Structure-assisted discovery of small molecules for targeting and transport across the blood-brain barrier

by J R Exequiel Timbol Pineda

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Chemistry

Montana State University

© Copyright by J R Exequiel Timbol Pineda (2002)

Abstract:

The existence of a blood-brain barrier (BBB) prevents more than 95% of therapeutic drugs from entering the brain. Several strategies are being utilized to target drugs to the Central Nervous System (CNS). Among the elegant approaches being pushed forward is the utilization of endogenous BBB transport system present in brain capillary endothelial cells. The transferrin receptor (TfR) is highly expressed in brain microvascular endothelial cells and is know to transcytose this layer of cells to facilitate the uptake the Fe3+-loaded transferrin (Tf). Furthermore, some antibodies against TfR have been shown to piggyback in the process of transporting Tf across the BBB. The structure of TfR has finally been solved in 1999 and this opened up a new possibility that can be explored for BBB drug delivery. A small molecule that specifically binds inside a vestigial active site of TfR can serve as a generic tag for molecules that need to cross the BBB, as long as it does not perturb the physiological function of TfR. In this project, several virtual screening programs were used to scan the Available Chemicals Database and other virtual chemical libraries for lead compounds. The gene for the ectodomain of the human transferrin receptor was cloned into a pFastBac plasmid. Recombinant protein was overexpressed in SF9 cells using a bacculovirus vector. Isothermal titration microcalorimetry was used to detect binding of the top scoring compounds from virtual screening. Hanging drops had been set up and good quality crystals were obtained for X-ray diffraction studies of the protein-ligand complexes.

STRUCTURE-ASSITED DISCOVERY OF SMALL MOLECULES FOR TARGETING AND TRANSPORT ACROSS THE BLOOD-BRAIN BARRIER

by

J.R. Exequiel Timbol Pineda

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

August 2002

N378
P653

APPROVAL

of a thesis submitted by

J.R. Exequiel Timbol Pineda

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. C. Martin Lawrence _____ 8/30/2002
(signature) Date

Approved for the Department of Chemistry and Biochemistry

Dr. Paul A. Grieco _____ 8-30-02
(signature) Date

Approved for the College of Graduate Studies

Dr. Bruce McLeod _____ 9-4-02
(signature) Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Ezequiel T Pineda

Date 8-30-2002

TABLE OF CONTENTS

1. INTRODUCTION.....1

What is the Blood Brain Barrier?.....1
Current Status of Neuropharmaceuticals.....3
Ways of Solving the BBB Drug Delivery Problem.....5
Invasive Strategies.....5
Non-invasive Strategies.....6
Drug Lipidization.....6
Utilize Endogenous BBB-Transport Systems.....6
Transferrin Receptors in the BBB.....8
OX26 Monoclonal Antibody-Mediated Delivery of Peptides Across the BBB.....10
Crystal Structure of the Ectodomain of Human TfR.....12

2. DESIGN OF TRANSFER VECTOR.....18

Design of Baculovirus Transfer Vectors.....20

3. GENERATION OF RECOMBINANT BACULOVIRUS, OVEREXPRESSION OF TFR FROM SF9 CELLS AND CRYSTALLIZATION OF TFR.....28

MOI Optimization and Small Scale Protein Expression.....31
Large Scale Protein Expression and Purification.....33
Crystallization of TfR.....34

4. VIRTUAL SCREENING OF THREE-DIMENSIONAL COMPOUND LIBRARIES.....36

Preparation of TfR and Ligand Coordinates for Docking.....39
Docking Virtual Chemical Database onto TfR.....48

5. ISOTHERMAL TITRATION CALORIMETRY.....51

Validation of Virtual Screening Results by ITC.....53

6. CONCLUSION AND FUTURE DIRECTIONS.....61

Conclusion.....61
Future Directions.....61

LITERATURE CITED.....63

LIST OF TABLES

Table Page

1. Ligation protocol to insert the gp67 secretion signal sequence and 6x His tag into pFBΔBH1. Volumes of each component are in microliter. dd-doubly digested.....22

2. Ligation protocol to insert TfR cDNA into pFBΔBH1-gp67.....24

3. Comparison of the number of transformants grown in Amp-containing agar plates.....25

4. Cell density and viability at each time point during the MOI optimization.....33

5. List of residues included in the active site definition.....40

LIST OF FIGURES

Figure Page

1. The anatomy of the Blood Brain Barrier.....2

2. Northern Blot of samples from different tissues. Lanes; (1) C6 rat glioma cells, (2) rat brain capillaries, (3) total rat brain, (4) rat heart, (5) rat kidney, (6) rat lung, and (7) rat liver, respectively.....8

3. Receptor-mediated uptake of diferric transferrin.....9

4. Receptor-mediated transport of diferric transferrin across the BBB endothelial layer.....11

5. Structure of the ectodomain of human transferrin receptor solved by x-ray crystallography.....13

6. Structural comparison of the protease-like domain of human TfR and an aminopeptidase.....14

7. Surface rendered image of the ectodomain of human TfR.....15

8. Digestion of pFBΔBH1 with (1) Not1, (2) EcoR1 and (3) BamH1. Lane 4 shows a 1kb ladder.....21

9. Lanes 1,2,3,5 and 6 are all clones that were isolated. Lane 4 contains the 100 bp ladder and lane 7 contains a negative control.....23

10. Triply digested plasmid DNA from isolated clones.....25

11. Lanes 1, 2, 6, 7, 8 are clones that the contain the right insert. Lane 3 is an aliquot of the undigested PCR product, that's why it is heavier. Lane 4, is some doubly digested PCR product and lane 5 is doubly digested pFBΔBH1-gp67-TfR$^{121-760}$.....27

12. Agar-plate with transformed DH10Bac *E.coli* cells (cell dilution 10^{-3}) showing blue and white colonies.....28

13. Agarose gel of PCR amplified portions of the recombinant bacmid. Lanes 1 and 2 -- pFBΔBH1-gp67-TfR$^{117-760}$; lanes 4 and 5 – pFBΔBH1-gp67-TfR$^{121-760}$; lane 3 -- mixture of 100 bp and 1 kbp MW ladders.....30

Figure Page

14. Western blot confirming the presence of His-tagged TfR.....31

15. Western blot showing the relative amounts and quality of TfR at each time point corresponding to different MOI.....33

16. Crystals of TfR observed three weeks after the drops were set up.....35

17. Residues lining the vestigial active site (depicted as red sticks).....41

18. A set of 100 spheres whose internal distances were used to find compounds that fit into the active site.....42

19. Scoring grids were calculated for the part of the receptor that is bounded by the box.....43

20. Structures of EGAETA, EDTA, and PTC.....54

21. Binding isotherm of EGAETA when titrated into a TfR solution.....55

22. Binding isotherm of EDTA when titrated into a TfR solution.....56

23. Binding isotherm of PTC when titrated into a TfR solution.....57

24. Predicted binding modes EDTA, EGAETA, and PTC superimposed in the active site of TfR.....58

25. Binding isotherm of PTC when titrated into a solution of preformed (TfR-diferric$Tf)_2$.....60

ABSTRACT

The existence of a blood-brain barrier (BBB) prevents more than 95% of therapeutic drugs from entering the brain. Several strategies are being utilized to target drugs to the Central Nervous System (CNS). Among the elegant approaches being pushed forward is the utilization of endogenous BBB transport system present in brain capillary endothelial cells. The transferrin receptor (TfR) is highly expressed in brain microvascular endothelial cells and is know to transcytose this layer of cells to facilitate the uptake the Fe^{3+}-loaded transferrin (Tf). Furthermore, some antibodies against TfR have been shown to piggyback in the process of transporting Tf across the BBB. The structure of TfR has finally been solved in 1999 and this opened up a new possibility that can be explored for BBB drug delivery. A small molecule that specifically binds inside a vestigial active site of TfR can serve as a generic tag for molecules that need to cross the BBB, as long as it does not perturb the physiological function of TfR. In this project, several virtual screening programs were used to scan the Available Chemicals Database and other virtual chemical libraries for lead compounds. The gene for the ectodomain of the human transferrin receptor was cloned into a pFastBac plasmid. Recombinant protein was overexpressed in SF9 cells using a bacculovirus vector. Isothermal titration microcalorimetry was used to detect binding of the top scoring compounds from virtual screening. Hanging drops had been set up and good quality crystals were obtained for X-ray diffraction studies of the protein-ligand complexes.

CHAPTER 1

Introduction

The number of incidents of diseases of the central nervous system dwarfs the combined mortality from cancer and heart disease. In the United States alone, there are an estimated 80 million people that have some disorder of the brain or spinal cord and who require neurotherapeutics. Based on the number of potential beneficiaries, the neuropharmaceutical area is the largest potential growth sector of the pharmaceutical industry. However, the presence of the so-called blood brain barrier (BBB), that excludes >95% of all drugs in the circulation from entering the brain, poses further challenge to that already difficult task of developing drugs. A dictum that became popular in brain-drug discovery is that a molecule must be lipid-soluble and it should have a molecular weight below the 500g/mole threshold to be able to diffuse freely across the BBB.

What is the Blood Brain Barrier?

The BBB, illustrated in Figure 1 (taken from 1) is a physical and metabolic barrier between the CNS and the systemic circulation, which serves to maintain equilibrium and protect the microenvironment of the brain. The anatomical site of the blood-brain barrier is the endothelial lining of the brain microvasculature. Although astrocyte foot processes invest more than 90% of the microvascular basement membrane, the luminal and antiluminal membrane of the endothelial cells form the main diffusion barrier for solutes crossing from the circulation to brain interstitial fluid. This thin membranous structure,

separated by about 0.3 μm of cytosol, significantly impedes the entry from the blood to brain of virtually all molecules, except those that are small and lipid-soluble. Brain capillaries stretch a length of approximately 400 miles, covering a surface area of approximately $20m^2$. The brain capillaries are approximately $40\mu m$ apart, and it takes around one second for a drug to diffuse $40\mu m$. This allows for almost instantaneous solute equilibration throughout the brain interstitial space once the endothelial barrier is overcome (1-2).

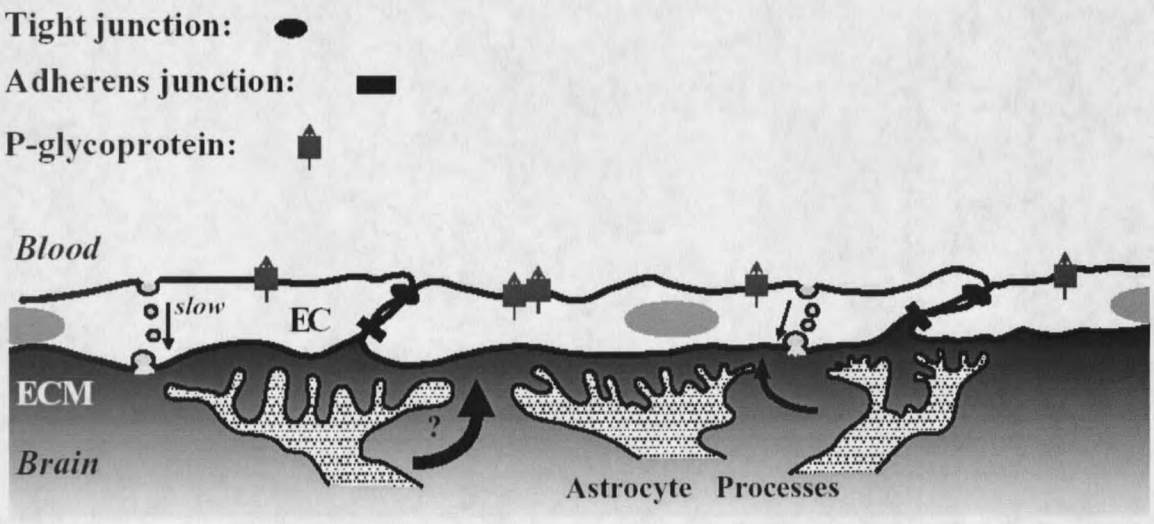


Figure 1. The anatomy of the Blood Brain Barrier

Endothelial cells in the brain differ from those in other peripheral organs in fundamental ways. Fewer endocytic vesicles are detectable by ultrastructural studies, hinting that there is reduced transcellular flux of free solute. Also, the cells are coupled by tight junctions, creating a rate-limiting barrier to paracellular diffusion of solutes between the endothelial cells. The tight junctions are the most apical elements of the junctional complex, which includes both tight and adherens junctions. Their presence is the predominant factor that results in high transendothelial electrical resistance (1500 to

2000 Ω*cm^2) of brain capillary endothelial membrane, a magnitude similar to that of epithelial membranes (1,3).

The high metabolic activity of cells in the central nervous system requires a constant supply of nutrients. There are sets of small and large molecules that can enter the brain via active transport. Membrane transporting proteins for glucose and certain amino acids are present in relatively high concentrations on brain capillaries. Systems capable of transporting macromolecules into the brain are also known, with some of them thought to be receptor-mediated. The best known of these is the transferrin receptor. Another important transporter present at relatively high concentration on brain capillaries is P-glycoprotein. It works in an opposite direction to those previously described in that it transports back into the blood a variety of lipophilic molecules that penetrate the brain or enter the endothelial cells. Circulating drugs and potential toxins concentrate at a higher-than-normal level in the brain of P-glycoprotein knock-out mice, indicating that this membrane protein is a functional component of the barrier (2).

Current Status of Neuropharmaceuticals

At present, the only diseases of the central nervous system that are treatable by small molecule drug therapy are the following:

- Obsessive-compulsive disorders
- Depressions
- Schizophrenia
- Epilepsy
- And Chronic pain

On the other hand, medicines for the following CNS patient groups are scant:

- Alzheimer's disease

- Brain and spinal cord trauma
- Huntington's disease and other neurodegenerative disorders
- Brain cancer
- Stroke
- AIDS
- Infection in the brain
- Genetic diseases that lead to metal retardation and premature death
- Ataxias (Inability to coordinate voluntary muscle movements), etc.

Clearly, only a few diseases of the central nervous system are amenable to small molecule drug therapy. This is, in part, because many promising small molecules are not able to cross the blood-brain barrier (BBB). But more than 98% of small molecules do not cross the blood brain barrier because they fail to satisfy both criteria of lipid-solubility and a molecular weight less than 500g/mole (4).

Trial-and-error methods employed in traditional CNS drug discovery invariably selected drugs that had appropriate drug-receptor interactions and transport properties. In contrast, modern receptor-based high throughout screening methods screen large collections of compounds for leads. But most drug leads that come out of receptor-based high throughput screening (HTS) programs lack the two criteria necessary for BBB transport. It's almost inconceivable that CNS drug discovery has evolved in the absence of a parallel maturation of an effective CNS drug delivery strategy. Indeed, knowing the presence of the BBB, it is quite odd that >99% of worldwide CNS drug development is devoted to CNS drug discovery with <1% of the effort devoted to CNS drug delivery (). Without an effective BBB drug delivery strategy, a HTS program for CNS drug development is destined for termination if it fails to identify compounds that satisfy the traditional standards for BBB permeability.

Ways of Solving the BBB Drug Delivery Problem

The currently proposed solutions to the BBB drug delivery problem can be categorized as (a) invasive or (b) non-invasive strategies. Invasive strategies attempt to either circumvent the BBB by neurosurgical approaches or induce transient disruption of the BBB by osmotic stress. Non-invasive approaches alter the molecular properties of the drug so that it will cross the BBB, with either pharmacologic or physiological strategies.

Invasive Strategies

Intra-cerebroventricular (ICV) infusion or intracerebral implants are the material science solution to the BBB problem in brain drug delivery. Controlled release of drugs in the brain is achieved by implanting a degradable polymer that encapsulates the drug or genetically engineered cells that are designed to secrete the required factor. With ICV, the drug is directly infused into the ventricular compartment with the goal of distributing the drug, mainly by diffusion, throughout the brain parenchyma. Sole reliance on diffusion from the local depot site is the principal problem of both ICV infusion and intracerebral implants because diffusion is a poor mode of drug delivery to the brain. In the case of ICV infusion, most of the drug is rapidly exported from the ventricles to the bloodstream. The drug that does enter the brain is confined only to the ipsilateral ependymal surface of the brain, because of the logarithmic fall in brain-drug concentration relative to the distance (mm) from the surface of the brain. Similarly, the effective diffusion distance of the drug from an intra-cerebral implant is only 1-2 mm. The other huge problem is the cost of the procedure, every ICV implant costs about

US$15,000, and there is no guarantee that implantation to a single locus will result to a complete cure (6).

Another strategy, that is also quite invasive, is transient BBB disruption by injecting a hypertonic solution, i.e. 1.7 M mannitol, or a vasoactive agent, i.e. bradykinin, directly into the carotid artery. The BBB cells act as osmometers, with the cell volume changing inversely as the osmolality of the environment. Cells shrink in a hypertonic environment as the intracellular water moves out. The cells of reduced volume must exert a tension at the tight junctions, thus creating openings, which temporarily modify the permeability of the BBB (7). However, this transient disruption also allows the entry of other molecules that are normally excluded by the healthy BBB, which could potentially cause more harm to the already compromised health of the patient.

Non-invasive Strategies

Drug Lipidization. A pharmacologic approach involve the 'lipidization' of drugs, whereby water-soluble molecules are conjugated to lipid carriers, such as free fatty acids, or replacement of highly polar functional groups with suitable bioisosteres. But although lipidization increases barrier permeability, it also increases drug uptake into the peripheral tissues.

Utilize Endogenous BBB-Transport Systems. An alternative to the invasive and pharmacologic approaches outlined above that is not dependent on lipid solubility and a molecular weight threshold is to reformulate the drug such that it can access the endogenous transport systems that exist in brain capillary endothelial membranes. Three

different classes of endogenous transport systems exist within the BBB. These are: (a) carrier-mediated transport systems (CMT), (b) receptor-mediated transcytosis (RMT) systems, and (c) active efflux transporters (AETs). Carrier-mediated transport systems include the glucose and amino acid carriers that allow bi-directional movement of small molecule nutrients and vitamins between the blood and the brain. The timescale of transport through carrier-mediated transport systems occur on the order of milliseconds. Among the receptor-mediated transcytosis systems are the BBB insulin receptor and the transferrin receptor. These systems mediate the bi-directional movement of large molecules between the blood and the brain and the process is completed within minutes. Active efflux transporters, such as P-glycoprotein, mediate the transport of small molecules from the brain to the blood (8).

The strict steric requirements of most CMT systems make them less attractive for use as delivery targets because a drug conjugated to an endogenous substrate will probably be exclude from the binding sites. Also, most drugs cannot be structurally altered to take on a molecular structure that mimics the endogenous nutrient, and thus most drugs will not have access to the carrier-mediated transport systems within the BBB. An alternative approach is to use the receptor-mediated transcytosis systems within the brain capillary endothelial plasma membrane. RMT is comprised of three steps: (a) receptor-mediated endocytosis at the luminal membrane of the capillary endothelial cells; (b) movement through the 300 nm of endothelial cytoplasm; and (c) exocytosis across the abluminal endothelial membrane into the brain interstitial fluid.

Transferrin Receptors in the BBB

It was demonstrated that monoclonal antibodies against rat and human transferrin receptors (TfR) preferentially label blood capillaries within the brain. Furthermore, in rats, labeling occurs after injection of antibody into the blood, implying that the receptors are accessible at the endothelial surface. From these results, it was suggested that the receptors might be expressed on these cells to allow transport of transferrin into brain tissues (9).

Recent results from a BBB genomics initiative further highlighted the abundance of TfR in rat brain capillary endothelial cells. The Northern blot analysis shown in Figure 2 (taken from 10) demonstrates expression of the 6.6kb transcripts in isolated rat brain capillaries. The mRNA for the full-length rat TfR contains 3413 bases. A recent study have also identified a second form of the transferrin receptor encoded by mRNA of 2.9 and 2.5 kb and is specific for liver (10).

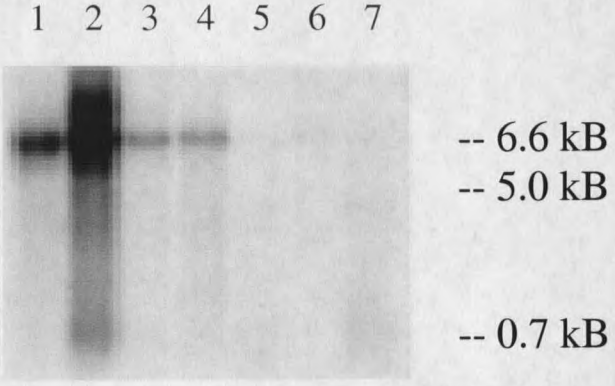


Figure 2. Northern Blot of samples from different tissues. Lanes; (1) C6 rat glioma cells, (2) rat brain capillaries, (3) total rat brain, (4) rat heart, (5) rat kidney, (6) rat lung, and (7) rat liver, respectively

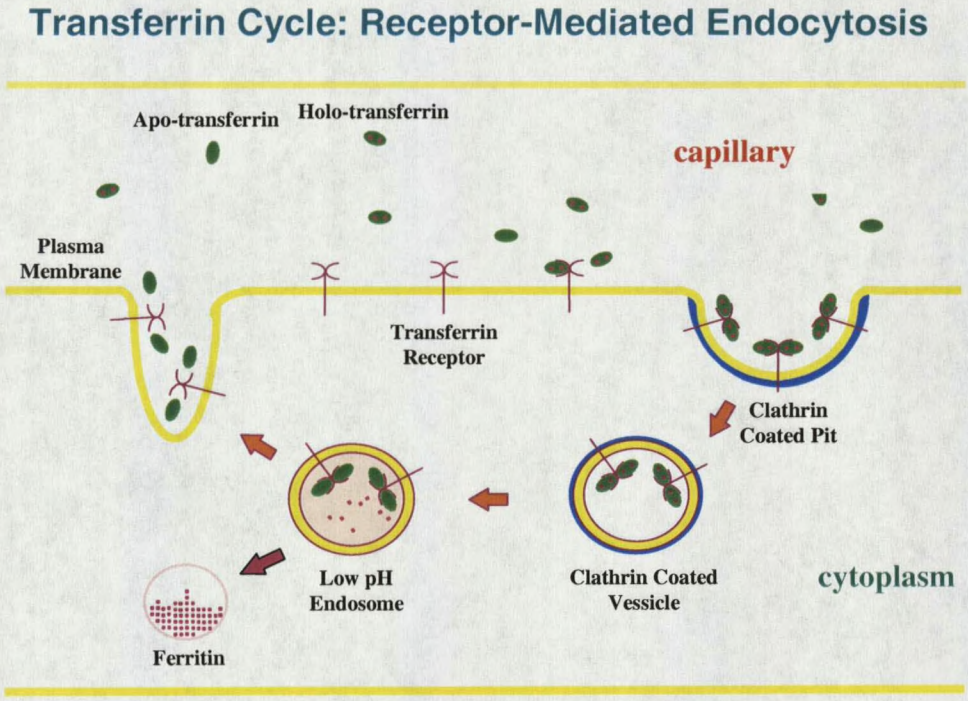


Figure 3. Receptor-mediated uptake of diferric transferrin

The mechanism of iron uptake of most vertebrate cells that require iron is depicted above. Extracellular diferric transferrin binds to the membrane-bound transferrin receptor. The complex is internalized via receptor-mediated endocytosis into an endosome. Acidification of the endosome causes iron to dissociate from transferrin, but apo-Tf remains bound to the receptor. The complex returns to the surface where apo-Tf is released and the receptor becomes free to bind another molecule of diferric transferrin from the serum (11).

OX26 Monoclonal Antibody-Mediated Delivery of Peptides Across the BBB

In one of the initial demonstrations of RMT of Tf across the BBB, radiolabeled Tf was perfused into rat brain. At first, the radiolabel accumulated in the endothelial cells of the BBB followed by a decrease of this intra-endothelial radioactivity associated with a concomitant increase in the nonvascular parts of the brain (12). Years later, an independent study by Skarlatos, et al validated these results. In these later experiments the transport of rat holo-transferrin was evaluated by in situ brain perfusion with the radiolabeled protein. Unlike before, the brain was postperfused with saline prior to decapitation of the rat. They again demonstrated a rapid and extensive distribution of the blood-borne Tf throughout the brain parenchyma. Furthermore the presence of competing plasma Tf in the perfusate reduced the amount radiolabeled Tf detected in the brain (13).

Broadwell, et al, carried out an example of an attempt to target a large biomolecule to the brain. They conjugated diferric-Tf to horseradish peroxidase, which by itself does not cross the BBB. The diferric-Tf-HRP conjugate labeled BBB vessels throughout the CNS without discernible disruption of the BBB or extravasation of the blood-borne probes into the brain parenchyma. After less than one hour, peroxidase activity was detected in brain compartments distal to the endothelial space. (14) These examples clearly illustrate that diferric-Tf is delivered to the brain tissues via receptor mediated transcytosis. The ensuing model is illustrated below.

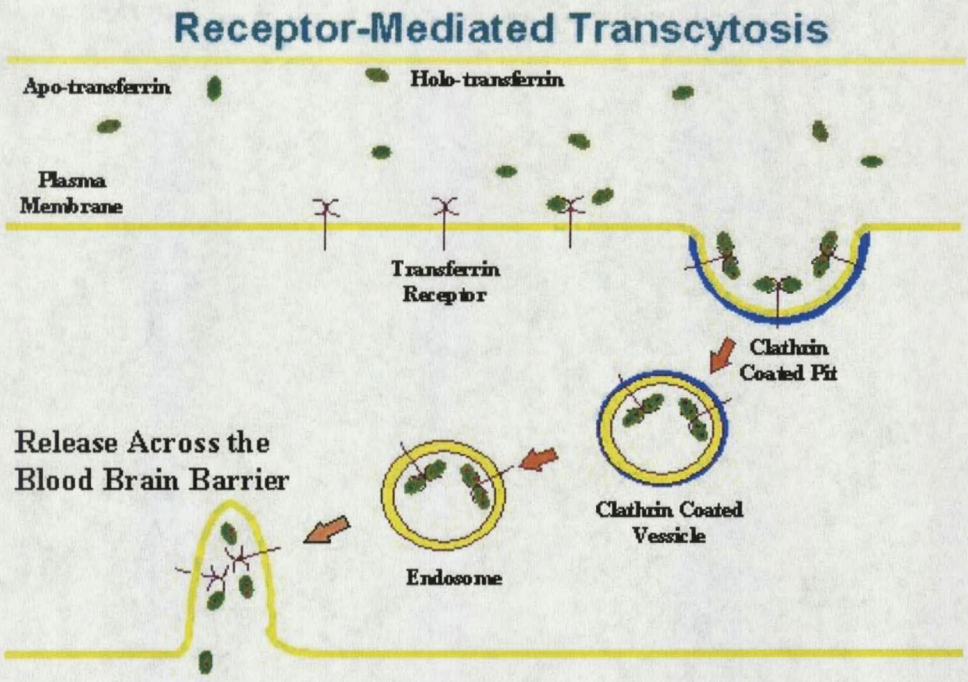


Figure 4. Receptor-mediated transport of diferric transferrin across the BBB endothelial layer

Although diferric-Tf was shown to transcytose better than the OX26 monoclonal antibody against rat TfR, it is still unwise to use it to deliver therapeutic molecules across the BBB. The main impediment to its use is the high level of endogenous diferric-Tf in the serum. A Tf-drug conjugate will have to compete with this pool to gain access to the RMT system in the BBB. To target molecules to BBB TfR the OX26 monoclonal antibody, or its Fab, is still the wiser choice. OX26 binds to an exofacial epitope on rat TfR that is removed from the Tf binding site. Therefore, in rat models of CNS diseases, conjugating this antibody to a therapeutic molecule was shown to be a promising approach to BBB drug delivery with minimal perturbation on the normal receptor function. In fact, peptide-based neuropharmaceuticals conjugated to such monoclonal

antibodies against unique cell surface receptors in the human BBB are subjects of intense studies and are very well documented (15)

Crystal Structure of the Ectodomain of Human TfR

Lawrence, et al in 1999, solved the crystal structure of the extracellular domain of human TfR, corresponding to residues 121 to 760. This fragment is equivalent to that released when TfR-containing membranes are treated with trypsin. The soluble receptor is a dimer and it binds two transferrin molecules. $TfR^{121-760}$ is 28% identical at the amino acid level to membrane glutamate carboxypeptidase II (mGCP) whose substrate, N-acetyl-α-L-aspartyl-L-glutamate, is the most prevalent mammalian neuropeptide (16-17). TfR, however, lacks peptidase activity presumably because it lacks three of the supposed Zinc ligands in the predicted protease-like domain.

The TfR monomer has three distinct domains, and the dimer assumes a butterfly-like shape, as depicted in the Figure 5.

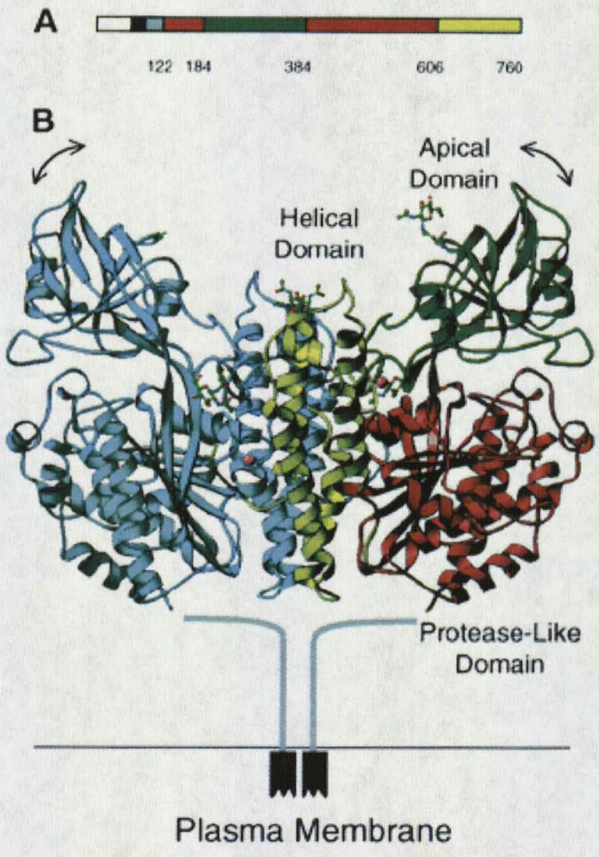


Figure 5. Structure of the ectodomain of human transferrin receptor solved by x-ray crystallography

The domain boundaries in one monomer are color-coded in for clarity, with the protease-like domain in read, helical domain in yellow and the apical domain in green. The protease-like domain has a fold that is closely related to that of carboxy- and aminopeptidase. Its central, seven-stranded, mixed β sheet is flanked by α helices. Eight β strands are present in carboxypeptidase itself, but in TfR the polypeptide chain traces a path away from the outside edge of the β sheet forming an extended loop (see Figure 6).

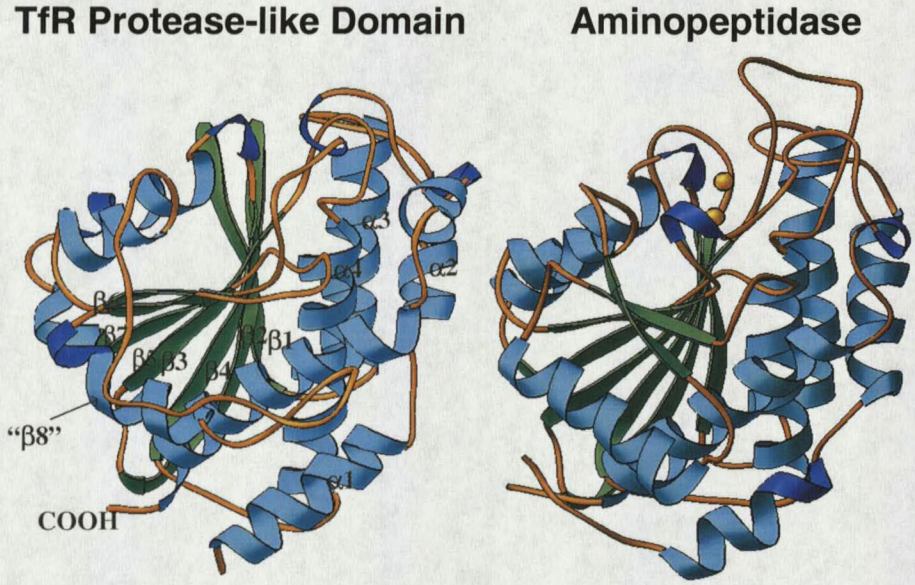


Figure 6. Structural comparison of the protease-like domain of human TfR and an aminopeptidase

The apical domain resembles a β sandwich in which the two sheets are splayed apart, and with a helix running along the open edge. The helical domain is responsible for dimerization. The helical domain of one monomer interacts with its counterpart across the molecular twofold axis, contacting each of the three domains. A significant sequence identity between corresponding domains in TfR and mGCP is revealed by a structure-based alignment. The protease-like, apical, and helical domains of TfR share 30.3%, 30.2%, and 24% identity, respectively, with those of mGCP. The apical domain covers the catalytic site, but the position of the β8 loop allows access through an interdomain channel (11). The location of this channel that leads to the vestigial active site in TfR is more obvious in the surface rendered image in Figure 7. Together with the sequence

similarity, this structural feature suggests that TfR may have evolved from a cell-surface protease, similar to mGCP.

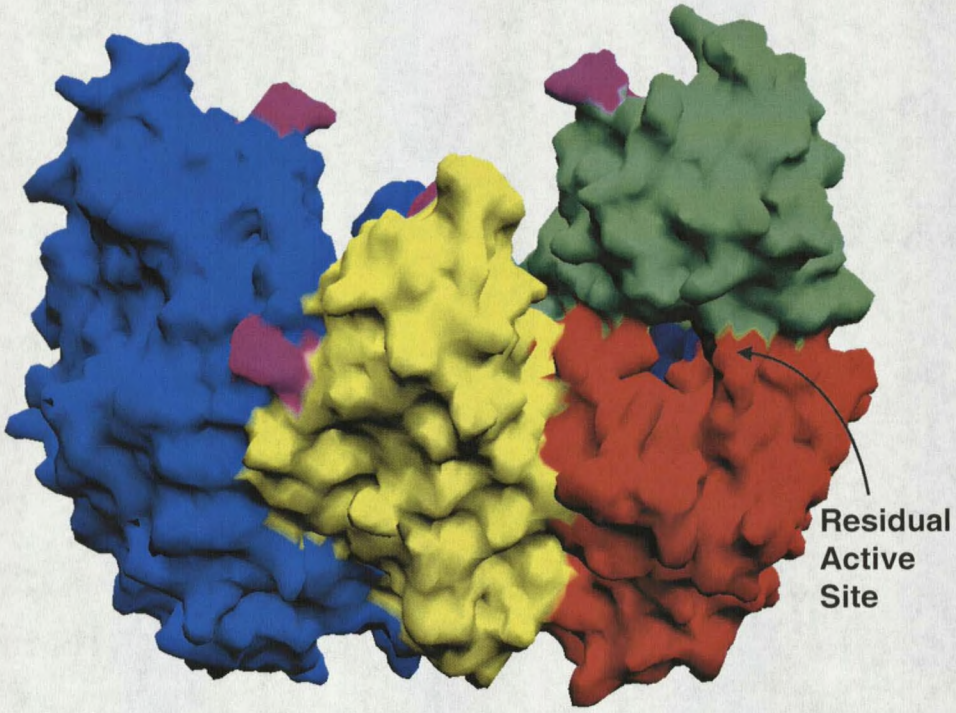


Figure 7. Surface rendered image of the ectodomain of human TfR

It is not clear if this vestigial active site has any functional role. However, its discovery has opened up a new possibility that can be explored for BBB drug delivery: a small molecule that specifically binds inside this vestigial "active site" pocket could serve as a molecular stamp for delivering drugs across the blood-brain barrier, provided that the chimeric molecule does not significantly perturb the physiological role of TfR.

Using the structural features of this pocket as constraints for virtual screening, the program DOCK (Version 4.0) was used to screen libraries of commercially available databases namely: Available Chemicals Database, Tripos and Chembridge. DOCK was

quick but not as good a predictor as more CPU intensive programs, such a FLO. However, the more CPU intensive programs are impractical, with our current computational resources, for screening such libraries that contain about 300,000 compounds. A good compromise is to screen, or filter, the entire database with DOCK. Then the top ranking compounds of approximately 2,500 molecules were evaluated with the more CPU intensive program FLO, using the DOCK-generated conformations as input for energy minimizations in the context of the vestigial active site. The ligands were allowed to move freely while the active site residues were restrained during the minimization. After a detailed examination of the calculated total binding energies and the corresponding component energies, together with visual inspection of the energy minimized conformation of the active site-ligand complexes, we selected a subset of 27 compounds for further in vitro screening.

While these virtual experiments were underway, we overexpressed the ectodomain of human TfR in a lytic insect cell/baculovirus expression system. Some of the protein was used to set up hanging drops for crystallization. The bulk of the isolated protein was used to assay for binding of the selected compounds using isothermal titration calorimerty. Ligands that were found to bind were diffused into TfR crystals for determination of the structure of the complex by x-ray diffraction. It is hoped that with the structure of the protein-ligand complex, we will be able to: (1) determine if the ligands bind to the site where the docking experiments were directed, (2) identify TfR residues that interact directly with the ligand, as well as ligand functional groups that contact the receptor and (3) determine if there are any significant movements on the receptor that are promote binding (induced-fit). These informations will allow us to

propose modifications in the ligand structures that will improve their affinity for TfR, and therefore start the cycle of design and optimization of the leads.

CHAPTER 2

Design of Transfer Vector

Bacterial expression systems are the most established and probably the oldest protein expression systems utilized, however they have an inherent limitation in that they do not have the necessary machinery to process proteins that require post-translational modifications to achieve their functional status. This is true for most eukaryotic proteins expressed on the cell surface. As a result, alternative protein production systems that more closely mimic the complex post-translational modification processes of mammalian cells are being sought for production of complex proteins (18). Insect cells perform post-translational modifications, such as glycosylation, phosphorylation, palmitoylation, myristylation and addition of glycosyl-phosphatidylinositol anchors (19). Glycosylation is the most extensive and necessary modification as it is important in secretion, antigenicity and clearance of glycoproteins (20).

Baculovirus expression vectors provide a versatile and reliable system for the production of recombinant protein in insect cells. One of the most widely used vectors is the *Autographa californica* multiple capsid nucleopolyhedrovirus (AcMNPV), a prototype member of *Baculoviridae*. Several cell lines can be used to propagate this virus; among them are *Estigmene acrea*, *Mamestra brassicae*, *Trichoplusia ni* (Tn) and *Spodoptera frugiperda* (Sf) (21). Previous work have been carried out to examine the extent of N-glycosylation of secreted placental alkaline phosphatase produced by recombinant baculovirus-infected *Trichoplusia ni* and *Spodoptera frugiperda* cell lines.

For this particular protein, it was concluded that *Spodoptera frugiperda* cells produced more fucosylated oligosaccharides than either of the *Trichoplusia ni* cell lines (22).

The early and late genes in the life cycle of AcMNPV are largely concerned with the production of virus particles, which bud from the cell to spread infection to new cells. In contrast, the very late genes, those encoding polyhedrin and p10, are required for the production of occlusion bodies that contain virus particles, in the nucleus of the host cell. Both of these genes are under the control of strong promoters, but can be deleted from the virus genome without affecting the production of infectious virus particles. This is exploited by inserting foreign coding genes in lieu of the polyhedrin and p10 sequences to derive expression vectors (21).

AcMNPV has a large (130-kb), circular, double-stranded DNA genome with multiple recognition sites for many restriction endonucleases. Therefore, it is not wise to construct the recombinant bacmid by traditional "cut and paste" approach. A way around this is to clone the gene of interest in a transfer vector containing a baculovirus promoter flanked by baculovirus DNA derived from a nonessential locus, for example the polyhedrin gene. The gene of interest is inserted into the genome of the parent virus by homologous recombination after transfection in insect cells. This strategy however, requires successive plaque purification to ensure isolation of the pure clone of recombinant virus. A quicker strategy uses site-specific transposition with Tn7 to insert the heterologous gene into bacmid DNA propagated in E. coli. This technique still requires cloning the desired gene into a transfer vector, but the transposition event is carried out in bacterial host cells. Antibiotic selection and blue/white screening identified colonies that contain the recombinant bacmid, because the transposition results in

disruption of the lacZα gene. High molecular weight DNA was prepared from selected clones and this DNA was used to transfect insect cells. Having prepared the DNA from a pure clone, this strategy avoids successive plaque purification to obtain recombinant virus (23).

Design of Baculovirus Transfer Vectors

TfR is a type II integral membrane protein. The ectodomain, composed of residues 121 to 760, contains an intramolecular disulfide bond between Cys^{556} and Cys^{558} (11). Previous structural work utilized protein overexpressed in CHO cells. These cells were expensive to maintain, both in terms of time and money. We wanted to know if we could reproduce the clone produced in CHO cells as closely as possible in a different expression system. We decided to express the protein in a lytic baculovirus/insect cell expression system. The pFASTBAC1 plasmid of the Bac-to-Bac baculovirus expression system was available in the laboratory. However, this transfer vector does not contain a secretion signal sequence and a 6x His tag. A plasmid for a different system, pAcGP67TfR of the BaculoGold system, that contains the secretion signal, 6x His tag, Factor Xa cleavage site and TfR residues 121-760, respectively, was obtained as a generous gift from Tony Gianentti and Prof. Pamela Bjorkman (CalTech). This construct contains a BamH1 endonuclease site between the secretion signal and the 6x His tag. pFASTBAC1 has a BamH1 site in its multiple cloning site.

To avoid having multiple BamH1 cleavage sites in our final construct, we decided to delete this site in pFASTBAC1. 10 μg of the plasmid was digested with 20 units of BamH1. The linearized vector was purified by agarose gel electrophoresis, sticky ends repaired by treating with 10 units of Klenow to generate blunt ends, heated at 60°C to inactivate Klenow and ligated overnight using T4 DNA ligase. MC1061 competent cells were transformed with this ligated DNA and selection was done by streaking several dilutions of the transformed cells on gentamycin-containing agarose plates. A single colony was picked for miniprep and the isolated plasmid was digested with several endonucleases that recognize the multiple cloning site. The difference in hydrodynamic behavior of the linearized as opposed to the circular plasmid DNA indicated the absence of a BamH1 site in the purified plasmid. We refer to this plasmid as pFBΔBH1.

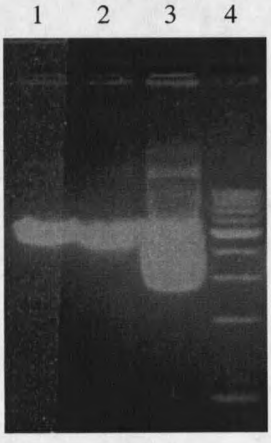


Figure 8. Digestion of pFBΔBH1 with (1) Not1, (2) EcoR1 and (3) BamH1. Lane 4 shows a 1kb ladder

Having pFBΔBH1, we are now ready to insert the DNA for the secretion signal sequence, 6x His and Factor Xa into pFBΔBH1. A BssHII restriction site before the

secretion signal sequence was introduced by PCR, using pAcGP67-TfR as template. The reverse primer spans the EcoR1 site between the 6x His tag and $TfR^{121-760}$. These are the sequences, in the 5' to 3' convention, of the primers used:

Forward: AAT GCG CGC ATG CTA CTA GTA AAT CAG

Reverse: CGC GAA TTC ACC ACG TCC CTC GAT

Due to incompatibility in optimum temperature, we sequentially digested both the plasmid (pFBΔBH1) and the PCR product with BssHII followed by EcoR1. For the plasmid it is important that the order indicated be followed because BssHII is not an efficient cutter when the site is close to the end of the DNA, while EcoR1 remains efficient.

After purifying the digested materials by agarose gel electrophoresis, ligation reactions were set up following the protocol below. The molar ratios are pFBΔBH1:insert. A 1:1 molar ratio contains approximately 26μg pFBΔBH1/μg insert. The concentrations of the starting materials are 250 ng/μl for pFBΔBH1and 4.5ng/μl for insert, as determined from O.D at 260 nm.

Molar Ratio	dd-pFBΔBH1	dd-insert	T4 buffer	T4 DNA ligase	Water
1:1	4	0.9	2	0.5	12.6
1:3	4	2.6	2	0.5	10.9
1:10	4	8.6	2	0.5	4.9
Vector only(1)	4	0	2	0.5	13.5
Vector only(2)	4	0	2	0	14
Insert only	0	2.6	2	0.5	14.9

Table 1. Ligation protocol to insert the gp67 secretion signal sequence and 6x His tag into pFBΔBH1. Volumes of each component are in microliter. dd-doubly digested

The ligation was allowed to proceed overnight at 4°C. MC1061 competent cells were transformed with the ligated products. A total of five clones were selected from different ampicillin plates, grown for twelve hours at 37°C in 2 ml Luria broth with ampicillin, and plasmid DNA was purified from each clone. To determine if the desired fragment was successfully inserted, 15 μl aliquot from each prep was digested with BssHII and EcoR1. An equivalent amount of pFBΔBH1 was treated the same way as a negative control. Figure 9 clearly shows that the insert is present in all clones isolated.

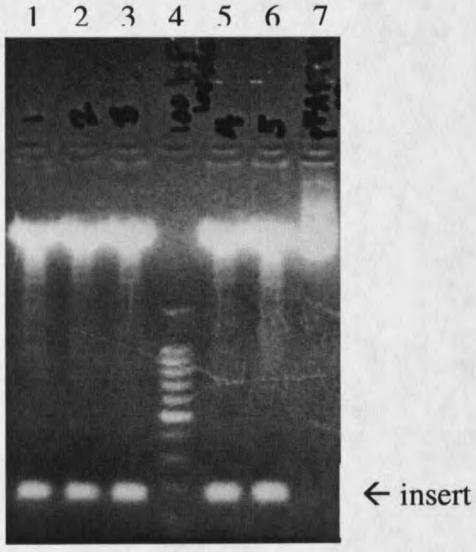


Figure 9: Lanes 1,2,3,5 and 6 are all clones that were isolated. Lane 4 contains the 100 bp ladder and lane 7 contains a negative control

With the secretion signal sequence, 6x His Tag and Factor Xa cleavage site in pFBΔBH1, now referred to as pFBΔBH1-gp67, we are now ready to incorporate the $TfR^{121-760}$ cDNA into the transfer vector. Although it was possible to have just PCR amplified the entire plasmid to do a single step incorporation of the insert into the new vector, we decided to take the longer path to avoid the risk of having PCR errors and the

need to sequence a 2 kb insert to validate our experiment. To excise the TfR-coding fragment from pAcGP67-TfR, 10 μg of the plasmid was digested with EcoR1 and Not1. pFBΔBH1-gp67 was also doubly digested with EcoR1 and Not1 to generate sticky ends complementary to those of the TfR-coding fragment from pAcGP67-TfR. The linearized vector, pFBΔBH1-gp67, and the TfR cDNA were both extracted by standard procedures from agarose gel-separated bands. Stock solutions were made with concentrations of 175 ng/μl for the TfR cDNA and 50 ng/μl for doubly digested pFBΔBH1-gp67. The ligation protocol was as follows:

Molar Ratio	dd-pFBΔBH1-gp67	TfR cDNA	T4 buffer	T4 DNA ligase	Water
1:1	4	0.5	2	0.5	13
1:3	4	1.5	2	0.5	12
1:10	4	5	2	0.5	8.5
Vector only(1)	4	0	2	0.5	13.5
Vector only(2)	4	0	2	0	14
Insert only	0	1.5	2	0.5	16

Table 2. Ligation protocol to insert TfR cDNA into pFBΔBH1-gp67.

These reactions were allowed to proceed overnight at 4°C. 100μl aliquots of freshly thawed MC1061 competent cells were transformed with 5 μl from one ligation condition and allowed to grow and express the selectable marker in 2 ml Luria broth. 100-, 1,000- and 10,000-fold dilutions were made for each and 150, 100 and 50 μl from the respective dilutions was streaked on ampicillin-containing agar plates. Colonies that grew in each of the 10,000-fold dilute samples were counted.

Description	Number of Colonies
1:1	4
1:3	4
1:10	2
Vector only(1)	0
Vector only(2)	1
Insert only	0

Table 3. Comparison of the number of transformants grown in Amp-containing agar plates.

Ten colonies were picked for further screening. Plasmid DNA was prepared from each colony and a triple digest with BssHII, EcoR1 and Not1 was carried out to determine the presence of the TfR-coding fragment and the fragment that codes for the gp67 secretion signal sequence, 6x His and Factor Xa cleavage site. The picture below shows the presence of all the expected fragments in seven out of ten clones that were picked. We refer to this construct as pFBΔBH1-gp67-TfR$^{121\text{-}760}$.

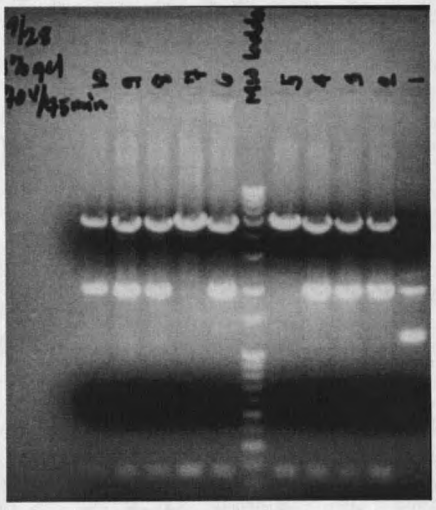


Figure 10. Triply digested plasmid DNA from isolated clones

Although we know that the complete insert we've made is identical to that of the source plasmid, pAcGP67-TfR, the protein they've isolated from this construct does not crystallize, at least not without another macromolecule. Removing the 6x His tag by Factor Xa digestion leaves three more residues (GFF) attached to the N-terminal of the ectodomain of TfR. We suspect that, aside from the difference in the host organism for overexpression, these extra residues drastically alter the crystallization conditions, because we know that the N-terminus is involved in crystal contacts.

The ectodomain of TfR, comprised of R121 to F760, is released when TfR-containing membranes are treated with trypsin. Trypsin cleaves between R120 and R121 but it is possible that the N-terminal residues may be important for recognition of this site. Thereafter, we decided to incorporate residue P117 to R120 into our construct, allowing Trypsin cleavage to produce TfR 121-760.

To do this we searched for an enzyme that cuts in the TfR-coding region but not anywhere else in the entire vector. We identified Nde1 as a single cutter that, together with EcoR1, will cut out a 98 base pair fragment from pFBΔBH1-gp67-TfR$^{121\text{-}760}$ that spans R121. This was replaced with a PCR amplified fragment from pCMVTfR, a plasmid used to overexpress TfR$^{117\text{-}760}$ in CHO cells that contained residues P117-R120. The primers used to amplify the fragment of interest from pCMVTfR are:

Forward primer: ATA GAA TTC CCT GCA GCA CGT CGC

Reverse primer: CGC ATC TTT TTG AGA TCC AGC CTC

We introduced an EcoR1 cleavage site before P117 in the forward primer. The reverse primer includes 33 bases more after the Nde1 site because Nde1 does not cut as well near the ends.

pFBΔBH1-gp67-TfR$^{121-760}$ and the PCR amplified fragment from pCMVTfR were doubly digested with EcoR1 and Nde1 and purified by standard techniques. A similar ligation protocol as those enumerated in table 1 was followed and five colonies were picked from the transformants. Plasmid DNA was purified from minipreps and digestion of small aliquots from each clone with EcoR1 and Nde1 confirmed the presence of the low molecular weight insert, as shown in Figure 11. This version was called pFBΔBH1-gp67-TfR$^{117-760}$. This plasmid, together with pFBΔBH1-gp67-TfR$^{121-760}$, was submitted for sequencing, and results confirmed the presence of the right insert.

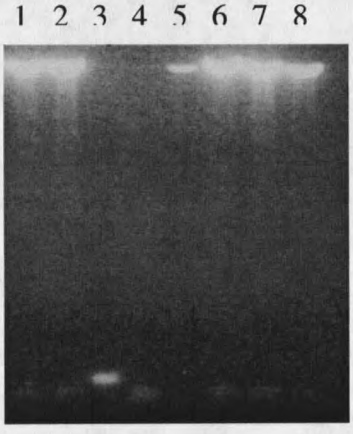


Figure 11: Lanes 1, 2, 6, 7, 8 are clones that the contain the right insert. Lane 3 is an aliquot of the undigested PCR product, that's why it is heavier. Lane 4, is some doubly digested PCR product and lane 5 is doubly digested pFBΔBH1-gp67-TfR$^{121-760}$.

CHAPTER 3

Generation of Recombinant Baculovirus and Overexpression of TfR from Sf9 Cells

Following the manufacturer's protocol, we transformed DH10Bac cells with pFBΔBH1-gp67-TfR$^{121-760}$ and pFBΔBH1-gp67-TfR$^{117-760}$. Cells that contain the recombinant bacmid are white in the presence of a chromogenic substrate, Bluo-gal, in contrast to blue colonies containing the wild type bacmid. White colonies were selected and restreaked for validation. One can clearly see in Figure 12 the blue colonies that contain the unaltered bacmid. The fewer white colonies the transposition of the gene of interest into the plasmid disrupted expression of the lacZα peptide. White colonies contain recombinant bacmids.

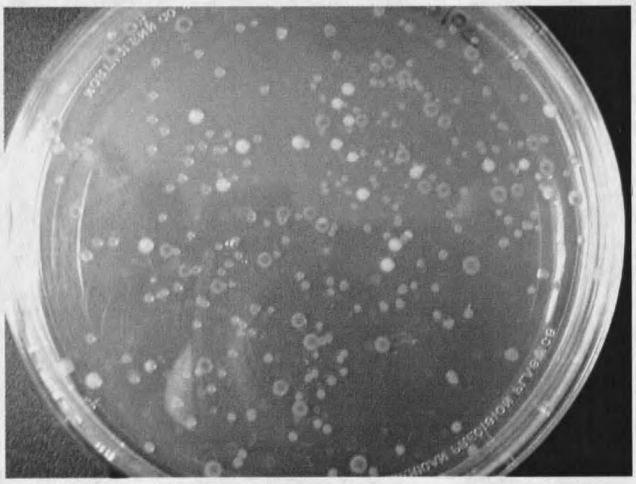


Figure 12 : Agar-plate with transformed DH10Bac *E.coli* cells (cell dilution 10^{-3}) showing blue and white colonies.

Using reagents from the Qiagen miniprep kit, high molecular weight bacmid DNA of each recombinant clone was isolated and the success of the site-specific transposition was confirmed by PCR. An agarose gel of the amplified fragments is shown in Figure 13. The smaller fragment is approximately 2 kbp. It is a portion of the recombinant bacmid that is amplified by a forward primer (5'-CGCCAGGGTTTTCCCAGTCACGAC-3') that anneals at a site upstream of the multiple cloning site of pFastBac1 and a reverse primer (5'-CGCATCTTTTTGAGATCCAGCCTC-3') that partly anneals at the region of TfR that codes residue E156 to D162. The larger fragment is approximately 3kbp. It contains the entire insert (gp67 secretion signal sequence-6xHis tag-Factor Xa site-TfR117-760), which is about 1900 bp, and another 1kb portion that is a part of pFastBac1 plasmid. The primers used to amplify the heavier fragment are:

Forward: 5'-AAT GCG CGC ATG CTA CTA GTA AAT CAG-3'

Reverse: 5'-AGC GGA TAA CAA TTT CAC ACA GGA-3'

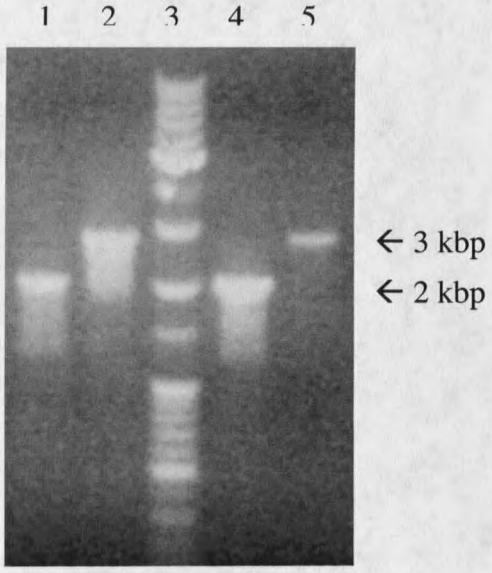


Figure 13: Agarose gel of PCR amplified portions of the recombinant bacmid. Lanes 1 and 2 -- pFBΔBH1-gp67-TfR$^{117-760}$; lanes 4 and 5 -- pFBΔBH1-gp67-TfR$^{121-760}$; lane 3 -- mixture of 100 bp and 1 kbp MW ladders.

To generate primary viral stocks of each construct, 5μL aliquots of the bacmid DNA was used to transfect adherent Sf9 cells, following manufacturer's protocol. After 72 hours of incubation at 27±1°C in an incubator, the supernatant was collected and clarified by centrifugation at 500 x g for five minutes. 45 μl of each clarified supernatant was loaded on an 8% SDS-PAGE gel. The bands were transferred to nitrocellulose for western blot analysis using mouse anti-6x His primary antibody and rabbit anti-mouse Ig antibody conjugated with alkaline phosphatase as secondary antibody. The blot for TfR$^{117-760}$ is shown as Figure 14 below, while that for TfR$^{121-760}$ did not show any sign of expression. At this point we decided to use TfR$^{117-760}$ for all succeeding work.

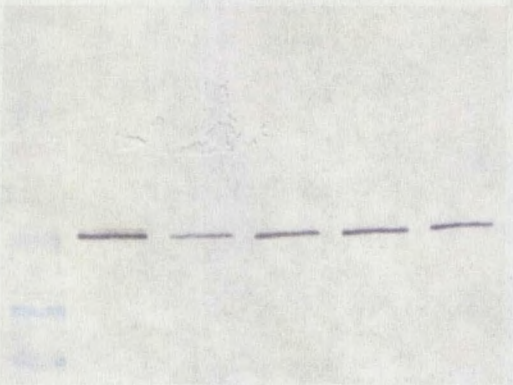


Figure 14. Western blot confirming the presence of His-tagged TfR.

Assuming an initial titer of $2x10^7$pfu/ml of the primary viral stock obtained above, a 50 ml suspension culture of Sf9 cells at mid log phase was infected at a multiplicity of infection (MOI) of 0.1. The amount of virus required was calculated as follows:

$$\text{Inoculum required (ml)} = \frac{\text{MOI (pfu/cell)x(total number of cells)}}{\text{Viral titer (pfu/ml)}}$$

The culture was maintained in the incubator at 27±1°C with stirring rate between 80 to 100 rpm. This secondary viral stock was harvested after 48 hours by centrifugation at 500x g for five minutes. A titer between 1 to $1.5x10^8$ pfu/ml was typically obtained at this step and this stock was used for large-scale viral amplification (>1L).

MOI Optimization and
Small Scale Protein Expression

To determine the optimum MOI for protein expression and the time required to obtain the best quality of protein we carried out a time course experiment at four different MOI: 1, 2, 5, 10. Four 100 ml spinner flasks containing 50 ml suspension culture with

densities of $1.58x10^6$ cells/ml were infected with a predetermined volume of virus. The volumes were equalized by adding fresh medium containing 1X penicillin and streptomycin such that the final volume in each spinner flask was 60 ml. 2 ml aliquots were taken out at each time point and the viability of the culture was determined by dividing the number of viable cells by the total number of cells. Viable cells exclude the staining dye Trypan blue while dead cells are completely stained. The aliquots were centrifuged for 10 minutes at 12,000x g and the supernatants were frozen.

When all required time points were obtained, all frozen aliquots were thawed and the quality and quantity of TfR at each MOI were compared by western blot. A rabbit polyclonal antibody raised against the ectodomain of human TfR was used as primary antibody. The bands were visualized by reacting the primary blot with mouse anti-rabbit Ig antibody conjugated with alkaline phosphatase. Figure 15 shows western blots of all time points and Table 5 shows the corresponding cell counts at each time point .

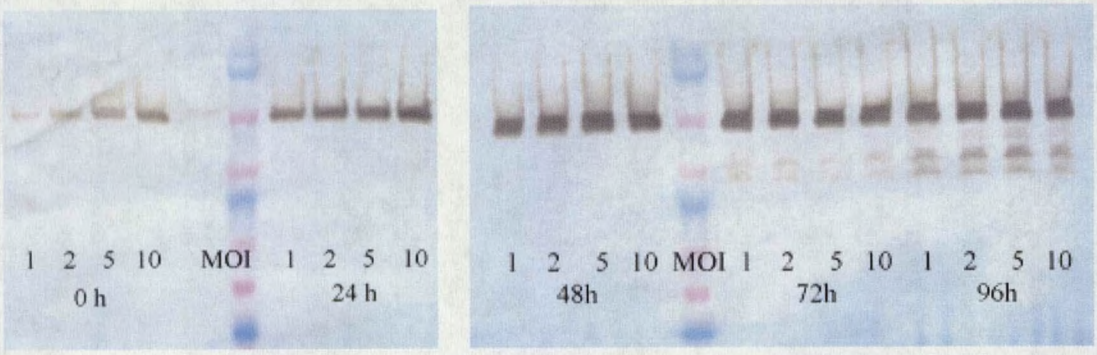


Figure 15. Western blot showing the relative amounts and quality of TfR at each time point corresponding to different MOI.

		MOI = 1	MOI = 2	MOI = 5	MOI = 10
0 hour	Viable Cells	$1.23x10^6$ cells/ml			
	Total Cells	$1.32x10^6$ cells/ml			
	Viability in %	93%			
24 hours	Viable Cells	$1.27x10^6$	$1.24x10^6$	$1.09x10^6$	$1.00x10^6$
	Total Cells	$1.41x10^6$	$1.46x10^6$	$1.31x10^6$	$1.25x10^6$
	Viability in %	90%	85%	83%	80%
48 hours	Viable Cells	$1.17x10^6$	$1.07x10^6$	$1.04x10^6$	$0.84x10^6$
	Total Cells	$1.46x10^6$	$1.35x10^6$	$1.65x10^6$	$1.46x10^6$
	Viability in %	80.1%	79%	63%	57.5%
72 hours	Viable Cells	$0.83x10^6$	$0.76x10^6$	$0.71x10^6$	$0.46x10^6$
	Total Cells	$1.54 x10^6$	$1.47 x10^6$	$1.62 x10^6$	$1.31 x10^6$
	Viability in %	54%	51.7%	43.8%	35%
96 hours	Viable Cells	$0.41x10^6$	$0.3x10^6$	$0.32x10^6$	$0.23x10^6$
	Total Cells	$1.6x10^6$	$1.44x10^6$	$1.49x10^6$	$1.47x10^6$
	Viability in %	25.6%	20.8%	21.4%	15.6%

Table 4. Cell density and viability at each time point during the MOI optimization.

A MOI between 2 to 10 yield about the same amount of protein. For later work we merely assumed that we had a titer of $1x10^8$ and infected our suspension cultures at an MOI of 7.5, based on the approximate viral titer.

Large Scale Protein Expression and Purification

For large-scale protein production we routinely maintain seven 3L spinner flasks, each flask containing suspension-adapted Sf9 cells grown to a density of $2x10^6$ cells/ml prior to infection. All seven 1L cultures were routinely infected in parallel with virus for every large scale protein purification that was done. The supernatant was collected typically after 48 to 60 hours by centrifugation at 7300x g for at least 50 minutes.

TfR was isolated by affinity chromatography on a Transferrin-sepharose column that was prepared by standard Cyanogen bromide chemistry (24). A total of 20 ml of the slurry was used for every large-scale protein purification that was carried out. After

binding, the resin was washed with at least 1L of running buffer (150 mM NaCl, 50 mM HEPES, pH 7.5). Then a low pH wash buffer consisting of 100mM KCl, 100mM NaCl, 50mM Citrate and 1mM Deferoxamine Mesylate, pH 5.0, was passed through the column to wash away the bound Fe^{3+} from the transferrin. The column is then washed with about 45 mL of running buffer to bring back the pH close to neutral before finally eluting TfR with a buffer containing 2M KCl and 50 mM HEPES at pH 7.5. The purification procedure effectively mimics the transferrin-TfR cycle. The typical yield was between 35-40 mg (~5mg/L). All TfR containing fractions were combined and stored in liquid nitrogen.

Crystallization of TfR

Freshly thawed TfR stocks were diluted with one volume of nanopure water to adjust the KCl concentration to 1M. Diluted samples were then concentrated to 1 to 2 mg/ml using Biomax-30 spin concentrators. 20µl of trypsin at 10mg/ml was added for every 4mg of TfR followed by incubation on ice for 1 hour. Passing the solution over a benzamidine sepharose column terminated the digestion. The resin was washed with at least 5 column volumes of running buffer consisting of 150mM NaCl, 50mM HEPES at pH 7.5. Phenylmethylsulfonyl fluoride was added to the flow-through to a final concentration of 1mM to inhibit any residual trypsin activity. Undigested TfR was sequestered from the bulk by passing the combined flowthrough from the previous step over a Ni-NTA column. The flowthrough from the Ni-NTA column was then concentrated to about 10 mg/ml then passed over a Superdex 200 size-exclusion column

equilibrated with 5mM potassium phosphate, 100mM KCl, 100mM NaCl at pH 6.7 (buffer A) to remove aggregated protein.

For crystallization, all fractions containing dimeric TfR were pooled and concentrated to about 12.5 mg/ml as determined by Bradford assay. Well solutions containing 2.25 to 2.5 M KCl and 1.1 to 1.4% polyethylene glycol 20K were set up. Hanging drops were assembled by mixing 2μl of TfR with 1μl of well solution. The crystals attain their maximum size after three weeks. These crystallization conditions are identical to those used to crystallize TfR produced in CHO cells, and the crystal morphology is also identical. This would suggest that the TfR produced in the baculovirus system crystallizes in the same space group ($P2_12_12_1$) as the CHO produced TfR, and is expected to have similar unit cell parameters: a=105.4Å, b=216.9Å, c=361.9Å, $\alpha=\beta=\gamma=90^{o}$.

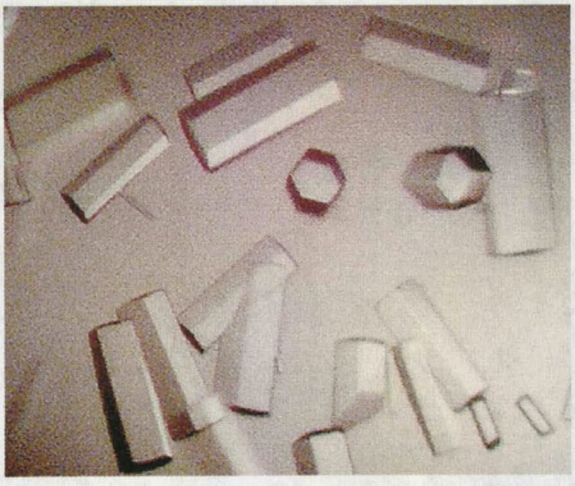


Figure 16. Crystals of TfR observed three weeks after the drops were set up.

CHAPTER 4

Virtual Screening of Three-Dimensional Compound Libraries

The "lock and key" hypothesis formulated by Emil Fisher about 100 years ago is a constantly recurring leitmotif in modern drug design. However, it is only recently that we have been in possession of detailed description of the "locks", i.e., the biochemical targets to which potent and selective drugs (the "keys") must fit. In the past, the structural features of these "locks" were inferred from structure/activity analysis of a series of experimental test ligands. This largely trial and error approach requires a tremendous amount of collective effort but can yield only limited information about the "lock". A three-dimensional structure obtained by crystallography obtained by X-ray crystallography and NMR spectroscopy contain a quantity of information that is orders of magnitude greater than that which can be obtained by building pharmacophores, even when a well studied set of ligands is available (25).

With the large number of three-dimensional structures of biomolecules becoming available and the continuing improvements in docking and scoring technologies, virtual lead screening (VLS) is becoming an attractive alternative to the traditional methods of lead discovery. In theory, VLS can examine infinite chemical diversity of drug-like molecules without synthesizing and experimentally testing every molecule to be screened. A typical corporate compound library may have between 200,000 to 1,000,000 molecules, but even with such a large sample collection, the experimental high-throughput screening (HTS) often does not result in viable leads. The high cost and technical complexity of such massive experimental testing are further motivations for the

theoretical alternative. A cost of US$10 per assay would result in a US$10,000,000 cost for a single HTS. And one very unique aspect, the virtual experiment, as opposed to a HTS assay, can be easily designed to select for a particular binding site or receptor specificity (26).

In our effort to find small molecules that, upon judicious optimization of structure and physical properties, will facilitate transport of drugs across the blood-brain barrier, we adapted a pathway similar to most structure-based drug discovery and design initiatives. In this particular phase of the project, our modest goal is to identify compounds that will bind to the target with affinity constants in the low to mid micromolar range.

We employed computational strategies to screen libraries of commercially available compounds to identify a small subset that will likely interact with the receptor. To carry out the initial filtering of the libraries we used the less CPU intensive program DOCK so that we can obtain a rank-ordered list with respect to goodness of fit in a reasonable amount of time.

DOCK solves the 3-D jigsaw puzzle of fitting putative ligands into appropriate sites on the receptor. A starting point is a high-resolution structure of the macromolecule. DOCK explores three important aspects of computational drug discovery: 1) creation of a negative image of the target site; 2) placement of putative ligands into the site; and 3) evaluation of the quality of fit (27).

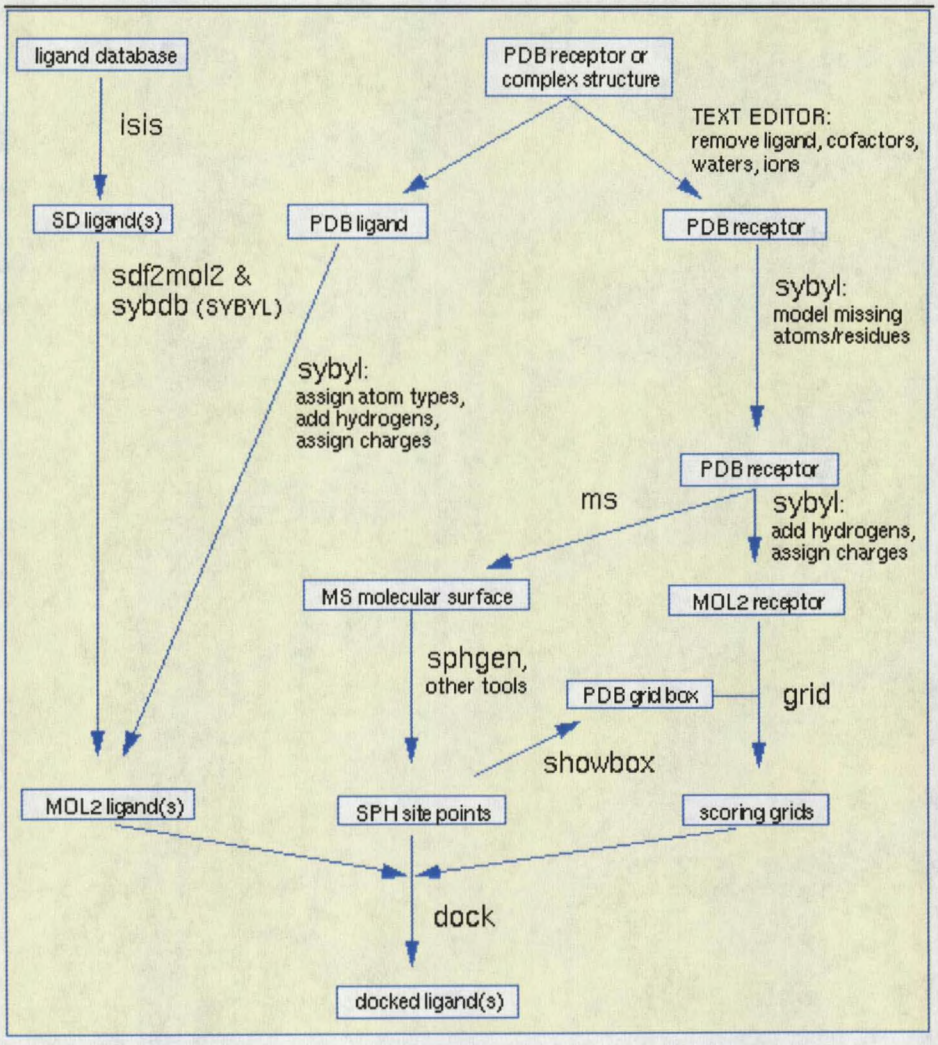
As a first step, DOCK characterizes the entire surface of the macromolecule, seeking the groves and invaginations in the surface that form the target sites. Alternatively, the user can define the active site where the screening is to be directed. The

default method of generating site points involves creating an inverse surface or negative image of the binding site that is derived from the molecular surface of the macromolecule. The active site is filled with a set of overlapping spheres, each of which touches the molecular surface at only two points. A set of sphere centers serves as the negative image of the site.

Next, the DOCK algorithm matches the x-ray of computer-derived structures of putative ligands to the negative image of the receptor on the basis of comparison of internal distances. Stated another way, ligand atoms are matched to the sphere centers to find matching sets (cliques) in which all the distances between the ligand atoms in the set are equal to the distances between the corresponding sphere centers, within some user-defined tolerance level. Each orientation generated is evaluated to measure the goodness of fit of the ligand to the site (28).

For evaluating goodness of fit, the compounds in the library are sorted based on the calculated ligand-receptor energy, taken to be approximately the sum of the van der Waal's attractive, van der Waal's dispersive, and Coulombic electrostatic energies. All evaluations are done on the pre-computed scoring grids that encompass the binding site in order to minimize the overall computational time. At each grid point, the receptor contribution to the score is stored. That is, receptor contributions to the score are calculated only once; the appropriate terms are simply fetched from memory. To generate the energy score, the ligand atom terms are combined with the receptor terms from the nearest grid point, or with a "virtual" grid point with interpolated receptor values. The program ranks each candidate on the basis of the intermolecular energy of the best orientation that was found (29).

The diagram below was reproduced from the DOCK User Manual to outline the tasks involved in docking a database of compounds to a receptor.



Preparation of TfR and Ligand Coordinates for Docking

The coordinates of TfR used in all docking simulations that we carried out were those of chain A from protein data bank entry 1CX8. Because a structure obtained by x-ray diffraction does not contain hydrogen atoms, we assigned all hydrogen atoms using the Biopolymer module of the Insight II software package. The pH was set to 7.2,

capping mode was off, and the ends were uncharged. Partial charges were assigned using the CVFF forcefield. After assigning all hydrogens and partial charges to the receptor, the structure was saved as a mol2 file.

The residues lining the vestigial active site of TfR were manually selected by visual inspection using the Insight II interactive viewer module. The Cartesian space coordinates of this subset were saved as a separate file in pdb format. Table 5 lists all residues included in the active site definition. All residues in green font belong to the apical domain while those in red belong to the protease-like domain. In Figure 17, the active site is highlighted.

TYR219	PHE288	ARG409	THR464	THR536	TYR565
VAL220	PHE297	ASP410	GLU465	LEU537	LEU566
ALA221	PHE298	ALA411	TRP466	ASP538	GLY567
TYR222	GLY299	TRP412	ASN483	ASN539	THR568
VAL257	HIS300	LYS418	LEU484	ALA540	
ARG258	ALA301	SER419	ASP485	ALA541	
ALA259	HIS302	PHE450	LYS486	PHE542	
GLY260	LEU303	ALA451	ALA487	ALA552	
LYS261	GLY304	SER452	SER492	VAL553	
ILE262	THR305	TRP453	ASN493	SER554	
THR263	GLY306	SER454	PHE494	PHE555	
PHE264	ASP307	ALA455	LYS495	CYS556	
ALA265	ILE386	GLY456	VAL496	PHE557	
GLU266	LEU387	ASP457	SER497	CYS558	
LYS267	ASN388	PHE458	ALA498	GLU559	
LEU280	ILE389	GLY459	LYS531	ASP560	
ILE281	PHE390	SER460	VAL532	THR561	
TYR282	GLY406	VAL461	GLU533	ASP562	
MET283	ALA407	GLY462	LYS534	TYR563	
LYS287	GLN408	ALA463	LEU535	PRO564	

Table 5. List of residues included in the active site definition.

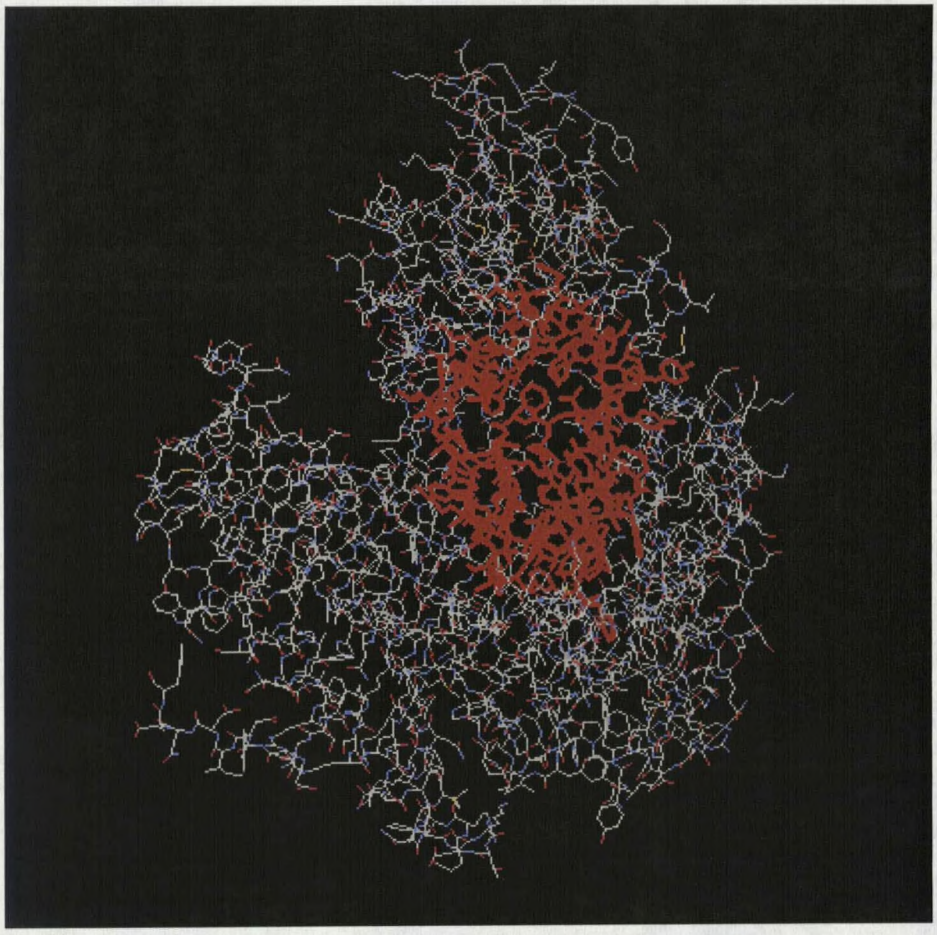


Figure 17. Residues lining the vestigial active site (depicted as red sticks).

To calculate the molecular surface of just the site where we want the putative ligands to bind, we used the accessory program `invertPDB` to generate an exclude.pdb file that defines receptor atoms away from the specified active site and should not be surfaced. The program `ms` from Quantum Chemistry Program Exchange (QCPE) was used to generate the molecular surface of the active site. We then used `SPHGEN` to fill up the ms-generated surface with overlapping spheres, visualized the first cluster from the output and selected 100 sphere centers that will be used for matching during the docking simulations.

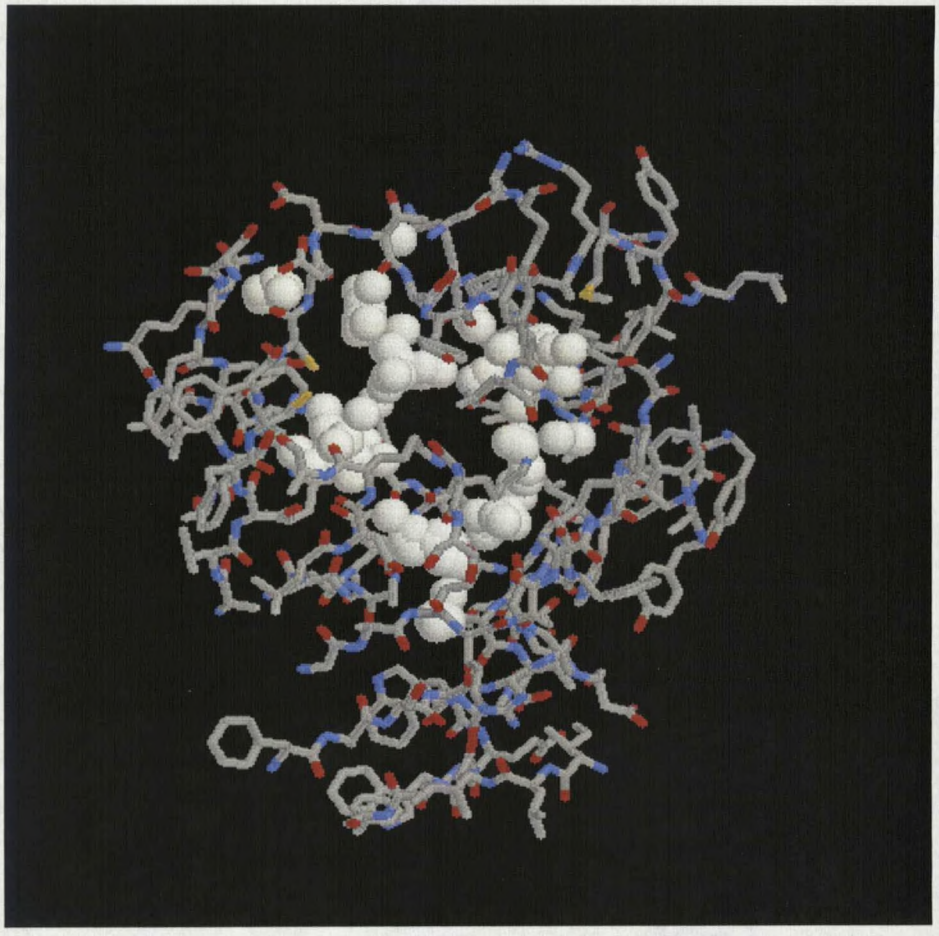


Figure 18. A set of 100 spheres whose internal distances were used to find compounds that fit into the active site.

Scoring girds for a subset of the receptor that "sees" the compound being docked were calculated using GRID prior to the docking run. The boundaries for calculating the grids were specified by a box (shown below). The input parameters for the GRID calculation are shown in the following lines:

```
compute_grids                   yes
grid_spacing                    0.3
output_molecule                 no
contact_score                   yes
contact_cutoff_distance         4.5
chemical_score                  yes
energy_score                    yes
energy_cutoff_distance          10
atom_model                      u
```

```
attractive_exponent             6
repulsive_exponent              12
distance_dielectric             yes
dielectric_factor               4
bump_filter                     yes
bump_overlap                    0.75
receptor_file                   LH_tfr.mol2
box_file                        box4_6A.pdb
vdw_definition_file
   /faculty2/Lawrence/Progs/DOCK/SGI/4.0.1/parameter/vdw.defn
chemical_definition_file
   /faculty2/Lawrence/Progs/DOCK/SGI/4.0.1/parameter/chem.defn
score_grid_prefix               lh_tfr_grid
```

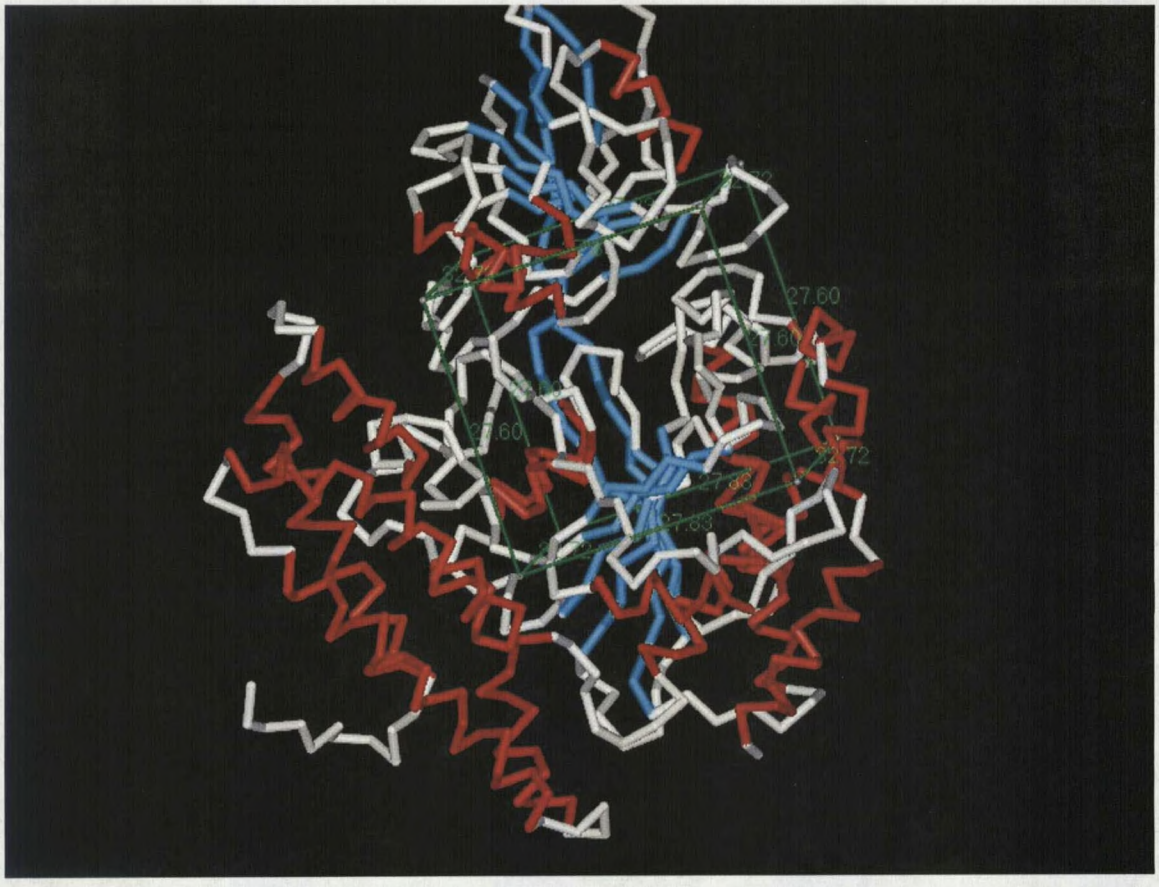


Figure 19. Scoring grids were calculated for the part of the receptor that is bounded by the box.

Prior to docking, the compound databases had to be prepared in the proper molecular file format. Version 4.0 of DOCK uses MOL2 as the primary molecule format

because this format has the advantage of storing all the necessary information for atom features, position, and connectivity. MOL2 files are also standardized formats that other modeling programs can read

Two of our databases, Chembridge and Tripos, were obtained in 2D-SD formats and the Available Chemicals Directory (ACD) was in a 3-D SD format. Hydrogen atoms were not explicit in all coordinate files but they can be inferred from the atom valence labels by a suitable program. MOL2 libraries with complete partial charge assignments for each atom in each molecule had to be generated.

We used the program Converter, also in the insight II package, to generate 3-D SD libraries from the two-dimensional libraries that we have. Converter recognizes both wedged bonds and atom stereo designators in the 2D file and converts the molecules to the appropriate configuration. One can also generate both R and S conformations for centers whose chirality is unspecified in the 2D file. Multiple molecules can be generated if more than one unspecified chiral centers are present. The three-dimensional conversion process can also be biased towards producing either all-chair or random conformations of 6-membered rings. We chose to generate just one conformation for every molecule in our file conversion, favoring extended conformation over random and all chair over random conformation. All hydrogens were written in the output 3-D SD file (30). To insert the proper names of the molecules in the 3-D SD libraries, we run an awk script provided to us by MSI.

```
#################################################################
#  file name: sd_add_name.awk
#
#  awk script to insert names into the first record of each
#  molecule in an SD file using the value from a data field
#  whose name is specified as input parameter
#
#  Usage:
#
#  nawk -f <this script> namefield=<value> <input_SD_file> >
#        <output_SD_file>
#################################################################

BEGIN {
  count = 1
  readname = 0
  isname = "<" namefield ">"
  name = ""
}

{ line[count++] = $0 }

/^>/ {
  if ( index($0,isname) ) {
    readname = 1
    next
  }
}

readname == 1 {
  name = $0
  readname = 0
}

/^\$\$\$\$/ {
  print name
  for ( i = 2; i < count; i++ )
    print line[i]
  count = 1
  name = ""
}
```

Now that we have the databases in 3-D SD formats, we can invoke yet another Insight II BCL macro to convert the 3-D SD file to a SYBYL MOL2 file with the charges needed for DOCK. An accompanying awk script was run assign the proper names to the molecules in the final mol2 libraries.

```
#################################################################
#  file name: sd_to_mol2.bcl
#
#  Macro to process all molecules from an SD file and write
#  them as a MOL2 file.
#  Marvin Waldman, Molecular Simulations Inc.  25-Jan-2000
#
#################################################################
Define_macro SD_to_MOL2 Lstring Input_SD_File Lstring Output_Mol2_File
    VLstring cmd
    Int count, Num_Frames, frame1, frame2
    Lstring *mol_list, mol
    Param_Add_File_List Input_SD_File sd
    cmd = "/bin/grep -c \$\$\$\$ " // $Input_SD_File // " > tmpxxx.tmp"
    bcl_unix $cmd
    Read tmpxxx.tmp "%d" $Num_Frames
    bcl_unix "/bin/rm tmpxxx.tmp"
    mol_list = { System_Fetch_Mol Temp_Mol* }
    count = sizeof($mol_list)
    if ( $count > 0 )
      Delete Object Temp_Mol*
    end
    frame2 = 0
    While ( $frame2 < $Num_Frames )
      frame1 = $frame2 + 1
      frame2 = $frame1 + 99
      if ( $frame2 > $Num_Frames )
        frame2 = $Num_Frames
      end
      Get Molecule MDL Range $frame1 $frame2 1 $Input_SD_File Temp_Mol
      Potentials Forcefield Fix -Print_Potentials Fix -
Print_Part_Chargs \
        Accept Temp_Mol*
      mol_list = { System_Fetch_Mol Temp_Mol* }
      foreach $mol in $mol_list
        Put Molecule Sybyl_Mol2 $mol "tmpxxx.mol2" -Transformed -
Displayed
        cmd = "/bin/cat tmpxxx.mol2 >> " // $Output_Mol2_File
        bcl_unix $cmd
        cmd = "/bin/echo ' ' >> " // $Output_Mol2_File
        bcl_unix $cmd
        bcl_unix "/bin/rm tmpxxx.mol2"
      end
      Delete Object Temp_Mol*
    end
    cmd = "/bin/nawk -f mol2_names_from_sd.awk sd_file=" //
$Input_SD_File // \
          " " // $Output_Mol2_File // " > tmpxxx.mol2"
    bcl_unix $cmd
    cmd = "/bin/mv tmpxxx.mol2 " // $Output_Mol2_File
    bcl_unix $cmd
End_Macro
```

Add_To_Pulldown SD_to_MOL2 File

```
##########################################################################
#  file name: mol2_names_from_sd.awk
#  awk script to replace the list of molecule names in a mol2 file
#  from the names in a corresponding SD file.
#
#  Usage:
#  nawk -f mol2_names_from_sd.awk sd_file=<input_sd_file>
#              <input_mol2_file> > <output_mol2_file>
#
#  Marvin Waldman
#  Molecular Simulations Inc., 26-Jan-2000
#
##########################################################################

Begin {
  get_cmpd_name = 0
  count = 0
}

{ line = $0 }

get_cmpd_name == 1 {
  while ( getline <sd_file > 0 ) {
    count++
    if ( count == 1 ) {
      line = $0
      get_cmpd_name = 0
      break
    }
    if ( $0 ~ /^\$\$\$\$/ )
      count = 0
  }
}


/^@<TRIPOS>MOLECULE$/ {
  get_cmpd_name = 1
}

{ print line }
```

For the Available Chemicals Directory, we skipped the 2-D to 3-D conversion, however we had to insert a few lines of code in the BCL macro to include adding hydrogens to each molecule in the database before charges are to be assigned.

Docking Virtual Chemical Database onto TfR

We screened the ACD, Tripos and Chembridge databases for compounds that potentially bind to TfR. In all, 250,455 compounds were read by DOCK while only 248,429 were actually scored because these were the only ones that satisfied the filters prior to scoring. Below is an example of an input to DOCK.

```
flexible_ligand                 no
orient_ligand                   yes
score_ligand                    yes
minimize_ligand                 no
multiple_ligands                yes
parallel_jobs                   no
random_seed                     1000
match_receptor_sites            yes
random_search                   no
automated_matching              yes
maximum_orientations            500
write_orientations              no
intermolecular_score            yes
gridded_score                   yes
grid_version                    4
bump_filter                     yes
bump_maximum                    3
contact_score                   no
chemical_score                  no
energy_score                    yes
atom_model                      u
vdw_scale                       1
electrostatic_scale             1
ligands_maximum                 500000
initial_skip                    0
interval_skip                   0
heavy_atoms_minimum             3
heavy_atoms_maximum             100
rank_ligands                    yes
rank_ligand_total               500
restart_interval                100
ligand_atom_file
/home/epineda/Libraries/ACD/acd00210_231_charged.mol2
receptor_site_file
/home/epineda/TfR/LH_TfR/LH_site/LH_tfr_ed.sph
score_grid_prefix
/home/epineda/TfR/LH_TfR/LH_grid/new_grid
vdw_definition_file
/home/epineda/TfR/dock_parameters/vdw.defn
quit_file                       acd0021_231_chrg.quit
dump_file                       acd0021_231_chrg.dump
ligand_energy_file              acd0021_231_chrg_nrg.mol2
```

```
info_file                       acd0021_231_chrg.info
restart_file                    acd0021_231_chrg.rst
```

We are aware that the 500 maximum conformations that will be tried for each compound in the library do not constitute an exhaustive conformational search in our docking simulations. In an effort to compensate for this shortcoming, we selected a total of 2075 top scoring compounds to be refined using more CPU intensive programs that are also independent of site descriptors (32).

We assumed that the DOCK-generated ligand conformations were close to the minimum bound conformation. Using the minimization utility in FLO, we calculated the estimated total energy of the protein-ligand systems. The total energy of the system is most related to the binding constant. FLO is easy to use and it outputs a detailed accounting of the different factors contributing to the total association energy. Among the numbers that it returns are: total estimated binding energy (Eass), internal strain energy of the ligand (Elig), binding site energy relative to local minimum for empty site (Eshe), electrostatic energy (Eest), contact energy (Ecnt), van der Waals dispersion energy (Evw+), van der Waals energy of interaction (Vdw), number of hydrophobic contacts (Nhph) and number of H-bonds (Nhbd). Ecnt provides a good measure of hydrophobic interactions.

After sorting the output of FLO, we closely examined all complexes that have a total estimated binding energy less than minus 30 kJ/mole because this number, according to the FLO manual, is characteristic of ligands with less than micromolar

binding affinity. Twenty seven representative compounds were purchased and their interaction with TfR was characterized by isothermal titration calorimetry.

CHAPTER 5

Isothermal Titration Calorimetry

What comes after all the computational prediction of putative ligands is the experimental validation of the rigorous, and sometimes alchemical, quantitative exercise. Our goal is to find compounds that exhibit significant binding to the target. In order to measure TfR ligand interaction, we chose to use isothermal titration calorimetry because it is the only technique that will provide the most number of thermodynamic information from a single optimized experiment.

Isothermal titration calorimetry (ITC) follows the heat change when a test compound binds a target protein. It allows precise measurement of the binding affinity. There is no requirement for competing molecules, making interpretation of the result more straightforward. Titration in the presence of other ligands rapidly provides information on the mechanism of action of the test compound, identifying the intermolecular complexes that are relevant for structure-based drug design. Because it gives an accurate measure of the stoichiometry of the complex being formed, ITC allows the evaluation of the population of the sample that is functional. Protein fragments as well as catalytically inactive mutants can be characterized, as long as they exhibit significant binding. It is the only technique that directly measures the enthalpy of binding (ΔH°). Knowing the ΔH° values under different buffer conditions allow characterization of proton movement linked to the association of protein and ligand, giving information about the ionization of the groups involved in binding, a very powerful application of a fundamental principle that is Hess's law. Biochemical systems are unique in that they

exhibit enthalpy-entropy compensation where increased binding is offset by an entropic penalty, reducing the magnitude of change in affinity. This also causes a lack of correlation between the free energy of binding ($\Delta G°$) and $\Delta H°$. When characterizing structure-activity relationships (SAR), most groups involved in binding can be detected as contributing to $\Delta H°$, but not to affinity. Large enthalpy changes may reflect a modified binding mode, or protein conformational changes. Thus, $\Delta H°$ values may highlight a potential discontinuity in SAR, so that experimental structural data are likely to be particularly valuable in molecular design (32).

Enthalpically dominated ligands do not behave the same as entropically dominated ligands. Current approaches in drug design usually generate relatively hydrophobic and conformationally constrained ligands. Such rigid ligands are characterized by an entropically dominated binding affinity often accompanied by an unfavorable binding enthalpy because more nonbond interactions with the solvent are lost by each binding partner than the number of nonbond interactions gained by the formation of the protein-ligand complex. However the large number of bound solvent molecules liberated to the bulk contributes a net higher disorder to the system, making the complex formation favorable. The disadvantage with conformationally constrained ligands is their inability to adapt to changes in geometry of the binding site. This can be a major problem, particularly with antiviral and antibacterial agents where the target molecule is usually capable of incurring point mutations. Such mutations can lead to drug resistance since the relatively rigid drug is not able to adapt to the new pocket. These problems also arise due to naturally occurring genetic polymorphisms that make drugs effective for many, but potentially fatal to the unlucky. The design of ligands that are capable of

adapting to a change in the structure of the target requires the introduction of certain elements of flexibility or the relaxation of some conformational constraints. Since these compounds pay a larger conformational entropy penalty upon binding, the optimization of their binding affinity requires the presence of a favorable binding enthalpy. This has been demonstrated in the case of HIV-1 protease inhibitors. It was shown that a thermodynamic guide to drug design permits the identification of drug candidates with a lower susceptibility to target mutations causing drug resistance (33).

Validation of Virtual Screening Results by ITC

Of the 2075 compounds refined by FLO, there were about 300 with estimated total binding energies less than minus 30 kJ/mol. After visual inspection of the compounds that we found interesting, unique, and affordable we purchased twenty seven candidates. Twelve of these compounds dissolve in aqueous buffer so we were able to carry out ITC experiments to characterize their interaction with TfR.

Because ITC is such a sensitive instrument, one has to make sure that solution conditions of the injectant (usually the ligand) and the receptor are as close as possible. In an ideal case the only difference is that the receptor is present in one while the ligand is in the other solution, but everything else must be the same.

In our case, Fast Protein Liquid Chromatography (FPLC) separates dimeric TfR from the any multimeric form or aggregates. If the column has been pre-equilibrated with the running buffer, this step is essentially equivalent to multiple dialyses. However, we still dialyze the FPLC purified sample in order to obtain a suitable solvent for the ligands. When a change in pH greater than 0.05 units is detected, back titration is necessary to

bring the pH of the ligand solution as close as possible to the pH of the receptor solution. For all compounds that dissolve well in aqueous buffer, we used a solution containing 150 mM NaCl and 20 mM HEPES at pH 7.5 as the generic solvent.

The interesting ligand showed saturable signals and therefore caught our attention. The compounds are Ethylene Glycol-bis(beta-aminoethyl ether)-N,N,N',N'-tetraacetic acid, Ethylenediaminetetraacetic acid, and Phenolphthalein complexon. All compounds are tetraacetic acids and we suspect, just like FLO predicts, that the interaction is predominantly electrostatic and may be specific to the pre-organized lysines that line the vestigial active site of TfR, if indeed they bind to that pocket. The structures of the compounds are shown in Figure 20. Figures 21-23 are representative binding isotherms of each compound when titrated into a solution of TfR.

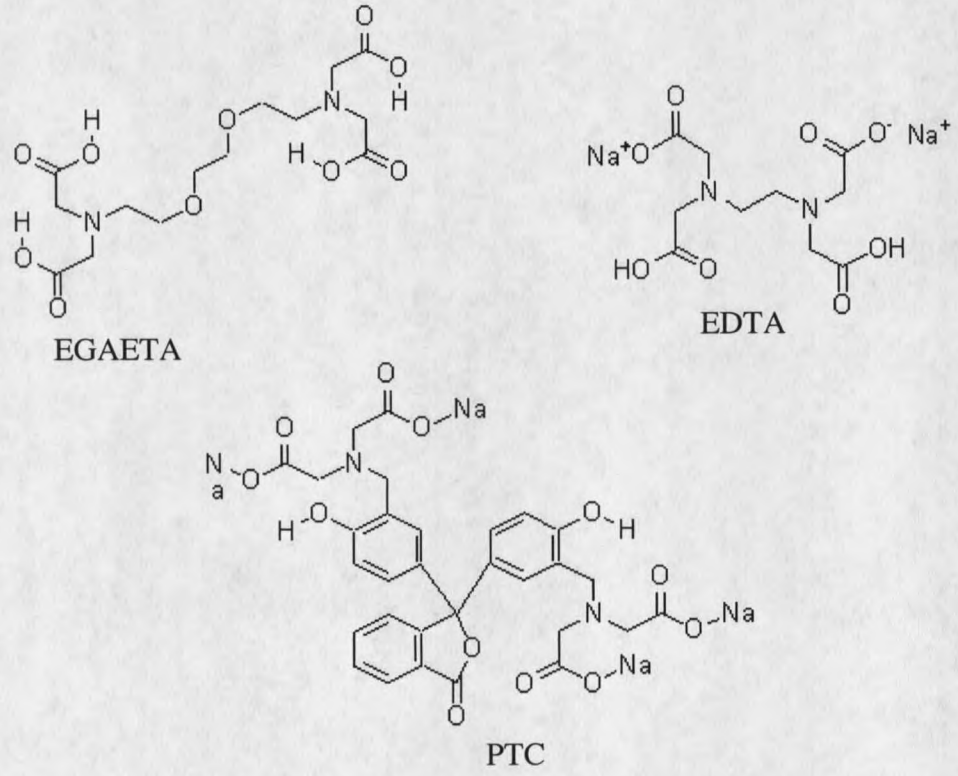


Figure 20. Structures of EGAETA, EDTA, and PTC.

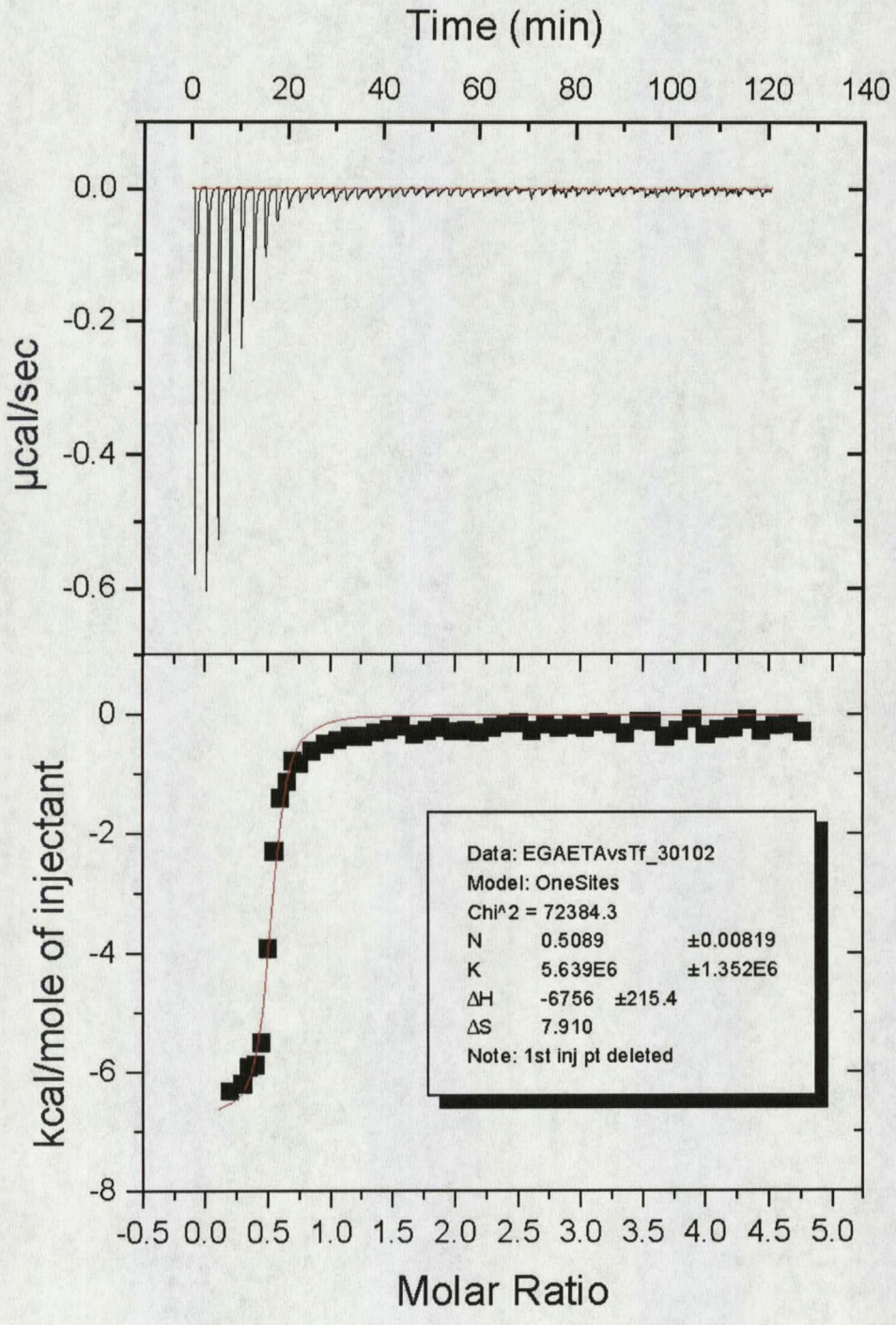


Figure 21. Binding isotherm of EGAETA when titrated into a TfR solution.

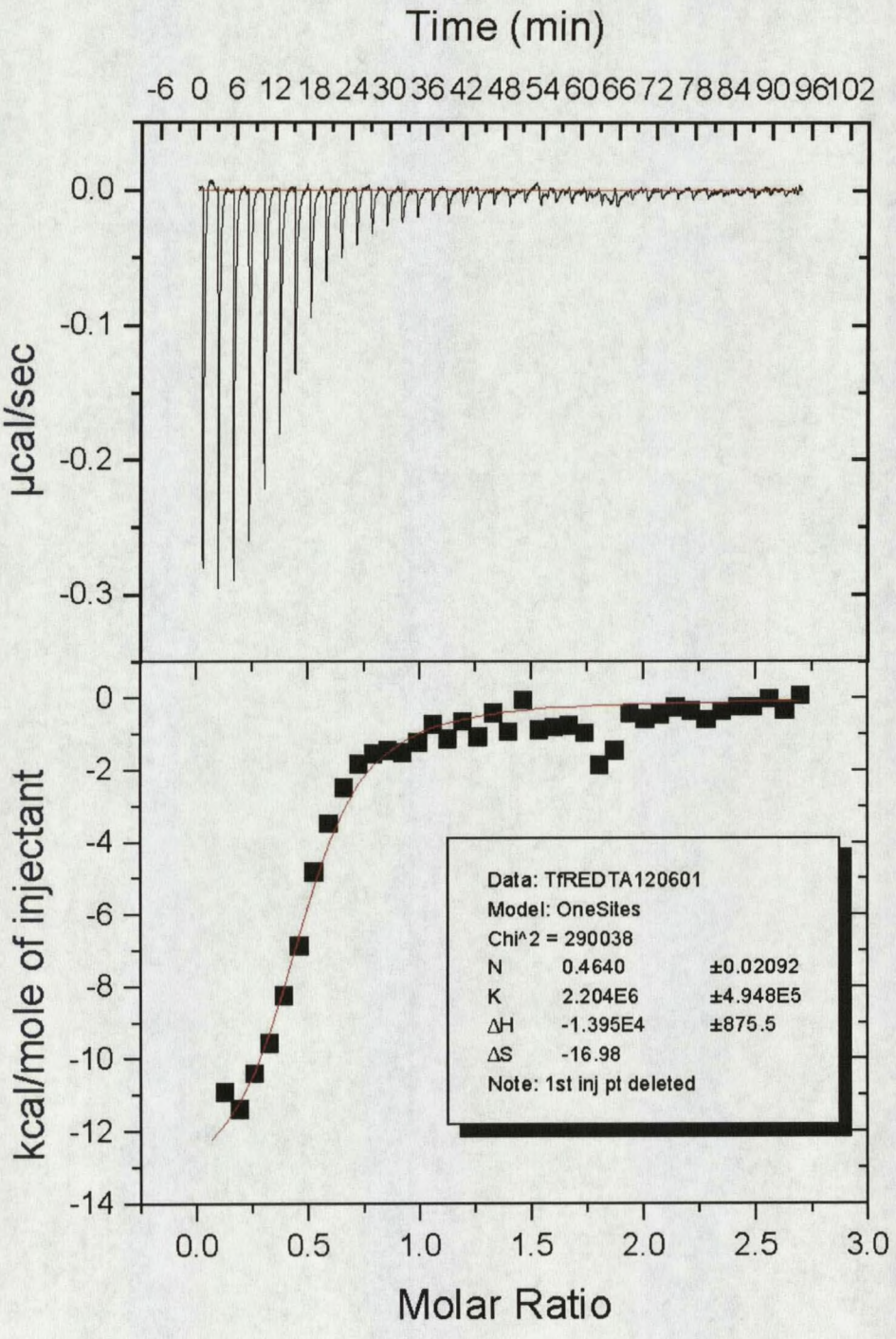


Figure 22. Binding isotherm of EDTA when titrated into a TfR solution.

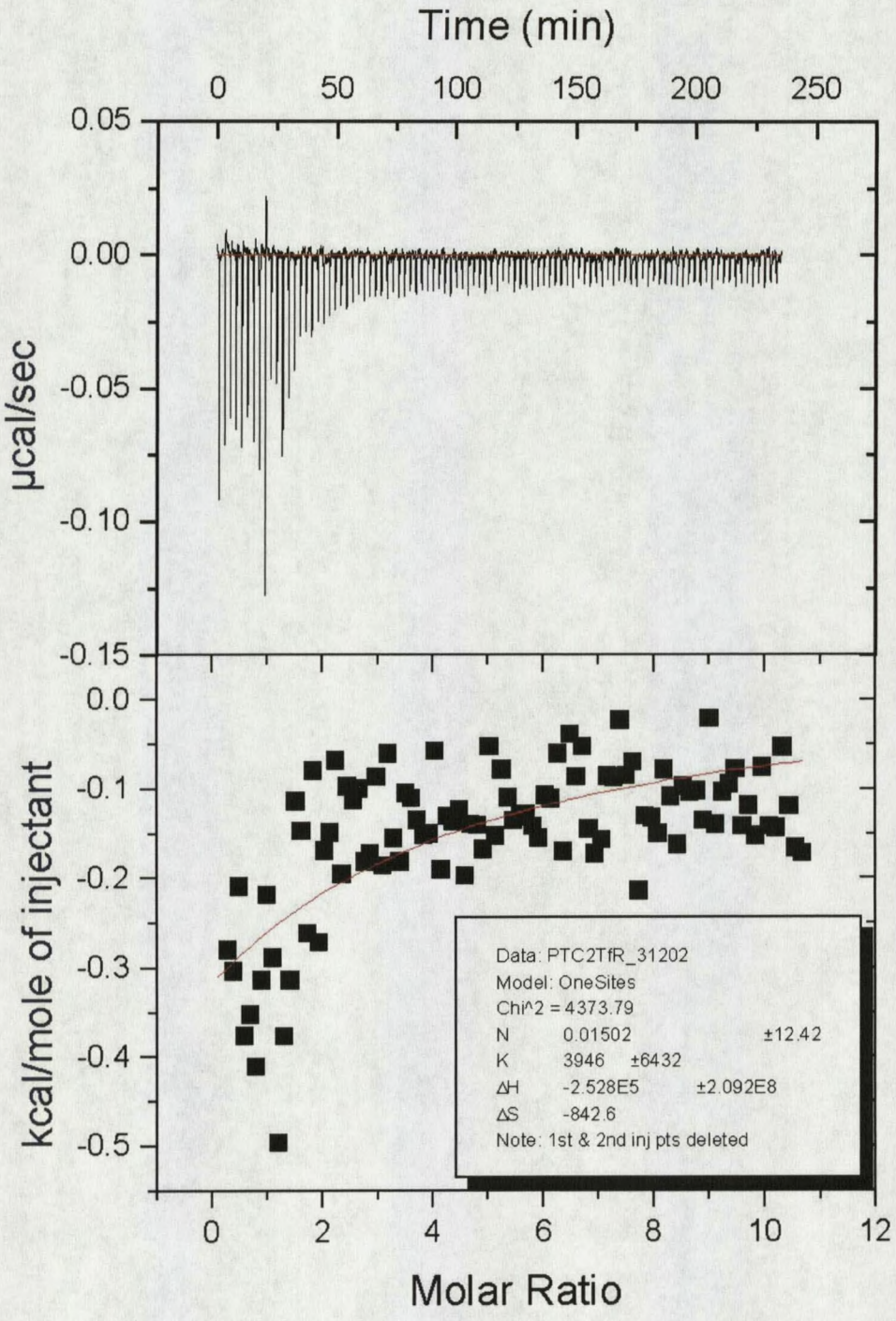


Figure 23. Binding isotherm of PTC when titrated into a TfR solution.

The three compounds can be superimposed in the TfR active site. The binding configurations are shown in the figure below.

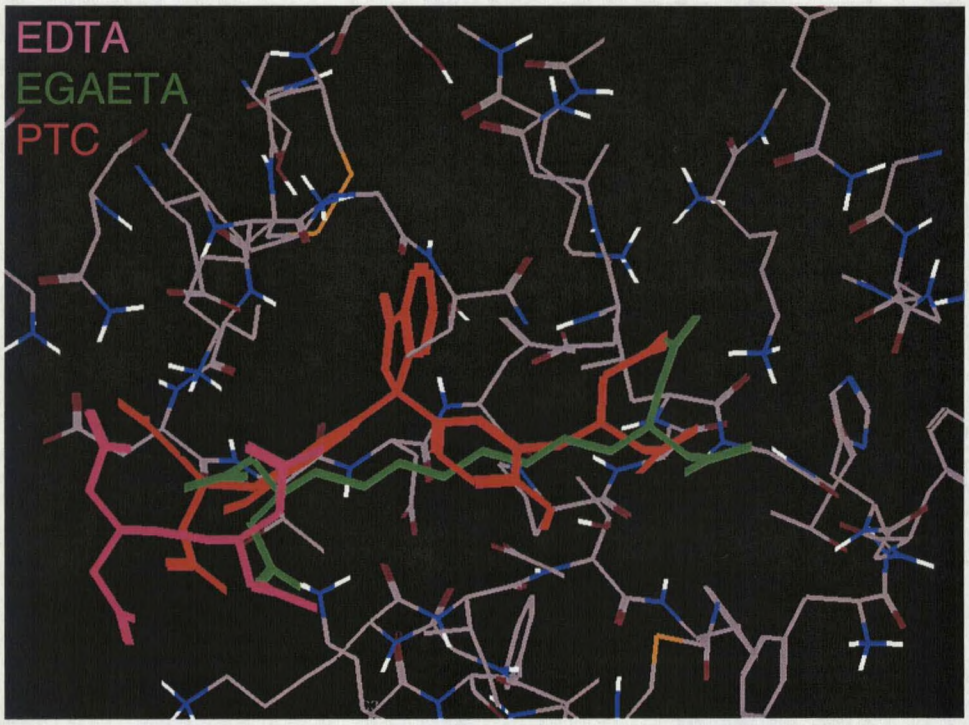


Figure 24. Predicted binding modes EDTA, EGAETA, and PTC superimposed in the active site of TfR.

The interpretation of the results just stated is not that clear to us. We are aware that Ca^{2+} ions were found in the electron density of the TfR-HFE complex (34). Some groups claimed that TfR loses its affinity for transferrin after extensive washing with EDTA (35).

However, in the crystal structure of the TfR dimmer solved by Lawrence, et al, there were no Ca^{2+} that were found the electron density maps. On the other hand some of

the Sm^{3+} that were soaked in to help solve phases localized in sites occupied by Ca^{2+} in the HFE:TfR structure.

Although we'd like to see the ligands dock into the vestigial active site of TfR, it is not a remote possibility that the signals detected by ITC were partly due to the dissociation of Ca^{2+} from TfR and its subsequent binding of Ca^{2+} to the ligand being titrated into the solution of the macromolecule.

Experiments are on the way to soak in PTC, EDTA and EGAETA into the TfR crystals that we obtained. Ultimately, it will be a structural analysis by x-ray diffraction that will answer whether or not we identified lead compounds that, through successive optimization of structure and physical property, will help solve the problem of drug delivery across the blood-brain barrier.

It is also worth noting that in an experiment that involved titrating PTC into preformed TfR:Tf complex, the overall behavior of the titration curve is the same as that when PTC was titrated into TfR alone. Several possibilities could explain this: 1) that there are at least two distinct sites where PTC binds to TfR and these sites do not overlap with the Tf binding site; 2) if PTC is stripping away TfR-bound Ca^{2+} then it probably binds Ca^{2+} that are equally accessible in the Tf-bound and Tf-unbound transferrin receptor.

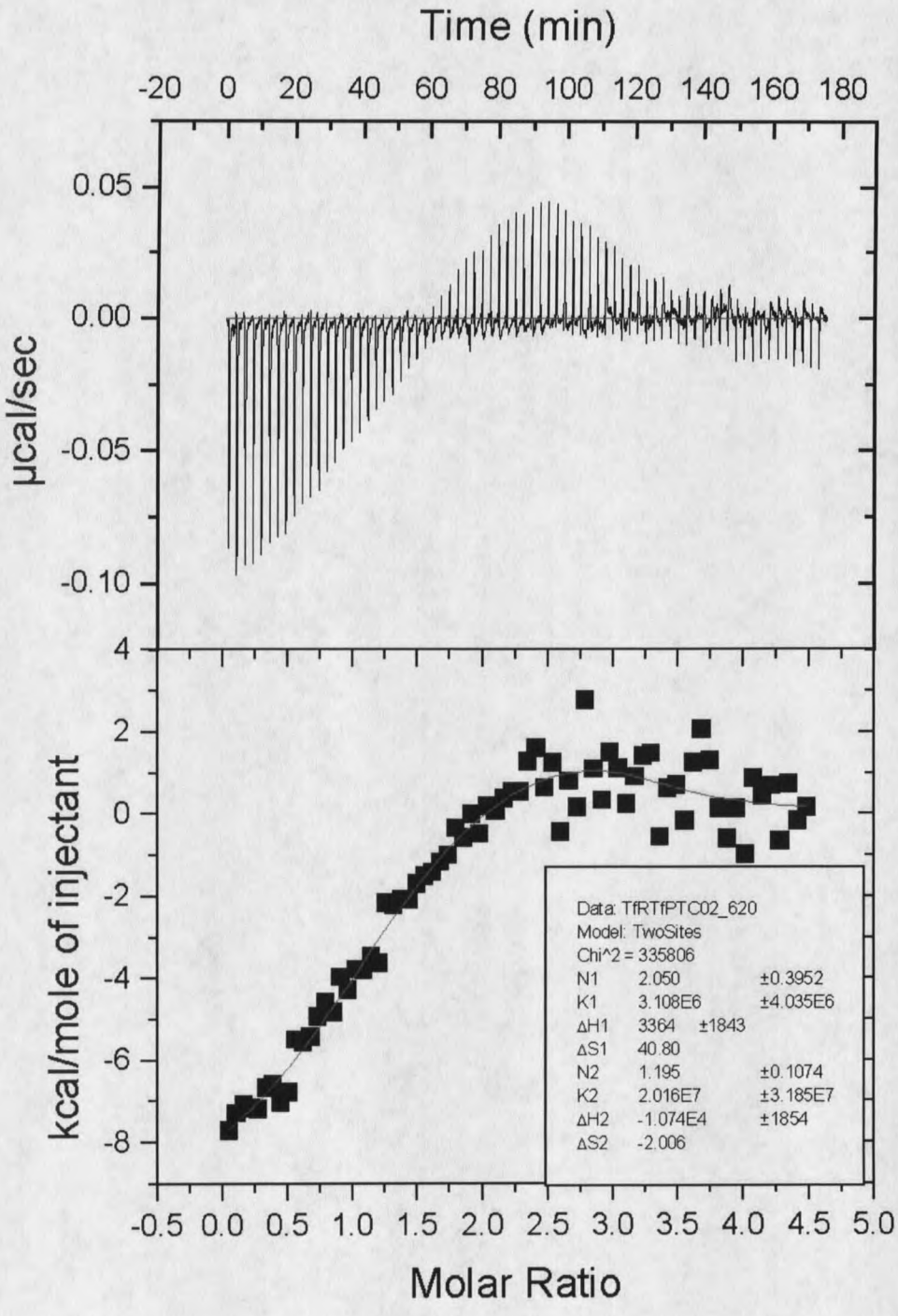


Figure 25. Binding isotherm of PTC when titrated into a solution of preformed (TfR-diferricTf)$_2$

Orberger et al interpreted the results of their experiments as indicating that the bound Ca^{2+} stripped away by incubation with EDTA are critical to Tf binding (35). There was no estimate nor were any attempts reported to determine the number of such "Ca^{2+}-binding sites" in TfR. It is worth noting from our purification protocol that to liberate TfR from the immobilized Tf one has to strip away the Tf-bound Fe^{3+} by washing the resin with a low pH buffer containing a strong Fe^{3+}-chelator (Deferoxamine Mesylate), that can also strip away divalent metal ions. The purification protocol that Orberger et al used was similar to ours and they claimed that the purified TfR is still competent to bind transferrin because the critical Ca^{2+} ions remain bound to TfR possibly because in the complex they are not accessible to the solvent .If only this/these tightly-bound and Tf-shielded Ca^{2+} remain with the purified TfR, what could account for the multiphasic titration that we observed for PTC to TfR and PTC to TfR:Tf? If Ca^{2+} is/are coming off from the protein we should be able to see a contribution of this process/es to the overall thermodynamic event being captured by ITC. We should also observe Tf:TfR dissociation if indeed the Ca^{2+} being stripped away from TfR are crucial in keeping a Tf-binding competent conformation, assuming the two macromolecules do not exhibit chaperone activity for one another.

CHAPTER 6

Conclusion and Future Directions

Conclusion

The overall goal of our experiments to identify compounds that could facilitate transport across the blood brain barrier is not without precedence. A huge amount of effort is devoted by other groups to find peptidomimetic monoclonal antibodies against cell surface receptors of the BBB endothelial layer. It has been clearly demonstrated that a molecule as large as an immunoglobulin can piggyback on TfR as it transcytoses the BBB endothelial layer.

The uniqueness of our approach comes from the fact that it is based on the structure of a known receptor capable of transcytosing the endothelial layer of the BBB. If strongly complements all efforts to find small molecule drugs that will work effectively against CNS diseases.

Future Directions

While we await the result of our own effort to determine by x-ray diffraction if the ligands we identified indeed bind to the site where we directed our searches, more thermodynamic analysis can be done to shed light on the issue of whether or not these divalent metal chelators are indeed interacting with a structural metal ion of the protein. It also remains to be determined if Ca^{2+} copurify with TfR in stoichiometric quantity. A small amount of TfR isolated using our protocol can analysed by Atomic Absorption Spectrophotometry to answer these conundrum. Extensive dialysis of TfR against a

strong and specific Ca^{2+} chelator followed by size exclusion chromatography to remove the chelator can be done prior to titrating TfR with our ligands. Such steps should clearly determine whether the heat evolved during our preliminary measurements was coming from association of these compounds with TfR or with a bound cation that copurify with TfR.

We need to continue our search for suitable co-solvents that will solubilize some of the compounds we initially purchased. All of these compounds scored very well with DOCK and with FLO. A wise approach to drug discovery entails that several leads should be followed. One of the best ways to solve a problem is to not force one approach that is destined to fail. We should really be looking for more scaffolds, specifically those that are amenable to synthetic modification, and also novel structures to give us a free hand in pursuing the project in several different directions.

LITERATURE CITED

1. Rubin LL, Staddon JM: **The cell biology of the blood-brain barrier.** *Annual Reviews in Neuroscience* 1999, **22**:11-28.

2. Bickel U, Yoshikawa T, Pardridge WM: **Delivery of peptides and proteins through the blood-brain barrier.** *Advanced Drug Delivery Reviews* 2001, **46**:247-27.

3. Huber JD, Egleton RD, Davis TP: **Molecular physiology and pathophysiology of tight junctions in the blood-brain barrier.** *Trend in Neuroscience* 2001, **24**:719-725.

4. Lawrence RN: **William Pardridge discusses the lack of BBB research.** *Drug Discovery* today 2002, **7**:223-226.

5. Pardridge WM: **Drug delivery to the brain.** *Journal of Cerebral Blood Flow and Metabolism* 1997, **17**:713-731.

6. Pardridge WM: **Non-invasive drug delivery to the human brain using endogenous blood-brain barrier transport systems.** *PSST* 1999, **2**:49-59.

7. Gregoire N: **The blood-brain barrier.** *Journal of Neuroradiology* 1989,**16**:238-250.

8. Pardridge WM. *Brain Drug Targeting: The Future of Brain Drug Development.* Cambridge, England: Cambridge University Press; 2001:36-81.

9. Jefferies WA, Brandon MR, Hunt SV, Williams AF, Gatter KC, Mason DY: **Transferrin receptor on endothelium of brain capillaries.** *Nature* 1984, **312**:162-163.

10. Li JY, Boado RJ, Pardridge WM: **Blood-Brain Barrier Genomics.** *Journal of Cerebral Blood Flow and Metabolism* 2001, **21**:61-68.

11. Lawrence CM, Ray S, Babyonyshev M, Galluser R, Borhani DW, Harrison SC: **Crystal Structure of the ectodomain of the human transferrin receptor.** *Science* 1999, **286**:779-782.

12. Fishman JB, Rubin JB, Handrahan JV, Connor JR, Fine RE: **Receptor-mediated transcytosis of transferrin across the blood-brain barrier.** *Journal of Neuroscience Research* 1987, **18**:299-304.

13. Skarlatos S, Yoshikawa T, Pardridge WM: **Transport of ^{125}transferrin through the rat blood-brain barrier.** *Brain Research* 1995, **683**:164-171.

14. Broadwell RD, Baker-Cairns BJ, Friden PM, Oliver C, Villegas JC: **Transcytosis of protein through the mammalian cerebral epithelium and endothelium. III. Receptor-mediated transcytosis through the blood-brain barrier of blood-borne transferrin and antibody against the transferrin receptor.** *Experimental Neurology* 1996, **142**:47-65.

15. Pardridge WM: **CNS Drug Design Based on Principles of Blood-Brain Barrier Transport.** *Journal of Neurochemistry* 1998, **70**:1781-1792.

16. Carter RE, Feldman AR, Coyle JT: **Prostate-specific membrane antigen is a hydrolase with substrate and pharmacologic characteristics of a neuropeptidase** *Proceedings of the National Academy of Sciences* 1996, **93**:749-753.

17. Luthi-Carter R, Berger U V, Barczak AK, Enna M, Coyle JT: **Isolation and expression of a rat brain cDNA encoding glutamate carboxypeptidase II.** *Proceedings of the National Academy of Sciences* 1998, **95**:3215-3220.

18. Pfeifer TA: **Expression of heterologous proteins in stable insect cell culture.** *Current Opinion in Biotechnology* 1998, **9**:518-521.

19. Lenhard T, Reilander H: **Engineering the folding pathway of insect cells: Generation of a stably transformed insect cell line showing improved folding of a recombinant membrane protein of special interest** . *Biochemical Biophysical Research Communications* 1997, **238:**823-830.

20. Jenkins N, Parekh RB, James DC: **Getting the glycosylation right: implications for the biotechnology industry**. *Nature Biotechnology* 1996, **14:**975-983.

21. Possee RD: **Baculoviruses as expression vectors.** *Current Opinion in Biotechnology* 1997, **8:**569-572.

22. Kulakosky PC, Shuler ML, Wood HA: **N-glycosylation of a baculovirus-expressed recombinant glycoprotein in three insect cell lines.** *In Vitro Cellular and Developmental Biology - Animal* 1998, **34:**101-108.

23. Instruction Manual: Bac-to-Bac Baculovirus Expression Systems. Invitrogen Life Technologies.

24. Turkewitz AP, Amatruda JF, Borhani D, Harrison SC, Schwartz AL: **A high yield purification of the human transferrin receptor and properties of its major extracellular fragment.** *Journal of Biological Chemistry* 1988, **263**: 8318-8325.

25. Bohacek RS, McMartin C, Guida WC: **The Art and Practice of Structure-Based Drug Design: A Molecular Modeling Perspective.** *Medicinal Research Reviews* 1996, **16**:3-50.

26. Abagyan R, Totrov M: **High-Throughput Docking for Lead Generation.** *Current Opinion in Chemical Biology* 2001, **5**:375-382.

27. Kuntz ID, Meng EC, Shoichet BK: **Structure-Based Molecular Design.** *Accounts of Chemical Research* 1994, **27**:117-123.

28. Kuntz ID: Structure-Based Strategies for Drug Design and Discovery. *Science* 1992, 25:1078-1082.

29. DOCK 4.0 User's Manual

30. Converter: from Insight II, March 2000, San Diego: Molecular Simulations Inc., 1999.

31. McMartin C, Bohacek R.S.: **QXP: Powerful, Rapid Computer Algorithms for Structure-Based Drug Design** *Journal of Computer-Aided Molecular Design* 1997, **11**:333-344.

32. Ward WHJ, Holdgate GA: **Isothermal titration Calorimetry in Drug Discovery.** *Progress in Medicinal Chemistry* 2001, **38**:309-376.

33. Velazquez-Campoy A, Luque I, Freire E: **The application of thermodynamic methods in drug design.** *Themochimica Acta* 2001, **380**:217-227.

34. Bennett M, Lebrón J, Bjorkman P: **Crystal Structure of the Hereditary Haemochromatosis Protein HFE Complexed with Transferrin Receptor.** *Nature* 2000, **403**:46-53.

35. Orberger G, Fuchs H, Geyer R, Geβner R, Köttgen E, Taubner R: **Structural and Funcational Stability of the Mature Tranferrin Receptor from Human Placenta.** *Archives of Biochemistry and Biophysics* 2001, **386**:79-88.

MONTANA STATE UNIVERSITY - BOZEMAN
3 1762 10372245 8