Antigen Chip:

Development and analysis;

An application to autoimmune diseases

Gad Elizur

M.Sc. Thesis submitted to the Scientific Council of the
Weizmann Institute of Science

Research conducted under the supervision of
Prof. Eytan Domany and
Prof. Irun R. Cohen

January 2004

<u>Acknowledgments</u>

Contents

Abstract

Autoimmune diseases affect 3% of the US population, and a similar percentage of the population of the industrialized world. Autoimmune diseases are usually caused by an immune attack, mostly by autoantibodies and T-cells, against body components. A very promising approach to the understanding of this group of diseases is studying the antibody-antigen interaction on large scale data using bioinformatic tools. We created large scale data of antibody-antigen interactions and analyzed it.

The first study dealt with the analysis of ELISA data obtained from patients suffering from various diseases. We looked for antibodies against specific antigens that are differentially expressed in healthy and sick patients over the different diseases. We found that the diseases were significantly differentiated by most of the antigens presented in the data, which suggested that we were dealing with artifacts.

In the second project we developed a new technique named the 'antigen chip'. The production process of this chip is described in detail in this work. We had to cope with broad aspects of the production process, starting with chip design, followed by choosing the right working protocol, scanning, quantifying and finally analyzing the data we had produced.

Finally, we carried out a biological experiment using the new chip and showed that the chip is a useful and reliable tool for studying autoimmune diseases via antibody-antigen interactions, or any other protein interaction.

1. Human autoimmune diseases

1.1 Introduction

Autoimmune diseases affect 3% of the U.S. population, and probably a similar percentage of the industrialized world's population [1]. Autoimmune diseases are usually caused by an immune attack against body components, mostly by autoantibodies and T-cells. Traditionally, immunologists have reduced the immune system, where possible, to one-to-one relationships between particular antigens and corresponding antibodies or T-cell clones [2, 3]. They also sought to establish one-to-one relationships between a particular autoantibody and a particular disease. Thus autoimmune diseases were usually investigated by following the response to single self-antigens.

Although science has made remarkable progress towards understanding immune function over the past four decades in terms of the role of the major histocompatibility complex (MHC) and the nature of lymphocyte antigen receptors that confer specificity to autoimmune responses, our understanding of the underlying dysregulation and specificity of autoimmune responses remains limited [4]. For certain autoimmune diseases, including Multiple Sclerosis (MS) and Systemic Lupus Erythematosus (SLE), candidate autoantigens have been identified, but their exact role in the initiation, perpetuation, and pathophysiology of the disease are not well understood. In other autoimmune diseases, including Rheumatoid Arthritis (RA) and Psoriasis, the targeted autoantigens remain unidentified despite extensive experimental efforts [4]. Moreover, the presence of autoantibodies in healthy persons [5, 6] complicates the serological diagnosis of autoimmune disease and confounds our understanding of how the immune system actually discriminates the self from the non-self [7]. Still a hallmark of an autoimmune disease is the production of high-affinity autoantibodies. Therefore, in order to overcome the complexity of the system, an appropriate way to further investigate autoimmune diseases is using large-scale autoantibody affinity data. The data will include information regarding the interaction between a variety of antigens and these antibodies. This will allow us to reveal complex connections between groups of antigens and diseases.

As a result of the above, one can say that microarrays and other multiplex screening technologies probably represent powerful tools to study the pathophysiology and specificity of autoimmune responses.

Here we are using an unbiased solid-phase ELISA antibody test in order to measure the antibodies reactivity against autoantigens. Serum was collected from healthy and sick patients. IgM and IgG antibodies were purified from the serum and were hybridized with the antigens in the ELISA plate. The antibodies from the serum were detected using ELISA with 80 different antigens, mostly self-antigens. We applied statistical tools in order to identify antigens that significantly discriminate the healthy and the patients that suffer from one of the autoimmune diseases we checked.

1.2 Autoimmune diseases

An autoimmune disease is usually caused by a self attack of the body's immune system, against one or more organs. The prefix "auto" means "self." Autoimmunity means that the immune system is reacting against normally occurring antigens in the person's body, as if these antigens were foreign.

There are many mysteries in autoimmune diseases. The mechanisms of all the diseases are not well known. Why does the body attack itself? Why now? Why this specific patient? etc.

On each cell surface there are antigens, which are fragments of inner proteins. The antigens are presented to the outside by another protein – MHC. Autoantigens are a way for the body's immune system to identify self-cells. Foreign cells or cells that are infected with foreign microbes proteins, present the foreign antigens on their surface, which generally stimulate the immune response [8, 9]. The immune system recognizes the antigens presented on the cell surface using several mechanisms such as T-cells and antibodies, which can lead to an antibody-antigen complex. By this antibody-antigen complex, the immune system identifies the problematic infected cells in the organ and deals with them in several ways. (I won't go into more details regarding the immune actions). In autoimmune diseases, the immune system usually identifies some self-cells as non-self cells which starts the immune response. In order to overcome these apparent non-self cells, the immune system produces antibodies or cytoxic T-cells (CTL's) which attack the self antigen, represented on these cells. Of course, every autoimmune disease has its own mechanism and its own immune response [8]. In diabetes for example T-cells (with some help from autoantibodies) attack β cells in the pancreas and kill them.

1.2.1 Diabetes

Diabetes is a chronic metabolic disease associated with elevated blood sugar levels in which the body does not produce, or improperly uses insulin. Insulin is a hormone manufactured and secreted by the pancreas, and it is a key player for keeping blood sugars balanced, for regulating metabolization of glucose and ensuring proper function of cells. Diabetes patients suffer from a multitude of complications such as blindness, heart disease, stroke, neuronal damage, amputation and kidney failure [10]. These complications usually result from high levels of circulating blood glucose which causes damage to blood vessels and nerves [11]. Roughly 120 million people world wide suffer from diabetes. There are three major types of diabetes, namely Type I diabetes, Type II diabetes and gestational diabetes.

1.2.1.1 Type I or Insulin-Dependent Diabetes Mellitus (IDDM)

Type I diabetes, also known as juvenile onset diabetes mellitus, ketosis-prone diabetes mellitus and immune-mediated diabetes, is an autoimmune disorder in which the immune system destroys the β cells in the pancreas which produce insulin.

Type I diabetes accounts for 5-10% of all cases of diabetes, affecting 0.3% of the world's population. It usually begins in childhood or adolescence, but is a lifelong disease and at present there is no cure which means patients are dependent on artificial insulin treatment for life. The autoimmune destruction is caused by auto-aggressive T cells that infiltrate the pancreas and eventually destroy the insulin-producing β-islet cells [12]. Thus, type I diabetes is part of the autoimmune disease group. The disease is characterized by increased levels of T-cells which respond to several autoantigens. The identity of the self proteins in the pancreatic islets that target the T-cells for autoimmune destruction is still being debated.

A cure for the disease is one of the main goals of researchers today. A DNA Vaccine has been developed to prevent and halt the progression of the disease [13]. This vaccine has been shown to cause the immune system to modulate the T cell response by injecting an antigen known as HSP60 and thereby "shifting" the attack of the immune system and preventing the progression of the disease.

1.2.1.2 Type II or Non-Insulin Dependent Diabetes Mellitus (NIDDM)

An estimated 15.7 million Americans have diabetes type II, which represent 90 to 95 percent of all diabetics. Almost one-third of them are unaware they have the disorder. Type II diabetes usually occurs in people over 30 years; and is therefore also called maturity-onset and adult onset diabetes. The onset of type II diabetes can be related to lifestyle (overweight, inactivity, certain drugs, excess alcohol). Obesity is considered a major risk factor for the disease. This type of diabetes has a strong familial occurrence, with a high risk of disease among relatives of first degree.

People with this condition usually have a combination of production of an inadequate quantity of insulin and the inability of the body to properly respond to the insulin's signals, called "insulin resistance". Normally, insulin binds with a receptor molecule on a cell membrane setting off signals that stimulate an array of events, including the induction of glucose and cell growth. In type 2 diabetes this signal induction is inefficient and the cells' ability to respond to insulin is impaired.

Obesity worsens Insulin resistance because fat cells prevent the insulin from moving the glucose from the blood into the body cells.

Type II diabetes treatment centers on a multidisciplinary treatment plan; Diet control is important for the prevention of obesity and for maintaining a glucose free diet, regular exercise, and oral medication therapy aimed at reducing the insulin resistance.

1.2.2 Multiple Sclerosis

MS is an autoimmune disease that affects the central nervous system (CNS). The CNS consists of the brain, spinal cord, and the optic nerves. Surrounding and protecting the nerve fibers of the CNS is a fatty tissue called myelin, which helps nerve fibers conduct electrical impulses [14].

In MS, myelin is lost in multiple areas, leaving scar tissue called sclerosis. These damaged areas are also known as plaques or lesions. Sometimes this results in damage or breakage of the nerve fiber.

Myelin not only protects nerve fibers, but makes their job possible. When myelin or the nerve fiber is destroyed or damaged, the ability of the nerves to conduct electrical

impulses to and from the brain is disrupted, and this produces the various symptoms of MS [15].

The exact cause of MS is unknown; yet most researchers believe that the damage to myelin results from an abnormal autoimmune response by the body's immune system to myelin.

The disease is multifactorial, some of the main factors thought to cause the disease are – Genetic factors, Gender and environmental triggers. Others, such as viruses, trauma to nerve tissue and toxic causes may also be a possibility [16, 17].

1.2.3 Arthritis

1.2.3.1 Rheumatoid Arthritis (RA)

Rheumatoid arthritis causes chronic inflammation in the cartilage of the joints and/or other internal organs.

RA is a one of the most common forms of arthritis; it is systemic disease that affects the entire body. It is characterized by the inflammation of the membrane lining the joint called synovium, which causes pain, stiffness, redness and swelling of the joint. The inflammatory cells in the synovium can invade and damage the surrounding bone and cartilage. These cells release enzymes that may digest bone and cartilage which cause damage and misalignment of the involved joint. Thus, the joint may loose its shape and alignment, resulting in pain and loss of movement.

The specific cause of rheumatoid arthritis is not yet known. However, it is known that RA is a debilitating, chronic, systemic, autoimmune disease [18]. The immune system does not operate as it should, resulting in the immune mediated attack of healthy, self joint tissue, leading to inflammation and joint damage.

Researchers suspect that agent-like viruses may trigger the autoimmune response in RA in people who have an inherited tendency for the disease. This is supported by the fact that many people with RA have a certain genetic marker called HLA-DR4, which seems to be a marker of high susceptibility for RA. Many other genes seem to have an influence on the development of RA, and researchers are still in search of these genes and other factors which cause the disease.

1.2.3.2 Osteoarthritis

OA is the most common form of arthritis in the United States, affecting more than 20 million Americans. Osteoarthritis (OA) begins with the breakdown of cartilage in joints, resulting in joint pain and stiffness.

OA is usually a result of injury or unusual stress caused to a specific joint. It commonly affects the joints of the fingers, knees, hips, and spine. Other joints affected less frequently include the wrists, elbows, shoulders, and ankles. OA is not an autoimmune disease.

1.3 Aims

In this study, we wish to investigate autoimmune diseases using semi-large scale techniques. More specifically, we study the differential autoantibody expression in healthy patients and those suffering from one of several autoimmune diseases. The research will be carried out on data that was collected from healthy and sick donors and produced by the ELISA method (see section 1.4.4 in the chapter on materials and methods). The analysis of the data will be conducted using known bioinformatics tools that are used in DNA chip data analysis. The data analysis will be done in several steps: first, we will use supervised methods in order to find antibodies that are differentially expressed against specific antigens in healthy and sick people over different diseases. We will then use unsupervised methods, such as cluster analysis, in order to identify groups of antibodies with similar expression patterns, in order to check if those antibodies are biologically related to any specific disease. The combination of the supervised and unsupervised methods is expected to give us a better understanding of the antigens that are related to different diseases.

1.4 Materials and Methods

1.4.1 Antigens

The 80 antigens used in this study are enumerated in Table 6 in the appendix. These antigens include proteins, peptides, nucleotides and phospholipids reported to interact with antibodies.

1.4.2 Antibodies

The secondary antibodies used in the ELISA assay were F(ab')2 goat anti-human IgG + IgM linked to alkaline phosphatase and goat anti-human IgM linked to horseradish peroxidase. These antibodies were purchased from Jackson ImmunoResearch Labs. Inc. (West Grove, Pennsylvania, USA), and were used at a final dilution of 1:1500 in bovine serum albumin 0.3 %.

1.4.3 Test samples

The serum samples were collected from 5 different groups:

1. Healthy donors (H.D.) – 30 samples.
2. Multiple sclerosis patients (MS) – 10 samples.
3. Osteoarthritis patients (OA) – 10 samples.
4. Rheumatoid arthritis patients (RA) – 10 samples.
5. Type II diabetes patients – 30 samples.

1.4.4 Solid-phase antibody assay (ELISA)

Standard ELISA assay was used.

Enzyme-Linked Immunosorbent Assay (ELISA) is a useful and powerful method for estimating concentrations in the range of ng/ml to pg/ml of materials in a solution (such as serum, urine and culture supernatant). In the current experiment our aim was to detect antibodies (IgM or/and IgG) which recognize specific antigens. Each well of a 96-wells plate is loaded with a different antigen that sticks to the plastic where it is then ready for the interaction with antibodies. Serum from a patient is taken and being diluted. After diluting the antibody's solution, a fixed amount is dropped into each well. If the antibodies in the solution recognize the specific antigen in a well, they bind to each other. After incubation the solution is washed in order to get rid of non-specific recognition. This step is followed by labeling; enzyme linked secondary antibodies (conjugate) are added to each well followed by the enzyme substrate. The enzyme linked secondary antibodies are specific to the constant part of the antibodies (IgM or IgG) and interact with the attached antibody. The final part of the ELISA is the Optical Density (O.D.) readings, which quantitavely measures the product formation and may determines the amount of antibodies from the patient's serum that have recognized each antigen. This reading is carried out by measuring the fluorescent signal emitted by each conjugate, using the ELISA apparatus (Figure 1).

(For more details see: "cluster analysis of human antibody reactivities in health and type I diabetes mellitus: A bio-informatic approach to immune complexity" – Quintana et al 2003). The ELISA OD reading was achieved for anti IgM antibodies (IgM data) and for anti IgM + anti IgG antibodies together (IgM+IgG data).

**Figure 1:** schematic description of the ELISA procedure.
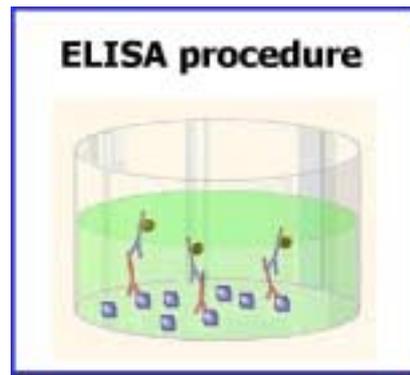
1.4.5 Pre-processing

The raw data was scaled using several methods. In spotted array chips or ELISA there is usually a problem of 'overall different intensity', due to a difference between the overall optical densities of each chip. This phenomenon can arise from human error caused by different serum concentrations, or different total liquid volume on the chip/well. In order to overcome this problem, scaling methods are used. In the following analysis three scaling methods were used in order to verify that the scaling does not dominate the effect of the raw data. The methods used are:

- No scaling – raw data as is.
- Divide each point on the 'chip' by the overall 'chip' mean.
- Divide each point on the 'chip' by the overall 'chip' geometric mean.
    - Log transformation (log10)
    - Subtract the mean
- Divide each point on the 'chip' by the overall 'chip' median.

The aim of this step was to scale all the samples in order to be able to compare the antigen reactivity results between different samples. In Figure 2 we see the expression matrix after scaling, using the divide by mean method. One can see that the RA group reacts differently in the two sub experiments. In the IgM data, the over all expression of the antibodies is relatively low and uniform compared to all other groups in the same experiment. By looking at other groups in the same experiment one can see a pattern and some change in antibody reactivity along the groups and antigens. This finding does not seem to be a real biological result, and so to confirm this we look at the RA group in the IgM+IgG data where we can see that the reactivity pattern of this group is similar to that of other groups in the same experiment. It is hard to tell which one of the two RA groups is corrupted, but one can say for sure that in at least one of the two groups there is an

artifact, which is not of biological origin. Another phenomenon is the dissimilarity of H.D. and MS patients from all the other groups in both sub experiments (IgM data & IgM+IgG data). – see the high reactivities to antigens 64-80.
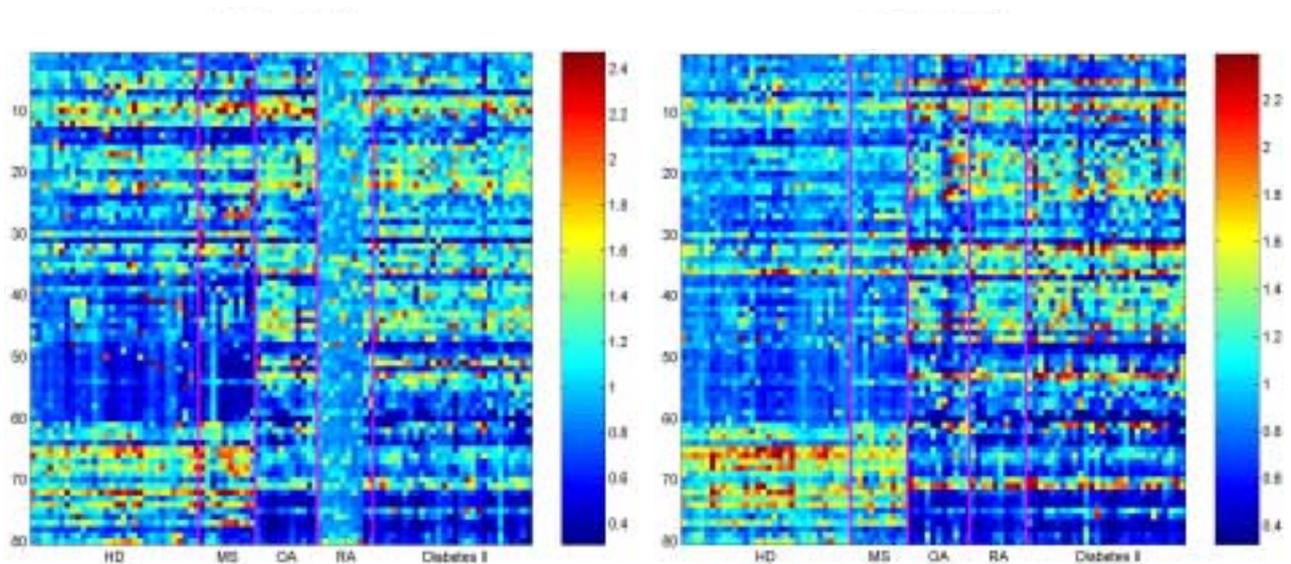
Figure 2: Raw data after scaling (divide by mean). Each Elisa plate is represented by one column and each antigen reactivity by one row. The samples are arranged according to their disease. Left – IgM data.   Right – IgM+IgG data.

1.4.6 Supervised analysis

The OD readings corresponding to the antibody reactivities of the five predefined patient groups were placed in matrices. Each row denoted an antigen and each column a patient. In order to present both IgM and IgM+IgG reactivities, two different matrices were produced from the OD readings. The first matrix contained the data collected from IgM antibody detection only (IgM data), and the second contained data collected from both IgM & IgG antibody detection (IgG+IgM data). A Wilcoxon Rank Sum test was conducted for each antigen separately in order to find the significant discrimination between pairs of subject groups, and a P-value was generated. The P-values were collected and sorted. In order to calibrate the P-values to multiple experiments (multiplicity problem) using 80 antibodies – the False Discovery Rate (FDR) method was used [19]. The selected FDR fraction was set to 0.05 – as expected false discovery rate. The antigens with p-values lower than the FDR-line, were selected as significantly discriminating antigens. A more restrictive and strict test was also conducted – the Bonferroni test. This test is much stricter than ranksum + FDR. In order to verify the results, the entire analysis was repeated using the T-test instead of ranksum.

In each statistical analysis test (t-test or ranksum), two groups of samples were compared. In each comparison the healthy donor group was tested versus one group of diseased subjects.

Both t-test and ranksum test were used in order to eliminate the possibility of biased results. While conducting t-test, one assumes that the data is normally distributed, but since no initial assumption was proved in the relevant data, the ranksum test was conducted as well.

The entire supervised analysis was conducted on each matrix/expression data separately – IgM and IgG+IgM.

1.4.7 FDR

1.4.7.1 The Multiplicity problem

Usually, each experiment checks one hypothesis. The regular reject α level is 0.05, which means that one can reject the null hypothesis, when the p-value calculated using some statistical test is less that 0.05. The chance of error (rejecting the null hypothesis while it is true) in such case is less than 5%, which is a reasonable result. In large scale data, such as DNA chips or ELISA plates, there is usually more than one gene/antigen that is being tested. Usually N, the number of gene/antigens, ranges between a few hundreds to a few thousand in a single experiment. Assuming statistical independence, in every big experiment there is a chance that a fraction α of the tested genes will have a p-value lower than α. If we will take the single test experiment threshold ($\alpha=0.05$), our results will contain mistaken genes/antigens, which received low p-values by chance. This is the multiplicity problem. In order to overcome it we need to find another method to quantify p-values. A well known method is the Bonferroni method. This method is quite strict, the null hypothesis will be rejected at a p-value lower than α/N. ($\alpha=0.05$, N=number of genes/antigens in the experiment or the number of tests performed). Methods like Bonferroni assure that there are no false positive samples. On the other hand, in order to find such a sample that will reject the null, very low p-values must be obtained. This can cause loss of samples which truly reject the null but got p-value higher than α/N. These are called false negatives. In order to overcome both problems, false positive and false negative, a new method was developed, which controls the false positive samples without losing the false negatives. This method is called FDR.

1.4.7.2 False discovery Rate (FDR) method

This method supplies a measure for the expected fraction of falsely discovered genes/antigens among the total list of genes/antigens that are identified as rejecting the null. The expected fraction is called the FDR [19].

Let R denote the number of hypotheses rejected by the procedure, and V the number of true null hypotheses that were wrongly rejected. Then FDR=E(V/R).

FDR places a weaker demand on the genes/antigens identified than Bonferroni; therefore less true genes/antigens are lost (less false negatives).

Protocol:

Let N be the number of null hypotheses tested. For each hypothesis a p-value P_g was calculated.

- Arrange the N genes/antigens according to their P_g values in ascending order.

- Set an upper bound Q, for the fraction of false positives.

- Create a list of $q_i = Q \times i/N$, where $Q/N \leq q_i \leq Q$.

- Starting from the max q_i – find j such that $q_i \leq p_i$ for all i<j.

- Reject the null hypothesis for all the antigens below this p_j. The expected fraction of false discoveries in this list is Q.

(See Benjamini and Hochberg, 1995 for further information regarding FDR).

1.4.8 Wilcoxon Rank Sum test

Rank sum is a non-parametric test; it does not assume the underlying source population to be normally distributed. (This is the major advantage it has to t-test).

The Rank-Sum test's assumptions are: 1) The two samples are randomly and independently drawn; 2) The dependent variable is continuous; 3) The measures within the two samples have the properties of an ordinal scale of measurement (e.g. there is an order relation on the elements).

The test involves assignment of ranks to the pooled values of both samples. The sum of ranks of the smaller group is considered to be the statistic, W, in order to obtain a p-value. This p-value is the probability to get such a rank distribution, between the two sample groups, under the null hypothesis that the values of both samples come from the same distribution [20].

Protocol:

Given two groups of samples *A* and *B*, with n_A and n_B sample sizes, where $n_A < n_B$ and $n_A + n_B = N$.

- Combine the two samples.
- Assign ranks from the lowest value to the highest one. (1, 2….*N*). If there are two equal values assign the average of the two consecutive ranks.
- The statistic W_A is the sum of ranks of the smaller sample.
- For sufficiently large n_A and n_B, the following normal approximation is used:

$$Z = \frac{|\mu - W| - 0.5}{\sigma}$$

Where

$$\mu = \frac{n_A(n_A + n_B + 1)}{2}$$

$$\sigma = \sqrt{n_B \times \mu / 6}$$

- A p-value is obtained from *Z* using the normal distribution.

1.5 Results and Discussion

1.5.1 Ranksum and t-test

We used the ranksum test by reviewing the antibody reactivity of the antigens, one by one, on two predefined groups. After the P-values were calculated, we arranged them in order and generated the FDR line of 0.05 false discovery rate. For each comparison/test we collected three numbers as our result:

1. Number of antigens that discriminate the two groups with a P-value lower than 0.05

2. Number of antigens that discriminate the two groups, with FDR of 0.05.

3. Number of antigens which discriminate the two groups with a P-value lower than 0.05/N (N-number of measured antigens = 80) – the Bonferroni test.

We looked for discriminating antigens between two groups, the healthy donors (HD) and the sick patients (each group separately). In the beginning we used the data which was collected using only IgM antibodies detector, and later on we conducted the same analysis on the data collected using detectors for both IgM+IgG antibodies.

The results are summarized in the following tables:

Table 1: **IgM** data - Number of antigens separating H.D. from other groups using **rank sum** and 5% FDR/Bonferroni

HD vs. Scaling method	Multiple Sclerosis	Osteoarthritis	Rheumatoid arthritis	Type II Diabetes
No scaling	55 / 8	66 / 48	2 / 2	69 / 54
Divide by mean	17 / 3	41 / 27	50 / 26	58 / 38
Divide by geometric mean	22 / 2	46 / 26	49 / 28	62 / 40
Log + subtract mean	22 / 2	46 / 26	49 / 28	62 / 40
Divide by median	18 / 3	39 / 28	50 / 25	58 / 40

Separating antigens between Healthy Donors and patients with different diseases.

We used 5 scaling methods and calculated the number of discriminating antigens using rank-sum test and FDR/Bonferroni methods.

Example: 50/25 (HD vs. RA, scaling - divide by median) means

- 50 - Antigens separate H.D. from R.A. as result of ranksum test and FDR of 0.05.

- 25 - Antigens separate H.D. from R.A. as result of ranksum test and Bonferroni.

Table 2: IgM data - Number of antigens separating H.D. from other groups using **t-test** and
FDR/Bonferroni

HD vs. Scaling method	Multiple Sclerosis	Osteoarthritis	Rheumatoid arthritis	Type II Diabetes
No scaling	47/24	68/55	0/0	65/44
Divide by mean	12/4	40/25	38/18	51/29
Divide by geometric mean	14/5	43/25	29/18	50/30
Log + subtract mean	14/5	43/25	29/18	50/30
Divide by median	14/4	37/25	27/17	42/31

Separating antigens between Healthy Donors and patients with different diseases.

Example: 27/17

- 27 - Antigens separate H.D. from R.A. as result of t-test and FDR of 0.05.

- 17 - Antigens separate H.D. from R.A. as result of t-test test and Bonferroni.
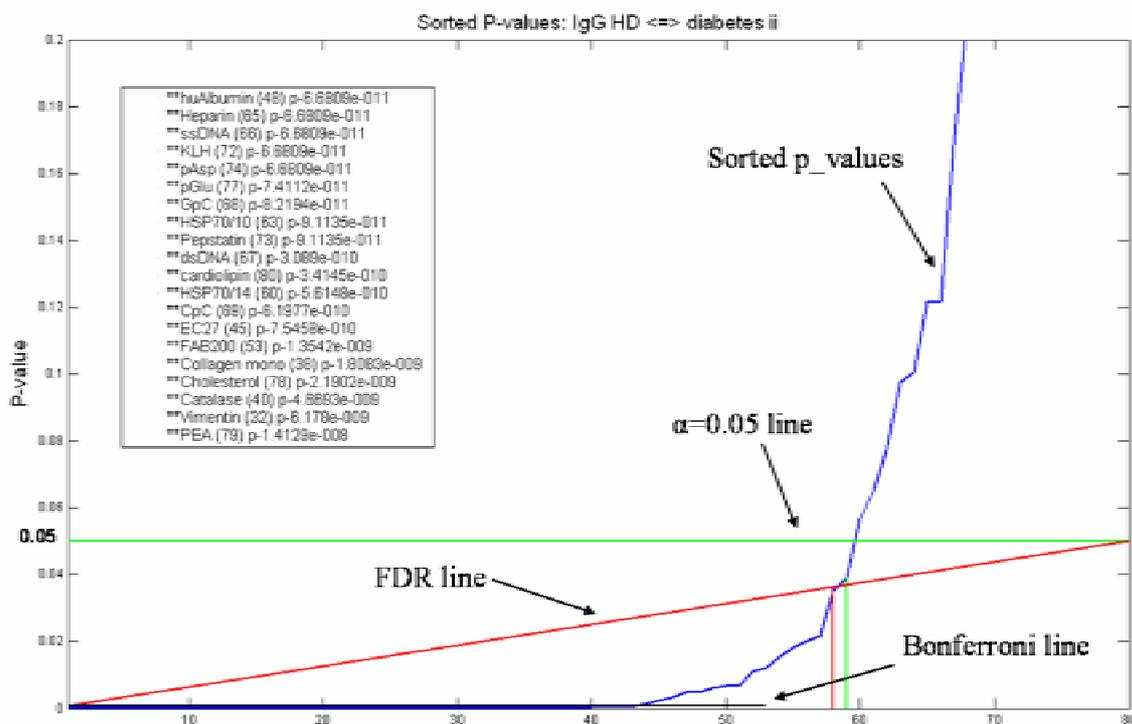
Figure 3: IgM+IgG data. Sorted p-values of 80 antigens after ranksum test. Scaling method is 'dividing by mean'. Number of separating antigens, with 0.05 FDR, is 58 out of 80. Number of separating antigens, passing Bonferroni, is 43! out of 80. On inset the top 20 separating antigens are listed.

Table 3: IgG+IgM data - Number of antigens separating Healthy Donors (H.D.) from other groups using **rank-sum** and FDR/Bonferroni

HD vs. Scaling method	Multiple Sclerosis	Osteoarthritis	Rheumatoid arthritis	Type II Diabetes
No scaling	14/2	74/62	72/63	75/70
Divide by mean	10 / 3	42/26	54/30	58/43
Divide by geometric mean	11/3	48/27	61/31	61/42
Log + subtract mean	11/3	48/27	61/31	61/42
Divide by median	7/3	47/25	50/34	60/42

Separating antigens between Healthy Donors and patients with different diseases.

Table 4: IgG+IgM data - Number of antigens separating H.D. from other groups using **t-test** and FDR/Bonferroni

HD vs. Scaling method	Multiple Sclerosis	Osteoarthritis	Rheumatoid arthritis	Type II Diabetes
No scaling	51/0	72/66	72/67	67/64
Divide by mean	7/3	50/33	53/35	47/33
Divide by geometric mean	6/3	54/39	55/38	53/41
Log + subtract mean	6/3	54/39	55/38	53/41
Divide by median	5/3	54/33	50/36	51/37

Separating antigens between Healthy Donors and patients with different diseases.

As one can tell by looking at the numbers, the data is slightly problematic. It is unreasonable that about 40 out of the 80 antigens tested discriminate between healthy donors and sick people, after using the strictest method of Bonferroni. We are dealing with quite small groups of 10 to 30 patients in each group. Although statistical analysis done on a small group can bias the results, it is still unreasonable that over half of the antigens appear as discriminating between healthy and sick patients. Among the p-values of the discriminating antigens one can find very low p-values, such as $10^{-10,}$ and even less.

Even in terms of low statistics, p-values of 10^{-10} are very significant. We know that most of the examined antigens are not related to autoimmunity in general or to these diseases in particular. We decided to look more carefully at diabetes type II patients in comparison to healthy donors. We chose this group since it has the most patients in it. We found it very strange that at least 58 antigens were found as significantly discriminating between the two groups, using rank sum test and FDR (Table 3). The p-values were very low; the top 20 antigens got p-values of 10^{-10} - 10^{-8}. One can look (Figure 4) at the scaled data, even without conducting the statistical tests, and notice the clear separation of these two groups of patients.
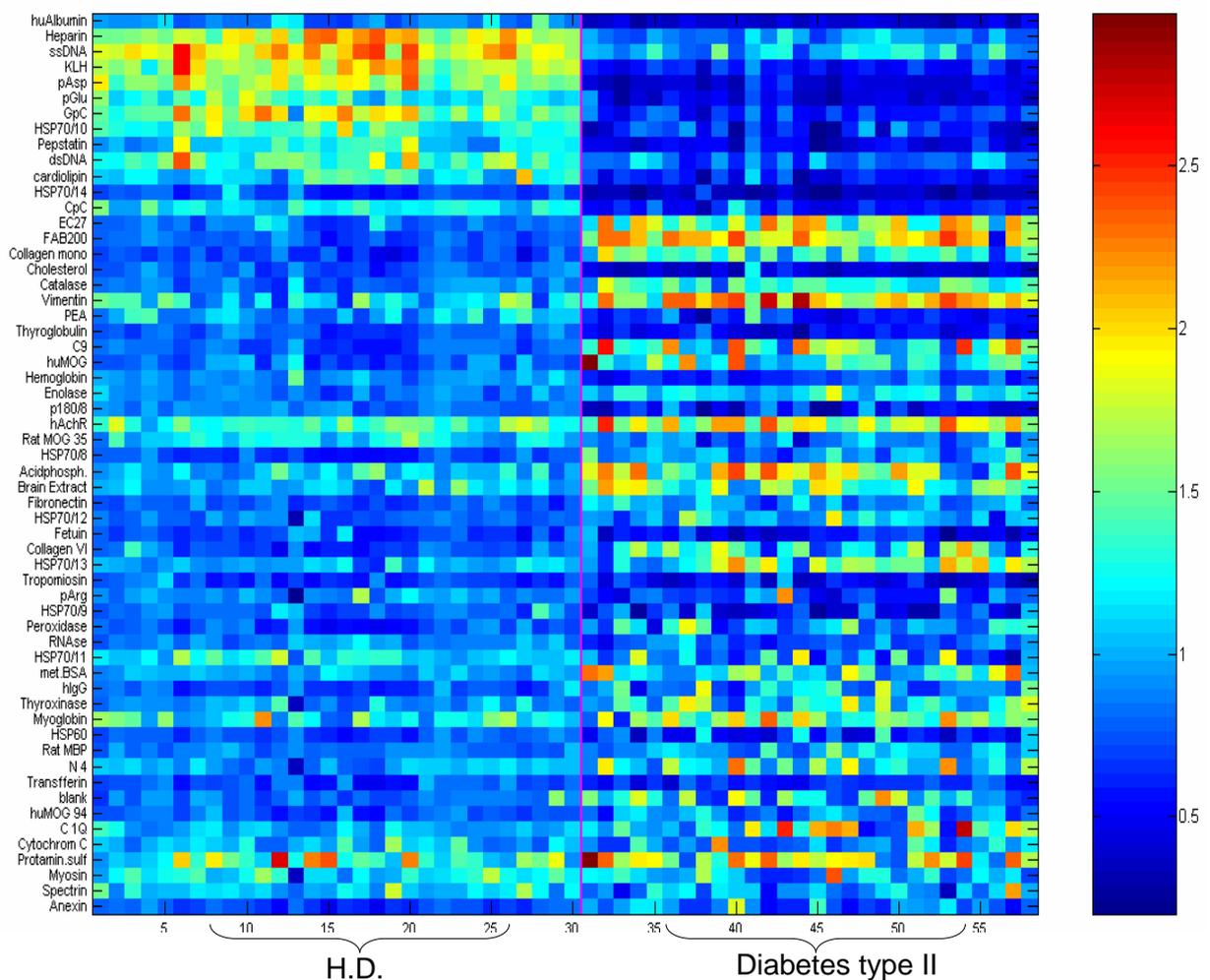
Figure 4: o.d. readings, using IgM+IgG data, of 58 antigens which separate significantly the H.D. patients and diabetes type II patients. The expression data appear after scaling using the divide by mean method. On the left – antigen names, sorted according to their p-values

- 19 -

1.5.2 P-values

In order to verify the results obtained from the statistical test, we collected the lowest p-values after each scaling method and ranksum test. The aim was to verify the consistency of the results regarding the scaling method. We plotted the mean of the lowest p-values in the different comparisons we made using different scaling methods (Figure 5). These lowest p-values are stable despite changing the scaling methods. The p-values which were obtained from our tests on the two different matrices showed high similarity and consistency. We can conclude from the above that the results are reliable and the high number of discriminating antigens is not the result of the scaling method or statistical test used.
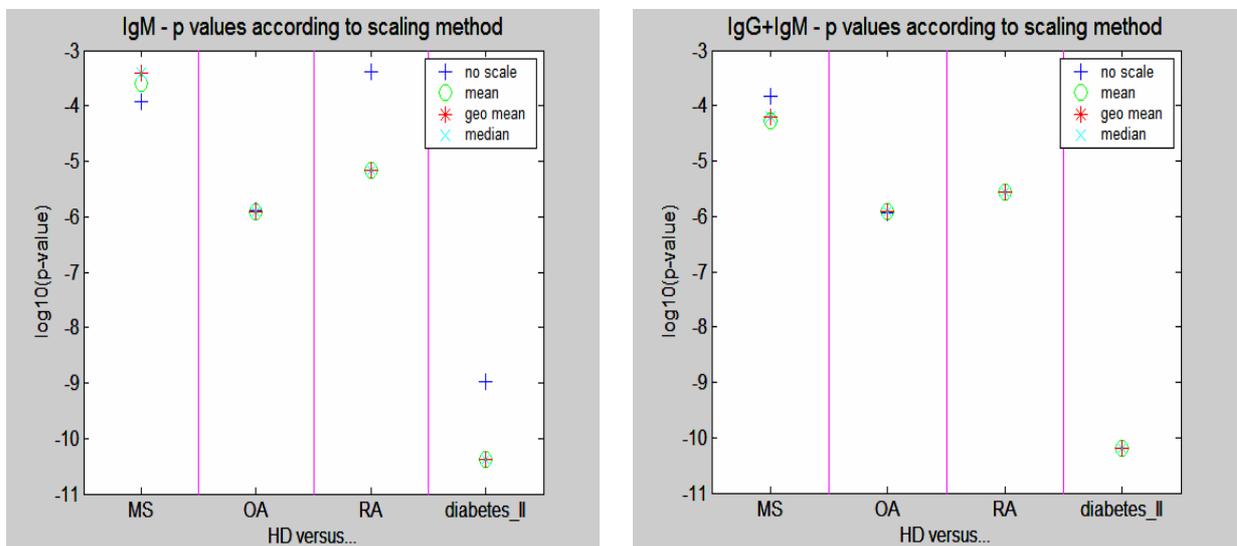

Figure 5: plot of lowest p-values received from the comparison of H.D. versus each disease using ranksum test. Each scaling method is denoted with different shape and color. On the left – results from IgM data. On the right – results from IgM+IgG data

1.5.3 Diabetes type I experiment

We conducted the same tests on other data from the same lab, which was published earlier (Quintana et al 2003). The data is composed of 88 antigens (approximately 60% appear in the current experiment as well). The data is related to H.D. and diabetes type I patients. The group included 20 H.D. and 20 diabetes patients. The results for IgM+IgG showed very few antigens (6-8) which significantly separate the H.D. from the diabetes type I patients (Figure 6). Using IgM data there were eight antigens which discriminate the group of H.D. from the sick patients, after conducting FDR. Conducting the Bonferroni test we found only two and one separating antigens for IgM+IgG and IgM

data, respectively. All six antigens which were found to be discriminators in IgM data were found in the IgM+IgG data as well. These results are compatible with the unsupervised analysis which was performed on this data previously (see Quintana et al 2003).
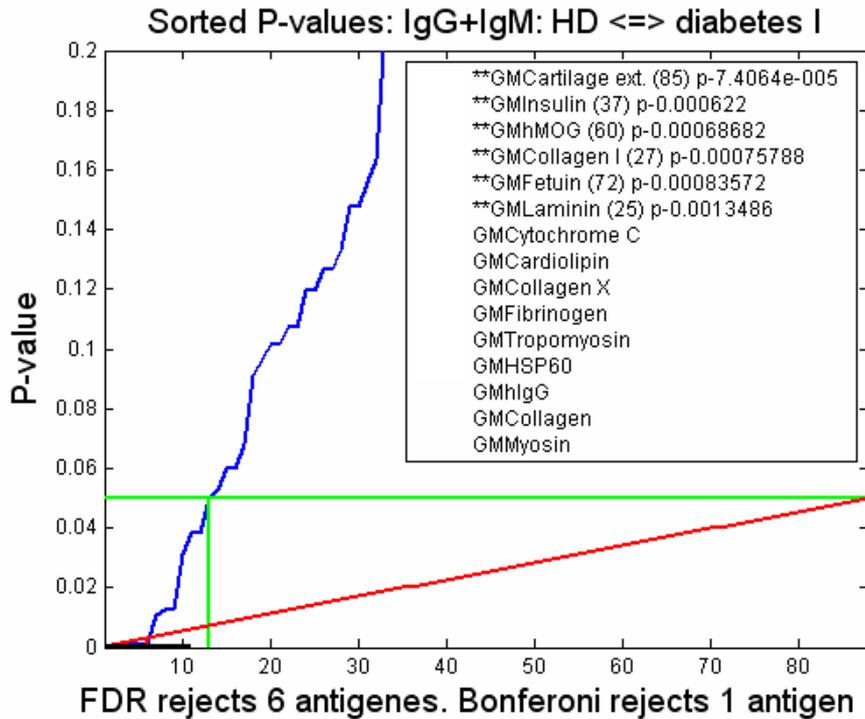

Figure 6: Diabetes type I experiment, IgG+IgM data. Sorted p-values of 80 antigens after ranksum test. Scaling method is 'dividing by geometric mean'. Number of separating antigens, with 0.05 FDR, is 6 out of 88. Number of separating antigens, passing Bonferroni, is 1 out of 88. On the legend are the separating antigens names marked with **.

1.5.4 Healthy donor's combination

We tried to increase the size of the H.D. group in order to increase the population size, and gain better statistical results. This was done by combining our data with the Diabetes type I experiment data, which was described above. We combined 30 H.D. samples from our experiment and 20 from the diabetes type I experiment. In order to combine the two groups we took the antigens that appear in both experiments and used the relevant measurements. We created a new matrix of the combined data and scaled the samples using different scaling methods (same as above). After scaling we conducted ranksum test and FDR in order to see whether the two groups of H.D. can be discriminated. Such a separation would be indicative of a serious problem with the data. We found that 75% (29

out of 38) of the overlapping antigens gave a good separation between the groups (Figure 7). We discarded the top 10 separating antigens from the raw data, in order to eliminate the possibility that we got defective antigens as a result of bad experiments, whereafter we started the whole process again. We did all the calculations of scaling and statistics, and we got approximately the same results – a very high percent of separating antigens. This time we took only part of the overlapping antigens in different cross-sections in order to avoid bias of specific antigens. We scaled and repeated the statistical tests again and the results were approximately the same. We didn't find any manipulation that could combine the two groups without any discrimination. This lead us once more to the conclusion regarding severe unexplained artifacts in the data.
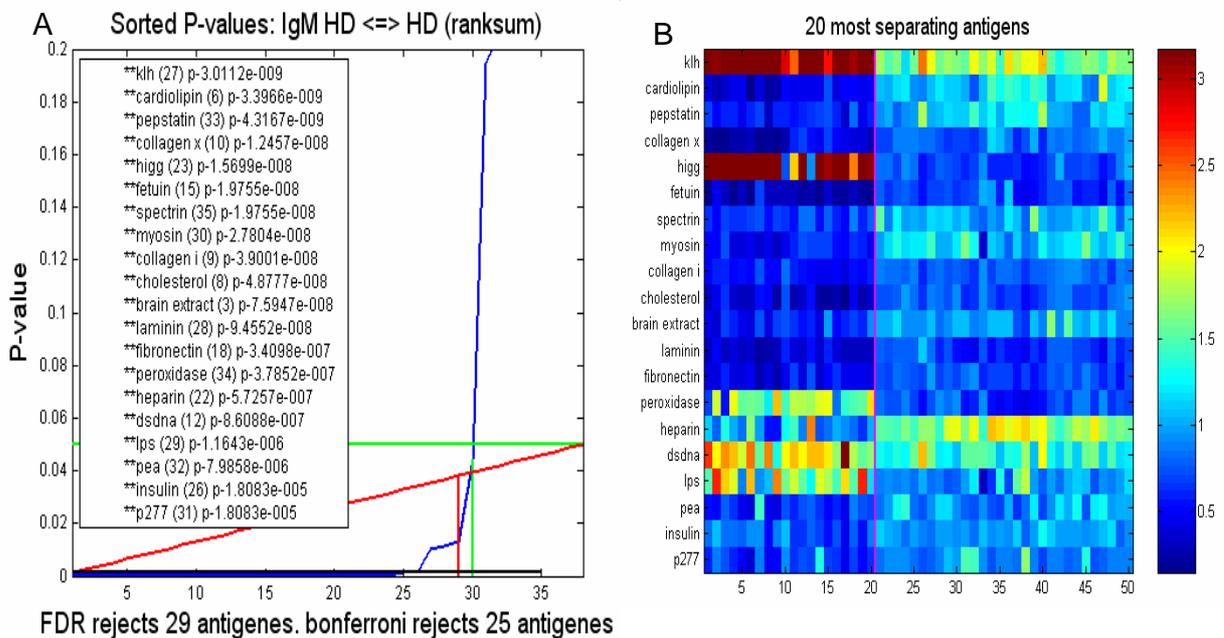
Figure 7: Healthy donor's combination between current experiment and diabetes type I experiment, IgM data. A-sorted p-values of 38 antigens after ranksum test. Scaling method is 'divide by mean'. Number of separating antigens, with 0.05 FDR, is 29 out of 38. Number of separating antigens, passing Bonferroni, is 25 out of 38. On the legend are the 20 top separating antigens names. B- O.D. readings of the 20 most separating antigens. On the left – antigen names, sorted according to their p-values.

1.6 Summary

We are dealing with a new kind of data. This data can leads us to a better understanding and better interpretation of autoimmune diseases. Better understanding of the etiology can lead us to new possible treatments of these diseases, and new targets for further investigation of the disease pathophysiology. It still seems that the data is unstable and contains a lot of noise. While working on the current data, we realized that there were severe artifacts that exclude a possible biological interpretation. We think that all the effects we found in the data were caused by technical problems. There may be a problem arising from low statistics, but we believe that the major problem, of too many separating antigens, is caused by reasons other than the statistical problems. The analysis we did on the diabetes type I data supports this hypothesis.

2. Antigen Chip

2.1 Introduction

The advent of DNA microarray technology during the last decade has led to an explosion of studies aimed at identifying novel mRNA transcripts, or patterns of gene expression whose transcriptional up or down regulation is characteristic of particular diseases or phenotypes [21]. DNA microarray analysis may lead to significant biological insights, including better understanding of biological processes, improved ability to identify diseases and cure them with higher efficiency. Despite the success of DNA microarrays in gene expression profiling, it is the activity of the encoded proteins that directly manifest gene function [22]. Moreover, use of RNA transcriptional profiling has important limitations and it is unsuitable to provide the comprehensive understanding needed to study complex diseases that operate on the organism level, such as autoimmune processes.

RNA transcriptional profiling alone is inadequate for studying human autoimmune diseases, for several reasons. First, these diseases do no manifest at the level of RNA transcription, but rather at the protein level. Second, the correlation between RNA expression and protein concentration and function is often weak [23]. Messenger RNA undergoes a variety of processing events that can profoundly affect cell phenotype, yet are not revealed in current transcriptional profiles. For example, there exist mRNA molecules that encode certain apoptosis-regulatory molecules in two or more alternatively spliced forms, encoding proteins with opposing functions. Translation of mRNA into protein is also regulated by translational regulatory elements. Third, protein function can be regulated by posttranslational modifications by enzymes such as kinases or proteases. Finally, autoimmune responses are regulated by autoantigen-specific T and B lymphocytes expressing distinct and heterogeneous antigen receptors that are not easily examined by transcriptional profiling [21]. Instead of dealing with gene expression, we should move towards the proteomics area, which involves antigens and antibodies with their interactions. The mechanisms of production of high-affinity autoantibodies that underlie most autoimmune diseases should be the target of research in the coming post-genomic era [24]. This type of research, integrated with new technologies, can lead to more efficient study of autoimmune diseases.

Identification of autoantigen targets of B and T lymphocytes has a central role in understanding the pathogenic mechanisms governing the initiation and propagation of autoimmune diseases. The utility of autoantibody profiling is based on the hypothesis that overall specificity of the autoantibody response correlates with helper T-cell autoreactivity, which drives many autoimmune diseases [25]. For many human autoimmune diseases, autoreactive B- and T-cells' responses are directed against the same targeted autoantigen, and in some cases the same immunodominant epitopes [24]. Even if there is discordance between the fine specificity of the B- and T-cell responses, the ability to use autoantigen arrays to identify the specific self-protein(s) targeted by the autoimmune response is of tremendous value for identifying and evaluating candidate autoantigens.

Moreover, we can utilize the fact that the antibodies repertoire/pattern can serve as the "fingerprint" of an individual's autoimmune response [3]. This fact can lead to new technologies to serve as a classification tool of antigen related diseases. One hopes to be able to classify patients according to their "fingerprint", determine their disease and give them personalized medicine, and even find new disease subgroups that are not yet known.

We developed a large scale data acquisition method, which can deal with hundreds of different antigens in different concentrations simultaneously in a single experiment. This is a big step towards better understanding and analysis of autoimmune diseases, compared to the currently used technologies. Till recently, research dealt with one antigen at a time, which was time and money consuming as well as not always being helpful. Miniaturized and highly parallelized immunoassays like our chip will reduce costs by decreasing reagent consumption and improve efficiency by greatly increasing the number of assays that can be performed with a single serum sample [26]. This system will significantly facilitate and accelerate the diagnosis of autoimmune diseases [26]. Protein array monitoring of autoantibody responses has the potential to improve care for patients with autoimmune diseases by permitting identification of 'biosignatures' for diagnosis, prognostics and guidance of tolerising therapy [27]. The rapidly advancing analysis tools for DNA-microarrays data can be helpful for the analysis of our antigen chip as well. This new technology puts us at the cutting edge of the proteomics area. In summary, there are many potential applications of autoantigen array technology, for example: 1) rapid screening for autoantibody specificities associated with autoimmune diseases to facilitate

early diagnosis and treatment, as well as follow up during therapy; 2) characterization of the specificity, diversity and epitope spreading of autoantibody responses; 3) determination of isotype subclass of antigen-specific autoantibodies to assess their potential pathogencity and the relative T-helper type I (Th1) versus T-helper type II (Th2) bias of autoimmune response against a particular antigen; 4) guiding development and selection of antigen-specific therapies for use in the clinic; and 5) as a discovery tool to identify novel autoantigens or epitopes [1, 21, 22, 24, 28]. In a broader view, protein microarrays can also provide a powerful tool for expression analysis of defined proteins or receptor ligand interactions [26], or even to determine a profile of a protein. One can print many different peptides on a chip with large variety of structures and characteristics. By hybridizing a protein to the chip one can characterize a protein in terms of its binding affinity to diverse peptides.

2.2 Proteomics - background

Several new methods have been introduced, trying to deal with the limitations of the information obtained at the transcription level. The new methods are trying to study and deal with the expression, function and interactions of proteins expressed in a tissue or organism [21]. This field is termed proteomics. The ideal assay for detecting proteins and their interactions should be sensitive, specific, and reproducible. There are few common and well known assays, including western blot, two-dimensional gel electrophoresis, mass spectrometry, enzyme-linked immunosorbent assay (ELISA) and the yeast two-hybrid system. The major problem with these assays is that they are not amenable to "multiplex analysis" whereby one can simultaneously screen thousands of individual proteins for their ability to interact with other molecules (except maybe the two hybrid system). I will give a short description of a few of these methods.

The two dimensional gel approach deals with proteins extracted from cells and loaded on the gel. The crude protein mixture is applied to the 'first dimension' gel strip that separates the proteins based on their isoelectric points (protein's charge). The 'second dimension' is SDS-PAGE gel, where proteins are denatured and separated on the basis of size [29]. The disadvantages of this method are poor reproducibility between labs, and the difficulty in exact identification of various proteins. Another disadvantage is the low capacity of this method.

The mass spectrometry approach combines the gel method with immunoprecipitation, which recognizes proteins on the gel using anti-phosphotyrosine antibodies. After the first step the protein band is excised and subjected to digestion with trypsin (cleavage after arginin or lysine) to peptides. The peptides are run through the mass spectrometer and the protein's components – peptides – can be identified very precisely (Figure 8). This is a high throughput system for protein identification, but doesn't give an answer to the problem of interactions.
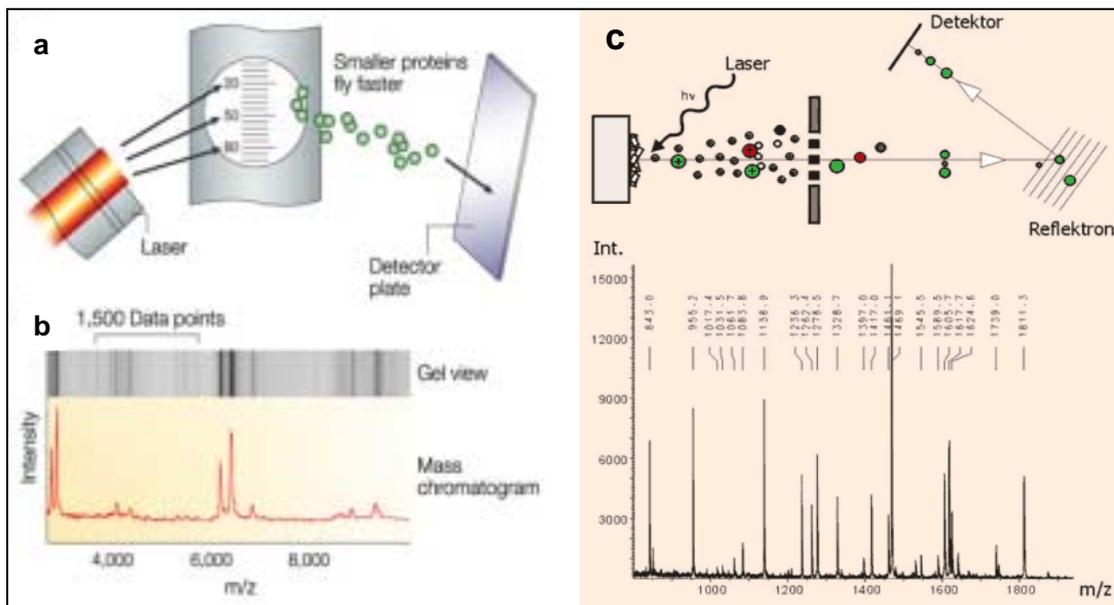
**Figure 8: a+b** - SELDI –TOF mass spectrometry. **c** – Maldi mass spectrometry.

The two hybrid system is a method which one can use to study protein-protein interactions. This method is applied extensively in the yeast organism and gives reasonable results. This is a genetic method based on the modular structure of transcription factors; the key player in this method is GAL4, a yeast transcriptional activator protein, which has a DNA-binding domain and an activation domain that induces increased transcription of a reporter gene containing a binding site for GAL4 [30]. In order to test the binding of proteins A and B, one uses two chimeric proteins; one made by fusing the DNA-binding domain of GAL4 with A, and the other by fusing the activation domain of GAL4 with B. If the two proteins A and B interact and bind, the GAL4 DNA-binding and activation domains get into proximity, and the complex acts as 'normal' GAL4 and enhances expression of a reporter gene. The reporter gene is then transcribed and translated and is used as a marker for the interaction. This method is

usually conducted using cDNA libraries, where each colony of yeast contains two ORF's fused to the components of GAL4. One of the main consequences of this is that once a positive interaction is detected, the ORF is identified simply by sequencing the relevant clones. For these reasons it is a generic method that is simple and amenable to high-throughput screening of protein–protein interactions. This method can detect also indirect interactions; it's main limitation is the need to use genetic manipulation [29].

ELISA – is briefly described in the methods 1.4.4, on 'Human diseases' analysis' chapter. This method is the most accurate [31], most relevant to our goals (which are to identify antibody-protein interactions in the immune system). The difficulty of this method is in its up-scaling; quite lot of work is required and one can not deal with thousands of antigens. Moreover, a lot of material is required for using ELISA. In order to get a result using animal serum, one should use approximately 100 μl for each test in order to examine dozens of antigens. This is considered a very high quantity for such small animals.

The only solution for high throughput screening of hundreds of antigens is protein arrays. There are two general types of protein microarrays. Firstly, analytical microarrays in which antibodies, antibody mimics or other proteins are arrayed and used to measure the presence and concentrations of proteins in a complex mixture. Secondly, functional protein microarrays, in which sets of proteins or even an entire proteome are prepared and arrayed for a wide range of biochemical activities [22].
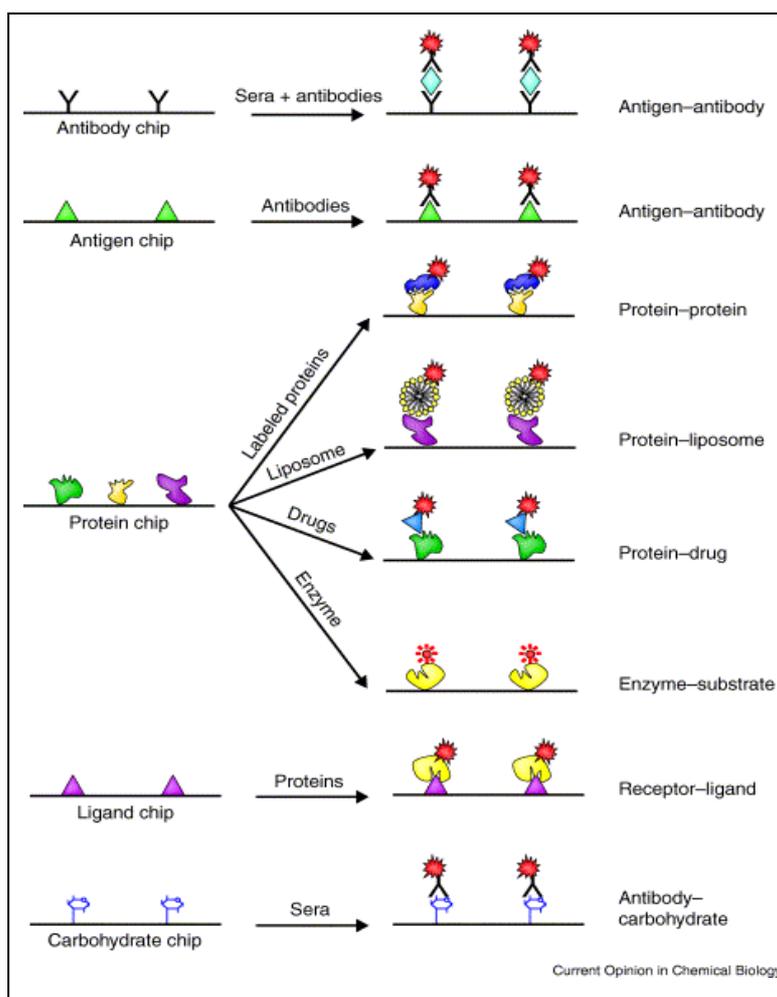
Figure 9. Applications of protein microarrays. There are two general types of protein microarray: analytical and functional protein microarrays. Analytical microarrays involve a high-density array of affinity reagents (e.g. antibodies or antigens) that are used for detecting proteins in a complex mixture. Functional protein chips are constructed by immobilizing large numbers of purified proteins on a solid surface. Unlike the antibody–antigen chips, protein chips have enormous potential in assaying for a wide range of biochemical activities (e.g. protein–protein, protein–lipid, protein–nucleic-acid, and enzyme–substrate interactions), as well as drug and drug target identification. Small molecule and carbohydrate microarrays are other types of analytical microarrays that have been demonstrated to be capable of studying protein binding activities to ligands and carbohydrates [22].

Analytical microarrays involve a high density array of affinity reagents that are used for detecting proteins in a complex mixture. The most common form of analytical arrays are antibody arrays in which antibodies (or similar reagents) that bind specific antigens are arrayed on a glass slide at high density. A lysate is passed over the array and the bound antigen is detected after washing. The detection can be made by using a pre-labeled lysate or secondary antibody which recognizes a specific antigen of interest. This method faces a challenge of producing reagents that identify the protein of interest with high enough specificity. Using this method one can gain a profile for each protein of interest. The most significant problem with antibody arrays is specificity. Proteins are

often present in a very large dynamic range (10^6) of concentration; thus, reagents that might have high affinity for one protein, but have a low affinity for another will still exhibit detection of the lower affinity protein if it is much more prevalent. To avoid this problem, many groups have turned to using sandwich assays, in which the first antibody is spotted on the array and then the antigen is detected with a second antibody that recognizes a different part of the protein. This approach dramatically increases the specificity of the antigen detection, but requires that at least two high-quality antibodies exist for each antigen to be detected [22].

Our method falls into this category of analytical microarrays.

The second category is the functional protein microarrays that are constructed by immobilizing large numbers of purified proteins on a solid surface. Protein chips have enormous potential in basic research, as well as drug and drug target identification (Figure 9). This kind of chip can also be used for conducting enzymatic assays or to detect protein-protein interactions in high-throughput system.

Several groups, including MacBeath and Schreiber at Harvad University [32] and Haab, Dunham, and Brown at Stanford University [33], and Robinson et al at Stanford University [1, 21, 24, 27], developed methods utilizing a robotic arrayer, such as the one which was developed in the laboratory of Dr. P.O. Brown to fabricate DNA microarrays[33]. Their methods where used to deposit proteins onto microscope slides in order to fabricate high-density protein arrays. The protein arrays described by these investigators contain hundreds of different proteins attached to the surface of microscope slides, where they are analyzed for interactions with other proteins, enzymes, or drugs.

MacBeath and Schrieber demonstrated that simple protein arrays can be used to detect interactions of immobilized proteins with antibodies, with cellular proteins such as p50 nuclear factor, and others. They also demonstrated use of protein arrays to detect the enzymatic activity of kinases, with a panel of specific immobilized protein substrates [32].

Haab et al used microarrays in a comparative fluorescence assay to measure the abundance of many specific proteins in complex solutions. Two complex protein samples, one serving as standard for comparative quantization, the other representing an experimental sample, were labeled by covalent attachment of spectrally resolvable fluorescent dyes. They characterized the reactivities of 115 antibody-antigen pairs, demonstrating that 50% of the arrayed antigens and 20% of the arrayed antibodies provided specific and accurate measurements of their cognate ligands. A small subset of

antibodies was capable of detecting antigens in complex solutions at a level that would allow their use in certain clinical applications, including detection of human serum proteins present at ng/ml quantities [33]. The spotted antigen arrays developed by Haab et al enabled sensitive detection of purified antibodies and antigens diluted into complex solutions of proteins.

Joos et al described the construction of autoantigen microarrays containing 18 prominent autoantigens spotted onto surfaces including silane-treated glass slides and nitrocellulose [26]. Their arrays proved to be both sensitive and specific for detection of autoantibodies to many spotted antigens. Bound antibodies were visualized using a secondary antibody conjugated to horseradish peroxidase prior to addition of a luminescent substrate, and imaged using a charge-coupled device chemiluminescence camera. Their array was a three-layer array with known antigens and known antibodies.

Robinson et al described the development of protein microarray technology to study the specificity of the autoantibody response in murine and human autoimmune diseases [24]. They used a robotic arrayer to attach peptides, proteins, nucleic acids, and protein complexes on microscope slides (196 distinct biomolecules). Individual arrays were incubated with serum from autoimmune disease patients and control. The detection of specific autoantigen-autoantibody binding was made using secondary antibodies which were covalently conjugated to spectrally resolvable fluorescent labels (such as Cy3 or Cy5). They demonstrated sensitive and specific detection of characteristic autoantibodies in serum derived from patients with eight distinct autoimmune diseases. Their technique demonstrated the accuracy of detection using tiny amounts of antibodies present in the serum, not alike other methods such as ELISA which require higher amounts of material.

Unlike DNA chips such as the Affymetrix microarray, the protein-chips field is not dominated by a broadly used commercial tool. Therefore, in order to produce such a chip, one must do some research and develop the technology almost from scratch, similarly to the technology of the cDNA chips in their beginning. This is our main goal.

The Antigen Chip we are trying to develop in this work has a resemblance to the antigen chip developed in parallel by Robinson et al. The technique used by both teams to produce the two chips is similar, both relying on cDNA robotic tools and dealing with antigen and antibody binding. Yet our idea and goals differ greatly. Robinson et al. tried to develop a system which could be termed as a "mini-ELISA", investigating and quantifying specific antigens which are known to be relevant to a specific disease. Our team inspired to investigate a broader scope of the autoimmune diseases, and wished to

print a wider range of antigens in order to get an overview picture of the autoantibodies present in these diseases, thereafter investigating the patterns that were formed without a preconceived vision of the results. Since no "chip producing kit" exists, we intend to build a system based on our specific requirements and calibrate it to fit our needs. There is quite lot of work that has to be done in order to obtain a satisfactory chip in our lab. Once the system is stable and reproducible, we will have a tool to further investigate biological questions of interest.

2.3 About the new technology

The Antigen chip technology is based on one central feature: the affinity between the antigen and the antibody; this fact dictates our chip development. The chip is glass made; the main idea resembles cDNA chips in the sense of high flexibility in chip design and development. Each biological question or experiment can lead to a totally different chip in terms of arrayed elements (antigens/peptides).

2.3.1 Printing, Hybridizing and Scanning

We used 'TeleChem International, Inc.' ('arrayit') company slides. The slides where made from glass with a surface that binds the antigen (Figure 10). We used two different surfaces, starting with 'Super Aldehyde' and switching to 'Super Epoxy' which gave better results [34, 35].
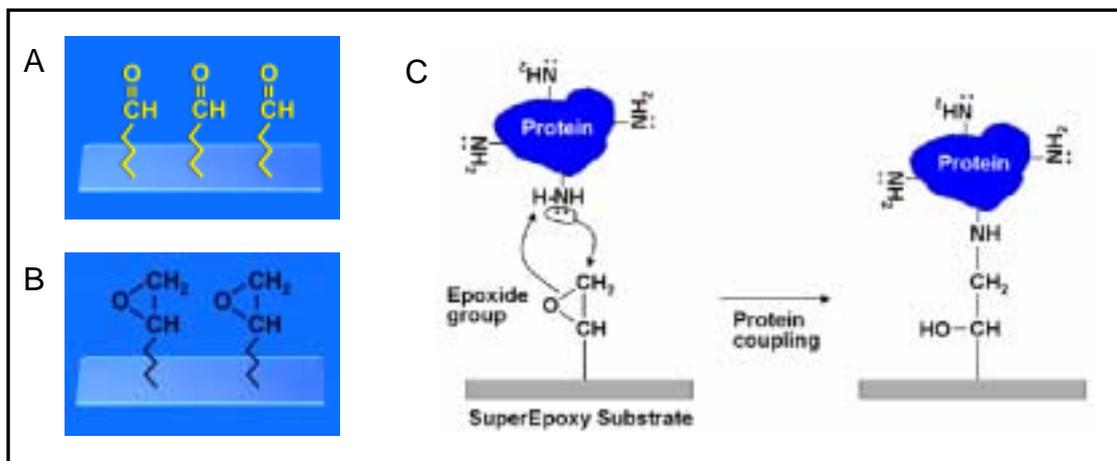

Figure 10: chip surfaces. A-Super Aldehyde substrate. B-Super Epoxy substrate. C-the coupling process of a protein/antigen to the slide surface.

We attached various antigens to the surface (Figure 12a). In each experiment we used different sets of antigens as our probes, depending on the biological question we investigated. In order to attach the antigens or peptides to the surface, we used a spotted array printing robot. We prepared our probes in a 384 wells plate; in each well we deposited a different probe or different concentrations of the same probe, depending on whether we needed to check greater variety of antigens or whether we were looking for the antigen titer. We programmed the robot to print one or more spots from each well. Two sub-slides were printed independently (top & bottom) on each slide, meaning that we could check different samples in each sub-slide. After spotting we continued with the blocking process. This process is composed of a few steps which include washing and blocking. First we washed the chip, mainly with PBS or PBST with low concentration (0.005%-0.05%), to get rid of dust or any other unspecific binding to the glass. Next, we blocked all the free sticky regions on the glass, using BSA 3% of unrelated animal. The aim of this step was to give us a uniform and low background without noise. The blocking was carried out overnight, followed by another washing series. This last washing series prepared the chip for the next step; incubation with the patient's serum (Figure 12b). During incubation with the patient's serum, antibodies that recognize an antigen/peptide, will attach to it and make an antigen-antibody complex/interaction. In order to avoid non specific interaction we repeated the washing procedure. The next step is the incubation of the chip with anti-IgM and anti-IgG antibodies (Figure 12c). The purpose of this step is to detect the antibodies that are bound to the first layer (the antigens). These specific antibodies are important since they are responsible for the first and general immune response (IgM), and for the specific secondary response (IgG) against the disease. The detection is made by antibodies labeled by Cy3 and Cy5 dyes. The labeled antibodies are specific to the IgM or IgG tail, each interacting with a different tail of the bound antibodies from the patient's serum. After detection we used a spotted array scanner to start the scanning step of the process. For each fluorescent dye we used different scan and gain power depending on the dye and the total signal (see details below). The scan power is the intensity of the laser illuminating the slide, while the gain power is the power of the detector that collects the light emitted by the fluorescent dye after excitation. After scanning, the chips were stored at -20$^{\mathrm{o}}$c.
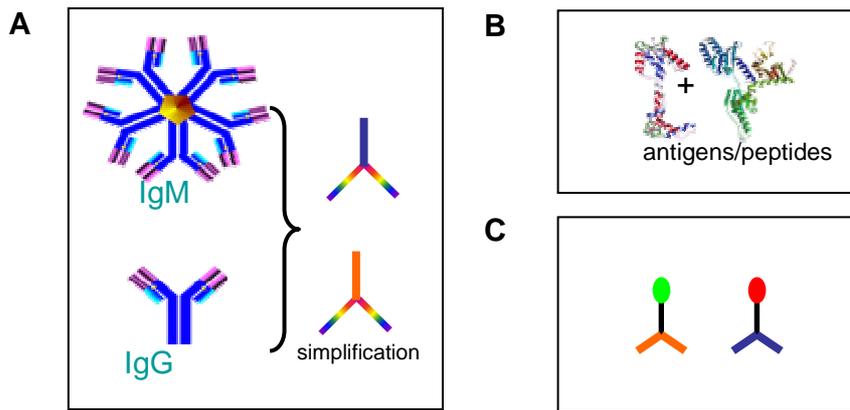
Figure 11: **A**-Structure of IgM antibody which reacts at the beginning of the immune response, and the structure of the IgG antibody that reacts in the late immune response. They recognize various antigens using their 'arms'. Each lineage has its own specificity in the arms, however all the lineages of IgM or IgG have a common fixed tail. **B**-Example of antigens/peptides which are printed on the slide surface. **C**- Labeled secondary antibodies. The green colored antibody's arms interact with the tail of the IgG antibody and the arms of the red **-** with the tail of the IgM antibody

Each scan ends with an array image in TIF format. This file is loaded to a software application called 'QuantArray'.
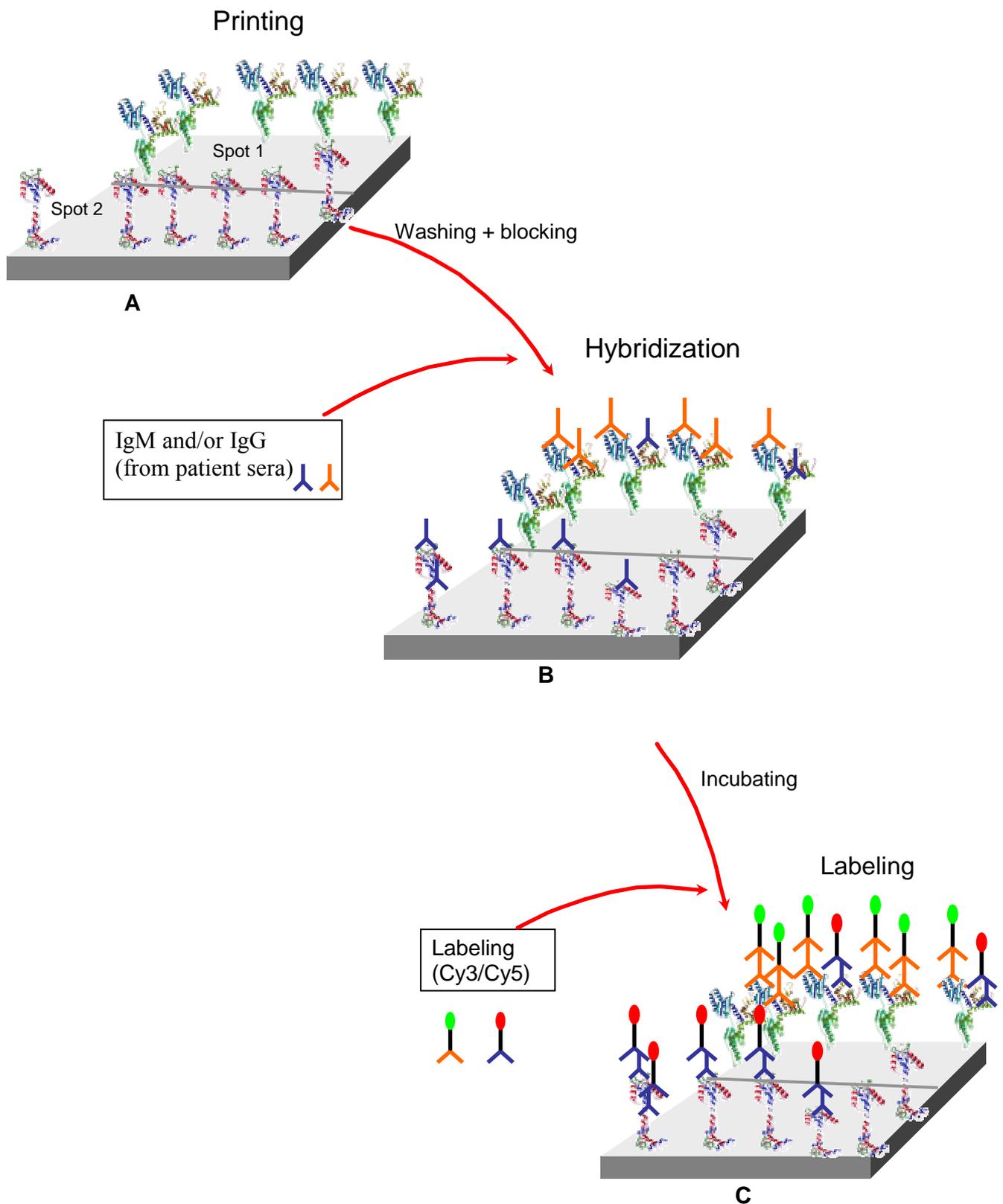
Figure 12: Antigen chip production. The surface in the figure represents two different spots on a chip. On each spot we printed one kind of antigen. A-Attaching the probes to the chip (antigens/peptides). B-Hybridization of the antibodies produced from patient sera, with the antigens on the chip. C-Labeling all bound antibodies, using labeled secondary antibodies (anti IgM + anti IgG).

2.3.2 Quantifying

The GSI Lumonics QuantArray Quantitative Microarray Analysis application program is designed to quantify the fluorescence intensity of spots on images that have been generated with the GSI Lumonics ScanArray Microarray Analysis System.

The program can be operated on two levels:

- As a stand-alone program that analyzes images under manual operation or as a semi-automated application, using a user generated protocol.

- In automated mode with the ScanArray Microarray Analysis System. In this mode of operation the ScanArray system sends an experiment file which consists of a series of images to the QuantArray program, which will then analyze the images according to the indicated protocol. This mode is an optional add-on to the basic program.

The analysis program is designed to be used in both research and production environments. Data can be exported to external programs for additional calculation and/or archiving.

We used the software as a stand-alone program.

The output of this program is a large excel file, includig the intensity of each spot on our chip. This data is converted to a simpler format and loaded into a MATLAB program. The rest of the analysis is carried out using this program.

2.3.3 First chip experiment

In the first chip the 'Super Aldehyde' surface was used. The results are shown in Figure 13. The first batch included four glass slides. Two subarrays were printed on each glass slide (total of eight different chips for the analysis), which are noted as 'up/top' and 'down/bottom'. We quantified these chips using QuantArray. During the analysis we saw and realized that the figures coming out of the chip were not reliable enough to supply biologically relevant quantitative information. The major problems with the chips were the following:

- There were a lot of smears in the spots, as well as merges between spots.

- The background was not uniform enough to accurately distinguish between the spot and its background, while quantifying it.

- The reproducibility was very low (each spot had 4 repeats).

- The correlation between chips was poor.

In one word –we had a "dirty" chip.
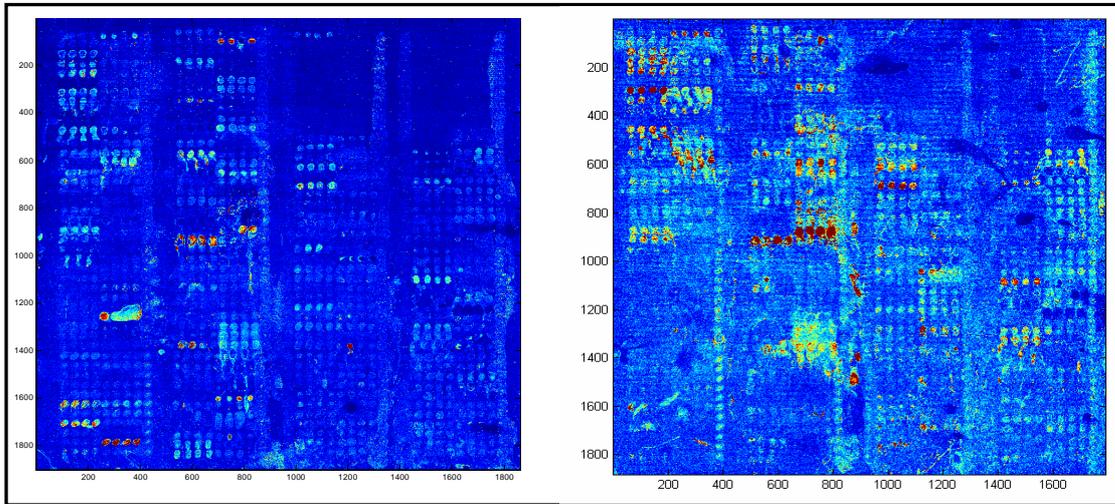
Figure 13: Two different chips pictures (in TIF format) from the first antigen-chip batch. The size of the glass slide is ~6 cm. On each glass slide two chips are spotted, each with size ~1.5 cm^2. Distance between spots' centers is 0.4 mm in the first chip and 0.6 mm in the new chip. Spot's radius using 1mg/ml protein concentration is ~0.2mm. The maximum number of spots using distance of 0.4mm is 1296 spots in 16 sub arrays and 786 spots using distance of 0.6 mm. (The number of spots relates to one plate of 384 wells)
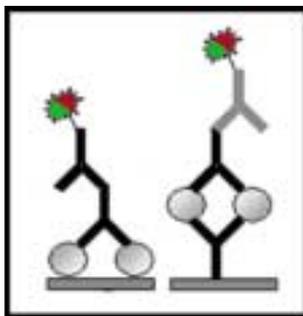
Figure 14: chip's layers abstraction. Left- Two layers chip; purified antibody IgM/IgG is being spotted on the chip, and secondary labeled antibody stick to it after recognizing the constant part of the IgM or IgG. Right- Three layers chip; on the bottom the antigen is spotted. The middle layer is an antibody from a serum which binds to it after recognition, and the top layer is the secondary labeled antibody which recognizes the antibody from the 2nd layer in the constant region.

In order to overcome the problems mentioned above, we decided to design a new experiment. The most important change in the new experiment was elimination of one layer in the chip production. In theory, an antigen chip is made of three layers. The first layer is composed of antigens bound covalently to the surface. The second layer is of antibodies from the patient's serum which hybridize with the antigens. The third is the detection layer using secondary labeled antibodies which attach to the second layer antibodies. We decided to 'skip' the first layer of antigens and print purified antibodies (IgM and/or IgG) directly on the chip's glass. This step was done in order to simplify the experiment and deal with each problem separately. Deleting the first layer reduces the noise in the experiment since we won't have the effect of the antibody specificity to the

antigens, which increases the noise and reduces the signal. After solving all the problems, we will return the antigen layer to the experiment in order to create a real biologically relevant chip. The simplification step changes our direction of study from biological questions and problems to technical and engineering issues.

2.4 The simplified chip

In this part I will describe in detail the steps in our new, simplified experiment, starting with its goals, continuing with our plan for the experiment and ending with the results.

2.4.1 Goals

Our main goal in the second experiment (second batch), was to produce a chip that has a reasonable and reliable appearance. We define a chip as reasonable if it produces a "nice looking" clean picture, and reliable – if when quantified, it yields reproducible numbers. This means that when we are looking at the chip we can identify each spot and quantify it properly. This main goal can be achieved through several sub goals listed below:

- Reduce background noise – the background should be more uniform and well separated from the spots.
- Reduce smears and leaking on the chip, so that each spot can be identified in a reliable way without any lost material or added material from others spots.
- Increase the reliability and reproducibility of the results between spots on the same chip and between different chips.
- Examine whether we can detect any concentration effect on the spot intensity or size.
- Examine whether the position on the chip affects the spot.
- Validation of the intactness of the materials used in the chip production process, as well as the production protocol.

2.4.2 Plan

In order to achieve our goals, firstly we decided to print only two chips (each of them with top and bottom parts) as a "sanity check". These chips were meant to verify that we were on the right track of improving the antigen chip. One chip was made of Super Aldehyde substrate (which was used in the previous experiment), and the second chip of Super Epoxy - a new substrate we decided to check. In these two chips we increased the distance between spots from 0.4 to 0.6 mm between spots' center. We printed several dilutions of the antibodies (IgM and IgG). The chip was divided into three section; (1) our experimental data, which consisted of mixtures of the two antibodies in different concentrations (IgM+IgG), (2) positive control, which was the fluorescent dye printed directly on the chip (pCy3 and pCy5), (3) negative control, spots that contained proteins/peptides that were irrelevant for the experiment. Each spot on the chip was composed of a fixed concentration which was determined as X=1mg/ml. In the spots which were assembled from antibodies, we used a mixture of IgM and/or IgG in different ratios in order to examine the total intensity of IgM and IgG in the same spot (as it is going to be in a real experiment with animal/human serum). The ratios between IgM and IgG where 0:100 (0% IgM and 100% IgG), 10:90 and going on in steps of 10 till 100:0 (100% IgM and 0% IgG). Furthermore, we checked different concentrations to determine the finest concentration of proteins in one spot, on the antigen chip. The detailed plan of the chip is mapped out in Figure 15. The plan was fitted to a 384 well plate; where in each well we placed an antibody with concentration of 1mg/ml or less, as described in the detailed map

Ratios	1	2	3	4	5	6	7	8	9	10	11
IgG	100	90	80	70	60	50	40	30	20	10	0
IgM	0	10	20	30	40	50	60	70	80	90	100

Dilution
X
x/4
x/16
x/64
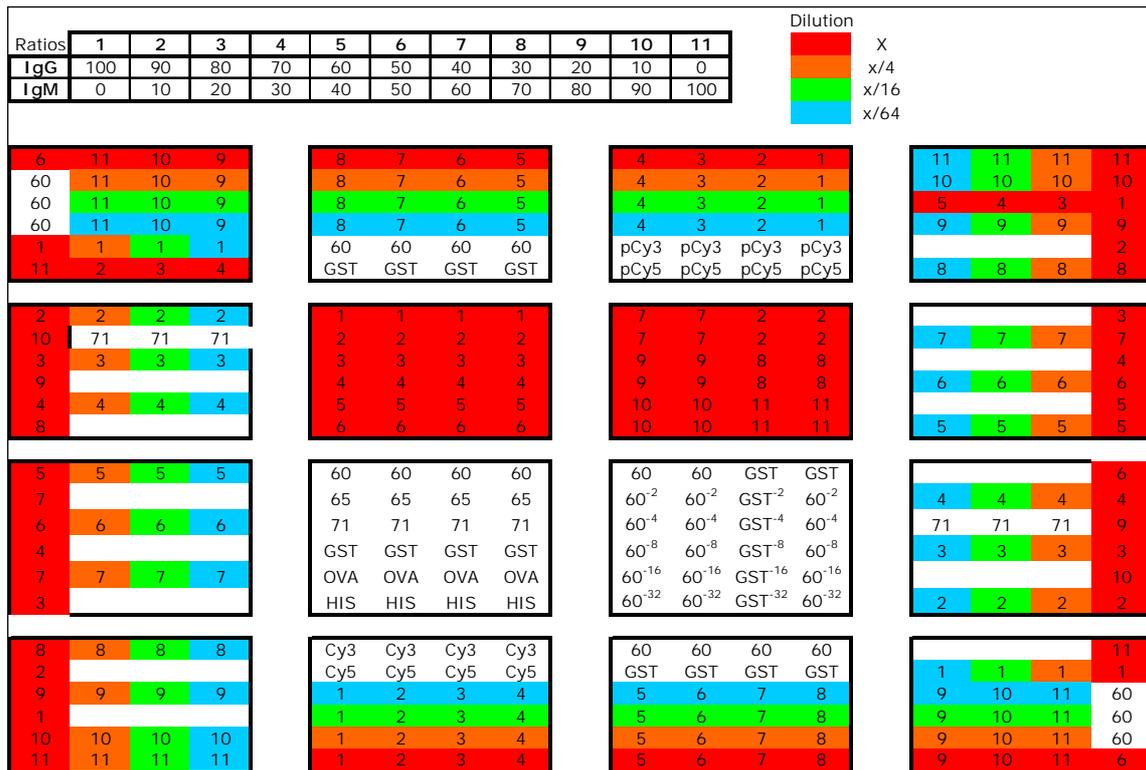
Figure 15: The chip map design. The chip is built of 16 sub arrays, where each sub array is built of six rows and four columns- maximum total of 24 spots in each sub array. Each number/letter in the map represents one spot on the chip. The chip contains three different groups of substrates; 1) The antibodies IgM+IgG, 2) Negative control which are several peptides (60, 65, 71, GST, OVA, HIS), peptides 60 and GST where spotted in several dilutions from 1mg/ml and decreasing in power of -2 3) Positive controls (pCy3, pCy5).The colors represent the concentration of the antibodies in the spot. The starting concentration was X=1mg/ml and the dilution is from X to X/64 in steps of 1/4. The spots that are presented as a number from one (1) to eleven (11) are combination of IgM and/or IgG in different ratios. Spot number one (1) contains 100% IgG and spot number eleven (11) contains 100% IgM. In the center of the chip the concentration of IgG decreases and the concentration of IgM increases in 10% steps (look in the small table in the top of the figure. The total concentration of each spot is constant and marked with the relevant color.

2.4.3 Results

The first astonishing fact was the huge difference between the two different coatings on the chip. The "Super Epoxy substrates" significantly improved the background uniformity and the leaking effect (Figure 16).
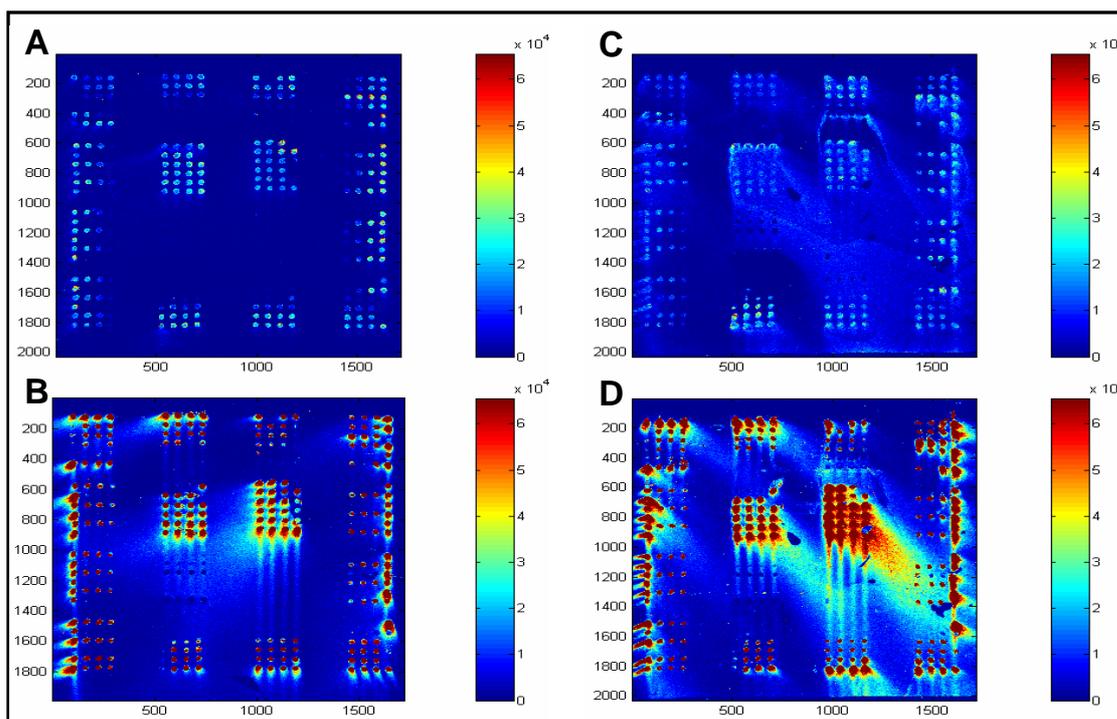
Figure 16 : four chip pictures from the 2nd batch. Chips A+C were illuminated with Cy3 wave length with laser power 70 and gain power 80. Chips B+D were illuminated with Cy5 wave length with laser power 80 and gain power 80. Chips A+B are - "Super Epoxy" chips with low and uniform background. Chips C+D are - "Super Aldehyde" with non uniform background and leaking.

The new grid spacing seemed to eliminate almost completely the spot merges, and each spot could be identified without any doubt. The overall shape of the spots was circular. The concentration seemed to directly affect the spot size. The more we reduced the concentration of the antibody/protein – the smaller the spot size became. The tendency of the spots was to be condensed in the middle of the printing area. The more the concentration decreases the more the spot tends to be smaller and located centrally. The negative controls gave no signal which was expected. The overall materials in the experiment seemed to be intact and can be used for the next experiments. Nevertheless some new difficulties and observations where raised in this experiment:

1. The positive controls (pCy3/pCy5), which where spotted on the chip, weren't detected during the development step (Figure 17B). A possible cause of this phenomenon is that the positive controls were stored in glycerol. Studies have shown that glycerol may not always stick to the chip coating, therefore causing the antibody (pCy3/pCy5) to be washed from the chip [36].
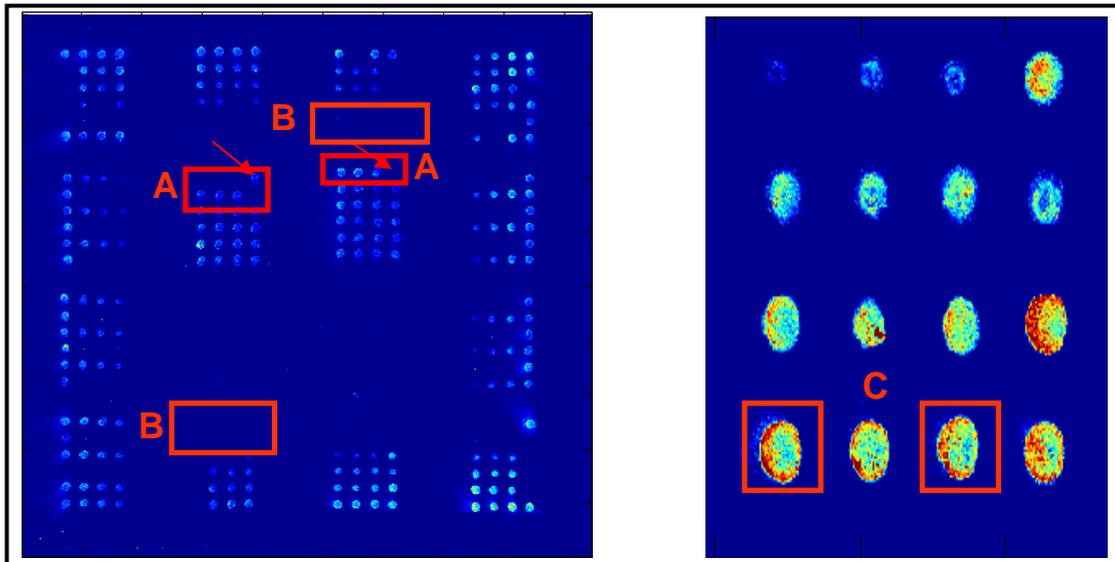
Figure 17: 2nd batch chip results. The picture demonstrates the new difficulties we observed in this batch. **A**- Misplaced spots – spots which were printed in the wrong location or missing spots, caused by human error while filling the 384 plate. On the left- the rightmost spot was spotted above the correct row. On the right- the fourth spot is completely missing. **B**- Absent positive control. In this location the positive control of pCy3 and pCy5 should have been seen. **C**- Spot rings – the spot is not uniform and tends to have rings.

2. There were some misplaced/missing spots on the chip (Figure 17A). In the central part of the chip each antigen was spotted four times; therefore there should be exactly four spots in each row. Human error was probably the cause for the misplaced spots. While filling the 384 wells plate manually, some of the initial materials may have been placed in the wrong well, which led to wrong position on the chip. Since it is important that one should know exactly what each spot represents during analyzing, misplaced spots shuffle the results, which may have a dramatic influence on the results of the biological data.

3. A new effect, of 'spot rings', was observed (Figure 17C). This is a ring containing a high concentration of protein which is translated into a red color during the scanning step. The ring is at the margin of the spot, although doesn't always encircle the whole spot. A spot with a 'spot ring' looks brighter in the center and darker in the margins. This effect can cause imprecise quantification in the step following scanning. Most likely, the cause of these spot rings is rapid drying of the spot after the printing stage. The same can be observed happens in coffee drops; when one spills some coffee on the table and lets it dry, one can see that the middle of the drop is hollow and the coffee can be seen in the drop's margins. The

same effect happens to the spot on the chip. The spot dries before the proteins uniformly stick to the slide, and a ring is created.

2.4.4 Simplified chip – next generation

In order to deal with the new phenomena and check a few new aspects, we designed a newer and simpler chip. The problem of human error was not dealt with here, but was addressed by a totally different approach, using a new robot in the bioinformatics center. I will describe the use of the new robot later.

In the 3rd batch we tried to address a few questions/issues:

- 'Spot rings' –
    - o We tried 5 different protocols (Table 7– appendix) that dealt with chip quality. (Spot rings, accurate quantification, low background etc.)
    - o We checked the interference of IgM on IgG and vice versa. Our assumption was that a possible cause of the ring was the interference of one antibody with another, causing aggregation and creation of spot rings. Therefore we repeated the ratio experiment (IgM:IgG) with another control. We created two new different sets of ratios. The first was IgM antibody mixed with an inert protein (OVA), and the second of IgG mixed with OVA. The mixture was the same as the ratios of IgM and IgG. 0%-100% of OVA and 100%-0% of IgM/IgG separately.
- Determination of laser power. In order to get best results, the chip should be scanned at a specific laser power. The laser power is combined of two parameters: scanning power and gain power (see Sec.2.3.1). In this experiment we tried to find the best scanning and gain power for optimal results.
- Picture improvement. We tried to improve the quality of the picture produced by the scanner. The rationale behind this was that picture quality improvement will lead to better quantification and more accurate results. In order to deal with this aspect we scanned a few chips 10 times and analyzed the results, looking for picture improvement and for bleaching. Bleaching occurs when a fluorescent dye has been illuminated several times. Each illumination can damage the fluorescent dye, which will result in a decrease of the signal with increasing number of illuminations.

- QuantArray parameters. The QuantArray program gives the user several optional methods of spot quantification. The parameters we checked were a 'fix' method versus an 'adaptive' method. The aim was to find a method that is best for a most accurate quantification.

The 3rd chip design is presented in Figure 18.

Ratios	1	2	3	4	5	6	7	8	9	10	11
IgG	100	90	80	70	60	50	40	30	20	10	0
PRO1	0	10	20	30	40	50	60	70	80	90	100

Ratios	1	2	3	4	5	6	7	8	9	10	11
PRO2	100	90	80	70	60	50	40	30	20	10	0
IgM	0	10	20	30	40	50	60	70	80	90	100

Ratios	1	2	3	4	5	6	7	8	9	10	11
IgG	100	90	80	70	60	50	40	30	20	10	0
IgM	0	10	20	30	40	50	60	70	80	90	100

Figure 18: The chip design. The chip is built of 16 sub arrays, where each sub array is built of six rows and four columns- total of maximum 24 spots in each sub array. Each number/letter in the map represents one spot on the chip. The spots that are presented as a number from one (1) to eleven (11) are combination of IgM and/or IgG and/or OVA in different ratios. Spot number one (1) contains 100% IgG and spot number eleven (11) contains 100% IgM. Orange color represents combination of OVA with IgG. Blue - OVA with IgM. Red - IgG with IgM. The central subarray consists of a decreasing concentration of IgM and IgG separately, followed by positive controls (pCy3/pCy5) and negative controls (OVA).

The chip contains

- "Marginal" sub-arrays with gradients of IgM and IgG combined with other protein, in the range of zero to 100 in steps of 10.
- "Central" sub-arrays with gradients of IgM or IgG combined of four repeats of exponential dilutions (x, x/2, x/4, x/8, x/16, x/32).
- Positive and negative controls were placed on the chip as well.

We spotted eight chips in the current batch and followed five protocols (Table 7 of Appendix).

2.4.4.1 Scanning power

After the labeling step we went on to the scanning phase. In order to find the suitable scanning power, we took two chips from the batch and scanned them 10 times. In one chip we used gain-power of 80, and in the second chip we used gain-power of 95. The different scans varied in the laser-power, starting at 10 and up to 100 in steps of 10 (10:80, 20:80 … 100:80). Our first conclusion concerned the gain power. Gain-power of 95 was shown to be too sensitive and gave rise to a noisy background. There were also quite a few background pixels which were not uniform with all the others. Therefore we decided to scan with gain-power of 80, at which the background is much more uniform and the entire picture visibility seems to have no fluctuating scanning errors. The next step was adjusting the scanning power. The best scanning power should give maximum variation in the spot's intensity. If we scan the chip with power that is too high, all the spots will become saturated, which means – no valuable data. On the other hand, power that is too low will lead to very weak signal, and again, no valuable data. The optimal power will produce a minimal number of saturated pixels, while the rest of the spots should be in the range between zero and saturation. In other words, we are looking for the power which will give us maximum variation on a maximum intensity scale. In order to do so, we took the entire chip picture in TIF format after each scan, and plotted a histogram of the intensity versus the number of pixels (Figure 19 A and B). Unfortunately we got very poor results since the majority of the chip area is covered by the background; therefore causing the histogram to be dominated by low intensity pixels. In order to overcome this problem, we took the last scan (laser-power = 100), where all the real spots are saturated, with each pixel getting the highest intensity level ~65000. Using this intensity level we found the spots' exact location on the chip, which covered about 4% of the entire chip area. We now used their location on the picture and replotted the histograms, one for each laser-power (Figure 19 C,D). The laser power which gave the best results was determined as:

- Cy5 – 55 (between 50-60 which were examined) (Figure 19 C)
- Cy3 – 75 (between 70-80 which were examined) (Figure 19 D)

In those histograms the intensity variation was maximal with a relatively low amount of saturation. The saturation can be determined by the line that arises at the end of the histogram plot. With the chosen parameters there are 1000+ saturated pixels. One should

emphasize that one pixel ≠ one spot. Each spot is constructed of ~1500 pixels. At this point of the experiment we can start quantifying the spots and search for their meaning. Having determined these parameters we used them to scan the rest of the chips, a total of 16 different chips (8 slides*2 {top/bottom}). Most of the chips gave high quality signals in terms of low background and spots visibility (Figure 32 - appendix). The mean background for Cy3 was ~45 units, and the mean background for Cy5 was 60 units (on a scale of 0-65000 units).
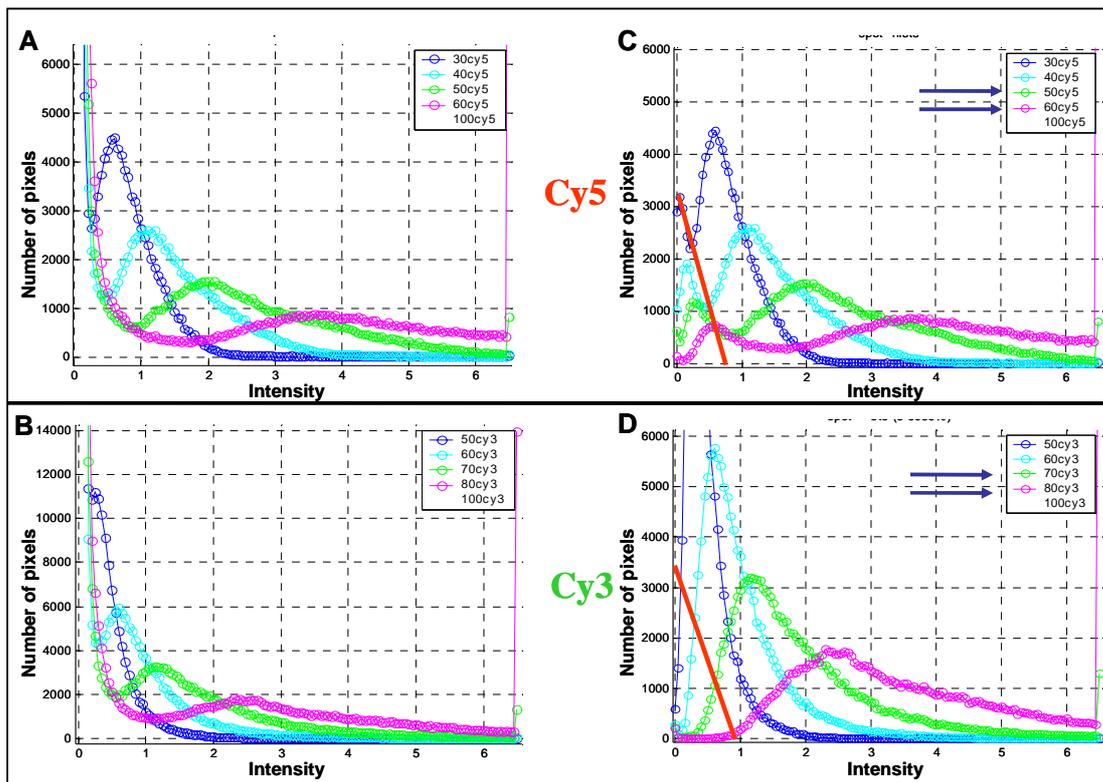
Figure 19: Histogram of scanning results. A,C refer to Cy5 dye with laser power of 30 up to 60. B,D refer to Cy3 dye with laser power of 50 up to 80. A,B present the histogram of the pixels intensity of the entire chip picture including the whole background area. Therefore most of the pixels tend to be between 0-10000. C,D present the histogram of the pixels intensity of the approximated spots on the chip. The blue arrows mark the best laser-power which gives maximum variation in spot intensity with minimal saturation. The intensity on the horizontal axes is from zero up to 65535. In Cy5 the finest laser power is 50-60. In Cy3 the finest laser power is 70-80

2.4.4.2 Multiple scanning and bleaching effect

Before quantifying the pictures and analyzing the numbers, we examined another important issue. It is known that scanning fluorescent dyes tend to have some fluctuations and errors. The scanning resolution is in pixels format where each pixel is represented by only one number. However, sometimes an error arises during the scanning process, or pixels are formed/reside exactly at the border of changing color. In that case, one scanning will not always give the best and most accurate picture. In order to overcome this phenomenon, one should repeatedly scan the chip and take the mean of all scans for each pixel. This idea seems simple, but one should take into account the bleaching effect. During repeated scanning, fluorescent dyes tend to become damaged, thereby producing a weaker signal. Thus we must measure the improvement of the picture visibility versus the bleaching effect which weakens and changes the signal. We scanned one chip ten times for Cy3 and ten times for Cy5. At the visual level, there was certain improvement. The picture was smoother and clearer (Figure 20). We developed the pictures and quantified them using QuantArray.
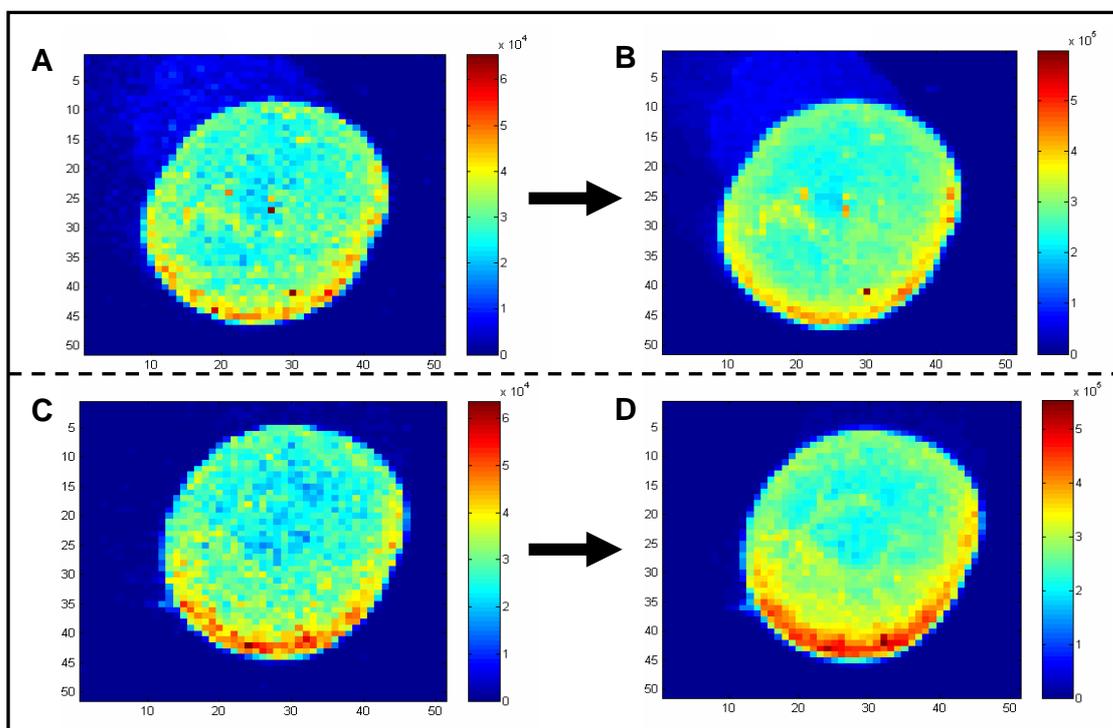
Figure 20: Multiple scanning. Results of first scan and sum of ten scans for two spots. **A,C** are spot's pictures after first scan. **C,D** are spot's pictures after summing up ten scans of the same chip. One can see the improvement of the picture quality and smoothness of the spots.

After quantification we examined a few spots which were not at a saturation point and also had a signal above minimum. We believed that in order to examine the bleaching effect, it is best to study a spot in the middle range of intensity, since there the signal is not saturated, and the bleaching can be seen in the middle intensity range. The middle intensity range is wide enough for detecting bleaching. We used only pixels where the spot could be easily distinguished from the background. We identified the spot's pixels from a square of 50x50 pixels, using the 'prcntil' function in matlab. For each spot we verified that all pixels related to it were found (Figure 33 - appendix). We plotted the pixels of the spot as we chose them and compared them to the original picture. After examining this effect both on Cy3 and Cy5, we concluded they each reacted differently to multiple scanning. In Cy5 we detected a 10%-15% decrease in the intensity after 10 scans at the same laser power (Figure 21). The intensity decrease was almost linear and the influence of each scan was more or less equal to the other scans. The behavior of Cy3 was different. The multiple-scanning effect was weaker; the total amount of bleaching was 2%-6%. The graph of the mean intensity wasn't linear, and was combined of decreasing and increasing intensity along the scannings. We were unable to determine the exact influence of multiple scanning on the Cy3 dye in terms of bleaching percent. The results weren't uniform over all the spots we checked, unlike Cy5, where most spots reacted similarly to multiple scanning. The quantifications of the chip after one scan and after ten scans did not differ. Multiple scans did not result in mean intensities that were more similar to expected values of the concentration gradient (Figure 22). Consequently we realized that although performing the multiple scanning will improve the visibility of the picture, the quantification aspect will remain unchanged. Therefore we concluded that it was both time and technology consuming to perform multiple scans and did not produce the expected results, so we scanned the rest of the chips only once, and quantified the single scan.
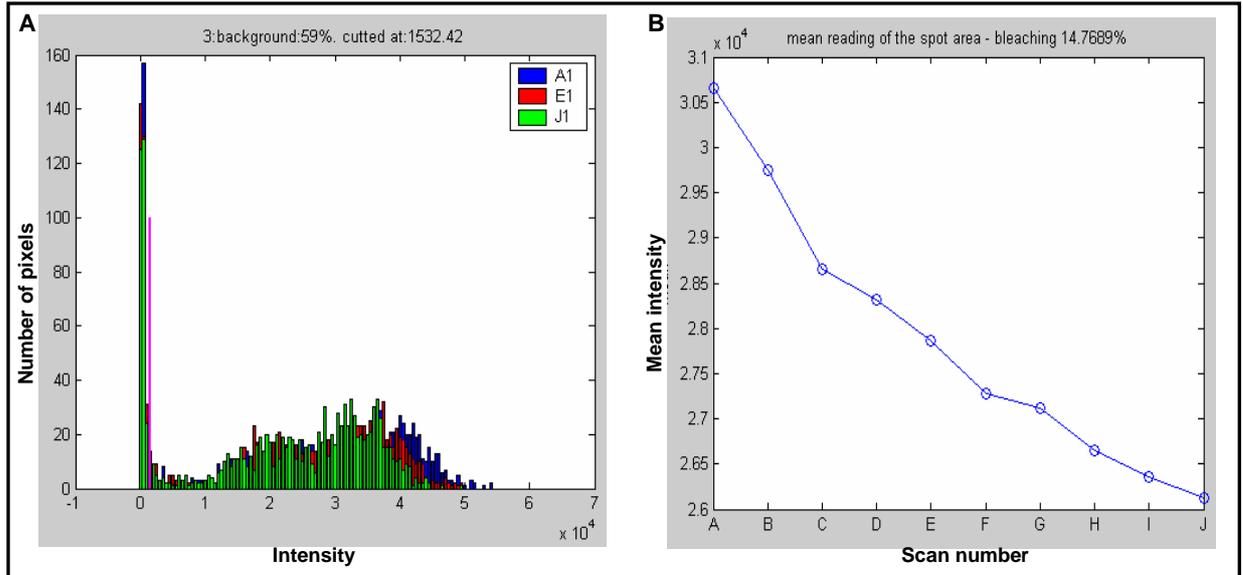
Figure 21: Bleaching effect of Cy5. **A-** Three histograms of the pixels intensity of a particular spot. The three histograms denote the intensity after the 1st scan (A1), after 5th scan (E1) and after the 10th scan (J1). The spot area was cut from the chip as a 50*50 square. The pink vertical line separates the background from the spot's pixels. The overall intensity of the spot's pixels tends to decrease after each scan. The bleaching effect can be seen very clearly by the shift of the histogram. In this particular spot the bleaching was 14.76% after ten scans. **B-** The mean spot intensity after each scan (only the spot's pixels, without the background).

2.4.4.3 Gradients' quantification

2.4.4.3.1 "Marginal" sub-arrays

The next step after having a matrix of numbers was to find a suitable way to test the chip. In order to do so we extracted each gradient separately. As explained above (Figure 18), we built the chip from groups of eleven spots, which should yield a gradient from 0 to 100 in steps of 10. We took the numbers, which represent the spot intensity, and manipulated them in a very simple way presented below, in order to fit them to a linear scale between 10 and 100.

$$\left(\left(\left(gradient\_vec - \min\right)\Big/\left(\max - \min\right)\right) \times 90\right) + 10$$

- gradient_vec : vector composed of ten numbers which represent the spot intensity.

- min : minimal number in the current gradient.

- max : maximal number in the current gradient.

We plotted the results and measured the correlation/distance from the optimal line. In this step we checked the improvement made due to multiple scanning. If multiple scanning greatly improves the spot quantification, we should have been able to detect it in this analysis. As mentioned above, multiple scanning unfortunately didn't improve the quantification dramatically. In Figure 22 we show examples taken from one chip, looking at the best gradients we could detect for IgG+Cy3. This was the best multiple scanning improvement we could detect. This improvement is not dramatic and will not have a great effect on the further analysis using clustering algorithms, since we will center and normalize the data causing small fluctuations and differences to be invisible and not critical for the analysis. Therefore, our opinion regarding the multiple scanning up to this point of the experiment was that it has a good potential of visualization improvement, but the quantification improvement does not justify the extra effort. Thus, we decided not to scan each chip ten times, but rather to improve the chip quality by other methods, such as refining the protocol, dealing with spot rings etc. Still, one can see (Figure 22 A,C) that we achieved great progress towards getting a quantitative chip. The gradient curve of IgG+Cy3 is very similar to the optimum curve. We identified the gradient effect in the chip with relatively high accuracy. On the other hand, the curves of IgM+Cy5 were less accurate (Figure 22 B,D). While decoding all the chips we made a few new observations. First, the gradient along the vertical axis of the chip gave poorer results compared to the gradient in the horizontal axis. (The figures present only the horizontal gradients). The quantified gradient in the horizontal axis was more accurate and resembled the expected linear gradient. In other words, the gradient along the rows in the chip gave reasonable results and was close to the optimal line, and the gradients in the columns didn't match the trend of the optimum line in over than 50% of the curves. As a result of this phenomenon we continued our analysis based on the horizontal gradients alone. This reduced the number of gradients from seven in each IgG/IgM to two IgM's and five relevant IgG's. We suspect that the poor vertical gradients could be due to the procedure the robot uses in order to print the spots. In this experiment, the robot printed the spots horizontally in each sub array, meaning that the consecutive spot was on the right. The time delay between two spots, one over the other, is much longer than spots next to each other. We think that some of the stock in the 384 plate evaporates during the printing. As a result of this, as printing proceeds, the concentration in the spot changes (usually increases). As result of this, spots in the horizontal order are more similar and accurate than in the vertical order. Thus detection of the gradient in the vertical direction is less accurate. Our conclusion

from this short discussion is that repetitions should be printed in the horizontal axis as opposed to the vertical axis, as long as the robot prints the spots in a horizontal order.
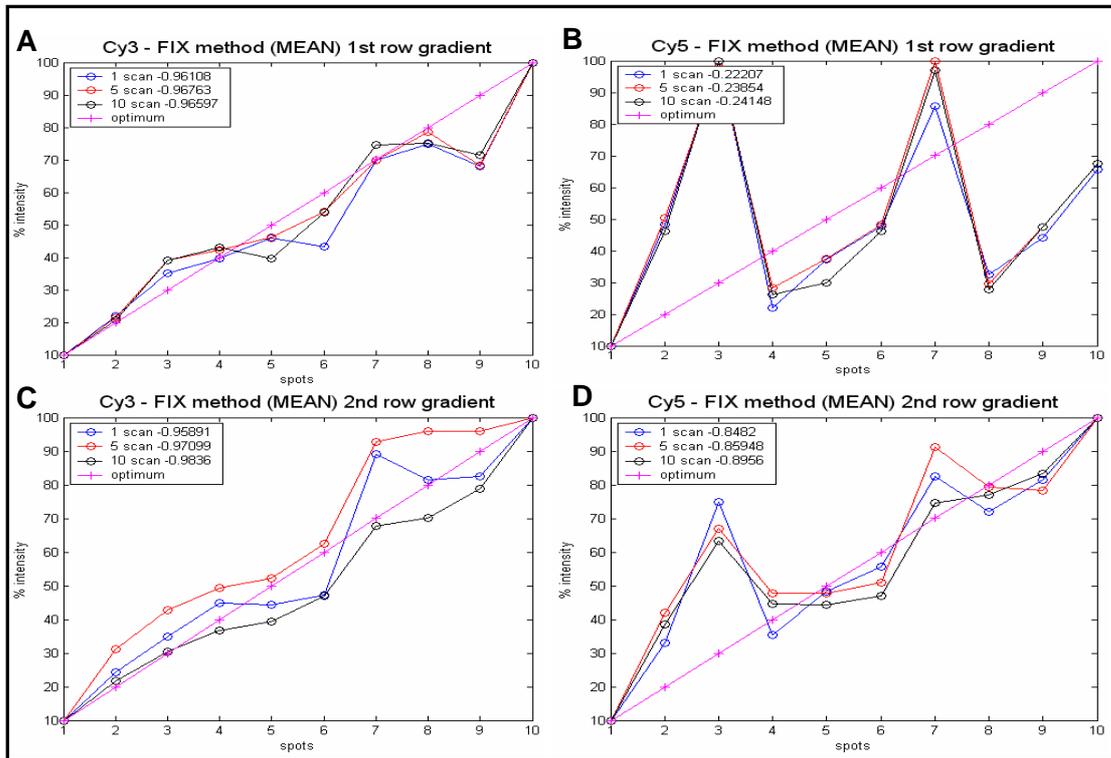
Figure 22: gradient quantification. Gradients of the top two rows of the chip number six are presented. Each gradient is combined of 10 numbers. We manipulated the numbers such they will be in the range of 10 to 100. A- 1st row in the chip, IgG+Cy3 gradient. B-1st row in the chip, IgM+Cy5 gradient. C-2nd row in the chip, IgG+Cy3 gradient. D-2nd row in the chip, IgM+Cy5 gradient. In each figure the optimum gradient is plotted in pink (linear line). And then three gradients are plotted. The blue one represents the gradient as it was detected after the 1st scan. The red represents the gradient after the 5th scan, and the black represents the gradient after 10 scans. A+C – represents high quality chip with gradients with high correlation (above 0.9) to the optimum. B+D represents low quality chip with gradients with relatively low and not stable correlation to the optimum (0.2 – 0.8).

2.4.4.3.2 "Central" sub-arrays

In addition to the "marginal" gradient analysis, we analyzed the gradients of the central sub-arrays. The central spots were combined of six dilutions composed of four spots each. We transformed the gradients such so that they decrease from 100, and compared it to the optimum/planned gradient. In order to minimize the variability we took the mean of the spots from each dilution. Once again we noticed that IgG+Cy3 gave much better results than IgM+Cy5 (Figure 23).
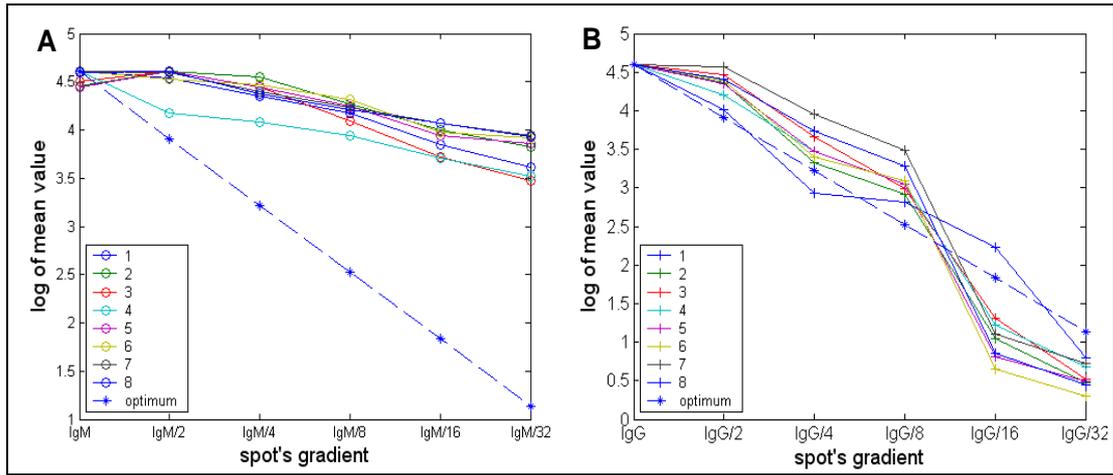
Figure 23: central exponential gradient. The gradient is composed of 6 exponential dilutions. In each dilution, four spots were printed on the chip. **A**- IgM gradient of 8 chips. Each line represents a chip, where each dilution corresponds to the mean of four repetitions in the current dilution. The linear line represents the optimum gradient that should appear. **B**- IgG gradient. The notation is the same as A.

In IgG+Cy3 the central gradients behave similarly to the optimum (Figure 23 B). Moreover, even by looking at the actual measured intensity (raw data) (Figure 34 B see Appendix) without performing any transformation on the data, one should clearly be able to detect the decay of the intensity along the gradient. In the entire set of chips we detected a very clear slope; the difference between them was the highest measurement of the intensity level. After taking log from the original measurement (Figure 23 B and Figure 34 D in Appendix) the curves from different chips tend to be more similar to one another while the overall gradient detection remains clear. Looking carefully at the slope, we can state that most of the decay in the intensity which was caused by the dilutions is detectable in the first four dilutions: x, x/2, x/4, and x/8. In the last two dilutions the signal decreases significantly from an actual value of three (3) to one (1), were it is no longer reliable. As opposed to the IgG data, the central gradient on the IgM+Cy5 data does not have the expected profile (Figure 23 A, Figure 34 A,C in Appendix). In some of the chips there is no decrease of the intensity with dilution, and the starting intensity point is fairly low (~ 20,000). This fact contributes to the poor marginal gradient in IgM.

2.4.4.4 Materials interference

We printed spots which are combination of IgM + IgG or IgM/IgG + inert protein. This experiment was meant to detect whether IgM "disturbs" IgG and vice versa, which could be a cause of the spot rings effect. We can say very clearly that this isn't the case here. We observed spot rings also in the spots where IgM/IgG were printed with the inert protein. Moreover, the best gradients were detected in the IgM+IgG mixture.

Another new fact we detected is that IgG+Cy3 gave better gradients than IgM+Cy5 (Figure 22 A,B). This fact was observed over the entire batch that we checked. We don't yet have any reasonable explanation for this fact.

We also paid attention to the fact that the fluorescent dye Cy3 is not specific only for IgG but also for IgM (for certain amounts - not yet quantified). Cy3 reacts as a polyclonal antibody, with different affinities to IgG and IgM, even though the producer claims that this secondary antibody is monoclonal to IgG only. Maybe this fact influences the gradients in IgM antibodies.

2.4.4.5 Human errors

In order to deal with the 'human errors problem' we mentioned above, we used a new robot of the TECAN Company called 'Genesis RSP'. It is composed of eight robotic pipettes with a sophisticated and accurate liquid control. One can determine with μl precision the amount of liquid to pump from a well, after programming the robot to find the liquid at a certain height; depending on the plate manufacturer (each manufacturer has its own small changes in the structure of the plate which influence the size and height of the well. In order to reach high accuracy, one should be aware to this fact while using the robot). One can also program the robot to pump in some air before and after pumping the liquid from the well. This prevents the robot from loosing any material while moving the pipettes, and helps to emit the total liquid that is placed in the pipette. The adsorption of the robot is minimal and constant for each program. Since it is much easier to fill manually a 96 wells plate according to a plan than to fill a complete 384-well plate, we programmed the robot to fill a 384-well plate from four 96-well plates.. We validated the robot's performance by measuring the liquid volume in a few wells. In each well we validated, we located 9μl out of the expected 10μl. This deviation was systematic and indicates that the error is constant and at approximately 1μl. Once all the wells in the plate are filled with the same amount of liquid the experiment is much more accurate,

especially while working with very small amounts of liquid. We believe this solution will reduce the human errors to minimum and solve the problems mentioned above.

2.4.4.6 Protocol determination

In this part we intended to determine a protocol which will give rise to the most precise results. The protocols differed from one another in the following aspects:

- Washing buffers. PBS / PBST-0.005% / PBST-0.05% (washing buffer with certain percent of detergent in it).
- Number of washings.
- Number of blockings.

The rationale of these three aspects was as follows:

- Different washing buffers clean the chip with varying strengths. Nonspecific bindings occur after each step in the chip production. The antibodies recognize the antigens and attach to them, but the contact is not always specific, and the interaction is relatively weak. We want to eliminate these interactions in order to detect the clearest and most specific signals. In order to do so, we washed the chip after each phase in the process. The washing material peels off molecules bound by the unspecific interactions, as well as left over material, and washes it off the chip. In order to control the washing strength we added detergent to the washing buffer. Caution was used regarding the amount of detergent, since adding too much detergent may wash out the signal. In order to find the most suitable washing buffer + detergent percent we tried several detergent concentrations.
- The number of washings in each "washing process" and their position/timing in the production process can affect the signal on the chip as well.
- In order to receive a clean signal with a light and uniform background there is a need to block all the free positions on the chip after the spotting step. The blocking can be accompanied either by washing or by another blocking step. We checked both alternatives.

The best protocol is the one that gives rise to results closest to those expected. We spotted a few marginal and central gradients on the chip, and calculated the distance of the results from the optimum. Previously we mentioned that we omitted the entire column's gradients since we believed that there is a problem not relevant to the experiment, so we were left with five IgG marginal gradients and two IgM marginal

gradients. For each spot of the gradient we calculated the distance from the optimum and summed up all the distances for all the gradients. We believe that the chip with the lowest distance from the optimum is the best and the protocol which was used to produce this chip should be adopted as the best. It is important to remember that this is only a two-layer chip, in which the binding of the antibodies is specific and consequently the signal obtained is pretty obvious and strong, whereas in a three-layer chip the washing buffer may have to be weaker in order to sustain a strong signal. After measuring the distances of IgG gradients and IgM gradients we discovered that the differences between the chips weren't pronounced. Therefore we decided to take only the four best gradients of IgG and one best gradient of IgM. The results were approximately the same. Chip number two gave the worse results in both measures. Chips four and six gave the best results (Figure 24 B), but this should be interpreted with care since the differences between the chips are small and can easily be affected by a few out-lyer spots.
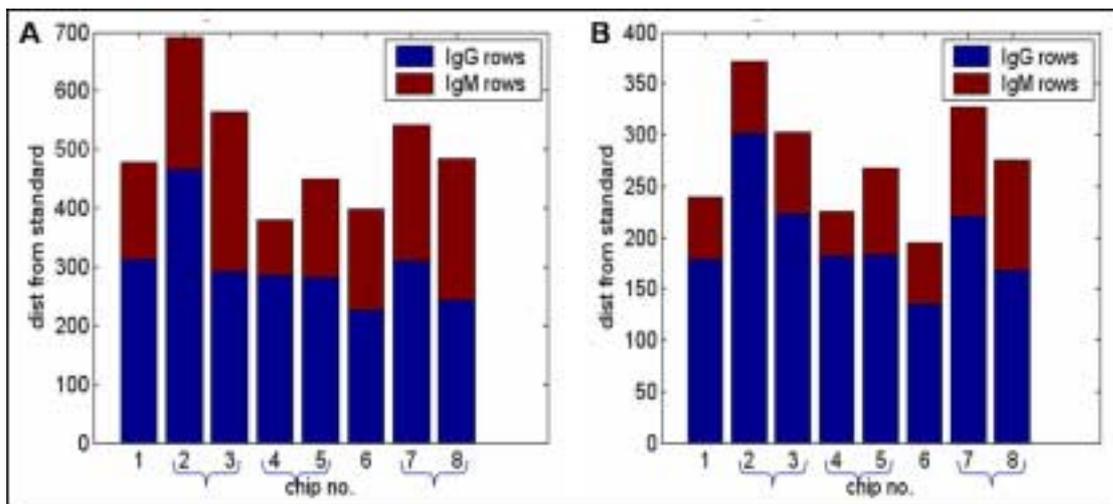
Figure 24: Protocol evaluation. A- Five IgG gradients and two IgM gradients were taken. For each spot in each gradient the distance from the optimal measure was collected and summarized in this histogram. B- Same as A, for the best four IgG gradients and only one IgM gradient. The distance was calculated versus the optimum 10-100 in steps of 10 units. The distance units are arbitrary.

To summarize this part, we can conclude that the choice of protocol doesn't greatly affect the results, which means that the parameters we varied were not crucial for good chip production.

2.4.4.7 QuantArray parameters

The relevant parameters are the spot's characteristics and the quantification method. We used the following spots' characteristics:

- Spot size = ~400
- bkgInner = ~600
- bkgOuter = ~800

The antigen spots seem to be bigger than regular DNA spots in cDNA chips; therefore we gave higher parameters than those used for cDNA chips. Another reason for this was that we have spot rings which probably affect the quantification. If the ring is not contained by the spot size, quantification will not reflect the actual intensity in the spot (even though we suspect that the spot rings damage the quantification accuracy).

The second aspect was the quantification method. There are two relevant options; the "fix" method and the "adaptive" method. In theory the adaptive method can distinguish less well between signal and background than the fix method. The "fix" method builds two histograms. One histogram is built of the spots' pixels and the other is built of the background pixels. The user must input the upper and lower boundaries of the histograms, where after the number within those boundaries is calculated. The method "checks" the separation between the two histograms and calculated the mean/median of the spot histogram after ignoring the boundaries. We used the following boundaries; signal: 30-95, background 5-80.

The "adaptive" method selects a few pixels that represent the signal and a few pixels which represent the background. The selection is based on the intensity of the pixels, the logic being that the signal from the background is always higher. Thus, a negative signal (no intensity) will result as wrong quantification as well as a spot with a minor scratch on it. (The full algorithm of this method is beyond the scope of this work.)

In reality both methods gave similar results. There were some spots where the "adaptive" version gave more accurate results and vice versa. Since each chip contains quite a few kinds of spots with different backgrounds, the total quantification is approximately the same. In our opinion there is no significant difference between the methods for the results of current study. Nevertheless, since the "fix" method was found to be more stable using chips with different intensities of background noise this method alone will be used.

2.4.4.8 Reproducibility

Another important aspect of a high quality chip is its reproducibility. We designed a new chip in order to study this aspect, as well as some new aspects in the production process. As we mentioned above, one of difficulties we faced, were the spot rings. We had some new ideas that we thought would help to attack this problem. One of them concerned the printing buffer. The antigens that were spotted up to this point in the experiment, where diluted in PBS, while the 'Arrayit' company claimed that they have a special printing buffer solution, which improves the spots' uniformity on the chip. Furthermore, the spot rings problem can be handled with a technical procedure such as heating the chip after spotting. In cDNA chips, this method of steaming is well known, and improves the spots' uniformity. In order to research these aspects we designed the following chip (Figure 25).
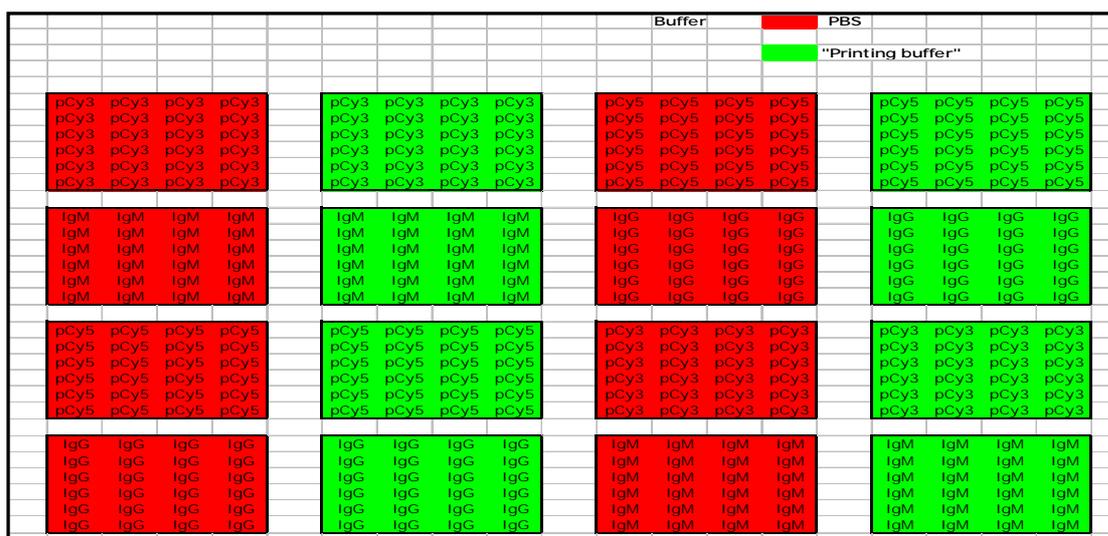
Figure 25: Chip design. Each subarray contains 24 identical spots. The antibodies are taken from different wells for each spot. pCy3 and pCy5 are the fluorescent labeled secondary antibodies spotted directly on the chip. IgM and IgG are purified antibodies which are detectable after incubation with the secondary antibodies. Red color – PBS as printing buffer. Green color – printing buffer bought from "Arrayit" company.

Two chips were scanned after spotting and were then treated using the same protocol that was described in the previous batches (1,2). The rest of the chips were passed over steam for different time duration, ranging from 30 seconds up to one minute. Unfortunately, the chips that were treated for more than 30 seconds were corrupted and therefore not analyzed. All the chips were scanned with laser power and gain of 55:80 for Cy5 and 75:80 for Cy3.

The first image that popped up from the pictures of chips 1,2 after spotting was the high uniformity of the spots that were printed using the special printing buffer. However, this phenomenon disappeared after a set of incubations with secondary antibodies and a washing procedure. The second finding was the variation of the intensity between spots in the subarray which was up to two fold change (Figure 26). This finding is extremely important for the accuracy and reproducibility of the next generation of chips. Chips 1,2 gave rise to less variation between repetitions than other chips. This evidence, combined with the fact that passing a chip over steam is a very devastating action, which doesn't eliminate the spot rings, lead us to the conclusion that the antigen chip shouldn't be treated with steam.
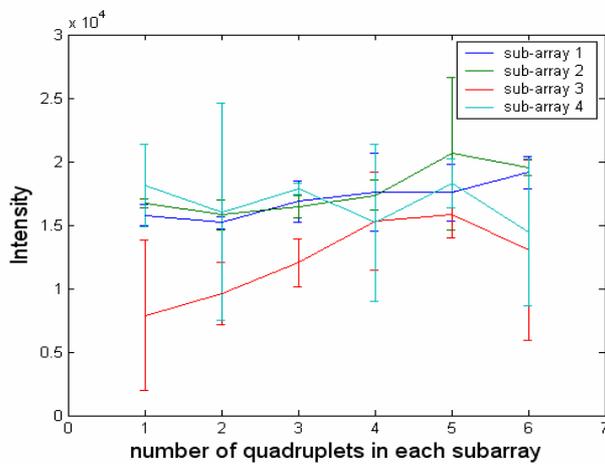
Figure 26: quantified results of IgM spots in chip number one. Each subarray is presented in a different line. The subarray was divided into six quadruples. Each quadruplet was composed of 4 spots on the same row and presented in the figure with mean and standard deviation. The fold change in each subarry is up to two fold.

2.4.5 Simplified Chip - Conclusions

The experiments done with the simplified chip assisted us to understand better the mechanism and production process of a quantified chip. We learnt to produce a clean and specific chip, how to scan one, and to quantify it properly. The clean and uniform background, quantified gradients and reasonable fold change has lead us to conclude that this chip is ready to be tested on biological data once again. Hence, we should design a new experiment with the involvement of a chip with three layers; antigen/peptides, serum with antibodies and labeled secondary antibodies. After succeeding in such an experiment and revealing biological meaning to the results, such a chip will be considered as credible and relevant for biological questions and biological utilizations of the kind which were described in detail in the introduction (2.1,2.2).

2.5 '3-layer' chip experiment – CAD disease

After developing the methodology of the antigen chip, we turned to testing it on a biological system. This experiment included the three-layer chip consisting of antigens, antibodies from patient's/mice serum and labeled secondary antibodies.

2.5.1 Biological background and experimental system

Type I diabetes is a metabolic disorder caused by the autoimmune destruction of the insulin-producing β cells of the pancreas [13]. Nonobese diabetic (NOD) mice can develop two types of autoimmune diabetes, spontaneous diabetes and cyclophosphamide-accelerated diabetes (CAD) [13, 37]. Spontaneous diabetes appears in NOD mice by 11 weeks of age, ~80% occurrence in females and ~40% in males [38]. Intraperitoneal injection of high dose of cyclophosphamide (CY) consistently induces the onset of diabetes in male and female NOD mice at an age when spontaneous diabetes rarely occurs [38]. This type of diabetes appears by six weeks of age.

In this research 30 NOD mice were included. Ten NOD mice served as control and did not get any treatment. The other 20 mice were injected twice with 200 mg/kg of CY at the age of four and five weeks. Blood samples were taken from each mouse at four time points. The first time point was at age of four weeks before the first injection. The second was at five weeks of age. The next were seven and nine weeks of age (Figure 27). At the end of the experiment the mice were tested for the development of cyclophosphamide-accelerated diabetes (CAD). Of the treated group 50% developed diabetes by nine weeks of age, the remaining 50% remained disease-free. None of the mice in the control group developed diabetes by nine weeks of age.
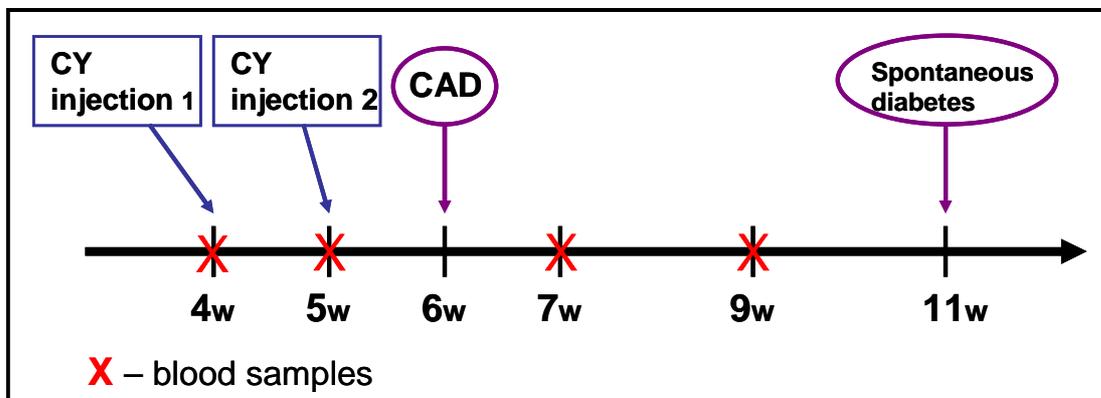
Figure 27: experimental scheme.

2.5.2 Goals

The aim of this study was to reveal potential antigens related to the difference between CAD and spontaneous diabetes in NOD mice. The experiments sub-goals were to; 1) Validate the intactness of the three-layer antigen chip. 2) Reveal the difference in antibody repertoire between healthy and diseased mice receiving cyclophosphamide treatment. 3) Try to predict which mouse will develop CAD based on the antibody repertoire at four weeks of age.

2.5.3 Experiment Design

We designed a new version of the chip. The chip was built of 16 sub-arrays. Each sub-array contained 49 spots in seven rows and seven columns. The first column in each sub-array served as a control. The control was composed of two spots of IgM+IgG mixture, one spot of PBS used as negative control, two spots of anti-IgG and two spots of anti-IgM. The anti-IgG/IgM interacts with the constant region of IgM or IgG antibodies taken from the serum. In addition we printed a set of IgM+IgG spots in different dilutions, similar to the gradients we printed in the two-layer chip (see Figure 37). The rest of the sub-array was utilized for antigens and peptides. More than 200 different antigens were printed on the chip, each in triplicate. The current chip using this design can hold up to 224 triplets of antigens (for chip design map and antigen list see Figure 37, Figure 38 and Table **8** in the appendix). Most of the antigens tested were already used in the previous ELISA experiments in the lab. Some of the antigens were known to be related to autoimmune diseases. We also printed a peptide library of HSP60 which is extensively studied in Prof. Cohen's laboratory.

We printed eight slides, such that each slide is composed of 'top' and 'bottom' parts. We used blood taken from the group of eight to ten mice injected with cyclophosphamide. The blood was drawn at nine weeks of age, and mixed in two pools of serum (dilution 1/5) – depending on whether the mice developed CAD or not. The serum was then incubated with the antigens on the chip, each pool on two slides (top, bottom), a total of four chips for each group. The chips were scanned in two different conditions, since the positive control on the chip (such as anti-IgM and anti-IgG) gave rise to a very high response compared to the antigens on the chip. The concentration of the control spots was too high and tended to saturate in relatively low laser power. On the other hand, laser power which was optimal for the control spots produced an extremely low antigen signal.

As a result, we scanned each chip twice for each fluorescent dye (Cy5 & Cy3). Cy5 (IgM) was scanned with laser and gain power of 55:80 for the control spots and 95:80 for the antigens (signal) spots. Cy3 (IgG) was scanned with 80:80 for control and 95:80 for signal. The laser power and gain were determined after conducting the same analysis described in section 2.4.4.1- Scanning power. The images were produced in the scanning phase were quantified using the QuantArray program. We executed the fix method with the following parameters:

1. Spots' parameters: Spot size = 410, bkgInner = 635, bkgOuter = 860.
2. Boundaries: Signal: 40-95, Background: 5-85.

2.5.4 Results and Conclusions

The 3-layer chip turned out to be a relevant, reliable and meaningful tool. Firstly, the chip had a very clean and uniform background compared to the first 3-layer chip experiment (Figure 28).
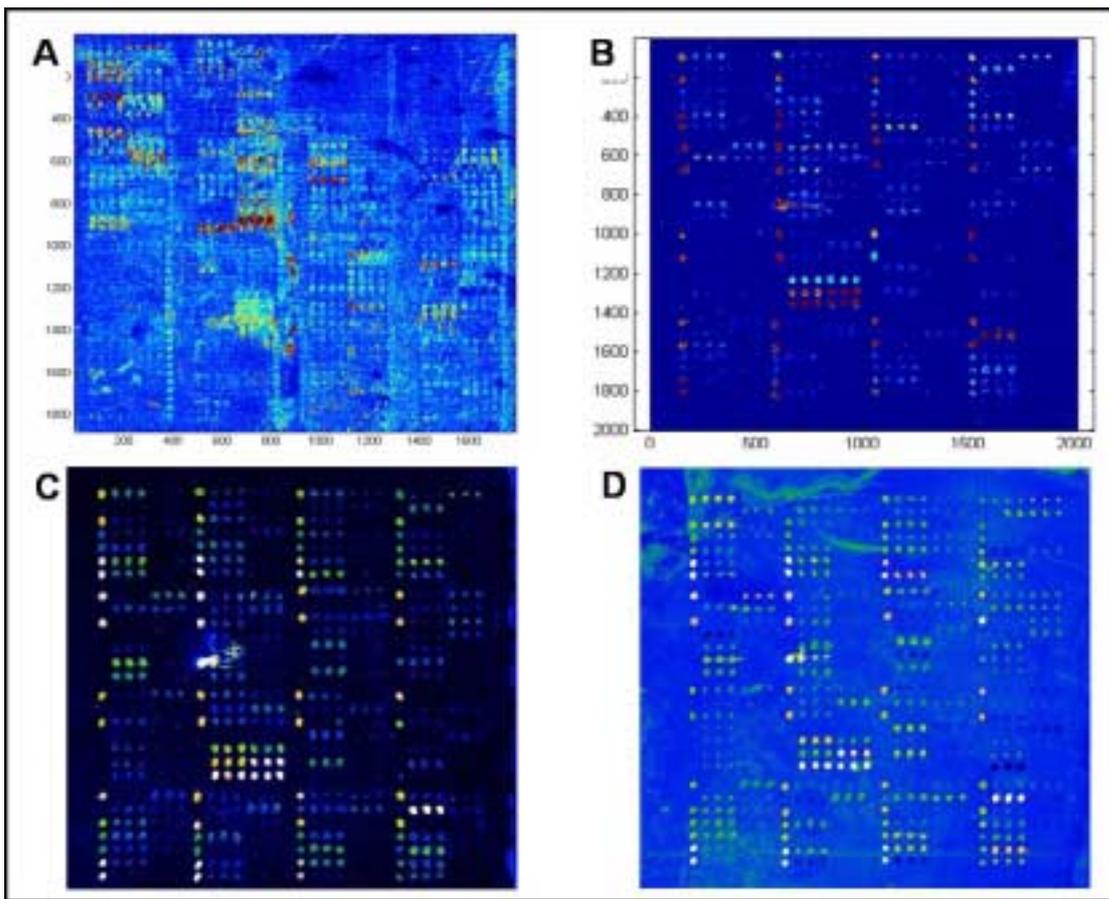
Figure 28: 3-layer chip examples. This figure summarizes the major progress of the antigen chip production. A- First experiment (2.3.3). The image was produced using the matlab program. B-D - CAD disease chip examples (2.5). B- Chip images using matlab program. C,D - the original images in BMP format, produced by the scanner apparatus, using different laser power. The background is low and uniform. There are almost no spills and merges.

Global normalization was used for each chip in such a way that the intensities on different chips were comparable to each other. We used two normalization methods: 1) Divide by mean/median. 2) Take log and subtract the mean after the log. The normalization step was done separately for Cy3 and Cy5 scanning. The chips were separated into two groups representing the CAD mice and the non CAD mice (healthy). Each group included 12 measurements for each antigen, resulting in four chips with three replicates per antigen. The groups were reduced after applying the ignore filter which was generated while quantifying the chips. The ignore filter determines the reliability of the spot. The reliability of each spot was determined visibly, by human eye. Any spot that we suspected to be false because of scratches, dirt on the chip or any other reason was marked as unreliable. Before performing any statistical test, data verification was done to ensure that there were at least six measurements per antigen for each group (after applying the ignore filter). The majority of antigens were represented by 10-12 repetitions (Figure 29) which gives results that are statistically significant.
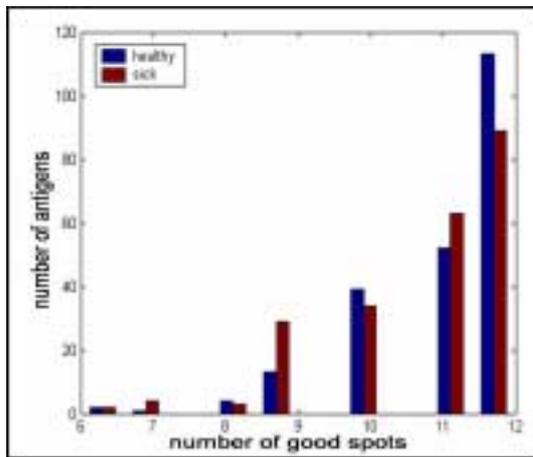
Figure 29: Histogram of number of spots per antigen in each group after applying the manual ignore filter. The maximum number of spots per antigen is 12 per group. More than 95% of the antigens have 9-12 replicates in both groups.

We looked for antigens which can help define two separate groups, healthy mice versus CAD mice. In order to do so we performed supervised analysis using two statistical tests, rank-sum test (see 1.4.8) and t-test. The tests were done on all antigens, one by one, and P-values were collected. In order to overcome the multiplicity problem (see 1.4.7), we used the FDR method with a threshold of 0.05 false discovery rate. In addition, we conducted the Bonferroni test. Results can be seen in Table 5.

Table 5: - Number of antigens separating CAD mice from healthy mice using various normalizations and statistical tests

CAD vs. Healthy Scaling Method + **test**	IgG	IgM
Divide by mean **Rank-sum**	53/8	31/11
Log + subtract Mean **Rank-sum**	39/8	28/11
Divide by Median **Rank-sum**	25/10	26/10
Divide by Median **t-test**	30/11	33/16

Separating antigens between CAD mice and healthy mice.
Example: 30/11
- 30 - Antigens separate CAD from healthy as result of ranksum/t-test and FDR of 0.05.
- 11 - Antigens separate CAD from healthy as result of ranksum/t-test and Bonferroni.

We used the rank-sum test in order to find antigens that discriminate between the two groups without assuming any normality of the data. Rank-sum will result in a low P-value even when two groups are distinguished by very small differences between them, whereas t-test will not yield a low P-value in these cases. Furthermore, rank-sum test will give the same P-value for antigens that share a common number of points in the data separating the two groups, even if the actual intensity significantly differs between those antigens. T-test on the other hand, will refer to the actual intensity of the two antigens and will give different weights, resulting in different P-values. However t-test assumes normality of the data, which we have no reason to assume. Therefore we combined the result of both statistical tests. Moreover, we calculated the fold change between CAD mice and healthy mice and set the threshold for separating antigens to be at least two-fold change. The combination of rank-sum, t-test, FDR and fold change should give us the most accurate and strict results. We validated the results by looking at the reactivity pattern of the candidate antigens. This validation was done manually one by one and cleaned the data further. We found few antigens that gave rise to low p-values and high fold change due to

one or more out-layers in the data. Additionally, we threw out antigens that gave rise to non consistent patterns along the 12 repeats, such that the fold change within the repeats was higher than two fold. After carrying out all the filtering we were left with seven separating antigens that reacted with IgG antibodies (Figure 30) and 16 separating antigens that reacted with IgM antibodies (Figure 31). These antigens gave rise to a very clear separation between the CAD and healthy mice following all our criteria. However, we believe and know that there are more antigens that separate these two groups of mice but did not pass our strict criteria. As the technique continues to develop and the measurements become more accurate, we will be able to cut back some of the criteria.
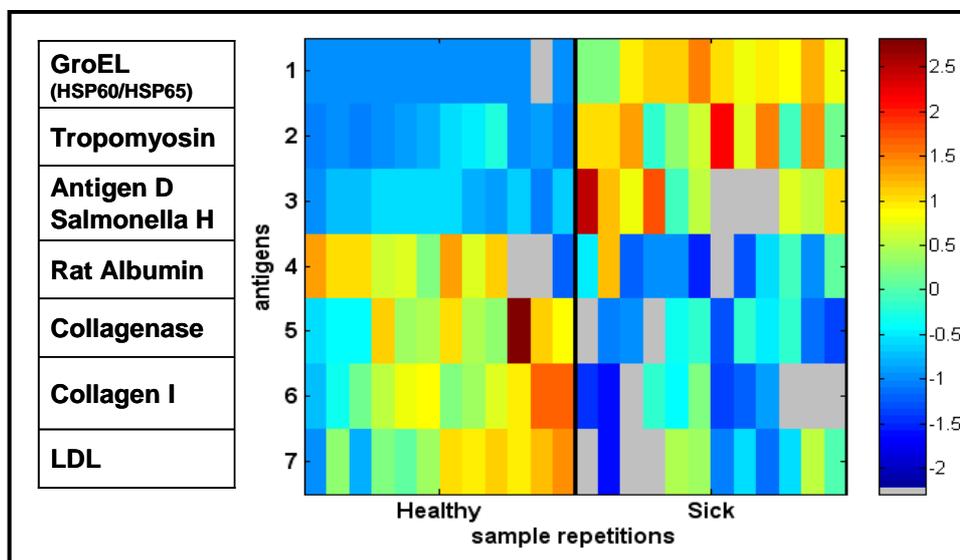
Figure 30: Seven Antigens which differentiate CAD and healthy mice. For each antigen the results of 12 repeats, for both pools of sera, are presented. The detection was of **IgG** antibodies. These seven antigens passed the following tests: ranksum test, t-test + FDR (5%) and two fold-change. Each row/antigen was centered (mean = 0) and normalized (std = 1). Gray squares represent missing values in the data (NaN).
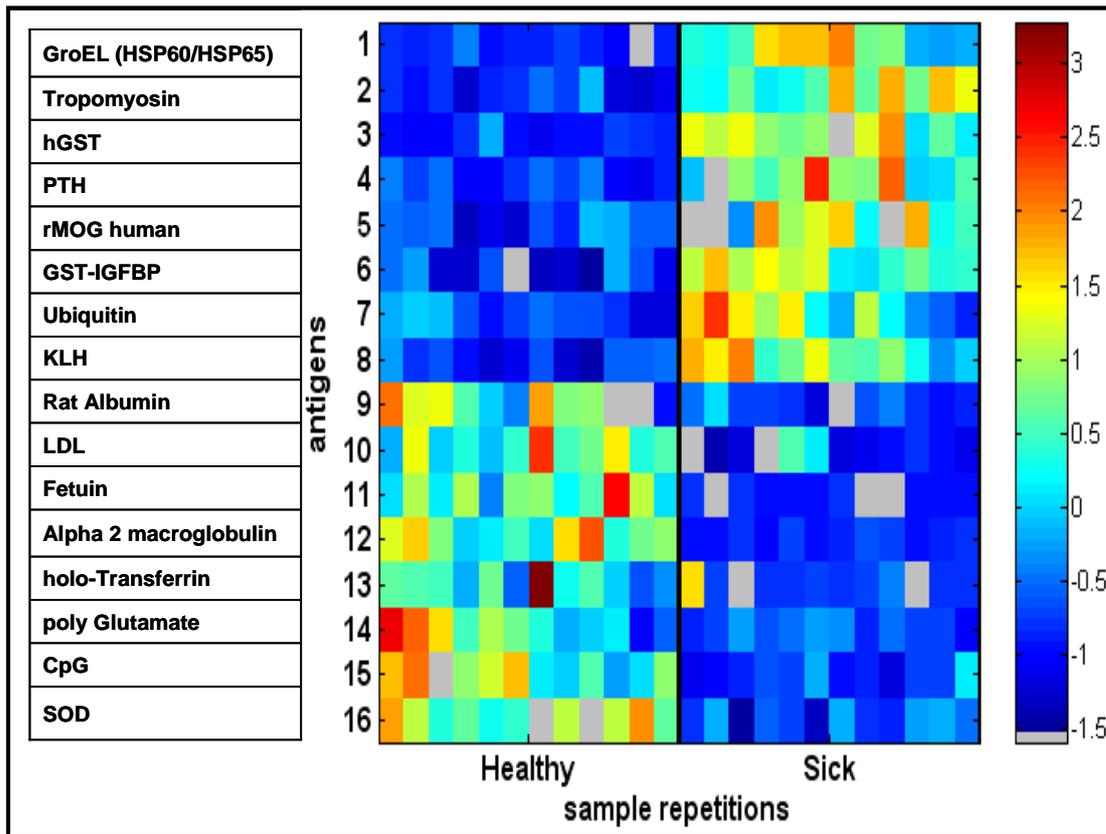
Figure 31: 16 Antigens which differentiate CAD and healthy mice. For each antigen the results of 12 repeats, for both pools of sera, are presented. The detection was of **IgM** antibodies. These 16 antigens passed the following tests: ranksum test, t-test + FDR and two fold-change. Each row/antigen was centered (mean = 0) and normalized (std = 1). Gray squares represent missing values in the data (NaN).

One should remember that the deviation in the reactivity between the two groups shown in Figure 30 and Figure 31 can be caused by an individual mouse within the group or by the entire group of mice. Due to the fact that we are using a pool of serum we are unable to ascertain the origin of the phenomena. Therefore, it is difficult to draw unambiguous biological conclusions upon this experiment. This experiment was meant to check the intactness of the 3-layer chip and to authenticate that there are candidate antigens on the chip that may be related to the difference between CAD mice and non CAD mice. Another issue concerns the control spots in this experiment which gave a very high signal at a relatively low laser power whereas the antigens hardly showed any signal using the same laser power, as we mentioned above. In order to scan the chip only once and to be able to normalize all the chips according to the control spots such as anti-IgM/IgG, we should decrease the concentration of the control spots from 1 mg/ml to ½ mg/ml or even ¼ mg/ml. This way the intensity of the control spots will be in the middle intensity range. Another technical issue is the development step; we realized that the chip

gets cleaner once all the washing steps are done in a specific order as follows: The washing starts with five minutes of PBS followed by five minutes of PBST 0.005% and ends with another five minutes of PBS.

2.5.5 Future plans

In the near future we are planning to print 40 slides with ~300 antigens; these slides will be used to further investigate the CAD disease. We used 40 samples of blood taken from 20 mice. Each mouse received cyclophosphamide and was sampled before and after treatment. Ten of the mice developed CAD. In the coming experiment, each mouse serum will be hybridized on an independent chip. This will help us reveal biological differences in the antibodies pattern between CAD mice and non CAD mice. The spread of the four groups of serum with different biological outcome can provide us with a powerful tool to characterize CAD disease, which can bring us one step forward in gaining new insight concerning autoimmune diabetes. Moreover, we will try to predict the disease by looking at the antibodies' pattern prior to the CY treatment.

Another project in the near future will deal with epitope mapping, in which we will try to map different antibodies specificities. It is known that each antibody binds to a specific epitope of a larger antigen. We will attempt to characterize these epitopes in order to get a closer look at the antibody specificity that is related to the autoimmune diseases. For example we will try to map HSP60 epitopes in diabetes patients using the antigen chip.

2.6 Summary

In this work we developed a new technique called the 'antigen chip'. We tried to deal with different technical difficulties and biological issues that were related to the production process. We constantly improved the chip production such that it finally became a quantitative chip which gave preliminary results regarding biological phenomena (CAD). The current data was analyzed using supervised methods. In the future, we will also apply unsupervised, methods while using large scale data, in order to reveal the maximal information hidden in the data. We have shown that the antigen chip can be used as a good tool for immunological studies of antibody-antigen interactions. We hope that the quantitative antigen chip will become another powerful tool in every scientist's suitcase. Every question that deals with antibody-antigen interaction could be solved using this tool. Moreover, this tool can be easily adapted to any other proteomic field that deals with protein interactions.

References

1.	Robinson, W.H., et al., *Millennium Award. Proteomics for the development of DNA tolerizing vaccines to treat autoimmune disease.* Clin Immunol, 2002. **103**(1): p. 7-12.
2.	Verge, C.F., et al., *Combined use of autoantibodies (IA-2 autoantibody, GAD autoantibody, insulin autoantibody, cytoplasmic islet cell antibodies) in type 1 diabetes: Combinatorial Islet Autoantibody Workshop.* Diabetes, 1998. **47**(12): p. 1857-66.
3.	Quintana, F.J., et al., *Cluster analysis of human autoantibody reactivities in health and in type 1 diabetes mellitus: a bio-informatic approach to immune complexity.* J Autoimmun, 2003. **21**(1): p. 65-75.
4.	Robinson, B., *http://www.stanford.edu/group/utzlab/autoantibodies.htm*. 2002.
5.	Lacroix-Desmazes, S., et al., *Self-reactive antibodies (natural autoantibodies) in healthy individuals.* J Immunol Methods, 1998. **216**(1-2): p. 117-37.
6.	Shoenfeld, Y. and D.A. Isenberg, *Natural Autoantibodies - Their physiological role and regulatory significance*. 1993.
7.	Stahl, D., et al., *Analysis of human self-reactive antibody repertoires by quantitative immunoblotting.* J Immunol Methods, 2000. **240**(1-2): p. 1-14.
8.	Cohen, I.R., *Tending adam's garden - evolving the cognitive immune self*. ACADEMIC PRESS ed. 2000.
9.	Cohen, I.R., *Discrimination and dialogue in the immune system.* Semin Immunol, 2000. **12**(3): p. 215-9; discussion 257-344.
10.	Diabetes, *http://clinical.diabetesjournals.org/*. 2001.
11.	Gale, E.A., *The discovery of type 1 diabetes.* Diabetes, 2001. **50**(2): p. 217-26.
12.	Winer, S., et al., *Autoimmune islet destruction in spontaneous type 1 diabetes is not beta-cell exclusive.* Nat Med, 2003. **9**(2): p. 198-205.
13.	Quintana, F.J., P. Carmi, and I.R. Cohen, *DNA vaccination with heat shock protein 60 inhibits cyclophosphamide-accelerated diabetes.* J Immunol, 2002. **169**(10): p. 6030-5.
14.	Lock, C., et al., *Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis.* Nat Med, 2002. **8**(5): p. 500-8.
15.	Zhou, X., et al., *Systemic sclerosis (scleroderma): specific autoantigen genes are selectively overexpressed in scleroderma fibroblasts.* J Immunol, 2001. **167**(12): p. 7126-33.
16.	MS, *http://www.nationalmssociety.org/*. 2002.
17.	Lock, C.B. and R.A. Heller, *Gene microarray analysis of multiple sclerosis lesions.* Trends Mol Med, 2003. **9**(12): p. 535-41.
18.	Walsmith, J. and R. Roubenoff, *Cachexia in rheumatoid arthritis.* Int J Cardiol, 2002. **85**(1): p. 89-99.
19.	Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* J Roy. Stat. Soc., 1995. **B**(57): p. 289-300.
20.	Ravid, O., *Gene expression analysis - Thesis*, in *Physics of complex systems*. 2003, Weizmann institute of science: Rehovot.
21.	Robinson, W.H., L. Steinman, and P.J. Utz, *Proteomics technologies for the study of autoimmune disease.* Arthritis Rheum, 2002. **46**(4): p. 885-93.

22.   Zhu, H. and M. Snyder, *Protein chip technology.* Curr Opin Chem Biol, 2003.
       7(1): p. 55-63.

23.   Futcher, B., et al., *A sampling of the yeast proteome.* Mol Cell Biol, 1999. **19**(11):
       p. 7357-68.

24.   Robinson, W.H., et al., *Autoantigen microarrays for multiplex characterization of
       autoantibody responses.* Nat Med, 2002. **8**(3): p. 295-301.

25.   Abbas, A.K., K.M. Murphy, and A. Sher, *Functional diversity of helper T
       lymphocytes.* Nature, 1996. **383**(6603): p. 787-93.

26.   Joos, T.O., et al., *A microarray enzyme-linked immunosorbent assay for
       autoimmune diagnostics.* Electrophoresis, 2000. **21**(13): p. 2641-50.

27.   Robinson, W.H., et al., *Protein microarrays guide tolerizing DNA vaccine
       treatment of autoimmune encephalomyelitis.* Nat Biotechnol, 2003. **21**(9): p. 1033-
       1039.

28.   Zhu, H., M. Bilgin, and M. Snyder, *Proteomics.* Annu Rev Biochem, 2003. **72**: p.
       783-812.

29.   Pandey, A. and M. Mann, *Proteomics to study genes and genomes.* Nature, 2000.
       405(6788): p. 837-46.

30.   Chien, C.T., et al., *The two-hybrid system: a method to identify and clone genes
       for proteins that interact with a protein of interest.* Proc Natl Acad Sci U S A,
       1991. **88**(21): p. 9578-82.

31.   Wulfkuhle, J.D., L.A. Liotta, and E.F. Petricoin, *Proteomic applications for the
       early detection of cancer.* Nat Rev Cancer, 2003. **3**(4): p. 267-75.

32.   MacBeath, G. and S.L. Schreiber, *Printing proteins as microarrays for high-
       throughput function determination.* Science, 2000. **289**(5485): p. 1760-3.

33.   Haab, B.B., M.J. Dunham, and P.O. Brown, *Protein microarrays for highly
       parallel detection and quantitation of specific proteins and antibodies in complex
       solutions.* Genome Biol, 2001. **2**(2): p. RESEARCH0004.

34.   Schaeferling, M., et al., *Application of self-assembly techniques in the design of
       biocompatible protein microarray surfaces.* Electrophoresis, 2002. **23**(18): p.
       3097-105.

35.   Angenendt, P., et al., *Next generation of protein microarray support materials:
       evaluation for protein and antibody microarray applications.* J Chromatogr A,
       2003. **1009**(1-2): p. 97-104.

36.   ArrayIt-TeleChem International, I., *http://www.arrayit.com*. 2003.

37.   Elias, D., *The NOD mouse: A model for autoimmune insulin-dependent diabetes*,
       in *Autoimmune Disease Models*, I.R. Cohen and A. Miller, Editors. 1994,
       Academic Press, Inc.: New York. p. 147-161.

38.   Yasunami, R. and J.F. Bach, *Anti-suppressor effect of cyclophosphamide on the
       development of spontaneous diabetes in NOD mice.* Eur J Immunol, 1988. **18**(3):
       p. 481-4.

Appendix

Table 6: Antigens/peptide list used in the human diseases experiment.

blank	Rat MBP	Peroxidase	HSP70/7
RNAse	Brain Extract	Thyroxinase	HSP70/11
Collagenase	HSP70/13	huMOG	HSP70/10
Hemoglobin	Collagen VI	p277	p180/8
Acidphosph.	Collagen X	EC27	Heparin
LPS	Collagen I	RatAlbumin	ssDNA
Fetuin	Acid.Collag.IX	OVA	dsDNA
Tuberculin	Laminin	huAlbumin	GpC
hAchR	Fibronectin	Thyroglobulin	CpC
Cytochrom C	Myosin	C1	met.BSA
C 1Q	Tropomiosin	hIgG	Protamin.sulf
Rat MOG 35	Vimentin	hIgM	KLH
Transfferin	Myoglobin	FAB200	Pepstatin
huMOG 94	Enolase	HSP70/12	pAsp
Cart.Extr.	Spectrin	HSP70/8	pLys
Fibrinogen	GAD	Actin	pArg
Fibrin	HSP60	Anexin	pGlu
N 4	Collagen mono	FP3	Cholesterol
Insulin	Histone II	HSP70/9	PEA
C9	Catalase	HSP70/14	cardiolipin

These 80 antigens/peptides were measured by ELISA. These antigens include proteins, peptides, nucleotides and phospholipids

Table 7 - protocols

Protocols / Steps	None	PBS	Blocking * 2	PBST - 0.005 %	PBST - 0.05 %
Wash 0	None	2 x 5' PBS	2 x 5' PBS	2 x 5' PBST 0.005 %	2 x 5' PBST 0.05 %
Blocking 1 O.N. in 4ºC.	BSA 3 % in PBS	BSA 3 % in PBS	BSA 3 % in PBS	BSA 3 % in PBS	BSA 3 % in PBS
Wash A	2 x 5' PBS	2 x 5' PBS	2 x 5' PBS	2 x 5' PBST 0.005 %	2 x 5' PBST 0.05 %
Blocking 2	None	None	BSA 3 % in PBS	None	None
Wash B	None	None	2 x 5' PBS	None	None
Hybridization (*)	Serum 1/50 in BSA 1 %, 2 hr 37ºC				
Wash C (*)	2 x 5' PBS	2 x 5' PBS	2 x 5' PBS	2 x 5' PBST 0.005 %	2 x 5' PBST 0.05 %
Detection	αIgG-Cy3 + αIgM-Cy5 1/500 in BSA 1 %, 30' 37ºC.				
Wash D	3 x 5' PBS	3 x 5' PBS	3 x 5' PBS	3 x 5' PBS	3 x 5' PBS
Scan (**)	Cy5 – 55:80 Cy3 – 75:80				
Number of chips	1 (1)	2 (2,3)	2 (4,5)	1 (6)	2 (7,8)

This table presents the 5 protocols which were used in the 3rd batch experiment of the simplified antigen chip. The protocol is combined of few steps. In all the protocols we followed all the steps, unless it is mentioned as 'None' in the table. The protocols differ also in the materials that were used for the washing process, in term of the detergent concentration. (notation: 3x5' => 3 times for 5 minutes each)

(*) - done only in the 3-layer chip

(**) – In each experiment may be different

A general protocol is listed bellow:

- **Spotting**
- **Wash 0**   **:** 2 x 5' PBS gentle rocking (washing-buffer accordingly to the relevant protocol). The PBS should be filtered ahead.
- Centrifuging   : 5' x 1000 rpm (drying the chip)
- **Blocking**   : BSA 3% in PBS, for Over-Night in 4ºC**.**
- **Wash A**   **:** 3 x 5' PBS gentle rocking (washing-buffer accordingly to the relevant protocol).
- Centrifuging   : 5' x 1000 rpm (drying the chip)
- **Hybridization (*):** Serum 1/50 in BSA 1 %, 2 hr 37ºC. 40μl/cover slide.
- **Wash B**   **:** 3 x 5' PBS gentle rocking (washing-buffer accordingly to the relevant protocol).
- Centrifuging   : 5' x 1000 rpm (drying the chip)
- **Detection**   **:** αIgM-Cy3 + αIgM-Cy5 1/500 in BSA 1 %, 30' 37ºC.
- **Wash C**   **:** 3 x 5' PBS gentle rocking (washing-buffer accordingly to the relevant protocol).
- Centrifuging   : 5' x 1000 rpm (drying the chip)
- **Scan the chip.**

(*) **-**   done only in the 3-layer chip.

(**) – The protocol was generally improved such that each washing process will be 3*5'. Starting with PBS, then PBST-0.005% and last with PBS again.
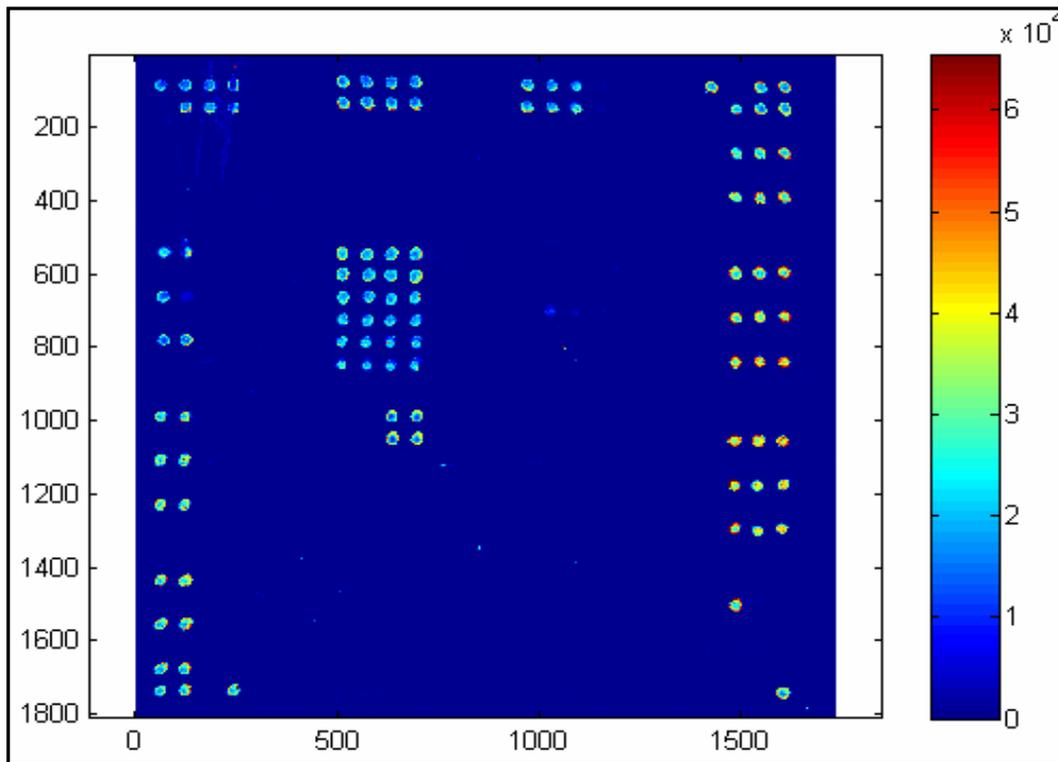
Figure 32: Chip picture from 3rd batch. Chip number seven, Cy5 scanning. Scanning parameters 50:80.
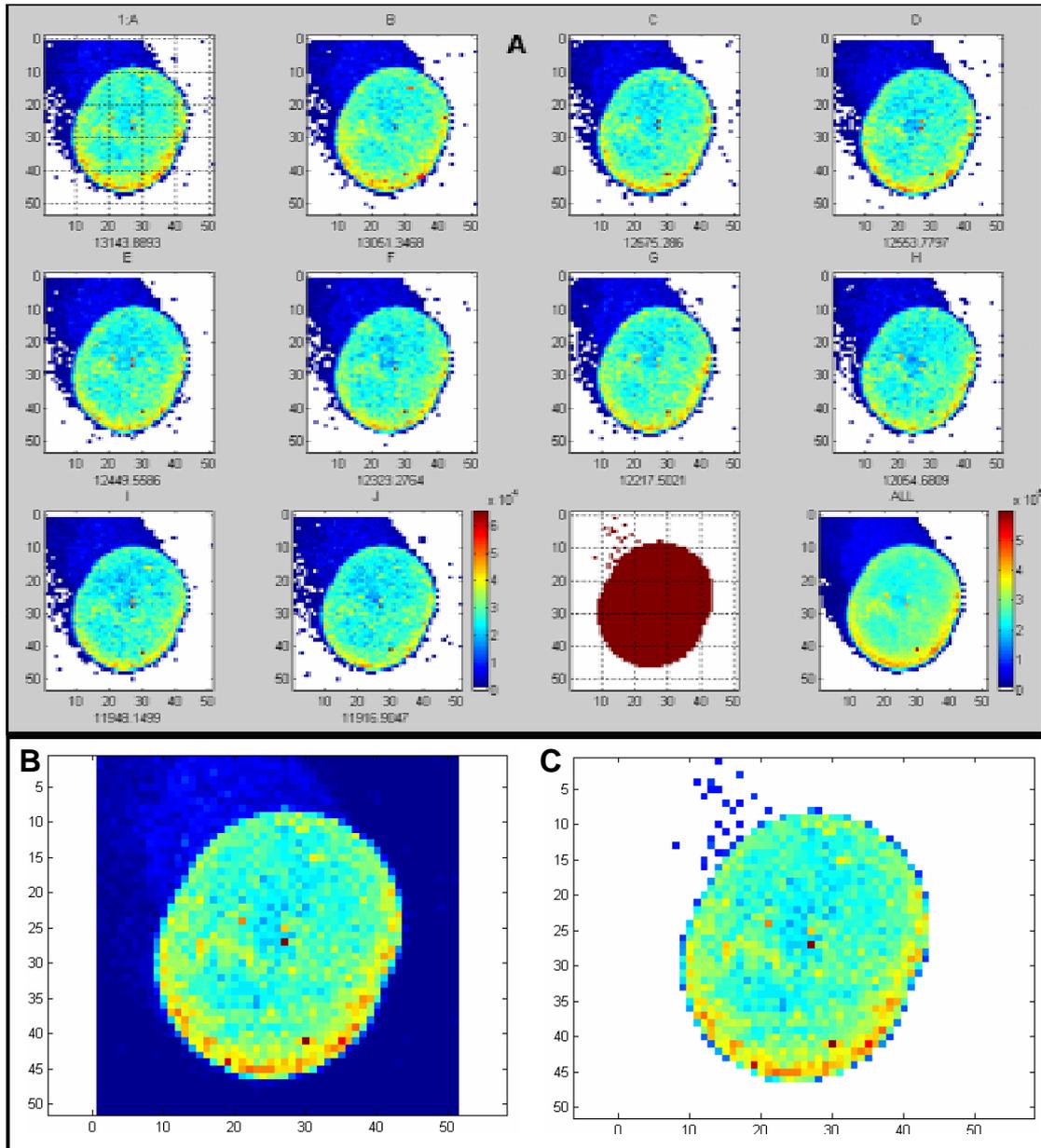
Figure 33: Spot calculation. A – Ten scanning of the same spot (from A-J), and the spot's pixels as were identified by our program. On the right-bottom figure we can see the sum of all the previous scans. The colorbar is matlab regular colorbar from zero up to ~65000. The number zero is presented as white color instead of dark blue. B- The original spot after one scan, C- the calculated spot location – output of our program. The calculation is very accurate which adds up to 25 pixels from the background to the spot.
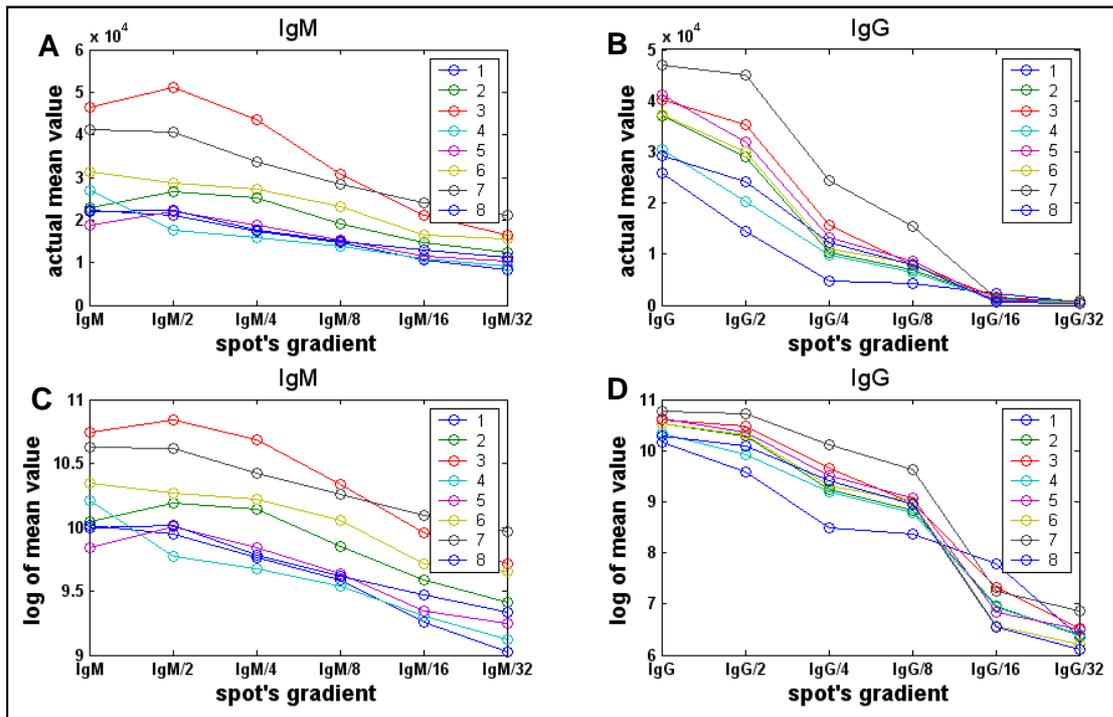
Figure 34: central gradient – original data. In each subfigure one line represents one chip. Each line combined of six measurements of the different dilutions in the central gradient. Each measurement out of the six is combined of the mean of four spots in the same dilution on the chip. A,B show the value of the intensity as it was obtained from quantarray (which meant to be exponential in optimum). C,D present the same numbers as in A,B after natural log (without any other manipulation). IgM gradient presented in A,C which exhibit very poor slope compared to the IgG slopes in B,D.
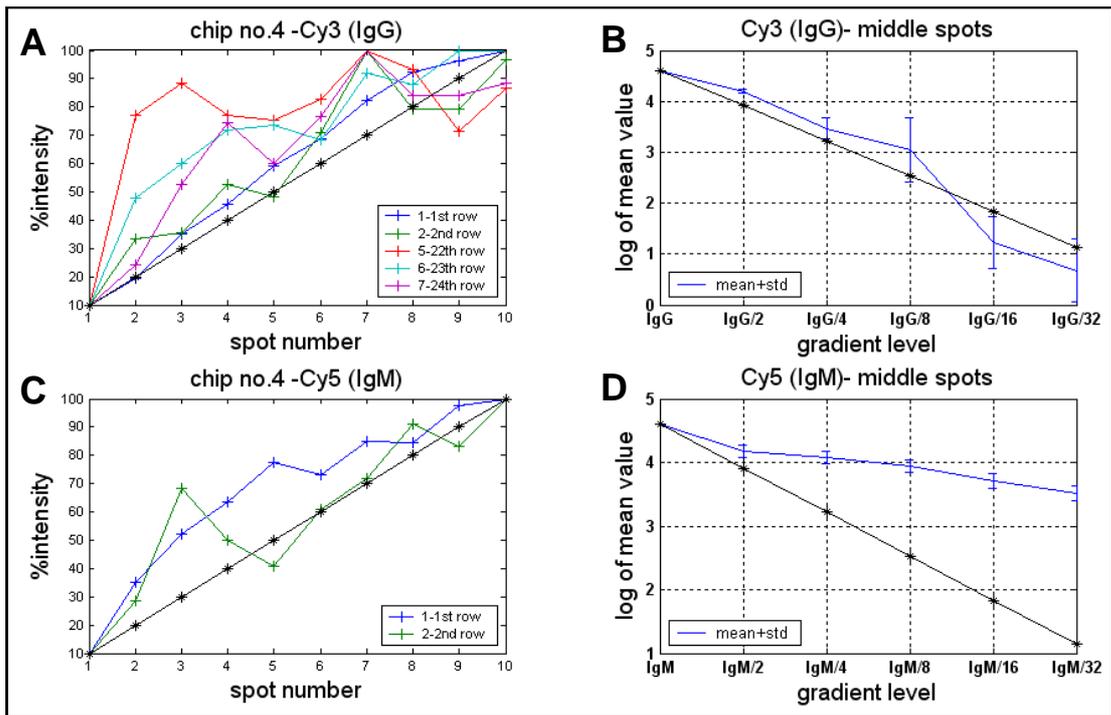
Figure 35: Quantified gradients of chip number four. The picture summarizes all the horizontal gradients of one chip. A- Five IgG gradients after mathematical manipulation that fits them to the range of 10-100. B- The central gradient of IgG. The gradient is composed of the mean of four spots in each dilution. The mean and std are presented in the blue line as log values. The gradient was manipulated such as the start point will be 100. C- Two IgM gradients which were manipulated as the gradients in A. D- IgM central gradient – same as B. in all subfigure the black line represent the optimum gradient. This chip gave rise to the best gradients among all chips.
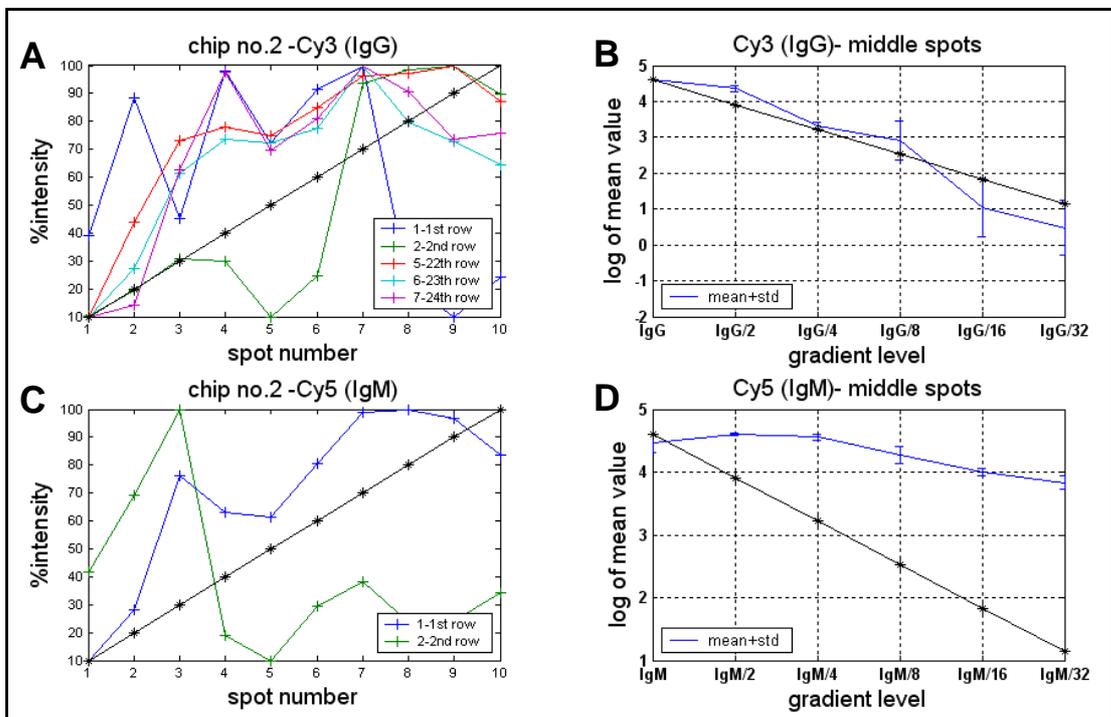
Figure 36: Quantified gradients of chip number two. The details are the same as **Figure 35**. This chip gave rise to the worse gradients among other chips.
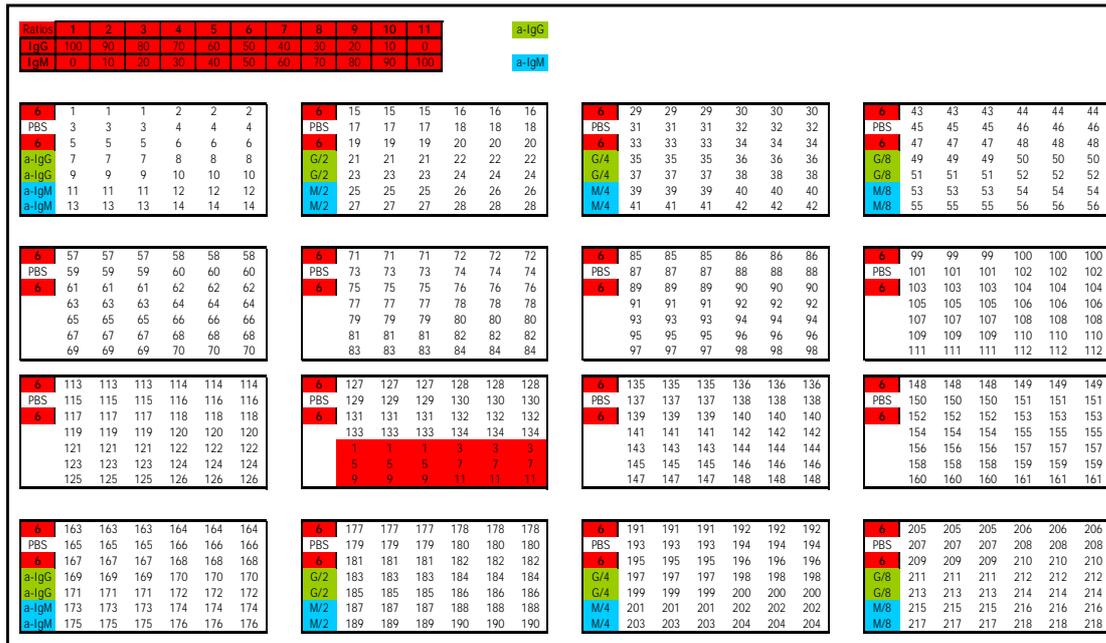
Figure 37: Three-layer chip design. The first column of each sub-array is control. The control is built of two spots with equal concentration of IgM+IgG, one spot of PBS as negative control, two spots of anti-IgG and two spots of anti-IgM. The anti-IgG/IgM were printed in different concentration in the sub-arrays. Each number represents a potential antigen/peptide on the chip. In the middle there is a gradient of IgM and IgG as a positive control.

Figure 38: Three-layer chip design with Antigen names. The map is the same as in **Figure 37**. each triplet is replaced with the fitting antigen. (St# equal to # in the previous figure).

Table 8 – Antigen/peptide list

C-protein	20514	Defensin	PTH
C1q-II	MM2	Insulin	Vitronectin
Insulin B	Tropomyosin	Peroxidase	Acid Phosphatase
SSDNA	Tyrocalcitonin	Fibrinogen	Thyroglobulin
Collagen IX	Caspase 8	Gelsolin	Big gastrin I
C1Q	OVA	Glucagon	GHRF
Fibronectin	Catalase	HDL	Plasmin
C1complement	Fibrin	poly Lysine	poly Aspartate
Substance P	Annexin V	Hemagglutin Yeda	Factor II
Neurotensin	Alpha 2 macroglobulin	Cartilage Extract	human IgM
poly Glutamate	Vasopresin	OVA	Rat Albumin
MOG 35-55	poly Arginine	Chorionic Gonadotrophin	gamma MSH
Fetuin	MBPrat	Factor II	alpha MSH
holo-Transferrin	Hemoglobin	ANP	Thrombin
Neuropep Y	human Albumin	KLH	Protamine Sulfate
Insulin A	beta MSH	Endoproteinase Clu-c Staphylococcus	Human Proinsulin
SOD	Troponin	Rat IgG	Tyrosinase
N4	CRF	met BSA	IL4
Antigen D Salmonella H	Heparin	Myeloperoxidase	VIP
human IgG	C-peptide	hGST	Ubiquitin
BSA	LDL	Vimentin	Caspase 3
Beta Amyloid	Enolase	Collagen I	Collagen X
Actin	Collagenase	ribonuclease A	Collagenase/II
Acetyl Choline Esterase	Pepstatin	myosin	Aldolase
Collagen VI	tubulin	IFNgamma	poly G
Endoproteinase Clu-c Staphylococcus/ II	IL12	polyT	GpC
vimentin	polyA	GST-CASPR2	GST-MUPP
IL2	GST-INAPC	MMP3	MMP1
polyC	MMP9	60/9	60/10
HSP90	60/8	60/32	60/33
Galactosyltransferase	60/30	ATTA	SSDNA
60/6	TAAT	MOBP 78-89	MOG 35-55 rat
60/29	MOG 94-111	Factor X	Oxytocin
CpG	Endothelin 1	60/16	60/18
GST-IGFBP	60/14	60/36	60/37
Endothelin 2	60/35	C9	EC27
60/12	Enolase	rMOG human	rMBP human
60/34	rMOG mouse	60/1	mtP3
DSDNA	p277	60/23	60/24
Brain Extract	60/22	LPS	p12/60
MISTAKEN	HSP70	GroEL	PPD
60/19	alpha Cristallin	beta Cristallin	PLP
HSP60	60/2	60/4	60/5
Histone IIA	60/26	60/27	60/28
gpMBP	Laminin	spectrin	
Kinetesin	Somatostatin	Beta2 microglobulin	
60/25	BNP	LHRH	

Antigen and peptides list that where printed on the three-layer chip.

שבב אנטיגנים: פיתוח ואנליזה
יישום למחלות אוטואימוניות

גד אליצור

תזה לשם קבלת התואר מוסמך במדעים
מוגש למועצה המדעית של
מכון ויצמן למדע

בהדרכת

פרופסור איתן דומאני
ופרופסור ירון כהן

שבט תשס"ד
ינואר 2004