

Original Article

A New Module for On-Line Manipulation and Display of Molecular Information in the Brain Architecture Management System

Mihail Bota* and Larry W. Swanson

The Neuroscience Research Institute, University of Southern California, 3641 Watt Way, Los Angeles, California 90089-2520, USA

Abstract

A new "Molecules" module of the Brain Architecture Management System (BAMS; <http://brancusi.usc.edu/bkms>) is described. With this module, BAMS becomes the first on-line knowledge management system to handle central nervous system (CNS) region and cell-type chemoarchitectonic data in the context of axonal connections between regions and cell types, in multiple species. The "Molecules" module implements a general knowledge representation schema for data and metadata collated from published and unpublished material, and allows insertion of complex reports about the presence of molecules collated from the literature. For different CNS neural regions and cell types, the module's database structure includes representation of molecule expression revealed by various techniques including *in situ* hybridization and immunohistochemistry, molecule coexpression and time-dependent level changes, and

physiological state of subjects. The metadata representation allows online comparison and evaluation of inserted experiments, and "Molecules" structure allows rapid development of data transfer protocols enabling neuroinformatics visualization tools to display gene expression patterns residing in BAMS, in terms of levels of expressed molecules and *in situ* hybridization data. The module's web interface allows users to construct lists of CNS regions containing a molecule (depending on physiological state), retrieve further details about inserted records, compare time-dependent data within and across experiments, reconstruct gene expression patterns, and construct complex reports from individual experiments.

Index Entries: Chemoarchitecture; coexpression; database; immunohistochemistry; inference engine; *in situ* hybridization; molecules; molecular neuroanatomy.

(Neuroinformatics DOI: 10.1385/N1:4:4:275)

*Author to whom all correspondence and reprint requests should be addressed.

E-mail: mbota@almaak.usc.edu

Introduction

The vast amount of neuroscience-related textual and graphical information—which actually extends back centuries—is one important factor that makes searching and interpretation a difficult task. In addition, the interpretation and integration of data are derived from different levels of central nervous system (CNS) organization. Any CNS region can be described with respect to multiple levels of organization: from patterns of gene expression under specific experimental conditions; to structural, chemical, and physiological characteristics of its constituent neuronal and glial cell types; to specific roles in functional networks of regions in a particular species (Swanson 2000a,b; Bota et al., 2003).

Database systems are complementary to the neuroinformatics visualization tools and applications developed by individual research groups, and are necessary for data storage and handling. The display of multiple modalities—including structural and functional data, and gene expression patterns (levels of expressed molecules and *in situ* hybridization data)—on three-dimensional brain reconstructions or two-dimensional atlas reconstructions of brains from various species requires the design and development of complex backend databases as information repositories.

To address these problems we are developing the on-line knowledge management system Brain Architecture Management System (BAMS) for systematizing, organizing, and processing neuroscience information relevant to different levels of CNS organization. The structure of BAMS allows insertion of data about CNS regions, cell types, neural (axonal) pathways, and molecules collated from the literature or recorded by users. BAMS includes inference engines that establish qualitative topological relations between CNS regions, reconstruct projection patterns from CNS regions of interest, and construct possible networks of CNS regions

from connectivity data resident in the system (Bota et al., 2005).

BAMS is also a data provider for neuroinformatics systems developed by other groups. For example, simple data transfer protocols have been implemented with several web-based and stand-alone applications created by the Laboratory of NeuroImaging (LONI; <http://www.loni.ucla.edu>) at UCLA that provide hierarchically organized neuroanatomical nomenclatures and definitions of CNS parts (MacKenzie-Graham et al., 2003, 2004).

Here we describe the structure of a new BAMS module, “Molecules,” that allows insertion of molecular (chemical) data pertaining to various CNS regions and cell types. We also describe here the user interfaces designed to allow searching for molecules in BAMS, whose database structure and inference engines are described elsewhere (Bota et al., 2005).

Methods

BAMS is hosted on a Dell 8200 desktop computer with a Pentium III processor running under the Windows XP operating system, with IIS5.1. The backend relational database of BAMS was created in MySQL, and PHP was used to construct web interfaces and inference engines. The architecture of BAMS is constructed on three levels: a set of tables constructed in MySQL that store data collated from the literature (or inserted by neuroanatomists), an intermediate level encoded in PHP that includes queries and algorithms for processing data populating BAMS, and an output level that is mainly in an HTML tabular format and graphics.

“Molecules” Module Database Structure

The BAMS’s “Molecules” module consists of two parts: a public part, and a “Personal account” part that was designed to store data entered by collators and restricted to their use. The knowledge representation schema of BAMS’s “Molecules” module (Fig. 1) follows

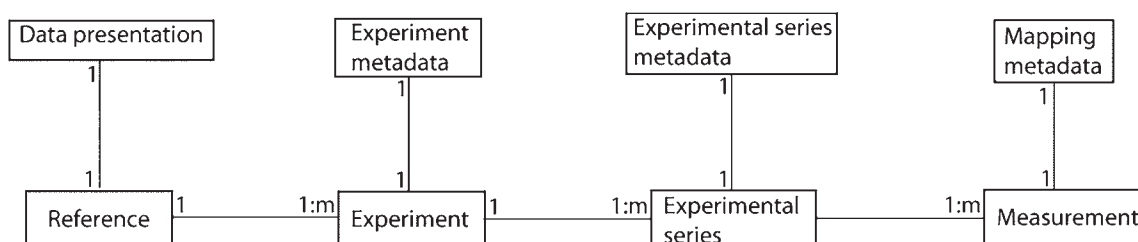


Fig. 1. The knowledge representation schema of "Molecules" module in BAMS.

the general organization and presentation of experimental data in published neuroscience research articles.

Any reference inserted in BAMS's "Molecules" module can be associated with multiple experiments. An experiment is defined as an experimental paradigm applied to a group of animals or human subjects. Each experiment may consist of several experimental series, which are defined by the specific procedures (e.g., different antibodies or nucleic acid probes) that were applied. Each experimental series is associated with a set of experimental data that include mapped brain regions and measured variables. The conceptual design of BAMS's "Molecules" module completely separates experimental data from metadata collated from the reference, or inserted by collators. Metadata classes associated with data and experiments inserted in BAMS's "Molecules" module include: mapping, experimental series, experiment, data presentation, and physiological state. Each of these metadata classes is described later.

The general object-relationship schema that stores data and physiological state metadata in BAMS's "Molecules" module is presented in Fig. 2. Each object and relation shown in the figure may be captured in more than one table. The backend MySQL relational database of the "Molecules" module public part now has 17 tables (listed in Table 1).

The table *Brain part* contains the unique identification of CNS parts (gray matter, fiber tracts, or ventricles) as collated from neuroanatomical

nomenclatures, and it was described in Bota et al. (2005). The table *Chemoarchitecture* allows insertion of a large range of experimental data: from qualitative assessments of the presence or absence of molecules in different CNS neural regions to quantitative measurements of such data. Qualitative attributes that describe the distribution of a molecule in a CNS region include staining strength, staining pattern, and topographical position of stained cells in the region. The assessment of qualitative strengths is similar to that for neuroanatomical projections between regions, as described in detail elsewhere (Bota et al., 2005). *Chemoarchitecture* also allows insertion of data pertaining to cell counts, percent of cells in terms of total area or per area, and average number of labeled cells (\pm standard deviation or standard error of the mean). Quantitative and semiquantitative data that can be inserted in this table include optical absorption, absorption level relative to background, standard deviation, and standard error of the mean.

The "Molecules" module also allows insertion of qualitative and quantitative reports of molecule expression in different cell types, as revealed by ligand autoradiography, immunocytochemistry, or *in situ* hybridization. BAMS's "Molecules" module includes the table *Coexpression* that contains reports of molecule coexpression in CNS regions or cell types under various physiological conditions, as collated from the literature. The presence or absence of a molecule in a CNS region or cell type often depends on the age, sex, and/or

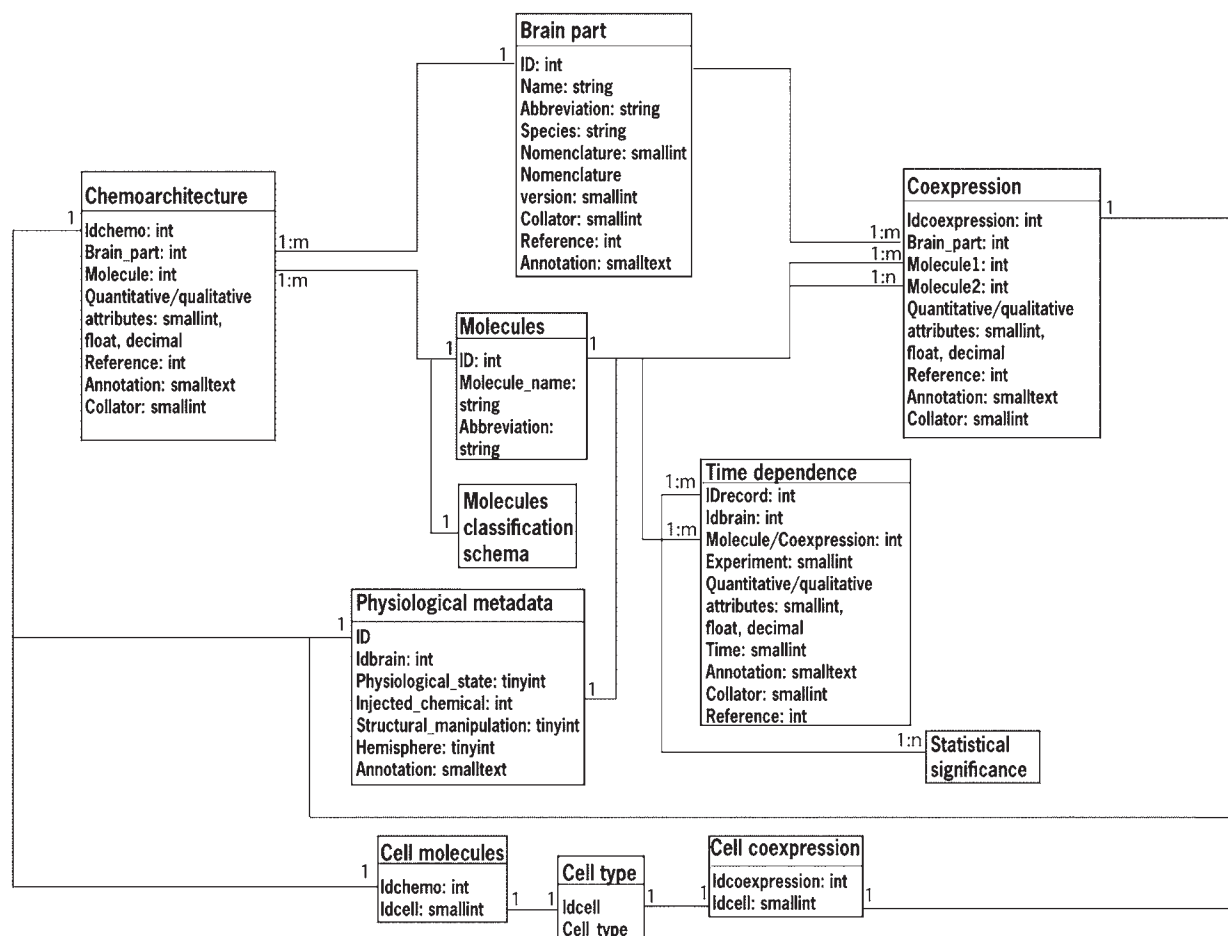


Fig. 2. The object-relationship schema of Molecules' public part for experimental data and physiological state metadata.

physiological state of the subjects. Therefore, any database that stores data related to the presence of molecules in different CNS regions should include physiological state metadata (see Fig. 2). The "Molecules" module includes a simple representation of the subject's physiological state encoded in the following tables: *Physiological_state*, *Manipulation_type*, and *Manipulation_data*. The table *Physiological_state* encodes for the values allowed in BAMS, which are "normal" and "manipulated." An experimental animal is considered to be in a "normal" physiological state, when the authors of the collated reference state that no special

procedure was applied. The physiological state "manipulated" currently refers either to injections of chemicals, or to structural lesions in different CNS regions or parts of the experimental animal. The structural lesions produced by different chemicals injected in the CNS are classified as chemical manipulations. Therefore, the physiological state "manipulated" is represented in BAMS by two possible values: "chemical manipulation" and "structural manipulation," which are encoded in the table *Manipulation_type*. The table *Manipulation_data* stores details about the lesioned CNS parts, the side of the brain where

Table 1
The Tables Included in the Public Part of BAMS's "Molecules" Module That Represent Experimental Data and Physiological State Metadata

Table in BAMS	Encodes for
Brain part	Unique identification and basic information about brain parts collated from different neuroanatomical atlases or nomenclatures
Molecule	Name and abbreviation of those chemicals inserted in BAMS
Molecule_type	The major classes of molecules allowed in BAMS: cell associated or releasable by neurons
Molecule_releasable	Subclasses of molecules considered in BAMS to be releasable by neurons
Cell_associated	Subclasses of molecules considered in BAMS to be found in cells and not released by neurons
Enzymes	Basic information about enzymes inserted in BAMS
Receptors	Basic information about receptors inserted in BAMS
Receptors subunits	Basic information about subunits of receptors inserted in BAMS
Chemoarchitecture	Chemoarchitecture reports as collated from the literature
Coexpression	Reports of molecule coexpression identified in different brain regions or cell types, and collated from the literature
Time dependence-chemoarchitecture	Reports of time-dependent chemical data
Time dependence-coexpression	Reports of time-dependent coexpression data
Statistical significance	Information about statistically relevant data measured across or within experiments
Physiological state	Allowed values of the subject's physiological state
Manipulation type	Information about the chemical or structural manipulations allowed in BAMS
Manipulation data	Information such as location, hemisphere, and associated annotation about the physical or chemical manipulation performed on subjects

the procedure was performed, the injected chemicals or lesioned CNS regions, and annotations from the collated references.

The BAMS "Molecules" module also includes several tables that allow insertion of time-dependent molecular data. The range of experimental data and variables that can be inserted in the tables, *time_dependence-chemoarchitecture* and *time_dependence-coexpression* are identical with those in *chemoarchitecture* and *coexpression*, respectively. Records inserted in either of these tables are grouped in experiments, as reported in collated references. This allows recording the statistical significance of experimental data calculated within or across experiments as reported in collated references.

"Molecules" Module Classification Schema

In this initial implementation, "Molecules" includes a simple classification schema for its namesake, with just two general classes of molecules: "releasable by neurons" and "cell associated," which are encoded in the table *Molecule_type*. These two classes partition completely the set of molecules that can be inserted in BAMS, because any chemical that can be identified in a CNS region is considered here for simplicity to be produced by neurons (glial cells will be incorporated in the future). Again for simplicity, the current partition of molecules in BAMS is of the disjoint type. That is, any molecule can be classified in either of the

general classes, but cannot be an instance of both. The design of the "Molecules" module allows each of these two classes to be further divided into subclasses. The class "cell associated" is subdivided at present into two subclasses: "enzymes" and "receptors." The class "receptors" is further subdivided into "receptor subunits." Each of these classes and subclasses is assigned to a table that encodes for relationships between parents and children, and names and abbreviations of the classified molecules. The classification schema included in "Molecules" allows new children to be inserted into either of the most general classes.

Metadata

The class "Experimental series metadata" shown in Fig. 1 allows description of procedures applied to series included in an experiment, and is divided in three subclasses that correspond to the techniques encoded in BAMS's "Molecules" module: ligand autoradiography, immunohistochemistry, and *in situ* hybridization. The fields that make each of these subclasses are listed in Table 2.

Ligand autoradiography metadata describe receptor mapping experiments, using radioactively labeled ligands. In constructing this metadata subclass we followed Burt (1985); Kuhar (1985); and Kuhar et al. (1986). The "equilibrium dissociation constant" and "maximum specific binding" fields are used to describe ligand-receptor interactions, and measure the ligand affinity and the maximal number of bound receptors, respectively. The "quantitation standard" and "quantitation method" fields describe the procedures used to quantify receptor autoradiograms (Burt 1985, Kuhar et al., 1986).

Immunohistochemistry metadata describe experiments using antibodies to reveal the molecular phenotypes of neurons. Fields included in this metadata class subtype have been discussed in Watts and Swanson (1989); Burry (2000); Saper and Sawchenko (2003); and Saper (2005). The field "performed control"

specifies what type of control was performed to ensure antibody specificity, and allowed values are: "knockout," "*in situ* hybridization," "immunoblot," "immunoprecipitation," "adsorption," "staining pattern," and "not specified" (Saper, 2005).

In situ hybridization metadata describe neuronal molecular phenotype mapping with nucleic acid probes. The fields included in this metadata class subtype are used extensively in the literature (see for e.g., Swanson and Simmons, 1989; Wada et al., 1989; Watts and Swanson, 1989; Watts and Sanchez-Watts, 2002). The "labeling method" field encodes for the technique used to label the probe sequence and the allowed values are "radiolabeled," "antigen labeling," "fluorescence," and "not specified."

All three subclasses of "Experimental series metadata" have several common fields: "source (provider)," "visualization method," "visualization support," and "annotation." The "source (provider)" field can be used to register the company or laboratory that provided the ligand, antiserum, or nucleic acid sequence used. This information can be very important in determining the specificity and affinity of a specific ligand or antiserum (Kuhar et al., 1986; Sawchenko and Saper, 2003; Saper, 2005), as well as for other specific information. The "visualization method" field specifies procedures used to visualize target molecules, and allowed values are: "autoradiogram," "fluorescence," "PET," "colorimetry," and "not specified." Allowed values for the "visualization support" field are "film," "slide," "none," and "not specified." Finally, the "annotation" field can be used to insert more details about each of the techniques.

The metadata associated with an experiment (see Fig. 1) are listed in Table 3. This metadata class is made up of three subclasses, and includes two fields that specify the method of neuron identification and the staining frequency, respectively. The "Animals (subjects)"

Table 2
 Metadata Subclasses and Fields Included in the "Experimental Series Metadata" Class

Experimental series metadata	Ligand autoradiography	Ligand Ligand source (provider) Equilibrium dissociation constant (K_D) Maximum specific binding (B_{max}) Quantitation standard Quantitation method Visualization method Visualization support Annotation
	Immunohistochemistry	Antibody source (provider) Antigen Antigen species Primary antibody Primary antibody species Primary antibody type Secondary antibody Secondary antibody species Immunoglobulin class Performed control Visualization method Visualization support Annotation
	<i>In situ</i> hybridization	Sequence type: mRNA, heterogeneous nuclear (hn)RNA Sequence source (provider) Sequence Sequence direction Species sequence Control direction Labeling method Visualization method Visualization support Annotation

subclass allows insertion of basic information about the group of animals or subjects, including number, sex, weight, age, and living conditions. The "Employed technique" subclass lists the experimental techniques used in the associated experiment. The "Anatomical metadata" subclass allows insertion of details associated with tissue preparation.

Finally, metadata classes associated with a reference in BAMS's "Molecules" module, "Data presentation" and "Mapping/nomenclature metadata," are listed in Table 4. The class "Data presentation metadata" includes three subclasses: "Coordinates," "Description types," and "Data types." The subclass "Coordinates" allows collators to insert metadata related to stereotaxic

Table 3
Subclasses and Fields Included in the “Experiment Metadata” Class

Experiment metadata	Animals (subjects)	Number Sex Weight Age Living conditions Annotation
	Neuron/glia identification method Staining frequency (sampling) Used technique	Ligand autoradiography Immunohistochemistry <i>In situ</i> hybridization
	Anatomical metadata	Sectioning method Section orientation Angle of cutting Section thickness Preservation method Staining method Staining frequency Annotation

or Talairach coordinates associated with data. The “Description types” subclass is similar to the precision description codes implemented in CoCoMac (Stephan et al., 2001), and its fields include possible ways of presenting data in a reference: from textual descriptions to complete sets of maps representing labeled CNS regions on templates of reference or standard neuroanatomical atlases. The “Mapping” metadata class allows collators to specify how a record abstracted from a reference was associated with CNS part names from nomenclatures inserted in BAMS. This metadata class is important for maintaining the accuracy and consistency of data in BAMS (see “Curation and Insertion of Data” below). The “Data types” metadata subclass contains types of data associated with the reference: quantitative, semiquantitative, or qualitative. This metadata subclass is not inserted by collators, but is inferred by the system from inserted information.

Curation and Insertion of Data

One of the most important problems in populating and maintaining a database is the

quality of inserted data. To help ensure accuracy of data entry in BAMS, we designed a multistep, comprehensive curation policy that is applied to all information inserted in public modules. This policy involves different actions by BAMS’s collators and curators, as well as by authors of inserted references. The curation policy schema implemented in BAMS is shown in Fig. 3.

The first step of BAMS’s curation methodology is related to data sources. All information inserted in BAMS’s public modules is collated from published literature and is documented by references—data inserted in BAMS’s public modules is collated from original research articles. Methods for handling the literature and BAMS database tables that hold information about inserted references are described in Bota et al. (2005) and Bota and Arbib (2004). The second step in BAMS’s curation methodology is data mapping onto one of the inserted neuroanatomical nomenclatures. Any record inserted in BAMS must be associated with one of its neuroanatomical nomenclatures. Allowed mappings of a statement in

Table 4
Subclasses and Fields Included in "Data Presentation" and "Mapping/Nomenclature" Metadata Classes

Data presentation metadata	Coordinates	Lambda Bregma Interaural distance X Talairach Y Talairach Z Talairach
	Description types	Text Representative labeled images Integral set of labeled images Representative labeled images with drawn boundaries Integral set of labeled images with drawn boundaries Representative drawings, or mapping on atlas templates Integral set of drawings or mappings on atlas templates
	Data types	Quantitative data Semiquantitative data (ratios) Qualitative data
	Mapping metadata	Brain region is captured in original published nomenclature, in BAMS Brain region is captured in a BAMS nomenclature other than that used in the original publication Brain region is not in a BAMS nomenclature, and the mapping to a BAMS nomenclature was performed by the author Brain region is not in a BAMS nomenclature, and the mapping to a BAMS nomenclature was performed by the collator

the general form "molecule X is present in CNS region Y"—in one of BAMS's nomenclatures—are listed in the "Mapping/nomenclature" metadata class. The first of two currently allowed mappings is based on a statement in the collated reference that refers explicitly to the neuroanatomical nomenclature used. Here, the collator simply checks whether the nomenclature is inserted in BAMS. If it is inserted, and the collator wishes to use this nomenclature, no further action is necessary. The second allowed mapping occurs when the collator wishes to map all or part of the data onto a different nomenclature resident in BAMS. This remapping can be performed by

the collator and/or original author, with the procedure stated clearly in the annotation associated with all records in BAMS. If neither of these mapping possibilities is met, the record cannot be inserted in BAMS. To circumvent this condition, the collator (or original author) may wish formally to enter a new, satisfactory nomenclature into BAMS, so that the second allowed mapping procedure can be followed.

In short, BAMS's curation policy has two rules for data entry in its public modules: data is collated from published original research articles, and each record has to be mapped onto one of BAMS's nomenclatures. Hence, association with the "Mapping metadata" class is mandatory for

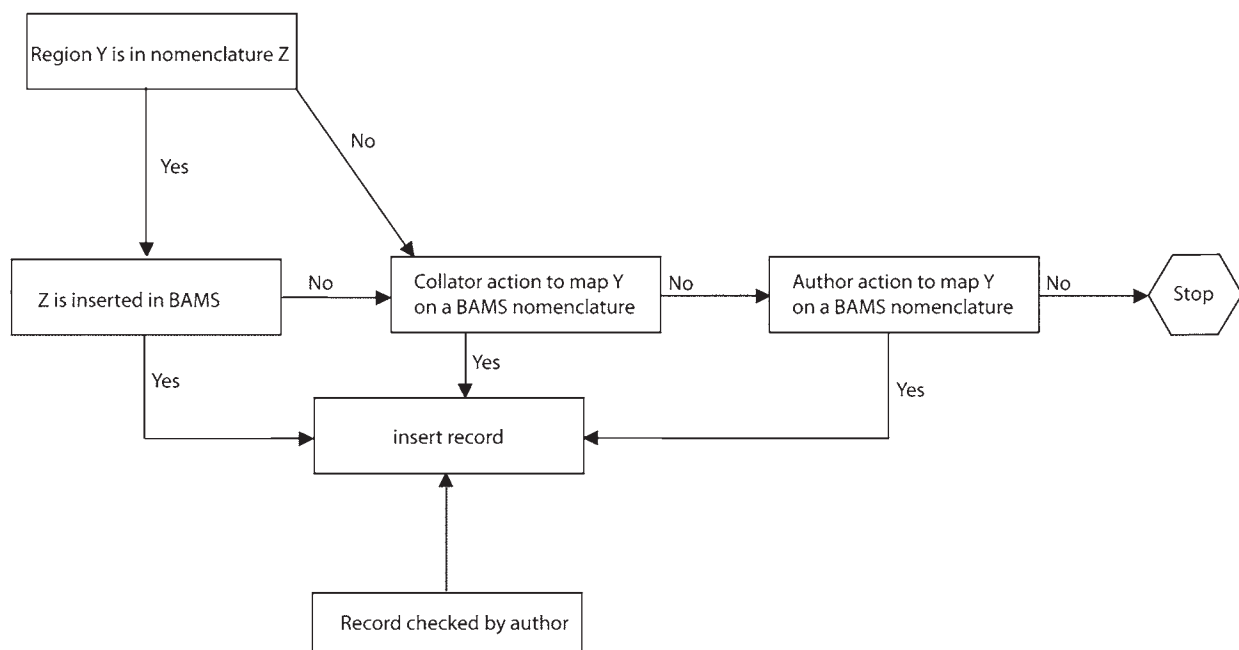


Fig. 3. The conceptual schema of the curation policy implemented in BAMS.

insertion of records in the “Molecules” module, whereas the other metadata classes are optional. The third step of BAMS’s curation policy consists of records and metadata checking performed by the reference’s author. Once information collated from a reference is completely inserted in BAMS, curators contact authors to verify the accuracy, consistency, and completeness of inserted data and metadata. All data and metadata inserted in BAMS and checked by authors of collated references is tagged and displayed accordingly in the backend database and the user interface, respectively.

BAMS’s “Molecules” module also includes a “Personal Account” part that allows registered users to insert data from collated references, or unpublished experimental data. The database structure of this part is similar to the public part of “Molecules,” and the insertion process follows the knowledge representation schema shown in Fig. 1. The conceptual schema of data and metadata insertion and update in the “Personal Account” section is shown in Fig. 4.

Results

Searching for Molecules in BAMS

Users can search for information related to molecules, compare the presence of molecules in different CNS regions, search for information by type of manipulation, and construct reports from multiple experiments. The search for information related to molecules identified in CNS parts is performed by using a form that displays all molecules inserted in BAMS, arranged according to the classification schema. The result of this search is a list of CNS parts where the searched molecule was identified, and associated experimental conditions. Details about the presence of a molecule in a CNS region can be accessed by clicking on the link associated with the experimental condition (Fig. 5). As shown in Fig. 5, users can access the metadata associated with a record. An example of metadata associated with retrieved records, and accessible by users is shown in Fig. 6.

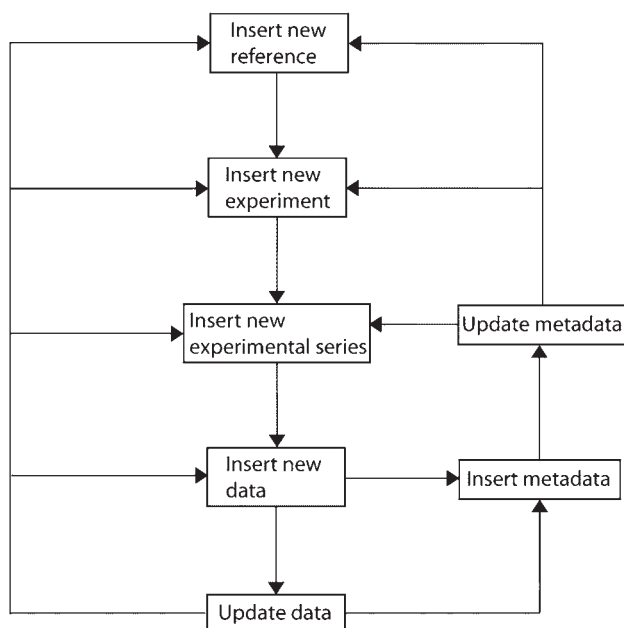


Fig. 4. The conceptual schema of the data entry part of Molecules' "Personal account" part.

Users can also view all molecules identified in a brain region when they search for CNS regions. The result of a search for CNS regions in BAMS is a page that summarizes data collected by researchers or inferred by the system and associated with the part of interest. A detailed description of BAMS web interfaces is provided in Bota et al. (2005). If a CNS region is associated in BAMS with several molecules, they are returned in the page summarizing the region, as shown in Fig. 7. Users can access further information about each retrieved molecule by clicking on the associated links (Fig. 8).

Records associated with the molecule of interest are grouped in four categories (molecule presence in normal and manipulated physiological state, and coexpression of other molecules in normal and manipulated state; see Fig. 8, by experimental condition, and by coexpression of other molecules in the same CNS region). The system first retrieves references corresponding to each category. Every reference is associated with a "Data" link that retrieves corresponding experimental records. If at least two experiments from different

references are retrieved in the same category and have metadata associated with them, a new link called "Methods comparison" becomes available. Users can therefore compare methods used in all experiments associated with metadata retrieved in one of the four categories. An example of methods comparison is shown in Fig. 9.

The methods comparison across references and experiments for the same CNS region, molecule, and experimental condition first retrieves metadata associated with each retrieved reference: data types, mapping approach, and data presentation. This allows users to compare rapidly and evaluate experiments and references in terms of data richness and mapping consistency. The second comparison is performed across experiment metadata. And the third comparison allows users to compare metadata associated with retrieved experimental series. In short, this BAMS user interface functionality can be used by users to compare the accuracy and reliability of molecular data by comparing associated metadata.

Species: Rat														
Data about presence of corticotropin-releasing hormone in PVHmpd manipulated state														
Brain regions where CRH is present	Experimental condition	Physiological condition	Cell pool position in region	Hemisphere	Qualitative density of CRH in PVHmpd	Labeled cells count	Percentage labeled cells	Average labeled cells	Standard deviation	Relative to basal	Annotation	Reference	Collator	
PVHiv	manipulated state													
PVHmpd	basal state													
PVHmpd	manipulated state	Chemical treatment: colchicine	everywhere	bilateral	not assigned	808	40.50	0	not measured	not measured	Collator note: see Figure 2 page 168. The percentage of labeled cells was calculated using the numbers provided in the legend of Figure 2 and represents the ratio between the labeled cells in the associated region and the total number of labeled cells in this experiment.	Swanson LW, Sawchenko PE, Rivier J, Vale WW, 1983	Mihail Bota	Metadata ✓
PVHmpd	manipulated state	Experiment P5439, received 200 micrograms of colchicine in the lateral ventricle 18 hours before perfusion.												
PVHm	manipulated state													
PVHmp	manipulated state													
PVHmp	manipulated state													
PVHmp	manipulated state	Chemical treatment: colchicine	unclear	bilateral	not assigned	0	0.00	446	50.00	not measured	Collator note: experiment CRH/AVP, Table 3 page 506. Page 505: eighteen to twenty hours after an injection with PEG, there was a significant 22% (P<0.01, Fig.1) increase of the CRH mRNA in the PVHmpd compared to vehicle-injected controls when measured using the 35S-cRNA probe.	Watts, A.G. & Sanchez-Watts, G., 1995	Mihail Bota	Metadata
PVHmp	manipulated state													
PVHmp	manipulated state	Structural manipulation: ablation												
PVHmp	manipulated state		not known	bilateral	exists	0	0.00	0	not measured	not measured	...90% of the brightly labeled cells were found in the medial part of the parvocellular division. Collator note: see Figure 1b page 167. We assigned the CRH label as significantly higher than the normal state according to Swanson 1991 in Progress in Brain Research.	Swanson LW, Sawchenko PE, Rivier J, Vale WW, 1983	Mihail Bota	Metadata
PVHmp	manipulated state	Manipulated structure: Adrenal gland.												
PVHmp	manipulated state	Chemical treatment: polyethylene glycol	everywhere	left hemisphere	very strong	0	0.00	0	not measured	significantly higher	Collator note: see Figure 2 page 6264 and Figure 3, page 6265.	Watts A.G. & Sanchez-Watts, G., 2002	Mihail Bota	Metadata

Fig. 5. The result of search of CNS regions where the neurotransmitter corticotropin-releasing hormone was identified. Users can view details of records associated with each retrieved CNS region, including qualitative density, cell counts, statistical measurements, spatial characteristics of cells expressing the molecule, and associated annotations by clicking on links associated with experimental conditions. For reports associated with manipulated state, users can also view type of manipulation, injected chemicals, and details about experimental procedure. Users may also access metadata associated with retrieved records. Records verified by original authors of collated references will also have a check mark associated with them.

If a molecule identified in a CNS region is associated in BAMS with time-dependent data, users can access this information by clicking on a link called "Time dependence" (Fig. 8). The time-dependent molecular data are grouped by experiment, reference, and experimental state of the animals. The statistical significance associated with the quantities being compared is also retrieved, and is graphically

coded. An example of time-dependent data retrieval is shown in Fig. 10.

BAMS's "Molecules" module includes an inference engine for reconstructing the chemoarchitectonic profile of a CNS region from the molecular data associated with its substructure. This engine can be accessed from the page describing CNS parts in BAMS, shown in Fig. 6. It is similar to the projections

Experiment acronym	Annotation	Collator	Author checked
Broberger-H/O-MCH	In situ hybridization study of the H/O and MCH neurons distributions in the rat hypothalamus	Mihail Bota	no

Animals (subjects) Sex: M Number of animals: 10 Age: not specified Mass: 250-300 Unit of mass: Housing conditions: normal housing Annotation: Male Sprague-Dawley rats (n = 10, 250–300 g; B&K Universal, Stockholm, Sweden) and male C57B1mice (n = 5, 20–40 g; B&K Universal) were used. The animals were kept under regular lighting conditions (lights on at 6.00 and off at 18.00) in a temperature-controlled environment and had free access to standard rodent chow and tap water.	Experimental method Experiment type: in situ hybridization Neuron/glia identification method: stain specific Staining frequency: 1:10 Experimental details: Measured nucleic acid: mRNA Source (producer): Scandinavian Gene Synthesis, Koping, Sweden). Probe sequence: 2-49 H/O nucleotides Sequence species: rat Probe sequence orientation: antisense Control: sense Labelling method: radiolabelling Visualization method: autoradiogram Visualization medium: slide Annotation: Probes complementary to nucleotides 479-527 of the rat MCH mRNA(Nahon et al., 1989) and to nucleotides 2-49 of the rat H/O mRNA(Sakurai et al., 1998) were synthesized (Scandinavian Gene Synthesis, Koping, Sweden). Probe sequences were controlled against other sequences in the GenBank database, and no homologies exceeding 75% were found.
---	--

Anatomy and histology				
Section plane: coronal	Angle: not specified	Cutting method: microtome	Preservation: freezing	Thickness: 14 micrometer
Staining type: toluidine blue	Sampling: 1:10	Annotation: The brains were cut serially at 14micrometers thickness on a cryostat (Microm, Heidelberg, Germany), and every tenth section was processed for double-labeling in situ hybridization.		

Mapping details	
Coordinates: none	Data presentation: text representative labeled images representative images with label and drawn boundaries representative drawings or mappings onto templates Mapping approach: brain region in a BAMS nomenclature

Fig. 6. Users can access metadata associated with retrieved experimental data, and with the corresponding experiment and experimental series.

profile inference engine described in Bota et al. (2005), and it displays gene expression data as a function of experimental conditions (normal and manipulated states). An example chemo-architectonic profile reconstruction is shown in Fig. 11.

BAMS's "Molecules" module also allows construction of reports from references and experiments of interest. Users can choose experiments individually and reconstruct the associated data in matrix format. The output of such customized reports is similar to the matrix shown in Fig. 11.

Comparing Molecules in BAMS

The "Molecules" module user interface includes two options for comparing the presence of various molecules in particular CNS regions from inserted data. The first option, called "Comparison of molecules existence in CNS regions," processes the query type "What are the CNS parts where all the molecules of interest have been identified?" This engine returns a list of CNS parts where all the chosen molecules have been identified, and the associated physiological states. This engine also returns records of coexpression data for any

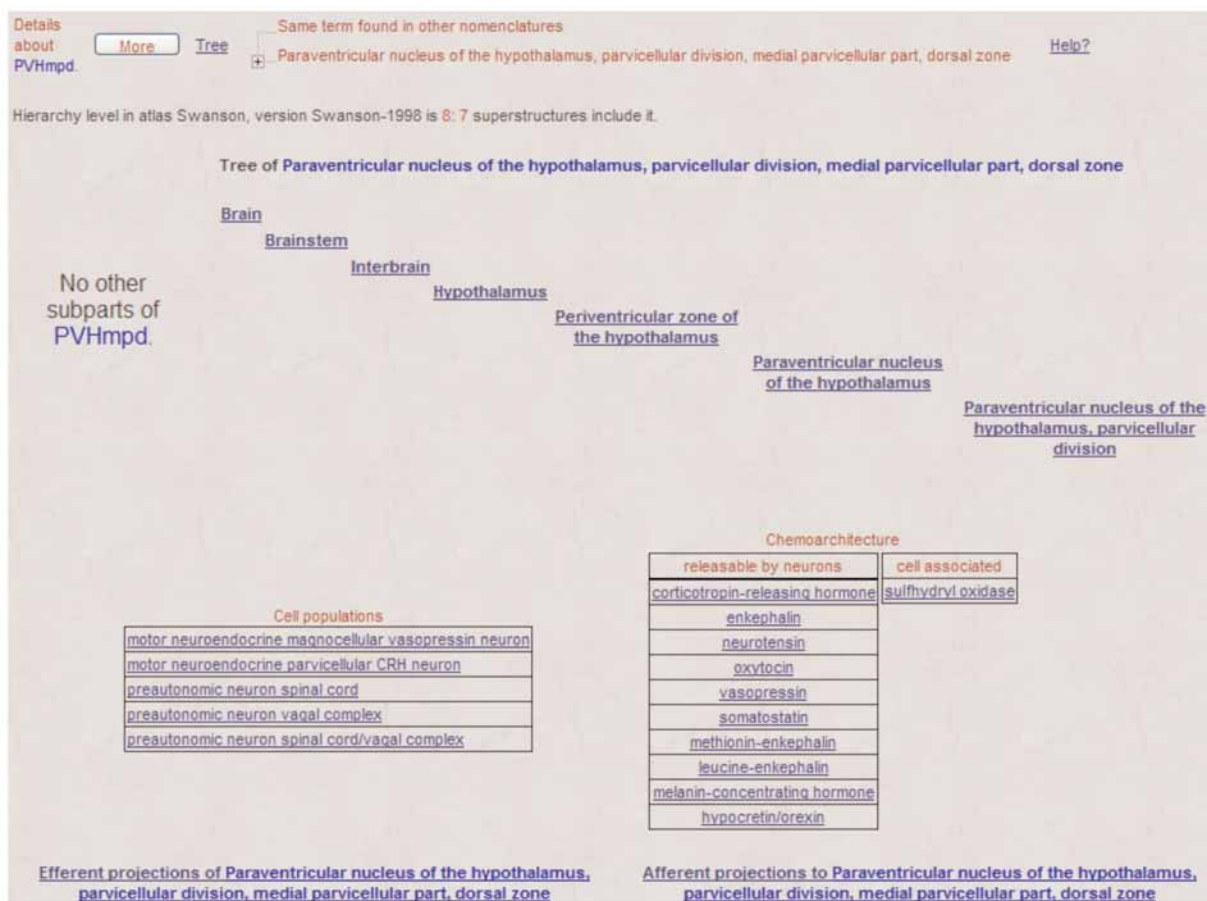


Fig. 7. The result of a CNS parts search in BAMS. If a CNS region is associated in BAMS with molecules, the latter are listed under the category “Chemoarchitecture” and divided according to Molecule’s classification schema.

pair of molecules in the chosen set. An example of the reconstruction of the molecular composition of CNS parts is shown in Fig. 12.

The second option, called “Search by type of manipulation,” processes the query type “Show identified molecules and CNS parts that are associated with a specific molecule or structural manipulation.” This engine returns the list of CNS parts and molecules associated with the searched type of manipulation, as well as basal state data for each retrieved molecule. If any retrieved experimental data is associated with statistical information, it also will be displayed in graphical format. An example of the

online retrieval of chemoarchitectonic data associated with systemic injection of polyethylene glycol is shown in Fig. 13.

The BAMS Workspace

The functionality of BAMS’s web interface was augmented with two additional features: registered users may now add comments to data reports, and they can also save their activity in a personal workspace. Registered users are allowed to attach comments to brain region, projection, and molecule reports inserted in the public part of BAMS. These comments may be accessed and viewed at any

Reports of presence of CRH in PVHmpd.				Time dependence of CRH				
Reports of CRH presence in PVHmpd, normal physiological state	Watts A.G. & Sanchez-Watts G., 2002 : Interactions between heterotypic stressors and corticosterone reveal integrative mechanisms for controlling corticotropin-releasing hormone gene expression in the paraventricular nucleus. Data						Methods comparison	
	Watts A.G. & Sanchez-Watts G., 1995 : Physiological regulation of peptide messenger RNA colocalization in rat hypothalamic paraventricular medial parvocellular neurons. Data							
	Swanson L.W., Sawchenko P.E., Lind R.W. & Rho J.-H., 1987 : The CRH Motoneuron: differential peptide regulation in neurons with possible synaptic, paracrine, and endocrine outputs. Data							
Reports of CRH presence in PVHmpd, manipulated physiological state	Watts A.G. & Sanchez-Watts G., 2002 : Interactions between heterotypic stressors and corticosterone reveal integrative mechanisms for controlling corticotropin-releasing hormone gene expression in the paraventricular nucleus. Data						Methods comparison	
	Watts A.G. & Sanchez-Watts G., 1995 : Physiological regulation of peptide messenger RNA colocalization in rat hypothalamic paraventricular medial parvocellular neurons. Data							
CRH coexpression reports in PVHmpd, normal physiological state	Watts A.G. & Sanchez-Watts G., 1995 : Physiological regulation of peptide messenger RNA colocalization in rat hypothalamic paraventricular medial parvocellular neurons. Data							
CRH coexpression reports in PVHmpd, manipulated physiological state	Watts A.G. & Sanchez-Watts G., 1995 : Physiological regulation of peptide messenger RNA colocalization in rat hypothalamic paraventricular medial parvocellular neurons. Data							
Sawchenko P.E., Swanson L.W. & Vale W.W., 1984 : Corticotropin-releasing factor: co-expression within distinct subsets of oxytocin-, vasopressin-, and neurotensin-immunoreactive neurons in the hypothalamus of the male rat. Data								
Data collated from: Swanson L.W., Sawchenko P.E., Rivier J., Vale W.W., 1983								
Physiological condition	Cell pool position in region	Hemisphere	Qualitative density of CRH in PVHmpd	Labeled cells count	Percentage labeled cells	Relative to basal	Annotation	Collator
Chemical treatment: colchicine	everywhere	bilateral	not assigned	808	40.50	not measured	Collator note: see Figure 2 page 168. The percentage of labeled cells was calculated using the numbers provided in the legend of Figure 2 and represents the ratio between the labeled cells in the associated region and the total number of labeled cells in this experiment.	Mihail Bota Metadata ✓
Structural manipulation: ablation	not known	bilateral	exists	0	0.00	significantly higher	...90% of the brightly labeled cells were found in the medial part of the parvocellular division. Collator note: see Figure 1b page 167. We assigned the CRH label as significantly higher than the normal state according to Swanson 1991 in Progress in Brain Research.	Mihail Bota Metadata
Manipulated structure: Adrenal gland								

immunohistochemistry data

Fig. 8. Users can view detailed reports about the presence of a molecule in a particular CNS region. Experimental data is grouped in four categories, which depend on the physiological state of the animal, and coexpression data. Records verified by original authors are indicated with a green checkmark.

time by the registered user who inserted them. The BAMS workspace now becomes the place where registered users can save and view reports of interest concerning CNS regions, their input and output axonal connections, customized connection matrices (Bota et al., 2005), molecules, and groups of gene expression

pattern experiments. The number of reports and matrices that can be used by registered users is unlimited. An example of the types of information that can be saved in the BAMS workspace is shown in Fig. 14.

The BAMS workspace, combined with the option of adding comments to data reports,

Experiments comparison			
Reference: Watts A.G. & Sanchez-Watts G. 2002	Reference: Watts A.G. & Sanchez-Watts G. 1995	Reference: Swanson L.W., Sawchenko P.E., Rivier J., Vale W.W. 1983	
Experiment acronym: Watts-CRH (PEG) Data types: qualitative Mapping approach: brain region in a BAMS nomenclature Mapping details: Coordinates: none Anatomical data presentation: text representative labeled images	Experiment acronym: Watts-CRH/AVP/PEG Data types: quantitative qualitative Mapping approach: brain region in a BAMS nomenclature Mapping details: Coordinates: none Anatomical data presentation: text representative labeled images representative drawings/mappings onto templates	Experiment acronym: Swanson-1983(CRH) Data types: quantitative qualitative Mapping approach: brain region in a BAMS nomenclature, constructed superior to the experiment Mapping details: Coordinates: none Anatomical data presentation: text representative labeled images drawings/template mappings of all sections	Experiment acronym: PS439 Data types: quantitative qualitative Mapping approach: brain region in a BAMS nomenclature, constructed superior to the experiment Mapping details: Coordinates: none Anatomical data presentation: text representative labeled images drawings/template mappings of all sections
Animals (subjects) Number: not specified Sex: M Weight: 225-250g Housing conditions: normal housing Annotation: Adult male Sprague Dawley rats (225-250 gm body weight at the beginning of the experiment) were maintained on a 12 hr light/dark photoperiod (lights on at 6:00 AM) with ad libitum access to water and rat chow and were allowed at least 5d of acclimation to the animal quarters.	Animals (subjects) Number: 6 Sex: M Weight: 280-320g Housing conditions: normal housing Annotation: Adult male Sprague-Dawley rats (280-320 g BW at injection) were maintained on a 12 hour light/ 12 hour dark photoperiod (lights on 0700 hours) with water and rat chow available ad libitum. They were allowed 7 days' acclimation to the animal quarters before we proceeded with the experiment.	Animals (subjects) Number: 10 Sex: M Weight: 0 Housing conditions: normal housing Annotation: The second group consisted of animals that received a single injection of colchicine (4-8 microgram/microliter) saline into either the lateral ventricle (100-200 micrograms, n=8) or the fourth ventricle (50 micrograms, n=2) the day before perfusion	Animals (subjects) Number: 3 Sex: M Weight: 0 Housing conditions: normal housing Annotation: The third group was bilaterally adrenalectomized and allowed to survive for 3 days, 1 month and 2 months before perfusion (n=3).
Method: Neuron/glia identification method: not specified Staining frequency: 1:8 Technique: in situ hybridization	Method: Neuron/glia identification method: not specified Staining frequency: 1:8 Technique: in situ hybridization	Method: Neuron/glia identification method: not specified Staining frequency: 1:4 or 1:8 Technique: immunohistochemistry	Method: Neuron/glia identification method: not specified Staining frequency: 1:4 or 1:8 Technique: immunohistochemistry
Measurement 1: Measured nucleic acid: not specified Source (producer): not specified Probe sequence: not specified Sequence species: not specified Probe sequence orientation: not specified Control: not specified Labelling method: not specified Visualization method: autoradiogram Visualization medium: not specified Annotation: not specified	Measurement 1: Measured nucleic acid: not specified Source (producer): not specified Probe sequence: 700 bp RsaI-RsaI CRH Sequence species: not specified Probe sequence orientation: not specified Control: not specified Labelling method: not specified Visualization method: DIG Visualization medium: not specified Annotation: not specified	Measurement 1: Antigen: not specified Antigen species: not specified Source (producer): not specified Primary antibody: not specified Primary antibody species: not specified Antibody type (monoclonal, polyclonal): not specified Secondary antibody: not specified Secondary antibody species: not specified Immunoglobulin class: not specified Control: not specified Visualization method: fluorescence Visualization medium: not specified Annotation: not specified	Measurement 1: Antigen: not specified Antigen species: not specified Source (producer): not specified Primary antibody: 24 (Try22, Gly23)-CRF (1-23) antiserum Primary antibody species: not specified Antibody type (monoclonal, polyclonal): not specified Secondary antibody: not specified Secondary antibody species: not specified Immunoglobulin class: not specified Control: not specified Visualization method: fluorescence Visualization medium: not specified

Fig. 9. Users may compare metadata associated with experimental data and experiments for the presence of a molecule in a CNS region of interest.

brings a new level of interactivity to the system and allows registered users to annotate reports according to their expertise.

Inserting Data in BAMS's "Molecules" Module

The "Molecules" module's "Personal account" functionality allows registered users to insert molecular data from published references or unpublished experiments, update inserted records, view inserted data, combine unpublished experiments with published references, and create reports from references and experiments of interest. Registered BAMS users

may thus be collators, and/or authors, and/or curators. Insertion of molecular data in a "Personal account" can be performed with a broad range of complexity: from very simple reports of presence or absence in a CNS region, to expression in different neuron types, to very complex reports when all metadata are specified. The structure of "Molecules" module's "Personal account" (see Fig. 3 for its knowledge representation) also allows updating previously inserted information when adding new records, and the addition of new fields. An example of data entry in a "Personal Account" is shown in Fig. 15.

Temporal dynamics data of **CBH** in: PVltprnd

Summary of data

male rats	manipulated state	collated from Swanson, L.W. & Simmons, D.M., 1989, J Comp Neurol
male rats	basal state	collated from Watts, A.G. & Swanson, L.W., 1989, Endocrinology
female rats	basal state	collated from Watts, A.G. & Swanson, L.W., 1989, Endocrinology

Basal state
Data collated from [Diurnal variations in the content of proopiomelanocortin-releasing hormone messenger ribonucleic acids in the hypothalamic paraventricular nucleus of rats of both sexes as measured by in situ hybridization](#), authors [Watts, A.G. & Swanson, L.W.](#), published in [Endocrinology](#).

Protocol details: Experiment Acronym: **expmale** (the acronym was inserted by collator)
Sex: Male

Employed technique: mRNA hybridization

	Position of the cell pool within the region	Hemisphere	Relative optical density	Standard error optical density	Statistical significance (go with the pointer over the symbols to view associated annotations)	Associated annotation	Collator
6-7 (h)	everywhere	bilateral	73.0	3.0	— compared to 12-13 h — compared to 24-1 h ★ compared to 24-1 h	Between 0600-0700 and 1200-1300 h in both sexes the level of hybridization was approximately 70% of that in adrenalectomized controls. Collator note: data shown in Figure 2 page 1736, distribution of CBH mRNA is shown in Figure 1, page 1736.	Mihail Bota
12-13 (h)	everywhere	bilateral	72.0	2.0	— compared to 18-19 h — compared to 6-7 h ★ compared to 24-1 h	Between 0600-0700 and 1200-1300 h in both sexes the level of hybridization was approximately 70% of that in adrenalectomized controls. Collator note: data shown in Figure 2 page 1736, distribution of CBH mRNA is shown in Figure 1, page 1736.	Mihail Bota
18-19 (h)	everywhere	bilateral	65.0	3.0	Statistical significance across experiments. Compared experiment: expfemale; Sex: female. Measured variables: Relative optical density: 34.0 Standard error optical density: 3.0 — compared to 6-7 h	The level of hybridization in male rats declined steadily throughout the day, until the midnight was 55% of the adrenalectomized control value and significantly lower (P<0.05) than 0600-0700 h.	Mihail Bota
24-1 (h)	everywhere	bilateral	55.0	5.0	— compared to 6-7 h — compared to 1(d), 6-7 (h) — compared to 1(d), 12-13 (h)	The level of hybridization in male rats declined steadily throughout the day, until the midnight was 55% of the adrenalectomized control value and significantly lower (P<0.05) than 0600-0700 h.	Mihail Bota

Fig. 10. The display of time-dependent chemoarchitectonic data is grouped by experiment, reference, and physiological state of the subject. Users can view qualitative assessments, quantitative measurements, and annotations associated with each retrieved record, as well as the statistical significance of each experimental point. The retrieved statistical significance is graphically coded and associated with quantities compared within or across experiments.

Transfer of inserted data and metadata from the “Personal account” to the public part of BAMS is performed according to the “Molecules” Module curation policy. On completing the insertion of a reference, the collator verifies its accuracy and completeness and then contacts the reference’s authors, if possible. After the curator checks the reference and its associated records and metadata, they are transferred to the public part of BAMS, without being erased from the collator’s “Personal account.” Once the data is transferred to the public part of BAMS, collators are notified of this action.

Collators may also combine unpublished experiments with data from published references. In this case, curators do not check the consistency and completeness of records inserted by authors. The “Personal account” functionality also allows creating reports from experiments of interest to an individual collator. This option is similar to the one implemented in BAMS’s public part.

Inserted Data

The BAMS “Molecules” module contains more than 3000 reports about the presence of 18 molecules in the rat CNS. Because all molecular

Inferred chemoarchitecture pattern of the Paraventricular nucleus of the hypothalamus (PVH)

Click on symbols to access detailed reports

	PVHp	PVHm	PVHd	PVHpV	PVHmpd	PVHap	PVHpm	PVHm	PVHam	PVHt	PVHlp	PVHdp	PVHmpv	PVHpmI	PVHpm
beta2 nAChR	+	+	+												
CRH		+		+	+	+	+	+	+	+	+	+	+		
alpha4-2 nAChR	-	+	+												
alpha4-1 nAChR	-	-	+												
Oxy				+	+	+	+	+	+	+		+			
VAS				+	+	+	+	+	+	+		+	+		
SS				+	+	+	+	+	+	+		+			
met-ENK				+	+	+	+	+	+	+		+			
leu-ENK				+	+	+	+	+	+	+					
GSOX				+	+	+			+	+	+	+	+	+	+
ENK					+										
NT				+	+	+									
MCH					+										
H/O					+						+				

+ molecule present in basal physiological state
 + molecule present in manipulated physiological state
 - molecule absent in normal physiological state

Fig. 11. The result of a BAMS gene expression pattern reconstruction dealing with the various subdivisions of the rat paraventricular nucleus of the hypothalamus. The reconstructed chemoarchitecture profile includes presence or absence of molecules in both normal and manipulated experimental conditions. Users can view details of the associated reports by clicking on the symbols in each cell. A similar output is generated for reports created by users from experiments of interest.

data inserted so far in BAMS originated from our laboratory, or from research groups that use the Swanson-1998 rat CNS nomenclature for mapping (Swanson, 1998), they were registered to this nomenclature in an internally consistent way.

Overall, BAMS contains 11 comprehensive CNS nomenclatures from five species: human, macaque (*Macaca fascicularis*), cat, rat, and mouse; approx 40,000 reports of neuroanatomical projections as collated from the literature since 1962, and related generally to

the visual and limbic systems of the rat; 4000 gene expression reports, and data referring to 175 neuronal cell types identified in 50 rat brain regions that are defined in the Swanson-1998 nomenclature (Swanson, 1998). This data was inserted in BAMS by five collators and two curators.

Conclusions

This article describes the structure and major features of BAMS's "Molecules" module, which has been designed for the online handling of

Species: Rat

Compared molecules: corticotropin-releasing hormone (CRH), oxytocin (Oxy), vasopressin (VAS)

Brain regions	Physiological condition	Data about presence of CRH in PVHmpd manipulated state:													
		Physiological condition	Cell pool position in region	Hemisphere	Qualitative density of CRH in PVHmpd	Labeled cells count	Percentage labeled cells	Average labeled cells	Standard deviation	Relative to basal	Annotation	Reference	Collator		
PVHmpd	CRH	basal state													
	Oxy	basal state													
	VAS	basal state													
	CRH/Oxy	manipulated state	Chemical treatment: colchicine	everywhere	bilateral	not assigned	808	40.50	0	not measured	not measured	Collator note: see Figure 2 page 168. The percentage of labeled cells was calculated using the numbers provided in the legend of Figure 2 and represents the ratio between the labeled cells in the associated region and the total number of labeled cells in this experiment.	Swanson LW, Sawchenko PE, Rivier J, Vale WW, 1983	Mihail Bota	Metadata ✓
	CRH/VAS	basal state													
PVHmp	CRH	manipulated state													
	Oxy	basal state													
	VAS	basal state													
	CRH/Oxy	manipulated state	Chemical treatment: colchicine	unclear	bilateral	not assigned	0	0.00	446	50.00	not measured	Collator note: experiment CRH/AVP, Table 3 page 506. Page 505: eighteen to twenty hours after an injection with FEG, there was a significant 22% (P<0.01, Fig 1) increase of the CRH mRNA in the PVHmpd compared to vehicle-injected controls when measured using the 35S-cRNA probe.	Watts, A.G. & Sanchez-Watts, G., 1985	Mihail Bota	Metadata
	CRH/VAS	manipulated state													
PVHsp	CRH	manipulated state													
	Oxy	basal state													
	VAS	basal state													
PVHsp	CRH/Oxy	manipulated state	Structural manipulation: ablation												
	CRH/VAS	manipulated state	Manipulated structure: Adrenal gland	not known	bilateral	exists	0	0.00	0	not measured	not measured	90% of the brightly labeled cells were found in the ... medial part of the parvocellular division. Collator note: see Figure 1b page 167. We assigned the CRH label as significantly higher than the normal state according to Swanson 1991 in Progress in Brain Research.	Swanson LW, Sawchenko PE, Rivier J, Vale WW, 1983	Mihail Bota	Metadata
	CRH	manipulated state													
PVHsp	Oxy	basal state													
	VAS	basal state													
	CRH	manipulated state	Chemical treatment: polyethylene glycol	everywhere	left hemisphere	very strong	0	0.00	0	not measured	significantly higher	Collator note: see Figure 2 page 6284 and Figure 3, page 6285.	Watts, A.G. & Sanchez-Watts, G., 2002	Mihail Bota	Metadata
PVHmp	Oxy	manipulated state													
	VAS	basal state													
CRH/Oxy	manipulated state														

Legend: immunohistochemistry data, in situ hybridization data

Fig. 12. Users may compare the existence of particular molecules in different CNS regions. The result is a list of regions that are associated in BAMS with reports of all the searched molecules (corticotropin-releasing hormone [CRH], oxytocin [Oxy], and vasopressin [VAS]), in either basal physiological state or manipulated state. Coexpression data is also returned for all pairs of molecules in the set of interest. Users can access reports of molecule presence in the retrieved CNS regions.

molecular and gene expression data associated with different CNS regions or cell types, and collated from the literature or from a collator's personal unpublished results.

The first major new contribution is the "Molecules" module knowledge representation, which completely separates data from metadata in BAMS and is detailed enough to allow association of different types of metadata

with a collated reference. At the same time, the knowledge representation now implemented in "Molecules" is general enough to allow extensions with other gene expression techniques, and to other types of neuroscientific experimental data. Thus, the classes "Experiment metadata," "Data presentation metadata," and "Mapping metadata" can be applied to a wide range of experiments on different parts of the

Region where PEG was injected	Brain region	Molecule	Manipulated state data	Basal state data
Systemic	Paraventricular nucleus of the hypothalamus, parvicellular division, medial parvicellular part, dorsal zone (PVHpmid)	enkephalin	Watts, A.G. & Sanchez-Watts, G., 1995 , Average number of labeled cells: 201 Standard error: 25.00	Watts, A.G. & Sanchez-Watts, G., 1995 , Average number of labeled cells: 105 Standard error: 15.00
		neurotensin	Watts, A.G. & Sanchez-Watts, G., 1995 , Average number of labeled cells: 104 Standard error: 15.00	no basal state data
		corticotropin-releasing hormone	Watts, A.G. & Sanchez-Watts, G., 2002 , Relative absorbance: 125 + Standard error: 12.00	Watts, A.G. & Sanchez-Watts, G., 1995 , Average number of labeled cells: 488 Standard error: 70.00
			Watts, A.G. & Sanchez-Watts, G., 2002 , Relative absorbance, basal state: 100, p value: 0.0200. Annotation: Collator note: see Figure 2 page 6284.	Watts, A.G. & Sanchez-Watts, G., 1995 , Average number of labeled cells: 390 Standard error: 76.00
		coexpression: neurotensin and corticotropin-releasing hormone	Watts, A.G. & Sanchez-Watts, G., 1995 , Percentage expression NT/CRH: 64.50 Standard error percentage expression NT/CRH: 7.00	Swanson, L.W., Savchenko, P.E., Lind, R.W. & Rho, J.H., 1987 , Qualitative strength: exists Watts, A.G. & Sanchez-Watts, G., 2002 , Relative absorbance: 100 Standard error: 5.00

Fig. 13. Users can retrieve chemoarchitectonic data associated with a chemical or structural manipulation and compare the presence of molecules in basal and manipulated states.

vertebrate CNS. The knowledge representation's hierarchical structure also allows extension of metadata that describe records about molecules inserted in BAMS, and association of new metadata classes for the description of new experimental techniques.

The curation policy now implemented in BAMS allows rigorous collation of data from published references, and full documentation of actions performed by collators. The problems of data accuracy and consistency, which

are very important in populating neuroinformatics databases, are addressed by the set of curation rules described here. This set requires the rigorous mapping of molecular data onto CNS nomenclatures in BAMS, and includes validation of inserted data by original authors.

However, this set of curation rules does not solve the problem of contradictory results, which are common in the neuroscience literature. Automatic evaluation of molecular experimental data is not yet possible, unlike the case

Workspace of Mihail Bota [Log out](#)

Menu

Saved brain parts, connections & molecules

Saved brain parts	Saved on		Saved output patterns	Saved on		Saved input patterns	Saved on		Molecules presence reports	Saved on	
AAA	2006-05-04 21:00:05	Delete	CEAm	2006-05-04 18:31:43	Delete	PVHmpv	2006-05-04 18:00:24	Delete	CRH	2006-05-05 12:51:07	Delete
Hi	2006-05-05 11:50:52	Delete	BSTRh	2006-05-04 18:31:25	Delete	BSTRh	2006-05-04 18:31:35	Delete	H/O	2006-05-05 12:54:20	Delete
HPF	2006-05-05 21:16:10	Delete	TR	2006-05-05 19:50:18	Delete	CEAm	2006-05-04 18:31:52	Delete			
						AAA	2006-05-04 20:44:36	Delete			

Saved projections matrices & gene expression experiments

Saved projections matrices	Saved on		Saved gene expression experiments	Saved on	
PVH-output	2006-05-07 14:18:46	Delete	watts-1	2006-05-05 19:32:54	Delete
PVH-input	2006-05-07 14:20:22	Delete	watts-2	2006-05-05 19:33:30	Delete
			Swanson-group	2006-05-05 21:23:49	Delete

Fig. 14. The BAMS workspace allows registered users to store and view reports of interest concerning CNS parts and cell types, input and output axonal connections, and a wide variety of molecules—as well as complex, axonal connection matrices and groups of gene expression patterns.

for neuroanatomical axonal connection reports (Bota and Arbib 2004), because experimental techniques are quite diverse and complicated, and results also depend on experimental conditions like time of day and behavioral state. Moreover, techniques and methods used for gene expression mapping continue to evolve. Thus, potential automatic reliability evaluation will have to change accordingly.

Nevertheless, accuracy evaluation for molecular experimental data can be performed in BAMS. For this, users compare metadata associated with experiments, according to their expertise. We present here a relatively comprehensive and extendable set of metadata that describe the most widely used experimental techniques, a range of experiment details, and the richness of data collated from the literature—all

of which can be accessed by users for the comparison and evaluation of gene expression and other molecular data. Moreover, users can generate reports from data recorded in BAMS by choosing only gene expression experiments that are accurate—according to their expertise. Finally, registered users may annotate data with comments, and save the results of this activity in their personal workspace for future use. BAMS to our knowledge is the first online neuroinformatics system that allows users to interact with the system by annotating inserted information and storing the results of their searches.

The structure of BAMS's "Molecules" module allows insertion of complex reports about molecules identified in CNS parts or expressed in different cell types. It also allows construction of a web interface enabling users to search for

BAMS Molecules Module

Update measurements	Insert metadata, this experiment	Insert new experiment, this reference	Insert new reference	Main menu	Log out
-------------------------------------	--	---	--------------------------------------	---------------------------	-------------------------

Reference: Organization of angiotensin II immunoreactive cells and fibers in the rat central nervous system > Experiment: Exp-colchicine rats > Inserted measurements

Expressed molecule: angiotensin

Inserted measurements

Brain Region	Physiological condition	Location within region	Hemisphere	Staining intensity (qualitative)
Lateral hypothalamic area(LHA)	Chemical treatment: colchicine	dorsal	left hemisphere	light/moderate
Paraventricular nucleus of the hypothalamus, magnocellular division, medial magnocellular part(PVH _{mm})	Chemical treatment: colchicine	everywhere	left hemisphere	moderate/strong
Paraventricular nucleus of the hypothalamus, parvicellular division, medial parvicellular part, dorsal zone(PVH _{mpd})	Chemical treatment: colchicine	everywhere	left hemisphere	moderate/strong
Paraventricular nucleus of the hypothalamus, magnocellular division, posterior magnocellular part(PVH _{pm})	Chemical treatment: colchicine	unclear	left hemisphere	exists
Paraventricular nucleus of the hypothalamus, magnocellular division, posterior magnocellular part, lateral zone(PVH _{pl})	Chemical treatment: colchicine	everywhere	left hemisphere	moderate/strong
Paraventricular nucleus of the hypothalamus, magnocellular division, posterior magnocellular part, medial zone(PVH _{pm})	Chemical treatment: colchicine	everywhere	left hemisphere	moderate/strong
Paraventricular nucleus of the hypothalamus, parvicellular division, periventricular part(PVH _{pv})	Chemical treatment: colchicine	everywhere	left hemisphere	exists
Supraoptic nucleus(SO)	Chemical treatment: colchicine	everywhere	left hemisphere	light/moderate

immunohistochemistry data

Fig. 15. The BAMS “Personal account” allows registered users to insert, update, and view molecular data and metadata from collated references (or the results of unpublished experiments); to group unpublished experiments with new published references; and to create reports about experiments of interest.

information related to the chemical makeup of CNS regions of interest, run complex queries that reconstruct gene expression patterns in different CNS regions, and compare the existence and levels of molecules under specified physiological conditions.

This is part of an expanding effort in the emerging field of neuroinformatics. Several research groups have developed on-line knowledge management systems that handle data at different levels of nervous system organization: brain

region nomenclatures (Bowden and Martin, 1997; Burns, 2001; Stephan et al., 2001; Bota and Arbib, 2004), neuroanatomical projections (Burns, 2001; Stephan et al., 2001; Bota and Arbib, 2004), and cytology (Marenco et al., 1999; Bota and Arbib, 2004). Neuroinformatics systems that share similar features with BAMS include NeuroNames (<http://braininfo.rprc.washington.edu>; Bowden and Dubach, 2003), CoCoMac (<http://cocomac.org>; Kötter, 2004), and the NeuronDB database ([Neuroinformatics](http://senselab.med.</p>
</div>
<div data-bbox=)

yale.edu/senselab/NeuronDB; Marenco et al., 1999). However, none of these systems include a representation of molecular data identified in various CNS parts and cell types. BAMS appears to be the first on-line neuroinformatics system that manipulates neuroscience data across four levels of CNS organization: molecules, cell types, brain regions, and networks of brain regions.

The present status of BAMS is “in progress” because we continue to add new data and extend its functionality. High priority future developments of the “Molecules” module include extension of the present molecule classification schema, molecule representation at the level of synapses (and thus terminal fields), and inference engines for automatic comparison of molecular data according to the age, sex, and other metadata of experimental animals. Finally, implementation of webservice protocols for providing data and process queries submitted in software applications developed by other research laboratories (MacKenzie-Graham et al., 2004) is a top priority. BAMS is already a data provider for a series of neuroinformatics systems through URL links (NeuroNames), middlelayer applications (LONI), and backend MySQL connections (NeuronDB, LONI, WebQTL). Our goal is to continue expanding the list of neuroinformatics and bioinformatics systems to which BAMS acts as data provider or client, including gene expression databases developed by the Allen Institute for Brain Science (<http://www.brainatlas.org/>) and Gensat (<http://www.gensat.org/>).

Acknowledgments

This work was supported by NIH Grants MH61223, NS16668, and NS050792-01. We especially thank Drs. Alan Watts and Arshad Khan for their suggestions and feedback.

References

Bota, M., Dong, H. W., and Swanson, L.W. (2003) From gene networks to brain networks. *Nature Neurosci.* 6, 795–799.

- Bota, M., Dong, H. W., and Swanson, L.W. (2005) The Brain Architecture Management System. *Neuroinformatics* 3, 15–48.
- Bota, M. and Arbib, M. A. (2004) Integrating Databases and Expert Systems for the Analysis of Brain Structures: Connections, Similarities and Homologies. *Neuroinformatics* 2, 19–58.
- Bowden, D. M. and Martin, R. F. (1997) A digital Rosetta stone for primate brain terminology. In: Bloom, F. E., Bjorklund, A., and Hokfelt, T. (eds.) *Handbook of Chemical Neuroanatomy vol. 13: The Primate Nervous System, part I*, Academic Press, San Diego, pp. 1–37.
- Bowden, D. M. and Dubach, F. M. (2003) NeuroNames 2002. *Neuroinformatics* 2, 63–83.
- Burry, R. W. (2000) Specificity controls for immunocytochemical methods. *J. Histochem. Cytochem.* 48, 163–165.
- Burns, G. A. P. C. (2001) Knowledge mechanics and the NeuroScholar project: a new approach to neuroscientific theory. In: Arbib, M. A. and Grethe, J. (eds.) *Computing the Brain: A Guide to Neuroinformatics*, Academic Press, San Diego, pp. 319–336.
- Burt, D. R. (1985) Criteria for receptor identification. In: Yamamura, et al. (eds.) *Neurotransmitter receptor binding*, 2nd ed., Raven Press, New York, pp. 41–60.
- Kötter, R. (2004) Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac Database. *Neuroinformatics* 2, 127–144.
- Kuhar, M. J. (1985) Receptor localization with the microscope, in Yamamura et al. (eds.) *Neurotransmitter receptor binding*, 2nd ed., Raven Press, New York, pp. 41–60.
- Kuhar, M. J., De Souza, E. B., and Unnerstall, J. R. (1986) Neurotransmitter receptor mapping by autoradiography and other methods. *Annu. Rev. Neurosci.* 9, 27–59.
- Marenco, L., Nadkarni, P., Skoufos, E., Shepherd, G., and Miller, P. (1999) Neuronal database integration, the Senselab EAV data model, *Proceedings of AMIA Symposium*, pp. 102–106.
- MacKenzie-Graham, A., Lee, E. F., Dinov, I. D., et al. (2004) A multimodal, multidimensional atlas of the C57BL/6J mouse brain. *J. Anat.* 204, 93–102.
- MacKenzie-Graham, A., Jones, E. S., Shattuck, D. W., Dinov, I. D., Bota, M., and Toga, A. W. (2003) The informatics of a C57BL/6J mouse brain atlas. *Neuroinformatics* 1, 397–410.
- Saper, C. B. (2005) An open letter to our readers on the use of antibodies. *J. Comp. Neurol.* 493, 477–478.

- Saper, C. B. and Sawchenko, P. E. (2003) Magic peptides, magic antibodies: guidelines for appropriate controls for immunohistochemistry. *J. Comp. Neurol.* 465, 161–163.
- Stephan, K. E., Kamper, L., Bozkurt, A., Burns, G. A. P. C., Young, M. P., and Kötter, R. (2001) Advanced database methodology for the Collation of Connectivity data on the Macaque brain (CoCoMac). *Philos. Trans. R. Soc. London B. Biol. Sci.* 1159–1186.
- Swanson, L. W. (1998) *Brain Maps: Structure of the Rat Brain*. Second Edition. Elsevier, San Diego.
- Swanson, L. W. (2000a) What is the brain? *Trends Neurosci.* 23, 519–527.
- Swanson, L. W. (2000b) A history of neuroanatomical mapping. In: *Brain Mapping: The Applications*, Toga A. W. and Mazziotta J. C. (eds.) Academic Press, San Diego, pp. 77–109.
- Swanson, L. W. and Simmons D. M. (1989) Differential steroid hormone and neural influences on peptide mRNA levels in CRH cells of the paraventricular nucleus: a hybridization histochemical study in the rat. *J. Comp. Neurol.* 285, 413–435.
- Wada, E., Wada, K., Boulter, J., et al. (1989) Distribution of alpha 2, alpha 3, alpha 4, and beta 2 neuronal nicotinic receptor subunit mRNAs in the central nervous system: a hybridization histochemical study in the rat. *J. Comp. Neurol.* 284, 314–335.
- Watts, A. G. and Sanchez-Watts, G. (2002) Interactions between heterotypic stressors and corticosterone reveal integrative mechanisms for controlling corticotropin-releasing hormone gene expression in the rat paraventricular nucleus. *J. Neurosci.* 22, 6282–6289.
- Watts, A. G. and Swanson, L. W. (1989) Combination of *in situ* hybridization with histochemistry and retrograde tract-tracing. *Methods in neurosciences*, vol. I, Academic Press, San Diego, pp. 127–136.