INTERSPECIES ANALYSIS OF THE NCAM HOMOPHILIC BINDING SITE

by

James Kuo

A thesis submitted in conformity with the requirements for the degree of Master of Science
Graduate Department of Biochemistry
University of Toronto

© Copyright by James Kuo 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-29197-9
Abstract

Previous studies have identified the sequence 243-KYSFNYDGSE-252 in the third immunoglobulin-like (Ig3) domain of chick NCAM as the homophilic binding site. Several reports suggest that the rodent NCAM Ig3 domain may not contain homophilic binding activity. To investigate these differences, a series of fusion proteins were created and tested in binding assays. A rat GST-Ig3 fusion protein showed negligible binding activity, confirming the absence of rat Ig3-mediated homophilic binding. Through the creation of chimeric fusion proteins in chick and rat Ig3, putative rat and human homophilic binding sequence have been identified. However, the rat sequence appeared to have a much lower binding affinity than the chick and human sequences. Further mutational analysis suggests that all NCAM species possess the amino acid residues necessary to mediate homophilic binding, but this ability may be hindered by neighbouring residues. We suggest histidine-244 to be responsible for disrupting Ig3-mediated homophilic binding in rat NCAM.
Acknowledgements

I would like to take this opportunity to thank Dr. Siu, my graduate supervisor. He pushed me to think, analyze and formulate my ideas clearly. He always had a gentle reminder to keep working hard. I am a better researcher for his efforts.

I would also like to thank my co-supervisors, Dr. Clarke and Dr. Rini. Over my graduate studies, they provided some key ideas and suggestions which helped to shape my project and to make it a better one.

I would also like to thank my family. My parents, Corrigan and Marisa, and my brother, Eugene, were all very supportive and understanding of the demands required of a graduate student. My family has always been a very strong influence in my life.

And finally, I thank Betty Wong, who has been more than patient and supportive through this intensive time. She was always there to bring me dinner when I forgot to bring it or to drive me home after a very long session at the lab.
# Table of Contents

Abstract ........................................................................................................ ii

Acknowledgments .................................................................................. iii

Table of Contents .................................................................................... iv

List of Figures and Tables......................................................................... v

List of Abbreviations ................................................................................ vi

Chapter 1: Introduction ........................................................................... 1

Chapter 2: Materials and Methods .................................................. 43

Chapter 3: Results .................................................................................. 58

Chapter 4: Discussion ............................................................................. 88

References ............................................................................................... 101
Chapter 1: Introduction

Figure 1 CAM families
Figure 2 Ig-like domain topology
Figure 3 NCAM isoforms
Figure 4 Topology of NCAM isoforms
Figure 5 NCAM-mediated neurite outgrowth response
Table 1 Interspecies sequence identity to chick NCAM
Figure 6 Partial sequence comparison among NCAM species

Chapter 2: Materials and Methods

Table 2 Primers used in fusion protein construction
Figure 7 Construction of recombinant rat GST NCAM Ig3 protein
Figure 8 Construction of chimeric chick NCAM Ig3 proteins
Figure 9 Covasphere binding assay protocol
Table 3 Plating efficiency in the Covasphere assay

Chapter 3: Results

Figure 10 Rat NCAM Ig3 fusion protein construct
Figure 11 Binding of rIg3 Covaspheres
Figure 12 Competition of rIg3-cIg3 Covasphere binding
Figure 13 NCAM chimeric fusion protein constructs
Figure 14 Purification of chimeric fusion proteins
Figure 15 Binding of c(r)Ig3 Covaspheres
Figure 16 Binding of cIg3 and c(h)Ig3 Covaspheres
Figure 17 Competition of c(h)Ig3 Covasphere binding
Figure 18 Rat r(c)Ig3 chimeric fusion protein construct
Figure 19 Binding of r(c)Ig3 Covaspheres
Figure 20 Rat PC12 cell binding
Figure 21 Mutant c(4)Ig3 construction
Figure 22 Binding of c(4)Ig3 Covaspheres

Chapter 4: Discussion

Figure 23 Predicted amino acid arrangement of chick Ig3
Figure 24 NCAM binding models
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Antigen presenting cell</td>
</tr>
<tr>
<td>BCIP</td>
<td>5-Bromo-4-chloro-3-indolyl phosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CAM</td>
<td>Cell adhesion molecule</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>GPI</td>
<td>glycosylphosphatidylinositol</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>HEV</td>
<td>High endothelial venules</td>
</tr>
<tr>
<td>ICAM</td>
<td>Intercellular cell adhesion molecule</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>kDa</td>
<td>KiloDalton</td>
</tr>
<tr>
<td>LFA</td>
<td>Lymphocyte functional antigen</td>
</tr>
<tr>
<td>MAG</td>
<td>Myelin-associated glycoprotein</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MIDAS</td>
<td>Metal ion-dependent adhesion site</td>
</tr>
<tr>
<td>MSD</td>
<td>Muscle specific domain</td>
</tr>
<tr>
<td>NBT</td>
<td>Nitro Blue Tetrazolium</td>
</tr>
<tr>
<td>NCAM</td>
<td>Neural cell adhesion molecule</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PECAM</td>
<td>Platelet endothelial cell adhesion molecule</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonylfluoride</td>
</tr>
<tr>
<td>PSA</td>
<td>Polysialic acid</td>
</tr>
<tr>
<td>PTK</td>
<td>Protein tyrosine phosphatase</td>
</tr>
<tr>
<td>PTP</td>
<td>Protein tyrosine kinase</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>VCAM</td>
<td>Vascular cell adhesion molecule</td>
</tr>
<tr>
<td>VLA</td>
<td>Very late antigen</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction
Cellular adhesion is crucial for establishing, maintaining and regulating tissue structure. Adhesive interactions underlie most major processes during development and are both specific and dynamic in nature. Through these contacts, cells may physically interact with one another and may also transfer specific information or messages. The first cell sorting phenomenon was observed when cells of two different marine sponge species were mixed and allowed to reaggregate (Wilson, 1907). These sponge cells, through cell-cell contacts, are able to undergo cell sorting, exchange information and distinguish between cells of different origin.

These adhesive interactions are mediated by adhesion molecules on the cell surface. Most cell adhesion molecules (CAMs) exhibit narrow specificities and tightly controlled expression patterns. It is through this specificity and regulation of expression that cells modulate their interactions with other cells. For example, the initial down-regulation, and later up-regulation of specific adhesion molecules (i.e. NCAM and cadherins) on the surface of neural crest cells allow these cells to detach from the neural tube, migrate to other sites in the embryo, and firmly attach at these sites (McClay and Ettensohn, 1987). In this manner, the temporal expression of adhesion molecules represents one level of control over cell-cell interactions.

The expression of individual CAMs may be further influenced by the microenvironment of the cell. Often, the expression of specific adhesion molecules in individual cells may be targeted to different regions of the plasma membrane. Expression of specific CAMs at the leading edge of migrating cells allows these CAMs to interact with environmental cues or with cells in its vicinity (Sandig et al., 1996). Similarly, at the tissue level, the expression of specific adhesion molecules (i.e. selectins) on the surface of peripheral lymph node high endothelial venules (HEV) allows homing of leukocytes to that area (Dunon et al., 1996). Therefore, the selective spatial expression of cell adhesion molecules serves as an additional mechanism that regulates cell-cell interactions.
to stabilize specific cell-cell interactions and allow the formation of proper contacts. For instance, zonula adherens junctions are formed by specific CAMs (i.e. cadherins) and represent areas of strong intercellular attachment (Takeichi, 1991). The expression of CAMs thus may contribute to the organization and stability of cell-cell binding and tissue formation.

Furthermore, the adhesive properties of CAMs may be modulated by other environmental factors. Divalent ions (Humphries, 1996), neighboring proteins (Kadmon et al., 1990) and intracellular proteins (Huber et al., 1996) are known to modulate the binding affinity of individual CAMs. In this way, cells expressing identical sets of surface adhesion molecules may further differentiate between similar target cells and substrates.

The fact that cells are able to temporally and spatially regulate their CAM expression patterns, as well as modulate their binding affinity or function allows cells to interact very specifically with their environments. The transfer of information through adhesion molecules and their many roles in developmental processes allows the creation of many diverse and intricate tissue architectures.

II. Cell Adhesion Molecules

A. Strategies to Identify Cell Adhesion Molecules

In order to study specific adhesion mechanisms, it is necessary to identify and isolate the adhesion molecules responsible for mediating the binding. A powerful approach to identify cell adhesion molecules has been developed based on the use of univalent antibody fragments to inhibit cell re-association (Beug et al., 1973; Gerisch, 1980; Edelman, 1985). First, antibodies are raised against intact cells that exhibit stage-specific cell-cell binding activity. Then, after absorption against non-adhesive cells, Fab fragments are prepared and assayed for their ability to inhibit the reassociation of adhesive cells. The major antigens recognized by the absorbed antibody are gel purified and tested for their
successfully to identify many CAMs, such as gp80 in *Dictyostelium* (Muller and Gerisch, 1978), NCAM (Brackenbury et al., 1977; Thiery et al., 1977), Ng-CAM (Grumet and Edelman, 1984), E-Cadherin (Yoshida-Noro et al., 1984; Bertolotti et al., 1980), L1 (Rathjen and Schachner, 1984; Moos et al., 1988) and PECAM-1 (Albelda et al., 1990).

A second approach has been developed for adhesion molecules with unknown ligands (Seed and Aruffo, 1987). In this approach, a non-adhesive cell line (e.g. COS cells) is transfected with the cDNA library of adherent cells and is used to screen for the ligand of a known adhesion molecule. The screening process consists of depositing the transformants on a substratum coated with the adhesion molecule and then washing away the non-adherent cells. The adherent cells are then amplified, the panning process is repeated, and the resultant clones are characterized. ICAM-2 was cloned in this manner, using LFA-1 as its substrate (Staunton et al., 1989).

**B. Binding Mechanisms**

The adhesion of cells of the same lineage or type is often referred to as homotypic cell-cell adhesion. Conversely, adhesion among cells of different lineages is termed heterotypic cell-cell adhesion. A similar nomenclature is used for individual CAM-to-CAM binding. In instances where the adhesion molecule on one cell binds to the same adhesion molecule on another cell, as is the case with gp80 in *Dictyostelium* (Siu et al., 1987; Kamboj et al., 1989), it is referred to as homophilic binding. When two different cell adhesion molecules mediate cell-cell binding, this is known as heterophilic binding. An example of this is the binding between T helper cells and antigen presenting cells (APC), which is mediated by ICAM-1 and LFA-1 on their respective cell surfaces (Staunton et al., 1990). A final type of interaction is exemplified in *Microciona prolifera* sponge cells (Henkart et al., 1973; Misevic and Burger, 1993). Here, cell adhesion molecules on the cell surface are linked by a proteoglycan that acts as a bridging molecule.
Cell adhesion molecules identified in vertebrates have been divided into four major classes: (1) the cadherin family, (2) the integrin family, (3) the selectin family and (4) members of the immunoglobulin (Ig) superfamily. Members of each family typically co-exist on the cell surface and co-operate in mediating cell contact formation.

1. Cadherin Family

Members of the cadherin family of adhesion molecules are responsible for mediating calcium-dependent cell-cell adhesion. Earlier work identified three distinct cadherins expressed on epithelial cells, neural cells and placenta which are designated E-, N- and P-cadherin, respectively (Takeichi, 1977). These classical cadherins show distinct antigenic specificity, tissue distribution and temporal expression, but are all capable of mediating homophilic Ca\(^{2+}\)-dependent cell adhesion (Takeichi, 1990). Additional members of this family have been identified. These include B-cadherin (Napolitano, 1991), R-cadherin (Inuzuka et al., 1991), OB-cadherin (Okazaki et al., 1994), PB-cadherin (Sugimoto et al., 1996), T-cadherins 1 and 2 (Ranscht and Dours-Zimmermann, 1991; Sacristan et al., 1993), and V-cadherin (Heimark et al., 1990).

Typically, classical cadherins contain five CAD repeats, a single transmembrane segment and a cytoplasmic tail. Sequence identity among the various cadherin family members is approximately 50%, with the highest identity observed in the cytoplasmic domain (Takeichi, 1991). Interspecies homology among cadherin members is highly conserved; for example, mouse and chicken N-cadherin is 92% identical at the amino acid level (Takeichi, 1990).

Cadherin-mediated cell-cell binding is homophilic in nature (Nagafuchi et al., 1987), with members exhibiting molecular specificity in this interaction (Nose et al., 1988). This homophilic binding is apparently mediated by a highly conserved HAV
Fig 1. Schematic diagrams representing typical members of the four major CAM families. Cadherins contain five repeated motifs (CAD domains) in their extracellular portion, followed by a transmembrane segment and a cytoplasmic tail (Takeichi, 1991). Members of the Ig-superfamily members typically contain multiple Ig-like domains (Ω) in tandem. Many Ig CAMs also contain fibronectin type three repeats (short open box). Integrins are heterodimers of α and β subunits (Hynes, 1992). Each of the subunits contains a large extracellular domain, a single transmembrane segment and a short cytoplasmic tail. The α subunit contains four sets of EF-hand-like sequences (shaded boxes) which help to form the integrin ligand binding site. The β subunit contains four cysteine rich motifs (diagonally striped boxes). Members of the selectin family contain a N-terminal Ca²⁺-dependent (C-type) lectin binding domain (shaded circle), followed by an EGF-like domain (horizontally striped box) and multiple short consensus repeats (SCR) (open ovals).
Cadherins  Ig-superfamily (i.e. NCAM)  Integrins  Selectins
this HAV sequence partially conferring inter-cadherin specificity (Nose et al., 1990). Recent NMR and crystallography studies have solved the structure of the amino-terminal domains of E-cadherin (Overduin et al., 1995; Nagar et al., 1996) and N-cadherin (Shapiro et al., 1995). The studies of Nagar et al. (1996) show that cadherins form cis-dimers, with their HAV sequences located in the middle of an external, outwardly pointing face. Formation of cell-cell contacts is likely to involve multiple cadherin dimers in anti-parallel interactions forming a "zipper" structure. Furthermore, three Ca\(^{2+}\) ions were shown to be sequestered between each of the CAD repeats and act to stabilize the cadherin into an active conformation. The classic cadherins have been shown to bind the actin filament network via catenin complexes (Aberle et al., 1996). The initial binding of either β- or γ-catenin to the cadherin cytoplasmic tail allows for α-catenins to bind and mediate anchoring to the cytoskeleton (Hinck, et al., 1994). The association of β-catenins in this complex is believed to be important in regulating cadherin function. It may also serve to transduce signals initiated by cadherins via a wingless/Wnt related pathway (Otmar et al., 1996; Gumbiner et al., 1996).

2. Integrin Family

Members of the integrin family serve as cell surface receptors to many different ligands and may mediate both cell-cell and cell-matrix interactions (Hynes, 1987). Integrins are αβ heterodimers which form as a result of restricted, noncovalent pairing of 15 α and 8 β subunits (Varner and Cheresh, 1996). Typically, each subunit contains a large extracellular domain (>100 kDa for the α subunit and >75 kDa for the β subunit), a single transmembrane segment, and a short cytoplasmic tail (Hogevorst et al., 1990; Suzuki and Naitoh, 1990; Tamura et al., 1990).
appears to require the participation of both α and β subunits to mediate high affinity binding (Humphries, 1996). Cross-linking experiments have implicated the EF-hand-like sequences in the α subunit and a region 100-200 residues from the amino-terminus of the β subunit in forming the binding pocket (D'Souza et al., 1990; Smith and Cheresh, 1990).

An additional domain, the I domain, is inserted into a subset of integrin α subunits and has been implicated in both adhesion and integrin activation (Jones, 1996). It is believed that this I domain may contribute an additional ligand binding site to the integrin molecule and may increase its overall binding avidity.

Integrin ligand binding has been shown to be dependent, and also modifiable, by divalent cations, such as Mg$^{2+}$, Ca$^{2+}$, and Mn$^{2+}$ (Mould et al., 1995; Smith, et al., 1994). These studies suggest that for specific ligand-CAM pairings, the presence of specific cation binding sites on the CAM is able to allosterically increase or decrease ligand binding affinity; in the case of α5β1 binding to fibronectin, both Mg$^{2+}$ and Mn$^{2+}$ are suggested to promote ligand binding, whereas Ca$^{2+}$ is believed to inhibit binding. The crystal structure of the Mac-1 I domain has also implicated divalent cations in the binding of ligands, through a proposed metal ion-dependent adhesion site (MIDAS) motif (Lee et al., 1995). From this structure, it was suggested that a divalent cation located in the integrin-binding pocket would act as a bridge between integrin and ligand molecules. Furthermore, the charged residue commonly found in most integrin binding sequences (i.e. RGD, QAGDV or LDV) would complete the coordination complex and bind the metal ion in the MIDAS motif (Humphries, 1996).

In addition to metal ions, both ligands and intracellular proteins have been shown to alter the conformation of integrins (Dedhar and Hannigan, 1996). This innate conformational flexibility serves as the basis for both 'outside-in' and 'inside-out' signal transduction and allows these integrin molecules to participate in many important cellular events. Integrins have been shown to be involved in cell survival and proliferation, cellular
3. Selectin Family

The three closely related members of the selectin family of adhesion molecules, L-, P-, and E-selectin, mediate leukocyte attachment to the endothelium (Bevilacqua et al., 1991). Each selectin molecule contains an amino terminal Ca\(^{2+}\)-dependent (C-type) lectin, an epidermal growth factor (EGF)-like domain, and at least 2 short consensus repeats (SCR). E-selectin contains six SCR, L-selectin contains only two SCR and P-selectin may contain either eight or nine SCR depending on the species. Members of this family of CAMs have been shown to play important roles in leukocyte homing (Mebius and Watson, 1993), rolling and adhesion (Dunon et al., 1996). All selectins family members are found in both transmembrane and soluble forms, with the soluble form arising from proteolytic cleavage of the transmembrane form. Since the soluble form retains binding activity, it is believed that this form may regulate the activity of the transmembrane species.

Selectins, through their C-type lectin domains, bind specific carbohydrates on the surface of other cells. All three selectins bind weakly to the tetrasaccharides sialyl Lewis\(^x\) and sialyl Lewis\(^a\), indicating that these carbohydrate structures are likely components of the physiologically relevant ligands. Ligands of higher binding affinity have been recently found for each of the selectin family members. The molecules GlyCAM-1, CD34, and MAAdCAM-1 all contain sulfated, sialylated fucosylated O-linked carbohydrate side chains specific for binding of L-selectin. Both E-selectin and P-selectin have been shown to bind the P-selectin glycoprotein ligand-1 (PSGL-1). Specific ligands for E-selectin include a 260 kDa glycoprotein on bovine γδ T cells, the 150 kDa E-selectin ligand (ESL-1), and a 250 kDa mouse neutrophil protein; the only known specific ligand for P-selectin is a 160 kDa glycoprotein found on mouse myeloid cells (Tedder et al., 1995; Stoolman, 1989; Lasky, 1992).
Cell surface receptors belonging to the Ig superfamily have wide temporal and spatial expression patterns. They are found in cells of the nervous, circulatory, neural and immune systems and at all stages of development (Brummendorf and Rathjen, 1994). Members may mediate either homotypic or heterotypic cell adhesion, in either homophilic or heterophilic manners, and with varying Ca^{2+}-dependence. Although this family of adhesion molecules is structurally and functionally diverse, all members contain one or more domains that share a conserved structure, called the Ig fold (Williams and Barclay, 1988). Since my thesis deals with the neural cell adhesion molecule NCAM, which is a member of the Ig superfamily, a more detailed discussion of this CAM family will follow.

(i) Common Structural Motifs in the Ig Superfamily

Members of the Ig superfamily are structurally diverse. Despite this, they all possess regions in their extracellular domains predicted to be structurally homologous to Ig domains. Another common motif among members are fibronectin type III-like repeats expressed in tandem with the Ig-like domains. Protein tyrosine kinase and phosphatase domains have also been found in the cytoplasmic tails of an increasing number of Ig superfamily members.

(a) Ig-like Domains

Identification of molecules with Ig-like domains is based on primary sequence analysis. Typically, sequences predicted to have homology to Ig domains are 70-110 residues long and possess two disulfide-bonded cysteine residues separated by 55-77
Fig 2. Schematic diagrams of the topology of Ig-like domains (adapted from Brummendorf and Rathjen, 1994). (A), Ig-like variable domains (V-type). (B), Ig-like constant domains (C1-set). (C), Ig-like constant domains (C2-set). Shaded arrows represent β-strands and are oriented in the amino-to-carboxy direction. Strands on the left of the dotted line form β-sheet I, while those on the right of the line form β-sheet II; together, β-sheets I and II form a tightly packed sandwich structure. Lettering above each β-strand is in accordance with Williams and Barclay (1988) and Jones et al. (1992). V-type domains (A) contain 9 β-strands, with the first β-strand often crossing from β-sheet I to β-sheet II. Strand ordering in this domain set is ABED on one β-sheet and A’GFCC’C” on the other. The I-set domains (D) have essentially the same strand ordering as the V-set, except for the absence of the C” strand. C1-set domains (B) contain two less β-strands than the V-set (less the C’ and C” strands) and does not display the A β-strand cross-over. C2-set domains (C) display a similar topology to the C1-set domains, except for the occurrence of β-strand D shifting over to the other β-sheet. Strand ordering in this domain is ABE on one β-sheet and C’CFG on the other. Often, a disulfide bond is present which links β-strands B and F in each of these Ig-like domains and stabilizes the sandwich structure.
terminal of the first cysteine residue (Vaughn and Bjorkman, 1996; Williams and Barclay, 1988). Some regions, such as domain 3 of CD4 or domain 1 of CD2, which lack the disulfide bond have also been shown to adopt the Ig fold (Wang et al., 1990; Jones et al., 1992).

Ig folds are formed by two β-pleated sheets, each of which contains 3 to 5 anti-parallel β-strands. Amino acid residues within these β-strands alternate between hydrophobic and hydrophilic residues, with the hydrophobic ones pointing inwards and the hydrophilic ones pointing outward. This creates a tightly packed hydrophobic core between β-sheets that is further stabilized by a disulfide bridge spanning the two β-sheets.

Based on sequence and predicted structural similarities to variable and constant Ig domains, Ig folds have been divided into V-, C1- and C2-sets (Williams and Barclay, 1988). The V-set, which includes Ig variable domains, predicted the presence of 9 β-strands; ordering of β-strands in this domain is A, B, E and D on one β-sheet, and (A'), G, F, C, C' and C'' on the other β-sheet. Both the C1-set, which includes Ig constant domains, and the C2-set contain two less β-strands than found in the V-set. In both V- and C- sets, the predicted folding topologies are similar, with the first β-sheet containing β-strands A, B, E and D and the second containing β-strands C, F, and G. Domains of the C2-set are thought to be intermediates of C1- and V-sets, having C1-set folding, yet sharing sequence patterns with V-set domains.

The crystal and solution structures of Ig-like domains have been solved for a number of CAMs involved in T-cell activation (CD2: Jones et al., 1992; Driscoll et al., 1991; CD4: Wang et al., 1990; Ryu et al., 1990; CD8: Leahy et al., 1992), leukocyte adhesion (VCAM-1: Wang et al., 1995; Jones et al., 1995) and neuronal adhesion (NCAM: Thomsen et al., 1996). These domain structures confirm and refine many earlier predictions. In all cases, the Ig-like domains possess the predicted common folding motif of two sandwiched β-sheets. The amino-terminal domain of CD2, CD4 and CD8 all
The second domains of CD2, CD4 and VCAM-1 are all of the C2-set and differ slightly from the originally proposed model. The C1-set β-strand D is switched to the other β-sheet and renamed β-strand C’. Strand ordering for C2-set domains is then revised to β-strands ABE in one β-sheet, and C’CFG in the other. A novel domain set, the I-set, has been recently proposed as another “intermediate” set between V- and C1-sets (Harpaz and Chothia, 1994). This I-set has a β-sheet alignment of β-strands ABED in one sheet and A’GFCC’ on the other. The first domains of both VCAM-1 and NCAM have been classified to this set (Thomsen et al., 1996).

The Ig-like domains are directly responsible for mediating adhesion in a number of Ig superfamily CAMs. Examples of this include the amino terminal domain of ICAM-1 which binds LFA-1 (Staunton et al., 1990), the second domain of L1 which mediates homophilic binding (Zhao et al., 1995), and CD2 which binds to LFA-3 through its Ig-like domain 1 (Recny et al., 1990). Due to limited structural data available for Ig-like domains responsible for ligand binding, potential binding motifs have not been well characterized. An aspartic acid on the C-D loop of VCAM-1 domain 1 was shown to be crucial in binding VLA-4; other important residues were found clustered on the CFG face (Wang et al., 1995). Similarly, residues involved in CD2 binding of LFA-3 are also centered on the C’C’CFG face (Driscoll et al., 1991). It has been suggested that Ig-like domains may bind through these CFG faces in analogy to the dimerization of Ig domains (Springer, 1991). There is still much more work which must be done before generalizations can be made on tertiary binding properties of Ig-like domains.

(b) Fibronectin-type III Repeats

A number of Ig CAMs, primarily those expressed on axonal surfaces, contain extracellular motifs which show homology to fibronectin type III (Fn-III) repeats (Brummendorf and Rathjen, 1994). Originally identified as a 90 amino acid motif which
These repeats are characterized by the presence of conserved tyrosine and tryptophan residues in C- and N-terminal regions, respectively, as well as the absence of conserved cysteine residues. The tenth Fn-III repeat in fibronectin contains an Arg-Gly-Asp (RGD) sequence involved in integrin-mediated cell-to-substratum adhesion (Pierschbacher and Ruoslahti, 1984); the majority of other Fn-III repeats, including most found in Ig-like CAMs, do not contain this RGD sequence and perform no established function (Leahy et al., 1992). The structure of Fn-III domains have been reported for fibronectin (Main et al., 1992; Leahy et al., 1996), neuroglian (Huber et al., 1994) and tenasin (Leahy et al., 1992). These studies indicate that Fn-III domains adopt a conformation similar to that found in the Ig C2-set domains. Furthermore, those Fn-III domains that contain the RGD sequence display this tri-peptide sequence prominently in the loop region between β-strands F and G.

(c) Protein Tyrosine Kinase Catalytic Domains

Receptor protein tyrosine kinases (PTKs) represent a large class of molecules able to participate in signaling pathways through highly conserved catalytic kinase domains (Fantl et al., 1993). Recently, members of two sub-families of Ig-like CAMs have shown kinase activity in their cytoplasmic domains; these proteins are encoded by the trk gene family (human proto-oncogene trk: Martin-Zanca et al., 1989; murine trkB: Klein et al., 1989, Drosophila Dtrk: Pulido et al., 1992; and porcine trkC: Middlemas et al., 1991) and ufo/axl gene family (human nyk/mer: Ling et al., 1996; ufo/axl: Janssen et al., 1991; and murine ark: Rescigno et al., 1991). In both families, members show autophosphorylation on tyrosine residues in response to both Ca^{2+}-independent homophilic binding and ligand binding. The downstream signaling pathway has been partially elucidated for the Nyk/Mer protein using Fms-Nyk/Mer chimeras (Ling and Kung, 1995).
Protein tyrosine phosphatases (PTPase) are believed to help regulate, in conjunction with PTKs, the tyrosine phosphorylation states of many proteins involved in signal transduction (Charbonneau and Tonks, 1992). Domains responsible for phosphatase activity are typically 300 residues long and contain about 40 highly conserved amino acids. Furthermore, a cysteine residue contained within the PTPase consensus sequence (I/V)HCXAGXXR(S/T)G located near the carboxy-terminus of PTPase domains appears to be directly involved in the catalytic process (Guan and Dixon, 1991; Pot and Dixon, 1992).

Two Ig superfamily members, PTPµ (Brady-Kalnay, 1993) and LAR (Streuli et al., 1988) have been shown to contain phosphatase activity. In vitro experiments have shown that these PTPases are able to mediate Ca²⁺-independent homophilic binding through their extracellular domains. Recent loss-of-function experiments performed on the Drosophila homologue of LAR, DLAR, implicate this PTPase in mediating motor axon guidance (Kuegar et al., 1996).

(ii) General Localization and Functions of Members of the Ig Superfamily

Members of the Ig superfamily show wide spatial and temporal expression patterns during development (Brummendorf and Rathjen, 1994). A number of CAMs are expressed in the neural system, such as F3, F11, L1, NrCAM, NgCAM, NCAM, DM-GRASP, MAG, neurofascin, P₀, contactin and TAG-1. With the exception of NCAM, which is also found in muscle cells, these CAMs are expressed predominantly on neurons. All of these neuronal CAMs mediate Ca²⁺-independent binding and most play various roles in neurite outgrowth promotion, guidance, and fasciculation (Brummendorf and Rathjen, 1994). Both P₀ and MAG are involved in myelination.

Many Ig CAMs are also found expressed in the immune system. These include CD2, CD4, CD8, ICAM-1, -2, and -3, CTLA-4, LFA-3, MAdCAM-1, PECAM-1 and
binding mechanisms. The Ig CAMs CD2, LFA-3, ICAM-1, CD4, CTLA-4 and CD8 play important roles in the processes of antigen presentation and T-cell activation. The ICAM-1 and LFA-3 molecules on antigen presenting B-cells (APC) bind to LFA-1 and CD2, their respective receptors on T helper cells and act to maintain cell-cell contact between the two cells. CD4 and CD8 are believed to stabilize the T-cell receptor (TCR)-MHC interaction in class II MHC and class I MHC APC, respectively. Once activated, T-cells upregulate their expression of CTLA-4. The binding of CTLA-4 to its ligand, either CD80 or CD86, is believed to mediate an inhibitory signaling pathway through the recruitment of the SHP-1 PTPase (Scharenberg and Kinet, 1996).

The Ig-like molecules PECAM-1, MAdCAM-1 and VCAM-1 have been shown to play important roles in the process of lymphocyte homing, rolling, adhesion and extravasation. MAdCAM-1 serves as a ligand to L-selectin, which plays a key role in the rolling process; VCAM-1 and its ligand VLA-4 are thought to be involved in all four of the above processes. PECAM, which shows both homophilic and heterophilic binding mechanisms, is also believed to be involved in leukocyte adhesion, but likely plays a key role in the extravasation process.

III. Structure and Function of the Neural Cell Adhesion Molecule NCAM

NCAM was one of the first cell adhesion molecules to be identified and characterized extensively (Thiery et al., 1977; Brackenbury et al., 1977; Cunningham et al., 1987). It is a membrane-bound glycoprotein expressed primarily on brain and muscle cells (Hoffman et al., 1982; Rutishauser et al., 1983), able to mediate homophilic binding in a Ca\(^{2+}\)-independent manner (Rutishauser et al., 1982).
A. Structural Features of NCAM

1. Primary Structure of NCAM

NCAM cDNA, originally cloned in chick (Cunningham et al., 1987), was later reported in human (Dickson et al., 1987), rat (Small et al., 1987), mouse (Barthels et al., 1988), Xenopus (Krieg et al., 1989), bovine and Drosophila. The NCAM protein is encoded by a single gene, containing at least 24 exons (Olsen et al., 1993). Alternative RNA splicing of the primary transcript generates diverse mRNAs (Dickson et al., 1987; Owens et al., 1987; Small et al., 1988; Thompson et al., 1989; Santoni et al., 1989). Transcription, translation and post-translational modifications are regulated in a developmental and cell-specific manner (Murray et al., 1986; Prediger et al., 1988; Dickson et al., 1987; Gennarini et al., 1986).

Three major NCAM isoforms are expressed in vertebrates: the 180 kDa form (NCAM-180), the 140 kDa form (NCAM-140) and the 120 kDa form (NCAM-120). These three isoforms have identical extracellular domains, consisting of 5 Ig-like domains followed by two Fn-III repeats. Each Ig-like domain contains approximately 100 amino acids and includes two conserved cysteine residues which are likely to form an intramolecular disulfide bond (Hemperly et al., 1986).

Differences in the three NCAM isoforms primarily reside in membrane attachment and the cytoplasmic tail. NCAM-120 is anchored to the plasma membrane via a glycosyl-phosphatidylinositol (GPI) moiety (He et al., 1986; Owens et al., 1987), whereas both the 180 kDa and the 140 kDa isoforms contain transmembrane domains. The cytoplasmic tail of NCAM-140 is shorter than that of the NCAM-180 isoform by approximately 200 residues (Murray et al., 1986). Two additional exons, VASE and MSD1, may also be
Fig 3. Schematic representation of the three major isoforms of NCAM (adapted from Brummendorf and Rathjen, 1994). Anchorage of the NCAM isoforms to the membrane is either via a GPI-anchor (large, shaded arrow), or by a transmembrane segment (black cylinder). The five Ig-like domains (Ω) are followed by two fibronectin type III repeats (FN III). A proline-rich region between the Ig-like domains and the fibronectin type-III repeats is predicted to be a hinge region in the NCAM molecule (black dot). In the fourth Ig-like domain of all three major isoforms, a ten amino acid sequence, termed the VASE sequence, may be alternatively spliced into the molecule (boxed V). An additional splice sequence may be inserted into NCAM-120; this 37 amino acid sequence (boxed M), termed MSD1, may be inserted between the two FN-III repeats. Predicted sites of N-linked oligosaccharides are shown by the black arrowheads.
NCAM-120  NCAM-140  NCAM-180
spliced into the extracellular domains of subsets of each of these isoforms (Small et al., 1988; Dickson et al., 1987). These exons will be discussed in the next sub-section.

Selective expression of various isoforms is evident in different cell populations. NCAM-180 is mainly expressed in neurons, astrocytes and oligodendrocytes (Murray et al., 1986); the 140 kDa and 120 kDa isoforms have also been detected in the heart, skin, breast muscle and gizzard (Prediger et al., 1988; Pollerberg et al., 1985). NCAM-120, expressing the MSD1 sequence, is predominantly expressed on myotubes (Thompson et al., 1989).

**Additional Splice Forms**

(a) **VASE Exon**

A 30 bp fragment, designated the VASE sequence, has been found alternatively spliced into the extracellular domains of NCAM (Small et al., 1988; Santoni et al., 1989). When inserted into the fourth Ig-like domain, predicted folding models suggest that this domain adopts a V-set fold, rather than a C2-set fold (Williams and Barclay, 1988). These facts are reflected in the name; it resembles a Variable domain, and it is an Alternatively Spliced Exon. Expression levels of the VASE sequence on NCAM increase during central nervous system development. It is expressed initially on 3% of embryonic neuronal NCAM and this rises to a level of 50% of total neuronal adult NCAM (Walsh et al., 1992; Forster and Frotscher, 1995). Transfection of this exon into 3T3 fibroblasts reduces NCAM ability to promote neurite outgrowth from cerebellar neurons (Doherty et al., 1992). Furthermore, in these cells, this inhibitory effect appears to be dominant over endogenously expressed NCAM; even at low levels of VASE expression, relative to the non-VASE containing NCAM, there is a significant reduction of neurite outgrowth promotion from these cells (Liu et al., 1993; Saffell et al., 1994).

Although the VASE sequence was not shown to significantly effect adhesion (Doherty et al., 1992), it does appear to effect NCAM-NCAM interactions in some manner.
those without the exon (Chen et al., 1994). Although more work is still necessary to understand the mechanisms by which this small exon regulates NCAM function, the current evidence indicates that it plays key roles in regulating NCAM function during development.

(b) MSD1 Exons

NCAM may also undergo alternative splicing to include four small exons collectively encoding the muscle-specific sequence domain (MSD1) (Dickson et al., 1987; Brummendorf and Rathjen, 1994). The insertion of these exons results in the additional expression of a 37 amino acid sequence between the two fibronectin type three repeats. The MSD1 sequence is found in a subpopulation of muscle cells. It is only expressed on GPI-linked, NCAM-120 found on myotubes. This sequence is proline, serine and threonine rich and introduces O-linked glycosylation sites into the NCAM molecule (Walsh et al., 1989). Since this region contains homology to the hinge region of Ig molecules, this sequence may add flexibility to NCAM (Thompson et al., 1989). Due to its tight developmental and tissue regulation, the MSD1 sequence is thought to play an important role in myogenesis; this will be discussed in a later section.

2. Post-translational Modifications of NCAM

(i.) Glycosylation

N-linked Glycosylation

NCAM contains six potential N-glycosylation sites; three sites are located in the fifth Ig-like domain, two sites are found in the fourth domain and the remaining site resides in the third Ig-like domain (Cunningham et al., 1987; Hemperly et al., 1986; Santoni et al., 1988). Two N-linked carbohydrates have specifically been identified which may affect NCAM functions: the HNK-1/L2 epitope and the polysialic acid moiety.
The HNK-1/L2 carbohydrate epitope is expressed on a number of CAMs of the Ig superfamily, including MAG, Po, L1, J1, contactin, and NCAM (Griffith et al., 1992). It has been implicated in mediating or aiding cell adhesion, but its role and location on NCAM have not been defined. The reactive carbohydrate epitope consists of a sulfate-3-glucuronyl moiety (Doherty et al., 1995) and is expressed on 15-20% of transmembrane NCAM-positive neurons (Krog and Bock, 1992). Studies have suggested that this epitope may facilitate NCAM adhesion (Keilhauer et al., 1985; Cole and Schachner, 1987), but due to its low overall expression, it likely only plays minor regulatory roles.

(b) Polysialic Acid Moiety

Homopolymers of α2-8 linked sialic acid, termed polysialic acid (PSA) are added to the fifth domain of NCAM and has been the topic of much study in recent years. The PSA moiety is unique in a number of aspects. During normal carbohydrate processing, a single sialic acid residue is usually added as a terminal, non-reducing sugar in a α2-3 or α2-6 linkage. In NCAM, multiple sialic acid units are added in tandem to form long chains. These chains utilize an unique α2-8 linkage and have been known to contain greater than 55 sialyl residues (Regan, 1991). NCAM is the only known carrier of the PSA moiety in mammals; this unique moiety also occurs in capsular polysaccharides of bacteria, on fish egg glycoproteins and on a sodium channel subunit in Electrophorus electricus (Alcaraz and Goridis, 1991; Reglero et al., 1993). Two eukaryote polysialyltransferase enzymes have recently been cloned (Rutishauser, 1996), and through molecular modeling and mutagenesis work, asparagine residues 430 and 459 have been shown likely acceptors of the PSA moiety (Cunningham et al., 1983; Nelson et al., 1995).

The expression of PSA on NCAM appears to be developmentally regulated. Embryonic forms of NCAM contain 30% (w/w) PSA and this is down-regulated during development to <10% in adult stages. PSA has been shown to modulate NCAM cell-cell
NCAM molecule and also by sterically hindering physical interactions with NCAM. Not only is the PSA moiety believed to interfere with NCAM adhesive properties (Yang et al., 1992), but also it is believed to affect other adhesive mechanisms (Acheson et al., 1991). The change from the high to low PSA form of NCAM during development suggests a role for PSA in attenuating adhesion during periods of high plasticity. For example, in embryonic stages, when PSA-NCAM is expressed, a large amount of cell migration and axonal outgrowth occurs. The subsequent loss of the PSA moiety in adults coincides with stabilization of tissue architecture and more permanent cell-cell contacts. This correlation between PSA expression and plasticity is further displayed in regions of the central nervous system, where the expression of high polysialic forms of NCAM persist throughout adult stages. These areas are usually associated with continued regeneration or growth (Doherty and Walsh, 1996).

**O-linked Glycosylation**

As previously mentioned the MSD1 sequence may be alternatively spliced into NCAM-positive muscle cells. During myogenesis, surface expression of NCAM switches from the 140 kDa transmembrane isoform to the 120 kDa MSD1-containing isoform. With this switch, O-linked carbohydrates are added to serine and threonine residues within the MSD1 sequence (Walsh et al., 1989). The functional significance of O-linked carbohydrates on NCAM is not well understood, but may influence NCAM adhesive properties or regulate NCAM turnover rates (Thompson et al., 1989).

**(ii.) Fatty Acid Acylation**

The 120 kDa NCAM isoform, through the use of exon 15, is attached to the membrane via a glycosyl-phosphatidylinositol (GPI) anchor (Owens et al., 1987). The function of this GPI linkage, although present in many other molecules such as Thy-1,
(iii.) Phosphorylation

The transmembrane isoforms of chick NCAM are phosphorylated on both serine and threonine residues on their cytoplasmic tails (Hoffman et al., 1982; Sorkin et al., 1984). Mouse NCAM, in contrast, is phosphorylated only on serine residues (Gennarini et al., 1984). Phosphorylation on tyrosine residues has not been detected. Although two protein kinases, glycogen synthase kinase 3 and casein kinase I, have been shown to phosphorylate NCAM on threonine residues in vitro (Mackie et al., 1989), the protein kinases able to phosphorylate serine residues in vivo have yet to be discovered. The role of these phosphorylated residues, in terms of NCAM function, remains to be determined.

(iv.) Sulfation

When embryonic chicken brain tissue was cultured in the presence of $^{35}$S, the isotope was incorporated in NCAM carbohydrates (Sorkin et al., 1984). Sulfation occurs on asparagine-linked carbohydrates, but not in the polysialic acid moiety.

3. Predicted Protein Folding Models for NCAM

The five amino-terminal domains of NCAM show sequence similarity to Ig domains and each of these domains is believed to adopt an Ig-fold. Based on sequence
NCAM-180  NCAM-140  NCAM-120

Fig 4. Topology of the three major isoforms of NCAM (adapted from Hall and Rutishauser, 1987). Numeric values represent distances shown and are derived from electron micrographs of NCAM molecules. The shaded circle represents a GPI-anchor.
Ig-like domains (Williams and Barclay, 1988). A later alignment to known tertiary structures of variable and C2-set domains suggested that these NCAM domains conform more closely to the V-set domains (Vaughn and Bjorkman, 1996). The crystal structure for domain 1 has been solved (Thomsen et al., 1996) and shows an I-set fold for this domain. The I-set fold is similar to the V-set, except for β-strand C" being absent. Further structural work is necessary to determine the folding patterns of the remaining domains.

4. Molecular Topology of NCAM

The three major NCAM isoforms have been studied by electron microscopy under low angle rotary shadowing and freeze replication (Hall and Rutishauser, 1987; Becker et al., 1989). In these studies, NCAM appears as a single rod with a flexible hinge region in its middle. As expected, the length from the hinge region to the cytoplasmic tail differs depending on the NCAM isoform. This region in NCAM-180, NCAM-140 and NCAM-120 is approximately 32 nm, 22 nm and 11 nm, respectively. The other length, which corresponds to the area from the hinge region to the N-terminus was about 18 nm. The thickness of the NCAM molecule was shown to be about 4 nm. This is one-half of that found in Fab' fragments and suggests that NCAM Ig-like domains remain unpaired.

B. Ligand Binding Properties of NCAM

The NCAM binding repertoire is relatively large. In addition to being able to mediate homophilic binding in a \textit{trans} interaction with itself, NCAM is also able to mediate \textit{cis} interactions, with L1. It has also been shown to bind to specific proteoglycans in the extracellular matrix and to various cytoskeletal components.
1. NCAM-NCAM Homophilic Interaction

Early studies using neural retinal cells reported that these cells could, in the absence of calcium, quickly re-aggregate after dissociation by mild trypsinization (Brackenbury et al., 1977). Furthermore, antibodies raised against cell surface molecules could inhibit this re-aggregation. A 140 kDa cell surface protein, later named NCAM (Bertolotti et al., 1980), was able to neutralize this antibody inhibition and was suggested to mediate cell-cell adhesion among neuronal cells (Thiery et al., 1977).

When NCAM was inserted into artificial lipid vesicles, these vesicles were able to bind to chick retinal cells, but not to retinal cells pre-coated with NCAM Fab'. (Rutishauser et al., 1982). Further vesicle studies showed a non-linear relationship between NCAM concentration and binding kinetics for NCAM interactions. A 2-fold increase in the NCAM-to-lipid ratio resulted in a 30-fold increase in aggregation (Hoffman and Edelman, 1983) and suggested a threshold concentration requirement for NCAM binding.

Transfection of NCAM cDNA into non-adhesive fibroblastic cell lines, such as mouse L cells and 3T3 cells, conferred NCAM-specific adhesive properties in these cells (Edelman et al., 1987; Pizzey et al., 1989). These transfection studies revealed that NCAM alone is sufficient to mediate cell-cell adhesion and that its adhesive properties are not isoform specific.

Initial attempts to localize the homophilic binding site showed that a 65 kDa proteolytic fragment retained homophilic binding activity (Cunningham et al., 1983). This region corresponded to the 5 amino-terminal Ig-like domains. By epitope mapping of blocking monoclonal antibodies, the NCAM homophilic binding site location was further localized to the amino terminal 40 kDa (Freelinger III and Rutishauser, 1986).

Specific mapping of the chick NCAM binding site was achieved in our laboratory. Employing strategies which included domain deletions, epitope mapping, and synthetic peptides, the NCAM homophilic binding site was localized to the third Ig-like domain (Rao
NCAM homophilic binding in cell binding assays. Synthetic peptides containing this sequence were also able to compete for NCAM binding in membrane aggregation, retinal cell aggregation and Covasphere binding assays. Transfection studies further confirmed the role of this decapeptide sequence in homophilic binding and identified potentially important residues in this interaction (Rao et al., 1993). This binding is believed to occur through an isologous interaction, with the decapeptide sequences on apposing NCAM molecules directly interacting to mediate homophilic binding (Rao et al., 1994).

A recent, independent study has also confirmed the prominence of Ig-like domain three in mediating homophilic binding (Ranheim et al., 1996). This study also implicated other domains in NCAM-NCAM interaction. The NCAM domains 1 and 2 were able to weakly bind to domains 5 and 4, respectively.

2. NCAM-L1 Interaction

L1 is a 200 kDa transmembrane glycoprotein expressed predominantly in the developing peripheral and central nervous systems (Faissner et al., 1985; Hortsch, 1996). L1 is a member of the Ig superfamily (Moos et al., 1988) and is able to mediate neuron-neuron (Keilhauer et al., 1985), neuron-glia (Seilheimer and Schachner, 1987; Grumet and Edelman, 1988) and cell-substratum (Montgomery et al., 1996) adhesion via both homophilic and heterophilic mechanisms (Hortsch, 1996). It is known to be involved in myelination, promotion of neurite outgrowth, fasciculation, and cerebellar cell migration (Lindner et al., 1983; Martini and Schachner, 1986; Lagenaour and Lemmon, 1987).

Early characterization of Ng-CAM, the chicken homologue of L1, indicated a possible interaction between NCAM and the Ng-CAM molecule (Grumet et al., 1984). In these studies, when Ng-CAM was specifically immunopurified from embryonic chicken brain, small amounts of NCAM were co-purified with Ng-CAM. Later NCAM capping experiments revealed L1 to be recruited into NCAM caps and suggested an in vivo cis-
Interaction between NCAM and L1 (Kadmon et al., 1990). The addition of castanospermine to these cells inhibited co-capping and further suggested that NCAM-L1 binding was mediated by protein-carbohydrate interactions. Sequence analysis of the fourth Ig-like domain of NCAM revealed limited homology with carbohydrate recognition domains of C-type lectins (Horstkorte et al., 1993). Studies based on this homology showed that a 19 amino acid peptide that formed part of the C-type lectin consensus sequence was able to inhibit the NCAM-L1 interaction. Furthermore, an oligomannosidic structure on L1 has been implicated in this interaction.

This *cis*-interaction between NCAM and L1 was also shown to influence a *trans*-interaction. The carbohydrate-dependent binding of NCAM to L1 greatly increased L1 homophilic binding affinity (Kadmon et al., 1990). It is believed that the *cis*-binding of NCAM causes L1 to adopt a more favorable conformation for the L1-L1 binding on apposing cell surfaces.

3. NCAM-proteoglycan Interactions

In addition to interacting with other cell surface molecules, NCAM may interact with components of the extracellular matrix (ECM). Two types of proteoglycans abundant in the ECM, heparan sulfate (HSPG) and chondroitin sulfate (CSPG), have shown positive binding to NCAM (Cole et al., 1985; Friedlander et al., 1994).

Early studies, using heparin-agarose columns and $^3$H heparin, showed that NCAM could directly bind heparin (Cole et al., 1985). It was suggested that this interaction could be important in mediating cell-substratum binding in cells expressing NCAM (Cole and Glasser, 1986). The *in vivo* association between NCAM and a heparan sulphate proteoglycan was demonstrated with the co-purification of a 400-520 kDa HSPG with NCAM immunoprecipitated from chicken retinal cells (Cole and Burg, 1989).

Limited proteolysis initially indicated that the heparin binding domain (HBD) was located in the amino 25 kDa of NCAM (Cole et al., 1986). Later studies identified a 17
directly bind heparin (Cole et al., 1989). The predicted heparin binding domain contains a high proportion of basic residues and agrees with heparin binding models and consensus sequences (Cardin and Weintraub, 1989; Ferran et al., 1992). These models predict the HBD to form an amphipathic α-helix containing a highly positively charged face which would bind the negatively charged heparin moiety. Mutagenesis and peptide studies confirmed the importance of the basic residues K$^{135}$, K$^{142}$, K$^{143}$, and R$^{146}$ in mediating heparin binding (Reyes et al., 1990). Molecular modeling of the second Ig-like domain suggests that the HBD is present at the end of β-strand C, in the loop region connecting β-strands C and D.

In addition to heparin sulfate, NCAM has shown positive binding to certain types of chondroitin sulfate. In bead aggregation and $^{125}$I labeled protein assays, NCAM was able to bind to neurocan, a 500 kDa chondroitin sulfate proteoglycan (Grumet et al., 1993; Friedlander et al., 1994). In contrast, binding was not observed with aggrecan, another chondroitin sulfate proteoglycan. This suggests that NCAM recognizes only certain structural motifs on chondroitin sulfate proteoglycans. In vivo studies using embryonic brain extracts also showed NCAM binding of chondroitin sulfate proteoglycans, but only in conjunction with the binding of heparan sulfate proteoglycans (Storms et al., 1996).

4. **NCAM-spectrin Interaction**

Photobleaching studies have indicated a more limited membrane mobility for the 180 kDa isoform of NCAM over that of the 140 kDa isoform (Pollerberg et al., 1986). This was suggested to be due to an interaction between the cytoplasmic tail of NCAM-180 and the cytoskeleton. Immunofluorescence studies revealed in vivo accumulation of actin and spectrin at sites of NCAM-180 clustering. In vitro experiments further indicated a positive binding between brain spectrin and NCAM-180 (Pollerberg et al., 1987). In these experiments, NCAM-180 was able to specifically bind $^{125}$I-labeled spectrin. Also,
C. Functions of NCAM

Cells may express NCAM at different times, in different areas, or with different modifications which allow them to fully utilize the unique functions of NCAM. The differences in NCAM expression have allowed the partial elucidation of the roles that NCAM plays in various processes. NCAM has been implicated in axonal growth and elongation, neuronal fasciculation, cell migration, and myoblast fusion.

1. Axonal Growth and Elongation

Developing neurons send out axons which follow specific pathways to their distant targets by detecting and interacting with cues from their environment (Sperry, 1963; Goodman and Shatz, 1993). In the retinotectal system, retinal ganglion cells send out axons which follow a stereotyped path. These axons grow unidirectionally along the retina to the optic nerve head, join the optic nerve, past the chiasm, along the optic tract and finally reach the optic tectum, where they arborize (Chien and Harris, 1994). Early immunostaining experiments showed NCAM expression along the entire axonal migration route to the tectum and suggested a role for NCAM-NCAM interactions in this process (Silver and Rutishauser, 1984). When NCAM antibodies were introduced into chicken embryo eyes, axons at the optic nerve showed defasciculation and minor mis-routing (Thanos et al., 1984; Silver and Rutishauser, 1984). Similar result were obtained when NCAM antibodies were introduced into the tectum of adult frogs, through agarose spikes. Axons were found to avoid these areas and retinotopic maps became slightly less precise (Fraser et al., 1984). Since axons were still able to find their tectal targets despite the presence of anti-NCAM antibodies, and due to its nongraded expression (Silver and
The ability of NCAM to promote axonal growth has been shown with chicken retinal ganglion cells cultured on either astrocytes or Muller cells (retinal glial cells) (Neugebauer et al., 1988; Drazba and Lemmon, 1990). In both cases, retinal cells extend long neurites which are inhibited by incubation with NCAM antibodies. This NCAM-mediated neurite outgrowth is inducible by either the transmembrane or GPI-linked NCAM isoform (Doherty et al., 1989), but may require presentation of NCAM in a membrane environment. When NCAM is coated on a substratum, in contrast to when expressed on a cell surface, it becomes a less potent promoter of neurite outgrowth (Lagenaur and Lemmon, 1987). NCAM-mediated neurite outgrowth also requires a functional homophilic binding site. L cells transfected with non-functional binding sites are unable to elicit neurite outgrowth responses from chick retinal cells, suggesting that NCAM homophilic interaction may be involved in the regulation of neurite outgrowth (Sandig et al., 1994). A threshold NCAM concentration also appears to be required to trigger neurite outgrowth (Doherty et al., 1990).

2. Neuronal Fasciculation

When neurons send out processes, these axons often have the option of either growing along the substrate, making axon branches, or joining other axons and forming fascicles. Both cultured spinal ganglion cells and retinal ganglion cells have been shown to send out highly fasciculated axons which partially defasciculate upon incubation with anti-NCAM Fab' (Rutishauser et al., 1978; Pollerberg and Sickinger, 1993). In addition, the removal of the PSA moiety, which modulates NCAM adhesion (Acheson et al., 1991), caused an enhancement of fasciculation (Rutishauser et al., 1985). These results indicate that NCAM-NCAM interactions on axons are directly involved in neuronal fasciculation.
3. Cell Migration

(i) Granule Precursor Cell Migration

Inactivation of the NCAM gene in mice resulted in reductions in olfactory bulb size and minor defects in spatial learning (Tomasiewicz et al., 1993; Cremer et al., 1994). During normal olfactory bulb development, granule cell precursors originate in the subventricular zone, migrate tangentially to the olfactory bulb and differentiate into granule and periglomerular cells (Hu et al., 1996). Since NCAM in the olfactory bulb maintains high PSA expression levels even in adult stages (Edelman and Chuong, 1982; Miragall et al., 1990), these NCAM inactivation studies suggested an important role for PSA in granule precursor cell migration. When PSA expression was abolished in developing mice, through injections of endoneuramidase, the resulting phenotypes were similar to those found in the NCAM inactivation studies (Ono et al., 1994). Further studies showed a PSA-NCAM requirement in the substrate along the entire migration route (Hu et al., 1996). It has been suggested that granule precursor cells may migrate over one another and serve as their own PSA-NCAM substrate as they move toward the olfactory bulb (Rousselot et al., 1995).

(ii) Neural Crest Cell Migration

Neural crest cells express many adhesion molecules, including N-cadherin, cytoactin, NCAM and other HNK-1 positive CAMs, which may regulate their ability to migrate during development (Okagawa et al., 1995). Although the role of NCAM is still being investigated, NCAM has been implicated in the induction of neural crest cell migration. NCAM is expressed prior to migration, becomes downregulated during migration and then reappears upon reaching its destination tissue area (Edelman, 1983; Akitaya and Bronner-Fraser, 1992). In addition, when NCAM is abnormally expressed or blocked with antibodies, neural crest cells do not undergo normal migration (Bronner-Fraser et al., 1992).
4. Myoblast Fusion

During myogenesis, mononucleate myoblast cells adhere to one another, fuse to form multinucleate myotubes and then further differentiate into primary and secondary myotubes. NCAM is expressed at all stages of myoblast fusion and may play key roles in the process. Although both NCAM and N-cadherin are likely to share the in vivo responsibility, NCAM expression is sufficient to mediate myoblast condensation in vitro (Knudsen et al., 1989). Prior to monoblast fusion, a switch from the 140 kDa to the 120 kDa isoform is known to occur. The expression of this GPI-linked isoform, as well as that of VCAM-1 and VLA-4, is necessary for myotube formation (Rosen et al., 1992; Knudsen et al., 1989). The MSD1 sequence, expressed in this GPI-linked isoform, is also likely required for the fusion process (Peck and Walsh, 1993; Dickson et al., 1990). At later stages, when primary and secondary myotubes separate, the highly sialylated form of NCAM is expressed at points of separation (Fredette et al., 1993) and implicates PSA-NCAM in the separation process.

D. Signaling Pathway Triggered by NCAM

The ability of NCAM to promote neurite outgrowth implicates NCAM in a signal transduction pathway regulating neurite elongation. To examine this pathway, antibodies were used to mimic NCAM binding on rat pheochromocytoma (PC12) cells and their effects on common second messengers were followed (Schuch et al., 1989). Antibody binding resulted in reduced intracellular levels of IP2, IP3, pH, and increased levels of Ca\(^{2+}\). Since pertussis toxin inhibited this influx of Ca\(^{2+}\), a G protein was also proposed upstream of the Ca\(^{2+}\) channel. Similar results were obtained when PC12 cells were cultured on NCAM-transfected 3T3 cells (Doherty et al., 1991); NCAM-NCAM interaction lead to a pertussis toxin-sensitive influx of Ca\(^{2+}\) through N- and L-type Ca\(^{2+}\) channels.

Tyrosine kinases have also been implicated in this signaling pathway. When neurons from \textit{fyn}\(^{-}\) mice were cultured on NCAM-transfected L-cells, neurite outgrowth
Fig 5. Schematic diagram representing proposed major components in the NCAM-mediated neurite outgrowth response (adapted from Beggs et al., 1994; Williams et al., 1994; Doherty et al., 1991). In one of the predicted pathways, NCAM activation leads to a direct interaction with the FGF receptor. The activated FGF receptor induces PLCγ to convert phospholipids (PL) to diacylglycerol (DG), which is further converted to arachidonic acid (AA) by DG lipase. The creation of AA induces the opening of N and L-type Ca²⁺ channels, and causes an influx of Ca²⁺ into the cell. This eventually leads to a NCAM-induced neurite outgrowth response from the cell. Other possible participants in the NCAM-induced neurite outgrowth signaling pathway are the non-receptor tyrosine kinase p59⁵⁵n and a pertussis toxin-sensitive G-protein.
outgrowth (Beggs et al., 1994). Furthermore, since p59fyn minus neurons show no outgrowth inhibition on L1 transfected L-cells, p59fyn appears to be unique to the NCAM pathway (Ignelzi Jr et al., 1994). A number of tyrosine kinase inhibitors have been tested for their effects on NCAM-mediated neurite outgrowth (Williams et al., 1994). An erbstatin analogue and two tyrphostins specifically inhibited NCAM stimulated outgrowth. It is believed that erbstatin inhibits a kinase upstream of the calcium channel, whereas the two tyrphostins inhibit a kinase downstream of the channel. Recent studies (Williams et al., 1994; Williams et al., 1994) suggest the FGF receptor as the erbstatin sensitive kinase and that NCAM promotes neurite outgrowth through the FGF-PLCy cascade.

IV. Description of the Project

Studies in our laboratory have resulted in the identification of a decapeptide sequence (243-KYSFNYDGSE-252) responsible for mediating homophilic binding in chick NCAM. This decapeptide sequence, located in Ig-like domain 3, was shown to be necessary and sufficient to mediate binding in adhesion studies. Results also indicated a direct and isologous interaction between decapeptide sequences in apposing NCAM molecules.

NCAM has been found in a number of species, including human, rat, mouse, bovine, *Xenopus*, and chick. The interspecies sequence identity, in comparison to the chick protein sequence, is relatively high in the first five Ig-like domains (Table 1). The percentage sequence identity ranges from 87-89% for bovine, rat and human to 74% for *Xenopus*. In contrast to these high values for overall sequence identity, the levels of identity in the region of the chick homophilic binding site is relatively low (Fig. 6). Values, in this case, range from a high of 70% identity for *Xenopus* to a low of 50% identity for both rat and human. This relatively low sequence identity in the region of the
Table 1. Interspecies sequence identity to chick NCAM. The Ig-like domains of human, mouse, bovine and *Xenopus* were manually aligned to the chick sequence, and residues which were identical were scored. Values were tabulated for all five Ig-like domains, and also in the region of the chick homophilic binding site.

<table>
<thead>
<tr>
<th></th>
<th>Ig1-5 (520 a.a's)</th>
<th>Binding Site (10 a.a's)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># residues identical</td>
<td>%</td>
</tr>
<tr>
<td>chick to human</td>
<td>461</td>
<td>89</td>
</tr>
<tr>
<td>chick to mouse/rat</td>
<td>459</td>
<td>89</td>
</tr>
<tr>
<td>chick to bovine</td>
<td>453</td>
<td>87</td>
</tr>
<tr>
<td>chick to <em>Xenopus</em></td>
<td>383</td>
<td>74</td>
</tr>
</tbody>
</table>
Figure 6. Partial sequence comparison between the region of chick NCAM containing the homophilic binding site with that of the analogous regions in other species. Numbering of amino acids contained within the third NCAM Ig-like domain of chick NCAM is according to Cunningham et al. (1987). The chick homophilic binding site sequence, as well as its proposed analogue in other species of NCAM, is contained within the boxed area. Residues that are shaded share sequence identity with that of the chick sequence at that position.
conserved in other vertebrate species.

Just before the start of my project, two studies were published on rodent NCAM. In the study of Zhou et al. (1993), using chimeric NCAM/CEA and full-length NCAM constructs to transform CHO cells, they demonstrate that cell-cell binding requires all five extracellular Ig-like domains of rat NCAM. Furthermore, the first 3 Ig-like domains are unable to mediate cell aggregation. In contrast to this study, the results of Frei et al. (1992) suggest that Ig1, Ig2 and Fn-III repeats 1 and 2 of mouse NCAM are efficient in mediating neuronal cell body attachment to protein-coated substrates. Both studies suggest that Ig3 alone is insufficient for NCAM homophilic binding.

To investigate differences between these results and ours found in chick NCAM, a series of fusion proteins were created for adhesion studies. Initially, rat Ig-like domain 3 (rIg3) was expressed in bacteria and tested for binding activity. Consistent with previous results in rodent NCAM, little binding activity was observed in this construct. To explore these discrepancies further, chimeric constructs were made to determine whether the putative rat and human homophilic binding sequences can function in the chick Ig3 environment. The first two chimeric proteins were based on a (His)$_x$-tagged chick Ig3 fusion protein (Rao et al., 1994) which previously exhibited positive binding ability. Using site directed mutagenesis, the chick binding site sequence was mutated to the putative rat or human binding sequences, creating a c(r)Ig3 and c(h)Ig3 fusion protein, respectively. These fusion proteins were tested in Covasphere-to-substratum binding assays. Interestingly, c(h)Ig3 displayed positive homophilic binding activity, while c(r)Ig3 showed only weak homophilic binding.

The replacement of the functional chick binding site in these constructs should, in theory, only retain binding activity if replaced with another functional binding site. Results from these experiments show the retention of binding activity with the chick/human chimera and suggests that the analogous decapeptide sequence in human NCAM is
sufficient to mediate homophilic binding activity. The chick/rat chimera, in contrast, displayed low binding activity. Since the GST-rIg3 fusion protein also exhibited only weak binding, an attempt was made to rescue the rat binding activity by constructing a rat/chick chimera, r(c)Ig3. This protein mutated the putative rat binding site to that of chick NCAM. In rat PC12 cell attachment assays, the r(c)Ig3 chimeric protein was able to mediate cell-substrate binding. Overall, these results suggest that human NCAM may use the same binding mechanism as seen in chick NCAM, while rat NCAM may not. Results using r(c)Ig3 further indicate that this lack of rat binding may be due to lower binding affinities exhibited by the putative rat binding sequence.

Further studies were also performed on the chick homophilic binding site to study the importance of various amino acid residues in the binding process. Of the ten amino acids in the region of the chick homophilic binding site, only four are conserved among all species of NCAM. In the same chick Ig3 fusion protein used for the chimera constructs, the decapeptide sequence was mutated to the sequence 243-KAAFAADASA-252. This mutant Ig3 fusion protein retained only those amino acid residues which were conserved among all five NCAM species. In Covasphere binding studies, this mutant Ig3 fusion protein retained binding activity, indicating that these remaining residues are sufficient to mediate homophilic binding.
Chapter 2: Materials and Methods
The pGEX-1λT, pGEX-2T plasmids and glutathione-Sepharose 4B were purchased from Pharmacia Biotech Inc. (Toronto, ON). The pQE-8 plasmid and nickel-nitrilotriacetic acid resin were purchased from Qiagen (Chatsworth, CA). Covaspheres were obtained from Duke Scientific Corp. (Palo Alto, CA). Peptides were synthesized by Procyon Biopharma Inc. (London, ON). The bicinechonic acid (BCA) and Coomassie Plus protein determination kits were purchased from Pierce Chemical Co. (Rockford, IL). Poly-L-lysine and BSA were purchased from Sigma (St. Louis, MO). Penicillin/streptomycin (10,000 unit penicillin, 10,000 µg streptomycin per ml stock solution) was obtained from Gibco (Toronto, ON). Fetal bovine serum was obtained from Hyclone (Logan, UT). Other reagents used were of highest chemical grade.

Cell Culture Conditions

PC12 cells, obtained from Dr. J. Reed (University of Toronto, ON), were cultured in media consisting of RPMI 1640 with 5% fetal bovine serum, 5% donor horse serum, 2 mM L-glutamate, and 1% penicillin/streptomycin. Media was changed every 3-4 days and cells were passaged when subconfluent. For passage, cells were washed once with new media, resuspended into 8 ml of fresh medium and then seeded at about 1/20 in a new, pre-coated 100 x 15-mm plate. Plates were coated with 6-8 ml of a 0.02% (w/v) poly-L-lysine solution for 1 hr at room temperature, aspirated and filled with 8 ml of media before seeding with resuspended cells. Cells were maintained at 37°C and under 5% CO₂.

Construction of Rat NCAM Domain 3 GST Fusion Protein

The vector pSRV-CAT, which contains the entire coding region for rat NCAM-140 (donated by Dr. A.Acheson), was used as a template for PCR amplification of the third NCAM Ig-like domain (Fig. 7). The oligonucleotides, 5'-GCGCGATCCACT
<table>
<thead>
<tr>
<th>Domain</th>
<th>Forward 5'</th>
<th>Reverse 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>chick domain 1</td>
<td>5' - GCGCGGATCCCCTACAAGTGATATTGT</td>
<td>3'</td>
</tr>
<tr>
<td>chick domain 1 + 2</td>
<td>5' - GCGCGGATCCCTACAAGTGATATTGT</td>
<td>3'</td>
</tr>
<tr>
<td>chick/human chimeric domain 3</td>
<td>5' - AATTCACACAGCATGCAAAAGAGG</td>
<td>3'</td>
</tr>
<tr>
<td>chick/rat chimeric domain 3</td>
<td>5' - AATTCACACAGCATGCAAAAGAGG</td>
<td>3'</td>
</tr>
<tr>
<td>rat domain 1 + 2</td>
<td>5' - GCGCGGATCCCCTCAAGTGATATTGT</td>
<td>3'</td>
</tr>
<tr>
<td>rat domain 3</td>
<td>5' - CCGGAATTCACTGCGAGGCCAGACAG</td>
<td>3'</td>
</tr>
<tr>
<td>rat/chick chimeric domain 3</td>
<td>5' - CCACGTTTGGTGTTGGCG</td>
<td>3'</td>
</tr>
</tbody>
</table>

Table 2. Primers used in the construction of various fusion proteins.
Rat NCAM cDNA

\[ \text{PCR} \]

\[
\begin{array}{c}
\text{Bam HI} \\
\text{Eco RI}
\end{array}
\]

\[
\begin{array}{c}
\text{Bam HI} \\
\text{Eco RI}
\end{array}
\]

\begin{center}
\begin{tikzpicture}
  \node (pGEX) [circle, draw, inner sep=1pt] at (0,0) {
    \begin{minipage}{3cm}
      \text{GST} \quad \text{Amp} \\
      \text{pGEX-1λT}
    \end{minipage}
  };

  \draw[->] (pGEX) -- ++(0,1) node[above] {\text{T293}};
  \draw[->] (pGEX) -- ++(0,-1) node[below] {\text{F286}};

  \draw[->] (pGEX) ++(1,0) -- ++(0,1) node[above] {\text{GST}};
  \draw[->] (pGEX) ++(-1,0) -- ++(0,-1) node[below] {\text{Rat NCAM Ig3}};
\end{tikzpicture}
\end{center}

Fig 7. Schematic diagram of the construction of a recombinant rat GST NCAM domain 3 protein. A PCR fragment containing the coding region of Ig3 was generated and inserted into the pGEX-1λT expression vector. The arrows indicate the positions where the oligonucleotides were annealed with the plasmid DNA in PCR amplification.
were used to amplify the cDNA fragment corresponding to the region between amino acids Thr-213 and Lys-308 of the translated protein. Temperatures and cycling times used in the PCR reaction for primer annealing, extension and melting were 52°C for 2 minutes, 72°C for 3 minutes and 94°C for 1 minute, respectively. BamHI and EcoRI cleavage sites, placed at the ends of the oligonucleotides, were used to cleave the PCR product and insert it into the BamHI/EcoRI site of the GST plasmid pGEX-1λT. Clones that incorporated the 300 bp insert and expressed a protein of the correct molecular weight, as judged by SDS-PAGE, were verified by DNA sequencing. The plasmid was then transfected into the Escherichia coli strain JM101 for large scale protein expression.

**Expression and Purification of Rat NCAM Domain 3 GST Fusion Protein**

A 100 ml overnight culture of transformed E. coli cells, which contained the GST-rat Ig3 plasmid, was used to inoculate 1 litre of LB containing 100 μg/ml ampicillin. The large culture was grown at 30°C for 1.5 hours before inducing protein synthesis with the addition of β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM. After 4.5 hours of induction, the cells were collected by centrifugation at 4,000 x g for 20 minutes at 4°C. The resulting pellet was resuspended into 20 ml of sonication buffer (PBS containing 10% glycerol, 1 mM EDTA, 0.1 mM PMSF and 1 mM DTT) and sonicated 5 times (setting = 5), in 20 second bursts. After a further centrifugation at 7,000 x g for 10 minutes at 4°C, Triton X-100 was added to the supernatant at a final concentration of 1%. The supernatant, containing the soluble rat Ig3 protein, was then mixed with pre-equilibrated glutathione-Sepharose 4B resin (2 ml of 50% slurry) and allowed to adsorb to the resin under slow rotation for 1 hour, at 4°C. The supernatant/resin mixture was packed into a column (1.6 cm x 8.5 cm) and then washed with 15 ml of 1 M NaCl in PBS to remove nonspecifically bound materials. Elution of the soluble GST-rat Ig3 was performed with 50 mM glutathione in 50 mM Hepes (pH 7.5). Column fractions of 2 ml were collected and those
dialyzed overnight against PBS. Samples were stored frozen at -70°C. Before use, all proteins were subjected to brief (5 pulses at setting = 3) sonication to disperse aggregates.

**Construction of Chimeric His-tagged Ig3 Proteins**

The chick NCAM His-tagged Ig3 recombinant protein (Rao et al., 1994) was used as a template for creating chimeric proteins to further study the putative binding site in rat and human NCAM (Fig. 8). The two mutagenic primers, 5'-GAGGATAACGAAGAGAAGTACATCTTCAGCGACGATAGTTCCCAGCTGATCATCAAGAAG-3' and 5'-GAGGATAACGAAGAGAAGCATACTTCAGTGACGACAGTTCGGAGCTGATCATCAAGAAG-3', contain the 30 bp nucleotide sequence encoding the proposed human and rat binding sites, respectively. The human and rat sequences are surrounded on either side by the 15 bp of chick sequence that normally flanks the chick homophilic binding site sequence. These primers, along with the selection primer 5'-AATTCACACAGCATGCATTAAAGAGG-3', were used in the Transformer™ site-directed mutagenesis kit from Clonetech (Palo Alto, CA) to mutate the chick homophilic binding site to that of the proposed human and rat sites.

In brief, the chick Ig3 plasmid, in the presence of both mutagenic and selection primers, was denatured at 100°C for 5 minutes and then immediately cooled to 4°C to allow for primer annealing. T7 polymerase and T4 ligase were added to the mixture and, upon incubation at 37°C for 2 hours, the primers were elongated along the parental strand and the nicked ends were ligated. The newly synthesized mutant strands incorporated both the selection primer, which changed a parental unique cut site (EcoRI) to another unique cut site (SphI), and the mutagenic primer, which changed the chick binding site sequence to that of either the rat or human analogues. The mixture was incubated with EcoRI for 2 hours, which linearized the parental strands, and then transformed into the repair-deficient
Fig 8. Schematic diagram of the construction of chimeric chick NCAM Ig3 proteins. Mutagenesis was used to replace the chick homophilic binding site in a chick domain 3 protein construct to that of the proposed human and rat homophilic binding site analogues.
round of EcoRI digestion and transformation into the E. coli strain JM101. Clones that incorporated the selection primer, as shown by the release of a 300 bp SphI/HindIII digestion fragment, were sequenced for mutagenic primer incorporation. DNA isolated from clones found expressing the chimeric proteins was then used to transform the E.coli strain M15, which harbored the plasmid pREP4 encoding the lacZ repressor, and cultures from these transformants were used for large scale expression.

Expression and Refolding of Chimeric Ig3 His-tagged Fusion Proteins

For the large scale expression of chimeric Ig3 proteins, 1 litre of LB medium containing 100 μg/ml ampicillin and 50 μg/ml kanamycin was inoculated with 10 ml of an overnight culture of M15 cells transformed with the chimeric plasmid. Cells were grown at 37°C with vigorous shaking until the A₆₀₀ reached 0.7-0.9 (~2.5 hours). Synthesis of fusion protein was induced by the addition of IPTG to a final concentration of 2 mM and cells were grown at 37°C for another 4 hours. Cells were then collected by centrifugation at 4,000 x g for 10 min at 4°C and resuspended in Buffer A (6 M guanidine HCl, 0.1 M NaH₂PO₄, 0.01 M Tris/HCl, pH 8.0) at 5 ml/g wet weight. The mixture was rotated for 1 hour at room temperature before centrifugation at 10,000 x g for 15 min at 4°C to collect the solubilized His-tagged fusion proteins, contained within the supernatant.

Nickel-nitrilotriacetic acid resin (8 ml of a 50% slurry) was equilibrated in Buffer A and then added to the collected supernatant. The mixture was rotated for 1 hour at room temperature and then packed into a column (1.6 cm x 8.5 cm). The column was first washed with 10 column volumes (~40 ml) of Buffer A and then with Buffer B (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris/HCl, pH 8.0) until the A₂₈₀ was < 0.01 (~25 ml). The column was further washed with Buffer C (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris/HCl, pH 6.3) until the A₂₈₀ was < 0.01 (~60 ml). The chimeric protein was eluted from the column with 20 ml of Buffer D (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris/HCl, pH 5.9),
Column fractions of 2 ml were collected. Those fractions which contained eluted chimeric proteins, as determined by SDS-PAGE, were pooled and adjusted to ~2 mg/ml.

To renature the chimeric protein, β-mercaptoethanol was added to the protein solution to a final concentration of 0.1 M. The mixture was incubated at 37°C for 1.5 hours for complete reduction, and then dialyzed against 4 changes of 1 litre of 8 M urea, 150 mM NaCl, 50 mM Tris/HCl, pH 8.0 at 4°C, over the next 4 days. The denatured protein solution was slowly diluted 1 to 20 in cold PBS and then concentrated to ~6 ml in an Amicon stirred ultrafiltration unit (operated under a pressure of 38 psi). The recombinant protein was dialyzed overnight against 2 litres of PBS, aliquoted and frozen at -70°C under dry ice. Samples were stored frozen and subjected to brief (5 pulses at setting = 3) sonication before use.

Construction of a Chimeric Rat Ig3-GST Fusion Protein

Using the rat Ig3-GST fusion protein as a template, the proposed rat binding site was mutated to the chick sequence by the four primer method (Higuchi, 1990). The primers, 5'-CCACGTTTGGTGTTGGCG-3' and 5'-GTCATTTTTGTCCACATT CCTGATGGTCAGCTCGGACCCTCGTAGTTAAAACGCTGTATT-3', amplified the first half of the third domain and also introduced the chick binding site sequence into the rat half-domain. Similarly, the primers, 5'-GAGGAAGATGACGAGAAATACAGTTTTAACTACGATGGGTCCGAGCTGACCATCAGGAAT-3' and 5'-CCGAAACGCGAGGCAG-3', were used to amplify the second half of the Ig3 domain, while at the same time introducing the same mutation into the rat sequence. Using these half-domains as templates in the PCR reaction resulted in the production of the complete rat Ig3 domain, which now contained the chick binding site sequence. The PCR product was inserted into the GST vector pGEX-2T, expressed and purified in accordance with protocols outlined above in the rat Ig3 construction.
**Construction of a Mutant Chick Ig3 His-tagged Fusion Protein**

To investigate the involvement of specific amino acid residues in chick to chick homophilic binding, a further chick Ig3 mutant was constructed. Using the same protocols outlined in the construction of the chimeric fusion proteins, the chick binding site was mutated to the sequence 243-KAAFAADASA-252. This construct retained the four residues conserved among the vertebrate species and the remaining six residues were mutated to alanine. The primer used for mutagenesis was 5'-GAGGATAACGAAGAGAGAGGCTGCTTTCGCTGCTGATGCTTCCGCTCTGATCATCAAGAAG-3'. The expression, purification and refolding of this fusion proteins was performed in accordance to procedures outlined above.

**DNA Sequencing**

After mutagenesis or PCR amplification, all constructs were sequenced to ascertain correct mutations or against PCR errors. DNA sequencing was performed using the T7 Sequencing™ kit obtained from Pharmacia Biotech.

**Electrophoresis and Immunoblotting**

Column fractions and samples from various protein purification steps were routinely separated on 7.5% or 14% (w/v) SDS-polyacrylamide gels to determine protein purity and size. Slabs were prepared according to Laemmli (Laemmli, 1970) and were either stained with Coomassie Brilliant Blue or electroblotted onto nitrocellulose filters. Protein blots were incubated with rabbit anti-NCAM antiserum, alkaline phosphatase-conjugated secondary antibody, and visualized with 5-bromo-4-chloro-3-phosphate(BCIP)/nitro blue tetrazolium (NBT) as substrates for colour development. Prestained molecular weight markers included in all gels and blots were obtained from Gibco (Toronto, ON) and contained lysozyme ($M_r \approx 15.4$ kDa), $\beta$-lactoglobulin ($M_r \approx 18.2$ kDa), carbonic anhydrase
Phosphorylase B \( (M_r = 110.8 \text{ kDa}) \) and myosin \( (M_r = 216.8 \text{ kDa}) \).

**Protein Determination**

The bicinchoninic acid (BCA) and/or Coomassie Plus assay was used for all protein determinations, with crystalline bovine serum albumin (BSA) serving as the standard.

**Conjugation of Fusion Proteins to Covaspheres**

To conjugate green MX Covaspheres to either GST or His-tagged fusion proteins, 50 µl of Covaspheres was sonicated for 20 seconds and then added to 10 µg of the fusion protein in 50 µl PBS. The mixture was incubated under slow rotation at room temperature for 75 minutes and with brief vortexing every 15 minutes. The protein-conjugated Covaspheres were then pelleted by centrifugation at 4°C for 10 minutes and the protein solution was aspirated. All remaining active sites on the Covaspheres were blocked by incubation with 2% BSA (w/v) for 30 minutes at room temperature. The Covaspheres were then pelleted by centrifugation at 4°C for 10 minutes, washed with 1 ml PBS, pelleted once again, and finally resuspended into 100 µl PBS. The protein-conjugated Covaspheres were stored in a dark container and subjected to mild (10 seconds in a chilled water sonicator) sonication before use.

**Covasphere-to-Substratum Adhesion Assay**

The various proteins were used to coat spots (~2.5 mm in diameter) on 35-mm plastic Petri dishes (Fig. 9). Protein solutions were used at a concentration of ~70 µM in 4 µl PBS and were allowed to adsorb to the dishes at room temperature for 1 hour before removal by aspiration. Each protein-coated spot was washed twice with of 4 µl
Covaspheres were conjugated to the Ig3 fusion protein of interest (D3) at room temperature, for 1 hour before blocking with a 1 % BSA (w/w) solution for a further 1 hour. The Covaspheres were then allowed to bind to a prepared substrate for 1 hour, washed and then counted for binding activity. Prepared substrates typically consisted of protein solutions adsorbed to plastic Petri dishes. Experiments were routinely performed in triplicates and all data was recorded to video tape and the number of Covaspheres bound per field on substrate-coated protein was measured using the NIH program.
The BSA was aspirated and two further washes of 7 µl PBS removed any excess BSA.

To assay for binding between protein-conjugated Covasphere and various substrate-coated proteins, 5 µl of the Covaspheres was added to 100 µl PBS and dispersed for 10 seconds in a chilled water sonicator. The Covaspheres were then added, at 15 µl per spot, to each protein-coated area and incubated at room temperature for 1 hour, with slow shaking. The dishes were gently washed 5 times with PBS, mounted with coverslips, and then assessed for Covasphere binding using epifluorescence microscopy. On average, 10 microscope fields per spot, at 4 seconds per field, were recorded to videotape using a 25X objective. Recordings of fields were played back using the Video Monitor and images were captured to disk as still images. Images were rendered black and white using the NIH Image program, inverted, and the number of Covaspheres per field was quantitated. All experiments were repeated three times, normalized against binding of Covaspheres to the positive control and expressed as relative binding. In competition experiments, protein-conjugated Covaspheres were incubated with the competing protein solution for 15 minutes prior to the binding assay.

**Determination of Plating Efficiency for Covasphere Assay**

To determine the efficiency of protein adsorption in the Covasphere assay (Table 3), the various proteins were adsorbed to the Petri dishes under standard assay conditions. Protein concentrations were the same as used in the Covasphere assay, but volumes were scaled up by a factor of ten (40 µl), to allow for better protein quantitation. Following adsorption, protein spots were washed twice with 40 µl PBS. The protein was removed from the Petri dish through incubation with 50 µl of a 1 % SDS solution for 15 minutes. Quantitation of removed protein was performed using the BCA protein determination kit.
Table 3. Plating efficiency in the Covasphere experiment. The efficiency of proteins being adsorbed to the surface of Petri dishes was examined. Equal molar quantities of each protein was added to a Petri dish in a ~ 10 mm spot (40 µl) and allowed to adsorb for 1 hour. The spot was washed twice with 40 µl PBS and the adsorbed protein was removed with a solution of 1% SDS in PBS. Adsorbed protein amount was quantitated using the BCA protein determination kit. All values represent the mean ± S.D. of 3 experiments.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amount Plated</th>
<th>Amount Bound</th>
<th>Binding Efficiency</th>
<th>µmoles Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>cIg1</td>
<td>8.6µg /715µmoles</td>
<td>0.827 ± 0.09µg</td>
<td>9.61 ± 1.1 %</td>
<td>69.9</td>
</tr>
<tr>
<td>cIg3</td>
<td>10µg /715µmol</td>
<td>0.780 ± 0.28µg</td>
<td>7.80 ± 2.8 %</td>
<td>55.7</td>
</tr>
<tr>
<td>c(h)Ig3</td>
<td>10µg /715µmoles</td>
<td>0.976 ± 0.41µg</td>
<td>9.77 ± 4.1 %</td>
<td>69.7</td>
</tr>
<tr>
<td>c(r)Ig3</td>
<td>10µg /715µmoles</td>
<td>0.857 ± 0.14µg</td>
<td>8.57 ± 1.4 %</td>
<td>61.2</td>
</tr>
<tr>
<td>c(4)Ig3</td>
<td>10µg /715µmoles</td>
<td>0.773 ± 0.08µg</td>
<td>7.73 ± 0.8 %</td>
<td>55.2</td>
</tr>
<tr>
<td>rIg3</td>
<td>28.6µg /715µmoles</td>
<td>0.773 ± 0.25µg</td>
<td>2.37 ± 0.9 %</td>
<td>19.3</td>
</tr>
<tr>
<td>r(c)Ig3</td>
<td>28.6µg /715µmoles</td>
<td>0.636 ± 0.24µg</td>
<td>2.22 ± 0.8 %</td>
<td>15.9</td>
</tr>
<tr>
<td>GST</td>
<td>18.8µg /715µmoles</td>
<td>0.717 ± 0.09µg</td>
<td>3.85 ± 0.4 %</td>
<td>27.6</td>
</tr>
</tbody>
</table>
To prepare 35-mm Petri dishes for protein coating, 1 ml of a nitrocellulose solution (5 cm² dissolved in 1 ml of methanol) was added to the dish and allowed to evaporate overnight. Protein solutions were then adsorbed to the nitrocellulose-coated dish for 1 minute, washed twice with PBS, and blocked for 1 minute with 1 % BSA in RPMI 1640 medium. Protein adsorption and blocking was performed just prior to the assay.

To prepare PC12 cells, cells were grown to subconfluence, washed with 8 ml RPMI 1640 medium, and then resuspended in medium to ~10⁶ cells per ml. A volume of 20 μl of cells was added to each of the protein coated spots and cells were allowed to settle for 1 hour at 37°C. Two gentle washes using 8 ml RPMI each time removed unbound cells. The remaining bound cells were counted under an inverted microscope. All experiments were repeated three times.
Chapter 3: Results
The Ig3 domain of chick NCAM has been shown to contain homophilic binding activity (Rao et al., 1992). To investigate whether rat NCAM uses the same domain to mediate homophilic interaction, a GST fusion protein containing the rat Ig3 domain (GST-rIg3) was expressed in bacteria (Fig. 10). The pGEX-1αT plasmid was chosen as the expression vector because it yielded GST-rIg3 in a soluble form. Conditions for growth, induction and expression of the recombinant protein were selected to enhance the expression of the soluble protein. GST-rIg3 was purified on a glutathione-Sepharose 4B column and gave a yield after column purification of ~3 mg/litre culture.

To check for purity, a sample of GST-rIg3 was analyzed by SDS-polyacrylamide gel electrophoresis and visualized by Coomassie Blue staining. Under reducing conditions, the protein migrated as a major band of apparent molecular weight of ~40 kDa, consistent with the predicted molecular weight for the fusion protein (Fig. 10c). Increased mobility of the protein band was observed under non-reducing conditions (data not shown), suggesting the presence of an intramolecular disulfide bond in the fusion protein. Molecular markers (M) are shown in the left lane.

**GST-rIg3 did not Exhibit Homophilic Binding Activity**

The soluble GST-rIg3 fusion protein was tested for homophilic binding activity in the Covasphere to substratum binding assay. This assay was selected due to positive binding observed for the cIg3 fusion protein in previous studies. The binding of rIg3-conjugated Covaspheres to substrate-coated rIg3 fusion protein was essentially at the background level, comparable to the result obtained with substrate-coated GST (Fig. 11). However, rIg3-conjugated Covaspheres exhibited a significant level of binding with the (His)$_6$-tagged chick Ig3 fusion protein (His-cIg3). This result was unexpected and a close examination of the assay revealed that although similar amounts of GST-rIg3 and His-cIg3 were adsorbed on the Petri dish, they were present in different molar concentration due to
Fig. 10. Schematic diagram of the rat NCAM Ig3 fusion protein. (A) The rat Ig3 construct was created based on the rat NCAM cDNA using standard PCR methods. Numbering of amino acids in the rat NCAM protein sequence is in accordance with Small et al. (1987). Dotted regions on the protein schematic represent Ig-like domains in the rat NCAM protein and fibronectin type-III repeats are located in the white area; the cross-hatched region shows the location of the transmembrane segment and the striped area represents the cytoplasmic domain of the protein. The cDNA encoding rIg3 was fused to the 3' end of the GST expression vector and was expressed in the soluble form. The GST tag is shown as a shaded box at the amino-terminus of the fusion protein. (B) Partial sequence of rIg3. The bolded and boxed area identifies the region of rIg3 which is analogous to the chick homophilic binding site. (C) Gel profile of purified GST-rIg3 fusion protein. Protein samples were separated on a 10% SDS-polyacrylamide gel under reducing conditions and then stained with Coomassie Brilliant Blue. Molecular weight markers (M) are shown on the left lane.
Attempts to adsorb more GST-rIg3 on the plastic were ineffective, indicating that the saturated level was reached. These results may suggest that rIg3 may have a weak binding site which requires a higher threshold concentration of substrate-coated protein for effective binding. As a negative control, these Covaspheres were seeded on substrate-coated (His)$_6$-tagged chick Ig1 fusion protein (His-cIg1). Although the concentration of His-cIg1 bound on the plastic was similar to that of His-cIg3 (Table 1), only a background level of binding was observed (Fig. 11).

The specificity of the interaction between GST-rIg3 and His-cIg3 was demonstrated by competition experiments using soluble rat and chick Ig3 fusion proteins as competitors. The rIg3-Covaspheres were preincubated with either soluble rIg3 or cIg3 fusion proteins before seeding on substrate-coated cIg3. Inhibition of Covasphere binding was dose dependent in both cases, with 50% inhibition at ~0.3 μM soluble protein competitor (Fig. 12). Binding was abolished at ~10 μM of either competitor.

**Expression of Chimeric Domain 3 Fusion Proteins Containing the Putative Human and Rat Homophilic Binding Site**

Although rIg3 does not appear to contain homophilic binding activity, failing to mediate Covasphere binding to substrate-coated protein, it does exhibit affinity for the cIg3 fusion protein. This suggests that rIg3 may contain a low affinity binding site. To determine whether such a site exists in rat NCAM and, if it does, whether it can function better in the chick environment, chimeric recombinant proteins were constructed and expressed in bacteria (Fig. 13). The chick/rat chimeric Ig3 (c(r)Ig3) was constructed and expressed as a (His)$_6$-tagged fusion protein. For comparison, a chick/human chimeric (His)$_6$-tagged Ig3 (c(h)Ig3) was also constructed. In both proteins, the chick homophilic binding sequence (between Lys-243 and Glu-252) was replaced by the analogous region in
Fig. 11. Binding of rIg3 protein-conjugated Covaspheres to various substrate-coated proteins. (A) Epifluorescence micrographs of the binding of rIg3 to substrate-coated cIg3 (a), rIg3 (b), GST (c), or cIgl (d). (B) Quantitative analysis of the binding of rIg3 Covasphere binding to substrate-coated cIg3, rIg3, GST and cIg1. All values represent the mean ± S.D. of 3 experiments.
Fig. 12. Competition of the binding of rIg3-Covaspheres to substrate-coated cIg3 by soluble cIg3 and rIg3 fusion proteins. Covaspheres conjugated with rIg3 were pre-incubated for 15 minutes with varying amounts of either cIg3 (O) or rIg3 (o) before adding to substrate-coated cIg3. All values were normalized to the number of Covaspheres bound per unit area in the absence of any competitor and data represent the mean ± S.D. of 3 experiments.
residues within the decapeptide sequence that differed from the wild-type chick sequence.

By replacing the entire decapeptide sequence of the chick binding site, binding activity should only be observed in constructs which contain sequences that act to rescue homophilic binding. Thus, if these chimeric proteins, substituted with the human or rat analogue decapeptide sequence, can still undergo homophilic binding, these results will provide evidence that the inserted sequences serve as homophilic binding sites in their respective species.

A chick Ig3 fusion protein construct (Rao et al., 1993), containing a (His)$_6$-tag at its amino-terminus, was used as a template for the chimeric constructions. This His-Ig3 fusion protein was previously shown to be active in a number of binding studies, including Covasphere binding assays. The insertion of the putative human and rat binding site sequences into the chick Ig3 domain construct was achieved by using the Transformer™ site-directed mutagenesis kit. Amino acid residues 243-252 of cIg3 were replaced with either residues 242-251 (KYIFSDDDSSQ) of the human sequence (Dickson et al., 1987), or residues 243-252 (KHIFSDDDSSE) of the rat sequence (Small et al., 1987). The chimeric proteins were expressed in E. coli and collected from the bacterial inclusion bodies. Purification of the recombinant proteins was performed on a nickel resin column under denaturing condition (Fig 14A). To renature the chimeric fusion proteins, the samples were initially reduced with β-mercaptoethanol, reoxidized by dialysis in 8 M urea and then diluted and dialyzed against PBS. For both proteins, the yield following column purification was ~20 mg/litre culture, and ~10 mg/litre culture after refolding.

To check for purity, a sample of each of the chimeric proteins was analyzed by SDS-PAGE and visualized by Coomassie Blue staining (Fig 14B). Under reducing conditions, both proteins migrated as a single band with an apparent molecular weight of ~16 kDa, which is slightly slower than the predicted molecular weight of ~14 kDa. The decreased mobility could be contributed to the presence of the (His)$_6$-tag added to the
Fig. 13. NCAM chimeric fusion protein constructs. (A) NCAM constructs were based on the chick NCAM cDNA. The domain construct cIg1 was created using standard PCR methods with the pEC1402 vector (Edelman et al., 1987) as the template. Chimeric fusion protein constructs were created by mutagenesis of the chick Ig3 construct, using the Clonetech Transformer™ site-directed mutagenesis kit. Numbering of amino acids in the chick NCAM protein sequence is in accordance with Cunningham et al. (1987). All chick constructs contain a (His)$_x$-tag at their NH$_2$-terminus (slim white box), were purified under denaturing conditions and were then refolded. Shaded regions in the NCAM protein represent Ig-like domains, while unshaded regions indicate where fibronectin type-III repeats are located; the stippled area represents the predicted transmembrane segment and the striped area shows the cytoplasmic domain. (B) Partial sequence alignment of the two chimeric fusion proteins with the wild-type chick Ig3 fusion protein. The bolded and boxed area shows the region of cIg3 in which mutations were incorporated. The chick decapeptide sequence was replaced with the putative human or rat homophilic binding site, with shaded residues representing those that were altered.
A

Chick NCAM-140

NH₂

L1  K98

V195  K288

human

rat

cIg1

cIg3

c(h)Ig3

c(r)Ig3

B

230  243  252  270

cIg3  ...TWT KDGEPIEQED NEE KYSFNYS GSE LIIKKVD KSDEAYICI AE ....
c(h)Ig3  ...TWT KDGEPIEQED NEE KYIFSDD SSO LIIKKVD KSDEAYICI AE ....
c(r)Ig3  ...TWT KDGEPIEQED NEE KHFPSDD SSE LIIKKVD KSDEAYICI AE ....
Fig. 14. Purification of chimeric fusion proteins. (A) A representative column elution profile of the c(r)Ig3 His-tagged fusion protein. Total cell lysate, solubilized in Buffer A (6 M guanidine HCl, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris/HCl, pH 8.0) was adsorbed on a nickel-NTA column and washed with Buffer B (8 M urea, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris/HCl, pH 8.0), Buffer C (8 M urea, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris/HCl, pH 6.3) and then eluted with Buffers D (8 M urea, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris/HCl, pH 5.9) and E (8 M urea, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris/HCl, pH 4.5). Column fractions of 2 ml were collected and subjected to SDS-PAGE and spectrophotometric analysis. Fractions containing the majority of the protein were pooled and refolded. (B) Gel profile of purified, refolded cIg3, c(h)Ig3 and c(r)Ig3 fusion proteins. Protein samples were separated on a 14 % SDS-polyacrylamide gel under reducing conditions and then stained with Coomassie Brilliant Blue.
NH₂-terminus of the fusion protein. An increased gel mobility of these two proteins was observed under non-reducing conditions (data not shown), suggesting the presence of an intramolecular disulfide bond. The overall gel profile of both chimeric proteins was consistent with that observed with the cIg3 fusion protein.

*The Chimeric Chick/Rat Protein, c(r)Ig3, did not Exhibit Homophilic Binding Activity*

To determine whether the chimeric c(r)Ig3 fusion protein retained homophilic binding activity, c(r)Ig3 was conjugated to Covaspheres and allowed to bind to various substrate-coated proteins (Fig 15). Very low binding of c(r)Ig3-conjugated Covaspheres was observed on substrate-coated c(r)Ig3 and only background level of binding was observed on substrate-coated cIg1. However, significantly higher binding was observed on substrate-coated cIg3. The introduction of the putative rat binding site sequence in cIg3 appears to cause a reduction in the binding affinity of the chimeric fusion protein. These results indicate that the c(r)Ig3 chimeric protein does not significantly support homophilic interaction in the Covasphere-to-substratum binding assay. However, since c(r)Ig3 was able to bind cIg3, it is possible that the putative rat binding site may have a very low binding affinity or that the flanking chick sequences are capable of mediating binding interactions with cIg3.

*The Chimeric Chick/Human Protein, c(h)Ig3, Exhibited Homophilic Binding Activity*

To investigate the homophilic binding activity of c(h)Ig3, the chimeric c(h)Ig3 fusion protein was conjugated to Covaspheres and allowed to bind to substrata coated with various protein solutions. In contrast to results obtained with c(r)Ig3, c(h)Ig3-conjugated
Fig. 15. Binding of c(r)Ig3 protein-conjugated Covaspheres to various substrate-coated proteins. (A) Epifluorescence micrographs of the binding of c(r)Ig3 protein-conjugated Covaspheres to substrates coated with cIg3 (a), c(r)Ig3 (b) and cIg1 (c). (B) Covaspheres conjugated with c(r)Ig3 fusion protein were assayed for its ability to bind to substrate-coated cIg3, c(r)Ig3 or cIg1. Data represent the mean ± S.D. of 3 experiments.
Ig1 (Fig 16). Positive binding was also observed on substrate-coated cIg3. As a comparison, the wild-type cIg3 fusion protein was also conjugated to Covaspheres and allowed to bind to the same substrate-coated proteins. As expected, cIg3-conjugated Covaspheres bound very well to substrate-coated cIg3 but poorly to cIg1. Significant binding was retained even when the amount of cIg adsorbed on the plastic was reduced by 1/10 (data not shown). There was also significant binding of the cIg3-Covaspheres to the c(h)Ig3-substrate. These results indicate that the c(h)Ig3 chimeric protein retains homophilic binding activity in the Covasphere-to-substrate binding assay and suggests that the putative human binding site is capable of mediating chimeric Ig3-to-Ig3 binding.

Interestingly, a 20% reduction in relative binding activity is observed in the cIg3-c(h)Ig3 interaction, relative to the cIg3-cIg3 interaction (Fig 16B). Similarly, a 20% difference is seen between the c(h)Ig3-cIg3 and c(h)Ig3-c(h)Ig3 interactions. These results suggest that the introduction of the human binding sequence may cause a reduction in the binding affinity of the cIg3 fusion protein. They also suggest that the relative binding strength between cIg3 and c(h)Ig3 may be ranked as follows: cIg3-cIg3 > cIg3-c(h)Ig3 ≡ c(h)Ig3-cIg3 > c(h)Ig3-c(h)Ig3. In general, the binding of c(h)Ig3-Covaspheres to substrate-coated cIg3 was 3- to 4-fold higher than that of c(r)Ig3-Covaspheres, suggesting that the interaction between c(r)Ig3 and cIg3 is a relatively weak one.

The specificity of the c(h)Ig3-c(h)Ig3 interaction was tested in competition assays using soluble chimeric proteins. Covaspheres conjugated with c(h)Ig3 were pre-incubated with varying concentrations of the competitor before assaying for binding to the substrate-coated c(h)Ig3 (Fig 17). Binding was inhibited in a dose dependent manner by both cIg3 and c(h)Ig3, with 50% inhibition occurring at ~0.8 μM of the competitor (Fig 17B). Complete inhibition was achieved at ~12 μM. Interestingly, both c(h)Ig3 and cIg3 competed equally in this assay despite the earlier observation that c(h)Ig3 binds better with
Fig. 16. Binding of chick Ig3 and c(h)Ig3 protein-conjugated Covaspheres to various substrate-coated proteins. (A) Epifluorescence micrographs of the binding of cIg3 and c(h)Ig3 protein-conjugated Covaspheres. Panels a to c are representative micrographs showing the binding of cIg3-conjugated Covaspheres to substrate-coated cIg3 (a), c(h)Ig3 (b), or cIg1 (c). In panels d to f, c(h)Ig3-conjugated Covaspheres are shown interacting with substrate-coated c(h)Ig3 (d), cIg3 (e) and cIg1 (f). (B) Quantitative analysis of the binding of the cIg3- and c(h)Ig3-conjugated Covaspheres to substrate-coated cIg3, cIg1 and c(h)Ig3. All values were normalized to the number of Covaspheres bound per unit area in self-binding experiments (i.e. cIg3-to-cIg3 or c(h)Ig3-to-c(h)Ig3) and data represent the mean ± S.D. of 3 experiments.
Fig. 17. Competition of c(h)Ig3 Covasphere to substrate-coated c(h)Ig3 binding.

(A) Epifluorescence micrographs of the binding of c(h)Ig3-conjugated Covaspheres to substrate-coated c(h)Ig3, in the absence (a) or presence of soluble protein competitors (b-d). In panel b, c(h)Ig3 Covaspheres were pre-incubated with ~10 μM soluble c(h)Ig3 for 15 minutes before seeding on the c(h)Ig3 substrate. In panels c and d, the c(h)Ig3 Covaspheres were pre-incubated with ~10 μM soluble cIg3 and cIg1 proteins, respectively.

(B) Dose effects of soluble competitors on the binding of c(h)Ig3-conjugated Covaspheres. Covaspheres conjugated with c(h)Ig3 were pre-incubated with varying amounts of cIg3 (o), c(h)Ig3 (□) or cIg1 (X) for 15 minutes, before adding to substrate-coated c(h)Ig3. All values were normalized to the number of Covaspheres bound per unit area in the absence of competitor and data represent the mean ± S.D. of 3 experiments.
differences in affinity between two proteins, or due to possible cIg3-cIg3 equilibrium interactions affecting the concentration of free cIg3 available for inhibition of the c(h)Ig3-c(h)Ig3 interaction.

Expression and Characterization of a Rat/Chick Chimeric Ig3 Fusion Protein

The above results suggest that the putative rIg3 binding sequence might be a low affinity binding site. It is also possible that the flanking sequences of rIg3 does not support the function of a homophilic binding sequence. To address this issue, a rat/chick chimeric fusion protein r(c)Ig3 was constructed using the GST expression system to obtain soluble protein (Fig. 18). This r(c)Ig3 fusion protein construct is the inverse of the c(r)Ig3 construct, with the rIg3 fusion protein containing the chick homophilic binding site sequence. Mutagenesis of the GST-rIg3 fusion protein was achieved by the 4 primer method (Higuchi, 1990), with the proposed rat binding site sequence being replaced with the chick sequence. All steps in the expression, purification and characterization of this protein were in accordance to those followed with the rat GST-Ig3 fusion protein. The substitution of the putative low affinity rat sequence with the high affinity chick sequence should confer homophilic binding activity in this construct.

Surprisingly, when tested in the Covasphere-to-substratum binding assay, the r(c)Ig3 protein displayed very low binding activity. It displayed a binding pattern very similar to that of the GST-rIg3 fusion protein. Although significant levels of binding were achieved with cIg3, only background levels of binding were observed with substrate-coated r(c)Ig3 and rIg3. These results suggest that the chick homophilic binding site may not be entirely compatible with the rat environment, resulting in a weak binding protein.
Fig. 18. Schematic diagram of the rat/chick Ig3 chimeric protein. (A) The r(c)Ig3 construct was created based on the rat NCAM cDNA using standard PCR methods. The chimeric r(c)Ig3 construct was based on this rIg3 construct and replaced the putative rat homophilic binding site with that of chick NCAM. Mutagenesis was performed using the four primer method (Higuchi, 1990). Numbering of amino acids in the rat NCAM protein sequence is in accordance with Small et al. (1987). Dotted regions on the protein schematic represent Ig-like domains in the rat NCAM protein and fibronectin type-III repeats are located in the white area; the cross-hatched region shows the location of the transmembrane segment and the striped area represents the cytoplasmic domain of the protein. The r(c)Ig3 cDNA was inserted into the GST expression vector and then expressed as a soluble protein. The GST-tag is shown as a shaded box at the amino-terminus of the fusion protein. (B) Partial sequence alignment of the rIg3 and r(c)Ig3 fusion proteins. The bolded and boxed area identifies the region of rIg3 which is analogous to the chick homophilic binding site and where mutations were incorporated. Shaded residues represent those which were mutated. (C) Gel profile of the purified r(c)Ig3 fusion protein. Protein samples were separated on a 10 % SDS-polyacrylamide gel under reducing conditions and then stained with Coomassie Brilliant Blue.
Fig. 19. Binding of r(c)Ig3-conjugated Covaspheres to various substrate-coated proteins. Covaspheres conjugated with r(c)Ig3 were allowed to bind to substrate-coated cIg3, r(c)Ig3, rIg3, and GST. All data were normalized to the number of Covaspheres bound per unit area on substrate-coated cIg3 and represent the mean ± S.D. of 3 experiments.
**Nitrocellulose-coated Petri Dish**

The binding activity of chimeric fusion proteins was further evaluated using the cell to substratum binding assay. Since we have shown that cIg3-conjugated Covaspheres are able to bind rat NCAM expressed on the rat pheochromocytoma PC12 cells (Rao et al., 1994), the attachment of PC12 cells to various substrate-coated fusion proteins was examined. Fusion proteins were adsorbed on to nitrocellulose-coated Petri dishes, blocked with 2% BSA and then assayed for their ability to mediate PC12 cell attachment. As expected, the cIg3 fusion protein was able to mediate PC12 cell binding at a relatively high level (Fig. 20). A GST-cIg3 fusion protein was also assayed. Since the molar concentration of the GST-cIg3 protein adsorbed on the plastic was lower that that of His-cIg3, it would predict a lower level of PC12 cell binding. Indeed, PC12 cell attachment was reduced by 20% in comparison to the His-cIg3 fusion protein. In contrast to the Covasphere binding results, both His-c(r)Ig3 and GST-r(c)Ig3 chimeric proteins were able to support the attachment of PC12 cells. However, only a marginal level of PC12 cell binding was achieved on substrate-coated GST-rIg3.

**Effects of Mutating Six Residues in the Chick Ig3 Fusion Protein Binding Site to Alanine**

Studies using the chimeric fusion proteins showed that substitution of several residues in the chick binding site could be tolerated, unless deleterious mutations are incorporated. To study the extent of this tolerance, six of the ten residues in the chick homophilic binding site were mutated to alanine, while the four highly conserved residues were not altered (Fig. 21A). The six amino acid residues within the chick homophilic binding site (KYSFNYDGSE) that were not conserved among vertebrate NCAM species were substituted with alanine and this resulted in the new sequence of KAAFAADASA in
Fig. 20. Binding of PC12 cells to various substrate-coated fusion proteins.

Proteins were adsorbed to small spots (~3 mm in diameter) on nitrocellulose-coated Petri dishes and then tested for their ability to mediate the attachment of PC12 cells. (A) Phase contrast micrographs of rat PC12 cells binding to substrates coated with cIg3 (Aa), c(r)Ig3 (b), GST-cIg3 (c), r(c)Ig3 (d), rIg3 (e), and BSA (f). (B) Quantification of PC12 cell attachment to various substrates. PC12 cells were allowed to attach to substrate-coated cIg3, BSA, r(c)Ig3, c(r)Ig3, and rIg3. Unbound cells were removed by gentle washing, and samples were then examined and photographed using an inverted microscope. The number of bound cells was counted from the pictures. All values represent the mean ± S.D. of 3 experiments.
A

B

Substrate-coated protein

Number of cells bound per unit area

(\text{His})_6
\text{clg3}

(\text{His})_6
\text{c(r)lg3}

\text{GST}
\text{clg3}

\text{GST}
\text{r(c)lg3}

\text{GST}
\text{rlg3}

\text{BSA}
To determine if the His-c(4)Ig3 fusion protein retained homophilic binding activity, the mutant was examined in a Covasphere-to-substratum binding assay. Covaspheres were conjugated with c(4)Ig3 and binding to various substrates was assessed. c(4)Ig3-conjugated Covaspheres bound well to a substrate-coated c(4)Ig3 (Fig. 22). However, the binding was relatively poor with other substrate-coated proteins. These results demonstrate that c(4)Ig3 retains some binding activity and suggest that the four conserved amino acid residues are sufficient to mediate homophilic binding. Interestingly, c(4)Ig3 does not bind significantly to cIg3. The six substituted residues might play a role in determining species specificity and possibly in the regulation of cIg3-cIg3 interaction.
Species | amino acid sequence
--- | ---
Chick | K Y S F N Y D G S E
Xenopus | K Y L F N E D Q S E
Bovine | K Y L F S D D S S E
Rat | K H I F S D D S S E
Human | K Y I F S D D S Q

Conserved residues | K - - F - - D - S -

Alanine substitutions | K A A F A A D A S A

**c(4)Ig3** (His)_6 S-S

Fig. 21. Construction of the mutant chick Ig3 fusion protein, c(4)Ig3. (A) Alignment of the chick homophilic binding site (between amino acid positions 243 and 252) with the analogous region in *Xenopus*, Bovine, rat, and human NCAM. The conserved residues are shown in boldface type and boxed. Residues which are not conserved among the various NCAM species, were substituted with alanine. The mutagenesis was performed on the His-cIg3 fusion protein construct, using a strategy similar to that adopted in the construction of the chimeric chick Ig3 fusion proteins. (B) Gel profile of the purified and refolded c(4)Ig3 protein. Protein samples were separated on a 14 % SDS-polyacrylamide gel under reducing conditions and then stained with Coomassie Brilliant Blue. Molecular weight markers (M) are shown on the left lane.
Fig. 22. Binding of c(4)Ig3-conjugated Covaspheres to various substrate-coated fusion proteins. (A) Round spots (~3 mm in diameter) on Petri dishes were coated with various proteins and c(4)Ig3-conjugated Covaspheres were added to these spots. After incubation for 1 hour at room temperature, the number of Covaspheres bound per unit area was estimated. (B) Competition experiments were carried out by preincubating c(4)Ig3-conjugated Covaspheres with 0, 1.5 or 15.5 μM of soluble c(4)Ig3 fusion protein before seeding on top of substrate-coated c(4)Ig3. All values were normalized to the number of Covaspheres bound to a substrate-coated c(4)Ig3 and data represent the mean ± S.D. of 3 experiments.
Chapter 4: Discussion
In this thesis, work was performed to establish and understand differences in binding observed in rat NCAM and chick NCAM. Single Ig-like domains, chimeras, and binding site mutants in both chick and rat Ig3 were made to address this question. Protein binding ability was tested in Covasphere-to-substratum and cell-to-substratum binding assays.

Our previous studies have shown that, in chick NCAM, homophilic binding is mediated by a decapeptide sequence in the third Ig-like domain (Rao et al., 1992). In contrast to these results, studies in rodent NCAM have indicated an absence of Ig3-mediated binding ability (Frei et al., 1992; Zhou et al., 1993). To address this apparent difference, a fusion protein containing rat NCAM Ig-like domain 3 (rIg3) was created and expressed in bacteria. When tested in the Covasphere-to-substratum binding assay, rIg3 showed negligible homophilic binding activity. Despite high sequence identity between the rat and chick NCAM suggesting a high degree of structural and functional similarities, these results suggest a fundamental difference between the rat and chick Ig3-associated homophilic binding ability. This may be due to rat NCAM possessing a limited or much reduced binding capability in its Ig3 domain. The fact that rIg3 is able to interact with chick Ig3 shows its capability, however small, to mediate binding. In this assay system, the saturating level of protein adsorbed on plastic was consistently ~0.8 µg per ~10 mm spot. For GST fusion proteins, including rIg3 and r(c)Ig3, the molar concentration of the adsorbed protein was equivalent to about one-third of the His-tagged proteins because of the differences in their molecular weights. It is possible that the lower molar concentration of substrate-coated protein might have contributed to the lower level of Covasphere binding. Furthermore, rIg3 showed only marginal activity in the more robust cell-substratum binding assay. Results obtained with both assays are consistent with the interpretation that rIg3 has a very low homophilic binding activity. To explore the possibility that the low rat Ig3 homophilic binding may be attributed to either a diminished
otherwise active binding site, a chick/rat chimera, c(r)Ig3, was constructed. This chimera inserted the putative rat binding sequence into the positive-binding environment of chick Ig3. In Covasphere-to-substratum binding assays, c(r)Ig3 exhibited very low homophilic binding. Since this protein was a His-tagged fusion protein, the molar concentration of protein adsorbed on plastic was comparable to cIg3 and other active chimeric proteins. These results are similar to those obtained with the rIg3 fusion protein, suggesting that the low rIg3 binding activity may be due to the putative rat binding sequence being unable to support strong homophilic binding. In comparison to the background levels of binding observed for rIg3, c(r)Ig3 shows binding activity consistently higher than the background level. This low level of binding activity may be attributed to the residual binding activity in the rat sequence or positive effects of the chick Ig3 environment on binding. On the other hand, when tested in the cell-to-substratum binding assay, c(r)Ig3 was capable of mediating a significant level of PC12 cell attachment. Conformational freedom, lateral membrane mobility and clustering of adhesion complexes are possible for rat NCAM on the PC12 cell surface and likely account for the more robust binding ability of cells on this c(r)Ig3-coated substrate.

The insertion of the putative rat decapeptide binding sequence into chick Ig3 shows a decrease in its binding ability in the Covasphere-to-substratum assay, which we attribute to the low affinity rat binding sequence. Although, our previous results in chick indicate that the homophilic binding site is essential and sufficient for chick Ig3 binding, regions flanking the binding site likely affect the binding site activity. By inserting the rat sequence into chick Ig3, local conformational changes may have resulted. It may be argued that the any non-chick sequence will result in the same low binding levels observed, and that the residual binding is due to the chick Ig3 domain surrounding the binding site sequence. To control for this possibility, another chimera, c(h)Ig3, was constructed. In contrast to c(r)Ig3, the c(h)Ig3 fusion protein showed significant binding in both Covasphere-to-
between the inserted putative binding sequences and binding activity. Thus, the putative rat homophilic binding sequence appears to be low affinity.

The ability of c(h)Ig3 to bind in the Covasphere-to-substratum binding assay implies that the putative human NCAM binding sequence retains homophilic binding activity. Based upon the ~80% amino acid sequence identity between the chick, rat and human NCAM Ig3, results obtained for the chimeric fusion proteins can be considered to be representative of the Ig3 in their respective species. When we compared the rIg3 and r(c)Ig3 fusion proteins in the Covasphere-to-substratum binding assay, this was shown to be largely true. Our results therefore are consistent with the interpretation that the human NCAM Ig3 possesses homophilic binding activity centered at the putative decapeptide binding sequence. Interestingly, when the putative human homophilic binding site is inserted into cIg3, lower levels of binding are observed for the c(h)Ig3-c(h)Ig3 interaction, relative to the cIg3-cIg3 interaction. This may suggest that the human binding site sequence mediates human NCAM homophilic binding at a lower affinity than that observed with the chick homologue.

Among the vertebrate NCAM species, only four of the ten binding site residues are highly conserved. To specifically examine the role of these residues in mediating binding, all other residues within the chick homophilic binding site were mutated to alanine. When tested in the Covasphere-to-substratum binding assay, c(4)Ig3 exhibited significant homophilic binding activity. These results suggest that these four conserved amino acid residues are sufficient to support homophilic binding. Also, the substitution of multiple residues to alanine may create a hydrophobic surface or region that can further enhance protein-protein interaction. Interestingly, the homophilic binding activity of c(4)Ig3 appears to be highly specific. These results suggest that the other six residues in the homophilic binding site may participate in modulating affinity, as seen in c(r)Ig3, or may
It has been suggested that the putative rat binding site is only able to mediate low affinity binding. Since the Ig3 domain of vertebrate NCAM species are highly conserved, one would expect the recovery of homophilic binding activity by replacing the putative low affinity rat binding site with the high affinity chick binding site. This possibility was tested in the r(c)Ig3 fusion protein. Unexpectedly, when tested in the Covasphere-to-substratum binding assay, r(c)Ig3 showed negligible homophilic binding ability. However, when tested in the cell-to-substratum binding assay, a significant level of cell binding was achieved. The latter results strengthen the notion that the rat NCAM Ig3 is unable to mediate homophilic binding, due to a low affinity rat homophilic binding sequence.

Differences in binding activity was observed when c(r)Ig3 and r(c)Ig3 were assayed in the binding assays. As mentioned above, since the rat NCAM on the cell surface possesses conformational freedom, lateral membrane mobility and ability to cluster, the cell-to-substratum assay is likely able to display lower affinity protein interactions. The Covasphere-to-substratum assay, in contrast, is a solid-state assay and likely relies on higher affinity interactions. When used together, these two assays appear to partially complement one another, with the Covasphere assay displaying higher affinity interactions, while the cell attachment assay displaying the lower affinity interactions.

In further analyzing the results obtained in the rat PC12 cell attachment assay, the flanking sequences surrounding the binding site also appear to contribute to cell-to-substratum binding. In these assays, the chick Ig3 sequences may provide a positive environment for binding, with interchanging of the binding site sequence affecting affinity in the domain. Conversely, when the high affinity chick binding site was inserted into the rat environment in r(c)Ig3, only partial recovery of binding activity was recovered, suggesting that the rat environment has a negative effect on homophilic binding. Thus, under the assay conditions tested, the combination of the low affinity binding site and the
assays. From these results, we conclude that Ig3 in rat NCAM probably play a much lesser role than its counter-part in human and chick NCAM in mediating NCAM-NCAM interactions.

In these studies, changes made to the homophilic binding site are reflected as changes in binding affinity and specificity in the chick NCAM Ig3 binding interactions. This was especially apparent in the case of the c(h)Ig3 and c(r)Ig3 protein constructs, where large differences are observed in the Covasphere binding assays. Although both proteins differ from the chick binding site sequence at five positions, they only differ from each other at two positions. Furthermore, only one residue, His-244 of the rat sequence, differs from both the chick and human binding site sequences. This single amino acid residue may, then, be responsible for the observed differences in binding affinity between the c(h)Ig3 and c(r)Ig3 fusion proteins. This could be due to possible conformational changes attributed to the presence of this residue in the binding site sequence.

Secondary structure predictions, based on the algorithm developed by Chou and Fasman (1978), indicate that the amino-terminal portion of the chick binding site sequence may adopt a β-strand conformation (Fig. 23). This region contains alternating hydrophilic and hydrophobic residues, which is characteristic of this type of secondary structure. The replacement of the chick sequence with the proposed human binding site sequence increases the calculated probability that this portion of the sequence will adopt a β-strand conformation. Conversely, when the rat sequence is introduced into the chick environment, the His-244 residue causes this amino-terminal region to adopt a predicted α-helical conformation. The presence of the His-244 residue in this area may destabilize the β-strand conformation necessary for homophilic binding. By analogy to the known crystal structures of C2-set Ig domains, the amino-terminal portion of the chick binding sequence would correspond to the C' strand and then continue into the C'-E β-turn region.
Fig. 23. Schematic diagram of chick NCAM Ig3 depicting the predicted position of the homophilic binding site sequence. Residues comprising the homophilic binding site (shown in bold) are located in the C' β-strand and the following C'-E β-turn.
presence of the His-244 residue could possibly alter the conformation required for specific interaction.

The majority of evidence supporting Ig3 in mediating homophilic binding has largely centered around chick NCAM (Watanabe et al., 1986; Rao et al., 1993; Ranheim et al., 1996). Rat/mouse is the next most utilized species (Kadmon et al., 1990; Zhou et al., 1993; Kiselyov et al., 1997), and human NCAM has seldomly been used in binding studies. In addition to work from our laboratory, Ranheim et al. (1996) have shown that chick NCAM-mediated homophilic binding is likely centered around Ig3 of NCAM. They also established that there are lower affinity interactions between other domains in the molecule, in an anti-parallel manner (i.e. Ig1 pairing with Ig5 and Ig2 with Ig4). With rat/mouse NCAM, Ig3 seems to play a lesser role in NCAM-mediated binding (Frei et al., 1992). Their studies show that Ig1, Ig2, and FnIII-1 and FnIII-2 together, all possess some ability to mediate attachment of neuronal cells to protein-coated substrates, with Ig3 showing little binding ability in this assay. In studies using chimeric constructs of NCAM and the CEA molecule (Zhou et al., 1993), it has been suggested that the first three domains of rat NCAM are not sufficient to mediate cell aggregation and that all five domains are necessary for NCAM binding. In the most recent studies on rat NCAM-mediated binding (Kiselyov et al., 1997), Ig1 and Ig2 are again implicated in rat NCAM binding. This study suggests direct binding between Ig1 and Ig2 on apposing NCAM molecules, in a double reciprocal bonding manner. The consensus among the rat/mouse studies appears to be a lack of binding activity associated with the third Ig-like domain. Our work is consistent with this view and further postulates the low affinity rat binding site and a partially non-favourable rat Ig3 environment as a possible explanation. This also explains the need for an alternative binding mechanism in rat NCAM.

The present evidence showing Ig3-mediated NCAM binding (in chick), and the lack of Ig3-mediated NCAM binding (in rat), has led to the generation of two separate molecular
Ig3 mediates the initial contact between apposing NCAM molecules, followed by anti-parallel interactions, with Ig domains 1, 2, 3, 4, and 5 interacting with Ig domains 5, 4, 3, 2, and 1, respectively, (Fig. 24A). Of the four residues which were shown to be sufficient to mediate homophilic binding, as in the case of c(4)Ig4, two residues (Lys-243 and Asp-249) are charged and may mediate initial electrostatic interactions between the two NCAM molecules. This may be similar to the homophilic binding mechanism of gp80, a cell adhesion molecule in *Dictyostelium discoideum*, known to mediate Ca$^{2+}$-independent cell-cell adhesion (Siu et al., 1995; Kamboj et al., 1988). In gp80, the homophilic binding site centers around an octapeptide sequence YKLNVDS, which is also predicted to adopt a β-strand conformation (Kamboj et al., 1989), and the two charged residues (Lys and Asp) have been shown to play vital roles in gp80-gp80 binding. The close apposition of these charged residues in NCAM would likely cause NCAM Ig3 to interact in an anti-parallel manner. These initial electrostatic interactions may be stabilized by subsequent hydrophobic interactions and/or the formation of hydrogen bonds between the two Ig3 domains and on other secondary sites along the two molecules.

In the other model (Kieselyov et al., 1997), based on rat NCAM, a cluster of basic residues in Ig2 are postulated to interact with a cluster of acidic residues in Ig1 to create reciprocal bonding between the Ig1 and Ig2 on apposing NCAM molecules (Fig. 24B). Although they do not observe anti-parallel, pair-wise interactions between the NCAM Ig-like domains depicted in the chick model, they attempt to reconcile the discrepancy by proposing a concurrent sliding model. They suggest that the developmentally regulated presence of the PSA moiety on Ig5 controls a switch between the reciprocal Ig1/Ig2 binding model and the 5 Ig domain contact model. In the absence of the PSA, NCAM is able to slide in or slide out using these two mechanisms. Since interactions between Ig1 and Ig2 have not been observed for chick NCAM, we suggest that two independent mechanisms have evolved for NCAM. Our results using cIg3 and rIg3 fusion proteins suggest a
mentioned above, rat NCAM may have developed other compensatory binding mechanisms after it has incurred a loss of Ig3-mediated homophilic binding ability.

In summary, our studies have investigated the homophilic binding activity of NCAM Ig3 in the chick, human and rat species. We have created chimeric fusion proteins, single domain constructs and a binding site mutant. Our results suggest that human NCAM may mediate homophilic binding in the same manner as chick NCAM and that the binding site sequence is located in the same region in human NCAM Ig3. We also present evidence suggesting that rat Ig3 possesses a reduced homophilic binding capability and that other binding mechanisms may be prevalent in this NCAM species. Furthermore, the initial chick Ig3 interaction is likely mediated by an electrostatic contact between the residues 243-Lys and Asp-249 of the homophilic binding site sequence and this suggests an anti-parallel orientation for apposing NCAM molecules.

For future consideration, mutation of the His-244 residue in the rIg3 fusion protein would be a high priority. In sequence analysis of Ig3 among human, rat and bovine NCAM, a high degree (87 to 95%) of sequence identity is observed. With the exception of His-244, all changes represent conservative substitutions as a result of single point mutations. Our prediction is that mutation of His-244 may restore binding activity to a similar, or even greater degree, than the r(c)Ig3 chimera. Choices of substituted residues may include alanine or tyrosine, the latter being present in chick and human NCAM.

To extend these results and to test the proposed rat and chick NCAM binding models, it might be of interest to express the chick and rat Ig1-2. By testing Ig1-2 versus Ig3 in binding studies, differences in binding mechanisms between these two species would be clearly evident.

Furthermore, results from this study suggest that the putative human binding site mediates homophilic binding in human NCAM. Direct evidence for this still needs to be
establish this point.

Ideally, the full biochemical analysis of the Ig3-mediated NCAM binding mechanism should include X-ray crystallography and NMR methods. The spatial arrangement of residues within the homophilic binding site and their roles in binding may be precisely defined when the structure is solved.
Fig. 24. Models for NCAM-NCAM interactions. (A) Model proposed for chick NCAM (Ranheim et al., 1996). The Ig-like domains are predicted to bind in pairs, with Ig1 interacting with Ig5, Ig2 interacting with Ig4, and Ig3 interacting with itself. Pairing and binding is initiated by Ig3-Ig3 interactions because of its higher binding affinity. (B) Model proposed for rat NCAM (Kieselyov et al., 1997). NCAM-NCAM interaction is strongest between Ig1 and Ig2 domains of rat NCAM, in a reciprocal binding manner. Preference for Ig1-Ig2 binding occurs in the presence of the PSA moiety on Ig5. In the absence of this moiety, the Ig-like domains slide along each other and adopt the binding model proposed for chick NCAM. Ig-like domains are shown as shaded ovals, Fn-III domains are shown as shaded squares, and the PSA moiety is represented by 🌿.
References


protein-protein interactions mediate the homophilic binding of the neural adhesion molecule P.


He, H. T., Barber, J., Chaix, J. C., and Goridis, C. (1986). Phosphatidylinositol is involved in the membrane attachment of NCAM-120, the smallest component of the neural cell adhesion molecule. EMBO J. 5, 2489-2494.


