

**UCSF**

**UC San Francisco Electronic Theses and Dissertations**

**Title**

Macromolecular structures of receptor-ligand complexes from developmental neurobiology and cancer biology

**Permalink**

<https://escholarship.org/uc/item/2tv575mz>

**Author**

Butte, Manish J.

**Publication Date**

2000

Peer reviewed|Thesis/dissertation

Macromolecular Structures of Receptor-Ligand Complexes  
From Developmental Neurobiology and Cancer Biology  
by

Manish J. Butte

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO



Date

University Librarian

Degree Conferred: .....

Copyright © 2000  
by  
Manish Janardhan Butte

*To my loving family:  
Amy and Quincy, and  
Chance, in memoriam*

*and my mother, father, and brother whose support of me has never wavered*

## Acknowledgements

Looking for an advisor, I remember my first meeting with Robert Fletterick in his office. He knew me not at all, save that I asked some good questions in a recent lecture. Without the slightest bit of hesitation, he began to offer me various crystallography projects, meetings with potential collaborators, and good advice. I realized immediately that his style, hands-off but completely available for everything, was what my personality needed to succeed. Thus, I learned a lot about leadership from Robert. I also learned about balance: skiing, sailing trips, well-timed vacations, professional meetings, entrepreneurship, and business consulting can all make an otherwise static academic life interesting and refreshing.

My time with Bill Mobley, my advisor on the neurotrophin project, has been equally educational. Bill was my first exposure to a pediatric sub-specialist who was also a scientist, a path I knew I wanted to be on, but how? Bill always had all the answers, and it's no wonder I called him a role model from early on. And despite the enormous workload Bill consistently creates for himself, he always found time to meet and chat, keep up to date on my crystallography progress, and ask good questions. I thank him sincerely for his countless hours of time, patient advice, and scientific curiosity.

Other faculty at UCSF have been very constructive for my development. Most important are Jonathan Weissman, who is the third member of my thesis committee, and Mark Shuman, my collaborator on the urokinase project. I appreciate the many hours Jonathan has spent with me, tossing out ideas and paper references from apparently thin air. Mark is a relatively new colleague in the scheme of graduate school. I only started to work with his lab in mid-1999 and yet he and his lab have made an enormous impact in my life. As a senior graduate student, I was allowed tremendous autonomy to run the uPA/uPAR crystallography project as I saw fit, with two incredible technicians for resources. Lest it go

unsaid, Marion Conn and Jennifer Garrison, my collaborators on the urokinase project, have been superb co-workers on this difficult journey. Other faculty with whom I have had excellent discussions and for whom I owe much of my intellectual development are David Agard, Robert Stroud, Roger Cooke, Charly Craik, and Thomas Kornberg.

While the faculty in graduate school have been great advisors and teachers, most of who I became in graduate school I owe to my friends. And at the top of this list comes Russ Huber. I will never forget the synchrotron trips, the long car ride discussions, the trips to Stanford and Santa Cruz and skiing and Scotland and Spain, and the coffee breaks that became a tradition. We share parallel desires to be good physicians, scientists, entrepreneurs, and family men.

My other 10<sup>th</sup> floor friends, Erin Cunningham, Chuck Sindelar, and Chris Reyes have been the finest of friends as well. There have been too many adventures and mis-adventures to catalog here. Let me only say that I had a lot of fun in graduate school because of these three. I also deeply respect their sense of balance, the knowledge that lab work is not all there is in life. Two other friends, Jennifer Harris and Keith Burdick, stand out among the crowd as well. Jennifer is the kind of person I always wanted to be: smart yet humble, dedicated and yet balanced, organized but not distracted. I want to grow up and be Jennifer someday. I have shared so many adventures with Keith, from towerball to mountaineering to the 30-mile death march through the snows of Yosemite. I will never forget our climbs of Mt. Shasta and Mt. Rainier. Your companionship and enthusiasm are inspirational.

Along these lines, I must thank the members of the Fletterick lab, the sine qua non of my graduate school life. Peter Hwang stands out as the most balanced and brilliant of the lab, by far. There never will be a smarter, nicer, and friendlier soul in the lab. He is a Jedi master of every aspect of life that I can discern, from setting up crystal trays to obtaining

great MAD data sets, from skiing and family life. It has been a great pleasure to know and work with him. Carolyn Sousa, now departed for greener pastures, our lab-manager-by-day and rock-climber-by-night turned a lab that was squabbling and mis-managed into the model of happiness we are today. Jennifer Turner, Linda Brinen, and Stephanie Wang have each influenced me in many positive ways, and I thank you all for wonderful conversations and friendship.

Other UCSF friends I cannot forget include Andrew Bogan, who helped take my appreciation of wine and finance beyond the spectator stage; Sarah (Gillmor) Hymowitz, who showed me how to ski and survive the world of crystallography; Mark Kaplan, the jokester at whose wedding I met my wife-to-be; and Bobby Otilar, who reminded me once that I could have whatever I was willing to work for. My good friends in the late night crew comprised Alan Derman, with whom countless late night pep talks and hilarious grammar sessions often pulled me out of despair; Andy Shiau, whose straight talk and gifted advice about crystallography kept me in a constant reality check; Sherry LaPorte, whose wry sense of humor helped on many tiring nights; Ted Mau, who despite beleaguement found time to encourage and support my efforts; and Julian Chen, an explorer at heart. I thank my friends on towerball team for keeping me healthy, Erin, Russ, Keith, Andy, Christa Nunes, Cari Whyne, and Savannah Partridge.

Many thanks to Lavenida Taqueria, Taqueria Cancun, and Starbucks Coffee.

I thank my family for support, financial, emotional, and spiritual. The most important person in my life deserves the greatest acknowledgement, my wife Amy. Her tolerance of my piteous lifestyle, including numerous nights and weekends in lab, is surpassed only by her incredible support of me finishing my thesis after having left UCSF to start my residency in Philadelphia. She reminds me why I do what I do when I can least

remember it myself. Her cat Chance and I were great friends until his premature demise, and I thank him for being the instrument of my proposal of marriage and for purring loudly when I needed it most. Our new cat Quincy had big shoes to fill, and he has done admirably.

The text of chapter 2 is a reprint of material in *Biochemistry* 1998 Dec 1;37(48):16846-52. The co-authors listed in this publication directed and supervised the research that forms the basis for part of this dissertation. The paper is Copyright 1998 American Chemical Society. The American Chemical Society Joint Committee on Copyrights and Publications automatically grants permission to include the article in this thesis.



# Macromolecular Structures of Receptor-Ligand Complexes from Developmental Neurobiology and Cancer Biology

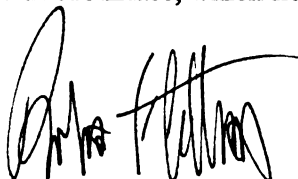
Manish Janardhan Butte

## Abstract

Growth factors play a critical role in the development, maturation, maintenance, and programmed cell death of most cells. For the developing nervous system, one family that bears these responsibilities is the neurotrophins. How the neurotrophins orchestrate these activities depends on the receptor proteins with which they interact. One major aim of my research is to better understand the activation process of the neurotrophins and their receptors by studying their atomic structures. I investigated the neurotrophin called neurotrophin-3, which plays a major role in the development of the peripheral nervous system and neural crest cells.

One process triggered by growth factors is remodeling the shape of the cell, for example to grow in a certain direction. This process appears to be controlled by a cell surface molecule called the urokinase receptor (uPAR). Like other receptors, uPAR sends a signal when activated by a ligand, in this case a protease called urokinase; this signal then informs the cell to reshape itself and migrate. Urokinase is a protease that can cut basement membrane and extracellular matrix, especially when restrained to the surface of a cell by uPAR. Cancerous cells appear to maliciously use these mechanisms to invade into blood vessels and metastasize.

Using X-ray crystallography, I solved the structure of neurotrophin-3 (NT-3) to 2.4 Å. The structure can be compared to those of other neurotrophins to clarify how NT-3 interacts specifically with its receptor. I also determined the low-resolution structure of uPAR with urokinase, which helps us understand how cells make use of this signaling system.



# Table of Contents

Acknowledgements .....	iv
Abstract .....	viii
Table of Contents .....	ix
List of Tables .....	xi
List of Figures .....	xii
Chapter 0. A Simple Introduction to Growth Factors.....	1
Introduction .....	1
Neurotrophins .....	2
Urokinase receptor.....	5
Chapter 1. Neurotrophic factor structures reveal clues to evolution, binding, specificity, and receptor activation .....	7
Abstract .....	8
Introduction .....	9
Sequence Alignments and Molecular evolution.....	11
Neurotrophin structures .....	14
GDNF structure.....	18
CNTF structure .....	20
Conclusion .....	22
Acknowledgements.....	23
Figure captions .....	25
References .....	36
Chapter 2. Crystal structure of neurotrophin-3 homodimer shows distinct regions are used to bind its receptors.....	49
Abbreviations and Textual Footnotes.....	50
Abstract .....	51
Introduction .....	52
Experimental Procedures.....	54
Protein purification and crystal growth.....	54
Data collection and processing.....	54
Molecular replacement.....	54
Refinement. ....	55
Analysis. ....	57
Results and Discussion.....	58
Overall Structure.....	58
Dimer interface. ....	59
Functional implications.....	60
Acknowledgment .....	64
References .....	65
Chapter 3. Urokinase and Urokinase Receptor.....	77
Abstract .....	78
Introduction.....	80
Primary Structure .....	84
Urokinase Receptor.....	84

Urokinase.....	89
Crystal Structures .....	92
Urokinase receptor.....	92
Urokinase.....	94
Protein Production.....	98
Purification.....	98
Urokinase receptor.....	98
Urokinase.....	108
Crystallization .....	113
Data collection.....	116
Processing.....	119
Heavy atom detection and refinement.....	119
Phasing.....	120
Molecular replacement.....	120
Multiple isomorphous replacement .....	120
Multiwavelength anomalous dispersion .....	121
Building uPA and uPAR .....	129
Discussion .....	134
References .....	136

## List of Tables

Table 1-1 .....	24
Table 2-1 .....	68
Table 2-2 .....	69
Table 2-3 .....	70
Table 3-1 .....	117
Table 3-2 .....	118
Table 3-3 .....	122
Table 3-4 .....	124
Table 3-5 .....	126
Table 3-6 .....	133

UC  
Francisco  
RARY

## List of Figures

Figure 1-1.....	25
Figure 1-2.....	26
Figure 1-3.....	28
Figure 1-4.....	29
Figure 1-5.....	30
Figure 1-6.....	31
Figure 1-7.....	33
Figure 1-8.....	34
Figure 1-9.....	35
Figure 2-1.....	71
Figure 2-2.....	72
Figure 2-3.....	73
Figure 2-4.....	74
Figure 2-5.....	75
Figure 3-1.....	86
Figure 3-2.....	88
Figure 3-3.....	90
Figure 3-4.....	91
Figure 3-5.....	93
Figure 3-7.....	102
Figure 3-8.....	103
Figure 3-9.....	104
Figure 3-10.....	105
Figure 3-11.....	106
Figure 3-12.....	107
Figure 3-13.....	111
Figure 3-14.....	112
Figure 3-15.....	114
Figure 3-16.....	115
Figure 3-17.....	127
Figure 3-18.....	128
Figure 3-19.....	131

# Chapter 0. A Simple Introduction to Growth Factors

## Introduction

An essential feature of multicellular organisms is the need to coordinate the growth of cells. Unlike bacteria, which grow and reproduce depending on the availability of nutrients, multicellular organisms evolved mechanisms to regulate growth and reproduction despite the presence of nutrients. The cells could then use the energy that would have otherwise been consumed to specialize. This leads to evolution of the organism as a whole.

For one cell to tell another to grow or to stop growing requires communication. If two cells directly touch, they can communicate via a direct connection of their cytoplasm or by secreting a molecule that the other detects. For longer distances in multicellular organisms, secretion of soluble, diffusible, long-lived molecules was essential for communicating the signal to grow and reproduce or not. Some proteins have these properties.

The developing nervous system in say, a mammal, is an incredibly complex network of interconnected cells of a variety of types. The size and number of neurons needs to be carefully regulated, so that for example the forebrain doesn't crowd out the hindbrain, or so that a nerve cell can grow from the brain all the way down to a touch-receptive cell in the foot. It was postulated in the early twentieth century by Victor Hamburger and others that there should be a way for a developing neuron to receive positive or negative feedback from the environment to keep growing or to stop growing. His neurotrophic factor hypothesis suggested that neurons will compete for a limited supply of positive signal, that the nerve cells perceive the signal as a reward to keep growing, and that the target cells of these neurons would secrete these signals to lure the neurons into developing a connection with them. In fact, all these things occur, and are mediated by proteins called growth factors.

Numerous growth factors participate in the development of the nervous system and other organs, but there are three families that are best understood: the neurotrophin family, the glial-derived neurotrophic factor family, and the ciliary neurotrophic factor family. Of these three, the neurotrophins have been best described both structurally and biochemically.

## Neurotrophins

The neurotrophins are a family of proteins that direct developing neurons to grow, differentiate, and undergo programmed cell death. The prototype member, nerve growth factor (NGF), was discovered in snake venom serendipitously by Rita Levi-Montalcini and Stanley Cohen while they were using the venom to degrade nucleotides in a solution found to stimulate nerve growth. The other members of the neurotrophin family include brain derived neurotrophic factor (BDNF), neurotrophin 3 (NT-3) and neurotrophin 4 (also called neurotrophin 5 or NT-4/5). Genes for other neurotrophins, including 3 human ones called NT-6, and many fish genes called NT-6 and NT-7, have also been discovered.

Neurotrophins are not found in the fruit fly *Drosophila melanogaster* (Butte MJ, unpublished observation) or presumably in more “simple” species than the fruit fly.

The neurotrophins belong to a larger superfamily of proteins called the cystine-knot growth factors. This family includes proteins such as platelet-derived growth factor (PDGF), chorionic gonadotropin (hCG), and many others. The cystine-knot is a set of six cysteines that are bonded pairwise to form a true knot in the protein chain. This knot presumably adds tremendous stability to the protein, an important feature of a growth factor as explained above. Another feature of cystine-knot growth factors is that the proteins are all dimeric: the molecule consists of two individual pieces that are associated. As I will show in the structure of NT-3, this serves to greatly stabilize the molecule over the individual pieces.



The neurotrophins convey their message to a cell via receptor proteins that are on the surface of the cell. There are two families of neurotrophin receptors, the Trk family, which includes TrkA, TrkB, and TrkC, and the p75 family, which is a family of one protein called p75 (also called p75NTR or NTR). The Trk receptors contain an intracellular tyrosine kinase domain that can phosphorylate tyrosine amino acids when the receptor is active. The p75 receptor does not possess any catalytic activity, but contains an intracellular region called the death domain that can interact with other intracellular proteins to convey a signal.

The Trk receptors are monomeric, that is, they exist alone rather than paired or in a larger complex. When a dimeric neurotrophin binds to a Trk protein, it facilitates a second Trk to also bind, as the two “faces” of the neurotrophin dimer can each associate with one Trk protein. When the two Trks are brought together in this manner, their intracellular kinase domains can phosphorylate each other. This phosphorylation is the hallmark of Trk receptor activation and leads to the phosphorylation of a multitude of other molecules, including other kinases, that lie nearby awaiting phosphorylation. A crescendo of phosphorylation results, culminating in activation of transcription factors that direct the transcription of various genes. These gene products cause the phenotypic effects seen in Trk activated nerve cells: growth and differentiation of the cell. One of the genes transcribed by Trk is *upar*, which I discuss in greater detail later.

The Trk receptors TrkA, TrkB, and TrkC are specific for particular neurotrophins NGF, BDNF, and NT-3, respectively. The receptors are all very similar in amino acid sequence to each other, as are the neurotrophins. Thus an intuitive way the specificity may come about is through the amino acids that are different between the molecules. For example, the keys in a building are usually very similar, but the particular different notches in each key allows particular doors to be unlocked. As I will show with the NT-3 structure,

sequence differences do not entirely explain the specificity of the neurotrophins for their receptors. The Trk receptors have a high affinity for their particular neurotrophins. The on-rate of the neurotrophins is much slower than the off-rate, which means that the neurotrophins are reluctant to bind but stick well after they bind.

These kinetic properties are troublesome for some cells, which need to detect minute quantities of neurotrophins in their environment. These cells make use of the p75 receptor to facilitate detection of the neurotrophin signal. P75 is not specific to any particular neurotrophin. It also has a high affinity to the neurotrophins, but unlike the Trks, it has a very fast on-rate. It has been speculated that p75 binds a neurotrophin and then hands it off to a nearby Trk receptor for signaling. As I will show with NT-3, the structure the molecule and experimental data regarding which parts of the molecule are important for binding to the receptors, support this possibility.

P75 can be found by itself on the cell-surface. When these solitary p75 molecules are activated by binding the neurotrophin, the signal sent to the cell is not for growth and differentiation but rather for death. Cell death in the nervous system occurs for a variety of reasons. Neurotrophin-mediated cell death may be related to the neurotrophic factor hypothesis, that neurons prosper or die based on competition for limited growth factor resources.

Curiously, when both p75 and Trk receptors are expressed on a cell surface, neurotrophins can be shown to activate the Trk receptors with a much higher affinity than with either receptor type alone. The on-rate of this combined interaction resembles that of p75, while the off-rate resembles that of the Trk receptors. This leads some to believe that p75 somehow initially binds the neurotrophin and then hands off the molecule to the Trk

receptor. As I will show in the NT-3 chapter, this possibility is supported by the crystal structure of the molecule.

### Urokinase receptor

When a cell is stimulated to grow, this growth comes about by a series of regulatory reactions that result in the boundary of the cell to be expanded in particular direction, for the cell in that direction to be removed, whatever is in the way of and for the contents of the cell to be pushed in that direction. The urokinase receptor plays a critical role in these reactions. One major function of the receptor is to bind urokinase, a rather-specific, extracellular serine protease whose primary specificity is for the protein plasminogen. Plasminogen is proteolyzed by urokinase to become plasmin, which is a broad-spectrum protease that can degrade basement membrane, the connective tissue that underlies most epithelial cells. By binding urokinase to a cell surface, urokinase receptor focuses the degradation of basement membrane at particular sites.

Urokinase receptor (uPAR) also appears has a direct affect on the growth and migration of a cell. This may result from signals sent by uPAR upon activation by its ligand, urokinase. Some evidence suggests that the activation of uPAR that leads to migration is in fact a cleavage of the receptor molecule between two of its domains. The structure of the linker between domain 1 and 2 would help clarify the role of this region in chemotaxis and migration.

Finally, uPAR may play an important role in regulating the activity of growth factors. Some evidence surrounding the neurotrophin NGF suggests that it cannot induce its differentiative effects unless uPAR is expressed on the cell surface. Moreover, uPAR must not be blocked by an antibody. This suggests that growth factor receptor activation triggers a downstream signal to uPAR, which may require a ligand to allow cell growth and migration

to occur. How uPAR might receive a signal from within the cell is not known. It is thought that receptors like uPAR can signal to within the cell by clustering intracellular tyrosine kinases of the Src family. How uPAR, which has no intracellular domain, can influence these molecules is still unknown. That they both cluster to lipid-dense regions (called rafts) in the plasma membrane offers a clue.

## **Abstract**

The neurotrophin family, the glial-derived neurotrophic factor family, and the ciliary-neurotrophic factor are the best-described growth factors specific for developing neurons and neural-crest cells. As might be expected for regulatory molecules of the complex central and peripheral nervous system, these factors show considerable receptor specificity and cross-talk. Thanks to a decade of intense research by numerous labs, the structures of many of these factors are now available. We discuss the structural bases of receptor binding, specificity, and activation in each of these systems. Using structure-based sequence alignments, we also address evolutionary implications of these molecules and their receptors. We also suggest further directions for research in the structure and function of these neurotrophic factors.

## **Keywords**

neurotrophin; glial derived neurotrophic factor; ciliary neurotrophic factor; structure; alignment

## Introduction

Many growth factors play a critical role in the development, maturation, maintenance, and programmed cell death of the nervous system. While dozens of proteins are implicated in these roles, there are currently three families of well-understood growth factors that are specific to the nervous system: the neurotrophin family, the glial-cell derived neurotrophic factor family, and the ciliary neurotrophic factor. The members of these families signal the developing and mature organism to grow, form differentiated cells, and survive disease or injury.

Nerve growth factor (NGF) is the prototype member of the neurotrophin family, which has at least five other members. NGF was discovered in the 1940s by Stanley Cohen and Rita Levi-Montalcini in moccasin snake venom [1]. NGF and its siblings, brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3), NT-4/5, NT-6, and NT-7, provide the signal for growth, differentiation, and programmed cell death for developing and mature neurons and other neural-crest-derived cells. Neurotrophin signaling underpins the development of numerous biological systems, including the brain and peripheral nervous system. Beyond these, neurotrophins have been found to play a role in retinal [2], cochlear [3], and heart development [4,5]. Moreover, neurotrophins are essential for maintaining function of neurons in the adult. In addition, derangements of the neurotrophin signaling system are implicated in Alzheimer's disease [6], peripheral neuropathy [7,8], neuroblastoma [9-11], and many of the neurodegenerative and neuronal storage diseases. Indeed, BDNF and NT-3 have been in clinical trials for diabetic peripheral neuropathy. The neurotrophins signal through two receptor types (Figure 1a). First are the Trk receptors, a family of three nanomolar-affinity receptors (called TrkA, TrkB, and TrkC) each specific for a neurotrophin (NGF, BDNF or NT-4, and NT-3, respectively) with an intracellular tyrosine kinase domain

that stimulates the MAP (mitogen-activated protein) kinase pathway when activated. The second receptor is p75, a nanomolar-affinity TNF-R (tumor necrosis factor receptor) family member that signals through an intracellular death domain to activate the cellular apoptosis machinery. Excess p75 receptors may also act in consort with Trk when neurotrophin concentrations are low, producing a sub-nanomolar binding site [12], though details of this interaction are not well understood.

The glial-derived neurotrophic factor (GDNF) was found in a search for secreted factors that promoted neuronal survival in primary culture by Leu-Fen H. Lin in 1993 [13]. GDNF was cloned from a glial cell line culture, and quickly established prominence as a potent stimulant of dopaminergic neurons. Three other members of the GDNF family have been identified: persephin [14], neurturin [15], and artemin (also called enovin and neublastin) [16]. GDNF has been shown to play a critical role in the development of the kidney [17-19] and the enteric nervous system [20]. The GDNF system has an important medical role, as mutations have been linked with medullary thyroid tumors [21], the multiple endocrine neoplasia type 2 (MEN 2A and 2B) syndrome [22], and Hirschprung's disease [23,24]. Like the neurotrophins, the GDNF family signals through two receptor families (Figure 1b): the GDNF family receptor alpha (GFR $\alpha$ ) family of glycosylphosphoinositol (GPI)-linked receptors and Ret, a tyrosine kinase receptor. GDNF family members predominantly signal in a two-step mechanism, first uniting two GFR $\alpha$  receptors, which then join with two Ret receptors, forming a heterohexameric complex. This then activates numerous intracellular signaling mechanisms including the MAP kinase and phosphatidylinositol 3-kinase pathways [25]. Each GDNF family member has a preferred GFR $\alpha$  receptor and vice versa, though there is considerable cross-talk between the receptors (Figure 1b) [26-28].

Ciliary neurotrophic factor (CNTF) was discovered in 1979 by Silvio Varon's group in a search for growth factors in the ciliary ganglion [29,30]. The primary role of CNTF remains elusive, though it contributes to the development and differentiation of the retinal rod neurons as well as other neurons. It may also contribute to regulation of weight [31] by its influence on the hypothalamic arcuate nucleus, and is currently in clinical trials for this purpose. CNTF is structurally similar to the interleukin-6 family of hematopoietic cytokines, and it signals using common cytokine receptor components. CNTF first associates with CNTFR $\alpha$ , a specific GPI-linked receptor, then binds to the cytokine receptors gp130 and LIFR (leukemia inhibitory factor receptor), which activate the JAK-STAT (Janus Kinase/Signal Transducer and Activator of Transcription) signaling pathway [32].

The result of a decade of intense research, the crystal structures of many neurotrophic factors are now available. In addition, the recent explosion of genomics research has given us dozens of amino acid sequences of neurotrophic factors from dozens of species. Evidence from these sources in addition to the results from biochemistry experiments reveal clues to the evolution, binding, and specificity of these molecules. This paper attempts to summarize these results and suggest further directions for research in the structure and function of the neurotrophic factors.

### Sequence Alignments and Molecular evolution

Comparing the amino acid sequences of the neurotrophic factors can give insight to the evolutionary history of these families and can help reveal functionally important regions of the molecules. A structure-based sequence alignment was prepared using all currently known neurotrophins (Figure 2), representing over 30 species. As the figure shows, there are six major subfamilies of neurotrophins, including NGF, BDNF, NT-3, NT-4/5, the human NT-6s, and the fish NT-6 and NT-7. Two NGF-like molecules from the fowlpox virus



genome are included as an aside, as it is not known if these proteins are even transcribed, though it is easy to propose competitive advantages conferred to a virus that can secrete NGF, including promoting survival of infected cells [33].

Previous works have divided neurotrophin residues into two categories, conserved and variable, based on early sequence alignments [34]. As we can see, most of the residues that participate in secondary structure are strongly conserved both within and between subfamilies, though there are numerous exceptions. The N- and C-termini and loops 2 and 3 show modest diversity between families. Loop 3 has a large insert in the NT-4 and NT-6 subfamilies. In fish NT-6, this region is thought to confer binding to heparin as a regulatory control mechanism [35]. Interestingly, two cysteines that make up the knot are missing from the human neurotrophin-6 molecules. Those residues implicated in binding that are conserved across families are thought to represent a common interface to the Trk receptors, while the unique ones may represent elements of specificity [36].

Neurotrophin evolution has been studied using phylogenetic trees that organize the relationships between their amino acid sequences [37]. Early works used limited subsets of the currently known neurotrophins [37-43], though the general conclusions have been consistent: The ancestral neurotrophin underwent a gene duplication to form the branches that would lead to NGF and NT-3 on one side and BDNF and NT-4 on another. Another round of gene duplications resulted in the current set of four principal neurotrophins in the tetrapods. No ortholog of the NT-4 gene has been found in teleost fishes, and no ortholog of the NT-6 or NT-7 genes has been found in tetrapods (although the names are similar, the human NT-6  $\alpha$ ,  $\beta$ , and  $\gamma$  human genes are closer to NGF). These findings place constraints on the timeline of gene duplications to before 400 million years ago [43]. If a similar tree is prepared for the Trk receptors, the general pattern of interrelationships is upheld: TrkA and

TrkC share a common branch of a tree, while TrkB lies on a separate branch [37]. That the receptor and ligand trees are “parallel” supports the notion of co-evolution of these two families, which has been seen in other protein families [44].

A similar sequence alignment was prepared for the known GDNF family members (Figure 3). The sequence alignment reveals strong conservation, especially within secondary structure elements. Within each member’s subfamily, most changes are conservative. GDNF shows a significantly larger N-terminal tail than is seen in the other members of the family. This tail has been shown to be unimportant to GFR $\alpha$ 1 binding [26], and its role is currently unknown. The helical region in persephin is likely one turn shorter than the other members, and the succeeding loop region is more glycine-rich, perhaps to accommodate more dramatic structural turns needed to return to the next beta strand.

Until very recently, CNTF was in a subfamily by itself. The dramatic contrast between the phenotypes of the gene deletion mice for CNTF and CNTFR $\alpha$  knockout mouse suggested the existence of another ligand for CNTFR $\alpha$ , which was recently discovered [45]. Figure 4 shows the structure-based sequence alignment of the known CNTFs as well as the newly identified molecule, cardiotrophin-like cytokine (CLC). The alignment shows strong conservation among the CNTFs, with the exception of the chicken homologue, which may reflect a different role for CNTF in that species [46]. The alignment also shows that CLC is only weakly related to the CNTF family, with 23-27% identity between CLC and the various CNTFs and with a significantly longer N-terminal region. A handful of charged and bulky hydrophobic residues are preserved, especially in helices B through D. More importantly, the residues identified by mutagenesis to be responsible for binding CNTFR $\alpha$  (site I, discussed below) appear to be well-conserved between CNTF and CLC, suggesting a common mechanism of binding.

## Neurotrophin structures

The neurotrophins and GDNF family members are members of a large superfamily of growth factors that contain a “cystine knot”. This family includes TGF- $\beta$ , human chorionic gonadotropin, platelet derived growth factor, vascular endothelial growth factor, and many others. The cystine knot consists of three disulphide bonds that form a true knot of the polypeptide chain. All the members of this family exist exclusively as dimers, though a variety of arrangements of the protomers are seen, including heterodimers and homodimers, covalent and non-covalent association of the protomers, and different spatial configurations of the protomers (head-to-toe, head-to-head, and skew). Oddly, there are other non-growth-factor proteins that have three disulphide bonds in a knotted arrangement, but with a different configuration of disulphide bonds. One example of this is the invertebrate toxin, omega-conotoxin, which curiously acts as a calcium channel antagonist.

At present the three-dimensional structures of most of the human neurotrophins have been determined (Table 1). These structures can be aligned to reveal that the regions of similarity are much larger than was suggested from sequence alignments [34]. The core structure consists of two pairs of intertwined two-strand beta sheets, joined by three disulphide bonds (Figure 5). There are also three shorter beta strands leading to beta turns and loops. The four core beta strands are virtually superimposable across all the structures with less than 1 Å deviation. Loop 3 is the most different among the four neurotrophin structures. NT-3 has a single loose helical turn in this region. NT-4 has a large insertion in this portion, including a small eight-residue beta turn, but the structure of much of this region was not clearly determined (PDB entry 1b8m chain B or 1b98). This region also appears unstructured in the TrkA-NGF co-crystal structure, and does not appear to lie on the interface with TrkA domain 5. It may play a minor role in the interface with the p75

receptor [47-49] in NT-3 and NGF. The N- and C-termini are highly variable in both sequence and structure among the neurotrophins, and the temperature factors of these loosely structured residues are comparatively high. The N-terminus assumes a helical structure upon binding the Trk receptor, as will be discussed below. The role of the C-terminal residues is not known, though they may play a role in p75 binding in NT-3 [48] and NGF [50].

Structural details can also be used in conjunction with alanine-scanning mutagenesis, chimeras, and deletion mutations to synthesize a coherent model of receptor binding. However, there are at least three major confounding factors relevant to the neurotrophin literature to simply mapping the results of loss-of-function mutation experiments to reveal functionally important regions. Numerous early experimenters produced but failed to purify mutant neurotrophins (for example, [50,51]). In a heterologous expression and assay system, such as transiently transfected PC12 cells, wild-type neurotrophins additionally produced by the cells could easily confound experiments [52]. Unpurified proteins might be mixed with other cellular products like kinase inhibitors, which could explain away otherwise significant kinetics results or biological assays. Second, some mutations, especially ones that cause significant structural rearrangement or instability, for instance by damaging the dimer interface, can result in limited expression of the mutant protein. Critical residues that lie on the binding interface might thus be unintentionally excluded from analysis merely because of low production yield. Finally, mutations in one part of a molecule may show loss-of-function, and thus be incorrectly included in the putative interface, by inducing structural changes in a receptor contact surface in another part of the molecule. For these reasons, the actual interface has always been found to be a subset of those residues identified by mutagenesis [53].

With these limitations in mind, site-directed and alanine-scanning mutagenesis can still be useful tools for inferring the binding interface in the absence of a co-crystal structure [53]. The results of several mutagenesis experiments over the past ten years have revealed a consistent pattern of spatially distinct portions of the neurotrophin dimer molecule that are critical to Trk and p75 receptor binding. The structure of the NGF dimer in complex with domain 5 of the TrkA receptor validates many of these experimental predictions (Figure 6a). By comparing the residues involved in the interface between TrkA-d5 and NGF with the corresponding sequence alignments for the neurotrophins and the Trk receptors, we can see that they form two groups of conserved and non-conserved residues.

One set of residues forming a large patch and centered on Arg 103 form a common set across all the neurotrophins and receptors (Figure 6b). As these residues are highly homologous and consist mainly of hydrophobic and aromatic side chains, they probably mediate the bulk of common binding affinity between neurotrophin and receptor. Another example is the one hydrophobic region is centered on Trp 21, a universally conserved residue that is principally buried by Phe 86 (which is either Phe or Tyr in all neurotrophins) and Phe 101 (which is either Phe or Trp in all neurotrophins) on the neurotrophin side and on the receptor side, His 353 (universally conserved), the side chain of Met 379 (either Met or Leu in TrkA, B, and C), and Pro 382 (universally conserved).

The other portion of the TrkA interface centers on the N-terminal residues of the neurotrophin. As mentioned before, there is little conservation of these residues, both on the neurotrophin side and on the receptor side. These residues, which were disordered in all the unbound neurotrophin structures, formed a one and a half turn helical arrangement in the complex between TrkA and NGF. When NT-3 and BDNF chimeras containing the NGF N-terminus were constructed, they showed enhanced ability to bind TrkA [54,55] supporting

the importance of this region for NGF specificity. NGF binding buries two helical hydrophobic residues and creates a salt bridge across the interface. BDNF, NT-3, and NT-4 do not share the same pattern of residue-types at the N-terminus, and TrkB and TrkC also differ in their corresponding interacting residues [36]. Together, these results suggest that the N-terminal residues help determine receptor binding specificity and that each neurotrophin probably uses a different specific interface with its cognate receptor in this region.

Other specificity-determining residues appear to lie scattered within the common binding site. For example, NGF point-mutants introduced from the NT-3 sequence that conferred TrkC binding were centered around Gly 23 in one study [56]. Another study suggested TrkC specificity is determined by broadly separated patches across the length of the NT-3 molecule [57].

The putative neurotrophin binding site for the p75 receptor has shared similar attention to that of the Trk receptors. Unlike the Trk receptors, p75 binds all the neurotrophins equally, with high affinity [58] and much faster on- and off-rates [12,59]. While the neurotrophins can clearly signal in the presence of either the Trk or p75 receptor alone, evidence suggests that excess p75 can improve Trk signaling by 25-fold when neurotrophin concentrations are low [12,60]. This may imply either simultaneous binding of p75 and Trk receptor to a neurotrophin (the hand-off model), which is not structurally excluded [36,49] (Figure 6a), or in serial steps that improve Trk kinetics by increasing the local concentration of neurotrophin. Supporting these models is recent evidence that suggests p75 and Trk receptors may lie in preformed heteroreceptor complexes on the cell surface [61-63].

One unique feature of the p75-neurotrophin interface is the contribution of charged residues rather than buried surface area to stabilization of the complex [64]. Primarily

positively charged residues on the neurotrophin side of the interface correspond to negative charges on p75, creating a network of paired charges. Mutating these charged residues significantly affects the local net charge and disrupts binding, which has been seen experimentally [65-67]. NGF, NT-3, and NT-4/5 display positively charged Arg, His, and Lys residues on Loop 1, while BDNF displays these residues on the spatially close Loop 4 (Figure 6a). Besides these charged clusters, residues in Loop 3 [47,68] and the C-terminus [48,50] have been shown to play a less significant role in p75 binding.

While neurotrophin structural research is the most advanced among the three families discussed here, numerous open questions remain. Though the structure of TrkA domain 5 in complex with NGF significantly advanced understanding of their interaction, evidence supporting the role of other TrkA domains is left to be explained. In addition, the determinants of specificity for BDNF and NT-3 could be elucidated by co-crystal structures of these neurotrophins with their receptors. The structural basis of p75's facilitation of Trk activation also remains unclear. The structure of the p75-neurotrophin complex would clarify the structural bases of binding and promiscuity of this receptor. Finally, the increasing interest in peptides [69-71] and small molecules [72] that mimic the neurotrophin-receptor interaction should drive structural examinations into these alternatives to define the minimal interface needed for receptor activation.

### GDNF structure

Like the neurotrophins, members of the GDNF family can interact with two different receptors, probably simultaneously. Within the GDNF family, only the structure of GDNF itself has been determined. It shows the usual cysteine-knot fold, with two pairs of anti-parallel twisted beta strands tightly joined by three disulphide bonds. The dimer arrangement, however, is anti-parallel, has an interchain disulphide bond, and due to a large

lack of overlap, has the appearance of two “fingers” on each side of the molecule (Figure 7). This places GDNF in the TGF- $\beta$  subfamily of cysteine-knot growth factors. Because the protomers are arranged anti-parallel, there is an explicit left-right symmetry created, suggesting symmetric binding sites for a dimerized receptor. Between strands three and four, where the neurotrophins have a large variable region, GDNF has a three-turn alpha helix. The large N-terminus was not seen in the crystal structure, suggesting that this region is flexible in solution.

Limited mutagenesis has been performed on the GDNF molecule to determine its interaction site with the Ret and GFR $\alpha$  receptors. Eketjäll and colleagues extensively changed surface-exposed GDNF residues to alanine and determined binding to Ret and GFR $\alpha$ 1 using a steady-state competitive binding assay [73]. These results improve on other works [74] to show that there are two principal modes of GDNF signaling. The first involves GDNF activating a preformed Ret-GFR $\alpha$ 1 complex. This model is supported by evidence that GDNF mutant Tyr 120->Ala is deficient in binding to GFR $\alpha$ 1 but can still activate Ret [73], and that ligand-independent Ret signaling is enhanced by the presence of GFR $\alpha$ 1 [75]. Acidic residues along finger 1 were shown to be critical for binding to GFR $\alpha$ 1 (Figure 7) [73]. Baloh et al. showed that two patches along finger 2 transplanted from GDNF into a persephin backbone are sufficient to induce activation of GFR $\alpha$ 1-Ret suggesting overlapping sites for each receptor [26]. Surprisingly, the positively charged band of residues at the center of the dimer appeared to play no role in GFR $\alpha$ 1 binding. Overlapping sites do not necessarily preclude simultaneous binding of both receptor types to GDNF, as the site on each end of the molecule could be used to associate with a different receptor. GFR $\alpha$ 2 and GFR $\alpha$ 3 binding and activation appear to require both the finger 2 regions needed in GFR $\alpha$ 1 activity as well as part of the loop and helix in the heel region [26].



A second structural model of GDNF signaling, launched via Ret-independent GFR $\alpha$ 1 activation, has been discussed [27,76]. Alanine-scanning mutagenesis revealed that acidic and hydrophobic residues along GDNF fingers 1 and 2 affected binding to GFR $\alpha$ 1 but not Ret activation. Increasing evidence supports the notion that GPI-linked receptors, which have no transmembrane or intracellular region, can signal via activation of Src-family tyrosine kinases by virtue of sharing a special, detergent-insoluble patch of membrane [77-79]. Clustering of the extracellular receptors is thought to induce activation of the intracellular kinases by proximity. Whether this functionality of GFR $\alpha$ 1 has a biological role remains to be demonstrated.

These results suggest that acidic and hydrophobic residues along the exposed fingers play a critical role in mediating GDNF activation of both the Ret-GFR $\alpha$ 1 complex and GFR $\alpha$ 1 alone. Further work is needed to clarify the biological roles of each of these signaling mechanisms. Structures of GDNF in complex with each of the receptor types, taking advantage of GDNF mutants that confer specific binding, would help reconcile apparently overlapping binding sites and inconsistencies in the biochemical evidence. Furthermore, structures of the other GDNF family members are needed to detail the structural basis of receptor specificity and cross-talk.

### CNTF structure

CNTF is not a cysteine-knot growth factor, but rather a four-helix bundle belonging to the interleukin-6 family of hematopoietic cytokines. The crystal structure revealed a dimer, thought to be an artifact of the high concentrations used in crystallization [80]. The basic structure shows four helices, named A – D, with two long cross-over loops (AB and CD) and one short loop (BC) (Figure 8). Much of the AB and the C-terminal part of the CD loops were not seen well in the electron density.

CNTF is thought to bind three separate receptor partners: CNTFR $\alpha$ , gp130, and LIFR. Both immunoprecipitation [81,82] and gel filtration with analytical ultracentrifugation [83] experiments performed on IL-6 and CNTF support the notion that the active CNTF receptor complex is hexameric with stoichiometric ratios of 2:2:1:1 of CNTF, CNTFR $\alpha$ , gp130, and LIFR. However, these experiments were performed at concentrations well above the relatively weak dimerization constant of CNTF (40  $\mu$ m), which could explain why large complexes were seen. There is also evidence from molecular modeling that the hexameric complex is inconsistent with the known size of CNTF [84]. In contrast, two additional models of the CNTF receptor complex have been proposed that are consistent with published mutagenesis data (Figure 9) [85]. The simplest of these suggests a tetrameric complex and still preserves the known interactions.

The CNTF molecule is thought to contain three binding sites, for each its three receptors. These sites are necessarily disjoint, as these receptors bind at once, and are numbered sequentially in analogy to the growth hormone receptor model. Careful regional mutagenesis and chimeras of CNTF and IL-6 have shown that site I, which is responsible for binding the respective alpha receptor (CNTFR $\alpha$  or IL-6 $\alpha$ ), comprises the C-terminal AB loop and the C-terminal D helix [80,84,86-91]. Site II is the gp130 binding site (the primary of two such sites in IL-6, and the only such one in CNTF) which has been localized to residues on the A and C helices [80,92-94]. While site III in IL-6 plays the role of an additional gp130-binding region, in CNTF it is responsible for binding LIFR. This surface is distributed across three regions that are spatially located at one end of the helical bundle, comprising the C-terminal A helix and N-terminal AB loop, the BC loop, and the C-terminal CD loop and N-terminal D helix [85]. The minimal surface needed to bind LIFR may be

smaller, as is seen in the leukemia inhibitory factor (LIF) interaction surface with LIFR [95,96].

Structural and biochemical research on CNTF is the least complete of the three families discussed in this paper. Open questions remain concerning the specific mechanism of receptor recruitment and activation. While IL-6-related results can be extrapolated to the CNTF system, obvious differences in gp130 and LIFR interaction highlight the need for specific investigations into CNTF binding. The structure of CNTF in complex with each of the receptors would refine understanding of the binding sites and lend support to particular models of receptor assembly. Furthermore, the discovery of a new CNTFR $\alpha$  ligand, cardiotrophin-like cytokine (CLC), demands a structural investigation. Common site I residues between CNTF and CLC should be validated by structural investigation, as these would circumscribe the surface needed for CNTFR $\alpha$  interaction.

## Conclusion

In this article we consider the three major families of neurotrophic growth factors: the neurotrophins, the glial-derived neurotrophic factor (GDNF) family, and the ciliary neurotrophic factor (CNTF). The crystal structures of the neurotrophins show the classic cysteine-knot growth factor structure with head-to-head subunits forming a non-covalently linked dimer. The only significant differences between neurotrophins are found in the loops and turns between beta strands, especially Loop 3. Mutagenesis and other biochemical techniques, as well as co-crystal structure of TrkA with NGF, suggest that neurotrophin-Trk binding revolves around a patch of common interactions centered around Arg103 and a patch of specific interactions primarily centered on the N-terminus. GDNF and its family members are cysteine-knot growth factor dimers arranged head-to-toe, covalently linked by a disulphide bond. Structural and sequence analysis of the GDNF family suggests that acidic

and hydrophobic residues along fingers 1 and 2 are necessary for binding and activation of the Ret-GFR $\alpha$ 1 preformed complex, the principal form of GDNF activity. A distinct site may be required for Ret-independent GFR $\alpha$ 1 activation, though how this signal is transduced is not well understood. CNTF is a four-helix bundle in the IL-6 family of hematopoietic cytokines. It binds to three different receptors simultaneously to activate a signal for neuronal survival. Drawing from mutagenesis and chimera experiments with IL-6, the three binding sites have been localized to distinct surfaces on the CNTF molecule. Open questions remain about the true stoichiometry of the receptor complex and the structural basis of coordinated receptor binding.

### Acknowledgements

The author thanks Dr. Charles Howe for evaluation of the manuscript, and Dr. David Christianson and Dr. Robert Fletterick for advice and space. MJB's research was supported by the Office of Naval Research.

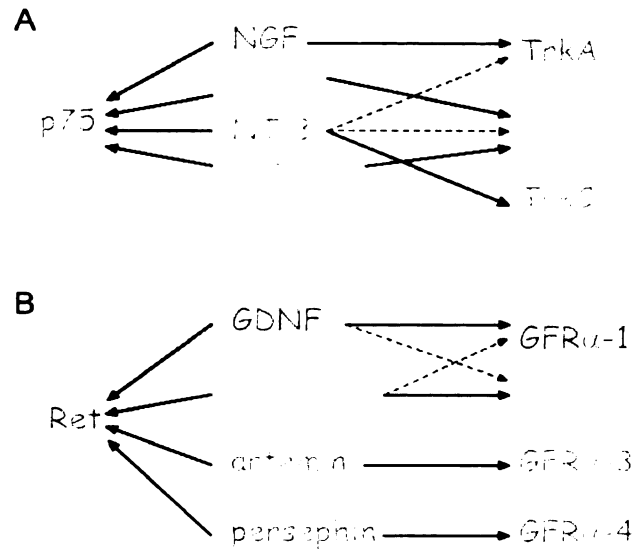
Table 1-1

List of all the neurotrophic factor structures determined to date. Protein data bank

(PDB) entries can be found at <http://www.pdb.org>.

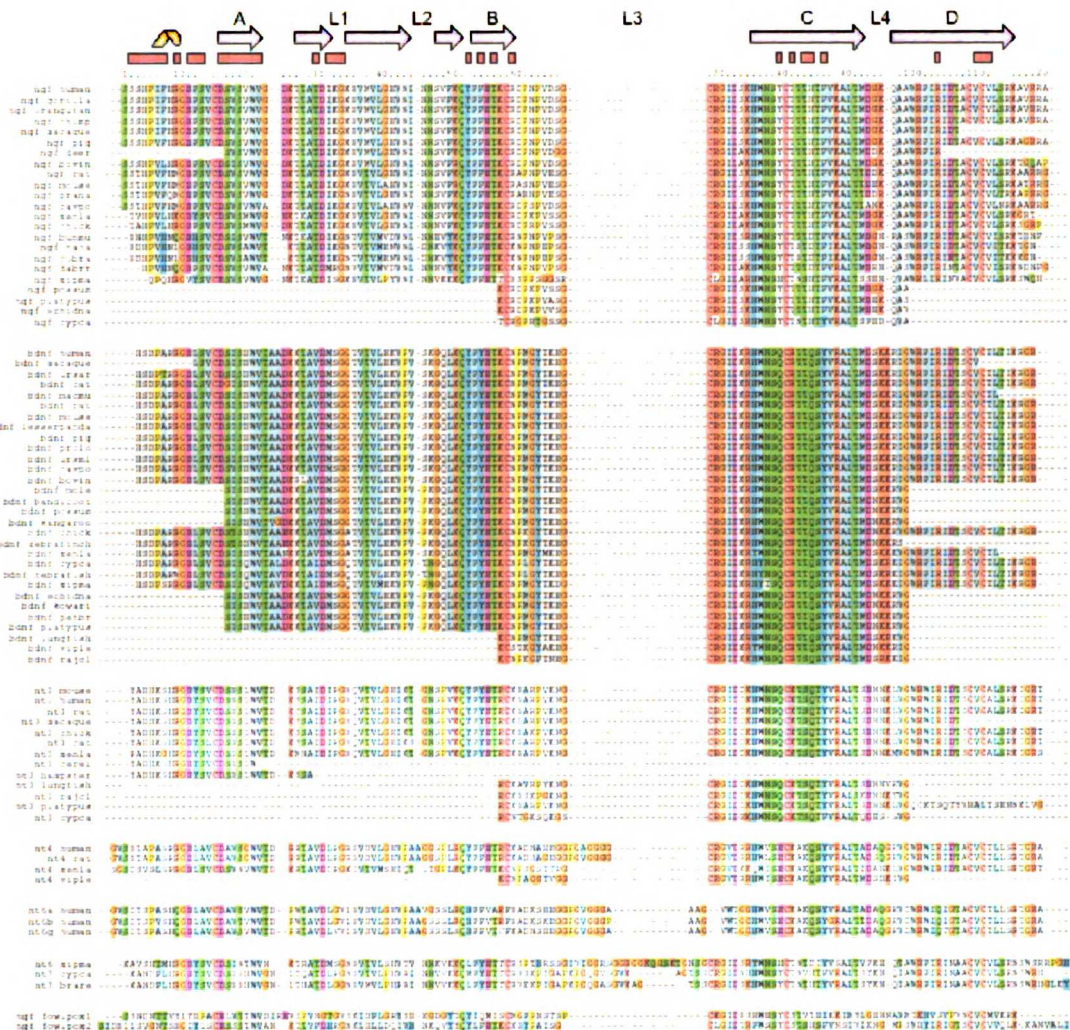
Molecule	PDB ID code	Highest Resolution	Reference
NGF homodimer	1bet	2.3	[97]
NGF homodimer	1btg	2.5	[98]
BDNF/NT3 heterodimer	1bnd	2.3	[99]
NT3 homodimer	1nt3	2.4	[49]
NT4 homodimer	1b98	2.75	[100]
NT3 homodimer	1b8k	2.15	[100]
BDNF/NT4 heterodimer	1b8m	2.75	[100]
NGF + TrkA domain 5	1www	2.2	[36]
GDNF	1agq	1.9	[101]
CNTF	1cnt	2.4	[80]

## Figure captions



### Figure 1-1

(a) The neurotrophins signal through two separate receptor systems, the Trk family of tyrosine kinase receptors, and p75, a death-domain containing TNF-R-like receptor. While significant specificity is seen among the Trk receptors (thick lines), there is some cross-talk that may be evolutionarily reconciled (dashed lines). p75 binds all the neurotrophins equally. (b) The GDNF family of neurotrophic factors also binds two separate receptor systems, Ret, a GPI-linked surface receptor, and the GFR $\alpha$  receptors. Like the neurotrophins, considerable specificity and cross-talk is seen.



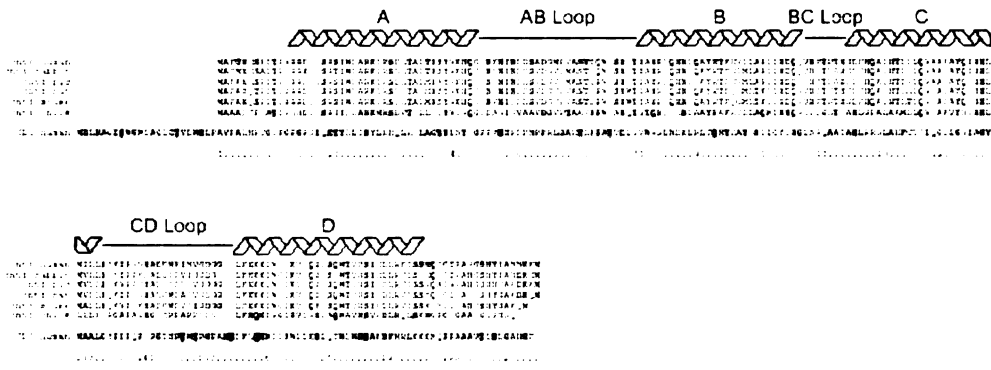
**Figure 1-2**

Structure-based sequence alignment of the known neurotrophins, grouped by subfamily. The secondary structure assignments from PDB structure 1www are at the top (orange: helix, purple arrow: beta strand). The red bars along the top indicate residues seen in the interface with TrkA. The four core beta strands are designated by letter, and the four loops are indicated as well. Sequence numbering is based on NGF at the top. Many of the

sequences are only fragments, and they are grouped at the bottom of each subfamily. The families clearly share similarities, especially along regions of secondary structure. Two NGF-like neurotrophins from the fowlpox virus are included at the bottom. The species is indicated next to each neurotrophin. Certain abbreviations include: prana is *Praomys natalensis* (African soft-furred rat), cavpo is *Cavia porcellus* (domestic guinea pig), xenla is *Xenopus laevis* (African clawed frog), bunmu is *Bungarus multicinctus* (many banded krait), dabrr is *Daboia russelli russelli* (Russell's viper), xipma is *Xiphophorus maculatus* (southern platypus), cypca is *Cyprinus carpio* (common carp), macmu is *Macaca mulatta* (Rhesus monkey), prolo is *Procyon lotor* (raccoon), petbr is *Petaurus breviceps* (sugar glider), rajcl is *Raja clavata* (thornback ray), and viple is *Macrovipera lebetina* (Levantine viper).

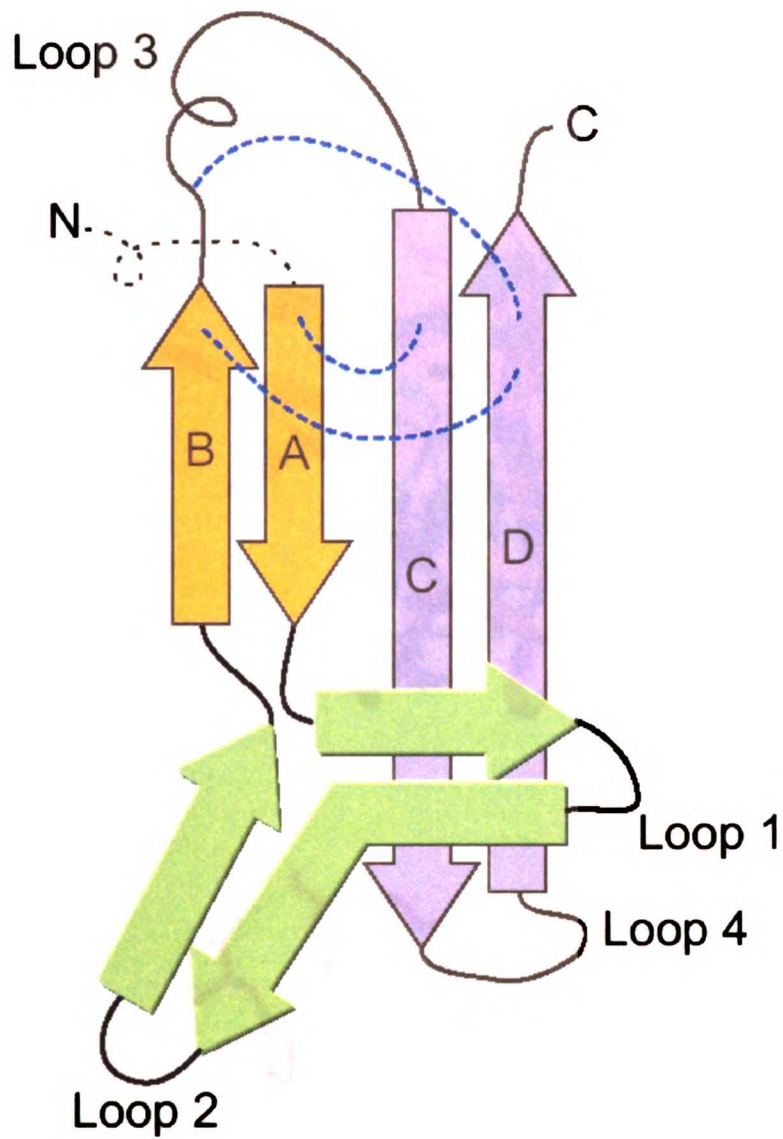






**Figure 1-4**

Structure-based sequence alignment of the known CNTFs, including the newly identified cardiotrophin-1-like cytokine (CLC).



**Figure 1-5**

Schematic of the neurotrophin molecule. Dashed blue lines represent the three disulphide bonds of the cystine knot. The N-terminus is disordered in the unbound structures and is shown by a dashed line.

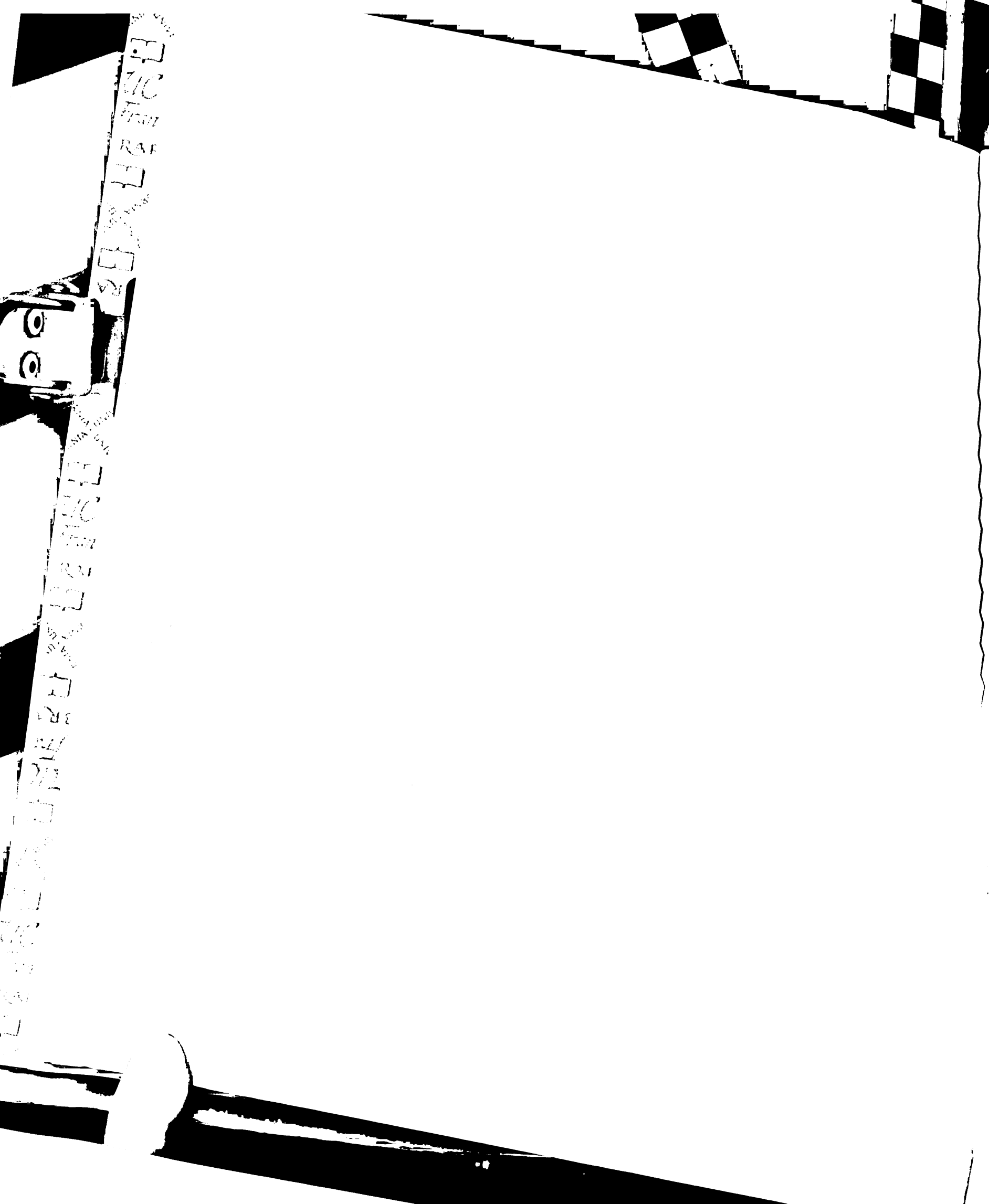


intimate spatial relationship with key residues in TrkA domain 5 (in red) and NGF (in blue). TrkA residue Phe 327 is conserved as Phe, Tyr, or Ile in TrkB and TrkC; Asn 349 is either Asn or Gln; and Gln 350 is either Asn or Lys. TrkA residue His 84 is conserved as Gln in BDNF, NT-3, and NT-4; Phe 86 is either Phe or Tyr; and Thr 29, which forms a mainchain hydrogen bond, is either Val or Ile with similar sized side chains.



**Figure 1-7**

The structure of GDNF, a cystine knot growth factor dimer arranged in head-to-toe manner. Residues from each protomer are highlighted in red and green if they affected binding to the GFR $\alpha$ 1 receptor upon mutagenesis.



TIC

RAF

VY

1000

TIC

RAF

VY

1000

TIC

RAF

VY

1000

TIC

RAF

VY

1000

TIC

RAF

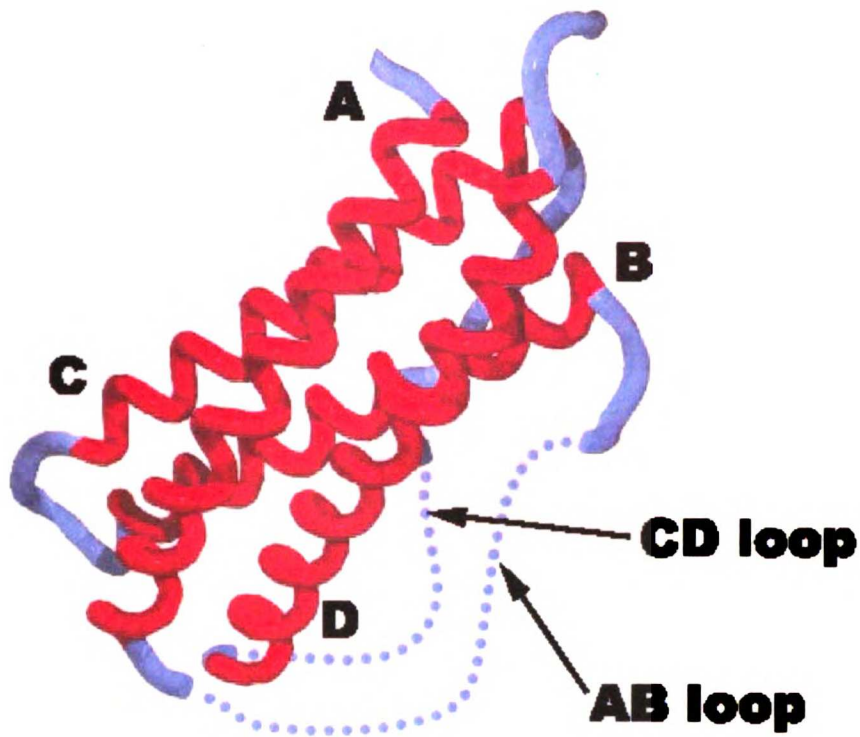
VY

1000

TIC

RAF

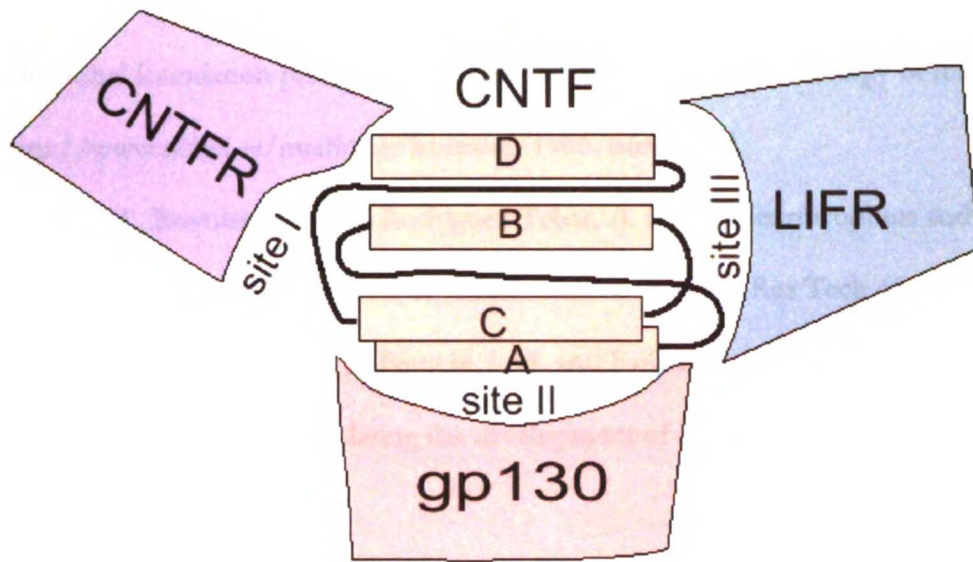
VY



**Figure 1-8**

The structure of CNTF, a four-helix bundle protein similar to the IL-6 family of hematopoietic cytokines. Letters are positioned near the N-terminus of each helix. Two long loops AB and CD were not resolved in the structure and are represented by dashed lines here.





**Figure 1-9**

Schematic showing one possible binding model of CNTF with its receptors CNTFR $\alpha$ , gp130, and LIFR in a 1:1:1:1 stoichiometry.

UCSF LIBRARY

## References

- 1 The nobel foundation press release: The 1986 nobel prize in physiology or medicine, <http://www.nobel.se/medicine/laureates/1986/press.html>
- 2 Frade, J.M., Bovolenta, P. and Rodriguez-Tebar, A. (1999) Neurotrophins and other growth factors in the generation of retinal neurons. *Microsc Res Tech* 45, 243-51.
- 3 Fritsch, B., Silos-Santiago, I., Bianchi, L.M. and Farinas, I. (1997) The role of neurotrophic factors in regulating the development of inner ear innervation. *Trends Neurosci* 20, 159-64.
- 4 Srivastava, D. and Olson, E.N. (1996) Neurotrophin-3 knocks heart off trk. *Nat Med* 2, 1069-71.
- 5 Tessarollo, L., Tsoulfas, P., Donovan, M.J., Palko, M.E., Blair-Flynn, J., Hempstead, B.L. et al. (1997) Targeted deletion of all isoforms of the trkc gene suggests the use of alternate receptors by its ligand neurotrophin-3 in neuronal development and implicates trkc in normal cardiogenesis. *Proc Natl Acad Sci U S A* 94, 14776-81.
- 6 Mufson, E.J., Kroin, J.S., Sendera, T.J. and Sobreviela, T. (1999) Distribution and retrograde transport of trophic factors in the central nervous system: Functional implications for the treatment of neurodegenerative diseases. *Prog Neurobiol* 57, 451-84.
- 7 McMahon, S.B. and Priestley, J.V. (1995) Peripheral neuropathies and neurotrophic factors: Animal models and clinical perspectives. *Curr Opin Neurobiol* 5, 616-24.
- 8 Apfel, S.C. (1999) Neurotrophic factors in the therapy of diabetic neuropathy. *Am J Med* 107, 34S-42S.
- 9 Nakagawara, A. and Brodeur, G.M. (1997) Role of neurotrophins and their receptors in human neuroblastomas: A primary culture study. *Eur J Cancer* 33, 2050-3.

- 10 Eggert, A., Ikegaki, N., Liu, X.G. and Brodeur, G.M. (2000) Prognostic and biological role of neurotrophin-receptor trka and trkb in neuroblastoma. *Klin Padiatr* 212, 200-5.
- 11 Brodeur, G.M. (1994) Molecular pathology of human neuroblastomas. *Semin Diagn Pathol* 11, 118-25.
- 12 Mahadeo, D., Kaplan, L., Chao, M.V. and Hempstead, B.L. (1994) High affinity nerve growth factor binding displays a faster rate of association than p140trk binding. Implications for multi-subunit polypeptide receptors. *J Biol Chem* 269, 6884-91.
- 13 Lin, L.F., Doherty, D.H., Lile, J.D., Bektesh, S. and Collins, F. (1993) Gdnf: A glial cell line-derived neurotrophic factor for midbrain dopaminergic neurons [see comments]. *Science* 260, 1130-2.
- 14 Milbrandt, J., de Sauvage, F.J., Fahrner, T.J., Baloh, R.H., Leitner, M.L., Tansey, M.G. et al. (1998) Persephin, a novel neurotrophic factor related to gdnf and neurturin. *Neuron* 20, 245-53.
- 15 Kotzbauer, P.T., Lampe, P.A., Heuckeroth, R.O., Golden, J.P., Creedon, D.J., Johnson, E.M., Jr. et al. (1996) Neurturin, a relative of glial-cell-line-derived neurotrophic factor. *Nature* 384, 467-70.
- 16 Baloh, R.H., Tansey, M.G., Lampe, P.A., Fahrner, T.J., Enomoto, H., Simburger, K.S. et al. (1998) Artemin, a novel member of the gdnf ligand family, supports peripheral and central neurons and signals through the gfralpha3-ret receptor complex. *Neuron* 21, 1291-302.

- 17 Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. and Pachnis, V. (1994) Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor *ret* [see comments]. *Nature* 367, 380-3.
- 18 Treanor, J.J., Goodman, L., de Sauvage, F., Stone, D.M., Poulsen, K.T., Beck, C.D. et al. (1996) Characterization of a multicomponent receptor for *gdnf* [see comments]. *Nature* 382, 80-3.
- 19 Moore, M.W., Klein, R.D., Farinas, I., Sauer, H., Armanini, M., Phillips, H. et al. (1996) Renal and neuronal abnormalities in mice lacking *gdnf*. *Nature* 382, 76-9.
- 20 Worley, D.S., Pisano, J.M., Choi, E.D., Walus, L., Hession, C.A., Cate, R.L. et al. (2000) Developmental regulation of *gdnf* response and receptor expression in the enteric nervous system [in process citation]. *Development* 127, 4383-93.
- 21 Grieco, M., Santoro, M., Berlingieri, M.T., Melillo, R.M., Donghi, R., Bongarzone, I. et al. (1990) *Ptc* is a novel rearranged form of the *ret* proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. *Cell* 60, 557-63.
- 22 Mulligan, L.M., Eng, C., Healey, C.S., Clayton, D., Kwok, J.B., Gardner, E. et al. (1994) Specific mutations of the *ret* proto-oncogene are related to disease phenotype in men 2a and *fmtc*. *Nat Genet* 6, 70-4.
- 23 Romeo, G., Ronchetto, P., Luo, Y., Barone, V., Seri, M., Ceccherini, I. et al. (1994) Point mutations affecting the tyrosine kinase domain of the *ret* proto-oncogene in hirschsprung's disease [see comments]. *Nature* 367, 377-8.
- 24 Chiariello, M., Visconti, R., Carlomagno, F., Melillo, R.M., Bucci, C., de Franciscis, V. et al. (1998) Signalling of the *ret* receptor tyrosine kinase through the c-jun nh2-terminal protein kinases (jnk): Evidence for a divergence of the erk and jnk pathways induced by *ret*. *Oncogene* 16, 2435-45.

- 25 Hayashi, H., Ichihara, M., Iwashita, T., Murakami, H., Shimono, Y., Kawai, K. et al. (2000) Characterization of intracellular signals via tyrosine 1062 in ret activated by glial cell line-derived neurotrophic factor. *Oncogene* 19, 4469-75.
- 26 Baloh, R.H., Tansey, M.G., Johnson, E.M., Jr. and Milbrandt, J. (2000) Functional mapping of receptor specificity domains of glial cell line- derived neurotrophic factor (gdnf) family ligands and production of gfralpha1 ret-specific agonists. *J Biol Chem* 275, 3412-20.
- 27 Saarma, M. (2000) Gdnf - a stranger in the tgf-beta superfamily? *Eur J Biochem* 267, 6968-6971.
- 28 Saarma, M. and Sariola, H. (1999) Other neurotrophic factors: Glial cell line-derived neurotrophic factor (gdnf). *Microsc Res Tech* 45, 292-302.
- 29 Adler, R., Landa, K.B., Manthorpe, M. and Varon, S. (1979) Cholinergic neuronotrophic factors: Intraocular distribution of trophic activity for ciliary neurons. *Science* 204, 1434-6.
- 30 Manthorpe, M., Skaper, S., Adler, R., Landa, K. and Varon, S. (1980) Cholinergic neuronotrophic factors: Fractionation properties of an extract from selected chick embryonic eye tissues. *J Neurochem* 34, 69-75.
- 31 Sleeman, M.W., Anderson, K.D., Lambert, P.D., Yancopoulos, G.D. and Wiegand, S.J. (2000) The ciliary neurotrophic factor and its receptor, cntfr alpha. *Pharm Acta Helv* 74, 265-72.
- 32 Ip, N.Y. and Yancopoulos, G.D. (1996) The neurotrophins and cntf: Two families of collaborative neurotrophic factors. *Annu Rev Neurosci* 19, 491-515.
- 33 Afonso, C.L., Tulman, E.R., Lu, Z., Zsak, L., Kutish, G.F. and Rock, D.L. (2000) The genome of fowlpox virus. *J Virol* 74, 3815-31.

- 34 Ibanez, C.F. (1995) Neurotrophic factors: From structure-function studies to designing effective therapeutics [published erratum appears in trends biotechnol 1995 aug;13(8):310]. Trends Biotechnol 13, 217-27.
- 35 Li, X., Franz, J., Lottspeich, F. and Gotz, R. (1997) Recombinant fish neurotrophin-6 is a heparin-binding glycoprotein: Implications for a role in axonal guidance. Biochem J 324, 461-6.
- 36 Wiesmann, C., Ultsch, M.H., Bass, S.H. and de Vos, A.M. (1999) Crystal structure of nerve growth factor in complex with the ligand-binding domain of the trka receptor. Nature 401, 184-8.
- 37 Hallbook, F., Lundin, L.G. and Kullander, K. (1998) Lampetra fluviatilis neurotrophin homolog, descendant of a neurotrophin ancestor, discloses the early molecular evolution of neurotrophins in the vertebrate subphylum. J Neurosci 18, 8700-11.
- 38 Barde, Y.A. (1994) Neurotrophic factors: An evolutionary perspective. J Neurobiol 25, 1329-33.
- 39 Ebendal, T. (1992) Function and evolution in the ngf family and its receptors. J Neurosci Res 32, 461-70.
- 40 Gotz, R., Raulf, F. and Schartl, M. (1992) Brain-derived neurotrophic factor is more highly conserved in structure and function than nerve growth factor during vertebrate evolution. J Neurochem 59, 432-42.
- 41 Hallbook, F., Ibanez, C.F. and Persson, H. (1991) Evolutionary studies of the nerve growth factor family reveal a novel member abundantly expressed in xenopus ovary. Neuron 6, 845-58.

- 42 Kullander, K., Carlson, B. and Hallbook, F. (1997) Molecular phylogeny and evolution of the neurotrophins from monotremes and marsupials. *J Mol Evol* 45, 311-21.
- 43 Hallbook, F. (1999) Evolution of the vertebrate neurotrophin and *trk* receptor gene families. *Curr Opin Neurobiol* 9, 616-21.
- 44 Goh, C.S., Bogan, A.A., Joachimiak, M., Walther, D. and Cohen, F.E. (2000) Co-evolution of proteins with their interaction partners. *J Mol Biol* 299, 283-93.
- 45 Elson, G.C., Lelievre, E., Guillet, C., Chevalier, S., Plun-Favreau, H., Froger, J. et al. (2000) Clf associates with clc to form a functional heteromeric ligand for the *cntf* receptor complex. *Nat Neurosci* 3, 867-72.
- 46 Finn, T. and Nishi, R. (1996) Does ciliary neurotrophic factor serve a different function in the rat versus the chicken? *Perspect Dev Neurobiol* 4, 91-9.
- 47 Ryden, M. and Ibanez, C.F. (1997) A second determinant of binding to the p75 neurotrophin receptor revealed by alanine-scanning mutagenesis of a conserved loop in nerve growth factor. *J Biol Chem* 272, 33085-91.
- 48 Urfer, R., Tsoulfas, P., Soppet, D., Escandon, E., Parada, L.F. and Presta, L.G. (1994) The binding epitopes of neurotrophin-3 to its receptors *trkc* and *gp75* and the design of a multifunctional human neurotrophin. *Embo J* 13, 5896-909.
- 49 Butte, M.J., Hwang, P.K., Mobley, W.C. and Fletterick, R.J. (1998) Crystal structure of neurotrophin-3 homodimer shows distinct regions are used to bind its receptors. *Biochemistry* 37, 16846-52.
- 50 Kruttgen, A., Heymach, J.V., Jr., Kahle, P.J. and Shooter, E.M. (1997) The role of the nerve growth factor carboxyl terminus in receptor binding and conformational stability. *J Biol Chem* 272, 29222-8.

- 51 Ilag, L.L., Lonnerberg, P., Persson, H. and Ibanez, C.F. (1994) Role of variable beta-hairpin loop in determining biological specificities in neurotrophin family. *J Biol Chem* 269, 19941-6.
- 52 Heymach, J.V., Jr., Kruttgen, A., Suter, U. and Shooter, E.M. (1996) The regulated secretion and vectorial targeting of neurotrophins in neuroendocrine and epithelial cells. *J Biol Chem* 271, 25430-7.
- 53 Bogan, A.A. and Thorn, K.S. (1998) Anatomy of hot spots in protein interfaces. *J Mol Biol* 280, 1-9.
- 54 Ilag, L.L., Curtis, R., Glass, D., Funakoshi, H., Tobkes, N.J., Ryan, T.E. et al. (1995) Pan-neurotrophin 1: A genetically engineered neurotrophic factor displaying multiple specificities in peripheral neurons in vitro and in vivo. *Proc Natl Acad Sci U S A* 92, 607-11.
- 55 Ibanez, C.F., Ilag, L.L., Murray-Rust, J. and Persson, H. (1993) An extended surface of binding to trk tyrosine kinase receptors in ngf and bdnf allows the engineering of a multifunctional pan-neurotrophin. *Embo J* 12, 2281-93.
- 56 Urfer, R., Tsoulfas, P., O'Connell, L. and Presta, L.G. (1997) Specificity determinants in neurotrophin-3 and design of nerve growth factor-based trkc agonists by changing central beta-strand bundle residues to their neurotrophin-3 analogs. *Biochemistry* 36, 4775-81.
- 57 Kullander, K., Kylberg, A. and Ebendal, T. (1997) Specificity of neurotrophin-3 determined by loss-of-function mutagenesis. *J Neurosci Res* 50, 496-503.
- 58 Dechant, G. and Barde, Y.A. (1997) Signalling through the neurotrophin receptor p75<sup>ntr</sup>. *Curr Opin Neurobiol* 7, 413-8.



- 59 Dechant, G., Biffo, S., Okazawa, H., Kolbeck, R., Pottgiesser, J. and Barde, Y.A. (1993) Expression and binding characteristics of the bdnf receptor chick trkb. *Development* 119, 545-58.
- 60 Chao, M.V. (1994) The p75 neurotrophin receptor. *J Neurobiol* 25, 1373-85.
- 61 Bibel, M., Hoppe, E. and Barde, Y.A. (1999) Biochemical and functional interactions between the neurotrophin receptors trk and p75ntr. *Embo J* 18, 616-22.
- 62 Gargano, N., Levi, A. and Alema, S. (1997) Modulation of nerve growth factor internalization by direct interaction between p75 and trka receptors. *J Neurosci Res* 50, 1-12.
- 63 Ross, G.M., Shamovsky, I.L., Lawrance, G., Solc, M., Dostaler, S.M., Weaver, D.F. et al. (1998) Reciprocal modulation of trka and p75ntr affinity states is mediated by direct receptor interactions. *Eur J Neurosci* 10, 890-8.
- 64 Shamovsky, I.L., Ross, G.M., Riopelle, R.J. and Weaver, D.F. (1999) The interaction of neurotrophins with the p75ntr common neurotrophin receptor: A comprehensive molecular modeling study [published erratum appears in *protein sci* 2000 mar;9(3):623]. *Protein Sci* 8, 2223-33.
- 65 Ibanez, C.F., Ebendal, T., Barbany, G., Murray-Rust, J., Blundell, T.L. and Persson, H. (1992) Disruption of the low affinity receptor-binding site in ngf allows neuronal survival and differentiation by binding to the trk gene product. *Cell* 69, 329-41.
- 66 Ryden, M., Murray-Rust, J., Glass, D., Ilag, L.L., Trupp, M., Yancopoulos, G.D. et al. (1995) Functional analysis of mutant neurotrophins deficient in low-affinity binding reveals a role for p75lngfr in nt-4 signalling. *Embo J* 14, 1979-90.

- 67 Ryden, M. and Ibanez, C.F. (1996) Binding of neurotrophin-3 to p75<sup>l</sup>ngfr, trka, and trkb mediated by a single functional epitope distinct from that recognized by trkc. *J Biol Chem* 271, 5623-7.
- 68 Urfer, R., Tsoulfas, P., Oconnell, L., Hongo, J.A., Zhao, W. and Presta, L.G. (1998) High resolution mapping of the binding site of trka for nerve growth factor and trkc for neurotrophin-3 on the second immunoglobulin-like domain of the trk receptors. *J Biol Chem* 273, 5829-5840.
- 69 Estenne-Bouhtou, G., Kullander, K., Karlsson, M., Ebendal, T., Hacksell, U. and Luthman, K. (1996) Design, synthesis, tandem mass spectrometric sequencing and biological activity of ngf mimetics. *Int J Pept Protein Res* 48, 337-46.
- 70 Longo, F.M., Vu, T.K. and Mobley, W.C. (1990) The in vitro biological effect of nerve growth factor is inhibited by synthetic peptides. *Cell Regul* 1, 189-95.
- 71 Maliartchouk, S., Feng, Y., Ivanisevic, L., Debeir, T., Cuello, A.C., Burgess, K. et al. (2000) A designed peptidomimetic agonistic ligand of trka nerve growth factor receptors. *Mol Pharmacol* 57, 385-91.
- 72 Owolabi, J.B., Rizkalla, G., Tehim, A., Ross, G.M., Riopelle, R.J., Kamboj, R. et al. (1999) Characterization of antiallodynic actions of ale-0540, a novel nerve growth factor receptor antagonist, in the rat. *J Pharmacol Exp Ther* 289, 1271-6.
- 73 Eketjall, S., Fainzilber, M., Murray-Rust, J. and Ibanez, C.F. (1999) Distinct structural elements in gdnf mediate binding to gfralpha1 and activation of the gfralpha1-c-ret receptor complex. *Embo J* 18, 5901-10.
- 74 Chen, Z.Y., He, Z.Y., He, C., Lu, C.L. and Wu, X.F. (2000) Human glial cell-line-derived neurotrophic factor: A structure-function analysis. *Biochem Biophys Res Commun* 268, 692-6.

- 75 Bordeaux, M.C., Forcet, C., Granger, L., Corset, V., Bidaud, C., Billaud, M. et al. (2000) The ret proto-oncogene induces apoptosis: A novel mechanism for hirschsprung disease. *Embo J* 19, 4056-63.
- 76 Trupp, M., Scott, R., Whittemore, S.R. and Ibanez, C.F. (1999) Ret-dependent and -independent mechanisms of glial cell line-derived neurotrophic factor signaling in neuronal cells. *J Biol Chem* 274, 20885-94.
- 77 Hooper, N.M. (1999) Detergent-insoluble glycosphingolipid/cholesterol-rich membrane domains, lipid rafts and caveolae (review). *Mol Membr Biol* 16, 145-56.
- 78 Koshelnick, Y., Ehart, M., Hufnagl, P., Heinrich, P.C. and Binder, B.R. (1997) Urokinase receptor is associated with the components of the jak1/stat1 signaling pathway and leads to activation of this pathway upon receptor clustering in the human kidney epithelial tumor cell line tcl-598. *J Biol Chem* 272, 28563-7.
- 79 Davy, A., Gale, N.W., Murray, E.W., Klinghoffer, R.A., Soriano, P., Feuerstein, C. et al. (1999) Compartmentalized signaling by gpi-anchored ephrin-a5 requires the fyn tyrosine kinase to regulate cellular adhesion. *Genes Dev* 13, 3125-35.
- 80 McDonald, N.Q., Panayotatos, N. and Hendrickson, W.A. (1995) Crystal structure of dimeric human ciliary neurotrophic factor determined by mad phasing. *Embo J* 14, 2689-99.
- 81 Paonessa, G., Graziani, R., De Serio, A., Savino, R., Ciapponi, L., Lahm, A. et al. (1995) Two distinct and independent sites on il-6 trigger gp 130 dimer formation and signalling. *Embo J* 14, 1942-51.
- 82 De Serio, A., Graziani, R., Laufer, R., Ciliberto, G. and Paonessa, G. (1995) In vitro binding of ciliary neurotrophic factor to its receptors: Evidence for the formation of an il-6-type hexameric complex. *J Mol Biol* 254, 795-800.

- 83 Ward, L.D., Howlett, G.J., Discolo, G., Yasukawa, K., Hammacher, A., Moritz, R.L. et al. (1994) High affinity interleukin-6 receptor is a hexameric complex consisting of two molecules each of interleukin-6, interleukin-6 receptor, and gp- 130. *J Biol Chem* 269, 23286-9.
- 84 Grotzinger, J., Kurapkat, G., Wollmer, A., Kalai, M. and Rose-John, S. (1997) The family of the il-6-type cytokines: Specificity and promiscuity of the receptor complexes. *Proteins* 27, 96-109.
- 85 Kallen, K.J., Grotzinger, J., Lelievre, E., Vollmer, P., Aasland, D., Renne, C. et al. (1999) Receptor recognition sites of cytokines are organized as exchangeable modules. Transfer of the leukemia inhibitory factor receptor-binding site from ciliary neurotrophic factor to interleukin-6. *J Biol Chem* 274, 11859-67.
- 86 Panayotatos, N., Radziejewska, E., Acheson, A., Somogyi, R., Thadani, A., Hendrickson, W.A. et al. (1995) Localization of functional receptor epitopes on the structure of ciliary neurotrophic factor indicates a conserved, function-related epitope topography among helical cytokines. *J Biol Chem* 270, 14007-14.
- 87 Ehlers, M., Grotzinger, J., deHon, F.D., Mullberg, J., Brakenhoff, J.P., Liu, J. et al. (1994) Identification of two novel regions of human il-6 responsible for receptor binding and signal transduction. *J Immunol* 153, 1744-53.
- 88 Kruttgen, A., Grotzinger, J., Kurapkat, G., Weis, J., Simon, R., Thier, M. et al. (1995) Human ciliary neurotrophic factor: A structure-function analysis. *Biochem J* 309, 215-20.
- 89 Leebeek, F.W., Kariya, K., Schwabe, M. and Fowlkes, D.M. (1992) Identification of a receptor binding site in the carboxyl terminus of human interleukin-6. *J Biol Chem* 267, 14832-8.

- 90 Saggio, I., Gloaguen, I., Poiana, G. and Laufer, R. (1995) Cntf variants with increased biological potency and receptor selectivity define a functional site of receptor interaction. *Embo J* 14, 3045-54.
- 91 Toniatti, C., Cabibbo, A., Sporena, E., Salvati, A.L., Cerretani, M., Serafini, S. et al. (1996) Engineering human interleukin-6 to obtain variants with strongly enhanced bioactivity. *Embo J* 15, 2726-37.
- 92 Savino, R., Ciapponi, L., Lahm, A., Demartis, A., Cabibbo, A., Toniatti, C. et al. (1994) Rational design of a receptor super-antagonist of human interleukin-6. *Embo J* 13, 5863-70.
- 93 Hudson, K.R., Vernallis, A.B. and Heath, J.K. (1996) Characterization of the receptor binding sites of human leukemia inhibitory factor and creation of antagonists. *J Biol Chem* 271, 11971-8.
- 94 Vernallis, A.B., Hudson, K.R. and Heath, J.K. (1997) An antagonist for the leukemia inhibitory factor receptor inhibits leukemia inhibitory factor, cardiotrophin-1, ciliary neurotrophic factor, and oncostatin m. *J Biol Chem* 272, 26947-52.
- 95 Layton, M.J., Owczarek, C.M., Metcalf, D., Clark, R.L., Smith, D.K., Treutlein, H.R. et al. (1994) Conversion of the biological specificity of murine to human leukemia inhibitory factor by replacing 6 amino acid residues. *J Biol Chem* 269, 29891-6.
- 96 Owczarek, C.M., Layton, M.J., Metcalf, D., Lock, P., Willson, T.A., Gough, N.M. et al. (1993) Inter-species chimeras of leukaemia inhibitory factor define a major human receptor-binding determinant. *Embo J* 12, 3487-95.
- 97 McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A. and Blundell, T.L. (1991) New protein fold revealed by a 2.3-Å resolution crystal structure of nerve growth factor. *Nature* 354, 411-4.

- 98 Holland, D.R., Cousens, L.S., Meng, W. and Matthews, B.W. (1994) Nerve growth factor in different crystal forms displays structural flexibility and reveals zinc binding sites. *J Mol Biol* 239, 385-400.
- 99 Robinson, R.C., Radziejewski, C., Stuart, D.I. and Jones, E.Y. (1995) Structure of the brain-derived neurotrophic factor/neurotrophin 3 heterodimer. *Biochemistry* 34, 4139-46.
- 100 Robinson, R.C., Radziejewski, C., Spraggon, G., Greenwald, J., Kostura, M.R., Burtnick, L.D. et al. (1999) The structures of the neurotrophin 4 homodimer and the brain-derived neurotrophic factor/neurotrophin 4 heterodimer reveal a common trk-binding site. *Protein Sci* 8, 2589-97.
- 101 Eigenbrot, C. and Gerber, N. (1997) X-ray structure of glial cell-derived neurotrophic factor at 1.9 Å resolution and implications for receptor binding [letter]. *Nat Struct Biol* 4, 435-8.

## Chapter 2. Crystal structure of neurotrophin-3 homodimer shows distinct regions are used to bind its receptors

Manish J. Butte<sup>‡</sup>, Peter K. Hwang<sup>§</sup>,

William C. Mobley<sup>□</sup>, Robert J. Fletterick<sup>\*,§</sup>

<sup>‡</sup> Graduate Group in Biophysics, University of California San Francisco, Box 0448, San Francisco, California 94143, USA, [manish@msg.ucsf.edu](mailto:manish@msg.ucsf.edu)

<sup>§</sup> Dept. of Biochemistry and Biophysics, University of California San Francisco, Box 0448, San Francisco, California 94143, USA

<sup>□</sup> Dept. of Neurology, Stanford University, Palo Alto, California 94304, USA

Running title: Crystal structure of neurotrophin-3

\*Corresponding Author: Robert J. Fletterick, Ph.D.

phone: 415-476-5080

fax: 415-476-1902

email: [flett@msg.ucsf.edu](mailto:flett@msg.ucsf.edu)

†This work is supported by the Office of Naval Research Predoctoral Fellowship (to MJB), the McGowan Charitable Fund (to WCM), and NIH DK32822 (to PKH and RJF). Atomic coordinates have deposited with the Brookhaven Protein Data Bank as 1nt3.

Reproduced with permission from Biochemistry 1998 Dec 1;37(48):16846-52.  
Copyright 1998 American Chemical Society

## Abbreviations and Textual Footnotes

<sup>1</sup>NT-3, neurotrophin-3; BDNF, brain derived neurotrophic factor; NGF, nerve growth factor; PDGF, platelet derived growth factor; TGF-beta, transforming growth factor beta; hCG, human choriogonadotropin; TNFR, tumor necrosis factor receptor.

<sup>2</sup>Analogous NT-3 and NGF residues differ in numbering by one until residue 94, as shown in Figure 2.



## Abstract

Neurotrophin-3 (NT-3)<sup>1</sup> is a cystine knot growth factor that promotes the survival, proliferation, and differentiation of developing neurons and is a potential therapeutic for neurodegenerative diseases. To clarify the structural basis of receptor specificity and the role of neurotrophin dimerization in receptor activation, the structure of the NT-3 homodimer was determined using x-ray crystallography. The orthorhombic crystals diffract to 2.4 Å, with dimer symmetry occurring about a crystallographic two-fold axis. The overall structure of NT-3 resembles that of the other neurotrophins, NGF and BDNF; each protomer forms a twisted four-stranded beta sheet, with three intertwined disulfide bonds. There are notable differences, however, between NT-3 and NGF in the surface loops and in three functionally important regions, shown in previous mutagenesis studies to be critical for binding. One such difference implies that NT-3's binding affinity and specificity depend on a novel hydrogen bond between Gln 83, a residue important for binding specificity with TrkC, and Arg 103, a residue crucial for binding affinity with TrkC. NT-3's extensive dimer interface buries much of the otherwise solvent-accessible hydrophobic surface area and suggests that the dimeric state is stabilized through the formation of this hydrophobic core. A comparison of the dimer interface between the NT-3 homodimer and the BDNF/NT-3 heterodimer reveals similar patterns of hydrogen bonds and nonpolar contacts, which reinforces the notion that the evolutionarily-conserved neurotrophin interface resulted from the need for receptor dimerization in signal initiation.

## Introduction

Neurotrophin-3 (NT-3) is a protein growth factor that belongs to the neurotrophin family, which includes nerve growth factor (NGF), brain derived neurotrophic factor (BDNF), neurotrophin-4 (also called neurotrophin-4/5), and neurotrophin-6. Neurotrophins act on specific developing neuronal populations to prevent or direct programmed cell death, promote differentiation, and regulate proliferation (1). In addition, neurotrophins act in the mature nervous system to maintain neuronal function (2), enhance neuronal function in animal models of neurological diseases (3), and may play a role in neural tumor progression (4). The neurotrophins belong to the cystine knot superfamily, a large group that shares a conserved core of three intertwined disulfide bonds. Other members of this family include platelet-derived growth factor (PDGF), transforming growth factor beta (TGF-beta), and human choriogonadotropin (hCG). All cystine knot growth factors are non-globular beta-sheets, lacking a well-defined hydrophobic core. In addition, they all exist exclusively as dimers, the interface of which is notable for its extensive buried hydrophobic surface. One hypothesis suggests that dimerization is required because the protomers alone are not structurally stable. Only upon dimerization is there a sufficient hydrophobic core to stabilize the structure (5). The NGF crystal structure supports this notion, but whether it applies to NT-3 or the other neurotrophins was uncertain.

At least two NT-3 cell surface receptors have been identified. TrkC, a receptor tyrosine kinase that shows selectivity and nanomolar affinity for NT-3, must be activated for the normal survival of certain peripheral nervous system and sensory neurons (6). p75<sup>NTR</sup>, a member of the tumor necrosis factor receptor (TNFR) superfamily, binds all the neurotrophins with nanomolar affinity and plays a role in neuronal apoptosis (7). Extensive mutagenesis has defined the sequences in NGF (7-10) and NT-3 (11) that mediate binding

affinity and specificity for their receptors. One puzzling conclusion is that each neurotrophin uses wholly different sequences to bind its receptors. Whether common themes may be apparent in the spatial arrangement of these residues is not yet known. Even more oddly, NT-3 is the only neurotrophin that can bind, albeit weakly, to the other Trk receptors, TrkA and TrkB (the cognate receptors for NGF and BDNF, respectively); the biological significance of this has not been established. In contrast to the promiscuity of its cognate ligand, TrkC is very stringent: it binds only NT-3. Are there structural clues to NT-3's binding specificity?

To explore NT-3's structural relationship with the other neurotrophins and cystine knot growth factors and to investigate the structural basis for receptor specificity, we determined the structure of the NT-3 homodimer using x-ray crystallography. The NT-3 protomer from the structure of a BDNF/NT-3 heterodimer (12) was used as a search model for molecular replacement phasing. The structure of the recombinant human NT-3 homodimer was compared to the NGF homodimer and the BDNF/NT-3 heterodimer. We propose a structural model for how NT-3 binds to TrkC and p75<sup>NTR</sup>. Our results also support a simple structural explanation for obligate dimerization of the neurotrophins.

## Experimental Procedures

### **Protein purification and crystal growth.**

Recombinant human NT-3 was expressed in *E. coli* and purified to greater than 95% purity by Amgen, Inc., and donated to us by Regeneron Pharmaceuticals. NT-3 was buffered in Tris HCl pH 8.0, and concentrated to 15 mg/ml. Edman sequencing from the N-terminus to Val 13 and silver-stained SDS PAGE were performed to ensure the protein's identity. Crystals were grown by vapor diffusion from a reservoir solution of ammonium sulfate 200 mM, PEG 400 32%, MES (2-[N-Morpholino]ethanesulfonic acid) 100 mM at pH 5.6, which was optimized from starting conditions appearing in (13) and (14). The crystals appeared as thin plates, and grew to their full size (approximately 0.1 mm in the largest dimension) in about one week at 18°C.

### **Data collection and processing.**

A crystal was captured in a nylon loop and transferred to a 15%-glycerol/85%-reservoir cryoprotectant solution briefly, then rapidly frozen in a liquid nitrogen bath. A 2.4 Å data set (half the reflections have an I/sigma of 3.0 at 2.4 Å resolution) was measured from a single crystal at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 7-1, with average completeness 85.3% to 2.4 Å (89.6% for the 2.5 – 2.4 Å shell). The space group is P2<sub>1</sub>2<sub>1</sub>2, and the unit cell has parameters a = 37.4 Å, b = 52.1 Å, c = 64.8 Å. The intensity data were processed and integrated with DENZO and Scalepack (15). Statistical analysis of the intensities with a Wilson plot (CCP4 program Wilson (16)) revealed a mean isotropic temperature factor of 33.

### **Molecular replacement.**

The structure was solved by molecular replacement, using the NT-3 protomer from the BDNF/NT-3 heterodimer structure (PDB 1bnd) (12) as the search model. This search

model was chosen because it should have 100% identity to our NT-3, except for two inconsistencies in the PDB entry: Val 11 should be a Tyr and Gln 44 should be a Gly according to the Swiss-Prot entry for human NT-3 (accession P20783) and according to the results of our N-terminal sequencing experiment. All 108 residues from the NT-3 protomer of the heterodimer structure were used, including all side chains. A number of candidate solutions were found with the CCP4 program AMoRe (17). The top solutions were visualized in a unit cell using INSIGHT II (version 95.0, Biosym Technologies, San Diego, California), and the second solution was found to have the best appearance, in terms of both packing with its symmetry mate and of global similarity to the search model. The dimer-mate falls across a crystallographic 2-fold symmetry axis. For this solution, the  $R_{\text{cryst}}$  was 41.6% and the correlation coefficient was 59.7.

### **Refinement.**

The molecular replacement solution model underwent a rigid-body minimization and a Cartesian simulated annealing during which a bulk solvent correction was applied. Repeated cycles of refinement using X-PLOR (18) and the developmental CNS (Crystallography and NMR System) program (A.T. Brunger, personal communication), and manual rebuilding using MOLOC (19) followed to ensure that the free R-factor decreased steadily. Refinement included positional refinement, Cartesian and torsional simulated annealing, and B factor refinement with the maximum likelihood target (20). The model was refined against 90% of the measured data between 30.0 and 2.4 Å using the Engh and Huber stereochemical dictionary. The remaining 10% of the data were excluded from all refinement calculations to cross validate the progress of refinement (the free R set).

After twenty cycles of refinement and manual rebuilding, water molecules were added to the structure provided that (1) density appeared  $> 2.0$  sigma on the  $2F_o - F_c$  map,

(2) they formed hydrogen bonds of reasonable geometry, and (3) their inclusion reduced the free R-factor. The refinement continued with 2.0 sigma or better structure factors in the 8.0 – 2.4 Å resolution range to reduce the effects of poorly measured low-resolution data. In later stages, solvent-flipped electron density maps were calculated using the CCP4 program SOLOMON (21). Phase probabilities from SOLOMON were also used as constraints in the MLHL target function (Pannu N.S. *et al.*, in the press: *Acta Cryst D*) implemented in CNS; this reduced the free R-factor further.

Many loop regions of the protomer had only sparsely visualizable density in a simulated annealing omit map, notably at the residues between 42 and 47, between 61 and 66, between 68 and 71, and between 93 and 95. A few residues at the N- and C-termini (8–10 and 115) also show poor electron density because they are inadequately constrained structurally. To improve the electron density maps of these loops, we attempted multi-crystal averaging with a lower-resolution data set (good to 3.2 Å) collected from a similar crystal, using the program MAVE (22). This did not significantly improve the density in the loops, and in fact suggested that the above loops may assume variable conformations. Many of the above loop regions were also noted to be disordered or variable in conformation in the two different NGF structures (23, 24) and in the BDNF/NT-3 heterodimer. Unfortunately, these disordered loop regions make up around 20% of the atoms of the protomer, so the R-factor and free R-factor are somewhat higher than might be expected for a similar sized, globular protein. Fortunately, all the residues that appear to play a biological role in mediating receptor binding were well-ordered, as will be discussed below. The final R-factor was 23.3 and the free R-factor was 26.7, with excellent geometry (for all data, the R-factors were less than 1% higher). A Ramachandran plot reveals 97% of the residues fall into allowed regions, with the only exceptions being residues in the aforementioned loops. The

overall statistics for data collection and refinement are summarized in Table 3. To check for search model bias, a high-temperature electron density omit map was calculated, a representative portion of which is shown in Figure 5. It revealed that the main chain positions were similar between our structure and the search model and that the side chain conformations clearly support our structure versus that of the search model.

### **Analysis.**

The dimer interface statistics in Table 2 and the accessible surface area calculations were computed using the ASC program (25), using a 1.4 Å probe. Molscript (26), Raster3D (27), and GRASP (28) were used to view the final structure and produce the figures. The program ALSCRIPT (29) was used for display of the sequence alignment. Superimposed neurotrophin structures were compared using the LSQMAN program (30).

## Results and Discussion

### Overall Structure.

The NT-3 dimer shares a similar fold with NGF and the BDNF/NT-3 heterodimer (Figures 1a, b, c, and d). Its tertiary structure consists of four anti-parallel beta strands, with the fourth beta strand twisted around the third, a feature conserved across all the cystine knot growth factors (5). The three disulfides that make up the cystine knot are virtually superimposable on those of NGF with a r.m.s. deviation of less than 0.3 Å (Table 1). They follow the pattern established in all cystine knot growth factors: Cys(III – VI) penetrates the ring formed by Cys(I – IV) and Cys(II – V). The hydrogen bonding pattern surrounding the cystine knot is remarkably well conserved, including the beta-bulge noted in hCG, NGF, and PDGF after Cys V (31).

The superpositions of the three dimers, as shown in Table 1, reveal that the overall structural similarity between neurotrophins is very high, about 1 Å in comparing positions of matched alpha carbon atoms. This was to be expected, given the sequence alignment of human NT-3 with the other known human neurotrophins, as shown in Figure 2. Human NT-3 shares 57% amino acid identity with human NGF and 56% with human BDNF. Variable regions, in both sequence and structure, are present at the N- and C-termini, in the loops between the beta strands, and in the fourth beta strand. As might be anticipated from the evolutionary variety of residues in these loops and from looking at other neurotrophin crystal structures, we found a few of these loops to be variable in conformation and poorly ordered in the electron density maps, as discussed above. Other regions, including the critical six cysteines and the spacing between them, appear well conserved. As an aside, we note that human NT-6, which has been poorly characterized thus far, appears to lack one of the three disulfides (Figure 2). Removing a cysteine in Norrie disease protein, a cystine knot growth



factor, destroys its function and causes an X-linked neurological disorder (32); how NT-6 compensates for the similar lack of cystine knot stabilization is unknown.

### **Dimer interface.**

The dimer is stabilized by over two hundred non-bonding interactions that bury 2110 Å<sup>2</sup> of hydrophobic surface area and 1250 Å<sup>2</sup> of polar and charged surface area. Table 2 summarizes these interfacial interactions for all the known neurotrophin structures, and confirms that over a quarter of the NT-3 protomer's surface area is buried in the dimer interface. The amount of non-polar surface area that is solvent inaccessible in the NT-3 protomer is low at 44.2%, which suggests a relative lack of hydrophobic core. Upon dimerization, this rises to 55.7%, which is more typical of globular proteins. The formation of an ample hydrophobic core is thought to help stabilize dimeric neurotrophins and cystine knot growth factor dimers (5).

The residues that lie on the dimer interface are remarkably well conserved among the neurotrophins; this fact has been used to facilitate manufacture of stable neurotrophin heterodimers (33). Heterodimers of NT-3 with NGF were not stable (33); this may be due to the significant differences in the polar and apolar makeup of NGF's and NT-3's dimer interfaces (Table 2). We compared the dimer-interface structures of the NT-3 homodimer and the NT-3 heterodimer with BDNF (12). While the amount of buried surface area in proportion to total is slightly higher in the homodimer (29% versus 25%), the hydrophobic contribution is roughly the same in both. The homodimer and heterodimer have a similar number of side-chain-to-side-chain hydrogen bonds across the interface, though the actual residues involved are different. Our results show a striking overall similarity in the number of hydrogen bonds and in the distribution of buried surface area between the NT-3 homodimer and the BDNF/NT-3 heterodimer, which reinforces the notion that an

evolutionarily-conserved dimer interface probably resulted from the necessity for receptor dimerization in signal transduction.

### **Functional implications.**

To gain insight into the structural determinants of binding to the Trk receptors and p75<sup>NTR</sup>, several investigators have analyzed NT-3 and other neurotrophins using chimeric constructs, alanine scanning mutagenesis, and site-directed mutagenesis (7, 9, 11, 34, 35). These experiments suggest that NT-3 and NGF differ in the regions that confer receptor binding with affinity and specificity. In comparing NT-3 to NGF, residues have been identified that contribute little to binding affinity but much to binding specificity, and vice versa. In the case of NGF's binding to TrkA, a single structural feature most significantly contributes to both binding affinity and specificity, the N-terminal six residues (36). On the other hand, NT-3's residues involved in binding TrkC are not confined to one particular section, but are distributed throughout the sequence. The key structural feature for binding to p75<sup>NTR</sup> appears to be basic residues in region I or in region V, two loops that are spatially close (34) (Figure 2). Positively charged residues in these regions are conserved across the neurotrophins. We discuss below the structural features of the residues mediating binding affinity and binding specificity to TrkC and p75<sup>NTR</sup>.

Urfer *et al.* analyzed surface residues of NT-3 for their contribution to binding affinity using alanine-substitution mutants. Binding to TrkC was measured *in vitro* using binding competition assays with radioactive wild-type neurotrophin and using biological activity assays monitoring neurite growth in cell culture. When mutated to alanine, the residues identified as most significantly affecting binding to TrkC were Arg 103 (beta strand 4); Lys 80 and Gln 83 (beta strand 3); Arg 56 and Glu 54 (beta strand 2); and Thr 22 (beta strand 1) (11). Since our model represents an NT-3 homodimer, we can identify binding

regions that involve both protomers, a task that could not be accomplished using the BDNF/NT-3 heterodimer structure, or on an individual protomer. The residues mentioned above can be seen in our NT-3 model as creating a discrete binding face (Figure 3). The data of Urfer *et al.* also help define a clearly marked boundary of residues that do not contribute to TrkC binding.

Discerning an analogous binding site on NT-3 for p75<sup>NTR</sup> is more difficult. Biological assays for p75<sup>NTR</sup> activity have not been performed; thus the only evidence for the interaction site is based on *in vitro* binding studies using radioactive wild-type neurotrophin competition. Nevertheless, residues were found (11) that form two discontinuous binding regions (Figure 3). One binding region involves residues Arg 31 and His 33, both in variable region I. More interestingly, a second region crosses both protomers; it is comprised of Tyr 11 and Arg 68 from one protomer and Lys 73 and Arg 114 from the other. This result suggests that the NT-3 dimer may be needed for p75<sup>NTR</sup> activation. All the residues implicated in binding p75<sup>NTR</sup> lie in mobile regions, either near the N- or C-termini or on one of the flexible loops. Strikingly, there is no overlap between the presumed binding regions for TrkC and p75<sup>NTR</sup>. This is consistent with the curious observation that the presence of p75<sup>NTR</sup> improves neurotrophin binding to a Trk receptor (37), possibly by p75<sup>NTR</sup> initially binding the neurotrophin (38).

To identify the NT-3 residues most responsible for TrkC specificity, variants of NGF, formed by replacing four key residues with their NT-3 counterparts, were constructed that bind TrkC well and yet do not sacrifice TrkA binding (35). Two distinct regions on the beta sheet contain these critical specificity-lending residues. The first region lies near Thr 22 (Figure 4), the residue which Urfer *et al.* show is the most important for TrkC specificity. Substituting Thr 22 in place of the corresponding residue, a glycine, in NGF, produces a

240-fold improvement in TrkC binding<sup>2</sup>. The addition of the hydroxyl moiety in threonine may allow a crucial hydrogen bonding contact with the receptor. As was predicted by Urfer *et al.*, one might expect that changing the Gly to Thr would significantly affect the backbone conformation, since glycine can sustain greater flexibility. Surprisingly, this is not borne out in our NT-3 model, the backbone of which is exactly superimposable on NGF's in this region (Figure 4). This region also contains another important specificity-lending residue, Leu 19, which replaces Val 20 in NGF. The specificity here may be mediated by tight constraints on the side chain size or branching.

The second cluster of critical residues lies near Arg 103. Arg 103 is the most critical residue for binding affinity to TrkC; the corresponding residue in NGF, also an arginine, was shown to be insignificant for binding affinity to TrkA (39). Our results suggest that stabilization of this arginine in a unique conformation may represent a key feature of TrkC specificity. Significantly, the mutagenesis experiments of Urfer *et al.* show that Gln 83, which replaces His 84 in NGF, improves TrkC specificity. In our NT-3 model, the carboxyl group of the Gln 83 side chain accepts a hydrogen bond from the side chain of Arg 103 (Figure 4); this pulls the Arg side chain into a dramatically different position than is seen in NGF, which lacks such stabilization. Finally, Tyr 85 in NT-3 is important for TrkC specificity when it replaces Phe 86 in NGF. Tyr 85 in our model is directly adjacent to Arg 103, again highlighting the importance of this region for binding specificity. Tyr 85 is also the only of these four residues that lies in a hydrophobic cluster in the dimer interface. Mutation of Phe 86/His 84 to Tyr 85/Gln 83 in NGF may direct the molecule to a conformation more suitable for TrkC binding because of reorientation of Arg 103 and other side-chains in this region.

In summary, we determined the x-ray crystal structure of the NT-3 homodimer. Clues to the specificity of NT-3 for its TrkC receptor were found when comparing the structure to NGF at functionally important residues. These clues suggest that both rigid and flexible regions are important for binding affinity and specificity. Rigid regions, such as the critical structures in the area of Arg 103 and the constant backbone structure near Thr 22 (Figure 4), likely represent scaffold areas that mediate binding to the TrkC receptor. Flexible loop regions, on the other hand, make up the binding site for p75<sup>NTR</sup> (Figure 3), which binds all the neurotrophins equally. TrkC specificity may depend on a hydrogen bond between Arg 103, the residue most crucial for binding affinity, and Gln 83, a residue very important for binding specificity. In addition, our results reveal that the dimer interface of the NT-3 homodimer is extensive and non-polar, which may explain why dimerization is required for this molecule to exist. The similarity of interfacial hydrophobic contacts and hydrogen bonding patterns between the NT-3 homodimer and those of the BDNF/NT-3 heterodimer helps substantiate the view that this well-conserved dimer interface resulted from the requirement for receptor dimerization in signal transduction. Finally, combined with the recent Trk receptor mutagenesis results of Urfer *et al.* (40), crystallization of the Trk receptors (Butte M.J. *et al.*, unpublished) with the neurotrophins promises to further define and confirm the interactions indicated.

## Acknowledgment

We thank George D. Yancopoulos of Regeneron Pharmaceuticals, Inc., Tarrytown, New York, USA, for generously providing the neurotrophin-3. MJB acknowledges Andy Shiau and the UCSF Macromolecular Structure Group for helpful discussions.

## References

1. Kaplan, D. R., and Miller, F. D. (1997) *Curr Opin Cell Biol* 9, 213-21.
2. Levi-Montalcini, R., Skaper, S. D., Dal Toso, R., Petrelli, L., and Leon, A. (1996) *Trends Neurosci* 19, 514-20.
3. Yuen, E. C., and Mobley, W. C. (1996) *Ann Neurol* 40, 346-54.
4. Ryden, M., Sehgal, R., Dominici, C., Schilling, F. H., Ibanez, C. F., and Kogner, P. (1996) *Br J Cancer* 74, 773-9.
5. Sun, P. D., and Davies, D. R. (1995) *Annu Rev Biophys Biomol Struct* 24, 269-91.
6. White, F. A., Silos-Santiago, I., Molliver, D. C., Nishimura, M., Phillips, H., Barbacid, M., and Snider, W. D. (1996) *J Neurosci* 16, 4662-72.
7. Ilag, L. L., Lonnerberg, P., Persson, H., and Ibanez, C. F. (1994) *J Biol Chem* 269, 19941-6.
8. Persson, H., and Ibanez, C. F. (1993) *Curr Opin Neurol Neurosurg* 6, 11-8.
9. Ryden, M., and Ibanez, C. F. (1996) *J Biol Chem* 271, 5623-7.
10. Ibanez, C. F., Ilag, L. L., Murray-Rust, J., and Persson, H. (1993) *Embo J* 12, 2281-93.
11. Urfer, R., Tsoulfas, P., Soppet, D., Escandon, E., Parada, L. F., and Presta, L. G. (1994) *Embo J* 13, 5896-909.
12. Robinson, R. C., Radziejewski, C., Stuart, D. I., and Jones, E. Y. (1995) *Biochemistry* 34, 4139-46.
13. Robinson, R. C., Radziejewski, C., Stuart, D. I., and Jones, E. Y. (1996) *Protein Sci* 5, 973-7.
14. Kelly, J. A., Singer, E., Osslund, T. D., and Yeates, T. O. (1994) *Protein Sci* 3, 982-3.

15. Otwinowski, Z., and Minor, W. (1996) in *Methods in Enzymology Vol. 276, Macromolecular Crystallography* (Carter, C. W., Jr., and Sweet, R. M., Eds.) pp 307-326, Academic Press.
16. Bailey, S. (1994) *Acta Crystallogr D-Biol Cryst* 50, 760-763.
17. Navaza, J. (1994) *Acta Crystallogr. A, Found. Crystallogr. (Denmark)* A50, 157-63.
18. Brunger, A. T. (1988) *J Mol Biol* 203, 803-16.
19. Muller, K., Amman, H. J., Doran, D. M., Gerber, P. R., Gubernator, K., and Schrepfer, G. (1988) *Bull. Soc. Chim. Belg.* 97, 655-67.
20. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997) *Proc Natl Acad Sci U S A* 94, 5018-23.
21. Abrahams, J. P., and Leslie, A. G. W. (1996) *Acta Crystallogr D-Biol Cryst* 52, 30-42.
22. Jones, T. A. (1992) in *Molecular Replacement* (Dodson, E. J., Gover, S., and Wolf, W., Eds.) pp pp. 91-105, SERC Daresbury Laboratory, Warrington.
23. McDonald, N. Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A., and Blundell, T. L. (1991) *Nature* 354, 411-4.
24. Holland, D. R., Cousens, L. S., Meng, W., and Matthews, B. W. (1994) *J Mol Biol* 239, 385-400.
25. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C., and Scharf, M. (1995) *J. Comput. Chem. (USA)* 16, 273-84.
26. Kraulis, P. J. (1991) *J Appl Cryst* 24, 946-950.
27. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr D-Biol Cryst* 50, 869-873.
28. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* 11, 281-96.
29. Barton, G. J. (1993) *Protein Eng* 6, 37-40.
30. Kleywegt, G. J., and Jones, T. A. (1994) *ESF/CCP4 Newsletter* 31, 9-14.



31. McDonald, N. Q., and Hendrickson, W. A. (1993) *Cell* 73, 421-4.
32. Strasberg, P., Liede, H. A., Stein, T., Warren, I., Sutherland, J., and Ray, P. N. (1995) *Hum Mol Genet* 4, 2179-80.
33. Radziejewski, C., and Robinson, R. C. (1993) *Biochemistry* 32, 13350-6.
34. Ryden, M., Murray-Rust, J., Glass, D., Ilag, L. L., Trupp, M., Yancopoulos, G. D., McDonald, N. Q., and Ibanez, C. F. (1995) *Embo J* 14, 1979-90.
35. Urfer, R., Tsoulfas, P., O'Connell, L., and Presta, L. G. (1997) *Biochemistry* 36, 4775-81.
36. Shih, A., Laramee, G. R., Schmelzer, C. H., Burton, L. E., and Winslow, J. W. (1994) *J Biol Chem* 269, 27679-86.
37. Mahadeo, D., Kaplan, L., Chao, M. V., and Hempstead, B. L. (1994) *J Biol Chem* 269, 6884-91.
38. Bothwell, M. (1995) *Annu Rev Neurosci* 18, 223-53.
39. Guo, M., Meyer, S. L., Kaur, H., Gao, J. J., and Neet, K. E. (1996) *Protein Sci* 5, 447-55.
40. Urfer, R., Tsoulfas, P., Oconnell, L., Hongo, J. A., Zhao, W., and Presta, L. G. (1998) *J Biol Chem* 273, 5829-5840.
41. McDonald, I. K., Naylor, D. N., Jones, D. T., and Thornton, J. M. (1993) ,  
Department of Biochemistry and Molecular Biology, University College, London.

Table 2-1

Root mean square differences (Å) in C-alpha superpositions<sup>a</sup>

Reference	Target for superposition		
	NT-3	NT-3/BDNF	NGF
NT-3	•	1.02 (198)	1.20 (200)
NT-3/BDNF	0.30	•	0.95 (202)
NGF	0.31	0.20	•

<sup>a</sup>Values above the diagonal are root mean square differences comparing the superpositions of alpha carbons (number of carbons used in the alignment is in parentheses) of the various neurotrophin dimers. Values below the diagonal represent the same for the six cysteine residues in each of those structures.

Table 2-2

Interactions that constitute the neurotrophin dimers<sup>a</sup>

Interaction	NT-3	NGF	NT-3 / BDNF
Hydrogen bonds	10	7	8
van der Waals	200	100	140
Contact surface area (Å <sup>2</sup> )			
Polar	1250	740	1030
Apolar	2110	1800	2060
Total	3360	2540	3090
Total surface area (Å <sup>2</sup> )	11490	11990	12360

<sup>a</sup>Hydrogen bonds were determined by the program HBPLUS (41); van der Waals

interactions are contacts closer than 4.0 Å.

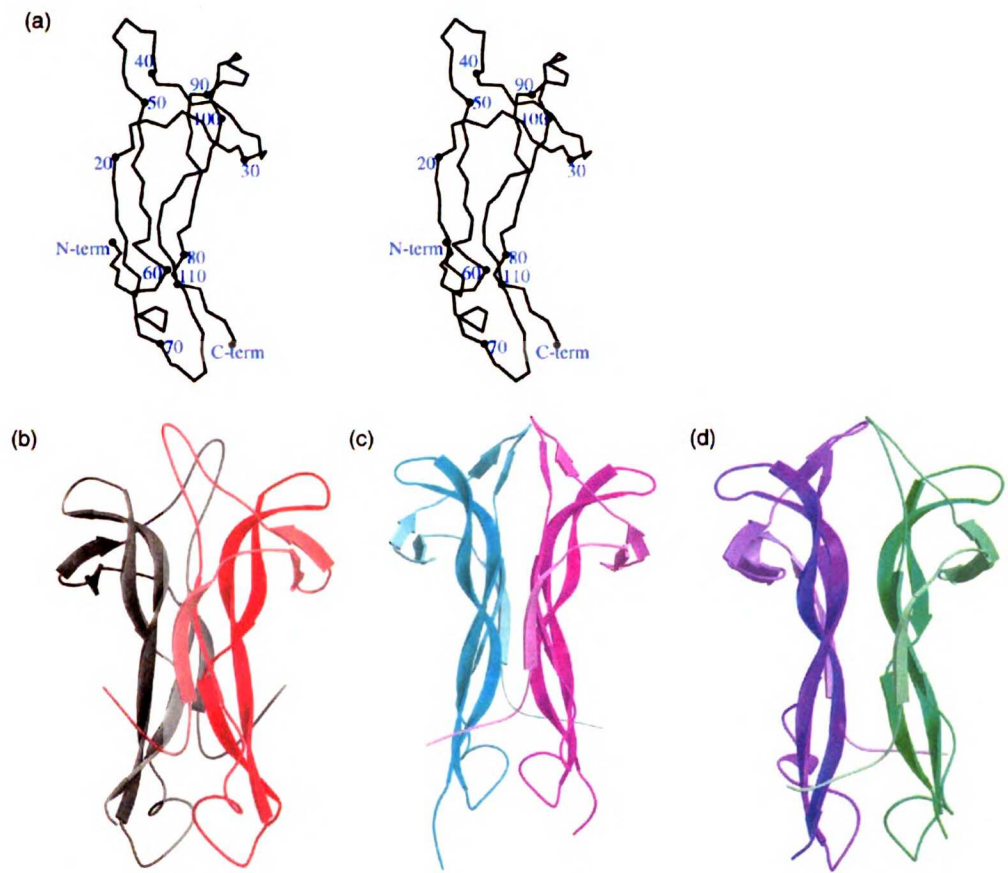
Table 2-3

**Crystal parameters, data collection and refinement statistics**

Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2
Protein number of residues / atoms	108 / 838
Total reflections / Unique reflections	20827 / 4535
Average coverage of data from scaling	
all data	30 – 2.4 Å : 85.3%
highest resolution shell	2.5 – 2.4 Å : 89.6%
Average redundancy of observations	4.6
R sym (%)	5.5
Molecular replacement	
R cryst	41.6
correlation coefficient	59.7
<b>Refinement</b>	
Resolution range (Å) / sigma cutoff / reflections used	30 – 2.4 / F > 0s / 4506 8.0 – 2.4 / F > 2s / 4263
Ordered waters seen	63
Free R-factor (%) <sup>a</sup>	26.7
R-factor (%) <sup>b</sup>	23.3
Bond r.m.s. deviation (Å)	0.007
Angle r.m.s. deviation (°)	1.56
Protein mean B factor (Å <sup>2</sup> )	49.4
Waters mean B factor (Å <sup>2</sup> )	53.7

$${}^a\text{R-factor} = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

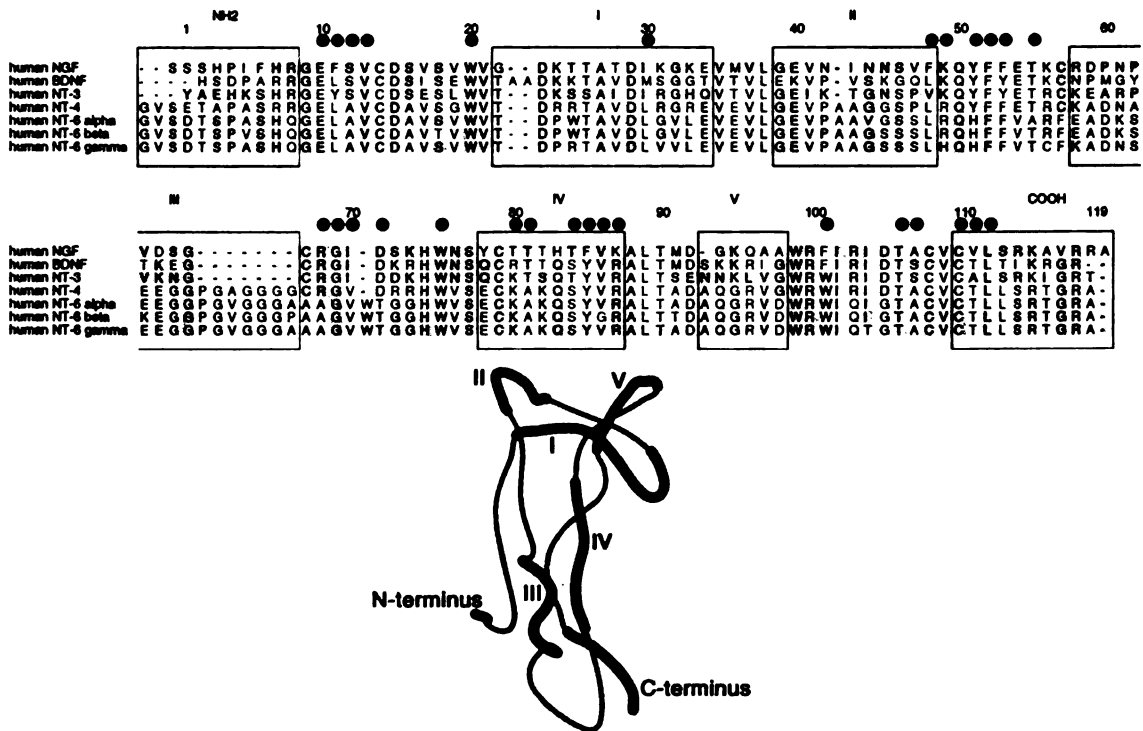
<sup>b</sup>Free R-factor is the R-factor calculated only on the 10% of the reflections that were set aside for cross validation and not used in refinement.



**Figure 2-1**

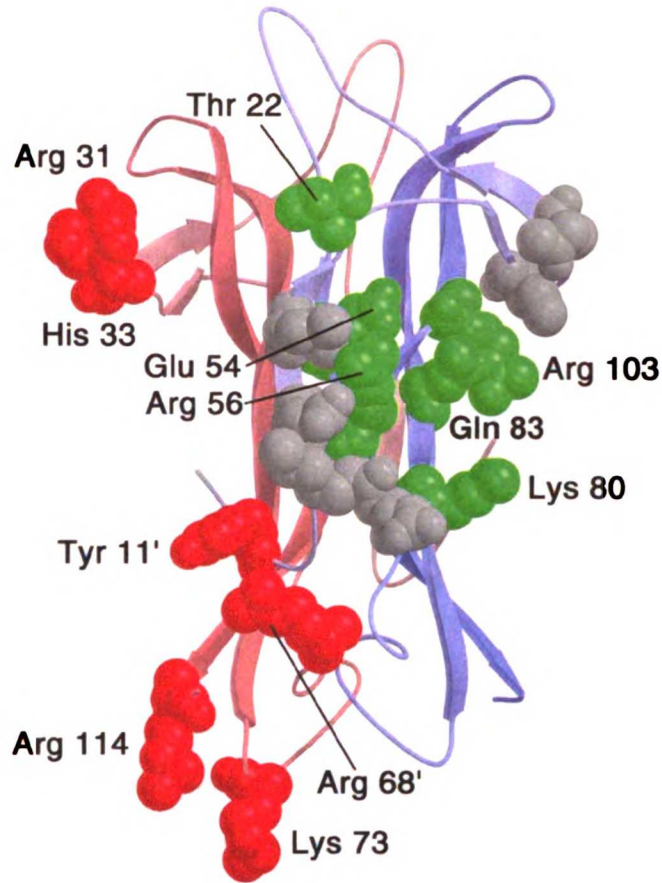
NT-3 shares a similar fold with NGF and the BDNF/NT-3 heterodimer. (a) stereo diagram of alpha carbon trace of the protomer. The three known neurotrophin structures, (b) NT-3 homodimer, (c) NGF dimer, and (d) BDNF/NT-3 heterodimer, with BDNF in purple and NT-3 in green.

www.lifescience.lipson.net



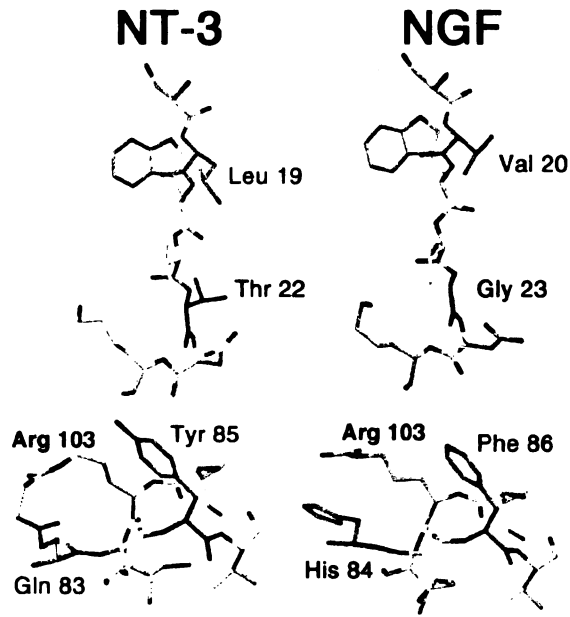
**Figure 2-2**

Sequence alignment of NT-3 with other known human neurotrophins: NGF, BDNF, NT-4, and three NT-6 sequences. The residue numbering is based on NT-3. The black circles represent residues in NT-3 that are in the dimer interface. The roman numerals identify the variable regions of the sequence alignment with the structure.



**Figure 2-3**

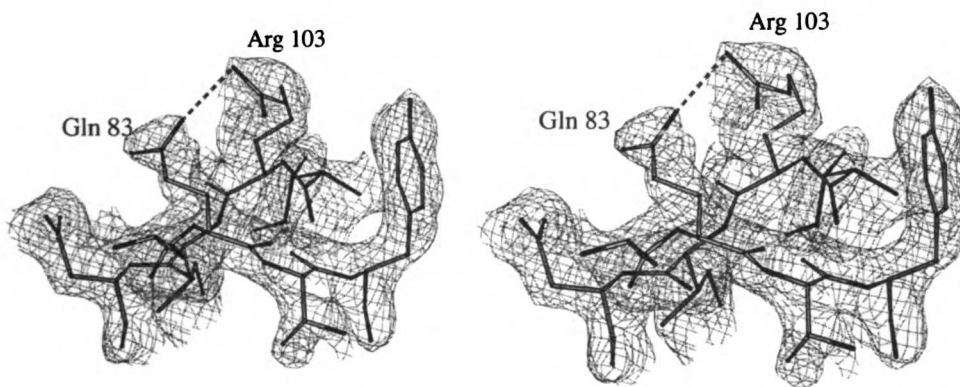
Residues involved in binding to TrkC and p75<sup>NTR</sup> appear to form discrete patches, as shown by alanine scanning mutagenesis. When mutated to alanine, residues in green cause a significant loss of binding affinity to TrkC, those in orange cause a loss of binding affinity to p75<sup>NTR</sup>, and those in gray caused no change in binding affinity to p75<sup>NTR</sup> or TrkC. The two NT-3 protomers are shown in light red and blue.



**Figure 2-4**

The four residues most important for TrkC specificity in NT-3 and their counterparts in NGF. (a) Thr 22 and Leu 19 in NT-3 compared to Gly 23 and Val 20 in NGF; (b) Gln 83 and Tyr 85 in NT-3 compared to His 84 and Phe 86 in NGF. Arg 103 is also shown in both NT-3 and NGF.





**Figure 2-5**

Stereo view of electron density in the region of Arg 103 and Gln 83. Shown is a  $2F_o - F_c$  composite omit map, contoured at 1.0 sigma and superimposed with the NT-3 homodimer model.

## Abstract

Proteolytic degradation processes, mediated by the serine protease urokinase (uPA), underlie the tissue remodeling in cancer metastasis. This degradation is localized by constraining uPA to the surface of tumor and tumor-related cells by the uPA receptor (uPAR). The structural basis of this localization is not understood, though mutagenesis has consistently suggested that the first of the three similar uPAR domains interacts with uPA. The role of the other two domains is not known, nor is the overall structure of uPAR. Besides coordinating the proteolytic activity of uPA, uPAR appears to play a role in signaling. Though the details are only recently being uncovered, binding of uPA to uPAR causes internalization of the complex and activation of second messenger systems that lead to metastasis. Finally, increasing evidence suggests that uPAR is necessary for growth-factor-mediated cell growth. For example, the presence of uPAR is a *sine qua non* for neurotrophin-mediated cell growth and differentiation. To better understand the biology of uPAR activation and uPA binding, I propose to solve the structure of uPAR bound to uPA. This structure will also provide a starting point for structure-based design of an inhibitor to block uPA binding to uPAR, with the eventual goal of interfering with tumor metastasis.

Using multiwavelength anomalous dispersion from platinum-derivative soaked uPAR/uPA crystals, I determined the low-resolution structure of the uPAR / uPA complex to 4 Å. This complex shows that the urokinase molecule is significantly distorted in the complex. It also shows uPAR bound near but not in the active site of uPA. Much of uPA and uPAR could not be specifically built due to poor electron density. Despite this limitation, the low-resolution structure of uPAR / uPA answers and opens new questions about the interaction of these two molecules.

## Introduction

Urokinase receptor (uPAR) is an extracellular protein linked to the cell membrane that plays a role in chemotaxis, cellular adhesion, and cellular migration. By coordinating urokinase (or urokinase-type plasminogen activator, uPA) in fixed, specific locations on the cell surface, the cell can orchestrate conversion of plasminogen to plasmin, which can degrade basement membrane and extracellular matrix and thus allow migration. Urokinase receptor also directly interacts with cell-surface integrin receptors, which are critical to mediating cell-cell adhesion and migration. In addition, urokinase is a necessary component of the signaling pathway of growth factors, though the mechanism of this is not fully understood. Because urokinase receptor ties into the cell growth and migration apparatus, it can be hijacked, as in the case of some cancers, to promote metastasis.

Urokinase was discovered in 1861 when von Bruke noted proteolytic activity of urine [1]. It was nearly 100 years before urokinase was purified in 1947 [2] and independently in 1951 [3]. In 1952, Sobel showed that the protein in urine was not directly fibrinolytic but instead activated endogenous plasminogen to form plasmin; he named this factor urokinase [4]. Urokinase receptor was discovered in 1985 when urokinase-binding sites were noted on the surface of fibroblasts in culture, and uPAR was identified and purified just three years later [5].

Urokinase receptor appears to be critical for chemotaxis of numerous cell types of the immune system [6,7]. A number of outstanding questions remain about uPAR's role in this process. Current evidence suggests that urokinase binding to uPAR causes a conformational change (either by limited proteolysis of the receptor or by a structurally-mediated effect) that exposes a hidden motif [8]. Interaction with an unknown receptor causes activation of a tyrosine kinase of the Src family, p56/p58 hck, which leads to

chemotaxis [9-11]. The critical piece of urokinase receptor that has the strongest chemotactic effect has been localized to the linker between domain 1 and domain 2, residue numbers 87 and 92. A peptide of this region alone, for example, is a very potent chemoattractant in the 0.1 pM range [12]. Within this region, the sequence Arg-Tyr-Leu-Glu (RYLE) is totally conserved across the known uPARs. The structure of these residues in the context of the uPAR linker would greatly facilitate understanding of uPARs role in chemotaxis and may provide some direction in designing small molecule adjuvants to chemotaxis.

Cell adhesion is important in cell attachment, spreading, and migration, activities that are essential for the development of many cells. Urokinase receptor appears to play a role in this process through its interactions with cell surface adhesion molecules. This has been shown in the case of the differentiating monocyte, a cell type in the immune system. Adhesion of differentiating monocytes can be inhibited by blocking urokinase binding to urokinase receptor with an anti-urokinase antibody. In addition, mice with uPAR gene deletion (knock-out) display a failure to recruit neutrophils into the peritoneum under inflammatory conditions. Urokinase receptor interacts with two major cell adhesion molecules, vitronectin and the integrins. The beta-1 integrins, in particular, have been strongly implicated in interacting with urokinase receptor. These integrins not only serve as adhesion molecules, but also mediate the binding, internalization, and degradation of fibrinogen, a major non-plasmin pathway for fibrin clearance. The interaction site for the beta-1 integrins is apparently distinct from that of vitronectin, as a peptide that blocks binding to one but not the other has been found [13]. Interestingly, urokinase receptor is not able to interact with all beta-integrins. The structural determinants of this specificity could be revealed by the structure of urokinase receptor in the region of binding to the integrins.

In addition to its primary roles in chemotaxis, adhesion, and cell spreading, uPAR plays a secondary role in assisting activated growth factor receptors to produce their desired effect. For example, the TrkA neurotrophin receptor requires urokinase receptor to be expressed and active on the cell surface to ensure its differentiative effect on developing neurons [14]. Urokinase receptor's critical role was shown in PC12 pheochromocytoma cells through advanced representational difference analysis [15], which allows cloning of mRNAs that differ between two cells, taking advantage of the difference between EGF and NGF stimulated effect. Farinas-Eisner used two lines of experimental evidence to suggest the connection between uPAR and the TrkA neurotrophin receptor. First, antisense RNA to uPAR blocked NGF-driven morphological differentiation of PC12 cells, while also blocking the genes that are induced by TrkA activation by NGF (such as COX-1 and a type II sodium channel), but did not affect non-NGF-driven signaling pathways. Second, antibodies to uPAR similarly blocked NGF-driven morphological differentiation and mRNA expression typically seen with NGF-activated TrkA. These experiments suggest that uPAR is required for NGF to drive differentiation of PC12 cells.

The role of urokinase receptor in cancer has been shown in numerous research areas. On a clinical level, high levels of urokinase or urokinase receptor are negative prognostic signs for a variety of neoplasms, including breast, pancreatic, gastric, and colorectal cancer [16-25]. In a mouse model, antagonizing the urokinase receptor results in a significant decrease in the metastatic ability of implanted tumor cells [26-29] and also on the growth of the primary tumor [30]. Furthermore, inhibiting either urokinase or urokinase receptor expression in cancer cells using antisense RNA decreases the invasiveness of the tumor [31-33].

One puzzling aspect of urokinase receptor's role in cancer is the fact that the tumor cells themselves do not always produce the urokinase receptor or urokinase themselves. The expression of these proteins, however, is always found to be abnormally high at the leading edge of certain tumor types, including breast cancer, prostate cancer, and melanoma [34]. This may result from the tumor cells inducing stromal cells to produce urokinase receptor. Activation of these urokinase receptors by proteolysis may release a soluble, chemotactic fragment as is discussed above. This could then stimulate the migration and invasiveness of nearby tumor cells. Cleavage of the urokinase receptor between domain 1 and domain 2, the putative location of the chemotactic fragment, is found in breast cancer cells [35,36].

The residues surrounding the active site determine the binding of the substrate that undergoes proteolysis. The amino acid N-terminal to the cleavage in the substrate is called P1, and its binding determinant in uPA is called S1. This nomenclature continues to P2, P3, P4, etc., and correspondingly S2, S3, and S4, while on the other side of the bond cleavage the nomenclature is P1', P2' with S1', S2', and so on. Using a variation on substrate phage display, Ke et al determined the P3 substrate specificity residues for tPA and uPA [37]. Harris et al improved this result by showing the extended substrate specificity of uPA using substrates that held the P1 residue constant (Arg) [38]. Each site (P2, P3, and P4) was tested combinatorially in three groups, using a large fluorogenic substrate library that was randomized along each of these residues while holding the other two constant. This analysis showed that the P2 position was best occupied by small amino acids, Gly, Ala, or Ser. The preferred residues at P3 were significantly different than those in tPA. In uPA, small polar amino acids (Thr and Ser) were seen, whereas tPA preferred the large aromatic residues Phe and Tyr in the substrate. No particular preference was seen in the substrate specificity of P4 in uPA.

Using gel filtration, dynamic light scattering and surface plasmon resonance between soluble-uPAR and sc-uPA (see uPA section below), Shilom et al showed that upon binding they form a 1:1 complex [39]. Surprisingly, however, uPAR alone formed higher molecular weight complexes, principally dimers. Addition of uPA to the oligomeric uPAR uniformly resulted in dissociation to stable 1:1 complexes. This suggests that the higher molecular weight complexes of uPAR were less amenable to binding than monomeric uPAR.

## Primary Structure

### **Urokinase Receptor**

Human uPAR is a 50-60 kDa glycoprotein whose gene (called *Plaur*) on chromosome 19q13.2 encodes for 313 amino acid residues with an N-terminal 22 residue export signal sequence. The 30 residue C-terminal tail is removed in the endoplasmic reticulum during post-translational modification to form a glycosyl-phosphatidylinositol (GPI) linkage, resulting in a mature protein with 283 amino acids. The GPI linkage attaches uPAR to the extracellular membrane without a transmembrane or cytoplasmic domain. The majority of the uPAR gene codes for three similar 90-100-residue domains belonging to the Ly-6/uPAR superfamily of glycolipid-anchored membrane proteins.

Also seen in the sequence are five potential N-linked glycosylation sites (Asn-X-Thr/Ser), and evidence suggests uPAR is not only heavily and heterogeneously glycosylated, but also that removal of the carbohydrate affects ligand binding [40]. Ploug et al performed limited proteolysis and MALDI-MS analysis to determine the carbohydrate heterogeneity patterns in uPAR. In domain I, Asn 52 is the only modified residue, and it typically (90%) shows modification by a small biantennary complex-type oligosaccharide. In domain II, Asn 162 and Asn 172 are modified with bi- and triantennary complex-type oligosaccharides, respectively. Although there are two potential glycosylation sites in domain III, only one was

found modified in the study, Asn 200, which predominantly carried tri- and tetraantennary complexes. Ploug et al showed by enzymatic deglycosylation-footprinting that Asn 52 showed the greatest protection from deglycosylation when bound by ligand. They speculated that domain I has the least complex, and least overall, glycosylation because of this dominant role in ligand binding.

While looking at the conserved pattern of cysteine residues and the gaps between them in the uPAR sequence, Ploug et al discovered three repeats of a Ly6/CD59 family [41,42]. There is very little conservation between members of the Ly6 family (at most 20% to human uPAR) as can be seen in the alignment (figure 3-1). Interestingly, uPAR is the only member of the family that contains more than one repeat. However, within the uPAR subfamily, the proteins are quite related. Between human and rat there is 76% identity, human and mouse, 61%, and between mouse and rat, 91.7%. The Prosite genome-wide search for motif patterns revealed a consensus amino acid code for the Ly6/CD59/uPAR repeat [43]:

$$\{EQR\} - C - \{LIVMFYAH\} - x - C - x(5, 8) - C - x(3, 8) - \{EDNQSTV\} - C - \{C\} - x(5) - C - x(12, 24) - C$$

where x(a,b) is any number (between a and b) of amino acids of any type.





Despite uPAR's solubility when its GPI-anchor is cleaved, soluble uPAR (s-uPAR) did not crystallize in repeated trials. Modifications were sought to improve solubility and increase stability. The construct used in this work has three modifications from the wild type urokinase receptor, all at the end of the first domain. Two of the changes are Arg->Ser mutations, and the third is a Tyr->Ile mutation. The full-length sequence is shown in Figure 3-2.

upar\_trunc

[1-303] mass = 33452.8

Small polar:	D(13)	E(23)	N(18)	Q(13)				
Large polar:	K(10)	R(17)	H(14)					
Small non-polar:	S(26)	T(21)	A(8)	G(27)				
Large non-polar:	L(31)	I(8)	V(13)	M(7)	F(5)	Y(7)	W(2)	
Special:	C(29)	P(13)						

```

1  M G H P P L L P L L L L L H T C V P A S Y G L R C M Q C K T 30
31 N G D C R V E E C A L G Q D L C R T T I V R L W E E G E E L 60
61 E L V E K S C T H S E K T N R T L S Y R T G L K I T S L T E 90
91 V V C G L D L C N Q G N S G S A V T I S S S R Y L E C I S C 120
121 G S S D M S C E R G R H Q S L Q C R S P E E Q C L D V V T H 150
151 W I Q E G E E G R P K D D R H L R G C G Y L P G C P G S N G 180
181 F H N N D T F H F L K C C N T T K C N E G P I L E L E N L P 210
211 Q N G R Q C Y S C K G N S T H G C S S E E T F L I D C R G P 240
241 M N Q C L V A T G T H E P K N Q S Y M V R G C A T A S M C Q 270
271 H A H L G D A F S M N H I D V S C C T K S G C N H P D L D V 300
301 Q Y R 303

```

Figure 3-2

Sequence of urokinase receptor used in crystallography. The signal sequence and the three uPAR domains are colored green, orange, blue, and tan, respectively. The three residues that required modification from wild type are highlighted in yellow.

## **Urokinase**

Urokinase belongs to the P1 family of serine proteases, also known as the trypsin/chymotrypsin family. The amino acid sequence of urokinase reveals three distinct domains: a growth factor receptor binding domain, a kringle-like domain, and a catalytic domain. The first two domains play a role in binding of urokinase to cell surfaces, though urokinase is primarily a soluble molecule found in the serum and interstitial spaces. The N-terminal domain, also called the amino terminal fragment (ATF), is principally involved in binding to uPAR.

As was mentioned previously, uPA's catalytic domain belongs to the trypsin/chymotrypsin superfamily. Within this family, uPA and tPA share common sequence characteristics that are not widely held, especially in the loops between secondary structure elements. This can be seen in the sequence alignment figure below (Figure 3-3).

Like uPAR, the sequence of urokinase used in our crystallography work differed from wild type in a few places. The catalytic functionality was made inactive by mutating the active site Ser to Ala. The second mutation was to mutate a methionine in the B-chain to isoleucine, to improve solubility.

```

      . . . . .
chymotrp  I V N G S D A V P K N P M V L Q D R S . . . . . P E R F C G E I I E E M V V A A E C G . . . . . V S E S D V Y V A G S F D C Q D K E R I Q V L K I A E V F K H P F P I L L V I N I I L L E C A P . . . . . A R P L Q V E A N C L P E A D D P F A C I L C A T G M C E E K Y N A K E
trypsin   I V G I Y M C E F N P V P Q V S L N . . . . . Y E F T D R I I N P M V T E A T E C . . . . . T K S I Q V R I G F K N I K V Z T R G P T N A A K I E R P Q V D R F I E N I E M I E E S R . . . . . A V T N A V S T S L P T A P P A T G E T I E I G C H T L S G A D
urokinase I I S G E F T I S M S P M A A I F R R I R G G E V T T V . . . . . R S S L S P . . . . . K I S A T R E P I D Y P E F E D E T Y L S P R H L N E M G S M Y F P S N L I G F Y T A S P S A S H I I A L E I V F E R T A G P S R I T C S L P S P Y D D P D A T S S T P P I E N E T D I L
tpa       I K S D L P A D I S P M A I F A R H A R S P O E R V L D G I L S A M I L S A A S P E R F P H L T V I L A G E V Y V G E F A P F F W K E I T E F Y F D S . . . . . T E N I A L L K E R A S V V S V C L P P A D I Q L P T A S T E L G Y C K E A S P F
ruler     1 . . . . . 10 . . . . . 20 . . . . . 30 . . . . . 40 . . . . . 50 . . . . . 60 . . . . . 70 . . . . . 80 . . . . . 90 . . . . . 100 . . . . . 110 . . . . . 120 . . . . . 130 . . . . . 140 . . . . . 150

      . . . . .
chymotrp  P D K L S A A L P L L E N A S C K E . . . . . W G R R I S V M I C A S . . . . . A G Q . . . . . V S E S H D G S P L V C Q K D S A M L V G I S M I S D C A T S S P Q V Y A R V F L I P N V E I L A A . . . . .
trypsin   I P D L G L D A P M L C A Y C S A . . . . . Y F P F T R H F P G F C G G . . . . . K D E T Q L E D S P Y V N G . . . . . Q L Q Q T S M I C G S G K N F D V Y F S V Y V M E N T I A A . . . . .
urokinase T F S G L M F T P F S I R S C E P R E T G S T T F K L A A C P G N F . . . . . T S E T S L S I P L C C S L Q R M E L T I S M I K E R L F S H P A T S S H L P H S K S E A S M O L A L . . . . .
tpa       S S E R K R A V R L E P S S R C T S D M L N R T V S M L C A G M S G G P P A E L I D A S G L G G P L V C L D G R S L V G I I S M L S C S E D P P V Y E K V S L L W I R E N N R P . . . . .
ruler     . . . . . 160 . . . . . 170 . . . . . 180 . . . . . 190 . . . . . 200 . . . . . 210 . . . . . 220 . . . . . 230 . . . . . 240 . . . . . 250 . . . . . 260

```

Figure 3-3  
Sequence alignment of the catalytic domains of urokinase, tissue plasminogen activator (tPA), trypsin, and chymotrypsin.

upa\_s378a

[1-411] mass = 46351.7

Small polar:	D(18)	E(20)	N(19)	Q(20)				
Large polar:	K(27)	R(22)	H(17)					
Small non-polar:	S(29)	T(26)	A(17)	G(35)				
Large non-polar:	L(31)	I(19)	V(19)	M(7)	F(12)	Y(18)	W(8)	
Special:	C(24)	P(23)						

```

1  S N E L H Q V P S N C D C L N G G T C V S N K Y F S N I H W 30
31 C N C P K K F G G Q H C E I D K S K T C Y E G N G H F Y R G 60
61 K A S T D T M G R P C L P W N S A T V L Q Q T Y H A H R S D 90
91 A L Q L G L G K H N Y C R N P D N R R R P W C Y V Q V G L K 120
121 P L V Q E C M V H D C A D G K K P S S P P E E L K F Q C G Q 150
151 K T L R P R F K I I G G E F T T I E N Q P W F A A I Y R R H 180
181 R G G S V T Y V C G G S L I S P C W V I S A T H C F I D Y P 210
211 K K E D Y I V Y L G R S R L N S N T Q G E M K F E V E N L I 240
241 L H K D Y S A D T L A H H N D I A L L K I R S K E G R C A Q 270
271 P S R T I Q T I C L P S M Y N D P Q F G T S C E I T G F G K 300
301 E N S T D Y L Y P E Q L K M T V V K L I S H R E C Q Q P H Y 330
331 Y G S E V T T K M L C A A D P Q W K T D S C Q G D A G G P L 360
361 V C S L Q G R M T L T G I V S W G R G C A L K D K P G V Y T 390
391 R V S H F L P W I R S H T K E E N G L A L 411

```

Figure 3-4

Sequence of urokinase used in this crystallography work. The EGF-like domain is colored orange, the kringle domain yellow, the A chain of the catalytic domain green, and the B chain blue. The three residues of the catalytic triad are highlighted in magenta, with the Ser->Ala mutation already incorporated. The second mutation, from Met->Ile is shown highlighted in purple.

## Crystal Structures

### **Urokinase receptor**

Although no structure of the urokinase receptor, or any protease receptor, currently exists, sequence similarity to the CD59 receptor suggests that this molecule may serve as a starting model. The CD59 receptor structure was solved by Kieffer et al in 1994 using NMR [44]. A similar structure by Fletcher and colleagues includes a glycosylated residue [45]. These structures show a uniform row of three beta strands below another smaller sheet of 2 strands and above a short 1.5 turn helix (Figure 3-5).

When Ploug et al compared these structures to urokinase receptor, they found a few differences. First, domain 1 appears to be missing a disulphide bond compared with CD59 and even compared with domains 2 and 3 of uPAR. Second, a key Asp, Asn, or Gln that supports anchors a loop to the core of the protein through a salt bridge appears to be missing from domains 2 and 3 of uPAR. The sequence alignment of uPAR with CD59 in Figure 3-1 also shows a large gap inserted into a beta strand in domains 2 and 3. This may significantly disrupt the secondary structure layout of these domains. These results show even though a model like CD59 is available, caution must be used in applying its domain structure to urokinase receptor.

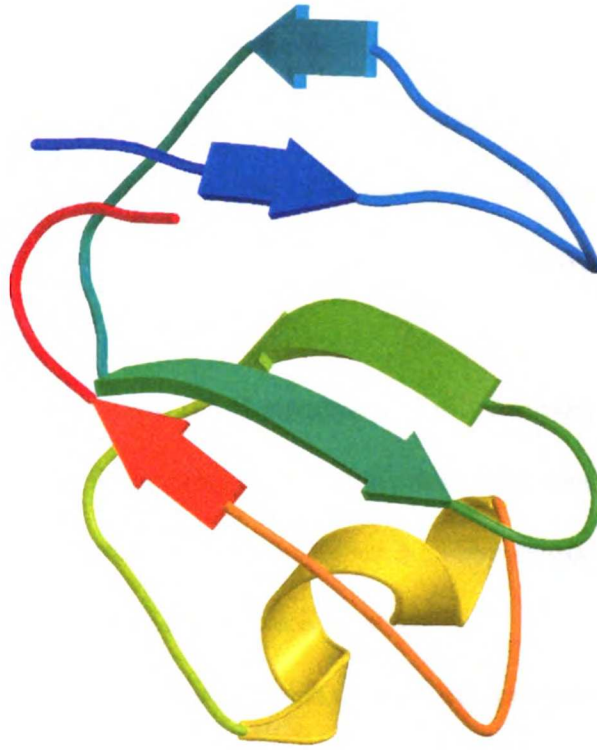


Figure 3-5

Average minimized NMR structure of CD59 extracellular domain (lacking the GPI-anchor).

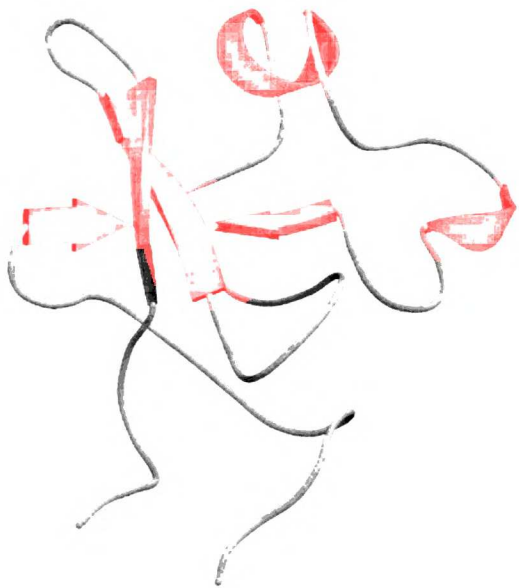


## **Urokinase**

A great deal of structural information can be inferred from mutagenesis, labeling, and deletion experiments. In addition, the X-ray crystal structure of urokinase has been determined by numerous groups, with a handful of inhibitors bound to the active site. Together, these results can help us better interpret the structure of the urokinase receptor-urokinase complex.

Xiang Li solved the structure of the N-terminal domain of urokinase in 1994 by NMR, revealing a typical kringle fold similar to that of the tissue plasminogen activator (tPA) [46]. The minimized average structure is shown in Figure 3-6. This domain, as mentioned above, is thought to play an important role in binding uPAR.

The structure of the catalytic domain, which contains the protease activity of urokinase, was determined by X-ray crystallography by Spraggon and colleagues [47]. This structure revealed that urokinase followed a trypsin-like serine protease fold, namely two perpendicular beta barrels with a central helix, and characteristic loops and helices connecting the beta strands (Figure 3-19 left). Significant differences between urokinase and chymotrypsin were seen in six such loops. One of these loops provided a new hydrophobic face to the substrate-binding surface. Many of the differences were also seen in tissue plasminogen activator (tPA), as can be seen in the sequence alignment of uPA with tPA (Figure 3-2).



**Figure 3-6**

Average minimized NMR structure of a kringle domain of urokinase.

Four recent papers have examined the interaction of uPA with a small molecule inhibitor. The initial structure of uPA, by Spraggon et al was determined in complex with the inhibitor Glu-Gly-Arg chloromethyl ketone (EGRCmk). The chloromethyl ketone forms an irreversible complex with the uPA. The structure showed its tetrahedral hemiketal shape mimics the transition state of the proteolysis reaction. Spraggon et al showed that the S1 pocket of uPA contains Asp189, which makes a charge interaction with the inhibitor Arg (the P1 residue). Also involved in coordinating this Arg are residues Ser 190 and Gly 218. The P2 residue in this structure was Gly, which packs against the His 99 side chain in the S2 pocket [47].

Nienaber et al showed through a novel screening approach the presence of a novel S1 subsite called S1 $\beta$  [48]. Using a shape-diverse library of compounds, urokinase crystals were soaked with mixtures of 32 compounds (though up to 10000 were possible) and the resulting crystal was diffracted and its structure determined. The identity of the bound compound during each iteration of ligand development, presumed to be the best binding compound of the batch, was determined by its electron density pattern. This approach led to use of the compound 2-naphthamide as a lead scaffold for structure-directed optimization ( $K_i = 5.9 \mu\text{M}$ ). The 8-position of this compound had access to a new subsite near the S1 pocket, called S1 $\beta$ . This subsite is lined with residues Gly 218 and Ser 146, the Cys 191 – Cys 220 disulfide bridge, the side chain of Lys 143 and part of Gln 192 [48]. Substitutions at the 8-position led to discovery of 8-methylcarbonyl 2-naphthamide, which bound 100-fold better than the parent compound ( $K_i = 0.04 \mu\text{M}$ ). The structure of this compound bound to uPA led to realization that a six-membered ring would fit better in the S1 $\beta$  site than the methylcarbonyl group. The resulting search led to the creation of 8-aminopyridine

naphthamidine, which bound with a  $K_i$  of 0.03  $\mu\text{M}$  and showed better selectivity against other serine proteases (plasminogen, tPA, kallikrein, trypsin, and thrombin).

Using uPA/benzamidine crystals that permitted exchange soaking of inhibitors, Sperl et al performed structure-based iterative derivatization of a lead compound to optimize binding into the S2, S3, and S4 pockets. Starting with 4-aminomethyl phenylguanidine and an examination of the original uPA structure, the P2 pocket was seen to be too narrow to allow optimization and a spacer was chosen to span the distance between the S1 and the S3/S4 pockets. Then various hydrophobic moieties were chosen for interaction with the S3/S4 residues, settling on the urethanyl derivatives of 4-aminomethyl phenylguanidine. This led to the realization that the spacer was significantly important for binding, and hydrogen-bonding interactions in the spacer were sought. Eventually, N-(1-adamantyl)-N'-(4-guanidiniobenzyl) urea was synthesized, which improved the binding by a factor of 5 ( $K_i = 2.4 \mu\text{M}$ ) over its parent, implicating the formation of a new hydrogen bond. This was confirmed by X-ray crystallography of the final compound with uPA at a resolution of 1.8 Å.

In a recent article, Katz and colleagues at Axys Pharmaceuticals showed the basis of structural specificity of a compound for uPA versus trypsin and thrombin. They determined the structure of 4-iodobenzo(b)thiophene-2-carboxamide in complex with uPA ( $K_i = 0.21 \mu\text{M}$ ), trypsin ( $K_i = 0.44 \mu\text{M}$ ), and thrombin ( $K_i = 20 \mu\text{M}$ ) and also determined its inhibition constant against tPA (16.8  $\mu\text{M}$ ), tryptase (1.5  $\mu\text{M}$ ), and Factor Xa (30  $\mu\text{M}$ ) [49]. Inhibition constants against these serine proteases fell into two groups based on residue 190: serine or alanine. In the Ser 190 class (uPA, trypsin, and tryptase), the inhibitor amidine formed a hydrogen bond at the S1 site to  $\text{O}_\gamma$  in Ser 190, which was not possible in the Ala 190 group of proteases. In addition, differences in the water-mediated hydrogen-bonding network led to unfavorable lengthening of a hydrogen bond in the Ala 190 class versus the Ser 190 class.

These results suggest the structural basis of selectivity of a uPA inhibitor versus other serine proteases.

## Protein Production

Urokinase receptor protein is produced from baculovirus culture. The baculovirus genetic construct was created by Marion Conn, and she perfected the growth and harvesting techniques as well. Essentially the virus-infected insect cells are grown for 96 hours starting at a low viral titer.

Urokinase protein is produced from constitutively expressing, stably transfected, Chinese hamster ovary (CHO) cells in a 10 layer Nunc Cell Factory apparatus. The cells are initially grown in Gibco CHO serum free media (SFM), and transferred to a mixture of half SMF and half Dulbecco's Modified Eagle Media/H16/F12.

## Purification

### Urokinase receptor

Upar protein is initially immunoaffinity purified from the CHO cell supernatant. The protocol is

1. Harvest fluid from CHO-uPAR cells
2. Centrifuge 30 minutes at 6000g and pass through a 0.2  $\mu$ M filter
3. Add PMSF to a final concentration of 1 mM
4. Add 1/10 (w/v) of 100 mM NaPhosphate, 1.4 M NaCl, pH 7.4
5. Pass over a 10 ml protein A pre-column
  - a. wash with 20 beds of wash buffer
  - b. load sample
  - c. collect flow through

- d. regenerate with 10 ml elution buffer followed by 50 ml wash buffer + azide
6. Then pass over the immunoaffinity column (10 ml Protein G Sepharose 4 Fast Flow coupled to purified monoclonal anti-uPAR IgG) using wash buffer #1 of 10 mM NaPhosphate, 140 mM NaCl pH 7.4; wash buffer #2 of 10 mM NaPhosphate, 1.0 M NaCl, pH 7.4; elution buffer of 0.1 M acetic acid pH 2.5 + 0.5 M NaCl
  - a. wash column with 20 volumes
  - b. load sample using 50 ml superloop and inject
  - c. wash with wash buffer 1 for 10 volumes, flow rate 1 ml/min
  - d. wash with wash buffer 2 for 10 volumes, flow rate 1 ml/min
  - e. elute collecting 2ml fractions and neutralize immediately with 0.4 ml, 100 mM NaPhosphate, 1 M NaHCO<sub>3</sub>, pH 9.0, flow rate 1 ml/min
  - f. regenerate with wash buffer
7. Pool fractions and concentrate with centricon-30
8. Check purity by 12.5% SDS-PAGE gel, silver stain
9. Optional Activity assay using soluble uPAR ELISA assay

The uPAR is at this stage concentrated to 3-5 mg/ml with a total volume around 3 ml and needs to be deglycosylated with the enzyme PNGaseF. The reaction is to incubate on a rotator for 12 hours at 37°C:

200 ul PNGaseF

3 ml uPAR

10x G7 buffer (NaPhosphate pH 7.5 buffer)

The uPAR is then concentrated in a centricon-10 to a small volume (less than 1 ml) and dialyzed simultaneously to buffer in preparation for gel filtration purification.

Gel filtration was initially performed on a Pharmacia Superdex 200 30 cm column, but substantially better results were seen using two TosohHaas G3000SW 30 cm columns in series (60 cm). The running buffer used was Tris Buffered Saline (TBS, 50 mM Tris, 150 mM NaCl, pH 7.5). The gel filtration run produced two peaks (Figure 3-7), which were identified by calibration to be approximately 100-120 kDa and 50-60 kDa. These were thought to be monomer and dimer forms of uPAR, so they were subjected to MALDI-TOF (matrix assisted laser desorption ionization – time of flight) mass spectrometry to reveal a mass of 34354 Da and 34475 Da respectively, similar to well within the error of the measurement (Figures 3-8 and 3-9).

To better determine which oligomer of uPAR would be better suited for crystallography, a binding study was undertaken. Monomeric uPAR and dimeric uPAR were mixed with uPA and incubated at room temperature for 30 minutes. Then each sample was run on a non-denaturing gel (Figure 3-10). The gel shows that more monomeric uPAR shifts upon mixing with uPA than does the dimeric uPAR. This suggests that monomeric uPAR binds better to uPA than does dimeric uPAR. At the time, dimeric uPAR had not yet been described in the literature. Subsequently, Shilom and colleagues have improved on these results, using dynamic light scattering, surface plasmon resonance, gel filtration, and electron microscopy to show that monomeric uPAR is in equilibrium with a dimeric form, which binds to uPA with lower affinity [39].

Monomeric uPAR was then purified and mixed with uPA and run over the gel filtration columns as described above. The resultant graph shows two peaks (Figure 3-11) at 60 kDa and another shifted up by 20-30 kDa. A native (non-denaturing) PAGE gel of these

fractions reveals monomeric uPAR and uPA fully shifted, with no residual trace of either species alone, in the early (larger molecular weight) peak, and monomeric uPAR alone in the second peak (Figure 3-12). This result confirms not only that monomeric uPAR was purified, but that the resulting uPAR/uPA complex was also purified and recovered for crystal trials.



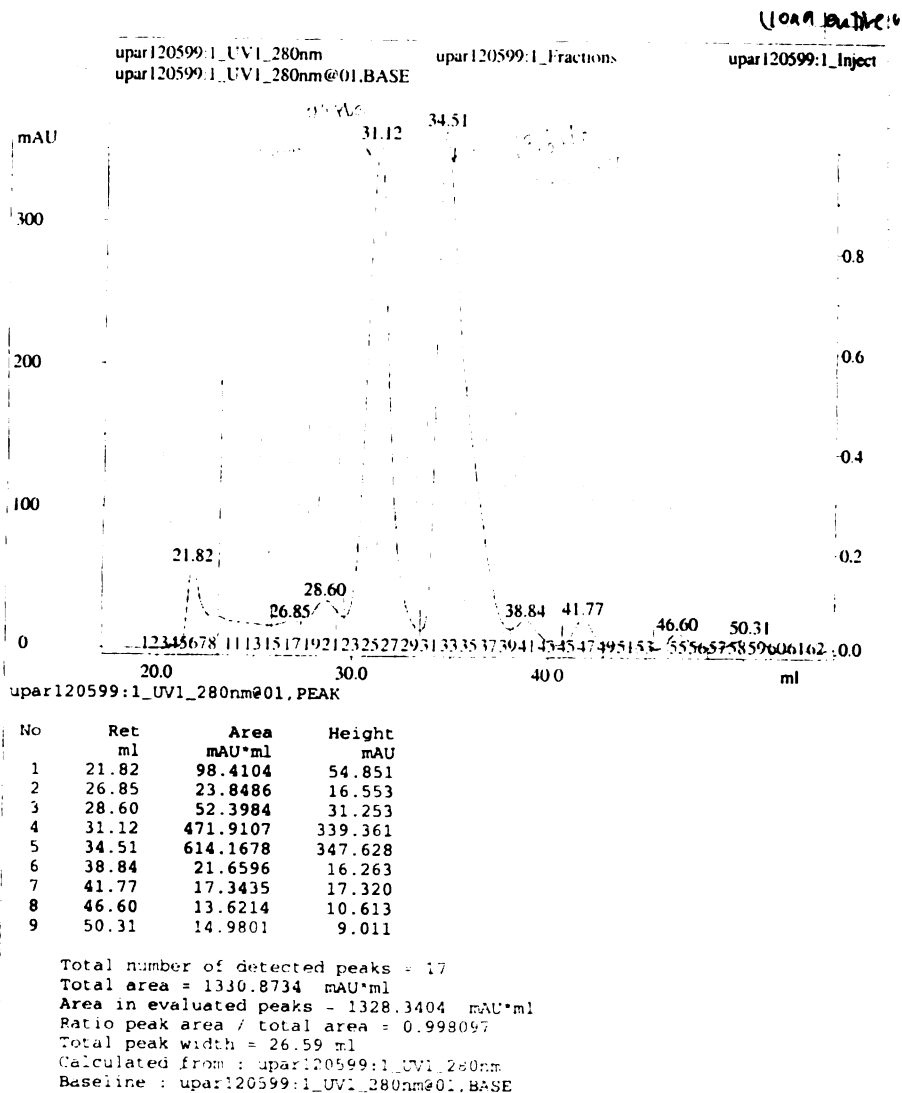


Figure 3-7

Gel filtration purification of uPAR using two TosoHaas G3000SW columns in series. Two peaks are seen at 99-120 kDa and a second at 50-60 kDa. These were shown to be monomer and dimer forms of uPAR.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\manish\upar\umono101.ms

This File # 1 : C:\VOYAGER\DATA\MANISH\UPAR\SMOOTH.LMS

Comment: prep 15-2 (36-38) fractions tp-01-288

Method: HCD1004	Laser: 2280
Accelerating Voltage: 25000	Scans Averaged: 64
Grid Voltage: 65.000 %	Pressure: 5.81e-07
Orbit Wire Voltage: 0.650 %	Low Mass Gate: 2000.0
Delay: 700 ON	Negative Ion: OFF
Sample: 68	Collected: 12/09/99 7:20 PM

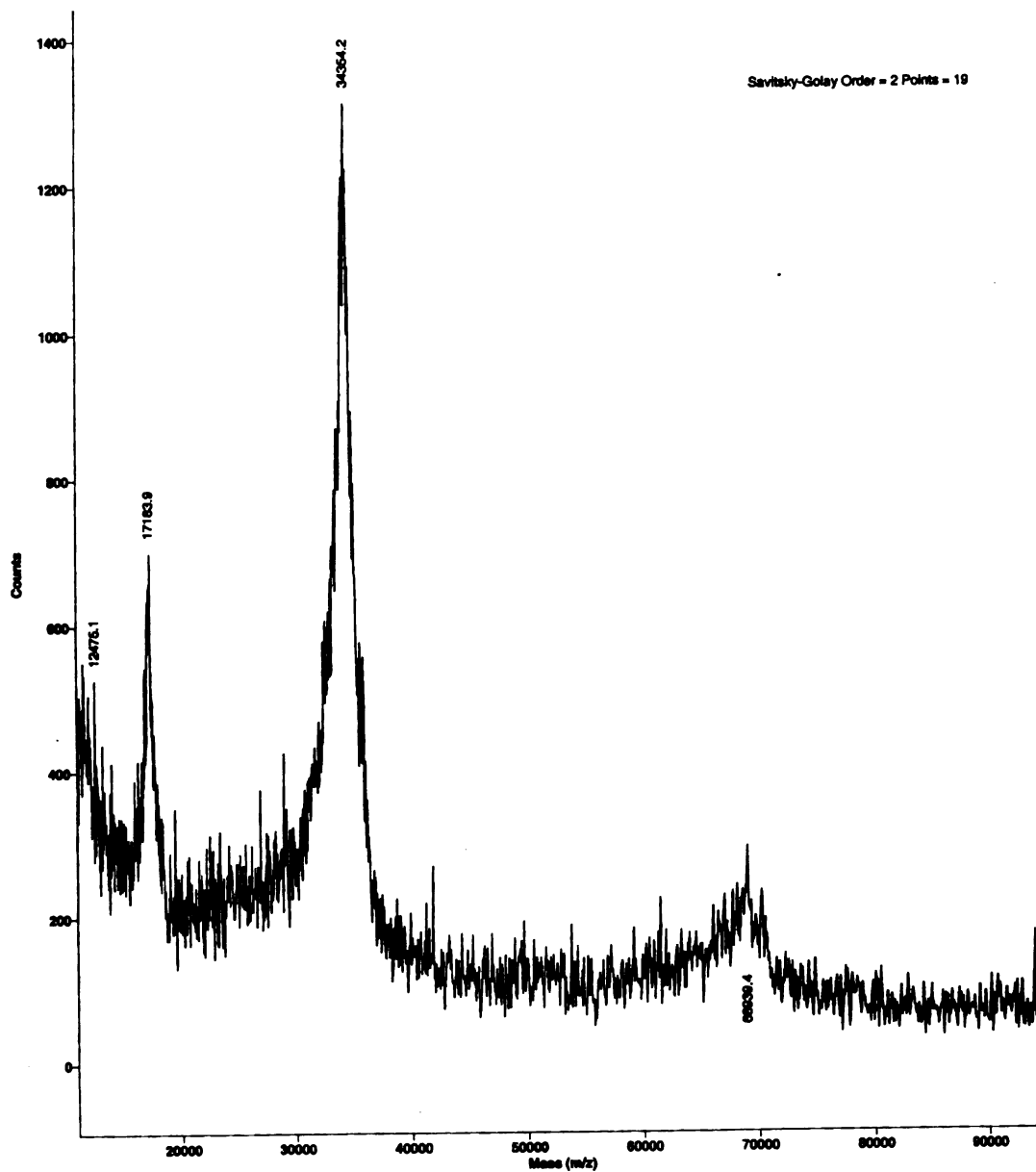


Figure 3-8

MALDI-TOF run of monomeric uPAR

PerSeptive Biosystems

Original Filename: c:\voyager\data\manish\upar\udim1\_02.ms  
The File is: C:\VOYAGER\DATA\MANISH\UPAR\SMOOTH.MS  
Comment: prep 15-2 135-38 fractions 1q-01-209

Method: HCD1004      Laser: 2280  
Accelerating Voltage: 25000      Scans Averaged: 64  
Grid Voltage: 50.000 %      Pressure: 3.57e-07  
Guide Wire Voltage: 0.050 %      Low Mass Gate: 2000.0  
Delay: 750 ON      Negative Ions: OFF  
Sample: 67      Collected: 12/19/99 7:54 PM

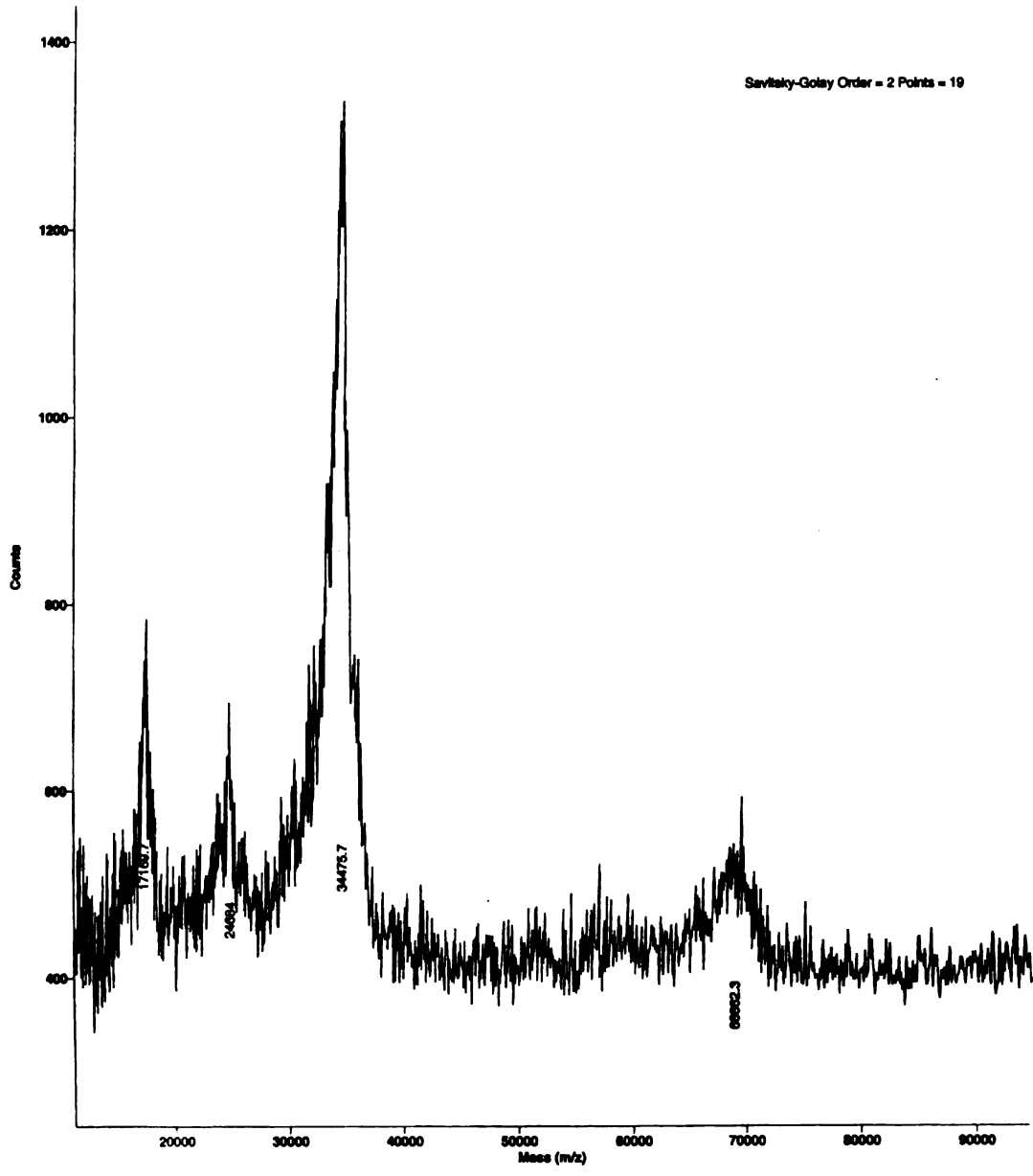


Figure 3-9

MALDI-TOF run of dimeric uPAR



Figure 3-10

Non-denaturing gel showing migration of monomeric and dimeric uPAR with and without uPA. More of the monomeric uPAR shifts upon mixing with uPA than does the dimeric uPAR.

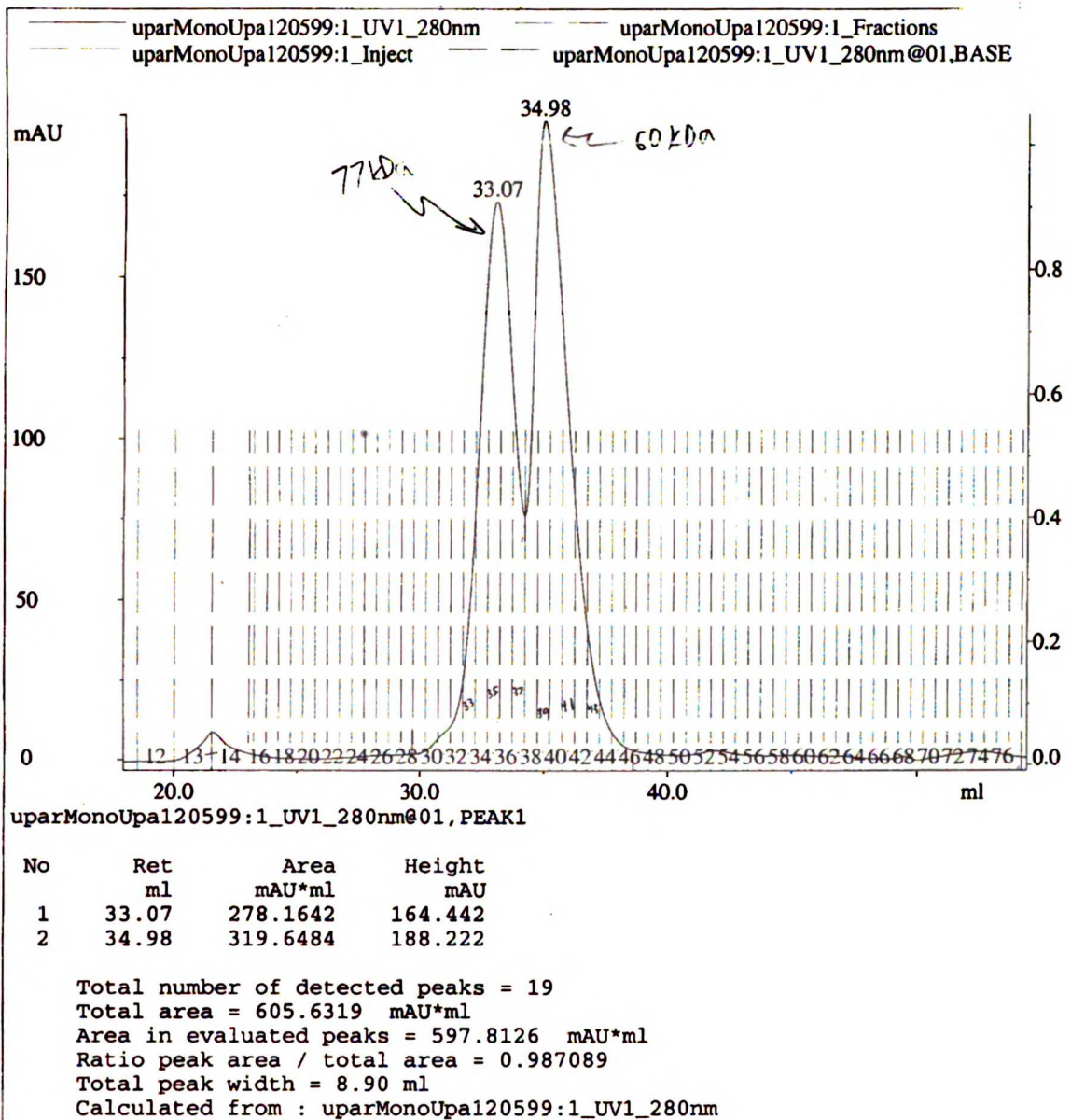


Figure 3-11

Gel filtration purification of monomeric uPAR bound to uPA using two Tosoh G3000SW columns in series. A native gel of these fractions is shown in Figure 3-12.

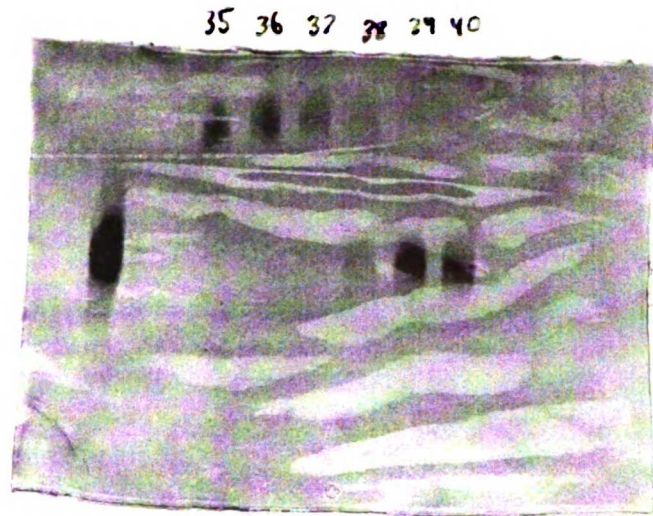


Figure 3-12

Non-denaturing of monomeric uPAR bound to uPA gel filtration (Figure 3-11) fractions. Monomeric uPAR alone is in the second lane as a control. The fractions show a shifted complex that reverts to monomeric uPAR during the second peak of the gel filtration graph.

## Urokinase

uPA was purified from the CHO cell supernatant. The solution (500 – 1000ml) was centrifuged and the supernatant was dialyzed overnight at 4°C against 4 liters of 50 mM NaPhosphate pH 6.0, changing the buffer at least once. The dialyzed solution was then filtered with a 0.22  $\mu$ M filter to remove any remaining cells or debris. The result was approximately 1 liter of dilute solution.

Purification was principally accomplishing using a cation exchange column. The initial protocol from Jennifer Garrison employed a single Pharmacia MonoS column, but this process was limited due to the slow flow rate (1 ml/min) required for loading the large sample. The sample loading stage alone took nearly 17 hours under the original protocol. I obtained faster purification using a two-stage approach.

The first-stage purification was run using Poros media. This media is designed with two types of pores to allow convective flow into the beads. I packed a Poros HS20 column in a 4.6mm x 100mm column, with a final bed volume of 1.7 ml. The packing pressure was 11 MPa, allowing an effective run pressure of 9.6 MPa, which translated to a rate of 8 ml/min of the protein-containing sample. The resultant solution contained significant salt from elution and thus dialysis was needed to return the sample to a low salt condition. The second-stage process relied on the same Pharmacia MonoS media, which has a smaller particle size (10  $\mu$ M), and thus a higher resolution. First stage purification reduced the sample from 1000 ml to approximately 10 ml, and this sped the loading time of the MonoS column by 100 fold. Using a two-stage process allowed me to take advantage of the best features of the two columns: speed and resolution. uPA purification time was reduced from approximately 18 hours to 2.5 hours.

Purification made use of two buffers, the load and wash buffer consisting of 50 mM NaPhosphate, pH 6, and the elution buffer consisting of 50 mM NaPhosphate, 1.0 M NaCl, pH 6.0.

The protocol was:

1. Wash column with 20 volumes wash buffer at 8 ml/min
2. Load protein onto column at the same rate
3. Wash with 10 volumes wash buffer
4. A four-stage elution protocol was used. This protocol resulted from numerous trials to separate the various contaminants from uPA.
  - a. An initial ramp from 0% elution buffer to 17% over 10 column volumes
  - b. Slow ramp from 17% to 22% over 20 column volumes, during which time impurities were eluted.
  - c. Slow ramp from 22% to 40% over 20 column volumes, during which time the uPA peak alone was seen
  - d. Rapid ramp from 40% to 100% over 10 column volumes, during which time other impurities were eluted.
5. The column was washed with 100% elution buffer, then 100% wash buffer in preparation for the next run.
6. Between runs of the PorosHS (first) and MonoS (second), dialyze the pooled fractions in NaPhosphate 50 mM NaCl 50 mM pH 6.0 overnight to decrease the salt left from the first run.

The uPA after the first PorosHS run was quite pure (Figure 3-13). The resulting uPA was then further purified using gel filtration. I used two Tosohaas G3000SW 30 cm columns in series to produce an column effectively 60 cm long. TBS pH 7.5 was used as the



buffer, and the flow rate was 0.5 ml/min. A final gel showing the completed uPA purification is shown in Figure 3-14.

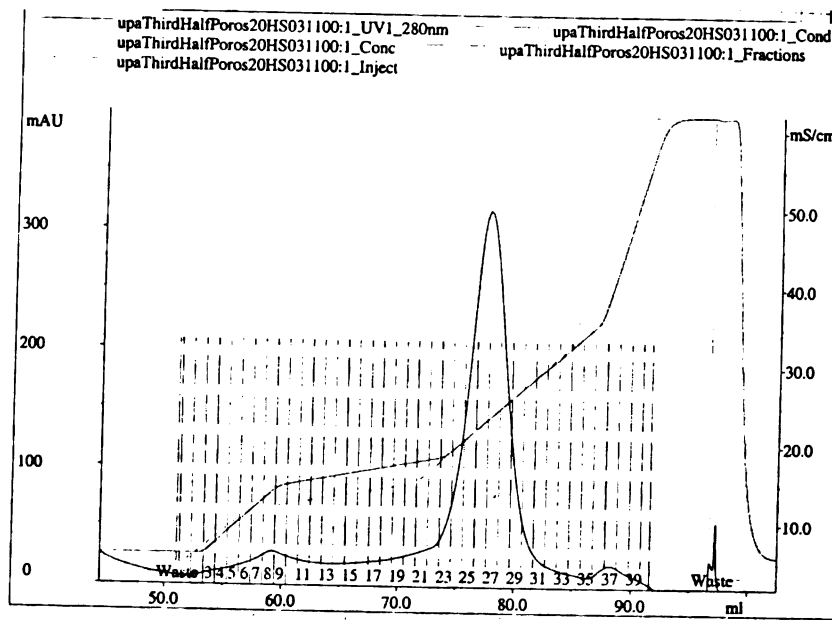


Figure 3-13

Purification of uPA with PorosHS anion exchange column. This shows a test run with the final buffer gradients in place.

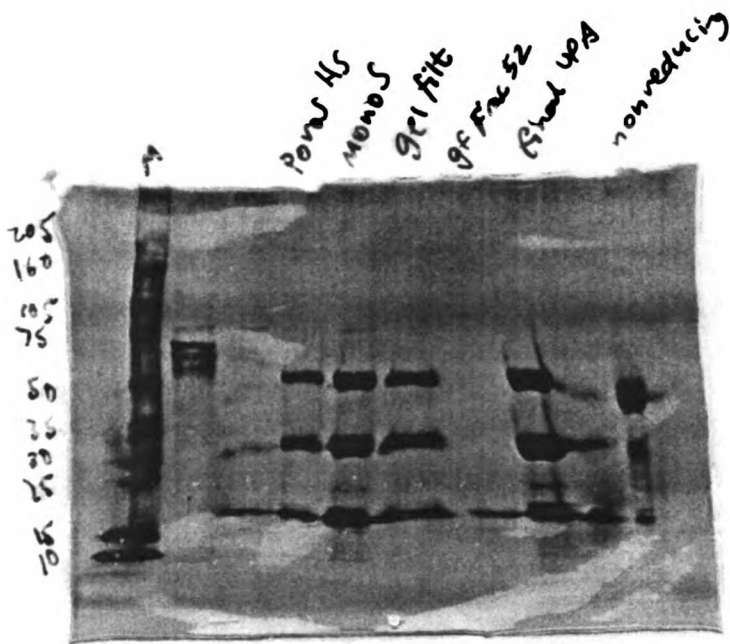


Figure 3-14

SDS-PAGE gel showing stages of purification for uPA. The final column was “final uPA” protein mixed with non-reducing sample buffer. The multiple bands seen in the uPA lanes are the A and B chains of uPA.

## Crystallization

Crystals were produced from two conditions.

- 1.5 M NaFormate ( $\text{NaCHO}_2$ ), 100 mM NaCacodylate ( $\text{NaC}_2\text{H}_6\text{AsO}_2$ ), pH 6.5
- 1.0 M LiSulfate ( $\text{Li}_2\text{SO}_4$ ), 100 mM Tris ( $\text{NH}_2\text{C}(\text{CH}_2\text{OH})_3$ ), pH 8.0

Crystals were best produced using the sitting drop method with a 4  $\mu\text{l}$  drop. The crystallization conditions were incubated at 18°C or 20°C, either worked equally well, for approximately 2 weeks, after which time the crystals would appear. Microseeding had variable success and could produce crystals in 3 days following a 1-2 day pre-seeding equilibration stage. Macroseeding was not successful in producing crystals.

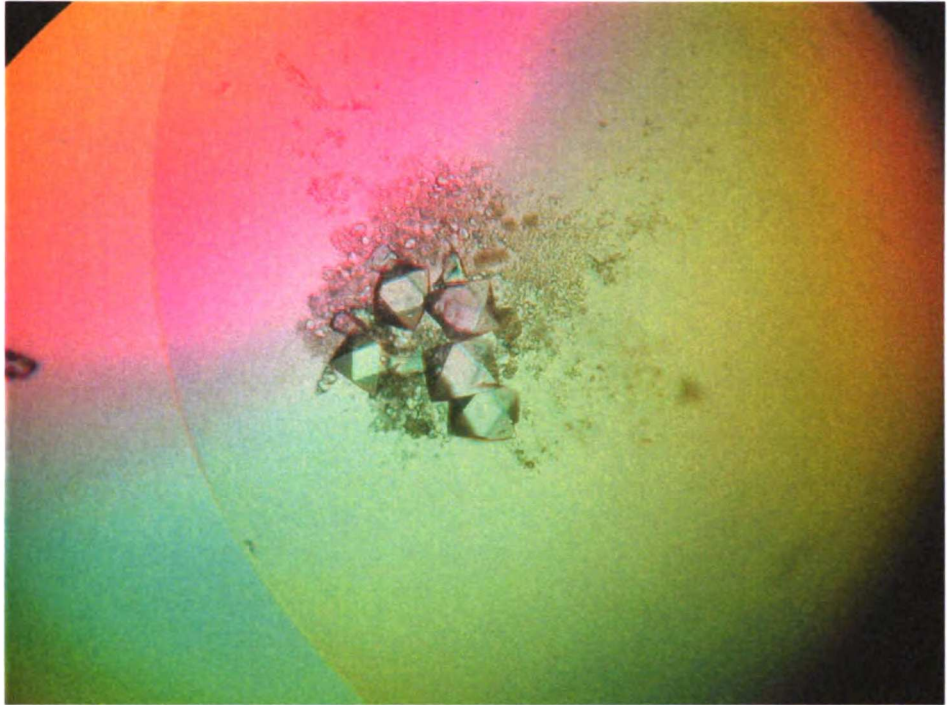


Figure 3-15

A small collection of uPAR + uPA crystals seen in one sitting drop.

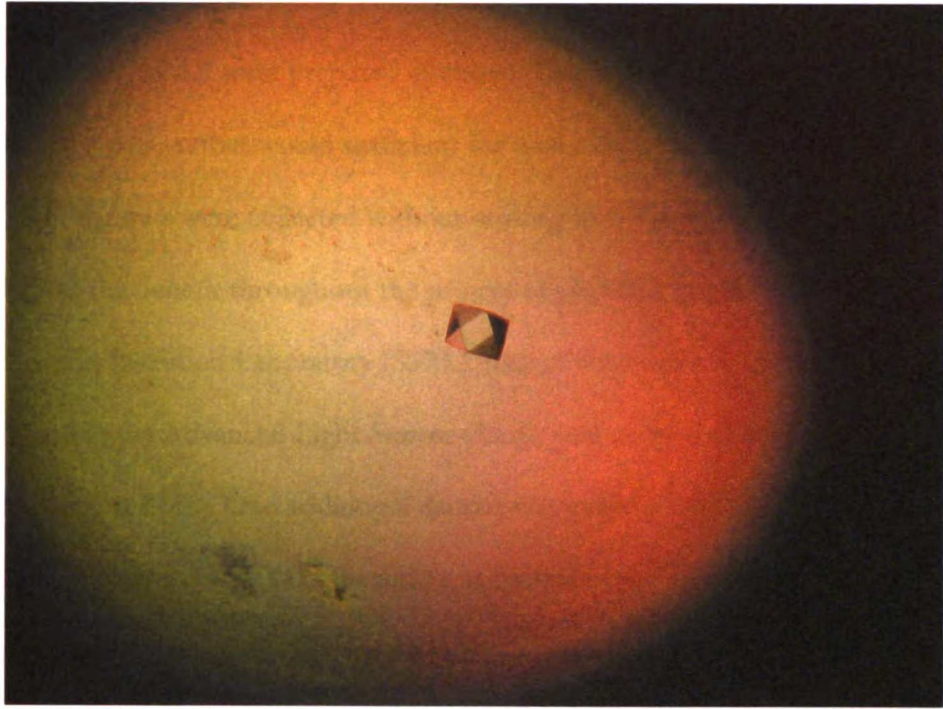


Figure 3-16

A single uPAR + uPA crystal, approximately 200  $\mu\text{M}$  x 200  $\mu\text{M}$ .

## Data collection

Crystals of uPA/uPAR were prepared as above. The initial goal was to collect a native (underivatized) dataset that would be sufficient for molecular replacement of the uPA. Thus the first eight datasets were collected without soaking heavy atoms and at a single wavelength. We had the benefit throughout the project of plentiful synchrotron time at the Stanford Synchrotron Radiation Laboratory (SSRL), part of the Stanford Linear Accelerator Center (SLAC), and at the Advanced Light Source (ALS), part of the Lawrence Berkeley National Laboratories (LBNL). One additional dataset was collected at the National Synchrotron Light Source (NSLS) X12C beamline at Brookhaven.

As can be seen in Table 3-2, subsequent datasets were collected with heavy atom derivatized crystals. Many of the platinum and mercury atom derivatives from the Hampton Research kits were attempted (see Table 3-1), most with little success. Many of the compounds resulted in fracture of the crystal, even in concentrations as low as 1 mM.

<ol style="list-style-type: none"> <li>1. Mersalyl Acid</li> <li>2. Ethyl Mercuric Phosphate</li> <li>3. Mercury (II) Chloride</li> <li><b>4. Mercury (II) Acetate</b></li> <li>5. Ethylmercurithiosalicylic Acid, sodium salt</li> <li>6. Methylmercuric Acetate</li> <li><b>7. Mercury (II) Potassium Iodide</b></li> <li>8. p-Chloromercuriphenylsulfonic Acid</li> <li>9. p-Chloromercuribenzoic Acid</li> <li>10. Ethylmercury Chloride</li> <li>11. 1,4-Diacetoxymercuri-2,3-dimethoxybutane</li> <li>12. para-Chloromercuribenzoic Acid, sodium salt</li> <li><b>13. Mercury (II) Bromide</b></li> <li>14. Mercury (II) Iodide</li> <li>15. Mercury (II) Nitrate monohydrate</li> <li>16. Mercury (II) Cyanide</li> <li>17. Mercury (II) Oxide</li> <li>18. Tetrakis(acetoxymercuri)methane</li> <li>19. Methylmercury (II) Bromide</li> <li>20. Methylmercury (II) Chloride</li> </ol>	<ol style="list-style-type: none"> <li><b>1. Potassium Tetrachloroplatinate (II)</b></li> <li>2. Ammonium Tetrachloroplatinate (II)</li> <li><b>3. Potassium Hexachloroplatinate (IV)</b></li> <li><b>4. Potassium Tetranitroplatinate (II)</b></li> <li><b>5. Potassium Tetracyanoplatinate (II)</b></li> <li><b>6. Dichloroethylenediamine Platinum (II)</b></li> <li>7. Diammino Platinum Dinitrite</li> <li>8. Potassium Tetrabromoplatinate (II)</li> <li><b>9. Potassium Hexabromoplatinate (IV)</b></li> <li><b>10. Platinum Potassium Iodide</b></li> <li>11. Platinum Potassium Thiocyanate</li> <li>12. Di-<math>\mu</math>-iodobis(ethylenediamine)diplatinum (II) Nitrate</li> <li><b>Platinum Potassium Cyanide (II)</b></li> </ol>
--	--

Table 3-1

Mercury and Platinum heavy atom compounds from Hampton Research kits. The ones indicated in bold were used in derivative soaking experiments.



Date	Dataset Name	Location	Crystal	Native/Deriv	Cryo	Highest Res. seen	Unit cell ( $\alpha,\beta,\gamma$ )	Mosaicity	Rmerge	Degrees collected	Completeness to highest useful res.
18 May 99		Raxis 4 (12)	990405-A1	Native	Yes	3.3 Å (60 min exposure)					
20-21 May 99	upar1	SSRL 7-1	990405-C1	Native	Yes	2.4 Å					
20-21 May 99	upar2hi, upar2low	SSRL 7-1	990405-C1	Native	Yes	High: 2.4 Å Low: 3.7 Å	100.56, 100.56, 104.28	0.43 ° / 0.52°	7.8%	110°, 88°	merged : 99.1% to 2.5Å
26 May 99	upar3	SSRL 7-1	990403-A6	Native	Yes	2.6 Å	101.0, 101.0, 104.89	0.46°	17%	100°	97.6% to 3.1Å
26 May 99	upar4	SSRL 9-1	990403-A5	Native	Yes	2.7 Å	100.954, 100.954, 105.257	0.58°	13%	30°	54% to 2.8Å
26 May 99	nb	SSRL 7-1	990403-B1	6 hour soak in 2 mM NB6Cl14	Yes	2.7 Å	100.61, 100.61, 102.21	0.75°	11.9%	125°	99.7% to 3.5Å
26 Jul 99	upar5	Brookhaven X12C	unknown	Native	Yes	3.0 Å	100.46, 100.46, 106.35	0.52°	12.3%	96°	98.1% to 3.2 Å
18 Sep 99	upar6	Raxis 4 (10)	Seed xtal 4	Native	Yes	2.8 Å	100.84, 100.84, 106.25	0.56°	18.7%	140°	100% to 3Å, but missing water ring
22 Oct 99	upar7	ALS	082799-seed2	Native	Yes	2.1 Å	99.880, 99.880, 101.351	1.31°	8.8%	80°	75.7% to 2.2 Å
22 Oct 99	upar8	ALS	082799-seed4	Native	Yes	2.7 Å	101.188, 101.188, 106.089	0.35°	7.5%	47°	89% to 2.9 Å
23 Nov 99	upar9	SSRL 7-1		3 hour soak in 5 mM K3PbCl4	Yes	~ 4 Å	101.591, 101.591, 106.717	3.2°	10.8%	2 x 72°	67.5% to 4 Å
27 Mar 00	upar10	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 7.5	1.5 hr soak in 1 mM K3PbCl4	Yes	3 Å	101.296, 101.296, 106.148	0.51°	4.7%	2 x 150°	97.8% to 4Å
28 Mar 00	upar11	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 8.0	3.5 hr soak in 1 mM K3P(CN)6	Yes	2.4 Å	100.78, 100.78, 106.40	0.3°	5.7%	2 x 130°	99.8% to 2.4Å
28 Mar 00	upar12	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 8.0	12 hr soak in 1 mM K3P(CN)6	Yes	2.4 Å	101.48, 101.48, 105.859	0.56°	6.5%	2 x 135°	99.4% to 3Å
28 Mar 00	upar13	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 8.0	4.5 hr soak in 1 mM K3PbCl4	Yes	3 Å	100.87, 100.87, 106.82	0.26°	3.4%	2 x 101°	99.9% to 3Å
28 Mar 00	upar14	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 8.0	4 hr soak in 1 mM K3P(NO3)4	Yes	2.7 Å	100.88, 100.88, 106.4	0.25°	4%	2 x 63°	99.9% to 2.75Å
28 Mar 00	upar15	SSRL 7-1	U6 (0399) 1.0M LiSO4 Tris 8.0	5 hr soak in 1 mM Fggt2	Yes	2.77 Å	101.14, 101.14, 101.74	0.43°	15.5%	2 x 62°	100% to 2.95Å
12 Apr 00	upar16	ALS 5.0.2	?	?? soak in 1 mM K3PbCl4	Yes	$\lambda_1 = 1.07225 \text{ Å}, 11563 \text{ eV}$ $\lambda_2 = 1.07255 \text{ Å}, 11559 \text{ eV}$ $\lambda_3 = 1.03640 \text{ Å}, 11963 \text{ eV}$	101.78, 101.78, 108.06 101.97, 101.97, 108.13 102.13, 102.13, 108.33	0.45° 0.45° 0.45°	4.8% 3.9% 4.1%	2 x 100° 2 x 100° 2 x 100°	94.2% to 2.55Å 93.5% to 2.9Å 92.8% to 3.7Å
12 Apr 00	upar17	ALS 5.0.2	040399-1.5M NaForm Cac 6.5	Native	Yes					62°	
12 Apr 00	upar18	ALS 5.0.2	0499-1.5M NaForm Cac 6.5	Native	Yes					60°	
12 Apr 00	upar16	ALS 5.0.2	0499, 1.5M NaForm Cac-6.5	3 hr soak in 1 mM K3P(NO3)4	Yes	$\lambda_1 = 1.07220 \text{ Å}, 11563 \text{ eV}$ $\lambda_2 = 1.07250 \text{ Å}, 11559 \text{ eV}$				2 x 60° 2 x 60°	

Table 3-2

## Processing

The datasets were all indexed and scaled with Denzo [50] and Scalepack [51].  $R_{\text{merge}}$ , mosaicity, degrees, and completeness statistics are listed in Table 3-2 for each dataset. The initial datasets were indexed and scaled in space groups P1, P2, and P4 to determine the space group (data not shown). Absent reflections along the axes signified the space group to be P4<sub>1</sub>2<sub>1</sub>2 or P4<sub>3</sub>2<sub>1</sub>2.

Significant anisomorphism was seen between datasets. As is seen in Table 3-2, the unit cell dimensions varied between sets, especially in the C dimension, though A and B are also affected. This significantly limited use of many datasets.

## Heavy atom detection and refinement

Heavy atoms were initially approximated using the programs SOLVE [52] and CNS [53] and refined using SHARP (Statistical Heavy-Atom Refinement and Phasing) [54]. Sites were tested by comparison of putative sites with the isomorphous and anomalous Patterson maps, though this was not always successful as noise often dominated these maps. A better technique for approving initial site was cross phasing. Specifically, a very good site was chosen from the highest peak on the Patterson map, and Fourier difference maps were calculated, leading to a new best peak. This process was iterated until subsequent peaks were below a cutoff threshold (5 sigma above the mean).

Cross phasing to locate sites was also applied from different derivatives. Specifically, a good site from one derivative was used to phase a different dataset. The highest peaks on this dataset (aside from the starting site) were picked as potential heavy atom sites. Then the above technique was applied. This technique of phasing from a single initial site to lead to subsequent sites in the same derivative and other derivatives has significant advantages. First, when considering isomorphous and anomalous contribution to phases, the same handedness is preserved across all the sites. Second, keeping a common reference point prevents using

equivalent sites in the unit cell mistakenly as a new point. Surprisingly, this second mistake is common, especially when switching between two programs that may use a different zero point. Third, the space group of the uPA/uPAR crystals could have been either  $P4_12_12$  or  $P4_32_12$ , as these are not distinguishable from the pseudoprocession images. Preserving the handedness of the 4-fold screw axis required using cross phasing to find additional sites.

## Phasing

### **Molecular replacement**

Initially, phasing was to be accomplished via molecular replacement of uPA. The uPAR molecular would then be iteratively built into the additional, ever-improving electron density. Molecular replacement of the 1LMW PDB model was unsuccessful using datasets upar2, upar3, upar5, and upar8 and the software CNS [53], AMORE [55], and EPMR [56]. Brad Katz of Axys Pharmaceuticals offered a model of uPA from a recently solved structure [49], which was used to repeat the above searches, again unsuccessfully. The eventual structure of uPA in the complex, significantly distorted from the unbound uPA structure, suggests why these initial molecular replacement runs were unsuccessful.

### **Multiple isomorphous replacement**

The next strategy involved using multiple isomorphous replacement (MIR) to solve for the phases. This involved using heavy atom soaked crystals, collecting a derivative dataset, and comparing that dataset against a native set using isomorphous Patterson maps to look for the heavy atoms. Isomorphous Fourier maps were calculated from the phases contributed by the heavy atoms to look for electron density. Many derivative datasets were found to have fundamental low-resolution non-isomorphism against numerous native datasets and were removed from further analysis.

## **Multiwavelength anomalous dispersion**

Due to the limitations mentioned above, the MAD (multiwavelength anomalous dispersion) technique was used to determine phases. The upar16 and upar19 datasets were collected at multiple wavelengths (see Table 3-2). These two datasets were largely complete and compatible with each other, which allowed combined phasing. Following the technique described above, one heavy atom site was determined in upar16 dataset, and this was used to both determine additional sites in upar16, as well as cross phase sites from the upar19 dataset. See Table 3-3 for a final list of sites used.

<b>Dataset</b>	<b>X</b>	<b>Y</b>	<b>Z</b>	<b>Occupancy</b>
upar16	0.533294	0.048289	0.079058	0.91
	0.191349	0.172241	0.121813	0.60
	0.016299	0.118451	0.107940	0.49
	0.751309	0.714714	0.067956	0.40
upar19	0.537624	0.826847	0.038820	0.95
	0.604805	0.693629	0.080270	0.45
	0.548473	0.002192	0.047723	0.30
	0.057720	0.532148	0.085819	0.54
	0.698319	0.318402	0.129422	0.30
	0.729500	0.708567	0.073198	0.54
	0.551084	0.025548	0.080217	0.34

Table 3-3

Heavy atom sites used in MAD phasing

The program SHARP, besides refining the heavy atom sites, also calculates phases using maximum-likelihood to simultaneously scale the data, model for the lack of isomorphism, and refine all heavy-atom parameters from MIR and MAD data, or any mixture of them [57]. Phasing statistics from SHARP revealed an overall figure of merit of 0.45, isomorphous R<sub>cullis</sub> of 0.98, and anomalous R<sub>cullis</sub> of 0.89. Phasing statistics for each dataset is shown in Table 3-4.

	upar16 peak	upar16 remote	upar16 inflection	upar19 peak	upar19 inflection
Rcullis iso	1.0	0.98	0.97	0.90	0.90
Rcullis ano	0.61	0.90	0.74	0.86	0.87
Phasing power iso	1.0	0.22	0.15	0.85	0.77
Phasing power ano	2.6	1.8	2.5	1.6	1.5

Table 3-4

Phasing statistics out of SHARP for each MAD wavelength. iso = isomorphous, ano  
= anomalous

The phases from SHARP were subjected to solvent flattening and flipping in Solomon [58], DM [55], and Squash [59] using solvent contents between 22% and 25%. Solomon and DM were found to over-flatten areas of interest, probably because the density of protein is much higher than in most protein crystals. Unlike DM and Solomon, which use only density histogram matching, Squash uses a 2D histogram matching method, and employs the joint probability distribution of electron density values and their gradients to constrain density modification. In addition, Squash can perform solvent flattening/flipping and phase extension. Squash was used to perform density modification to 2.5 Å, though analysis of resulting maps showed that the phases were only reliable to 3.5 – 4 Å. The figure of merits and mean phase angle changes for various resolution ranges after the final Squash iteration are shown in Table 3-5.



Correlation coefficient b/w F\_observed and F\_modified is 0.809

Resolution range	Number of reflections	Lack of closure	Mean FOM(Obs)	Mean FOM(Modi)	Mean FOM(Comb)	Mean  Phi(Obs)-Phi(Modi)	Mean  Phi(Obs)-Phi(Comb)
Infinity - 10.00	324	2301941.3	0.635	0.474	0.785	50.92	45.85
10.00 - 7.07	627	729372.5	0.650	0.631	0.841	50.00	43.56
7.07 - 5.77	789	398434.6	0.691	0.659	0.851	47.57	41.54
5.77 - 5.00	920	470012.1	0.642	0.636	0.830	52.89	44.63
5.00 - 4.47	1022	503956.5	0.586	0.680	0.836	50.77	45.08
4.47 - 4.08	1122	494749.0	0.540	0.642	0.806	53.20	48.04
4.08 - 3.78	1204	356270.3	0.714	0.671	0.864	45.32	33.96
3.78 - 3.54	1277	302453.7	0.386	0.682	0.779	69.67	62.62
3.54 - 3.33	1358	205287.2	0.457	0.623	0.742	69.70	63.45
3.33 - 3.16	1308	150412.9	0.336	0.579	0.668	85.28	82.03
3.16 - 3.02	1192	138308.5	0.331	0.451	0.577	86.04	82.90
3.02 - 2.89	1396	89914.6	0.328	0.537	0.624	87.65	83.86
2.89 - 2.77	1607	63484.4	0.290	0.590	0.644	88.72	88.07
2.77 - 2.67	1645	50249.4	0.275	0.592	0.635	94.62	94.12
2.67 - 2.58	1694	43607.9	0.284	0.574	0.620	90.31	92.42
2.58 - 2.50	529	39323.7	0.308	0.534	0.586	94.65	96.48

Table 3-5

Figures of merit (FOM) versus resolution ranges after the final iteration of histogram matching, density modification, and phase extension in Squash. A sharp fall-off can be seen in the observed FOMs below 4 Å. The final two columns show mean phase angle difference between the observed phases and the modified phases.

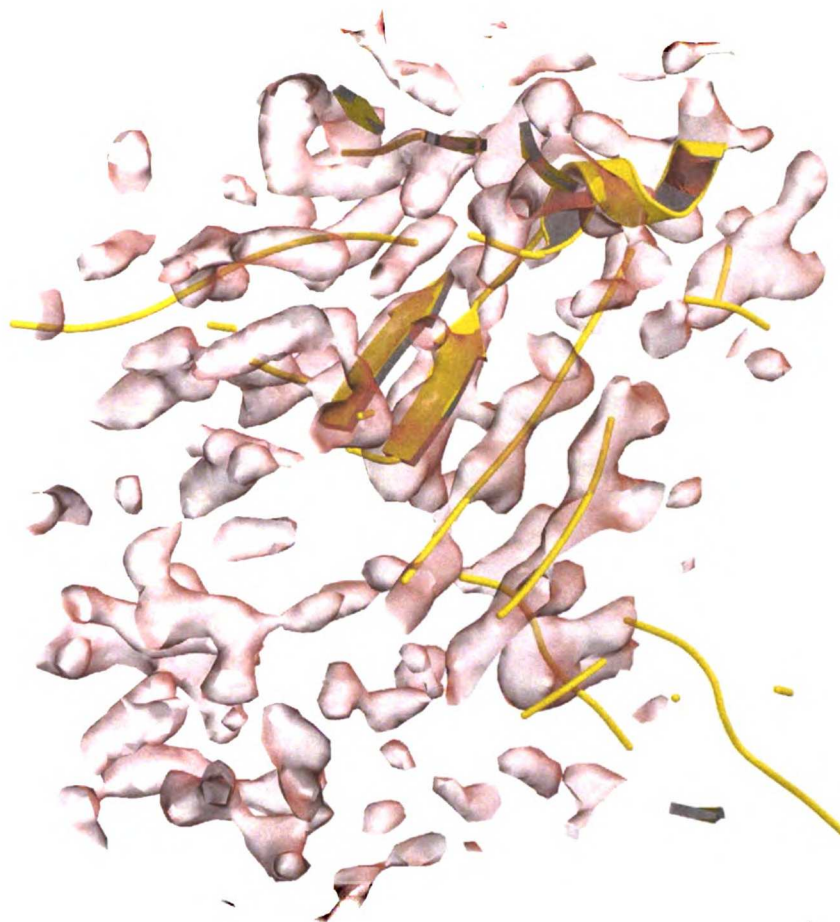


Figure 3-17

Density modified experimental  $2F_o - F_c$  electron density in the region of urokinase, contoured at 2.0 sigma. Visible are four strands of one beta barrel and the main alpha helix at the bottom of the figure. Figure was prepared with Molscript [60] and Conscript [61].



Figure 3-18

2Fo-Fc electron density at 1.8 sigma contour for urokinase receptor shown alongside urokinase. Circled roughly in red are three domains of uPAR. Figure was prepared with Molscript [60] and Conscript [61].

## Building uPA and uPAR

The electron density maps from density modification to 4 Å showed globular-looking regions. This was used as a cue to attempt phased molecular replacement of uPA into the datasets, using only low-resolution phases (10 – 4 Å). Phases from SHARP were converted to Hendrickson-Lattman coefficients and imported into CNS for the cross rotation and phased translation searches. Because the total amount of scattering material used for the search was under 50%, it was not expected that the R-factor would significantly change if the proper solution were found. Thus, an alternate scoring system for the searches was employed, based on the packing score. The translation solutions with the best packing scores were further analyzed using a custom program to determine if a packing solution geometrically intersected a symmetry copy of itself. This software constructs a convex hull around each solution symmetry mate and looks for intersections of the hulls, speeding comparisons of molecular intersection from a manual time of 5 – 10 minutes to an automatic time of 1.2 seconds on average (Manish Butte, unpublished software). This led to a solution that appeared to coincide with features of electron density very well. Specifically, beta strands from urokinase could be seen lined together in the electron density (see Figure 3-17 for an example).

Though the general position of the uPA model in density was correct, the model required some modifications to fit into density. Specifically, the two beta barrels were slightly spread apart and twisted, possibly making room for new electron density that is associated with uPAR (Figure 3-19). Unfortunately, the low-resolution phases were not adequate to complete building the uPA model, and the extent of these changes is not known. Loops connecting the beta strands were not seen, and many of the beta strands themselves were incomplete. Similarly, refinement using CNS of the urokinase model either using

experimental phases with a maximum likelihood target or density modified phases with the MLHL target did not significantly improve the structure.



Figure 3-19

Left, urokinase model used for phased molecular replacement. Highlighted are the three residues of the catalytic triad. Right, urokinase as seen in the low-resolution complex structure. Clearly visible are the two beta barrels, though distorted.

Plentiful electron density was seen apart from the urokinase model, attributed to uPAR. The quality of this electron density was too low for secondary structure elements to be detected, so phase combination was attempted using the urokinase model as it was built and the density modified experimental phases. One challenge was to down-weight the role of the urokinase model, which might otherwise swamp out the signal from the relatively weaker uPAR contribution, by using phase blurring as implemented in CNS. The phase combination was performed using the CCP4 program SigmaA [55] and the results from the run can be seen in Table 3-6. The mean figures of merit for low-resolution shells were slightly improved.

Quanta (Molecular Simulations, Inc.) was used to detect bones and secondary structure in the urokinase regions of electron density, and these were developed into three small regions. Although much electron density remains unfit, the quality of these areas is too low to properly support fitting. Protein/solvent boundaries, however, were greatly improved by both squash and by phase combination, resulting in three clear blobs of electron density where one would expect uPAR (Figure 3-18).

Resolution Limits	Number of reflections	Combined	Experimental	Model
1000.00-8.06	409	0.8690	0.8136	0.4629
8.06-5.70	878	0.8240	0.7780	0.3954
5.70-4.66	1187	0.7502	0.6919	0.3777
4.66-4.03	1430	0.7013	0.6164	0.3785
4.03-3.61	1613	0.5868	0.5375	0.2385
3.61-3.29	1820	0.4961	0.4693	0.1536
3.29-3.05	1648	0.3868	0.3633	0.1326
3.05-2.85	1924	0.3599	0.3433	0.1092
2.85-2.69	2264	0.3061	0.2949	0.0768
2.69-2.55	2266	0.2989	0.2885	0.0710

Table 3-6

Figures of merit from the urokinase model (after phase blurring), experimental phases, and the phase combination.



## Discussion

Urokinase receptor (uPAR) is a cell-surface GPI-linked receptor that binds urokinase (uPA), a soluble serine protease found in the serum, urine, and interstitial spaces. Activation of uPAR by uPA triggers a number of processes. Primarily, urokinase is restrained on the surface of the cell to coordinate proteolytic activity against basement membrane or extracellular matrix components. This facilitates a cell to grow or migrate in a particular direction. In addition, uPAR appears to signal through spatial association with Src-family tyrosine kinases, which in turn activate signaling cascades directing a cell to grow or differentiate. Curiously, growth-factor-directed cell growth appears to require the presence of functional uPAR, as if uPAR were also the target of a signal to grow a cell in a particular direction. Because of the close ties uPAR has with the cellular apparatus for migration and growth, numerous cancers have hijacked its functionality to direct neoplastic metastasis to the extent that the presence of uPAR often portends an adverse prognosis.

Numerous open questions remain about the basis of uPAR activation, and uPA binding to uPAR. In addition, a significant effort has developed to develop small molecules to interrupt the activation of uPAR, the proteolytic function of uPA, or the binding of uPA. These issues have compelled us to determine the structure of uPAR in complex with uPA. We produced tetragonal crystals that diffracted to below 3.0 Å, and with multiwavelength anomalous dispersion phasing, we were able to determine the structure of the complex to 4.0 – 4.5 Å.

Though the double beta barrel structure of urokinase appears mildly disturbed in the bound structure, the active site, which sits just below the central helix and between the two barrels, appears not be involved with uPAR density (Figure 3-18). Because of phasing limitations, however, the extent of disruption of the uPA is not known. This supports the

notion that urokinase is free to proteolyse extracellular targets while bound to uPAR. Higher resolution data would help clarify if uPAR is itself cut in its interaction with uPA (though our urokinase was proteolytically dead) as is suggested from the literature on immune cell chemotaxis. We were unable to resolve the structure of the linker between domain 1 and domain 2, and thus the chemotactic role of the peptide therein remains elusive.

While the structure of uPA bound with uPAR does not shed any light on the putative interaction between uPAR and intracellular Src-family kinases, it may on uPAR's interaction with the integrins. A detailed analysis of the surface of uPAR to look for commonalities with other integrin partners could help clarify their roles.

Finally, the structure of the complex can help elucidate the role of inhibitors on this signaling system. If uPAR-directed cell migration is activated by proteolysis, and a urokinase-specific proteolytic inhibitor is discovered, such a compound might have a place in the therapeutic anti-cancer arsenal. Moreover, these results may usher in a new class of inhibitors designed to target the interface rather than the urokinase active site.

## References

- 1 von Bruke, E. (1861) Die verdauende substanz im urin. SB Akad Wiss (Vienna) 43, 601-615.
- 2 Macfarlane, R.G. and Piling, J. (1947) Fibrinolytic activity in normal urine. Nature 159, 779.
- 3 Williams, J.R.B. (1951) The fibrinolytic activity of the urine. Br J Exp Pathol 32, 530-537.
- 4 Sobel, G.W., Mohler, S.R., Jones, N.W., Dowdy, A.B.C. and Guest, M.M. (1952) Urokinase: A potent activator of plasma profibrinolysin extracted from urine. 171, 768-769.
- 5 Nielsen, L.S., Kellerman, G.M., Behrendt, N., Picone, R., Dano, K. and Blasi, F. (1988) A 55,000-60,000 mr receptor protein for urokinase-type plasminogen activator. Identification in human tumor cell lines and partial purification. J Biol Chem 263, 2358-63.
- 6 Gyetko, M.R., Todd, R.F., 3rd, Wilkinson, C.C. and Sitrin, R.G. (1994) The urokinase receptor is required for human monocyte chemotaxis in vitro. J Clin Invest 93, 1380-7.
- 7 Gyetko, M.R., Sitrin, R.G., Fuller, J.A., Todd, R.F., 3rd, Petty, H. and Standiford, T.J. (1995) Function of the urokinase receptor (cd87) in neutrophil chemotaxis. J Leukoc Biol 58, 533-8.
- 8 Mondino, A., Resnati, M. and Blasi, F. (1999) Structure and function of the urokinase receptor. Thromb Haemost 82 Suppl 1, 19-22.

- 9 Bohuslav, J., Horejsi, V., Hansmann, C., Stockl, J., Weidle, U.H., Majdic, O. et al. (1995) Urokinase plasminogen activator receptor, beta 2-integrins, and src- kinases within a single receptor complex of human monocytes. *J Exp Med* 181, 1381-90.
- 10 Resnati, M., Guttinger, M., Valcamonica, S., Sidenius, N., Blasi, F. and Fazioli, F. (1996) Proteolytic cleavage of the urokinase receptor substitutes for the agonist-induced chemotactic effect. *Embo J* 15, 1572-82.
- 11 Fazioli, F., Resnati, M., Sidenius, N., Higashimoto, Y., Appella, E. and Blasi, F. (1997) A urokinase-sensitive region of the human urokinase receptor is responsible for its chemotactic activity. *Embo J* 16, 7279-86.
- 12 Trigwell, S., Wood, L. and Jones, P. (2000) Soluble urokinase receptor promotes cell adhesion and requires tyrosine- 92 for activation of p56/59(hck) [in process citation]. *Biochem Biophys Res Commun* 278, 440-6.
- 13 Wei, Y., Lukashev, M., Simon, D.I., Bodary, S.C., Rosenberg, S., Doyle, M.V. et al. (1996) Regulation of integrin function by the urokinase receptor. *Science* 273, 1551-5.
- 14 Farias-Eisner, R., Vician, L., Silver, A., Reddy, S., Rabbani, S.A. and Herschman, H.R. (2000) The urokinase plasminogen activator receptor (upar) is preferentially induced by nerve growth factor in pc12 pheochromocytoma cells and is required for ngf-driven differentiation. *J Neurosci* 20, 230-9.
- 15 Lisitsyn, N. and Wigler, M. (1993) Cloning the differences between two complex genomes. *Science* 259, 946-51.
- 16 Buo, L., Meling, G.I., Karlsrud, T.S., Johansen, H.T. and Aasen, A.O. (1995) Antigen levels of urokinase plasminogen activator and its receptor at the tumor-host interface

- of colorectal adenocarcinomas are related to tumor aggressiveness. *Hum Pathol* 26, 1133-8.
- 17 Zheng, Q., Tang, Z.Y., Xue, Q., Shi, D.R., Song, H.Y. and Tang, H.B. (2000) Invasion and metastasis of hepatocellular carcinoma in relation to urokinase-type plasminogen activator, its receptor and inhibitor [in process citation]. *J Cancer Res Clin Oncol* 126, 641-6.
- 18 Yang, J.L., Seetoo, D., Wang, Y., Ranson, M., Berney, C.R., Ham, J.M. et al. (2000) Urokinase-type plasminogen activator and its receptor in colorectal cancer: Independent prognostic factors of metastasis and cancer-specific survival and potential therapeutic targets [in process citation]. *Int J Cancer* 89, 431-9.
- 19 McCabe, N.P., Angwafo, F.F., 3rd, Zaher, A., Selman, S.H., Kouinche, A. and Jankun, J. (2000) Expression of soluble urokinase plasminogen activator receptor may be related to outcome in prostate cancer patients. *Oncol Rep* 7, 879-82.
- 20 Rha, S.Y., Yang, W.I., Gong, S.J., Kim, J.J., Yoo, N.C., Roh, J.K. et al. (2000) Correlation of tissue and blood plasminogen activation system in breast cancer. *Cancer Lett* 150, 137-45.
- 21 Ho, C.H., Yuan, C.C. and Liu, S.M. (1999) Diagnostic and prognostic values of plasma levels of fibrinolytic markers in ovarian cancer. *Gynecol Oncol* 75, 397-400.
- 22 Miyake, H., Hara, I., Yamanaka, K., Arakawa, S. and Kamidono, S. (1999) Elevation of urokinase-type plasminogen activator and its receptor densities as new predictors of disease progression and prognosis in men with prostate cancer. *Int J Oncol* 14, 535-41.
- 23 Nozaki, S., Endo, Y., Kawashiri, S., Nakagawa, K., Yamamoto, E., Yonemura, Y. et al. (1998) Immunohistochemical localization of a urokinase-type plasminogen

- activator system in squamous cell carcinoma of the oral cavity: Association with mode of invasion and lymph node metastasis. *Oral Oncol* 34, 58-62.
- 24 Sier, C.F., Stephens, R., Bizik, J., Mariani, A., Bassan, M., Pedersen, N. et al. (1998) The level of urokinase-type plasminogen activator receptor is increased in serum of ovarian cancer patients. *Cancer Res* 58, 1843-9.
- 25 Cantero, D., Friess, H., Deflorin, J., Zimmermann, A., Brundler, M.A., Riesle, E. et al. (1997) Enhanced expression of urokinase plasminogen activator and its receptor in pancreatic carcinoma. *Br J Cancer* 75, 388-95.
- 26 Crowley, C.W., Cohen, R.L., Lucas, B.K., Liu, G., Shuman, M.A. and Levinson, A.D. (1993) Prevention of metastasis by inhibition of the urokinase receptor. *Proc Natl Acad Sci U S A* 90, 5021-5.
- 27 Ignar, D.M., Andrews, J.L., Witherspoon, S.M., Leray, J.D., Clay, W.C., Kilpatrick, K. et al. (1998) Inhibition of establishment of primary and micrometastatic tumors by a urokinase plasminogen activator receptor antagonist. *Clin Exp Metastasis* 16, 9-20.
- 28 Xing, R.H., Mazar, A., Henkin, J. and Rabbani, S.A. (1997) Prevention of breast cancer growth, invasion, and metastasis by antiestrogen tamoxifen alone or in combination with urokinase inhibitor b-428. *Cancer Res* 57, 3585-93.
- 29 Guo, Y., Higazi, A.A., Arakelian, A., Sachais, B.S., Cines, D., Goldfarb, R.H. et al. (2000) A peptide derived from the nonreceptor binding region of urokinase plasminogen activator (upa) inhibits tumor progression and angiogenesis and induces tumor cell death in vivo. *Faseb J* 14, 1400-10.

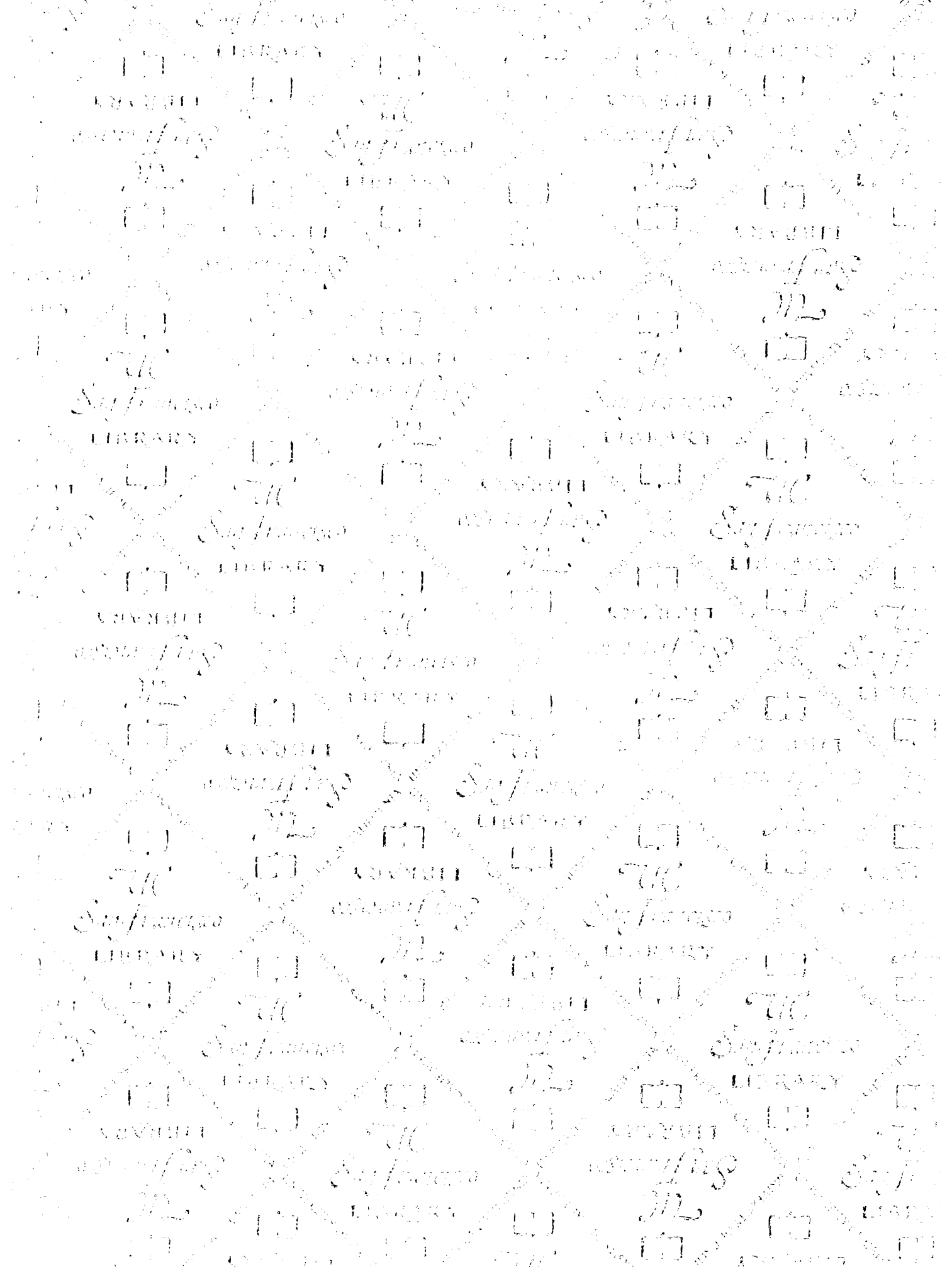
- 30 Min, H.Y., Doyle, L.V., Vitt, C.R., Zandonella, C.L., Stratton-Thomas, J.R., Shuman, M.A. et al. (1996) Urokinase receptor antagonists inhibit angiogenesis and primary tumor growth in syngeneic mice. *Cancer Res* 56, 2428-33.
- 31 Kook, Y.H., Adamski, J., Zelent, A. and Ossowski, L. (1994) The effect of antisense inhibition of urokinase receptor in human squamous cell carcinoma on malignancy. *Embo J* 13, 3983-91.
- 32 Ossowski, L. and Reich, E. (1983) Antibodies to plasminogen activator inhibit human tumor metastasis. *Cell* 35, 611-9.
- 33 Hearing, V.J., Law, L.W., Corti, A., Appella, E. and Blasi, F. (1988) Modulation of metastatic potential by cell surface urokinase of murine melanoma cells. *Cancer Res* 48, 1270-8.
- 34 Bastholm, L., Nielsen, M.H., De Mey, J., Dano, K., Brunner, N., Hoyer-Hansen, G. et al. (1994) Confocal fluorescence microscopy of urokinase plasminogen activator receptor and cathepsin d in human mda-mb-231 breast cancer cells migrating in reconstituted basement membrane. *Biotech Histochem* 69, 61-7.
- 35 Hoyer-Hansen, G., Ploug, M., Behrendt, N., Ronne, E. and Dano, K. (1997) Cell-surface acceleration of urokinase-catalyzed receptor cleavage. *Eur J Biochem* 243, 21-6.
- 36 Solberg, H., Romer, J., Brunner, N., Holm, A., Sidenius, N., Dano, K. et al. (1994) A cleaved form of the receptor for urokinase-type plasminogen activator in invasive transplanted human and murine tumors. *Int J Cancer* 58, 877-81.
- 37 Ke, S.H., Coombs, G.S., Tachias, K., Corey, D.R. and Madison, E.L. (1997) Optimal subsite occupancy and design of a selective inhibitor of urokinase. *J Biol Chem* 272, 20456-62.

- 38 Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A. and Craik, C.S. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc Natl Acad Sci U S A* 97, 7754-9.
- 39 Shliom, O., Huang, M., Sachais, B., Kuo, A., Weisel, J.W., Nagaswami, C. et al. (2000) Novel interactions between urokinase and its receptor. *J Biol Chem* 275, 24304-12.
- 40 Ploug, M., Rahbek-Nielsen, H., Nielsen, P.F., Roepstorff, P. and Dano, K. (1998) Glycosylation profile of a recombinant urokinase-type plasminogen activator receptor expressed in chinese hamster ovary cells. *J Biol Chem* 273, 13933-43.
- 41 Ploug, M., Behrendt, N., Lober, D. and Dano, K. (1991) Protein structure and membrane anchorage of the cellular receptor for urokinase-type plasminogen activator. *Semin Thromb Hemost* 17, 183-93.
- 42 Behrendt, N., Ploug, M., Patthy, L., Houen, G., Blasi, F. and Dano, K. (1991) The ligand-binding domain of the cell surface receptor for urokinase- type plasminogen activator. *J Biol Chem* 266, 7842-7.
- 43 Bucher, P. and Bairoch, A. (1994) A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation. *Ismb* 2, 53-61.
- 44 Kieffer, B., Driscoll, P.C., Campbell, I.D., Willis, A.C., van der Merwe, P.A. and Davis, S.J. (1994) Three-dimensional solution structure of the extracellular region of the complement regulatory protein cd59, a new cell-surface protein domain related to snake venom neurotoxins. *Biochemistry* 33, 4471-82.



- 45 Fletcher, C.M., Harrison, R.A., Lachmann, P.J. and Neuhaus, D. (1994) Structure of a soluble, glycosylated form of the human complement regulatory protein cd59. *Structure* 2, 185-99.
- 46 Li, X., Bokman, A.M., Llinas, M., Smith, R.A. and Dobson, C.M. (1994) Solution structure of the kringle domain from urokinase-type plasminogen activator. *J Mol Biol* 235, 1548-59.
- 47 Spraggon, G., Phillips, C., Nowak, U.K., Ponting, C.P., Saunders, D., Dobson, C.M. et al. (1995) The crystal structure of the catalytic domain of human urokinase-type plasminogen activator. *Structure* 3, 681-91.
- 48 Nienaber, V.L., Davidson, D., Edalji, R., Giranda, V.L., Klinghofer, V., Henkin, J. et al. (2000) Structure-directed discovery of potent non-peptidic inhibitors of human urokinase that access a novel binding subsite. *Structure Fold Des* 8, 553-63.
- 49 Katz, B.A., Mackman, R., Luong, C., Radika, K., Martelli, A., Sprengeler, P.A. et al. (2000) Structural basis for selectivity of a small molecule, s1-binding, submicromolar inhibitor of urokinase-type plasminogen activator. *Chem Biol* 7, 299-312.
- 50 Otwinowski, Z. and Minor, W. The hkl program suite.
- 51 Otwinowski, Z. and Minor, W. (1996) Processing of x-ray diffraction data collected in oscillation mode. In: *Methods in enzymology vol. 276, macromolecular crystallography* (Carter, C.W., Jr. and Sweet, R.M., eds.), pp. 307-326, Academic Press.
- 52 Terwilliger, T.C. and Berendzen, J. (1999) Automated mad and mir structure solution. *Acta Crystallogr D Biol Crystallogr* 55, 849-61.
- 53 Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W. et al. (1998) Crystallography & nmr system: A new software suite

- for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-21.
- 54 La Fortelle, E. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement in the mir and mad methods. In: *Methods in enzymology: Macromolecular crystallography* (Carter, C.W. and Sweet, R.M., eds.), pp. 472-494, Academic Press, San Diego.
- 55 Bailey, S. (1994) The ccp4 suite - programs for protein crystallography. *Acta Crystallogr D-Biol Cryst* 50, 760-763.
- 56 Kissinger, C.R., Gehlhaar, D.K. and Fogel, D.B. (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* 55, 484-91.
- 57 La Fortelle, E., de I., J. J. and Bricogne, G. (1997) Sharp: A maximum-likelihood heavy-atom parameter refinement and phasing program for the mir and mad methods. In: *Crystallographic computing 7* (Watenpaugh, P.B.a.K., ed.).
- 58 Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial f-1 atpase. *Acta Crystallogr D-Biol Cryst* 52, 30-42.
- 59 Nieh, Y.P. and Zhang, K.Y. (1999) A two-dimensional histogram-matching method for protein phase refinement and extension. *Acta Crystallogr D Biol Crystallogr* 55, 1893-900.
- 60 Kraulis, P.J. (1991) Molscript - a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24, 946-950.
- 61 Lawrence, M. and Bourke, P. (2000) *Conscript.*, Biomolecular Research Institute, Parkville, Victoria, Australia.



# For reference

Not to be taken  
from the room.

628393



3 1378 00628 3934

