

Developmental neurobiology: A genetic Cheshire cat?

Jerold Chun

In the wake of evidence that essential neurogenic processes might involve aspects of DNA rearrangement, recent discoveries about the unusual arrangement of genes encoding neuronal adhesion molecules known as protocadherins are very intriguing. But is this just a coincidence?

Address: Department of Pharmacology, Neurosciences Program, Biomedical Sciences Program, School of Medicine, University of California, San Diego, California 92093-0636, USA.
E-mail: jchun@ucsd.edu

Current Biology 1999, 9:R651–R654
<http://biomednet.com/elecref/09609822009R0651>

0960-9822/99/\$ – see front matter © 1999 Elsevier Science Ltd.
All rights reserved.

“Curiouser and curiouser!” Alice’s reaction to the strange events encountered during her *Adventures in Wonderland* is being echoed by biologists puzzling over recent tantalizing hints that somatic DNA rearrangement — of a kind well-known to be of fundamental importance in the immune system — may operate in the nervous system. The first of these hints, reviewed recently in these pages [1], came from close examination of mutant mice deficient for components of the enzymatic machinery essential for the DNA rearrangements that create the genes for mature antigen receptors of B and T lymphocytes. These mutant mice, deficient for either DNA ligase IV [2,3] or its dimerization partner XRCC4 [2], were found to exhibit a striking increase in the death of very young neurons and neuroblasts. This phenotype has subsequently been found to be shared by mice deficient in other components of the so-called ‘non-homologous end-joining system’ (F. Alt, personal communication).

These results have excited interest in the possibility that some form of somatic DNA rearrangement might be associated with the development of the nervous system, although this is certainly not the only explanation for the observations. If DNA rearrangement does turn out to be important in the nervous system, what are the relevant genetic targets? Such a target is, of course, likely to be expressed in the nervous system, probably in specific subpopulations of neurons; by analogy with the antigen-receptor genes, it is likely to consist of a large gene family, and the genes might well encode cell-surface molecules [4]. An attractive candidate target that fits these criteria has been identified. This is the family of genes encoding the protocadherins [5–7], a subfamily of the cadherin adhesion molecules. The protocadherin genes in humans have recently been found [8] to have a genomic structure

that is rather suggestive of a locus that might undergo DNA rearrangements.

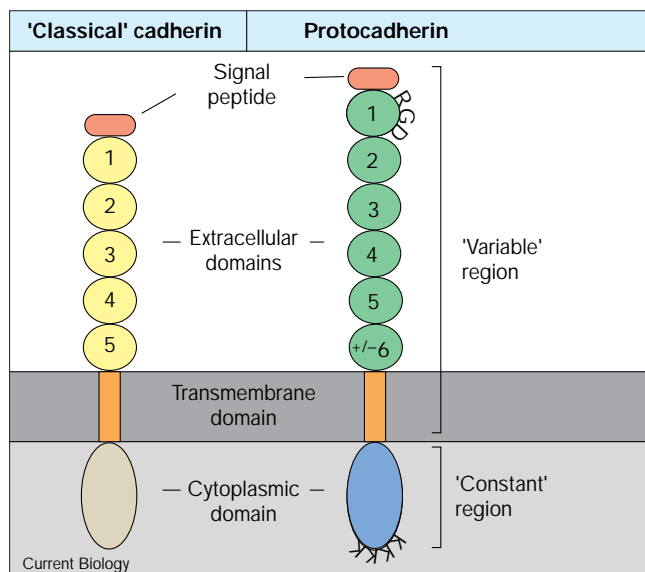
Protocadherins and cadherin-related neuronal receptors

The protocadherins are structurally related to the better-known ‘classical cadherins’ [9] (Figure 1). These are cell adhesion molecules that engage in calcium-dependent homophilic interactions, and have been shown to have important roles in development. They are oriented in the membrane with the amino-terminus outside the cell; the amino-terminal signal peptide is followed by a characteristic grouping of five tandem repeats, known as ‘ectodomains’, a single transmembrane region, and then an intracellular carboxy-terminal region that interacts with the cytoskeleton. Protocadherins have a similar extracellular organization, with five or six ectodomains, but they clearly differ in the cytoplasmic region.

A distinct subgroup related to the protocadherins consists of the cadherin-related neuronal receptors [7]. These all have six ectodomains — the first of which has an unusual RGD (putative integrin-binding) motif — and a distinctive cytoplasmic region that includes a characteristic lysine-rich segment. Notably, the cadherin-related neuronal receptors and some protocadherins show enriched or specific expression in the nervous system [7,10]; and like some of the classical cadherin subtypes, the cadherin-related neuronal receptors can also show differential expression in subpopulations of synapses [11].

The *Cnr* genes that encode the cadherin-related neuronal receptors were first cloned on the basis of their protein products’ ability to bind to the tyrosine kinase Fyn, a molecule that has been shown to function in both the immune system — where it is involved in signaling from the T-cell receptor [12] — and the nervous system — *fyn* deletion has been associated with behavioral deficits [13]. The mouse *Cnr* genes were initially found [7] to be part of a multi-gene family, with the additional interesting feature that when genomic DNA was hybridized with a 5′ *Cnr* cDNA probe, approximately 20 different genes were recognized, whereas a 3′ *Cnr* cDNA probe detected only three or so genes. This indicated that the 5′ coding portion of the *Cnr* cDNA may have originated from a diverse locus, whereas the 3′ portion had a simpler and clearly different genomic origin. This observation was further remarkable in that multiple *Cnr* genes appeared to reside in the same chromosomal region, suggesting a rather localized and perhaps tandem gene organization, and hinting at interesting mechanisms of gene regulation, likely including mRNA splicing or even DNA rearrangement.

Figure 1



An outline of the domain organization of cadherins (left) and protocadherins or cadherin-related neuronal receptors (right). (See text for details.)

Organization of human protocadherin genes

Wu and Maniatis [8] have now analysed the human counterpart of the protocadherin genes, including the *Cnrs*. By a combination of genomic database analyses and cDNA cloning, they identified a multigene family, subgroups of which encoded proteins homologous to the mouse protocadherins and cadherin-related neuronal receptors. They identified approximately 52 genes, spread over nearly 700 kilobases of the genome (Figure 2a). Within this locus, which was mapped to chromosome 5q31, three sub-families of 'protocadherin human' (*Pcdh*) genes — α , β and γ — were identified.

Most remarkable was the organization of genes within each subfamily. In the 5' upstream — part of each group, multiple, homologous 'variable' gene segments were found to be tandemly arrayed. The last of these was followed by three small exons encoding a single constant region, the cytoplasmic portion of the protein. The constant regions encoded by *Pcdh α* and *Pcdh γ* differ in sequence; for *Pcdh β* , the constant region exons have not yet been identified, but it seems likely that they will eventually be found. Each variable gene segment encoded the remaining — extracellular and transmembrane — domains of the protocadherin. At some level — RNA splicing or DNA recombination — the coding sequences of the variable segment and the constant region exons are joined, so that the final mRNA encodes a full protocadherin protein. This organization can explain why, in the mouse, a 5' *Cnr* probe recognized a much larger group of gene segments than a 3' *Cnr* probe. It is

quite different from the more conventional organization of the genes encoding the classical cadherins, such as P-cadherin or L-CAM [14].

Similarities to antigen receptor genes

This general organization of tandemly arranged gene segments is reminiscent of the distributed germline organization of antigen receptor genes (Figure 2b). So far, physiological somatic DNA rearrangements have been demonstrated only in the immune system, where they bring together the gene segments needed to produce mature genes coding for immunoglobulin heavy and light chains [15] and the T-cell receptor [16]. In the immune system, this process of recombination generates the huge diversity of receptors needed to guard against similarly diverse foreign antigens. Taking the immunoglobulin heavy chain locus as an example, the component segments — *V* ('variable'), *D* ('diversity') and *J* ('joining') — are brought together by specific '*V(D)J* recombination' reactions; multiple exons coding for the protein's constant region domains are located downstream of the *J* segments [17]. *V(D)J* recombination requires *cis* elements, known as 'recognition signal sequences', and the specific recombinase proteins, Rag1 and Rag2 [18].

A number of the key hallmarks of *V(D)J* recombination are clearly not shared by the protocadherin loci. The entire protocadherin variable region is encoded by a single exon, so any genetic rearrangement must involve joining of a variable gene segment to the constant region exons. In the case of the antigen receptor genes, this particular joining event involves either splicing or a non-site-specific DNA rearrangement (see below). Any sequences significantly similar to the recognition signal sequences of the antigen receptor genes should have been recognizable; Wu and Maniatis [8] did not find any such sequences within the protocadherin loci, though as they noted this does not rule out the presence of a different type of *cis*-acting recognition sequence for recombination.

The possibility that a site-specific DNA rearrangement of the kind that occurs in the immune system operates at the protocadherin loci is thus a remote one. As alternatives, Wu and Maniatis [8] have suggested three different types of RNA splicing mechanism that could bring together the protocadherin coding sequences. As an aside, it seems likely, given the availability of DNA sequence data and the sophistication of polymerase chain reaction (PCR) techniques, that any obvious rearrangement at the protocadherin loci would have already been identified. Given these considerations, are there any other aspects of protocadherin genes suggestive of DNA rearrangement?

Gene rearrangement?

Several intriguing aspects of the protocadherins remain for one to ponder. First, the cadherin-related neuronal

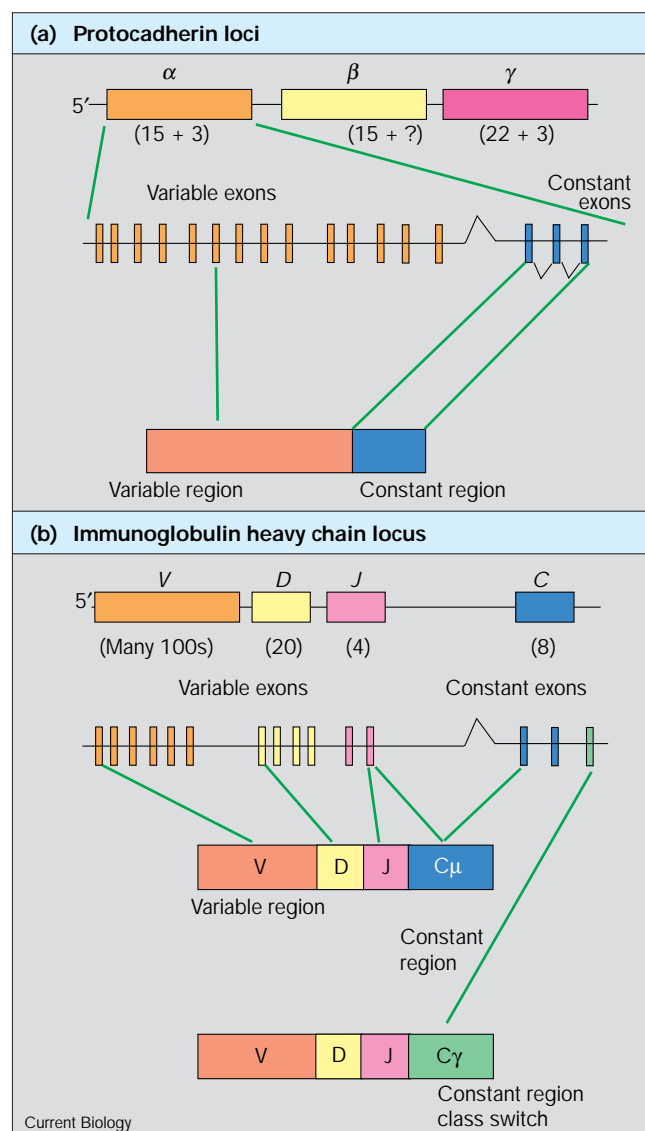
receptors — generally considered a subfamily of the protocadherins — were first identified by their ability to bind Fyn, which in the immune system interacts with the product of genes that clearly undergo DNA rearrangement — the T-cell receptor [12]. Second, earlier studies suggested that, if DNA rearrangement does occur in the nervous system, it would be of a distinct kind from $V(D)J$ recombination, as although Rag1 has been observed in neurons the requisite Rag2 has not [1,19,20]. And third, the genomic organization of the protocadherin loci — with a tandem array of intact variable exons, rather than multiple gene segments that require joining to create complete exons — itself suggests that any kind of DNA rearrangement would be significantly different from $V(D)J$ recombination.

Another type of DNA recombination that operates on the immunoglobulin heavy chain gene is that mediating ‘class switching’ [21]. Following $V(D)J$ recombination, production of a heavy chain protein involves RNA splicing that joins the variable region exons to either μ or δ constant region exons; subsequent maturation of the B cell bearing the immunoglobulin involves switching to production of a different type of heavy chain, typically an immunoglobulin G with a constant region encoded by γ exons. This class switching involves DNA rearrangement that is not site-specific. Another type of DNA rearrangement, known for many years as the basis of yeast mating-type switching, but likely to occur in other contexts (such as the chicken immune system), is gene conversion [22]. This process also is not site-specific, but exploits sequence similarity within tandem gene arrays — a condition that does exist at the protocadherin loci.

A very speculative possibility is that one or other, or both, of these mechanisms might be used to produce different combinations of variable and constant exons, and thereby recombining the extracellular adhesion and intracellular signaling specificities in different ways. In the nervous system, where neurons can be very long-lived, a genomic alteration may provide advantages over a reliance on non-genomic ways of controlling gene expression. Classical cadherins interact homophilically. How protocadherins interact *in vivo* is unclear, but combining extracellular and intracellular domains in different ways could achieve a substantial degree of combinatorial complexity that could be very useful for the nervous system.

All of these musings lead back to the original problem of identifying a rearrangement locus. It might prove fruitful to investigate whether non-homologous end joining is relevant to protocadherin expression, though it should be noted that deficiency of XRCC4 or DNA ligase IV [2,3] causes neuronal death and embryonic lethality, whereas deficiency of Fyn, which signals downstream of the cadherin-related neuronal receptors at least, causes non-lethal

Figure 2



A comparison of the genomic organization of (a) protocadherin loci [8] and (b) an immunoglobulin heavy chain locus. (See text for details.)

behavioral abnormalities [11]. One could, however, imagine that protocadherins interact with a range of other intracellular signaling molecules that can partially rescue the Fyn deficiency, but not the more pervasive defects caused by a loss of non-homologous end joining capacity. If protocadherin genes do undergo DNA rearrangement, it should not take long to demonstrate the fact. One should not be too surprised, however, if strange beasts still remain to be encountered.

Acknowledgements

I thank F. Alt, D. Schatz and T. Yagi for their insightful discussions. This review was written with support from the NIH and NIMH.

References

- Chun J, Schatz DG: Developmental neurobiology: alternative "ends" to a familiar story? *Curr Biol* 1999, 9:R251-R253.
- Gao Y, Sun Y, Frank KM, Dikkes P, Fujiwara Y, Seidl KJ, Sekiguchi JM, Rathbun GA, Swat W, Wang J, et al.: A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 1998, 95:891-902.
- Barnes DE, Stamp G, Rosewell I, Denzel A, Lindahl T: Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr Biol* 1998, 8:1395-1398.
- Dreyer WJ, Gray WR, Hood L: The genetic, molecular and cellular basis of antibody formation: some facts and a unifying hypothesis. *Cold Spring Harbor Symp Quant Biol* 1967, 32:353-367.
- Obata S, Sago H, Mori N, Rochelle JM, Seldin MF, StJohn T, Taketani S, Suzuki ST: Protocadherin Pcdh2 shows properties similar to, but distinct from, those of classical cadherins. *J Cell Sci* 1995, 108:3765-3773.
- Sago H, Kitagawa M, Obata S, Mori N, Taketani S, Rochelle JM, Seldin MF, Davidson M, StJohn T, Suzuki ST: Cloning, expression, and chromosomal localization of a novel cadherin-related protein, protocadherin-3. *Genomics* 1995, 29:631-640.
- Kohmura N, Senzaki K, Hamada S, Kai N, Yasuda R, Watanabe M, Ishii H, Yasuda M, Mishina M, Yagi T: Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron* 1998, 20:1137-1151.
- Wu Q, Maniatis T: A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* 1999, 97:779-790.
- Takeichi M: Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* 1991, 251:1451-1455.
- Hirano S, Yan Q, Suzuki ST: Expression of a novel protocadherin, OL-protocadherin, in a subset of functional systems of the developing mouse brain. *J Neurosci* 1999, 19:995-1005.
- Yagi T: Molecular mechanisms of Fyn-tyrosine kinase for regulating mammalian behaviors and ethanol sensitivity. *Biochem Pharmacol* 1999, 57:845-850.
- Samelson LE, Phillips AF, Luong ET, Klausner RD: Association of the fyn protein-tyrosine kinase with the T-cell antigen receptor. *Proc Natl Acad Sci USA* 1990, 87:4358-4362.
- Grant SG, O'Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER: Impaired long-term potentiation, spatial learning, and hippocampal development in Fyn mutant mice. *Science* 1992, 258:1903-1910.
- Hatta M, Miyatani S, Copeland NG, Gilbert DJ, Jenkins NA, Takeichi M: Genomic organization and chromosomal mapping of the mouse P-cadherin gene. *Nucleic Acids Res* 1991, 19:4437-4441.
- Hozumi N, Tonegawa S: Evidence for somatic rearrangements of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci USA* 1976, 73:3628-3632.
- Hedrick SM, Cohen DI, Nielsen EA, Davis MM: Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature* 1984, 308:149-153.
- Alt F, Yancopoulos GD, Blackwell TK, Wood C, Thomas E, Boss M, Coffman R, Rosenberg N, Tonegawa S, Baltimore D: Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J* 1984, 3:1209-1219.
- Schatz DG: V(D)J recombination moves *in vitro*. *Semin Immunol* 1997, 9:149-159.
- Chun JJ, Schatz DG, Oettinger MA, Jaenisch R, Baltimore D: The recombination activating gene-1 (RAG-1) transcript is present in the murine central nervous system. *Cell* 1991, 64:189-200.
- Chun J, Schatz D: Rearranging views on neurogenesis: neuronal death in the absence of DNA end-joining proteins. *Neuron* 1999, 22:7-10.
- Lutzker SG, Alt FW: Immunoglobulin heavy-chain class switching. In *Mobile DNA*. Edited by Berg DE, Howe MM. Washington, DC: American Society for Microbiology; 1989:693-714.
- Carlson LM, Oettinger MA, Schatz DG, Masteller EL, Hurley EA, McCormack WT, Baltimore D, Thompson CB: Selective expression of RAG-2 in chicken B cells undergoing immunoglobulin gene conversion. *Cell* 1991, 64:201-208.

If you found this dispatch interesting, you might also want to read the August 1999 issue of

Current Opinion in Genetics & Development

which included the following reviews, edited by Norbert Perrimon and Claudio Stern, on Pattern formation and developmental mechanisms:

Cell polarity in the early *Caenorhabditis elegans* embryo

Bruce Bowerman and Christopher A Shelton

The polarisation of the anterior-posterior and dorsal-ventral axes during *Drosophila* oogenesis

Fredericus van Eeden and Daniel St Johnston

Wnt signaling and dorso-ventral axis specification in vertebrates

Sergei Y Sokol

Establishment of anterior-posterior polarity in avian embryos

Rosemary F Bachvarova

Polarity in early mammalian development

Richard L Gardner

Diverse initiation in a conserved left-right pathway?

H Joseph Yost

Extracellular modulation of the Hedgehog, Wnt, and TGF- β signalling pathways during embryonic development

Javier Capdevila and Juan Carlos Izpisua Belmonte

Fringe, Notch, and making developmental boundaries

Kenneth D Irvine

Polarity determination in the *Drosophila* eye

Helen Strutt and David Strutt

Wnt signalling: pathway or network?

Alfonso Martinez Arias, Anthony MC Brown and Keith Brennan

Epithelial cell movements and interactions in limb, neural crest and vasculature

Cheryll Tickle and Muriel Altabel

Cell movement in the sea urchin embryo

Charles A Etensohn

Cell migration in *Drosophila*

Alexandria Forbes and Ruth Lehmann

The full text of *Current Opinion in Genetics & Development* is in the BioMedNet library at

<http://BioMedNet.com/cbiology/gen>