (Narruse et al., 1996). A search of the YAC databases at CEPH-Genethon (http://www.cephb.fr/ceph-genethon-map.html) and the Whitehead Institute Genome Center/MIT (http://www-genome.wi.mit.edu) revealed that the 1.8Mb YAC contig constructed was positioned between D17S933 and D17S1995. It partially overlapped the published YAC contig and extended towards the telomeric side of the cluster region (Figure 4.5). This contig contained both MIP-1α and -1β genes, but no other CC chemokines were identified in this region yet. The construction of this YAC contig laid the groundwork for the identification of new human CC chemokine gene(s) from this region.
| HYBRID/DNA     | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15  | 16 | 17 | 18  | 19  | 20  | 21  | 22  | X   | Y   |
|---------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|------|----|----|------|------|------|------|------|------|-----|-----|
| GM/NA09925    | 74 | 24 | 0  | 74 | 76 | 60 | 82 | 78 | 0  | 0  | 4  | 68 | 6  | 86 | 78  | 14 | 98 | 96   | 46   | 84   | 0    | 76  | 0   | 0   |
| GM/NA09926    | 69 | 75 | 75 | 65 | 2  | 88 | 85 | 69 | 0  | 68 | 0  | 2  | 77 | 73 | 93  | 2   | 81 | 75   | 84   | 96   | 2    | 4   | 2   | 0   |
| GM/NA09927    | 69 | 83 | 75 | 77 | 0  | 93 | 79 | 73 | 0  | 82 | 0  | 0  | 77 | 79 | 90  | 0   | 81 | 73   | 87   | 89   | 0    | 0   | 0   | 0   |
| GM/NA09928    | 0  | 84 | 58 | 0  | 48 | 32 | 0  | 66 | 0  | 2  | 0  | 0  | 4  | 76 | 92  | 0   | 98 | 0    | 28   | 0    | 70   | 82  | 0   | 78  |
| GM/NA09929    | 0  | 0  | 61 | 59 | 0  | 43 | 2  | 49 | 0  | 0  | 33 | 49 | 0  | 59 | 2   | 0   | 96   | 2    | 31   | 0    | 0   | 2   | 0   |
| GM/NA09930A   | 0  | 34 | 62 | 4  | 12 | 0  | 26 | 4  | 0  | 0  | 6  | 22 | 56 | 82 | 12  | 0   | 86   | 78   | 0    | 22   | 82  | 76  | 6   | 8   |
| GM/NA09931    | 0  | 0  | 0  | 12 | 0  | 78 | 0  | 0  | 46 | 0  | 64 | 0  | 100 | 0   | 0   | 100  | 0    | 0    | 78   | 90  | 0   | 0   | 14  |
| GM/NA09932    | 0  | 0  | 0  | 26 | 0  | 78 | 0  | 0  | 46 | 0  | 64 | 0  | 100 | 0   | 0   | 100  | 0    | 0    | 78   | 90  | 0   | 0   | 14  |
| GM/NA09933    | 0  | 0  | 0  | 68 | 86 | 46 | 0  | 80 | 0  | 2  | 28 | 26 | 0  | 0   | 0    | 0    | 96   | 2    | 0    | 92   | 0   | 0   | 0   |
| GM/NA09934    | 50 | 0  | 84 | 16 | 54 | 76 | 92 | 54 | 0  | 6  | 0  | 50 | 84 | 78 | 92  | 0   | 88   | 70   | 80   | 32   | 94  | 88  | 0   | 32  |
| GM/NA09934A   | 0  | 50 | 0  | 0  | 83 | 79 | 4  | 87 | 0  | 0  | 77 | 87 | 0  | 2  | 89  | 0   | 90   | 89   | 0    | 91   | 89  | 2   | 0   |
| GM/NA09935A   | 0  | 0  | 52 | 10 | 28 | 12 | 0  | 0  | 0  | 0  | 0  | 22 | 74 | 72  | 0   | 0    | 93   | 59   | 0    | 9   | 91  | 71  | 0   |
| GM/NA09936    | 0  | 0  | 0  | 18 | 0  | 46 | 70 | 10 | 0  | 16 | 34 | 0  | 2  | 88 | 2   | 0    | 100  | 0    | 44   | 24  | 0   | 18  | 0   |
| GM/NA09937    | 0  | 0  | 54 | 38 | 0  | 62 | 54 | 70 | 0  | 4  | 0  | 42 | 0  | 70 | 60  | 0   | 96   | 66   | 0    | 0    | 0   | 0   | 0   |
| GM/NA09938    | 0  | 0  | 2  | 88 | 60 | 88 | 86 | 4  | 0  | 0  | 36 | 92 | 0  | 80 | 4   | 0    | 92   | 0    | 4    | 80  | 76  | 60  | 0   |
| GM/NA09940    | 0  | 0  | 46 | 0  | 0  | 0  | 84 | 62 | 0  | 0  | 0  | 0  | 0  | 0   | 62  | 0    | 100  | 0    | 0    | 0   | 0   | 0   |
| GM/NA10324    | 0  | 0  | 0  | 0  | 0  | 0  | 84 | 62 | 0  | 0  | 0  | 0  | 0  | 0    | 0    | 0    | 0    | 0    | 90   | 0   | 0   | 0   |
| GM/NA10567    | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0    | 0    | 98   | 0    | 0    | 0    | 0   | 0   | 0   |
| GM/NA10611    | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0    | 0    | 0    | 0    | 0    | 0    | 0   | 0   | 0   |

Table 4.1 Nigms human/rodent somatic cell hybrid mapping panel #1 from ATCC (Rockville, MD). The numbers in the table indicate percentage of cells with human chromosomes, which are the averages of the results obtained for the cytogenetic analysis of a minimum of 25 cells examined both at first passage and at final harvest.
Table 4.2 Mapping analysis of YAC ends. Both ends of each YAC were subcloned and sequenced. Their chromosomal locations were determined by PCR, using the derived primers to amplify hybrid DNA samples from the human/rodent somatic hybrid mapping panel #1 (ATCC, Rockville, MD).
Figure 4.1 Determination of YAC sizes by Pulse Field Gel Electrophoresis (PFGE).

A. Ethidium bromide staining of PFGE agarose gel. Yeast chromosomes from clones without YAC vectors serve as MW markers.

B. Southern hybridization of YAC DNA with total human genomic probe.
Figure 4.2 Schematic representation of Vectorette PCR. YAC DNA is denoted as bar, filled bar for YAC vector sequence and open bar for human DNA insert. "H" refers to Hind III sites within YAC DNA. "L", "R", "V" represent PCR primers of YAC left-arm, YAC right-arm and vectorette adapter, respectively.
Figure 4.3 Assembly of a YAC contig. The contig is approximately 1.8Mb in size and encompasses human MIP-1α and MIP-1β genes. The filled bars denote individual YAC clones. The numbers above them show the clone address and size of the YAC. The physical map of this contig is according to RH mapping database from the Whitehead Institute at MIT (http://www-genome.wi.mit.edu). “Cen” and “Tel” denote the centromeric and telomeric sides of the contig, respectively.
Figure 4.4 Southern analysis of the YAC contig. DNA isolated from each YAC clones were digested with EcoRI. The blots were probed with human CC chemokine gene fragments, MIP-1α, MIP-1β, MCP-1, MCP-3, RANTES, and I-309 respectively (negative results not shown).
Figure 4.5. Comparison of the contig map with the contig published by Hieshima et al (Hieshima et al., 1997). The horizontal lines denote the contig and the labels above them refer to STS markers.

A. The contig map of the YAC contig constructed
B. The contig map according to the paper by Hieshima et al (Hieshima et al., 1997).
5.1 Introduction

In order to isolate CC chemokine candidate genes from the region covered by this YAC contig, the experimental strategy was based on the following hypothesis.

1. Since the average interval between human CC chemokine genes is about 20Kb, this contig (~1.8Mb), which had been shown to contain MIP-1α and -1β genes, most likely contained some unidentified chemokine genes.

2. Chemokine genes have a conserved and simple two-intron-three-exon genomic structure. Therefore, it is reasonable to isolate individual exons by using exon as well as 3'-exon trapping systems.

3. Taking full advantage of human EST database, the complete cDNA of a chemokine gene can be identified by searching the database using individual exon sequence.

4. Chemokine genes share high homology with each other. Therefore, novel gene(s) can be identified through sequence comparison with the known chemokine genes.
Exon trapping is an effective means for isolating coding sequences from genomic DNA of relatively high complexity. This approach, based on the selection for functional splice sites in genomic DNA, has been successfully used to identify many human disease genes, including a copper transporter gene mutated in Menkes disease (Vulpe et al., 1993), the neurofibromatosis type 2 tumor suppressor gene (Trofatter et al., 1993), the Huntington’s disease candidate genes (Ambrose et al., 1992), and a potassium channel gene KVLQT1 for cardiac arrhythmias (Wang et al., 1996).

5.2 Materials and Methods

5.2.1 Partial digestion of YAC chromosomal DNA

The high molecular weight YAC DNA was isolated as described above. The partial digestion of the DNA was carried out according to Wirth et al. (Wirth et al., 1995). Agarose plugs, containing about 7.5ug of YAC DNA per plug, were washed three times in 1X TE at 50°C, then partially digested with Sau3AI (about 0.06U/plug). For the analytic digestion, one plug was melted and its final volume was brought to 250ul with 1X enzyme reaction buffer. An aliquot of 50ul sample was taken out and directly transferred to a tube containing 5ul of 0.5M EDTA (pH 8.0) at the time courses of 0 min, 3 min, 5 min, 7 min, 10 min, 15 min and 20 min. The samples were separated on a 0.3% agarose gel at 30 V over night. The digested fragments ranging from 35kb to 40kb were chosen for cosmid library construction. The preparative digestion was performed using 3 agarose plugs under the digestion time that gave the optimal products. The enzyme was immediately inactivated by adding EDTA to a final concentration of 15mM and heated at 68°C for 10 min. Dephosphorylation and
β–agarase treatment were carried at the same time. Alkaline phosphatase (5U/plug) (Boehringer Mannheim Biochemicals, Indianapolis, IN) and β–agarase (2U/plug) (USB, Cleveland, OH) were added to the digested products and incubated at 37°C for 2 hours. The enzymes were inactivated by EDTA (final concentration of 15mM) and heated at 68°C for 10 min. NaCl was added to the samples to a final concentration of 100mM. Then the mixes were extracted with buffered phenol (without chloroform) twice, followed by ether (saturated with double distilled sterile water) extraction twice. The DNA was finally precipitated out with 2.5 volume of ice-cold ethanol and left on ice for 5 min. The samples were pelleted down by spinning at 5000rpm for 20 min in a microcentrifuge. DNA was resuspended in 6ul of TE (pH 7.6). All three aliquots were pooled together and used for cosmid library construction.

5.2.2 Preparation of cosmid vector DNA

Cosmid vector SuperCos I was prepared according to the manufacturer’s protocol (Stratagene, La Jolla, CA). Vector DNA (1ug/ul) was digested with 9U/ug of Xba I enzyme in a total volume of 200ul at standard buffer conditions for 1 hour at 37°C. When the digestion was carried to the completion (a single linear band at 7.6kb on gel), DNA was extracted once with Tris-HCl (pH8.0) buffered phenol-chloroform-isoamyl alcohol (25:24:1) and once with chloroform. The aqueous phase was adjusted to 0.3M Na2OAc (pH 5.5) and DNA was precipitated with 2.5 volume of ethanol. The DNA pellet was washed with 70% ethanol once and resuspended in distilled, deionized sterile water at a concentration of 1ug/ul. The vector DNA was dephosphorylated with calf intestinal alkaline phosphatase (CIAP) (Boehringer Mannheim Biochemicals,
Indianapolis, IN). 2.5U of CIAP was used for 50pmole of 5'-termini in 1XCIAP buffer for 1 hour at 37°C. CIAP was inactivated by adding EDTA to 15mM and incubation at 68°C for 10 min. DNA was extracted with phenol-chloroform, precipitated with ethanol and resuspended in TE buffer (pH 7.6) to a concentration at 1ug/ul. The linearized vector DNA was finally digested with BamHI (5U/ug) for 1 hour at 37°C. The digestion completion was checked out on a 0.8% agarose gel. Two cosmid bands should be seen at 1.1kb and 6.5kb. DNA was phenol-chloroform extracted, precipitated and resuspended in distilled, deionized sterile water at a concentration of 1ug/ul and stored at -20°C.

5.2.3 Construction of cosmid libraries from YAC 737-d-3 and 803-a-2

About 2 to 4ug of Sau3A I-digested YAC DNA and 1 to 2ug of SuperCos I (Xha I and BamH I double digested) were ligated in a total volume of 8 to 12ul overnight at 14°C using 2U T4 ligase (New England BioLabs, MA). Half of each ligation mix was packaged with GigaPack III Gold Packaging kit (Stratagene, La Jolla, CA). Packaging extracts from -80°C was quickly thawed between figures and then placed on ice. The cosmid-YAC DNA ligation product was added to the packaging extract immediately and mixed well by gently pipetting. The mix was incubated at room temperature (22°C) for 2 hours. 500ul of SM buffer (50mM Tris-HCl, pH 7.5, 100mM NaCl, 10mM MgSO₄, and 0.01% (w/v) gelatin) and 20ul of chloroform were added to the packed stock. Debris was removed after a brief spin and the supernatant was ready to be titered or stored at 4°C.
Bacterial host Ecoli XL 1-Blue MR cells were grown up on LB (1% NaCl, 1% tryptone, 0.5% yeast extract, pH7.0) agar plate overnight at 37°C. A 50ml of NZY (2.1% (w/v) NZY broth), supplemented with 10mM MgSO₄ and 0.2% maltose, was inoculated with a single colony. The culture was grown at 30°C overnight with gentle shaking. Cells were pelleted at 2000rpm for 10 min at 4°C and resuspended in about 12.5ml of 10mM MgSO₄ (adjust to OD₆₅₀=0.5). The final packaged reactions were diluted in SM buffer to 10⁻¹–10⁻⁵. One hundred microliters of each diluted package mix was added to 100ul of bacterial host cells (in 10mM MgSO₄) and incubate for 30 min at room temperature. 150ul of LB broth was then added to each sample and incubated for 1 hour at 37°C, plated on LB plates with 50ug/ml ampicillin and grown overnight at 37°C. The number of the colonies was counted and the titer of the cosmid library was determined according to the following formula: the number of colonies X dilution factor X 1000/100 (cfu/ml).

Cosmid library was further amplified. LB Plates with at least 1X10⁶ cosmid clones were pooled together. 3ml of LB broth was added to each plate and colonies were collected into the liquid using a sterile spreader. All LB broth was pooled together and glycerol was added to a final concentration of 18% with ampicillin (50 ug/ml). An aliquot and was titered and the rest was stored at -80°C.

5.2.4 Selection of cosmid clones for putative CC chemokine gene sequences

Two degenerate oligos were synthesized based the conserved human CC chemokine amino acid sequences: oligo I 5’-TGGGTYCRGRIYHY-3’ (15bp, 192-fold degeneracy); oligo II 5’-TGYGCYDWISMM-3’ (12bp, 144-fold degeneracy).
The oligos were labeled using T4PNK and 32-P γ-ATP. Briefly, 20 pmole of oligo DNA was used in a 20ul of reaction with 1X T4 PNK buffer, 5ul of 32-P γ-ATP (Amersham, Clearbrook, IL), and 5U of T4 PNK (USB, Cleveland, OH). The mix was incubated for 45 minutes at 37°C. 180ul of fresh-made EDTA-Tris-DNA (160ul of 50mM EDTA-Tris, pH6.0 and 40ul of sonicated λ–DNA (range from 0.6kb to 5kb, 0.25ug/ul)) was added to the labeling mix followed by addition of 22ul of 1%(w/v) CPB (final concentration to 0.1%). The mix was frozen in dry-ice/ethanol bath then thawed at room temperature. Oligo was pelleted down by centrifugation at full speed for 5 minutes in a microcentrifuge at 4°C. DNA pellet was washed once with 500ul of distilled water then twice with 500ul of 80%ethanol and 20% 0.1M Na2OAc (pH5.2) (vortex the sample for 15 minutes before centrifugation). The labeled oligo was air-dried then resuspended in 100ul of distilled water. The oligos were used in the screening of cosmid clones for CC chemokine sequences.

The filters prepared from the cosmid clones containing human DNA inserts were probed with 32-P labeled generated oligo I and oligo II, respectively. The filters were pre-hybridized in 1X oligo buffer (6XSSPE, 0.5%SDS, 1% dry milk) 2 hour at room temperature, then replaced with 50ml of fresh 1X oligo buffer and the labeled oligo at 1X10^6 cpm. The hybridization was carried out over night at room temperature with gentle agitation. The filters were washed at 37°C for 15 minutes in each of following buffers: 6XSSC-0.1%SDS, 4XSSC-0.1%SDS, 2XSSC-0.1%SDS, 1XSSC-0.1%SDS, 0.5XSSC-0.1%SDS. The filters were autoradiographed 16hours at -70°C. The positive cosmid clones were collected for exon trapping.
5.2.5 Exon trapping

Exon trapping system (GIBCOL BRL, Gaithersburg, MD) provides a simplified means for isolating expressed sequences from cloned genomic DNA. Cosmid DNA from ten clones, which are positive for chemokine gene sequences, were pooled together for trapping.

5.2.5.1 Subcloning of genomic DNA into trapping vectors

The trapping was carried out according to the manufacturer’s protocol (GIBCOL BRL, Gaithersburg, MD). Trapping vector pSPL3B was digested with BamHI at its MCS and dephosphorylated with CIAP as described. Cosmid DNA was digested with BamHI, BglII and BamHI/BglII, respectively and combined together for trapping subcloning. Both cosmid DNA and vector DNA were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) once and precipitated with ethanol. The purified DNA was resuspended in TE (pH 7.6) to 0.25µg/µl. The ligation was performed in 10ul volume with 0.25µg of digested pSPL3B, 0.5µg of digested cosmid DNA, 1XT4 ligation buffer, 1U of T4 ligase (GIBCOL BRL, Gaithersburg, MD) and incubated for 16 hours at 14°C. A control reaction without cosmid DNA was done in parallel. 2ul of each ligation mix was transformed into 50ul of MAX EFFICIENCY HB101 cells following the manufacturer’s recommendation (GIBCOL BRL, Gaithersburg, MD). 50ul of each transformation was plated onto a LB plate (100µg/ml of ampicillin) to check non-recombination rate (<10% is acceptable). The remaining cells were grown in 5ml of LB with 100µg/ml of ampicillin over night at 37°C. DNA was prepared using
the QIAprep Spin Mini kits followed the manufacturer's recommendation (QIAGEN, Chatsworth, CA) for transfection.

5.2.5.2 Transfection

COS-7 cells (ATCC, Rockville, MD) were grown in 5ml of supplemented D-MEM medium (1X D-MEM, 10% fetal bovine serum, 2mM L-Glutamine, 1X Penicillin-Streptomycin, 0.1mM MEM non-essential amino acids solution) in 6-well tissue culture dish at 37°C and in a 5% CO₂ incubator until 60% to 80% confluent. Then the medium was replaced with 2ml of serum free D-MEM medium. The COS-7 cells were ready for transfection.

6ul of LipofectACE reagent was added to 100ul of OPTI-MEM I without serum (GIBCOL BRL, Gaithersburg, MD), mixed well and incubated at room temperature for 5 min. 1ug of trapping DNA prepared as before was diluted in 100ul of OPTI-MEM I medium without serum and combined with the lipid mixture prepared just before. The mixture was incubated at room temperature for 10 min followed by addition of 0.8ml of serum free D-MEM medium. After the D-MEM medium was removed, the COS-7 cells were layered with the 1ml of DNA/lipid/OPTI-MEM I mixture and incubated for 5 to 6 hours at 37°C in a 5% CO₂ incubator. The DNA/lipid/OPTI-MEM I was replaced with 2ml of supplemented D-MEM medium and the COS-7 cells were incubated for an additional 24 hours at 37°C in a 5% CO₂ incubator before RNA was isolated.
5.2.5.3 Isolation of total RNA

Total RNA from COS-7 cells was isolated using TRIzol reagent (GIBCOL BRL, Gaithersburg, MD) according to the manufacturer’s protocol. Medium was removed before 1ml of TRIzol reagent was added to each COS-7 cell sample. The cells were lysed by passing through a pipette several times and collected into a Eppendorf tube. The cell lysates were held for 5 min at room temperature before 0.2ml of chloroform was added. The mixture was vortexed for 15 sec, held for 2 min at room temperature, and centrifuged at 12,000g for 15 min at 4°C. The upper aqueous phase was transferred to a fresh Eppendorf tube. Total RNA was precipitated by adding isopropanol (0.5ml per ml of TRIzol reagent used) and holding for 5 to10 min at room temperature, followed by centrifugation at 12,000g for 10 min at 4°C. RNA was washed once with 75% ethanol, briefly dried at room temperature and resuspended in 50ul of DEPC-treated water by 5 min incubation at 65°C followed by vortexing.

5.2.5.4 cDNA synthesis

2ug of total RNA was annealed to 1ul of primer SA2 (20uM, ) by heating 5 min at 70°C followed by chilling 1 min on ice. The first strand cDNA was synthesized in 20ul of final volume with 1X first strand buffer, 10mM DTT and 500uM dNTPs. The mixture was incubated for 2 min at 42°C. After addition of 1ul of superscript II RT (GIBCOL BRL, Gaithersburg, MD) (200U/ul), the mixture was incubated for additional 30 min at 42°C before the enzyme was inactivated at 55°C for 5 min. Then, 1ul of Rnase H was added and the mixture was continually incubated for 10 min at 55°C. The cDNA products were ready for PCR amplification.
5.2.5.5 Primary PCR amplification

The first PCR amplification was carried out in a total volume of 40ul with 8ul of each cDNA product, 1X PCR buffer, 200uM dNTPs, 1uM of each primer SA2 and SD6, and 2.5U Taq (GIBCOL BRL, Gaithersburg, MD). Total of 6 cycles of PCR amplification were followed: denaturization at 94°C for 1 min, annealing at 60°C for 1 min, extension at 72°C for 5 min with a final extension for 10 min at 72°C.

5.2.5.6 Digestion with BstX I to eliminate vector-derived products

The primary PCR products were treated with BstX I to eliminated vector-derived products. 10ul of BstX I (10U/ul) was added to each primary products. The mixtures were incubated at 55°C over night. The next morning, additional BstX I was added (4U) for another 2 hours digestion at 55°C.

5.2.5.7 Secondary PCR amplification

The secondary PCR amplification was carried out in a total volume of 50ul with 5ul of primary PCR product, 1X PCR buffer, 200uM dNTPs, 400pM of each primer dUSA4 and dUSD2. The PCR mixtures were hot-started before 2.5U Taq was added. Amplification was done for 30 cycles of denaturization at 94°C for 1 min, annealing at 60°C for 1 min, extension at 72°C for 3 min with a final extension for 10 min at 72°C. The PCR products were checked in a 2% agarose gel for the presence of fragments >177bp but <500bp (177bp-band is the vector-derived fragment).
5.2.5.8 Cloning the secondary PCR products

Because the secondary PCR primers have dUTPs incorporated into their 5' ends, the secondary PCR products could be easily cloned by uracil DNA glycosylase (UDG) (GIBCO BRL, Gaithersburg, MD). 100ng of secondary PCR products (typically 1 to 2ul) was added to a final volume of 10ul of subcloning mix with 1X PCR buffer, 50ng of pAMP10 vector DNA, and 1U of UDG (GIBCO BRL, Gaithersburg, MD). The mix was incubated at 37°C for 30 min. 5ul of each products was transformed into 50ul of competent DH5α cells (GIBCO BRL, Gaithersburg, MD) and plated out on LB plates with 100ug/ml ampicillin.

5.2.5.9 PCR evaluation of transformants

Colonies from each transformation were screened with PCR using secondary amplification primers. Colony PCR was performed using a standard protocol and in 30 cycles of denaturization at 94°C for 45 sec, annealing at 55°C for 30 sec, extension at 72°C for 1 min with a final extension for 10 min at 72°C. The PCR products were checked on a 2% agarose gel for the presence of fragments ranging from 177bp to 500bp.

5.2.5.10 Sequencing trapped fragments

Plasmid DNA was purified using QIAGEN Spin miniprep kit (QIAGEN, Chatsworth, CA) and sequenced using plasmid vector specific primer SD2 and Sequenase 2.0 sequencing kit (Amersham, Cleveland, OH) according to the manufacturer's protocol.
5.2.6 3'-exon trapping

3'-exon trapping follows a similar scheme as in exon trapping. The trapping vector for 3'-exon, pATG4, selects 3'-exon because the vector itself lacks a 3'-polyadenylation (poly(A)) signal for the transcription unit. Cosmid DNA was digested with BamHI, Bgl II and BamHI/Bgl II and subcloned into pATG4 at BamHI site. The total RNA was isolated from transfected COS-7 cells. The primary PCR products were digested with EcoRI to eliminate cryptic splicing before being re-amplified with nested primers. PCR products, ranging from 100bp - 400bp were subcloned for sequencing (vector derived sequence is about 100bp).

5.2.7 Homology analysis

Sequence homology of the trapped fragments with human chemokines at both nucleotide and amino acid levels were analyzed through NCBI databases using BLASTN and BLASTX programs (http://www.ncbi.nlm.nih.gov). Sequences with significant homology with the known human CC chemokines were used to search human EST databases at TIGR for full length cDNA clones.

5.2.8 Southern analysis of cosmid clones used in exon trapping:

Cosmid clones used in exon trapping were analyzed for the presence of the chemokine candidate genes. Cosmid DNA was digested with EcoRI and probed with 32-P labeled THC-14098_1 fragment, following the protocol described. The X-ray film was exposed at -70°C for 3 days before being developed.
5.2.9 Analysis of expression patterns of the trapped exons

From TIGR databases, two EST contigs, THC-14098_1 and THC-79282, were identified as showing high homology to human MIP-1 genes. Two primer sets were designed to amplify the putative coding regions of these two clones. Primers 5'-TGCCTGCCAGCATCATG-3' and 5'-CCTCAGGCATTCAGC TTCA-3' amplified THC-140198_1 coding region. Primers 5'-AGCTTCCCACA GCATGAAGA-3' and 5'-TGAGAGTTAGCGGTGGGTG-3' amplified THC-79828 coding region. Total of 15 cDNA samples were amplified with these two primer sets under the cycling conditions: 35 cycles of denaturization at 94°C for 45 sec, annealing at 52°C for 30 sec, extension at 72°C for 30 sec with a final extension for 10 min at 72°C. The cDNA samples were from monocyte (LPS-stimulated), monocyte (3-hour culture), monocyte (72-hour culture), MonoMacs, T cells, activated T cells, T cell blasts, HL-60, HL-60 (PMA/PHA-stimulated), HL-60 (cAMP-stimulated), human neutrophils, THP-1, eosinophils, U937, U937 (DMSO stimulated) (They were kindly provided by Dr. Kuldeep Neote from Pfizer Central Research, Gorton, CT). RT-PCR products were subcloned directly into pTA2.1 vectors and sequenced. They were used as probes in Northern hybridization.

Human Multiple Tissue Northern (MTN) blots H3 (immune system) and H (general tissues) were purchased from Clontech (CLONTECH Laboratories, Inc., Palo Alto, CA). cDNA fragments isolated as before were 32-P labeled and hybridized to the MTN blots according to manufacturer's protocol (CLONTECH Laboratories, Inc., Palo Alto, CA). MTN blots were prehybridized in 20ml of Solution 2 (5X SSPE, 10X Denhardt's solution, 100µg/ml sheared salmon sperm DNA, 2.0%SDS, 50%
formamide) for 6 hours at 42°C while rotating in a Hybaid oven (Labnet International, Westbourne, UK). 32-P labeled cDNA probes were added to 10ml of fresh Solution 2 at a concentration of $1 - 2 \times 10^6$ cpm/ml and mixed thoroughly. The prehybridization buffer was replaced with the fresh hybridization/probe solution and the blots were hybridized at 42°C for 18 – 24 hours. The blots were rinsed with wash solution 1 (2XSSC, 0.05%SDS) at room temperature then washed several times for 30 to 40 min in wash solution 1 with shaking. After being washed two times in solution 2 (0.1XSSC, 0.1%SDS) at 50°C with shaking, the blots were exposed to X-ray film at -70°C for two days.

5.2.10 Determination of the genomic structure of the chemokine candidate gene

The genomic structure of the candidate gene was defined by subcloning of the cosmid clones that had been used for the exon trapping. The cosmid DNAs were digested with EcoR I then cloned into the pBluescript-SK vector (Stratagene, La Jolla, CA). Positive clones were identified by hybridization with the THC cDNA fragments. The intron-exon boundaries of the candidate gene were defined in sequencing reactions of the these genomic clones using the forward and reverse primers of the exons: 5'-TGCCTGCCCAGCATCATG-3' (Exon-1.For), 5'-TTGGTACCAACAAAGAGCT-3' and 5'-GGCTTGGGGGACTGGGGG-3' (Exon-2.For and Rev), 5'-CCTCAGGGCATCAGCTTCAGG-3' (Exon-3.Rev). The 5' promoter region was sequenced using a series of primers generated from the 5'-sequence. The whole gene fragment and each intron were amplified and sized using TAKARA Ex-Taq (Oncor, Gaithersburg, MD) in long range PCR with primers Exon-1.For and Exon-3.Rev (whole gene), Exon-1.For
and Exon-2.Rev (intron 1), and Exon-2.For and Exon-3.Rev (intron 2) under the conditions of 2 min of initial denaturization at 94 °C, followed by 35 cycles of amplification at 94°C for 20 sec, 60°C for 30 sec, 68°C for 4 min, ended with final extension at 68°C for 7 min.

5.2.11 Chromosomal location of the chemokine candidate gene

The cosmid clone which contains the candidate gene was found to contain MIP-1α also. This indicates that these two gene are located within 30kb distance. The gene location of MIP-4 relative to MIP-1α on 17q11.2-q12 was determined using TAKARA LA-Tag (Oncor, Gaithersburg, MD) in extra-long range PCR amplifications. Primers from one gene were paired with primers from the other gene to amplify the junction fragment between them, MIP-1α 5'-Reverse 5'-ACGGAATGTGGGCTCGAGTG-3' and 3'-Forward 5'-AAACATGCGTGTGACCTCCACA-3', MIP-4 5'-Reverse 5'-GAAGGGTTCAGAGTGAGCTGGG-3' and 3'-Forward 5'-AATGGCCCTGACCCGCTGAC CCTG-3'. Cosmid DNA was treated with Not I to release the insert before amplification. Total human genomic DNA was used as template for a positive control reaction. The PCR amplification was performed under the following conditions: 2 min of initial denaturization at 94 °C, followed by 35 cycles of amplification at 94°C for 20 sec, 60°C for 30 sec, 68°C for 7 min, ended with final extension at 68°C for 7 min.

5.3 Results

To isolate the human inserts from the YACs, one approach is to construct a cosm id library from the clone and then screen the library with human genomic DNA
probes. Cosmid vector SuperCosI (Stratagene, La Jolla, CA) (Figure 5.1) is commonly used for subcloning YAC DNA. YAC DNA, partially digested with Sau3AI for 5 min to 10 min, was subcloned into SuperCosI at the BamHI site. The two cosmid libraries, generated from the non-chimeric YAC clones 737.d.3 and 803.a.2, contain about 3,000 clones, twice as many as the clones required for 99% representation of the YAC DNA in the cosmid library. Since the DNA source used for the cosmid library construction was total YAC DNA that includes yeast DNA, the cosmid libraries were screened with human genomic probes to isolate the clones that have human DNA inserts. A number of clones were isolated from each library: 162 clones from 737.d.3 (870kb) and 132 from 803.a.2 (1520kb). Theoretically, full representation of both YAC chromosomes was met.

It has long been recognized that human chemokines share high homology with each other. By sequence alignment analysis at the amino acid level, two highly conserved regions were determined among human CC chemokines. The degenerate oligonucleotides derived from these regions were then used to screen cosmid clones for chemokine sequences. About 20 clones from each cosmid library were isolated. These chemokine sequence-enriched clones were pooled together (10 clones in one pool) to increase the efficiency in trapping experiments.

Exon trapping vector pSPL3B (Figure 5.2), which has a transcription unit with a 5’exon and a 3’-exon and driven by the SV40 promoter, was employed in isolating exon sequences. The vector’s multiple cloning site (MCS) is flanked by a splicing donor site (SD) and a splicing acceptor site (SA). When the cloned genomic fragment contains an intact exon in the correct orientation, splicing can occur between the vector and the
insert sequence and a hybrid mRNA is generated in the transfected COS-7 cells. First round RT-PCR was followed by digestion with BstXI to remove PCR products, which resulted from the vector self-splicing. Using nested PCR primers, the secondary RT-PCR products were generated to ensure the specific amplification of trapped sequences. All of the RT-PCR primers were vector sequences derived. The PCR product is 177bp if there is no trapped sequence between vector exons (Figure 5.3).

3'-exon trapping system works in a similar way. Vector pATG4 (Figure 5.4) contains a truncated transcription unit, which lacks a 3' polyA signal. Subcloned genomic DNA containing a 3'exon which has a polyA signal will compensate the truncated vector transcription unit and produce a mature and stable mRNA in COS-7 cells. The first-strand cDNA was initialized with the adapter primer (AP2) which relies on the presence of a poly(A) tail. To eliminate cryptic splicing products, EcoRI restriction digestion was followed after primary RT-PCR. Nested PCR primers were used in secondary amplification.

From the trapped clones that have been sequenced, two clones showed very high homology to known human CC chemokines, especial the MIP-1 genes. These two putative chemokine exons identified two EST contig clones from the TIGR human cDNA database (HCD), THC-140198-1 and THC-79282 (Figure 5.5).

According to the TIGR HCD database, these EST contigs were assembled from sequences generated from human aorta, uterus, and fetal liver. To determine which cells express these sequences, a series of RT-PCR reactions were performed using a variety of different cell types from the immune system. Out of the 15 cell samples checked, monocyte (LPS-stimulated), monocyte (3-hour culture), monocyte (72-hour culture),
MonoMacs, T cells, activated T cells, T cell blasts, HL-60, HL-60 (PMA/PHA-stimulated), HL-60 (cAMP-stimulated), human neutrophils, THP-1, eosinophils, U937, U937 (DMSO stimulated), THC-140198_1 was expressed in monocyte (72-hour culture), T cell blasts and U937 while THC-79282 was only expressed in LPS-stimulated monocyte (Figure 5.6).

Northern hybridization results showed that THC-140198_1 transcript is about 800bp long and expressed abundantly in lung (Figure 5.7). A smaller amount was detected in bone marrow, lymph node and placenta. A very low amount was found in heart but no expression was observed in muscle, liver, brain or kidney. The expression pattern of THC-79282 is quite different from that of THC-140198_1. It is expressed as a 650bp message and was detected in liver, heart and occipital pole of human brain (Figure 5.8). At this stage, THC-79282 was identified as HCC-1 (Knapp et al., 1996). Therefore, it was excluded from further studies.

Sequence analysis indicates that THC-140198_1 has a putative open reading frame (ORF) of 297bp, which encodes a protein with a 20-amino acid signal peptide. Its predicted amino acid sequence shares very high homology to the known human CC chemokines, especially to MIP-1α with 59.6% identity (Figure 5.9). It has the typical 4-cysteine motif and a conserved C-terminal domain of CC chemokines. Sequence homology analysis at the nucleotide level indicates a range of 30% (Eotaxin) to 49% (MIP-1α) of similarity between THC-140198_1 and other known human CC chemokine genes. Since THC-140198_1 shows the closest sequence homology to huMIP-1α (with 49.5% identity at the nucleotide level and 59.6% at the predicted
amino acid level). In view of this similarity to MIP-1α, and of the existence of MIP-2α and β, and MIP-3α and β, we have called this gene MIP-4.

The exon trapping system is supposed to isolate an intact internal exon from a gene. The CC chemokine gene family has a conserved three-exon-two-intron genomic structure. The trapped exon, isolated by this exon trapping study, is the exon 2 of a CC chemokine gene and therefore, the exon boundaries within the full length cDNA fragment could be deduced. The primers, designed from the exon boundary sequences, generated intron sequences when they were used in sequencing the plasmid clones containing the MIP-4 gene fragment. All the introns have the conserved splicing consensus sequences (Table 5.1). Within 500bp of the 5' upstream region, a TATA box, and a TFII site were identified (Figure 5.10).

The result of the southern hybridization of the genomic DNA from the cosmid clones has shown that the MIP-4 cDNA hybridizes to a 3.4kb EcoRI fragment. One cosmid clone showed a smaller band at 1.0kb (Figure 5.11). This could be due to that the gene exists in the end of the cosmid insert. The clone that contains MIP-1α gene is also positive for the MIP-4 cDNA probe. This suggests that these two genes exist within a 40kb distance.

To further define the gene structure, the whole gene and its two introns were amplified in long-range PCR reactions (Figure 5.12). The complete MIP-4 gene is about 7.1 kb long with intron 1 of 5.8 kb and intron 2 of 0.5 kb. One cosmid clone was found to be positive for both the MIP-1α and MIP-4 genes, suggesting that these two genes are located close to each other. Long-range PCR was performed to determine the distance between the two genes. A single fragment of 16 kb was amplified using the
primers MIP-1α 3'-For and MIP-4 3'. For from both cosmid and human genomic DNA templates. This result suggests that the two genes are approximately 16kb apart and reside in a "tail to tail" fashion (Figure 5.13).

5.4 Discussion

From the cluster region on 17q11.2-q12, two novel human CC chemokines were isolated. One was reported shortly after as Human CC chemokine 1 (HCC-1) (Knapp et al., 1996). We named the other as MIP-4 based on its high sequence homology with human MIP-1α gene. The overall sequence homology at the nucleotide level is 49.6%.

The genomic structure of the MIP-4 gene was determined by long-range PCR amplification, and the gene was found to consist of three exons. This is very similar to the MIP1α, MCP1, I-309 and MIP 1β genes which all have three exons and two introns. The CXC chemokines are also similar to each other in gene structure. As gene mapping studies indicated that the same cosmid contained MIP-4 and MIP 1α, we used long-range PCR to determine the distance between these two genes. They are located 16kb apart and they are in opposite orientations. As reported previously, the MIP-1β gene is located 14 kb away from the MIP-1α gene in a "head to head" direction (Irving et al., 1990). The chemokine HCC-1 is not in the same cosmid as MIP1α and MIP-4 are, but does lie in the same 1.8Mb YAC contig that has been constructed. Based on the radiation hybrid map generated at the Whitehead Genome center, some other CC chemokines were also located within the 17q11.2 region. Given this remarkable clustering of the genes, it can be speculated that these chemokines arose by duplication. The conservation of exon/intron structure between the CC chemokine genes including
those in the two separate clusters would also be consistent with this origin of the chemokines genes. Therefore, from an evolutionary point of view, it will be interesting to determine whether the same number of CC chemokine genes exists in other species.
<table>
<thead>
<tr>
<th>Exon</th>
<th>AA number</th>
<th>Nucleotide num.</th>
<th>5'-intron</th>
<th>Exon sequence</th>
<th>3'-intron</th>
<th>exon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-24</td>
<td>1</td>
<td>CCAGCTCACTCTGA...CTCCTGTGCACAAG</td>
<td>gtagtgctgtcag</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>24-59</td>
<td>105</td>
<td>cctacctcttgag</td>
<td>TTGGTACCAACAA...AAGCCAGGTGTCAT</td>
<td>gtaagtgcaagtg</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>59-89</td>
<td>217</td>
<td>tccctctccacag</td>
<td>CATCCTCCTAAACCA...TAAAGGTAACCACG</td>
<td></td>
<td>518</td>
</tr>
</tbody>
</table>

Table 5.1 Exon-intron boundaries of the human MIP-4 gene. The exon sequences are in upper case, showing the flanking sequences of the exon. Intron sequences are shown in lower case with flanking sequences. The numbers in “Exon” denote the order of the exons. “AA number” refers to the starting and ending amino acids of the peptide encoded by the exon. If the codon is divided by the exon boundary, the amino acid is counted in both exons. “Nucleotide num.” indicates the starting nucleotide of the exon.
Figure 5.1 Map of SuperCos I cosmid vector.
The vector can accommodate genomic DNA fragments ranging in size from 30kb to 42kb. The NotI sites, flanking the vector’s poly linker (CS), allow the excision of the genomic inserts as a single large fragment in most cases.
Figure 5.2 Vector map of pSPL3B (GIBCO BRL, Gaithersburg, MD). The vector contains a SV40 promoter, a splice donor and a splice acceptor which flank the MCS.
Figure 5.3 Schematic diagram of exon trapping procedure. Exons are denoted by bars (open bars for vector exons and the filled bar for subcloned exon) and introns by horizontal lines. Two potential splicing events ($\alpha$ and $\beta$) are shown. RNA transcripts from transformed COS-7 cells are isolated and nested amplified by RT-PCR. The PCR products which contain trapped exon are larger than 177bp.
Figure 5.4 Map of the 3'-exon trapping vector pTAG4. The vector contains a SV40 promoter. The MCS follows a vector splice donor site. A poly(A) signal within the subcloned fragment is essential to generate a stable transcript.
THC 140198_1 (MIP-la isolog): 734bp

GCTGTTACCTYTTATATATAAGTACTGCCAATATATTAAAAAGGTGTGCTGA
GCAAACCATTTCAATAAAAGATTGTTGATACATATGGCACAATGTCTGCT
GAGAAAGCTCAATCACACTGATTGAGCCTATGCAATGATGAAGAGTTGAAG
GGAAAGGGGAAAGGATGATAAAGACAATGTCATGAGATGTCAGAGAAATACTA
ATTTCATAATGTCAGCTCACAATAGAATAAATCAATTTCAAAATGAAATATGATGCA
TAAACTAAGATATAGTTAATGTGCTTGCCATAGCATGTGGACTCTTTA
GAAGACGTGCTCCAGGTCGATGATTTCTGGGACCACCTCTTTTATGAGGTCAGC
ACAGATCTCCGCCCTCTCTTTGGTATGGAGATGACACCTGGCTTGGGGCACTG
GGGGCTGGTTTCAGAATAGTCAACTATGAACTTTTGAGTAGACACCTGGCCTGCTGCT
THC 79282 (LD78β isolog): 485bp

CCTCACCGCAGaGcTCTGAAGCTCCACCAAGGCAGCTCTCCCTCCACACACAG
CTCCCCACAGCATGAGATCTCCGTCGCTTACCTTTCTCCCTGCTCTCCTTCTCAA
CCATGGCCCTAGGGACAGACGACTGGAATCTCTCCTCAGGAGCCACTTACACCCCT
CAGAGTGCTGCTTCTTACACTACTAAGAAATCCTCAGGCTCATAGATTGAGATT
ACTATGAGACCAACAGCATGCTCCACAGCCCGAGATTGTCTTCCATACACAAACAG
GGGGCCATTCCCTCCACGATCAACTGAAGTGAGGGTGCGAAGGCAACAGTCAGA
GACATAAAGGAAAGTGCCAGCGCCCTCTCCCTCCACAGCTCATCAGGCCC
CAGTCACCCTCTTTGAGCCTCCCTCTGCTTTGAAATAGACACACTTACATGCTCCTT

Figure 5.5 Nucleotide sequences of THC-140198-1 and THC-79282 from the human cDNA database (HCD) at TIGR (http://www.tigr.org).
Figure 5.6 RT-PCR amplification of cDNA from different cell lines. PCR primers were derived from THC sequences. cDNA samples include:
1. Monocytes (LPS-stimulated)  
2. Monocytes (3-hour culture)  
3. Monocytes (72-hour culture)  
4. MonoMacs  
5. T cells  
6. Activated T cells  
7. T blastcells  
8. HL-60  
9. HL-60 (cAMP stimulated)  
10. HL-60 (PMA/PHA-stimulated)  
11. Neutrophils  
12. THP-1  
13. Eosinophils  
14. U937  
15. U937 (DMSO stimulated)  
16. Blank control
Figure 5.7  Expression profile of THC-140198-1 in human tissues. Human muscles included placenta, liver, skeletal muscle, kidney, pancreas, stomach, thyroid, spinal cord, lymph node, trachea, adrenal gland, bone marrow, and brain. The expression levels of THC-140198-1 were quantified using Western blotting.
Figure 5.8 Expression profile of THC-79282 in human tissues. Human multi-tissue Northern blots (MTN) were purchased from Clonetech and probed with $^{32}$P labeled THC-79282 fragment.
Figure 5.9  Alignment of human CC chemokine sequences at amino acid level. MIP-4 sequence was aligned against 9 known chemokines by Clustal Align method from DNASTAR package. The four cysteines are shown in bold face. The consensus sequence is shown in shadow.
Figure 5.10 Genomic sequence of the human MIP-4 gene. The promoter sequence, 3' & 5' untranslated region and intron sequences are shown in lower case. The exon sequences are shown in upper case while the initiation and termination codons are in bold face. A potential TATA box is double underlined. The poly(A) signal is underlined in bold face.
Figure 5.11  Southern analysis of cosmid DNA.
Cosmid DNA was digested with EcoRI and probed with
$^{32}$P labeled MIP-4 cDNA fragment.
Samples 1 to 6 are cosmids subcloned from YAC 737.d.3.
Clone 3 also contains MIP-1$\alpha$ gene.
Sample 7 is a cosmid subcloned from YAC 803.a.2.
Figure 5.12 Genomic structure of the human MIP-4 gene. Exons are diagrammed as bars, the solid portions of the bars indicating the translated regions while the hashed portion indicating 5' & 3' UTR. The lines between exons indicate introns and their sizes. Below the diagram, the amplification products obtained using the following primer combination are shown.

A. Ex-1-For and Ex-3-Rev
B. Ex-2-For and Ex-3-Rev
C. Ex-1-For and Ex-2-Rev
Lane 1: PCR negative control
Lane 2: human genomic DNA template
Lane 3: cosmid DNA template
Figure 5.13  The chromosomal location of the MIP-4 gene related to the MIP-1α gene. The fragment between MIP-4 and MIP-1α was amplified in a long range PCR reaction using the primers from 3'-region of MIP-1α and 3'-region of MIP-4. The amplified fragment is shown below and is approximately 16kb.
Lane 1: PCR negative control
Lane 2: genomic DNA template
Lane 3: cosmid DNA template
Chapter 6

Generation of Recombinant MIP-4 Protein and Studies of Its Chemotaxic Activities

6.1 Introduction

The chemokines are a superfamily of small proteins (8,000 kDa to 14,000 kDa) secreted primarily by leukocytes and associated with each other by a conserved 4-cysteine motif. These proteins have been implicated in a wide range of acute and chronic inflammatory processes as well as other immunoregulatory functions.

The distinguishing feature of the chemokines is that they act as chemoattractants for various cell types, especially those from the immune system. CC chemokines are mainly monocyte-activators. They also act on basophils, eosinophils, lymphocytes, as well as some non-leukocytes such as fibroblasts, macrophage stem cells but not neutrophils (Baggiololini et al., 1993). There are a few in vitro assays which are commonly used in determining a protein's chemotactic activity. These include chemotaxis chamber assay which measures the migration of cells attracted by a particular chemokine, Ca++ influx assay which monitors calcium mobilization that usually leads the chemotaxis movement of the cells. Actin polymerization which
reflects cellular shape changes, and histamine release which is prior to cell activation are also the means to monitor chemokine activities.

Three CC chemokines MIP-1α, -1β and RANTES have been recently identified as the major HIV suppressive factors. The discovery of distinct chemokine receptors that support entry of T-cell tropic (CXCR-4) and macrophage tropic HIV-1 strains (CCR-5 and CCR-2b) explains the differences in cell tropism between viral strains, the inability of HIV-1 to infect most non-primate cells, and the resistance of a small percentage of the population to HIV-1 infection.

Studying chemokine functions in vivo has led to the assessment of the significant roles of chemokines in immune responses. Intradermal injection of CC chemokine proteins MCP-1, MCP-2, MCP-3 into small animals such as the mouse and rabbit, induced almost exclusively monocyte infiltration (Van Damme et al., 1992). Injection of MIP-1α into mouse peritoneal cavity recruited both peritoneal young macrophages and exudate macrophages from the circulation to the inflammation sites (Chen et al., 1993).

6.2 Materials and Methods

6.2.1 Cloning into expression vector pRSET C

Bacterial expression vector pRSET-C (Invitrogen, Carlsbad, CA) was double-digested with BamHI and EcoRI. The cDNA fragment from codon 22 (Ala) to codon 89 (Ala) was PCR-amplified. A BamHI site was incorporated into the forward primer 5'-TGGATCCGATGTGCACAAGT-3' and an EcoRI site into the reverse primer 5'-CG AATTCCTCAGGCATTCAG-3'. The PCR product was purified, digested
completely with EcoRI and BamHI and cloned into pRSET_C at BamHI/EcoRI sites. The reading frame and the sequence of the construct were verified by sequencing before being transformed into the host cell BL_21(DE3).

6.2.2 Generation of 6XHis-MIP-4 fusion protein

A pilot expression experiment was done to determine induction kinetics of the fusion protein and the optimal induction condition. 1ml cultures were grown up and samples were collected at different time points of IPTG induction. The cell samples were lysed and checked by SDS-PAGE (12% gel) for the induction of the MIP-4 fusion proteins, compared with non-induced samples. Different temperatures (37°C and 30°C) and different IPTG concentrations (0.1mM to 1mM) were tested. For the preparative induction, 50ml of LB-ampicillin culture from a single colony was grown overnight at 37°C. The next day, a one liter culture was inoculated with 25ml overnight culture and continuously grew at 30°C until OD₆₀₀=0.3. IPTG was added to the culture to a final concentration of 0.3mM and the induction was lasted for 1.5 hours at 30°C. Cells were harvested by centrifugation 4500xg for 15 min at 4°C.

6.2.3 Generation of MIP-4 fusion protein from expression vector pET-32c

To increase the solubility of the MIP-4 fusion protein, another expression system pET-32c (Novagen, Madison, WI) was used. The MIP-4 cDNA insert from pRSET_C construct was released and subcloned into pET-32c at BamHI/EcoRI sites in frame. The construct was transformed into BL_21 (DE3) cells and pilot experiments were performed to determine the optimal induction and reasonable solubility of the
fusion protein. A one liter LB-Ampicillin culture was grown up at 30°C until OD600=0.6 and induced with IPTG (final concentration at 0.4mM) for 2.5 hours at 30°C. Cells were harvested by centrifugation 4500xg for 15 min at 4°C.

6.2.4 Purification of MIP-4 fusion protein using TALON resin

The MIP-4 fusion protein was purified using TALON metal affinity column essentially as described (Clontech Laboratories, Inc., Palo Alto, CA). Native and denaturing fusion protein purification protocols were used for pRSET_C or pET-32c systems, respectively.

6.2.4.1 Preparation of denaturing 6XHis protein for purification

Cell pellets from one liter culture were resuspended in 50ml of lysis buffer (50mM NaH₂PO₄, 10mM Tris-HCl (pH8.0), 100mM NaCl, 6M Guanidinium-HCl), agitating until completely lysed. Cell debris was removed after centrifugation at 12,000xg or 15 min at 4°C.

6.2.4.2 Preparation of native 6XHis protein for purification

Cell pellets from one liter culture were resuspended in 50 ml of sonication buffer (50mM NaH₂PO₄, 10mM Tris-HCl(pH8.0), 100mM NaCl) and sonicated thoroughly using of a series of short, repetitive bursts in a ice-water bath. Cell debris was removed after centrifugation at 12,000xg or 15 min at 4°C.
6.2.4.3 Purification of 6XHis fusion protein by TALON resin

TALON resin (5ml for 1 liter culture prep) was washed with 5x volume of lysis buffer (for denaturing protein) or sonication buffer (for native protein). The protein samples were added to the resin and incubated for 20 minutes at room temperature with gentle agitation. Supernatant was removed after centrifugation for 5 min at 2,000xg and the resin-protein complex was washed three times in 10x volume of lysis or sonication buffer by incubation at room temperature for 10 min with gentle agitation. The protein was eluted with 1x resin bed volume of elution buffer (20mM Tris-HCl pH8.0, 100mM NaCl, 50mM imidazole, plus 8M Urea for denaturing protein) by incubation at room temperature for 10 min with gentle agitation. The elution steps were repeated three times and all supernatants were collected. The protein was concentrated to a final volume of 2ml using Centricon-30 concentrators essentially as described (Amicon, Inc., Beverly, MA).

6.2.5 Cleavage of 6Xhis tag using Enterokinase

Fusion proteins generated from both pRSET-C and pET-32c systems contain an enterokinase recognition site immediately prior to the recombinant protein sequence. The fusion tags were removed by EK digestion, according to the manufacturer's recommendation with a slight modification (Invitrogen, Carlsbad, CA). Pilot experiments were performed to determine the optimal enzyme concentration and digestion conditions. Fusion tags and recombinant MIP-4 proteins were checked through 18% SDS-PAGE. The recombinant MIP-4 protein is about 7kDa, Xpress fusion tag is 4kDa (from pRSET_C) and Thioredoxin fusion tag is 14.6kDa. For
preparative treatment of the purified 6Xhis-MIP-4 fusion proteins, the protein samples were dialyzed against 1X EKMAX buffer (50mM Tris-HCl pH8.0, 1mM CaCl₂, 0.1% Tween-20) and incubated with 2U EKMAX per 10ug of protein for 2 hours at room temperature. The free tags and EKMAX itself were removed by passing through the TALON column as described above.

6.2.6 Measurements of intracellular calcium using the spectrofluorometer and flow cytometry

Cells were harvested, resuspended at 1X10⁶ per ml in RPMI 1640 and incubated with 2 μM Fluo-3 AM (Molecular Probes, Eugene, Oregon) for 20 min at 37°C. Following incubation with Fluo-3 AM, the cells were washed in RPMI, allowed to rest for 5 min, resuspended in 100 μl RPMI and stained for 10 min at room temperature with 20 μl anti-phycoerythrin conjugated CD4, CD8, CD45RA or CD45RO (Coulter /Immunotech, Hialeah, FL) per 2x10⁶ cells. Cells were then washed in RPMI and resuspended in 500 μl of RPMI. Flow cytometry was performed on a FACScan (Becton Dickinson, San Jose, CA) as follows: the sample run was initiated and the samples were then removed, challenged with 100 nM chemokines and the run was continued as quickly as possible for 60 to 90 seconds. The parameters collected were FSC, SSC, FL1 and FL2 versus time. The data was analyzed over the entire 90 seconds by gating on either CD4- or CD8-positive cells. Dot plots were generated using time vs. FL1 and the increase in mean channel fluorescence of FL1 over background was calculated.
6.2.7 Analysis of peritoneal cell influx

Six to eight week-old male BALB/c mice were obtained from the Central Animal House at the University of Adelaide, South Australia. Synthetic MIP-4 (10 μg/ml) or diluent (endotoxin-free PBS) was injected into the peritoneal cavity. The cellular exudate in the peritoneal cavity was collected 24 hours later and flow cytometry was performed as previously described (Tessier et al., 1997). The exudate cells were centrifuged at 100xg for 10 minutes, the supernatant was discarded and the cells were resuspended in RPMI supplemented with 10% FCS. The cells were stained using supernatants from the hybridoma P3X63Ag8 as a negative control, or using the CD4+ T cell marker GK1.5, or the CD8+ T cell marker 53-6.7.

6.2.8 Isolation of natural killer (NK) cells from peripheral blood

6.2.8.1 Isolation of peripheral blood mononuclear cells (PMBC) through Ficoll-Hyphaque density gradient centrifugation

PMBC were isolated from normal donor fresh leukopacs (American Red Cross, Columbus, OH) as described (Fehniger et al., 1998). Leukopac from a single donor (~60ml) was diluted (1:1) with RPMI-1640, then layered with half volume of Ficoll (Histopaque-1077) (Sigma, St. Louise, MO). After centrifugation for 30 min at 2000 rpm, PMBC layer (“buffy” coat) was collected, washed once with RPMI-1640 and the cells were resuspended in 5ml of RPMI-1640 supplemented with 10% human AB serum (HAB, C-six Diagnostics, Mequon, WI), antibiotics, and anti-PPLO (GIBCO BRL, Grand Island, NY). Red blood cells were lysized by adding 25ml of lysis buffer
(150mM NH₄Cl, 10mM KHCO₃, 1mM EDTA) and incubation for 5 min at room temperature. The remaining cells were seeded out in plastic tissue culture dishes in RPMI-1640 with 10% HAB and cultured for at least 2 hours to remove adherent monocytes. The non-adherent and loosely adherent peripheral blood lymphocytes (PBL) were gently washed off from the dishes and the cell pellets were resuspended in 5ml of 10%HAB.

6.2.8.2 Depleting non-NK cells: T cells, B cells, and remaining monocytes

Non-NK cells in the PBL preparation were depleted using a combination of mouse MoAb reactive against CD3 (T cells), CD20 (B cells), HLA-DR (MHC class II) (Coulter, Hialeah, FL) and goat anti-mouse immunomagnetic beads (Perspective Biosystems, Framingham, MA) as described previously (Matos et al., 1993). Four different mouse hybridoma supernatants were used: 3F5B11 (anti-HLA-DR), Leu4 (anti-CD3), OKT3 (anti-CD3), IF5 (anti-CD20) and their titers were all around 10ul/10⁶ cells. The appropriate amount of each hybridoma supernatant was calculated according to numbers of the cells bearing the particular antigens in the PBL preparation. The cell distribution within PBL population is: 50% HLA-DR+, 60% CD3+, 40% CD20+. Cell-MoAb mixture was incubated for 30 min at 4°C with rotation. After the unbound antibodies were washed away, the cells were resuspended in 5ml of 10%HAB. Appropriate amount (total number of PBL divided by 10⁷ then multiplied by 0.32 gives the volume (in ml) of the beads required) of pre-equilibrated (with RPMI-1640) Goat anti-mouse immunomagnetic beads were added to the cell suspension and incubated for 30 min at 4°C with rotation. The beads were then
removed while adhering to a magnetic plate and the cells in the supernatant were pelleted down and resuspended in 1 ml of 10% HAB. These are enriched NK cells.

6.2.8.3 Sorting CD56+ NK cells

Enriched NK cells were stained with fluorescence-conjugated anti-CD56 MoAb (Becton Dickinson, San Jose, CA) for sorting on a Coulter ELITE flow cytometer (Coulter Immunology, Hialeah, FL). 5X10^5 cells were stained with PE-control and the rest cells were stained with CD56-PE (NK H1-RD1) by incubation on ice for 15 min to 30 min. Excess MoAb were washed off and the cells were resuspended in 500ul of 10% HAB (filtered) for PE-control, in 2ml of the media for CD56-PE. The sorting was performed in the Analytical Flow Cytometry Lab at the Comprehensive Cancer Center of the Ohio State University (Columbus, OH). The sorted CD56+ NK cells were collected in 30% HAB and were usually more than 97% pure by flow cytometric analysis.

6.2.9 Preparation of NK cell cultures

Sorted NK cells (about 1X10^6) were seeded in 10% HAB at a density of 2X10^6 cells/ml in 24-well flat bottom plates (Costar, Cambridge, MA) and incubated at 37°C in 5% CO_2, stimulated with IL-12 (10U/ml), IL-15 (10ng/ml), IL12 + IL-15, or no stimulus (medium alone). At the time points of 0, 1, 4, 12 and 24 hours of incubations, cells were collected and lysed thoroughly by vortexing in 300ul of lysis buffer (Ambion, Austin, TX). Cell lysates were kept frozen at -80°C until they were ready for RNA isolation.
6.2.10 Determination of MIP-4 expression in human NK cells

6.2.10.1 Isolation of total RNA from NK cells

Total RNA from NK cells was isolated using RNAqueous kit according to the manufacturer's protocol (Ambion, Austin, TX). Cell lysates were completely thawed at room temperature before an equal volume of 64% ethanol was added. The lysate/ethanol mixture was applied to a filter and centrifuged for 30 sec at 14,000xg at room temperature. The filter cartridge was washed once with 700ul of Solution #1, twice with 500ul of Solution #2, followed by a final spin for 1 min at 14,000xg to remove trace amounts of wash solutions. RNA was eluted with 60ul of preheated (to 95°C) DEPC-water by centrifugation for 1 min at 14,000xg. The final volume of RNA eluates was reduce to ~16ul using a Speed-Vac at room temperature. RNA samples were stored at -80°C.

6.2.10.2 RT-PCR amplification

First strand cDNA was synthesized using AMV reverse transcriptase first-strand cDNA synthesis kit (Life Sciences, Inc., St. Petersburg, FL). For each reaction, 16ul of RNA and 1ul of pd(N)$_6$ (50ng/ul) was mixed together and heated up to 70°C for 10 min then chilled on ice for 5 min. The following reagents were added in order: 1.0ul of 0.25M DTT, 1.0ul of RNasin (12.5U/ul), 5.0ul of 5X reaction buffer, 1.0ul of AMV-RT (25U/ul). Reactions without AMV-RT were included as no-RT controls. The reactions were incubated for 1 hour at 41°C then stopped by heating for 5 min at 95°C. cDNA samples were stored at -20°C.
RT-PCR amplification using HPRT primers were performed as positive control to check the quality of cDNA samples. The amplification was carried out in a total volume of 50ul with 2ul of each cDNA product, 1X PCR buffer, 200uM dNTPs, 1uM of each primer and 2.5U Taq (GIBCOL BRL, Gaithersburg, MD). Total of 30 cycles of PCR amplification were followed: denaturization at 94°C for 1 min, annealing at 60°C for 1 min, extension at 72°C for 1 min with a final extension for 10 min at 72°C. The PCR products were run in a 2% agarose gel to check the presence of a ~300bp fragment.

The MIP-4 RT-PCR amplifications were carried out in two steps. The primary PCR amplification was performed using MIP-4 primers 5'-TGCCTGCCCAGCATG-3' (For-1) and 5'-CCTCAGGCATTCAAGCTTCAGG-3' (Rev-3). The cycling conditions were as follows: 30 cycles of PCR amplification were followed: denaturization at 94°C for 45 sec, annealing at 55°C for 30 sec, extension at 72°C for 30 sec with a final extension for 7 min at 72°C. The primary PCR product was diluted (1:50) and 1ul of each diluted sample was used in secondary PCR amplification with nested primers from MIP-4 gene: 5'-GGTACCAACAAAGAGCTCTGCTGC-3' (For-2) and 5'-GTCGCTGTATGTATTCTGGAC-3' (Rev-3). The cycling conditions were as follow: 25 cycles of PCR amplification were followed: denaturization at 94°C for 45 sec, annealing at 60°C for 30 sec, extension at 72°C for 30 sec with a final extension for 7 min at 72°C. The PCR products were run in a 2% agarose gel to check for the presence of ~300bp fragment.
6.3 Results

To study the biological function of this novel chemokine gene, MIP-4 recombinant proteins were generated. Initially, the bacterial expression vector pRSET_C was used (Figure 6.1). The cDNA sequence coding for the predicted mature MIP-4 protein was subcloned into pRSET_C at BamHI/EcoRI sites. The vector encodes a 6xHis tag attached to the N-terminal end of the protein and an enterokinase recognition site which allows the cleavage of His tag from the fusion protein, leaving an intact mature chemokine protein. After IPTG induction, the fusion protein was produced as a 12 kDa band in SDS-PAGE (Figure 6.2). However, the vast majority of the fusion protein was insoluble. After purification using TALON affinity resin in the denatured state, the protein could not be resuspended in EK MAX buffer for enterokinase treatment. Most chemokine proteins tested so far, they lose their chemotactic activities if the N-terminals are modified (Holzer et al., 1996). Therefore, another vector system pET-32c was used (Figure 6.3). This system is designed for cloning at high-level expression of proteins fused to the thioredoxin(trxA) tag. Studies have shown that a number of mammalian cytokines and growth factors stayed remarkably soluble in the E.coli cytoplasm under certain conditions (LaVallie et al., 1993). Under mild induction condition with IPTG (final concentration at 0.4mM) for 2.5 hours at 30°C, about half of the fusion protein was soluble. From one liter culture, approximate three milligrams of fusion protein was generated. After cleavage of the TrxA tag, the purified protein was resuspended in PBS at 1ug/ml for bioassays (Figure 6.4).
Because the recombinant protein purified from bacterial hosts might be contaminated with bacterial endotoxin (LPS), which is a potential activator for Ca++ mobilization, chemically synthesized human MIP-4 protein was also included in the in vitro assay. As reported previously, MIP-4 protein was a potent activator of naïve T cells (Adema et al., 1997, Hieshima et al., 1997). The activity of the recombinant human MIP-4 protein in comparison with other chemokines was tested on both naïve and memory T cells with respect to its ability to mobilize intracellular calcium (This assay was performed at Pfizer Central Research under the direction of Dr. Kudeep Neote). Thapsigargin, a non-specific stimulator of calcium mobilization was used as a positive control and for comparison of the level of activation of the leukocyte subpopulations (Figure 6.5). Both SDF-1 (a CXC chemokine) and MIP-4 induced calcium mobilization in the total cell population (no gate), although it was apparent that SDF-1 was more potent and overall the responsive cells were predominantly monocytes (comparing lymphocyte gate with monocyte gate). Analysis of the CD45RA (naïve) and CD45RO (memory) T cell populations revealed that both SDF-1 and MIP-4 induced calcium mobilization in both T cell subpopulations. The CD4+ and CD8+ T lymphocyte subpopulations were then analyzed. Both SDF-1 and MIP-4 induced an increase in the level of intracellular free calcium in both the CD4+ and CD8+ subpopulations (Figure 6.6). The potency of both recombinant and synthetic MIP-4 chemokines was approximately equivalent.

The results of previous studies have demonstrated that MIP-4 is a chemoattractant for T lymphocytes in vitro but exerts no effect on either monocytes or granulocytes (Adema et al., 1997, Hieshima et al., 1997). To determine the target cell
specificity of MIP-4 in vivo, synthetic MIP-4 was injected into the peritoneal cavity of mice, and the cellular exudate was analyzed by differential cell analysis and by flow cytometry (This study was done by Dr. Shaun McColl's lab at University of Adelaide, Australia). Synthetic rather than recombinant MIP-4 was used to avoid the possibility of contamination with LPS in this assay. The concentration of MIP-4 and the stimulation times used were derived from the results of similar studies on other chemokines (Lee et al., 1995). Injection of MIP-4 into the peritoneal cavity had no statistically significant effect on the total number of leukocytes migrating into the peritoneal cavity at either 12 hours or 24 hours post-injection. The majority of cells comprising the exudate were mononuclear phagocytes and granulocytes and MIP-4 did not induce an increase in either the percentage or total number of either of these two cell populations (Figure 6.7). In contrast, flow cytometric analysis of the CD4+ and CD8+ T lymphocyte populations indicated a statistically significant doubling of the number of both of these cell populations (Figure 6.8). This was apparent at 24 hours but not at shorter time points.

Recent studies have suggested that CC chemokines produced by human natural killer (NK) cells suppress HIV-1 entry and replication (Oliva et al., 1998; Fehniger et al., 1998). After depletion of the chemokines MIP-1α, MIP-1β, and RANTES which have been shown to inhibit HIV-1 infection (Cocchi et al., 1995) by immunoprecipitation with their specific antibodies, the NK cell supernatant still retained some of its inhibitory activity. This indicated that additional HIV suppressive factor(s) are present in the NK cell supernatant. Since MIP-4 is highly homologous to MIP-1α, it is reasonable to suggest that MIP-4 might be involved in HIV inhibition.
Under the same stimulation conditions as described (Fehniger et al., 1998), NK cells expressed MIP-4 gene shown by RT-PCR amplification using MIP-4 specific primers (Figure 6.9). The MIP-4 mRNA was present four hours after stimulation with IL-12, IL-15, and IL-12 plus IL-15 cytokines. The expression level increased with incubation time (up to 24 hours tested) and the strongest expression was induced by the combined stimulation with IL-12 and IL-15.

Since the high homology was found between MIP-4 and MIP-1α, which has been identified as one of the major HIV suppressive factors, MIP-4 protein was tested for a potential role in HIV infection. The preliminary results generated by Dr. A. Oliva at NIH have shown that the human MIP-4 protein did not inhibit the uptake of either T-tropic or M-tropic HIV-1 virus. Indeed, it appeared to enhance the HIV uptake.

6.4 Discussion

Growing evidence has suggested a critical role for chemokines in the regulation of immune responses. However, little is presently known concerning the biological significance of MIP-4. Previous studies have examined the target cell specificity of MIP-4 in vitro. Both PARC and DC-CK-1 were shown to be highly specific chemoattractants for naïve T lymphocytes, but not for any other leukocytes tested, including activated T lymphocytes, monocytes or granulocytes (Adema et al., 1997, Hieshima et al., 1997). The results of our studies have significantly extended these previous results by showing that MIP-4 activates both the CD4+ and CD8+ T lymphocyte subpopulations.
Using synthetic MIP-4 protein to avoid possible contamination of recombinant material by LPS, we were also able to extend these previous in vitro observations to the whole animal and confirm that MIP-4 functions as a chemoattractant in vivo for both CD4+ and CD8+ T lymphocytes, and that it exerts no effect on granulocytes or monocytes. This is important since it suggests that MIP-4 might be involved in adaptive immunity at the level of both humoral and cell-mediated immunity. Other chemokines also exhibit the ability to attract both CD4+ and CD8+ T lymphocytes in vitro, including non-ELR CXC chemokines, most of the CC chemokines, as well as lymphotactin and fractalkine and, although there are exceptions, there is no obvious preference for either CD4+ or CD8+ T cells amongst the chemokines, at least under the in vitro conditions tested. It must be stressed however, that our study is the first examining the effect of a chemokine on T lymphocyte subset recruitment in vivo.

Further insights into the possible biological function of MIP-4 derived from studies of its pattern of expression. Unlike many chemokine genes, MIP-4 appears to have a highly restricted tissue and cellular distribution. The results of earlier studies indicate that alveolar macrophages and dendritic cells are a major cellular source of MIP-4 (Adema et al., 1997; Hieshima et al., 1997) and the results of the present study show that macrophages derived from peripheral blood monocytes also express high levels of MIP-4 mRNA. Moreover, lung, lymph nodes, placenta and bone marrow are the major tissues where MIP-4 is expressed. It seems likely therefore that a major function of this chemokine is to attract naïve T lymphocytes towards dendritic cells and macrophages in lymph nodes. This, combined with the documented ability of some chemokines to enhance cellular interactions and lymphocyte proliferation (Taub et al.,
1996) suggests a major role for this chemokine in the generation of primary immune responses.
Figure 6.1  Map of the prokaryotic expression vector pRSET. The system expresses recombinant protein at high level, with 6XHis fusion tag for purification.
Figure 6.2 SDS-PAGE (12%) of the MIP-4 fusion protein generated from pRSET_C.
Lane 1: cell lysate after incubation with TALON resin
Lane 2: cell lysate before incubation with TALON resin
Lanes 3 to 7: washing through
Lanes 8 to 12: eluates with imidazole
Figure 6.3 Map of the prokaryotic expression vector pET32_C. The vector is designed for cloning and high-level expression of protein fused with thioredoxin protein which increases the solubility of the fusion protein dramatically. Cloning sites are available for producing fusion protein containing cleavable 6XHis-Tag and S-Tag for detection and purification.
Figure 6.4 SDS-PAGE (18%) of the MIP-4 fusion protein generated from pET-32C.
Lane 1: cell crude lysate
Lane 2: the fusion protein purified using TALON resin
Lane 3: the protein after the treatment with EK
Lanes 4 and 5: purified MIP-4 protein.
Figure 6.5 MIP-4 mobilizes intracellular calcium in fresh human T lymphocytes. T lymphocytes were purified from the peripheral blood of healthy donors and sorted through fluorescence gates. The cells were then treated with synthetic MIP-4 protein. Two-color flow cytometric analysis was conducted. Results are from three separate donors.
Figure 6.6 MIP-4 protein mobilizes intracellular calcium in activated human T lymphocytes. T lymphocytes were purified from the peripheral blood of healthy donors and incubated with either diluent control or flu antigen for 13 days in the presence of IL-2. The cells were then harvested, resuspended at $1 \times 10^6$ per ml in RPMI 1640, incubated with 2 mM Fluo 3-AM for 20 min at 37°C and then labelled with either anti-CD4 or anti-CD8. Two colour flow cytometric analysis was then conducted. Results are from three separate donors.
Figure 6.7 MIP-4 has no effects on the total number of leukocytes *in vivo*. MIP-4 (10 µg) or PBS were injected into the peritoneal cavity of mice. The peritoneal cells were collected 24 hours later and analyzed. The numbers of mononuclear cells, and granulocytes in the peritoneal cavity as determined by differential staining of cytospins.
Figure 6.8 MIP-4 attracts both CD4$^+$ and CD8$^+$ T lymphocyte *in vivo*. MIP-4 (10 μg) or PBS were injected into the peritoneal cavity of mice. The numbers of CD4$^+$ and CD8$^+$ T lymphocytes in the peritoneal cavity as determined by flow cytometry. These data represent the mean ± s.e.m. of replicates from 7 mice.
Figure 6.9 Expression of MIP-4 in NK cells under cytokine stimulation.

Freshly sorted human NK cells were cultivated either in RPMI-1640 medium alone or with cytokines IL-12 + IL-15. Gene expression was determined by RT-PCR.

A. MIP-4 expression
B. IFNγ expression
C. HPRT expression

Lane 1: NK in medium at time point 0
Lane 2: NK in medium, lane 3 with IL-12 + IL-15 at time point 1 hr
Lane 4: NK in medium, lane 5 with IL-12 + IL-15 at time point 4 hr
Lane 6: NK in medium, lane 7 with IL-12 + IL-15 at time point 12 hr
Lane 8: NK in medium, lane 9 with IL-12 + IL-15 at time point 24 hr
Lane 10: NK with cytokines IL-12 + IL-15 at time point 12 hr
Lane 11: Blank control
Chapter 7

Cloning of the Murine Homologue of the MIP-4 Gene

7.1 Introduction

The strategy of cloning mouse MIP-4 gene was based on the hypothesis that Hu-MIP-4 and Mu-MIP-4 genes have:

(1) a homology which is high enough for cross-species hybridization under low stringency condition;

(2) similar tissue distributions of MIP-4 gene expression;

(3) syntenic chromosomal location.

7.2 Materials and Methods

7.2.1 cDNA library screening

Mouse lung cDNA Uni-ZAP XR library (Stratagene, La Jolla, CA) was screened with human MIP-4 cDNA probe.

The primary library was titered, and plated out, essentially following the manufacturer's protocol (Stratagene, La, Jolla, CA). 8X10⁵ pfu/plate was plated in each 250X250cm plate (Baxter, ) for 12 plates. Plaque lifts was hybridized to 32-P labeled human MIP-4 cDNA (1 to 2X10⁶ cpm/ml) under the same low stringency conditions.
used in the Southern hybridization described above. The positive plaques were cored out and each was resuspended in 500ul of SM (100mM NaCl, 80mM MgSO₄, 50mM Tris-HCL (pH7.5), 2% galatin) and 20ul of chloroform. The supernatants were titered and plated out as secondary libraries. The rest supernatants were kept at 4°C as original stocks.

The secondary, tertiary and quandary screenings were performed following the same procedure. The single plaque was amplified for direct in vivo excision of the pBluescript phagemid sequence from the Uni-ZAP XR vector as the manufacture's recommendation (Stratagene, La Jolla, CA). Phagemid DNA as well as its total phage DNA were double-digested with EcoR I and Xho I. The products were separated in a 1.5% agarose gel and probed with human MIP-4 cDNA under the low stringency condition used in the screening procedure.

7.2.2 Degenerate PCR amplification

Amino acid sequences of human MIP-4, human MIP-1a and mouse MIP-1a were analyzed through Blocks WWW server (http://www.blocks.fhcrc.org). Three pairs of degenerate PCR primers were deduced. One pair of primer which would amplify genomic DNA templates for both MIP-4 and MIP-1a genes was selected for PCR amplification. The forward primer was (from Exon-2) 35bp with sequence as 5'-ATTC(G/C)C(A)A(C)CA(G)GAATTT(C)CA(T)TAGCTGA(C)CTACCTTGAGAC-3' and reverse primer of 32bp 5'-AGC(T)TCC(T)AGGTCGC(A)T(G)GACA(T)T(G)ATT T(C)CTGGACCACCA-3'. Genomic DNA from human and mouse were used as templates and amplified under the following conditions: initial heat denaturization at 94°C for 3
min, one cycle low stringency amplification at 94°C for 30 sec, 45°C for 30 sec and 72°C for 45 sec followed by 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 45 sec and one final elongation at 72°C for 7 min. The PCR products were checked in a 1.5% agarose gel for the presence of 600bp ~ 700bp products. To identify the putative mouse MIP-4 gene sequence, the degenerate PCR products were also cloned into TA vectors for sequencing analysis.

7.2.3 Genomic cloning

7.2.3.1 Screening mouse BAC genomic library with a PCR approach

Primers derived from mouse MIP-1α exon 2 and exon 3 were sent to Genome System Inc. (Genome System, Inc., St. Louis, MO) to screen mouse ES-129/SvJ library Release I DNA pools. The forward primer and reverse primer are as follows: 5'-CG GAAGATTCCACGCCAATTC-3' and 5'-GCATTCAGTTCCAGGTCAGTGATG-3'. The PCR reactions were performed under the conditions of 3 min of heat denaturization at 94°C followed by 35 cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 45 sec and one final elongation at 72°C for 7 min. The expected product is 403bp. The only positive clone from Release I DNA pool, BAC114K21, was ordered directly from Genome System, Inc.

7.2.3.2 Southern hybridization

BAC DNA was prepared using QIAGEN Plasmid Midi Kit (QIAGEN, CA) and digested with three different restriction enzymes: PstI, BamHI and EcoRI. The blots were probed with the 32-P labeled mouse MIP-1α cDNA probe under the standard
stringency and the human MIP-4 cDNA probe under the low stringency, respectively. The low stringency hybridization was performed as follows. The blot was prehybridized in 50 ml of prehybridization buffer (7% SDS, 10% PEG 8000, 1.5X SSPE, 5X Denhart's Buffer, and 100 ug/ml of salmon sperm DNA (sonicated and boiled before adding to the solution)) at 4 hours at 50°C, then hybridized (7% SDS, 10% PEG 8000, 1.5X SSPE, and 100 ug/ml of salmon sperm DNA) to the human MIP-4 probe for 16 hours at 50°C. The blot was washed for twenty minutes at the following conditions: 6X SSC-0.1% SDS at room temperature, 6X SSC-0.1% SDS at 55°C, 4X SSC-0.1% SDS at 55°C, 2X SSC-0.1% SDS at 55°C, 1X SSC-0.1% SDS at 55°C, 0.5X SSC-0.1% SDS at 55°C. The blot was exposed to X-ray film for 3 days at -70°C with intensifying screen.

7.2.3.3 Subcloning of BAC into pBluescript vector for sequencing

PstI bands which were positive to the human MIP-4 probe were isolated and subcloned into pBluescript vector. MIP-4 positive clones were isolated and their DNA were prepared for sequencing using Sequenase 2.0 sequencing kit (Amersham, Cleveland, OH).

7.3 Results

The mouse lung cDNA phage library was screened with a human MIP-4 cDNA probe under low stringency. However, the screening generated no positive clones even after four rounds of selections.
Based on sequence conservation, multiple MIP degenerate PCR primers were synthesized according to the consensus sequences between human MIP-4, human MIP-1α and mouse MIP-1α. Two different primer sets did not produce reliable products and were not analyzed any further. One set of primers amplified fragments at expected sizes so that the PCR products were subcloned into pCRII and sequenced. Most of the clones analyzed turned out to be the mouse MIP-1α gene. Those that were not MIP-1α products were further analyzed by sequencing but they turned to be PCR artifacts.

Based on the proximity of the human MIP-4 gene to the MIP-1α gene, one would expect that the BAC clone 114K21 which is positive for the mouse MIP-1α gene also contains the mouse MIP-4 gene if it exists. The BAC clone contains an insert of about 200kb. The PstI digest of the BAC DNA was probed with the human MIP-4 cDNA. Two bands were detected at 3kb and 4kb while the mouse MIP-1α probe hybridized to 3kb and 5kb bands, respectively (Figure 7.1). As a control, PstI digests of human genomic DNA and mouse genomic DNA were also included in the hybridization but they did not generate convincing signals. The 3kb and 4kb PstI fragments from BAC 114K21 were band-isolated and subcloned into pBluescript. Seven clones were isolated with 3kb, 4kb or 7kb inserts. They were partially sequenced but none has shown any sequence homology to the human MIP-4 gene so far. Since only short sequences have been analyzed, I am continuing to sequence these clones.

7.4 Discussion

cDNA library screenings did not generate any positive clones for the human MIP-4 probe. This could be due to the loss of the clone(s) during the amplification
steps while the library was made, especially when the transcript was expressed at a low
dose.

Degenerate PCR amplification generated MIP-1α gene products mostly. The
degenerate primers failed to amplify any fragments which could be candidates of the
mouse MIP-4 gene. This may be caused by the degeneracy bias toward to the MIP-1α
gene.

Sequencing analysis of the BAC 114K21 subclones might tell us whether or not
the mouse MIP-4 gene exists.
Figure 7.1 Southern hybridization of BAC 114K21 DNA with mouse MIP-1α cDNA probe and human MIP-4 cDNA probe. BAC114K21 DNA was digested with PstI and BamHI respectively. The blots were probed separately with $^{32}$P labeled mouse MIP-1α cDNA under standard stringency or human MIP-4 cDNA under low stringency.

A. probed with mouse MIP-1α
B. probed with human MIP-4

Lane 1: BAC DNA digested with BamHI
Lane 2: human genomic DNA digested with PstI
Lane 3: Mouse genomic DNA digested with PstI
Lane 4: BAC DNA digested with PstI

mw: 1kb molecular weight markers
Chapter 8

Conclusions

This thesis describes the molecular cloning and functional characterization of the human novel CC chemokine gene MIP-4. Using a positional cloning approach, two novel human CC chemokines MIP-4 and HCC-1 were isolated from a YAC contig which is about 1.8 Mb encompassing MIP-1α gene in the cluster region of 17q11.2-q12. The MIP-4 gene shares the highest homology to the MIP-1α gene with 49.5% identity at the nucleotide level and 59.5% identity at the predicted amino acid level. Similar to the other cloned human CC chemokine genes, the genomic structure of the MIP-4 gene consists of three exons and two introns. The gene is spread over a 7.1kb fragment and is located 16kb away from the MIP-1α gene in “tail to tail” orientation, compared with that human MIP-1 β gene is located 14 kb away from MIP-1α in a “head to head” fashion. The remarkable clustering of the human chemokine genes and the high homology between them strongly supports the hypothesis that chemokine gene family arose by duplication followed by divergent evolution.

The expression profile of the MIP-4 gene was investigated using Northern hybridization and RT-PCR amplifications. The gene was transcribed into a message of about 800bp and is present in lung and lymph node at high level. Monocytes cultivated
for 72 hours and T blast cells were identified as the cellular sources of the MIP-4 gene expression. In *in vitro* calcium flux assays, both the recombinant and the synthetic human MIP-4 proteins were shown to be active on both naïve and activated T cells (CD4+ and CD8+) at equivalent potency. Direct injection of the MIP-4 protein into mice has shown that the MIP-4 protein is a potent activator of CD4+ and CD+ T cells *in vivo*. Taken together, it is likely that MIP-4 plays an important role in priming of both cell-mediated and humoral immune responses. The cloning of the human MIP-4 gene and the generation of MIP-4 proteins lays the foundations for the studies of the MIP-4 receptor(s), which will be essential to fully understand the biological role that MIP-4 plays in the immune system. Blocking the MIP-4 signal pathway by neutralizing antibodies against MIP-4 and generating MIP-4 transgenic animal models will be very useful in elucidating the host defense mechanisms that is initiated by MIP-4 and the pathogenesis of diseases that MIP-4 might be associated with.

The fact that chemokines inhibit HIV entry and replication *in vitro* has promoted further investigations of the chemokine-HIV interaction at the signal transduction level. The enhanced uptake of T-tropic HIV-1 virus by the treatment with CC chemokines MIP-1α, MIP-1β or RANTES was blocked by Herbimycin A—an inhibitor of the PTK pathway, but not by Pertussis toxin-Gi protein inhibitor (Dolei, et al 1998). This suggests that the chemokines transduce their chemotactic signals and interact with HIV-1 virus by using different pathways. The preliminary results showed that MIP-4 did not inhibit the HIV infection either by M-tropic or by T-tropic viruses. One explanation is that the MIP-4 protein cannot bind to chemokine receptors CCR5 or CXCR4 which are the major HIV-1 virus fusion cofactor. Further analysis has found that the MIP-4
incubation increased T-tropic HIV-1 virus uptake by PBL 2 to 4 fold. The enhancement of T-tropic HIV-1 entry might be caused by the upregulation of CXCR4 expression or by mononuclear cell proliferation induced by the MIP-4 treatment. Flow cytometric analysis of the target cells in chemokine receptor expression and in cell growth may help to elucidate the mechanism(s) of MIP-4 enhancement to HIV infection. Further investigations in the structural characteristics of MIP-4 and MIP-1α proteins will identify the domain(s) which are critical to inhibition/enhancement of HIV-1 infection and provide important information for drug development in AIDS treatment.
References

1. Ahuja, S.K., Murphy, P.M. 1993 Molecular piracy of mammalian interleukin-8 receptor type B by herpesvirus saimiri. Journal of Biological Chemistry 268:20691-20694


28. Farber J.M. 1990 A macrophage mRNA selectively induced by gamma-interferon encodes a member of the platelet factor 4 family of cytokines. Proceedings of the National Academy of Sciences of the United States of America 87:5238-5242


33. Garcia-Zepeda E.A., Combadiere, C., Rothenberg, M.E., Sarafi, M.N., Lavigne, F., Hamid, Q., Murphy, P.M., Luster A.D. 1996 Human monocyte chemoattractant protein (MCP)-4 is a novel CC chemokine with activities on monocytes, eosinophils, and basophils induced in allergic and nonallergic inflammation that signals through the CC chemokine receptors (CCR)-2 and -3. Journal of Immunology 157:5613-5626


39. Holzer, W., Petersen, F., Strittmatter, W., Matzku, S., von Hoegen, I. 1996 A fusion protein of IL-8 and a Fab antibody fragments binds to IL-8 receptors and induces neutrophil activation. Cytokine 8:214-221


C-C chemokine with a unique 37-amino acid carboxyl-terminal extension. Journal of Immunology 159:2554-2558


84. Rot, A. 1993 Chemokines link the two steps of leukocyte adhesion to endothelium. Immunologist 1:145-147


86. Schall, T.J. and Bacon, K.B. 1994 Chemokine, leukocyte trafficking, and inflammation. Current Opinion in Immunology 6:865-873


activating factor gene expression induced in endothelia cells by IL-1 and tumor necrosis factor. Journal of Immunology 144:3034-3038


