Mammalian Neuroproteomics and Neuropeptidomics: Analysis by Mass Spectrometry

by

James A Dowell

A dissertation in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Pharmaceutical Sciences)

at the

University of Wisconsin-Madison

2008

Acknowledgements

I would like to thank my advisor and mentor Lingjun Li for all for always being there. She has always supported me in whatever crazy thing I wanted to do and works harder than anyone I have ever met. She is truly a passionate scientist and a kind person. I would like to thank my co-mentor Jeff Johnson for his financial support and taking a chance on me—he's still waiting for it to pay off. He has a great scientific mind and has always been a great intellectual resource. I would like to thank Ann E. Kelley for being an inspiration to a young graduate student. She was truly a role model. She was a gifted scientist and a wonderful person. We miss her. I would like to thank my committee members, Warren Heideman, Arash Bashirullah, and Jerry Yin, for reading my thesis, listening to me talk, and giving me valuable feedback. I would like to thank all of my lab mates past and present and all of the people I have met along the road through graduate school. You have made the experience: wonderful, horrible, memorable, forgettable, funny, sad, bearable and unbearable. I would like to thank my parents for their support both emotional and financial, especially my mother who has always been there to encourage me in my times of doubt and misgiving. And last but not least, I would like to thank my wife Oana Martin for putting up with me.

Table of Contents

Abstract

The development of new analytical methods in neuroscience is important for our understanding of the mammalian brain. Mass spectrometry is a powerful tool for the identification and quantification of both proteins and peptides; however, methods for analyzing the mammalian neuroproteome and neuropeptidome have been lacking. Using a combination of neuroscience and mass spectrometry, we have developed a method for the analysis of mammalian neuropeptides *in vivo*. This method was subsequently used to analyze changes in neuropeptide release in fed and unfed rats. Among the neuropeptides analyzed, only the opioid peptides exhibited a change in release upon feeding. The opioid peptides are important regulators of hedonistic feeding; however, the exact site of opioid release was not known. In this study, modulation of opioid release in the reward/pleasure center of the brain was found to be feeding dependent. In addition to neuropeptide analysis, a method was developed for the analysis of proteins secreted from primary murine astrocytes. Astrocytes are the support cells in the brain, providing a matrix for neuron growth as well as playing important roles in maintaining neuronal health. Neurons and astrocytes communicate via a complex cell-to-cell language, modulated largely by secreted proteins. This communication is vital to the astrocyte maintenance of neuronal survival. In order to identify important players in this complex interaction, we developed a method for the global analysis of astrocyte secreted proteins. Using a combination of mass spectrometry and bioinformatics, over 170 astrocyte secreted proteins were identified, most of which were not known to be secreted by astrocytes.

Chapter 1: Introduction to Neuropeptides and Feeding

1.1 Neuropeptides

1.1.1 General Aspects

Peptides are expressed throughout the central nervous system and modulate important physiological functions, including feeding and body weight regulation, ciracadian rhythms, fear, pain, anxiety, sleep, learning and memory, and reward mechanisms.[1] Like proteins, peptides are composed of amino acids but, unlike proteins, are generally less than 100 amino acids in length. However, in order for a peptide to be classified as a neuropeptide, it must exhibit bioactivity in the central nervous system with many neuropeptides acting locally in cell-to-cell communication and others acting systemically as hormones.[1]

1.1.2 Neuropeptide Synthesis, Processing, and Packaging

Neuropeptides are synthesized in the rough endoplasmic reticulum as immature prepropeptides.[1] The transformation of a prepropeptide into a mature neuropeptide requires a variety of processing enzymes, including peptidases, sulfatases, and acetyltransferases. The most important of these enzymes are the endopeptidases, prohormone convertase 1 and 2 (PC1 and PC2), which typically cleave the propeptide at dibasic sites, e.g. Lys-Arg or Arg-Arg, to yield a semi-mature neuropeptide with a dibasic extension at the *C*-terminal.[2] This basic residue is then removed by carboxypeptidase E.[3] After the removal of the *C*-terminal basic residue, about half of the known neuropeptides subsequently undergo *C*-terminal amidation via peptidyl-α-amidating monooxygenase (PAM), a modification thought to

enhance receptor binding and protect the neuropeptide from extracellular peptidases.[4] A large number of other modifications can also occur in the secretory vesicles, including sulfation, acetylation, phosphorylation, and glycosylation, which can affect the biological properties of the neuropeptide (see Figure 1.1).[1] For example, *N*-terminal acetylation of β-endorphin eliminates its ability to bind the opioid receptor.[5]

1.1.3 Extracellular Processing

Once the immature neuropeptide reaches the *trans*-Golgi, it is sorted and packaged into secretory vesicles and then transported to the inner leaf of the cell membrane.[6] Upon the proper stimulus, usually a rise in cytoplasmic calcium, the secretory vesicle fuses with the cell membrane, releasing the neuropeptide into the extracellular space.[7] Once released into the extracellular space, the neuropeptide binds its receptor and elicits a response in the postsynaptic cell.[1] Unlike classic neurotransmitters, no known re-uptake or transport mechanisms exist for nueropeptides and thus termination of the neuropeptide signal is elicited by the cleavage of the neuropeptide by extracellular peptidases.[7] A variety of extracellular neuropeptide processing peptidases exist in the brain, with the majority belonging to the metallopeptidase family of peptidases, including angiotensin-converting enzyme (ACE), endothelin-converting enzyme (ECE), neutral endopeptidase (NEP, EP3.4.24.11), thimet oligopeptidase (EP3.4.24.15), and neurolysin (EP3.4.24.16) (see Figure 1.1).[8, 9] Each of these peptidases has specific substrates and can exhibit contrasting functions depending on the identity of the cleaved neuropeptide. For example, when ACE cleaves angiotensin I, it produces the bioactive peptide angiotensin II; while the cleavage of bradykinin by ACE renders bradykinin inactive.[10]

1.2 Neuropeptides and Feeding

1.2.1 General Aspects

As mentioned above, neuropeptides play a critical role in many physiological processes. My research has focused on the involvement of neuropeptides in feeding behavior. We can begin by asking a simple question—why do we eat? Because we're hungry, you answer, which is absolutely right. But what makes us hungry? At its most basic level, hunger is the manifestation of our need for energy. We are hungry because we need energy and we eat to fulfill that need. This is true; however, the whole truth is more complicated. If we always ate exactly what we needed exactly at that moment, then perhaps we would be continually walking around munching on a small amount of cracker or cheese or chocolate to fulfill our constant energy demands. However, this would be horribly inconvenient and so, thankfully, we have meals—larger amounts of food consumed all at once. Having meals makes things much easier; it gives us idle time between meals to do other things like work and play. These meals generally occur at a regular time each day and if the specified meal time passes and you haven't eaten then you become very hungry, even if you, like the majority of us, possess enough fat stores to survive a couple of weeks without food. So again we have to ask, what makes us hungry? We have enough energy stored for a much longer time than the interval between meals so we have no immediate need for food but we are hungry anyway. For an answer, we turn to our evolutionary history.

We, as a species, have not always had it quite this good. In times past, food was scarcer than it is today and so when our ancestors happened upon a good meal (i.e. a high energy food), they would gorge themselves in an attempt to build up energy stores for the times they

did not have food. These energy stores would increase their chances of survival and thus the opportunity to pass on their genes. So in our evolutionary history, binge eating and fat storage was actually a genetic advantage. Unfortunately, we have retained this 'advantage' into modern times when energy rich food is plentiful and easily obtained, resulting in a world-wide obesity problem.

1.2.2 Energy Homeostasis and the Hyphothalamus

So what are the biochemical underpinnings of this evolutionary drive to eat? Conceptually, the drive to eat can be broken down into two parts: the drive for energy homeostasis and the drive for reward/pleasure. The first drive, energy homeostasis, is perhaps the most basic and, as mentioned above, is simply the drive to maintain enough energy for our organism to function. In the mammalian brain, a structure called the hypothalamus is the main processing conduit for feeding and a specialized area within the hypothalamus, called the arcuate nucleus (ARC), is the energy sentinel for body. The ARC relays energy information from the peripheral organs, such as the gut, liver and adipose tissue, to higher structures in the brain involved in decision making and motivation.[1] In turn, the ARC is also responsible for sending signals back to the periphery via a specialized, hormone secreting organ called the pituitary (see Figure 1.2). This system affords a back-and-forth communication between the body and higher brain centers about the energy state of the organism. The second drive to eat, the reward/pleasure drive, is also partially controlled by the hypothalamus in a region called the lateral hypothalamus (LH). The LH relays information about energy homeostasis from the hypothalamus to the reward/pleasure centers, including the nucleus accumbens and the ventral tegmental area, and in a reciprocal manner

is responsible for relaying information back to the hypothalamus from the pleasure centers.[11, 12]

Neuropeptides are intimately involved in both feeding paradigms. In the maintenance of energy homeostasis, the peripheral peptide signals insulin, leptin, ghrelin, and cholecystokinin (CCK) relay information from the liver, gut, and adipose tissue to the ARC (see Figure 1.2).[13-15] ARC neurons are in close anatomical proximity to the fenestrated capillaries at the base of the hypothalamus, giving the ARC neurons direct access to blood circulating peptide signals, such as ghrelin and leptin, and nutritional signals, such as fatty acids, glucose, and insulin.[16] Ghrelin and leptin exhibit opposing bioactivities, with ghrelin stimulating feeding and leptin suppressing feeding. Ghrelin is produced in the fundus of the stomach and as the stomach becomes full, it produces less ghrelin, signaling to the brain that the stomach is full.[17-20] Leptin also acts as a satiety signal but instead of being produced in the stomach is produced in the adipose tissue. Circulating leptin is an indicator of adipose levels in the body and a reduction in adipose tissue causes a decrease in ghrelin release which in turn stimulates feeding.[21, 22] Both of these peripheral signals modulate feeding via neuropeptide signaling in the ARC.

The ARC is the 'master' controller of feeding in mammals. Virtually all the bodily systems which modulate feeding do so via the feeding circuitry of the ARC. The ARC feeding circuit is primarily composed of the opposing actions of two groups of neuropeptides: the stimulatory orexigenic neuropeptides, including neuropeptide Y (NPY) and agouti-related peptide (AgRP), and the inhibitory anorexigenic neuropeptides, including the pro-opiomalnocortin (POMC) and cocaine-amphetamine-regulated transcript (CART). In the ARC, neurons expressing NPY co-express AgRP while neurons expressing POMC co-

express CART and collectively these neurons are referred to as the melanocortin system—the central regulator of energy homeostasis in the mammalian brain.[23] NPY had been known to be involved in feeding before the discoveries of leptin and ghrelin;[24] however, much of what is known about the intricacies of the melanocortin system has been discovered from leptin- and ghrelin-induced activity in the ARC.[23, 25] Both the NPY/AgRP and POMC/CART neurons express both leptin and ghrelin receptors. Leptin increases activity in POMC/CART cells while inhibiting activity in the NPY/AgRP cells while ghrelin does the opposite.[26] Both NPY/AgRP and POMC/CART cells project to various areas of the hypothalamus, including the paraventricular nucleus, the ventromedial nucleus, and the lateral hypothalamus (LH).[27] The LH is of particular importance in regulating the rewarding, pleasurable aspects of feeding, a fact we will examine more closely in the following paragraphs.

1.2.3 Hedonistic Aspects of Feeding and Nucleus Accumbens

If eating were solely a function of energy homeostasis then the world would not be suffering from an obesity epidemic. However, eating is not only about energy maintenance but, more pointedly, it is about pleasure. Yes, eating generally is a pleasurable activity—good food tastes good. But why does certain food taste good? We have an evolutionary drive to eat the most energy dense food possible. When we eat energy dense foods, our bodies reward us by releasing endogenous opioids. In this manner, food is given valence with the more energy dense foods, such as fatty, sugary foods, given a higher value than the lower energy foods. But what are the biological underpinnings of this reward/pleasure valuation of food?

Motivation and reward have been studied most extensively in regard to drugs of

addiction. However, a number of studies have suggested an overlap of neurobiological mechanisms between drug addiction and the hedonisitic aspects of feeding.[28] Electrical stimulation of the LH, an area associated with feeding, is inherently rewarding and LH probe-implanted rats work vigorously to obtain stimulation. This simulation also caused a concurrent increase in food intake.[29] Alternatively, food deprivation increases the self-administration of non-food rewards, such as psychostimulants, intracranial stimulation, and heroin, suggesting a significant link between food reward and drug rewards.[30, 31]

Since the discovery of the endogenous opioids and their receptors in the 1970s, there has been a large body of research dedicated to understanding their functional roles in the brain.[32, 33] Subsequent to their discovery, the endogenous opioids were found to be involved in a wide variety of physiological processes, including the modulation of pain, stress, endocrine responses, motivational states, feeding, and addiction.[34] Early studies examining the effects of opiates on rats found morphine injections dramatically increased food intake.[35] These initial experiments led to a number of follow-up studies using both opioid agonists and antagonists, with the agonists increasing food consumption and antagonists decreasing food intake, solidifying the specific involvement of opioids in the regulation of feeding.[36] More recent studies using specific μ-opioid agonists, such as D-Ala2,Nme-Phe4,Gyol5-enkephalin (DAMGO), have shown the increase in food intake to be modulated by the μ-opioid receptors in the NAc.[37] These experiments were performed with satiated rats which had just eaten, and upon infusion of DAMGO into the NAc the rats resumed eating, consuming up to 400% of their normal intake of food.[38] Additionally, the DAMGO rats displayed a preference for high-fat food.[38, 39] These experiments demonstrate the powerful influence opioids have over feeding behavior which leads us to the question—how do opioids modulate feeding?

The neurons in the NAc are primarily GABAergic medium spiny neurons which express either preproenkephalin or prodynorphin.[34] The neuropetide products of preproenkephalin include both Met- and Leu-enkephalin (μ-opioid receptor agonists) while prodynorphin yields dynorphin A and B (κ-opioid agonists).[1] Both types of medium spiny neurons project to various parts of the brain, including connections to the ARC and the lateral hypothalamus (LH) (see Figure 1.3). This is important in regards to feeding behavior due to the ARC's position as the 'master' feeding switch. DAMGO infusion into the NAc causes μ-opioid receptor activation and consequent inhibition of GABA release from the affected medium spiny neurons.[34] So how is DAMGO infusion in the NAc affecting feeding behavior? The NAc has projections to both the LH and the ARC, as well as to the ventral pallidum (VP) and the ventral tegmental area (VTA), and infusion of DAMGO into the NAc results in *c-fos* activation in all of these areas.[12]

1.2.4 Orexin and the Lateral Hypothalamus

The LH is perfectly positioned to relay information between the ARC and the NAc, with projections both to and from these areas (see Figure 1.3). In addition, many studies have shown the importance of the LH in modulating the addictive effects of illicit drugs.[40] The LH possesses two main types of neuropeptide expressing cells—melanin concentrating hormone (MCH) and orexin expressing cells. It seems that the orexin cells are particularly important in the modulation of hedonistic feeding.[40] Using the established paradigm of DAMGO-induced feeding, Berthoud and co-workers infused an orexin antagonist in a number of brain areas, including the VTA, followed by the infusion of DAMGO into the NAc. Interestingly, the stimulatory effects of DAMGO on feeding were abolished with the

pre-treatment of the VTA with the orexin antagonist.[40] This VTA-specific denouement of opioid-induced feeding indicates the involvement of dopamine in the hedonistic valuation of food and thus we have stumbled into one of the great debates within the drug and addiction community—is dopamine the reward or are endogenous opioids the reward?

1.2.5 Dopamine and Endogenous Opioids: The Chicken or the Egg

In regards to feeding, both dopamine and opioids seem to be important in different ways: the opioids affect the palatability of food (the reward valuation) while dopamine affects the degree of motivation or the desire to obtain the reward. Or in the words of Kent Berridge, dopamine is the agent of 'wanting' and opioids are the 'liking.'[41] Much of the work done in opioids and palatability revolves around responses to bitterness and sweetness. In both humans and rats, the perception of palatability seems to be dependent on the opioid system and independent of the dopamine system.[41] Disruption of the dopamine-striatal pathway does not affect the results of taste testing nor food intake but instead affects the motivational and locomotor responses in the seeking of food reward.[41]

In addiction research, operant conditioning is an important read-out for the effects of various treatments on drug/reward seeking behaviors. The most common form of operant conditioning is positive reinforcement. Positive reinforcement pairs a sensory stimulus with a positive reward and as a result the animal learns to associate the stimulus with the reward. The most common form of positive reinforcement is called conditioned place preference (CPP); an animal receives a reward in a specific area of a maze and learns to associate that area with the reward. After entrainment, the animal will then actively seek out that area in anticipation of a reward. However, after the initial conditioning phase of the entrainment,

dopamine is only released when the reward exceeds the animal's expectations, i.e. the amount of drug or juice is larger than normal.[42] From these data, it seems that dopamine is important for the initial acquisition of the cue-reward association and in encoding novel, unexpected information regarding the cue-reward pairing but is not necessarily involved in the actual rewarding aspects of the drug or food. In contrast, drug and food CPP has been shown to depend on the opioid system with μ-opioid antagonists blocking cocaine induced CPP and μ-opoid agonists inducing CPP.[43] In addition, dopamine deficient mice can still display CPP for morphine and food, indicating that dopamine is not necessary for food reward associative learning.[44-46] On the other hand, as mentioned above, DAMGO food seeking behavior seems to be dependent on orexin and specifically on orexin receptors in the VTA, indicating a dopamine component in food reward.[40] Thus there continues to be a great deal of controversy surrounding the role of dopamine in addiction and reward—some researchers believe in a complete dopamine centrality, with dopamine being necessary for all aspects of drug reward, including motivation, acquisition, and reward; others believe dopamine only to be involved in the motivational and associative/learning phases; and the last group who believe dopamine is secondary to the opioids in both learning/acquisition and reward. Regardless of the chicken or the egg debate, which seems to polarize the addiction research community, everyone agrees there is an intimate association between dopamine, opioids, and addiction.

1.3 Hypothesis and Aims

Our goal is to develop mass spectrometry-based methods to analyze the mammalian neuropeptidome. We then plan to apply these methods to the analysis of the opioid peptides in the rat nucleus accumbens. We hypothesize that feeding induces opioid peptide release within the nucleus accumbens. In the following chapters, I will present my research findings, including the development of novel dissection and sample preparation methods for mammalian neuropeptides and the application of these methods to the analysis of feeding induced neuropeptide changes in rats.

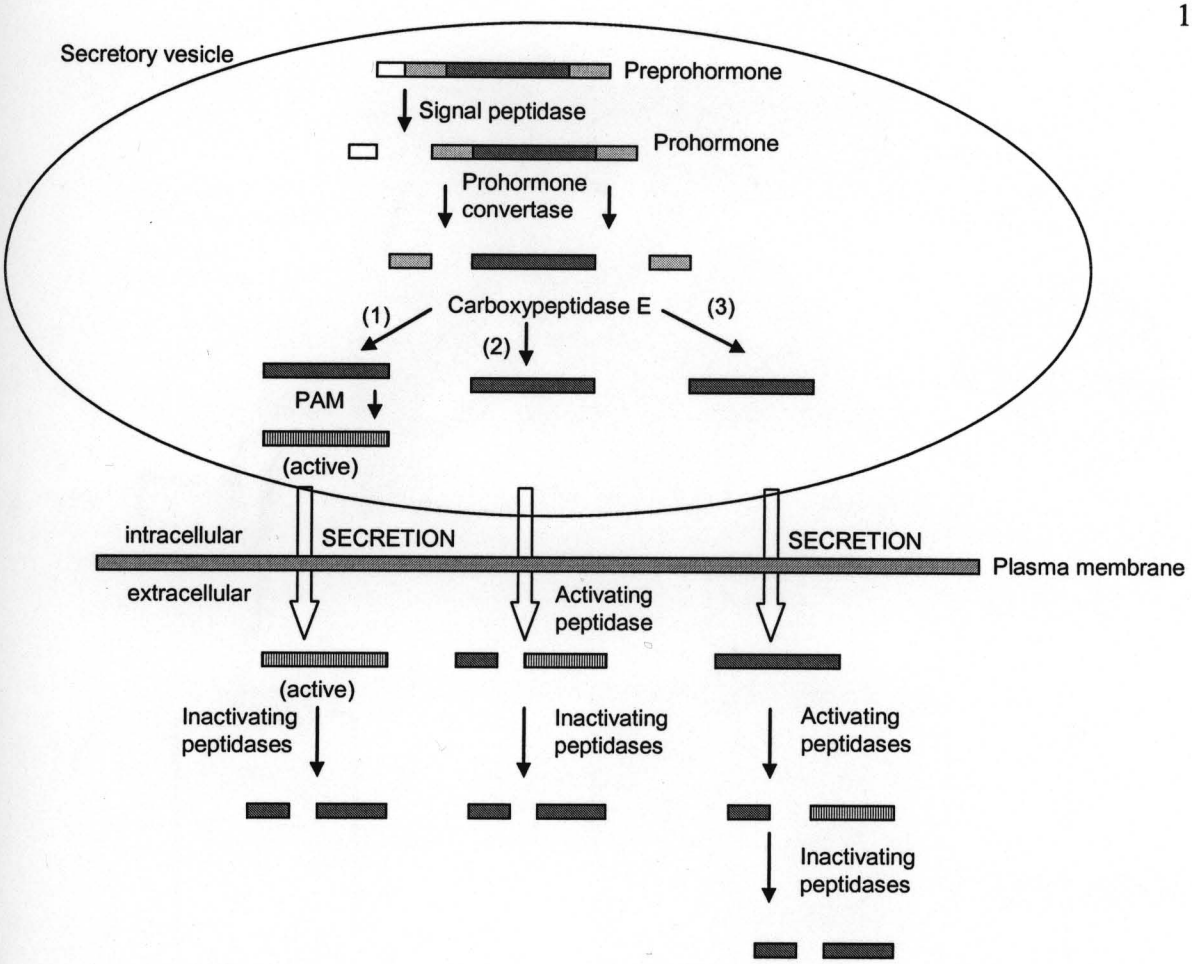

Figure 1.1 Neuropeptide processing pathways. First, the signal peptide is removed from the prepropeptide. The propeptide is then cleaved at dibasic residues by prohormone convertase (PC1/2), followed by carboxypeptidase E which removes the *C*-terminal dibasic residues to yield the mature peptide. Once the peptide is secreted into the extracellular space it is deactivated (or activated) by extracellular peptidases.

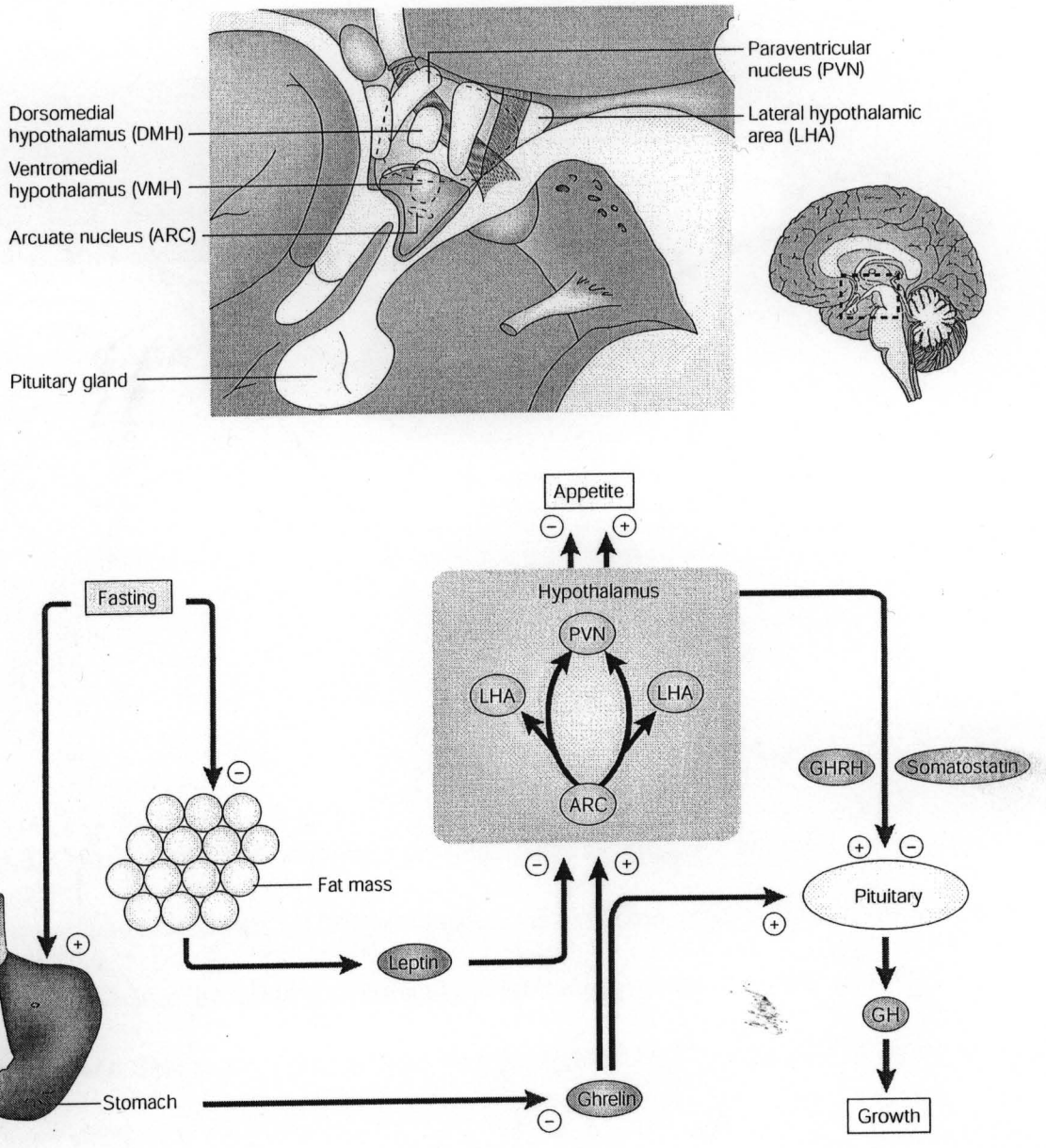

Figure 1.2 Top panel: anatomy of hypothalamus and pituitary. Bottom panel: peripheral regulation of arcuate nucleus 'master' feeding circuitry. Gherlin stimulates feeding. Leptin inhibits feeding.[15]

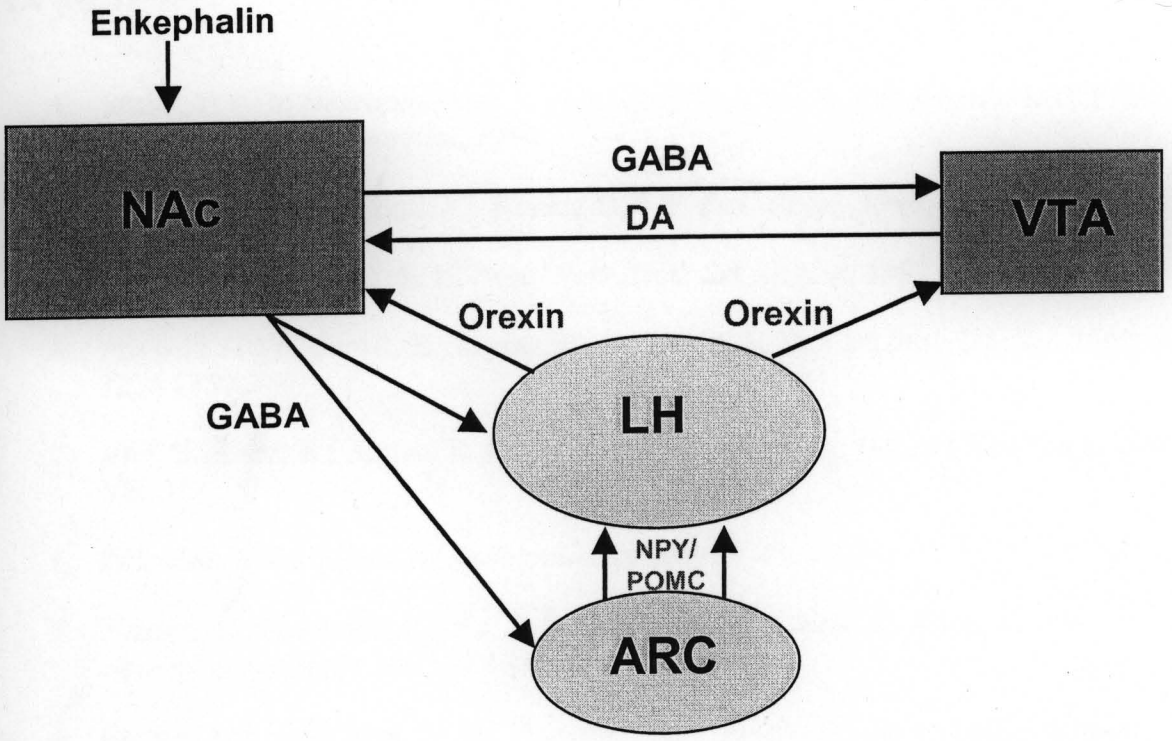

Figure 1.3 Opioid control of hedonistic feeding. After feeding, enkephalins are released within the nucleus accumbens (NAc). This enkephalin release inhibits GABA release from the NAc, disinhibiting the feeding circuits of the lateral hypothalamus (LH) and ARC (arcuate nucleus). Dopamine (DA) release from the ventral tegmental area (VTA) modulates motivational state.

1.4 References

1. Strand, F. L. In *Neuropeptides: Regulators of Physiological Processes.* MIT Press: Cambridge, Massachusettes, 1999.

2. Zhou, A.; Webb, G.; Zhu, X.; Steiner, D. F. *J. Biol. Chem.* **1999**, *30,* 20745-20748.

3. Fricker, L. D.; Snyder, S. H. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *12,* 3886-3890.

4. Prigge, S. T.; Mains, R. E.; Eipper, B. A.; Amzel, L. M. *Cell Mol. Life Sci.* **2000**, *8-9,* 1236-1259.

5. Akil, H.; Herz, A.; Simon, E. J. *Opioids;* Springer-Verlag: Berlin ; New York, 1993; Vol. 104.

6. Pritchard, L. E.; White, A. *Endocrinology* **2007**, *9,* 4201-4207.

7. Hokfelt, T.; Broberger, C.; Xu, Z. Q.; Sergeyev, V.; Ubink, R.; Diez, M. *Neuropharmacology* **2000**, *8,* 1337-1356.

8. Barrett, A. J.; Rawlings, N. D.; Woessner, J. F. *Handbook of proteolytic enzymes;* Elsevier Academic Press: Amsterdam ; San Diego, 2004; .

9. Shrimpton, C. N.; Smith, A. I.; Lew, R. A. *Endocr. Rev.* **2002**, *5,* 647-664.

10. Davis, T. P.; Konings, P. N. *Crit. Rev. Neurobiol.* **1993**, *3-4,* 163-174.

11. Harris, G. C.; Wimmer, M.; Aston-Jones, G. *Nature* **2005**, *7058,* 556-559.

12. Kelley, A. E.; Baldo, B. A.; Pratt, W. E. *J. Comp. Neurol.* **2005**, *1,* 72-85.

13. Zhang, Y.; Proenca, R.; Maffei, M.; Barone, M.; Leopold, L.; Friedman, J. M. *Nature* **1994**, *6505,* 425-432.

14. Hillebrand, J. J.; de Wied, D.; Adan, R. A. *Peptides* **2002**, *12,* 2283-2306.

15. Inui, A. *Nat. Rev. Neurosci.* **2001**, *8,* 551-560.

16. Benoit, S.; Schwartz, M.; Baskin, D.; Woods, S. C.; Seeley, R. J. *Horm. Behav.* **2000**, *4,* 299-305.

17. Bagnasco, M.; Kalra, P. S.; Kalra, S. P. *Endocrinology* **2002**, *2,* 726-729.

18. Horvath, T. L.; Diano, S.; Sotonyi, P.; Heiman, M.; Tschop, M. *Endocrinology* **2001**, *10,* 4163-4169.

19. Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H.; Kangawa, K. *Nature* **1999**, *6762,* 656-660.

20. Nakazato, M.; Murakami, N.; Date, Y.; Kojima, M.; Matsuo, H.; Kangawa, K.; Matsukura, S. *Nature* **2001**, *6817,* 194-198.

21. Ahima, R. S.; Saper, C. B.; Flier, J. S.; Elmquist, J. K. *Front. Neuroendocrinol.* **2000**, *3,* 263-307.

22. Frederich, R. C.; Lollmann, B.; Hamann, A.; Napolitano-Rosen, A.; Kahn, B. B.; Lowell, B. B.; Flier, J. S. *J. Clin. Invest.* **1995**, *3,* 1658-1663.

23. Cone, R. D. *Nat. Neurosci.* **2005**, *5,* 571-578.

24. Clark, J. T.; Kalra, P. S.; Crowley, W. R.; Kalra, S. P. *Endocrinology* **1984**, *1,* 427-429.

25. Boston, B. A.; Blaydon, K. M.; Varnerin, J.; Cone, R. D. *Science* **1997**, *5343,* 1641-1644.

26. Baskin, D. G.; Hahn, T. M.; Schwartz, M. W. *Horm. Metab. Res.* **1999**, *5,* 345-350.

27. Abizaid, A.; Gao, Q.; Horvath, T. L. *Neuron* **2006**, *6,* 691-702.

28. Saper, C. B.; Chou, T. C.; Elmquist, J. K. *Neuron* **2002**, *2,* 199-211.

29. Wise, R. A. *Brain Res.* **1974**, *2,* 187-209.

30. Cabeza de Vaca, S.; Carr, K. D. *J. Neurosci.* **1998**, *18,* 7502-7510.

31. Carr, K. D. *Neurochem. Res.* **1996**, *11,* 1455-1467.

32. Lord, J. A.; Waterfield, A. A.; Hughes, J.; Kosterlitz, H. W. *Nature* **1977**, *5611,* 495-499.

33. Pert, C. B.; Pasternak, G.; Snyder, S. H. *Science* **1973**, *119,* 1359-1361.

34. Kelley, A. E.; Bakshi, V. P.; Haber, S. N.; Steininger, T. L.; Will, M. J.; Zhang, M. *Physiol. Behav.* **2002**, *3,* 365-377.

35. MARTIN, W. R.; WIKLER, A.; EADES, C. G.; PESCOR, F. T. *Psychopharmacologia* **1963**, , 247-260.

36. Majeed, N. H.; Przewlocka, B.; Wedzony, K.; Przewlocki, R. *Peptides* **1986**, *5,* 711-716.

37. Zhang, M.; Kelley, A. E. *Psychopharmacology (Berl)* **1997**, *4,* 350-360.

38. Zhang, M.; Gosnell, B. A.; Kelley, A. E. *J. Pharmacol. Exp. Ther.* **1998**, *2,* 908-914.

39. Zhang, M.; Kelley, A. E. *Neuroscience* **2000**, *2,* 267-277.

40. Zheng, H.; Patterson, L. M.; Berthoud, H. R. *J. Neurosci.* **2007**, *41,* 11075-11082.

41. Berridge, K. C. *Psychopharmacology (Berl)* **2007**, *3,* 391-431.

42. Fields, H. L.; Hjelmstad, G. O.; Margolis, E. B.; Nicola, S. M. *Annu. Rev. Neurosci.* **2007**, , 289-316.

43. Imaizumi, M.; Takeda, M.; Sawano, S.; Fushiki, T. *Behav. Brain Res.* **2001**, *1-2,* 129-136.

44. Robinson, S.; Sandstrom, S. M.; Denenberg, V. H.; Palmiter, R. D. *Behav. Neurosci.* **2005**, *1,* 5-15.

45. Cannon, C. M.; Palmiter, R. D. *J. Neurosci.* **2003**, *34,* 10827-10831.

46. Hnasko, T. S.; Sotak, B. N.; Palmiter, R. D. *Nature* **2005**, *7069,* 854-857.

Chapter 2: Rat Neuropeptidomics by LC-MS/MS and MALDI-FTMS

2.1 Abstract

Recently developed sample preparation techniques employing microwave irradiation have enabled the comprehensive study of endogenous mammalian neuropeptides. These methods reduce interference from post-mortem protein degradation by deactivating proteases via heat denaturation. Alternatively, we have developed a protocol using cryostat dissection and a boiling extraction buffer to achieve a similar effect. This novel methodology greatly reduces post-mortem protein contamination and increases neuropeptide identification without the use of specialized equipment. In addition, a 2D HPLC scheme employing differential pH selectivity in the first and second dimension has been used to enhance neuropeptidome coverage. By using our novel dissection protocol in tandem with 2D RP-RP HPLC, we were able to identify a total of 56 peptides from known neuropeptide precursors, including 17 previously unidentified peptides. The use of cryostat dissection and two-dimensional RP-RP HPLC enhances the detection of novel neuropeptides by deactivating proteases and reducing sample complexity.

2.2 Introduction

Peptides are functionally important molecules in the central nervous system that exhibit diverse regulatory functions in reproduction, food intake, sleep, learning, and memory.[1] A large number of neuropeptides have been studied employing radioimmunoassay, immunohistochemistry, and Edman degradation.[2] These techniques continue to be important but recent advances in mass spectrometry (MS) have provided a new and powerful analytical platform for studying neuropeptides.[3-17] In contrast to immuno-based techniques, mass spectrometry enables neuropeptide detection without *a priori* knowledge of peptide identity and enables "on the fly" sequencing of novel neuropeptides. The strength of mass spectrometry resides in the ability to sensitively and rapidly detect multiple neuropeptides in a non-biased manner. However, these strengths also form the Achilles heel of MS-based neuropeptidomics. This ability to "see" all the ionizable species, which is so powerful, also makes MS highly susceptible to signal suppression due to the presence of unwanted contaminants, such as lipids, salts, and surfactants. Highly purified and enriched peptide samples are a prerequisite for the success of any MS-based neuropeptide experiment and thus, the development of effective sample preparation protocols is of utmost importance.

In mammals, non-specific protein degradation occurs rapidly after death. This post-mortem protein degradation produces a large abundance of non-neuropeptide protein fragments which effectively mask the signal of the less abundant endogenous neuropeptides. In order to curtail post-mortem protein degradation, a number of methods have been employed, including the use of transgenic mice lacking carboxypeptidase E activity, focused

microwave irradiation for animal sacrifice, and rapid post-sacrificial microwave irradiation.[18-24] Microwave irradiation heat deactivates endogenous proteases and arrests post-mortem protein degradation, resulting in a clean neuropeptide sample for MS analysis.

Microwave irradiation is a proven technique to reduce post-mortem protein degradation. However, the equipment employed in focused microwave irradiation sacrifice is expensive and requires the restraint of the animal during sacrifice. In contrast, post-sacrifice microwave irradiation employs an inexpensive household microwave. However, this technique still requires the time and animals to develop consistent protocols.[18] Based on experiments using non-microwave protease heat deactivation such as boiling extraction,[25-28] our lab has developed a neuropeptide extraction protocol which combines snap freezing, cryostat dissection and a boiling extraction buffer to curtail protein degradation without the use of microwave irradiation.

In addition to the protein degradation issues, the reduction of sample complexity by enhanced fractionation techniques is essential to the success of MS-based peptidomic experiments. In proteomics, multi-dimensional liquid chromatography (MDLC) is routinely used to improve the resolution of complex mixtures of proteins or peptides.[29-32] In a complex sample, the improved resolution of two-dimensional separations enhances the detection of low abundance proteins. A number of different MDLC methodologies exist, including two-dimensional protein-peptide or peptide-peptide chromatographic schemes. Yates et al. popularized a particular form of MDLC, termed Multidimensional Protein Identification Technology (MudPIT), for proteomic analyses.[33] MudPIT employs the online coupling of two orthogonal chromatographic steps, strong cation exchange (SCX) and reverse phase (RP), to improve sample resolution and increase proteome coverage. Recently, this

commonly used proteomics methodology has been successfully applied to the peptidomic analysis of both *C. elegans* and *D. melanogaster*; the use of MudPIT in these experiments resulted in enhanced detection of low abundance, novel peptides.[34, 35]

Only one study has employed MudPIT in a mammaliam neuropeptidomic study; however, the experiment yielded very few neuropeptides due to a lack of protease deactivation during sample preparation.[36] In our study, we have applied a newly developed two-dimensional separation scheme to effectively enhance the detection of neuropeptides in rat hypothalamus and striatum. In contrast to MudPIT, this approach employs an off-line high pH reverse phase separation followed by a standard on-line low pH reverse phase separation, as recently developed by Gilar et al.[37, 38] The first and second dimensions exhibit different selectivities due to charge changes in acidic and basic amino acids under different pH conditions. During the high pH dimension, the basic amino acids are neutral and the acidic amino acids are negatively charged. Conversely, during the low pH dimension, the basic amino acids are positively charged and the acidic amino acids are neutral. The resulting selectivity change is sufficient to produce orthogonality between the first and second RP separations (see Figure 2.1). This 2D RP-RP LC/MS/MS method coupled with our novel cryostat dissection and enhanced extraction procedure has resulted in the identification of 56 previously identified neuropeptides and 17 novel neuropeptides.

2.3 Experimental Section

2.3.1 Animal Sacrifice and Cryostat Dissection

Rats were anesthetized with halothane from Sigma (St. Louis, MO) and then sacrificed by decapitation. The brain was then rapidly removed (<90 s) and snap frozen in 2-methylbutane from Sigma cooled by dry ice. The frozen brain was then sectioned in 300 μm thick slices on a cryostat from Leica (Wetzlar, Germany) with a compartment temperature of –15 °C. The hypothalamus and striatum were removed with a 3 mm tissue punch and stored in 1.5 mL tubes at –80 °C until extraction.

2.3.2 Extraction with Boiling

Hypothalamus and striatum punches (20-30 mg per rat) were removed from –80 °C freezer. Immediately, 300μL of boiling water was added to each sample and then placed in a boiling water bath. After 10 minutes, the samples were removed, placed on ice, and the water was decanted and saved. 300 μL extraction buffer of ice cold dilute acetic acid (0.25% acetic acid) was added to the tissue. The samples were then homogenized with a sonic dismembrator from Mixsonix (Farmingdale, NY) for 5 seconds on a setting of 3, followed by 30 seconds of vortexing. The homogenization/vortexing was repeated every ten minutes for an hour. The homogenized sample was then spun at 20,000 *g* for 20 minutes at 4 °C to pelletize the insoluble material. The supernatant was decanted and combined with the boiling water fraction (see above) and the combined sample was filtered through a 10 kD

molecular weight cut-off (MWCO) tube from Sartorius (Goettingen, Germany) at room temperature. MWCO filtration took between 20 and 30 minutes.

2.3.3 Extraction on Ice

Striatum punches were extracted with a sonic dismembrator (as above) into 300 μL ice cold dilute acetic acid without pre-boiling in water. The sample was spun-down and passed through a 10 kD molecular weight cut-off tube at room temperature. MWCO filtration took between 20 and 30 minutes.

2.3.4 Comparison of Cold and Boiling Extraction Procedures

Striatum extracts from either the boiling or on ice procedures (see above) were concentrated and desalted with a C_{18} spin column from Sartorius to 4 μL. 0.4 μL of the concentrated/desalted sample was then spotted with 0.4 μL 2,5-dihydroxybenzoic acid (DHB) from ICN (Costa Mesa, CA) on a MALDI probe and analyzed by MALDI-FTMS from IonSpec (Lake Forest, CA). Mass spectra were collected using an in-house developed in-cell accumulation method, as previously described.[39]

2.3.5 Comparison of Extraction Buffers

In order to assess the extraction efficiency of different buffers, rat punches from a single animal (n=2 experiments) were extracted by boiling and microsonication (see above) in the following buffers: dilute acetic acid (0.25% acetic acid in water), acidified MeOH (9% acetic acid in 90% MeOH), or SDS+KCl (8 mM SDS + 145 mM KCl + 0.25% acetic acid in water), as previously reported.[21] The flow through from the 10 kD molecular weight cut-off

tubes was dried-down to ~50 μL by vacuum drying. 5 μL of each extraction was analyzed by LC/MS and the extracted ion chromatogram (EIC) for selected neuropeptides was compared (see Table 2.1).

2.3.6 Neuropeptide Analysis by MALDI-FTMS

Hypothalamus and striatum punches from a single animal (n=3 experiments) were extracted in dilute acetic acid by boiling and microsonication (see above). The 10 kD molecular weight cut-off flow through was concentrated and desalted with a C_{18} spin column from Sartorius to 4 μL. 0.4 μL of the concentrated/desalted sample was then spotted with 0.4 μL DHB on a MALDI probe and analyzed by MALDI-FTMS. Mass spectra were collected using an in-house developed in-cell accumulation method, as previously described.[39]

2.3.7 Neuropeptide Analysis by 1D LC/MS/MS

Hypothalamus and striatum punches from a single animal (n=2 experiments) were extracted in dilute acetic acid by boiling and microsonication (see above). The 10 kD molecular weight cut-off flow through was dried down to ~20 μL and then adjusted to pH 2 with formic acid. 6 μL of sample was analyzed by LC/MS/MS.

2.3.8 Neuropeptide Analysis by 2D RP-RP LC/MS/MS

Hypothalamus and striatum punches were pooled from three animals (n=2 experiments, a total of 6 animals) and extracted in acetic acid by boiling and microsonication. The 10 kD molecular weight cut-off flow through was dried down to ~30 μL. Triethylamine (TEA) from Sigma was used to adjust the sample to pH 11. 20 μL of the concentrated

sample was injected via a Rainin HPLC onto high pH stable RP column (Phenomenex Gemini C_{18}; 2.1 x 150 mm) at a flow rate of 150 μL/minute. The peptides were eluted with a mobile phase B: 5-45% gradient over 60 minutes (A: 20 mM TEA in water pH 11.5, B: 20 mM TEA in acetonitrile). Fractions were collected every 4 minutes for 60 minutes (15 total fractions).

Collected fractions were evaporated under vacuum to dryness and then reconstituted into 10 μL of 0.1% formic acid. 5 μL of each reconstituted fraction was delivered to the ESI-QTOF instrument according to the LC/MS/MS instrument set-up (detailed below).

2.3.9 LC/MS/MS Analysis

A capillary HPLC (Waters, Milford, MA) system was used to deliver the specified volume of sample to a trap column (LC Packings PepMap C_{18}, 300 μm x 5 cm; Amsterdam, The Netherlands) via an isocratic flow of mobile phase A: (0.1% formic acid in water) at a rate of 30 μL/minute for 3 minutes. The flow rate was then switched to ~250 nL/minute and the peptides were flushed onto the analytical column (Microtech C_{18}, 75 μm x 15 cm; Vista, CA) and eluted via a mobile phase B: 5-45% (0.1% formic acid in acetonitrile) gradient over 30 minutes into a nanoelectrospray ionization (nESI) quadrupole time-of-flight (QTOF) mass spectrometer from Waters (QTOF Micro).

The peptides were detected in positive mode. Data was collected in both MS-only mode and MS/MS data-dependent acquisition (DDA) mode. The DDA survey and MS-only mass-to-charge range was from *m/z* 400 to 2000. The intensity threshold for switching from the survey scan to MS/MS was set at 15 ion counts. The scan time was: 0.9 s, inter-scan time: 0.1 s, capillary voltage: 3800 V, cone voltage: 35 V.

2.3.10 Data Analysis

Micromass ProteinLynx 2.1 was used to process the LC/MS/MS data. Deisotoping was performed using the "slow" function. The resulting .pkl files were searched against the SwissProt database using the on-line version of Mascot. The enzyme was specified as "none". The peptide mass tolerance was set at 200 ppm and the MS/MS mass tolerance was set at 0.2 Da. A cascade of three separate searches was performed for variable modifications, including: C-terminal amidation and N-terminal acetylation; methionine oxidation and C-terminal amidation; phosporylation of tyrosine, threonine, and serine and C-terminal amidation. Peptides of interest which did not yield significant Mascot scores were verified by *de novo* sequencing, using Micromass MassLynx 4.0 PepSeq software. A *de novo* sequence matching at least three consecutive amino acids of the presumed sequence was considered a positive identification.

MALDI data was analyzed by direct comparison of measured molecular weights of the putative peptide peaks to an in-house generated database of known neuropeptides. An experimental mass matching within 25 ppm of a known neuropeptide mass was considered a positive identification.

2.4 Results and Discussion

2.4.1 Comparison of Cold and Boiling Extraction Buffers

The MALDI-FTMS spectra from the boiled extracts were compared to spectra from the cold extractions. As can be seen from Figure 2.2, the boiled extraction yields an uncomplicated spectrum composed mostly of known neuropeptides (Figs. 2.2A and 2.2B),

whereas the cold extraction yields a complex spectrum with few identifiable neuropeptides (Figs. 2.2C and 2.2D). The complex spectrum of the cold extraction is the result of protease-produced protein degradation, possibly due to the molecular weight filtration step performed at room temperature. It is obvious from these results and previous results that the endogenous protease deactivation is critical to the success of neuropeptidomic studies in mammals.[18, 21, 22]

As mentioned in the introduction, microwave irradiation has become the technique of choice to circumvent the problems of protein degradation. These techniques deactivate proteases during sacrifice or shortly after death. In contrast, the technique presented in this study introduces a delay between death and protease deactivation (<90 s). The only study addressing effects of delayed protease deactivation found an insignificant amount of protein degradation during a 20-second delay between sacrifice and microwaving.[18] It is unknown the degree of protein degradation produced with a longer delay, however, the approximately 90-second delay encountered in this study does not seem to affect the downstream quality of MS data and may in fact allow the study of extracellular neuropeptide processing (see discussion below).

Additionally, the cryostat technique has an advantage over microwave irradiation in regards to simplicity. Microwave irradiation either requires specialized equipment (focused microwave sacrifice) or the optimization of post-mortem microwave settings. The cryostat technique can be performed using equipment commonly available in any neuroscience lab.

2.4.2 Comparison of Overall Extraction Efficiency

To assess the relative effectiveness of various extraction techniques and buffers, a comparison was made between the relative amounts of five neuropeptides extracted by three different conditions: dilute acetic acid, acidified MeOH, and SDS+KCl. Using MassLynx software, the peak area of each peptide was determined by extracted ion chromatogram (EIC). These areas were averaged (n=2) and then combined for a total peak area (see Table 2.1).

As can be seen from Table 1, the dilute acetic acid gave the highest overall extraction yields and the highest yields for all five individual neuropeptides, acidified MeOH was half as efficient as dilute acetic acid, and SDS+KCl gave the worst results with an overall yield almost ten-fold less than the dilute acetic acid.

2.4.3 Comparison of 1D and 2D LC/MS/MS Experiments

MDLC methodologies enhance the detection of low abundance peptides by improving the resolution of a complex sample. The effectiveness of MDLC separations depends on the degree of difference, or orthogonality, between the separation chemistries of the first and second dimensions. The higher the peak capacity and resolution of the two dimensions the more effective they are at reducing sample complexity. Traditionally, strong cation exchange (SCX), which separates on peptide charge, followed by reverse phase (RP) has been the conventional set-up for MDLC experiments. However, SCX has certain limitations, such as low resolution, poor overall peak capacity, and poor compatibility with subsequent MS analysis due to high salt concentrations. In addition, unlike tryptic peptides, neuropeptides do not commonly carry basic residues at the C-terminus and, since SCX

separates based on charge, this lack of charge decreases the separation selectivity and efficiency.

In contrast to SCX, RP exhibits high resolution, good peak capacity and MS compatibility.[37, 38] However, until the recent development of high pH stable RP columns, orthogonality could not be achieved by RP in tandem. Gilar et al. exploited the high pH stability of these new columns to perform RP-RP separations employing a high pH mobile phase in the first dimension and a standard low pH mobile phase in the second dimension.[37, 38] As mentioned in the introduction, the pH change produces a concurrent change of selectivity between the first and second dimensions. This change in selectivity produces orthogonality between the first and second RP separations, resulting in an effective two-dimensional separation. This RP-RP method has several advantages over SCX methods: RP is a more robust, user- friendly system, and exhibits both higher peak capacity and superior resolution.

We employed this two-dimensional RP-RP approach to improve resolution and reduce sample complexity in our neuropeptide extracts. Punches from three rats were pooled and fractionated by high pH off-line RP followed by low pH on-line LC/MS/MS. The higher loading capacity of the 2.1 mm RP column in the first dimension allowed for the injection of three times more extraction solution than in the one dimensional separation. The resulting fractions from the high pH fractionation exhibited a reduced complexity and enhanced MS/MS data in comparison to the one dimensional analysis. This resulted in a significant increase in the total number of neuropeptides identified in the two-dimensional analysis compared with that from the one-dimensional analysis (see Table 2.2). However, five peptides were detected exclusively in the 1D experiment, possibly due to sample loss in the

2D experiment during offline fractionation. As shown in Table 2.2, several peptides that were not detected by LC ESI MS/MS methodologies were identified by MALDI FTMS accurate mass measurements. This enhanced neuropeptidome coverage highlights the advantages of using a combined approach that involves complementary analytical methods.

2.4.4 Identified Neuropeptides

In the current study, we identified 56 putative neuropeptides, 17 of which were novel (see Table 2.2). As outlined, three different procedures were used to identify neuropeptides: MALDI-FTMS, 1D LC/MS/MS and 2D LC/MS/MS. MALDI-FTMS offers very high mass accuracy and good sensitivity. However, due to the less efficient fragmentation of singly charged ions, MALDI does not produce sequence quality MS/MS spectra unless the peptide exhibits exceptionally strong signal intensity. As a result, the peptides identified by MALDI were determined by mass measurement alone with a 25 ppm error threshold. Due to the high sensitivity of the instrument, we were able to identify a large number of known neuropeptides. However, it was impossible to identify novel neuropeptides without MS/MS data.

In addition to the known neuropeptides identified by MALDI, one novel peptide and a number of known neuropeptides were identified by 1D LC/MS/MS experiments via database searching and *de novo* sequencing. However, the 2D LC/MS/MS experiments identified the highest number of known and novel neuropeptides. As with the one dimensional experiment, both the novel and known neuropeptides were identified by database searching and verified with *de novo* sequencing (see Figure 2.3 for representative spectra).

Many known neuropeptides from various neuropeptide precursors were identified, including: cocaine- and amphetamine-related transcript protein (CART), proenkephalin, prochromogranin, promelanocortin (POMC), protachykinin, proSAAS, pituitary adenylate cyclase-activating polypeptide (PACAP), provasopressin, vascular endothelial growth factor (VGF) (see Table 2.2). Also, a number of peptides from non-specific protein degradation were also identified, including peptides from: hemoglobin, cytochrome c, thymosin, fibrinogen, aldolase, calmodulin and VAP-33 (see Table 2.3). Many of these, including the peptides from calmodulin, fibrinogen, and aldolase, are derived from C-terminal side cleavage of aspartic acid. Aspartic acid bonds are known to be extremely acid labile and these peptides are probably due to acid exposure during extraction.[18] Peptide fragments from hemoglobin and fibrinogen are greatly reduced by both microwave fixed tissue[18] and cryostat dissection coupled with heat deactivation (see Figure 2.2).

Seventeen novel peptides were identified from five different proneuropeptide precursors (see Table 2.2). These include peptides from prochromogranin, proenkephalin, proSAAS, diazepam-binding inhibitor protein (DBI), and phosphatidylethanolamine-binding protein (PEBP). Of these, only three are cleaved at classical dibasic or monobasic neuropeptide processing sites. A number of these novel peptides are cleaved at the C-terminal side of leucine or tryptophan.

Leucine specificity has been reported in the processing of the endorphins as well as in the processing of CLIP.[18] Furthermore, these neuropeptides are not detected in samples which have been microwave irradiated immediately after sacrifice. Presumably, microware irradiation arrests the secretory machinery necessary for vesicle fusion and neuropeptide release. Che et al. hypothesize that the final processing of these neuropeptides occurs

extracellularly by an unidentified Leu-X specific enzyme.[18] Upon microwaving, these neuropeptides are no longer released and therefore unavailable for extracellular processing. In comparison to the studies by Che et al., our experiments have an increased time delay between animal sacrifice and protease heat deactivation which would allow for the release and extracellular processing of these peptides. This would explain our discovery of a number of Leu-X processed peptides derived from proSAAS, PEBP, and vasopressin. However, this is speculatory and neither the previously identified Leu-X processed endorphin nor CLIP peptides were found in this study.

Obviously, it is impossible to infer bioactivity from the sequence of a neuropeptide. However, there are a couple of signature modifications which are commonly found in bioactive neuropeptides. The most common modification is C-terminal amidation and the other is N-terminal acetylation.[1] This study identified a number of C-terminally amidated neuropeptides. In addition, we also identified a novel N-terminally acetylated peptide from a known neuropeptide precursor: diazepam-binding inhibitor (DBI). DBI is known to block the activity of diazepam (an anxiety relieving compound) by competing at the diazepam binding site on the GABA receptor. This binding has been shown to cause anxiety-like effects in rats and has been implicated in drug and alcohol addiction.[40] The actions of DBI are produced by the proteolytic processing of the 9 kD propeptide into two bioactive peptides—octadecaneuropeptide (ODN) and triakontatetaneuropeptide (TTN).[41] TTN is produced by the cleavage on the C-terminal side of Lysine-16 and Lysine-50. ODN is produced by the cleavage on the C-terminal side of Lysine-32 and Lysine-50. We did not identify either of these neuropeptides; however, we did identify three different fragments of

the acetylated N-terminal fragment of DBI from C-terminal side cleavage of Lysine-16 (the same cleavage that produces TTN), Lysine-13, and Leucine-15 (see Figure 2.4).

Although it is impossible to infer bioactivity to this peptide, the N-terminal acetylation may imply bioactivity of the peptide. Acetylation is known to increase peptide-receptor interactions and is necessary for the bioactivity of many neuropeptides.[42] Unfortunately, acetylation is also the most commonly known post-translational protein modification in mammalian cells, so it is not a definitive indication of neuropeptide activity.[42] Also, as can be seen from Figure 2.4, DBI does not possess a signal peptide and thus can not be secreted through the regulated pathway. This contrasts with the majority of classic neuropeptides which are secreted in the regulated pathway. In addition, DBI is known to perform housekeeping duties in the cell by delivering acyl-CoA to the mitochrondrial membrane. Thus, this peptide could either be a neurotransmitter or simply a fragment from a ubiquitous housekeeping protein.

In contrast to the N-terminal fragment of DBI, the majority of the other novel neuropeptides are fragments of known protease/peptidase inhibitors. ProSAAS is known to be a substrate and inhibitor of prohormone convertase 1 (PC1) [43]; PEBP is known to be a substrate and inhibitor of several extracellular serine proteases [44]; chromogranin C belongs to a family of serine protease inhibitors which are potentially involved in amyotrophic lateral sclerosis (ALS) pathology.[45] A large number of ProSAAS, chromogranin, and PEBP peptides have been reported by a number of groups.[18, 21, 22] The large number peptides and diversity of cleavage sites produced from these propeptide precursors is consistent with their activity as protease inhibitors.

The novel peptide encoded by proenkephalin is produced by cleavage at alanine, a residue not commonly associated with classic neuropeptide processing. It is unknown if this cleavage is produced by extracellular processing or is simply a sample preparation artifact.

2.5 Conclusions

In summary, we have shown the analytical power of coupling a novel neuropeptide extraction technique with two-dimensional RP-RP LC/MS/MS. The combined use of high resolution high accuracy MALDI FTMS and two-dimensional LC ESI MS/MS enabled improved coverage of the rat brain neuropeptidome. As a result, both previously known and novel neuropeptides were identified, including a novel N-terminal fragment of diazepam binding inhibitor (DBI) and numerous novel processing products from neuropeptide prohormones. These results, together with further improvements in sample preparation and tandem mass spectrometric capabilities, promise to accelerate the pace of neuropeptide discovery and better understanding of neuropeptide signaling.

Table 2.1 Comparison of Overall Extraction Efficiencies from Extracted Ion Chromatograms

Extract Conditions	**Leu-Enk: 556.3 m/z**	**Met-Enk: 574.2 m/z**	**Pro-Enk: 693.8 m/z**	**Pro-MCH: 724.3 m/z**	**PACAP: 820.9 m/z**	**Total:**
Acetic Acid	6620	10700	48100	3910	3780	14622
Acid MeOH	3550	4490	23610	3780	980	7282
SDS/KCl	1670	2360	5590	670	230	2104

Table 2.2 Peptides from known neuropeptide precursors[a,b,c,d]

Precursor	Sequence	m/z	Cal. Mass	2D LC	1D LC	MALDI
CART	R.IPIYE.K	634.26	634.34		X	
CART	**R.APGAVLQIEALQEVLKKL.K***	**640.64**	**1919.15**	**X**		
CART	**R.APGAVLQIEALQEVLKKLKS.K***	**712.34**	**2134.27**	**X**		
Chromogranin B	R.SFAKAPHLDL.K	549.74	1097.59	X	X	
Chromogranin B	R.LLDEGHDPVHESPVDTA.K	915.91	1829.84	X		
Chromogranin B	R.LGALFNPYFDPLQWKNSDFE.K	1201.03	2400.14	X	X	
Chromogranin C	**R.IPAGSLKNEDTPN.R ***	**678.33**	**1354.67**	**X**		
Chromogranin C	**F.PLMYEENSRENPF.K ***	**813.29**	**1624.72**	**X**		
Chromogranin C	R.TNEIVEEQYTPQSL ATLESVFQELGKLTGPSNQ.K	1217.59	3649.80		X	
DBI	**.SQADFDKAAEEVK.R + Acetyl (N-term) ***	**740.32**	**1478.69**	**X**		
DBI	**.SQADFDKAAEEVKRL.K + Acetyl (N-term) ***	**583.57**	**1747.87**	**X**		
DBI	**.SQADFDKAAEEVKRLK.T + Acetyl (N-term) ***	**626.25**	**1875.97**	**X**		
Endorphin	R.YGGFLRKYPK.R	1228.69	1227.68			X
Enkephalin	R.YGGFMRF.-	877.40	876.40			X
Enkephalin	R.SPQLEDEAKELQ.K	1386.66	1385.67	X	X	X
Enkephalin	R.VGRPEWWMDYQ.K	1466.65	1465.65	X	X	X
Enkephalin	**Y.PVEPEEEANGGEILA.K ***	**1553.70**	**1552.73**		**X**	
Enkephalin	R.YGGFL.K	556.24	555.27	X	X	
Enkephalin	K.YGGRFM.K	574.17	573.23	X	X	
MCH	R.EIGDEENSAKFPI.G + Amide (C-term)	1447.71	1446.68	X	X	X
PACAP	R.GMGENLAAAAVDDRAPLT.K	886.34	1770.86	X		
PEBP	**W.DDSVPKLHDQLAGK.- ***	**508.21**	**1521.79**	**X**		
PEBP	W.AGPLSLQEVDEPPQHAL.R	900.87	1799.91	X		
PEBP	V.DYGGVTVDELGKVLTPTQV.M	995.92	1990.03	X		
PEBP	**W.AGPLSLQEVDEPPQHALRV.D ***	**685.95**	**2055.07**	**X**		
POMC	R.SYSMEHFRWGKPV.G +Amide (C-term)	1622.78	1621.77			X
POMC	R.SYSMEHFRWGKPV.G +Acetyl (N-term) +Amide (C-term)	1664.78	1663.79			X
POMC	R.SYSMEHFRWGKPV.G +Diacetyl (N-term) +Amide (C-term)	1706.80	1705.81			X
POMC	R.YGGFMTSEKSQTPLVT.L +Acetyl (N-term)	1787.86	1786.85			X
POMC	R.AEEETAGGDGRPEPSPRE.G Amide (C-term)	1882.89	1881.85			X
POMC	R.YGGFMTSEKSQTPLV TLFKNAIIKNA +Acetyl (N-term)	2900.50	2899.51			X
POMC/CLIP	R.RPVKVYPNVAENESAEAFP.L	2117.05	2116.06			X
POMC/CLIP	R.RPVKVYPNVAENESAEAFPLE.F	2359.23	2358.19			X
POMC/CLIP	R.RPVKVYPNVAENESAEAFPLEF.K	2506.31	2505.26		X	X
SAAS	S.APLAETSTPLRL.R	634.82	1267.71	X		
SAAS	A.SAPLAETSTPLRL.R	678.35	1354.75	X		
SAAS	**R.AVPRGEAAGAVQEL.A ***	**684.33**	**1366.72**	**X**		
SAAS	A.ASAPLAETSTPLRL.R	713.85	1425.78	X		
SAAS	**R.AVPRGEAAGAVQELA.R ***	**719.84**	**1427.76**	**X**		
SAAS	**L.SAASAPLAETSTPLR.L ***	**736.33**	**1470.77**	**X**		
SAAS	**S.AASAPLAETSTPLRL.R***	**749.41**	**1496.82**	**X**		
SAAS	L.SAASAPLAETSTPLRL.R	792.89	1583.85	X		
SAAS	R.SLSAASAPLAETSTPLR.L	836.42	1670.88	X		
SAAS	R.LENSSPQAPARRLLPP.-	582.62	1744.96	X		
SAAS	R.SLSAASAPLAETSTPLRL.R	892.98	1783.97	X	X	X
SAAS	**V.PRGEAAGAVQELARALAHLLEAE RQE.R***	**697.00**	**2784.46**	**X**		

SAAS	R.AVPRGEAAGAVQELARALAHLLEA ERQE.R	739.54	2954.57	X		
Somatostatin	R.SANSNPAMAPRE.R	622.76	1243.56	X	X	
Somatostatin	K.AGCKNFFWKTFTSC.-	1637.71	1636.74			X
Stathmin	-.ASSDIQVKELEKRASGQAF.E + Acetyl (N-term)	702.68	2105.08		X	
Tachykinin A	R.RPKPQQFFGLM.G + Amide (C-term)	1347.74	1346.73	X		X
Vasopressin	L.VQLAGTQESVDSAKPRVY.-	649.93	1947.01	X		
Vasopressin	A.CYFQNCPRG.G +Amide (C-term)	1084.45	1083.44			X
Vasopressin	R.ELYENKPRRPYIL.K + Pyroglutamic acid (N-term)	1672.90	1671.91			X
VGF	**R.VPERAPLPPSVPSQFQ.A***	**874.95**	**1747.93**	**X**		
VGF	**R.VPERAPLPPSVPSQFQA.R***	**910.48**	**1818.96**	**X**		

[a]X Indicates type of MS used to identify peptide.
[b]**Bold** with * indicates novel peptides.
[c]Amino acids separated by "." indicate cleavage sites.
[d]Amino acid ending with "-" indicate the C-terminal end of the protein sequence.

Table 2.3 Peptides from Protein Fragments [a,b,c]

Protein Fragments	Sequence	m/z	Cal. Mass	2D LC	1D LC	MALDI
Albumin	K.LGEYGFQNAILVR.Y	740.37	1478.79	X		
Alpha-1-antiproteinase	M.IVESETQSPLFVGKVIDPTR.-	739.00	2214.19	X		
Calmodulin	D.GQVNYEEFVQMMTAK.-	887.83	1773.81	X		X
Calmodulin	D.GDGQVNYEEFVQMMTAK.-	973.89	1945.85	X		
Cytochrome c oxidase	L.GISTPEELGLDKV.-	679.29	1356.71	X		
Cytochrome c oxidase	L.NELGISTPEELGLDKV.-	857.35	1712.88	X		
Fibrinogen	A.DTGTTSEFIEAGGDIR.G	834.85	1667.76	X		
Fibrinogen	T.ADTGTTSEFIEAGGDIR.G	870.39	1738.80	X		
Fructose-bisphosphate aldolase	D.GGAAAQSLYVANHAY.-	746.79	1491.71	X		
Hemoglobin	A.SVSTVLTSKYR.-	620.81	1239.68	X		
Hemoglobin	-.VHLTDAEKAAVN.G	634.32	1266.66	X		
Hemoglobin	L.ASVSTVLTSKYR.-	656.33	1310.72	X		
Hemoglobin	L.LVVYPWTQRY.F	662.85	1323.70		X	
Delta3 isomerase	S.LVNELTFTAR.K	582.30	1162.63	X		
Peroxiredoxin 5	M.APIKVGDTIPSVEVF.E	786.40	1570.86	X		
Proteasome endopeptidase	C.RSGSAADTQAVADAVTY.Q	841.87	1681.79	X	X	
Thymosin beta-4	-.SDKPDMAEIEKFDK.S + Acetyl (N-term)	565.6	1693.79	X		

[a]X Indicates type of MS used to identify peptide.
[b]Amino acids separated by "." Indicate cleavage sites
[c]Amino acid ending with "-" indicates the C-terminal end of the protein sequence.

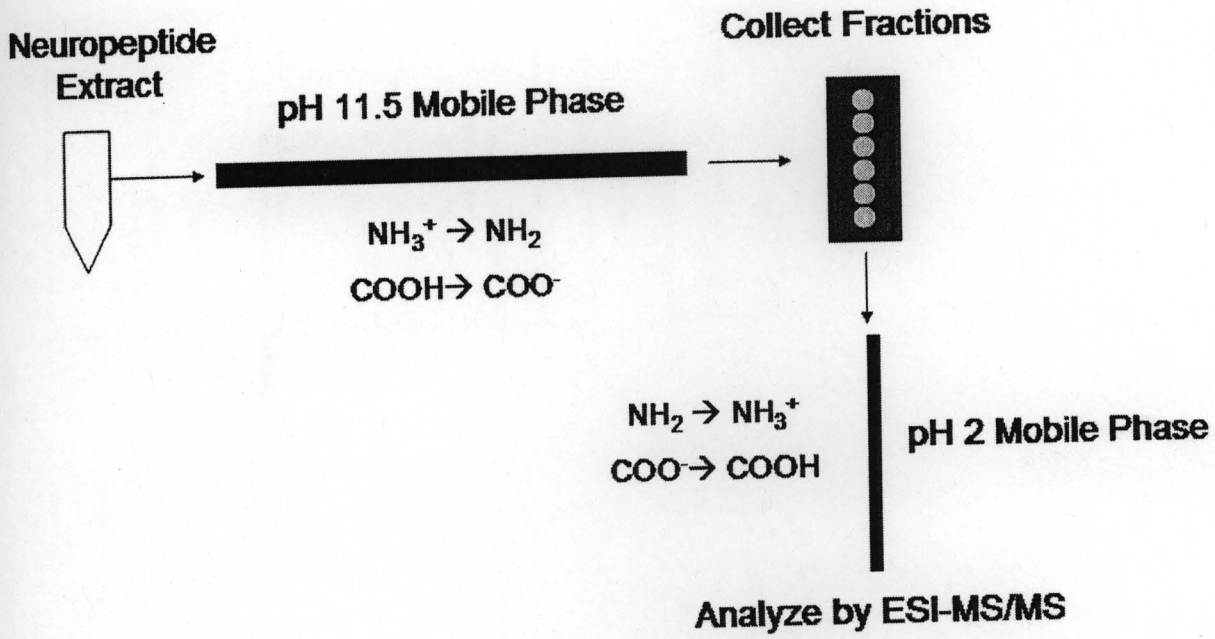

Figure 2.1 Schematic diagram showing the two-dimensional reverse phase RP-RP LC coupled to electrospray ionization tandem mass spectrometry.

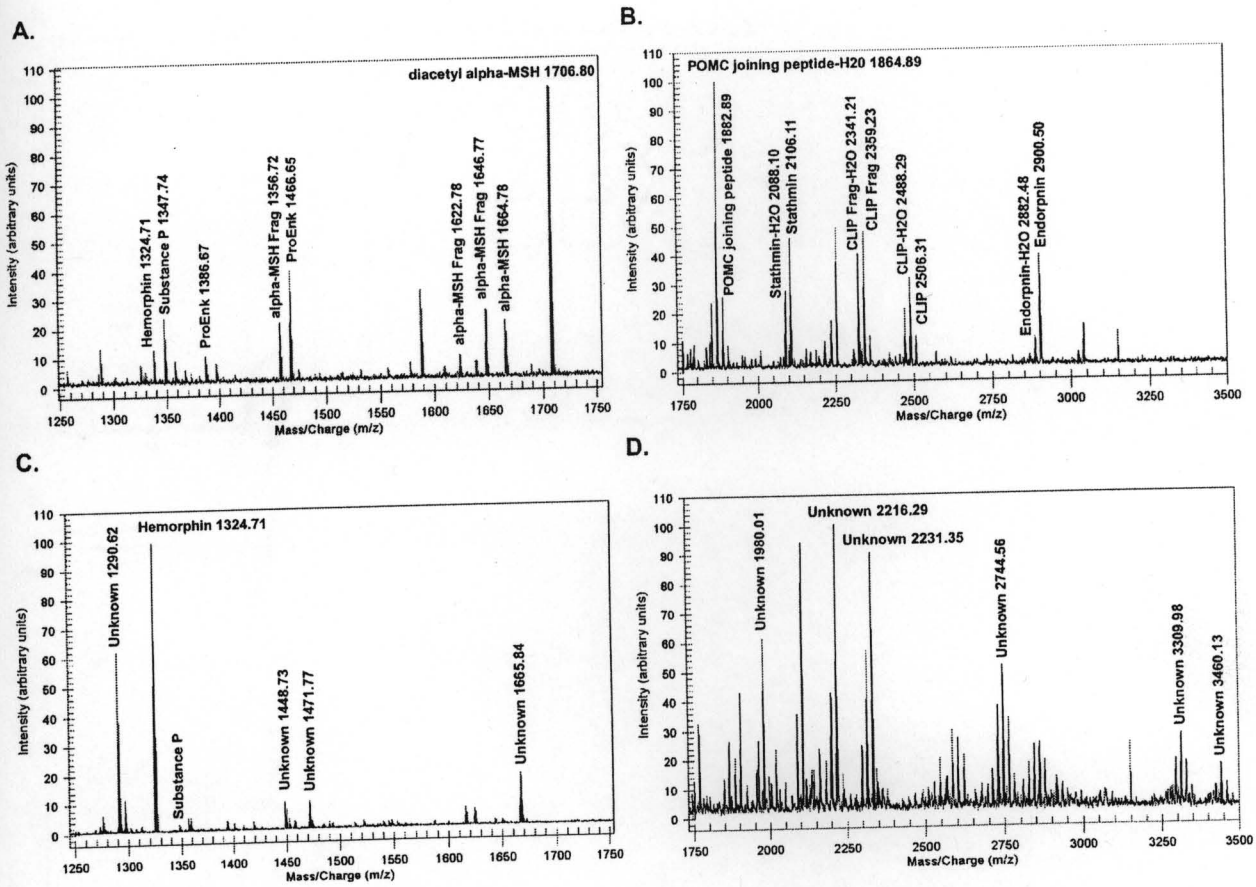

Figure 2.2 MALDI FTMS spectral comparison of different neuropeptide extraction protocols. (**A**) Boiled extraction in acetic acid (low molecular weight, *m/z* 1250-1750), (**B**) Boiled extraction in acetic acid (high molecular weight, *m/z* 1750-3500), (**C**) Cold extraction in acetic acid (low molecular weight, *m/z* 1250-1750), and (**D**) Cold extraction in acetic acid (high molecular weight, *m/z* 1750-3500).

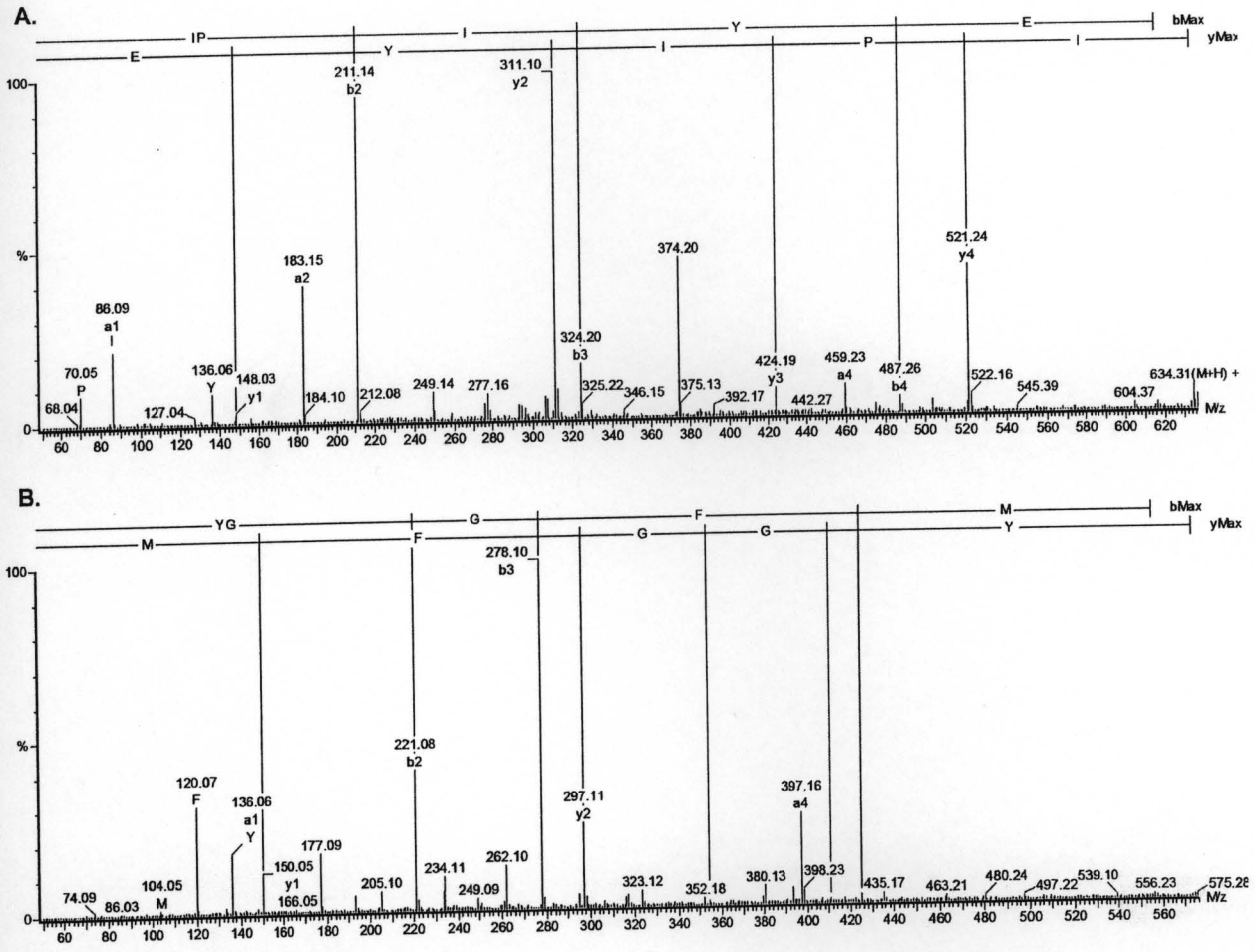

Figure 2.3 MS/MS *de novo* sequencing of (**A**) CART fragment, IPIYE (634.26^{+}) and (**B**) Met-enkephalin, YGGFM (574.34^{+}). Major sequence-specific fragment ions such as b-type and y-type ions are labeled in the fragmentation spectra. The peptide sequences are deduced from the mass differences between consecutive b-type or y-type ions, and are shown on the top of the fragmentation spectra. Several immonium ions indicative of Pro (70), Ile (86), Met (104), Phe (120), Tyr (136), and several a-type ions are labeled in the fragmentation spectra. The identity of Ile was determined based on known CART prohormone sequence.

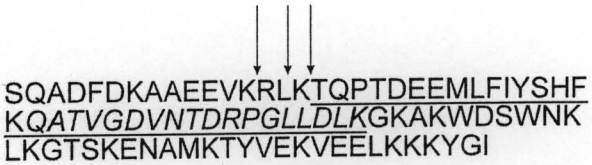

Figure 2.4 Amino acid sequence of Diazepam-binding inhibitor (DBI) protein. Arrows indicate the processing sites observed in this study. Underline indicates sequence of TTN. Italicized portion indicates sequence of ODN.

2.6 References

1. Strand, F. L. *Neuropeptides: regulators of physiological processes*; The MIT Press: Cambridge, MA, 1999.

2. Hokfelt, T.; Broberger, C.; Xu, Z. Q.; Sergeyev, V.; Ubink, R.; Diez, M. *Neuropharmacology* **2000**, *39*, 1337-1356.

3. Desiderio, D. M. *Life Sci* **1992**, *51*, 169-176.

4. Nilsson, C. L.; Karlsson, G.; Bergquist, J.; Westman, A.; Ekman, R. *Peptides* **1998**, *19*, 781-789.

5. Li, L.; Garden, R. W.; Sweedler, J. V. *Trends Biotechnol* **2000**, *18*, 151-160.

6. Kennedy, R. T.; Watson, C. J.; Haskins, W. E.; Powell, D. H.; Strecker, R. E. *Curr Opin Chem Biol* **2002**, *6*, 659-665.

7. Clynen, E.; De Loof, A.; Schoofs, L. *Gen Comp Endocrinol* **2003**, *132*, 1-9.

8. Fricker, L. D.; Lim, J.; Pan, H.; Che, F. Y. *Mass Spectrom Rev* **2006**, *25*, 327-344.

9. Hummon, A. B.; Amare, A.; Sweedler, J. V. *Mass Spectrom Rev* **2006**, *25*, 77-98.

10. Andren, P. E.; Caprioli, R. M. *Brain Research* **1999**, *845*, 123-129.

11. Skold, K.; Svensson, M.; Kaplan, A.; Bjorkesten, L.; Astrom, J.; Andren, P. E. *Proteomics* **2002**, *2*, 447-454.

12. Skold, K.; Svensson, M.; Nilsson, A.; Zhang, X.; Nydahl, K.; Caprioli, R. M.; Svenningsson, P.; Andren, P. E. *J Proteome Res* **2006**, *5*, 262-269.

13. Nachman, R. J.; Russell, W. K.; Coast, G. M.; Russell, D. H.; Predel, R. *Peptides* **2005**, *26*, 2151-2156.

14. Li, L.; Kelley, W. P.; Billimoria, C. P.; Christie, A. E.; Pulver, S. R.; Sweedler, J. V.; Marder, E. *J Neurochem* **2003**, *87*, 642-656.

15. Jimenez, C. R.; Li, K. W.; Dreisewerd, K.; Spijker, S.; Kingston, R.; Bateman, R. H.; Burlingame, A. L.; Smit, A. B.; van Minnen, J.; Geraerts, W. P. *Biochemistry* **1998**, *37*, 2070-2076.

16. Fu, Q.; Li, L. *Anal Chem* **2005**, *77*, 7783-7795.

17. Jimenez, C. R.; Spijker, S.; de Schipper, S.; Lodder, J. C.; Janse, C. K.; Geraerts, W. P.; van Minnen, J.; Syed, N. I.; Burlingame, A. L.; Smit, A. B.; Li, K. *J Neurosci* **2006**, *26*, 518-529.

18. Che, F. Y.; Lim, J.; Pan, H.; Biswas, R.; Fricker, L. D. *Mol Cell Proteomics* **2005**, *4*, 1391-1405.

19. Fricker, L. D.; Berman, Y. L.; Leiter, E. H.; Devi, L. A. *J Biol Chem* **1996**, *271*, 30619-30624.

20. Wei, H.; Nolkrantz, K.; Parkin, M. C.; Chisolm, C. N.; O'Callaghan J, P.; Kennedy, R. T. *Anal Chem* **2006**, *78*, 4342-4351.

21. Parkin, M. C.; Wei, H.; O'Callaghan, J. P.; Kennedy, R. T. *Anal Chem* **2005**, *77*, 6331-6338.

22. Svensson, M.; Skold, K.; Svenningsson, P.; Andren, P. E. *J Proteome Res* **2003**, *2*, 213-219.

23. Theodorsson, E.; Stenfors, C.; Mathe, A. A. *Peptides* **1990**, *11*, 1191-1197.

24. Nylander, I.; Stenfors, C.; Tan-No, K.; Mathe, A. A.; Terenius, L. *Neuropeptides* **1997**, *31*, 357-365.

25. de Quidt, M. E.; Emson, P. C. *Neuroscience* **1986**, *18*, 545-618.

26. Dockray, G. J. *Nature* **1976**, *264*, 568-570.

27. Marley, P. D.; Rehfeld, J. F. *J Neurochem* **1984**, *42*, 1515-1522.

28. Muller, J. E.; Straus, E.; Yalow, R. S. *Proc Natl Acad Sci U S A* **1977**, *74*, 3035-3037.

29. Delahunty, C.; Yates, J. R., 3rd *Methods* **2005**, *35*, 248-255.

30. Neverova, I.; Van Eyk, J. E. *J Chromatogr B Analyt Technol Biomed Life Sci* **2005**, *815*, 51-63.

31. Issaq, H. J.; Chan, K. C.; Janini, G. M.; Conrads, T. P.; Veenstra, T. D. *J Chromatogr B Analyt Technol Biomed Life Sci* **2005**, *817*, 35-47.

32. Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J Proteome Res* **2003**, *2*, 43-50.

33. Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd *Nat Biotechnol* **2001**, *19*, 242-247.

34. Baggerman, G.; Boonen, K.; Verleyen, P.; De Loof, A.; Schoofs, L. *J Mass Spectrom* **2005**, *40*, 250-260.

35. Husson, S. J.; Clynen, E.; Baggerman, G.; De Loof, A.; Schoofs, L. *Biochem Biophys Res Commun* **2005**, *335*, 76-86.

36. Holm, A.; Storbraten, E.; Mihailova, A.; Karaszewski, B.; Lundanes, E.; Greibrokk, T. *Anal Bioanal Chem* **2005**, *382*, 751-759.

37. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Journal of Separation Science* **2005**, *28*, 1694-1703.

38. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Analytical Chemistry* **2005**, *77*, 6426-6434.

39. Kutz, K. K.; Schmidt, J. J.; Li, L. *Anal Chem* **2004**, *76*, 5630-5640.

40. Ohkuma, S.; Katsura, M.; Tsujimura, A. *Life Sci* **2001**, *68*, 1215-1222.

41. Costa, E.; Guidotti, A. *Life Sci* **1991**, *49*, 325-344.

42. Polevoda, B.; Sherman, F. *Genome Biol* **2002**, *3*, reviews0006.

43. Mzhavia, N.; Berman, Y.; Che, F. Y.; Fricker, L. D.; Devi, L. A. *J Biol Chem* **2001**, *276*, 6207-6213.

44. Hengst, U.; Albrecht, H.; Hess, D.; Monard, D. *J Biol Chem* **2001**, *276*, 535-540.

45. Urushitani, M.; Sik, A.; Sakurai, T.; Nukina, N.; Takahashi, R.; Julien, J. P. *Nat Neurosci* **2006**, *9*, 108-118.

Chapter 3: Quantitative Neuropeptidomic Analysis of Fed and Unfed Rats

3.1 Introduction

As mentioned in the Chapter 1, opioids regulate the hedonistic aspects of feeding. The infusion of μ-opioid receptor agonists into the nucleus accumbens (NAc) of rats causes an increase in hedonistic feeding.[1-3] However, all of this work used indirect methods, such as pharmacological manipulation, to study the effects of opioids on feeding without direct examination of the endogenous opioid response to different motivational states, i.e. fed versus satiated. In light of this, Kelley and co-workers measured the preproenkephalin changes within the NAc by Northern blotting and *in situ* hybridization in fed and unfed rats.[4] Interestingly, preproenkephalin expression was significantly down-regulated in the rats which had eaten (the satiated rats) when compared to the food-denied rats, implying causality between endogenous enkephalin expression and feeding. However, while these data equate endogenous opioid expression changes with specific motivational states, using mRNA changes to examine the effects of endogenous opioids is still an indirect method of monitoring spatial and temporal enkephalin release.

In regards to temporal resolution, there is a time lag between mRNA transcription and mature neuropeptide secretion and thus monitoring mRNA expression changes is only indicative of changes in transcriptional activity and not actual neuropeptide release. Consideration of this time lag is important when trying to link a physical event, such as feeding, with a biochemical event, such as opioid release. Furthermore, the monitoring of

mRNA expression for spatial information, i.e. an *in situ* hybridization, is also problematic due to a physical separation between the site of transcription and translation (the cell body) from the point of release (the axon terminal). In light of these limitations, we wanted to use a method which allowed both spatial and temporal resolution.

Traditionally, microdialysis coupled with radioimmunoassay (RIA) has been used to this end, with a number of studies examining various treatments on *in vivo* enkephalin release, including morphine, alcohol, and corticotrophin releasing factor.[5-7] While microdialysis coupled with RIA is an effective method for *in vivo* monitoring, it requires specialized equipment and lacks the ability to monitor expression changes in multiple neuropeptides simultaneously. Instead, we employed a whole tissue extract coupled with LC-MS/MS for a sensitive, multiplexed analysis.

As mentioned in Chapter 2, MS-based quantitative neuropeptidomics can be technically challenging in mammalian systems due largely to post-mortem protease activity. In order to avoid this problem, we employed the same sample preparation technique as used in Chapter 2. In addition, quantifying neuropeptide expression changes with MS-based neuropeptidomics requires special considerations to avoid run-to-run instrument variations and signal suppression. To overcome these difficulties, many labs, including ours, have turned to stable isotope labeling. Stable isotope labeling consists of tagging one sample with a "light" (non-deuterated) tag and the other sample with a "heavy" (deuterated) tag.[8] Accurate relative abundances can then be deduced by analyzing the difference in signal intensity between the "light" and "heavy" tagged neuropeptides. In this manner, the run-to-run instrument variations and the signal suppression effects are removed. However, the systematic bias introduced by differences in initial sample amounts and/or extraction

efficiencies can not be controlled for by isotope tagging. In order to control for this bias, we developed a normalization scheme which will be outlined in the Methods section.

3.2 Experimental Section

3.2.1 Animals

Adult male Sprague-Dawley rats (total $n = 8$, Harlan Sprague Dawley, Indianapolis, IN) weighing 300–400 g were maintained in a temperature- and humidity-controlled animal colony on a 12:12-h light-dark cycle (lights off at 1800). All subjects were naive and were allowed a minimum of a week adaptation followed by 2 days of daily handling before the beginning of the experiment. Subjects had free access to normal laboratory chow (24% protein, 4% fat) and drinking water was available *ad libitum*. On the day of the experiment, at 1700, food (chow) was removed from half of the subjects, while a measured amount of food was given to the other subjects (18 g/cage). The rats were then killed at 1930, 2.5 h after food was given. The time point was chosen based on the previous mRNA experiment performed in the lab[4]. All experimental procedures were in accordance with protocols approved by the University of Wisconsin Institutional Animal Care and Use Committee.

3.2.2 Animal Sacrifice and Cryostat Dissection

Rats were anesthetized with halothane from Sigma (St. Louis, MO) and then sacrificed by decapitation. The brain was then rapidly removed (<90 s) and snap frozen in 2-methylbutane cooled by dry ice. The frozen brain was then sectioned in 300 μm thick slices on a crytostat from Leica (Wetzlar, Germany) with a compartment temperature of −15 °C.

The nucleus accumbens was removed by punching with 3 mm micro-punch and stored in 1.5 mL tubes at –80 °C until extraction (see Figure 3.1).

3.2.3 Extraction with Acidified Methanol

Nucleus accumbens punches (20-30 mg per rat) were removed from –80 °C and immediately extracted into 0.2 μg/μL of ice-cold acidified methanol (90:10:1 / MeOH:Water:Acetic Acid). The samples were then homogenized by hand with a glass-glass Dounce homogenizer. The homogenized sample was then spun at 20,000 *g* for 20 minutes at 4 °C to remove insoluble material. The supernatant was decanted and then dried in a vacuum centrifuge. Extracts were resuspended in 200 μL 50mM Sodium Acetate (pH 6.5) by vortexing and brief sonication. Samples were spun again at 20,000 *g* for 20 minutes at 4 °C and supernatant decanted.

3.2.4 Differential Isotope Labeling

Extracts from unfed rats were labeled with formaldehyde and extracts from fed rats were labeled with *d*-formaldehyde (see Figure 3.2 for experimental outline). To each extract, 3 μL of 2 M $NaCNBH_4$ was added, followed by 6 μL of 4% formaldehyde or 4% *d*-formaldehyde from Sigma. Samples were incubated at 37 °C for 1 hour with occasional vortexing. The labeled extracts from two rats per condition were pooled, yielding two replicates containing the pooled extracts from four paired samples (see Figure 3.2). Each paired extract was then filtered through a 10 kDa molecular weight cut-off tube from Sartorius (Goettingen, Germany). The flow-through was dried in a vacuum centrifuge and then reconstituted in 25 μL of 0.1% formic acid in water.

3.2.5 LC/MS/MS Analysis

A capillary HPLC from Waters (Milford, Massachusetts) system was used to deliver 15 μL of sample to a trap column from LC Packings (PepMap C_{18}, 300 μm x 5 cm; Amsterdam, The Netherlands) via an isocratic flow of mobile phase A: (0.1% formic acid in water) at a rate of 30 μL per minute for 3 minutes. The flow rate was then switched to ~250 nL per minute and the peptides were flushed onto the analytical column from Microtech (C_{18}, 75 μm x 15 cm; Vista, CA) and eluted via a mobile phase B: 5–45% (0.1% formic acid in acetonitrile) gradient over 30 minutes into a quadrupole time-of-flight mass spectrometer from Waters (QTOF Micro).

Data was collected in positive ion mode from *m/z* 400 to 2000, followed by 3 data-dependent MS/MS acquisitions from *m/z* 50 to 2000. The intensity threshold for switching from the survey scan to MS/MS was set at 15 ion counts. The scan time was 0.9 s; inter-scan time, 0.1 s; capillary voltage, 3200 V; and cone voltage, 35 V.

3.2.6 Neuropeptide Identification

Micromass ProteinLynx 2.1 was used to process the LC/MS/MS data. Deisotoping was performed using the *fast* function. The resulting *.pkl* files were searched against the SwissProt rat database using the on-line version of Mascot. The enzyme was specified as *none*. The peptide mass tolerance was set at *200 ppm* and the MS/MS mass tolerance was set at *0.5 Da*. Two separate searches were performed for each sample, one with the static modification for dimethylation (+28.0 Da) and one for deuterated dimethylation (+32.0 Da) at both the N-terminal and lysine side-chain. The variable modifications were set to include

C-terminal amidation, N-terminal acetylation, and methionine oxidation. Peptides of interest which did not yield significant Mascot scores were verified by *de novo* sequencing, using Micromass MassLynx 4.0 PepSeq software (see Figure 3.3). A *de novo* sequence matching at least three consecutive amino acids of the presumed sequence was considered a positive identification.

3.2.7 Neuropeptide Quantitation

Relative expression changes were calculated for each identified peptide pair (formaldehyde versus *d*-formaldehyde) by comparing extracted ion chromatograms (XIC). The *m/z* for each identified neuropeptide and its isotope partner was used to create XICs in the Masslynx 2.1 chromatogram viewer (see Figure 3.3). Each XIC was then integrated to obtain the peak area for each labeled peptide pair. The area values were entered into an Excel spreadsheet and the ratios between XIC areas were calculated. To control for the bias associated with unequal amounts of tissue and differences in extraction efficiencies a normalization scheme was employed. The normalization factor was determined from the average isotope ratio from all identified peptides and then re-applied to each individual ratio to obtain the normalized ratios (see Table 3.4).

3.3 Results and Discussion

As can be seen from Table 3.1, the relative expression changes of 10 neuropeptides were compared between fed and unfed rats. Of the10 neuropeptides compared, only Leu- and Met-enkephalin showed a substantial expression change. Both Leu- and Met-enkephalin were up-regulated approximately 30% in the unfed versus fed rats, mirroring the previously

published mRNA data.[4]

At first glance, an enkephalin increase in the unfed group might seem counterintuitive given that opioid agonists increase feeding behavior. However, in our current experiment, we are assuming the release of endogenous opioids occurs upon feeding and once released, the enkephalins are degraded. There is a large amount of evidence from *in vivo* microdialysis studies supporting this assumption.[5, 9-11] In all of these experiments, released enkephalin is relatively short-lived in the extracellular space and is typically completely degraded within 60 minutes of release. In regards to our experiment, we sacrificed our rats approximately 150 minutes after feeding which, according to all of the microdialysis data, is long after the released enkephalin is degraded. Thus the enkephalin we extract is still contained within the large-dense-core vesicles. If this logistical caveat is taken into account, then the relative changes in enkephalin expression fit within our paradigm—upon eating, the fed rats release enkephalin into extracellular space where it is degraded, while the food-deprived rats retain the enkephalin that we subsequently extract. So while our previous data showed transcriptional changes in relation to feeding, this data shows actual enkephalin release within the NAc is directly related to food consumption. The importance of this subtle difference is two fold: temporal and spatial.

In regards to the temporal link between feeding and enkephalin expression changes, the mRNA data is only reflective of transcriptional changes affected by feeding and not changes in the release of mature neuropeptides. In contrast, the current neuropeptidomic approach directly monitors peptide release in the same time frame as the feeding event, providing a more direct temporal link between feeding and peptide release. However, in considering the temporal aspects of feeding and enkephalin release, one has to consider

diurnal variations in enkephalin release.

Enkephalin release varies with diurnal rhythms and the rewarding and/or feeding effects of opiates can also be influenced by diurnal rhythms.[12, 13] Rats forage for food directly after waking-up (they are nocturnal and wake-up at the beginning of the dark cycle) and μ-opioid receptor membrane trafficking within the NAc increases during this time.[13] These effects are not only important in the interpretation of data but also in experimental design. In the case of our experiment, we carefully picked the beginning of the dark cycle to food deprive the rats—their natural time to eat and presumably the peak time for enkephalin release.

In regards to the spatial resolution, μ-opioid receptor agonists infused directly into the NAc induce feeding, clearly indicating the importance of the NAc opioid system in hedonisitic feeding;[1-3] however, to date, there has been no direct evidence of feeding-induced enkephalin release in the NAc. Our approach, employing a tissue punch for dissection, allows a very specific region of the brain to be removed and extracted and thus the difference in enkephalin signal that we observe is a direct result of release within the NAc and not some other area of the brain (see Figure 3.1). So what does this mean? Once released, Leu- and Met-enkephalin bind the μ-opioid receptor which inhibits GABA release from the NAc. This decrease in GABA release disinhibits the feeding circuit within the ARC (see Figure 1.4), inducing a more robust, hedonistic feeding behavior, mirroring the effects of DAMGO administration into the NAc. The exact mechanism involved in this disinhibition is not fully elucidated but seems to involve both the ventral tegmental area and the lateral hypothalamus (see Figure 1.4).

As can be seen from Table 3.1, out of the ten peptides identified, only Leu- and Met-

enkephalin exhibited consistent changes. Interestingly, a number of the other neuropeptides (or proneuropeptide fragments) have been implicated in feeding, including cocaine-amphetamine regulated transcript (CART), melanin concentrating hormone (MCH), and SAAS.[14] However, none of these peptides exhibited expression changes within our experimental paradigm. However, the majority of these peptides are thought to modulate feeding via hypothalamic circuitry, downstream of the endogenous opioid system within the NAc. In conclusion, we were able to demonstrate a change in enkephalin release with the NAc upon feeding, further emphasizing the importance of this system on hedonistic feeding.

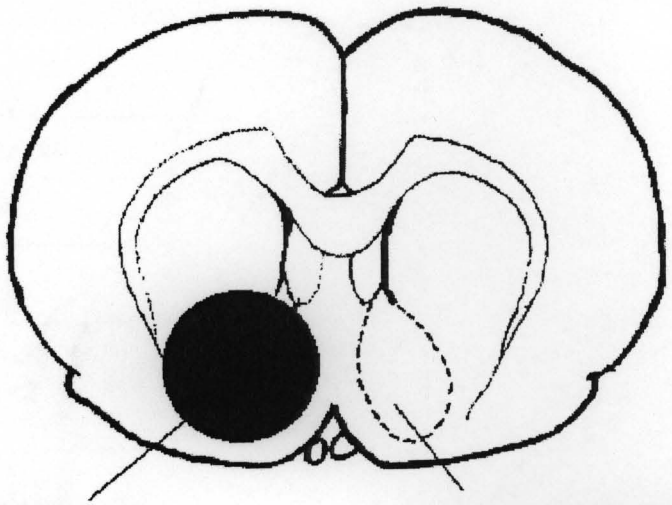

Figure 3.1 Anatomical location of tissue punch used for MS-analysis.

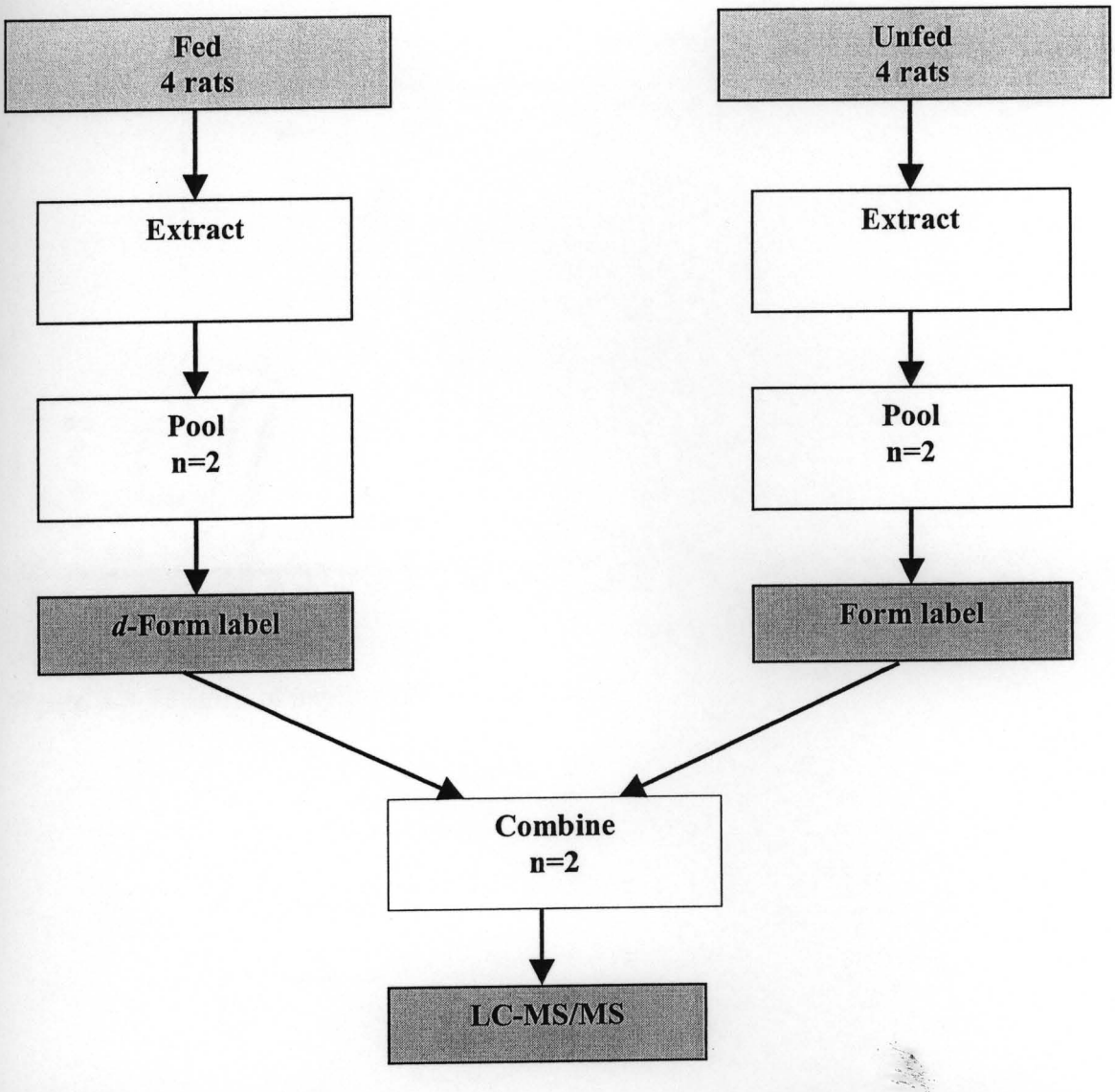

Figure 3.2 Experimental Workflow. 8 total animals, 4 per condition. Punches from 2 animals were pooled after extraction, labeled, and then pooled with their opposite labeled partner for LC-MS/MS analysis.

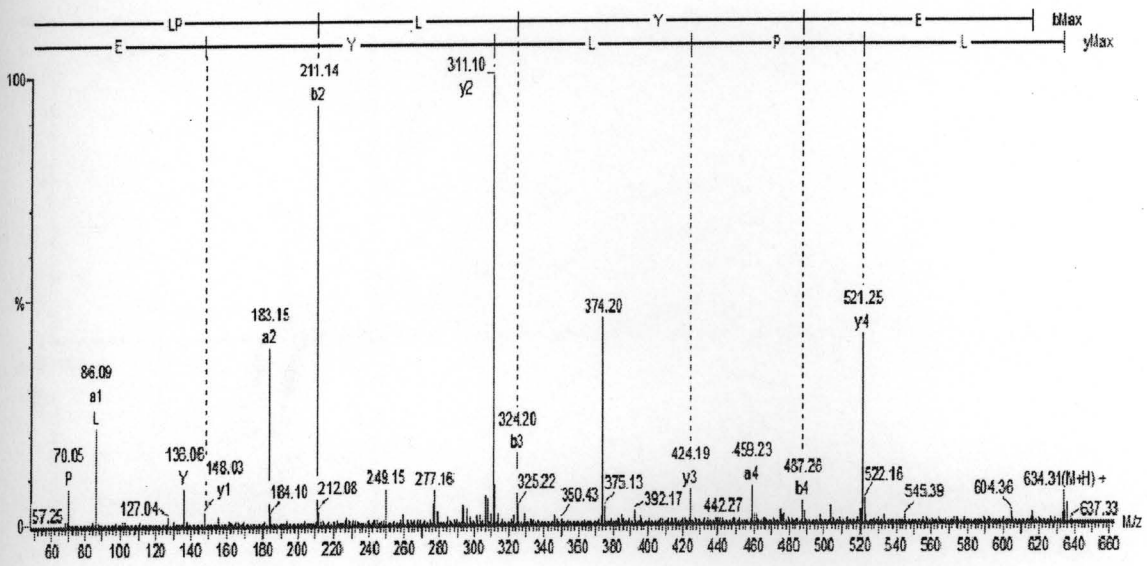

Figure 3.3 MS/MS of peptide IPIYE from CART.

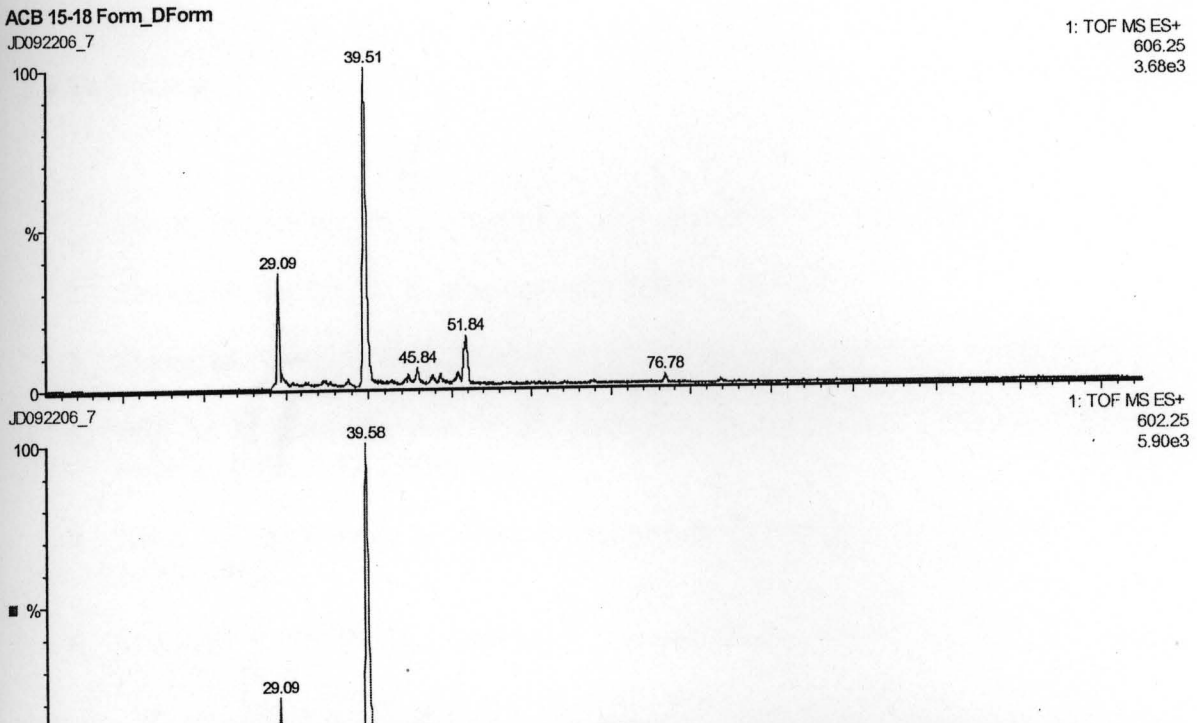

Figure 3.4 XICs for Met-enkephalin labeled with formaldehyde (602.25 *m/z*) and *d*-formaldehyde (606.25 *m/z*).

Propeptide	Sequence	Unfed/Fed 1	Unfed/Fed 2	AVG
Proenkephalin	R.YGGFL.K	1.26	1.32	1.29
Proenkephalin	K.YGGFM.K	1.31	1.35	1.33
CART	R.IPIYE.K	0.84	1.01	0.93
Proenkephalin	R.SPQLEDEAKELQ.K	1.01	1.00	1.01
ProSAAS	A.ASAPLAETSTPLRL.R	0.96	0.75	0.86
ProSAAS	.AASAPLAETSTPLRL.R	0.80	0.88	0.84
Proenkephalin	Y.PVEPEEEANGGEILA.K	1.03	0.84	0.94
ProSAAS	R.SLSAASAPLAETSTPLRL.R	0.91	0.87	0.89
Pro-MCH	R.EIGDEENSAKFPI.G + Amidated (C-term)	0.88	0.98	0.93
Somatostatin	R.SANSNPAMAPRE.R	1.00	0.99	1.00

Table 3.1 Ratio of normalized neuropeptide XICs from fed and unfed rats.

3.4 References

1. Zhang, M.; Kelley, A. E. *Psychopharmacology (Berl)* **1997**, *4*, 350-360.
2. Zhang, M.; Kelley, A. E. *Neuroscience* **2000**, *2*, 267-277.
3. Zhang, M.; Gosnell, B. A.; Kelley, A. E. *J. Pharmacol. Exp. Ther.* **1998**, *2*, 908-914.
4. Will, M. J.; Vanderheyden, W. M.; Kelley, A. E. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2007**, *1*, R217-26.
5. Nieto, M. M.; Wilson, J.; Cupo, A.; Roques, B. P.; Noble, F. *J. Neurosci.* **2002**, *3*, 1034-1041.
6. Marinelli, P. W.; Bai, L.; Quirion, R.; Gianoulakis, C. *Alcohol. Clin. Exp. Res.* **2005**, *10*, 1821-1828.
7. Sirinathsinghji, D. J.; Nikolarakis, K. E.; Herz, A. *Brain Res.* **1989**, *2*, 276-291.
8. Che, F. Y.; Lim, J.; Pan, H.; Biswas, R.; Fricker, L. D. *Mol. Cell. Proteomics* **2005**, *9*, 1391-1405.
9. Maidment, N. T.; Brumbaugh, D. R.; Rudolph, V. D.; Erdelyi, E.; Evans, C. J. *Neuroscience* **1989**, *3*, 549-557.
10. Shen, H.; Lada, M. W.; Kennedy, R. T. *J. Chromatogr. B Biomed. Sci. Appl.* **1997**, *1-2*, 43-52.
11. Emmett, M. R.; Andren, P. E.; Caprioli, R. M. *J. Neurosci. Methods* **1995**, *1-2*, 141-147.
12. Barbano, M. F.; Stinus, L.; Cador, M.; Ahmed, S. H. *Pharmacol. Biochem. Behav.* **2005**, *3*, 569-574.
13. Naber, D.; Wirz-Justice, A.; Kafka, M. S. *Neurosci. Lett.* **1981**, *1*, 45-50.
14. Hokfelt, T.; Broberger, C.; Xu, Z. Q.; Sergeyev, V.; Ubink, R.; Diez, M. *Neuropharmacology* **2000**, *8*, 1337-1356.

Chapter 4: Introduction to Astrocyte Secreted Factors

4.1 Secreted Astrocyte Factors and Synaptic Signaling

4.1.1 History and Background

When astrocytes were first discovered in the mid-nineteenth century, they were called 'glia' (which literally means 'glue' in Greek) and were thought to provide a structural matrix on which neurons to grow. However, as early as the late nineteenth century, Santiago Ramòn y Cajal realized astrocytes might have important physiological properties, including modulation of synaptic signaling and local blood flow[1]. Cajal's speculations would not be validated until nearly a hundred years later with the discovery of astrocyte mediated synaptic transmission and neuronal-vascular coupling.

After Cajal's initial discoveries, astrocytes faded into relative obscurity with the advent of modern electrophysiology. Unlike neurons which exhibit action potentials, astrocytes are not electrically active and thus were thought to be passive by-standers in the lexicon of neurotransmission. A renewed interest in astrocytes began in the 1980s with the discovery of astrocyte-espressed voltage-gated channels and neurotransmitter receptors, raising the possibility of astrocyte participation in neurotransmission. However, it was not until the mid-1990s with the discovery of astrocyte excitability did the importance of astrocytes in synaptic communication become apparent. This renaissance of interest in astrocyte function has resulted in a rapid expansion of information regarding the roles of astrocytes in the central nervous system (CNS), including roles in synaptic signaling, extracellular ionic homeostasis, control of local blood flow, and neuropathological disease.[2]

4.1.2 General Aspects

Astrocytes are characterized by irregular cell bodies with branched processes, very negative membrane potentials, low input resistances, voltage-dependent K^+ currents, gap junctions and glutamate transporters. Within the mammalian brain, astrocytes are arranged in a non-overlapping manner with each astrocyte covering a discrete area. This arrangement allows an individual astrocyte to modulate synaptic transmission and local blood flow to specific areas of the brain.[3, 4] Furthermore, an individual astrocyte displays differential control over individual microdomains within its territory by changing the degree of synaptic coverage. Amazingly, the astrocytic filopodia surrounding individual synapses can extend and retract in a matter of minutes and effectively change the synaptic coverage of a particular synapse.[5, 6] In the hypothalamic supraoptic nucleus, these changes occur in response to lactation. Astrocytes reduce their coverage of synapses which impinge on oxytocin-secreting neurons in lactating animals.[7] These dramatic changes are thought to modulate neuronal excitability through changes in glutamate signaling.

4.1.3 Glutamate Signaling

One of the main methods by which astrocytes affect synaptic signaling is through the modulation of extracellular glutamate.[2] Astrocytes express excitatory amino acid transporters 1 and 2 (EAAT1/2) which transport glutamate out of the extracellular space into the cell, effectively acting like a glutamate sponge.[8] Regulation of extracellular glutamate concentrations is important for the maintenance proper synaptic transmission and neuronal health. Initially, astrocyte modulation of extracellular glutamate was thought to occur strictly

through the actions of the excitatory amino acid transporters; however, this view has been challenged by the discovery of glutamate induced excitatory calcium waves in astrocytes.[9] In hippocampal slices, stimulation of neuronal afferents release glutamate into the synapse which causes Ca^{2+} oscillations in the astrocytic microdomains surrounding the synapse.[9, 10] Interestingly, in experiments using alternative methods to induce Ca^{2+} spikes within astrocytes, the astrocytic Ca^{2+} spikes produce a concurrent increase in Ca^{2+} in surrounding neurons, an event which was later proven to be invoked by glutamate release from the astrocyte.[11-16] Thus there exists bi-directional glutamate signaling from neurons-to-astrocytes and vice versa.[9] In astrocytes, glutamate binds metabotropic glutamate receptors (mGluR) which in turn activate phopholipase C (PLC). PLC activation produces a rise in inositol triphosphate (IP_3) that causes Ca^{2+} release from internal stores.[17] In neurons, synaptic glutamate binds a different set of NMDARs than the astrocytic released glutamate, eliciting distinct effects in the post-synaptic neuron (see Figure 4.1). The discovery of astrocyte modulation of synaptic transmission has marked a new era in our understanding of astrocyte function and has led to the formulation of the tripartite synapse theory—pre- and post-synaptic neurons and the surrounding astrocyte all participate in synaptic transmission.[18]

4.1.4 ATP and D-serine Release from Astrocytes

Although astrocyte modulation of synaptic transmission was first demonstrated with glutamate, recent studies have also shown the importance of D-serine and ATP in modulating synaptic transmission. Interestingly, ATP, and its derivative adenosine, were both found to have the opposite effect of glutamate on synaptic transmission, with glutamate being excitatory and the purines being inhibitory.[19-21] In contrast to the purines, D-serine is a

stimulatory gliotransmitter like glutamate; however, unlike glutamate, it acts at the glycine-binding site of the NMDA receptor.[22] Thus astroctyes can exhibit a bimodal effect, either inhibitory or excitatory, on the same synapse depending on the identity of released gliotransmitter.

4.1.5 Gliotransmitter Exocytosis

The release of neurotransmitters from neurons and neurosecretory cells occurs via the fusion of a neurotransmitter-containing vesicle with the cell membrane, a process termed exocytosis. Although the release of bioactive compounds from astrocytes is well-documented, it is unknown whether these gliotransmitters are released through exocytosis. However, astrocytes express a large variety of essential components for exocytosis, including SNARES and vesicular glutamate transporters (VGLUT), indicating a possible exocytotic mechanism,[19, 23-27] In addition, Ca^{2+}-dependent exocytotic fusion of clear synaptic-like microvesicles (SLMV) has been shown to occur in cultured astrocytes.[28] In comparison to neuronal exocytosis, the astrocytic event seems to occur at a much slower rate and at a lower Ca^{2+} concentration. [28, 29] In addition to SLMVs, larger heterogeneous organelles which label for secretogranin II, a marker for neuronal dense core vesicles, are also present in astrocytes.[25, 30] Thus there is substantial evidence that astrocytes can release gliotransmitters via exocytosis; however, there is one problem with the exocytotic mechanism—the low number of SLMVs in astrocytes in comparison with neurons. If normal exocytosis were occurring, it would soon deplete all of the astrocytic SLMVs and the astrocyte would no longer be able to modulate synaptic signaling. In response to this logistical problem, a kiss-and-run mechanism of release has been proposed in which the vesicle only releases a portion

of its contents at a time.[31] This theory is supported by presence of synaptotagmin IV, instead of synaptotagmin I, which favors a kiss-and-run mechanism in other cell types.[32, 33] Also, the astrocytic SLMVs appear larger than their neuronal counterparts and are assumed to contain more neurotransmitter.[34] If indeed the vesicles are larger and prefer a kiss-and-run mechanism, then the relative paucity of astrocytic vesicles could be sufficient for synaptic signaling. However, there are many unanswered questions and the exact mechanism of gliotransmitter release remains unclear.

4.1.6 Astrocyte Control of Synaptogenesis

Another aspect of synaptic function which is reliant on astrocytes is synapse formation and elimination. Retinal ganglion cells (RGCs) cultured with astrocytes or astrocyte conditioned media (ACM) form substantially more synapses than RGCs cultured alone in serum-free media.[35] Recently, astrocyte-derived thrombospondins (TSPs), which are large extracellular matrix proteins,[36] were found to significantly modulate synapse formation in RGCs.[37, 38] Mice without TSPs were found to have 30% less synapses than wild-type mice. TSP-1 and -2 are only expressed during development and are turned-off in adulthood; however, TSP-4 is highly upregulated in adult human brain in comparison with monkey brain, presenting the possibility of TSP-4 mediated synapse formation in adult humans.[39] While TSPs increase synapse formation *in vitro*, these synapses do not have AMPA receptors and are functionally 'quiet,' indicating the need for other astrocyte secreted factors, possibly cholesterol, to make functional synapses.[40]

Synapse elimination also seems to be under the control of an astrocyte secreted factor. Activity dependent synapse elimination is critical for proper development of the CNS and

seems only to occur in a critical postnatal window. In an attempt to elucidate the mechanisms of synapse elimination in the developing eye, Barres and co-workers examined the gene expression changes in RGCs co-cultured with immature astrocytes.[41] The immature astrocytes induced the expression of the complement cascade protein C1q almost 30-fold in the co-cultured RGCs and experiments with C1q KO mice verify an increase in synapse number. However, the astrocyte secreted factor responsible for the C1q induction remains unidentified.

4.1.7 Astrocyte Control of Neurovascular Coupling

In the late nineteenth century, Ramòn y Cajal proposed a role of astrocytes in controlling cerebral blood flow. Based on the intimate contact between astrocytic endfeet and blood vessels, he formulated the idea that astrocytes could control vasodilatation and vasoconstriction through the movement of their endfeet. In addition, Mosso, a contemporary of Cajal, described activity dependent increases blood flow, in which local blood flow increased in relation to local neuronal activity.[31] This phenomenon was later confirmed by Roy and Sherrington.[42] However, it would be nearly a hundred years before there was a substantial body of scientific evidence to support Cajal's and Mosso's ideas.

Recent results have confirmed the idea that blood flow is directly linked to neuronal activity[43, 44] and that astrocytes are responsible for modulating local blood flow.[45] In particular, Carmignoto and co-workers have shown activity dependent astrocytic Ca^{2+} fluctuations propagate to astrocytic endfeet.[46] Since the endfeet directly contact the endothelial cells surrounding the blood vessels, activity dependent Ca^{2+} fluctuations in the astrocytic endfeet could modulate blood flow. The exact mechanism by which this occurs is

still under investigation but seems to depend on cyclooxgenase (COX) activity, implicating a prostaglandin dependent mechanism.[45] According to the results of multiple experiments, a mechanistic model of astrocyte blood flow modulation has been proposed: Ca^{2+}-encoded signals, which reflect synaptic activity, propagate into the astrocytic endfeet. These Ca^{2+} waves cause a release both vasodilating agents, such as epoxyeicosatrienoic acids (EETs),[47-49] and prostaglandins (PGE_2),[45, 46] as well as vasoconstricing agents, such as 20-hydroxyeicosatetraenoic acid (20-HETE),[50] from the astrocytic endfeet (see Figure 4.2).

4.2 Reactive Astrocytes, Inflammation, and Neurodegenerative Diseases

4.2.1 Background

In addition to regulating synaptic signaling and maintenance, astrocytes play an important role monitoring the health of the surrounding neurons. In this sense, astrocytes are the sentinels of the brain and react to neuronal stress or injury by providing neuronal metabolic and trophic support as well as activating the innate immune response. This astrocytic response to neuronal injury is mediated largely through secreted factors, including neurotrophins, neuropeptides, cytokines, and chemokines. However, these astrocytic responses are not always beneficial and there is much debate about the relative benefits and detriments of the astrocytic response to neuronal stress and injury.

In the event of extreme or chronic neuronal stress, astrocytes undergo a stereotypical response termed 'reactive gliosis' which involves morphological and metabolic changes, including hypertrophy, accumulation of intermediate filaments, and large increases in

metabolism.[51, 52] In addition, these 'reactive' astrocytes express and secrete a large number of cytokines and neurotrophins which are usually not present in 'resting' astrocytes. However, while the astrocytic response is ostensibly to protect neurons from injury and death, the results of astrocyte activation are not always beneficial. This paradoxical nature of astrocyte activation has been the focus of intensive research.

While reactive gliosis is a stereotyped response in the CNS, reactive astrocytes exhibit large differences in their relative reactivities. The degree of reactivity is generally dictated by the extremity of the neuronal insult; physical trauma (e.g. stroke) elicits a strong astrocytic response while less extreme insults, such as neurodegenerative diseases, elicit less of a response. Depending on the degree of reactivity, astroctyic reactivity is classified as either anisomorphic or isomorphic. The anisomorphic state is generally associated with the more deleterious effects of activation, including astroglial scar formation, inhibition of neurite outgrowth and the release of neurotoxic compounds.[53] On the other hand, the isomorphic state has been shown to be neuroprotective and enhance neuronal survival.[53] However, the delineation between the neuroprotective and neurotoxic mechanisms of isomorphic and anisomorphic reactivity remain murky.

Activated astrocytes in the isomorphic stage of reactivity express and secrete a large variety of neurotrophic agents as well as factors involved in the inflammatory response, including cytokines and chemokines.[51, 54, 55] Interestingly, many of these factors are important in embryonic development and are only expressed in adult astrocytes during toxic insult, indicating a recapitulation of neonatal astrocyte functions.[54] In the following paragraphs, the effects of some of these secreted factors will be summarized.

4.2.2 Neurotrophins

Nerve growth factor (NGF) was the first identified neurotrphin and is prominently expressed in astrocytes along with brain-derived neurotrophic factor (BDNF) and neurotrophin-3 (NT-3). All three of these neurotrophins are expressed in resting adult astrocytes and can be upregulated during toxic insult and astrocyte activation.[51, 52] The direct neuroprotective properties of the neurotrophins are well-documented.[56] However, under certain circumstances, such as oxidative stress, NGF becomes an extracellular death signal instead of a neuroprotectant (see section 4.2.7).

4.2.3 TGF-β

Transforming growth factor beta (TGF-β) is expressed mainly by astrocytes and is induced during toxic brain insult.[52] Astrocyte secreted TGF-β exhibits direct neuroprotective effects on serum deprived neurons via the c-Jun/AP-1 pathway.[57] In addition to direct neuroprotective effects, TGF-β is also a cytokine which exhibits anti-inflammatory effects. When TGF-β expression was induced in transgenic mice overexpressing hAPP (a model for Alzheimer's disease), the mice exhibited a lower rate of plaque deposition. However, this lower rate of plaque deposition depends on the presence of activated microglia, indicating a possible mechanism mediated via TGF-β's cytokine activity.[55]

4.2.4 Pro-inflammatory Cytokines

IL-1β, TNF-α, and IL-6 are all considered inflammatory cytokines. All three are induced in reactive gliosis.[51] Interestingly, IL-6 and TNF-α can be either neurotoxic or neuroprotective depending on the context.[53] However, microglia produce both of these

cytokines at a much higher rate than astrocytes and it is possible that these cytokines elicit different responses depending on amounts released and cell-type origin. In addition, IL-1β and IL-6 stimulated astrocytes (IL-1β produces a reactive phenotype) induce the release of NGF and enhance neuronal survival during toxic insult.[53]

4.2.5 Fibroblast Growth Factors

FGF-1 and -2 are both expressed in neonatal and mature astrocytes. Both are induced by a number of noxious insults and are associated with reactive astrocytes and are known to be neuroprotective.[54] However, FGF-2 induces a reactive pheonotype in astrocytes, indicating a possible astrocytic autocrine or paracrine activity which induces neuroprotection via an indirect mechanism.[56]

4.2.6 Erythropoietin

Erthropoietin (EPO) is induced in astrocytes after ischemic insult. It seems to be instrumental in ischemic preconditioning, possibly through the induction of the metallothioneins, which are also released from astrocytes. Neurons express EPO receptors and EPO binding causes activation of the JAK2/PI3K pathway, indicating a direct protective effect.[56].

4.2.7 Neurotrophins and Motor Neuron Apotosis

Astrocyte secreted factors can be either neuroprotective or neurotoxic, depending on the context, including the cell's oxidative state, the concentration and source of released factor, and the identity of expressed receptors. An example of this phenomenon is the effects

of NGF on motor neuron survival. Under normal circumstances, NGF exerts its neuroprotective effects via the tyrosine kinase receptor (TrkA); however, in the proper context NGF becomes an extracellular death signal. Using an *in vitro* cell culture model, Barbeito and co-workers induced motor neuron apoptosis with nitrosylated NGF in a p75 neurotrophin receptor ($p75^{NTR}$) dependent manner.[58] Normally, adult neurons do not express the $p75^{NTR}$; however, when neurons experience oxidative stress, they re-express this receptor which makes them susceptible to nitro-NGF regulated cell death.[59]

This example demonstrates the importance of context in neurotrophic signaling: a compound which was normally neuroprotective was transformed into a neurotoxin by an oxidative modification and the concurrent expression of a receptor. Interpreting the complex interaction of the neurotrophic factors and their effects is a daunting task requiring specialized tools. Ideally, these tools would afford a global view of protein expression changes and post-translational modifications. To this end, a number of groups, including ours, have developed a combination of cell culture and mass spectrometry techniques to identify secreted protein factors.

4.3 Secretome Proteomics

Marin and co-workers were the first to use modern proteomics techniques to identify proteins secreted from cultured mouse astrocytes.[60] Employing a combination of two-dimensional gel electrophoresis (2DE) and matrix-assisted laser desorption ionization (MALDI), they identified a total of 36 putative secreted proteins. Before this study, there had not been a comprehensive examination of proteins secreted from astrocytes and, from a

technical perspective, this was a major advance. The problem with the traditional cell culture techniques was the use of either serum- or protein-containing media. The high concentrations of protein in the media interfere with the downstream proteomics analysis; however, astrocytes will not grow and proliferate without serum-containing media. Marin and co-workers cleverly side stepped this problem by growing their astrocytes in a serum-containing media until confluent, washing their cells with phosphate buffered saline (PBS), incubating the cells in a protein-free media for 18 hours and then collecting the media. In this manner, the collected media was not contaminated by proteins from the media supplement. Another technical hurdle was developing a technique to discriminate physiologically secreted proteins from cytosolic proteins released by dead or dying cells. In addition to assessing cell death in their cell culture by MTS assay, Marin and co-workers employed a whole cell lysate to control for non-secreted protein contamination. By comparing the difference in relative spot intensity between the media and the lysate 2DE gels, they were able to identify media enriched proteins. Working under the assumption that media enrichment indicated physiological secretion, they were able to discern the secreted proteins from the cytosolic proteins from dead or dying cells.

Even though this study was a major technical advance, the relatively small number of proteins identified limits the utility of this technique in the global analysis of the astrocyte secretome. In addition to the small total number of protein identifications, not a single trophic factor or cytokine was among the identified proteins. The low number of protein identifications was due largely to the downstream proteomics analysis—2DE is relatively insensitive and requires a large amount of protein relative to 2D LC-MS/MS shotgun proteomics.[61] Fortunately, since the publication of this manuscript, there have been a number

of advancements in proteomics and bioinformatics, allowing for a more comprehensive and accurate analysis of secreted proteins; however, none of these techniques have been applied to primary astrocyte cultures.

A number of studies have analyzed the secretomes of various cell types, including endothelial,[62] myeloid,[63] adipocytes,[64, 65] and various cancer cell lines.[66-68] Many of these studies have developed novel techniques to address the issues mentioned above. Many studies employed 2D LC-MS/MS shotgun proteomics which generally led to hundreds of total protein identifications.[64, 66-68] Each study employed cell washing techniques to remove serum proteins; however, some groups only used a single 1X wash with PBS[68] while other groups found a 3X wash to be optimal.[62] As for the actual protein precipitation and extraction, two studies found a hydrophobic resin extraction to improve recovery[66, 67] while another study found the optimal precipitation conditions to be trichloroacetic acid (TCA) precipitation which included sodium lauroyl sarcosinate (NLS).[63] Finally, Rabilloud and co-workers found the use of a cytosolic extract was superior to a whole cell extract because it was more reflective of protein contamination from dead and dying cells.[63]

4.4 Research Aims

Our goal is develop effective analytical chemistry techniques to analyze astrocyte secreted proteins. To this end, we plan to optimize 2D LS-MS/MS fractionation methods in order to increase our proteome coverage. After optimization of 2D LC-MS/MS methodologies, we plan to apply the best method to the analysis of astrocyte secreted proteins. In regards to cell culture technique, the optimization of protein-free media conditions, including length of serum-free incubation and cell washing procedures, is

important to minimize cell death; it is also important to monitor the cell viability of individual cultures. In regards to the post-culture procedures, many trophic factors and cytokines are present at low ng/mL concentrations, requiring the rigorous optimization of precipitation/concentration methods as well as downstream analytical and MS methods.[63] To this end, we have undertaken two separate studies: the first (Chapter 5) is the optimization of the general conditions for 2D LC-LC/MS/MS and the second (Chapter 6) is a combination of these optimized analytical techniques, improved primary astrocyte culture procedures, and bioinformatics.

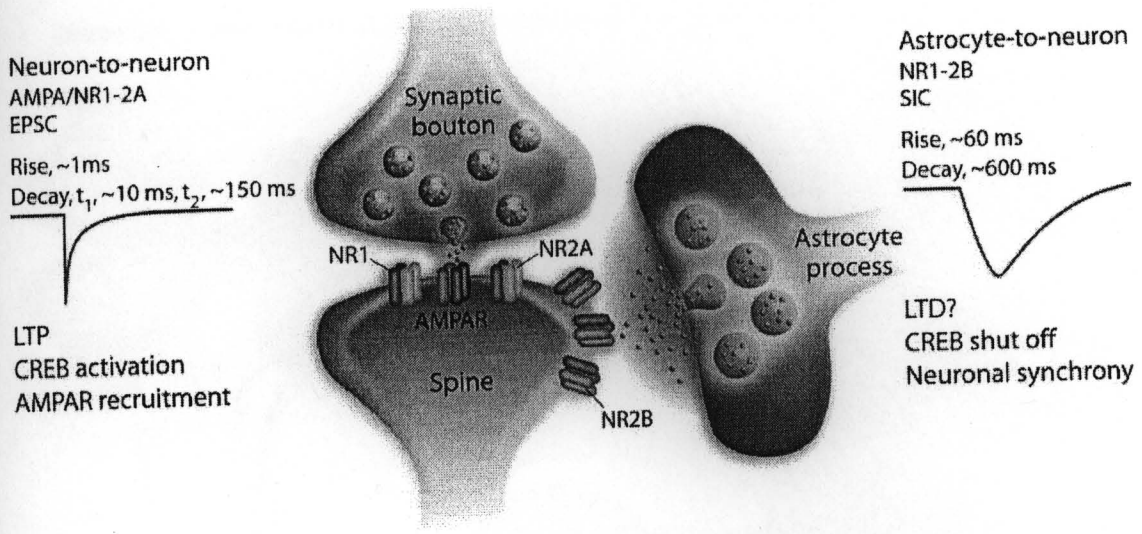

Figure 4.1 Presynaptic glutamate release exhibits different characteristics than extrasynaptic glutamate release. Astrocyte glutamate activates different receptors and exhibits a different release profile.[69]

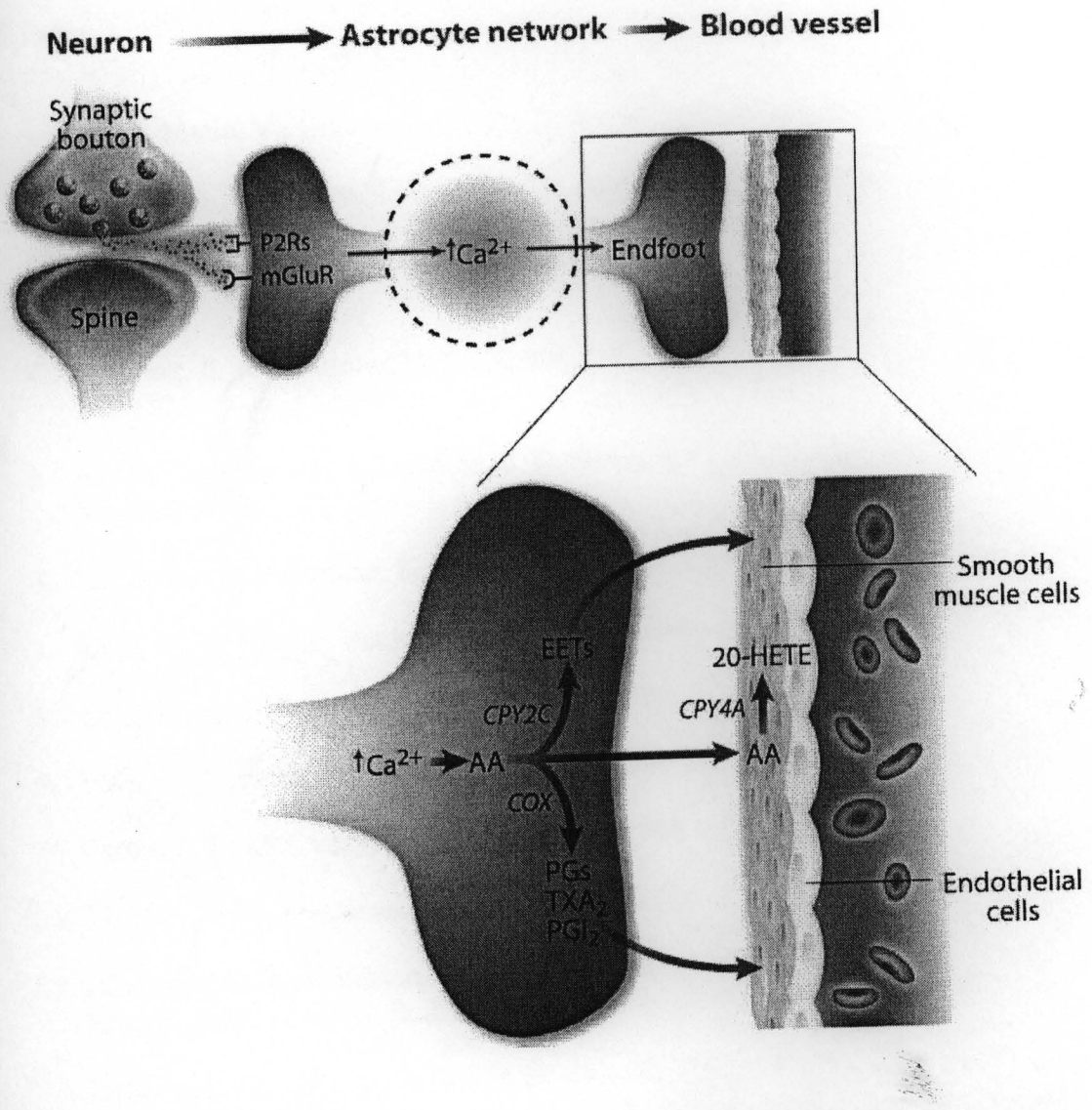

Figure 4.2 Synaptic activity can regulate blood flow via astrocyte interactions with blood capillaries. Synaptic glutamate release causes a rise in astrocytic intercellular which produces a concurrent increase in the COX-2 production of vasomodulatory compounds.[69]

4.5 References

1. Garcia-Marin, V.; Garcia-Lopez, P.; Freire, M. *Trends Neurosci.* **2007**, *9*, 479-487.
2. Volterra, A.; Meldolesi, J. *Nat. Rev. Neurosci.* **2005**, *8*, 626-640.
3. Bushong, E. A.; Martone, M. E.; Ellisman, M. H. *Int. J. Dev. Neurosci.* **2004**, *2*, 73-86.
4. Bushong, E. A.; Martone, M. E.; Jones, Y. Z.; Ellisman, M. H. *J. Neurosci.* **2002**, *1*, 183-192.
5. Benediktsson, A. M.; Schachtele, S. J.; Green, S. H.; Dailey, M. E. *J. Neurosci. Methods* **2005**, *1*, 41-53.
6. Hirrlinger, J.; Hulsmann, S.; Kirchhoff, F. *Eur. J. Neurosci.* **2004**, *8*, 2235-2239.
7. Oliet, S. H.; Piet, R.; Poulain, D. A. *Science* **2001**, *5518*, 923-926.
8. Danbolt, N. C. *Prog. Neurobiol.* **2001**, *1*, 1-105.
9. Pasti, L.; Volterra, A.; Pozzan, T.; Carmignoto, G. *J. Neurosci.* **1997**, *20*, 7817-7830.
10. Porter, J. T.; McCarthy, K. D. *J. Neurosci.* **1996**, *16*, 5073-5081.
11. Charles, A. C. *Dev. Neurosci.* **1994**, *3-4*, 196-206.
12. Hassinger, T. D.; Atkinson, P. B.; Strecker, G. J.; Whalen, L. R.; Dudek, F. E.; Kossel, A. H.; Kater, S. B. *J. Neurobiol.* **1995**, *2*, 159-170.
13. Jeftinija, S. D.; Jeftinija, K. V.; Stefanovic, G.; Liu, F. *J. Neurochem.* **1996**, *2*, 676-684.
14. Nedergaard, M. *Science* **1994**, *5154*, 1768-1771.
15. Parpura, V.; Basarsky, T. A.; Liu, F.; Jeftinija, K.; Jeftinija, S.; Haydon, P. G. *Nature* **1994**, *6483*, 744-747.
16. Parpura, V.; Liu, F.; Brethorst, S.; Jeftinija, K.; Jeftinija, S.; Haydon, P. G. *FEBS Lett.* **1995**, *3*, 266-270.
17. Kawabata, S.; Tsutsumi, R.; Kohara, A.; Yamaguchi, T.; Nakanishi, S.; Okada, M. *Nature* **1996**, *6595*, 89-92.
18. Araque, A.; Parpura, V.; Sanzgiri, R. P.; Haydon, P. G. *Trends Neurosci.* **1999**, *5*, 208-215.

19. Zhang, Q.; Fukuda, M.; Van Bockstaele, E.; Pascual, O.; Haydon, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *25,* 9441-9446.

20. Bowser, D. N.; Khakh, B. S. *J. Neurosci.* **2004**, *39,* 8606-8620.

21. Fiacco, T. A.; McCarthy, K. D. *J. Neurosci.* **2004**, *3,* 722-732.

22. Stevens, E. R.; Esguerra, M.; Kim, P. M.; Newman, E. A.; Snyder, S. H.; Zahs, K. R.; Miller, R. F. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *11,* 6789-6794.

23. Crippa, D.; Schenk, U.; Francolini, M.; Rosa, P.; Verderio, C.; Zonta, M.; Pozzan, T.; Matteoli, M.; Carmignoto, G. *J. Physiol.* **2006**, *Pt 3,* 567-582.

24. Hepp, R.; Perraut, M.; Chasserot-Golaz, S.; Galli, T.; Aunis, D.; Langley, K.; Grant, N. J. *Glia* **1999**, *2,* 181-187.

25. Maienschein, V.; Marxen, M.; Volknandt, W.; Zimmermann, H. *Glia* **1999**, *3,* 233-244.

26. Parpura, V.; Fang, Y.; Basarsky, T.; Jahn, R.; Haydon, P. G. *FEBS Lett.* **1995**, *3,* 489-492.

27. Wilhelm, A.; Volknandt, W.; Langer, D.; Nolte, C.; Kettenmann, H.; Zimmermann, H. *Neurosci. Res.* **2004**, *3,* 249-257.

28. Bezzi, P.; Gundersen, V.; Galbete, J. L.; Seifert, G.; Steinhauser, C.; Pilati, E.; Volterra, A. *Nat. Neurosci.* **2004**, *6,* 613-620.

29. Kreft, M.; Stenovec, M.; Rupnik, M.; Grilc, S.; Krzan, M.; Potokar, M.; Pangrsic, T.; Haydon, P. G.; Zorec, R. *Glia* **2004**, *4,* 437-445.

30. Coco, S.; Calegari, F.; Pravettoni, E.; Pozzi, D.; Taverna, E.; Rosa, P.; Matteoli, M.; Verderio, C. *J. Biol. Chem.* **2003**, *2,* 1354-1362.

31. Haydon, P. G.; Carmignoto, G. *Physiol. Rev.* **2006**, *3,* 1009-1031.

32. Wang, C. T.; Grishanin, R.; Earles, C. A.; Chang, P. Y.; Martin, T. F.; Chapman, E. R.; Jackson, M. B. *Science* **2001**, *5544,* 1111-1115.

33. Zhang, J. M.; Wang, H. K.; Ye, C. Q.; Ge, W.; Chen, Y.; Jiang, Z. L.; Wu, C. P.; Poo, M. M.; Duan, S. *Neuron* **2003**, *5,* 971-982.

34. Kang, N.; Xu, J.; Xu, Q.; Nedergaard, M.; Kang, J. *J. Neurophysiol.* **2005**, *6,* 4121-4130.

35. Pfrieger, F. W.; Barres, B. A. *Science* **1997**, *5332,* 1684-1687.

36. Adams, J. C. *Annu. Rev. Cell Dev. Biol.* **2001**, , 25-51.

37. Ullian, E. M.; Sapperstein, S. K.; Christopherson, K. S.; Barres, B. A. *Science* **2001**, *5504,* 657-661.

38. Christopherson, K. S.; Ullian, E. M.; Stokes, C. C.; Mullowney, C. E.; Hell, J. W.; Agah, A.; Lawler, J.; Mosher, D. F.; Bornstein, P.; Barres, B. A. *Cell* **2005**, *3,* 421-433.

39. Preuss, T. M.; Caceres, M.; Oldham, M. C.; Geschwind, D. H. *Nat. Rev. Genet.* **2004**, *11,* 850-860.

40. Allen, N. J.; Barres, B. A. *Curr. Opin. Neurobiol.* **2005**, *5,* 542-548.

41. Stevens, B.; Allen, N. J.; Vazquez, L. E.; Howell, G. R.; Christopherson, K. S.; Nouri, N.; Micheva, K. D.; Mehalow, A. K.; Huberman, A. D.; Stafford, B.; Sher, A.; Litke, A. M.; Lambris, J. D.; Smith, S. J.; John, S. W.; Barres, B. A. *Cell* **2007**, *6,* 1164-1178.

42. Roy, C. S.; Sherrington, C. S. *J. Physiol.* **1890**, *1-2,* 85-158.17.

43. Attwell, D.; Iadecola, C. *Trends Neurosci.* **2002**, *12,* 621-625.

44. Sandor, P. *Neurochem. Int.* **1999**, *3,* 237-259.

45. Zonta, M.; Sebelin, A.; Gobbo, S.; Fellin, T.; Pozzan, T.; Carmignoto, G. *J. Physiol.* **2003**, *Pt 2,* 407-414.

46. Zonta, M.; Angulo, M. C.; Gobbo, S.; Rosengarten, B.; Hossmann, K. A.; Pozzan, T.; Carmignoto, G. *Nat. Neurosci.* **2003**, *1,* 43-50.

47. Alkayed, N. J.; Narayanan, J.; Gebremedhin, D.; Medhora, M.; Roman, R. J.; Harder, D. R. *Stroke* **1996**, *5,* 971-979.

48. Bezzi, P.; Carmignoto, G.; Pasti, L.; Vesce, S.; Rossi, D.; Rizzini, B. L.; Pozzan, T.; Volterra, A. *Nature* **1998**, *6664,* 281-285.

49. Lovick, T. A.; Brown, L. A.; Key, B. J. *Exp. Physiol.* **2005**, *1,* 131-140.

50. Mulligan, S. J.; MacVicar, B. A. *Nature* **2004**, *7005,* 195-199.

51. Eddleston, M.; Mucke, L. *Neuroscience* **1993**, *1,* 15-36.

52. Ridet, J. L.; Malhotra, S. K.; Privat, A.; Gage, F. H. *Trends Neurosci.* **1997**, *12,* 570-577.

53. Liberto, C. M.; Albrecht, P. J.; Herx, L. M.; Yong, V. W.; Levison, S. W. *J. Neurochem.* **2004**, *5,* 1092-1100.

54. Nakagawa, T.; Schwartz, J. P. *Neurochem. Int.* **2004**, *2-3,* 203-242.

55. Wyss-Coray, T.; Mucke, L. *Neuron* **2002**, *3*, 419-432.

56. Trendelenburg, G.; Dirnagl, U. *Glia* **2005**, *4*, 307-320.

57. Dhandapani, K. M.; Hadman, M.; De Sevilla, L.; Wade, M. F.; Mahesh, V. B.; Brann, D. W. *J. Biol. Chem.* **2003**, *44*, 43329-43339.

58. Pehar, M.; Vargas, M. R.; Robinson, K. M.; Cassina, P.; England, P.; Beckman, J. S.; Alzari, P. M.; Barbeito, L. *Free Radic. Biol. Med.* **2006**, *11*, 1632-1644.

59. Pehar, M.; Vargas, M. R.; Cassina, P.; Barbeito, A. G.; Beckman, J. S.; Barbeito, L. *Neurodegener Dis.* **2005**, *3-4*, 139-146.

60. Lafon-Cazal, M.; Adjali, O.; Galeotti, N.; Poncet, J.; Jouin, P.; Homburger, V.; Bockaert, J.; Marin, P. *J. Biol. Chem.* **2003**, *27*, 24438-24448.

61. Baggerman, G.; Vierstraete, E.; De Loof, A.; Schoofs, L. *Comb. Chem. High Throughput Screen.* **2005**, *8*, 669-677.

62. Pellitteri-Hahn, M. C.; Warren, M. C.; Didier, D. N.; Winkler, E. L.; Mirza, S. P.; Greene, A. S.; Olivier, M. *J. Proteome Res.* **2006**, *10*, 2861-2864.

63. Chevallet, M.; Diemer, H.; Van Dorssealer, A.; Villiers, C.; Rabilloud, T. *Proteomics* **2007**, *11*, 1757-1770.

64. Kratchmarova, I.; Kalume, D. E.; Blagoev, B.; Scherer, P. E.; Podtelejnikov, A. V.; Molina, H.; Bickel, P. E.; Andersen, J. S.; Fernandez, M. M.; Bunkenborg, J.; Roepstorff, P.; Kristiansen, K.; Lodish, H. F.; Mann, M.; Pandey, A. *Mol. Cell. Proteomics* **2002**, *3*, 213-222.

65. Zvonic, S.; Lefevre, M.; Kilroy, G.; Floyd, Z. E.; DeLany, J. P.; Kheterpal, I.; Gravois, A.; Dow, R.; White, A.; Wu, X.; Gimble, J. M. *Mol. Cell. Proteomics* **2007**, *1*, 18-28.

66. Mbeunkui, F.; Fodstad, O.; Pannell, L. K. *J. Proteome Res.* **2006**, *4*, 899-906.

67. Mbeunkui, F.; Metge, B. J.; Shevde, L. A.; Pannell, L. K. *J. Proteome Res.* **2007**, *8*, 2993-3002.

68. Kulasingam, V.; Diamandis, E. P. *Mol. Cell. Proteomics* **2007**, *11*, 1997-2011.

69. Haydon, P. G.; Carmignoto, G. *Physiol. Rev.* **2006**, *3*, 1009-1031.

Chapter 5: Comparison of 2D Fractionation Techniques for Shotgun Proteomics

5.1 Abstract

Two-dimensional (2D) fractionation is a commonly used tool to increase dynamic range and proteome coverage for bottom-up, shotgun proteomics. However, there are few reports comparing the relative separation efficiencies of 2D methodologies using low microgram sample quantities. In order to systematically evaluate 2D separation techniques, we fractionated microgram quantities of *E. coli* protein extract by seven different methods. The first dimension of separation was performed with either reverse phase high pressure liquid chromatography (RP-HPLC), gel electrophoresis (SDS-PAGE), or strong cation exchange (SCX-HPLC). The second dimension consisted of a standard reverse phase capillary HPLC coupled to an electrospray ionization quadrupole time-of-flight mass spectrometer (ESI-QTOF MS) for tandem mass spectrometric analysis. The overall performance and relative fractionation efficiencies of each technique were assessed by comparing the total number of proteins identified by each method. The protein-level RP-HPLC and the high pH RP-HPLC peptide-level separations performed the best, identifying 281 and 266 proteins, respectively. The on-line pH variance SCX and the SDS-PAGE returned modest performances with 178 and 139 proteins identified, respectively. The off-line SCX had the worst performance with 81 proteins identified. We also examined various chromatographic factors which contribute to separation efficiency, including resolving power, orthogonality, and sample loss.

5.2 Introduction

The identification of proteins from complex biological matrices has traditionally been performed using two-dimensional gel electrophoresis (2-DE). 2-DE separates proteins both by their isoelectric point and molecular weight, producing a high resolution protein map from which individual protein spots are picked, excised, and sequenced.[1] While 2-DE remains a powerful tool, shotgun proteomics utilizing LC-MS/MS has emerged as the technique of choice for large scale protein studies due to its superior throughput and sensitivity.[2, 3]

In a typical shotgun proteomics experiment, a complex protein sample is enzymatically digested into peptides which are then separated by HPLC, introduced into a mass spectrometer, fragmented, sequenced and used to identify the parent protein via database searching. Each peptide introduced into the mass spectrometer must be fragmented individually and the rate at which a mass spectrometer can perform this fragmentation is called MS/MS duty cycle. Duty cycle is an important determinant of sampling depth and dynamic range; state-of-the-art, high duty cycle instruments can fragment hundreds of peptides in a single 1D LC-MS/MS experiment. However, a typical biological sample contains thousands of tryptic peptides, resulting in multiple co-eluting peptide species at a given LC elution time which quickly overwhelm the MS/MS acquisition speed of even the highest duty cycle instruments. This duty cycle overload results in a preferential fragmentation of high abundance peptides, introducing an overall sequencing bias towards more highly expressed proteins at the expense of less abundant proteins.[4]

In contrast to LC-MS/MS shotgun proteomics, 2-DE employs a divide-and-conquer strategy by resolving proteins into individual spots which can then be selectively excised and

sequenced. The high resolution of 2-DE allows the researcher to "hand-pick" the proteins of interest, including those that are not highly expressed, while by-passing the more abundant or less interesting proteins. So while shotgun proteomics generates a much larger number of protein identifications via high throughput sequencing, 2-DE exhibits better dynamic range via high resolution fractionation and selective sequencing. In order to couple both the high throughput sequencing of LC-MS/MS and the dynamic range of 2-DE, researchers have coupled multiple fractionation methods in tandem with mass spectrometry, including various HPLC techniques, gel electrophoresis, and capillary electrophoresis (see Washburn and co-workers for an extensive review).[5] However, only a few reports have compared separation efficiencies between different fractionation modes both at the protein and peptide level.[6, 7]

All proteins/peptides exhibit different physicochemical attributes based on their amino acid sequences and by exploiting these differences proteins/peptides can be fractionated. The extent to which the proteins/peptides are fractionated depends on the resolving power of the separation technique and the orthogonality of the two separation modes. For example, in 2-DE proteins are first separated by isoelectric point and then by size. These two separation modes have very different selectivities and each exhibits high resolving power, resulting in a very efficient separation. Given that proteins generally exhibit much different physicochemical characteristics from their constitutive peptides, an obvious way of gaining orthogonality in a 2D shotgun experiment is by first fractionating on the protein-level, digesting the fractions, and then fractionating the resulting peptides in the second dimension. A variety of separation modes have been employed to achieve protein level separation, including size exclusion chromatography (SEC),[8, 9] ion exchange

chromatography (IEC),[2, 10] isoelectric focusing (IEF),[11-15] 1D gel separation (SDS-PAGE),[16, 17] and reverse phase chromatography (RP).[18- 25]

In contrast to a protein-level separation followed by a peptide separation, a two-dimensional peptide-level separation can be performed by employing two or more methods with different separation selectivities. A number of separation modes have been implemented to this end, including strong cation exchange (SCX),[26- 31], isoelectric focusing (IEF),[32, 33] capillary electrophoresis (CE),[34, 35] capillary isoelectric focusing (CIEF),[36, 37] and mixed mode pH reverse phase (RP-RP).[38-41]

The protein level and peptide level separations have relative advantages and disadvantages. Proteins are sensitive to precipitation upon exposure to high salt concentrations, pH extremes, and organic solvents. Peptides, on the other hand, are relatively well-behaved in solution and generally do not exhibit solubility issues. In addition, peptides are amenable to stable isotope labeling while whole proteins are not. However, peptide-level separations also have limitations, including the scattering of tryptic peptides from a single parent protein into multiple fractions which can potentially reduce protein identification scores.

The goal of this work is to systematically compare the factors that impact separation efficiencies, such as peak capacity and orthogonality, as evidenced by the total number of protein identifications of the following separation techniques: 1D SDS-PAGE, offline SCX, online SCX, mixed mode pH RP-RP, and protein RP.

5.3 Experimental Section

5.3.1 Escherichia coli Whole Cell Lysate

50mL of CircleGrow media from MP Biomedicals (Solon, OH) was inoculated with *E. coli* (DH5-alpha) from Invitrogen (Carlsbad, CA) and grown overnight to saturation. The cells were pelleted by centrifugation at 15,000g for 5 minutes and frozen at -80°C. For protein extraction, the cells were thawed, re-suspended in 1mL of cold water, placed on ice, and extracted with five, 10 second bursts at 30% power with an ultrasonic dismembrator from Microson (Farmingdale, NY). Cell debris was removed by centrifugation at 15,000g for 15 minutes at 4°C. The supernatant was decanted and subjected to protein determination by bichoninic acid (BCA).

5.3.2 Protein Solubility

150µL of *E. coli* protein extract was diluted to 0.3, 0.6, 0.9 and 1.2µg/µL in the following buffers: 6M urea with 0.1% and 0.05% trifluoroacetic acid (TFA), 0.1% and 0.05% TFA without urea, and 1% acetic acid without urea. For all conditions containing urea, samples were pre-treated with urea before acidification. After treatment, samples were allowed to incubate for 15 minutes at room temperature and then centrifuged at 15,000g for 15 minutes at 4°C. The Eppendorf tubes were then visually inspected for precipitate. Solubility was judged on degree of precipitation and called either soluble (S, no precipitation), partially soluble (S/I, slight precipitation), or insoluble (I, total precipitation). Results are summarized in Table 5.1.

5.3.3 Solvent-Assisted Protein Digestion

Proteins were enzymatically digested as described.[42] Protein extracts were diluted into 80% acetonitrile/ 20% 200mM ammonium bicarbonate at pH 8. Cysteinyl disulfides were reduced by the addition of 2mM Tris[2-carboxyethyl] phosphine (TCEP) for 30 minutes at 37°C. Reduced disulfides were then alkylated by the addition of 10mM iodoacetamide (IAA) for 30 minutes in the dark. The pH was checked with pH paper and adjusted to pH 8. Trypsin from Promega (Madison, WI) was added at a 1:20 weight-to-weight ratio and incubated for 12 hours at 37°C. After digestion, solvent was removed by vacuum centrifugation.

5.3.4 Off-line Strong Cation Exchange Chromatography (Off-SCX)

100µg of protein digest was reconstituted in 100µL of 25% acetonitrile, 5mM K_2HPO_4 at pH 3 (HPLC Buffer A). A Rainin Dynamax HPLC equipped with a 100µL sample loop, binary pump, UV detector, and fraction collector was used to deliver the reconstituted digest to a PolySulfoethyl A SCX column (2.1 x 250mm, 5µm, 300Å) from PolyLC (Columbia, MD). Peptides were eluted with a linear gradient of 25% ACN, 500mM K_2HPO_4 at pH 3 (HPLC Buffer B) from 5-100% over 60 minutes. Fractions were collected every 5 minutes for a total of 12 fractions. Each fraction was vacuum-centrifuged to dryness and reconstituted in 20µL of 0.1% formic acid (FA); 2µL of each reconstituted fraction was analyzed by LC/MS/MS (see below for details).

5.3.5 Off-line High pH Reverse Phase Chromatography (pH-RP)

100μg of protein digest was reconstituted in 100μL of 200mM NH_4Formate at pH 10 (HPLC Buffer A). A Rainin HPLC (see Off-SCX for description) was used to deliver the reconstituted digest to a Gemini C_{18} RP column (2 x 150mm, 3μm, 110Å) from Phenomenex (Torrance, CA). Peptides were eluted with a linear gradient of 100% acetonitrile (HPLC Buffer B) from 5-35% over 60 minutes. Fractions were collected every 5 minutes for a total of 12 fractions. Each fraction was vacuum-centrifuged to dryness and reconstituted in 20μL of 0.1% FA; 2μL of each reconstituted fraction was analyzed by LC/MS/MS (see below for details).

5.3.6 On-line pH Variance Strong Cation Exchange (On-SCX)

10μg of protein digest was reconstituted in 20μL of citric acid buffer at pH 3 from Column Technology (Fremont, CA). The reconstituted peptides were delivered to a SCX trapping column (0.32 x 100mm, 5μm) from Column Technology via a Waters (Milford, MA) capillary HPLC (see LC-MS/MS for description). Ten discrete buffers from Column Technology were used to elute peptides from the trap column via autosampler injection. The buffers consisted of 10mM citric acid adjusted to pH 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, and 8.0 by ammonium hydroxide (see Dai et al. for details[30]). After each injection, the eluted peptides were delivered directly to a capillary reverse phase C_{18} column and analyzed by MS/MS (see LC-MS/MS section for description).

5.3.7 Off-line Conventional Protein Reverse Phase (cProt)

40μg of undigested protein extract in 100μL of water was delivered to a Macrosphere C_{18} RP column (2.1 x 250mm, 5μm, 300Å) from Alltech (Lexington, KY) using a Rainin HPLC (see Off-SCX for description). To increase recovery and decrease protein adsorption, the column was heated to 60°C throughout the separation. Proteins were eluted with a linear gradient of 0.1% TFA in acetonitrile (HPLC Buffer B) from 15-55% over 60 minutes (HPLC Buffer A was 0.1% TFA in water). Fractions were collected every 5 minutes for a total of 12 fractions. Each fraction was vacuum-centrifuged to ~50μL and then diluted to 80% acetonitrile/20% 200mM ammonium bicarbonate and digested according to the Solvent-Assisted Protein Digestion procedure. The digested fractions were reconstituted in 20μL of 0.1% FA; 5μL of each reconstituted fraction was analyzed by LC/MS/MS.

5.3.8 Off-line High Recovery Protein Reverse Phase (hrProt)

The procedure for cProt was followed exactly, except an Agilent (Santa Clara, CA) macroporous mRP-C_{18} column (2.1 x 75mm, 5μm) heated to 80°C was employed instead of the conventional C18 from Alltech.

5.3.9 Off-line High Recovery Protein Reverse Phase + Urea (hrProt+Urea)

The procedure for hrProt was followed exactly, except urea was added to the protein extract to a concentration of 6M, followed by the addition of TFA to a concentration of 0.1%.

5.3.10 SDS-PAGE (Gel) Separation

40μg of undigested protein extract in 15μL of water was denatured and reduced via the addition of 2.5μL NuPAGE loading buffer and 1μL of NuPAGE reducing agent followed by heating at 70°C for 10 minutes (all NuPAGE products are from Invitrogen). The sample was loaded onto a 12-well pre-cast NuPage Novex 10% Bis-Tris gel. The gel was run at 200V for approximately 45 minutes with NuPAGE MOPS SDS running buffer. The gel was rinsed with water and then stained for 1 hour with Simple Blue SafeStain from Invitrogen (Carlsbad, CA). The gel was then washed with water overnight. After washing, the entire gel lane was excised with a razor blade and then cut into 12 equal size pieces. The individual pieces were then cut into smaller 1mm cubes and placed in an Eppendorf tube and subjected to the In-Gel Digestion protocol (see below). After the in-gel digestion, the peptides were reconstituted in 20μL of 0.1% FA; 5μL of each reconstituted fraction was analyzed by LC/MS/MS.

5.3.11 In-Gel Digestion

Gel pieces were destained twice with 250μL of 50% methanol/ 50% 100mM ammonium bicarbonate and then dehydrated with 200μL of 50% acetonitrile/ 50% 100mM ammonium bicarbonate followed by vacuum centrifugation. Proteins were reduced by the addition of 25mM dithiothretiol (DTT) for 30 minutes at 56°C. After the removal of DTT, 25mM iodoacetamide was added for 30 minutes in the dark. The gel pieces were then dehydrated as before and then rehydrated with 25μL of 20ng/μL of trypsin in 25mM ammonium bicarbonate. After digesting overnight at 37°C, the digest solution was transferred to an Eppendorf tube. The gel bound peptides were extracted with 50μL 0.1%

formic acid (FA) followed by two additional extractions with 70% acetonitrile/0.1% FA. The digest solution and extract buffers were then combined and dried-down by vacuum centrifugation.

5.3.12 LC-MS/MS

Using a capillary HPLC from Waters, tryptic peptides were delivered to a trap column (PepMap C_{18}, 0.3 x 50mm) from LC Packings (Sunnyvale, CA) via an isocratic flow of 0.1% formic acid in water (LC Buffer A) at a rate of 30μL/min for 3 min. The flow rate was then reduced to 250nL/min, and the peptides were flushed onto an in-house packed capillary column (C18, 75μm x 150mm) and eluted via a 5-45% linear gradient of 0.1% formic acid in acetonitrile (LC Buffer B) over 40 minutes into a nanoelectrospray ionization (nESI) quadrupole time-of-flight (QTOF) mass spectrometer (QTOF Micro) from Waters. Data was collected in positive ion mode from *m/z* 400 to 2000, followed by 3 data-dependent MS/MS acquisitions from *m/z* 50 to 2000. The intensity threshold for switching from the survey scan to MS/MS was set at 15 ion counts. The scan time was 0.9 s; inter-scan time, 0.1 s; capillary voltage, 3200 V; and cone voltage, 35 V.

5.3.13 Database Search

Micromass ProteinLynx 2.1 was used to convert the .raw files into .pkl text files for database searching. The .pkl files from individual fractions were combined into one file and searched against the Swiss-Prot database using both Mascot and X!Tandem. The following parameters were used for both search engines: taxonomy was limited to *E. coli*, parent mass tolerance was 800ppm, fragment mass tolerance was 0.8Da, a maximum of two missed

cleavages was allowed, and carbamidomethylation was set as a fixed modification. For the Mascot search results, the significance threshold was set at $P<0.05$ for protein scores (i.e. scores >25 indicate identity or extensive homology). For the X!Tandem search results, protein expectation values (log E) were required to be <-0.9. Total peptides, unique peptides, protein sequence coverage, and the false discovery rates (FDR) were all derived from the X!Tandem search engine.

5.4 Results and Discussion

In order to systematically compare the relative efficiencies of 2D separations for low microgram quantities of protein extract, we employed a consistent protein load in the second dimension (10μg), consistent number of collected fractions (10-12), consistent LC-MS/MS conditions, and consistent database search parameters (see Figure 1 for a workflow overview and Table 1 for a summary of method conditions). The total number of proteins identified from each 2D separation scheme was used as the prime indicator of separation efficiency (see Table 5.2). We also examined the various reasons for the observed disparities between the separation efficiencies, including resolving power, orthogonality, and sample loss, as well as the physicochemical characteristics of the proteins identified by each method.

Table 5.2 summarizes the total proteins identified by Mascot and X!Tandem for each workflow. As can be seen from the table, high pH reverse phase (pH-RP) and high recovery protein reverse phase (hrProt) gave the highest number of proteins identified, followed closely by high recovery protein reverse phase with urea (hrProt+Urea), and conventional protein reverse phase (cProt). The gel electrophoresis (Gel) and on-line strong cation exchange (On-SCX) gave average results while the off-line strong cation exchange (Off-

SCX) gave the worst results.

In addition to total proteins identified, we examined the total number of peptides identified (including redundant identifications), the total number of unique peptides (only non-redundant identifications), the average protein sequence coverage, and the protein E-value (the probability score). The protein-based HPLC separations consistently yielded the highest number of total peptides and the highest protein scores. In contrast, the pH-RP had a relatively low number of total peptides, yielding less than half as many of hrProt (870 versus 1848), but had an equal or greater number of protein identifications compared to the protein-based separations. At first glance, this seems counterintuitive, since logically a higher number of identified peptides should yield a higher number of identified proteins. However, if you compare the average protein sequence coverage between the two methods, the hrProt method exhibits much better sequence coverage than the pH-RP. In addition, the hrProt gives a higher average protein probability score. Both of these values reflect the higher number of peptides identified per protein in the hrProt fractionation compared to the pH-RP method.

Although the hrProt method gives higher average protein scores and sequence coverage, the pH-RP still yields a similar or better number of identified proteins. The reason for this can be seen in the percentage of unique/total peptides given in Table 5.2. The ratio of unique to total peptides is a measure of how often the instrument fragments the same peptide (MS/MS re-sampling). This re-sampling causes a loss in duty cycle (the instrument spends valuable MS/MS time re-fragmenting the same peptides) and results in a lower number of protein identifications. The substantially higher percentage of unique/total given by the pH-RP method is indicative of a low MS/MS re-sampling rate. So while the hrProt method

yields more total peptides than the pH-RP method, the pH-RP method has a higher ratio of unique/total peptides. This is due to the hrProt having a higher rate of re-sampling events which results in a lower number of protein hits per peptide sequenced. But why do certain workflows exhibit more re-sampling events? Re-sampling is partly dependent upon LC-MS/MS instrument parameters, such as dynamic exclusion, however, in this study the LC-MS/MS parameters have been held constant and thus the re-sampling differences are the result of distinct characteristics of the first dimension separations.

In order to explore the nature of these re-sampling events, we examined the resolution of the seven methods by comparing the elution profiles of three 'signature peptides' from three different proteins (see Figure 5.2). These proteins were picked based on their high abundance in all seven methodologies. Extracted ion chromatograms (XICs) were created across all fractions of every workflow. Using the height of the XICs in each fraction, we were able to generate a virtual elution profile of the protein/peptides in the first dimension of separation. From these XIC chromatograms we compared the relative resolution of each method. As can be seen in Figure 5.2, the XICs reveal two separate events—peak broadening and peak multiplicity (elution of the same analyte in multiple fractions). Both of these events contribute to re-sampling by introducing the same peptide back to the LC-MS/MS in multiple fractions. The pH-RP workflow exhibits very high resolving power, as evidenced by the discrete distribution of the signature peptides into singular fractions, and thus a low rate of re-sampling and a higher percentage of unique/total peptides.

While pH-RP yields a high resolution separation, SCX yields a much lower resolution separation. Both the On-SCX and the Off-SCX exhibit fairly extreme peak broadening. In addition, the On-SCX exhibits peak multiplicity which results in a high rate of peptide re-

sampling. It is not known specifically why this occurs and why it appears to be more pronounced in the On-SCX, which is a pH elution method, relative to the Off-SCX, which is a salt elution method. According to Dai et al., the elution of peptides in the pH mode occurs according to the isoelectric point (pI) of the peptide.[30] However, a peptide possessing multiple functional groups with different pKa's could interact with the SCX resin in a more complex manner, resulting in a peak elution profile which is not strictly dictated by isoelectric point. If this were the case, then peptides containing fewer ionizable functional groups would elute in a single peak while more complex peptides with multiple ionizable groups would elute in multiple peaks. However, the interaction between individual peptides and the SCX resin is complex and depends on both the pH and ionic strength of the elution buffer.

In regards to the protein based HPLC workflows, both the peak broadening and peak multiplicity are fairly pronounced. These phenomenon have been documented in the literature and have contributed to the general reluctance of researchers to use protein reverse phase in 2D shotgun proteomics.[21, 43, 44] Many factors seem to be involved in the production of peak broadening and multiplicity, including column temperature, mobile phase composition, and the conformation of the protein itself, i.e. native or denatured. Due to the inherent complexity of their structures, proteins exhibit very unpredictable interactions with the hydrophobic stationary phase in reverse phase HPLC. In contrast to peptide-based reverse phase HPLC separation, these interactions seem to be a non-partitioning type of separation in which proteins do not actually adsorb into the stationary phase but instead 'stick' to the stationary phase due to hydrophobic interactions. As can be imagined, the degree to which a protein's hydrophobic core is exposed dictates how well it 'sticks' to the

stationary phase. One of the most important factors in determining hydrophobic core exposure is the extent to which the protein is denatured. Using protein standards, Cohen et al. demonstrated an elution dependence on protein conformation—proteins in their native state exhibit different elution times and peak broadening compared to the same protein in its denatured state.[44] In order to examine the effects of protein denaturation, we used a combination of 6M urea and 0.1% TFA as a sample buffer to denature the whole protein extract before fractionation (hrProt+Urea). However, we observed precipitation at higher protein concentrations (1μg/μL) and deemed it necessary to systematically examine the effects of different acid concentrations on protein solubility. As can be seen from Table 5.3, the addition of acid to our protein extract causes protein precipitation at all tested concentrations. The addition of 6M urea as a denaturant/solubilizing agent alleviates some of the solubility issues but protein precipitation still occurs at a concentration of 0.6μg/μL. From these experiments, we concluded the maximum concentration we could use with denaturing conditions (6M urea + 0.1% TFA) was 0.4μg/μL. Using a 100μL loop, this concentration yielded a load of 40μg which is 2.5-times lower than that employed for the Off-SCX and pH-RP. In order to adjust for this disparity, we injected 5μL of the digested hrProt and cProt fractions in second dimension instead of the 2μL employed in the Off-SCX and pH-RP workflows. As can be seen from Figure 5.2, the addition of urea and TFA as denaturants did not significantly diminish peak broadening or multiplicity, nor did it yield more protein identifications and in fact, produced a lower number of protein identifications due possibly to protein precipitation. While we examined some aspects of protein denaturation on separation efficiencies, we did not include disulfide bond reduction or sample pre-heating, both of which enhance denaturation, however, Jorgensen and co-workers

report the reduction of disulfide bonds does not mitigate peak broadening or multiplicity.[45]

In addition to resolution, the other factor affecting separation efficiency in 2D experiments is the orthogonality of the separations. As mentioned in the introduction, orthogonality is a measure of the relative difference between the separation modes in the first and second dimensions, or the difference in the selectivities of the two modes.[39] Figure 5.3 exhibits the orthogonality of three workflows by comparing the distribution of peptides in the second dimension RP-HPLC. The chromatograms are base peak intensity (BPI) traces of eluting peptides in three discrete fractions from the beginning, middle, and end of the first dimension separation. As can be seen from all three of the hrProt-Urea BPI chromatograms, the peptides are evenly distributed throughout, indicating good orthogonality. A protein separation followed by digestion and a peptide level separation is inherently orthogonal since the parent proteins exhibit completely different physicochemical properties than their constituent tryptic peptides. All of the protein-based workflows, including the Gel workflow, exhibit good orthogonality (data not shown). The consideration of orthogonality is more important when separating the same species in subsequent fractionations, as with the separation of peptides first by ion exchange and then by reverse phase. Ion exchange and reverse phase selectivities are very different—ion exchange separates by charge while reverse phase separates by hydrophobicity—and thus they exhibit a large degree of orthogonality and a constant elution profile throughout fractions (see Figure 5.3). In contrast, the pH-RP exhibits a staggered elution profile when examined from the early fraction to the late fraction. The first fraction (F1) exhibits an obvious retention time bias towards early elution; the middle fraction exhibits no bias; and the last fraction (F12) exhibits a bias

towards late elution. The pH-RP method employs the same reverse phase stationary phase in the first and second dimensions; the selectivity is changed from the first to second dimensions by simply varying the mobile phase pH. While the pH difference does change the charge on certain functional groups, such as primary amines and carboxylic acids, this change does not constitute a monumental change in hydrophobicity, especially for peptides lacking a large number of these functional groups, and thus the selectivity of the high pH first dimension is not grossly different from the selectivity of the conventional acidic second dimension, resulting in a semi-orthogonal separation.

The last factor affecting protein identifications via 2D shotgun proteomics is sample loss. As mentioned above, the Off-SCX mode exhibited good orthogonality and decent resolution, however, due to extreme sample loss this method resulted in the lowest number of protein identifications, total peptides, and unique peptides. Most of the previous studies employing 2D LC-MS/MS with SCX-RP utilized a large amount of protein (1-2mg).[26-30] The use of such a large amount of proteins makes it difficult to assess sample loss and the relative sensitivity of these 2D methods. In the current study, we did not assess protein loss/recovery directly but instead inferred recoveries based on total proteins and peptides identified. As can be seen from Table 5.2, Off-SCX produced a pronounced amount of sample loss compared to the other workflows. In contrast, the On-SCX, pH-RP, and protein-based separation methods exhibited a fairly low amount of sample loss. The Gel separation exhibited a fair degree of sample loss as a result of the in-gel digestion which is less efficient than in-solution digests.

To examine the proteome coverage of different techniques, we compared the number of unique protein identifications in the On-SCX, hrProt, and pH-RP workflows. As shown in

Figure 5.4, hrProt, and pH-RP yield a large number of unique protein identifications and On-SCX a significantly lower number. Both the hrProt and pH-RP cover the majority of the On-SCX identified proteins but demonstrate a significant independence from each other. Thus the hrProt and pH-RP are highly complementary to each other, each yielding a high number of unique proteins, while the majority of the proteins identified by On-SCX are encompassed by the other two methods. Employing both hrProt and pH-RP yields a greater proteome coverage compared to either workflow taken independently.

In addition to proteome coverage, resolution, and orthogonality, we also examined the physicochemical characteristics of the identified proteins, including protein abundance, hydrophobicity, molecular weight, and isoelectric point. As can be seen in Figure 5, there seems to be no significant bias for any one method for proteins of a specific hydrophobicity or isoelectric point. In addition, no significant bias is observed in protein abundance, except perhaps a small bias exhibited by the pH-RP workflow in determining low copy number proteins (log copy number below 2.5). However, as can be seen in the molecular weight comparison, the hrProt exhibited a fairly pronounced advantage in identifying low molecular weight species (under 20kD) in comparison with the other methods. It is unknown why this occurs but could be important in regards to experimental design for studies specifically targeting low molecular weight proteins.

5.5 Conclusions

In conclusion, all of the techniques examined in the current study are suitable for 2D shotgun proteomics but each has its advantages and limitations depending on specific experimental design, available equipment, expertise, and monetary resources. The simplest and most cost effective method is the gel-based workflow—this method is cheap and can be performed using equipment commonly available in most biology labs. The off-line HPLC-based methods are more costly and require specialized equipment—an HPLC and column. However, the off-line HPLC methods are more effective, flexible, and higher throughput than the gel-based method. Within the off-line methods, the peptide separations are easier than the protein-based methods owing to simple sample processing, i.e. *en masse* digestion versus digestion of individual fractionations. In addition, protein level separations generally preclude the use of stable isotope labeling for quantitative shotgun proteomics. However, protein level separations are superior for spectral counting due to the high number of total peptides identified. Also, protein level separations seem to be superior for smaller proteins and are generally more effective at removing residual abundant proteins after immunodepletion of serum than peptide HPLC.[25] The on-line methods are the most technically challenging of all of the methods and require the most complicated and expensive equipment. However, sample loss is mitigated when comparing the same separation mode, i.e. SCX.

	Type of Separation	1st Dimension Column Type	1st Dimension Mobile Phase A:	1st Dimension Mobile Phase B:	Gradient Percent B:
hrProt	Off-line Protein	High Protein Recovery C18	0.1% TFA H_2O	0.1% TFA ACN	15-55
hrProt + Urea	Off-line Protein	High Protein Recovery C18	0.1% TFA H_2O	0.1% TFA ACN	15-55
cProt	Off-line Protein	Conventional 300Å C18	0.1% TFA H_2O	0.1% TFA ACN	15-55
pH-RP	Off-line Peptide	High pH Stable C18	200mM NH_4Formate pH 10	ACN	5-35
On-SCX	On-line Peptide	Strong Cation Exchange	Ammonium Citrate 10mM	Ammonium Citrate 10mM	pH Buffers
Off-SCX	Off-line Peptide	Strong Cation Exchange	25% ACN, 5mM K_2HPO_4	25% ACN, 500mM K_2HPO_4	5-100
Gel	Protein	NA	NA	NA	NA

Table 5.1 Summary of Separation Methods.

	Total Proteins		Unique Peptides	Total Peptides	% Unique/Total	Avg % Sequence Coverage	Avg Protein Scores (log E)	%FDR
	Mascot	X!Tandem						
pH-RP	266	299	749	870	86.1	13.1	-16.3	1.16
cProt	241	252	731	1036	70.6	19.4	-19.3	1.03
hrProt+Urea	255	287	977	1509	64.7	19.3	-24.9	1.05
hrProt	281	271	1110	1848	60.1	19.8	-27.3	1.07
Offline SCX	81	105	172	232	74.1	8.3	-8.1	1.35
Online SCX	178	199	369	857	43.1	9.4	-9.7	1.17
Gel	139	157	354	476	74.4	9.5	-12.2	1.35

Table 5.2 Comparison of Separation Methods According to the Number of Total Proteins, Unique Peptides, Total Peptides, Average Percent Sequence Coverage, Average Protein Scores (log E), and False Discovery Rate (FDR)

	Protein mg/mL			
Sample buffer	**0.3**	**0.6**	**0.9**	**1.2**
Urea + 0.05% TFA	S	S/I	S/I	I
Urea + 0.1 % TFA	S	S/I	S/I	I
0.05 % TFA	S/I	S/I	I	I
0.1 % TFA	S/I	I	I	I
1% acetic acid	I	I	I	I

Table 5.3 Protein Solubility in Various Sample Buffers: Soluble (S), Partially Soluble (S/I), and Insoluble (I)

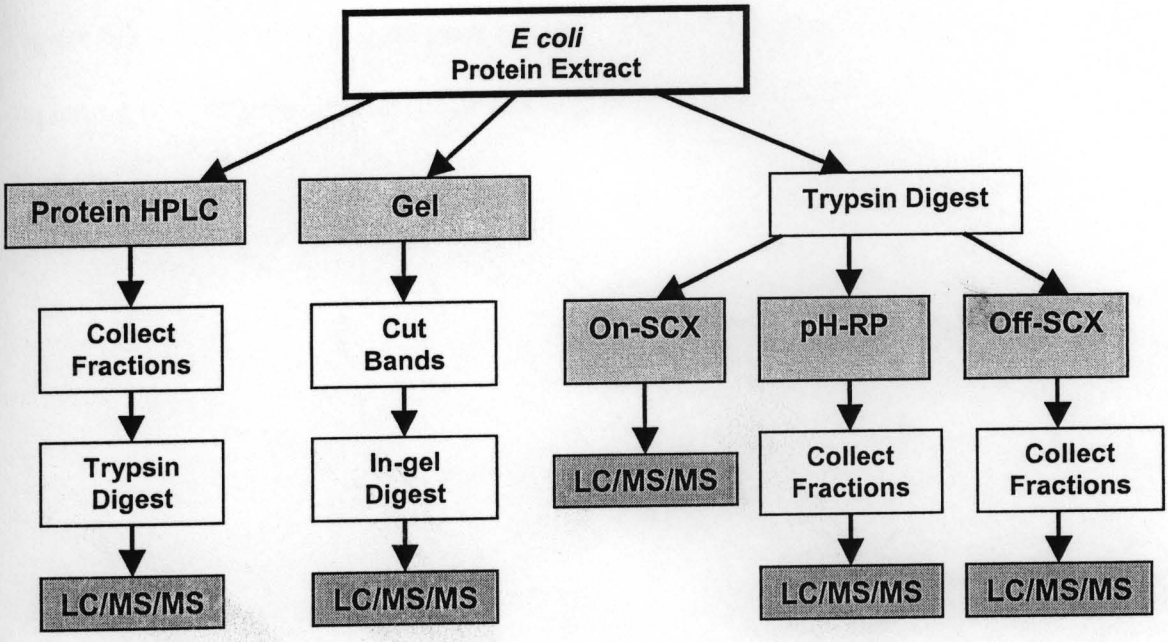

Figure 5.1 Overview of Gel, On-SCX, pH-RP, Off-SCX, and Protein HPLC (includes cProt, hrProt, and hrProt+Urea) workflows.

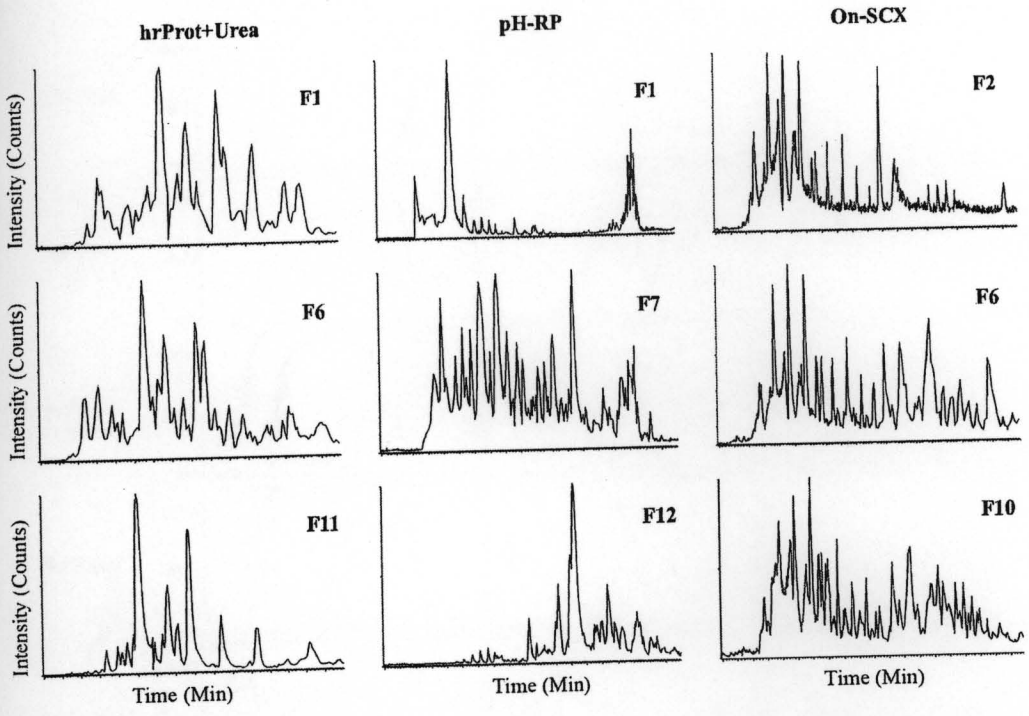

Figure 5.2 Orthogonality: Base peak intensity (BPI) chromatograms for fractions from the beginning (F1, F2), the middle (F6, F7) and the end (F10-F12) of hrProt+Urea, pH-RP, and On-SCX.

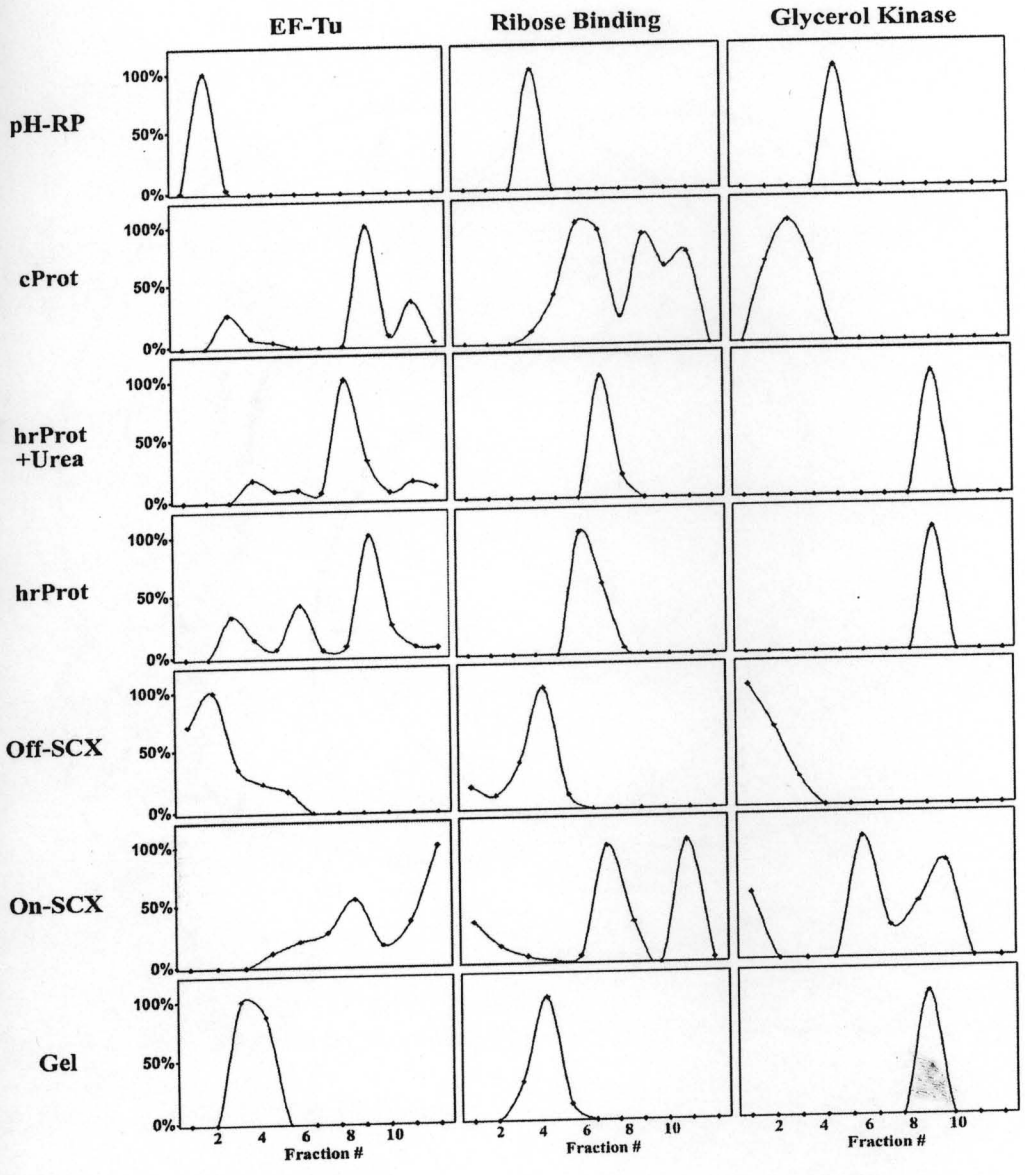

Figure 5.3 Resolution: Elution profiles of three different proteins (Elongation Factor Tu (P02290), Ribose Binding Protein (P02925), and Glycerol Kinase (P08859)) were reconstructed by taking the signal intensity for signature peptides across all collected fractions. The signal intensity is the height of the extracted ion chromatograms (XICs) normalized to 100%.

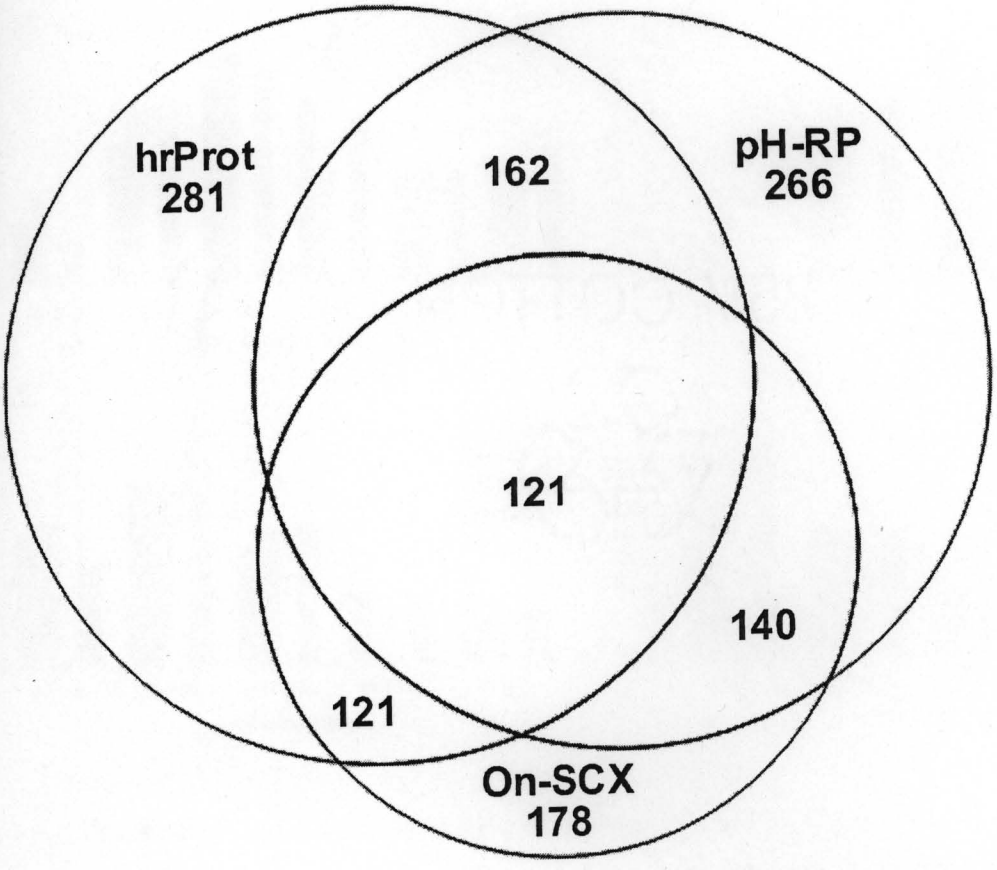

Figure 5.4 Venn diagram representing the overlap of Mascot identified proteins in the pH-RP, hrProt, and On-SCX workflows.

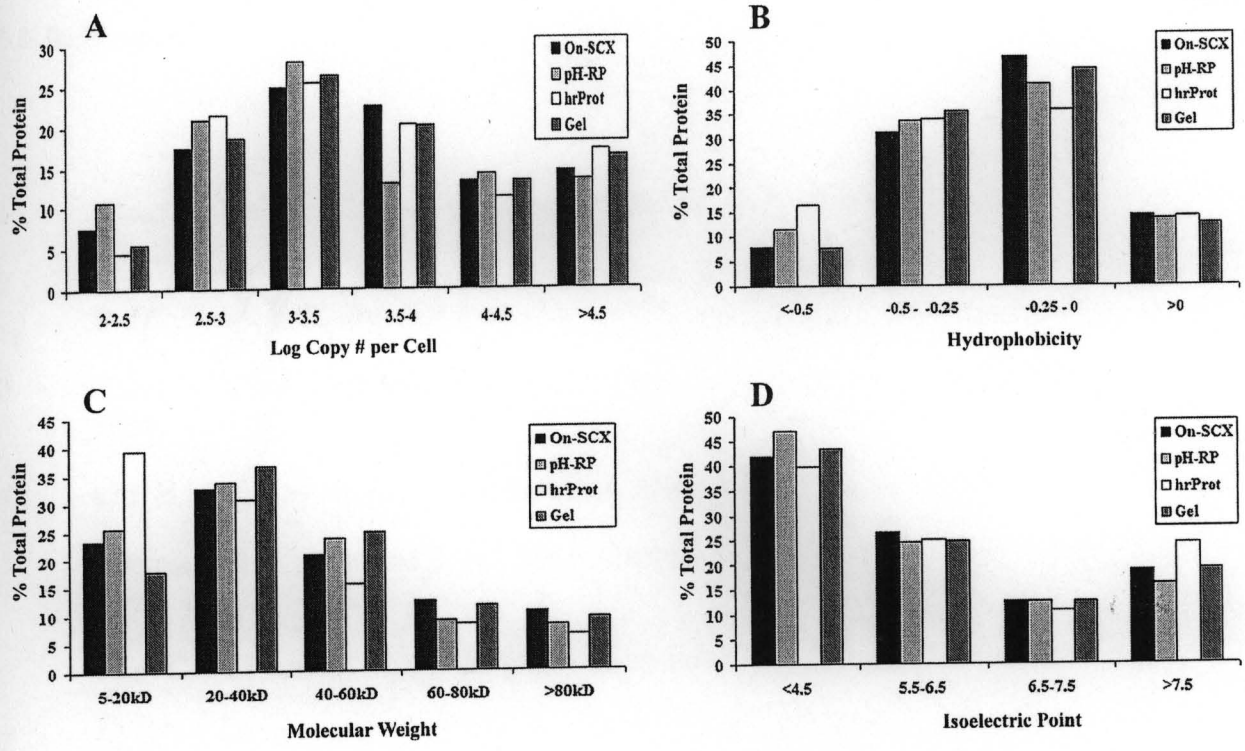

Figure 5.5. Comparison of protein abundance (A), hydrophobicity (B), molecular weight (C), and isoelectric point (D) between On-SCX, pH-RP, hrProt, and Gel workflows. Protein abundance is reported as the log copy number per cell as determined by Frishman and co-workers[46]. Hydrophobicity was determined using the Kyte-Doolittle scale[47]. The molecular weight and isoelectric point were calculated using the ExPASy Server tool: Compute MW/pI.

5.6 References

1 Klose, J.; Kobalz, U. *Electrophoresis.* **1995**, *16*, 1034-1059.

2 Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R.,3rd. *Nat.Biotechnol.* **1999**, *17*, 676-682.

3 Yates, J. R.,3rd; Carmack, E.; Hays, L.; Link, A. J.; Eng, J. K. *Methods Mol.Biol.* **1999**, *112*, 553-569.

4 Liu, H.; Sadygov, R. G.; Yates, J. R.,3rd. *Anal.Chem.* **2004**, *76*, 4193-4201.

5 Fournier, M. L.; Gilmore, J. M.; Martin-Brown, S. A.; Washburn, M. P. *Chem.Rev.* **2007**, *107*, 3654-3686.

6 Gan, C. S.; Reardon, K. F.; Wright, P. C. *Proteomics.* **2005**, *5*, 2468-2478.

7 Li, X.; Gong, Y.; Wang, Y.; Wu, S.; Cai, Y.; He, P.; Lu, Z.; Ying, W.; Zhang, Y.; Jiao, L.; He, H.; Zhang, Z.; He, F.; Zhao, X.; Qian, X. *Proteomics.* **2005**, *5*, 3423-3441.

8 Opiteck, G. J.; Jorgenson, J. W. *Anal.Chem.* **1997**, *69*, 2283-2291.

9 Gao, J.; Opiteck, G. J.; Friedrichs, M. S.; Dongre, A. R.; Hefta, S. A. *J.Proteome Res.* **2003**, *2*, 643-649.

10 Butt, A.; Davison, M. D.; Smith, G. J.; Young, J. A.; Gaskell, S. J.; Oliver, S. G.; Beynon, R. J. *Proteomics.* **2001**, *1*, 42-53.

11 Giorgianni, F.; Desiderio, D. M.; Beranova-Giorgianni, S. *Electrophoresis.* **2003**, *24*, 253-259.

12 Walker, A. K.; Rymar, G.; Andrews, P. C. *Electrophoresis.* **2001**, *22*, 933-945.

13 Wall, D. B.; Kachman, M. T.; Gong, S.; Hinderer, R.; Parus, S.; Misek, D. E.; Hanash, S. M.; Lubman, D. M. *Anal.Chem.* **2000**, *72*, 1099-1111.

14 Kachman, M. T.; Wang, H.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Anal.Chem.* **2002**, *74*, 1779-1791.

15 Yan, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal.Chem.* **2003**, *75*, 2299-2308.

16 Lee, C. L.; Hsiao, H. H.; Lin, C. W.; Wu, S. P.; Huang, S. Y.; Wu, C. Y.; Wang, A. H.; Khoo, K. H. *Proteomics.* **2003**, *3*, 2472-2486.

17 Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. *Proc.Natl.Acad.Sci.U.S.A.* **2004**, *101*, 12130-12135.

18 Evans, C. R.; Jorgenson, J. W. *Anal.Bioanal Chem.* **2004**, *378*, 1952-1961.

19 Zolla, L.; Bianchetti, M. *J.Chromatogr.A.* **2001**, *912*, 269-279.

20 Lee, S. W.; Berger, S. J.; Martinovic, S.; Pasa-Tolic, L.; Anderson, G. A.; Shen, Y.; Zhao, R.; Smith, R. D. *Proc.Natl.Acad.Sci.U.S.A.* **2002**, *99*, 5942-5947.

21 Swanson, R. V.; Glazer, A. N. *Anal.Biochem.* **1990**, *188*, 295-299.

22 Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal.Chem.* **1999**, *71*, 4397-4402.

23 Wang, Y.; Balgley, B. M.; Rudnick, P. A.; Lee, C. S. *J.Chromatogr.A.* **2005**, *1073*, 35-41.

24 Marshall, J.; Jankowski, A.; Furesz, S.; Kireeva, I.; Barker, L.; Dombrovsky, M.; Zhu, W.; Jacks, K.; Ingratta, L.; Bruin, J.; Kristensen, E.; Zhang, R.; Stanton, E.; Takahashi, M.; Jackowski, G. *J.Proteome Res.* **2004**, *3*, 364-382.

25 Martosella, J.; Zolotarjova, N.; Liu, H.; Nicol, G.; Boyes, B. E. *J.Proteome Res.* **2005**, *4*, 1522-1537.

26 Washburn, M. P.; Wolters, D.; Yates, J. R.,3rd. *Nat.Biotechnol.* **2001**, *19*, 242-247.

27 Wolters, D. A.; Washburn, M. P.; Yates, J. R.,3rd. *Anal.Chem.* **2001**, *73*, 5683-5690.

28 Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J.Proteome Res.* **2003**, *2*, 43-50.

29 Vollmer, M.; Horth, P.; Nagele, E. *Anal.Chem.* **2004**, *76*, 5180-5185.

30 Dai, J.; Shieh, C. H.; Sheng, Q. H.; Zhou, H.; Zeng, R. *Anal.Chem.* **2005**, *77*, 5793-5799.

31 Shen, Y.; Jacobs, J. M.; Camp, D. G.,2nd; Fang, R.; Moore, R. J.; Smith, R. D.; Xiao, W.; Davis, R. W.; Tompkins, R. G. *Anal.Chem.* **2004**, *76*, 1134-1144.

32 Cargile, B. J.; Talley, D. L.; Stephenson, J. L.,Jr. *Electrophoresis.* **2004**, *25*, 936-945.

33 Xiao, Z.; Conrads, T. P.; Lucas, D. A.; Janini, G. M.; Schaefer, C. F.; Buetow, K. H.; Issaq, H. J.; Veenstra, T. D. *Electrophoresis.* **2004**, *25*, 128-133.

34 Tong, W.; Link, A.; Eng, J. K.; Yates, J. R.,3rd. *Anal.Chem.* **1999**, *71*, 2270-2278.

35 Figeys, D.; van Oostveen, I.; Ducret, A.; Aebersold, R. *Anal.Chem.* **1996**, *68*, 1822-1828.

36 Chen, J.; Balgley, B. M.; DeVoe, D. L.; Lee, C. S. *Anal.Chem.* **2003**, *75*, 3145-3152.

37 Wang, Y.; Rudnick, P. A.; Evans, E. L.; Li, J.; Zhuang, Z.; Devoe, D. L.; Lee, C. S.; Balgley, B. M. *Anal.Chem.* **2005**, *77*, 6549-6556.

38 Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *J.Sep.Sci.* **2005**, *28*, 1694-1703.

39 Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Anal.Chem.* **2005**, *77*, 6426-6434.

40 Delmotte, N.; Lasaosa, M.; Tholey, A.; Heinzle, E.; Huber, C. G. *J.Proteome Res.* **2007**, *6*, 4363-4373.

41 Nakamura, T.; Kuromitsu, J.; Oda, Y. *J.Proteome Res.* **2008**, *7*, 1007-1011.

42 Strader, M. B.; Tabb, D. L.; Hervey, W. J.; Pan, C.; Hurst, G. B. *Anal.Chem.* **2006**, *78*, 125-134.

43 Cohen, K. A.; Schellenberg, K.; Benedek, K.; Karger, B. L.; Grego, B.; Hearn, M. T. *Anal.Biochem.* **1984**, *140*, 223-235.

44 Cohen, S. A.; Benedek, K.; Tapuhi, Y.; Ford, J. C.; Karger, B. L. *Anal.Biochem.* **1985**, *144*, 275-284.

45 Eschelbach, J. W.; Jorgenson, J. W. *Anal.Chem.* **2006**, *78*, 1697-1706.

46 Ishihama, Y.; Schmidt, T.; Rappsilber, J.; Mann, M.; Hartl, F. U.; Kerner, M. J.; Frishman, D. *BMC Genomics.* **2008**, *9*, 102.

47 Kyte, J.; Doolittle, R. F. *J.Mol.Biol.* **1982**, *157*, 105-132.

Chapter 6: Identification of Astrocyte Secreted Proteins by LC-MS/MS

6.1 Introduction

As mentioned in Chapter 4, astrocyte secreted factors are essential for normal brain function and play important roles in synaptic signaling, synapse maintenance, blood flow control, and neuronal health and survival. The bioactive factors released from astrocytes, including trophic factors, cytokines, chemokines, proteases, protease inhibitors and bioactive small molecules, compose a complex system of extracellular signaling in astrocyte-to-neuron signaling as well as astrocytes-to-astrocyte signaling. As mentioned in Chapter 4, the impact of this signaling is heavily dependent on the physiological context with certain factors exhibiting completely opposing effects under different conditions of cellular stress (see discussion of NGF in Chapter 4 for an example). In order to understand this signal complexity, one must have a global view of expression changes and post-translational modifications and, to date, this is sorely lacking. In addition to the immense complexity of interactions, many of these factors are secreted in very low amounts (low ng/ml), making their direct analysis difficult. In light of the importance of astrocyte secreted factors and the technical challenges associated with their analysis, I have developed a combination of advanced mass spectrometry techniques and bioinformatics to analyze the astrocyte secretome.

To my knowledge there have been only two papers which have applied modern proteomics techniques to the astrocyte secretome; both are from Marin and co-workers.[1, 2]

While representing a major technical advance, both papers resulted in a relatively small number of protein identifications (less than 40). As mentioned in Chapter 4, this small number was due largely to the small amount of starting material (it takes many large cultures to produce enough media for proteomic analysis) and the use of 2DE which is relatively insensitive. In addition to the low number of total proteins identified in these original studies, many of the putative secreted proteins do not possess a signal peptide and/or are not predicted to be secreted in non-classical way. However, in their defense, at the time of publication, the bioinformatics tools that predict protein secretion were not available, making it difficult for Marin and co-workers to discern which proteins were physiologically secreted versus those released by dying or dead cells. To alleviate this problem, I have employed the bioinformatics tools SignalP and SecretomeP, which predict the presence of signal peptides and non-classical secretion, to help identify secreted proteins[3, 4]. In addition to the application of bioinformatics tools, I used quantitative proteomics to compare the relative abundances of proteins in the media versus the lysate control and in doing so was able to discern proteins enriched in the media. This extra information adds significant analytical leverage and security in the identification of secreted proteins.

6.2 Experimental Section

6.2.1 Experimental Overview

Proteins secreted from primary astrocyte cultures were analyzed by LC-MS/MS. A cytosolic extract was used as a control and also analyzed by LC-MS/MS. Three different LC-MS/MS methods were used to analyze the secreted proteins, including a standard 1D LC-MS/MS experiment (3 technical replicates per n; n=2 biological replicates), a 2D LC-MS/MS experiment comprised of a high pH peptide separation (pH-RP) followed by a standard 1D LC-MS/MS (no technical replicates; n=2 biological replicates), and a 2D LC-MS/MS experiment comprised of a protein separation (Prot-RP) followed by a standard 1D LC-MS/MS run (no technical replicates; n=2 biological replicates). The lysate control was only analyzed once with the 2D pH-RP method.

6.2.2 Materials for Cell Culture

The Hank's balanced salt solution (HBSS), Dulbecco's modified Eagle's medium (DMEM), Neurobasal medium (NB), fetal bovine serum (FBS), penicillin/streptomyocin were obtained from Invitrogen (Carlsbad, CA). The sodium bicarbonate, HEPES, trypsin, DNase, and 70-μm cell strainer were obtained from Sigma (St. Louis, MO). 25-cm^2 culture flasks (T25) were obtained from Nunc (Naperville, IL).

The culture medium (DMEM+FBS) consisted of Dulbecco's modified Eagle's medium (DMEM) with glutamine and supplemented with: 2.4 mg/ml sodium bicarbonate, 7.2 mg/ml HEPES, 10% fetal bovine serum (FBS), penicillin/streptomyocin at 100 IU/ml and 100μg/ml, respectively.

6.2.3 Primary Astrocyte Cultures

Primary astrocyte cultures were prepared from postnatal day 1 (P1) mice. The cerebral cortices from individual P1 pups were removed, placed in ice-cold HBSS, minced with a scalpel blade, and then transferred to a tube with 3 ml of 0.25% trypsin at 37 °C. After 25 minutes, 1 ml of culture medium (DMEM+FBS) was added to deactivate the trypsin. DNase was added to a final concentration of 0.05 mg/mL and then the tube was spun at 300 *g* for 5 minutes. The supernatant was removed and the tissue pellet was re-suspended in 1 ml of fresh DMEM+FBS. After triturating 15-20 times with a 1 ml pipette, 4 ml of DMEM+FBS was added to the disassociated cells and then the suspension was filtered through a 70 μm cell strainer. The cells from a single pup were plated onto an uncoated, plastic T25 flask. The DMEM+FBS was changed after the first day and then every 3 days thereafter.

6.2.4 Preparation of Astrocyte Conditioned Media

After the astrocytes reached confluency (Day 7-8), the flasks were placed on a rotary shaker (200 rpm) for 24 hours. After shaking, the DMEM+FBS was removed and discarded (shaking removes the majority of microglia and neurons). The astrocytes were then washed twice with DMEM without FBS and then 5 ml of NB medium was added. After 24 hours, the astrocyte conditioned medium was removed with a pipette and centrifuged at 3,000 *g* for 15 minutes to remove cells and debris. The conditioned media was stored at –80 °C until processed. After thawing, the conditioned media was concentrated 20-40X with a 2 kD

MWCO tube (15mL) at 3,000 *g* in a refrigerated centrifuge (4 °C). Total protein concentration was determined by Bradford assay.

6.2.5 Preparation of Cytosolic Protein Extract

After removal of the astrocyte conditioned media, the confluent astrocyte layer was washed 1X with phosphate buffered saline (PBS). The cells were then scraped from the flask in 1 mL of PBS and pipetted into a Dounce glass-glass homogenizer. After 10 strokes with the Dounce homogenizer, the cell extract was spun at 20,000 *g* for 30 minutes in a refrigerated centrifuge (4 °C). The supernatant (the cytosolic protein extract) was decanted and stored at –80 °C until processed.

6.2.6 Protein Digestion

Protein samples or Prot-RP fractions (see below) were diluted into 6 M urea/200 mM ammonium bicarbonate (pH 8). Cysteinyl disulfides were reduced the addition of 2 mM Tris[2-carboxyethyl] phosphine (TCEP) for 30 minutes at 37 °C. Reduced disulfides were then alkylated by the addition of 10 mM iodoacetamide (IAA) for 30 minutes in the dark. The urea was then diluted to <1 M with 50 mM ammonium bicarbonate and the pH was checked and adjusted to pH 7.8. Trypsin from Promega (Madison, WI) was added at a 1:20 weight-to-weight ratio and incubated for 18 hours at 37°C. After digestion, the solvent was removed by vacuum centrifugation.

6.2.7 Off-line High pH Reverse Phase Chromatography (pH-RP)

Approximately 20 μg of protein digest was reconstituted in 100 μL of 200 mM ammoniun formate at pH 10 (HPLC Buffer A). A Rainin Dynamax HPLC equipped with a 100 μL sample loop, binary pump, UV detector, and fraction collector was used to deliver the reconstituted digest to a Gemini C_{18} RP column (2 x 150 mm, 3 μm, 110 Å) from Phenomenex (Torrance, CA). Peptides were eluted with a linear gradient of 100% acetonitrile (HPLC Buffer B) from 5-35% over 60 minutes. Fractions were collected every 5 minutes for a total of 12 fractions. Each fraction was vacuum-centrifuged to dryness and reconstituted in 20μL of 0.1% FA; 10μL of each reconstituted fraction was analyzed by LC/MS/MS (see below for details).

6.2.8 Off-line Protein Reverse Phase (Prot-RP)

40 μg of undigested protein extract in 100 μL of water was delivered to a macroporous mRP-C_{18} column (2.1 x 75mm, 5μm) from Agilent (Santa Clara, CA) Rainin HPLC (see *pH-RP* for description). To increase recovery and decrease protein adsorption, the column was heated to 80 °C throughout the separation. Proteins were eluted with a linear gradient of 0.1% TFA in acetonitrile (HPLC Buffer B) from 15-55% over 60 minutes (HPLC Buffer A was 0.1% TFA in water). Fractions were collected every 5 minutes for a total of 12 fractions. Each fraction was vacuum-centrifuged to ~50 μL and then diluted to 80% acetonitrile/20% 200 mM NH_4CO_3 and digested according to the *Protein Digestion* procedure. The digested fractions were reconstituted in 20μL of 0.1% FA; 10 μL of each reconstituted fraction was analyzed by LC/MS/MS.

6.2.9 LC-MS/MS

Using a nanoflow HPLC from Eskigent, tryptic peptides were delivered to a trap column (PepMap C_{18}, 0.3 x 50 mm) from LC Packings (Sunnyvale, CA) via an isocratic flow of 0.1% formic acid in water (LC Buffer A) at a rate of 10 μL/min for 3 min. The flow rate was then reduced to 250 nL/min, and the peptides were flushed onto an in-house packed capillary column (C_{18}, 75 μm x 150 mm) and eluted via a 5-35% linear gradient of 0.1% formic acid in acetonitrile (LC Buffer B) over 100 minutes into a nanoelectrospray ionization (nESI) linear ion trap mass spectrometer (LTQ) from Thermo. The MS survey scan was performed in positive ion mode from *m/z* 400 to 2000, followed by data-dependent MS/MS acquisition from *m/z* 100 to 2000. The intensity threshold for switching from the survey scan to MS/MS was set at 15 ion counts. The scan time was 0.9 s; inter-scan time, 0.1 s; capillary voltage, 3200 V; and cone voltage, 35 V.

6.2.10 Database Search

Thermo Xcaliber was used to convert the .RAW files generated during the LC-MS/MS runs into .dta text files for database searching. The .dta files from each sample were then combined into a single file using the NCBI's OMSSA search engine. Using OMSSA, the combined .dta files were searched against the Swiss-Prot database, with the following parameters: taxonomy was limited to *mus musculus*, parent mass tolerance was 2.0 Da, fragment mass tolerance was 1.0 Da, a maximum of two missed cleavages was allowed, carbamidomethylation was set as a fixed modification, and the protein expectation value (log E) was required to be <-0.9. After database searching, proteins were filtered by *Top Hit per Spectrum Only* in OMSSA browser.

6.2.11 Quantitative Analysis by Spectral Counting

From the OMSSA browser, the database search results, including the spectral counts for each protein, were exported to Microsoft Excel (see Figure 6.1 for an overview). The proteins from each run, including the lysate control, were all combined into a master list of proteins. This master list was then re-combined and sorted with each individual run to generate a master list for each run. Each master list contained an identical list of proteins; the proteins which were not originally identified in a run were included as 'zero' counts in the master list (Note: all zero counts were converted to 0.1 to allow fold change calculations—a zero can not be a denominator). In addition, to correct for inconsistencies in sample preparation and run-to-run instrument variation, a total intensity normalization scheme was employed in which each media master list was normalized against the lysate control by total spectral counts (the sum of spectral counts for all the proteins within that list). By generating a master list across all runs and normalizing on total spectral counts, a direct comparison of protein expression across all runs was achieved.

6.2.12 Fold Enrichment and Rank Analysis

After normalization, the spectral counts for each protein in a media run were compared against the spectral counts in the lysate control. From this comparison, a fold enrichment was determined which reflects the relative enrichment of a protein in the media versus the lysate. In order to make the analysis more practical, a $\log_2$ transform was performed. Using the fold enrichment across media runs, two different analyses were performed: average fold enrichment (AFE) and rank analysis (RA). The average fold

enrichment analysis (AFE) takes the average enrichment across all media runs and applies a cut-off, in this case a 4 fold average enrichment (log_2=2), to which all proteins below this value are assumed to be contaminating cytosolic proteins from dead and dying cells in the culture. The rank analysis also uses a 4 fold cut-off value but instead of averaging the fold changes across runs the rank analysis ignores the actual fold change and simply assigns a '1' or a '0' according to a protein's enrichment being above or below the 4 fold cut-off (1=above the cut-off, 0=below the cut-off). The ranks (1 or 0) for the individual runs are then summed to give a total rank analysis score for each protein.

Note: The 4 fold cut-off value for the AFE and RA analyses was determined empirically based on what the author perceived as a reasonable degree of enrichment and by examining the identity and subcellular location (i.e. extracellular) of the identified proteins.

6.2.13 Gene Ontology, SignalP, and SecretomeP

Utilizing the available bioinformatics tools, including gene ontology searches, SignalP, and SecretomeP, was important in validating many of the putative 'hits' obtained from the quantitative analyses. For each protein in the master list, a gene ontology search was performed with emphasis on subcellular location. If a protein was reported as secreted in the literature, it was marked 'yes' in Table 6.1. If, however, the protein yielded a high enrichment factor, as revealed by the fold enrichment or rank analysis, but was not reported to be a secreted protein, then SignalP and SecretomeP were employed to determine if the protein could either be secreted via the classical pathway (protein must possess a signal peptide (SignalP)) or a non-classical route (no signal peptide (SecretomeP)). If the protein

was not reported to be secreted but had a signal peptide it was reported as so in Table 6.1. If the protein was not reported to be secreted and had no signal peptide then SecretomeP returns a score (a score above 0.5 indicates a high probability of being secreted via a non-classical route) which is reported in Table 6.1.

6.3 Results and Discussion

6.3.1 Results and Scoring

Using the secretion prediction software SignalP and SecretomeP, 172 astrocyte secreted proteins were identified, including growth factors, proteases, protease inhibitors, detoxifying enzymes, transport proteins, transcription factors, and even RNA binding and processing enzymes (see Table 6.1). However, many of these proteins do not meet the 4 fold enrichment cut-off for either the average fold enrichment (AFE) or rank analyses (RA). If the 4 fold AFE cut-off is applied to this data, only 108 proteins meet the criteria of a SecretomeP score >0.5 and a 4 fold average enrichment (see Table 6.2). The RA yields similar results, culling the identified proteins from 172 to 127 (RA score of 2 and a SecretomeP score >0.5 required). If all three filters are applied, then only 108 proteins meet the combined criteria (the same number as the AFE/SecretomeP combined filtering). Furthermore, only applying the AFE and RA criteria yield a greater number of proteins than with the co-application of SecretromeP, with the AFE yielding 117 total proteins (see Table 6.2) and the RA criteria yielding 150 total proteins (data not shown). From these data, it is apparent that the AFE analysis is the most stringent of the two analyses and that the application of AFE and SignalP/SecretomeP yields a high quality list of secreted proteins

(see Table 6.2). However, it is also important to consider the proteins in Table 6.2 which have high AFE scores but are not predicted to be secreted; these proteins could be secreted but in a different manner than the known classical and non-classical pathways.

As can be seen from Tables 6.1 and 6.2, the majority of proteins identified as secreted by both the 4 fold cut-off in the AFE and RA criteria are either known to be specifically secreted from astrocytes or from other cell types. As mentioned in the introduction, only two other papers have employed modern proteomics to examine astrocyte secreted proteins, resulting in about 40 identified proteins[1,2]. Out of these 40 proteins 38 were identified in the current study; however, out of these 38 proteins only 14 passed the criteria for enrichment and predictive secretion. Furthermore, these 14 proteins were all brefeldin A sensitive (brefeldin A selectively inhibits secretory vesicle assembly) in Marin and co-workers original paper, indicating a vesicular pathway of secretion[1]. The overlap of the brefeldin A data with the data here serves as a validation for both sets of data. In the interest of time and space only a few identified proteins will be profiled, including some which are known to be secreted from astrocytes as well as some which are novel.

6.3.2 Complement System

The complement system is part of the innate immune response and is important in modulating the inflammatory response. Activated astrocytes secreted many components of the complement system, inducing local inflammation and possibly contributing to neurodegenerative diseases.[5] However, more recent evidence indicates a possible neuroprotective role of certain members of the complement system, such as C1q, and its induction in the initial stages of neurotrauma could be beneficial.[6] Also, as mentioned in the

introduction, neuronal C1q is necessary for synaptic pruning, a process mediated by a C1q induced response in astrocytes.[7] Both complement C3 and C4 were found to be media enriched in the current study.

6.3.3 Insulin-like Growth Factor Binding Proteins (IGFBP)

IGFBPs are known to be secreted from astrocytes and regulate the physiological availability of insulin-like growth factors (IGF-1/-2), neurotrphic factors which promote neuron survival. In addition to regulating the availability of IGF, some of the IGFBPs exhibit IGF-independent bioactivity and can localize to the nucleus.[8] In this study, neither IGF-1 nor -2 were detected; however, IGFBP-2,-3,-4,-5, and -7 were all detected. This bias could be due to the smaller size of the IGFs in comparison with the IGFBPs; mass spectrometry-based proteomics exhibits a sequencing bias for larger proteins. However, it could also be due to differences in secretion or extracellular degradation rates. This trend was also observed for TGF-β whose binding protein, latent-transforming growth factor-beta binding protein (LTBP), was detected while TGF-β was not.

6.3.4 Metallothionein (MT)

MTs do not encode signal peptides; however, it is well documented that these proteins are secreted from astrocytes and exhibit strong neuroprotective properties. Originally, these neuroprotective properties were thought to originate from their heavy metal binding and free radical scavenging abilities; however, there is recent evidence that MTs elicit neuroprotection via the extracellular binding of the low density lipoprotein receptor and downstream activation of MAPK and PI3 kinase/Akt pathways. Exogeneous MT-I/-II

protects dopaminergic neurons from 6-hydroxydopamine toxicity and hippocampal neurons from beta-amyloid toxicity.[9] Three of the four MTs (-I, -II, and –III) were identified and found to be media enriched in the current study.

6.3.5 Superoxide Dismutase (SOD)

Superoxide dismutase is an important detoxification enzyme; it converts superoxide, a highly reactive radical, into oxygen and hydrogen peroxide. However, certain mutations in the SOD gene cause a toxic gain-of-function which is thought to underlie the inheritable cause of amytrophic lateral sclerosis (ALS). In addition to the cytosolic SOD, there is also an extracellular SOD; however, they are two distinct protein species and the cytsolic SOD does not contain a signal sequence. Interestingly, both were found to be secreted by astrocytes in the current experiment, corroborating other experiments in which cytosolic SOD was found to be secreted by astrocytes and contribute to the pathogenesis of ALS.[10-13]

6.3.6 Prothymosin

Extracellular prothymosin was recently found to be an extracellular death switch which inhibits necrosis and induces apoptosis. However, in a rodent model of stroke, systemic administration of prothymosin protected against ischemia-reperfusion injury.[14] In addition, prothymosin can liberate the neuroprotective transcription factor Nrf2 from its binding to Keap1, allowing Nrf2 to localize to the nucleus and activate gene transcription.[15] Prothymosin is not predicted to be secreted by either classical or non-classical pathways.

6.3.7 RNA Associated Proteins

Two RNA binding proteins and ribonuclease (RNase) were also found to be enriched in the astrocyte conditioned media. While the two heterogeneous nuclear ribonucleoproteins (hnRNPs) were not predicted to be secreted, they did show significant enrichment scores in the AFE and RA analysis. On the other hand, RNase is known to be secreted and passed both the RA and AFE filters. It is unexpected to discover astrocyte secreted RNA binding and degrading enzymes; however, there is precedent for the existence of extracellular RNA with discovery of extracellular RNA involvement in the blood coagulation pathway.[16] In addition, extracellular RNA has also been found in squid axon preparations and is thought to be derived from the surrounding astrocytes.[17, 18]

6.4 Conclusion

I have identified over 170 putative astrocyte secreted proteins using a combination of LC-MS/MS techniques, quantitative proteomics, and bioinformatics. The use of both a rank analysis (RA) and average fold enrichment (AFE) enabled a flexible and robust analysis of protein secretion. In addition, the information provided by SignalP/ SecretomeP added another degree of analytical leverage and security in my results. Although I have only highlighted a few of the identified proteins, there is a wealth of opportunity for future research—we are only at the beginning.

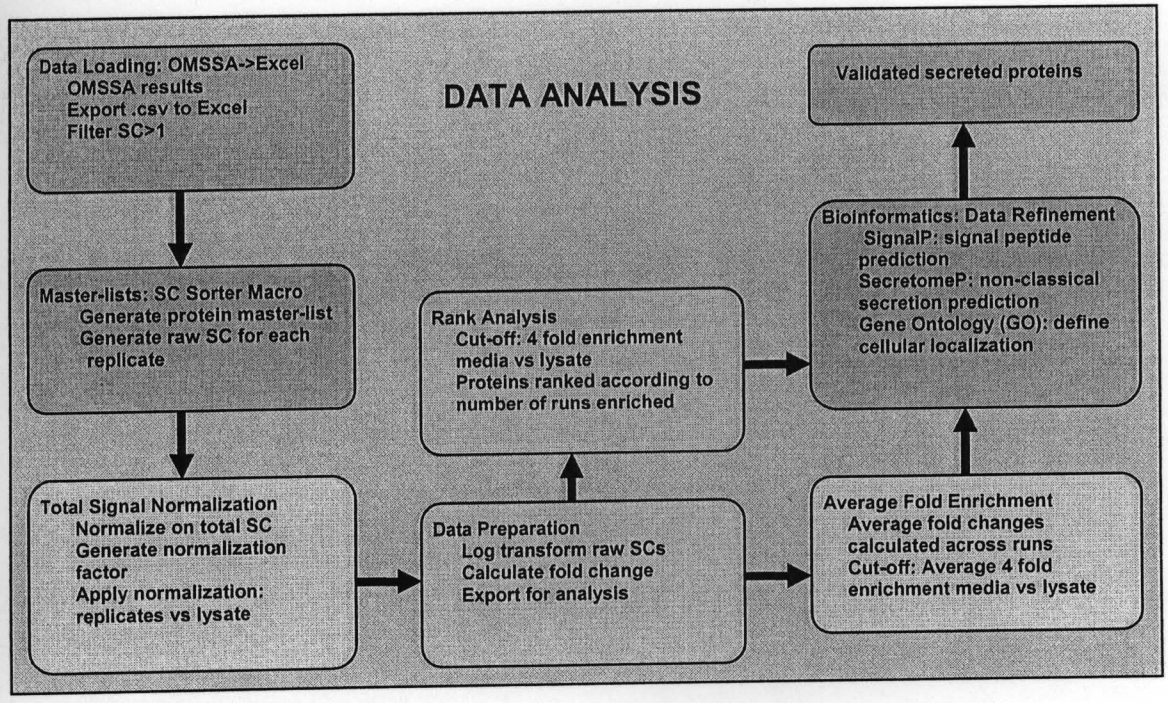

Figure 6.1 Data Analysis Overview. Database search results from OMSSA are exported to MS Excel. A master list is generated from the proteins identified in all the runs, including the lysate control. This master list is then re-sorted against each original list using an in-house generated Excel macro, generating an identical list of proteins for each run. The media runs are then normalized against the lysate control and then analyzed by Average Fold Enrichment (AFE) or Rank Analysis (RA).

Table 6.1 Total Identified Proteins with Significant SecretomeP/SignalP Score

Protein Name	AVG Enrichment (Log2)	Rank Analysis	Secreted
Follistatin-related protein 1 precursor (Follistatin-like 1) (TGF-beta-inducible protein TSC-36)	9.9	6	Yes
Insulin-like growth factor-binding protein 5 precursor (IGFBP-5) (IBP-5) (IGF-binding protein 5)	9.8	6	Yes
Secretogranin-3 precursor (Secretogranin III) (SgIII)	9.5	6	Yes
Collagen alpha-2(I) chain precursor (Alpha-2 type I collagen)	9.0	6	Yes
Cathepsin B precursor (Cathepsin B1) [Contains: Cathepsin B light chain; Cathepsin B heavy chain]	8.7	6	Yes
Calumenin precursor (Crocalbin)	8.6	6	signal peptide
Pigment epithelium-derived factor precursor (PEDF) (Serpin-F1) (Stromal cell-derived factor 3) (SDF-3) (Caspin)	8.5	6	Yes
Collagen alpha-1(I) chain precursor (Alpha-1 type I collagen)	8.3	5	Yes
Chitinase-3-like protein 1 precursor (Cartilage glycoprotein 39) (GP-39) (BRP39 protein)	7.6	5	Yes
Nucleobindin-1 precursor (CALNUC)	7.3	5	signal peptide
Serine protease inhibitor A3N precursor (Serpin A3N)	7.2	5	Yes
Calsyntenin-1 precursor	7.0	6	Yes
Dickkopf-related protein 3 precursor (Dkk-3) (Dickkopf-3) (mDkk-3)	6.9	5	Yes
Calmodulin (CaM)	6.6	6	0.676
Collagen alpha-2(V) chain precursor	6.5	5	Yes
Macrophage colony-stimulating factor 1 precursor (CSF-1) (MCSF) [Contains: Processed macrophage colony-stimulating factor 1]	6.4	4	Yes
Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)	6.4	4	Yes
Alpha-2-macroglobulin-P precursor (Alpha-2-macroglobulin)	6.0	6	Yes
SPARC precursor (Secreted protein acidic and rich in cysteine) (Osteonectin) (ON) (Basement-membrane protein 40) (BM-40)	5.9	6	Yes
Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin)	5.8	6	Yes
ProSAAS precursor (pro-SAAS) (Proprotein convertase subtilisin/kexin type 1 inhibitor)	5.8	4	Yes
Nidogen-2 precursor (NID-2) (Entactin-2)	5.6	4	Yes
Metallothionein-2 (MT-2) (Metallothionein-II) (MT-II)	5.5	4	Yes
Latent-transforming growth factor beta-binding protein 2 precursor (LTBP-2)	5.4	4	Yes
Proheparin-binding EGF-like growth factor precursor	5.3	4	Yes

Esophageal cancer-related gene 4 protein precursor	5.2	4	Yes
Extracellular matrix protein 1 precursor (Secretory component p85)	5.2	4	Yes
SPARC-like protein 1 precursor (Matrix glycoprotein Sc1) (Extracellular matrix protein 2)	5.0	4	Yes
Biotinidase precursor (Biotinase)	5.0	4	Yes
Sulfated 50 kDa glycoprotein precursor (SGP50) (Endothelial ligand FOR L-selectin)	4.9	4	Yes
Serine protease HTRA1 precursor (Serine protease 11)	4.9	4	Yes
Fibronectin precursor (FN)	4.9	4	Yes
Carboxypeptidase E precursor (CPE) (Carboxypeptidase H) (CPH) (Enkephalin convertase	4.9	4	Yes
Serine protease inhibitor A3G (Serpin A3G) (Serine protease inhibitor 2A) (Serpin 2A)	4.8	4	Yes
WNT1-inducible-signaling pathway protein 2 precursor (WISP-2)	4.8	4	Yes
Insulin-like growth factor-binding protein 2 precursor (IGFBP-2) (IBP-2) (IGF-binding protein 2)	4.8	5	Yes
Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh) (C1 esterase inhibitor) (C1-inhibiting factor)	4.8	4	Yes
Transcobalamin-2 precursor (Transcobalamin II) (TCII) (TC II)	4.7	4	Yes
Complement C3 precursor (HSE-MSF)	4.7	6	Yes
Cadherin-2 precursor (Neural cadherin) (N-cadherin) (CD325 antigen)	4.7	6	Yes
Collagen alpha-1(XI) chain precursor	4.6	4	Yes
Insulin-like growth factor-binding protein 4 precursor (IGFBP-4) (IBP-4) (IGF-binding protein 4)	4.5	4	Yes
Complement C4-B precursor	4.4	6	Yes
Neuroendocrine protein 7B2 precursor (Secretogranin-5) (Secretogranin V)	4.3	3	Yes
Insulin-like growth factor-binding protein 7 precursor (IGFBP-7) (IBP-7) (IGF-binding protein 7) (MAC25 protein)	4.3	3	Yes
Clusterin precursor (Sulfated glycoprotein 2) (SGP-2) (Clustrin) (Apolipoprotein J) (Apo-J)	4.2	6	Yes
Phospholipid transfer protein precursor (Lipid transfer protein II)	4.1	3	Yes
Biglycan precursor (Bone/cartilage proteoglycan I) (PG-S1)	4.0	3	Yes
Protein CYR61 precursor (Cysteine-rich angiogenic inducer 61) (Insulin-like growth factor-binding protein 10) (3CH61)	4.0	3	Yes
Fibulin-5 precursor (FIBL-5) (Developmental arteries and neural crest EGF-like protein) (Dance)	3.9	3	Yes
Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	3.9	3	0.812
Plasma glutamate carboxypeptidase precursor (Hematopoietic lineage switch 2)	3.9	3	Yes

Collagen alpha-1(IV) chain precursor	3.8	3	Yes
Fibromodulin precursor (FM) (Collagen-binding 59 kDa protein)	3.6	3	Yes
Scotin precursor	3.3	3	membrane
Serine protease inhibitor A3M precursor (Serpin A3M)	3.3	3	Yes
Metallothionein-1 (MT-1) (Metallothionein-I) (MT-I)	3.2	2	0.839
Lysosome-associated membrane glycoprotein 2 precursor (LAMP-2)	3.1	6	Yes
Neutrophil gelatinase-associated lipocalin precursor (NGAL) (p25) (SV-40-induced 24P3 protein) (Lipocalin-2)	3.0	2	Yes
Antithrombin-III precursor (ATIII)	3.0	3	Yes
Ceruloplasmin precursor (Ferroxidase)	2.8	2	Yes
Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERp60)	2.8	2	signal peptide
Granulins precursor (Proepithelin) (PEPI) (PC cell-derived growth factor) (PCDGF)	2.7	2	Yes
Epididymal secretory protein E1 precursor (Niemann Pick type C2 protein homolog) (mE1)	2.6	2	Yes
Acyltransferase-like 1-A	2.6	2	membrane
Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) (S-cyclophilin) (SCYLP) (CYP-S1)	2.6	2	Yes
BDNF/NT-3 growth factors receptor precursor (Neurotrophic tyrosine kinase receptor type 2) (TrkB tyrosine kinase) (GP145-TrkB/GP95-TrkB) (Trk-B)	2.6	2	membrane
Mimecan precursor (Osteoglycin)	2.6	3	Yes
Aspartate aminotransferase, mitochondrial precursor (Transaminase A	2.5	2	signal peptide
Elastin precursor (Tropoelastin)	2.5	2	Yes
Protein S100-A11 (S100 calcium-binding protein A11) (Protein S100C) (Calgizzarin) (Endothelial monocyte-activating polypeptide) (EMAP)	2.5	5	0.812
Extracellular superoxide dismutase [Cu-Zn] precursor (EC-SOD)	2.5	3	Yes
Histone H2A type 1	2.5	2	0.536
Peptidyl-prolyl cis-trans isomerase C (PPIase) (Rotamase) (Cyclophilin C)	2.5	2	Yes
Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane protein NAP-22)	2.5	2	membrane bound
Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer disease amyloid protein homolog)	2.4	2	membrane
Calsequestrin-1 precursor (Calsequestrin, skeletal muscle isoform)	2.4	2	signal peptide
EGF-containing fibulin-like extracellular matrix protein 2 precursor (Fibulin-4) (FIBL-4)	2.4	2	Yes
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 precursor (Lysyl hydroxylase 1) (LH1)	2.4	2	Yes

Insulin-like growth factor-binding protein-like 1 precursor (Insulin-like growth factor-binding-related protein 4)	2.4	2	Yes
Metalloproteinase inhibitor 2 precursor (TIMP-2) (Tissue inhibitor of metalloproteinases 2)	2.4	3	Yes
14-3-3 protein sigma (Stratifin)	2.4	2	Yes
Intercellular adhesion molecule 1 precursor (ICAM-1) (CD54 antigen) (MALA-2)	2.4	2	Yes
Neuroblastoma suppressor of tumorigenicity 1 precursor (Zinc finger protein DAN) (N03)	2.4	2	Yes
Collagen alpha-1(II) chain precursor (Alpha-1 type II collagen) [Contains: Chondrocalcin]	2.3	2	Yes
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 precursor (Lysyl hydroxylase 2) (LH2)	2.3	2	Yes
Ribonuclease pancreatic precursor (RNase 1) (RNase A)	2.3	2	Yes
Transthyretin precursor (Prealbumin)	2.3	2	Yes
Phosphoinositide-3-kinase-interacting protein 1 precursor (Kringle domain-containing protein HGFL)	2.3	2	Yes
Prostaglandin-H2 D-isomerase precursor (Lipocalin-type prostaglandin-D synthase)	2.3	2	Yes
Spondin-1 precursor (F-spondin)	2.3	2	Yes
Coiled-coil domain-containing protein 80 precursor (Up-regulated in BRS-3 deficient mouse)	2.3	2	Yes
FXYD domain-containing ion transport regulator 6 precursor (PLM-like protein)	2.3	2	membrane
Metalloproteinase inhibitor 1 precursor (TIMP-1)	2.3	2	Yes
CD44 antigen precursor (Phagocytic glycoprotein 1)	2.2	2	Yes
Collagen alpha-1(V) chain precursor	2.2	2	Yes
Metallothionein-3 (MT-3) (Metallothionein-III) (MT-III) (Growth inhibitory factor) (GIF)	2.2	2	0.693
Palmitoyl-protein thioesterase 1 precursor (PPT-1) (Palmitoyl-protein hydrolase 1)	2.2	2	Yes
Serine protease inhibitor A3C precursor (Serpin A3C) (Kallikrein-binding protein) (KBP)	2.2	2	Yes
Serine protease inhibitor A3F (Serpin A3F)	2.2	2	Yes
Serine protease inhibitor A3K precursor (Serpin A3K) (Contrapsin) (SPI-2)	2.2	2	Yes
Prothymosin alpha [Contains: Thymosin alpha]	2.2	2	Yes
Platelet-derived growth factor A chain precursor (PDGF A-chain) (Platelet-derived growth factor alpha polypeptide)	2.2	2	Yes
Cathepsin Z precursor	2.1	2	Yes
Plasminogen activator inhibitor 1 precursor (PAI-1) (Endothelial plasminogen activator inhibitor) (PAI)	2.1	2	Yes
72 kDa type IV collagenase precursor (72 kDa gelatinase) (Matrix metalloproteinase-2) (MMP-2) (Gelatinase A)	2.0	2	Yes
N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase precursor	2.0	2	signal peptide
Cell adhesion molecule 4 precursor (Immunoglobulin	1.8	3	membrane

superfamily member 4C) (Nectin-like protein 4)			
Insulin-like growth factor-binding protein 3 precursor (IGFBP-3) (IBP-3) (IGF-binding protein 3)	1.8	4	Yes
Superoxide dismutase [Cu-Zn]	1.8	2	0.76
Elongation factor 1-beta (EF-1-beta)	1.5	3	0.655
Tenascin precursor (TN) (Tenascin-C) (TN-C) (Hexabrachion)	1.5	1	Yes
Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy chain H2) (Inter-alpha-inhibitor heavy chain 2)	1.4	1	Yes
Thymosin beta-4 (T beta 4) [Contains: Hematopoietic system regulatory peptide (Seraspenide)]	1.4	2	0.864
Apolipoprotein E precursor (Apo-E)	1.4	3	Yes
CD109 antigen precursor (GPI-anchored alpha-2 macroglobulin-related protein)	1.3	2	signal peptide
Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein 1) (SPP-1)	1.3	1	Yes
Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180) (NCAM-180) (CD56 antigen)	1.3	1	membrane
Cathepsin D precursor	1.3	2	Yes
Uncharacterized protein C19orf10 homolog precursor (Stromal cell-derived growth factor SF20) (Interleukin-25)	1.3	1	Yes
Connective tissue growth factor precursor (Protein FISP-12) (Hypertrophic chondrocyte-specific protein 24)	1.3	1	Yes
Vitamin K-dependent protein S precursor	1.3	1	Yes
40S ribosomal protein S10	1.3	2	0.685
N-acetylglucosamine-6-sulfatase precursor (G6S) (Glucosamine-6-sulfatase)	1.2	2	Yes
Reticulocalbin-2 precursor (Taipoxin-associated calcium-binding protein 49) (TCBP-49)	1.2	2	signal peptide
ADAMTS-like protein 1 precursor (Punctin-1)	1.2	1	Yes
Basement membrane-specific heparan sulfate proteoglycan core protein precursor (HSPG) (Perlecan) (PLC)	1.2	1	Yes
Matrix-remodeling-associated protein 7 (Transmembrane anchor protein 1)	1.2	1	membrane
PRKC apoptosis WT1 regulator protein (Prostate apoptosis response 4 protein) (Par-4)	1.2	1	0.799
Thy-1 membrane glycoprotein precursor (Thy-1 antigen) (CD90 antigen)	1.2	1	membrane bound
Vascular cell adhesion protein 1 precursor (V-CAM 1) (CD106 antigen)	1.2	1	membrane
Phosphatidylethanolamine-binding protein 1 (PEBP-1) (HCNPpp) [Contains: Hippocampal cholinergic neurostimulating peptide (HCNP)]	1.2	2	Yes
Delta-like protein precursor (DLK) (Preadipocyte factor 1) (Pref-1) (Adipocyte differentiation inhibitor protein)	1.2	1	membrane
DNA-binding protein A (Cold shock domain-containing protein A) (Y-box protein 3)	1.2	1	0.676

Fibulin-1 precursor (Basement-membrane protein 90)	1.2	1	Yes
Reticulocalbin-1 precursor	1.2	2	signal peptide
Y-box-binding protein 2 (Germ cell-specific Y-box-binding protein) (FRGY2 homolog)	1.2	1	0.71
Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120) (NCAM-120) (CD56 antigen)	1.2	1	membrane
Sulfhydryl oxidase 1 precursor (Quiescin Q6) (Skin sulfhydryl oxidase) (mSOx)	1.2	1	Yes
Vascular endothelial growth factor B precursor (VEGF-B) (VEGF-related factor) (VRF)	1.2	1	Yes
Cathepsin L1 precursor (Major excreted protein) (MEP) (p39 cysteine proteinase)	1.1	1	Yes
Amyloid-like protein 2 precursor (CDEI box-binding protein) (CDEBP)	1.1	1	membrane
Beta-hexosaminidase beta chain precursor (N-acetyl-beta-glucosaminidase) (Beta-N-acetylhexosaminidase)	1.1	1	signal peptide
Chordin-like protein 1 precursor (Neuralin-1) (Ventroptin) (Neurogenesin-1)	1.1	1	Yes
FMRFamide-related peptides precursor (Neuropeptide VF)	1.1	1	Yes
Putative RNA-binding protein 3 (RNA-binding motif protein 3)	1.1	1	0.745
Nuclease sensitive element-binding protein 1 (Y-box-binding protein 1) (Y-box transcription factor) (YB-1)	1.1	2	0.73
3'(2'),5'-bisphosphate nucleotidase 1 (Bisphosphate 3'-nucleotidase 1) (PAP-inositol-1,4-phosphatase) (PIP)	1.1	1	0.652
Lysosomal protective protein precursor (Cathepsin A)	1.1	1	signal peptide
Lysyl oxidase homolog 3 precursor (Lysyl oxidase-like protein 3) (Lysyl oxidase-related protein 2)	1.1	1	Yes
Probable ATP-dependent RNA helicase DDX56 (DEAD box protein 56) (ATP-dependent 61 kDa nucleolar RNA helicase)	1.1	1	0.534
Tyrosine-protein phosphatase non-receptor type substrate 1 precursor (SHP substrate 1) (SHPS-1)	1.1	1	membrane
Cystatin-C precursor (Cystatin-3)	1.0	2	Yes
Prolargin precursor (Proline-arginine-rich end leucine-rich repeat protein)	1.0	1	Yes
Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3) (ITI-HC3) (Inter-alpha-inhibitor heavy chain 3)	1.0	1	Yes
Prostaglandin E synthase 3 (Cytosolic prostaglandin E2 synthase) (cPGES)	1.0	1	0.706
Small ubiquitin-related modifier 2 precursor (SUMO-2) (Ubiquitin-like protein SMT3B)	0.9	2	0.883
Small ubiquitin-related modifier 3 precursor (SUMO-3) (Ubiquitin-like protein SMT3A) (SMT3 homolog 1)	0.9	2	0.883
Brain-specific polypeptide PEP-19 (Brain-specific antigen PCP-4) (Purkinje cell protein 4)	0.9	1	0.704
Collagen alpha-2(IV) chain precursor (Canstatin)	0.9	1	Yes

Ribonuclease T2 precursor (Ribonuclease 6)	0.9	1	Yes
Serine protease inhibitor A3A precursor (Serpin A3A)	0.9	1	Yes
Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal-binding globulin)	0.9	1	Yes
Serpin H1 precursor (Collagen-binding protein) (Colligin) (47 kDa heat shock protein) (Serine protease inhibitor J6)	0.8	2	Yes
Latexin (Endogenous carboxypeptidase inhibitor) (ECI) (Tissue carboxypeptidase inhibitor) (TCI)	0.7	2	0.554
Legumain precursor (Asparaginyl endopeptidase) (Protease, cysteine 1)	0.6	1	signal peptide
Ubiquitin-conjugating enzyme E2 L3 (Ubiquitin-protein ligase L3) (Ubiquitin carrier protein L3) (UbcM4)	0.6	1	0.582
Apolipoprotein A-I-binding protein precursor (AI-BP)	0.5	1	Yes
Drebrin (Developmentally-regulated brain protein)	0.4	1	0.72
Protein kinase C and casein kinase substrate in neurons protein 2	0.4	1	0.608
Platelet-activating factor acetylhydrolase precursor (PAF acetylhydrolase) (PAF 2-acylhydrolase)	0.3	0	Yes

Table 6.2 Proteins identified with an average fold enrichment of at least 4 fold

Protein Name	AVG Enrichment (Log2)	Rank Analysis	Secreted
Follistatin-related protein 1 precursor (Follistatin-like 1) (TGF-beta-inducible protein TSC-36)	9.9	6	Yes
Insulin-like growth factor-binding protein 5 precursor (IGFBP-5) (IBP-5) (IGF-binding protein 5)	9.8	6	Yes
Secretogranin-3 precursor (Secretogranin III) (SgIII)	9.5	6	Yes
Collagen alpha-2(I) chain precursor (Alpha-2 type I collagen)	9.0	6	Yes
Cathepsin B precursor (Cathepsin B1) [Contains: Cathepsin B light chain; Cathepsin B heavy chain]	8.7	6	Yes
Calumenin precursor (Crocalbin)	8.6	6	signal peptide
Pigment epithelium-derived factor precursor (PEDF) (Serpin-F1) (Stromal cell-derived factor 3) (SDF-3) (Caspin)	8.5	6	Yes
Collagen alpha-1(I) chain precursor (Alpha-1 type I collagen)	8.3	5	Yes
Chitinase-3-like protein 1 precursor (Cartilage glycoprotein 39) (GP-39) (BRP39 protein)	7.6	5	Yes
Nucleobindin-1 precursor (CALNUC)	7.3	5	signal peptide
Serine protease inhibitor A3N precursor (Serpin A3N)	7.2	5	Yes
Calsyntenin-1 precursor	7.0	6	Yes
Dickkopf-related protein 3 precursor (Dkk-3) (Dickkopf-3) (mDkk-3)	6.9	5	Yes
Calmodulin (CaM)	6.6	6	0.676
Collagen alpha-2(V) chain precursor	6.5	5	Yes
Macrophage colony-stimulating factor 1 precursor (CSF-1) (MCSF	6.4	4	Yes
Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)	6.4	4	Yes
Alpha-2-macroglobulin-P precursor (Alpha-2-macroglobulin)	6.0	6	Yes
SPARC precursor (Secreted protein acidic and rich in cysteine)	5.9	6	Yes
Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin)	5.8	6	Yes
ProSAAS precursor (pro-SAAS) (Proprotein convertase subtilisin/kexin type 1	5.8	4	Yes
Nidogen-2 precursor (NID-2) (Entactin-2)	5.6	4	Yes
Metallothionein-2 (MT-2) (Metallothionein-II) (MT-II)	5.5	4	Yes
Latent-transforming growth factor beta-binding protein 2 precursor (LTBP-2)	5.4	4	Yes
Proheparin-binding EGF-like growth factor precursor	5.3	4	Yes
Esophageal cancer-related gene 4 protein precursor	5.2	4	Yes
Extracellular matrix protein 1 precursor (Secretory component p85)	5.2	4	Yes
SPARC-like protein 1 precursor (Matrix glycoprotein Sc1) (Extracellular matrix protein 2)	5.0	4	Yes

Biotinidase precursor (Biotinase)	5.0	4	Yes
Sulfated 50 kDa glycoprotein precursor (SGP50) (Endothelial ligand FOR L-selectin	4.9	4	Yes
Serine protease HTRA1 precursor (Serine protease 11)	4.9	4	Yes
Fibronectin precursor (FN)	4.9	4	Yes
Carboxypeptidase E precursor (CPE) (Carboxypeptidase H) (CPH) (Enkephalin convertase)	4.9	4	Yes
Serine protease inhibitor A3G (Serpin A3G) (Serine protease inhibitor 2A) (Serpin 2A)	4.8	4	Yes
WNT1-inducible-signaling pathway protein 2 precursor (WISP-2) (Connective tissue growth factor-like protein)	4.8	4	Yes
Insulin-like growth factor-binding protein 2 precursor (IGFBP-2) (IBP-2) (IGF-binding protein 2)	4.8	5	Yes
Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh) (C1 esterase inhibitor) (C1-inhibiting factor)	4.8	4	Yes
Transcobalamin-2 precursor (Transcobalamin II)	4.7	4	Yes
Complement C3 precursor (HSE-MSF)	4.7	6	Yes
Cadherin-2 precursor (Neural cadherin) (N-cadherin) (CD325 antigen)	4.7	6	Yes
Collagen alpha-1(XI) chain precursor	4.6	4	Yes
Insulin-like growth factor-binding protein 4 precursor (IGFBP-4) (IBP-4) (IGF-binding protein 4)	4.5	4	Yes
Complement C4-B	4.4	6	Yes
Neuroendocrine protein 7B2 precursor (Secretogranin-5) (Secretogranin V) (Secretory granule endocrine protein I)	4.3	3	Yes
Insulin-like growth factor-binding protein 7 precursor (IGFBP-7) (IBP-7) (IGF-binding protein 7) (MAC25 protein)	4.3	3	Yes
Clusterin precursor (Sulfated glycoprotein 2) (SGP-2) (Clustrin) (Apolipoprotein J) (Apo-J)	4.2	6	Yes
Phospholipid transfer protein precursor (Lipid transfer protein II)	4.1	3	Yes
Biglycan precursor (Bone/cartilage proteoglycan I) (PG-S1)	4.0	3	Yes
Protein CYR61 precursor (Cysteine-rich angiogenic inducer 61) (Insulin-like growth factor-binding protein 10) (3CH61)	4.0	3	Yes
Fibulin-5 precursor (FIBL-5) (Developmental arteries and neural crest EGF-like protein) (Dance)	3.9	3	Yes
Glyceraldehyde-3-phosphate dehydrogenase, testis-specific (Spermatogenic glyceraldehyde-3-phosphate dehydrogenase)	3.9	3	0.812
Plasma glutamate carboxypeptidase precursor (Hematopoietic lineage switch 2)	3.9	3	Yes
Collagen alpha-1(IV) chain precursor	3.8	3	Yes
Fibromodulin precursor (FM) (Collagen-binding 59 kDa protein) (Keratan sulfate proteoglycan fibromodulin)	3.6	3	Yes
Scotin precursor	3.3	3	membrane
Serine protease inhibitor A3M precursor (Serpin A3M)	3.3	3	Yes
Metallothionein-1 (MT-1) (Metallothionein-I) (MT-I)	3.2	2	0.839
Lysosome-associated membrane glycoprotein 2 precursor (LAMP-2) (Lysosomal membrane glycoprotein type B)	3.1	6	Yes
Neutrophil gelatinase-associated lipocalin precursor (NGAL) (p25) (SV-40-induced 24P3 protein) (Lipocalin-2)	3.0	2	Yes

Antithrombin-III precursor (ATIII)	3.0	3	Yes
Ceruloplasmin precursor (Ferroxidase)	2.8	2	Yes
Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERp60)	2.8	2	signal peptide
Granulins precursor (Proepithelin) (PEPI) (PC cell-derived growth factor) (PCDGF)	2.7	2	Yes
Histone H4	2.6	2	0.408
Epididymal secretory protein E1 precursor (Niemann Pick type C2 protein homolog) (mE1)	2.6	2	Yes
Acyltransferase-like 1-A	2.6	2	membrane
Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) (S-cyclophilin	2.6	2	Yes
BDNF/NT-3 growth factors receptor precursor (Neurotrophic tyrosine kinase receptor type 2) (TrkB tyrosine kinase)	2.6	2	membrane
Mimecan precursor (Osteoglycin)	2.6	3	Yes
Aspartate aminotransferase, mitochondrial precursor (Transaminase A) (Glutamate oxaloacetate transaminase 2) (mAspAT) (Fatty acid-binding protein) (FABP-1) (FABPpm)	2.5	2	signal peptide
Elastin precursor (Tropoelastin)	2.5	2	Yes
Tropomyosin alpha-4 chain (Tropomyosin-4)	2.5	5	0.485
Protein S100-A11 (S100 calcium-binding protein A11) (Protein S100C) (Calgizzarin)	2.5	5	0.812
Extracellular superoxide dismutase [Cu-Zn] precursor (EC-SOD)	2.5	3	Yes
Histone H2A type 1	2.5	2	0.536
Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Single-strand-binding protein)	2.5	2	0.087
Peptidyl-prolyl cis-trans isomerase C (PPIase) (Rotamase) (Cyclophilin C)	2.5	2	Yes
Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane protein NAP-22	2.5	2	membrane bound
Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer disease amyloid protein homolog)	2.4	2	membrane
Calsequestrin-1 precursor (Calsequestrin, skeletal muscle isoform)	2.4	2	signal peptide
EGF-containing fibulin-like extracellular matrix protein 2 precursor (Fibulin-4) (FIBL-4)	2.4	2	Yes
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 precursor (Lysyl hydroxylase 1) (LH1)	2.4	2	Yes
Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2 / hnRNP B1)	2.4	3	0.081
Insulin-like growth factor-binding protein-like 1 precursor (Insulin-like growth factor-binding-related protein 4)	2.4	2	Yes
Metalloproteinase inhibitor 2 precursor (TIMP-2) (Tissue inhibitor of metalloproteinases 2)	2.4	3	Yes
14-3-3 protein sigma (Stratifin)	2.4	2	Yes
Intercellular adhesion molecule 1 precursor (ICAM-1) (CD54 antigen) (MALA-2)	2.4	2	Yes
Neuroblastoma suppressor of tumorigenicity 1 precursor (Zinc finger protein DAN) (N03)	2.4	2	Yes

Collagen alpha-1(II) chain precursor (Alpha-1 type II collagen) [Contains: Chondrocalcin]	2.3	2	Yes
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 precursor (Lysyl hydroxylase 2) (LH2)	2.3	2	Yes
Ribonuclease pancreatic precursor (RNase 1) (RNase A)	2.3	2	Yes
Transthyretin precursor (Prealbumin)	2.3	2	Yes
Cysteine and glycine-rich protein 1 (Cysteine-rich protein 1) (CRP1) (CRP)	2.3	2	0.27
Dedicator of cytokinesis protein 7	2.3	2	0.305
Phosphoinositide-3-kinase-interacting protein 1 precursor (Kringle domain-containing protein HGFL	2.3	2	Yes
Prostaglandin-H2 D-isomerase precursor (Lipocalin-type prostaglandin-D synthase)	2.3	2	Yes
Spondin-1 precursor (F-spondin)	2.3	2	Yes
FXYD domain-containing ion transport regulator 6 precursor (PLM-like protein)	2.3	2	membrane
Metalloproteinase inhibitor 1 precursor (TIMP-1) (Erythroid-potentiating activity) (EPA)	2.3	2	Yes
Coiled-coil domain-containing protein 80 precursor (Up-regulated in BRS-3 deficient mouse)	2.3	2	Yes
CD44 antigen precursor (Phagocytic glycoprotein 1) (PGP-1) (HUTCH-I) (Extracellular matrix receptor III)	2.2	2	Yes
Collagen alpha-1(V) chain precursor	2.2	2	Yes
Metallothionein-3 (MT-3) (Metallothionein-III) (MT-III) (Growth inhibitory factor) (GIF)	2.2	2	0.693
Palmitoyl-protein thioesterase 1 precursor (PPT-1) (Palmitoyl-protein hydrolase 1)	2.2	2	Yes
Prothymosin alpha [Contains: Thymosin alpha]	2.2	2	yes
Serine protease inhibitor A3C precursor (Serpin A3C) (Kallikrein-binding protein) (KBP)	2.2	2	Yes
Serine protease inhibitor A3F (Serpin A3F)	2.2	2	Yes
Serine protease inhibitor A3K precursor (Serpin A3K) (Contrapsin) (SPI-2)	2.2	2	Yes
Platelet-derived growth factor A chain precursor (PDGF A-chain	2.2	2	Yes
60S acidic ribosomal protein P2	2.1	2	0.26
Lamina-associated polypeptide 2 isoforms beta/delta/epsilon/gamma	2.1	2	0.306
Cathepsin Z precursor	2.1	2	Yes
Plasminogen activator inhibitor 1 precursor (PAI-1) (Endothelial plasminogen activator inhibitor) (PAI)	2.1	2	Yes
Protein ECT2 (Epithelial cell-transforming sequence 2 oncogene)	2.1	2	0.16
72 kDa type IV collagenase precursor (72 kDa gelatinase) (Matrix metalloproteinase-2) (MMP-2) (Gelatinase A)	2.0	2	Yes
N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase precursor (Glycosylasparaginase)	2.0	2	signal peptide

6.5 References

1. Lafon-Cazal, M.; Adjali, O.; Galeotti, N.; Poncet, J.; Jouin, P.; Homburger, V.; Bockaert, J.; Marin, P. *J. Biol. Chem.* **2003**, *27,* 24438-24448.

2. Delcourt, N.; Jouin, P.; Poncet, J.; Demey, E.; Mauger, E.; Bockaert, J.; Marin, P.; Galeotti, N. *Mol. Cell. Proteomics* **2005**, *8,* 1085-1094.

3. Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. *J. Mol. Biol.* **2004**, *4,* 783-795.

4. Bendtsen, J. D.; Jensen, L. J.; Blom, N.; Von Heijne, G.; Brunak, S. *Protein Eng. Des. Sel.* **2004**, *4,* 349-356.

5. Wyss-Coray, T.; Mucke, L. *Neuron* **2002**, *3,* 419-432.

6. Pisalyaput, K.; Tenner, A. J. *J. Neurochem.* **2008**, *3,* 696-707.

7. Stevens, B.; Allen, N. J.; Vazquez, L. E.; Howell, G. R.; Christopherson, K. S.; Nouri, N.; Micheva, K. D.; Mehalow, A. K.; Huberman, A. D.; Stafford, B.; Sher, A.; Litke, A. M.; Lambris, J. D.; Smith, S. J.; John, S. W.; Barres, B. A. *Cell* **2007**, *6,* 1164-1178.

8. Russo, V. C.; Gluckman, P. D.; Feldman, E. L.; Werther, G. A. *Endocr. Rev.* **2005**, *7,* 916-943.

9. Chung, R. S.; Hidalgo, J.; West, A. K. *J. Neurochem.* **2008**, *1,* 14-20.

10. Atkin, J. D.; Farg, M. A.; Turner, B. J.; Tomas, D.; Lysaght, J. A.; Nunan, J.; Rembach, A.; Nagley, P.; Beart, P. M.; Cheema, S. S.; Horne, M. K. *J. Biol. Chem.* **2006**, *40,* 30152-30165.

11. Turner, B. J.; Atkin, J. D.; Farg, M. A.; Zang, D. W.; Rembach, A.; Lopes, E. C.; Patch, J. D.; Hill, A. F.; Cheema, S. S. *J. Neurosci.* **2005**, *1,* 108-117.

12. Urushitani, M.; Sik, A.; Sakurai, T.; Nukina, N.; Takahashi, R.; Julien, J. P. *Nat. Neurosci.* **2006**, *1,* 108-118.

13. Urushitani, M.; Ezzi, S. A.; Matsuo, A.; Tooyama, I.; Julien, J. P. *FASEB J.* **2008**, *7,* 2476-2487.

14. Fujita, R.; Ueda, H. *Cell Death Differ.* **2007**, *10,* 1839-1842.

15. Karapetian, R. N.; Evstafieva, A. G.; Abaeva, I. S.; Chichkova, N. V.; Filonov, G. S.; Rubtsov, Y. P.; Sukhacheva, E. A.; Melnikov, S. V.; Schneider, U.; Wanker, E. E.; Vartapetian, A. B. *Mol. Cell. Biol.* **2005**, *3,* 1089-1099.

16. Kannemeier, C.; Shibamiya, A.; Nakazawa, F.; Trusheim, H.; Ruppert, C.; Markart, P.; Song, Y.; Tzima, E.; Kennerknecht, E.; Niepmann, M.; von Bruehl, M. L.; Sedding, D.; Massberg, S.; Gunther, A.; Engelmann, B.; Preissner, K. T. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *15,* 6388-6393.

17. Giuditta, A.; Chun, J. T.; Eyman, M.; Cefaliello, C.; Bruno, A. P.; Crispino, M. *Physiol. Rev.* **2008**, *2,* 515-555.

18. Giuditta, A.; Chun, J. T.; Eyman, M.; Cefaliello, C.; Bruno, A. P.; Crispino, M. *Riv. Biol.* **2007**, *2,* 203-219.

Chapter 7: Future Directions

7.1 Specific Aims

Oxidative stress has been implicated as a causative agent in a number of neurodegenerative diseases, including Parkinson's, Alzheimer's, and amyotrophic lateral sclerosis (ALS)[1]. Enhancing the resistance of neurons to oxidative stress holds great therapeutic potential for all of these diseases. In light of this, the discovery the antioxidant response element (ARE), which acts like a master switch in a cell's response to oxidative stress, has generated hope for new treatment options for neurodegenerative diseases[2].

The ARE is a regulatory sequence present in the promoter region of many genes responsible for detoxifying the cell[3]. Upon oxidative insult, a transcription factor called NF-E2-related factor (Nrf2) binds to ARE sites within the genome and activates the cell's antioxidant defenses[4]. Up-regulation of Nrf2 enhances neuronal resistance to oxidative insult[5]. Our lab has recently discovered that Nrf2 is expressed at higher endogenous levels in astrocytes than in neurons, possibly indicating a prominent role of astrocytes in Nrf2-ARE mediated neuroprotection. In addition, astrocyte-specific overexpression of Nrf2 confers neuroprotection to co-cultured neurons[6, 7]. How these neuroprotective effects are transmitted from astrocyte to neuron is unknown, although glutathione synthesis and secretion seems to be an important element[6, 8]. In addition to enhanced neuroprotection, astrocyte-specific Nrf2 overexpression produces gene expression changes within the co-cultured neuronal population. This gene induction would seem to indicate some form of astrocyte-to-neuron communication. Experiments performed in our lab in which the co-cultured astrocytes and neurons were physically separated (membrane-delimited co-culture), indicate that the gene

inducing factor is secreted[6]. **We propose to combine mass spectrometry-based expression proteomics, molecular weight fractionation and bioactivity assays to identify the astrocyte-secreted neuroprotective factor.**

Aim 1: Combine quantitative neuroproteomics and bioactivity assays to identify secreted neuroprotective factors from Nrf2 over expressing astrocytes. Mass spectrometry-based expression proteomics is a powerful technique. However, without a focused experimental goal, the large amount of data generated by these types of experiments can be overwhelming and quite difficult to analyze. In order to circumvent this problem, we propose to combine molecular weight fractionation, neuroprotection assays, and quantitative proteomics. We will fractionate astrocyte conditioned media from both Nrf2 KO and GFAP-Nrf2 transgenic mice using molecular weight cut-off tubes and then assay each fraction for neuroprotection. The fractionation(s) which exhibit differential neuroprotection will then be subject to quantitative mass spectrometry to identify those proteins which are differentially secreted. **The combination of expression proteomics, fractionation, and bioactivity assays allow us to identify the important protein(s) which are responsible for the GFAP-Nrf2 astroctye mediated neuroprotection.**

Aim 2: Validation of Neuroprotection. Once we have identified potential neuroprotective candidates, we will validate the candidates using western blotting, antibody depletion, and direct application of the proteins to neuronal cultures. **The discovery of novel**

neuroprotective factors would expand potential treatment avenues for neurodegenerative diseases.

7.2 Background and Significance

Oxidative Stress, Neurodegenerative Diseases, and the ARE: Neurodegenerative disorders affect millions of people worldwide and the treatment options for those affected are limited[9]. Oxidative stress has been implicated in a number of these diseases, including Parkinson's Alzheimer's, and ALS, and potential treatment strategies have focused on ways to curtail oxidative stress[1]. After the neuroprotective effects of Nrf2/ARE activation were discovered, Nrf2 emerged as a promising drug target to treat neurodegenerative diseases[10]. **However, in order to develop effective therapeutics, elucidation of Nrf2-dependent signaling between astrocytes and neurons is vital.**

The ARE is a *cis*-acting regulatory sequence in the promoter region of many genes responsible for the protection of cells against oxidative stress[3, 5]. This response is dependent on the binding of Nrf2, a transcription factor, to a consensus sequence in the ARE[4]. When oxidative insult is absent, Nrf2 is kept inactive in the cell's cytoplasm by its cysteine-rich binding partner Keap1. Upon oxidative insult specific cysteine residues in Keap1 are oxidized, resulting in the release and nuclear translocation of Nrf2[11]. Activation of the ARE by Nrf2 results in the activation of a number of antioxidant genes and neuroprotection[5, 12]. However, a recent finding in our lab has shown this neuroprotection to be more strongly potentiated by Nrf2 overexpression in astrocytes than overexpression in neurons. In addition, for unknown reasons, in viral-mediated Nrf2 overexpression, astrocytes were selectively

infected with the ad-Nrf2 construct over neurons. Interestingly, the astrocyte specific overexpression of Nrf2 still confers neuroprotection to the co-cultured neurons and induces neuronal gene expression changes[6, 7].

In order to assess the specific genetic changes associated with astrocyte-specific Nrf2 overexpression in astrocytes and neurons separately, oligonucleotide microarray analysis was performed on either mixed neuronal/astroglial cultures or astroglial-enriched cultures[6]. A number of genes showed similar expression changes in both cultures. However, many genes exhibited unique expression changes in either the glial-enriched culture or the mixed culture. Many of the Nrf2-dependent expression changes were in secreted proteins/peptides (see Preliminary Data). In addition to the above mentioned experiments, a membrane delimited co-culture was performed. This system, which physically isolates Nrf2 overexpressing astrocytes from the neurons, exhibited the same neuroprotective effects as the standard co-culture system. **The logical inference from these results is the existence of a diffusible factor which induces neuroprotection and gene expression changes in the non-Nrf2 overexpressing neuronal population.**

In a separate study, the effects of *tert*-butylhydroquinone (tBHQ) (a chemical inducer of the ARE) treatment on gene expression were examined[7]. A mixed culture of neurons and astrocytes was infected with a virus containing a GFP construct (ad-GFP). Once again, the astrocytes were selectively infected and thus, only the astrocytes exhibited fluorescence. The infected mixed culture was then sorted by fluorescent-activated cell sorting (FACS), yielding pure samples of both the fluorescent astrocytes and non-fluorescent neurons. To identify Nrf2-regulated gene changes, both vehicle and tBHQ treated cultures were sorted and analyzed by DNA microarray. In a separate experiment, a mixed culture was infected with a dominant negative Nrf2 construct (ad-DN-Nrf2). Once again, the astroctyes were selectively

infected, resulting in a wild-type neuronal population and an astrocytic population overexpressing DN-Nrf2. This culture system was highly susceptible to oxidative stress, implicating an Nrf2 mediated neuroprotective mechanism. After FACS sorting and microarray analysis of this sample, the majority of tBHQ induced gene changes seen in the ad-GFP cultures were now unchanged, including the gene changes in the co-cultured neurons. **We hypothesize that the neuroprotection and neuronal gene changes associated Nrf2 overexpression is the result of a diffusible factor or factors of astrocytic origin.**

Astrocyte and Secreted Neurotrophic Factors: The role of astrocytes in the central nervous system (CNS) was traditionally thought to be 'brain glue'[13]. Astrocytes were considered boring cells which simply provided structural and nutrient support to neurons. However, with the discovery of astrocytic voltage-gated ion channels and complex intracellular calcium fluctuations, these 'support cells' are becoming increasingly important to our view of neuronal signaling and overall brain function[13]. One of the most intriguing developments has been the recognition of neuronal-astrocytic cross-talk and the discovery of signaling not only from neurons to astrocytes but from astrocytes to neurons[14].

The regulation of extracellular glutamate by astrocytes was one of the first discoveries that indicated astrocytes were participating in synaptic signaling[15]. Astrocytes regulate glutamate by both uptake and release. Astrocytes themselves respond to glutamate with intracellular calcium fluctuations. These calcium fluctuations affect the release of both glutamate and neuropeptide containing vesicles from astrocytes[16, 17]. A variety of astrocyte released signaling molecules, including D-serine, ATP, cytokines, trophic factors, neuropeptides, glutathione, and cholesterol, are integral players in: neuronal survival, blood flow control, synaptic transmission, learning and memory, and circuit integration[14].

Our lab is primarily interested in the effects of these released factors on neuronal survival. Astrocyte conditioned media (ACM) is necessary for neuronal survival[18] and many astroctye secreted factors are known to be neuroprotective. Putative neuroprotective factors expressed in astrocytes, include trophic factors, cytokines, neuorpeptides and small molecules. A partial list of these factors include: nerve growth factor (NGF), brain-derived growth factor (BDNF), glial cell line-derived neurotrophic factor (GDNF), fibroblast growth factor-2 (FGF-2), interleukin-6 (IL-6), transforming growth factor-β (TGF-β), S100β, activity-dependent neurotrophic factor (ADNF), and insulin-like growth factor (IGF).

In many circumstances, neuron-to-astrocyte communication is critical in regulating astrocytic secretion of cytokines and neurotrophic factors. An example of this paradigm is the activation of astrocytic neuroprotective mechanisms by the neuronal release of vasoactive intestinal peptide (VIP) and pituitary adenylate-cyclase activating peptide (PACAP)[19]. Under oxidative duress, neuronal release VIP induces astrocyte secretion of two neuroprotective species: activity-dependent neurotrophic factor (ADNF) and activity-dependent neuroprotective peptide (ADNP)[20, 21]. Both of these species enhance neuronal survival in vivo and in vitro[22]. Neuronal release of PACAP induces a similar response from astrocytes, causing the release of RANTES, macrophage inflammatory protein-6, and IL-6, all of which have neuroprotective properties[23-25].

In addition to the neuroprotective effects of the larger molecular weight species mentioned above, astrocyte-derived glutathione is also thought to play an important role in neuroprotection[6, 8, 26]. It is not fully understood how the astroctye-specific synthesis and secretion of glutathione enhances neuronal survival but It is hypothesized that either glutathione protects neurons via extracellular detoxification of reactive oxygen species or that neurons uptake the Cys-Gly byproduct of glutathione breakdown for *de novo* synthesis

of glutathione[27, 28]. Certainly, glutathione plays a role in the neuroprotective effects of astrocytes overexpressing Nrf2; activation of the ARE by Nrf2 potently upregulates expression of the cellular machinery responsible for the synthesis of glutathione[5-7]. However, more recent findings in our lab indicate neuroprotection only occurs after a long incubation period (1+ days) with Nrf2 astrocyte conditioned media (Nrf2-ACM) (unpublished data). Extracellular glutathione-mediated detoxification and Cys-Gly neuronal uptake occur on a timescale of hours (<4 hours)[27, 28]. This observation indicates a more complicated neuroprotective mechanism than glutathione-mediated extracellular detoxification or Cys-Gly uptake by neurons. **We hypothesize that other factors beyond glutathione contribute largely to the neuroprotective effects of ARE induction and that astrocyte-to-neuron signaling is a vital element of Nrf2 mediated neuroprotection.**

Molecular Weight Fractionation: Fractionation based on size can be achieved using ultrafiltration devices (molecular weight cut-off tubes (MWCO)). In contrast to SEC, MWCO tubes concentrate samples instead of dilute them, a highly desirable trait for an initial purification step. MWCO tubes can also be a quick, easy way to assess the general identity of a biologically active molecule. Using MWCO tubes as the initial fractionation step is crude but can be quite effective in directing downstream separations. We have concentrated astrocyte conditioned media from Nrf2 KO, WT, and GFAP-Nrf2 cultures. This concentration step is performed using a 2 kD MWCO tube. The subsequent fractions, the retentate (>2 kD) and the flow through (< 2 kD), were subjected to a neuroprotection assay. The retentate showed protection while the flow-through did not, indicating a factor larger than 2 kD being the primary agent of neuroprotection, most likely a neuropeptide or protein. Additionally, the GFAP-Nrf2 conditioned media exhibited better protection than the

WT or KO media. **The neuroprotective agent(s) is larger than 2 kD and is more highly secreted by the GFAP-Nrf2 cultured astrocytes.**

Mass Spectrometry-based Expression Proteomics: MS-based expression proteomics is a relatively new technique. With the advent of electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) methods, the ability to directly identify proteins and peptides from mass spectrometry (MS) became possible[29]. This significantly enhanced both the throughput and sensitivity of protein characterization, rendering traditional sequencing techniques, such as Edman degradation, virtually obsolete. More recently, the development of mass spectrometry-based methods for protein expression studies has emerged[30]. While the technology is not as mature, robust, or high throughput as DNA microarray technologies, proteomic expression studies have significant advantages over microarrays: direct interrogation of protein expression and identification of post-translational modifications, such as phosphorylation [31]. **MS-based proteomics is particularly useful in the analysis of secreted factors; the amount of secreted protein can be radically different than the expression levels indicated by mRNA.** However, in most cases DNA microarray and proteomic technologies are complementary techniques and when used in tandem can yield powerful data sets.

Generally, ESI- and MALDI-MS instruments used for proteomics yield two types of information: the mass of the peptide and the amino acid sequence of the peptide[32]. The mass is determined by calculating the mass-to-charge ratio and then converting this into an actual mass. The amino acid sequence of the peptide is determined by fragmenting the peptide. Peptides fragment in a very predictable manner which allows the reading of the amino acid sequence from the resultant fragment masses. Most MS-based proteomics experiments rely

on what's called a 'bottom-up' approach in which the proteins are first enzymatically digested (usually with trypsin) into peptides and then the protein's identity is inferred from the peptides' sequences matched to a protein database[32]. The reason for this 'bottom-up' method is due to the technical limitations of the instrumentation: the majority of currently available mass spectrometers have a limited mass range (5-9kDa and below). However, with recent advances in instrumentation, 'top-down' strategies, in which the protein mass and sequence are determined directly, have been developed[33]. Top-down experiments require highly specialized mass spectrometers (Fourier Transform (FT) instruments) which are capable of very high mass accuracy and resolution. These FT instruments are extremely expensive and relatively rare, although they are becoming more commonplace as the technology matures and the instrument prices drop.

The ability to interrogate protein expression directly with mass spectrometry is a powerful tool. However, many experiments employing expression proteomics are overambitious and attempt to analyze the expression patterns of the entire proteome. The massive amounts of data created in these experiments are incredibly time consuming to analyze and are difficult to interpret. **Our experimental design utilizes expression analysis with a very specific goal. The focused nature of our experiment should yield small and easily interpretable data sets.**

7.3 Preliminary Data:

Microarray Data: A number of differentially expressed secreted factors in cultures treated with tBHQ or infected with ad-Nrf2 were previously identified by our group using microarray analysis[6, 7]. These include glutathione-related genes, chromogranin B, superoxide dismutase, tumor necrosis factor (TNF), insulin-like growth factor binding protein (IGBP),

and tachykinin. Glutathione is known to be neuroprotective and its regulation by the ARE is well documented (see Background for discussion) but the significance of the other differentially expressed factors remains unknown. Both chromogranin B and superoxide dismutase have been implicated in ALS. TNF-α has a putative role in learning and memory[34-36] as well as conflicting putative effects as a neuroprotective and neurotoxic agent[37, 38]. Insulin-like growth factor (IGF) which is regulated by IGBP has been shown to be neuroprotective by a number of different studies[39-41].

MS Data: Using previously harvested ACM, we were able to identify a number of known astrocyte secreted factors by mass spectrometry as a proof of principle. Comparing our results to a previous proteomic study of astrocyte secreted factors, we identified a significant number of secreted factors which corroborated the identifications made by Lafon-Cazal et al.[42], including: macrophage stimulating colony factor (MCSF), IGF, IGBP, superoxide dismutase, fibroblast growth factor (FGF-9), stathmin 3, apolporptein (ApoE), metalloproteinase inhibitor 1&2 (TIM-1,2), and neurophil gelatinase-associated protein (NGAL). In the microarray data, both IGBP and superoxide dismutase were up-regulated in Nrf2 overexpressing cultures. **The identification of these factors by mass spectrometry is a proof of principle and indicates that we can identify secreted protein factors with our current cell culture system and MS instrumentation.**

7.4 References

1. Mattson, M. P. *Nat. Rev. Mol. Cell Biol.* **2000**, *2,* 120-129.

2. Rushmore, T. H.; King, R. G.; Paulson, K. E.; Pickett, C. B. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *10,* 3826-3830.

3. Rushmore, T. H.; Morton, M. R.; Pickett, C. B. *J. Biol. Chem.* **1991**, *18,* 11632-11639.

4. Moi, P.; Chan, K.; Asunis, I.; Cao, A.; Kan, Y. W. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *21,* 9926-9930.

5. Li, J.; Lee, J. M.; Johnson, J. A. *J. Biol. Chem.* **2002**, *1,* 388-394.

6. Shih, A. Y.; Johnson, D. A.; Wong, G.; Kraft, A. D.; Jiang, L.; Erb, H.; Johnson, J. A.; Murphy, T. H. *J. Neurosci.* **2003**, *8,* 3394-3406.

7. Kraft, A. D.; Johnson, D. A.; Johnson, J. A. *J. Neurosci.* **2004**, *5,* 1101-1112.

8. Chen, Y.; Vartiainen, N. E.; Ying, W.; Chan, P. H.; Koistinaho, J.; Swanson, R. A. *J. Neurochem.* **2001**, *6,* 1601-1610.

9. Spires, T. L.; Hannan, A. J. *FEBS J.* **2005**, *10,* 2347-2361.

10. Chen, X. L.; Kunsch, C. *Curr. Pharm. Des.* **2004**, *8,* 879-891.

11. Wakabayashi, N.; Dinkova-Kostova, A. T.; Holtzclaw, W. D.; Kang, M. I.; Kobayashi, A.; Yamamoto, M.; Kensler, T. W.; Talalay, P. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *7,* 2040-2045.

12. Lee, J. M.; Calkins, M. J.; Chan, K.; Kan, Y. W.; Johnson, J. A. *J. Biol. Chem.* **2003**, *14,* 12029-12038.

13. Volterra, A.; Meldolesi, J. *Nat. Rev. Neurosci.* **2005**, *8,* 626-640.

14. Haydon, P. G. *Nat. Rev. Neurosci.* **2001**, *3,* 185-193.

15. Haydon, P. G.; Carmignoto, G. *Physiol. Rev.* **2006**, *3,* 1009-1031.

16. Bezzi, P.; Gundersen, V.; Galbete, J. L.; Seifert, G.; Steinhauser, C.; Pilati, E.; Volterra, A. *Nat. Neurosci.* **2004**, *6,* 613-620.

17. Krzan, M.; Stenovec, M.; Kreft, M.; Pangrsic, T.; Grilc, S.; Haydon, P. G.; Zorec, R. *J. Neurosci.* **2003**, *5,* 1580-1583.

18. Takuma, K.; Baba, A.; Matsuda, T. *Prog. Neurobiol.* **2004**, *2,* 111-127.

19. Shioda, S.; Ohtaki, H.; Nakamachi, T.; Dohi, K.; Watanabe, J.; Nakajo, S.; Arata, S.; Kitamura, S.; Okuda, H.; Takenoya, F.; Kitamura, Y. *Ann. N. Y. Acad. Sci.* **2006**, , 550-560.

20. Brenneman, D. E.; Gozes, I. *J. Clin. Invest.* **1996**, *10,* 2299-2307.

21. Gozes, I.; Bassan, M.; Zamostiano, R.; Pinhasov, A.; Davidson, A.; Giladi, E.; Perl, O.; Glazner, G. W.; Brenneman, D. E. *Ann. N. Y. Acad. Sci.* **1999**, , 125-135.

22. Gozes, I.; Brenneman, D. E. *J. Mol. Neurosci.* **2000**, *1-2,* 61-68.

23. Vaudry, D.; Pamantung, T. F.; Basille, M.; Rousselle, C.; Fournier, A.; Vaudry, H.; Beauvillain, J. C.; Gonzalez, B. J. *Eur. J. Neurosci.* **2002**, *9,* 1451-1460.

24. Silveira, M. S.; Costa, M. R.; Bozza, M.; Linden, R. *J. Biol. Chem.* **2002**, *18,* 16075-16080.

25. Dejda, A.; Sokolowska, P.; Nowak, J. Z. *Pharmacol. Rep.* **2005**, *3,* 307-320.

26. Drukarch, B.; Schepens, E.; Jongenelen, C. A.; Stoof, J. C.; Langeveld, C. H. *Brain Res.* **1997**, *1-2,* 123-130.

27. Dringen, R.; Pfeiffer, B.; Hamprecht, B. *J. Neurosci.* **1999**, *2,* 562-569.

28. Drukarch, B.; Schepens, E.; Stoof, J. C.; Langeveld, C. H.; Van Muiswinkel, F. L. *Free Radic. Biol. Med.* **1998**, *2,* 217-220.

29. Aebersold, R.; Mann, M. *Nature* **2003**, *6928,* 198-207.

30. Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *5,* 252-262.

31. Kalume, D. E.; Molina, H.; Pandey, A. *Curr. Opin. Chem. Biol.* **2003**, *1,* 64-69.

32. McDonald, W. H.; Yates, J. R.,3rd *Curr. Opin. Mol. Ther.* **2003**, *3,* 302-309.

33. Kelleher, N. L. *Anal. Chem.* **2004**, *11,* 197A-203A.

34. Beattie, E. C.; Stellwagen, D.; Morishita, W.; Bresnahan, J. C.; Ha, B. K.; Von Zastrow, M.; Beattie, M. S.; Malenka, R. C. *Science* **2002**, *5563,* 2282-2285.

35. Stellwagen, D.; Beattie, E. C.; Seo, J. Y.; Malenka, R. C. *J. Neurosci.* **2005**, *12,* 3219-3228.

36. Stellwagen, D.; Malenka, R. C. *Nature* **2006**, *7087,* 1054-1059.

37. Turrin, N. P.; Rivest, S. *J. Neurosci.* **2006**, *1,* 143-151.

38. Zou, J. Y.; Crews, F. T. *Brain Res.* **2005**, *1-2,* 11-24.

39. Digicaylioglu, M.; Garden, G.; Timberlake, S.; Fletcher, L.; Lipton, S. A. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *26,* 9855-9860.

40. Torres Aleman, I. *Adv. Exp. Med. Biol.* **2005**, , 243-258.

41. Vincent, A. M.; Feldman, E. L.; Song, D. K.; Jung, V.; Schild, A.; Zhang, W.; Imperiale, M. J.; Boulis, N. M. *Neuromolecular Med.* **2004**, *2-3,* 79-85.

42. Lafon-Cazal, M.; Adjali, O.; Galeotti, N.; Poncet, J.; Jouin, P.; Homburger, V.; Bockaert, J.; Marin, P. *J. Biol. Chem.* **2003**, *27,* 24438-24448.