UNIVERZA V LJUBLJANI

FAKULTETA ZA FARMACIJO

MARKO VIŠIČ

MAGISTRSKA NALOGA

ENOVITI MAGISTRSKI ŠTUDIJ FARMACIJE

LJUBLJANA, 2014

UNIVERZA V LJUBLJANI

FAKULTETA ZA FARMACIJO

MARKO VIŠIČ

NOVA METODA ZA DOLOČEVANJE NAVZKRIŽNE REAKTIVNOSTI PROTITELES NA OSNOVI VEZAVNIH RAVNOVESIJ

A NEW METHOD FOR DETERMINING THE CROSS-REACTIVITY OF ANTIBODIES BASED ON BINDING EQUILIBRIUM DATA

LJUBLJANA, 2014

This Master's thesis was done at Uniklinik Köln, Germany, department of Experimentelle Anästhesiologie und Schmerzforschung Klinik für Anästhesiologie und Operative Intensivmedizin, under supervision of Prof. Dr. Tim Hucho.

Mathematical modeling and computer simulation part of Master's thesis was done at Institut für Automatisierungstechnik (IFAT) Magdeburg, Germany, under supervision of Jr. Prof. Dr. Steffen Waldherr.

ACKNOWLEDGEMENT

I would like to extend my special appreciation to Prof. Dr. Tim Hucho, who has truthfully an accomplished and inspiring mentor. Herewith I would like to thank you for all your hard work and putting your trust in me. I am truly honored to have worked under your tactful supervision. I would also like to thank to the research team at Uniklinik Köln from whose extensive knowledge and friendship I have benefited more than words can say.

My deepest appreciation goes to Jr. Prof. Dr. Steffen Waldher for maneuvering me through the mathematical and programming part of this thesis.

I would particularly like to express my gratitude to Prof. Dr. Iztok Grabnar, who helped me bring this thesis to a successful close. Quickly gasping the concept of the thesis, he provided a very on point insight on the paper.

Last of all, thanks to Johann Wolfgang von Goethe. For all words he wrote, but especially for these ones:

"Knowing is not enough; we must apply.
Willing is not enough; we must do."

Statement

I hereby declare that this Master's thesis was done by me under supervision of Prof. Dr. Izidor Grabnar, Prof. Dr. Tim Hucho and Jr. Prof. Dr. Steffen Waldherr.

Student's signature: _____

Head of Committee: Prof. Dr. Borut Štrukelj

Member of Committee: Asist. Dr. Stane Pajk

Table of Contents

List of Tables

List of Figures

ABSTRACT

Antibody-based techniques are central to life science research and therapy. Specificity is the fundamental requirement of antibodies to be used in such techniques. One of the central pitfalls of immunoassay techniques is the assurance of antibody specificity vital for immunostaining procedures, which are heavily employed by all different kinds of cell biologists. Current methods for specificity evaluation such as the use of knock-out mice and preabsorption with immunogen are expensive and not applicable to all antibodies. To address the antibody specificity assurance and detect cross-reactivity by obtaining affinity distribution of antibodies a new method based on antibody binding was devised and implemented in a standard laboratory use of immunofluorescence staining.

A mathematical model based on pseudo-first order antibody-antigen binding was created. Computer simulations were used to validate the model and pinpoint its parameters in detail. Model was further tested and evaluated in a wet lab environment. A dilution series of monoclonal and polyclonal antibodies were used for immunofluorescence staining of HeLa cell line and corresponding fluorescent signals in equilibrium were measured. By applying the model the number of binding sites was obtained from antibody affinity distribution. Moreover, the models ability to detect antibody cross-reactivity was examined.

Modeling results have shown that it is possible to differentiate between antibody binding sites with a 10 fold difference in binding affinity even at 10% measurement uncertainty. Wet lab experiments confirmed that pseudo-first order model validly describes antibody-antigen binding interaction. Using the model, we were able to accurately recognize the number of binding sites for monoclonal antibodies, while this approach was found to be ineffective for polyclonal antibodies. Furthermore, it was shown that major antibody cross-reactivity is detected with developed method.

EXTENDED ABSTRACT IN SLOVENE LANGUAGE (RAZŠIRJEN POVZETEK V SLOVENSKEM JEZIKU)

Tehnike, ki vključujejo uporabo protiteles, so osrednjega pomena v sodobnih biološko-baziranih znanostih in terapiji. Specifičnost je ključna in nujna lastnost protiteles, ki se uporabljajo pri izvedbi imunoloških tehnik. Ena izmed poglavitnih pomanjkljivosti imunoloških tehnik, ki jih celični biologi uporabljajo širom sveta, je pomankanje tehnik za zagotavljanje specifičnosti protiteles oziroma za odkrivanje navzkrižne reaktivnosti. V celični biologiji se uporabljajo predvsem za imunološko obarvanje in posledično identificiranje tkiv in celic. Obstoječe metode za določevanje specifičnosti protiteles, kot recimo uporaba mišk z izbitimi geni ali predadsorpcija protiteles z imunogenom, so zelo draga in ne morejo biti uporabljena na vseh vrstah protiteles.

V uvodu je predstavljena standarna proizvodnja monoklonskih in poliklonskih protiteles. Temu sledi struktura protiteles in razumevanje območij na protitelesu, ki so odgovorna za specifično vezavo antigenov. Izpostavljene so sile, ki tvorijo vez med vezavnim mestom na protitelesu z epitopom na antigen. Za uporabo protiteles je ključno, da je vez specifična – torej da se veže zgolj na določen antigen. V nasprotnem primeru, kjer se protitelo nespecifično veže tudi na druge tarče, govorimo o navzkrižni reaktivnosti. Predstavljene so imunološke tehnike, kot so fluorescenčna mikroskopija in radijska imunologija, kjer je odsotnost navzkrižne reaktivnosti uporabljenih protiteles bistvenega pomena.

Cilj moje magistrske naloge je izdelava nove metode za zagotavljanje specifičnosti protiteles. Metoda temelji na razumevanju dinamike vezave protiteles na antigene, na podlagi te je bil zgrajen model vezave. Le-ta je zmožen napovedati afinitetno distribucijo protitelesa na osnovi eksperimentalnih meritev pridobljenih s standardno laboratorijsko uporabo imunofluorescenčnega barvanja. Iz znane afinitetne distribucije je moč razbrati, ali je protitelo navzkrižno reaktivno v uporabljenem eksperimentalnem sistemu.

Magistrska naloga je razdeljena na dva dela. Prvi del obsega matematično modeliranje in računalniške simulacije vezave protiteles. Primarni cilj tega dela je določitev parametrov, ki vplivajo na zmožnost modela, da iz simuliranih podatkov vezave protiteles pravilno napove afinitetno distribucijo protitelesa. Poleg tega smo ovrednotili tudi, kako parametri vplivajo na

zmožnost zaznave navzkrižne reaktivnosti, če je le-ta prisotna. Namen drugega dela je validacija nastalega modela na podlagi eksperimentalnih meritev vezave protiteles z antigeni. Zmožnost modela, da zazna navzkrižno reaktivnost, je bila proučena tako na meritvah monoklonskih kot tudi poliklonskih protiteles.

Pri matematičnem modeliranju kinetike smo izhajali iz sledeče preproste enačbe vezave protitelesa (Ab) z antigenom (Ag) in tvorbe imunskega kompleksa (AbAg):

$$Ab + Ag \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AbAg \tag{1}$$

Predpostavljena kinetika, s katero smo opisali zgornjo reakcijo, je kinetika psevdo prvega reda, katere različica je tudi bolj poznana Michaelis-Mentenova kinetika. V ravnovesnem stanju, iz katerega smo izhajali, lahko rešitev te kinetike zapišemo kot:

$$S_{eq}(K_D, c) = \frac{S_{max}}{1+\frac{K_D}{c}} \tag{2}$$

Kjer je Kd vrednost disociacijske konstante, c je koncentracija protitelesa, S_{eq} je izmerjen signal iz označenega protitelesa v ravnovesju in S_{max} je signal ob doseženi kapaciteti vezave protitelesa na epitop antigena.

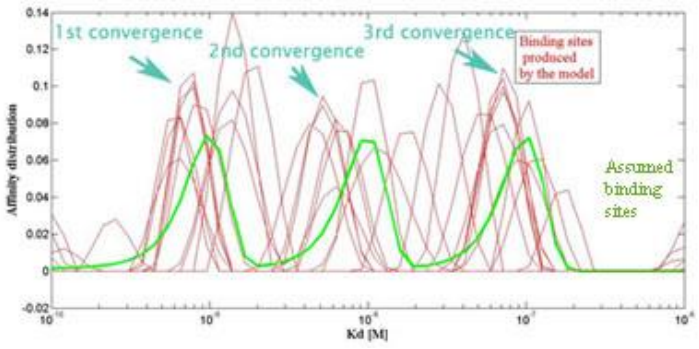
V sklopu računalniških simulacij smo na podlagi izbrane afinitetne distribucije protitelesa, izračunali točke oblike (c, S_{eq}) in jim dodali 1 % ali 10 % naključno napako. Napisali smo program v MATLAB-u, ki s pomočjo linearnih kombinacij enačbe (2) poišče najbolje prilegajočo se krivuljo simuliranim podatkov, in, kljub dodani napaki, izračuna prvotno afinitetno distribucijo. S tako postavljenim programom smo lahko analizirali podatke iz eksperimentalnega dela s protitelesi.

Osnovni princip eksperimenalnega dela je bil sledeč – vzgojili smo celice celične linije HeLa. Antigene, prisotne v celicah, smo obarvali s pomočjo direktno ali indirektno fluorescenčno označenih protiteles. Fluorescenčni signal, ki je izhajal iz protiteles, vezanih na antigene v celicah, smo izmerili s pomočjo avtomatiziranega fluorescenčnega mikroskopa za visoko rešetanje.

V prvem sklopu eksperimentalnega dela smo celice HeLa transfektirali z GFP-jem (green fluorescent protein) s pomočjo Lipofektamine-a. S tem smo vcepili antigen GFP v visokih količinah in lahko spremljali variabilnost antigena v vdolbinah 96-mikrotiterske ploščice. Na GFP smo najprej vezali mešanico dveh monoklonih protiteles proti GFPju, nato pa še poliklonsko protitelo. Za tvorbo fluorescenčnega signala smo te imunske komplekse obarvali direktno ali indirektno.

V drugem sklopu eksperimentalnega dela smo testirali zmožnost modela, da zazna navzkrižno reaktivnost, če je prisotna. Najprej smo testirali dve različni protitelesi na celicah HeLa in izračunali afinitetno distribucijo vsakega. Nato smo naredili mešanico teh dveh protiteles in izračunali afinitetno distribucijo obeh skupaj. Če je le-ta ista seštevku distribucij obeh posamičnih protiteles, potem vemo, da je model zmožen detektirali navzkrižno reaktivnost. V primeru z navzkrižno reaktivnim monoklonskim protitelesom bi lahko napovedali, da se to protitelo veže na več kot zgolj en antigen.

Rezultati simulacij so pokazali, da je s postavljenim modelom mogoče izračunati afinitetno distribucijo protitelesa iz imunofluorescenčnih meritev. Glavna parametra, ki vplivata na kvaliteto in validnost izračuna, sta količina dodane eksperimentalne napake in količina regularizacija, ki optimizira iskanje najbolje prilegajoče se krivulje.

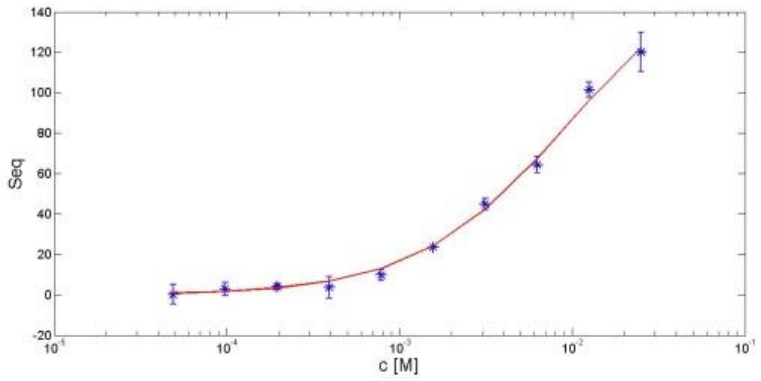


Graf 1: Izračunana afinitetna distribucija (rdeče) pri 10% eksperimenalni napaki v primerjavi z izvorno afinitetno distribucijo (zeleno)

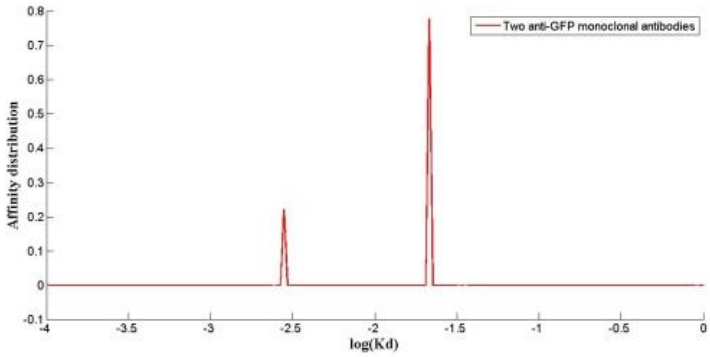
Iz grafa 1 je razvidno, da je možno razbrati tri izračunana vezavna mesta (konvergence rdečih vrhov) tudi pri 10 % eksperimentalni napaki, če je razlika v njihovi Kd vrednosti 10.

Zelena črta predstavlja izvorno afinitetno distribucijo, na podlagi katere smo izračunali simulirane točke in jim dodali napako.

Iz eksperimentalnega dela je razvidno, da predpostavljen model, baziran na kinetiki psevdo prvega reda, pravilno opiše vezavo protiteles, kot je prikazano na grafu 2.



Graf 2: Prileganje krivulje izmerjenih podatkov vezave dveh monoklonskih proti-GFP protiteles



Graf 3: Izračunana afinitetna distribucija vezave dveh monoklonskih proti-GFP protiteles na podlagi Grafa 2

Na grafu 3 vidimo, da je model napovedal dve vezavni mesti za dve monoklonsko protitelesi.

Rezultati drugega eksperimentalnega dela so prav tako potrdili, da je psevdo prvi red kinetike primeren za opis vezave protiteles. Poleg tega je bilo pokazano, da afinitetna distribucija mešanice dveh protiteles pravilno zajame večji del obeh posamičnih protiteles v

mešanici. Na podlagi tega lahko trdimo, da lahko z metodo uspešno zaznamo navzkrižno reaktivnost monoklonih protiteles.

Kar se tiče matematičnega modela, lahko zaključimo, da je zmožen izračunati afinitetno distribucijo protitelesa na podlagi s pomočjo imunofluorescenčne mikroskopije pridobljenih meritev. Določili in ovrednotili smo vpliv pomembnejših parametrov (količina eksperimentalne napake in količina regularizacije). Prav tako smo identificirali manj pomembna vpliva kot sta izbira S_{max} in števila krivulj, s katerimi izračunamo najbolje prilagajočo se krivuljo za meritve signala pri seriji koncentracij protiteles. Pri tem je vredno omeniti, da smo zaradi standardizacije simulacij in analize podatkov uporabljali konstantno vrednost parametrov regularizacije in števila krivulj, ter s tem zmanjšali variabilnost med rezultati. Dodati je treba še, da ima model težave pri razpoznavanju različnih vezavnih mest, ki imajo zelo podobno vrednost Kd.

Eksperimentalni del je potrdil, da je vezavo protiteles moč opisati s kinetiko psevdo prvega reda. Pokazali smo, da je metoda zmožna določiti afinitetno distribucijo monoklonih protiteles, ki se uporabljajo v bioloških laboratorijih. Prikazano je bilo tudi, da metoda lahko razkrije prisotnost navzkrižne afinitete v primeru monoklonskih protiteles, če imajo vezavna mesta vsaj za faktor 10 različne vrednosti Kd. V primeru poliklonskih protiteles smo ugotovili, da njihove afinitetne distribucije ni moč odkriti z razvito metodo, saj le-te v afinitetni distribuciji ne izkazujejo več vezavnih mest, pričakovanih za protitelo, ampak le eno vezavno mesto. Kd vrednost tega mesta je enaka obteženi povprečni vrednosti vseh vrednosti Kd, s katerimi se različna protitelesa v poliklonskem protitelesu vežejo na isti antigen. To je rezultat kompetitivne vezave zaradi vezave na isti antigen, ki pa ni prisoten v primeru monoklonskih protiteles.

LIST OF ABBREVIATIONS

Abbreviation	Description
Ab	Antibody
Ag	Antigen
BSA	Bovine Serum Albumin
CDR	Complementary Determining Region
DAPI	4,6-diamidino-2-phenylindolem
DMEM	Dulbecco's Modified Eagle Medium
ERK	Extracellular Signal-regulated Kinase
FBS	Fetal Calf Serum
GFP	Green Fluorescent Protein
K_D	Dissociation Constant
NF200	Neurofilament subunit with molecular mass of 200 kDa
P	Differential distribution factor of binding sites
PBS	Phosphate Buffered Saline
pERK	Phosphorylated Extracellular Signal-regulated Kinase
PFA	Paraformaldehyde
PMA	12-O-tetradecanoylphorbol-13-acetate
PS	Penicillin-streptomycin
S_{eq}	Signal intensity at equilibrium
S_{max}	Maximum binding capacity of an epitope

1. INTRODUCTION

Antibodies are a valuable tool both in modern day pharmaceutical and biological analytics and as therapeutics. Mastering antibody characteristics and properties is of vital importance for their application and use.

A problem encountered in cell biology is how to correctly track and quantify species involved in cell's biochemical processes. The solution is represented by the use of antibody based techniques such as immunocytochemistry by the means of which biological species are labeled and observed. However, for such techniques to produce valid and reproducible results antibodies must be properly characterized and evaluated.

In order to characterize antibodies, interaction with antigens can be exploited to reveal antibody's nature and bring better understanding of how they may influence the validity of immunoassay results.

1.1. ANTIBODY

Immune system has a unique way for identification and neutralization of foreign object such as bacteria and viruses. It uses a series of large Y-shaped proteins called antibodies or immunoglobulins, created by beta cells with a unique characteristic of specific recognition of antigen (11). They possess an ability to elicit an immune response against immunogen by the recruitment of other cells and molecules. Specific recognition derives from interaction of antibody's antigen binding site with an epitope on antigen. Antigen binding site is a section of antibody responsible for antigen binding. On the other hand, epitope is a section of antigen which is specifically recognized by antibody's antigen binding site. When bounded together they form an antibody-antigen complex, on which detection immunoassay techniques are based.

Antibodies provide three analytically significant properties, which are made use of in immunoassay techniques:

- RANGE: Capability of binding to an extremely wide range of chemicals, biomolecules, cells and viruses, gives antibodies a vast spectrum of application possibilities. This derives from antibodies being proteins comprised of amino acids, each of them having the unique binding and orientation properties. On top of that,

sequences of amino acids have an ability of taking different kind of three-dimensional forms, making multiple-site binding possible.

- STRENGTH OF BINDING: Despite bonds between antibody and antigen being non-covalent, their net effect creates a strong enough connection for antibody-antigen containing samples to undergo processing and signal generation stages of immunoassay techniques. High binding affinity provides the ability for immunoassays to be accurate and precise even at low concentrations of antigen in biological samples.
- SPECIFICITY: Antibodies express extraordinary binding specificity for its antigen. This phenomenon makes detection of extremely low concentrations (picomolar) of antigen amidst many closely related substances possible. It is a necessary fundamental property of antibodies to be used in immunoassays.

1.1.1. ANTIBODY PRODUCTION

The source of all initial antibody production is an active immunization of laboratory animals via immunogen. As a molecule capable of inducing an immune response when injected into an animal, immunogen is not to be confused with an antigen which does possess antibody binding capabilities yet it does not by default induce an immune response. Both immunogen and antigen have many different epitopes. Immunogen is bound to have certain characteristics such as sufficient molecular mass. If, however, a small molecule is used, it is usually conjugated to carriers such as albumin to be effective. The most effective immunogens are proteins and polysaccharides, as well as nucleic acids, lipids and synthetic polypeptides.

Antibody production can yield two kinds of antibody reagents – monoclonal antibodies produced by a single clone of beta cells and polyclonal antibodies produced by several clones of beta cells. Polyclonal antibody or polyclonal antisera is a heterogeneous mixture of antibodies with variety of binding affinities and specificities. This is due to antibody recognizing epitopes not only on immunogen, but also on any impurities injected with it. Variable quality of polyclonal antisera derives from numerous factors influencing the antibody production, such as impurity content of immunogen, choice of animal species, site and timing of injections. Nonetheless, many analytical techniques do not require high quality polyclonal antisera to produce successful results.

However, in some applications antisera purification is necessary. Four different techniques may be used – salting out with ammonium or sodium sulfate and separation with dialysis or gel filtration, gradient ion-exchange techniques, and purification using specific and reversible binding of antibodies to lectins. Not only do these techniques offer lower than desired yields or are very expensive, they also non-specifically isolate total IgG fraction, which is comprised of whole array of antibodies. The forth method, immunoaffinity purification takes advantage of specific antibody-antigen binding by attaching an antigen of interest to an inert support and passing heterogeneous antibody mixture through such column. After extensive washing, only antibodies that bind with the highest affinity are left bound to antigens. Eluting these, polyclonal antibodies of high quality are collected.

Monoclonal antibodies are represented as a homogeneous antibody reagent. They were first created by isolating and propagating antibody producing beta cells clones *in vitro*. A procedure combining beta lymphocites with immortal myeloma cells was devised in order to create hybrids inbred with ability to secrete specific antibodies and the capability of proliferating indefinitely. These so called hybridoma cells undergo an extensive selection procedure before a particular clone is isolated, yielding monoclonal antibodies with outstanding specificity in theoretically limitless quantities. Antibodies created in such procedure are able to differentiate between very similar molecules, cells or microorganism. They are also easier to store than polyclonal and have as such proven to be very useful both in fundamental biological research as in medical and diagnostics applications of antibodies (1).

1.1.2. ANTIBODY STUCTURE

Antibodies are complex protein molecules consisting of four polypeptide chains revealed by reductive splitting of their disulfide bonds. The two heavier ones (50 kDa) are known as heavy chains which form the basic antibody structure. The two lighter ones (25kDa) are known as light chains which are together with heavy chains responsible for forming antigen binding site.

On the other hand, another antibody structure representation is in use. Reduction of antibody with an enzyme papain yields three similar fragments. Two of these fragments, known as Fab (fragment antigen binding), are identical and each one has antigen binding site. The third one does not express the ability of antigen binding and is known as Fc (fragment cristallizable). It was discovered that Fc is responsible for antibody's effector functions, such as complement or

cell binding. Inability of papain to further split fragments suggested an existence of a hinge region. Taking advantage of X-ray crystallography, a three dimensional crystalized antibody structure was revealed to be in the shape of letter Y and that both Fab regions are flexibly bound to Fc region, as seen from Figure 1.

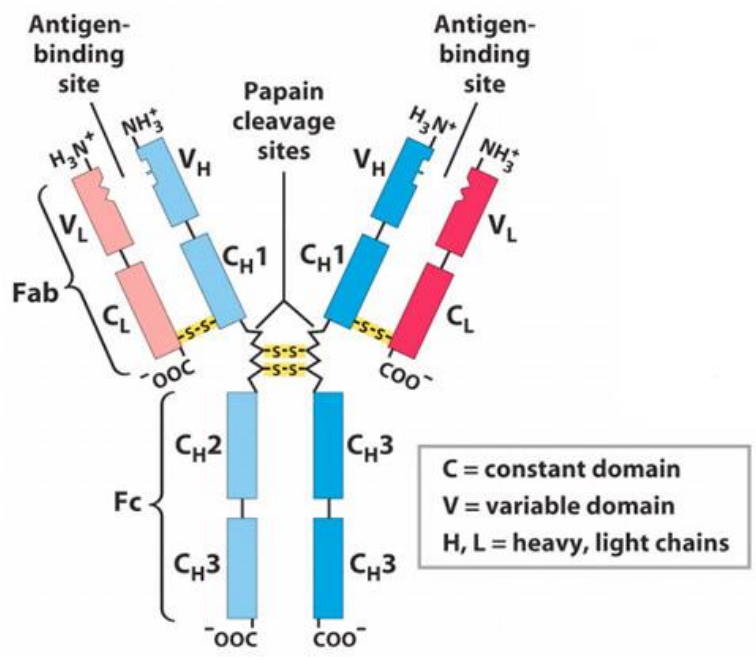


Figure 1: Antibody structure (21)

Both light (211-217 amino acids) and heavy (450-550 amino acids) chains consist of repeating structural domains, which are about 110 amino acids long. In turn, the light chain is known to possess two structural domains, while heavy chain is organized into 4-5 domains, depending on heavy chain class (human immunoglobulin are classified on the basis of theirs heavy chain antigen properties into five classes, namely IgG, IgA, IgM, IgD and IgE). Domains that terminate with amine group express a significant variability, which is vital for specific antigen binding process. Light chain has one of these variable domain and one constant, while heavy chain has one of variable domain as well and three constant ones. Digging deeper into variable domain, a parameter reflecting the degree of variability was introduced. As a result, three short hypervariable regions (10 amino acids or less) or CDRs (complementary determining regions) were found in each light and heavy chain.

1.1.3. ANTIGEN BINDING SITE

The CDRs represent antibody's antigen binding site. Highly variable amino acid sequences in CDRs result in variable dimension and binding grooves of antigen binding site which fundamentally influences the variety and specificity of antibodies (5). Majority of antibody-antigen binding interactions are a direct result of epitope on antigen binding to CDRs of antibody. However, recent studies have suggested that as much as 22% of antigen binding residues fall out of CDRs, proposing the influence of binding site framework. Disconcertingly, modern tools for identification of antigen binding regions rely solemnly on recognition of CDRs (12). Theory of key-and-lock antigen binding considers that antibody-antigen interaction derives almost exclusively from CDRs interaction with epitope. This, nonetheless, is based on experiments using small carbohydrate molecules as antigens. Most antigens, however, are big molecules, and only a small part of these molecules is interaction with CDRs. New aspects revealed how, in addition to CDRs, antibody framework is also of significance in interaction process which explains how bigger antigen molecules can interact with non-CDR antibody parts as well.

An antibody can, in addition to specific binding to epitope, also bind to other non-specific epitopes. This is biologically of great importance. However, antibody binding to different epitopes and thus many different antigens makes the use of antibodies in biological laboratories challenging. Nonetheless, even though multiple binding is present, in majority of cases binding to specific epitope is high in relation to binding to non-specific epitopes. Effect of non-specific binding tends to be under level of detection in many immunological tests (5, 13).

1.2. FORCES GOVERNING ANTIBODY-ANTIGEN INTERACTION

Antigen binding sites interaction with epitope is governed by different attractive forces, the strength of which is largely dependent on distance between interaction species. CDRs, primarily responsible for antibody's specific recognition, are largely complementary to epitope. In effect, short distance between CDR and epitope yields higher binding strength produced by attractive forces. High binding affinity between antibody and antigen is not a result of strong covalent bonds, but an effect of numerous weaker interactions working

together. These attractive forces consist of hydrogen bonds, electrostatic bonds, Van der Waals forces and hydrophobic forces.

From physical standpoint, the strength of these forces is fundamentally determined by distance (d) between groups responsible for binding. Namely, electrostatic forces are proportional to $1/d^2$ and van der Waals forces exhibit $1/d^7$ proportion. This means that for the binding to take place, CDR of antibody and antibody binding region on antigen must be in close proximity. This represents the need for structural similarity between interacting species, not only in shape, but in charge distribution and hydrophobicity as well. The sum of attractive and repulsive forces acting on monovalent Fab fragment and antigen epitope is referred to as affinity, which serves as indication of interaction strength. Made up of non-covalent bonds, the antibody-antigen binding is a reversible reaction, where at equilibrium the Law of Mass Action may be applied and equilibrium constant can be determined (14).

1.3. DYNAMICS OF ANTIBODY-ANTIGEN BINDING

The dynamics of antibody binding to antigen quantifies how antigen binding site is interacting with epitope to form an antibody-antigen complex. Such dynamics is comprised of three interacting species, namely antibody (Ab), antigen (Ag) and antibody-antigen complex (AbAg).

Basics of antibody-antigen binding are represented as a standard reversible interaction of antibody and antigen to form antibody-antigen complex, according to the following equation:

$$Ab + Ag \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AbAg \quad (1)$$

During the reaction, concentration of unbound Ab and Ag decreases, while concentration of AbAg increases. At dynamic equilibrium, rate constant of forward reaction k_1 is the same as rate constant of reverse reaction k_2. The overall rate – quantity of molecules which react in a certain time unit – is heavily dependent on species concentration. Equilibrium concentrations are a function of binding energy. Thermodynamically speaking, equilibrium is achieved when Gibbs free energy is equal to 0. As stated by Law of Mass Action, the ratio of products of the concentrations on both sides of the equation will at equilibrium be constant. If we denote such a constant with Kd (dissociation constant), the following sentence hold true:

$$\frac{[AbAg]}{[Ab][Ag]} = Kd \quad (2)$$

K in such sense is known as affinity constant, with 1/mol units. Incorporating k_1 and k_2, the following is apparent:

$$Kd = \frac{k_1}{k_2} \quad (3)$$

Dissociation constant Kd serves as a measure of energy of reaction. Its use in immunoassay techniques is the ability to quantify the ratio of free ligand (Ab or Ag) to ligand bound in AbAg complex, given that the concentration of one of the ligand is known. This is the basic principle of all binding immunoassays (2).

1.4. ANTIBODY SPECIFICITY

Antibody specificity within the content of immunoassay use is defined as a *'degree to which an assay responds to substances other than that for which the assay was designed'* (2). Generally, immune reactions are notoriously specific. Particular antibody excessively binds to antigen of interest and little else, albeit antigens structurally very similar to antigen of interest will have increased likelihood of exhibiting an antibody binding abilities (15).

It may come as a surprise that high affinity antibodies tend to be less specific than low affinity antibodies (5). This is a consequence of high affinity antibodies producing measurable reaction with molecules similar to antigen. Even though such binding is of lower affinity than specific binding to the antigen of interest, it can nonetheless produce a measurable effect. On the other hand, antigen of interest itself may be of low affinity and thus barely measurable. Antigen similar molecules with even lower affinity will thus be more likely to fall under the level of detection. Different affinities of antibody solutions with variable specificity can be useful for some technical approaches. High specificity of low affinity antibody binding (for example to carbohydrates) is in practice used for typization of blood groups.

While high affinity of antibodies is vital for immune system to be able to perform, a high specificity plays a most fundamental property in immunoassays used in basic scientific research and is as such of paramount importance (5).

1.5. ANTIBODY CROSS-REACTIVITY

Antibody binds specifically to antigen of interest. However, antibody can also non-specifically bind to another antigen. This is, in context of an immunoassay, referred to as cross-reactivity. Cross-reactivity of an antibody is a measure of a degree to which an antibody

can interfere with an assay due to non-specific binding. This may happen for the following reasons:

a) A monoclonal antibody can bind to spectra of closely related antigens with various affinity constants.
b) Polyclonal antibody antisera can contain one or more antibodies aimed at a site on primary antigen. This site can also be present in closely related antigens.
c) Polyclonal antibody antisera can contain antibodies directed at antigen contaminants which are also present in related antigens.

1.5.1. DETECTING CROSS-REACTIVITY

Cross-reactivity may experimentally be determined using standard curves, representing the percentage of antibody or antigen bounded at different dilutions of its counterpart. From graph theory, we are to distinguish between parallelism and non-parallelism of different standard curves (2). In essence, if we target a single antigen in a complex systems, the standard binding curve will always be parallel, be it if we target it with one or numerous antigen specific antibodies of variety of affinity constants. That is to say, the shape of the standard curve in the case of one antibody of certain binding affinity or ten antibodies of various binding affinity will be the same and thus parallel in both cases. Observable from the curve, in both cases antibodies react with antigen as if they had a single binding constant. In case with one antibody this is actually the case, while the seeming one binding affinity constant for ten antibodies will simply be a weighted-average affinity of ten different binding constants. Thus antibody or antibody mixture which standard curve can be expressed with a single binding affinity constant will only target one specific antigen, and said antibodies will in such a case exhibit no cross-reactivity (16). The conclusion is valid only if the binding capacity of antigen of interest with relation to any other antigen that may be responsible for cross-reactivity is not significantly higher (2). On the other hand, if beside targeted antigen, antibody or antibodies bind to anything else, or even to the other epitopes on the targeted antigen, the curve will be influenced by many different binding affinity constants and will as a result be non-parallel. Different binding affinity constants constitute the very nature of cross-specificity (16).

In clinical studies, the relevance of lack of cross-reactivity is connected with the likelihood of interference with the practical operation in used assays. For example, in radioimmunoassays

the thyroid hormone tyhroxine (T_4) shows substantial cross-reactivity (>10%) with triiodothyronine (T_3) (2). Nonetheless, due to circulating levels of T_3 being for a factor of at least hundred times less than that of T_4, the cross-reactivity produces insignificant disturbance. On the other hand, a radioimmunoassay for T_3 may show 1% cross-reactivity with T_4, which is much more abundant. Thus T_4 could influence results very significantly (2).

1.6. IMMUNOASSAYS

Antibodies have wide spectra of applications. Immunoassays represent the core of analytical application of antibodies. Immunoassays are analytical tests heavily based upon harnessing the unique properties of antibodies. Being able to serve as a sophisticated chemical toolbox enabling scientist to investigate and manipulate minute concentrations of complex biomolecules, they proved to be one of the most important technological contribution in medicine and fundamental life science research in 20^{th} century. Since its inception, immunoassays demonstrated the potential to be applied in wide range of areas, primarily in pharmaceutical, veterinary, forensic, military and life sciences. Novel and ingenious assay designs were spread and implemented in clinical research and diagnosis. They were utilized as a means to simplify and automatize broad range of supporting technologies. Thus they are thought-of as fundamental and invaluable tool in life science research by being able to pinpoint the nature of basic biological phenomena, whilst at the same time providing us with the means of simplification they incorporate to make everyday applications such as pregnancy test for use in one's household possible.

Principle concept of immunoassays is based on generating a measurable signal from a sample (1). Signal strength is the basis for quantification of antigen (analyte) in underlying reaction of an antigen and an antibody. Generally, immunoassays are bioanalytical methods for evaluating a competitive binding reaction between fixed amount of labeled form of analyte and a variable amount of unlabeled sample analyte for limited amount of binding sites on antibody specific for analyte used. In the process of immunoanalytical reagents mixing and incubating, the immune complex is formed by antibody binding to antigen. Produced complex undergoes a separation procedure, where unbound reagent fraction is washed away by physical or chemical separation techniques. To acquire a signal, a so called tracer is implemented in either of the bound or free fraction. By measuring tracer activity (e.g.

radiation, fluorescence, or enzyme) analysis of the sample can be made. For this purpose, a standard curve representing measured signal as a function of concentration of the unlabeled analyte in the sample is generally constructed. This calibration curve represents a tool with which an unknown analyte concentration is determined (3).

1.6.2. IMMUNOFLUORESCENCE

Being one of most important modern immunoassay designs, immunofluorescence is a common laboratory technique used in all aspects of biology. Its wide applicability ranges from basic research to clinical diagnostics. Signal detection indicating the characteristics of immunoassay reaction of antibody and antigen is obtained by using a fluorescent tracer. Tracer fluorescent dyes such as fluorescein isothiocyanate or tetramethyl rhodamine isothiocyanate are chemically conjugated to antibodies, which upon reaction with antigen of interest allows for antigen detection using fluorescence techniques. Quantification of fluorescence and thus antigen content can be achieved with the use of automated imaging instrument, array scanner, flow cytometer or by visual inspection with fluorescence microscope.

When it comes to attaching fluorescent tracer or immunofluorescence labeling, there are two distinct approaches. Less frequently used is **direct immunofluorescence labeling** where a reacting antibody is conjugated directly to fluorescent tracer. This approach offers not only shorter staining protocols, but more importantly a simpler multiple labeling procedures. Also in some cases, where antibodies used are raised in the same species (such as two mouse monoclonal antibodies), a direct labeling is a preferred choice. On the other hand, even though tracer should enable higher signals, direct labeling produces low signals and it generally represents a substantial financial burden. A lack of commercially available direct conjugates for some antibodies renders this procedure less useful.

On the contrary, with **indirect immunofluorescence labeling**, an antibody specific for antigen of interest is not labeled. To generate a signal, a secondary anti-immunoglobulin antibody labeled with a fluorescence tracer is bound to constant portion of the first antibody. Because many secondary antibodies can attach to the primary, the signal obtained is amplified which results in greater sensitivity in comparison with direct labeling. In addition, secondary antibodies are produced on a greater scale, rendering them relatively inexpensive and quality controlled, while also having an additional flexibility in concern to wider array of colors

available. However, with additionally antibody-antibody reaction the chance for cross-reactivity is increased. Moreover, primary antibodies must not be raised in the same species as the secondary, while there is also a concern of endogenous immunoglobulin exhibiting a high background.

Of the limitations of immunofluorescence techniques one of the most concerning is photobleaching. Due to photochemical destruction of tracer fluorescence dyes, the signal generated from a sample depends on the extent of photobleaching effect. To achieve stronger signals, intensity and duration of excitation light during storage and experimentation is decreased. Another potential difficulty arises from autofluorescence of the sample, which can be minimized by the selection of probes and optical filters that maximize the fluorescence to autofluorescence ratio. What is more, in cases where multiple tracers are used simultaneously a potential complication due to fluorescence overlapping is present. Fluorescence dyes absorb light at certain wavelength and in turn emit light of higher wavelength. The emission is, however, continuous over considerable range of wavelengths. Ranges of different fluorescence tracer can overlap each other, creating a higher than expected signal in overlapping regions. With process called fluorescence compensation these overlaps are electronically removed on the basis of using a positive control (6).

1.6.3. CROS-REACTIVITY IN IMMUNOASSAYS

The use of immunoassays in basic research science has offered a gateway to a lot of new information. However, modern day assays in immunohistochemistry as an example take an advantage of even ten or more antibodies for cell colocalization and identification. Recent discoveries put the credibility of such immunoassay designs to question due to questionable antibody staining specificity. Most notably, some antibodies were found to stain cells from which antigen of interest, for which antibody is supposed to be specific, were knocked out. Due to this cross-specificities not only were scientific papers withdrawn (17), but the very integrity of scientific research was put in danger. New standards of validating the staining specificity, claimed by antibody manufacturer, are being put into practice. As dr. Clifford B. Saper, an editor-in-chief of The Journal of Comparative Neurology (17), writes, there are predominately three ways to ensure antibody staining specificity:

1) Most credible is the assurance of antibody not staining cell of interest from which molecule of interest is absent. This golden standard can be achieved by targeting molecules not usually present in cells (such as green fluorescence protein) or by the use of knock-out mice. However, existence of mice with the desired knock-out is limited.

2) Less certain but practically useful technique for antibody specificity assurance, especially for antibodies raised against synthetic peptides, is the preabsorption of antisera against immunizing peptide. The idea is that staining will disappear since antibody targeting antigen of interest has been effectively removed from the sample by being preabsorped. This, however, does not exclude non-specific interactions via other parts of the antibody. Many a paper was withdrawn when antisera passed the test with preabsorption but failed the knockout test. Test is also not applicable to monoclonal antibodies, which will always be absorbed out by their antigen.

3) Lacking the antigen, staining pattern can be referenced to an identical pattern previously reported, which has better characterization (preabsorption test for example). In this case, assurance is largely dependent of test the previously reported staining pattern has undergone.

It is of utmost importance for any antibody in use to clearly show the lack of cross-reactivity and assure staining specificity (17).

2. RESEARCH AIM AND OBJECTIVES

2.1. AIM

Antibody-based techniques are central to life science research and therapy. One of the central pitfalls in modern immunoassay techniques is the assurance of staining specificity. Antibody applications in basic scientific research share a common weakness in a form of antibody cross-reactivity. Existing methods for cross-reactivity detection are limited. Method with knock-out mice is applicable only if the knock-out exists, while the preabsorption techniques are constrained to polyclonal antibodies. A new, simple and time efficient method is to be devised that can not only identify potential cross-reactivity, but possess the ability to characterize antibodies even further by capturing the affinity distribution of antibody itself.

Aim of this master's thesis is to create a novel method for antibody characterization with the use in detection of antibody cross-reactivity.

2.2. OBJECTIVES

Approach taken will be based on understanding the nature of antibody binding to heterogeneous epitopes in a standard laboratory use of immunofluorescence assays. Project will be divided into two distinct parts – simulation of a mathematical model of antibody-antigen binding. Secondly, a wet lab experimentation of the binding will be conducted to verify the created model, and evaluate the likelihood of cross-reactivity detection.

Objectives of mathematical modeling and computer simulations:

- Theoretically determine the mathematical model of interaction by which heterogeneous antibodies are binding to heterogeneous epitopes, and determine the parameters of binding process.
- From proposed model, computationally solve the indirect problem of acquiring the affinity distribution from binding process using simulated experimental binding data.
- Simulate how likelihood of determining cross-reactivity detection is being influenced by varying the parameters of binding and level of cross-reactivity.

Objectives of wet lab experimentation:

- Confirm theoretically determined mathematical model of antibody-antigen binding with monoclonal and polyclonal antibodies targeting specific antigen.
- Implement and evaluate the mathematical parameters of binding process on monoclonal antibodies by creating an artificial cross-reactivity experimental system.
- Determine the strengths and weaknesses of the model.

3. MATERIALS AND METHODS

3.1. MATHEMATICAL MODELING

The goal was to characterize antibody through the use of antibody-antigen binding equilibrium signals.

A time-course that best describes a process involving two entities, with the value of one constant and the value of other dependent on time, is pseudo-first-order differential equation. In essence, Michaelis-Menten kinetics used in enzyme dynamics is a variation of such a differential equation. In this case, it was stated as follows:

$$\frac{dS}{dt} = k_{on} \cdot c \cdot (S_{max} - S) - k_{off} \cdot S \tag{4}$$

Time of binding is denoted by t and concentration of antibody is denoted by c, while S_{max} is the signal produced at saturation. Focus of the research was on equilibrium data - the reaction was allowed to achieve dynamic equilibrium, at which the signal S_{eq} was collected. In order to calculate the signal intensity at equilibrium (denoted by S_{eq}), the equation was integrated with respect to S from 0 to S_{eq} and with respect to t from 0 to infinity as follows:

$$\int_0^{S_{eq}} \frac{dS}{k_{on} \cdot c \cdot (S_{max} - S) - k_{off} \cdot S} = \int_0^{\infty} dt \tag{5}$$

At this point known expression $K_D = k_{off}/k_{on}$ was substituted into equation *5* and integrating it the following expression for S_{eq} in terms of K_D and c was obtained:

$$S_{eq}(K_D, c) = \frac{S_{max}}{1 + \frac{K_D}{c}} \tag{6}$$

Signal at equilibrium S_{eq} depends on concentration of antibody c, binding affinity K_D between them, while the maximum binding capacity of an epitope S_{max} in certain setting was constant. Theoretically, S_{eq} is equal to S_{max} when concentration of antibody c is infinite.

The model was expanded to a realistic antibody cross-reactivity situation where an antibody does not only bind to one specific type of epitopes, but to at least one more type of epitopes. Each of such epitopes represents a specific binding site for antibody with unique binding affinity K_D and binding capacity S_{max}, and thus produces unique equilibrium signal S_{eq} at each concentration of antibody. Proposed model for such binding was as follows:

$$S_{eq} = \frac{S_{max1}}{1+\frac{K_{D1}}{c}} + \frac{S_{max2}}{1+\frac{K_{D2}}{c}} + \frac{S_{max3}}{1+\frac{K_{D3}}{c}} + \cdots + \frac{S_{max\,n}}{1+\frac{K_{D\,n}}{c}} = \sum_{i=1}^{n} \frac{S_{max\,i}}{1+\frac{K_{D\,i}}{c}} \quad (7)$$

In equation *7* equilibrium signals from individual epitopes were summed into a net signal S_{eq}. Evaluation of contribution to net signal by each individual binding site was essential to obtain the affinity distribution of antibody. In order to obtain such K_D distribution as a result, a total capacity of all epitopes $S_{max\ tot}$ was introduced. This was done by formulating a differential distribution factor of binding sites denoted by $P(K_D)$, which effectively captures the contribution of each binding site to the net signal. The defining property of $P(K_D)$ is that an integral over any K_D range is equal to binding capacity of all binding sites within that range. Such a property was written in this way:

$$S_{max\,tot} = \int_{K_{D,min}}^{K_{D,max}} P(K_D) \cdot dK_D \quad (8)$$

Herewith all the basic parameters in mathematical model were defined and expressed.

3.2. COMPUTER SIMULATIONS

All computer simulations were done using MATLAB 8.1. Released on 7th of March 2013 under release name R2013a by the company The MathWorks, Inc., Matlab is a numerical computing environment.

Simulations were based on equation *7*. The purpose of computer simulations was to program and evaluate the mathematical model with the use of assumed affinity distribution scenarios.

3.2.1. INVERSE PROBLEM

Reverse engineering approach was used by solving a so called inverse problem. First, we designed the artificial K_D distribution by proposing one or several binding sites. Secondly, data points were calculated based on the underlying K_D distribution and a level of uncertainty was introduced for each point calculation in order to better make simulation closer to realistic web lab scenario. Lastly, simulated data points were be fitted by a least-square based fitting algorithm.

Assumed Kd distribution peaks were normally distributed. Based on such K_D distributions, data points of form (c, S_{eq}) were calculated with equation *7*. A constant random noise was added into calculations and arbitrary values of S_{max} for each binding site were used.

In the last part, noise-including calculated data points were fitted using least-square fitting algorithm. Algorithm uses custom made fitting curves with linear combination of curves based on equation *6* with different K_D values. Best fitting curve is the one where the squared net of distances between data points and points on the curve at certain concentration value is the lowest. In order to obtain best fitting curve, the following minimization equation was numerically solved:

$$\min_{P \geq 0} [\overrightarrow{DP} - A \cdot \vec{P}]^2 \quad (9)$$

DP is a vector of data points, A is matrix containing all possible K_D and c combinations (ie. A is vector product of $\mathbf{K_D}$ and **c**). **P** is the differential distribution factor of binding sites, otherwise known as affinity distribution factor or K_D distribution factor. Its purpose is to capture the effect of an individual curve contribution to the net signal over a selected K_D range.

In order to computationally solve equation *9* a MATLAB's inbuilt function *quadprog* was used. *Quadprog* is an algorithm that solves minimization equation in quadratic programming.

To numerically stabilize the resulting K_D distribution, a regularization factor was added to minimization equation *9*. This factor stabilizes equation by the way of reducing the importance of more deviating data points in curve fitting. A Tikhonov-Phillips regularization was used which was applied in minimization equation by adding a stabilizing term in the following way:

$$\min(_{P \geq 0} [\overrightarrow{DP} - A \cdot \vec{P}]^2 + a \cdot B(\vec{P})) \quad (10)$$

Whereas B is a regularization constraint preventing curve overfitting by choosing the smoothest possible curve fit. a is a weighting parameter for B, arbitrary chosen to strike a right compromise between overfitting and smoothness of the curve. In terms of P and K_D, the constraint B was expressed in the following way:

$$B(\vec{P}) = \int (\frac{d^2 \vec{P}}{d \log(K_D)^2})^2 dK_D \quad (11)$$

On the basis of mathematical model and this simulations approach various simulations scenarios were devised and evaluated (16).

3.3. WET LAB EXPERIMENTATION

In order to validate the created mathematical model, a laboratory based data were obtained in order to feed the computer simulation with realistic (c, S_{eq}) data points. Basic layout of experimental design was comprised of the following parts:

- Cells containing specific epitopes.
- Antibodies targeting specific epitopes in cells.
- Quantification of antibody-epitope interactions via directly/indirectly fluorescence antibody labeling and measuring the signal intensities by a high screening fluorescence microscope.

3.3.1. PREPARATION, PASSAGING AND PLATING OF HELA CELLS

An immunocytochemistry experiment implementing standard laboratory procedure of immunofluorescence staining was used for evaluation of antibodies.

a) PREPARATION OF HELA CELLS

The purpose of preparation of cells was to prepare them for cell passaging in order to be later used for immunostaining experiments.

Table 1: Materials used for HeLa cell preparation

Material name	*Company name*
DMEM (1x) (Dulbecco's Modified Eagle Medium) - 4.5 g/L D-glucose - L-glutamine - Pyruvate	Life Technologies
PS (Penicillin-Streptomycin) Solution (10.000 U/ml)	Life Technologies
FBS-Superior (Fetal Calf Serum)	Biochrom AG
Cell culture flask T75	Costar
HeLa cell line (frozen)	DSMZ

I. <u>Protocol for prepreparation of cell medium</u>

Take out 445 ml of DMEM and put it into 500 ml sterile flask. Add to the medium 5 ml of PS Solution and 50 ml of FBS-Superior. Equilibrate everything to 37^{o}C in a water bath. This mixture is referred to as 'cell medium' from here on.

II. Protocol for defreezing HeLa cells

Put 5 ml of prewarmed cell medium in a 50 ml tube. Thaw the vial with frozen HeLa cells carefully in the water bath until there are only small pieces left in the vial. Put defrozen HeLa cells into 50 ml tube with prewarmed cell medium and centrifuge for 8 min at 1200 RPM. Discard the supernatant and resuspend the cells in 10 ml fresh prewarmed cell medium.

III. Protocol for incubating HeLa cells

Put the cells into cell culture flask. Place the flask into incubator with 37 ^{o}C and 5 % of CO_2.

b) PASSAGING AND PLATING OF HELA CELLS

Passaging of HeLa cells served as a supply of cells for plating in order to be used in antibody binding experiments.

Table 2: Materials used for HeLa cell passaging and plating

Material name	*Company name*
Cell medium	
Trypsin-EDTA (1x) - 0.05% Solution	Life Technologies
Dulbecco's **PBS** (1x) (without Mg or Ca)	GE Healthcare
96-well **microplates**, black/clear Imaging **Plate**	BD Falcon
Cell culture **flask** T75	Costar
Neubauer chamber	
Light microscope	Leica

I. Protocol for cell passaging and cell plating

Prewarm cell medium, Trypsin and PBS in water bath to 37 ^{o}C. Take cell culture flask containing HeLa cells out of the incubator. Take of the medium in the flask. Wash with 10 ml of PBS. Add 2 ml of Trypsin and incubate for 3 min at 37 ^{0}C, 5 % CO_2 in the incubator. Take flask from the incubator and gently shake it for cells to detach. Confirm cell detachment under light microscope. Stop enzyme function of trypsin by adding 8 ml of prewarmed cell medium. Resuspend the cells, put them into a 50 ml tube and centrifuge for 8 min at 1200 RPM. Take off supernatant and homogenously resuspend the cells in the sediment in 30 ml of cell medium. Take 10 µl of the homogenous cell suspension and put it onto the Neubauer chamber. Wait for 1-2 min for cells to settle down. Count the cells in the 4 corner quadrants and calculate the cell number per ml (No) using the following equation:

$$\frac{\text{sum of cells in 4 quadrants}}{4} \cdot 10^4 = \text{No} \qquad (12)$$

Take 1-5 ml of homogenous cell suspension with known cell concentration and dilute it to desired concentration.

CELL PASSAGING: Put10 ml of diluted cell suspension into cell culture flask and place it in incubator (37 ^{0}C, 5 % CO_2). Concentration for cell passaging was 5.000 cells/ml.

CELL PLATING: Place 100 µl of diluted cell suspension in each well of 96-well plate and place it in incubator (37 ^{0}C, 5 % CO_2). Concentration for cell plating was 20.000 cells/ml.

3.3.2. GFP TRANSFECTION AND EVALUATION OF ANTI-GFP ANTIBODIES AFFINITY DISTRIBUTION

Cells were transfected with green fluorescent protein (GFP), which served as an antigen for anti-GFP antibodies used in antibody binding experiments. GFP is a protein comprised of 239 amino acid residues (26.9 kDa) that exhibits green fluorescence when stimulated by blue to ultraviolet light (22). GFP transfection was done in the interest of accomplishing two things. Firstly, to create large quantities of antigen antibodies are going to target in order to obtain measurable signals even at very low antibody concentrations. In addition, GFP served as a control of well-per-well antigen content variability.

3.3.2.1. GFP TRANSFECTION AND OPTIMIZATION

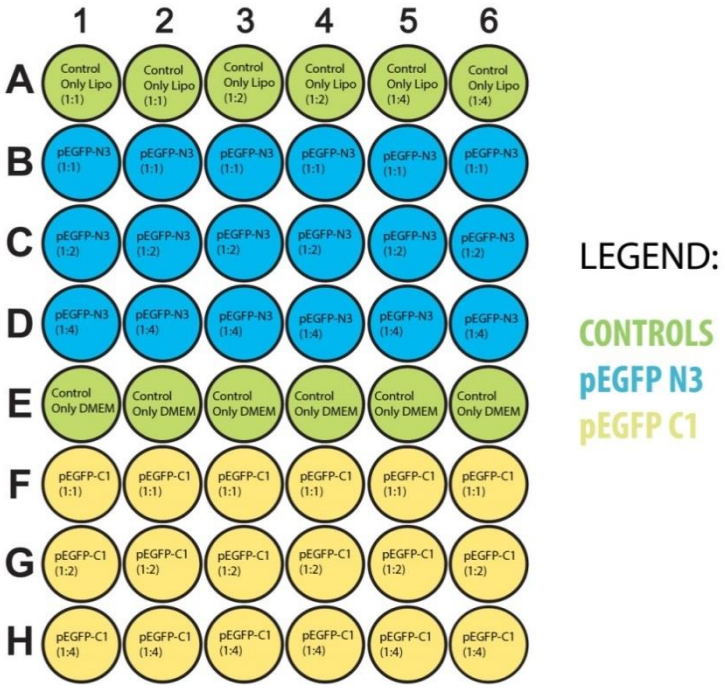
For optimum cell transfection, two different plasmids containing GFP regions were used in three different concentration ratios with transfection reagent Lipofectamine. The purpose of this experiment was to determine most successful transfection environment.

Table 3: Materials used for GFP transfection of HeLa cells

Material name	*Company name*
Half of 96-well **microplate** containing app. 10.000 cells/well	
Lipofectamine 2000	Invitrogen
pEGFP-C1 plasmid 1 µg/µl (22)	Clontech
pEGFP-N3 plasmid 1 µg/µl (23)	Clontech
DMEM (1x) (Dulbecco's Modified Eagle Medium) - 4.5 g/L D-glucose - L-glutamine - Pyruvate	Life Technologies
Fluorescence microscope	Leica

I. Protocol for cell transfection

Table 4: GFP transfection plate plan



Day before transfection experiment plate the cells, putting 100 µl of cell suspension containing app. 2.000 cells in each of 96 wells. Before transfection, replace cell medium above cells with 50 µl of pure DMEM (antibiotics can, interfere with transfection) prewarmed to 37 ^{o}C. Dilute 4.0 µg of each of the two plasmids in two separated vials each containing 496 µl of DMEM and mix gently.

Mix Lipofectamine 2000 before use and make 3 different dilutions: 4.0 µl of Lipofectamine in 496 µl of DMEM, 8.0 µl of Lipofectamine in 492 µl and 12 µl of Lipofectamine in 488 µl. Incubate Lipofectamine dilutions for 5 min at room temperature.

Combine prepared plasmids and Lipofectamine in ratios of 1:1, 1:2 and 1:4 by putting 150 µl of plasmid solution in each of 3 different vials and adding 150 µl of one Lipofectamine solution. Using two plasmids and three Lipofectamine solution we obtain 6 vials. Incubate 6 vials for 20 min at room temperature. Add 50 µl of solutions per well according to plate plan illustrated below. Mix gently by rocking the plate back and forth. Incubate cells at 37 ^{o}C and 5 % CO_2. Change medium with 100 µl of DMEM after 6 hours. 24 h after transfection visually test with fluorescent microscope for transfection.

3.3.2.2. INDIRECT STAINING WITH A MIXTURE OF TWO MONOCLONAL ANTIBODIES

GFP transfected cells were indirectly stained with mixture of two monoclonal antibodies in order to validate the pseudo-first order antibody binding model.

Table 5: Materials used for indirect staining with a mixture of two monoclonal antibodies

Material name	*Company name*
96-well **microplate** containing app. 10.000 cells/well. Transfected as seen from Table 6	
Anti-GFP mouse $IgG_1\kappa$ monoclonal antibodies (clones 7.1 and 13.1.), respective dilutions 1:200 and 1:1000 (9)	Roche
Alexa 647 donkey anti-mouse IgG antibodies, dilution 1:1000	Invitrogen
Fluorescence microscope	Leica
8% **PFA** (paraformaldehyde) in DPBS	Sigma-Aldrich
Dulbecco's **PBS** (1x) (without Mg or Ca)	GE Healthcare
NGSB (normal goat serum blocking): 2% goat serum, 1% BSA, 0.1% Triton X-100, 0.05% Tween 200)	Dianova, Roth, Sigma-Aldrich
5 mg/ml **DAPI** stock solution in DMF	Molecular Probes
Aluminum sealing	Costar
High Content Screening Automated Fluorescence Microscope	Leica
1% **BSA** (bovine serum albumin)	Sigma-Aldrich

I. Protocol for indirect immunofluorescence staining with mixture of two monoclonal antibodies (Anti-GFP)

Take GFP transfected 96-well plate out of incubator and observe it under fluorescence microscope. Make sure that only columns 2, 4, 6, 8, 10, 12 are GFP transfected. Other columns are not transfected and serve as various controls. Defreeze PFA and fix cells with 4% PFA solution for 10 min at room temperature by adding 100 µl of 8% PFA per well to existing 100 µl of medium. Rinse 2 times with 100 µl PBS for 10 min each. Defreeze NGSB, block cells with 50 µl per well for 1h at room temperature.

Make a dilution series of Anti-GFP antibody with 1% BSA from 1:40 to 1:20.480 dilutions with 2x incremental dilution. Incubate with 50 µl of Anti-GFP antibody of certain dilution (according to plate plan) for 1 hour at room temperature. Rinse 3 times with 100 µl PBS for 10 min each. Incubate with 50 µl of secondary antibodies Alexa 647 diluted 1:1000 in PBS for 1h at room temperature. Include DAPI stock solution (1:100). Rinse 3 times with 100 µl PBS for 10 min each. Fill wells with 200 µl PBS, seal with aluminum sealing.

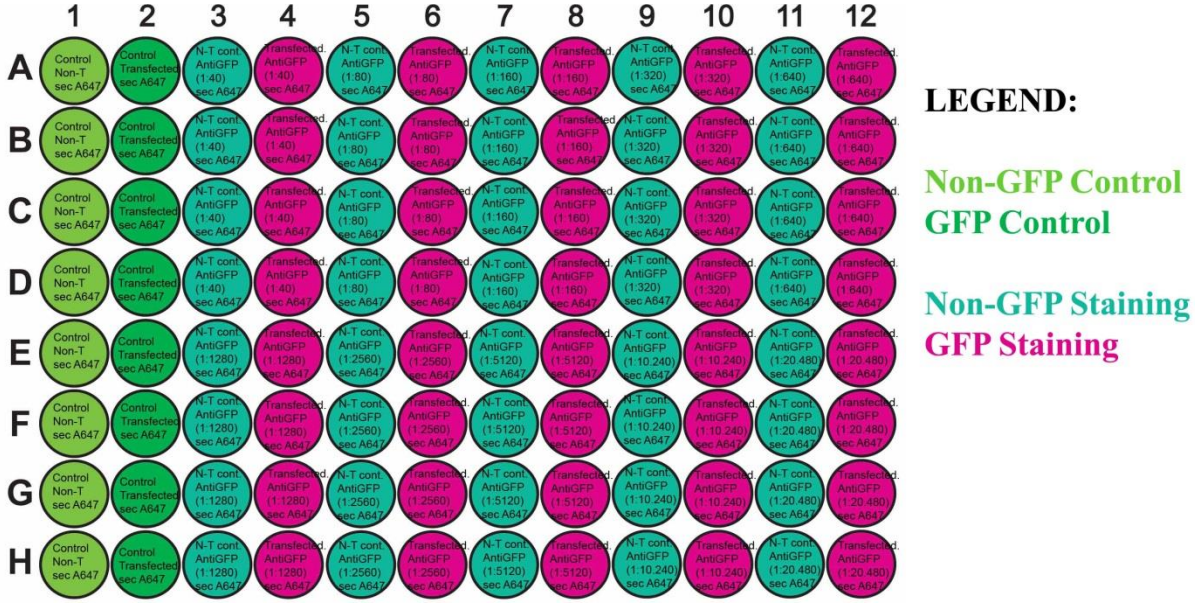
After 10 minutes, scan with high screening automated fluorescence microscope in the following wavelengths (channels):

- CH1: Wavelength of 386 nm for DAPI detection
- CH2: Wavelength of 485 nm for GFP detection

- CH3: Wavelength of 650 nm for Alexa Fluor 647 detection

Extract the data and feed it into computer model to get resulting affinity distribution.

Table 6: Plate plan for indirect staining with mixture of two monoclonal antibodies



3.3.2.3. DIRECT STAINING WITH A MIXTURE OF TWO MONOCLONAL ANTIBODIES

GFP transfected cells were directly stained with mixture of two monoclonal antibodies in order to validate the pseudo-first order antibody binding model and to evaluate potential differences with indirect staining.

Materials used were similar as with indirect staining, with an addition of *Mix-n-Stain CF633 antibody labeling kit* (18) by Biotium, used to label Anti-GFP antibodies with fluorescence marker CF633.

Staining protocol was similar as with indirect staining, with omission of steps 8 and 9. In step 6, DAPI stock solution (1:100) was included. In step 5, starting dilution was 1:20 and 2x incremental dilutions to 1:10.240 were made, in same manner as seen in plate plan in Table 6, omitting secondary A647 marked antibodies.

3.3.2.4. DIRECT STAINING WITH POLYCLONAL ANTIBODY

GFP transfected cells were directly stained with polyclonal antibody in order to validate the pseudo-first order antibody binding model and to evaluate expected differences with monoclonal antibodies.

Materials and staining protocol were similar as in indirect staining with a mixture of the two monoclonal antibodies. The difference was the use of a rabbit polyclonal anti-GFP antibody conjugated with Alexa Fluor 594 and produced by Life Technologies with catalogue number A21312 (8). This was used in lieu of monoclonal antibodies, starting with 1:80 dilution and ending with 1:20.480 dilution. Also wavelength gathered in channel 3 (CH3) was 549 nm for better detection of Alexa Fluor 594 signal. Plate plan was in same manner as plate plan seen in Table 6, with omission of secondary A647 marked antibody.

3.3.3. CROSS-REACTIVITY DETECTION EXPERIMENTS

The purpose of cross-reactivity detection experiments was to determine if antibody cross-reactivity can be determined if present. This was done by experimentally simulate cross-reactive system.

3.3.4.1. CROSS-REACTIVITY DETECTION EXPERIMENT WITH ERK AND pERK

Anti-ERK (extracellular signal-regulated kinase) and anti-pERK (phosphorylated ERK) antibodies were used to target ERK and pERK, respectively. ERK and pERK are protein kinases involved in extracellular signal-regulated kinase cascade, which is a central pathway that transmits the signals from many extracellular agents to regulate cellular processes such as proliferation, differentiation and cell cycle progression. ERK phosphorylation is the key mechanics by which signal is transmitted (23). Anti-ERK (targeting ERK1 and ERK2 protein kinases) and Anti-pERK antibodies were used because most of the cells, including HeLa cell line, contain ERK and pERK to act as an antigen. In addition, the content of ERK is sufficient enough to produce measurable signals at very low antibody concentration. On the other hand, quantities pERK as antigen were stimulated by PMA in order to obtain measurable signals at the same very low antibody concentration.

Table 7: Materials used for cross-reactivity detection experiment with ERK and pERK

Material name	*Company name*
Anti-ERK mouse monoclonal antibody: Anti-Erk1 (pT202/pY204) + Erk2 (pT185/pY187) [MAPK-YT] antibody (ab50011) (19)	Abcam
Anti-pERK mouse monoclonal antibody: Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (E10) Mouse mAb #9106 (20)	Cell Signaling

96-well **microplate** containing app. 10.000 cells/well	
Alexa 647 donkey anti-mouse IgG antibodies, dilution 1:1000	Invitrogen
8% **PFA** (paraformaldehyde) in DPBS	Sigma-Aldrich
Dulbecco's **PBS** (1x) (without Mg or Ca)	GE Healthcare
NGSB (normal goat serum blocking): 2% goat serum, 1% BSA, 0.1% Triton X-100, 0.05% Tween 200)	Dianova, Roth, Sigma-Aldrich
5 mg/ml **DAPI** stock solution in DMF	Molecular Probes
Aluminum sealing	Costar
High Content Screening Automated Fluorescence Microscope	Leica
1% **BSA** (bovine serum albumin)	Sigma-Aldrich
Stimulation 96-well **plate** with glass bottoms	Greiner
PMA, P 8139-1MG	Sigma-Aldrich
Cell medium	

I. Protocol for cell stimulation and staining with ERK and pERK

CELL STIMULATION**:** Dilute PMA in PBS 1:1000 and add 12.5 µl of diluted PMA in each well on stimulation plate. Take cells out of incubator and remove 50 µl of medium from each well and add it to PMA on stimulation plate. Mix 2x up and down and slowly return 50 ml of mixture on stimulation plate back on top of cells. Incubate plate with cells for 30 min at 37^{o}C, 5% CO_2.

Defreeze PFA and fix cells with 4% PFA solution for 10 min at room temperature by adding 100 µl of 8% PFA per well to existing 100 µl of medium. Rinse 2 times with 100 µl PBS for 10 min each. Defreeze NGSB, block cells with 50 µl per well for 1h at room temperature. Make a dilution series of both anti-pERK and anti-ERK antibody with 1% BSA from 1:50 to 1:25.600 dilutions with 2x incremental dilution. Plate the top half of the plate according to plate plan in Table 8 with only anti-pERK or only anti-ERK antibody, by adding 50 µl in each well. Then mix same volumes and dilutions of both antibodies and plate the bottom half of the plate by adding 50 µl in each well. Incubate 1 hour at room temperature.

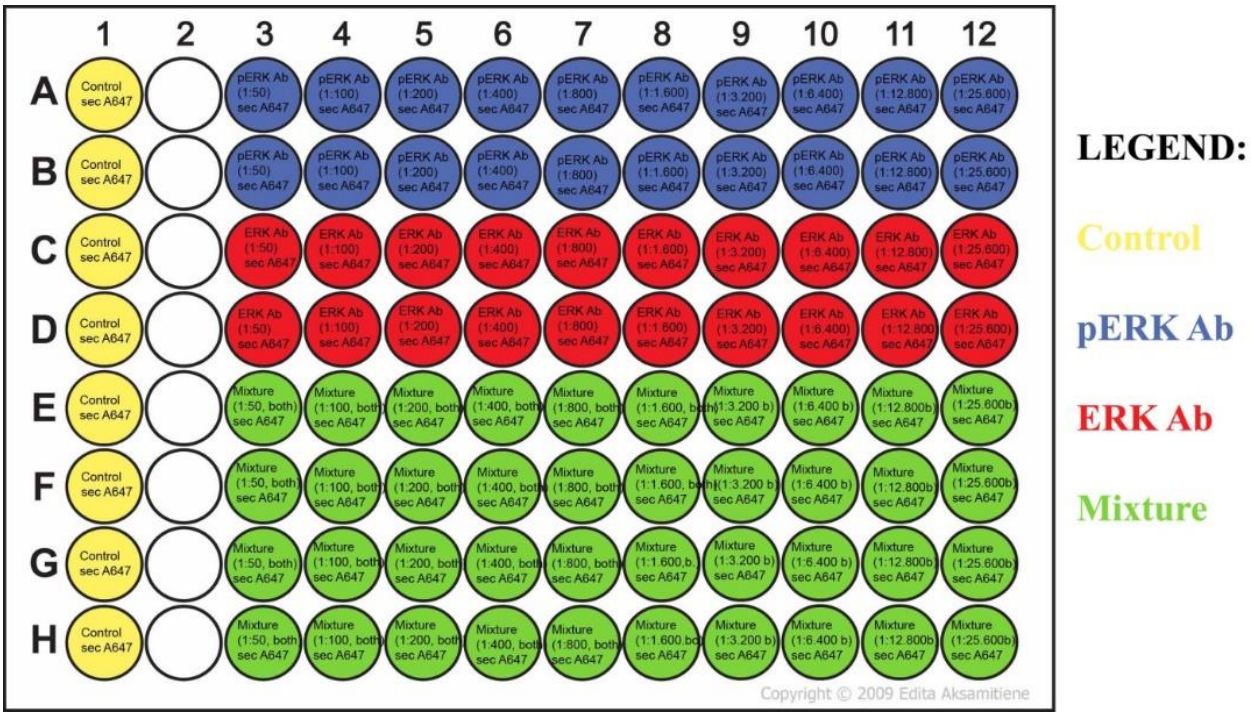
Rinse 3 times with 100 µl PBS for 10 min each. Incubate with 50 µl of secondary antibodies Alexa 647 diluted 1:1000 in PBS for 1h at room temperature. Include DAPI stock solution (1:100). Rinse 3 times with 100 µl PBS for 10 min each. Fill wells with 200 µl PBS, seal with aluminum sealing. After 10 minutes, scan with high screening automated fluorescence microscope in the following wavelengths (channels):

- CH1: Wavelength of 386 nm for DAPI detection

- CH2: Wavelength of 650 nm for Alexa Fluor 647 detection

Extract the data and feed it into computer model to get resulting affinity distribution.

Table 8: pERK and ERK antibody experiment plate plane



3.3.4.2. CROSS-REACTIVITY DETECTION EXPERIMENT WITH pERK AND NF200

Anti-pERK and anti-NF200 (neurofilament subunit with molecular mass of 200 kDa) antibodies were used to target pERK and NF200, respectively.

Materials used were similar as in ERK and pERK experiment, only anti-ERK antibody was replaced with anti-NF200 mouse monoclonal antibody (Monoclonal Anti-Neurofilament 200 (Phos. and Non-Phos.) antibody, clone N52, N0142, Sigma-Aldrich).

I. Protocol for cell stimulation and staining with pERK and NF200

Cell stimulation and most of staining protocol were similar as in experiment with ERK and pERK, except in step 5 where the following changes were made:

- Dilution series of anti-pERK antibody started with 1:50 and went with 2x increments to 1:25.600
- Dilution series of anti-NF200 started with 1:400 and went with 2x increments to 1:204.800

4. RESULTS

Results were divided in two parts. In computer simulations part (see 4.1) the model was programmed and parameters of created model to determine cross-reactivity were evaluated. In wet lab experimentation results (see 4.2) the dynamics of antibody-antigen binding were validated and the possibility of the method to determine cross-reactivity was put to the test.

4.1 COMPUTER SIMULATIONS

The purpose of computer simulations was to test if pseudo-first order model is able to distinguish between different antibody binding sites. This was done via calculating the antibody affinity distributions based on binding data. In this section, all binding data are fictive and simulated by pseudo-first order model.

Computer simulations served as a rough estimate of antibody binding situations. The basic experimental concept simulated was comprised of introducing a series of dilutions of free antibody solution on top of fixed cells. These cells are comprised of specific epitopes and a myriad of different cell parts and products, all potentially viable for antibody binding. Antibodies were directly or indirectly stained by a fluorescence marker, specific signal of which was measured after binding process had taken place and all the free antibody was washed away. Hence the measured signal was directly proportional to the amount of bound antibodies.

4.1.1 LINEAR COMBINATIONS OF DIFFERENT DISCRETE BINDING SITES

Applying pseudo-first order model, we wanted to demonstrate that each specific antibody affinity distribution forms a specific signal curve. Thus by calculating the signal curve of an antibody, antibody's affinity distribution can be theoretical obtained with the use of equation *7*. Furthermore, by visualizing signal curves produced by discrete binding sites or linear combinations of these, the observable difference between these lines is seen from Figure 2.

Dark blue and yellow curve represent binding curve of a single binding site with Kd = 0.1 nM and Kd = 1000 nM, respectively. Green, red, light blue and violet curve lines are linear combinations of several different binding sites (Kd values) as seen from the legend in Figure 2. Net binding capacity S_{max} = 100 is evenly distributed between all Kd values generating the signals curves.

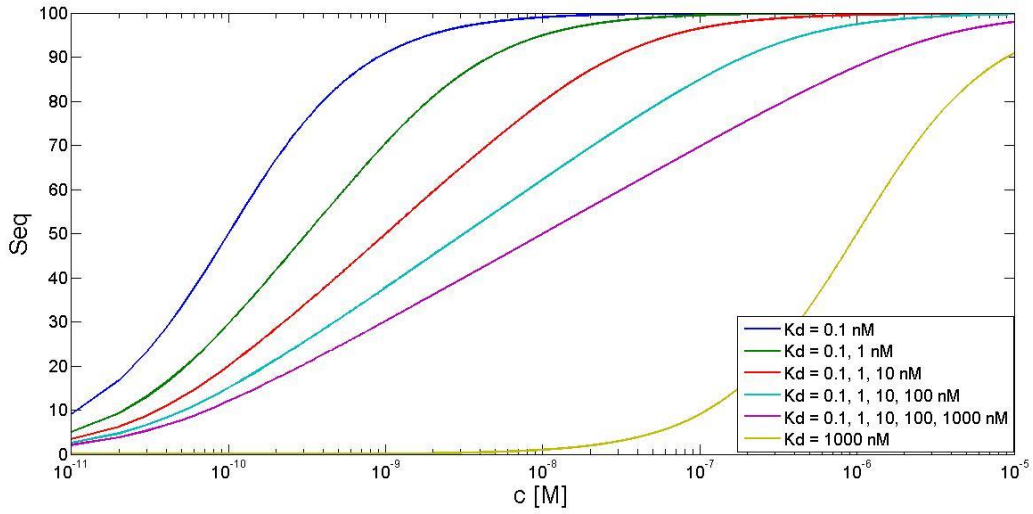


Figure 2: Linear combination of discrete binding sites

By looking at dark blue and yellow curve it is obvious that distinction between single binding sites with great enough antibody affinity difference will be detectable. In this case with great Kd difference (10^5), it is clear that at antibody concentration c = 10 nM, the binding site with Kd = 0.1 nM is almost completely saturated, while binding site with Kd = 1000 nM show very little binding. However, if the affinity difference between single binding sites is decreased, distinguishing between binding curves is increasingly difficult.

Green, red, light blue and violet curve lines represent a situation where antibody has many different binding sites with different affinities. Linear combinations calculated binding curves for two or more binding sites generate the green, red, light blue and violet curve lines in Figure 2. Even though additional binding sites have an affinity increment of only 10, it is obviously still possible to distinguish between curves with 2, 3, 4 or 5 different binding sites. What is more, curve with single binding site is clearly distinguishable from curve with two in affinity very similar binding sites. The best place to see this difference is at low antibody concentrations. At c = 10^{-11} M, the signal from single binding site curve (dark blue line) is approximately 100% higher than the signal from two binding sites curve (green line). As antibody concentration is increased in vicinity of Kd, the absolute difference in saturation is at first increased, yet relative percentage difference is persistently decreasing after Kd value, as seen in Figure 2.

It is also noteworthy to look at the steepness of the curves. It is evident that single binding site curves are the steepest, producing the sharpest increase and most sigmoidal shape of the curve. On the other hand, the more binding sites are added to the curve generation, the less steep is the binding curve, eventually becoming less sigmoidal in shape and becoming more linear.

4.1.2. SIMULATED THREE BINDING SITES DISTRIBUTION WITH 1% NOISE

In the end we will try to extract the Kd value distribution from measured binding studies. Necessarily, the respective experiments will test discrete concentrations. To test if Kd value distribution can be measured from discrete measured datasets, an inverse problem of generating binding curves via simulated (c, S_{eq}) data points from proposed antibody affinity distribution was solved and programmed in MATLAB (*code in Appendix*).

In the first case three normally distributed binding sites were proposed. Program was written and its ability to obtain antibody affinity distribution from (c, S_{eq}) dataset was put to the test. Average affinity of each binding site was separated by a factor of 10. Binding capacity ratios for proposed binding sites with Kd values 1, 10, 100 nM were 1:10:100, respectively. 10 measurements were simulated. A randomized constant noise of 1% in average for each simulated data point was added.

This simulation is presented in detail to get a better understanding of how mathematical model was implemented in the program. Computer simulation is comprised of the following steps:

1. Assume binding affinity distribution (3 normally distributed binding sites in this case).

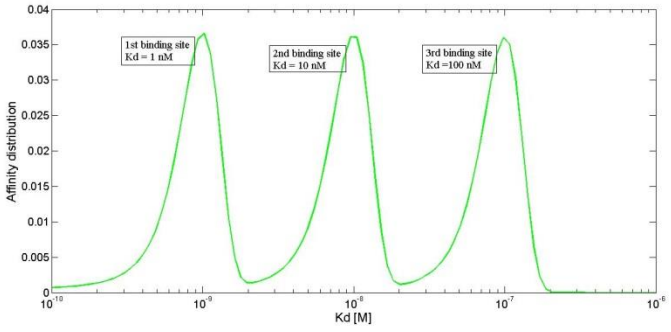


Figure 3: Assumed 3 binding site affinity distribution

2. Based on assumed distribution and using equation *7*, calculate 10 data points of form (c, S_{eq}) and add 1% of normally randomized (measurement uncertainty) – these are simulated data points.

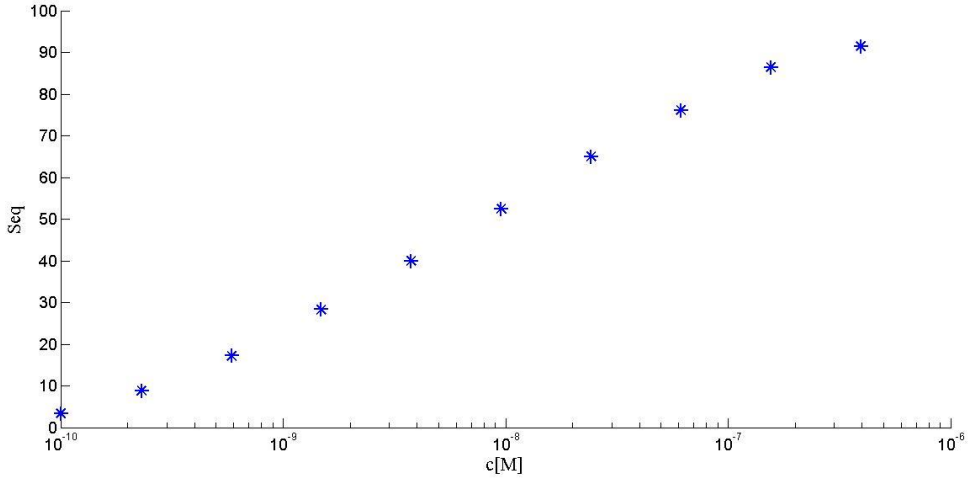


Figure 4: 3 binding site data simulated with 1% noise

3. Fit the simulated data points with linear combination of 100 signal curves (based on equation *7*) with different Kd values.

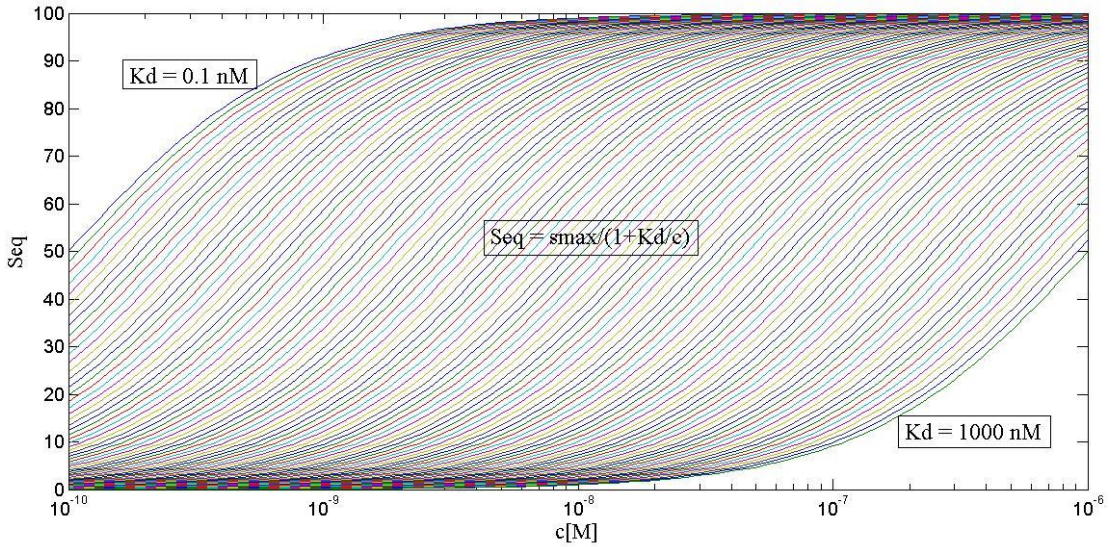


Figure 5: Functions used for fitting the simulated data points

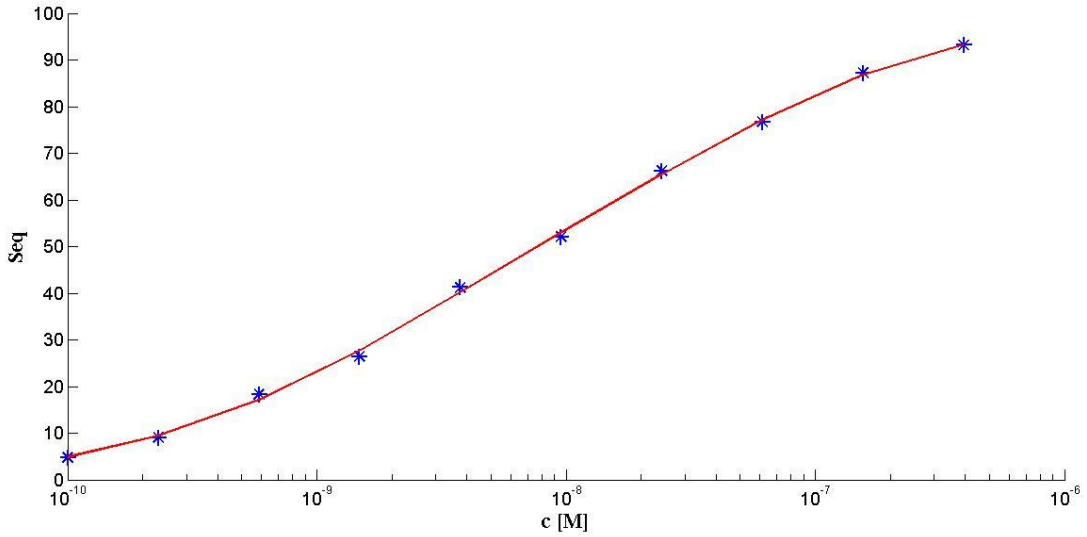


Figure 6: Best fit curve for 3 binding sites data simulated with 1% noise

4. Using the coded program, calculate and visualize binding affinity distribution. Signal curves that best fit the simulated data will attain highest values, as seen by red line in Figure 7.

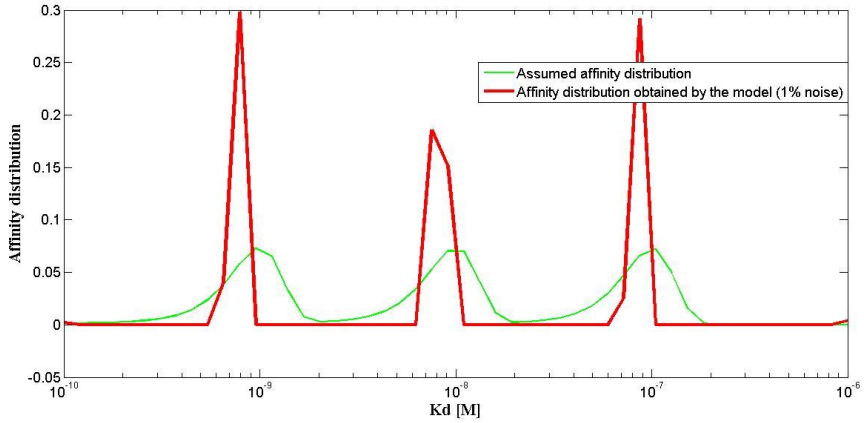


Figure 7: Resulting affinity distribution with 1% noise

If assumed affinity distribution from step 1 (green line) is captured in modeled affinity distribution in step 4 (red line), then model and simulation are successful in determining the binding affinity distribution of an antibody. From Figure 7 we see that the location of the resulting binding peaks (red line) roughly coincides with assumed affinity distribution (green line). This is true in shown simulation with 1% noise in mind. However, noise is randomly distributed, making each simulation unique due to unique noise signature. In Figure 8, we see

that even after many simulations runs the majority of produced binding sites concentrate at Kd values where in fact binding sites really are. Nonetheless, we also see peaks created in between assumed binding sites. Furthermore, a single simulation sometimes produces 4 or 5 peaks instead of 3 expected peaks. On the other hand, peaks are also very narrow in comparison with assumed binding sites peaks.

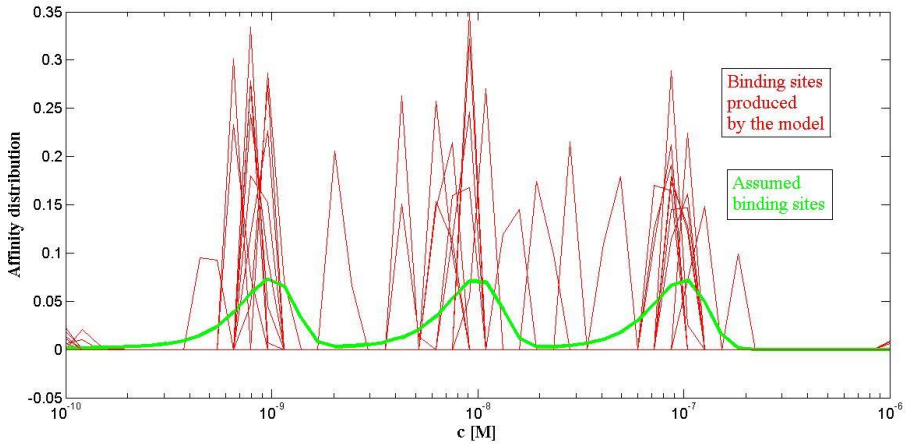


Figure 8: 10 simulation runs for 'Resulting affinity distribution with 1% noise'

Extra peaks are produced as a result of fitting algorithm overfitting the simulated dataset. When algorithm is trying to incorporate a more deviated data point, it usually takes a single Kd value signal curve to capture it, which in turn produces an extra peak. In order to eliminate extra peaks, we introduced a stabilizing regularization parameter in the minimization part of the algorithm. In any biological experiment some measurements will deviate significantly more than others. Current algorithm makes a best fit curve heavily based on these deviated data points, producing extra peaks. Regularization in fitting favors smoothness of the curve in lieu of fitness by decreasing the influence of heavily deviated data points. Result is a decrease in number of peaks as well as increase in broadness of peaks as seen in Figure 9.

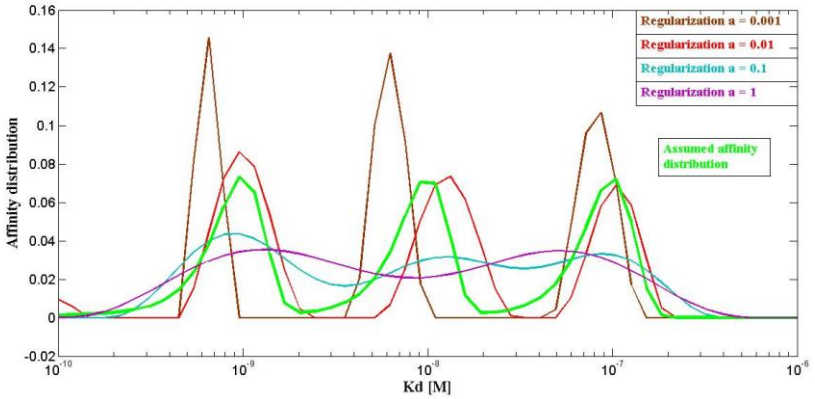


Figure 9: Resulting affinity distribution with regularization

Regularization, or in other words the influence of heavily deviated points to curve fitting, can be manually adjusted by changing the factor 'a' in the code. Increasing the regularization broadens the peaks as seen in Figure 9. It also decreases the number of peaks, which represents a difficulty. Regularization is not selective and will, if increased high enough extent, not only eliminate noisy unwanted peaks but also valid binding site peaks as seen from Figure 9 in case with a=1(violet line) - we only get 2 peaks out of 3 valid peaks. Based on this simulation, the amount of regularization that reflects the assumed affinity distribution (green line) the best is in the case of a=0.01 (red line). This 'a' value was used throughout the experimental part.

Overall, from Figure 10 it is clearly seen that 3 binding site curve is distinguishable from any single binding site curve. The model is well capable of capturing the affinity distribution in the case with 1% noise.

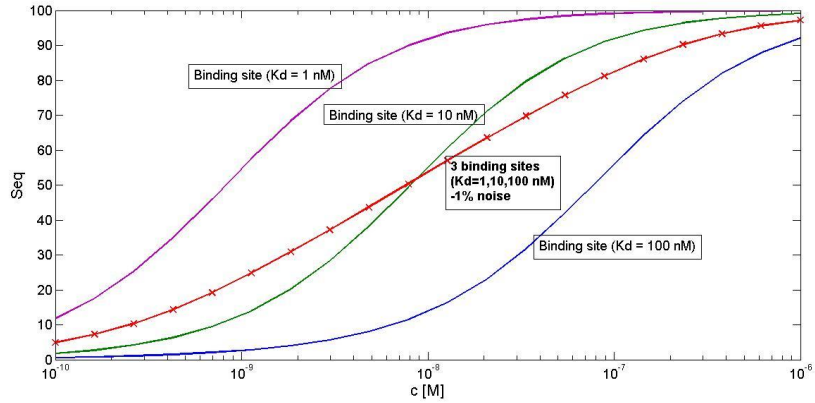


Figure 10: Three binding sites with 1% noise

4.1.3. SIMULATED THREE BINDING SITES DISTRIBUTION WITH 10% NOISE

In order to more accurately simulate the realistic experimental situation, a 10% noise was added to the simulated signal measurements before calculating out of these discrete data the underlying respective Kd distribution.

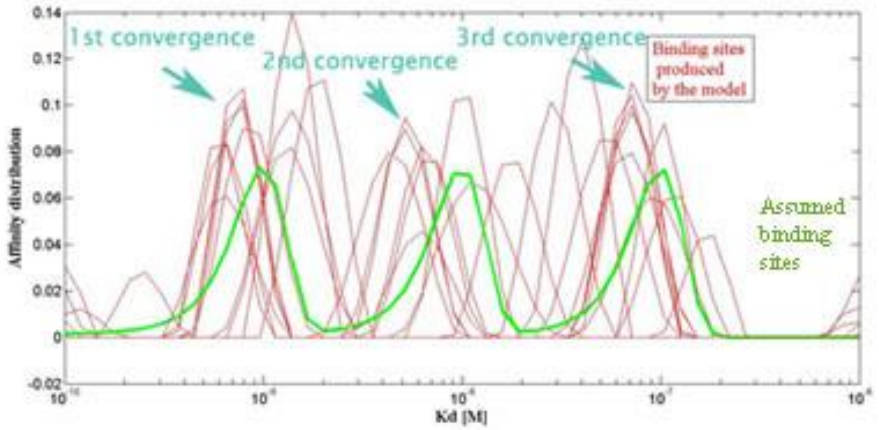


Figure 11: 10 simulation runs for 'Resulting affinity distribution with 10% noise'

As seen in Figure 11 it is still possible to see the convergence of peaks into three separated binding sites, despite 10% of measurement uncertainty. However, due to higher uncertainty we also observed peaks that did not correspond with assumed binding sites. Fortunately they do not significantly interfere with majority of peaks. An inclining to lower Kd values was observed at all three peak convergences. This phenomenon was most probably present due to high uncertainty level, as it does not appear in the case with 1% noise.

4.2. WET LAB EXPERIMENTATION

In the modeling part the affinity distribution of an antibody was assumed and the ability of computer simulation to capture affinity distribution was evaluated. In wet lab experimentation we were now testing, if we could calculate the underlying Kd distribution also from real measured data.

4.2.1. GFP TRANSFECTION AND EVALUATION OF ANTI-GFP ANTIBODIES AFFINITY DISTRIBUTION

GFP TRANSFECTION

In order to create a defined antibody target, HeLa cells were transfected with GFP. Two different plasmids were used and three different plasmid-to-Lipofectamine ratios were evaluated, figuring out which transfection condition results in best transfection.

Highest GFP transfection was achieved using pEGFP-C1 plasmid in 1:2 ratio with Lipofectamine. Plasmid pEGFP-N3 was also successful in achieving transfection, but the amount of GFP transfected was considerably lower than with pEGFP-C1. GFP transfection was visually observed and evaluated.

In cases of both plasmids, 1:2 ratios with Lipofectamine were most successful. 1:1 ratio proved to achieve high amount of transfection as well, while 1:4 ratio was less successful. In all controls a lack of transfection was observed.

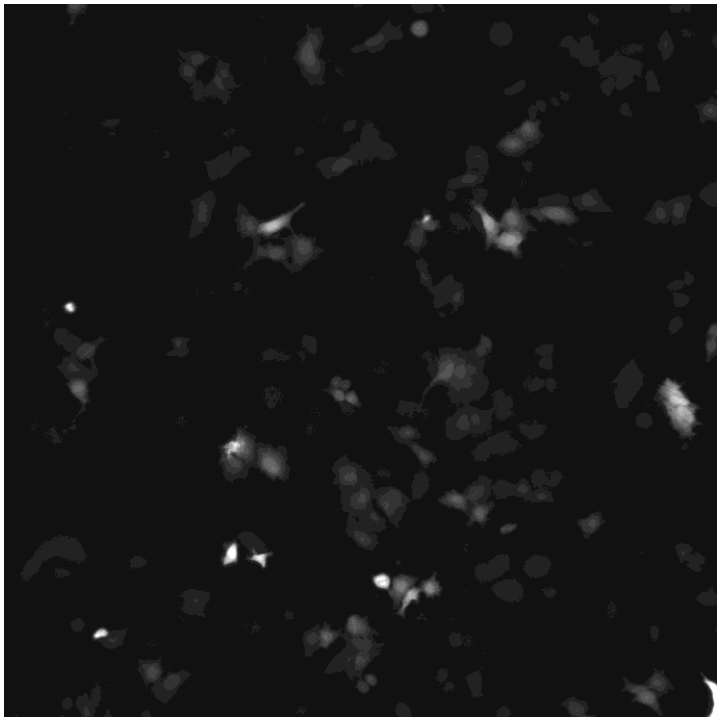


Figure 12: pEGFP-C1 with 1:2 ratio to Lipofectamine

DIRECT STAINING WITH A MIXTURE OF TWO MONOCLONAL ANTIBODIES

In order have a defined antibody affinity distribution, a mixture of two monoclonal antibodies were used. Roche's Anti-GFP antibodies were used to target transfected GFP. By

having two monoclonal antibodies, we were able to roughly determine the expected affinity distribution. There are 3 possible solutions in terms of number of binding sites:

1. A single binding site because either:
   a) Kd values of both binding sites are very close together and model isn't able to distinguish between them.
   b) Both antibodies compete for same binding site – competitive binding doesn't produce a specific Kd values of each binding antibody, but it rather creates a single average Kd value of all competing antibodies.
2. 2 binding sites – one for each monoclonal antibody.
3. 3 or more binding sites due to monoclonal antibody cross-reactivity.

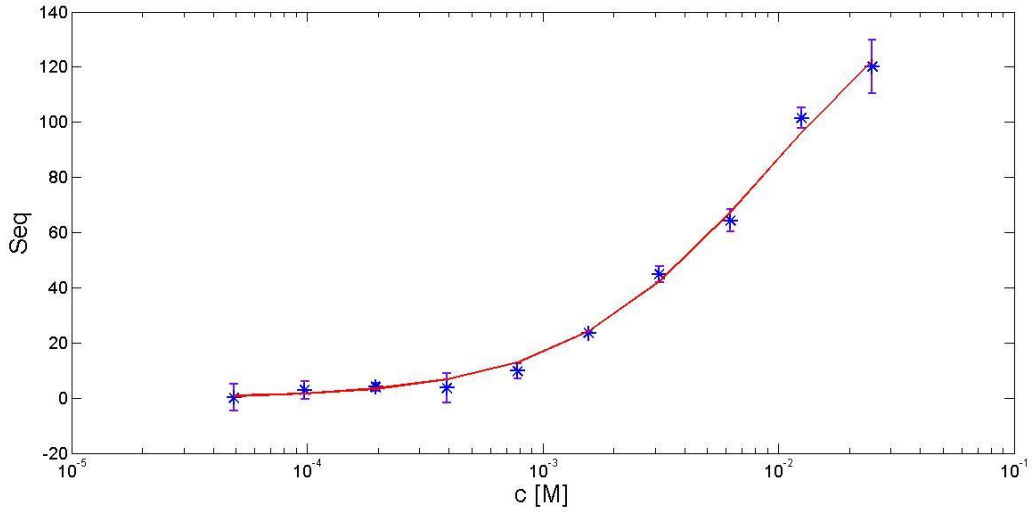


Figure 13: Two anti-GFP monoclonal antibodies (direct staining)

Data measurements are represented in Figure 13. The concentration c on the x-axis is the antibody dilution, since antibody concentration are unknown (not provided by manufacturer). However, dilution series still captured the relative concentration ratio between the samples. In such a way antibody dilution served as parameter of concentration throughout this paper.

As seen from Figure 13, measurements were fitted very well with linear combination of pseudo-first order model curve. This was the first experimental result confirming that antibody binding indeed follows such binding kinetics. Most of the error bars were captured by the curve, which is a good result for biological experiment with theoretically generated binding curve. GFP signal produced a very little (< 5%) variation in GFP well-per-well content.

Figure 14: HeLa cells nuclei stained by DAPI (left), GFP signal (middle), Antibody A647 staining signal (right)

As seen from the Figure 13, the highest concentration of antibodies (the lowest dilution of 1:40) was still unable to achieve saturation plateau. To reach that point, a higher concentration of very expensive antibodies could be used, but we have taken a different approach. Namely, if we can roughly predict the maximum S_{eq} (S_{max}), and input it into the model, we can derive such a curve fit as seen in Figure 14, where a S_{max} of 200 was input. This was determined by visually detecting the curve's inflection point and/or by observing the result of minimization equation – the S_{max} at which result was the lowest is the most valid prediction.

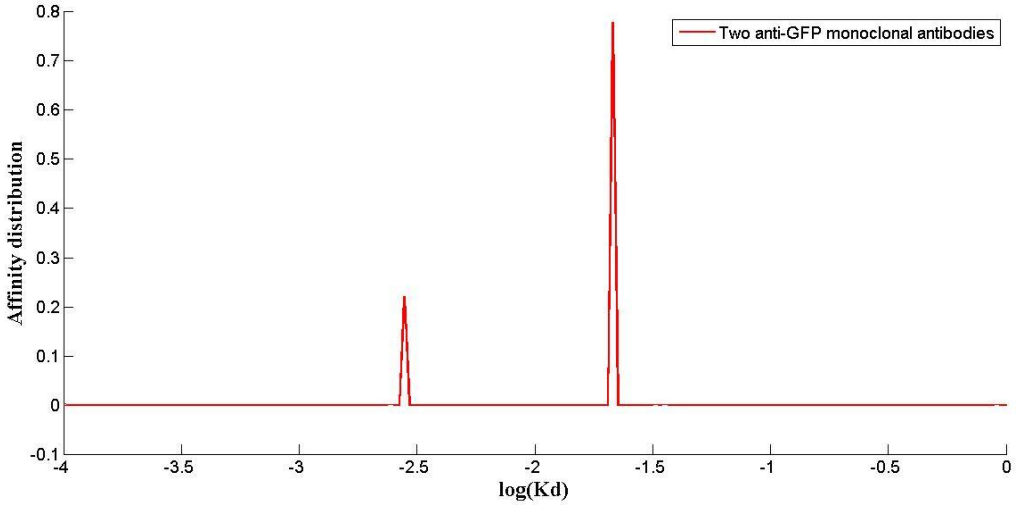


Figure 15: Two anti-GFP monoclonal antibodies affinity distribution (direct staining)

The calculated affinity distribution of the two anti-GFP monoclonal antibodies based on fitting curve in Figure 13 is represented in Figure 15. The calculated affinity distributions indeed result for two mixed monoclonal antibodies in two distinct distribution peaks separated

in Kd value by a factor of about 10. It is important to understand that this Kd values are only known in relative terms, based on antibody dilution.

Beyond the ability to detect that the detected signal was based on two different Kd values this result also indicates how the two different antibodies bind to GFP. The result indicates that the two monoclonal antibodies target different parts of GFP and do not bind competitively in relation to each other. It also means that each has a distinct Kd value describing individual binding kinetics.

As seen from Figure 15, peak at higher Kd is equal to 3-4 times the height of peak at lower Kd. This indicates that binding capacity of higher Kd binding site is 30-40 times the binding capacity of binding site at lower Kd which derives from the logarithmic scale on the x-axis. Mathematically speaking, for two peaks in affinity distribution plot separated by a Kd value of factor 10, the height of both discrete peaks will be equal only if capacity ratio is 10:1 in favor of peak with lower Kd value. We see that in order for binding site at higher Kd to be recognized in affinity distribution plot as a peak, it has to have appropriately higher binding capacity in relation to peak with low Kd.

It is worth mentioning that S_{max} value is assumed. Around chosen value of 200 (150-300) did not radically affect two peak affinity distribution. However, the number of curve signals used for curve fitting did influence affinity distribution substantially. We used 200 curve signals with 200 distinct Kd values to fit the curve in Figure 13. When lower curve signals were used (50), a single binding site affinity distribution was produced. At higher number of signal curves (300), a multiple binding sites affinity distribution was obtained. Nonetheless, curve fitting was poor in cases with 50 and 300 fitting signal curves, while the curve fit was best in the range of 150-220 fitting signal curves, where no variations in number of binding sites were seen and only little variations of relative Kd values and peak heights ratios were observed.

INDIRECT STAINING WITH A MIXTURE OF TWO MONOCLONAL ANTIBODIES

We evaluated if the technique of staining interferes with affinity distribution, namely if exactly the same number of peaks are discovered in experimental setting regardless of staining technique being direct or indirect. Theoretically, secondary antibody could target something other than primary antibodies, which would produce a measurable signal and distort calculated

Kd distribution. In Figure 16 the results of indirect staining with the same antibody mixture as in the previous case are presented.

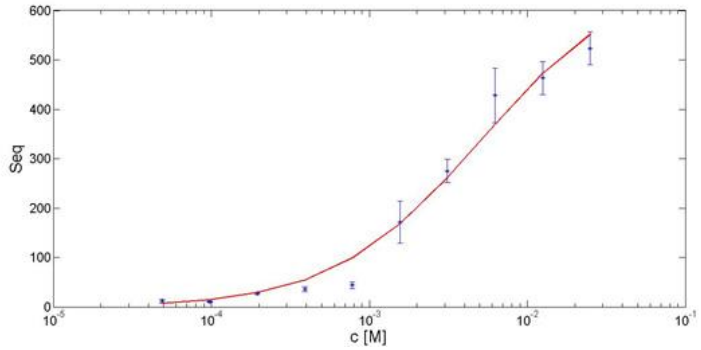


Figure 16: Two anti-GFP monoclonal antibodies (indirect staining)

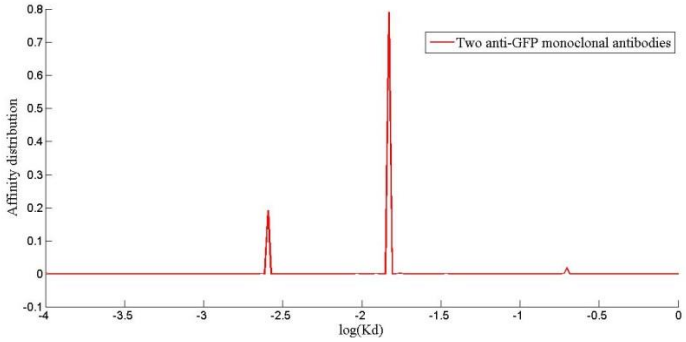


Figure 17: Two anti-GFP monoclonal antibodies affinity distribution (indirect staining)

From Figure 16 we see that fit with pseudo-first order model curves is good, but not as ideal as in the case with direct staining. On the other hand, signal intensities are stronger, which is a consequence of indirect staining amplifying the signal. As a result absolute measurement error is more substantial.

Similar to previous case, we obtained two major peaks in affinity distribution in Figure 17. Again the Kd difference between them was approximately of factor 10 and height ratio of 4:1 in favor of peak at higher Kd value. In addition, Kd values of peaks were also very similar to the Kd values of peaks in the case with direct staining. All of these things were expected because same antibody mixture was used. However, an additional minor peak was seen at very high Kd value. Even with the use of heavy regularization, the peak persisted thus making it a viable peak. This additional peak might be a result of non-specific binding of secondary antibodies to some target in HeLa cell system. Concentration of secondary antibodies used

was substantial enough for a low affinity binding with enough binding capacity to form a measurable signal.

As to the modeling parameters chosen, curve was fitted with 200 signal curves with distinct Kd values and S_{max} was set to 700. Similar variations as in the case with direct staining were experiences if number of signal curves for fitting was different or S_{max} was varied.

DIRECT STAINING WITH POLYCLONAL ANTIBODY

What happens if we combine more and more distinct antibodies? Will we still be able to calculate distinct affinities? One example of multiple combined distinct antibodies are polyclonal antibody sera. Thus, similar as in the cases with monoclonal antibody, a polyclonal anti-GFP antibody was tested. As oppose to a somewhat determined affinity distribution of monoclonal antibodies, polyclonal antibody sera contain many different antibody clones each targeting a specific epitope on GFP with a specific affinity. In terms of number of binding sites the following possible solution were viable:

1. A single binding site because all antibodies compete for same binding space – structural competitive binding doesn't produce specific Kd values of each binding antibody, it rather results in a single weighted average Kd value of all competing antibodies.
2. A multiple binding sites because many antibodies of polyclonal antibody bind non-competitively to specific epitopes.

Results are presented in Figures 18 and 19.

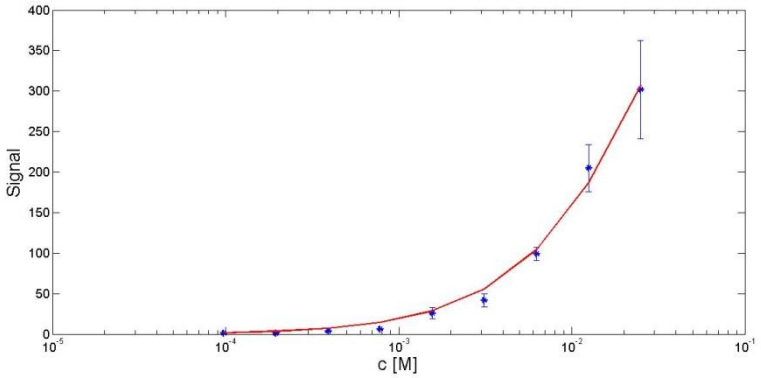


Figure 18: Anti-GFP polyclonal antibody (direct staining)

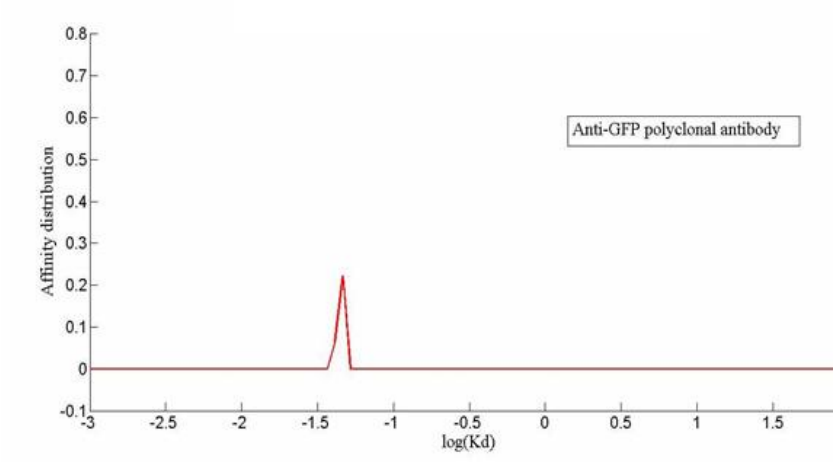


Figure 19: Anti-GFP polyclonal antibody affinity distribution

From Figure 18 it is evident that pseudo-first order model validly describes the binding of polyclonal antibody to GFP as antigen. However, expected multiple binding site affinity distribution was not attained. On the contrary, distribution peak resulting in Figure 19 suggest a single binding site. Interestingly, the single peak in Figure 19 is broader than previously obtained peaks in the case of monoclonal antibodies.

Position of affinity peak in Figure 19 is at higher Kd than in the case of monoclonal antibodies. This indicated lower affinity of polyclonal antibody. However, despite lower affinity, signals seen from Figure 18 tend to have roughly comparable strengths at signals from two indirectly stained monoclonal antibodies. This indicated that binding capacity of polyclonal antibody is substantially higher than that of monoclonal antibodies, which is also an expected result.

Modeling parameter S_{max} = 1500 was chosen. Increasing S_{max} even more (even up to 10000) did not significantly affect the affinity distribution. Number of signal curves used for curve fitting was 200. Lower numbers (50) have proven to be insignificant in affecting affinity distribution, while higher numbers, even 250 produced an additional peak. However, the level of fitness above curve number of 200 was lower than at 200.

4.2.2. CROSS-REACTIVITY DETECTION EXPERIMENTS

Cross-reactivity detection experiments were devised. The basic experimental design was comprised of taking two monoclonal antibodies, measuring signal curve of each of them and then of both of them together. We thus calculated three different Kd distribution – one for the

first antibody, the second for second antibody, and the third for the mixture of first and the second antibody. If two individual signals give the same net affinity distribution as combined signal, then the method is valid for capturing the underlying affinity distribution. On the contrary to the first set of wet lab experiments, here we were sure what kind of affinity distribution the individual antibodies had because we measured it first, whereas in the first set the nature of antibody affinity distribution was a rough estimate. With this approach we demonstrated that cross-reactivity can be detected in the following way - the idea is that if we had a single monoclonal antibody, which would bind to two targets, we could see that antibody in fact was bound by two, and not one, targets from affinity distribution – thus making a monoclonal antibody cross-reactive in that environment.

CROSS-REACTIVITY DETECTION EXPERIMENT WITH ERK AND pERK

In the first experiment antibody targets chosen were ERK (extracellular signal-regulated kinese) and pERK (phospho-ERK). pERK is phosphorylated derivate of ERK. Both are known to be present in HeLa cells (27). PMA (Phorbol 12-myristate 13-acetate) was used to stimulate ERK and pERK expression (28) in order to increase targets content and attain stronger fluorescence signals. ERK and pERK mouse monoclonal antibodies were used to target ERK and pERK. Alexa 647 donkey anti-mouse antibodies were used as secondary antibodies implementing an indirect staining technique.

Similar as in previous experiments, fluorescence signals at different dilutions of three antibody species were measured:

1. pERK monoclonal antibody.
2. ERK monoclonal antibody.
3. Mixture (pERK and ERK combined).

Data points were fitted using pseudo-first order model and resulting affinity distribution were obtained.

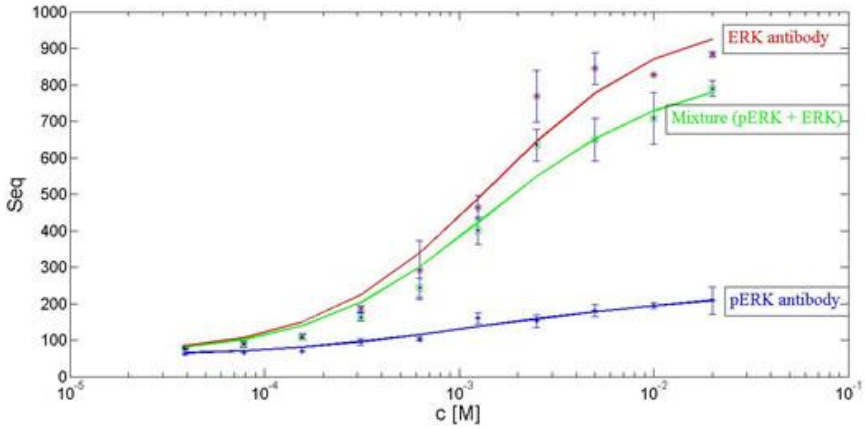


Figure 20: pERK and ERK antibody

In Figure 20 all the signal curves are well fitted with pseudo-first order model. Surprisingly, the mixture of both antibodies does not produce the strongest signal. ERK signal is the strongest, while pERK signal seems to be weak but is nonetheless of measurable intensity.

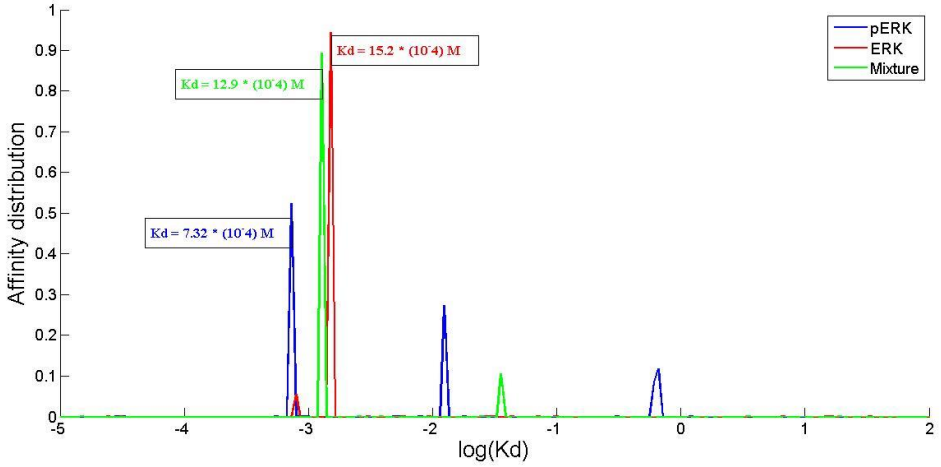


Figure 21: Affinity distribution for pERK, ERK antibody and pERK-ERK Mixture

Judging from Figure 21, we observed that ERK antibody has a major single binding site expected of a monoclonal antibody. However, at higher affinity we see a minor peak corresponding with major pERK peak. Nonetheless, the height of major ERK peak is substantial in relation to high affinity ERK peak, indicating that, given small difference in Kd value, it possesses high binding capacity and is thus responsible for majority of the signal produced by ERK antibody.

On the other hand, pERK antibody produced 3 distinct peaks, which is unexpected of monoclonal antibody. The peak with lowest Kd seem to be dominant, having at least 10 times higher affinity then second peak and 1000 times more than peak with highest Kd. In comparison with ERK antibody peaks, pERK binding sites have a considerably larger affinity difference between them.

Mixture of two antibodies produced two distinct peaks. Major high affinity peak is located between major ERK and pERK peaks, while minor low affinity peaks is located between two pERK minor peaks. At first sight it seems that affinity distribution of individual antibodies is not captured by affinity distribution of the mixture of both of them. However, if we presume that ERK and pERK antibodies bind competitively due to structural interference, these results seem reasonable. In such a case, the Kd value of major mixture peak is a weighted-average of both pERK and ERK peaks. This would also explain why signal curve from ERK in Figure 20 is higher than signal curve from the mixture – mixture's Kd value is higher (as a result of averaging) that ERK's Kd value.

As for modeling parameters, 200 signal curves were used and the following S_{max} values – S_{max} (ERK) = 930, S_{max} (pERK) = 220, S_{max} (mixture) = 820.

CROSS-REACTIVITY DETECTION EXPERIMENT WITH pERK AND NF200

In order to eliminate the possibility for structural interference between binding antibodies, two completely separate targets – pERK and NF200 (neurofilament subunit with molecular mass of 200 kDa) - were chosen for cross-reactivity detection experiment.

HeLa cells are known to contain NF200. (31) Stimulation with PMA was used to amplify pERK content and signal. Already known anti-pERK antibody was used and anti-NF200 mouse monoclonal antibody. Staining was indirect, with previously used secondary antibodies marked with A647, and identical signal generation and quantification same as in experiment with ERK and pERK.

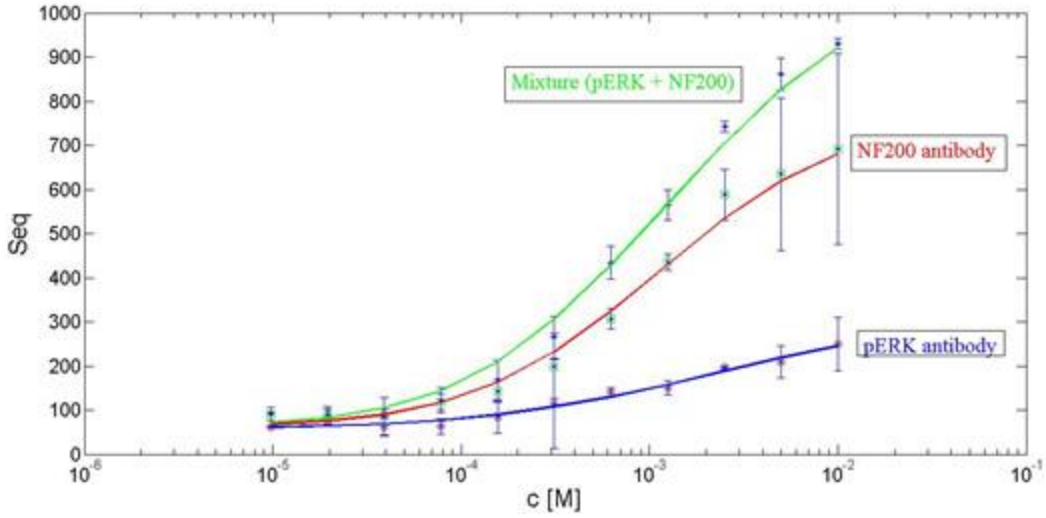


Figure 22: pERK and NF200 antibody and pERK-NF200 Mixture

From Figure 22 we can see that mixture of both antibodies produced the strongest signals as expected. pERK antibody signal curve is very similar to the one in previous experiment. It is obvious that if we add signal curve of NF200 and pERK, the resulting signal is very similar to signal curve of the mixture.

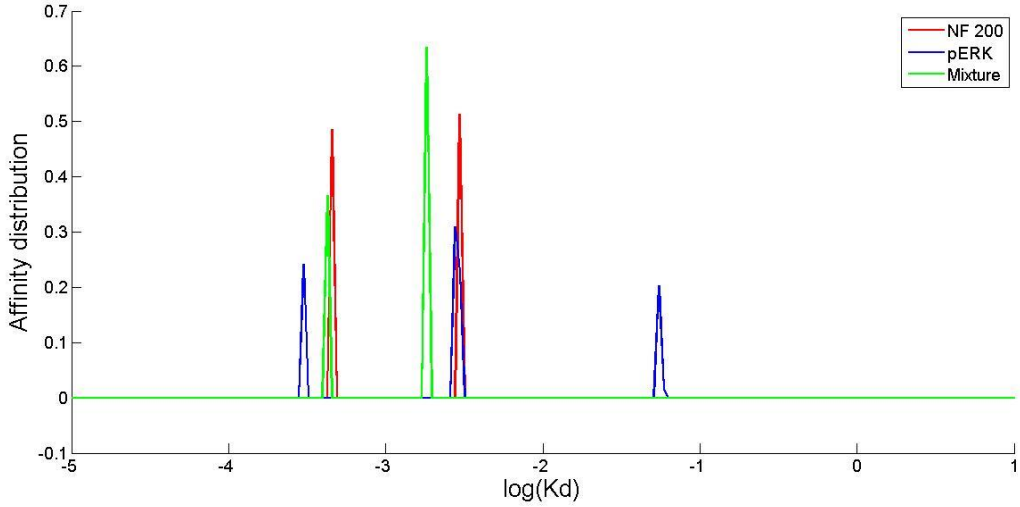


Figure 23: Affinity distribution of pERK, NF200 antibody and pERK-NF200 Mixture

Affinity distribution for pERK antibody in Figure 23 consists of three peaks, which is the same result as in previous experiment. Moreover, the relative difference between peaks is also similar. The relative affinity ratio between peaks, however, is inconsistent with previous

experiment – In Figure 23 the highest peak is at middle affinity, while the highest peak in previous experiment was at highest affinity.

On the other hand, NF200 produced two distinct peaks as seen from Figure 23. Peaks are of roughly same height, difference in Kd value is approximately of factor 10.

The mixture's affinity distribution consists of two distinct peaks. Roughly speaking, high affinity peak is overlapping high affinity peaks of both pERK and NF 200 antibody as expected. Low affinity peak is in proximity of NF200's low affinity peak and pERK's middle peak. Low affinity pERK binding peak, however, is not captured by affinity distribution of a mixture. From Figure 22 it is clear that major antibody creating mixture's signal is NF200 which is roughly captured by mixture's affinity distribution. On the other hand, pERK's influence is minor and only two higher affinity peaks of its three peaks can be suggested to be captured by mixture's affinity distribution.

As for modeling parameters, 200 signal curves were used and the following S_{max} values – S_{max} (pERK) = 270, S_{max} (NF200) = 700, S_{max} (mixture) = 1000.

5. DISCUSSION

To put immunostaining assurance to the test, antibody specificity was characterized based on their affinity distribution in a complex HeLa cell environment. The detection of antibody cross-reactivity a new method was designed based evaluation of antibody affinity distribution results. To achieve this, mathematical model of antibody binding was designed and tested via computer simulations, followed by a wet lab experimentation using standard laboratory immunofluorescence techniques.

5.1. MATHEMATICAL MODELING AND COMPUTER SIMULATIONS

Pseudo-first order kinetics model was chosen as a fundamental dynamics principle governing the antibody binding. Equilibrium points of such a binding were chosen to evaluate antibody affinity distribution via constructed mathematical model. In order to solve a problem of applying the model to acquire antibody affinity distribution from equilibrium points from series of antibody dilution, numerical programming approach was used and computer simulations were devised based on created program.

In Figure 2 we demonstrated that, if pseudo-first antibody binding model holds true, it is possible to differentiate between affinity distributions (namely having different combinations of binding sites) of different antibodies based solemnly on signal curves. Equilibrium points forming these signal curves can be experimentally acquired by staining a biological system with different concentrations of investigated antibody and measuring the fluorescent signal coming from bound antibodies.

MODELING PARAMETERS

To evaluate the model, parameters affecting the calculated antibody affinity distribution were identified and investigated. Major parameter is the **amount of noise** (measurement uncertainty) in determination of antibody binding equilibrium points. As seen from Figures 7 and 8, a 1% normally randomized constant noise was not significant in obtaining the expected affinity distribution. On the other hand, a 10% normally randomized constant noise observed in Figure 11 posed a problem. Even though we observed 3 peak convergences, they were nonetheless all moved to higher Kd values as expected. This is a direct results of the manner noise was applied to equilibrium data points in computer simulations – namely low signal

points were more affected by such a noise than high signal points. Nevertheless, this effect is not relevant for workings of the model on real-time experimental data, where noise is not added by the model itself, but is inherent in the measurements.

By applying random noise, additional unwanted binding sites or peaks in affinity distributions were formed as seen in Figure 8. This is a direct result of fitting algorithm overfitting the data points, which was resolved with the introduction of regularization into fitting algorithm. Regularization favors smoothness of the best fit line over fitness by devaluing the effect of more deviating data points. The **amount of regularization** added, signified by factor 'a', is a parameter in acquisition of antibody affinity distribution. As seen from Figure 9, the amount of regularization that offered the best results was at a=0.01, a constant parameter that was used throughout the experimental part.

Another two parameters were identified – namely $\mathbf{S_{max}}$ (maximum binding capacity of an epitope) and **number of signal curves used for curve fitting**. In computer simulations, S_{max} was arbitrary selected, while it has proven to be more problematic to determine in wet lab results. By not reaching the saturation plateau, S_{max} had to be determined by visually detecting the curve's inflection point and/or by observing the result of minimization equation – the S_{max} at which result was the lowest is the most valid prediction. On the other, varying the number of signals curves used for curve fitting has proven to affect affinity distribution, namely the number of peaks, considerably. However, an arbitrary number of 200 was chosen because the such a number produces the best curve fits. Lower number produced lower number of peaks, while higher numbers resulted in higher number of peaks, yet the curve fits in both cases of lower fitness than in cases with 200 signals curves.

It was demonstrated that by determining the affinity distribution, the antibody cross-reactivity can be determined if difference in Kd values is at least 10 and binding sites have binding capacity ratio similar to their Kd ratio, despite 10% measurement uncertainty. Nonetheless, this method cannot differentiate binding sites that have same or very similar Kd values. Moreover, it is worth nothing that by changing modeling parameters a number of different affinity distributions can be deduced for the same antibody. This process can be partly mended by using a uniform set of modeling parameters.

Looking back at according objectives, we see that all of them were fulfilled.

Objectives of mathematical modeling and computer simulations:

- Theoretically determine the mathematical model of interaction by which heterogeneous antibodies are binding to heterogeneous epitopes, and determine the parameters of binding process. ✅
- From proposed model, computationally solve the indirect problem of acquiring the affinity distribution from binding process using simulated experimental binding data. ✅
- Simulate how likelihood of determining cross-reactivity detection is being influenced by varying the parameters of binding and level of cross-reactivity. ✅

5.2. WET LAB EXPERIMENTATION

In order to confirm pseudo-first order model to be suitable for describing antibody binding dynamics, series of dilutions of several different antibodies and antibody combinations were used in immunostaining experiments. Judging from Figures 13, 16, 18, 20, 22 such a model was confirmed to be valid in all the cases.

In two different experiments with two monoclonal anti-GFP antibodies, the affinity distribution of such a mixture produced very similar results as seen by comparing Figure 15 with Figure 17. This confirms that the model can, using uniform set of parameters, produce repetitive results. In addition, the technique of staining was direct for one experiment and indirect for the other. Nonetheless, the difference in results was minimal and did not affect the ability to determine cross-reactivity (ie. major binding sites), thus this method is not influenced by the technique of staining.

The experiment with polyclonal anti-GFP antibody resulted in a single peak as seen in Figure 19. Polyclonal antibody is composed of many different antibodies targeting the same antigen and should as such produce a multiple peak affinity distribution. Nonetheless, solution may be suggested by the very fact that all of them are targeting epitopes of the same antigen - GFP. In the case of antibodies competitive binding, the affinity distribution of many different antibodies binding to a target is indistinguishable from binding of a single antibody to the same target, with Kd value equal to weighted-average affinity of Kd values of many different antibodies (16). In this case, a competitive binding may come from polyclonal antibody targeting only one epitope, or because of structural interference of already bound antibody. If

an antibody is already bound to a certain GFP epitope, another antibody binding to GFP's different epitope is increasingly difficult because of the structural interference of the already bound antibody. Given the fact that antibodies are huge molecules, the later may be a valid conclusion. In affinity distribution, more difficult but not impossible binding may be observed as increase in broadness of the binding peak due to different levels of binding interference.

Cross-reactivity detection experiments were designed to test the ability of the model to determine monoclonal antibody cross-reactivity if present. The basic principle was to immunostain the system with two monoclonal antibodies individually and then as a mixture to evaluate if affinity distribution of a mixture is able to capture the sum of affinity distribution of both individual antibodies.

As seen from Figure 21, in the case with anti-ERK and anti-pERK monoclonal antibody, the peaks of the mixture do not correspond with peaks resulting from individual antibody peaks. The major mixture peak in Figure 21 is positioned between major ERK and major pERK peak, albeit in more proximity to ERK peak than to pERK peak. This is consistent outcome for competitive antibody binding in the case of the mixture. As a result in the affinity distribution of antibody mixture, two major peaks contribution to ERK and pERK are not seen observed. On the other hand, what is observed is a single peak with a weighted-average Kd value of ERK and pERK peaks. Mixture's single peak is in closer proximity of ERK peak because it has greater binding capacity as seen from Figure 20. Therefore we can deduce that anti-ERK and anti-pERK antibodies competitively bind to the same antigen.

Moreover, evaluating Figure 21, both highest affinity peaks of individual ERK and pERK have a very similar, if not identical, Kd value. This suggest anti-ERK antibody may be cross-reactive with anti-pERK antibody, both targeting an almost identical molecule (pERK is phosphorylated ERK). Such cross-reactivity may pose a problem in standard experimental procedures where low antibody concentrations are used. In lower concentrations binding to highest affinity peak will contribute significantly to the measured signal, supposedly coming solemnly from ERK molecule. In experimental design where for example one would want to measure the ratio of ERK phosphorylation with the use of these antibodies a major error can be made as a consequence of antibody cross-reactivity.

In the other cross-reactivity detection experiment with anti-pERK and anti-NF200 antibodies it is clearly seen that individual signal curves in Figure 22 can be summed into signal curve of

a mixture. Figure 23 suggests that mixture's affinity distribution can roughly capture the affinity distribution of both individual antibodies.

In determining affinity distribution of several antibodies, be it mixture of monoclonal antibodies or polyclonal antibody, it is vital that antibodies are not binding competitively. In competitive binding, the resulting affinity distribution will merely show a single binding site with a Kd value equal to weighted-average of Kd values of antibodies competing for binding. This in turn means that affinity distribution is very hard to determine for polyclonal antibodies which compete for same epitopes or structurally interfere with the binding process.

Nonetheless, the modeling parameters such as number of curves used for curve fitting and the amount of regularization can significantly influence resulting affinity distribution. Also each antibody maximum capacity has to be visually determined and manually inputted. Unified parameter values were used to put all data to as similar analysis as possible. However, were we to change parameters the resulting affinity distribution would change accordingly.

Looking back at according objectives, we see that all of them were fulfilled.

Objectives of wet lab experimentation:

- Confirm theoretically determined mathematical model of antibody-antigen binding with monoclonal and polyclonal antibodies targeting specific antigen. ✅
- Implement and evaluate the mathematical parameters of binding process on monoclonal antibodies by creating an artificial cross-reactivity experimental system. ✅
- Determine the strengths and weaknesses of the model. ✅

6. CONCLUSION

The new method of obtaining antibody affinity distribution based on pseudo-first order model has proven to be successful in the use of detecting cross-reactivity of monoclonal antibodies. Such detection is limited to a specific environment (in our case HeLa cells) and cannot be generalized onto other environments.

Capturing the whole affinity distribution of an antibody has eluded us mainly because the measurement uncertainty is too high and concentration range of antibodies used is practically limited. However, the practically relevant part of affinity distribution is high affinity part, which was successfully covered.

The most accurate way of detecting cross-reactivity is by the use of knock-out mice. In comparison, the developed method is in testing stages but can deliver cross-reactivity detection with less, yet sufficient accuracy with monoclonal antibodies. On the other hand, it is also quicker (can be done in a day) and more cost effective. More importantly, it can detect cross-reactivity of most of monoclonal antibodies, whereas knock-out mice technique can only be used if specific knock-out mice exist. Moreover, it is able to detect cross-reactivity in any environment (different cell lines, tissues) while in the case with knock-out mice one is limited to mouse cell or tissue environment.

Whereas detecting cross-reactivity with preabsorption to immunizing peptide is a valid technique for polyclonal antisera, it is of little use in case of monoclonal antibodies. The developed method has proven to be primarily successful in detecting monoclonal antibody cross-reactivity, and can be as such applied to evaluate different antibodies as preabsorption technique.

Another advantage of developed method is that it is applied in exactly the same way as the antibodies are used in a standard laboratory staining practice. Techniques such as western blot employ detergents, salts and buffers for cell lysis and to solubilize proteins, thereby destroying their natural 3-D conformation. These findings are then transferred into completely new experimental design. It is obvious that potential for cross-reactivity of certain antibody depends on the environment in which it is used, due to different antigens present in individual environments. With developed method the detection of cross-reactivity and immunofluorescence staining are done in the same way and in the same environment, thereby

eliminating potential experimental differences. Furthermore, potential antigens only undergo cell fixation and thereby remain in their native 3-D form, preserving binding sites and binding affinities.

Generally there seems to be a lack of antibody specificity assurance techniques, which are vital for immunoassays validity. Immunofluorescence staining assurance can be evaluated with developed method. By detecting a cross-reactivity of monoclonal antibodies in such a way, we would produce a parameter by which such antibodies can be evaluated, based both on antibody nature and the environment in which antibody is to be used.

6.1. FURTHER RESEARCH

In order for pseudo-first order model to produce consistent results, the modeling parameters need to be further tested by varying them and monitoring the output. This output needs to be connected with a proven realistic situation about antibody affinity distribution. We propose a further validation test of the model by using antibodies with proven cross-reactivity antibodies, based on knock-out mice findings. A valid model is expected to be consistent with these findings in regards to cross-reactivity detection.

7. LITERATURE

1 Wild D: The Immunoassay Handbook, Third Edition, Elsevier, Oxford UK, 2005: 3-5

2 Chard T: An introduction to radioimmunoassay and related techniques, revised enlarged edition, Elsevier Biomedical Press, Amsterdam, 1982, 5-6

3 Kellner R, Mermet JM, Otto M, Widmer HM: Analytical Chemistry, Wiley-VCH, New York, 1988: 405-429

4 Deshpande SS: Immunoassay Classification and Commercial Technologies, Springer US, 1996: 231-232

5 Vozelj M: Temelji imunologije, DZS, First Edition, Ljubljana, 2000: 100-111

6 Robinson JP, Sturgis J, Kumar GL: Immunohistochemical Staining Methods Education Guide, Fifth Edition, Dako North America, Carpinteria CA, 2009: 61-65

7 Crowther JR: The ELISA guidebook, First Edition, Humana Press, Totowa NJ 2001: 421

8 Life Technologies's Anti-GFP Polyclonal antibody, conjugated by A647 - http://www.lifetechnologies.com/order/catalog/product/A21312

9 Roche's Anti-GFP antibody – http://www.roche-applied-science.com/shop/products/anti-gfp#tab-0

10 Koivunen ME, Krogsrud RL: Principles of Immunochemical Techniques Use in Clinical Laboratories. LabMedicine 2006; 37: 490-497

11 Janeway C: Immunobiology, Fifth Edition, Garland Publishing, New York, 2001, Chapter 3: Antigen Recognition by B-cell and T-cell Receptors

12 Kunik V, Ashkenazi S, Ofran Y: Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. Nucleic Acid Research 2012; 40:521-524

13 Sela-Culang I, Kunik V, Ofran Y: The Structural basis of antibody-antigen recognition. Frontier in Immunology 2012; 4: Article 302

14 Male D, Brostoff J, Roth DB, Roitt I: Immunology, Elsevier, Seventh Edition, Canada, 2006: 59-87

15 Frank SA: Immunology and Evolution of Infections Disease, Princeton University Press, Princeton NJ, 2002: 33-43

16 Svitel J et al.: Combined Affinity and Rate Constant Distributions of Ligand Populations from Experimental Surface Binding Kinetics and Equilibria. Biophysical Journal 2003; 84: 4062-4077

17 Saper CB: An Open Letter to Our Readers on the Use of Antibodies. Editorial. The Journal of Comparative Neurology 2005; 493:477–478

18 Mix-n-Stain CF633 Antibody Labeling Kit - http://biotium.com/product/mix-n-stain-cf633-antibody-labeling-kit/

19 Anti-ERK Monoclonal antibody - http://www.cellsignal.com/products/9106.html

20 Anti-pERK Monoclonal antibody - http://www.abcam.com/erk1-pt202py204--erk2-pt185py187-mapk-yt-antibody-ab50011.html
21 Antibody structure in Figure 1 - http://www.news-medical.net/image.axd?picture=2012%2F11%2Fparts+of+the+antibody.jpg
22 Tsien R: Annual Review of Biochemistry, Annual Reviews, 1998; 67: 509-512
23 Yoon S, Seger R: The extracellular-regulated kinase: multiple substrates regulate diverse cellular functions. Growth Factors 2006; 24: 21-44

8. APPENDIX

Basic structure of the code written in MATLAB for computer simulations and data analysis of wet lab experimental measurements are herewith presented.

1. Code for computer simulations:

```
% Set modeling parameters
smax = 100;
nme = 10;
c = logspace(-10,-6,nme);
noise = awgn(c,10);
Kd = logspace(-10,-6,nme);
% 3 binding sites distribution
u1 = 1e-9;
q1 = 3e-10;
P1 = (1./(sqrt(2*pi)*q1))*exp(-((Kd-u1).^2)./(2*(q1.^2)));
P1norm = (P1)./norm(P1,1);
u2 = 1e-8;
q2 = 3e-9;
P2 = (1./(sqrt(2*pi)*q2))*exp(-((Kd-u2).^2)./(2*(q2.^2)));
P2norm = (P2)./norm(P2,1);
u3 = 1e-7;
q3 = 3e-8;
P3 = (1./(sqrt(2*pi)*q3))*exp(-((Kd-u3).^2)./(2*(q3.^2)));
P3norm = (P3)./norm(P3,1);
% Measured net signals
for i=1:length(P1norm)
s1(i) = sum(P1norm.*(smax./(1+(Kd./c(i)))));
s2(i) = sum(P2norm.*(smax./(1+(Kd./c(i)))));
s3(i) = sum(P3norm.*(smax./(1+(Kd./c(i)))));
end
figure(2)
semilogx(c,s1,c,s2,c,s3)
hold on
% Capacity adjustment
B3 = 100;
B2 = 10;
Pr = (B3*P3 + P1 + B2*P2)./(norm(B3*P3 + P1 + B2*P2,1));
figure(1)
semilogx(Kd,Pr,'g')
hold on
% Input noise parameter
for i=1:length(Pr)
sr1(i) = sum(Pr.*(smax./(1+(Kd./c(i))))+noise);
end
for i=1:length(sr1)
    sr(i)= sr1(i)+noise(i);
end
figure(2)
semilogx (c,sr,'x')
% Formating signal curves for fitting without regularization
```

```
Kd1 = logspace(-10,-6,100);
for i=1:1:length(c)
for j=1:1:length(Kd1)
    A(i,j) = sum(smax./(1+Kd1(j)./c(i)));

end
end
Aeq = ones(1,j);
beq = [1];
lb = zeros(1,length(Kd1));
% quadprog algorithm for solving minimization equation
H = 2*A'*A;
f = -2*sr*A;
[p, fval] = quadprog(H,f,[],[],Aeq,beq,lb,[]);
% parameter of fitness
C = fval + sr*sr'
% Plot results
figure(2)
sc = A*p;
semilogx(c,sc)
figure(1)
semilogx (Kd1,p, 'r')
```

2. Code for data analysis of wet lab measurements:

```
smax1 = 200;% Manually choosen
% Input signal and dilution measurements
sr1= [120.068 101.518 64.285 44.77 23.56 9.735 3.635 4.03 2.845 0.0875];
c= [1/20 1/40 1/80 1/160 1/320 1/640 1/1280 1/2560 1/5120 1/10240];
% Input error bars
figure(1)
E1 = [9.74 3.64 4.03 2.85 0.088 2.69 5.36 1.50 3.32 4.81];
errorbar(c,sr1,E1,'.')
h1 = errorbar(c,sr1,E1,'.');
set(get(h1,'Parent'), 'XScale', 'log')
hold on
% Plot
figure(1)
semilogx(c,sr1,'*')
hold on
% Fitting
% Formating of signal curves for fitting
nKdL1 = 200;
KdL1 = linspace(-4,0,nKdL1);
h = 1;
% Create matrix A
for i=1:1:length(c)
for j=1:1:length(KdL1)
    A1(i,j) = sum(smax1./(1+(10.^KdL1(j))./c(i)));
end
end
Aeq = ones(1,j);
beq = [1];
lb = zeros(1,length(KdL1));
```

```
% Input regularization
a=0.01;
D = (sqrt(a)./(h.^2))*(-2.*eye(nKdL1)+diag(ones(nKdL1-
1,1),1)+diag(ones(nKdL1-1,1),-1));
figure(2)
hold on
% Use quadprog algorithm to solve minimization fitting equation
H = 2*(A1'*A1+a*D'*D);
f = -2*sr1*A1;
[p1, fval1] = quadprog(H,f,[],[],Aeq,beq,lb,[]);
% Plot results
figure(1)
sc1 = (A1)*p1;
h1 = semilogx(c,sc1,'b', 'LineWidth',2)
figure(2)
plot(KdL1,p1, 'b', 'LineWidth',2)
hold on
```