Review

Toward a molecular catalogue of synapses

Seth G.N. Grant⁎

Genes to Cognition Programme, The Wellcome Trust Sanger Institute, Hinxton, Cambridge, Cambridgeshire CB10 1SA, UK

Abstract

1906 was a landmark year in the history of the study of the nervous system, most notably for the first ‘neuroscience’ Nobel prize given to the anatomists Ramon Y Cajal and Camillo Golgi. 1906 is less well known for another event, also of great significance for neuroscience, namely the publication of Charles Sherrington’s book ‘The Integrative Action of the Nervous system’. It was Cajal and Golgi who debated the anatomical evidence for the synapse and it was Sherrington who laid its foundation in electrophysiological function. In tribute to these pioneers in synaptic biology, this article will address the issue of synapse diversity from the molecular point of view. In particular I will reflect upon efforts to obtain a complete molecular characterisation of the synapse and the unexpectedly high degree of molecular complexity found within it. A case will be made for developing approaches that can be used to generate a general catalogue of synapse types based on molecular markers, which should have wide application.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Synapse Proteomic

1. Neuron diversity

The development of microscopy and novel cell staining methods during the 19th century led to the all important cell theory and the neuron doctrine (Finger, 1994) (and other papers in this issue). The extraordinary diversity of neuronal architecture revealed by Cajal’s illustrations led to the realisation that the brain comprises a vast array of different neurons, which was in stark contrast to some simpler organs, such as the liver, with a fairly small number of similar shaped cells. This diversity of neuronal types and their connections remain of great interest and subject to new molecular and imaging approaches.

The realisation that it may be technically possible to define the complete circuitry and neuronal morphology of a mam-
malian brain has been suggested by the success from a simpler, and thus more tractable model organism, the worm *C. elegans* (Wood, 1987). Comprehensive electron microscopic images as well as transgenic fluorescent labelling methods were used to define the development and organisation of the nervous system. The long-term ambition of creating a definitive map of neural connections in a mammal is underway with electron microscopic strategies being developed to map the connectivity in the rodent brain (Denk and Horstmann, 2004).

Morphological descriptions of neuronal circuitry will not be sufficient in the longer term, and it will be necessary to include molecular information for the particular neurons. There are several large scale programs that have produced valuable data on the molecular composition of the mouse brain using complementary methods and are reviewed elsewhere (Anderson and Grant, 2006; Hatten and Heintz, 2005; Sunkin, 2006). These programs include the Allan Brain Atlas (www.brainatlas.org) that has completed in situ hybridisations of 16,000 genes in the adult mouse brain (Wong et al., 2007); the Brain Gene Expression Map (Magdaleno et al., 2006) that includes adult and developing mouse brain; GENSAT (www.gensat.org) (Gong et al., 2003) that uses bacterial artificial chromosome (BAC) based transgenesis to drive expression of green fluorescent protein (GFP) into hundreds of different neuronal populations. The GENSAT approach has the advantage that live cells can be visualised using GFP and targeted for electrophysiological recordings. Moreover, the BAC transgenic lines can be used for further studies expressing other kinds of markers or proteins that alter neuronal function.

These atlas programs generally have the feature of identifying a single gene in an individual neuron. They are limited by the number of molecules that can be detected simultaneously in any given neuron. For example, the in situ hybridisation is currently useful for only one gene at a time although techniques are in development to increase this to several probes. The BAC transgenesis approach is similarly limited to single or few genes as it requires interbreeding of lines of reporter mice. Complementary to these approaches are single cell methods, which have limited throughput in terms of numbers of cells that can be assayed, but have much higher coverage of mRNAs in that neuron. Single neuron patch-clamping methods that serve both to record the electrical profile of the neuron as well as remove the mRNA for RT-PCR (Reverse Transcriptase polymerase chain reaction) amplification have been used to profile dozens of mRNAs and correlate it with the properties of the nerve cells (Toledo-Rodriguez et al., 2005).

It is clear that morphologically distinct neurons express distinct sets and combinations of genes and these combinations form a signature that identifies that neuron. An important insight that has emerged from these gene expression studies is that there are relatively few genes that have a highly restricted pattern of gene expression (Wong et al., 2007). Several years ago it was hoped that each morphologically distinct neuron would express a unique gene. This would make transgenesis experiments very simple, as mice could be created that would use this unique gene promoter and enhancer sequences to express reporters or biologically active molecules in that neuron type. It appears that a given neuron is defined not by single genes but by the combination of genes expressed within it. An important functional corollary is that the properties of the neuron are indeed the sum of a common mixture of ingredients and it is the proportions that define the neuronal properties (Magdaleno et al., 2006). In addition, this molecular dissection has been shown to distinguish neurons that are morphologically indistinguishable (Wong et al., 2007) and thereby raise the possibility that there are more types of neurons to be described than can be uncovered with anatomical methods alone. It is also worth pointing out that two neurons might have identical molecular marker profiles and could show different anatomies or connectivities, a problem that further reinforces the value of combining morphology with molecular markers.

A further level of molecular characterisation, which is outside the scope of the gene expression approaches, is the subcellular neuronal architecture and the pattern of protein expression in the different parts of the neuron. The striking diversity of dendritic architectures of different classes of neurons has become intensively studied at the electrophysiological level with patch methods in recent years (Davie et al., 2006). Dendrites and their branches are considered to be functional units for processing information, whether by modifying post synaptic potentials or for translating synaptic proteins (Stuart et al., 1999). It is clear that some proteins are not equally distributed across the dendritic tree (Lai and Jan, 2006; Lorincz et al., 2002; Sheng et al., 1992; Varga et al., 2000). The problems of studying the distribution of dendritic proteins in a quantitative manner over a single neuron have not been resolved or shown to be scalable to large numbers of proteins.

### 2. Developing a synapse signature of molecular markers

The computational functions of the brain are conducted at many ‘anatomical’ levels: brain regions and local circuits, and as mentioned in the previous section, also at the level of dendrites and their branches. It is now well established that the molecular circuitry within synapses plays an important role in computational functions. For example, synaptic plasticity and signalling to the nucleus require particular sets of synaptic proteins. It is therefore necessary to consider the need to define the number of synapse types and their respective functions.

Unlike the diversity of anatomically distinct neurons, much less is known about the diversity of synapses (Wang et al., 2006). Anatomical methods have described dendritic spines and synapses of different shapes and lengths (Harris, 1999). A molecular marker based system of synapse categorisation would be complementary to a morphological approach. Clearly the pharmacological definitions based on neurotransmitter receptor specificities (e.g. glutamatergic, GABAergic) were a valuable step in this direction. In addition to these neurotransmitter based markers, there are hundreds of other synaptic proteins that could contribute to the cataloguing system. A signature for a given synapse based on a combination of molecular markers could be used to develop a synapse catalogue.
Despite a century of research on the morphology and function of the synapse, a complete list of synapse protein components is still unavailable. Estimates of the number of protein components have been made from proteomic studies of the postsynaptic terminal of rodent forebrain synapses (Collins et al., 2005; Husi et al., 2000; Jordan et al., 2004; Li et al., 2004; Peng et al., 2004; Wallkonis et al., 2000; Yamauchi, 2002; Yoshimura et al., 2004; Cheng et al., 2006; Okabe, 2007). In these experiments, the postsynaptic density and other postsynaptic neurotransmitter receptor complexes (together referred to as the postsynaptic proteome) were biochemically isolated and their components identified using mass spectrometry and antibody detection methods. The number of proteins described in these studies exceeds 1000, which was much more than expected (Grant, 2006). This information is a useful resource for building a catalogue of synapses that can be defined according to their expression profiles of these proteins. It is already clear that not all synapse proteins are expressed equally at all synapses as some of the expression patterns in different classes of neurons are quite distinct.

Like neuronal patterns of gene expression, many differences between synapses will likely reflect different combinations of protein expression. These combinations can be used as a signature through the use of clustering algorithms (Zapala et al., 2005). To illustrate this, consider a simple model where the postsynaptic compartment of the synapse comprises 4 proteins (A, B, C, D) that can be expressed at 2 different levels (1,2 and indicated as A1 or A2) and these levels are measured in a CA1 pyramidal cell and a dentate gyrus granule neuron. For these two cells there may exist a signature such as A1B1C2D2 for CA1 neurons and A2B2C2D1 for dentate gyrus neurons. This type of ‘barcode’ can be extended to more than 2 levels of expression (4 is feasible) and to hundreds of proteins with the result that an enormous number (~10^15) of potentially different synaptic signatures can be defined.

The main technical obstacle to a synapse catalogue is the quantitative measurement of protein levels at individual synapses. Several very elegant studies have quantified specific proteins at individual synapses (Sugiyama et al., 2005; Masugi-Tokita et al., 2007). Currently, it is difficult to see how these approaches could be applied to all synapse of a given neuron or all synapses in a piece of brain tissue. Before considering tackling this problem on a sample that is complex and heterogeneous, such as a piece of tissue, it would be better to focus on the simplest possible system. This could be a preparation of purified synapses (synaptosomes) that have identical proteomes and is available in limitless amounts. This is analogous to a clonal population of uniform cells and here will be referred to as the ‘clonal synapse’. Two different clonal synapse preparations would show a different signature of protein marker expression (e.g. a glutamatergic and a GABA synapse).

The advantage of the clonal synapse is that existing mass spectrometry based proteomic methods could identify and quantify all proteins, splice variants and posttranslational modifications. In other words, a comprehensive pattern of marker expression. This pattern of expression would then be confirmed using single synapse staining of as many markers as possible to confirm their homogeneity. A preparation of such homogeneous synapses would be an excellent model system for a range of fundamental studies of the synapse.

How might clonal synapses be prepared? There are no established methods although it could be developed along the following lines. The first is to purify a population of neurons that is as uniform as possible and then from those neurons isolate their synaptosomes. The second approach is to clonally expand a subtype of neuron from neuronal stem cells or embryonic stem cells (Trounson, 2006). The caveat of these methods is that the population of synapses in a given nerve cell are not homogeneous, which is well known for many neurons. A simple example would be a rodent hippocampus CA1 pyramidal neuron that has both excitatory and inhibitory synapses. Thus a cell purification or cloning method may be insufficient and might only comprise a first step. Another approach may be to create transgenic mice that have synapses tagged with affinity labels that could allow purification of the synapses. For example a particular class of ion channel or extracellular receptor could be tagged so that synaptosomes in suspension could be sorted.

The above discussion is based on a simple model where synapses are composed of different sets of proteins and these represent markers for the types of synapses. Although this is a useful framework, at some stage it will be important to consider the dynamics of proteins. The activation of a receptor could lead to degradation or synthesis of new synaptic proteins and thus produce diversity in the population of synapses as a result of their activation state. Given that this is known to occur, this only increases the potential for synapse diversity in the nervous system.

3. Applications of a synapse catalogue and cataloguing techniques

There are a wide range of applications for a synapse catalogue based on synapse proteome markers. It would be of interest to simply define the diversity of synapses on a given neuron, and then brain region and ultimately the entire brain. One exciting outcome of this investigation might be to ask for humans, which have an estimated 10^14–15 synapses: how many types (as defined by their synapse proteome signatures) are there? In line with the earlier discussion on the combinations of protein levels and the concept of a barcode, it is not inconceivable that simply varying the levels of synaptic proteins could produce more synapse proteome combinations than there are synapses in the human brain.

Examining the synapse types in different species would also be of great interest. Do different species have synapses with the equivalent composition? Although we know that mammalian species share similar receptors and some other proteins, there has not been a comparative proteomic profile of synapses between species. It remains possible that some human synapses are quite different to other species and this would have functional consequences. Comparative methods could be used to examine those parts of the nervous system that have expanded or become anatomically specialised in different species.

Synapse cataloguing tools could be used to examine the dynamics of synapses. How much does the synapse
composition change with various forms of stimulation, behaviour or developmental states? Do all synaptic proteins turnover with a similar time constant? What are the different molecular mechanisms regulating the turnover of different classes of proteins? These questions are very actively being studied using a variety of methods. For example, differential quantitative protein expression and morphological studies including single molecule labelling techniques described dynamics of shape and composition and proteins that change (Chen et al., 2005; Ehlers, 2003; Ebihara et al., 2003; Inoue and Okabe, 2003; Kurisu et al., 2006; Petersen et al., 2003). These approaches form a basis for larger scale systematic surveys of synapses across the nervous system. A recent study of the global protein composition of synaptic vesicles shows that calculation of numbers of proteins and their organisation into a subcellular organelle can be achieved and provide a valuable model (Takamori et al., 2006).

Of great interest is the issue of diseases of the synapse (synapsopathies). Cataloguing tools could be used to identify the subtypes of synapses affected during the course of disease processes. In neurodegenerative diseases it is well known that specific regions of the brain deteriorate before others and given that these parts of the brain have distinct synaptic signatures (or barcode of expression) then these changes could be detected using cataloguing tools, perhaps leading to new diagnostics.

Returning to Sherrington and electrophysiology; functional studies of synapses in experimental situations are essential for understanding how synapses of different composition work. It is well known that different receptors and downstream signalling proteins provide different signalling properties as revealed by synaptic plasticity studies. A molecular catalogue could be an important starting point for systematic electrophysiological studies in that different synapses could be identified by molecular tools and then selected for electrophysiological study.

4. Functional significance of different synapse compositions

The levels of expression of individual synapse proteins are critical to the function of the synapse. This is very clearly seen in studies of ∼100 mouse knockouts where the reduction of expression of a single synaptic protein leads to changes in synaptic physiology (see G2Cdb database for repository of synaptic physiology phenotypes: www.genes2cognition.org/db). The synapse may be rather sensitive to these changes in levels of expression since heterozygous mice with 50% changes in protein expression show differences in synaptic physiology. It is much more difficult to predict the functional differences between two (or more) synapses where many proteins are changed in levels. In other words, the functional effects may reflect the combinational effects of several proteins that work together.

Molecular network studies of the postsynaptic proteome may be useful in interpreting the effects of differences in multiple proteins (Pocklington et al., 2006). Pocklington and colleagues have constructed protein interaction maps of glutamate receptor associated proteins and describe modules or groupings of proteins that have common functional properties. They also modeled the effects of different levels of protein expression that occurs in disease states and suggest that there are different distributions for certain types of diseases. Not all proteins were of equivalent functional significance with respect to synaptic plasticity; some mutations have small and others large effects on long-term potentiation. Thus the levels of expression found in synapse cataloguing could be mapped onto synapse protein–protein interaction maps and used to map onto functional properties. It may be that those proteins (or classes of proteins) that vary the most between synapses fall into different functional groups with different distributions on the network.

5. Concluding comments

We are indebted to Cajal and Golgi for revealing the extraordinary and complex world of neuronal diversity. Molecular techniques in the realm of gene expression and proteomic methods are beginning to follow Cajal’s systematic histological methods with systematic molecular profiling of neurons and synapses. The definition of the molecular expression profile for all neurons and synapses is within reach and will provide a definitive molecular composition of the brain.

Many scientists believe that systematic science is dull and uninteresting and that only a hypothesis lead research (sometimes referred to as discovery science) is important. Anyone familiar with the work of Cajal, would certainly place his extensive cataloguing of neuronal morphology into the category of systematic science. This would also be said of Sulston, White and Brenner who followed in his footsteps and mapped the nervous system of the worm. Profound new insights that have shaped all biology have emerged from these systematic studies. Systematic large scale studies are an important area of future research and can make a major contribution toward our understanding of the brain. Systematic science and discovery science are not separate as each feeds the other. There is an important need to specifically fund and develop large scale neuroscience programs that link molecular, anatomical and functional data sets.

Finally, the brain is the most complex organ and requires systematic studies to simplify this complexity. Experimentally aligning the wonderful anatomical diversity of Cajal’s neurons with the molecular diversity of synapse proteomes remains as one challenge for the future.

Acknowledgments

SG is supported by the Wellcome Trust. Thanks to J. Turner for manuscript preparation.

References

Chen, X., Vinade, L., Leapman, R.D., Petersen, J.D., Nakagawa, T.,


