



ICGEB International Centre for Genetic
Engineering and Biotechnology

Developing
Knowledge



ICGEB RESEARCH GRANTS PROGRAMME

RESEARCH GRANTS
COMPLETED
IN 2012

CRP - ICGEB RESEARCH GRANTS COMPLETED IN 2012

COUNTRY	PRINCIPAL INVESTIGATOR	PROJECT TITLE
ARGENTINA	O. Mario AGUILAR	Sustainable agriculture: molecular studies on the genetic basis of nodulation preference in common bean-Rhizobium coevolution
BRAZIL	José Ernesto BELIZÁRIO	Production of transgenic rabbits for hybridome-free generation of monoclonal antibodies
CHILE	Claudio HETZ	Role of the Unfolded Protein Response (UPR) in tumour suppression by caveolin-1
CHINA	Paul Kay Sheung CHAN	Genomic sequence diversity and oncogenic potential of human papillomavirus type 58 variants across the world
COLOMBIA	Silvia RESTREPO RESTREPO	Identification of pathogenicity factors in <i>Batrachochytrium dendrobatidis</i> by <i>in silico</i> comparative genomics
CROATIA	Gordana MARAVIC VLAHOVICEK	Mechanisms of aminoglycoside resistance by ribosomal RNA methylation
EGYPT	Mahmoud Khairy TAHOUN Mohamed	<i>plg-1</i> gene expressing propionin by lactic starters in dairying
HUNGARY	András PERCZEL	Assessing the protein folding - aggregation paradigm with NMR and molecular modeling
MEXICO	Dimitrios GEORGELLIS	Signal transduction by the BarA/UvrY two-component system
MEXICO	Miguel Angel GÓMEZ LIM	Human health and melon biotechnology
PAKISTAN	Sheikh RIAZUDDIN	Genetic basis of hereditary hearing impairment
PERU	Mirko J. ZIMIC	An ELISA assay for cysticercosis based on a recombinant cathepsin-L
POLAND	Jaroslav POZNANSKI	TFE as the 3D-fold inducer. Application for domain 4 of <i>E. coli</i> RNAP σ^{70}
RUSSIA	Alexander A. MAKAROV	IsoAsp7-Abeta as potential trigger of Alzheimer's disease
SERBIA	Jelena KUSIC	Structural alterations of DNA induced by initiator protein
SLOVAKIA	Peter FILIPCIK	Misfolded tau protein and oxidative defense system in tauopathies
SLOVENIA	Mateja MANCEK KEBER	Molecular mechanism of the recognition of double stranded viral RNA mediated by TLR3

SOUTH AFRICA	Bryan Trevor SEWELL	Structural studies of enzymes in the context of anti-malarial drug development
SRI LANKA	Menaka D. HAPUGODA	Performance of recombinant proteins as diagnostic intermediates for Chikungunya, Dengue and Leptospirosis
UNITED ARAB EMIRATES	Ahmed H. HASSAN	Studying chromatin-modifying proteins in transcription regulation
URUGUAY	Eduardo OSINAGA	Design of anti-tumour vaccine by glycoprotein engineering

ARGENTINA

Title: Sustainable agriculture: Molecular studies on the genetic basis of nodulation preference in common bean-Rhizobium coevolution

Principal Investigator: O. Mario Aguilar, Instituto de Biotecnología y Biología Molecular (IBBM), Facultad de Ciencias Exactas, Universidad Nacional de La Plata – CONICET, 1900 La Plata, Argentina. Tel: +54-221-4229777, Fax: +54-221-4229777, E-mail: aguilar@biol.unlp.edu.ar

ICGEB Contract No.: CRP/08/003

ICGEB Reference No.: CRP/ARG08-03

Abstract: The symbiotic association between *Phaseolus vulgaris* (common beans) and rhizobio displays mutual affinity between host and bacteria from the same geographical region of diversification. The project was addressed to identify genes of the host that are differentially expressed in response to infection by sympatric strains of *R. etli*, based in the use of cDNA analysis, gene expression, reverse genetics. In addition, the molecular signal for nodulation produced by different lineages of *R. etli* was examined by chemical analysis of the nodulation factors. We have shown that a gene coding for a small GTP binding protein, PvRabA2a acts in polar growth of root hair, and it is also required for reorientation of root hair growth axis during bacterial infection in the early stages of interaction. In addition, we have identified a gene coding for the C subunit of the transcription factor NGF-Y, and demonstrated to play a key role in determining preference for bacteria that occupy nodules. On the other hand, we found no differences in chemical structure of nodulation factors produced by lineages of *R. etli* belonging to the Mesoamerican and Andean gene pool.

Objectives:

We aim to discover and characterise common beans genes involved in the preferential nodulation towards the definition of molecular indicators of coevolution. For this, we propose to study the pattern of expression and investigate the role of selected host genes we have identified in our previous work based on a differential screening. Specifically, we will further characterise common bean genes coding for a small GTP binding protein Rab11, a transcription factor Hap5, and the chemical structure of nodulation factors produced by lineages of species *Rhizobium etli*.

Results Obtained:

Common beans (*Phaseolus vulgaris*) are the most important legume for direct human consumption. In countries such as Brazil, Mexico and African countries Burundi, Rwanda, Kenya and Tanzania, represent the major source of protein in the diet. It is the source of protein for more than 500 million people in Africa and Latin America. Latin America produces about 50% of the world production however in many cases low yields are reported due to different causes that include pests, fertility of soil, drought. As most legume plants, common beans establish a nitrogen fixing association with soil bacteria collectively known as rhizobia. This symbiotic interaction results in the formation of a new organ, the nodule, where the bacteria are allocated, differentiate into bacteroids and fix atmospheric dinitrogen into reduced forms, which are readily available to sustain plant growth. To improve our understanding of the basic processes involved in biological nitrogen fixation by the association common beans-rhizobia can potentially be useful to improve nitrogen nutrition in this crop. Argentina produces about 300,000 ton per year in the Northwest region of the country, which about 95% is exported.

This legume is believed to have been domesticated in the Americas, with two centres of diversification, one of them in Meso-America and the other in the Southern Andes dispersed in Ecuador, Peru, Bolivia and Argentina.

We have demonstrated nodulation preference between common beans and lineages of *R. etli* from the same host region. Our results invited us to speculate that the major genetic pools of *P. vulgaris* and rhizobia coevolved independently after geographic separation (Aguilar, et al., PNAS 2004).

By applying a suppressive subtractive hybridisation approach in which cDNA from a Mesoamerican cultivar inoculated either with the more or the less efficient strain was used as driver and tester, a set of genes were identified as been involved in the early preferential symbiotic interaction (Peltzer-Meschini, et al., Mol. Plant. Microb. Interac. 2008). This analysis identified candidate genes to be further investigation of their roles in the specific interaction with lineages of species *Rhizobium etli*, which build up the objectives of this proposal.

The characterisation of PvRabA2a, a common bean cDNA previously isolated as differentially expressed in root hairs infected with *R. etli*, which encodes a protein highly similar to small GTPases of the RabA2 subfamily. This gene is constitutively expressed in roots, particularly in the elongation zone and in root hairs, where RabA2 protein accumulates in the tip of the expanding region. The role of this gene in nodulation has been studied in *Phaseolus vulgaris* composite plants using a reverse genetic approach. Examination of root morphology in PvRabA2a RNAi plants revealed that the number and length of the root hairs were severely reduced. Upon inoculation with *R. etli*, nodulation was completely impaired and induction of early nodulation genes (ENODs) was undetectable in silenced hairy roots. Moreover, PvRabA2a RNAi plants failed to induce root hair

deformation and to initiate ITs, indicating that morphological changes that precede bacterial infection are compromised in these plants. Early molecular responses, as ERN1, ENOD40 and Hap5 induction are abolished in PvRabA2a RNAi roots. We propose that PvRabA2a, which acts in polar growth of root hair, also is required for reorientation of root hair growth axis during bacterial infection (Blanco, et al. Plant Cell 2009).

In addition, we identified in *Phaseolus vulgaris* a C subunit of the heterotrimeric nuclear factor NF-Y, and demonstrated to play a key role in determining preference for bacteria that occupy nodules. This gene is part of the signal transduction triggered in Mesoamerican cultivars of *Phaseolus vulgaris* in response to *Rhizobium etli* strains predominant in soils of this domestication center. Knock-down of NF-YC produces excessive root hairs deformation in response to bacterial infection and decreases nodule number and occupancy. In addition, over-expression of NF-YC in roots leads to a more efficient colonisation by less effective bacteria in single inoculation and competence experiments. NF-YC silencing affected negatively the induction of cell cycle genes, providing a link between strain perception and cell divisions required for nodule organogenesis. Our results show that the coevolution between legumes and rhizobia is accompanied by a specific recognition followed by a more efficient infection process. Understanding molecular mechanisms that determine the genetic compatibility between legumes and rhizobia is the key to define breeding programs designed to optimise nitrogen fixation (Zanetti, et al., Plant Cell 2010). Attempts to identify the interactor of NF-YC are in progress.

We have further characterised the interaction between beans from the host center of diversification and lineage of *R. etli* and found that that inoculated roots showed earlier response to infection with cognate strains as compared to alopatric strains. These responses include root hair curling, formation of infection threads and number of nodule per plant. In order to gain insight on the differential response, the nodulation factor molecules (NF) produced by strains from the Mesoamerican center of diversification and from the Andean center of diversification were analysed. The molecule signal of one out of three Andean *R. etli* strains (strain 55N1) we examined showed a chitolipooligosaccharidic molecule that is devoid of the acetyl residue on the fucosyl moiety at the reducing end of the molecule. The structure of the other Andean strains as well as the Mesoamerican strains were found to be similar each other. We assessed the competitiveness for nodule occupancy of Mesoamerican beans by an *R. etli* nolL mutant derived of the Mesoamerican strain CE3, which fails to acetylate the fucosyl residue. Several combinations of two *R. etli* strains were used in co-inoculation assays. Results indicated that although mutation in gene nolL affects negatively the competitiveness in co-inoculations with any of the two Mesoamerican *R. etli* wild type strains we assayed (CE3 and SC15), it does not affect the large nodule occupancy in co-inoculation experiments with the alopatric strain *R. etli* 55N1 from the Southern Andean region. These results suggested that factors other than NFs could play a role in determining affinity for nodulation by certain lineages of *R. etli*. Within this frame, we have performed competition experiments in bean varieties from both centers of host diversification by using a combination of strains that included *Rhizobium tropici* CIAT899, a bean-nodulating genotype other than *R. etli*. Nodules of the Mesoamerican cultivar occupied by strains CIAT899 and *R. etli* 55N1 were found to be comparable to each other, whereas in the Andean bean cultivar Alubia, occupancy by the sympatric strain 55N1 clearly overcomes strain CIAT899. This result indicated that competitiveness is expressed in the host x strain combination of the same centre of genetic diversification.

Results Unforeseen in the Original Project:

The results meet the hypothesis we have proposed in our objectives, and indeed provided important insight in the biology of symbiotic interaction. What is more important, are the new avenues for future research, now open, and matter of current investigation in our laboratory, which brings new challenges for young and future scientists.

Publications:

Blanco, F.A., Peltzer Meschini, E., Zanetti, M.E., Aguilar, O.M. A monomeric G-protein of the Rab family is required for root hair formation and pre- infection stages in the common bean-rhizobium symbiotic association. 2009. Plant Cell **21**, 2797-2810

Zanetti, M.E., Blanco, F.A., Beker, M.P., Battaglia, M., Aguilar, O.M. A C-subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean-*Rhizobium etli* symbiosis. 2010. Plant Cell **22**, 4142-4157

Mazziotta, L., Reynoso, M., Aguilar, O.M., Blanco F., Zanetti, M.E. Transcriptional and functional variation of NF-YC1 in genetically diverse accessions of *Phaseolus vulgaris* during the symbiotic association with *Rhizobium etli*. 2012. Plant Biol. (submitted)

BRAZIL

Title: Production of transgenic rabbits for hybridoma-free generation of monoclonal antibodies

Principal Investigator: José Ernesto Belizário, Department of Pharmacology, Institute of Biomedical Sciences, University of São Paulo, Av. Lineu Prestes, 1524 Cidade Universitária, CEP 05508-900, São Paulo, Brazil. Tel: +55-11-30917318, Fax: +55-11-30917438, E-mail: jebeliza@usp.br

ICGEB Contract No.: CRP/08/020

ICGEB Reference No.: CRP/BRA08-01

Abstract: The Simian Virus 40 (SV40) large T antigen is multifunctional protein with DNA helicase, RNA helicase and ATPase activities, which contribute to tumorigenesis in rodents and humans. The Immortomouse mouse strain expresses a mutated large T antigen tsA58 oncogene under the control of the interferon inducible murine H-2K^b MHC (Class I antigen) promoter on chromosome 16. We backcrossed H-2K^b-SV40-tsA58 mice, which is a F1 hybrid of CBA/Ca x C57BL/10 strains with BALB/c mice to obtain a transgenic colony with unique BALB/c background. We have used two pre-validated PCR genotype assays that can distinguish between wild-type, hemizygous, and homozygous animals. Enlargement of thymus is a phenotypic abnormality of immortomouse. We characterised macroscopically and by immunohistochemistry thymic hyperplasia only in the offspring hemizygous females (2:5 ratio). Studies are underway to typing T cell (CD4/8) populations in the thymuses. Moreover, we have routinely observed offspring with undersized males (1:10 ratio) displaying abnormalities after birth. This BALB/c strain of H-2K^b-tsA58 immortomouse can be utilised to investigate specific pathological and physiological patterns associated SV-40 oncogenicity. This transgenic mouse strain will help to isolate immortalising cell lines growing under the permissive 33°C temperature including T cell lymphomas and B cell plasmacytoma capable of producing monoclonal antibodies.

Objectives:

Construction of transgene cassette into the FUGW plasmid, production of transgenic rabbits and genotype and phenotype of founders.

Results Obtained:

Since the creation of rabbit hybridoma technology, efforts to improve efficiency and stability of rabbit monoclonal antibody-producing plasmacytoma lines as rabbit antibodies have the best properties compared to mouse monoclonal antibodies. Aiming to establish the techniques and skills for creation genetically modified rabbits, along the project we have superovulated does and plated rabbit embryos in KSOM medium in Petri dishes under mineral oil and analysed the influence of various conditions and times regarding quality, number and embryo development with and without manipulation. We observed rabbit embryos to differentiate from zygotes, two-cell, four-cell, eight-cell, morula, and early blastocyst stage in CO₂ incubator at 38°C. We purified the H-2K^b-SV40tsA58 transgene using GeneClean III kit (Bio101). Over 200 fertilised zygotes were microinjected and transferred into recipient New Zealand white rabbits. The live pups from 2 recipient does were checked for transgene expression using PCR specific primers and no product was observed suggesting they were negative for the transgene. To address this problem we have been using FUGW vector construction that expresses GFP report to validate its application to rabbit transgenesis by lentiviral co-infection method. So far, we did not succeed to obtain live-born rabbit pups. In parallel, we have done studies for the establishment and characterisation of a primary rabbit fibroblast cell lines Rb1 and Rb2 expressing both H-2K^b-SV-40-tsA58 transgene and eGFP marker gene. We confirmed the expression of SV40tsA58 protein in rabbit fibroblast cell line under the regulation of H-2K^b promoter after induction by human IFN-γ at 33°C permissive temperature. The fibroblast cell lines will be used in future experiments aiming to evaluate the somatic cell nuclear transfer technique for producing genetically engineered rabbits. In 2004, Pasqualini and Arap's paper in PNAS showed that splenocytes derived from the spleen of hyper-immunised immortomouse could yield high titers of monoclonal antibodies against a filamentous phase capsid antigen by growing them at 33°C in the presence of IFN-γ. This study was carried out with mice from with CBA/Ca x C57BL/10 F1 strain which have a mixed background and weak immunological response to antigenic stimulation as compared to BALB/c strain. We have successfully crossbreeding, genotyping and phenotyping mice from offspring of CBA/Ca x C57BL/10 hybrid with BALB/c mice. In functional studies, we will carry out the immunisation of experimental groups of BALB/c immortomice with specific antigens, check out for antibody titers and collect spleens for *in vitro* screening and generation of clonal antibody-producing plasmacytomas under the temperature of 33°C in presence of IFN-γ. This new technology may be more reliable and convenient for large-scale manufacture of antibodies *ex vivo*.

Results Unforeseen in the Original Project:

We had technical problems with the construction of the 6.9 kb containing H-2K^b-SV40tsA58 transgene, which exceeds the estimated size for an insertion fragment into FUGW lentivirus vector. Most viral vectors can carry no more than 8-10 kb of DNA payload taking into account the cis-acting

sequences from those trans-acting factors that are absolutely required for viral particle production, infection, and integration. Transgenic rabbits are much more difficult to produce than mice. Using the pronuclear microinjection method, we obtained only three live-born pups from one out of 6 does. The rabbit embryo has a thick mucin layer that forms around the zona pellucida, which strongly influences implantation in the uterus and pregnancies. We assume that thickness of mucin layer might be contributed to low rates and efficiency of pregnancies.

Publications:

Belizario, J.E. Immunodeficient mouse models: an overview. 2009. *Open Immunol.* **2**, 79-85

Belizario, J.E. Pathogens and dead cells cooperate with cytokines in activating the innate and adaptive response. 2011. *Eur. J. Inflamm.* **9(1)**, 1-11

Belizario, J.E., Sayami, P., Wolf, P., Strauss, B., Xavier-Neto, J. New routes for transgenesis of the mouse. 2012. *J. App. Genet.* **53(3)**, 295-315

CHILE

Title: Role of the Unfolded Protein Response (UPR) in tumour suppression by caveolin-1

Principal Investigator: Claudio Hetz, Institute of Biomedical Sciences, University of Chile, Independencia 1027, Santiago, Chile. Tel: +56-2978-6506 (direct), 2978-6876 (assistant), Fax: +56-2-9786871, E-mail: chetz@hsph.harvard.edu

ICGEB Contract No.: CRP/08/004

ICGEB Reference No.: CRP/CHI08-03

Abstract: The mechanisms underlying the ability of caveolin-1 to suppress tumour formation remains highly controversial. In particular, the possibility that caveolin-1 presence in the endoplasmic reticulum (ER) may modulate stress responses in this organelle relevant to tumour formation remains unexplored. Tumour hypoxia causes a robust increase in the expression of molecular chaperones and other components of an ER stress reaction termed the Unfolded Protein Response (UPR). Several UPR components are directly implicated in the generation of solid tumours. In this project, we developed a joint investigation with Dr. Andrew Quest to define the possible role of the UPR as a downstream target of caveolin-1 in cancer. An *in vivo* cancer model was developed using B16F10 mouse melanoma cells that lack caveolin-1 expression. Tumour progression correlated with the activation of a strong ER stress response, a phenomenon that was significantly attenuated by caveolin-1 expression. Our collaborative results suggest that caveolin-1 expression suppresses survival pathways triggered by ER stress and hypoxia in solid tumours. Our results contributed to understand the functional relationship between caveolin-1 expression and suppression of the UPR, as well as define the possible therapeutic benefits of targeting the UPR pathway in cancer models *in vivo*.

Objectives:

- (i) To define the role of the UPR in tumour suppression by caveolin-1;
- (ii) To assess the effects of caveolin-1 on UPR signalling and cell survival *in vitro* in experimental models of ER stress and hypoxia;
- (iii) To investigate the correlation between caveolin-1 levels and markers of ER stress in human tumour samples derived from Chilean cancer patients.

Results Obtained:

In this project we have studied the role of the UPR as a possible target to uncover the known tumour suppressor activities of caveolin-1. We have performed a series of *in vivo* experiments to provide *in vivo* validation of the major results obtained in cell culture. Our results indicate that caveolin-1 expression significantly reduces the signalling of two major UPR signalling branches initiated by the stress sensors IRE1a and PERK both in tumours and ex-tumour cultures. Moreover, we were able to demonstrate a functional role of caveolin-1 in the suppression of ER stress signalling in cell culture experiments in models of ER stress and hypoxia.

Publications:

Valenzuela, V., Collyer, E., Armentano, D., Parsons, G., Court, F.A., Hetz, C. Activation of the Unfolded Protein Response enhances motor recovery after spinal cord injury. 2012. *Cell Death Dis.* (in press)

Rojas-Rivera, D., Armisén, R., Colombo, A., Eguiguren, A.L., Martínez, G., Díaz, A., Kiviluoto, S., Rodríguez, D., Patron, M., Rizzuto, R., Bultynck, G., Concha, M.L., Sierralta, J., Stutzin, A., Hetz C. TMBIM3/GRINA is a novel Unfolded Protein Response (UPR) target gene that controls apoptosis through the modulation of ER calcium homeostasis. 2011. *Cell Death Differ.* (in press)

Barrientos, S.A., Martinez, N.W., Yoo, S., Jara, J.S., Zamorano, S., Hetz, C., Twiss, J.L., Alvarez, J., Court, F.A. Axonal degeneration is mediated by the mitochondrial permeability transition pore. 2011. *J. Neurosci.* **31**, 966-978

Castillo, K., Rojas-Rivera, D., Lisbona, F., Caballero, B., Nassif, M., Court, F., Schuck, S., Ibar, C., Walter, P., Sierralta, J., Glavic, A., Hetz, C. BAX inhibitor-1 regulates autophagy by controlling the IRE1a/JNK branch of the unfolded protein response. 2011. *EMBO J.* **30(21)**, 4465-4478

Hetz, C., Martino, F., Rodriguez, D., Glimcher, L.H. The Unfolded Protein Response: Integrating stress signals through the stress sensor IRE1a. 2011. *Physiol. Rev.* **91**, 1219-1243

Klein, A., Mosqueira, M., Martinez, G., Robledo, F., González, M., Caballero, B., Cancino, G.I., Alvarez, A.R., Hetz, C., Zanlungo, S. Lack of activation of the Unfolded Protein Response (UPR) in mouse and cellular models of Niemann-Pick type C disease. 2011. *Neurodegener. Dis.* **8(3)**, 124-128

Matus, S., Glimcher, L.H., Hetz, C. Protein folding stress in neurodegenerative diseases: a glimpse into the ER. 2011. *Curr. Op. Cell Biol.* **23**, 239-252

Nassif, M., Hetz, C. Targeting autophagy in ALS: a complex mission. 2011. *Autophagy* **7(4)**, 450-453

Rodriguez, D., Rojas-Rivera, D., Hetz, C. Integrating stress signals at the endoplasmic

- reticulum: The BCL-2 protein family rheostat. 2011. *Biochim. Biophys. Acta* **1813(4)**, 564-574
- Torres, M., Encina, G., Soto, C., Hetz, C.** Abnormal calcium homeostasis and protein folding stress at the ER: a common factor in familial and infectious Prion disorders. 2011. *Commun. Integr. Biol.* **4**, 258-261
- Vidal, R., Caballero, B., Couve, A., Hetz, C.** Converging pathways in the occurrence of endoplasmic reticulum (ER) stress in Huntington's disease. 2011. *Curr. Mol. Med.* **11**, 1-12
- Woehlbier, U., Hetz, C.** Modulating stress responses by the UPRosome: A matter of life and death. 2011. *Trends Biochem. Sci.* **36**, 329-337
- Gupta, S., Deepti, A., Deegan, S., Lisbona, F., Hetz, C., Samali, A.** HSP72 protects cells from ER stress-induced apoptosis via enhancement of IRE1-XBP1 signalling through a physical interaction. 2010. *Plos Biol.* **8(7)**, e1000410
- Nassif, M., Matus, S., Castillo, K., Hetz, C.** Amyotrophic lateral sclerosis pathogenesis: a journey through the secretory pathway. 2010. *Antioxid. Redox. Signal.* **13**, 1955-1989
- Torres, M., Matus, S., Arminsen, R., Stutzin, A., Soto, C., Hetz, C.** Prion protein misfolding affects calcium homeostasis and sensitises cells to endoplasmic reticulum stress. 2010. *Plos One* **5(12)**, e15658
- Hetz, C.** The UPR as a survival factor of cancer cells: More than folding proteins? 2009. *Leuk. Res.* **33(7)**, 880-882
- Hetz, C., Glimcher, L.** Fine tuning of the Unfolded Protein Response: Assembling the IRE1 interactome. 2009. *Mol. Cell* **35**, 551-561
- Hetz, C., Thielen, P., Matus, S., Nassif, M., Court, F., Cuervo, A.M., Martinez, G., Kiffin, R., Brown, R.H., Glimcher, L.** XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. 2009. *Genes Dev.* **23**, 2294-2306 (Comment in: *Genes Dev.* 2009. **23(19)**, 2253-2259)

CHINA

Title: Genomic sequence diversity and oncogenic potential of human papillomavirus type 58 variants across the world

Principal Investigator: Paul Kay-Sheung Chan, Department of Microbiology, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong Special Administrative Region, China. Tel: +86-852-2632-3333, Fax: +86-852-2647-3227, E-mail: paulkschan@cuhk.edu.hk

ICGEB Contract No.: CRP/08/005

ICGEB Reference No.: CRP/CHN08-03

Abstract: This project analysed the sequence variation, phylogeny and geographical distribution of human papillomavirus (HPV) 58 variants from a global perspective. It assessed the epidemiological risk association and characterised the *in vitro* oncogenic properties of E7 proteins derived from naturally occurring HPV58 variants. The data are crucial for designing diagnostic assays and vaccines, and are instrumental for further studies on HPV58, a variant with strong ethnogeographical predilection in attribution to cervical cancer.

Objectives:

Ultimate goal is to achieve a better control on human papillomavirus (HPV) 58-related cervical neoplasia which constitutes a substantial health burden in East Asia and parts of Latin America.

- (i) Comprehensive analysis on genomic sequences of HPV58 isolates collected worldwide;
- (ii) Elucidate the oncogenicity of representative HPV58 strains.

Results Obtained:

We examined 401 HPV58 isolates collected from 15 countries/cities, with the full length of L1, LCR, E2, E4, E5, E6 and E7 gene regions accounting for ~57% (4416-4462 bp) of the whole viral genome sequenced.

(i) Lineage classification: A lineage classification system was established based on phylogenetic analysis of the concatenated E6-E7-E2-E5-L1-LCR sequences. HPV58 variants could be classified into 4 distinct lineages designated as A, B, C and D. In order to facilitate large-scale epidemiological studies, we have identified 3 short fragments that can serve as surrogate targets for constructing phylogenetic trees for lineage identification.

(ii) Geographical distribution of lineages: The distribution of HPV58 variant lineages were found to vary worldwide. Lineage A was significantly less frequent in Africa (49.3%) than other regions (85.7%-95.8%). Lineage C was significantly more frequent in Africa than Asia, America or Europe ($P \leq 0.001$ for each comparison).

(iii) Sequence variation of PCR primer binding sites on L1: We analysed the sequence variation of HPV58 L1 regions correspond to the binding sites of five commonly used consensus primer sets. As a result, 19 mismatches were found within the primer binding sites of 5 sets of commonly used primers. GP5+/6+ had the most, whereas PGMY09/11 had the least number of mismatches.

(iv) Sequence variation of E6 and E7 ORF: Although the degree of nucleotide sequence variation among HPV58 variants was small (maximum 2% for E6, 3% for E7), the variants could be grouped into 4 clades that were identifiable by distinct signatures of sequence variations. These molecular signatures are not only convenient markers for large-scale epidemiological studies, the substitutions themselves may carry biological implications.

(v) Geographical distribution of E6 and E7 variants: The distribution of E6/E7 clades varied geographically. Africa had significantly more E6 clade 3, but significantly less E6 clade 1 than other regions. E7 clades 2 and 3 co-circulated in Asia, whereas clades 2 and 4 co-circulated in Africa. E7 clade 3 was more frequent in Asia, and clade 4 was more frequent in Africa. E6 WW01 was the predominant variant detected in Asia (78%), Europe (91%) and Americas (79%), but not in Africa (47%). E6 WW02 was significantly more frequent in Africa compared to other regions. E7 WW01 was the most common variant found across the four regions, and with a significantly higher prevalence in Europe (77%) compared to Asia (42%) and Africa (47%). The other variants E7 WW02 / 03 / 04 showed a more distinct geographical predilection.

(vi) Epidemiological risk association: E7 clade 2, E7 WW01, E7 G41R and G63D were associated with a lower risk; whereas E7 clade 3, E7 WW02, E7 T20I and G63S were associated with a higher risk for cervical neoplasia. Similar associations were reproduced for Hong Kong and Korean which provided sufficient sample for subgroup analysis.

(vii) Transformation potential of E7 proteins: Results of epidemiological analyses indicated that the most important amino acid substitutions were T20I and G63S located in the E7 ORF. E7 proteins expressed from plasmids containing the prototype and three naturally occurring variants, 58E7-V1 (with nucleotide changes of C632T [T20I], T744G [silent] and G760A [G63S]), 58E7-V2 (G694A [G41R], T744G [silent] and G761A [G63D]) and 58E7-V3 (T744G [silent], A793G [T74A], C798T [D76E], C801A [D76E] and C840T [silent]) were compared. HPV58 E7-V1 showed a marked increase in ability to transform BRK cells compare to prototype; whereas the other two variants were similar to prototype.

(viii) Binding with proteasome subunits: Proteasome subunit S2 shows a negative interaction with all the E7 proteins. S4 interacted efficiently with the prototype E7 but less efficiently with the three variants. S5a showed the strongest interaction with E7, especially with 58E7-V1. S7 and S8 also interacted efficiently with HPV58 E7 proteins.

(ix) pRb degradation: pRB was more efficiently degraded by HPV58 E7-V1 protein *in vivo* compared to prototype.

(x) Suppression of cellular senescence: The HPV58 E7-V1 protein was found to be most efficient in suppressing the cellular senescence, which was in concordance with the results obtained from the pRB degradation assay.

Publications:

Chan, P.K., Luk, A.C., Park, J.S., Smith-McCune, K.K., Palefsky, J.M., Konno, Y., Giovannelli, L., Coutlée, F., Hibbitts, S., Chu, T.Y., Settheetham-Ishida, W., Picconi, M.A., Ferrera, A., De Marco, F., Woo, Y.L., Raiol, T., Piña-Sánchez, P., Cheung, J.L., Bae, J.H., Chirenje, M.Z., Magure, T., Moscicki, A.B., Fiander, A.N., Di Stefano, R., Cheung, T.H., Yu, M.M., Tsui, S.K., Pim, D., Banks, L. Identification of human papillomavirus type 58 lineages and the distribution worldwide. 2011. *J. Infect. Dis.* **203**, 1565-1573

Chan, P.K., Cheung, J., Luk, A., Ho, W. Sequence variations of human papillomavirus type 58 E2, E4 and E5 genes across the world. In: 26th International Papillomavirus Conference & Clinical and Public Health Workshop, Montreal, Canada, 3-8 July 2010. Abstract No. P-119

Cheung, J., Luk, A.C., Ho, W.C., Chan, P.K. Sequence variation and lineage classification system for HPV58: A worldwide perspective. In: 26th International Papillomavirus Conference & Clinical and Public Health Workshop, Montreal, Canada, 3-8 July 2010. Session: Basic Science 6, Abstract No. 345

Ho, W.C., Cheung, J.L., Luk, A.C., Chan, P.K. Sequence variation of Human Papillomavirus Type 58 E7 across the world. In: 26th International Papillomavirus Conference & Clinical and Public Health Workshop, Montreal, Canada, 3-8 July 2010. Abstract No. P-125

Chan P.K.S. Worldwide distribution of human papillomavirus 58 variant lineages. In: The Christophe Merieus Conference – Trends in Virology, Annecy, France, 17-19 January 2010

Cheung, J., Luk, A., Chan, P.K. Mismatch of commonly used primers for HPV58 sequences collected worldwide. In: 27th International Papillomavirus Conference and Clinical Workshop, Berlin, 17-22 September 2011, Abstract No. P-32.23

COLOMBIA

Title: Identification of pathogenicity factors in *Batrachochytrium dendrobatidis* by *in silico* comparative genomics

Principal Investigator: Silvia Restrepo Restrepo, Mycology and Plant Pathology Laboratory, Biological Sciences Department, Faculty of Sciences, Universidad de Los Andes, Cra 1 No 18A-10, Bogota, Colombia. Tel: +57-1-3394949, ext. 2777, Fax: +57-1-3394949, ext. 2817, E-mail: srestrep@uniandes.edu.co

ICGEB Contract No.: CRP/09/005

ICGEB Reference No.: CRP/COL09-01

Abstract: Global amphibian diversity is at great risk, and Chytridiomycosis is one of the main causes linked to this massive extinction. To develop future control strategies for this threat is paramount to understand what molecular determinants mediate the infection process and allow *Batrachochytrium dendrobatidis* (*Bd*) to cause disease. We developed a comparative genomics strategy using both available sequenced *Bd* strains and ESTs libraries of 44 organisms (9 parasites, 9 non-pathogenic Fungi, 26 pathogenic fungi) to independently determine potential pathogenicity determinants. Among the candidates recovered, five are particularly promising. To test these pathogenicity hypotheses, a controlled infection system with tolerant host skin (*Rheobates plamatus*) was developed in which we assessed expression patterns of three of the candidates at early stages of infection (0, 12, 24, and 36hpi). For these three genes, we found experimental evidence of strong variations in transcript in very early stages in the establishment of infection (0-12hpi), suggesting a possible role in the initial interactions between *Bd* and its host. On the other hand, due to the importance of the studying of *Bd* in Colombia as a way to understand the epidemiology and potentially the biology of the disease, here we also did the genetic and morphological characterisation of the first Colombian isolate of *Bd* (strain EV001).

Objectives:

- (i) Prediction of putative pathogenicity factors by comparative genomics;
- (ii) Isolation and characterisation of a Colombian strain of *Batrachochytrium dendrobatidis*;
- (iii) Development of a co-culture system of *Bd* and host skin;
- (iv) Determination of the expression patterns of the putative pathogenicity genes.

Results Obtained:

We developed a comparative genomics strategy using both available sequenced *Bd* strains and ESTs libraries of 44 organisms (9 parasites, 9 non-pathogenic Fungi, 26 pathogenic fungi) to independently determine potential pathogenicity determinants based on two criteria: 1) Proteins shared between *Bd*, parasites and pathogenic fungi, 2) Proteins shared only with pathogenic fungi. Around 20% of the predicted proteins were shared with pathogenic fungi. Of these, only about 3% were absent in non-pathogenic, but more than 60% of these were also shared with Parasites. Among these and of special interest we obtained a complex with putative Ascorbic acid-dependant ferric reductase activity; a receptor possibly involved in lipophilic ligands detoxification (LMBRL); a Serine peptidase (POP1) possibly related to skin defense peptide degradation; a Metazoa-like possibly secreted PHACTR protein and a PhsF homolog maybe obtained by horizontal transfer. These candidates provide lights towards understanding *Bd* pathogenesis and future control strategies.

As part of the required analysis to complement and test the pathogenicity hypothesis obtained through the bioinformatic approach, we developed the first controlled system for the study of this infection. The system consists of a co-culture system in which frog skin microorgans are inoculated with *Bd* zoospores. Based on extant expression data, we suggest possible reference genes for further expression analyses and tested three of them experimentally. The transcript profile of the early stages of infection (0, 12, 24, and 36hpi) of tolerant host (*Rheobates plamatus*) was followed for three of pathogenicity factors (POP1, LMBRL and PhzF) and for a melanin synthesis protein. For these three genes, we found experimental evidence of strong variations in transcript in very early stages in the establishment of infection (0-12hpi), suggesting a possible role in the initial interactions between *Bd* and its host.

In Colombia, which hosts the second most diverse Amphibian fauna in the world, the introduction of *Bd* seems to be very recent and is hypothesised to be from two of the three suggested waves of spread of *Bd* for South America. Due to the importance of the studying of *Bd* in Colombia as a way to understand the epidemiology and potentially the biology of the disease, here we present the genetic and morphological characterisation of the first Colombian isolate of *Bd* (strain EV001). Our aim was to locate the Colombian gene pool in the global context and to gather initial evidence to test the hypothesis of introduction of the disease to Colombia. For this we used Light and Scanning electron microscopy and an MLST approach using 17 markers previously developed, providing 20 loci for further analysis, given that intralocus recombination was identified for three loci. Although the morphological characteristics are congruent with the general features of *Bd*, we found broad distance between our strain and the only extant South American isolate. Our results suggest a

frequency of recombination that might be higher than the observed for strains described previously and provide evidence, although circumstantial, for the multiple introduction or sex for Northern South America.

The comparative genomic exploratory analyses provided other proteins that although were not recovered directly by our approach, are of great interest in understanding pathogenesis and the evolutionary history of early diverging Fungi. Among these there are a Polyphenol oxidase (PPO1) involved in melanin synthesis and a possible light-sensing rhodopsin. We were able to determine the expression pattern of PPO1 using the co-culture system and found very low variation in mRNA levels along the studies time frame. Research on this proteins and their corresponding family is ongoing, includes collaboration with renowned members of the field (see supplementary research).

Results Unforeseen in the Original Project:

The comparative genomic exploratory analyses provided other proteins that although were not recovered directly by our approach, are of great interest in understanding pathogenesis and the evolutionary history of early diverging Fungi. Among these there are proteins involved in melanin synthesis and members of the rhodopsin family We were able to determine the expression pattern of the melanin synthesis protein using the co-culture system and found very low variation in mRNA levels along the studied time frame. Research on these proteins and their corresponding families is ongoing. Also, due to the training internship and conferences one of the researchers attended, we were able to collaborate in research in which we applied approaches developed during this project to another system (see Parallel molecular evolution in an herbivore community).

Publications:

Zhen, Y., Aardema, M.L., Medina, E.M., Schumer, M., Andolfatto, P. Parallel molecular evolution in an herbivore community. 2012. Science (to be resubmitted)

Medina, E.M., Flechas, S.V., Sarmiento, C., Cárdenas, M., Crawford, A.J., Amézquita, A., Restrepo, S. Highly recombinant Colombian strain of *Batrachochytrium dendrobatidis*. 2011. Ecohealth (to be resubmitted)

Medina, E.M., Ahrendt, S., Restrepo, S., Stajich, J.E. Because they were blind but now they see!: animal-like light-sensing rhodopsins in early diverging Fungi. 2012. (in preparation)

Medina, E.M., Restrepo, S., Stajich, J.E. Melanin-mediated stress resistance in the amphibian-killing pathogen *Batrachochytrium dendrobatidis*. 2012. (in preparation)

CROATIA

Title: Mechanisms of aminoglycoside resistance by ribosomal RNA methylation

Principal Investigator: Gordana Maravić Vlahoviček, Department of Biochemistry and Molecular Biology, Faculty of Pharmacy and Biochemistry, University of Zagreb, Ante Kovačića 1, Zagreb, Croatia. Tel: +385-1-639-4448, Fax: +385-1-639-4400, E-mail: gordana@pharma.hr

ICGEB Contract No.: CRP/08/006

ICGEB Reference No.: CRP/CRO08-02

Abstract: Increasing and improper use of antibiotics in humans, animals, and agriculture has resulted in emergence and spreading of multiply resistant bacterial pathogens. As a consequence, many infectious diseases are increasingly difficult to treat, making the antibiotic resistance problem one of the top priorities of public health. This project investigates the scarcely explored mechanism of resistance to aminoglycoside antibiotics, found in both aminoglycoside producers and in growing number of clinical strains. This type of resistance is generated through the action of methyltransferases from the Arm and Kam family, which methylate the aminoglycoside interaction site in 16S rRNA (Arm - m⁷G1405; Kam - m¹A1408) using the S-adenosyl-methionine as a methyl donor. Research objectives are directed towards the functional characterisation of the Arm and Kam enzymes and their RNA substrate. Obtained results provide a notable insight into molecular basis of aminoglycoside resistance, which will directly lead the future research initiatives towards the finding of specific inhibitor of the Arm and Kam enzymes and thus contribute to the worldwide efforts to combat antibiotic resistance in bacterial pathogens.

Objectives:

- (i) to define the core motif of Arm and Kam RNA substrate;
- (ii) to define RNA-binding modules for both enzyme families;
- (iii) to propose the reaction mechanism of 16S rRNA:m⁷G and rRNA: m¹A methylation.

All individual objectives of the project have been met and accomplished. Obtained results have been published in four papers and two additional papers are currently in preparation.

Results Obtained:

In order to analyse the binding pattern of Arm and Kam enzymes to their cognate rRNA substrate, we did the mutagenesis of the ribosomal A-site within the helix 44 of 16S rRNA using *E. coli* strains with a single rRNA operon. We analysed the impact of the mutations on the function of all members of Arm and Kam family by determining the minimal inhibitory concentration of kanamycin and analysing the target nucleotide methylation with primer extension. These results enabled us to define the core motif of Arm and Kam RNA substrate and to discover the differences among the individual representatives of both families, as well as to improve the model of enzyme interaction with the ribosomal subunit 30S.

We carried out extensive structural and functional analysis of the Sgm enzyme, as well as sequence analysis of entire Arm family and related RNA methyltransferases (MTases). We showed that the Sgm methyltransferase confers resistance to 4,6-disubstituted deoxystreptamine aminoglycosides by introducing the 16S rRNA modification m⁷G1405 within the ribosomal A site. This region of *Escherichia coli* 16S rRNA already contains several methylated nucleotides including m⁴Cm1402 and m⁵C1407. Modification at m⁵C1407 by the methyltransferase RsmF is impeded as Sgm gains access to its adjacent G1405 target on the 30S ribosomal subunit. An Sgm mutant (G135A), which is impaired in S-adenosylmethionine binding and confers lower resistance, is less able to interfere with RsmF methylation on the 30S subunit. The two methylations at 16S rRNA nucleotide m⁴Cm1402 are unaffected by both the wild-type and the mutant versions of Sgm. The data indicate that interplay between resistance methyltransferases and the cell's own indigenous methyltransferases can play an important role in determining resistance levels confirmed the guanine 1405 to be the methylation site of Sgm methyltransferase. We then solved the structure of Sgm in complex with cofactors S-adenosyl-methionine and S-adenosyl-homocysteine at 2.0 and 2.1 Å resolution, respectively. We also analysed a number of mutants by isothermal titration calorimetry, performed the rRNA footprinting by chemical probing of Sgm-30S subunit complex and protein-substrate docking. We proposed the mechanism of methylation of G1405 by Sgm and compared it with other m⁷G methyltransferases, revealing a surprising diversity of active sites and binding modes for the same basic reaction of RNA modification. The results of our analyses provide a stepping stone for the design of inhibitors against the medically important Arm family of MTases. In particular, the guanosine-binding site identified in the Sgm structure (and by extension, in homologous structures of ArmA and RmtB MTases) can be used as receptor for structure-based virtual screening.

Furthermore, we conducted the extensive structural and functional analysis of the NpmA enzyme, as well as sequence analysis of entire Kam family and related RNA MTases. We solved the structure of NpmA enzyme and of its complex with cofactors S-adenosyl-methionine and S-adenosyl-homocysteine at 2.4, 2.7 and 1.68 Å resolution, respectively. We also analysed a number of mutants by isothermal titration calorimetry, methylation test *in vitro*, studied their ability to generate

kanamycin resistance *in vivo* and performed the rRNA footprinting by chemical probing of Sgm-30S subunit complex and protein-substrate docking. We proposed the mechanism of methylation of A1408 by NpmA and compared it with other m¹A methyltransferases. Helices 24, 42 and 44 were found to be the main NpmA-binding site. Both experimental and theoretical analyses suggest that NpmA flips out the target nucleotide A1408 to carry out the methylation. Obtained results will assist in the development of specific NpmA inhibitors that could restore the potential of aminoglycoside antibiotics.

Results Unforeseen in the Original Project:

In our studies of the Sgm methyltransferase, where we confirmed the guanine 1405 to be the methylation site of Sgm methyltransferase, we discovered the interconnection of Sgm enzyme with the housekeeping methyltransferase RsmF.

We extended our research to specific application of the MTases from the Arm and Kam family in the studies of bacterial quorum sensing. We investigated whether tobramycin sensing/response at sub-inhibitory concentrations was affecting the two independent AHL QS systems in and environmental *Pseudomonas aeruginosa* strain. It was established that sub-inhibitory concentrations of tobramycin inhibited the RhlI/R system by reducing levels of C4-HSL production. This effect was not due to a decrease of rhlI transcription and required tobramycin-ribosome interaction, as confirmed by the action of RmtC, RmtA and NpmA MTases.

Networking:

The project enabled and promoted intensive collaboration with the following groups: (i) Dr. Janusz Bujnicki (International Institute of Molecular and Cell Biology, Warsaw, Poland), (ii) Dr. Jayaraman Sivaraman (National University of Singapore), (iii) Dr. Stephen Douthwaite (University of Southern Denmark, Odense, Denmark) and (vi) Dr. Vittorio Venturi (ICGEB, Trieste, Italy).

Publications:

Čubrilo, S., Babić, F., Douthwaite, S., Maravić Vlahoviček, G. The aminoglycoside resistance methyltransferase Sgm impedes RsmF methylation at an adjacent rRNA nucleotide in the ribosomal A site. 2009. *RNA*. **15**, 1492-1497

Babić, F., Venturi, V., Maravić Vlahoviček, G. Tobramycin at subinhibitory concentration inhibits the RhlI/R quorum sensing system in a *Pseudomonas aeruginosa* environmental isolate. 2010. *BMC Infect. Dis.* **10**,148

Husain, N., Tkaczuk, K., Tulsidas, S., Kaminska, K., Čubrilo, S., Maravić Vlahoviček, G., Bujnicki, J., Sivaraman, J. Structural basis for the methylation of G1405 in 16S rRNA by aminoglycoside resistance methyltransferase Sgm from an antibiotic producer: A diversity of active sites in m⁷G methyltransferases. 2010. *Nucleic Acids Res.* **38(12)**, 4120-4132

Husain, N., Obranić, S., Koscinski, L., Seetharaman, J., Babić, F., Bujnicki, J., Maravić Vlahoviček, G., Sivaraman, J. Structural basis for the methylation of A1408 in 16S rRNA by a panaminoglycoside resistance methyltransferase NpmA from a clinical isolate. 2011. *Nucleic Acids Res.* **39(5)**, 1903-1918

EGYPT

Title: *plg-1* gene expressing propionin by lactic starters in dairying

Principal Investigator: Mahmoud Khairy Tahoun, Department of Dairy Science and Technology, Faculty of Agriculture, Alexandria University, Aflatoon Street, El-Shatby 21545, Alexandria, Egypt. Tel: +20-3-5925405, Fax: +20-3-5922780, E-mail: samehgene@yahoo.com

ICGEB Contract No.: CRP/07/012

ICGEB Reference No.: CRP/EGY07-03

Abstract: The expression of propionin PLG-1 by lactic acid starter culture is the target of the present research program, hence propionin PLG-1 expressed by *plg-1* gene of *Propionibacterium thoenii* P127 was reported to inhibit several food-borne pathogens such as *Campylobacter jejuni*, *Escherichia coli*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Vibrio parahaemolyticus*, *Trichoderma reesi*, *Fusarium tricinctum*, *Candida lipolytica*, *Aspergillus* spp., *Yersinia enterocolitica* and a strain of *Corynebacterium* sp. Such program depends on *plg-1* gene transformation to *Lactococcus lactis* subsp. *lactis* MG1613 and *Lactobacillus plantarum* TF103 as models of dairy starters. This approach can be accomplished using pFG200, pEM76, pINT125 and pLEB590 vectors derived from lactococcal DNA. The multicopy vector pFG200 contains the replicon fragment of pCT1138, the amber suppressor *supD* from *Lactococcus* strain NJ1 as a food-grade selectable marker, and a versatile polylinker. The vector pEM76 is based on an integrase (*Int*), a resolve-invertase (β - recombinase) and their respective target sites (*attP* – *attB* and six, respectively). It contains heterologous origin and antibiotic resistance markers surrounded by two directly oriented six sites, a multiple cloning site, the *int* gene and the *attP* site of phage A2. pINT125 is a food-grade vector system that allows stable integration of multiple plasmid copies into the chromosome of *Lactococcus lactis*. The vector consists of the plus origin of replication (Ori⁺) of the lactococcal plasmid pWV01, the sucrose genes of the lactic acid bacterium *Pediococcus pentosaceus* PPE1.0 as a selectable marker, a multiple-cloning site, and a lactococcal DNA fragment of a well-characterised chromosomal region. Later single-crossover integration system could be a simple procedure for the engineering of stable food-grade strains carrying multiple copies of a gene of interest. pLEB590 is a food-grade cloning vector for lactic acid bacteria that constructed using the nisin immunity gene *nisI* as a selection marker. It is constructed entirely of lactococcal DNA containing the pSH71 replicon, the *nisI* gene, and the constitutive promoter P45 for *nisI* expression. Such food-grade expression system has potential for expression of foreign genes in lactic acid bacteria in order to construct improved starter bacteria for food applications. *Propionibacterium thoenii* P127 (propionin PLG-1 producer strain) was screened for its antimicrobial activity against a wide range of microorganisms including, 14 *Escherichia coli* strains, 7 *Propionibacterium* strains, 3 yeast strains, 3 *Lactococcus* strains, 2 *Lactobacillus* strains, and one strain of each of *Bacillus* sp., *Clostridium* sp., *Streptococcus* sp., *Oenococcus* sp., *Listeria* sp., *Salmonella* sp. and *Aspergillus* sp. The results indicated inhibition of above microbial strains with different degrees. Above findings were found to contradict the results obtained by other investigators. Therefore, total crude protein of the cell-free supernatant was partially purified and retested again against above strains. The Results obtained indicated that total crude protein strongly inhibited the growth of all *Escherichia coli*, *Listeria* sp., *Salmonella* sp. and yeasts strains investigated. However, it moderate affected the growth of all propionibacteria strains expect *Propionibacterium freudenreichii* and weakly inhibited the growth of all lactic acid bacterial strains. However it showed negative effect on the growth of *Aspergillus niger*, *Bacillus subtilis* DB100 host, and *Clostridium sporogenes* DSM1446. Therefore, the total crude protein extract was exposed to SDS-PAGE and the band containing propionin PLG-1 (~10 kDa) was extracted and retested against all culture strains investigated at the beginning and revealed similar data to the total crude protein extract. This was followed by biochemical tests to confirm the characterisation of the propionin PLG-1. On the genetic level, *plg-1* gene expressing propionin PLG-1 was isolated using specific bioinformatically designed primers and its sample was sent overseas to be sequenced for the first time. The resulted PCR product was visualised on DNA agarose gel (1.5%) and migrated beside 1 KB DNA marker. After gel electrophoresis, the PCR product size was calculated to be ~ 500 bp. Such product was isolated from the gel and purified and transferred to different lactic acid bacterial strains (*Lactococcus lactis* subsp. *lactis* MG1613 and *Lactobacillus plantarum* TF103 as models of dairy starters) using pLEB590 (as the best examined lactic acid bacterial vector selected from a collection of different vectors including pFG200, pEM76, pINT125 and pLEB590 vectors originated from lactococcal DNA) to give the modified vector pLEBPLG-1. The lactic acid bacterial transformants showed strong antimicrobial activity against *Escherichia coli* DH5a, *Listeria monocytogenes* 18116, and *Salmonella enterica* 25566. Such lactic acid bacterial transformants can be used in dairy industry to control above food-borne pathogens that are largely distributed in worm climates and to feed school children in the poor countries where epidemic diseases and diarrhea prevails.

Objectives:

- (i) Screening for antimicrobial activity of *Propionibacterium thoenii* P127 against some microorganisms including bacteria, yeast and fungi;
- (ii) Propionicin PLG-1 isolation, partially purification and characterisation;
- (iii) Isolation of *Propionibacterium thoenii* P127 genomic DNA using different method;
- (iv) Isolation of *plg-1* gene of *P. thoenii* P127. Isolation of several cloning vectors for *plg-1* transfer into some strains of lactic acid bacteria;
- (v) Digestion of PCR product using appropriate restriction endonucleases. Digestion of pLEB590 cloning vector with the same restriction endonucleases;
- (vi) Ligation of digested *plg-1* gene and digested pLEB590 cloning vector to construct the new vector pLEBPLG-1;
- (vii) Electroporation of pLEBPLG-1 to *Lactococcus lactis* subsp. *lactis* MG1613as well as *Lactobacillus plantarum* TF103;
- (viii) Examine the new transformants for their antimicrobial activity against some food-borne pathogens. Examine the use of the transformants for their capacity to affect the growth of above food-borne pathogens in milk environment.

Results Obtained:

- (i) Propionicin PLG-1 of *Propionibacterium thoenii* P127 was isolated and partially purified and characterised;
- (ii) Such bacteriocin strongly inhibited different microorganisms including the pathogens *Escherichia coli* DH5 α , *Listeria monocytogenes* 18116, *Salmonella enterica* 25566;
- (iii) Isolation of genomic DNA of *P. thoenii* P127 using different isolation methods;
- (iv) Designing different primers for the isolation of *plg-1* gene using different bioinformatics programs;
- (v) PCR product and the cloning vector pLEB590 were digested using appropriate restriction endonucleases. Ligation of digested PCR product containing *plg-1* gene and digested pLEB590 cloning vector in order to construct the new vector pLEBPLG-1 (3600 bp);
- (vi) Electroporation of pLEBPLG-1 to *Lactococcus lactis* subsp. *lactis* MG1613 as well as *Lactobacillus plantarum* TF103 as models of lactic acid bacteria was successfully performed;
- (vii) The new lactic acid bacterial transformants showed strong antimicrobial activity against above pathogens;
- (viii) The use of above new transformants in milk environment and their capacity to affect the growth of above food-borne pathogens specially in milk was successfully performed.

Publications:

Tahoun, M.K., Mohamed, S.E. Propionicin PLG-1, screening and partial characterisation of the protein. 2010. Int. J. Biotech. Biochem. 6, 325–338

Mohamed, S.E., Tahoun, M.K. Heterologous expression of *pctA* gene expressing propionicin T1 by some lactic acid bacterial strains using pLEB590. 2011. J. Agric. Biol. Sci. (in press)

Tahoun, M.K., Mohamed, S.E. Heterologous expression of *pctA* gene expressing propionicin T1 by some lactic acid bacterial strains using pINT125. 2011. Adv. Biosci. Biotechnol. (in press)

HUNGARY

Title: Assessing the protein folding-aggregation paradigm with NMR and molecular modeling

Principal Investigator: András Perczel, Eötvös Loránd University, Institute of Chemistry, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary. Tel: +36-1-3722500, ext. 1653, Fax: +36-1-3722620, E-mail: perczel@chem.elte.hu, Website: www.chem.elte.hu/departments/protnmr

ICGEB Contract No.: CRP/08/007

ICGEB Reference No.: CRP/HUN08-03

Abstract: We have investigated issues related to protein folding and aggregation. Our model systems comprised (mini)proteins and peptides of general and special biological interest, investigated at atomic level by CD and NMR spectroscopy, compiled with theoretical works on molecular modeling, quantum chemical calculations and bioinformatics. We have completed the rational design of selected Trp-cage miniproteins and have further optimised their stability. The variants were selected to exhibit either more stable structures, inducible aggregation or aggregation-prone behavior. Their structure, dynamics and folding was investigated by homo- and heteronuclear NMR and theoretical methods. Furthermore, additional model systems were studied such as intrinsically disordered proteins or IDPs (PP25, calpastatin fragments, highly charged sequences, etc.) as well as protein domains of variable sizes (CCP and DLC). In addition, supramolecular assemblies of both alpha- and beta-peptides were studied. Our present results are hoped to contribute to the general understanding of protein folding and stability, as well as to the specific description of several biologically important molecular systems spelled out above.

Objectives:

Our aims were to investigate several of the principles involved in determining the structure and dynamics of folded and intrinsically disordered natural proteins and foldamers, as well as to gain insight into the structure-dynamic-function relationships of selected biomolecules at the atomic level.

Results Obtained:

(A) Secondary structure elements and foldamers

Quantum chemical calculations contributed to the theoretical aspects of structure analyses. On one hand, the collagen triple-helix model was studied and its stabilisation and elasticity was reported to be due to the presence of bridging water molecules (Pálfi and Perczel, 2010). On the other hand, the structure and stability of sheet-like conformers of beta-peptides were studied and novel foldamers were reported, which secondary structural elements are structural analogs of beta-pleated sheets of proteins and are enzymatically resistant biomaterials (Beke, et al., 2009).

NMR is a powerful tool for analysing proteins especially when well folded. However, traditional techniques based on NOEs, *J*-coupling constants, H-bonds etc. fail to perform well for highly dynamic proteins, especially for IDPs. In these cases secondary chemical shifts are of great help to decipher information on the backbone fold (if any). These $\Delta\delta$ -type experimental techniques however require more computational efforts. Thus, additional *ab initio* NMR chemical shift calculations were completed on major types of backbone folds and subsequently compared to the corresponding experimental counterparts retrieved from selected protein databases and adjusted carefully. We have found acceptable/good agreement for the most common backbone conformers (e.g. helices, β -sheets, PPII). Experimental and computed $\Delta\delta^{\text{H}^a}$ - $\Delta\delta^{\text{C}^a}$, $\Delta\delta^{\text{H}^a}$ - $\Delta\delta^{\text{C}^b}$, and $\Delta\delta^{\text{C}^a}$ - $\Delta\delta^{\text{C}^b}$ correlation maps provided qualitatively realistic pictures, yielding help to estimate the dominant backbone fold of even dynamic proteins from CSIs (Czajlik, et al., 2011).

Using first principles calculations, we have characterised the hydration layers of collagen in detail. Our results suggest that the water binding places on the surface of the triple helix can provide explanation on how an almost liquid-like hydration environment exists between the closely packed tropocollagens. It seems that these water reservoirs or buffers can provide space for "hole conduction" of water molecules and thus contribute to the elasticity of collagen (Pálfi and Perczel, 2010).

We have conducted a matrix isolation conformational study on model beta peptides, *N*-acetyl-3-aminopropionic acid-*N'*-methylamide (Ac- β -HGly-NHMe) and *N*-acetyl-3-aminobutanoic acid-*N'*-methylamide (Ac- β -HAla-NHMe). The IR spectra recorded suggest that although β -peptides having an "extra" backbone torsion should be more flexible than α -peptides, fewer backbone conformers are favorable based on their relative energies. Thus, from a larger conformational arsenal, only a lower number of backbone conformers can emerge, which possibly had a fundamental effect on their applicability during prebiotic evolution (Beke, et al., 2009).

Foldamers and the replacement of α -amino acid residues by β - or γ -residue types is a fascinating and modern approach to stabilise peptide backbones. In order to predict the most probable secondary structures of β -peptides, a systematic QM modeling project was launched early on in our laboratory. More recently, calculations were carried out on artificial β -peptides in order to deeply understand their conformational behavior. A systematic study on the effect of chirality on "apolar"

sheet stability was completed on disubstituted [HCO- β -Ala- $\beta^{2,3}$ -hAla- β -Ala-NH₂]₂ model peptides calculated at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) and B3LYP/6-311++G(d,p)//B3LYP/6-31G(d) levels of theory, both in vacuum and in polar and apolar solvents. In addition, both types of "apolar" sheets were investigated; the one with two strands of identical (AA) and enantiomeric (AB) backbone structure. Our results show that not only heterochirally disubstituted sheets have great preference for sheet formation ($\Delta G \sim -11$ kcal mol⁻¹), but unexpectedly "homochiral disubstitution" itself could result in a stable fold ($\Delta G \sim -5$ kcal mol⁻¹). Thus, it does not necessarily disrupt the sheet structure. The predictive nature of the effect of chirality on sheet stability offers an easy-to-use summary on how to design β -peptide sheet structures. It was found that heterochirally disubstituted models ("lego elements") are the best candidates for sheet formation, if the two strands are substituted in a way to create identical torsional angle sets on the two backbones for an ideal hydrogen-bonding pattern. With adequately selected side chains, homochirally disubstituted derivatives may also form sheet structures, and the position of methyl groups would prevent the assembly of more than two strands making it ideal to create hairpin folds (Pohl, et al., 2012, Pohl, et al., in press).

(B) Miniproteins

We have proceeded with the design, production and characterisation of Trp-cage miniprotein variants. One of the concepts applied is to generate longer variants of the model protein Tc5b according to its 'parent molecule', exendin 4. We have prepared a series of Tc5b variants elongated at its N-terminus by one residue at each step. Detailed characterisation (NMR measurements, structure calculations etc.) of these molecules is in progress. In addition, other Tc5b variants have been characterised using state-of-the-art dynamic structure calculation methods (MUMO protocol as implemented in GROMACS). Evaluation of the results is in progress. Additionally, we started to redesign a Trp-cage miniprotein by one or two mutations to obtain a genuine IDP. Based on rational considerations and extensive bioinformatical investigations, the biotechnological production of selected sequences has been started.

One of the variants, Tc5b-D9Aad has been extensively analysed with CD spectroscopy using CCA+, a program developed in our group. Convolution analysis revealed that the CD spectrum this variant can be adequately described by two components, one corresponding to the ordered and the other to the disordered state (Jákli and Perczel, 2009).

We have significantly proceeded with the design and analysis of novel Tc5b variants. We have detailed dynamical data for selected variants at different temperatures, for which the calculation and analysis of dynamically restrained structural ensembles is in progress. In addition, we have several double mutants at hand that behave like intrinsically unfolded/dynamic proteins. Detailed characterisation of these variants and the design and preparation novel ones is under way. Our present data promise understanding of the role of interaction networks within Tc5b from a dynamic perspective.

We have continued our extended research on miniproteins by investigating the folding and stability role of a salt-bridge (Asp⁹-Arg¹⁶) within the molecule. The point mutations identified that both of the above side-chains are also part of a dense interaction network composed of various types of nature. The fold stabilising effects were judged by electronic circular dichroism, ECD, and NMR data of these mutants at different pHs. Our present results indicate that the folding of Trp-cage miniproteins is a complex process via an intermediate state related to the salt-bridge stabilisation. In this molecular framework, Arg¹⁶ stabilises more than Asp⁹ does because of its higher degree of 3D-fold cooperation. The Xxx⁹ \leftrightarrow Yyy¹⁶ salt-bridge is not an isolated entity of this fold; rather it is an integrated part of a complex interaction network (Rovó et al. 2011).

(C) Modular folded protein(s)

(C.i) Small serine protease inhibitors: Using NOE and backbone S² data characteristic of ps-ns time-scale motions, we have generated dynamically restrained protein structural ensembles of small serine protease inhibitors and investigated the relation of the conformer sets to the available crystal structures of the enzyme-bound forms of the inhibitors. We found that the dynamically restrained conformer sets reproduce most NMR experimental data better than structures obtained using "conventional" structure calculation methods. Moreover, they also include conformers highly similar to those observed in the enzyme complexes, indicating that enzyme binding of small, 35-residue inhibitors is actually a dynamic conformer selection process on the nanosecond-timescale. Thus, fast timescale dynamics enables the association rate to be solely diffusion-controlled just like in the rigid-body model traditionally used to describe the mechanism of canonical serine protease inhibitors (Gáspári, et al., 2010a).

(C.ii) Complement control modules: We have focused on the dynamical determination of protein molecules especially those of the two CCP modules from human C1r of innate immunity by NMR spectroscopy. We have also identified the flexible parts of the CCP1-CCP2 module pair that might be responsible for interactions with third molecules and those regions whose flexibility could be modulated by partner binding to the neighboring module. Furthermore, we showed that the two separated modules interact with each other, which observation might have consequences regarding the solution-state dimerisation of full C1r molecules and thus can contribute to further understanding the details of C1 activation (Láng, et al., 2010a, 2010b).

(C.iii) Modules in protein evolution: We have participated in providing an overview of the SBASE database and server, capable of identifying domains in modular proteins from sequence. SBASE is of particular interest in our current and planned studies on protein evolution as its remarkable feature is its capability to detect atypical members of domain families, which exhibit only a moderate similarity to most other members (Dhir, et al., 2010).

(C.iv) dUTPase: We have solved the 1.8 Å-resolution structure of a mycobacterial dUTPase enzyme with and Asp/Asn mutation. Our results show that the role of this residue is not in Mg²⁺ coordination, as previously thought, rather contributing to other types of intra- and intersubunit contacts. Our results contribute to the understanding of elaborate residue-residue interaction networks in this protein family (Takács, et al., 2010).

(D) *Intrinsically disordered proteins*

To explore the evolutionary significance of protein disorder and aggregation, we have started to generate random sequences systematically and to select those with no detectable similarity to known proteins but which are predicted to be well-folded or on the contrary, could be expected to be genuine IDPs. This work is in progress.

We have also conducted a model study to assess the extent and relevance of cross-predictions between folded and unfolded sequences. As a model for the folded state, we have chosen coiled-coil sequences as these adopt a well-defined structured state and are relatively straightforward to predict based on their regularity. Based on our benchmark tests and survey involving the Swiss-Prot database, we propose the simultaneous use of the programs Coils and IUPred to achieve acceptable prediction accuracy and minimise the extent of cross-predictions. The relevance of observed cross-predictions might be that disordered sequences can adopt coiled-coil conformation relatively easily during protein evolution. (Szappanos, et al., 2010).

A newly recognised protein structural motif, the charged single alpha-helix (CSAH) shows similarities both to coiled coils and intrinsically disordered segments. We have investigated these segments in more detail, set up a web server for their detection (<http://csahserver.chem.elte.hu>) and provided a discussion on their possible evolutionary transitions in relation to coiled coils (Gáspári, et al., 2012, Gáspári and Nyitray, 2011).

Structure elucidation by NMR was completed for several of the envisaged proteins. The main goal of these studies is to provide a link between atomic resolution structural information and biological function. Concerning the intrinsically disordered proteins (IDPs) the most recent target of us was dehydrin, termed as ERD14, from *Arabidopsis thaliana*, which protein is highly expressed in somatic cells upon dehydration elicited by water stress, high salinity or cold. The backbone ¹H, ¹⁵N and ¹³C resonance assignment of the 95% of the 185 amino acid-long protein was completed. Its secondary chemical shifts as well as NMR relaxation data were determined and showed that ERD14 is indeed a fully disordered protein under near native conditions, with very short regions of somewhat restricted motion and about 5–15% helical propensity. These results suggested that ERD14 may have partially preformed segmental elements for functional interaction with its partner(s) (Ágoston, et al., 2011).

(E) *Analysis of protein structural ensembles reflecting experimentally determined dynamics*

Our previous work on the use of structural ensembles reflecting the internal dynamics of proteins on different time scales is further elaborated by implementing a web server for the evaluation of such ensembles. The CoNSEnsX (Consistency of NMR-derived Structural Ensembles with experimental data) method is freely available at <http://consensx.chem.elte.hu> (Ángyán, et al., 2010). We have also performed a pilot study to analyse the effect of using dynamically restrained ensembles (instead of single conformers) for structural bioinformatics analysis (CX, DPX, PRIDE and PRIDE-NMR methods). All these investigations prompted us to invoke the idea termed Dynamics–Structure–Activity Relationships (DSARs) of proteins (Gáspári, et al., 2010b, Gáspári and Perczel, 2010).

Dai, X.Y., Gu, X.C., Luo, M., Zheng, X.F. Protein expression, crystallisation and preliminary X-ray crystallographic studies on HSCARG from *Homo sapiens*. 2006. *Protein Pept. Lett.* **13**, 955-957

Publications:

Beke, T., Somlai, C.S., Magyarfalvi, G., Perczel, A., Tarczay, G.Y. Chiral and achiral fundamental conformational building units of β-peptides: a matrix isolation conformational study on Ac-β-HGly-NHMe and Ac-β-HAla-NHMe. 2009. *J. Phys. Chem. B.*, **113**, 7918-7926

Jákli, I., Perczel, A. The inherent flexibility of peptides and protein fragments quantified by CD in conjunction with CCA+. 2009. *J. Pept. Sci.* **15**, 738-752

Ángyán, A.F., Szappanos, B., Perczel, A., Gáspári, Z. CoNSEnsX: an ensemble view of protein structures and NMR-derived experimental data. 2010. *BMC Struct. Biol.* **10**, 39

Dhir, S., Pacurar, M., Franklin, D., Gáspári, Z., Kertész-Farkas, A., Kocsor, A., Eisenhaber, A., Pongor, S. Detecting atypical examples of known domain types by sequence similarity searching: the SBASE domain library approach. 2010. *Curr. Prot. Pept. Sci.* **11**, 515-549

Gáspári, Z., Perczel, A. Protein dynamics as reported by NMR. 2010. In: *Annual reports on NMR spectroscopy*. Academic Press **71**, 35-75

Gáspári, Z., Várnai, P., Szappanos, B., Perczel, A. Reconciling the lock-and-key and dynamic views of canonical serine protease inhibition. 2010a. *FEBS Lett.* **584**, 203-206

- Gáspári, Z., Ángyán, A.F., Dhir, S., Franklin, D., Perczel, A., Pintar, A., Pongor, S.** Probing dynamic protein ensembles with atomic proximity measures. 2010b. *Curr. Prot. Pept. Sci.* **11**, 515-522
- Láng, A., Szilágyi, K., Major, B., Gál, P., Závodszy, P., Perczel, A.** Intermodule cooperativity in the structure and dynamics of consecutive complement control modules in human C1r: structural biology. 2010a. *FEBS J.* **277**, 3986-3998
- Láng, A., Major, B., Szilágyi, K., Gáspári, Z., Gál, P., Závodszy, P., Perczel, A.** Interaction between separated consecutive complement control modules of human C1r: implications for dimerisation of the full-length protease. 2010b. *FEBS Lett.* **584**, 4565-4569
- Pálfi, V.K., Perczel, A.** Stability of the hydration layer of tropocollagen: a QM study. 2010. *J. Comput. Chem.* **31**, 764-777
- Pohl, G., Beke, T., Csizmadia, I.G., Perczel, A.** Extended apolar beta-peptide foldamers: the role of axis chirality on beta-peptide sheet stability. 2010. *J. Phys. Chem. B.* **114**, 9338-9348
- Szappanos, B., Süveges, D., Nyitray, L., Perczel, A., Gáspári, Z.** Folded-unfolded cross-predictions and protein evolution: the case study of coiled-coils. 2010. *FEBS Lett.* **584**, 1623-1627
- Takács, E., Nagy, G., Leveles, I., Harmat, V., Lopata, A., Tóth, J., Vértessy, B.G.** Direct contacts between conserved motifs of different subunits provide major contribution to active site organisation in human and mycobacterial dUTPases. 2010. *FEBS Lett.* **584**, 3047-3054
- Ágoston, B.S., Kovács, D., Tompa, P., Perczel, A.** Full backbone assignment and dynamics of the intrinsically disordered dehydrin ERD14. 2011. *Biomol. NMR Assign.* **5**, 189-193
- Czajlik, A., Hudáky, I., Perczel, A.** Local protein backbone folds determined by calculated NMR chemical shifts. 2011. *J. Comput. Chem.* **32**, 3362-3382
- Gáspári, Z., Nyitray, L.** Coiled coils as possible models of protein structural evolution. 2011. *BioMol. Concepts* **2**, 199-210
- Rovó, P., Farkas, V., Hegyi, O., Szolomajer-Csikós, O., Tóth, G.K., Perczel, A.** Cooperativity network of Trp-cage miniproteins: probing salt-bridges. 2011. *J. Pept. Sci.* **17**, 610-619
- Gáspári, Z., Süveges, D., Perczel, A., Nyitray, L., Tóth, G.** Charged single alpha-helices in proteomes revealed by a consensus prediction approach. 2012. *Biochem. Biophys. Acta* **1824**, 637-646
- Pohl, G., Jakli, I., Csizmadia, I.G., Papp, D., Matias, G.F., Perczel, A.** The role of entropy in initialising the aggregation of peptides: a first principle study on oligopeptide oligomerisation. 2012. *Phys. Chem. Chem. Phys.* **14**, 1507-1516
- Pohl, G., Beke-Somfai, T., Csizmadia, I.G., Perczel, A.** Exploiting diverse stereochemistry of β -amino acids: toward a rational design of sheet-forming β -peptide systems. 2012. *Amino Acids* (in press)

MEXICO

Title: Signal transduction by the BarA/UvrY two component system

Principal Investigator: Dimitrios Georgellis, Lab 226N, Molecular Genetics Department, Universidad Nacional Autónoma de México, Instituto de Fisiología Celular, Circuito Exterior S/N, Col. Copilco, México D.F., Mexico. Tel: +52-55-56225738, Fax: +52-55-56225611, E-mail: dimitris@ifc.unam.mx

ICGEB Contract No.: CRP/08/008

ICGEB Reference No.: CRP/MEX08-02

Abstract: The animal and plant pathogens of the gamma subdivision of *Proteobacteria* cause vast amounts of agricultural damage and human disease. These pathogens include members of the genera *Pseudomonas*, *Erwinia*, *Escherichia*, *Vibrio*, and *Salmonella*. The individual species among these genera are very diverse, as they include free-living species, symbionts, plant pathogens, and animal pathogens. However, they also have striking similarities. For instance, in all five genera the BarA/UvrY two-component system (TCS) has been shown to be required for virulence.

Despite this realisation, two observations motivate us to propose that the primary function of the BarA/UvrY TCS has not yet been discovered. Firstly, this system is found in both pathogens and non-pathogens. Secondly, although *uvrY* is encoded within an evolutionarily conserved region of the various genomes, the virulence genes that it regulates are specific to each pathogen. This strongly suggests that *uvrY* was present in these genomes before the acquisition of the virulence genes that it now controls, and therefore the regulation of these genes must be a relatively new function for the UvrY orthologs. Hence, UvrY must have a more ancient and evolutionarily conserved function(s) that remains to be discovered.

We propose to probe for the structural requirements of BarA for signal reception, to elucidate the routes of phosphoryl-group transfer for signal transmission and signal decay, to search for the member genes of the UvrY regulon, to identify the conserved functions of the various UvrY orthologs, and to screen libraries of small molecules for potential antibacterials that specifically target TCSs. To address these questions we will primarily use *E. coli* as the model organism, and a wide array of biochemical and genetic techniques.

Objectives:

- (i) Probe for the physiological signals that stimulate BarA: We intend to use biochemical and genetic analyses in our searches for the physiological signal(s) that stimulate BarA;
- (ii) Probing for the structural requirements of BarA for signal reception;
- (iii) Structure function relationships in the BarA/UvrY two component system: Biochemical and genetic analyses will be undertaken to probe for the structural requirements of BarA for signal reception, and to elucidate the routes of phosphoryl-group transfer for signal transmission and decay;
- (iv) Search for BarA/UvrY controlled genes: For identifying genes that belong to the UvrY regulon we plan to use a modified protocol of the chromatin immuno-precipitation (ChIP) genome based search. In order to optimise the output of this method the ChIP will be combined with a microarray (ChIP on chip);
- (v) Conserved functions of the various ortholog systems: The UvrY binding site on the promoter of *csrB*, *csrC*, and other genes identified under (d) will be located by foot-printing assays, validated by site directed mutagenesis, and a consensus sequence will be derived. Finally, the derived motif will be used for searches of potential UvrY controlled promoters on different genomes;
- (vi) Antibacterial agents that inhibit signal transduction systems: Two-component systems are considered potential targets for antibacterials. Therefore, we plan to use two independent methods to screen libraries of small molecules with the intention of identifying antimicrobial agents that specifically target these systems.

Results Obtained:

The BarA/UvrY two-component system activates transcription of CsrB and CsrC noncoding RNAs, which act by sequestering the RNA-binding global regulatory protein CsrA. Previously, we demonstrated that pH lower than 5.5 provides an environment that does not allow activation of the BarA/UvrY signaling pathway (Mondragón, V., et al. 2006. J. Bacteriol. 188, 8303-8306). Based on this finding and using genetic and microbial physiology approaches we searched for the physiological signal for the BarA sensor kinase. We discovered that the key metabolites, formate, acetate as well as other aliphatic short chain fatty acids (propionate butyrate, valerate and caproate) act as stimuli that activate the BarA/UvrY two-component system. We also showed that acetate might stimulate UvrY directly via the formation of acetyl-P, thereby linking BarA/UvrY signaling as another target of the "acetate switch". Finally, we demonstrate that acetate signalling via the BarA/UvrY system is physiologically relevant since extracellular acetate accumulation in the growth medium correlated with BarA/UvrY activation. Thus, our results link post-transcriptional regulation by the Csr system to the metabolic state of the cell (Gonzalez Chavez, R., et al. 2010. J. Bacteriol. 192, 2009-2012).

Furthermore, in two subsequent reports, we demonstrate that the BarA/UvrY TCS is integrated in a complex regulatory cascade involving the Csr and stringent response global regulatory systems, and also HilD, which controls expression of the Salmonella SPI-1 and SPI-2 virulence regulons (Edwards, A.N., et al. 2011. *Mol. Microbiol.* 80, 1561-1580; Martínez, L.C., et al. 2011. *Mol. Microbiol.* 80, 1637-1656).

Publications:

Alvarez, A., Georgellis, D. *In vitro* and *in vivo* analysis of the ArcB/A redox signaling pathway. 2010. *Methods Enzymol.* **471**, 205-228

Gonzalez Chavez, R., Alvarez, A., Romeo, T., Georgellis, D. The physiological stimulus for the BarA sensor kinase. 2010. *J. Bacteriol.* **192**, 2009-2012

Barba-Ostria, C., Lledias, F., Georgellis, D. The *Neurospora crassa* histidine kinase DCR-1 is required for normal sexual and asexual development, carotenogenesis, and circadian rhythmicity. 2011. *Eukaryot. Cell* **10**, 1733-1739

Edwards, A.N., Patterson-Fortin, L.M., Vakulskas, C.A., Mercante, J.W., Potrykus, K., Vinella, D., Camacho, M.I., Fields, J.A., Thompson, S.A., Georgellis, D., Cashel, M., Babitzke, P., Romeo, T. Circuitry linking the Csr and stringent response global regulatory systems. 2011. *Mol. Microbiol.* **80**, 1561-1580

Martínez, L.C., Yakhnin, H., Camacho, M.I., Georgellis, D., Babitzke, P., Puente, J.L., Bustamante, V.H. Integration of a complex regulatory cascade involving the SirA/BarA and Csr global regulatory systems that controls expression of the Salmonella SPI-1 and SPI-2 virulence regulons through HilD. 2011. *Mol. Microbiol.* **80**, 1637-1656

Luévano-Martínez, L.A., Barba-Ostria, C., Georgellis, D., Rial, E., Uribe-Carvajal, S. A critical tyrosine residue of the mitochondrial oxalacetate carrier determines its uncoupling protein (UCP)-like function in yeast. 2012. *Biochemical J.* (in press)

MEXICO

Title: Human health and melon biotechnology

Principal Investigator: Miguel A. Gómez Lim, CINVESTAV-Irapuato, Km 9.6 Carretera Irapuato-León, Irapuato, Guanajuato, Mexico. Tel: +52-462-6239679, E-mail: mgomez@ira.cinvestav.mx
ICGEB Contract No.: CRP/09/009

ICGEB Reference No.: CRP/MEX09-01

Abstract: Many epidemiology studies have shown that there exists a relationship between dietary habits and disease risk, indicating that the type of food consumed has a direct impact on health. Biofortification, the process of enriching the nutrient content of crops as they grow, provides a sustainable solution to worldwide malnutrition, because other methods, such as diversifying people's diets or providing dietary supplements, have proved impractical, especially in developing countries. Carotenoids provide essential phytonutrients, such as provitamin A and confer many health benefits, such as reduction in incidence of certain diseases, including cancer, cardiovascular diseases, and age-related eye diseases. Developing carotenoid-enriched food offers a sustainable avenue for improving human nutrition and health. Melon fruit is one of the most important fruits of the world. They contain a cohort of antioxidant enzymes. In this project we have transformed melon with two genes that increase the contents of carotenoids. Five lines were analysed by Southern blot, northern blot and by GUS staining. The molecular assays performed indicate that they contain and express both genes.

Objectives:

- (i) Construction of vectors for expression of CHYB and Or genes under the control of the CaMV promoter;
- (ii) Expression of the constructs from in transgenic melon plants.

Results Obtained:

We have identified five melon lines that contain and express both genes.

PAKISTAN

Title: Genetic basis of hereditary hearing impairment

Principal Investigator: Sheikh Riazuddin, Centre for Molecular Biology, 87-West Canal Bank Road, Lahore-53700, Pakistan & Iqbal Medical College, Jinnah Hospital Complex, Lahore-54550, Pakistan. Tel: +92-42-35161710, Fax: +92-42-5164155, E-mail: riaz@lhr.comsats.net.pk; riazuddin@aimrc.org

ICGEB Contract No.: CRP/08/009

ICGEB Reference No.: CRP/PAK08-01

Abstract: The prevalence of genetic deafness has been estimated on a worldwide basis as 1 per 1000; however, in Pakistan and Middle East, where consanguinity is a social requirement, the prevalence is 40-50% higher than the world average. Mapping of loci and identification of genes are helpful in understanding the functioning of the inner ear, where sound waves are converted into electrical pulses, which are translated in the brain as hearing perception. This is accomplished by a mosaic of supporting cells and sensory hair cells, which are interconnected by uniquely complex tight junction proteins such as Tricellulin.

One hundred large consanguineous families were identified and enrolled from ethnically distinct areas of Pakistan bordering Afghanistan and Iran. DNAs isolated from 1250 blood samples were genetically analysed to identify 25 new/novel mutations, and five new genetic loci. A deaf mouse model was developed to study the malfunctioning of tight junction protein, tricellulin. Clinical examination of the deaf mouse revealed that mutated tricellulin gene affects the functioning of salivary gland, thyroid follicles, heart functioning, epithelia and inner hair cells in the cochlea.

Hearing impairment can lead to impaired ability to develop speech, language and effective communication skills. Identification of the auditory gene, characterisation of pathogenic mutations in these genes and genotype/ phenotype correlation in the deaf mouse is expected to elucidate the molecular basis of recessive deafness caused by malfunctioning of the inner ear. This information will be helpful in developing preventive strategies.

Objectives:

Aim 1: Search for new loci and pathogenic mutations associated with human hearing impairment

- (1.i) Family identification and enrolment;
- (1.ii) Clinical evaluation;
- (1.iii) Sample collection;
- (1.iv) Genotyping to find the linkage to known loci;
- (1.v) Mutations identified;
- (1.vi) Genome wide linkage analysis/high through put sequencing.

Aim 2: Generation of DFNB49 gene knockin and knockout transgenic mouse model and elucidation of the role of tricellulin in hearing impairment

- (2.i) DFNB49 knock-in construct;
- (2.ii) DFNB49 conditional knockout construct;
- (2.iii) Electroporation of the above constructs into 129vsJES cells of the mouse;
- (2.iv) Selection of positive ES cells clones;
- (2.v) Microinjection of positive ES cells into blastocysts to produce chimeric animals;
- (2.vi) Evaluation of the hearing and balance function in *Tric* R497X mice.

Results Obtained:

One hundred large consanguineous families were identified and enrolled in these studies. All members were clinically evaluated to determine if the deafness is syndromic or non-syndromic. Blood samples (1250) were collected from all affected individuals, unaffected siblings, parents and grandparents, if alive. Further, 300 normal random samples were collected from the same population for comparative purposes. Genes were sequenced for mutational analysis and 25 new mutations were discovered. Ten selected unlinked families were screened by whole genome scan and high throughput sequencing, resulting in the discovery of five new loci.

DFNB49 Knock-in and conditional knock-out constructs were developed and used in raising Tricellulin deaf mice. The said two constructs were electroporated into specific mouse cells. Positive clones were selected and introduced by microinjection into blastocysts to produce chimeric animals. The resultant animals were examined for the growth of heart, spleen, liver and kidney, which were all found heavier than those wild type animals. The level of total protein, phosphorous, albumin and triglycerides were also higher in heterozygotes. Enlarged nuclei were seen in a large number of myocytes, as a consequence, the animal suffered from myocardial hypertrophy.

Results Unforeseen in the Original Project:

The project went smoothly. There was no serious constraint. All the family members were fully cooperative and enrolled voluntarily. They all signed consent forms approved by the institutional Ethical Review Board.

Publications:

Anwar, S., et al. SLC26A4 mutation spectrum associated with DFNB4 deafness and Pendred's

Syndrome in Pakistanis. 2009. *J. Hum. Genet.* **54(5)**, 266-270

Rehman, A.U., et al. Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. 2010. *Am. J. Hum. Genet.* **86**, 378-388

Borck, G., et al. Loss-of-function mutations of ILDR1 cause autosomal-recessive hearing impairment DFNB42. 2011. *Am. J. Hum. Genet.* **88(2)**, 127-137

Rehman, A.U., et al. Mutations of GIPC3 cause nonsyndromic hearing loss NB72 but not DFNB81 that also maps to chromosome 19p. 2011. *Hum. Genet.* **130(6)**, 759-765

Waryah, A.M., et al. Molecular and clinical studies of X-linked deafness among Pakistani families. 2011. *J. Hum. Genet.* **56(7)**, 534-540

Ali, R.A., et al. DFNB86, a novel autosomal recessive non-syndromic deafness locus on chromosome 16p13.3. *Clin. Genet.* (in press)

Nayak, G.D., Belyantseva, I.A., Yousaf, R., Edelman, S., Lee, S., Van Itallie, C.M., Rafeeq, M., Sinha, G., Anderson, J.M., Forge, A., Frolenkov, G.I., Riazuddin, S. Tricellulin is required to maintain the barrier function of the reticular lamina and protect the cochlear hair cells from potassium toxicity (in preparation)

Nayak, G.D., Rafeeq, M., Trincot, C.E., Khan, S.N., Riazuddin, S., Ahmed, Z.M., Riazuddin, S. Genotype-phenotype correlation in DFNB49 families with *TRIC* mutations (in preparation)

PERU

Title: An ELISA assay for cysticercosis based on a recombinant cathepsin-L

Principal Investigator: Mirko Zimic, Laboratorio de Bioinformática y Biología Molecular, Departamento de Ciencias Celulares y Moleculares, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, SMP Lima, Peru. Tel: +51-1-3190000, ext. 2604, E-mail: mirko.zimic@upch.pe

ICGEB Contract No.: CRP/08/010

ICGEB Reference No.: CRP/PER08-02

Abstract: Human neurocysticercosis (NCC), caused by the cestode *Taenia solium*, is responsible for a significant amount of neurological morbidity and epilepsy in developing countries. The disease remains highly endemic in many areas, despite several efforts and interventions to control it.

Cathepsin L-like proteases are secreted by several parasites including *Taenia solium*. The mechanism used by *T. solium* oncospheres to degrade and penetrate the intestine and infect the host is incompletely understood. It is assumed that intestinal degradation is driven by the proteolytic activity of enzymes secreted by the oncosphere. Blocking the proteolytic activity by an antibody response would prevent the oncosphere penetration and further infection. Serine and cysteine proteases including chymotrypsin, trypsin, elastase, and cathepsin L, are secreted by *T. solium* and *Taenia saginata* oncospheres when cultured *in vitro*, being potential vaccine candidates. However, the purification of a sufficient quantity of proteases secreted by oncospheres to conduct a vaccine trial is costly and lengthy.

A 53/25kDa cathepsin L-like fraction partially purified from *T. solium* cyst fluid was described previously by our group as an important antigen for immunodiagnostics in a Western blot and in an ELISA.

We initially explored the functional role of the 53/25 kDa cathepsin L-like fraction in the cysticercus. Immunohistochemistry and confocal microscopy were used to localise the 53/25 kDa cathepsin L-like fraction in the brain tissue of naturally infected swine. The effect of this fraction on the proliferation of human peripheral blood mononuclear cells and its capacity to degrade human IgG and bovine serum albumin were tested.

The 53/25 kDa cathepsin L-like fraction was detected in the internal face of the cyst membrane and was seen to migrate to the external face during progression of inflammation. The 53/25 kDa cathepsin L-like fraction did not induce proliferation of human peripheral blood mononuclear cells and was unable to degrade human IgG or BSA.

We confirmed that the 53/25 kDa cathepsin L-like fraction purified from the cysticercus fluid is also present in the *T. solium* oncosphere and is secreted by the cysticercus. This protein fraction was tested for its ability to protect pigs against an oral challenge with *T. solium* oncospheres in a vaccine trial. IgG antibodies against the 53/25 kDa cathepsin L-like protein fraction were elicited in the vaccinated animals but did not confer total protection (i.e., total absence of viable parasites in the pig carcass after a necropsy) although we found a significant trend to reduced number of cysts in the vaccinated group.

The 53/25 kDa cathepsin-L-like protein fraction purified from the cysticercus fluid performed well in immunodiagnostics based on Western immunoblots or standard ELISA. These tests, however, depend on the availability of infected pigs as a source of antigens. In addition, a Western blot is a lengthy and complicated test not able to be performed in rural settings. A simple, cheap and fast diagnostic assay that is suitable for use in field conditions is highly desired.

We tested the 53/25-kDa antigen in a dot-ELISA that could be employed for mass screenings under field conditions. The assay was standardised and tested not only against sera from a large group of NCC cases but also against sera from patients with other common parasitic infections, so that sensitivity and specificity could be assessed.

For NCC, the assay gave better sensitivity in the detection of individuals with extraparenchymal cysts (94.4%–100%) or multiple parenchymal cysts (74.6%–80.0%) than in the detection of individuals with single parenchymal cysts (29.4%–45.1%). The assay also showed a high specificity for NCC (99.0%–100%), with a very low level of cross-reactivity with other parasitic infections. The dot-ELISA developed in this study is a highly specific, simple, cheap and rapid test for NCC that could be used under field conditions, even in the low-resource settings that are common in developing countries.

We analysed the composition and stability of the 53/25 kDa cathepsin L-like purified fraction using mass spectrometry. Treatment with type-O and type-N glycosidases and lectin blots were used to analyse the carbohydrate content. A total of 165 serum samples from patients with neurocysticercosis were evaluated to ascertain the sensitivity of the subunits for diagnosis, and 103 serum samples from negative controls were evaluated to estimate the specificity. Ninety-nine sera from patients with other helminth infections were included to determine cross-reactions.

The 53/25 kDa fraction was shown to be a protein complex comprising four subunits with type-N glycosylation, and to include several lectin-type carbohydrates. The proteomic analysis revealed

that the four subunits comigrated with and showed motifs similar to the 8 kDa lectin-lentil glycoproteins used in the EITB from the USA Centres of Disease Control (CDC). The sensitivity of each of the subunits individually as well as of the combinations was lower than that of the native complex.

The highly antigenic 53 kDa band, associated with cathepsin L-like activity, is a protein complex that comprises four subunits with type-N glycosylation that show motifs of the 8 kDa family antigens. The lack of evidence of a known proteolytic domain in the proteomic analysis, suggests the possibility of the existence of a novel proteolytic domain associated to the 53/25 kDa fraction. None of the single subunits attained the high sensitivity of the native complex, suggesting the presence of quaternary structural epitopes.

Although the mechanism of infection is not completely understood, it is likely driven by proteolytic activity that degrades the intestinal wall to facilitate oncosphere penetration and further infection.

We analysed the publicly available *T. solium* EST/DNA library and identified two contigs comprising a full-length cDNA fragment very similar to *Echinococcus granulosus* Ag5 protein. The *T. solium* cDNA sequence included a proteolytic trypsin-like-domain in the C-terminal region, and a thrombospondin type-1 adherence-domain in the N-terminal region. Both the trypsin-like and adherence domains were expressed independently as recombinant proteins in bacterial systems. TsAg5 showed marginal trypsin-like activity and high sequence similarity to Ag5. The purified antigens were tested in a Western immunoblot assay to diagnose human neurocysticercosis. The sensitivity of the trypsin-like-domain was 96.36% in patients infected with extraparenchymal cysts, 75.44% in patients infected with multiple cysts, and 39.62% in patients with a single cyst. Specificity was 76.70%. The thrombospondin type-1 adherence-domain was not specific for neurocysticercosis.

Objectives:

The goal of this project is to further analyse the 53/25 kDa protein fraction with cathepsin L-like activity purified from *T. solium* cysticercus fluid. This antigen will be used to standardise an Enzyme Linked Immunosorbent Assay (ELISA), for immunodiagnosics of human cysticercosis.

(i) To clone, express and purify an active recombinant cathepsin L included in the 53/25 kDa protein fraction, from cysticercus stage mRNA. Recombinant proteins will be expressed in prokaryote cells.

(ii) To standardise an ELISA assay for human cysticercosis based on the recombinant antigen, and to estimate the sensitivity, specificity and positive/negative predictive values of this test.

Results Obtained:

The 53/25 kDa cathepsin L-like fraction was shown to be a protein complex comprised of four subunits with cysteine proteinase activity with type-N glycosylation, and several lectin-type carbohydrates. The proteomic analysis revealed that the four showed motifs similar to the 8 kDa lectin-lentil glycoproteins used in the CDC EITB. This fraction was detected mainly in the internal face of the cyst and did not induce proliferation of human PBMC. We confirmed that the 53/25 kDa cathepsin L-like fraction was also present in the *T. solium* oncosphere and was secreted by the cysticercus. Although this fraction elicited IgG antibodies in vaccinated pigs, it did not confer total protection. However, a significant trend to reduced number of cysts in the vaccinated group was observed. The 53/25 kDa cathepsin-L-like protein fraction purified from the cysticercus fluid performed well in immunodiagnosics based on a dot-ELISA. Other proteolytic antigens, including a cathepsin L-like and a trypsin-like protease have been cloned and expressed recombinantly, and are being tested in immunoassays.

Results Unforeseen in the Original Project:

We confirmed that the 53/25 kDa antigen is actually a protein complex comprising four subunits that are likely to be the 8kDa LLGP family proteins of the CDC EITB assay.

Publications:

Piña, R., Gutiérrez, A.H., Gilman, R.H., Rueda, D., Sifuentes, C., Flores, M., Sheen, P., Rodriguez, S., García, H.H., Zimic, M. A dot-ELISA using a partially purified cathepsin-L-like protein fraction from *Taenia solium* cysticerci, for the diagnosis of human neurocysticercosis. 2011. Ann. Trop. Med. Parasitol. **105(4)**, 311-318

Rueda, A., Sifuentes, C., Gilman, R.H., Gutiérrez, A.H., Piña, R., Chile, N., Carrasco, S., Larson, S., Mayta, H., Verástegui, M., Rodriguez, S., Gutiérrez-Correa, M., García, H.H., Sheen, P., Zimic, M. TsAg5, a *Taenia solium* cysticercus protein with a marginal trypsin-like activity in the diagnosis of human neurocysticercosis. 2011. Mol. Biochem. Parasitol. **180(2)**, 115-119

Zimic, M., Pajuelo, M., Gilman, R.H., Gutiérrez, A.H., Rueda, L.D., Flores, M., Chile, N., Verástegui, M., Gonzalez, A., García, H.H., Sheen, P.; for the Cysticercosis Working Group in Peru. The highly antigenic 53/25kDa *Taenia solium* protein fraction with cathepsin-L like activity is present in the oncosphere/cysticercus and induces non-protective IgG antibodies in pigs. 2012. Vet. Immunol. Immunopathol. **145(1-2)**, 171-178

POLAND

Title: TFE as the 3D-fold inducer. Application for domain 4 of *E. coli* RNAP σ^{70}

Principal Investigator: Jarosław Poznański, Department of Biophysics, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a, 02-106 Warszawa, Poland. Tel: +48-22-5925783, Fax: +48-22-5922190, E-mail: jarek@ibb.waw.pl

ICGEB Contract No.: CRP/08/011

ICGEB Reference No.: CRP/POL08-03

Abstract: Domain 4 of *E. coli* (σ^{70}_4) RNA polymerase holoenzyme (RNAP) is responsible for the recognition of the -35 element of cognate promoters during the transcription initiation. Its structure in a free form has not been yet determined, and the only accessible data come from the homologous domains of thermophilic bacteria, and for σ^{70}_4 from the two complexes with its natural cofactors. We have found that our construct of σ^{70}_4 (rECo $^{70}_4$) is almost insoluble in the neutral aqueous solution, but in acidic conditions (pH \sim 2.8) it displays propensities towards a low-populated structural arrangement that generally agrees with the known structures of σ^{70}_4 homologues (Biochemistry 2003, **42**, 13438). We have also proved that 2,2,2-trifluoroethanol (TFE), which is a well known protein-folding inducer preferentially stabilising α -helical protein structures, significantly increases the protein solubility and induces not only the helical secondary structure, but also the canonical HLTH fold, which we have successfully determined by the combination of NMR data and molecular modelling. Screening for solution composition combined with the detailed analysis of the collected ^{15}N relaxation data enabled us to propose the system suitable for the study of the dynamics of the DNA-recognising HLTH motif in a solution. However, the limited solubility of the engineered protein sample at neutral pH enables the acquisition of basic 2D heteronuclear NMR spectra, at the moment it precludes NMR determination of the protein structure or detailed relaxation study. In the progress, we have additionally demonstrated how heteronuclear NMR spectroscopy supplemented by molecular dynamics methods can be applied for the structural studies of the separated protein domains, which lost their structure when cut from the whole protein context.

Objectives:

The main objectives of this project were:

- (i) Determination of the stability of the TFE-induced fold using CD-monitored unfolding in various solvent systems;
- (ii) Estimation of the local mobility of the protein NMR with the aid of ^{15}N relaxation estimated for protein in rationally-selected solvents;
- (iii) Determination of at least low/resolution structure of the protein;
- (iv) To propose protein sequence and solvent composition that enables the determination of protein solution structure;
- (v) Comparison of the determined solution structure of *E. coli* σ^{70}_4 with those of the homologous subdomains of thermophilic bacteria of a known structure.

Results Obtained:

An Intensive Circular Dichroism (CD) study demonstrated that the aqueous solution of TFE significantly increases both the contents of the helical structure of rECo $^{70}_4$ and its thermal stability, but also decreases the free energy of protein unfolding, thus, acting simultaneously as a structure inducer and denaturant. The wide screening pointed at some promising compositions of solution that may be applicable for structural studies. Thus, rECo $^{70}_4$ in the presence of 1.8M MgSO $_4$ in 10mM sodium cacodylate buffer at pH \sim 7 remains a good starting point for solvent optimisation directed at protein crystallisation. 100mM NDSB-195 at pH \sim 7 was found to be a good folding agent, and the free energy difference between the folded and unfolded states was estimated to equal 1(0.3) kcal/mol. This solvent system could be used in an NMR application, although, besides the acquisition of the basic heteronuclear HSQC spectra, the limited protein solubility (50 μM achieved for the best system) precluded at the moment the application of any NMR structural and/or relaxation studies.

We have demonstrated that TFE does not act as a non-specific helical structure inducer or denaturant, but also that for the aqueous solution of rECo $^{70}_4$ the stepwise addition of TFE causes, in its 10% content, an almost cooperative reorganisation of secondary structure of the limited set of residues. Interestingly, their location within the protein sequence coincides with the location of non-helical elements in the structures of homologous proteins. Complete NMR studies were carried out for the aqueous solution of protein at pH \sim 4.4 in the presence of either 10 or 30% of TFE. In lower TFE content, the TFE-induced helical structures were found to be loosely interacting, and on the basis of ^{15}N relaxation and CSI data (chemical shift index), only conformational preferences and an internal motion of the individual residues in protein could have been modeled (Proteins 2010, **78**, 754). In high-TFE solution, the NMR studies concerned both the structure and the internal dynamics of the protein. Using the limited set of distance restraints deduced from 3D NOESY spectra, extended by the additional set of structural constraints obtained from CSI data

with the use of TALOS+ software, we succeeded in the determination of the low-resolution solution structure of the DNA-recognising HLTH motif (manuscript in preparation). The experimental data enabled the estimation of protein backbone arrangement with the accuracy better than 2Å. However, it should be stressed that, according to ¹⁵N relaxation data, the proteins experience conformational exchange, on the μs to ms timescale, between the proposed folded state and an ensemble of unfolded forms, and only the structure of the folded component could be determined. The backbone architecture of rECo⁷⁰₄ that is built of the three helices is similar to that of the HLTH motif in the closest homologue of *Thermotoga maritima* (TMσ^A₄). RMSD difference between these two subunits, calculated for the CA backbone atoms, is high (11Å), although for the framework of H1 and H3 helices it is reduced to 1.5 Å. This confirms the high functional homology of these two regions. Moreover, the H1 and H3 helices in both proteins have the same arrangement. However, the helix H1 of rECo⁷⁰₄ is slightly longer than that of TM σ^A₄. The H2 region, which is involved in DNA sequence recognition, displays differences in the relative orientation in both proteins. The loop separating H1 and H2, as well as the short turn between H2 and H3, are longer in rECo⁷⁰₄, and their arrangement is visibly different from that observed in TMσ^A₄. Only H3 region seems to be of a similar length. At the moment, on the basis of the performed comparative analysis of HLTH motifs, it should be pointed out that secondary structure elements are not of an identical length, while the general organisation and secondary structure distribution along the sequence remain the same, however, the promoter sequence preferences may be different due to divergence in spatial location of the H2 helix.

The detailed analysis of the structural data followed by relaxation analysis showed that the large N-terminal part of rECo⁷⁰₄ displays substantially increased conformational flexibility. We have tested several oligopeptides corresponding to the rationally chosen C-terminal parts of ECo⁷⁰₄. The most promising results were obtained for a 64-residue fragment Gly550-Asp613 of the whole *E. coli* sigma factor, although the solubility of the protein sample in non-TFE systems was, even at moderately acidic pH, limited to 50 μM. However, C-terminal His₆-tag improves protein solubility. Despite the fact that the actual solubility is very limited, the application of a combination of salts and some low-mass protein stabilisers may help in preparing a sample for NMR structural studies.

Results Unforeseen in the Original Project:

All the rationally chosen protein fragments displayed extremely low solubility at neutral pH. The application of protein stabilisers resulted in a very limited improvement of protein solubility. The complex formed with DNA precipitates.

Publications:

Kaczka, P., Polkowska-Nowakowska, A., Bolewska, K., Zhukov, I., Poznański, J., Wierzchowski, K.L. Backbone dynamics of TFE-induced native-like fold of domain 4 of *Escherichia coli* RNA polymerase σ70 subunit. 2010. *Proteins* **78**, 754-768

Kaczka, P., Polkowska-Nowakowska, A., Bolewska, K., Zhukov, I., Wierzchowski, K.L., Poznanski, J. A native-like fold of region 4 of initially unfolded sigma(70) subunit of *E. coli* RNA polymerase induced by TFE. 35th Congress of the Federation of European Biochemical Societies, Gothenburg, Sweden, 26 June – 1 July 2010. *FEBS J.* **277(1)**, 255

Wąsik, R., Wińska, P., Poznański, J., Shugar, D. Isomeric mono-, di-, and tri-bromobenzo-1H-triazoles as inhibitors of human protein kinase CK2α. 2012. *PLoS One* **7(11)**, e48898

Sitkiewicz, E., Tarnowski, K., Poznanski, J., Kulma, M., Dadlez, M. Oligomerization interface of RAGE receptor revealed by MS-monitored hydrogen deuterium exchange. 2013. *PLoS One* **8(10)**, e76353

Kaczka, P., Winiewska, M., Zhukov, I., Rempola, B., Bolewska, K., Lozinski, T., Ejchart, A., Poznańska, A., Wierzchowski, K.L., Poznański, J. The TFE-induced transient native-like structure of the intrinsically disordered σ704 domain of *Escherichia coli* RNA polymerase. 2014. *Eur. Biophys. J.* **43**, 581-594

Maciejewska, A.M., Poznanski, J., Kaczmarska, Z., Krowisz, B., Nieminuszczy, J., Polkowska-Nowakowska, A., Grzesiuk, E., Kusmierek, J. AlkB dioxygenase preferentially repairs protonated substrates. 2013. *J. Biol. Chem.* **288(1)**, 432-441

RUSSIA

Title: IsoAsp7-A β as potential trigger of Alzheimer's disease

Principal Investigator: Alexander A. Makarov, Laboratory of Conformational Stability of Proteins and Physical Methods of Analysis, Engelhardt Institute of Molecular Biology RAS, Vavilov str. 32, Moscow 119991, GSP-1, Russia. Tel: +7-499-1354095, Fax: +7-499-1351405, Email: aamakarov@eimb.ru

ICGEB Contract No.: CRP/08/012

ICGEB Reference No.: CRP/RUS08-02

Abstract: Alzheimer's disease (AD) is characterised by the formation of amyloid plaques as a result of the extracellular accumulation of the amyloid- β peptide (A β) in the brain of AD patients. The A β aggregates contain high concentrations of zinc ions bound to A β , and most of A β aspartate-7 amino-acid residues undergo a specific chemical modification resulting in the formation of isoaspartate-7. The project has addressed the role of the isoAsp7-containing A β (isoA β) and its interactions with metal ions in AD pathogenesis. We have developed and tested an analytical MS-based method for measuring the relative abundance of isoA β in blood samples. *In vitro* experiments have allowed us to establish the following A β structural determinants: (i) the minimal zinc-binding site, (ii) the primary zinc-binding site, and (iii) the site mediate zinc-induced dimerisation. We also have shown that isoA β has significant differences from intact A β in terms of interactions with copper ions. Additionally, 3D structures of the isoA β and rat A β zinc-binding domains are solved. Finally, results of *in vivo* experiments have shown that intravenously injected isoA β triggers cerebral amyloidosis in transgenic mice.

Objectives:

(i) To develop an appropriate analytical method for measuring the relative abundance of isoA β in blood samples derived from healthy individuals and AD patients in order to determine whether this parameter could be used as a biomarker for diagnostic purposes.

(ii) To determine the 3D structures of the human isoA β and rat A β zinc-binding domains and to investigate the molecular mechanism of interactions between zinc/copper ions and various natural variants of the A β metal-binding domain 1-16.

(iii) To investigate whether isoA β is the active pathogenic agent of amyloid plaque extracts, which induce cerebral amyloidogenesis in transgenic mice, so as to determine if isoA β could be used as a target molecule for AD therapy.

Results Obtained:

(i) To evaluate levels of isoA β in human blood samples, a quantitative method for determining of the relative abundance of the isomerised and the native forms of A β in their binary mixtures was developed. This method is based on the differences in the MS fragmentation spectra of the two isoforms of A β . IsoA β demonstrates fragments c6+57 and z10-57 in the ECD spectra, and fragment b6+H₂O in the CID spectra, which are not present in the fragmentation spectra of the normal peptide. Using synthetic domain 1-16 of A β (A β 16) and its isoaspartate-7-containing isoform (isoA β 16) ESI MS calibration curves of characteristic ECD and CID fragments of isoA β were obtained and then used for quantitative analysis of the relative abundance of the isomerised form in the A β fraction extracted from plasma samples. It was determined that for the 4 patients with Alzheimer's disease in heavy stages the relative abundance of the isomerised fragment 1-16 of the A β peptide is about 6% (using CID LTQ and CID FT methods), while in the samples from other donors only the native A β fragments were detected.

(ii) The thermodynamics of the binding of Zn²⁺ to A β fragments and A β (1-16) mutants was characterised by isothermal titration calorimetry (ITC) in order to determine the minimal Zn²⁺-binding site of A β under physiological conditions. The amino acid region 6-14 of A β was determined as the minimal Zn²⁺-binding site wherein the ion is coordinated by His6, Glu11, His13, and His14. We also used the quantum mechanics/molecular mechanics (QM/MM) method to model Zn²⁺ recognition by A β . Both ITC and QM/MM showed that three of four residues from the A β region 11-14 (EVHH) contribute to zinc binding, and that this tetrapeptide readily forms dimers linked through a zinc ion similarly to a model proposed for A β aggregation in a recent molecular dynamics study by Miller et al. (2010). These data allow to us consider A β (11-14) tetrapeptide as a primary Zn²⁺-recognition site of A β and an important drug target candidate to prevent Zn²⁺-induced aggregation of A β .

(iii) The interactions of the immobilised A β metal-binding domain with soluble peptides corresponding to the A β (1-16) and its fragments were characterised using surface plasmon resonance biosensing in order to localise the interface of metal-induced A β dimerisation. We found that zinc induces dimerisation of the A β domain 1-16, while copper does not influence this process. The amino acid region 11-14 has been determined as the critical site for Zn²⁺-induced dimerisation of A β (1-16). The stoichiometry of zinc-bound complexes mimicking the A β (1-16) dimerisation interface was studied by electrospray ionisation mass spectrometry (ESI-MS). ESI-MS experiments have demonstrated that only one zinc ion participates in binding of the interacting A β molecules.

So, our results have shown that the primary zinc recognition site of A β (residues 11-14) controls Zn²⁺-induced dimerisation of A β (1-16). Since this process may be the trigger for further formation of pathogenic A β oligomers and aggregates, inhibition of this site to prevent its interactions with zinc ions may be a potential strategy in prevention of Alzheimer's disease onset and progression.

(iv) Effect of isomerisation of aspartate 7 in chelating of the copper ions by the metal-binding domain of A β was examined by isothermal titration microcalorimetry. It was found that the N-terminal amino group, which is the copper chelator in the native A β , does not participate in the coordination of the copper ion by the A β fragment 1-16 containing isoaspartate instead of aspartate 7. Thus, the isomerisation of aspartate 7 significantly changes the structure of the complex of A β with copper and consequently might play a pathogenic role in Alzheimer's disease.

(v) Solution structure of isoA β (1-16) in the free-state was determined by NMR spectroscopy and molecular modeling. The data show that in the absence of zinc ions the 3D structures of isoA β (1-16) and native A β (1-16) are quite similar, whereas zinc binding significantly changes zinc coordination by these peptides.

(vi) In an attempt to reveal the mechanism of rats' resistance to Alzheimer's disease, the structure of the metal-binding domain 1-16 of rat amyloid- β (rat A β (1-16)) in solution in the absence and presence of zinc ions was determined. Zinc induced dimerisation of the domain was detected. The zinc coordination site was found to involve residues His-6 and His-14 of both peptide chains. We used experimental restraints obtained from analyses of NMR and isothermal titration calorimetry data to perform structure calculations. The calculations employed an explicit water environment and a simulated annealing molecular-dynamics protocol followed by quantum-mechanical/molecular-mechanical optimisation. We found that the C-tails of the two polypeptide chains of the rat A β (1-16) dimer are oriented in opposite directions to each other, which hinders the assembly of rat A β dimers into oligomeric aggregates. Thus, the differences in the structure of zinc-binding sites of human and rat A β (1-16), their ability to form regular cross-monomer bonds, and the orientation of their hydrophobic C-tails could be responsible for the resistance of rats to Alzheimer's disease.

(vii) In order to test our hypothesis about potential pathogenicity of isoA β , *in vivo* investigations of the seeded aggregation of A β were carried out by using commercially available APP/PS1 mice of strain B6C3-Tg(APP^{swe}, PSEN1^{dE9})85Dbo/j received from the Jackson Laboratory (USA) (stock number 004462). These double transgenic mice manifest characteristic cognitive features of AD-like pathology and possess significant amounts of dense congophilic amyloid plaques starting from 6-7 month age. To test whether inoculation of synthetic peptide isoA β might trigger A β misfolding and deposition in the brain, we administered each month intravenous injections (100 μ l each; 50 μ g of isoA β) to two cohorts of young (2-month-old) female APP/PS1 tg mice. After performing injections (5 for one cohort and 9 for the other), the host mice were analysed. IsoA β induced robust cerebral β -amyloidosis in all intravenously inoculated mice compared with untreated littermate controls, whereas intravenous inoculation with phosphate-buffered saline or synthetic peptide A β was ineffective.

Results Unforeseen in the Original Project:

A method for structure simulation of small peptides in aqueous environment has been developed using the Amber-03 force field and Gromacs program after modification of its code.

Publications:

Istrate, A.N., Tsvetkov, P.O., Mantsyzov, A.B., Kulikova, A.A., Kozin, S.A., Makarov, A.A., Polshakov, V.I. NMR solution structure of rat a β (1-16): Toward understanding the mechanism of rats' resistance to Alzheimer's disease. 2012. *Biophys. J.* **102(1)**, 136-143

Shelkovnikova, T.A., Kulikova, A.A., Tsvetkov, P.O., Peters, O., Bachurin, S.O., Buchman, V.L., Ninkina, N.N. Proteinopathies - forms of neurodegenerative disorders with protein aggregation-based pathology. 2012. *Mol. Biol.* **46(2)**, 172-183

Indeykina, M.I., Popov, I.A., Kozin, S.A., Kononikhin, A.S., Kharybin, O.N., Tsvetkov, P.O., Makarov, A.A., Nikolaev, E.N. Capabilities of MS for analytical quantitative determination of the ratio of α - and β Asp7 isoforms of the amyloid- β peptide in binary mixtures. 2011. *Anal. Chem.* **83(8)**, 3205-3210

Kozin, S.A., Mezentsev, Y.V., Kulikova, A.A., Indeykina, M.I., Golovin, A.V., Ivanov, A.S., Tsvetkov, P.O., Makarov, A.A. Zinc-induced dimerisation of the amyloid- β metal-binding domain 1-16 is mediated by residues 11-14. 2011. *Mol. Biosyst.* **7(4)**, 1053-1055

Istrate, A.N., Mantsyzov, A.B., Kozin, S.A., Pol'shakov, V.I. Optimisation of the methods for small peptide solution structure determination by NMR spectroscopy. 2010. *Mol. Biol.* **44(6)**, 1075-1085

Tsvetkov, P.O., Kulikova, A.A., Golovin, A.V., Tkachev, Y.V., Archakov, A.I., Kozin, S.A., Makarov, A.A. Minimal Zn⁽²⁺⁾ binding site of amyloid- β . 2010. *Biophys. J.* **99(10)**, L84-86

Tsvetkov, F.O., Makarov, A.A., Archakov, A.I., Kozin, S.A. Effect of isomerisation of aspartate-7 on the binding of copper (II) ion by the beta-amyloid peptide. 2009. *Biofizika* **54(2)**, 197-201

SERBIA

Title: Structural alterations of DNA induced by initiator protein

Principal Investigator: Jelena Kusic, Laboratory for Molecular Biology, Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Vojvode Stepe 444a, P.O. Box 23, 11010 Belgrade, Serbia. Tel: +381-11-3976658, Fax: +381-11-3975808, E-mail: jkusic@imgge.bg.ac.rs
ICGEB Contract No.: CRP/08/013

ICGEB Reference No.: CRP/YUG08-01

Abstract: To ensure faithful replication of the genome in a limited time period, replication of chromosomal DNA in eukaryotes initiates at multiple DNA sites called replication origins. Initiation of DNA replication is highly coordinated biological process involving numerous protein-DNA and protein-protein interactions. Nevertheless, it is still poorly understood how particular locations in the chromosomes are selected as origins of replication, especially in the metazoan where the mode of interactions between initiator proteins and DNA are largely unclear.

In eukaryotes, DNA replication is initiated by the Origin Recognition Complex (ORC), which binds DNA at replication origins and promotes the assembly of the pre-Replication Complex (pre-RC). In all eukaryotes homologues of ORC have been identified and biochemical studies revealed their requirement for the initiation of DNA replication, suggesting that the mechanism of initiation is conserved.

Although there is striking conservation of ORC, sequence requirements for ORC binding are not well understood. Even in budding yeast, in which well-defined sequences serve as origins, specific sequences are not required for ORC binding *in vitro*. Low sequence specificity of metazoan ORC is consistent with the lack of common sequence motifs within metazoan origins, but does not match the selectivity of initiation observed in somatic cells. Origin transfer studies, in combination with mutational analysis of native and ectopic initiation sites, demonstrate that even metazoan origins function as replicators, and therefore must contain required genetic elements. The most common property of metazoan replication origins is frequent occurrence of the clustered adenines and thymines distributed in either alternating or homogenous fashion. Since different pairs of clustered residues relate to each other as either mirror or glide reflection images, they could interact in alternative manners and force interacting DNA to assume unorthodox structure. The capacity of origin DNA to build specific structural features could thus define replicator function.

Structural alteration of origins could be induced or stimulated by initiators. All six ORC subunits belong to the initiator subgroup of the AAA+ superfamily of proteins, which undergoes conformational changes or induces changes in interacting partners after binding of ATP. As suggested by the mechanism of prokaryotic origin remodelling and emerging analogy between eukaryotes and prokaryotes in initiator structure and function, eukaryotic initiators are expected to remodel DNA in nucleotide-dependent manner.

Previously, we have identified specific unorthodox structures in DNA fragments IboI and SphI-EcoRV, isolated from the essential and functionally interchangeable regions of the human LMNB2 and the hamster DHFR replication origins and we have observed the ability of HsOrc4 to stimulate formation of these structures. To better characterise HsOrc4's action towards different DNA substrates, we have used oligonucleotides that mimic origin regions involved in alternative interactions and we have registered the action of the wild type HsOrc4 or its mutants, deficient in ATP binding or hydrolysis. We have observed that HsOrc4 stimulated formation of previously identified DNA structures such as TAT triplexes and created novel structures such as AA duplexes or quadruplexes. For catalytic action of HsOrc towards oligoadenine substrates an intact ATP binding site was strictly required.

In this project we have focused on the role of HsOrc4 and human ORC holocomplex in origin restructuring. In addition to the above described assays, we used circular, origin containing molecules as substrates for HsOrc4. In circular molecules DNA unwinding is promoted by negative supercoiling, which is a far superior and more natural manner of strand separation than thermal denaturation. Based on our preliminary analysis with different agents we have chosen the plasmid pMCD as substrate for protein mediated origin restructuring. This plasmid makes a convenient substrate because it contains the whole DHFR origin and one unorthodox element identified in the SphI-EcoRV region.

In addition to assays on linear DNA fragments (different electrophoretic techniques, chemical assays with agents sensitive or resistant to single stranded DNA, substitution of 7-Deaza-dATP for dATP), we have established methods specific for circular, origin containing DNA molecules - mung bean nuclease sensitivity assays of the chosen origin regions and topoisomerase footprints on chloroquine gels. Mung Bean nuclease sensitivity assays enable identification of DNA sequences in which protein mediated structural alteration occurred, whereas topoisomerase footprints register topological changes of DNA induced by protein. Combined results points to the capacity of essential *ori* elements to adopt altered DNA structures and to the ability of HsOrc4 to direct structural alterations of DNA, as potentially important elements of origin activation.

Objectives:

The primary goal of our research project was better understanding of the sequence elements involved in origin restructuring during initiation of DNA replication and the potential of initiator proteins to direct this structural alterations. The DNA replication origins of metazoan genomes are the sites of complex protein-DNA interactions. Previously, we have observed that one subunit of the Origin Recognition Complex, Orc4, could mediate formation of unorthodox structures from linear, single-stranded DNA fragments. Since (-) supercoiling is more natural manner of strand separation we investigated how different DNA topology conformations and Orc4 binding affect origin structure.

Results Obtained:

MB nuclease assay demonstrated that, in negatively supercoiled DNA and at neutral pH, AT-rich regions of hamster DHFR origin of replication adopt unorthodox DNA structure. Following addition of initiation protein HsOrc4, this structure partly disappears indicating either protein binding to unorthodox DNA structure or structural DNA remodelling mediated by protein. Analysis of MB sensitivity in entire plasmid revealed that in topologically constrained DNA structural alteration in *ori* sequences could be transmitted across DNA after binding of initiator protein HsOrc4. Topoisomerase relaxation assays has confirm our assumption that HsOrc4 mediates topology changes in pMCD. DNA remodelling by protein depends on (-) supercoiled state of plasmid DNA and are stabilised by topoisomerase I action. Taken together our results bring us a step closer to understanding the mode of interactions between initiator proteins and *ori* sequences.

Results Unforeseen in the Original Project:

We assumed that HsOrc4 will induce topology changes in constrained DNA molecules, but direction of this change was unknown. The topoisomers distribution toward lower mobility compared to reactions without the protein indicated introduction of positive writhe by initiation protein HsOrc4.

Publications:

Divac, A., Tomic, B., Kusic, J. The role of ATP in human ORC4 protein function. 2010. J. Serb. Chem. Soc. **75**, 317-322

Kusic, J., Tomic, B., Divac, A., Kojic, S. Human initiation protein Orc4 prefers triple stranded DNA. 2010. Mol. Biol. Rep. **37**, 2317-2322

SLOVAKIA

Title: Misfolded tau protein and oxidative defense system in tauopathies

Principal Investigator: Peter Filipcik, Laboratory of Cellular and Molecular Neurobiology, Institute of Neuroimmunology, Dubravská Cesta 9, Bratislava, Slovakia. Tel: +421-2-54788100, Fax: +421-2-54774276, E-mail: peter.filipcik@savba.sk

ICGEB Contract No.: CRP/08/014

ICGEB Reference No.: CRP/SVK08-01

Abstract: Oxidative stress has been implicated in the pathogenesis of many neurodegenerative diseases including different forms of tauopathy and Alzheimer's disease. We have investigated the role of human tau protein in the process of neurodegeneration in context with oxidative stress. Previously, we observed significant accumulation of ascorbyl free radicals in brains of transgenic animals expressing pathological form of tau protein. We have, therefore, focused to more detailed understanding of relationships between tau protein modifications and oxidative stress. Here we discovered that Vitamin C was found to be selective in the free radicals scavenging activity in neurons, suggesting that expression of truncated tau protein preferentially leads to increases in aqueous phase oxidants and free radicals. Our transcriptional analysis of brain tissue using PCR arrays revealed several dysregulated genes coding for the proteins in redox signaling pathways. We have also pointed to possible mechanisms that can lead to generation of pathological forms of tau protein, which subsequently can induce the oxidative stress in neurons. We hypothesise that environmental stress can lead to generation of pathological forms of tau protein, either hyperphosphorylated or truncated, which may precipitate neurodegeneration, induce oxidative stress and inflammation in vulnerable populations of brain neurons.

Objectives:

(i) Analysis of the pathogenic events associated with tau protein modifications and oxidative stress, the common risk factor of neuronal degeneration:

(i.a) at the level of recombinant proteins;

(i.b) "in vitro" model of cultured cells (modified neuroblastoma cells and rat primary neurons) and;

(i.c) "in vivo" system using transgenic animals expressing pathologically modified tau protein.

(ii) Determination of temporal order and sequence of pathogenic events at early stage of pathogenesis in human tauopathies and Alzheimer's disease.

Results Obtained:

Within the project we have achieved the following results:

(i) Oxidative stress strongly increases oligomerisation of recombinant truncated tau proteins, which may accelerate accumulation of aberrant tau in neurons when exposed to oxidative conditions;

(ii) Expression of truncated tau protein leads to formation of ROS in the aqueous phase and vitamin C is selective scavenger of ROS generated in truncated tau-protein expressing neurons;

(iii) The pathogenic modifications of tau protein may precede oxidative damage in tauopathies and Alzheimer's disease.

Results Unforeseen in the Original Project:

(i) Our data suggest that hypoxic stress alters 3R-4R ratio in cerebro-cortical neurons, which may lead to precipitation of neuropathogenic changes typical for specific tauopathies.

(ii) Immobilisation stress leads to induction of pathologically phosphorylated forms of tau protein in different brain areas of acutely stressed experimental animals, which may subsequently lead to redox disbalance in neuronal tissue.

Publications:

Cente, M., Filipcik, P., Mandakova, S., Zilka, N., Krajciová, G., Novak, M. Expression of a truncated human tau protein induces aqueous-phase free radicals in a rat model of tauopathy - implications for targeted antioxidative therapy. 2009. J. Alzheimers Dis. **17**, 913-920

Filipcik, P., Novak, P., Mravec, B., Ondicová, K., Krajciová, G., Novak, M., Kvetnansky, R. Tau protein phosphorylation in diverse brain areas of normal and CRH deficient mice: Up-regulation by stress. 2012. Cell. Mol. Neurobiol. (in press)

SLOVENIA

Title: Molecular mechanism of the recognition of double stranded viral RNA mediated by TLR3

Principal Investigator: Mateja Manček-Keber, Department of Biotechnology, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia. Tel: +386-1-4760393, Fax: +386-1-4760300, E-mail: mateja.mancek@ki.si

ICGEB Contract No.: CRP/08/015

ICGEB Reference No.: CRP/SLO08-01

Abstract: The innate immune response is an essential component of our defense against pathogenic microorganisms. TLRs recognise molecules characteristic for pathogenic microorganisms (PAMPS). Endosomal TLRs (TLR3, TLR7, TLR8, and TLR9) recognise different classes of bacterial, viral and endogenous nucleic acids. TLR3 is activated by ds RNA, which is characteristic for viruses.

In order to design potential inhibitory molecules of endosomal TLRs, a resolved mechanism of ligand binding is of high importance. Although a crystal structure of TLR3 has been resolved this dimer cannot explain why molecules smaller than 46bp such as siRNA can also activate TLR3. That is why we proposed a second binding site on the N-terminal side of TLR3 ectodomain and a different mechanism of TLR3 positioning after ligand binding. Moreover, we showed additional important feature corresponding only to TLR3 that is cell surface expression when chaperon UNC93B1 is coexpressed. This surface expressed TLR3 might enable cells to respond to released RNA and thus amplifying antiviral response.

Small interfering RNA molecules (siRNA) are short dsRNA molecules, which are widely used for silencing and also in a variety of therapeutic applications. Their off-target activity can potentially lead to unwanted toxicities. Our goal was also to study one of the side effects of siRNA-based methods, namely activation of the TLR3, which limits the use of siRNA in therapeutic applications.

Objectives:

After the crystal structure of human TLR3 had been solved and site-directed mutagenesis of individual amino acid residues within the ectodomain were done, the C-terminal binding site was discovered. It includes H539 and N541. The resulting binding model showed dsRNA binding to the C-terminal site on the nonglycosylated face of the ectodomain. According to this model dsRNA shorter than 46 bp should not activate TLR3, but reports stated that the activation limit was 22 bp. We proposed the existence of a second, N-terminal binding site. For this part we structured the experimental plan where we selected, prepared and tested point mutations of TLR3. To evaluate the mutations also binding studies of dsRNA to TLR3 mutants were performed. Additionally, we proposed a binding model in which the ectodomains are shifted for the length of one dsRNA turn closer together. In our model, TLR3 can be activated by ligands of a minimal length of 21 bp such as siRNA.

Additional work as proposed was done in this project. TLR movement to endosomes is driven by UNC93B1 chaperon. Surprisingly, we observed that UNC93B1 also mediated a translocation of TLR3 to the cell surface, so we tried to determine, which segment of TLR3 governs this translocation. Cell surface or endosomal localisation of TLR3 might have impacts on its signalling properties, so we also tested how would surface expression influence poly (I:C) as well as siRNA recognition and signaling. We also tried to explain what could be the physiological consequences of cell surface expression in antiviral response.

Smaller part of the project was also extended to other endosomal TLR receptors (TLR7, TLR8 and TLR9), which recognise nucleic acids other than dsRNA. The antimalarial drugs chloroquine, hydroxychloroquine, and quinacrine have been used for a long time to treat immune-mediated inflammatory disorders such as SLE, rheumatoid arthritis, and Sjogren's syndrome. Chloroquine and its analog quinacrine inhibit CpG DNA-driven cellular activation. With the aim to clarify the mechanism of endosomal TLR inhibition, we therefore investigated direct binding of antimalarials to nucleic acid TLR ligand.

Results Obtained & Unforeseen in the Original Project:

(i) A second binding site for double-stranded RNA in TLR3 and consequences for interferon activation

We showed that mutations of H39 and H60 completely abrogated activation of TLR3. The presence of two binding sites in the ectodomain probably determines the substrate specificity of TLR3. This distance of 50 Å corresponds to two turns of a double-stranded RNA duplex, allowing differentiation between nucleic acids in the A- or B-type conformation as we showed.

According to the results we proposed that dsRNA molecules of longer than 45 bp, which arise from viral replication, form the most stable complex with TLR3, and consequently lead to robust interferon production. Shorter RNA duplexes of between 21 bp and 30 bp can form less stable complexes, resulting in weaker activation of the TLR3 signalling pathway.

(ii) UNC93B1 directs TLR3 towards the cell membrane

UNC93B1 is a chaperon, which was shown to deliver intracellular Toll-like receptors (TLRs) to endolysosomes. We observed that overexpression of UNC93B1 elevated the abundance of TLR3, but not TLR9 at the surface of the plasma membrane as well. In cells overexpressing UNC93B1 a hyperglycosylated form of TLR3 was found, which was located exclusively on the cell surface and was sensitive to deglycosylation. UNC93B1 also led to upregulation of intracellular TLR3 and TLR9, but not surface TLR4 protein expression by affecting retention time of associated proteins in cells as prevention of de novo synthesis with cycloheximide had no influence.

With knowledge that intracellular TLR receptors interact with UNC93B1 via their transmembrane segments and cytosolic domains, we set the study of these segments for their importance for localisation of TLR3 and TLR9. We prepared chimeric constructs, where the transmembrane segments or cytosolic domains of TLR3 and TLR9 had been exchanged. The results showed that the luminal domain of TLR3 governs UNC93B1-dependent localisation of TLR3 receptor to the plasma membrane. Additionally, physiological relevance of UNC93B1 governed TLR3 translocation to the cell surface was investigated. Stimulation of human primary cell line HUVEC with TLR3 ligand poly(I:C) elevated transcription of UNC93B1 and TLR3 and increased the amount of TLR3 on the cell surface. So, our results show that surface expressed TLR3 might enable cells to respond to released RNA thus amplifying antiviral response.

(iii) Differential activation of TLR3 by short and long dsRNA

Small interfering RNA molecules (siRNA) are short dsRNA molecules, which are widely used for silencing and also in a variety of therapeutic applications. Their off-target activity can potentially lead to unwanted toxicities. Our goal was to study one of the side effects of siRNA-based methods, namely activation of the innate immune system, which limits the use of siRNA in therapeutic applications. We showed that siRNA activated a TLR3-dependent inflammatory response on the surface of primary endothelial cells. The NF- κ B response was completely inhibited by antibodies, reactive to the N-terminal binding site of the receptor's ectodomain. On the other hand, the interferon type I response started predominantly by activating TLR3 in endosomal compartments, as in the case of poly(I:C).

Additionally, we showed that poly(I:C) activated a stable IFN- β and IL-8 response in human lymphatic endothelium cells, which was inhibited by bafilomycin A1 but not by anti-TLR3 antibodies. On the contrary, bafilomycin A1 could not inhibit activation with siRNA. According to our results we believe that the outcome of TLR3 activation depends on receptor localisation and the length of dsRNA. Activation of TLR3 on the cell surface by short dsRNA leads to a predominantly proinflammatory response, while TLR3 activation in endosomal compartments by long dsRNA mainly tends to activate type I IFN.

(iv) Mechanism of endosomal TLR inhibition by antimalarial drugs and imidazoquinolines

Endosomal Toll-like receptors (TLRs) play an important role in innate immune response as well as in autoimmune processes. In the therapy of systemic lupus erythematosus, some antimalarial drugs, such as chloroquine, hydroxychloroquine and quinacrine, have been used for a long time. Their suppression of endosomal TLR activation has been attributed to the inhibition of endosomal acidification, which is a prerequisite for the activation of these receptors. We discovered that chloroquine inhibits only activation of endosomal TLRs by nucleic acids, while it augments activation of TLR8 by a small synthetic compound R848. We detected direct binding of antimalarials to nucleic acids by spectroscopic experiments and determined their cellular colocalisation. Further analysis revealed that other nucleic acid binding compounds, such as propidium iodide also inhibited activation of endosomal TLRs and colocalised with nucleic acids to endosomes. We found that imidazoquinolines, which are TLR7/8 agonists, inhibit TLR9 and TLR3 even in the absence of TLR7 or TLR8 and their mechanism of inhibition is similar to the antimalarials. In contrast to bafilomycin, none of the nucleic acid binding compounds inhibited endosomal proteolysis, confirming that inhibition of pH acidification in the presence of antimalarials and imidazoquinolines is not the underlying cause of inhibition. We conclude that the direct binding of inhibitors to nucleic acids changes their TLR binding epitope and may explain the efficiency of those compounds in the treatment of autoimmune diseases.

Publications:

Pirher, N., Ivicak, K., Pohar, J., Bencina, M., Jerala, R. A second binding site for double-stranded RNA in TLR3 and consequences for interferon activation. 2008. *Nat. Struct. Mol. Biol.* **7**, 761-763

Kuznik, A., Bencina, M., Svajger, U., Jeras, M., Rozman, B., Jerala, R. Mechanism of endosomal TLR inhibition by antimalarial drugs and imidazoquinolines. 2011. *J. Immunol.* **186**, 4794-4804

Pirher, N., Pohar, J., Mancek-Keber, M., Jerala, R. Differential activation of TLR3 by short and long dsRNA. (in preparation)

Pohar, J., Pirher, N., Bencina, M., Mancek-Keber, M., Jerala, R. Determinants of trafficking TLR3 towards the cell membrane. (ready for submission)

SOUTH AFRICA

Title: Structural studies of enzymes in the context of anti-malarial drug development

Principal Investigator: Bryan Trevor Sewell, Electron Microscope Unit, University of Cape Town, Private Bag, Rondebosch, Cape Town, South Africa. Tel: +27-21-650-2817, Fax: +27-21-689-1528, E-mail: Trevor.Sewell@uct.ac.za

ICGEB Contract No.: CRP/08/016

ICGEB Reference No.: CRP/ZAF08-04

Abstract: It is well established that patients with acute malarial infection suffer substantial depletion of blood glutamine. This glutamine is utilised by the parasite that also has an endogenous glutamine synthetase. There are, in addition, several unusual features of the metabolic pathways concerned with nitrogen metabolism in *Plasmodium falciparum*. This applies particularly to enzymes involved in maintaining the NAD⁺/NADH redox pair where three enzymes: NAD⁺ kinase, NAD phosphorylase and nicotinamide phosphoribosyl transferase, which are part of the normal pathway are absent from the parasite genome. The amidase domain of NAD⁺ synthetase (*PfNADS*) was found to have an insertion and unusual sequence surrounding the active site cysteine which could be evidence of specialisation required by the parasite. The expression of full-length NAD⁺ synthetase was found to be toxic to *E. coli* cells and thus far the protein has not been expressed recombinantly. A method for preparing *Plasmodium falciparum* glutamine synthetase (*PfGS*) and maintaining the enzyme stably in solution at concentrations high enough for crystallisation trials was developed. The enzyme was visualised and shown to be homogeneous by both negative stain and cryo-electron microscopy. A preliminary three-dimensional reconstruction has been made from images of negatively stained protein.

Objectives:

- (i) Clone, express, purify and crystallise NAD⁺ synthetase and glutamine synthetase from *Plasmodium falciparum*;
- (ii) Determine their structures by x-ray crystallography and determine whether they interact by electron microscopy.

Results Obtained:

A method for preparing *Plasmodium falciparum* glutamine synthetase and maintaining the enzyme stably in solution at concentrations high enough for crystallisation trials was developed. The enzyme was visualised and shown to be homogeneous by both negative stain and cryo-electron microscopy. A preliminary three-dimensional reconstruction has been made from images of negatively stained protein.

Results Unforeseen in the Original Project:

The sequence of *PfGS* that was deposited in the PlasmoDB was incorrect. We experienced considerable difficulty with the recombinant expression of both *PfGS* and *PfNADS* in *E. coli*. These difficulties were overcome in the case of *PfGS* by using pCold expression plasmids and a good purification procedure was developed for this enzyme. Further difficulties which have not yet been overcome were encountered in crystallising the protein.

Networking:

The project enabled Prof. Trevor Sewell and Dr. Brandon Weber to visit the laboratory of Dr. Amit Sharma of the ICGEB New Delhi Component for the crystallisation trials and to present the outcome of the project at the MALSIG meeting. Dr. Jason van Rooyen, who was involved with this project, is continuing the collaboration with Dr Sharma.

Publications:

Patel, S.I., Weber, B., Sewell, B.T. Recombinant expression and reconstruction of *Plasmodium falciparum* glutamine synthetase. 2012. Protein Expr. Purif. (in preparation)

SRI LANKA

Title: Performance of recombinant proteins as diagnostic intermediates for Chikungunya, Dengue and Leptospirosis

Principal Investigator: Menaka Hapugoda, Molecular Medicine Unit, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka. Tel: +94-112-960483, E-mail: menakaha@yahoo.com
ICGEB Contract No.: CRP/08/017

ICGEB Reference No.: CRP/SRI08-02

Abstract: Chikungunya (CHIK), dengue and leptospirosis are important diseases with explosive outbreaks occurring in Sri Lanka. Many clinicians as well as general practitioners find it difficult to confirm diagnosis of these three infections only on clinical grounds, particularly when they exist as co-infections. Confirmation of disease outbreaks is important for clinicians for proper management of patients. Laboratory diagnosis of these three diseases in Sri Lanka is hindered by the high cost of commercial diagnostic kits and inaccessibility of reagents. Because the burden of CHIK, dengue and leptospirosis is greatest in Sri Lanka, there is a need to develop tests which can be discriminated these three diseases and produced at low cost and also easily standardised for use in field settings. Further, differential rapid, optimised diagnostic assay with high sensitivity and specificity of each disease is essential for clinical management and epidemiological studies. Laboratory diagnosis depends on Enzyme-Linked Immunosorbent Assay (ELISA) based on whole viral/bacterial antigens which cause biohazard risk, high production cost and cross reactivity with other organisms of the same genus/family. A diagnostic intermediate using a single recombinant protein antigen to detect both IgM and IgG antibodies of each disease is important to overcome problems associated with whole viral/bacterial antigen/lysate is important.

Objectives:

The overall objective of this project was to assist confirmation of CHIK, dengue and leptospirosis outbreaks through developing rapid laboratory diagnostic assays.

Methodology:

Novel recombinant protein antigens for all three diseases were prepared. Here, synthetic genes of CHIK were designed and expressed in both bacteria and yeast vector systems. Synthetic genes for dengue and leptospirosis were designed and expressed in bacteria vector systems. Resulted proteins for each disease were purified using a single affinity chromatographic step. Then, potential use of each protein as a diagnostic tool for the detection of both IgM and IgG antibodies produced against the particular disease organism was evaluated using a large panel of well characterised serum samples. Indirect ELISAs using each novel recombinant protein as a capture antigen for detection of both anti-IgM and IgG antibodies of each disease were developed using reference serum samples obtained from a reference laboratory. Specificity, and sensitivity of IgM and IgG ELISAs developed for each disease were analysed with currently available diagnostic assays including the Gold standard assay using a large panels of well characterised serum samples.

Results Obtained:

The resulted proteins which show the highest sensitivity and specificity with currently available diagnostic assay were selected. In order to synergise activities, the work was carried out jointly by the Molecular Medicine Unit, Faculty of Medicine, University of Kelaniya, Sri Lanka (i.e., main counterpart Institution) and by the Mammalian Biology Recombinant Gene Products (RGP) Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi Component. These two Institutions worked with each other through exchanging knowledge and resources (both human and technical). In particular, the main counterpart Institution worked together with clinicians, general practitioners and epidemiologist in notifying and locating cases at the regional level when they collected clinical samples from suspected patients of each disease. The Sri Lankan Institute collaborated with national and international reference centres to characterise three panels of serum samples by the Gold standard assay for each disease. At national level networking was established with: (i) clinicians from the Ministry of Health; (ii) Medical Research Institute; (iii) National Science Foundation (RG/2009/BT/01). At international level collaboration was established with: (i) Department of Virology (WHO Reference Centre for viral reference and research), Institute of Tropical Medicine, Nagasaki University, Japan; (ii) Department of Biomedical Research, WHO/FAO/OIE and National Collaborating Centre for Reference and Research on Leptospirosis, Royal Tropical Institute, Amsterdam.

During the operation period, suspected patients for each disease warded in hospital/s in outbreak area/s in Sri Lanka were benefited by having results of laboratory diagnostic assays (molecular and currently available serological assays) within 48 hours of collection of the sample.

Funds received by the main counterpart institution through the project were used to meet the cost of equipment, consumables, training and travel. The main counterpart institution has set up facilities for molecular diagnosis and preparation of recombinant proteins for each disease.

One under graduate and three post graduate students were directly benefited by the project. Project staff were benefited by obtaining local and international training through the project.

Results of the project were disseminated through a number of presentations made at national and international conferences and were part of a post graduate thesis. In the near future, dissemination of project results will be made through journal publications and other two post graduate thesis.

Resulted proteins are available at two counterpart Institutions for field use and further studies. At the end of the project target group beneficiaries will be the infected people and healthy persons in outbreak area/s of these three diseases by confirming each disease using ELISAs developed using these recombinant protein antigens. This type of study on development of diagnostic intermediates has a significant effect for rapid confirmation of outbreaks and limit spread of such outbreaks from one geographical area to another. Confirmation of disease outbreaks will avoid loss of working hours and will also have socio-economical impact on individuals and government.

Publications:

In preparation.

UNITED ARAB EMIRATES

Title: Studying chromatin-modifying proteins in transcription regulation

Principal Investigator: Ahmed H. Hassan, Department of Biochemistry, Faculty of Medicine and Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates. Tel: ++971-3 713-7478, Fax: +971-3-767-2033, E-mail: ahmedh@uaue.ac.ae

ICGEB Contract No.: CRP/08/018

ICGEB Reference No.: CRP/ARE08-01

Abstract: It is clear that the targets for signalling pathways regulating gene expression include both the DNA and protein components of chromatin. In the last decade, a number of transcriptional regulators that function by modifying the structure of chromatin have been discovered. The *Saccharomyces cerevisiae* SWI/SNF is the founding member of the chromatin-remodelling family. Phylogenetic analysis shows that chromatin-remodelling proteins share several common features including the presence of a distinct ATPase domain. Based on sequence homology to this domain of the Swi2/Snf2 subunit of the SW/SNF, we have identified some new candidate remodelling proteins including the Fun30 (Function unknown now 30) protein. With purified complexes in hand, a growing panel of *in vitro* and *in vivo* assays, a new wealth of knowledge regarding the mechanisms of actions of many chromatin-remodelling complexes, and the ease of genetic manipulations in yeast, we have carried out mechanistic biochemical and some genetic investigations of their functions and their inter-relationships. Our overall objective was to better understand the mechanisms by which the Fun30 protein is involved in chromatin remodelling and transcription regulation. Toward this end, we have carried out the biochemical characterisation of the yeast Fun30 protein and have performed a series of *in vitro* assays as well as deletion studies. We are now working on understanding the *in vivo* roles of Fun30 in gene regulation.

Objectives:

The overall objectives of this proposal were to better understand the mechanisms of action of Fun30, a newly identified chromatin remodeller in *Saccharomyces cerevisiae*, in regulating gene expression.

Results Obtained:

We have been able to purify Fun30 from the yeast and have shown that it is present as a homodimer in our purifications. We had also performed a variety of *in vitro* experiments to elucidate the mechanism of action of this chromatin remodeller. We observed that Fun30 binds to DNA and nucleosomes, and has a variety of ATP-dependent remodelling activities. Since we have shown that Fun30 can be targeted to promoters via direct interactions with acidic activation domains, a new line of research on identifying the subunits or domains within fun30 that can bind the acidic activators has now opened up and we will be conducting studies to identify the domain(s) within Fun30 that interact with transcription activation domains. In addition to these experiments, we have extended our dimer exchange studies and have shown that Fun30 exhibits specificity towards H2AZ. We have also compared the ability of Fun30 to transfer H2B and H2AZ dimers from HeLa donor chromatin to a tetrasome acceptor. We observed that the transfer of H2AZ was slightly less efficient than H2B transfer. It would be interesting to investigate the effects of the CUE motif or ATPase domain deletions on the activity of Fun30.

To explore the physiological functions of Fun30, we have also investigated the phenotypes of the *fun30* null mutant (*fun30Δ* strain). These phenotypic studies should help identify genes whose promoters are targets of the Fun30 complex, which is an important part of our proposal. First we investigated the involvement of Fun30 in the functions of metabolic pathways. The strain was propagated on a variety of different carbon sources (rafinose, lactose and galactose). We found that the *fun30Δ* strain grows on rich media with a growth rate comparable to wt strains. It has been suggested that Fun30 is involved in chromosome stability and potentially DNA repair based on the pattern of genomic instability. Thus, we investigated growth under conditions that could highlight the role of Fun30 in the maintenance of DNA structure or DNA repair. To do this, the wild-type (wt), *snf2Δ*, and *fun30Δ* strains in the same genetic background were spotted onto YPD plates after their OD600 reached 1.5, and then different treatments to induce DNA damage. We first investigated resistance/sensitivity to UV irradiation. UV causes cross-linking between adjacent cytosine and thymine bases resulting in pyrimidine dimers which consequently inhibit polymerases and arrest replication. While the wild-type and *snf2Δ* strains are very sensitive to even low doses of UV irradiation, the *fun30Δ* yeast cells confer resistance and grew well even at the highest UV exposure time interval. This supports a previous report of UV resistance in the absence of Fun30. We have also found that the cells of the *fun30 Δ* strain were more resistant to increasing amounts of radiation.

We have recently found that *fun30* null mutant cells were slightly sensitive to caffeine. This finding indicated that Fun30 may be important for the maintenance of cell viability during S phase arrest. To explore any links between Fun30 and genes involved in progression through the cell cycle and DNA replication, we have investigated the ability of *fun30* deletion to progress through the cell

cycle. For this, we have used flow cytometry following propidium iodide staining to monitor the DNA content of cultures synchronised using α -factor. We have observed a delay in the progression through the G1/S phase of the cell cycle in the *fun30 Δ* strain. We have now examined the expression of several checkpoints marker proteins.

In addition, more recently, we have been concentrating our efforts on performing ChIP-Seq experiments to identify the targets (localisation) of the Fun30 complex at the whole genome level, which should help us understand the *in vivo* role(s) of this remodeller. These experiments will help in identifying the involvement of the Fun30 protein in gene activation as well as repression.

Results Unforeseen in the Original Project:

We have faced some technical challenges over the years on this project that we have overcome and the results obtained thus far are very promising and would lead to increased understanding of the role of this novel chromatin remodeller.

URUGUAY

Title: Design of anti-tumour vaccine by glycoprotein engineering

Principal Investigator: Eduardo Osinaga, Departamento de Inmunobiología, Facultad de Medicina, Universidad de la República, Av. General Flores 2125, Montevideo, Uruguay. Tel: +598-29249562, Fax: +598-29249563, E-mail: eosinaga@fmed.edu.uy

ICGEB Contract No.: CRP/08/019

ICGEB Reference No.: CRP/URU08-03

Abstract: The project aimed at developing a new strategy for the production of glycosylated hemi-synthetic vaccines for anti-tumour immunotherapy. Our strategy was based on the directed enzymatic glycosylation of mucin-like proteins from parasitic origin, in order to produce glycoproteins containing the tumour-associated Tn antigen. Tn-based glycopeptides or glycoproteins were prepared by *in vitro* glycosylation of mucin-like peptides from parasite origin, either by using several recombinant ppGalNAc-Ts, or by solid-phase synthesis. The obtained data during the performance of this project suggest that extensive glycosylation of peptides modifies the immunological properties, such as the type and intensity of T cell responses, as well as the antigen capture by dendritic cells and their maturation. Also, we found that parasite-derived peptides could induce a cytotoxicity capacity by *in vivo*-primed splenocytes that might be mediated by dendritic cells. This work has provided new insights on the rationale of the design of Tn-based vaccine candidates against cancer. Moreover, considering that the Tn antigen has also been identified in numerous helminth parasites, a Tn vaccine produced from a parasite protein might be useful in the induction of protective immune responses against these microorganisms.

Objectives:

(i) To synthesise different O-glycosylated isoforms of the C317 parasitary protein bearing the Tn antigen using recombinant ppGalNAc-Ts.

(i.a) C317 recombinant protein production in *Escherichia coli*;

(i.b) *In vitro* GalNAc transfer to C317 proteins and physico-chemical characterisation of the resulting glycoconjugates;

(i.c) Antigenicity of C317-Tn glycoproteins.

(ii) To evaluate the anti-tumour properties of the immune responses induced by the C317-Tn glycoproteins in animal models (TA3/Ha transplantable tumour in mice and NMU-induced breast cancer in rats).

Results Obtained:

(i) Production of parasite-derived glycopeptides carrying the Tn antigen for tumour immunotherapy

This project is based on an innovative approach for the production of glycosylated hemi-synthetic vaccines, combining recent discovered mucin-like proteins from helminth parasites and an enzymatic glycosylation strategy. Three different peptides from helminth-derived mucin-like proteins were subjected to glycosylation: Egmuc-I and Egmuc-II (derived from *Echinococcus granulosus*) and Fhmuc (derived from *Fasciola hepatica*). In order to obtain peptides carrying the Tn antigen, we used the glycosyltransferases (ppGalNAc-T) that transfer GalNAc residues *in vivo* to a Thr or Ser in a protein. ppGalNAc-T2, -T3, and -T6 enzymes have been produced in insect cells infected with recombinant baculovirus. Fhmuc and Egmuc-II were efficiently glycosylated by ppGalNAc-Ts. The presence of the Tn antigen on these glycopeptides was confirmed by Western blotting. We next carried out a semi-preparative scale glycosylation of Fhmuc and EgmucII (500 µg). Additionally, we carried out the identification of lymphocyte T CD4 immunodominant sequences by epitope mapping by immunising mice with the Egmuc peptides. The peptide designed EM5 was the only one capable of stimulating T cell proliferation, indicating that this sequence contains a T CD4 epitope.

(ii) Evaluation of the immunological properties of parasite-derived glycopeptides carrying the Tn antigen

Tn-EgmucII and Tn-Fhmuc were highly recognised by the anti-Tn mAb 83D4. In contrast, the 15-mer di-glycosylated Egmuc₃₃₋₄₇-2Tn peptide was not recognised by 83D4. The immunogenicity of peptides and glycopeptides was evaluated by their ability to produce antibodies in mice and to induce specific T cell proliferation of *in vivo*-primed lymph node cells. We were not able to detect by ELISA, IgG or IgM antibodies in sera from mice immunised with the glycosylated peptide Egmuc₃₃₋₄₇-2Tn against none of the evaluated (glyco)peptides. Additionally, Egmuc *in vivo* primed-splenocytes cross-react with Egmuc₃₃₋₄₇, while recognition is abolished when incorporating GalNAc residues, suggesting that glycosylation alters T cell epitope recognition. In contrast, Fhmuc-Tn induced much higher levels of cell proliferation and IFN γ than non glycosylated Fhmuc, suggesting that, for this peptide backbone, the incorporation of GalNAc residues confers more cellular immunogenicity to the peptide. These results indicate that glycosylation of T or B epitopes may alter their immunogenicity, clearly depending on the peptide backbone, as well as the number of GalNAc incorporated.

We analysed the capacity of bone marrow derived dendritic cells (BMDC) to internalise the Egmuc₃₃₋₄₇ peptide and glycopeptide. A strong capture and endocytosis by BMDC was observed for both the non-glycosylated Egmuc and Egmuc₃₃₋₄₇ peptides and the Egmuc₃₃₋₄₇-2Tn glycopeptide, indicating that the presence of GalNAc in the peptide did not modify internalisation by BMDC. The internalisation was not modified in the presence of EDTA, suggesting that MGL (Macrophage Gal/GalNAc lectin) does not intervene in the capture or internalisation of the Egmuc₃₃₋₄₇-2Tn glycopeptide, which needs several carbohydrate residues to allow cross-linking. In order to localise the internalised peptides, the cells were analysed by confocal microscopy. In contrast with results obtained with Egmuc₃₃₋₄₇-2Tn, that was internalised at the same rate as Egmuc₃₃₋₄₇, glycosylated Fhmuc-Tn presented an enhanced uptake by DCs, as compared to Fhmuc, suggesting that, for Fhmuc peptide backbone, the incorporation of GalNAc residues potentiates antigen internalisation. We also analysed the presence of IL6, IL10 and IL12 in dendritic cell culture supernatants. The studied peptides modify the expression of co-stimulatory molecules on dendritic cells and induce the production of pro-inflammatory cytokines. However, this effect on dendritic cell maturation would not be dependent on the Tn antigen on the Egmuc₃₃₋₄₇ peptide backbone.

(iii) Anti-tumour immune response induced by parasite-derived glycopeptides carrying the Tn antigen

Antibodies induced by any of Egmuc peptides and glycopeptides did not recognise antigens derived from lysates of TA3/Ha or Panc02 tumour cell lines. Egmuc₃₃₋₄₇ and Egmuc₃₃₋₄₇-2Tn primed splenocytes proliferated when stimulated with TA3/Ha or Panc02 protein lysates and produced modest levels of INF γ and IL-5. These *in vivo* primed-splenocytes were also incubated with tumour cells, in order to evaluate their cytotoxic ability. Splenocytes from mice immunised with Egmuc₃₃₋₄₇ or Egmuc-2Tn peptides were capable of killing both tumour cells, while spleen cells from mice immunised with Egmuc were not. The fact that both spleen cells from mice immunised with Egmuc₃₃₋₄₇ or Egmuc₃₃₋₄₇-2Tn were capable of killing tumour cells, suggests that the recognition is not Tn-dependent. One possibility could be that the Egmuc₃₃₋₄₇ peptide and glycopeptide are able to activate dendritic cells, inducing a release of certain activating cytokines (such as IL12 that activates natural killer cells) or inducing the over-expression of certain co-stimulatory molecules that favors the development of a cytotoxic response.

(iv) *In vivo* studies in experimental cancer models

Taking into account the *in vitro* capacity of Egmuc₃₃₋₄₇ primed-splenocytes to mediate tumour cell killing, we next intended to carry out some *in vivo* experiments in order to evaluate whether Egmuc₃₃₋₄₇ peptide or glycopeptide were able to protect mice from tumour growth. In a preliminary evaluation, we were not able to protect from tumour growth Egmuc-immunised mice. We are currently working on different mice experimental cancer models.

Results Unforeseen in the Original Project:

We characterised the Fhmuc as a good peptide substrate in order to produce glycoproteins containing the tumour-associated Tn antigen, using several recombinant ppGalNAc-Ts. We confirm that the incorporation of several residues of GalNAc in long peptides of Fhmuc considerably changes its immunological properties. We found that Fhmuc-Tn was much immunogenic than non-glycosylated Fhmuc, being capable of inducing *in vitro* proliferation of splenocytes that produced more than 10 times more of IFN γ than non-glycosylated Fhmuc. Additionally, we observed that Fhmuc-Tn is more efficiently internalised by dendritic cells. We are currently working on this aspect.

Publications:

We are in the process of preparing a manuscript that will be submitted soon. This paper will contain the data regarding the anti-tumour immune response induced by Egmuc peptides. On the other hand, we will very soon start the preparation of another paper, which still requires some experimental data. This paper will include the glycosylation and study of the immunogenicity of Fhmuc peptide and glycopeptides, stressing the fact that cellular immunogenicity is enhanced when attaching GalNAc residues on this peptide backbone.