Next generation microfluidic platforms for high-throughput protein biochemistry

Sebastian J Maerkl

DNA technologies such as cloning, DNA microarrays, and next generation sequencing have transformed the life sciences. Protein technologies on the other hand have not seen such explosive progress. This is mainly due to the inherent difficulty of working with proteins because of their manifold physical characteristics as opposed to the well behaved and well understood DNA polymer. Recent technological advancements have increased the throughput of protein biochemistry to levels where it is becoming of interest to systems biology. Here I review methods for high-throughput in situ synthesis and characterization of proteins and their integration with microfluidic devices. In the near future, the use of gene synthesis, microfluidic based protein synthesis and characterization will give rise to a resurgence of protein biochemistry in the current world of high-throughput genomics.

The difference between protein and DNA is well exemplified by comparing DNA and protein microarrays and the respective impact each has had on the life sciences. The first DNA arrays were generated by physically spotting cDNAs generated with PCR and synthesized oligos [1]. These oligos were obtained by synthesis in high-throughput in situ synthesis was performed by comparing DNA and protein microarrays and the respective impact each has had on the life sciences. The first DNA arrays were generated by physically spotting cDNAs generated with PCR and synthesized oligos [1]. These oligos were obtained by synthesis in high-throughput and yield, and the PCR could produce large quantities of DNA for spotting. The PCR product could also be purified in sufficient amounts to generate hundreds of DNA microarrays. Other advances to generating DNA microarrays included the photolithographic synthesis of oligos directly on the glass substrate at high spot densities [2,3]. For the detection of the mRNA sample, fluorescent nucleotides are incorporated into cDNA during a reverse transcription reaction. Because oligos can be synthesized with high fidelity and in high quantities, and because DNA is a stable molecule, DNA microarrays can be fabricated robustly and cheaply. A large number of biologists were thus able to apply DNA microarrays to their research and find new innovative uses for them including gene expression profiling, SNP detection, tiling arrays, and ChIP-chip applications.

Protein arrays were predicted to impact our understanding of protein structure and function, in the same way that DNA arrays provided insight into gene expression and regulation [4,5]. Unfortunately generating and applying protein arrays had one significant impediment: proteins! Thousands of purified proteins were required for a protein array. The spotting process itself was also suboptimal as proteins are prone to unfolding, especially when deposited and dried on a substrate. Protein arrays not only required an exorbitant amount of work, but also

Introduction

Proteins are the arbiters of cellular function; they control cell division, transduce internal and external signals, and regulate the expression state of the genome, in addition to performing a plethora of other functions. All these functions require proteins to bind to other proteins, DNA, RNA, or small molecules. The resulting complexes can either be directly functional, or lead to functional modifications such as phosphorylation of the components of the complex. It is these interactions and enzymatic activities that allow proteins to perform such diverse functions. Proteins are uniquely suited to these tasks as they are built from 20 different building blocks, that can be combined in an astronomically large number of ways, giving rise to proteins with many different folds and physical characteristics. It is this enormous plasticity that makes proteins such fantastic all-purpose machines, but is also the very reason why working with proteins is difficult.

Many useful methods have been developed for working with and manipulating DNA (incidentally almost all of them are based on proteins). DNA can be cut at precise positions using restriction enzymes, combined together with ligases, and amplified with polymerases. DNA can also be chemically synthesized, purified, and sequenced. For proteins these methods are either much more labour intensive (purification, synthesis, and sequencing) or do not exist at all (digestion, ligation, amplification). This lack of simple and scalable methods for protein biochemistry severely impacts the amount of time and effort researchers need to invest in characterizing a protein, leaving protein biochemistry in the low-throughput, pre-genome era.
yielded less reliable interaction measurements. In DNA arrays, two ssDNA probes need to anneal on the array to be detected. These interactions generally have a $K_d$ in the picomolar range and a kinetic dissociation rate on the order of $10^{-4} \text{s}^{-1}$ to $10^{-5} \text{s}^{-1}$ ($t_{1/2} = 96 \text{ min to 2 hours}$) even for relatively short 20 bp long oligos [6]. Protein interactions on the other hand can range from micromolar to picomolar affinities and kinetic dissociation rates can reach up to $10 \text{s}^{-1}$ ($t_{1/2} = .07 \text{ s}$) [7,8]. Interactions with low-affinity and high off-rates are lost during stringent wash steps necessary to reduce non-specific background signal. Even though protein arrays in principle are a powerful tool, the difficulty of generating the arrays, and the fact that protein interactions are significantly more difficult to measure than DNA interactions, have thus far prohibited protein arrays from reaching their full potential.

Recent technological developments have addressed the issues associated with first generation protein arrays. The invention of DNA to protein arrays (DTPAs) using in vitro transcription/translation (ITT) circumvents many of the difficulties associated with cloning, expressing, purifying, and spotting of proteins. DTPAs have been shown to be able to produce hundreds to thousands of proteins in parallel, making it possible for a single researcher to generate large protein arrays. The second technological advance is the integration of novel and existing methods for measuring molecular interactions allowing proteins to be characterized quantitatively and with better sensitivity. The integration of DTPAs and novel detection mechanisms into a single microfluidic device platform is a significant step toward automating these methods, and increasing throughput in protein biochemistry.

**In vitro protein synthesis**

In 2004 Ramachandran et al. developed a method for transforming spotted DNA arrays into protein arrays, termed nucleic acid programmable protein array (NAPPA) [9,10**]. Expression-ready biotinylated plasmid DNA is co-spotted on a glass substrate together with a biotinylated antibody. Both the plasmid DNA and antibody are immobilized to the substrate surface by streptavidin. To transform the DNA array into a protein array the entire substrate is submerged in ITT reaction, which transcribes and translates the plasmid DNA into protein. The synthesized protein is then locally immobilized by the co-spotted antibody. This approach has multiple advantages over spotted protein arrays: firstly, plasmid DNA can be generated using established molecular biology methods amenable to high-throughput; secondly, spotted DNA/antibody arrays are more robust than spotted protein arrays; and thirdly, the proteins are synthesized and purified in one step, eliminating the most labor intensive aspects of generating spotted protein arrays.

NAPPA effectively solved the problem of generating protein arrays, but it also has a few shortcomings. The use of plasmids as the template is suboptimal as plasmid generation also requires long and tedious cloning and transformation steps. Angenendt et al. have shown that PCR products can serve directly as templates for protein synthesis and protein array production, further streamlining the approach [11]. A second problem associated with the NAPPA method is that synthesis is not compartmentalized. This gives rise to large spots and requires large distances between spots to avoid cross-contamination in turn limiting spot density. Tao and Zhu developed an interesting variation based on spotted mRNA molecules and puromycin based capture of the synthesized proteins [12]. This method has been shown to produce very small spots, and contamination is eliminated as only protein synthesized locally from the spot is effectively captured by the puromycin. Of course mRNA is not the most optimal molecule to be used in the NAPPA approach, but the approach by Tao and Zhu may give rise to interesting alternative applications.

Integration of DTPAs and microfluidics has recently been demonstrated for the synthesis and characterization of S. pneumoniae proteins [13*] and synthetic transcription factor mutants [14]. In microfluidic DTPAs, DNA templates are spotted on an epoxy coated slide. The DNA array is aligned to a microfluidic device with up to 2400 unit cells [14,15**] so that each DNA spot is enclosed by a microfluidic chamber. To generate the DTPA the device is loaded with ITT reaction. The proteins are ultimately captured by antibodies in an adjacent chamber, generating a protein array (Figure 1). There are several advantages to combining DTPAs with microfluidics. First, all synthesis reactions are compartmentalized and completely segregated from one another, eliminating cross-contamination. Compartmentalization also gives rise to the possibility for combinatorics and multiplexed expression of proteins. For example, two proteins can be co-expressed in a single chamber and consequently tested for interaction [13*]. Alternatively it is also possible to express a protein in the presence of a small target molecule such as dsDNA oligos or drugs to assess the effect of these small compounds [14,16]. Most importantly, integration of DTPAs with microfluidics also allows integration with advanced detection mechanisms, and therefore simultaneously addresses both shortcomings of the classical protein arrays, namely protein synthesis and detection. The various methods for integrated detection are discussed in the next section.

**Integrated interaction measurements**

Sensitive and scalable methods for protein characterization are equally important aspects for high-throughput protein biochemistry, aside from large-scale protein synthesis. These generally focus on measuring molecular interactions and characterizing the parameters governing these interactions. Detection methods can be roughly divided into two categories. The first category of methods...
requires the fluorescent labeling of at least one component being measured. Fluorescent resonance energy transfer (FRET) [17] and fluorescence correlation spectroscopy (FCS) or fluorescent cross-correlation spectroscopy (FCCS) [18] are the best known examples. A recently developed method based on the mechanically induced trapping of molecular interactions (MITOMI) [15**] is another method which requires fluorescent labeling of one component and surface localization of the interaction. The methods in the second category do not require labeling but do require that the interactions are localized to a surface. Surface plasmon resonance (SPR) is the most well-established method of this kind [19,20]. Other promising new technologies are based on nanowires, optical microcavities, and nanomechanical resonators.

FRET

FRET occurs between two fluorophores, a donor and acceptor pair, that are in close proximity (distance of 10 nm or below). If the donor is excited and the acceptor is in close proximity, the energy can be transferred radiationless to the acceptor, which then emits a photon at a red shifted wavelength (Figure 2a). FRET can thus be used as a molecular ruler, as the transfer between donor and acceptor is strongly dependent on the distance between the two. Ridgeway et al. have used a microfluidic device and FRET to measure the kinetics of binding of a rRNA and ribosomal protein [21*] (the authors also used FCS, which is described in the next section). Because FRET is very sensitive to distance it generally requires precise knowledge of the molecule and is much more readily applied to DNA or RNA containing complexes than pure protein–protein interactions. FRET is unique in its ability to provide information on the conformational state of a protein or protein complex and thus makes it an interesting method for the high-throughput interrogation of protein dynamics as a function of sequence space.

Another FRET method that is suitable for measuring protein interactions is the use of two fluorescent proteins, each coupled to a different target protein (A and B). Interaction of the proteins ‘A’ and ‘B’ brings the two fluorescent proteins together leading to FRET (Figure 2a). The advantage of this method is that the fluorescent proteins can be genetically encoded, and large libraries of chimeric proteins can be generated. The drawback of FRET based methods is the inherent sensitivity of FRET to distance. It is thus possible that two interacting proteins are geometrically oriented in such a way that the two fluorescent proteins are not in close enough proximity for efficient FRET.

FRET is thus somewhat suboptimal for large-scale exploratory methods for detecting new protein interactions. Nonetheless it remains an exceedingly sensitive and informative method for characterizing known and well defined interactions, and in the right system could very well be scaled to high-throughput for interrogating protein sequence–structure relationships.

FCS/FCCS

FCS and FCCS are methods based on measuring the dwell time of a fluorescent molecule in a small illuminated volume (Figure 2b). The dwell time of molecules...
in the illumination spot and therefore the autocorrelation function of the sample depend on the diffusion coefficient ($D$) of the molecule. FCS can therefore measure the concentration of a sample, as well as determine the diffusion coefficient of a molecule. Binding of another molecule to the labeled molecule causes $D$ to decrease which is detected by FCS. In FCCS both molecules are labeled, and can be said to be interacting if their fluorescence signals are temporally correlated with one another. A few reports show that FCS or FCCS can be coupled to microfluidic devices. Chou et al. quantitated epidermal growth factor receptor (EGFR) and its interaction with Src and STAT3 [22]. Ridgeway et al. implemented a two-photon detection mechanism to reduce the amount of photobleaching of the sample and with single molecule sensitivity [21*].

Unlike with FRET based methods, steric constraints on the distance between the fluorescent proteins is not important in FCCS. The drawback is that FCS and FCCS require fairly dilute samples, which in turn increase the time required to acquire a statistically sufficient number of counts leading to lengthy interrogation times. Even a 1 min interrogation time per sample would require a total of 16 hours for 1000 samples, and interrogation times of 10 min and more are not uncommon. Spatial parallelization of FCS and FCCS measurements may be possible through diffractive optical elements (DOEs) [23], or the use of EMCCD cameras as detector arrays [24,25]. The latter method may profit in particular from the integration with microfluidics, as the z-dimension of the sample depth can be precisely controlled in a microfluidic environment without the need for specialized optics. FCS and FCCS are promising optical methods that could see increased application to high-throughput protein biochemistry.

MITOMI
MITOMI is a novel opto-mechanical method for measuring molecular interactions [15**] (Figure 2c). The method is based on capturing interacting molecules between two surfaces on a highly integrated microfluidic device fabricated by multilayer soft lithography [26,27], followed by optical quantitation of the trapped molecules. To measure the interaction of a two component system, say between protein A and protein B, protein A is localized to the surface. Protein B is added to the solution and
is allowed to interact with protein A. At this point the roof of the microfluidic channel in which the interaction takes place is collapsed, bringing it in direct contact with the surface to which protein A has been localized. Consequently, all unbound solvent and solute molecules, including unbound protein B, are excluded from the detection area, leaving only protein A and the bound fraction of protein B behind. For detection purposes protein A and B can be labeled with fluorescent antibodies, or if synthesized by ITT, through residue specific incorporation of a fluorescently labeled lysine, or fluorescent proteins. MITOMI is able to derive absolute binding affinities through multiple dilution series. MITOMI also can capture interactions between weakly associating molecules and transient complexes. The kinetics of association and dissociation can also be measured by MITOMI, through the rapid and precise temporal control of the time that the detection area is available for interaction. Using this approach hundreds of on-rates and off-rates can be measured in parallel on a single device [28, 29].

MITOMI is exceedingly easy to scale. No complex optical setups are required to readout the devices. In the first MITOMI experiments 2400 interactions were measured in parallel on a single device, and multiple experiments could be run in a short amount of time by a single researcher. It was also shown that MITOMI is applicable to a wide variety of interaction measurements including protein–DNA [14, 15], protein–RNA [16], and protein–protein [13]. In one experiment it was shown that 100 transcription factor variants could each be measured against 64 target DNA sequences requiring over 19 200 independent interaction measurements [14]. In another example 43 S. pneumoniae proteins were characterized in a total of 14 792 interaction experiments [13]. The current throughput of MITOMI is orders of magnitude above any other currently available method, speaking for its ease of integration and simplicity.

**SPR**

SPR is probably the best known method for characterizing the kinetics of protein interactions [19] (Figure 2d). SPR requires a gold surface which is derivatized with a protein layer. Binding of molecules to the gold layer can be detected as a change in incident angle or wavelength, which is dependent on the refractive index of the interface. Mass transfer to the surface can be detected in real-time, allowing on-rates and off-rates to be measured. SPR was integrated with microfluidics numerous times [30–32]. The most recent example comes from Ouellet et al., who adopted a complex microfluidic device [26] to perform 264 independent SPR measurements [33].

**Nanowires, optical microcavities, nanomechanical resonators**

A number of novel methods are showing potential for characterizing molecular interactions and integration with microfluidic devices. Many of the methods mentioned in this section are still in early development, but will likely reach adequate maturity to assess whether they can be applied to large-scale protein interaction measurements.

Nanowires are a promising new method for detecting protein interactions in high-throughput [34, 35] (Figure 2e). Binding of molecules to a nanowire changes the conductivity of the wire which can be readily detected. As the readout is purely electrical and does not require any optics, nanowires promise to be a very effective method for integration with microfluidic devices.

Optical sensors based on ultrahigh quality factor whispering-gallery microcavities have recently been demonstrated to be able to detect label-free single molecule binding events [36]. This potential for ultra-high sensitivity makes microcavities interesting candidates where sensitivity is the most important parameter [37]. Integration with microfluidics is somewhat more complicated and highly parallel measurements are not yet feasible.

Nanomechanical resonators have recently been applied to measuring the binding of molecules to the surface of the resonator. Nanomechanical resonators measure changes in their weight through changes in their resonance frequency. High sensitivity generally requires the resonators to resonate at high-frequency, which excludes their use in liquids (or air). Burg et al. have developed an ingenious method that circumvents the damping problems associated with liquid, by routing the liquids inside of the cantilever [38]. As with optical microcavities integration and fabrication of these devices is more complicated, but also may soon see sufficient improvement to make the technique accessible.

**Conclusion**

Recent technological and methodological advances are enabling protein biochemistry by drastically increasing the throughput of protein synthesis and protein characterization. DTPA methods make it possible to synthesize thousands of proteins in parallel. MITOMI microfluidic platforms integrate DTPAs and sensitive detection methods to both synthesize and characterize protein in high-throughput. MITOMI is currently the only high-throughput, integrated platform that has been applied to actual systems biology problems, such as measuring the binding energy landscape of transcription factors, generating protein-interaction networks, and measuring the functional consequence of mutations in transcription factors. Other detection methods such as FRET, FCS, and SPR, will also likely soon be scaled up and applied to pertinent problems in the field, but current examples are either low-throughput, or have only been applied to proof of principle systems. Other promising methods based on nanowires, optical microcavities, and nanomechanical resonators will likely require more time to reach the same
6 Analytical biotechnology

level of integration, but useful high-sensitivity platforms for diagnostics and related problems can be expected sooner. It is an exciting time for the development of high-throughput methods for protein synthesis and characterization. Genomic sequences have already become a commodity, and gene synthesis costs are expected to fall as drastically as sequencing costs, making it soon possible to order hundreds to thousands of synthetic genes (microfluidics may also prove to be the enabling technology in that field). Together, these two methods will make it possible to obtain expression ready DNA templates for genes from any organism, including unculturable and pathogenic ones. Gene synthesis will also enable the generation of defined sequence variants of a given protein. These methods in combination with novel microfluidic platforms such as MITOMI, will make it possible to express and characterize proteins from synthetic DNA, providing a unique opportunity to answer a wide variety of questions related to protein structure and function that were previously inaccessible.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
●● of outstanding interest


A recent example of the DNA to protein microrray technology developed by Ramachandran et al. [9]. The authors show the feasibility of arraying 1000 human cDNA templates, which were transcribed and translated into a protein array in situ.


The authors demonstrate an integrated microfluidic platform for on-chip protein synthesis and detection of protein interactions.


The authors describe an integrated high-throughput microfluidic platform for the quantitative characterization of protein-DNA interactions, using a novel detection mechanism based on mechanically induced trapping of molecular interactions.


A good example of the use of both FCS and FRET on a microfluidic reactor for studying RNA-protein interactions.


Demonstration of a parallel nanowire array for the detection of various cancer markers.

The first demonstration of the use of high-Q optical microcavities for label-free biomarker detection at the single-molecule level.


The authors apply nanomechanical cantilevers to the detection of biological molecules and cells.