

**NANOGENOMICS AND NANOPROTEOMICS ENABLING PERSONALIZED,
PREDICTIVE AND PREVENTIVE MEDICINE**

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

dem Fachbereich Chemie der Philipps-Universität Marburg

vorgelegt von

MD Nicola Luigi Bragazzi

aus Carrara

Marburg/Lahn 2014

Erstgutachter: - Prof. Dr. Norbert Hampp

Zweitgutachter: - Prof. Dr. Claudio Nicolini

Prüfungstermin: 16.01.2014

Hochschulkenziffer: 1180

Vom Fachbereich Chemie

der Philipps-Universität Marburg (Hochschulkennziffer: 1180)

als Dissertation am ____.____.201_ angenommen am:

Tag der mündlichen Prüfung: _____

Erstgutachter: Prof. Dr. N. Hampp

Zweitgutachter: Prof. Dr. C. Nicolini

Die vorliegende Dissertation wurde unter Betreuung von Herrn Prof. Dr. Norbert Hampp (Fachbereich Chemie, Philipps-Universität Marburg) in der Zeit von Juli 2011 bis Dezember 2013 an den Laboratories of Biophysics and Nanobiotechnologies, Department of Experimental Medicine (DIMES), Faculty of Medicine, University of Genoa, Via Antonio Pastore 3, Genoa (Italy) angefertigt.

E r k l ä r u n g

gemäß § 10, Abs. 1 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fachbereiche und des Medizinischen Fachbereichs für seine mathematisch-naturwissenschaftlichen Fächer der Philipps-Universität Marburg vom 28.4.2010

Ich versichere, dass ich meine Dissertation „**Nanogenomics And Nanoproteomics Enabling Personalized, Predictive And Preventive Medicine**“ selbstständig, ohne unerlaubte Hilfe angefertigt und mich keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient, sowie alle vollständig oder sinngemäß übernommenen Zitate als solche gekennzeichnet habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 4.1.2014

Nicola Luigi Bragazzi

*“Not only is the Universe stranger than we think,
it is stranger than we can think.”
Werner Heisenberg, Across the Frontiers*

TABLE OF CONTENTS

1. Acknowledgments7

2. Abstract8

 2.1 Abstract in German8

 2.2 Abstract in English10

3. Nanomedicine: definition and main goals12

4. Nanogenomics21

 4.1. DNASER26

 4.2. Leader Gene Algorithm29

5. Nanoproteomics34

 5.1 Structural Proteomics and Protein Nanocrystallography38

 5.2 Functional nanoproteomics66

 5.2.1. NAPPA66

 5.2.2. Nanoconductometric QCM_D.....69

 5.2.3. APA73

 5.2.4. Mass Spectrometry.....76

6. Medical Applications81

 6.1. A state-of-art of the literature81

 6.2. Bioinformatics *ab initio* medical applications85

 6.2.1. Applications in the field of Nanonephrology85

 6.2.2. Applications in the field of Oral Pathology104

Contents

6.3. NAPPA QCM_D based applications	123
7. Conclusions	151
8. Appendixes	153
8.1. Appendix I	153
9. Bibliography	162
10. <i>Curriculum Vitae et Studiorum</i> of the PhD candidate	202
11. List of publications	203

Acknowledgments

1. ACKNOWLEDGMENTS

First of all, I would like to acknowledge my PhD Thesis advisors Professor Claudio Nicolini who gave me the chance to work in his team, trusted in my skills, made me love the fascinating world of the nanobiotechnologies and supported me during the difficulties and the hard time, and Professor Norbert Hampp, who accepted me in his team, trusting too in my attitudes. I want to thank him particularly for his patience, suggestions and tips that have improved a lot my PhD thesis and presentation. His Arbeitsgruppe (AG) in Marburg represents a unique and incredibly stimulating environment for making science and research.

I would like to acknowledge also Professor Eugenia Pechkova that guided me through every project giving me precious suggestions on how to improve my scientific skills and made me know and love protein crystallography.

Finally, I would like to acknowledge every single co-author of the projects in which I was involved.

Dr. Nicola Luigi Bragazzi,
Genoa-Marburg 2011-2014

2. ABSTRACT

2.1 ZUSAMMENFASSUNG IN DEUTSCH

Seit der Entdeckung von Nucleic Acid hat die Molekularbiologie enorme Fortschritte gemacht und vieles erreicht. Trotz all dem, besteht immer noch eine Lücke zwischen dem klassisch-traditionellen medizinischen Ansatz, und der Molekularen Welt.

Inspiriert durch die riesigen Datenmengen, die durch die "*omics*"-geführten Techniken und die "high-throughput" Techniken (*HTTs*) entstanden sind, habe ich versucht ein Protokoll zu entwickeln, welches die Hürde zwischen der phenomenological Medizin und der Molekularmedizin überbrücken soll, und somit das Labor dem Patientenbett etwas näher bringen wird. Ausserdem finde ich, dass es dringend nötig ist die wichtigsten "*omics*" Wissenschaften - Genomics und Proteomics - miteinander zu integrieren. Nucleic Acid Programmable Protein Arrays (NAPPA) können dies erreichen, durch die Verwendung eines "*complex mammalian cell free expression*" Systems um die Proteine *in situ* zu erzeugen. Als Alternative zu "*fluorescent-labelled*" Ansätzen, kann eine neue kennzeichnungsfreie Methode, die durch den kombinierten Einsatz von drei unabhängigen und komplementären nanobiotechnologischen Ansätzen entsteht, verwendet werden.

Diese Methode mag fähig sein das Zusammenspiel von Gen- und Proteinfunktionen, Gen-Proteine, Gen-Medikamente, Protein-Proteine und Protein-Medikamente zu analysieren. Dies wäre sehr Vorteilhaft für die personalisierte Medizin. Quartz Micro Circuit Nanogravimetry (QCM),

Abstracts

basierend auf Frequenz- und Zerstreuungsfaktoren, Massenspektrometrie (MS) und Anodic Porous Alumina (APA) überwinden fürwahr die Grenzen von "*correlated fluorescence detection*", welches darunter leidet, dass es trotz gründlichem Waschen im Hintergrund weiterhin detektiert werden kann. Ausserdem wird momentan daran gearbeitet ein "*homogeneous and well defined bacterial cell free expression system*" weiter zu optimieren, welches die ambitionierten Ziele der Quantifizierung der bestimmenden Gene und Protein Netzwerke im menschlichen Körper verwirklichen kann. Implikationen für die personalisierte Medizin der oben erwähnten kennzeichnungsfreien Protein Arrays mit dem Einsatz verschiedener Test-Gene und Proteine, werden in dieser PhD-Arbeit besprochen.

2.2 ABSTRACT IN ENGLISH

Since the discovery of the nucleic acid, molecular biology has made tremendous progresses, achieving a lot of results.

Despite this, there is still a gap between the classical and traditional medical approach and the molecular world. Inspired by the incredible wealth of data generated by the "omics"-driven techniques and the “*high-throughput technologies*” (HTTs), I have tried to develop a protocol that could reduce the actually extant barrier between the phenomenological medicine and the molecular medicine, facilitating a translational shift from the lab to the patient bedside. I also felt the urgent need to integrate the most important omics sciences, that is to say genomics and proteomics.

Nucleic Acid Programmable Protein Arrays (NAPPA) can do this, by utilizing a complex mammalian cell free expression system to produce proteins *in situ*. In alternative to fluorescent-labeled approaches a new label free method, emerging from the combined utilization of three independent and complementary nanobiotechnological approaches, appears capable to analyze gene and protein function, gene-protein, gene-drug, protein-protein and protein-drug interactions in studies promising for personalized medicine. Quartz Micro Circuit nanogravimetry (QCM), based on frequency and dissipation factor, mass spectrometry (MS) and anodic porous alumina (APA) overcomes indeed the limits of correlated fluorescence detection plagued by the background still present after extensive washes. Work is in progress to further optimize this approach a homogeneous and well defined bacterial cell free expression system able to realize the ambitious objective to quantify the regulatory gene and protein networks in humans.

Abstracts

Abstract

Implications for personalized medicine of the above label free protein array using different test genes and proteins are reported in this PhD thesis.

3. NANOMEDICINE: DEFINITION AND MAIN GOALS

“Personalized nanomedicine” is *“broadly defined as the approach of science and engineering at the nanometer scale”* toward individualized, *“biomedical applications”* [Lee et al., 2009]. In the next decades, it is expected to be a promising and formidable tool for a targeted and more rationale therapeutics, that really fits individual patient's needs, emerging from the combination of *“personalized medicine”* and *“nanomedicine”*.

“Personalized medicine” can be defined as *“as the prescription of specific treatments best suited for an individual, taking into consideration both genetic and other factors that influence the response to therapy”* [Jain, 2008].

“Nanomedicine” can be defined as *“the interplay of Bioinformatics and Nanotechnology to a previously unforeseeable level”* [Nicolini, 2006; Nicolini, 2010; Nicolini and Pechkova, 2010a; Nicolini et al., 2012b].

The integration of information obtained from both *“ex vivo”* and *“in vivo”* techniques [Lammers et al., 2012] enables a precise and rich picture of the disease, which is completely unprecedented.

The size of the functional elements of the biological systems is the fundamental nanometer scale, from whence the term *“nanomedicine”*, being the level at which nanomaterials interact with molecules. These nanomaterials offer many advantages with respect to the classical and traditional materials, such as the small size, larger volume-to-surface ratio, and can be finely tuned and manipulated down to the scale of a cluster of atoms, in order to have specifically designed properties and unique features. The main objectives of personalized nanomedicine are: diagnosis (chip-based bioassays, molecular imaging), therapy (drug delivery and drug

Chapter 3. Nanomedicine: definition and main goals

release), and prognosis (biomarkers, molecular monitoring and risk assessment).

These phases are not necessarily separate, but can be harmonized into a coherent framework, and for example diagnosis and therapy can be integrated into a unique step, termed as “*theranostics*” (which is a portmanteau of “*therapy*” and “*diagnostics*”) [Özdemir and Cho, 2012; Ozdemir and Dove, 2012].

It should be emphasized that, before, therapeutics was based upon and followed the diagnosis (for a certain kind of disease, a drug or a treatment), while nowadays, considering the fact that the same drug can exert different effects on different patients, the relationship between therapeutics and diagnosis becomes more complex and mutual, rather than univocal. Another ambitious goal is prevention, which can be attained diagnosing the disease at an earlier stage or delivering chemopreventive drugs in an effective way.

High-throughput technologies are able to produce a wealth of data and information, that act as a link between the molecular level (molecular medicine) and the classical and traditional medical approach (the clinical and phenomenological medicine).

Thus, personalized nanomedicine differs from molecular medicine, in that it exploits a systems biology approach and a broader, holistic method, in order to capture the networks and the pathways which lie behind a disease.

For this reason, personalized nanomedicine is called also “*network medicine*” [Barabási, 2007; Barabási et al., 2011; Chen and Butte, 2013; Chen et al., 2013; Piro, 2012; Roukos, 2013; Silverman and Loscalzo, 2012;

Chapter 3. Nanomedicine: definition and main goals

Silverman and Loscalzo, 2013; Smith et al., 2013] or “*systems medicine*” [Emmert-Streib and Dehmer, 2013; Jain, 2008].

While molecular medicine focuses and relies on point and discrete pieces of information (thus providing static information), omics science can capture the activity and expression variations.

It differs of course also from the traditional approach, which is one-size-fits-all, while nanomedicine can be predictive, personalized and preventive (the so-called “*P4 model*”, introduced by Leroy Hood) [Auffray et al., 2010; Hood, 2013; Hood and Friend, 2011].

Within this new framework, data integration plays a major role, since integrated genomics–proteomics data are useful for finding disease biomarkers [Boja et al., 2011; Faratian et al., 2009; Hanash and Taguchi, 2011] and for providing a more precise diagnosis and prognosis in the frame of translational medicine, shifting from the lab to the patient bedside: they can give an added value to the clinical assessment of the patient, as well as they can even predict the risk of a not yet clinically observable disease.

A very recent and complex publication [Allen et al., 2012] has advocated a more sophisticated approach and shown that iPOP (integrative personal Omics profile) resulting from the combination of different Omics-based data (namely, genomics, proteomics, transcriptomics, metabolomics or metabonomics, and autoantibodyomics), extensively collected and covered for a period of 14 months, can be a promising approach to manage proactively individual health: dynamic information of how genes interact with environment and change their expression throughout the time could be extracted and correlated with clinical implications. In the specific case type-2 diabetes was predicted much earlier than its clinical onset: thus medicine

Chapter 3. Nanomedicine: definition and main goals

can become quantitative, predictive and preventative [Marko-Varga et al., 2007]. However, the traditional anamnesis and record of the medical history are not negligible.

The two approaches are in fact complementary, rather than exclusive, and the interpretation of omics-based data without properly incorporating them within a clinical assessment can be misleading.

Personalized nanomedicine “*has the power of combining nanomedicine with clinical and molecular biomarkers ("OMICS" data) achieving an improved prognosis and disease management as well as individualized drug selection and dosage profiling to ensure maximal efficacy and safety*” [Rosenblum and Peer, 2013].

Unfortunately, at the moment, further integration among the different omics-based and enabled disciplines is still under development since a complete and coherent taxonomy of all the omics sciences is missing (from an ontological and semantic point of view), and the concept of metadata for omics has been only recently introduced [Ozdemir et al., 2014].

Diseases such as cancer, obesity and diabetes, transplanted organ rejection and hypertension are complex biological phenomena, being sometimes highly heterogeneous among patients despite similar clinical features. In the last decades, patients can be affected by many chronic diseases [Lee et al., 2008; Neale and Kendler, 1995] throwing classical therapeutic approaches into crisis. “*Classical medicine*” fails to achieve its goal to cure individuals, and also molecular approach-based medicine risks to be unsuccessful if it does not take into account individual variations of the same disease [Pearson and Fabbro, 2004].

Chapter 3. Nanomedicine: definition and main goals

A network approach for molecular medicine is thus emerging [Faratian et al., 2009; Janjić and Pržulj, 2012; Loscalzo, 2007; Loscalzo et al., 2007; Piro, 2012] in order to find precisely the target for an optimized drug therapy. An alternative strategy for cancer biomarker validation and cancer treatment based on the decoding of the serum proteins patterns is indeed emerging and the technical advancement in measurements and the information derived from system biology should permit to develop tailored treatment of personalized medicine in the future. Moreover, new developments in proteomics tools emerging from the combination of both microfluidic chips (like the microfluidic integrated approach for identifying protein isoforms, [Hughes et al., 2012]) and an ad hoc designed protein-capture agents (like the SOMAmers in the SOMAscan technology, [Gold et al., 2012]) can make the dream of a comprehensive proteomics become true. Inspired by these recent technical and technological advancements, in *Figure 1* we introduce a protocol starting from the collection of patients' samples (saliva, blood, urine, biopsies or others) and from the analysis of the differentially expressed genes (DEGs) using both statistical and not-statistical analyses with the Leader Genes algorithm [Sivozhelezov et al., 2006] and our *in-house* microarray scanner termed DNASER [Nicolini et al., 2006] approaches, that we introduced as tools for molecular genomics enabling the selection of few important highly interconnected genes. These *ab initio* targeted microarray-based bioinformatics and nanogenomics [Nicolini, 2006] lead to a panel of few identified genes that can be expressed in a cell-free environment thanks to Nucleic Acid Programmable Protein Array (NAPPA) technologies [LaBaer et Ramachandran, 2005; Nicolini et al., 2012a; Nicolini et al., 2012b; Ramachandran et al., 2008]

Chapter 3. Nanomedicine: definition and main goals

which utilizes a complex mammalian cell free expression system to produce proteins *in situ*. In alternative to fluorescently-labeled approaches, a wide range of nanotechnological approaches have been recently introduced to characterize NAPPA microarrays. Label-free NAPPA technology, in combination with drug nanodelivery [Garibaldi et al., 2006] and protein nanobiocrystallography [Pechkova and Nicolini, 2004; Pechkova and Riekel, 2011] and its possible ongoing developments using anodic porous alumina (APA) and bacterial cell-free expression system, appear to form a single approach, named “*nanoproteomics*” [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b], capable of effectively solving the numerous problems still present in medical diagnosis and therapy.

The emerging label free NAPPA method [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b] consists of the combined utilization of Quartz Crystal Microbalance based on the D-factor monitoring (QCM_D) nanogravimetry, mass spectrometry (MS) and APA as the last step of the protocol for personalized medicine (*Figure 1*) being optimal in analyzing protein function and protein–protein interaction. Target protein microarrays provide a powerful tool for high-throughput technology-based (*HTT*) analysis of protein–protein or small molecules–protein interactions. “*Common protein microarrays are generated by separately spotting purified protein on the arrays with a linkage chemistry*” [Spera et al., 2013].

“However, due to the complicated work associated with HTT protein production and purification, and the concerns regarding protein stability during storage before and after being spotted on the array, a significant improvement to the target protein micro-arrays is represented by NAPPA

Chapter 3. Nanomedicine: definition and main goals

which allows for functional proteins to be synthesized in situ directly from printed complementary DNAs (cDNAs) just in time for assay” [LaBaer and Ramachandran, 2005; Nicolini et al., 2012a; Nicolini et al., 2012b; Ramachandran et al., 2008].

“Printing cDNA, rather than proteins, eliminates the need to express and purify proteins separately and produces proteins “just-in-time” for the assay, abrogating concerns about protein stability during storage” [Andersson et al., 2011; Ramachandran et al., 2004; Ramachandran et al., 2005]. Label free nanotechnologies, previously used to study molecular interaction and adsorption of different biomolecules, are here applied to functionalized NAPPAs surfaces [Garibaldi et al., 2006; Nicolini and LaBaer, 2010; Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b; Pechkova and Nicolini, 2004; Pechkova and Riekel, 2011] to enable the detection of proteins and their interactions without the need for labeling reagents with tags.

Chapter 3. Nanomedicine: definition and main goals

The detection of low-abundant biomarkers from biological fluids (e.g., blood, urine, saliva, or biopsies) requires HTT detection techniques and in this sense nanoproteomics provides the robust analytical platform for personalized medicine, as needed for real-time and label-free sensitive detection of low-abundant proteins [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b]. Usually, in traditional labeling methods used in both nanogenomics and nanoproteomics a fluorescent molecule is bound to the query protein by a linking molecule, mainly an antibody; the labeled molecule is then detected via a fluorescence microscope, flow cytometer or some other fluorescence reading instrument such as DNASER [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b; Nicolini et al., 2002; Nicolini et al., 2006; Troitsky et al., 2002].

“This kind of protein array fluorescence analysis presents limitations particularly in the context of measuring naturally occurring ligands, such as in clinical studies where it is not possible to produce fusion proteins: namely the need to obtain a capture antibody for each analyzed protein, and the concern that a labeling molecule may alter the properties of the query protein. The label free methods here utilized for the analysis of NAPPA micro-array then represent for personalized medicine a major advancement” [Nicolini et al., 2013b].

Chapter 3. Nanomedicine: definition and main goals

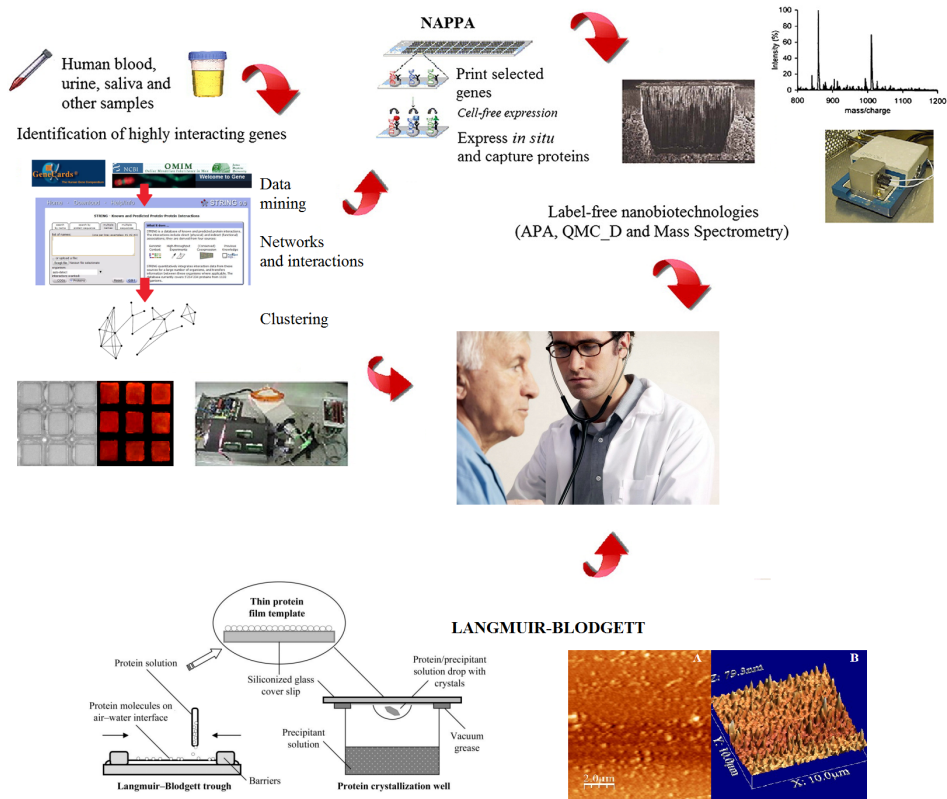


Figure 1: the experimental approach developed for targeting issues in the field of personalized nanomedicine up to the level of a single physician/patient relationship (modified from [Nicolini, Bragazzi and Pechkova. Nanoproteomics enabling personalized nanomedicine. Advanced Drug Delivery Reviews, 2012])

4. NANOGENOMICS

Nanogenomics can be defined as a discipline combining nanobiotechnologies and genomics, that is to say the use of the former techniques in order to get insights about gene regulation [Nicolini, 2006; Nicolini, 2010; Nicolini and Pechkova, 2010].

Nanogenomics is emerging as a promising discipline in the field of biomedicine, being the interplay of nanobiotechnologies and bioinformatics. Gene microarrays and the other biomolecular high-throughput tools of the omics sciences (such as linkage and linkage disequilibrium, GWAS or genome-wide association study, QTL or quantitative trait loci and its derivative, such as eQTL, expression QTL, phenome profile, transcriptome analysis for genome mapping, and so on) [Jiang et al., 2012], can produce an incredible wealth of data.

A microarray is a high-throughput nanobiotechnology developed at Stanford in the nineties, which enables the simultaneous measurement of a vast number of genes or the genotyping of some genomic regions, and exploits the reverse hybridization process. It has as its precursor the Northern blot (developed in the seventies by James Alwine, David Kemp, and George Stark at Stanford University). It is a solid-phase assay and as materials, glass, nylon, plastic or silicon chips can be used.

If on one hand, these “*big data*” [Marx, 2013; Vivar et al., 2013] are necessary for a deep and profound underpinning of the entire molecular picture, which is extremely complex, on the other hand, researchers risk to be overwhelmed by these data, without a proper data reduction [Bragazzi and Nicolini, 2013].

For this reason, gene prioritization- or gene-ranking-based approaches have

been advocated as a fundamental task in bioinformatics, resulting in a list of few selected strong candidate genes, that can be further validated with *ad hoc* targeted experiments. Underpinning genetic/genomics components and traits is particularly demanding and challenging, but also useful when studying complex and multi-factorial diseases [Cohen and Hobbs, 2013; van Nas et al., 2013], such as Alzheimer's dementia and other neurological pathologies [Hallock and Thomas, 2012], diabetes and other metabolic disorders [Jain et al., 2013], cancer [Liu et al., 2012], neurological and psychiatric diseases [Bragazzi, 2013], and cardiovascular disorders [King et al., 2005]. Usually these diseases are chronic, present a relatively mild phenotype, are slowly progressive, and have a tremendous burden for the society and impact on the population in term of quality of life (QOL).

The physio-pathology of complex heterogeneous pathologies is characterized by various biologic pathways and networks, dependent upon the contribution of a large number of genes and gene products, and therefore, the knowledge of molecular mechanisms of complex multi-factorial diseases must include and deal with a large array of genes. These genes form complex networks of interactions, which may be direct (that is to say, physical interactions between the proteins, confirmed by experimental techniques, such as NMR or crystallography), or indirect (involvement in the same metabolic pathway or co-expression under different conditions).

There are a lot of gene-ranking methods, which for example, exploit the Pearson correlation with clinical information (such as the Kaplan-Meier survival time) ([Jiang et al., 2012] and references therein; [Goh and Kasabov, 2005; Muthuswami et al., 2013]), or are based upon gene

Chapter 4. Nanogenomics

expression time series or gene co-expression series [Odibat and Reddy, 2012], as derived from genome-wide association studies (GWAS), microarrays, quantitative trait loci (QTL)/expression QTL (eQTL) analysis, single nucleotide polymorphisms (SNPs) and other mutations or from biomedical literature mining (*ab initio* approach), regulatory networks (protein-protein interaction, PPI [Gonçalves et al., 2012; Guney and Oliva, 2012; Re et al., 2012; Wang et al., 2013; Zheng and Zhao, 2012], gene-protein interaction, or gene-gene interaction maps [Vi and Ningron, 2012]) (for further explanations and references, the reader is referred [Re and Valentini, 2012; Winter et al., 2012], ortholog mapping, or integrating and combining the above mentioned approaches [Chen et al., 2013, with the proposal of the BRIDGE algorithm; Jiang et al., 2012].

The techniques that are usually employed are: support vector machine (SVM) [Keerthikumar et al., 2009; Li et al., 2013], neural networks, fuzzy algorithms, kernelized score functions [Shin et al., 2012], and other learning machine procedures, statistical approaches such as meta-analysis, biomedical annotation [Feng et al., 2012] (such as LSI, Latent Semantic Indexing [Roy et al., 2011], PosMed or Positional PubMed [Makita et al., 2013], GeneSniffer and SUSPECTS) [Thornblad et al., 2007; Yoshida et al., 2009], and clustering [Bragazzi and Nicolini, 2013; Bragazzi et al., 2011; Jay et al., 2012]. Moreover, gene prioritizing approaches can be integrated with more classical, statistics-based techniques [Jiang et al., 2014].

In the following chapters, we will explore selected clinically relevant examples of nanogenomics for personalized nanomedicine.

Molecular gene signature of operationally tolerant patients from kidney allograft has been discovered via gene microarray and DEGs analysis. In

collaboration with Nantes University, France we used a previously described algorithm named “Leader-Gene” to study the interaction networks and find the most important genes [Bragazzi et al., 2011] taking into account both interconnections and fold changes in expression, factors that we named statistical and biological significance parameters. Operational tolerance was extensively studied at a cellular level using RNA interfering silencing technology [Braud et al., 2008; Racapé et al., 2012; Sivozhelezov et al., 2008], with the goal being to study and dissect the molecular basis of the tolerance which would enable patients to avoid taking immunosuppressive drugs.

As already described in literature we found that differential genes tend to form a highly compact and dense core network. We also found that first-ranked genes tend to be connected forming forward loops [Mangan et al., 2003] and other network modules that are significantly over-represented in respect to randomly-generated networks [Brazhnik et al., 2002].

Integrated genomics–proteomics data was also the idea inspiring a research devoted to the identification of those genes responsible of the malignant transformation of T-lymphocytes [Bragazzi et al., 2011; Braud et al., 2008; Brazhnik et al., 2002; Mangan et al., 2003; Racapé et al., 2012; Sivozhelezov et al., 2008; Sivozhelezov et al., 2010]. Some genes were found promoting this biological process, some preserving the normal lymphocyte functions, and others being neutral. The interaction map appeared to be dominated by kinases and cytokines, and above all by LCK, playing a key role in two regulatory complexes. The interplay of genomics, proteomics, transcriptomics, and metabolomics is emerging as a powerful tool to analyze human disorders and providing the patient a personalized

Chapter 4. Nanogenomics

cure [Jain, 2008; Jain, 2009; Jain, 2010; Nicolini et al., 2012]. In 2007, Barabasi introduced the concept of “diseasome” in the frame of the network medicine [Barabasi, 2007].

In all of the above the goal is to investigate the medical properties of this supernetwork comparing the theoretical results with real electronic patient registers (EPR) data [Jain, 2008].

4.1. DNASER

DNASER, which is an acronym staying for DNA analySER, is a newly conceived microarray scanner and reader [Nicolini and Pechkova, 2010; Nicolini et al., 2002; Nicolini et al., 2006; Nicolini et al., 2012; Sivozhelezov et al., 2010; Troitsky et al., 2002].

A white light beam produced by a Xenon lamp emitting a light source of 150 W, accurately filtered and focused, illuminates the target sample enabling the acquisitions of images of high sensitivity. DNASER makes use of a wide-band charge-coupled device (CCD) camera (ORCA II Dual Scan Cooled 12–14 bit B/W, Hamamatsu), a progressive scan interline designed to eliminate the need for a mechanical shutter. Software routines that implement the image processing algorithms for the spot analysis are written in C language [Sivozhelezov et al., 2010].

This special high-performance set-up combines high-resolution (an effective 1280×1024 (1.3 M pixel) pixel array), low noise, high quantum efficiency, as well as a wide spectral response (300–700 nm) and a high dynamic range (full well capacity of 16,000 electrons and of 40,000 with binning).

Microarrays are grids consisting of built spots laying equidistantly and while the orientation error of the grid is negligible, its spatial position can significantly vary.

Furthermore, depending on the specific arrays used, since the distance among the different wells is different, the geometry is variable and this should taken into account while analyzing the micro-array experiment. In order to develop a robust system, the image processing algorithm was designed to perform the spot analysis in a fully automated way, without any

additional information about the DNA micro-array geometry [Nicolini et al., 2002; Sivozhelezov et al., 2010].

This device permits to acquire wider images, in a faster way (in a single shot) and of higher quality than the traditional systems, even using samples with very low fluorescence intensity.

The input data to the spot analysis can be DNA micro-array images, as well as other kind of micro-arrays. The output data of this processing are all the spot features:

- Brightness features (chip spots intensities);
- geometric features (spatial distribution, area, circularity and other parameters) (Nicolini et al., 2002).

DNASER makes use also of two interchangeable interference filters, which being situated between two achromatic doublets, can work in parallel illumination conditions. These two filters are different because are tuned respectively on the excitation and emission frequency of the DNA micro-array fluorescent spots.

This geometry is used both for illumination of the sample and collection of the fluorescence signal. The light of desired and selected spectral properties is then focused onto the sample at 45 degrees of incidence. *Figure 2* shows the main hardware blocks of the instrumentation. The illumination system, the optical filter system and the motorized holder blocks constitute the opto-mechanical components.

Three motorized linear translation stages are used to fine position the sample on the focal plane (along the z axis) of the CCD camera and to select the appropriate portion of area to be imaged (along the x, y axis). An *ad hoc*

Chapter 4. Nanogenomics

software has been developed in order to position the sample with micrometer precision.

The acquired images can be analyzed using different image processing software, such as GenePix Software. The image was acquired above an intensity threshold, considered significant. The identification of the spots is usually realized by comparing the image with a grid, which contains every information needed (e.g., coordinates and gene names) to recognize the spots (*Figure 2*).

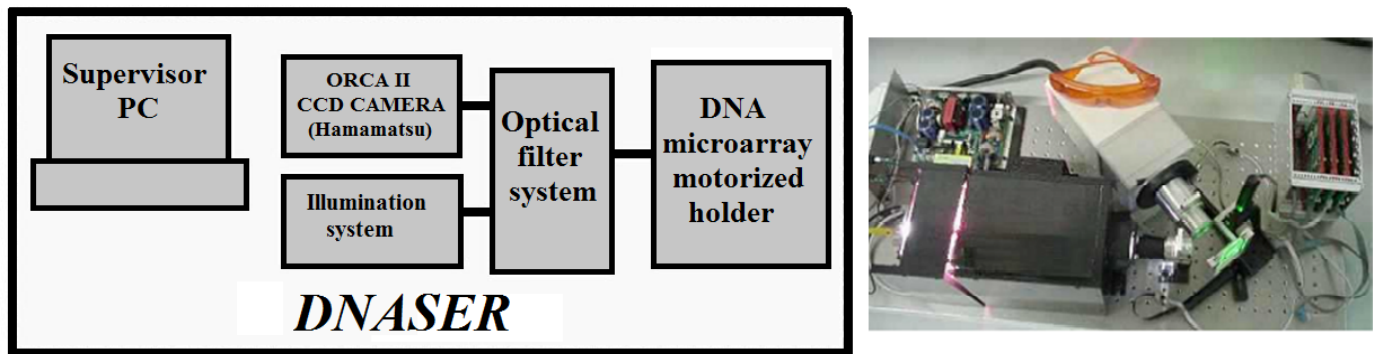


Figure 2: a scheme of the DNASER (left) and a picture of it (right (taken from [Troitsky et al., 2002] and [Sivozhelezov et al., 2010]).

4.3. LEADER GENE ALGORITHM

Pieces of information about gene expression and regulation in a given biological system (such as a biochemical or biophysical reaction) or in a disease are the main goal of functional genomics. In the last years, the search has shifted from studying a single gene deregulation to the investigation of all the gene-gene interactions and networks, finally arriving to the concept of genomic signature [Lammers et al., 2012].

The main purposes of gene-ranking or gene prioritization [Masoudi-Nejad et al., 2012] are to provide robust molecular signatures of the studied events, in order to discover reliable, sensitive and reproducible biomarkers, as well as to predict gene function (GFP), or further biological targets of potential pharmaceutical interest [Xu and Wang, 2013].

The Leader Gene approach exploits a search-dependent statistical algorithm, which is based on the systematic, exhaustive and recursive search for the genes and their products, on the computation of their interaction map, and on their ranking according to the number of all experimentally established interactions, as derived from publicly available Web-servers and databases, such as STRING (Search Tool for the Retrieval of Interacting Genes, Heidelberg, Germany) [von Mering et al., 2005].

The comprehensive search is usually carried out mining different databases and repositories containing DNA micro-arrays data, integrating the obtained hits.

The computation of an interaction map of the obtained list of genes is computed using the weighted number of links (WNLs). This measure is calculated for each gene using the program STRING, and this value is derived from the weighed sum of three types of interactions:

1. literature co-occurrence of the names of genes, and/or their products in abstracts of papers available on the Internet. The scores assigned are derived from benchmarked and validated scoring system based on the frequencies and distributions of gene/gene products names in the aforementioned abstracts. The benchmarks themselves are set from manual evaluation of predictions of gene and protein interactions by experts, and are typically below 0.5;
2. scores derived from databases dedicated to gene networks, containing data on induction and expression of a particular genes by other genes derived from microarray experiments, or other high-throughput omics techniques. The score of 1 is assigned if the link is already present in the database, while putative links have lower values (typically in the range 0.6–0.8);
3. the same range of scores is assigned to gene interactions via physically observed protein-protein interactions. The software does not discriminate between *in vivo* or *in vitro* experiment derived data. Generally, the scores are close to those of interaction type 2, but links of this type occur much rarely than of type 2.

The combined association scores S_{ij} were summed for each gene i over its neighbors (i,j), giving the final weighted number of links for the gene i . Further, we applied clustering methods to the weighted number of links in order to identify the group of leader genes. Cluster analysis, also called segmentation analysis or taxonomy analysis, is a way to partition a set of objects into homogeneous and separated groups or clusters, in such a way that the profiles of objects in the same cluster are very similar and the profiles of objects in different clusters are quite distinct. In particular,

cluster analysis can be defined as follows:

Given a set S of n objects $\{x_1, x_2, \dots, x_n\}$, where each object is described by m attributes $x_i=(x_{i1}, x_{i2}, \dots, x_{im})$, determine a classification that is most likely to have generated the observed objects.

Genes belonging to the highest rank are defined as “*leader genes*” or “*hub genes*” because they may be assumed to play an important role in the analyzed processes. The “*Leader Gene approach*” can suggest a list of few, but strong candidate genes potentially relevant within a given cellular process or a pathology, according to the already available experimental data. Moreover, the interaction map among all the genes involved in the same process may be useful in interpreting the experimental and clinical results, and in planning new *ad hoc* targeted experimentation. Interestingly, such experimentation may be simpler to be analysed than mass-scale molecular genomics, whose wealth of details may raise problems and complications [Nicolini, 2006; Nicolini, 2010]. This computational method gave promising results, when applied to the human T lymphocyte cell cycle, human kidney transplant, oral lichen planus and periodontitis. These results were also integrated with a targeted experimental analysis, to draw an overall picture of these processes, and are reviewed in the following paragraphs.

This interactive, automatic and user-friendly stand-alone tool has been written in house in Java, Javascript, PHP and HTML. The completely automated pipeline is performed via NCBI e-utilities (e-search, e-fetch, for further information the author is referred to the NCBI site), and other similar facilities.

A pictorial scheme of this tool is given in *Figure 3*, while the algorithm is shown in *Figure 4*.



Figure 3: screen-shot of the software GenEthics based on the Leader Gene algorithm.

The clustering techniques the user can choose are: Clustering K-means and Chinese whispers (specifically designed for graph clustering); as far as the number of clusters is concerned, the user can choose from heuristic number or provided by the user himself.

FLOWCHART	INPUT	OUTPUT
Identification of genes list	Organism Keywords Databases Keywords for cross-check Confidence threshold	List of relevant genes
Calculation of WNL	List of relevant genes Confidence threshold Interaction map Map as list of pairs in MEDUSA format	Relevant genes + WNL Interaction map (optional)
Genes clustering	Relevant genes + WNL	Genes clustered

Figure 4: the algorithm of the GenEthics software.

Moreover, it is possible also to compute different topological properties of the bio-graphs obtained, such as the number of nodes and edges, the mean connectivity, the graph degree, diameter and size, the density and heterogeneity, the clustering coefficient, and the percentage of some particular sub-networks [Racapé et al., 2012].-

5. NANOPROTEOMICS

Nanoproteomics can be defined as that broad, multidisciplinary science that studies protein structure and functions in an innovative way with respect to the previous methods [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b].

Here, the prefix “*nano-*”, once again, has a lot of nuances and meanings: it refers to the fact that nanoproteomics exploits the nanobiotechnologies, but also to the fact that it enables the study of peptides using very small amounts of them and despite the low-consumption it offers a high sensitivity and specificity in the protein detection and analysis.

Throughout the years, different approaches have been developed and tested to assay the protein structure and function. In the former case, many examples can be quoted: from the single-molecules techniques such as atomic force microscopy (AFM), MS, the biomembrane force probes (BFPs), the optical and magnetic tweezers, microneedles and flow-chamber devices, to the fluorescence-based devices (among the others, the fluorescence resonance energy transfer or FRET techniques) [Dufrière, 2009] and to the X-ray crystallography and the recent third-generation synchrotrons., including the advanced Microbeam grazing-incidence small-angle X-ray scattering (MICROGISAXS) techniques [Nicolini et al., 2014] (*Figures 12, 13*).

Nanocrystallography (*Figure 5*) is, of course, an important component of the nanoproteomics-based approaches.

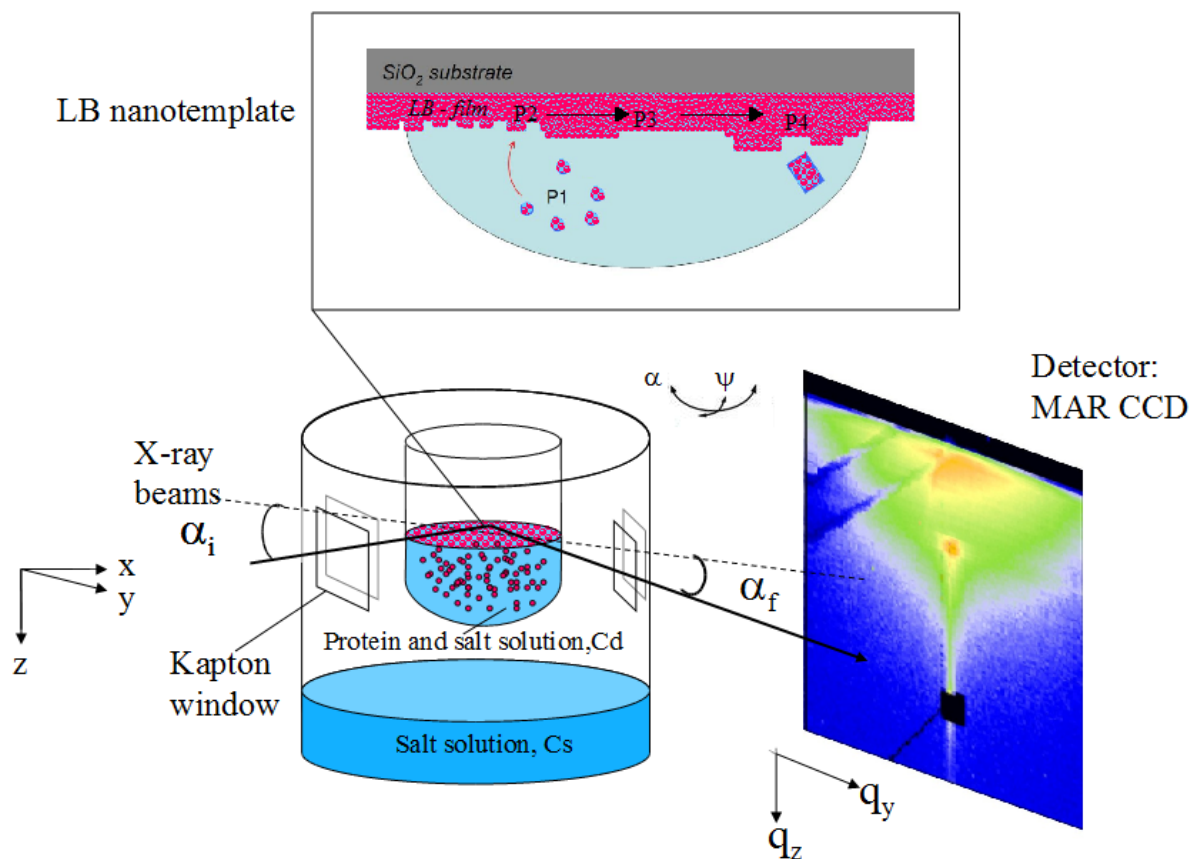


Figure 5: Experimental lay-out for in situ μ -GISAXS simulation of protein crystallization by LB nanotemplate method. From [Pechkova et al., 2010a; Pechkova et al., 2010b].

As far as the function is concerned, scholars have developed and exploited the nanofluidic devices [Okanda and El Rassi, 2006; Pasa-Tolić et al., 2004]. Gold, iron oxide and silver nanoparticles, carbon nanotubes,

nanowires and porous metal-organic-frameworks prove to be effective platforms for biomarker discovery [Agrawal et al., 2013; Chen and Li, 2013; Dasilva et al., 2012; Jia et al., 2013].

MS can be exploited both in qualitative and quantitative ways (Isotope-coded affinity tags or ICAT, Isobaric labeling Tandem mass tags or TMT, Isobaric tags for relative and absolute quantization or iTRAQ), metal-coded tags or MeCAT, N-terminal labeling, Stable isotope labeling with amino acids in cell culture or SILAC, Terminal amine isotopic labeling of substrates or TAILS, and Selected Reaction Monitoring or SRM) [Ong and Mann, 2005].

A key technology in the nanoproteomics field is represented by the protein microarrays, that constitute the natural evolution of the concept of gene/DNA microarrays. Moreover, the protein array represents a major advancement in the field of proteomics and of the previous biomolecular tools, such as the protein purification, expression and the protein engineering.

Purifying a protein is a time-consuming and difficult task. Usually a Ni-NTA high affinity chromatography is needed in order to capture the recombinant protein that carries a poly-histidine tag. It requires a cellular system (usually bacterial, but recently also mammalian, or fungal). Yeast (*Saccharomyces cerevisiae*) or baculovirus/insects may also be used.

However eukaryotic expression vectors should be preferred since a) they produce higher yields of recombinant protein, b) the produced protein remains soluble (instead, in the bacterial system it may precipitate within the so-called “inclusion bodies”), and c) they enable to study proteins in their natural milieu and environment (since eukaryotic proteins undergo

Chapter 5. Nanoproteomics

extensive post-translational modifications). There are many cell-free protein microarrays (such as the in-situ puromycin capture assay, the Protein In-Situ Array or PISA, the DNA Array to Protein Array or DAPA, and the nanowell format assay), but we chose to focus on the NAPPA approach because of its advantages which will be presented and described in-depth in the following paragraphs.

5.1 STRUCTURAL NANOPROTEOMICS AND PROTEIN NANOCRYSTALLOGRAPHY

In the last decades, a major progress in the field of protein structure determination was achieved thanks to advancements in X-ray crystallography combined with the third generation synchrotron. This unique method based on a combined approach is likely to remain the most important structure determination method for the foreseeable future. In the literature there is a dearth of both structural and functional studies devoted to the investigation of therapeutically and clinically relevant proteins, because of the initial crystal production reluctance and the X-ray radiation induced damage, that limits the quality of the collected diffraction data. Moreover, crystallization is a highly demanding and time-consuming task, being a real bottle-neck in the nowadays basic research. Several efforts have been made in order to understand which factors and parameters can influence this process.

Langmuir-Blodgett (LB) nano-template crystallization method has proven to give prominent results in target proteins crystallization. The advantage of using this nanobiotechnology is the higher quality of crystals both in terms of X-ray diffraction and radiation stability when employing the high energy X-ray source and focused beams, such as the third generation synchrotrons and the micro-diffraction beamlines. Differences in protein crystal nucleation, formation and growth between classical hanging-drop and LB-facilitated crystallization were been established by a variety of methods including *in situ* Grazing-Incidence Small-Angle Scattering (GISAXS), laser-micro-dissection combined with nano- and micro-focus beamlines,

Raman spectroscopy and Atomic Force Microscopy (AFM) [Pechkova et al., 2012].

Introduction of LB film affects protein stability as in the case of lysozyme, which turned out quite thermally stable. Previously, on the basis of a mass-scale analysis of crystal structures, we have established a role of the aqueous surroundings of a protein in its thermal stability. In another previous study we noted that introduction of LB film affects the aqueous environment of the protein leading to smaller numbers of water molecules. Recently, we proved that the introduction of thin biofilms directly affects the aqueous environment of the protein.

Different crystallization techniques have been proposed during the decades, like hanging drop vapor diffusion method and its variant (sitting drop), dialysis, cryo-temperature, batch, and even in space (using techniques like free interface diffusion or FID, counter-ion diffusion or CID). Recently, LB-based crystallization technique has been proposed in order to improve crystal growth and for this reason in this study is carried their systematic comparison out with in space ones and classical ones without LB.

Here, using both bioinformatics approaches, namely clustering algorithm and protein alignment and Molecular Dynamics (MD), we show how LB-based crystals can be compared with those obtained in space.

Protein selection

We decided to focus our study selecting those proteins that are considered models in structural crystallography and for which a comparison with microgravity-obtained structures was possible. Namely, these proteins are: lysozyme (EC 3.2.1.17, of *Gallus gallus*), thaumatin (plant sweet protein, of *Thaumatococcus daniellii*), proteinase K (EC 3.4.21.64, of *Engyodontium album*, formerly known as *Tritirachium album*) and human insulin (*Homo sapiens*). For each one protein, we downloaded all the structures present in the PDB database, then we iteratively refined the choice excluding proteins belonging to other *taxa* or radiation damaged structures, and finally we subdivided the structures using the crystallization procedure as variable. Since protein selection is a crucial and key step for this study, we clarified the protocol and the flow-chart for each studied protein.

Clustering algorithm and protein alignment

The most accurate method of 3D protein structure alignment algorithm based on the TM-align topological score was used.

Root mean square deviation (RMSD) for C- α atoms was used as the similarity measure for all structures. All calculations were performed using the web-based protein structure comparison tool ProCKSI.

The clustering algorithm was used selecting the Ward distance, both as All-against-all and as All-against-target options.

Clustering software

The files for tree diagrams building were prepared according to the Newick format (file extension *.newick*), which exploits a standard that makes use of the correspondence between trees and nested parentheses. The trees were visualized using a tree editor, Dendroscope.

Biostatistical analysis

Statistical analysis was performed by means of Statpages (Technical University of Denmark), Excel version 2011 14.1.3 (Microsoft Corporation – Redmond, Washington) and SPSS18 (IBM Corporation, USA; <http://www-01.ibm.com/software/>). ANOVA-one way (Analysis of Variance) was performed, and *p*-values less than 0.05 were considered statistically significant.

Molecular dynamics

All proteins were solvated and neutralized adding Cl⁻ or Na⁺ depending on the charge of the protein. OPLS-aa (optimized potentials for liquid simulations all atom) force field

For the long-range electrostatic interactions, we used Particle-Mesh Ewald (PME) method, which calculates direct-space interactions within a finite distance using a modified Coulomb's Law, and in reciprocal space using a Fourier transform to build a "mesh" of charges, interpolated onto a grid. It is from this charge interpolation that long-range forces can be calculated and incorporated into the non bonded interactions in a simulated system. For the van der Waals interactions, a typical 12 Å cut-off was used.

Chapter 5. Nanoproteomics

As a result, lysozyme was put in boxes of 6X6X6 nm, thaumatin in boxes of 7.4X7.4X7.4 nm and proteinase K in boxes of 7.23X7.23X7.23 nm.

A standard equilibration procedure was used and adopted for all systems. We started with a conjugate gradient minimization procedure and then with a two-stages equilibration, each step consisting of 100 picoseconds molecular dynamics simulation. In the first step protein was frozen in space and only solvent molecules were equilibrated, in the second step all atoms were equilibrated.

The obtained configurations were used as the starting point of other 1,000 picoseconds NpT (isothermal-isobaric) simulation, using a statistical mechanical ensemble maintaining constant temperature and pressure.

Visualization software

Chimera software (Chimera LLC, UCSF) and VMD software (Visual Molecular Dynamics, available on the World Wide Web at www.ks.uiuc.edu/Research/vmd/) were used for visualizing MD simulations snapshots.

Radius of gyration

The radius of gyration of an object describes its dimensions, calculated as the root mean square distance between its center of gravity and its ends. In analyzing proteins, the radius of gyration is indicative of the level of compaction in the structure, i.e. how folded or unfolded the polypeptide chain is. The comparison was carried out on four different model proteins

using both biostatistical analysis, bioinformatics approaches and molecular dynamics (MD).

Biostatistical analysis

We looked for an explanation of the findings obtained from clustering analysis from a quantitative and bio-statistical point of view: as can be seen in Fig. 3, LB insulin crystal follows the trends of microgravity crystals, that is to say higher resolution, lower content of water (in term of the Matthew coefficient), lower B-factor (i.e. thermal noise and atom displacement), higher number of reflections (i.e. the crystals diffract better).

ANOVA-one way (analysis of variance) in fact shows statistically significant difference among the groups for B factor (*p-value* <0.05, with *p*=0.024) and for water content (*p-value* <0.05, with *p*=0.032). In *Figures 5, 6 and 7*, trends comparisons for Resolution, B factor and solvent percentage among different crystallization techniques are shown, with the vertical bar showing the sigma level of the data and the yellow line the average behavior of the structures. As we can see the behavior of LB-based crystals is comparable with the behavior of space-grown crystals and thus, even though clustering results are positive for lysozyme and human insulin and partially negative for thaumatin and proteinase K, the statistical analysis of some parameters (like Resolution, B factor, solvent content) showed a common trend shared by the two crystallization techniques.

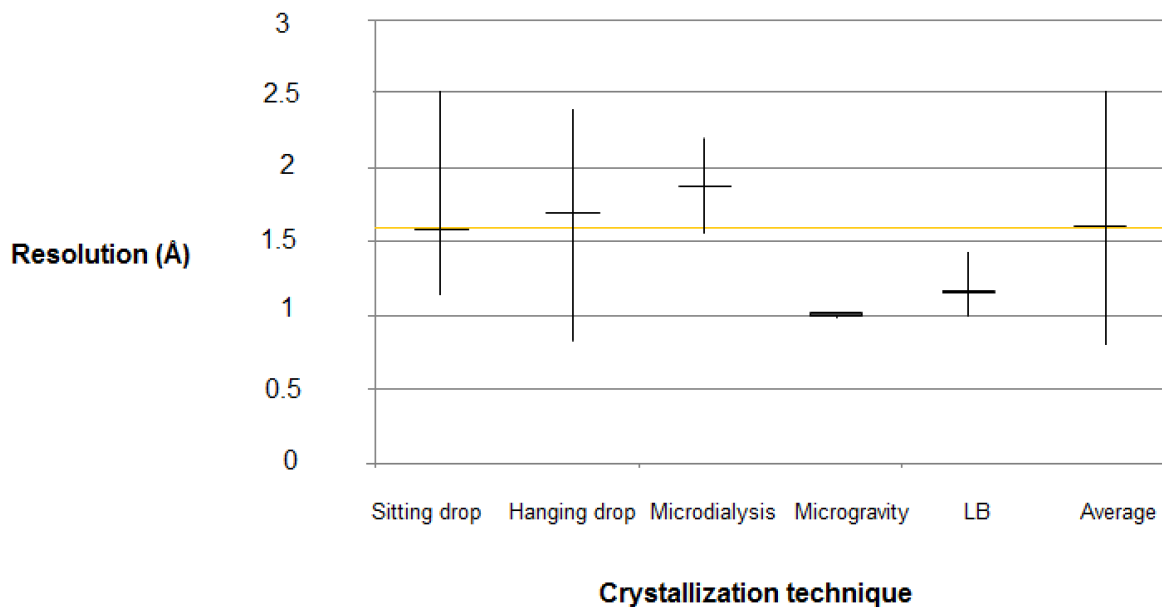


Figure 5: trends comparisons for resolution among different crystallization techniques for all the four studied target proteins are shown, with the vertical bar showing the sigma level of the data and the yellow line the average behavior of the structures.

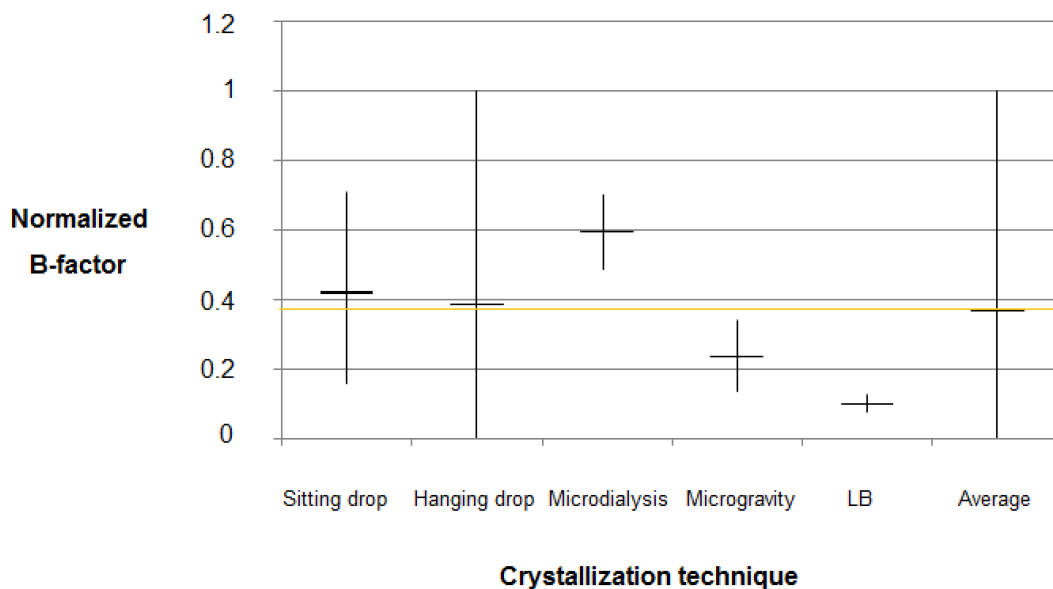


Figure 6: trends comparisons for the normalized Debye-Waller factor (DWF, known also as B factor in the field of macromolecular crystallography) among different crystallization techniques for all the four studied target proteins are shown, with the vertical bar showing the sigma level of the data and the yellow line the average behavior of the structures.

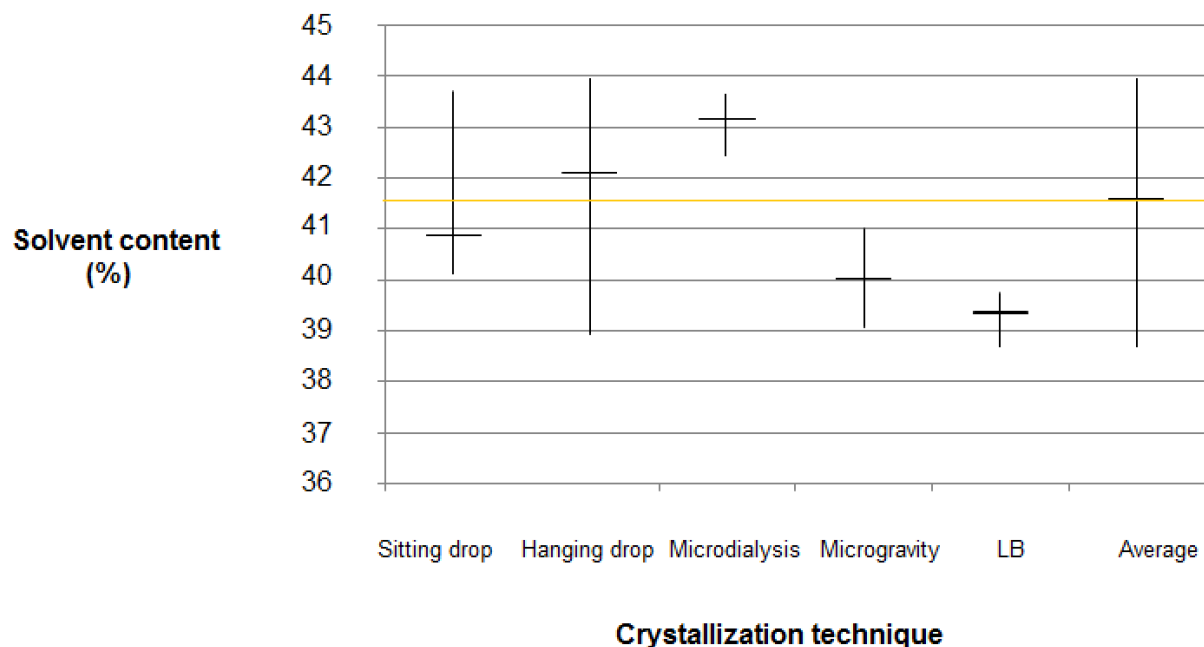


Figure 7: trends comparisons for solvent content (in percentage) among different crystallization techniques for all the four studied target proteins are shown, with the vertical bar showing the sigma level of the data and the yellow line the average behavior of the structures.

Human insulin

LB nanobiotemplate-based human insulin was first solved (PDB ID 4ihn) and the resolution distribution for Human Insulin structures makes evident that the LB insulin structure (4ihn) is the best among the 465 structures be-

ing deposited by all the different crystallization methods, existing world-wide.

Using ProCKSI server and our LB insulin as target structure we compared our LB structure with those structures already deposited in PDB database, taking into account similar experimental condition of crystallization.

A summary of the investigated structures is in *Table 1*. We found that the most similar structures to LB insulin are space-grown obtained crystals (as can be seen from *Figures 7* and *8*, and in *Appendix 1*), pointing to the fact that our LB insulin structure was clustered together with the microgravity cluster.

INSULIN METHOD	PDB IDS
HANGING DROP	1B9E, 1GUJ, 1HTV, 1LPH, 1TYL, 1TYM, 1UZ9, 1W8P, 1XDA, 2C8Q, 2C8R, 2CEU, 2G54, 2G56, 2OLY, 2OLZ, 2OM0, 2OM1, 2OMG, 2OMH, 2OMI, 2QIU, 2R34, 2R35, 2R36, 2VJZ, 2VK0, 2W44, 2WBY, 2WC0, 2WRU, 2WRV, 2WRW, 2WRX, 2WS0, 2WS1, 2WS4, 2WS6, 2WS7, 3BRR, 3EXX, 3ILG, 3INC, 3IR0, 3KQ6, 3P2X, 3P33, 3Q6E, 3ROV, 3TT8, 3UTQ, 3UTS, 3UTT, 3V1G, 3V19, 3ZQR, 3ZS2, 3ZU1
LB	4IHN
BATCH	1QIY, 1QIZ, 1QJ0, 1ZEG, 1ZEH, 1ZNJ
SITTING DROP	3U4N
MICROGRAVITY	1BEN, 1G7A, 1G7B, 1MSO, 1OS3, 1OS4, 3E7Y, 3E7Z, 3I3Z, 3I3Z, 3I40

Table 1: all the Human Insulin PDB structures that have been used for the clustering tree, divided according to their crystallization technique.

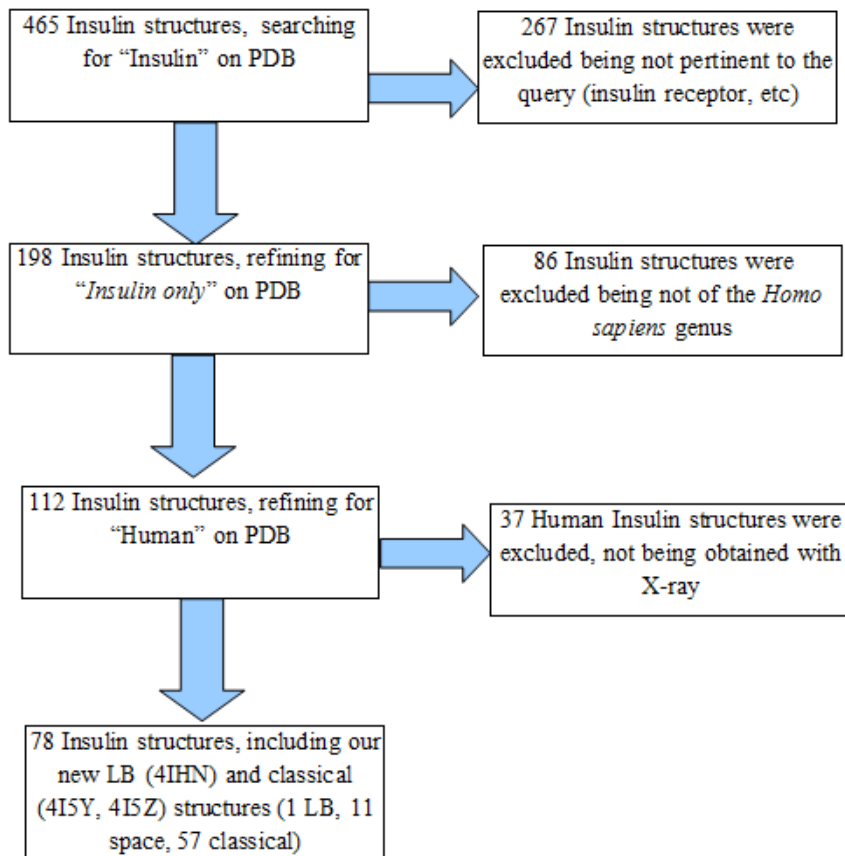


Figure 7: inclusion/exclusion flowchart for PDB insulin structures.



The ProCKSI-Server
Protein (Structure) Comparison, Knowledge, Similarity and Information

Human insulin structural alignment and clustering

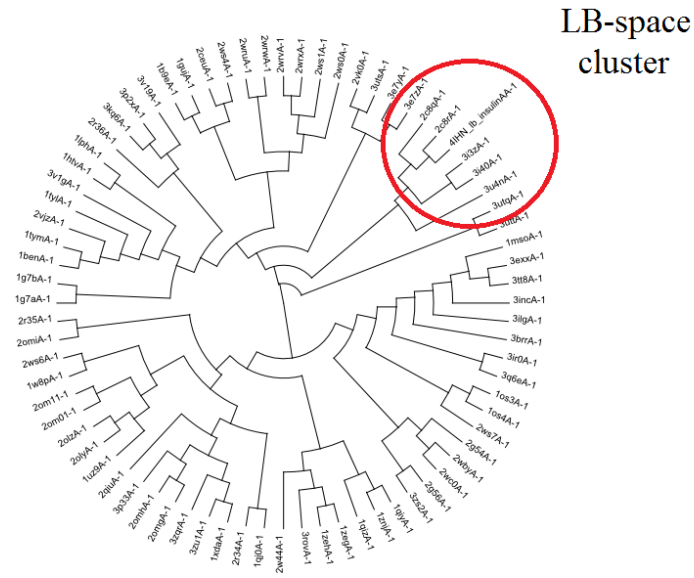


Figure 8: clustering diagram for Human insulin structures deposited on PDB. The clustering analysis has been performed via ProCKSI server.

Lysozyme

Once again, confirming the results obtained with lysozyme structures we compared all the Hen Egg-White Lysozyme (HEWL) structures in PDB database.

Our LB-obtained structures are PDB ID 2aub, 4ias, and 4iat, which have been recently acquired. Using clustering algorithm we found an exceptional quality structure space-grown lysozyme structure (PDB ID 1iee) clustering with 2aub, with the RMSD of only 0.09 Å, together with other high quality lysozyme crystals among the structures most similar to the LB-based lysozyme (as shown in *Fig. 9*). The LB clusters are well reproducible, as well as the LB-microgravity ones.



The ProCKSI-Server
Protein (Structure) Comparison, Knowledge, Similarity and Information

HEWL structural alignment and clustering

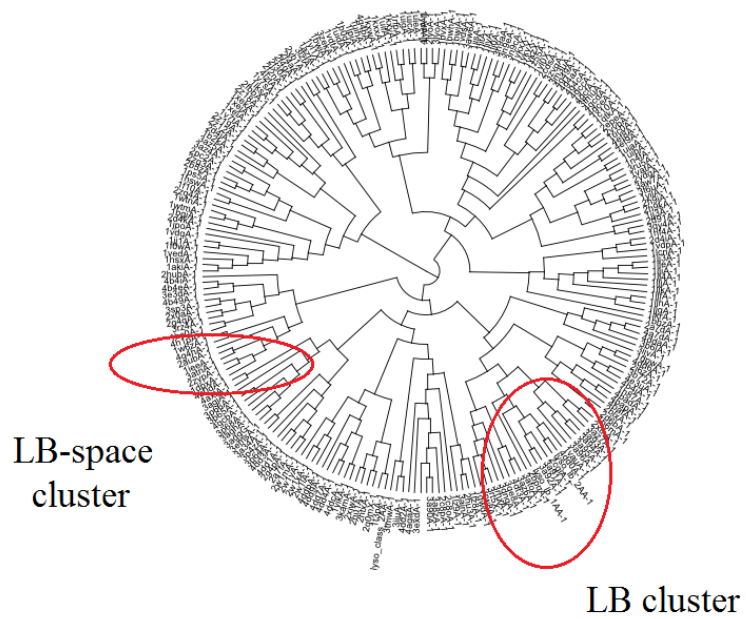


Figure 9: clustering diagram for HEWL lysozyme structures deposited on PDB. The clustering analysis has been performed via ProCKSI server.

Molecular dynamics

Figure 12 shows the RMSD and Figure 13 shows the normalized radius of gyration's behavior of the analyzed structures at 300° K, where in both cases strikingly differences can be seen.

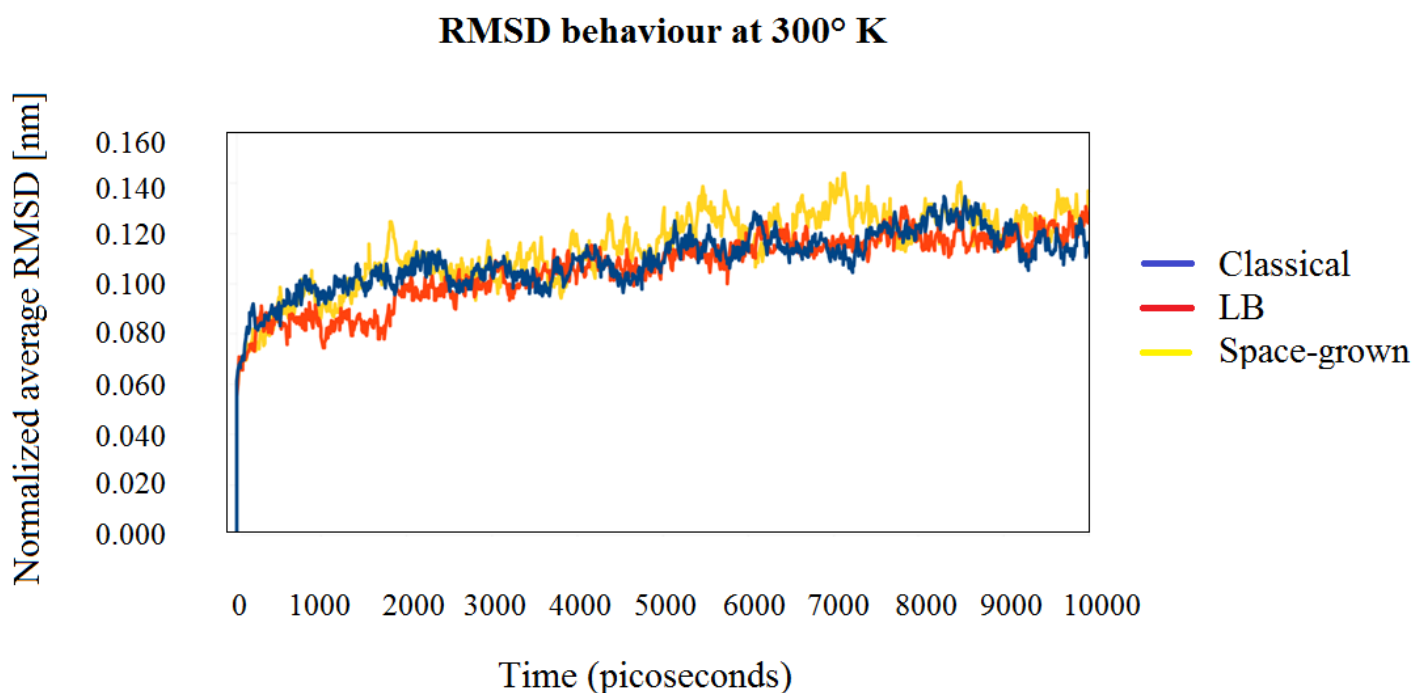


Figure 12: root mean square deviation (RMSD) of the selected crystal structures simulated with molecular dynamics at 300 K.

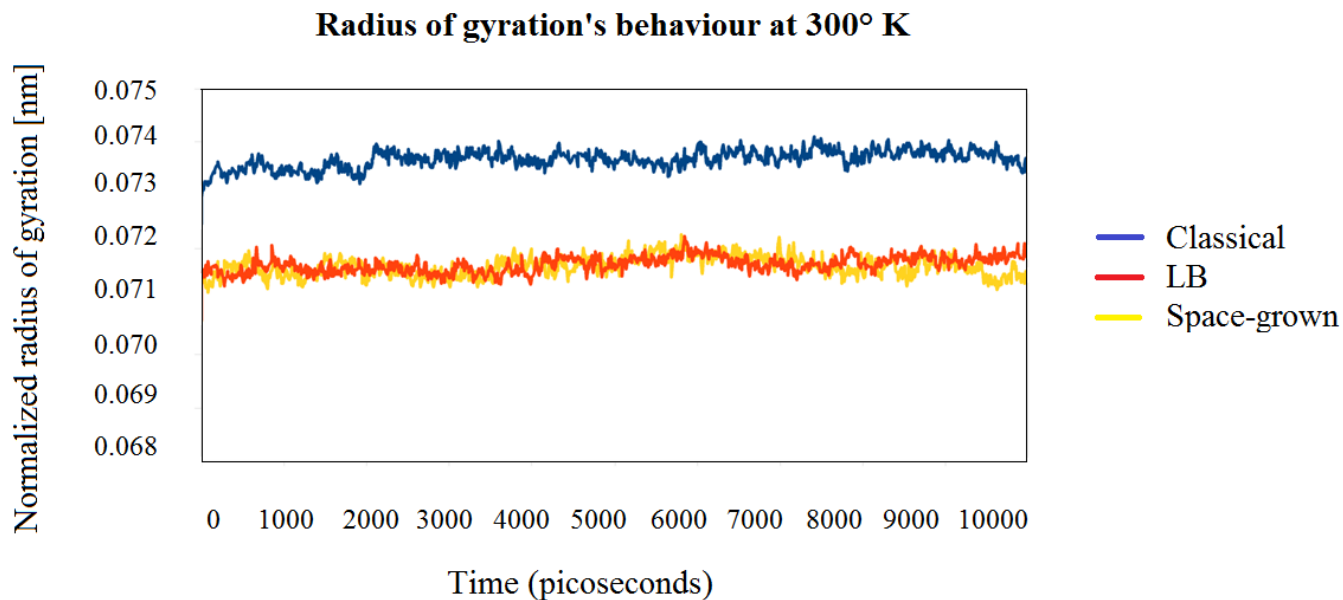


Figure 13: radius of gyration of the selected crystal structures simulated with molecular dynamics at 300 K.

In conclusion, space-grown crystals is a topic that has fostered to produce a huge body of research and has raised also some controversies. While some scholars claim that in absence of gravity crystal growth is better, other studies show no difference between ground-grown and space-grown crystals.

However, it has been claimed that microgravity environment should minimize “*buoyant convection flows, limiting nucleation sites, and eliminating crystal sedimentation*” [Ng, 2002].

Chapter 5. Nanoproteomics

We found that for lysozyme and human insulin structures were absolutely comparable, while similar evidence was collected also for Thaumatin and proteinase K.

From clustering algorithm, bioinformatics, and biostatistics, we can conclude that :

- 1) bioinformatics and clustering algorithms can be applied to the study and modeling of proteins crystallized according to different crystallization techniques;
- 2) according to the clustering algorithm and statistical analysis of parameters like resolution, B-factor, solvent content, LB-based and microgravity proteins are comparable and different from proteins crystallized with other techniques;

All these findings obtained with bioinformatics and MD approaches, if taken together are converging and prove that LB-proteins and space-grown ones are similar and appear the best methods to carry out protein crystallization .

In the second part of our structural proteomics study we used an advanced technique in order to get insight on protein crystallization with the LB nanobiotemplate.

Grazing-incidence Small Angle X-ray scattering (GISAXS) is an advanced scattering technique that can be used to unravel these processes and to investigate large-scale structures in thin films, including biofilms.

A combination of this technique with third-generation synchrotron radiation micro-beams (μ -GISAXS) [Pechkova et al., 2010a; Pechkova et al., 2010b] has been used for studying different kinds of surfaces, like surface gradients or confined surfaces. The potential for studying thin protein films by μ -GISAXS during nanotemplate-assisted crystallization experiments was previously demonstrated in a study about the effect of temperature on long-range order and during both *ex situ* and *in situ* experiments. *Ex situ* data obtained using both Cytochrome P450_{sc} and Lysozyme were extremely complex to interpret and needed to be studied in parallel with microscopy characterization, using either the classical method or the nanotemplate hanging-drop method, as well as *in situ* data.

Moreover, *ex situ* experiments were partially limited by the discontinuous nature of the studies themselves and for addressing this issue, *in situ* experiments were carried out, designing an *ad hoc* apparatus, namely the flow-through crystallization cell, which, besides the conventional crystallization set-up, is made up by two kapton windows being inserted into the outer cell walls (*Figure 5*). For obtaining a rapid and proper buffer exchange, the reservoir was connected via Teflon tubes to two Harvard syringe pumps. The main advantage of *in situ* experiments was the possibility of monitoring the different processes of crystal growth in real time.

The aim of this work is to shed light on the processes of nucleation and crystal growth starting from a LB-protein film by quantitatively analyzing previously acquired *in situ* μ -GISAXS experiments with suitable models [Pechkova et al., 2010].

Experimental Thaumatin and Lysozyme μ -GISAXS

In situ scattering experiments were performed at the ID13 microfocus beamline facility at the European Synchrotron Radiation Facility in Grenoble, France.

A typical μ -GISAXS pattern is indexed as function of the wave-vector transfer $Q=(Q_x, Q_y, Q_z)$ which parallel Q_y and perpendicular Q_z components scale with the in-plane 2θ and exit α_f scattering angles. Knowing α_i is the grazing incidence angle, Q_y and Q_z can be computed as follows:

$$Q_y = 2\pi/\lambda \sin(2\theta) \cos(\alpha_f)$$

$$Q_z = 2\pi/\lambda [\sin(\alpha_f) + \sin(\alpha_i)] \text{ while } Q_x \text{ is negligible in most conditions.}$$

The specular and the so called Yoneda Peak occurs as characteristic features in the scattering pattern.

Specular scattering is observed for $Q_x = Q_y = 0$, $Q_z > 0$, with the specular peak appearing when the specular condition is fulfilled ($\alpha_i = \alpha_f$), and while diffuse scattering is observed for Q_z , $Q_y \neq 0$. The so called Yoneda Peak $\alpha_c = \alpha_f$ occurs at the critical angle α_c of the sample. Correlations vertical to the sample surface can be probed along Q_z at $Q_y = 0$.

Data reduction

The Fit2D software package was used for data reduction. GISAXS patterns were analyzed with the *IsGISAXS* software which is dedicated to simulation of scattering from supported nanostructures. The scattering cross section is

expressed in terms of island form factor and interference function and the specificity of the grazing-incidence geometry is stressed, in particular in the evaluation of the island form factor in the distorted-wave Born approximation (DWBA). A full account of size and possible shape distributions is given in the decoupling approximation (DA), where sizes and positions are not correlated, and in the local monodisperse approximation (LMA). Two types of island repartitions on the substrate can be implemented: disordered systems characterized by their particle-particle pair correlation functions, and bidimensional crystalline or paracrystalline systems of particles.

TABLE 2 . Modeling Parameters of the simulation.	Value
Wavelength (nm)	0.0991
Incident angle (deg)	0.71
Delta – Glass (index of refraction)	3.336e-6
Delta – Thaumatin (index of refraction)	2.19e-6
Delta- Lysozyme (index refraction)	
Beta – Glass (absorption)	1.68e-8
Beta – Protein (absorption)	~0
LB 2 Layers thickness Thaumatin (nm)	7.4
LB 2 Layers thickness Lysozyme (nm)	6.4

Choice of parameters

Chapter 5. Nanoproteomics

In *table 1*, the parameters which have been used in our simulation for modeling the crystal growth have been reported.

Proteins have been modeled as cylinders, LB film thickness, found with the best fit, was fixed at 7.4 nm for Thaumatin and at 6.4 nm for Lysozyme, while the wavelength was experimentally known (0.0991 nm). Critical incident angle for Thaumatin and for Lysozyme was computed to be 0.71 deg. The delta refraction coefficients were $3.336 \cdot 10^{-6}$ for glass, $2.19 \cdot 10^{-6}$ for proteins. The beta absorption coefficients were about 0 for proteins, $1.68 \cdot 10^{-8}$ for glass.

Curves have been fitted with a χ^2 Levenberg–Marquardt minimization procedure.

Unique combination of GISAXS, synchrotron microfocus beam and computer simulation has enabled us to follow the nucleation and growth of the Lysozyme and the Thaumatin crystals.

At 100 minutes, the particle radius of Thaumatin is about 5.62 nm, while the LB-film layer thickness is about 2.25 nm, with height ratio of 1 nm (*Figure 14*, where is shown the perpendicular cut along Q_z).

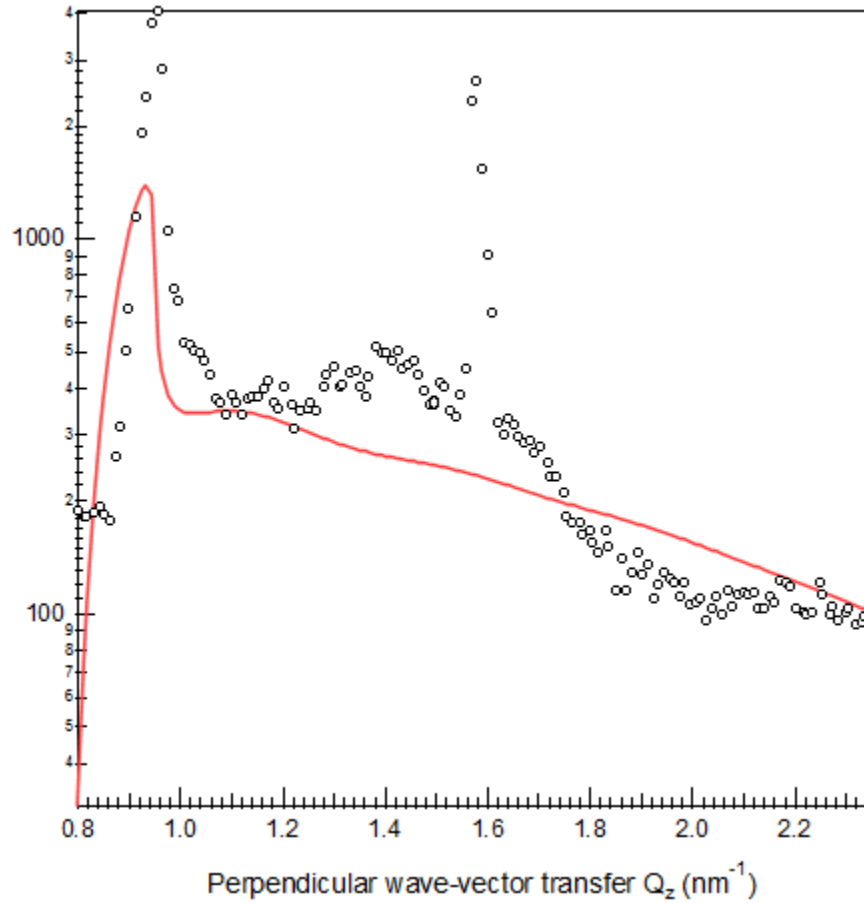


Figure 14: Thaumatin – 100min – Modeling crystal growth. Cylinder: Size = Fixed – Height = 1.7nm – Fixed LB thickness = 7.4 nm.

At 900 minutes from the start of the experiment, the particle radius has increased up to 40.89 nm, while on the contrary the LB-film layer thickness has decreased down to 0 nm, with a decreased height ratio of $6.30 \cdot 10^{-2}$ nm. In *figure 7* the cut at Q parallel is shown at right, while at left of the same *Figure 15* the cut at Q perpendicular is shown.

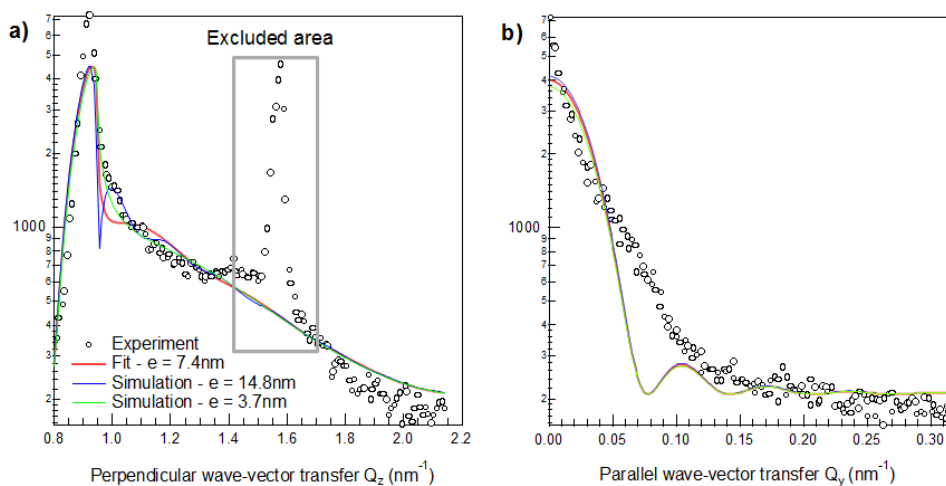


Figure 15: thaumatin – 900min – Modeling crystal growth. Cylinder: Size = 95 nm – Height = 5.6 nm – Fixed LB thickness = 7.4 nm. Blue line = Simulation with fixed LB thickness = 14.8 and 3.7nm. Perpendicular wave-vector transfer Q_z (nm^{-1}) (left), parallel wave-vector transfer Q_y (nm^{-1}) (right).

Previously, we hypothesized that the protein appears to transfer directly from the nanobiostructured film into the drop to directly trigger the formation of the crystal, therefore highlighting the physical interpretation of the mechanism for nanobiotemplate-facilitated protein crystallization [Pechkova et al., 2010a; Pechkova et al., 2010b]. This working hypothesis however was only qualitative and not quantitative, since the GISAXS spectra were complex and difficult to interpret. Indeed only computer simulation can help us in understanding the biological processes at sub-micron and nano-scale level, thus shedding light on protein nanocrystallography and paving the way for further research in the field.

For Lysozyme, a comparison between with and without LB-thin film was indeed possible in a different way, because the acquired data were quite less optimal than the Thaumatin one due to a poor alignment.

The process of crystal growth previously found for Thaumatin was found also for Lysozyme: in *Figure 16*, the Gaussian fit of the Yoneda peak region of LB-Lysozyme monitored at start and at end of the crystallization process, showing the gradual increase in R (which is the ratio of crystal protein scattering volume to glass scattering volume elevated to the square). At Q perpendicular cut, R is 0.27 at the beginning (*Figure 16, left*) and gradually shifts to 1.13 (*Figure 16, right*).

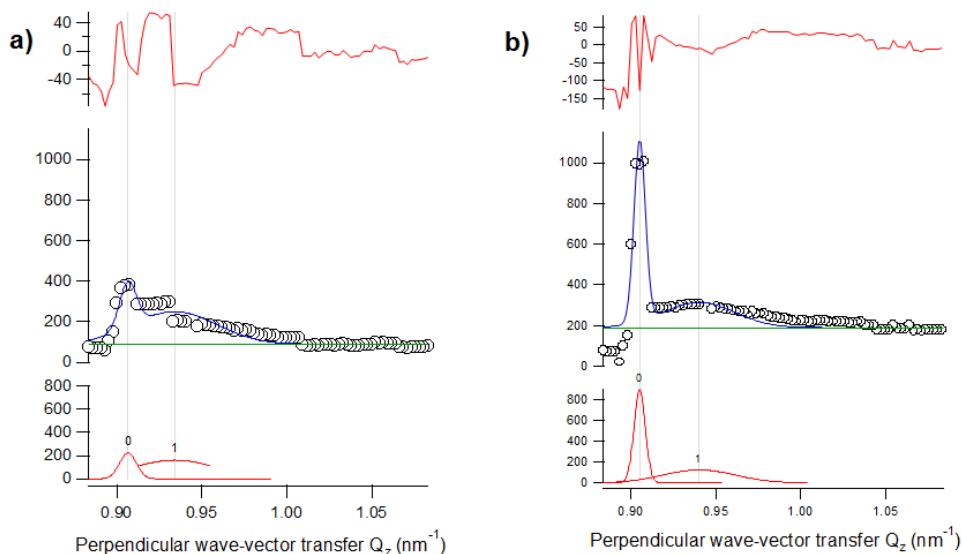


Figure 16: Gaussian fit of Yoneda peak region of LB-Lysozyme at the start (left) and at the end (right) of the crystallization process. R is the ratio of crystal protein scattering volume to glass scattering volume to the square. .

Taken together the data of Lysozyme and of Thaumatin, we have proved that LB-thin film acts a transferring nanobiotemplate and can really enhance and facilitate crystallization growth, which is often difficult and demanding.

Moreover, the quantitative findings we obtained are in perfect agreement with the model previously hypothesized.

The models here introduced are based on the submicron GISAX experiments being previously carried out at the ID13 beamline at ESRF on Thaumatin and Lysozyme and point to an highly unstable time dependent growth for Lysozyme crystals contrary to what is happening for Thaumatin growth. In addition the lysozyme experimentation appear compatible with a

Chapter 5. Nanoproteomics

model where the LB crystals keep growing at a constant rate while the classic crystals interrupt their growth in the same time interval.

LB-grown proteins show peculiar and unique features at the submicron size. LB-assisted crystal growth was previously studied by means of AFM, *ex situ* and *in situ* μ -GISAXS but obtained data were complex and difficult to reduce. IsGIXAS has enabled to model the crystal growth.

5.2. FUNCTIONAL NANOPROTEOMICS

5.2.1. NAPPA

Ab initio/targeted bioinformatic analysis leads to the selection of few genes (Section 2 of Fig. 1) and these can be the proper target queries of NAPPA, a novel tool for studying proteins and protein–protein interactions in which proteins are produced in a mammalian environment with natural post-translational modifications occurring and just in time, thus preserving natural folding, stability and freshness. Moreover NAPPA avoids the need of having purified proteins and instead uses plasmid-produced ones. NAPPA slides were printed. building upon the successful use of *in vitro* translated protein in standard scale applications, we replaced the use of purified proteins with the use of cDNAs. The proteins are translated adding a T7-coupled rabbit reticulocyte lysate (*in vitro* transcription– translation, *IVTT*) system [Ramachandran et al., 2004; Ramachandran et al., 2005; Sivozhelezov et al., 2010]. The engineering of each protein by adding a C-terminal glutathione S-transferase (GST) tag enabled its capture to the spot through an anti-GST antibody, printed on the array simultaneously with the expression plasmid. NAPPA technology is promising, as it overcomes the problems like protein purification, storage, misfolding and denaturation that can occur during the generation of the micro-arrays [Sivozhelezov et al., 2010].

Typically a fluorescent molecule is bound to the query protein by the linking antibody, with the labeled molecule being then detected via a fluorescence microscope, flow cytometer or some other fluorescence reading instrument such as the DNASER [Nicolini et al., 2006]. In our

protocol (*Figure 1*) we eliminate the need for a labeling molecule that may alter the properties of the query protein, utilizing QCM_D nanogravimetry, mass spectrometry and electrochemistry in conjunction with Anodic Porous Alumina [Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b]. According to the protocol each sample is printed with 10×10 NAPPA spots of 300 μm diameter, spaced 350 μm center-to-center, using the chosen genes and corresponding expressed proteins for the required medical diagnosis, prognosis or treatment. As validation of the label-free NAPPA implemented in the protocol we use the following test genes with 4×4 spots:

- Cdk2_Human Gene of 897 bp expressing 33.93 kDa protein;
- Jun_Human Gene of 996 bp expressing 35.67 kDa protein;
- P53_Human Gene of 1182 bp expressing 43.65 kDa protein;
- CdKN1A_Human Gene (called also P21) of 495 bp expressing 18,119 kDa protein.

The number of test genes is different for QMC-D and QMC-F (three of them), APA (2 of them) and MS (all four of them; next chapter). Typically the first label-free measure is carried out at 22 °C for blank slide to quantify the initial value in order to calculate the variation at the end of each step (i.e., in QMC), and then on each blank slide NAPPA spots were printed [Nicolini and LaBaer, 2010; Nicolini and Pechkova, 2010a; Nicolini and Pechkova, 2010b; Pechkova and Nicolini, 2004; Pechkova and Riekel, 2011]. The first step of NAPPA expression protocol was the blocking, at room temperature (22 °C) followed by a washing process. Gene expression and protein synthesis took place at 30 °C for about 1.5 h. Then, the temperature was decreased to 15 °C for a period of 0.5 h to facilitate the proteins binding on the spot surface [Spera et al., 2013].

Chapter 5. Nanoproteomics

The last step was the washing: the quartzes were washed at room temperature (22 °C), in pure water since the frequency stabilized or in PBS with this step being repeated for five times. The protocol described above was followed identically for both control and working reference slide in QCM_D.

To confirm protein expression and immobilization, parallel experiments were led earlier on several genes by fluorescence microscopy on NAPPA spotted on microscope glass [Bragazzi et al., 2011; Ramachandran et al., 2004; Ramachandran et al., 2008; Spera et al., 2013]. The NAPPA function and selectivity with respect to protein–protein interactions were indeed confirmed [Ramachandran et al., 2004].

For the four target plasmid DNAs (Jun, Cdk2, P53 and CdkN1A) here immobilized onto the glass slide binary interactions between several well characterized interacting pairs (including Jun-Fos and P53-MDM2) were tested and confirmed. The target proteins were expressed with rabbit reticulocyte lysate supplemented with T7 polymerase and the signals were detected with antibody to GST and tyramide signal amplification (TSA) reagent (Perkin Elmer). To verify that the detected proteins were the expected target proteins target protein-specific antibodies were used, which detected only their relevant spots. The four genes were queried for potential interactions with different query proteins: Jun, Fos and MDM2. The query proteins were added to the IVT lysate that was spotted on the NAPPA.

The interaction was detected using the appropriate protein-specific antibodies.

5.2.2. NANOCONDUCTOMETRIC QCM_D

From sound waves generated by the vibration of a quartz crystal microbalance (QCM) a sensor based on a quartz crystal coated with gold oscillating at a specific frequency was introduced time ago [Bradshaw, 2000; Höök et al., 2002; Rodahl and Kasemo, 1996a; Rodahl and Kasemo, 1996b]. Binding events change the frequency by varying the amount of mass linked to the surface. While with a competitive label-free approach like Surface Plasmon Resonance (SPR) [Feley et al., 2008; Matarraz et al., 2011; Wang et al., 2010] only kinetics and mass changes are detected, with the QCM-based method here described in addition to mass changes and kinetics can be detected also changes in viscosity on the surface; are indeed the protein conformational changes to appear causally related to the changes in viscosity.

In *Fig. 17* our quartz crystal microbalance being recently developed [Adami et al., 2010; Spera et al., 2013] is shown to monitor both the variation in frequency (QCM-F) and dissipation (QCM-D). This nanogravimetric device exploits acoustic waves generated by oscillating piezoelectric single crystal quartz plate to measure the frequency at which the amplitude of the oscillation driven by AC field will be strongly enhanced. The crystal will be in resonance at this characteristic frequency and the resonance frequency decreases at the increasing mass, provided it is firmly attached to the quartz crystal surface. By monitoring the frequency in real time (QCM-F), the temporal change in mass on the crystal can be investigated in a label free fashion.

The viscoelastic properties of the NAPPA quartzes are measured by the dissipation factor (QCM-D) jointly developed with Elbatech srl (Marciana,

Italy). Our system detects the dissipation factor of the quartz crystal by means of a detailed method usually known as the “*half-width half-height*” of its impedance curve [Spera et al., 2013]. Other methods are possible, such as measuring by fast acquisition the decay curve of the oscillation once it is suddenly interrupted [Bradshaw, 2000; Höök et al., 2002], but the impedance way has the advantage to maintain the crystal powered and to offer the whole picture of its oscillation mode.

In our case the quartz is connected to an RF gain-phase detector (Analog Devices, Inc., Norwood, MA, USA) and is driven by a precision Direct Digital Synthesizer (DDS) (Analog Devices, Inc., Norwood, MA, USA) around its resonance frequency, thus acquiring a voltage vs. frequency curve which shows a typical Gaussian behavior. The curve peak is at the actual resonance frequency while the shape of the curve indicates how damped is the oscillation, i.e. how the viscoelastic effects of the surrounding layers affect the oscillation. In order to have a stable control of the temperature the experiments need to be conducted in a chamber having the temperature maintained at 30 °C.

We have designed miniature flow-cell to follow washing water directly on the quartz with a Milli-Q water flow of 10 µl/s.

The scheme of the experimental set up is reported in *Fig. 17* for QCM-F and QCM-D. The QCM-D nanogravimeter has been positioned in a temperature chamber and monitored at the same time for frequency variation and dissipation factor variation.

“The nanogravimetric sample consists of 9.5 MHz, AT-cut quartz crystal of 14 mm blank diameter and 7.5 mm electrode diameter, the electrode material is 100 Å Cr and 1000 Å Au, and the quartz is embedded into glass-

like structures for easy handling” [Spera et al., 2013]. The quartzes were produced by ICM (Oklahoma City, USA). We are presently upgrading and finalizing the version of the QCM_D prototype dedicated to NAPPA. The QCM_D instrument, besides being connected to the PC, has two analogical displays, one for frequency shift and the second for the D factor, in such a way that allows to visualize, at the same time, information on the mass and the visco-elasticity of the sample. It proves to be of great help to place the quartz in a controlled temperature chamber. In this way it is possible to perform the NAPPA expression protocol directly in the nanogravimeter giving a further innovation to our prototype.

Another great advantage proves to have at least two inputs to attach at least two quartzes at the same time at the instrument (as for QCM-R3), one quartz as control and the other as working element performing also a direct match of different measurements controlled in real time. The progress is thereby not only in the sample holder, but in the actual implementation of a novel highly sensitive instrument dedicated to NAPPA, including a 3D image display of the simultaneous variation in impedance, frequency and time. The assembled QCM-D prototype has been optimized and validated with glycerol and various protein solutions [Adami et al., 2010].

The above quartz crystal microbalance techniques QCM-F and D-Factor QCM-D are used to investigate protein expression in NAPPA arrays when proteins can be synthesized directly on the substrate (*in situ*) and subsequently attached to the surface via capturing antibodies that bind the proteins by means of an antigen tag synthesized at one of their ends. The data confirm the higher sensitivity of QMC-D also revealing with the impedance curve CdK2, which instead cannot be detected by QCM-F.

The most efficient strategy coupled a psoralen-biotin conjugate to the expression plasmid DNA with the use of ultraviolet light, which was then captured on the surface by avidin. Finally, QCM-F/QCM-D along with Atomic Force Microscopy [Nicolini et al., 2013; Sartore et al., 2010] is a technique that not only validates protein expression occurring on the array surface, but also quantitatively measures the level and uniformity of protein expression, comparing directly and quite favorably to fluorescent distribution and intensity as evaluated via DNASER on the same protein [Nicolini et al., 2006; Sivozhelezov et al., 2010].

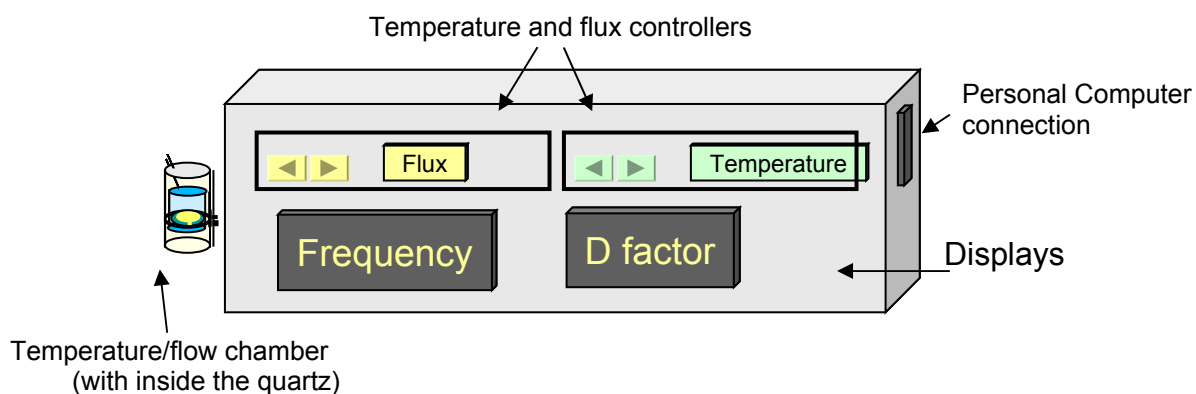


Figure 17: the experimental set-up of QCM.

5.2.3. APA

The APA surface can be prepared [Grasso et al., 2006; Stura et al., 2007; Stura et al., 2010] by a suitable electrolytic process designed to obtain a regular distribution of deep micrometric holes. The task of evaporating aluminum over glass has been accomplished by avoiding its detachment during the anodization process, a typical problem due to the incompatibility of cold borosilicate glass to the vapors of aluminum. This phenomenon was contrasted by means of a thin layer of chromium (deposited by sputtering) as medium element between glass and aluminum.

The high aspect ratio (depth/width ratio) of the pores makes this material also a natural wave guide for any fluorescent molecule present on the bottom of the pores, avoiding crosstalk of many point-light sources too close as frequently in fluorescent NAPPAs. The dielectric properties of Al₂O₃ makes this structure optimal for the realization of an electrically anisotropic system; the electrochemical reactions occurring on the bottom of the well (caused by the interaction between the biological probe molecule and the test molecule) induce variation on the electrical response to alternating voltage signal, and they can be quantified by means of scanning electrodes moving on the surface of the array, placed in a proper solution. The walls of the pore behave like insulator decoupling from the electrochemical events occurring few microns/millimeters away from the measurement place. This option constitutes the APA label-free approach to the analysis of protein arrays, since no fluorescent/marked molecule is utilized, while the alternative one still linked to APA requires a fluorescent molecule to spread the luminous signal through the APA wave guide. The potential of label-free approaches to complement and even to improve other

detection technologies of proteins being expressed in NAPPA microarrays has never been higher.

The application of APA [Nicolini et al., 2013; Sartore et al., 2010; Stura et al., 2010] as an advanced “on-chip laboratory” for an electrochemical detection of gene expressions is investigated with a NAPPA based on JUN and P53 gene sequences contained in an engineered plasmid vector.

At the same time the NAPPA printing on APA has proven:

- The ability to spot a colored fluid on the APA surface in discrete spots
- The ability to rapidly exchange that fluid with a different fluid
- The ability to repeat these manipulations as needed
- The production of the APA slides in a format compatible also with a fluorescence reader and not only with electrochemistry [Nicolini et al., 2013; Sartore et al., 2010; Stura et al., 2010]
- The structural integrity of APA slide format (either alone or over aluminum) has been proven acceptable for “routine” manipulation and especially during fluorescence readings, mainly relying for the stand alone configuration also on alternative printing which exploits capillary forces.

Last but not least the usefulness of APA substrate was shown [Nicolini et al., 2013; Sartore et al., 2010; Stura et al., 2010] as a candidate to isolate protein expression in a defined space followed by detection of the photoluminescence signal emitted from the complex secondary antibody anchored to Cy3 with a CCD microscope. The fluorescence of these labeled proteins was clearly evident in circular shaped arrangements on a limited surface of APA where the proteins were confined in the pores. APA surface appears to allow a label-free analysis using electric impedance spectroscopy (EIS). It is known that with EIS it is possible to detect different amounts of

Chapter 5. Nanoproteomics

organic materials deposited even indirectly over conducting surfaces. After the hybridization/expression experiment, using a scanning electrode controlled by a manipulator (MPC 200 by Sutter Technologies) via a PC, different EIS measurements were performed in different spots on the surface of the array in phosphate buffered saline solution (PBS); the experimental setup is shown in *Fig.18* emphasizing the feasibility of using APA for building microarrays..

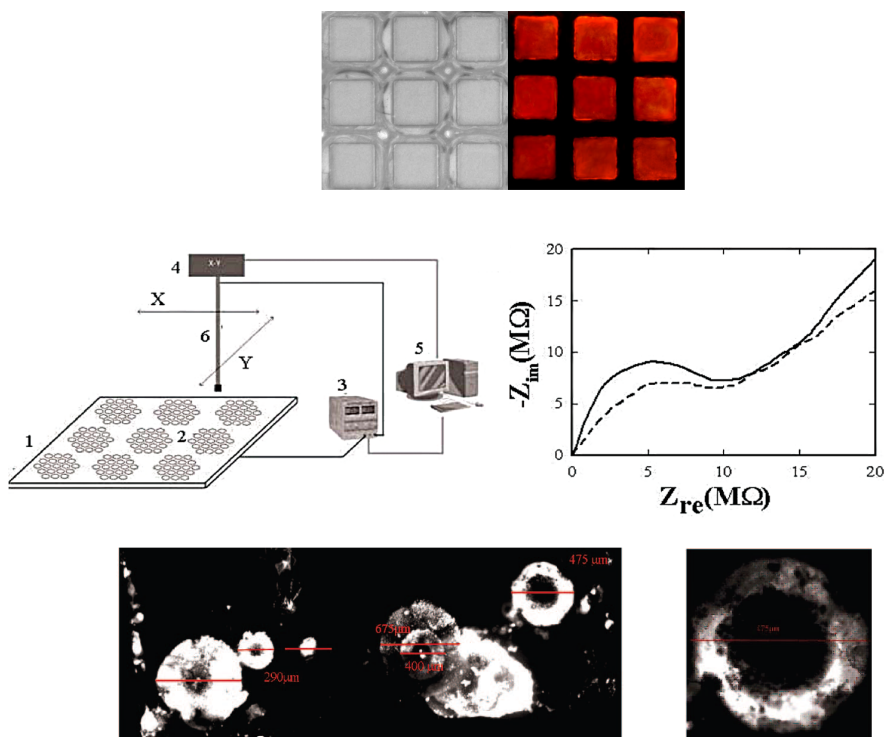


Figure 18: APA as a proper biomaterial for biomolecular microarray.

5.2.4. MASS SPECTROMETRY

In the present technique, we coupled the Matrix Assisted Laser Desorption Ionization Time-of-flight Mass Spectrometry (MALDI-TOF MS) to NAPPA technology and bioinformatic analysis for the detection of the proteins translated from the cDNA on the surface of the array. However, the development of a MALDI-TOF MS-compatible protein microarray is complex since existing methods for forming protein microarrays do not transfer readily into a MALDI target. For use with laser desorption/ionization mass spectrometry, a support having an electroconductive target surface was essential [Spera et al., 2011].

One of the challenges in evaluating the mass spectra obtained from NAPPA was the extra biological material present on the NAPPA together with the target proteins, such as the BSA protein. This also required interpreting the spectra of synthesized proteins in the context of the additional peptide chain, the GST tag, and the anti-GST antibody. Fortunately, these proteins are present in all features of the array, and represent a “common background” In this context, then, we used bioinformatic tools to better interpret the obtained results, namely we developed a matching algorithm to identify in the spectra of the given expressed protein [Spera and Nicolini, 2008; Spera et al., 2011] those peaks corresponding to the “background peaks” (*Figures 19, 20*).

The current NAPPA chemistry and the recent advances in mass spectrometry [Anderson et al., 2011; Ho et al., 2002; Zhu et al., 2010] allow us to potentially validate this label-free technology even in clinical settings through the correlation with the tremendous amount of fluorescence data already acquired over the years [LaBaer and Ramachandran, 2005; Nicolini

and LaBaer, 2010; Ramachandran et al., 2004; Ramachandran et al., 2005; Ramachandran et al., 2008].

In this case, the end game is to demonstrate that we can identify proteins, in particular proteins that bind to the target proteins on the array. Towards this end we have shown that we can identify the expressed proteins, printing on gold slides 4 different genes (each one with 16×300 μm spots) — first in a known configuration, then in an unknown configuration [Spera et al., 2011]. MS analysis was conducted successfully by searching peptides on a database. Key to success was to do tryptic digestion and get the peptides to fly and to be identified. Once dried the array was placed on the MALDI target and analyzed. The analysis was performed with an Autoflex MALDI-TOF mass spectrometer (Bruker Daltonics, Leipzig, Germany) operating in linear and reflector mode.

The resulting mass accuracy for peptides was b20 ppm. MALDI-TOF mass spectra were acquired with a pulsed nitrogen laser (337 nm) in positive ion mode, using two software programs to acquire and process mass spectra: FlexControl, designed to configure and to operate time of flight mass spectrometer of the Bruker's flex-series, and FlexAnalysis, designed for data analysis of spectra acquired with Bruker's TOF MS.

We acquired spectra for each sample (P53, JUN, CdK2, CdKN1A, A, B, C, D), and in order to identify the A, B, C and D samples we matched, with the aid of a “matching algorithm” implemented by us and described in [Spera et al., 2011], their experimental mass lists with that of the known samples (P53, JUN, CdK2, CdKN1A). For the successful protein fingerprint analysis, as shown in *Fig. 20* for p53, we utilized a specific tool, Biotools (Bruker Daltonics, Leipzig, Germany), that allowed an automated protein

identification via library search with fully integrated MASCOT software (Matrix Sciences, Ltd. www.matrixscience.com).

We can then conservatively conclude that the implemented chemistry and analysis for the first time demonstrate the successful use of mass spectrometry for the characterization of proteins immobilized on Nucleic Acid Programmable Protein Arrays. Further development is in progress to bring this label-free procedure to practice as an adjunct to fluorescence NAPPA work, which has already seen significant clinical applications in the last decade [Anderson et al., 2011]. The background generated by the reticulocytes lysate is however still significant and need to be reduced to make this approach routinely applicable in the clinics. This reduction might be instead achieved by the use of a bacterial cell-free expression system with respect to the traditional mammalian lysate [Pechkova et al., 2010a; Pechkova et al., 2010b], particularly required by the highly sensitive nanotechnologies being here utilized. The application of bacterial PURExpress to NAPPA (work in progress) consists of a template double stranded DNA (*dsDNA*) containing the gene of interest fused to a SNAP tag and the upstream T7 promoter. By adding the PURExpress reconstituted cell-free translation system, the template DNA is transcribed into mRNA, and then translated into a fusion protein containing the N-terminal SNAP tag and the C-terminal target protein. In the same spot, the SNAP tag allows the synthesized protein to bind to its own template DNA via the BG linkage, thus immobilizes the target protein. The rest of the reaction mixture can be washed away and the immobilized target protein is allowed to interact with a mixture of query proteins. After the binding reaction, the unbound proteins

are washed away and the target protein complex is released by cleaving the template DNA.

To compare the backgrounds of NAPPA between PURExpress (bacterial lysate) and RRL (rabbit reticulocyte lysate) by mass spectrometry and fluorescence we are presently utilizing NEB *in vitro* system and SNAP fusion as an alternative to *in vitro* system and GST tag. Their advantages are higher expression level and cleaner for downstream analysis, making possible and really effective label free quantitative analysis at the nanoscale (work in progress in cooperation with the Arizona State University, ASU, and New England Biolabs, NEB).

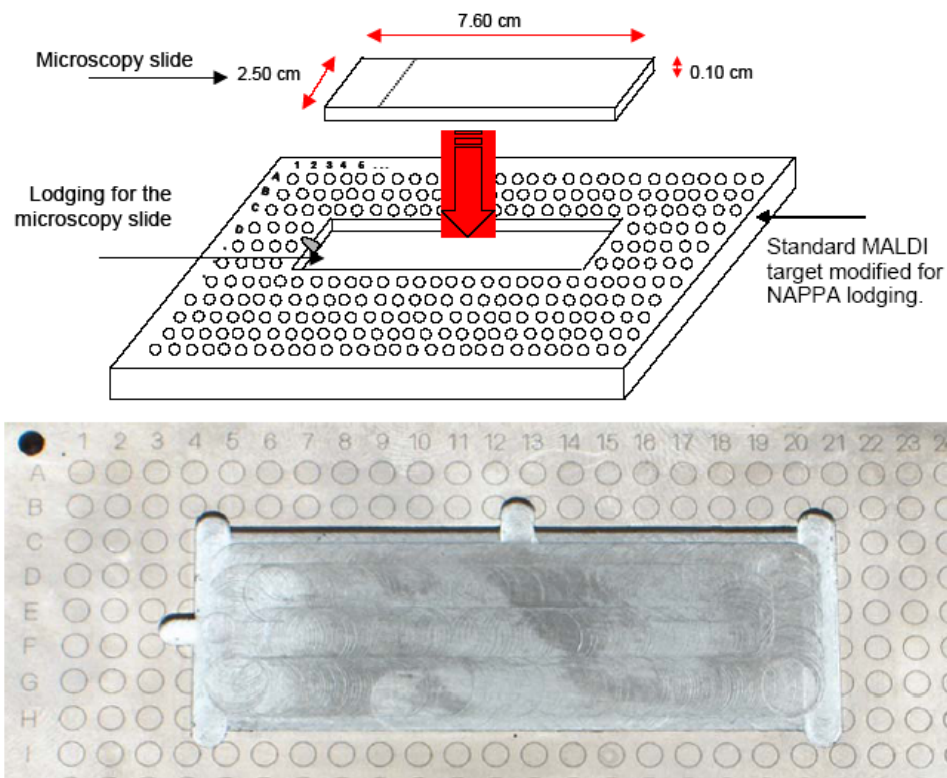


Figure 19: MS MALDI standard target adapted to carry NAPPA. From [Nicolini, Bragazzi and Pechkova, 2013].

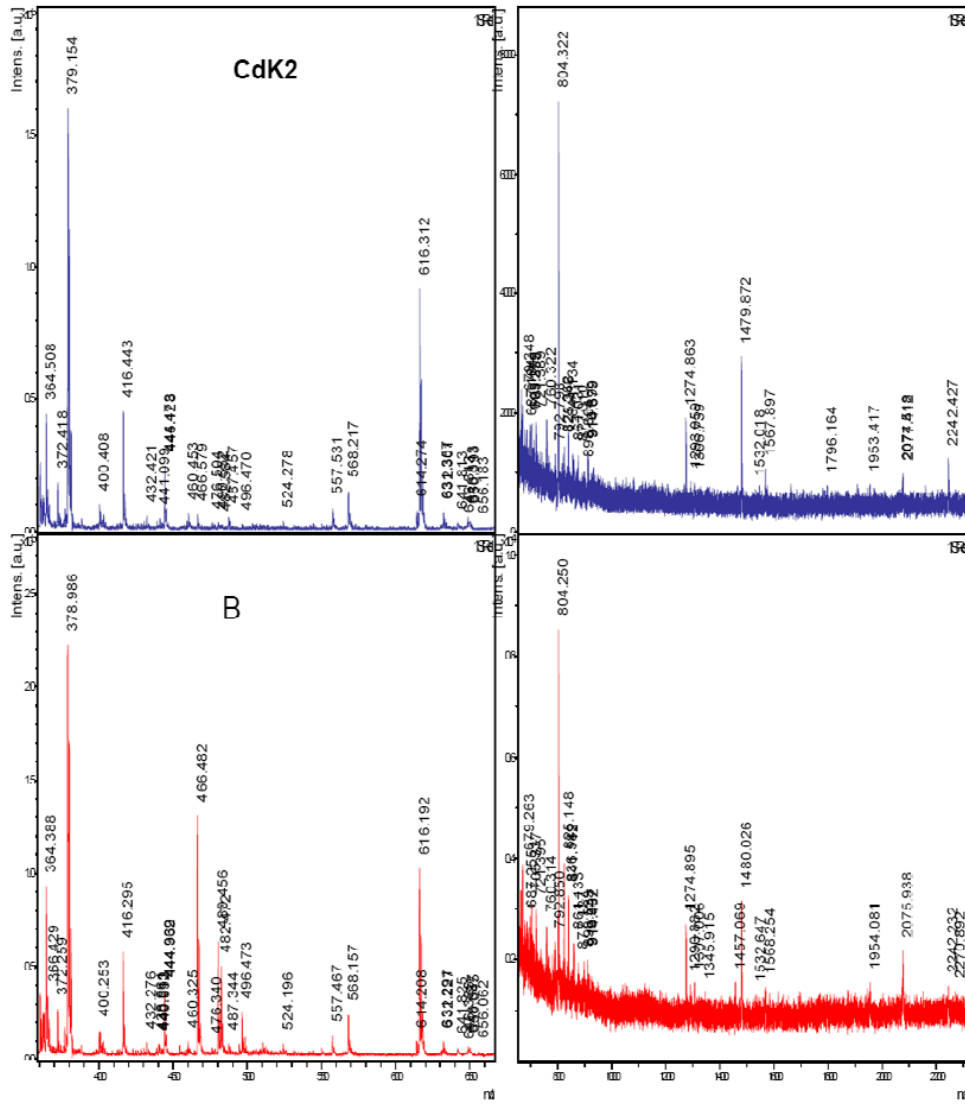


Figure 20: MALDI TOF Spectra of NAPPA after protein tryptic digestion, 5–20 kDa range, for optimal identification of single unknown genes. From [Nicolini, Bragazzi and Pechkova, 2013].

6. MEDICAL APPLICATIONS

6.1.A REVIEW AND A STATE-OF-ART OF THE LITERATURE

In Personalized Nanomedicine protein arrays are primarily employed to identify new molecular targets in toxicology studies, for the analysis of clinical diagnostics to monitor the evolution of a given disease (through studies of drug efficacy) and for pharmacokinetics. In the context of basic research, however, protein arrays are used for the analysis of receptor–ligand, protein–protein and protein–DNA interactions and for the analysis of post-translational modifications analysis.

Label-free nanoproteomics can then be used either as discovery tool with thousands of genes in very few samples for screening properties [Collings and Vaidya, 2008; Gold et al., 2006; Madu and Lu, 2010; Umehara et al., 2009; Wang, 2005] or for prognosis/diagnosis tool [Chikkaveeraiah et al., 2011; Hueber and Robinson, 2006; Uetz et al., 2000] but not yet for therapy with single few genes in many samples for various types of cancer having those genes uncommon. This represents our major objective for the future. Indeed even if presently adapting fluorescent NAPPA to luminex appears to have “*per se*” a strong clinical impact in diagnosis, only label-free nanobiotechnologies can overcome the limitations of the fluorescence-based approaches for mass proteomics. The three nanotechnological approaches proposed in the protocol for Personalized Nanomedicine can be used in an independent and complementary fashion, with occasional integration of fluorescent labeling utilizing DNASER and APA.

Recently the high relevance of NAPPA to medicine, namely to cancer, has been strikingly proven by a large cooperation among leading American

Chapter 6. Medical applications

institutions and hospitals, and at the same time a complete body of novel work with label free NAPPA, has seen the press with unique experimental data here presented for the first time. Cancer patients indeed spontaneously generate autoantibodies (AAb) or tumor-derived proteins and to detect them the recent novel NAPPA has been successfully developed expressing 4988 candidate tumor antigens with sera from patients with early stage breast cancer (IBC), and bound IgG was measured.

Using a three-phase serial screening approach NAPPA array proved capable to develop potential biomarkers for the early detection of breast cancer [Anderson et al., 2011]. Label-free sensing technologies have many other advantages that could be exploited in new developments capable to identify clinically relevant proteins without the need for fluorescent tags.

Novel high-density custom NAPPA, which was recently used to detect these autoantibodies in sera from patients with early stage breast cancer, represents the first demonstration that the use of programmable protein microarrays has strong medical implication and high relevance for the proteomic detection of novel disease biomarkers. While these studies were done using fluorescent molecular tags, our label free approaches simplify signal detection and provide improvements in sensitivity, reproducibility, dynamic range, and ease of use. To date, two major techniques have been used for the HT analysis of protein–protein interactions: yeast two-hybrid [Chikkaveeraiah et al., 2011; Hueber and Robinson, 2006] and immunoprecipitation followed by mass spectrometry.

A key advantage of the mass spectrometry method described here is the ability to identify new ligands through fingerprinting without the need to use antibodies. Nanogravimetry for the dynamic flow characterization warrants

Chapter 6. Medical applications

high screening, and Anodic Porous Allumina for the high throughput at the highest gene and protein densities. Among the label free detection methods, mass spectrometry, APA and QMC_D have the unique advantage of being able to determine with HT not only the presence but also the identification of a given ligand without the introduction of extraneous molecules. Other needs, not HT, appear to be properly addressed by Atomic Force Microscopy for geometric static determination of the gene expressed protein and its complex formation with the query protein. A new integrated approach for medical diagnostics and therapy, termed nanoproteomics [Nicolini and Pechkova, 2010], is thereby emerging also from the interplay between protein nanocrystallography and Nucleic Acid Programmable Protein Arrays (NAPPAs) on a previously unseen scale.

In the clinics the challenges that are still open are systemic inflammatory diseases such as rheumatoid arthritis and aggressive cancers with the worst prognosis and fast metastasis, such as pancreatic ductal adenocarcinoma.

Unfortunately nanotechniques still face several limitations to be resolved for a wide spread application in biomarker discovery, but currently new nanoproteomics is being progressively adopted by clinical researchers due to increased knowledge of the pathophysiology of unresolved diseases. An example actively pursued is the real-time label-free detection of protein interactions in high-throughput format by NAPPA-SPRi, where the indicated proteins having C-terminal GST *E. coli* tags were expressed according to the NAPPA method, and placed on the SPRi instrument and the CCD recording initiated. The proteins were probed with anti-GST monoclonal antibody at the time indicated and the individual sensorgrams were identified with protein-specific antibodies. Using SPRi when purified

Chapter 6. Medical applications

recombinant Fos protein was used as query to probe a NAPPA array the binding to Jun was monitored, and when purified recombinant protein MDM2 was used as query to prove the binding to p53 [Pechkova et al., 2010]. Nanobiocrystallography via LB is a promising technique for crystallography already proven useful for drug design and nanotubes and fullerene (simulations using GROMACS software) can represent a promising approach for drug delivery [Bozdaganyan et al., 2014; Garibaldi et al., 2006].

Nanogenomics and nanoproteomics allow us to study and compare the huge amount of genes and proteins involved in the cell-cycle progression of human T lymphocytes and in its transformation in lymphoma. However, nanogenomics suffers many pitfalls that only functional proteomics, in the form of NAPPA, can overcome, probing the native in situ protein–protein interaction with unique sensitivity. This facilitates the identification of the key proteins involved in the control of cell function and enables their further characterization by structural proteomics at the atomic level using synchrotron radiation. We believe that only the interplay of new genomics, proteomics and other high-throughput tools as well as of nanotechnologies (as discussed in [Dasilva et al., 2012; Gonzalez-Gonzalez et al., 2012; Ray et al., 2011; Walhout et al., 2000; Zhu et al., 2001], with clinical correlations and examples) can be the successful key to personalized medicine. It is in fact the combination of protein nanocrystallography and NAPPA technologies that provides the focus of nanoproteomics for Personalized Medicine.

6.2. APPLICATIONS IN THE FIELD OF NANONEPHROLOGY AND KIDNEY TRANSPLANT

In the broader field of nanomedicine, nanonephrology [Jain, 2008] is a superspecialized branch, indicating a new approach to the kidney pathologies. It includes a comprehensive molecular and systems biology knowledge of the kidney genes and genes products up to the atomic level, along with a nano-imaging based diagnostics in order to study cellular processes and events in the kidney cells [D'Agati and Mengel, 2013].

Moreover, it includes all those novel medical treatments that utilize nanoparticles and to treat various kidney diseases, the creation and use of nanobiomaterials and devices at the molecular and atomic levels that can be used for the diagnosis and therapy of renal diseases.

An interesting application of Nanonephrology is the nanobiotechnology-based dialysis filters: human nephron filters (HNFs), blood compatible nanoporous filters, ceramic filters able to reduce the time of the treatment.

Chronic kidney disease (CKD) is an important public health problem with a tremendous burden and social-economical implications, affecting only in the USA more than 20 million people and with an expected increasing prevalence in the near future.

It is a condition in which the kidneys damage leads to a loss of their physiological functions (filtering blood, regulating the pressure, maintaining acids and bases homeostasis, regulating fluids and electrolytes). Thus, wastes are not properly removed, their level tends to increase and this leads to other complex health problems and comorbidities, including above all complications such as cardiovascular disease (CVD), hypertension, anemia, malnourishment (in particular, protein-energy malnutrition, or PEM),

Chapter 6. Medical applications

endocrinological disorders (impaired balance of electrolytes, erythropoietin deficiency, secondary hyperparathyroidism, hyperinsulinism, growth hormone impairment, sexual dysfunctions), bone diseases (osteoporosis, osteomalacia, osteitis fibrosa, osteodystrophy, adynamic bone and other bone abnormalities that are included in the complex clinical syndrome termed as renal bone disease, or renal osteodystrophy, or CKD-mineral and bone disorder), infections (especially, hepatitis C), cancer, and mental disorders like dementia and cognitive impairment.

People with early CKD tend not to feel any kind of symptom and the only ways to detect CKD as earlier as possible are through a blood test to estimate kidney function, and a urine test to quantitatively assess the entity of kidney damage. For this reason, CKD is often diagnosed when patients need replacement therapy (that is to say, kidney transplantation, chronic haemodialysis or peritoneal dialysis), since CKD is usually an irreversible and progressive disease and can lead to kidney failure, also called CRF (Chronic Renal Failure), whose last stage is termed End Stage Renal Disease (ESRD), which is fatal if it is not properly treated. Once detected, CKD can be managed through medication and lifestyle changes to slow down the disease progression, and to prevent or delay the onset of kidney failure as much as possible.

CKD is more common among women than men, more than 35% of people aged 20 years or older with diabetes have CKD, more than 20% of people aged 20 years or older with hypertension have CKD.

The state of operational tolerance in humans has been detected sporadically in some renal transplanted patients who stopped immunosuppressive drugs and retained a good renal function, demonstrating that allograft tolerance

Chapter 6. Medical applications

might exist [for a review, kidney is less susceptible to successful immunosuppressive drug withdrawal than the liver [Orlando et al., 2009; Lozano et al., 2011], there is now mounting evidence showing that kidney transplant recipients too can become operationally tolerant [Roussey-Kesler et al., 2006; Newell et al., 2010; Sagoo et al., 2010]. In search of biological signatures of “operational tolerance” in humans, we previously identified a list of 49 genes which were able to discriminate operationally tolerant patients from other cohorts of transplant patients and healthy individuals [Brouard et al., 2007].

The gene SMILE/TMTC3 was one of the genes found to be upregulated in the blood of operationally tolerant patients and whose function was unknown [Racapé et al., 2011]. SMILE was found to interact with PDIA3, which has a crucial role in glycoprotein folding in endoplasmic reticulum (ER) [Zhang and Kaufman, 2008], in the loading of peptide on MHC class I [Santos et al., 2007] and which is over-expressed during ER stress. We showed that siRNA-mediated SMILE knock-down in HeLa cells induces a decrease in several types of transcripts involved in protein catabolism and proteolysis, particularly immunoproteasome subunits, suggesting that SMILE exerts its function via the proteasome pathway [Racapé et al., 2011]. This was confirmed by the facts that SMILE down-regulation and/or treatment with proteasome inhibitor (Bortezomib) induced dramatic ER enlargement and features of cellular injury and that Bortezomib inhibition of long-term cellular growth was strongly enhanced in SMILE siRNA-transfected cells. SMILE silencing was found to directly increase XBP-1 transcript expression after 6h of Bortezomib treatment, while DNA

Chapter 6. Medical applications

microarray analysis revealed that SMILE down-regulation in HeLa cells affects secretory pathways, such as vesicle-mediated transport.

Earlier, we outlined a microarray-based identification of key leader genes found respectively in blood from kidney transplanted patients with chronic rejection or operational tolerance by utilizing a non-statistical bioinformatics approach based on the identification of “key genes” either as those mostly changing their expression, or having the strongest interconnections [Sivozhelezov et al., 2008].

An informative picture emerges on the genes controlling the human transplant from the detailed comparison of these findings with the traditional statistical significance analysis of microarrays (SAM) and with the clinical study [Braud et al., 2008]. Some of these genes, such as BANK-1, a modulator of B-cell hyperactivation, have been confirmed through further studies using different techniques [Pallier et al., 2010]. In parallel, in another study, in a rodent model of kidney allograft, DNA microarrays were used to compare gene patterns in the kidney transplant from anti-donor anti-class II tolerated or untreated syngeneic rat in the Lewis 1W and Lewis 1A rat strain combinations [Jovanovic et al., 2010]. Statistical and nonstatistical analysis combined with ab initio analysis, using the leader gene approach [Bragazzi et al., 2011] suggested that tolerance and rejection outcomes may be in large part correlated to the low expression variations of some genes, which can form a core gene network around specific genes, such as Rac1, NFKB1, RelA, AKT1, IKBKB, BCL2, BCLX, and CHUK. Through this model, we showed that AKT1 gene, WNT pathway and NO synthesis were strictly connected to each other, with AKT1 gene being the center of this

Chapter 6. Medical applications

complex network of interactions in kidney tolerance and rejection processes [Jovanovic et al., 2010].

The aim of the present study was to define networks of genes either modified by SMILE knock-down compared to control knockdown cells or phorbol 12-myristate 13-acetate (PMA)-activated cells compared to non-activated cells. Then, we compared these two networks with those found in our microarray studies in human [Sivozhelezov et al., 2008] and rat [Jovanovic et al., 2010]. Gene expression microarray data representing three independent experiments from HeLa cells transfected 24 h with negative control or SMILE siRNA, or activated or not with 20mM PMA during 6 h as previously described [Racapé et al., 2011] were submitted for our analysis.

For each analysis, we arbitrarily fixed the FDR less than 0.5%. LOWESS normalization was applied, we made a first visual inspection using the Volcano Plot and its variants. Interaction networks from the genes involved in SMILE knockdown selected by SAM analysis (or alternatively Fisher Z-score) resulted in large intractable networks. It is well-known that different statistical analysis can yield different results [Kim et al., 2006] and a need for combining statistical and biological significance is emerging [Draghici et al., 2007].

We performed nonstatistical analysis on the different conditions of culture based mostly on fold-changes [Racapé et al., 2011]. The columns containing d-scores and fold-changes were thus multiplied to select both statistically and biologically significant genes. We considered genes with both statistical and biological significant changes in expression: we performed a scatter-plot with the number of genes versus SAM d-score and log of fold-change

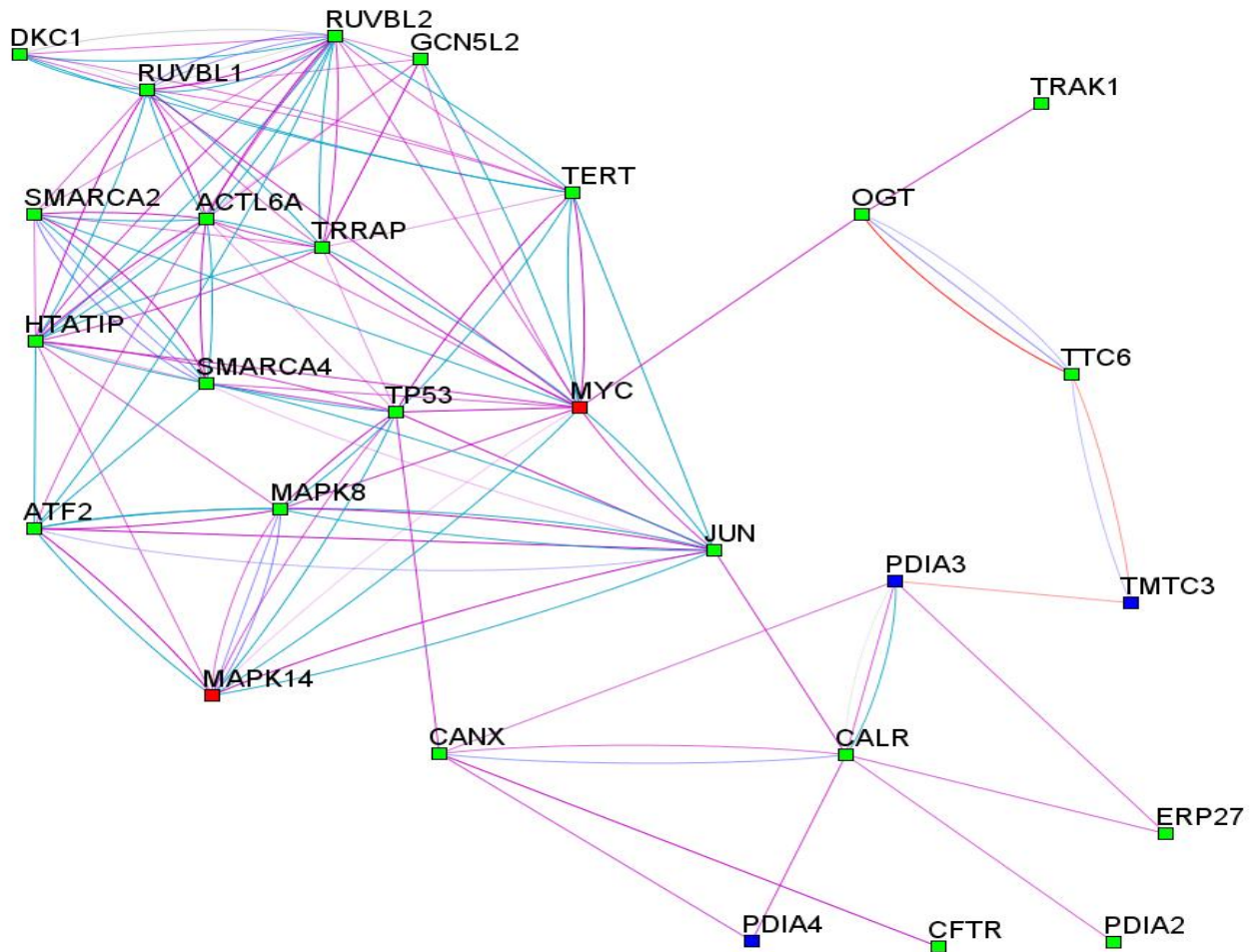


Figure 22: semi-theoretical network of SMILE silencing.

Networks have been proved a very useful tool for understanding cell functions [Barabasi and Oltvai, 2004] and in the frame of the systems

biology they have been widely exploited for characterizing different cellular processes, among which also kidney allograft [see for a review, Perkins et al., 2011]. As far as we know, this is the first time that topological and mathematical analysis have been used to compare the changes of two biological networks after a gene has been knocked-down.

Genes selected with the SAM analysis form no visible physical interaction networks after SMILE silencing (Fig. 21). Therefore, we formed a semi-theoretical interaction network starting from the known interaction of SMILE with PDIA3 and added their neighbors based on family-wide and cross-species analogs (COG mode of STRING). The resulting map is shown in Figure 22. Genes significantly up-regulated are shown in red, down-regulated in blue, not changing their expression in green. Magenta lines denote physically observed interactions between the proteins encoded by the respective genes, cyan lines neighborhood in known regulatory or metabolic pathways. Using this method, we confirmed that indirect neighbors of TMTC3 are related to ER functioning and stress response genes (CANR, CALX, PDIA4, PDIA2, and ERP27) due to TMTC3 interaction with PDIA3 [Racapé et al., 2011].

In contrast to SMILE silencing, gene highlighted from the SAM analysis resulted in a well-defined network following PMA activation. This is likely due to extensive studies on PMA activation, leading to a more extended knowledge of genes involved in this activation than those involved in SMILE knock-down. The two networks, semi-theoretical for SMILE silencing and expression-derived for PMA activation, were compared to the “pro-tolerance” and “pro-rejection” networks inferred from our previous studies in blood from patients with a kidney transplant [Braud et al., 2008;

Sivozhelezov et al., 2008]. SMILE gene (TMTC3) is linked via PDIA3 to the network previously identified in the blood from human transplanted kidney recipients with a profile of chronic rejection (referred above as the pro-rejection network): namely STAT3, TNFRSF1A, IL2RB, LCK, VAV1, HNRPU, and their neighbors, as described in more details by Sivozhelezov and collaborators [Sivozhelezov et al., 2008]. SMILE gene relation to the network identified in the blood from human transplanted kidney recipients with a profile of operational tolerance (referred above as the pro-tolerance network) is less direct, although one gene, IFN γ is linked to SMILE again via the PDIA3 gene. There are five overlapping genes between SMILE knock-down network and the pro-tolerance and pro-rejection networks (four genes pro-tolerance: TP53, ATF2, JUN, and MAPK14 and one gene pro-rejection: HTATIP). We found that those five genes form their own network. For PMA activation and the human kidney transplantation, there is a sub-network of PMA activation most tightly connected to the pro-tolerance network, namely ERBB2, CD44, and IL4R.

There is the same situation for the subnetwork formed by genes RELA, NFKB1, and NFKBIA in comparison with the pro-rejection network.

Before performing our topological analysis, we verified the scale-free behavior of our networks, since this feature has emerged as a peculiar and distinguishing characteristic of biological networks [Barabasi et al., 2001].

From the “Density,” the “Mean connectivity” and the “Clustering coefficient” rows of Table 3, it appears that SMILE networks are denser and more clustered than PMA networks, each SMILE node having more neighbors than PMA. In contrast, PMA networks are more heterogeneous than SMILE networks. The centrality (tendency to form star-like rather than

Chapter 6. Medical applications

mesh-like topology) is small for all the networks (maximum is 1), which is typical of biological networks that rarely have a single superhub.

Finally, size of the network is higher for PMA networks. Other parameters (like diameter and degree) show little statistical variance or show an ambiguous trend (namely, degree is higher for PMA in the case of tolerance network, while being lower for PMA in the case of rejection).

Frequency distributions of node connectivity for rejection network SMILE knock-down versus rejection network PMA activation showing statistically significant difference are also analyzed. Hub genes occur more frequently in rejection network SMILE knock-down than in rejection network PMA activation. Instead, frequencies for tolerance network SMILE knock-down versus tolerance network PMA activation show no statistically significant difference. We conclude that SMILE knock-down more significantly affects the kidney transplant rejection network than PMA activation, while the tolerance network is affected to the same extent by SMILE knock-down and PMA activation (*Table 3*).

Parameters, samples	Tolerance SMILE	Tolerance PMA	Rejection SMILE	Rejection PMA	Rat SMILE	Rat PMA
Nodes	79	92	78	83	52	61
Density	0.06	0.53	0.64	0.47	0.87	0.05
Heterogeneity	0.83	0.93	0.69	0.76	0.64	0.69
Clustering coefficient	0.3	0.26	0.3	0.25	0.4	0.27
Diameter	8	8	8	9	6	9
Centrality index	0.17	0.16	0.12	0.14	0.17	0.14
Number of neighbors	4.86	4.85	4.95	3.83	4.46	3.02
Size	2.5	2.65	2.54	3.12	2.24	3.19
Mean connectivity	4.7	4.76	4.99	3.75	4.45	2.98
Degree	11.27	13.36	11.81	9.93	10.31	7.94
FFL subgraphs	20.00%	19.50%	17.00%	16.30%	18.20%	13.7%

Table 3: topology Parameters of SMILE-silencing and PMA-activation Networks.

As far as the local topology is concerned, we studied the small motifs composing the subgraphs of the network—using a well established approach, comparing them against motifs in random networks and selecting only those motifs for which Z-score was higher than 2, whose P-value was lower than 0.05 and whose frequency was higher than five times. Our study showed that FFLs motifs [Mangan and Alon, 2003] are the core motifs of our networks and they are abundant in SMILE networks, suggesting that SMILE calls for more fine-tuned regulation. For the rat kidney transplant study, too few genes notably change expression to analyze pro-rejection or

pro-tolerance network separately. Besides, the setup of the microarray analysis (TOL/SYN and TOL/REJ) would render such analysis of separate networks biased. Therefore, we merged the tolerance and rejection gene lists.

There is a sub-network of rat kidney transplant related genes interacting strongly with the SMILE network modulation (Fig. 8a, underlined in dotted line). These include RelA, CHUK, NFKB-1, NFKB-1A, and AKT1, having a significantly high number of links with SMILE sub-network. The same observation is even more pronounced for PMA activation. Three of the genes mentioned above (NFKB-1, NFKB-1A, and RelA) even constitute an overlap between the rat transplant network and the PMA activation network. Topological analysis of joint SMILE knock-down and PMA activation with tolerance and rejection networks from the rat data from Table 3, also allowed to conclude that SMILE networks are topologically quite different than PMA activation networks. SMILE networks are denser, more clustered and connected, while being also more distributed. In contrast, PMA networks are less dense and more localized around specific genes.

Node connectivity also shows significant difference as for the human data. Comparison graphs for average connectivities and clustering coefficient follow the same pattern.

Considering that the rat networks are largely dominated by prorejection genes [Jovanovic et al., 2010], this leads to the same conclusion that SMILE knock-down more significantly affects the kidney transplant rejection network than PMA activation. Again from local topology analysis, we can see that SMILE network contains more FFL motifs than PMA network and this difference is even more strikingly true for rat compared to humans. In

this study, a microarray study of SMILE knock-down and PMA activation in HeLa cells was compared to our earlier analysis based on microarray data of kidney allograft tolerance and rejection in humans and in a rat model of heart allograft transplantation to determine possible new genes and gene networks involved in kidney transplantation. Based on bioinformatics analysis, we showed that at the intersection of the SMILE silencing gene network and earlier identified microarray-based pro-rejection versus pro-tolerance gene networks, five expression leader genes appear that form their own sub-network. These genes encode one transcription regulator, namely the kinase MAPK14, three transcription factors (JUN, ATF2, and TP53), and a histone acetyltransferase (HTATIP). This network has been expanded via STRING to its nearest neighbors.

Also interesting are the two immediate neighbors of this subnetwork, Nephrin and Arrestin b2. Nephrin is a member of the immunoglobulin family of cell adhesion molecules that function in the glomerular filtration barrier in the kidney, being the structural component of the glomerular slit diaphragm [Ruotsalainen et al., 1999]. Arrestin b2 is known to have a significantly reduced expression in monocytes during kidney graft rejection as it has been recently demonstrated in human [Zakrzewicz et al., 2011]. Nephrin fosters activation of stress-activated protein kinase 38 and JUN which acts in complex with FOS [Huber et al., 2001]. In competition with the binding of Nephrin to the podocin, Arrestin b2 mediates Nephrin endocytosis and therefore its functioning reduction [Quack et al., 2006]. Note that the Arrestin b2 activation is via MAP kinase and down-regulates the TGF- β signaling pathway [Shenoy and Lefkovitz, 2011]. But most importantly, the network includes the protein XBP-1 already reported by

Racape' et al. [2011] to be up-regulated in response to SMILE down-regulation, reinforcing these previous findings. FOS together with JUN forms a functional heterodimer with XBP-1 [Ono et al., 1991].

The comparison between the PMA activation network and the pro-tolerance network revealed that visually, the pro-tolerance network appears to be more tightly connected with the PMA activation network than the SMILE knock-down network, suggesting that tolerance may be associated to an activation profile. This will have, however, to be proven numerically, for example, by modifying our Leader Gene software [Bragazzi et al., 2011] to assign scores to networks and not to individual genes.

The other finding is that SMILE knock-down networks, irrespective of the context, show topological features strikingly different from the PMA activation networks. Indeed, analysis of the topological parameters of SMILE knock-down and PMA activation with transplantation networks in human shows increased density and clustering coefficient (that measures the degree to which nodes in a graph tend to cluster together, for example, average number of links between each gene's neighbors) of SMILE networks, meaning that SMILE networks are better connected and more clustered than PMA networks. Denser networks should be expected for SMILE because SMILE is a single gene while PMA activation may simultaneously activate or suppress many genes possibly unlinked to each other. In contrast, heterogeneity of PMA networks compared to SMILE networks suggests that there is a larger fraction of higher connected (hub) genes in PMA than in SMILE networks. In other words, SMILE knock-down leads to more evenly distributed effects on tolerance/rejection gene networks while PMA activation focuses on fewer but more highly

Chapter 6. Medical applications

connected genes. Finally, PMA networks display a bigger size of the networks (higher number of links). Larger number of links, however, does not mean better connection because links may be members of (partially) isolated subnetworks.

Another interesting observation is that SMILE networks contain more FFL motifs than PMA which means a more sophisticated genetic regulation. Interestingly, genes were found to be common between the two analyzed network combinations, SMILE knockdown network rat transplant network and PMA activation network rat transplant network (NFKB-1, NFKB-IA, and RelA). This allows concluding that common molecular mechanisms are involved in all three phenomena or that they are linked together through the same network. Besides, these results confirm our earlier finding on AKT1 gene in rat kidney transplant tolerance [Jovanovic et al., 2010], and this could be better investigated in further bioinformatics studies.

Subsequently, we validated the list of found genes, namely HTATIP/KAT5, C-JUN, TP53, ATF2, MAPK14, ARRB2, XBP1, and NPHS1.

We accessed Gene Expression Omnibus (GEO), a public repository of micro-array experiments, looking for the expression profiles of the previously identified Leader Genes.

In case of different existing experiments studying the expression of the chosen gene, the data were normalized and then pooled together. Student's t-test for independent samples, Analysis of Variance (ANOVA)-one way and Pearson's correlation between gene expression in kidney and in peripheral blood were calculated with SPSS software V21.0.0 package (IBM) and using R environment. p-values were computed with SPSS, and values equal or less than 0.05 were considered statistically significant.

Chapter 6. Medical applications

We found that 5 out of 8 genes (62.5%) are able to distinguish between rejection and tolerance to kidney transplantation, being differentially expressed between the two groups of patients in a statistically significant way. Some of these genes have been rarely described as predictors of clinical outcomes to renal graft in the literature.

Pooled Gene expression profiles analysis revealed that HTATIP/KAT is able to distinguish between rejection and tolerance outcome, being differentially expressed between the two groups of patients with a p-value <0.001, both in kidney and in peripheral blood. Similar results were obtained with TP53 (p-value <0.001), C-JUN (p-value <0.001 in blood, p-value <0.05 in kidney). ARRB2 yielded a statistically significant result only in kidney (p-value <0.05), such as ATF2 (p-value <0.05) (Figures 22 and 23).

On the contrary, MAPK14, XBP1 and NPHS1 expression profiles did not differ in a statistically significant way.

The lack of statistical significance for these genes may be due to the limited existing number of micro-array experiments. Further experiments are needed to confirm and replicate our findings.

Correlation between expression profiles in peripheral blood and in kidney was also studied, using Pearson's correlation. Only ATF2 tolerance profile in kidney and in blood yielded statistically significant results ($r=-0.68$, p-value <0.001) (Figure 25).

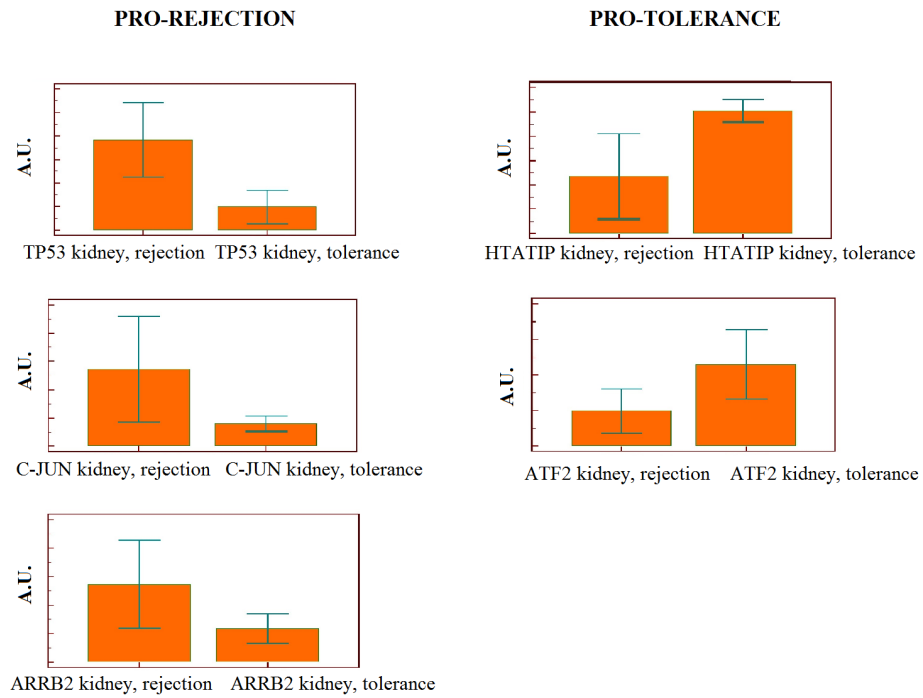


Figure 23: pro-rejection and pro-tolerance gene expression in kidney tissue. p -value <0.001 .

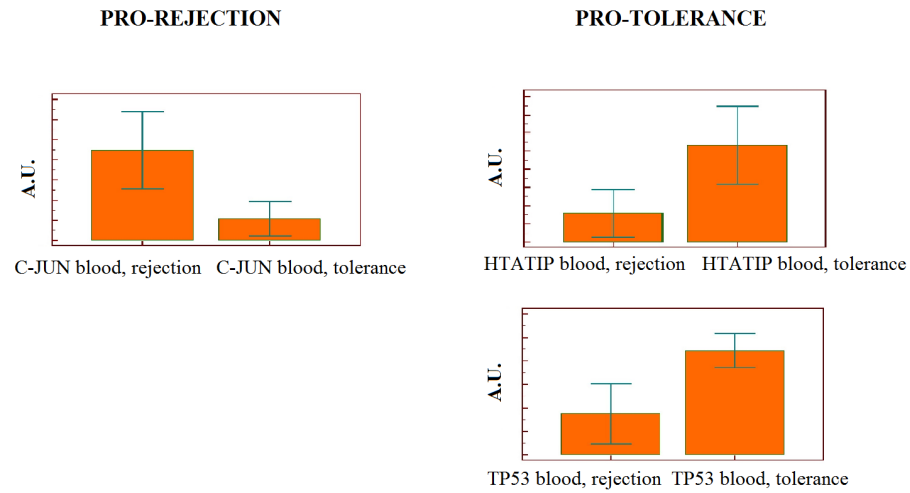


Figure 24: pro-rejection and pro-tolerance gene expression in blood. p -value <0.001 .

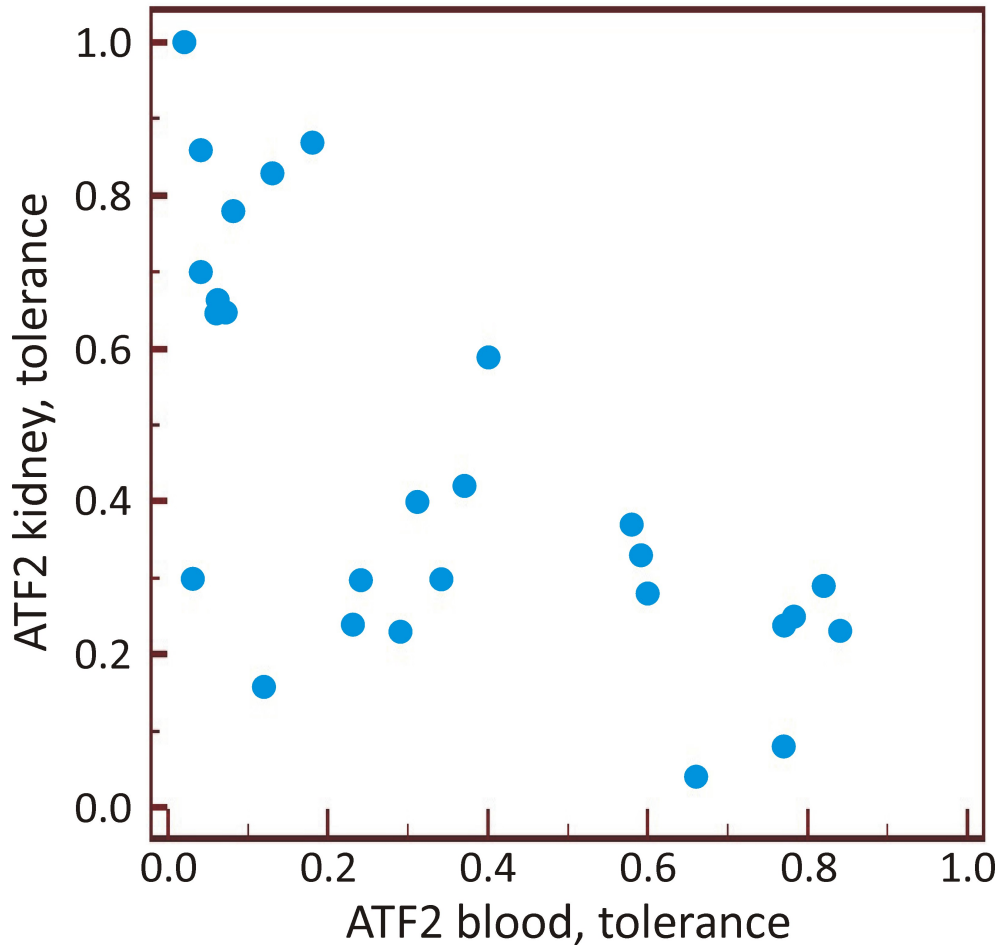


Figure 25: Pearson scatterplot showing correlation between ATF2 gene expression in blood and in kidney tissue.

6.3. APPLICATIONS IN THE FIELD OF ORAL DISEASES

Oral Lichen Planus (OLP), first identified by Doctor Erasmus Wilson in 1869, is a chronic inflammatory oral mucosal disease affecting about 0.5–2% of the world population, more often females, with a mean age-of-onset usually in the fourth decade [Anuradha et al., 2008]. The histopathologic features of OLP are hydropic degeneration of the basal cell layer, hyperkeratosis, acanthosis, irregular ridges and band-like infiltration of T lymphocytes mainly in the *lamina propria*. Although the aetiology of OLP is still unknown, it has been widely accepted that immunological aberrances are very critical among the multiple aetiological factors. Previous studies have suggested that it represents a cell-mediated immunologic response to an induced antigenic change in the oral mucosa [Ismail et al., 2007; Mollaoglu, 2000; Sugerman and Savage, 2002]. The inflammatory response in OLP is characterized by the accumulation and expansion of T helper (Th) lymphocytes.

Clinical features and accompanying symptoms of OLP lesions range from asymptomatic reticular white lesions in atrophic mucosa to erosive-ulcerative areas accompanied with pain and discomfort. The histological lesion is characterized by a sub-epithelial inflammatory infiltrate, composed of different mononuclear cells, dominated by T lymphocytes, and limited to the basal keratinocyte layer [Epstein et al., 2003]. OLP lesions are characterized by a higher degree of cell turnover than healthy tissue [Scully and el-Kom, 1985]. In fact, the proliferation, the maturation and the apoptosis of basal keratinocytes require a fine regulation at a genomic level [Gandarillas, 2000; Hirota et al., 2002]. Moreover, the World Health Organization (WHO) defined OLP as a pre-malignant condition making

Chapter 6. Medical applications

“oral mucosa more sensitive to exogenous carcinogens and thus to develop oral carcinoma” [Pindborg et al., 1997]. However, the molecular mechanisms determining the possible development of cancer in OLP lesions are not well understood yet [Lodi et al., 2005a].

From these observations, OLP may be seen as a complex multi-factorial disease [Chiappelli et al., 2011; Hirschhorn, 2005; Lodi et al., 2005b; Orlando et al., 2013; Tabor et al., 2002].

Cluster analysis of the WNL identified 5 genes belonging to the highest cluster, i.e., the leader genes: JUN, EGFR, FOS, IL2, ITGB4. The analysis of variance (ANOVA) revealed a statistically significant difference in the WNL. In particular, the post hoc test revealed that leader genes had a significantly higher WNL when compared to class B genes (p-value < 0.001), and that class B genes differed significantly from other classes (p-value < 0.001 versus class C). The established or putative role of leader genes in OLP is summarized in *Table 3*.

Table 3– leader genes (JUN, FOS, EGFR, IL2, ITGBA2) and their molecular functions (modified from Wikipedia and NCBI annotation).

JUN The locus 1p32-p31, whose translocations and mutations cause several human malignancies, encodes an oncoprotein which, interacting directly with specific DNA sequences, plays a crucial role in gene expression.

EGFR This gene encodes a transmembrane glycoprotein that belongs to the protein kinase superfamily. It binds to epidermal growth factor (EGF), thus leading to cell proliferation.

FOS This gene encodes a leucine zipper protein that interacting with JUN forms the transcription factor complex Activator Protein type 1 (AP-1), implicated in cell proliferation, differentiation transformation, and apoptosis.

IL2 It encodes a cytokine that regulates the proliferation of T and B lymphocytes.

expression; in yellow: neutral genes in OLP disease; in red: up-regulated genes in OLP disease; in blue: down-regulated genes in OLP disease..

Topological analysis is summarized in Table 4. Interestingly all leader genes were widely distributed in the network (in term of topological parameters such as stress, eccentricity and radiality) and showed higher topological coefficients at the topological analysis.

OLP network parameter	Value
Clustering coefficient	0.33
Network centralization	0.07
Average number of neighbors	3.22
Network density	0.02
Shortest paths (in percentage)	52
Average path in length	4.66
FFL ^a	17.83
^a Feed-forward loops.	

Table 4: global topological analysis for OLP network.

The bibliographic research revealed that among the 5 genes identified as leader genes, only 3 were specifically associated with OLP. This search confirmed that every gene identified as a leader gene can be supposed to play a major role in OLP at a molecular level. The analysis of the interaction map allowed the identification of different groups of genes corresponding to the typical aspects of OLP lesion: changes of the epithelial basement,

regulation of cell cycle regulation and interleukine/chemokine signalling (see *Figure 27*) [Giacomelli et al., 2009; Lodi et al., 2005a; Lodi et al., 2005b; Oluwadara et al., 2009]. Interestingly, there is at least one leader gene for each area.

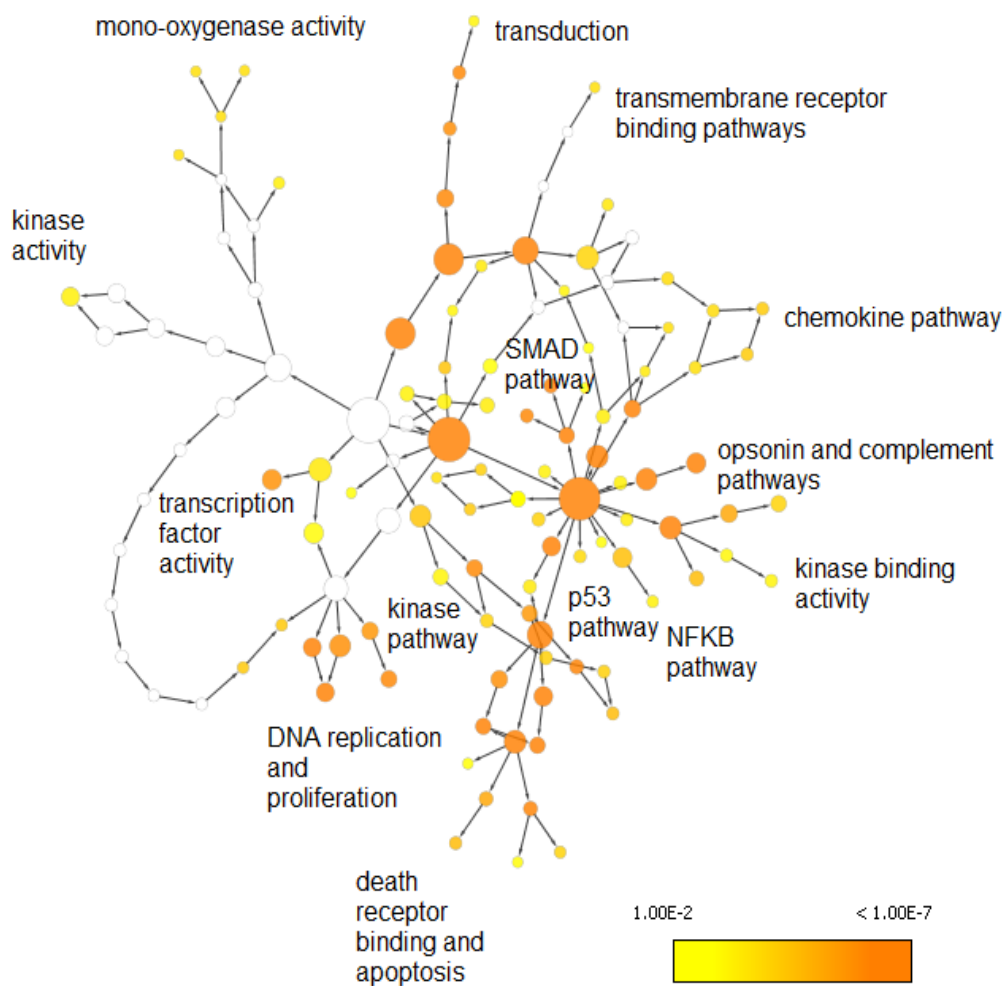


Figure 27: ontological analysis of OLP network showing the most important pathways over-represented in the graph versus whole set

annotation, carried out with BinGO software (p-value < 0.01, Benjamini-Hochberg correction, hypergeometric clustering).

Epithelial basement membrane changes are common in OLP and include cellular damage as well as the degeneration of the basal keratinocyte anchoring elements, which may result in a weakened epithelial-connective tissue interface and histological cleft formation (the so-called Max-Joseph space) [Ishimura et al., 2006; Karatsaidis et al., 2003; Zhou et al., 2001].

Noteworthy, it was suggested that dysregulation in integrin pathways and inhibition of the TGF β 1/Smad pathway could play a major role in the pathogenesis of OLP [Giannelli et al., 2006].

The α 4 integrin encoded by the gene ITGB4 has different functions, ranging from epithelial cell adhesion (by interacting with the basement membrane component laminin) to cell motility (interacting with F-actin) and also to immunity regulation. Reflecting this wide array of functions, ITGB4 has been involved in a variety of oral diseases (being a potential biomarker of the Warthin's tumour of the parotid gland, the tongue squamous cell carcinoma, the ameloblastoma and other oral cavity tumours, and playing as well a role in the Sjogren's syndrome disease and in most pre-malignant lesions, from oral leukoplakia to oral pemphigoid). Moreover, ITGB4 expression may reflect the response to a dental implant, determining the success of it together with osseointegration, a well-known parameter of the clinical outcome [An et al., 2012; Andreadis et al., 2005; Hamidi et al., 2000; Jones et al., 1995; Kurokawa et al., 2008; Modolo et al., 2004; Sami et al., 2002; Velozo et al., 2009].

Chapter 6. Medical applications

EGFR too plays a key role in OLP even if its precise function is still not understood [Ebrahimi et al., 2008; Kumagai et al., 2010; Martin-Ezquerria et al., 2010].

EGFR has been linked to other many oral pathologies and has been found to have a role in oral infections and to be as well a good biomarker for both the diagnosis and the prognosis of oropharyngeal and oral cavity squamous cell carcinomas, salivary gland cancer and other oral tumours. EGFR may play a key role in the malignant transformation of oral pre-malignant lesions and the contradictory experimental findings about its expression and regulation in OLP may reflect different clusters of OLP in term of prognosis, considering that a fraction of OLP lesion transforms into a cancer [Chandarana et al., 2012; Ettl et al., 2012a; Ettl et al., 2012b; Monteiro et al., 2012; Wali et al., 2012; Xia et al. 1999; Zhu et al., 2012]. Some scholars have proposed EGFR-targeted therapeutics as an emerging treatment for oral tumours [Wali et al., 2012], while other authors have suggested to use anti- EGFR monoclonal antibodies for monitoring oral cancers in vivo [Yang et al., 2011].

FOS and JUN control cell cycle regulation but their role in OLP is still very little known. Also these genes have been related to a variety of oral disorders, from oral infections to oral cancers, and interestingly they may play an important role in the transition from pre-malignant to malignant lesions [Gutierrez-Venegas et al., 2008; Mishra et al., 2010; Sachdev et al., 2008; Soejima et al., 2007; Turatti et al., 2005].

The third area in the interaction map is represented by interleukine and chemokine signalling and T lymphocytes/macrophages infiltration. There is

Chapter 6. Medical applications

a large evidence on a role for immune misregulation, specifically involving the cellular immune system [Simark-Mattsson et al., 1999].

IL2 has been found to play a major role in oral cancers, thus becoming an important drug target [Cheriyana et al., 2011; Covani et al., 2008; Khan et al., 2003]. It has been also linked with oral infections, periodontitis, oral autoimmune diseases and other pathologies [Gaffen and Liu, 2004; Reichert et al., 2009].

Noteworthy, at present no direct link at a genomic level is identified between this area and the basal membrane alteration and cell cycle control areas.

Our topological analysis showed that our network exhibits a power law behaviour in agreement with the Scale-free theory of bionetworks [Barabasi et al., 2001; Shimbela, 1953; Stelzl et al., 2005] and has more FFL (feed-forward loops) than one would expect to find in a random graph [Mangan and Alon, 2003]. The topological properties of leader genes and their role in controlling each pathway emerged from ontological analysis confirm our results and the linear regression shows that OLP-related genes have a higher disease-related connectivity than a general global connectivity (*Figure 28*).

Chapter 6. Medical applications

In the second part of our work we investigated the link between microRNAs and leader genes.

microRNAs (miRNAs) are a family of small and short (usually 19-25 nucleotides long), single-stranded, endogenous, non-coding RNA molecules. They intervene at a post-transcriptional level of the gene regulation and expression, binding with a certain degree of complementarity to the 3'-untranslated region (UTR) of different target mRNAs. This results in degradation of the target molecule via site cleavage or protein translation inhibition or perturbation, or both [Lee, 2013]. Because of their unique regulatory properties, miRNAs are supposed to play a major role in embryogenesis and throughout the different developmental stages, as well as in the regulation of the cell cycle and growth, proliferation, differentiation, and apoptosis. For this reason, they can be considered as important biomarkers and new potential therapeutic biotargets in oncology [De Cecco et al., 2013] and in pathology.

The role of miRNAs in oral diseases is emerging, since evidences are growing and accumulating in favor of this hypothesis. For example, miRNAs circulating in peripheral blood and in serum [Maclellan et al., 2012], as well tissutal and salivary miRNAs (Gallo et al., 2013) can be used as helpful biomarkers of early-stage oral cancer [Yoshizawa et al., 2013]. Salivary miRNAs are transported in vesicles called exosomes, that preserve their stability [Michael et al., 2010; Ogawa et al., 2013]. In particular, the interest in studying salivary microRNAome is increasing because of the non invasivity and effectiveness of the procedure, and also because of the advancement in salivary diagnostic technologies [Baum et al., 2011; Burbelo et al., 2012; Patel et al., 2011; Wong et al., 2011]. Several scholars

Chapter 6. Medical applications

maintain the importance of integrating it with other diagnostic alphabets (namely transcriptomics, proteomics, metabolomics and oral microbiomics), that constitute the five pillars of a new, recently emerging approach to oral diseases called “salivaomics” [Wong, 2012] and more generally speaking “oralomics” [Arrais et al., 2013; Rosa et al., 2012].

The importance of the omics studies in medicine and in dentistry is that they could enable a personalized treatment of the disorders [Eng et al., 2012; Garcia et al., 2013; Glick et al., 2012; Nicolini et al., 2012].

However, notwithstanding the importance of incorporating the oral microRNAome in the omics-based study of oral pathologies, so far only few miRNAs related to them have been discovered and described in the literature. Most published studies focus on oral cancer, while for example gingival and periodontal disorders are less studied.

The aim of the data-mining based approach that we introduce in this second part of our research dedicated to OLP is to accelerate and facilitate discoveries in the field of oral diseases.

We use our previously found Leader genes in order to predict putatively associated miRNAs, exploiting the miRGene database..

miRNAs indeed could act as important biomarkers of OLP.

Many biomarkers useful for predicting the evolution to cancer are emerging, including the pentaspan transmembrane glycoprotein CD133 (Sun et al., 2013), matrix metalloproteinases like MMP1, MMP2 (Chen et al., 2008; Sutinen et al., 1998), MMP7 and their inhibitors, TIMP1, TIMP2 (Chen et

Chapter 6. Medical applications

al., 2008; Sutinen et al., 1998), interleukins and cytokines such as IL1-alpha (Rhodus et al., 2005), IL6 and IL8 (Lisa Cheng et al., 2013; Rhodus et al., 2005).

Also apoptosis and cellular proliferation markers have been extensively investigated: the oncogenes p53 [Agha-Hosseini and Mirzaii-Dizgah, 2013; de Sousa et al., 2009; Ebrahimi et al., 2008, 2011; Leyva-Huerta et al., 2012; Ogmundsdóttir et al., 2009a; Ogmundsdóttir et al., 2009b] and c-erbB2 [Kilpi et al., 1996], CDK4, p16 [Poomsawat et al., 2011], BAX [de Sousa et al., 2009], BCL-2 [de Sousa et al., 2009; Leyva-Huerta et al., 2012], TNF and its receptor [Rhodus et al., 2005; Younes et al., 1996] and Proliferating-cell nuclear antigen or PCNA [Martinez-Lara et al., 1996].

Integrins and integrin-like proteins, growth factors are other molecules that could be involved in the process of carcinogenesis: alpha-9-integrin [Häkkinen et al., 1999], laminin-5 or kalinin [Kainulainen et al., 1997], endothelin-1 [Cheng et al., 2011], podoplanin and ABCG2 [Shi et al., 2010], desmocollin type 1 [Mattila et al., 2008], and TGF-beta-1 [Chen et al., 2008].

Other interesting and promising markers include COX2 [Li and Cui, 2013; Pandey et al., 2008], OAZ1, S100P, and DUSP1 [Cheng et al., 2013], aldehyde dehydrogenase type 1 [Xu et al., 2013], and BMI1 [Ma et al., 2013].

LCK, survivin and PI-3K are other emerging biomarkers of the pre-malignant nature of OLP [Oluwadara et al., 2009], suggesting that LCK-inhibitors could be useful .

Chapter 6. Medical applications

Besides genetic imbalances, alterations and polymorphisms, behavioral lifestyle factors such as tobacco smoking and chewing, nutrition, infectious agents like HPV [Syrjänen et al., 2011] or HCV [Carrozzo and Gandolfo, 2003], environmental and occupational factors, psychological state like stress [Prolo et al., 2002; Valter et al., 2013] may play a role in the pathogenesis of OLP.

Annual screening for early oral cancer detection [Mattsson et al., 2002; Nagao and Warnakulasuriya, 2003] and personalized plaque controls (Stone et al., 2013) have been proposed as preventive approaches since a complete healing from OLP is very rare, and so far there insufficient evidences about the best treatment for OLP, even though corticosteroids are wide use for the management of this disease [Lodi et al., 2012]. Other treatments include retinoids, photochemotherapy [Pavlic and Vujic-Aleksic, 2013], and laser therapy [Boorghani et al., 2010].

In order to predict the potential miRNAs network related to OLP, we used our previously identified “leader genes” [Orlando et al., 2012], namely both the Class A genes (JUN, EGFR, FOS, IL2, and ITGB4) and the Class B genes (CASP3, CD247, IL2RA, IFNG, MMP2, LAMC2). Then, we mined the miRGen database using the “Targets” option (available at <http://www.microna.gr/mirgen>). This software relies on a relational database technology and enables a comprehensive research, integrating both the most widely used databases and repositories with the list of experimentally validated miRNAs and the bioinformatics tools for the microRNAs prediction (namely, DIANA-microT, miRanda, miRBase, PicTar, and TargetScanS) [Alexiou et al., 2010; Megraw et al., 2006].

Chapter 6. Medical applications

Topological properties of the obtained graphs portraying the OLP-related microRNAomes have been also studied.

As mentioned before, only few studies in the literature report OLP-related miRNAs. These articles (Dang et al., 2013; Danielsson et al., 2012a, 2012b; Gassling et al., 2013; Hu et al., 2013; Zhang et al., 2012) have been reviewed and summarized in Table 5. All the *in silico* predicted miRNAs are reported in Table 6 for Class A leader genes-related molecules and in Table 7 for the Class B leader genes-related miRNAs.

miRNAs that have been already experimentally validated represent the 32.26% (10/31) of the entire list.

Class A leader genes-related miRNAs predict the 42.19% (81/192) of the OLP + oral diseases-associated miRNAs validated in the literature.

miRNAs that have been already found to be associated with OLP are the 32.26% (10/31) of the entire list.

Class B leader genes-related miRNAs predict the 46.03% (58/126) of the OLP + oral diseases-associated miRNAs validated in the literature.

Taken together the Class A and Class B leader genes-related miRNAs, they can predict the 48.39% (15/31) of the OLP and the of OLP + oral diseases-associated miRNAs validated in the literature.

The entire OLP-related microRNAome is too complex to be properly studied (shown in Figure 29), since the number of nodes and edges of the graph is overwhelming. For this reason, after using a holistic and highly integrated approach, a reductionist methodology should be exploited in

order to underpin a more essential panel of miRNAs, that could be better investigated and elucidated via *ad hoc* targeted experiments.

From the topological analysis (summarized in Table 5), the two graphs representing the Class A and Class B-related microRNAomees appear to be the most significant part of the entire microRNAome graph, being the bulk of it in term of network diameter and radius, network centralization, topological density and clustering coefficients, while preserving other topological parameters (such as the scale-free behavior, the characteristic path length and the percentage of shortest paths) (Figures 30, 31).

In conclusion, microRNAs (miRNAs) are a family of small and short (usually 19-25 nucleotides long), single-stranded, endogenous, non-coding RNA molecules. They play a key role both in physiology and in pathology, and their role in oral diseases is emerging. Here we introduce a biomolecular strategy for predicting diseases-related miRNAs, based upon our previous published “Leader Genes algorithm”. We managed to find most of the already established OLP-related miRNAs. Moreover, we found also other miRNAs that have not been directly linked with OLP yet but have a role in other oral diseases. For the remaining miRNAs their role is still unknown. This biomolecular strategy can foster further research in the field of the biology of oral diseases, suggesting to the researchers and molecular biologists some targets to focus on and to explore with *ad hoc* targeted experiments.

TOPOLOGICAL PARAMETER	ENTIRE OLP- RELATED microRNAome	CLASS A Leader Genes-RELATED microRNAome	CLASS B Leader Genes-RELATED microRNAome
Clustering coefficient	0.19	0.96	0.73
Network diameter	3	2	2
Network radius	2	1	1
Network centralization	0.94	0.99	0.99
Shortest paths (%)	99	99	99
Characteristic path length	1.99	1.99	1.99
Average number of neighbors	41.49	5.78	7.7
Network density	0.04	0.01	0.01

Table 5: comparison of the topological properties among the entire OLP-related microRNAome, and the ones associated to Class A and B leader genes.

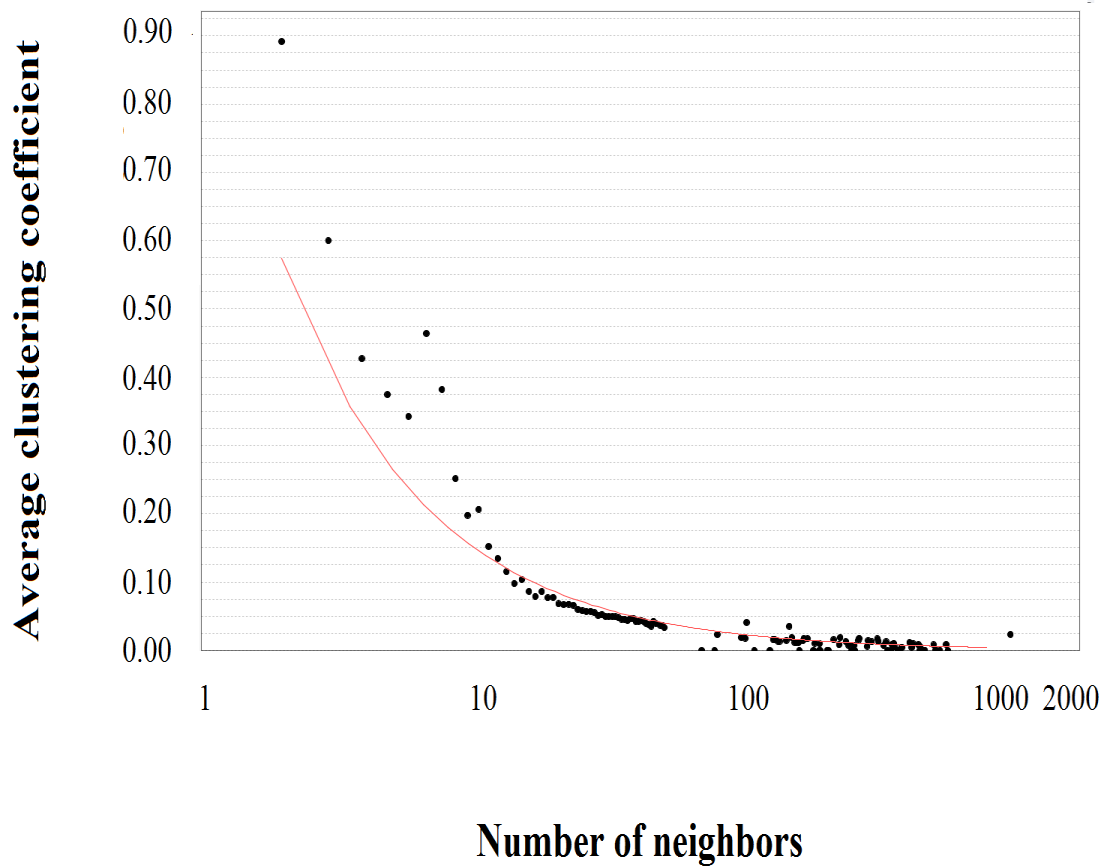


Figure 30: the topological properties of the OLP-related microRNAome.

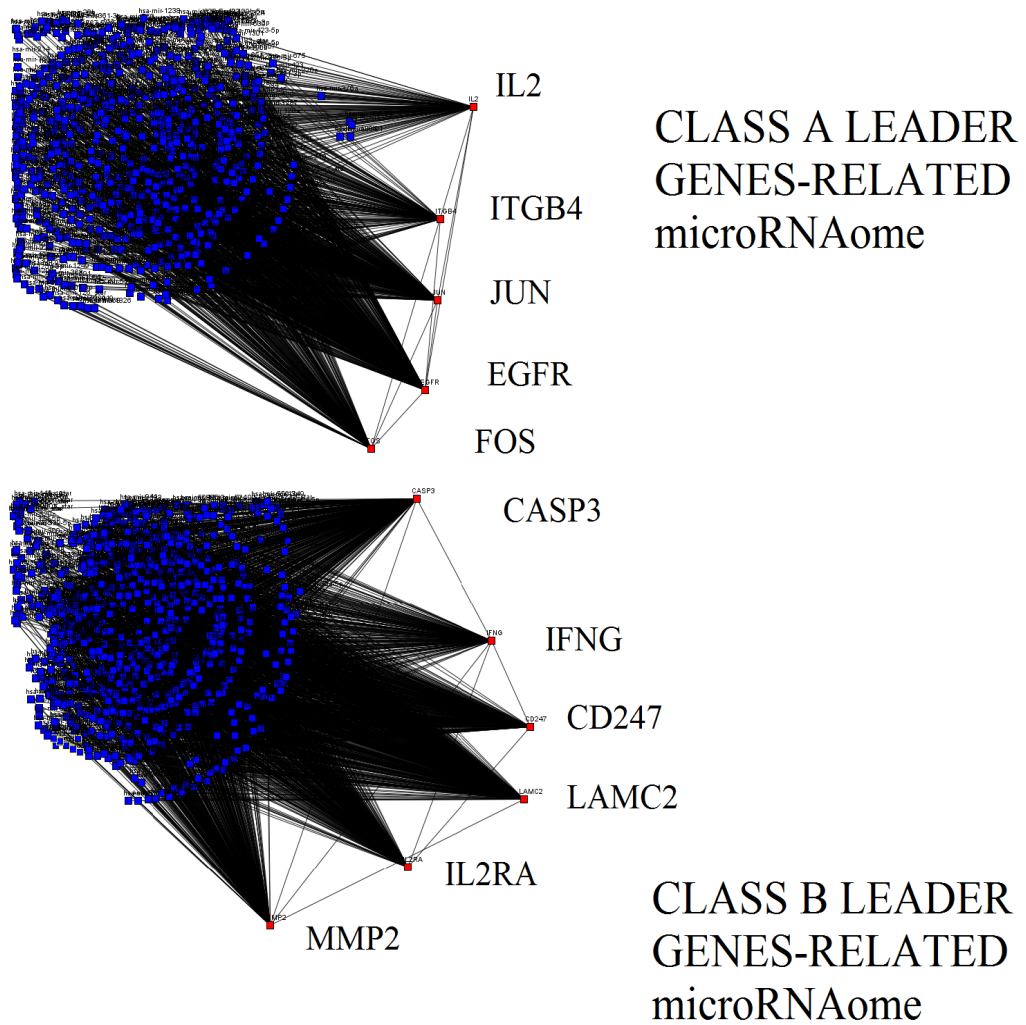


Figure 31: the leader genes-based OLP-related microRNAome.

6.3. NAPPA-BASED QCM APPLICATION IN ONCOLOGY

We built a prototype version of our NAPPA-based biosensor with each quartz containing 16 (4x4) NAPPA spots of 300-micron square in order to synthesize functional proteins *in situ* directly from printed cDNAs (complementary DNAs) just before the assay [Nicolini et al., 2012a; Nicolini et al., 2012b; Spera et al., 2013].

Quartzes' gold surfaces were coated with cysteamine to allow the immobilization of the NAPPA printing mix.

Briefly, quartzes were washed three times with ethanol, dried with Argon and incubated over night at 4°C with 2mM cysteamine. Quartzes were then washed three times with ethanol to remove any unbound cysteamine and dried with Argon.

Plasmids DNA coding for GST tagged proteins were transformed into *E. coli* and DNA were purified using the NucleoPrepII anion exchange resin (Macherey Nagel). NAPPA printing mix was prepared with 1.4µg/ul DNA, 3.75µg/ul BSA (Sigma-Aldrich), 5mM BS3 (Pierce, Rockford, IL, USA) and 66.5µg polyclonal capture GST antibody (GE Healthcare).

Negative controls, named master mix (hereinafter abbreviated as “MM”), were obtained replacing DNA for water in the printing mix. Samples were incubated at room temperature for 1 hour with agitation and then printed on the cysteamine-coated gold quartz using the Qarray II from Genetix.

As test proteins, we used p53 and MDM2. These monitored proteins are of fundamental importance in the molecular mechanisms leading to malignant cell transformation and cancer; p53 is in fact a 53-kiloDalton phosphoprotein oncosuppressor, encoded by a 20-kilobases gene situated on the short arm of human chromosome 17 and termed as the “guardian of the

genome” and the “policeman of oncogenes” [Efeyan and Serrano, 2007; Levine, 1997]. Mutated, it is involved in up to 70% of human tumors, being responsible of cell growth arrest, senescence, apoptosis in response to an array of stimuli such as DNA damages (DSB, or double-strand-breaks), hypoxia, telomeres shortening, cell adhesion, oncogene activation and other molecular and cellular stresses [Vousden and Lu, 2002].

MDM2, a p53-specific E3 ubiquitin ligase, inhibits p53 functions by binding to its N-terminal transcription-promoting domain, thus having an oncogenic activity. The MDM2-p53 plays a major role in cancer, being also a molecular therapeutic target and its monitoring is of crucial importance in cancer diagnosis and treatment [Shangary and Wang, 2008].

Table 6: peak frequency (f), half-width at half-height (Γ), maximum conductance increment (Y_{RE}), and normalized D factor for the impedance curves shown in Figure 90 left.

Impedance Curve	f (Hz)	Γ (Hz)	Y_{RE} (mS)	D_N (Hz/mS)
beginning	9475435	2220	0,42	10699
After protein expression and washing	9470905	3705	0,31	23903
After MDM2 addition	9467860	3600	0,14	53333

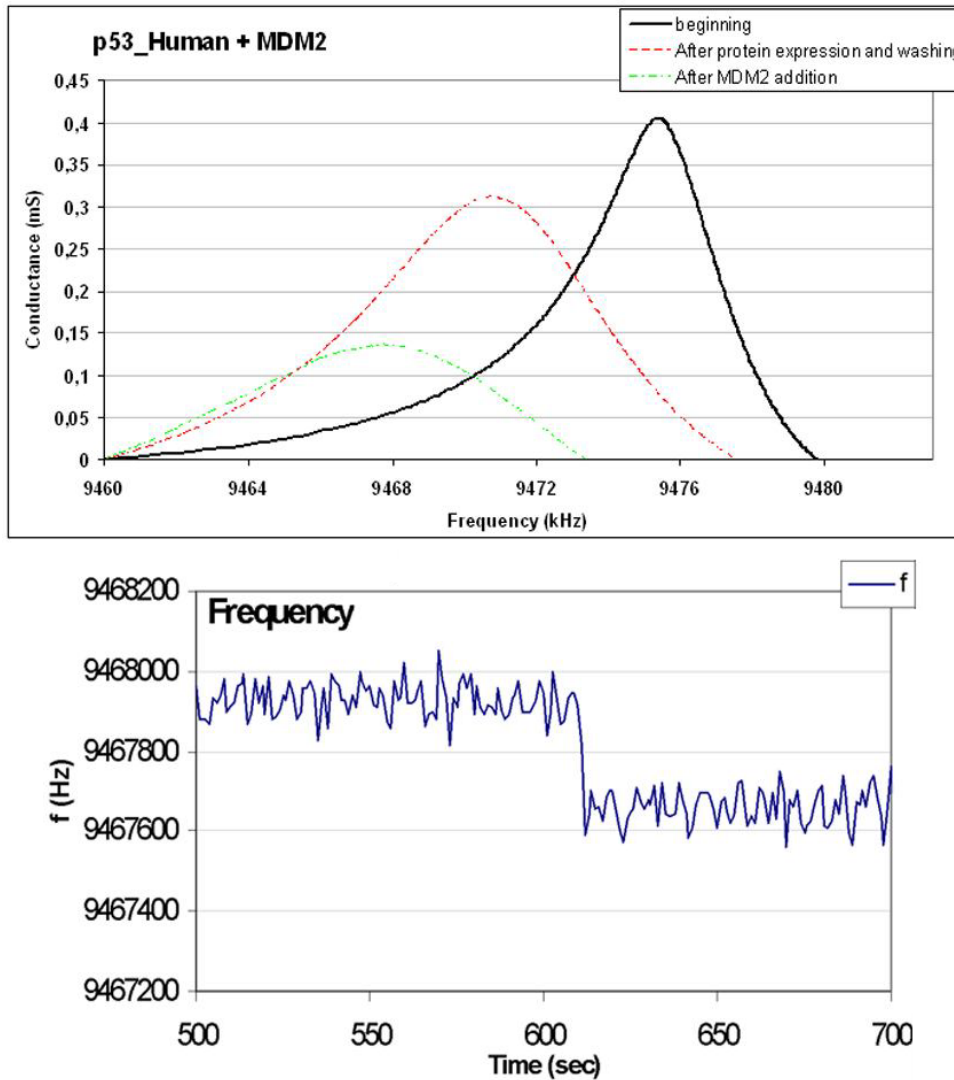


Figure 32: (Left) p53 quartz impedance curves recorded before the gene expression (solid curve), after the protein expression and the washing process (dashed curve) and after MDM2 addition (dotted curve); (Right) p53 quartz frequency variation in real time; at time $t = 600$ s a MDM2 solution ($50 \mu\text{M}$ in PBS) was injected in the flow chamber.

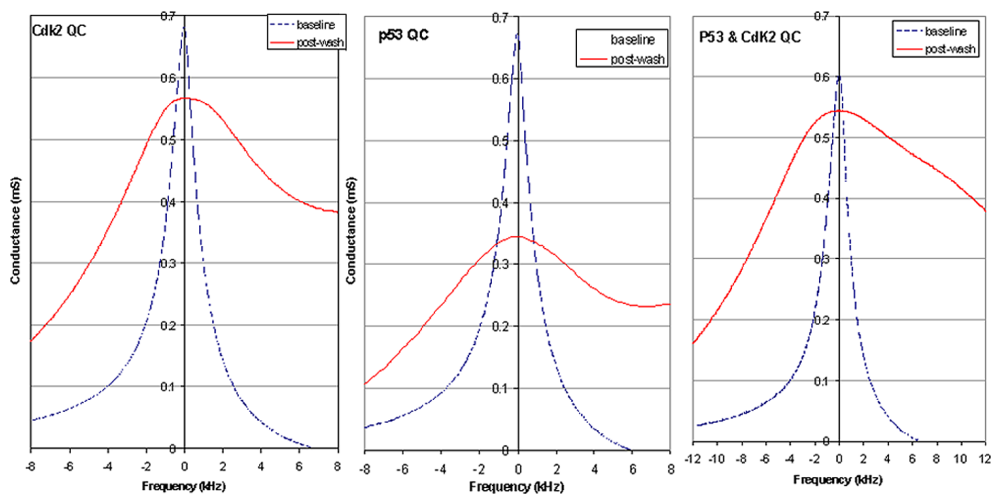


Figure 33: NAPP expression for the individual CDK2 (left), P53 (middle) and combined P53 and CDK2 genes (right) on each individual quartz. The blue curves were acquired before the expression of NAPP, while the red curves were acquired after the protein expression/capture and washing process. The curves have been centered to their maximum frequency to better visualize the changing in bandwidths and maximum of conductance.

In order to enhance the sensitivity, each quartz was printed with 100 identical features of 300 microns diameter each, spaced by 350 microns center-to-center.

The human cDNAs immobilized on the NAPP-QC were: CDK2 (Cyclin-Dependent Kinase 2), CYP11A1 (Cytochrome P450, Family 11, Subfamily A, Polypeptide 1), Jun, MLH1 (mutL homolog 1), p53, and polD1

Chapter 6. Medical applications

(Polymerase (DNA Directed), Delta 1, Catalytic Subunit). The size of those proteins ranged from 61 kDa to 150 kDa, including the 27 kDa GST tag.

Gene expression was performed immediately before the assay, following the protocol described as modified from [Festa et al., 2013].

Briefly, *in vitro* transcription and translation were performed using human IVTT mix (1-Step Human Coupled IVTT Kit, Thermo Fisher Scientific Inc.), prepared according to the manufacturers' instructions. The quartz, connected to the nanogravimeter inside the incubator, was incubated for 1.5 hr at 30°C with 40µl of HeLa lysate mix for proteins synthesis and then, the temperature was decreased to 15°C for a period of 30 min to facilitate the proteins binding on the capture antibody (anti-GST). After the protein expression and capture, the quartz was removed from the instrument and washed at room temperature, in PBS for 3 times. The quartz was then placed in the flux chamber for protein-protein interaction analysis. The protocol described above was followed identically for both negative control QC (the one with only MM, i.e, all the NAPPA chemistry except the cDNA) and protein displaying QC.

After protein expression and capture, the QCs were washed and used for the interaction studies as follows: QC displaying Jun was exposed to a flow of a 33 µM ATF2 (Sigma-Aldrich) solution in PBS (for flow interaction at 0.02 ml/min flow rate) for 10 minutes at 22°C; alternatively 60 µl of 33 µM ATF2 solution in PBS was added on the QC surface (for static interaction) for 10 minutes at 22°C.

We also tested the possibility to analyze protein-protein interactions in QC displaying multiple proteins. For this aim, we co-printed cDNA for Jun&CDK2 and Jun&CDK2&p53 on a single QC and after the expression a

Chapter 6. Medical applications

mixture of these proteins was displayed on the QC surface. We analyzed the interaction response to ATF2 on both NAPPA-expressed QCs. It is well known that the proto-oncoprotein c-Jun is a major dimerization partner of ATF2, and c-Jun–ATF2 heterodimers are important for many cellular and neoplastic processes [Hai and Curran, 1991; Liu et al., 2006].

On the contrary, no interactions are known between ATF2 and CDK2 or p53.

We analyzed the interaction between CYP11A1 and cholesterol, both in solution and in blood, to acquire information on the binding kinetics⁶. After protein expression and capture CYP11A1 expressing QC was positioned in the flow chamber and exposed to a flow of a 50 μ M Cholesterol (Sigma-Aldrich) solution in 30% sodium cholate (Sigma-Aldrich), at 0.02 ml/min flow rate, for 10 minutes at 22°C. We used cytochrome P450_{sc} (CYP11A1) for the detection of Cholesterol (Sigma-Aldrich) because of its specificity. As negative control we analyzed the interaction between Clomipramine and polD1 and MLH1. MLH1 is a protein involved in DNA mismatches repair and polD1 is a DNA polymerase. Noone of these protein should interact with Clomipramine.

Various kinetic models have been developed to describe protein-protein and protein-small molecules interactions.. For our data, on the base of the results of [Turon et al., 2008], we performed mathematical modelling of the protein-protein interaction, assuming a saturation like kinetics, that is to say a quasi-steady-state condition model. For this purpose, we used the following experimental equation [Turon et al., 2008] based on an exponential decay function instead of using deterministic models, focusing on frequency variation:

Chapter 6. Medical applications

$$\Delta f = M_{max}(1 - e^{-t/\tau}),$$

where Δf is the frequency variation (in Hz), M_{max} is the maximum binding value in hertz (corresponding to the minimum frequency measured, or plateau value), τ is the relaxation time (expressed in reciprocal minutes) that is the reciprocal of the enzyme protein adsorption/interaction rate, and t is the time (in minutes).

The sensitivity of our NAPPA based biosensor was determined by the QC characteristics and the sensitivity of the nanogravimeter. At the moment, the minimum frequency shift detectable is of 0.05 Hz that corresponds to about 0.3 ng of detected molecules [Spera et al., 2013].

QCM_D measures were calibrated for frequency and for D factor shifts. The frequency calibration was performed using different amounts of thaumatin. The calibration curve equation (obtained with Ordinary Least Squares methods, OLS) is:

$$\Delta f = -7.16 - 231.18 m;$$

with $r^2 = 0.9986$.

The D factor calibration in function of the viscosity was obtained using fructose samples at different concentrations. The calibration curve equation (obtained with OLS methods) is:

$$D = 0.831 + 0.286 \eta; \text{ (equation 5)}$$

with $r^2 = 0.9990$. For both the curves, a good correlation with the linear slope was found^{3,6}.

We analyzed the conductance curves acquired in NAPPA-QCs in different steps of the expressing and capturing process. Coefficients of variation (C_V) have been computed for each conductance curve, according to the following equation:

$$C_V = \sigma / \mu;$$

where σ is the standard deviation and μ the mean.

Figure 33 shows the conductance curves for three NAPPA-QCs expressing p53, CDK2 or a mixture of p53 and CDK2 (both cDNAs were co-immobilized in each feature). The blue curves were acquired before the expression of NAPPA while the red curves were acquired after the protein expression/capture and washing process. The curves have been centered to their maximum frequency to better visualize the changing in bandwidths and conductance. These data pointed to a unique conductance curve shape for each protein and suggested the possibility to identify the expressed proteins by QCM-D even when combined on the same expressing QC (Figure 3).

To test the possibility to acquire information on the kinetic constant of a protein-small molecule interaction we generated a NAPPA-QC expressing CYP11A1 to be tested for cholesterol interaction as proof of principle. The P450_{scc}-cholesterol interaction, in fact, is well characterized and the results obtained have been satisfactorily compared with those in literature [Antonini et al., 2004; Arya et al., 2008; Nicolini et al., 2001; Shumyantseva et al., 2004; Stura et al., 2007].

In Figure 4 are reported the conductance curves for the negative control (MM) or CYP11A1 expressing NAPPA-QC arrays. The conductance curves were acquired in different steps of NAPPA expression process. In particular:

- Before the beginning of the gene expression process (“Baseline”, the NAPPA-QC is dry);
- After the addition of human IVTT lysate at 30°C (“IVTT addition”), *i.e.* prior protein expression;

Chapter 6. Medical applications

- After 90 minutes from the addition of human IVTT lysate, *i.e.* after protein expression (“90 min IVTT addition”);
- After 120 minutes from the addition of human IVTT lysate, *i.e.* at the end of QC incubation at 15°C, which is the capture step (“Post-capture (30 min)”);
- After the final washing process with PBS (“Post-wash”).

At the end of the entire expression process, the CYP11A1 expressing QC was positioned in the flow chamber and exposed to a flow of a 50 μ M Cholesterol solution in 30% sodium cholate. The conductance curve “MM+CYP11A1+cholesterol” (see Figure 4b) was recorded in this condition.

In table 6 are reported the main parameters of the conductance curves of Figure 32, 33.

Table 7: Main parameters of QC-NAPPA displaying no proteins (MM) or CYP11A1. Conductance curves were collected in different steps of NAPPA protocol.

	f (Hz) ^a	Γ (Hz) ^a	G_{max} (mS) ^a	$D * 10^{3b}$	D_N (Hz/mS) ^b
MM Conductance curves					
Baseline	9497485	2422	0.432	0.51	11218
IVTT addition	9493570	4562	0.370	0.96	24653
90 min IVTT addition	9493495	4500	0.390	0.95	23071
Post-capture (30 min)	9493120	4875	0.378	1.03	25814
Post-wash	9489085	6255	0.046	1.32	273144

CYP11A1 Conductance curves

Baseline	9487615	2265	0.440	0.48	10286
IVTT addition	9481540	4198	0.399	0.87	21069
90 min IVTT addition	9479935	3984	0.401	0.84	19895
Post-capture (30 min)	9479790	3980	0.400	0.84	19895
Post-wash	9475975	5625	0.150	1.19	75050
MM+CYP11A1+ cholesterol	9478300	6410	0.268	1.35	47782

^a f is peak frequency, Γ is the half-width half-maximum and G_{max} is the max conductance.

^b D factor and $D_N = 2\Gamma/G_{max}$ normalized D factor.

It is evident the decrease in the frequency (f) due to the human IVTT lysate addition. There is also a change in viscoelastic properties of the quartzes after the human IVTT lysate addition, leading to a measurable increase of the bandwidth (Γ). During the incubation, on the contrary, the frequency and bandwidth variations were minimal. This effect could be related to two main effects: first merely due to the IVTT lysate addition on the QC surface - when the QC comes in contact with a solution the frequency decreases depending upon the viscosity and the density of the solution, and there is a decrement in damping the resonant oscillation ⁴⁷- and the second due to the change of the composition of both QC surface and IVTT lysate after the gene expression and the protein synthesis and immobilization. The conductance curves acquired after PBS washing evidenced the further changes of solution in contact with the QC.

In tables 7, 8 is reported the shift of the main parameters normalized against the “Baseline” conductance curve and the conductance curve acquired after

human IVTT lysate (“IVTT addition”). It is also reported the mass of molecules immobilized on the QC surface at the end of immobilization protocol and after the interaction with Cholesterol (m), estimated through the calibration coefficients.

Table 8 Shift of the main parameters of MM and CYP11A1 conductance curves after lysate addition and the corresponding mass of immobilized protein on the QC surface.

MM Conductance curves	C_v^a	ΔD^b	$\frac{\Delta D_N}{(Hz/mS)^b}$	Δf (Hz) ^c	$\Delta f'$ (Hz) ^c	m (μg) ^c	m' (μg) ^c
IVTT addition	3.3%	-	-	-3915	-	17,0	-
90 min IVTT addition	4.1%	-0.01	-1582	-3990	-75	17,3	0,4
Post-Capture (30 min)	4.0%	0.07	1161	-4365	-450	18,9	2,0

CYP11A1 Conductance curves	C_v^a	ΔD^b	$\frac{\Delta D_N}{(Hz/mS)^b}$	Δf (Hz) ^c	$\Delta f'$ (Hz) ^c	m (μg) ^c	m' (μg) ^c
IVTT addition	4.8%	-	-	-6075	-	26.3	-

^a Coefficient of variation of three independent experiments..

^b D factor and normalized D factor shifts (ΔD and ΔD_N) respect the values immediately after lysate addition

^c Frequency shifts respect the initial frequency (Δf) and respect the frequency immediately after lysate addition ($\Delta f'$) and corresponding molecular masses (m and m').

Chapter 6. Medical applications

90 min IVTT addition	4.8%	-0.03	-1174	-7680	-1605	33.3	7.0
Post-Capture (30 min)	4.6%	-0.03	-1174	-7825	-1750	33.9	7.6
MM+CYP11A1+ cholesterol	6.9%	0.48	26713	-9315	-3240	40.3	14

^a Coefficient of variation of three independent experiments..

^b D factor and normalized D factor shifts (ΔD and ΔDN) respect the values immediately after lysate addition

^c Frequency shifts respect the initial frequency (Δf) and respect the frequency immediately after lysate addition ($\Delta f'$) and corresponding molecular masses (m and m').

To estimate the amount of proteins anchored on the QC surface after the NAPPA expression we had to account for the human IVTT lysate molecules aspecifically adsorbed on the quartz surface. Assuming that on each QC surface there was the same aspecific adsorption we could estimate it from reference quartz conductance curves. In particular we considered the frequency shift between the reference QC conductance curves acquired immediately after the human IVTT lysate addition ("IVTT addition") and that acquired at the end of the protein anchorage ("90 min IVTT addition"); this value is 450 Hz (see table 1), that correspond to 2 μg of molecules aspecifically adsorbed (see tables 7, 8). The values of immobilized CYP11A1 molecules, therefore, was 5.6 μg – obtained subtracting 2 μg of molecules aspecifically adsorbed to the 7.6 μg of molecules adsorbed in correspondence of "Post-Capture (30 min)" curve (see table 2) - while the amount of Cholesterol was 6.4 μg – ob-

Chapter 6. Medical applications

tained subtracting 5.6 μg of CYP11A1 proteins captured on the array and 2 μg of molecules aspecifically adsorbed from the 14 μg of molecules adsorbed in correspondence of “MM+CYP11A1+ cholesterol” curve. We verified (data not reported) that the exposure of the QC to sodium cholate solution affected only the bandwidth of the curve, considerably increasing the D factor, leaving the resonance frequency unchanged. This result also served as control to verify the stability of the biosensor.

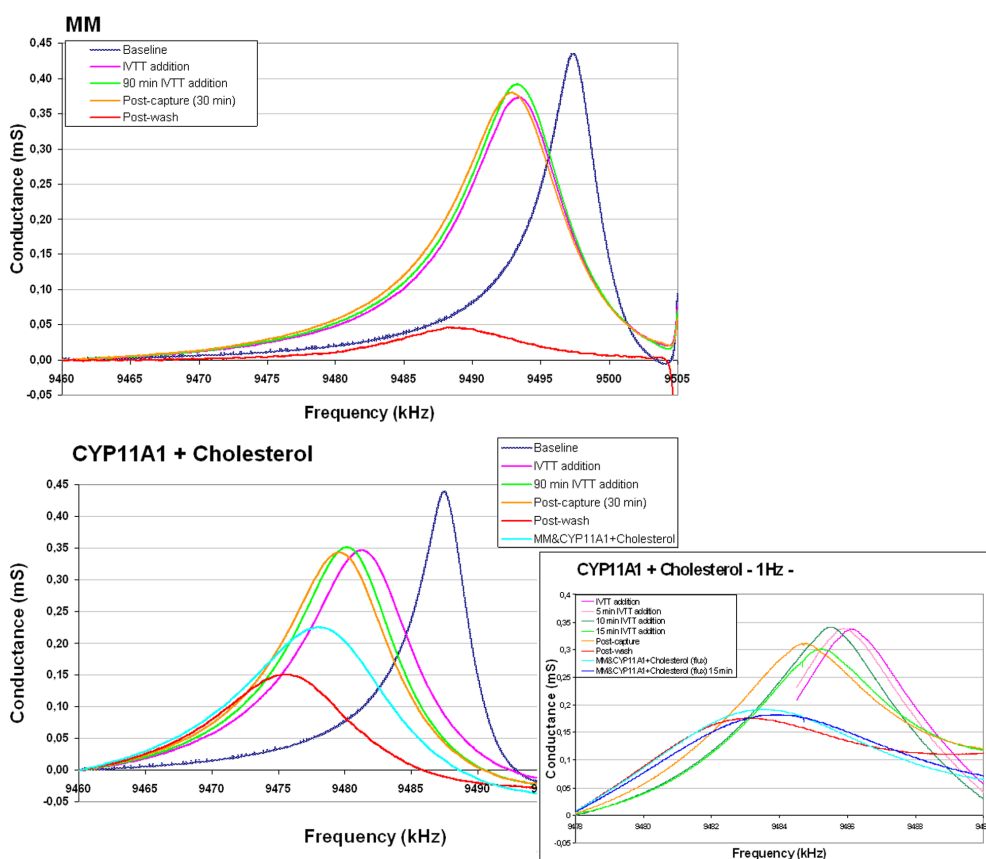


Figure 34: conductance curves of MM QC (upper) and of CYP11A1 QC (lower). The curves were collected in different steps of NAPPA protocol, as reported in the legend, and after the addition of cholesterol. In the box

the MM+CYP11A1+cholesterol conductance curves acquired with frequency acquisition steps of 1 Hz.

Our QCM_D instrument gave us the opportunity to monitor in real-time the trend of D factor and f . In Figure 34 we reported frequency *vs.* time during the interaction between CYP11A1 and Cholesterol, both in solution (as previously reported) and in blood. We obtained a K of about 100 μM , a value in good agreement with the values found in the literature.

In the first experiments conductance curves were acquired every 5 minutes (data non reported) and it was noticed that 15 minutes after IVTT lysate addition at 30°C the position and shape of the curves did not change until we decrease the temperature to 15°C. After few minutes at 15°C again position and shape of the curves did not change until the washing step. Deducing from these results that the protein expression took place in the first minutes and that also their capture needed only few minutes we decided, then to perform next experiments reducing the expression time from 90 minutes to 15 minutes (at 30°C) and the capture time from 30 to 5 minutes (at 15°C).

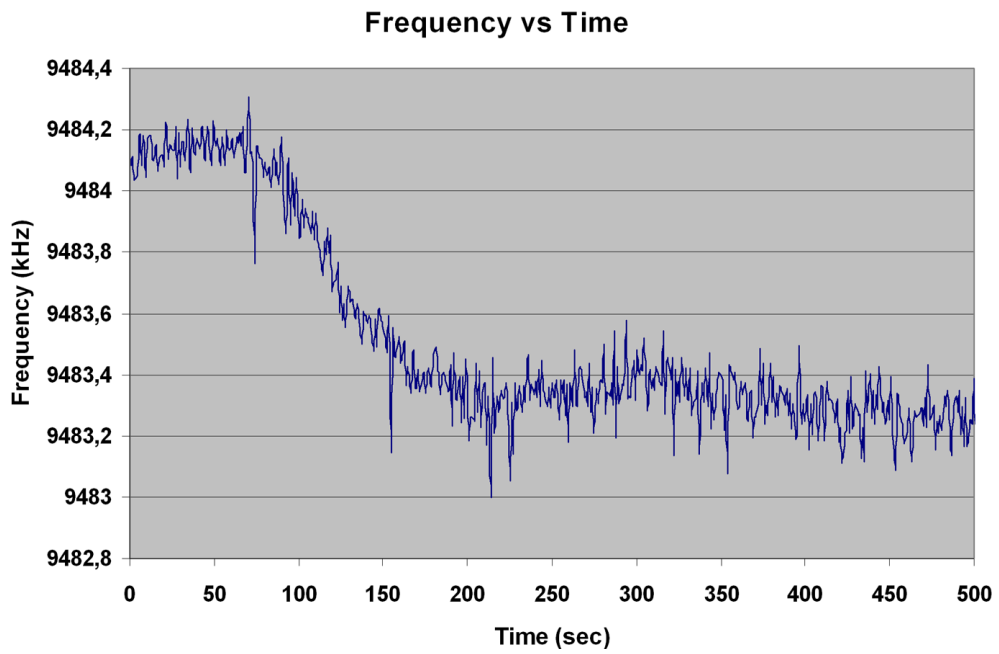
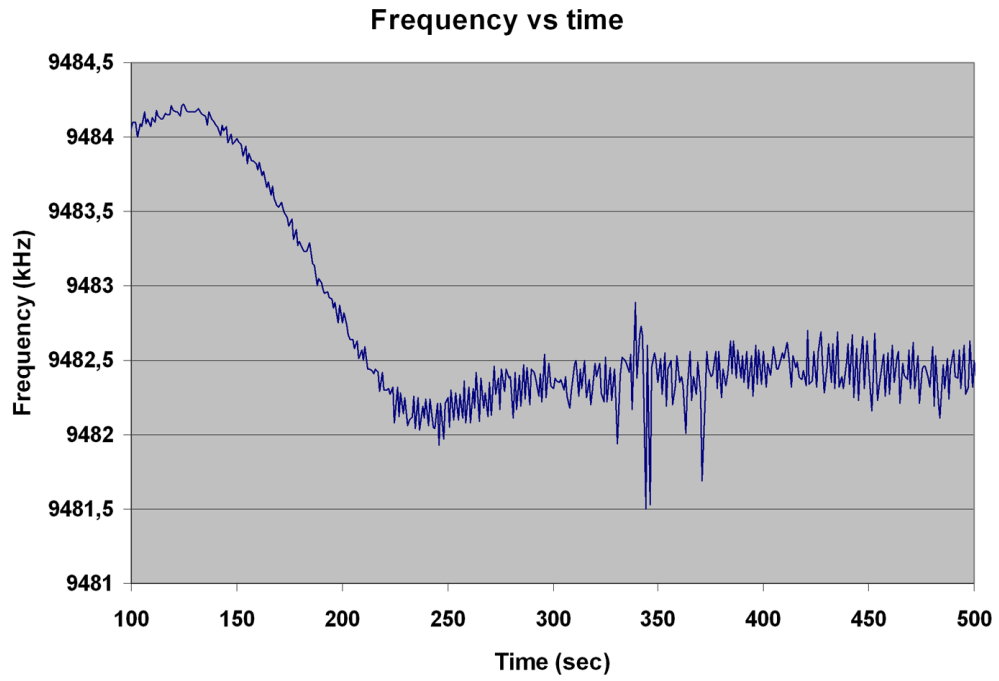


Figure 35: FLOW interaction CYP11A1 with HDL cholesterol in blood (upper panel) and in solution (lower panel).

To test the possibility to analyze multi-protein interactions we immobilized on the same QC both Jun and CDK2 cDNAs and tested the response to ATF2 interaction. c-Jun is a partner of ATF2 in many cellular processes³⁸, however no interaction are known between CDK2 and ATF2.

The conductance curves of NAPPA-QC expressing Jun&CDK2 acquired in different steps of NAPPA expressing process are reported in Figure 6a. After protein expression and capture the Jun&CDK2 expressing QC was positioned in the flow chamber and exposed to a flow of 33 μ M ATF2 solution in PBS. The conductance curve “MM+Jun&CDK2+ATF2” (Figure 35) was recorded after 3 minutes of flux. In Table 9 are reported the main parameters of conductance curves of Figure 6a.

Table 9: main parameters of QC-NAPPA displaying Jun&CDK2. Conductance curves were collected in different steps of NAPPA protocol.

Jun&CDK2 Conductance curves	f (Hz) ^a	Γ (Hz) ^a	G_{max} (mS) ^a	$D * 10^3$ ^b	D_N (Hz/mS) ^b
IVTT addition	9482473	3073	0.461	0.65	13336
5 min IVTT addition	9482145	2945	0.457	0.62	12885
15 min IVTT addition	9481600	2800	0.447	0.59	12528
Post-capture (5 min)	9481018	2836	0.410	0.60	13822
Post-wash	9479200	2400	0.261	0.51	18391
PBS-flux	9480764	2909	0.313	0.61	18577
MM+Jun&CDK2+ATF2	9481309	2982	0.294	0.63	20284

^a f is peak frequency, Γ is the half-width half-maximum and G_{max} is the max conductance.

^b D factor and $D_N = 2\Gamma/G_{max}$ normalized D factor.

In table 10 is reported the shift of the main parameters normalized against the conductance curve acquired immediately after human IVTT lysate addition (“IVTT addition”). Through equation 5 we evaluated the mass of immobilized proteins on the quartz surface at the end of capture protocol and after the interaction with ATF2 (m').

Table 10: shift of the main parameters of Jun&CDK2 conductance curves after lysate addition and after protein-protein interaction, and relative mass of immobilized proteins on the quartz surface.

Jun&CDK2 Conductance curve	C_v ^a	DD ^b	DD_N (Hz/mS) ^b	Df' (Hz) ^b	m' (μ g) ^c
5 min IVTT addition	5.5%	-0.03	-455	-456	2.0

Chapter 6. Medical applications

15 min IVTT addition	5.7%	-0.06	-660	-807	3.5
Post-capture (5 min)	5.3%	-0.06	509	-1432	6.2
Post-wash	4.8%	-0.28	45965	-3895	17.0
PBS-flux	7.3%	-0.21	11182	-2561	-5.8*
MM+Jun&CDK2+ATF2	8.0%	-0.20	9644	-2491	-6.1*

^a Coefficient of variation. We performed three independent experiments.

^b D factor, normalized D factor and frequency shifts (DD , DD_N and Df°) with \square respect the values immediately after IVTT lysate addition (see table 3).

^c Mass of molecules immobilized on the quartz surface for each step (calculated by equation 5).

* These values were obtained considering the frequency shifts respect “Post-wash” curve, since the QC was in contact with PBS solution.

The amount of molecules captured on the QC 5 minutes after protein capture corresponded to about 6 μg . A further frequency decrease was recorded after the washing. This decrease was not due to further molecules capture on the QC surface but rather to the change of the solution in contact with the QC surface, since the frequency is dependent upon the viscosity and the density of the solution. This effect is mostly evident considering the shift in D and normalized D factor. In particular post wash DD_N increase of about 100 times. What is meaningful is the frequency shift after QC exposure to ATF2 flux. We recorded, in fact, a frequency increase, rather than a decrease as expected. Evidently the flux, even if very gentle, caused spoliation of the QC, removing some captured molecules and making impossible to assess the interaction between Jun and ATF2.

In order to overcome the spoliation problem, the interaction between Jun and AFT2 was tested in a static way, by simply adding the ATF2 solution

Chapter 6. Medical applications

on the QC surface. For this experiment, QC with immobilized cDNA of Jun&CDK2&p53 were expressed for a period of 90 minutes and captured for 30 minutes. After protein expression and capture, the QC was washed with PBS, and incubated for 10 minutes with 60 μ l of 33 μ M ATF2 solution in PBS.

In Table 11 are reported the main parameters of conductance curves of Figure 35 and in Table 12 is reported the shift of these parameters with respect those of the conductance curve acquired immediately after lysate human IVTT lysate addition (“IVTT addition”).

Table 11: Main parameters of QC-NAPPA displaying Jun&CDK2&p53. The conductance curves were collected in different steps of NAPPA protocol.

Jun&CDK2&p53 conductance curves	f (Hz) ^a	Γ (Hz) ^a	G_{max} (mS) ^a	$D * 10^3$ ^b	D_N (Hz/mS) ^b
Baseline	9489250	2375	0.436	0.50	10900
IVTT addition	9483250	4625	0.347	0.98	26629
10 min IVTT addition	9482625	4675	0.339	0.99	27543
80 min IVTT addition	9481750	4625	0.342	0.98	27038
Post-capture (30 min)	9481250	4688	0.300	0.99	31250
Post-wash	9477250	5375	0.111	1.13	97262
MM+Jun&CDK2&p53+ATF2	9475125	6000	0.188	1.27	63687
MM+Jun&CDK2&p53+ATF2 (10 min)	9475625	6063	0.187	1.28	64894

^a f is peak frequency, Γ is the half-width half-maximum and G_{max} is the max conductance.

^b D factor and $D_N = 2\Gamma/G_{max}$ normalized D factor.

Table 12: Shifts of the main parameters of Jun&CDK2&p53 conductance curves after lysate addition and relative mass of immobilized protein on the quartz surface.

Jun&CDK2&p53 conductance curves	C_v	ΔD	$\frac{\Delta D_N}{(Hz/mS)}$	$\Delta f'$ (Hz)	m' (μg)
10 min IVTT addition	5.4%	0.01	914	-625	2.7
Post-expression (80 min)	6.0%	0.00	409	-1500	6.5
Post-capture (30 min)	5.5%	0.01	4621	-2000	8.7
Post-wash	5.2%	0.15	70633	-6000	26.2
MM+Jun&CDK2&p53+ATF2	6.5%	0.29	37058	-8125	9.2*
MM+Jun&CDK2&p53+ATF2 (10 min)	8.0%	0.30	38265	-7625	7.1*

^a Coefficient of variation. We performed two different experiments.

^b D factor, normalized D factor and frequency shifts (DD , DD_N and Df') with respect the values immediately after IVTT lysate addition (see table 5).

^c Amount of molecules immobilized on the QC surface (calculated through equation 5 respect the frequency recorded immediately after IVTT addition).

* These values were obtained considering the frequency shifts respect post-wash curve, since the QC was in contact with PBS solution.

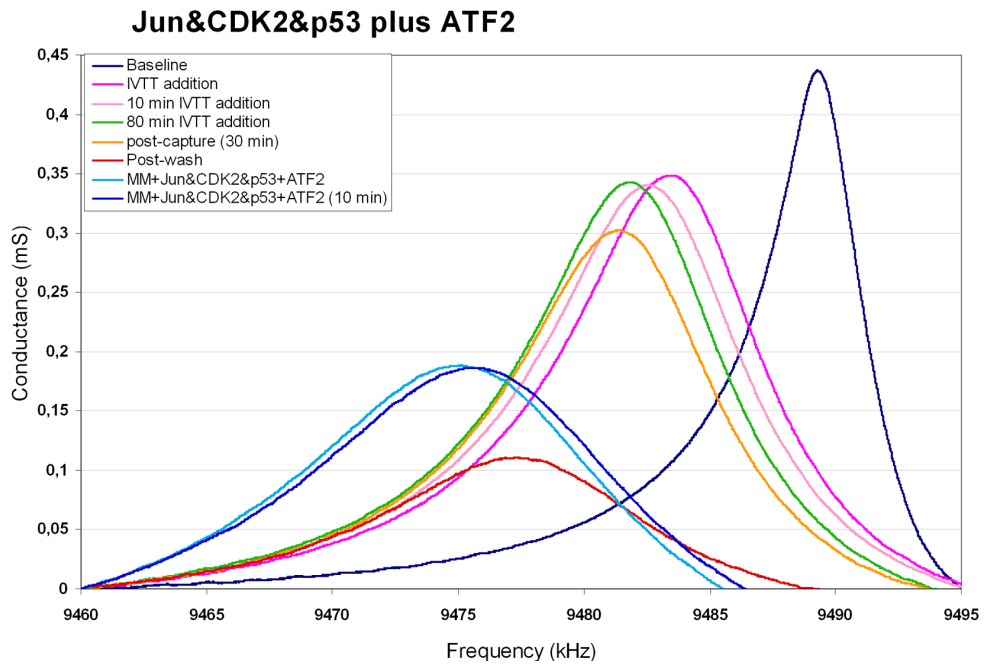
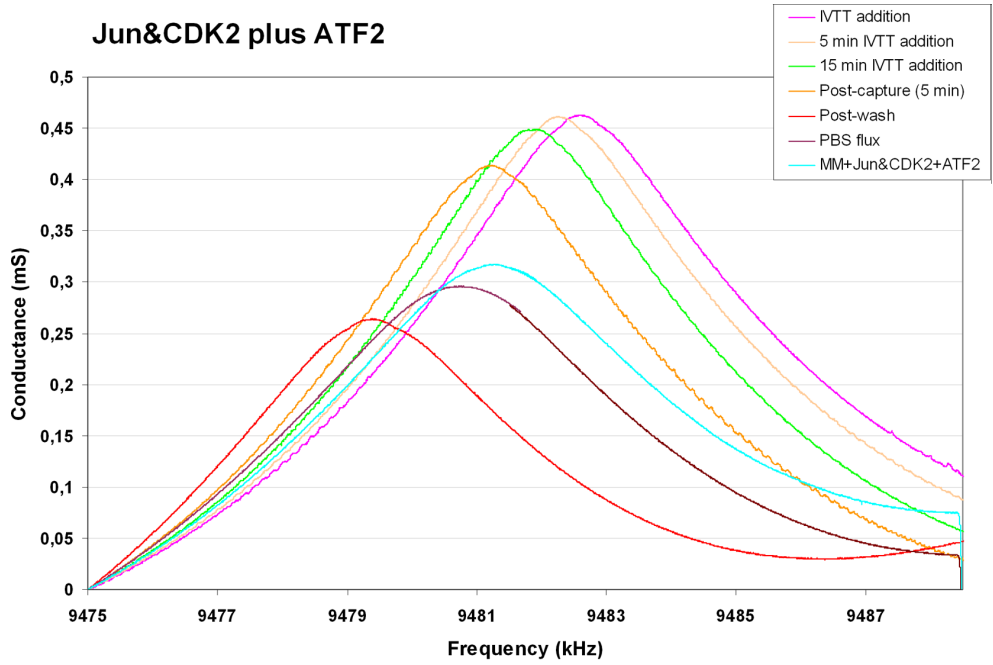


Figure 36: conductance curves of Jun and CDK2 expressing QC (upper panel). Conductance curves of Jun, CDK2 and p53 expressing QC (lower panel). The curves were collected in different steps of NAPPA process, as reported in the legends, and after the addition of ATF2 solution.

The unique conductance curve shape for each of the three proteins confirms the possibility to identify the three expressed proteins by QCM_D even when combined on the same expressing QC (Figure 37).

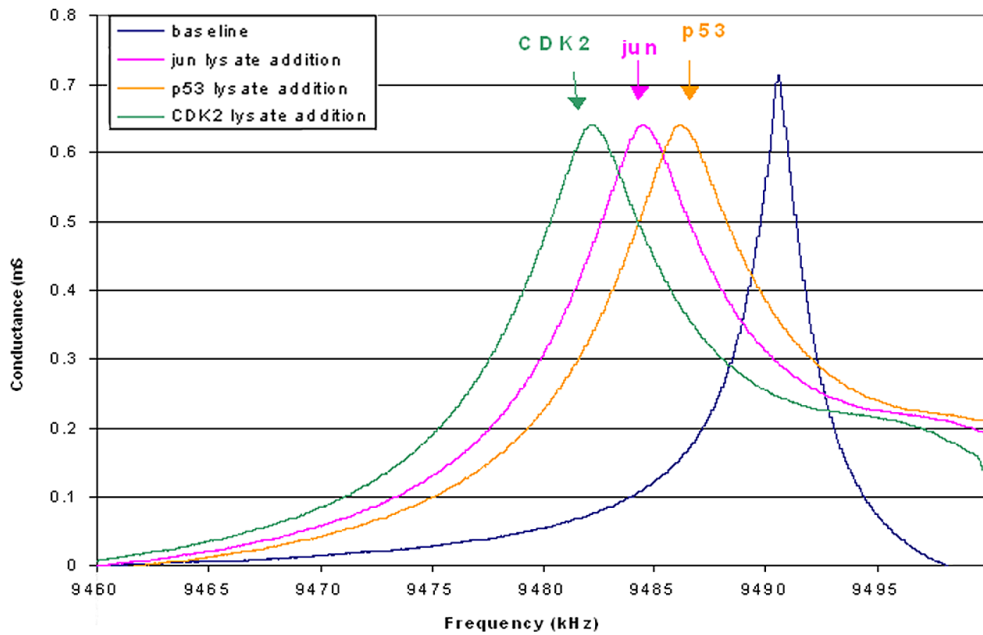


Figure 37: matching of conductance curves of Jun expressing QC, CDK2 expressing QC and p53 expressing QC after lysate addition.

The amount of molecules captured on the QC 90 minutes post-expression and 30 minutes post-capture corresponded to about 8 μg , not much different from 6 μg obtained in the previous experiments. Again a frequency decrease

was recorded after the washing (due to PBS solution). In contrast to what previously obtained, the frequency further decreased after QC exposure to ATF2 flux, suggesting that the interaction between Jun and ATF2 was favored by the static process. To evaluate the amount of ATF2 bound to Jun we considered the frequency shift between “post-wash” and “MM+Jun&CDK2&p53+ATF2” curves, because in both conditions the QC was in the same solution (PBS) and any changes in the frequency is due to the binding of ATF2. The amount of ATF2 molecules bound to the QC was 9.2 μg .

As negative control a NAPPA-QC co-expressing polD1 and MLH1 was tested for Clomipramine interaction. After protein expression and immobilization and after PBS washing 60 μl of 110 μM Clomipramine solution in PBS was spotted on the polD1&MLH1 expressing QC. The conductance curve “MM+ polD1&MLH1+Clomipramine” (Figure 38) was recorded 1 minute after Clomipramine solution addition.

In Table 13 are reported the main parameters of conductance curves of Figure 38 and in Table 14 are reported the shifts of these parameters with respect to those of “IVTT addition” curve.

Table 13: main parameters of polD1&MLH1 conductance curves collected in different steps of NAPPA expression protocol.

polD1&MLH1 Conductance curves	f (Hz) ^a	Γ (Hz) ^a	G_{max} (mS) ^a	$D * 10^3$ ^b	D_N (Hz/mS) ^b
Baseline	9486895	2860	0.450	0.60	12710
IVTT addition	9486439	2702	0.441	0.57	12254
5 min IVTT addition	9486088	2596	0.431	0.55	12050

15 min IVTT addition	9485463	2596	0.393	0.55	13219
Post-capture (5 min)	9483000	1509	0.051	0.32	58674
Post-wash	9484333	1877	0.157	0.40	23892
MM+polD1& MLH1+Clomipramine	9484404	1895	0.170	0.40	22354

^a f is peak frequency, Γ is the half-width half-maximum and G_{max} is the max conductance.

^b D factor and $D_N = 2\Gamma/G_{max}$ normalized D factor.

The amount of molecules captured on the QC 15 minutes after the expression and 5 minutes after the capture corresponded to about 15 μg (Table 14). A frequency decrease was recorded after the washing (due to PBS solution). After Clomipramine addition no significant frequency shift was recorded.

Table 14: Shifts of the main parameters of polD1&MLH1 conductance curves after lysate addition and after protein-protein interaction, and relative amount of molecules immobilized on the quartz surface.

polD1&MLH1 conductance curves	C_v^a	ΔD^b	ΔD_N (Hz/mS) ^b	Δf^c (Hz) ^b	m' (μg) ^c
5 min IVTT addition	4.8%	-0.02	-204	-351	1.5
15 min IVTT addition	5.3%	-0.02	965	-976	4.3
Post-capture (5 min)	6.0%	-0.25	46420	-3439	15.0
Post-wash	5.2%	-0.17	11638	-2106	9.2
MM+polD1&MLH1+Clomipramine	7.1%	-0.17	10100	-2035	-0.3*

^a Coefficient of variation. We performed two different experiments.

^b D factor, normalized D factor and frequency shifts (ΔD , ΔD_N and $\Delta f'$) with respect the values immediately after IVTT lysate addition (see table 9).

^c Amount of molecules immobilized on the QC surface (calculated through equation 5 respect the frequency recorded immediately after IVTT addition).

* These values were obtained considering the frequency shifts respect post-wash curve, since the QC was in contact with PBS solution.

This result confirms that no interaction occurs between polD1 or MLH1 and Clomipramine, as expected.

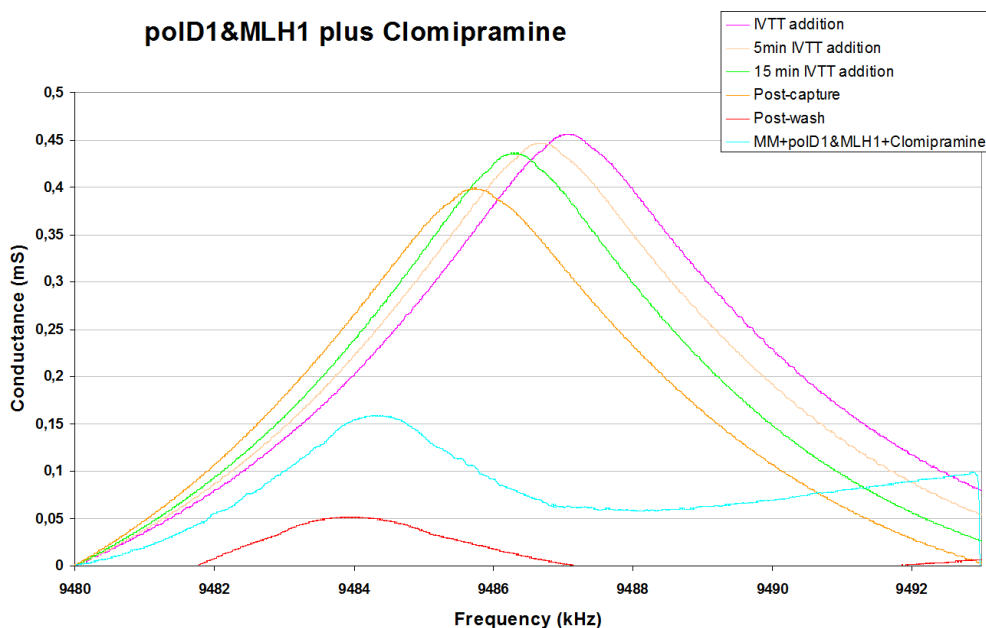


Figure 38 Conductance curves of polD1 and MLH1 expressing QC. The curves were collected in different steps of NAPPA process, as reported in the legend, and for polD1 and MLH1 expressing QC, after the addition of Clomipramine solution, as negative control.

Chapter 6. Medical applications

As reported in the literature C_V values are usually very low, confirming the repeatability of the experiments and the validity and portability of the technique. In our hands, NAPPA-based QCM_D proved to have an intra-assay overall C_V of 5 % (range 3.3 -8.0 %).

In this second part, we report and discuss some preliminary results of protein expression of genes related to cancer. Experiments have been carried out coupling Nucleic Acid Programmable Protein Array (NAPPA) with a recently improved nanogravimetric apparatus which exploits the quartz crystal microbalance with frequency (QCM_F) and quartz crystal microbalance with dissipation monitoring (QCM_D) technologies. The selected proteins are BRIP1, JUN and ATF2 and their role and biological roles will be discussed further in this chapter.

We selected BRIP1, JUN and ATF2 for their clinical relevance.

FANCF/BRIP1/BACH1 (Fanconi anemia complementation group J protein/BRCA1-interacting protein 1/BRCA1-associated C-terminal helicase 1) is a 1249 aminoacids protein involved in breast, cervical and ovarian cancers, as well as in Fanconi anemia, which is a genetic cause of anemia, leading to acute myelogenous leukemia or to bone marrow failure. It is a DNA-dependent ATPase and helicase of the family RecQ DEAH, which interacts with the BRCT repeats domain of BRCA1. As helicase, it has an important role in preserving the chromatin structure and in facilitating and driving the replication process. It is encoded by the locus 17q22–q24, which is 184 kilobases long and has 19 exons, and is situated near to the locus encoding the BRCA1 protein.

JUN encodes the c-Jun protein, which is a well-known oncogene, being the putative transforming gene of avian sarcoma virus 17. It encodes indeed

a protein with a strong similarity to the viral protein. It is encoded by the locus 1p32-p31 and its alteration leads to several malignancies .

ATF2 gene (chromosome 2q32) encodes a 505 amino-acids long transcription factor, that belongs to the leucine zipper family of DNA-binding proteins. Also ATF2 is an important component of signaling pathways whose alterations are a cause of malignant transformation .

We analyzed the conductance curves acquired in NAPPA-QCs in different steps of the expressing and capturing process, for each conductance curve for NAPPA-QCs expressing BRIP1 and JUN (before and after ATF2 addition). The curves were acquired both before the expression of NAPPA and after the protein expression/capture and washing process. The curves have been centered to their maximum frequency to better visualize the changing in bandwidths and conductance. These data pointed to a unique conductance curve shape for each protein and suggested the possibility to identify the expressed proteins by QCM-D even when combined on the same expressing QC. The conductance curves were acquired in different steps of NAPPA expression process.

Using the calibration coefficient, we estimated the amount of molecules immobilized on the quartz surface at the end of immobilization protocol (m).

The values of immobilized BRIP1 molecules, therefore, was 2.28 μg while the amount of JUN was 1.69 μg . In terms of molar concentration, we found 0.02 nM for BRIP1 (having a molar weight of $\sim 130\text{kDa}$) and 0.04 nM for JUN (having a molar weight of $\sim 39\text{kDa}$).

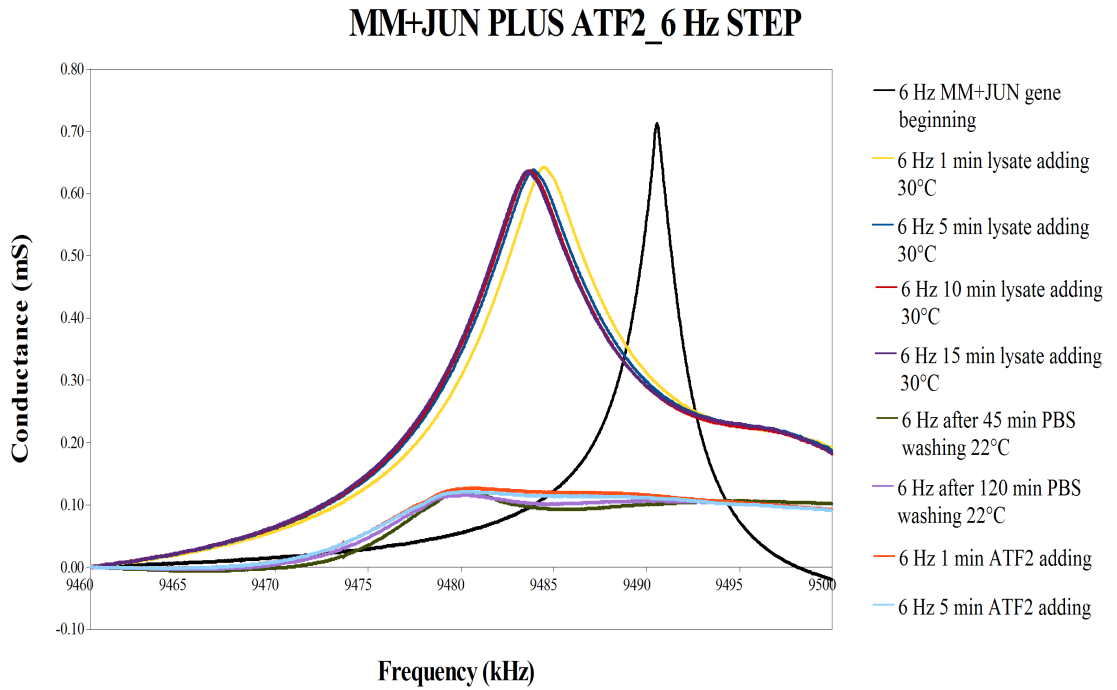


Figure 39 Conductance curves of JUN expressing QC. The curves were collected in different steps of NAPPA process, as reported in the legend, and for JUN expressing QC, after the addition of ATF2 protein.

7. CONCLUSIONS.

In this thesis, we have presented the results obtained from bioinformatics and from applying our innovative conductometer, realized by combining NAPPA technology with QCM_D, to the characterization and analytical quantification of protein expression and protein-protein interactions in a multiparametric way, taking advantage of the multiple information provided by the analysis of the conductance curves (i.e. conductance, viscoelasticity and adsorbed mass).

In the case of the bioinformatics work we were able to properly prioritize genes, in order to select few of them, that being hub genes may prove to be clinically relevant and significant. Further clinical studies will be needed to confirm our findings.

In the case of the NAPPA-based biosensor, we were able to find and obtain pieces of information that, as stated by Guerrini and collaborators [39], until today very few quantitative data oncogenes expression are present in the literature, which presents only relative changes in magnitude orders during the carcinogenesis.

The multi-spot feature of our sensor could be further exploited for investigating the multi-protein expression and protein-protein interactions, becoming more clinically relevant. In fact, not single biomarkers are able to capture the diagnosis and prognosis of cancer, but a composite panel of proteins that could be properly detected via high-throughput sensors.

Future perspectives of this biosensor regard its clinical introduction for assessing cancer prognosis and make a personalized diagnosis and/or to deliver an individualized treatment. In particular, different mutations of the

Conclusions

selected genes will be evaluated in terms of protein instability, functioning, and clinical implications for cancer and other diseases prognosis.

Appendix

8. APPENDIX I: TABLE OF RMSD OF PDB INSULIN STRUCTURES FROM 4IHN LB INSULIN.

INSULIN PDB CAPTION	RMSD DEVIATION FROM LB INSULIN	METHOD	RESOLUTION
2C8R	0.06	Insuline (60 sec) and UV laser excited fluorescence classical vapor hanging drop	1.50
2C8Q	0.10	Insuline (1 sec) and UV laser excited fluorescence classical vapor hanging drop	1.95
3I3Z	0.12	Human insulin microgravity	1.60
3I40	0.13	Human insulin microgravity	1.85
3U4N	0.17	A novel covalently linked insulin dimer	1.98
2CEU	0.53	Despentapeptide insulin in acetic acid (pH 2)	1.80
1BEN	0.63	Human insulin microgravity	1.40
1GUJ	0.65	Insulin at pH 2: structural analysis of the conditions	1.62

Appendix

1HTV		promoting insulin fibre formation Crystal structure of destriptide (B28-B30) insulin	1.90
1B9E	0.69	Human insulin mutant SERB9GLU	2.50
3 ^E 7Y	0.70	Structure of human insulin	1.60
1TYL		The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide	1.90
1TYM	0.71	The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide	1.90
3E7Z	0.72	Structure of human insulin	1.70
3V1G		Forestalling insulin fibrillation by insertion of a chiral clamp mechanism-based application of protein engineering to global health	2.20
2VJZ		Crystal structure form ultalente insulin microcrystals	1.80
2VK0	0.73	Crystal structure form ultalente insulin	2.20

Appendix

2WS4		microcrystals Semi-synthetic analogue of human insulin ProB26-DTI in monomer form	1.90
2WS0	0.74	Semi-synthetic analogue of human insulin NMeAlaB26-insulin at pH 7.5	2.10
1LPH		Lys(B28)Pro(B29)-human insulin	2.30
1G7B	0.75	1.3 A structure of T3R3 human insulin at 100 K	1.30
2WRU		Semi-synthetic highly active analogue of human insulin NME AlaB26-DTI-NH2	1.57
2WRX		Semi-synthetic analogue of human insulin NMeAlaB26-insulin at pH 3.0	1.50
2WS1		Semi-synthetic analogue of human insulin NMeTyrB26-insulin in monomer form	1.60
2WRV	0.76	Semi-synthetic highly active analogue of human insulin NMeHisB26-DTI- NH2	2.15

Appendix

2WRW		Semi-synthetic highly active analogue of human insulin NMeHisB26-DTI- NH2	2.41
1G7A	0.77	1.2 A structure of T3R3 human insulin at 100 K	1.20
3P2X	0.81	Insulin fibrillation is the Janus face of induced fit. A chiral clamp stabilizes the native state at the expense of activity	2.00
2W44	0.82	Structure DELTAA1-A4 insulin	2.00
3V19	0.83	Forestalling insulin fibrillation by insertion of a chiral clamp mechanism-based application of protein engineering to global health	2.00
2R36	0.84	Crystal structure of ni human ARG-insulin	2.00
3KQ6	0.85	Enhancing the therapeutic properties of a protein by a designed zinc-binding site, structural principles of a novel long-acting insulin analog	1.90
2OMI	0.91	Structure of human insulin cocrystallized with protamine	2.24
3IR0	0.94	Crystal Structure of Human	2.20

Appendix

		Insulin complexed with Cu ⁺² metal ion	
2QIU	1.01	Structure of Human Arg-Insulin	2.00
1QIY	1.02	Human insulin hexamers with chain B His mutated to Tyr complexed with phenol	2.30
2OMG	1.03	Structure of human insulin cocrystallized with protamine and urea	1.52
2R35		Crystal structure of RB human arg-insulin	2.08
2OM1		Structure of human insulin in presence of thiocyanate at pH 6.5	1.97
3Q6E	1.04	Human insulin in complex with cucurbit[7]uril	2.05
1ZNJ		Insulin, monoclinic crystal form	2.00
3P33	1.06	Insulin fibrillation is the Janus face of induced fit. A chiral clamp stabilizes the native state at the expense of activity	2.30

Appendix

2R34		Crystal structure of MN human arg-insulin	2.25
2OM0	1.07	Structure of human insulin in presence of urea at pH 6.5	2.05
1OS3		Dehydrated T6 human insulin at 100 K	1.95
3BRR	1.10	Crystal Structure of Insulin in Complex with Sulfatide	1.70
3ILG	1.12	Crystal structure of human insulin Sr+2 complex	1.90
3ZQR		NME PheB25 insulin analogue crystal structure	1.90
1MSO		T6 Human Insulin at 1.0 A Resolution	1.00
3INC	1.13	Crystal structure of human insulin with Ni+2 complex	1.85
1ZEG		Structure of B28 Asp insulin complex with phenol	1.60
3TT8		Crystal Structure Analysis of Cu Human Insulin	1.12

Appendix

3EXX		Derivative Structure of the T6 human insulin derivative with nickel at 1.35 A resolution	1.35
2OMH	1.14	Structure of human insulin cocrystallized with ARG-12 peptide in presence of urea	1.36
1QJ0	1.16	Human insulin hexamers with chain B His mutated to Tyr	2.40
3ZU1	1.17	Structure of LysB29(N epsilon omega-carboxyheptadecanoyl)des(B 30) human insulin	1.60
2WS6	1.19	Semi-synthetic analogue of human insulin NME TyrB26-insulin in hexamer form	1.50
3ROV		Insulin's biosynthesis and activity have opposing structural requirements: a new factor in neonatal diabetes mellitus	2.30
1QIZ		Human insulin hexamers with chain B His mutated to Tyr complexed with resorcinol	2.00
1XDA	1.21	Structure of insulin	1.80

Appendix

2OLZ	1.22	Structure of human insulin in presence of thiocyanate at pH 7.0	1.70
2OLY	1.23	Structure of human insulin in presence of urea at pH 7.0	1.70
2WS7		Semi-synthetic analogue of human insulin ProB26-DTI	2.59
1UZ9	1.26	Crystallographic and solution studies of N-lithocholyl insulin: a new generation of prolonged acting insulins	1.60
1ZEH	1.28	Structure of insulin	1.50
1OS4	1.34	Dehydrated T6 human insulin at 295 K microgravity	2.25
3ZS2	1.37	TyrB25,NMEPheB26,LysB28,ProB29-insulin analogue crystal structure	1.97
1W8P	1.41	Structural properties of the B25Tyr-NME-B26he insulin mutant	2.08
2G56	1.70	Crystal structure of human insulin-degrading enzyme in complex with insulin B chain	2.20
3UTT	1.72	1E6-A*0201-ALWGPDPAAA Complex, Monoclinic	2.60

Appendix

3UTQ		1E6-A*0201- ALWGPDPA AAA Complex, Monoclinic	1.67
2WC0	1.81	Crystal structure of human insulin degrading enzyme in complex with iodinated insulin	2.80
2WBY	1.89	Crystal structure of human insulin degrading enzyme in complex with iodinated insulin	2.60
2G54	1.90	Crystal structure of Zn- bound human insulin- degrading enzyme in complex with insulin B chain	02.25.00
3UTS	2.02	1E6-A*0201- ALWGPDPA AAA Complex, Monoclinic	2.71

9. BIBLIOGRAPHY

1. Adami M, Sallam S, Eggenhoffner E, Sartore M, Hainsworth E, LaBaer J, Nicolini C. Label-free NAPPA via nanogravimetry. In: C. Nicolini, J. LaBaer (Eds.): *Functional Proteomics and Nanotechnology-based Microarrays*. Pan Stanford Series on Nanobiotechnology, vol. 2 (Chapter 4), Published 1st July 2010 by Pan Stanford Publishing, London–New York–Singapore, 2010, pp. 95–108.
2. Agha-Hosseini F, Mirzaii-Dizgah I. p53 as a neoplastic biomarker in patients with erosive and plaque like forms of oral lichen planus. *J Contemp Dent Pract*. 2013 Jan 1;14(1):1-3.
3. Agrawal GK, Timperio AM, Zolla L, Bansal V, Shukla R, Rakwal R. Biomarker discovery and applications for foods and beverages: Proteomics to nanoproteomics. *J Proteomics*. 2013 Nov 20;93:74-92.
4. Alexiou P, Vergoulis T, Gleditsch M, Prekas G, Dalamagas T, Megraw M, Grosse I, Sellis T, Hatzigeorgiou AG. miRGen 2.0: a database of microRNA genomic information and regulation. *Nucleic Acids Res*. 2010 Jan;38(Database issue):D137-41.
5. Allen M, Spiegel AM, Hawkins M. ‘Personalized medicine’ to identify genetic risks for type 2 diabetes and focus prevention: can it fulfill its promise? *Health Aff*. 2012;31:143–149.
6. Anderson KS, Sibani S, Wallstrom G, Qiu J, Mendoza EA, Raphael J, Hainsworth E, Montor WR, Wong J, Park JG, Lokko N, Logvinenko T, Ramachandran N, Godwin AK, Marks J, Engstrom P, Labaer J. Protein microarray signature of autoantibody

References

- biomarkers for the early detection of breast cancer. *J Proteome Res.* 2011 Jan 7;10(1):85-96.
7. Andreasen JO. Oral lichen planus: a clinical evaluation of 115 cases. *Oral Surg Oral Med Oral Pathol.* 1968; 25:31-42.
 8. An N, Rausch-fan X, Wieland M, Matejka M, Andrukhov O, Schedle A. Initial attachment, subsequent cell proliferation/viability and gene expression of epithelial cells related to attachment and wound healing in response to different titanium surfaces. *Dental Materials* 2012;28:1207–14.
 9. Anderson K, Sibani S, Wallstrom G, Mendoza E, Hainsworth E, Montor W, Mong J, Park I, Lokko N, Logvinenko T, Ramachandran N, Engstrom P, Labaer J. Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer. *J. Proteome Res.* 2011;10(1):85–96.
 10. Andreadis D, Epivatianos A, Pouloupoulos A, Nomikos A, Christidis K, Papazoglou G, Antoniadis D, Barbatis C. Immunohistochemical detection of the expression of the cell adhesion molecules E-cadherin, desmoglein-2, beta4-integrin, ICAM-1 and HCAM (CD44s) in Warthin's tumour of the parotid gland. *Oral Oncology.* 2005;41:799–805.
 11. Anuradha CH, Reddy BV, Nandan SR, Kumar SR. Oral lichen planus. A review. *New York State Dental Journal.* 2008;74:66–68.
 12. Antonini M, Ghisellini P, Paternolli C, Nicolini C. Electrochemical study of the interaction between cytochrome P450sccK201E and cholesterol. *Talanta.* 2004 Apr 19;62(5):945-50.

References

13. Arrais JP, Rosa N, Melo J, Coelho ED, Amaral D, Correia MJ, Barros M, Oliveira JL. OralCard: a bioinformatic tool for the study of oral proteome. *Arch Oral Biol*. 2013 Jul;58(7):762-72.
14. Arya SK, Datta M, Malhotra BD. Recent advances in cholesterol biosensor. *Biosens Bioelectron*. 2008 Feb 28;23(7):1083-100.
15. Auffray C, Charron D, Hood L. Predictive, preventive, personalized and participatory medicine: back to the future. *Genome Med*. 2010 Aug 26;2(8):57.
16. Barabási AL. Network medicine from obesity to the “diseasome”. *N. Engl. J. Med*. 2007;357:404–407.
17. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet*. 2011 Jan;12(1):56-68.
18. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet*. 2004 Feb;5(2):101-13.
19. Barabasi AL, Ravasz E, Vicsek T. 2001. Deterministic scale-free networks. *Phys A Stat Mech Appl* 299(3–4):559–564.
20. Baum BJ, Yates JR 3rd, Srivastava S, Wong DT, Melvin JE. Scientific frontiers: emerging technologies for salivary diagnostics. *Adv Dent Res*. 2011 Oct;23(4):360-8.
21. Boja E, Hiltke T, Rivers R, Kinsinger C, Rahbar A, Mesri M, Rodriguez H. Evolution of clinical proteomics and its role in medicine. *J Proteome Res*. 2011 Jan 7;10(1):66-84.
22. Boorghani M, Gholizadeh N, Taghavi Zenouz A, Vatankhah M, Mehdipour M. Oral lichen planus: clinical features, etiology,

References

- treatment and management; a review of literature. *J Dent Res Dent Clin Dent Prospects*. 2010 Winter;4(1):3-9.
23. Bozdaganyan M, Bragazzi N, Pechkova E, Orekhov P, Shaytan K, Nicolini C. Molecular Dynamics for nanobiotechnology. In Nicolini C (Ed): *Nanobiotechnology in energy, environment and electronics: methods and applications*. Pan Stanford Series on Nanobiotechnology, vol. 4 (Chapter 5). In press by Pan Stanford Publishing, London-New York-Singapore, 2014.
 24. Bradshaw L. Understanding piezoelectric quartz crystals. RF time and frequency. 2000;8:50–58.
 25. Braud C, Baeten D, Giral M, Pallier A, Ashton-Chess J, Braudeau C, Chevalier C, Lebars A, Léger J, Moreau A, Pechkova E, Nicolini C, Soullillou JP, Brouard S. Immunosuppressive drug-free operational immune tolerance in human kidney transplant recipients: Part I. Blood gene expression statistical analysis. *J Cell Biochem*. 2008 Apr 15;103(6):1681-92.
 26. Bragazzi NL. From P0 to P6 medicine, a model of highly participatory, narrative, interactive, and "augmented" medicine: some considerations on Salvatore Iaconesi's clinical story. *Patient Prefer Adherence*. 2013 Apr 24;7:353-9.
 27. Bragazzi NL, Nicolini C. A Leader Genes Approach-based Tool for Molecular Genomics: From Gene-ranking to Gene-network Systems Biology and Biotargets Predictions. *J Comput Sci Syst Biol*. 2013; 6:165-176.

References

28. Bragazzi NL, Giacomelli L, Sivozhelezov V, Nicolini C. Leader Gene: A fast data-mining tool for molecular genomics. *J Proteomics Bioinform.* 2011;4(4):83-86.
29. Braud C, Baeten D, Giral M, Pallier A, Ashton-Chess J, Braudeau C, Chevalier C, Lebars A, Léger J, Moreau A, Pechkova E, Nicolini C, Soullillou JP, Brouard S. Immunosuppressive drug-free operational immune tolerance in human kidney transplant recipients: Part I. Blood gene expression statistical analysis. *J Cell Biochem.* 2008 Apr 15;103(6):1681-92.
30. Brouard S, Mansfield E, Braud C, Li L, Giral M, Hsieh SC, Baeten D, Zhang M, Ashton-Chess J, Braudeau C, Hsieh F, Dupont A, Pallier A, Moreau A, Louis S, Ruiz C, Salvatierra O, Soullillou JP, Sarwal M. Identification of a peripheral blood transcriptional biomarker panel associated with operational renal allograft tolerance. *Proc Natl Acad Sci U S A.* 2007 Sep 25;104(39):15448-53.
31. Burbelo PD, Bayat A, Lebovitz EE, Iadarola MJ. New technologies for studying the complexity of oral diseases. *Oral Dis.* 2012 Mar;18(2):121-6.
32. Carrozzo M, Gandolfo S. Oral diseases possibly associated with hepatitis C virus. *Crit Rev Oral Biol Med.* 2003;14(2):115-27.
33. Chandarana SP, Lee JS, Chanowski EJ, Sacco AG, Bradford CR, Wolf GT, Prince ME, Moyer JS, Eisbruch A, Worden FP, Giordano TJ, Kumar B, Cordell KG, Carey TE, Chepeha DB. Prevalence and predictive role of p16 and epidermal growth factor receptor in

References

- surgically treated oropharyngeal and oral cavity cancer. *Head Neck*. 2013 Aug;35(8):1083-90.
34. Chen B, Butte AJ. Network medicine in disease analysis and therapeutics. *Clin Pharmacol Ther*. 2013 Dec;94(6):627-9.
35. Chen SH, Lin MY. Quantitative NanoProteomics approach for protein complex (QNanoPX) using gold nanoparticle-based DNA probe. *Methods Mol Biol*. 2013;977:183-92.
36. Chen JY, Piquette-Miller M, Smith BP. Network medicine: finding the links to personalized therapy. *Clin Pharmacol Ther*. 2013 Dec;94(6):613-6.
37. Chen Y, Zhang W, Geng N, Tian K, Jack Windsor L. MMPs, TIMP-2, and TGF-beta1 in the cancerization of oral lichen planus. *Head Neck*. 2008 Sep;30(9):1237-45.
38. Cheng YS, Jordan L, Rees T, Chen HS, Oxford L, Brinkmann O, Wong D. Levels of potential oral cancer salivary mRNA biomarkers in oral cancer patients in remission and oral lichen planus patients. *Clin Oral Investig*. 2013 Jul 28. [Epub ahead of print]
39. Cheng YS, Rees T, Jordan L, Oxford L, O'Brien J, Chen HS, Wong D. Salivary endothelin-1 potential for detecting oral cancer in patients with oral lichen planus or oral cancer in remission. *Oral Oncol*. 2011 Dec;47(12):1122-6.
40. Cheriyan VT, Thomas C, Balaram P. Augmentation of T-cell immune responses and signal transduction proteins in oral cancer patients: potential for IL-2-mediated immunotherapy. *Journal of Cancer Research and Clinical Oncology*. 2011;137(10):1435-44.

References

41. Chiappelli F, Covani U, Giacomelli L. Proteomics as it pertains to oral pathologies and dental research. *Bioinformatics*. 2011 Jan 6;5(7):277.
42. Chikkaveeraiah BV, Mani V, Patel V, Gutkind JS, Rusling JF. Microfluidic electrochemical immunoarray for ultrasensitive detection of two cancer biomarker proteins in serum. *Biosens Bioelectron*. 2011 Jul 15;26(11):4477-83.
43. Cohen JC, Hobbs HH. Genetics. Simple genetics for a complex disease. *Science*. 2013 May 10;340(6133):689-90
44. Collings FB, Vaidya VS. Novel technologies for the discovery and quantitation of biomarkers of toxicity. *Toxicology*. 2008 Mar 20;245(3):167-74.
45. Covani U, Marconcini S, Giacomelli L, Sivozhelevov V, Barone A, Nicolini C. Bioinformatic prediction of leader genes in human periodontitis. *J Periodontol*. 2008 Oct;79(10):1974-83.
46. D'Agati VD, Mengel M. The rise of renal pathology in nephrology: structure illuminates function. *Am J Kidney Dis*. 2013 Jun;61(6):1016-25.
47. Dang J, Bian YQ, Sun JY, Chen F, Dong GY, Liu Q, Wang XW, Kjems J, Gao S, Wang QT. MicroRNA-137 promoter methylation in oral lichen planus and oral squamous cell carcinoma. *J Oral Pathol Med*. 2013 Apr;42(4):315-21.
48. Danielsson K, Ebrahimi M, Wahlin YB, Nylander K, Boldrup L. Increased levels of COX-2 in oral lichen planus supports an autoimmune cause of the disease. *J Eur Acad Dermatol Venereol*. 2012 Nov;26(11):1415-9.

References

49. Danielsson K, Wahlin YB, Gu X, Boldrup L, Nylander K. Altered expression of miR-21, miR-125b, and miR-203 indicates a role for these microRNAs in oral lichen planus. *J Oral Pathol Med*. 2012 Jan;41(1):90-5.
50. Dasilva N, Díez P, Matarraz S, González-González M, Paradinas S, Orfao A, Fuentes M. Biomarker discovery by novel sensors based on nanoproteomics approaches. *Sensors*. 2012;12:2284–2308.
51. De Cecco L, Dugo M, Canevari S, Daidone MG, Callari M. Measuring microRNA expression levels in oncology: from samples to data analysis. *Crit Rev Oncog*. 2013;18(4):273-87.
52. de Sousa FA, Paradella TC, Carvalho YR, Rosa LE. Comparative analysis of the expression of proliferating cell nuclear antigen, p53, bax, and bcl-2 in oral lichen planus and oral squamous cell carcinoma. *Ann Diagn Pathol*. 2009 Oct;13(5):308-12.
53. Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R. A systems biology approach for pathway level analysis. *Genome Res*. 2007 Oct;17(10):1537-45.
54. Dufrêne YF. Atomic force microscopy: a powerful molecular toolkit in nanoproteomics. *Proteomics*. 2009 Dec;9(24):5400-5.
55. Ebrahimi M, Boldrup L, Coates PJ, Wahlin YB, Bourdon JC, Nylander K. Expression of novel p53 isoforms in oral lichen planus. *Oral Oncol*. 2008 Feb;44(2):156-61.
56. Ebrahimi M, Boldrup L, Wahlin YB, Coates PJ, Nylander K. Decreased expression of the p63 related proteins betacatenin, E-cadherin and EGFR in oral lichen planus. *Oral Oncology* 2008;44:634–8.

References

57. Ebrahimi M, Nylander K, van der Waal I. Oral lichen planus and the p53 family: what do we know? *J Oral Pathol Med.* 2011 Apr;40(4):281-5.
58. Efeyan A, Serrano M. p53: guardian of the genome and policeman of the oncogenes. *Cell Cycle.* 2007 May 2;6(9):1006-10.
59. Emmert-Streib F, Dehmer M. Enhancing systems medicine beyond genotype data by dynamic patient signatures: having information and using it too. *Front Genet.* 2013 Nov 19;4:241.
60. Eng G, Chen A, Vess T, Ginsburg GS. Genome technologies and personalized dental medicine. *Oral Dis.* 2012 Apr;18(3):223-35.
61. Epstein JB, Wan LS, Gorsky M, Zhang L. Oral lichen planus: progress in understanding its malignant potential and the implications for clinical management. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2003 Jul;96(1):32-7.
62. Ettl T, Baader K, Stiegler C, Müller M, Agaimy A, Zenk J, Kühnel T, Gosau M, Zeitler K, Schwarz S, Brockhoff G. Loss of PTEN is associated with elevated EGFR and HER2 expression and worse prognosis in salivary gland cancer. *Br J Cancer.* 2012 Feb 14;106(4):719-26.
63. Ettl T, Stiegler C, Zeitler K, Agaimy A, Zenk J, Reichert TE, Gosau M, Kühnel T, Brockhoff G, Schwarz S. EGFR, HER2, survivin, and loss of pSTAT3 characterize high-grade malignancy in salivary gland cancer with impact on prognosis. *Hum Pathol.* 2012 Jun;43(6):921-31.

References

64. Faratian D, Clyde RG, Crawford JW, Harrison DJ. Systems pathology--taking molecular pathology into a new dimension. *Nat Rev Clin Oncol*. 2009 Aug;6(8):455-64.
65. Feley K, Shan X, Tao N. Surface impedance imaging technique. *Anal. Chem*. 2008;80:5146–5151.
66. Festa F, Rollins SM, Vattem K, Hathaway M, Lorenz P, Mendoza EA, Yu X, Qiu J, Kilmer G, Jensen P, Webb B, Ryan ET, LaBaer J. Robust microarray production of freshly expressed proteins in a human milieu. *Proteomics Clin Appl*. 2013;7:372-377.
67. Gaffen SL, Liu KD. Overview of interleukin-2 function, production and clinical applications. *Cytokine*. 2004 Nov 7;28(3):109-23.
68. Gallo A, Alevizos I. Isolation of circulating microRNA in saliva. *Methods Mol Biol*. 2013;1024:183-90.
69. Gandarillas A. Epidermal differentiation, apoptosis, and senescence: common pathways? *Exp Gerontol*. 2000 Feb;35(1):53-62.
70. Gandolfo S, Richiardi L, Carrozzo M, Broccoletti R, Carbone M, Pagano M, Vestita C, Rosso S, Merletti F. Risk of oral squamous cell carcinoma in 402 patients with oral lichen planus: a follow-up study in an Italian population. *Oral Oncol*. 2004 Jan;40(1):77-83.
71. Garcia I, Kuska R, Somerman MJ. Expanding the foundation for personalized medicine: implications and challenges for dentistry. *J Dent Res*. 2013 Jul;92(7 Suppl):3S-10S.
72. Garibaldi S, Brunelli C, Bavastrello V, Ghigliotti G, Nicolini C. Carbon nanotube biocompatibility with cardiac muscle cells. *Nanotechnology*. 2006;17:391–397.

References

73. Gassling V, Hampe J, Açil Y, Braesen JH, Wiltfang J, Häsler R. Disease-Associated miRNA-mRNA Networks in Oral Lichen Planus. *PLoS One*. 2013 May 27;8(5):e63015.
74. Gebhardt R, Pechkova E, Riekkel C, Nicolini C. In situ muGISAXS: II. Thaumatin crystal growth kinetic. *Biophys J*. 2010 Aug 9;99(4):1262-7
75. Georgakopoulou EA, Troupis TG, Troupis G, Gorgoulis VG. Update of the cancer-associated molecular mechanisms in oral lichen planus, a disease with possible premalignant nature. *J BUON*. 2011 Oct-Dec;16(4):613-6.
76. Giacomelli L, Nicolini C. Gene expression of human T lymphocytes cell cycle: experimental and bioinformatic analysis. *J Cell Biochem*. 2006 Dec 1;99(5):1326-33.
77. Giacomelli L, Oluwadara O, Chiappe G, Barone A, Chiappelli F, Covani U. Relationship between human oral lichen planus and oral squamous cell carcinoma at a genomic level: a datamining study. *Bioinformatics*. 2009 Dec 31;4(6):258-62.
78. Giannelli G, Brassard J, Foti C, Stetler-Stevenson WG, Falk-Marzillier J, Zambonin-Zallone A, Schiraldi O, Quaranta V. Altered expression of basement membrane proteins and their integrin receptors in lichen planus: possible pathogenetic role of gelatinases A and B. *Lab Invest*. 1996 Jun;74(6):1091-104.
79. Giannelli G, Milillo L, Marinosci F, Lo Muzio L, Serpico R, Antonaci S. Altered expression of integrins and basement membrane proteins in malignant and pre-malignant lesions of oral mucosa. *J Biol Regul Homeost Agents*. 2001 Oct-Dec;15(4):375-80.

References

80. Glick M. Personalized oral health care: providing '-omic' answers to oral health care queries. *J Am Dent Assoc.* 2012 Feb;143(2):102-4.
81. Gold DV, Modrak DE, Ying Z, Cardillo TM, Sharkey RM, Goldenberg DM. New MUC1 serum immunoassay differentiates pancreatic cancer from pancreatitis. *J Clin Oncol.* 2006 Jan 10;24(2):252-8. Epub 2005 Dec 12.
82. Gold L, Walker JJ, Wilcox SK, Williams S. Advances in human proteomics at high scale with the SOMAscan proteomics platform. *New Biotechnol.* 2012;29(5):543–549.
83. Gonzalez-Gonzalez M, Jara-Acevedo R, Matarraz S, Jara-Acevedo M, Paradinas S, Sayagües JM, Orfao A, Fuentes M. Nanotechniques in proteomics: protein microarrays and novel detection platforms. *Eur J Pharm Sci.* 2012 Mar 12;45(4):499-506.
84. Grasso V, Lambertini V, Ghisellini P, Valerio F, Stura E, Perlo P, Nicolini C. Nanostructuring of a porous alumina matrix for a biomolecular microarray. *Nanotechnology.* 2006;17:795–798.
85. Gutiérrez-Venegas G, Castillo-Alemán R. Characterization of the transduction pathway involved in c-fos and c-jun expression induced by *Aggregatibacter actinomycetemcomitans* lipopolysaccharides in human gingival fibroblasts. *Int Immunopharmacol.* 2008 Nov;8(11):1513-23.
86. Hai T, Curran T. Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. *Proc Natl Acad Sci U S A.* 1991 May 1;88(9):3720-4.
87. Häkkinen L, Kainulainen T, Salo T, Grenman R, Larjava H. Expression of integrin alpha9 subunit and tenascin in oral

References

- leukoplakia, lichen planus, and squamous cell carcinoma. *Oral Dis.* 1999 Jul;5(3):210-7.
88. Hamidi S, Salo T, Kainulainen T, Epstein J, Lerner K, Larjava H. Expression of alpha(v)beta6 integrin in oral leukoplakia. *British Journal of Cancer.* 2000;82:1433–40.
89. Hanash S, Taguchi A. Application of proteomics to cancer early detection. *Cancer J.* 2011 Nov-Dec;17(6):423-8.
90. Hirota M, Ito T, Okudela K, Kawabe R, Yazawa T, Hayashi H, et al. Cell proliferation activity and the expression of cell cycle regulatory proteins in oral lichen planus. *Journal of Oral Pathology and Medicine.* 2002;31:204–12.
91. Hirschhorn JN. Genetic approaches to studying common diseases and complex traits. *Pediatr Res.* 2005 May;57(5 Pt 2):74R-77R.
92. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sørensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature.* 2002 Jan 10;415(6868):180-3.
93. Hood L. Systems biology and p4 medicine: past, present, and future. *Rambam Maimonides Med J.* 2013 Apr 30;4(2):e0012.

References

94. Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nat Rev Clin Oncol*. 2011 Mar;8(3):184-7.
95. Höök F, Vörös J, Rodahl M, Kurrat R, Böni P, Ramsden JJ, Textor M, Spencer ND, Tengvall P, Gold J, Kasemo B. A comparative study of protein adsorption on titanium oxide surfaces using in situ ellipsometry, optical waveguide lightmode spectroscopy, and quartz crystal microbalance/dissipation, *Colloids Surf. B Biointerfaces*. 2002;24:155–170.
96. Hu JY, Zhang J, Cui JL, Liang XY, Lu R, Du GF, Xu XY, Zhou G. Increasing CCL5/CCR5 on CD4+ T cells in peripheral blood of oral lichen planus. *Cytokine*. 2013 Apr;62(1):141-5.
97. Huber TB, Kottgen M, Schilling B, Walz G, Benzing T. Interaction with podocin facilitates nephrin signaling. *J Biol Chem*. 2001 Nov 9;276(45):41543-6.
98. Hueber W, Robinson WH. Proteomic biomarkers for autoimmune disease. *Proteomics*. 2006 Jul;6(14):4100-5.
99. Hughes AJ, Lin RK, Peehl DM, Herr AE. Microfluidic integration for automated targeted proteomic assays. *Proc Natl Acad Sci U S A*. 2012 Apr 17;109(16):5972-7.
100. Ichimura M, Hiratsuka K, Ogura N, Utsunomiya T, Sakamaki H, Kondoh T, Abiko Y, Otake S, Yamamoto M. Expression profile of chemokines and chemokine receptors in epithelial cell layers of oral lichen planus. *J Oral Pathol Med*. 2006 Mar;35(3):167-74.

References

101. Ismail SB, Kumar SK, Zain RB. Oral lichen planus and lichenoid reactions: etiopathogenesis, diagnosis, management and malignant transformation. *Journal of Oral Science*. 2007;49:89–106.
102. Jain KK. *Handbook of Nanomedicine*, Humana/Springer, Totowa, NJ, 2008.
103. Janjić V, Pržulj N. The Core Diseasesome. *Mol Biosyst*. 2012 Oct;8(10):2614-25.
104. Jia L, Lu Y, Shao J, Liang XJ, Xu Y. Nanoproteomics: a new sprout from emerging links between nanotechnology and proteomics. *Trends Biotechnol*. 2013 Feb;31(2):99-107.
105. Jiang L, Sørensen P, Thomsen B, Edwards SM, Skarman A, Røntved CM, Lund MS, Workman CT. Gene prioritization for livestock diseases by data integration. *Physiol Genomics*. 2012 Mar 1;44(5):305-17.
106. Jones J, Downer CS, Speight PM. Changes in the expression of integrins and basement membrane proteins in benign mucous membrane pemphigoid. *Oral Diseases*. 1995;1:159–65.
107. Jovanovic V, Giacomelli L, Sivozhelezov V, Degauque N, Lair D, Soulillou JP, Pechkova E, Nicolini C, Brouard S. AKT1 leader gene and downstream targets are involved in a rat model of kidney allograft tolerance. *J Cell Biochem*. 2010 Oct 15;111(3):709-19.
108. Kainulainen T, Autio-Harminen H, Oikarinen A, Salo S, Tryggvason K, Salo T. Altered distribution and synthesis of laminin-5 (kalinin) in oral lichen planus, epithelial dysplasias and squamous cell carcinomas. *Br J Dermatol*. 1997 Mar;136(3):331-6.

References

109. Karatsaidis A, Schreurs O, Axell T, Hegeland K, Schenck K. Inhibition of the transforming growth factor-B/Smad signaling pathway in the epithelium of the oral lichen. *Journal of Investigative Dermatology*. 2003;121:1283–90.
110. Khan A, Farah CS, Savage NW, Walsh LJ, Harbrow DJ, Surgeman PB. Th1 cytokines in oral lichen planus. *Journal of Oral Pathology and Medicine*. 2003;32:77–83.
111. Kilpi A, Rich AM, Kontinen YT, Reade PC. Expression of c-erbB-2 protein in keratinocytes of oral mucosal lichen planus and subsequent squamous cell carcinoma. *Eur J Oral Sci*. 1996 Jun;104(3):278-84.
112. Kim SY, Lee JW, Sohn IS. Comparison of various statistical methods for identifying differential gene expression in replicated microarray data. *Stat Methods Med Res*. 2006 Feb;15(1):3-20.
113. Kumagai K, Horikawa T, Gotoh A, Yamane S, Yamada H, Kobayashi H, Hamada Y, Suzuki S, Suzuki R. Up-regulation of EGF receptor and its ligands, AREG, EREG, and HB-EGF in oral lichen planus. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2010 Dec;110(6):748-54.
114. Kurokawa A, Nagata M, Kitamura N, Noman AA, Ohnishi M, Ohyama T, Kobayashi T, Shingaki S, Takagi R; Oral, Maxillofacial Pathology, and Surgery Group. Collaborators. Diagnostic value of integrin alpha3, beta4, and beta5 gene expression levels for the clinical outcome of tongue squamous cell carcinoma. *Cancer*. 2008 Mar 15;112(6):1272-81.

References

115. LaBaer J, Ramachandran N. Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* 2005;9:14–19.
116. Lammers T, Rizzo LY, Storm G, Kiessling F. Personalized nanomedicine. *Clin Cancer Res.* 2012 Sep 15;18(18):4889-94.
117. Lee HJ. Exceptional stories of microRNAs. *Exp Biol Med (Maywood)*. 2013 Apr;238(4):339-43.
118. Lee K-B, Solanki A, Kim JD, Jung J. Nanomedicine: dynamic integration of nanotechnology with biomedical science. In: Zhang M, Xi N (Eds). *Nanomedicine: a Systems Engineering Approach*. Published by Pan Stanford Publishing; London–New York–Singapore, 2009, pp. 1–38.
119. Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell.* 1997 Feb 7;88(3):323-31.
120. Li TJ, Cui J. COX-2, MMP-7 expression in oral lichen planus and oral squamous cell carcinoma. *Asian Pac J Trop Med.* 2013 Aug;6(8):640-3.
121. Leyva-Huerta ER, Ledesma-Montes C, Rojo-Botello RE, Vega-Memije E. P53 and bcl-2 immunoexpression in patients with oral lichen planus and oral squamous cell carcinoma. *Med Oral Patol Oral Cir Bucal.* 2012 Sep 1;17(5):e745-50.
122. Lisa Cheng YS, Jordan L, Gorugantula LM, Schneiderman E, Chen HS, Rees T. Salivary Interleukins 6 and 8 in Oral Cancer Patients and in Patients With Chronic Oral Inflammatory Diseases. *J Periodontol.* 2013 Oct 23. [Epub ahead of print]

References

- 123.Liu H, Deng X, Shyu YJ, Li JJ, Taparowsky EJ, Hu CD. Mutual regulation of c-Jun and ATF2 by transcriptional activation and subcellular localization. *EMBO J*. 2006 Mar 8;25(5):1058-69.
- 124.Lodi G, Carrozzo M, Furness S, Thongprasom K. Interventions for treating oral lichen planus: a systematic review. *Br J Dermatol*. 2012 May;166(5):938-47.
- 125.Lodi G, Scully C, Carrozzo M, Griffiths M, Sugerman PB, Thongprasom K. Current controversies in oral lichen planus: report of an international consensus meeting. Part 1. Viral infections and etiopathogenesis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2005 Jul;100(1):40-51.
- 126.Lodi G, Scully C, Carrozzo M, Griffiths M, Sugerman PB, Thongprasom K. Current controversies in oral lichen planus: report of an international consensus meeting. Part 2. Clinical management and malignant transformation. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2005 Aug;100(2):164-78.
- 127.Loscalzo J. Association studies in an era of too much information: clinical analysis of new biomarker and genetic data. *Circulation*. 2007 Oct 23;116(17):1866-70.
- 128.Loscalzo J, Kohane I, Barabasi AL. Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Mol Syst Biol*. 2007;3:124.
- 129.Lozano JJ, Pallier A, Martinez-Llordella M, Danger R, López M, Giral M, Londoño MC, Rimola A, Soullillou JP, Brouard S, Sánchez-Fueyo A. Comparison of transcriptional and blood cell-phenotypic

References

- markers between operationally tolerant liver and kidney recipients. *Am J Transplant*. 2011 Sep;11(9):1916-26.
130. Ma L, Wang H, Yao H, Zhu L, Liu W, Zhou Z. Bmi1 expression in oral lichen planus and the risk of progression to oral squamous cell carcinoma. *Ann Diagn Pathol*. 2013 Aug;17(4):327-30.
131. Maclellan SA, Lawson J, Baik J, Guillaud M, Poh CF, Garnis C. Differential expression of miRNAs in the serum of patients with high-risk oral lesions. *Cancer Med*. 2012 Oct;1(2):268-74.
132. Madu CO, Lu Y. Novel diagnostic biomarkers for prostate cancer. *J Cancer*. 2010 Oct 6;1:150-77.
133. Manuel Gándara Rey J, Diniz Freitas M. High rate of malignant transformation in atypical oral lichen planus lesions. *Med Oral*. 2003 Nov-Dec;8(5):309.
134. Mangan S, Alon U. Structure and function of the feedforward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(21):11980–5.
135. Marconcini S, Covani U, Barone A, Vittorio O, Curcio M, Barbuti S, Scatena F, Felli L, Nicolini C. Real-time quantitative polymerase chain reaction analysis of patients with refractory chronic periodontitis. *J Periodontol*. 2011 Jul;82(7):1018-24.
136. Marko-Varga G, Ogiwara A, Nishimura T, Kawamura T, Fujii K, Kawakami T, Kyono Y, Tu HK, Anyoji H, Kanazawa M, Akimoto S, Hirano T, Tsuboi M, Nishio K, Hada S, Jiang H, Fukuoka M, Nakata K, Nishiwaki Y, Kunito H, Peers IS, Harbron CG, South MC, Higenbottam T, Nyberg F, Kudoh S, Kato H. Personalized

References

- medicine and proteomics: lessons from non-small cell lung cancer. *J Proteome Res.* 2007 Aug;6(8):2925-35.
137. Martín-Ezquerria G, Salgado R, Toll A, Gilaberte M, Baró T, Alameda Quittlet F, Yébenes M, Solé F, Garcia-Muret M, Espinet B, Pujol RM. Multiple genetic copy number alterations in oral squamous cell carcinoma: study of MYC, TP53, CCND1, EGFR and ERBB2 status in primary and metastatic tumours. *Br J Dermatol.* 2010 Nov;163(5):1028-35.
138. Martínez-Lara I, González-Moles MA, Ruiz-Avila I, Bravo M, Ramos MC, Fernández-Martínez JA. Proliferating cell nuclear antigen (PCNA) as a marker of dysplasia in oral mucosa. *Acta Stomatol Belg.* 1996 Mar;93(1):29-32.
139. Marx V. Biology: The big challenges of big data. *Nature.* 2013 Jun 13;498(7453):255-60.
140. Masoudi-Nejad A, Meshkin A, Haji-Eghrari B, Bidkhorji G. Candidate gene prioritization. *Mol Genet Genomics.* 2012 Sep;287(9):679-98.
141. Matarraz S, Gonzalez-Gonzalez M, Jara M, Orfao A, Fuentes M. New technologies in cancer protein microarrays for biomarker discovery. *Clin. Transl. Oncol.* 2011;13:156–161.
142. Mattila R, Alanen K, Syrjänen S. Desmocollin expression in oral atrophic lichen planus correlates with clinical behavior and DNA content. *J Cutan Pathol.* 2008 Sep;35(9):832-8.
143. Mattsson U, Jontell M, Holmstrup P. Oral lichen planus and malignant transformation: is a recall of patients justified? *Crit Rev Oral Biol Med.* 2002;13(5):390-6.

References

144. Megraw M, Sethupathy P, Corda B, Hatzigeorgiou AG. miRGen: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res.* 2007 Jan;35(Database issue):D149-55.
145. Michael A, Bajracharya SD, Yuen PS, Zhou H, Star RA, Illei GG, Alevizos I. Exosomes from human saliva as a source of microRNA biomarkers. *Oral Dis.* 2010 Jan;16(1):34-8.
146. Mishra A, Bharti AC, Saluja D, Das BC. Transactivation and expression patterns of Jun and Fos/AP-1 super-family proteins in human oral cancer. *International Journal of Cancer.* 2010;126:819–29.
147. Mithani SK, Mydlarz WK, Grumbine FL, Smith IM, Califano JA. Molecular genetics of premalignant oral lesions. *Oral Dis.* 2007 Mar;13(2):126-33.
148. Modolo F, Martins MT, Loduca SV, de Araujo VC. Expression of integrin subunits alpha2, alpha3, alpha5, alphav, beta1, beta3 and beta4 in different histological types of ameloblastoma compared with dental germ, dental lamina and adult lining epithelium. *Oral Diseases* 2004;10:277–82. 7.
149. Mollaoglu N. Oral lichen planus: a review. *Br J Oral Maxillofac Surg.* 2000 Aug;38(4):370-7.
150. Monteiro LS, Diniz-Freitas M, Garcia-Caballero T, Warnakulasuriya S, Forteza J, Fraga M. Combined cytoplasmic and membranous EGFR and p53 overexpression is a poor prognostic marker in early stage oral squamous cell carcinoma. *Journal of Oral Pathology and Medicine.* 2012;41:559–67.

References

151. Nagao T, Warnakulasuriya S. Annual screening for oral cancer detection. *Cancer Detect Prev.* 2003;27(5):333-7.
152. Neale MC, Kendler KS. Models of comorbidity for multifactorial disorders. *Am J Hum Genet.* 1995 Oct;57(4):935-53.
153. Newell KA, Asare A, Kirk AD, Gisler TD, Bourcier K, Suthanthiran M, Burlingham WJ, Marks WH, Sanz I, Lechler RI, Hernandez-Fuentes MP, Turka LA, Seyfert-Margolis VL; Immune Tolerance Network ST507 Study Group. Identification of a B cell signature associated with renal transplant tolerance in humans. *J Clin Invest.* 2010 Jun;120(6):1836-47.
154. Ng JD. Space-grown protein crystals are more useful for structure determination. *Ann N Y Acad Sci.* 2002 Oct;974:598-609.
155. Nicolini C. Nanogenomics for medicine. *Nanomedicine (Lond).* 2006 Aug;1(2):147-52.
156. Nicolini C. Nanogenomics in medicine. *Wiley Interdiscip Rev Nanomed Nanobiotechnol.* 2010 Jan-Feb;2(1):59-76.
157. Nicolini C, Adami M, Sartore M, Bragazzi NL, Bavastrello V, Spera R, Pechkova E. Prototypes of newly conceived inorganic and biological sensors for health and environmental applications. *Sensors (Basel).* 2012 Dec 12;12(12):17112-27.
158. Nicolini C, Bragazzi N, Pechkova E. Nanoproteomics enabling personalized nanomedicine. *Adv Drug Deliv Rev.* 2012 Oct;64(13):1522-31.
159. Nicolini C, Bragazzi NL, Pechkova E, Lazzari R. Ab Initio Semi-Quantitative Analysis of Micro-Beam Grazing-Incidence Small-

References

- Angle X-Ray Scattering (M-GISAXS) during Protein Crystal Nucleation and Growth. *J Proteomics Bioinform.* 2014;7:064-070.
160. Nicolini C, Bruzzese D, Cambria MT, Bragazzi NL, Pechkova E. Recombinant laccase: I. Enzyme cloning and characterization. *J Cell Biochem.* 2013 Mar;114(3):599-605.
161. Nicolini C, LaBaer J. Functional Proteomics and Nanotechnology-based Microarrays (vol. 2). Pan Stanford Series on Nanobiotechnology, vol. 2, Published 1st July 2010 by Pan Stanford Publishing, London–New York–Singapore, ISBN: 978-981-4267-76-2, 2010, pp. 1–306.
162. Nicolini C, Malvezzi AM, Tomaselli A, Sposito D, Tropiano G, Borgogno E. DNASER I: layout and data analysis. *IEEE Trans Nanobioscience.* 2002 Jun;1(2):67-72.
163. Nicolini C, Pechkova E. Nanoproteomics for nanomedicine. *Nanomedicine (Lond).* 2010 Jul;5(5):677-82.
164. Nicolini C, Pechkova E. An overview of nanotechnology-based functional proteomics for cancer and cell cycle progression. *Anticancer Res.* 2010 Jun;30(6):2073-80.
165. Nicolini C, Spera R, Festa F, Belmonte L, Chong S, Pechkova E, LaBaer J. Mass Spectrometry and Florescence Analysis of Snap-Nappa Arrays Expressed Using E. coli Cell-Free Expression System. *J Nanomed Nanotechnol.* 2013;4(5):181-194.
166. Nicolini C, Spera R, Stura E, Fiordoro S, Giacomelli L. Gene expression in the cell cycle of human T-lymphocytes: II. Experimental determination by DNASER technology. *J Cell Biochem.* 2006 Apr 1;97(5):1151-9.

References

167. Ogawa Y, Taketomi Y, Murakami M, Tsujimoto M, Yanoshita R. Small RNA transcriptomes of two types of exosomes in human whole saliva determined by next generation sequencing. *Biol Pharm Bull.* 2013;36(1):66-75.
168. Ogmundsdóttir HM, Björnsson J, Holbrook WP. Role of TP53 in the progression of pre-malignant and malignant oral mucosal lesions. A follow-up study of 144 patients. *J Oral Pathol Med.* 2009 Aug;38(7):565-71.
169. Ogmundsdóttir HM, Hilmarsdóttir H, Björnsson J, Holbrook WP. Longitudinal study of TP53 mutations in eight patients with potentially malignant oral mucosal disorders. *J Oral Pathol Med.* 2009 Oct;38(9):716-21.
170. Okanda FM, El Rassi Z. Affinity monolithic capillary columns for glycomics/proteomics: 1. Polymethacrylate monoliths with immobilized lectins for glycoprotein separation by affinity capillary electrochromatography and affinity nano-liquid chromatography in either a single column or columns coupled in series. *Electrophoresis.* 2006 Mar;27(5-6):1020-30.
171. Oluwadara O, Giacomelli L, Christensen R, Kossan G, Avezova R, Chiappelli F. LCK, survivin and PI-3K in the molecular biomarker profiling of oral lichen planus and oral squamous cell carcinoma. *Bioinformation.* 2009 Dec 31;4(6):249-57.
172. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol.* 2005 Oct;1(5):252-62.
173. Ono SJ, Liou HC, Davidon R, Strominger JL, Glimcher LH. Human X-box-binding protein 1 is required for the transcription of a subset

References

- of human class II major histocompatibility genes and forms a heterodimer with c-fos. *Proc Natl Acad Sci U S A*. 1991 May 15;88(10):4309-12.
174. Orlando B, Bragazzi N, Nicolini C. Bioinformatics and systems biology analysis of genes network involved in OLP (Oral Lichen Planus) pathogenesis. *Arch Oral Biol*. 2013 Jun;58(6):664-73.
175. Orlando G, Soker S, Wood K. Operational tolerance after liver transplantation. *J Hepatol*. 2009 Jun;50(6):1247-57.
176. Özdemir V, Cho WC. Theranostics: rethinking postgenomic diagnostics. *Expert Rev Mol Diagn*. 2012 Nov;12(8):783-5.
177. Özdemir V, Dove ES. Theranostics, the 21st century bioeconomy and 'one health'. Interview by Tikki Pang. *Expert Rev Mol Diagn*. 2012 Nov;12(8):807-9.
178. Ozdemir V, Kolker E, Hotez PJ, Mohin S, Prainsack B, Wynne B, Vayena E, Coşkun Y, Dereli T, Huzair F, Borda-Rodriguez A, Bragazzi NL, Faris J, Ramesar R, Wonkam A, Dandara C, Nair B, Llerena A, Kılıç K, Jain R, Reddy PJ, Gollapalli K, Srivastava S, Kickbusch I. Ready to put metadata on the post-2015 development agenda? Linking data publications to responsible innovation and science diplomacy. *OMICS*. 2014 Jan;18(1):1-9.
179. Pallier A, Hillion S, Danger R, Giral M, Racapé M, Degauque N, Dugast E, Ashton-Chess J, Pettré S, Lozano JJ, Bataille R, Devys A, Cesbron-Gautier A, Braudeau C, Larrose C, Soullillou JP, Brouard S. Patients with drug-free long-term graft function display increased numbers of peripheral B cells with a memory and inhibitory

References

- phenotype. *Kidney Int.* 2010 Sep;78(5):503-13. doi: 10.1038/ki.2010.162.
180. Pandey M, Prakash O, Santhi WS, Soumithran CS, Pillai RM. Overexpression of COX-2 gene in oral cancer is independent of stage of disease and degree of differentiation. *Int J Oral Maxillofac Surg.* 2008 Apr;37(4):379-83.
181. Pasa-Tolić L, Masselon C, Barry RC, Shen Y, Smith RD. Proteomic analyses using an accurate mass and time tag strategy. *Biotechniques.* 2004 Oct;37(4):621-4, 626-33, 636 passim.
182. Patel RS, Jakymiw A, Yao B, Pauley BA, Carcamo WC, Katz J, Cheng JQ, Chan EK. High resolution of microRNA signatures in human whole saliva. *Arch Oral Biol.* 2011 Dec;56(12):1506-13.
183. Pavlic V, Vujic-Aleksic V. Phototherapy approaches in treatment of oral lichen planus. *Photodermatol Photoimmunol Photomed.* 2014 Feb;30(1):15-24.
184. Pearson MA, Fabbro D. Targeting protein kinases in cancer therapy: a success? *Expert Rev Anticancer Ther.* 2004 Dec;4(6):1113-24.
185. Pechkova E, Chong S, Tripathi S, C. Nicolini. Cell free expression and APA for NAPPA and protein crystallography. In: Nicolini C, LaBaer J (Eds). *Functional Proteomics and Nanotechnology-based Microarrays*. Pan Stanford Series on Nanobiotechnology, vol. 2 (Chapter 7). Published 1st July 2010 by Pan Stanford Publishing, London–New York–Singapore, 2010, pp. 121–147.
186. Pechkova E, Gebhardt R, Riekel C, Nicolini C. In situ muGISAXS: I. Experimental setup for submicron study of protein nucleation and growth. *Biophys J.* 2010 Aug 9;99(4):1256-61.

References

187. Pechkova E, Innocenzi P, Malfatti L, Kidchob T, Gaspa L, Nicolini C. Thermal stability of lysozyme Langmuir-Schaefer films by FTIR spectroscopy. *Langmuir*. 2007 Jan 30;23(3):1147-51.
188. Pechkova E, Nicolini C. Protein nanocrystallography: a new approach to structural proteomics. *Trends Biotechnol*. 2004 Mar;22(3):117-22.
189. Pechkova E, Nicolini C. Domain organization and properties of LB lysozyme crystals down to submicron size. *Anticancer Res*. 2010 Jul;30(7):2745-8.
190. Pechkova E, Nicolini C. In situ study of nanotemplate-induced growth of lysozyme microcrystals by submicrometer GISAXS. *J Synchrotron Radiat*. 2011 Mar;18(Pt 2):287-92.
191. Pechkova E, Riekel C (Eds). *Synchrotron Radiation and Structural Proteomics*. Pan Stanford Series on Nanobiotechnology, vol. 3. Published 11th November 2011 by Pan Stanford Publishing, London–New York–Singapore, 2011, pp. 1–447.
192. Pechkova E, Sartore M, Giacomelli L, Nicolini C. Atomic force microscopy of protein films and crystals. *Rev Sci Instrum*. 2007 Sep;78(9):093704.
193. Pechkova E, Sivozhelezov V, Belmonte L, Nicolini C. Unique water distribution of Langmuir-Blodgett versus classical crystals. *J Struct Biol*. 2012 Oct;180(1):57-64.
194. Pechkova E, Sivozhelezov V, Nicolini C. Protein thermal stability: the role of protein structure and aqueous environment. *Arch Biochem Biophys*. 2007 Oct 1;466(1):40-8.

References

195. Pechkova E, Tripathi S, Nicolini C. MicroGISAXS of Langmuir-Blodgett protein films: effect of temperature on long-range order. *J Synchrotron Radiat*. 2009 May;16(Pt 3):330-5.
196. Pechkova E, Tripathi S, Ravelli RB, McSweeney S, Nicolini C. Radiation stability of proteinase K crystals grown by LB nanotemplate method. *J Struct Biol*. 2009 Dec;168(3):409-18.
197. Pechkova E, Vasile F, Spera R, Fiordoro S, Nicolini C. Protein nanocrystallography: growth mechanism and atomic structure of crystals induced by nanotemplates. *J Synchrotron Radiat*. 2005 Nov;12(Pt 6):772-8.
198. Perkins D, Verma M, Park KJ. Advances of genomic science and systems biology in renal transplantation: a review. *Semin Immunopathol*. 2011 Mar;33(2):211-8.
199. Pindborg JJRP, Smith CJ, van der Waal I. Histological typing of cancer and precancer of the oral mucosa. World Health Organization international histological classification of tumours. 2nd ed. Berlin: Springer; 1997.
200. Piro RM. Network medicine: linking disorders. *Hum Genet*. 2012 Dec;131(12):1811-20.
201. Poomsawat S, Buajeeb W, Khovidhunkit SO, Punyasingh J. Overexpression of cdk4 and p16 in oral lichen planus supports the concept of premalignancy. *J Oral Pathol Med*. 2011 Apr;40(4):294-9.
202. Prolo P, Chiappelli F, Cajulis E, Bauer J, Spackman S, Romeo H, Carrozzo M, Gandolfo S, Christensen R. Psychoneuroimmunology

References

- in oral biology and medicine: the model of oral lichen planus. *Ann N Y Acad Sci.* 2002 Jun;966:429-40.
203. Quack I, Rump LC, Gerke P, Walther I, Vinke T, Vonend O, Grunwald T, Sellin L. beta-Arrestin2 mediates nephrin endocytosis and impairs slit diaphragm integrity. *Proc Natl Acad Sci U S A.* 2006 Sep 19;103(38):14110-5.
204. Racapé M, Bragazzi N, Sivozhelezov V, Danger R, Pechkova E, Duong Van Huyen JP, Soulillou JP, Brouard S, Nicolini C. SMILE silencing and PMA activation gene networks in HeLa cells: comparison with kidney transplantation gene networks. *J Cell Biochem.* 2012 Jun;113(6):1820-32.
205. Racapé M, Duong Van Huyen JP, Danger R, Giral M, Bleicher F, Foucher Y, Pallier A, Pilet P, Tafelmeyer P, Ashton-Chess J, Dugast E, Pettré S, Charreau B, Soulillou JP, Brouard S. The involvement of SMILE/TMTC3 in endoplasmic reticulum stress response. *PLoS One.* 2011;6(5):e19321.
206. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, Walter JC, LaBaer J. Self-assembling protein microarrays. *Science.* 2004;305:86–90.
207. Ramachandran N, Larson DN, Stark PRH, Hainsworth E, LaBaer J. Emerging tools for real-time label-free detection of interactions on functional protein microarrays. *FEBS J.* 2005;272:5412–5425.
208. Ramachandran N, Raphael JV, Hainsworth E, Demirkan G, Fuentes MG, Rolfs A, Hu Y, LaBaer J. Next-generation high-density self-assembling functional protein arrays. *J. Nat. Methods.* 2008;5(6):535–538.

References

209. Ray S, Reddy PJ, Choudhary S, Raghu D, Srivastava S. Emerging nanoproteomics approaches for disease biomarker detection: a current perspective. *J Proteomics*. 2011 Nov 18;74(12):2660-81.
210. Re M, Mesiti M, Valentini G. A fast ranking algorithm for predicting gene functions in biomolecular networks. *IEEE/ACM Trans Comput Biol Bioinform*. 2012 Nov-Dec;9(6):1812-8.
211. Reichert S, Machulla HK, Klapproth J, Zimmermann U, Reichert Y, Gläser C, Schaller HG, Schulz S. Interleukin-2 -330 and 166 gene polymorphisms in relation to aggressive or chronic periodontitis and the presence of periodontopathic bacteria. *J Periodontal Res*. 2009 Oct;44(5):628-35.
212. Rhodus NL, Cheng B, Myers S, Miller L, Ho V, Ondrey F. The feasibility of monitoring NF-kappaB associated cytokines: TNF-alpha, IL-1alpha, IL-6, and IL-8 in whole saliva for the malignant transformation of oral lichen planus. *Mol Carcinog*. 2005 Oct;44(2):77-82.
213. Rodahl M, Kasemo B. A simple setup to simultaneously measure the resonant frequency and the absolute dissipation factor of a quartz crystal microbalance. *Rev. Sci. Instrum*. 1996;67:3238-3241.
214. Rodahl M, Kasemo B. Frequency and dissipation-factor responses to localized liquid deposits on a QCM electrode. *Sensors and Actuators B: Chemical*. 1996;37:111-116.
215. Rosa N, Correia MJ, Arrais JP, Lopes P, Melo J, Oliveira JL, Barros M. From the salivary proteome to the OralOme: comprehensive molecular oral biology. *Arch Oral Biol*. 2012 Jul;57(7):853-64.

References

216. Rosenblum D, Peer D. Omics-based nanomedicine: The future of personalized oncology. *Cancer Lett.* 2013 Aug 11. pii: S0304-3835(13)00549-1. doi: 10.1016/j.canlet.2013.07.029. [Epub ahead of print]
217. Roukos DH. Genome network medicine: new diagnostics and predictive tools. *Expert Rev Mol Diagn.* 2013 Sep;13(7):643-6.
218. Roussey-Kesler G, Giral M, Moreau A, Subra JF, Legendre C, Noël C, Pillebout E, Brouard S, Souillou JP. Clinical operational tolerance after kidney transplantation. *Am J Transplant.* 2006 Apr;6(4):736-46.
219. Ruotsalainen V, Ljungberg P, Wartiovaara J, Lenkkeri U, Kestilä M, Jalanko H, Holmberg C, Tryggvason K. Nephricin is specifically located at the slit diaphragm of glomerular podocytes. *Proc Natl Acad Sci U S A.* 1999 Jul 6;96(14):7962-7.
220. Sachdev R, Mandal AK, Singh I, Agarwal AK. Progressive rise of c-fos expression from premalignant to malignant lesions of oral cavity. *Medicina Oral Patologia Oral y Cirugia Bucal.* 2008;13:E683-6.
221. Sagoo P, Lombardi G, Lechler RI. Relevance of regulatory T cell promotion of donor-specific tolerance in solid organ transplantation. *Front Immunol.* 2012 Jul 13;3:184.
222. Sagoo P, Perucha E, Sawitzki B, Tomiuk S, Stephens DA, Miqueu P, Chapman S, Craciun L, Sergeant R, Brouard S, Rovis F, Jimenez E, Ballow A, Giral M, Rebollo-Mesa I, Le Moine A, Braudeau C, Hilton R, Gerstmayer B, Bourcier K, Sharif A, Krajewska M, Lord GM, Roberts I, Goldman M, Wood KJ, Newell K, Seyfert-Margolis V, Warrens AN, Janssen U, Volk HD, Souillou JP, Hernandez-

References

- Fuentes MP, Lechler RI. Development of a cross-platform biomarker signature to detect renal transplant tolerance in humans. *J Clin Invest*. 2010 Jun;120(6):1848-61.
223. Sami N, Bhol KC, Ahmed AR. Treatment of oral pemphigoid with intravenous immunoglobulin as monotherapy. Longterm follow-up: influence of treatment on antibody titres to human alpha6 integrin. *Clinical and Experimental Immunology*. 2002;129:533–40.
224. Santos SG, Campbell EC, Lynch S, Wong V, Antoniou AN, Powis SJ. Major histocompatibility complex class I-ERp57-tapasin interactions within the peptide-loading complex. *J Biol Chem*. 2007 Jun 15;282(24):17587-93.
225. Sartore M, Eggenhoffner R, Bezerra Correia Terencio T, Stura E, Hainsworth E, LaBaer J, Nicolini C. Label-free detection of NAPPA via atomic force microscopy. In: Nicolini C, LaBaer J (Eds.). *Functional Proteomics and Nanotechnology-based Microarrays*. Pan Stanford Series on Nanobiotechnology, vol. 2 (Chapter 6). Published 1st July 2010 by Pan Stanford Publishing, London -New York -Singapore, 2010, pp. 109–120.
226. Scully C, el-Kom M. Lichen planus: review and update on pathogenesis. *J Oral Pathol*. 1985 Jul;14(6):431-58.
227. Shangary S, Wang S. Targeting the MDM2-p53 interaction for cancer therapy. *Clin Cancer Res*. 2008 Sep 1;14(17):5318-24.
228. Shenoy SK, Lefkowitz RJ. β -Arrestin-mediated receptor trafficking and signal transduction. *Trends Pharmacol Sci*. 2011 Sep;32(9):521-33.

References

229. Shi P, Liu W, Zhou ZT, He QB, Jiang WW. Podoplanin and ABCG2: malignant transformation risk markers for oral lichen planus. *Cancer Epidemiol Biomarkers Prev.* 2010 Mar;19(3):844-9.
230. Shimmel A. Structural parameters of communication networks. *Bull Math Biophys* 15 (1953) 501-507.
231. Shumyantseva V, Deluca G, Bulko T, Carrara S, Nicolini C, Usanov SA, Archakov A. Cholesterol amperometric biosensor based on cytochrome P450sc. *Biosens Bioelectron.* 2004 Apr 15;19(9):971-6.
232. Silverman S Jr, Gorsky M, Lozada-Nur F. A prospective followup study of 570 patients with oral lichen planus: persistence, remission, and malignant association. *Oral Surg Oral Med Oral Pathol.* 1985;60:30-4.
233. Silverman EK, Loscalzo J. Network medicine approaches to the genetics of complex diseases. *Discov Med.* 2012 Aug;14(75):143-52.
234. Silverman EK, Loscalzo J. Developing new drug treatments in the era of network medicine. *Clin Pharmacol Ther.* 2013 Jan;93(1):26-8.
235. Simark-Mattsson C, Bergenholtz G, Jontell M, Eklund C, Seymour GJ, Sugeran PB, Savage NW, Dahlgren UI. Distribution of interleukin-2, -4, -10, tumour necrosis factor-alpha and transforming growth factor-beta mRNAs in oral lichen planus. *Arch Oral Biol.* 1999 Jun;44(6):499-507.
236. Sivozhelezov V, Braud C, Giacomelli L, Pechkova E, Giral M, Soulillou JP, Brouard S, Nicolini C. Immunosuppressive drug-free operational immune tolerance in human kidney transplants

References

- recipients. Part II. Non-statistical gene microarray analysis. *J Cell Biochem.* 2008 Apr 15;103(6):1693-706.
237. Sivozhelezov V, Giacomelli L, Tripathi S, Nicolini C. Gene expression in the cell cycle of human T lymphocytes: I. Predicted gene and protein networks. *J Cell Biochem.* 2006 Apr 1;97(5):1137-50.
238. Sivozhelezov V, Spera R, Giacomelli L, Hainsworth E, Labaer J, Bragazzi NL, Nicolini C. Bioinformatics and fluorescence DNASER for NAPPA studies on cell transformation and cell cycle. In: Nicolini C, LaBaer J (Eds). *Functional Proteomics and Nanotechnology-Based Microarrays*. Pan Stanford Series on Nanobiotechnology, vol. 2 (Chapter 1). Published 1st July 2010 by Pan Stanford Publishing, London -New York -Singapore, 2010, pp. 31–59.
239. Smith RE, Tran K, Vocque RH. Network medicine and high throughput screening. *Curr Drug Discov Technol.* 2013 Sep;10(3):182-94.
240. Soejima K, Nakamura H, Tamai M, Kawakami A, Eguchi K. Activation of MKK4 (SEK1), JNK, and c-Jun in labial salivary infiltrating T cells in patients with Sjogren's syndrome. *Rheumatology International.* 2007;27:329–33.
241. Spera R, Bezerra Correia TT, Nicolini C. NAPPA based nanogravimetric biosensor: Preliminary characterization. *Sens Actuators B:* 2013;182:682–688.
242. Spera R, Labaer J, Nicolini C. MALDI-TOF characterization of NAPPA-generated proteins. *J Mass Spectrom.* 2011 Sep;46(9):960-5.

References

243. Spera R, Nicolini C. NAPPA microarrays and mass spectrometry: new trends and challenges. In: "Essentials in Nanoscience Booklet Series". Published by Taylor & Francis Group, LLC, 15th January 2008.
244. Sugerman PB, Savage NW. Oral lichen planus: causes, diagnosis and management. *Aust Dent J.* 2002 Dec;47(4):290-7.
245. Sun L, Feng J, Ma L, Liu W, Zhou Z. CD133 expression in oral lichen planus correlated with the risk for progression to oral squamous cell carcinoma. *Ann Diagn Pathol.* 2013 Aug 1. doi:pii: S1092-9134(13)00078-6. 10.1016/j.anndiagpath.2013.06.004. [Epub ahead of print]
246. Sutinen M, Kainulainen T, Hurskainen T, Vesterlund E, Alexander JP, Overall CM, Sorsa T, Salo T. Expression of matrix metalloproteinases (MMP-1 and -2) and their inhibitors (TIMP-1, -2 and -3) in oral lichen planus, dysplasia, squamous cell carcinoma and lymph node metastasis. *Br J Cancer.* 1998 Jun;77(12):2239-45.
247. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzlaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksöz E, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE. A human protein-protein interaction network: a resource for annotating the proteome. *Cell.* 2005 Sep 23;122(6):957-68.
248. Stone SJ, McCracken GI, Heasman PA, Staines KS, Pennington M. Cost-effectiveness of personalized plaque control for managing the

References

- gingival manifestations of oral lichen planus: a randomized controlled study. *J Clin Periodontol.* 2013 Sep;40(9):859-67.
249. Stura E, Larosa C, Bezerra Correia Terencio T, Hainsworth E, Ramachandran N, LaBaer J, Nicolini C. Label-free NAPPA: anodic porous alumina. In: Nicolini C, LaBaer J (Eds). *Functional Proteomics and Nanotechnology-based Microarrays*. Pan Stanford Series on Nanobiotechnology, vol. 2 (Chapter 5), Published 1st July 2010 by Pan Stanford Publishing, London–New York–Singapore, 2010, pp. 95–108.
250. Stura E, Bruzzese D, Valerio F, Grasso F, Perlo P, Nicolini C. Anodic porous alumina as mechanical stability enhancer for LDL-cholesterol sensitive electrodes. *Biosens. Bioelectron.* 2007;23:655–660.
251. Syrjänen S, Lodi G, von Bültzingslöwen I, Aliko A, Arduino P, Campisi G, Challacombe S, Ficarra G, Flaitz C, Zhou HM, Maeda H, Miller C, Jontell M. Human papillomaviruses in oral carcinoma and oral potentially malignant disorders: a systematic review. *Oral Dis.* 2011 Apr;17 Suppl 1:58-72.
252. Tabor HK1, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet.* 2002 May;3(5):391-7.
253. Troitsky V, Ghisellini P, Pechkova E, Nicolini C. DNASER II: novel surface patterning for biomolecular microarray. *IEEE Trans Nanobioscience.* 2002 Jun;1(2):73-7.

References

254. Turatti E, da Costa Neves A, de Magalhães MH, de Sousa SO. Assessment of c-Jun, c-Fos and cyclin D1 in premalignant and malignant oral lesions. *Journal of Oral Science*. 2005;47:71–6.
255. Turon X, Rojas OJ, Deinhammer RS. Enzymatic kinetics of cellulose hydrolysis: a QCM-D study. *Langmuir*. 2008 Apr 15;24(8):3880-7.
256. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*. 2000 Feb 10;403(6770):623-7.
257. Umehara S, Karhanek M, Davis RW, Pourmand N. Label-free biosensing with functionalized nanopipette probes. *Proc Natl Acad Sci U S A*. 2009 Mar 24;106(12):4611-6.
258. Valter K, Boras VV, Buljan D, Juras DV, Susić M, Pandurić DG, Verzak Z. The influence of psychological state on oral lichen planus. *Acta Clin Croat*. 2013 Jun;52(2):145-9.
259. van Nas A, Pan C, Ingram-Drake LA, Ghazalpour A, Drake TA, Sobel EM, Papp JC, Lusk AJ. The systems genetics resource: a web application to mine global data for complex disease traits. *Front Genet*. 2013 May 20;4:84.
260. Velozo J, Aguilera S, Alliende C, Ewert P, Molina C, Pérez P, Leyton L, Quest A, Brito M, González S, Leyton C, Hermoso M, Romo R, González MJ. Severe alterations in expression and localisation of $\alpha_6\beta_4$ integrin in salivary gland acini from

References

- patients with Sjogren syndrome. *Ann Rheum Dis.* 2009 Jun;68(6):991-6.
261. Vivar JC, Pemu P, McPherson R, Ghosh S. Redundancy control in pathway databases (ReCiPa): an application for improving gene-set enrichment analysis in Omics studies and "Big data" biology. *OMICS.* 2013 Aug;17(8):414-22.
262. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* 2005 Jan 1;33(Database issue):D433-7.
263. Vousden KH, Lu X. Live or let die: the cell's response to p53. *Nat Rev Cancer.* 2002 Aug;2(8):594-604.
264. Walhout AJ1, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, Thierry-Mieg N, Vidal M. Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science.* 2000 Jan 7;287(5450):116-22.
265. Wali RK, Kunte DP, De La Cruz M, Tiwari AK, Brasky J, Weber CR, Gibson TP, Patel A, Savkovic SD, Brockstein BE, Roy HK. Topical polyethylene glycol as a novel chemopreventive agent for oral cancer via targeting of epidermal growth factor response. *PLoS One.* 2012;7(6):e38047.
266. Wang J. Carbon-Nanotube Based Electrochemical Biosensors: A Review. *Electroanalysis.* 2005;17(1):7-14.
267. Wang S, Shan X, Patel U, Huang X, Lu J, Li J, Tao N. Label-free imaging, detection, and mass measurement of single viruses by

References

- surface plasmon resonance. *Proc Natl Acad Sci U S A*. 2010 Sep 14;107(37):16028-32.
268. Wong DT. Salivaomics. *J Am Dent Assoc*. 2012 Oct;143(10 Suppl):19S-24S.
269. Xia W, Lau YK, Zhang HZ, Xiao FY, Johnston DA, Liu AR, Li L, Katz RL, Hung MC. Combination of EGFR, HER-2/neu, and HER-3 is a stronger predictor for the outcome of oral squamous cell carcinoma than any individual family members. *Clinical Cancer Research*. 1999;5:4164–74.
270. Xu Z, Shen Z, Shi L, Sun H, Liu W, Zhou Z. Aldehyde dehydrogenase 1 expression correlated with malignant potential of oral lichen planus. *Ann Diagn Pathol*. 2013 Oct;17(5):408-11.
271. Yang K, Zhang FJ, Tang H, Zhao C, Cao YA, Lv XQ, Chen D, Li YD. In-vivo imaging of oral squamous cell carcinoma by EGFR monoclonal antibody conjugated near-infrared quantum dots in mice. *Int J Nanomedicine*. 2011;6:1739-45.
272. Yoshizawa JM, Wong DT. Salivary microRNAs and oral cancer detection. *Methods Mol Biol*. 2013;936:313-24.
273. Younes F, Quartey EL, Kiguwa S, Partridge M. Expression of TNF and the 55-kDa TNF receptor in epidermis, oral mucosa, lichen planus and squamous cell carcinoma. *Oral Dis*. 1996 Mar;2(1):25-31.
274. Zakrzewicz A, Krasteva G, Wilhelm J, Dietrich H, Wilker S, Padberg W, Wygrecka M, Grau V. Reduced expression of arrestin beta 2 by graft monocytes during acute rejection of rat kidneys. *Immunobiology*. 2011 Jul;216(7):854-61.

References

275. Zhang K, Kaufman RJ. From endoplasmic-reticulum stress to the inflammatory response. *Nature*. 2008 Jul 24;454(7203):455-62.
276. Zhang WY, Liu W, Zhou YM, Shen XM, Wang YF, Tang GY. Altered microRNA expression profile with miR-27b down-regulation correlated with disease activity of oral lichen planus. *Oral Dis*. 2012 Apr;18(3):265-70.
277. Zhou XJ, Sugarman PB, Savage NW, Walsh LJ. Matrix metalloproteinases and their inhibitors in oral lichen planus. *Journal of Cutaneous Pathology*. 2001;28:72–82.
278. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, Mitchell T, Miller P, Dean RA, Gerstein M, Snyder M. Global analysis of protein activities using proteome chips. *Science*. 2001;293:2101–2105.

CV and records list

10. Curriculum Vitae et Studiorum of the PhD candidate

Dr. Nicola Luigi Bragazzi was born on the 2nd of March in 1986 in Carrara (MS), Tuscany (Italy) and is currently a MD, a MSc, a PhD candidate and a resident in Public Health.

He got his MD (medical degree) on the 15th of July in 2011 with a final mark of 110/110 *cum laude* with a thesis on Personalized Nanomedicine ("Nanomolecular aspects of medicine, at the cutting-edge of the nanobiosciences in the field of health-care") and the joint Italo-Russian MSc (Master of Science) in nanobiotechnologies at Lomonosov Moscow State University (27th April 2012).

Besides being a PhD candidate, he is resident in Public Health at University of Genoa, Italy, 2nd year.

AWARDS

Alfiere della Repubblica Italiana (Young Knight of Italian Republic), awarded by President Carlo Azeglio Ciampi (2nd June 2005)

Medal Degree at University of Genoa (2011)

11. LIST OF PUBLICATIONS

PUBLICATIONS DIRECTLY PERTINENT TO THE PhD THESIS

- 1) Nicolini C, Bragazzi N, Pechkova E. Nanoproteomics enabling personalized nanomedicine. *Adv Drug Deliv Rev.* 2012 Oct;64(13):1522-31. **(I.F.: 12.888)**
- 2) Spera R, Festa F, Bragazzi NL, Pechkova E, Labaer J, Nicolini C. Conductometric monitoring of protein-protein interactions. *J Proteome Res.* 2013 Dec 6;12(12):5535-47. **(I.F.: 5.056)**
- 3) Racapé M, Bragazzi N, Sivozhelezov V, Danger R, Pechkova E, Duong Van Huyen JP, Soulillou JP, Brouard S, Nicolini C. SMILE silencing and PMA activation gene networks in HeLa cells: comparison with kidney transplantation gene networks. *J Cell Biochem.* 2012 Jun;113(6):1820-32. **(I.F.: 3.062)**
- 4) Bozdaganyan M, Bragazzi NL, Pechkova E, Shaytan K, Nicolini C. Identification of best protein crystallization methods by Molecular Dynamics. *Crit Rev Eukaryot Gene Expr.* 2014, in press. **(I.F.: 2.065)**
- 5) Bragazzi NL, Pechkova E, Scudieri D, Terencio TB, Adami M, Nicolini C. Recombinant laccase: II. Medical biosensor. *Crit Rev Eukaryot Gene Expr.* 2012;22(3):197-203. **(I.F.: 2.065)**
- 6) Pechkova E, Bragazzi NL, Bozdaganyan M, Belmonte L, Nicolini C. A review of the strategies for obtaining high quality crystals utilizing nanotechnologies and space. *Crit Rev Eukaryot Gene Expr.* 2014, in press. **(I.F.: 2.065)**

CV and records list

7) Nicolini C, Adami M, Sartore M, Bragazzi NL, Bavastrello V, Spera R, Pechkova E. Prototypes of newly conceived inorganic and biological sensors for health and environmental applications. *Sensors (Basel)*. 2012 Dec 12;12(12):17112-27. **(I.F.: 1.953)**

8) Bragazzi NL, Nicolini C. A biomolecular data-mining strategy for predicting human diseases-related microRNAs (miRNAs), based on Leader Genes algorithm: an application to Oral Lichen Planus (OLP). *Arch Oral Biol*. 2013. Accepted with revision. **(I.F.: 1.549)**

9) Orlando B, Bragazzi N, Nicolini C. Bioinformatics and systems biology analysis of genes network involved in OLP (Oral Lichen Planus) pathogenesis. *Arch Oral Biol*. 2013 Jun;58(6):664-73. **(I.F.: 1.549)**

10) Bragazzi NL, Nicolini C. A Leader Genes Approach-based Tool for Molecular Genomics: From Gene-ranking to Gene-network Systems Biology and Biotargets Predictions. *J Comput Sci Syst Biol*. 2013;6:165-176 **(U.I.F.: 2.53)***.

11) Bragazzi NL, Giacomelli L, Sivozhelezov V, Nicolini C. Leader Gene: A Fast Data-mining Tool for Molecular Genomics. *J Proteomics Bioinform*. 2011;4(4):083-086. **(U.I.F.: 2.22)***.

Total I.F.: 32.252

(including also the U.I.F.: 37.002)

CV and records list

PATENTS

Nicolini C, Bragazzi NL, Pechkova E. GE2012A000080. APA-SNAP NANOARRAYS PER PROTEOMICA LB-BASATA (STRUTTURALE E FUNZIONALE) E PER VACCINOLOGIA

FURTHER ACTIVITIES

Further activities that have been made during the PhD thesis include attendances to international conferences (both as poster presenter and oral speaker). Moreover, four crystallographic protein structures have been done and deposited in PDB data-bank.

FURTHER PUBLICATIONS NOT DIRECTLY PERTINENT TO THE THESIS

Further publications have been made during the Medical Degree and the joint Italian-Russian Master.

**Unofficial 2012 Impact Factors (U.I.F.) were established by dividing the number of articles published in 2010 and 2011 which were cited in 2012 based on a search of the Google Scholar Citation Index database, by the number of articles published in the previous two years (2010 and 2011).*

A = the number of times, articles published in 2010 and 2011, which were cited by indexed journals during 2012

B = the total number of "citable items" published by that journal in 2010 and 2011.