Contract: LSHG-CT-2006-037277

VALAPODYN

Validated Predictive Dynamic Model of Complex Intracellular Pathways Related to Cell Death and Survival

FP6 - Specific Targeted Research Projects (STREP)

Priority T1 – Life Sciences

D6.5 Publishable Final Activity Report

Period covered: from October 1st, 2006 to March 31st, 2010
Date of preparation: May June 2010
Start date of project: October 1st, 2006 Duration: 42 Months

Project coordinator name: Dr. Antoine DEPAULIS
Project coordinator organization name: INSERM

Revision: final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Approvals

	First name & Name	Organisation	Date	Visa
Coordinator	Antoine Depaulis	INSERM	30/06/2010	OK
Quality manager	Raffaella Catena	ALMA	30/06/2010	OK

Document History

Revision	Author	Modification	Date
R0	R.Catena	Creation	01/05/10
R1	Antoine Depaulis	Creation and WP5	02/05/10
R2	Olga Kel-Margoulis	WP1 and WP3	10/06/10
R3	David Greenberg	WP4	14/06/10
R4	Raffaella Catena	WP6	15/06/10
R5	Despina Sanoudou	WP2	18/06/10
R6	Todor Vujasinovic	WP1 and WP3	22/06/10
R7	Jean-Baptiste Dumas	All+ IP	23/06/10
Final	Antoine Depaulis	Final revision	30/06/10

TABLE OF CONTENT

1 Summary of project results

The main objective of Valapodyn was to show the feasibility of developing a new Systems Biology approach to model the dynamics of Molecular Interaction Networks (MIN) related to neuronal death and survival, and to perform *in silico* simulations to identify new potential drug targets to treat neurodegeneration (NDG). The experimental model used was obtained by intracerebral injection in mice of kainate, a neurotoxic compound, known to result in NDG. The development of the MIN model required the production of accurate experimental data that were obtained after high-throughput RNA and protein analysis of brain samples at 10 different time-points. This data was fed into the model, along with data collected from an exhaustive compilation of current molecular knowledge on NDG derived from the literature and organized in established databases.

This led to the first version of the dynamic MIN model of kainate-induced neuronal death which described the behaviour of a large signalling network including 521 molecules (genes/proteins) connected by 3,069 direct and oriented interactions that are all downstream of the kainate signal. The model obtained has over 4,500 parameters and is one of the largest dynamic models of a signalling network ever published, with an overall relative error of 16%. The simulations performed with this model identified a list of 9 potential therapeutic targets. Out of these targets, 3 were selected for either (i) their coherence with the existing literature (i.e., brain derived neurotrophic factor or BDNF), (ii) the possibility to inhibit their action with already existing chemical compounds (Vala09) or (iii) their innovative aspects and potential "added value" for development as a therapeutic target (Vala01). Inhibition of Vala09 by a chemical compound led to an aggravation of kainate-induced NDG. In contrast, inhibition of the expression of BDNF or of Vala01 by local application of shRNA-lentivirus, reduced kainate-induced NDG, as indicated by 3 different biomarkers of neuronal death.
These results validate the proof of principle of the relevance of the Systems Biology Approach in modelling the dynamics of MIN and identifying innovative therapeutic targets in Neurology research. They will allow us to improve the model and generate more appropriate targets.

2 Contractors involved

The VALAPODYN consortium consists of 7 partners, from 5 European countries. It utilises expertise developed during many years of research in different complementary fields including genomics (Academy of Athens FBRAA), proteomics (University of Liege ULg), molecular interactions (HELIOS Biosciences SARL, BIOBASE GmbH), RNA interference (Hebrew University of Jerusalem HUJI), Neuroscience (Institut National de la Santé et de la Recherche Médicale INSERM) and management expertise (Alma Consulting Group ALMA).

3 Scientific approach

To achieve its ambitious objectives, the VALAPODYN work programme has been organized into 6 scientific work packages:

- 3 work packages (WP1-3) were dedicated to large scale data collection, the preparation and construction of the dynamic models as well as bioinformatics analysis of high throughput data generated by the consortium.
- 1 work package (WP4) was dedicated to the validation of the dynamic models
- 1 work package (WP5) was dedicated to the preparation of the biological material for the construction of the dynamic models and for their validation.

- 1 work package (WP6 – Project Management) was dedicated to the management of the consortium and of knowledge (intellectual property rights, dissemination activities and exploitation plan).

The objectives of the technical WPs are detailed below:

WP1: Adaptation / construction of intracellular cascades:
WP1 was aimed to adapt the network of interactions to the pathology studied. The molecular interaction networks (MIN) of both signal transduction events as well as transcription regulatory events occurring in the central nervous system was investigated. Relevant information in the literature was collected, including the scanning of appropriate and curated databases. The objective was to build a network reflecting the biological process which will be modelized.

WP2: Collection of the relevant biological data for construction of the dynamic model and design of an integrated database:
This WP included the generation of gene and protein expression profiles and evaluation of the activity of selected cascade proteins.

WP3: Construction of the dynamic models & simulation / selection of therapeutic targets:
This work package aimed at integrating known signalling cascades & novel biological data to train the dynamic model at selecting therapeutic targets from different simulation conditions. The output of this model was a selection of putative therapeutic targets that was then validated in WP4.
Another aim of this work package was to provide bioinformatics support for the Consortium, to organize genomics and proteomics data generated by the consortium in a database and to perform advanced bioinformatical analysis.

WP4: Validation of predictive value of model
The aim of this WP was to conduct RNA interference in cell culture and *in vivo* in order to verify the effect of limiting or suppressing expression on neurodegeneration - i.e., the neuroprotective potential

WP5: Biological Material
This WP was dedicated to provide the most appropriate biological material for the realization of WP2 and WP4.

4 Work performed and achievements

4.1 WP1 – Adaptation / construction of intracellular cascades:

1 State-of-the art at project start

Experimental data on molecular cascades in neuronal cells are being actively accumulated by the scientific community. However, published data are dispersed through different journals, and not systematically available in a computational format suitable for modelling.

2 Project objectives

Therefore, the aim of WP1 was to systematically collect and collate relevant information as molecular interaction networks in a database, and to make it easy to use for the consortium members.

3 Project achievements

Information was systematically collected on genes and proteins that are already known to be connected with brain diseases associated with cell loss such as some forms of epilepsies (e.g., mesiotemporal lobe epilepsy), Alzheimer and Parkinson diseases. In particular, the following major topics were added or highly enriched in the TRANSPATH database during the course of the project: e.g., signaling via the ionotrophic glutamate receptors (NMDARs, AMPARs, kGluRs); neurotrophic signaling induced by ligands like BDNF, NGF and NT-3, -5 which activate the trk-A, -B, -C and p75NTR receptors; Alzheimer Precursor Protein (APP) regulation; cell death *versus* survival processes.

Data collection from literature (curation process) at BIOBASE was done through manual extraction of information upon full text paper reading by scientists. In-house curation technology involved a dedicated program called TRANSTOOL, which provides special input forms for different entities and is equipped by controlled vocabularies, consistency checks along with a number of auxiliary functions to enable effective and correct data input. The curation strategy privileged "canonical" pathway-oriented molecules and their interactions. Data were collected on direct protein interactions, modification reactions such as phosphorylation, acetylation, ubiquitination, and some other reaction types.

Published information was also collected for genes/molecules coming from the experiments performed by FBRAA and ULg in WP2 from biological samples provided by INSERM (WP5). Biobase curation strategy involved focusing on molecules that were not yet or not well curated among the differentially expressed genes and proteins from the literature.

During the project, Biobase has curated more than **600 molecules**, more than **5,000 reactions** and **8 pathways**, **with more than 50 chains** linked to them. This allowed the creation of **10** corresponding "hand-drawn" maps that were updated regularly.

Further extension of MIN by Helios

Despite the compilation of a large network (over 11,000 interactions), when the whole set of gene expression data was provided, the initial analysis of the MIN showed that it was insufficient, in particular with regards to the early time-points. Moreover, a significant percentage of the molecules covered were poorly connected (or even inputs/outputs of the network). Even when the coverage was calculated only in reference to the annotated genes (as genes with no annotation cannot be integrated in the MIN), the coverage was too low for any model to be credible: between 16% and 33% of the annotated genes modulated over 2-fold were present in the MIN. The situation was similar for the proteomics data: there was a low coverage of the modulated proteins in the initial MIN of NDG. Furthermore, there was only a small overlap between the modulated genes and

the modulated proteins, probably due to the techniques used. Only about 10% of the modulated proteins were already represented in the first version of the NDG dynamic model.

To solve this issue, Helios compiled from the literature a set of about 1,200 additional interactions between 400 molecules that are specific to neuronal intracellular signalling with regards to neurodegeneration, based on the experimental data generated during the project. Depending on the time-points, between 26% and 81% of the modulated genes at early time points were integrated in the MIN. 75% of the modulated signaling proteins at early time points (0 to 9 hours) and 40% of the modulated signaling proteins at late time points (6 to 24 hours) have been also integrated. This strategy complemented the work of Biobase on "canonical" pathways, produced a MIN that fits to the screening data of kainate NDG. This brought several non-canonical pathways and provided a modelling-grade MIN. 700 of these interactions are present in the Modelling-grade MIN used to build the dynamic model in WP3.

4.2 WP2 - Collection of the relevant biological data for construction of the dynamic model and design of an integrated database:

1 State-of-the art at project start

In order to obtain an unbiased, global view of the transcriptome and proteome of kainate injected mouse hippocampi, the consortium selected cutting-edge high-throughput genomics and proteomics approaches. Specifically, global gene expression changes were evaluated using the latest version of the Affymetrix Mouse GeneChips (containing 45,100 probe sets) together with advanced, well established, bioinformatical tools for the initial stages of image and statistical analysis.
For proteome analysis, 2D gel-based approaches were very well established technologies at ULg. The technical controls and accurate experimental tools of differential protein expression on multiple gels were set up. Modulated proteins spots were visualized with the Decyder image analysis software's using manual studies of each spots. Protein spots identification was performed using the Ultraflex II TOF/TOF MALDI mass spectrometer.

2 Project objectives

The aim was to study dissected hippocampi from kainate injected mice at a wide range of time-points (ranging from 0 to 24 hours) in a series of parallel genomics and proteomics analyses. This data served for the generation of genomic and proteomic databases depicting the global detected changes that would form the basis for the work of the other WPs and the building of the MIN.

3 Project achievements

At the gene expression level, kainate was found to have a significant effect with increasing post-injection time-intervals. The number of highly and significantly changed probe sets per time point ranged from 1 to 929 with the maximum number of changes being observed at 12 hours post-injection. The fold changes of individual probe sets ranged from 69.3 to -20.35, which were observed at 6 and 12 hours respectively. Interestingly, although the levels of over- and under-expressed probe sets were approximately equal during the earlier time points (~55% genes were under-expressed), most of the significant changes involved over-expression (~85%) at the later time points. Cross comparison of the significantly changed probe set lists from each time-point revealed that 39 probe sets were consistently changed in at least 3 time points. These data were used to generate a detailed database containing all the gene expression changes that were detectable at the specific time-points following local kainate injection.
At the proteomic level, using gel-based and non gel-based approaches, the expression of 232 different proteins were identified as modulated during the first 24 hours after local kainate injection. In the gel-based approach, 2D-DIGE modulated proteins spots were prepared according to 2 different methods and visualized by 2 images analysis software's. A spot picker workstation was used to automatically pick proteins spots from the gel thus avoiding hand contaminations. "User-

friendly" documents were constructed to efficiently communicate information. The obtained lists of proteins were extensively analysed with public databases (UniProtKB, NCBI, Pfam etc) to review interactions, biological functions, protein families with isoforms and cellular localisation.
This state-of-the-art genomics and proteomics strategy led to the identification of the molecular pathways affected during the investigated time-course and in collaboration with the WPs of the VALAPODYN project, it enabled the selection of promising new therapeutic targets to validate the dynamic modeling of NDG.

4.3 WP3 - Construction of the dynamic models & simulation / selection of therapeutic targets

1 State-of-the art at project start

The state of the art in dynamic modelling has not really changed during the course of Valapodyn and has been recently reviewed by Vujasinovic *et al.* (2010)[1] Briefly, models usually address small signalling modules of 10 to 30 genes/proteins, and only one other systematic initiative of large-scale dynamic modelling by the U.S. Company Entelos have been reported (although not published). Helios had already built middle to large scale dynamic models of lymphocyte intracellular signalling (338 molecules (Genes/proteins) and 1,498 interactions. With regards to Neuroscience and neurodegeneration, there were some very limited attempts to model small signalling modules, the most noticeable being a model of neuronal hypoxia[2]. There was no model of kainate-induced neuronal death and the underlying signalling events were not exhaustively described.

2 Project objectives

The main objective was to show the feasibility of developing a new Systems Biology approach to model the dynamics of Molecular Interaction Networks (MIN) related to neuronal death and survival, and perform *in silico* simulations to identify new potential drug targets to treat NDG. This required the production of accurate experimental data that would feed the modelling work, as well as the exhaustive compilation of current molecular knowledge on neurodegeneration. More specifically, the objective was to build such a dynamic model of kainate-induced neuronal death, and validate, at least in part, its predictive power, that would stand as a basis for the future development of models more specifically dedicated to brain pathologies such as Alzheimer and Parkinson diseases.

3 Project achievements

The first version of the dynamic model of kainate-induced neuronal death described the behaviour of a large signalling network including 521 molecules (genes/proteins) connected by 3,069 direct and oriented interactions that are all downstream of the kainate signal and has over 4,500 parameters. This made it one of the largest dynamic models of a signalling network ever built. It faithfully fits the experimental data with an overall relative error of 16%.

To specify the MIN modelled from the global MIN developed in WP1, it was assumed that the genes/proteins whose expression is significantly modulated in the experimental data determined the core of the signalling network involved in kainate-induced NDG. As these genes/proteins were not all directly connected, gaps in the network were filled by additional literature analysis (WP1) and then by systematically analysing the possible signal propagation within the global network using the specific algorithms of the PROPAGAPATH software. All feed-backs identified in the global MIN are also present. This provides a 'static network' of the signalling events underlying kainate-triggered NDG. This way of defining the biological system directly based on the experimental data generated by the consortium also highlighted additional signalling mechanisms that were not readily

[1] Vujasinovic T, Zampera AS, Jackers P, Sanoudou D, Depaulis A.: *In Silico* Dynamic Molecular Interaction Networks for the Discovery of New Therapeutic Targets. Curr Pharm Des. 2010 May 12. [Epub ahead of print] PMID: 20459389
[2] Cakir T, Alsan S, Saybasili H, et al. Reconstruction and flux analy-sis of coupling between metabolic pathways of astrocytes and neu-rons: application to cerebral hypoxia. Theor Biol Med Model 2007; 4: 48.

identified from the compilation of the literature on NGD. It also directly highlighted several mechanisms that most probably explain the toxicity of kainate in this experimental paradigm.

The kinetics of expression of each gene/protein in the MIN as described in the expression studies were used to calculate the parameters of the NDG 'dynamic' model, based on Helios Biosciences DYNAPATH software platform. During the parameter determination procedure, all kinetics of each molecule in the model were "learned" from an initial kainate trigger lasting 2 hours. The simulations then searched systematically for the genes/proteins the inhibition of which was predicted to inhibit the kainate-induced dynamic molecular cascades as defined by the screening data (WP2 and WP5). These simulations identified a set of 9 potential therapeutic targets to counteract kainate-induced neuronal death and, by extension, to potentially treat neurodegenerative diseases. This achievement validates the proof of principle of the relevance of applying dynamic systems biology to therapeutic research in Neurology.

To complete this dynamic modelling work, additional bioinformatics analyses were performed. In particular, BIOBASE used the previously developed ExPlain software platform to load, store, classify and analyse microarray and proteomics data generated in the course of WP2. Incorporated software tools include mapping putative TF binding sites on the promoter sequences, searching promoter models and finding key-nodes in the networks. Analysis of the differentially expressed genes and proteins at three time points: 6h, 12h, 24h, kainate-injection *versus* saline injection, was performed.

First, downstream analysis was done, including functional classifications of differentially expressed genes by GO terms and by the associated disease terms. Next, upstream analysis was performed, including finding common promoter models, suggesting involved transcription factors (TFs), network analysis upstream of TFs and revealing potential key nodes in networks.

Mapping differentially-regulated proteins identified by proteomics approaches in WP2 (46 proteins) plus statistically significant key-nodes to the TRANSPATH pathways suggested the following affected pathways Htt pathway, glucose and lactose metabolism, RhoA and EGF pathways, Parkin-associated pathways and neurotrophic signalling. These pathways, identified by independent bioinformatics analysis, are already known to be involved at various levels in NDG, thus validating our approach of performing a high-throughput molecular screen on a dense time series and analyse the results by Systems Biology methods.

4.4 WP4 - Validation of predictive value of model

1 State-of-the art at project start

Knock-down of the expression of the target candidate genes was done by short interfering RNA (siRNA) introduction into the central nervous system. siRNA are short 20-22 base single stranded RNA sequences that are processed by cellular protein machinery and are able to hybridize the mRNA to initiate specific degradation. siRNA are effective tools for the modulation of the mRNA expression and show relatively few side-effects. There are now widely used in molecular biology in vitro for the specific reduction of gene expression. Therefore, it was decided to evaluate the efficacy of 2 modes of delivery: (i) infection by replication of incompetent lentivirus vectors that infect the cell and express short hairpin RNA molecules that are processed by the cellular machinery to give mature 22 base-pair siRNA; (ii) the use of in vivo electroporation to deliver synthetic 22 bases siRNA to the target tissue.

2 Project objectives

The goal of this WP was to validate the gene candidates generated by the MIN models as outlined in WP1 and WP3. These candidates are the genes selected as having an important role in the neurodegenerative cascade. The working hypothesis was that reducing the level of expression of one of these genes should reduced NDG seen after intrahippocampal kainate injection in the same mouse model used in WP5.

3 Project achievements

siRNA and lentivirus vectors targeted against BDNF and acetylcholinesterase (AChE) were first developed by HUJI and tested on tissue culture in order to test (i) the efficiency of the synthetic siRNA and (ii) the ability to generate high-titre biologically active viral preparations. These vectors were able to reduce both RNA and protein levels of the target genes *in vitro*. In addition lentivirus vectors expressing GFP were produced by HUJI and injected into mice by INSERM to measure *in vivo* infectivity and knockdown effects. In parallel, experiments carried out at INSERM were done to test the efficacy of electroporation for *in vivo* transfection of synthetic siRNA into mouse hippocampus. The results of these two parallel experimental approaches showed that lentivirus infection is the most effective delivery mode to efficiently knockdown targeted protein.
Lentivirus vectors were produced by HUJI against one of the candidate genes (VALA001). These lentivirus were sent to INSERM and injected in mice treated with kainate 4 days later. The mice were killed 24h after kainate and analysed for the extent of NDG. In addition, chemical approaches were used to inhibit the product of specific genes (VALA09). It was found that chemical inhibition of the target VALA09 induced a 30% *aggravation* of NDG, whereas the suppression of VALA001 by shRNA lentivirus tended to reduce NDG. Interestingly, the reduction of expression of BDNF by shRNA-lentivirus also showed a clear reduction of NDG, suggesting that this neurotrophin may facilitate cell loss during the early processes that follow kainate injection in the hippocampus.

4.5 WP5 – Biological Material

1 State-of-the art at project start

Before the start of Valapodyn, very few studies had reported transcriptomic analysis during neurodegenerative processes. In addition, only two studies reported partial transcriptomic analysis of genes involved in epileptogenesis in a rat model[3][4][5]. However, in both cases the gene analysis was performed at very few time-points precluding any dynamic analysis of the processes involved. In addition, there was no information on protein analysis. Finally, the predictive value of the animal models used to study epileptogenesis remains questionable.

2 Project objectives

The aim of this workpackage was to provide tissue samples for mRNA and protein analysis from an animal model of local and rapid cell degeneration after intrahippocampal injection of kainate. Using this model, it was thus possible to collect tissue samples (i) of a limited part of the brain, (ii) at fixed time-points, (iii) in a limited period of time. Because neuronal death occurs mainly during the first 24h after the injection of KA, it was decided to focus on this period during the course of Valapodyn.

3 Project achievements

INSERM first provided FBRAA and ULg with hippocampal samples from mice injected in the right hippocampus with kainate or saline, as well as samples from non-injected mice. A rapid dissection protocol and a reliable shipping procedure were first implemented to allow optimal RNA extraction or protein analysis. In addition, several quality controls were implemented to ensure reliable delivery to FBRAA and ULg.
INSERM prepared and shipped samples collected at 0, 1, 2, 4, 6, 9, 12, 18h and 24h post kainate. In addition, samples collected in animals without anesthesia or injection (non-treated controls) were also analyzed. Altogether, samples from a total of 10 different time-points were collected for both mRNA and protein analyses with a total of 314 samples sent to FBRAA or Ulg.

[3] Elliott RC, Miles MF, Lowenstein DH. Overlapping microarray profiles of dentate gyrus gene expression during development- and epilepsy-associated neurogenesis and axon outgrowth. J Neurosci. 2003 Mar 15;23(6):2218-27.
[4] Lukasiuk K, Kontula L, Pitkänen A. cDNA profiling of epileptogenesis in the rat brain. Eur J Neurosci. 2003 Jan;17(2):271-9.
[5] Gorter JA, van Vliet EA, Aronica E, Breit T, Rauwerda H, Lopes da Silva FH, Wadman WJ. Potential new antiepileptogenic targets indicated by microarray analysis in a rat model for temporal lobe epilepsy. J Neurosci. 2006 Oct 25;26(43):11083-110.

New samples from kainate-injected mice were later provided to HUJI for rtPCR analysis for non-injected mice, 6h post-saline, 4h, 6h, 9h, 12 and 18h post kainate.

4.6 WP6 – Management

1 Project objectives

- Achieve the project's objectives by efficiently using the planned resources
- Fulfill contractual requirements towards the EC
- Develop and facilitate efficient communication between partners and thus optimize indirect benefits of the project

2 Project achievements

The resources of the project were efficiently managed with a regular financial and technical follow up done by the Coordinator with ALMA's support. All the contractual deliverables were sent to the EC respecting the schedule and all the project goals were successfully achieved. The Consortium asked and obtained a 6-month extension to be able to validate the *in silico* model with the identification of specific candidate genes. An intranet web-platform was provided by Alma to ensure efficient transfer of the different documents (https://www.myndsphere.com).

WP6 - Exploitation and Dissemination

1 Project objectives

- Management of knowledge and intellectual property and establishment of exploitation and dissemination strategies

2 Project achievements

A public web site (http://www.valapodyn.eu/) was created and regularly updated. The statistics revealed that it was visited by internet users from 56 different countries among which there are United States, Israel, India and different countries from Europe.
A brochure with the VALAPODYN description can be downloaded from the web site and was distributed to the main conferences/congresses of the field by all the partners.

- 23 abstracts were presented to internationally and national conferences and congresses.
- 27 peer-reviewed articles were published in international journals acknowledging VALAPODYN
- 4 publications are in preparation in high impact factor scientific journals.
- Our knowledge dissemination activities have been targeting seven different audiences: 1) scientists, 2) graduate students (n=365), 3) undergraduate students (n=209), 4) high-school students (n=683), 5) high school teachers (n=125), 6) lay public, and 7) industry. These oral presentations were given at local, national and international meetings (e.g. France, Germany, Belgium, Greece, Austria, USA, India, Japan) and they have distributed the VALAPODYN leaflet or presented posters of our work at scientific meetings (e.g. Greece, Belgium, France, Israel, Japan, USA, Spain).
- The consortium had regular meetings to discuss the exploitation strategy. As of June 30th, 2010, the Vala01 target requires further validation before being considered for exploitation.