Spring 2015

Development of experimental and instrumental systems to study biological systems

Amanda J Hemphill
Purdue University

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Amanda J. Hemphill

Entitled   DEVELOPMENT OF EXPERIMENTAL AND INSTRUMENTAL SYSTEMS TO
STUDY BIOLOGICAL SYSTEMS

For the degree of   Doctor of Philosophy

Is approved by the final examining committee:

Chittaranjan Das

Brian Todd co-chair

Garth J. Simpson

Marcy H. Towns

Ken Ritchie

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Chittaranjan Das

Approved by Major Professor(s): _____

_Brian Todd__co-chair_____

Approved by: R. E. Wild   04/20/2015

Head of the Department Graduate Program   Date

DEVELOPMENT OF EXPERIMENTAL AND INSTRUMENTAL SYSTEMS TO
STUDY BIOLOGICAL SYSTEMS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Amanda J Hemphill

In partial Fulfillment of the

Requirements of the Degree

of

Doctor of Philosophy

May 2015

Purdue University

West Lafayette, Indiana

"Remember to look up at the stars and not down at your feet. Never give up work. Work gives you meaning and purpose and life is empty without it. If you are lucky enough to find love, remember it is there and don't throw it away."

― Stephen Hawking

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure   Page

ABSTRACT

Hemphill, Amanda J. Ph.D. Purdue University, May 2015. Development of Experimental and Instrumental Systems to Study Biological Systems. Major Professor: Brian A. Todd.

Chapters 1-4 of this thesis describes the development of an experimental system to measure diffusion-limited reaction kinetics in a biological environment. About 100 years ago, the relationship between reaction rate and diffusion in homogenous solution, ie water or buffer, was described as a linear relationship by Smoluchowski. Applying this theory naively would suggest that since the diffusion coefficients drop by factors of 4-100 then the rates of reaction would drop by the same amount. However, recent theory and simulations suggest that this does not hold. Even though biological diffusion coefficients drop to 0.1-20% of that in buffer, these recent studies show that the reaction kinetics are much more weakly affected by the biological environment. Due to the lack of experimental evidence for biological diffusion, there is a great need for information in this area. Here, I describe a protein system, exogenous to *E. coli*, that will form a dimer in the presence of a small molecule.

I also describe the development of a new type of multivariate hyperspectral Raman instrument (MHI); the instrument is developed for use to study biological tissues and for high speed cell sorting applications. The new instrument design has a large speed

advantage over traditional Raman instrumentation for rapid chemical imaging. While the MHI can reproduce the functionality of a traditional Raman spectrometer, its true speed advantage is realized after pre-training on known sample components. The MHI makes use of a spatial light modulator as a programmable optical filter that can be programmed with filters based on multivariate signal processing algorithms, such as PLS, in order to rapidly detect chemical components and create chemical maps. Chapters 5-8 of this thesis describe the development and construction of the MHI, as well as provide proof-of-concept experimental results demonstrating its functionality.

CHAPTER 1. INTRODUCTION TO BIOLOGICAL DIFFUSION AND THE DIFFUSION-LIMITED REACTION RATE

Physical measurements of biological systems are frequently performed in test tubes, removing the biomolecules from their native environment, Figure 1.1. When biomolecules are isolated from their environments, they often lose some of their essential features, such as 3-dimensional structure and biological function [1-5]. A large area of research today is underway to perform physical measurements inside living cells where molecules retain their structure and function [6-16]. The project described in Chapters 1-4 of this thesis, is the development of an experimental system to measure reaction rates inside living cells.
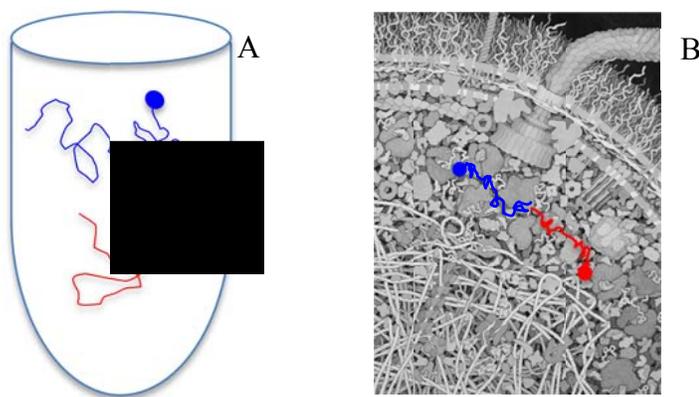
Figure 1.1 Diffusion of two macromolecules (blue and red) in a A) homogeneous environment, such as buffer in a test tube, and B) the heterogeneous environment of an *E. coli* cell. The cartoon of the *E. coli* cell is from reference [17].

1.1 The Diffusion-Limit

In order for a reaction to take place, the reactants must first be transported to a common location; this transport is achieved through either diffusion or active transport. The speed of this transport, consequently, sets an upper bound on the kinetics of the reaction. Many biological reactions are transport-limited, meaning that their reaction kinetics closely approach their transport kinetics [18]. Active transport plays a large role in some scenarios such as transporting molecules across large cells [19], enriching regions of the cell with particular molecules [20], and the transport of vesicles [21]. However, even when active transport mechanisms are present, diffusion still provides the final search mechanism for biomolecules in their local space to find their reaction partners. In many cases however, diffusion is the exclusive mechanism of transport. Since diffusion always plays a role in transport, its contribution is an important factor in the kinetics of biological reactions.

Often times, biological reactions involve many intermediate steps. Typically, the reactants must first be transported to a common location, then they must rotate into the proper orientation so that the reaction centers of the molecules are properly aligned. The overall reaction rate can be affected by interactions between the reactants in these intermediate steps [22]. Specific reaction mechanisms vary from reaction to reaction but generally, reactions can be classified into two categories based on their rate-limiting step: diffusion-limited and reaction-limited. If the overall rate is limited by the chemical process, the rate is reaction-limited. When the overall rate is limited by diffusional transport, the rate is diffusion-limited.

The simplest reaction mechanism of two reactants forming a single product is given by,

$$A + B \xrightarrow{k} C \tag{1.1}$$

where A and B are reactants forming product C with rate constant, k. The reaction rate of this second order reaction is

$$Rate = k[A][B] \tag{1.2}$$

where k is the rate constant and [A] and [B] are concentrations of reactants A and B. In a diffusion-limited reaction, the time for A and B to react and produce C is negligible compared to the diffusion of A and B. Therefore, k is primarily affected by the diffusion time and can be noted as the diffusion-limited rate constant.

Since many biological reactions are diffusion-limited, a prediction of the diffusion limit gives a good approximation of the reaction rate. For other biological reactions, in which diffusion plays a crucial role, the diffusion-limit sets the upper bound for the reaction rate since molecules cannot react any faster than they encounter one another.

1.2 Homogeneous Diffusion

1.2.1 Random Walk

Microscopically, diffusion is the random migration of molecules as a result of their kinetic energy. In liquids, the mean free path is much shorter than that for gases; diffusion in liquids follows Brownian motion. For an isotropic random walk, the average position of the molecules does not change as a function of time; the average displacement

of the particles is always zero. A way to estimate how much the particles spread is using the mean square displacement, $\langle x^2 \rangle$. In one dimension, the mean-square displacement is given by

$$\langle x^2 \rangle = 2D_0 t \tag{1.3}$$

where D_0 is the translational diffusion coefficient for the molecule in a particular solvent at a particular temperature. For two or three dimensions, the mean-square displacement can be derived, using the fact that the square of the displacement from the origin r^2 equals $x^2 + y^2$ in two dimensions and $x^2 + y^2 + z^2$ in three dimensions. The mean-square displacement therefore is given by

$$\langle x^2 \rangle = 4D_0 t \tag{1.4}$$

in 2D, and

$$\langle x^2 \rangle = 6D_0 t \tag{1.5}$$

in 3D.

1.2.2 Fickian Diffusion

Fick's law is a classical theory that describes diffusion. Fick's first law states that the net flux of a solute is proportional to the spatial concentration gradient with proportionality constant, $-D_0$.

$$J_x = -D_o \frac{\partial \phi}{\partial x} \tag{1.6}$$

where J is the net diffusion flux, D_0 is the diffusion coefficient, and ϕ is the concentration.

Fick's second law predicts how the concentration changes with time as a result of diffusion.

$$\frac{\partial \phi}{\partial t} = D_0 \frac{\partial^2 \phi}{\partial x^2} \tag{1.7}$$

The solution to Fick's law describes how diffusing particles distribute themselves in space and time in a homogeneous environment. Diffusion that obeys Fick's laws is referred to as "Fickian diffusion".

1.2.3 Translational and Rotational Diffusion Coefficients

Both the translational diffusion coefficient, D_0, and the rotational diffusion coefficient, D_r, can be obtained from the Einstein-Smoluchowski equation

$$D_0 = \frac{k_B T}{f_t} \tag{1.8}$$

and

$$D_r = \frac{k_B T}{f_r} \tag{1.9}$$

where k_B is the Boltzmann constant, T is the temperature, and f_x is the frictional drag coefficient. According to Stoke's relation, for a particle that can be represented as a uniform sphere in an incompressible, viscous solvent, the translational frictional drag coefficient is

$$f_t = 6\pi \eta a \tag{1.10}$$

where η is the viscosity of the solvent and a is the radius of the sphere. The frictional drag coefficient for rotation, for the same conditions listed above, is

$$f_r = 8\pi \eta a^3 \tag{1.11}$$

For both instances, the frictional drag coefficient, is directly proportional to the viscosity of the solvent. Therefore, changes in viscosity result in proportional changes in both the translational and rotational diffusion coefficient.

1.2.4 The Smoluchowski Equation

Smoluchowski's theory gives an upper bound for the second order rate constant for a diffusion-limited reaction. The equation was derived by Smoluchowski in 1917 [23] by solving Fick's laws for diffusion. The classical Smoluchowski equation is

$$k = 4\pi D_0 R \qquad\qquad (1.12)$$

This equation considers the diffusion-limited reaction of two reacting molecules A and B; both are considered to be spheres with a reaction radius, R. Any barrier, due to the formation of a transition state, will result in a slower reaction rate. Therefore, for any reaction, the Smoluchowski equation is an upper bound, and the rate constant can be expressed as an inequality,

$$k_0 \leq 4\pi D_0 R. \qquad\qquad (1.13)$$

1.3 Biological Diffusion

After reviewing measurements of biological diffusion, Verkman wrote "The view of the cell interior has evolved from that of a viscous gel to that of a watery but crowded compartment" [24]. In this view, at short length scales, rapid, water-like, diffusion predominates and at longer length scales, as obstacles are encountered, diffusion becomes slowed. This heterogeneity in diffusion cannot be explained using Fickian diffusion relations.

Experimental data has shown that translational diffusion coefficients measured *in vivo* are reduced 3-100 fold as compared to those measured *in vitro* [6-15]. This large reduction in diffusion was once characterized by a high viscosity of the biological medium [14]. However, this interpretation is now precluded by Verkman's view of the biological environment as a "watery but crowded" compartment. Direct evidence for the heterogeneity of biological diffusion is provided by comparing the translational and rotational diffusion coefficients. Equations 1.8 and 1.9 show that both the translational and rotational diffusion coefficients will be proportionally effected as a result of a more viscous medium. However, rotational diffusion coefficients measured in cells are maintained at 70-90% of their values in buffer [25-27], whereas translational diffusion coefficients drop to 0.1-20% of their values in buffer [6-15].

This contrast between translational and rotational diffusion coefficients can be interpreted as meaning that the solution seen by a molecule in its immediate vicinity has the same viscosity as water. For rotation, this "watery" environment is all that matters. However, as the molecule translates over longer length scales, it encounters obstacles that hinder its translational diffusion. The obstacles in biological environments are usually macromolecules, cytoskeletons, filaments, microtubules, and membranous boundaries.

The size dependence for biological diffusion also provides evidence that translational diffusion in biological environments is impeded by obstacles. Small molecules are more weakly affected by macromolecules in the biological environment than larger ones.

1.3.1 Diffusion-Reaction Coupling in Biological Systems

Smoluchowski's equation sets an upper limit for reaction rates for systems diffusing in a homogeneous environment (equation 1.12). If this equation is applied to biological diffusion, k, the biological reaction rate, bound by Smoluchowski's inequality (equation 1.13), D_0 is replaced with D, the biological diffusion coefficient. This predicts the biological diffusion limit *in vivo* as,

$$k \leq 4\pi DR. \tag{1.14}$$

This version of Smoluchowski's equation predicts that the 3-100 fold reduction in the diffusion coefficient will cause a proportional decrease in the rate constant. However, Brownian dynamics simulations find that this equation does not hold [28]. Also, a recent study of protein dimerization in HeLa cells and *in vitro* found that the rate constants were similar [29]. This recent data shows that Smoluchowski's does not apply in biological systems. This is a result of the fact that Smoluchowski's equation was derived under the assumption that diffusion is homogeneous at all length scales. However, the diffusion limited reaction rate is not linearly related to the diffusion coefficient. Some theory has been developed to predict reaction rates in biological environments [30-32]. Even fewer experiments have been done to correlate diffusion and reaction rates in living cells [29]. There is thus a great need for more experiments to develop this relation.

1.4 References

1. Gething, M. and J. Sambrook, *Protein Folding in the Cell.* Nature, 1992. **355**: p. 33-45.

2. Hartl, F., *Molecular Chaperones in Cellular Protein Folding.* Nature, 1996. **381**: p. 571-580.

3. Zimmerman, S.B. and A.P. Minton, *Macromolecular crowding: biochemcial, biophysical, and physiological consequences.* Annual Review of Biophysics and Biomolecular Structure, 1993. **22**: p. 27-65.

4. Dobson, C., *Protein Folding and misfolding.* Nature, 2003 **426**: p. 884-890.

5. Ellis, R.J., *Macromolecular Crowding: Obvious but Underappreciated.* Trends Biochem Sci, 2001 **26**(10): p. 597-604.

6. Dix, J.A. and A.S. Verkman, *Crowding effects on diffusion in solutions and cells.* Annu Rev Biophys, 2008. **37**: p. 247-63.

7. Luby-Phelps, K., et al., *Hindered Diffusion of inter tracer particles in the cytoplasm of mouse 3t3 cells.* Proc Natl Acad Sci U S A, 1987. **84**: p. 4910-4913.

8. Arrio-Dupont, M., et al., *Diffusion of fluorescently labeled macromolecules in cultured muscle cells.* Biophysical Journal, 1996. **70**: p. 2327-2332.

9. Seksek, O., J. Biwersi, and A.S. Verkman, *Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus.* Journal of Cell Biology, 1997. **138**: p. 131-142.

10. Elowitz, M.B., et al., *Protein mobility in the cytoplasm of E coli.* Journal of Bacteriology, 1999. **181**: p. 197-203.

11. Mullineaux, C.W., et al., *Diffusion of green fluorescent protein in three cell environments in escherichia coli.* Journal of Bacteriology, 2006. **188**: p. 3442-3448.

12. Dayel, M.J., E.F.Y. Hom, and A.S. Verkman, *Diffusion of green fluorescent protein in the aqueous-phase lumen of endoplasmic reticulum.* Biophysical Journal, 1999. **76**(5): p. 2843-2851.

13. Partikian, A., et al., *Rapid diffusion of green fluorescent protein in the mitochondrial matrix.* Journal of Cell Biology, 1998. **140**: p. 821-829.

14.   Wojcieszyn, J.W., et al., *Diffusion of injected macromolecules within the cytoplasm of living cells.* Proc Natl Acad Sci U S A, 1981. **78**: p. 4407-4410.

15.   Slade, K.M., et al., *Effects of recombinant protein expression on green fluorescent protein diffusion in escherichia coli.* Biochemistry, 2009. **48**: p. 5083-5089.

16.   Meacci, G., et al., *Mobility of min-proteins in escherichia coli measured by fluorescence correlation spectroscopy.* Physical Biology, 2006. **3**: p. 255-263.

17.   Goodsell, D.S., *Escherichia coli.* Biochem Mol Biol Educ, 2009. **37**(6): p. 325-32.

18.   Berg, O. and P. von Hippel, *Diffusion-controlled macromolecular interactions.* Annual Review of Biophysics and Biophysical Chemistry, 1985. **14**: p. 131-160.

19.   Zhang, H.L.L., et al., *Active transport of the survivla motor of neuron protein and the role of exon-7 in cytoplasmic localization.* Journal of Neuroscience, 2003. **23**: p. 6627-6637.

20.   Andrei, M.A., et al., *A role for eif4e and eif4e-transporter in targeting mrnps to mammalian processing bodies.* Rna-a Publication of the RNA Society, 2005. **11**: p. 717-727.

21.   Zenisek, D., J.A. Steyer, and W. Almers, *Transport, capture and endocytosis of single synaptic vesicles at active zones.* Nature, 2000. **406**: p. 849-854.

22.   Noyes, R.M., *Effects of diffusion rates on chemical kinetics.* Progress in Reaction Kinetics, 1961. **1**: p. 129-161.

23.   von Smoluchowski, M., *Versuch einer mathematischen theorie der koagulation-skinetik kolloider losungen.* Zeitschrift fuer Physikalische Chemie (Leipzig), 1917. **92**: p. 129.

24.   Verkman, A.S., *Solute and Macromolecule diffusion in cellular aqueous compartments.* Trends Biochem Sci, 2002. **27**(1): p. 27-33.

25.   Periasamy, N., M. Armijo, and A.S. Verkman, *Picosecond rotation of small polar fluorophores in the cytocol of sea urchin eggs.* Biochemistry, 1991. **30**: p. 11836-11841.

26.   Fushimi, K. and A.S. Verkman, *Low viscosity in the aqueous domainof cell cytoplasm measured by picosecond polarization microfluorimetry.* Journal of Cell Biology, 1991. **112**(4): p. 719-725.

27.   Swaminathan, R., C.P. Hoang, and A.S. Verkman, *Photobleaching recovery and anisotropy decay of green fluorescent protein gfp-s65t in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion.* Biophys J, 1997. **72**(4): p. 1900-1907.

28.   Klann, M.T., A. Lapin, and M. Reuss, *Agent-based simulation fo reactions in the crowded and structured intracellular environment: Influence of mobility and location of the reactants.* BMC Systems Biology, 2011. **5**.

29.   Phillip, Y., V. Kiss, and G. Schreiber, *Protein-binding dynamics imaged in a living cell.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1461-1466.

30.   Li, R., J.A. Fowler, and B.A. Todd, *Calculated Rates of Diffusion-Limited Reactions in a Three-Dimensional Network of Connected Compartments: Application to Porous Catalysts and Biological Systems.* Physical Review Letters, 2014. **113**.

31.   Haugh, J.M., *Analysis of reaction-diffusion systems with anomolous subdiffusion.* Biophyscial Journal, 2009. **97**: p. 435-442.

32.   Hellmann, M., D.W. Heermann, and M. Weiss, *Anomolous reaction kinetics and domain formation on crowded membranes.* Epl, 2011. **94**.

CHAPTER 2. SYSTEM DEVELOPMENT FOR MEASUREMENT OF BIOLOGICAL
DIFFUSION-LIMITED REACTION KINETICS

Traditionally biochemical reaction parameters, such as reaction rates, are studied

in vitro, in controlled buffer solutions. However, real biochemistry takes place *in vivo,*

which is a crowded, complex environment [1]. Attempts have been made to bridge the

gap between the *in vitro* and *in vivo* environments by conducting studies in the presence

of crowding agents, such as synthetic polymers or proteins [2]. Still, the actual cellular

environment is too complex to be able to be modeled appropriately by this approach.

Therefore, measurements inside living cells are required to fully understand the

biochemistry that occurs [3]. With only one notable exception, little work has been done

to measure reaction dynamics inside living cells [1].

2.1 System Exploration

One of the simplest biological reactions to be thought of is a protein dimerization.

As reaction candidates were considered, there were two key factors to be considered.

First, this reaction should be exogenous to the biological environment in which it is going

to be measured; this allows only proteins introduced into the environment to react and

thus no need to correct for possible reactions with intracellular protein. Second, there

needs to be a way for the reaction to be initiated once the system has reached equilibrium; this prevents the reaction from occurring prior to observation.

An initial concept for studying this type of reaction was to cage the monomeric states of the dimer forming protein. These "cages" disrupt the structure of the peptide and upon removal the protein folds back to its native state [4-7]. After incubation in the biological material, a pulse of ultra-violet (UV) light can be used to "uncage" the monomer and initiate dimerization. One reaction of this type was initially considered as a candidate for kinetics measurements; the Alzheimer Aβ peptide, which has previously been "caged" in an unfolded state [7]. However, this protein is known to form higher order oligomers [8, 9] and would thus not be useful for studying a simple dimerization.

Another approach to the caging strategy is to choose an inducible dimerization and cage the induction agent. This way, the proteins as well as the caged inducer can be introduced into the biological environment and allowed to equilibrate before initiating the dimer reaction by uncaging and monitoring dimer formation. Common dimerization induction partners are the Ca^{2+} ion and acid, or H^+. One reaction considered as a candidate for the measurement of kinetics, was the acid induced dimerization of Troponin C from rabbit skeletal muscle; Troponin C is known to be a monomer at pH 7.5, but at pH 5.4, it exists as a dimer [10]. This same protein will also form a dimer in the presence of excess Ca^{2+} ions [11]. Two other calcium-induced reactions were also candidates; the dimerization of isoforms of phospholipase A_2 isolated from snake venom [12, 13] and the calcium induced dimer formation of calmodulin [14]. Both caged calcium and caged H^+ are commercially available; however experimental biological systems for measurement of Troponin C, phospholipase A_2, and calmodulin dimerization have not yet been

explored, as well as whether the caging strategy will be successful for prevention of their dimer formation.

A more promising candidate was discovered which is based on small-molecule induced dimerization. The proteins FKBP12, the FK506 binding protein 12, and FRB, the rapamycin binding domain of mTOR, form a dimer in the presence of the small molecule rapamycin. Previously, this reaction has been demonstrated to respond to the caging effect of rapamycin, with a slight modification of FKBP12 [15, 16]. This reaction has also been shown to function in *E. coli* cells [17]. Table 2.1 is a summary of reaction candidates for the study of biological diffusion-limited kinetics. The heterodimerization of FKBP12 and FRB in the presence of rapamycin was selected for use in this study. This reaction has the ability to be exogenously expressed in *E. coli* cells and the reaction can be initiated by the UV uncaging of rapamycin.

Table 2.1 Possible dimerization reactions considered as candidates for biological kinetic measurements

Protein	Induction Method
Alzheimer peptide Aβ	caged peptide
Troponin C	H^+ or Ca^{2+}
Phospholipase A_2	Ca^{2+}
Calmodulin	Ca^{2+}
FKBP/FRB	rapamycin

2.2 Measures of Dimerization

Once the dimer system was selected, a method to study the dynamics of the reaction, ie formation of the dimer, was required. The dimerization of FKBP12 and FRB has previously been monitored using a variety of fluorescence measurements [16, 18, 19]. In one example, the FRB monomer is tethered to the cell membrane and FKBP12 is fused to a variant of yellow fluorescent protein (YFP). Upon dimer formation, FKBP12 is localized to the membrane; as is the YFP variant and dimerization is monitored by following the localization of YFP fluorescence to the membrane [18]. Similarly, Karginov et. al. use fluorescence localization of GFP fluorescence to kinase receptors inside cells to monitor dimer formation [16]. Finally, FKBP12 and FRB are fused to fluorescent proteins Cerulean and Venus, respectively, and their dimerization is monitored by Förester Resonance Energy Transfer (FRET) [19].

A most idealistic mechanism for measuring the diffusion-limited reaction kinetics is when both dimer-forming proteins are allowed to diffuse freely through the biological environment. For this reason, the reaction system was developed using fluorescent proteins Cerulean and Venus for monitoring dimer formation by FRET.

2.3 FKBP-FRB System Description

The system chosen to study the diffusion-limited reaction kinetics in living cells was the rapamycin induced, heterodimerization of FKBP12 and FRB. Rapamycin is a product of the bacterium *Streptomyces hygroscopicus* that was found in a soil sample from Easter Island that is used as an immunosuppressive drug often administered to organ transplant patients to prevent transplant rejection [20-24]. FKBP12 is a protein in

the immunophilin family. This family is capable of binding the immunosuppressive

agent FK506 and serves as a cytosolic receptor for rapamycin [25]. FRB is the

rapamycin binding domain of mTOR, which acts as a target for the immunosuppressive

effects of the FKBP12-rapamycin complex. The structure of the FKBP-rapamycin-FRB

complex is shown in Figure 2.1.
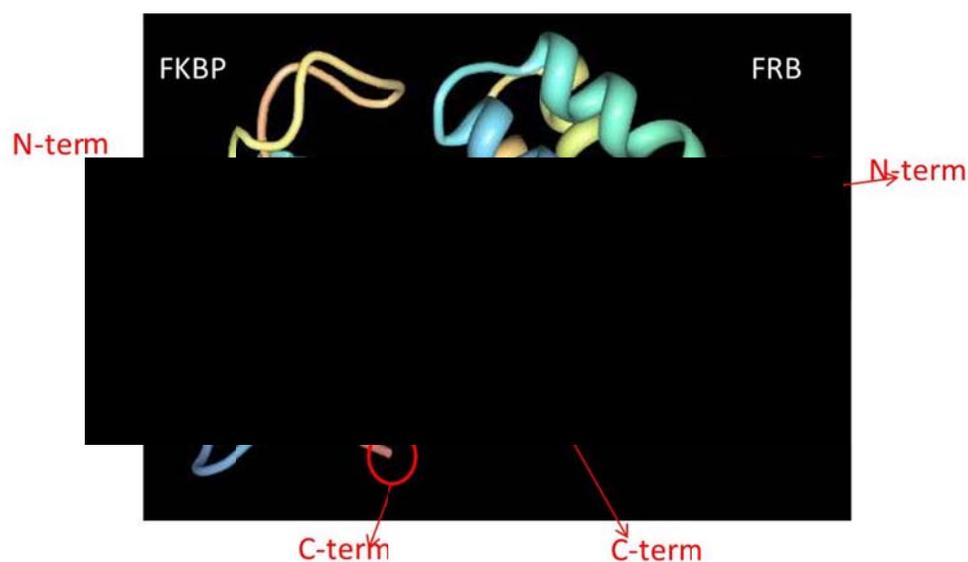
Figure 2.1. The crystal structure of the FKBP12-rapamycin-FRB complex with labeled
N- and C-termini, represented as ribbons. PDB file 1FAP [26].

If caged rapamycin is to be used to inhibit dimer formation, an engineered version

of FKBP12 is required [15]; without this altered form of the FKBP12 protein, the dimer

between the latter and FRB will still form. A simple truncation of the FKBP12 protein

leads to a form that will not bind caged rapamycin. While caged rapamycin is not

commercially available, its synthesis has been reported [16]. Prior to use of caged

rapamycin, experiments are to be carried out with the standard dimer pair and uncaged rapamycin, which is commercially available.

The rate of dimer formation will be monitored by FRET between fluorescent proteins Cerulean and Venus, fused to FKBP12 and FRB, respectively. Cerulean is a variant of cyan fluorescent protein and Venus is a variant of yellow fluorescent protein; this pair is known to be a good FRET pair for measurement [27].

2.4 References

1. Phillip, Y., V. Kiss, and G. Schreiber, *Protein-binding dynamics imaged in a living cell.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1461-1466.

2. Zhou, H.X., G. Rivas, and A.P. Minton, *Macromolecular crowding and confinement: Biochemical, biophysical and potential physiological consequences.* Annual Review of Biophysics, 2008. **37**: p. 375-397.

3. Gierasch, L.M. and A. Gershenson, *Post-reductionist protein science, or putting Humpty Dumpty back together again.* Nat Chem Biol, 2009. **5**: p. 774-777.

4. Chen, R.P., et al., *Measuring the refolding of beta-sheets with different turn sequences on a nanosecond time scale.* Proc Natl Acad Sci U S A, 2004. **101**(19): p. 7305-10.

5. Hansen, K.C., et al., *A method for photoinitiating protein folding in a nondenaturing environment.* J Am Chem Soc, 2000. **122**: p. 11567-11568.

6. Rock, R.S. and S.I. Chan, *Preparation of a water soluble "cage" based on 3',5'-dimethoxybenzoin.* JACS, 1998. **120**: p. 10766-10767.

7. Chen, H.L., et al., *Studying submicrosecond protein folding kinetics using a photolabile caging strategy and time-resolved photoacoustic calorimetry.* Proteins, 2010. **78**(14): p. 2973-83.

8. Mandal, P., N. Eremina, and A. Barth, *Formation of two oligomers in the early phase of pH-induced aggregation of the Alzheimer Abeta(12-28) peptide [corrected].* J Phys Chem B, 2012. **116**(41): p. 12389-97.

9. Peralvarez-Marin, A., A. Barth, and A. Graslund, *Time-resolved infrared spectroscopy of pH-induced aggregation of the Alzheimer Abeta(1-28) peptide.* J Mol Biol, 2008. **379**(3): p. 589-96.

10. Wang, C., J. Lebowitz, and H.C. Cheung, *Acid-Induced Dimerization of Skeletal Troponin C.* Proteins: Struct., Funct., Genet., 1989. **6**: p. 424-430.

11. Slupsky, C.M., et al., *Calcium-Induced Dimerization of Troponin C: Mode of Interaction and Use of Trfluoroethanol as a Denaturant of Quaternary Structure.* Biochemistry, 1995. **34**: p. 7365-7375.

12. Jabeen, T., et al., *Crystal structure of a calcium-induced dimer of two isoforms of cobra phospholipase A2 at 1.6 A resolution.* Proteins, 2005. **59**(4): p. 856-63.

13. Sanchez, S.A., et al., *Solution and Interface Aggregation Stated of Cotalus atrox Venom Phospholipase A$_2$ by Two-Photon Excitation Fluorescence Correlation Spectroscopy.* Biochemistry, 2001. **40**: p. 6903-6911.

14. Shishido, H. and S. Maruta, *Engineering of a novel Ca(2)(+)-regulated kinesin molecular motor using a calmodulin dimer linker.* Biochem Biophys Res Commun, 2012. **423**(2): p. 386-91.

15. Karginov, A.V., et al., *Engineered allosteric activation of kinases in living cells.* Nat Biotechnol, 2010. **28**(7): p. 743-7.

16. Karginov, A.V., et al., *Light regulation of protein dimerization and kinase activity in living cells using photocaged rapamycin and engineered FKBP.* J Am Chem Soc, 2011. **133**: p. 420-423.

17. Davis, J.H., T.A. Baker, and R.T. Sauer, *Small-molecule control of protein degradation using split adaptors.* ACS Chem Biol, 2011. **6**(11): p. 1205-13.

18. Umeda, N., et al., *A photocleavable Rapamycin conjugate for spatiotemporal control fo small GTPase Activity.* J Am Chem Soc, 2011. **133**: p. 12-14.

19. Chen, H., H.L. Puhl, 3rd, and S.R. Ikeda, *Estimating protein-protein interaction affinity in living cells using quantitative Forster resonance energy transfer measurements.* J Biomed Opt, 2007. **12**(5): p. 054011.

20. Banaszynski, L.A., C.W. Liu, and T.J. Wandless, *Characterization of the FKBP-Rapamycin-FRB Ternary Complex.* J Am Chem Soc, 2005. **127**: p. 4715-4721.

21. Findlay, J.A. and L. Radics, *On the chemistry and high field nuclear magnetic resonance spectroscopy of rapamycin.* Canadian Journal of Chemistry, 1980. **58**: p. 579-590.

22. Swindells, D.C.N., P.S. White, and J.A. Findlay, *The X-ray crystal structure of rapamycin, C51H79NO12.* Canadian Journal of Chemistry, 1978. **56**: p. 2491-2492.

23. Sehgal, S.N., H. Baker, and C. Vezina, *Rapamycin, a new antifungal antibiotic II. Fermentation, isolation and characterization.* Journal of Antibiotics, 1975. **28**(10): p. 727-732.

24. Vezina, C., A. Kudelski, and S.N. Sehgal, *Rapamycin, a new antifungal antibiotic I. Taxonomy of the producing streptomycete and isolation of the active principle.* Journal of Antibiotics, 1975. **28**(10): p. 721-726.

25.   *FKBP12, GST-Fusion, Human, recombinant, E coli*. [cited 2012 November 2012].

26.   Choi, J., et al., *Structure of the FKBP12-Rapamycin Complex Interacting with the Binding Domain of Human FRAP.* Science, 1996. **273**: p. 239-242.

27.   Chen, H., et al., *Measurement of FRET efficiency and ratio of donor to acceptor concentration in living cells.* Biophys J, 2006. **91**(5): p. L39-41.

CHAPTER 3. STOICHIOMETRIC COEXPRESSION OF FKBP-CERULEAN AND
FRB-VENUS IN *E. COLI*

Co-expression of FKBP-cerulean and FRB-venus in *E. coli* at stoichiometric

levels is the first key step of being able to study their diffusion and reaction rate inside a

biological environment. The ratio of expression was determined using both fluorescence

spectroscopy and microscopy. Achieving a near 1 to 1 expression ratio of FKBP-

Cerulean and FRB-Venus is necessary to perform further experiments measuring

diffusion and reaction rates. Without a close to 1 to 1 stoichiometry, the reaction could

be limited by the concentration of the reactants and therefore not produce an accurate rate

determination.

3.1 Experimental Methods

All work toward protein expression and purification was done in the laboratory of

Dr. Chittaranjan Das in the Purdue Department of Chemistry. *E. coli* cells were always

cultured in his lab and taken to a secondary location for either microscopy or

spectroscopy.

3.1.1 Genes and Expression Vector selection

While expression of a single protein in an individual cell is somewhat trivial, co-expression of 2 different proteins has proven to be more difficult. After initially achieving individual expression of FKBP-cerulean and FRB-venus using custom-made genes (Life Technologies), it was discovered that the selected expression vectors were not compatible in a single *E. coli* cell; each cell favored one of the plasmids and would only express one of the proteins. This incompatibility was due to a site on each plasmid called an origin of replication; a section of the genetic sequence on the plasmid where replication is initiated for the propagation of the DNA plasmid in living *E. coli* cells [1]. Upon careful consideration of this and other key factors, such as antibiotic resistance and plasmid copy number, expression plasmids, pET-45(b) and pCDF-1b, were chosen. The pET-45(b) vector has a copy number of ~40, an ampicillin antibiotic resistance, and the origin of replication is pBR322. The pCDF vector has a copy number of 20-40, with a streptomycin resistance and the replication of origin, CDF (which is compatible with pBR322). New genes for the fusion proteins previously expressed in HeLa cells [2] were obtained, as well as expression plasmids pET-45(b) and pCDF-1b.

Prior to protein expression, the genes for FKBP-Cerulean and FRB-Venus were subcloned into pET-45(b) and pCDF-1b, respectively. Briefly, the genes are amplified using a polymerase chain reaction (PCR) procedure. The products from PCR, along with the expression vectors are digested using appropriate restriction enzymes. This step cuts the expression vectors leaving "sticky" ends, or DNA base overhangs that are complimentary to those on the genes. Once both the genes and the expression vectors have sticky ends, they undergo a ligation reaction which will cause the complimentary

ends to "stick" together and the gene will be inserted into the proper expression vector. This reaction is allowed to progress overnight at 4°C; when the ligation is transformed into *E. coli* cells which are plated onto an agar plate supplemented with the corresponding antibiotic. To confirm successful subcloning, a PCR test was performed on select colonies from the agar plate. The PCR product was then run on an agarose DNA gel to confirm the length of the cloned gene. Figure 3.1 shows the DNA gels for both FKBP-Cerulean cloned into pET-45(b) and FRB-Venus cloned into pCDF-1b. The genes were also confirmed by DNA sequencing using the Purdue University Genomics Core Facility.
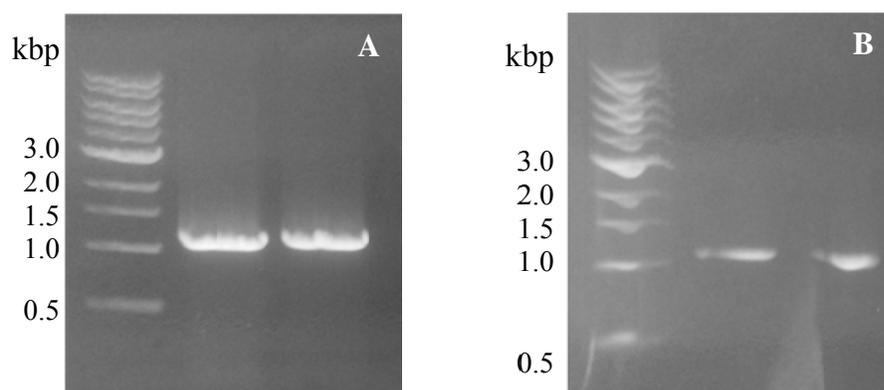
Figure 3.1 DNA gels confirming successful subcloning of A) FKBP-Cerulean into pET-45(b) and B) FRB-Venus into pCDF-1b. The first lane is a standard DNA ladder (labeled from 3.0kbp to 0.5kbp. Lanes 2 and 3 are the PCR products for 2 separate colonies. The FKBP-Cerulean gene is 1,113 base-pairs and FRB-Venus is 1,083 base-pairs.

3.1.2 Protein Expression and Co-expression

Once the genes have successfully been subcloned into the expression vectors, their ability to express the proteins needed to be tested. Expression of each protein

individually was tested prior to co-transformation. *E. coli* cells (Rosetta strain) were

transformed with 2μL plasmid DNA and plated on an agar plate containing the

appropriate antibiotic and allowed to grow at 37°C overnight. A single colony from the

plate was then added to a tube of 3mL of LB media supplemented with antibiotics and

again grown overnight on an orbital shaker (220rpm) at 37°C. From this overnight

culture, a fresh culture was innoculated, in LB media containing antibiotics, and allowed

to grow at 220 rpm and 37°C until optical density, OD, 0.5. Then 1mM Isopropyl β-D-1-

thiogalactopyranoside (IPTG) was added to the culture to induce protein expression and

the culture grows overnight at 220 rpm and 18°C. These cultures were then harvested by

centrifugation at 2900rpm for 10 minutes and resuspended in either 1mL phosphate

buffered saline for lysis by sonication or 1mL BugBuster lysis buffer (EMD Millipore

Corp.), which will lyse the cells at room temperature in 30 minutes. Protein expression

was confirmed in the lysates using a Cary Eclipse fluorescence spectrophotometer (see

section 3.1.4). The spectra for lysates containing FKBP-Cerulean and FRB-Venus are

shown in figure 3.2.

FKBP-cerulean & FRB-venus Expression
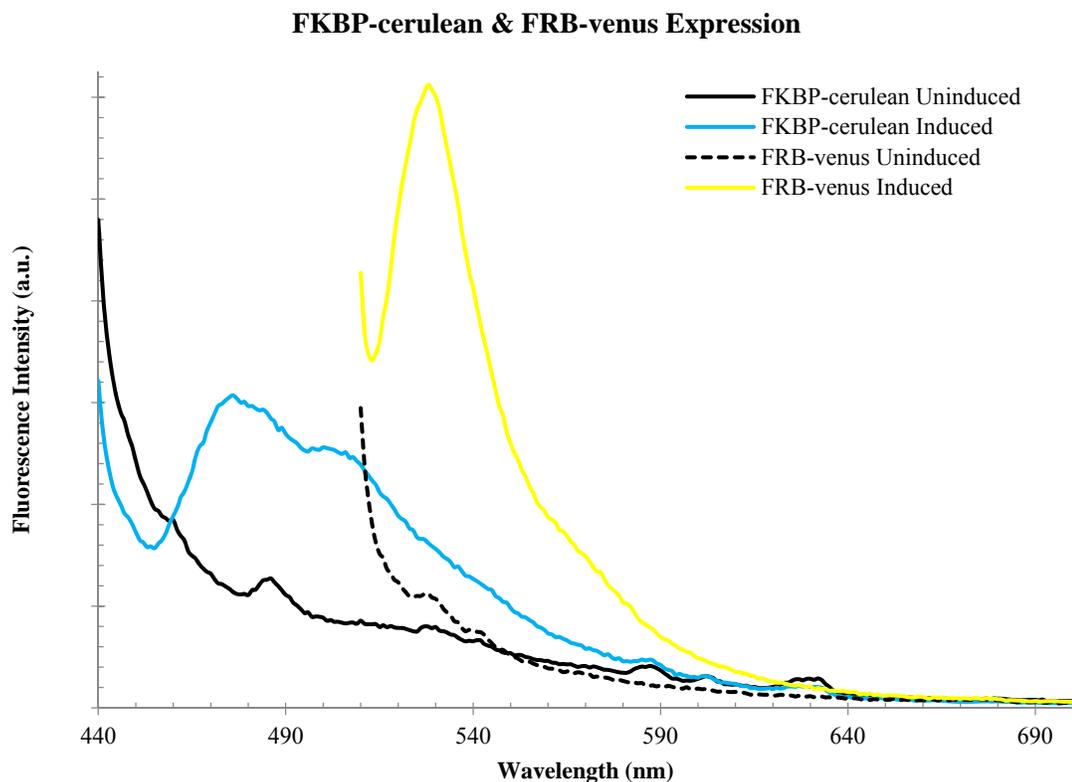
Figure 3.2 Fluorescence spectra of separate *E. coli* lysates testing for FKBP-Cerulean (blue curve) and FRB-Venus (yellow curve) expression. Also shown are cultures that were not induced with IPTG (black curves). Cerulean excitation was 430nm, and an emission peak is expected at 475nm. Venus excitation was 500nm and an emission peak is expected at 528nm.

For co-expression of both FKBP-cerulean and FRB-venus, the same procedure for

DNA transformation and cell culture described above is followed except: 5μL of each

DNA strain is added to the cells and a single agar plate is supplemented with both

antibiotics. Fluorescence spectroscopy is not sufficient, however, to determine successful

co-expression. A culture may contain a mixture of cells expressing either FKBP-cerulean

or FRB-venus – a spectrum for this type of mixture will look the same as a culture in

which individual cells are co-expressing both proteins. Fluorescence microscopy (see

section 3.1.5) is required to prove co-expression in individual cells by visualizing both fluorescent protein in each cell. Cultures grown for fluorescence microscopy measurements are taken to Dr. Ken Ritchie's lab for measurement.

3.1.3 Protein Purification

Expression vectors pET-45(b) and pCDF-1b both contain gene sequences such that the expressed proteins will contain an N-terminal 6x-Histidine tag for purification. FKBP-Cerulean and FRB-Venus were individually expressed as described in section 3.1.2, except the first overnight culture was 50mL and the expression cultures were 3x 1L cultures to achieve a higher protein yield. Cells from the 3 large cultures were harvested by centrifugation and resuspended in buffer for lysis by pressure using a French press. The lysate from the French press was then centrifuged in an ultra-centrifuge at 50,000rpm for 1 hour. The clean supernatant was then applied to a clean, charged Ni^{2+} column, which binds the histidine tag. An imidazole concentration gradient was then used to elute the protein; imidazole competed with histidine for column binding. All elutions were collected in 5mL fractions beginning with an initial wash to remove any un-adhered remains from the supernatant. The collected fractions were tested for protein and all those containing the protein, either FKBP-Cerulean or FRB-Venus, were pooled together and concentrated down to approximately 1.5mL. After concentration, a buffer exchange was performed to remove imidazole from the purified protein samples. Protein was then flash frozen in 100μL aliquots and stored at -80°C.

3.1.4 Fluorescence Spectroscopy

Fluorescence spectroscopic measurements were done using a commercial Cary Eclipse fluorescence spectrophotometer. Figure 3.3 shows a diagram of a Cary Eclipse fluorometer [3]. The excitation monochrometer automatically adjusts to its fixed position depending on the given excitation wavelength and the emission monochrometer scans the range of wavelengths selected for the emission spectrum. For all measurements, both the excitation and emission slits are set to 5nm; the slit widths determine the resolution of the spectrum. For fluorescent proteins Cerulean and Venus, which produce a reasonably high fluorescence signal, these slits can be set quite narrow. When measuring the fluorescence spectrum of Cerulean, the excitation wavelength was set to 430nm and the emission range was 440 to 650nm. For Venus, the excitation wavelength was 500nm with an emission range of 510 to 650nm. Sample cuvettes were filled with a minimum of 1mL of sample, either cell lysate or pure protein. Peaks indicating Cerulean and Venus expression are expected at 475nm and 528nm, respectively.
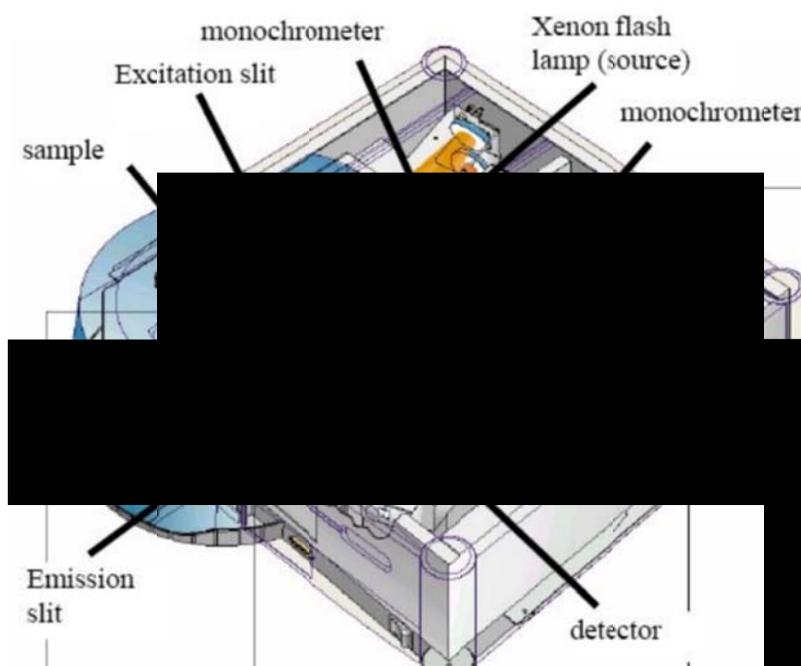
Figure 3.3. Diagram of a Cary Eclipse Fluorescence Spectrophotometer [3]. Once the excitation wavelength and emission range is selected, the excitation monochrometer (along the yellow light path) shifts to its appropriate fixed position for collection of the emission spectrum and the emission monochrometer (along the blue light path) will scan across the emission range selected for the spectrum. Both slits are set to a narrow 5nm, since cerulean and venus are highly fluorescent proteins.

3.1.5 Fluorescence Microscopy

All fluorescence microscopy was done using the optical setup in the lab of Dr.

Ken Ritchie in the Purdue Department of Physics and Astronomy.

3.1.5.1 Fluorescence Microscopy Optical Set-up

The optics of the microscopy system can be divided into 3 groups: an illumination

group, which contains all the optical components for sample excitation including the

lasers; the microscope group which includes primarily the inverted microscope and the

lenses which direct the light into the objective and back out to the detector; and the

imaging group which contains the optics that direct the light from the microscope group

into the CCD detector [4]. A schematic of the optical set-up is shown in figure 3.4

In the illumination group, excitation of Cerulean and Venus was achieved using

diode and argon-ion lasers with center wavelength lines of 445nm and 488nm,

respectively. Two lenses, labeled L1 and L2 in figure 3.2, and two periscopic mirrors

guide the collimated laser beams into the microscope. Two manually controlled shutters

control the length of time the sample is exposed. Notch, or single-band pass, filters,

labeled ExF1 and ExF2, isolate the excitation wavelength (removing any light that may

include shorter or longer wavelengths than what the filter is centered around). The cyan

laser, or 445nm, line includes a single edge dichroic mirror (instead of a plane mirror like

the blue laser, or 488), labeled Di-M1, which transmits the blue laser into the same

optical path as the cyan laser into the microscope.

The microscope group is made up primarily of the microscope itself which is an

inverted type microscope with a high NA (1.45) 100x, oil immersion objective. The

incident laser light is focused into the microscope with a lens, L3. Two more dichroic mirrors, Di-M2 and Di-M3, reflect the incident laser into the objective while allowing the fluorescent light from the sample to pass through. The fluorescent light from the sample is then focused and reflected with a plane mirror and another lens, L4, at an image plane outside of the microscope.

Finally, the imaging group is made up of the optics which receive the light from the microscope and focus it onto the CCD detector. The emitted light is collimated by a lens, L5, is split into Cerulean and Venus emission by a long-pass dichroic mirror, Di-M4, and is recombined by a short-pass dichroic mirror, Di-M5. While the emitted signal is split into its separate colors, the signal from Cerulean and Venus fluorescence emission is isolated using respective emission filters, EmF1 and EmF2.   A final lens focuses the collimated light onto the CCD detector.
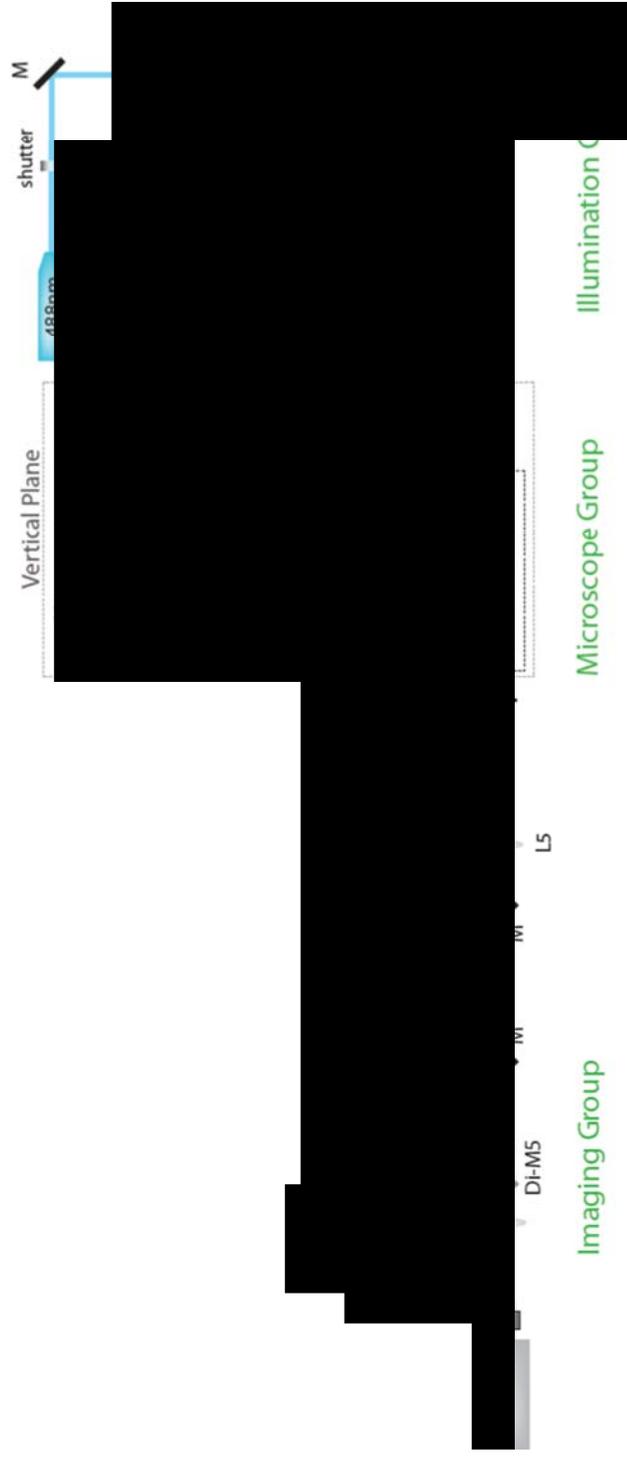
Figure 3.4. Schematic depiction of the fluorescence microscopy optical setup.[4] The system is made up of three groups: an illumination group, a microscope group and an imaging group.

3.1.5.2 Fluorescence Microscopy Sample and Chamber Preparation

Cultures were grown fresh and brought to the microscope on the day of the experiment as described in section 3.1.2.

Sample chambers for microscopy experiments were constructed using a commercially available flexiPERM® silicone ring that is reversibly adhered to a clean glass coverslip. Once the silicone ring is adhered, the clean glass is treated with poly-L-lysine (Sigma) which electrostatically adheres the cells to the glass. *E. coli* cells are then deposited into the sample chamber and allowed to settle for approximately 20 minutes prior to observation. To prevent a large excess signal coming from outside of the focal volume of the objective, unadhered cells are washed away with phosphate-buffered saline solution.

The flexiPERM® silicone rings are washed with 50% ethanol solution between samples and allowed to air dry before a new piece of glass is adhered. For added sterilization, the rings are autoclaved weekly at 250°F for 55 minutes.

Before use, the glass coverslips are cleaned thoroughly. They are first soaked in a 5% Contrad detergent solution for 24 hours followed by a high power sonication for 30 minutes. The glass is then washed 10 times with distilled water followed by a triplicate wash with ultra-high quality water. The same soak, sonicate, wash cycle is repeated using 1 M hydrochloric acid instead of Contrad detergent. After the final water wash, they are washed with methanol. The glass coverslips are stored in methanol until use.

The poly-L-lysine layer is deposited by pipetting 100μL of 0.1mg/mL poly-L-lysine solution onto the glass surface. The poly-L-lysine solution incubates on the glass

for 30 minutes to allow the polymer to coat the glass. Excess polymer is washed away with ultra-high quality water.

Once the surface is coated with poly-L-lysine, 200μL of the cell culture was pipetted into the chamber and allowed to settle for approximately 20 minutes. Excess cell culture with unadhered cells were washed away by aspirating with PBS three times. A final 400μL aliquot of fresh PBS is added to the chamber to prevent the cells from drying out during image collection.

<u>3.2 Data Collection and Analysis</u>

The expression levels of FKBP-Cerulean and FRB-Venus in *E. coli* were determined using both fluorescence spectroscopy and microscopy. While an absolute concentration scale was not determined, the concentration ratio was determined by normalizing the measured value with the ratio from a 1 to 1 mixture of the purified proteins.

3.2.1 Fluorescence Spectroscopic Data

Fresh samples were centrifuged and lysed as described in section 3.1.2. Spectra were collected, according to the settings described in section 3.1.4, of each of the lysates under both excitation wavelengths. The spectrum of a mixture of pure protein at a ratio of 1 to 1 FKBP-Cerulean to FRB-Venus is collected at each excitation for normalization. Figure 3.3 shows emission spectra of 4 cell lysates expressing protein, a strain of wild type *E. coli*, as well as the spectra of the pure proteins for both excitations. The concentration ratio of the expressed proteins in the lysate was determined by dividing the

area under the Cerulean emission curve by that of the Venus curve. Ratios are reported

as the area of FKBP-Cerulean emission to the area of FRB-Venus emission, normalized

by the same ratio of the purified protein mixture.
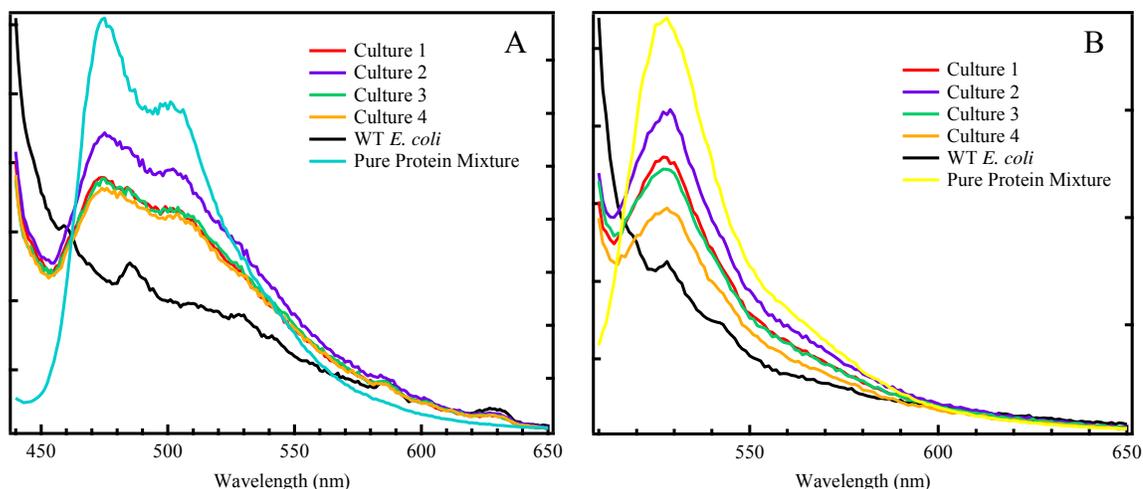
Figure 3.5. Fluorescence spectra of cell lysates and purified protein mixture collected at
A) 430 nm excitation and 440-650 nm emission range and B) 500 nm excitation and 510-
650 nm emission range using a commercial Cary Eclipse Fluorescence
Spectrophotometer. Cultures 1-4, shown in red, purple, green and orange, are induced
cultures expressing proteins FKBP-Cerulean and FRB-Venus. Wild type *E. coli,* shown
with the black curve, express no protein. The spectra of the purified protein mixture is
shown in blue and yellow, in A and B respectively.

3.2.2 Fluorescence Microscopic Data

Daily cultures were grown for each measurement, so that only fresh cells are

sampled to determine the expression levels. Samples were prepared and images were

collected according to the details outlined in section 3.1.5. Images were collected under a

single laser excitation at a time; no movement of the sample occurred between. Figure

3.6 shows example fluorescent images of cells transformed with DNA plasmids encoding

for expression of FRBP-Cerulean and FRB-Venus, as well as wild type *E. coli*, used as the control to compare fluorescence from protein expression to that of cellular auto-fluorescence. The normalization values of the protein mixture were obtained by pipetting a drop of the 1 to 1 mixture onto a clean glass slide (the same as used for the sample chamber) and leaving the focus of the objective set to the same position as required for cell imaging. This is simply used as a metric to correct for the differences in optics and inherent intensity levels of Cerulean and Venus. The normalization factor is also a ratio, so there is once again not an actual concentration measurement, but a ratio of intensities. The calculated intensity ratio of each cell analyzed is divided by the ratio of the pure mixture intensities.

Image analysis is carried out in Image J software (National Institute of Health). For each image, the background is removed using the available background subtraction tool in Image J, which removes smooth, continuous backgrounds from images[5]. The algorithm used is based on the "rolling ball" concept described in Stanley Sternberg's Biomedical image processing [5, 6]. Images obtained with Cerulean (445nm) and Venus (488nm) excitation are compared side by side and the intensity of cells expressing both FKBP-Cerulean and FRB-Venus was measured by manually outlining each cell in each image and using ImageJ's measure function to record the mean brightness calculated. Since cells in the images exhibit a variety of appearances and do not all look like what a "typical" *E. coli* is expected to, and some formed clumps of cells stacked with undefinable edges, the spots were categorized into 3 categories. The first category includes those which have a "typical" *E. coli* shape; having an oblong shape with the longest dimension being 1.5 times or greater than the short dimension with definable

edges. Category 2 included those cells that still show a definable edge, but do not exhibit

a large enough aspect ratio, i.e. they are more round than oblong. The last category

includes spots of any size that can be contributed to clumped cells, with no definable

edges.
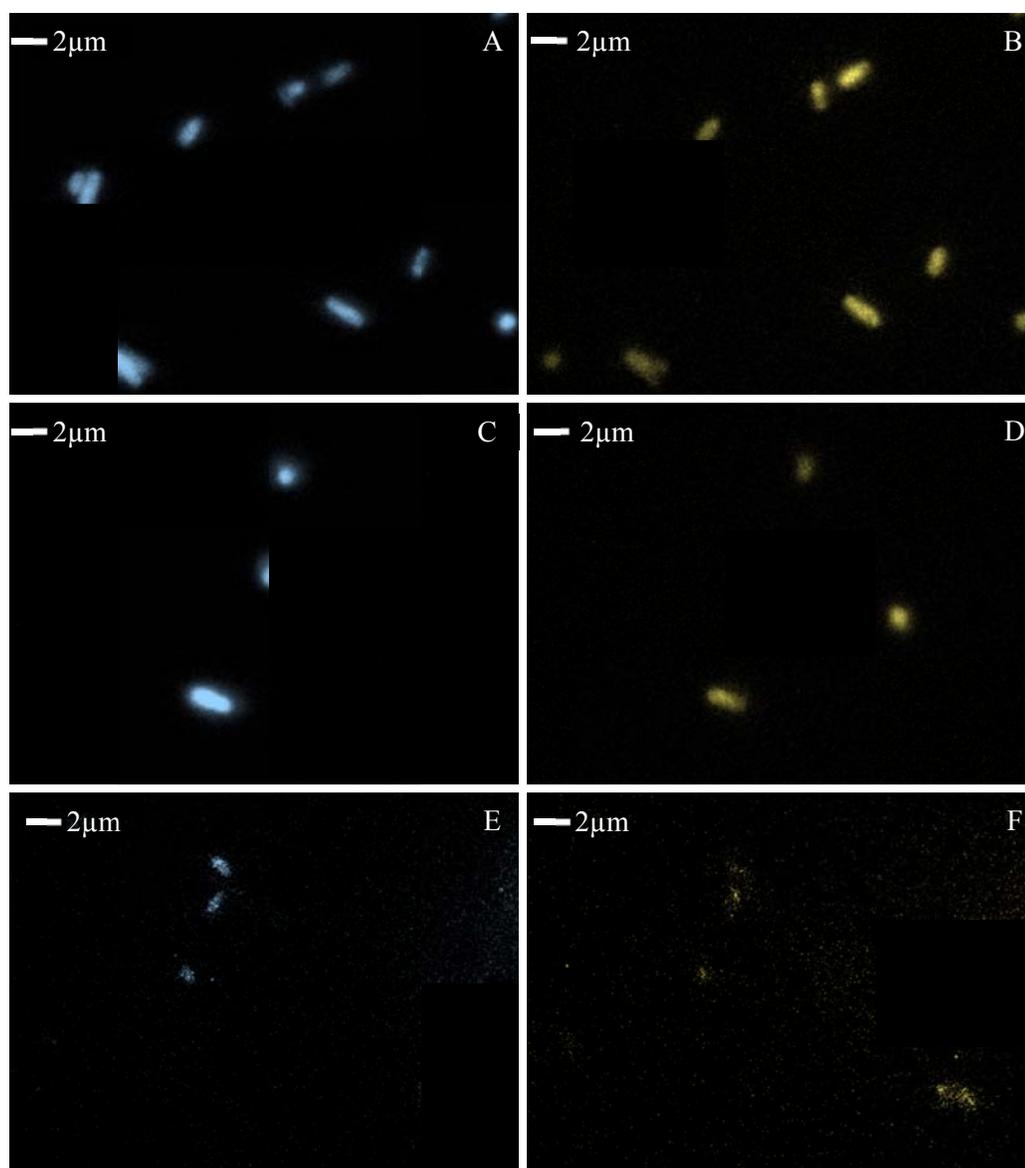
Figure 3.6. Fluorescent images of two sets of *E. coli* transformed with DNA plasmids encoding FKBP-Cerulean and FRB-venus (A-D) and a set of control, wild type, *E. coli* (E-F). Images in the left column (A, C, E) were excited with 445nm laser and collected through appropriate Cerulean emission filters, and are therefore showing expression of FKBP-Cerulean. Images in the right column (B, D, F) were excited with 488nm laser and collected through the Venus emission filters. These are therefore showing expression of FRB-Venus. Control images E and F were not transformed with plasmids containing the gens for FKBP-Cerulean and FRB-Venus and are thus not expressing any protein; the fluorescence observed here is from the cellular auto-fluorescence. Images are colored only for clarity purposes, blue shading for Cerulean and yellow shading for Venus.

3.3 Stoichiometric Ratio of FKBP-Cerulean and FRB-Venus Expressed in *E. coli* Determined by Fluorescence Spectroscopy

Fluorescence spectroscopy was used to determine the ratio of the concentration of expressed FKBP-Cerulean and FRB-Venus in *E. coli*. Table 3.1 shows the normalized ratio values obtained from four separate *E. coli* cultures transformed with DNA plasmids encoding the proteins (see section 3.2.1 for an explanation of how these values were determined). According to fluorescence spectroscopy, in a culture of *E. coli* transformed with DNA including the genes for FKBP-Cerulean and FRB-Venus, there is 1.721 ± 0.125 times as much Cerulean protein expression than Venus protein. This value is not expected to be as accurate as that obtained by fluorescence microscopy (result in section 3.4). The values obtained using spectroscopy were determined for an entire culture which may contain cells expressing only one of FKBP-Cerulean or FRB-Venus. These cells are contributing to the fluorescence emission spectra of only the protein they are expressing, resulting in a falsely high fluorescence emission. Without microscopy, there is no way to determine how many *E. coli* cells only expressing one protein are present in a given culture.

Table 3.1 Expression Ratio of FKBP-Cerulean to FRB-Venus in *E. coli* as determined by fluorescence spectroscopy

Normalized Ratio of Cerulean/Venus	
Average	1.721
Standard Deviation	0.125

3.4 Stoichiometric Ratio of FKBP-Cerulean and FRB-Venus Expressed in *E. coli*
Determined by Fluorescence Microscopy

The ratio of expression levels of FKBP-Cerulean and FRB-Venus in *E. coli* was
also determined using fluorescence microscopy. This method was expected to give a
more accurate concentration ratio since only cells evident to be expressing both FKBP-
Cerulean and FRB-Venus were evaluated. Cells in all images were divided into 3
categories (outlined in section 3.2.2). Histograms of each of the 3 categories are shown
in figure 3.7. Each of the histograms were fitted with a Gaussian. Both category 1 and 3
display Gaussian properties and the population in category 2 displays a plateau from
ratios 0.75-1.25; however, all three can be fit to a reasonable Gaussian. These Gaussian
fits are centered at 1.014 ± 0.016, 0.972 ± 0.031, and 0.963 ± 0.022 for cells of category
1, 2 and 3, respectively. When all three categories are combined into a single histogram,
shown in figure 3.8, a Gaussian was fit with a center of 0.995 ± 0.012. These Gaussian
centers were used as the determined ratio of expression of FKBP-Cerulean to FRB-
Venus. For all cell types, this value is approximately 1; specifically, the ratio of
expression levels of FKBP-Cerulean to FRB-Venus in *E. coli* was determined to be 1.014
± 0.419 for cells in category 1, 0.972 ± 0.350 for cells in category 2, 0.963 ± 0.317 for
cells in category 3, and 0.995 ± 0.391 when all cells are included (errors are one standard
deviation value). Table 3.2 summarizes these concentration ratios of FKBP-Cerulean to
FRB-Venus expressed in *E. coli* determined by fluorescence microscopy.
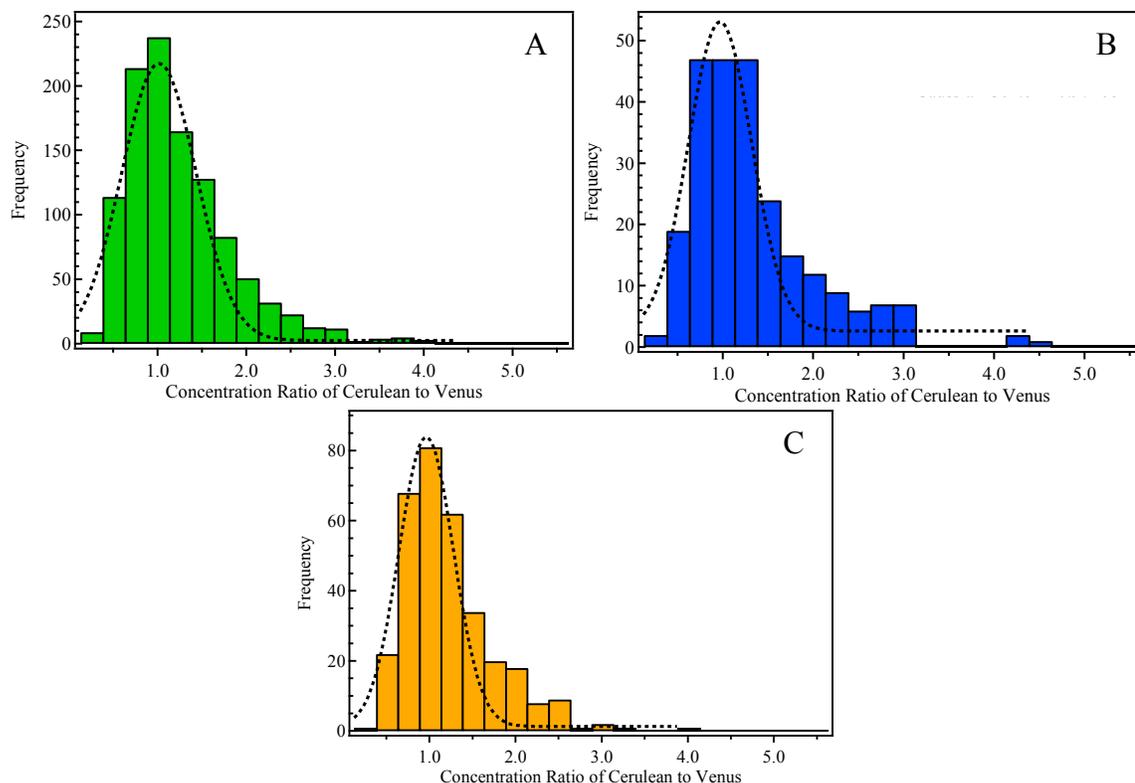
Figure 3.7 Histograms showing the distribution of normalized concentration ratios for expression of FKBP-Cerulean to FRB-Venus in *E. coli*, for A) image spots that meet the criteria for category 1; spots having well-defined edges that are oblong in shape, having the longest dimension being at least 1.5 times greater than the shortest dimension. The dotted line is a Gaussian fit to the histogram data. The Gaussian is centered 1.014 ± 0.016. B) Image spots that meet the criteria for category 2; spots having well-defined edges that do not meet the aspect ratio (≥1.5:1) criteria for category 1, and are thus more round in shape. The dashed curve is a Gaussian fit to the histogram data; the Gaussian fit is centered at 0.972 ± 0.031. C) Image spots that meet the criteria for category 3; spots without well-defined edges with no aspect ratio criteria – clumps of cells. The dashed curve is a Gaussian fit to the histogram data; the Gaussian fit is centered at 0.963 ± 0.022.
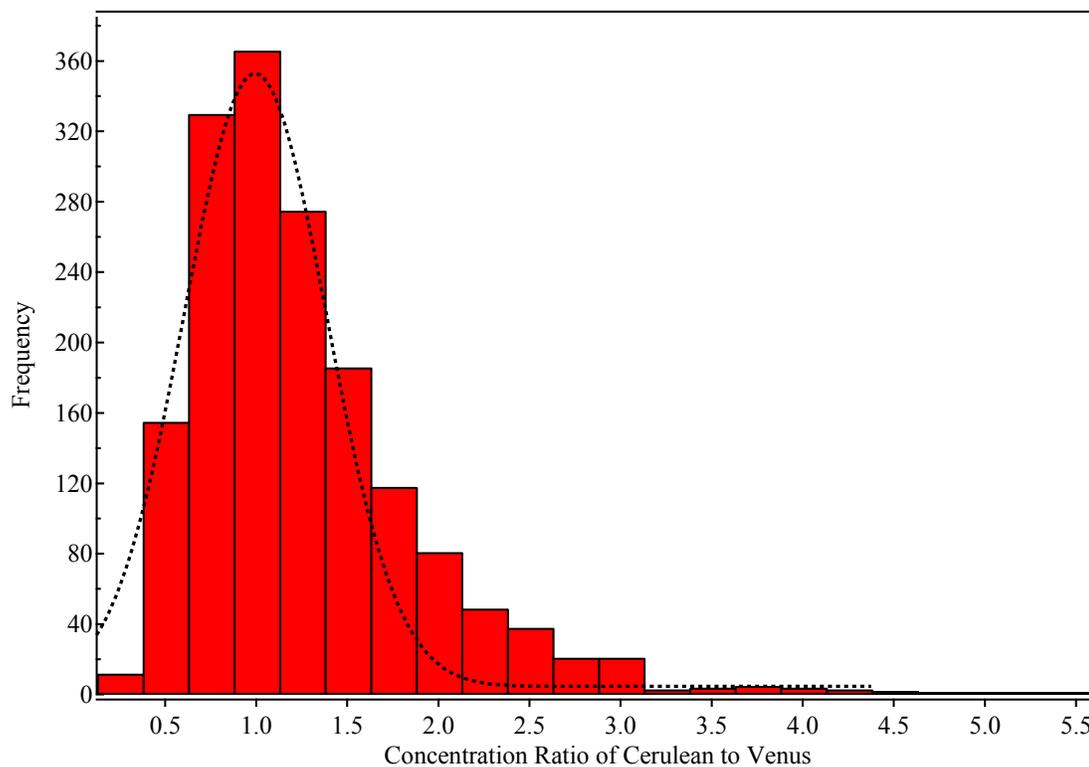
Figure 3.8 Histogram showing the distribution of normalized concentration ratios for expression of FKBP-Cerulean to FRB-Venus in *E. coli* for all cells in all images. The dashed curve is a Gaussian fit to the histogram data; the Gaussian fit is centered at 0.995.

Table 3.2. Summary of the concentration ratios of FKBP-Cerulean to FRB-Venus express in *E. coli* determined by fluorescence microscopy.

Cell Type	Concentration Ratio of Cerulean to Venus Expressed in *E. coli*
Category 1	1.014 ± 0.419
Category 2	0.972 ± 0.350
Category 3	0.963 ± 0.317
All Types	0.995 ± 0.391

3.5 Conclusion

Tables 3.1 and 3.2 are the main results. The concentration ratio of FBKP-Cerulean to FRB-Venus expressed in *E. coli* was determined by fluorescence spectroscopy and microscopy. While the ratios determined using microscopy are more accurate, since only cells expressing both proteins are considered, both methods demonstrate a nearly 1 to 1 stoichiometric concentration ratio of FKBP-Cerulean to FRB-Venus. These results indicate that the rapamycin induced dimer-forming system of FKBP-FRB with fusion to Cerulean and Venus, respectively, is a good candidate to study biological diffusion and reaction rates. Both fusion proteins were able to be purified, thus measurements could be made both *in vitro* and *in vivo*. Development of this dimer forming system for measurement in *E. coli* provides a means to test the previously developed theory on reaction rates in biological systems [7].

3.6 References

1.   Wagner, E.K., et al., *Technical glossary*, in *Basic Virology, Third Edition*. 2007, Blackwell Publishing: Malden, MA.

2.   Chen, H., H.L. Puhl, 3rd, and S.R. Ikeda, *Estimating protein-protein interaction affinity in living cells using quantitative Forster resonance energy transfer measurements*. J Biomed Opt, 2007. **12**(5): p. 054011.

3.   Braun, G. and S. Kibbee, *Laurier Research Instrumentation. User Guidelines and Standard Operating Procedure for the Cary Eclipse Fluorescence Spectrophotometer*. 2007.

4.   Hochan, L., *A FRET Srudy of PhoR-PhoB Two Component system in E. coli*, in *Physics*. 2013, Purdue University.

5.   Ferreira, T. and W. Rasband, *ImageJ User Guide*. 2012.

6.   Sternberg, S.R., *Biomedical Image Processing*. Computer, 1983. **16**(1): p. 22-34.

7.   Li, R., J.A. Fowler, and B.A. Todd, *Calculated Rates of Diffusion-Limited Reactions in a Three-Dimensional Network of Connected Compartments: Application to Porous Catalysts and Biological Systems*. Physical Review Letters, 2014. **113**.

CHAPTER 4. SUMMARY AND FUTURE WORK RELATED TO FKBP-CERULEAN
AND FRB-VENUS

An experimental system was developed to measure reaction rates inside living *E. coli* cells. One of the simplest reactions to measure inside a cell is a protein dimerization. The system developed here is a protein dimer that is induced by a small molecule; the dimer between FKBP12 and FRB forms in the presence of the small molecule rapamycin. The rate of increase in the concentration of the product is to be monitored by FRET (Förester Resonance Energy Transfer) between fluorescent proteins Cerulean and Venus fused to FKBP12 and FRB, respectively.

Using *E. coli* expression vectors that encode an N-terminal histidine tag, both fusion proteins were able to be expressed and purified for measurement in buffer; this allows comparison between *in vivo* and *in vitro* diffusion and kinetics. A key factor for the success of the FKBP-Cerulean/FRB-Venus dimerization system is coexpression of both proteins in single *E. coli* cells at a nearly 1 to 1 stoichiometry. This was achieved and verified using fluorescence spectroscopy and fluorescence microscopy. The stoichiometric ratio of the expression level of FKBP-Cerulean to FRB-Venus in *E. coli* determined for all cells was 1.721±0.125 and 0.995±0.391 by spectroscopy and microscopy, respectively.

Previously, FKBP12 and FRB have been observed to form a dimer in *E. coli* in the presence of rapamycin [1]; and the fusion protein dimer has been demonstrated to

form in HeLa cells [2]. Currently, the dimer has not been observed for this specific

system (FKBP and FRB fused to Cerulean and Venus) in *E. coli.* Possible reasons for the

inability to visualize the dimer could be: an error in protein folding of either FKBP or

FRB, a disruption of the binding pocket, or perhaps the rapamycin is not able to permeate

the *E. coli.* The latter of these possibility is unlikely however, since dimer formation has

been observed in *E. coli* previously [1]. Future work would include troubleshooting the

inability to observe dimer formation, followed by diffusion and reaction rate

measurements *in vivo* and *in vitro*.

4.1 References

1.   Davis, J.H., T.A. Baker, and R.T. Sauer, *Small-molecule control of protein degradation using split adaptors.* ACS Chem Biol, 2011. **6**(11): p. 1205-13.

2.   Chen, H., H.L. Puhl, 3rd, and S.R. Ikeda, *Estimating protein-protein interaction affinity in living cells using quantitative Forster resonance energy transfer measurements.* J Biomed Opt, 2007. **12**(5): p. 054011.

CHAPTER 5. INTRODUCTION TO RAMAN SPECTROSCOPY AND THE MULTIVARIATE HYPERSPECTRAL RAMAN INSTRUMENT

Raman scattering provides molecular, vibrational spectra that can be used to distinguish and identify a variety of chemical components of many types of systems, especially biological systems [1]. A known advantage of Raman over infrared spectroscopy is that there is little interference from water, which is important when studying biological systems since they typically contain large amounts of water. Raman is increasingly becoming recognized as a powerful tool for biological and biomedical analysis [2-8].

Recent studies of biological systems have shown that Raman spectroscopy can be used to classify and quantify intracellular chemical components [6, 9-14]. While Raman can provide a unique view of biological systems, the time required using current methods is too long to be useful for high-resolution imaging, high-throughput cell sorting/screening, and cellular dynamics. Chapters 5-8 of this thesis describe the development of a new type of Raman spectrometer which employs a multivariate, hyperspectral imaging technique, using a low-noise single channel compressive detection strategy to rapidly perform the aforementioned tasks.

5.1 Fundamentals of Raman Spectroscopy

Raman is an inelastic light scattering process in which an incident photon exchanges energy with molecular vibrations. This leads to either a decrease (Stokes Raman scattering) or increase (anti-Stokes Raman scattering) in the energy of the photon, by an amount that is exactly equal to the corresponding vibrational transition. Figure 5.1 shows the possible shifts, Stokes Raman, Anti-Stokes Raman, as well as Rayleigh scattering (when the photon is elastically scattered, without producing a molecular vibrational energy change). When Stokes scattering occurs, the molecules will absorb some of the incident photons energy causing the outgoing photon to have a lower energy. Anti-stokes scattering arises when the molecules are initially in an excited state and they lose some of their energy to the incident photon causing the emitted photon to have a higher energy than the incident photon.

Stokes scattering provides information about the vibrational normal mode frequencies of molecules in the fingerprint region from $400 - 2000$ cm^{-1}. Raman scattering can occur due to changes in polarizability of the molecules upon excitation. Therefore, structural information about the molecules of interest can be obtained by measuring shifts in vibrational frequencies.

Virtual energy states

Rayleigh Scattering E=hv$_{in}$

Stokes Scattering E=hv$_{in}$ - ΔE

Anti-Stokes Scattering E=hv$_{in}$ + ΔE

v_3

v_2

v_1

ΔE

v_0

Ground state
Vibrational energy
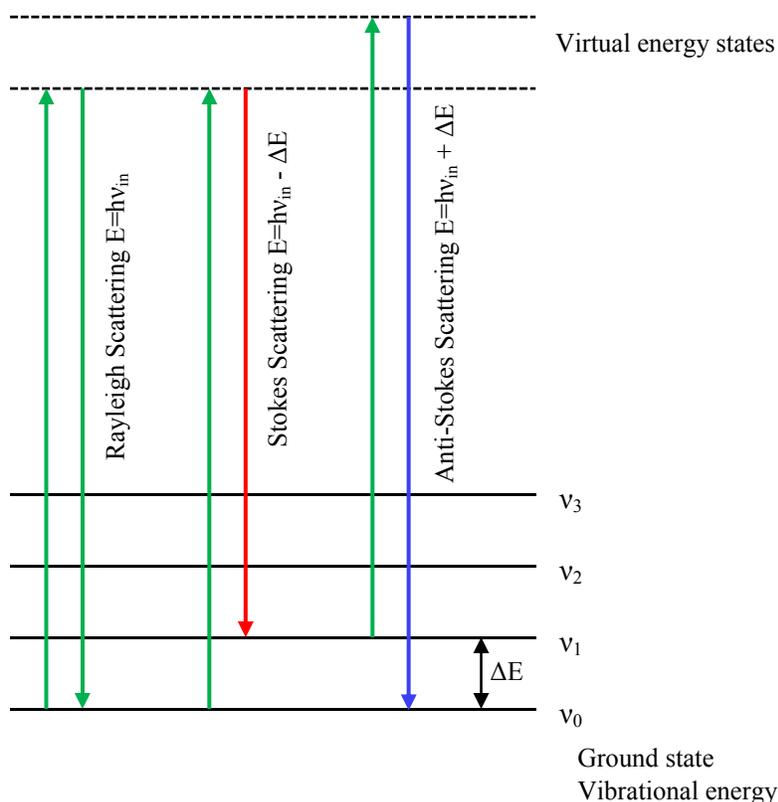
Figure 5.1. Energy diagram for Raman and Rayleigh excitation and emission. Green lines represent the incident beam of photons and, for Rayleigh scattering the emitted photons. The red line signifies the emitted photons for Stokes scattering and the blue line represents the emission for Anti-stokes scattering. The energy of the incident light is denoted by E = hv$_{in}$.

5.2 Motivation for the Multivariate Hyperspectral Imaging Instrument

Increasing Raman imaging speed requires a shift in methodology – the last major advances in this area came with the introduction of multiplexing [15, 16] and optical array detectors [17]. Despite ongoing improvements in these detectors, the time required for hyperspectral image collection is still the main obstacle for wider implementation of Raman chemical imaging. Typical Raman instruments used for imaging generally

operate as raster scanning instruments. Imaging is performed by collecting spectra at each x, y and spectral (wavenumber) dimension producing "data cubes" which are converted into an image via a univariate or multivariate analysis. For a typical point scanning, CCD-based spectrometer, at each image pixel there would be a 1 second spectra collected, making large images nearly impossible; a 1 megapixel image at this speed would take approximately 12 days. There have been a few attempts at global illumination with a tunable wavelength detection band [18-20] , and while these greatly speed up the data collection at each pixel, they do so at the expense of discarding any Raman photons outside the detection band. A superior alternative to the global illumination strategy is line scanning [17, 18], which distributes the laser power and simultaneously collects full spectra at each point along a line; however, this method suffers due to the dispersion of the full laser power across the wavelength range.

5.2.1 Single-Channel Detection

One of the advantages of the new MHI method described in the following chapters is that it makes use of a single channel detector which simultaneously detects all of the scattered Raman photons transmitted through a programmable optical filter. The single channel detector provides a fundamental signal to noise advantage over traditional CCD-based scanning Raman spectroscopy. For example, 100 Raman scattered photons detected across a CCD with many wavelength channels (>100) would barely be detectable with a signal to noise ratio of ~1. These same photons detected on a single channel detector would have a higher signal to noise ratio of ~10. Because of this intrinsic advantage of the single channel detector, the MHI would be able to function at

higher speeds requiring less accumulation time while still collecting the same quality information as a CCD-based spectrometer. However, it is important to note that the full speed advantage of the MHI detection strategy is only realized after training to determine the optimal optical filters for each application. While this is the advantage of the new detection method, the spectrometer is also able to reproduce the speed, quality and functionality of a traditional spectrometer.

5.2.2 Spectral Data Compression

The MHI detection strategy enables the single channel detector to collect spectral information by making use of hardware based spectral data compression. A few studies have illustrated the use of hardware based spectral compression that make use of a single channel detector to collect all the light transmitted through a specialized optical filter [21-23]. One of the earliest implementations of this method [21] used static optical interference filters that were customized for sets of transmission spectra. This method, however, requires manufacturing new filters for each sample set since they are chemical specific. These previous studies often make use of chemometric techniques to optimize the filters for each chemical system. Some of these systems may result in optimal filters that are quite complex, producing optical filters which are difficult to manufacture and result in a much less efficient instrument. An attractive alternative to these complex static filters would be a programmable optical filter which can be programmed with a filter of any spectral shape.

5.2.3 Spatial Light Modulators

Spatial light modulators (SLMs) provide an attractive alternative to static optical filters because they are a means of producing variable programmable filters. There are 2 main types of SLMs – liquid crystal (LC) based and digital micromirror devices (DMDs). LC based SLMs operate based on polarization of light and can use either phase or amplitude modulation to produce variable filters [24, 25]. DMDs provide binary filtering since each mirror is either in an "on" state (reflecting toward the detector) or an "off" state (reflecting away from the detector). Since this new detection method works based on variable, programmable filters, a LC-SLM is the best fit. The LC-SLM is able to provide not only "on" and "off" states, but also a gray-scale that can be used to produce any filter shape. Previously, an application of a transmissive LC-SLM suffered from low signal throughput (<20%) [22]. Another type of LC-SLM, reflectance based LC-SLMs, are reported to have a much higher signal throughput [26] and has therefore been selected for use.

5.3 The MHI

The unmatched speed of this new detection strategy is derived from the high throughput, reflectance based LC-SLM and the low noise single channel detector. The MHI can be used for either full spectral detection or high-speed compressive detection by implementing optimized filter functions. The following 3 chapters of this thesis (Chapters 6, 7, and 8) will discuss the design of the new type of Raman spectrometer as well as provide preliminary results illustrating its ability to rapidly classify samples, quantitate mixtures and produce high quality chemical images.

5.4 References

1.	Batchelder, D.N., C. Cheng, and G.D. Pitt, *Molecular Imaging by Raman Microscopy.* Advanced Materials, 1991. **3**(11): p. 566-568.

2.	Kalasinsky, K.S., et al., *Raman chemical imaging spectroscopy reagentless detection and identification of pathogens: Signature development and evaluation.* Analytical Chemistry, 2007. **79**: p. 2658-2673.

3.	Escoriza, M.F., et al., *Raman spectroscopy and chemical imaging for quantification of filtered waterborne bacteria.* Journal of Microbiological Methods, 2006. **64**: p. 63-72

4.	Maquelin, K., et al., *Rapid epidemiological analysis of Acinetobacter strains by Raman spectroscopy.* Journal of Microbiological Methods, 2006. **64**: p. 126-131.

5.	Mourant, J.R., et al., *Comparison of vibrational spectroscopy to biochemical and flow cytometry methods for analysis of the basic biochemical composition of mammalian cells.* Journal of Biomedical Optics 2006. **11**.

6.	Yu, C., et al., *Characterization of human breast epithelial cells by confocal raman microspectroscopy.* Cancer Detection and Prevention, 2006. **30**: p. 515-522.

7.	Ortiz, C., et al., *Identification of insulin variants using Raman spectroscopy.* Analytical Biochemistry, 2004. **332**: p. 5703-5709.

8.	Ortiz, C., et al., *Analysis of insulin amyloid fibrils by Raman spectroscopy.* Biophys Chem, 2007. **128**: p. 151-155.

9.	Yang, H. and J. Irudayaraj, *Rapid detection of foodborne microorganisms on food surfaces using Fourier transform Raman spectroscopy.* Journal of Molecular Structure, 2003. **646**: p. 35-43.

10.	Swain, R.J. and M.M. Stevens, *Raman microspectroscopy for non-invasive biochemical analysis of single cells.* Biochemical Society Transactions, 2007. **35**(3): p. 544-579.

11.	Gierlinger, N. and M. Schwanninger, *The potential of Raman micorscopy and Raman imaging in plant research.* Spectroscopy: an International Journal, 2007. **21**: p. 69-89

12.   Notingher, I., et al., *Multivariate analysis of Raman spectra for in vitro non-invasive studies of living cells.* Journal of Molecular Structure, 2005. **744-747**: p. 179-185.

13.   Notingher, I., et al., *In situ characterisation of living cells by Raman spectroscopy.* Spectroscopy: an International Journal, 2002. **16**: p. 43-51.

14.   Notingher, L. and L.L. Hench, *Raman microspectroscopy: a noninvasive tool for studies of individual living cells.* Expert Review of Medical Devices, 2006. **3**: p. 215-234.

15.   Decker, J.A. and M.O. Harwitt, *Sequential Encoding with Multislit spectrometers.* Applied Optics, 1968. **7**(11): p. 2205-2209.

16.   Ibbett, R.N., D. Aspinall, and J.F. Grainger, *Real-time multiplexing of dispersed spectra in any wavelength region.* Applied Optics, 1968. **7**(6): p. 1089-1093.

17.   Golcuk, K., et al., *Is photobleaching necessary for Raman imaging of bone tissue using a green laser?* Biochimica et Biophysica Acta - Biomembranes, 2006. **1758**(7): p. 868-873.

18.   Delhaye, M. and P. Dhamelincourt, *Raman microprobe and microscope with laser excitation.* Journal of Raman Spectroscopy, 1975. **3**(1): p. 33-43.

19.   Morris, H.R., et al., *Liquid crystal tunable filter Raman chemical imaging* Applied Spectroscopy, 1996. **50**(6): p. 805-811.

20.   Morris, H.R., C.C. Hoyt, and P.J. Treado, *Imaging spectrometers for fluorescence and Raman Microscopy: Acousto-optic and liquid-crystal tunable filters.* Applied Spectroscopy, 1994. **48**(7): p. 857-866.

21.   Nelson, M.P., et al., *Multivariate optical computation for predictive spectroscopy.* Analytical Chemistry, 1998. **70**: p. 73-82.

22.   Uzunbajakava, N., et al., *Low-cost spectroscopy with a variable multivariate optical element.* Analytical Chemistry, 2006. **78**: p. 7302-7308.

23.   Quyen, N.T., et al., *New Raman spectrometer using a digital micromirror device and a photomultiplier tube detector for rapid on-line industrial analysis. Part I: Description of the prototype and preliminary results.* . Applied Spectroscopy, 2008. **62**(3): p. 273-278.

24.   Armitage, D., et al., *Ferroelectric liquid crystal spatial light modulator.* Molecular Crystals and Liquid Crystals, 1987. **144**: p. 309-316.

25.   *Spatial Light Modulators*, in *Boulder Nonlinear Systems* 2006.

26.   Harriman, J., S. Serati, and J. Stockley, *Comparison of transmissive and reflective spatial light modulators for optical manipulation applications.* Proceedings of SPIE, 2005. **5930**.

CHAPTER 6. MHI INSTRUMENTATION AND DETECTION

6.1 MHI Instrument Design

The MHI is built around a microscope with backscattering geometry. A 785nm

single mode laser diode (Innovatie Photonic Solutions Inc.) is focused on to the sample

through a 20X (NA 0.4) NIR objective (Olympus). Light scattered from the sample is

collected back through the objective and Raman (Stokes-shifted) photons are collected

through a 45° dichroic long pass filter (Semrock Inc.). Additionally, Rayleigh photons

are rejected by a 785nm notch filter (Semrock Inc.) placed outside of the microscope.

After the microscope, the Raman photons pass into the path of the MHI detection optics.

The optical path of the MHI is illustrated in the schematic shown in Figure 6.1

and a detailed parts list is given in Table 6.1. The MHI instrument operates by

manipulating the polarization of light. First, the light passes through a half-wave plate

(Edmund Optics), which rotates the light and determines which polarization is detected.

The second component is a Glan-Laser polarizing beam splitting cube (ThorLabs), which

only passes p-polarized light toward the SLM. The p-polarized light then passes through

the volume phase holographic (VPH) grating (Edmund Optics) where it is dispersed by

wavelength and focused onto individual pixels of the SLM by an achromatic lens

(Edmund Optics). A second half-wave plate, between the focusing lens and the SLM, is

used to align the polarization of the incoming light to the polarization of the liquid

crystals of the SLM to allow for maximum reflection and rejection. A key factor for the success of the SLM-based detection is that the distances between the VPH grating and the achromatic lens and the SLM are matched. This ensures that the light reflected off the SLM will return on the same optical path to the VPH grating. Depending on the filter function loaded into the SLM, some of the light focused onto the SLM will have its polarization flipped. Any light whose polarization has been rotated away from p-polarization will be reflected by the Glan-Laser beam splitting cube and directed to the detector.

The key feature of the MHI is the LC-SLM. It is a programmable filter which allows for high speed imaging of many systems. The SLM selected for use in the MHI is reflectance, rather than transmission based, for its higher throughput (>80%). It is an array of 12,288 individually addressable linear pixels. Figure 6.2 is a diagram of the SLM used in the MHI. Each pixel is an individual optical phase modulator that can rotate the light from p to s polarization. Only light that is rotated away from p-polarization will be reflected to the detector. Neighboring pixels have a pixel pitch, center to center distance, of 1.6 μm with a pixel size of 1.0 μm and gap of 0.6 μm. The digital filters programmed into the SLM determine the degree of rotation of the polarization of light hitting each pixel thus determining how much of the incoming light will be reflected back through the Glan-Laser cube and detected.

This reflected light is then collected with an aspheric lens (Edmund Optics) and focused into a 200μm core fiber optic cable (ThorLabs) coupled to an avalanche photo-diode (APD) single channel detector (Hamamatsu). The APD detector was selected for its high photosensitivity and low noise characteristics in the wavelength range of interest,

800-950 nm. The APD outputs a voltage to a USB-data acquisition device (National

Instruments) which digitizes the information for computer readout.
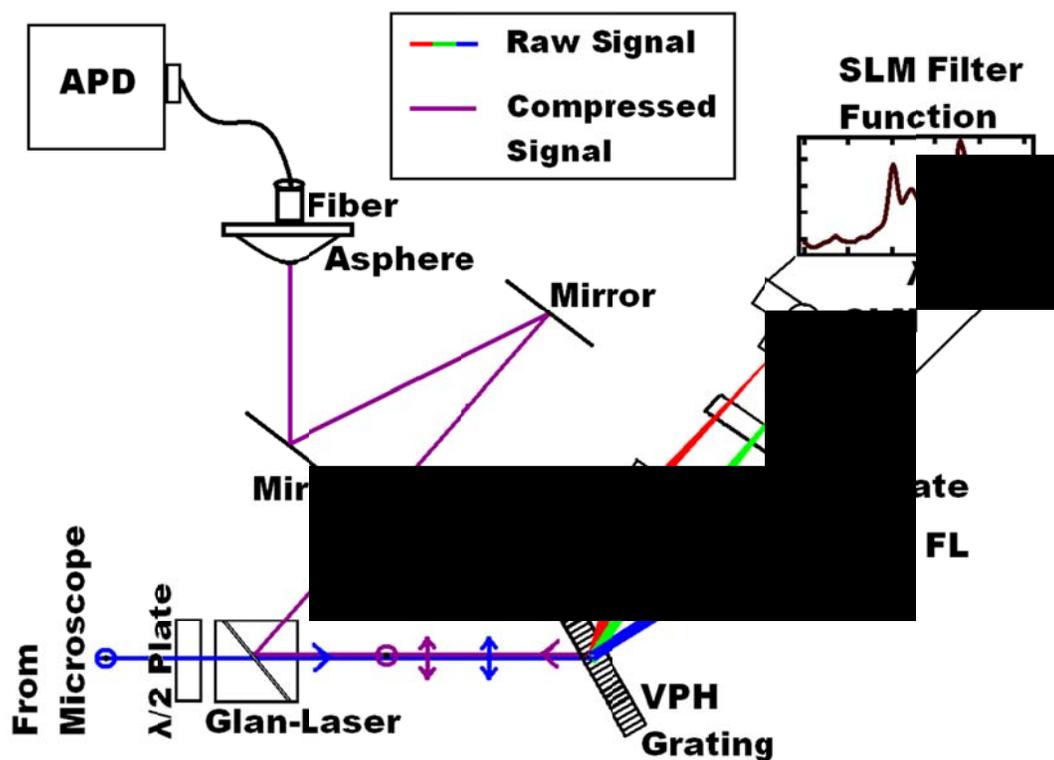
Figure 6.1. Schematic of the MHI detection optics [1]. Raman scattered light comes from the microscope and enters into the MHI detection optics. Major parts to the new type of detection include the Glan-Laser polarization beam splitting cube, the high-throughput volume holographic grating (VPH), the liquid-crystal SLM, and the single-channel, Avalanche Photodiode detector.
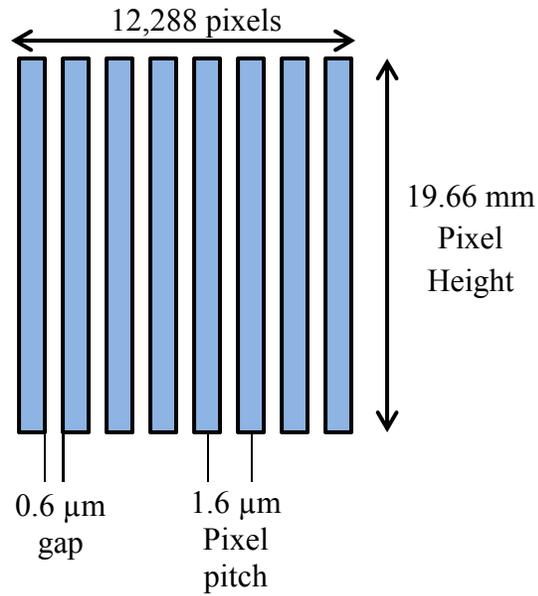
Figure 6.2. A schematic representation of the LC-SLM used in the MHI [2]. The SLM is made of 12,288 individually addressable linear pixels (represented by the blue rectangles). Each pixel has a width of 1.6 μm and a gap between them of 0.6μm, resulting in a center-to-center distance, or pixel pitch, of 1.6 μm.

Table 6.1. MHI Parts List

PART DESCRIPTION	MANUFACTURER	PART NUMBER
785 nm Diode Laser Module	Innovative Photonic Solutions	IO785SU0100B
12,288 Linear Pixel LC-SLM	Boulder Nonlinear Systems	P12,288-0785
Avalance Photodiode Module	Hamamatsu	C4777-01
Triple Output Power Supply, +5V, ± 20V	Agilent Technologies	E3630A
Upright Research Microscope	Olympus	BX-51
NI USB-6211 Bus Powered M Series Multifunction DAQ	National Instruments	779676-01
Motorized Microscope Stage	Prior Scientific	ProScan
Volume Phase Holographic Grating, 1200 l/mm, 830 nm	Edmund Optics	48-590
λ/2 achromatic waveplates, 30 mmØ	Edmund Optics	48-500
Achromatic Lens f=100mm, NIR coated	Edmund Optics	47-317
Aspheric Lens 25 mmØ, 0.66 N.A., NIR Coated	Edmund Optics	49-113
Glan Laser Calcite Polarizers, 600-1050 nm Coating, 1"	ThorLabs	GL15-B
Broadband Dielectric Mirrors, 700-1100nm, 1"Ø	ThorLabs	BB1-E03-10
0.22 NA, 200 μm core SMA, Multimode Patch Cable, 2 meter	ThorLabs	M25L02
785 nm MaxLine Bandpass Filter, 1" dia.	Semrock	LL01-785-25
785 nm RazorEdge Dichroic 45° Beamsplitter	Semrock	LPD01-785RU-25X36X2.0
785 nm StopLine Notch Filter	Semrock	NF01-785U-25
20X NIR objective, 9 mm FL, 8.1 mm WD, NA 0.40	Olympus	LMPL20XIR

<u>6.2 MHI Detection Strategies</u>

6.2.1 Conventional Raman Spectroscopy

While the full potential of the MHI instrument is only realized after pre-training based on system components, it is also fully functional as a conventional Raman spectrometer. The MHI spectrometer can collect full Raman spectra in one of two main ways. Both involve programming a series of filters into the SLM. First, the MHI can imitate functionality of a simple scanning CCD-based Raman spectrometer by using a series of programmable notch, or band-pass, filters for the SLM. For this method, only a short section of pixels are turned "on", or set to a full polarization rotation from p to s, for each filter. Each filter will then have a different short section of pixels active, and this section will scan sequentially across the SLM. Figure 6.3 is a diagram of this filter method for a 28 pixel system. This method is, however, much less efficient than the traditional scanning CCD-based spectrometer since most of the Raman photons are reflected away from the detector.
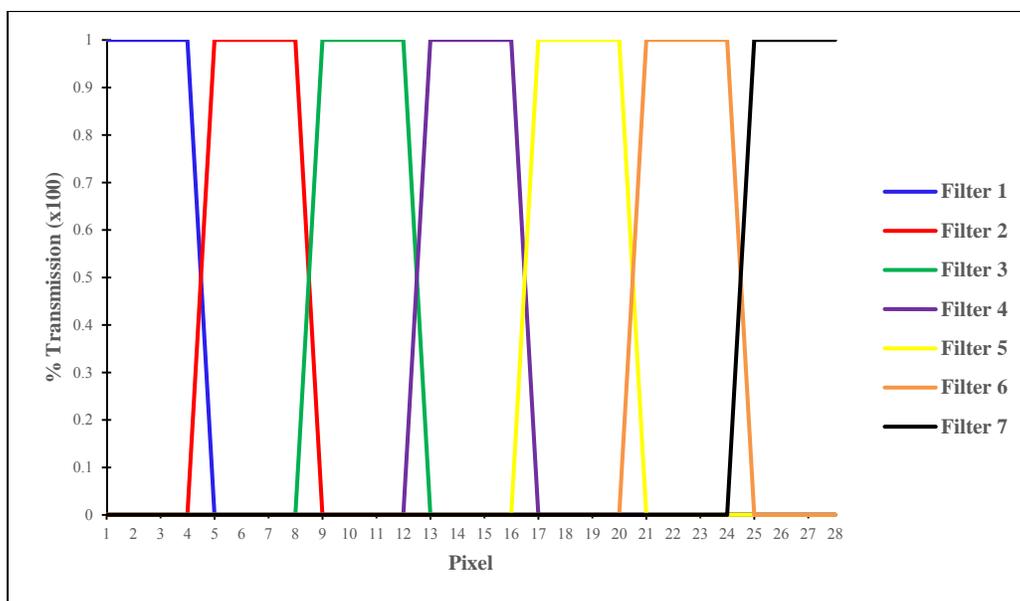
Figure 6.3. Example of a set of sequential scanning notch-filters for a 28-pixel spectra.

Another more efficient method for collecting full Raman spectra using the MHI is to employ Hadamard transform. Hadamard transform is analogous to Fourier transform; however, instead of sines and cosines, Hadamard uses square waves. Few Hadamard based systems have been reported in the literature, but these are not commonly available due to the high cost of limited mechanical devices to produce these filters. These systems require moving mechanical parts to sequence through the necessary Hadamard filters for data collection. By programming these filters into the SLM, the MHI can overcome this issue.

Both band-pass scanning and Hadamard require the same amount of time; for both methods, N resolution elements (or N wavelength points) require N consecutive measurements. However, Hadamard measurements are advantageous due to a higher signal to noise ratio giving more accurate spectra. While during a band-pass scan, only a

small section of the SLM is rotating the polarization at time, during Hadamard collection, exactly half the SLM elements are "on" at a time. This allows for the collection of many more photons at a time, while the noise remains approximately the same.

      The Hadamard matrix is translated to a series of filters for the SLM using 1s and 0s. For an individual filter, a pixel encoded with a 1 is set to 100% reflectance, or the polarization is fully rotated. Conversely, a pixel coded with a 0 is set to 0% reflectance and the polarization remains the same. A simple 3 point Hadamard matrix application is shown in Figure 6.4. For a Hadamard measurement, S represents the set of Hadamard filters programmed into the SLM and Y is the measured voltage from the APD. The spectrum is obtained by multiplying the APD voltage response by the inverse of the Hadamard matrix. A typical Hadamard spectrum taken using the MHI will have 128 or 256 spectral points, requiring 128 or 256 Hadamard filters.

Y: Filter Response
S: Hadamard Matrix
X: Spectrum

$$Y = S\ X$$

$$
\begin{matrix}
\mathbf{Y} & & \mathbf{S} & & \mathbf{X}
\end{matrix}
$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} =
\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
$$

$$X = S^{-1}\ Y$$

Figure 6.4. Example Hadamard transformation for a 3-element system [3]. Hadamard transform is typically done with 128 elements when using the MHI for spectroscopy.

Raw data is recorded from the APD detector as a voltage versus pixel; each voltage measurement is the MHI response to a single Hadamard filter programmed into the SLM. Figure 6.5(a) shows the plot of the APD voltage responses to the Hadamard filters for the organic compound *n*-hexane. Using the Hadamard transform algorithm, which multiplies this voltage response by the inverse Hadamard matrix, yields the spectrum of *n*-hexane (Figure 6.5(b)). Also, to illustrate the improvement in signal to noise from the band-pass scanning SLM filters to the Hadamard method, Figure 6.5(c) shows the notch scan spectrum of *n*-hexane.
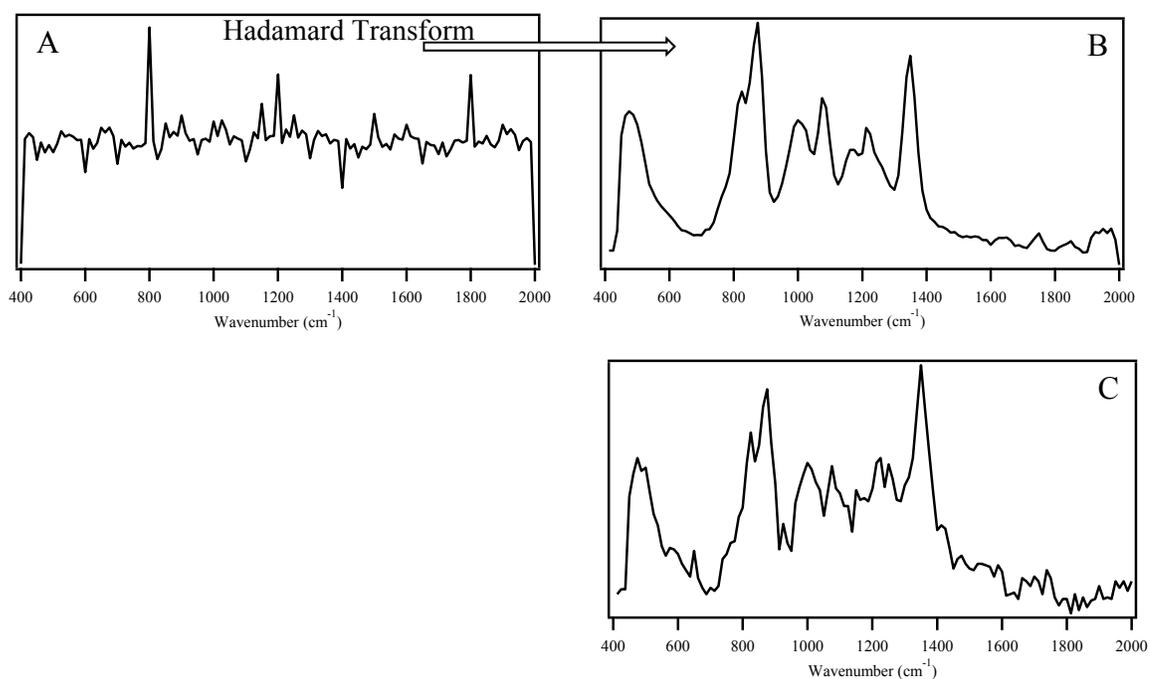
Figure 6.5. Hadamard transform spectra versus notch-scan spectra of *n*-hexane with 128 resolution elements. A) Response of the MHI to the set of 128 Hadamard filters. B) Hadamard transform of A. C) Notch-scan spectra of *n*-hexane with 128 notch filters.

6.2.2 Filter Methods for High-Speed Detection

The true efficiency of the MHI instrument is realized using SLM filter functions that contain information about the compound of interest. These filters can be functions whose shapes match that of the components or whose shapes match that of eigenvectors optimized for each component. In other words, the high speed advantage of the MHI instrument requires pre-training using filters tuned to the components of interest. One method is to use spectral angle mapping [4], or using the component spectra themselves as filters, and another is to use a partial least squares (PLS) regression [5] to obtain loading vectors which can be translated into SLM filters.

Of these two methods, spectral angle mapping is the simplest to implement, as it only requires the analyte spectra to be converted into an SLM filter function [4]. The spectra are scaled to an intensity range of 0 to 1 which will convert to 0 to 100% reflectance on the SLM. Both the SLM filter function and the Raman scattered light at different wavelengths can be represented as n-dimensional vectors. By measuring the amount of light directed toward the detector from the SLM, this method is effectively measuring the dot-product, or correlation coefficient, of those vectors. The spectral information obtained from the sample can then be classified according to the correlation coefficient of the spectral vector, or the scattered light from the analyte, with the SLM filter function [1].

 Although spectral angle mapping can be quite useful for classifying simple systems, it is not always sufficient for more complex systems with more complex spectra, such as biological samples. For these more complex samples, SLM filter functions are made using the PLS regression technique [5]. Although it is more computationally

intensive to obtain the filters, it is advantageous since PLS maximizes the variance

between each component of the system allowing for increased selectivity [1].

6.2.2.1 Principles of the PLS Regression Method

Partial least squares is a multiple linear regression(MLR) model [5]. Multiple

linear regression can be used with many factors, however, if the number of factors

becomes too large, then the data become over-fitted. This means that the model will fit

the sampled data perfectly, but will fail to predict accurately. There are actually only a

small number of underlying factors that account for most of the variance in the data. The

general goal of PLS is to extract these "latent" factors, while accounting for as much of

the variation as possible [6]. In other words, PLS finds the regression vector that defines

the components in a sample by relating a matrix of sample spectra (X) to a concentration

vector (Y). In a very simplified sense, the use of the PLS regression vector is similar to a

simple chemical concentration calibration curve. The regression vector accounts for

variance in the sample spectra while relating it to the concentration matrix. The resulting

PLS model can then predict the concentration (Y) of an unknown sample by applying the

regression coefficient [7].

When using PLS to produce SLM filter functions, training spectra are recorded

using the MHI in order to produce the sample spectra matrix (X). These spectra should

include samples of known components of known concentration. The concentrations

should span the range of possible concentrations of the components in the unknown

samples. The PLS algorithm will then compute a set of regression coefficients (or

loadings) that will account for the variance among the spectra, given the information about the concentration matrix. Each regression vector will account for a certain degree of the variance among the spectra. The one which captures the largest amount of variance is chosen for use as an SLM filter. For a 2 component system on the MHI, this is typically the second eigenvector.

The PLS method can also be used for classification problems. In this case, instead of having a set of training spectra of varying concentrations, the spectra matrix will contain only spectra of the pure components. For the concentration matrix, one component should be chosen to be encoded as 100% and the other as 0%. For example, for a simple classification of *n*-hexane and *n*-hexanol, the spectra matrix would contain perhaps 5 spectra of pure *n*-hexane and 5 spectra of pure *n*-hexanol. Then the corresponding concentration matrix would encode the *n*-hexane spectra as 0 and the *n*-hexanol spectra as 100.

Regardless of the application, quantitation or classification, the second PLS regression coefficient, chosen for use as an SLM filter function, has both positive and negative portions. In order to use this as a filter, the vector needs to be split into 2 parts, a positive part and a negative one. Both the positive and the negative portions of the filter are then scaled to an intensity range from 0 to +1. Figure 6.6 shows a PLS eigenvector from an *n*-hexane/*n*-hexanol quantitation example being divided into its positive and negative portions.
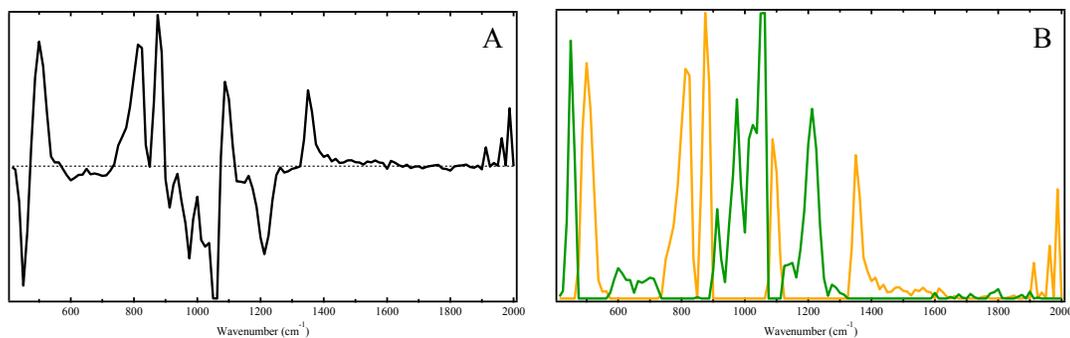
Figure 6.6. PLS-derived *n*-hexane/*n*-hexanol filter process.
A) Second PLS regression vector for quantitation of *n*-hexane and *n*-hexanol. B)
Regression vector from B deconstructed into its positive (yellow) and negative (green)
portions and scaled from 0 to +1.

6.2.3 Post-Acquisition Analysis

Once the SLM filters are used to collect information about a sample, some post-acquisition analysis is necessary to reconstruct the response of the sample to the filter. This analysis is quite minor for the spectral angle mapping method, as this only requires taking into account the scaling factor that was necessary in order to scale the spectra to an intensity range from 0 to 1.

In order to regenerate the response of the samples to the PLS filters shown in Figure 6.6B, there is some post-acquisition analysis required. Let R, P and N be the MHI responses to the SLM filters r_i (Figure 6.6A), the positive portion p_i (Figure 6.6B – yellow) and the negative portion n_i (Figure 6.6B – green), respectively.

The normalized signal arising from each wavelength pixel (i) is

$$I_i = \frac{S_i}{\Sigma_i S_i};$$

(6.1)

where S_i is the raw APD output voltage and $\Sigma_i S_i$ is the APD output voltage when the SLM is set to "all on" or 100% reflectance. The APD signal when both the positive and negative filters are applied to the SLM can be expressed as follows:

$$P = c^+ \Sigma_i p_i I_i$$

(6.2)

$$N = c^- \Sigma_i n_i I_i$$

(6.3)

where c^+ and c^- are the scaling factors necessary to translate each filter into a full intensity range from 0 to 1. The response to R can then be reconstructed simply by summing P and N, $R = P + N$. This final output, R, relates to the classifier, 0 or 1, or concentration scale, percent composition from 0 – 100%, used in the PLS algorithm.

6.3 References

1.   Davis, B.M., et al., *Multivariate hyperspectral Raman imaging using compressive detection.* Analytical Chemistry, 2011. **83**(13): p. 5086-92.

2.   *Spatial Light Modulators*, in *Boulder Nonlinear Systems* 2006.

3.   Treado, P.J. and M.D. Morris, *A Thousand Points of Light: The Hadamard Transform in Chemical Analysis and Instrumentation.* Analytical Chemistry, 1989. **61**(11): p. 723A-734A.

4.   Yuhas, R.H., A.F.H. Goetz, and J.W. Boardman, *Discrimination Among Semi-arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm*, in *Third Annual JPL Airborne Geoscience Workshop*, R.O. Green, Editor. 1992: Pasadena, CA. p. 147-149.

5.   Geladi, P. and B.R. Kowalski, *Partial Least-Squares Regression: A Tutorial.* Analytica Chimica Acta, 1986. **185**: p. 1-17.

6.   Tobias, R.D., *An Introduction to Partial Least Squares Regression.*

7.   Cebeci, D., *Multivariate Hyperspectral Raman and CCD-based Raman spectroscopy applications for pharmaceutical analysis*, in *Chemistry*. 2011, Purdue University.

CHAPTER 7. CLASSIFICATION, QUANTITATION AND IMAGING USING THE
MHI INSTRUMENT

The MHI instrument was used to perform "proof-of-concept" measurements to

classify and quantitate liquid mixtures, as well as to rapidly image a pharmaceutical

sample. Liquids whose Raman spectra are quite different in the fingerprint region (400 to

2000 cm^{-1}), cyclohexane and n-hexane, are used to demonstrate the classification ability

of the spectral angle mapping filter method. Conversely, liquids whose Raman spectra

are similar in this region, n-hexane and n-hexanol, are used to demonstrate the ability of

the PLS regression filter method to correctly identify the concentration of each

component in a mixture. Finally, the PLS regression filter method is used to rapidly

produce an image of a composite aspirin tablet.

7.1 Classification of Hexane and Cyclohexane

Spectral angle mapping, which uses the component spectra themselves as filter

functions for the SLM [1], was used to classify liquid samples of cyclohexane and n-

hexane.   Figure 7.1 shows the spectra of liquid cyclohexane and n-hexane obtained

using the Hadamard detection method for collecting full spectra using MHI (see section

6.2.1). These spectra were then converted to SLM filters by scaling their intensities from

0 to 1.
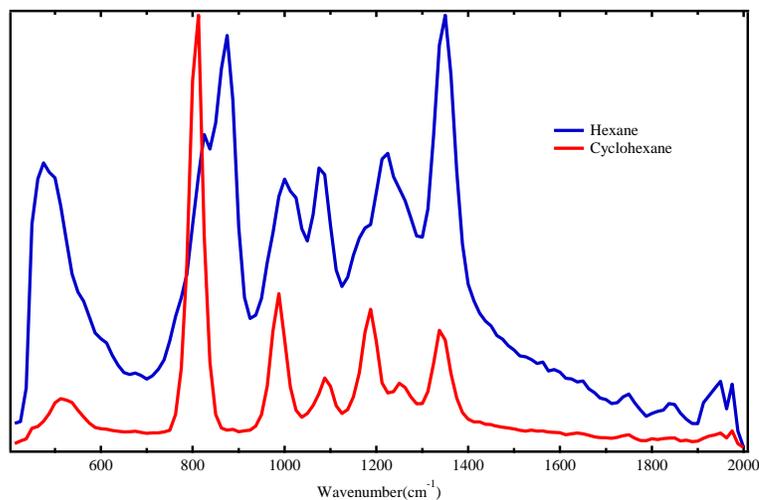
Figure 7.1. MHI (Hadamard transform) spectra of cyclohexane (red) and *n*-hexane (blue). These spectra were used as SLM filters for spectral angle mapping classification application of the MHI.

Figure 7.2 shows the MHI classification results using the spectral angle mapping method on cyclohexane and *n*-hexane. These results were obtained by applying the spectral filters, along with an "all on" filter for normalization, to the SLM for various integration times. Results are plotted as the response to the cyclohexane filter versus the response to the *n*-hexane filter. Samples that were cyclohexane were expected to be in the upper left corner at (0,1) of the plot and samples that were *n*-hexane were expected to be in the lower right corner at (1,0). Using these results, it is evident that accurate classification of these 2 representative liquids was achieved at as low as 100µs.
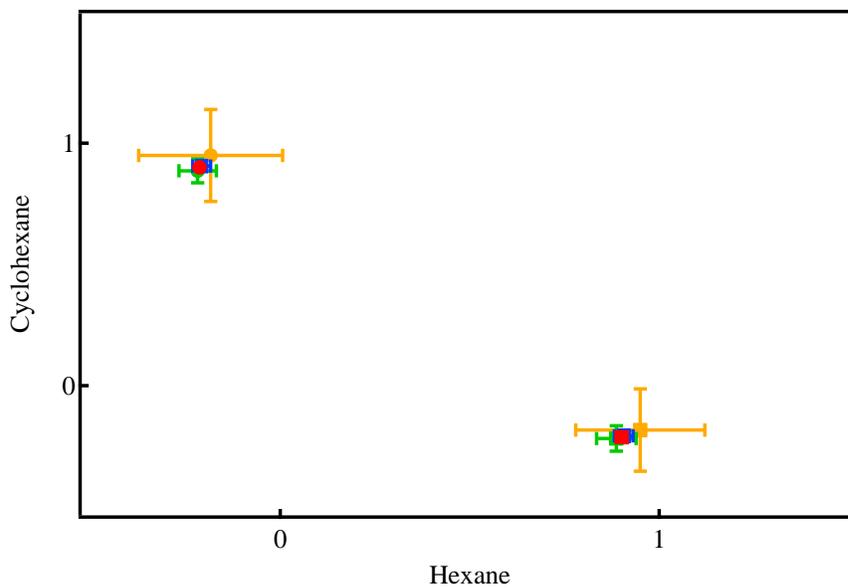
Figure 7.2. Classification results for cyclohexane *n*-hexane for spectral equivalent filters at integration times of 100ms (red points), 10ms (blue points), 1ms (green points) and 100μs (yellow points) [2]. The results are plotted as the response to the cyclohexane filter (left axis) versus the response to the *n*-hexane filter (bottom axis). Samples classified as cyclohexane are represented as circles and samples classified as *n*-hexane are represented as squares. Error bars are based on the standard deviation of 15 separate measurements at each integration time. (Image adapted from reference 2).

7.2 Quantitation of Hexane and Hexanol Mixtures

Like spectral angle mapping, PLS regression can also easily classify two components. PLS also has the added benefit of being able to determine the concentration of a known component of a mixture. Liquid samples of *n*-hexane and *n*-hexanol, whose Raman spectra are similar in the fingerprint region from 400-2000 cm^{-1}, were used to demonstrate the capability of the MHI to accurately quantify the concentration of their mixtures.

Figure 7.3 shows the Raman spectra and SLM filter functions for *n*-hexane and *n*-hexanol, whose Raman spectra look similar in the fingerprint region of 400-2000 cm^{-1}.

The spectra of *n*-hexane and *n*-hexanol in figure 7.3a were obtained using the Hadamard transform collection technique for the MHI (see section 6.2.1). These spectra, along with spectra of a range of concentrations (0-100% *n*-hexane), were used as input into the SIMPLS [3] algorithm, which performs PLS (using MATLAB along with the PLS toolbox, Eigenvector research inc.). The output of the PLS algorithm is a set of regression vectors; since there are two components in this system the one that captures the largest amount of variance to be used as an SLM filter is the second vector, shown in figure 7.3b. In order for this regression vector to be used as an SLM filter function, it must be split into its positive and negative parts and scaled to an intensity range of 0 to 1. Figure 7.3c illustrates the splitting of the regression vector from figure 7.3b.

The PLS-derived SLM filters were then used to measure test samples to determine the capability of the PLS method for determining the concentration of the components in mixtures. Like for the spectral angle mapping classification method, the PLS method for quantitation was tested at multiple APD integration times. Figure 7.4 shows the concentration quantification results for *n*-hexane/*n*-hexanol mixtures using the PLS-derived filter functions described above with APD integration times ranging from 100ms to 1ms per filter. These results are quite good, considering the similarity of their spectra in the fingerprint region. The correlation coefficient of a linear fit to the data points in figure 7.3 increases from 0.921 to 0.995 as the integration time increases from 1ms to 100 ms respectively (the linear fit of the 100ms data is the line shown in figure 7.4). As expected, the longer integration time of 100ms provides the most accurate concentration measurements.

Figure 7.3. PLS-derived SLM component filter algorithm. (A) Pure component spectra of *n*-hexane (red curve) and *n*-hexanol (blue curve) taken using the MHI [2]. These as well as spectra of mixtures were used as input into the PLS algorithm. (B) PLS regression output eigenvector used to for the SLM filter functions. (C) The PLS eigenvector from B is split into a positive portion (yellow curve) and negative portion (green curve) and scaled to an intensity range of 0 to 1 for use as SLM filter functions. (Image adapted from reference 2)
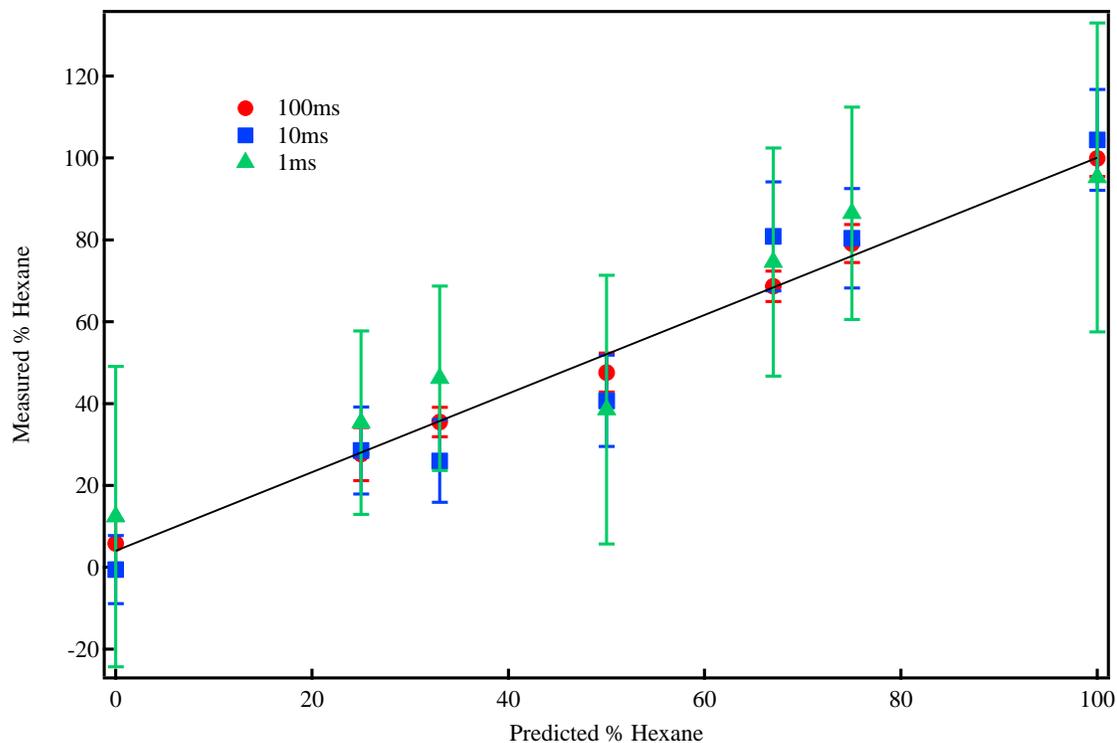
Figure 7.4. MHI concentration measurements using the PLS-derived SLM filter functions as compared with the known mixture concentrations of *n*-hexane/*n*-hexanol mixtures for 3 different APD integration times per filter [2]: 100 ms (red circles), 10ms (blue squares), and 1ms (green triangles). The black curve is the linear fit for the 100ms data points, which has a correlation coefficient of 0.995. Error bars represent the standard deviation of 10 individual measurements at each integration time. (Image adapted from reference 2)

7.3 Pharmaceutical Imaging application – Composite Aspirin Tablet

The high-speed imaging capability of the MHI was tested using a composite aspirin tablet; a store bought aspirin tablet was drilled with 3 holes that were then filled with theophylline (a common pharmaceutical excipient). This application was tested to show that the MHI could potentially be used to rapidly determine the distribution of active ingredients and excipients in a pharmaceutical sample. The PLS regression method was used to produce SLM filter functions as described for the hexane/hexanol quantitation. Figure 7.5 shows the PLS-derived SLM filter algorithm for a sample containing aspirin and theophylline. Spectra of pure aspirin and theophylline, figure 7.5a, were taken using the Hadamard transform method (see section 6.2.1). These pure spectra were used as input into the PLS algorithm with aspirin coded as 1 and theophylline coded as 0. Since this is a two-component system, the second PLS regression vector is used for the SLM filter, figure 7.5b. Figure 7.5c shows the second regression vector split into its positive and negative portions and scaled from 0 to 1 intensity range.

Figure 7.6 shows the chemical image created using the PLS-derived SLM filters for aspirin and theophylline as described above. The total APD integration time was 1 ms per filter, resulting in a 30 second total signal collection. It is evident from the image that the MHI can create a chemical map of the aspirin tablet with theophylline packed craters. If the integration time is increased, the accuracy of the map is expected to increase and contain less uncertainty.
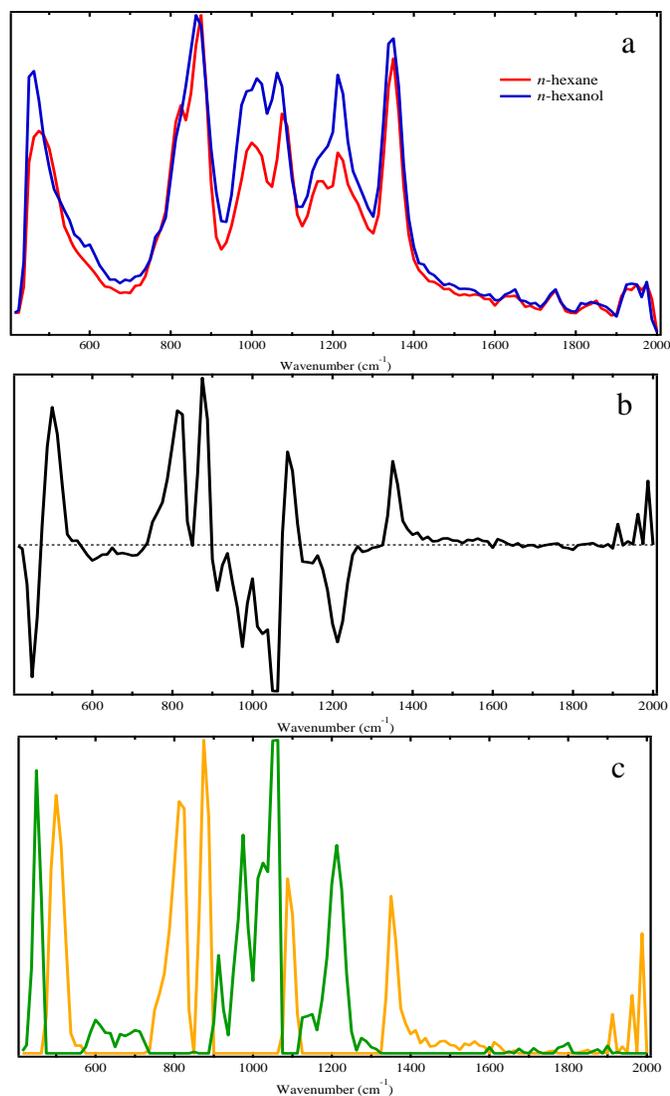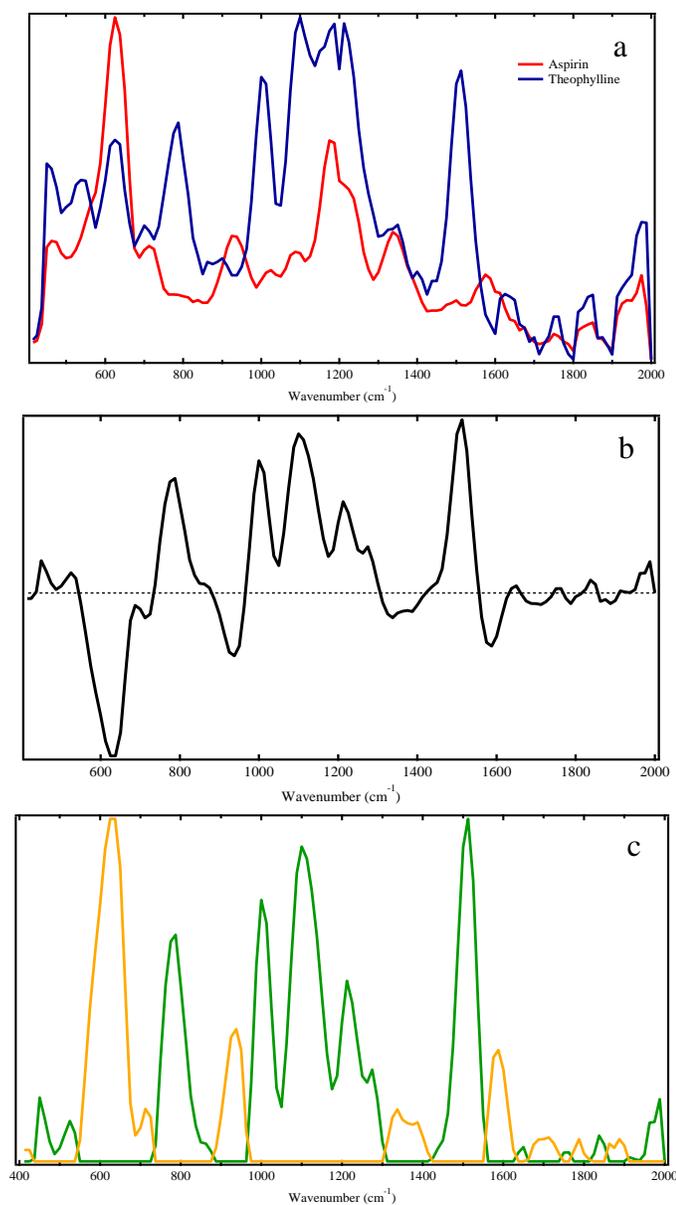
Figure 7.5. PLS-derived SLM component filter algorithm [2]. (A) Pure component spectra of aspirin (red curve) and theophylline (blue curve) taken using the MHI. (B) PLS regression output eigenvector used to for the SLM filter functions. (C) The PLS eigenvector from B is split into a positive portion (green curve) and negative portion (yellow curve) and scaled to an intensity range of 0 to 1 for use as SLM filter functions. (Image adapted from reference 2)
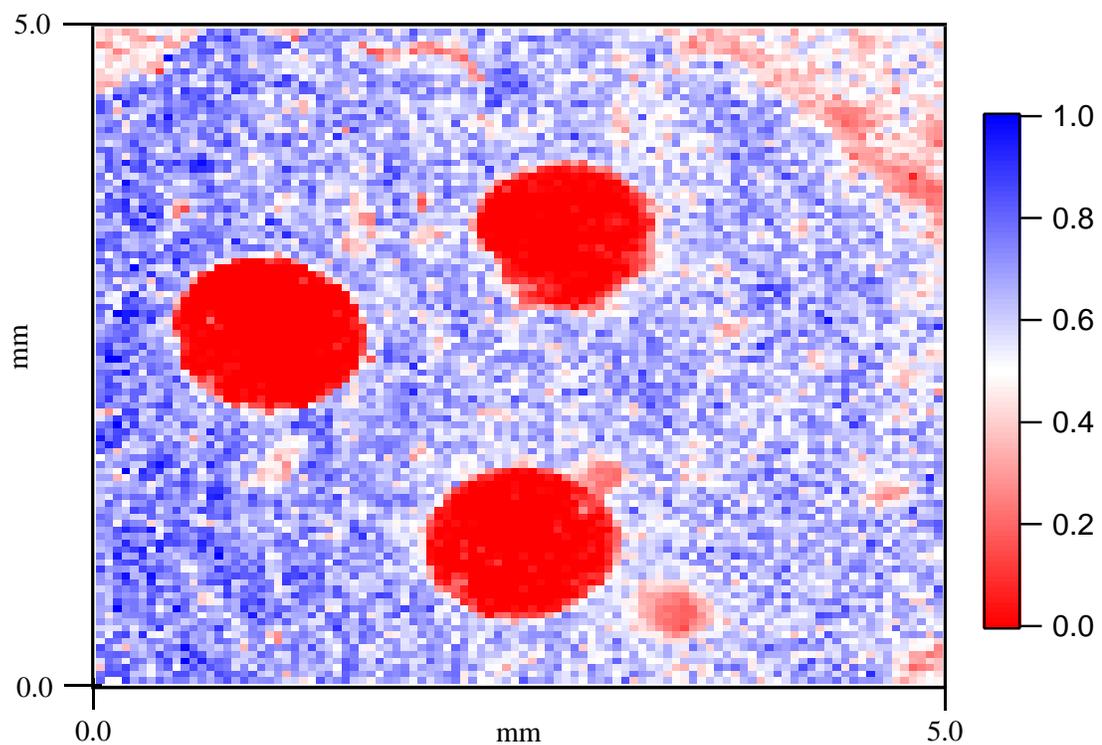
Figure 7.6. MHI chemical image (100x100 pixels) of an aspirin tablet (blue) with theophylline (red) packed holes at an APD integration time of 1ms per filter [2]. Total signal collection time was 30s. The image color map is continuous from red (0) to white (0.5) to blue (1). (Image adapted from reference 2)

7.4 References

1.   Yuhas, R.H., A.F.H. Goetz, and J.W. Boardman, *Discrimination Among Semi-arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm*, in *Third Annual JPL Airborne Geoscience Workshop*, R.O. Green, Editor. 1992: Pasadena, CA. p. 147-149.

2.   Davis, B.M., et al., *Multivariate hyperspectral Raman imaging using compressive detection.* Analytical Chemistry, 2011. **83**(13): p. 5086-92.

3.   De Jong, S., *SIMPLS: an alternative approach to partial least squares regression.* Chemometrics and Intelligent Laboratory Systems, 1993. **18**(3): p. 251-263.

CHAPTER 8. MHI SUMMARY AND CONCLUSIONS

The MHI instrument described in chapter 6 makes use of a programmable optical filter to produce a high-throughput Raman spectrometer and hyperspectral imaging system. The instrument described is based on a Raman spectrometer with specially designed detection optics. The detection optics are based on a reflectance type liquid crystal spatial light modulator; a type of programmable optical filter that is programmed with filter functions optimized for analytes of interest, as well as filters to reproduce the functionality of a traditional spectrometer.

The results discussed in chapter 7 demonstrate that the MHI instrument can reproduce the functionality of a traditional Raman spectrometer by making use of Hadamard transform filter functions, as well as produce rapid chemical classifications, concentration measurements and chemical images by making use of specialized filter functions from PLS regression analysis or spectral angle mapping. Test samples of cyclohexane and *n*-hexane were used to illustrate the rapid classification ability of the MHI using spectral angle mapping. Mixtures of *n*-hexane and *n*-hexanol were also used to demonstrate the ability of the MHI to use PLS regression coefficients as optimized filter functions to determine the concentration of each component. Finally, a chemical map of an aspirin tablet with theophylline packed craters was produced rapidly using PLS-derived filters after pre-training with the pure components. This is when the true

advantage of the MHI is realized; after pre-training on components of interest to produce optimized filter functions.

Although the instrument described in the previous chapters focuses on Raman spectroscopy, the MHI detection strategy could be adapted to a variety of other spectral detection methods. For instance, one could think of a fluorescence based application where filter functions are optimized for different chromophores with highly overlapping emission spectra for high-speed sorting. The MHI detection method could also be adapted to study high-speed chemical kinetics. Filters could be optimized for variance between reactant, intermediate and product species and used to track the changes in their concentrations as a function of time in a variety of reaction types.

VITA

VITA

Amanda J. Hemphill was born on August 21, 1986 in Butler, Pennsylvania. Throughout her schooling, she was always drawn to science and mathematics and in 2004 she graduated from Butler Area Senior High School in Butler, Pennsylvania. Amanda then entered Slippery Rock University of Pennsylvania initially majoring in mathematics with the intent to teach secondary mathematics. After just one semester, she realized her true interest lie in the field of chemistry and switched majors. In 2008, she graduated *summa cum laude* with a bachelor of science in chemistry; she then entered the graduate program for a Ph.D. in physical chemistry at Purdue University in West Lafayette, Indiana. In 2010, Amanda received the National Science Foundation Graduate Research Fellowship. Her research involved Raman spectroscopy, as well as the study of biological diffusion and developing an experimental system to measure biological reaction kinetics. Amanda received her Ph.D. in Chemistry from Purdue University in May of 2015. Upon graduation, she intends to find a job in industry.

PUBLICATION

analytical chemistry

Multivariate Hyperspectral Raman Imaging Using Compressive Detection

Brandon M. Davis, Amanda J. Hemphill, Derya Cebeci Maltaş, Michael A. Zipper, Ping Wang, and Dor Ben-Amotz*

Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

S *Supporting Information*

ABSTRACT: A multivariate hyperspectral imaging (MHI) instrument has been designed and constructed to achieve greatly increased Raman imaging speeds by utilizing a compressive spectral detection strategy. The instrument may be viewed as a generalized spectrometer, which can function either as a conventional monochromator or in a wide variety of other hyperspectral modalities. The MHI utilizes a spatial light modulator (SLM) to produce programmable optical filters to rapidly detect and map particular sample components. A sequence of Hadamard-transform or random filter functions may be used to regenerate full Raman spectra. Compressive detection is achieved either using multivariate signal processing filter functions or the actual component spectra. Compressive detection is shown to be capable of achieving sampling speeds exceeding 1 ms per image pixel and the collection of chemical images in less than a minute.
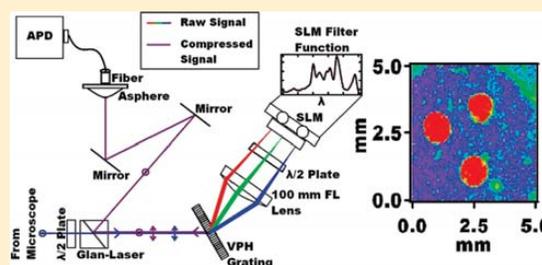
I ncreasing the speed at which Raman chemical imaging is performed requires a shift in paradigm. The last major advances in Raman imaging speed occurred with the introduction of multiplexing[1-4] and optical array detectors.[5] The recent development of compressive sampling data collection strategies,[6-12] which have been used to create such devices as single pixel cameras,[13] has set the stage for the creation of fast hyperspectral imaging instruments. Although previous studies have demonstrated the feasibility of such a hyperspectral detection paradigm,[6,8,11,12,14-16] none of these have realized its potential speed advantages. Here we describe the design and performance of a new multivariate hyperspectral imaging (MHI) instrument which is optimized to facilitate the rapid collection of chemical images using a low-noise single-channel compressive detection strategy.

Despite ongoing improvements in optical array detector technologies (such as charge-coupled device, CCD, and electron-multiplying-CCD cameras, as well as time-delayed integration, TDI, strategies),[5,17] the time required to collect hyperspectral images remains an obstacle to the wider application of Raman spectroscopy for chemical imaging. More specifically, point scanning with CCD-based spectral detection methods typically requires of the order of 1 s per spectrum and thus is impractical for the collection of large spectral images (as the collection of a 1 megapixel image would require 10^6 s or ~12 days). Global illumination[18-21] strategies with a tunable detection band-pass filter and a two-dimensional (e.g., CCD camera) detector can be used to rapidly collect single wavelength images but do so at the expense of discarding the Raman scattered photons outside the wavelength detection band. Thus, tunable band-pass filter imaging is fundamentally less efficient (slower)

than full-spectral detection. Line scanning[5,18] can provide a superior alternative, by distributing the laser power and simultaneously collecting full spectra from each point along a line. Both point and line scanning methods simultaneously collect full spectral information from each image pixel with a signal-to-noise (S/N) that is limited only by the sample's Raman scattering cross section, excitation laser power, and integration time.

A key advantage of the MHI detection strategy described in this work is provided by using a single channel detector to simultaneously detect all the photons transmitted by a multivariate optical filter of arbitrary programmable spectral shape. Thus, for example, a Raman spectrum with a total intensity of the order of 100 detected photons (counts) would have a S/N ~ 10 if all the photons were detected on an ideal single channel detector. However, if those same photons were spread over >100 wavelength channels (using a CCD detector), they would be practically undetectable, given that the signal (~1 count/channel) is less than the typical read-noise (of a few counts/channel). This is the fundamental reason why the MHI detection strategy can outperform conventional optical array detection methods. Moreover, the MHI detection strategy could potentially be implemented using low-cost SLM and single channel detector components[11] (although the research grade components used in the present system are comparable in price to a CCD/spectrograph, as further described in the Experimental Section and the Supporting Information).

It is important to stress that the full speed advantages of the MHI are only realized after pretraining to establish the optimal programmable filter functions for a given imaging application, with no pretraining the MHI can reproduce the functionality and speed of a conventional array detection scheme. Moreover, the MHI detection strategy is most advantageous (relative to full-spectral detection methods) under very high speed (or low signal) conditions, such that the total number of detected photons are comparable to (or smaller than) the number of full-spectral wavelength channels.

A few previous studies have demonstrated the implementation of hardware based spectral compression strategies.[6,10,11,16] Such instruments all use a single channel detector to integrate the light transmitted through optical filters of various spectral shapes. For example, the filters may be designed to reproduce the eigenfunctions obtained from chemometric techniques such as principal component analysis (PCA)[22,23] or partial least-squares (PLS).[24] In other words, such hardware spectral compression methods utilize spectral eigenvectors (or loadings) for detection rather than for postprocessing of full-spectral data (or for compressed data storage). An early implementation of this strategy by Myrick and co-workers used static optical interference filters with customized transmission spectra[6] for multivariate imaging and optical computing.[25] However, since the required filters are sample specific, this approach requires manufacturing different filters for each chemical imaging application. Moreover, with dependence on the complexity of the filter function, many dielectric layers may be needed, thus increasing the cost and decreasing the maximum optical throughput of the required filters.

Spatial light modulators (SLM) offer an attractive alternative, as they provide a means of producing variable programmable filter functions. Recent implementations of this strategy have utilized various types of SLM devices, such as liquid crystal panels[10−12,15] and digital micromirror device (DMD) arrays.[14,16,26] LC-based SLMs use optical polarization to produce either phase or amplitude modulated variable spectral filter functions.[27] DMD arrays provide binary filtering states, as each mirror is either "on" (reflecting toward the detector) or "off" (reflecting away). Liquid crystal (LC) based SLMs, on the other hand, provide a variable transmittance (or reflectance) gray scale and thus may readily be used to produce spectral filter functions of arbitrary shape (although a DMD can also reproduce this functionality by controlling the number and duration of mirrors that are in the "on" state during the data collection time). Previous applications using transmissive LC SLMs for spectral compression have suffered from low signal throughput (~20%).[11] Recently developed reflectance based LC-SLMs provide much higher throughput (>80%) by increasing the reflectivity and fill factor of the LC array[28] and thus have been selected for use in the present MHI instrument.

The unprecedented hyperspectral imaging speed of the MHI derives not only from the increased throughput and high contrast of the SLM but also from the use of a high-throughput volume holographic grating and a low noise avalanche photodiode detector. The MHI can be used either for full spectral detection, using band-pass or Hadammard-transform filters, or high speed compressive detection, using PLS or spectral angle mapping calibration and/or classification.

■ EXPERIMENTAL SECTION

Instrument Description. The MHI is built around a Raman microscope with backscattering collection geometry. The schematic
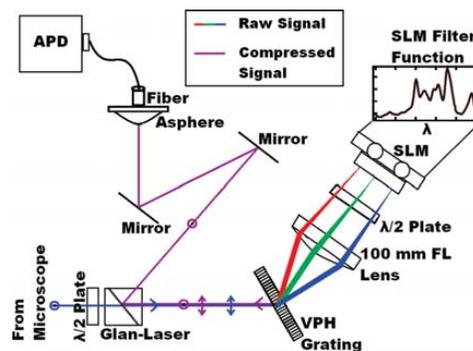

Figure 1. Schematic diagram of the MHI detection optics (see text and Supporting Information for details).

of the MHI detection optics is shown in Figure 1, and a detailed parts list is provided as Supporting Information. A 785 nm (single mode, 100 mW, $\Delta\lambda = 0.026$ nm) laser diode module (Innovative Photonic Solutions, Inc.) is used as the excitation source. The laser is directed through a 20× (NA 0.40) NIR objective (Olympus, LMPL20XIR) using a 45° dichroic long pass filter (Semrock, LPD01-785RU-25 × 36 × 2.0) which passes the collected (Stokes-shifted) Raman scattered light from the sample. Additional Rayleigh scatter rejection is achieved using a 785 nm notch filter (Semrock, NF01-785U-25) placed outside of the microscope.

After scattering from the sample, the Raman signal passes into the MHI detection optics, as illustrated in Figure 1. The first component in the detection optical path is a half waveplate, which is used to rotate the polarization of the Raman light and thus determining whether the V or H Raman scattering component is detected. The second component is a Glan-Laser polarizing beam splitter which transmits the p-polarized light toward the SLM, and reflects s-polarized light coming back from the SLM (with polarization contrast of at least >1000:1). The second half-waveplate, between the lens and SLM is required in order to align the light's polarization to match the liquid crystal axis of the SLM for optimal reflectance modulation. The MHI filter functions are produced by using the SLM phase modulation to control the degree to which the input p-polarized signal is rotated to s-polarization and thus reflected by the Glan-Laser polarizer into the detection optical path. The volume holographic grating (1200 L/mm, center wavlength 830 nm, Edmund Optics, 48-590) and an achromatic lens with a focal length of $f =$ 100 mm (Edmund Optics, 47-317) are used to disperse light and focus different wavelengths onto different pixels of the SLM. Note that the distances between the holographic grating, lens, and SLM are matched to f (see Figure 1) to ensure that the light reflected by the SLM retraces the same optical path as the incoming light, although its polarization is modulated in accordance with the filter function applied to the SLM.

The SLM (Boulder Nonlinear Systems Inc.) is composed of a linear array of 12 288 pixels. Each pixel is a separately addressable optical phase modulator which is used to rotate the polarization of the detected light between 0° (p-polarization) and 90° (s-polarization), only light that is rotated away from 0° will be reflected into the detection optical path by the Glan-Laser polarizing cube. The degree of polarization rotation of each SLM pixel is set with an eight-bit control voltage. The eight-bit
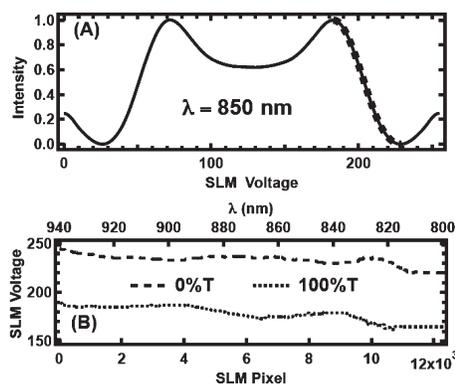
Figure 2. SLM reflectance calibration. (A) The measured collected light intensity at 850 nm (normalized to unit full scale) is plotted as a function of the voltage applied to the SLM. The thick dashed segment highlights the voltage range used for setting the % *T* of the corresponding SLM pixel. (B) Calibrated 100% *T* (dotted) and 0% *T* (dashed) SLM voltage settings are plotted as a function of both SLM pixel number (bottom axis) and wavelength (top axis).

reflectance scale of the SLM pixels allows the MHI to produce programmable spectral filter functions of variable shape, as illustrated in Figure 1. The birefringence of each SLM pixel is temperature and wavelength dependent.[29] The relation between the SLM voltage setting and percent transmission as a function of wavelength is shown in Figure 2 (and further details concerning the SLM transmission vs wavelength calibration are provided as Supporting Information).

The s-polarized signal which is reflected into the detection path is collected with an 18 mm focal length aspheric lens (Edmund Optics) and focused into a 200 μm core optical fiber (NA 0.22, with SMA terminals, Thor Laboratories Inc.) that is directly coupled to the APD detector (Hamamatsu, C4777-01). The APD module has a built-in temperature control and amplifier and was chosen for its low noise and high photosensitivity between 800 and 950 nm. The APD has an active sensor diameter of 3 mm and a measured photosensitivity of ~1.2 e9 V/ W. This current-integrating detector was selected for the MHI (rather than a photon-counting APD) as MHI Raman signals may exceed 10 MHz photon count rates. However, a photon-counting APD can provide lower noise for MHI signals corresponding to detection rates of the order 1 MHz or less (as further described in the Supporting Information).[30,31]

The voltage output of the APD is connected to a USB data acquisition device (National Instruments USB-DAQ 6211) for digitization and computer readout. The DAQ unit has a maximum read rate of 250 kHz and can store 250 000 samples before analog-to-digital conversion. The integration time of the measurements using the APD-DAQ is set by controlling the number of samples that are collected and signal averaged. The spatial, spectral, and time resolution performance of the MHI is further described in the Results and Discussion section (as well as in the Supporting Information).

MHI Detection Strategies. The MHI may be viewed as a generalized spectrometer which is capable of functioning either as a conventional scanning spectrometer or using other more efficient spectral detection strategies. Different detection

modalities utilize different spectral filter functions. For example, one may reproduce the functionality of a simple scanning spectrometer by using the SLM to produce band-pass filters with variable center wavelengths. The efficiency of the latter strategy is lower than that of a CCD-based full spectral instrument, since most of the Raman scattered light (away from the band-pass) is rejected by each filter. Alternatively, one may more efficiently collect full spectra using either Hadamard[3,4,32] or random (compressive sampling)[13] filter functions. The efficiency (speed) of the Hadamard transform detection method can approach that of a CCD-based point scanning, or line scanning, instrument (as half of the Raman scattered light is detected by each filter). By using random (compressive sampling) filters, one can in principle achieve higher speeds than when using Hadamard filters, since fewer filters are required to regenerate a complete spectrum.[33] Far more efficient (and faster) MHI spectral imaging data collection may be achieved using trained filter functions which are tuned to optimally detect particular components of interest. For example, this can be achieved either using filter functions whose shape is the same as that of the component spectra, or using filter functions whose shape reproduces loadings (eigenfunctions) optimized for a particular imaging application. In other words, high speed chemical imaging requires pretraining to construct MHI filter functions. Two such methods are demonstrated in this paper. One is spectral angle mapping using component spectral filter functions,[34] and the other utilizes PLS loading vectors as filter functions.[24] In both cases, the full Raman spectra used for training were obtained using the MHI with a Hadamard transform detection scheme.

For most applications, the measured SLM signals (i.e., APD voltages obtained when using a particular SLM filter function) are normalized by the integrated "all-on" signal, obtained by setting all of the SLM pixels to 100% *T* (in order to maximize the reflectivity of the SLM). In other words, the all-on measurement is used to normalize each SLM signal so that it represents the fraction of the total number of Raman scattered photons which are detected when using the corresponding SLM filter function.

Spectral angle mapping uses the analyte spectrum itself as the filter function.[34] This method is quite simple to implement as it requires only scaling the spectrum to a full intensity range of 0 to 1 for use as an SLM transmittance filter function. Note that any such SLM filter function can be represented by an *n*-dimensional vector, and the same is true for the Raman scattered light intensity at different wavelengths. Spectral angle mapping effectively measures the correlation coefficient (dot-product) of the latter two vectors by measuring the amount of the Raman scattered light which is reflected toward the detector by the SLM. The Raman scattered light (spectral vector) emerging from a given point in a sample is thus classified based on its correlation coefficient with each SLM filter function (detection vector), with appropriate classification cutoff values.

Alternatively, pretraining using PLS regression may be used to determine optimal MHI filter functions.[24] This method requires more computation effort prior to data acquisition than spectral angle mapping but can be advantageous as PLS regression maximizes the covariance between the spectra of each species and their concentrations, thus establishing a quantitative concentration metric built into the scaling of the loadings (filter spectral vectors). PLS also has increased selectivity since the PLS loadings are constructed such that they maximize the variance between the components of interest.[24]

Materials and Samples. The following liquid samples were used in order to validate the chemical classification and quantitation performance of the MHI: *n*-hexane (Mallinckrodt, 5189, 99.4%), *n*-hexanol (Sigma Aldrich, reagent grade, 98%, H13303), and cyclohexane (Sigma Aldrich, 99.9+% HPLC, 27 062-8). The high-speed chemical imaging performance of the MHI is demonstrated using the following pharmaceutical composite and a powder containing two types of sugar microcrystals. The pharmaceutical composite sample was produced using an aspirin tablet (Equate, lot no. 3CE0649) in which three small craters were created and packed with theophylline anhydrate (AMEND Drug and Chemical Co., lot no. Z52258K16). The theophylline powder was heated at 80 °C for 24 h to ensure that it was in the anhydrate form. The surface of the aspirin-theophylline tablet was shaved with a razor blade to create a flat imaging surface. The sugar imaging sample was created by distributing powders of D(−) fructose (Sigma-Aldrich, F-0127) and sucrose (Mallinckrodt Chemicals, 8360-04) granules over a glass microscope slide. The spatial resolution of the MHI was determined using a standard 1951 USAF test target (Edmund Optics Inc.), consisting of a 2 in. square clear (soda-lime glass) substrate with a chrome USAF test target pattern.

■ RESULTS AND DISCUSSION

Validation of MHI Imaging Performance. *Signal-to-Noise and Resolution.* The total signal throughput of the MHI detection optics is greater than 50%, as determined by comparing the intensity of incoming horizontally polarized incoming light before the MHI Glan-Laser polarizer, with that emerging from the detection optical fiber (which is coupled to the APD detector), as shown in Figure 1. These throughput measurements were performed at 785, 850, and 915 nm using diode-laser light sources. The resulting MHI throughput of >50% is significantly higher than that previously reported when using an transmission LC-SLM (with maximum transmittance of ~20%, before including the additional losses associated with other components in the detection path).[11,28]

The following measurements were performed in order to quantify the sensitivity and noise trade-offs associated with the MHI detection system, under typical experimental Raman detection conditions. The USB-DAQ device has a noise of ±12 μV per read, as measured using a constant input voltage (obtained from a battery with resistive voltage-divider). The dark voltage and noise (standard deviation) of the APD is ±0.21 mV, as measured with the APD completely shielded from light. The latter dark noise is equivalent to an optical power of approximately ±182 fW (at ~850 nm). The response time of the APD was determined to be approximately consistent with the manufacturer specifications of 67 μs. More specifically, the measured APD response time is 55 ± 16 μs (determined as described in the Supporting Information). The manufacturer specified APD photosensitivity of 2.7 × 10^{-10} V s/photon (at 850 nm) was also found to be consistent with our independent experimental measurements.

The actual S/N performance of the APD was compared to that expected using the following expression.

$$S/N_{APD} = \sqrt{n}\,\frac{S_L}{\sqrt{\sigma_D{}^2 + \sigma_L{}^2}} \qquad (1)$$

$S_L = V_{APD}\tau_{APD}/C_{APD}$ represents the light induced APD signal, expressed as a number of photons detected by the APD during its response time of $\tau_{APD} \sim 55\ \mu$s, where V_{APD} is the APD output

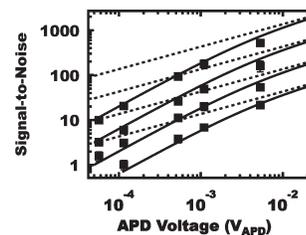**Figure 3.** MHI/APD S/N versus signal voltage at different integration times: τ_R = 100 ms, 10 ms, 1 ms, and 100 μs, from top to bottom. Experimental data points are compared to predictions obtained using eq 1 (solid curves) and to the performance of an ideal (dark noise free) electron–hole detector (dashed lines).
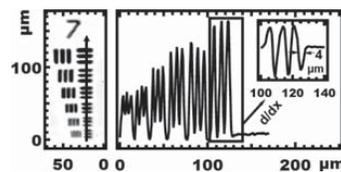

Figure 4. (Left) MHI reflectance image of 1951 USAF test target Group 7 with a 20× IR objective. (Right) Cross section of image along the arrow. (Inset) first derivative of cross section from within the rectangular box. The width of the negative-going peak in the first derivative represents the width of the edge of a target line, giving a spot size of 4 μm.

voltage and $C_{APD} \sim 2.7 \times 10^{-10}$ V s/photon is the APD photosensitivity. The total detection noise includes both the APD dark noise $\sigma_D \sim 43$ (expressed as an equivalent number of photons) and photoelectron–hole (Poisson) noise $\sigma_L = (2S_{APD})^{1/2}$, where the factor of 2 is included because each detected photon produces an electron–hole pair.[35] In order to obtain the optimal signal-to-noise, the APD must be sampled at a rate that is greater than $1/\tau_{APD} \sim 20$ kHz (and we typically sampled the APD at 250 kHz). Under these conditions $n = \tau/\tau_{APD}$ represents the number statistically independent APD output voltages which are measured in a total read time.

The points in Figure 3 were each obtained from 30 replicate APD signal measurements of Raman scattered light from a liquid *n*-hexanol sample using different excitation laser intensities (ranging from 1 to 100 mW) and integration times (from bottom to top) of 100 μs, 1 ms, 10 ms, and 100 ms. The experimental S/N of these measurements was then compared to predictions obtained using eq 1, which was used to generate the solid curves in Figure 3. The dashed lines in Figure 3 represent predictions for an ideal electron–hole detector (with no dark noise $\sigma_D = 0$). These results indicate that for APD voltages of 1 mV < V_{APD} < 10 mV, which are typical of the MHI Raman signal levels, the MHI signal-to-noise ratio is no more than about a factor of 2 less than that of an ideal electron–hole Poisson detector. A photon-counting APD may be used to further improve the MHI signal-to-noise performance, particularly under low signal conditions (as further discussed in the Supporting Information).

The spatial resolution of the MHI instrument is ~4 μm using the 20× objective, as illustrated in Figure 4 (see Supporting Information for further details). The MHI spectral resolution is
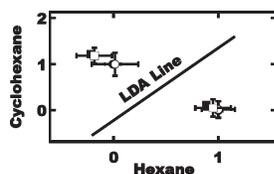
Figure 5. PLS (square points) and spectral equivalent (circle points) filter classification of *n*-hexane (upper-left points) and cyclohexane (lower-right points) samples at integration times 1 ms (solid points) and 100 μs (open points). Spectral equivalent filter measurements are plotted as the response to the cyclohexane filter (left axis) versus the response to the *n*-hexane filter (bottom axis). PLS filters are plotted as the predicted % *n*-hexane (right axis) versus the actual % *n*-hexane (top axis).

dictated by the sample excitation spot size imaged on the SLM. The resolution was determined to be ~18 cm^{-1} by measuring the fwhm of the 991 cm^{-1} benzene peak and the 811.7 nm line from an argon lamp. The latter spectra were measured with the MHI using the Hadamard transform spectral collection procedure. This resolution corresponds to a spot size on the SLM of ~190 μm, which is about a factor of 4 larger that expected under ideal imaging conditions (given the 4 μm spatial resolution at the sample and the 100/9~11× magnification of the collection optics at the SLM surface). The spectral range of the MHI is ~240 cm^{-1} to 2 100 cm^{-1}, which is dictated by the 19.6 mm width of the SLM (as well as the 1200 g/mm groove density of the holographic volume grating and the f = 100 mm focal length of the lens between the grating and the SLM, see Figure 1). Thus, the MHI effectively has ~103 independently addressable wavelength channels (as dictated by its full spectral range of 1 860 cm^{-1} and resolution of ~18 cm^{-1}). Further improvement in spectral resolution may be obtained by introducing a vertical slit in the detection path (as further described in the Supporting Information).

Liquid Classification and Concentration Measurements. Figure 5 shows the MHI classification results obtained using liquid cyclohexane and *n*-hexane samples and either PLS or spectral equivalent (spectral angle mapping) MHI filters, with an APD signal averaging time of either 1 or 0.1 ms per filter. The MHI signals from each pure liquid have all been scaled to an average value of ~1. The line in Figure 5 was obtained using linear discriminant analysis (LDA), which provides a classifying metric for distinguishing these two compounds based on their MHI signals. Although similar performance was in this case obtained using both the PLS and spectral angle mapping detection strategies, the same is not the case for samples whose Raman spectra are much more similar to each other, such as *n*-hexane and *n*-hexanol. Similar tests performed using the latter two liquids demonstrated that accurate classification (with standard deviation error bars smaller than the distance between the *n*-hexane and *n*-hexanol points) is possible using PLS filters with a 10 ms APD signal averaging time but not when using spectral angle mapping with spectral equivalent filters (as further described below).

Figure 6 shows Raman spectra and MHI filter functions for liquid *n*-hexane and *n*-hexanol, whose Raman spectra look quite similar (in the fingerprint region from 400 to 2000 cm^{-1}). These results were obtained using PLS filter functions derived from the Hadamard transform training spectra of *n*-hexane and *n*-hexanol
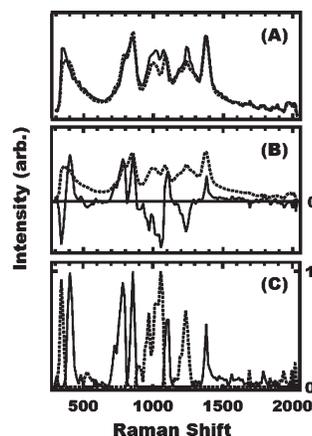

Figure 6. PLS-derived SLM analyte filter algorithm. (A) Pure component Hadamard spectra of *n*-hexanol (solid) and *n*-hexane (dashed). (B) PLS output regression vectors using spectra from part A. (C) Splitting of the second PLS regression vector (solid curve in part B) into a positive part (solid) and the absolute value of the negative portion (dashed).

shown in Figure 6A. The latter training spectra for the PLS algorithm are normalized to unit area (with a concentration vector that codes *n*-hexane as 0 and *n*-hexanol as 1). The SIMPLS[36] algorithm was used to perform the PLS analysis of the input spectra (using MATLAB along with the PLSToolbox, eigenvector Research, Inc.). The output from the SIMPLS algorithm for a two component system is a set of two regression vectors, as shown in Figure 6B; the solid curve in Figure 6B resembles the sum of the spectra of the two liquids, and the dashed curve in Figure 6B resembles the difference between the two spectra. Figure 6C illustrates the way that the latter PLS regression vector is split into two parts, each of which are nonnegative functions.[10] The latter two regression vectors are scaled to a maximum value of 1 (which corresponding to a maximum SLM transmittance) and the corresponding scaling constants, c^+ and c^-, are subsequently used to regenerate the full PLS response. Further details concerning the generation of MHI filter functions are provided as Supporting Information (which also includes comparisons of MHI and CCD spectra of various liquid and solid samples in Figure S-1 in the Supporting Information).

Figure 7 shows MHI concentration quantification results obtained for various *n*-hexane/*n*-hexanol liquid mixtures, using PLS filter functions defined as described above (and shown in Figure 6), with APD signal averaging times ranging from 1 to 100 ms. These MHI based concentration measurement results are remarkably good considering the similarity of the fingerprint spectra of these two liquids as well as the fact that only the pure liquid spectra were used to train the PLS filter functions. The correlation coefficient of the linear fit to the data points shown in Figure 7 increased from 0.959 to 0.997, as the integration time (per SLM filter measurement) increased from 1 to 100 ms, respectively.

Imaging of Solid Composites and Powders. The chemical image shown in Figure 8 was obtained from the aspirin/theophylline composite sample. Initial (training) measurements were performed by collecting five Hadamard spectra of each of the two pure components with a high S/N (~50:1) using the
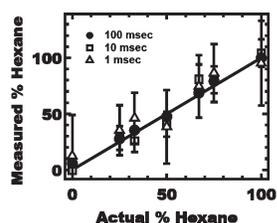
Figure 7. MHI concentration measurements are compared with the actual mixture concentrations results for n-hexane/n-hexanol liquid mixtures, with filter functions obtained using PLS and three different MHI integration times (per filter signal). The error bars (standard deviations of 10 repeated measurements) of the 100 ms integration results are similar to the size of the corresponding data points.
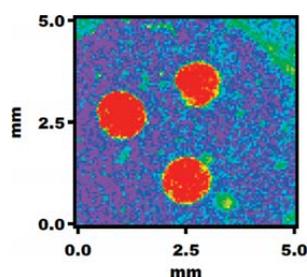

Figure 8. MHI chemical image (100 × 100 pixels) of an aspirin (blue) tablet with theophylline (red) packed craters at 1 ms per pixel per filter with a total integration time of 30 s. The image color ramp is continuous from blue (0) to green (0.5) to red (1).

Hadamard spectral collection strategy. The pure component spectra were then processed using PLS and classified with theophylline coded as 1 and aspirin coded as 0.

The chemical map shown in Figure 8 was obtained using a 1 ms APD signal averaging time with two PLS filters (for the positive and negative components of the second PLS loading vector) and one all-on filter. The total signal collection time was 30 s for the tablet chemical map in Figure 8. The average voltage response to the two PLS filters was typically 2—3 mV (with S/N ratios between 20:1 and 30:1, see Figure 4). When the integration time per filter was decreased to 100 μs, the total signal collection time decreased to 3 s, with a S/N ratio of ~10. The actual data measurement time may be longer than the signal collection time, depending on the dead-time associated with moving the sample stage and reading the signal. When reading the MHI signal continuously while linearly raster-scanning the stage we have collected MHI images in a total time that is less than 1.5 times longer than the associated total signal collection time (and further improvement is undoubtedly possible).

Sugar Crystal Mixture Imaging. Sugars have relatively weak Raman scattering intensities. The average APD voltage reading obtained from powders of sucrose and fructose (with the PLS filters) was ~300 μV, which is about 10 times lower than that obtained for aspirin and theophylline (as well as various liquid samples). A powder composed of mixture of sucrose and fructose microcrystals was dusted onto a glass microscope slide. The region that was imaged contained two large sucrose crystals surrounded by smaller crystals composed primarily of fructose,
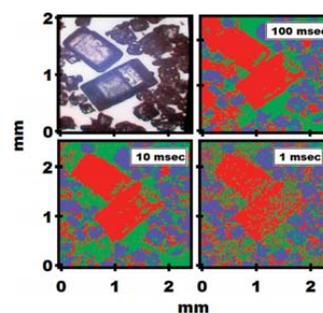

Figure 9. MHI chemical image (100 × 88 pixels) of sucrose (red) and fructose (violet) crystals spread on a glass microscope slide (green) at different integration times.

separated by regions revealing the glass substrate. Three sets of MHI filters, obtained using PLS training, were used to obtain the classification image shown in Figure 9. The top-left panel in Figure 9 contains a white-light image of the sugar microcrystal sample, and the other three panels contain MHI chemical classification maps obtained using different integration times (per pixel, for each SLM filter function). Good chemical images were obtained with integration times down to 1 ms, although the accuracy of classification clearly improved with increasing integration time.

■ CONCLUSIONS

The MHI instrument design described in this work utilizes a programmable optical filter to produce a high-throughput near-infrared micro-Raman spectrometer and hyperspectral imaging system. The results demonstrate that this instrument can either reproduce the functionality of optical array based spectrometers (using Hadamard transform filter functions) or far more rapidly collect hyperspectral images using either PLS or spectral angle mapping filter functions derived from the sample components of interest. The latter speed advantage can be of the order of 100 or more, as it is approximately equal to the ratio of the number of full-spectral wavelength channels and the number of MHI filters that are required for a particular imaging application.

Although we have focused on hyperspectral Raman imaging applications, the MHI detection strategy can readily be adapted to a wide variety of other high speed spectral detection applications. For example, the MHI may be used to increase the multiplexing capability of fluorescence-based bioarray sensing and high-speed sorting applications. More specifically, programmable fluorescence detection filters trained using multivariate signal processing algorithms may be used to distinguish chromophores with highly overlapping emission spectra and thus increase the multiplexing capability relative to that obtained using conventional fluorescence band-pass detection strategies.[37−39] Moreover, the MHI compressive detection strategy may be used for chemical kinetics measurements with millisecond (or faster) time resolution. More specifically, programmable filters trained to project reactant, product, and/or intermediate species may be used to track time dependent concentration changes in liquids, solids, polymers, glasses, or biological samples.

The fundamental advantages of the MHI detection strategy, relative to optical array (CCD) based micro-Raman detection schemes, derive from the lower noise and higher speed of the

MHI. More specifically, the detection limit of a CCD-based point- or line-scanning micro-Raman system is limited by the CCD read-out speed and noise. The MHI detection scheme is expected to outperform a conventional micro-Raman detection strategy for signals which approach or are below the latter CCD detection limits. For example, if a Raman signal consisting of a total of ~3000 counts measured in an integration time of 1 ms were distributed over ~1000 pixels of a CCD with a read noise of ~3 counts/pixel, the resulting spectrum would have an average S/N of ~1. Equation 1 (and Figure 3) indicates that the same signal would produce an MHI APD voltage of ~0.8 mV and a S/N of ~15. Moreover, a photon counting APD detector may be used to further improve the MHI signal-to-noise and detection limit (as described in the Supporting Information).

■ ASSOCIATED CONTENT

ⓢ Supporting Information. Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: bendor@purdue.edu.

■ ACKNOWLEDGMENT

■ REFERENCES

(1) Fellgett, P. *Journal De Physique Et Le Radium* **1958**, *19*, 187.
(2) Michelson, A. A. *Astrophys. J.* **1913**, *37*, 190.
(3) Decker, J. A.; Harwitt, M. O. *Appl. Opt.* **1968**, *7*, 2205.
(4) Ibbett, R. N.; Aspinall, D.; Grainger, J. F. *Appl. Opt.* **1968**, *7*, 1089.
(5) Golcuk, K.; Mandair, G. S.; Callender, A. F.; Sahar, N.; Kohn, D. H.; Morris, M. D.; , *Biochim. Biophys. Acta—Biomembr.* **2006**, *1758*, 868, DOI: 10.1016/j.bbamem.2006.02.022.
(6) Nelson, M. P.; Aust, J. F.; Dobrowolski, J. A.; Verly, P. G.; Myrick, M. L. *Anal. Chem.* **1998**, *70*, 73.
(7) Duarte, M. F.; Davenport, M. A.; Takhar, D.; Laska, J. N.; Sun, T.; Kelly, K. F.; Baraniuk, R. G. *IEEE Signal Process. Mag.* **2008**, *25*, 83.
(8) Kearney, K. J.; Corio, M.; Ninkov, Z. In *Sensors and Camera Systems for Scientific, Industrial and Digital Photography Applications*; SPIE: Bellingham, WA, 2000; Vol. 3965, p 11.
(9) Myrick, M. L.; Soyemi, O. O.; Haibach, F. G.; Zhang, L.; Greer, A. E.; Li, H.; Priore, R. J.; Schiza, M. V.; Farr, J. R. Application of multivariate optical computing to near-infrared imaging. In *Vibrational Spectroscopy-Based Sensor Systems*; SPIE: Bellingham, WA, 2002; Vol. 4577, p 148.
(10) Sweatt, W. C.; Boye, C. A.; Gentry, S. M.; Descour, M. R.; Stallard, B. R.; Grotbeck, C. L. In *Imaging Spectrometry IV*; SPIE: Bellingham, WA, 1998; Vol. 3438, p 98.
(11) Uzunbajakava, N.; de Peinder, P.; Hooft, G. W. t.; van Gogh, A. T. M. *Anal. Chem.* **2006**, *78*, 7302.
(12) 2005.
(13) Baraniuk, R. G. *IEEE Signal Process. Mag.* **2007**, *24*, 118.
(14) DeVerse, R. A.; Hammaker, R. M.; Fateley, W. G.; Graham, J. A.; Tate, J. D. *Am. Lab.* **1998**, *30*, 112S.
(15) Prakash, A. M. C.; Stellman, C. M.; Booksh, K. S. *Chemom. Intell. Lab. Syst.* **1999**, *46*, 265.
(16) Quyen, N. T.; Da Silva, E.; Dao, N. Q.; Jouan, M. D. *Appl. Spectrosc.* **2008**, *62*, 273.
(17) Jacak, J.; Hesch, C.; Hesse, J.; Schutz, G. J. *Anal. Chem.* **2010**, *82*, 4288.
(18) Delhaye, M.; Dhamelincourt, P. *J. Raman Spectrosc.* **1975**, *3*, 33.
(19) Morris, H. R.; Hoyt, C. C.; Treado, P. J. *Appl. Spectrosc.* **1994**, *48*, 857.
(20) Morris, H. R.; Hoyt, C. C.; Miller, P.; Treado, P. J. *Appl. Spectrosc.* **1996**, *50*, 805.
(21) Turner, J. F.; Treado, P. J. *Infrared Technology and Applications XXIII, Parts 1 and 2*; SPIE: Bellingham, WA, 1997; Vol. 3061, p 280.
(22) Pearson, K. *Philos. Mag.* **1901**, *2*, 559.
(23) Hotelling, H. *J. Educ. Psychol.* **1933**, *24*, 417.
(24) Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta* **1986**, *185*, 1.
(25) Simcock, M. N.; Myrick, M. L. *Appl. Opt.* **2007**, *46*, 1066.
(26) Hanley, Q. S.; Verveer, P. J.; Jovin, T. M. *Appl. Spectrosc.* **1998**, *52*, 783.
(27) Armitage, D.; Thackara, J. I.; Clark, N. A.; Handschy, M. A. *Mol. Cryst. Liq. Cryst.* **1987**, *144*, 309.
(28) Harriman, J.; Serati, S.; Stockley, J. *Proc. SPIE* **2005**, *5930*, 59302D.
(29) Zhang, Y.; Wu, L. Y.; Zhang, J. In *4th International Symposium on Instrumentation Science and Technology (ISIST' 2006)*, 2006; Vol. 48, p 790.
(30) Lawrence, W. G.; Varadi, G.; Entine, G.; Podniesinski, E.; Wallace, P. K. In *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues VI*; 2008; Vol. 6859, p M8590.
(31) Stewart, C. C.; Woodring, M. L.; Podniesinski, E.; Gray, B. *Cytometry, Part A* **2005**, *67A*, 104.
(32) Treado, P. J.; Morris, M. D. *Anal. Chem.* **1989**, *61*, A723.
(33) Donoho, D. L. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289.
(34) Yuhas, R. H.; Goetz, A. F. H.; Boardman, J. W. Discrimination Among Semi-Arid Landscape Endmembers Using The Spectral Angle Mapper (SAM) Algorithm. In *Summaries of the 3rd Annual JPL Geoscience Workshop*, 1992; Vol. 1, p 147.
(35) Schottky, W. *Annalen Der Physik* **1918**, *57*, 541.
(36) Dejong, S. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 251.
(37) Watson, D. A.; Brown, L. O.; Gaskill, D. R.; Naivar, M.; Graves, S. W.; Doorn, S. K.; Nolan, J. P. *Cytometry, Part A* **2008**, *73A*, 119.
(38) Perfetto, S. P.; Chattopadhyay, P. K.; Roederer, M. *Nat. Rev. Immunol.* **2004**, *4*, 648.
(39) Baumgarth, N.; Roederer, M. *J. Immunol. Methods* **2000**, *243*, 77.