

## ABSTRACT

### Studies on Bovine $\gamma$ -Glutamylamine Cyclotransferase

Maryuri Roca

Mentor: Mary Lynn Trawick, Ph.D.

The purification and study of proteins are cooperative processes because at least partially purified protein is needed in order to study its properties, and certain information about the protein's properties is required in order to design its purification. Particularly difficult to purify is  $\gamma$ -glutamylamine cyclotransferase ( $\gamma$ GACT) which catalyzes the cyclization of the  $\gamma$ -glutamyl moiety in L- $\gamma$ -glutamylamines, notably N <sup>$\epsilon$</sup> -( $\gamma$ -glutamyl)lysine. From this last activity the function of the enzyme is speculated to be related to the catabolism of transglutaminase products; although, there is no direct evidence of this.

Electrophoretically pure bovine  $\gamma$ GACT was obtained using preparative ultracentrifugation, anion exchange chromatography on DEAE-Sepharose, ammonium sulfate fractionation and precipitation, size exclusion chromatography on Sephacryl S100, anion exchange chromatography on Mono-Q under reducing conditions, isoelectric focusing of the alkylated

sample, electroelution, electrophoresis, ultrafiltration, and lyophilization. The enzyme was purified more than 2,000 fold to a specific activity of more than 1,300U/mg of enzyme. A monomeric enzyme of molecular mass of 22,000 Daltons was observed. Anion exchange chromatography on a Mono Q GL column revealed two forms of the enzyme with pIs of 6.86 and 6.62 under non-reducing conditions, and a single form of pI 6.62 under reducing conditions.  $\gamma$ GACT was then subjected to analytical isoelectric focusing and the active fraction appeared as a single band on SDS-PAGE.

Amino acid sequencing of the tryptic digest of the band from SDS-PAGE corresponding to the enzyme was carried out by microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry; 42 proteins and protein fragments of similar mass and pI as that of  $\gamma$ GACT were obtained. Analysis of their properties indicates that the unknown protein for MGC:134378 is the most likely protein to be the bovine  $\gamma$ GACT enzyme. However, expression of the active enzyme from the cloned gene has to be done in order to assure that this is indeed the  $\gamma$ GACT enzyme.

An affinity column based on the inhibitor glutarylhexylamine showed binding of the enzyme when loaded at low ionic strength and elution when salt was increased.

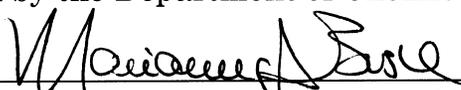
Studies on Bovine  $\gamma$ -Glutamylamine Cyclotransferase

by

Maryuri Roca, B.S.

A Dissertation

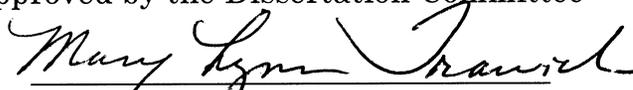
Approved by the Department of Chemistry and Biochemistry



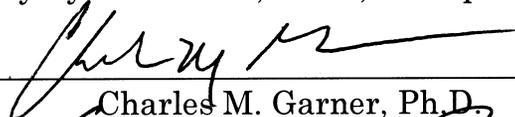
Marianna A. Busch, Ph.D. Chairperson

Submitted to the Graduate Faculty of  
Baylor University in Partial Fulfillment of the  
Requirements for the Degree  
of  
Doctor of Philosophy

Approved by the Dissertation Committee



Mary Lynn Trawick, Ph.D., Chairperson



Charles M. Garner, Ph.D.



Gouri S. Jas, Ph.D.

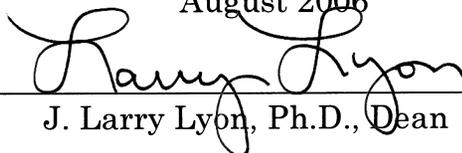


Robert R. Kane, Ph.D.



Erich J. Baker, Ph.D.

Accepted by the Graduate School  
August 2006



J. Larry Lyon, Ph.D., Dean

Copyright © 2006 by Maryuri Roca

All rights reserved

## TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xv
ACKNOWLEDGMENTS	xvii

### PART I

#### GENERAL BACKGROUND

CHAPTER ONE	2
$\gamma$ -GLUTAMYLAMINE CYCLOTRANSFERASE	2
$\gamma$ -Glutamylamino Acid Cyclotransferase ( $\gamma$ GAACT)	2
$\gamma$ -Glutamylamine Cyclotransferase ( $\gamma$ GACT)	3
Cyclotransferase Product, 5-Oxoproline	4
Transglutaminases and N <sup><math>\epsilon</math></sup> -( $\gamma$ -Glutamyl)lysine	7
Substrate Specificity of $\gamma$ GACT	10
General Characteristics of Cyclotransferases and Related Enzymes	15
CHAPTER TWO	30
PURIFICATION	30
Purification Considerations	30
Native vs Recombinant Protein Purification	31
Tracking $\gamma$ GACT During Purification	32
$\gamma$ GACT Purification Procedures	32
Considerations for the Purification of Unstable Proteins	34
Chromatography Improvements	36
Isoelectric Focusing	39

CHAPTER THREE	41
SEQUENCING OF PROTEINS	41
Cyclotransferase and Genetics	41
From the Protein to Its Gene	41
Sequencing Techniques	43
Gene Search	47
CHAPTER FOUR	50
DESIGN OF AN AFFINITY COLUMN FOR $\gamma$ GACT BASED ON THE	50
GLUTARYLHEXYLAMINE INHIBITOR	50
Affinity Chromatography	50
Reversible Inhibition	51
$\gamma$ GACT Inhibitors	56
Glutarylhexylamine Affinity Chromatography	
Column	57
Purification of $\gamma$ GACT in Glutarylhexylamine	
Sephrose 4B	58
CHAPTER FIVE	60
STATEMENT OF PURPOSE	60
PART II	
EXPERIMENTAL WORK	
CHAPTER SIX	65
EXPERIMENTAL PROCEDURE FOR $\gamma$ GACT PURIFICATION	65
Tracking of Enzymes during Purification	65
Bradford Assay for Total Protein Concentration	71
Gross Purification	73
Intermediate Purification	75
Intermediate-Polishing Purification Studies on	
Superdex HR 75 SEC	76
Intermediate-Polishing Purification Studies on	
Mono-Q IEC	77
Isoelectric Focusing	82

CHAPTER SEVEN	86
EXPERIMENTAL PROCEDURE FOR AMINO ACID SEQUENCING	86
N-Terminal Determination	86
Preparation of the Sample for MS Sequencing	90
MS Sequencing by Harvard Microchemistry and Proteomics Analysis Facility	91
Analysis of Reported Proteins	92
CHAPTER EIGHT	94
EXPERIMENTAL PROCEDURE FOR THE AFFINITY COLUMN DESIGN	94
Determination of Ligand Inhibition Constant	94
Synthesis of Glutarylhexylamine Affinity Column	96
Purification of $\gamma$ GACT in Glutarylhexylamine Sephacrose 4B	101
PART III	
RESULTS AND DISCUSSION	
CHAPTER NINE	104
RESULTS AND DISCUSSION OF THE PURIFICATION OF $\gamma$ GACT	104
Results of the Gross Purification	104
Results of the Polishing Purification Studies on Superdex S75	108
Results from the Intermediate-Polishing Purification Studies on Mono Q GL	111
Results from Isoelectric Focusing	129
Summary of $\gamma$ GACT Purification	132
CHAPTER TEN	135
RESULTS AND DISCUSSION OF THE SEQUENCING	135
N-Terminal Determination	135
MS Sequencing	138
Analysis of Candidate Proteins	138

CHAPTER ELEVEN	157
RESULTS AND DISCUSSION OF AFFINITY COLUMN DESIGN	157
Determination of Ligand Inhibition Constant	157
Results of the Two Coupling Methods	157
Ligand Leaching under Different Conditions	161
Results of the Purification of $\gamma$ GACToin Glutarylhexylamine Sepharose 4B	162
CHAPTER TWELVE	166
DISCUSSION AND CONCLUSIONS	166
Purification	166
Sequencing	172
Affinity Column	174
Future Work	176
APPENDICES	178
Appendix A	179
CHARACTERISTICS AND USE OF CHROMATOGRAPHIC COLUMNS	179
DEAE Sepharose Fast Flow Anion Exchange Column	180
Sephacryl S100 Size Exclusion Column	182
Superdex HR 75HR 10/30 Size Exclusion Column	185
Mono Q GL 5/50 Anion Exchange Column	188
EAH Sepharose 4B, Affinity Chromatography Solid Support	190
Appendix B	191
SUMMARY OF PURIFICATION OF $\gamma$ GACT FROM BOVINE KIDNEY	191
Appendix C	201
SEQUENCING REPORT FROM HARVARD MICROCHEMISTRY AND PROTEOMIC ANALYSIS FACILITY	201
REFERENCES	222

## LIST OF FIGURES

Figure 1.1: $\gamma$ -Glutamyl cycle. $\gamma$ -Glutamyl transpeptidase (1), $\gamma$ -glutamyl cyclotransferase (2), 5-oxoprolinase (3), $\gamma$ -glutamylcysteine synthetase (4), glutathione synthetase (5), intracellular protease (6)	6
Figure 1.2: Proposed mechanism of $\gamma$ GACT action (Gonzalez, 2005)	17
Figure 1.3: Crystal structure of glutaminyl cyclase from A) human (Huang et al., 2005) secondary structure obtained using the Swiss View 3.7 and B) plant (Wintjens et al., 2006)	19
Figure 1.4: Proposed mechanism for glutamyl cyclization by mammalian QC (Huang et al., 2005)	20
Figure 1.5: Proposed interactions of active site amino acids of plant QC and substrate (Wintjens et al., 2006)	21
Figure 1.6: Possible binding of $\gamma$ -glutamyl substrate by $\gamma$ GTP (Taniguchi ad Ikeda, 1998)	23
Figure 1.7: Crystal structure of $\gamma$ GTP dimer with glutamate substrate (Structure 2DG5) (From Okada et al., 2006)	23
Figure 1.8: Crystal structure of Blood Coagulation Factor XIII transglutaminase (monomer) (Structure 1GGU) (Fox et al., 2000)	24
Figure 1.9: Proposed mechanism for Factor XIII transglutaminase reaction. (Pedersen et al., 1994)	25
Figure 1.10: Crystal structure of pyroglutamidase (Structure 1AUG) (Odagaki et. al., 1999)	26
Figure 2.1: Organic osmolytes (Galinski, 1995)	36
Figure 3.1: Transcription and translation are not conservative processes	42
Figure 3.2: Edman Degradation reaction	44

Figure 3.3: Mass spectrum of [Glu <sup>1</sup> ] fibrinopeptide B, used as standard for the verification of MS/MS performance; the masses of the C-terminal fragments are shown in the top sequence (Yates, 1996)	47
Figure 4.1: Separation by affinity chromatography	51
Figure 4.2: General scheme for inhibition; where, E = Enzyme, S= substrate, I = inhibitor, P= product, and EI and ES are the complexes enzyme-inhibitor and enzyme-substrate, respectively.	52
Figure 4.3: A) Michaelis-Menten plot, B) Lineweaver-Burke plot	55
Figure 4.4: A) $K_{app}$ vs inhibitor concentration plot for the determination of $K_i$ for competitive inhibition	55
Figure 4.5: Coupling reaction for the production of a glutaryl-hexylamine affinity column for $\gamma$ GACT	59
Figure 6.1: Sketch of the results in TLC	67
Figure 7.1: Dansylation and hydrolysis of the N-terminal primary amine	88
Figure 7.2: Scheme of the MS sequencing of $\gamma$ GACT	92
Figure 8.1: Glutaric acid and hexylamine coupling employing a carbodiimide	99
Figure 8.2: Glutaryl anhydride and hexylamine coupling	99
Figure 8.3: Possible crosslink in affinity column synthesis	100
Figure 9.1: Typical results from DEAE-Sepharose. A) Elution profile of $\gamma$ GACT purification on DEAE-Sepharose. B) TLC assay	105
Figure 9.2: A) Elution profiles of activity eluted from Sephacryl S100.. B) Electrophoresis of the fractions from Sephacryl S100; molecular masses of the bands are shown.	107
Figure 9.3: Effect of reducing agent on the elution profile of sample in Superdex S75. A) non reducing conditions, B) reducing conditions.	110

Figure 9.4: SDS-PAGE of the Superdex HR 75 fractions; molecular mass of the bands are shown.	111
Figure 9.5: Elution profile of sample run in Mono Q GL with 2 mM Phosphate buffer pH 7.5. ** conductivity, ___ absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- specific activity	113
Figure 9.6: Elution profile of sample run in Mono Q GL with 0.5 mM Phosphate buffer pH 7.5. ** conductivity, ___ sample absorption at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- specific activity	114
Figure 9.7: Effect of Tris buffer on the activity, protein measurement and specific activity, in comparison to KPi. $\blacklozenge$ 4.3 U/mL, $\blacksquare$ 22 U/mL, $\blacktriangle$ 39U/mL	116
Figure 9.8: Elution profile of $\gamma\text{GACT}$ sample on Mono run at recommended conditions; ** conductivity, ___ absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- specific activity	120
Figure 9.9: Elution profile of $\gamma\text{GACT}$ sample on Mono Q GL, low concentration of Tris, shallow gradient; without reducing agent (Condition #1). ** Conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ )	121
Figure 9.10: Elution profile of $\gamma\text{GACT}$ sample on Mono Q GL; at high concentration of Tris, shallow gradient; without reducing agent (condition #2). **Conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- Specific activity	122
Figure 9.11: Elution profile of $\gamma\text{GACT}$ sample on Mono Q GL at high concentration of Tris, steeper gradient; without reducing agent (Condition #3).** Conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- Specific activity	123
Figure 9.12: Elution profile of $\gamma\text{GACT}$ sample on Mono Q GL at high concentration of Tris, steeper gradient; with reducing agent (condition #4). ** Conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- Specific activity	124

Figure 9.13: SDS-PAGE of Mono Q GL fractions. A) peak of pI 6.9 at condition #2 and B) peak of pI 6.6 condition #4	125
Figure 9.14: Optimization of $\gamma$ GACT purification on Mono Q GL column	125
Figure 9.15: Elution profile of $\gamma$ GACT sample on Mono Q GL with EGTA; ** conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h), -+- Specific activity	127
Figure 9.16: Elution profile of $\gamma$ GACT alkylated sample on Mono Q GL; ** conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h), -+- Specific activity	128
Figure 9.17: Isoelectric focusing of the non alkylated sample (A) and Alkylated sample (B), with measured pH. The migration distances are measured with the scale shown at the left. Arrows indicate the position of the bands.	130
Figure 9.18: SDS-PAGE of the protein from IEF. Molecular masses of the bands are marked. Load is the sample loaded onto the IEF, ppt was the precipitated protein in water before IEF gel loading; the numbers correspond to the migration distances (cm) on the IEF gel	131
Figure 9.19: LogMW vs migration distance (cm). Calibration curve of SDS-PAGE gel for the measurement of the molecular mass of proteins.	131
Figure 9.20: Purification followed by SDS-PAGE; molecular mass of the bands are indicated	132
Figure 9.21: Scheme of $\gamma$ GACT purification	133
Figure 10.1: Dansyl amino acid standard resolved by HPLC, detected by fluorescence. A) 20 nmol of the indicated standards, B) 0.1 nmol of $\epsilon$ -dansyl-lysine standard	136
Figure 10.2: Elution profile of dansylated and hydrolyzed carbonic anhydrase standard protein	137
Figure 10.3: Elution profile of dansylated and hydrolyzed gel extracted $\gamma$ GACT	137

Figure 10.4: A and B) Crystal structure of PPIL3b with polypeptide HAGPILA (PDB code: 1AWQ from Vajdos et al., 1997). C) Predicted secondary structure of similar to PPIL3b protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)	144
Figure 10.5: Mechanism of methyl transfer by GAMT (Komoto et al., 2004)	146
Figure 10.6: A) Crystal structure of GMAT with substrates SAM and GAA (crystal structure 1XCJ) (Komoto et al., 2004); B) predicted secondary structure for similar to GAMT protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)	147
Figure 10.7: Predicted secondary structure of Umf1 protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)	148
Figure 10.8: A) Human Vh1-Related Dual-Specificity Phosphatase (crystal structure 1J4X) (Yuvaniyama et al., 1996). B) Predicted secondary structure of unknown protein for MGC :134378 obtained using the Swiss View modeler (Guex and Peitsch, 1997)	149
Figure 10.9: Mechanism of Dual-specificity phosphatase. The catalytic triad is the catalytic thiol (Cys124) general acid (Asp92) and hydroxyl group (Ser131) (Denu and Dixon, 1995)	150
Figure 10.10: Predicted secondary structure of Prostaglandin H2 D-isomerase complexed with carboxymycobactin T obtained using the Swiss View modeler (Guex and Peitsch, 1997)	151
Figure 10.11: Predicted secondary structure of unknown protein for MGC:134378 showing the unbridged thiol groups. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)	154
Figure 10.12: Predicted secondary structure of Similar to PPIL3b protein showing the probable active site formed by His84, Tyr89, and Lys88. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)	155

Figure 10.13: Predicted secondary structure of Similar to GAMT protein showing the probable active site formed by His144, Tyr136, and Lys180. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)	155
Figure 10.14: Predicted secondary structure of unknown protein for MGC :134378 showing the probable active site formed by His139, Lys105, Tyr94, and Tyr101. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)	156
Figure 11.1: A) Lineweaver-Burke plot with glurarylhexylamine as $\gamma$ GACT inhibitor, B) $K_{app}$ s versus [I] for the determination of $K_i$	158
Figure 11.2: Tracking of the coupling reaction using a carbodiimide. From left to right: Negative control, positive control, EAH Sepharose 4B, 1h coupling, 12 h coupling.	159
Figure 11.3: Scheme of the carbodimide reactions; 1) new peptide bond formation, 2) intermediate rearrangement into a stable N-acylurea	160
Figure 11.4: Coupling reaction using glutaryl anhydride. From left to right: Negative control, positive control, EAH Sepharose 4B, 1h coupling, 12 h coupling, 24h coupling	160
Figure 11.5: Temperature effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, 4, 15, 25, 35, and 45 °C	161
Figure 11.6: pH effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, pH 2, 5, 7, 9, and 12	161
Figure 11.7: Salt concentration effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, 0, 0.1, 0.25, 0.5, and 1 M NaCl	162
Figure 11.8: Weekend storage of GH Sepharose 4B stability. From left to right: Negative control, positive control, EAH Sepharose 4B, 20% methanol stored, and 20% methanol with 0.5 M NaCl stored	162

Figure 11.9: Elution profile of 13Units of $\gamma$ GACT separation in Glutarylhexyalmine Sepharose 4B. The arrow indicates the introduction of buffer with 1 M NaCl. ** conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h)	163
Figure 11.10: Elution profile of 3 Units of $\gamma$ GACT separation in Glutarylhexyalmine Sepharose 4B. The arrow indicates the introduction of buffer with 1 M NaCl. ** conductivity, ___ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h)	164
Figure A1: Calibration of Sephacryl S100	184
Figure A2: Calibration of Superdex S75	187
Figure A3: Calibration curve of Mono Q GL	189

## LIST OF TABLES

Table 1.1: Catalytic properties of cyclotransferases and related enzymes	28
Table 1.2: Characteristics of cyclotransferases and related enzymes	29
Table 6.1: Optimization of Mono Q GL parameters in Tris buffer	82
Table 7.1: HPLC Condition of Dansyl amino acids determination	89
Table 8.1: Concentration of substrate (N <sup>ε</sup> -(γ-glutamyl)lysine) and inhibitor (glutarylhexylamine) in the reaction for the determination of $K_i$	96
Table 9.1: Effect of reduction and alkylation on the enzyme activity	117
Table 9.2: Purification of γGACT, from 50 g of bovine kidney tissue	134
Table 10.1: Reported Proteins	139
Table 10.1: Reported Proteins (continuation)	140
Table 10.2: Candidate proteins summary	141
Table 10.3: Amino acid sequence of the candidate proteins	142
Table 12.1: Comparison of the results for the purification of 50 g of bovine enzyme	167
Table A1: DEAE Sepharose column characteristics (From Amersham)	180
Table A2: Sephacryl S100 column characteristics (From Amersham)	182
Table A3: Superdex HR 75column characteristics (From Amersham)	185
Table A4: Mono Q GL column characteristics (From Amersham)	188
Table A5: Results for the calibration standards of Mono Q GL	189
Table A6: EAH Sepharose 4B matrix characteristics (From Amersham)	190

## LIST OF ABBREVIATIONS

2-ME	2-Mercaptoethanol
BCCA	British Columbia Cancer Agency
BSA	Bovine Serum Albumin
DEAE	Diethylaminoethyl
DTT	DL- Dithiothreitol
EDC	N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid
ESI	Electrospray Ionization
FPLC	Fast Protein Liquid Chromatography
$\gamma$ GACT	$\gamma$ -Glutamylamine cyclotransferase
$\gamma$ GAACT	$\gamma$ -Glutamylamino acid cyclotransferase
$\gamma$ GTP	$\gamma$ -Glutamyl transpeptidase
GAA	Guanidinoacetate
GAMT	Guanidinoacetate N-methyltransferase
GDC	Glutamyl dansylcadaverine
GDP	Guanosine 5'-diphosphate
HPLC	High Performance Liquid Chromatography
IAA	Iodoacetamide
IEC	Ion Exchange Chromatography
IEF	Isoelectric Focusing

KPi	Potassium phosphate
MDC	Monodansylcadaverine
MES	4-Morpholineethane sulfonate
MGC	Mammalian Gene Collection
MS/MS	Mass spectrometry/Mass spectrometry
NaPi	Sodium phosphate
NCBI	National Center for Biothecnology Information
OPA	o-Phthalaldehyde
PCR	Polymerase Chain Reaction
PPIL3b	Peptidylpropyl isomerase-like protein 3
PTC	Phenylthiocarbamyl
PTH	Phenylthiohydantoin
QC	Glutaminy cyclase
RCSB PDB	Research Collaboratory for Structural Bioinformatics Protein Data Bank
SAM	S-adenosylmethionine
SAH	S-adenosylhomocysteine
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
SCOP	Structural Classification of Proteins
TCA	Trichloroacetic acid
TLC	Thin layer chromatography
Ufm1	Ubiquitin fold modifier-conjugated enzyme 1

## ACKNOWLEDGMENTS

I would like to thank Dr. Mary Lynn Trawick for all her teaching and constant motivation, support, patience, and supervision. I would like to thank my committee members Dr. Robert Kane, Dr. Charles Garner, and Dr. Gouri Jas not only for their willingness to review this work but also for the knowledge I gained in their classes. I especially want to thank Dr. Erich Baker for his contribution to this dissertation in working on the similarities studies and for donating his time to review my work.

I would like to thank the chemistry department and the graduate school professors and staff for all their collaboration. I would like to thank my fellow graduate students especially to Milenka, Sam, and John for their advice, friendship, and scientific support.

I want to thank Baylor University for supporting this research.

*To my family for all their love and support*

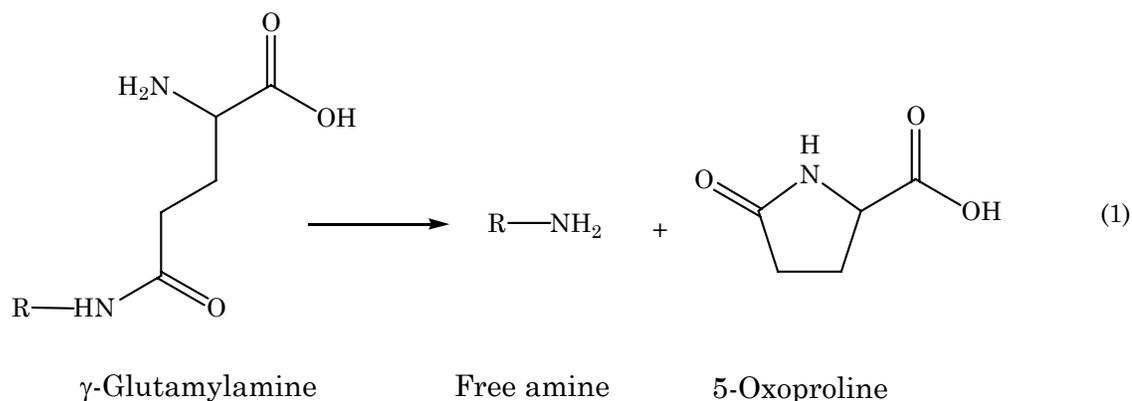
PART I

GENERAL BACKGROUND

## CHAPTER ONE

### $\gamma$ -Glutamylamine Cyclotransferase

$\gamma$ -Glutamyl cyclotransferases are enzymes that catalyze peptide bond cleavage between a  $\gamma$ -glutamyl and an amino portion not by hydrolysis but rather by cyclization of the glutamyl moiety of the substrate, releasing 5-oxoproline (pyroglutamic acid) and a free amino portion, as reaction 1 shows.

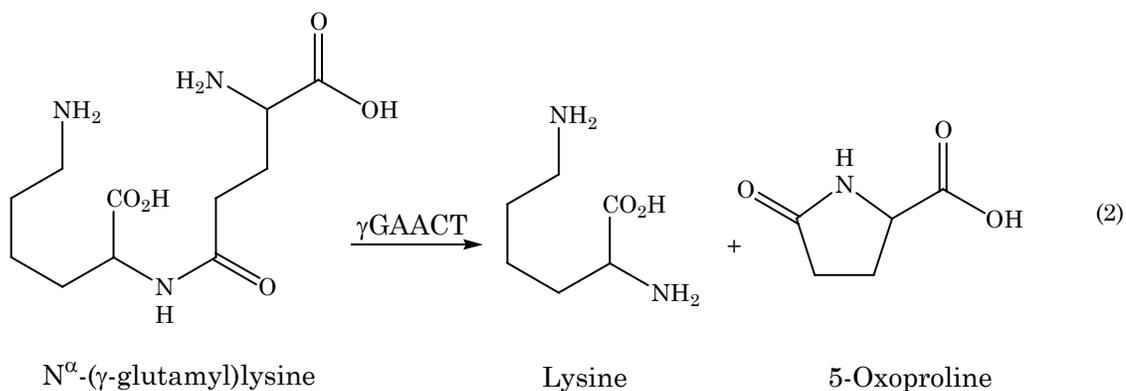


This enzymatic activity has been identified in a variety of tissues and organisms, like human and sheep brain (Orlowski et al., 1969), yeast (Mooz and Wigglesworth, 1976), and with different  $\gamma$ -glutamyl derivatives.

#### *$\gamma$ -Glutamylamino Acid Cyclotransferase ( $\gamma$ GAACT)*

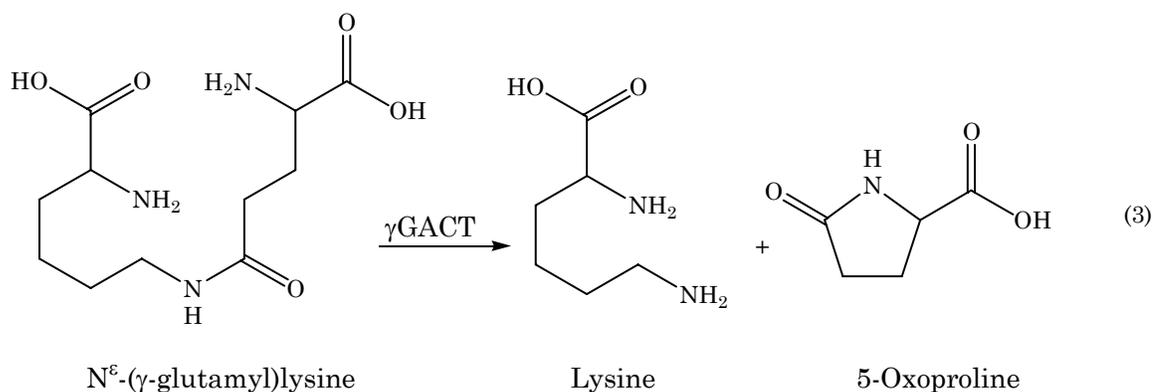
$\gamma$ -Glutamyl cyclotransferase activity was first identified in liver by Connell and Hanes (1956), where 5-oxoproline was the product of the reaction of  $\gamma$ -glutamylamino acid cyclotransferase ( $\gamma$ GAACT,  $\gamma$ -L-glutamyl-

cyclotransferase; L-glutamic cyclase, EC 2.3.2.4) acting on  $\gamma$ -glutamyl amino acid and producing free amino acid, reaction 2.



*$\gamma$ -Glutamylamine Cyclotransferase ( $\gamma$ GACT)*

In 1980, Fink and coworkers reported a  $\gamma$ -glutamyl cyclotransferase from rabbit kidney, and named it  $\gamma$ -glutamylamine cyclotransferase ( $\gamma$ GACT) based on its specificity toward the isopeptide  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine and other  $\gamma$ -glutamylamines; reaction 3 shows  $\gamma$ GACT catalysis.



$\gamma$ GACT activity was found in different tissues and organisms like kidney, liver, pancreas, mammalian blood cells (Fink et al., 1980), chinese

hamster ovary cells (Fesus and Tarcsa, 1989), and soybean seeds (Kang et al., 1997).

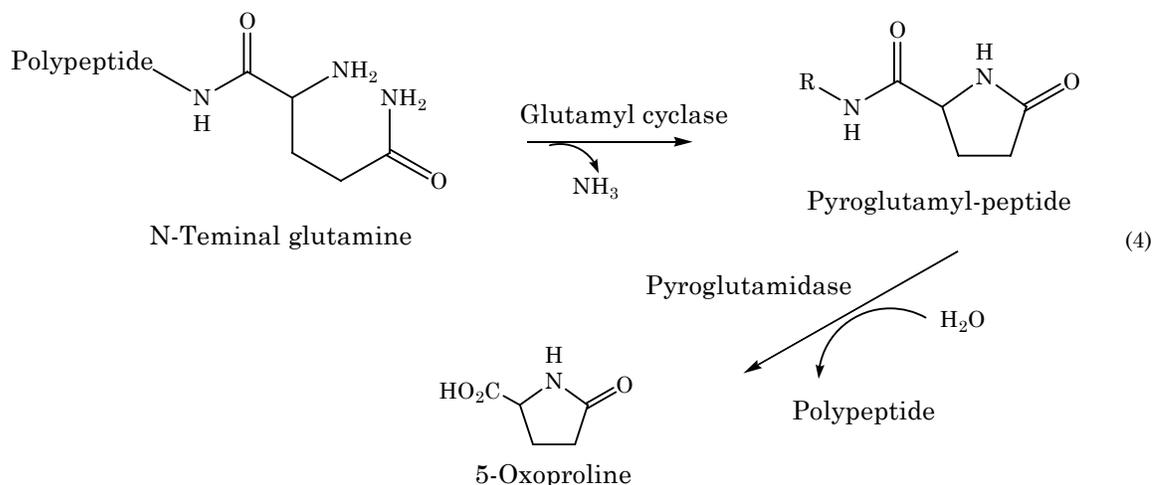
### *Cyclotransferase Product, 5-Oxoproline*

5-Oxoproline (or pyroglutamate) was first described as a derivate of glutamic acid lacking a water molecule. The cyclization of N-terminal glutamic acid was observed by Sanger and coworkers in 1955. 5-Oxoproline can be produced non-enzymatically from glutamate or glutamate derivatives (Van der Werf et al., 1971) and from spontaneous cyclization of glutamate-5-phosphate.

On the other hand, the enzymatic formation of 5-oxoproline was first observed in rat kidneys by Woodward and Reinhart (1942) as a product of the enzymatic degradation of glutathione; but only much later, cyclotransferase activity was identified in rat liver by Connell and Hanes (1956), and in papaya by Messer and Ottesen (1965).

In liver, 5-oxoproline was the product of the reaction of  $\gamma$ GAACT, but in papaya 5-oxoproline was the product of glutaminyl cyclase (QC, glutaminyl-peptide cyclotransferase, EC 2.3.2.5) acting on free glutamine and producing ammonia. 5-Oxoproline can also be formed by the complementary action of QC, which modifies N-terminal glutaminyl residues into 5-oxoprolyl ones (Busby et al., 1987; Fischer and Spiess, 1987), and pyroglutamidase (pyroglutamyl aminopeptidases, pyrased, EC 3.4.19.3), which hydrolyses 5-

oxoproline from these N-terminal modified polypeptides (Doolittle and Armentrout, 1968) as seen in reaction 4.



5-Oxoproline is an intermediate in the  $\gamma$ -glutamyl cycle (Meister, 1988), shown in figure 1.1. The  $\gamma$ -glutamyl cycle is one of the transport mechanisms for cell intake of amino acids.  $\gamma$ -Glutamyl transpeptidase ( $\gamma$ GTP, EC 2.3.2.2) exchanges  $\gamma$ -glutamyl from glutathione to an incoming amino acid forming a  $\gamma$ -glutamyl isodipeptide. The amino acid is released in the interior of the cell after cleavage by  $\gamma$ -glutamyl cyclotransferase, with the concomitant production of 5-oxoproline. 5-Oxoproline is hydrolyzed by 5-oxoprolinase (EC 3.5.2.9) to form glutamate.

At the N-terminus of proteins, 5-oxoprolinyl residues can occur not only to minimize degradation of the polypeptide but also to provide proteins with particular functions (Awade et al., 1994). A N-terminal pyroglutamyl residue is involved in the thermal stability and catalytic activity of frog cytotoxic ribonuclease by maintaining its structural integrity (Liao et al., 2003; Lou et

al., 2006); pyroglutamic acid derivative of melanoma vaccine “failed to elicit cytotoxic T lymphocyte activity” (Beck et al., 2001).

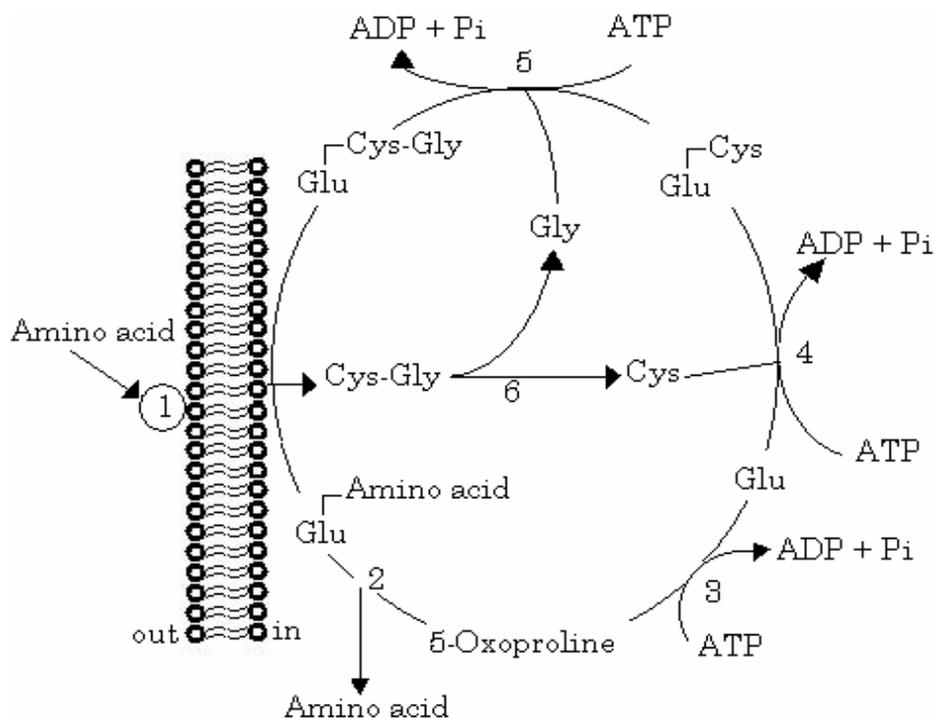
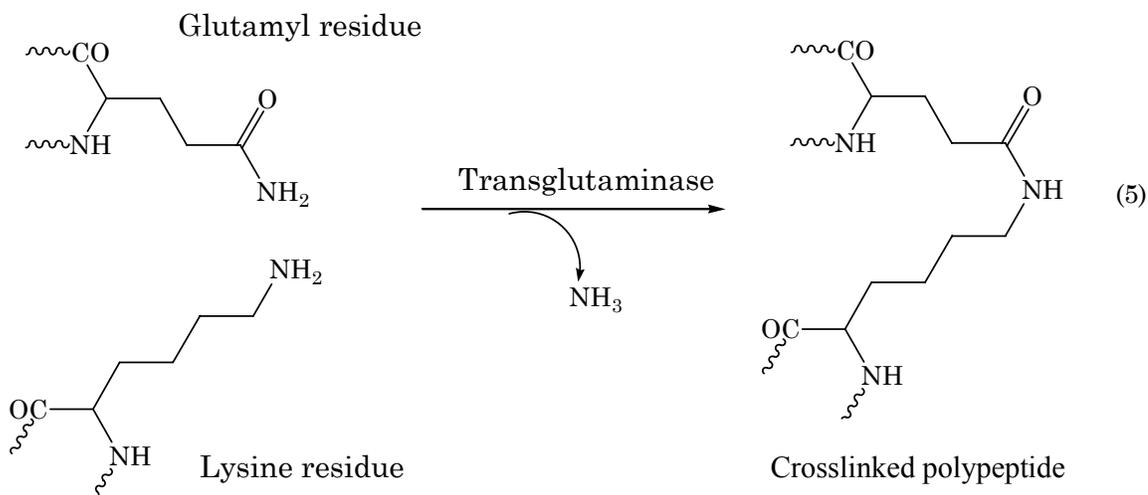


Figure 1.1:  $\gamma$ -Glutamyl cycle.  $\gamma$ -Glutamyl transpeptidase (1),  $\gamma$ -glutamyl cyclotransferase (2),  $\gamma$ -oxoprolinase (3),  $\gamma$ -glutamylcysteine synthetase (4), glutathione synthetase (5), intracellular protease (6)

The function of the enzymes has been proposed based on their substrate specificity.  $\gamma$ -GAAC shows activity with substrates that produce products that are good substrates of  $\gamma$ -glutamyl transpeptidase (Orlowski and Meister, 1973). This observation is consistent with the expected complementary functions of  $\gamma$ -GAAC and  $\gamma$ -GTP in the  $\gamma$ -glutamyl cycle

*Transglutaminases and N<sup>ε</sup>-(γ-Glutamyl)lysine*

γGACT has been hypothesized to play a role in the metabolism of the N<sup>ε</sup>-(γ-glutamyl)lysine crosslink found in several proteins. The crosslink is made by transglutaminase (EC 2.3.2.13) enzymes, as reaction 5 shows.



Transglutaminases are enzymes that catalyze posttranslational modification of proteins, forming the N<sup>ε</sup>-(γ-glutamyl)lysine crosslink and linkages of other γ-glutamyl derivatives with a variety of amines (Folk, 1980). The transglutaminase reaction goes through a calcium-dependent acyl transfer mechanism with the formation of a high energy thioester intermediate (Folk, 1973).

The action of these enzymes is found intracellularly and extracellularly in a number of tissues and important physiological processes like bone and skin formation, and wound healing (Raghunath et al., 1996), blood clotting and development of the heart, lung, saliva gland, and the central peripheral

nervous system (Lorand and Graham, 2003), cell survival (Fesus and Szondy, 2005), apoptosis marking (Fabbi et al., 1999), and cellular recognition (Akimov and Belkin, 2001). Their capabilities have been advantageous in the improvement of food (Gerrard and Sutton, 2005) and protein chemistry (Taki et al., 2004).

The isopeptide has been found intracellularly and extracellularly in prokaryotes, eukaryotes, birds, and mammals (Matacic and Loewy, 1979). It is essential for the function and structural integrity of a number of mammalian proteins; for example, Pisano and coworkers in 1968 found that the human cross linked fibrin, formed during blood clotting, is actually linked through a  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine bridge formed by the transglutaminase Factor XIIIa. Waibel and Carpenter (1972) observed that the isopeptide, which can substitute lysine in a lysine deficient diet in chicks and rats, is degraded *in vivo* in order to satisfy the lysine demand for growth of animals.

The degradation of the isopeptide is unlikely to happen by means of the reverse reaction of transglutaminases, since this reaction is slow and inefficient (Folk, 1969); Fesus and Tarcsa (1989) found that the radiolabeled isodipeptide has a half-life time of 3 hours in Chinese-hamster ovary cells grown in [ $^3\text{H}$ ]-lysine labeled media; it was also observed that such degradation can be prevented by  $\gamma$ -glutamyl dansylcadaverine and  $\gamma$ -glutamylputescine, both good substrates of  $\gamma$ GACT (Fink and Folk, 1981).

In the same experiment of Fesus and Tarcsa (1989), the lysosomal inhibitor methylamine prevented the appearance of the isopeptide and increased the protein-bond isopeptide, demonstrating that proteolysis is required for the liberation of the isopeptide and that the isopeptide is resistant to proteolysis; this resistance to hydrolysis was the basis of isolation of the isopeptide in 1968 by Pisano and coworkers from human fibrin, and by Maticic and Loewy from bovine fibrin. The result of all this is that despite the fact that the isopeptide is resistant to proteolysis by proteolytic enzymes the isodipeptide is degraded *in vivo*.

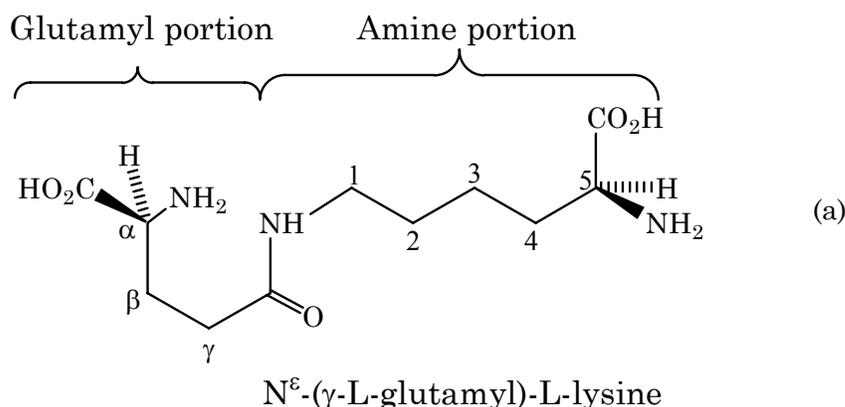
Deregulation of transglutaminases activity is associated with diseases, including neurodegenerative diseases, neoplastic diseases, autoimmune diseases, tissue fibrosis, and diseases related to the epidermis (Griffin et al., 2002). Also, the addition of transglutaminases to food is associated with inflammatory disorders in the intestine (coeliac response) (Gerrard and Sutton, 2005).

Lorand (1996) commented on the relation between neurodegenerative diseases and transglutaminases; he remarks the transglutaminase reaction with neuron filaments that can impact unfavorably the function of the neuron, the increase in reactivity of transglutaminases with repeated glutamine residues similar to those found in several neurodegenerative diseases, and the relation among  $\text{Ca}^{2+}$  pumps, transglutaminases, and Alzheimer and Huntington diseases. Nemes et al. (2001) reported that the

isopeptide is in high concentration in the cerebrospinal fluid of patients with both Alzheimer's and vascular type dementias; also, he found that aggregated proteins may branch not only by  $N^\epsilon$ -( $\alpha$ -glycyl)lysine or  $\alpha$ -glycyllysine but also by  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine crosslinks as the one found between ubiquitin and HSP27 or  $\alpha$ -synuclein (Nemes et al., 2004).

### *Substrate Specificity of $\gamma$ GACT*

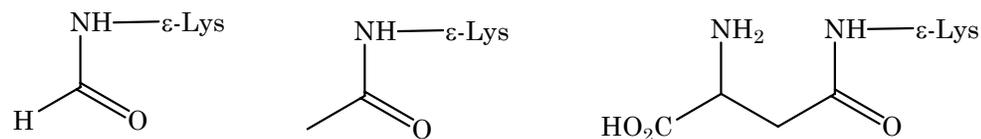
For rabbit  $\gamma$ GACT, Fink and coworkers (1980) found a  $K_M$  of  $0.23 \pm 0.1$  mM with  $N^\epsilon$ -( $\gamma$ -L-glutamyl)lysine, compound (a). A similar  $K_M$  value was obtained by Bowser in 1997, ( $K_M = 0.26 \pm 0.02$  mM).



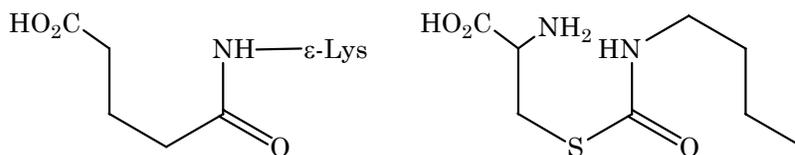
$\gamma$ GACT acts on L- $\gamma$ -glutamylamines and is highly specific for the L- $\gamma$ -glutamyl moiety, but its activity is not limited by the amide portion; this was concluded when the protein was tested toward various  $\gamma$ -glutamylamines and related compounds (Fink et al., 1980 and Fink and Folk, 1981).

$\gamma$ GACT does not show activity toward any substrate such as formyl, acetyl, aspartyl, glutaryl, or other  $\epsilon$ -lysine derivatives, compounds (b), (c), (d),

and (e), respectively. In contrast, Bowser found that  $\gamma$ -GACT displays high activity toward S-(n-butylcarbamyl)-L-cysteine, which is a substrate without the  $\gamma$ -glutamyl moiety, compound (f).

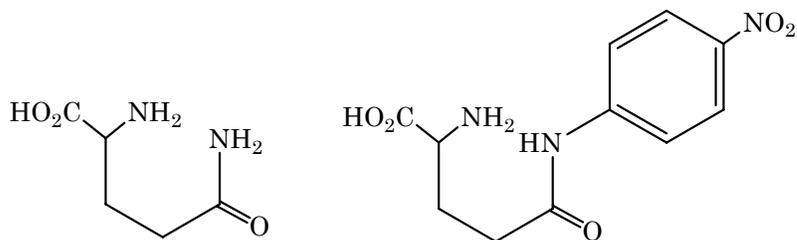


(b) N $\epsilon$ -Formyllysine    (c) N $\epsilon$ -Acetyllysine    (d) N $\epsilon$ -Aspartyllysine



(e) N $\epsilon$ -( $\gamma$ -Glutaryl)lysine    (f) S-(n-Butylcarbamyl)cysteine

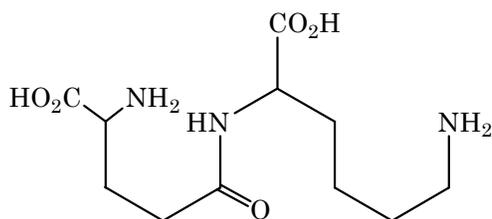
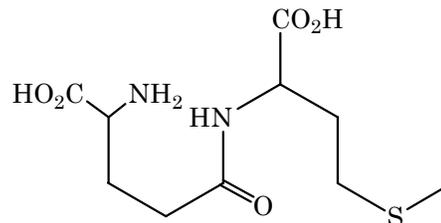
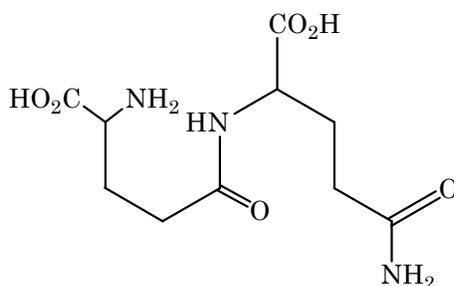
$\gamma$ GACT does not recognize as a substrate free glutamine, and  $\gamma$ -glutamyl-p-nitroanilide is a very poor substrate, compounds (g) and (h), showing specificity toward the  $\gamma$ -glutamylamine linkage.



(g) Glutamine

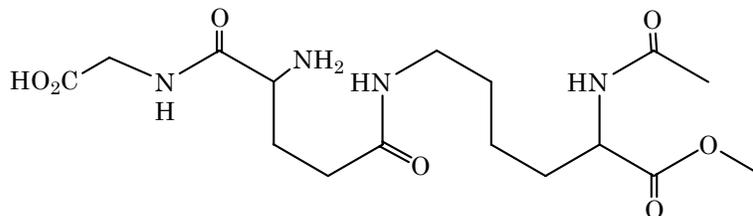
(h)  $\gamma$ -Glutamyl-p-nitroaniline

The following  $\gamma$ -glutamyl- $\alpha$ -amino acids did not show reaction with  $\gamma$ GACT:  $N^\alpha$ -( $\gamma$ -glutamyl)lysine,  $\gamma$ -glutamylmethionine, and  $\gamma$ -glutamylglutamine, compounds (i), (j), and (k), respectively.

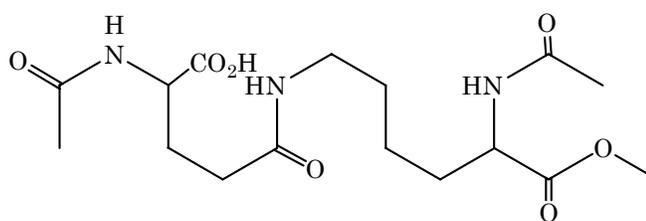
(i)  $N^\alpha$ -( $\gamma$ -Glutamyl)lysine(j)  $\gamma$ -Glutamylmethionine(k)  $\gamma$ -Glutamylglutamine

Additionally, the glutamyl  $\alpha$ -carboxylic and  $\alpha$ -amino groups should be free, as a requirement for cyclization; neither compound (l)  $N^\alpha$ -acetyl- $N^\epsilon$ -( $\gamma$ -glutamylglycine)lysine methyl ester, where the  $\alpha$ -carboxylic group of the glutamyl is in a linkage, nor  $N^\alpha$ -acetyl- $N^\epsilon$ -( $N$ -acetyl- $\gamma$ -glutamyl)lysine methyl ester, where the  $\alpha$ -amino group of  $\gamma$ -glutamyl is linked, compound (m), showed reaction with  $\gamma$ GACT. Quite the opposite, the enzyme shows activity toward  $N^\alpha$ -acetyl- $N^\epsilon$ -( $\gamma$ -glutamyl) lysine methyl ester, compound (n), that is a

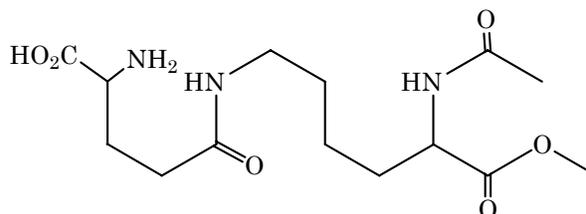
substrate with both amine and carboxylic groups of the amide portion blocked, which implies that the amino portion does not have to be free.



(l)  $N^{\alpha}$ -Acetyl- $N^{\epsilon}$ -( $\gamma$ -glutamylglycine)lysine methyl ester

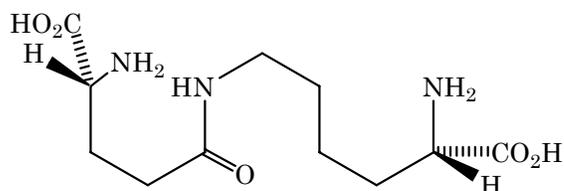
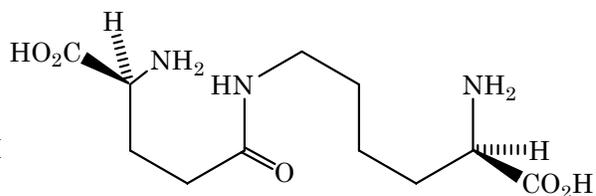


(m)  $N^{\alpha}$ -Acetyl- $N^{\epsilon}$ -( $N$ -acetyl- $\gamma$ -glutamyl)lysine methyl ester

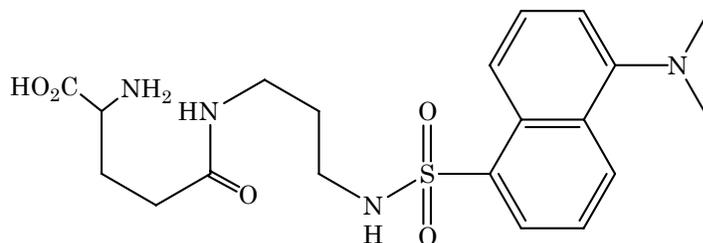
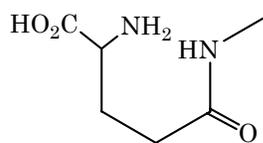
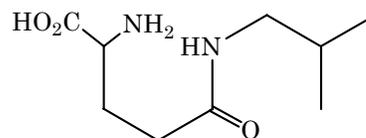


(n)  $N^{\alpha}$ -Acetyl- $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine methyl ester

Even more, the enzyme is highly stereospecific toward the L-glutamyl isomer, since D-glutamylamines, like  $N^{\epsilon}$ -( $\gamma$ -D-glutamyl)-L-lysine, compound (o), is not a substrate of  $\gamma$ GACT. But the enzyme is not stereospecific for the amide portion of the substrate because  $N^{\epsilon}$ -( $\gamma$ -L-glutamyl)-D-lysine, compound (p), works as an excellent substrate, as well as other compounds.

(o) N<sup>ε</sup>-(γ-D-Glutamyl)-L-lysine(p) N<sup>ε</sup>-(γ-L-Glutamyl)-D-lysine

$\gamma$ GACT specificity toward the amide portion of the substrate is less restrictive, since the enzyme shows activity toward  $\gamma$ -glutamylamines when the amide portion is dansylcadaverine, compound (q), methylamine, compound (r), isobutylamine, compound (s), and others.

(q)  $\gamma$ -Glutamyl dansylcadaverine(r)  $\gamma$ -Glutamylmethylamine(s)  $\gamma$ -Glutamylisobutylamine

Bowser (1997) concluded that  $\gamma$ GACT specificity is toward  $\gamma$ -glutamylamines with unbranched and extended alkylamide chains, which extend at least 4 carbons, and the first position (in compound (a), position 1) is not branched or occupied by a carboxylic group; consequently, the enzyme does not react with  $\gamma$ -glutamylamino acids.

The specificity of the enzyme toward the different portions of the substrate, and its significant inhibition by  $N^{\epsilon}$ -glutaryllysine, compound (e) and  $N^{\epsilon}$ -( $\gamma$ -D-glutamyl)-L-lysine, compound (o), allow the design of active-site inhibitors.

Inhibitors of  $\gamma$ GACT are required for the study of the enzyme activity, with the final objective of elucidation of  $\gamma$ GACT metabolic function in the organism.

### *General Characteristics of Cyclotransferases and Related Enzymes*

#### *$\gamma$ -Glutamylamine Cyclotransferase ( $\gamma$ GACT)*

From the work of Fink et al. (1980), Gowda (1985), and Bowser (1997) it is known that  $\gamma$ GACT is a soluble cytosolic protein with an optimum pH between 6.8 and 8.5 and is inactive at pH 4. Rabbit  $\gamma$ GACT was reported to have a molecular weight of  $\sim 25,000$  D, determined by size exclusion chromatography (Fink et al., 1980); while, Gowda (1985) found a dimeric molecule of 27,000 D and a pI of 6.15 for bovine  $\gamma$ GACT; she reported a  $K_M$  of the enzyme reaction with  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine of 0.385 mM.

Instability of  $\gamma$ GACT was observed during its purification by Bowser (1997); diluted samples of  $\gamma$ GACT were well behaved in Tris/HCl buffer, but concentrated enzyme precipitated in Tris with loss of activity.

Fink and Folk (1981), observed that inhibitors of serine proteases do not affect the activity of  $\gamma$ GACT, the enzyme has no esterase or glutaminase

activity, and activity is lost when the rabbit enzyme was treated with sulfhydryl modifiers 5,5'dithio-bis(2-nitrobenzoic acid) and 4-hydroxymercuribenzoic acid, but not with iodoacetamide, cysteamine, or N-ethylmaleimide. Gowda's (1985) treatment of bovine  $\gamma$ GACT with iodoacetamide gave a shift of activity from pI 6.15 to 7.10; she observed inactivation of the bovine enzyme by treatment with iodoacetamide; she also observed increase or decrease of activity by treatment with disulfide compounds that can form disulfides with sulfhydryl groups of the enzyme. This suggested the participation of sulfhydryl groups in the catalytic action of the enzyme. Physiological disulfides such as homocysteine, oxidized glutathione, and cysteamine, also tested by Gowda, may be the natural disulfide activators of  $\gamma$ GACT.

pH-dependence studies carried out by Gonzalez (2005) showed that ionizable groups in the active site of rabbit  $\gamma$ GACT or in the N<sup>ε</sup>-( $\gamma$ -glutamyl)lysine substrate were responsible for activity changes in the pH range of 6.0 to 7.5 and 7.5 to 9.0. The presence of an imidazole ring of histidine,  $\alpha$ -amino group of the N-terminal residue, the  $\epsilon$ -amine group of lysine, or the phenol ring of tyrosine residue are likely to be present in the active site of the enzyme; the proposed mechanism is the one shown in figure 1.2.

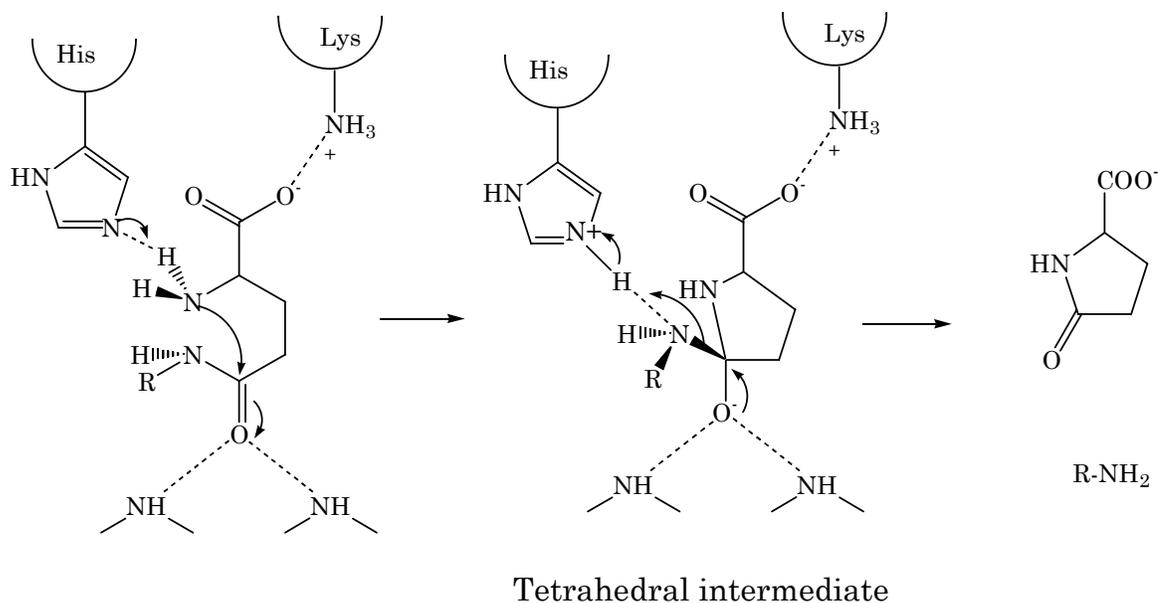


Figure 1.2: Proposed mechanism of  $\gamma$ GACT action (Gonzalez, 2005)

#### *$\gamma$ -Glutamylamino Acid Cyclotransferase ( $\gamma$ GAACT)*

In 1978, Taniguchi and Meister reported the presence of several forms of  $\gamma$ GAACT, distinguished by ion exchange chromatography and isoelectric focusing, as modifications of 5 to 7 sulfhydryl groups. Two isoforms were purified; both of them have molecular weights of 27,000 D, optimal pH of 7.5 to 8 in Tris/HCl, similar E1% of 13.1, and similar amino acid composition, but pI's of 4.6 and 5.1. However, reduction and alkylation of the mixture, using dithiothreitol (DTT) and idoacetamide (IAA), gave a single protein of pI 4.6. A conclusion of this work is that during the purification and storage the enzyme is unstable and undergoes considerable changes in its physical and catalytic properties.

For  $\gamma$ GAACT, 4-hydroxymercurybenzoate inhibits an isoform of pI 5.1 but not one of pI 4.5; also, disulfides inhibit the activity of the isoform 5.1,

while treatment with cysteamine causes its heterogeneity. In contrast, disulfide compounds cause a decrease of  $\gamma$ GAACT activity (Taniguchi and Meister, 1978). Although,  $\gamma$ GAACT showed no inhibition by EDTA and a metal does not seem to be necessary for catalysis, it has been reported that its mechanism should be similar to that of QC based on the isotopic exchange effect of the reactions catalyzed by each enzyme (Gololobov et al., 1994 and York et al., 1984).

### *Glutaminyl Cyclase (QC)*

Glutaminyl cyclase enzymes present good stability, which has allowed their purification, cloning, and structural studies. The glutaminyl cyclase enzymes from plant and human differ greatly in their structure, figure 1.3. The crystal structure of the plant enzyme has an all- $\beta$  conformation (Oberg et al., 1998), while the mammalian enzyme has equal  $\alpha$  and  $\beta$  structures.

For the human QC, Busby et al. (1987) found two isoforms of the enzyme one with pI 5.7 and the other with pI 7.2. Human QC isoforms were inhibited by ammonium (Fischer and Spiess, 1987) and by transition metals, 1,10-orthophenanthroline, and N-ethylmaleimide, but stimulated by EDTA (Busby et al., 1987).

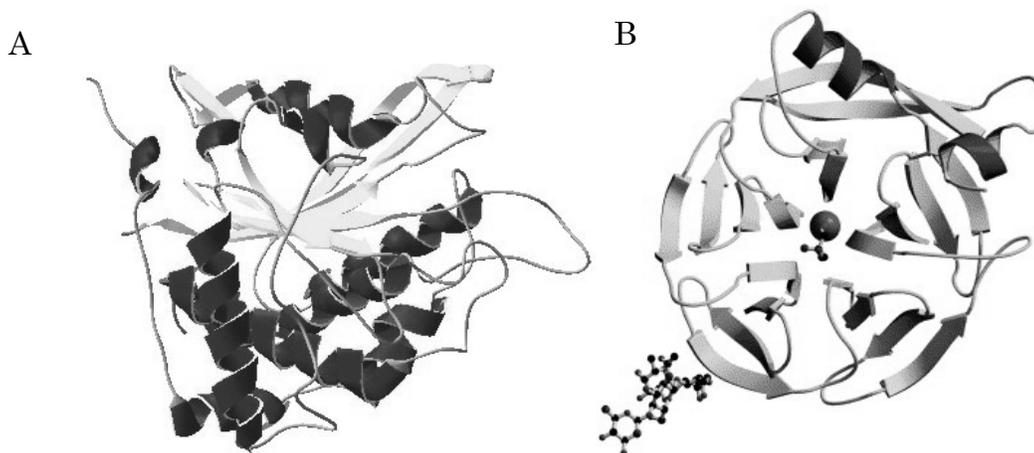


Figure 1.3: Crystal structure of glutaminyl cyclase from A) human (Huang et al., 2005) secondary structure obtained using the Swiss View 3.7 and B) plant (Wintjens et al., 2006)

In 1991 Pohl et al. reported that QC from different tissues showed different sensitivity toward sulfhydryl reagents, metals, and salts. Finally, Schilling et al. (2003) reported the human enzyme as a metallo-protein requiring one zinc ion and two histidine residues for catalysis; although, the enzyme was inhibited by 1,10-phenantroline but not by EDTA. Schilling et al. (2003) said about the mechanisms of the human QC:

“it seems likely that a metal ion in the active site of QC acts by polarizing the  $\gamma$ -amide group of the substrate glutaminyl residue, simultaneously stabilizing the oxianion formed by the nucleophilic attack of the  $\alpha$ -nitrogen of the scissile  $\gamma$ -carbonyl carbon.”

Huang et al. (2005) published the crystal structure of the human QC; showing that the zinc ion is tetrahedrally coordinated to water, histidine, and the side chain oxygen of glutamate and aspartate; on the other hand, aspartate and glutamate residues are needed for catalysis. Glu201

deprotonates the substrate  $\alpha$ -amino group while Asp248 binds the  $\gamma$ -amide group and Asp159 binds the metal that binds the  $\gamma$ -carbonyl; the  $\alpha$ -nitrogen attacks the  $\gamma$ -carbonyl and ammonia is released.

The proposed mechanism for catalysis of mammalian QC is shown in figure 1.4; Glu201, Glu202, Asp159, and Asp248 in combination with one zinc atom promote the intramolecular nucleophilic attack of the  $\alpha$ -amino group on the  $\gamma$ -glutamylcarbonyl, forming a tetrahedral intermediate and followed by the release of ammonia.

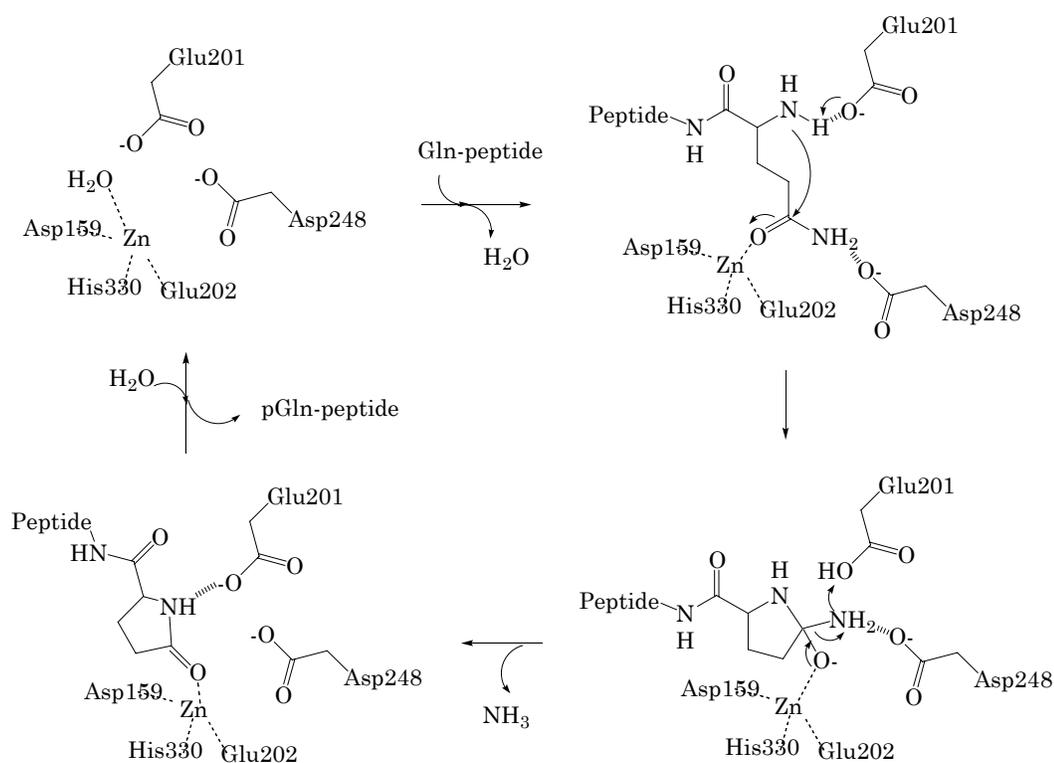


Figure 1.4: Proposed mechanism for glutamyl cyclization by mammalian QC (Huang et al., 2005)

Messer and Ottesen (1965) reported that the plant QC does not have a catalytic metal because it was not inhibited by EDTA or 1,10-phenantroline;

neither does it have catalytic cysteine residues because it was not inhibited by iodoacetamide or mercuric chloride. Later, Wintjens et al. (2006) found that plant QC has a zinc metal ion that does not participate in the catalysis of the enzyme, and that the enzyme is slightly inhibited by tris buffer as a competitive inhibitor.

Gololobov et al. (1994) found that the plant QC reaction does not go thru an acyl intermediate but rather a concerted cyclization; an isotopic effect of same magnitude as the one observed for human  $\gamma$ GAACT was observed by York and coworkers (1984) for plant QC. In a comparison of the human and plant QC, the role of Glu201 is conserved, but the action of zinc in the enzyme from humans is undertaken in plants by Lys225 with the participation of Glu24 and Asp155 side chains (Wintjens et al., 2006), as observed in figure 1.5.

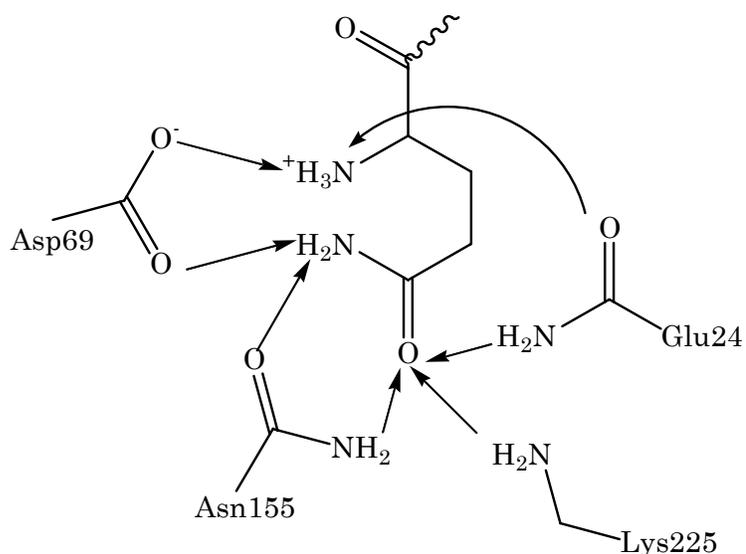


Figure 1.5: Proposed interactions of active site amino acids of plant QC and substrate (Wintjens et al., 2006)

### *$\gamma$ Glutamyl Transpeptidase*

Iodoacetamide inactivates  $\gamma$ GTP, although it only possesses a single free thiol that when replaced by alanine or serine results in fully activity of the enzyme (Ikeda et al., 1995a). Modification of the enzyme by iodo[<sup>14</sup>C]-acetamide showed the labeling of Asp422 (Smith and Meister, 1995). Phenylglyoxal, which modifies guanidino groups, inactivates the enzyme by modifying Arg110 (Stole and Meister, 1991). Mutation of Ser451 or Ser452 results in loss of activity (Ikeda et al., 1995b). Taniguchi and Ikeda (1998) reported that the binding of the human  $\gamma$ GTP could involve Asp423, Arg107, and Ser 451 and Ser452, where serine residues can act in two possible ways as nucleophile or as components of an oxianion hole for the stabilization of the transition state; these structures are shown in figure 1.6.

Figure 1.7 shows the crystal structure of two of the subunits of the tetrameric enzyme.

### *Transglutaminases*

Alkylation of the active site thiol group of Factor XIIIa affects the total activity of the enzyme (Seelig and Folk, 1980). The formation of N<sup>ε</sup>-( $\gamma$ -glutamyl)lysine linkage by transglutaminase requires the formation of an intermediate thioester (Folk, 1983). Transglutaminase has a catalytic triad similar to cysteine proteases, but instead of cleaving, the enzyme links polypeptides.

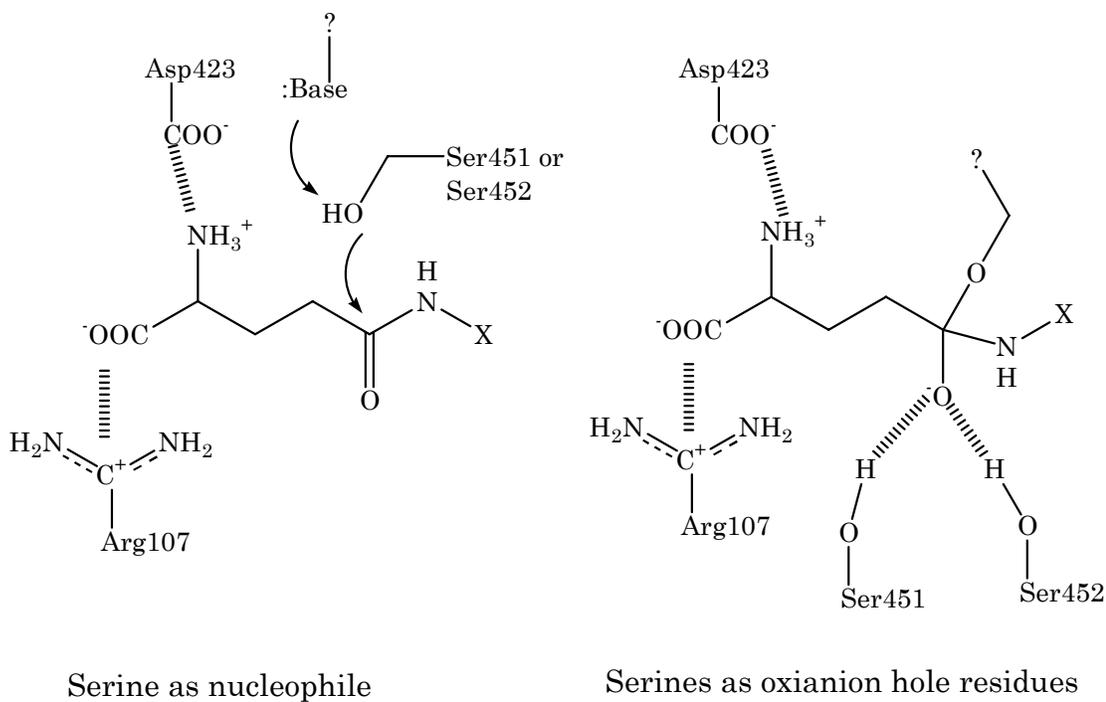


Figure 1.6: Possible binding of  $\gamma$ -glutamyl substrate by  $\gamma$ GTP (Taniguchi and Ikeda, 1998)



Figure 1.7: Crystal structure of  $\gamma$ GTP dimer with glutamate substrate (Structure 2DG5) (From Okada et al., 2006)

The crystal structure of one of the subunits of the dimeric transglutaminase blood coagulation Factor XIII is shown in figure 1.8.

The mechanism of action of Factor XIII is as shown in figure 1.9; the thiol group of Cys314 attacks the  $\gamma$ -carbonyl of the glutamyl reactant; ammonia release reforms the carbonyl of the reactant forming a thioester intermediate that now is attacked by the  $\epsilon$ -amino group of a peptide bound lysine residue reactant; the product is released by the reformation of the carbonyl and release of the enzyme thiol group (Pedersen et al. 1994).

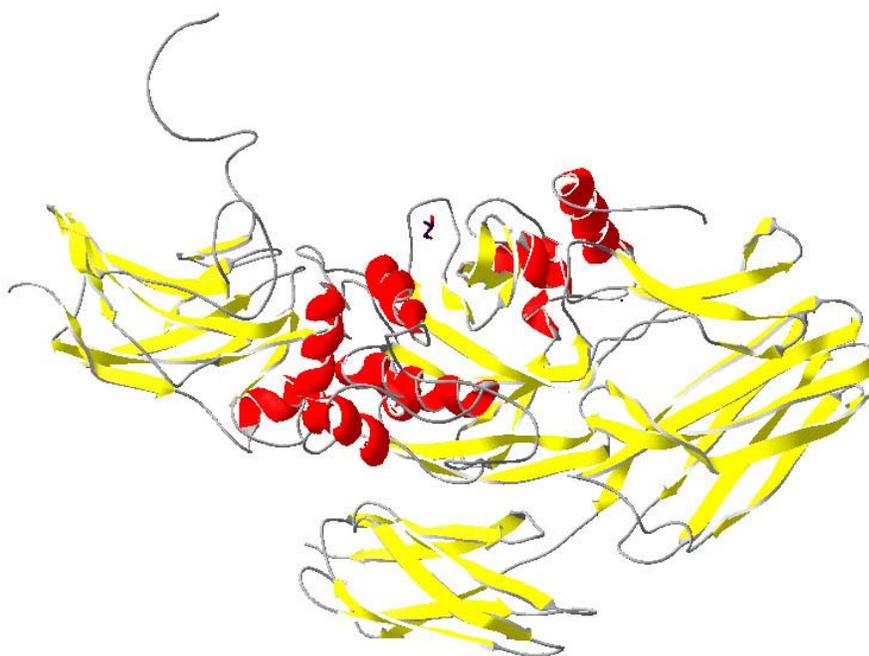


Figure 1.8: Crystal structure of Blood Coagulation Factor XIII transglutaminase (monomer) (Structure 1GGU) (Fox et al., 2000)

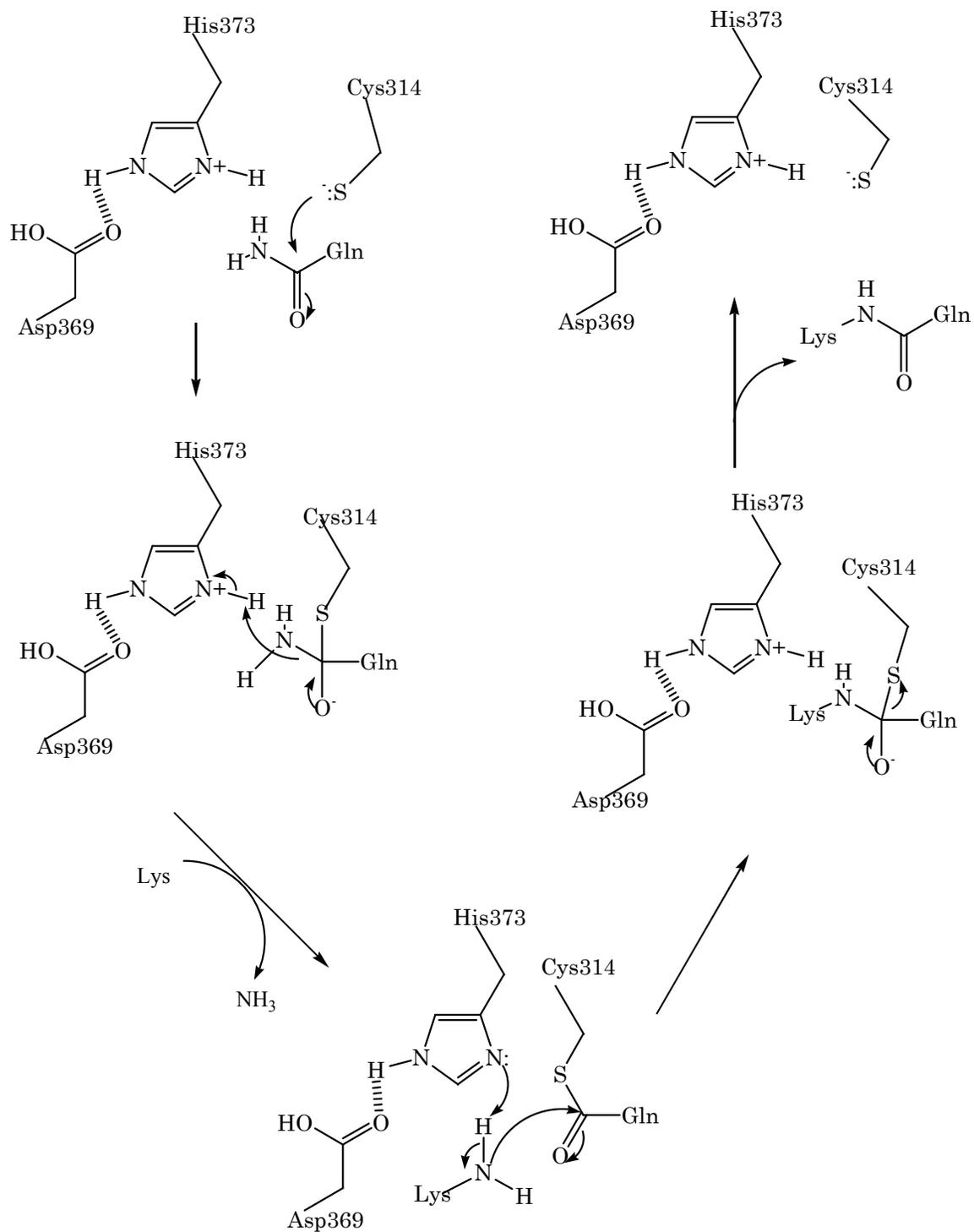


Figure 1.9: Proposed mechanism for Factor XIII transglutaminase reaction. (Pedersen et al., 1994)

### *Pyroglutamidase*

The purification of pyroglutamidase was initially attempted by Doolittle and Armentrout (1968); nonetheless, the difficulties in this work hindered the purification of the enzyme for several years. Pyroglutamidase is a cysteine protease; the mechanism of cleavage of the 5-oxoprolyl residue by pyroglutamidase involves the formation of an intermediate thioester. The recognition of the substrate (pyroglutamyl-peptide) by pyroglutamidase requires the formation of two hydrogen bonds with the main chain of the enzyme; the hydrogen bonds help to orient the pyroglutamyl residue for Cys144 attack. A hydrophobic pocket formed by Phe10, Phe132, Thr45, Ile92, Pro142, and Val143 seems to be essential for catalysis (Ito et al., 2001). The crystal structure of pyroglutamidase is shown in figure 1.10.

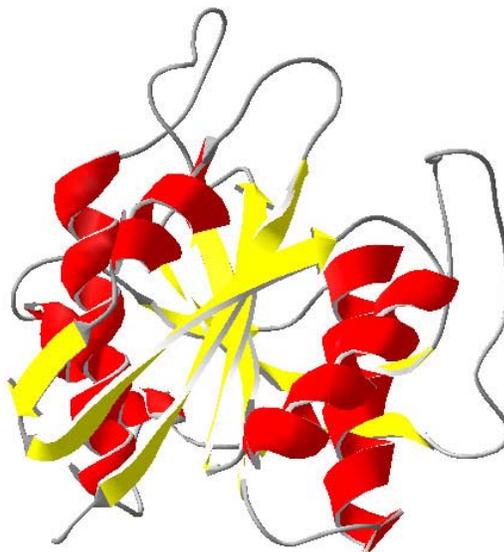


Figure 1.10: Crystal structure of pyroglutamidase (Structure 1AUG) (Odagaki et. al., 1999)

### *General Observation of Enzyme Characteristics*

What is known about enzymes depends on the facility of their purification, which in turn depends on the stability of the enzymes; for example, for pyroglutamidase Doolittle and Armentrout (1968) said; “Further purification attempts have been hindered by the instability of the more purified enzyme preparations.” Structural and genetic studies on  $\gamma$ GAACT and  $\gamma$ GACT have not been done because these enzymes have not been purified due to their instability.

About  $\gamma$ GAACT, Orlowski and Meister (1973) said:

“The possibility that there are different  $\gamma$ -glutamyl cyclotransferases exhibiting different substrate specificities must be considered, but the present findings indicates that the occurrence of the enzyme modification during purification presents a serious complication that would hinder studies on this point.” “It would appear that purification and study of the native  $\gamma$ -glutamyl cyclotransferases must await the development of procedures that will minimize or prevent altogether enzyme alteration of the type observed here”

Partially purified enzyme can be use to study some properties and for kinetic studies as long as the contaminants present do not alter the catalytic activity of the enzyme. Tables 1.1 and 1.2 summarize the characteristics of the cyclotransferases and their related enzymes.

The fact that  $\gamma$ GAACT substrates are not  $\gamma$ GACT substrates demonstrates that they are indeed two different enzymes.

Table 1.1: Catalytic properties of cyclotransferases and related enzymes

Enzyme	Substrate	Product	Mechanism	Active site	Reference
$\gamma$ GACT	N- $\gamma$ -glutamylamine N <sup><math>\epsilon</math></sup> ( $\gamma$ -glutamyl)lysine	5-Oxoproline and Lysine or amine			
$\gamma$ GAACT	N <sup><math>\alpha</math></sup> ( $\gamma$ -glutamyl)amino acid	5-Oxoproline and amino acid			EC 2.3.2.4
Transglutaminase	$\gamma$ -glutamyl residues variety of amines	N <sup><math>\epsilon</math></sup> ( $\gamma$ -glutamyl)lysine crosslink	calcium-dependent acyl transfer, high energy thioester intermediate	Cys314, His373, Asp396	EC 2.3.2.13 (Pedersen et al. 1994; Folk, 1983)
QC Human	Glutamine or N-terminal glutaminyl	5-Oxoproline and ammonium or 5-oxopropyl	$\alpha$ -amino intramolecular nucleophilic attack to $\gamma$ - glutamyl-carbonyl, tetrahedral intermediate	Zn, Asp159, Asp248, Glu201, Glu202, His330	EC 2.3.2.5 (Huang et al., 2005)
QC plant				Lys225, Glu69, Glu24, Asn155	EC 2.3.2.5 (Wintjens et al., 2006)
Pyroglutamidase	5-oxopropyl- polypeptide	5-Oxoproline and polypeptide	Cysteine protease intermediate thiol-ester	Cys144, His168, Glu81	EC 3.4.19.3 (Ito et al., 2001)
$\gamma$ -Glutamyl transpeptidase	Glutathione and amino acid	$\gamma$ -glutamyl amino acid and Cys-Gly	nucleophilic attack, tetrahedral intermediate	Asp423, Ser451, Ser452, Arg107	EC 2.3.2.2 (Taniguchi and Ikeda, 1998)

Table 1.2: Characteristics of cyclotransferases and related enzymes

Protein (reference)	Total amino acids	MW (Daltons)	pI	Amino acid composition																			
				A	R	N	D	Q	E	G	H	I	L	K	M	F	P	S	T	Y	V	W	C
Rabbit $\gamma$ GAACT (Taniguchi and Meister, 1978)	237	27,000	5.1 4.6	20	9	<u>23</u>		<u>29</u>	20	5	8	20	17	2	10	13	12	12	7	15	8	7	
Human QC (2AFZ)	329	37,515	6.12	23	16	15	19	16	19	16	16	17	44	10	7	15	22	28	12	13	7	2	
Bovine QC (NP_803472)	361	41,224	6.53	23	25	14	25	15	16	20	16	50	50	8	8	18	23	26	15	11	22	8	3
Papaya QC (GI:75220038)	288	33,375	8.48	9	18	18	19	7	15	21	9	24	37	19	3	11	7	13	17	13	18	7	3
Human Pyroglutamidase (Q9NXJ5)	209	23,138	5.90	11	8	5	12	10	14	18	9	10	19	10	4	5	10	13	11	9	21	2	8

## CHAPTER TWO

### Purification

#### *Purification Considerations*

During the initial steps of purification of proteins, they may be attacked by proteolytic enzymes; in consequence, the early steps of protein purification should be carried out at low temperature, where proteases are less active and as fast as possible to prevent degradation. Hence, the initial techniques employed should allow the handling of a high amount of sample and quick processing at low temperature; later, smaller amounts of proteins with similar characteristics to that of the target protein are recovered.

The best strategy for purification is the one that combines resolution, capacity, speed, and recovery in the minimum number of steps. In general, the first steps are capture or gross purification with the objective of isolation, concentration, and stabilization of the proteins; critical in these steps are speed and capacity because the quick handling of high amounts of sample is required. Intermediate purifications have as an objective the removal of most of the bulk impurities; the most important considerations in these steps are resolution and capacity. Finally, polishing steps are carried out to remove those contaminants with similar characteristics to those of the protein of interest; these steps should allow the handling of small amounts of sample

looking for the best resolution and recovery. The protein is considered pure when no extra purification can be obtained by more powerful techniques.

### *Native vs Recombinant Protein Purification*

Purification of protein from its natural source is complicated because the target protein is present at a very low concentration and all other proteins are potential contaminants. Purification of recombinant proteins is much easier because the target protein is expressed in a simpler organism like yeast or bacterium, where less contamination can be found. The usual target of 1000-fold purification of native proteins can be decreased to 100-fold purification for recombinant proteins (Stein, 1991).

Recombinant proteins are obtained by introduction of their genes in the genome of the simpler organism, which will express those recombinant genes as if they were its own genes. The target protein can be conjugated with another known protein to be separated on the basis of the affinity of the second protein, and sometimes the target proteins can be over expressed in those organisms.

For example, in the purification of human sphingosine kinase (SK) done by Pitson and coworkers (2000) the native purification employed a 13-step long procedure including enzymatic digestion, centrifugations, ammonium sulfate precipitation, dialysis, and a combination of seven chromatography columns. On the other hand, the recombinant protein was purified in a 6-step long procedure; SK was conjugated with glutathione S-

transferase (GST-SK), extracted from cell lysate in a GSH-Sepharose 4B, and eluted out using glutathione; GST-SK was cleaved using thrombin and separated in 2-column steps.

### *Tracking $\gamma$ GACT During Purification*

Given that  $\gamma$ GACT is an enzyme, it is followed during the purification steps by observation of one of the products of the reaction that the enzyme catalyses. Qualitative and quantitative procedures, developed by Fink and coworkers (1980), are used to track the enzyme. The quantitative method quantifies the amount of lysine released in the reaction of  $\gamma$ GACT on  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine (reaction 2) during 15 min at 37°C. The amount of lysine released by the reaction of  $\gamma$ GACT on  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine is used to define the activity of the enzyme; one unit equals one micromole of lysine produced per hour of reaction at 37°C.

In parallel, others assays are carried out for enzymes that can interfere with  $\gamma$ GACT activity measurement. Yet,  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine is also a substrate for  $\gamma$   $\gamma$ GTP; the presence of this enzyme is monitored by tracking the formation of products from their specific substrates.

### *$\gamma$ GACT Purification Procedures*

In brief, the purification done by Fink and Folk (1983) consisted on the homogenization of rabbit kidney tissue in sucrose solution by breaking open the cells mechanically; then ultracentrifugation is employed to separate the

soluble proteins from the insoluble cell material and organelles. The supernatant, which contains  $\gamma$ GACT, is introduced onto an anion exchange chromatography column in order to remove those proteins with pI values different than that of  $\gamma$ GACT; a DEAE cellulose matrix is used; this step is fast and large amount of sample can be introduced in the column. The fractions that show activity are pooled together and precipitated using ammonium sulfate; proteins precipitated this way are stable. The purification gave ~250-fold purification; the obtained protein had a specific activity of 32 U/mg and the total yield of the purification was 7 %.

Gowda (1985) worked with the bovine enzyme; she added a size exclusion chromatography step on Sephadex G100 and G50 in order to discard those proteins with molecular masses different than that of  $\gamma$ GACT and an isoelectric focusing step after lyophilization and dialysis of the protein after size exclusion chromatography. The purification gave ~900-fold purification; the obtained protein had a specific activity of ~300 U/mg and the total yield of the purification was ~1%. Not enough protein was recovered for further studies.

Bowser (1997), working with the rabbit enzyme, used a similar purification to that of Fink and Folk (1983) and the size exclusion from Gowda (1985) with the addition of ion exchange and size exclusion chromatography on Mono Q and Superdex HR 75 columns, respectively.

Recovery and purification gained by those last chromatography columns were very poor.

### *Considerations for the Purification of Unstable Proteins*

The removal of sensitive proteins from their physiological environment into an artificial environment during purification can have a negative effect on the protein stability. In order to retain this stability, chemical factors like pH, ion strength, reducing conditions, cofactors, protein concentration, as well as physical factors like temperature, time, surface effect, and pressure, have to be taken into account in the design of a purification protocol.

Although the main goal for purifying a protein is to remove all other proteins, contaminant proteins are needed sometimes. During the early steps of purification proteases are the principal threat for the target proteins, and other contaminants can act as competitive protease inhibitors; also, additional proteins are sometimes added to the pure protein in order to mimic the environment of the cell. Moreover, pure protein in low concentration can easily be lost by absorption onto surfaces; in this case a contaminant protein prevents this loss by coating the surface of the vial.

The stability of the protein in an immobilized state is also important. The addition of sugars to proteins before lyophilization has proven to preserve the native structure of proteins in the absence of water molecules. Chang et al. (2005) showed that the stability of the protein correlates with the conservation of its native structure. There are two hypotheses for the

stabilization effect of sugars on sensitive proteins: the kinetic “glass dynamic hypothesis”, which states the formation of a rigid matrix where the protein is limited in mobility minimizes unfolding attempts; and the thermodynamic “water substitute hypothesis” that states the formation of specific hydrogen bonds on the surface of the proteins replaces the thermodynamic stabilization from the lost water molecules.

The stabilization of proteins is even more challenging when the unstable proteins are subjected to separation by chromatography. Kaufmann (1997) analyzed the effect of some chromatography parameters on the stability of proteins, showing that the purification is compromised by the stability of the protein.

While enzymatic activity is minimized by working at temperatures close to 4°C, backpressure is increased and diffusion is decreased for chromatography separations, causing the decrease of flow rate and the increase of elution time. Pressures like in a FPLC separation showed no negative effect on the stability of proteins; however, pressures in centrifugation can cause inactivation, especially in a complex of proteins.

The composition of the buffer for separation depends on the nature of the solid matrix and the stability of the protein. Salt concentration should prevent non specific interaction with the matrix and avoid salting out of the protein. Reducing agents prevent oxidation, and while they can help to maintain the native structure of proteins by preventing the formation of

disulfide bridges, they can also disrupt disulfide bridges required for the native structure. Osmolyte additives such as sugars, polyols, amino acids, betaines, and ectoines (figure 2.1), which are neutral, highly soluble in water, and do not interact with the proteins, oppose the unfolding and denaturation of proteins (Galinski, 1995) by inducing the formation of better clusters of hydration because proteins will tend to hydrate more in order to avoid contact with these osmolytes (Gekko and Timasheff, 1981).

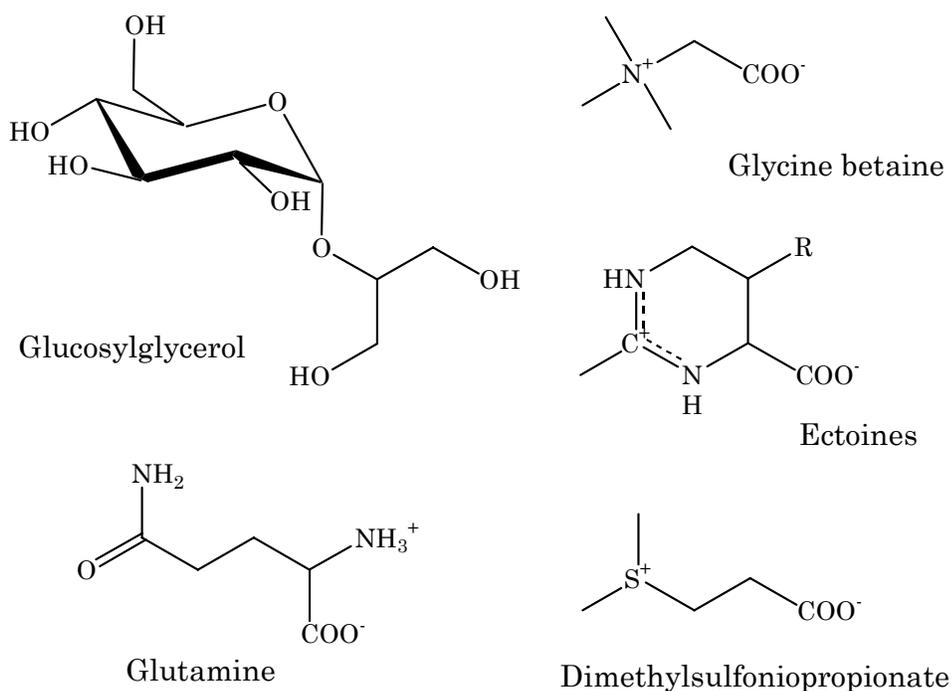


Figure 2.1: Organic osmolytes (Galinski, 1995)

### *Chromatography Improvements*

Capture of charges is the basis of separation in ion exchange chromatography. In this technique the matrix is modified with an ion

exchanger group that confers the solid support with a net charge, so ions of opposed charge will bind to the solid support and ions with the same charge as the solid support will pass through unretained.

The elution of the retained ions is obtained by increasing the surrounding ionic strength or by changing the pH. By increasing the ionic strength the interaction between the retained ions and the matrix is no longer the most favorable; by changing the pH amphoteric molecules will change their charge and be repelled from the solid support.

When purifying proteins, their charge is determined by the pH. If the pH is below a protein's isoelectric point (pI), the protein will be charged positively because both acidic and basic groups will be protonated. On the contrary, if the pH is above a protein's pI, the protein will be negatively charged because acidic and base groups are deprotonated.

The original ion exchange column employed for the purification was based on a cellulose matrix; the one proposed in this work is Sepharose. The Sepharose matrix offers better attachment for substituents than cellulose (Baeseler et al., 1992). Cellulose matrices have poor flow capacity due to their irregular shape, while Sepharose media, based on cross-linked chains of agarose, have the advantages of better capacity, faster flows, and lower non-specific adsorption. For the anion exchange in the first chromatographic step, the matrix is modified with the anion exchanger diethylaminoethyl (DEAE).

Mono Q is a matrix employed for anion exchange chromatography; the solid support is made of monodispersed, rigid, polystyrene/divinyl benzene particles. The matrix is modified with quaternary ammonium (-O-(CH<sub>2</sub>)-N<sup>+</sup>(CH<sub>3</sub>)<sub>3</sub>); because this exchanger is not capable of exchanging a proton, the pH is not altered due to the salt changes during elution; this kind of exchanger is called “strong”. In contrast, DEAE exchanger (-O-CH<sub>2</sub>CH<sub>2</sub>N<sup>+</sup>H-(CH<sub>2</sub>CH<sub>3</sub>)<sub>2</sub>) that can exchange a proton with the buffer is called “weak” because the pH can suddenly vary with ionic strength changes.

Mono Q has a great resolution power; it has been extensively employed for the purification of very similar proteins; for example, different forms of protein kinase C from rabbit platelets (Pelech et al., 1991) were resolved by Mono Q, as well as, the subunits of cytochrome C oxidase (Liu et al., 1995).

Since ion exchange chromatography is a capture technique, it shows two great advantages: a large volume of diluted sample can be loaded before inducing elution and flow rates can be fast. Quite the opposite, size exclusion chromatography is not a capture technique because the molecules do not bind to the solid support and are always moving with the flow rate.

Size exclusion chromatography, or gel filtration chromatography, is based on the ability of particles to enter the pores of a solid support; this ability is inversely related to their sizes; the bigger the particles the less they will be retained in the pores of the solid support. Hence slow flow rates are required to allow the small particles to be held in the pores of the solid

support and be separated from the bigger particles that have to pass through. Also, because slow flows are required, small volumes of samples are required to prevent cross diffusion.

Sephadex matrices employed in the original purification are cross-linked dextran by reaction with epichlorohydrin, while Sephacryl is actually a composition of two polymers, cross-linked support dextran with N,N'-methylene bisacrylamide. This more developed matrix can maintain faster flows. In addition, Sephacryl present more resolution power than Sephadex and is still appropriate for large scale work.

### *Isoelectric Focusing*

Isoelectric focusing (IEF) is a separation technique based on the migration of amphoteric molecules in an amphoteric gel under an electric field. In this technique, amphoteric buffers in a large range of pHs called ampholites are prepared in a gel, generally polyacrylamide.

To generate the amphoteric gel, acidic and basic strips are placed on opposed ends of the gel; the acidic end will support the positive electrode (anode), and the basic end will support the negative electrode (cathode). When a voltage is applied, all ions will move toward their opposite charges; cations will move toward the cathode, and anions will migrate toward the anode. In their way, ampholites will pick up  $H^+$  or  $OH^-$  and become neutral, so they will remain steady, giving that position a specific pH.; in conjunction ampholites will grant the gel with a continuous pH gradient, increasing from

anode to cathode. Similarly, for the separation of a mixture of proteins, these are loaded onto the ampholyte gel; because proteins have specific charges determined by the pH, they will migrate toward their opposed charge when the voltage is applied. In their way, proteins will pick up  $H^+$  or  $OH^-$  and become neutral. Any attempt to migrate farther or its way back will grant the protein a charge that forces it to return to its isoelectric point; this is the infamous focusing effect.

In 1961, Svensson invented IEF as a preparative technique. This technique evolved as an analytical tool into two dimensional electrophoresis (O'Farrell, 1975) with isoelectric focusing on an immobilized pH gradient (Bjellqvist et al., 1982). IEF has been employed as an analytical and preparative technique for the separation of the most similar mixtures (Radola, 1973); it can separate proteins that differ only in 0.01 pI units (Righetti and Bossi, 1998). IEF shows how a target protein and its contaminants behave under different pH's; consequently, the technique can be employed to predict the optimal conditions for the separation of proteins in ion exchange chromatography (Haff et al., 1983). In the work of Lindblom and Axio-Fredriksson (1983) they separated urine proteins on a Mono Q column after finding optimal conditions for separation by IEF.

## CHAPTER THREE

### Sequencing of Proteins

#### *Cyclotransferase and Genetics*

For  $\gamma$ GACT nothing is known about its gene. On the contrary, structural and genetic studies on glutaminyl-peptide cyclotransferase are advanced; the bovine (Pohl et al., 1991), human (Song et al., 1994) and plant (Dahl et al., 2000) enzymes were cloned and expressed. Several amino acid sequences of the enzyme from animals, plants, and bacteria have been deposited at the GenBank.  $\gamma$ GAACT's alleles were found in the chromosome 6 of mouse (Tulchin and Taylor, 1981); however, its primary structure has not been reported.

#### *From the Protein to Its Gene*

In order to find a gene, the order in which the amino acids of a polypeptide are linked must be known because this is what is related to the DNA sequence of the gene that produces the polypeptide. The searching of a gene is done by reverse translation of the protein amino acid sequence into the corresponding mRNA, and reverse transcription of the latter into the complementary DNA (cDNA). However, transcription and translation are not conservative processes, figure 3.1; this means that by knowing the amino acid sequence of a protein only a probable paternal DNA can be guessed. Fortunately, at least one of the putative DNAs will find a match in the known

genome.  $\gamma$ GACT gene should be available for production of recombinant  $\gamma$ GACT.

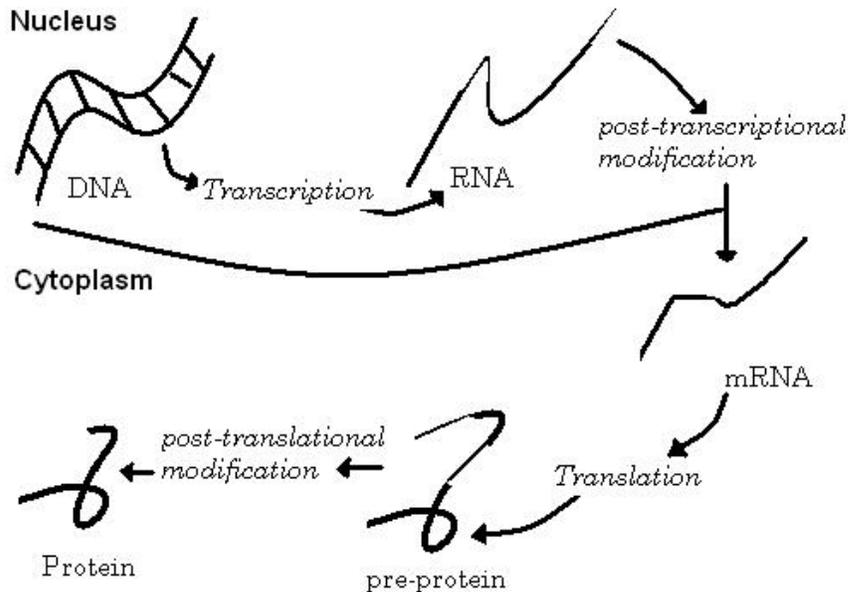


Figure 3.1: Transcription and translation are not conservative processes

The sequence of the whole protein is not usually required; a sequence of 6 to 10 amino acids long is needed in order to propose the possible cDNA and identify a gene. With this little sequence of amino acids the possible paternal cDNAs are proposed, and bioinformatics searches are started in order to find the gene in the bovine genome that originates  $\gamma$ GACT. If the gene is found, then a library with this gene can be bought or prepared, and the nucleotides for screening of the library can be fabricated; the  $\gamma$ GACT gene can be amplified using the polymerase chain reaction (PCR), inserted in a cell and expressed in a future work. In this scene, a relative large amount of protein can be obtained for structural and functional studies.

### *Sequencing Techniques*

Sequencing of pure proteins is done in two ways: Edman degradation or *de novo* sequencing mass spectrometry. The first one is believed to be more accurate because amino acid residues are analyzed consecutively; however, sequencing by mass spectrometry can be more sensitive, can cope better with protein mixtures and N-terminal modified proteins, and requires minimal assistance of genomic information (Standing, 2003).

#### *Polypeptide Sequencing by Edman Degradation*

Edman degradation, named after its inventor Perh Edman (Edman and Begg, 1967) uses phenylisothiocyanate to form a phenylthiocarbamyl adduct (PTC) with primary amines in mildly alkaline conditions; the PTC product is treated with anhydrous, strong acid that only cleave the N-terminal residue as a thiazolinone amino acid but does not hydrolyze the other residues of the polypeptide. The derivatized amino acid is extracted into an organic solvent and converted into a stable phenylthiohydantion (PTH) amino acid, which can be identified by comparison with PTH-amino acid standards using HPLC with UV detection. The polypeptide can be treated successively with the same method for determination of its sequence; this method has been automated. The Edman degradation cycle is shown in figure 3.2.

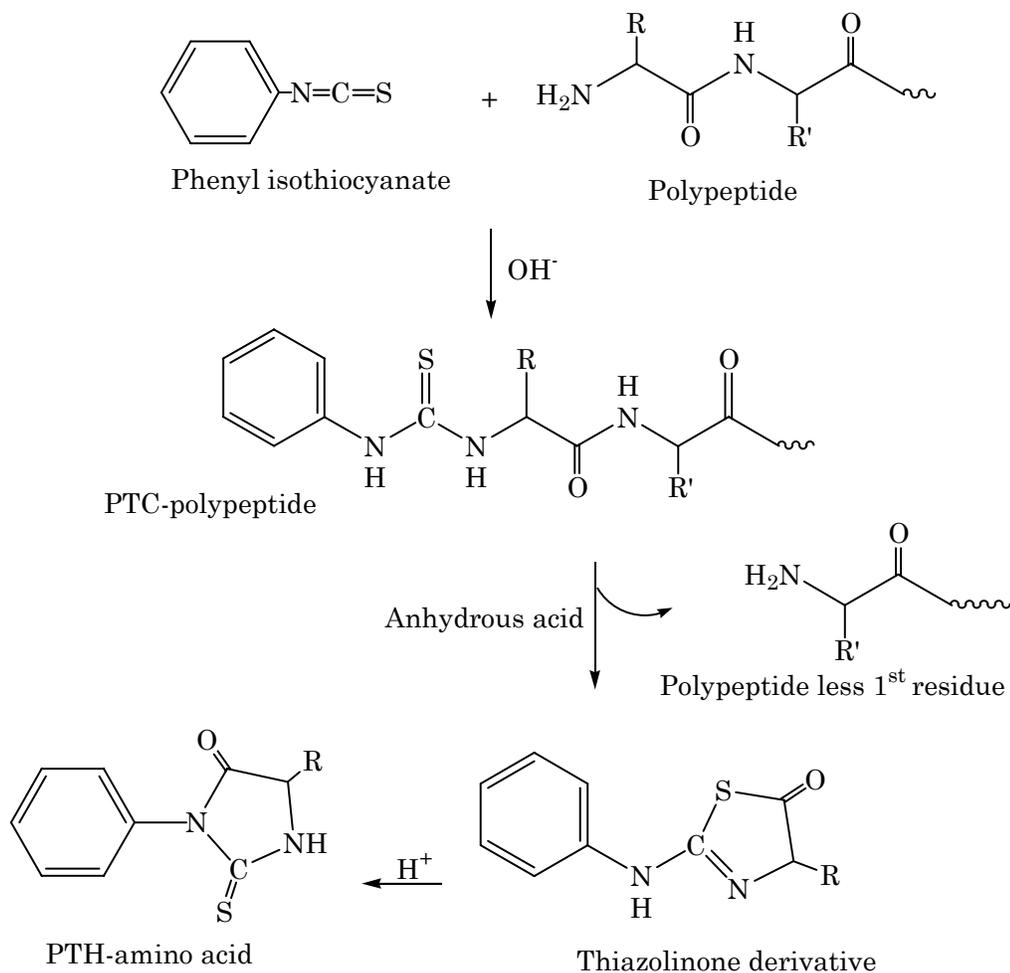


Figure 3.2: Edman Degradation reaction

The requirement for Edman degradation sequencing is a free N-terminal. However, almost all cytosolic proteins are blocked at their N-terminus (Brown and Roberts, 1976). To check for a free N-terminus the whole protein is subjected to dansylation that labels only free primary amines, and hydrolysis that releases the free amino acids; then, the individual amino acids are resolved by HPLC and detected by fluorescence; only the free terminal amines are observed, except for N- $\epsilon$ -dansyl lysine.



enzymatic digestion the protein are introduced in the mass spectrometer. In both approaches the identification of the protein can be done by comparison of the fragments masses with databases or by *de novo* peptide sequencing.

For *de novo* peptide sequencing in mass spectrometry (Hunt et al., 1986), polypeptides of fragments are fragmented in a tandem MS/MS at low-energy, so that a limited number of ions are obtained, generally breakage of the peptide bond. The sequence of the peptide is found by matching the differences of masses among peaks with the standard masses of the amino acids. For example in the spectrum shown in figure 3.3, the difference of mass between the peak 813.0 and 684.2 is 128.8, which is the mass of glutamic acid; the difference of the peak 684.2 and 627.8 is 56.4 that corresponds to the mass of glycine; because the observed peaks are the c-terminal ions, the sequence gly-glu can be concluded. Obviously the mass difference between 684.2 and 642.6 that is 41.6 does not correspond to the mass of any amino acid, so the fragment 642.6 should be compared against other fragments.

Nonetheless for protein identification using mass spectrometry sequencing of the individual peptides is not always necessary if the genome of the target protein is known. Peptide mass fingerprinting is a technique that *in silico* compares the mass fragmentation pattern of a digested protein by a specific protease with the predicted mass pattern of the protein obtained by computational translation of the genome and digestion with the same

protease (Griffin et al., 1995). The disadvantage of this method is that the peptide map of the protein of interest has to be in a database.

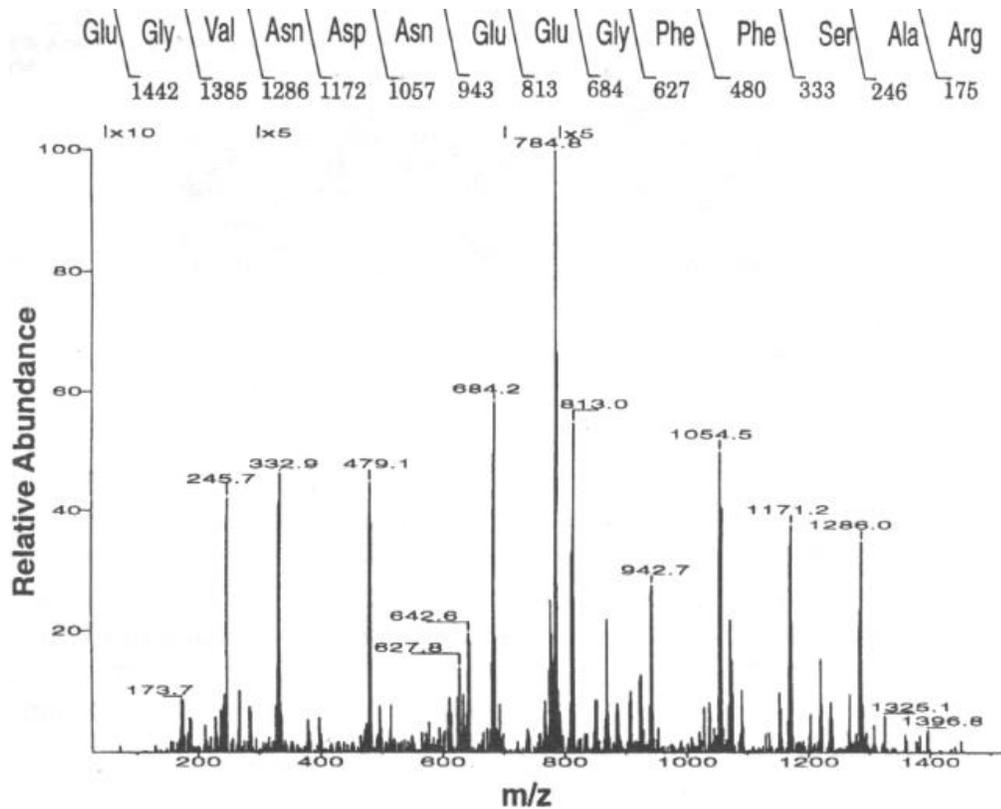


Figure 3.3: Mass spectrum of [Glu<sup>1</sup>] fibrinopeptide B, used as standard for the verification of MS/MS performance; the masses of the C-terminal fragments are shown in the top sequence (Yates, 1996)

### *Gene Search*

As soon as part of the primary structure of a protein is identified, the searching for similarities is the next step. If the protein can be explained as of contamination, the sample can be discarded, and further useless work be prevented; on the other hand, the protein and its purification would be

validated if the protein is found to match with similar proteins of other organisms.

A look to the genome-translated proteins can also be useful in discriminating a putative gene. The correlation between the properties of these model proteins and the actual properties observed during purification and study of the protein can be use to narrow down the list of possible genes. For example if one of the model proteins has a transmembrane sequence this should be discarded if the actual protein is a cytosolic one. Similarly, molecular mass, amino acid composition, and post-translational modification are properties that can be employ in discriminating possible genes.

The Basic Logical Alignment Search Tools (BLAST), introduced in 1990, is a program based on an algorithm to compare a sequence in a database, looking for similarities of the query sequence by alignment of portions of the query and database sequences; it is able to compare amino acid sequences against a protein database, nucleotide sequences against a nucleotide database, all reading frames of a nucleotide sequence against a protein database and vice versa.

When the genome of the organism is poorly known, scientists not only have to sequence the protein, but also the DNA in order to propose the gene. The bovine QC sequence was determined by Pohl and coworkers in 1991. Using Edman degradation of the pure protein, they found a 28-peptide long N-terminus; they prepared a cDNA library from the RNA purified from the

same cells in a  $\lambda$ gt11 vector with the librarian XI kit from Invitrogen. The cDNA was expressed in COS-7 cells in a pCDM8 vector. The enzyme was extracted from the cells 48 hours later. Based on the sequence of the N-terminus, the following oligonucleotides were proposed 5'-GG(C/G/T/A)-GC(T/C)-GT(GC)-GA(T/C)-TGG-AC(A/C)-CA(A/G)-GA(A/G)-AA(A/G)-AA(T/C)-TA(T/C)-(C/A)G(G/A)-CA(A/G)-CC(T/C)-GC(C/T)-CT(G/C)-CT-3'; the library was screened with the nucleotides, and the positive clones were isolated, subcloned and sequenced. No matches of this sequence were found in the databank at the time, so they proposed the gene. The whole sequencing of the cloned DNA allows them to find the gene of QC in the bovine genome.

Nonetheless, if the genome of the organism is known the actual gene can be used to design the cloning and amplification experiments. The Bos taurus sequencing project is being led by researchers at the Baylor College of Medicine Human Genome Sequencing Center and the Genome British Columbia Sequencing and Mapping Platform at the British Columbia Cancer Agency (BCCA); a first draft of this genome was made public in October 2004. The bovine genome is similar in size to the genomes of human and other mammals, with an estimated size of 3 billion base pairs.

## CHAPTER FOUR

### Design of an Affinity Column for $\gamma$ GACT based on the Glutarylhexylamine Inhibitor

#### *Affinity Chromatography*

Chromatography separations are based on the differences of the individual interactions of the analytes with the solid support and the mobile phase under given conditions. These interactions are generally physical, but in affinity chromatography these interactions are based on chemical and biological properties of the molecules and their ligands; some common interactions employed in affinity chromatography are antigen-antibody, hormone-receptor, lectin-polysaccharide, nucleic acid-complementary base, and enzyme-substrate.

There are two requirements for a suitable ligand in affinity chromatography: the ligand should have a chemically modifiable group that allows its attachment to the solid support and must exhibit a reversible binding of the analyte. In the present work, the interaction of enzyme with substrate is fundamental to the affinity chromatography. Nonetheless, the enzyme substrate can not be used because this reaction is not reversible, so a reversible inhibitor of the enzyme is employed.

The process of separation (figure 4.1) can be described in two phases: first, capture of the enzyme and wash out of contaminants; second, elution

and recovery of the enzyme. In the first step the conditions are so that the binding of the enzyme to the ligand is promoted, binding step; in the second step the conditions are changed to promote elution of the enzyme, elution step.

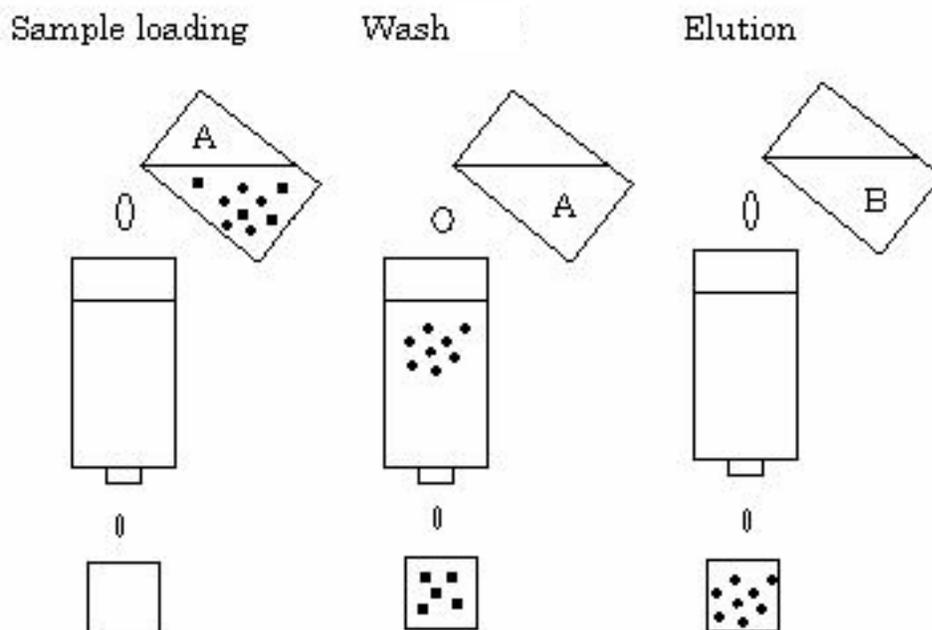


Figure 4.1: Separation by affinity chromatography

### *Reversible Inhibition*

The following figure (4.2) shows the reactions of an enzyme with its substrate in the presence of a reversible inhibitor.

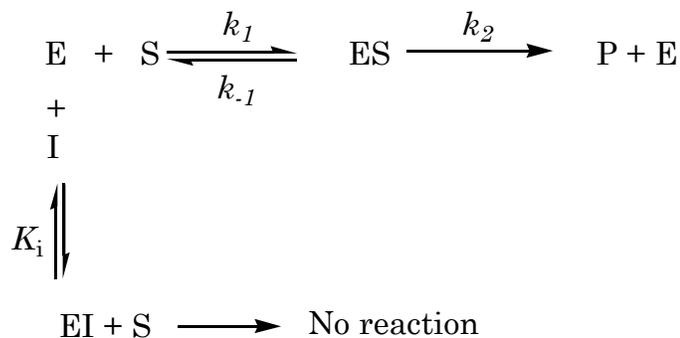


Figure 4.2: General scheme for inhibition; where, E = Enzyme, S= substrate, I = inhibitor, P= product, and EI and ES are the complexes enzyme-inhibitor and enzyme-substrate, respectively.

For this system without inhibitor when  $k_{-1} \gg k_2$ , this is for initial velocity steady state conditions under rapid equilibrium

$$K_M = \frac{[\text{E}][\text{S}]}{[\text{ES}]} \quad (\text{I})$$

$$k_{\text{cat}} = \frac{V_{\text{max}}}{[\text{E}]_T} \quad (\text{II})$$

where  $K_M$  represents the concentration of substrate at which the velocity of the reaction equals a half of the maximum velocity; thus, a small  $K_M$  means a an efficient reaction. The catalytic constant,  $k_{\text{cat}}$ , is called the turnover number because it represents the maximum of substrate molecules converted to product per active site per unit time; hence, a high  $k_{\text{cat}}$  correspond to a fast reaction. The specificity constants  $k_{\text{cat}}/K_M$  relates the reaction rate to the concentration of actual free enzyme and determines the specificity for competing substrates, so the higher the  $k_{\text{cat}}/K_M$  the better is the competing substrate for the enzyme.

When a competitive inhibitor is present, equation (III) describes the dissociation constant for the enzyme inhibitor complex

$$K_i = \frac{[E][I]}{[EI]} \quad (\text{III})$$

A smaller  $K_i$  means a higher  $[EI]$ , which is better inhibition. This is very important when the inhibitor is to be used *in vivo*, because a high concentration of inhibitor can cause parallel reactions in other pathways.

For an affinity column the dissociation constant of the enzyme-inhibitor complex ( $K_i$ ) should be in the range of  $10^{-4}$  to  $10^{-8}$  M during the capture step; a greater  $K_i$  will offer weak interactions with no retention of the enzyme; while, lower  $K_i$ 's will be too strong an interaction that the conditions required for elution will most likely inactivate the enzyme. During the elution step  $K_i$  should increase close to  $10^{-2}$  M by changing external parameters like pH, temperature, or ionic strength. In many cases a stronger competitive inhibitor can be used to displace the enzyme from the solid support.

For the determination of the kinetic constants  $K_i$ , and  $K_M$ , the Michaelis-Menten description of enzyme kinetics is employed assuming that the enzyme is the only catalyst of the reaction and there is no cooperativity or allosteric regulation, there is no spontaneous formation of products, sub product formation, or accumulation of products. The substrate is in a larger

concentration than the enzyme, so the enzyme-substrate complex is not altered by the consumption of the substrate.

For the determination of the  $K_i$ , the reaction velocity is measured at different concentrations of substrate and inhibitor; based on the double-reciprocal form of the Michaelis-Menten equation or the Lineweaver-Burke equation, equation IV; the reciprocal velocities are plotted against the concentration of substrate (figure 4.3) to obtain an apparent  $K_M$ , which includes the inhibition constant as expressed in equation (V).

$$\frac{1}{v} = \frac{K_{app}}{V_{max}} \times \frac{1}{[S]} + \frac{1}{V_{max}} \quad (IV)$$

$$K_{app} = K_M + K_M \frac{[I]}{K_i} \quad (V)$$

From equation (V)  $K_i$  can be obtained by plotting the different  $K_{app}$ s obtained at different concentrations of inhibitor against the concentration of inhibitor, figure 4.4.

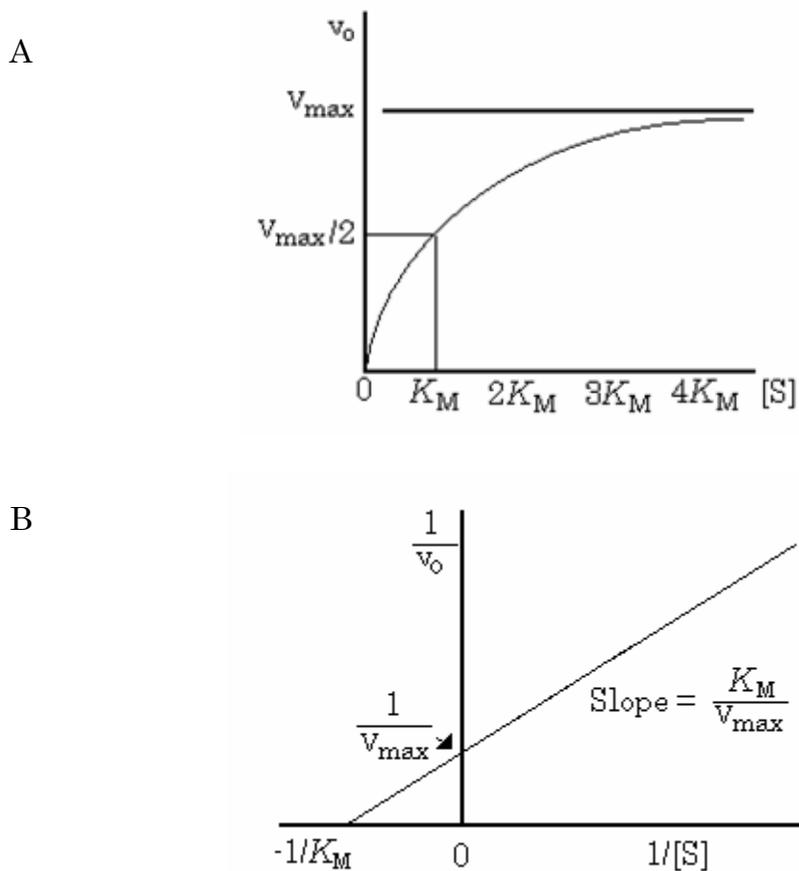


Figure 4.3: A) Michaelis-Menten plot, B) Lineweaver-Burke plot

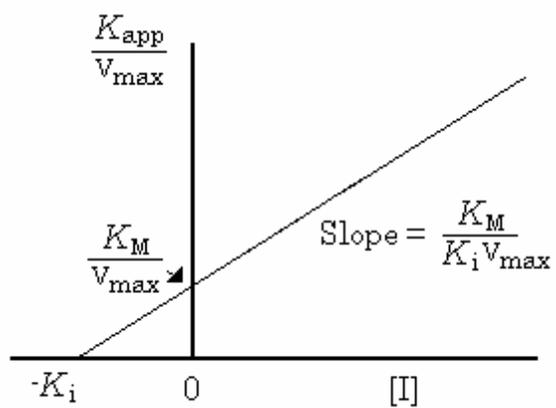
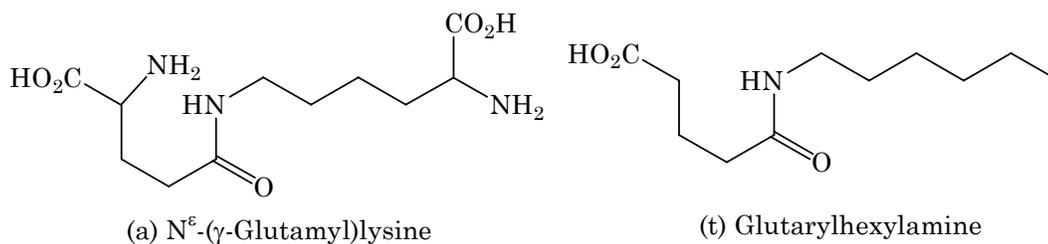


Figure 4.4: A)  $K_{app}$  vs inhibitor concentration plot for the determination of  $K_i$  for competitive inhibition

*$\gamma$ GACT Inhibitors*

Compounds that resemble the substrate but can not be turned over may be competitive inhibitors of the enzyme. Inhibitors of  $\gamma$ GACT were studied by Bowser (1997); he concluded that compounds that mimic a cyclic intermediate formed during the catalytic reaction of  $\gamma$ GACT are potential inhibitors; in addition, the removal of the  $\alpha$ -amino group of the glutamyl portion of the substrate allowed interaction at the active site, increasing inhibition power.

When  $N^\epsilon$ -( $\gamma$ -glutaryl)lysine (compound e) is the inhibitor and  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine (compound a) is the substrate for  $\gamma$ GACT, competitive inhibition was observed giving a graph similar to figure 4.4 with an inhibition constant ( $K_i$ ) of  $30 \times 10^{-6}$  M. Based on these results alkyl analogs of  $N^\epsilon$ -( $\gamma$ -glutaryl)lysine are proposed as good inhibitors of the enzyme; the inhibitor chosen for this work is glutarylhexylamine (compound t) with the objective of making an affinity column. It must be remarked that glutarylhexylamine lacks the  $\alpha$ -amino group required for cyclization of the of the glutamyl moiety; in consequence, the enzyme will bind the molecule but will not be able to turn it over.



Since the kinetic constants are expected to be similar to that of N<sup>ε</sup>-(γ-glutaryl)lysine, the kinetic assay of this inhibitor follows the assay done by Bowser (1997). It must be noted that the dissociation constant of the enzyme-inhibitor interaction is expected to increase when the inhibitor is bound to a solid support (Graves and Wu, 1974).

#### *Glutarylhexylamine Affinity Chromatography Column*

Glutarylhexylamine attached to a solid support can be obtained by modification of the Amersham commercial affinity matrix EAH Sepharose 4B, which is made by the covalent linkage of 1,6-diaminohexane to Sepharose solid support using an epoxy coupling method, see Appendix A for the characteristics of this matrix. This matrix possesses a hexylamine derivative arm spacer to which the glutaryl portion of the inhibitor can be attached by reaction of the amine terminus of the arm spacer and glutaric anhydride, as shown in figure 4.5.

It must be recalled that the enzyme shows less restriction for the amide portion of the substrate, so the connector portion between glutarylhexylamine and the Sepharose matrix (-CH<sub>2</sub>-CH(OH)-CH<sub>2</sub>-O-) is expected not to influence the enzyme-inhibitor interaction to a major extent.

The coupling is monitored by observing the disappearance of free amine group using the fluorescamine reagent. The coupling can be obtained by the method of Bethell et al (1979) using a carbodiimide molecule to

activate one of the reactant groups, or by direct reaction of the anhydride and the amino group.

After attachment of the ligand, the stability of this must be characterized. Ligand leaching is a problem even with the most sophisticated coupling methods; non-covalent binding of the ligand, protein degradation, support degradation, and bond solvolysis contribute to the detachment of the immobilized ligands, even by peptide bonds (Sudesh et al., 1997). In this particular case the whole ligand, glutarylhexylamine, will not be lost, but only the glutaryl portion will leach. The workable pH and temperature range as well as the presence of additive will affect the lifetime of the attached ligand.

#### *Purification of $\gamma$ GACT in Glutarylhexylamine Sepharose 4B*

The separation in affinity chromatography is straightforward, just binding of the target molecule, when contaminants are washed out followed by elution of the target molecule. Conditions for binding must be found; the most desirable procedure for elution is the one that promotes elution of the protein just by adding salt and not other additives or competitive inhibitors.

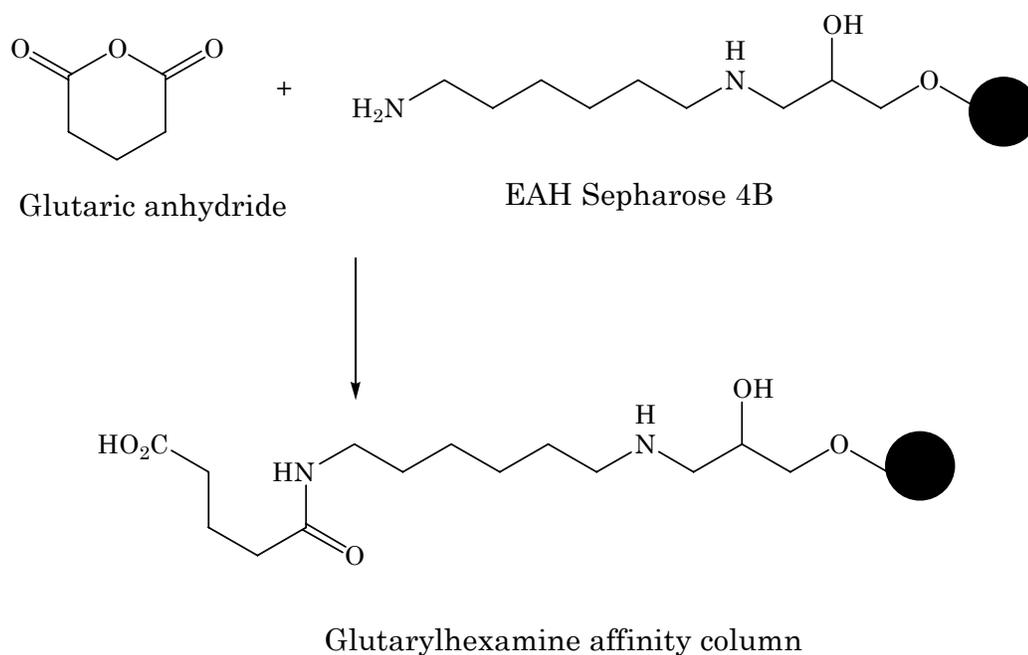


Figure 4.5: Coupling reaction for the production of a glutaryl-hexylamine affinity column for  $\gamma$ GACT

The development of an affinity column for  $\gamma$ GACT not only will make the purification easier but also physical chemical properties of the inhibition can be studied.

## CHAPTER FIVE

### Statement of Purpose

In the study of proteins the feasibility of protein purification limits what can be known. Nevertheless, during the purification valuable information can be obtained about protein properties that can be used to design better purification methodologies; for example, often the catalytic activity of enzymes can be studied with a partially purified protein as long as there is no interference from contaminants. When the purification of a protein from a direct source is too difficult, the tendency is to express the protein in a simpler organism; however, this requires the genetic information of the protein, a knowledge that is most readily obtained when the protein is pure.

Particularly difficult in its purification is  $\gamma$ -glutamylamine cyclotransferase ( $\gamma$ GACT). Partially purified bovine  $\gamma$ GACT was obtained using techniques like preparative ultracentrifugation, anion exchange and size exclusion chromatography, ammonium sulfate precipitation, electrophoresis, isoelectric focusing, ultrafiltration, and lyophilization (Fink and Folk, 1983; Gowda, 1985; Bowser, 1997). Despite numerous attempts,  $\gamma$ GACT has never been purified to homogeneity in sufficient quantities for further analysis. The information gained during the purification, mainly in the polishing steps, is used to obtain the pure protein.

The main objective of this work is to obtain electrophoretically pure protein  $\gamma$ GACT in single band observed by SDS-PAGE. This will be done by improving the existing purification methodology on the basis of the information obtained during the initial purification of the enzyme.

For this work, the purification procedure is similar to that of Fink and Folk (1983) with contributions from Gowda's (1985) and Bowser's (1997) procedures. The following improvements are added to the purification procedure: first, chromatography columns are more powerful, DEAE-Sephacryl S100 instead of DEAE cellulose and Sephacryl S100 instead of Sephadex G100 or G50; second, optimization of intermediate purification steps using Mono Q column for ion exchange chromatography; and finally, isoelectric focusing and SDS-PAGE. Handling among techniques is minimized by using electroelution, cell ultrafiltration, and lyophilization. An electrophoretically pure enzyme is the highest purity that can be obtained using these techniques.

Some properties of the enzyme such as the stability and behavior in presence of reducing agent will be studied in order to improve the purification, especially in the polishing steps of Mono Q ion exchange chromatography and isoelectric focusing, where the enzyme has been shown to be more unstable and susceptible to degradation.

The second objective of this work is to propose a putative gene of bovine  $\gamma$ GACT in the known bovine genome by sequencing the enzyme

primary structure by mass spectrometry and by using bioinformatics tools to determine the best possible candidate for the enzyme gene.

In this work, the N-terminal analysis of the enzyme will be attempted in order to select the best sequencing process (Edman degradation or MS sequencing) for the determination of the primary amino acid sequence of the highest purity enzyme obtained by the described purification procedure.

The pure protein can be used to obtain its amino acid primary sequence, which can be translated into the DNA sequence that will allow the finding of the  $\gamma$ GACT gene in the known bovine genome. In future work, cloning and expression of the  $\gamma$ GACT gene can be attempted for easier purification and study of the enzyme in cells.

The electrophoretically pure enzyme will be sequenced by an external laboratory. Based on the amino acid sequence report the protein most likely to be  $\gamma$ GACT will be proposed by looking at the properties of known and genome-translated proteins and comparing them against the actual properties observed during purification. However, if a gene is found for  $\gamma$ GACT, cloning experiments must be carried out to prove that the protein is indeed the product of the expression of the proposed gene.

Partially purified enzymes have been useful for the study of the catalytic properties of an enzyme. These properties can also be used for the design of better purification methodologies. For example, the design of an affinity column using a reversible inhibitor can shorten the polishing

purification after assuring  $\gamma$ GAACT removal, due to the highly specific interaction of the ligand and the enzyme.

The final objective of this work is the design of an affinity column for  $\gamma$ GACT based on one of its inhibitors, glutarylhexylamine, which resembles the substrate  $N^\epsilon$ -( $\gamma$ -glutamyl) lysine but does not have the  $\alpha$ -amino of the glutamyl moiety. This  $\alpha$ -amine is required for nucleophilic attack on the  $\gamma$ -carbonyl group, which is the proposed first step in the cyclotransferase reaction.

In this work, the percentage of inhibition and the inhibition constant of glutarylhexylamine for bovine  $\gamma$ GACT catalysis using  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine substrate will be calculated. An affinity column with a glutarylhexylamine ligand attached to the solid support, Sepharose 4B, will be synthesized for the capture of  $\gamma$ GACT. The physical properties of this column will be determined in order to know the physical limitations of the column.

PART II  
EXPERIMENTAL WORK

## CHAPTER SIX

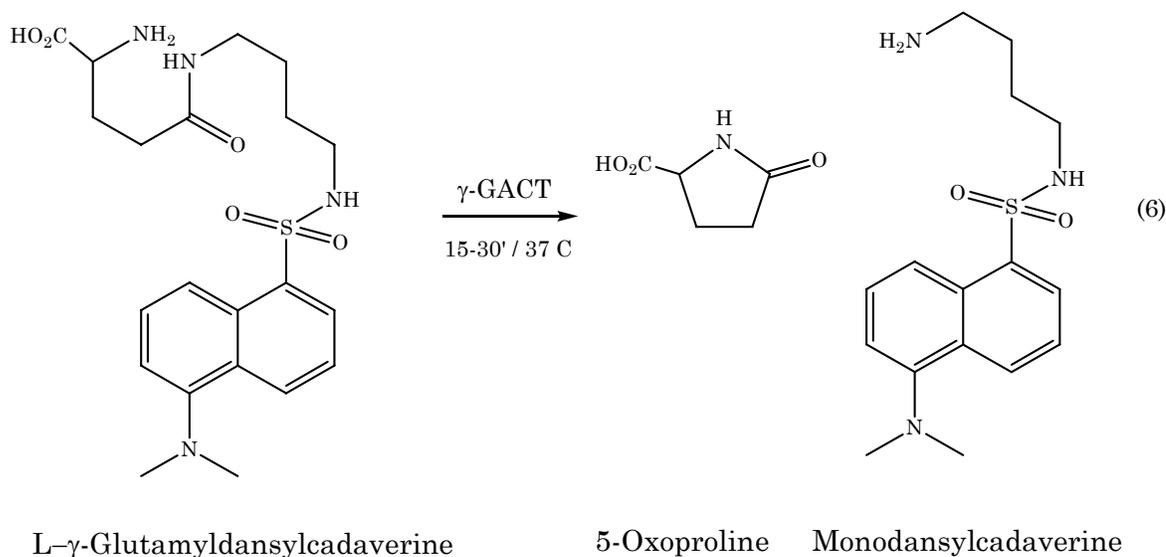
### Experimental Procedure for $\gamma$ GACT Purification

#### *Tracking of Enzymes during Purification*

Given that  $\gamma$ GACT is an enzyme, it was followed during the purification steps by observation of one of the products of the reaction that the enzyme catalyzed. Qualitative and quantitative procedures developed by Fink and coworkers (1980) were used to track the enzyme. The assays of  $\gamma$ -glutamyl transpeptidase ( $\gamma$ GTP) and  $\gamma$ -glutamylamino acid cyclotransferase ( $\gamma$ GAACT), enzymes that can interfere with the measurement of  $\gamma$ GACT activity, were carried out in parallel.

#### *Qualitative Tracking of $\gamma$ GACT*

The qualitative method for the fast visualization of the  $\gamma$ GACT presence employs the fluorescence substrate  $\gamma$ -glutamyl dansylcadaverine. The action of the enzyme on this substrate will release monodansylcadaverine and 5-oxoproline, reaction 6. The product and unreacted substrate are resolved by TLC since the product is more polar than the substrate. This procedure is fast and was used to follow the elution of the enzyme during the chromatography runs. Six samples could be analyzed in approximately 15 minutes.



The following reagents and equipment were employed: saturated  $\gamma$ -glutamyl dansylcadaverine (GDC, synthesized in previous works) in water, saturated monodansylcadaverine (MDC, from Sigma) in water, 1% Pyridine (Sigma) buffer pH 5.4, 1x3 cm polyamide TLC sheets (Avocado), capillary pipettes, 0.2 mL microcentrifuge tubes, UV lamp (UVP model UVL-21 long wave), automatic pipette 0.5 to 10  $\mu$ L with tips (Mettler Toledo, model Volumate), water bath (Thermolyne model 165000 dri-bath) at 37°C, and TLC chamber.

To assay  $\gamma$ GACT, 2  $\mu$ L of dansylcadaverine and 2  $\mu$ L of enzyme were mixed in a microcentrifuge tube and incubated for 10 min at 37°C; using a capillary micropipette, a spot of the mixture was applied at the bottom of the TLC plate. The plate was placed in the TLC chamber and developed until the buffer reached the top; the plate was air dried and visualized under UV light (366 nm).

The result from a typical assay is sketched in figure 6.1. The substrate appears at the bottom of the plate (Rf 0.13); while, the product appears at the top of the plate (Rf 0.3). Active protein is present if a spot of monodansylcadaverine product is observed.

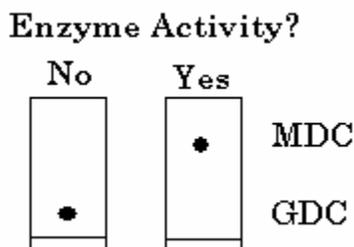
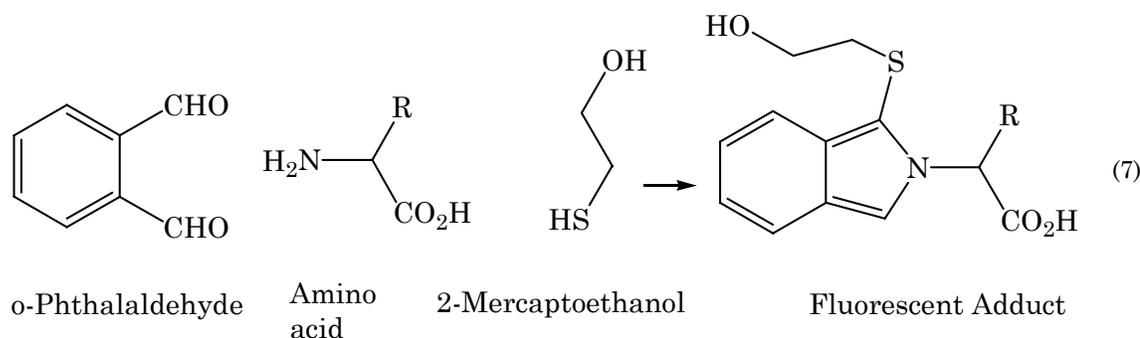


Figure 6.1: Sketch of the results in TLC

#### *Quantification of $\gamma$ GACT*

The quantitative method quantifies the amount of lysine released in the reaction of  $\gamma$ GACT on  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine (reaction 2) during 15 min at 37°C; then, the reaction is quenched by addition of trichloroacetic acid. The quantification was carried out in an amino acid analyzer, where the produced lysine was separated in 6 minutes from the unreacted substrate  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine on an ion exchange chromatography column, post column derivatized with o-phthalaldehyde (OPA) as seen in reaction 7, and detected by fluorescence. The amount of lysine released by the reaction of  $\gamma$ GACT on  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine is used to define the activity of the enzyme; one unit of activity equals one micromole of lysine produced per hour of reaction at 37°C.



The following reagents and equipment were employed: amino acid analyzer with column DC-6A DIONEX (sulfonated polystyrene resin cross-linked 8% divinyl benzene 0.4 x 7 cm), on line fluorometer with excitation filter of 305-395 and emission filter of 430-470 nm (Gilson 121), integrator (Hewlett-Packard, HP 3396A), 0.6 M sodium citrate (Sigma) buffer (with 1ml/L Brij 35 (30% from Sigma) and 1ml/L phenol (Sigma)), 0.6 M potassium borate buffer (with 2.5 mL/L of mercaptoethanol, 2mL/L Brij 35 (30%), and 100 mg/L of o-phthalaldehyde prepared in 5 mL ethanol), reaction buffer 0.2 M NaPi pH 7.5, 10 mM substrate in 0.2 M NaPi pH 7.5 ( $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine (Sigma), 1 mM lysine (Sigma) standard, 20% TCA (Sigma) solution, 0.5 mL microcentrifuge tubes (VWR), microcentrifuge (Eppendorf model 5415 R), syringe of 100  $\mu$ L (Hamilton) and micropipettes 2 to 100  $\mu$ L with tips (Mettler Toledo, model Volumate).

For the assay all buffers were filtered through a 0.2  $\mu$ m nylon membrane (Millipore). Column was at 65  $^{\circ}$ C, citrate buffer at 0.7 mL/min flow rate and 0.4 mL/min borate flow rate; detector was at 0.02 range, and recorder at attenuation 2 and chart speed 0.2; then, 40  $\mu$ L of substrate was

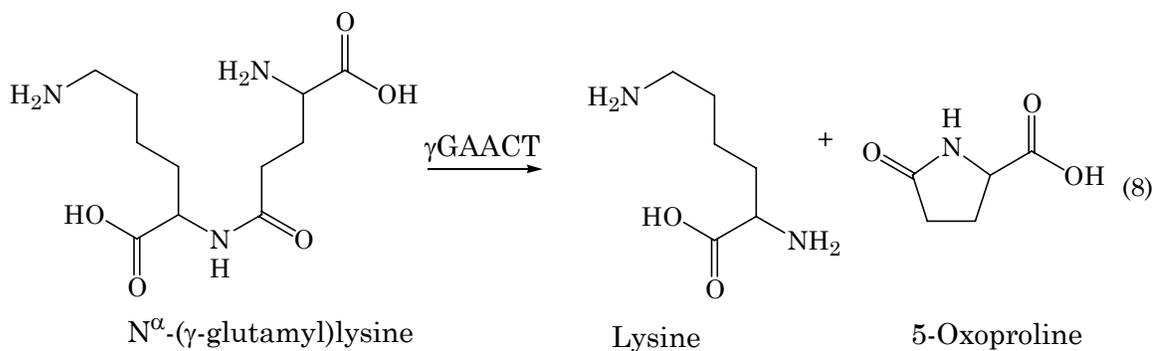
preheated at 37°C for 5 minutes; 20  $\mu\text{L}$  of enzyme was added, and then the mixture was incubated for 15 minutes at 37°C; next, 60  $\mu\text{L}$  of ice cold TCA 20% was added to quench the reaction; this solution was kept on ice for 30 minutes and centrifuged at 13,000 rpm at 4°C for 2 minutes; 25  $\mu\text{L}$  of reaction supernatant were injected as well as 25  $\mu\text{L}$  of lysine standard (10 nmols).

Units are defined as the amount of enzyme able to release 1  $\mu\text{mol}$  of lysine in one hour. Equation (VI) was used to obtain the total units; the calculation was corrected for dilution when needed.

$$\frac{\text{Units}}{\text{mL}} = \frac{\mu\text{mols Lys}}{\text{h x mL}} = \frac{\text{Area Sample}}{20 \mu\text{L Enzyme}} \times \frac{20 \text{ nmols Lys}}{\text{Area Std}} \times \frac{120 \mu\text{L rxn}}{20 \mu\text{L inj}} \times \frac{60 \text{ min}}{\text{h x 15min}} \quad (\text{VI})$$

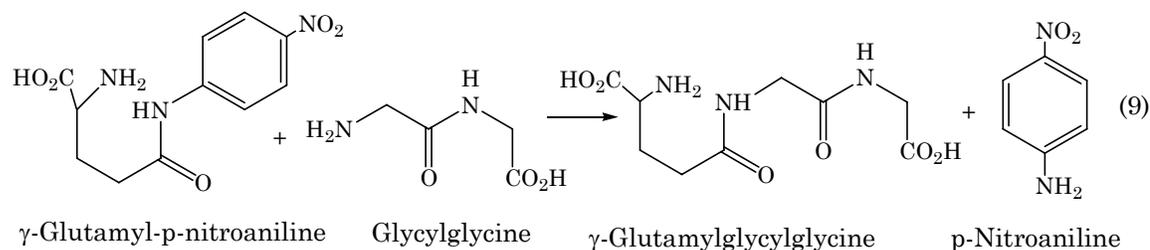
### *Quantification of $\gamma\text{GAACT}$*

The procedure for the determination of  $\gamma\text{GAACT}$  was the same than the one for the quantitative tracking of  $\gamma\text{GACT}$  but the substrate employed for this enzyme was  $\text{N}^\epsilon$ -( $\alpha$ -glutamyl)lysine, reaction 8.



### Quantification of $\gamma$ GTP

This enzyme can act on N<sup>ε</sup>-( $\gamma$ -glutamyl)lysine; the presence of this enzyme is observed by the formation of p-nitraniline from  $\gamma$ -glutamyl-p-nitroaniline as reaction 9 shows.



The following reagents and equipment were employed: 0.2 M Tris/HCl (Sigma) buffer pH 8.5, 0.15 M glycylglycine (Sigma) in Tris buffer, 3 mM  $\gamma$ -glutamyl-p-nitroaniline (Sigma) in Tris/HCl buffer, 1 M HCl (Fisher), spectrophotometer (Beckman DU520), cuvettes, and pipettes 2 to 100  $\mu$ L with tips (Mettler Toledo, model Volumate).

For the assay, the spectrophotometer was set at the wavelength 405 nm. A mixture of 1 mL  $\gamma$ -glutamyl-p-nitroaniline, 0.4 mL glycylglycine, 1.5 mL Tris buffer, and 0.1 mL enzyme (or buffer for the blank) was incubated for 1 minute at room temperature; next, the reaction was quenched by addition of 0.6 mL of HCl; the absorption was measured.

Units are defined as the amount of protein able to release 1  $\mu$ mol of p-nitroaniline in one minute. The molar extinction coefficient ( $\epsilon$ ) of p-nitroaniline at 405 nm is 8800 AU\*cm<sup>-1</sup>\*M<sup>-1</sup>. Equations (VII) and (VIII) were

employed to obtain the total units; samples were corrected for dilution when necessary.

$$\text{Units} = \frac{\mu\text{mols p-Nitroaniline}}{\text{min}} \quad (\text{VII})$$

$$\text{Units} = \frac{\text{cm.M}}{8800 \text{ (AU)}} \times \frac{\text{AU}}{1 \text{ min}} \times \frac{3.6 \text{ mL}}{0.1 \text{ mL ER}} \times \text{EV}(\mu\text{L}) \quad (\text{VIII})$$

where AU, is the absorption of the product at 405 nm, and EV is the total the enzyme volume in microliters. ER is the volume, in milliliters, of enzyme in the reaction that generates the measured product

#### *Bradford Assay for Total Protein Concentration*

The measurement of the total concentration of protein was done using the Bradford assay (Bradford, 1976). Bradford reagent is a Brilliant Blue G dye that shifts its absorption from 465 to 595 nm when bound to proteins; the adsorption at 595nm is proportional to the concentration of protein.

The following reagents and equipment are employed: 50 mM KPi (Sigma) buffer pH 7.5, bovine serum albumin (BSA) stock standard 2 mg/mL (Sigma), and Bradford reagent (Sigma), pipettes 2 to 100  $\mu\text{L}$  with tips (Mettler Toledo, model Volumate) and disposable rest tubes (WVR).

For the assay, 50  $\mu\text{L}$  of BSA standards were prepared at the following concentrations: 0.1 mg/mL (2.5  $\mu\text{L}$  stock BSA + 47.5  $\mu\text{L}$  buffer), 0.5 mg/mL (12.5  $\mu\text{L}$  stock BSA + 37.5  $\mu\text{L}$  buffer), 1mg/mL (25  $\mu\text{L}$  stock BSA + 25  $\mu\text{L}$  buffer). A mixture of 50  $\mu\text{L}$  of sample, blank ( buffer alone), or standard with

1.5 mL Bradford reagent was incubated at room temperature for 15 minute; visible absorption at 595 nm was measured in a plastic disposable cuvette; samples were diluted to the appropriate concentration when needed.

To analyze the results: a calibration curve was obtained by graph of the absorption against the concentration of standard in order to obtain a linear equation; the values of sample absorption were introduced in the calibration equation and solved for the concentration of sample. The sample concentrations were corrected for dilution factors and multiplied by the total volume of sample to obtain the total milligrams of protein.

### *Electrophoresis of Proteins*

Polyacrylamide gel electrophoresis (PAGE) was run in presence of sodium dodecylsulfate (SDS) in order to gain an idea of the composition of the samples. Polyacrylamide 12% gels of 10 x 10 cm (Invitrogen) were run following the manufacture's protocol. Briefly, 0.5 µg of sample was digested with 2 µL of loading buffer (lithium dodecylsulfate from Invitrogen) for 10 min at 75 °C; for samples run under reducing conditions the digestion included 1 µL of reducing agent (DTT from Invitrogen); next, the digests were loaded into the gels, and set in an electrophoretic unit (Amersham, model Hoefer miniVE)) using MES as the electrophoretic buffer; for samples run under reducing conditions, antioxidant (Invitrogen) was added to the cathodic chamber in order to avoid oxidation of reduced samples during the electrophoresis.

The electrophoretic unit was connected to a power supply (Bio-Rad model power pac 300) and run at 200 V for 45 min. Visualization of the proteins was gained by staining the gel with commassie blue dye (Simply Blue safe stain from Invitrogen) following the quick stain instruction from the manufacture. The destained gel was washed with deionized water and stored at 4°C in a well sealed Ziploc plastic bag; gels stored under these conditions remained unaltered for as long as a year. The molecular mass of the protein was calculated by comparison against molecular mass markers (MultiMark from Invitrogen).

### *Gross Purification*

#### *Extraction of the Enzyme from Kidney Tissue*

Frozen kidney was thawed on ice overnight; working at 4°C in a cold room, 50 g of cortex of bovine kidney (from a local slaughter house) were diced and blended with 150 mL of 0.25 M sucrose (Sigma) solution in 5 mM phosphate buffer pH 7.5, for less than 1 min, using a polytron and a 600 mL plastic beaker.

The resulting homogenate solution was poured into 4 polycarbonate tubes (60 mL capacity), which were balanced by mass and placed in a precooled rotor Ti45 ( $r = 6.98$  cm); the homogenate solution was centrifuged in an ultracentrifuge (Beckman model) at 4 °C, at 100,000xg (35,000 rpm) for 55

minutes. The resulting supernatant was filtered through cheesecloth and kept on ice, while the pellets were discarded.

#### *Anion Exchange Chromatography in DEAE-Sepharose*

In a cold room at 4°C, the supernatant from the previous step was adjusted to pH 7.50 with 20% NaOH (Sigma), so  $\gamma$ GACT will be negatively charged since its pI is less than 7.50; this solution was introduced onto a 26 x 4.8 cm DEAE-Sepharose column (Amersham). See Appendix A for detailed characteristics and procedure for the use of this column. The matrix is activated with 10 mM phosphate buffer pH 7.5 with 0.3 M NaCl (buffer B, ~ 28 mS/cm), and then equilibrated in 5 mM phosphate buffer pH 7.5 (buffer A, ~ 1.1 mS/cm). The supernatant was loaded onto the column at the maximum speed of the peristaltic pump (Amersham) ~ 8 mL/min. The sample was washed with ~ 0.5 CV of buffer A. Just when unbound proteins eluted, the fraction collector (Gilson FC203 with an 80 tube rack) and the gradient maker (Amersham) were started; bound proteins were eluted with 1 CV of a linear gradient from 0 to 100 % buffer B at maximum flow rate. Approximately 90 fractions of 9 mL/tube were collected at 1 min/tube; these fractions were kept on ice while measuring pH and conductivity, and testing the presence of activity by the qualitative method; the active fractions were pooled together and kept on ice.

### *Ammonium Sulfate Fractionation and Precipitation*

The active pool from ion exchange chromatography was taken to ammonium sulfate fractionation in order to separate the very insoluble proteins and to ammonium sulfate precipitation in order to store the proteins in a stable way. For the 40% fractionation, 24.5 g of ammonium sulfate (Fisher Scientific) per 100 mL solution was slowly added to the pool of proteins from the DEAE-Sepharose ion exchange chromatography and kept on ice. The solution was left to precipitate for 1h with gentle stirring, and then centrifuged at 28,000xg (15,500 rpm) for 30 min in 10 polycarbonate tubes (capacity 38 mL) in a Sorval RC5B plus Centrifuge precooled at 4° C and equipped with a rotor SG600 (r= 10.45 cm). The resulting pellets were discarded, and the supernatant was made 90% saturated by slowly adding 36.5 g of ammonium sulfate per 100 mL solution. The solution was kept on ice and gently stirred for 4 hours. After this time the solution was centrifuged as previously; pellets were collected and stored at -70°C; while the supernatant was discarded.

### *Intermediate Purification*

#### *Size Exclusion Chromatography on Sephacryl S100*

In a cold room at 4° C, pellets from the ammonium sulfate precipitation were washed into 50 mM phosphate buffer pH 7.5 with 0.15 M NaCl, using ultrafiltration in an ultrafiltration cell (50 mL capacity from Amicon) with a

10K membrane (Millipore) and argon as the pressure gas. This was done in order to decrease the density of the sample; too dense samples do not enter evenly onto the size exclusion column, causing loss of resolution. The retentate from ultrafiltration was added to a 2.6 x 82 cm Sephacryl S100 column (Amersham); see Appendix A for detailed characteristics, calibration, and procedure for the use of this column. The elution of the sample was carried out at 0.2 mL/min for 36 h with collection of 4 mL/tube at 20min/tube in a Gilson fraction collector (model FC203 with an 80 tube rack). The collected fractions were kept on ice and their specific activity was measured; fractions with high specific activity were pooled and lyophilized (Freeze Dryer Virtis) to obtain a powder, which was stored at -70° C.

The sample from the gross purification was subjected to size exclusion and ion exchange chromatography under a variety of conditions.

#### *Intermediate-Polishing Purification Studies on Superdex HR 75 SEC*

In Superdex HR 75 size exclusion chromatography the sample was tested under reducing and non reducing conditions. Under reducing condition polymeric proteins will separate into their lighter momomers. If  $\gamma$ GACT forms a dimer, this will elute as a heavy peak under non-reducing conditions and as a lighter peak under reducing conditions; similarly, the shift of contaminants can be observed.

Under non-reducing conditions, the powder from Sephacryl S100 was restored in cold water and washed by ultrafiltration into running buffer (50

mM KPi pH 7.5 with 0.15 M NaCl). Concentrated sample was centrifuge for 5 min at 13,000 rpm at 4° C, and supernatant solution was injected in a FPLC system (AKTA Amersham) equipped with a Superdex HR 75 (1x30 cm column from Amersham) size exclusion column equilibrated in running buffer; see Appendix A for detailed characteristics, calibration, and procedure for the use of this column. A volume of 0.1 mL of sample (38 units of activity) was loaded and run at 0.2 mL/min; fraction of 0.2 mL were collected on ice and tested for specific activity and purity by SDS-PAGE.

For reducing conditions, similar units of concentrated sample were incubated with 1% 2-mercaptoethanol for 5 min at 37° C, centrifuged, and loaded and onto a column equilibrated in running buffer containing 0.2% 2-mercaptoethanol.

#### *Intermediate-Polishing Purification Studies on Mono-Q IEC*

In order to optimize the best purification of  $\gamma$ GACT on Mono Q GL anion exchange chromatography in a FPLC system (AKTA from Amersham), the sample was tested under different conditions that affect the chromatography and the sample nature. For anion exchange a positive buffer like Tris buffer is recommended; however, in the past (Bowser, 1997) the enzyme proved to be unstable in this buffer; for this reason, the separation on this column was initially attempted in phosphate buffer, which is a negative buffer.

The stability of the sample was also studied in Tris buffer, and different parameters like gradient, flow rate, presence of chelating agent, and reduction and alkylation of cysteine residues were studied in order to obtain a good separation on the Mono Q GL column.

#### *Mono Q GL Separation of $\gamma$ GACT in Phosphate Buffer*

The lyophilized powder from size exclusion in Sephacryl S100 was washed by ultrafiltration into initial buffer; the concentrated sample was centrifuged for 5 min at 13,000 rpm at 4 °C, and the supernatant injected in to a FPLC system equipped with a Mono Q GL (0.5 x 5 cm column from Amersham) strong anion exchanger column equilibrated in initial buffer; see Appendix A for detailed characteristics and procedure for the use of this column. Approximately 0.5 mL of sample was injected, and the elution of the sample was obtained by washing out the unbound protein with 5 CV of initial buffer, followed by a linear gradient from 0 to 100 % elution buffer in 20 CV, which is the recommended gradient for the separation on Mono Q GL column. The runs were carried out at 0.5 mL/min; fractions of 0.2 mL were collected on ice and tested for specific activity.

The phosphate concentration at pH 8 employed was 5 mM of initial buffer (all elution buffers contained 0.3 M NaCl); the phosphate concentration was decrease to 2 mM and then to 0.5 mM KPi. Also, a run was carried out where the concentration of phosphate, and not NaCl, was increased from 1 mM to 100 mM in order to induce elution.

### *Stability of $\gamma$ GACT in Tris Buffer*

In order to run a sample on a Mono Q GL column using Tris as running buffer, the stability of the sample was tested in Tris buffer. Different concentrations of enzyme (39 U/mL, 22 U/mL, and 4.3 U/mL) were incubated on ice with 10 mM Tris pH 7.5 during 15, 45, 75, and 105 min; activity was measured and compared against enzyme controls in 5 mM KPi pH 7.5.

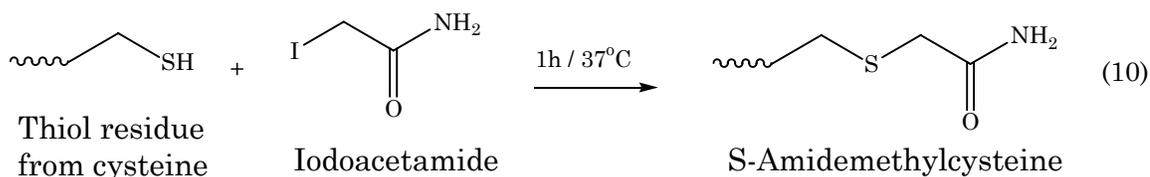
### *Reduction and Alkylation Effects on the Activity of $\gamma$ GACT*

The environment inside the cell is less oxidizing than outside of the cell, so sensitive residues like cysteine are threatened by air oxidation during the purification procedure. The oxidation of cysteine forms disulfide bridges; pI heterogeneity is observed for a protein when this is partially oxidized. In order to homogenize the sample, cysteine residues must be blocked to prevent their oxidation and reduction. Modification of cysteine residues consists in the preparation of the sample with reduction or denaturation, reaction with excess of alkylating reagent, removal of unreacted alkylating reagent, and removal of salts and by-products (Current Protocols in Proteins Science, 2000).

Reduction of the sample is achieved by interchange of the protein disulfide bonds with sulfhydryl groups of a reducing agent such as 2-mercaptoethanol (2-ME) and dithiothreitol (DTT); DTT forms a

intramolecular disulfide ring when oxidized, making it a better reducing agent than 2-ME.

Blockage of cysteine is made by alkylation with alkyl halide; iodoacetamide (IAA) is preferred over iodoacetic acid because the first produces neutral S-carboxamidomethyl cysteine without alteration of the protein environment, while the second produces negatively charged S-carboxymethyl cysteine. Iodoacetamide can modify thiols, carboxylic acids, and the nitrogen of the imidazole ring of histidine. The procedure from (Current Protocols in Proteins Science, 2000) was used for the alkylation of cysteine; this is reaction 10.



The following reagents and equipment are used: 0.2 M iodoacetamide (Sigma), 40 mM DTT (Sigma), 2-ME (sigma), 0.1 M Tris pH 8 (Sigma), amber microcentrifuge tube, water bath (Thermolyne model 165000 dri-bath) at 37°C, 50 – 500  $\mu\text{L}$  of protein solution (100  $\mu\text{g}/\text{mL}$ ) in Tris buffer.

For the assay the protein was mixed with 16  $\mu\text{L}$  of 50 mM DTT to give a 10 fold excess of reducing agent supposing that the protein is ~25 KD and has ~8 disulfide bridges; the mixture was flushed with nitrogen and incubated 1h at 37° C. In the dark, 44  $\mu\text{L}$  of 0.2 M iodoacetamide (a 5 fold excess over total thiols) was added; again the mixture was flushed with

nitrogen and incubated 1 h at 37° C (light can promote the production of iodine from iodide ions). Finally, 5 µL of 2-ME was added to more than 10 fold excess of alkylating agent; the alkylated protein was desalted by washing it with water or buffer, using ultrafiltration. Gentle reduction with 2-ME was done by adding 2-ME to the sample up to 1%, and incubating the mixture at 37° C for 5 min.

#### *Mono Q GL Separation in Tris Buffer*

In order to optimize the separation on the Mono Q GL column, a few screening experiments were carried out; first the sample was loaded at the recommended conditions for separation on this column; afterward, different gradients, buffer concentrations, pHs, and the effect of additives were studied.

The powder from the size exclusion separation on Sephacryl S100 was washed by ultrafiltration into the initial buffer; the concentrated sample was centrifuged for 5 min at 13,000 rpm at 4 °C, and the supernatant injected in a FPLC system equipped with a Mono Q GL column (0.5x5 cm from Amersham) equilibrated in initial buffer. The runs were carried out at the different conditions summarized in table 6.1, flow rate was 1.5 mL/ min and injection volume ~ 0.5 mL; fractions of 0.2 mL were collected on ice and tested for specific activity and SDS-PAGE.

Table 6.1: Optimization of Mono Q GL parameters in Tris buffer

Initial Conditions	Final conditions
Buffer A= 10 mM Tris pH 8, Buffer B=A + 0.3 M NaCl	Buffer A= 20 mM Tris pH 8, Buffer B= A + 0.3 M NaCl
Gradient: 10% B for 5CV, then to 30% B in 30 CV	Gradient: 10% B for 5CV, then to 30% B in 10 CV
Non reduced Sample, 0% 2-ME 0 % EGTA	Reduced Sample, 0.2 % 2-ME 1 % EGTA
Reduced sample	Alkylated sample

### *Isoelectric Focusing*

Isoelectric focusing was run for samples from Mono Q GL and for alkylated samples after Mono Q GL. The samples were washed by ultrafiltration into water; samples should be loaded without salts or other substances that can cause the irregular separation of proteins generating waving pattern.

The following reagents and equipment are employed: Isoelectric focusing apparatus from LKB, water bath (Neslab RTE7 from Thermo corporation), power supply (LKB), ultrafiltration mini cell (3 mL capacity from Amicon) with 10K cutoff membranes (Millipore), tweezers, Ampholine PAGplate gel 5.5-8.5 from Amersham (245 x 11 x 1 mm, with 5% polyacrylamide, 3% crosslinked, and 2.4% ampholine concentration), 5 mL of kerosene, 5 mL of 0.5 M NaOH (Sigma), 5 mL of 0.5 M acetic acid (EM Science), electrodes and sample loading pieces (Amersham), washing solution

(0.1%  $\text{CuSO}_4$  + 10% acetic acid + 30% methanol), staining solution (0.02% Commassie Blue R250 (from emprotech) in washing solution), fixation solution 20% TCA (Sigma), impregnating solution (5% Glycerol (Sigma) in 10% acetic acid), cellophane sheet (Amersham), Centrilon electroelutor (Millipore), centricon with 10 K cutoff filter (Millipore), electrophoretic power supply (power pac 300 from Bio-Rad), electroelution buffer 10 mM KPi pH 7.5.

To prerun: First the electrofusing apparatus was assembled, and the ceramic plate was pre cooled at 10°C for 30 min, after this time kerosene was applied onto the ceramic plate; the gel was placed onto the plate, and all trapped air bubbles were removed. Electrode strip soaked in electrode solutions were applied on the narrow borders of gel: cathode (-) with NaOH and anode (+) with acetic acid solution; electrodes were placed on the strips and plugged to their corresponding socket; the safety lid was replaced. The gel was prerun to generate a pH gradient increasing from anode to cathode, the prerun conditions were 800 V, 50 mA, and 30 W for 1h after this time the pH was verify using a flat pH electrode (Omega model PHE4272).

The centrifuged samples were applied onto the gel using cotton-cellulose loading pieces; 20  $\mu\text{L}$  of 20mg/mL sample. Samples were loaded near their isoelectric point at the cathodic region. Sample was allowed to enter the gel at low voltage to avoid streaking patterns caused by contaminants or precipitation of poorly soluble proteins; the sample entrance

conditions were 800 V, 50 mA, and 30 W for 10 min. Next, the voltage was increased to the running voltage; running conditions were 900 V, 50 mA, and 30 volts for 5 hours; after this time the power was disconnected, and pH was measured with a flat pH electrode. Finally the power was reconnected, and the bands were sharpened at sharpening conditions: 1000 V, 50 mA, 30 W, for 10 min.

After focusing one strip along the long side of the gel was cut for developing; the rest of the gel was taken for extraction of the protein. For developing, the gel strip was placed in TCA 20 % for 10 min to precipitate and fix the protein; then it was placed for 2 min in washing solution, and 15 min in staining solution with microwaving for 40 seconds. The gel was destained overnight in washing solution. After this, gel was prepared for storage by impregnating for 30 min with impregnating solution followed by wrapping with a cellophane sheet wetted with impregnating solution; the gel in this condition was left to air dry at room temperature.

For extraction of the pure enzyme the rest of the gel was cut into 0.5 cm wide strips along the individual pH regions; the gel strips were placed in a sample tube for electroelution, this last in turn was placed in a centricon device and loaded in the electroelutor; the gel strips were electroeluted at 20 V for 2 h in phosphate buffer.

In order to remove ampholytes, the extracted solution recovered in the centricon was washed by ultrafiltration for a final volume of 0.1 mL. The

extracted proteins were tested for specific activity and on SDS-PAGE. The pI of the protein is the pH measured with the pH electrode.

## CHAPTER SEVEN

### Experimental Procedure for Amino Acid Sequencing

#### *N-Terminal Determination*

The initial step in the N-terminal determination was the determination of the detection limit of dansylated amino acids by reversed-phase HPLC. Dansylated amino acids have different fluorescence, so the detection limit of each amino acid is different. The observation of a particular amino acid is also limited by the hydrolysis step; longer hydrolysis time is need for the liberation of aliphatic amino acids like valine, leucine, and isoleucine; while serine, threonine, and tryptophan are easily destroyed at hydrolysis conditions; also glutamine and asparagine are converted into their respective acids.

The single band obtained from isoelectric focusing extracted from the gel (LeGender and Matsudaira, 1989) was subjected to N-terminal determination, and run in parallel with carbonic anhydrase (Sigma) as a positive control protein.

#### *Dansylation*

The dansylation procedure was adapted from the one reported by Negro et al. in 1987. The dansylation reaction is shown in figure 7.1.

The following reagents and equipment are employed: 0.5 M NaHCO<sub>3</sub> (Fisher), 20 mM Dansylchloride (1-Dimethylamino-naphthalen-5-sulfonyl chloride, from Sigma), 40 µg amino acid standards (Sigma) in water, amino acids sample (free amino acids or polypeptides) amber microcentrifuge tubes, water bath (Thermolyne model 165000 dri-bath) at 65° C, and micropipettes with tips (Mettler Toledo model Volumate).

Assay: in an amber microcentrifuge tube, mix 0.1 mL of each reagent: NaHCO<sub>3</sub>, dansylchloride solution, and amino acid solution. Let the mixture react in the dark at 65° C for 40 min. To analyze the results, dansylated amino acids are resolved by HPLC; while, the dansylated polypeptides are subjected to hydrolysis before HPLC.

#### *Hydrolysis of Dansylated Polypeptide*

Hydrolysis of the dansylated sample was done using the protocol recommended by Pierce. Figure 7.1 shows the hydrolysis as the second step in N-terminal determination.

The following reagents and equipment are used: hydrolysis tubes with teflon plug (Pierce), vacuum and nitrogen purge lines, micropipettes with tips, Pasteur pipettes, di and tripeptides standards (Sigma), and Centrivap (LABCONCO).

To assay: mix 1 to 2 mg of peptide (or dansylated polypeptide) with 50 µL of 6 M HCl; using a long Pasteur pipette, introduce the mixture into the hydrolysis tube; insert the Teflon plug but do not tight it. Evacuate oxygen

by alternating nitrogen and vacuum; leave the tube in vacuum and close the plug tightly. Heat the tube in an oven at 110°C for 24 hours. After this time, let the tube cool down to room temperature. Take the solution into an amber microcentrifuge tube, and centrifuge-dry in a centrivap. To analyze the results, restore the dried sample in running buffer and analyze by HPLC. Tryptophan is destroyed during this procedure.

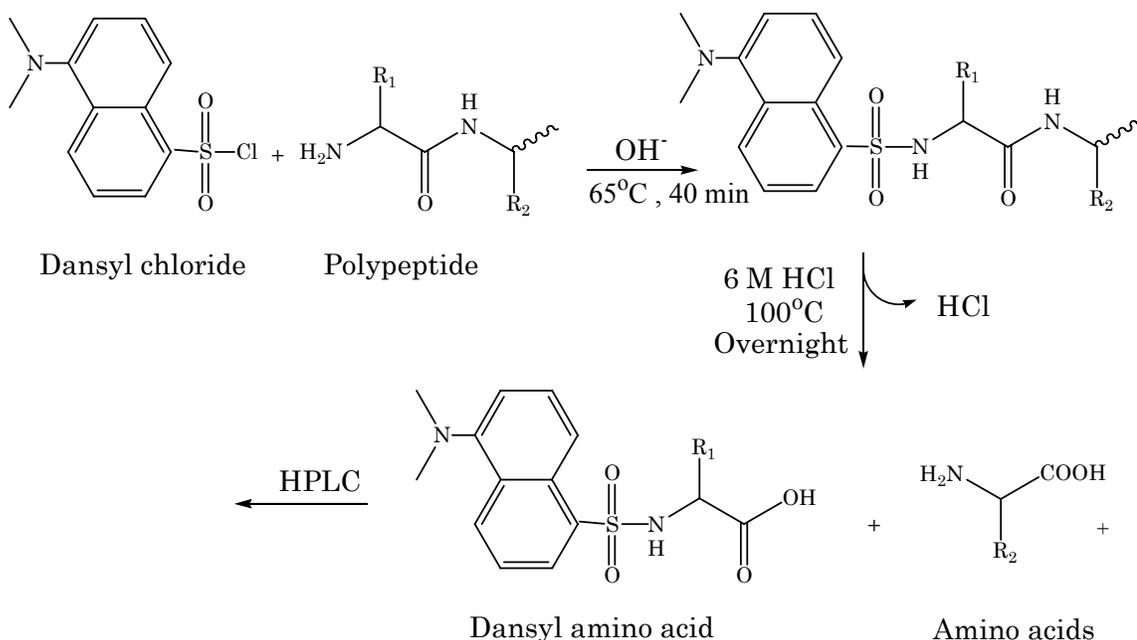


Figure 7.1: Dansylation and hydrolysis of the N-terminal primary amine

### *Dansylation Analysis by HPLC*

The analysis of dansylated amino acids was adapted from the work of Oray and coworkers (1983). For this procedure the following reagents and materials were employed: HPLC with UV detector (Hewlett-Packard), on line fluorometer with excitation filter of 305-395 and emission filter of 430-470

nm (Gilson 121), C18 reverse phase column (Beckman Ultrasphere ODS 4.6 x 150 mm), 50  $\mu$ L syringe with blunt tip (Hamilton), 200 pmoles of dansylated amino acid standards in acetone (Sigma), solvent A = 30 mM NaPi buffer pH 6.5, solvent B = Acetonitrile 100 %, microcentrifuge tubes, bench top microcentrifuge, and micropipettes with tips.

Table 7.1 summarizes the parameters for the determination of dansylamino acids. The column was equilibrated in 10 % solvent B; the sample was centrifuged for 5 minutes and the supernatant was injected. The retention times of the sample peaks were compared against the retention times of dansylamino acids standards (Sigma).

Table 7.1: HPLC Conditions for the determination of Dansyl amino acids

Dansylamino acids standards	1-10 pmols
Solvent A	30 mM NaPi buffer pH 6.5
Solvent B	Acetonitrile 100 %
Injection volume	20 $\mu$ L
Gradient	from 10% B to 22% B in 45 minutes, keep 22% B for 13 minutes, from 22% B to 40% B in 20 minutes and keep 40% B for 12 minutes
Flow rate	1 mL/min
Temperature	Room temperature
Detector settings	Range 0.005
Recorder Settings	Attenuation <sup>1</sup> , chart speed 0.2 Threshold 3, peak width 0.60 Area reject 0

For the run, column should be equilibrated in 10 % solvent B. Centrifuge the sample for 5 minutes and inject 20  $\mu$ L. Run under the following conditions: gradient: from 10% B to 22% B in 45 minutes, keep 22% B for 13 minutes, from 22% B to 40% B in 20 minutes and keep 40% B for 12 minutes; flow rate: 1 mL/min; fluorometer: range 0.2, recorder: attenuation 3, chart speed: 0.2; room temperature.

For the analysis of the results: compare the retention time of the sample peaks against the retention time of the standards. If a polypeptide is dansylated and then hydrolyzed, only the N-terminal amino acid would show a fluorescent peak; however, the following peaks are also observed due to the nature of the mixture: dansylsulfonic acid, dansylammonium, o-dansylthreonine, and  $\epsilon$ -dansyllysine. This last peak can also be employed to verify that dansylation has taken place in a protein sample because the N-terminal amino acid can be blocked and show no dansylation, but at least one lysine residue is present in most proteins.

#### *Preparation of the Sample for MS Sequencing*

In order to verify its purity, the sample after isoelectric focusing was run in a SDS-PAGE as described in the section electrophoresis of proteins of the purification experimental procedure. The enzyme band from SDS-PAGE was submitted for MS sequencing. The following are the considerations for the preparation of a good gel band for MS sequencing: as much as possible protein should be committed in a gel well in order to obtain a high protein/gel

ratio; the heavier the protein the darker the band should be because the molar quantities would be smaller. The preferred gel thickness is 1 mm, and non-fixing staining methods are recommended; also destaining of the gel must be done until a clear background is obtained. Contamination should be minimized by employing reagents of the highest purity available; also, keratin contamination, which can not be avoided, must be minimized by minimizing the handling steps after the gel has been run.

The electrophoretic band of the enzyme obtained taking into account all these considerations was excised tightly from the gel, washed twice with 50% acetonitrile and frozen at  $-70^{\circ}\text{C}$  for storage. These washing steps were also done for an equivalent protein-free area of the same gel in order to obtain a background control.

#### *MS Sequencing by Harvard Microchemistry and Proteomics Analysis Facility*

The sample submitted to Harvard Microchemistry and Proteomics Analysis Facility Proteolytic was sequenced using mass spectrometry (Taniguchi et al. 2002). First, proteins were in gel digested with Trypsin; the obtained fragments were separated by microcapillary reverse-phase HPLC and directly injected into the mass spectrometer; the fragments were subjected to nano electrospray ionization (ESI) and analyzed in an ion trap and/or orbitrap mass spectrometer. The obtained MS/MS sequence spectra were analyzed using the algorithm Sequest (Eng et al. 1994) and programs developed by Harvard microchemistry (Chittum et al. 1998), and manually

reviewed in order to obtain the possible proteins present in the sample.

Figure 7.2 summarizes the procedure for MS sequencing in this work.

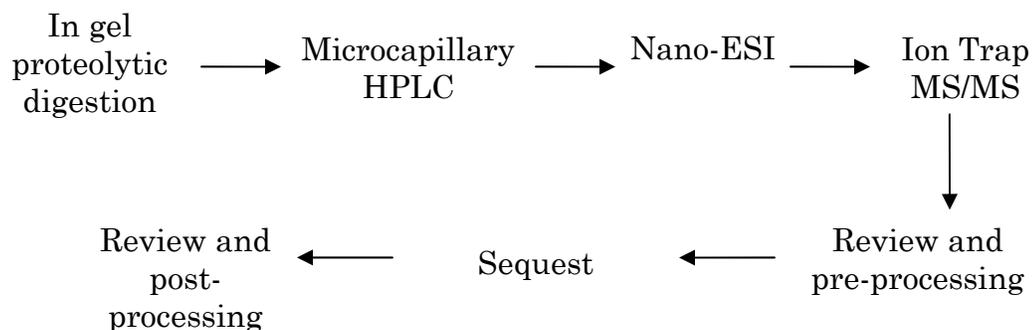


Figure 7.2: Scheme of the MS sequencing of  $\gamma$ GACT

### *Analysis of Reported Proteins*

A report of all possible proteins present in the sample was the result of the MS sequencing; from these reported proteins the one most likely to be  $\gamma$ GACT was selected. To accomplish this, the characteristics of the reported proteins were compared with those of  $\gamma$ GACT.

Initially the complete sequence of the proteins was withdrawn from the National Center for Biotechnology Information (NCBI), and with these sequences the predicted molecular mass and pI were calculated using calculators like the Protein Calculator V3.3 from the Scripps Research Institute. This protein calculator estimates the charge of the protein at a given pH using the Henderson-Hasselbalch equation with the following pKa values for charged residues: N-terminus 8.0, C-terminus 3.1, Lys 10.0, Arg 12.0, His 6.5, Glu 4.4, Asp 4.4, Tyr 10.0, and Cys 8.5. As the number (n) of

these amino acids present in the protein is known, the charge is calculated at different pH values until the pH at which the total charge of the protein is zero, the pI value. After this, the obvious contaminations were discarded and the best enzyme candidates were chosen.

The secondary structure for the best candidates was predicted using the SwissModel (Schwede et al., 2003 and Guex and Peitsch, 1997). The structure of the proteins employed by the SwissModel to predict the candidate's structure was obtained from the NCBI and the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (Berman et. al., 2000). The secondary structure was described using the Structural Classification of Proteins (SCOP) (Murzin et. al., 1995). Finally, structural similarities and the mechanism of these candidate enzymes were compared to that of  $\gamma$ GACT and related enzymes (table 1.1 and 1.2).

The chosen candidates were sent to Dr. Baker (Baylor University) to compare with the human and plant glutaminy cyclase enzymes based on the observed similarities of their catalytic mechanism with that of  $\gamma$ GACT. Various pairwise alignments, BLAST and PSI-BLAST, and EST surveys as well as a 3D similarity tool for secondary and tertiary structure similarity were run for these proteins.

## CHAPTER EIGHT

### Experimental Procedure for the Affinity Column Design

#### *Determination of Ligand Inhibition Constant*

An estimated  $K_i$  of the inhibitor of the enzyme is required in order to verify that the inhibitor will be a suitable ligand for affinity chromatography. The dissociation constant of the enzyme-inhibitor complex ( $K_i$ ) should be in the range of  $10^{-4}$  to  $10^{-8}$  M during the capture and close to  $10^{-2}$  M during the elution.

#### *Percent of Inhibition by Glutarylhexylamine*

To determine the inhibition power of glutarylhexylamine, the enzyme was quantified in presence of this inhibitor at an equal concentration to that of the substrate. Glutarylhexylamine was synthesized (Sam Chen); the reagents and equipment for the quantitative method for tracking the enzyme were employed. A mixture of 20  $\mu$ L of 20 mM phosphate buffer pH 7.50, 12  $\mu$ L of 5 mM inhibitor (glutarylhexylamine), and 12  $\mu$ L of 5 mM substrate (N<sup>ε</sup>-( $\gamma$ -glutamyl)lysine) was preheated at 37°C for 5 minutes; 5  $\mu$ L of 1 U/mL of enzyme was added, and then the mixture was incubated for 20 minutes at 37°C; after this time, 60  $\mu$ L of ice cold TCA 20% was added to quench the reaction; this solution was kept on ice for 30 minutes and centrifuged at 13,000 rpm at 4° C in a microcentrifuge (eppendorf 5414R) for 2 minutes.

Aliquots of the reaction were assayed as describe in the quantification of  $\gamma$ GACT method.

The percentage of inhibition was calculated using equation (IX)

$$\% \text{ Inhibition} = \left( 1 - \frac{\text{Activity with Inhibitor}}{\text{Activity of substrate alone}} \right) \times 100 \quad (\text{IX})$$

#### *Determination of $K_i$ for Glutarylhexylamine*

For the determination of the  $K_i$  of glutarylhexylamine in the inhibition of  $\gamma$ GACT, the reaction of the enzyme was carried out over six concentrations of the substrate at five different concentration of inhibitor. The range of inhibition concentrations employed was the same employed for Bowser (1997) for a similar inhibitor, glutaryl- $\epsilon$ -lysine, and by Gonzalez (2005) for glutarylhexylamine; these inhibitor concentrations are  $0.2K_i$  to  $5K_i$ . Table 8.1 summarizes the concentrations of substrate and inhibitor employed.

The reagents and equipment for the quantitative method for tracking the enzyme were employed. A mixture of 20  $\mu$ L of 20 mM phosphate buffer pH 7.50, 15  $\mu$ L of inhibitor (glutarylhexylamine), and 20  $\mu$ L of substrate (N <sup>$\epsilon$</sup> -( $\gamma$ -glutamyl)lysine) was preheated at 37°C for 5 minutes; 5  $\mu$ L of 1 U/mL of enzyme was added, and then the mixture was incubated for 20 minutes at 37°C; after this time, 60  $\mu$ L of ice cold TCA 20% was added to quench the reaction; this solution was kept on ice for 30 minutes and centrifuged at 13,000 rpm at 4° C in a microcentrifuge (eppendorf 5414R) for 2 minutes.

Aliquots of the reaction were assayed as describe in the quantification of  $\gamma$ GACT method.

Table 8.1: Concentration of substrate ( $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine) and inhibitor (glutarylhexylamine) in the reaction for the determination of  $K_i$

[I] (mM)	I <sub>0</sub> (0)	I <sub>1</sub> (0.03)	I <sub>2</sub> (0.75)	I <sub>3</sub> (0.15)	I <sub>4</sub> (0.225)	I <sub>5</sub> (0.3)
S1 (mM)	0.05	0.05	0.05	0.05	0.1	0.1
S2 (mM)	0.10	0.1	0.1	0.1	0.5	0.5
S3 (mM)	0.5	0.5	0.5	0.5	0.75	1
S4 (mM)	1	0.75	0.75	1	1.2	2
S5 (mM)	2	1.2	1.5	2	2.5	3

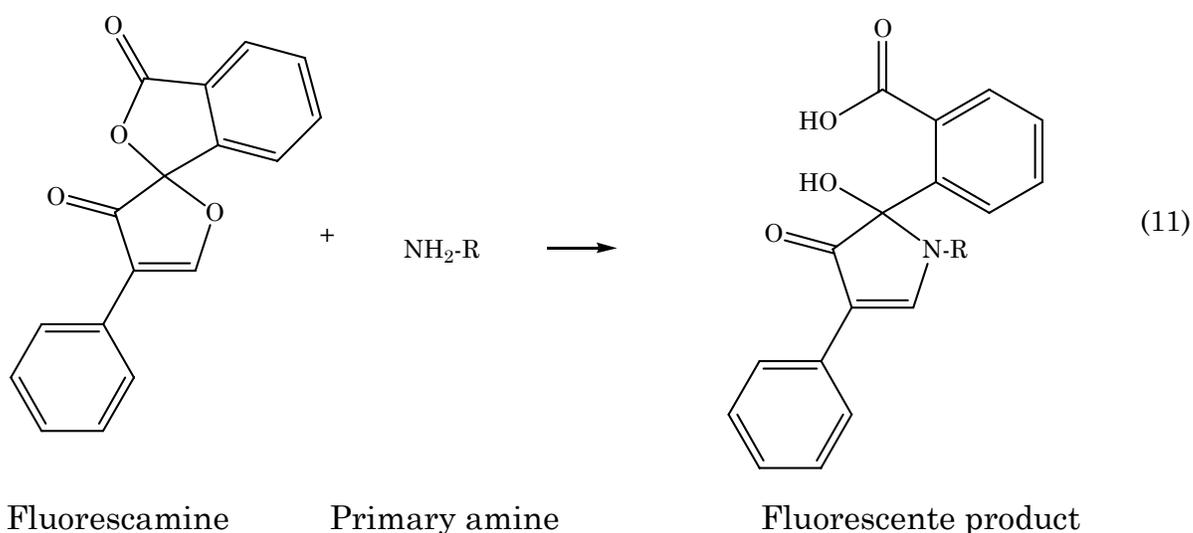
The concentration of product per hour represents the velocity of the reaction; the obtained velocities were analyzed by plotting the data as a double-reciprocal form of the Michaelis-Menten equation or the Lineweaver-Burke equation, while  $K_i$  can be obtained by plotting the different  $K_{app}$ s obtained at different concentrations of inhibitor against the concentration of inhibitor.

### *Synthesis of Glutarylhexylamine Affinity Column*

#### *Tracking of the Free Primary Amine*

The formation of glutarylhexylamine affinity column was monitored by the disappearance of the free primary amine using the fluorescamine reagent. The formation of the product consumes free primary amine; this fact

can be used to attack the advance of the reaction. One of the most sensitive and practical methods for observing primary amines is the use of fluorescamine (Udenfriend et al., 1972). This non-fluorescent reagent forms an instantaneous fluorescent product (390 nm excitation and 475 nm emission) with primary amines, reaction 11 with sensitivity in the range of picomoles.



The following reagents and equipment were employed: 0.03% fluorescamine (Roche) in acetone, 0.2 mL microcentrifuge tubes, UV lamp (UVP model UVL-21), automatic pipette 0.5 to 10  $\mu\text{L}$  with tips (Mettler Toledo model Volumate), sucrose (Sigma), and glycylglycine (Sigma).

To assay, the sample was mixed in a microcentrifuge tube with 0.1 mL of fluorescamine reagent; the mixture was let stand at room temperature for 5 min, and then visualized under UV light (366 nm). The observable

fluorescence indicates the presence of free primary amine; sucrose was used for a negative blank and glycyglycine as a positive blank.

### *Coupling Methods*

Two coupling methods were compared; the first one was the manufacture recommended coupling between a carboxylic acid and the terminal amine from the solid support using the carbodiimide method (Seghal and Vijay, 1994). The second method tried was the coupling of the anhydride with the free amine from the solid support (Sudesh and Lyddiatt, 1997). Figures 8.1 and 8.2 show the coupling using the carbodiimide method and the anhydride method, respectively.

A carbodiimide can activate a carboxylic acid and make it suitable for attack by the amine; carbodiimide should be in excess over the ligand, and the obvious disadvantage of this method when using a diacid molecule is the possibility of a second carbodiimide activation and crosslinkage of the solid support, figure 8.3.

The coupling method that employs an anhydride will form the glutarylhexylamine ligand in a nucleophilic attack of the amine on one of the carbonyl carbons of the anhydride; the formed carboxylic acid is not active enough to react with another amino group, so crosslink formation is less likely to happen.

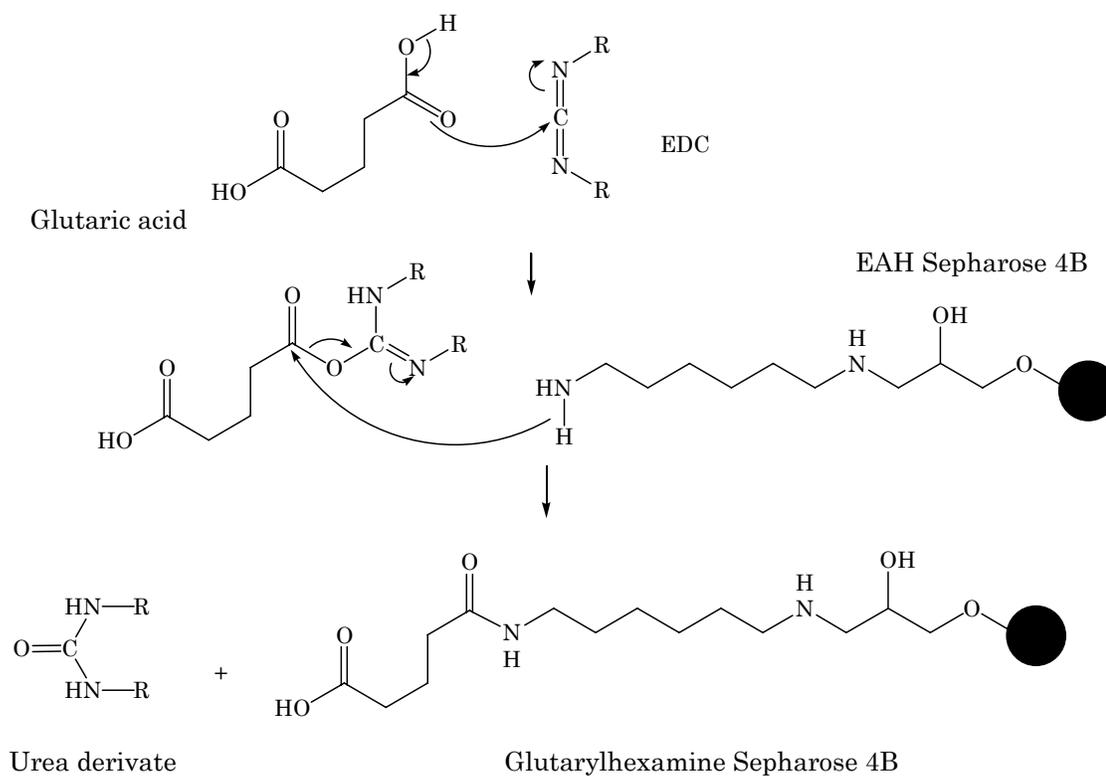


Figure 8.1: Glutaric acid and hexylamine coupling employing a carbodiimide

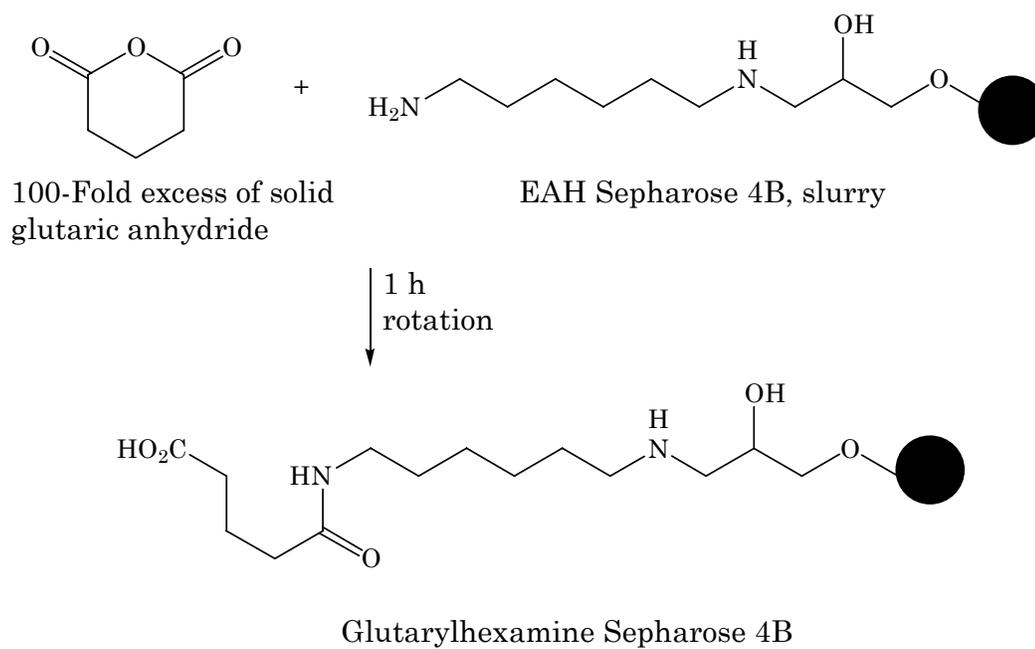


Figure 8.2: Glutaryl anhydride and hexylamine coupling

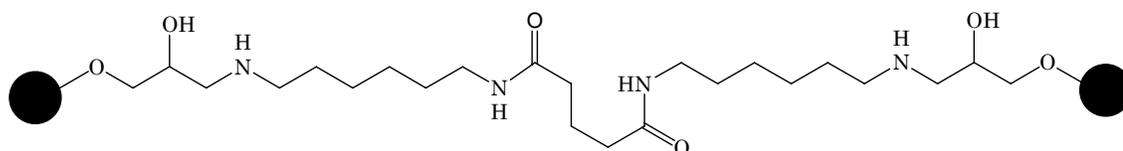


Figure 8.3: Possible crosslink in affinity column synthesis

The following reagents and equipment were employed: Roto-shaker (Scientific Industries), disposable chromatography column 30  $\mu\text{m}$  porous (Bio-Rad) connected to suction, 15 mL plastic tube (WVR), 1 mL EAH Sepharose 4B (Amersham) equal to 10  $\mu\text{mol}$ s of active groups (see Appendix A for the characteristics of this matrix), 100-fold excess of glutaryl anhydride (Acros) over active groups on matrix, this is 1.2 mmols (137 mg), 100-fold excess of glutaric acid and solid N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC, from) for a final 0.1 M concentration in reaction volume, distilled water pH 4.5 (HCl adjusted), 80 mL NaCl 0.5 M, 0.1 M and 0.5 M NaOH, 0.1 M acetate buffer pH 4 with 0.5 M NaCl, 0.1 M tris buffer pH 8 with 0.5 M NaCl.

For the assay, 1 mL of EAH Sepharose 4B slurry was taken into a plastic disposable column and washed with distilled water pH 4.5 followed by 80 mL of 0.5 M NaCl. The washer matrix was transferred into a plastic tube with 1 column volume of 0.5 M NaCl to prevent ionic adsorption. For the coupling using the EDC method the glutaric acid solution was added and the solid EDC was added in parts; while for the coupling using the anhydride method, solid glutaryl anhydride was added. The mixture was gently rotated

at room temperature for 12 h; the pH was measured every 5 min during the first hour and adjusted to pH 6 using 20 % NaOH. Aliquots of the reaction were removed for tracking of the reaction with fluorescamine reagent. The matrix was washed in a disposable column with alternating buffers pH 4 and pH 8, then with water. For storage the matrix was kept in 20% ethanol with 0.5 M NaCl.

#### *Characterization of the Glutarylhexylamine Affinity Column*

After coupling, ligand leaching must be characterized in order to know the workable range of the parameters during separation as well as for storage and cleaning of the column. For the determination of these safe conditions, the derivatized solid support was incubated for 1 h at pH 2, 5, 7, 9, and 12; temperatures of 4, 15, 25, 35, and 45 °C; and NaCl concentrations of 0, 0.1, 0.25, 0.5, and 1 M. The release of ligand was observed as the fluorescence formed by addition of fluorescamine reagent.

#### *Purification of $\gamma$ GACT in Glutarylhexylamine Sepharose 4B*

The following reagents and equipments were employed: Glutarylhexylamine Sepharose 4B affinity, empty column shirt of 0.5 x 5 cm, 0.05 to 1 M NaCl in 10 mM phosphate buffer pH 7.5, 20 % methanol with 0.5 M NaCl, FPLC system (AKTA Amersham), 1 mL syringe (Hamilton).

In order to bind the protein it is needed to know if the protein will be able to be recovered; hence, in the initial attempt of chromatography

separation the sample was loaded in a buffer containing 1M NaCl. For the binding of the enzyme, the sample was loaded in buffer containing 0.05 M NaCl and for elution the salt concentration was increasing to 1 M. Fractions of 0.2 mL were collected on ice and assayed for enzyme activity as described in the method for quantification of  $\gamma$ GACT.

PART III

RESULTS AND DISCUSSION

## CHAPTER NINE

### Results and Discussion of the Purification of $\gamma$ GACT

#### *Results of the Gross Purification*

An average 4,498 units of activity and 7.2 g of proteins were extracted from bovine kidney. After centrifugation approximately 1,000 units were found in the pellet; this activity is attributed to the action of  $\gamma$ -glutamyl transpeptidase on the substrate N<sup>ε</sup>-( $\gamma$ -glutamyl)lysine; while the 3,262 units recovered in the supernatant are attributed to the action of  $\gamma$ GACT. The main loss in activity at this point is attributed to  $\gamma$ GACT degradation by the action of proteases.

From DEAE-Sepharose ion exchange chromatography 1,350 units and 520 mg of proteins were recovered. A typical activity elution profile in this column and the quantitative assay of this chromatography run are shown in figure 9.1; the recovery from this column was 41%.

Fractionation and precipitation of the enzyme with ammonium sulfate showed complete precipitation of the protein because no activity could be detected in the supernatant. 87 % was the average recovery from this step.

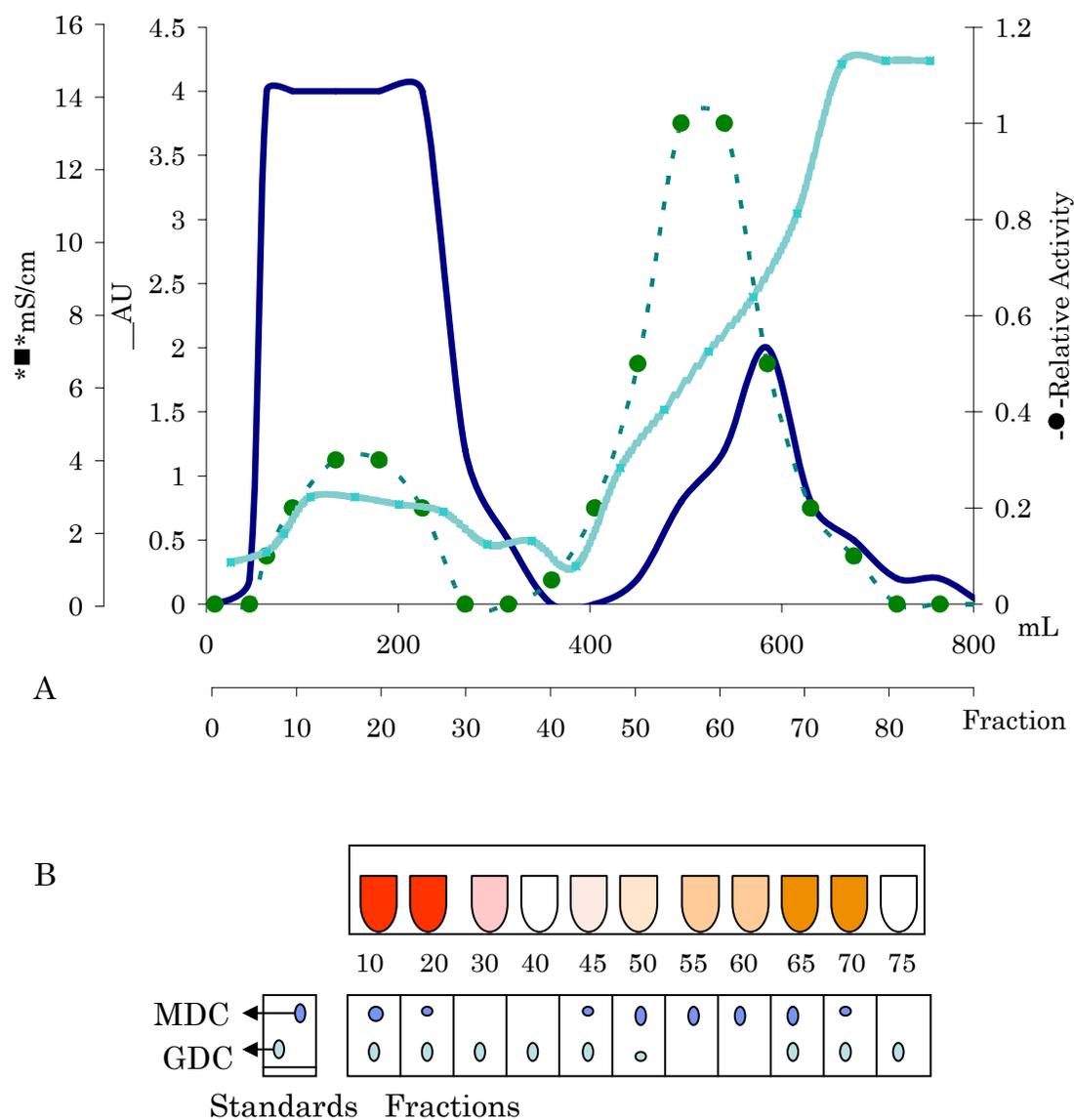


Figure 9.1: Typical results from DEAE-Sepharose. A) Elution profile of  $\gamma$ GACT purification on DEAE-Sepharose. Solid line represents the UV absorption at 280 nm; broken line represents the units of activity of  $\gamma$ GACT. The squares show that the activity elutes close to 5 mS/cm. B) TLC assay

The sample recovered from size exclusion chromatography showed an average of 694 units of activity and 12 mg of protein. The recovery from this column was 59%. This pool showed no activity of  $\gamma$ GAACT; the activity

elution profile from this column is shown in figure 9.2. It can be seen that the major colored contaminants elute completely before activity begins to elute. Sephacryl S100 is described by equation (X); see Appendix A for the details on this calibration.

$$MW_{S100} = 10^{(-0.0069 \text{ mL} + 3.3347)} \quad (\text{X})$$

Using this equation, it can be seen that contaminants removed by Sephacryl S100 chromatography have molecular masses between 200 and 56 KD, since their elution volumes are between 150 and 230 mL. Activity was observed in the range of molecular masses from 35 to 17 KD, with maximum activity for the molecular mass of 25 KD, which corresponds to a elution volume of 280 mL.

Although heavy proteins should elute first during size exclusion, sometimes non specific interactions between the sample and the matrix of the column occurs; in order to diminish this effect 0.15 M NaCl was added to the buffer. However, in this experiment a heavy molecular mass protein of ~ 40 KD was observed to elute at low molecular masses; figure 9.2 shows the electrophoresis of the fractions from size exclusion in Sephacryl S100. Even in fraction 105 where the main protein size is 8 KD, the 40 KD contaminant is observed.

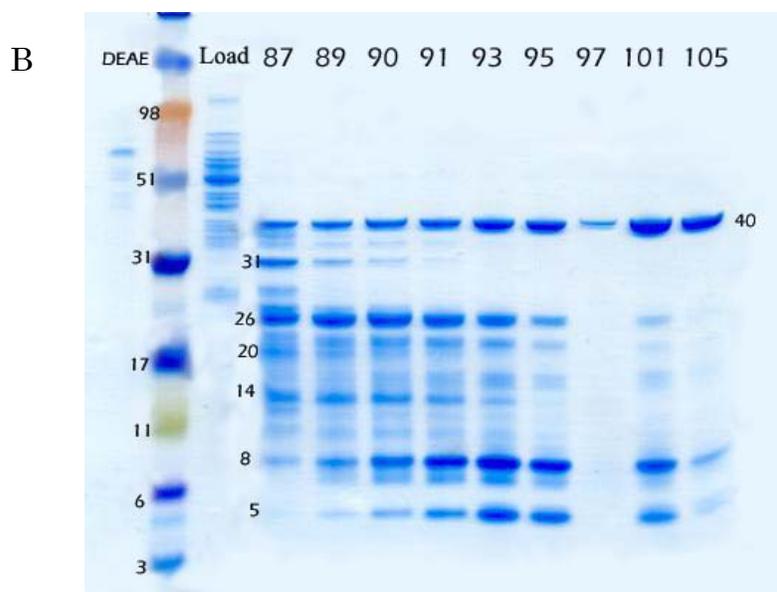
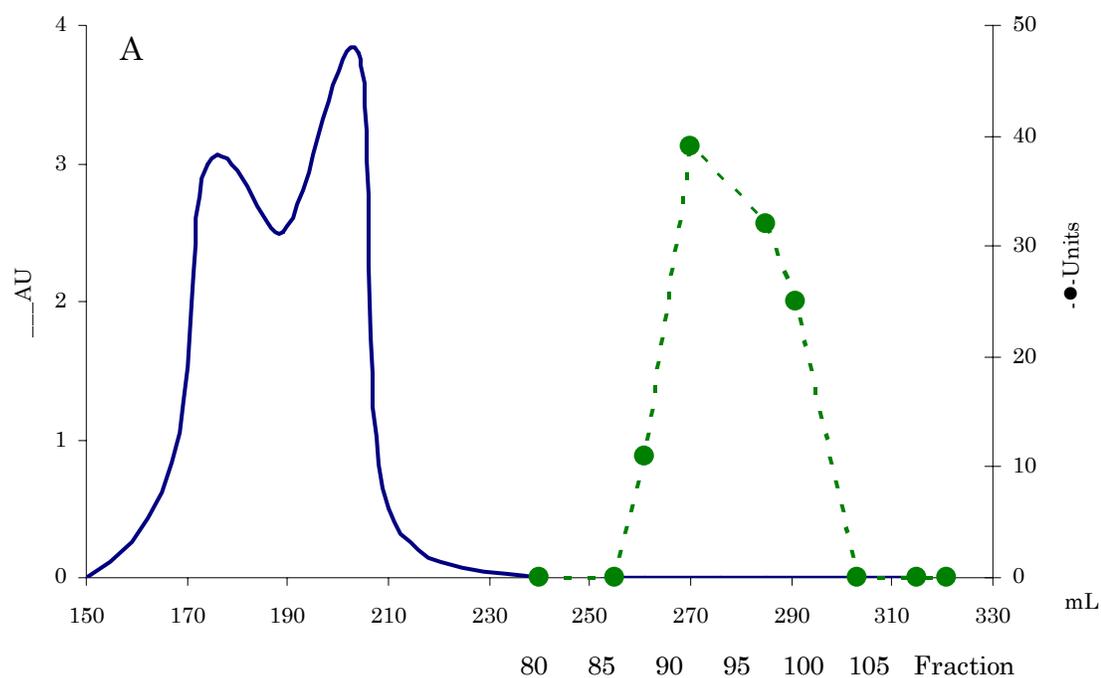


Figure 9.2: A) Elution profiles of activity eluted from Sephacryl S100; solid line represents the UV absorption at 254 nm, broken line represents the units of activity of  $\gamma$ GACT. B) Electrophoresis of the fractions from Sephacryl S100; molecular masses of the bands are shown.

After the gross purification presented in this work 50 U/mg was obtained; this result doubles the values reported by Gowda and Bowser, who obtained 20 U/mg and 25 U/mg respectively. In addition, size exclusion chromatography on Sephacryl S100 removed  $\gamma$ GAACT; this result was not observed with Sephadex G50; instead, the contaminant enzyme was removed by Bowser by Mono-Q ion exchange chromatography.

*Results of the Polishing Purification Studies on Superdex S75*

Figure 9.3 shows the elution profile of the sample from Sephacryl S100 run on a Superdex HR 75 size exclusion chromatography. Three peaks are differentiated in each graph; it is notable that heavy peaks elute faster in the presence of reducing agent. Similarly, the enzyme activity eluted approximately at 11.8 mL under non-reducing conditions and at 11.5 under reducing conditions. The reading of these elution volumes in equation (XI) obtained for the calibration curve for Superdex S75, see Appendix A for detailed calibration of this column, gives 19.2 and 22.3 KD as the range of molecular masses for  $\gamma$ GACT.

$$MW_{S75} = 10^{(-0.2154 \text{ mL} + 3.8251)} \quad (\text{XI})$$

If the protein were a dimer of molecular mass 22 KD, the activity peak that is observed at 11.5 mL under non reducing conditions, would be observed to elute at 13 mL under reducing conditions, indicating a molecular mass of 11 KD. Because no shift in activity was observed by changing the reducing

conditions of the run, it can be concluded that the protein is not a dimer in contrast to Gowdas' observations. It must be remarked that no dimer associated with the enzyme activity was observed by SDS-PAGE at any step of the purification.

In each case 38 U and 0.6 mg (60 Units/mg) were loaded into the column. Under non reducing conditions the activity peak was broader than under reducing condition; however, the total recovery is approximately the same, recovery without reducing agent was 32.7U (86 %) and with reducing agent was 36.7 U (97 %).

Electrophoresis of the fractions from this column, figure 9.4, shows that the addition of reducing agent has a sharpening effect in the peak; however, the Superdex HR 75 column is unable to resolve proteins in the molecular mass range of 26 to 14 KD, fractions #12 and #11.

The molecular mass of the protein observed in Sephacryl S100 was of 25 KD and of 22 KD in Superdex HR 75. The sample does not look to be a dimer because no peak of half mass was observed under reduction conditions in size exclusion chromatography.

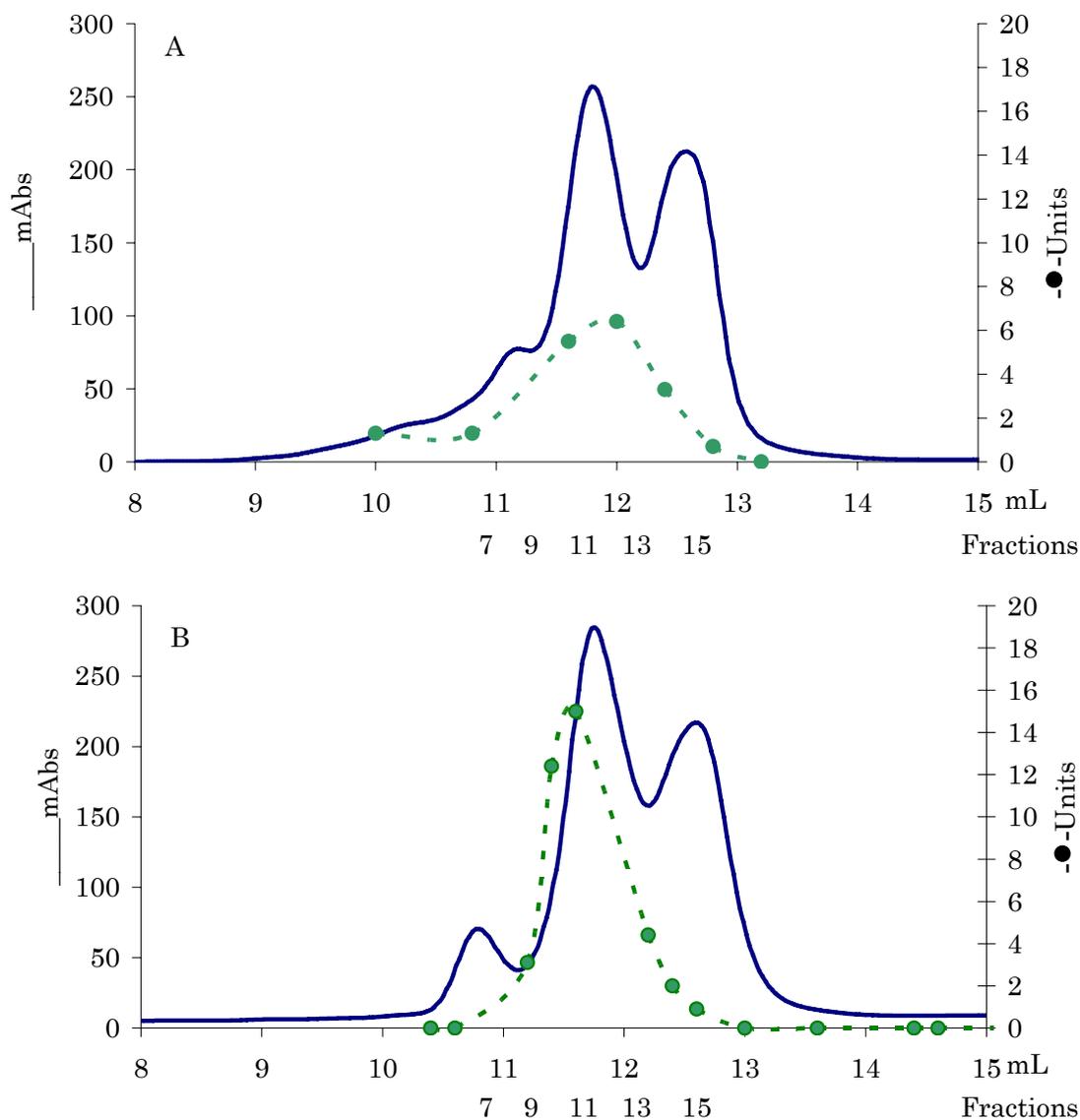


Figure 9.3: Effect of reducing agent on the elution profile of sample in Superdex S75. Solid line represents the UV absorption at 254 nm, broken line represents the units of activity of  $\gamma$ GACT. A) non reducing conditions, B) reducing conditions.



preferentially retained by the column since it eluted with the contaminants. When elution was attempted by increasing the phosphate concentration, and not by introducing NaCl, not even the standards bound to the column; this result shows that phosphate is not an appropriate buffer to run the Mono Q GL column.

Base on this result the enzyme must be resolved on Mono Q GL using Tris buffer, so the stability of the enzyme was tested in this buffer.

#### *Stability of $\gamma$ GACT in Tris Buffer*

The sample from size exclusion in Sephacryl S100 was concentrated to approximately 40 units /mL, and the stability of this concentration of sample was tested at different times in 10 mM Tris buffer pH 7.5. The activities obtained were normalized (Results in Tris/ Results in KPi) using the controls in 5 mM KPi pH 7.5.

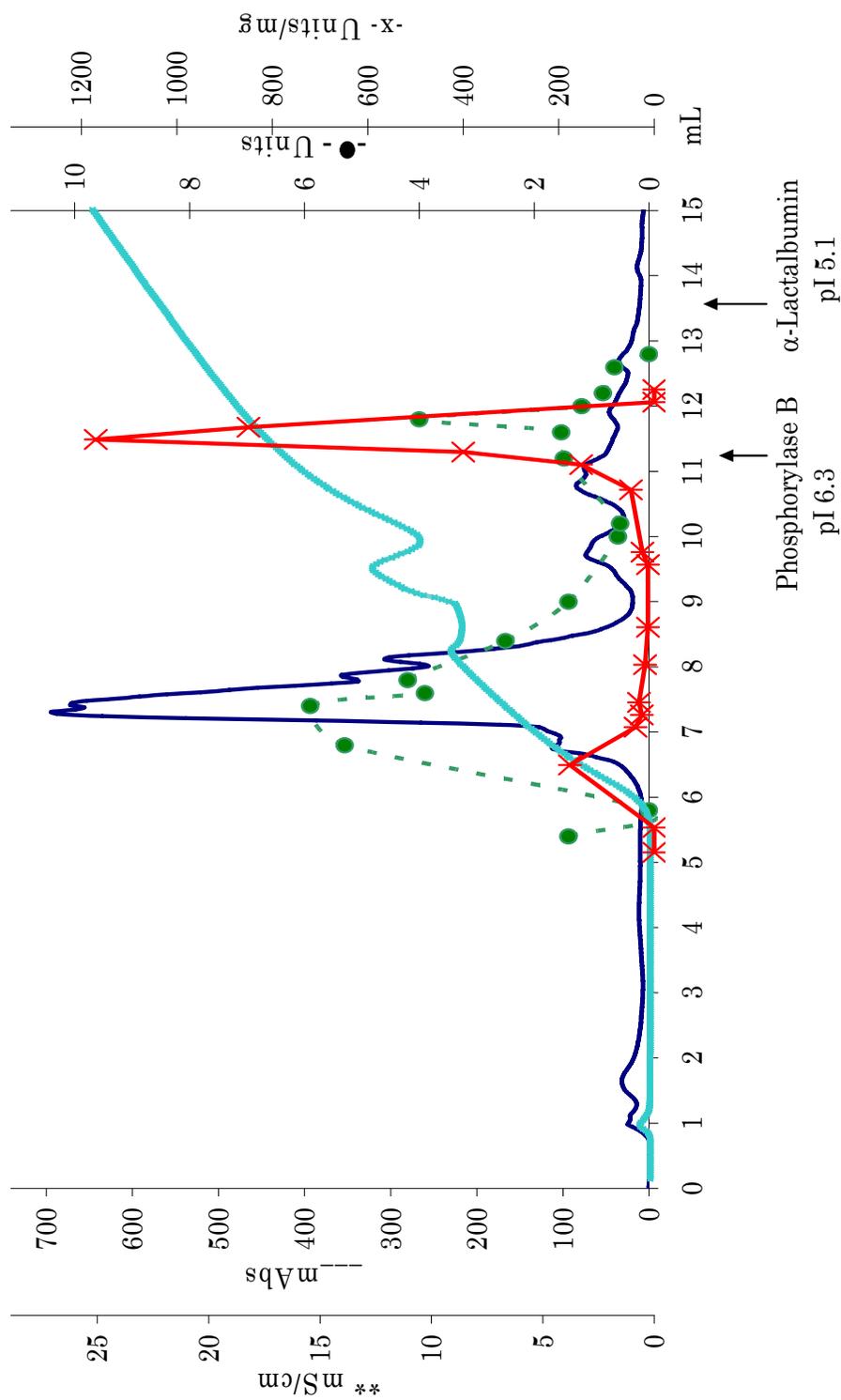


Figure 9.5: Elution profile of sample run in Mono Q GL with 2 mM Phosphate buffer pH 7.5. \*\* conductivity, — absorption of the sample at 254 nm, -o- units of activity (μmol Lys/h), -+- specific activity

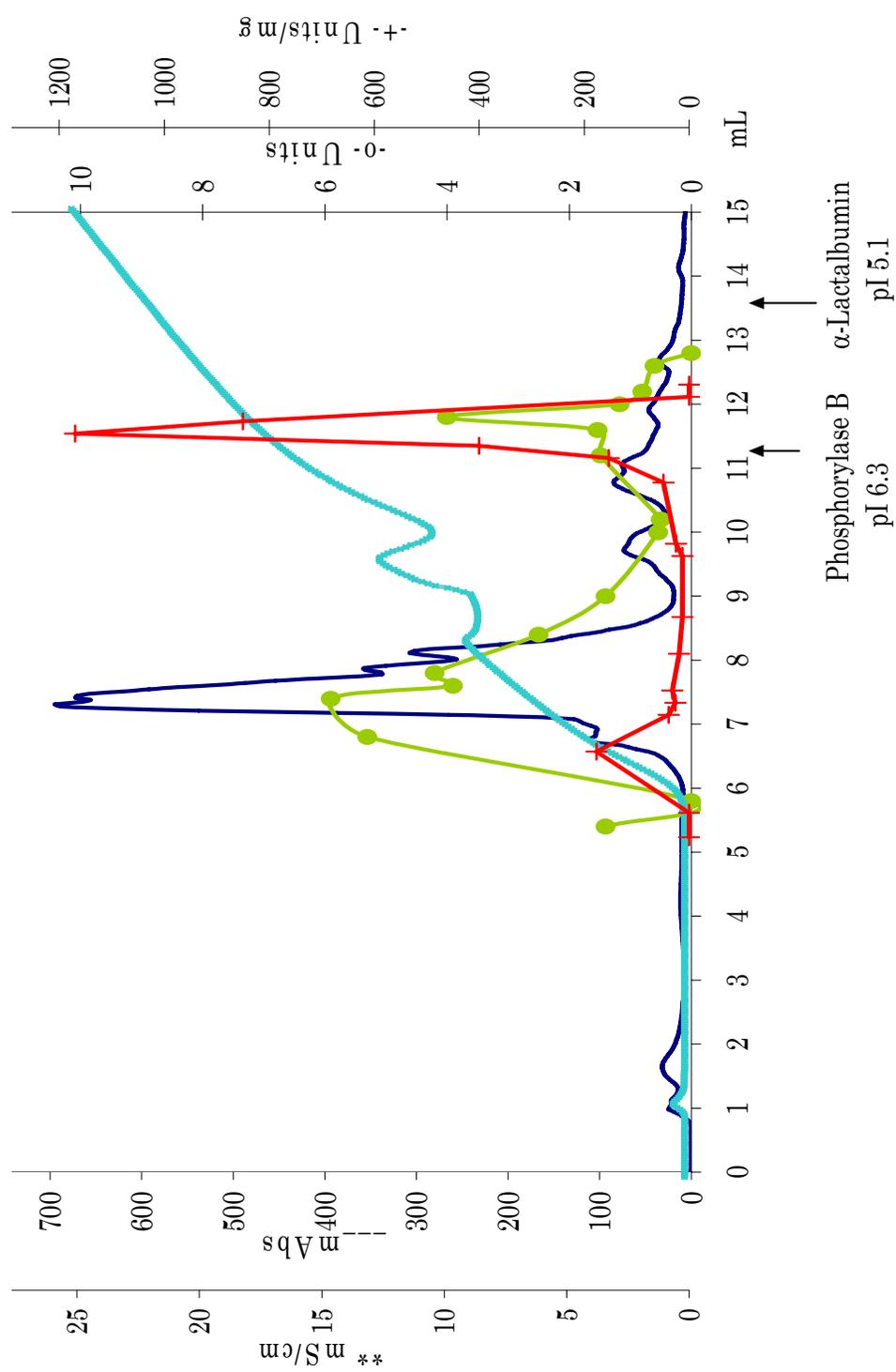


Figure 9.6: Elution profile of sample run in Mono Q GL with 0.5 mM Phosphate buffer pH 7.5. \*\* conductivity, — sample absorption at 254 nm, -o- units of activity ( $\mu$ mol Lys/h), -+ - specific activity

The graphs in figure 9.7 show the effect of incubating the sample in Tris buffer during different times. It was observed that enzyme activity (graph A), total concentration of proteins (graph B), and specific activity (graph C) remained unaltered in 10 mM Tris buffer pH 7.5 for samples as concentrated as 39 U/mL. Only a slightly deviation was observed at low concentration of enzyme, but it must be remembered that at the lower the concentration the higher the error of the assays.

#### *Effect of Alkylation on the Enzyme Activity*

Reduction with 1% of 2-mercaptoethanol showed approximately 20% decrease in activity; however, the specific activity slightly increased.

Table 9.1 summarizes the effect of reduction by DTT and alkylation with IAA of the sample from Sephacryl S100; the values from previous work are compared. It can be seen that in general reduction affects more significantly the lost of activity than alkylation.

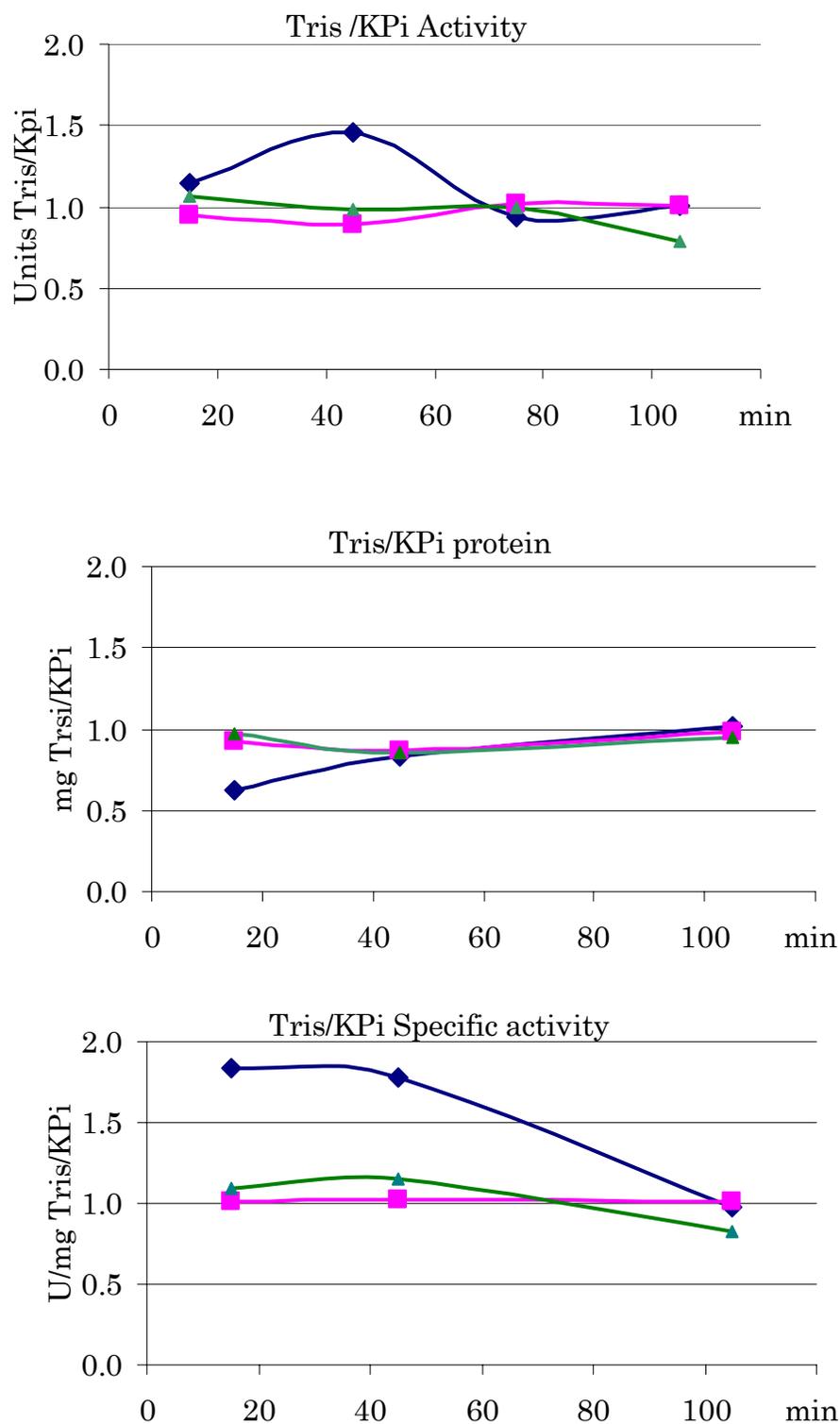


Figure 9.7: Effect of Tris buffer on the activity, protein measurement and specific activity, in comparison to KPi.  $\blacklozenge$  4.3 U/mL,  $\blacksquare$  22 U/mL,  $\blacktriangle$  39U/mL

Table 9.1: Effect of reduction and alkylation on the enzyme activity

Experimental conditions (reference)	Relative activity (%)	
	DTT	IAA
1 mM DTT, 15 mM IAA (bovine enzyme in this work)	35	54
10 mM of each reagent (bovine enzyme by Gowda, 1985)	81	44
1 mM of each reagent (bovine enzyme by Gowda, 1985)	85	82
2.5 mM of each reagent (Rabbit enzyme by Fink and Folk, 1981)	73	96

#### *Mono Q GL Separation in Tris Buffer*

The recommended conditions for the separation in this column, linear gradient from 0 to 100 % 1M NaCl in 20 mM Tris buffer pH 8 in 20 CV at 1 mL/min gave no resolution of the activity from the contaminants; loaded specific activity was the same as the collected one. The elution profile of this separation is shown in figure 9.8.

After several screening experiments, the initial conditions chosen for the purification were the following: 0.5 mL of sample injected, flow rate of 1.5 mL/min, initial buffer 10 mM Tris pH 8, elution buffer 10 mM Tris pH 8 with 0.3 M NaCl, gradient program: wash out of unbound proteins with 5 CV of 10 % elution buffer followed by a linear gradient to 30 % elution buffer in 30 CV (condition #1 are low Tris concentration, slow gradient, without reducing agent). Gradient, concentration of Tris, and presence of reducing agent were

optimized in specific activity increase and recovery of activity by changing the buffer concentration, gradient length, and adding reducing agent.

Under condition #1 (figure 9.9), the activity eluted at 6.6 mS/cm which corresponds to pI 6.89; see in Appendix A the calibration of Mono Q GL column. The increase of buffer concentration from 10 mM to 20 mM Tris at pH 8 showed narrowing the activity peak width. As can be seen in figure 9.10, with 20 mM buffer (condition #2 corresponds to high Tris concentration, slow gradient, without reducing agent) activity elutes between 14 and 17 mL at ~ 6.2 mS/cm; specific activity rose from 32 to 250 (8 fold); while, with 10 mM Tris (condition #1), activity elutes between 18 and 23 mL at ~ 6.6 mS/cm; specific activity rose from 39 to 69 (2 fold). On the other hand, recovery with 10 mM Tris was 37 % and with 20 mM Tris was only 17%.

Under condition #2, the activity eluted at 6.2 mS/cm which corresponds to a pI 6.96. A steeper gradient increases the recovery of activity although the purification fold decreases. A shallower gradient (figure 9.11), from 10 to 30% elution buffer in 30 CV (condition #2), showed a specific activity increased from 32 to 250 (8 fold) and a recovery of 17%; in contrast, a steeper gradient, from 10 to 30% elution buffer in 10 CV (condition #3 corresponds to high Tris concentration, steeper gradient, without reducing agent), showed activity increased from 71 to 230 (3 fold), and recovery of 30 %; figure 9.12 shows this result. Under condition #3 activity eluted at 6.86 mS/cm which

corresponds to a pI 6.86 and the presence of a second peak of activity of lower pI begun to be evident at 8 mS/cm which corresponds to a pI of 6.65.

By addition of reducing agent the activity eluted at lower pI, and the second peak increased in presence of 2-ME; the peak eluted at 8.2 mS/cm, which corresponds to pI 6.62. When the sample was treated with 2-ME and run in buffer containing 2-ME (condition #4 are high Tris concentration, stepper gradient, with reducing agent), specific activity and purification fold increased; specific activity increased from 22 to 173 (8 fold) and recovery was 47%; figure 9.13 shows this result.

The oxidation of cysteine forms disulfide bridges; pI heterogeneity is observed for a protein when it is partially oxidized. This is observed for  $\gamma$ GACT activity in the presence and absence of reducing agent when run on the Mono Q GL column. Two main forms of the protein were observed one with pI 6.6 presumed to be the reduced forms (reduced form needs more acid to neutralize thiol groups) and one with pI 6.9 presumed to be the oxidized form (oxidized form needs less acid for neutralization because disulfide bridges are neutral), both with molecular mass of  $\sim 22$  KDa (figure 9.14). Although, both forms show activity, the intermediate species are less active. It is also observed that the enzyme becomes less stable while it becomes more pure; oxidized cysteine groups are known to contribute in the stability of the protein.

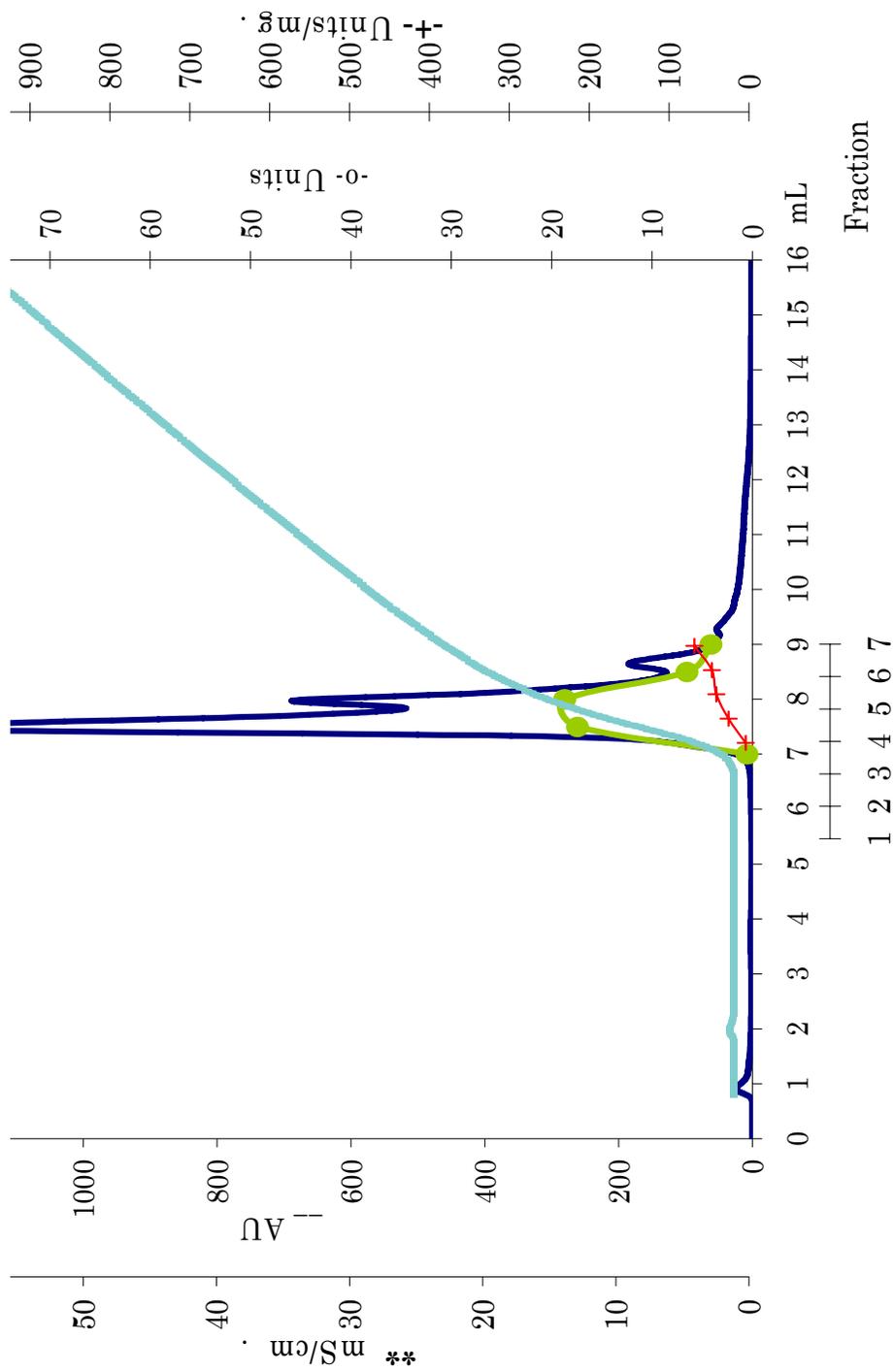


Figure 9.8: Elution profile of  $\gamma$ GACT sample on Mono run at recommended conditions; \*\* conductivity, -- absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h), -+- specific activity

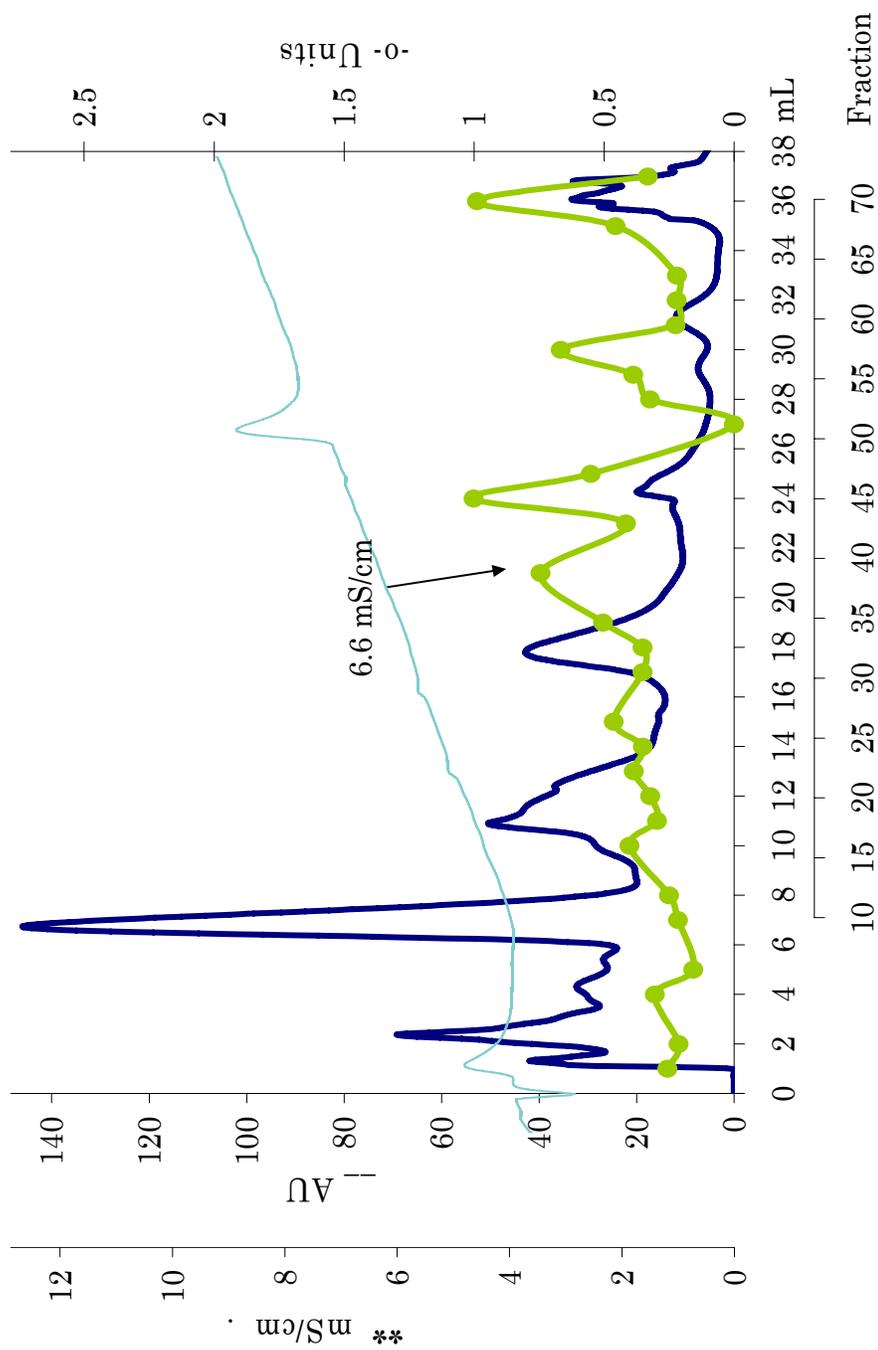


Figure 9.9: Elution profile of  $\gamma$ GACT sample on Mono Q GL, low concentration of Tris, shallow gradient; without reducing agent (Condition #1). \*\* Conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ )

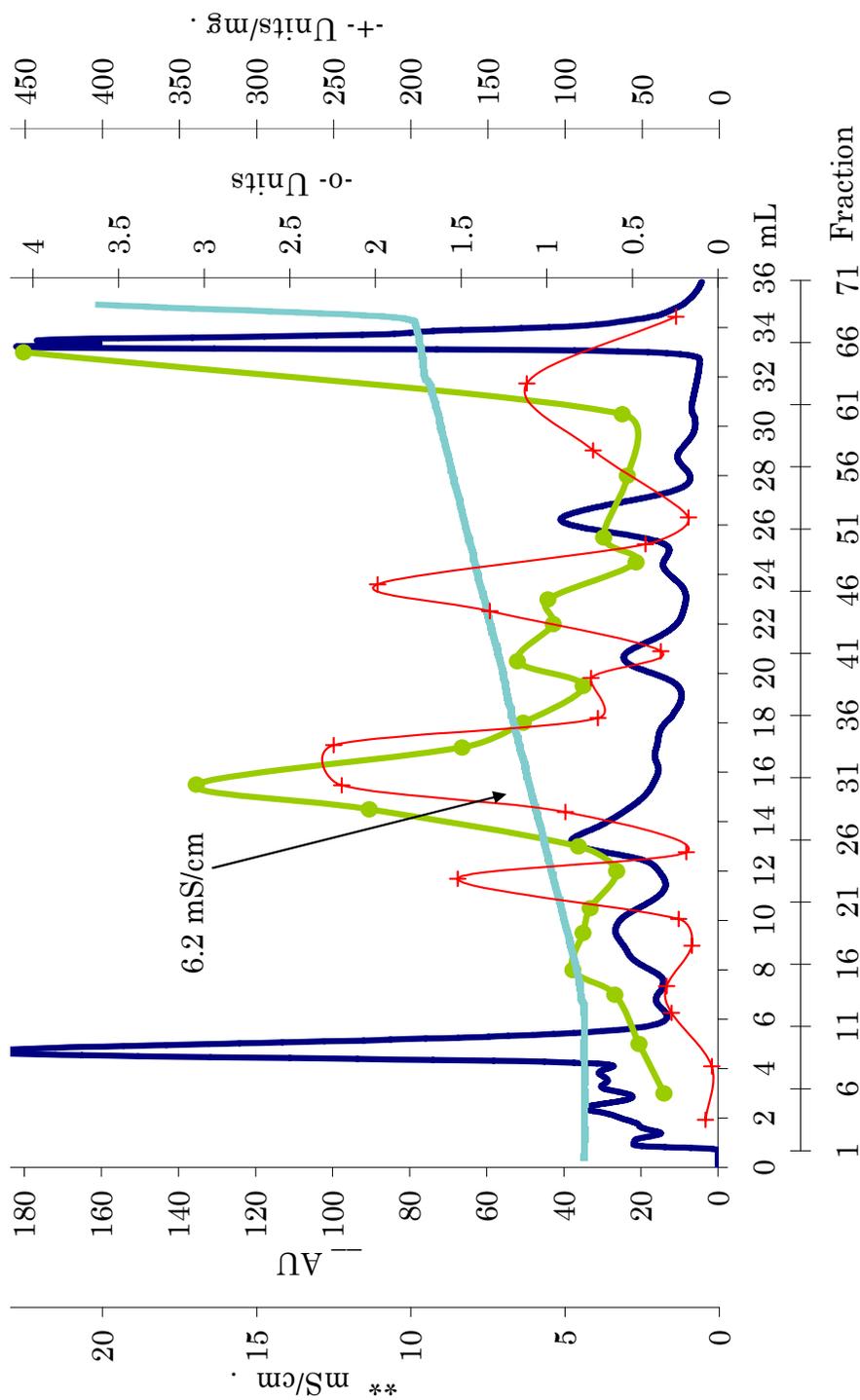


Figure 9.10: Elution profile of  $\gamma$ GACT sample on Mono Q GL; at high concentration of Tris, shallow gradient; without reducing agent (condition #2). \*\*Conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), -+- Specific activity

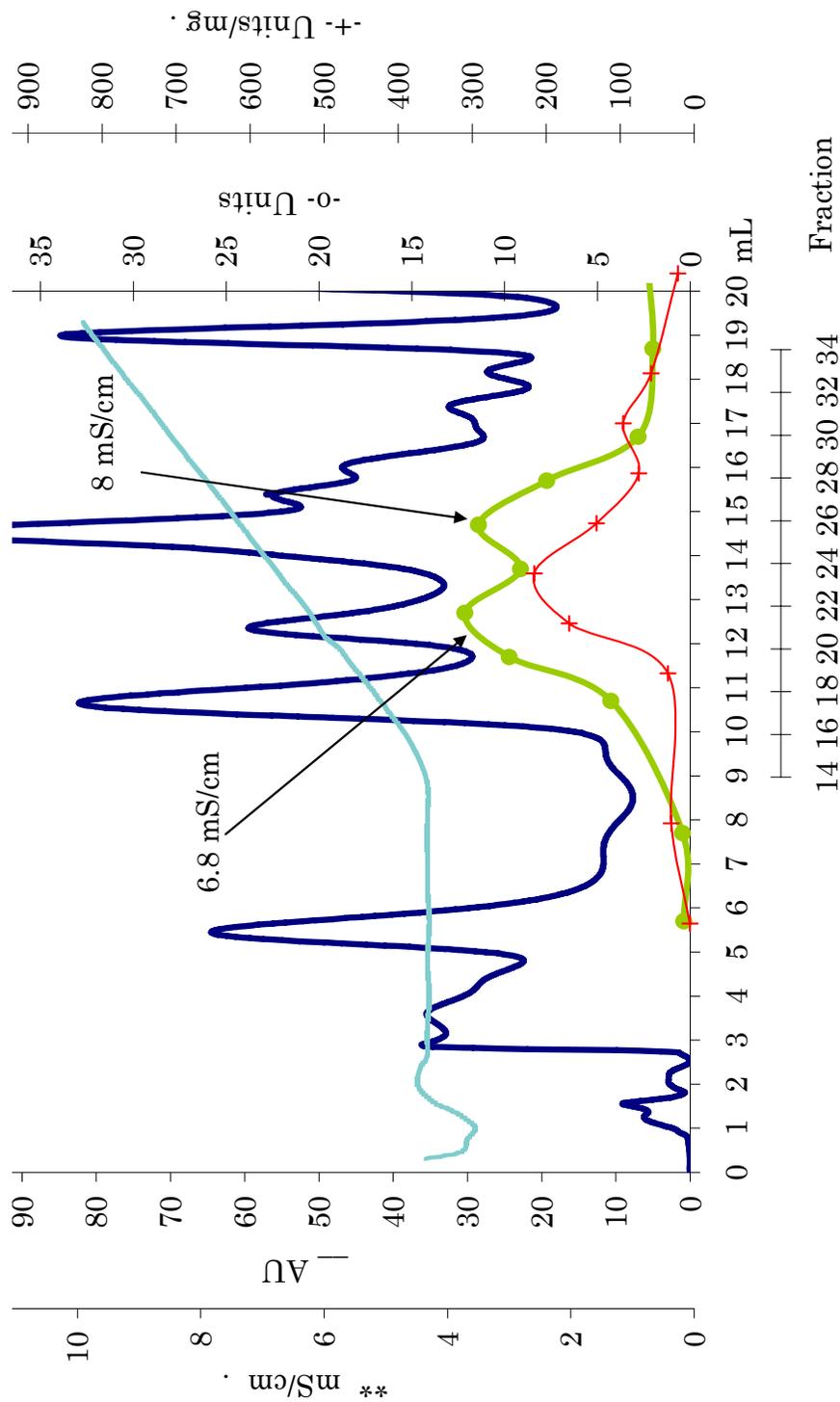


Figure 9.11: Elution profile of  $\gamma$ GACT sample on Mono Q GL at high concentration of Tris, steeper gradient; without reducing agent (Condition #3). \*\* Conductivity, — Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), -+- Specific activity

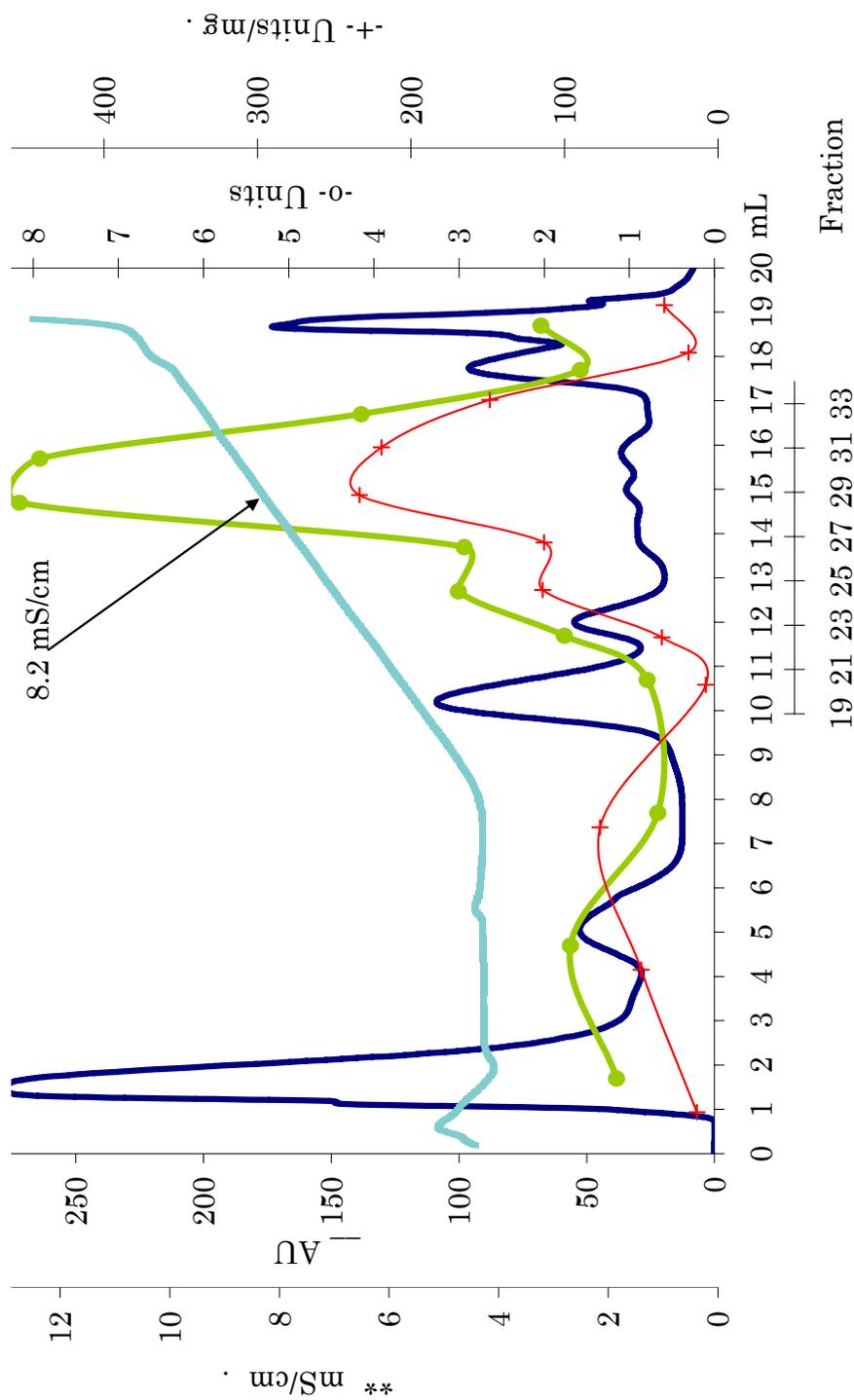


Figure 9.12: Elution profile of  $\gamma$ GACT sample on Mono Q GL at high concentration of Tris, steeper gradient; with reducing agent (condition #4). \*\* Conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), -+- Specific activity

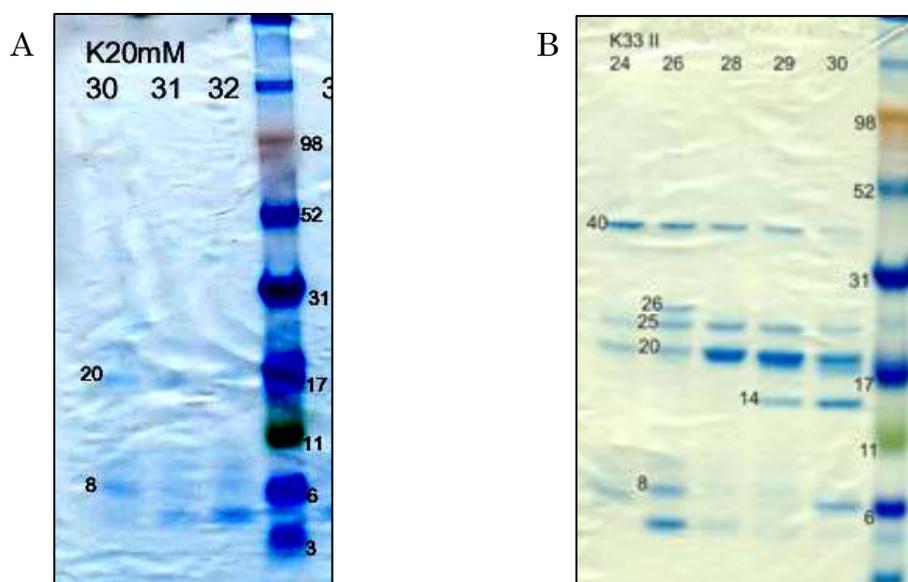


Figure 9.13: SDS-PAGE of Mono Q GL fractions. A) peak of pI 6.9 at condition #2 and B) peak of pI 6.6 condition #4

Figure 9.14 is a representation (recovery\*purification fold) of the improvement in specific activity and recovery gained by changing the buffer concentration, gradient speed, and reduction of the sample.

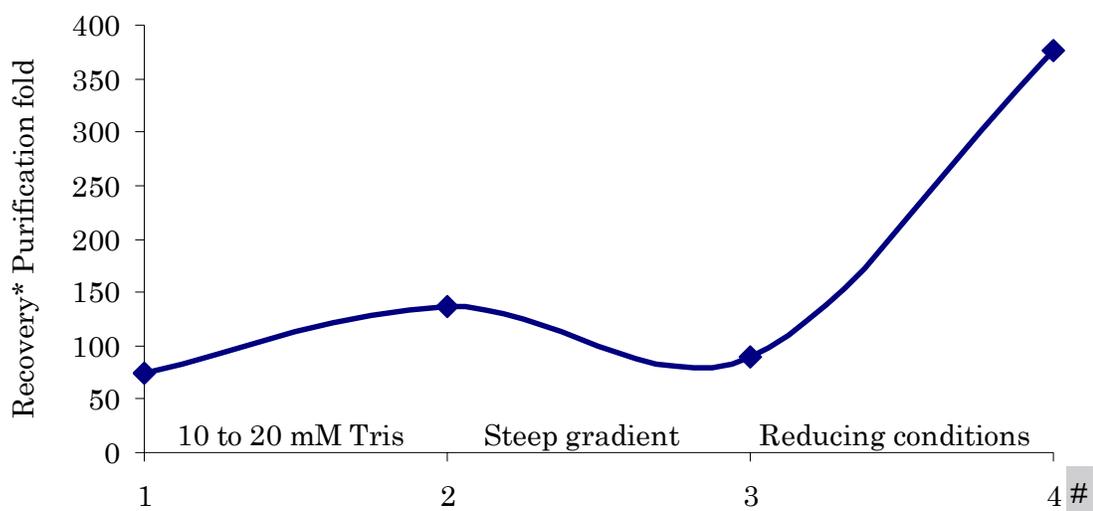


Figure 9.14: Optimization of  $\gamma$ GACT purification on Mono Q GL column

Ion exchange chromatography on Mono-Q done by Bowser showed 7 % recovery and only three fold in the increase of specific activity; while in this work, by optimizing the conditions of the separation a 47 % recovery and up to eight fold in purification by Mono-Q were obtained.

The addition of chelating agents as well as the alkylation of cysteine groups did not improve the purification. The addition of EGTA to the sample prevented the binding of the enzyme to the column, as can be seen in figure 9.15. Activity eluted at 15 mL and contaminants at 10 mL, when the sample was alkylated with iodoacetamide contaminants and activity eluted at 12 mL. The effect of alkylation of the sample brought together the enzyme and the contaminants, as can be seen in figure 9.16.

Based upon these observations, it is believed that the enzyme has cysteine groups responsible for the stability of the enzyme; however, it is likely that the thiol groups are not involved in the enzyme activity.

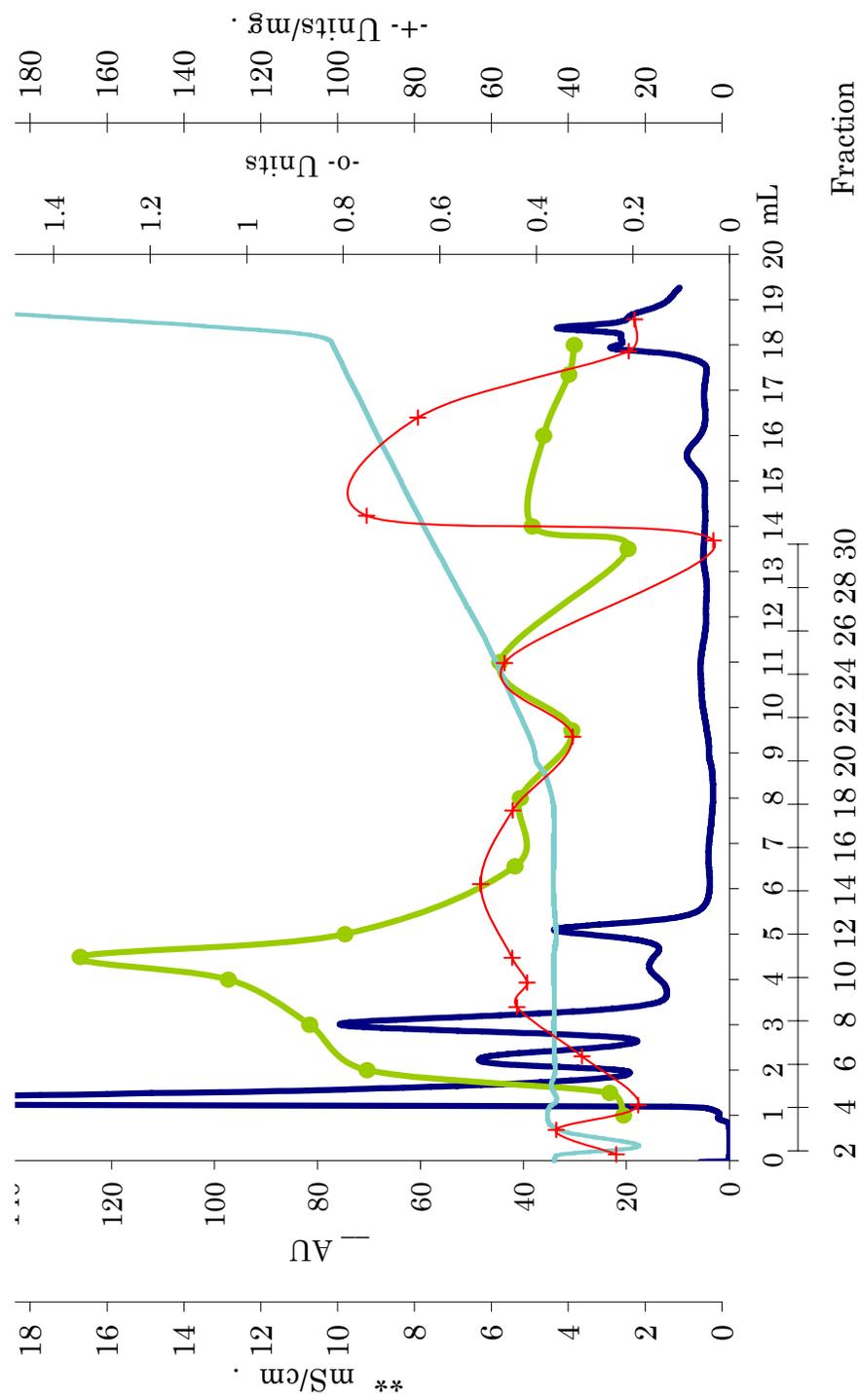


Figure 9.15: Elution profile of  $\gamma$ GACT sample on Mono Q GL with EGTA; \*\* conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), +- Specific activity

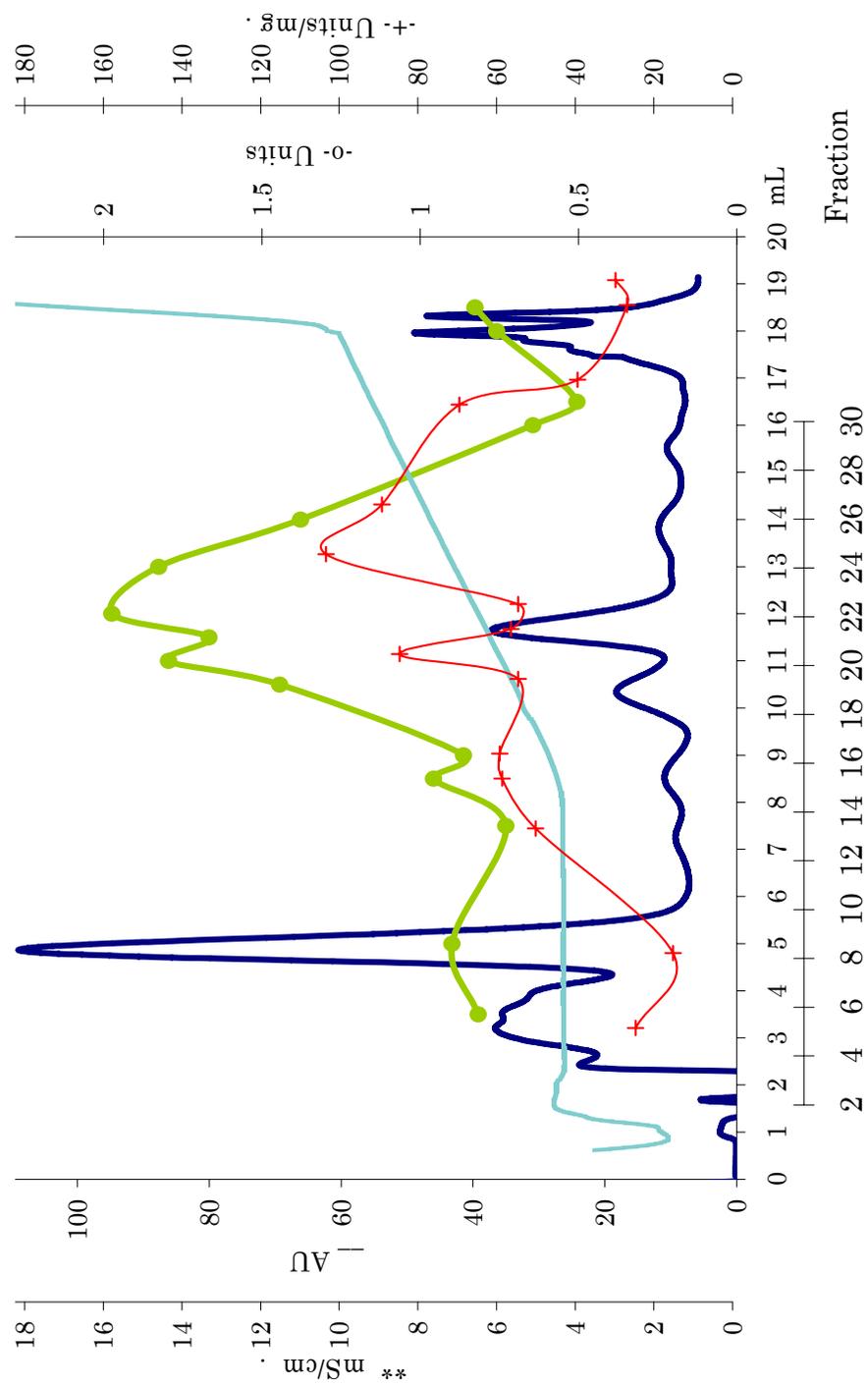


Figure 9.16: Elution profile of  $\gamma$ GACT alkylated sample on Mono Q GL; \*\* conductivity, — Absorption of the sample at 254 nm, -o- units of activity ( $\mu\text{mol Lys/h}$ ), -+- Specific activity

*Results from Isoelectric Focusing*

Isoelectric focusing was carried out for the sample from the Mono Q GL column run under the optimal conditions and for the same sample after alkylation with iodoacetamide. The stained IEF gels from these runs are shown in figure 9.17; the samples recovered from isoelectric focusing were tested for activity and run on SDS-PAGE.

Figure 9.18 shows the SDS-PAGE after IEF. Ampholites are observed as smear stain at the top of lanes 11, 12.5, 13 and 13.5; when ampholites were not removed the observation of the protein bands was impossible. The sample loaded onto the IEF gel showed a band at ~22 KD.

For the non-alkylated sample, activity was observed at pH 6.18, 6.54, and 6.76; while, for the alkylated sample the main activity was observed at pH 6.45 and a trace of activity at pH 6.72. This result is comparable to that of Gowda; she observed a pI shift from 6.15 to 7.10 by alkylating the sample before IEF. By alkylation of cysteine groups a shift towards higher pI is expected because acidic thiols become neutral thioesters.

The molecular mass of the protein corresponds to 22 KD as calculated with equation (XII) obtained from the calibration curve of standard molecular masses and their migration distances in the gel, Figure 9.19.

$$MW_{\text{SDS-PAGE}} = 10^{(-0.192 \text{ cm} + 2.3129)} \quad (\text{XII})$$

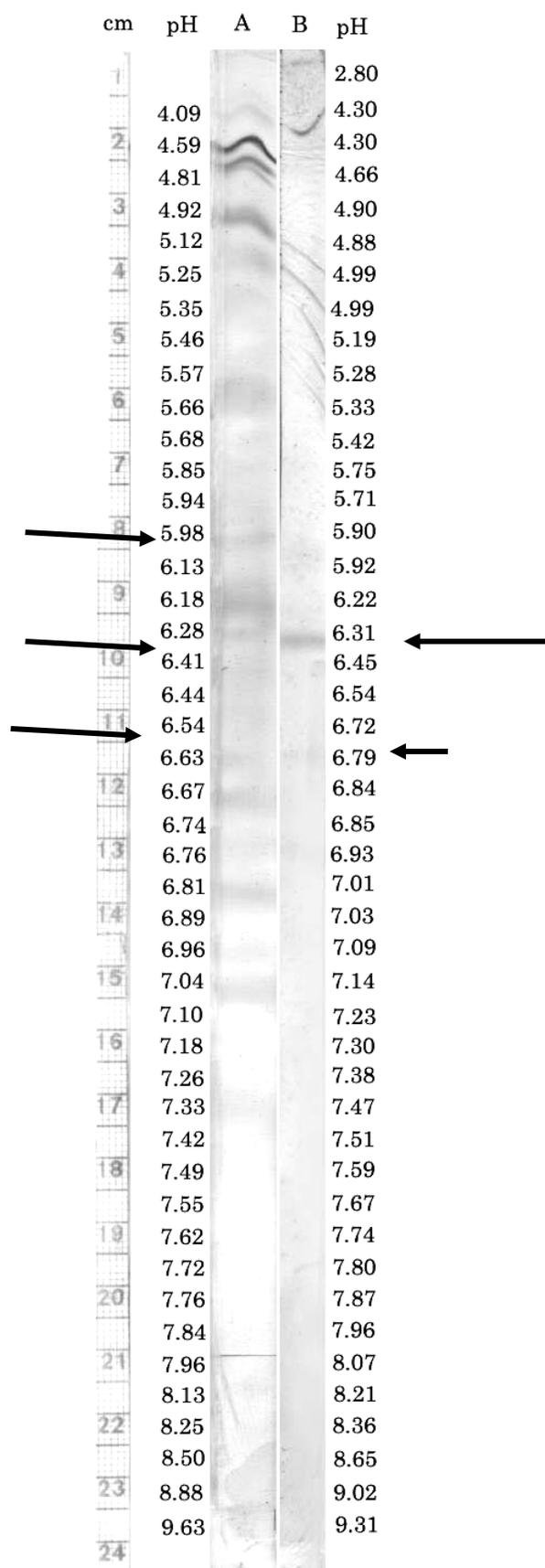


Figure 9.17: Isoelectric focusing of the non alkylated sample (A) and Alkylated sample (B), with measured pH. The migration distances are measured with the scale shown at the left. Arrows indicate the position of the bands.

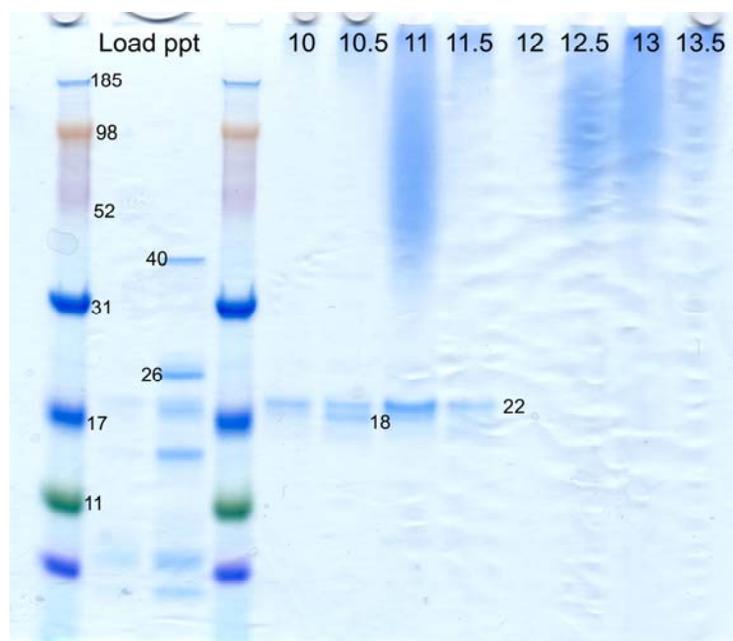


Figure 9.18: SDS-PAGE of the protein from IEF. Molecular masses of the bands are marked. Load is the sample loaded onto the IEF, ppt was the precipitated protein in water before IEF gel loading; the numbers correspond to the migration distances (cm) on the IEF gel

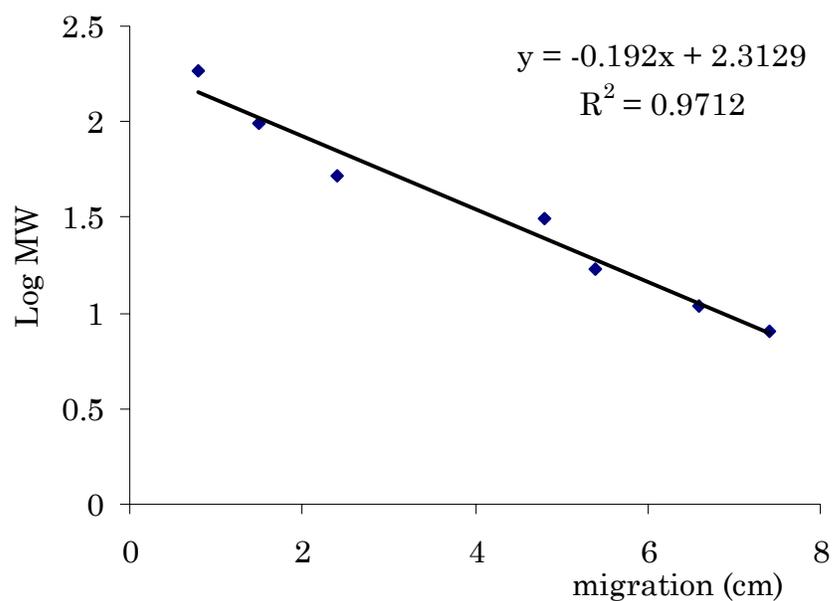


Figure 9.19: LogMW vs migration distance (cm). Calibration curve of SDS-PAGE gel for the measurement of the molecular mass of proteins.

*Summary of  $\gamma$ GACT Purification*

The purification of  $\gamma$ GACT was continued until an single electrophoretic band was observed; figure 9.20 shows the advance of the purification observed by SDS-PAGE; and figure 9.21 and table 9.2 summarize  $\gamma$ GACT purification in this work.

The behavior of the protein under reduction conditions was useful to observe other properties of the enzyme; great information was gained about the possible isoforms of the enzyme. This information was used to improve resolution and recovery on ion exchange chromatography and IEF

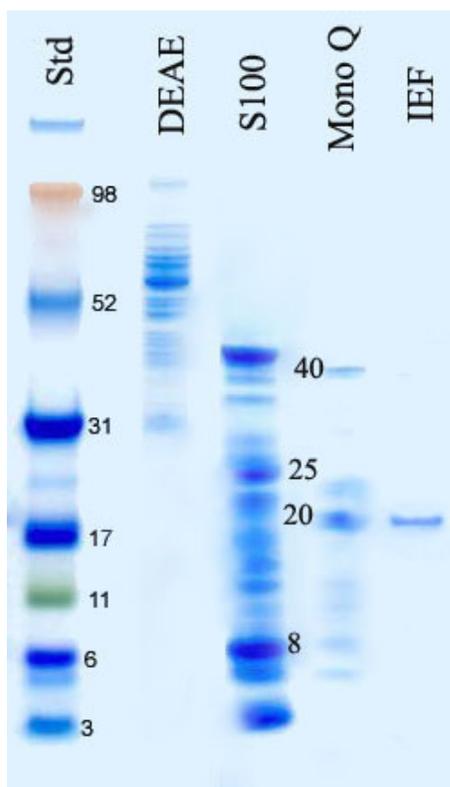


Figure 9.20: Purification followed by SDS-PAGE; molecular mass of the bands are indicated

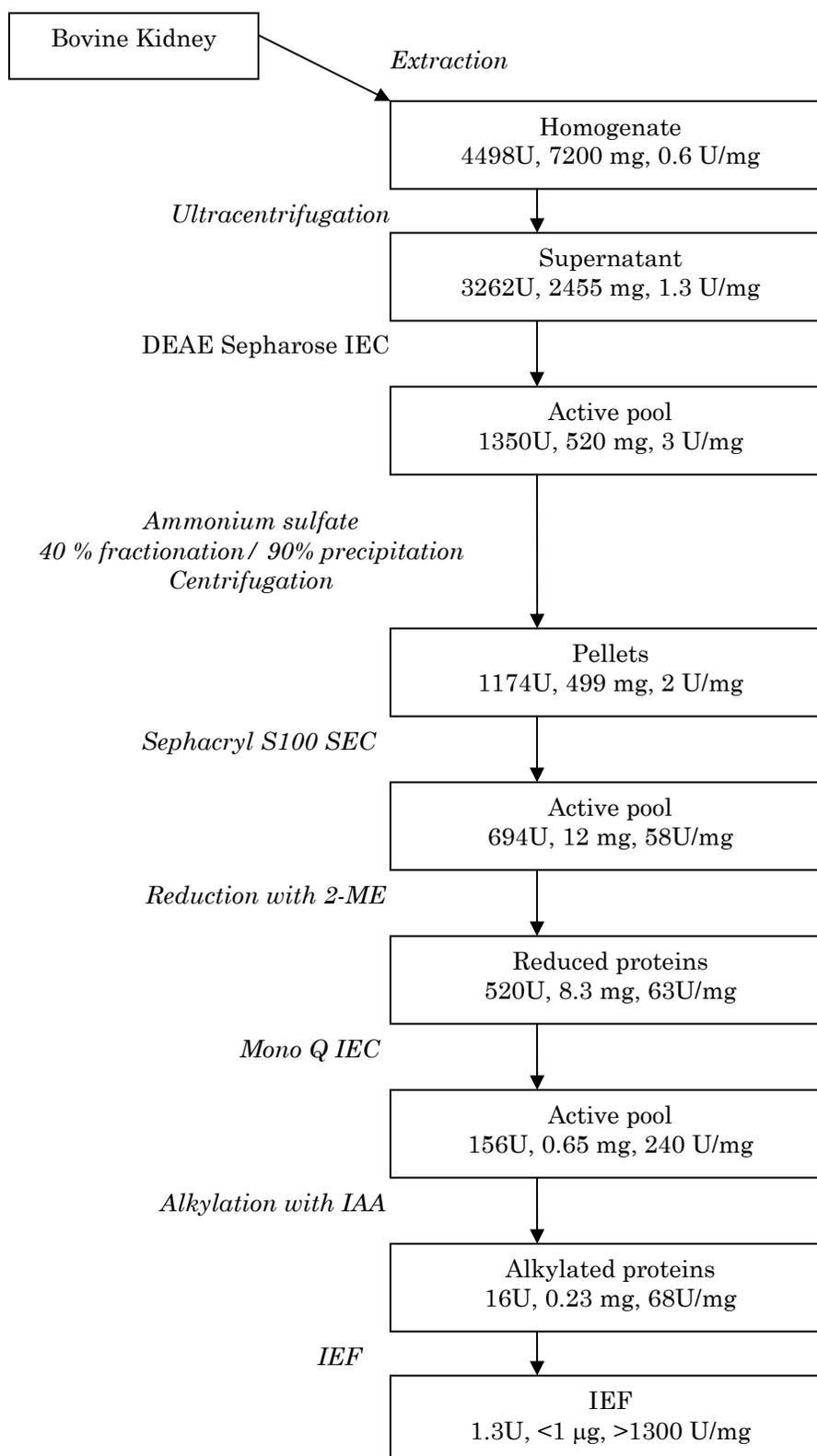
Figure 9.21: Scheme of  $\gamma$ GACT purification

Table 9.2: Purification of  $\gamma$ GACT, from 50 g of bovine kidney tissue

Step	Volume (mL)	Total protein (mg)	Total Units ( $\mu$ mol/h)	Specific Activity (U/mg)	Recovery (%)	Fold
Homogenate	200	7,200	4,498	0.6	100	
Ultracentrifugation Supernatant	150	2,455	3,262	1.3	72	2
DEAE Pool	250	520	1,350	3	30	5
Ammonium sulfate	10	499	1,174	2	26	3
Sephacryl S100	80	12	694	58	15	97
Mono Q GL (of 160U)	2	0.2	48	240	5*	400
Alkylation (of 48U)	3.3	0.072	4.9	68	0.11	113
IEF	0.05	<0.001	1.3	>1300	0.03	2166

\* Recovery from the Mono Q GL chromatography of 160 U is 1.1 %

## CHAPTER TEN

### Results and Discussion of the Sequencing

#### *N-Terminal Determination*

With the conditions described in table 7.1, the resolution and observation of 1 to 10 pmol of dansylamino acids was possible; figure 10.1 shows the separation of dansylamino acid standards by HPLC. Before attempting dansylation and hydrolysis of the sample, a protein control was run to verify that the procedure was working. Bovine carbonic anhydrase, which has a close molecular mass (29,000 D) and isoelectric point (5.8) to that of  $\gamma$ GACT, was dansylated and hydrolyzed; the observation of dansyl- $\epsilon$ -lysine (figure 10.2) at 42.7 min and probably dansyl-arginine at 32.9 minutes proves that the procedure worked.

The dansylation and hydrolysis of the sample extracted from SDS-PAGE showed no peaks (figure 10.3). Since not even the dansyl- $\epsilon$ -lysine peak was observed, it can be said that not enough protein was recovered from the gel in order to measure any dansylation; consequently, that the protein is N-terminal blocked can not be assumed.

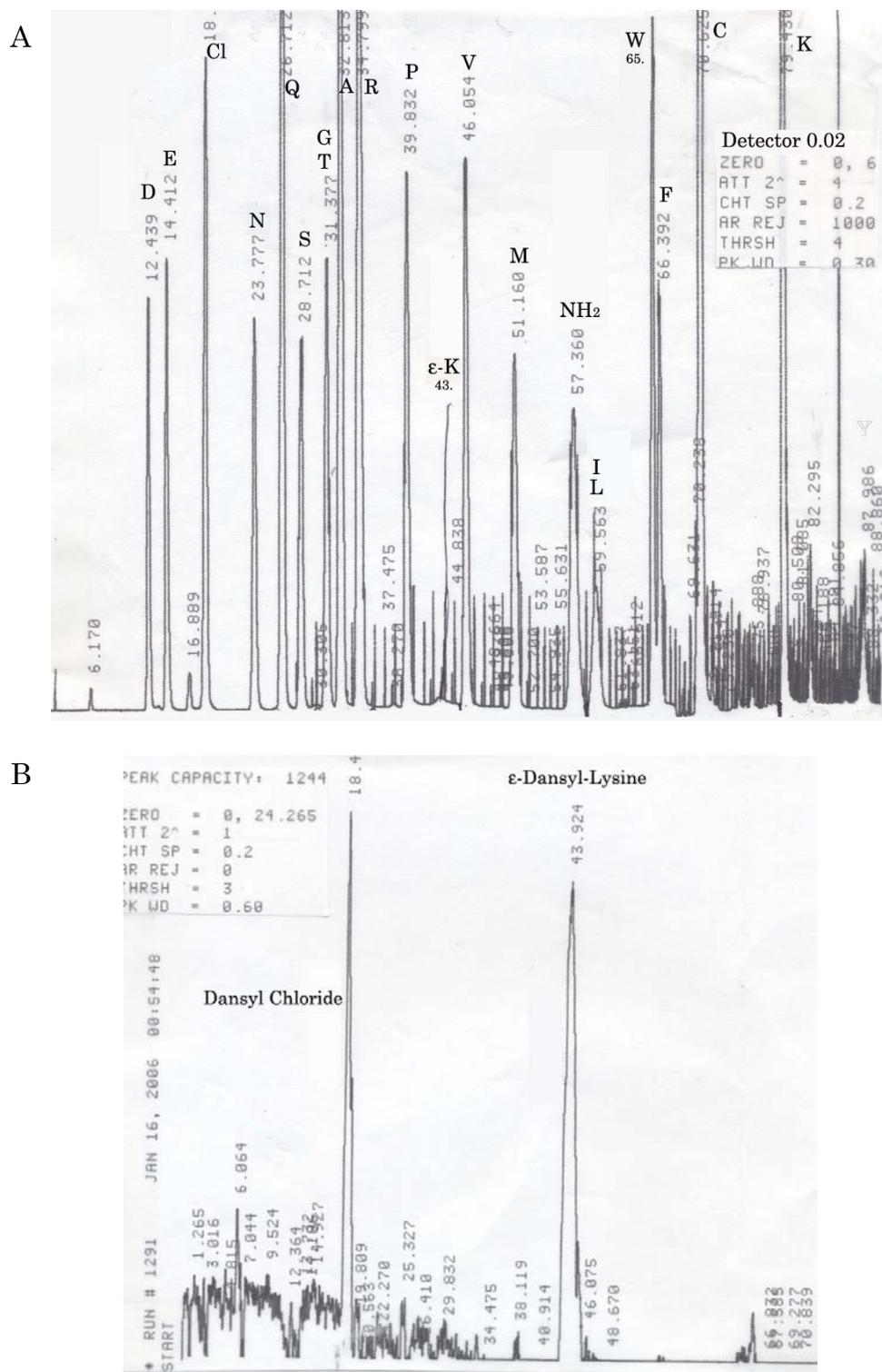


Figure 10.1: Dansyl amino acid standard resolved by HPLC, detected by fluorescence. A) 20 nmol of the indicated standards, B) 0.1 nmol of  $\epsilon$ -dansyl-lysine standard

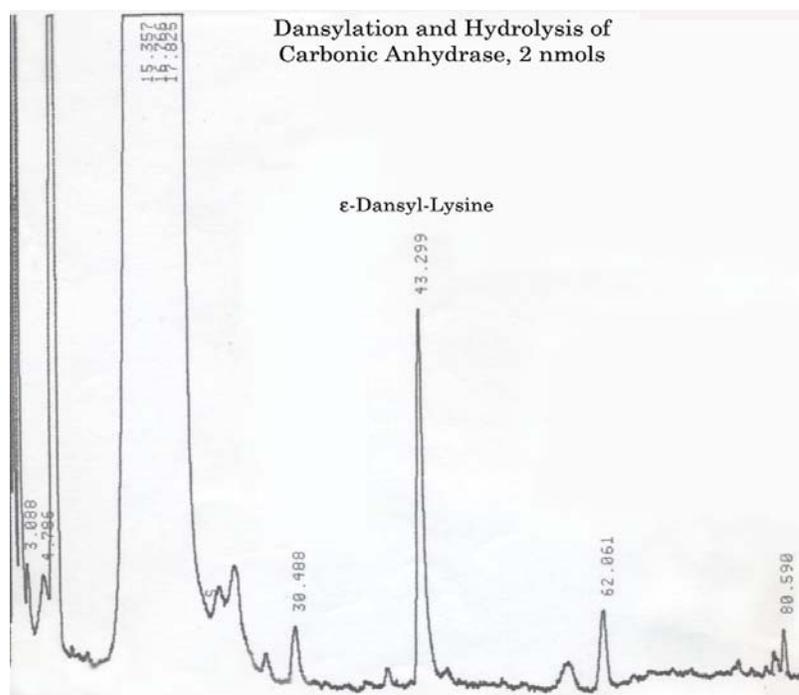


Figure 10.2: Elution profile of dansylated and hydrolyzed carbonic anhydrase standard protein

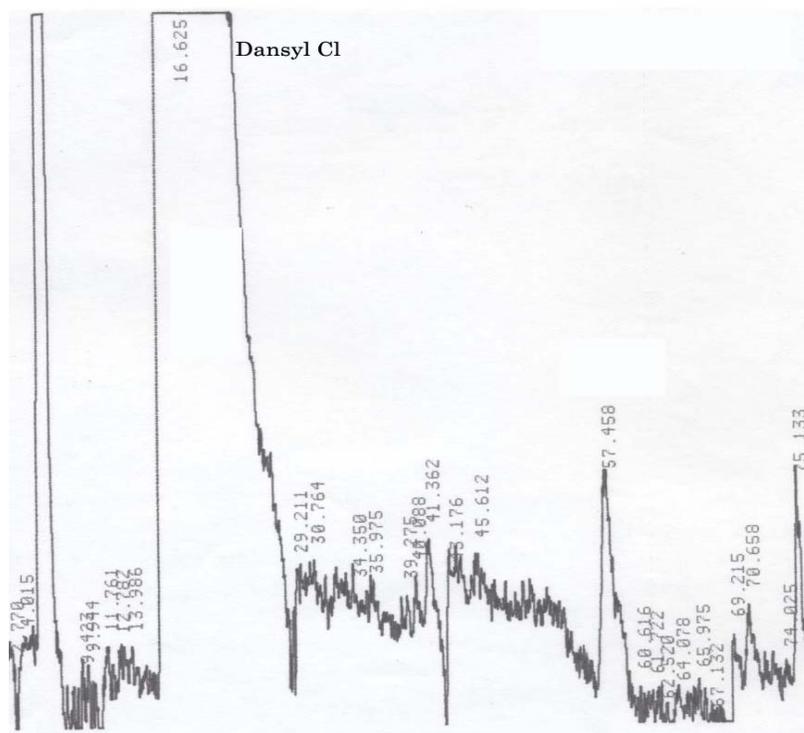


Figure 10.3: Elution profile of dansylated and hydrolyzed gel extracted  $\gamma$ GACT

### *MS Sequencing*

The determination of protein by MS spectrometry has a sensitivity of less than a fmol, so almost any contamination can be observed. A list of 42 proteins was reported (table 10.1 and Appendix C) by Harvard Microchemistry and Proteomic Analysis Facility to be present in the electrophoretically pure protein band of SDS-PAGE.

The reported identification number of the protein was used to search the sequence of the protein in the National Center for Biotechnology Information; with the obtained sequences the predicted molecular masses and isoelectric points were obtained. The obvious contaminations like protein “C” albumin bovine or protein “I” human keratin, were discarded. Other proteins were discarded base on those characteristics that do not correlate with  $\gamma$ GACT behavior; for example, protein “Q” Biliverdin reductase, was discarded because  $\gamma$ GACT does not present reductase activity.

Similarly, special attention was given to those proteins that act on peptide bonds or whose substrate or product resemble  $\gamma$ GACT's ones; for example, protein “P” Similar to peptidylpropyl isomerase-like protein 3 (PPIL3b), which acts on a peptidylproline substrate.

### *Analysis of Candidate Proteins*

Five candidate proteins, from Bos Taurus, were selected as the most probable  $\gamma$ GACT; these are protein “P” similar to peptidylpropyl isomerase-like protein 3 , protein “U” hypothetical protein LOC515270 and similar to

Guanidinoacetate N-methyltransferase , protein “a” hypothetical protein LOC537221, “1” protein unknown (protein for MGC:134378), and “5” Prostaglandin H2 D-isomerase. Tables 10.2 and 10.3 summarize the characteristics of these candidate proteins.

Table 10.1: Reported Proteins

#	Protein	ID number
A	Similar to Troponin C akin-1-protein [Bos taurus]	XP_590139.1
B	RAB11A member RAS oncogene family	AAV38958.1
C	Albumin [Bos taurus]	NP_851335.1
D	Chain B, Human Adp-Ribosylation factor 1	1HURB
E	Similar to cytoglobin [Bos taurus]	XP_58727.2
F	Transgelin 2 [Bos taurus]	NP_001013617.1
G	Hypothetical protein LOC613749 [Bos taurus]	NP_001029915.1
H	Similar to SAR1a gene homolog 1 isoform 2	XP_536379.1
I	Keratin 1 [Homo sapiens]	NP_006112.2
J	Galactose mutarotase	AAI02447.1
K	Similar to retinol binding protein 5	XP_587041.2
L	Rab7 [Mus musculus]	CAA61797.1
M	Similar to core-binding factor	XP_871348.1
N	Similar to Mark3 homolog	XP_595378.1
O	Similar to ADP-ribosylation factor-like 3	AAI09566.1
P	Similar to peptidylpropyl isomerase-like protein 3	XP_87272.1
Q	Biliverdi reductase B	AAI02270.1
R	Chain I, P13 alanine variant of antithrombin	1OYH
S	Chain B, Human Vh1-related dual-specificity phosphatase	1VHR

Table 10.1: Reported Proteins (continuation)

#	Protein	ID number
T	Unknown protein for MGC:127550 [Bos taurus]	AAI05348.1
U	Similar to Guanidinoacetate N-methyltransferase	AAI09826.1
V	Ribohorin II, precursor [Homo sapiens]	AAH03560.1
W	Similar to oxidation resistance 1	XP_584759.2
X	Cytokeratine 9 [Homo sapiens]	CAA82315.1
Y	Carbonyl reductase 1	AAI02944.1
Z	Protein tyrosine kinase 9	NP_00102049.1
a	Ufm1-conjugating enzyme 1	NP_001015663.1
b	Similar to Y55F3AM.10	XP_872147.1
c	Similar to alpha 1 type XVII collagen isoform 1 precursor	XP_592628.2
d	Similar to cytoskeleton-associated protein 4	AAH25341.1
e	Unnamed protein product [Homo sapiens]	CAA32649.1
f	Rho GDP dissociation inhibitor beta	AAI02110.1
g	Angiogenin-2	ANG2_BOVIN
h	Beta-amylase	AAA33898.1
i	Similar to disabled homolog 2 isoform b isoform 1	XP_586135.2
1	Unknown for MGC:133592	NP_001032563.1
2	Galacto mutarotase [Sus scrofa]	GALM_PIG
3	Unknown protein from MGC:134378	gi:83405412
4	Inter-alpha-trypsin inhibitor	AMBP_BOVIN
5	Prostaglandin H2 D-isomerase	PTGDS_BOVIN
6	Epidermal cytokeratin 2 [Homo sapiens]	K22E_HUMAN
7	Ig AH	gi:229537+71
8	Similar to cathepsin S preprotein	AAI02246.1

Table 10.2: Candidate proteins summary

Protein	Total amino acids	MW	pI	Amino acid composition																			
				A	R	N	D	Q	E	G	H	I	L	K	M	F	P	S	T	Y	V	W	C
Similar to peptidyl isomerase-like 3 protein	135	15,193	7.47	5	4	12	7	4	7	14	5	8	9	10	3	8	6	6	9	7	8	1	2
Similar to Guanidinoacetate N-methyltransferase	236	26,610	6.09	27	12	6	9	9	18	13	9	11	20	8	8	10	19	8	17	8	13	7	4
Ufm1 or Hypothetical LOC537221	167	19,431	7.38	11	10	7	10	5	14	10	4	11	16	14	3	6	10	4	8	6	9	6	3
Unknown MGC:134378	203	22,555	7.23	14	11	13	12	12	11	15	3	11	22	12	6	7	7	15	7	7	12	0	6
Protanglandin H2 D-isomerase	191	21,229	6.91	16	9	5	6	6	14	13	4	1	23	10	5	10	13	15	17	6	10	4	4

Table 10.3: Amino acid sequence of the candidate proteins

#	Similar to PPIL3b	Similar to GAMT	Ufm1	Unknown protein for MGC :134378	Prostaglandin H2 D- isomerase
1	matpnrlwma	msapaatpif	madeatrrvv	msgpfelsvq	matpnrlwma
11	llllgvlgl	apgenspaw	seipvlktna	dlndllsdgs	llllgvlgl
21	qtpapaqaal	raapaaydas	gprdrelwvq	gcyslpsqpc	qtpapaqaal
31	qpnfeedkfl	dthlqilgkp	rlkeeyqsli	nevtpriyvg	qpnfeedkfl
41	grwftsglas	vmerwetpym	ryvennknd	naiestmlde	grwftsglas
51	nsswflekkk	halaaaaasr	ndwfrlesnk	geirkccqgr	nsswflekkk
61	vismcksvva	ggrvlevfg	egtrwfgkcw	svaqdipklq	vismcksvva
71	paadgglnlt	maiaatkvqe	yihdllkyef	klgithvlna	paadgglnlt
81	stflrkdqce	apieehwiie	diefdipity	aegrsmhvn	stflrkdqce
91	trlllrpag	cnegvfqlq	pttapeiavp	tnanfykds	trlllrpag
101	ppgcysyts	dwalqqphkv	eldgktakmy	itylgikand	ppgcysyts
111	hwssthevs	vplkglweev	rggkicldh	tqefnlsayf	hwssthevs
121	aetdyetyal	aptlpdshfd	fkplwarnvp	ekaadfidqa	aetdyetyal
131	lytegvrgpg	gilydtypls	kfglahlmal	laqkngrvlv	lytegvrgpg
141	qdfrmatlys	eetwhthqfn	glgpwlavei	hcregysrsp	qdfrmatlys
151	rsqnpraevk	firdhafrll	pdliqkgviq	tlviaylmmr	rsqnpraevk
161	ehfttfaksl	kpggvltycn	hkekesq	qkmdvksals	ehfttfaksl
171	gfteegivfl	ltswgelmkt		ivrqnreigp	gfteegivfl
181	pktdkcmeeh	kysdittmfe		ndgflaqleq	pktdkcmeeh
191	p	etqvpallea		lndrlvkegk	p
201		gfrdnirtq		lkl	
211		vmelvppanc			
221		ryyafprmit			
231		plvtkh			

### *Reactions, Mechanism, and Structures of Candidate Enzymes*

*Similar to PPIL3b:* Peptidylpropyl isomerase-like protein 3 (PPIL3b) was initially known as cyclophilin, which binds the immunosuppressant cyclosporine A; the enzyme has also been related to the facilitation of protein folding.

PPIL3b catalyzes the cis-trans isomerization of peptidylproline polypeptides (reaction 12). The mechanism for this reaction was reported by Hur and Bruice in 2002; they showed that the substrate forms three hydrogen bonds with Gln63, Arg55, and Asn102, while substrate amide bond to become modified forms hydrophobic contacts with His126, Phe113, and Phe60, and electrostatic interaction with Arg55 and Asn102.

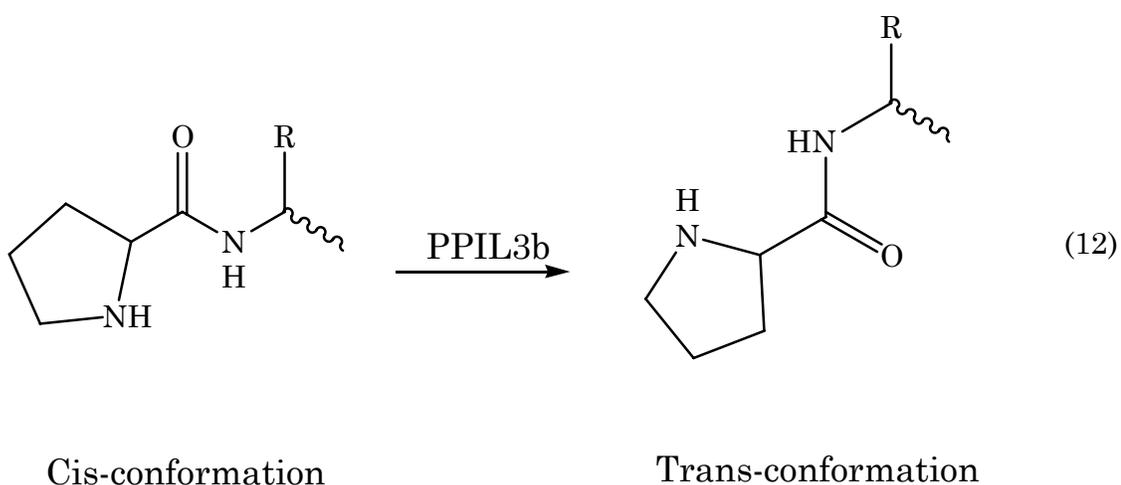


Figure 10.4 shows the crystal structure of PPIL3b and the binding of the polypeptide HAGPILA to the strand and surface representation of the enzyme, as well as the structure for protein similar to PPIL3b believed to be  $\gamma$ GACT. The structure of the protein is of the class alpha-beta protein with a folding called cyclophilin-like; the structure shows a closed barrel of 8 beta strands.

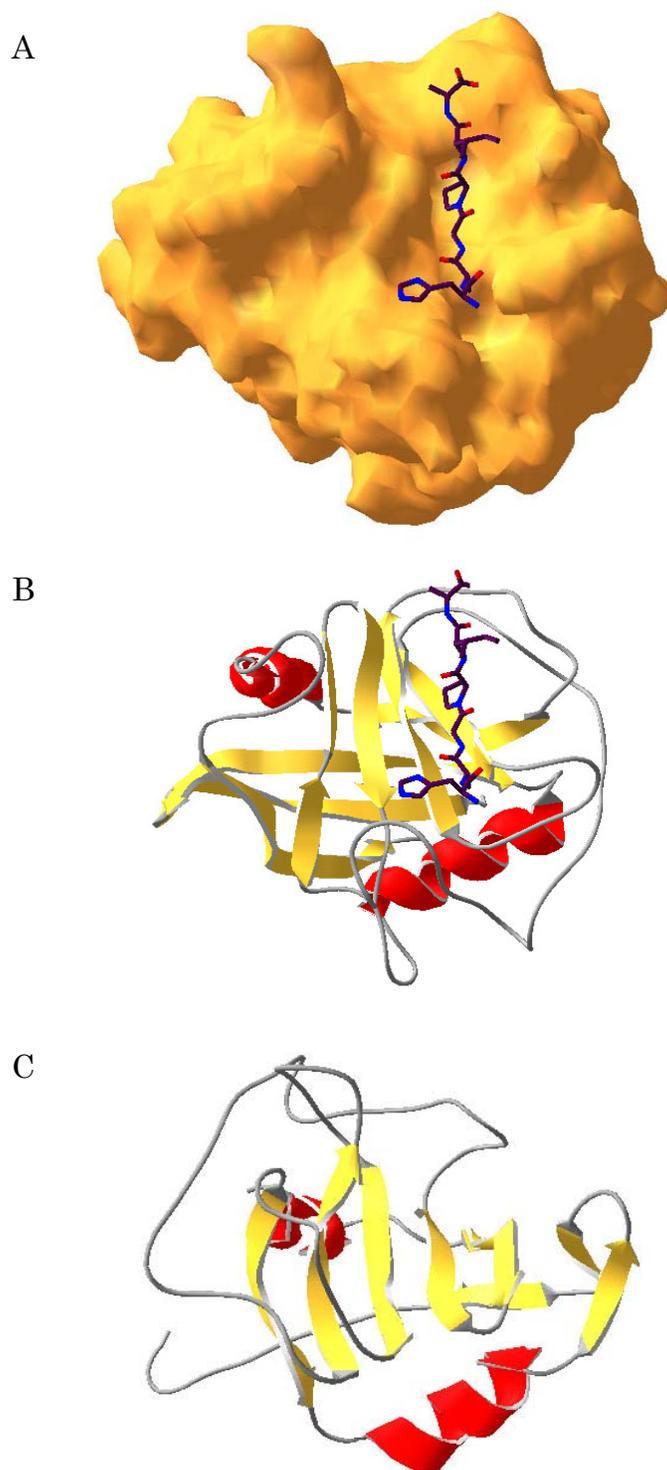
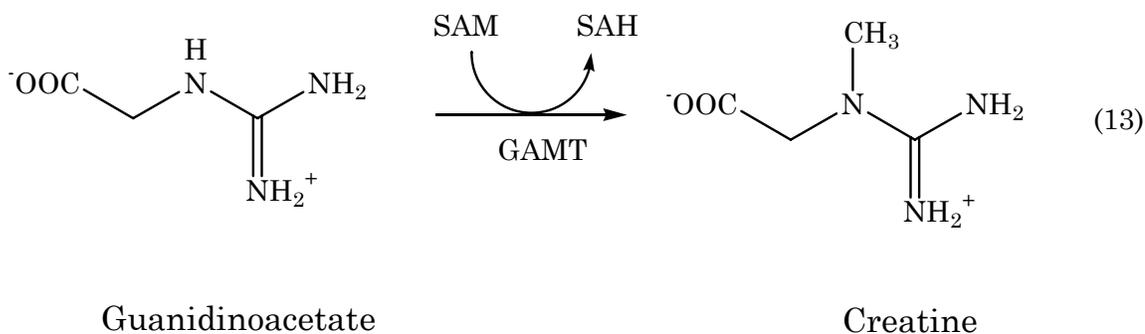


Figure 10.4: A and B) Crystal structure of PPIL3b with polypeptide HAGPILA (PDB code: 1AWQ from Vajdos et al., 1997). C) Predicted secondary structure of similar to PPIL3b protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)

*Similar to GAMT:* Guanidinoacetate N-methyltransferase (GAMT, EC 2.1.1.2) participates in the urea cycle and metabolism of amino groups, glycine, serine, and threonine metabolism, arginine and proline metabolism.

GAMT catalyzes the last step in creatine biosynthesis (reaction 13); the enzyme converts S-adenosylmethionine (SAM) into S-adenosylhomocysteine (SAH).



The mechanism for this reaction (figure 10.5) was reported by Komoto et al. in 2004. The guanidine group of guanidinoacetate forms two pairs of hydrogen bonds with Glu45 and Asp134, while the carboxylate group of the substrate interacts with the backbone amide groups of Leu170 and Thr171.

The crystal structures of GAMT and of similar to GAMT protein are shown in figure 10.6. The enzyme is of the class of alpha-beta proteins with mainly antiparallel  $\alpha$ - $\beta$ - $\alpha$  units.

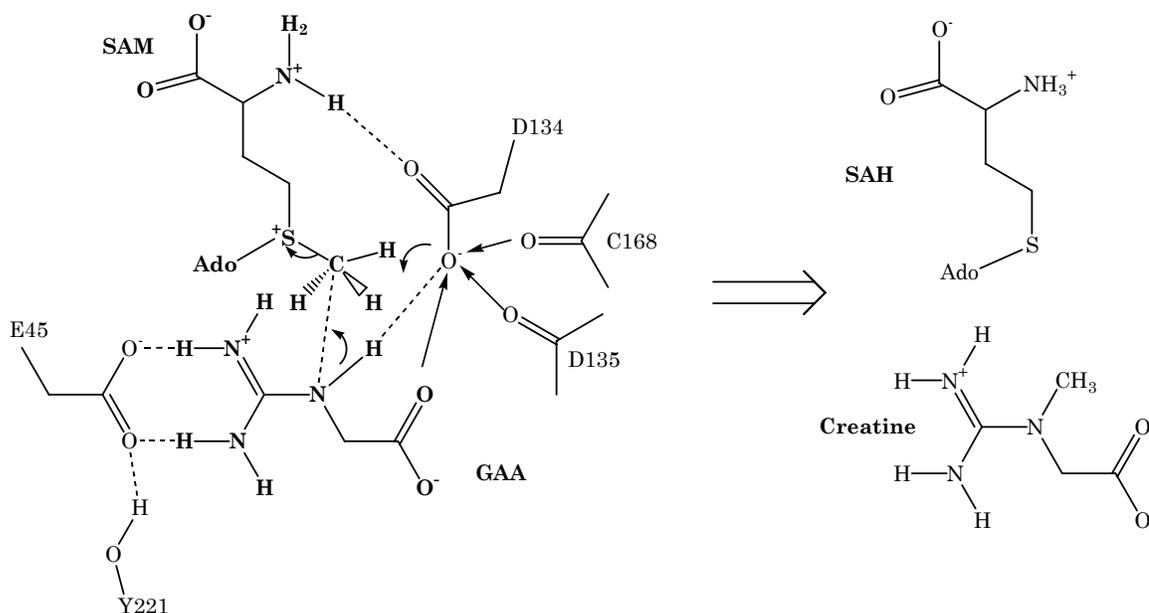
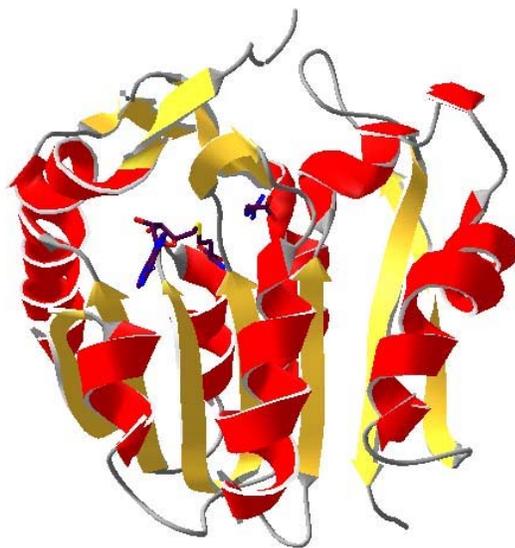


Figure 10.5: Mechanism of methyl transfer by GAMT (Komoto et al., 2004)

*Ufm1*: The eukaryotic ubiquitin-fold modifier-conjugating enzyme 1 (Ufm1) targets molecules in a similar manner to that of the ubiquitin molecule in the ubiquitylation pathway. Ubiquitin is covalently conjugated to the target protein by isopeptide linkage between the carboxyl termini and  $\epsilon$ -amino group of lysine of the target protein.

The bovine enzyme was reported by Sonstegard and coworkers in 2002. In figure 10.7 the crystal structure of this protein is shown. The predicted structure for the unknown protein is obviously just a fragment of Ufm 1.

A



B



Figure 10.6: A) Crystal structure of GMAT with substrates SAM and GAA (crystal structure 1XCJ) (Komoto et al., 2004); B) predicted secondary structure for similar to GAMT protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)

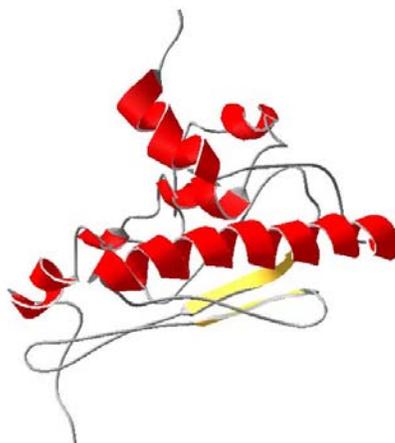
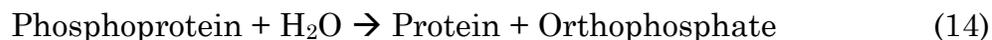


Figure 10.7: Predicted secondary structure of Umf1 protein obtained using the Swiss View modeler (Guex and Peitsch, 1997)

*Unknown Protein for MGC :134378*: Unknown protein for MGC:134378, MGC (Mammalian gene collection), was obtained by computational translation of the bovine genome by the BCCA, Canada. The Vh1-related dual-specificity phosphatase (EC 3.1.3.16) was employed to predict the structure of this protein; these secondary structures are shown in figure 10.8. The structure is of the class of alpha beta proteins, mainly antiparallel beta sheets and  $\alpha/\beta/\alpha$  units. Vh1-related dual-specificity phosphatase was reported as one of the proteins present in the sequenced sample, as protein S.

Vh1-related dual-specificity phosphatase belongs to a group of enzymes that removes the serine- or threonine-bound phosphate group from a wide range of phosphoproteins, as shown in reaction (14).



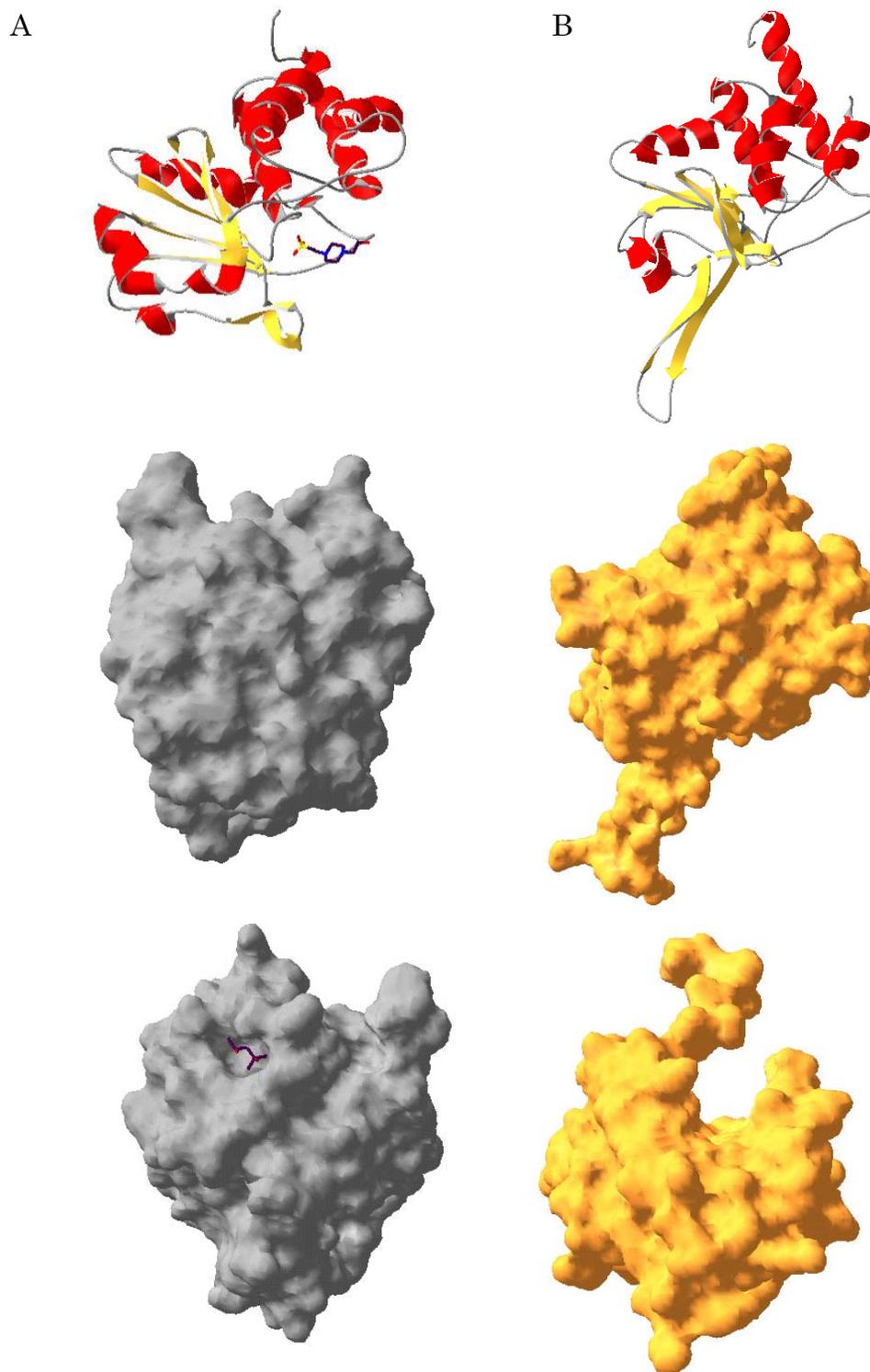


Figure 10.8: A) Human Vh1-Related Dual-Specificity Phosphatase (crystal structure 1J4X) (Yuvaniyama et al., 1996). B) Predicted secondary structure of unknown protein for MGC :134378 obtained using the Swiss View modeler (Guex and Peitsch, 1997)

The mechanism for the hydrolysis catalyzed by dual-specificity phosphatase is shown in figure 10.9. The active group is Cys124, which directly attacks the phosphate atom of the phosphorylated protein and releases the protein.

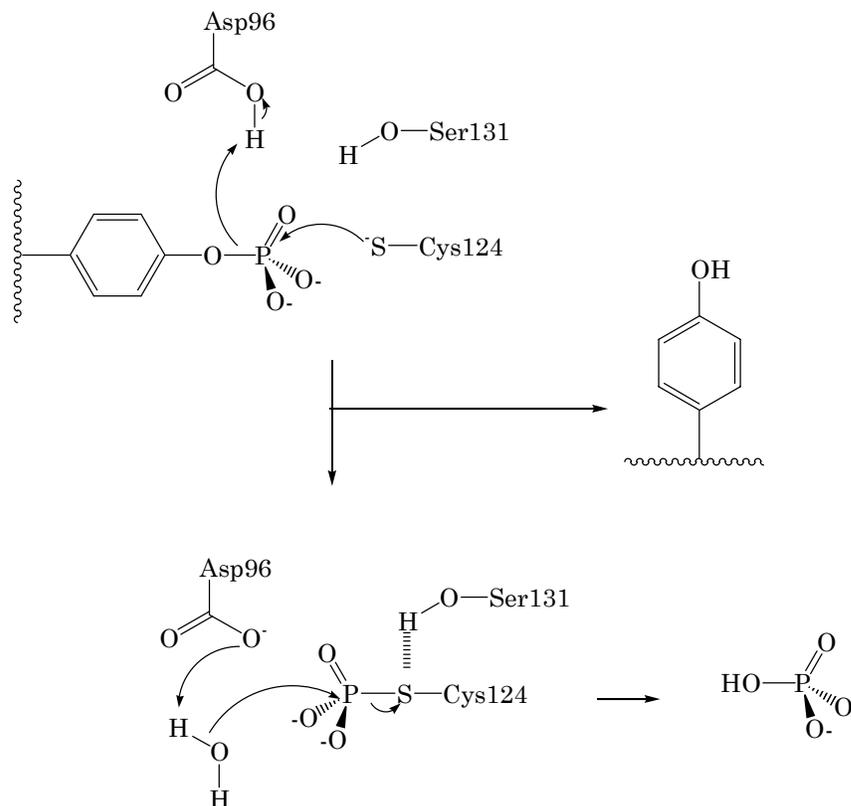


Figure 10.9: Mechanism of Dual-specificity phosphatase. The catalytic triad is the catalytic thiol (Cys124) general acid (Asp92) and hydroxyl group (Ser131) (Denu and Dixon, 1995)

*Prostaglandin H2 D-isomerase*: Prostaglandin H2 D-isomerase catalyzes reaction 15. The predicted secondary structure of the bovine enzyme is of the class of all beta proteins; its fold is classified as lipocalins; a barrel of 8 beta strand binds hydrophobic ligands in its interior, figure 10.10.

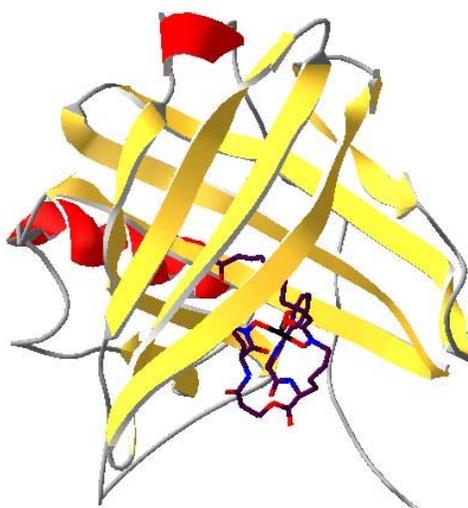
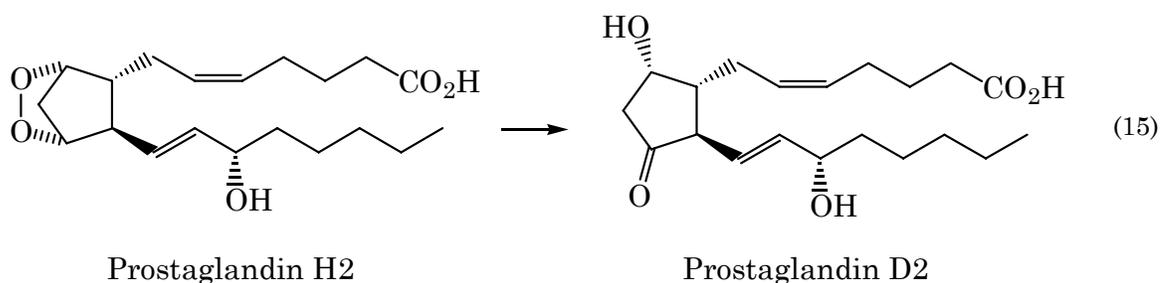


Figure 10.10: Predicted secondary structure of Prostaglandin H2 D-isomerase complexed with carboxymycobactin T obtained using the Swiss View modeler (Guex and Peitsch, 1997)

It should be noticed that the main contaminant is Similar to Troponin C-akin-1 protein; Troponin C-akin-1 complex senses calcium increase and triggers muscles contraction. The single protein has a molecular mass of 19 KD and a pI of 6.72; it is possible that dimerization of this protein is the ~ 40 KD contaminant observed during intermediate and polishing steps.

Also many fragments of high molecular mass proteins like BSA were observed; this indicates that proteolytic degradation creates contaminants of similar characteristics to that of  $\gamma$ GACT during the purification procedure.

Based on molecular mass calculated for these five candidates it can be said that the Unknown protein for MGC :134378, of molecular mass 22.5 KD, has the closest molecular mass to the one observed for  $\gamma$ GACT, 22KD.

Based on the calculated pI of the candidate proteins it can be said that Prostaglandin H2 D-isomerase, of pI 6.91 has the closest pI to the one observed for  $\gamma$ GACT; although, this pI is just a prediction and can not be given a specific degree of confidence, because pI will be influenced by the folding of the protein; consequently, the calculated pI value can not be use to discriminate between the five candidate proteins.

Based on number of cysteine residues, groups that have been of great important in the purification of  $\gamma$ GACT, for these five candidates it can be said that the Unknown protein for MGC :134378, which has 6 cysteine residues, may be closer to the number of cysteine of  $\gamma$ GACT; it must be recalled that  $\gamma$ GAACT was reported to have 7 cysteine groups (Taniguchi and Meister, 1978)

Based on molecular mass calculated for these five candidates it can be said that the Unknown protein for MGC:134378, of molecular mass 22.5 KD, has the closest molecular mass to the one observed for  $\gamma$ GACT, 22KD.

Based on the reaction that the enzyme catalyzes it can be said that Unknown protein for MGC:134378, which has the structure of Vh1-Related Dual-Specificity Phosphatase, is the one whose reaction is closer to  $\gamma$ GACT. Similar to PPIL3b and Prostaglandin H2 D-isomerase catalyze isomerization reaction on substrates that possess a five member ring. The one similar to GAMT catalyzes a transfer reaction from two substrate to form two products, while  $\gamma$ GACT acts on a single substrate; similarly, Ufm1 must catalyze the binding of two proteins acting on two substrates. Even though, Unknown protein for MGC:134378 acts on two substrates, the phosphorylated protein and water the reaction is a covalent bond breakage and not an isomerization or formation of new bonds; for this reason the reaction of Unknown protein for MGC :134378 is matched better over all other four candidates to the reaction of  $\gamma$ GACT.

Finally, based on the structure of the enzyme related to  $\gamma$ GACT it can be said that Prostaglandin H2 D-isomerase, which is of the class of all beta proteins, is the candidate with the poorest structure similarity to the probable structure of  $\gamma$ GACT because the  $\gamma$ GACT related enzymes, mammalian QC,  $\gamma$ GTP, and pyroglutamidase, all are of the alpha-beta class. Prostaglandin H2 D-isomerase is not further considered as a candidate protein.

None of the four remaining candidate proteins used histidine or lysine for their catalysis, except for Ufm1, whose mechanism has not been reported.

Trying to compare the active site of the predicted structure will be more a speculation than a valid result because the actual structure of the active protein should be optimized. The fact that in the predicted structure cysteine residues do not form disulfide bridges, which contribute to the stability of the protein, indicates that the actual structure of these predicted proteins must rotate and become more compact in order to bring together the thiol groups. For example for Unknown protein for MGC:134378, which has 6 cysteine residues, none of them are linked as it can be seen in figure 10.11.

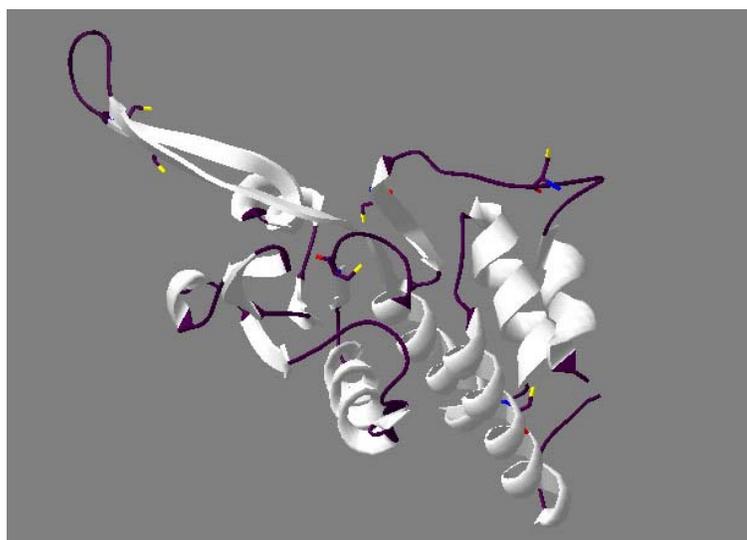


Figure 10.11: Predicted secondary structure of unknown protein for MGC:134378 showing the unbridged thiol groups. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)

The active site of each candidate protein was searched for histidine, lysine, N-terminal amino and tyrosine, which form the proposed active site residues for  $\gamma$ GACT (Gonzalez, 2005). Figures 10.12, 10.13, and 10.14 show the probable active site of the candidate proteins with these residues.

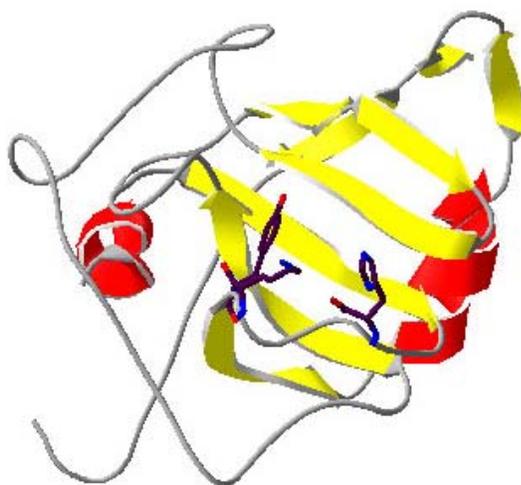


Figure 10.12: Predicted secondary structure of Similar to PPIL3b protein showing the probable active site formed by His84, Tyr89, and Lys88. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)

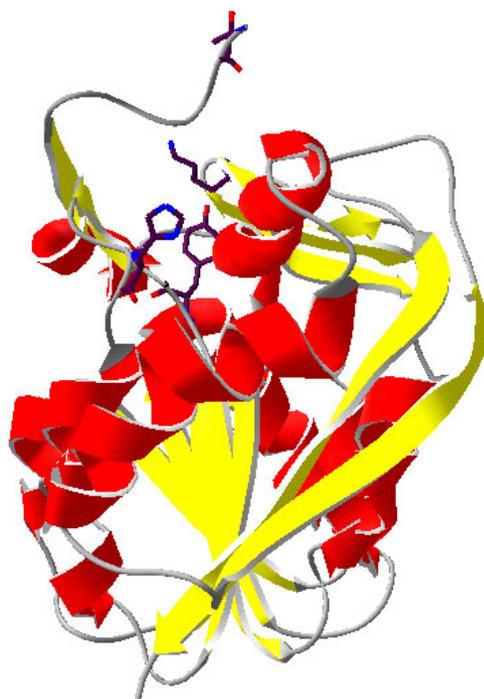


Figure 10.13: Predicted secondary structure of Similar to GAMT protein showing the probable active site formed by His144, Tyr136, and Lys180. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)

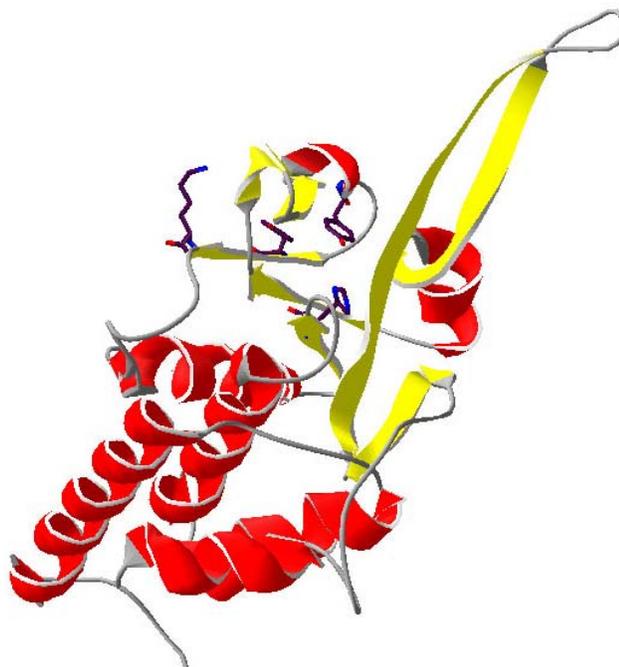


Figure 10.14: Predicted secondary structure of unknown protein for MGC :134378 showing the probable active site formed by His139, Lys105, Tyr94, and Tyr101. Structure obtained using the Swiss View modeler (Guex and Peitsch, 1997)

Most likely the N-terminal amine of the protein is modified and can not participate in the catalysis. The predicted structure of Ufm1 is too poorly assembled and that active site is not proposed for this candidate.

A similarities search between the candidate proteins and bovine QC gave not positive results; this result can only be used to say that  $\gamma$ GACT and QC are no related enzymes although they catalyze a similar reaction.

At this point unknown protein for MGC:134378 is chosen as the most likely enzyme to be  $\gamma$ GACT based on the different characteristics and the finding of a putative active site. However, similar to PPIL3b protein and similar to GAMT protein also must be considered.

## CHAPTER ELEVEN

### Results and Discussion of Affinity Column Design

#### *Determination of Ligand Inhibition Constant*

It was observed that 93 % inhibition of  $\gamma$ GACT activity occurs when equal concentrations (1mM) of substrate and inhibitor were present in the reaction. Figure 11.1A shows the Lineweaver-Burke plot with glutarylhexylamine as inhibitor. This graph shows a common intersection of each line at the  $1/v$  axis, which corresponds to the  $1/V_{\max}$  value. As was expected glutarylhexylamine appeared to be a competitive inhibitor; its inhibition is caused by decreasing the amount of available enzyme for reaction, so  $K_M$  seems to be higher but  $V_{\max}$  is not affected because the enzyme-substrate complex is not altered. From figure 11.1B the  $K_i$  and  $K_M$  can be calculated; based on equation V, the slope of this graph is  $K_M/K_i$ , while the interception at the  $K_{app}$  axis is  $K_M$ . The calculated  $K_i$  for glutarylhexylamine was of 11.6  $\mu$ M of inhibitor and the  $K_M$  was 0.183 mM with  $N^{\epsilon}$ -( $\gamma$ -glutamyl)lysine.

#### *Results of the Two Coupling Methods*

The manufacture recommended coupling method employing a carbodiimide as an activator shows incomplete coupling when observed with fluorescamine even after 24 h, figure 11.2.

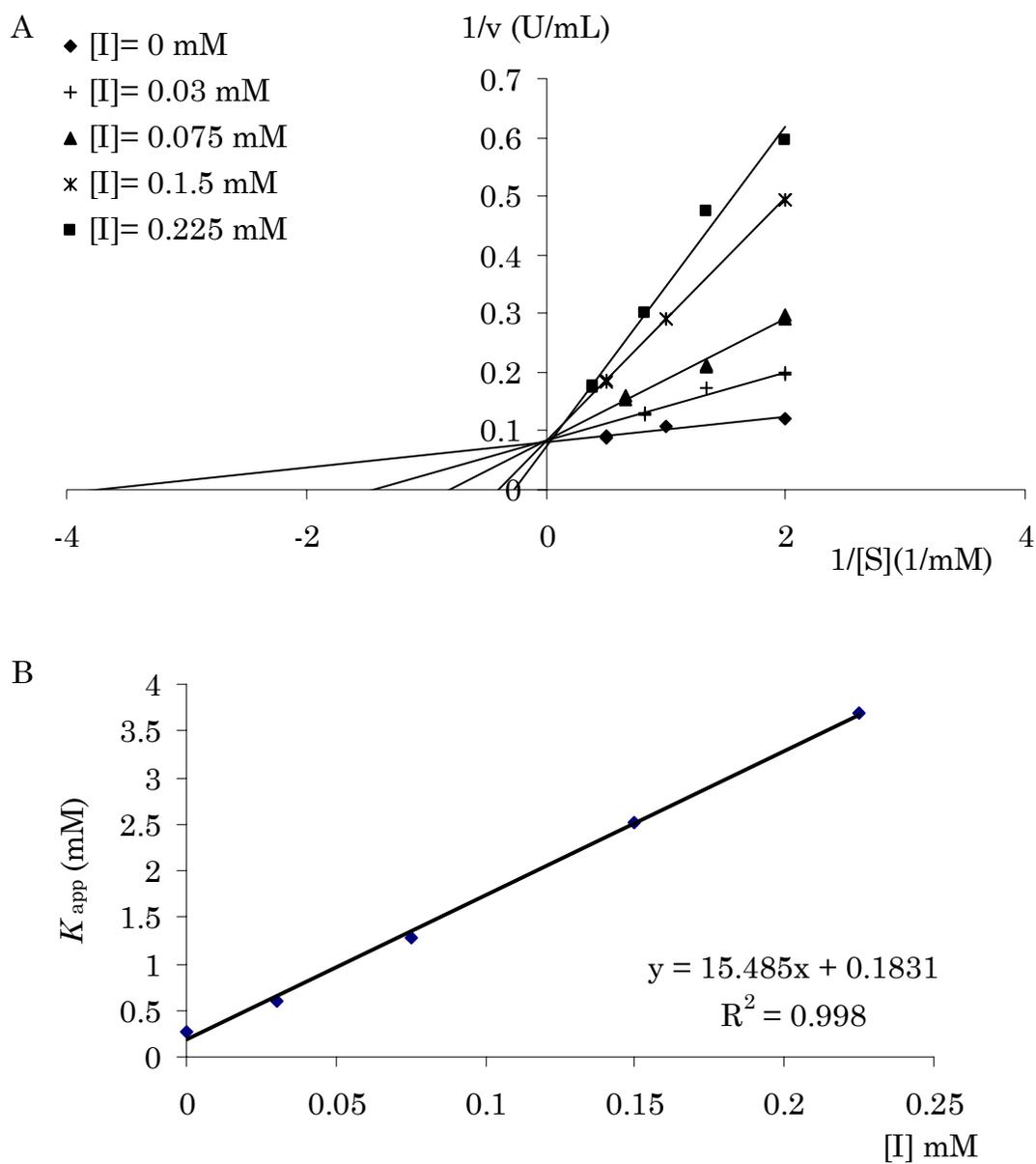


Figure 11.1: A) Lineweaver-Burke plot with glurarylhexylamine as  $\gamma$ GACT inhibitor, B)  $K_{app}$ s versus [I] for the determination of  $K_i$

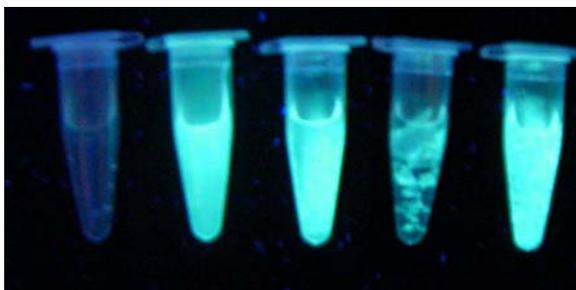


Figure 11.2: Tracking of the coupling reaction using a carbodiimide. From left to right: Negative control, positive control, EAH Sepharose 4B, 1h coupling, 12 h coupling.

Although the carbodiimide reaction is straightforward there are several side reactions that complicate it. The carboxylic acid will react with the carbodiimide to produce an activated carboxylic ester; the reaction of this intermediate with the amine will form the amide and the urea product, route 1 in figure 11.3. However, the rearrangement of the intermediate will form a stable N-acylurea, route 2 in figure 11.3.

The coupling of acids and amines using a carbodiimide activator has been reported to be extremely variable and to produce low yields not only because the o-acylurea intermediate rapidly undergoes hydrolysis or rearrangement into a more stable N-acylurea but also because the formation of the o-acylurea occurs at pH 4-5; at this pH primary amines are protonated and act poorly as nucleophiles (Sehgal and Vijay, 1994).

When the reaction was carried out by the direct reaction of the anhydride on the solid support amine, total coupling was observed in 1 h because no fluorescence was observable, figure 11.4.

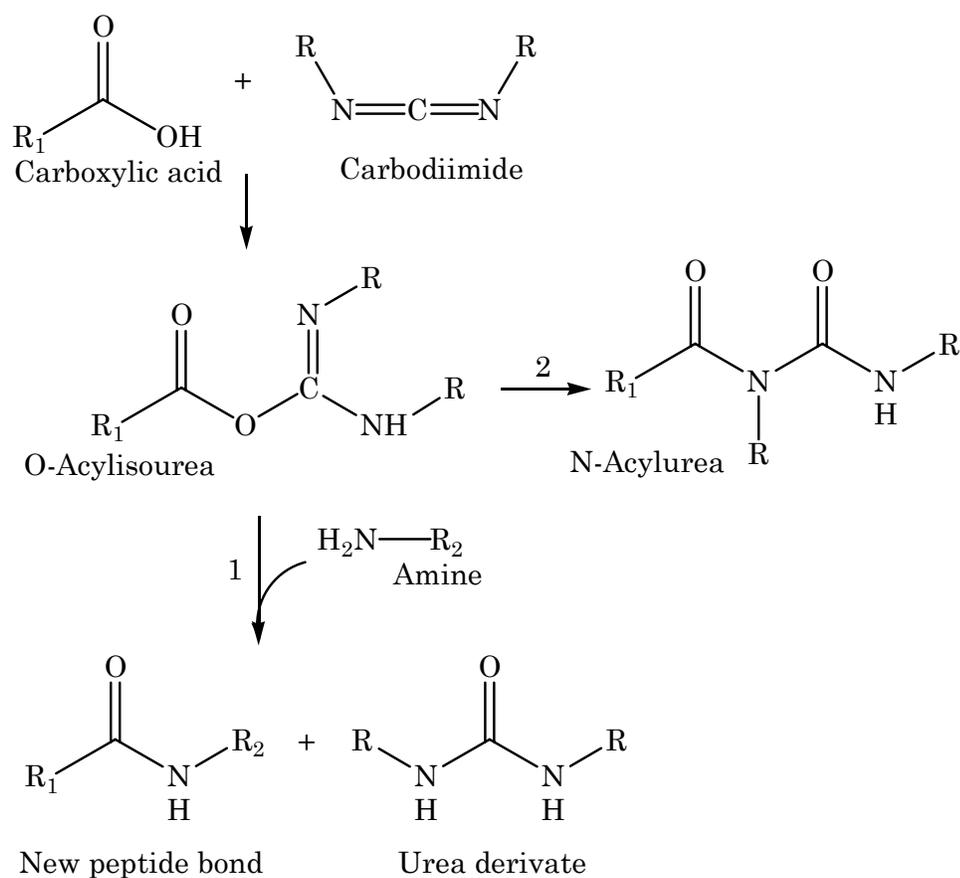


Figure 11.3: Scheme of the carbodiimide reactions; 1) new peptide bond formation, 2) intermediate rearrangement into a stable N-acylurea

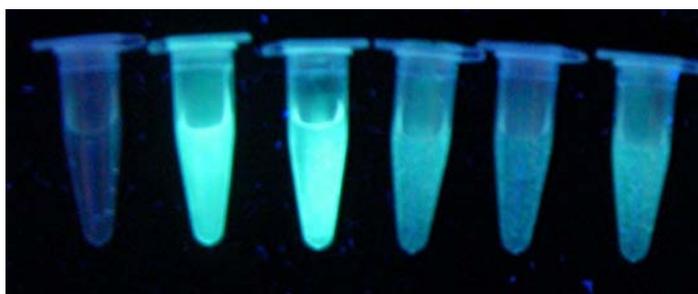


Figure 11.4: Coupling reaction using glutaryl anhydride. From left to right: Negative control, positive control, EAH Sepharose 4B, 1h coupling, 12 h coupling, 24h coupling

*Ligand Leaching under Different Conditions*

After incubation of 1 h of the derivatized matrix at different temperatures, pHs and salt concentrations, it can be concluded that ligand leaching is minimized at low temperatures, acidic pH, and at any concentration of salt, figures 11.5, 11.6 and 11.7, respectively, show this result. The derivatized column was also stable during weekend storage at 4°C in 20% methanol containing 0.5 M NaCl, figure 11.8.

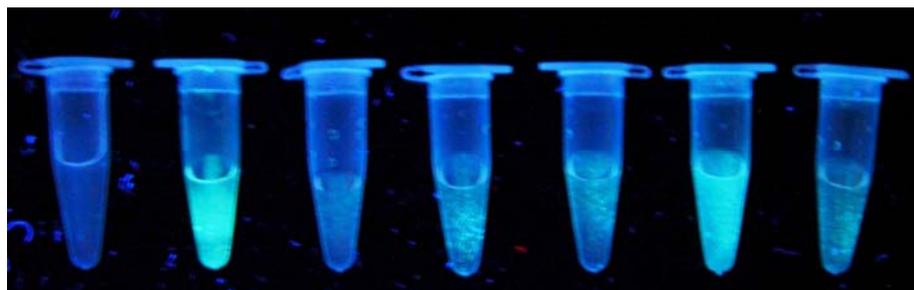


Figure 11.5: Temperature effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, 4, 15, 25, 35, and 45 °C

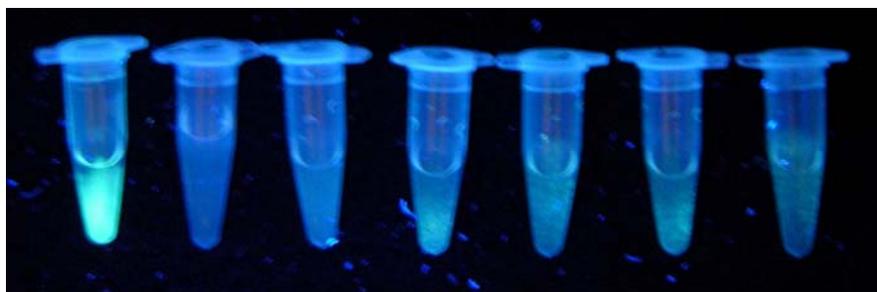


Figure 11.6: pH effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, pH 2, 5, 7, 9, and 12



Figure 11.7: Salt concentration effect on GH Sepharose 4B stability. From left to right: Negative control, positive control, 0, 0.1, 0.25, 0.5, and 1 M NaCl

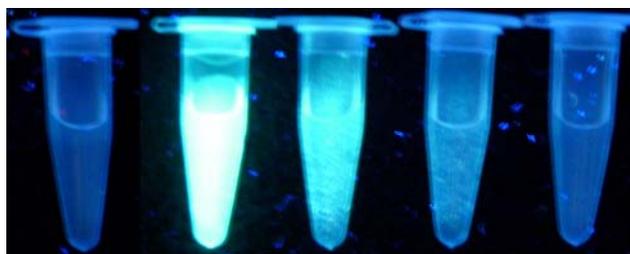


Figure 11.8: Weekend storage of GH Sepharose 4B stability. From left to right: Negative control, positive control, EAH Sepharose 4B, 20% methanol stored, and 20% methanol with 0.5 M NaCl stored

#### *Results of the Purification of $\gamma$ GACToin Glutarylhexylamine Sepharose 4B*

The column was packed in a 0.5 x 5 cm column shirt and connected to the FPLC apparatus; no retention of the sample was observed in 1 M NaCl, so this concentration of salt can be use to removed the enzyme from the column.

When the enzyme was run initially with buffer with 0.05 M NaCl followed by salt increase to 1 M of salt, the elution profile figure 11.9 was observed for 13 Units and figure 11.10 for 3 Units.

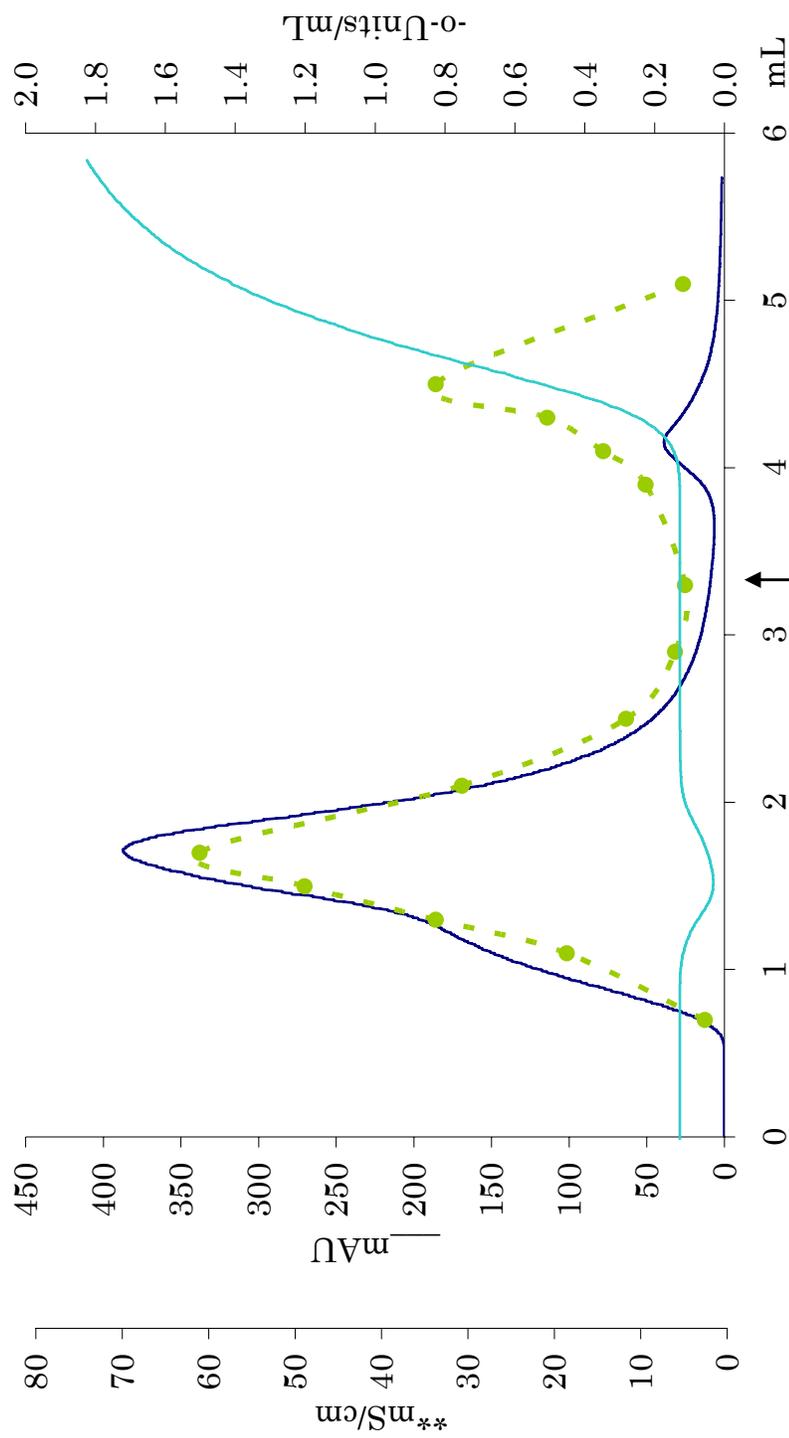


Figure 11.9: Elution profile of 13Units of  $\gamma$ GACT separation in Glutarylhexylamine Sepharose 4B. The arrow indicates the introduction of buffer with 1 M NaCl. \*\* Conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h)

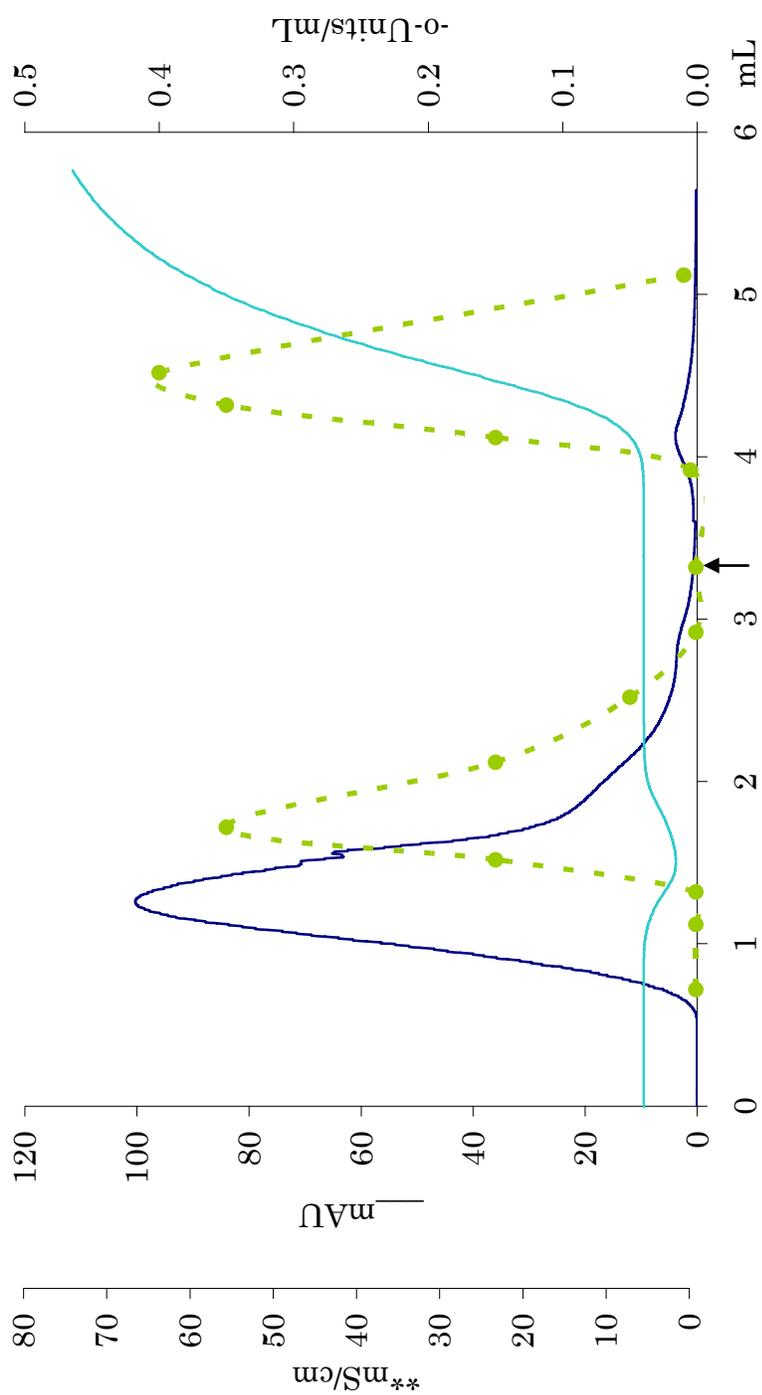


Figure 11.10: Elution profile of 3 Units of  $\gamma$ GACT separation in Glutarylhexylamine Sepharose 4B. The arrow indicates the introduction of buffer with 1 M NaCl. \*\* conductivity, \_\_\_ Absorption of the sample at 254 nm, -o- units of activity ( $\mu$ mol Lys/h)

The area for the first absorption peak is approximately 313 mAU\*mL, while the second peak area is ~ 17 mAU\*mL; this mean that 95 % of the total protein is unretained by the column, while only 5 % is retained. On the other hand, the activity area of the first peak is 7.1 Units, and for the second peak is 2.7 Units; this result says that more than 27% of the enzyme is retained by the column. A total recovery of 75% was observed for this column.

When the sample was run diluted, only 3 Units a smaller total recovery of 57% was observed; however, the amount of enzyme retained by the column was of 56% of the total recovered. This result shows that the enzyme binds to the column; however, no increase in the specific activity was gained because the protein still remains attached to other contaminants.

## CHAPTER TWELVE

### Discussion and Conclusions

#### *Purification*

In this purification of the bovine  $\gamma$ GACT, the enzyme was purified more than 2,000 fold for a final specific activity of more than 1,300U/mg of enzyme; which is the highest reported specific activity for the bovine enzyme. Table 12.1 contrasts the results obtained in this work and those obtained by Gowda (1985) in the purification of the 50 g of bovine enzyme.

The ion exchange chromatography step in the gross purification as well as the size exclusion chromatography step in the intermediate purification showed significant improvement of the purification. During the gross purification, Gowda (1985) used DEAE cellulose and recovered 1230 units of enzyme of specific activity 0.79 U/mg, while in this work using DEAE-Sepharose the recovered enzyme was 1350 Units of specific activity of 3 U/mg. Similarly, for the intermediate purification size exclusion chromatography carried out by Gowda (1985) on a Sephadex G100 the recovered protein was 915 Units of activity of a 20.2 U/mg specific activity; in this work the 694 Units of activity were recovered of a three times more pure enzyme of 58 U/mg specific activity.

Table 12.1: Comparison of the results for the purification of 50 g of bovine enzyme

Gowda purification (1985)	Volume (mL)	Total protein (mg)	Total Units ( $\mu\text{mol/h}$ )	Specific Activity (U/mg)	Recovery (%)	Fold
Homogenate	152	8,360	2,882	0.34	100	
Ultracentrifugation Supernatant	95	2,867	1,714	0.6	59.5	2
DEAE Cellulose	410	1,590	1,230	0.79	43	2
Sephadex G100	44	45.3	915	20.2	32	59
IEF	0.04	0.012	5.84	487	-	-

This purification	Volume (mL)	Total protein (mg)	Total Units ( $\mu\text{mol/h}$ )	Specific Activity (U/mg)	Recovery (%)	Fold
Homogenate	200	7,200	4,498	0.6	100	
Ultracentrifugation Supernatant	150	2,455	3,262	1.3	72	2
DEAE Sepharose	250	520	1,350	3	30	5
Sephacryl S100	80	12	694	58	15	97
IEF	0.05	<0.001	1.3	>1300	0.03	2166

Gowda (1985) found a dimeric protein of molecular mass 27,000 D for the enzyme under non-denaturing conditions and of 12,700 D under denaturing conditions. The molecular mass of the protein determined here by size exclusion chromatography on a calibrated Superdex HR 75 and SDS-PAGE was of 22,000 D under reducing and non-reducing conditions.

In this purification  $\gamma$ GAACT is removed in the intermediate purification by Sephacryl S100; the fact that the enzyme can be separated by size exclusion chromatography indicates that they have different molecular masses; the molecular mass of  $\gamma$ GAACT was reported by Taniguchi and Meister (1978) to be 27,000 D, which should elute before  $\gamma$ GACT on the Sephacryl S100 column.

The use of phosphate buffer for ion exchange chromatography is not appropriate because the buffer and the enzyme will have the same charge to bind to the column; both phosphate buffer and the enzyme will compete for binding sites in the column, in consequence the protein does not preferentially bind to the column and purification can not be obtained. Although the purification on Mono Q GL column presented many difficulties like the preparation of the sample in a buffer that has shown to cause problems not only to  $\gamma$ GACT (Browser, 1997) but also to glutamyl cyclase (Wintjens et al., 2006) the enzyme as concentrated as 39 U / mL was stable in Tris buffer for as long as 100 minutes, which is enough time for a separation on a Mono Q GL column and assay of specific activity.

Although EGTA does not have a negative effect on the activity of the enzyme, it prevents the binding of the enzyme to the Mono Q GL column, so this additive can not be used in this conditions. The reduction of the sample using 1% 2-ME caused less activity lost than reduction with 1 mM DTT; activity lost was ~20% with 2-ME and 65% with DTT. Hence, 2-ME was

employed for the study of the effect of reducing agents in Mono Q GL column separation.

For the optimization of the separation on a Mono Q GL column a progressive change of the parameters available, pH, gradient, injection volume, flow rate, concentration of buffer, and presence of reducing agent were varied one at the time to observe the effect.

Some parameters can not be changed too drastically without altering the nature of the separation; for example, pH can not be increased or decreased to far from pH 8 because at higher pH the protein is inactivated, and a lower pH the protein become neutral or positive; also the increase of enzyme concentration can cause salting out of the enzyme, while dilution can cause lost of the enzyme due to its binding to the contained. Other parameter as flow rate did not produce any change, as expected for an ion exchange chromatography.

The best purification in Mono Q GL column using Tris buffer pH 8 as running buffer and salt increased to induce the elution was the one containing reducing agent. Under non-reducing conditions a broad peak of activity is observed; in contrast, under reducing conditions several peaks of activity were observed; the fact that the best purification was gained under reducing condition, where one peak of activity grows over the other, indicates that the protein is been concentrated in one form.

The main form of the enzyme of pI 6.86 and 6.62 were under non-reducing conditions; by adding reducing agent the enzyme of pI 6.62 is enriched. This observation agrees with the rationalization that the addition of reducing agent that breaks disulfides bridges into thiol groups creates a more acidic protein; consequently, the pI of the protein should decrease. A more acidic protein will bind stronger to the anion exchange and require more salt concentration to be eluted than a less acidic protein.

Alkylation of the sample with a low concentration of alkylating agent showed poor effect on the inactivation of the enzyme (Gowda, 1985); however, a greater concentration of alkylating agent decreased the activity. Iodocadetamide alkylates cysteine thiol groups, and at a much smaller rate, it alkylates the imidazole ring of histidine.

The fact than inactivation is obtained at high concentration of IAA is not enough to think that histidine and not cysteine is present in the active site of the enzyme; it must be also be considered that cysteine groups are involved in the stability of the enzyme, and this can be the cause of activity loss. This second though is supported by the fact that 2-ME does not caused as much activity loss as DTT, which is a stronger reducing agent due to the formation of an intramolecular disulfide ring when oxidized.

Although alkylation did not improved the purification with ion exchange chromatography this was useful for the recovery of a single band

from IEF. The non-alkylated sample showed activity in IEF at pH 6.18, 6.54, and 6.76; while, the alkylated sample showed activity at pH 6.45.

Electroelution of the IEF gel allowed the recovery of the sample that Gowda (1985) tried to obtain by dialysis. Electroelution has many more advantages over dialysis because the protein is not allowed to diffuse and equilibrate in a medium, action that takes days, instead the protein is forced out from the gel by a current in a electric field. Electroelution takes only 1 hour, but it does not remove ampholites; nevertheless, ampholites are washed out from the sample in minutes by ultrafiltration, which can concentrate the sample into a volume as small as 0.1mL.

DEAE Sepharose and Sephacryl S100 really improved the purification procedure of Fink and Folk (1983) with contributions from Gowda's (1985) and Bowser's (1997) procedures. Additionally the use of electroelution, and ultrafiltration for the handling of proteins also helps to increase the efficiency of the purification. This work proves again that the purification of an enzyme is dependent of the information that can be gained during its development; all the work is done only to come back to the initial point and repeat it again but this time taking into account the knowledge gained in the previous purification.

The protein was obtained as pure as the techniques allows us to determine its purity; the electrophoretically pure band that seems to be a single protein is actually composed of a number of proteins. From these

estimates, the pure protein is expected to be more active, perhaps 10,000 - 60,000 U/mg.

### *Sequencing*

Edman degradation is the preferred sequencing methodology since it sequentially counts the amino acids from the N-terminus of the protein toward the C-terminus; although the requirement of a free N-terminal is not accomplished by many cytosolic enzymes.

No free N-terminal amino group for  $\gamma$ -GACT was found by dansylation of the gel extracted enzyme, followed by hydrolysis and finally resolution by HPLC with detection of the fluorescent dansyl group. It can be said that the methodology works because the standard amino acids and the standard protein, carbonic anhydrase, showed signal of typical amino acid dansylation of proteins.

The fact that gel extracted  $\gamma$ GACT did not show peaks of typical dansylation can only be used to conclude that there was not enough protein to do the assay; the conclusion of a blocked or free N-terminal for  $\gamma$ GACT can not be concluded from this experiment.

The enzyme of higher purity was sent for sequencing by an external laboratory, Harvard Microchemistry and Proteomic Analysis Facility; the report from the amino acid sequencing using internal trypsin digestion of the proteins, microcapillary HPLC, nano ESI, and Ion Trap MS/MS revealed that

the electrophoretically pure band contained 42 proteins and protein fragments of similar mass and pI as that of  $\gamma$ GACT.

A list of the physical characteristics of the reported proteins was compiled and the proteins were analyzed as obvious contaminations, or selected for further studies if they had similarities to  $\gamma$ GACT. From this selection the following five candidate proteins, from *Bos Taurus*, were selected: (1) similar to peptidylprolyl isomerase-like protein 3, (2) hypothetical protein LOC515270 also similar to Guanidinoacetate N-methyltransferase, (3) hypothetical protein LOC537221 also Ufm1-conjugating enzyme 1, (4) unknown protein for MGC:134378, and (5) Prostaglandin H2 D-isomerase.

Information on the sequences of these candidate proteins were sent to Dr. Baker (Baylor University) to compare with the human and plant glutamyl cyclase enzymes based on the observed similarities of their catalysis with  $\gamma$ GACT catalysis.

Similarities were not found between the candidates and QC enzyme, but considering that the mammalian and the plant QC enzymes are so poorly related, this is not an unexpected result. Mammalian QC used zinc for its catalysis while the plant enzyme activity is independent of the metal.

From the analysis of the properties of these candidate enzymes and  $\gamma$ GACT the unknown protein for MGC:134378 is proposed as the most likely

protein in the mixture to be  $\gamma$ GACT and as second options similar to PPIL3b protein and similar to GAMT protein.

### *Affinity Column*

The calculated  $K_M$  of the substrate reaction was 0.183 mM with  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine; this  $K_M$  is almost half the value reported by Gowda (1985) of 0.385 mM; although, the protein purified by her was not proved  $\gamma$ GAACT free.

Glutarylhexylamine showed 93 % inhibition of  $\gamma$ GACT at equalmolar concentration of substrate and inhibitor, 1 mM; the calculated  $K_i$  for this inhibitor was of 11.6  $\mu$ M; this value is in the range of the one reported by Gonzalez (2005) of 7.2  $\mu$ M for the action of rabbit  $\gamma$ GACT on the same inhibitor. Glutarylhexylamine shows to be a good inhibitor of the enzyme; its structure resembles the one of the substrate  $N^\epsilon$ -( $\gamma$ -glutamyl)lysine but it does not have the  $\alpha$ -amine of the glutamyl moiety. This  $\alpha$ -amine is required for nucleophilic attack onto the  $\gamma$ -carbonyl group, which is the proposed first step in the cyclotransferase reaction. This high binding affinity of the enzyme toward glutarylhexylamine implies that the  $\alpha$ -amino group is not very important for binding of the substrate.

The  $K_i$  of glutarylhexylamine is suitable for the interaction required in affinity chromatography, which is in the range of  $10^{-4}$  to  $10^{-8}$  M during the binding step.  $K_i$  of free glutarylhexylamine is of  $1.2 \times 10^{-5}$  M, but it must be

kept in mind that the dissociation constant is expected to increase when the ligand is attached to the column due to steric effect.

The fluorescamine procedure was employed to follow the synthesis of the column by the disappearance of free amine groups was shown to be quick and easy, but insufficient for the quantification of the reaction. The synthesis of the matrix was obtained by direct reaction of glutaryl anhydride and the amine group of the EAH Sepharose 4B column in aqueous media. Glutaryl anhydride will be readily hydrolyzed in water to form the diacid that will not react with the amine; since hydrolysis can not be prevented excess glutaryl anhydride was added to ensure reaction. Even though, the recommended methodology for the coupling of the acid ligand and the amine solid support employs the carbodiimide reaction, this reaction was not efficient for this preparation.

Ligand leaching has been observed even for the very stable amide bond (Sudesh and Lyddiatt, 1997). Ligand leaching was studied under different conditions of pH, temperature, and salt concentration, which are the parameters to modify in order to obtain elution of the protein from the affinity column just by changing the affinity interaction. Elution can be obtained by adding the free ligand because the dissociation constant of the enzyme with the free ligand is always smaller than the dissociation constant of the enzyme with the immobilized ligand; however, elution with ligand will require the removal of the ligand for further studies of the enzyme.

The matrix showed to be more stable at low temperatures, acidic pH, and at any concentration of salt. The addition of 0.05 M NaCl in 20% ethanol was enough to conserve the column for 4 weeks. However, the leaching of the ligand so readily in an hour seems suspicious; perhaps not all the ligand was covalently attached to the solid support and it may have masked the fluorescamine signal.

A trial experiment for the future application of this glutarylhexylmamine affinity column in the purification of  $\gamma$ GACT was to verify the binding and release of the enzyme to the column. The column has a negative charge because the terminal group is a carboxylic acid; salt concentration in the buffer is always required to prevent cation exchanger behavior of the column. When the sample was loaded a big UV peak of protein eluted unbound to the column; this peak showed the majority of the activity; a second peak of activity was observed to elute when ion strength was increased. Diluted sample still showed elution of unbound activity, but more relative retention of the enzyme showing that the enzyme binds to the column. The fact that specific activity did not increase shows that contaminants are also retained with  $\gamma$ GACT, perhaps the column act on them as a cation exchanger.

#### *Future Work*

The purification of the enzyme to an electrophoretically pure band does not ensure that the enzyme is completely pure. This work reveals the

identity of many contaminants of  $\gamma$ GACT. The characteristics and nature of these contaminants must be studied in order to revise the purification procedure and introduce the pertinent modifications.

Also, the characteristics of the candidate enzymes likely to be  $\gamma$ GACT must be taken into account for the design of discriminatory experiments; for example, the assay of the enzyme activity with substrates of these candidate enzymes. Nevertheless, only the cloning and expression of the candidate enzyme can assure that  $\gamma$ GACT gene has been identified.

For future work, much care has to be taken about proteolytic degradation since during the purification procedure contaminants of similar characteristics to that of  $\gamma$ GACT are introduced. Inhibitor cocktails can be added during the gross purification and much care has to be taken when keeping the enzyme on ice during the polishing steps carried out at room temperature like the chromatographic steps in the FPLC system and the digestion procedures like reduction and alkylation.

Glutarylhexylamine affinity column showed binding of the enzyme although contaminants are also retained. Conditions for the selective binding of  $\gamma$ GACT have to be optimized in order to use this column not only for the purification of the enzyme but also for the study of the interaction enzyme-inhibitor.

## APPENDICES

## APPENDIX A

### Characteristics and Use of Chromatographic Columns

*DEAE Sepharose Fast Flow Anion Exchange Column*

Table A1: DEAE Sepharose column characteristics (From Amersham)

DEAE	Dimethylaminoethyl
Dimensions	26 x 4.8 cm
Column volume (CV)	470 mL
Total ionic Capacity	0.11-0.16 mmols / mL gel
Bead structure	6% highly cross-linked agarose
Bead size range	45-165 $\mu$ m
Mean particle size	90 $\mu$ m
Maximum linear flow rate	750 cm/h at 25 °C, 100 KPa
Maximum operating pressure	0.3 MPa (3 bar, 42 psi)
pH working range	3 - 9
Chemical stability	All commonly used aqueous buffers, 1M acetic acid, 1 M NaOH, 8 M Urea, 8M Guanidine HCl, ethanol, methanol

*Reagents and Equipment*

DEAE Sepharose anion exchanger column (Amersham), buffer A = 5 mM KPi pH 7.5 of conductivity  $\sim$  1.1 mS/cm, buffer B = 10 mM phosphate buffer pH 7.5 with 0.3 M NaCl of conductivity  $\sim$  28 mS/cm, 20% Ethanol, 2 M NaCl (Sigma), 1 M NaOH (Sigma), peristaltic pump (Amersham), fraction collector (Gilson FC203), gradient maker (Amersham)

### *Running and Cleaning Procedure*

The column is stored in 20% methanol; ethanol is removed by passing 8 CV of water; then the matrix is activated with 1 CV of buffer B, and equilibrate with 40 CV of buffer A (until eluent's pH and conductivity are the same than the pH and conductivity of buffer A).

In order to load the sample, remove the liquid at the top of the column without letting the column run dried; add the sample without disturbing the top of the column, and introduce this at the maximum flow rate of the peristaltic pump, ~ 8mL/min when the thick pump hose is employed. When the entire sample has entered into the matrix, add buffer without disturbing the top of the column and replace the cap; wash out unbound proteins with 1CV of buffer A. In order to elute the sample, connect the gradient maker with 250 mL of buffer A and 250 mL of buffer B; begin gradient and add more buffer B after the gradient has finished. Collect 8 mL/tube at 1min/tube.

For cleaning of the column after elution of the sample, reverse the flow rate and introduce 0.5 CV of NaCl solution followed by 0.5 CV of NaOH; next wash the column with water until neutral pH elutes. For storage of the column, pass 3 CV of methanol solution (eluent should smell like ethanol); if the column is going to be reused, reverse flow rate and activate and equilibrate as previously mentioned.

*Sephacryl S100 Size Exclusion Column*

Table A2: Sephacryl S100 column characteristics (From Amersham)

Matrix	Cross-linked copolymer of allyl dextran and N,N-methylenebisacrylamide
Dimensions	2.6 x 82 cm
Column volume (CV)	435 mL
Void Volume (Vo)	174 mL
Separation range ( $M_r$ )	$1 \times 10^3 - 1 \times 10^5$
Mean particle size	47 $\mu\text{m}$
Theoretical plates	> 5000 $\text{m}^{-1}$
Recommended flow rate	15 cm/h at room temperature
Maximum operating pressure	0.15 MPa (1.5bar, 21 psi)
pH working range	3 - 11

*Reagents and Equipment*

Sephacryl S100 column, running buffer= 50 mM KPi pH 7.5 with 0.15 M NaCl), 0.2 M NaOH (Sigma), peristaltic pump (Amersham), fraction collector (Gilson FC203), 10 % Blue dextran (sigma) and 1% of the following standards from sigma: Hemoglobin (66 KD), Lactic dehydrogenase (36 KD), Carbonic anhydrase (29 KD), Myoglobin (17 KD), Cytochrome C (12.4 KD), refrigerated microcentrifuge (eppendorf 5415R).

### *Running and Cleaning Procedure*

The column is stored in 20% methanol; remove ethanol by passing 1.5 CV of water then equilibrate with 2 CV of running buffer (until eluent pH and conductivity equal the pH and conductivity of running buffer).

In order to load the sample, remove the liquid at the top of the column without letting the column run dried; add the sample without disturbing the top of the column at 0.5 mL/min. When the entire sample has entered into the matrix, add buffer without disturbing the top of the column, and replace the cap; set flow rate at 0.2 mL/min and collection and 20 min/tube.

For cleaning of the column after elution of the sample pass 0.5 CV NaOH at 0.5 mL/min; next wash the column with water until neutral pH elutes. For storage of the column, pass 2 CV of methanol solution (eluent should smell like ethanol); if the column is going to be reused requilibrate as previously mentioned.

### *Calibration*

Blue dextran elution volume defines the void volume. Standards are loaded into the column and run. The calibration curve of molecular mass versus elution volume gave the linear equation for the separation in this column:  $MW_{S100} = 10^{(-0.0069 \text{ mL} + 3.3347)}$

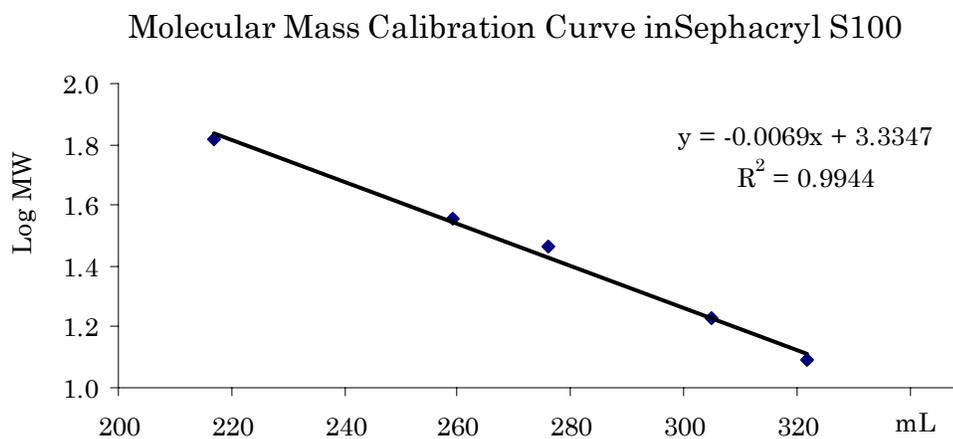
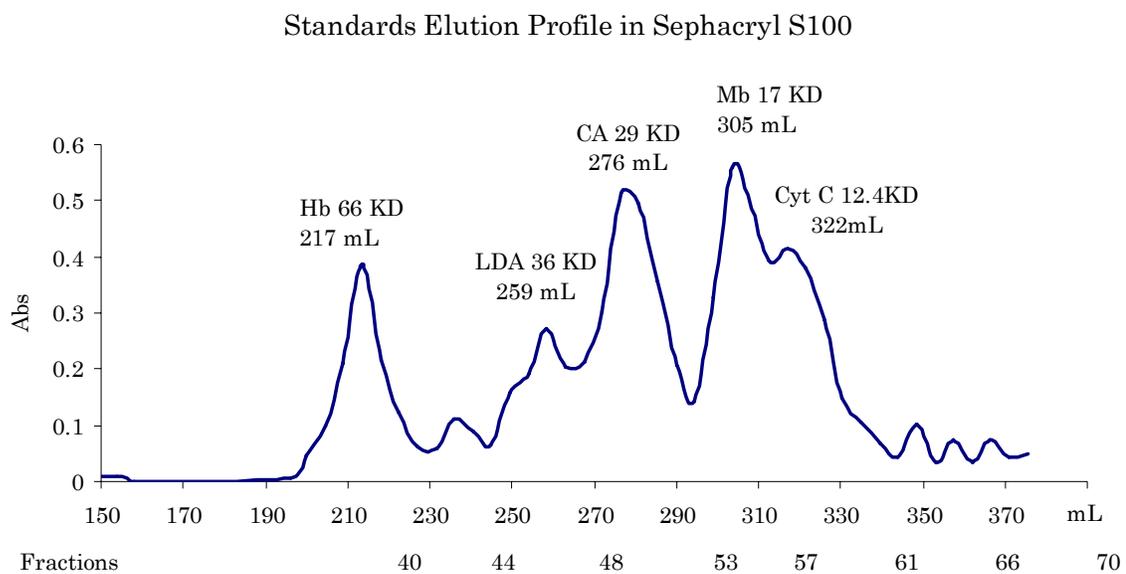


Figure A1: Calibration of Sephacryl S100

*Superdex HR 75HR 10/30 Size Exclusion Column*

Table A3: Superdex HR 75column characteristics (From Amersham)

Matrix	Dextran covalent bound to highly cross-linked porous agarose
Dimensions	1 x 30 cm
Column volume (CV)	25 mL
Void Volume (V <sub>0</sub> )	7 mL
Exclusion limit (M <sub>r</sub> )	100000
Separation range (M <sub>r</sub> )	3000 - 70000
Mean particle size	13 μm
Theoretical plates	> 30000 m <sup>-1</sup>
Maximum flow rate	110 cm/h (1mL/min)
Maximum operating pressure	1.8 MPa (18 bar, 260 psi)
pH working range	3 - 12

*Reagents and Equipment*

Superdex HR 75HR 10/30 column, running buffer= 50 mM KPi pH 7.5 with 0.15 M NaCl), 0.2 M NaOH (Sigma), FPLC system with UV detector AKTA (Amersham), 1% Blue dextran (sigma) and 0.25 % of the following standards from sigma: Carbonic anhydrase (29 KD), Myoglobin (17 KD), Aprotinin (6.5 KD), refrigerated microcentrifuge.

### *Running and Cleaning Procedure*

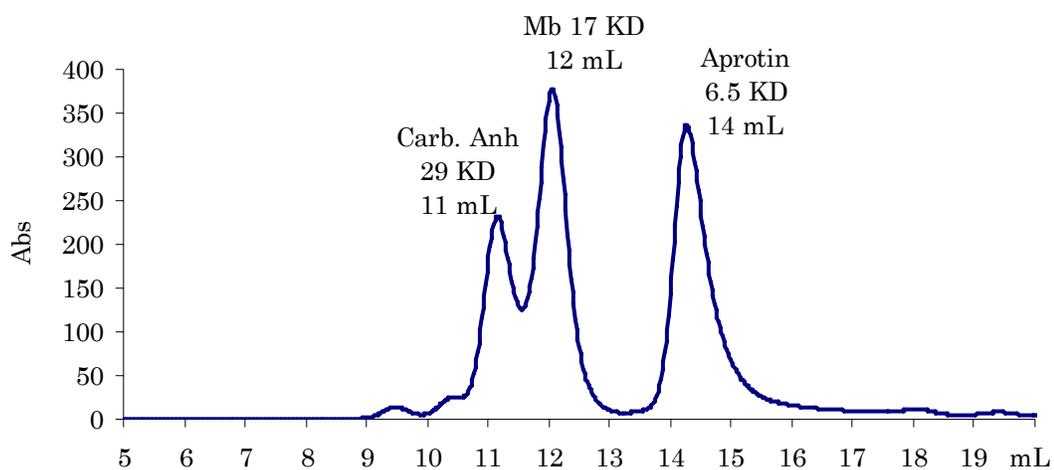
The column is stored in 20% methanol; remove ethanol by passing 1.5 CV of water; then equilibrate with 2 CV of running buffer (until eluent pH and conductivity equal the pH and conductivity of running buffer).

0.1 mL of sample is injected and run at 0.2 mL/min; for size exclusion everything should elute in 1CV. For cleaning of the column after elution of the sample pass 0.5 CV NaOH at 0.5 mL/min; next wash the column with water until neutral pH elutes. For storage of the column, pass 2 CV of methanol solution (eluent should smell like ethanol); if the column is going to be reused requilibrate as previously mentioned.

### *Calibration*

Standards are loaded into the column and run. The calibration curve of molecular mass versus elution volume gave the linear equation for the separation in this column:  $MW_{S75} = 10^{(-0.2154 \text{ mL} + 3.8251)}$

## Standards Elution Profile in S75



## Molecular Mass Calibration Curve in Superdex S75

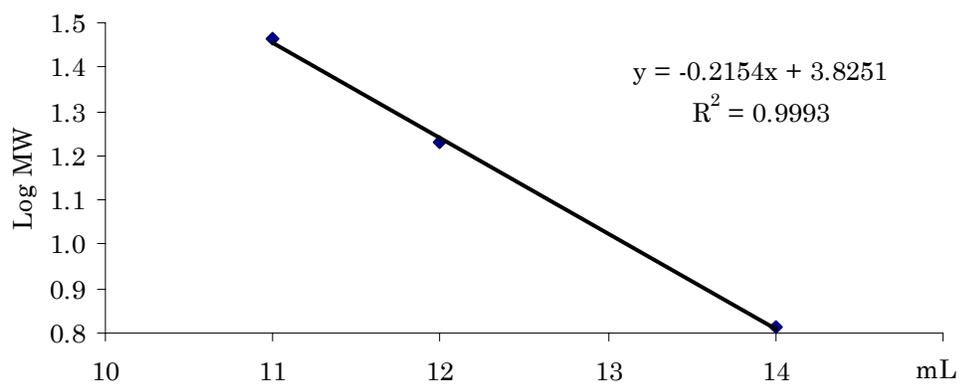


Figure A2: Calibration of Superdex S75

*Mono Q GL 5/50 Anion Exchange Column*

Table A4: Mono Q GL column characteristics (From Amersham)

Matrix	Polystyrene/divinyl benzene
Dimensions	0.5 x 5 cm
Column volume (CV)	1 mL
Total ionic Capacity	0.27-0.37 mmols Cl <sup>-</sup> / mL gel
Bead structure	Rigid, spherical, porous monodisperse
Mean particle size	10 µm
Maximum linear flow rate	3 mL/min at 25 °C
Maximum operating pressure	4 MPa (40 bar, 580 psi)
pH working range	2-12

*Reagents and Equipment*

Mono Q GL strong anion exchanger column (Amersham), initial buffers: 5 mM KPi pH 7.5 or 20 mM Tris pH 8, elution buffers: 10 mM phosphate buffer pH 7.5 with 0.3 M NaCl or 20 mM Tris pH 8 with 0.3 M NaCl, 20% Ethanol, 2 M NaCl (Sigma), 1 M NaOH (Sigma), 75% acetic acid, FPLC system AKTA (Amersham).

*Calibration*

One mg of individual standard was injected in the column and run under condition # 1: buffer A=10 mM Tris buffer pH 8 and buffer B=A+0.3 M NaCl. Gradient: 10% B for 5CV, then to 30% B in 30 CV.

Table A5: Results for the calibration standards of Mono Q GL

Standard	pI	Conductivity (mS/cm)
Carbonic Anhydrase (human)	7.3	4.3
Phosphorylase B	6.3	10
$\alpha$ -Lactalbumin	5.1	17

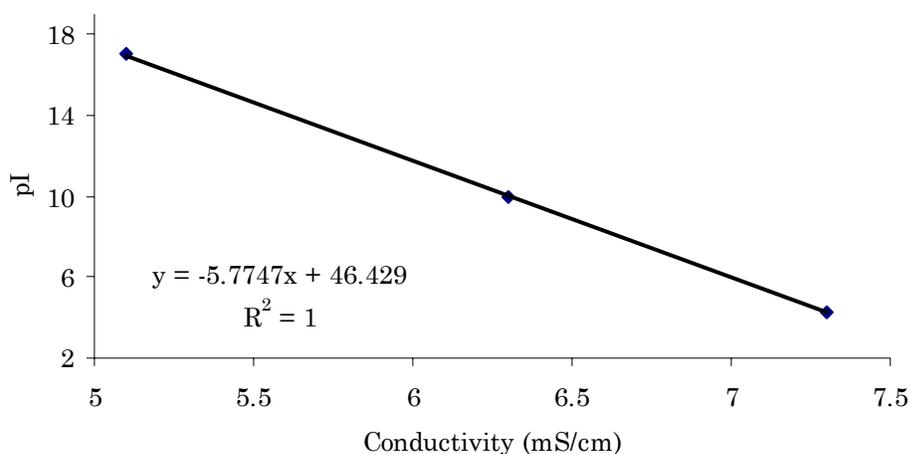


Figure A3: Calibration curve of Mono Q GL

### *Running and Cleaning Procedure*

The column is stored in 20% methanol, for running remove ethanol by passing 5 CV of water; then activate the matrix with 5 CV of elution buffer, and equilibrate with 5 CV of initial buffer (until eluent pH and conductivity equal the pH and conductivity of initial buffer).

0.5 to 2 mL of sample is injected and run at 0.5 to 1.5 mL/min. The sample is loaded and unbound proteins are washed out with initial buffer; when baseline returns to zero lineal or step gradients are started in order to elute the bound proteins. For cleaning of the column after elution of the

sample, flow rate is reversed and the column is cleaned with 2 mL of NaCl, water, NaOH, water, acetic acid, 10 CV of water. For storage of the column, pass 10 CV of methanol solution (eluent should smell like ethanol); if the column is going to be reused requilibrate as previously mentioned.

*EAH Sepharose 4B, Affinity Chromatography Solid Support*

Table A6: EAH Sepharose 4B matrix characteristics (From Amersham)

---

Active concentration	7-12 $\mu$ moles of amino group/mL drained matrix
Active Group	Amino group
Spacer	1,6-diaminohexane (10 atoms)
Coupling Capacity	Application dependent
Bead structure	4% Agarose
Mean bead size	90 $\mu$ m
Max linear flow rate	75 cm/h at 25°C, in 5 cm height column
pH stability	3-14
Chemical stability	Stable to all commonly use aqueous buffers

---

## APPENDIX B

### Summary of Purification of $\gamma$ GACT from Bovine Kidney

Sample	Measure	2	3	4	5	6
Homogenate	Units			896	2520	4460
	mg			675	13500	
	U/mg			1.3	0.2	
Supernatant	Units	3588	4716	*621	3888	2500
	mg			*553	3791	4110
	U/mg			1.1	1.0	0.6
Pellets	Units	278		0	1000	
	mg				2100	
	U/mg				0.5	
DEAE Sephacrose	Units			*315	2150	1650
	mg			747	857	364
	U/mg			0.4	2.5	4.5
Pellets 40%	Units			80	14	64
	mg			4	30	25
	U/mg			20	0.5	2.3
Supernatant 40%	Units	1499		1150	2958	1794
	mg			273	1000	320
	U/mg			4.2	3.0	5.6
Pellets 90%	Units	2224	1390	599	711	994
	mg			20	860	244
	U/mg			30	0.8	4.1
Supernatant 90%	Units	0	394	1900	0	0
	mg			114	0	89.6
	U/mg			0		2.4
S100 pool	Units			156	638	978
	mg			4	100	
	U/mg			39.0	6.4	

Sample	Measure	7	8	9	10	11
Homogenate	Units	4320	2600	4820	4956	5080
	mg	7020	6500	10060	13677	12007
	U/mg	0.6	0.4	0.5	0.4	0.4
Supernatant	Units	3030	1334	2783	3532	1533
	mg	2688	1306	4075	4476	3013
	U/mg	1.1	1.0	0.7	0.8	0.5
Pellets	Units	883	3910	439	480	985
	mg	1560	570	1070	2480	468
	U/mg	0.6	6.9	0.4	0.2	2.1
DEAE Sephrose	Units	1862	709	707	1936	1763
	mg	490	290	400	836	587
	U/mg	3.8	2.4	1.8	2.3	3
Pellets 40%	Units	29	27.5	7.1	61	
	mg	44	26	3.6	162	
	U/mg	0.7	1.1	2.0	0.4	
Supernatant 40%	Units	1733	844	546	1200	
	mg	432	307	175	600	
	U/mg	4.0	2.7	3.1	2.0	
Pellets 90%	Units	1550	693	398	1701	418
	mg	304	367	72	596	622
	U/mg	5.1	19	5.5	2.9	0.7
Supernatant 90%	Units	215	0	197	330	0
	mg	143	143	41	55	46
	U/mg	0	0	4.8	6.0	0
S100 pool	Units	823	420	400	785	
	mg	160		100	100.8	
	U/mg	5.1		4.0	8	

Sample	Measure	12	13	14	15	16
Homogenate	Units	5460	3552	2549	1537	3402
	mg	10478	9570	4294	7182	7800
	U/mg	0.5	0.4	0.6	0.2	0.4
Supernatant	Units	4057	1855	1626	1777	3354
	mg	3123	2367	3328	3692	2548
	U/mg	1.3	0.8	0.5	0.5	1.0
Pellets	Units	806	289	310	360	557
	mg	770	1470	852	1450	1850
	U/mg	1.0	0.2	3.6	0.2	0.3
DEAE Sephacrose	Units	1800	724	1073	1043	1756
	mg	792	500	346	620	784
	U/mg	2.3	1.4	3.1	1.7	2.0
Pellets 40%	Units					
	mg					
	U/mg					
Supernatant 40%	Units					
	mg					
	U/mg					
Pellets 90%	Units	1065	867	1732	1481	1354
	mg	563	500	372	551	480
	U/mg	0.9	1.7	4.7	2.7	2.8
Supernatant 90%	Units	0	0	0	0	151
	mg	428		0	84	32
	U/mg	0	0	0	0	4.7
S100 pool	Units	372	552	222	486	960
	mg					
	U/mg					

Sample	Measure	18	19	20	21	22
Homogenate	Units	8230	5635	6114	2460	11500
	mg	7016	5379	6000	6420	5130
	U/mg	1.2	1.0	1.0	0.4	2.2
Supernatant	Units	2726	4173	4180	1730	6125
	mg	2622	2843	1980	2129	1628
	U/mg	1.0	1.5	2.1	0.8	3.8
Pellets	Units	882	0	107	0	162
	mg	1059	1075	1000	0	1736
	U/mg	0.8	0.0	0.1	0	0.9
DEAE Sephrose	Units	2411	1500	1050	1045	1597
	mg	1240	714	370	437	400
	U/mg	1.9	2.1	2.8	2.4	4.0
Pellets 40%	Units					
	mg					
	U/mg					
Supernatant 40%	Units					
	mg					
	U/mg					
Pellets 90%	Units	2322	1395	1400	978	1400
	mg	1412	798	240	269	301
	U/mg	1.6	1.7	5.8	3.6	4.7
Supernatant 90%	Units	741	94	101	111	
	mg	49	9	306	16	
	U/mg	15.1	10.4	3.3	6.9	
S100 pool	Units	919	750	1212	842	
	mg					
	U/mg					

Sample	Measure	23	24	25	26	27
Homogenate	Units	4651	1975	14100	6603	583
	mg	5657		6097	10884	595
	U/mg	0.88		2.3	0.6	1.1
Supernatant	Units	3945	1514	9624	3303	3718
	mg	2118	1026	1026	2591	2123
	U/mg	1.9	1.5	3.7	1.1	1.8
Pellets	Units	248		1030	988	699
	mg	1324		2725	2186	448
	U/mg	0.2		0.4	0.5	1.6
DEAE Sephrose	Units	1441	686	1781	500	605
	mg	244	182	644	520	307
	U/mg	5.9	3.8	2.8	1.0	2.0
Pellet 40%	Units					
	mg					
	U/mg					
Supernatant 40%	Units					
	mg					
	U/mg					
Pellets 90%	Units	970	2300	11517	927	679
	mg	480	500	717	790	370
	U/mg	2.0	4.4	2.1	1.2	1.8
Supernatant 90%	Units	0				
	mg	47				
	U/mg	0				
S100 pool	Units	850		408	410	440
	mg	21		13	21	8
	U/mg	40.5		31.4	19.5	55

Sample	Measure	28	29	30	31	32
Homogenate	Units	4764		11700	5337	3374
	mg	8136		8410	8434	11148
	U/mg	0.3		1.4	0.6	0.3
Supernatant	Units	2237	2600	675	3212	3298
	mg	2240	2400	32541	3212	3912
	U/mg	1.0	1.1	2.1	1.0	0.8
Pellets	Units	308	817	1518	1281	552
	mg	240	550	1648	2796	913
	U/mg	1.3	1.5	0.9	0.5	0.6
DEAE Sephrose	Units		907		1130	1705
	mg		350		314	183
	U/mg		2.6		3.6	0.9
Pellets 40%	Units					
	mg					
	U/mg					
Supernatant 40%	Units					
	mg					
	U/mg					
Pellets 90%	Units	663		893	953	1008
	mg	416		348	395	266
	U/mg	1.6		2.6	2.4	3.8
Supernatant 90%	Units				147	97
	mg				20	0
	U/mg				7.4	
S100 pool	Units	1219	646	366	144	305
	mg	22	15	15	2	5.5
	U/mg	55.4	43.1	24	72	55

Sample	Measure	33	34	35	36	37
Homogenate	Units	6879	5727	3829	5185	2984
	mg	8249	8192	5987	6312	6175
	U/mg	0.8	0.7	0.6	0.8	0.5
Supernatant	Units	3476	3863	2838	3450	1780
	mg	2894	3141	2052	1906	1599
	U/mg	1.2	1.2	1.4	1.8	1.1
Pellets	Units	1813	1557	800	953	768
	mg	3763	3798	1616	1251	2040
	U/mg	0.5	0.4	0.5	0.8	0.4
DEAE Sephacrose	Units	2036	2793	1451	1576	684
	mg	649	175	119	488	333
	U/mg	3.1	1.6	1.2	3.2	2.1
Pellets 40%	Units					
	mg					
	U/mg					
Supernatant 40%	Units					
	mg					
	U/mg					
Pellets 90%	Units	1534	2445	1147	1379	546
	mg	669	1550	848	367	318
	U/mg	2.3	1.6	1.4	3.8	1.7
Supernatant 90%	Units	232	110	0	56	55
	mg	67	30	95	40	11
	U/mg	3.5	3.7	0	1.4	5
S100 pool	Units	706	834	466	370	161
	mg	12	26	8	9	3.4
	U/mg	58.8	32	58	41	47

Sample	Measure	38	39	40	41	42
Homogenate	Units	3397	4650	3565	4057	5805
	mg	5454	3680	5671	6150	3828
	U/mg	0.6	1.3	0.6	0.7	1.5
Supernatant	Units	2253	2470	2778	2492	4530
	mg	2075	1608	2083	2136	3014
	U/mg	1.1	1.5	1.3	1.2	1.5
Pellets	Units	753	768	441	511	825
	mg	1604	1367	1109	1941	20145
	U/mg	0.5	0.6	0.4	0.3	0.4
DEAE Sephacrose	Units	1208	1333	1275		421
	mg	319	422	391		490
	U/mg	3.8	3.2	3.3		9.2
Pellets 40%	Units					39
	mg					21
	U/mg					1.9
Supernatant 40%	Units					538
	mg					335
	U/mg					1.6
Pellets 90%	Units	891	403	1851	662	
	mg	574			299	
	U/mg	1.6			2.2	
Supernatant 90%	Units	119	298			
	mg	54				
	U/mg	2.2				
S100 pool	Units	279	719	641		
	mg		30	7.9		
	U/mg		21	81		

Sample	Measure	43	44	45	46
Homogenate	Units	5155	5554	3185	11673
	mg	8788	7757	6251	5565
	U/mg	0.6	0.7	0.5	2.1
Supernatant	Units	4088	4008	2008	5641
	mg	4699	2874	2675	2297
	U/mg	0.9	1.4	0.8	2.5
Pellets	Units	1791	1762	646	
	mg	2172	1729	1611	
	U/mg	0.8	1.0	0.4	
DEAE Sephacrose	Units	2995	2662	1205	1775
	mg	938	648	334	590
	U/mg	3.2	4.1	3.6	3.6
Pelet 40%	Units	95	106	96	
	mg	64	81	39	
	U/mg	1.5	1.3	2.5	
Supernatant 40%	Units	2388	3637	552	
	mg	775	494	275	
	U/mg	3.1	7.4	2	
Pellets 90%	Units	650	318	705	993
	mg	70	339	357	172
	U/mg	0.9	0.9	2	6
Supernatant 90%	Units	31	0	157	
	mg	19	4	16	
	U/mg	2.1	0	9.6	
S100 pool	Units	172	378	291	
	mg	5.7	15.7	5.7	
	U/mg	30	24	51	

APPENDIX C

Sequencing Report from Harvard Microchemistry  
and Proteomic Analysis Facility

## EXPLANATION OF MS/MS PEPTIDE SEQUENCE SUMMARY

The sequences summarized on the following pages are the result of analysis by microcapillary reverse-phase HPLC, directly coupled to the nano-electrospray ionization source of an ion trap and/or orbitrap mass spectrometer. These instruments are capable of acquiring individual sequence (MS/MS) spectra **on-line** at high sensitivity ( $<<<1$  femtomole) for multiple peptides in the chromatographic run. These MS/MS spectra are then correlated with known sequences using the algorithm Sequest developed at the Univ. of Washington (Eng *et al.*, 1994), and programs developed in our laboratory (Chittum *et al.*, 1998). MS/MS peptide sequences are then reviewed by a scientist for consensus with known proteins and the results manually confirmed for fidelity.



Application of this strategy and its technologies to the solution of biological problems can be found in the following publications or on our website at <http://mcb.harvard.edu/microchem/publications.html>:

1. Taniguchi, T, Garcia-Higuera, I, Xu, B, Andreassen, PR, Gregory, RC, Kim, ST, Lane, WS, Kastan, MB, D'Andrea, AD. Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. (2002) *Cell* **109**, 459-472.
2. Chittum, HS, Lane, WS, Carlson, BA, Roller, PP, Lung, FT, Lee, BJ, and Hatfield, DL. Rabbit  $\beta$ -Globin Is Extended Beyond Its UGA Stop Codon by Multiple Suppressions and Translational Reading Gaps. (1998) *Biochemistry* **37**, 10866-10870.
3. Eng, JK, McCormick, AL, and Yates, JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. (1994) *J. Am. Soc. Mass Spectrom.* **5**, 976-989.

Published results should state that this sequence analysis was performed at the Harvard Microchemistry and Proteomics Analysis Facility by microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry ( $\mu$ LC/MS/MS) on a Thermo LTQ-Orbitrap mass spectrometer.

Leu and Ile	This pair of amino acids are isobaric (identical mass), and thus cannot be unambiguously identified by mass analysis alone. In these cases, other information, such as the specificity of the enzyme used or a database comparison result, is used to determine the most probable assignment. It should be understood that this does not rule out the alternative isobaric amino acid.
-------------------	--

Header and footer: Sample name, user name, sample ID, file references, database and enzyme specificity used to support the interpretation.



Sequence	Reference	TIC	Sf	Scan
(R) FLDEFSCPDMYQR	gi 61831912	9.9e5	0.95	5005
(R) VFGEVYEVDER	gi 61831912	2.3e4	0.95	7080
(R) VFGEVYEVDER	gi 61831912	2.4e5	0.95	4583
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	3.3e4	0.95	5629
(R) FLDEFSCPDM*YQR	gi 61831912	8.6e4	0.95	4514
(R) FLDEFSCPDMYQR	gi 61831912	4.2e4	0.95	5655
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.4e4	0.95	5456
(R) VFGEVYEVDER	gi 61831912	2.9e4	0.95	6237
(R) VFGEVYEVDER	gi 61831912	8.9e3	0.94	8345
(R) VFGEVYEVDER	gi 61831912	2.2e4	0.94	7170
(R) VFGEVYEVDER	gi 61831912	3.6e4	0.94	6077
(K) WSTTETGAPCGTDDSSGR	gi 61831912	1.8e6	0.94	3531
(R) VFGEVYEVDER	gi 61831912	2.2e4	0.94	6822
(R) VFGEVYEVDER	gi 61831912	2.2e4	0.94	6659
(R) VFGEVYEVDER	gi 61831912	3.0e4	0.94	5996
(R) VFGEVYEVDER	gi 61831912	3.6e4	0.94	4327
(R) VFGEVYEVDER	gi 61831912	4.3e4	0.94	5662
(R) FLDEFSCPDM*YQR	gi 61831912	1.6e5	0.94	4874
(R) VFGEVYEVDER	gi 61831912	2.3e4	0.94	6990
(R) VFGEVYEVDER	gi 61831912	1.9e4	0.94	7709
(R) LAPVFFYGLKTGQPNHR	gi 61831912	4.7e4	0.94	4534
(R) VFGEVYEVDER	gi 61831912	3.9e4	0.94	5913
(R) VFGEVYEVDER	gi 61831912	2.0e4	0.94	7619
(R) VFGEVYEVDER	gi 61831912	1.4e4	0.94	8069
(R) VFGEVYEVDER	gi 61831912	4.4e3	0.94	9475
(R) VFGEVYEVDER	gi 61831912	1.2e4	0.94	8254
(R) VFGEVYEVDER	gi 61831912	5.1e7	0.94	4372
(R) VFGEVYEVDER	gi 61831912	2.3e4	0.94	6318
(R) VFGEVYEVDER	gi 61831912	2.3e6	0.94	4415
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.4e4	0.93	7225
(R) VFGEVYEVDER	gi 61831912	8.5e3	0.93	8437
(R) VFGEVYEVDER	gi 61831912	3.5e4	0.93	5411
(R) VFGEVYEVDER	gi 61831912	1.9e4	0.93	7529
(R) VFGEVYEVDER	gi 61831912	4.3e3	0.93	9192
(R) VFGEVYEVDER	gi 61831912	2.1e4	0.93	7260
(R) VFGEVYEVDER	gi 61831912	6.1e3	0.93	6579
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.4e6	0.93	4902

Sequence	Reference	TIC	Sf	Scan
(R) VFGEVYEVDER	gi 61831912	2.5e4	0.93	6736
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	9.7e3	0.93	7769
(R) VFGEVYEVDER	gi 61831912	1.8e4	0.92	7439
(R) VFGEVYEVDER	gi 61831912	4.3e4	0.92	5578
(R) VFGEVYEVDER	gi 61831912	2.0e4	0.92	7799
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.1e4	0.92	6874
(R) VFGEVYEVDER	gi 61831912	5.1e3	0.92	4865
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.2e4	0.92	5713
(R) VFGEVYEVDER	gi 61831912	4.7e3	0.92	9357
(R) VFGEVYEVDER	gi 61831912	6.8e3	0.91	8621
(R) VFGEVYEVDER	gi 61831912	4.4e3	0.91	9643
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	3.9e4	0.91	4662
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	8.8e3	0.91	8185
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	8.7e3	0.90	7137
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.2e4	0.90	7315
(R) VFGEVYEVDER	gi 61831912	9.0e3	0.90	5746
(R) VFGEVYEVDER	gi 61831912	3.6e3	0.90	9915
(R) VFGEVYEVDER	gi 61831912	3.7e3	0.90	9781
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.2e4	0.90	6018
(R) VFGEVYEVDER	gi 61831912	6.6e3	0.90	8524
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.9e4	0.89	5911
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	6.8e3	0.89	7960
(R) VFGEVYEVDER	gi 61831912	6.3e3	0.89	5323
(R) VFGEVYEVDER	gi 61831912	5.2e3	0.89	9025
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	8.5e3	0.89	7405
(R) TRLHVALEGVR	gi 61831912	2.5e6	0.88	3624
(R) VFGEVYEVDER	gi 61831912	5.0e3	0.88	8719
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	7.4e3	0.87	7872
(R) LAPVIFYGLK	gi 61831912	1.1e4	0.87	7286
(R) VFGEVYEVDER	gi 61831912	4.4e3	0.87	8827
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	6.0e3	0.86	8486
(R) VFGEVYEVDER	gi 61831912	4.9e3	0.86	4777
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.1e4	0.86	7584
(K) WSTTETGAPCGTDDSSGR	gi 61831912	1.2e5	0.85	3419
(R) LAPVIFYGLK	gi 61831912	9.1e6	0.84	5184
(R) LAPVIFYGLK	gi 61831912	7.4e3	0.84	7825

Sequence	Reference	TIC	Sf	Scan
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	7.6e3	0.84	8086
(R) LAPVYVYGLK	gi 61831912	3.9e6	0.84	5015
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.1e5	0.83	4722
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	5.2e3	0.83	8720
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.7e7	0.83	4843
(K) WSTTETGAPCGTDDSSGR	gi 61831912	1.6e5	0.83	3434
(R) LAPVYVYGLK	gi 61831912	1.0e4	0.82	7464
(R) LAPVYVYGLK	gi 61831912	4.3e4	0.82	5436
(R) LAPVYVYGLK	gi 61831912	5.6e7	0.82	5100
(R) VLLDGAHGR	gi 61831912	3.6e4	0.82	3194
(R) TRLHVALEGVR	gi 61831912	9.2e6	0.82	3658
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.1e4	0.81	5370
(R) LHVALEGVR	gi 61831912	1.0e7	0.80	3830
(R) LAPVYVYGLK	gi 61831912	1.0e4	0.80	7196
(R) LAPVYVYGLK	gi 61831912	8.9e3	0.80	7554
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.8e5	0.80	5016
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.2e8	0.80	4754
(K) WSTTETGAPCGTDDSSGR	gi 61831912	6.7e6	0.80	3506
(R) FLDEFESCPDMYQR	gi 61831912	1.2e5	0.79	5018
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	3.6e5	0.79	4668
(R) LAPVYVYGLK	gi 61831912	9.3e3	0.79	7016
(K) WSTTETGAPCGTDDSSGR	gi 61831912	3.4e5	0.78	3442
(R) LAPVYVYGLK	gi 61831912	6.5e3	0.77	8590
(K) WSTTETGAPCGTDDSSGR	gi 61831912	2.7e4	0.77	3891
(R) LAPVYVYGLK	gi 61831912	3.7e4	0.77	5352
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	8.3e3	0.77	7681
(R) VFGEVYVDER	gi 61831912	4.4e3	0.77	5494
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.1e4	0.76	7495
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	7.0e3	0.76	8309
(R) VLLDGAHGR	gi 61831912	2.2e4	0.75	3057
(R) LHVALEGVR	gi 61831912	1.6e7	0.75	3761
(R) LAPVYVYGLK	gi 61831912	8.2e3	0.74	7915
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.8e4	0.74	5816
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.8e3	0.73	8997
(R) LHVALEGVR	gi 61831912	2.0e6	0.73	3743

Sequence	Reference	TIC	Sf	Scan
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.1e7	0.73	4838
(R) LAPVFGYGLK	gi 61831912	6.7e3	0.72	8006
(R) LHVALEGVR	gi 61831912	5.9e5	0.72	3856
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	8.9e4	0.72	4805
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.7e6	0.72	4927
(R) LAPVFGYGLK	gi 61831912	5.3e3	0.71	8122
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	3.1e3	0.71	9532
(R) LAPVFGYGLK	gi 61831912	7.3e3	0.71	8400
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.2e4	0.70	7047
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	5.4e3	0.69	8585
(R) VLLDGAHGR	gi 61831912	1.4e5	0.69	3100
(R) LHVALEGVR	gi 61831912	4.2e4	0.69	4160
(R) LAPVFGYGLK	gi 61831912	1.1e4	0.69	7106
(K) WSTTETGAPCGTDDSSGR	gi 61831912	3.7e4	0.68	3521
(R) LAPVFGYGLK	gi 61831912	7.6e3	0.68	8221
(K) WSTTETGAPCGTDDSSGRLAPVFGYGLK	gi 61831912	1.1e5	0.68	5313
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.4e4	0.67	5185
(R) LAPVFGYGLK	gi 61831912	7.1e3	0.67	7734
(R) LHVALEGVR	gi 61831912	6.7e4	0.67	3944
(R) LAPVFGYGLK	gi 61831912	8.2e3	0.66	6929
(R) LAPVFGYGLK	gi 61831912	8.5e3	0.66	7644
(R) LAPVFGYGLK	gi 61831912	6.9e3	0.66	8769
(R) LAPVFGYGLK	gi 61831912	1.6e4	0.66	5856
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.4e4	0.65	8804
(R) LAPVFGYGLK	gi 61831912	6.2e3	0.65	9229
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.3e4	0.63	5449
(R) LAPVFGYGLK	gi 61831912	6.8e3	0.63	9788
(R) LAPVFGYGLK	gi 61831912	7.3e3	0.63	8679
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.6e4	0.61	8444
(R) VFGEVYEDER	gi 61831912	1.3e5	0.61	4290
(K) WSTTETGAPCGTDDSSGR	gi 61831912	3.8e5	0.59	3615
(R) LAPVFGYGLK	gi 61831912	6.1e3	0.57	9141
(R) FLDEFESCPDMYQR	gi 61831912	3.4e4	0.57	5888
(R) LAPVFGYGLK	gi 61831912	1.1e4	0.57	5608
(R) LAPVFGYGLK	gi 61831912	5.2e3	0.56	9879
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	4.3e3	0.56	9710
(R) LAPVFGYGLK	gi 61831912	6.4e3	0.55	9049

Sequence	Reference	TIC	Sf	Scan
(R) LAPVYVYGLTK	gi 61831912	1.3e4	0.54	5946
(R) LAPVYVYGLTK	gi 61831912	5.9e3	0.53	9596
(R) LAPVYVYGLTK	gi 61831912	7.3e3	0.53	8960
(R) LAPVYVYGLTK	gi 61831912	9.0e3	0.53	6839
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.1e4	0.53	8626
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.3e4	0.53	8624
(K) WSTTETGAPCGTDDSSGR	gi 61831912	4.1e4	0.52	3534
(R) LAPVYVYGLTK	gi 61831912	1.1e5	0.51	5189
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	3.2e5	0.51	4854
(R) LAPVYVYGLTK	gi 61831912	1.8e6	0.50	5020
(R) LAPVYVYGLTK	gi 61831912	7.0e3	0.50	8491
(R) LAPVYVYGLTK	gi 61831912	5.2e3	0.50	9321
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.1e4	0.47	5982
(R) LAPVYVYGLTK	gi 61831912	5.0e3	0.46	9411
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.7e4	0.46	5533
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.8e4	0.42	8714
(R) LAPVYVYGLTK	gi 61831912	6.5e3	0.41	8312
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	1.4e4	0.41	5186
(R) LAPVYVYGLTK	gi 61831912	5.3e3	0.41	9701
(R) LAPVYVYGLTK	gi 61831912	1.4e4	0.38	6023
(R) LHVALEGVR	gi 61831912	8.5e3	0.37	3767
(R) LAPVYVYGLTK	gi 61831912	2.2e6	0.36	5105
(R) LAPVYVYGLTK	gi 61831912	7.6e3	0.35	5686
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.7e4	0.34	5697
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	2.5e4	0.32	5790
(K) TGQPNHR	gi 61831912	7.2e3	0.32	339
(K) TGQPNHR	gi 61831912	1.3e4	0.29	437
(R) FLDEFESCPDM*YQR	gi 61831912	9.5e3	0.27	4432
(R) VFGEVYEVDERM*LR	gi 61831912	4.1e4	0.25	4524
(R) AHTLEPYPLVIAGEHNIPR	gi 61831912	6.1e3	0.24	9524
(R) LHVALEGVR	gi 61831912	7.8e3	0.21	3860

**B** gi|54697172|gb|AAV38958.1| **MS/MS Spectra:** 33 **Sum TIC:** 2.7e7 **Avg TIC:** 8.0e5 **Cov:** 48.8% **Uniq:** 9  
 RAB11A, member RAS oncogene family [synthetic construct] gi|54697166|gb|AAV38955.1| RAB11A, member RAS oncogene family [synthetic  
 construct] gi|30584069|gb|AAP36283.1| Homo sapiens RAB11A, member RAS oncogene family [synthetic construct]  
 gi|60653915|gb|AAX29650.1| RAB11A member RAS oncogene family [synthetic construct] gi|61365507|gb|AAX42719.1| RAB11A member RAS  
 oncogene family [synthetic const

Sequence	Reference	TIC	Sf	Scan
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	5.3e4	0.98	8084
(R) DHADSNIVIMLVGNK	gi 54697172 +31	1.2e5	0.97	5373
(R) GAVGALLVYDIK	gi 54697172 +60	9.8e6	0.96	5518
(R) DHADSNIVIMLVGNK	gi 54697172 +31	5.2e4	0.96	5453
(R) DHADSNIVIM*LVGNK	gi 54697172 +31	1.8e5	0.96	4819
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	1.4e4	0.96	7994
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	1.3e4	0.96	8004
(R) DHADSNIVIMLVGNK	gi 54697172 +31	1.1e5	0.93	5444
(K) ELRDHADSNIIMLVGNK	gi 54697172 +31	3.5e5	0.93	5214
(K) STIGVEFATR	gi 54697172 +99	10.0e6	0.93	4241
(R) DHADSNIVIMLVGNK	gi 54697172 +31	3.7e4	0.93	5364
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	2.4e4	0.92	8464
(R) DHADSNIVIMLVGNK	gi 54697172 +31	2.7e4	0.92	5536
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	9.2e3	0.92	8094
(K) ELRDHADSNIIM*LVGNK	gi 54697172 +31	6.8e5	0.92	4685
(R) GAVGALLVYDIK	gi 54697172 +60	4.6e4	0.91	5800
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	5.7e4	0.90	8489
(K) STIGVEFATR	gi 54697172 +99	9.0e4	0.88	4194
(R) GAVGALLVYDIK	gi 54697172 +60	2.0e4	0.85	5490
(R) GAVGALLVYDIK	gi 54697172 +60	1.1e5	0.78	5634
(K) AQIWDTAGQER	gi 54697172 +99	3.6e4	0.76	3709
(K) HLTyenVER	gi 54697172 +52	4.3e6	0.76	3285
(R) AITSAYYR	gi 54697172 +99	6.0e4	0.75	3472
(R) GAVGALLVYDIK	gi 54697172 +60	2.1e4	0.71	5945
(K) ELRDHADSNIIMLVGNK	gi 54697172 +31	1.0e4	0.70	5132
(K) AQIWDTAGQER	gi 54697172 +99	4.8e4	0.65	3938
(K) NGLSFIETSAIDSTNVEAAFQITLTIYR	gi 54697172 +9	1.4e4	0.58	8554
(R) NEFNLESK	gi 54697172 +70	1.2e5	0.57	3648
(K) AQIWDTAGQER	gi 54697172 +99	6.4e4	0.57	3918
(R) GAVGALLVYDIK	gi 54697172 +60	1.3e4	0.55	5488
(K) ELRDHADSNIIMLVGNK	gi 54697172 +31	2.1e4	0.51	5227
(K) AQIWDTAGQER	gi 54697172 +99	1.9e4	0.27	3923
(K) HLTyenVER	gi 54697172 +52	3.5e4	0.26	3287

**C** gi|30794280|ref|NP\_851335.1| **MS/MS Spectra:** 22 **Sum TIC:** 1.6e6 **Avg TIC:** 7.3e4 **Cov:** 38.2% **Uniq:** 20  
 albumin [Bos taurus] gi|23307791|gb|AAN17824.1| serum albumin [Bos taurus] gi|3336842|emb|CAA76847.1| bovine serum albumin [Bos  
 taurus] gi|2190337|emb|CAA41735.1| serum albumin [Bos taurus]

Sequence	Reference	TIC	Sf	Scan
(R) RHPEYAVSVLLR	gi 30794280 +5	7.5e5	0.97	4286
(K) DAFLGSFLYEYSR	gi 30794280 +4	6.3e4	0.97	6131
(R) FKDLGEEHFK	gi 30794280 +5	1.6e4	0.95	3489
(R) MPCTEDYLSLILNR	gi 30794280 +4	3.3e4	0.95	6140
(R) M*PCTEDYLSLILNR	gi 30794280 +4	2.7e4	0.95	5956
(K) TVMENFVAFVDK	gi 30794280 +4	3.8e4	0.92	5831
(K) EYEATLEECCAK	gi 30794280 +4	2.4e4	0.91	3602
(K) YICDNQDTISSK	gi 30794280 +3	4.7e4	0.90	3327
(K) CCTESLVNR	gi 30794280 +37	3.1e4	0.89	3274
(K) LVNELTEFAK	gi 30794280 +3	8.4e4	0.86	4677
(K) SLHTLFGDELCK	gi 30794280 +4	4.5e4	0.84	4697
(K) HLVDPEQNLIK	gi 30794280 +5	8.5e4	0.79	4016
(K) KQTALVELLK	gi 30794280 +11	2.2e4	0.78	4472
(R) KVPQVSTPTLVEVSR	gi 30794280 +34	7.1e4	0.77	4246
(R) RPCFSALTPDETYVPK	gi 30794280 +4	6.5e4	0.74	4523
(R) HPYFYAPELLEYANK	gi 30794280 +3	2.5e4	0.72	5554
(K) GLVLIAFSQYLQQCPFDEHVK	gi 30794280 +3	1.5e4	0.72	6422
(R) RPCFSALTPDETYVPK	gi 30794280 +4	4.4e4	0.66	4544
(K) YLYEYAR	gi 30794280 +30	6.0e4	0.65	4122
(K) LKECCDKPILLEK	gi 30794280 +7	3.3e4	0.61	3256
(K) AEFVEVTK	gi 30794280 +4	1.9e4	0.60	3593
(R) RHPYFYAPELLEYANK	gi 30794280 +3	9.0e3	0.50	5245

**D** gi|1065362|pdb|1HUR|B **MS/MS Spectra: 18** **Sum TIC: 9.5e5** **Avg TIC: 5.3e4** **Cov: 75.0%** **Uniq: 11**  
 Chain B, Human Adp-Ribosylation Factor 1 Complexed With Gdp, Full Length Non-Myristoylated gi|1065361|pdb|1HUR|A Chain A, Human Adp-Ribosylation Factor 1 Complexed With Gdp, Full Length Non-Myristoylated

Sequence	Reference	TIC	Sf	Scan
(R) MLAEDELRDVLLVFANK	gi 1065362 +27	3.5e4	0.97	6926
(R) MLAEDELRDVLLVFANK	gi 1065362 +27	3.9e4	0.96	6846
(R) NWYIQATCATSGDGLYEGLDWLSNQLR	gi 1065362 +9	1.6e4	0.96	6888
(R) MLAEDELRDVLLVFANK	gi 1065362 +27	2.5e4	0.96	6861
(K) QDLPNAMAAEITDK	gi 1065362 +83	2.3e4	0.95	4426
(K) NISFTVVDVGGQDK	gi 1065362 +99	5.2e4	0.95	5415
(R) M*LAEDELRDVLLVFANK	gi 1065362 +27	6.0e4	0.93	6627
(K) LGEIVTTIPTIGFNVETVEYK	gi 1065362 +99	5.4e4	0.93	6076
(R) ILMVGLDAAGK	gi 1065362 +99	1.2e5	0.90	4654
(K) LKLGEIVTTIPTIGFNVETVEYK	gi 1065362 +99	6.1e4	0.89	6120
(R) M*LAEDELRDVLLVFANK	gi 1065362 +27	9.8e4	0.89	6564
(R) M*LAEDELR	gi 1065362 +42	1.5e5	0.79	3469
(K) LGEIVTTIPTIGFNVETVEYK	gi 1065362 +99	3.4e4	0.79	6006
(K) QDLPNAMAAEITDKLGLSLR	gi 1065362 +74	2.8e4	0.67	5430
(K) QDLPNAM*AAEITDKLGLSLR	gi 1065362 +74	5.5e4	0.58	5162
(R) HYFQNTQGLIFVVDSDRDR	gi 1065362 +72	4.2e4	0.56	4919
(K) LGLHSLR	gi 1065362 +94	4.7e4	0.54	3463
(K) LGEIVTTIPTIGFNVETVEYK	gi 1065362 +99	1.4e4	0.30	6088

**E** gi|76645762|ref|XP\_587427.2| **MS/MS Spectra: 14** **Sum TIC: 5.9e5** **Avg TIC: 4.2e4** **Cov: 59.5%** **Uniq: 9**  
 PREDICTED: similar to cytoglobin [Bos taurus]

Sequence	Reference	TIC	Sf	Scan
(R) LYANCEDVGVAILVR	gi 76645762 +7	7.7e4	0.98	5226
(K) ILSGVILEVIAEEFANDFPETQR	gi 76645762 +1	1.9e4	0.97	7704
(R) VM*GALNTVVENLHDPEKVVSLSLVGK	gi 76645762	1.5e5	0.89	5977
(R) VM*GALNTVVENLHDPEKVVSLSLVGK	gi 76645762 +1	2.7e4	0.87	6057
(K) ILSGVILEVIAEEFANDFPETQR	gi 76645762 +1	7.6e3	0.86	7617
(R) SEELSEAER	gi 76645762 +8	1.5e4	0.85	3040
(R) ERSEELSEAER	gi 76645762 +8	2.1e4	0.77	2976
(K) HKVEPVYFK	gi 76645762 +8	2.7e4	0.76	3339
(R) VM*GALNTVVENLHDPEKVVSLSLVGK	gi 76645762	3.1e4	0.76	5983
(K) ILSGVILEVIAEEFANDFPETQR	gi 76645762 +1	1.2e4	0.73	7636
(R) FVFNPSAK	gi 76645762 +15	9.5e4	0.70	5011
(K) HM*EEPLEM*ER	gi 76645762	2.0e4	0.58	3037
(R) VMGALNTVVENLHDPEKVVSLSLVGK	gi 76645762	7.6e4	0.50	6149
(K) AVQATWAR	gi 76645762 +4	1.2e4	0.43	3455

F		gi 6188874 ref NP_001013617.1	MS/MS Spectra: 10	Sum TIC: 5.1e5	Avg TIC: 5.1e4	Cov: 53.8%	Uniq: 7
		transgelin 2 [Bos taurus] gi 73919851 sp Q5E9F5 TAGL2_BOVIN Transgelin-2 gi 59858295 gb AAX08982.1  transgelin 2 [Bos taurus]					
Sequence	Reference	TIC	Sf	Scan			
(R) YGINTTDIFQTVDLWEGK	gi 6188874 +16	1.3e5	0.98	6444			
(K) QYDADLEQILIQWITTQCR	gi 6188874 +15	1.6e4	0.97	7162			
(K) QMEIQSFLQAAER	gi 6188874 +15	4.9e4	0.96	5590			
(R) DDGLFSGDPNWFPK	gi 6188874 +15	6.2e4	0.94	5873			
(K) QM*EQISQFLQAAER	gi 6188874 +15	7.6e4	0.91	5165			
(K) QMEIQSFLQAAER	gi 6188874 +15	1.1e4	0.84	5583			
(K) DGTVLCELINGLYPEGQAPVKK	gi 6188874 +2	5.0e4	0.82	6193			
(R) GPAYGLSR	gi 6188874 +11	5.3e4	0.80	3314			
(K) DGTVLCELINGLYPEGQAPVKK	gi 6188874 +2	6.0e4	0.60	5480			
(R) TLM*NLGGLAVAR	gi 6188874 +15	3.6e3	0.42	4481			
G		gi 77736431 ref NP_001029915.1	MS/MS Spectra: 13	Sum TIC: 4.6e5	Avg TIC: 3.5e4	Cov: 39.8%	Uniq: 5
		hypothetical protein LOC613749 [Bos taurus] gi 61553191 gb AAX46365.1  RAB5C, member RAS oncogene family isoform b [Bos taurus]					
Sequence	Reference	TIC	Sf	Scan			
(R) AVEFQEAQAYAEEDNSLLFM*ETSAK	gi 77736431	3.5e4	0.98	5562			
(R) GAQAATVYDITNTDFFAR	gi 77736431 +12	1.8e4	0.98	5559			
(R) AVEFQEAQAYAEEDNSLLFMETSAK	gi 77736431	1.4e4	0.96	5940			
(K) TAMNVN EIFMAIAK	gi 77736431 +18	6.6e4	0.95	6159			
(R) GAQAATVYDITNTDFFAR	gi 77736431 +12	2.1e4	0.94	5560			
(R) AVEFQEAQAYAEEDNSLLFMETSAK	gi 77736431	1.8e4	0.94	5914			
(K) TAMNVN EIFMAIAK	gi 77736431 +18	6.0e4	0.91	6151			
(K) GQFHEYQESTIGAFLTQTVCLDDTTVK	gi 77736431 +17	2.7e4	0.91	5716			
(K) TAM*NVN EIFM*AIK	gi 77736431 +18	3.4e4	0.90	5426			
(K) RAVEFQEAQAYAEEDNSLLFMETSAK	gi 77736431	3.7e4	0.90	5640			
(R) AVEFQEAQAYAEEDNSLLFMETSAK	gi 77736431	1.5e4	0.90	5934			
(K) TAM*NVN EIFMAIAK	gi 77736431 +18	5.8e4	0.73	5657			
(K) GQFHEYQESTIGAFLTQTVCLDDTTVK	gi 77736431 +17	5.4e4	0.39	5737			

<b>H</b>		<b>MS/MS Spectra: 7</b>		<b>Sum TIC: 1.7e5</b>		<b>Avg TIC: 2.4e4</b>		<b>Cov: 23.2%</b>		<b>Uniq: 4</b>	
<p>gi 57085091 ref XP_536379.1  <b>PREDICTED: similar to SAR1a gene homolog 1 isoform 2 [Canis familiaris]</b> gi 72535188 ref NP_001026956.1  <b>GTP-binding protein SAR1a [Sus scrofa]</b> gi 74354052 gb AAI02444.1  <b>Hypothetical protein LOC517171 [Bos taurus]</b> gi 62868640 gb AAI17508.1  <b>GTP-binding protein SAR1a [Sus scrofa]</b> gi 77735989 ref NP_001029693.1  <b>hypothetical protein LOC517171 [Bos taurus]</b></p>											
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>
(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11	2.6e4	0.96	6257	(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11	2.4e4	0.91	6568	(K) VELNALMTDETISNPILILGNK	gi 57085091 +11
(K) VELNALMTDETISNPILILGNK	gi 57085091 +11	2.4e4	0.91	6568	(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11	5.2e4	0.88	6243	(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11
(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11	5.2e4	0.88	6243	(K) VELNALM*TDETISNPILILGNKIDR	gi 57085091 +11	3.1e4	0.85	6229	(K) IDRTDAISEEK	gi 57085091 +10
(K) VELNALM*TDETISNPILILGNKIDR	gi 57085091 +11	3.1e4	0.85	6229	(K) VELNALM*TDETISNPILILGNK	gi 57085091 +11	9.9e3	0.65	3045	(K) VELNALMTDETISNPILILGNK	gi 57085091 +11
(K) IDRTDAISEEK	gi 57085091 +10	9.9e3	0.65	3045	(R) EIFGLYGGTTGK	gi 57085091 +12	4.7e3	0.40	6553	(R) EIFGLYGGTTGK	gi 57085091 +12
(K) VELNALMTDETISNPILILGNK	gi 57085091 +11	4.7e3	0.40	6553			2.1e4	0.35	4679		
(R) EIFGLYGGTTGK	gi 57085091 +12	2.1e4	0.35	4679							
<b>I</b>		<b>MS/MS Spectra: 6</b>		<b>Sum TIC: 2.5e5</b>		<b>Avg TIC: 4.1e4</b>		<b>Cov: 10.7%</b>		<b>Uniq: 6</b>	
<p>gi 17318569 ref NP_006112.2  <b>keratin 1 [Homo sapiens]</b> gi 11935049 gb AAG41947.1  <b>keratin 1 [Homo sapiens]</b> gi 39794653 gb AAH63697.1  <b>KRT1 protein [Homo sapiens]</b></p>											
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>
(R) FLEQQNQVLQTK	gi 17318569 +25	4.6e4	0.97	3942	(R) THNLEPYFESFINLR	gi 17318569 +3	5.1e4	0.95	6135	(R) SLDLDSIIAEVK	gi 17318569 +48
(R) THNLEPYFESFINLR	gi 17318569 +3	5.1e4	0.95	6135	(K) YEELQITAGR	gi 17318569 +8	2.0e4	0.95	5960	(K) LALDLEIATYR	gi 17318569 +3
(R) SLDLDSIIAEVK	gi 17318569 +48	2.0e4	0.95	5960	(K) LALDLEIATYR	gi 17318569 +3	6.8e4	0.94	4013	(K) IEISELNR	gi 17318569 +6
(K) YEELQITAGR	gi 17318569 +8	6.8e4	0.94	4013			2.7e4	0.91	5417		
(K) LALDLEIATYR	gi 17318569 +3	2.7e4	0.91	5417			3.7e4	0.75	3962		
(K) IEISELNR	gi 17318569 +6	3.7e4	0.75	3962							
<b>J</b>		<b>MS/MS Spectra: 6</b>		<b>Sum TIC: 2.4e5</b>		<b>Avg TIC: 4.0e4</b>		<b>Cov: 18.7%</b>		<b>Uniq: 4</b>	
<p>gi 74267735 gb AAI02447.1  <b>Galactose mutarotase (aldose 1-epimerase) [Bos taurus]</b> gi 59857745 gb AAX08707.1  <b>galactose mutarotase (aldose 1-epimerase) [Bos taurus]</b> gi 84028319 sp Q5EA79 GALM_BOVIN <b>Aldose 1-epimerase (Galactose mutarotase)</b> gi 77736588 ref NP_001029967.1  <b>galactose mutarotase (aldose 1-epimerase) [Bos taurus]</b></p>											
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	<b>Sequence</b>	<b>Reference</b>
(R) ASDVVLGFDELEGYLQK	gi 74267735	6.5e4	0.98	6261	(R) ASDVVLGFDELEGYLQK	gi 74267735	3.6e4	0.97	6258	(R) VSPDGEEGYPGELK	gi 74267735 +6
(R) ASDVVLGFDELEGYLQK	gi 74267735	3.6e4	0.97	6258	(R) VLEVYTTQPGVQFYTGNFNLDGTLK	gi 74267735 +3	6.3e4	0.83	3782	(K) FQLQSDQLR	gi 74267735 +1
(R) VSPDGEEGYPGELK	gi 74267735 +6	6.3e4	0.83	3782	(R) VLEVYTTQPGVQFYTGNFNLDGTLK	gi 74267735 +3	1.1e4	0.74	6116	(R) VLEVYTTQPGVQFYTGNFNLDGTLK	gi 74267735 +3
(R) VLEVYTTQPGVQFYTGNFNLDGTLK	gi 74267735 +3	1.1e4	0.74	6116			4.3e4	0.63	4067		
(K) FQLQSDQLR	gi 74267735 +1	4.3e4	0.63	4067			2.0e4	0.31	6092		
(R) VLEVYTTQPGVQFYTGNFNLDGTLK	gi 74267735 +3	2.0e4	0.31	6092							

K		gi 76616433 ref XP_587041.2	MS/MS Spectra: 7	Sum TIC: 2.2e5	Avg TIC: 3.2e4	Cov: 40.0%	Uniq: 4
PREDICTED: similar to retinol binding protein 5, cellular [Bos taurus]							
Sequence	Reference	TIC	Sf	Scan			
(R) LWLEEEMLYQEVTR	gi 76616433	2.8e4	0.97	6142			
(R) LWLEEEEM*LYQEVTR	gi 76616433	2.0e4	0.96	5820			
(K) NLEDYLQALNVMALR	gi 76616433 +1	1.6e4	0.95	6383			
(K) NLEDYLQALNVM*ALR	gi 76616433 +1	5.5e4	0.94	6051			
(R) LWLEEEEM*LYQEVTR	gi 76616433	1.9e4	0.75	5827			
(R) DAVCQCVR	gi 76616433	3.7e4	0.68	3912			
(K) IALLKPKDEIDQR	gi 76616433	5.0e4	0.40	4266			
L		gi 1050551 emb CAA61797.1	MS/MS Spectra: 5	Sum TIC: 2.0e5	Avg TIC: 4.0e4	Cov: 29.5%	Uniq: 4
rab7 [Mus musculus] gi 6679599 ref NP_033031.1  RAB7, member RAS oncogene family [Mus musculus]							
Sequence	Reference	TIC	Sf	Scan			
(R) GADCCVLVFDVTAPNTRK	gi 1050551 +17	4.7e4	0.96	5637			
(K) EAINVEQAFQTIAR	gi 1050551 +22	4.4e4	0.96	5580			
(R) DPENFPFVVLGNK	gi 1050551 +21	4.5e4	0.93	5750			
(R) DPENFPFVVLGNK	gi 1050551 +21	4.3e4	0.88	5667			
(K) TLDNRWDEFLLIQASPR	gi 1050551 +19	2.1e4	0.83	5353			
M		gi 76640732 ref XP_871348.1	MS/MS Spectra: 5	Sum TIC: 2.0e5	Avg TIC: 4.0e4	Cov: 13.0%	Uniq: 3
PREDICTED: similar to core-binding factor, beta subunit isoform 1 [Bos taurus]							
Sequence	Reference	TIC	Sf	Scan			
(R) AQQEDALAQQAFFEAR	gi 76640732 +12	4.1e4	0.98	4576			
(R) AQQEDALAQQAFFEAR	gi 76640732 +12	2.4e4	0.92	4574			
(R) SKFENEEFFR	gi 76640732 +19	9.5e4	0.76	4256			
(R) SKFENEEFFR	gi 76640732 +19	3.2e4	0.52	4226			
(R) SEIAFVATGNTLSLQFFPASWQGEQR	gi 76640732 +15	5.6e3	0.43	6439			
N		gi 61872816 ref XP_595378.1	MS/MS Spectra: 5	Sum TIC: 1.9e5	Avg TIC: 3.7e4	Cov: 34.9%	Uniq: 4
PREDICTED: similar to Mak3 homolog, partial [Bos taurus]							
Sequence	Reference	TIC	Sf	Scan			
(K) FGFEIETK	gi 61872816 +15	1.0e4	0.90	5171			
(R) IELGDVTPHNIK	gi 61872816 +16	4.7e4	0.76	4166			
(R) LNQVIFPVSYNDKFKYKDVLEVGLAK	gi 61872816 +15	3.6e4	0.39	6152			
(K) MLNHLNICEK	gi 61872816 +14	2.0e4	0.39	4175			
(R) IELGDVTPHNIK	gi 61872816 +16	7.2e4	0.26	4159			

<b>O</b>	gi 83638644 gb AAI09566.1  Similar to ADP-ribosylation factor-like 3 [Bos taurus] gi 84370167 ref NP_001033656.1  hypothetical protein LOC540040 [Bos taurus]	<b>MS/MS Spectra:</b> 5 Sum TIC: 1.7e5 Avg TIC: 3.4e4 Cov: 22.0%	<b>Uniq:</b> 2	
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(R) NYFENTDILYVIDSADR	gi 83638644 +6	3.6e4	0.97	6758
(R) NYFENTDILYVIDSADR	gi 83638644 +6	3.8e4	0.95	6762
(K) QDLLTAAPASEIAEGLNLHTIR	gi 83638644 +13	5.7e4	0.94	5723
(K) QDLLTAAPASEIAEGLNLHTIR	gi 83638644 +13	2.9e4	0.90	5805
(K) QDLLTAAPASEIAEGLNLHTIR	gi 83638644 +13	1.2e4	0.86	5747
<b>P</b>	gi 76610205 ref XP_872672.1  PREDICTED: similar to peptidylprolyl isomerase-like protein 3 isoform PPIL3b, partial [Bos taurus]	<b>MS/MS Spectra:</b> 4 Sum TIC: 1.7e5 Avg TIC: 4.2e4 Cov: 32.6%	<b>Uniq:</b> 3	
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(K) GFM*VQTGDPGTGTR	gi 76610205 +8	5.1e4	0.95	3654
(K) GFMVQTGDPGTGTR	gi 76610205 +8	5.1e4	0.93	3964
(K) VIDGLETLDELEKLPVNEK	gi 76610205 +11	2.7e4	0.91	5763
(K) TYRPLNDVHIK	gi 76610205 +8	4.1e4	0.90	3653
<b>Q</b>	gi 74354774 gb AAI02270.1  Biliverdin reductase B (flavin reductase (NADPH)) [Bos taurus] gi 27806297 ref NP_776676.1  biliverdin reductase B (flavin reductase (NADPH)) [Bos taurus] gi 516594 gb AAC37323.1  flavin reductase gi 1706869 sp P52556 BLVRB_BOVIN Flavin reductase (FR) (NADPH-dependent diaphorase) (NADPH-flavin reductase) (FLR) (Biliverdin reductase B) (BYR-B) (Biliverdin-IX beta-reductase) (Green heme-binding pro	<b>MS/MS Spectra:</b> 3 Sum TIC: 1.6e5 Avg TIC: 5.3e4 Cov: 24.8%	<b>Uniq:</b> 3	
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(K) TVAGQDAVIVLLGTR	gi 74354774 +4	4.5e4	0.95	5607
(R) LPSEGPQAHVWVVDVR	gi 74354774 +1	7.3e4	0.87	3920
(K) VVACTSAFLLWDPKVPVPR	gi 74354774 +2	4.2e4	0.37	5598
<b>R</b>	gi 48425163 pdb 1OYH I Chain I, Crystal Structure Of P13 Alanine Variant Of Antithrombin	<b>MS/MS Spectra:</b> 6 Sum TIC: 1.6e5 Avg TIC: 2.6e4 Cov: 22.2%	<b>Uniq:</b> 6	
<b>Sequence</b>	<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(R) EVPLNTIIFMGR	gi 48425163 +14	2.6e4	0.71	5773
(-) HGSPVDICTAKPR	gi 48425163 +5	1.5e4	0.61	3215
(K) KATEDEGSEQKIPEATNR	gi 48425163 +12	8.4e3	0.57	3047
(R) ITDVPSEAINELTVLVLVNTIYFK	gi 48425163 +9	7.7e3	0.56	8276
(K) ATEDEGSEQKIPEATNR	gi 48425163 +12	4.4e4	0.40	3177
(R) VAEGTQVLELFPKGGDITMVLILPKPEK	gi 48425163 +15	5.5e4	0.30	6212

<b>S</b>	gi 1633322 pdb 1VHR B Chain B, Human Vh1-Related Dual-Specificity Phosphatase	<b>MS/MS Spectra:</b> 2	<b>Sum TIC:</b> 1.4e5	<b>Avg TIC:</b> 7.2e4	<b>Cov:</b> 14.7%	<b>Uniq:</b> 2
			gi 1633321 pdb 1VHR A Chain A, Human Vh1-Related Dual-Specificity Phosphatase			
	<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(R)	IYGNASVAQDIPK		gi 1633322 +5	4.2e4	0.95	4443
(K)	LGITHVLNAEGR		gi 1633322 +10	1.0e5	0.82	4282
<b>T</b>	gi 86823974 gb AAI05348.1  Unknown (protein for MGC:127550) [Bos taurus]	<b>MS/MS Spectra:</b> 4	<b>Sum TIC:</b> 1.4e5	<b>Avg TIC:</b> 3.6e4	<b>Cov:</b> 16.5%	<b>Uniq:</b> 4
	<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(K)	FGYQFTQQGPR		gi 86823974 +14	5.3e4	0.94	4252
(R)	SNDPVATAFADMLR		gi 86823974	2.8e4	0.93	6086
(R)	QQLGTAVELEMANMLEENTNILK		gi 86823974 +1	1.5e4	0.88	6669
(K)	YFSLAATR		gi 86823974 +6	4.7e4	0.71	4214
<b>U</b>	gi 83638578 gb AAI09826.1  Similar to Guanidinoacetate N-methyltransferase [Bos taurus]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 1.3e5	<b>Avg TIC:</b> 4.3e4	<b>Cov:</b> 18.2%	<b>Uniq:</b> 2
			gi 84370113 ref NP_001033633.1  hypothetical protein LOC515270 [Bos taurus]			
	<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(K)	YSDITTMFEETQVPALLEAGFR		gi 83638578 +1	1.9e4	0.97	6796
(K)	YSDITTM*FEETQVPALLEAGFR		gi 83638578 +1	2.5e4	0.92	6313
(K)	VQEAPIEEHWHIECNEGVFQR		gi 83638578	8.4e4	0.89	5339
<b>V</b>	gi 13097708 gb AAH03560.1  Ribophorin II, precursor [Homo sapiens]	<b>MS/MS Spectra:</b> 5	<b>Sum TIC:</b> 1.1e5	<b>Avg TIC:</b> 2.3e4	<b>Cov:</b> 17.0%	<b>Uniq:</b> 5
			gi 35493916 ref NP_002942.2  ribophorin II precursor [Homo sapiens]			gi 4730801 emb CAB41763.1  RPN2 [Homo sapiens]
			gi 18044228 gb AAH20222.1  Ribophorin II, precursor [Homo sapiens]			gi 5834424 emb CAB54801.1  ribophorin II [Homo sapiens]
			gi 9297108 sp P04844 RIB2_HUMAN Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit precursor [Homo sapiens]			
	<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>
(K)	FSSGYDFLVEVEGDNR		gi 13097708 +5	2.2e4	0.96	5903
(K)	TSFTPVGDVFEINFM*NVK		gi 13097708 +10	1.9e4	0.91	6083
(K)	TGQEVVFAEPDNK		gi 13097708 +15	3.1e4	0.90	4170
(K)	ASLDRPFTNLESAFYISIVGLSSLGAQVPDAK		gi 13097708 +8	9.0e3	0.59	7158
(R)	LSKETVLTATVQALQTASHLSQQADLR		gi 13097708 +13	3.4e4	0.57	5931

<b>W</b>	gi 76634736 ref XP_584759.2  PREDICTED: similar to oxidation resistance 1 [Bos taurus]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 1.1e5	<b>Avg TIC:</b> 3.7e4	<b>Cov:</b> 6.3%	<b>Uniq:</b> 3
<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K) KEDFCIQIEIWAFK		gi 76634736	3.5e4	0.96	6181	
(K) ADLESEFRPNLSDPSHLLLPDQIIK		gi 76634736	5.4e4	0.74	5658	
(R) TIGYPWTLVYGTR		gi 76634736	2.2e4	0.73	5822	
<b>X</b>	gi 435476 emb CAA82315.1  cytokeratin 9 [Homo sapiens] gi 81175178 sp P35527 K1C9_HUMAN Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9)	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 6.5e4	<b>Avg TIC:</b> 2.2e4	<b>Cov:</b> 9.5%	<b>Uniq:</b> 3
<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K) DIENQYETQITQIEHEVSSSGQEVQSSAK		gi 435476 +4	4.6e4	0.98	5791	
(K) VQALEEANNLDLENK		gi 435476 +4	1.7e4	0.98	3902	
(R) HGVQEIEIEIQSLSK		gi 435476 +3	2.5e3	0.29	5340	
<b>Y</b>	gi 74354825 gb AAI02944.1  Carbonyl reductase 1 [Bos taurus] gi 77735973 ref NP_001029685.1  carbonyl reductase 1 [Bos taurus]	<b>MS/MS Spectra:</b> 4	<b>Sum TIC:</b> 9.4e4	<b>Avg TIC:</b> 2.3e4	<b>Cov:</b> 23.5%	<b>Uniq:</b> 4
<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K) EYGGDLVLVNAGIAFK		gi 74354825 +11	1.7e4	0.96	5924	
(K) FSGDVLTAR		gi 74354825 +6	2.7e4	0.92	4163	
(R) AAVQQLQAEGLSPLFHQLDIDDR		gi 74354825	3.7e4	0.91	5869	
(R) DVCTELLPLIKPQGR		gi 74354825 +9	1.4e4	0.33	5427	
<b>Z</b>	gi 70778950 ref NP_001020491.1  protein tyrosine kinase 9 [Bos taurus] gi 58760463 gb AAW82139.1  PTK9 protein tyrosine kinase 9 [Bos taurus]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 8.9e4	<b>Avg TIC:</b> 3.0e4	<b>Cov:</b> 12.3%	<b>Uniq:</b> 2
<b>Sequence</b>		<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(R) KIEIDNGDELTDLFYEEVHPK		gi 70778950 +15	2.7e4	0.98	5539	
(K) HQTLDGVAFPISEQEAFQALEK		gi 70778950	3.2e4	0.93	5742	
(R) KIEIDNGDELTDLFYEEVHPK		gi 70778950 +15	3.0e4	0.88	5645	

<b>a</b>	gi 62751620 ref NP_001015663.1  hypothetical protein LOC537221 [Bos taurus] gi 59858499 gb AAX09084.1  Ufm1-conjugating enzyme 1 (Ubiquitin-fold modifier-conjugating enzyme 1)	<b>MS/MS Spectra:</b> 4	<b>Sum TIC:</b> 7.5e4	<b>Avg TIC:</b> 1.9e4	<b>Cov:</b> 40.1%	<b>Uniq:</b> 3
	<b>Sequence</b> (R) YVENNKNADNDWFR (K) YFDIEFDIPITYPTTAPEIAVPELDGK (K) FGLAHLM*ALGLGPWLAVEIPDLIQK (K) YFDIEFDIPITYPTTAPEIAVPELDGK	<b>Reference</b> gi 62751620 +8 gi 62751620 +4 gi 62751620 +7 gi 62751620 +4	<b>TIC</b> 5.0e4 7.0e3 1.2e4 5.8e3	<b>Sf</b> 0.70 0.67 0.57 0.31	<b>Scan</b> 4154 6727 7118 6811	
<b>b</b>	gi 76648962 ref XP_872147.1  PREDICTED: similar to Y55F3AM.10 [Bos taurus]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 7.9e4	<b>Avg TIC:</b> 2.6e4	<b>Cov:</b> 8.9%	<b>Uniq:</b> 3
	<b>Sequence</b> (R) EGAIQVQGQSLFFR (K) INAADYAR (R) EALPGGGQAAR	<b>Reference</b> gi 76648962 gi 76648962 +1 gi 76648962	<b>TIC</b> 4.1e4 1.8e4 2.1e4	<b>Sf</b> 0.90 0.63 0.30	<b>Scan</b> 5206 3345 3048	
<b>c</b>	gi 76691419 ref XP_592628.2  PREDICTED: similar to alpha 1 type XVIII collagen isoform 1 precursor, partial [Bos taurus]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 7.9e4	<b>Avg TIC:</b> 2.6e4	<b>Cov:</b> 7.6%	<b>Uniq:</b> 3
	<b>Sequence</b> (R) AATGQASSLLAGR (R) ATLPVNLRL (K) SVWHGSDPSGR	<b>Reference</b> gi 76691419 gi 76691419 gi 76691419 +15	<b>TIC</b> 4.4e4 1.6e4 1.9e4	<b>Sf</b> 0.86 0.38 0.26	<b>Scan</b> 3784 4525 3110	
<b>d</b>	gi 19263767 gb AAH25341.1  Similar to cytoskeleton-associated protein 4 [Homo sapiens]	<b>MS/MS Spectra:</b> 3	<b>Sum TIC:</b> 7.8e4	<b>Avg TIC:</b> 2.6e4	<b>Cov:</b> 7.3%	<b>Uniq:</b> 3
	<b>Sequence</b> (K) VQEQVHTLLSQDQAQAAAR (K) ASVSQVEADLK (K) IETNENNLESAK	<b>Reference</b> gi 19263767 +4 gi 19263767 +4 gi 19263767 +11	<b>TIC</b> 4.5e4 1.2e4 2.1e4	<b>Sf</b> 0.96 0.58 0.57	<b>Scan</b> 3883 4174 3226	

e		gi 28317 emb CAA32649.1  unnamed protein product [Homo sapiens]	MS/MS Spectra: 3	Sum TIC: 6.9e4	Avg TIC: 2.3e4	Cov: 9.1%	Uniq: 3
<b>Sequence</b>			<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(R)	NVQALEIELQSQLALK		gi 28317 +5	1.8e4	0.94	5910	
(K)	TIDDLKNQILNLLTTDNANILLQIDNAR		gi 28317 +6	3.5e4	0.94	6729	
(K)	IRLENEIQTYR		gi 28317 +8	1.6e4	0.30	4102	
<b>f</b>							
f		gi 73587245 gb AAI02110.1  Rho GDP dissociation inhibitor (GDI) beta [Bos taurus]	MS/MS Spectra: 3	Sum TIC: 7.5e4	Avg TIC: 2.5e4	Cov: 25.5%	Uniq: 2
gi 6007524 gb AAF00938.1  D4-GDP-dissociation inhibitor [Bos taurus] gi 13626951 sp Q9TU03 GDIS_BOVIN Rho GDP-dissociation inhibitor 2 (Rho GDI 2) (Rho-GDI beta) (Ly-GDI) (D4-GDP-dissociation inhibitor) (D4-GDI)							
<b>Sequence</b>			<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K)	ATFM*VGSYGPRPEEYFLTPIEEAPK		gi 73587245 +1	3.2e4	0.87	5617	
(R)	LTLVCESAPGITMDLTGDLEALKK		gi 73587245 +4	1.5e4	0.62	6052	
(K)	ATFMVGSYGPRPEEYFLTPIEEAPK		gi 73587245 +1	2.9e4	0.50	5758	
<b>g</b>							
g		gi 2500561 sp P80929 ANG2_BOVIN Angiogenin-2	MS/MS Spectra: 2	Sum TIC: 4.3e4	Avg TIC: 2.1e4	Cov: 21.1%	Uniq: 2
<b>Sequence</b>			<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K)	DTNTFIHGNSDDIR		gi 2500561	3.4e4	0.88	3813	
(R)	AVCDDRRNGEPYR		gi 2500561	8.8e3	0.20	3038	
<b>h</b>							
h		gi 169777 gb AAA33898.1  beta-amylase	MS/MS Spectra: 2	Sum TIC: 4.6e4	Avg TIC: 2.3e4	Cov: 5.3%	Uniq: 2
<b>Sequence</b>			<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K)	LFGFTYLR		gi 169777 +17	8.1e3	0.66	5653	
(R)	NIEYLTGVDVDDQPLFHGR		gi 169777 +33	3.8e4	0.50	5550	
<b>i</b>							
i		gi 76646481 ref XP_586135.2  PREDICTED: similar to disabled homolog 2 isoform b [Bos taurus]	MS/MS Spectra: 2	Sum TIC: 3.6e4	Avg TIC: 1.8e4	Cov: 3.8%	Uniq: 2
<b>Sequence</b>			<b>Reference</b>	<b>TIC</b>	<b>Sf</b>	<b>Scan</b>	
(K)	LIGIDVDPAR		gi 76646481 +33	2.5e4	0.92	4555	
(K)	DLFQVYNTK		gi 76646481 +6	1.1e4	0.88	5643	

Single, ungrouped sequences		MS/MS Spectra: 9	Sum TIC: 2.6e5	Avg TIC: 2.8e4
Sequence	Reference	TIC	Sf	Scan
(R) TGEVEGANQLLELFDLVR Unknown (protein for MGC:133592) [Bos taurus] gi 82697401 ref NP_001032563.1  hypothetical protein LOC613626 [Bos taurus]	gi 79157775	8.6e3	0.97	7000
(K) VWVYTLDDGGELVNVYR galactose mutarotase [Sus scrofa] gi 84028320 sp Q9GKX6 GALM_PIG Aldose 1-epimerase (Galactose mutarotase) gi 11611545 dbj BAB18973.1  aldose 1-epimerase [Sus scrofa]	gi 47523866 +2	2.1e4	0.97	5767
(K) ANDTQEFNLSAYFEK Unknown (protein for MGC:134378) [Bos taurus]	gi 83405412 +1	1.7e4	0.97	5595
(R) SYIQLWAFDAVK inter-alpha-trypsin inhibitor (protein HC), light [Bos taurus] gi 2506821 sp P00978 AMBP_BOVIN AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha-trypsin inhibitor light chain (ITI-LC) (Bikunin) (HI-30) (BI-14) (Cumulus extracellular matrix-stabilizing	gi 27806743	4.0e4	0.96	5988
(R) TGEVEGANQLLELFDLVR Unknown (protein for MGC:133592) [Bos taurus] gi 82697401 ref NP_001032563.1  hypothetical protein LOC613626 [Bos taurus]	gi 79157775	1.4e4	0.95	6965
(K) SLGFTEEGIVFLPK prostaglandin H2 D-isomerase [Bos taurus] gi 3914330 sp O02853 PTGDS_BOVIN Prostaglandin-H2 D-isomerase precursor (Lipocalin-type prostaglandin-D synthase) (Glutathione-independent PGD synthetase) (Prostaglandin-D2 synthase) (PGD2 synthase) (PGDS2) (PGDS) gi 2196434 dbj	gi 27807521 +1	3.5e4	0.91	5824
(K) VDLLNQEIEFLK epidermal cyokeratin 2 [Homo sapiens] gi 2565257 gb AAB81946.1  keratin 2e [Homo sapiens] gi 547754 sp P35908 K22E_HUMAN Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK 2e)	gi 181402 +1	2.0e4	0.91	5855
(R) WLQGSQELPR Ig A H	gi 229537 +71	3.2e4	0.88	4184
(R) YIELPFGSEELK PREDICTED: similar to cathepsin S preproprotein [Bos taurus] gi 74353837 gb AAI02246.1  Cathepsin S [Bos taurus] gi 75812934 ref NP_001028787.1  cathepsin S [Bos taurus]	gi 61858888 +2	6.8e4	0.82	5473
<b>Dir:</b> mrocar473-gt	<b>Files:</b> 060419Ymirr473-gt (04/20/06-04/20/06)	<b>Enz:</b> Trypsin		
<b>Oper:</b> wsl	<b>Date:</b> 5/6/06	<b>Report:</b> R473-GTsnr (wsl)		<b>MR:</b> 2

**Legend:**

**Sequence:** The isobaric residue pairs Leu/Ile, Gln/Lys and Phe/Msx are displayed with the assignment as in the known sequence. Either residue within the pair may be possible: the displayed assignment does not connote a defined assignment to one or the other. The amino acid N-terminal to the known sequence is displayed in parentheses ().

**Ions:** The number of fragment ions (b-, y- and/or a-ions) experimentally observed / number of fragment ions possible. This fraction is a crude estimate of the minimum percentage of the sequence represented by the spectrum.

**Reference:** The database reference which contains the displayed sequence. A reference followed by a plus sign and number (e.g. +2) indicates the displayed sequence is also present in that number of additional database references

**TIC:** The intensity (total ion current) of the MS/MS spectrum. This is a unitless number. Note: mass spectrometry of different peptide analytes is not quantitative, because each peptide's ionization is dependent on structure. Caution should be used in using these values to ascertain major vs. minor components.

**Histograms:** Each set of peptide sequences grouped A, B, C...etc. by common database identifier have three values in their header: Sum TIC, Avg TIC and number of MS/MS Spectra. Histograms plot the percent of the total that each of these is.

**Sum TIC:** Sum of all peptides' TIC in the group

**Avg TIC:** Average of all peptides' TIC in the group

**MS/MS Spectra:** Number of peptides in the group

**Scan:** The scan number(s) of the acquired MS/MS spectrum.

**Modifications:**

\* = +15.99491 on the preceding M

The residue weight used for C was 160.03064

## REFERENCES

- Affinity Chromatography Handbook, Amersham Biosciences: Sweden, 2001
- Akimov, S.; Belkin, A. Cell Surface Tissue Transglutaminase is Involved in Adhesion and Migration of Monocytic Cell on Fibronectin. *Blood*. **2001**, *98*, 1567-1576.
- Awade, A.; Cleuzait, P.; Gonzales, T.; Robert-Baudouy, J. Pyrrolidone carboxyl peptidase (Pcp): An Enzyme that Removes Pyroglutamic Acid (pGlu) from pGlu-peptides and pGlu-proteins. *Proteins*. **1994**, *20*, 31-51.
- Baeseler, M.; Boeden, H.; Koelsch, R.; Lasch, J. Cellulose Beads: a Weak Leaking Affinity Support. *J. Chromatogr. A*. **1992**, *589*, 93-100.
- Beck, A.; Bussat, M.; Klinguer-Hamour, C.; Goestsh, L.; Aubry, J.; Champion, T.; Julien, E.; Haeuw, J.; Bonnefoy, J.; Corvaia, N. Stability and CTL Activity of N-terminal Glutamic Acid Containing Peptides. *J. Peptide Res.* **2001**, *57*, 528-538.
- Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.
- Bethell, G.; Ayers, J.; Hancock, W.; Hearn, M. A Novel Method of Activation of Cross-linked Agaroses with 1,1'-Carbonyldiimidazole which Gives a Matrix for Affinity Chromatography Devoid of Additional Charged Groups. *J. Biol. Chem.* **1979**, *254*, 2572-2574.
- Bjellqvist, B.; Ek, K.; Righetti, P.; Gianzza, E.; Gorg, A.; Westermeier, R.; Postel, W. Isoelectric Focusing in Immobilized pH Gradients: Principle, Methodology and Some Applications. *J. Biochem. Biophys. Methods*. **1982**, *6*, 317-339.
- Bowser, T.E., Active-site Inhibitors of  $\gamma$ -Glutamylamine Cyclotransferase.. Ph.D. Dissertation, Baylor University, TX, 1997.

- Bradford, M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding. *Anal. Biochem.* **1976**, 72, 248-254.
- Brown, J.; Roberts, W. Evidence that Approximately Eighty per Cent of the Soluble Protein from Ehrlich Ascites Cell are N<sup>α</sup>-Acetylated. *J. Biol. Chem.* **1976**, 251, 1009-1014.
- Busby, W.; Quackenbush, G.; Humm, J.; Youngblood, W.; Kizer, J. An enzyme(s) that converts glutaminyl-peptides into pyroglutamyl-peptides. *J. Biol. Chem.* **1987**, 262, 8532-8536.
- Chang, L.; Shepherd, D.; Sun, J.; Ouellette, D.; Grant, K.; Tang, X.; Pikal, M. Mechanism of Proteins Stabilization by Sugars During Freeze-drying and Storage. *J. Pharm. Sci.* **2005**, 94, 1427-1444.
- Chittum, H.; Lane, W.; Carlosn, B.; Roller, P.; Lung, F.; Lee, B.; Hatfield, D. Rabbit β-Globin is Extended Beyond its UGA Stop Codon by MultipleSuppressions and Translational Reading Gaps. *Biochem.* **1998**, 37, 10866-10870.
- Connell, G.; Hanes, C. Enzymic Formation of Pyrrolidone Carboxylic Acid from γ-Glutamyl Peptides. *Nature.* **1956**, 177, 377-378.
- Current Protocols in Protein Science; Coligan, J., Ben, D., Ploegh, H., Speicher, D., Wingfield, P., Ed.; John Wiley & Sons, Inc: New York, 1999; Units 6, 11, 15, 16.
- Dahl, S.; Slaughter, C.; Lauritzen, C.; Beteman, R.; Connerton, I.; Pedersen, J. Carica Papaya Glutamine Cyclotransferase Belongs to a Novel Plant Enzyme Subfamily: Cloning and Characterization of the Recombinant Enzyme. *Prot. Expres. Purif.* **2000**, 20, 27-36.
- Denu, J.; Dixon, J. A Catalytic Mechanism for the Dual-Specific Phosphatase. *Proc. Natl. Acad. Sci. USA.* **1996**, 92, 5910-5914.
- Doolittle, R.; Armentrout, R. Pyrrolidonyl Peptidase. An Enzyme for Selectove Removal of Pyrrolidonecarboxylic Acid Residues from Polypeptides. *Biochem.* **1968**, 7, 516-521.
- Edman, P.; Begg, G. A Protein Sequenator. *European J. Biochem.* **1967**, 1, 80-91.

- Eng, J.; McCormack, A.; Yates, J. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-89.
- Fabbi, M.; Marimpietri, D.; Martini, S.; Brancolini, C.; Amoresano, A.; Scaloni, A.; Bargellesi, A.; Cosulich, E. Tissue Transglutaminase is a Caspase Substrate During Apoptosis. Cleavage Causes Loss of Transamidating Function and is a Biochemical Marker of Caspase 3 Activation. *Cell Death Differ.* **1999**, *6*, 992-1001.
- Fesus, L.; Szondy, Z. Transglutaminase 2 in the Balance of Cell Death and Survival. *FEBS Lett.* **2005**, *579*, 3297-3302.
- Fesus, L.; Tarcsa, E. Formation of  $\epsilon$ -( $\gamma$ -glutamyl)-lysine Isodipeptide in Chinese-Hamster Ovary Cells. *Biochem. J.* **1989**, *263*, 843-848.
- Fink, M.; Folk, J.  $\gamma$ -Glutamylamine Cyclotransferase. *Method Enzymol.* **1983**, *94*, 347-51.
- Fink, M.; Folk, J.  $\gamma$ -Glutamyl Cyclotransferase. An Enzyme Involved in the Catabolism of  $\epsilon$ -( $\gamma$ -Glutamyl)lysine and other  $\gamma$ -Glutamylamines. *Mol. Cell. Biochem.* **1981**, *39*, 59-67.
- Fink, M.; Chung, S.; Folk, J.  $\gamma$ -Glutamylamine Cyclotransferase: Specificity toward  $\epsilon$ -(L- $\gamma$ -glutamyl)-L-lysine and Related compounds. *Proc. Natl. Acad. Sci. USA.* **1980**, *77*, 4564-4568.
- Fischer, W.; Spiess, J. Identification of a Mammalian Glutamyl Cyclase Converting Glutamyl into Pyroglutamyl Peptide. *Proc. Natl. Acad. Sci. USA.* **1987**, *84*, 3628-3632.
- Folk, J. Mechanism and Basis for Specificity of Transglutaminases-catalyzed  $\epsilon$ -( $\gamma$ -glutamyl)lysine Bond Formation. *Adv. Enzymol. RAMB.* **1983**, *54*, 1-56.
- Folk, J. Transglutaminases. *Ann. Rev. Biochem.* **1980**, *49*, 517-531.
- Folk, J.; Chung, S. Molecular and Catalytic Properties of Transglutaminases. *Adv. Enzymol. RAMB.* **1973**, *38*, 109-191.
- Folk, J. Mechanism of Action of Guinea Pig Liver Transglutaminase. *J. Biol. Chem.* **1969**, *244*, 3707-3713.

- Fox, B.; Yee, V.; Pedersen, L.; Le, T.; Bishop, P.; Stenkamp, R.; Teller, D. Identification of the Calcium Binding Site and a Novel Ytterbium Site in Blood Coagulation Factor XIII by X-ray Crystallography. *J. Biol. Chem.* **2000**, 274, 4917-4923.
- Galinski, E. Osmodaptation in Bacteria. *Adv. Microbio. Physiology.* **1995**, 37, 273-319.
- Gekko, K.; Timasheff, S. Mechanism of Protein Stabilization by Glycerol: Preferential Hydration in Glycerol-Water Mixtures. *Biochem.* **1981**, 20, 4667-4676.
- Gel Filtration Chromatography Handbook, Amersham Biosciences: Sweden, 2001
- Gerrard, J.; Sutton, K. Addition of Transglutaminase to Cereal Products may Generate the Epitope Responsible for Coeliac Disease. *Trends Food Sci Tech.* **2005**, 16, 510-512.
- Graves, D.; Wu, Y. On predicting the Results of Affinity Procedures. *Method Enzymol.* **1974**, 34, 140-162.
- Griffin, M.; Casadio, R.; Bergamini, C. Transglutaminases: Natural's Biological Glues. *Biochem. J.* **2002**, 368,377-396.
- Griffin, P.; MacCoss, M.; Eng, J.; Blevins, R.; Aaronson, J.; Yates, J. Direct Database Searching with MALDI-PSD Spectra of Peptides. *Rapid Commun. Mass Sp.* **1995**, 9, 1546-51.
- Gololobov, M.; Song, I.; Wang, W.; Bateman, R. Steady-state of Glutamine Cyclotransferase. *Arch. Biochem. Biophys.* **1994**, 309, 300-307.
- Gonzalez, F. Purification, Characterization and Inhibition Studies of  $\gamma$ -Glutamylamine Cyclotransferase from Rabbit Kidney. Ph.D. Dissertation, Baylor University, TX, 2005.
- Gowda, L. Studies on  $\gamma$ -Glutamylamine Cyclotransferase and  $\gamma$ -Glutamyl Transpeptidase. Ph.D. Dissertation, Baylor University, TX, 1985.
- Guex, N.; Peitsch, M. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis.* **1997**, 18, 2714-2723.

Guide to Isoelectric focusing. Amersham Biosciences: Sweden, 1998.

Haff, L.; Fagerstam, L.; Barry, A. Use of Electrophoretic Titration Curves for Predicting the Optimal Chromatographic Conditions for the Fast Ion Exchange Chromatography of Proteins. *J. Chromatogr.* **1983**, 266, 409-425.

Huang, K.; Liu, Y.; Cheng, W.; Ko, T.; Wang, A.; Crystal Structure of Human Glutaminyl Cyclase, an Enzyme Responsible for the Protein N-terminal Pyroglutamate Formation. *Proc. Natl. Acad. Sci. USA.* **2005**, 102, 13117-13122.

Hunt, D.; Yates, J.; Shabanowitz, J.; Winston, S.; Hauer, C. Protein Sequencing by Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. USA.* **1986**, 83, 6233-6237.

Hur, S.; Bruice, T. The Mechanism of Cis-trans Isomerization of Prolyl Peptides by Cyclophilin. *J. Am. Chem. Soc.* **2002**, 124, 7303-7313.

Ikeda, Y.; Fujii, J.; Taniguchi, N.; Meister, A. Human gamma-Glutamyl Transpeptidase Mutants Involving Conserved Aspartate Residues and the Unique Cysteine Residue of the Light Subunit. *J. Biol. Chem.* **1995 a**, 270, 12471-12475.

Ikeda, Y.; Fuji, J.; Anderson, M.; Taniguchi, N.; Meister, A. Involvement of Ser-451 and Ser-452 in the Catalysis of human gamma-Glutamyl Transpeptidase. *J. Biol. Chem.* **1995 b**, 270, 22223-22228.

Ion Exchange Chromatography and Chromatofocusing: Principles and Methods, Amersham Biosciences: Sweden, 2004.

Ito, K.; Inoue, T.; Takahashi, T.; Huang, H.; Esumi, T.; Hatakeyama, S.; Tanaka, N.; Nakamura, K.; Yoshimoto, T. The Mechanism of Substrate Recognition of Pyroglutamyl-peptidase I from *Bacillus Amylolyquefaciens* as Determined by X-ray Crystallography and Site-Directed Mutagenesis. *J. Biol. Chem.* **2001**, 276, 18557-18562.

Kang, H.; Park, S.; Cho, Y. Identification of  $\gamma$ -Glutamylamine Cyclotransferase, as the Preform Enzyme at the Dormant Stage, from Soybean (*Glycine max*) Seeds. *J. Biochem. Mol. Biol.* **1997**, 30, 438.

- Kaufmann, M. Review. Unstable Proteins: How to Subject Them to Chromatographic Separation for Purification Procedures. *J. Chromatogr. B.* **1997**, 699, 347-369.
- Kelleher, N.; Lin, H.; Valaskovic, G.; Aaserud, D.; Fridriksson, E.; McLafferty, F. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, 121,806–812.
- Komoto, J.; Yamada, Y.; Konishi K.; Ogawa, H.; Gomi, T.; Fuhioka, M.; Takusagawa, F. Catalytic Mechanism of Guanidinoacetate Methyltransferase: Crystal Structure of Guanidinoacetate Methyltransferase Ternary Complexes. *Biochem.* **2004**, 43, 14385-14394.
- LeGender, N.; Matsudaira, P. Purification of Protein and Peptides by SDS-PAGE. In *A Practical Guide to Protein and Peptide Purification for Microsequencing*; Matsudaira, P., Ed.; Academic Press, INC: San Diego, 1989; p 49-69.
- Liao, Y.; Wang, S.; Leu, Y.; Wang, C.; Chang, S.; Hong, Y.; Pan, Y.; Chen, C. The Structural Integrity Exerted by N-terminal Pyroglutamate is Crucial for the Cytotoxicity of Frog Ribonuclease from *Rana pipiens*. *Nucleic Acids Res.* **2003**, 31, 5247-5255.
- Lindblom, H.; Axio-Fredriksson, U. Separation of Urine Protein on the Anion-exchanger Resin Mono Q. *J. Chromatogr.* **1983**, 273, 107-116.
- Liu, Y.; Sowdal, L.; Robinson, N. Separation and Quantitation of Cytochrome C Oxidase by Mono-Q Fast Protein Liquid Chromatography and C18 Reverse-phase High Performance Liquid Chromatography. *Archive Biochem. Biophys* **1995**, 324, 135-142.
- Lorand, L.; Graham, R. Transglutaminases: Crosslinking Enzymes with Pleiotropic Functions. *Nature.* **2003**, 4, 140-156.
- Lorand, L. Commentary. Neurodegenerative Diseases and Transglutaminase. *Proc. Natl. Acad. Sci. USA.* **1996**, 93, 14310-14313.

- Lou, Y.; Huang, Y.; Pan, Y.; Chen, C.; Liao, Y. Roles of N-terminal Pyroglutamate in Maintaining Structural Integrity and pKa Values of Catalytic Histidine Residues in Bullfron Ribonuclease 3. *J. Mol. Biol.* **2006**, 355, 409-421.
- Matacic, S.; Loewy, A. Presence of the  $\epsilon$ -( $\gamma$ -Glutamyl)lysine Crosslink in Cellular Proteins. *Biochi. Biophys. Acta.* **1979**, 576, 263-268.
- Matacic, S.; Loewy, A. The Identification of Isopeptide Crosslinks in Insoluble Fibrin. *Biochem. Biophys. Res. Comm.* **1968**, 30, 356-362.
- Meister, A. Glutathione Metabolism and its Selective Modification. *J. Biol. Chem.* **1988**, 263, 17205-17208.
- Messer, M.; Ottesen, M. Isolation and Properties of Glutamine Cyclotransferase of Dried Papaya Latex. *C. R. Trav. Lab. Carlsberg.* **1965**, 35, 1-24.
- Mooz, D.; Wigglesworth, L. Evidence for the  $\gamma$ -Glutamyl Cycle in Yeast. *Biochem. Biophys. Res. Comm.* **1976**, 68:1066-1072.
- Murzin, A.; Brenner, S.; Hubbard, T.; Chothia, C. SCOP: a Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *J. Mol. Biol.* **1995**, 247, 536-540.
- National Center for Biotechnology Information. <http://www.ncbi.nih.gov> (accessed May, 2006)
- National Human Genome Research Institute. <http://www.genome.gov> (accessed May, 2006)
- Negro, A.; Garbisa, S.; Gotte, L.; Spina, M. The Use of Reverse-phase High-Performance Liquid Chromatography and Precolumn Derivatization with Dansyl Chloride for Auantitation of Specific Amino Acids in Collagen and Elastin. *Anal. Biochem.* **1987**, 160, 39-46.
- Nemes, Z.; Devreese, B.; Steinert, P.; Van Beeumen, J.; Fesus, L. Corss-linking of Ubiquitin, HSP27, and  $\alpha$ -Synuclein by  $\gamma$ -Glutamyl- $\epsilon$ -lysine Bonds in Alzheimer's Neurifibrillary Tangles. *FASEB J.* **2004**, 18, 1135-1137.

- Nemes, Z.; Fesus, L.; Egerhazi, A.; Keszthelyi, A.; Degrell, I. Nε(γ-Glutamyl)lysine in Cerebrospinal Fluid Marks Alzheimer Type and Vascular Dementia. *Neurobiol. Aging*. **2001**, *22*, 403-406.
- Oberg, K.; Ruyschaert, J.; Azarkan, M.; Smolders, N.; Zerhouni, S.; Wintjens, R.; Amrani, A.; Looze, Y. Papaya Glutamine Cyclase, a Plant Enzyme Highly Resistant to Proteolysis Adopts an All-β-Conformation. *Eur. J. Biochem.* **1998**, *258*, 214-222.
- Odagaki, Y.; Hayashi, A.; Okada, K.; Hirotsu, K.; Kabashima, T.; Ito, K.; Yoshimoto, T.; Tsuru, D.; Sato, M.; Clardy, J. The Crystal Structure of Pyroglutamyl Peptidase I from *Bacillus amyloliquefaciens* Reveals a New Structure for a Cysteine Protease. *Structure*. **1999**, *7*, 399-411
- O'Farrell, P. High Resolution Two-dimensional Electrophoresis of Proteins. *J. Biol. Chem.* **1975**, *250*, 4007-4021.
- Okada, T.; Suzuki, H.; Wada, K.; Kumagai, H.; Fukuyama, K. Crystal Structures of gamma-Glutamyltranspeptidase from *Escherichia coli*, a Key Enzyme in Glutathione Metabolism, and Its Reaction Intermediate. *Proc. Natl. Acad. Sci. U S A*. **2006**, *103*, 6471-6476.
- Oray, B.; Lu, H.; Gracy, R. High-performance Liquid Chromatography Separation of Dns-Amino-acid Derivative and Applications to Peptide and Protein Structural Studies. *J. Chromatogr.* **1983**, *270*, 253-266.
- Orlowski, M.; Meister, A. γ-Glutamyl Cyclotransferase. Distribution, Isozymic Forms, and Specificity. *J. Biol. Chem.* **1973**, *248*, 2836-2844.
- Orlowski, M.; Richman, P.; Meister, A. Isolation and Properties of γ-L-Glutamylcyclotransferase from Human Brain. *Biochem.* **1969**, *8*, 1048-1055.
- Pedersen, L.; Yee, V.; Bishop, P.; Le T.; Teller, D.; Stenkamp, R. Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. *Protein Sci.* **1994**, *3*, 1131-1135.
- Pelech, S.; Samiei, M.; Charest, D.; Howard, S.; Salari, H. Characterization of Calcium-independents Forms of Protein Kinase C-β in Phobol Ester-Treated Rabbit Platelets. *J. Biol. Chem.* **1991**, *266*, 8696-8705.
- Pierce. Hydrolysis Protocol. <http://www.pierce.com/index2.html>

Pisano, J.; Finlayson, J.; Peyton, M. Cross-link in Fibrin Polymerized by Factor XIII:  $\epsilon$ -( $\gamma$ -Glutamyl)lysine. *Science*. **1968**, 16, 892-893.

Pitson, S.; D'Andrea, R.; Vandeleur, L.; Moretti, P.; Xia, P.; Gamble, J.; Vadas, M.; Wattenberg, B. Human Spingosine Kinase: Purification, Molecular Cloning and Characterization of the Native and Recombinant Enzymes. *Biochem. J.* **2000**, 350, 429-441.

Pohl, T.; Zimmer, M.; Mugele, K.; Spiess, J. Primary Structure and Functional Expression of a Glutaminyl Cyclase. *Proc. Natl. Acad. Sci. USA*. **1991**, 88, 10059-10063.

Protein Calculator V3.3. <http://www.scripps.edu/~cdputnam/protcalc.html> (accessed May 2006)

Protein Purification Handbook, Amersham Biosciences: Sweden, 2001.

Radola, B. Analytical and Preparative Isoelectric Focusing in Gel-stabilized Layers. *Ann. NY Acad. Sci.* **1973**, 209:127-143.

Raghunath, M.; Hopfener, B.; Aeschlimann, D.; Luthi, U.; Meuli, M.; Altermatt, S.; Gobet, R.; Buckner-Tuderman, L.; Steinmann, B. Cross-linking of the Demo-epidermal Junction of Skin Regenerating from Keratinocyte Autograft. *J. Clin. Invest.* **1996**, 98, 1174-1184.

Research Collaboratory for Structural Bioinformatics Protein Data Bank. <http://www.rcsb.org/pdb/Welcome.do> (accessed May, 2006)

Reid, G.; McLuckey, S. "Top down" Protein Characterization Via Tandem Mass Spectrometry. *J. Mass Spectrom.* **2002**, 37, 663-675.

Righetti, P.; Bossi, A. Isoelectric Focusing and Peptides in Gel Slabs and in Capillaries. *Anal. Chim. Acta* **1998**, 372, 1-19.

Sanger, F.; Thompson, E.; Kitai, R. The Amide Groups of Insulin. *Biochem. J.* **1955**, 59, 509-518.

Schilling, S.; Niestroj, A.; Rahfeld, J.; Hoffmann, T.; Wermann, M.; Zunkel, K.; Wasternack, C.; Demuth, H. Identification of Human Glutaminyl Cyclase as a Metalloenzyme. *J. Biol. Chem.* **2003**, 278, 49773-49779.

- Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. SWISS-MODEL: an Automated Protein Homology-modeling Server. *Nucleic Acids Res.* **2003**, 31, 3381-3385.
- Seelig, G.; Folk, J. Half-of-the-sites and All-of-the-sites Reactivity in Human Plasma Blood Coagulation Factor XIIIa. *J. Biol. Chem.* **1980**, 255, 9589-9593.
- Sehgal, D.; Vijay, I. A Method for the High Efficiency of Water-soluble Carbodiimide-mediated Amidation. *Anal. Biochem.* **1994**, 218, 87-91.
- Smith, T.; Meister, A. Chemical Modification of Active Site Residues in  $\gamma$ -Glutamyl Transpeptidase. Aspartate 422 and Cysteine 453. *J. Biol. Chem.* **1995**, 270, 12476-12480.
- Song, I.; Chuang, C.; Bateman, R. Molecular Cloning, Sequence Analysis and Expression of Human Pituitary Glutamyl Cyclase. *J. Mol. Endocrinol.* **1994**, 1377-1386.
- Standing, K. Peptide and proteins *de novo* sequencing by mass spectrometry. *Curr. Opin. Struc. Biol.* **2003**, 13, 595-601.
- Stein, S. Isolation of Natural Proteins. In *Fundamentals of Protein Biotechnology*; Stein, S., Ed.; Marcel Dekker: New York, 1991; p137-160.
- Stole, E.; Meister, A. Interaction of  $\gamma$ -Glutamyl Transpeptidase with Glutathione Involves Specific Arginine and Lysine Residues of the Heavy Subunit. *J. Biol. Chem.* **1991**, 266, 17850-17857.
- Structural Classification of Proteins. <http://scop.mrc-lmb.cam.ac.uk/scop/index.html> (accessed May, 2006)
- Sudesh, M.; Lyddiatt, A. Recent Development in Affinity Chromatography Separation Technology. In *Affinity Separation a Practical Approach*. Matejtschuk, P., Ed.; OIRL Press: New York, 1997; Vol. 179; p 24-28.
- Svensson, H. Isoelectric Fractionation, Analysis, and Characterization of Ampholytes in Natural pH Gradients. I. The Differential Equation of Solute Concentrations at a Steady State and its Solution for Simple Cases. *Acta Chem. Scand.* **1961**, 15, 325-41.

SWISS-MODEL. <http://www.expasy.org/spdbv> (accessed May, 2006)

Taki, M.; Shiota, M.; Taira, K. Transglutaminase-mediated N- and C-Terminal Fluorescein Labeling of a Protein can Support the Native Activity of the Modified Protein. *Protein Eng. Des. Sel.* **2004**, 17, 119-126.

Taniguchi, N.; Ikeda, Y.  $\gamma$ -Glutamyl Transpeptidase: Catalytic Mechanisms and Gene Expression. *Adv. Enzymol. RMB.* **1998**, 72, 239-278.

Taniguchi, N.; Meister, A.  $\gamma$ -Glutamyl Cyclotransferase from Rat Kidney. Sulfhydryl Groups and Isolation of a Stable Form of the Enzyme. *J. Biol. Chem.* **1978**, 253:1799-1806.

Tulchin, N.; Taylor, B.  $\gamma$ -Glutamyl Cyclotransferase: A New Genetic Polymorphism in the Mouse (*Mus Musculus*) Linked to *Lyt-2*. *Genetics.* **1981**, 99, 109-116.

Udenfriend, S., Stein, S.; Bohlen, P.; Dairman, W.; Leimgruber, W.; Weigele, M. Fluorescamine: A Reagent for Assay of Amino Acids, Peptides, Proteins, and Primary Amines in the Picomole Range. *Science.* **1972**, 178, 871-872.

Vajdos, F.; Yoo, S.; Houseweart, M.; Sundquist, W.; Hill, C. Crystal Structure of Cyclophilin A Complexed with a Binding Site Peptide from the HIV-1 Capsid Protein. *Protein. Sci.* **1997**, 6, 2297-2307.

Van der Werf, P.; Orłowski, M.; Meister, A. Enzymatic Conversion of 5-Oxoproline (L-Pyrrolidone Carboxylate) to L-Glutamate Coupled with Cleavage of Adenosine Triphosphate to Adenosine Diphosphate, a Reaction in the  $\gamma$ -Glutamyl Cycle. *Proc. Nat. Acad. Sci. USA.* **1971**, 68, 2982-2985.

Voet, D. and Voet, J. *Biochemistry*, 2<sup>nd</sup> ed.; John Wiley and Sons Inc: 1995.

Waibel, P.; Carpenter, K. Mechanisms of Heat Damage in Proteins. *Brit. J. Nutr.* **1972**, 27, 509-515.

Wintjens, R.; Belrhali, H.; Clantin, B.; Azarkan, M.; Bompard, C.; Baeyens-Volant, D.; Looze, Y.; Villeret, V. Crystal Structure of Papaya Glutaminyl Cyclase, and Archetype for Plant and Bacterial Glutaminyl Cyclase. *J. Mol. Biol.* **2006**, 357, 457-470.

- Woodward, G.; Reinhart, F. The Effect of pH on the Formation of Pyrrolidone-carboxylic Acid and Glutamic Acid during Enzymatic Hydrolysis of Glutathione by Rat Kidney Extract. *J. Biol. Chem.* **1942**, 145, 471-480.
- Yates, J. Protein Structure Analysis by Mass Spectrometry. *Method Enzymol.* **1996**, 271,351-377.
- York, M.; Kuchel, P.; Chapman, B. A Proton Nuclear Magnetic Resonance Study of  $\gamma$ -Glutamyl-amino Acid Cyclotransferase in Human Erythrocytes. *J. Biol. Chem.* **1984**, 259, 15085-15088.
- Yuvaniyama, J.; Denu, J.; Dixon, J.; Saper, M. Crystal Structure of the Dual Specificity Protein Phosphatase VHR. *Science.* **1996**, 272, 1328-1331.