

The Quantitative Human Interactome:

Genome-wide approaches to gather quantitative parameters for mechanistic models; Relevance for Systems Pharmacology

Executive Summary

The goal of the project is to experimentally obtain kinetic constants for cellular interactions between all of the proteins encoded by the human genome and construct a database of these parameters. Such a database of quantitative parameters will enable the development of differential equation-based kinetic models of cellular processes. Such models are necessary for the understanding of drug action and will serve as a major driver for the discovery of new drugs for many complex diseases. The development of quantitative multi-scale models can provide a mechanistic understanding of the therapeutic action and adverse effects of drugs at an organismal level.

This is the report from the one-day workshop by the Systems Biology Center New York to initiate discussion on the feasibility of a Quantitative Human Interactome Project. In addition to a general discussion, the focus in this first workshop was on the most appropriate experimental approaches to obtaining equilibrium and rate constants on a global scale. The availability of a large number of human gene clones that can be expressed and biochemically studied makes it feasible to conduct a large-scale effort to obtain the quantitative parameters for most of the interactions between proteins in the human genome. The lack of a genome-wide library of experimentally determined kinetic constants presents a serious barrier to the development of large scale computational models of cell physiology necessary to understand drug action.

This project will integrate large-scale quantitative biochemistry with the development and population of a carefully constructed, well-curated database. A fully populated quantitative interaction database (QID) will have for each interaction (non-covalent or enzymatic) both the reaction rates and estimates of initial concentrations of reactants in a precisely defined state. Such quantitative interaction database (QID) will be a universal resource for researchers in many fields. Starting from this quantitative information, researchers will be able to develop cell type-specific quantitative models for a comprehensive understanding of complex human diseases and the development of new therapeutic approaches. The overall project would have three components:

Measurement of Interaction Constants: Expression, purification, post-translational modification of proteins (or functional fragments) and measurement of interaction constants (K_D) and interaction rates (k_f and k_b) using surface plasmon resonance technology. A separate set of experiments will measure enzymatic activities to determine K_m , V_{max} and K_{cat} when appropriate.

Measurement of levels of cellular components: Development of a library of antibodies from commercial and non-commercial sources to quantitatively estimate cellular levels of proteins in 3-5 human cell types. Mass spectrometric analyses will be conducted to quantify key ions, nucleotides, sugars and lipids.

Informatics: A vigorous informatics program that conducts gap analyses to identify currently available resources and information to drive the experimental measurement projects; development of a well-curated quantitative interaction database (QID). The QID will also contain precise information defining the biological conditions under which quantitative information was collected. The QID will be set up in a format so as to enable exporting of the quantitative information into standard modeling programs such as MatLab.

**SBCNY Workshop Wednesday March 26, 2008
held at Mount Sinai School of Medicine**

<u>Participants</u>	
Robert Blitzer	MSSM
Yibang Chen	MSSM
Lakshmi Devi	MSSM
Heidi Hamm	Vanderbilt
Ravi Iyengar	MSSM
Susana Neves	MSSM
Björn Öbrink	Karolinska
Suzanne Scarlata	SUNY Stonybrook

This workshop was held as part of the ongoing activities of the Systems Biology Center New York and coincided with the visit of Professor Bjorn Obrink from the Karolinska Institute as Senior Visiting Scholar to the Center. The purpose of the workshop was to analyze the current state and develop a plan to obtain kinetic constants for interactions on a genome wide basis to permit the development of large scale quantitative models.

Long Term Goal

The goal of the project is to experimentally obtain kinetic constants for interactions between cellular components. This will ultimately enable the development of ordinary or partial differential equation-based kinetic models of cellular processes. Such models will be valuable for *in silico* drug discovery and the development of multi-scale models that can provide a mechanistic understanding of how organ and tissue levels functions arise from reactions between cellular components.

Current Status and State of Problem

Currently (August, 2008), 17,518 non-redundant human gene clones are available through the NIH Mammalian Gene Collection Project (<http://mgc.nci.nih.gov/>). Clones for expression are also readily available from other sources (Temple G. et al Human Mol. Gen 15: R31-43). Two large scale experiments have yielded detailed protein-protein interactions networks (Stelzl U., et al (2005) Cell 122:957 and Rual J. F. et al (2005) Nature 437: 1173). However, up to this date there have been no systematic attempts to obtain kinetic constants for the interactions. The lack of a library of experimentally determined kinetic constants presents a serious barrier to the development of large scale computational models. These models would be useful in the understanding of basic biological mechanisms and for drug discovery for complex diseases.

Relevance to Pharmacology and Drug Discovery

Quantitative kinetic models would be useful not only for the understanding of basic biological mechanisms but may also be essential for drug discovery for many complex diseases including cancer, immune disorders, psychiatric disorders, diabetes and other metabolic syndromes. The effectiveness of drugs for their therapeutic targets arises from the affinity of the drug for its molecular substrate as compared to interactions with cellular components that are not therapeutic targets. Adverse effects arise from the affinity of the drug for these non-therapeutic targets and the resultant biological effects. Thus, in order to understand the mechanisms of drug action at a systems level, all drug interactions (therapeutic and adverse) need to be characterized quantitatively. For this, it is necessary to develop large-scale kinetic models. To develop such models, it is imperative to experimentally obtain the

kinetic constants for all of the interactions that make up the human interactome and have them in a well-curated database. Such a quantitative interaction database (QID) will allow a wide range of researchers to use these parameters to build and analyze quantitative models for many physiological functions as well as pathophysiology.

Rationale

Why do we need these numbers?

If the long-term goal is to have a “druggable” genome whereby drug targets can be selected rationally and drugs developed on a genome wide basis it is essential that quantitative, molecularly specified models of physiological processes be developed. Such models will be major drivers for the three key phases of the drug discovery and development process. These include target identification and drug development, estimation of dosage and development of a therapeutic regime, and prediction and characterization of adverse effects.

Approach

1. Choice of Organism

There was consensus among all participants that human proteins would be the best choice since the kinetic models can be used for drug discovery. Currently, estimated constants that are mostly labeled “mammalian” come from protein purification studies that often started with bovine or porcine tissues. There are many examples where studies in animal models do not reflect the quantitative effects seen in humans. Since human clones are available it would be best to use human proteins.

2. Approach to Expression of Protein and Protein Fragments

The pros and cons of expressing full proteins or just interacting domains to determine the rate constants were discussed. It was noted that often whole proteins are needed to study allosteric effect. There was general consensus that initially, the use of full proteins would be preferable but the numerous pros and cons of using full proteins were recognized. Depending on the initial purification a full protein to functional fragment approach might work the best. It was noted that among the most successful approaches the crystallographers have developed, is to remove the floppy bits and crystallizing the rest. For example the soluble parts of adenylyl cyclase has been crystallized and studied to obtain substantial insights, even the full intrinsic membrane protein has never been crystallized. The simplest approach would be to subclone and express the relevant functional fragments that these be used to obtain the kinetic parameter. In moving towards such a functional fragment approach a number of issues require further consideration.

- Multiple domains (such as SH2 domains) in a protein may not be equivalent.
- Unstructured (proteolytically sensitive) regions may be sites of interaction too and/or regulation (steric hinderance).
- Allosteric effects may be missing
- Some domains are unstable without the rest of the protein
- Some proteins may have multiple interactions, that could be missing if only using domains

Given all of these issues, it was suggested computational structural biology approaches could be considered to identify the functional fragments that can be used for purification for functional measurements. As a general approach it was suggested that in the pilot project the full protein be used for purification and use functional fragments only if problems arise during purification, storage or determination of kinetic constants.

3 Protein Purification and Storage

All cDNAs would have an affinity tag to permit 1-2 step purifications. Proteins would be expressed in mammalian cells or insect cells so that they have appropriate post-translational modification. Proteins would be stored in glycerol or ethylene glycol and appropriate salts and reducing agents to preserve full activity at the appropriate temp (-20 or -80°C).

Appendix 1 lists some of the tags that can be used for affinity purification, Appendix 2 the issues that need to be addressed to set up a standard operation protocol for purification and Appendix 3 a list of interaction domains.

4 General Strategy Gathering Quantitative Interaction Data

There was consensus that surface plasmon resonance (SPR) methods were the best current approach for gathering kinetic parameters. The need for development of methods to make SPR measurements medium to high throughput was discussed. The availability of 386-well SPR machines would be very useful in the developing the scale-up methodology. The major limitation of these high throughput machines is that it determines equilibrium constants (K_D) rather the kinetic rate constants (k_{forward} and k_{reverse}). Nevertheless the 386 well machines will allow for the fast exploration of differing conditions for the measurement of K_D s and hence should be useful as a general starting strategy. It was agreed that we determine the equilibrium K_D s first. This can be done in a real high throughput manner for a large number of combinations of protein protein interactions. Then we could select protein-protein interactions pairs for $k_{\text{forward}}/k_{\text{reverse}}$ determinations, based on the equilibrium K_D values. This can be done in a low to semi-high throughput manner, with appropriate planning.

5 Detailed plan for gathering kinetic parameters

5a. Determination of K_D

A high-throughput approach appears feasible: The project would determine equilibrium K_D for pairwise interactions of purified, recombinant, soluble human proteins and/or functional fragments. This could be done with the Corning Epic System, which is based on SPR-technology, performed in 384-well microplates under static conditions. This instrument can be integrated with External HTS (high throughput screen) Robotics, for automated analysis. It can read 40,000 wells per 8 hours, with a sensitivity of 5 pg/mm². If 10 concentrations (to enable Schatchard analysis) of each soluble protein binding partner is analyzed simultaneously, both on protein-binding-partner-coated wells, and on control reference wells (for subtracting non-specific binding) it means that 2000 different protein-protein binding interactions can be determined in 8 hours. If the instrument is working 16 hours a day, it will be possible to determine the K_D for 4000 different protein-protein interactions per day, or **20,000 protein-protein interactions per week** (5 working days). The binding determinations should preferably be performed at 37°C.

This means that **40,000 proteins can be completely analyzed in 32 weeks**, i.e. less than a year, based on the following assumptions and calculations: Each protein can occur as 4 different isoforms or post-translational states, which gives 160,000 proteins to analyze. Each protein can have 4 different binding partners. This gives 640,000 protein-protein interactions to analyze. 640,000 divided by 20,000 equals 32 weeks. To make this possible, it is necessary to get information from other sources, about the plausible binding partners for the respective proteins. Hence a qualitative mammalian interactome is a prerequisite starting pointing for these quantitative measurements. A non-curated human interactome

of 12,000 components 72,000 interactions developed from merging the various publicly available data bases is available in the SBCNY Bioinformatics Core.

The calculations for K_D determination are based on the output of a single SPR machine, and would be the theoretical maximal throughput for one year, also allowing working out of immobilization conditions, service and unforeseen stand-still of the instrument. It would thus theoretically be possible to analyze 1,920,000 protein-protein interactions in one year with 3 machines, or 6,400,000 protein-protein interactions per year with 10 machines.

5b Determination of $k_{\text{forward}}/k_{\text{reverse}}$

Determination of kinetic constants has to be done with a flow-cell based SPR-biosensor instrument that allows monitoring in real time of the association and dissociation reactions. Commercially available instruments that allow this kind of analysis are the BIAcore instruments. This is much more cumbersome and time-consuming than determination of equilibrium K_D s. The selection of pairwise protein-protein interactions to be analyzed in this way will be based on the determined K_D s, and on knowledge from other sources on crucial binding interactions in the human interactome. Further criteria for selecting binding interactions for kinetic evaluation will be provided by computational modeling analyses of signaling pathway modules, in order to identify those binding interactions, the kinetics of which will be crucial for the outcome, i.e. the end effect, of the signaling pathway.

For each protein pair, one of the proteins should be selected for immobilization (ligand), and the other as the binding partner in the soluble phase (analyte). Generally, amine-coupling should be chosen for immobilization. The coupling conditions and regeneration conditions will first be worked out. Appropriate treatments of control reference lanes have to be developed. Principally, two different ligands can be immobilized on each chip, which has four flow cells each. Two lanes will be used for the ligands, and two lanes will serve as reference lanes. Five different analyte concentrations will be analyzed in triplicate. The analyte concentration ranges will easily be chosen from the K_D information. Preferably, the binding analysis should be done at 37°C.

Given “normal” behavior of the proteins it should be possible to run 4 pair-wise protein-protein interactions per week with one BIAcore instrument (coupling two different ligands on the same chip). If the instrument can be operated smoothly 50 weeks a year, it is thus possible to analyze 200 protein-protein interactions. With 10 BIAcore machines, one can thus determine the binding kinetics for 2000 pair-wise protein-protein interactions in one year with current technologies. If about 10,000 or so the reactions in the cell require rate parameters in addition to equilibrium constants then this should be 5 year project.

Analysis of the binding sensorgrams for determination of $k_{\text{on}}/k_{\text{off}}$ values will be done by global curve fitting to appropriate binding algorithms. While many binding interactions may apply to a single 1:1 binding, it is expected that many interactions will follow more complex kinetics. Choosing appropriate binding models will be facilitated by information from Schatchard analyses of the equilibrium K_D determinations. Most likely, new algorithms for curve fitting of complex binding interactions will have to be generated.

Appendix 4 lists a set of studies that have determined kinetic parameters for signaling molecules. This is an initial list that may grow.

6. Changes in interaction constants with post-translational modifications

Since interactions between proteins are often regulated by phosphorylation, it will become necessary to determine how changes in phosphorylation state affect the interaction. Such an effort that would entail comparison of affinities between binding partners upon phosphorylation should be considered a separate sub-project that would run along the main project with a delayed start time. The effects of phosphorylation on the equilibrium constants (K_D) will be determined. To start this project a preparatory bioinformatics effort is needed to bring together all of the currently available experimental and theoretical knowledge regarding phosphorylation states and interaction changes for human protein. A separate work group focused on this comparative study will be needed.

7 Multicomponent complexes

Since scaffolds and anchors bind multiple proteins and are numerous within the cell, a systematic cataloging of the known information about multi-protein complexes both from the functional literature and current IP-MS experiments are needed to determine how many complexes will need to be studied as groups of interactions. The binary interactions between the binding partners will be characterized in the large scale study above and this part of the project will focus on change in affinity due to multiple interactions. A focused workgroup is necessary to design how affinity measurement experiments for complexes will be needed

8 Measurement of enzyme activities

It was agreed that we should set up in parallel enzyme assays under a standard format to determine K_m and K_{cat} for recombinant enzymes. So as to be consistent with the affinity measurements purified recombinant enzymes will be used. For most enzymes natural substrates will be used. For a majority of the enzyme substrate binding affinities will also be determined in the equilibrium binding assays.

9 Verification of interactions/activity in the intact cell

It was recommended that a sampling algorithm be set to verify the affinity constants obtained from the in vitro assays in the intact cell. For this 1-3 human cell lines can be used. Cells will be transfected with YFP and CFP labeled interacting partners and the change in interaction with respect to time in response to a stimulus will be imaged by confocal microscopy. The time course experiments will be used for curve fitting to obtain rate constants for the reaction within the cell. The calculated in-cell constants will be listed along with SPR measurements in the data base.

10 Measuring protein levels in a cell

To build dynamical models both rates constants and concentrations of reactants are required. This would entail measuring the level of proteins in a cell. There was discussion as to whether this is a separate project. It was felt that there are number of distinct technological issues that need to be addressed to obtain cellular concentrations of proteins. The two most feasible approaches are mass spectrometry and quantitative immunoblotting. Both of these methods will require technology development including standardization of antibody production for quantitative immunoblotting and hence this is likely to be a large project that should run separately but in close collaboration with the project to measurement kinetic constants. The two projects together will constitute the experimental part of the Human Interactome Project.

It was recommended that a separate workshop be held to consider in depth approaches needed to obtain quantitative measurements of levels of cellular proteins.

Bioinformatics

The discussion on bioinformatics and its integration with the experiments centered around two themes.

1. Organization of current data and GAP analysis

It was recommended that prior to embarking on the Quantitative Human Interactome Project it would be useful to develop a searchable data base of our current experimental knowledge regarding components. Data regarding the molecular weight of the components, post-translational modification and any directly determined or estimated kinetic information so that this information can help drive the experimental plan. All proteins should be clearly associated with a human Gene ID. Such data currently exists in various forms in several databases. These could be combined to develop a meta-database similar to what has been done for the interactions. Using this data we can conduct a GAP analysis of what quantitative information is available and what is missing. Such a GAP analysis could be viewed as the first phase of a large-scale project

2. Development of a Quantitative Interaction database (QID)

The data obtained from this project should be in a database that is readily compatible with standard numerical solvers such as MatLab or Virtual Cell. It may be useful to have the data in the context of reactions that can be directly used to assemble a model. Such a system should be SBML compatible. The development of the QID will a critically important part of a quantitative human interactome project and requires substantial analysis and discussion.

A separate bioinformatics workgroup focused on QID should develop a detailed plan A workshop focused on bioinformatics could be useful for this.

Pilot project

To establish the feasibility of the proposed approach it was suggested that a pilot project focus one or two signaling pathways or a small network of about 100-200 component. The pilot project should establish the utility of obtaining genome wide quantitative interaction data for drug development. Although currently technology exist for obtaining quantitative measurement it was felt that the throughput was not sufficiently high enough. A well-constructed pilot project could be an engine to drive the needed technology development.

Outreach

It was recommended that SBCNY conduct further outreach to evaluate the ideas developed in this workshop. It is recommended that feedback be solicited from Mark Vidal, Peter Sorger, Doug Lauffenberger, Boris Kholodenko, Tony Pawson, Mathias Mann, Rudy Aebersol and Walter Kolch.

Appendix 1

Various types of tags for purifying proteins /functional fragments

Fusion Tag	Ligand	Binding conditions	Available formats	Pros/ Cons
Glutathione S-transferase(GST)	Reduced glutathione	Neutral (physiologic) pH, and non-denaturing; glutathione must be reduced and GST must be active	Prepacked column kits, spin cup column kits, coated microplates	Size 26 kDa
Histidine -tagged	Chelated Nickel or Cobalt	Neutral (physiologic) pH without reducing or oxidizing agents; small tag must be accessible in fusion protein structure;	Prepacked column kits, spin cup column kits, coated microplates	6-10 histidines Prone to contaminants - other proteins (especially those containing several histidines) may bind nonspecifically to the nickel matrix . Can be used directly to immobilize protein on a Biacore using a Ni-NTA-chip (Khan et al., 2006 Anal Chem 78:3072)
Maltose Binding Protein (MBP)	Dextrin	Neutral (physiologic) pH and non-denaturing; NaCl added to reduce nonspecific binding	Gel slurry, coated microplates	Size 40KDa Hearsay: more soluble than GST tags
Green Fluorescent Protein (GFP)	Anti-GFP antibody	Neutral (physiologic) pH and non-denaturing	Coated microplates	Size 27 kDa Keep track of folding via fluorescence monitoring
3xFLAG	Anti-	Neutral	Resin	22 amino acids

	FLAg antibody	(physiologic) pH and non-denaturing		Improved version of original FLAG (8 aa)
--	---------------	-------------------------------------	--	--

Appendix 2

Issues to deal with during purification

Orientation of constructs:

- N-terminal tag
- C-terminal tag
- both

Expression:

- Mammalian
 HeLa
 293T
- Insect

Treatment of cells with +/- inhibitors to account for changes in binding affinities due to post translational modifications

Scale Purification

- 50-100 ug

Affinity Purification

- column
- slurry
- antibody

Tag removal needed?

Dialysis/desalting?

Contaminant Removal?

Quality Controls

Appendix 3

Interaction Domains and their targets

(Adapted from Pawson and Nash 2003 Science 300:445)

Domain-Post-translationally modified target

SH2	p-Tyr
PTB	p-Tyr
FHA	p-Thr
14-3-3	p-Ser
WW	p-Ser
WD40	p-Ser
MH2	p-Ser
Chr	Me-Lys
Bromo	Ac-Lys
UIM	Ub
UBA	Ub _n

Domain-short amino acid sequence

PTB	NPXY
SH3	RXXK & PXXP
WW	PPXY
EVH1	FPPPP
GYF	Pro
VHS	D/E-XXLL
PDZ	Val-COOH

Domain- Nucleic Acid

PUM	RNA
Tubby	DNA

Domain-Domain

PDZ	PDZ
SAM	SAM
DD	DD
DED	DED
CARD	CARD
PyD	PyD
PB1	PB1
BRCT	BRCT

Domain- Phospholipid

PH	PIP ₂ , PIP ₃
FYVE	PI3P
C1	DAG
FERM	PIP ₂
C2	PA, PS
Tubby	PIP ₂

PX PI3P
ENTH PIP₂, PIP₃

Post-translational modifications and their inhibitors:

- Phosphorylation
- all sodium fluoride, sodium orthovanadate, sodium pyrophosphate and b-glycerophosphate.
- ser/threo Okadaic Acid or Microcystin-LR (PP1 and PP2A inhibitor)
- tyrosine Sodium Orthovanadate

- acylation
- glycosylation,
- prenylation
 - myristoylation
 - farnesylation FPT Inhibitor II (Ras)
 - geranylgeranylation GGTI-2147 (Rap1)

Appendix 4

BiaCore Kinetic Data Available on Signaling Molecules

- CD2-CD48 interaction
(van der Merwe, Brown et al. 1993 Affinity and kinetic analysis of the interaction of the cell adhesion molecules rat CD2 and CD48)
- IL5- IL5R
(Morton, Bennett et al. 1994 Analysis of the interaction between human interleukin-5 and the soluble domain of its receptor using a surface plasmon resonance biosensor)
- Calmodulin-like domains on Calpain- Calpastatin
(Yang, Ma et al. 1994 Analysis of calcium-dependent interaction between amino-terminal conserved region of calpastatin functional domain and calmodulin-like domain of mu-calpain large subunit)
- P-Shc binding to Grb2
(Pelicci, Giordano et al. 1995 The motogenic and mitogenic responses to HGF are amplified by the Shc adaptor protein)
- Grb2 and Sos1-c-terminus (SH3)
(Sastry, Lin et al. 1995 Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity)
- Neurocan and N-CAM
(Retzler, Gohring et al. 1996 Analysis of neurocan structures interacting with the neural cell adhesion molecule N-CAM)
- S100A1 and RYR complex
(Treves, Scutari et al. 1997 Interaction of S100A1 with the Ca²⁺ release channel (ryanodine receptor) of skeletal muscle)
- Ran and RanBP1
(Kuhlmann, Macara et al. 1997 Dynamic and equilibrium studies on the interaction of Ran with its effector, RanBP1)
- EGF and EGFR
(Domagala, Konstantopoulos et al. 2000 Stoichiometry, kinetic and binding analysis of the interaction between epidermal growth factor (EGF) and the extracellular domain of the EGF receptor)
- p53 and Bcl2
(Tomita, Marchenko et al. 2006 WT p53, but not tumor-derived mutants, bind to Bcl2 via the DNA binding domain and induce mitochondrial permeabilization)
- BMP-3 and Activin Receptor type II (ActRII)
(Allendorph, Isaacs et al. 2007 BMP-3 and BMP-6 structures illuminate the nature of binding specificity with receptors)

References:

Allendorph, G. P., M. J. Isaacs, et al. (2007). "BMP-3 and BMP-6 structures illuminate the nature of binding specificity with receptors." Biochemistry **46**(43): 12238-47.

Bone morphogenetic proteins (BMPs) are extracellular messenger ligands involved in controlling a wide array of developmental and intercellular signaling processes. To initiate their specific intracellular signaling pathways, the ligands recognize and bind two structurally related serine/threonine kinase receptors, termed type I and type II, on the cell surface. Here, we present the crystal structures of BMP-3 and BMP-6, of which BMP-3 has remained poorly understood with respect to its receptor identity, affinity, and specificity. Using surface plasmon resonance (BIAcore) we show that BMP-3 binds Activin Receptor type II (ActRII) with K_d approximately 1.8 μM but ActRIIb with 30-fold higher affinity at K_d approximately 53 nM. This low affinity for ActRII may involve Ser-28 and Asp-33 of BMP-3, which are found only in BMP-3's type II receptor-binding interfaces. Point mutations of either residue to alanine results in up to 20-fold higher affinity to either receptor. We further demonstrate by Smad-based whole cell luciferase assays that the increased affinity of BMP-3S28A to ActRII enables the ligand's signaling ability to a level comparable to that of BMP-6. Focusing on BMP-3's preference for ActRIIb, we find that Lys-76 of ActRII and the structurally equivalent Glu-76 of ActRIIb are distinct between the two receptors. We demonstrate that ActRIIbE76K and ActRII bind BMP-3 with similar affinity, indicating BMP-3 receptor specificity is controlled by the interaction of Lys-30 of BMP-3 with Glu-76 of ActRIIb. These studies illustrate how a single amino acid can regulate the specificity of ligand-receptor binding and potentially alter biological signaling and function in vivo.

Domagala, T., N. Konstantopoulos, et al. (2000). "Stoichiometry, kinetic and binding analysis of the interaction between epidermal growth factor (EGF) and the extracellular domain of the EGF receptor." Growth Factors **18**(1): 11-29.

The kinetics, binding equilibria and stoichiometry of the interaction between epidermal growth factor and the soluble extracellular domain of the epidermal growth factor receptor (sEGFR), produced in CHO cells using a bioreactor, have been studied by three methods: analytical ultracentrifugation, biosensor analysis using surface plasmon resonance detection (BIAcore 2000) and fluorescence anisotropy. These studies were performed with an sEGFR preparation purified in the absence of detergent using a mild two step chromatographic procedure employing anion exchange and size exclusion HPLC. The fluorescence anisotropy and analytical ultracentrifugation data indicated a 1:1 molar binding ratio between EGF and the sEGFR. Analytical ultracentrifugation further indicated that the complex comprised 2EGF:2sEGFR, consistent with the model proposed recently by Lemmon et al. (1997). Global analysis of the BIAcore binding data showed that a simple Langmuirian interaction does not adequately describe the EGF:sEGFR interaction and that more complex interaction mechanisms are operative. Furthermore, analysis of solution binding data using either fluorescence anisotropy or the biosensor, to determine directly the concentration of free sEGFR in solution competition experiments, yielded Scatchard plots which were biphasic and Hill coefficients of less than unity. Taken together our data indicate that in solution there are two sEGFR populations; one which binds EGF with a K_D of 2-20 nM and the other with a K_D of 400-550 nM.

Kuhlmann, J., I. Macara, et al. (1997). "Dynamic and equilibrium studies on the interaction of Ran with its effector, RanBP1." Biochemistry **36**(40): 12027-35.

Ran, a small nuclear GTP-binding protein, is one of the most abundant Ras-related proteins in eucaryotic cells. Ran is essential for nucleo-cytoplasmic transport and is primarily localized in the nucleus and at the nuclear pore complex. Here, we characterize the kinetics and equilibrium of the interaction between Ran and RanBP1 by two independent biophysical approaches: fluorescence spectroscopy using analogues of guanine nucleotides and surface plasmon resonance in the BIAcore system. Both approaches result in kinetic and equilibrium data which are in good agreement with each other. Affinities of RanBP1 for Ran in the GTP-bound state were in the nanomolar range, while Ran.GDP bound RanBP1 with a dissociation constant around 10 micromM. Interestingly, the difference in affinity of RanBP1 for Ran.GDP was mostly due to a dramatic increase of the dissociation rate constant. Mutant Ran protein lacking the last five amino acids of the C-terminus (RanDeltaC) is unable to facilitate nuclear import in vitro and does not bind to RanBP1. Here, we show that RanBP1 binds RanDeltaC.mGppNHp with KD values around 10 micromM, as is the case for its association with full-length Ran.GDP. The loss of affinity of RanBP1 for the triphosphate form of RanDeltaC was a result of both a decrease of the association rate and a moderately increased dissociation of the RanDeltaC.RanBP1 complex. Circular dichroism spectra indicate significant changes in the secondary structure of either Ran.GppNHp, RanBP1, or both proteins upon forming a stable complex with each other.

Morton, T. A., D. B. Bennett, et al. (1994). "Analysis of the interaction between human interleukin-5 and the soluble domain of its receptor using a surface plasmon resonance biosensor." J Mol Recognit 7(1): 47-55.

A surface plasmon resonance (SPR) biosensor was used to study the interaction of human interleukin-5 (hIL5) with its receptor. IL5 is a major growth factor in the production and activation of eosinophils. The receptor for IL5 is composed of two subunits, alpha and beta. The alpha subunit provides the specificity for IL5 and consists of an extracellular soluble domain, a single transmembrane region and a cytoplasmic tail. We expressed the soluble domain of the human IL5 receptor alpha subunit (shIL5R alpha) and human IL5 (hIL5) in *Drosophila*. Both hIL5 and shIL5R alpha were immobilized separately through amine groups onto the carboxylated dextran layer of sensor chips of the BIAcore (Pharmacia) SPR biosensor after N-hydroxysuccinimide/carbodiimide activation of the chip surface. Interactions were measured for the complementary macromolecule, either shIL5R alpha or hIL5, in solution. Kinetics of binding of soluble analyte to immobilized ligand were measured and from this the association rate constant, dissociation rate constant and equilibrium dissociation constant (Kd) were derived. With immobilized shIL5R alpha and soluble hIL5, the measured Kd was 2 nM. A similar value was obtained by titration calorimetry. The Kd for *Drosophila* expressed receptor and IL5 is higher than the values reported for proteins expressed in different systems, likely due to differences in the methods of interaction analysis used or differences in protein glycosylation. Receptor-IL5 binding was relatively pH independent between pH 6.5 and 9.5. Outside this range, the dissociation rate increased with comparatively little increase in association rate.(ABSTRACT TRUNCATED AT 250 WORDS)

Pellicci, G., S. Giordano, et al. (1995). "The motogenic and mitogenic responses to HGF are amplified by the Shc adaptor protein." Oncogene 10(8): 1631-8.

The receptor of Hepatocyte Growth Factor-Scatter Factor (HGF) is a tyrosine kinase which regulates cell motility and growth. After ligand-induced tyrosine phosphorylation, the HGF receptor associates with the Shc adaptor, via the SH2 domain. Site-directed mutagenesis of the HGF receptor indicates that phosphotyrosines Y1349VHV and Y1356VNV can work as docking sites for Shc. The Kd of this interaction, measured in real time using synthetic

phosphopeptides and recombinant Shc on a BIAcore biosensor, is 150 nm for both sites. After stimulation of the HGF receptor, Shc is phosphorylated on Y317VNV, generating an high affinity binding site for Grb2 (Kd = 15 nM). This duplicates the high affinity binding site for Grb2 present on the HGF receptor (Y1356VNV). Thus HGF stimulation can trigger the Ras pathway by recruiting Grb2 both directly through the receptor, and indirectly, through Shc. Overexpression of wild-type Shc, but not of the Y317-->F mutant, enhances cell migration and growth in response to HGF. These data show that Shc is a relevant substrate of the HGF receptor, and works as an 'amplifier' of the motogenic as well as of the mitogenic response.

Retzler, C., W. Gohring, et al. (1996). "Analysis of neurocan structures interacting with the neural cell adhesion molecule N-CAM." *J Biol Chem* **271**(44): 27304-10.

Neurocan is a brain-specific chondroitin sulfate proteoglycan, which has been shown to bind to the neural cell adhesion molecule N-CAM and to inhibit its homophilic interaction. To study in more detail the structures of neurocan responsible for this interaction, various recombinant neurocan fragments were generated. The ability of these fragments to interact with N-CAM was investigated in several different in vitro assay systems, enzyme-linked immunosorbent assay-type binding assays, Covasphere-aggregation assays, and assays based on an optical biosensor (BIAcore™) system. The analysis of the homophilic N-CAM interaction in the BIAcore system revealed a KD of 64 nM. This homophilic interaction could be reduced by preincubation of soluble N-CAM with neurocan. Direct binding of N-CAM to immobilized neurocan core protein and recombinant neurocan fragments could also be demonstrated, and KD values between 25 and 100 nM were obtained. In addition, direct binding of N-CAM to chondroitin sulfate could be demonstrated. Binding of N-CAM to the immobilized neurocan core protein could be inhibited with all recombinant fragments containing chondroitin sulfate or major parts of the mucin-like central region of neurocan. For the inhibition of homophilic N-CAM interactions, however, a combination of globular and extended structures was required.

Sastry, L., W. Lin, et al. (1995). "Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity." *Oncogene* **11**(6): 1107-12.

Grb2 is an adaptor protein that links receptor and cytoplasmic tyrosine kinases to the Ras signalling pathway by binding the Ras-specific guanine nucleotide exchange factor, Sos1, through its SH3 domains. The Grb2-SH3 domain binding has been localized to the carboxy-terminal two hundred amino acids of Sos1 (Sos1-c). By using real time biospecific interaction analysis (BIAcore), we studied the kinetic parameters and binding affinity of the Grb2-Sos1-c interaction. The binding of Grb2 to Sos1-c is a high affinity interaction with a moderate association rate (9.45×10^4 per M per s), a slow dissociation rate (13.8×10^{-5} s), and an affinity constant of 1.48 nM. BIAcore measurements on isolated N-terminal and C-terminal SH3 domains (NSH3 and CSH3) further indicate that the high affinity Grb2-Sos1-c interaction is primarily mediated through the NSH3 domain (Kd = 1.68 nM). The CSH3 domain shows substantially reduced binding to Sos1-c in these measurements. Inhibition studies with BIAcore using proline rich peptides derived from the C-terminus of Sos1 show that there is a single major binding site for Grb2 in Sos1. This binding site is contained within the peptide N20, which corresponds to amino acids 1143-1162 of Sos1. This peptide completely blocks the Grb2-Sos1-c and NSH3-Sos1-c interactions with IC50 values of 8 microM and 4 microM respectively. The discrete interaction between the NSH3 domain and the N20 peptide may be amenable for drug discovery through screening or peptidomimetic approaches.

Tomita, Y., N. Marchenko, et al. (2006). "WT p53, but not tumor-derived mutants, bind to Bcl2 via the DNA binding domain and induce mitochondrial permeabilization." *J Biol Chem* **281**(13): 8600-6.

The induction of apoptosis by p53 in response to cellular stress is its most conserved function and crucial for p53 tumor suppression. We recently reported that p53 directly induces oligomerization of the BH1,2,3 effector protein Bak, leading to outer mitochondrial membrane permeabilization (OMMP) with release of apoptotic activator proteins. One important mechanism by which p53 achieves OMMP is by forming an inhibitory complex with the anti-apoptotic BclXL protein. In contrast, the p53 complex with the Bcl2 homolog has not been interrogated. Here we have undertaken a detailed characterization of the p53-Bcl2 interaction using structural, biophysical, and mutational analyses. We have identified the p53 DNA binding domain as the binding interface for Bcl2 using solution NMR. The affinity of the p53-Bcl2 complex was determined by surface plasmon resonance analysis (BIAcore) to have a dominant component KD 535 +/- 24 nm. Moreover, in contrast to wild type p53, endogenous missense mutants of p53 are unable to form complexes with endogenous Bcl2 in human cancer cells. Functionally, these mutants are all completely or strongly compromised in mediating OMMP, as measured by cytochrome c release from isolated mitochondria. These data implicate p53-Bcl2 complexes in contributing to the direct mitochondrial p53 pathway of apoptosis and further support the notion that the DNA binding domain of p53 is a dual function domain, mediating both its transactivation function and its direct mitochondrial apoptotic function.

Treves, S., E. Scutari, et al. (1997). "Interaction of S100A1 with the Ca²⁺ release channel (ryanodine receptor) of skeletal muscle." *Biochemistry* **36**(38): 11496-503.

In the present report we studied the interaction between the skeletal muscle ryanodine receptor and the ubiquitous S100A1 Ca²⁺ binding protein. S100A1 did not affect equilibrium [3H]ryanodine binding to purified rabbit skeletal muscle terminal cisternae at 100 microM free [Ca²⁺]. At nanomolar free [Ca²⁺], however, S100A1 activated by 40 +/- 6.7% (mean +/- SE, n = 5) the [3H]ryanodine binding activity; the half-maximal concentration for stimulation of [3H]ryanodine binding was approximately 70 nM, a value well below the estimated S100A1 concentration in skeletal muscle fibers. Scatchard analysis of [3H]ryanodine binding performed in the presence of 100 microM EGTA indicates that S100A1 increases the apparent affinity of the receptor for ryanodine (Kd = 191 vs 383 nM in the presence and in the absence of 100 nM S100A1, respectively). The effect of S100A1 was also tested on the single-channel gating properties of the purified ryanodine receptor after reconstitution into a lipid planar bilayer. Currents carried by purified ryanodine receptor channels were modulated by both cis Ca²⁺ and ruthenium red. In the presence of nanomolar [Ca²⁺], S100A1 activated the channel by increasing (6.0 +/- 2.8)-fold (mean +/- SE, n = 3) the normalized open probability. The interaction between S100A1 and the purified RYR was verified using the optical biosensor BIAcore: we show that the two proteins interact directly both at millimolar and at nanomolar calcium concentrations. We next mapped the regions of the skeletal muscle RYR involved in the interaction with S100A1 by performing ligand overlays on a panel RYR of fusion proteins in the presence of 100 nM S100A1. Our results indicate that the skeletal muscle RYR contains three potential S100A1 binding domains. Binding of S100A1 to the RYR fusion proteins occurred at both nanomolar and millimolar free [Ca²⁺]. S100A1 binding domain 1 binds the ligand in the presence of 1 mM free [Ca²⁺] or 1 mM EGTA. Maximal binding to S100A1#2 was achieved in the presence of 1 mM free [Ca²⁺]. The S100A1#3 domain, which overlaps with calcium-dependent calmodulin binding domain 3 (CaM 3), exhibits weak and strong S100A1 binding activity in the presence of either millimolar or nanomolar Ca²⁺, respectively. The interaction between S100A1 and the purified RYR complex was also investigated by affinity chromatography: in the presence of nanomolar Ca²⁺, we observed binding of native RYR complex to S100A1-conjugated Sepharose. This interaction could be inhibited by the

presence of RYR polypeptides encompassing S100A1 binding sites S100A1#1, S100A1#2, and S100A1#3.

van der Merwe, P. A., M. H. Brown, et al. (1993). "Affinity and kinetic analysis of the interaction of the cell adhesion molecules rat CD2 and CD48." *Embo J* **12**(13): 4945-54.

CD2 is a plasma membrane glycoprotein present on T lymphocytes that functions as a cell adhesion molecule (CAM). The CD2 counter-receptor in rodents is the structurally-related CAM CD48. Intercellular adhesion involves the formation of multiple CAM complexes between adhering cells and de-adhesion requires disruption of these complexes. To gain an insight into the initiation and termination of intercellular adhesion, the kinetics and affinity of the rat CD2-CD48 interaction was analysed using a BIAcore instrument, which enables the monitoring of protein binding in real time. A soluble chimeric protein, comprising the extracellular portion of rat CD48 and domains 3 and 4 of rat CD4 (sCD48-CD4), bound to immobilized soluble CD2 (sCD2) with a KD of 90 microM. The affinity was also determined in the reverse orientation and sCD2 was shown to bind immobilized sCD48-CD4 with a comparable KD of 60 microM. sCD48-CD4 bound to immobilized deglycosylated sCD2 with a KD of 125 microM, indicating that glycosylation of sCD2 has little effect on the affinity of the interaction. The low affinity was the result of an extremely rapid off-rate constant ($K(\text{off}) > \text{or} = 6 \text{ s}^{-1}$), whereas the on-rate constant was unremarkable ($K(\text{on}) > \text{or} = 10(5) \text{ M}^{-1}\text{s}^{-1}$). The kinetic analysis revealed that small amounts of multimeric aggregates of sCD48-CD4 formed in concentrated preparations. Our experience suggests that even low concentrations (< 2%) of these aggregates may be a cause of artifactually high affinity values when analysing low-affinity protein interactions. In conclusion, this study provides the first detailed analysis of the kinetics and affinity of monomeric CAM interactions and suggests that binding between CAMs may be weaker than anticipated.

Yang, H. Q., H. Ma, et al. (1994). "Analysis of calcium-dependent interaction between amino-terminal conserved region of calpastatin functional domain and calmodulin-like domain of mu-calpain large subunit." *J Biol Chem* **269**(29): 18977-84.

Calpain requires Ca^{2+} both for proteolysis of its substrates and for interaction with its endogenous inhibitor, calpastatin. Although calmodulin-like domains (CaMLDs) of large and small subunits of calpain have been suggested to be the sites for Ca^{2+} -dependent interaction with calpastatin, specificity and molecular basis of the interaction have remained unclear. We investigated the interaction between the CaMLD of human mu-calpain large subunit expressed in *Escherichia coli* and a 19-residue synthetic oligopeptide corresponding to the region A (the amino-terminal conserved acidic region) of one of the four repetitive functional domains of calpastatin. The recombinant CaMLD bound to the oligopeptide immobilized on Sepharose beads in a Ca^{2+} -dependent manner. The CaMLD failed in binding to a mutant oligopeptide with one amino acid substitution. Enhancement of fluorescence intensity of a hydrophobic probe, 2-(p-toluidino)naphthalene-6-sulfonate, was observed upon incubating with the CaMLD and further increased by Ca^{2+} . The Ca^{2+} -dependent enhancement of fluorescence intensity was strongly suppressed by the wild type oligopeptide, but not by the mutant one. Kinetic experiments were performed with BIAcore where binding of the CaMLD to the oligopeptide immobilized on a biosensor chip was detected as real time signals of surface plasmon resonance. The determined dissociation constant (KD) was $3.1 \times 10^{-9} \text{ M}$. These results suggest that the region A of calpastatin binds to the CaMLD in a specific manner similar to interactions between calmodulin-binding peptides and calmodulin where hydrophobic properties are known to be important.

