ENGINEERING FOR IMPROVED FOLDING OF A HUMAN PROLACTIN ANTAGONIST

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Laura DePalatis, B.S.

The Ohio State University
2005

Master's Examination Committee:

Dr. Charles Brooks, Advisor

Dr. Gary Means

Dr. Richard Swenson

Approved by

Advisor
Department of Biochemistry
ABSTRACT

Human prolactin (hPRL) is a 23 kilodalton protein secreted by the anterior pituitary gland. Its structure is a four alpha helix bundle with an up-up-down-down motif. The molecule contains three disulfide bonds. The main function of hPRL is to aid in the development and maintenance of mammary cells and to initiate lactation. Human prolactin has also been recognized for its involvement in the growth and proliferation of many types of cancers, predominately breast cancer. For this reason, there is much interest in creating prolactin analogs that antagonize the growth of cancer cells.

Several prolactin antagonists have been developed in various labs. The prolactin antagonist most important to this document is Δ41-52 hPRL, in which the amino acid residues 41-52 have been deleted. This molecule has been proposed to work by interfering with one of the receptor binding sites on the prolactin molecule. Preparation of this molecule is difficult because the purified protein contains a large fraction of biologically inactive dimer (usually greater than 50% of the protein expressed is in the dimer form). This is presumably due to strain within the molecule resulting from such a large deleted region. The dimers can be reduced to their monomeric forms using mild reducing conditions.
on polyacrylamide gel leading to the assumption that the dimeric are held together by an intermolecular disulfide linkage. If the cysteine residues that would normally be involved in an intramolecular disulfide bond are instead free for other types of interactions, they can form disulfide bonds with similar cysteine residues from other prolactin molecules, resulting in a dimeric species.

The goal of this project was to engineer the Δ41-52 hPRL molecule so that it would recover some of the folding capabilities of the wild-type compound that would allow correct disulfide formation while retaining the antagonist biological activity of Δ41-52 hPRL. To do this, several mutant proteins were created by inserting various amino acids into the 41-52 region. Four initial mutants added a series of glycine residues (2, 3, 4, and 5 glycines were added to the Δ41-52 hPRL molecule). Another mutation strategy involved adding two alanine residues instead of glycines and yet another mutation added a beta turn (serine-proline-glycine-glycine) to mimic the larger turn created by residues 41-52. Two other mutations involved returning native residues to the deleted region in an attempt to take advantage of native folding.

To study the folding of the proteins, the ultraviolet absorbences at 280 nm and 260 nm were observed while purifying the proteins using anion-exchange chromatography. The relative amounts of monomeric and dimeric species were then measured by purifying the protein over a gel filtration column to separate the two species. The identity and relative amount of contamination in the monomer species was studied using SDS-containing gel electrophoresis and the identity of each protein was further elucidated using total protein mass spectrometry. The
folding and structure was also assessed using ultraviolet absorbance, fluorescence emission, and circular dichroism spectroscopy. Finally, the biological activity of each of the mutant proteins was assessed using an agonist biological assay with cells expressing the prolactin receptor.

While many of the proteins that had increased folding capability resembled wild-type hPRL in the biological assay, the 4G mutant not only proved to be a well-folded molecule but also had a similar biological activity to Δ41-52 hPRL. This information shows some promise in designing second-generation Δ41-52 hPRL antagonists for possible therapeutic use.
Dedicated to my parents
ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Charles L. Brooks, for his patience, guidance, and contagious enthusiasm for this and all projects in the lab. Discussions with him have shaped my approach to scientific thinking and problem solving and his encouragement has taught me patience and persistence when confronted with problems in the lab. I am also grateful for his sincere interest and support in my future career goals.

I thank Dr. Umasandri Sivaprasad, Jeff Voorhees, Dr. Colleen Almgren, Mark Troyer, Toni Hoepf, Veronica Dawson and other lab members for their collaborations in projects guidance and for providing a pleasant and comfortable working environment. Time spent in the lab during this project was made much more enjoyable by the presence of friends.

I would also like to thank Gordon Renkes for his help with CD Spectroscopy Instrument and Ben Jones at the Campus Chemical Instrumentation Center Mass Spectrometry and Proteomics Facility for performing the mass spectrometry and data analysis.

This research was funded by The Department of Veterinary Biosciences at The Ohio State University.
VITA

July 30, 1981............................................. Born, Dallas, Texas

Summer 2002............................................ Research Assistant
The Dow Chemical Company
Freeport, Texas

2002 – 2004............................................. Student Associate
Chemistry Department
The Ohio State University

2004....................................................... B.S. Biochemistry
The Ohio State University

2004 – present........................................... Graduate Research Associate
The Ohio State University

FIELDS OF STUDY

Major Field: Biochemistry
# TABLE OF CONTENTS

Abstract....................................................................................................................... ii  
Dedication.................................................................................................................. v  
Acknowledgments..................................................................................................... vi  
Vita.............................................................................................................................. vii  
List of Tables.............................................................................................................. xi  
List of Figures........................................................................................................... xii  

Chapters:  
1. Introduction........................................................................................................... 1  
   1.1 Background: Human Prolactin (hPRL)......................................................... 1  
   1.2 Background: Human Prolactin Receptor (hPRLr)....................................... 2  
   1.3 Prolactin/Prolactin Receptor Mechanism of Binding................................. 3  
   1.4 The Role of Prolactin in Breast Cancer..................................................... 4  
   1.5 Current Prolactin Antagonists................................................................. 5  
   1.6 Mutant Proteins Designed For This Project.............................................. 8  
   1.7 Protein Expression, Folding, and Characterization................................. 9  
   1.8 Assessment of Biological Activity.......................................................... 10  
2. Materials and Methods....................................................................................... 17  
   2.1 The pT7-7 Phagemid................................................................................... 17  

viii
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Human prolactin mutant sequences</td>
<td>16</td>
</tr>
<tr>
<td>2.1 Dalton Mark VII-L for SDS Gel Electrophoresis (Sigma)</td>
<td>31</td>
</tr>
<tr>
<td>3.1 280:250 ratios from the gel filtration data for each protein</td>
<td>51</td>
</tr>
<tr>
<td>3.2 Yields of Prolactin and Prolactin Variants</td>
<td>54</td>
</tr>
<tr>
<td>3.3 Molecular Weights of Prolactin and Variants</td>
<td>65</td>
</tr>
<tr>
<td>3.4 ED$_{50}$ values for hPRL and hPRL mutants</td>
<td>75</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Amino acid sequence of hPRL</td>
<td>11</td>
</tr>
<tr>
<td>1.2</td>
<td>NMR structure of hPRL</td>
<td>12</td>
</tr>
<tr>
<td>1.3</td>
<td>The two binding sites of hPRL</td>
<td>13</td>
</tr>
<tr>
<td>1.4</td>
<td>Proposed binding mechanism of prolactin and prolactin receptor</td>
<td>14</td>
</tr>
<tr>
<td>1.5</td>
<td>hPRL with deleted residues 41-52 shown in black</td>
<td>15</td>
</tr>
<tr>
<td>3.1</td>
<td>Ion Exchange Chromatograms</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Gel filtration profiles</td>
<td>52</td>
</tr>
<tr>
<td>3.3</td>
<td>SDS-PAGE of WT hPRL</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>SDS-PAGE of G129R hPRL and Δ41-52 hPRL</td>
<td>56</td>
</tr>
<tr>
<td>3.5</td>
<td>SDS-PAGE of 2G mutant</td>
<td>57</td>
</tr>
<tr>
<td>3.6</td>
<td>SDS-PAGE of 3G mutant</td>
<td>58</td>
</tr>
<tr>
<td>3.7</td>
<td>SDS-PAGE of 4G mutant</td>
<td>59</td>
</tr>
<tr>
<td>3.8</td>
<td>SDS-PAGE of 5G mutant</td>
<td>60</td>
</tr>
<tr>
<td>3.9</td>
<td>SDS-PAGE of A mutant</td>
<td>61</td>
</tr>
<tr>
<td>3.10</td>
<td>SDS-PAGE of Turn mutant</td>
<td>62</td>
</tr>
<tr>
<td>3.11</td>
<td>SDS-PAGE of Δ42-48 hPRL</td>
<td>63</td>
</tr>
<tr>
<td>3.12</td>
<td>SDS-PAGE of Δ42-50 hPRL</td>
<td>64</td>
</tr>
<tr>
<td>3.13</td>
<td>UV spectra of prolactin and its variants</td>
<td>66</td>
</tr>
</tbody>
</table>
3.14 Fluorescence spectra of prolactin and its variants..........................69
3.15 CD spectra of prolactin and its variants........................................72
3.16 Alamar Blue dose-response curves..............................................76
CHAPTER 1

INTRODUCTION

1.1 Background: Human Prolactin (hPRL)

Human prolactin is a 23 kilodalton, 199 amino acid protein that is secreted by the pituitary gland. Physiologically, hPRL is involved in mammary gland growth and development, induction and regulation of lactation, immune regulation, osmoregulation, cell differentiation, and behavioral modification (1-3). The main regulator of circulating prolactin is dopamine, an inhibitor of pituitary prolactin secretion, which is produced in the hypothalamus. Levels of circulating prolactin have been controlled by treatment with dopamine and dopamine analogs (1).

The gene encoding for human prolactin is found on chromosome 6 (1). Prolactin is made up of a four alpha helix bundle with an up-up-down-down helical motif (1, 4). The molecule contains three disulfide bonds between cysteine residues 4 and 11, 58 and 174, and 191 and 199, respectively (3). The structure of the molecule has been solved using NMR (5) (please see figure 1.2 for the NMR structure of prolactin).
Prolactin is closely related to growth hormone (GH) and placental lactogen (PL), and these molecules have been grouped into the hematopoietic cytokine family (1). Cytokines are small secreted proteins produced de novo in response to external stimuli. They are used by the body to regulate immune response, inflammation, and hematopoiesis. The cytokines are grouped into families based on the type of receptor that they bind to. The human prolactin receptor and the properties of hematopoietic cytokine receptors are described in more detail in section 1.2.

1.2 Background: Human Prolactin Receptor (hPRLR)

The prolactin receptor is a membrane-bound protein that has both an extracellular and an intracellular domain (1, 2, 6). The receptor is found in long, short, and intermediate isoforms. Within a species only the length of the intracellular domain differs while the length of the extracellular domain remains constant (1, 2).

The gene encoding for human prolactin receptor is found on chromosome 5 (1). The prolactin receptor is considered a cytokine receptor and structurally resembles other receptors in this family including growth hormone receptor, although the two have very low sequence homology. Both receptors are made up of two antiparallel beta sheet extracellular domains called S1 and S2, each of which is composed of seven beta strands (6). Both PRL and GH receptors also contain four disulfide bonds in the N-terminal domain.
Prolactin receptor contains a WS motif (Trp-Ser-X-Trp-Ser, where X is any amino acid) in the C-terminal domain (6), which is a common feature of the hematopoietic cytokine family. In addition, the human prolactin receptor contains a free cysteine near the C-terminus of the extracellular domain.

The prolactin receptor can bind to prolactin, placental lactogen, and growth hormone. Binding of these ligands activates the receptor. Once activated, the prolactin receptor uses various associated kinases to relay its message (7) leading to somatotropic and lactogenic cellular response. Some of the most well studied kinase cascades involved with the prolactin receptor are the Jak/Stat pathway, the Ras-Raf-MAPK pathway, and the Scr tyrosine kinases (1, 3, 6-12), but little is known about the intracellular domain signaling of the prolactin receptor.

1.3 Prolactin/Prolactin Receptor Mechanism of Binding

Prolactin binds to the prolactin receptor in a 1:2 stoichiometry (2, 4, 13). Two distinct binding sites, referred to as binding site 1 and binding site 2, have been identified in prolactin (13). Binding site 1 is located in a pocket created by helix 1, helix 4, and loop 1 (4). Binding site 2 is located between helices 1 and 3 (32). Important residues for binding at binding site 1 are lysine 181 and arginine 177 and the most important residue for site 2 binding is glycine 129 (4, 14, 15) (see figure 1.3).

The currently accepted theory is that prolactin binds one receptor molecule through binding site 1 before undergoing a conformational change which exposes binding site 2 for binding, and then binds a second receptor
molecule via binding site 2 (4, 6, 16). Please see figure 1.4 for a representation of this binding mechanism. This method of receptor binding is similar to other receptors in the cytokine family (6, 13, 14, 17-22).

1.4 The Role of Prolactin in Breast Cancer

Various human breast cancer cell lines have been identified which contain prolactin receptors on the cell surface (2, 23, 24). Knockout mice lacking the prolactin gene have been shown to have a lower incidence of mammary tumors (25). Additionally, studies in humans have shown that breast cancer risk is correlated with high serum prolactin levels (8, 26, 27). Prolactin has been found to up-regulate the expression of \textit{BRCA1}, a breast cancer susceptibility gene (28). These observations led to the belief that prolactin may be an important mitogen for breast cancer cells. The fact that treatment with dopamine agonists, which drastically lower levels of circulating prolactin, did not affect breast cancer growth seemed to disprove the involvement of prolactin in breast cancer (10, 29). It was not until it was discovered that breast cells and breast tumor cells produce their own prolactin (2, 8, 10, 23, 24, 29) that prolactin was again considered to stimulate tumor growth. This discovery led to the development of prolactin antagonists which attempt to starve the cancer cells of prolactin in an effort to kill the cells.

Prolactin and the prolactin receptor are also present in other types of cells including lymphocytes and prostate cells and may have mitogenic affects on these types of cells and the associated cancers as well (1, 3, 7, 30).
1.5 Current Prolactin Antagonists

The most popular method for designing prolactin antagonists involves creating a mutant form of the protein with an intact binding site 1 and a non-functional binding site 2. The most common antagonist in use contains a substitution of an arginine for glycine 129 (G129R hPRL). G129R hPRL was designed based on a corresponding mutation in human growth hormone, G120R hGH, originally proposed and patented by Genentech, Inc. (South San Francisco, CA) (31). The rationale for this mutation is that glycine 129 is important in site 2 binding based on its small size and neutral charge, so replacing it with a large positively charged residue should interfere with its ability to bind.

Another important prolactin antagonist was designed by Dr. Francis Peterson in Dr. Charles Brooks laboratory (32). This antagonist also employs the strategy of interfering with site 2 binding, but instead of replacing a residue it contains a twelve amino acid deletion that prevents the molecule from undergoing the conformational change necessary to make site 2 available for binding. This mutant was designed using sequence comparisons to a naturally occurring human growth hormone deletion mutant that had lost lactogenic activity but retained somatotrophic activity. The homologous sequence was deleted in hPRL, resulting in Δ41-52 hPRL which had reduced lactogenic and activity. The deleted amino acids are residues 41-52. The Δ41-52 hPRL mutant is pictured in figure 1.5.

One problem with Δ41-52 hPRL is that its expression results in a large incidence of biologically inactive dimer (making up greater than 50% of the yield).
This is believed to be caused by the formation of a disulfide linkage between two monomers. The deleted region of this protein lies on the region C-terminal to helix 1 and N-terminal to the loop region following helix 1. The loop region contains cysteine 58, which is involved in a disulfide bond with cysteine 174 on helix 4. It is thought that the deletion of these residues causes a strain on this central disulfide bond, making it difficult to form during the refolding process, and therefore leaving cysteine residues that would normally be involved in a disulfide bond available for dimer formation. A much greater monomer yield and more easily folded protein would be highly desirable if this molecule were ever to be used in therapy. Mass spectrometric analysis studying the nature of disulfide bonds in dimeric Δ41-52 hPRL is currently being pursued (Troyer and Brooks, unpublished).

Both G129R hPRL and Δ41-52 hPRL have lowered agonist activity and increased antagonist activity on breast cancer cells compared to WT hPRL. G129R hPRL has been shown in several studies to antagonize the prolactin receptor both in vivo and in vitro (33-36). One drawback associated with this mutant is that while G129R hPRL acts as an antagonist, it still maintains 1-10% of the wildtype compound's agonist activity which makes the remaining agonist activity significant when doses sufficient to bind to all the prolactin receptors on the cell surface are used (Brooks, unpublished). However, G129R hPRL is extremely important because it shows that, despite residual agonist activity, it is possible to engineer prolactin antagonists for possible use in breast cancer therapy.
It has been shown by measuring elevated cleaved caspase levels that Δ41-52 hPRL induces apoptosis in human breast cancer cells (Almgren and Brooks, unpublished). Both G129R hPRL and Δ41-52 hPRL have also been shown to lower ATP levels in breast cancer cells with respect to WT hPRL (Almgren and Brooks, unpublished). In addition, Δ41-52 hPRL and G129R hPRL have been shown to induce apoptosis in FDC-P1 cells, with Δ41-52 hPRL inducing at least 50% more apoptosis than G129R hPRL (Almgren and Brooks, unpublished). Δ41-52 hPRL has also been shown to have antagonistic properties in the NB2 rat lymphoma cell line (Patmastan and Brooks, unpublished). Additionally, hPRL antagonists have been shown to effect other types of cells besides breast cancer cells. Research has shown that Δ41-52 hPRL induced apoptosis in the Jurkat human lymphocyte cell line (Woodrich and Brooks, unpublished).

A third prolactin antagonist, S179D hPRL, has also been designed but with a different strategy than the other antagonists. This antagonist replaces serine 179 with an aspartic acid residue and was designed to mimic the phosphorylated form of prolactin (37). Serine 179 is located on helix 4 and its location would predict that it is involved in the binding site 1 interface, so the mode of action of this mutant protein is unclear. Whether S179D hPRL actually acts as a prolactin antagonist or not is a subject of much debate. Some studies have shown that it functions as an antagonist (38-42) while others have contradicted this fact, showing not only that S179D hPRL has no antagonist activity, but it has increased agonist activity (43). Regardless of the activity of
this particular mutant, the fact that prolactin has been mutated into several different potential antagonists is a positive step towards its use as treatment for breast cancer.

1.6 Mutant Proteins Designed For This Project

The goal of this project was to make further modifications to the Δ41-52 hPRL molecule that would enable it to fold more easily, contain less dimer contamination, and still maintain the antagonist activity of Δ41-52 hPRL. To do this I designed eight proteins. I operated under the assumption that the deletion of such a large number of amino acids was straining the cysteine 58 to cysteine 174 disulfide bond so I made several mutations that involved adding residues to this region.

The first series of mutants involved adding a series of glycine residues to act as a spacer in the deleted region. I made mutants that had 2, 3, 4, and 5 glycine residues added to the deleted region and named them Δ41-52 + 2G hPRL, Δ41-52 + 3G hPRL, Δ41-52 + 4G hPRL, and Δ41-52 + 5G hPRL, respectively. (These will sometimes be referred to as 2G, 3G, 4G, and 5G in this manuscript.) Glycine was chosen as a spacer because it is a small, neutrally charged molecule and should not cause any steric problems in the molecule or have any negative electrostatic effects on the protein structure.

Two other mutants were made with similar spacers. One involved adding two alanine residues to the deleted region to compare the difference between using alanine and glycine as a spacer. This mutant was called Δ41-52 + 2A hPRL, but will sometimes be referred to as A mutant. Another mutant had a beta
turn added as a spacer, with the idea being that the turn should mimic the larger turn in this region in the wildtype protein. The beta turn consisted of adding the residues serine-proline-glycine-glycine to the amino acid 41-52 region. This mutant was named Δ41-52 + Turn hPRL and will also be referred to as turn mutant or T mutant.

Finally, two mutants were created by adding residues present in the wildtype hPRL molecule back to the deleted region. This was done by looking at the prolactin molecule using molecular modeling and choosing residues that were spatially closer together than residues 40 and 53. It was hoped that this would cause less strain on the molecule. The first mutant added residues 41, 49, 50, 51, and 52 (connecting residues 41 and 49). This mutant was named Δ42-48 hPRL. The other mutant added residues 41, 51, and 52 (connecting residues 41 and 51), and was named Δ42-50 hPRL. All of these mutants are represented in table 1.1. In this table, residues 35 – 55 of the wildtype molecule are shown with the corresponding amino acid changes for each mutant.

1.7 Protein Expression, Folding, and Characterization

The mutant proteins were expressed in *E. coli* BL21 cells using the lac Z promoter induced with IPTG (isopropylthiogalactoside) and pelleted as an inclusion body. They were resolubilized in urea, air oxidized, and then allowed to refold during dialysis against TRIS (tris (hydroxymethyl) aminomethane). They were purified using anion exchange chromatography. Proteins exhibiting a 280 / 250 ratio of between 1.8:1 and 2.2:1 was collected from the column (the expected ratio for human prolactin is 2:1). The next step was to separate the
monomers from the dimers using gel filtration chromatography. The monomer and dimer peaks were then dried and run on an SDS-PAGE gel under both reducing and non-reducing conditions to determine the relative amount of dimer contamination, if any, in the monomer fractions. The monomers with the least amount of contamination were then scanned over a range of wavelengths for ultraviolet absorbance and fluorescence to determine proper folding. Finally, a circular dichroism spectrum was obtained for each monomer to confirm the presence of the correct secondary structure.

1.8 Assessment of Biological Activity

The agonist activity of the prolactin mutants was measured using an Alamar Blue biological assay (see section 2.10). The $\text{ED}_{50}$, or effective does to obtain half maximal cellular response, was obtained from this assay. A protein was considered a poorer agonist than wild-type prolactin if it had an increased $\text{ED}_{50}$. The shape of the dose-response curve obtained from this assay is also indicative of the mechanism of action of the prolactin antagonist.
Figure 1.1: Amino acid sequence of hPRL (from the RCSB Protein Data Bank, protein ID# 1N9D). Amino acids involved in helices are underlined.
Figure 1.2: NMR structure of hPRL (PDB # 1N9D)
Helix 1
(residues 1-43)

Helix 2
(residues 77-105)

Helix 3
(residues 113-136)

Helix 4
(residues 163-189)

Receptor Binding Site 1

Receptor Binding Site 2

Figure 1.3: The two binding sites of hPRL (courtesy of Jeff Voorhees)
Figure 1.4: Proposed binding mechanism of prolactin and prolactin receptor
Figure 1.5: hPRL with deleted residues 41-52 shown in black
<table>
<thead>
<tr>
<th>Protein</th>
<th>Native Residues Numbers 35 - 55 with additions in italics</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT hPRL</td>
<td>EMFSEFDKRYTHGRGFITKAI</td>
</tr>
<tr>
<td>Δ41-52 hPRL</td>
<td>EMFSEFKAI</td>
</tr>
<tr>
<td>2G mutant</td>
<td>EMFSEFGGKAI</td>
</tr>
<tr>
<td>3G mutant</td>
<td>EMFSEFGGGKAI</td>
</tr>
<tr>
<td>4G mutant</td>
<td>EMFSEFGGGGKAI</td>
</tr>
<tr>
<td>5G mutant</td>
<td>EMFSEFGGGGGKAI</td>
</tr>
<tr>
<td>A mutant</td>
<td>EMFSEFAAKAI</td>
</tr>
<tr>
<td>Turn mutant</td>
<td>EMFSEFSPGGKAI</td>
</tr>
<tr>
<td>Δ42-48 hPRL</td>
<td>EMFSEFDGFI TKAI</td>
</tr>
<tr>
<td>Δ42-50 hPRL</td>
<td>EMFSEFDI TKAI</td>
</tr>
</tbody>
</table>

Table 1.1: Human prolactin mutant sequences. Residues that were added to the sequence of Δ41-52 hPRL are shown in italics.
CHAPTER 2

MATERIALS AND METHODS

This chapter contains the materials and methods that were used in the following chapters. The details of general techniques such as storage and growth of E. coli, transformation, agarose gels, common media preparation, and preparation of competent cells can be found in Current Protocols of Molecular Biology (44) and will not be described in this chapter. Phagemid DNA amplification was done using Promega amplification kits. All methods discussed in this chapter have been performed by myself unless otherwise noted.

2.1 The pT7-7 Phagemid

All protein expression, DNA expression and site-directed mutagenesis on human prolactin (hPRL) and its variants were done using the pT7-7 phagemid. This phagemid was originally provided by S. Tabor (Harvard Medical School, Boston, MA) in the form of a plasmid. Later, Dr. Francis Peterson incorporated a f1 origin of replication into the plasmid to facilitate production of single-stranded DNA needed for the Kunkel method of site-directed mutagenesis (45). The original plasmid with the new ability to produce single-stranded DNA is now
called the pT7-7f(-) phagemid (32). The phagemid contains a bacterial origin of replication, a gene for ampicillin resistance, a T7 RNA promoter, a multiple cloning site (where the prolactin gene was inserted), and an f1 origin of replication. The DNA encoding for hPRL was inserted into the multiple cloning site by Peterson (45).

2.2 Single-Stranded DNA Preparation

In order to perform site-directed mutagenesis by the Kunkel method, the DNA to be mutated needed to be obtained in a single-stranded uracilated form. The DNA that was used for site-directed mutagenesis was the human prolactin variant Δ41-52 hPRL. This variant was designed by Dr. Peterson, and contains a twelve amino acid deletion of residues 41-52 of hPRL (32). This deletion was done using the Kunkel method, which is described in section 2.3. The Δ41-52 hPRL DNA resides on the pT7-7 phagemid described in section 2.1.

To obtain uracilated single-stranded DNA, the pT7-7 phagemid containing the Δ41-52 hPRL DNA was transformed into the Escherichia coli strain RZ1032. This strain is deficient in the genes dut and ung (in other words, the strain is dut ung). The dut gene is responsible for expressing the dUTPase enzyme, which hydrolyzes dUTP in the cell so it cannot be misincorporated into the DNA strand in place of dTTP (46). The ung gene codes for the enzyme N-glycolsylase, which is responsible for removing uracils that have been incorporated into DNA (46). The RZ1032 strain, therefore, has the ability to incorporate uracils into DNA.
Once colonies of RZ1032 E. coli bearing the Δ41-52 hPRL containing pT7-7 phagemid were obtained, one of these colonies was inoculated into 10 ml of LB media and grown overnight at 37°C with shaking. This culture was inoculated into 50 ml of 2xTY media and grown at 37°C with shaking until an OD$_{600}$ of 0.3 was obtained. At this point, R408 helper phage at an M.O.I. of 20 (purchased from Fisher) was added and the culture was allowed to continue shaking overnight. The phage is added to induce expression of the inserted DNA in a single-stranded form. The single-stranded Δ41-52 hPRL DNA is secreted by the bacteria into the solution. The solution was centrifuged at 10,000 X g for 15 minutes for clarification and incubated on ice with polyethylene glycol/sodium chloride solution (150 grams of polyethylene glycol and 146 grams of sodium chloride in 1 liter of water) for one hour before centrifuging at 5000 X g for 15 minutes to pellet the phage. The supernatant was discarded and the phage was treated with a high-salt extraction buffer containing 100 mM TRIS, pH 8.0, 300 mM NaCl, and 1 mM EDTA to extract the single-stranded DNA from the phage. The DNA (in solution) was separated from the phage (pellet) by centrifugation at 14,000 rpm in an Eppendorf 5415C model centrifuge for 15 minutes and then extracted using a phenol solution containing 50 ml buffer saturated phenol, 48 ml chloroform, 2 ml iso-amyl alcohol, and 0.1 g 8-hydroxyquinoline. The DNA was precipitated using ice-cold ethanol, dried, resuspended in deionized nuclease-free water, and stored at -20°C. The purity and concentration of the single-stranded phagemid were determined by absorbance at 260 and 280 nm.
2.3 Primer Design

Several primers were designed to insert various amino acids into the deleted region of Δ41-52. This was done by designing the added nucleic acid sequence, then flanking this sequence by 10-12 nucleotides on either side to anchor it down to either side of the added codons during the hybridization step of mutagenesis.

The following mutagenic primers were designed. The inserted nucleotides are indicated by bold lettering.

The first group of mutations adds a series of glycines. Primers were made to add 2 glycines, 3 glycines, 4 glycines, and 5 glycines between residues 40 and 53.

2 glycine (2G) primer: GTT GAT GGC CTT GCC GCC GAA TTC GCT GAA
3 glycine (3G) primer: GTT GAT GGC CTT GCC GCC GCC GAA TTC GCT GAA
4 glycine (4G) primer: GTT GAT GGC CTT GCC GCC GCC GCC GAA TTC GCT GAA
5 glycine (5G) primer: GTT GAT GGC CTT GCC GCC GCC GCC GCC GAA TTC GCT GAA

Two other mutations involved adding other types of sequences. One, named "A mutant", added a series of two alanine residues into the deleted
A mutant:  GTT GAT GGC CTT CGC CGC GAA TTC GCT GAA
Turn mutant:  GTT GAT GGC CTT GCC GCC CGG GCT GAA TTC GCT GAA

The last group of mutations added several native residues to this region based on observing the structure on molecular modeling programs and finding residues that were closer together in space than residues 40 and 53, and therefore may be less likely to cause stress to the molecule when refolding. The first of these mutants, Δ42-48 hPRL, added residues 41 (aspartic acid), 49 (glycine), 50 (phenylalanine), 51 (isoleucine), and 52 (threonine). The other mutant, Δ42-50, added residues 41 (aspartic acid), 51 (isoleucine), and 52 (threonine).

Δ42-48 hPRL primer:  GTT GAT GGC CTT GGT GAT GAA GCC ATC GAA TTC GCT GAA
Δ42-50 hPRL primer:  GTT GAT GGC CTT GGT GAT ATC GAA TTC GCT GAA

2.4 Site-Directed Mutagenesis

All mutagenesis was performed using the Kunkel method (46). In this method, the primers containing the desired mutation are first phosphorylated. The phosphorylated primer is annealed to the single-stranded DNA to be
mutated. The single-stranded DNA was made to contain uracils in place of thymines. The primer is elongated in vitro. The now double-stranded DNA, containing the uracilated parental strand and the non-uracilated mutated strand, is transformed into a dut∗ung⁺ (see section 2.2 for definition of the functions of these genes). The dut∗ung⁺ E. coli will destroy the uracilated strand, preferably extending the mutated strand.

The mutagenesis was done in four steps. In the first step, the primer was phosphorylated on its 5' end to aid in ligation by DNA ligase in a subsequent step. This was done by mixing 1 µl of 10X PNK (polynucleotide kinase) Buffer (New England Biolabs), 1 µl of 4 mM ATP, 1 µl of 5 µM primer (a separate reaction was done for each primer described in section 2.3), and 0.5 µl of 10 U/µl T4 phosphonucleotide kinase (New England Biolabs). The mixture was brought to 10 µl with water and incubated for 1 hour at 37°C before heat inactivating the enzyme by incubating the mixture at 55°C for 5 minutes.

The second step is annealing the primer to the single-stranded DNA described in section 2.2. This was done by mixing 2.5 µl of the phosphorylated primer from the first step with 1 µl of T7 Polymerase buffer (New England Biolabs) and 125 ng of single-stranded DNA and bringing the mixture to a total volume of 5 µl with water. This mixture was heated to 94°C and slowly cooled to 35°C using an MJR thermocycler.

Once the primer was annealed to the single-stranded DNA template, it could be elongated. This was done by mixing 5 µl of the annealed
template/primer, 1 µl of T7 Polymerase buffer, 0.5 µl of 10mM dNTP’s, 0.25 µl of 5U/µl T7 DNA Polymerase (New England Biolabs), 1 µl of 1U/µl T4 DNA Ligase (New England Biolabs), and 12.25 µl of water to provide a 20 µl reaction volume, and incubating this reaction mixture at 37°C for 2 hours. At the end of this step, there was double stranded DNA containing one mutated (non-uracilated) strand and one parental (uracilated) strand.

The final step of this procedure was to transform the double-stranded DNA into a dut* ung* bacteria strain. The E. coli strain used was Gold Competent cells from Stratagene. Colonies were selected on the basis of ampicillin resistance and expanded in LB-media. Plasmid DNA was prepared using a Promega amplification kit.

The DNA sequences of the hPRL mutants were confirmed by DNA sequencing by The Ohio State University Plant-Microbe Genomics Facility to assure that the desired mutations were correctly added to the DNA encoding the mutant proteins.

2.5 Protein Expression

All proteins were expressed using BL21(DE3) E.coli. The specific plasmid DNA was transformed into the E. coli and cloned on LB-agar plates containing ampicillin and colonies were selected and inoculated into a 10 ml starter culture of LB media at 37°C with shaking. After growth was apparent, the starter culture was transferred to a 1 L culture of LB media and allowed to grow to an OD₆₀₀ of 0.3 at which point expression was induced by IPTG for four hours. The bacteria
were collected by centrifugation for 10 minutes at 6371 X g and resuspended in 45 ml of 100 mM TRIS, pH 7.5, 25 mM DTT (DL-dithiothreitol), and 1 mM PMSF (Phenylmethylsulphonylfluoride). The resuspended pellet was pressed twice through a French Press at 5000 psi to lyse the cells. This suspension was centrifuged at 23975 X g for 30 minutes to collect the protein in the form of an inclusion body, which was suspended in 100 ml of 4.5 M urea and 100 mM TRIS (solution adjusted to pH 11.5) and allowed to solubilize for 2 hours. The solubilized protein was centrifuged for 90 minutes at 24371 X g to pellet protein aggregates and remaining cell matter. The protein was allowed to air oxidize for two days with stirring at 4°C. It was dialyzed against 4 L of 20 mM TRIS (first exchange at pH 11.5 and the rest at pH 7.5) for six exchanges total to remove the urea.

2.6 Protein Purification (Ion Exchange Chromatography and Gel Filtration Chromatography)

After TRIS dialysis, the protein was centrifuged at 23975 X g for 15 minutes to remove any particulate matter. It was purified by anion exchange chromatography using a DEAE fast flow Sepharose column with 20 mM TRIS pH, 7.5 and developed in the same buffer with a zero to 0.5 M NaCl gradient. The UV absorbance at 250 nm and 280 nm was monitored. The protein was collected when there was a 280/250 ratio of 2, although this ratio was not always achieved. After the protein was collected, it was dialyzed against 10 mM ammonium bicarbonate to remove the sodium chloride ions for 1 to 2 exchanges.
The protein was loaded onto a gel filtration column (Superdex 75 resin in a 5 cm diameter, 45 cm length column) to separate the monomer form from the dimer form. The buffer was 10 mM ammonium bicarbonate and 10 ml fractions were collected. The 280 nm, 250 nm, and 340 nm UV absorbances were read and graphed. Two distinct peaks could be drawn from all of the prolactin variants. The fractions in each peak were combined, frozen completely, lyophilized, and stored at 4°C under desiccation.

2.7 SDS-PAGE (Sodium dodecyl sulfate polyacrylamide gel electrophoresis) Analysis

All proteins were characterized for monomer and dimer content using SDS-containing polyacrylamide gel electrophoresis in a Hoeffer gel apparatus SE600 (14 x 12 cm gel). The dried protein was first resuspended in 10 mM ammonium bicarbonate and the protein concentration was determined using the BCA (bicinchoninic acid) method (47). 20 µg of protein were loaded onto each well. Each monomer and dimer sample from the gel filtration column was analyzed on the polyacrylamide gel using reducing (6% 2-mercaptoethanol, 15% glycerol, 9% SDS, pyronin Y) and non-reducing (15% glycerol, 9% SDS, pyronin Y) sample buffers. Each protein sample was mixed with an equal volume of the appropriate loading buffer and incubated in boiling water for five minutes before being loaded onto the gel. A molecular weight marker was also loaded onto every gel to aid in molecular weight determination. Please see table 2.1 for the proteins in the molecular weight marker. The stacking gel was 4% acrylamide mixed with 0.125 M TRIS, pH 6.8 containing 0.4% SDS.
Polymerization was initiated using ammonium persulfate and catalyzed by TEMED (N,N,N',N'-tetramethyl-ethylenediamine). The separating gel used was 15% acrylamide mixed with 0.375 M TRIS, pH 8.8 containing 0.4% SDS. Again, polymerization was initiated using ammonium persulfate and catalyzed by TEMED. Gels were run at 75 milliamperes per plate, stained with Coomassie Blue, and destained with 10% acetic acid, 10% isopropanol and visualized on an Alpha Imager.

2.8 Mass Spectrometry

To prepare each protein for a whole protein mass spectrometric analysis, the protein was suspended in 10 mM ammonium bicarbonate pH 7.0. The protein concentration was measured using the BCA method (47) and brought to a concentration of 40 µg protein per 20 µl solution with 0.1% trifluoroacetic acid. A peptide microtrap (Michrom BioResources, Inc, 50 kilodalton molecular weight cutoff) was washed three times with 25 µl each of methanol, 95% acetonitrile, 50% acetonitrile, and 0.1% trifluoroacetic acid, in that order. 10 µl of the protein sample were passed through the trap and washed three times with 25 µl of 0.1% trifluoroacetic acid. The protein was eluted using two 20 µl injections of 50% acetonitrile followed by a 10 µl injection of 95% acetonitrile. The whole protein molecular weight was measured by electrospray induction quadrupole time of flight mass spectrometry by Ben Jones at The Ohio State University Campus Chemical Instrumentation Center, Mass Spectrometry and Proteomics facility.
2.9 Ultraviolet Absorbance Spectroscopy

The UV absorbance spectrum of hPRL and its variants was used to characterize the protein. The dried protein was solubilized in a spectroscopy buffer consisting of 10 mM TRIS pH 8.2 and 150 mM NaCl and the concentration was measured using the BCA assay (47). The protein was brought to a concentration of 20 µM and a wavelength scan was run on Perkin Elmer Lambda 45 UV/Vis Spectrometer from 220 nm to 350 nm.

2.10 Fluorescence Spectroscopy

The fluorescence spectrum of hPRL and its variants was also used to characterize the protein. The same buffer was used as in section 2.7, but a 1µM protein concentration was used. The fluorescence spectrum was measured on a Perkin Elmer LS55 Luminescence Spectrometer, using 285 nm as an excitation wavelength and scanning the emission spectrum from 300 nm to 400 nm.

2.11 Circular Dichroism

Circular Dichroism (CD) was also used to observe the secondary structures of the various prolactin mutants. The CD spectrum was measured on a 25 µM solution in the same buffer as was used in sections 2.7 and 2.8. The protein solution was measured in a 1 mm cuvet using an Aviv Model 20 Circular Dichroism Spectrometer. Water was used as a blank. The wavelengths were scanned from 200 nm to 260 nm.

2.12 Alamar Blue Cell Assay

To measure the agonist activity of the prolactin variants, a cell assay measuring the effect of various doses of the proteins on FDC-P1 cells expressing
the human prolactin receptor was used. These cells were originally received as a generous gift from Genentech, Inc. (South San Francisco, CA) and have the human prolactin receptor tranfected in so the cells express the receptor on their membranes. The cells have been selected for prolactin response through growth in media containing 1 nM hPRL and lacking other mitogens. This assay measured the relative amount of Alamar Blue reduced by the cells in each treatment. Live cells metabolize Alamar Blue in the mitochondria, so the amount of Alamar Blue metabolized is directly related to the number of living cells for each treatment.

The FDC-P1 cells were grown in maintenance media of RPMI 1640 with phenol red containing 10% fetal bovine serum, 50,000 U penicillin-streptomycin antibiotic, and 1 nM human prolactin. Upon achieving log phase growth, the cells were starved for 24 hours in starvation media (RPMI 1640 without phenol red containing 10% Gelding horse serum from Hyclone Laboraties in Ogden, UT, and 50,000 U penicillin-streptomycin). The cells were washed with wash media (RPMI 1640 without phenol red), resuspended in starvation media, and diluted to 15,000 cells/50 µl (300,000 cells/ml). 50 µl were carefully pipetted into the middle 60 wells of a sterile 96 well tissue culture plate with a lid. The bordering 36 wells were filled with 100 µl of water because when this assay had been done in the past the wells bordering the edges of the plate had suffered some evaporation during incubation. The protein treatments and one blank were each done in triplicate wells. Protein that had originally been suspended in ammonium bicarbonate was diluted to the appropriate concentrations with starvation media.
and added in 50 µl doses to each well. Final hormone concentrations varied between 0.01 nM and 10,000 nM using three doses per order of magnitude. Once the protein was added, the plates were mixed by placing on a vortex set at high speed to distribute the protein and break up cell clumps. The plates were incubated for 48 hours at 37°C in a 5% carbon dioxide, 95% air environment. Each day during the incubation, the plates were mixed as on day 1. After the 48 hour incubation, 10 µl of Alamar Blue were added to each well (excluding the wells containing water only). The plates were again mixed and placed back into the incubator. Readings of the OD$_{570}$ and OD$_{600}$ were taken after 2, 3, and 4 hours of incubation. The data for the 3 hour time point only is presented in chapter 3. These readings were used to determine the percent of Alamar Blue reduced using the following equations.

Alamar Blue absorbs light at both 570 nm and 600 nm, with the peak of interest for these calculations being 570 nm. However, the 570 nm absorption and the 600 nm absorption overlap, so one needs to take this into account and subtract the absorption at 600 nm from the absorption at 570 nm. The extinction coefficient for reduced Alamar Blue at 570 nm is 155677 cm$^{-1}$M$^{-1}$ and for oxidized Alamar Blue at 570 is 80586 cm$^{-1}$M$^{-1}$. The extinction coefficient for reduced Alamar Blue at 600 nm is 14652 cm$^{-1}$M$^{-1}$ and for oxidized Alamar Blue is 117216 cm$^{-1}$M$^{-1}$. The equation for determining the concentration of reduced Alamar Blue at each hormone dose is as follows:

\[ C_{\text{red}} = \frac{(117216 \times A_{570} - 80586 \times A_{600})}{(155677 \times 117216 - 80586 \times 14652)} \]
The equation for the concentration of oxidized alamar at each hormone concentration is:

\[ C_{ox} = \frac{155677 \cdot A_{600} - 14652 \cdot A_{570}}{155677 \cdot 117216 - 80586 \cdot 14652} \]

The percent of Alamar Blue in the reduced state at each hormone concentration can then be calculated using the concentration of Alamar Blue in the reduced state and the average concentration of Alamar Blue in the oxidized state for the negative control using the following equation.

\[ \% \text{ Alamar Reduced} = \frac{C_{\text{red}}}{\text{avg } C_{ox} \text{ of negative control}} \]

To obtain the \( ED_{50} \) measurement, or the dose of hormone at which the reduced Alamar Blue is at half maximal response, the percent of Alamar Blue reduced was graphed against the log of the hormone dose. The SigmaPlot four parameter fit was used to extrapolate the \( ED_{50} \) values.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Approximate Molecular Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin, bovine</td>
<td>66,000 Daltons</td>
</tr>
<tr>
<td>Albumin, egg</td>
<td>45,000 Daltons</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate Dehydrogenase, rabbit muscle</td>
<td>36,000 Daltons</td>
</tr>
<tr>
<td>Carbonic Anhydrase, bovine</td>
<td>29,000 Daltons</td>
</tr>
<tr>
<td>Trypsinogen, bovine pancreas</td>
<td>24,000 Daltons</td>
</tr>
<tr>
<td>Trypsin Inhibitor, soybean</td>
<td>20,100 Daltons</td>
</tr>
<tr>
<td>α-Lactalbumin, bovine milk</td>
<td>14,200 Daltons</td>
</tr>
</tbody>
</table>

Table 2.1: Dalton Mark VII-L for SDS Gel Electrophoresis (Sigma)
CHAPTER 3

RESULTS

3.1 Anion Exchange Chromatography

The first piece of data collected for the mutant proteins was the UV absorbance spectrum observed during purification of the protein using ion exchange chromatography. The protein was collected when the 280 nm : 250 nm absorbance ratio was between 1.8:1 and 2.2:1. The protein ideally should have a 280:250 ratio of 2:1 based on previous studies with this protein. The height of the 2:1 peak should correlate to the relative amount of protein in the sample. Any other peaks are regarded as contaminant proteins. Also, it is important to note that both the monomeric and dimeric forms of hPRL will give the same 280:250 ratio. Figure 3.1 shows a representative chromatogram for each protein.

The expected behavior for WT hPRL is described above. Most of the prolactin variants also had a clear 280:250 peak with a ratio of 2:1. Δ41-52 PRL has a less distinct 280:250 peak (see figure 3.1). The 3G and A mutants had peaks that mimicked Δ41-52 hPRL. The Turn mutant and the 2G mutant had even smaller useable peaks, and also had irregularly shaped peaks.
These irregularities could be interpreted as poor folding in the deviant proteins evidenced by the 250 nm reading. The 250 nm absorbance is sensitive to disulfide bond stretching, bending, and turning. On the other hand, the deviant proteins could simply be contaminant proteins. Their true identity was further resolved using SDS-PAGE and Mass Spectrometric analysis.

Additionally, the maximum height of the 280:250 peak can be related to the amount of protein that is being eluted from the column. The WT hPRL had a maximum reading of 300 milliAbsorbance units (mAU). In general, the other proteins stayed near to this value with the maximum absorbance for any of the proteins being 600 mAU by Δ42-48 hPRL and the minimum being an outlier of 10 mAU by the 2G mutant. It is important not to read too much into the maximum absorbence as each of the proteins were purified on separate columns which could act slightly different due to differences in the matrix of each column.

3.2 Gel Filtration Chromatography

The protein that was purified from the ion exchange column was dialyzed against ammonium bicarbonate and then loaded onto a gel filtration column to separate the monomeric protein from the dimeric protein. The UV absorbencies for the fractions that were collected were read at 280 nm and 250 nm. If a protein sample contains a mixture of monomeric and dimeric protein, it is expected to have two separate peaks, the first peak being the larger (dimeric) protein and the second peak being the smaller (monomeric) protein. If a protein is pure monomer it is expected to have only one single peak.
The gel filtration data for WT hPRL shows that WT hPRL consistently contains one major protein species, presumably monomer (see figure 3.2 for the gel filtration profiles). An extremely small dimer peak was collected for analysis as well. Section 3.4 will use SDS-PAGE to demonstrate that this single species is indeed monomeric protein. Additionally, table 3.1 displays the 280:250 ratio of each protein's monomeric and dimeric species as calculated from the graphs in figure 3.2. As previously stated, this ratio should ideally be approximately 2 and indicates proper folding of the protein. For wildtype hPRL the monomer ratio is 3 according to this data so the rest of the proteins will be compared to this value.

G129R hPRL is similar to WT hPRL in its gel filtration profile. It also has one species present that will be analyzed by SDS-PAGE in section 3.4 as well as a miniscule dimer peak that will be analyzed. The 280:250 ratio for the G129R hPRL monomer is 2.8 which is extremely close to the value for WT hPRL.

Δ41-52 hPRL contains two distinct protein species. Since the dimeric protein is larger than the monomer protein it is contained in the first peak. The gel filtration profile shows that Δ41-52 hPRL has about 20-50% more dimeric species present than monomeric species. The 280:250 ratio for Δ41-52 hPRL is 2, which is 2/3 of the value for WT hPRL. This may indicate strain within the molecule.

During the time that the 2G mutant data was being collected, the gel filtration column was experiencing several problems including dry matrix and bad tubing. The irregularities in the peaks from the 2G mutant samples were attributed to these problems (see figure 3.2D). The fractions from the various
peaks were collected and analyzed by SDS-PAGE to determine which peaks contained the monomeric and dimeric portions (see section 3.4). Although the 2G gel filtration profiles are difficult to interpret alone, combining them with the results from SDS-PAGE analysis shows that the 2G mutant contains up to 75% more monomer than dimer. Additionally, the 280:250 ratio is about the same as that for Δ41-52 hPRL, indicating that the 2G molecule is just as strained as Δ41-52 hPRL.

The 3G mutant contained two distinct species and contained at least 50% as much monomer as dimer. However, the monomer had a 280:250 ratio closer to WT hPRL than 2G, so perhaps the folding is improved on this molecule.

The 4G mutant appeared to contain mostly monomer, and had a small dimer peak. The 280:250 ratio is close to the value for Δ41-52 hPRL, however, which could indicate a strained or poorly folded protein.

For the most part, the 5G mutant contains one single predominant 280 over 250 peak for analysis using SDS-PAGE but at this point, it was assumed to contain monomeric protein. The 280:250 ratio was similar to WT hPRL indicating a well-folded molecule.

The A mutant varied in amount of dimer versus monomer present but all of the samples contained at least 50% dimer. The monomer had a ratio similar to WT hPRL.

The T mutant contained two distinct forms of protein but each sample showed that approximately 50% more monomer was present than dimer. The 280:250 ratio was low, similar to Δ41-52 hPRL, indicating a strained molecule.
Δ42-48 hPRL was inconsistent, but overall there were approximately equal amounts of monomer and dimer present. In addition, the 280:250 ratio resembles that of Δ41-52 hPRL more than WT hPRL.

Δ42-50 hPRL had two distinct yet overlapping peaks but overall contained at least 50% more monomer than dimer and had a 280:250 ratio very similar to WT hPRL. Additionally, the percent yields of dimer and monomer were calculated by integrating the areas under the curves for the gel filtration profiles (see table 3.2).

3.3 Protein Yields

The yield of each prolactin variant is reported in milligrams and as a percent of the total yield in table 3.2. The proteins with the highest yields were WT hPRL and G129R hPRL. These two also both contained approximately 90% monomer, which is greater than any of the monomer yields of the Δ41-52 hPRL compounds. The best of these were the 5G mutant with a 67% monomer yield, and Δ42-50 hPRL, with a 70.4% monomer yield.

Any protein with a monomer percent yield greater that of the Δ41-52 hPRL can be considered an engineering improvement. This includes G129R hPRL, 4G mutant, 5G mutant, Turn mutant, Δ42-48 hPRL, and Δ42-50 hPRL. Table 3.2 contains the percent yield for each protein as calculated by the dry mass and the area under the curve in the gel filtration profile. Most of the values are similar between the two types of calculations. Those that contain discrepancies could be due to the fact that the monomer and dimer peaks were often overlapping in the gel filtration profile and the samples actually collected did not contain the
entire peak so this calculation may not be the most accurate way to discern the differences in yield.

3.4 SDS-PAGE

After the monomeric and dimeric proteins were separated using gel filtration, they were run on an SDS-PAGE gel to determine their approximate molecular weights and assure that the samples were correctly assigned as monomeric and dimeric. Each protein was run in non-reducing and reducing conditions. The non-reducing gel was useful in seeing differences between sizes of monomer and dimer and detecting monomeric proteins that still contained a large amount of dimer contamination even after separation by gel filtration chromatography. The reducing gel was run to see if the dimeric protein was reduced to a monomeric form, which gave some indication of the type of force holding the two monomeric units together.

The molecular weight marker that was used is detailed in table 2.1. The monomeric prolactin protein has an approximate molecular weight of 23 kilodaltons, so it should ideally run somewhere between bovine pancrease trypsinogen and soybean trypsin inhibitor. Likewise, the dimeric form of prolactin has an approximate molecular weight of 46 kilodaltons, so it would ideally run between bovine albumin and egg albumin.

WT hPRL monomer was extremely clean and ran in the expected location relative to the molecular weight marker, although the marker is difficult to see in this picture (see figure 3.3).
G129R hPRL monomer showed very little dimeric contamination (see figure 3.4). It ran near the 24.5 kilodalton marker, which is close to its predicted molecular weight. △41-52 hPRL ran slightly lower, which makes sense since it is a smaller molecule. The gel shows that after separating the monomer and dimer using gel filtration, the monomeric portion of △41-52 hPRL still contains some dimer contamination (see figure 3.4).

Since the 2G mutant had so many more peaks on the gel filtration graph than the other proteins that it was difficult to determine which of these peaks was the monomeric or dimeric protein. As a result, all of the peaks were collected and run on an SDS-PAGE gel to determine which contained the protein. From the gel pictures in figure 3.5, it was determined that the useable monomers were 2G sample 1 monomer, 2G sample 2 Peak 3, 2G sample 3 Peak 4, and 2G sample 4 Peak 3. These will be referred to as 2G1 monomer, 2G2 monomer, 2G3 monomer and 2G4 monomer, respectively, for the remainder of the document. The dimers were determined to be in 2G sample 1 dimer peak, 2G sample 2 Peak 2, 2G sample 3 Peak 4, and 2G sample 4 Peak 3. The rest of the peaks had very small yields and were assumed to be small contaminants left on the column. 2G sample 5 was run on the same gel as the 4G protein (see figure 3.7), but its yield was so small that it was not used for further experiments. The 2G gel showed that the 2G monomers suffered very little dimer contamination and appear to be the correct size (see figure 3.5). The image for the reducing gel was lost due to a corrupt disk. The gel looked as expected for a reducing gel; all of the bands were around the size of a monomeric protein.
For 3G mutant sample 1, two separate samples were collected from the ion exchange column and named 3G mutant sample 1-1 and 3G mutant sample 1-2. From the gel pictured in figure 3.6, only 3G samples 1-1 and 2 monomers were deemed adequate for future experiments. Both contain a slight dimer contamination. It is also important to note that this protein ran between the 14,200 Dalton and the 20,100 Dalton marker bands, which is on the low side for this protein. However, since SDS-PAGE is not an extremely accurate mode of measuring mass this can be disregarded unless it is supported by mass spectrometry data. This argument can be used for other mutants that have this same problem as well.

Based on the gel pictured in figure 3.7, only 4G mutant sample 2 monomer was used for assays because the other samples either did not show up on the gel or had an overwhelming dimer contamination.

The 5G monomers for samples 1 and 3 both have no dimer contamination and appear to be the correct molecular weight (see figure 3.8).

The alanine mutant monomers contained significant dimer contamination but appear to be the correct mass according to this gel (see figure 3.9).

The T mutant monomers all contained minimal dimer contamination and run at the expected molecular weight (see figure 3.10).

Δ42-48 hPRL monomers had minimal dimer contamination and appear to be the correct size, although they run a little bit light with respect to the markers (see figure 3.11).
△42-50 hPRL has minimal dimmer contamination and appears to be the correct mass, although it also runs a little light with respect to the markers (see figure 3.12).

In summary, the SDS-PAGE data shows that the proteins are the correct size. Some of the proteins ran a little bit faster than expected on the gel but this can be accounted for by the fact that SDS-PAGE is not a very sensitive mode of detecting protein mass and also that different shapes of the marker and the prolactin could play a role. For instance, if the prolactin is packed together more tightly than the marker protein, then it will run further on the gel and appear to have a smaller mass. The non-reducing gels were useful in determining if the monomers and dimers were separated during gel filtration and if any dimer contamination remained in the monomer fraction. The reducing gels were useful for showing that the reduced dimers were the same size as the monomers, leading to the conclusion that a reducible force such as a disulfide bond must hold the dimer together.

3.5 Mass Spectrometry

Mass spectrometry was used to verify that each protein was, in fact, the expected protein. The molecular weight of each protein was calculated using Expasy (http://us.expasy.org/tools/). The total mass of each protein was then measured using electrospray time of flight mass spectrometry. Table 3.3 lists the theoretical molecular weights and the actual molecular weights obtained by mass spectrometry as well as the percent error. Since all of the percent error values are less than 0.1, these proteins are accepted as the expected proteins. It is
important to note that while the values for the expected molecular weight and the actual molecular weight are extremely close to one another, the actual molecular weight is always slightly greater. This could be due to ions associating with the protein during the mass spectrometry reading or due to hydrogen ions associating with the protein under the extremely acidic conditions used to prepare the protein for this analysis.

Additionally, the difference between the molecular weight for Δ41-52 hPRL and all of the insertional mutants was calculated (see table 3.3). The molecular weight of glycine is 75 g/mol, so after accounting for the loss of water for each peptide bond formed (18 g/mol per peptide bond) it is expected that the 2G mutant will have an increase of 114 g/mol, the 3G mutant will have an increase of 171 g/mol, the 4G mutant will have an increase of 228 g/mol, and the 5G mutant will have an increase of 285 g/mol relative to Δ41-52 hPRL. The molecular weight of alanine is 89, so the A mutant is expected to have a difference of 142 g/mol after the loss of water is accounted for. The molecular weight of serine is 105 g/mol and proline is 115 g/mol, so the turn mutant (with an addition of serine, proline, and two glycines) is expected to have an increase of 298 g/mol relative to Δ41-52 hPRL. The molecular weights of the added amino acids aspartic acid, glycine, phenylalanine, isoleucine, and threonine are that are added to Δ42-48 hPRL are 133 g/mol, 75 g/mol, 165 g/mol, 131 g/mol, and 119 g/mol, respectively, so the difference between the mass of Δ41-52 hPRL and Δ42-48 hPRL is expected to be 533 g/mol. Likewise, for Δ42-50 hPRL which has the residues aspartic acid, isoleucine, and threonine added, the expected molecular
weight increase is 329 g/mol. All of these values are equal to or extremely close to the values obtained by mass spectrometry. The data suggest that the intended proteins were in fact correctly prepared and do not contain additional modifications.

3.6 Ultraviolet Absorbance Spectroscopy

Monomeric forms of each hPRL variant that was to be used for future assays was characterized using UV absorbance spectroscopy. This was done to gain information regarding the fold of each prolactin variant by comparing the absorbance spectra of each mutant with that of WT hPRL. The data in this section is the average UV spectra for each prolactin variant. The data was first normalized to the absorbance at 285 nm. This was done because each sample will vary slightly due to slight variations in protein concentration. Normalizing the spectra allows for comparisons in shape to be made between the different proteins since each should have a peak near 285 nm. Each prolactin variant is graphed with WT hPRL. Additionally, the raw data for each protein is included as an inset on the normalized graph (see figure 3.13).

All of the spectra of mutant human prolactins were measured to compare to the WT hPRL spectrum to ascertain whether they had the same structure as WT hPRL as reflected by their absorbance spectra. The WT hPRL spectrum has a characteristic 277 nm peak indicative of its aromatic amino acids. It also has a slight peak at 290 due to tryptophan absorbance. As mentioned in section 3.1, it also has a characteristic 280/250 ratio of approximately 2. This ratio can be approximated from the UV spectrum.
All of the prolactin mutants have a UV spectrum with the same characteristics as WT hPRL. The only difference is a slight difference in the 245 nm to 265 nm range in some of the variants (G129R hPRL, 2G mutant, 3G mutant, 4G mutant, and A mutant). This could either be an artifact of the normalization of the spectrum or could indicate some stretching of the disulfide bonds in these proteins relative to the wild-type compound.

The UV spectrum of the TRIS-NaCl buffer was taken to assure that the peak at 285 nm was not due to absorbance by the buffer (see figure 3.13 L). No absorbance was measured so it can be assumed that the above spectra are due to protein rather than buffer. Please note that in this figure, the scale in the y-axis has a maximum lower than that in the protein graphs. This is because there was such a low absorbance for the buffer.

3.7 Fluorescence Spectroscopy

The fluorescence spectrum was measured for each monomer species of hPRL. Fluorescence spectroscopy is another method of assessing proper folding of a protein because it is indicative of the environment around the aromatic residues, in particular tryptophan and tyrosine. Similar to the UV spectra, the fluorescence spectrum for each mutant were normalized, this time to the intensity at 340 nm. The average normalized spectrum for each mutant was then graphed with WT hPRL for comparison. Additionally, each graph contains the raw data as an inset. The raw data is important in this case because the intensity of the signal can be attributed to the protein’s packing due to the exclusion of water. All of the proteins show the 340 nm peak characteristic of
prolactin. Each protein also contains a small peak around 320 nm, which can probably be attributed to fluorescence emission by the buffer because it fluoresces in this region (see figure 3.14). While it is more significant in some proteins than others, all of the prolactin variants except for G129R hPRL exhibit a red shift in the higher wavelengths relative to the WT hPRL spectrum. This shift indicates a stress on the protein’s folded structure. Δ41-52 hPRL has the largest red shift, 5G and A have a significant shift and the rest of the proteins have a less significant shift. Additionally, delta 41-52 hPRL has an exaggerated extra peak around 320 nm, which could be a result of folding problems as well.

The signal intensities of all of the mutants except for 2G and Δ41-52 hPRL show an increased signal in the raw data with respect to WT hPRL. This is indicative of a more tightly packed protein and the exclusion of water. It is interesting to note that for the 3G, 4G, and 5G mutants, the intensity of the signal increases with the number of added glycines. Also, Δ42-48 hPRL (5 added amino acids) and 5G have very similar increases in signal intensity while Δ42-50 hPRL (3 added amino acids) and 3G do not.

The spectrum for the buffer was also read to ensure that the spectrum was not characteristic of the buffer (see figure 3.14 L). Please note the difference in scale since the buffer spectrum is not normalized.

3.8 Circular Dichroism (CD)

To assess the secondary structure of the proteins, the CD spectrum for each was measured from 200 to 260 nm. This was done to ensure that the mutant prolactins were still alpha helical in structure. The characteristic shape of
an alpha helix CD spectrum has a negative peak at 222 nm (due to the $\pi\pi^*$ electron orbital transition) and 208 nm (due to the $\pi\pi^*$ electron orbital transition). The molar ellipticity ($\Theta$) was calculated for each in units of g cm$^{-2}$ mol$^{-1}$. Each spectrum was normalized to its 222 nm peak to account for differences in intensities of the different spectra due to variations in concentration. In addition, the raw data is included as an inset. Please see figure 3.15 for the CD spectra of prolactin and its variants.

Since prolactin is a four alpha helix bundle it is expected that its CD spectrum would be characteristic of alpha helix secondary structure. All of the proteins have a negative peak at 222 nm and 208 nm, indicating that their secondary structures are not significantly different from the parent compound.

Most of the mutants have a larger negative 208 nm peak which could indicate a larger $\pi\pi^*$ energy transition and could reflect stress on the molecule. However, the main information sought from the CD spectra was the fact that all of the proteins remain alpha helical in structure.

Additionally, the raw data shows that all of the mutants except for G129R hPRL have a significantly weaker signal than WT hPRL. This indicates that the mutants are less helical in structure than WT hPRL. Interestingly, the protein with the least discrepancy is A42-48 hPRL. It is important to remember, though, that the concentration of the protein may play a role in these discrepancies as well. The spectrum of the buffer was also read to assure that the characteristic alpha helix curve was not due to the buffer (data not shown).
3.9 Cell Agonist Assay

The previous sections of this chapter have examined the structural features of each protein. It is also important to assess the biological response that each protein evokes to determine if each mutant behaves more like WT hPRL or Δ41-52 hPRL *in vivo*. To do this, the Alamar Blue cell agonist assay was used to determine the relative amount of agonist activity left in each protein relative to WT hPRL.

The data from the Alamar Blue cell agonist assay (described in section 2.12) was graphed using Sigma Plot. The data was graphed as percent Alamar reduced versus the log of the hormone concentration. Each prolactin variant was graphed with the data for WT hPRL. The ED$_{50}$ was calculated for each protein using a four parameter fit (48) (this data is included at the end of the this section in table 3.4). A shift to a higher ED$_{50}$ than that of WT hPRL corresponds to less agonist activity because it requires more hormone to invoke the same biological response. The particular binding site that contains the mutation and the type of mutation have an effect on the shape of the dose-response curve (49-51).

Hormones that stimulate cells through receptor dimerization, such as the hematopoietic cytokines, classically display bell-shaped dose-response curves with agonist activity at low concentrations and antagonist activity at high concentrations. Site 1 mutants typically either increase or decrease the affinity of the ligand for the receptor and will shift the entire dose-response curve to the left or the right, respectively. Site 2 mutants typically lower the maximum of the dose-response curve as well as shifting the curve if agonist activity is affected.
Ligands with a connective motif block (or a mutation within a motif that couples sites 1 and 2), such as Δ41-52 hPRL, typically shift the curve if agonist activity is affected, as well as exceeding the maximum response of the wildtype ligand at high doses. This information is helpful in interpreting the dose-response curves presented in figure 3.16.

The data for each mutant is a result of the average of three cell assays. The ED$_{50}$ reading for each mutant is included in table 3.4. It is important to note that while G129R hPRL only increased the ED50 value by one order of magnitude, Δ41-52 hPRL, 4G mutant, and A mutant increased it by three to four orders of magnitude, which is significantly higher. This means that these three mutants are much poorer agonists and therefore have the potential to be much greater antagonists than G129R hPRL.

The shape of the curves is also useful data to interpret. While it is important to note that the cell assays were run at different times so the behavior of the cells may play a role in this phenomenon, the dose-response curves do display the expected behavior. For instance, G129R hPRL, which is a simple site 2 block mutant, displays a classic dose-response curve for a site 2 block with decreased agonist activity and a lower maximal response than WT hPRL (see figure 3.16 B). Δ41-52 hPRL, which has a connective motif mutation, shows a decreased agonist activity and an increased maximal response (see figure 3.16 C). All of the other mutations that involve an addition to Δ41-52 hPRL act in a similar matter except that they have recovered some of the wildtype hormone's agonist activity (see figure 3.16 D-K). In addition, the 4G mutant curve is atypical
in that the entire curve is has a higher value of Alamar Blue reduced than the wildtype curve. This could be due to a loss in responsiveness of the FDC-P1 cells at the time of this assay (see figure 3.16 F). Before this data is submitted for publication, this cell assay will be repeated to assure that this information is correct.
Figure 3.1: Ion Exchange Chromatograms for WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), Turn mutant (I), Δ42-48 hPRL (J), and Δ42-50 hPRL (K). The blue line indicates the absorbance at 280 nm and the red line indicates the absorbance at 260 nm. All other lines in this figure can be ignored. The protein was collected when the value of the 280 nm absorbance was greater than the 260 nm absorbance (ideally a 2:1 ratio).
Figure 3.1 continued

G

H

I

J

K

The graphs in Figure 3.1 show the results of various experiments. Each graph represents different conditions and mutants. The data suggests that certain mutations affect the expression or activity of the proteins under study.
<table>
<thead>
<tr>
<th>Protein</th>
<th>monomer 280:250</th>
<th>dimer 280:250</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT hPRL</td>
<td>3</td>
<td>peak too small to measure</td>
</tr>
<tr>
<td>G129R hPRL</td>
<td>2.8</td>
<td>peak too small to measure</td>
</tr>
<tr>
<td>Δ41-52 hPRL</td>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>2G mutant</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td>3G mutant</td>
<td>3.6</td>
<td>4.3</td>
</tr>
<tr>
<td>4G mutant</td>
<td>2.3</td>
<td>2.6</td>
</tr>
<tr>
<td>5G mutant</td>
<td>2.9</td>
<td>3</td>
</tr>
<tr>
<td>A mutant</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Turn mutant</td>
<td>2.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Δ42-48 hPRL</td>
<td>2.2</td>
<td>2.5</td>
</tr>
<tr>
<td>Δ42-50 hPRL</td>
<td>2.9</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 3.1: 280:250 ratios from the gel filtration data for each protein.
Figure 3.2: Gel filtration profiles for WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), Turn mutant (I), Δ42-48 hPRL (J), and Δ42-50 hPRL (K).
Figure 3.2 continued
<table>
<thead>
<tr>
<th>Protein</th>
<th>Average Monomer Yield (mg)</th>
<th>Average Dimer Yield (mg)</th>
<th>Average monomer percent yield (as calculated from dry mass)</th>
<th>Average dimer percent yield (as calculated from gel filtration profile)</th>
<th>Average monomer percent yield (as calculated from dry mass)</th>
<th>Average dimer percent yield (as calculated from gel filtration profile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT hPRL</td>
<td>15.4</td>
<td>1.8</td>
<td>90.1</td>
<td>100</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>G129R hPRL</td>
<td>18.3</td>
<td>2.1</td>
<td>90</td>
<td>96</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Δ41-52 hPRL</td>
<td>9.4</td>
<td>13.3</td>
<td>42</td>
<td>12</td>
<td>58</td>
<td>88</td>
</tr>
<tr>
<td>2G mutant</td>
<td>2.7</td>
<td>3.6</td>
<td>39</td>
<td>61</td>
<td>61</td>
<td>39</td>
</tr>
<tr>
<td>3G mutant</td>
<td>7.9</td>
<td>10.2</td>
<td>40</td>
<td>43</td>
<td>60</td>
<td>57</td>
</tr>
<tr>
<td>4G mutant</td>
<td>13.5</td>
<td>10.2</td>
<td>57</td>
<td>61</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>5G mutant</td>
<td>11.4</td>
<td>5</td>
<td>67</td>
<td>89</td>
<td>33</td>
<td>11</td>
</tr>
<tr>
<td>ΔA mutant</td>
<td>6.2</td>
<td>11.5</td>
<td>34</td>
<td>44</td>
<td>66</td>
<td>56</td>
</tr>
<tr>
<td>Turn mutant</td>
<td>8.7</td>
<td>8.8</td>
<td>49</td>
<td>65</td>
<td>51</td>
<td>35</td>
</tr>
<tr>
<td>Δ42-48 hPRL</td>
<td>17.5</td>
<td>15.7</td>
<td>52</td>
<td>57</td>
<td>48</td>
<td>43</td>
</tr>
<tr>
<td>Δ42-50 hPRL</td>
<td>13.4</td>
<td>5.4</td>
<td>70</td>
<td>74</td>
<td>30</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 3.2: Yields of Prolactin and Prolactin Variants
Figure 3.3: Reducing (A), and non-reducing (B) SDS-PAGE of WT hPRL. Lane 1 contains the molecular weight markers, lanes 2, 4, and 6 contain the monomers for samples 1, 2, and 3, respectively. Lanes 3, 5, and 7 contain the dimers for samples 1, 2, and 3, respectively.
Figure 3.4: Reducing (A) and non-reducing (B) SDS-PAGE of G129R hPRL and Δ41-52 hPRL. Lane 1 contains the molecular weight markers, lanes 2, 4, and 6 contain G129R samples 1, 2, and 3 monomer, respectively. Lanes 3, 5, and 7 contain G129R hPRL dimer for samples 1, 2, and 3, respectively. Lanes 8 and 10 contain Δ41-52 hPRL samples 1 and 2 monomers, and lanes 9 and 11 contain Δ41-52 hPRL samples 1 and 2 dimers.
Figure 3.5: Non-reducing SDS-PAGE of 2G mutant. The reducing gel image was lost due to a corrupt disk. Lane 1 contains the molecular weight markers, lane 2 contains sample 1 peak 1, lane 3 contains sample 1 monomer, lane 4 contains sample 1 dimer, lane 5 contains sample 2 peak 1, lane 6 contains sample 2 peak 2, lane 7 contains sample 2 peak 3, lane 8 contains sample 3 peak 3, lane 9 contains sample 3 peak 4, lane 10 contains sample 3 peak 5, lane 11 contains sample 4 peak 1, lane 12 contains sample 4 peak 1, lane 13 contains sample 4 peak 3, lane 14 contains sample 5 peak 1 and lane 15 contains sample 5 peak 2.
Figure 3.6: Reducing (A) and non-reducing (B) SDS-PAGE of 3G mutant. Lane 1 contains the molecular weight markers and lanes 2, 3, and 4 contain the monomers for samples 2, 3, and 4, respectively. Lanes 5, 6, 7, and 8 contain the dimers for samples 1-1, 1-2, 2 and 3, respectively.
Figure 3.7: SDS-PAGE of 4G mutant. Lane 7 contains the molecular weight markers. Lanes 1-4 are non-reducing and lanes 9-12 are reducing. Lanes 1 and 9 contain sample 1, lanes 2 and 10 contain sample 2 monomer, lanes 3 and 11 contain sample 3 monomer, and lanes 4 and 12 contain sample 3 dimer.
Figure 3.8: Reducing (A) and non-reducing (B) SDS-PAGE of the 5G mutant.
Lanes 1-6 contain protein from another experiment and can be ignored. Lanes 7 and 10 contain the monomers for samples 1 and 3, respectively. Lanes 8, 9, and 11 contain the dimers for samples 1, 2, and 3 respectively. Lane 12 contains the molecular weight markers.
Figure 3.9: Reducing (A) and non-reducing (B) SDS-PAGE of A mutant. Lane 1 contains the molecular weight markers. Lanes 2, 4, 6, 8, 10, and 12 contain the monomers for samples 1, 2, 3, 4, 5, and 6, respectively. Lanes 3, 5, 7, 9, and 11 contain the dimers for samples 1, 2, 3, 4, 5, and 6 respectively. Lanes 13-15 contain protein for another experiment and can be ignored.
Figure 3.10: Reducing (A) and non-reducing (B) SDS-PAGE of Turn mutant.
Lane 1 contains the molecular weight markers. Lane 2, 4, and 6 contain the monomers for samples 1, 2, and 3 respectively. Lanes 3, 5, and 7 contain the dimers for samples 1, 2, and 3 respectively.
Figure 3.11: Reducing (A) and non-reducing (B) SDS-PAGE of Δ42-48 hPRL.

Lane 1 contains the molecular weight markers. Lanes 2, 4, 6, 8, 10, and 12 contain monomers for samples 1-6, respectively. Lanes 3, 5, 7, 9, and 11 contain dimers for samples 1-5, respectively.
Figure 3.12: Reducing (A) and non-reducing (B) SDS-PAGE of Δ42-50 hPRL.
Lane 1 contains the molecular weight markers. Lanes 2, 4, 6, 8, and 10 contain monomer from samples 1-5 respectively. Lanes 3, 5, 7, 9, and 11 contain the dimers from samples 1-5 respectively.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Theoretical Molecular Weight (g/mol)</th>
<th>Actual Molecular weight (g/mol)</th>
<th>Percent Error</th>
<th>Mass Difference between Δ41-52 and each mutant (g/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT hPRL</td>
<td>23011</td>
<td>23028</td>
<td>0.07</td>
<td>-</td>
</tr>
<tr>
<td>G129R hPRL</td>
<td>23111</td>
<td>23123</td>
<td>0.05</td>
<td>-</td>
</tr>
<tr>
<td>Δ41-52 hPRL</td>
<td>21579</td>
<td>21595</td>
<td>0.08</td>
<td>0</td>
</tr>
<tr>
<td>2G mutant</td>
<td>21693</td>
<td>21705</td>
<td>0.06</td>
<td>110</td>
</tr>
<tr>
<td>3G mutant</td>
<td>21750</td>
<td>21762</td>
<td>0.06</td>
<td>167</td>
</tr>
<tr>
<td>4G mutant</td>
<td>21807</td>
<td>21822</td>
<td>0.07</td>
<td>227</td>
</tr>
<tr>
<td>5G mutant</td>
<td>21864</td>
<td>21879</td>
<td>0.07</td>
<td>284</td>
</tr>
<tr>
<td>A mutant</td>
<td>21721</td>
<td>21736</td>
<td>0.07</td>
<td>141</td>
</tr>
<tr>
<td>Turn mutant</td>
<td>21877</td>
<td>21893</td>
<td>0.07</td>
<td>298</td>
</tr>
<tr>
<td>Δ42-48 hPRL</td>
<td>22112</td>
<td>22128</td>
<td>0.07</td>
<td>533</td>
</tr>
<tr>
<td>Δ42-50 hPRL</td>
<td>21908</td>
<td>21924</td>
<td>0.07</td>
<td>329</td>
</tr>
</tbody>
</table>

Table 3.3: Molecular Weights of Prolactin and Variants
Figure 3.13: UV spectra for WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), Turn mutant (I), Δ42-48 hPRL (J), Δ42-50 hPRL (K), and spectroscopy buffer (L).

Each graph has been normalized to the absorbance at 285 nm. The raw data is included as an inset.
Figure 3.13 continued
Figure 3.13 continued

Relative UV Absorbance vs Wavelength (T mutant)

Relative UV Absorbance vs Wavelength (delta 42-42)

Relative UV Absorbance vs Wavelength (delta 42-8)

UV Spectrum for TRIS-NaCl

Relative Absorbance

wavelength (nm)

Relative Absorbance

wavelength (nm)

Relative Absorbance

wavelength (nm)
Figure 3.14: Fluorescence spectra of WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), Turn mutant (I), Δ42-48 hPRL (J), Δ42-50 hPRL (K) and spectroscopy buffer (L). The data shown has been normalized to the signal intensity at 340 nm. The raw data is included as an inset.
Figure 3.14 continued

Continued
Figure 3.15: CD spectra of WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), Turn mutant (I), Δ42-48 hPRL (J), and Δ42-50 hPRL (K). The data shown has been normalized to the molar ellipticity at 222 nm. The raw data is included as an inset.
Figure 3.15 continued

Continued
Figure 3.15 continued

Relative Molar Ellipticity vs Wavelength (Turn mutant)

Relative Molar Ellipticity vs Wavelength (delta 42-48 HPK)

Relative Molar Ellipticity vs Wavelength (delta 42-58 HPK)
<table>
<thead>
<tr>
<th>Protein</th>
<th>ED50 (nM)</th>
<th>Fold decrease in agonist activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT hPRL</td>
<td>0.23</td>
<td>1</td>
</tr>
<tr>
<td>G129R hPRL</td>
<td>5.99</td>
<td>26</td>
</tr>
<tr>
<td>Δ41-52 hPRL</td>
<td>702.42</td>
<td>3041</td>
</tr>
<tr>
<td>2G</td>
<td>5.25</td>
<td>23</td>
</tr>
<tr>
<td>3G</td>
<td>18.91</td>
<td>82</td>
</tr>
<tr>
<td>4G</td>
<td>1277.16</td>
<td>5312</td>
</tr>
<tr>
<td>5G</td>
<td>1.36</td>
<td>6</td>
</tr>
<tr>
<td>A</td>
<td>163.38</td>
<td>707</td>
</tr>
<tr>
<td>Turn</td>
<td>0.44</td>
<td>2</td>
</tr>
<tr>
<td>Δ42-48 hPRL</td>
<td>47.93</td>
<td>207</td>
</tr>
<tr>
<td>Δ42-50 hPRL</td>
<td>45.94</td>
<td>199</td>
</tr>
</tbody>
</table>

Table 3.4: ED\textsubscript{50} values for hPRL and hPRL mutants
Figure 3.16: Alamar Blue dose-response curves. Alamar Blue reduced versus the log of the hormone concentration for WT hPRL (A), G129R hPRL (B), Δ41-52 hPRL (C), 2G mutant (D), 3G mutant (E), 4G mutant (F), 5G mutant (G), A mutant (H), T mutant (I), Δ42-48 hPRL (J), and Δ42-50 hPRL (K).
Figure 3.16 continued
Figure 3.16 continued

Alamar Reduced vs Hormone Concentration
Turn mutant

\[ \text{Turn mutant} \]

Alamar Reduced vs Hormone Concentration
delta 42-48 hPRL

\[ \text{delta 42-48 hPRL} \]

Alamar Reduced vs Hormone Concentration
delta 42-50 hPRL

\[ \text{delta 42-50 hPRL} \]
4.1 Project Aims Summary

The main problem addressed in this project was that of protein folding. Δ41-52 hPRL has great potential as a prolactin antagonist in cancer fighting applications, but its ability to refold into monomeric units is poor. The aim of this project was to design a prolactin antagonist based on Δ41-52 hPRL that maintained its reduced agonist and increased antagonist activity but also had an improved ability to refold into the monomeric form. The mutant proteins designed for this project were analyzed by measuring their 280 nm : 250 nm ratio, ultraviolet absorbance, fluorescence emission, and circular dichroism spectra to determine if the folding more closely resembled the wild-type protein or the Δ41-52 hPRL protein. The relative amounts of monomer and dimer were also measured using gel filtration chromatography as well as yield measurements in the form of milligrams of product obtained. Finally, the proteins were tested for their agonist activity in FDC-P1 cells transfected with the human prolactin receptor. This section of the thesis will analyze the data presented in section 3
as it pertains to the success in folding each protein and the maintenance of the bioactivity of Δ41-52 hPRL.

4.2 Anion Exchange Chromatography

WT hPRL and G129R hPRL are by far the best folded proteins as judged by the 280/250 peak with a ratio of 2. Several of the mutant proteins had elution profiles similar to WT hPRL. These were 4G, 5G, Δ42-48 hPRL, and Δ42-50 hPRL. Δ41-52 hPRL has a much broader peak eluted from the column and the 3G and A mutants seem to mimic the folding of Δ41-52 hPRL in this regard. The two other notably poor folders using ion exchange as a measurement were the 2G and T mutants. However, since there are so many variables when running an ion exchange column (matrix, buffer pH, and system settings), using ion-exchange chromatography to judge folding is probably not the best way to assign the success of folding for each protein. The folding success will be further discussed in the sections pertaining to gel filtration chromatography and spectroscopy.

4.3 Gel Filtration Chromatography and Protein Yields

Gel filtration chromatography is useful in that it not only displays the presence of multiply sized particles but it shows the amount of each type of particle relative to the others. It was especially useful in this project because the size of the dimer peak could be directly related to correct folding of the molecule. If the protein has been strained during the refolding process, the disulfide bond will be weakened or not form and have the opportunity to bond with a free cysteine from another molecule, creating a dimeric species. Those molecules
that have little to no dimer peak on the gel filtration column can be deemed as good folders.

The proteins that were the best at folding using the criteria of a high proportion of monomeric species were WT hPRL and G129R hPRL, which each contained one large monomeric protein peak and a small, negligible dimeric protein peak. Based on this data alone, these proteins can be deemed correctly folded because little to no dimer formed through free cysteine residues. Δ41-52 hPRL, on the other hand, displays between a 1:1 and a 1:2 ratio of monomer to dimer species. This displays poor folding. This data is interpreted to mean that there is more strain within these molecules that does not allow proper intermolecular dimer formation. Three other molecules: the A mutant, 3G mutant and Δ42-48 hPRL folded in essentially the same way as Δ41-52 hPRL. On the other hand, the 5G mutant folds in essentially the same manner as WT hPRL. Also, 2G, Turn, and Δ42-50 hPRL all have larger monomer peaks than dimer peaks, showing that the folding problem has been at least partially solved for these proteins. It is interesting to note that while the 5G mutant, which has five residues added to the deleted region of Δ41-52 hPRL folds like the WT hPRL molecule, Δ42-48 hPRL, which also has five residues added to the deleted region, has extremely poor folding more reminiscent to Δ41-52 hPRL. Additionally, while Δ42-50 hPRL (which has three residues added to the deleted region) folds in such a way that there is more monomer than dimer present, the 3G mutant which also has three amino acids added folds more like Δ41-52 hPRL. This data shows that the number of amino acids added to a region does
not have a direct relationship to folding potency, but the identity of the amino acids does make a difference. It is also interesting to note that the sequential addition of glycines does not have a consistent effect. For instance, the 2G and 5G mutants fold with more monomer than dimer, with the 5G mutant folding almost exactly like the WT hPRL. At the same time, the 3G folds more like Δ41-52 hPRL than WT hPRL and the 4G mutant does not seem to have a consistent folding pattern. This could be due to both the number and nature of the added amino acid residues. It may be the case that five glycines provide enough space for the added glycines to form a loop connecting residues 40 and 53, while a smaller number of glycines still has some strain in this region and instead makes no contribution to stability. This theory is supported by the UV spectroscopy data which shows the 5G mutant as a well-folded molecule and the 3G mutant as strained. However, fluorescent spectroscopy shows very little difference between the 3G, 4G, and 5G mutants. This theory does not explain the success of the 2G mutant, however, which is a properly folded molecule according to all spectroscopy data.

The protein yields were not very promising as to the success of the folding project in making a prolactin antagonist which folds similarly to wild-type prolactin. The two most successful mutants in producing higher monomer yields were the 5G mutant and Δ42-50 hPRL. This is consistent with the gel filtration data in which both of these proteins resembled WT hPRL more than Δ41-52 hPRL.
4.4 SDS-PAGE and Mass Spectrometry

SDS-PAGE was used to determine both the approximate molecular weight of each protein and the dimer contamination in the monomer portion of each protein. According to the SDS-PAGE data, all of the proteins were approximately the correct molecular weight. Mass Spectrometric analysis confirmed that the proteins were the proper molecular weight.

The main point of interest to come out of the SDS-PAGE analysis was the amount of dimer contamination that was still left in the monomer portion of each sample. The peaks collected as monomer peaks from the gel filtration column were collected rather conservatively to avoid collecting dimer protein with the monomer as much as possible. The proteins that came off of the column with dimer and monomer peaks very close together could still have some contamination in the monomer peak. These proteins probably either had a lot of dimer contamination to begin with or still had dimer forming after the monomer peak was collected from the column. Either way, these proteins can be seen as more poorly folded than the proteins with a single monomer peak or more distinct monomer and dimer peaks.

Another useful piece of information is the fact that in all cases the dimeric species of the proteins could be converted to the monomeric form under reducing conditions. This is important because it shows that a reducible force such as a disulfide bond is holding the two monomers together. It also shows that this bond indeed forms a dimer rather than a 46 kilodalton contaminant protein.
because it is reduced into smaller 23 kilodalton units that are the size of the monomeric protein when it is reduced.

WT hPRL can successfully fold into monomer without straining the molecule and forming dimers as indicated by the single tight peak formed from the gel filtration of this protein and the fact that this monomer peak had virtually no dimer contamination. G129R hPRL and the 5G mutant most resembled WT hPRL in that they had low dimer contamination in the monomer bands and very little dimer elute off of the column at all and had similar spectroscopic profiles. Other proteins that had low dimer contamination were 2G, 4G, T and Δ42-50 hPRL. Of these four proteins, 2G, 4G, and Δ42-50 hPRL also had a very low amount of dimer present even in the dimer bands according to the gel. T had a distinct dimer band in the dimer peak. According to this data, 2G, 4G, and Δ42-50 hPRL can be considered better folders than the T mutant. Finally, Δ41-52 hPRL contains a large dimer contamination in the monomer band which confirms our observation that led to this project. Δ42-48 hPRL and the A mutant resemble Δ41-52 hPRL in this manner, and so these three proteins can be considered less successful at folding than the wild-type molecule according to the SDS-PAGE data.

4.5 Spectroscopy

The UV absorbance, fluorescence, and CD spectra show that the mutant prolactins are overall structurally similar to WT hPRL. When normalized, all of the proteins had a UV absorbance spectra that had a similar shape as that of WT hPRL. The proteins most similar to WT hPRL by this assessment were G129R
hPRL, 2G, 4G, and Turn mutant. Δ41-52 hPRL showed a large deviance from the WT hPRL protein in the raw data indicating stress within the structures. The proteins that most closely resemble Δ41-52 hPRL in the UV absorbance spectra were 3G, Δ42-48 hPRL, and Δ42-50 hPRL, although none of these had as extreme of a deviance from the WT hPRL spectrum as Δ41-52 hPRL.

The fluorescent spectrum provided another indication of proper structure. As was the case for the UV measurement, G129R hPRL and 2G most closely resembled WT hPRL. The rest of the proteins showed more deviance from the WT hPRL spectra, however. Δ41-52 hPRL was the most unusual, displaying an irregularly shaped fluorescent profile. This indicated that Δ41-52 hPRL has the most strain on the molecule and that all of the mutations that add residues to the deleted sequence are an improvement in this regard. The raw data for Δ41-52 hPRL is puzzling, however, because despite its irregular shape, the signal intensity is similar to WT hPRL. This could lead to the interpretation either that this concentration of the Δ41-52 hPRL protein was closer to the concentration of the WT hPRL protein than the other mutants were or that the Δ41-52 hPRL is for some reason a more tightly packed molecule than the other mutants, which does not seem likely due to its behavior discussed in previous sections. The UV and fluorescence spectra show that there is a good indication that most of the Δ41-52 hPRL mutants have significant structural stress according to the variance in 250 nm absorbance in the ultraviolet spectra and the red shift in the fluorescent spectra. This is consistent with the idea that a deletion causes strain on the Cys58/Cys174 central disulfide bond and demonstrates that while the mutants
described above have improved the Δ41-52 hPRL compound’s monomer formation ability, they still have some stress due to the deleted region.

The CD data is useful in that it shows that all of the proteins retain an alpha helical secondary structure. All of the proteins have a reduced signal relative to WT hPRL, however, with G129R hPRL being the closest to the WT hPRL signal. For the deletion proteins, this could simply be a result of the loss of a number of amino acids, which could be forming partial helical structure. The loss of these residues would reduce the absorption of circularly polarized light, which would in turn reduce the CD signal. It could also be coincidental and simply be an artifact of slight differences in protein concentration between the samples, which seems to be a more likely explanation.

4.6 Biological Activity

The most relevant test for any compound with promise as a treatment for disease is its activity in a biological assay. For this project, the Alamar Blue cell agonist assay was used to test the changes in biological response between prolactin and its variants. An agonist assay is important because if the compound retains agonist activity or increases the agonist activity with respect to the parent compound, doses would have to be extremely high to have its antagonist activity have an effect on the cells. This also speaks to the effectiveness of the mode of antagonism. If a hormone retains agonist activity it means that the mechanistic block being used is only partially effective and the molecule will still allow receptor dimerization under some conditions. A more effective mechanism would block receptor dimerization in all cases causing all
receptors to be bound to only one molecule of hormone and thus block the biological response. In this experiment the G129R hPRL compound had only a modest increase in $ED_{50}$, so its agonist activity is only slightly decreased with respect to WT hPRL. $\Delta 41-52$ hPRL had over 100 times the increase in $ED_{50}$ as G129R hPRL which means it has a much more reduced agonist activity. Two of the second-generation $\Delta 41-52$ hPRL mutants (4G and A mutant) also showed some promise in reduced agonist activity, with a $10^4$ and a $10^3$ fold reduction in agonist activity, respectively. This shows a similar effect on the cells as $\Delta 41-52$ hPRL, and shows much promise in these two proteins as possible prolactin antagonists. Considering the fact that gel filtration, SDS-PAGE, and spectroscopy showed that the 4G mutant is a well-folded molecule, this compound should be further studied as a possible new form of $\Delta 41-52$ hPRL.  

4.7 Summary and Conclusions

For human prolactin, success of folding can be measured by the amount of monomer formed during the refolding process. In this study this was measured by separating the monomeric and dimeric species using gel filtration chromatography and measuring the relative amounts of monomer and dimer using dry mass and SDS-PAGE. The most successfully folded prolactin mutants (aside from WT hPRL and G129R hPRL) according to the above analysis are 5G, 2G, Turn, 4G, and $\Delta 42-50$ hPRL. The worst folders are A, 3G, and $\Delta 42-48$ hPRL. Again, it is interesting to note that 3G and $\Delta 42-50$ hPRL, which contain the same number of residues added to the deleted region, and likewise 5G and $\Delta 42-48$ hPRL act in opposite ways as far as folding is concerned. This shows
that the number of residues added to the region is not as important as the identity of the residues. From this it can be concluded that the spacing is not as important as interactions between side chains for the stability of this molecule.

The most promising mutants in the biological activity test were the 4G and A mutants. Since the A mutant was such a poor folder, it is probably not worth pursuing any further since it folds in the same manner as Δ41-52 hPRL but does not have as much of a decrease in agonist activity. The 4G mutant should be studied further since it demonstrated better folding than Δ41-52 hPRL and similar agonist activity. It is extremely important to point out the difference in agonist activity between G129R hPRL and the Δ41-52 hPRL proteins. The extreme decrease in agonist activity in Δ41-52 hPRL as compared to the moderate decrease in G129R hPRL shows that Δ41-52 hPRL has more potential as a cancer treatment.

Other strategies could be used to develop more successful second generation prolactin antagonists. One idea would be to truncate the N-terminus of the protein. Bernichtein et al. made a mutation in WT hPRL in which they deleted the first thirteen amino acids (52). They found that this decreased the formation of protein aggregates, presumably because they deleted the first disulfide bond from the protein. It would be interesting to see if such a deletion would also decrease the dimer formation of Δ41-52 hPRL and if it would have any effect on biological activity. This would also provide insight as to whether or not the central disulfide bond is the one being strained in this molecule, because if this deletion solves the problem of dimer formation, the dimer is probably being
formed between the first disulfide bond rather than the central bond. Another possibility would be to re-engineer the central disulfide bond itself using molecular modeling to move either Cys58 or Cys174 so the two would be closer together, which should help with stability of the disulfide bond.

The next logical step for this type of a project would be to develop a successful antagonist assay to study the amount of increase in prolactin antagonism for each of the mutants. Studies are currently underway to develop a flow cytometry protocol to assess the amount of apoptosis induced when varying concentrations of protein are added to Jurkat cells. This study is based on previous findings in the Brooks lab that Jurkat cells are responsive to prolactin and when treated with Δ41-52 hPRL respond with an increased incidence of apoptosis.

Another interesting study would be to determine the binding mechanism for the prolactin mutants. It is assumed that each of the mutants discussed in this project act as antagonists by preventing site 2 binding from occurring, but this is based on theory and not experimental evidence. Future plans for the lab are to use surface plasmon resonance to elucidate the exact binding mechanism and relative binding affinities for each receptor binding site.
LIST OF REFERENCES


