Anatomical funneling, sparse connectivity and redundancy reduction in the neural networks of the basal ganglia

Genela Morris, Alon Nevet, Hagai Bergman *

Department of Physiology, the Interdisciplinary Center for Neural Computation and the Eric Roland Center for Neurodegenerative Diseases, Hadassah Medical School, The Hebrew University, P.O. Box 12272, Jerusalem 91120, Israel

Abstract

The major anatomical characteristics of the main axis of the basal ganglia are: (1) Numerical reduction in the number of neurons across layers of the feed-forward network, (2) lateral inhibitory connections within the layers, and (3) neuro-modulatory effects of dopamine and acetylcholine, both on the basal ganglia neurons and on the efficacy of information transmission along the basal ganglia axis. We recorded the simultaneous activity of neurons in the output stages of the basal ganglia as well as the activity of dopaminergic and cholinergic neurons during the performance of a probability decision-making task. We found that the functional messages of the cholinergic and dopaminergic neurons differ, and that the cholinergic message is less specific than that of the dopaminergic neurons. The output stage of the basal ganglia showed uncorrelated neuronal activity. We conclude that despite the huge numerical reduction from the cortex to the output nuclei of the basal ganglia, the activity of these nuclei represents an optimally compressed (uncorrelated) version of distinctive features of cortical information.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Striatum; Globus pallidus; Acetylcholine; Dopamine; Parkinson’s disease

1. Introduction: main anatomical constraints

Information processing in the brain is bounded by the underlying anatomical substrate. Within the limits set by anatomy, the actual physiological parameters (e.g., firing rate, patterns and synchronization among groups of neurons) set the modus operandi for the computational processes. In this manuscript we will use the results of recent anatomical and physiological investigations of the basal ganglia to shed light on the possible computational processes and tactics used by these structures.

The detailed anatomy of the basal ganglia (see reviews in [1–3]) is beyond the scope of this manuscript. Moreover, we fear that as is often the case in discussion of anatomical detail, it may cloud the overall view of the circuit. From a broad perspective, the striatum (input stage of the basal ganglia) receives excitatory projections from most cortical areas as well as from several thalamic nuclei [4]. Subsequent direct and indirect projections link striatal neurons to the output stage of the basal ganglia, i.e., the internal segment of the globus pallidus (GPI) and the substantia nigra pars reticulata (SNr). The GABAAergic inhibitory projections of GPI and SNr neurons control the activity of the excitatory thalamo–cortical connections (Fig. 1). In this report we will only consider the “main axis” of the basal ganglia: cortex, striatum and GPI/SNr. Each level of this cortex–striatum–GPI/SNr pathway is characterized by a high degree of numerical reduction in the number of neurons. The number of striatal neurons is one to two orders of magnitude smaller than that of cortical neurons projecting to the striatum [5], and an additional reduction of the same magnitude occurs from the striatum to the GPI/SNr [6–8]. Most anatomical studies concur that this axis is in fact comprised of a number of domains [2,9]. However, the degree to which the different domains overlap is still under debate [10–15]. The feed-forward frame of the cortex–striatum–GPI/SNr is complicated by lateral connectivity, as well as by the action of neuro-modulatory substances. Most striatal and pallidal neurons form massive collateral GABAAergic connections within their nuclei of origin [16]. Furthermore, the collateral inhibitory system of the striatum is augmented by
the parvalbumin positive GABAergic interneurons [3,17]. Finally, dopamine (mainly from the substantia nigra pars compacta, SNC) and intrinsic acetylcholine modulate both the activity of striatal neurons (and to a lesser degree also of GPi/SNr) [18] and the efficacy of cortico-striatal transmission [19,20].

To summarize, our aim is to provide physiological data and perhaps some insight into the computational processes carried out by the main axis (cortex, striatum, GPi/SNr) of the basal ganglia. To do so, we will assume the following anatomical characteristics of this axis: (1) Feed forward connectivity between the structures. (2) Considerable reduction in the number of neurons along the axis. (3) Lateral inhibitory connections in the striatum and (4) Modulation of cortico-striatal transmission by acetylcholine and dopamine.

2. Materials and methods

2.1. Animals, behavioral task and surgical procedures

Three macaque (Macaca fasicularis) monkeys were trained to perform a self initiated probabilistic delayed visual-motor task, in which the probability of receiving reinforcement for correct performance depended on the presented visual cue. During all training and recording sessions monkeys were seated across of a screen with a panel consisting of three keys in front of them. Trials were initiated when the monkey touched the central key. After a variable delay (1.5–2.5 s in monkeys C and E; 2–4 s in monkey Y), a visual cue appeared for a short period (0.3 in monkeys C and E; 0.45 s in monkey Y) on a randomly chosen side of the screen. The monkeys were well acquainted with a set of five possible cues. Each cue was associated with a different probability of reward (0, 0.25, 0.5, 0.75 and 1.0). The cue presentation was followed by a fixed hold period of 2 s, after which a go signal appeared, upon which the monkeys were required to press either the left or right key, according to the location of the memorized cue. Correct response was followed (with an interval of 700 ms) by liquid reward at the probability associated with the visual cue. All trials (correct, incorrect, rewarded and unrewarded) were followed by a variable inter-trial-interval (3–6 s in monkeys E and C; 5–7 s in monkey Y).

After training, a square recording chamber with a 27 mm (inner) side was attached to the skull to allow access to the basal ganglia targets. The recording chamber was tilted 50° laterally in the coronal plane, with its center targeted at stereotaxic coordinates of the GPi or the SNr. The chamber’s coordinates were adjusted according to MRI imaging (Biospec Bruker 4.7 T animal system, fast-spin echo sequence; effective TE = 80 ms and TR = 2.5 s, 13 coronal slices 2 mm wide). All surgical and MRI procedures were performed under general and deep anesthesia. The monkeys’ care and surgical procedures were in accordance with the NIH Guide for the Care and Use of Laboratory Animals (1996), and with the Hebrew University guidelines for the use and care of laboratory animals in research, supervised by the institutional animal care and use committee.

2.2. Data acquisition and analysis

During recording sessions the monkeys’ heads were immobilized, and eight glass-coated tungsten micro-electrodes (impedance 0.3–1.2 MΩ at 1000 Hz), confined...
within a cylindrical guide (1.65 mm inner diameter), were advanced separately (EPS, Alpha-Omega Engineering, Nazareth, Israel) into the recording targets (Fig. 1). The signal from the electrodes was amplified with a gain of 10 K and band-pass filtered with a 300–6000 Hz 4-pole Butterworth filter (MCP+, Alpha-Omega Engineering, Nazareth, Israel). This electrical activity was sorted and classified on-line using a template-matching algorithm (MSD, Alpha-Omega Engineering, Nazareth, Israel). The sampling rate of spike detection pulses and behavioral events was 12 kHz (AlphaMap, Alpha-Omega Engineering, Nazareth, Israel). Eye movements were recorded with monocular infrared coulometer (Dr. Bouis, Karlsruhe, Germany) and sampled at 0.8 kHz. The analog output of all electrodes was continuously sampled at 24 kHz after 5 K amplification and 1–6000 Hz band-pass filtering (Figs. 2 and 4).

Only spike trains considered during real-time sorting to be emitted by a single cell were subjected to rate stability analysis. In the rate stability analysis a smoothed estimate of the instantaneous rate of a neuron as a function of time was displayed for the entire period of recording, and the largest segment of stable data was selected for further analysis. Subsequently, the cells’ responses to behavioral events, autocorrelograms and pairwise cross-correlograms were calculated. The cross-correlograms were calculated only for pairs of cells recorded by different electrodes to avoid artifacts due to sorting shading [21].

3. Results

3.1. Correlation studies of the spiking activity in the output stages of the basal ganglia

Taking into account the abundant lateral inhibitory connections in the striatum one might expect strong lateral physiological interactions between striatal neurons. This prediction, however, has not been supported by previous physiological intra-cellular studies. No evidence has been shown for functional synaptic interactions between striatal projection neurons [22]. Recent in vitro studies, using spike triggered average techniques, revealed a small number of weak inhibitory connections between striatal projection neurons, and those found were all uni-directional [23].

It is extremely difficult to perform intra-cellular studies in the awake behaving animal. Therefore, we used multiple extra-cellular electrodes to record the simultaneous spiking activity of several neurons in the output stages of the basal ganglia. Our working assumption is that functional lateral inhibitory connections should result in troughs in the cross-correlation functions. Functional connectivity in the form of striatal or pallidal functional assemblies, created by shared inputs, should be evidenced as positive peaks in the cross-correlation functions of neurons in the respective nuclei. Previous studies have failed to reveal correlations between the spiking activity of simultaneously recorded pallidal neurons [24,25]. It should be noted, however, that the lateral connectivity, as well as the number of interneurons and the direct dopamine effects in the SNr show distinct differences to those observed in the GPi [26,27]. We therefore set out to study the correlation between simultaneously recorded SNr neurons (Figs. 1 and 2). Our preliminary analysis reveals that as with GPi, the pairwise correlation of SNr neurons is relatively flat (Fig. 3). Such correlation functions suggest that the amount of functional connectivity within the SNr is minimal, as is the degree of convergence of common input to SNr neurons.
3.2. Teaching signals in the basal ganglia

It has consistently been shown that the striatum shows the densest tracing in the central nervous system for dopaminergic [1] and cholinergic markers [28]. While most of the brain dopamine is generated by midbrain dopaminergic neurons and merely projected in the striatum, striatal acetylcholine is probably generated by the striatal cholinergic interneurons. There are several studies that indicate that the tonically active neurons (TANs) of the striatum are the cholinergic interneurons of the striatum [29–31].

The central role of dopamine and acetylcholine in control of motivation and learning has been known for many years [32]. Recent studies [33,34] revealed that the dopaminergic signal is best characterized as related to the discrepancy between the animal’s predictions and reality [35]. Thus, dopamine neurons respond only to the first cue in a trial that predicts reward. However, they do not respond to aversive stimuli [36].

Initial physiological studies of TANs revealed similar properties to those of the dopaminergic neurons, with inverse polarity of responses [37]. However, several recent studies indicate that the functional message of the cholinergic system may be different from that of the dopaminergic system. Thus, the TANs show robust responses to aversive stimuli [38] and respond to more than one event in a trial [39].

In our recordings, most TANs exhibited a stereotypical response to the visual cue (a pause in firing, flanked on both sides by a brief elevation of the firing rate). The same cells displayed a similar response to the reward (Figs. 4 and 5). The TANs’ response was not significantly modified by the different cues (right or left side and different probabilities for future reward). Thus, unlike dopaminergic neurons that code for the difference between the animal’s prediction and reality [35] the TANs provide a more general message, probably indicating, or alternatively instructing the occurrence of an attention shift to salient events.

4. Discussion

4.1. Reinforcement learning models of the basal ganglia

The “actor–critic” architecture is an abstract learning system [40], in which there is an “actor”, acting in a certain environment, and a “critic”, providing reinforcing signals which are used by the “actor” in order to maximize the weighted sum of all future reinforcement values.

The apparent activity in the basal ganglia teaching systems resembles that of the “critic” in reinforcement temporal delay learning models [41]. Therefore, in computational models of the basal ganglia [33,42–44], the cortex–striatum–GPi/SNr axis is frequently modeled as the “actor”, while the dopaminergic (and cholinergic) neurons are presented as the “critic” or the provider of an error signal in a learning network.

Actor–critic network models postulate that the teaching signal will modulate synaptic transmission in the actor. Indeed, it has been shown that plastic changes in the morphology of striatal synapses occur after dopamine depletion [45]. Physiological studies show that the dopaminergic [46,47] and the cholinergic [48] signals modulate the cortical input to striatal projection neurons. Moreover, as predicted by reinforcement learning models, striatal and pallidal neurons significantly change their discharge as a function of the prediction of future reward [49–51]. Furthermore, their discharge patterns vary significantly during different phases of learning [52].
The balance between the dopaminergic and the cholinergic messages in the basal ganglia has been studied extensively since the discovery of the beneficial therapeutic effects of anti-cholinergic drugs for parkinsonian patients [53]. Anatomical studies of the striatum revealed direct contact of dopamine terminals and striatal cholinergic neurons [54]. Furthermore, numerous neurochemical studies demonstrated that dopamine application inhibits acetylcholine release within the striatum. Intriguingly, it has recently been demonstrated that acetylcholine has opposing effects on the release of dopamine in the striatum [55]. Experimental procedures that abolish dopamine input to the striatum (Haloperidol, MPTP [56,57]; local application of D2 antagonists [58]) abolish the TANs’ response. However, re-establishment of the TAN activity by treatment with non specific (post-synaptic) dopamine replacement therapy [56] indicates that the dopaminergic system enables (but does not drive) the TANs to express their stereotypical responses.

Our results show that the functional message of the cholinergic system is not identical to that of the dopaminergic system. Unlike the dopaminergic activity, which is proportional to the difference between prediction and reality, the TAN message does not depend on the probability of future reward. TANs show a similar response for both unpredicted positive and negative (disappointment) events (data not shown). Further studies are needed to establish whether the dopaminergic and cholinergic projections converge on the same population of striatal neurons or rather on separate ones, and to establish their specific role in controlling the local learning rules in the cortico-striatal synapses.

4.2. Redundancy reduction neuronal networks

The complex anatomical and physiological setting of the “actor” elements in the basal ganglia (i.e., the main axis) can be combined within a computational model of local competitive learning rules [59], controlled by reinforcement or critic signals [60–62]. The models assume that the basal ganglia perform efficient dimensionality (redundancy) reduction [59,60,63–65] and decorrelation of the large information space spanned by the activity of the cortical–striatal neurons. It has been proven that neural networks can perform such efficient coding using local cellular competitive learning rules [63]. Most sensory systems that have been shown to perform dimensionality reduction [66], do so solely in relation to the input statistics. Due to the added value of the “critic” elements in the basal ganglia, dimensionality reduction performed in this structure is a function not only of the statistical properties of the cortical (input) patterns, but also of their behavioral significance. This is achieved by a triple striatal synapse, in which the teaching (dopaminergic or cholinergic) signal controls the feed-forward cortico-striatal (and striatal–pallidal) Hebbian learning. Thus, decorrelation of basal ganglia activity is achieved by a dynamic process rather than by fixed cortico-striatal–GPi/SNr connectivity. More
importantly, this network enables a discriminative redundancy reduction process, performing better for reward related inputs than for irrelevant events.

4.3. Sparse connectivity and information in the basal ganglia axis

Alternatively to the dimensionality reduction models, the apparent lack of temporal correlation between the neuronal activity within the GPi and SNr could be accounted for by sparse cortico-striatal and striatal-pallidal connectivity. Recent studies indicate that the anatomy of the cortico-striatal pathways is patchy and discontinuous, and that individual cortical foci give rise to multiple and separated sites of striatal innervation [15,37,67]. Quantitative analysis of single neuron tracing reveals that the degree to which cortical input is shared by nearby striatal neurons is low [5,68,69]. Wilson and colleagues have shown that the dendritic field of striatal projection neurons (which is, incidentally, of an equal size to that of the focal axonal arborization of cortico-striatal axons) contains $4 \times 10^7$ cortico-striatal axons and 2850 other projection neurons. Each striatal projection neuron receives 5300 cortico-striatal synapses and therefore shares only about 70 synapses with any given striatal neuron within this small (0.4 mm diameter) volume.

The morphology of pallidal neurons is very different from that of the striatal neurons. In contrast to the spherical shape of the dendritic field of a typical striatal neuron, the dendritic field of pallidal neurons resembles a disk oriented orthogonally to the incoming striatal axons. Moreover, the striatal dendritic tree is very dense, whereas the pallidal dendritic field is very sparse, with only a few dendrites (on average 4 main and 13 tips). A quantitative study [70] has shown that a striatal axon provides 240 synapses in the primate pallidum and makes ten contacts with one pallidal neuron on average. Considering the total number of striatal and pallidal (primate) neurons ($\sim 10^7$ and $\sim 10^5$, respectively) the

![Figure 5. Example of TAN responses to visual signals and reward in probability task. Each row represents one cell of the electrodes depicted in Fig. 4. Each column represents a different probability associated with reward: (A) responses to visual signal, (B) responses to reward.](image-url)
probability of two pallidal cells to share common striatal inputs is also very low [71].

4.4. Redundancy reduction vs. sparse connectivity

The number of cortico-striatal neurons exceeds the number of striatal neurons by a factor of 10 [5]. It was therefore concluded that unless the cortico-striatal output is overwhelmingly redundant, the striatum cannot compress the full cortical information. However, dimensionality (redundancy) reduction networks do not always perform compression without information loss. In many cases, (especially in those where the number of elements in the output structure is significantly smaller than the number of elements in the input layer), the compression process will result in information loss. However, the network will keep the “most important” parts of the input patterns.

Regardless of the underlying reason, it is clear that the cortico-striatal physiological message is not a simple read-out of the cortical state [72,73]. Multiple lines of evidence suggest that cortico-striatal and cortico-spinal (cortico-peduncular) neurons belong to distinct populations, and that the signal transmitted by cortico-striatal neurons is distinct from that sent to the spinal cord or the brainstem. It seems that the firing of cortico-striatal neurons is more selective than that of cortico-spinal neurons. Although, there is no data regarding the redundancy or correlation between cortico-striatal neurons, one may conclude that only sparse and selective cortical information is transmitted to the striatum, and therefore this information can be compressed in the much smaller number of striatal, and subsequently, pallidal, neurons, only if some of it is lost.

Even more important is the finding of sparse connectivity within the basal ganglia. Each single striatal neuron receives ~5000 cortico-striatal synapses. A single striatal cell is therefore exposed only to less than 0.01% of the cortical information. Standard competitive networks have “all to all” connectivity, i.e., all neurons in the output layer receive projections from all input neurons (Fig. 6(A)). Such all-to-all connectivity enables the system to adapt to the changes in the input pattern, and to create optimal (with minimal loss of information) representation of the input patterns. However, more realistic redundancy reduction models of the basal ganglia must assume that each striatal neuron receives information from a limited number of cortical neurons (Fig. 6(B) and (C)). This subset of connections could either constitute one of a number of fully segregated cortico-striatal pathways (Fig. 6(B)), or of partially overlapping circuits (Fig. 6(C)). Finally, one does not have to assume a symmetric network. Lateral directionality (Fig. 6(D)) may enable the sharing of information in the basal ganglia in “ascending” order [74,75], e.g., from limbic to cognitive to motor domains.

Fig. 6. Architecture of the feed-forward connectivity in the basal ganglia. The figure shows the possible configuration for the feed-forward connections (cortex → striatum or striatum → GPi/SNr): (A) all-to-all connectivity, (B) fully segregated convergence network, (C) mixed convergence–divergence network, (D) asymmetric (from left to right) convergence–divergence network.

5. Conclusions

The physiological evidence, indicating lack of correlation between neurons in the output nuclei of the basal ganglia, and the massive reduction in the number of neurons from the input to the output structures, suggest that the basal ganglia output is a compressed form of certain aspects of cortical activity. However, this compression probably involves information loss. The teaching signals of the basal ganglia, delivered by cholinergic and dopaminergic neurons may enable the system to keep the most important aspects of the information. We showed that these two “critics” are different in their sensitivity. What is the significance of each of these signals in the information processing in the basal ganglia, and how the cortex uses the compressed output of this system, are yet to be answered.

Acknowledgements

This center of excellence (8006/00) was supported by the Israel Science foundation. This research was also supported in part by the United States–Israel Binational Science Foundation, the German–Israel Binational foundation (GIF) and the BMDE Israel–Germany collaboration in medical research. G.M. and A.N. have made equal contribution in this work.

References

A. Nini, A. Feingold, H. Slovin, H. Bergman, Neurons in the
D. Jaeger, H. Kita, C.J. Wilson, Surround inhibition among
P. Calabresi et al., Synaptic transmission in the striatum: from
J.N. Reynolds, B.I. Hyland, J.R. Wickens, A cellular mechanism
S.M. Nicola, J. Surmeier, R.C. Malenka, Dopaminergic modula-
T. Koos, J.M. Tepper, Inhibitory control of neostriatal projection
588
F.A. Middleton, P.L. Strick, Basal ganglia and cerebellar loops:
D.E. Oorshcot, Total number of neurons in the neostriatal,
pallidal, subthalamic, and substantia nigral nuclei of the rat basal
ganglia: a stereological study using the cavaliere and optical
C.D. Hardman et al., Comparison of the basal ganglia in rats,
marmosets, macaques, baboons, and humans: volume and
neuronal number for the output, internal relay, and striatal modulating nuclei, J. Comp. Neurol. 445 (2002) 238–255.
F.A. Middleton, P.L. Strick, Basal ganglia and cerebellar loops:
D. Joel, Open interconnected model of basal ganglia–thalamo-
A. Parent et al., Organization of the basal ganglia: the importance of axonal collateralization, Trends Neurosci. 23 (2000) S20–S27.
A. Raz, E. Vaadia, H. Bergman, Firing patterns and correlations of spontaneous discharge of pallidal neurons in the normal and the tremulous 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine ver-
Y. Kawaguchi, C.J. Wilson, S.J. Augood, P.C. Emson, Striatal interneurons: chemical, physiological and morphological character-
T. Aosaki, M. Kimura, A.M. Graybiel, Temporal and spatial cha-
R.S. Sutton, A.G. Barto, Reinforcement Learning—An Introduc-
R.E. Suri, W. Schultz, A neural network model with dopamine-


