Molecular investigation of the rat jejunal active nucleoside transport protein rCNT1

Stephen Robin Hamilton

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Biochemistry and Molecular Biology

September 1998

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others

Abstract

Active, sodium-dependent nucleoside transport systems have been kinetically described in a number of different mammalian tissues. The recent cloning of several active, mammalian nucleoside transporters belonging to a novel gene family has provided an opportunity to examine both the structure/function relationships and the tissue distributions of these proteins in detail, and thus to elucidate their physiological roles. The first of these transporters to be cloned was rCNT1, a protein from rat jejunal epithelium which exhibits pyrimidine-selective, *cit*-type transport activity. The present study involved a bi-directional approach to the investigation of this protein by the analysis of its distribution in rat tissues at both the cellular and subcellular levels, in addition to structural investigation of rCNT1 by protein chemistry and gene manipulation techniques.

A model originally proposed for the topology of this protein, based largely upon hydropathy analysis of its sequence, suggested that the *N*- and *C*-terminal hydrophilic regions of this protein were cytoplasmic. However the present study has shown, from the effect of endoglycosidase F treatment of the protein expressed in *Xenopus* oocytes, that the *C*-terminal hydrophilic domain of the protein is *N*-glycosylated and so must in fact be extracellular. Furthermore, mutagenesis of rCNT1 cDNA allowed for the production of an aglyco-rCNT1 mutant into which novel glycosylation sites were introduced as a means of probing topology. Evidence from this mutagenesis has indicated that the *N*-terminus of rCNT1 is cytoplasmic, therefore suggesting that rCNT1 exists in the lipid bilayer with thirteen putative transmembrane domains and not fourteen as originally proposed. Analysis of the

protein sequence using predictive algorithms suggests that the originally-predicted transmembrane helix 6 is in fact extramembranous.

Analysis of the distribution of rCNT1 in rat tissues by immunoblotting has revealed that this active transporter is more widespread than originally suggested by the results of Northern blotting experiments. Moreover, it has been possible by immunocytochemical analysis not only to identify the cell types in which rCNT1 is expressed but also to show that its distribution is restricted to certain domains of the plasma membrane. rCNT1 in the jejunum and kidney has been localized at the intestinal and renal brush-borders respectively. While in the liver it is abundant in the bile canalicular membranes. This location is consistent with roles for the concentrative transporter in the absorption of nucleosides from the diet or their salvage from the extracellular breakdown of nucleotides. Unexpectedly, rCNT1 also appears to be abundant in the hippocampus, and in the actively dividing seminiferous epithelium of the testis. Furthermore, its presence has also been suggested in cardiac muscle, skeletal muscle and the spleen, although these observations await further confirmation.

To Mum and Dad

Contents

Abstract ii
Contents v
List of tables ix
List of figures x
Abbreviations for amino acids xii
Abbreviations xiii
Acknowledgements xv

Chapter 1 Introduction

1.1 Membrane Proteins: An Introduction 2
1.2 Nucleoside Transport 4
1.2.1 Physiological and synthetic substrates 4
1.2.2 Nucleoside transport and metabolism 6
1.2.3 Nucleoside transport and regulation 7
1.2.4 Classification of nucleoside transporters 9
1.3 A novel gene family 13
1.3.1 Establishment of homologues 13
1.3.2 Cloning of the first active nucleoside transporter: rCNT1 15
1.3.3 A comparative study of rCNT1 and the mammalian homologues 22
1.4 Aims of the study 27

Chapter 2 Materials and methods

2.1 Materials and Suppliers 30
2.2 Synthetic peptide synthesis and analysis 32
2.2.1 Peptide synthesis 32
2.2.2 Peptide cleavage 32
2.2.3 Peptide mass spectrometric analysis 33
2.2.4 Peptide high performance liquid chromatography 33
2.3 Production and characterisation of anti-peptide antibodies 34
2.3.1 Conjugation of peptides to ovalbumin 34

2.3.2 Affinity purification of anti-peptide antibodies using immobilised synthetic peptides 34
2.3.3 Enzyme-linked immunosorbent assay 36
2.4 General Protein techniques 37
2.4.1 Membrane preparation 37
2.4.2 Protein determination 37
2.4.3 SDS-polyacrylamide gel electrophoresis 38
2.4.4 Immunoblotting 38
2.5 Molecular biology techniques 38
2.5.1 Small scale isolation of plasmid DNA 39
2.5.2 Large scale isolation of plasmid DNA 40
2.5.3 Restriction analysis and agarose gel electrophoresis of plasmid DNA 41
2.5.4 Preparation and transformation of competent bacterial cells with plasmid DNA 41
2.5.5 Purification of DNA fragments 42
2.5.6 Dephosphorylation of linearised plasmid DNA with Calf Intestine Alkaline Phosphatase (CIP) 43
2.5.7 Ligation of DNA fragments 43
2.5.8 Automated DNA sequencing 44
2.6 GST fusion protein expression 45
2.6.1 Small scale In vivo expression of fusion protein 45
2.6.2 Large scale In vivo expression of fusion protein 45
2.6.3 Affinity purification of GST fusion protein 46
2.6.4 Antibody production, purification and analysis 46

Chapter 3 Generation of anti-peptide antibodies

3.1 Introduction 49
3.2 Criteria for the design of antigenic peptides 51
3.3 The synthetic peptide approach 54
3.3.1 An introduction to peptide synthesis 54
3.3.2 Results and discussion 58
3.3.2.1 Peptide synthesis, cleavage and purification 58
3.3.2.2 Peptide analysis 59
3.3.2.3 Antibody production 60
3.3.2.4 Affinity purification of anti-peptide antibodies 61

3.4 The fusion protein approach 73
3.4.1 An introduction to fusion proteins 73
3.4.2 PCR amplification 75
3.4.3 Strategy for production of ΔrCNT1-GST fusion protein 76
3.4.4 Results and discussion 81
3.4.4.1 Production of ΔrCNT1-GST fusion protein construct 81
3.4.4.2 In vivo expression of fusion protein 84
3.4.4.3 Antibody production, purification and analysis 85
3.5 Summary 93

Chapter 4 Distribution of rCNT1 in rat tissues

4.1 Introduction 95
4.2 Immunocytochemistry: an introduction 97
4.3 Methods for tissue localization 101
4.3.1 Membrane preparation and Western blotting 101
4.3.2 Tissue extraction and processing for immunocytochemistry 101
4.3.3 Antigen retrieval pretreatment of formaldehyde fixed tissue 102
4.3.4 Immunocytochemistry of unfixed frozen tissue 103
4.3.5 Immunocytochemistry of paraformaldehyde-fixed tissue 103
4.4 Results and discussion 104
4.4.1 Antibody usage 104
4.4.2 Fixation vs. fresh tissue 105
4.4.3 Tissue distribution of rCNT1 122
4.4.4 Subcellular distribution of rCNT1 125
4.5 Summary 129

Chapter 5 Topological analysis of rCNT1

5.1 Introduction 132
5.2 Methods for topological analysis 137
5.2.1 Endoglycosidase digestion to determine the glycosylation state of membrane proteins 137
5.2.2 Mutagenic removal of rCNT1 endogenous glycosylation sites 137
5.2.3 Incorporation of potential glycosylation sites by single residue modification 139

5.2.4 Amplification of cDNA encoding the glycosylated hydrophilic loop of GLUT4 140
5.2.5 Incorporation of exogenous potential glycosylation sites 141
5.3 Results and discussion 152
5.3.1 Deglycosylation of wild-type rCNT1 152
5.3.2 Extent of wild-type rCNT1 glycosylation 153
5.3.3 Topological analysis by glycosylation scanning mutagenesis 155
5.4 Summary 161

Chapter 6 General discussion

6.1 Distribution of rCNT1 in rat tissues 168
6.2 Elucidating the topology of rCNT1 172
6.3 Epilogue 175

References 176
Appendix 1 187
Appendix 2 189

List of tables

Table 1.1 Functional properties of nucleoside transporter subclasses 18

Table 3.1 Side chain protecting groups for FMOC chemistry 57

Table 3.2 Synthetic peptide production yields 63

Table 3.3 Side-products of peptide synthesis 64

Table 3.4 Immunogenic response to synthetic peptides 65

Table 3.5 Summarization of antibody purification 72

Table 4.1 Comparison of unfixed and paraformaldehyde-fixed tissue processing 110

Table 4.2 Comparison of antigen retrieval techniques 111

Table 4.3 Comparison of post-sectioning fixation reagents 112

Table 5.1 Mutagenic primers used in the topological analysis of rCNT1 145

List of figures

Figure 1.1 Phylogenetic analysis of rCNT1 gene family 19

Figure 1.2 Putative model of rCNT1 secondary structure in the lipid bilayer 20

Figure 3.1 MS and rpHPLC analysis of synthetic peptide A 66

Figure 3.2 MS and rpHPLC analysis of synthetic peptide B 67

Figure 3.3 MS and rpHPLC analysis of synthetic peptide C 68

Figure 3.4 MS and rpHPLC analysis of synthetic peptide D 69

Figure 3.5 Development of immune response against peptide B conjugate 70

Figure 3.6 Affinity purification of anti-peptide D antibodies 71

Figure 3.7 Alignment of 5' regions of rCNT1 and hCNT1 cDNAs..... 79

Figure 3.8 Strategy for ΔrCNT1-GST fusion protein construction..... 80

Figure 3.9 Screening and confirmation of the ΔrCNT1-GST construct..... 88

Figure 3.10 SDS-PAGE analysis of ΔrCNT1-GST expression in *E. coli*..... 89

Figure 3.11 Immunoblot analysis with anti-ΔrCNT1 purified antibodies 90

Figure 3.12 Location of antibodies raised within rCNT1 91

Figure 4.1 Integrity of unfixed tissue 109

Figure 4.2 Western blot analysis of rat tissues with anti-peptide C antibodies.... 113

Figure 4.3 Western blot analysis of rat tissues with anti-peptide D antibodies ... 114

Figure 4.4 Subcellular distribution of rCNT1 in the rat jejunum..... 115

Figure 4.5 Subcellular distribution of rCNT1 in the rat kidney 116

Figure 4.6 Subcellular distribution of rCNT1 in the rat liver..... 117

Figure 4.7 Cellular distribution of rCNT1 in the rat brain 118

Figure 4.8 Subcellular distribution of rCNT1 in the rat heart 119

Figure 4.9 Subcellular distribution of rCNT1 in the rat skeletal thigh muscle 120

Figure 4.10 Subcellular distribution of rCNT1 in the rat testis..... 121

Figure 5.1 Amplification of human GLUT4 cDNA encoding a potential *N*-glycosylated loop 136

Figure 5.2 *QuikChange*™ mutagenesis strategy 142

Figure 5.3 Summary of mutation sites located within the putative model for rCNT1 143

Figure 5.4 Glycosylation state of wild-type rCNT1 and homologues 147

Figure 5.5 Kinetic comparison of rCNT1 deglycosylation mutants 148

Figure 5.6 Extent of wild-type rCNT1 glycosylation 149

Figure 5.7 Kinetic comparison of rCNT1 topological mutants..... 150

Figure 5.8 Glycosylation state of rCNT1 mutants 151

Figure 5.9 Three new putative models for the topology of rCNT1 in the lipid bilayer 165

Abbreviations for amino acids

Amino acid	three-letter code	one-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine or glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Abbreviations

The following abbreviations are not listed as standard as outlined in the Biochem. J. (1997) **329**, 1-16.

A_x	Absorbance at x nm
AZT	Azidothymidine
BBM(V)	Brush-border membrane (vesicle)
BCA	Bicinchoninic acid
BLM(V)	Basolateral membrane (vesicle)
CDM	Counterion distribution monitoring
cfu	Colony forming units
cib	Concentrative nucleoside transport with broad specificity
cif	Concentrative nucleoside transport specific for purines
CIP	Calf intestinal alkaline phosphatase
cs	Concentrative nucleoside transport sensitive to NBMPR
cit	Concentrative nucleoside transport specific for pyrmidines
citg	Concentrative nucleoside transport specific for pyrmidines and guanosine
DCM	Dichloromethane
ddC	Dideoxycytidine
Dhbt	Dihydrohydroxyoxobenzotriazole
DIPCDI	Diisopropylcarbodiimide
DMSO	Dimethylsulfoxide
DTNB	Dithiobisnitrobenzoic acid
DTT	Dithiothreitol
ei	Facilitative nucleoside transport not inhibited by NBMPR
ER	Endoplasmic reticulum
es	Facilitative nucleoside transport inhibited by NBMPR
FMOC	Fluorenylmethoxycarbonyl
GET	Glucose, EDTA, Tris buffer
GST	Glutathione S-transferase
HPLC	High performance liquid chromatography

IPTG	Isopropylthiogalactoside
KGB	Potassium glutamate buffer
LB	Luria-Bertani
M_r	Relative molecular mass
mRNA	Messenger RNA
MS	Mass spectrometry
NBMPR	Nitrobenzylmercaptopurineriboside
OCT	Tissue embedding medium supplied by BDH
o/n	Overnight (16 h)
PBS-T	PBS containing 0.2 % Tween-20
PEG	Polyethylene glycol
PEG-PS	Polyethylene glycol-polystyrene
Pfp	Pentafluorophenol
rpHPLC	reverse phase HPLC
rpm	Revolutions per minute
ΔrCNT1	Residues 21 to 78 of rCNT1
r.t.	room temperature (22°C)
tBu	t-Butyl
TBS(-T)	Tris buffered saline (containing 0.2 % Tween-20)
TE	Tris, EDTA buffer
TFA	Trifluoroacetic acid
TMD	Transmembrane domain
Tris	Tris(hydroxymethyl)aminomethane
Tween-20	Polyoxyethylene sorbitan monolaurate
v/v	Volume for volume
w/v	Weight for volume

Acknowledgements

I would like to thank Professor Steve Baldwin for his immense encouragement and invaluable advice throughout my time under his supervision. His enthusiasm for research has been awe-inspiring and should be an example to all researchers alike.

I am also extremely grateful to all members of the Baldwin and Henderson labs, both past and present, for sharing their knowledge and advice on technical matters as well as extending such friendliness and making my time at Leeds so enjoyable. In particular, thanks to Mrs. Jean Ingram for her technical support and unlimited patience over the period of this project.

Thanks are also due to our collaborators in Professor Jim Young's lab in Canada and Dr. Maurice Gallagher's in Edinburgh, with special thanks to Dr. Sylvia Yao (Canada) for her expertise and help with *Xenopus* oocyte expression.

I must also thank my spiritual advisors Andy, Gary, Graeme and Patrick, who so often persuaded me away from the turmoil's of bench-work to a place of calm and enlightenment, that being the Fav..

Last but not least, I would like to thank my family for all of their support and encouragement, especially my mother and father who have always been there for me.

Chapter 1

Introduction

1.1 Membrane Proteins: An Introduction

The basic unit of life, the cell, has to produce an internal milieu contrasting from its surrounding environment in order to survive. To achieve this it must possess an impermeable barrier to enclose its components. One of the earliest proposals concerning the state of this barrier was put forward by Gorter and Grendel in 1925 who suggested that this barrier was in the form of a bimolecular leaflet, or lipid bilayer [1]. This was further elaborated in the paucimolecular model of Davson and Danielli (1935) who proposed that on either side of the bilayer resided a protein monolayer and it was these protein layers that were responsible for the enzymatic properties exhibited by membranes [2]. This proposal remained relatively unchanged for the following thirty years until experimentation using freeze fracture electron microscopy showed that there were apparently globular proteins embedded in the membrane [3]. Spectroscopic evidence indicating the presence of an appreciable amount of α-helix likewise supported the presence of globular proteins rather than a monolayer of extended polypeptide chain [4,5]. Finally, appreciation of the non-polar characteristics of membrane proteins [6] and of the fluid nature of the lipid bilayer [7] culminated in the development of the fluid mosaic model by Singer and Nicolson in 1972 [8].

The globular proteins found in membranes showed great diversity in their function, being enzymes, receptors, pores, transporters etc., but the insoluble characteristics of these proteins and of the hydrophobic peptides derived from them made primary structure determination difficult. However, investigation of the sialoglycoprotein glycophorin, found in the erythrocyte membrane, indicated the presence of a short

stretch of 23 nonpolar amino acids near the middle of the molecule [9]. Topological and other studies indicated that glycophorin extended completely through the membrane and that this stretch was α-helical [10]. This discovery led to the concept of membrane spanning α-helical domains in membrane proteins. Many membrane proteins in fact are predicted to contain multiple membrane-spanning α-helices. For example, electron microscopy image reconstruction studies of bacteriorhodopsin, the predominant protein found in the purple membrane of *Halobacterium halobium*, the presence of seven α-helical segments traversing the bilayer [11].

While the hydrophobic nature of membrane proteins rendered their sequencing by conventional means difficult, deduction of these protein sequences could be obtained from cDNA and gene sequence data. Thus leading to an explosion in the amount of information available on transporter sequences. An example of which is the major facilitator superfamily, with the first member to be sequenced being the *lac* permease, but comprises greater than fifty other transporters including the bacterial proteins GalP, AraE and XylE in addition to the mammalian erythrocyte glucose transport protein GLUT1 [12,13]. Subjecting the amino acid sequences of members of this superfamily to hydropathic analysis produced profiles that identified typically that its members contained hydrophobic regions of approximately 21 amino acid residues in length which are predicted to traverse the membrane in the form of α-helices. In addition, application of secondary structure algorithms and the positive inside rule led to the prediction that the *N*- and *C*- termini are cytoplasmic with the fore-mentioned 12 hydrophobic regions traversing the membrane between them. Of course, such a predicted topology must be backed by experimentation before it can be confirmed that the model is a true reflection of the protein structure [14]. In the

case of GLUT1 there is much evidence to support the model described above, based on circular dichroism [15,16]; Fourier Transform Infra-red (FTIR) spectroscopy [17]; proteolytic digestion [18]; site-specific antibodies [19,20]; biotinylation [21]; and glycosylation scanning mutagenesis [22]. Thus suggesting that the predicted secondary structure for the major facilitator superfamily as a whole is correct. However, while the topology and secondary structure of the major facilitator superfamily members is now well established, this is not the case for many other unrelated transport protein families that have emerged from genome sequencing and other approaches over the past few years. An example is provided by the rat active nucleoside transport protein rCNT1, which forms the subject of this thesis and which will be discussed in more detail later in this chapter.

1.2 Nucleoside Transport

1.2.1 Physiological and synthetic substrates

Nucleosides perform a wide variety of specialized roles in different cells and tissues. For example, adenosine is involved in regulation of platelet and neutrophil functions, of blood flow in the heart, kidneys and other tissues, and functions as a local hormone in the regulation of lipolysis, as well as acting as a neuromodulator [23], whilst inosine acts as an *in vivo* energy supply for adult pig erythrocytes, cells which cannot transport and hence metabolize glucose [24]. Similarly, both inosine and guanosine are also energy sources for embryonic and adult chicken erythrocytes [25]. Moreover, not all cells can synthesize these molecules *de novo*. Cells lacking *de novo* biosynthetic pathways include enterocytes, erythrocytes, leukocytes, certain brain cells and bone marrow cells [26,27]. Such cells rely upon dietary nucleosides or upon

the production by other cells of an excess of nucleosides which can then be transported via the circulation to sites where they are needed. Such is one of the specialized functions of the liver parenchymal cells.

Due to the importance of nucleosides many synthetic analogues have been made over the years with the aim of developing useful clinical drugs. Indeed some of these compounds have proven useful in the treatment of viral diseases and haematologic cancers including the use of azidothymidine (AZT) and dideoxycytidine (ddC) as anti-viral nucleosides, with the former being used in the treatment of acquired immuno-deficiency syndrome (AIDS) [28,29]. The majority of these drugs act intracellularly by interfering with DNA synthesis following chemical modification, namely phosphorylation. More recently interest has expanded to the development of agonists and antagonists which act on the adenosine receptors as a means of preventing reperfusion injury after ischaemia and in the treatment of hypertension, epilepsy and renal failure [30].

In relation to the efficiency of such drugs, the presence of nucleoside transporters can greatly affect this. For instance, in experimental systems using cultured cells, which have been genetically modified to eliminate nucleoside transport, resistance to a variety of nucleoside analogs with anticancer activity has been shown [31,32]. This therefore demonstrates that the absence or presence of these transporters plays a key role in the pharmacokinetics, disposition and *in vivo* biological activity of nucleoside analogues. The specificity and activity of these transporters in different cell types varies greatly, but there are certain features that are common to several of them which allows for their allocation into distinct groups or subclasses, see **section 1.2.4**.

1.2.2 Nucleoside transport and metabolism

Uptake of nucleosides by cells is followed by modification of these substrates in the majority of cases, although cells such as enterocytes and hepatocytes are capable of releasing unmodified nucleosides into the circulation for use by other cells within the organism. The reason for this modification is to prevent the loss of the acquired substrate while altering its chemical properties to make it more readily metabolized. This modification may be either anabolic, as in the phosphorylation by nucleoside kinases, or catabolic, as in phosphorolytic cleavage or deamination by nucleoside phosphorylases or nucleoside deaminases respectively. Thus when considering the rate limiting step for uptake, not only is the rate of translocation important but also the relative activity of the metabolizing enzyme.

For the purposes of exemplification a few cases of nucleoside fate following transport will be discussed but for more in-depth coverage see the review by Boss and Seegmiller [33]. One case where post-translocational modification of the nucleosides is essential is in cell types that are unable to synthesize their own nucleotides *de novo*, such as is the case with erythrocytes and hematopoietic bone marrow cells. Therefore they rely on other cells, namely hepatocytes, for a supply of nucleosides which they generally modify in one of two ways. Pyrimidine nucleosides are generally phosphorylated by nucleoside kinases whereas purines are modified by the sequential action of a phosphorylase followed by a phosphoribosyltransferase. Both of these salvage processes are simpler and much less energetically costly than *de novo* nucleotide synthesis.

Another area of interest concerning metabolism of translocated nucleosides is that regarding therapeutic analogues. In the majority of cases these only become active

when modified *in situ*. One such case is that of arabinosylcytidine which is phosphorylated by deoxycytidine kinase so producing the clinically active anti-cancer analogue [34]. Another such case is that of AZT used in the treatment of AIDS. This molecule is translocated prior to being phosphorylated by thymidine kinase which allows subsequent phosphorylation to the tri-phosphate form which is the active compound. This acts by two mechanisms, firstly it competes with endogenous nucleotides while secondarily acting as a chain terminator in the generation of proviral DNA [35].

1.2.3 Nucleoside transport and regulation

Understanding of the regulation of nucleoside transport is beginning to develop. Evidence now exists that factors including hormones, steroids, adenine nucleotides and key intracellular signaling pathways, including the protein kinase A and C systems, are important in the regulation of both facilitative and active nucleoside transport protein expression at the plasma membrane. In addition, cell growth and differentiation have also been shown to cause changes in the expression of these proteins, even showing the replacement of one transporter by another. Considering the effects of hormones first, it is understandable that such molecules should regulate more than one specific function depending on the physiological condition of the organism. This is exemplified by the transient and stable increase in sodium-dependent uridine uptake in liver parenchymal cells, caused by glucagon and insulin respectively [36]. The former of these phenomena involves the hyperpolarization of the cell membrane while the latter involves the synthesis and insertion of new transporters into the plasma membrane. Similar to this insulin effect thyroid hormone has been demonstrated to increase the number of *es*-type transporters present in the

cell membrane of cultured chromaffin cells by the synthesis and insertion of new transport protein molecules [37], (see subsequent section for classification of nucleoside transporter types). Interestingly this effect is antagonized by steroids and retinoic acid suggesting that these may act to return nucleoside transport to steady state levels. The nucleotide adenosine triphosphate (ATP) has also been established to increase *es*-type transport, thus indicating that ATP and possibly other nucleotides are capable of possessing regulatory roles in relation to nucleoside transport [38].

The role of intracellular signaling pathways in the regulation of nucleoside transport has recently been investigated, although the results obtained at present are only preliminary and require much more in-depth investigation. In chromaffin cells transport has been demonstrated to be under the control of protein kinase A or C whereas in bovine adrenomedullary endothelial cells transport was shown to be independent of such signaling mechanisms [39,40]. In the former of these activation of the kinases resulted in inhibition of transport which coincides with the observation that a cAMP-dependent protein kinase is involved in the reduction of *es*-type transport by phosphorylation of the transporter molecule or of a protein that interacts with it in cultured pig kidney cells [41].

Another factor effecting the regulation of the transporters is the stage of the cell within the cell cycle. Work carried out on quiescent S1 macrophages indicated that the majority of nucleoside transport is via the *ei* system but on the addition of colony-stimulating factor 1 (CSF-1) a 3-fold increase in activity was shown to be produced by *es*-type transport as the cells synchronously entered the S phase of the cell cycle [42]. Likewise *es*-type activity in haemopoietic cells and sodium-dependent nucleoside transport in the liver have been shown to increase in response

to rapid cell division [43-45]. In addition, differentiation of cells can also result in changes in levels of nucleoside transport observed while the actual type of transport itself varies. For example, in human HL60 leukemic cells induced to differentiate using dimethyl sulfoxide exposure to *N*-formyl-Met-Leu-Phe (FLMP) resulted in an increase in the V_{max} of *cib*-type nucleoside transport ~~proteins~~ present while parallel reductions in *ei*- and *es*-type transport were observed [46,47]. The fate of these transporters that disappear may be direct molecular modification of the transporters or translocation between the plasma membrane and a microsomal pool. The latter of these possibilities has been demonstrated using chromaffin cells, where *es*-type transporters were demonstrated to internalize with 50 to 60 percent being destroyed [48]. However the remaining transporters were later recycled but were once again internalized. The presence of cycloheximide, a protein synthesis inhibitor, resulted in a faster rate of internalization, suggesting that although originally internalization was occurring and a high percentage were being destroyed concomitantly synthesis of new transport proteins was occurring, thus preventing the total loss of this transporter from the cell.

1.2.4 Classification of nucleoside transporters

Until recently classification of nucleoside transport proteins was based solely on their functional and pharmacological characteristics with no indication of structural similarities. Based on these characteristics two obvious nucleoside transport classes have been observed, those which are equilibrative or concentrative. Each of these can be further subdivided into a number of more defined subclasses. The characteristics used to distinguish these nucleoside transport processes are, the ability to transport a non-metabolized nucleoside against its concentration gradient utilizing sodium

[†]The active nucleoside transport systems mentioned here refer to those systems found in mammalian cells. Other active nucleoside transporters exist in bacteria and protozoa which couple the translocation of substrate to that of the proton gradient [184].

gradients; sensitivity to inhibition by the nucleoside analogue nitrobenzylthioinosine (NBMPR); and the preference for purine and/ or pyrimidine nucleosides as permeants [49]. Generally, the facilitative transporters exhibit the aspects of any other equilibrative transport system in that transport appears to be bi-directional depending on which way the concentration gradient of the substrate is orientated. In contrast the concentrative transport systems use sodium ions as the driving force to allow the movement of the substrate from areas of low to higher concentration.† These differences in transport properties may explain why the equilibrative transporters are more widely dispersed than their concentrative counter-parts.

Equilibrative nucleoside transport falls into two subclasses based solely on the sensitivity to the nucleoside analogue NBMPR. Those transporters that are equilibrative and sensitive to ≤ 1 nM concentrations of this compound are designated *es* (*e*quilibrative and *s*ensitive) while transporters which are equilibrative and insensitive to NBMPR, only being inhibited by > 10 μM concentrations, are designated *ei* (*e*quilibrative and *i*nsensitive). Both of these subclasses have similar substrate specificities, transporting all of the natural nucleosides in addition to many analogues which have been modified both at the base and/ or sugar moiety levels. Each of these subclasses is widely distributed throughout different cells and tissues. By kinetic analyses they have been shown either independently or collectively to be present in erythrocytes, enterocytes, choroid plexus and cardiocytes to name but a few in addition to many cell lines [50-55]. More recently Griffiths *et al.* cloned the first of these transporters from a human placental cDNA library [56]. The cDNA obtained encoded a 456-residue glycoprotein with functional characteristics of an *es*-type transporter. Algorithmic analysis predicted this protein to have 11 membrane

spanning regions and to show significant homology to a number of yeast, nematode, plant and mammalian proteins of unknown function. This cloning opened up the gateway for further facilitative proteins to be cloned including that of the first *ei*-type transporter from the rat jejunum [57]. Although comparison of the sequences of the cloned *ei*- and *es*-type transporters indicated that they were approximately 50 percent identical it was clear that they were different proteins and not the same protein that existed in two differing conformations, a possibility suggested by Plagemann and Wohlhueter to explain the existence of *es* and *ei* transporters [58].

Concentrative nucleoside transport at present has been divided into five subclasses depending on the substrates transported. However, two differing nomenclatures for these subclasses have been employed. The first and most common system of nomenclature is based on the fact that these proteins are concentrative and substrate selective whereas the second system is numeric and based on the order of discovery. For the purposes of the current description the former system will be used with only a comparative summary of the numerical classification being included at the end. The first two subclasses characterised were the *cif*- and *cit*-type nucleoside transport proteins. The former of these showed preference for purine nucleosides transporting the analogue formycin-B as its model substrate (*c*oncentrative *i*nsensitive to inhibition by NBMPR capable of transporting *f*ormycin-B) [59-61] while the latter preferred pyrimidine nucleosides transporting thymidine as its model substrate (*c*oncentrative *i*nsensitive to inhibition by NBMPR capable of transporting *t*hymidine) [62-64]. However these subclasses showed overlapping specificity with each transporting adenosine and uridine. The third subclass showed broad substrate specificity, transporting both purine and pyrimidine nucleosides, thus being

designated *cib* (concentrative insensitive to inhibition by NBMPR and possessing broad permeant selectivity for both purine and pyrimidine nucleosides) [65,66]. The fourth type is a form of *cit*-type nucleoside transport with the difference that it also transports guanosine [67,68]. Within this form of classification no unique title has been given to this subclass so for the purpose of description it shall be termed *citg* (concentrative insensitive to inhibition by NBMPR capable of transporting thymidine and guanosine). The final and most recently identified subclass has the unique feature that it is a concentrative system inhibited by NBMPR and thus has been designated *cs* (concentrative and sensitive to NBMPR) [69,70]. The second method of classifying is solely concerned with active transport and has been designated N1 to N5 corresponding to *cif*, *cit*, *cib*, *citg* and *cs* respectively. A summary of the functional properties of both equilibrative and concentrative nucleoside transport processes can be seen in **Table 1.1**.

Of the concentrative nucleoside transporters those of the *cif*- and *cit*-type have been most extensively investigated. Vijayalakshmi and Belt in 1988 made the momentous step of identifying that there were at least two types of concentrative nucleoside transport subclass [62]. Using freshly isolated mouse intestinal epithelial cells they demonstrated that two populations of transporters were present, one group with characteristics of the *cif*-type subclass while the other showed features of the *cit*-type subclass. Since then each of these subclasses have been demonstrated in a number of different species in a variety of tissues. *cif*-type transport has been established in mouse enterocytes, splenocytes and activated peritoneal macrophages [60,62,71]; in rat kidney brush border membrane vesicles (BBMVs), hepatocytes, and macrophages [72,73]; in rabbit small intestinal BBMVs and renal BBMVs [64,74]. In comparison

cit-type transport has been observed in mouse intestinal enterocytes and activated peritoneal macrophages [60,62]; in rat renal BBMVs [75]; in bovine renal BBMVs [76]; in rabbit choroid plexus, small intestine BBMVs and renal BBMVs [52,64,77]; and in guinea pig enterocytes [78].

The distribution of the three remaining subclasses has proved more restricted. Jarvis in 1989 was the first to demonstrate *cib*-type transport activity while working with rabbit intestinal BBMVs [77]. Here purine and pyrimidine nucleosides were shown to be transported while obeying Michaelis-Menten kinetics corresponding to a single transport mechanism. Such a mechanism was later shown again by Huang *et al.* by *Xenopus* expression of rat intestinal mRNA isolates and Gutierrez *et al.* in human kidney BBMVs [79,80]. Reports describing the properties of the *citg*-type transport have been restricted to research carried out by the Gutierrez group working with human renal BBMVs and this type of transport has yet to be observed by other researchers [67,68]. Finally, *cs*-type nucleoside transport has only been exhibited in leukemic cells [69,70]. In addition to these tissues nucleoside transport has been observed in a number of cultured cell lines, for a more information see the review by Cass [49].

1.3 A novel gene family

1.3.1 Establishment of homologues

For many years active nucleoside transport had been studied using functional techniques. Although this did provide evidence for the existence of a number of different nucleoside transport subclasses it was impossible to confirm whether a transport

activity observed in one tissue was due to the same protein as that observed in another tissue, or indeed if that observed in one organism was identical to that observed in another. Therefore when the first active sodium/ pyrimidine nucleoside transporter, rCNT1, was cloned by Huang *et al.* in 1994 this allowed for a better understanding of these proteins [81]. Database analysis indicated that there were a number of proteins that showed significant homology. Of these the closest were a number of *E. coli* proteins including the recently cloned proton/ pyrimidine nucleoside transporter NupC [82] and two proteins of unknown function, YeiM and YeiJ. These proteins showed between 27 and 34 percent identity to the amino acid sequence of rCNT1. This finding showed that rCNT1 was a member of a new gene family, which has been verified by the further cloning of several more proteins including its human homologue from the human kidney, hCNT1 [83]; the rat liver sodium/ purine nucleoside transporter, formerly designated SPNT1 for *s*odium *p*urine selective *n*ucleoside *t*ransporter type-*1* but which has now been renamed rCNT2 to standardize nomenclature [84]; the human kidney transporter hCNT2 with similar characteristics to the rat liver homologue [85]; and more recently the pig kidney transporter pCNT1 which has similar characteristics to the two sodium/ pyrimidine transporters mentioned earlier. To date this is the total number of mammalian homologues detected although there are a number of homologous proteins in other organisms including the hagfish, hfCNT3. The importance of the latter discovery is that the hagfish is one of the evolutionally oldest vertebrates on the planet and thus indicates that members of this gene family have existed for millions of years. Moreover, this protein appears to be *cib-type* and has thus been designated as the first CNT3-type transporter to be cloned. Of the remaining homologues so far identified

two very similar *Caenorhabditis elegans* proteins bear close homology to the active mammalian nucleoside transporters while the remaining homologues, from bacteria, are relatively shorter, appearing to lack at least 130 amino acid residues from the *N*-terminus as indicated in **appendix 2**. Additionally a partial EST from Drosophila has demonstrated significant homology but has been omitted from this discussion due to the lack of complete sequence. A summary of the phylogenetic relationships between all known prokaryotic and eukaryotic concentrative nucleoside transport proteins can be seen in **Figure 1.1**.

1.3.2 Cloning of the first active nucleoside transporter: rCNT1

As mentioned above rCNT1 was the first active nucleoside transport protein to be cloned [81], thus confirming its identity and allowing more in-depth investigation into its structure and function. Cloning of this protein involved expressional screening in *Xenopus* oocytes to isolate a 2.4 kb cDNA clone encoding properties of a *cit*-type sodium dependent nucleoside transporter. That is, it transported primarily the pyrimidine nucleosides thymidine, cytidine and uridine in addition to the purine nucleoside adenosine. Uptake was saturable having an apparent K_m of 37 μM for uridine and a V_{max} of 21 pmol/ oocyte min^{-1}. Of pharmacological importance was the observation that the therapeutic nucleoside analogues AZT and ddC were also transported. In accordance with this being active transport rCNT1 was shown to couple sodium ion transport to nucleoside transport in a ratio of 1 to 1 as expected for *cit*-type transport.

The cDNA sequence predicted rCNT1 to be 648 amino acid residues in length giving it an unmodified molecular mass 71 kDa [81]. Further analysis indicated that there were three potential *N*-linked glycosylation sites at positions 543, 605 and 643; three

potential *O*-linked glycosylation sites, as predicted by the World Wide Web available application NetOglyc (web site of which is http://www.cbs.dtu.dk/), which compares the protein sequence of interest against a database which currently contains approximately 160 *O*-glycosylated proteins [86]; four potential protein kinase C-dependent phosphorylation sites at residues 5, 203, 421 and 527; a -S-S-S-S- motif at residues 609 to 612; and that the protein possessed a relatively high cysteine content of 3.1 percent. Fourteen potential transmembrane domains (TMDs) were identified using physiochemical (Goldman, Engelman, Steitz) and statistical (von Heijne) hydropathy scales [14,87], which in addition to the lack of an *N*-terminal signal sequence suggested that the *N*-terminus was cytoplasmic and assuming that all fourteen putative TMDs resided in the membrane this also placed the *C*-terminus on the cytoplasmic side of the membrane. However, using alternative algorithms allowed the number of TMDs to vary between 10 and 14 indicating that this may not be the true representation of the 2-dimensional model of this protein, which can only truly be determined by experimentation. The majority of this information has been represented in **Figure 1.2**.

In vivo expression of rCNT1 as indicated by Huang *et al.* was restricted to the rat jejunum, from where it was isolated, and the kidney as indicated by Northern blotting [81]. However, more recently mRNA for this transporter has been demonstrated to be widely distributed in the rat brain, being observed in superior colliculus, choroid plexus, striatum, hypothalamus, cerebellum, cortex and hippocampus. In this latter region, due to the low presence of *ei*- and *es*-type activity, it is predicted that rCNT1 and the co-expressed rCNT2 play major roles in nucleoside regulation [88].

Furthermore, Felipe *et al.* have also demonstrated the presence of rCNT1 in the rat liver [89].

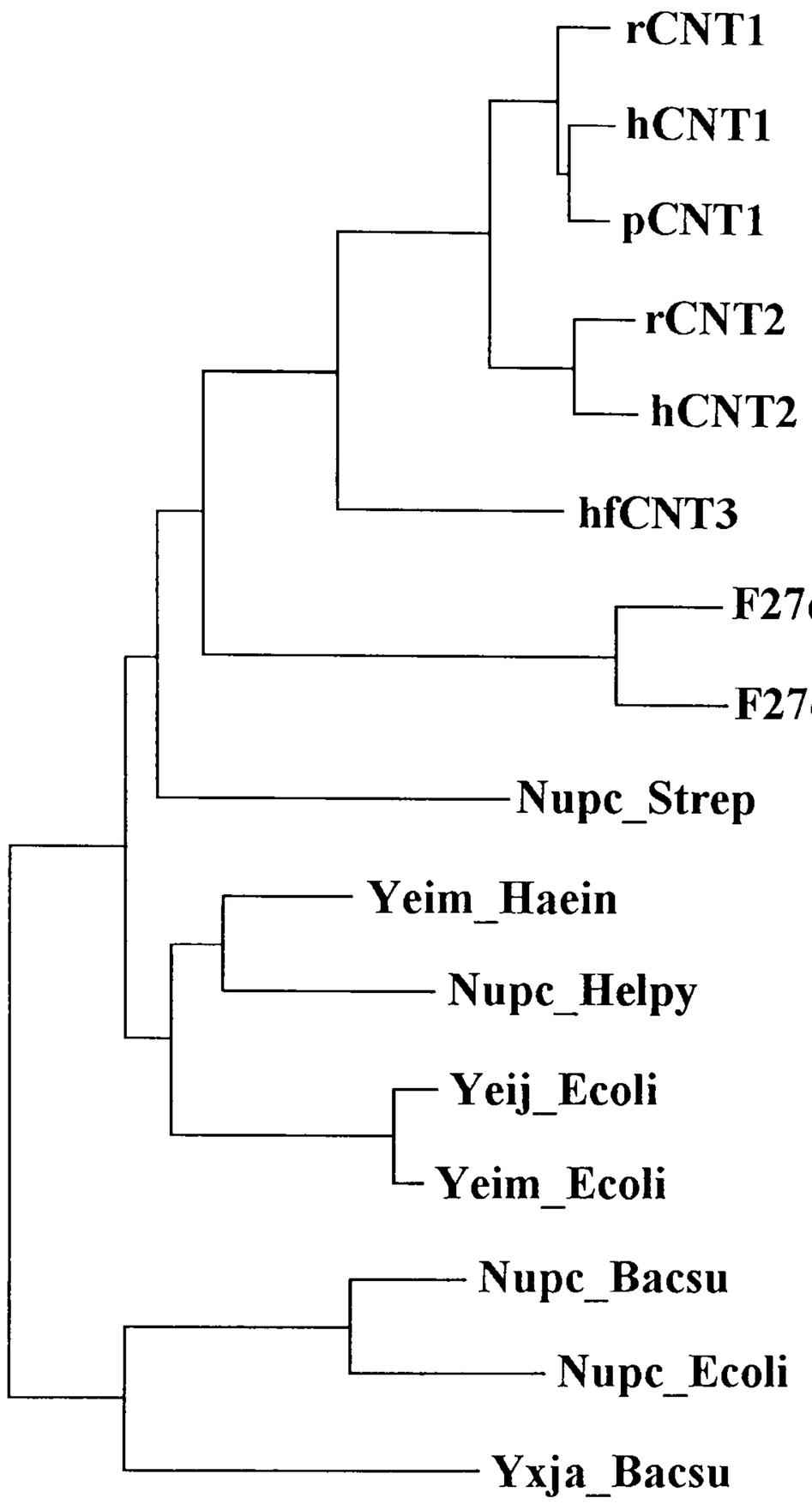
Table 1.1 Functional properties of nucleoside transporter subclasses

	Equilibrative		Concentrative				
Descriptive abbreviation	*es*	*ei*	*cif*	*cit*	*cib*	*citg*	*cs*
Numerical abbreviation			N1	N2	N3	N4	N5
Na^+-dependent	-	-	+	+	+	+	+
Na^+/ nucleoside stoichiometry			1:1	1:1	2:1	1:1	nd[a]
Inhibited by:							
NBMPR	+	-	-	-	-	-	+
Dipyridamole	+	+	-	-	-	-	+
Dilazep	+	+	-	-	-	-	+
Permeants:							
Adenosine	+	+	+	+	+	+	+
Uridine	+	+	+	+	+	+	nd
Guanosine	+	+	+	-	+	+	nd
Inosine	+	+	+	-	+	-	nd
Formycin B	+	+	+	-	+	-	+
Tubercidin	+	+	-	-	+	nd	nd
Thymidine	+	+	-	+	+	+	nd

[a] nd, not determined.

Table adapted from review by Cass [49]

Figure 1.1 Phylogenetic analysis of rCNT1 gene family

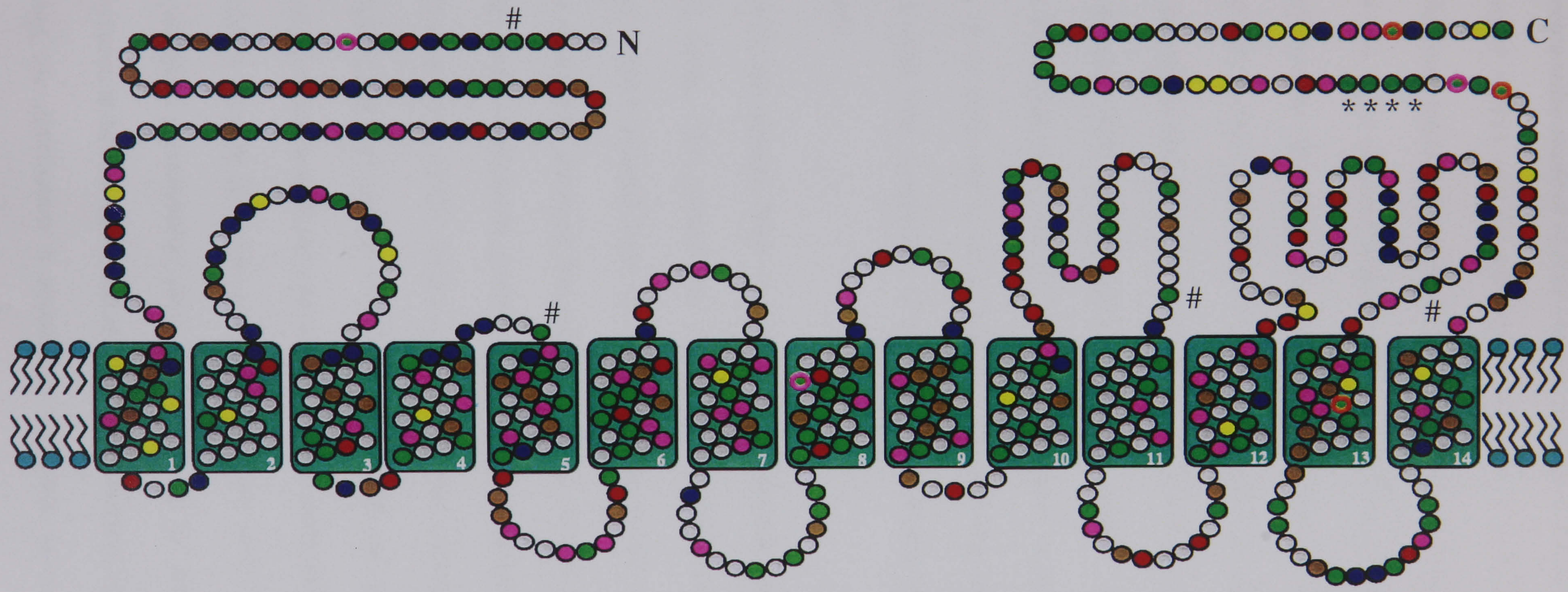


Phylogenetic analysis of the rCNT1 gene family members was performed to provide a multiple sequence alignment file using the software program ECLUSTALW which was subsequently analyzed to obtain evolutionary distance information in the GCG DISTANCES program, based upon the method of Kimura [90]. This information was subsequently interpreted by the GCG program GROWTREE, based the original method of Saitou and Nei [91], to produce the graphical representation observed above, where the horizontal axis represents sequence differences and thus evolutionary distances. Accession numbers for each member are provided in **appendix 1**. The original alignment output obtained was further modified to include further observations and analyses, and is represented in **appendix 2.**

Figure 1.2 Putative model of rCNT1 secondary structure in the lipid bilayer

Image on opposite page

Putative topological model of rCNT1 in the lipid bilayer, illustrating the presence of fourteen potential TMDs as predicted by the physio-chemical Goldman, Engelman and Steitz analysis and von Heijne statistical hydropathy scales [14,87]. The lack of an *N*-terminal signal sequence suggested that the *N*-terminus was cytoplasmic and assuming that all fourteen putative TMDs resided in the membrane, this also placed the *C*-terminus on the cytoplasmic side of the membrane. The amino acid residues in the model have been colour coded based on the classification used by SOMAP [92]; general non-polar residues are A, M, I, L & V (*grey*); general polar residues are S, T, N & Q (*green*); positive residues are H, K & R (*blue*); negative residues are D & E (*red*); aromatic residues are F, W & Y (*purple*); structure breaking residues are G & P (*brown*); and cysteine residues (*yellow*). Also shown in the model are the three potential *N*-linked glycosylation sites at positions 543, 605 and 643 although the former of these is located within predicted helix thirteen (*residues with orange borders*); the three potential *O*-linked glycosylation sites at S13, T320 and T607 (*residues with purple borders*); the four potential protein kinase C-dependent phosphorylation sites at T5, S203, S421 and S527 (*indicated by* #); and the four serine motif (*indicated by* ****).



Key:

- General non-polar
- General polar
- Positive
- Negative
- Aromatic
- Cysteine
- Structure breaker
- **** Serine motif
- Putative *N*-linked glycosylation site
- Putative *O*-linked glycosylation site
- # Putative protein kinase C phosphorylation site

1.3.3 A comparative study of rCNT1 and the mammalian homologues

The cloning of rCNT1 and the subsequent isolation of a cDNA encoding rCNT2 allowed additional nucleoside transport proteins within these two subclasses, *cit* and *cif* respectively, to be obtained with greater ease with the aid of degenerate primers and the polymerase chain reaction (PCR). As a result the cDNA sequences of hCNT1, hCNT2, pCNT1 and hfCNT3 have been isolated in addition to some of the other proteins mentioned in **section 1.3.1**. However little is known about the majority of these proteins with the exception of rCNT1, hCNT1, rCNT2 and hCNT2. Therefore discussion about rCNT1 and its homologues in the earlier part of this section is to be restricted to these four proteins with any exceptions specifically mentioned while later comparison of sequences will involve all the fore-mentioned homologues.

Since rCNT1 has already been discussed in the previous section rCNT2 will be considered firstly. This protein was cloned by functional expression in *Xenopus* oocytes of mRNA isolated from the rat liver. Like rCNT1 it exhibited sodium-dependent nucleoside transport but unlike rCNT1 it showed preference for purine nucleosides while additionally transporting uridine. The apparent K_m for adenosine was 6 μM with a V_{max} of 457 fmol/ oocyte min^{-1} in this expression system. The size of this protein was very similar to that of rCNT1 with the cDNA encoding a 659 amino acid polypeptide with the unmodified relative molecular mass of 72 kDa and showing 64 % identity in amino acid sequence. Like rCNT1 it has a number of potential *N*-linked glycosylation sites, consensus sites for protein kinase A and C phosphorylation in both termini, and an ATP/ GTP binding motif in the *N*-terminus. Considering the distribution it appears in relatively more tissues having been

demonstrated in the jejunum, spleen, heart and brain in addition to the liver [84,88]. Indeed, although both rCNT1 and rCNT2 have been isolated from the same tissues it is obvious from evidence to date that the levels of expression of each are not equivalent and that their specificities differ. For example, comparative studies on hepatic nucleoside transport by Felipe *et al.* indicated that rCNT2 is greatly up-regulated in response to insulin in regenerating rat livers, while the amount of rCNT1 remained relatively constant [89]. This provides evidence for the suggestion that rCNT2 was responsible for the uridine uptake observed in plasma membrane vesicles during regeneration of rat livers [45].

Having cloned the rat transporters rCNT1 and rCNT2 the human homologues could be obtained with relative ease. Both hCNT1 and hCNT2 were cloned by hybridization cloning and reverse transcriptase PCR amplification strategies. The former has been demonstrated to be 83 % identical in amino acid sequence to the rat isoform showing similar substrate specificities in that it was *cit*-type and transported AZT [83]. The cloning of this protein from the human kidney indicated that members of the CNT family existed in human tissues and provided evidence for their involvement in renal transport of physiological nucleosides and nucleoside drugs. Similarly the human homologue of rCNT2, hCNT2, showed 81 % identity to the rat isoform but differed markedly at the *N*-terminus in primary structure. Due to the specificities of the proteins being identical it is foreseeable that this difference may be a means of targeting the protein to a specific membrane surface and that in the two organisms this is achieved by different means. Alternatively, this may be a region of little functional importance which is consequently not conserved during evolution. Interestingly on comparison of the aligned sequences of rCNT1, rCNT2, hCNT1 and

hCNT2 in **appendix 2**, although the primary structure is variable there is the trait of the *N*-terminus being very hydrophilic with approximately 25 percent of the residues being either acidic or basic with a net charge of between plus 2 to plus 5. Therefore it is possible that relative conservation of these charged residues may play an important factor in protein activity or expression. Considering the remaining four transporters in the alignment with the extra *N*-terminal region, when compared to the facilitative homologues, the pig kidney CNT1 homologue, pCNT1, follows much the same pattern of charged residues in contrast to the hagfish transporter which is uniquely acidic (data not shown for hfCNT3 in alignment as this sequence was a personal communication from J. D. Young and has not yet been entered into Genbank or published). The final two clones come from *C. elegans* coding for two nucleoside transporters with clone F27e11_2 encoding an extra 79 amino acid residues at the *N*-terminus compared to clone F27e11_1. Interestingly both clones conform to the net positive terminus similar to all the other fore-mentioned homologues, with the exception of the hagfish protein.

Comparison of the shorter homologues with the concentrative nucleoside transporters suggested that the extra three helices located at the *N*-terminus of the concentrative transporters may be responsible for the co-transport of the sodium ion. Such localized sodium binding and transport activity has been shown previously with the *E. coli* melibiose transport protein which possesses four aspartate residues shared between helices one, two and four which when mutated decrease if not abolish sodium coupled transport [93]. However this does not appear to be the case with the active nucleoside transporters for two reasons. Firstly, on investigation for similar residues in the nucleoside transporter alignment only an aspartate residue within predicted

TMD 3 was conserved throughout the larger of the homologues. Secondly, bacterial transporters, lacking these helices, do nonetheless transport cations. Therefore sodium coupling must occur within the final eleven putative helices of this family of concentrative nucleoside transporters. Possible locations for the site(s) of sodium ion interaction could include regions containing either histidine or cysteine residues, which have been reported to participate in metal-protein interactions. A sodium-binding (SOB) motif, G-A-X-X-X-X-L-X-X-X-G-R where X is any amino acid, has been demonstrated to participate in sodium binding in glutamate, taurocholate and phosphate cotransporters [94]. However, examination of the concentrative transporter sequences indicated that this sequence was absent and therefore sodium-coupling cannot be via this motif.

Examination of the aligned sequences for potential residues involved in substrate binding indicated the presence of a number highly conserved single residues or motifs. These motifs were considered significant when nine or more of the fifteen sequences contained the same or similar residues. The first significant observation is the relatively high conservation of glycine residues throughout the sequences. Nineteen of these have been identified in the alignment suggesting that their presence is significantly important, these residues have been highlighted in **appendix 2**. Moreover twelve out of the nineteen are located within predicted helices suggesting that their presence may be necessary to allow correct folding. Also of interest is the absolute conservation of a proline residue within putative TMD twelve, and the prevailing presence of other prolines within putative helix ten and towards the end of putative TMDs eight and ten. Unlike soluble proteins where prolines are rarely found in α-helices a number of integral membrane proteins have been shown to possess this

residue within these secondary conformations. For example, bacteriorhodopsin has proline residues in the middle of three of its putative seven TMDs while rhodopsin has prolines in five of its seven putative helices. Other studies have also indicated a bias toward prolines in the presumed membrane-spanning segments of membrane proteins, especially those involved in transport functions [95,96]. One suggestion is that proline residues may act as a hinge, thus allowing change of the protein from outward facing to inward facing or vice-versa, depending on the local conditions including substrate binding. However, this suggestion has tended to be discounted as a result of mutagenesis studies, exemplified by GLUT1 [97], which have demonstrated such proline residues to be non-essential.

The cluster of basic residues located at the *C*-terminal end of putative TMD four appears important. In all cases at least two basic residues are located within close proximity of each other, with at least one lysine residue always being present. These residues may therefore be important in either substrate binding or in signaling the end of this helix. Further conservation of either acidic or basic residues can be seen throughout the sequences although to a higher degree around TMD eight and in the hydrophilic loops between TMDs 8 & 9, 10 & 11,12 & 13 and 13 & 14. These therefore may play a similar role in substrate recognition or structure determination.

Apart from single residues being conserved a number of short sequences appear important. The first of these is located in putative helix six encoding the consensus G-X-S-F-V-F-G-X-A-L followed almost immediately by F-A-F-Q-V-L-P-I-I-V-F-F-S-X-V-M-S-V-L-Y-Y-L-G, where X can be any residue, suggesting that this region may be important as this is a very high level conservation and in those regions where X has been displayed it normally contains an amino acid of similar physio-chemical

properties throughout the aligned molecules. Inspection of further regions of the alignment indicated that areas *C*-terminal to this contained relatively high conservation, particularly within the predicted TMDs. Taking into consideration the more continuous of these indicated the G-X-A-T-I-A-G-S-L-L-G-A-Y sequence within putative helix nine; A-A-X-L-I-A-F-L-A-V-L-A-L-V-N in potential TMD eleven; the conservation of G-X-K-X-X-X-N-E-F-V-A-Y in the hydrophilic loop connecting TMDs twelve and thirteen; and T-F-A-L-C-G-F-A-N-F-S-S-I-G-I-M-L-G in the thirteenth TMD. Although all of these sequences may play important roles in structure or substrate binding the absolute conservation of the N-E-F-V sequence is particularly interesting because it is found within a predicted hydrophilic loop. As a general concept of membrane proteins the hydrophilic regions tend to be very variable between homologues unless that region is of particular importance. This information in addition to the evidence discussed in **chapter 5**, indicating that this loop is provisionally extracellular, suggests that these residues may act in initial substrate recognition or form a predominant part in regulating access of the substrate to the substrate binding site.

Although the majority of the discussion in this sub-section is purely speculative based on sequence similarity it provides suitable locations to start the mutagenic investigation of this group of proteins to determine features that are both important in conformation and substrate recognition.

1.4 Aims of the study

The objectives of this study were two-fold in that distribution of the transporter rCNT1 in rat tissues was to be investigated while performing parallel investigations

into the conformation of this protein in the lipid bilayer. A means that would allow the specific detection of rCNT1 would be the production of polyclonal antibodies raised against this protein. To achieve this two strategies were adopted, principally this involved the generation of four synthetic site-directed peptides to predicted hydrophilic regions of the protein, which were be used for immunization following conjugation. Alternatively genetic manipulation of the rCNT1 cDNA allowed the production of a glutathione S-transferase (GST)-fusion protein encoding a fragment from the *N*-terminus of this former protein, which was then be used in immunization. These anti-peptide antibodies enabled the detection of rCNT1 expressed at the protein level in rat tissues to confirm and extend the original data published by Huang *et al* [81]. Furthermore, this was be taken a stage further by the immunocytochemical investigation of these tissues to localize the protein at the subcellular level, and thus enhance the understanding of the physiological role of this protein *in vivo*. The second objective of this study was to investigate the basic topology of rCNT1. This was to be achieved by initially investigating the glycosylation state of the native protein, followed by genetic manipulation of the cDNA to remove any accessible *in vivo* *N*-linked glycosylation sites to allow further analysis of the protein conformation by introduction of artificial glycosylation sites at defined locations within the peptide sequence.

Chapter 2

Materials and methods

2.1 Materials and Suppliers

Analytical grade chemicals were used throughout in this study with all solutions being prepared using deionised water. Materials were obtained from the following suppliers:

General laboratory reagents and equipment

Amersham International plc, Amersham, Buckinghamshire

Aldrich Chemical Company Ltd, Gillingham, Dorset

BDH Laboratory Supplies Merck Ltd, Lutterworth, Leicestershire

Bio-Rad Microscience Ltd, Hemel Hempstead, Hertfordshire

Pierce and Warriner (UK) Ltd, Chester, Cheshire

Pharmacia Biotech Ltd, St. Albans, Hertfordshire

Sigma Chemicals Co., Poole, Dorset

Peptide Synthesis

Calbiochem-Novabiochem (UK) Ltd, Beeston, Nottingham

Fisons Ltd., Loughborough, Liecestershire.

Fluka Chemicals Ltd, Glossop, Derbyshire

PerSeptive Biosystems (UK) Ltd, Hertford, Hertfordshire

Immunological reagents

Boehringer-Mannheim, Lewis, East Sussex

Sigma Chemicals Co., Poole, Dorset

Vector laboratories Inc., Burlingame, CA, USA

Molecular biology reagents

Invitrogen, British Bio-technology Ltd, Abingdon, Oxon

MBI Fermentas, Immunogen International Ltd, Sunderland, Tyne and Wear

New England Biolabs (UK) Ltd, Hitchin, Hertfordshire

Promega UK, Southhampton

Qiagen, Hybaid Ltd, Teddington, Middlesex

Stratagene Ltd, Cambridge

2.2 Synthetic peptide synthesis and analysis

2.2.1 Peptide synthesis

Site-directed peptides were synthesised on a Milligen/ Biosearch 9050 PepSynthesiser according to the diisopropylcarbodiimide (DIPCDI) method described in the operator's manual. Syntheses were performed with 1 g polyethylene glycol-polystyrene (PEG-PS) resin with a 4-fold excess of fluorenylmethyloxycarbonyl (FMOC) amino acid derivatives (PerSeptive Biosystems). Preformed FMOC amino acid-pentafluorophenol (Pfp) esters were used in all cases with the exceptions of Ser (t-butyl, tBu) and Thr (tBu) when dihydrohydroxyoxobenzotriazole (Dhbt) esters were utilised. Efficiency of coupling was monitored by the counterion distribution monitoring, CDM™, system (**section 3.3.1**).

2.2.2 Peptide cleavage

Resin-bound peptides were washed 5 times with dichloromethane (DCM), four times with HPLC-grade methanol and then lyophilised. Cleavage of the peptide from the resin and removal of protecting groups was carried out using trifluoroacetic acid (TFA) containing a mixture of scavenger reagents (81.5 % (v/v) TFA, 5 % (v/v) phenol, 5 % (v/v) water, 2.5 % (v/v) 1,2-ethanedithiol (EDT), 5 % (v/v) thioanisole, 1 % (v/v) triisopropylsilane (TIS); the latter three being obtained from Fluka chemicals) for 4 h at room temperature (r.t.) with shaking every 30 min. The cleavage mixture was separated from the resin by filtration and the resin then washed with 60 ml TFA, with the peptide/ TFA/ scavengers being collected in a round bottom flask. The TFA was removed by rotary evaporation using a dry ice/ ethanol cold finger trap. Scavengers were removed by ether precipitation following which the

peptides were dissolved in Milli-Q water or 2 % (v/v) dimethylsulfoxide (DMSO), lyophilised and finally stored at -70°C.

2.2.3 Peptide mass spectrometric analysis

Peptide fractions for analysis were diluted to 20 pmol/µl in water or 2 % (v/v) DMSO. 1 µl of a saturated solution of α-cyano-4-hydroxy-cinnamic acid (Aldrich) in 60 % (v/v) water / 40 % (v/v) acetonitrile and 1 µl of diluted peptide were mixed and air-dried (approx. 30 min) onto steel discs pre-washed with water and methanol. Measurements were performed by matrix-assisted laser desorption ionisation time-of-flight (TOF) mass spectrometry (MS) using a TOFspec mass spectrometer (VG Organics) by Dr J. N. Keene. Where necessary, the quantities of peptides used were adjusted to improve signal-to-noise ratio.

2.2.4 Peptide high performance liquid chromatography

Peptides were solubilised at 0.5 mg/ml by vortexing in either water or 2 % DMSO in water. Debris was removed at 11600*g* for 2 min and the supernatant transferred to a fresh tube. Reverse-phase high performance liquid chromatography (rpHPLC) was performed using a Dynamax 300Å, 5 µm, C18 analytical column attached to a Gilson 715 HPLC apparatus. Sample was applied using a 20 µl sample loop. Two reservoirs were used to create the elution gradient, reservoir A contained 0.1 % trifluoroacetic acid (TFA, from Fisons) in deionised water while reservoir B contained 0.1 % TFA in HPLC far UV grade acetonitrile (Fisons). Initial analysis was achieved by running a gradient from 0 to 40 % reservoir B solution against reservoir A. After each sample was analysed a control run was carried out using the solution in which the peptide was resuspended. This was subtracted from the sample profile to eliminate the

background signal. If required the gradient was altered to allow greater resolution over the region at which the peptide was eluted.

2.3 Production and characterisation of anti-peptide antibodies

2.3.1 Conjugation of peptides to ovalbumin

Peptides were conjugated to ovalbumin with m-maleimidobenzoyl-sulfo-succinimide ester using a method adapted from LaRochelle *et al* [98] with the exception that ovalbumin was used as the carrier protein in place of Keyhole limpet haemocyanin, which tends to come out of solution readily. Each conjugate (200µg at 1 mg/ml in 10 mM sodium phosphate, 145 mM NaCl, pH 7.2; PBS) was mixed with 200µl complete Freund's adjuvant and subcutaneously injected into New Zealand white laboratory rabbits. The rabbits were given a second injection of 100 µg conjugate in incomplete Freund's adjuvant after 4 weeks and serum samples analysed 2 weeks later by enzyme-linked immunosorbent assays (ELISAs), **section 2.3.3**. Boosting continued for a further four months if necessary, at intervals of not less than 1 month, after which the rabbits were bled out under terminal anaesthesia. Cross-reactivity tests of each antiserum with each of the other three peptides was carried out by ELISA. Peptide-specific antibodies were then affinity purified from each antiserum using peptide columns as described in **section 2.3.2**.

2.3.2 Affinity purification of anti-peptide antibodies using immobilised synthetic peptides

Peptide affinity columns were prepared by dissolving 6 mg peptide in 500 µl 50 mM Tris-HCl, 5 mM ethylenediaminetetraacetic acid (EDTA) pH8.5 plus a minimal amount of DMSO, if required, to solubilise the peptide. Any disulfide bonds in the

mixture were then reductively cleaved by addition of 25 µl 1 M dithiothreitol (DTT), flushing with nitrogen and incubating for 1 h at r.t.. The reduced peptide was then separated from DTT by passage down a 1 x 20 cm Sephadex G10 column (Pharmacia) pre-equilibrated in Tris-HCl/ EDTA pH 8.5. The absorbance of fractions was measured at 230 nm and the fractions corresponding to the peak eluting at the void volume were pooled (3-4 ml). Duplicate 50 µl samples of this pooled sample were taken to determine coupling efficiency at a later stage by the dithiobisnitrobenzoic acid (DTNB) assay [99]. Meanwhile 6 ml Sulfo-link coupling gel (Pierce) was washed with 6 volumes Tris/EDTA buffer, composition as described above. The solution of reduced peptide was added to the Sulfo-link gel and incubated at r.t. for 25 min, rotating slowly in the dark, followed by incubation for 30 min without rotation. The mixture was centrifuged as briefly as possible in a clinical centrifuge and the supernatant removed, keeping duplicate 100 µl samples of the latter to determine coupling efficiency by the DTNB assay. Unbound peptide was removed by resuspending the gel in Tris/EDTA buffer and centrifuging as previously. After this procedure had been repeated twice more, excess iodoacetyl groups on the coupling gel were blocked by the addition of 5 ml 50 mM cysteine in Tris/EDTA and then rotated slowly for 25 min at r.t. followed by incubation for 30 min without rotation. After brief centrifugation as before, the supernatant was removed and the gel washed with 50 ml 1 M NaCl. The gel was equilibrated with PBS by washing with 50 ml of this buffer, centrifuged to remove most of the supernatant and loaded into a 5 ml econo-column (Bio-Rad). PBS was then passed through the column at a flow rate of 15 ml/h for 1 h. For long term storage columns were equilibrated in PBS containing 0.02 % sodium azide.

Purification of antisera samples using the affinity columns involved the addition of 5-10 ml antiserum to the PBS pre-equilibrated column at a flow rate of 10 ml/h and measuring the absorbance of the eluate at 280nm (A_{280nm}). After circulation of the serum through the column for 1-2 h non-specific and weakly-bound antibody was removed by elution with 10 mM sodium phosphate, 800 mM NaCl, pH 7.2 until the A_{280nm} of the eluate fell to near zero, with 2 ml fractions being collected. Tightly-bound, peptide-specific antibodies were then eluted with 50 mM diethylamine pH 11. The peak fractions (4-6 ml) were pooled, rapidly neutralised by addition of a small volume of 0.5 M acetic acid, and then dialysed overnight (o/n) versus 4 l PBS. Subsequent to alkaline elution, acid elution of the column with 0.2 M glycine-HCl pH 2.4 was also carried out and the peak fractions pooled and neutralized with 2 M Tris prior to dialysis o/n as described for alkaline elution. Following dialysis bicinchoninic acid (BCA) protein determination was carried out on each fraction, in addition to ELISA to determine recovery efficiency.

2.3.3 Enzyme-linked immunosorbent assay

ELISAs were performed by placing 80 µl 0.5 µg/ml peptide, in 50 mM sodium carbonate pH 9.6, into each well of a 96-well microtitre plate and incubating o/n at 4°C . The plate was washed 5 times with phosphate-buffered saline Tween-20, PBS-T, (140 mM NaCl, 5 mM KCl, 3 mM Na_2HPO_4, 3 mM KH_2PO_4, 6 mM NaN_3, 0.05 % (v/v) Tween-20, pH 7.2) and then blocked for 2 h at 37°C with blocking buffer (5 % (w/v) milk powder in PBS-T). After washing the plates 5 times with PBS-T, 100 µ l samples of primary antibody were added to the wells at successive 3-fold dilutions, starting with a 200-fold diluted stock in antibody buffer (1 % milk powder in PBS-T), and incubated o/n at 37°C . The plate was washed as before and 100 µl of goat

anti-rabbit IgG conjugated to alkaline phosphatase (Bio-Rad), diluted 3000-fold in antibody buffer, was added to each well for 2 h at 37°C. After washing the plate as before 100 µl *p*-nitrophenol phosphate (PNPP) solution (5 mM *p*-nitrophenyl phosphate disodium salt in 10 mM diethanolamine, 1 mM $MgCl_2$, pH 9.8) was added and incubation continued for 2 h at r.t.. Finally, the absorbance of each well at 405 nm was measured using a Titretex plate reader.

2.4 General Protein techniques

2.4.1 Membrane preparation

Brain, heart, jejunum, kidney, liver, muscle, spleen and testis were obtained from male Wistar rats (200-250 g). Membranes were prepared by chopping the tissues to allow homogenization with a Kinematica CH-6010 polytron at setting 3 for 2-4 min in 4 volumes 250 mM sucrose, 100 mM sodium phosphate pH 7.5, 1 mM EDTA, 2 µg/ml aprotonin, 1 mM benzamidine, 10 µM leupeptin, 0.1 mM phenylmethylsulfonyl fluoride, 10 µg/ml pepstatin A. Homogenates were centrifuged at 4°C for 30 min at 7500*g*. The supernatants were passed through 150 µm nylon mesh prior to centrifuging at 100,000*g* for 1 h at 4 °C . Supernatants were removed and the pellets were resuspended in phosphate buffered saline pH 7.2 (PBS), prior to homogenization by 10 strokes of a Teflon pestle. Samples were stored at -70°C until required.

2.4.2 Protein determination

Protein concentrations were determined with the BCA assay (Pierce) according to the manufacturer's instructions, using BSA as a standard.

2.4.3 SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli [100]. 1.5 mm 10 % (w/v) resolving gels were run on vertical mini gel apparatus for 1 h at 30 mA per gel. All samples were diluted 4:1 in SDS-sample buffer (0.2 M Tris-HCl pH 6.8, 4 mM EDTA, 4 % (w/v) SDS, 4 % (v/v) glycerol, 40 mM DTT). The amount of protein present varied depending on the source of the sample, therefore the protein content will be stated in each case. Gels were stained with Coomassie blue (0.25 % (w/v) coomassie brilliant blue in 25 % methanol/ 10 % acetic acid) and destained with 10 % (v/v) acetic acid.

2.4.4 Immunoblotting

Transfer to nitrocellulose membranes was preformed as described by Towbin [101]. Membranes were washed in tris buffered saline (TBS, 20 mM Tris, 500 mM NaCl pH 7.5) for 10 min prior to blocking with 5 % milk powder in TBS containing 0.2 % Tween-20 (TBS-T) for 2 h at room temperature (r.t.). Primary antibody (2 μg/ml) in antibody buffer, TBS-T containing 1 % milk powder, was added overnight at r.t.. Membranes were washed 3 times in TBS-T after which goat anti-rabbit peroxidase labelled secondary antibody, from Boehringer-Mannheim, in antibody buffer was added at 1:10,000 for 1 h at r.t.. The nitrocellulose was washed 3 times in TBS-T and peroxidase activity detected using Enhanced Chemical Luminescence (Amersham), according to the manufacturer's instructions.

2.5 Molecular biology techniques

The procedures used for DNA manipulations were essentially as described by Sambrook *et al* [102] with alterations mentioned where appropriate. Plasticware,

glassware and media were sterilised by autoclaving at 121°C for 15 min. Chemicals were made from molecular biology grade reagents and either autoclaved or filter sterilised prior to use.

2.5.1 Small scale isolation of plasmid DNA

Plasmid DNA was prepared from the host strain based on a modified form of the alkali lysis method of Ish-Horowicz and Burke [103]. Single colonies were taken from Luria-Bertani (LB) agar (1 % (w/v) NaCl, 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1.5 % (w/v) bacterial agar, pH 7) plates containing the appropriate antibiotics and used to inoculate 5 ml LB broth containing the correct antibiotics in 50 ml Falcon tubes. These cultures were propagated overnight at 37°C with shaking at 200 rpm. The following day cultures were sedimented at 1000*g* for 5 min and the pellets resuspended in 250 µl GET buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0). The cells were lysed by addition of 250 µl lysis buffer (0.2 M NaOH, 1 % (w/v) SDS) with gentle inversion 15 times followed by neutralisation with 350 µl neutralising buffer (3 M NaOAc pH 4.8) after which samples were left on ice for 5 min. Protein and chromosomal DNA precipitates were removed by transferring samples to 1.5 ml eppendorfs followed by sedimentation at 11600*g* for 10 min in a bench-top microfuge. The supernatant was removed to a fresh eppendorf to which 750 µl 96 % ethanol was added to allow precipitation of plasmid DNA while standing on ice for 10 min. The DNA was sedimented at 11600*g* for 30 min and the pellet washed with 70 % (v/v) ethanol re-pelleted for 5 min and air-dried. The DNA was finally resuspended in 50 µl TE (10 mM Tris-HCl pH 8, 1 mM ethylenediamine tetraacetic acid (EDTA)) containing 50 µg/ml RNase A incubated at

37°C for 30 min and stored at -20°C. Yields from 5 ml overnight cultures ranged from 0.5 to 5 μg depending on the plasmid copy number.

Production of ultra-pure plasmid DNA for automated DNA sequencing was performed by culturing the cells as described above followed by plasmid DNA extraction using the QIAGEN plasmid mini-prep kit which is also based on alkali lysis.

2.5.2 Large scale isolation of plasmid DNA

A large scale culture was prepared using a 10 ml overnight culture in LB broth, containing the appropriate antibiotics, to inoculate 1 L of the same medium followed by o/n incubation at 37°C with agitation at 220 rpm. Cells were harvested at 3500*g* for 10 min. The pellet was resuspended in 50 ml GET buffer and the cells lysed by the addition of 100 ml lysis buffer followed by inversion 10 times. The solution was neutralised by addition of 50 ml 3 M potassium acetate, 5 M acetic acid and inverted as before. Cell debris was removed at 3500*g* for 30 min at 4°C with the supernatant being passed through a 0.45 μm sterile filter prior to DNA precipitation by mixing with 0.6 volumes isopropanol. DNA and RNA were pelleted at 3500*g* for 30 min at 4°C, washed with 20 ml 70 % (v/v) ethanol and air-dried. The pellet was resuspended in 4.5 ml TE, transferred to a 50ml Falcon tube to which 6 ml 5 M LiCl was added and left on ice for 10 min. RNA was pelleted at 3500*g* for 10 min at 4°C and the supernatant transferred to a 50 ml Falcon tube containing 2.5 volumes 96 % ethanol and incubated on ice for 30 min. Plasmid DNA was pelleted at 3500*g* for 30 min at 4°C, washed with 5 ml 70 % ethanol and air-dried. The pellet was resuspended in 800 μl TE containing 50 μg/ml RNase A, transferred to an eppendorf and incubated at 37°C for 30 min. 400 μl 20 % (w/v) polyethylene glycol 6000, 2.5 M NaCl was

added and incubated on ice for 15 min. Plasmid DNA was pelleted at 11600*g* for 15 min at 4°C and resuspended in 500 µl TE. The DNA was further cleaned by phenol/ chloroform extraction as described in Sambrook *et al* [102] with the pellet being resuspended in 500 µl TE following ethanol precipitation and stored at -20°C.

2.5.3 Restriction analysis and agarose gel electrophoresis of plasmid DNA

DNA (100 to 500 ng) was digested with 0.5-2 U of the appropriate restriction endonuclease in its corresponding buffer, as recommended by the supplier, with the addition of 100 µg/ml BSA for 3 h in a final volume of 10 µl. 2 µl of 6x DNA sample buffer (40 % (w/v) sucrose in Milli-Q containing 0.1 % bromophenol blue) was then added and 6 µl of each sample was subjected to 0.7 % agarose slab gel electrophoresis at 80 V using a 1x TAE (40 mM Tris-acetate pH 8.3, 1 mM EDTA) buffer system. For determining the molecular size of restriction products either Lambda/ *Bst*EII or Phi X147/ *Hae*III DNA markers were run along side the samples.

Quantitation of plasmid DNA was carried out as for restriction analysis but with serial dilutions 1: 2, 1: 4, 1: 8 and 1: 16 of the DNA to allow visual comparison with 250 ng of the fore-mentioned DNA markers.

2.5.4 Preparation and transformation of competent bacterial cells with plasmid DNA

Competent *E. coli* were produced by a procedure based on the $CaCl_2$ method of Hanahan (1983) [104]. LB (5 ml) overnight cultures grown at 37°C were used to inoculate 250 ml LB the following morning. This culture was incubated at 37°C with agitation at 220 rpm until A_{600nm} = 0.4-0.5 for *RecA*$^-$ strains. The cells were then incubated on ice for 10 min, and all subsequent stages performed at 4°C. Cells were sedimented at 2500*g* for 10 min followed by gentle resuspension in 125 ml sterile,

ice-cold 100 mM $CaCl_2$ and incubated on ice for 10 min. The cells were re-pelleted as before, gently resuspended in 10 ml 100 mM $CaCl_2$ and 400 µl aliquots prepared in eppendorfs. Aliquots were incubated at 4°C for 2 h prior to use or snap-freezing in liquid nitrogen.

Competent cells (100 µl) thawed on ice or freshly prepared cells were transformed by the addition of 40-200 ng plasmid DNA, including pBluescript control (Promega) to test competency. Following the addition of DNA the cells were incubated on ice for 30 min in prechilled 15 ml Falcon tubes after which they were heat shocked at 42°C for 45 sec. Pre-warmed LB (0.9 ml, 37°C) was added to each tube which was incubated at 37°C with agitation for 2 h. Cells were pelleted 2500*g* for 10 min, 900 µl removed and the pellet gently resuspended in the remaining 100 µl LB which was spread onto an LB agar plate containing the appropriate antibiotics. Plates were incubated overnight at 37°C and colonies counted the following day. Typically DH5α *E. coli* gave 10^7 colony forming units (cfu) per µg of pBluescript while recombinant DNA tended to produce ten-fold lower cfu/ µg of DNA.

2.5.5 Purification of DNA fragments

Digested DNA was mixed with 6x DNA sample buffer (**section 2.5.3**) and run on the appropriate percentage agarose gel to obtain sufficient separation of the desired fragment. This fragment was removed using a sterile scalpel and weighed. The DNA was purified using the QIAGEN gel extraction kit, with elution of the DNA in 50 µl Tris-HCl, pH 8.5. This eluted DNA was quantitated and if necessary concentrated by ethanol precipitation, followed by resuspension in 10 µl Tris-HCl, pH 8.5.

2.5.6 Dephosphorylation of linearised plasmid DNA with Calf Intestine Alkaline Phosphatase (CIP)

Prior to ligation of restriction digested vector DNA with insert the vector was routinely dephosphorylated with CIP when either both cohesive ends were complementary or blunt.DNA (5 µg)in 45 µl Tris-HCl pH 8.5 was mixed with 5 µl 10x CIP reaction buffer (supplied with CIP from NEB) to which 1U CIP was added followed by incubation at 37°C for 30 min for cohesive ends; or 37°C for 15 min followed by 55°C for 45 min for blunt ends. The CIP was then inactivated by addition of EDTA, pH 8, to a final concentration of 5 mM and heated at 75°C for 10 min. Vector DNA was then cleaned by phenol / chloroform extraction as described by Sambrook *et al* [102].

2.5.7 Ligation of DNA fragments

To 100 ng vector DNA a 5-fold molar excess of insert DNA was added in addition to 1.5 µl of both 5x KGB (500 mM potassium glutamate, 125 mM Tris-acetate pH 7.5, 50 mM magnesium acetate, 0.25 mg/ml BSA, 2.5 mM 2-mercaptoethanol) and 5x PEG-6000 solution (50 % (w/v) polyethylene glycol 6000 in Milli-Q) with the volume being made up to 13 µl with sterile Milli-Q water. These components were heated at 45°C for 5 min to melt any complementation followed by cooling on ice for 5 min. ATP was added to a final concentration of 1 mM and 1U of T4 DNA ligase added. For cohesive end fragments the reaction was incubated at room temperature for 3 h; whereas blunt end fragments were incubated at 16°C overnight in the presence of 50 µM ATP and 1 µM hexamine chloride. Controls were set up as described in Sambrook *et al.* [102].

2.5.8 Automated DNA sequencing

All DNA sequence analysis was carried out using an Applied Biosystems Inc. 373A system, a departmental facility. This was based on *Taq* dyedeoxy™ terminator cycle sequencing. Plasmid DNA (1 μg) was mixed with 3.2 pmol primer and a reaction premix containing 80 mM Tris-HCl, 2 mM $MgCl_2$, 20 mM $(NH_4)_2SO_4$ pH 9, 4U AmpliTaq® DNA polymerase, deoxynucleotides (G, A, T and C) and Dyedeoxy™ G, A, T and C terminators in 20 μl volume. The reaction mix was overlaid with 20 μl mineral oil and placed in a Hybaid thermal cycler. Tubes were subjected to 25 cycles of the following conditions;

rapid thermal ramp to 96°C,

96°C for 30 sec (denaturation),

rapid thermal ramp to 50 °C,

50°C for 15 sec (annealing),

rapid thermal ramp to 60°C,

60°C for 4 min (extension).

Tubes were then allowed to come to room temperature prior to purification of extension products to remove unreacted dye terminator. This was carried out as for plasmid DNA phenol/ chloroform extraction except that it was necessary to use phenol/ chloroform without 8-hydroxyquinolone, since its fluorescence would interfere with detection of products. The samples were left in dry pellet form and submitted to Ms D. Ashworth for electrophoresis and data generation.

2.6 GST fusion protein expression

2.6.1 Small scale In vivo *expression of fusion protein*

From a 5 ml LB overnight culture containing carbenicillin two aliquots of 9 ml LB / carbenicillin were inoculated with 1 ml each and incubated with shaking for 2 h at 37°C. 10 µl of 0.2 M isopropyl-β-D-thiogalactoside (IPTG) was added to one of the cultures and incubated for a further 2 h at 37°C. Cells were harvested at 1000*g* for 5 min at 4°C. The pellets were resuspended in 200 µl sonication buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM phenylmethylsulfonylfluoride (PMSF)). 10 µl of 40 mg/ml lysozyme in TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) was added and incubated at 37°C for 15 min. Samples were sonicated in an ice-water bath three times for 1 min at a time with 1 min on ice between each. Cell disruption was checked using oil immersion microscopy. Following confirmation of cell disruption 20 µl of 10 % Triton X-100 was added and the samples mixed at 4°C for 30 min. Cell debris was removed by centrifugation at 13000*g* for 10 min at 4°C. 10 µl of the uninduced and induced supernatants where analysed on an SDS-PAGE minigel. (DH5α strain of *E. coli* was used in all studies unless stated otherwise.)

2.6.2 Large scale In vivo *expression of fusion protein*

1 ml of a 10 ml LB carbenicillin overnight culture was used to seed 100 ml of LB carbenicillin which was incubated overnight at 37°C with shaking. This culture was used to inoculate 900 ml LB carbenicillin and grown at 37°C until A_{600nm} of 0.6-0.8 was reached. The culture was induced by the addition of 1 ml 0.2M IPTG and incubated at 37°C for 3 h. The cells were harvested at 5000*g* for 5 min at 4°C and the pellet resuspended in 10 ml sonication buffer prior to incubation with 250 µl lysozyme at 40 mg/ml in TE buffer for 30 min at room temperature. The cells were

sonicated three times for 20 sec each on ice, with 1 min interval on ice between each. The cells were checked with an oil immersion microscope for disruption and sonicated further if required. 1 ml 10 % Triton X-100 was added to the sonicated cells and mixed on a spiramix for 30 min at 4°C prior to pelleting the debris at 11,600*g* for 10 min at 4°C. The resulting supernatant was frozen at -20°C until required.

2.6.3 Affinity purification of GST fusion protein

The GST affinity column was prepared by placing 1.5 ml of the glutathione-sepharose (Pierce) suspension in a 10 ml column and washing with 50 ml sonication buffer containing 1 % Triton X-100 at 20 ml/h at 4°C. The solubilised cell extract was passed through a 0.45 μm filter and applied to the column followed by 100 ml of sonication buffer containing Triton. The column was subsequently washed with 50 ml 50 mM Tris-HCl pH 8.0 but on reaching the end of this wash the column was capped and brought to room temperature. 2.5 ml 10 mM glutathione in 50 mM Tris-HCl pH 8.0 was added to the column and 0.5 ml allowed to drip through before the column was capped and left for 10 min at room temperature. The eluate was collected and dialysed against thrombin cleavage buffer (10 mM Tris-HCl pH 8.0, 20 mM NaCl, 2.5 mM $CaCl_2$). The column was regenerated by washing with 10 ml 0.1 M Tris-HCl pH 8.5, 0.5 M NaCl followed by 10 ml 0.1 M$Na(CH_3CO_2)$ pH 4, 0.5 M NaCl which was repeated prior to washing with 10 ml PBS and storage in 20 % (v/v) ethanol.

2.6.4 Antibody production, purification and analysis

Immunization with 200 μg of the fusion protein was carried out as described in **section 2.3.1**. The animals were boosted as described in this same section. Anti-GST antibodies were removed from the antisera (5 ml) raised against the fusion protein by

Miss E. Cox using a 1 ml immobilized GST column as directed by the manufacturers (Pierce). The resultant antiserum was further purified by passage through a column containing the GST fusion protein coupled to sepharose. The column was prepared by rehydrating 1 ml freeze-dried cyanogen bromide-activated sepharose (Sigma) in 10 ml 1 mM HCl for 15 min at room temperature and washed with 200 ml 1 mM HCl in a sintered funnel. The gel was washed quickly with coupling buffer (100 mM $NaHCO_3$ pH 8.3, 500mM NaCl) and added to 10 mg fusion protein, previously dialysed three times against 2 l coupling buffer over 48 h, in a universal and mixed on a Spiramixer for 2 h at room temperature. Unbound protein was removed by washing with coupling buffer through a sintered funnel followed by resuspending and incubating the gel in 10 ml blocking buffer (200 mM glycine pH 8.0) on a spiramixer for 2 h at room temperature. The gel was subsequently washed in a sintered funnel with 10 ml coupling buffer and 10 ml acetate buffer (100 mM $Na(CH_3CO_2)$ pH4, 500 mM NaCl) alternately five times each prior to packing in 5 ml column. The antiserum from the immobilized GST column was purified by Mrs J. Ingram by passage through this immobilized fusion protein column as described for the purification of site-specific antibodies using a peptide affinity column (**section 2.3.2**).

Chapter 3

Generation of anti-peptide antibodies

3.1 Introduction

Antibodies are some of the most useful tools available to the biochemist due to the specificity which they show towards a defined epitopic region. Owing to this property they have been used to study many aspects of the structure and function of membrane proteins. For example, they have been employed in the investigation of the human erythrocyte glucose transporter, GLUT1 [20], and to probe the substrate binding/ translocation sites of transporters and channels, such as the P-type calcium channel of human small-cell lung carcinoma cells [105]. They have also been extensively used to identify the tissue and subcellular locations of transporters, such as the sodium-dependent glucose transporter SGLT2 in human kidney [106] and to immuno-affinity purify a number of membrane proteins, including the human erythrocyte nucleoside transporter [107].

Conventional procedures for production of antibodies have typically involved utilizing purified proteins as immunogens. For instance, following the isolation of the angiotensin 2 receptor from a murine neuroblastoma cell line it was used in the generation of polyclonal antibodies [108]. The advantage of this method is that it frequently yields antibodies capable of recognizing the native protein. However, because some epitopes are more immunogenic than others, the majority of the antibodies produced tend to be directed against only a few regions of the intact protein. Moreover, a major disadvantage of this methodology so far as the membrane proteins are concerned is that it is usually difficult to purify sufficient amounts of the proteins to the homogeneity required for the production of highly specific antisera.

Many of the difficulties associated with raising antibodies against intact membrane proteins can be overcome by instead using as immunogens either protein fragments generated by recombinant DNA techniques or chemically-synthesized peptides. Knowing the amino acid sequence allows the identification of potential antigenic regions (**section 3.2**). Synthetic peptides corresponding to these regions can then be used to generate either monoclonal or polyclonal antibodies (reviewed by Bailyes *et al.*, [109]). Alternatively, genetic manipulation of the protein can be performed to excise the desired region and fuse it to a carrier protein in a bacterial expression vector, thus allowing over-production of the desired antigen. For example, Craig and Kumar [110] expressed and purified milligram quantities of the ligand binding domain of the human 1,25-dihydroxyvitamin D-3 receptor using a glutathione S-transferase (GST) fusion protein expression system. Unfortunately, in some instances the recombinant protein has proven toxic to the host strain, as demonstrated by Otto *et al.* [111]. In the latter study the tumour necrosis factor (TNF) gene was fused to the pFLAG-1 *E. coli* expression vector, and expression resulted in arrest of cell growth. Occasionally this toxicity may be overcome by changing host strain.

Having decided on the antigenic region and how it is to be presented to the humoral system the type of antibodies required now becomes a factor. Monoclonal antibodies are generated *in vitro* by immortalizing the antibody production of B lymphocytes through fusing the antibody producing cells with myeloma cells. Subsequently, clonal selection provides an infinite supply of antibodies to a defined epitope. Unfortunately, whilst monoclonals are specific for one epitope they sometimes have a low affinity [112]. Furthermore, their specificity for a single epitope may prevent their recognition of even closely related proteins, for example from different animal

species, if the epitope concerned lies in a region of sequence variation. In contrast, polyclonal antisera will contain a population of antibodies with differing specificities and affinities and so may exhibit greater cross-reactivity with related but non-identical antigens. However, a disadvantage of polyclonal antisera is that unless the immunogen used is pure the antisera may contain antibodies to contaminants as well as to the protein of interest. Even if the conditions for their generation have been favorable it is unlikely that more than 5 % of the total serum immunoglobulin content will be specific. However, if the pure antigen is available in sufficient quantity this can be used in affinity purification of the specific antibodies. Therefore, in general polyclonal antibodies tend to be more useful as biochemical tools than monoclonal antibodies, with the exception of immuno-affinity purification where the specificity of monoclonals can be advantageous.

3.2 Criteria for the design of antigenic peptides

Considering the importance of antibodies in research the factors which define an efficient immunogen have remained elusive. Only recently have we begun to understand how to predict a suitable sequence or domain within the target protein that will yield avid antibodies. Scheidtmann [113] in reviewing the topic suggested that several factors have to be carefully considered when designing an immunogenic peptide. The fore-most aspect is that the antigenic site must be accessible to the immunoglobulin on the surface of the protein, i.e. the more polar and hydrophilic a region is the more likely it is that the antibody can gain access. Conversely, the more hydrophobic a region is the more likely it is to be buried within the protein thus

restricting access. Therefore when considering proteins associated with the lipid bilayer this factor takes precedence. For example, when the protein has several transmembrane regions the most suitable sites for peptide design will lie outside of the bilayer in the surrounding inner and outer hydrophilic environments. For these reasons the *N*- and *C*- termini in the majority of cases represent the best candidates for epitopes from which to generate immunogenic peptides, since they are frequently not only hydrophilic but are predicted to lie at some distance from the bilayer surface and may be conformationally less restrained than, for example, the shorter loops connecting transmembrane helices.

Atassi [114] classified epitopes into two types: either continuous, consisting of a single, peptide bond-linked, region of primary structure; or discontinuous, formed by the juxtaposition in the secondary or tertiary structure of non-contiguous residues within the primary structure. Either type tend to be limited in size, typically covering 6-7 amino acid residues at the protein surface, although they are not necessarily the most hydrophilic areas. Within synthetic peptides epitopes tend to be continuous, with short peptide sequences often only possessing a single epitope, whereas longer peptides may contain a series of overlapping epitopes. In consequence, the probability of successfully generating polyclonal antisera is greater for longer peptides [115]. Furthermore, in longer peptides one or more of the epitopes may adopt the native conformation of that sequence within the native protein and therefore be of more use in techniques where the native protein is being investigated, for example immunocytochemistry. For this reason peptides encoding complete domains are good candidates for successful immunogens. However, it has been demonstrated that peptides as small as 13 residues can possess all the structural

information required to attain the correct native-like conformation in free solution and elicit a sufficient immunogenic response [116].

The immunogenic response appears to favor specific amino acid residues as perceived by Scheidtmann [113]. Here antigenic epitopes were observed to contain several amino acids residues with relatively high incidence. Apparently the order of antigenicity of these residues is His > Lys > Ala > Leu > Asp > Arg. Consequently, it may be advantageous to design peptides including at least one of these residues. However, this must not detract from the importance of the conformation of the peptide which also plays a major factor in antigenicity. Therefore it may also be beneficial to include amino acid residues which constrain the neighbouring conformation, such as proline.

The final consideration is how the peptide is to be introduced to the immune system. The general consensus is to couple synthetic peptides to a carrier protein, such as keyhole limpet haemocyanin, with the hope that the proximity of the carrier molecule may induce a conformation in the peptide which is close to that present in the parent protein. However cases have been reported where free peptides of 14-25 residues have led to moderate antibody response using the method described by Muller *et al* [117]. More surprisingly reports of immunization with free peptides consisting of only 6-8 residues have indicated satisfactory results [118]. However, in general the larger the immunogen the better the humoral response resulting in a population of antibodies which recognize more epitopes on the desired protein.

Based on the criteria in this and the previous section two strategies for antibody production were used. The first approach involved the generation of site-directed

peptides against four regions of rCNT1. The second involved the production of a GST-ΔrCNT1 fusion protein. Each of these strategies will be considered independently with a final summarizing section.

3.3 The synthetic peptide approach

3.3.1 An introduction to peptide synthesis

Historically the peptide hormone oxytocin was the first peptide to be synthetically produced, by Du Vigneaud *et al.* [119], with the synthesis of insulin being reported nine years later by Meienhoffer *et al.* [120]. In both cases synthesis was performed in solution, which required purification and characterization of each intermediate at every step. Hence the chemical synthesis of peptides was very time-consuming and complicated until the introduction by Merrifield [121] of the revolutionary approach of solid-phase peptide synthesis. In this approach the generating peptide chain is bound covalently through its *C*-terminus to an insoluble solid support. Synthesis is then carried out by the successive addition of *N*-terminal protected amino acids to give the desired sequence. The formation of a peptide bond, known as coupling, between the *C*-terminal of the newly added amino acid and the deprotected *N*-terminal of the preceding amino acid is achieved by use of a reactive *C*-terminal group such as an ester or anhydride. The advantages of solid-phase peptide synthesis came from the peptide remaining attached to the resin throughout the synthesis, making the loss of the intermediate products minimal. Moreover, by using excess soluble reagents the efficiency of each step could be increased dramatically. Over the years this process of peptide synthesis has become more automated with the result that today reagents only have to be added to a pre-programmed synthesizer to obtain

the required peptide, which then has only to be cleaved from the support resin and purified from the side-chain protecting groups.

Several criteria have to be fulfilled when considering which solid-phase technique to use. The system should allow flexibility of synthesis, maximal coupling efficiency under relatively mild conditions and a means to monitor the processes of deprotection and coupling to ensure efficient incorporation of amino acids into the growing peptide chain. For these reasons fluorenylmethoxycarbonyl (FMOC) solid-phase synthesis has become one of the most popular approaches. The favored solid supports for FMOC chemistry are polyethylene-polystyrene resins, which have replaced the more traditional inert Kieselguhr polyacrylamide matrix. Attached to these are linkers which extend the peptide out from the matrix to allow access for the reagents. Such linkers can vary in length and in terminal reactive group, some may even have the initial amino acid pre-coupled. The synthesis occurs in an environment of a dipolar solvent, such as dimethylformamide, which is ideally suited to the requirements of both the solid support and the peptide with its protecting side chains. Selection of the amino acid derivative is also an important factor, affecting both purity and yield of the final peptide. A great number of different amino acid derivatives are available, but those used in the present study are listed in **Table 3.1**. The importance of side-chain protecting groups is that they prevent chemical modification of the amino acid side-chains during peptide synthesis. The FMOC moiety at the α-amino position of each residue is removed by the secondary base piperidine, which provides mild alkaline conditions but does not interfere with the side chain protecting groups. The *C*-terminal activation of the amino acid is achieved by formation of pentafluorophenol (Pfp) esters, or in the case of Ser and Thr, 3,4-

dihydroxy-3-oxobenzotriazole (Dhbt) esters. This reduces the local electron density of the *C*-terminal carbon and increases the susceptibility of the residue to α-amino group attack. The efficiency of this process is increased by the presence of hydroxybenzotriazole (HOBt) which acts as a catalyst in the presence of carbodiimide.

Throughout the synthesis, deprotection and coupling have traditionally been monitored by following the distribution of the FMOC moiety with pre- and post-column UV detectors. More recently, the counterion distribution monitoring system CDM™ (PerSeptive Biosystems) has been adopted. This measures the distribution of the acidic dye quinoline yellow between the free amine groups of the growing peptide on the resin and the base in solution. As the activated amino acid reacts with the deprotected amine group on the resin or peptide, the dye is displaced from the solid phase thus indicating coupling efficiency.

Prior to using the synthetic peptide it has to be cleaved, deprotected, purified and analyzed. Cleavage and deprotection for FMOC chemistry involves the use of TFA in the presence of scavengers such as 1,2-ethanedithiol, triisopropylsilane, H_2O and thioanisole which prevent undesirable modification of the amino acids by the reactive cleavage products. Purification of the peptide is by filtration, rotary evaporation and precipitation to remove the resin, TFA and scavengers respectively. The resultant product can then be analyzed by rpHPLC and/ or MS.

Table 3.1 Side chain protecting groups for FMOC chemistry

RESIDUE	SIDE CHAIN PROTECTING GROUP
Ala	None
Arg	Pmc
Asn	Tmob
Asp	OtBu
Cys	Trt
Gln	Tmob
Glu	OtBu
Gly	None
Ile	None
Leu	None
Lys	t-Boc
Met	None
Phe	None
Ser	tBu
Thr	tBu
Trp	None
Tyr	tBu
Val	None

Abbreviations used are: **Pmc**, pentamethylchromansulfonyl; **Tmob**, trimethoxycarbonyl; **Trt**, triphenylmethyl; **OtBu**, t-butyl ester; **tBu**, t-butyl; **t-Boc**, t-butoxycarbonyl.

3.3.2 Results and discussion

3.3.2.1 Peptide synthesis, cleavage and purification

Based on the criteria discussed in **section 3.2** the four rCNT1 peptides illustrated below, and in **Figure 3.12**, were synthesized on a Milligen/ Biosearch 9050 Pepsynthesizer by solid phase peptide synthesis using FMOC chemistry, as described in **section 2.2.1**, with the side chain protecting groups indicated in **Table 3.1**. The resultant peptides were cleaved from the resin and purified to remove the side-chain protection groups giving a peptide which was lyophilized and stored at -70^{0}C, as described in **section 2.2.2**. Yields of peptide synthesis can be seen in **Table 3.2**.

Peptide A *Residues 1-15*, mol. wt: 1853

M A D N T Q R Q R E S I S L T *C*

Peptide B *Residues 389-406*, mol. wt: 2161

V E E S K F R S E N G V K L T Y G D *C*

Peptide C *Residues 505-524*, mol. wt: 2495

***C* S Q Y K Q R R L A G A E E W L G D K K Q**

Peptide D *Residues 618-637*, mol. wt: 2146

C R Q V F Q S T S S E F S Q V A L D N

In the sequences shown above *C* represents the incorporation of an additional cysteine residue to allow the conjugation of the peptide to the carrier. Molecular weights are given for each peptide. Yields for the synthetic production of the peptides indicated in **Table 3.2** show that percentage yields varied between 47 and 74 percent. This variance was attributed to practical experience of the techniques used as with more experience the yields were observed to increase.

3.3.2.2 Peptide analysis

Peptides B and C were dissolved in Milli-Q H_2O whereas peptides A and D were dissolved in 2 % DMSO prior to rpHPLC and MS analysis. Each peptide was then analyzed by MS and rpHPLC as described in **sections 2.2.3** and **2.2.4**, respectively.

Mass spectrometric analysis of the peptides indicated that in each case a molecule of molecular mass essentially identical (within the limits of accuracy of the spectrometer) to that of the desired peptide was obtained (**Figures 3.1a-3.4a**). However, in addition other molecules were present with molecular masses either larger or smaller than the required product. Comparison of the molecular masses of these species with those of the desired peptide in most cases led to their putative identification as peptide products into which either an extra amino acid had been incorporated, or from which an amino acid residue was lacking, or from which removal of the side-chain protection groups was incomplete. A summary of these putative identifications can be seen in **Table 3.3**. An explanation for the first of these side-product types is that the incorporation of an additional residue probably occurred when there was originally inefficient coupling during synthesis resulting in that residue being recoupled, thus allowing for duplication of the same residue in a small proportion of cases. The lack of a residue also derives from inefficient coupling where even on recoupling of that residue a second time a small proportion of synthesizing peptides do not conjugate the residue. Finally, the remainder with residual protecting groups attached are indicative of inefficient deprotection suggesting that a longer period of peptide treatment with TFA would have been beneficial.

While the peak height in the MS traces shown in **Figures 3.1a-3.4a** might suggest that the undesirable side-products were relatively abundant for some of the syntheses, in fact there is no simple relationship between peak height and abundance in such MS analyses. Indeed, further analysis by rpHPLC to determine the relative quantities of the desired product and side-products indicated that in each case only one major product was obtained (**Figures 3.1b-3.4b**). Considering this and the information in **Table 3.3** it was concluded that in each case the correct molecules had been synthesized with only minor contamination by the side-products. Therefore, based on both MS and rpHPLC analysis it was considered that these four peptides synthesized were of satisfactory quality to proceed with conjugation to ovalbumin and subsequent antibody generation.

3.3.2.3 Antibody production

Each peptide was conjugated to ovalbumin as described in **section 2.3.1**. This involved blocking the endogenous thiol sites in ovalbumin with N-ethylmaleimide followed by its incubation with the linker m-maleimidobenzoyl-N-hydroxsulfo-succinimide ester (sulfo-MBS). Once derivatised in this fashion, the ovalbumin was mixed with the peptide (previously reduced by DTT treatment) in order to allow conjugation.

Each peptide was used to immunize two New Zealand white laboratory rabbits by injection of 200 μg conjugate as described in **section 2.3.1**. Following boosting test bleeds were removed and tested by ELISA to determine the titre, as described in **section 2.3.3**. Titre is described as the dilution required to achieve 50 % maximal

immunogenic activity. During the immunization program several rabbits developed abscesses and therefore were destroyed.

An example of the development of an immune response can be observed in **Figure 3.5** where the peptide B conjugate has been used for immunization. The initial bleed was obtained after two weeks while subsequent bleeds were obtained at four week intervals following this. **Table 3.4** summarizes the development of the immune response in each of the rabbits immunized. This variance in immune response between different antigens and the same antigen in different animals is one of the disadvantages in producing antibodies from animal hosts.

3.3.2.4 Affinity purification of anti-peptide antibodies

In order to eliminate plasma proteins, including any non-specific IgG , the antibodies obtained were affinity purified by passing the antisera down the desired peptide affinity column. These affinity columns were synthesized by covalently coupling the Cys-containing peptides to Sulfolink™ by their terminal thiol groups, as described in **section 2.3.2**. Following DTT treatment, assay for free thiol groups indicated 70-90 % of the peptide cysteine had become reduced, indicating that of the 5 mg of starting peptide 3.5 to 4.5 mg were capable of covalent attachment to the resin. The efficiencies of coupling were 59 %, 74 %, 51 % and 73 % for peptides A to D respectively. These columns were subsequently used in the affinity purification of anti-peptide specific antibodies from 5 ml of serum from each of the final bleeds. An example of such a purification can be seen in **Figure 3.6**. Following passage of the antiserum through the column, washing under high salt conditions elutes non-specifically bound IgG and other plasma proteins, leaving specific polyclonal

antibodies which can subsequently be eluted with alkali followed by acid to remove the more avid IgG. A summary of the purification of each peptide-directed antiserum can be seen in **Table 3.5.** This indicates that the extent of purification varies greatly from 2-fold to 200-fold when compared to the antisera. An explanation for this is that an avid antibody will bind strongly to the column, thus allowing washing to remove the majority of the non-specific IgG and plasma protein, while the specific antibody is retained for subsequent elution having been purified greatly. In comparison, an antibody with low avidity to the column will result in a proportion of this antibody being removed on washing, thus giving a lower degree of purification.

Table 3.2 Synthetic peptide production yields

Synthetic Peptide	Theoretical Yield (mg)	Observed Yield (mg)	Percentage Yield (%)
A	324	153	47
B	378	210	55
C	359	260	74
D	365	257	70

Theoretical yields were obtained from the software supplied with the Milligen/ Biosearch 9050 Pepsynthesizer. Observed yields represent the actual amount of peptide obtained after cleavage, purification and lyophilization.

Table 3.3 Side-products of peptide synthesis

Peptide	Mass	Difference from desired product	Identity
A	1779	-76	?
	1912	+57	+ tBu
	2175	+320	+ tBu + Pmc
B	1811	-346	-2(Glu / Lys) + Ser
	1941	-216	-(Glu / Lys) + Ser
	2028	-129	- (Glu / Lys)
	2287	+130	+ (Glu / Lys)
	2388	+231	+(Glu / Lys) + Val
C	2552	+56	+ tBu
D	-	-	-

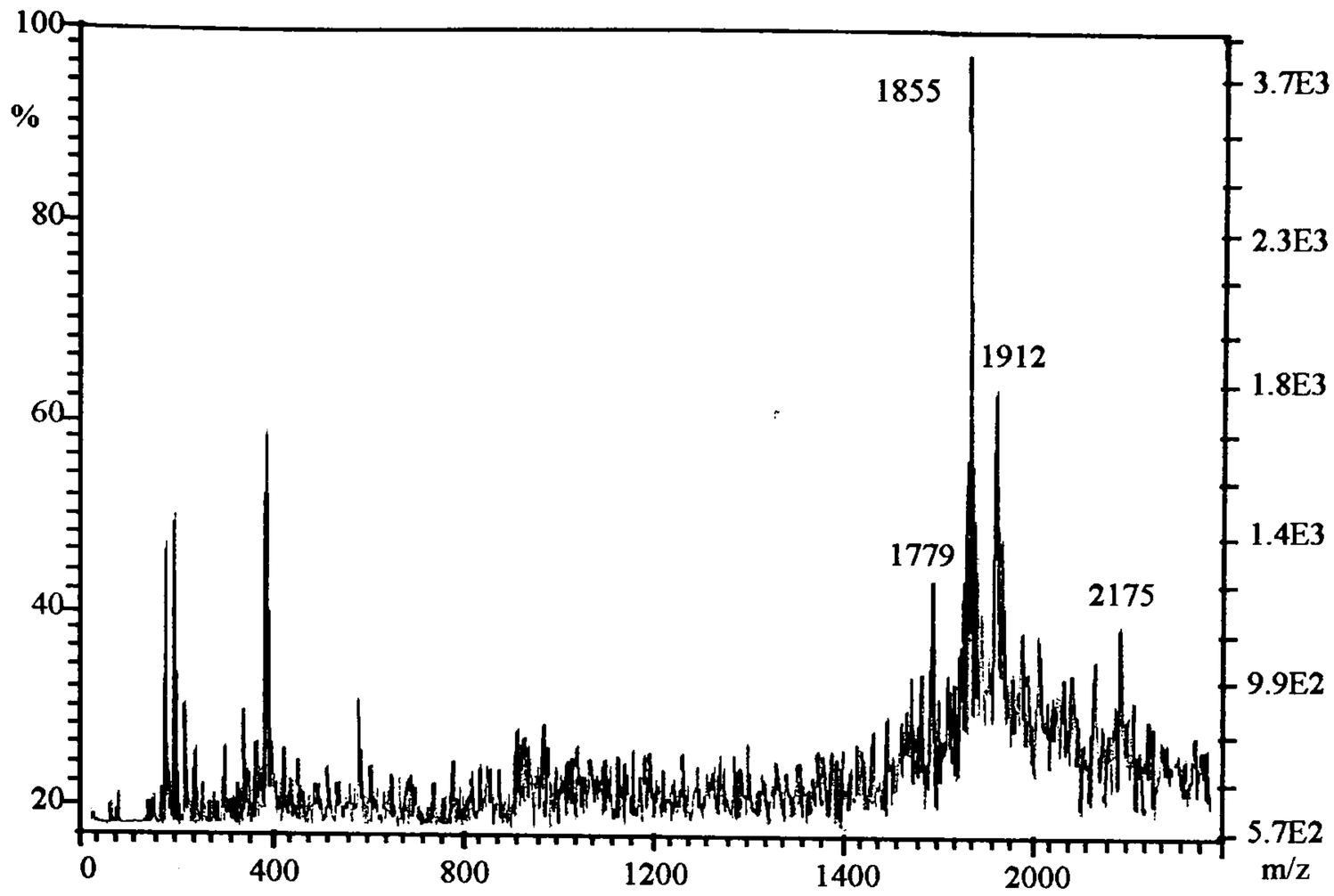
MS analysis, **Figures 3.1a to 3.4a**, indicated that side-products of synthesis and cleavage were present in the final peptide samples. The MS peak values of these contaminants have been analyzed and their putative relationships to the desired peptide products summarized in the above table. Miscalibration of the instrument has resulted in minor differences of some contaminants.

Table 3.4 Immunogenic response to synthetic peptides

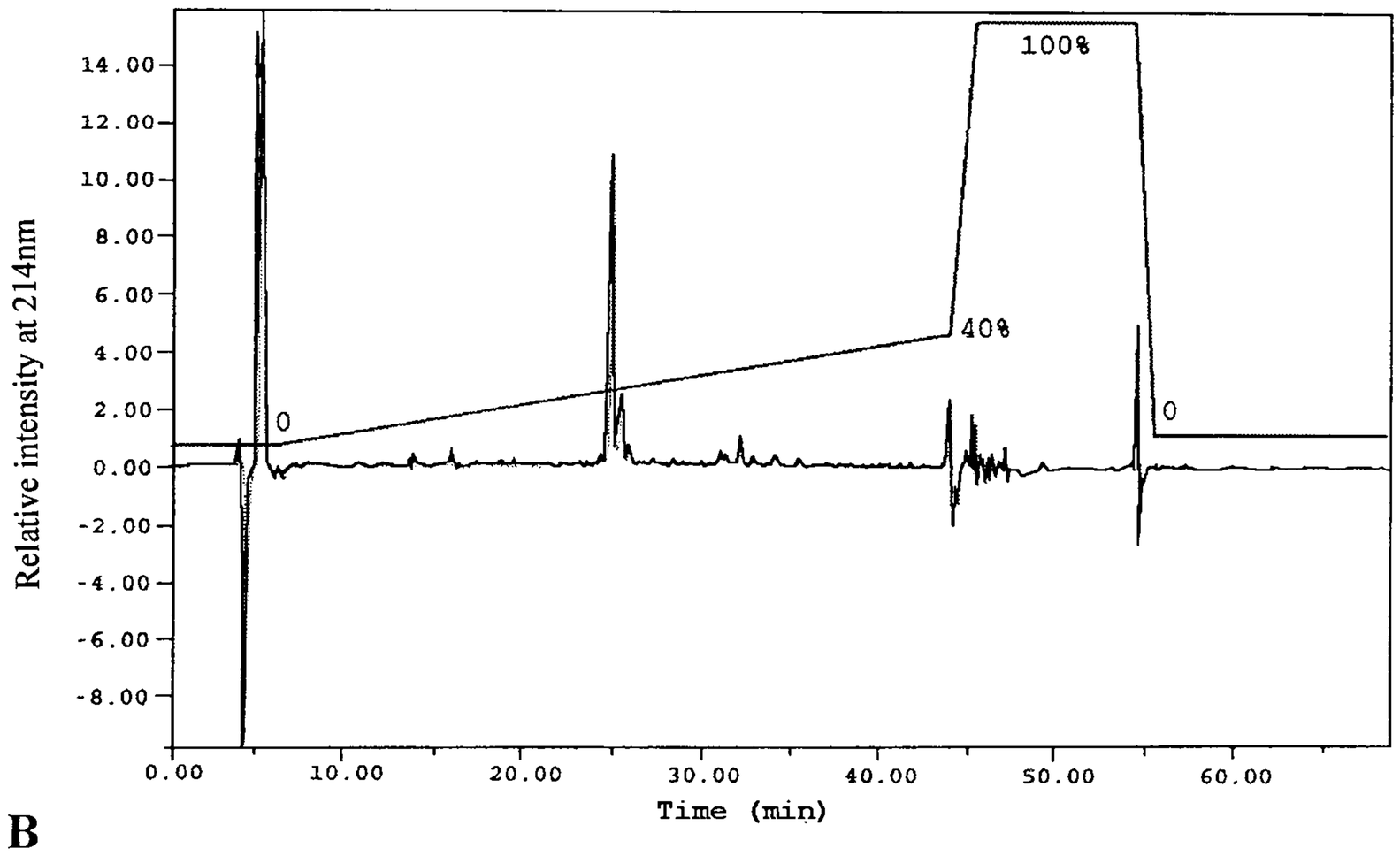
Peptide	Rabbit	Titre			
		Bleed 1	Bleed 2	Bleed 3	Bleed 4
A	3795	**T**			
	3895	1:100	1:500	1:800	**1:2000**
B	3995	1:900	1:1050	1:2000	**1:20000**
	4095	**1:250**	**T**		
C	4195	1:9000	1:10000	**1:40000**	
	4295	**T**			
D	4395	1:9000	1:10000	**1:10000**	
	4495	1:26000	1:40000	**1:30000**	

For each peptide two rabbits were immunized with the peptide conjugate. Titre values were obtained by ELISA with the corresponding peptide. Final bleed titres are indicated (*bold face*) and T indicates where the rabbit developed side effects, necessitating its premature killing. Titre is defined as dilution for 50 % maximum activity.

Figure 3.1 MS and rpHPLC analysis of synthetic peptide A



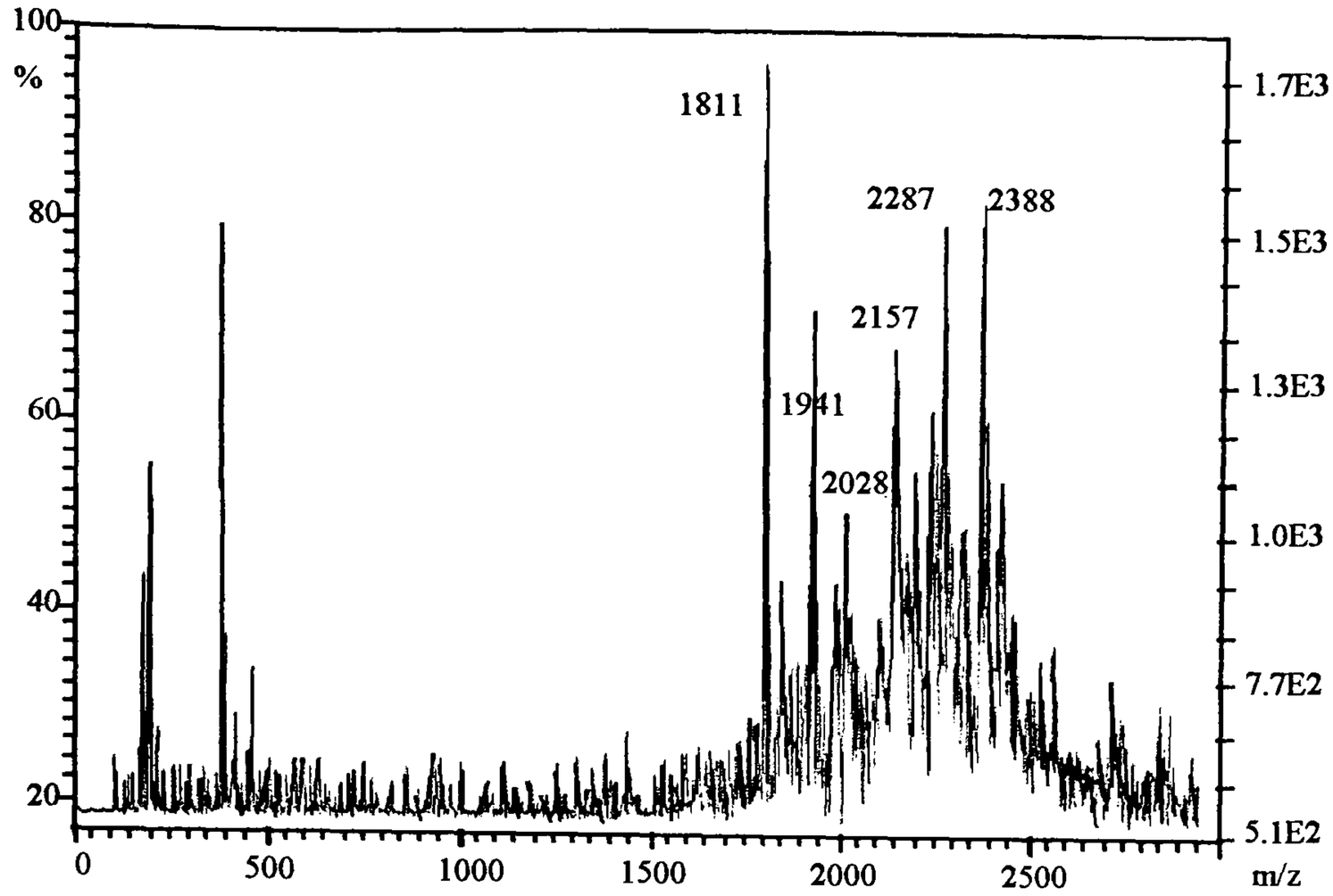
A



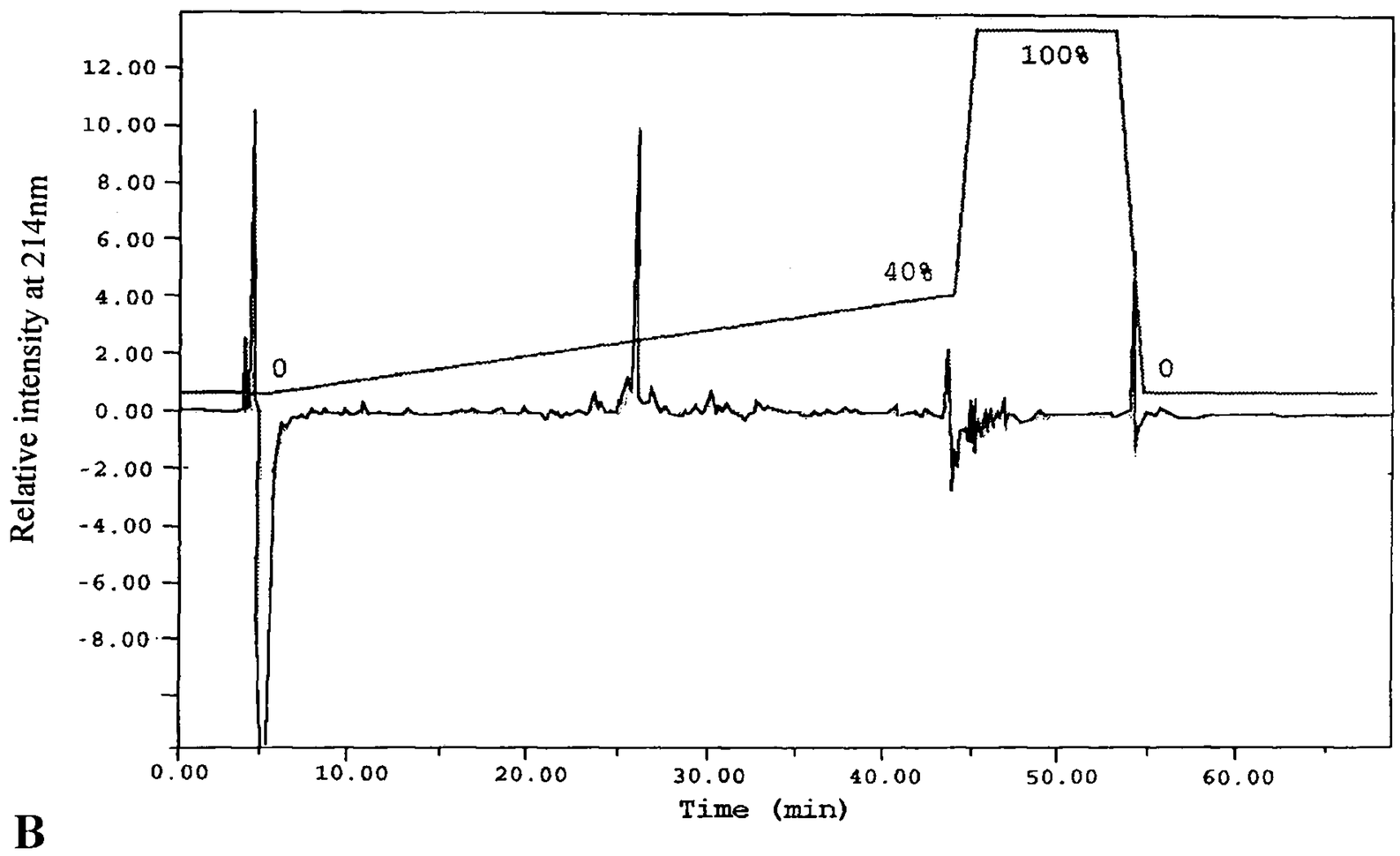
B

Peptide A dissolved in 2 % DMSO was subjected to MS (A) and rpHPLC (B) as described in **sections 2.2.3** and **2.2.4** respectively. The discrepancy between the theoretical and observed molecular masses, 1853 and 1855 respectively, may be due to miscalibration of the spectrometer. In the rpHPLC plot the grey line represents percentage of solvent B present during elution and washing of the column.

Figure 3.2 MS and rpHPLC analysis of synthetic peptide B



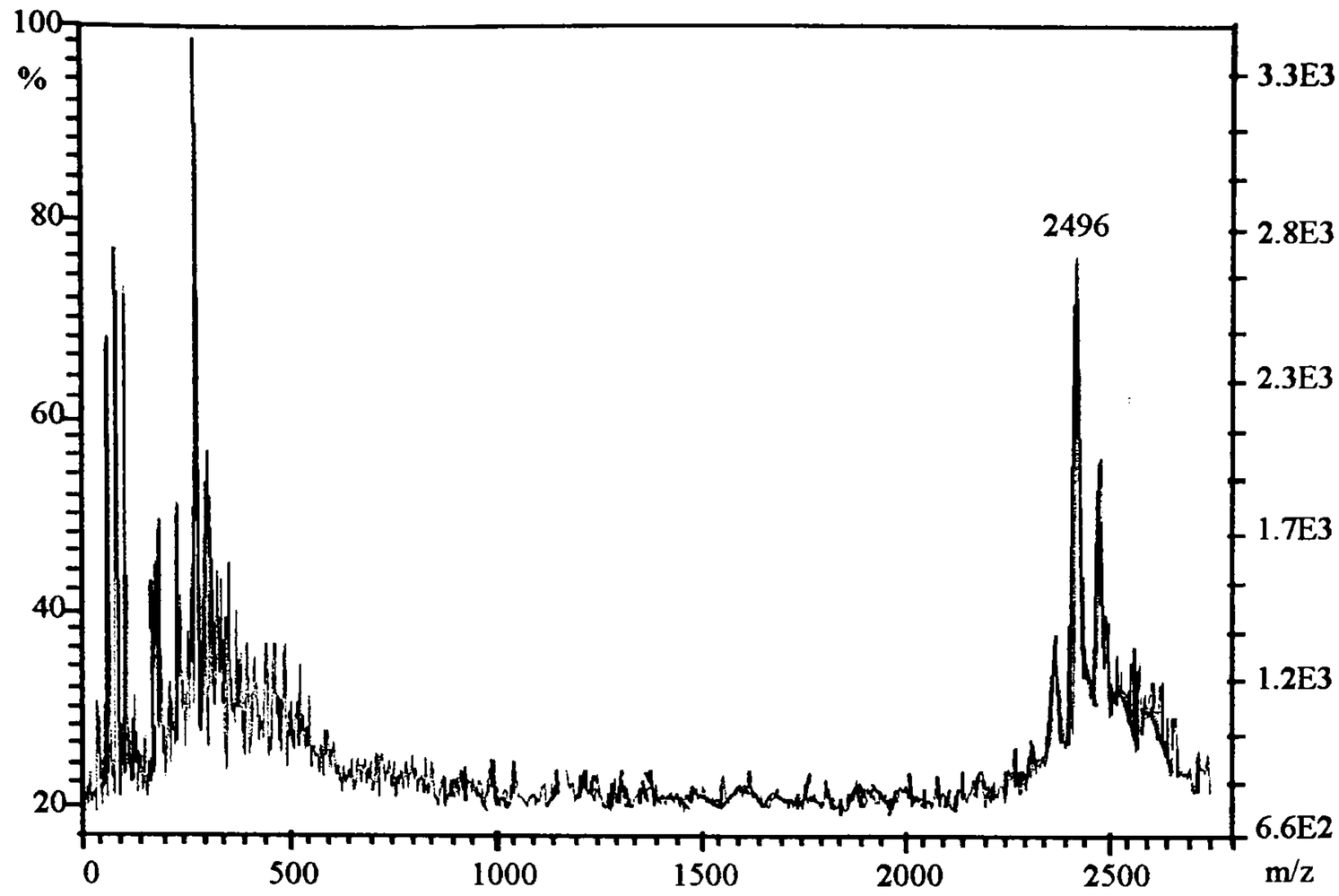
A



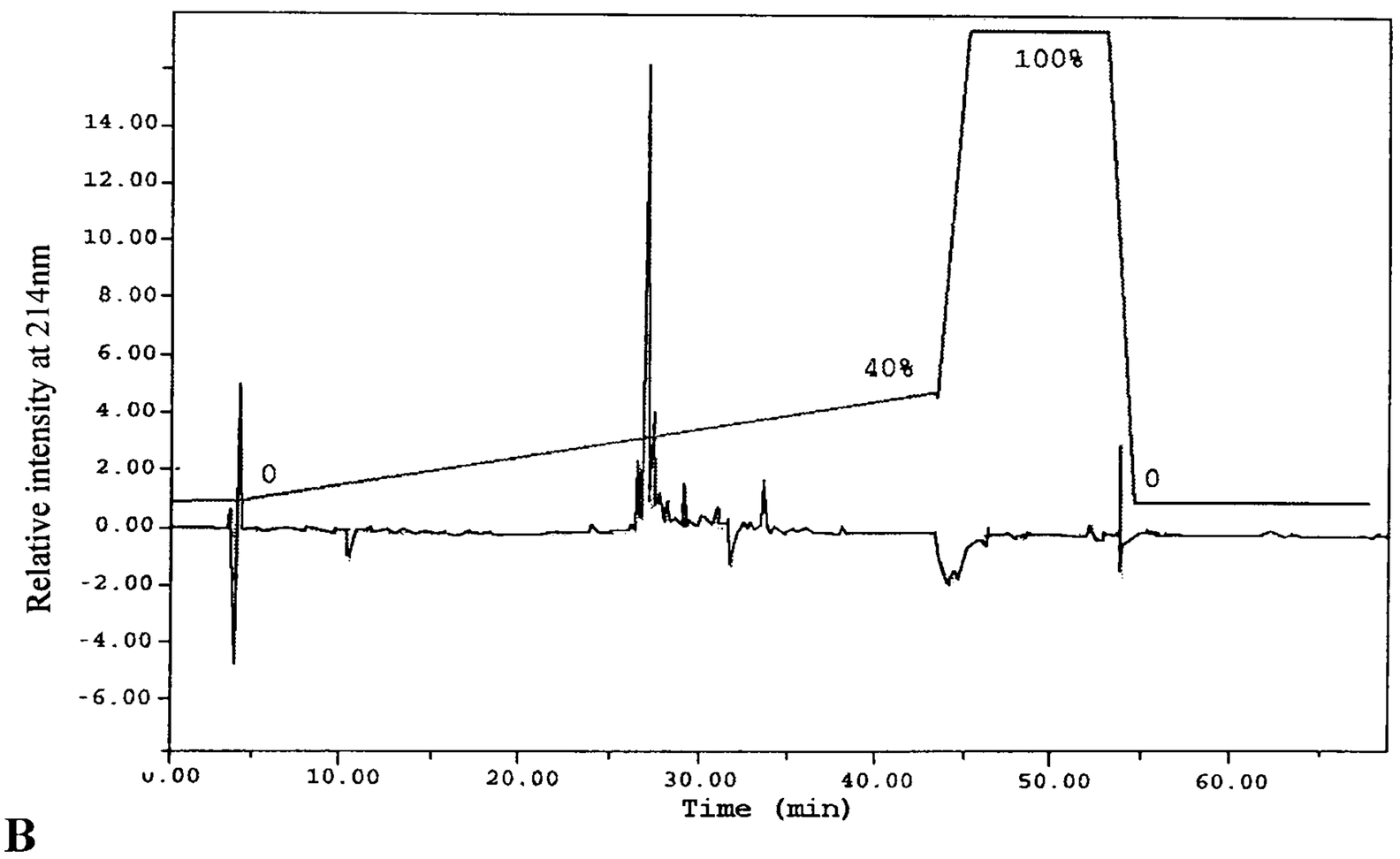
B

Peptide B dissolved in Milli-Q was subjected to MS (A) and rpHPLC (B) as described in **sections 2.2.3** and **2.2.4** respectively. The discrepancy between the theoretical and observed molecular masses, 2161 and 2157 respectively, may be due to miscalibration of the spectrometer. In the rpHPLC plot the grey line represents percentage of solvent B present during elution and washing of the column.

Figure 3.3 MS and rpHPLC analysis of synthetic peptide C



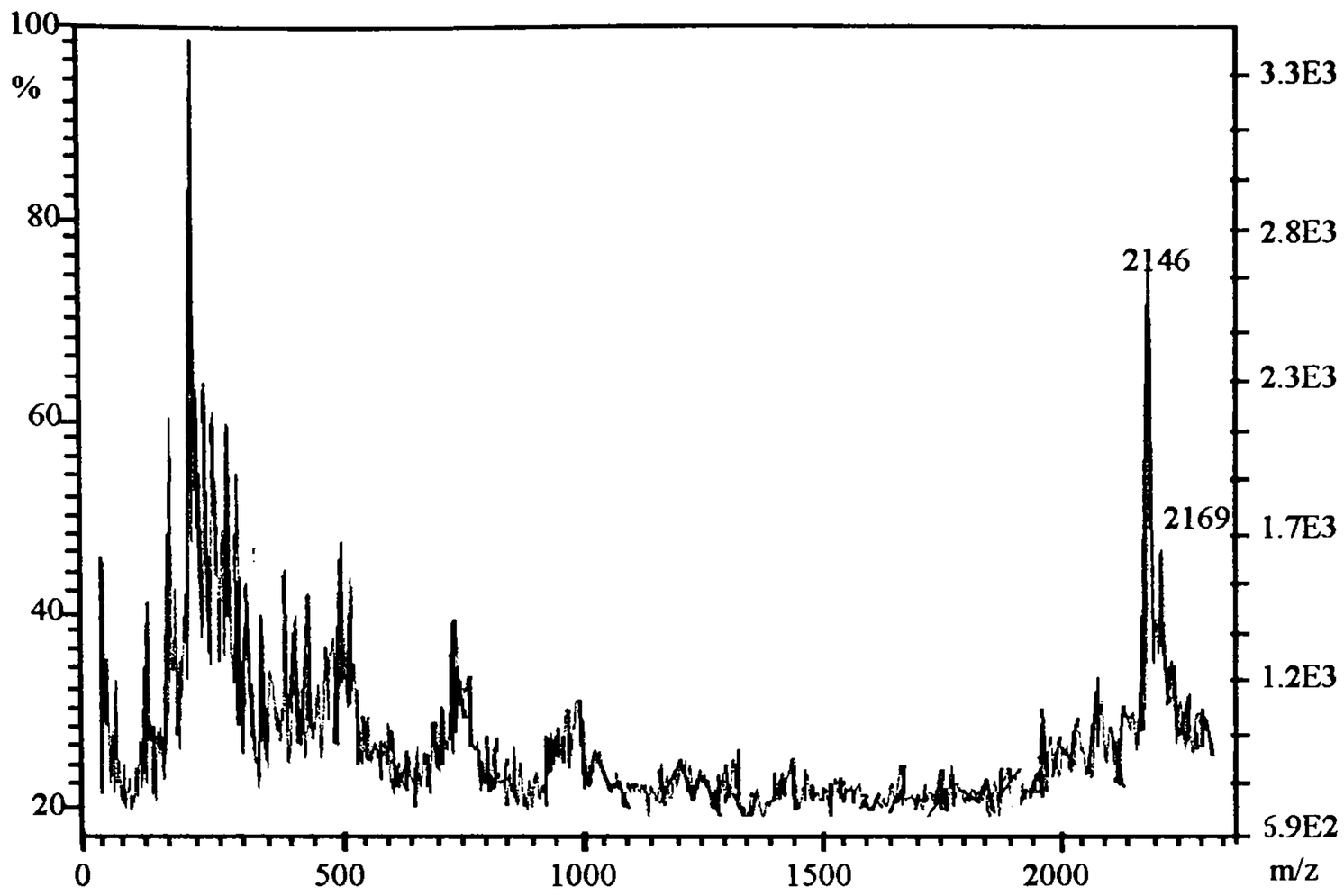
A



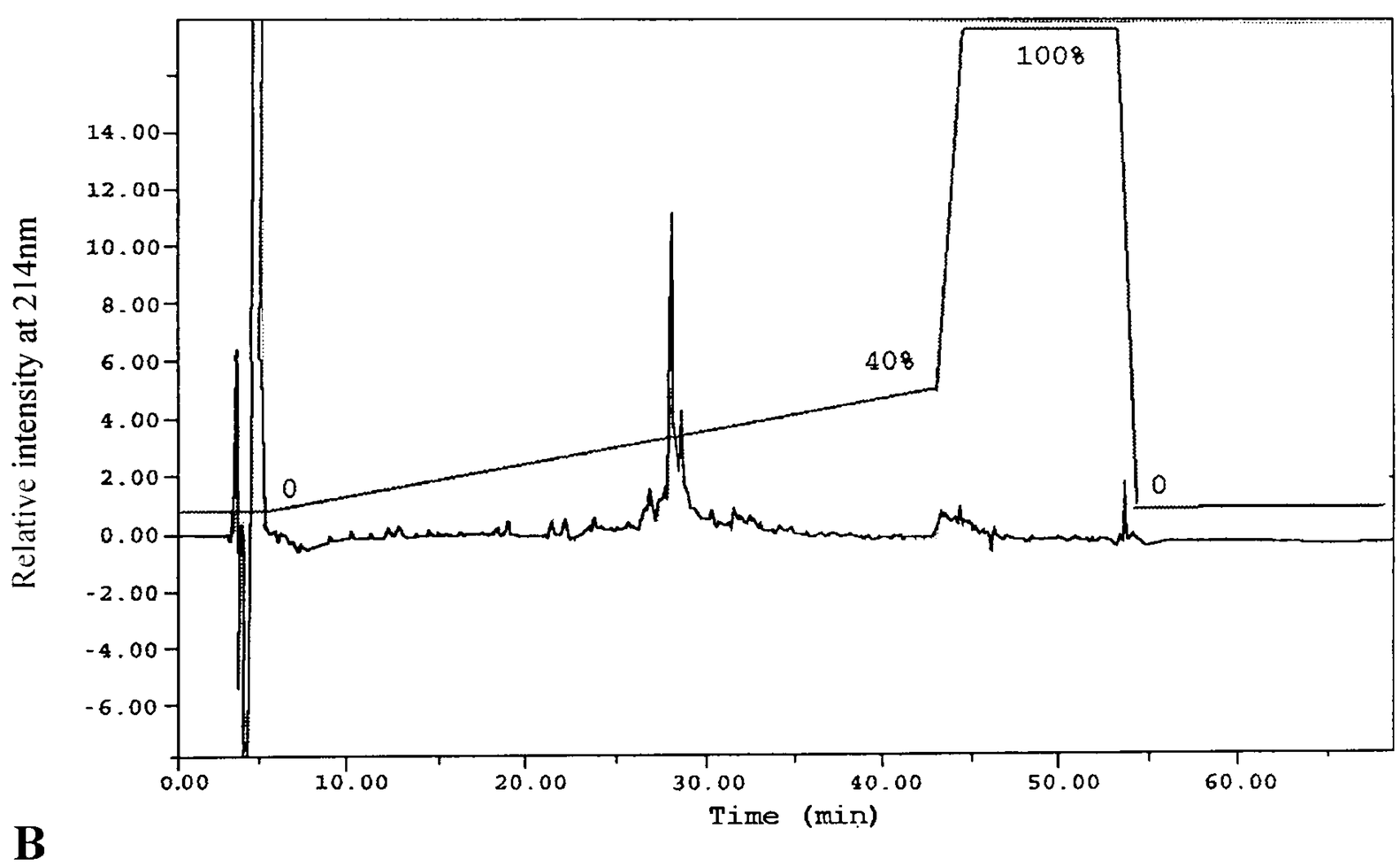
B

Peptide C dissolved in Milli-Q was subjected to MS (A) and rpHPLC (B) as described in **sections 2.2.3** and **2.2.4** respectively. The discrepancy between the theoretical and observed molecular masses, 2495 and 2496 respectively, may be due to miscalibration of the spectrometer. In the rpHPLC plot the grey line represents percentage of solvent B present during elution and washing of the column.

Figure 3.4 MS and rpHPLC analysis of synthetic peptide D



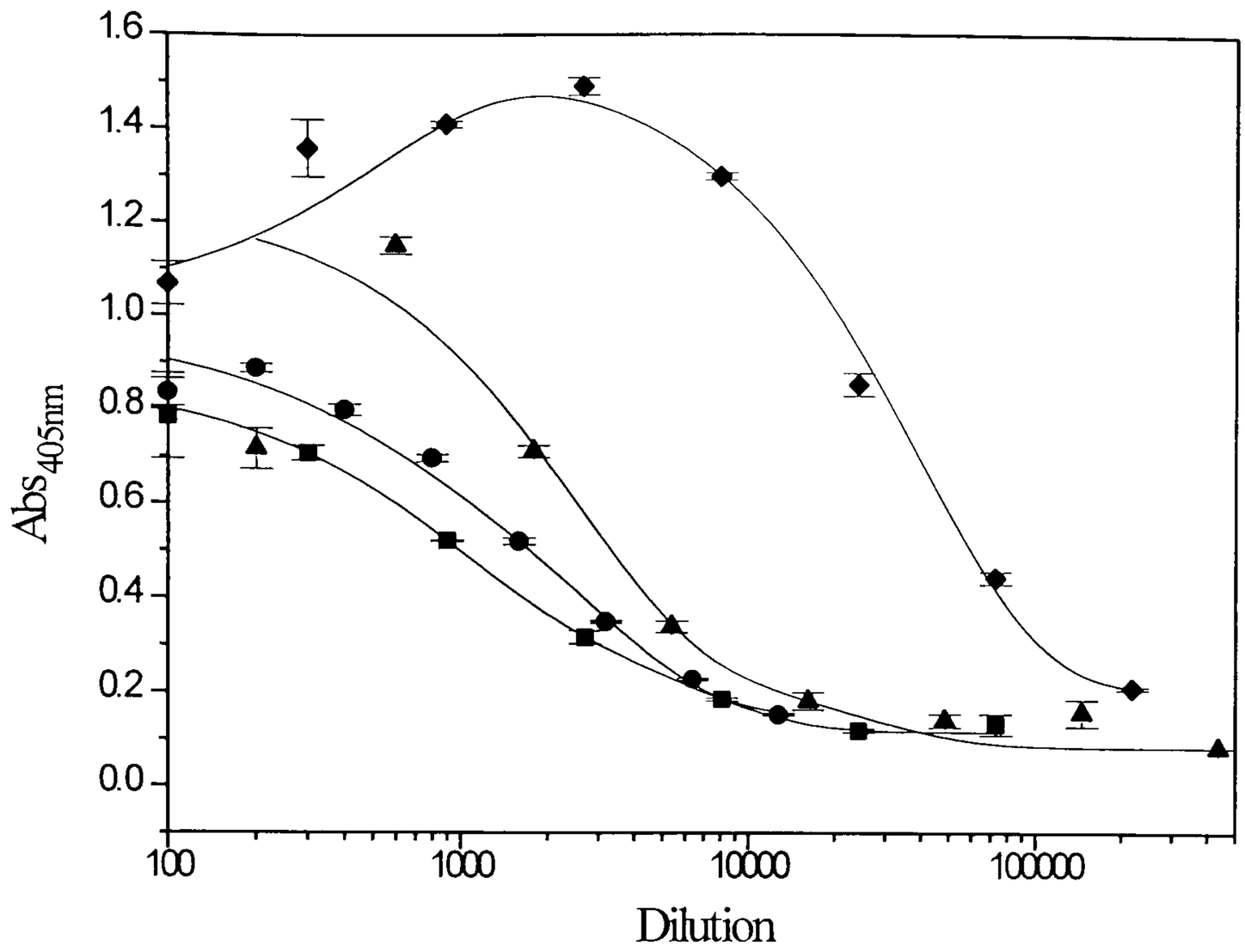
A



B

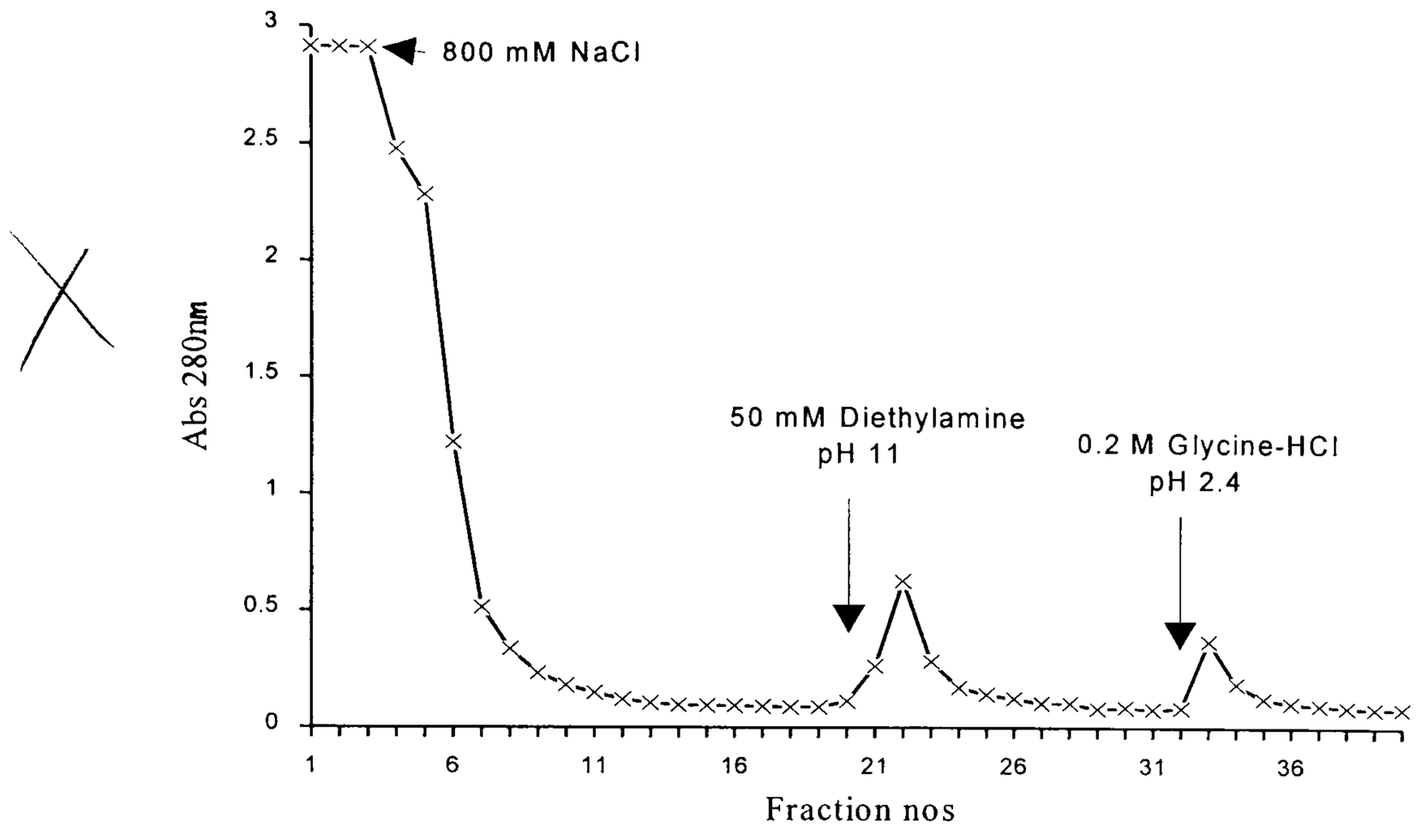
Peptide D dissolved in 2 % DMSO was subjected to MS (A) and rpHPLC (B) as described in **sections 2.2.3** and **2.2.4** respectively. The theoretical and observed molecular masses give the same value of 2146. In the rpHPLC plot the grey line represents percentage of solvent B present.

Figure 3.5 Development of immune response against peptide B conjugate



The antibody titre of rabbit 3995 following immunisation with an ovalbumin conjugate of peptide B was assessed by ELISA as described in **section 2.3.3**. Samples were taken 2 (■), 6 (●), 10 (▲) and 14 (◆) weeks after the initial immunisation, and the rabbit was given booster immunisations one week after the removal of each of the first three samples.

Figure 3.6 Affinity purification of anti-peptide D antibodies



The elution profile of anti-peptide D antibodies from a peptide affinity column, as described in **section 2.3.2**, demonstrates the elution of non-specific IgG on addition of NaCl followed by the release of specific IgG under alkali and acid conditions successively. Comparison of the relative concentrations of alkali and acid eluted fractions of all four anti-peptide antibodies with their antisera is summarized **Table 3.5**.

Table 3.5 Summarization of antibody purification

Anti-peptide IgG	Fraction	Titre	Approxi-mate yields (%)	Protein conc. (mg/ml)	Ratio: Titre/[Protein]	Fold purificatio n
A	Antiserum	1:2000	-	67.8	30	-
	Alkali	1:200	11.6	0.1	2000	**67**
	Acid	1:600	39.0	0.1	6000	**200**
B	Antiserum	1:20000	-	56.3	113	-
	Alkali	0	0	0.1	0	**0**
	Acid	1:300	12.6	0.1	3000	**27**
C	Antiserum	1:40000	-	46.5	860	-
	Alkali	1:4500	14.0	0.1	45000	**52**
	Acid	1:5500	12.4	0.2	27500	**32**
D	Antiserum	1:30000	-	59.7	503	-
	Alkali	1:700	2.0	0.7	1000	**2**
	Acid	1:3000	5.0	0.1	30000	**60**

Fractions referred to as alkali or acid apply to samples eluted under alkaline or acidic conditions respectively. Approximate yields were calculated by comparison of the (titre*volume) value of the purified samples with that of the antiserum from which they were purified. The values obtained for the purified antibodies can be compared with those of the antisera to give the fold purification (*bold face*). Titres of antisera were established by ELISA while protein concentration was determined by the BCA protein assay, as described in **section 2.4.2**.

3.4 The fusion protein approach

3.4.1 An introduction to fusion proteins

Fusion proteins have been widely used in the investigation of biological systems. For example, they have been used for the over-expression and purification of proteins as in the case of the soluble ligand binding domain of the human vitamin D3 receptor [110], for investigation of protein/ protein interactions such as those involved in the secretion of bacterial proteins [122] or the association of Ran GTPase with nuclear-pore proteins [123], and in the study of cellular signaling mechanisms including the role of tyrosine kinase in cell surface expression of the facilitative glucose transporters GLUT1 and GLUT4 [124,125]. They have also been extensively used in vaccination against diseases such as parasitism by tape worm in sheep [126] and lyme disease in mice [127], and in the generation of antibodies for investigative procedures, for example the immunocytochemical analysis of Na^+/Cl-dependent RXT1 transporter of the rat central nervous system [128] and immuno-detection of the human herpes virus 7 glycoprotein B [129].

The expression of foreign genes in *E. coli* presents a number of potential problems. For example, the foreign polypeptide may be toxic to the host cell, it may be degraded *in vivo* or incorrectly folded, and it may be difficult to purify. Toxicity can often be avoided by placing the gene under the control of a strong promoter which is normally repressed but that can be induced during the terminal stages of cell growth, thus allowing the polypeptide synthesis. Degradation can be overcome by fusing the foreign polypeptide to a native protein which acts as a vector in order to stabilize the introduced protein. Such fusion to a vector protein can also facilitate the purification of the polypeptide if the vector keeps its activity in the hybrid and allows for easy

purification, for example, by affinity chromatography. Alternatively the fusion protein can be selectively exported from the cytoplasm by the fusion to a signal peptide as a means of avoiding toxicity, although this in itself may prove toxic to the host cell.

Several vector proteins have been used with varying success. For example, β-galactosidase has been fused to both the N- and C-termini [130,131] of foreign proteins, the phosphate binding protein, PhoS, has been fused to the human growth-hormone-releasing factor [132], and the *S. aureus* protein A to alkaline phosphatase allowing a single-step purification by IgG affinity chromatography [133]. Additionally, a polyarginine tail has been fused to the C-terminus of human urgastrone to allow simple ion exchange purification [134] while the maltose binding protein has permitted an inexpensive means of purification of fusion proteins on an amylose affinity column [135,136]. Purification of the fusion protein products from each of the fore-mentioned systems varies greatly, as exemplified by the protein A fusion proteins requiring acidic elution which can be expected to alter the antigenicity and functional activity of the purified product whereas the maltose binding protein only requires competition with free maltose and is therefore relatively mild.

A widely used fusion partner for protein expression in *E. coli* is the enzyme glutathione S-transferase (GST), a 26 kDa protein from the parasitic helminth *Schistosoma japonicum*. Use of GST fusions have in particular been promoted by Pharmacia through its range of pGEX vectors, which are designed for inducible, high-level intracellular expression of the hybrid proteins. These vectors offer high level expression under the control of the strong *tac* promoter, which is IPTG-

inducible. In the absence of IPTG potent inhibition of expression is ensured by the presence of an internal *lac* I^q gene, and so problems of toxicity of the fusion protein are minimised. Following induction of their expression by IPTG, fusion proteins can be purified under mild conditions by adsorption to a column of immobilised glutathione, and then elution with free glutathione. Many vectors also encode proteolytic cleavage sites to facilitate subsequent release of the foreign polypeptide from the GST. Such a system has been used successfully in many applications including the production of vaccines [126,127], molecular immunology [137], and studies involving protein-protein [138] and DNA-protein [139] interactions.

3.4.2 PCR amplification

The polymerase chain reaction (PCR) allows the selective amplification of a chosen region of a DNA molecule using two oligonucleotide primers that flank the DNA fragment to be amplified. The process involves repeated cycles of heat denaturation of the template DNA, annealing of the primers to their complementary sequences and extension in the presence of a thermostable DNA polymerase [140]. The amplification takes place by extension between the two 3'-termini of the primer pair in such a way that after the initial cycle the first generation PCR product becomes the template for successive rounds of amplification and hence only the specific fragment is amplified.

Successful PCR results in the generation of a specific fragment of DNA. Frequently non-specific products are produced which can interfere with subsequent manipulations. If these by-products occur early on in the amplification then they will be amplified similarly to the target DNA. Thus, in order to maximize the specificity of a PCR, it is necessary to ensure correct annealing of the primers to their desired

positions on the DNA template. Numerous factors influence the extent to which PCR is successful, including optimization of the magnesium ion concentration, primer concentration, template concentration, optimization of the annealing temperature and the design of the primers. The factor often given the strongest weighting is in the design of the primers in which the following simple rules allow the design of efficient primers. The 3' regions of the primers should comprise 15-25 nucleotides of sequence complementary to the target site, to allow high-temperature annealing; incorporation of restriction sites should occur at the 5' end, thus allowing efficient priming at the 3' end; restriction sites located adjacent to the ends of PCR products are normally inefficiently cleaved, so it is recommended to add 3-4 extra bases outside of the site; the incorporation of complementary regions which may lead to formation of 'primer-dimers' should be avoided; the melting temperatures (T_m) of both primers should be similar; and their (G+C) content should be around 50 %. These rules are incorporated in the Apple-Mac program Amplify v1.2 which can be used to perform a theoretical PCR reaction indicating if the primers are suitable. If during the practical PCR false-priming occurs then the set annealing temperature can be increased to ensure more specific annealing, failing this the magnesium concentration is modified.

3.4.3 Strategy for production of ΔrCNT1-GST fusion protein

The aim of the work described in this section was to produce a GST fusion protein encoding a relatively large fragment of the rCNT1 *N*-terminus for immunization purposes. The vector used was pGEX-KT [141], which was derived from the pGEX series vector pGEX-1 (see above). pGEX-KT contains a glycine-rich kinker region, containing the sequence S-G-G-G-G-G, to increase the efficiency of cleavage at a

thrombin cleavage site. The latter is located *C*-terminal to the kinker, such that the foreign protein bears only an additional glycine and serine residue at its *N*-terminus following cleavage of the fusion protein.

The template for amplification of the rCNT1 *N*-terminal fragment was the plasmid pSRH1. This plasmid was derived from pQQH [81], obtained as a gift from Professor J. D. Young (University of Alberta, Edmonton, Canada). Plasmid pQQH contains the complete sequence of rCNT1 cDNA inserted as a 2.4 kb fragment between the *Eco*RI and *Xba*I restriction sites of pGEM-3Z. To facilitate excision of the coding region of rCNT1 for other purposes, an internal *Eco*RI restriction site at position 245 in pQQH was eliminated by PCR mutagenesis, through conversion of the sequence GAATTC to GAGTTC, thus creating pSRH1 (results not shown). The mutation was silent, the glutamate residue at position 27 of rCNT1 being retained.

For the present purposes, PCR primers were designed for amplification of the region of pSRH1 encoding residues 21 to 78 of rCNT1, the location of this peptide fragment in rCNT1 is illustrated in **Figure 3.12**. This peptide fragment was designated ΔrCNT1. The choice of primer locations was made not only because this region of the protein was predicted to be immunogenic, but also because the nucleotide sequences of the primers were conserved in the cDNA sequence encoding the homologous transporter hCNT1 (see **Figure 3.7**). Hence the primers could if necessary also be used for amplification of the corresponding region of the latter cDNA.

The locations of the 5' and 3' primers chosen for the amplification are shown in **Figure 3.7**. The 5' primer sequence was designed to incorporate a 5' *Bam*HI

restriction site, yielding the primer CNT1_21*Bam*HI, whose sequence is shown below (restriction site is italicized). The 3' primer sequence was designed to include a 5' *Eco*RI restriction site, producing the primer CNT1_78*Eco*RI, whose sequence is shown below (restriction site is italicized). In addition, the 3' primer was further modified to incorporate a stop codon (ATG, as indicated by the arrows in the primer sequence below). The resultant primer sequences were checked using the Mac program Amplify v1.2.

CNT1-21*BamHI* ⇒

5′ gcc*ggatcc*CTGGAGAACATGGGG 3′ Tm ~80°C

⇐ CNT1-78*EcoRI*

3′ AAGACGTCCCTCGTGact*cttaag*ccg 5′ Tm ~84°C

↑↑↑

Amplification of rCNT1 using these primers, followed by restriction digestion with *Bam*HI and *Eco*RI, was predicted to yield a cDNA fragment encoding residues 21 to 78, that could be ligated in frame to *Bam*HI/ *Eco*RI digested pGEX-KT. The resultant nucleotide and protein sequences of the construct and fusion protein in the region of the fusion junction are illustrated below.

Δ **rCNT1/ hCNT1** ⇒

TGGGATCTGTCTGGTGGTGGTGGTGGTCTGGTTCCGCGTG**GATCCCTGGAGAACATGGGGCA**

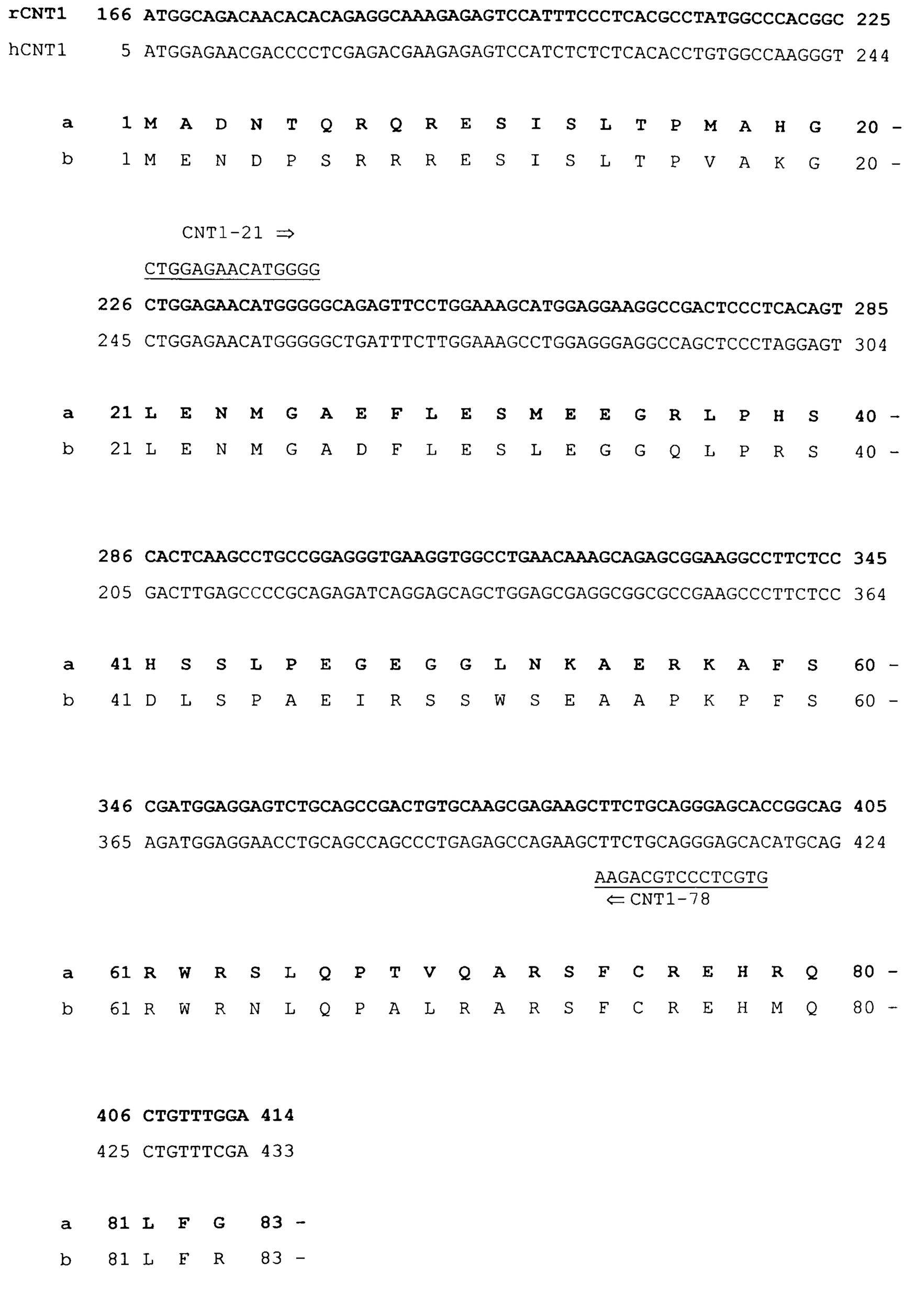
ACCCTAGACAGACCACCACCACCACCAGACCAAGGCGCACCTAG**GGACCTCTTGTACCCCCGT**

S D L **S G G G G G** L V P R G S **L E N M G A**

Kinker ↑ Thrombin cleavage site

The overall strategy for production of the fusion protein construct is illustrated in **Figure 3.8**.

Figure 3.7 Alignment of 5' regions of rCNT1 and hCNT1 cDNAs



Locations of primers used for amplification of a region encoding a portion of the hydrophilic *N*-terminus of rCNT1 are shown.

Figure 3.8 Strategy for ΔrCNT1-GST fusion protein construction

Residues 226 to 399 of pSRH1 are amplified with the oligonucleotides CNT1_20*Bam*HI and CNT1_78*Eco*RI, represented by **A*** and **B*** respectively. This amplified fragment possesses 5' *Bam*HI and 3' *Eco*RI restriction sites in addition to a stop codon 5' to the *Eco*RI site. Parallel digestion of pGEX-KT and the PCR product with these enzymes provides two fragments of 4959 and 183 base pairs respectively which after ligation produce the construct pSRH2, containing the gene for the fusion protein ΔrCNT1-GST under the control of the *tac* promoter and inducible by IPTG.

3.4.4 Results and discussion

3.4.4.1 Production of ΔrCNT1-GST fusion protein construct

Amplification of a DNA sequence coding for residues 21 to 78 of rCNT1, and incorporating flanking restriction sites, was achieved using the strategy outlined in **section 3.4.3**. The overall strategy for production of a construct encoding an *N*-terminal fragment of rCNT1 fused to GST is illustrated in **Figure 3.8**. The conditions used for rCNT1 fragment amplification are indicated below.

Reaction mix:

10x Amplification buffer	10 µl
10 mM $MgSO_4$	1 µl
2 mM dNTPs	10 µl
Forward primer	100 pmol
Reverse primer	100 pmol
Template DNA	20 ng
Made up to 100 µl with sterile Milli-Q water	
Vent™ DNA polymerase	1 µl
(overlaid with 100 µl light mineral oil)	

The pSRH1 template was used in conjunction with CNT1_20*Bam*HI and CNT1_78*Eco*RI primers. Vent™ DNA polymerase (NEB) is a high-fidelity thermophilic polymerase with fidelity 5-15 fold higher than that observed with *Taq* DNA polymerase. This fidelity is derived in part from an integral 3'→5' proof-reading exonuclease activity.

Protocol:

1.	96 ^{0}C	4 min	1 cycle
2.	96 ^{0}C	30 sec	25 cycles
	55 ^{0}C	45 sec	
	72 ^{0}C	30 sec	
3.	72 ^{0}C	2 min	1 cycle

Amplification using these conditions yielded a PCR product of the desired size of 195 bp. The fact that no other fragments were produced (results not shown) indicated that the conditions chosen for the reaction were suitable and that the primers had been designed efficiently.

Restriction digestion of pGEX-KT with *Bam*HI and *Eco*RI excised a 10 bp fragment from the parent plasmid producing a 4959 bp linearised DNA fragment. This fragment was subsequently dephosphorylated by CIP treatment (**section 2.5.6**) and the resultant 4959 bp product purified by gel extraction (**section 2.5.5**). The PCR product was restriction digested in the presence of the same two enzymes and the resulting 183 bp fragment was purified from the terminal two small fragments using Qiagen PCR columns as indicated by the manufacturers. Ligation of the two resulting DNA fragments was performed as described in **section 2.5.7**. 10 µl of this ligation mixture was used the following day to transform 100 µl of competent *E. coli* DH5α cells, prepared and transformed as described in **section 2.5.4**. This transformation resulted in no colonies on the either water transformed or vector

minus insert control plates but several hundred were obtained on each of the sample plates. (The 761 bp fragment has been visualized when sufficient loading occurred.)

30 colonies from one sample plate were screened for the presence of an insert encoding the rCNT1 fragment by PCR using the reaction conditions described previously in this section but scaled down to a final volume of 20 µl. For this screening procedure, one colony was lifted from the agar plate with a sterile P10 tip and swirled in the reaction mix for 5 sec. The colony was subsequently cultured by placing the tip placed in 1.5 ml of LB broth, containing carbenicillin, and incubating overnight at 37^0C with agitation. Meanwhile the reaction mix was overlaid with 20 µ l of mineral oil and subjected to the same amplification program as used in the original fragment amplification described above. The results of this screening are illustrated in **Figure 3.9a**. Five of the 30 colonies tested (colonies 9, 13, 22, 23 and 30 in **Figure 3.9a**) yielded PCR products of the desired size of 195 bp, indicating the presence of an insert encoding the *N*-terminal fragment of rCNT1.

To confirm the reliability of the PCR screening the positive colonies were further analyzed by restriction enzyme digestion singly with *Bam*HI, *Eco*RI, or *Eco*NI and doubly with *Bam*HI and *Eco*RI digestion. Digestion with *Eco*NI was chosen since the presence of a successful construct would result in elimination of a 761 bp fragment due to double digestion. However, the lack of an insert would result in single digestion with no resultant change in size of the plasmid. **Figure 3.9 b** shows the results for colony 9 but similar results were obtained for the other 4 PCR positive colonies. Lanes 3 and 4 show the linearised construct produced a band of 5142 bp as predicted; lane 5 indicates that digestion with *Eco*NI produced a fragment of 4381 bp and another of 761 bp is produced but is not visible in the figure due to under loading

of the sample; lane 6 demonstrates the elimination of the insert with the double digestion of *Eco*RI and *Bam*HI gave the 4959 bp fragment of the parent pGEX-KT vector. These findings confirmed that ligation of the PCR fragment into pGEX-KT had been successful. Finally, the construct from colony 9 was sequenced as described in **section 2.5.7** confirming that the PCR-amplified sequence was the desired product and that no errors had been introduced. This construct was designated pSRH2 and was used for the subsequent *in vivo* expression of the encoded fusion protein in the *E. coli* strain DH5α.

3.4.4.2 In vivo expression of fusion protein

Expression of the fusion protein within the *E. coli* strain DH5α is under the control of the *tac* promoter. This is a hybrid promoter engineered from the *lac* promoter and the *trp* promoter in order to provide strong ribosomal binding for efficient translation while keeping transcription tightly regulated by the *lac* I^q repressor, also encoded in the vector. During the initial stages of culture growth the expression of the fusion protein is repressed. However, once the culture has reached the log phase on the growth curve, IPTG is added to release the promoter from the inhibitory effect of the *lac* I^q gene product and thus allow induction of the fusion protein expression. After an incubation period of several hours the culture is processed and tested for fusion protein expression.

Prior to carrying out a large scale preparation of the fusion protein a small scale pilot preparation of the fusion protein was carried out as described in **section 2.6.1**. Following growth and induction with IPTG, or culture under identical conditions in the absence of IPTG, the cell lysates were analysed by SDS-PAGE. **Figure 3.10**

shows an additional band with an apparent M_r of 33 000 was present in the induced sample but not present in the uninduced, lanes 2 and 3 respectively. However, in addition to the desired product with an apparent M_r of 33 000, a protein with an apparent M_r of 26 000 was also expressed in the induced sample, an equivalent size to the GST carrier protein. This finding appeared to reflect the proteolytic degradation of the fusion protein to produce GST, because the addition of a higher concentration (1.5 mM) of the protease inhibitor cell lysis resulted in elimination of this band (results not shown). This now demonstrated that the complete fusion protein was being expressed successfully in DH5α. Therefore indicating that large scale expression was viable. Subsequently, large scale expression of the ΔrCNT1-GST fusion protein was performed as described in **section 2.6.2** with the resulting cell lysate being affinity purified using an immobilized-GST column. Following washing of the column free glutathione was used to elute the fusion protein, resulting in a yield of 17.5 mg from an initial I l overnight culture. This provided sufficient amounts of the 33 kDa fusion protein (**Figure 3.10**, lane 5) to allow further studies, namely immunization.

3.4.4.3 Antibody production, purification and analysis

The use of fusion proteins in the development of antisera is a common technique with the GST system being one of the most frequently used, as applied in this case. Having successfully produced the fusion protein, immunization and antibody generation can be undertaken. At this stage a number of options were available for the generation of anti-rCNT1 *N*-terminal antibodies. Firstly, the complete fusion protein could have been used but this would open up the opportunity for the generation of anti-GST antibodies in addition to anti-rCNT1 antibodies. Secondly,

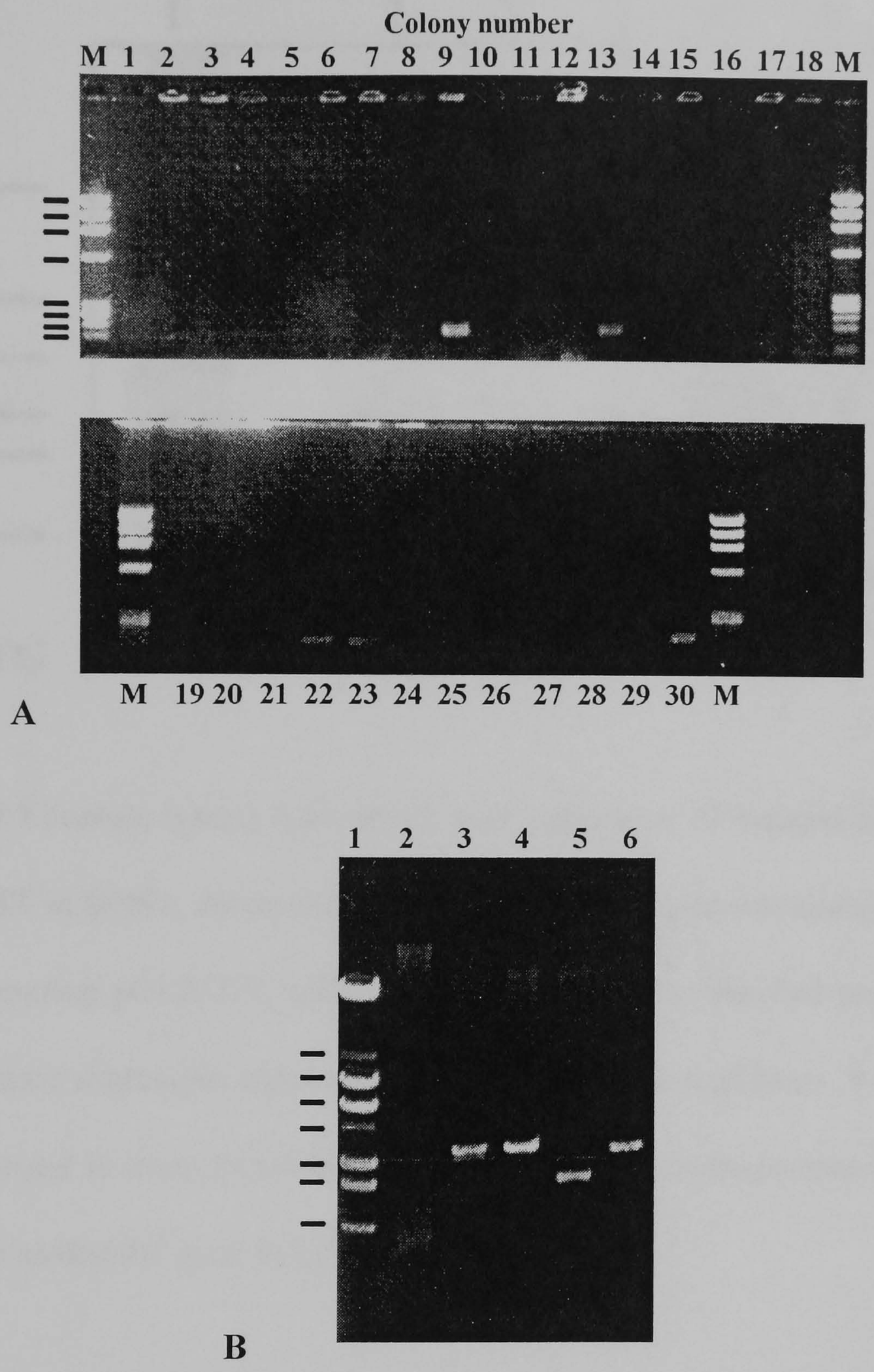
the *N*-terminal fragment could have been used for immunization by cleavage from the fusion protein to produce a 60 residue peptide of which 58 residues were from rCNT1 and 2 from the *C*-terminal region of the RGS thrombin cleavage site. Finally, the cleaved rCNT1 fragment could have been coupled to another carrier protein such as ovalbumin, as was the case for the synthetic peptides. This approach would certainly generate antibodies against the carrier protein as was suggested in the first option. Considering each alternative the former method was adopted based upon GST being reported by Pharmacia to have a low immunogenic response.

The intact ΔrCNT1-GST fusion protein containing 58 residues of rCNT1 fused to the soluble GST protein was used to perform immunization as described in **section 2.3.1** where 200 μg of the fusion protein was used in the initial immunization followed by boosting regularly using 100 μg. After a period of 2 months a sample bleed from rabbit A060 was used in an ELISA which resulted in a titre of 1:1,000,000 against the fusion protein. Further boosting was carried out monthly, the terminal bleed occurring 4 months after the initial immunization. ELISA showed that the titre of this final sample had decreased to 1:400,000. Such a decrease in titre is observed occasionally and its cause is unknown.

Western blot analysis against the fusion protein and GST with the polyclonal antiserum unfortunately demonstrated a significant cross-reactivity with GST was present (data not shown). Therefore it was necessary to affinity purify the anti-rCNT1 *N*-terminal antibodies. This was performed in two stages, firstly the antiserum was passed down an immobilized GST column, as described in **section 2.6.4**, to remove any anti-GST antibodies. Secondly, anti-rCNT1 antibodies were purified from the anti-GST antibody-depleted serum by passage down an

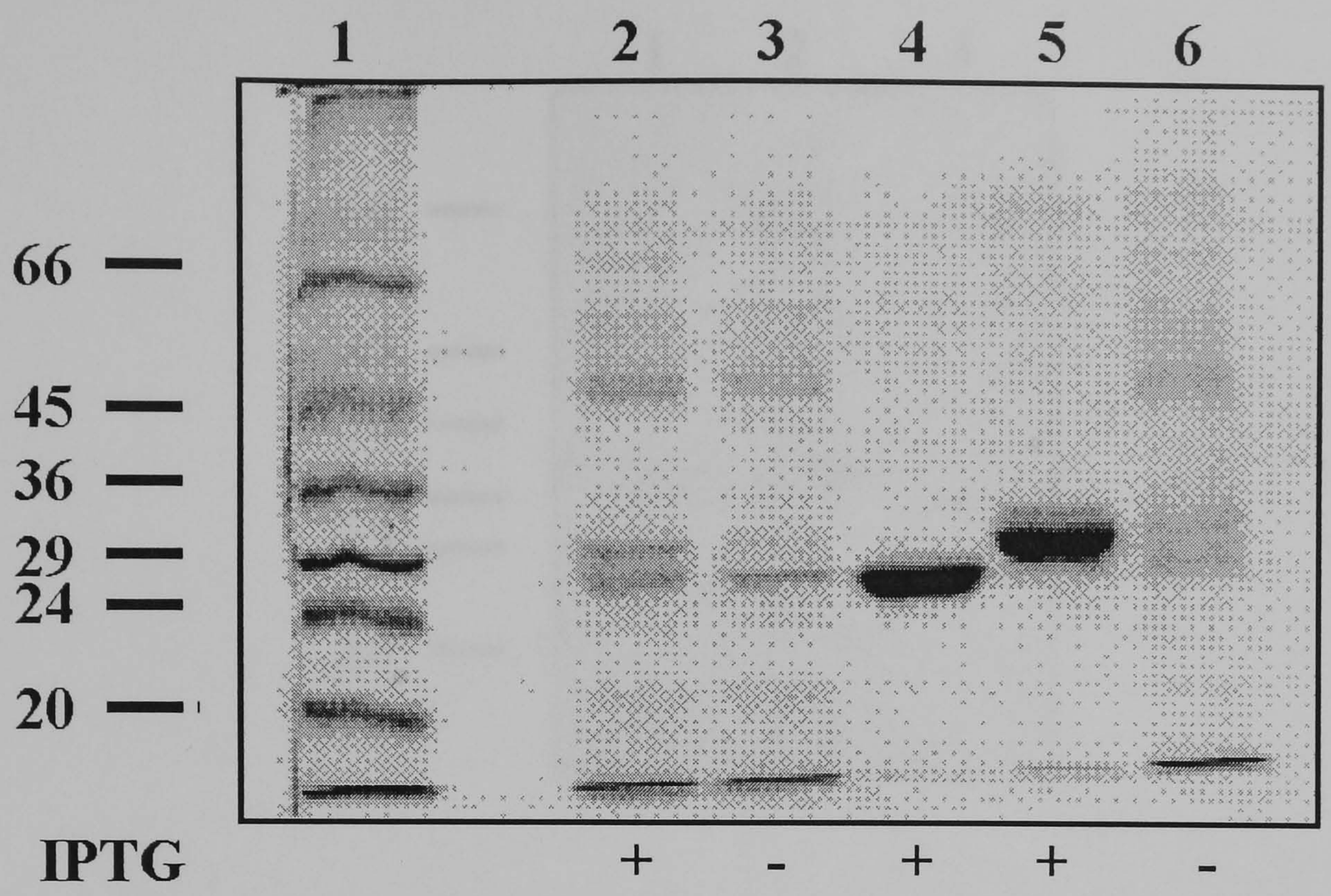
immobilized ΔrCNT1-GST fusion protein column and eluted with alkali, as described in **section 2.3.2**. Western blotting of this affinity-purified antibody (**Figure 3.11**) indicated that although the majority of the non-specific antibodies had been removed, as demonstrated in lane 2, there was still cross-reactivity of the affinity purified antibody with GST, lane 3. However, the amount of GST protein added to the gel was 10-fold more than fusion protein, and considering that the signal obtained was significantly lower than that against the fusion protein, indicated that the antibody had been significantly purified. With the aim of totally eliminating the anti-GST antibodies that remained the affinity purified antibody was passed through the immobilized GST column a second time but on Western blot analysis showed no significant improvement over the first round purification. Therefore having produced an antibody that cross-reacts with two different proteins an anti-GST control will have to be carried out in parallel whenever this antibody is used. Interestingly, the affinity-purified antibody also cross-reacted with the uninduced sample, lane 1, giving a signal at an equivalent position to the fusion protein. This result suggests that there has been some slippage of the promoter, allowing some expression of the fusion protein prior to induction with IPTG.

Figure 3.9 Screening and confirmation of the ΔrCNT1-GST construct



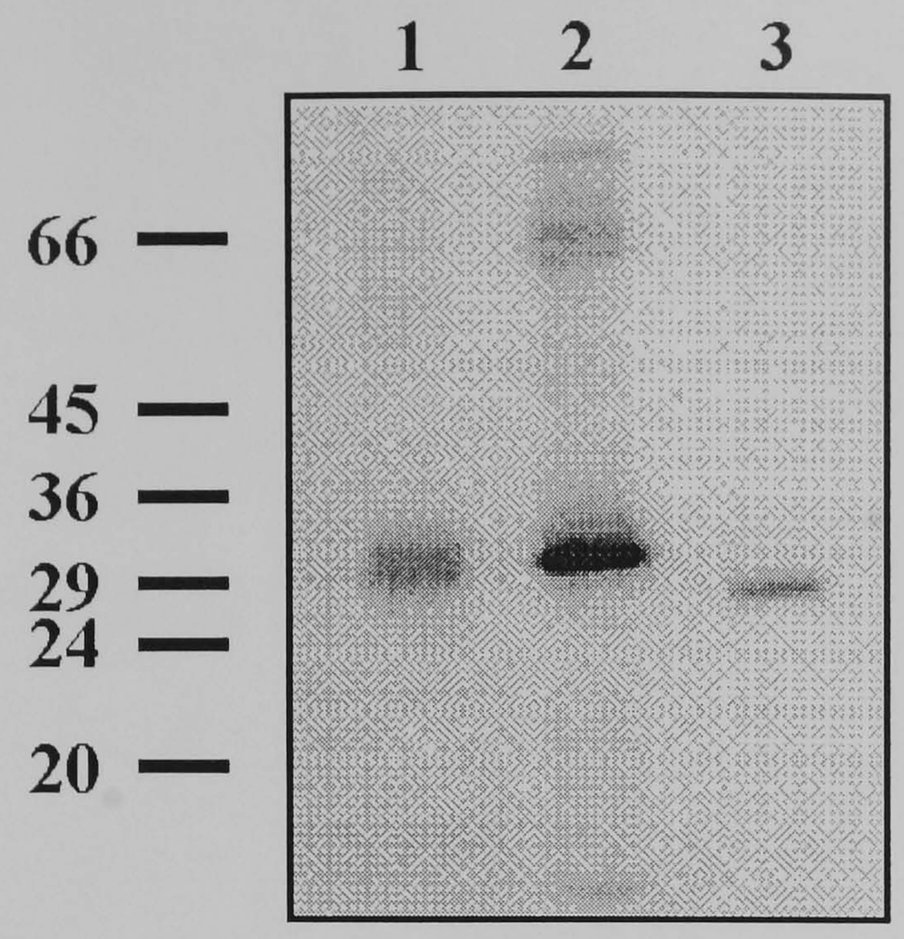
(A) PCR screening of ΔrCNT1-GST transformed colonies. ϕX174 DNA/ *Hae*III markers are present in lanes marked M. Their mobilities are represented by the bars located at the side of the image, which correspond to sizes of 1353, 1078, 872, 603, 310, 281, 234 and 194 base pairs, from the top down. (B) Restriction digestion analysis of colony 9. Lane 2 undigested control, lane 3 *Eco*RI digest, lane 4 *Bam*HI digest, lane 5 *Eco*NI digest, lane 6 *Eco*RI/ *Bam*HI digest. The mobilities of λ DNA/ *Bst*EII markers, lane 1, are represented at the side of the image by bars corresponding to 8454, 7242, 6369, 5686, 4822, 4324 and 3675 base pairs respectively.

Figure 3.10 SDS-PAGE analysis of ΔrCNT1-GST expression in *E. coli*



Lanes 2 and 3 contain lysates from small scale expression of induced and uninduced ΔrCNT1-GST in DH5α, respectively. Lane 4 contains lysate containing GST from a culture harbouring pGEX-KT, while lanes 5 and 6 contain purified protein samples from large scale expression under induced and uninduced conditions. 9 μg of sample protein is present in lanes 2 to 6. Molecular weight markers are present in lane 1 with their relative molecular mass in kDa indicated.

Figure 3.11 Immunoblot analysis with anti-ΔrCNT1 purified antibodies



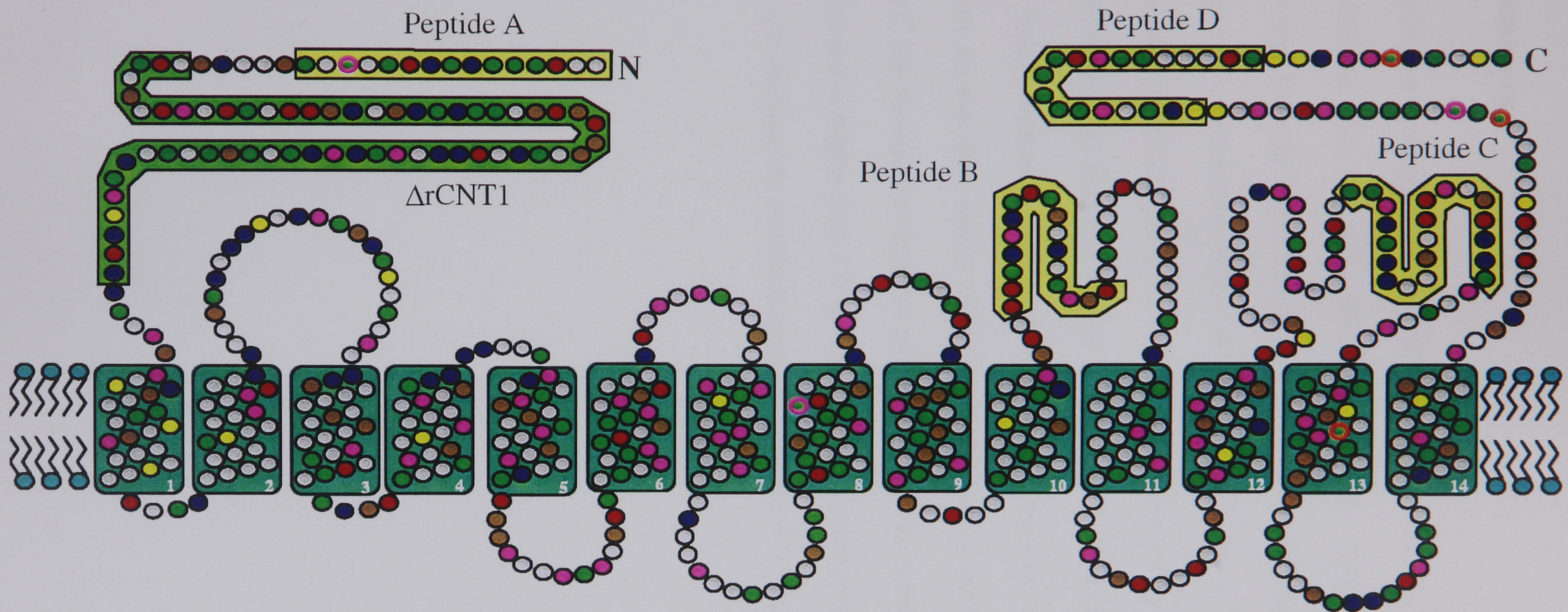
Affinity purified antibody (4 μg/ml) used to probe uninduced and induced ΔrCNT1-GST in *E. coli*, lanes 1 and 2 respectively. Lane 3 represents the signal obtained with the GST vector protein. Lanes 1 and 2 contained 0.1 μg while lane 3 contained 1 μg total protein. Molecular weight markers are represented at the side of the image with their relative molecular mass in kDa indicated.

Figure 3.12 Location of antibodies raised within rCNT1

Image on opposite page

Regions within rCNT1 to which site-directed antibodies have been generated by the synthetic peptide approach are indicated by a *yellow background box* while that generated with the fusion protein approach is illustrated by a *lime background box*.

For further information on components of the diagram refer to **Figure 1.2**.



Key:

- General non-polar
- General polar
- Positive
- Negative
- Aromatic
- Cysteine
- Putative *N*-linked glycosylation site
- Putative *O*-linked glycosylation site
- Structure breaker

3.5 Summary

Four antibodies have been raised against synthetic peptides corresponding to residues 1-15, 389-406, 505-524 and 618-637 of rCNT1, designated peptides A-D respectively. These have been affinity purified with the resulting antibodies showing varying levels of reactivity with rCNT1 peptides. Additionally, a fifth antibody towards residues 21-78 of rCNT1 (ΔrCNT1) was generated by the production of a GST fusion protein. However, affinity purification of this antibody was unable to totally remove antibodies raised towards the GST component of the fusion protein. Therefore an anti-GST control should be carried out in parallel when using this antibody. The locations of these five antibodies within the putative model of rCNT1 are illustrated in **Figure 3.12**. The generation of these antibodies now provides a tool which will facilitate the characterization of expressed rCNT1 (see **Chapters 4 and 5**).

Chapter 4

Distribution of rCNT1 in rat tissues

4.1 Introduction

Nucleosides play a variety of physiological roles in many mammalian tissues, as described in detail in **section 1.2.1**. Briefly, they are involved in neurotransmitter release, cardiac contractility, vasodilation, lipolysis and platelet aggregation [23]. Additionally, in cell types that lack the pathways for *de novo* nucleotide synthesis, such as haematopoietic stem cells, nucleoside salvage is essential for the production of nucleotides in the quantities required for intermediary metabolism and specialized cellular functions [33]. For example, inosine is an *in vivo* energy source for adult pig erythrocytes, cells that are unable to transport and hence metabolize glucose [24]. Therefore, due to the importance of these molecules it is expected that most if not all mammalian cells possess at least one type of nucleoside transporter. For example, in the rat *cif*, *cit* and *cib* transporters have been shown to be present in the kidney [63,72,75]; in hepatocytes and macrophages *cif* transport has been demonstrated [61,142]; and in the rabbit concentrative transport has been displayed in the choroid plexus, small intestine and the kidney [52,64,74]. These are to name but a few, however for a more in-depth discussion see the review by Cass [49].

An individual cell is unlikely to express all members of the facilitative and active nucleoside transporter families, and those expressed may reflect the particular metabolic requirements of the cell. Moreover, where two or more transporters are simultaneously present, these may occupy different regions of the cell surface, in particular in polarized cells. A comparable situation in the better understood field of glucose transport is presented by the epithelial cells of the small intestine. The polarization of these cells provides two contrasting environments, namely that of the

intestinal lumen and that of the interstitial fluid. Considering the facilitative glucose transporters, GLUT's 1-7, and the active transporter SGLT1, only GLUT2 and SGLT1 have been demonstrated to be present, located at the basolateral for the former and brush border membranes for the latter [143,144]. Additionally GLUT5 has been demonstrated to be inducibly present at the brush border, although it is primarily concerned with fructose uptake [145]. Therefore it can be assumed that not every cell will possess all the nucleoside transporters although if they do the distribution within the cell may vary depending on the local environment present. Furthermore, the expression of these proteins may not be permanent but may be regulated depending on the stage of cellular development or by the receiving of stimulatory or inhibitory signals. One such case showed that hypothyroidism reduces the adenosine transport in rat brain stem synaptosomes by decreasing the number of both the *es* transporters and A_1-receptors present [146], the latter of which may regulate some, if not all, nucleoside transporter activity. However, factors regulating the expression of such transporters have not been investigated to date extensively.

The expression of rCNT1 in the small intestine and kidney has been demonstrated by Northern blotting [81]. Here it is thought to be involved in the intestinal absorption of nucleosides from the diet and in renal salvage of nucleosides from the renal filtrate. Although the brain, heart, liver, lung, skeletal muscle, spleen and testis were also tested no signal on Northern blotting was observed. This however is not definitive proof that rCNT1 is not expressed in these or other tissues. Since Northern blotting is the detection of mRNA, which is very unstable and thus its existence may be short lived prior to being degraded, protein may exist depending on both the initial levels of expression and the turnover rate of that protein within the cell. Therefore it

is possible that if the protein is not constitutively expressed it may be regulated, with the protein being expressed either due to developmental changes or environmental conditions, as mentioned above, in which case rCNT1 may be present without the existence of its mRNA. Therefore, based on this information the aim of the work described in this chapter was to screen for the expression of rCNT1 in adult Wistar rats using antibodies as a means of detecting the presence of the protein. Detection was two-fold with initial screening by Western blotting to indicate in which tissues the protein was present, followed subsequently by immunocytochemical detection to confirm these results and further identify the subcellular localization of the protein. For the purposes of understanding the approaches used and results obtained this technique of immunocytochemistry is described in the following section.

4.2 Immunocytochemistry: an introduction

The technique of immunocytochemistry was developed by Coons *et al.*. who first labeled an antibody with a fluorescent probe and used this to identify an antigen in tissue sections [147,148]. This method of directly labeling the antigen-specific antibody was further improved by Coons *et al.* by the introduction of a second layer [149]. The primary antibody in this procedure is unmodified and is specific for the desired epitope while a secondary antibody is directed against the Fc fragment of an IgG raised in the same host as the primary antibody. Thus conjugation of a fluorophore to this secondary antibody allows indirect labeling of the desired epitope. The advantages of this system are three-fold; firstly, the anti-IgG serum used as a second layer usually has a very high avidity and is hyper-immune; secondly, at least two secondary layer antibodies can bind to each primary antibody thus amplifying

the signal; and thirdly, by only labeling one general secondary antibody and not each specific primary antibody the process is made more economical. Furthermore, developments in indirect labeling have resulted in tertiary level systems, such as the peroxidase-anti-peroxidase (PAP) system [150] or the avidin-biotin methods [151], but these require more time due to the incorporation of an extra layer and thus may be more expensive so will not be discussed but are reviewed by S.J. Carlton [152].

Successful immuno-staining requires a number of conditions to be met. The tissue antigens must be made insoluble and yet their antigenic sites must be available to allow antibody binding. Concurrently the tissue architecture must be preserved so that the immuno-reactive target may be identified in context. Originally fixation of tissue to maintain its structure was thought to reduce antigen availability, due to the strong cross-linking of tissue proteins by the conventional aldehyde fixatives. However advances in the antibody production, more rigorously controlled techniques and the correct pretreatment of samples have all allowed for such fixatives to be used in some cases, but not in all.

Several types of fixative are available but it is best to start by following published methods with fixatives which have worked for similar tissues or molecules and if necessary adapt these to the antigen in question. Fixatives available include, cross-linking agents which react with reactive end-groups of adjacent protein regions, such as formaldehyde which reacts primarily with amine groups of lysine; or precipitant fixatives which denature proteins by destroying the hydrophobic bonds which hold together the tertiary structure, such reagents are acetone or alcohols. Generally cross-linkers work better with smaller soluble proteins while precipitants work better with larger molecules. Therefore it may be beneficial to use a cocktail of both.

Alternatively, fresh samples may be frozen or freeze-dried in order to maintain the antigen. Furthermore, fixed or unfixed specimens can be embedded to maintain tissue structure while sectioning by providing support to reduce distortion due to the abrasion and pressure of the microtome blade. However, embedding in certain substances, such as paraffin wax, can itself reduce reactivity of the antibody with the antigen.

Recent advances have been made in the recovery of epitopes in tissues fixed with aldehyde fixatives. In this form of fixation the formation of hydroxy-methylene bridges between amino acid chains of proteins may mask the antigenic site. In the quest to overcome this problem a number of pretreatment processes have been developed and collectively termed antigen retrieval techniques. One such technique involves simply washing the tissue to reverse the effects of fixation, while a second method involves incubating the sections in a solution of protease to break down the cross-linking bonds between the fixative and the protein [153]. For the latter of these methods the concentration of the enzyme, the incubation time, thickness of section and temperature are all important and must be controlled tightly. The most common enzymes used are trypsin and pronase, although the latter is more harsh on the tissue integrity. Another method involves the use of heat and is therefore termed heat-mediated antigen retrieval. Within this category there are a number of methods which have been reviewed by Taylor *et al.* [154]. Briefly, these include heating sections either in an oven, microwave or pressure cooker in the presence of a heavy metal solution or chelating solution such as citrate or EDTA. An explanation on how these procedures might work was put forward by Morgan *et al.* [155]. These workers suggested that not only did heating provide energy to rupture the hydroxyl bonds

formed by aldehyde fixation but it also released tissue-bound calcium ions which contribute to tighter bonds with the fixative, thus exposing some antigenic sites previously masked. Considering these three pre-processing techniques for antigen retrieval it is possible that epitopes previously hidden may now be recovered, although even these techniques have not proven universally successful.

Another important consideration is how the antigen is to be visualized. For this there are a number of methods available with the label conjugated at either the primary, secondary or tertiary levels. Fluorescent labels are the most commonly used due to their immediate reporting when excited at the correct wavelength. Numerous such labels are available with the most common being fluorescein isothiocyanate (FITC) and tetrarhodamine isothiocyanate (TRITC) which fluoresce green and red respectively. Alternatively, enzyme labels can be used which produce a permanent result, as opposed to the unstable fluorophores, by conversion of soluble substrates to insoluble residues which can be observed with a standard microscope. An example of this method is the use of horseradish peroxidase, which in the presence of hydrogen peroxidase and a chromogen, such as diaminobenzidine, produces a brown resin at the site of the antigen. A further method of labeling is with colloidal gold particles conjugated to the secondary or tertiary antibody which allows detection using an electron microscope. For a more detailed description of these labeling techniques and others refer to the review by Polak and Van Noorden [156].

4.3 Methods for tissue localization

4.3.1 Membrane preparation and Western blotting

Total cell membranes from the brain, heart, jejunum, kidney, liver, skeletal thigh muscle, spleen and testis were prepared as described in **section 2.4.1** from a packed tissue volume of 20 ml. It was noted, particularly with the jejunum, that it was necessary to keep the samples on ice throughout the processing in addition to using protease inhibitors to minimize degradation of the membrane proteins. Prior to snap freezing and storage of the samples at -70°C an aliquot was removed for protein determination by the BCA method described in **section 2.4.2**.

Western blot analysis was performed by the electrophoresis of 70 μg of each membrane preparation by SDS-PAGE analysis on a 10 % polyacrylamide gel followed by transfer to a nitrocellulose membrane and treatment as described in **sections 2.4.3** and **2.4.4**. Affinity purified antibodies raised against synthetic peptides were used at 2 μg/ml, unless stated otherwise, while antisera raised against the ΔrCNT1-GST fusion protein was applied at a 1:1000 dilution or at 2 μg/ml when affinity purified. The Enhanced Chemical Luminescence (ECL) kit from Amersham was used in detection in conjunction with blue sensitive autoradiography film.

4.3.2 Tissue extraction and processing for immunocytochemistry

Tissues required were obtained from male Wistar rats (200-250g) by cervical dislocation and subsequent dissection. This was performed as rapidly as possible to reduce any tissue degradation. Once removed each tissue sample was washed with PBS and then embedded in OCT compound, frozen in isopentane and stored at -70°C. Alternatively, samples were incubated 16 h in 4 % paraformaldehyde at 4°C

subsequently followed by incubation for 16 h in 0.5 M sucrose in PBS at 4°C. Such fixed tissue samples could either be embedded in OCT compound and frozen as before or paraffin-wax embedded. Additionally, whole kidneys were perfused by Miss J. Smith with PBS prior to perfusing with 4 % paraformaldehyde in PBS until the tissue surface was evenly blanched. The kidney was subsequently removed for o/n incubation in 0.5 M sucrose at 4°C prior to freezing in isopentane and storage at -70°C.

Sectioning of unfixed and paraformaldehyde-fixed OCT-embedded tissues to yield sections of thickness 8 microns was performed using a Leitz Wetzlar Kryostat 1720 pre-chilled to -18°C. Cut sections were lifted onto Vectabond™-coated slides. These were subsequently stored at -20°C until required although immediate usage was preferred. Paraffin-wax embedded tissues were sectioned at r.t. to a thickness of 4 microns and lifted onto untreated glass slides. Prior to use these sections were dewaxed by washing twice in xylene for 3 min each followed by subsequent rehydration by sequential washing twice in 100 % ethanol, once in 90 %, once in 80 %, once in 70 % and twice in PBS for 3 min each.

4.3.3 Antigen retrieval pretreatment of formaldehyde fixed tissue

Four methods of antigen retrieval were used independently. Tryptic digestion was performed using 0.1 % w/v trypsin in PBS for 1 h at 37°C [153]; wet autoclave in 0.01 M sodium citrate, pH 6, for 5 min [157]; conventional oven heating at 70°C in 0.01 M sodium citrate, pH 6, overnight [158]; and microwave oven heating in 0.1 M sodium citrate, pH 6, for 10 min [159].

4.3.4 Immunocytochemistry of unfixed frozen tissue

Sections were allowed to air-dry for 1 h and fixed with acetone for 3 min at r.t.. Subsequently sections were washed twice for 10 min with PBS and blocked for 1 h with 100 µl blocking buffer (5 % w/v bovine serum albumin (BSA) in PBS containing 0.1 % Tween-20 (PBST)) in a humidified chamber at r.t.. The blocking buffer was then replaced by addition of 100 µl blocking buffer containing immune or pre-immune antisera at a dilution of 1:100; or either affinity-purified antibody or non-specific rabbit IgG at 10 µg/ml. Following overnight incubation in a humidified chamber at 4°C, sections were washed 3 times for 10 min each with PBST. Anti-rabbit IgG FITC-conjugate (Sigma) was then added at a dilution of 1:100 in blocking buffer and incubated for 1 h at r.t.. Following washing 3 times for 10 min each in PBST the sections were mounted with Vectashield™ anti-fade mountant and sealed with DPX. Confocal microscopy was performed using an inverted Nikon Diaphot adapted for confocal laser scanning microscopy.

4.3.5 Immunocytochemistry of paraformaldehyde-fixed tissue

Fixed OCT- or rehydrated wax-embedded sections were washed twice for 5 min each in PBS followed by washing twice for 5 min each with 1 mg/ml sodium tetra-borohydride to reduce auto-fluorescence. Subsequently, sections were either treated as for unfixed tissue, as described above, or blocked using an avidin/ biotin blocking kit (Vector Laboratories Inc.). With the exception that the antiserum raised against the ΔrCNT1-GST fusion protein was used at a dilution of 1:400. Subsequent steps were carried out as for frozen sections until addition of second antibody where biotin-conjugated IgG was used in combination with FITC-conjugated Extravidin™ (Sigma) to amplify the signal, according to the manufacturer's instructions.

4.4 Results and discussion

4.4.1 Antibody usage

Immuno-blotting using antibodies raised against peptides A and B gave no signal in any tissue. Similarly, neither antibody yielded specific staining of any tissue when used in immunocytochemistry. Therefore these two antibodies will not be discussed in the following sections.

Since anti-peptide C and D antibodies did recognize rCNT1, as described in **section 4.4.3**, the former of these antibodies was used, either as antiserum or affinity purified antibody, in initial investigations aimed at optimizing methods of tissue preparation and immuno-staining. However, for subsequent tissue screening both anti-peptide C and D antibodies were used.

Antibodies raised against the ΔrCNT1-GST fusion protein were developed towards the end of this study and thus their use has been more limited, only being used on paraformaldehyde fixed kidney. The advantage that has occurred using these antibodies was that they had the ability to recognize tissues which had undergone mild paraformaldehyde fixation and thus provided sections with better integrity. The reason why these antibodies raised against the fusion protein proved more successful than the anti-peptide antibodies in recognizing the mildly fixed tissue containing rCNT1 was that the synthetic peptides on average were approximately twenty residues in length compared to the rCNT1 fragment of the fusion protein which was nearly three times this size. Therefore having raised the antiserum against a larger portion of rCNT1 it contained antibodies against more epitopes and thus some of these antibodies have proven to recognize mildly fixed tissue. However this

development was not achieved until much work had already been carried out on tissue preparation techniques, as described in the following section.

4.4.2 Fixation vs. fresh tissue

Initial immuno-staining of unfixed frozen rat jejunal sections indicated disintegration of the villi, predominantly at the villi tips (**Figure 4.1**). Therefore, suggesting that fixation of the tissue was necessary. **Table 4.1** summarizes the comparison of both unfixed and paraformaldehyde-fixed intestine after each stage of the immuno-staining process as observed by phase contrast microscopy. The results demonstrated that unfixed tissue was relatively fragile with significant disruption occurring during both sectioning and washing while paraformaldehyde-fixed tissue was much more stable, showing negligible tissue disruption. However on comparison of the end results of immuno-staining the sections fixed with paraformaldehyde demonstrated no staining above background levels in comparison to unfixed tissue which indicated specific staining at the enterocyte brush border. As mentioned in **section 4.2** paraformaldehyde cross-links adjacent amino acid residues in the same or neighboring peptides thus in some situations blocking the antigenic site. It was therefore reasoned that a lower degree of cross-linking might leave more epitopes available to react with the antibody. Therefore the period of time during which the tissue was exposed to the fixative was modified in order to test times of fixation ranging from 20 min to 16 h. Unfortunately, decreasing the time of fixation yielded a negligible increase in the intensity of staining while decreasing the tissue stability. In all cases no significant differences were seen between paraffin wax- and OCT-embedded paraformaldehyde-fixed samples.

As mentioned in **section 4.2** a number of techniques have been developed to improve the immuno-reactivity of paraformaldehyde-fixed tissue. In the present studies four such techniques were used, involving the methods detailed in **section 4.3.3**. The results are summarized in **Table 4.2**. In each case the integrity of the tissue remained good but very little improvement was observed in the specific signal obtained. Where there was an improvement in signal, as for both wet autoclave and conventional oven heating, this was very weak when compared to unfixed tissue. In order to increase this signal streptavidin-biotin amplification was performed but in addition to increasing the signal at the brush border the background signal was also correspondingly increased. Later attempts to reduce this background staining with differing blocking techniques and reducing the primary antibody concentration failed to eliminate this non-specific staining.

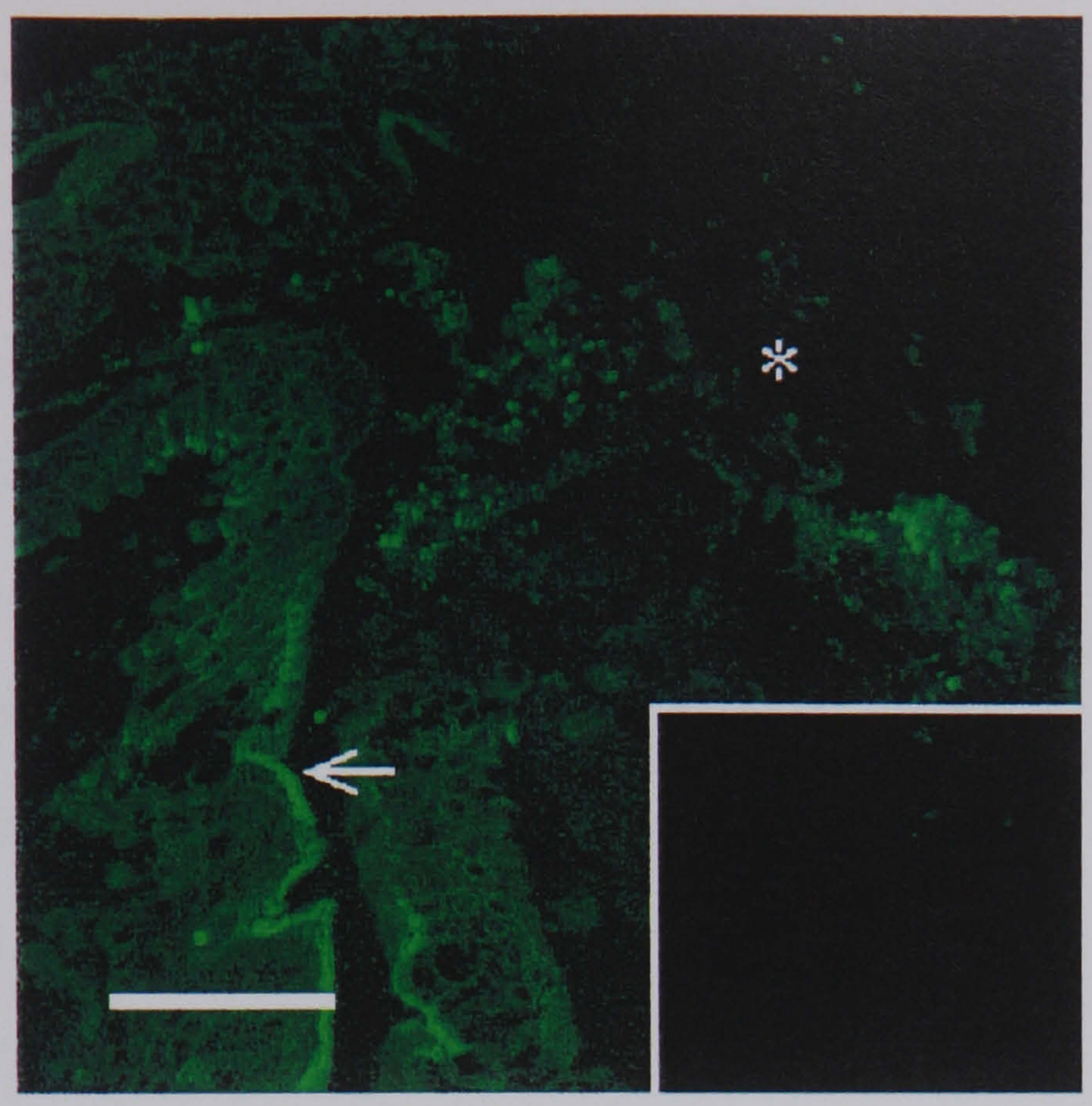
Although antigen retrieval has proved a useful technique with some antigens it is not always successful. Trypsin treatment by Huang *et al.* was found to decrease the nonspecific background fluorescence through the digestion of the tissue while possibly unmasking a number of epitopes [153]. Here it has been shown that the epitopic regions of rCNT1 investigated with such pretreatment proved unsuccessful in producing a specific signal with either paraffin wax- or OCT-embedded tissue. Similar results with the other attempts at antigen retrieval showed that rCNT1 appears to be resistant to such techniques. Comparable results have been found in other investigations with numerous other proteins. Tanimoto *et al.* tested 27 anti-lymphocyte antibodies with porcine sections using antigen retrieval to improve results but this only resulted in 5 samples being successful [160]. Similarly, Piffko *et al.* screened 22 fixed carcinomas with four antibodies using both wet autoclave and

microwave oven pre-treatment showing that recovery of the epitopic site was unpredictable, with at maximum only forty percent recovery in the best case [157]. They also concluded that wet autoclave gave better recovery. On the whole antigen retrieval does appear to be a useful technique by increasing the number of epitopic sites made available following paraformaldehyde fixation but does not always prove successful. However, a rapid assessment technique by Shi *et al.*, termed the 'test battery' approach, provides a strategy for testing a number of factors involved with antigen retrieval [161], including the use of buffered solutions at pH 1, 6, and 10 with three different heating conditions (temperatures 120°C, 100°C and 90°C) and would therefore prove to be a good starting point for further antigen retrieval studies.

Having shown that no significant recovery of antigenic sites could be obtained from paraformaldehyde-fixed tissue for the rCNT1 antigenic sites investigated, post-sectioning fixation of tissue was next tested on the reasoning that if disruption caused by sectioning could not be prevented then at least further disruption due to washing could. Therefore, following sectioning sections were exposed to PBS, methanol, ethanol, acetone, 4 % paraformaldehyde (in PBS) and calcium formol (4 % paraformaldehyde in PBS containing 1 % w/v calcium chloride) for 10 min at r.t. and subsequently air dried for 30 min prior to processing as described in **section 4.3.4**. Results from this investigation are summarized in **Table 4.3**. Although no post-fixation gave perfect integrity paraformaldehyde fixation was taken as the standard due to this giving the best fixation. Considering these results it is obvious that no treatment gives optimal fixation while allowing efficient staining so a compromise between both had to be reached. Considering this post-sectioning fixation with

acetone was adopted as a standard method for further studies, unless mentioned otherwise.

Figure 4.1 Integrity of unfixed tissue



Immunocytochemical analysis of an unfixed frozen rat jejunal section probed with antiserum raised against peptide D at a dilution of 1:100. rCNT1 localization at the brush border is indicated by the *arrow* while disintegration of the villus tip is marked by an *asterisk* (*). The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 40μm.

Table 4.1 Comparison of unfixed and paraformaldehyde-fixed tissue processing

Stage	Unfixed	Fixed
Tissue integrity:		
Post-sectioning	+++	++++
Washing	+++	++++
Blocking	+++	++++
Primary antibody	+++	++++
Washing	++	++++
Secondary antibody	++	++++
Washing	++	++++
Mounting	++	++++
Immuno-staining:	++++	-

Tissue integrity after each stage was assessed by phase-contrast microscopy setting paraformaldehyde-fixed tissue as the standard (++++), due to its providing the best tissue integrity. Immuno-staining was achieved after complete immuno-staining, with '-' indicating no specific staining.

Table 4.2 Comparison of antigen retrieval techniques

Pretreatment	Integrity	Immuno-staining
None (frozen tissue)	++	++++
None (PF fixed tissue)	++++	-
Tryptic digestion	++++	-
Wet autoclave	+++	+
Conventional oven	++++	+
Microwave oven	++++	-

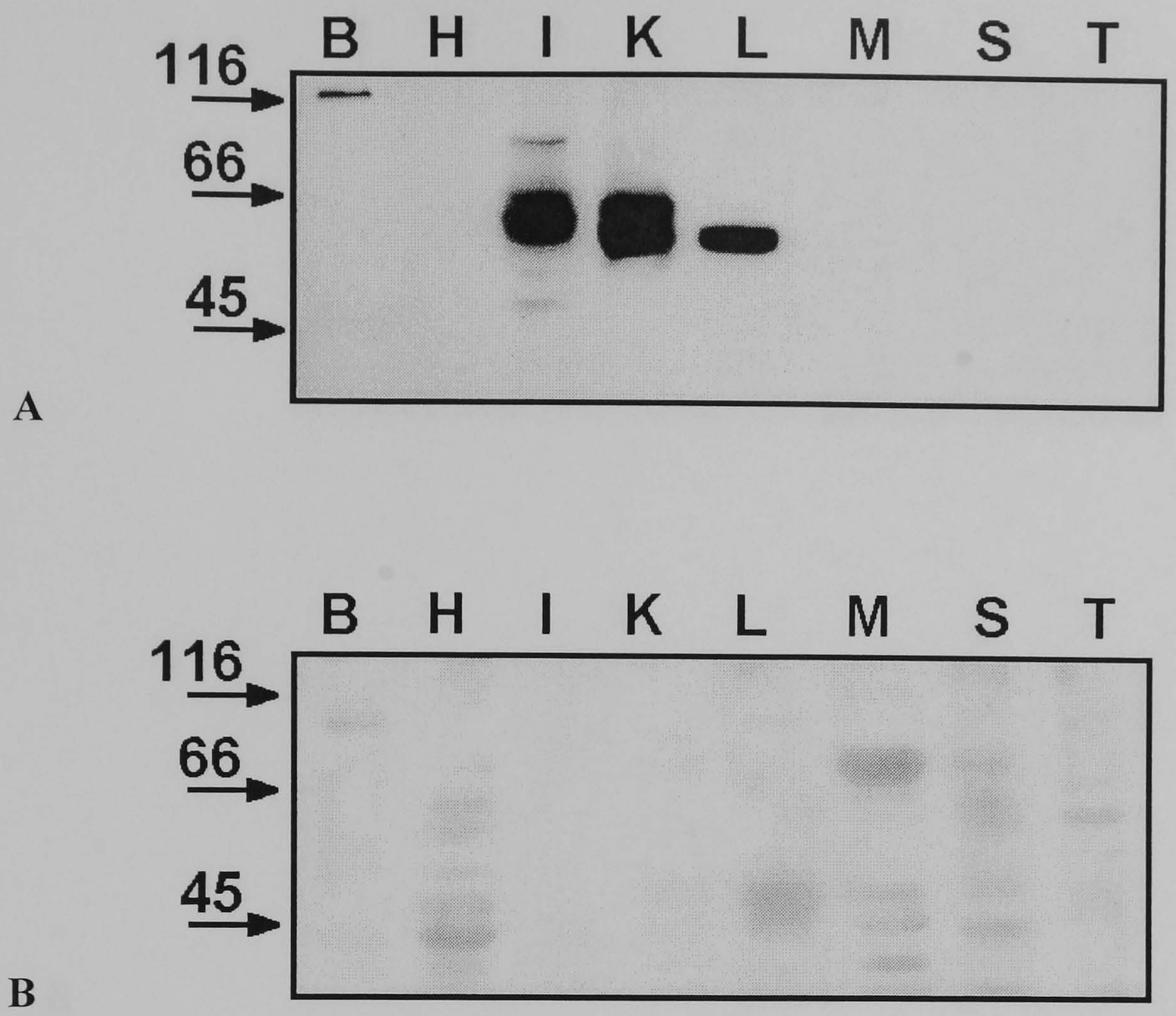
Paraformaldehyde (PF) fixed tissue was taken as standard (++++) for tissue integrity while unfixed tissue was taken as standard (++++) for immuno-staining. Where no signal was obtained '-' is used. Processing of tissue was performed as described in **sections 4.3.2-4.3.4**.

Table 4.3 Comparison of post-sectioning fixation reagents

Fixative	Integrity	Immuno-staining
PBS	++	++++
Methanol	+++	++
Ethanol	+++	+
Acetone	+++	+++
Paraformaldehyde	++++	-
Calcium formol	+++	-

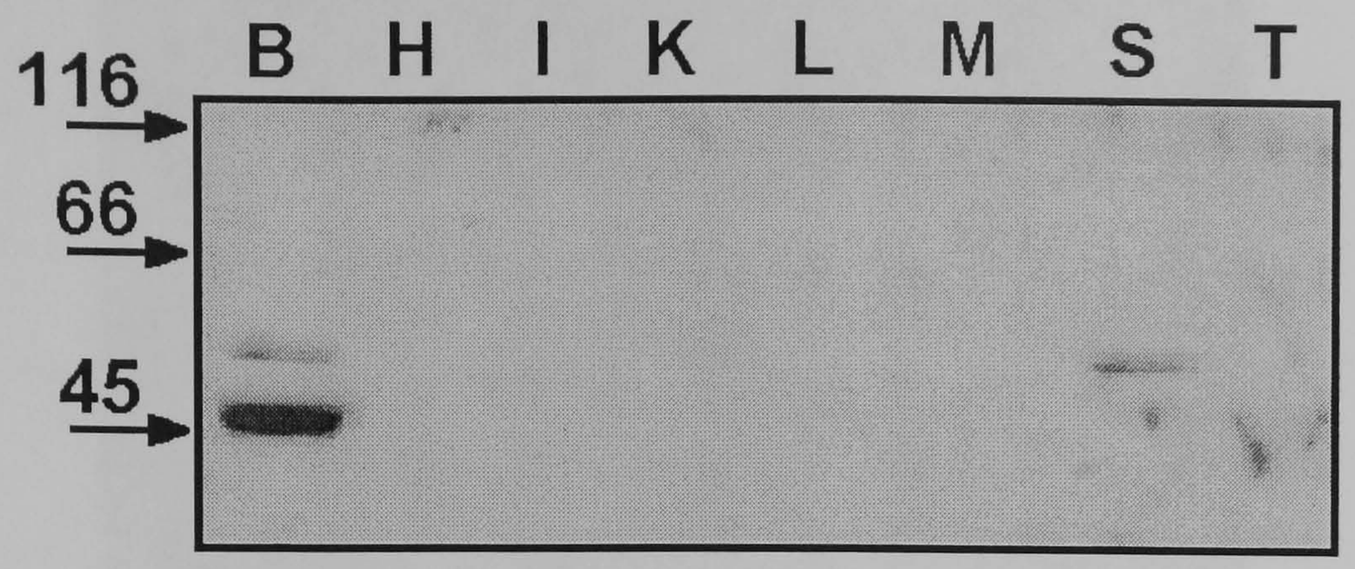
Although tissue integrity was not perfect due to sectioning disruption paraformaldehyde post-fixation gave the best tissue integrity and was thus used as a standard (++++). Unfixed, PBS treated tissue gave the best immuno-staining signal and was thus assigned as standard (++++). Where no stain was observed '-' is used.

Figure 4.2 Western blot analysis of rat tissues with anti-peptide C antibodies



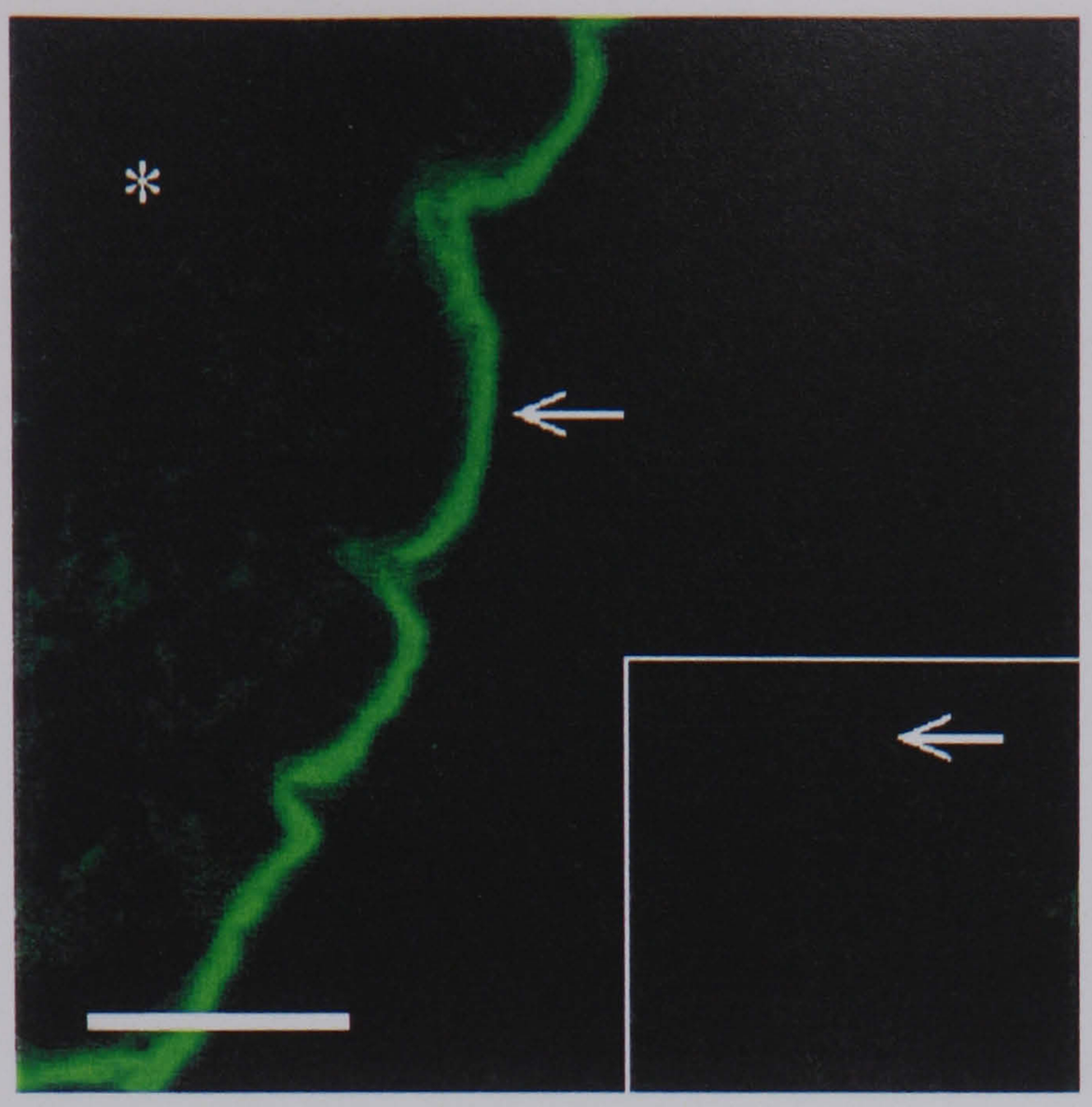
Analysis of rat brain (B), heart (H), intestine (I), kidney (K), liver (L), skeletal muscle (M), spleen (S) and testis (T) membranes containing 70 μg total protein using affinity purified antibodies (2 μg/ml) raised towards peptide C and non-specific rabbit IgG as a control, *plates* A and B respectively. Molecular mass markers on *left* are in kDa.

Figure 4.3 Western blot analysis of rat tissues with anti-peptide D antibodies



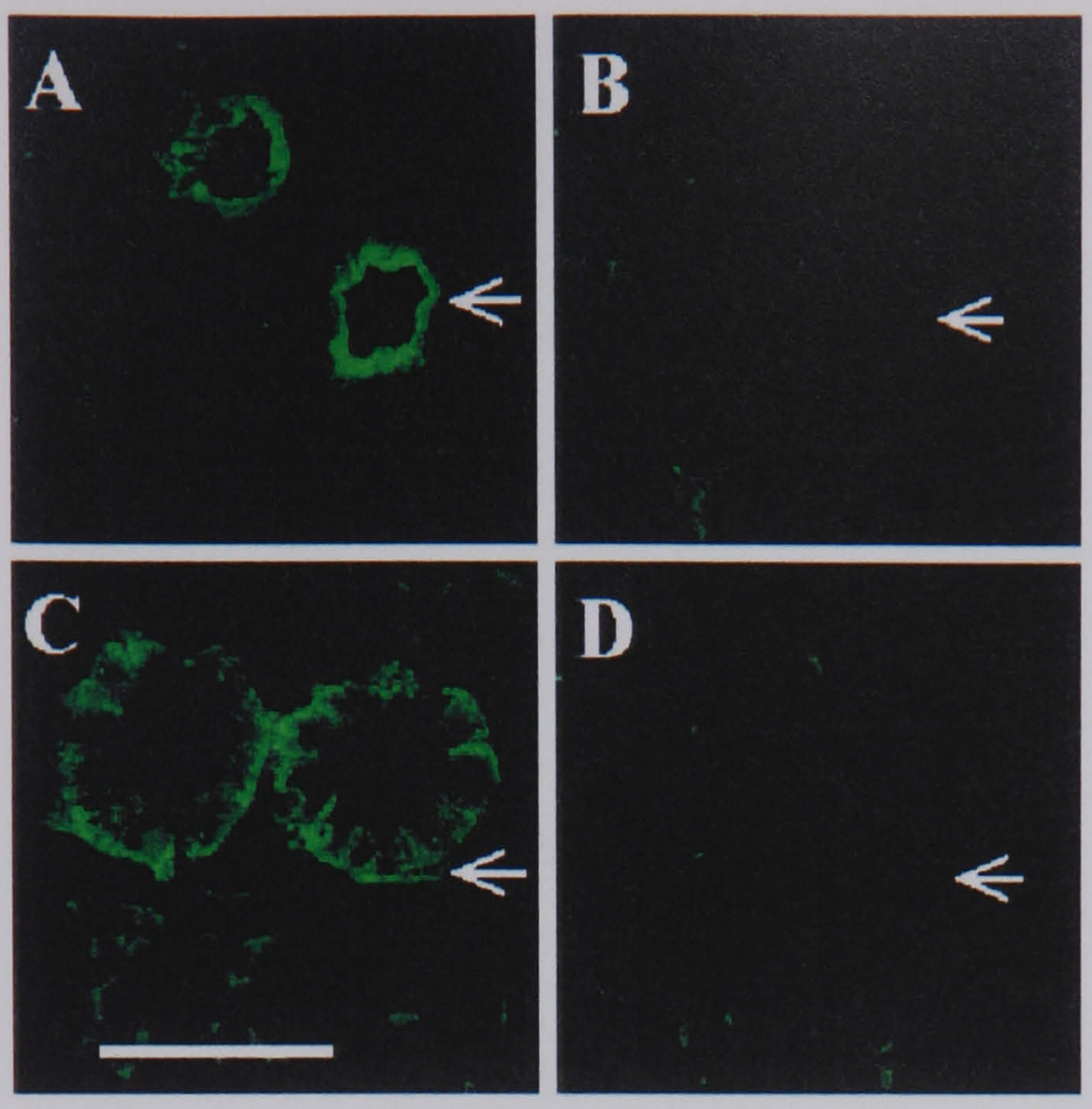
Analysis of rat brain (B), heart (H), intestine (I), kidney (K), liver (L), skeletal muscle (M), spleen (S) and testis (T) membranes containing 70 μg total protein using affinity purified antibodies (2 μg/ml) raised towards peptide D and non-specific rabbit IgG as a control, *plates* A and B respectively. Molecular mass markers on *left* are in kDa.

Figure 4.4 Subcellular distribution of rCNT1 in the rat jejunum



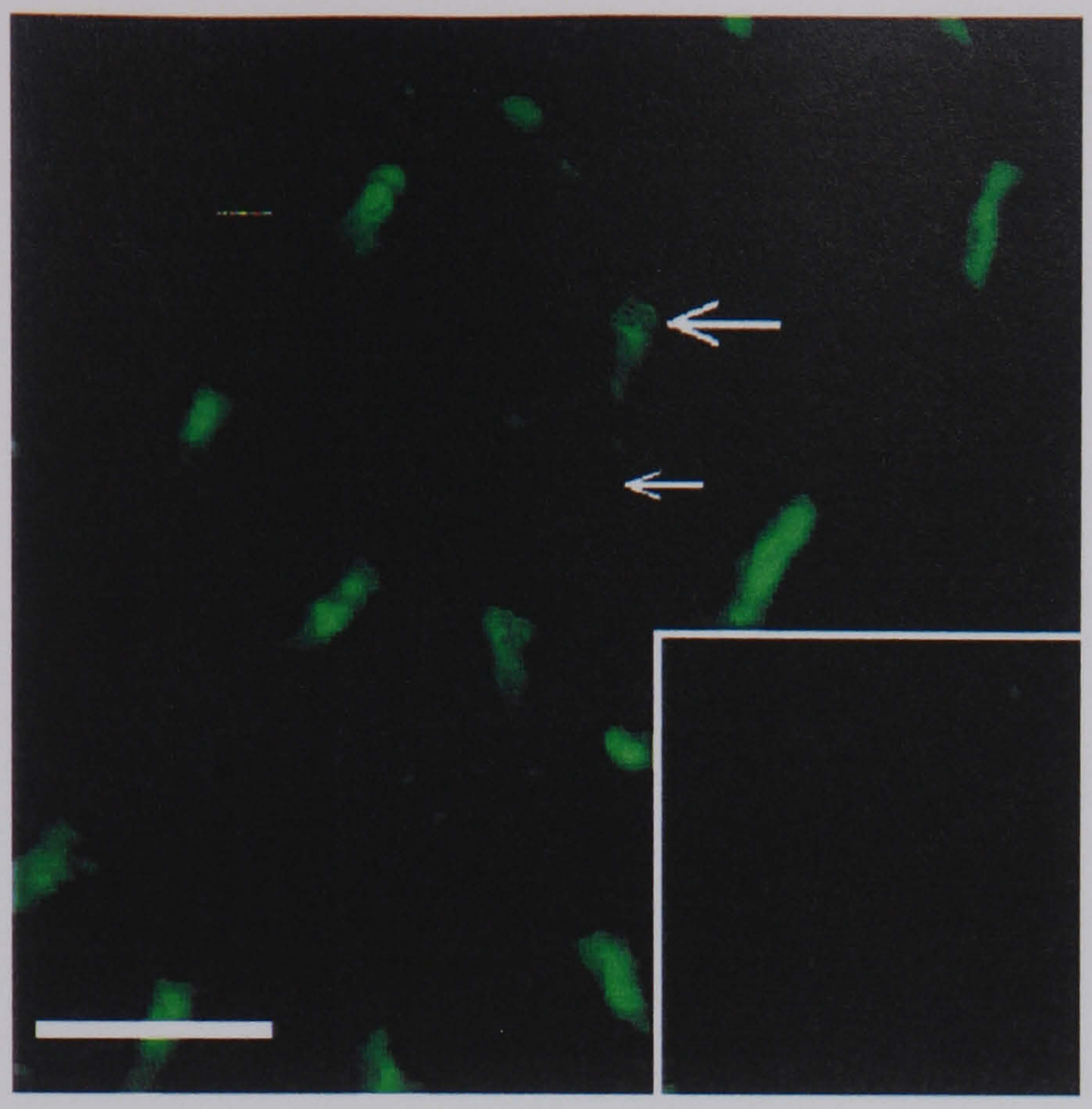
Immunocytochemical analysis of an acetone post-fixed rat jejunum section probed with antiserum raised towards peptide C, used at a dilution of 1:100. rCNT1 localization at the brush border is indicated by the *arrow* with the basolateral membrane being marked by an *asterisk* (*). The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 25μm.

Figure 4.5 Subcellular distribution of rCNT1 in the rat kidney



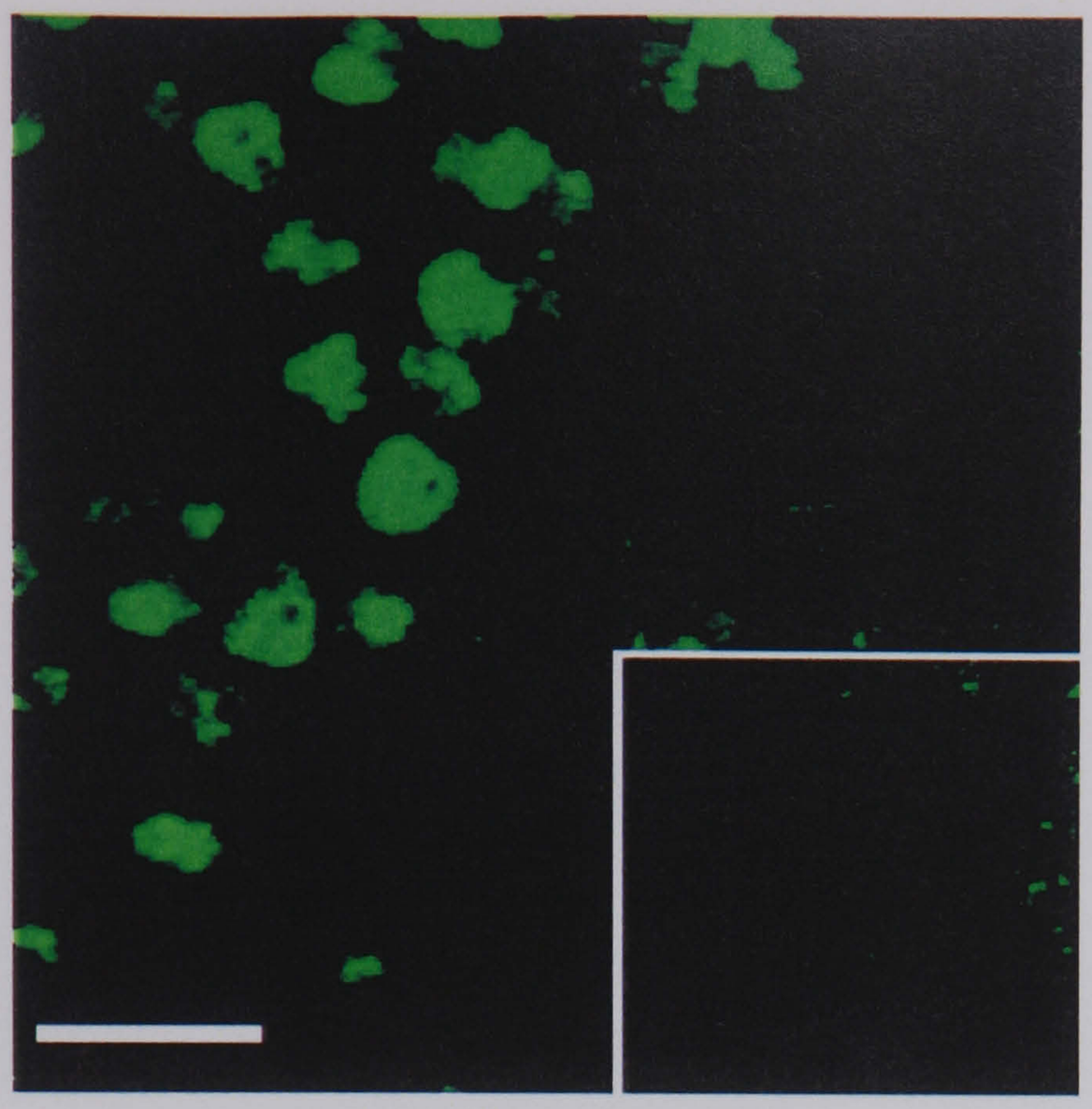
Immunocytochemical analysis of perfused paraformaldehyde-fixed rat kidney sections. *Plate* A probed with antiserum raised to the ΔrCNT1-GST fusion protein (1:400 dilution); *plate* B probed with antiserum raised against GST (1:400 dilution); *plate* C probed with antiserum raised against GLUT1 (1:400 dilution) identifying the proximal tubule, with pre-immune GLUT1 control, *plate* D. *Arrows* in *plates* A & B and C & D indicate brush border and basolateral membranes, respectively. Bar = 40 μm.

Figure 4.6 Subcellular distribution of rCNT1 in the rat liver



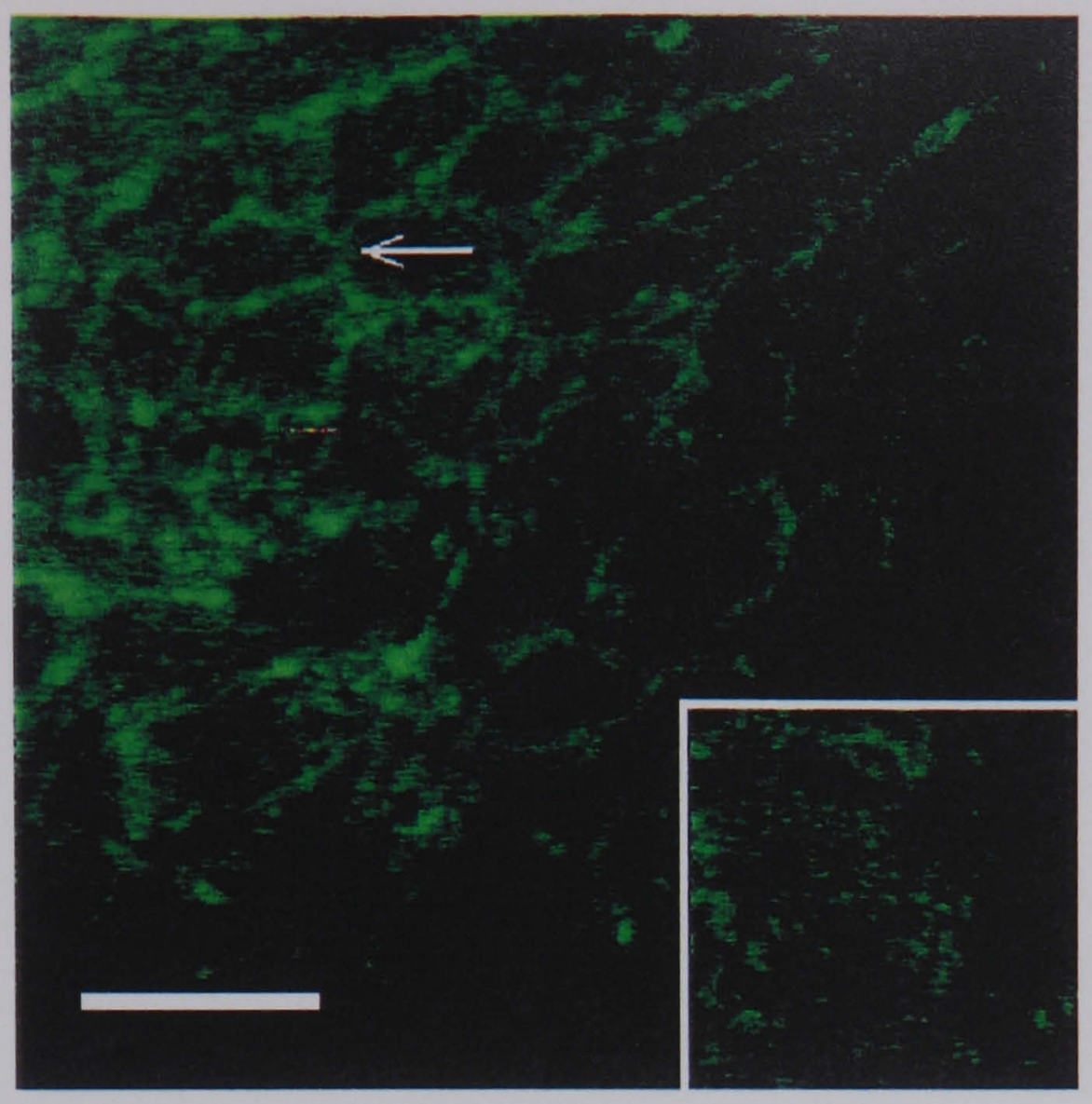
Immunocytochemical analysis of an acetone post-fixed rat liver section probed with antiserum raised towards peptide C, used at a dilution of 1:100. rCNT1 localization at the bile canalicular membrane of hepatic parenchymal cells is indicated by the *large arrow* while its absence from the sinusoidal membrane is indicated by the *small arrow*. The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 80 μm.

Figure 4.7 Cellular distribution of rCNT1 in the rat brain



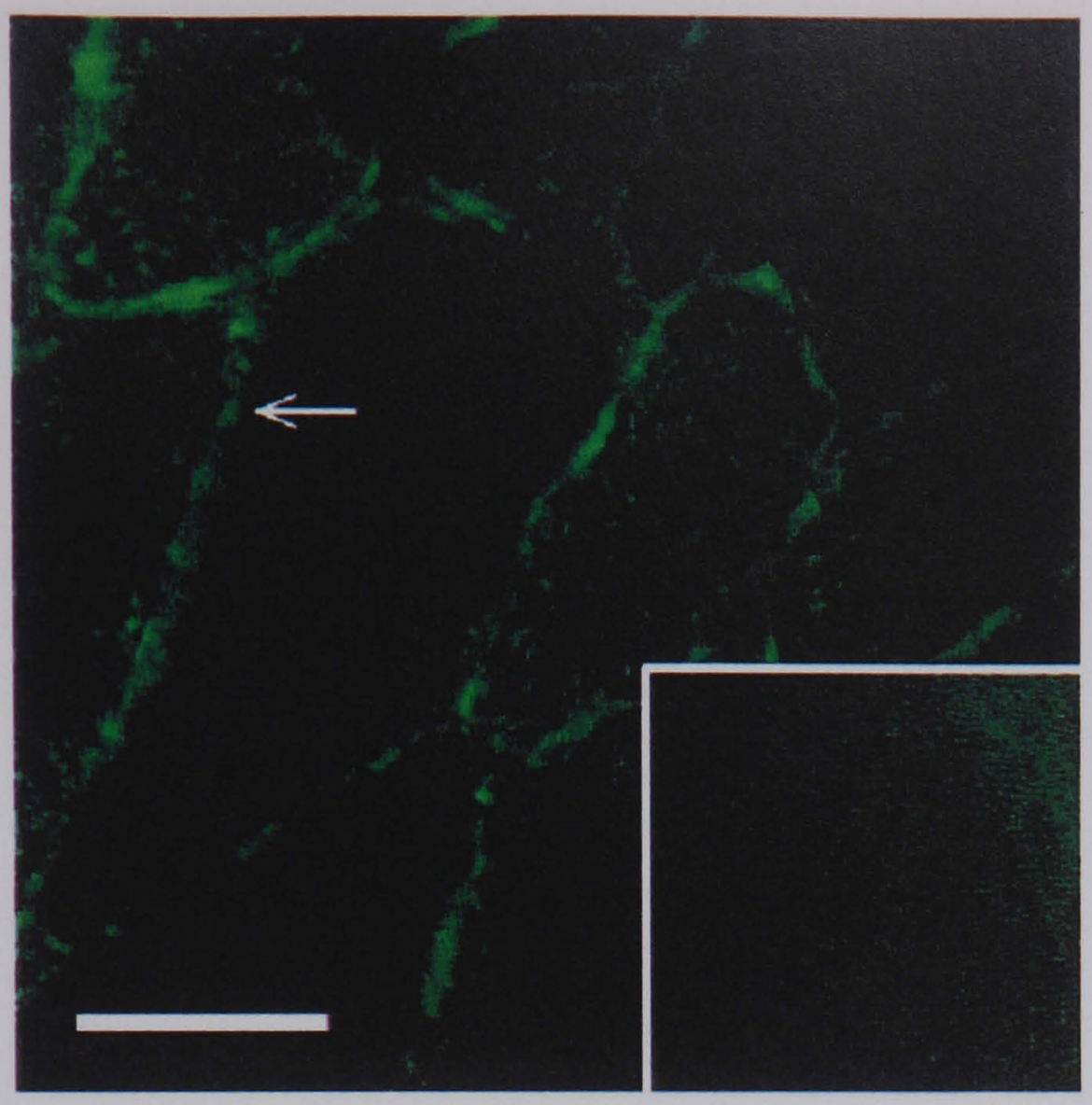
Immunocytochemical analysis of acetone post-fixed rat brain section probed with antiserum raised towards peptide D, used at a dilution of 1:100. rCNT1 localization is within the hippocampal region. The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 40 μm.

Figure 4.8 Subcellular distribution of rCNT1 in the rat heart



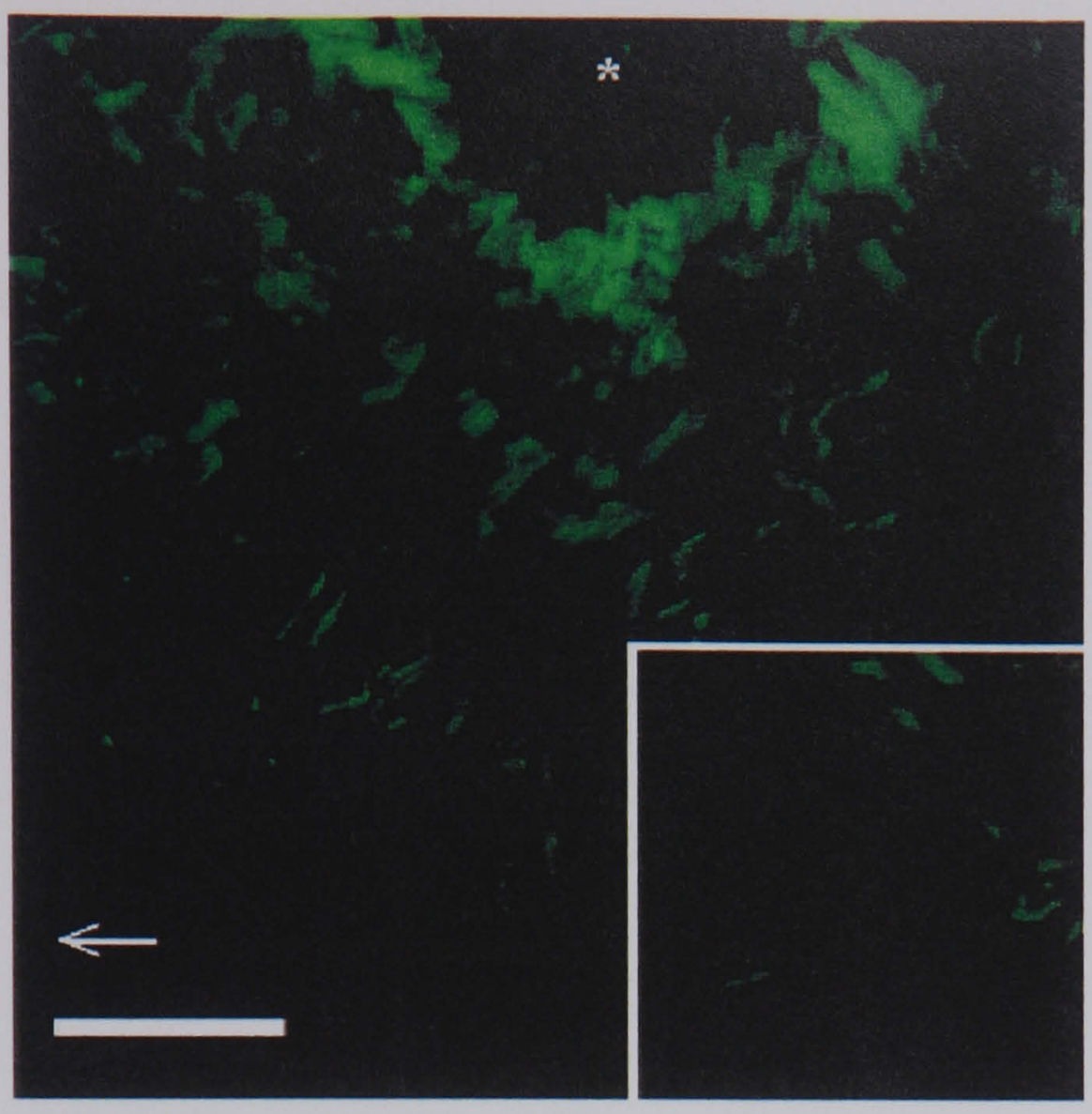
Immunocytochemical analysis of an acetone post-fixed rat cardiac section probed with antiserum raised towards peptide D, used at a dilution of 1:100. rCNT1 localization at the plasma membrane is indicated by the *arrow*. The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 40 μm.

Figure 4.9 Subcellular distribution of rCNT1 in the rat skeletal thigh muscle



Immunocytochemical analysis of an acetone post-fixed rat skeletal thigh muscle section probed with antiserum raised towards peptide D, used at a dilution of 1:100. rCNT1 localization at the plasma membrane is indicated by the *arrow*. The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 40 μm.

Figure 4.10 Subcellular distribution of rCNT1 in the rat testis



Immunocytochemical analysis of an acetone post-fixed rat testicular section probed with antiserum raised towards peptide C, used at a dilution of 1:100. rCNT1 localization at the central region of the seminiferous tubule is indicated by the *asterisk* (*). The *inset* shows a comparable section stained with pre-immune serum at the same concentration, as a control. Bar = 40 μm.

[†]Data discussed in **chapter 5** indicates that rCNT1 can exist *in vivo* in varying glycosylation states, with either one or both of the asparagines at positions 605 and/ or 643 glycosylated. This observation may explain the differences in the banding patterns observed between the intestine, kidney and liver. That is, in the intestine and kidney either one or both of these sites appear to be glycosylated resulting in the presence of two rCNT1 isoforms. In contrast, in the liver it appears that either one or other of the glycosylation sites are occupied, resulting in a protein, or mixture of proteins, with a single apparent M_r of 60 000. Furthermore, in the intestinal sample proteins were detected with apparent M_r of 100 000 and 50 000. These appeared to be degradation products of rCNT1 with the latter representing truncated protein while the former may represent aggregation of this truncated form.

4.4.3 Tissue distribution of rCNT1

Immunoblot analysis of eight rat tissues indicated that antibodies raised against rCNT1 peptide C reacted with a protein which migrated with an apparent M_r of between 60 000 and 63 000 in jejunum, kidney and liver (**Figure 4.2a**). Interestingly, although the rabbit non-specific IgG blot indicated that these bands were specific it also demonstrated that these control antibodies cross-reacted with several other proteins present (**Figure 4.2b**). The observation that the rCNT1 antibodies specifically labelled proteins of apparent M_r of between 60 000 and 63 000 was unexpected because according to the amino acid sequence rCNT1 is predicted to have an unmodified M_r of 71 000. However such a discrepancy between predicted and observed M_r is a known phenomenon when investigating membrane proteins by SDS-PAGE related techniques. The basis of this technique is that in the presence of excess SDS proteins bind this molecule in relation to their size, giving the protein a constant negative charge per unit mass. When such protein-SDS complexes are then added to a polyacrylamide gel the proteins present will move towards the anode during electrophoresis and separate owing to the molecular-sieving properties of the gel. However in the case of highly hydrophobic proteins, such as membrane proteins, SDS binds to the hydrophobic regions to a much higher extent thus giving the proteins a higher unit charge per unit mass than normal resulting in such proteins running faster on SDS-PAGE than expected. Alternatively, it may be that hydrophobic proteins fail to completely unfold and therefore will pass through the SDS gel faster due to their reduced size. An example where a decrease between predicted and observed apparent M_r is observed is with the facilitative glucose transporter GLUT1, which has an apparent M_r of 54 000 based on the unmodified

protein sequence [162] but routinely migrates with an apparent M_r of 45 000 on deglycosylation. The question then arises, how confident are we that the protein observed is rCNT1? There are two pieces of evidence that confirm the band observed, in each of the three tissues, with an apparent M_r of between 60 000 and 63 000 is rCNT1. Firstly, when a parallel blot was carried out using pre-immune or non-specific rabbit IgG, as control, no bands were observed migrating at an equivalent size to those mentioned above (**Figure 4.2b**). Secondly, anti-peptide C antibodies specifically reacted with a protein of similar apparent M_r when blotted against *Xenopus* oocyte membranes expressing rCNT1 but not water injected control membranes, see **chapter 5**. Additionally, this work indicated that this antibody did not cross-react with the homologue rCNT2, a purine specific active nucleoside transporter which has been shown to be present in the liver [84].

As mentioned previously, **section 4.1**, the detection by Northern blotting of mRNA for rCNT1 had suggested that this protein was present in the jejunum and kidney [81]. However the present study was the first to reveal that rCNT1 was also present in the liver until Felipe *et al.* (1998) demonstrated that both rCNT1 and rCNT2 were present in liver parenchymal cells and that the subcellular distribution was different with only rCNT2 appearing in basolateral membrane preparations [89].

Screening with anti-peptide D antibodies did not produce a signal of the expected apparent M_r in any tissue although two bands in the brain were observed with apparent M_r of 45 000 and 56 000 (**Figure 4.3a**). Furthermore, this latter band was also observed in the spleen. Although the control blot demonstrated significant cross-reactivity of the non-specific rabbit IgG with other proteins present (**Figure 4.3b**), no bands were observed with an equivalent apparent M_r similar to the fore-mentioned

bands demonstrated in the brain and spleen, thus indicating that these bands were specific. Considering information discussed in **chapter 5**, where the deglycosylated protein is shown to produce a signal with an apparent M_r of 56 000, this suggests that in both the brain and spleen at least a small percentage of the rCNT1 population is present in the unglycosylated state. Indeed, it may be that all the rCNT1 present is unglycosylated as a means of intra-cellular targeting, although this has not been proven. Alternatively, it is possible that rCNT1 is present in the glycosylated state but due to the relative proximity of the oligosaccharide moiety to the epitopic region the saccharide may cause steric-hindrance, thus preventing the antibodies raised against the rCNT1 peptide D from gaining access to the antigenic site, therefore suggesting that it is a degradation product or immature form of rCNT1 that is being observed. Furthermore, the second band observed in the brain, with an apparent M_r of 45 000, is possibly a truncated version of rCNT1, which in conjunction with the fact that the antibodies react towards the *C*-terminus of the protein suggests that the truncation occurs at the *N*-terminus. Further evidence for this is provided by a private communication from Prof. J. D. Young (University of Alberta, Canada), which reported that elimination of the three putative *N*-terminal helices did not affect transport activity of rCNT1, suggesting that the band observed with an apparent M_r of 45 000 in the brain is a truncated isoform of rCNT1 that may still be functional.

To investigate further the subcellular localization of rCNT1 in rat tissues immunocytochemistry was carried out in the present study. These immunocytochemical studies were initially focused on the tissues which produced a signal on Western blotting but were later extended to the remaining tissues initially screened.

4.4.4 Subcellular distribution of rCNT1

Immunocytochemical investigation of the rat jejunum indicated that rCNT1 was present on the brush border membrane but absent from the basolateral membrane, **Figure 4.4**. This observation is consistent with the distribution of other transporters in the intestinal epithelium. In the case of the glucose transporters, the active sugar transport protein SGLT1 is located at the brush border but is absent from the basolateral membrane. Complementary to this the facilitative transporter GLUT2 is present at the latter membrane but is absent from the former [163]. These distributions reflect the roles of the two transporters: SGLT1 is required for the active uptake of glucose from the intestinal lumen, where glucose is present at lower concentrations than in the cell, whereas GLUT2 allows the equilibrative transport of the cellular glucose down the concentration gradient into the interstitial fluid and hence blood supply. It is therefore possible that a similar situation is true for nucleoside uptake from the intestinal cavity and release into the blood supply, i.e. rCNT1 present at the brush border is responsible for the active uptake of pyrimidine nucleosides from the lumen while another transporter may be responsible for the release of modified or intact nucleosides from the epithelial cells into the blood. One case where facilitative nucleoside transport has been demonstrated to be present at the basolateral membrane was that of the human *es*-type transporter in T84 intestinal epithelial cells [164]. Therefore, this suggests that a facilitative nucleoside transporter may also be present in the rat BLM. The observation that rCNT1 is present at the brush border is consistent with previous work carried out on intestinal epithelial cells. Terasaki *et al.* in 1993 isolated the mRNA from rabbit small intestinal cells and showed that on expression in *Xenopus* oocytes this gave similar

[†]Further evidence for the presence of a purine specific nucleoside transporter in the intestinal brush border was demonstrated by Bronk *et al.* [185]. Working with isolated loops of jejunum they demonstrated that purine in addition to pyrimidine nucleosides were transported into epithelial cells.

characteristics to a *cit*-type nucleoside transport protein previously reported in brush border vesicles [165], suggesting that they were the same transporter. Later it was suggested that rat intestinal brush border membrane vesicles contained both facilitative and concentrative nucleoside transport systems [166]. This information in conjunction with the evidence for the human intestine containing both *cif*- and *cit*-type activity, and the cloning of a rat intestinal *cif*-type transporter, suggests that if a pyrimidine transporter is present at the brush border then a complementary purine specific transporter would also be expected [167,168].†

Immunocytochemical analysis of the rat kidney, using antibodies raised against the ΔrCNT1-GST fusion protein, demonstrated not only the expression of rCNT1 at the brush border but its absence from the basolateral membrane as distinguished by the use of a GLUT1 control, **Figure 4.5**. Additionally, since GLUT1 was restricted to the proximal tubules [169] and rCNT1 was localized to these same tubules, this indicated that rCNT1 expression was not throughout the kidney but was also restricted to the proximal tubules of the renal cortex. Active nucleoside transport has previously been demonstrated in the rat and rabbit renal brush borders [64,75] but this is the first case where an active *cit*-type transport protein has been visualized to be present only at the brush border membrane of the renal cortical proximal tubule. The location of rCNT1 at this membrane and its absence elsewhere in the kidney is due to its predicted physiological role of nucleoside salvage from the renal filtrate. A similar situation is also observed with the active glucose transporter SGLT1 which is present at the renal proximal tubule brush border but absent from the basolateral membrane.

Subcellular investigation of the rat liver indicated that the presence of rCNT1 was restricted to the bile canalicular membrane of parenchymal cells and thus it was

absent from the sinusoidal membrane, **Figure 4.6**. Once again this localization of rCNT1 suggests that its physiological role here in the liver is one of salvaging nucleosides from the bile following hepatic excretion. The recent cloning of the *cif*-type nucleoside transporter rCNT2 from the rat liver suggests that a complementary pyrimidine transporter is also required. More recently this has been proven by Felipe *et al* .who demonstrated that both rCNT1 and rCNT2 are present in the rat liver [89], as mentioned in **section 4.4.3**.

Although the presence of glycosylated rCNT1 was not detected by Western blotting in the brain using antibodies raised against peptide C a signal was obtained using antibodies raised against peptide D. However, the protein appeared to be either unglycosylated, or unglycosylated and truncated, as described in **section 4.4.3**. This previously observed Western blot signal, as described in **section 4.4.3**, was confirmed by the recognition of rCNT1 in what appeared to be the cell membrane of hippocampal pyramidal cells, as demonstrated in **Figure 4.7**. Unfortunately, at the time of carrying out the initial immunocytochemistry on various tissues the ΔrCNT1-GST fusion protein antibody was not available so frozen sections were used. Due to this and the inability to prefix the tissue this meant that the specific cell type in which the signal was observed could not be identified due to partial tissue disruption. However this will be possible in future studies now that a method for detection with the ΔrCNT1-GST fusion protein antibody has been established. Other evidence for the presence of rCNT1 in the brain comes from recent work carried out by Anderson *et al.* who demonstrated that both rCNT1 and rCNT2 mRNA were present in the brain [88].

Surprisingly, immuno-localization of rCNT1 was demonstrated in both cardiac and skeletal thigh muscle, as illustrated in **Figures 4.8** and **4.9** respectively. Wherein the presence of the transporter was localized to the plasma membrane in each case, although this was very weak when compared to the levels of detection mentioned in the other tissues to date, however it did appear specific. These results were obtained using antibodies raised against peptide D with the signal being much weaker with those raised towards peptide C. A possible explanation for this is that glycosylation confers structural constrains on the rCNT1 peptide backbone when present and that when this glycosylation is removed the protein alters to a conformation to which the antibodies raised against peptide D are avid. Therefore if rCNT1 is present mostly as an aglyco-protein in these two tissues then anti-peptide D antibodies will produce a stronger signal. Also, since the antibodies raised against peptide D produce a signal in the intestine, it may be that a proportion of rCNT1 in the intestine is also unglycosylated. Returning to the heart, the presence of active nucleoside transport has not been reported to as great an extent as the existence of facilitative transport, which has been widely investigated [146,170]. The only existence to date of active transport in cardiac muscle is in human myocytes where the presence of one active transporter has been demonstrated in addition to two facilitative transporters by physiological techniques [171]. Skeletal muscle has also had very little work carried out on it in relation to nucleoside transport, in so far that no reported cases could be found for skeletal muscle nucleoside transport although facilitative transport has been demonstrated in the smooth muscle cell line DDT1 MF-2 [172].

On immunocytochemical analysis of the testis a signal, indicating the presence of rCNT1, was observed at the center of the seminiferous tubule (**Figure 4.10**) in the

region where late development of the gametes occurs. On closer inspection it would appear that it was the central body region of the gametes where the signal was produced. Interestingly the mitochondrial fraction from rat testis has been demonstrated to possess facilitative nucleoside transport [173]. Coincidentally, the body region of the gamete is rich in mitochondria. Therefore it may be true that active nucleoside transporters are present on the gamete plasma membrane which transport nucleosides into the cell so that they are available for uptake from the cytoplasm by the mitochondria. Also of interest is that rat Sertoli cell-conditioned medium inhibited the uptake of thymidine by cells in culture [174]. This therefore may be a method of controlling gamete activity.

Finally, although a signal was produced in the spleen with antibodies raised against peptide D on Western blotting indicating the presence of rCNT1, no signal was observed on immunocytochemical analysis. Thus it may be that the antibodies only recognize the spleen isoform of rCNT1 when denatured and unglycosylated.

4.5 Summary

Immunoblot analysis has confirmed the presence of rCNT1 in the rat brain, jejunum, kidney, liver and spleen. Further investigation by immunocytochemistry localized this protein to the brush border of both the jejunum and kidney, although in the latter case distribution was restricted to the brush border of the proximal tubules of the renal cortex while being absent from other tubules present. Within the liver, rCNT1 was found at the bile canalicular membranes of hepatic parenchymal cells but was absent from the sinusoidal membranes. These results coincide with the predicted physiological roles of this protein in the absorption of nucleosides from the intestinal

lumen and in the salvage of nucleosides from both the renal filtrate and hepatic bile fluid. In the brain preliminary studies suggested that rCNT1 was present in the plasma membrane of hippocampal pyramidal cells, although this remains to be confirmed using fixed tissue and antibodies raised against the ΔrCNT1-GST fusion protein. However a possible role for the transporter in this tissue is in the uptake of the neurotransmitter adenosine. Interestingly, having demonstrated rCNT1 is present in the spleen by Western blotting this could not be confirmed by immunocytochemical analysis. However, rCNT1 was also demonstrated by immuncytochemistry to be present in the heart, skeletal thigh muscle and testis, although these were not confirmed by Western blotting, in addition to its presence in the fore mentioned tissues. Possible reasons for the differences in Western blot and immunocytochemical analyses have been discussed in detail in the previous section but briefly, they appear to have been due to differences in the levels of expression of the glycosylated and unglycosylated forms of rCNT1, and the preference of the different antibodies for one or other. Stemming from the observations made in this chapter, it is clear that further work is required to conclusively elucidate the reasons for the differences observed between the two techniques employed in this study. Therefore, presenting incitement for future investigation into the distribution of rCNT1 within these tissues.

Chapter 5

Topological analysis of rCNT1

5.1 Introduction

Information on the topology of rCNT1 is very limited, with the majority of it based on predictive computer algorithms as discussed in **section 1.3.2**. This has led to the development of the putative model represented in **Figure 1.2**, where both the *N*- and *C*-termini were predicted to be cytoplasmic with fourteen hydrophobic transmembrane domains located between them. As a first step towards understanding how this protein functions it is essential that this tentative model be directly tested by experimental procedures. Techniques which can be used include circular dichroism, FTIR spectroscopy, proteolytic digestion, site-specific antibodies, biotinylation or glycosylation scanning mutagenesis. All of these techniques have been used previously in topological investigation of membrane proteins as discussed in **section 1.1**, where the development of methods to elucidate the conformation of GLUT1 was considered.

One of the most powerful methods for determination of membrane protein topology is glycosylation scanning mutagenesis, which exploits the fact that many membrane proteins bear oligosaccharide chains on their extracellular domains. The attachment of such chains occurs as the result of post-translational glycosylation reactions as the proteins are transported through the secretory pathway. There are two basic forms of glycosylation, either *N*- or *O*-linked where an oligosaccharide is conjugated to an asparagine residue or to either serine or threonine respectively. *N*-linked glycosylation involves the transfer of a high mannose oligosaccharide structure *en bloc* from a dolichol donor to the nascent peptide in the endoplasmic reticulum [175]. The oligosaccharide chains are attached to an asparagine residue within the

consensus sequence Asn-Xaa-Ser/ Thr, where Xaa is any amino acid other than proline. However, a consensus sequence followed by a proline residue is unlikely to be glycosylated [176]. Following addition of the carbohydrate moiety the glycoprotein is translocated to the Golgi complex where the sugar component is modified by a number of different enzymes. Furthermore, it is in the Golgi where *O*-linked glycosylation of serine and threonine residues takes place. The specificities of the acetylgalactosaminyltransferase enzymes involved in *O*-linked glycosylation are relatively poorly understood. Although, the specificity appears to be modulated by sequence context, secondary structure and surface accessibility.

The majority of membrane proteins, such as receptors, channels or transporters, are *N*-linked glycosylated suggesting that this post-translational modification plays an important role in processing or expression of the protein. However there are exceptions such as the α-subunits of cation-transporting ATPases and the α_1-subunit of the calcium channel which, although they themselves are unglycosylated, are tightly associated with other *N*-linked glycosylated subunits [177]. A number of roles have been proposed for this type of modification including directing protein folding, regulating cell-surface expression and increasing the half-life of proteins [178]. Interestingly, if more than one glycosylation site is present in a transmembrane protein then it appears that usually only the first is glycosylated, an exception being the α_1-subunit of the sodium channel which due to a duplication in structure contains two glycosylated sites [179]. The mechanisms that restrict glycosylation to a single loop in the majority of proteins is unclear. However, the size of the loop and the position of the site within this segment are important: it has been demonstrated that for an extracellular loop to be efficiently glycosylated it must be larger than 25

residues, with at least 12 residues *N*-terminal and 14 residues *C*-terminal to the asparagine that is to be modified [180].

Taking into account the features mentioned above glycosylation can be used as a means of probing the topology of integral membrane proteins. This can be achieved by a technique known as glycosylation scanning mutagenesis where an aglyco-protein is made by the mutagenic elimination of any endogenous *N*-linked glycosylation sites. This therefore provides a backbone for the incorporation of exogenous glycosylation sites at specific positions in predicted hydrophilic segments within the peptide. If such sites are incorporated into extracellular loops of the protein they are expected to become glycosylated, provided that the loops are sufficiently large to allow access of the glycosylation machinery. However, if the sites are incorporated into intracellular loops they will not be glycosylated. Thus, analysis of the glycosylation state of the mutant protein can be utilized to determine the orientation of each segment with respect to the plasma membrane. Such a strategy has been used to investigate a number of membrane proteins and thus provide data that has led to production of a two-dimensional model for these proteins in the membrane [22,180,181].

rCNT1 has three potential *N*-linked glycosylation sites at positions 543, 605 and 643. The first of these is located in putative TMD thirteen and so is unlikely to be accessible to the glycosylating machinery of the endoplasmic reticulum. However, the sites at positions 605 and 643 lie in the hydrophilic *C*-terminal region of the protein, which was predicted by Huang *et al.* to be cytoplasmic [81], hence suggesting that these two sites are unglycosylated. Therefore, the initial aim of this investigation was to establish if rCNT1 was glycosylated. This was achieved by

investigating the glycosylation state of wild-type rCNT1 in *Xenopus* oocytes by comparing the protein in the presence and absence of the deglycosylating enzyme endoglycosidase F. Having established if rCNT1 was present as a glycoprotein the accessible consensus sequences were eliminated to produce an aglyco-rCNT1 mutant that could be used as the basis for mutagenesis. This aglyco-protein allowed the generation of a number of mutants with either single residue mutations to introduce a glycosylation site or unique *Kpn*I restriction sites at which cDNA fragments encoding a potentially *N*-glycosylated hydrophilic loop region could be inserted. In the present study the extracellular glycosylated hydrophilic loop of human GLUT4, residues 46 to 78, was employed for the latter purpose, a suitable cDNA fragment bearing terminal *Kpn*I restriction sites being produced by PCR amplification of GLUT4 cDNA (**Figure 5.1**).

Figure 5.1 Amplification of human GLUT4 cDNA encoding a potential *N*-glycosylated loop

Primers designed to amplify the cDNA encoding the potential *N*-glycosylated hydrophilic loop of human GLUT4 are indicated above with 5' *Kpn*I restriction site tags (shown in *italics*). The translated protein is illustrated with the region amplified identified in **bold**.

5.2 Methods for topological analysis

5.2.1 Endoglycosidase digestion to determine the glycosylation state of membrane proteins

Xenopus oocyte membrane preparations expressing the nucleoside transporters rCNT1, rCNT2, hCNT1 and hCNT2, in addition to water injected control membranes, were obtained from Dr J. D. Young (University of Alberta, Canada). In each case oocytes were injected with 10 ng of cRNA encoding each of the transporters prior to incubation at 18°C for 3 days as described by Huang *et al.* [81]. Total protein membrane preparation (10μg) in 50 μl 100 mM sodium acetate, pH 7.2, 50 mM EDTA, 0.2 % octyl-β-D-glucopyranoside, 1 % mercaptoethanol were digested with 0.125 U endoglycosidase F/ N-glycosidase F from *Flavobacterium meningosepticum* (Boehringer-Mannheim) overnight at 22°C. Controls were performed under identical conditions with water being substituted for the deglycosylating enzyme. Protein determination was performed with the BCA protein assay as described in **section 2.4.2**. SDS-PAGE and Western blot analyses were performed as described in **sections 2.4.3** and **2.4.4** respectively.

5.2.2 Mutagenic removal of rCNT1 endogenous glycosylation sites

Numerous strategies for mutagenesis are available but recently a non-PCR based mutagenesis technique has been developed by Stratagene, known as the QuikChange™ site-directed mutagenesis technique. This technique has a number of advantages over traditional PCR mutagenesis and will be used in the present study. The procedure starts with double-stranded DNA and two complementary primers containing the desired mutation. These are extended in the presence of the thermostable enzyme *Pfu* DNA polymerase. The incorporation of the primers

produces a mutated plasmid with staggered nicks. The mutated plasmid is then selected for by the digestion of the *dam* methylated parental DNA, which has been produced in *E. coli* by the methylation of adenine residues in the sequence GATC, leaving the nicked mutant plasmid containing the desired mutation which can then be transformed into *E. coli*. An overview of this strategy can be seen in **Figure 5.2**. The main advantages of this procedure are that subcloning of the final product is eliminated, thus speeding the synthesis of mutants; *Pfu* DNA polymerase has a high fidelity, thus decreasing the number of mis-matches; the technique is reportedly 150 times more efficient than standard PCR-based methods; and the complete procedure can be accomplished in one day. The efficiency of QuikChange™ site-directed mutagenesis is reported to produce greater than 80 % of the resulting colonies containing the desired mutation. This is due to the small amount of starting DNA template required for this technique, the high fidelity of *Pfu* DNA polymerase and the low number of thermal cycles.

cDNA encoding rCNT1 in the vector pGEM-3Z, designated pQQH by Huang *et al* [81], was mutated using the QuikChange™ site-directed mutagenesis kit (Stratagene). For mutagenesis of asparagine 605 to threonine the forward primer rCNT1_N605Tf was used in conjunction with the reverse primer rCNT1_N605Tr (see **Table 5.1** for sequences of all primers used in this chapter). Asparagine 643 was mutated to threonine using the forward primer rCNT1_N643Tf in combination with the reverse primer rCNT1_N643Tr. The reactions were set up to a final volume of 50 µl according to the manufacturer's recommendations using 125 ng of each primer, a range of template DNA between 5 and 50 ng, 0.2 mM dNTPs, 1x amplification buffer (supplied with the kit), and 2.5 U *Pfu* DNA polymerase. Conditions used for

thermo-cycling were: stage 1, one cycle of 95°C for 90 sec; stage 2, twelve cycles of 95°C for 30 sec, 55°C for 60 sec and 68°C for 12 min; and stage 3, one cycle of 68°C for 10 min. Following mutagenesis the parental DNA was digested by addition of 20 U *Dpn*I for 1 h at 37°C prior to transformation of competent XL10-Gold *E. coli* cells, as described in **section 2.5.4**, with 100 pg of the resulting mutated DNA. The N605T/ N643T double mutant was generated using the N605T mutant as template in conjunction with the N643T primers, using the same method used to obtain the single mutants. *Bam*HI-*Xba*I cassettes of 477 bp, containing the mutated sequences, were subcloned into the original pQQH construct and then this region in each mutant was completely sequenced with the ABI 373A DNA sequencer using dye-deoxy terminator chemistry (Perkin Elmer) in order to confirm the presence of the desired mutation and the absence of any other mutations. The resultant constructs were designated pSRH3, pSRH4 and pSRH5 for the 605, 643 and 605/ 643 mutants respectively. The locations of the mutated residues within rCNT1 are illustrated in **Figure 5.3**, with the relative positions of restriction sites used in subcloning of the mutated cassettes, described above, indicated.

5.2.3 Incorporation of potential glycosylation sites by single residue modification

To generate potential *N*-linked glycosylation sites by the replacement of single amino acid residues, the construct pQQH was mutated by the QuikChange™ site-directed mutagenesis kit (Stratagene) as described in **section 5.2.2**. Three such mutations were attempted, at position 6 where glutamine was converted to threonine using the forward primer rCNT1_Q6Tf in conjunction with the reverse primer rCNT1_Q6Tr, at position 54 where alanine was mutated to threonine with the primers rCNT1_A54Tf and rCNT1_A54Tr, and at residue 503 where glutamic acid was

replaced by asparagine using the complementary primers rCNT1_E503Nf and rCNT1_E503Nr. Following transformation into XL10-Gold *E. coli* cells, colonies were screened by automated sequencing to confirm generation of the correct mutations. Subsequently the successful Q6T mutant was subcloned into pSRH5 by ligation of the 194 bp *Nde*I-*Eco*NI fragment from the Q6T mutant into pSRH5 formerly digested with the two fore-mentioned enzymes. The resultant construct was designated pSRH22. The locations of the mutated residues within rCNT1 are illustrated in **Figure 5.3**, with the relative positions of restriction sites used in subcloning of the mutated cassettes, described above, indicated.

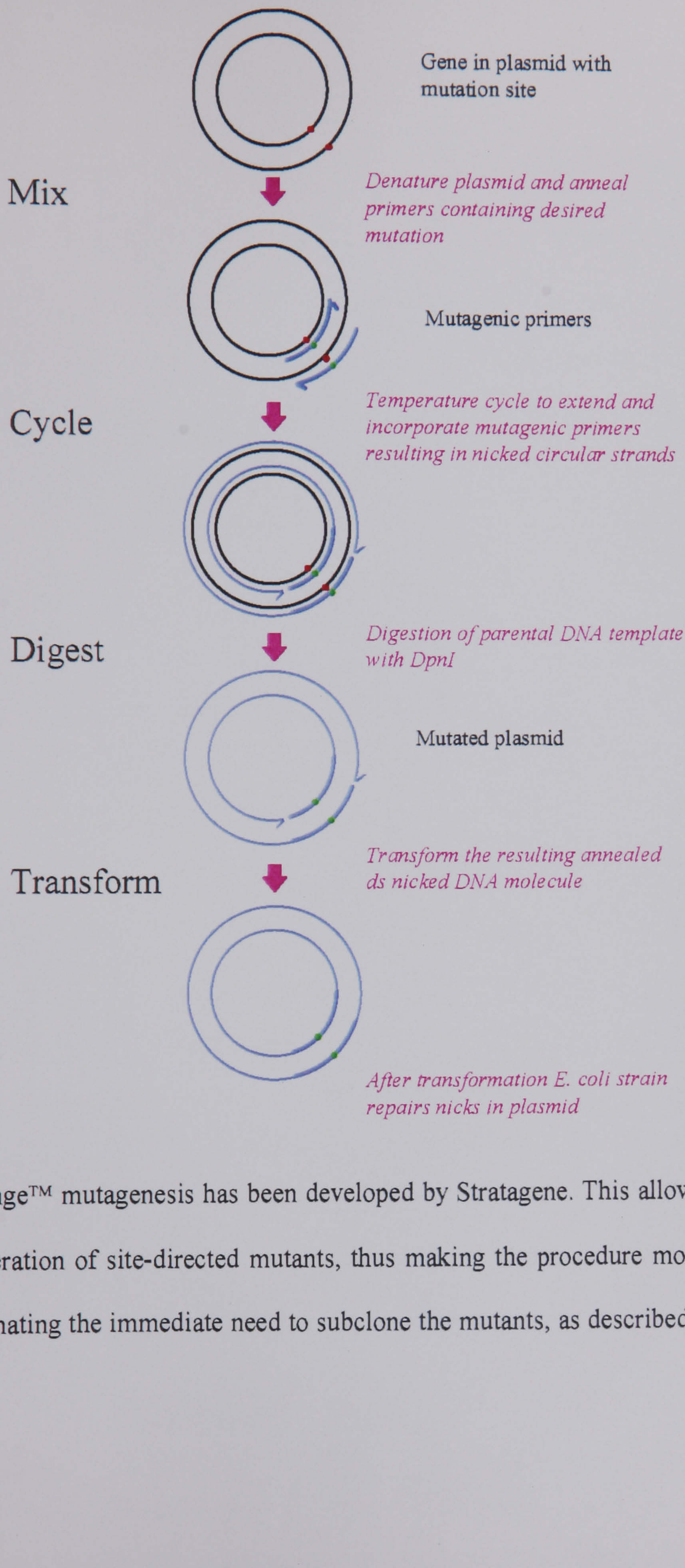
5.2.4 Amplification of cDNA encoding the glycosylated hydrophilic loop of GLUT4

The region encoding residues 46 to 78 was PCR amplified from human GLUT4 cDNA using primers hG4KpnIf and hG4KpnIr, which were designed to incorporate terminal *Kpn*I restriction sites. The 50 µl reaction was set up containing each primer at a concentration of 1 µM, 20 ng of template DNA, 0.2 mM dNTPs, 0.5 U $Vent_R$ DNA polymerase (NEB) and 1x amplification buffer (supplied with the polymerase). Conditions used for thermo-cycling were: stage 1, one cycle of 95°C for 90 sec; stage 2, twelve cycles of 95°C for 30 sec, 55°C for 45 sec and 68°C for 2.5 min; and stage 3, one cycle of 68°C for 7 min. The resulting product was blunt-end ligated into *Eco*RV pre-digested pBluescript II KS (+) as described in **section 2.5.7** with 100 pg of the resulting DNA being used to transform competent XL1-blue *E. coli* cells as described in **section 2.5.4**. The resultant construct was sequenced over the complete region of the insert as described in **section 2.5.8**. This construct was designated pSRH6.

5.2.5 Incorporation of exogenous potential glycosylation sites

The construct pQQH, containing the full-length coding region of rCNT1, was mutated by the QuikChange™ site-directed mutagenesis kit (Stratagene) as described in **section 5.2.2** in order to incorporate unique *Kpn*I restriction sites following residues 106, 200 and 406 of the transporter, using the primers listed in **Table 5.1**. Following transformation into XL10-Gold *E. coli* cells, colonies were screened by automated sequencing to confirm generation of the correct mutations. Subsequently *Nde*I/ *Eco*NI, *Apa*I/ *Nco*I and *Bam*HI/ *Bsp*EI cassettes bearing the 106, 200 and 406 mutations respectively were subcloned into the aglyco-mutant, construct pSRH5, by replacement of the corresponding regions, as described in **section 2.5.7**. Finally each of these constructs was digested with *Kpn*I to allow the insertion of the GLUT4 loop from the construct pSRH6, previously digested with *Kpn*I. The resultant GLUT4 106, 200 and 406 constructs were designated pSRH17, pSRH19 and pSRH21 respectively. The resultant constructs were sequenced over the complete regions of the inserts as described in **section 2.5.8** to confirm that the inserted cDNA was in the correct orientation. The locations within rCNT1 where the GLUT4 hydrophilic loops were inserted are illustrated in **Figure 5.3**, with the relative positions of restriction sites used in subcloning of the mutated cassettes, described above, indicated.

Figure 5.2 *QuikChange*™ mutagenesis strategy

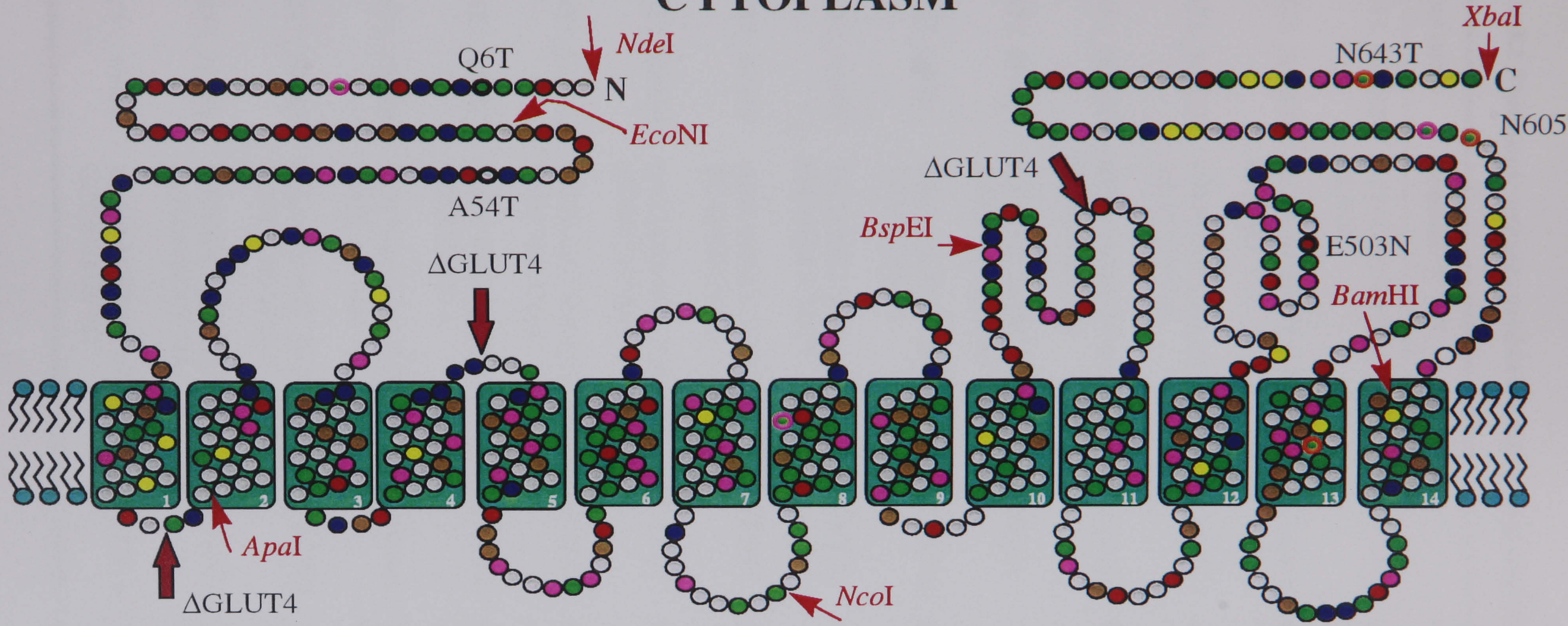


QuikChange™ mutagenesis has been developed by Stratagene. This allows the non-PCR generation of site-directed mutants, thus making the procedure more efficient and eliminating the immediate need to subclone the mutants, as described in **section 5.1.1**.

Figure 5.3 Summary of mutation sites located within the putative model for rCNT1

Image on opposite page

Mutations performed on rCNT1, described in **chapter 5**, are illustrated in the following diagram. Single mutations are indicated adjacent to the amino acid residue modified while locations where the GLUT4 hydrophilic loop was substituted are represented by *arrows* labeled ΔGLUT4. In addition, the relative locations of restriction enzyme sites used in subcloning of mutated cassettes are indicated by the name of the enzyme in *italics*. For further information on components of the diagram refer to **Figure 1.2**.



Key:

- General non-polar
- General polar
- Positive
- Negative
- Aromatic
- Cysteine
- Putative *N*-linked glycosylation site
- Putative *O*-linked glycosylation site
- Structure breaker

Table 5.1 Mutagenic primers used in the topological analysis of rCNT1

Primer name	Sequence 5'→3'	Comments
rCNT1_N605Tf	GCGTGTCCCTTCTGACCCAAACTG TCAGC	Asn 605 to Thr fp
rCNT1_N605Tr	GCTGACAGTTTGGGTCAGAAGGGA CACGC	Asn 605 to Thr rp
rCNT1_N643Tf	CTGCTGTCGATTTTACACCCACAC AGTCTGCAC	Asn 643 to Thr fp
rCNT1_N643Tr	GTGCAGACTGTGTGGGTGTAAAAT CGACAGC AG	Asn 643 to Thr rp
rCNT1_Q6Tf	CATGGCAGACAACACAACGAGGC AAAGAGAGTCC	Gln 6 to Thr fp
rCNT1_Q6Tr	GGACTCTCTTTGCCTCGTTGTGTTG TCTGCCATG	Gln 6 to Thr rp
rCNT1_A54Tf	GGTGGCCTGAACAAAACAGAGCG GAAGGCCTTCT	Ala 54 to Phe fp
rCNT1_A54Tr	AGAAGGCCTTCCGCTCTGTTTTGT TCAGGCCACC	Ala 54 to Phe rp
rCNT1_E503Nf	GTTTGTGGCCTATCAAAACCTTTC CCAGTACAAG	Glu 503 to Asn fp
rCNT1_E503Nr	CTTGTACTGGGAAAGGTTTTGATA GGCCACAAAC	Glu 503 to Asn rp
rCNT1_106*K*f	CTCCTGGACCTC*GGTACC*CAGAGG GCCCTAGCA	Introduction of KpnI site after residue 106 fp
rCNT1_106*K*r	TGCTAGGGCCCTCTG*GGTACC*GAG GTCCAGGAG	Introduction of KpnI site after residue 106 rp
rCNT1_200*K*f	CAAAGCATCACCGT*GGTACC*GCGG TGTCATGGCGAG	Introduction of KpnI site after residue 200 fp
rCNT1_200*K*r	CTCGCCATGACACCGC*GGTACC*AC GGTGATGCTTTG	Introduction of KpnI site after residue 200 rp
rCNT1_406*K*f	CAGAACCTCTTG*GGTACC*GAAGCA GCCAGTGCTG	Introduction of KpnI site after residue 406 fp

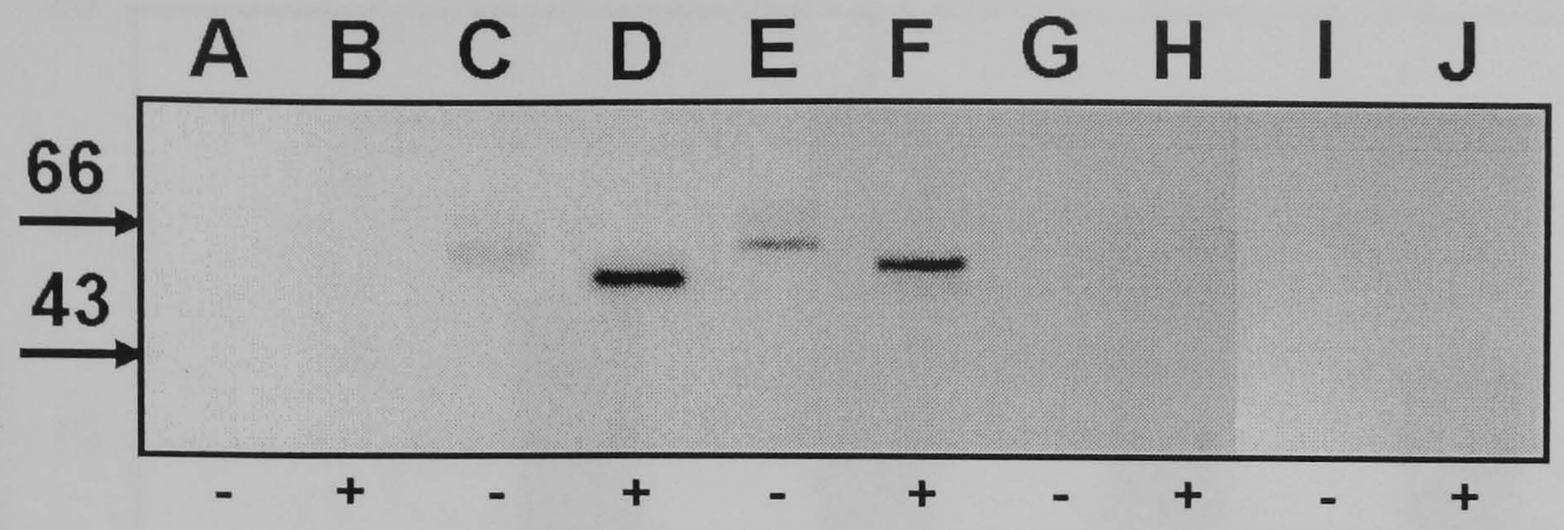
Table 5.1 (cont.)

Primer name	Sequence 5'→3'	Comments
rCNT1_406*K*r	CAGCACTGGCTGCTTC*GGTACC*CA AGAGGTTCTGAGC	Introduction of KpnI site after residue 406 rp
hG4KpnIf	cgg*GGTACC*AATGCCCCTCAGAAGG TG	Generation of KpnI site in GLUT4 loop fp
hG4KpnIr	gcc*GGTACC*GGTGAGGGTGCCTGGA GGGA	Generation of KpnI site in GLUT4 loop rp

fp= forward primer, rp= reverse primer

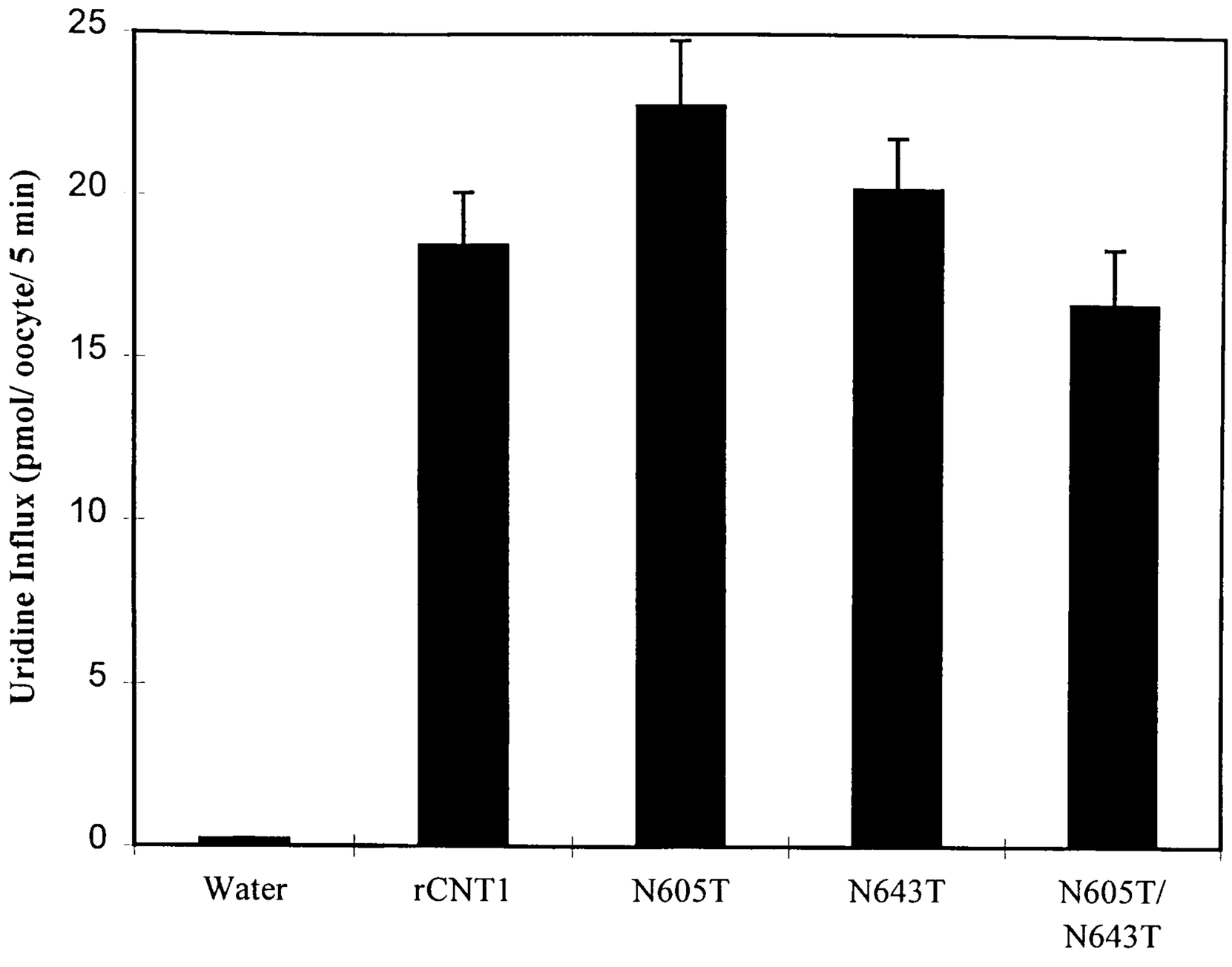
Mutated bases are represented by underlining while inserted bases incorporating restriction sites are *italicized*. Furthermore, additional bases that have been added to facilitate restriction digestion are represented in non-capitals.

Figure 5.4 Glycosylation state of wild-type rCNT1 and homologues



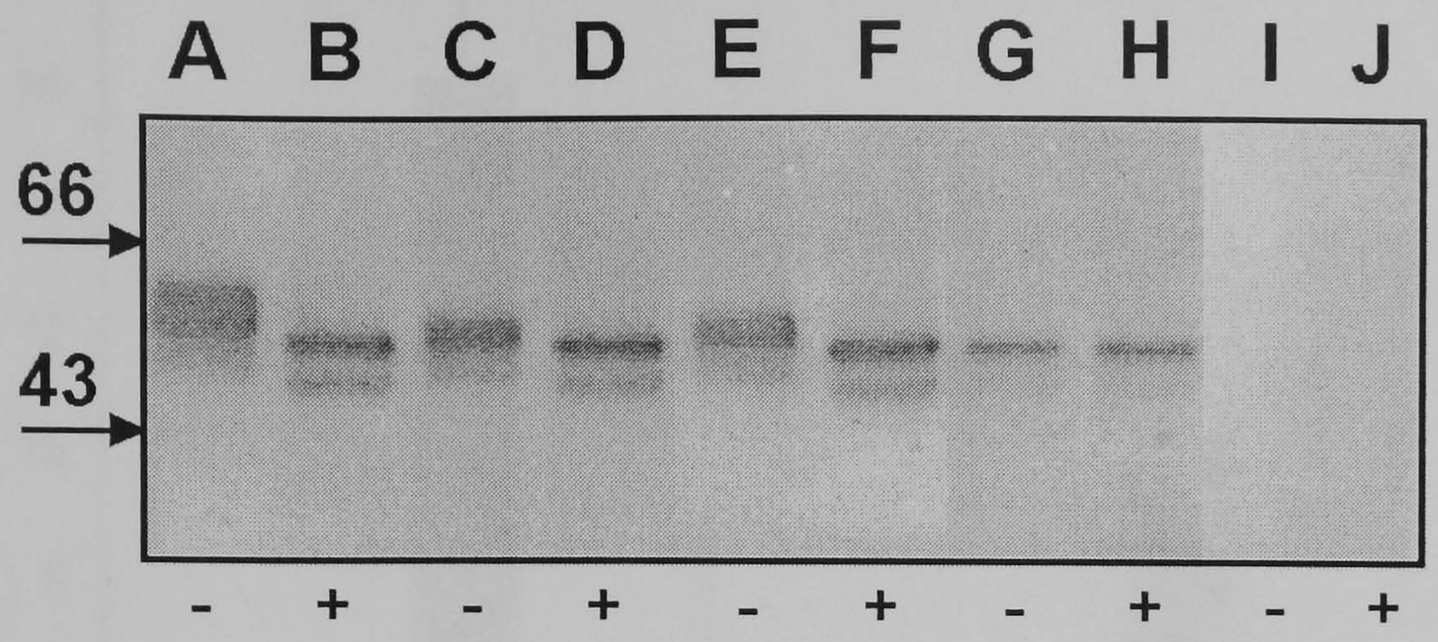
Western blot analysis of the *Xenopus* oocyte membranes untreated (-) and treated (+) with endoglycosidase F/ N-glycosidase F, as described in **section 5.2.1**. Each lane contained 5 μg total membrane protein from oocytes injected with water (*lanes* A and B), or with 10 ng of cRNA encoding rCNT1 (*lanes* C and D), hCNT1 (*lanes* E and F), rCNT2 (*lanes* G and H) and hCNT2 (*lanes* I and J). Oocytes were incubated at 18°C for 3 days following injection, as described in **section 5.2.1**. Detection employed affinity-purified antibodies (2 μg/ml) raised against peptide C as described in **chapter 3**. Molecular mass markers on *left* are in kDa.

Figure 5.5 Kinetic comparison of rCNT1 deglycosylation mutants



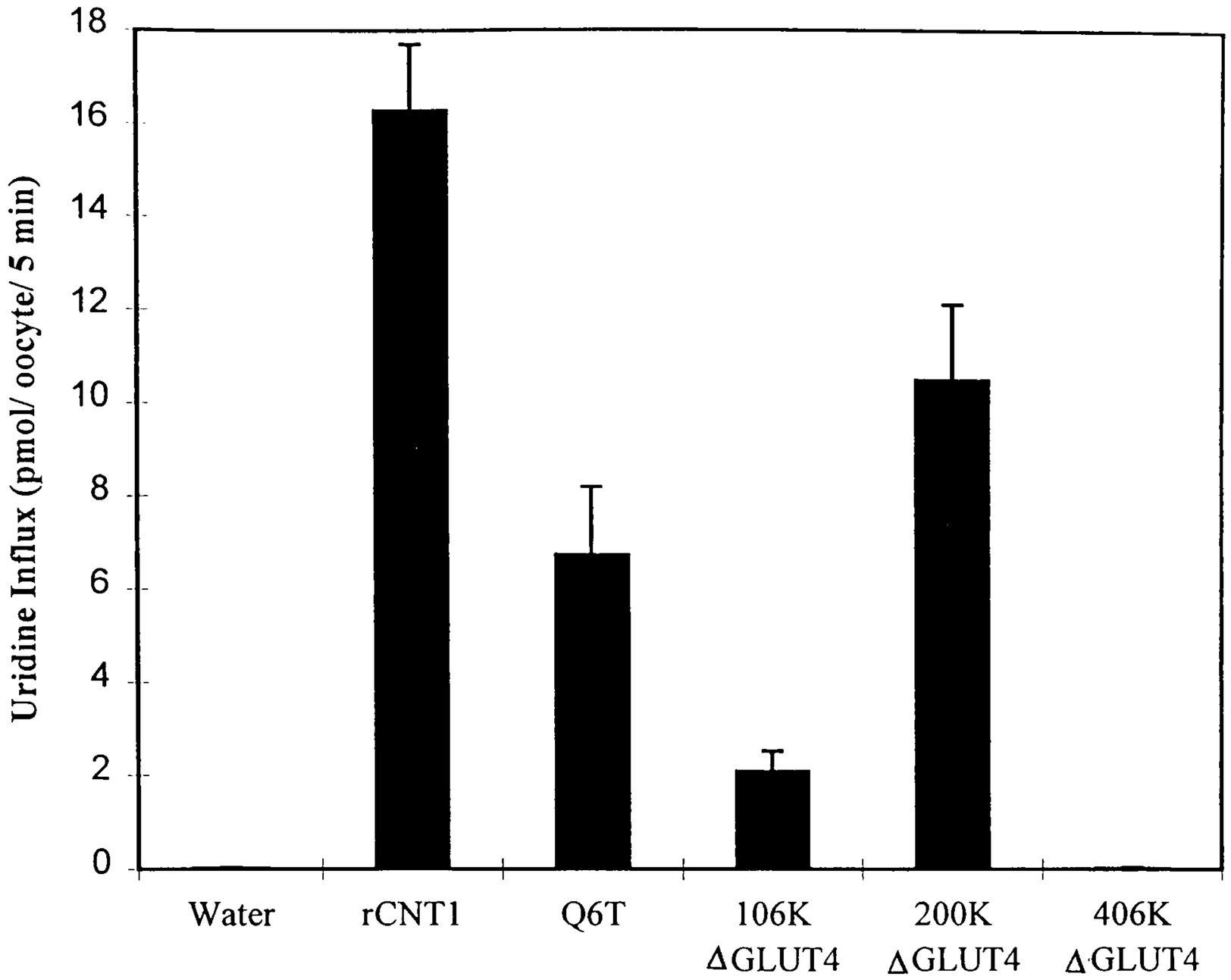
Oocytes were injected with 10 ng of cRNA encoding each of the transporters, including water injected control, prior to incubation at 18°C for 3 days. Fluxes of 10 μM uridine for 5 min at 20°C were determined as described by Huang *et al.* [81]. The data illustrated above was obtained from Dr S. Y. M. Yao (University of Alberta, Canada).

Figure 5.6 Extent of wild-type rCNT1 glycosylation



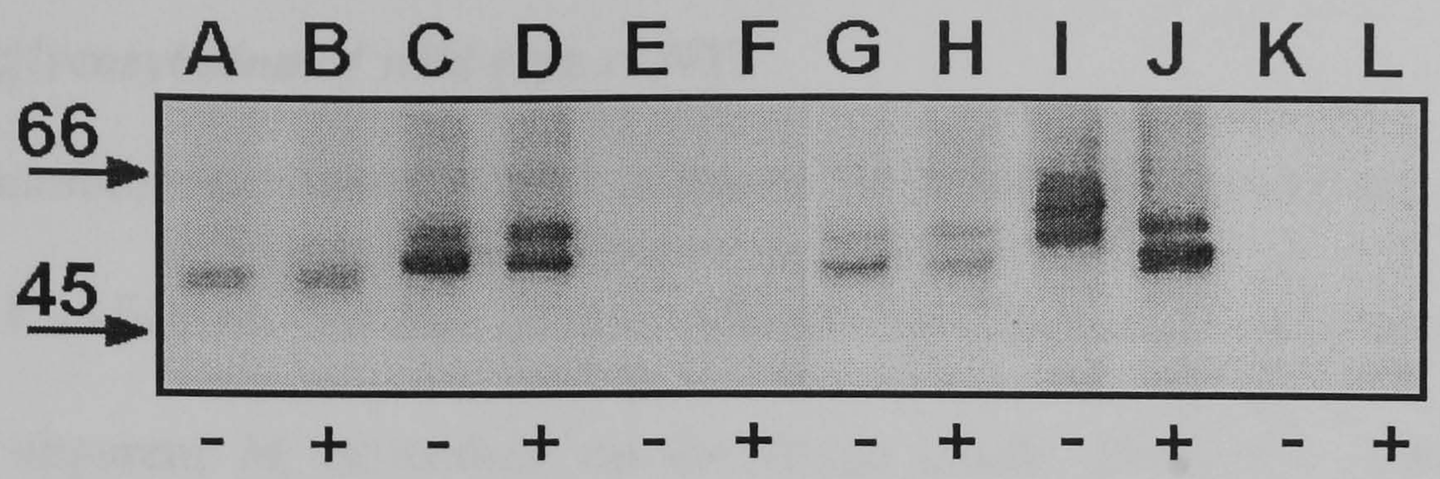
Western blot analysis of the *Xenopus* oocyte membranes untreated (-) and treated (+) with endoglycosidase F/ N-glycosidase F, as described in **section 5.2.1**. Each lane contained 5 μg total membrane protein from oocytes injected with 10 ng of cRNA encoding wild-type rCNT1 (*lanes* A and B), N605T mutant (*lanes* C and D), N643T mutant (*lanes* E and F), N605T/ N643T mutant (*lanes* G and H), and water (*lanes* I and J). Oocytes were incubated at 18°C for 3 days following injection, as described in **section 5.2.1**. Detection employed affinity-purified antibodies (2 μg/ml) raised against peptide C as described in **chapter 3**. Molecular mass markers on *left* are in kDa.

Figure 5.7 Kinetic comparison of rCNT1 topological mutants



Oocytes were injected with 10 ng of cRNA encoding each of the transporters, including water injected control, prior to incubation at 18°C for 3 days. Fluxes of 10 μM uridine for 5 min at 20°C were determined as described by Huang *et al.* [81]. The data illustrated above was obtained from Dr S. Y. M. Yao (University of Alberta, Canada).

Figure 5.8 Glycosylation state of rCNT1 mutants



Western blot analysis of the *Xenopus* oocyte membranes untreated (-) and treated (+) with endoglycosidase F/ N-glycosidase F, as described in **section 5.2.1**. Each lane contained 5 μg total membrane protein from oocytes injected with 10 ng of cRNA encoding the 106-rCNT1/ GLUT4 insert (*lanes* A and B), 200-rCNT1/ GLUT4 insert (*lanes* C and D), 406-rCNT1/ GLUT4 insert (*lanes* E and F), Q6T mutant (*lanes* G and H), wild-type rCNT1 (*lanes* I and J), and water (*lanes* K and L). Oocytes were incubated at 18°C for 3 days following injection, as described in **section 5.2.1**. Detection employed affinity-purified antibodies (2 μg/ml) raised against peptide C as described in **chapter 3**. Molecular mass markers on *left* are in kDa.

5.3 Results and discussion

5.3.1 Deglycosylation of wild-type rCNT1

The immunoblot analyses of rCNT1 in the rat intestine, kidney and liver described in **chapter 4** indicated that this protein was present as a number of isoforms with differing apparent M_r depending on the tissue under observation. Based on this finding it was postulated that these differences might be due to variations in the extent of rCNT1 glycosylation. To test this theory rCNT1 mRNA was expressed in *Xenopus* oocytes and the resulting membranes treated with the deglycosylating enzyme endoglycosidase F, as described in **section 5.2.1**, before examination of the mobility of the transporter on Western blots stained with antibody against rCNT1 peptide C. **Figure 5.4** demonstrates that rCNT1, which normally migrates with an apparent M_r of 63 000, decreases in size to a protein with an apparent M_r 56 000 upon enzyme treatment, thus indicating that this protein exists in a glycosylated form. In addition to testing rCNT1 the three homologues hCNT1, rCNT2 and hCNT2 were treated with the same enzyme to observe if similar effects were achieved. The human homologue of rCNT1, hCNT1 was detected on the blots as a band that migrated with an apparent M_r of 66 000 that decreased to 61 000 on enzyme treatment. This finding indicated that like rCNT1, hCNT1 bears one or more *N*-linked oligosaccharides. However, no cross-reactive bands were seen on the blots of membranes expressing rCNT2 and hCNT2, indicating that the anti-peptide C antibodies were specific for the CNT1 homologues. This latter information confirmed that the observations made in **chapter 4** using these antibodies reflected the presence of rCNT1 and not the purine-specific homologue rCNT2.

5.3.2 Extent of wild-type rCNT1 glycosylation

As discussed in the previous section rCNT1 has been demonstrated to bear at least one *N*-linked oligosaccharide. Although there are three *N*-linked glycosylation consensus sites present in the peptide sequence the predicted arrangement of rCNT1 in the membrane [81] suggests that only those at positions 605 and 643 would be accessible to the glycosylating machinery of the endoplasmic reticulum. In order to determine which, if either, of these sites is in fact glycosylated, they were removed either independently or collectively by site-directed mutagenesis as described in **section 5.2.2**. The resultant mutants expressed in *Xenopus* oocytes showed no significant differences in transport activity from that of wild-type rCNT1 (**Figure 5.5**). Subsequently the glycosylation states of these mutants were assessed by endoglycosidase F treatment to determine if *N*-linked glycosylation was present, as illustrated in **Figure 5.6**. The elimination by mutation of either asparagine, singly, resulted in a decrease in apparent M_r from the wild-type-value of 63 000 to one of 60 000. Furthermore, after treatment with endoglycosidase F the mutants both migrated with an apparent M_r of 56 000, similar to that of the deglycosylated wild-type protein. In addition to the major band with an apparent M_r of 56 000, a minor band of apparent M_r 50 000 was also evident on blots of the deglycosylated protein samples (**Figure 5.4**). As can be seen from the comparison of **Figures 5.4** and **5.6**, this was not always present, and may correspond to either a degradation product of rCNT1 or to a unique conformational state of the SDS-denatured protein, dependent upon the exact conditions for sample preparation. Elimination of both of the potential glycosylation sites collectively in the double mutant N605T/ N643T resulted in a protein which migrated with the same apparent M_r as the deglycosylated wild-type

transporter and which was not affected by endoglycosidase F treatment. This lack of effect confirmed the hypothesis, mentioned earlier, that the potential glycosylation site at position 543 is not in fact glycosylated, presumably because of its location in a membrane-spanning portion of the transporter. As observed in **chapter 4** wild-type rCNT1 appears *in vivo* as two isoforms, with apparent M_r values of 60 000 and 63 000 respectively. On comparison with the single mutants it would appear that in the natural state rCNT1 is present as a heterogeneous population with glycosylation occurring at either or both of the asparagine residues at positions 605 and 643.

The data described above indicated that in the wild-type rCNT1 molecule one or both of the asparagine residues at positions 605 and 643 are glycosylated. This finding was interesting in two respects. Firstly, these residues lie within the hydrophilic *C*-terminal region of the rCNT1 molecule, that in the published model of Huang *et al.* for the protein topology was predicted to be cytoplasmic [81]. Clearly, glycosylation of these sites indicates that the *C*-terminus is extracellular and that the model, at least in this respect, is incorrect. Such a finding exemplifies the caution with which models produced by use of predictive algorithms must be treated: while such models provide a useful starting point it is essential that their predictions are tested by direct experimentation.

A second interesting outcome of the glycosylation experiments was the finding that rCNT1 can apparently be simultaneously glycosylated at two sites. It has been suggested that the majority of *N*-linked glycosylated proteins are only glycosylated at one extracellular consensus site, the only known exception to date being the α-subunit of the sodium channel [182]. This exception probably results from the fact that the channel structure has arisen from an internal gene duplication event, such

that two glycosylation sites are present in distant regions of the protein. In contrast the two accessible *N*-linked glycosylation sites of rCNT1, which the present study has demonstrated can be simultaneously glycosylated, are present within the same extracellular segment. While other membrane proteins that are known to possess more than one potential *N*-linked glycosylation site in the same region, usually only one of these is glycosylated. Such a situation can be exemplified by bovine rhodopsin which has two potential glycosylation sites in its *N*-terminus, only one of which is glycosylated [183]. A possible explanation for this phenomenon may be that if two potential glycosylation sites occur in the same extracellular hydrophilic segment then they must be a defined distance apart in order for both to be glycosylated. In the case of rCNT1 the two sites are 37 residues apart, whereas in bovine rhodopsin the two consensus sites are only thirteen residues apart, a distance that may be insufficient to allow access of the glycosylating machinery of the endoplasmic reticulum to each.

5.3.3 Topological analysis by glycosylation scanning mutagenesis

The aglyco-mutant of rCNT1 described in **section 5.3.2**, being functional and thus presumably arranged in the membrane in the same fashion as the wild-type rCNT1, provided a suitable template for further investigation of the transporter topology by glycosylation scanning mutagenesis. Two separate approaches were used to incorporate consensus *N*-linked glycosylation sites into the N605T/ N643T double mutant, either the modification of a single amino acid residue or the introduction of the glycosylated hydrophilic loop of human GLUT4. The regions of rCNT1 selected on which to perform these mutations were within several of the predicted hydrophilic loops, based on the model of Huang *et al.* [81]. For the first approach, an attempt was made to construct three consensus glycosylation sites by making the mutations Q6T,

A54T and E503N as described in **section 5.2.3**. Unfortunately, sequence analysis of the constructs showed that only the Q6T mutant had in fact been successfully generated, but this did result in the incorporation of a consensus glycosylation site in the *N*-terminus of the aglyco-rCNT1 mutant. Expression of the Q6T mutant in *Xenopus* oocytes indicated that this mutant retained approximately 40 % transport activity of that displayed by wild-type rCNT1 (**Figure 5.7**). Subsequently, *Xenopus* oocyte membranes prepared from the Q6T mutant were treated with endoglycosidase F (**Figure 5.8**). The result of which indicated that there was no change in apparent M_r of this mutant when treated with endoglycosidase F, demonstrating that the introduction of an *N*-linked glycosylation site into the *N*-terminus of the aglyco form of rCNT1 did not change the glycosylation state of the aglyco-mutant. Deriving from this two interpretations could be reached. Either the *N*-terminus was extracellular but was not glycosylated or that it was cytoplasmic and hence could not be glycosylated. The former of these interpretations was possible but unlikely as this mutant contains a glycosylation site which is located 68 residues from where the peptide backbone is first predicted to enter the membrane and thus would be freely accessible to the glycosylating machinery of the endoplasmic reticulum. Furthermore, based on sequence analyses discussed in **chapter 1**, the *N*-terminus of rCNT1 lacks a signal sequence and is very hydrophilic. This hydrophilicity indicates that it would be very energetically expensive to translocate this region of rCNT1 across the lipid bilayer. Therefore, together this information suggested that the lack of glycosylation observed with the Q6T mutant was due to the *N*-terminus residing on the cytoplasmic face of the membrane. Based on this suggestion that the *N*-terminus is cytoplasmic and former data indicating that the *C*-terminus is extracellular this dictates that rCNT1

must have an odd number of TMDs and not fourteen as predicted by the original model by Huang *et al.* [81]. As previously discussed in **chapter 1** rCNT1 was predicted to have between 10 and 14 TMDs. Analysis of the rCNT1 amino acid sequence using a number of hydrophobicity scales indicated that the prediction of helices 6, 8 and 12 being TMDs was weaker than for the remaining eleven helices. Thus the absence of any one of these helices would result in a predicted model with thirteen TMDs with both termini in the correct orientation when considering the fore-mentioned evidence. Therefore, based on this information rCNT1 has been proposed to exist in the lipid bilayer as one of the three new putative models illustrated in **Figure 5.9**. Of these models, the one which predicts that the "helix" 6 region resides outside of the lipid bilayer is the most likely since while the predictions that helices 8 and 12 are TMDs were weaker than for the other regions of the protein, they were predicted to be membrane-spanning depending on whatever hydrophobicity scale was used in the analyses. In contrast, putative "helix" 6 was not identified as a TMD in any of the analyses performed. Further evidence to support any of these three models came from the insertion of the human GLUT4 hydrophilic loop between predicted TMDs 4 and 5. Normally the GLUT4 loop *in vivo* is glycosylated so by placing the cDNA encoding the GLUT4 hydrophilic loop in the cDNA of another protein, in the region of a predicted extracellular hydrophilic loop, should result in that protein becoming glycosylated. Doing this with rCNT1 at the fore-mentioned position did not produce a glycosylated mutant protein. Therefore it has been concluded that the hydrophilic loop between predicted TMDs 4 and 5 was cytoplasmic, as suggested by either of the two new putative models.

Two further mutants were attempted to be generated by single residue mutation. These involved the conversion of alanine 54 to threonine and glutamate 503 to asparagine. However following genetic manipulation of the rCNT1 cDNA no such mutants could be detected on sequencing. Analysis of the primer sequences for these two mutations (**Table 5.1**) did not identify any problems with the exception of the primers having the ability to form primer-dimers. However this is a standard event using this strategy of mutagenesis and has not proved to a problem in the generation of other mutants. This information suggested that the inability to generate these two mutants was a practical problem. Unfortunately time did not permit this to be rectified. Although altering both the primer concentrations and the annealing temperatures may prove suitable starting points.

The second approach used to introduce *N*-linked glycosylation consensus sequences into the rCNT1 aglyco-mutant involved mutagenesis to incorporate a foreign glycosylation site from another protein. This was achieved by the incorporation of the hydrophilic loop of human GLUT4 which possesses an *N*-linked glycosylation site, known to be glycosylated *in vivo*. Locations chosen for introduction of the exogenous GLUT4 hydrophilic loop were following leucine 106, arginine 200 and leucine 406 since these residues were predicted in the original model by Huang *et al.* to be located in the first, fourth and tenth predicted hydrophilic loops respectively [81]. A diagrammatic representation of these locations can be seen in **Figure 5.3**. Synthesis of the three mutants incorporating this exogenous loop following residues 106, 200 and 406 (designated 106*K*ΔGLUT4, 200*K*ΔGLUT4 and 406*K*ΔGLUT4 respectively), resulted in successful generation of the former two mutants but a mutant containing a single insert of the loop could not be obtained for the 406

mutant, even after several attempts at subcloning with various amounts of insert present. Due to this repetitive failure to isolate the 406 mutant containing a single insert a 406 construct containing two tandem copies, in the same orientation, of the cDNA region encoding the GLUT4 loop was isolated and used in subsequent investigations. Incorporation of the loop succeeding the leucine at position 106 resulted in a mutant that demonstrated approximately 12 % transport activity of the wild-type rCNT1 (**Figure 5.7**), indicating that although this mutation resulted in the insertion of 58 residues into rCNT1, the protein was still partially functional. Western blotting of *Xenopus* oocyte membranes expressing this mutant indicated an apparent *N*-terminal truncation of the aglyco-rCNT1, as observed by the decrease in apparent M_r from 56 000 to 50 000 (**Figure 5.8**). Such a decrease was equivalent to loss of the *N*-terminus of rCNT1 to the region of the inserted loop, although the exact location could not be determined without amino acid sequencing which was not performed. Alternatively, the mutation generated in the 106*K*ΔGLUT4 mutant may facilitate the protein in adopting the conformational isoform of rCNT1 with an apparent M_r of 50 000 which has previously been observed (**Figure 5.6**). More important than the 106*K*ΔGLUT4 protein migrating with a reduced apparent M_r of 50 000 was the observation that this protein was unaffected by endoglycosidase F treatment. This initially suggests that the hydrophilic loop connecting putative helices 1 and 2, in the original model by Huang *et al.* [81], is cytoplasmic, although it may be that the insertion of the GLUT4 loop into this very small loop between putative helices 1 and 2 has prevented the correct insertion of the protein into the membrane. Therefore a hydrophilic loop that is predicted to be extracellular, in the

three new models, has not been translocated correctly, resulting in the protein being unglycosylated.

Introduction of the exogenous GLUT4 loop into the aglyco-rCNT1 after arginine 200 resulted in a mutant that demonstrated approximately 65 % of the transport activity demonstrated by wild-type rCNT1 (**Figure 5.7**), thus indicating that this protein was functional. However expression of this mutant in *Xenopus* oocytes resulted in a protein whose apparent M_r did not change from 56 000 on treatment with endoglycosidase F (**Figure 5.8**). Considering both of these pieces of evidence it was assumed that the 200*K*ΔGLUT4 mutant had folded correctly and that it was not glycosylated. This confirmed the predicted changes in conformation discussed previously in this section. Briefly, if the *N*-terminus was cytoplasmic and the *C*-terminus extracellular with either helix 6, 8 or 12 not residing in the lipid bilayer then this would place the hydrophilic loop, between helices 4 and 5 into which the exogenous hydrophilic loop was placed, on the cytoplasmic side of the membrane resulting in this loop not being glycosylated as observed.

The final mutant made by this strategy involved the insertion of the exogenous loop following leucine 406. Unfortunately, attempts at expressing this mutant in *Xenopus* oocytes were unsuccessful: no transport activity was measurable and no rCNT1 protein was detectable on Western blots stained with antibody raised against rCNT1 peptide C (**Figures 5.7** and **5.8** respectively). An explanation for these observations may stem from the duplication of the GLUT4 hydrophilic loop within this mutant causing excessive distortion on the protein conformation which has subsequently resulted in the lack of any significant expression of functional rCNT1. Western blot analysis further confirmed that the lack of transport activity was not due to incorrect

folding of rCNT1, since rCNT1 would still have been detected. However this was not the case thus confirming that if rCNT1 was expressed then the levels were insignificant.

5.4 Summary

Initial evidence for the topology of rCNT1 came from the observation that rCNT1 was present in rat tissues as a number of different isoforms, based on the variation observed in apparent M_r from 60 000 to 63 000 (**chapter 4**). Assuming that it was *N*-linked glycosylation that was being observed then treatment of rCNT1 with the deglycosylating enzymes endoglycosidase F/ *N*-glycosidase F should result in a decrease in apparent M_r. To achieve this rCNT1 was expressed in *Xenopus* oocytes and the membranes subsequently treated with the fore-mentioned enzyme. The result of which was a shift in apparent M_r from 63 000 to 56 000, thus indicating that rCNT1 was indeed glycosylated. This result was very significant since the amino acid sequence of rCNT1 only contained three potential glycosylation sites at asparagines 543, 605 and 643. The former was predicted to reside within putative TMD thirteen and therefore would not be accessible to the glycosylating machinery of the endoplasmic reticulum, consequently indicating that either or both of the asparagines at positions 605 and 643 were glycosylated. The significance of this observation was that both of these residues are located in the *C*-terminus of rCNT1 which has previously been predicted to be cytoplasmic [81], however contrary to this the present study has proven that the *C*-terminus is extracellular. Thus with the first piece of experimental evidence to the topology of rCNT1 the original model, predicted by computer algorithmic analyses, has already been dramatically modified.

To further investigate the topology of rCNT1 the technique of glycosylation scanning mutagenesis was adopted. The basis of this strategy is that an *N*-linked glycosylation site is introduced into the predicted hydrophilic loop of the protein backbone with subsequent enzymatic cleavage of any oligosaccharide moieties indicating if that hydrophilic loop is cytoplasmic or extracellular [22]. However prior to adopting this methodology, endogenous *N*-linked glycosylation sites had to be eliminated. This was achieved by mutagenic elimination of asparagines 605 and 643 either independently or collectively. Following expression in *Xenopus* oocytes deglycosylation of these mutants indicated that in wild type rCNT1 either one or both of these residues were glycosylated. Furthermore, as discussed in the previous section, not only is the occurrence of two glycosylated asparagines in the same protein rare but for both of these to occur in the same hydrophilic region, as they do in rCNT1, is unique in that no reported cases of this glycosylation pattern could be located. Furthermore, the absence of glycosylation with the N605T/ N643T double mutant confirmed that asparagine 543 was not glycosylated.

Having produced the aglyco-rCNT1 mutant subsequent mutagenesis was performed to investigate its topology in the lipid bilayer. The first such piece of evidence which aided in the determination of rCNT1 structure was produced by the introduction of a potential *N*-linked glycosylation site into the *N*-terminus of the protein at residue 6. Deglycosylation studies indicated that the Q6T mutant was not glycosylated, which in addition to sequence analyses indicating that this terminus is very hydrophilic and possesses no identifiable signal sequence. This suggests that the *N*-terminus of rCNT1 is cytoplasmic and furthermore indicates that it is an odd number of TMDs that reside in the lipid bilayer, and not fourteen as prediced in the original model [81].

Considering information discussed in the previous section concerning transmembrane prediction algorithms, one of predicted helices 6, 8 or 12 is not a TMD. Taking this into consideration in conjunction with the protein having an odd number of TMDs indicates that any one of these hydrophobic regions may not reside in the lipid bilayer. Collectively the information to date suggests the *C*-terminus is extracellular, the *N*-terminus is cytoplasmic and that one of the predicted helices 6, 8 or 12 does not reside in the membrane, thus allowing for three new models for rCNT1 in the lipid bilayer to be proposed, as illustrated in **Figure 5.9**. Further evidence for the existence of rCNT1 in one or other of these predicted conformations comes from the generation of the introduction of a glycosylation site following arginine 200 which was originally predicted to be extracellular [81] but in this study was demonstrated to be unglycosylated, thus suggesting that it is cytoplasmic as predicted by the three new putative models of rCNT1. Unfortunately, other mutations of rCNT1 generated in this study did not produce any definitive information into the topology of rCNT1.

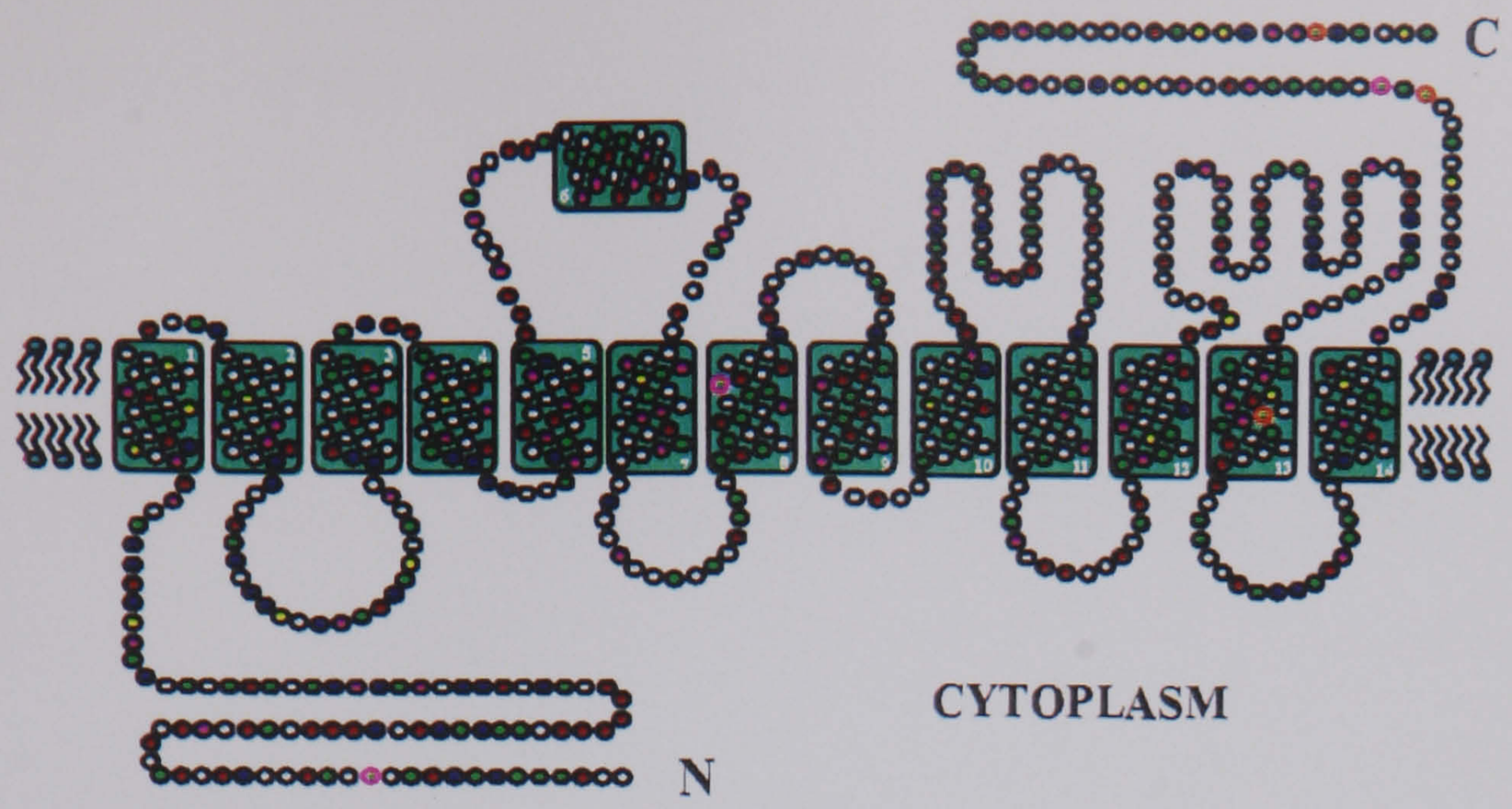
In conclusion, this study has involved a number of preliminary studies into the topology of rCNT1 in the lipid bilayer and has demonstrated that the putative model originally proposed [81] was incorrect. Furthermore, the evidence obtained in the present study has led to the proposal that rCNT1 exists in the lipid bilayer as one of the three new putative models illustrated in **Figure 5.9**. These models now provide a basis for further investigation of rCNT1 topology with the overall aim of elucidating the protein's 3-dimensional structure. In addition, this work illustrates that although computer predicted models provide a starting point for the investigation of

membrane protein topology, such predictive models must be confirmed by experimental evidence.

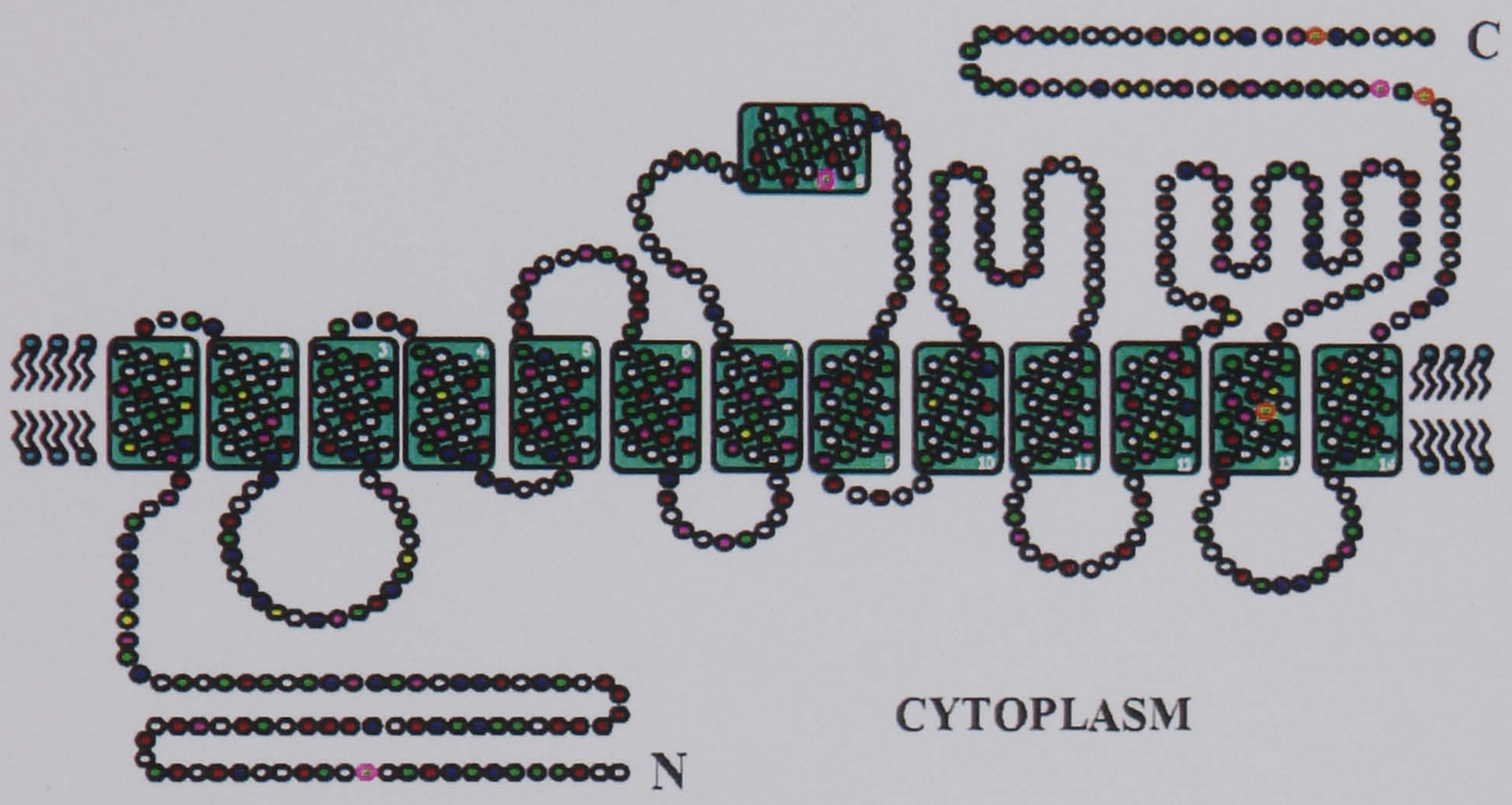
Figure 5.9 Three new putative models for the topology of rCNT1 in the lipid bilayer

Image on opposite page

Evidence discussed in **chapter 5** has led to the adaptation of the original model proposed by Huang *et al.* [81] so that rCNT1 is actually present as one of the three forms represented in *plates* A to C. Residues in the three models have been colour coded as described previously (**Figure 1.2**).



A



B

C

Chapter 6

General discussion

6.1 Distribution of rCNT1 in rat tissues

Recent cloning of the active nucleoside transport protein rCNT1 has allowed for a in-depth study of this protein [81]. In the investigation by Huang *et al.* it was suggested that this protein was present in both the jejunum and kidney. However this conclusion was based on the presence of rCNT1 mRNA which was only evidence that the gene encoding this protein was transcribed and not conclusive proof that rCNT1 was expressed. Therefore to confirm these earlier observations and to further localize the protein at the subcellular level, site-directed polyclonal antibodies were raised to identify rCNT1 as described in **chapter 3**. Of the antibodies raised against the synthetic peptides, A to D, only those antibodies raised against peptides C and D produced significant signals against the protein in the native or denatured state. Additionally, the production of antibodies raised against a rCNT1 *N*-terminal/ GST fusion protein proved essential for investigation of fixed rat tissue. Overall the polyclonal antibodies raised against synthetic peptide C were considered the most reliable as these cross-reacted strongly with a protein in both the jejunum and kidney on Western blotting where rCNT1 mRNA was previously observed by Northern blotting [81]. Additionally these antibodies cross-reacted with a protein in the liver of similar relative molecular mass. In these three tissues the relative molecular mass of rCNT1 was demonstrated to be approximately 63 000 although this varied marginally depending on the tissue. This variation in size suggested that the protein could be present in a number of differing glycosylation states. Evidence shown in **chapter 5**, indicated that in the jejunum and kidney the protein was present with either one or both of its predicted accessible asparagines being glycosylated, residues

605 and 643. In comparison, in the liver only one or the other of these sites is glycosylated, as indicated by the presence of only one apparent isoform.

To further identify the location of rCNT1 expression within the fore-mentioned tissues immunocytochemical analysis of these tissues was performed. In the jejunum rCNT1 was demonstrated to be present at the brush border and not at the basolateral membrane. This observation suggested that within this tissue rCNT1 is involved in the active uptake of nucleosides from the intestinal lumen but is not involved the translocation of nucleosides out of these enterocytes at the basolateral membrane. As discussed in **chapter 4** this distribution parallels that reported for other sodium symport systems that are present in the intestinal membranes. For example, the sodium-dependent glucose transporter SGLT1 is located in the brush border membrane where it is involved in the active uptake of glucose from the intestinal lumen. Like rCNT1, it is absent from the basolateral membrane, where the complementary protein GLUT2 is present for the facilitative translocation of glucose down its concentration gradient into the nearby capillary network [163]. The restriction of active nucleoside transporter to the brush border suggests that if the nucleosides translocated into the enterocyte by rCNT1 are not all metabolized, then there may be a facilitative nucleoside transporter in the basolateral membrane which allows the transport of nucleosides down the concentration gradient and into the nearby capillaries, paralleling the basolateral facilitative transport demonstrated in human intestine [164], previously discussed in **section 4.4.4**.

Localization of rCNT1 in the kidney proved to show similar distribution within the polarized cells as observed in the intestine. Once again this protein was confined to brush border membranes but only of proximal tubules and was absent from the basolateral membranes. Like the case of SGLT1 with glucose, rCNT1 is thought to

be involved in the active transport of nucleosides from the renal filtrate back into the proximal tubule cells from where they can return to the blood, thus conserving pyrimidine nucleosides.

Within the liver rCNT1 was demonstrated in the bile canalicular membrane of parenchymal cells but absent from the sinusoidal membrane suggesting that here too rCNT1 is involved in the conservation of pyrimidine nucleosides. Recently, work carried out by Felipe *et al.* has verified that rCNT1 is present in the liver in addition the purine specific nucleoside transport protein rCNT2 [89]. Here they suggest that in regenerating rat liver it is rCNT2 that is responsible for the increase in the uptake of uridine due increases in the amount of this protein present, while no such increase is observed with rCNT1. It must be noted however that this increase is observed at the basolateral membrane and not at the brush border suggesting that the uridine uptake is from the blood.

Interestingly, although the antibodies raised against synthetic peptide D did not identify the presence of rCNT1 in the jejunum, kidney or liver, they did identify two proteins in the rat brain with apparent M_r of 45 000 and 56 000 with the latter of these also being present in the spleen. As discussed in **chapter 5**, rCNT1 when unglycosylated has an apparent M_r of 56 000. Therefore indicating that the anti-peptide D antibodies showed preference for unglycosylated rCNT1, which only occurred in the brain and spleen. Furthermore, a private communication from Dr J. D. Young has indicated that an *N*-terminal truncated isoform of rCNT1 showed normal transport activity, suggesting that the protein observed in the brain with an apparent M_r of 45 000 may be a naturally occurring *N*-terminal truncated rCNT1 isoform. Cellular distribution within the brain indicated that rCNT1 was present in

the hippocampus but unfortunately due to poor tissue integrity the subcellular distribution could only be suggested to have been on the cell membrane of pyramidal cells. Definitive subcellular distribution of rCNT1 in the brain will shortly be known due to the establishment of a protocol which can detect the protein in mildly fixed tissue, using the antibodies raised against the rCNT1 *N*-terminal/ GST fusion protein (**chapter 4**). Recently work carried out by Anderson *et al.* has demonstrated that the mRNAs for rCNT1 and rCNT2 are both present in the brain [88]. Furthermore, this work confirmed that rCNT1 was present in the hippocampus. Although Western blotting has demonstrated unglycosylated rCNT1 to be present in the spleen, subcellular localization of this protein could not be identified. A possible explanation for this was that if the protein was present in all cells in this tissue the low levels of expression observed on blotting may have been too low to detect by the method of immunocytochemistry used. Although no *cit*-type transport has been demonstarted in the spleen, *cif*-type activity has been demonstrated in murine splenocytes [71].

In addition to the fore-mentioned tissues rCNT1 was also suggested to be present in the heart, skeletal muscle and testis. Although due the absence of a signal on Western blotting of a protein of the expected apparent M_r, these results require further investigation, as suggested in **chapter 4**. Furthermore, reports by other researchers has only identified active nucleoside transport in human cardiac myocytes [171] but not in skeletal muscle or the testis.

Summarizing, information obtained in this study has demonstrated conclusively that rCNT1 is present in the rat jejunum, kidney and liver, where it is proposed to be involved in the active uptake and salvage of nucleosides from the intestinal lumen, renal filtrate and hepatic bile respectively. Furthermore, unglycosylated rCNT1 has

been demonstrated to be present in the brain in addition to initial indications that rCNT1 is also present in the spleen, heart, skeletal muscle and testis, although these latter observations remain to be confirmed by future studies.

6.2 Elucidating the topology of rCNT1

Cloning of rCNT1 by Huang *et al.* allowed for the first time for the amino acid sequence of this protein to be deduced. Initial analysis of this protein indicated the absence of an *N*-terminal signal sequence, while hydrophobicity analysis suggested that the protein had fourteen potential transmembrane domains. In the light of this information Huang *et al.* proposed that both the *N*- and *C*-termini resided on the cytoplasmic side of the cell membrane while the fourteen potential TMDs transverse the membrane between the two termini. Since this initial model was proposed in 1994 no further information has been published relating to the structure of rCNT1. The present study therefore provides the first experimental evidence relating to the actual topological arrangement of rCNT1 within the lipid bilayer.

The first experimental clue to a topological feature of rCNT1 was derived from the observation that rCNT1 in the rat jejunum, kidney and liver was present as different isoforms, varying in their apparent M_r (**chapter 4**). Based on this observation it was suggested that rCNT1 was glycosylated. According to the protein sequence rCNT1 possessed only three potential *N*-linked glycosylation sites at asparagines 543, 605 and 643. The former of these was predicted to be in the center of a predicted TMD while the remaining two were present in the *C*-terminus, which was predicted to be cytoplasmic. To investigate if rCNT1 was glycosylated its mRNA was injected into

Xenopus oocytes and the resultant membranes treated with an enzyme mixture containing the deglycosylating enzyme endoglycosidase F. The results confirmed the initial observation that rCNT1 was glycosylated due to a decrease in apparent M_r from 63 000 to 56 000. Furthermore, this information indicated that the *C*-terminus of rCNT1 was extracellular and not cytoplasmic as originally predicted [81], since the only predicted accessible *N*-linked glycosylation sites were located in the *C*-terminus.

Having determined the location of the *C*-terminus the most useful information that could be obtained next was the location of the *N*-terminus. This was achieved by the incorporation of an *N*-linked glycosylation site into the *N*-terminus by mutating the glutamine at position 6 to threonine after elimination of the two accessible endogenous glycosylation sites at positions 605 and 643 (**chapter 5**). Following *Xenopus* oocyte expression deglycosylation studies were performed indicating that this Q6T mutant was not glycosylated. It may have been that this terminus was extracellular but was not glycosylated. However, a study performed by Landolt-Marticorena and Reithmeier, involving the known glycosylation state of many membrane proteins, indicated that in all reported cases if an *N*-linked glycosylation site was present at the *N*-terminus of a protein it would be glycosylated [182]. Furthermore, sequence analysis indicated that the *N*-terminus of rCNT1 lacked a signal sequence and that it was very hydrophilic. This hydrophilicity indicated that it would be very energetically expensive to translocate this region of rCNT1 across the lipid bilayer. Therefore it has been suggested that that *N*-terminus of rCNT1 was located on the cytoplasmic face of the cell membrane, although conclusive experimental verification of the location of this terminus remains to be established.

The result of positioning the *N*-terminus on the cytoplasmic face of the membrane and the *C*-terminus extracellular, dictated that there must be an odd number of TMDs located within the bilayer. Analysis of the rCNT1 amino acid sequence using a number of hydrophobicity scales indicated that the prediction of helices 6, 8 and 12 being TMDs was weaker than for the remaining eleven helices. Therefore, based on this information rCNT1 has been proposed to exist in the lipid bilayer as one of the three new putative models illustrated in **Figure 5.9**. Of these models, the one which predicts that the "helix" 6 region resides outside of the lipid bilayer is the most likely since while the predictions that helices 8 and 12 are TMDs were weaker than for the other regions of the protein, they were predicted to be membrane-spanning depending on whatever hydrophobicity scale was used in the analyses. In contrast, putative "helix" 6 was not identified as a TMD in any of the analyses performed. Further evidence to support any of these three models came from the insertion of the human GLUT4 hydrophilic loop between predicted TMDs 4 and 5. Normally the GLUT4 loop *in vivo* is glycosylated so by placing the cDNA encoding the GLUT4 hydrophilic loop in the cDNA of another protein, in the region of a predicted extracellular hydrophilic loop, should result in that protein becoming glycosylated. Doing this with rCNT1 at the fore-mentioned position did not produce a glycosylated mutant protein. Therefore it has been concluded that the hydrophilic loop between predicted TMDs 4 and 5 was cytoplasmic, as suggested by either of the three new putative models.

Interestingly, the present study has also demonstrated rCNT1 to be unique in reference to its glycosylation state. This is the first protein to demonstrate simultaneous glycosylation of two *N*-linked glycosylation sites within the same

extracytosolic hydrophilic region of a membrane protein (**chapter 5**). Although previous work by other researchers has indicated that proteins may possess more than one accessible *N*-linked glycosylation site, generally only one of these sites is glycosylated. For example, bovine rhodopsin has two potential glycosylation sites in its *N*-terminus but only one site can be glycosylated [183]. The only reported case known to date of another protein which is simultaneously glycosylated more than once is the α-subunit of the skeletal muscle sodium channel, although this arises from a duplication in structure with the two sites occurring on separate extracytosolic regions of the protein [182].

6.3 Epilogue

In essence, the present study has not only demonstrated the existence of rCNT1 in the rat jejunum and kidney, as previously suggested, but has also provided conclusive evidence for the expression of this protein the liver and brain. In addition, preliminary data suggested that rCNT1 was also be present in cardiac muscle, skeletal muscle, testis and possibly spleen. Furthermore, the present study has established three novel putative models for the topology of rCNT1 in the lipid bilayer, from which further investigations will be able to elucidate the complete 3-dimensional structure of this protein. There is no doubt that due to the close homology of this rat protein to its human counterpart there will be further extensive research on rCNT1, both at the physiological and structural levels.

References

1. Gorter, E. and Grendel, F. (1925) *J.Exp.Med.* **41**, 439-443

2. Danielli, J.F. and Davson, H. (1935) *J.Cell.Comp.Physiol.* **5**, 495-508

3. Branton, D. (1966) *Proc.Natl.Acad.Sci.USA* **55**, 1048-1056

4. Lenard, J. and Singer, S.J. (1966) *Proc.Natl.Acad.Sci.USA* **56**, 1828-1835

5. Wallach, D.F.H. and Zahler, P.H. (1966) *Proc.Natl.Acad.Sci.USA* **56**, 1552-1559

6. Richardson, S.H., Hultin, H.O., and Fleischer, S. (1964) *Arch. Biochem. Biophys.* **105**, 254-260

7. Frye, L.D. and Edidin, M. (1970) *J.Cell.Sci.* **7**, 319-335

8. Singer, S.J. and Nicolson, G.L. (1972) *Science* **175**, 720-731

9. Tomita, M. and Marchesi, V.T. (1975) *Proc.Natl.Acad.Sci.USA* **72**, 2964-2968

10. Furthmayr, H. (1978) *Nature* **271**, 519-524

11. Henderson, R. and Unwin, P.N.T. (1975) *Nature* **257**, 28-32

12. Maiden, M.C.J., Davis, E.O., Baldwin, S.A., Moore, D.C.M., and Henderson, P.J.F. (1987) *Nature* **325**, 641-643

13. Griffith, J.K., Baker, M.E., Rouch, D.A., Page, M.G.P., Skurray, R.A., Paulsen, I., Chater, K.F., Baldwin, S.A., and Henderson, P.J.F. (1992) *Curr.Opin.Cell Biol.* **4**, 684-695

14. von Heijne, G. (1992) *J.Mol.Biol.* **225**, 487-494

15. Chin, J.J., Jung, E.K.Y., Chen, V., and Jung, C.Y. (1987) *Proc.Natl.Acad.Sci.USA* **84**, 4113-4116

16. Chin, J.J., Jung, E.K.Y., and Jung, C.Y. (1986) *J.Biol.Chem.* **261**, 7101-7104

17. Alvarez, J., Lee, D.C., Baldwin, S.A., and Chapman, D. (1987) *J.Biol.Chem.* **262**, 3502-3509

18. Cairns, M.T., Alvarez, J., Panico, M., Gibbs, A.F., Morris, H.R., Chapman, D., and Baldwin, S.A. (1987) *Biochim.Biophys.Acta* **905**, 295-310

19. Davies, A., Meeran, K., Cairns, M.T., and Baldwin, S.A. (1987) *J.Biol.Chem.* **262**, 9347-9352

20. Davies, A., Ciardelli, T.L., Lienhard, G.E., Boyle, J.M., Whetton, A.D., and Baldwin, S.A. (1990) *Biochem.J.* **266**, 799-808

21. Preston, R.A. and Baldwin, S.A. (1993) *Biochem.Soc.Trans.* **21**, 309-312

22. Hresko, R.C., Kruse, M., Strube, M., and Mueckler, M. (1994) *J.Biol.Chem.* **269**, 20482-20488

23. Belardinelli, L., Linden, J., and Berne, R.M. (1989) *Prog.Cardiovasc.Dis.* **32**, 73-97

24. Jarvis, S.M., Young, J.D., Ansay, M., Archibald, A.L., Harkness, R.A., and Simmonds, R.J. (1980) *Biochim.Biophys.Acta* **597**, 183-188

25. Mathew, A., Grdisa, M., and Johnstone, R.M. (1993) *Biochem.Cell Biol.* **71**, 288-295

26. Murray, A.W. (1971) *Annu.Rev.Biochem.* **40**, 811-826

27. Fox, I.H. and Kelley, W.N. (1978) *Annu.Rev.Biochem.* **47**, 655-686

28. Perigaud, C., Gosselin, G., and Imbach, J.L. (1992) *Nucleosides and Nucleotides* **11**, 903-945

29. Clumeck, N. (1993) *J.Antimicrob.Chemother.* **32**, 133-138

30. Jacobson, K.A., VanGalen, P.J.M., and Williams, M. (1992) *J.Med.Chem.* **35**, 407-422

31. Cohen, A., Ullman, B., and Martin, Jr.D.W. (1979) *J.Biol.Chem.* **254**, 112-116

32. Cass, C.E., Kolassa, N., Uehara, Y., Dahlig-Harley, E.R., Harley, E.R., and Paterson, A.R.P. (1981) *Biochim.Biophys.Acta* **649**, 769-777

33. Boss, G.R. and Seegmiller, J.E. (1982) *Annu.Rev.Genet.* **16**, 297-382

34. Plunkett, W. and Saunders, P.P. (1991) *Pharmacol.Ther.* **49**, 239-268

35. Yarchoan, R., Mitsuya, H., Myers, C.E., and Broder, S. (1989) *N.Engl.J.Med.* **321**, 726-738

36. Gomezangelats, M., Delsanto, B., Mercader, J., Ferrermartinez, A., Felipe, A., Casado, J., and Pastor-Anglada, M. (1996) *Biochem.J.* **313**, 915-920

37. Fideu, M.D. and Miras-Portugal, M.T. (1992) *Neurochem.Res.* **17**, 1099-1104

38. Jones, K.W., Rylett, R.J., and Hammond, J.R. (1994) *Brain.Res.* **660**, 104-112

39. Miras-Portugal, M.T., Sen, R.P., and Delicado, E.G. (1991) *Nucleosides and Nucleotides* **10**, 965-973

40. Sen, R.P., Sobrevia, L., Delicado, E.G., Yudilevich, D., and Miras-Portugal, M.T. (1996) *American Journal Of Physiology-Cell Physiology* **40**, C 504-C 510

41. Sayos, J., Blanco, J., Ciruela, F., Canela, E.I., Mallol, J., Lluis, C., and Franco, R. (1994) *Amer.J.Physiol-Ren.Flu.Elect.Phys.* **36**, F758-F766

42. Meckling-Gill, K.A., Guilbert, L., and Cass, C.E. (1993) *J.Cell.Physiol.* **155**, 530-538

43. Smith, C.L., Pilarski, L.M., Egerton, M.L., and Wiley, J.S. (1989) *Blood.* **74**, 2038-2042

44. Wiley, J.S., Snook, M.B., and Jamieson, G.P. (1989) *Br.J.Haematol.* **71**, 203-207

45. Ruiz-Montasell, B., Martinez-Mas, J.V., Enrich, C., Casado, F.J., Felipe, A., and Pastor-Anglada, M. (1993) *FEBS Lett.* **316**, 85-88

46. Goh, L.B., Sokoloski, J.A., Sartorelli, A.C., and Lee, C.W. (1993) *Biochem.J.* **294**, 693-697

47. Lee, C.W. (1994) *Biochem.J.* **300**, 407-412

48. Torres, M., Delicado, E.G., Fideu, M.D., and Miras-Portugal, M.T. (1992) *Biochim.Biophys.Acta* **1105**, 291-299

49. Cass, C.E. (1995) in *Drug transport in antimicrobial and anticancer chemotherapy* (Georgopapadakou, N.H., ed) pp. 403-451, Marcel Dekker, New York

50. Cass, C.E., Gaudette, L.A., and Paterson, A.R.P. (1974) *Biochim.Biophys.Acta* **345**, 1-10

51. Kwan, K.F. and Jarvis, S.M. (1984) *Am.J.Physiol.* **246**, H710-H715

52. Spector, R. (1982) *Arch.Biochem.Biophys.* **216**, 693-703

53. Geisbuhler, T.P., Johnson, D.A., and Rovetto, M.J. (1987) *Am.J.Physiol.* **253**, C645-C651

54. Lynch, T.P., Paran, J.H., and Paterson, A.R.P. (1981) *Cancer Res.* **41**, 560-565

55. Belt, J.A. and Noel, L.D. (1985) *Biochem.J.* **232**, 681-688

56. Griffiths, M., Beaumont, N., Yao, S.Y.M., Sundaram, M., Boumah, C.E., Davies, A., Kwong, F.Y.P., Coe, I., Cass, C.E., Young, J.D., and Baldwin, S.A. (1997) *Nature Medicine* **3**, 89-93

57. Yao, S.Y.M., Ng, A.M.L., Muzyka, W.R., Griffiths, M., Cass, C.E., Baldwin, S.A., and Young, J.D. (1997) *J.Biol.Chem.* **272**, 28423-28430

58. Plagemann, P.G.W. and Wohlheuter, R.M. (1984) *Biochim.Biophys.Acta* **773**, 39-52

59. Plagemann, P.G., Aran, J.M., and Woffendin, C. (1990) *Biochim. Biophys. Acta* **1022**, 93-102

60. Baer, H.P. and Moorji, A. (1990) *Biochim. Biophys. Acta* **1026**, 241-247

61. Che, M., Nishida, T., Gatmaitan, Z., and Arias, I.M. (1992) *J.Biol.Chem.* **267**, 9684-9688

62. Vijayalakshmi, D. and Belt, J.A. (1988) *J.Biol.Chem.* **263**, 19419-19423

63. Lee, C.W., Cheeseman, C.I., and Jarvis, S.M. (1990) *Am.J.Physiol.* **258**, F1203-F1210

64. Williams, T.C., Doherty, A.J., Griffith, D.A., and Jarvis, S.M. (1989) *Biochem.J.* **264**, 223-231

65. Belt, J.A., Marina, N.M., Phelps, D.A., and Crawford, C.R. (1993) *Advances in Enzyme Regulation* **33**, 235-252

66. Lee, C.W., Sokoloski, J.A., Sartorelli, A.C., and Handschumacher, R.E. (1991) *Biochem.J.* **274**, 85-90

67. Gutierrez, M.M., Brett, C.M., Ott, R.J., Hui, A.C., and Giacomini, K.M. (1992) *Biochim.Biophys.Acta* **1105**, 1-9

68. Gutierrez, M.M. and Giacomini, K.M. (1993) *Biochim.Biophys.Acta* **1149**, 202-208

69. Paterson, A.R.P., Gati, W.P., and Vijayalakshmi, D. (1993) *Proc.Amer.Assoc.Cancer Res.* **34**, 14

70. Flanagan, S.A. and Meckling-Gill, K.A. (1997) *J.Biol.Chem.* **272**, 18026-18032

71. Darnowski, J.W., Holdridge, C., and Handschumacher, R.E. (1987) *Cancer Res.* **47**, 2614-2619

72. Franco, R., Centelles, J.J., and Kinne, R.K. (1990) *Biochim.Biophys.Acta* **1024**, 241-248

73. Plagemann, P.G. (1991) *Biochem.Pharmacol.* **42**, 247-252

74. Betcher, S.L., Forrest, J.N., Knickelbein, R.G., and Dobbins, J.W. (1990) *Am.J.Physiol.* **259**, G504-G510

75. Le Hir, M. (1990) *Renal.Physiol.Biochem.* **13**, 154-161

76. Williams, T.C. and Jarvis, S.M. (1991) *Biochem.J.* **274**, 27-33

77. Jarvis, S.M. (1989) *Biochim.Biophys.Acta* **979**, 132-138

78. Schwenk, M., Hegazy, E., and Del Pino, V.L. (1984) *Biochim.Biophys.Acta* **805**, 370-374

79. Huang, Q.Q., Harvey, C.M., Paterson, A.R., Cass, C.E., and Young, J.D. (1993) *J.Biol.Chem.* **268**, 20613-20619

80. Wu, X., Gutierrez, M.M., and Giacomini, K.M. (1994) *Biochim.Biophys.Acta* **1191**, 190-196

81. Huang, Q.-Q., Yao, S.Y.M., Ritzel, M.W.L., Paterson, A.R.P., Cass, C.E., and Young, J.D. (1994) *J.Biol.Chem.* **269**, 17757-17760

82. Craig, J.E., Zhang, Y.B., and Gallagher, M.P. (1994) *Mol.Microbiol.* **11**, 1159-1168

83. Ritzel, M.W.L., Yao, S.Y.M., Huang, M.Y., Elliott, J.F., Cass, C.E., and Young, J.D. (1997) *Am.J.Physiol.* **41**, C707-C714

84. Che, M.X., Ortiz, D.F., and Arias, I.M. (1995) *J.Biol.Chem.* **270**, 13596-13599

85. Wang, J., Su, S.F., Dresser, M.J., Schaner, M.E., Washington, C.B., and Giacomini, K.M. (1997) *American Journal Of Physiology-Renal Physiology* **42**, F1058-F1065

86. Hansen, J.E., Lund, O., Rapacki, K., and Brunak, S. (1997) *Nucl.Acid.Res.* **25**, 278-282

87. Engelman, D.M., Steitz, T.A., and Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem.* **15**, 321-353

88. Anderson, C.M., Xiong, W., Young, J.D., Cass, C.E., and Parkinson, F.E. (1996) *Mol.Brain Res.* **42**, 358-361

89. Felipe, A., Valdes, R., Delsanto, B., Lloberas, J., Casado, J., and Pastor Anglada, M. (1998) *Biochem.J.* **330**, 997-1001

90. Smith, J.M. (1983) *Nature* **306**, 713-714

91. Saitou, N. and Nei, M. (1987) *Mol.Biol.Evol.* **4**, 406-425

92. Parrysmith, D.J. and Attwood, T.K. (1991) *Computer Applications in the Biosciences* **7**, 233-235

93. Leblanc, G., Pourcher, T., and Zani, M.L. (1993) *Society of General Physiologists Series* **48**, 213-227

94. Wright, E.M., Hager, K.M., and Turk, E. (1992) *Curr.Opin.Cell Biol.* **4**, 696-702

95. Brandl, C.J. and Deber, C.M. (1986) *Proc.Natl.Acad.Sci.USA* **83**, 917-921

96. Deber, C.M., Brandl, C.J., Deber, R.B., Hsu, L.C., and Young, X.K. (1986) *Arch. Biochem. Biophys.* **251**, 68-76

97. Wellner, M., Monden, I., Mueckler, M.M., and Keller, K. (1995) *Eur.J.Biochem.* **227**, 454-458

98. La Rochelle, W.J., Wray, B.E., Sealock, R., and Froehner, S.C. (1985) *J.Cell Biol.* **100**, 684-691

99. Ellman, G.C. (1959) *Arch.Biochem.Biophys.* **82**, 70-77

100. Laemmli, U.K. (1970) *Nature* **227**, 680-685

101. Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc.Natl.Acad.Sci.U.S.A* **76**, 4350-4354

102. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning (A laboratory manual)*, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY

103. Ish-Horowicz, D. and Burke, J.F. (1981) *Nucleic Acids Res.* **9**, 2989-2989

104. Hanahan, D. (1983) *J.Mol.Biol.* **166**, 557-580

105. Barry, E.L., Viglione, M.P., Kim, Y.I., and Froehner, S.C. (1995) *J.Neurosci.* **15**, 274-283

106. Kanai, Y., Lee, W.S., You, G., Brown, D., and Hediger, M.A. (1994) *J.Clin.Invest.* **93**, 397-404

107. Kwong, F.Y.P., Davies, A., Tse, C.M., Young, J.D., Henderson, P.J.F., and Baldwin, S.A. (1988) *Biochem.J.* **255**, 243-249

108. Reagan, L.P., Theveniau, M., Yang, X.D., Siemens, I.R., Yee, D.K., Reisine, T., and Fluharty, S.J. (1993) *Proc.Natl.Acad.Sci.USA* **90**, 7956-7960

109. Bailyes, E.M., Richardson, P.J., and Luzio, J.P. (1987) in *Biological membranes: a practical approach* (Findlay, J.B.C. and Evans, W.H., eds) pp. 73-101, IRL Press, Oxford

110. Craig, T.A. and Kumar, R. (1996) *Biochem.Biophys.Res.Commun.* **218**, 902-907

111. Otto, C.M., Niagro, F., Su, X.Z., and Rawlings, C.A. (1995) *Clinical And Diagnostic Laboratory Immunology* **2**, 740-746

112. Siddle, K. and Soos, M. (1981) in *Monoclonal antibodies and developments in immunoassay* (Albertini, A. and Ekins, R., eds) pp. 53-53, Elsevier, Amsterdam, New York and Oxford

113. Scheidtmann, K.H. (1988) in *Antibodies: A practical approach* (Catty, E., ed) pp. 93-105, IRL Press, Oxford

114. Atassi, M.Z. and Young, C.R. (1985) *CRC Crit.Rev.Immunol.* **5**, 387-401

115. Sutcliffe, J.G., Shinnick, T.M., Green, N., and Lerner, R.A. (1983) *Science* **219**, 660-668

116. Niman, H.L., Houghten, R.A., Walker, L.A., Reisfeld, R.A., Wilson, I.A., Hogle, J.M., and Lerner, R.A. (1981) *Proc.Natl.Acad.Sci.USA* **80**, 4949-4958

117. Muller, S., Plaue, S., Couoppez, M., and Van Regenmortel, M.H.V. (1986) *Mol.Immunol.* **23** , 593-601

118. Young, C.R., Schmitz, H.E., and Atassi, M.Z. (1983) *Mol.Immunol.* **20**, 567-575

119. Du Vigneaud, V., Ressler, C., Swan, J.M., and Roberts, C.W. (1954) *J.Am.Chem.Soc.* **76**, 3115-3125

120. Meienhoffer, J., Schnabel, E., Brenner, H., Brinkhoff, O., Zabel, R., Sroka, W., Klostermeyer, H., Brandenburg, D., Okuda, T., and Zahn, H. (1963) *Z.Naturforsch* **18b**, 1120-1128

121. Merrifield, R.B. (1963) *J.Am.Chem.Soc.* **85**, 2149-2163

122. Letoffe, S., Delepelaire, P., and Wandersman, C. (1996) *EMBO J.* **15**, 5804-5811

123. Saitoh, H., Cooke, C.A., Burgess, W.H., Earnshaw, W.C., and Dasso, M. (1996) *Mol.Biol.Cell* **7**, 1319-1334

124. Haruta, T., Morris, A.J., Rose, D.W., Nelson, J.G., Mueckler, M., and Olefsky, J.M. (1995) *J.Biol.Chem.* **270**, 27991-27994

125. Hausdorff, S.F., Bennett, A.M., Neel, B.G., and Birnbaum, M.J. (1995) *J.Biol.Chem.* **270**, 12965-12968

126. Johnson, K.S., Harrison, G.B.L., Lightowlers, M.W., Ohoy, K.L., Cougle, W.G., Dempster, R.P., Lawrence, S.B., Vinton, J.G., Heath, D.D., and Rickard, M.D. (1989) *Nature* **338**, 585-587

127. Fikrig, E., Barthold, S.W., Kantor, F.S., and Flavell, R.A. (1990) *Science* **250**, 553-556

128. Masson, J., Langlois, X., Lanfumey, L., Gerard, C., Aidouni, Z., Giros, B., Hamon, M., and Elmestikawy, S. (1995) *Journal of Neuroscience Research* **42**, 423-432

129. Hata, A., Mukai, T., Isegawa, Y., and Yamanishi, K. (1996) *Virus Research* **46**, 125-137

130. Ullmann, A. (1984) *Gene* **29**, 27-31

131. Ruther, U. and Mullerhill, B. (1983) *EMBO J.* **2**, 1791-1794

132. Anba, J., Baty, D., Lloubes, R., Pages, J.M., Josephliauzun, E., Shire, D., Roskam, W., and Lazdunski, C. (1987) *Gene* **53**, 219-226

133. Nilsson, B., Abrahmsen, L., and Uhlen, M. (1985) *EMBO J.* **4**, 1075-1080

134. Sassenfeld, H.M. and Brewer, S.J. (1984) *Bio-Technology* **2**, 76-81

135. Bedouelle, H. and Duplay, P. (1988) *Eur.J.Biochem.* **171**, 541-549

136. Sharma, S. and Rose, D.R. (1995) *J.Biol.Chem.* **270**, 14085-14093

137. Toye, B., Zhong, G.M., Peeling, R., and Brunham, R.C. (1990) *Infection And Immunity* **58**, 3909-3913

138. Kaelin, W.G., Pallas, D.C., DeCaprio, J.A., Kaye, F.J., and Livingston, D.M. (1991) *Cell* **64**, 521-532

139. Chittenden, T., Livingston, D.M., and Kaelin, W.G. (1991) *Cell* **65**, 1073-1082

140. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988) *Science* **239**, 487-491

141. Hakes, D.J. and Dixon, J.E. (1992) *Anal.Biochem.* **202**, 293-298

142. Plagemann, P.G. and Aran, J.M. (1990) *Biochim.Biophys.Acta* **1028**, 289-298

143. Thorens, B., Cheng, Z.Q., Brown, D., and Lodish, H.F. (1990) *Am.J.Physiol.* **259**, C279-C285

144. Semenza, G., Kessler, M., Hosang, M., Weber, J., and Schmidt, U. (1984) *Biochim. Biophys. Acta* **779**, 343

145. Davidson, N.O., Hausman, A.M.L., Ifkovits, C.A., Buse, J.B., Gould, G.W., Burant, C.F., and Bell, G.I. (1992) *Am.J.Physiol.* **262**, C795-C800

146. Fideu, M.D., Arce, A., Esquifino, A.I., and Mirasportugal, M.T. (1994) *American Journal Of Physiology-Cell Physiology* **36**, C1651-C1656

147. Coons, A.H., Creech, H.J., and Jones, R.N. (1941) *Proc.Soc.Exp.Biol.Med.* **47**, 200-202

148. Coons, A.H. and Kaplan, M.H. (1950) *J.Exp.Med.* **91**, 1-13

149. Coons, A.H., Leduc, E.H., and Connolly, J.M. (1955) *J.Exp.Med.* **102**, 49-60

150. Sternberger, L.A., Hardy, P.H., Cucclis, I.J., and Mayer, H.G. (1970) *J.Histochem.Cytochem.* **18**, 315-333

151. Coggi, G., Dell'Orto, P., and Viale, G. (1986) in *Immunocytochemistry, Modern Methods and Applications* (Polak, J.M. and Van Noorden, S., eds) pp. 54-70, Butterworth-Heinemann, Oxford

152. Carlton, S.J. (1996) in *Theory and practice of histological techniques* (Bancroft, J.D., Stevens, A., and Turner, D.R., eds) pp. 421-434, Churchill Livingstone, New York

153. Huang, S.N., Minassian, H., and More, J.D. (1976) *Lab.Invest.* **35**, 383-390

154. Taylor, C.R., Shi, S.-R., and Core, R.J. (1996) *App.Immunohistochemistry* **4**, 144-166

155. Morgan, J.M., Navabi, H., Schmid, K.W., and Jasani, B. (1994) *J.Pathol.* **174**, 301-307

156. Polak, J.M. and Van Noorden, S. (1988) in *An introduction to immunocytochemistry: current techniques and problems* (Hammond, C., ed) pp. 8-13, Oxford University Press, Oxford

157. Piffko, J., Bankfalvi, A., Ofner, D., Joos, U., Bocker, W., and Schmid, K.W. (1995) *J.Pathol.* **176**, 69-75

158. Man, Y.G. and Tavassoli, F.A. (1996) *App.Immunohistochemistry* **4**, 139-141

159. Imam, S.A., Young, L., Chaiwun, B., and Taylor, C.R. (1995) *Anticancer.Res.* **15**, 1153-1158

160. Tanimoto, T. and Ohtsuki, Y. (1996) *Amer.J.Vet.Res.* **57**, 853-859

161. Shi, S.-R., Cote, R.J., Yang, C., Chen, C., Xu, H.-J., Benedict, W.E., and Taylor, C.R. (1996) *J.Pathol.* **179**, 347-352

162. Mueckler, M., Caruso, C., Baldwin, S.A., Panico, M., Blench, I., Morris, H.R., Allard, W.J., Lienhard, G.E., and Lodish, H.F. (1985) *Science* **229**, 941-945

163. Wright, E.M. (1993) *Annu.Rev.Physiol.* **55**, 575-589

164. Mun, E.C., Tally, K.J., and Matthews, J.B. (1998) *American Journal Of Physiology-Gastrointestinal And Liver Physiology* **37**, G261-G269

165. Terasaki, T., Kadowaki, A., and Higashida, H. (1993) *Biol.Pharm.Bull.* **16**, 493-496

166. Iseki, K., Sugawara, M., Fujiwara, T., Naasani, I., Kobayashi, M., and Miyazaki, K. (1996) *Biochimica Et Biophysica Acta-Biomembranes* **1278**, 105-110

167. Patil, S.D. and Unadkat, J.D. (1997) *Am.J.Physiol.* **35**, G1314-G1320

168. Schaner, M.E., Wang, J., Zevin, S., Gerstin, K.M., and Giacomini, K.M. (1997) *Pharmacol.Res.* **14**, 1316-1321

169. Hoffman, B.B. and Ziyadeh, F.N. (1995) *Current Opinion in Nephrology & Hypertension* **4**, 406-411

170. Parkinson, F.E. and Clanachan, A.S. (1991) *Br.J.Pharmacol.* **104**, 399-405

171. AbdElfattah, A.S.A., Hoehner, J., and Weschsler, A.S. (1998) *Mol.Cell.Biochem.* **180**, 105-110

172. Foga, I.O., Geiger, J.D., and Parkinson, F.E. (1996) *Eur.J.Pharmacol.* **318**, 455-460

173. Camins, A., Jimenez, A., Sureda, F.X., Pallas, M., Escubedo, E., and Camarasa, J. (1996) *Life Sci.* **58**, 753-759

174. Lamb, D.J., Shubhada, S., Baker, K., and Lee, S.K. (1994) *Journal Of Andrology* **15**, 117-124

175. Kornfeld, R. and Kornfeld, S. (1985) *Annu.Rev.Biochem.* **54**, 631-664

176. Roitsch, T. and Lehle, L. (1989) *Eur.J.Biochem.* **181**, 525-529

177. Reithmeier, R.A.F. and Deber, C.M. (1992) in *The structure of biological membranes* (Yeagle, P., ed) pp. 337-393, CRC Press, Boca Raton, Florida

178. Paulson, J.C. (1989) *Trends.Biochem.Sci.* **14**, 272-276

179. Landolt, C. and Reithmeier, R.A.F. (1994) *Biophys.J.* **66**, A 299-A 299

180. Popov, M., Tam, L.Y., Li, J., and Reithmeier, R.A.F. (1997) *J.Biol.Chem.* **272**, 18325-18332

181. Baldwin, A.N. and Shooter, E.M. (1995) *J.Biol.Chem.* **270**, 4594-4602

182. Landoltmarticorena, C. and Reithmeier, R.A.F. (1994) *Biochem.J.* **302**, 253-260

183. Khorana, H.G. (1992) *J.Biol.Chem.* **267**, 1-4

184. De Koning, H.P., Watson, C.J., and Jarvis, S.M. (1998) *J.Biol.Chem.* **273**, 9486-9494

185. Stow, R.A. and Bronk, J.R. (1993) *J.Physiol.Lond.* **468**, 311-324

Appendix 1

Sequence accession numbers for rCNT1 gene family

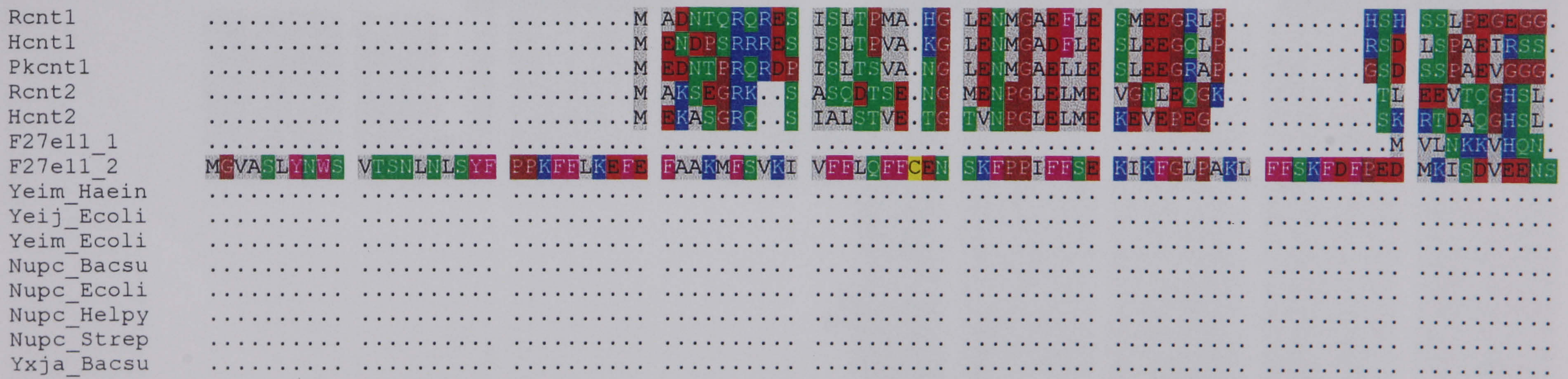
Homologue	Source	Database	Accession number
rCNT1	Rat	gb	U10279
hCNT1	Human	gb	U62968
pCNT1	Pig	gb	AF009673
rCNT2	Rat	gb	U25055
hCNT2	Human	gb	AF036109
hfCNT3	Hagfish	-	-
F27e11_1	*Caenorhabditis elegans*	gb	AF016413
F27e11_2	*C. elegans*	gb	AF016413
YeiM_Haein	*Haemophilus influenzae*	sp	P44742
YeiJ_Ecoli	*Escherichia coli*	sp	P33021
YeiM_Ecoli	*E. coli*	sp	P33024
Nupc_Bacsu	*Bacillus subtilis*	sp	P39141
Nupc_Ecoli	*E. coli*	sp	P33031
Nupc_Helpy	*Helicobacter pylori*	gb	AE000623
Nupc_Strep	*Streptococcus pyogenes*	-	-
Yxja_Bacsu	*B. subtilis*	sp	P42312

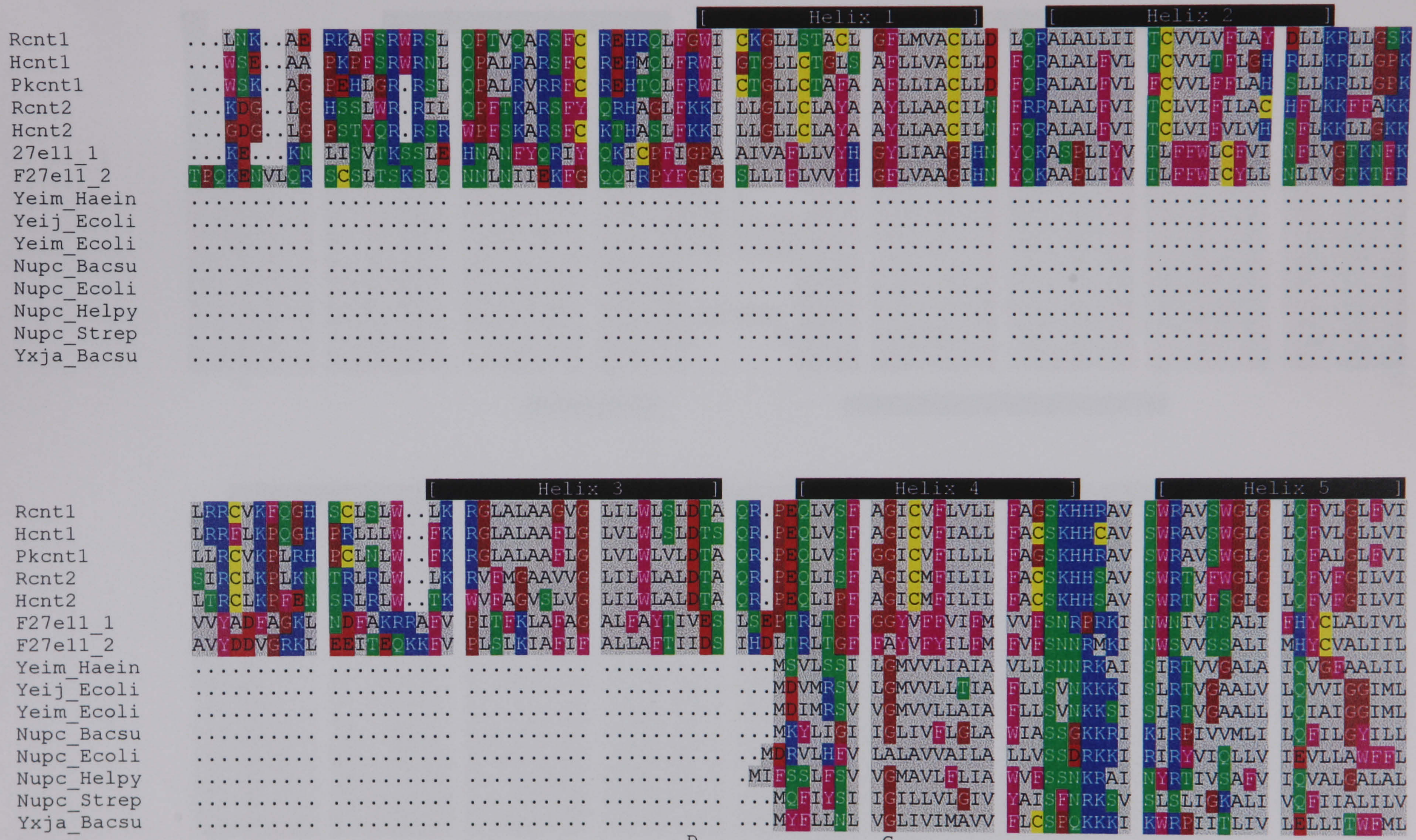
Database accession information for obtaining rCNT1 and its homologues from Genbank (gb) or Swissprot (sp) database services. The sequence for hfCNT1 was supplied by Dr J.D. Young (Canada) while Nupc_Strep was obtained as an open reading frame from contig188 from the *S. pyogenes* sequence database at Oklahoma University.

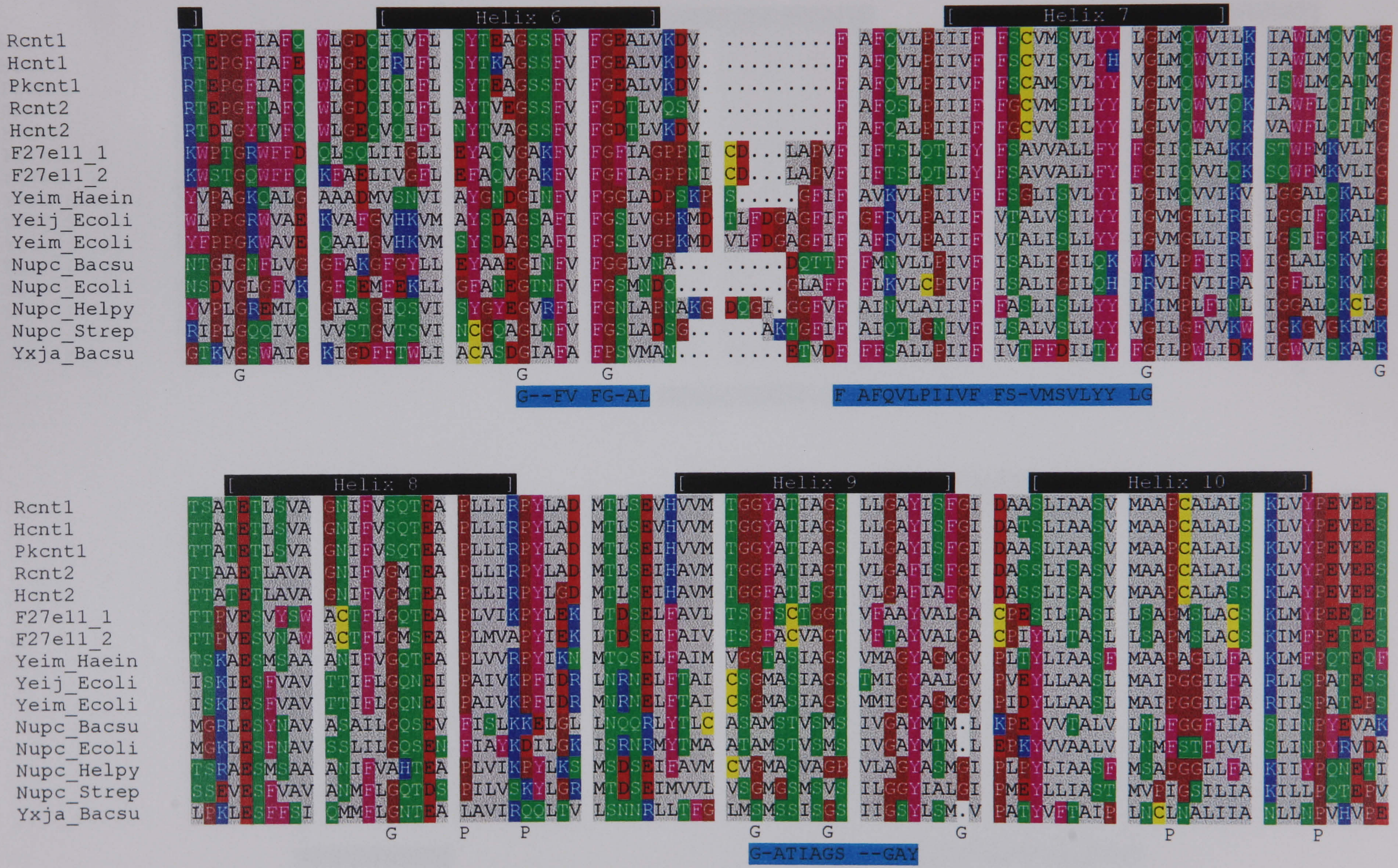
Appendix 2

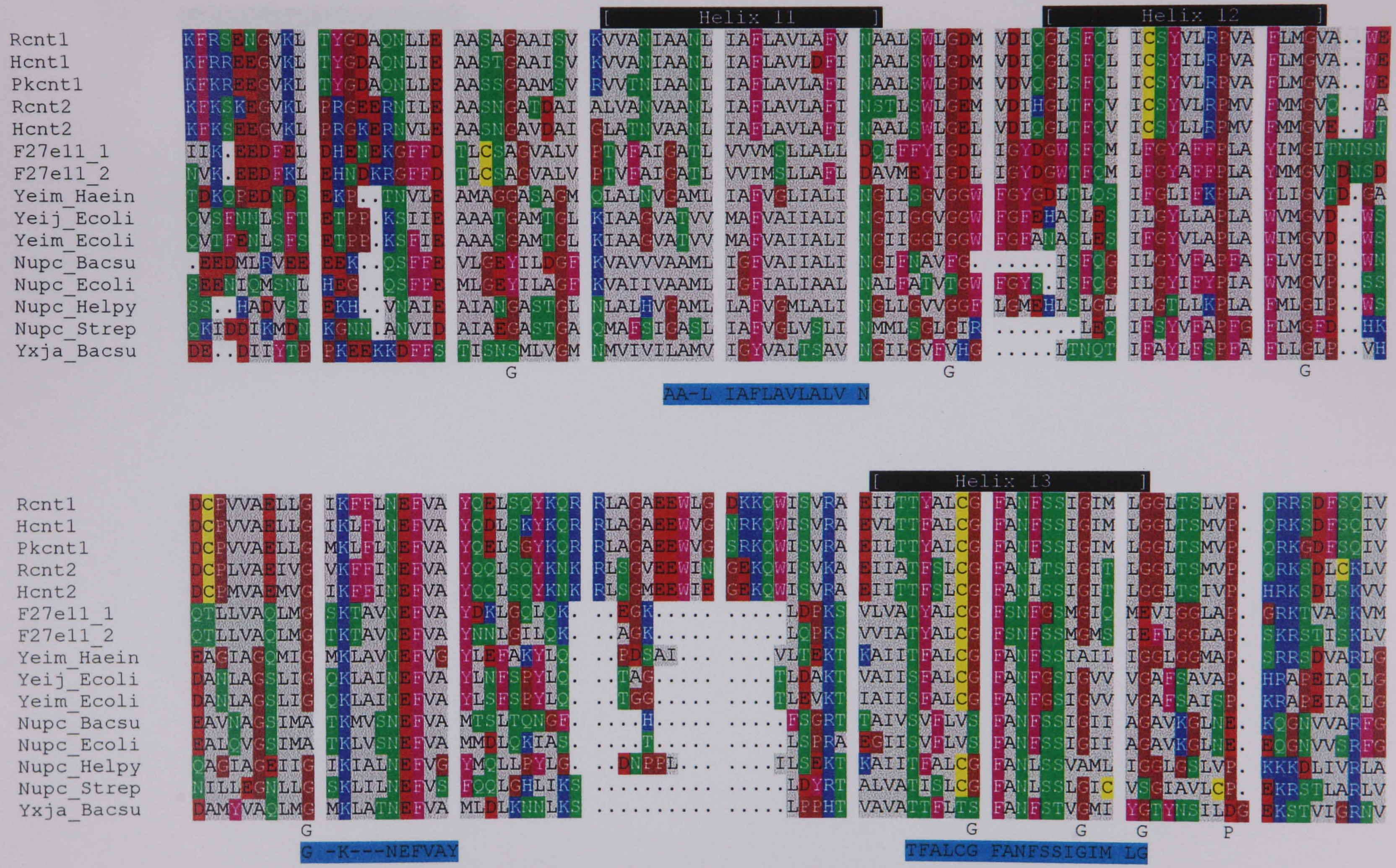
Protein sequence alignment of the rCNT1 gene family

Analysis of the rCNT1 gene family members was performed to provide a multiple sequence alignment file using the software program ECLUSTALW. This was further modified to show amino acid residues with respect to physio-chemical nature based on the SOMAP rules [92], with general nonpolar (A,M,I,L,V), general polar (S,T,N,Q), negatively charged (D,E), positively charged (H,K,R), aromatic (F,W,Y), structure breakers (G,P) and cysteine (C) side-chains. Predicted transmembrane domains are represented above the alignment (*black bars*). Regions of significant sequence homology, where nine or more of the protein sequences demonstrate conservation, are indicated below the alignment (*cyan bars*) in addition to the conserved aspartate (D) in helix three, and the conserved glycine (G) and proline (P) residues.




Helix 1
Helix 2
Helix 3
Helix 4
Helix 5
Rcnt1
Hcnt1
Pkcnt1
Rcnt2
Hcnt2
27e11_1
F27e11_2
Yeim_Haein
Yeij_Ecoli
Yeim_Ecoli
Nupc_Bacsu
Nupc_Ecoli
Nupc_Helpy
Nupc_Strep
Yxja_Bacsu
D
G


Helix 6
Helix 7
Helix 8
Helix 9
Helix 10
Rcnt1
Hcnt1
Pkcnt1
Rcnt2
Hcnt2
F27e11_1
F27e11_2
Yeim_Haein
Yeij_Ecoli
Yeim_Ecoli
Nupc_Bacsu
Nupc_Ecoli
Nupc_Helpy
Nupc_Strep
Yxja_Bacsu
G--FV FG-AL
F AFQVLPIIVF FS-VMSVLYY LG
G-ATIAGS --GAY


[Helix 11]
[Helix 12]
[Helix 13]
Rcnt1
Hcnt1
Pkcnt1
Rcnt2
Hcnt2
F27e11_1
F27e11_2
Yeim_Haein
Yeij_Ecoli
Yeim_Ecoli
Nupc_Bacsu
Nupc_Ecoli
Nupc_Helpy
Nupc_Strep
Yxja_Bacsu
AA-L IAFLAVLALV N
G -K---NEFVAY
TFALCG FANFSSIGIM LG

```
                                [      Helix 14      ]
Rcnt1           LRALITGAFV SLLNACVAGI LYVPRGVEVD CVSLLNQTVS SS.....SFE VYLCCRQVFQ STSS..... .........E FSQVALDNCC
Hcnt1           LRALFTGACV SLVNACMAGI LYMPRGAEVD CMSLLNTTLS SS.....SFE IYQCCREAFQ SVNP..... .........E FSPEALDNCC
Pkcnt1          LRALCTGACV SLVNACVAGI LYVPRGAEVD CVSFLNTTLS SS.....SFE VYQCCRQFFQ STSL..... .........E FSPEALDNCC
Rcnt2           VRALFTGACV SFISACMAGI LYVPRGAETD CVSFLNTNFT NR.....TYE TYVCCRELFQ STLLNGTNMP SFSGPWQDKE SSLRNLAKCC
Hcnt2           VRALFTGACV SLISACMAGI LYVPRGAEAD CVSFPNTSFT NR.....TYE TYMCCRGLFQ STSLNGTNPP SFSGPWEDKE FSAMALTNCC
F27e11_1        LRALCAGAIA CFMNATVAGI LISD..... .......... PV..... ..VCN..... ..SA..... .......... ....NSNSTC
F27e11_2        LRALCAGSIA CFMNATVAGI LISS..... .......... PV..... ..ICQ..... ..SS..... .......... ....QKSDTC
Yeim_Haein      IKAVIAGTLA NLMSATIAGL FIGLGAAAL. .......... .......... .......... .......... .......... ..........
Yeij_Ecoli      LRALAAATLS NLMSATIAGF FIGLA..... .......... .......... .......... .......... .......... ..........
Yeim_Ecoli      LRALAAATLS NLMSATIAGF FIGLA..... .......... .......... .......... .......... .......... ..........
Nupc_Bacsu      LKLLYGATLV SFLSAAIVGL IY........ .......... .......... .......... .......... .......... ..........
Nupc_Ecoli      LKLVYGSTLV SVLSASIAAL VL........ .......... .......... .......... .......... .......... ..........
Nupc_Helpy      LKAVLVGTLS NFMSATIAGL FIGLNAH... .......... .......... .......... .......... .......... ..........
Nupc_Strep      FRAMIGGIAV SMLSAFIVGI VTLF..... .......... .......... .......... .......... .......... ..........
Yxja_Bacsu      WKLLVSGIAV SLLSAAIVGL FVW...... .......... .......... .......... .......... .......... ..........
                      G           G
```

```
Rcnt1           RFYNHTVCT. .....
Hcnt1           RFYNHTICAQ .....
Pkcnt1          RFYNHTICV. .....
Rcnt2           DLYTSTVCA. .....
Hcnt2           GFYNNTVCA. .....
F27e11_1        FRIPS..... .....
F27e11_2        FRIPSG.... .....
Yeim_Haein      .......... .....
Yeij_Ecoli      .......... .....
Yeim_Ecoli      .......... .....
Nupc_Bacsu      .......... .....
Nupc_Ecoli      .......... .....
Nupc_Helpy      .......... .....
Nupc_Strep      .......... .....
Yxja_Bacsu      .......... .....
```