SMP: Protein

Project: Implementation and Application of Chemical Microarrays

Ronald Frank – GBF (German Research Centre for Biotechnology), Braunschweig - frank@gbf.de

Introduction

During the last few years, chemical genomics has emerged from a pharmacological genetics approach to a powerful concept for the study of the individual steps and pathways of complex phenotypic expressions such as the cell proliferation cycle, cell differentiation or molecular mechanisms underlying various diseases [1,2]. Chemical genomics uses small organic molecules as tools to interfere selectively with protein function (chemical knock-down). Highly parallel and combinatorial chemical synthesis nowadays provide libraries of compounds with diverse chemical structures for biological activity screening comparable in number with molecular biology based nucleic acid or protein libraries. Screening of these chemical libraries is carried out with either phenotype directed cell assays (*forward* chem. genetics) or *in vitro* protein binding assays followed by cell-based phenotype validation (*reverse* chem. genetics).

The chemical genomics approach complements other genetic approaches (random and directed mutagenesis, gene knock-out cellular and animal models, antisense and RNA-mediated interference with transcripts (RNAi); see Fig.1) by interfering with gene function at the level of the gene products, the proteins. Small synthetic compounds have several unique features which make these tools extremely valuable: they can be selected to pass membrane barriers; they can be applied and withdrawn in a timely and spatially controlled manner; they are small and thus more selective; they have no principle interference with other types of cellular molecules as do siRNAs and RNA-aptamers; and finally, they are drug-like and most suitable for studies on gene function *in vivo*! The small molecule approach allows high quality conditional gene knock down experiments *in vitro*, in cell culture and animal models and has been shown to address targets that are otherwise difficult to study, such as multifunctional proteins.

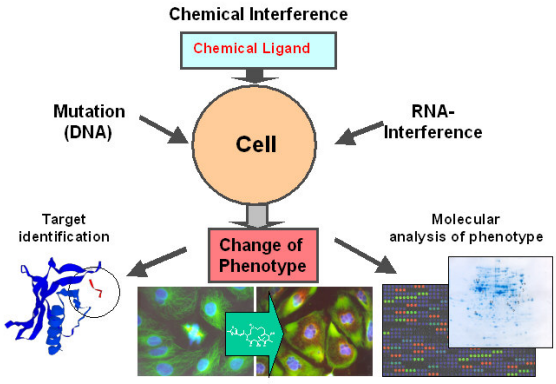

Fig 1: *Three complementary approaches to functional genome analysis*

The empirical search in large compound collections or combinatorial libraries for small molecules with suitable biological activity is a multi-disciplinary challenge with many diverse technical and strategic approaches. The successes and efficiencies of these depend on competent compound collections and intelligent assay configurations, and correlate with the degree of integration, miniaturization and automation of both, chemical synthesis and biological assay performance.

Concerning the high-throughput screening of chemical compounds in a high density immobilized format, chemical microarrays have become available only very recently [3]. We have pioneered the array-directed parallel miniaturized chemical synthesis by the development of the SPOT-synthesis (Fig.2). During NGFN-1 funding we have advanced this technology towards further miniaturized formats analogous to nucleic acid and protein micro-arrays on glass slides. For the goals of this SMP Protein our chemical micro-arrays will be applied to further analyze and validate protein interactions as discovered by the other SMP activities. Production of the chemical libraries will also allow the synthesis of soluble compounds which are being applied in screening high-density protein arrays bound to nitrocellulose membranes or glass slides.

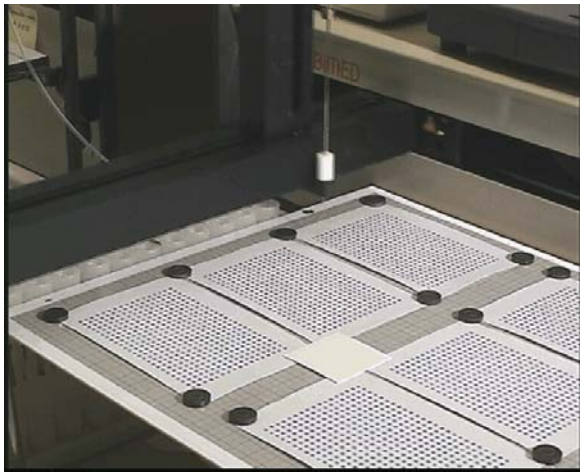

Fig 2: *Parallel miniaturized chemical synthesis of compound libraries on membrane supports by the SPOT method.*

Project Status

The general objective of this subproject of the Chemical Proteomics approach for "SMP Protein" is the implementation and application of chemical micro-array technology for rapid systematic identification of small molecule ligands for protein targets.

Work package 1: Compound library synthesis

A general synthetic strategy will be applied which allows us to use only one standard operation procedure (SOP) per assembly of a library scaffold which then can be flexibly varied to generate compounds suitable either for chip production or as soluble tagged molecules for screening with protein arrays/filters, cell based assays or target competition experiments. Post screening modification of a compound to achieve these options can be quite problematic because this could interfere with the compounds activity. This requires, however, to incorporate a suitable orthogonal reactive functionality into the compounds structure which allows to tether them to a support surface, a label or an affinity tag. In view of the application of compound libraries in these chemical genomics/proteomics studies, this feature is becoming more and more important because it is essential in post screening methods for target fishing. We will start using the SOPs for those library scaffolds established in NGFN-1 and other projects. These include currently small peptidomimetics, cyclopeptides, diketopiperazines, hydantoines, triazines and benzimidazoles. These already could provide a total of several million potential structures. Our compound collection will be extended continuously during

the project with novel SOPs.

Work package 2: Micro-array production
Compound libraries are prepared by solid phase synthesis on a special type of cellulose membrane following our combinatorial synthesis technologies like the SPOT synthesis [4] or the Cut&Combine synthesis [5]. These will produce compounds anchored to individual segments of the cellulose support and may be tested *in situ* for protein binding (Macro-array).The cellulose segments can be segregated into distinct wells of a microtiter plate and dissolved by treatment with e.g. a trifluoroacetic acid cocktail to yield solutions of cellulose-compound conjugates. Further handling allows for the precipitation and cleansing of the cellulose-conjugates which are finally dissolved in DMSO as stock solutions. From a dilution of these DMSO stocks, small aliquots are transferred to the surface of special glass carriers such as microscope slides either in a mini-format or a micro-format (Fig.3). We can print large numbers of copies from the initial library preparations. No high-cost slide materials are required. Therefore, we will be in a position to distribute the generic library chips also to partners of other

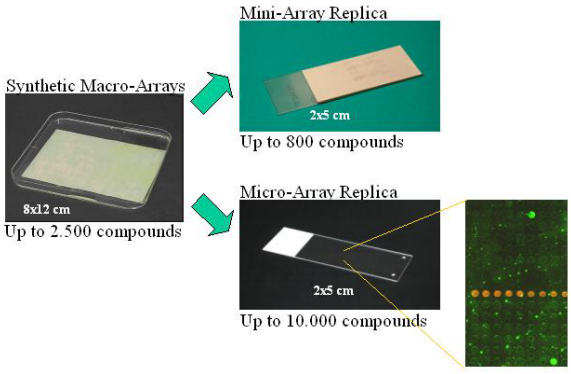

Fig 3: *Transformation of synthetic cellulose-bound compound macro-arrays to cellulose-bound mini-and micro-arrays printed onto microscope glass slides*

Work package 3: Protein screening with chemical micro-arrays
We generate compound micro-arrays from the libraries prepared in WP1 to be used in screening for selective ligands to the set of proteins from the SMP Index. One array of currently up to 10.000 features can be probed with either a single protein acceptor or with a pool of proteins. The latter then requires an iterative deconvolution of the active component in the pool. Protein preparations need to be provided with a label such as fluorescent dye, immuno-tag plus fluorescently labelled antibody, etc.

Figure 4 shows an example of such a micro-array experiment. 384 diketopiperazines coupled to solubilized cellulose were printed together with W7-cellulose and biotin-cellulose as positive controls and incubated with a mixture of calmodulin-Alexa-488(green) and streptavidin-Cy5(red). Biotin is recognized by streptavidin under both conditions whereas calmodulin weakly binds to W7 and to two of the diketopiperazines only in the presence of calcium ions.

Work package 4: Compound screening with protein arrays
As detailed above (WP1), we will prepare the same compound libraries as printed on micro-arrays in work package 2 also in a soluble tagged (e.g. biotin) or labeled (e.g. Cy3/5) form to be used for probing protein micro-arrays and membrane retardation screens in collaboration with the other projects of this SMP as well as for subsequent cell based assays and other applications that require soluble compounds. The prefered tag is biotin as this is chemically robust, economic and can be tetramerized through binding to streptavidin. Multivalent binding will give an avidity advantage for detection of weak interactions.

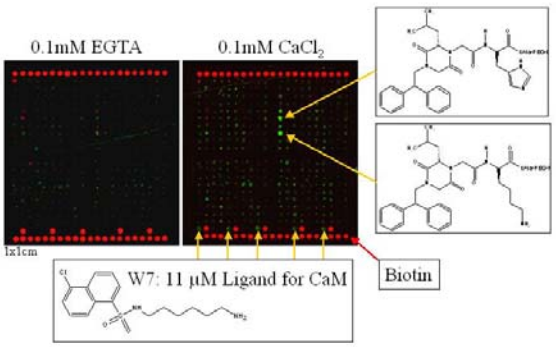

Fig 4: *Discovery of novel calmodulin ligands by probing a diketopiperazine-library array.*

For library synthesis we will utilize the same solid phase combinatorial synthesis technology but implementing a cleavable anchor which allows to release the compounds into solution after having separated the support segments. We will implement a concept of safety-catch linkers which will allow an integrated solid phase synthesis and purification scheme and a final release of the compounds directly into physiological buffers or cell cultures [6]. This will significantly simplify the logistics of syntheses and assays.

Outlook

Proteins are the primary targets for “chemical interference” in disease by drug therapy. Therefore, the identification of chemical compounds that bind with high affinity to disease proteins is paramount and most challenging. In view of the major focus of this SMP “Protein” for NGFN-2 which is on the systematic analysis of human and mouse proteins involved in disease processes, signaling cascades and gene regulation, we will combine the proteomics and chemical genomics activities in one unified platform. The goal of this integrated chemical proteomics platform project is to rapidly provide selective small molecule inhibitors/activators of protein function and protein-protein interactions (PPIs) for as many entries in the SMP Target Index as possible. These will be made available to the whole NGFN to support e.g. functional genome analyses in disease focused studies with e.g. cellular or animal models at the KGs, SMPs and EPs. There our compounds will reach their true value as selective tools to interfere with gene function. The hit compounds will also be offered to industrial partners for their drug development programs.

Chemical genetics studies focussing on disease relevant proteins and phenotypes will lead directly to new drug targets that can be further developed into therapies. Particularly for pharmaceutical drug development, the chemical genomics approach is the most relevant concerning the utilization of genomic information and resources. This very promising approach and its tools must be available for all NGFN partners.

*Lit.: **1.** Zheng XF et al. Chemical genomics in the global study of protein functions. Drug Discovery Today. 2002;7: 197-205. **2.** Frank R. BioSpektrum, Sonderausgabe “Proteomics & Drug Development”. 2002;474-7. **3.** Kuruvilla FG et al. Dissecting glucose signalling with diversity-oriented synthesis and small-molecule micro-arrays. Nature. 2002;416:653-6. 4. Frank R. Spot-Synthesis: An easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. Tetrahedron. 1992;48:9217-9232. **5.** Dittrich F et al. “Cut and Combine”: An easy membrane-supported combinatorial synthesis technique. Bioorg Med Chem Lett. 1998;8:2351-6. **6.** Hoffmann S et al. A new safety-catch peptide-resin linkage for the direct release of peptides into aqueous buffers. Tetrahedron Leters.1994;35:7763-6.*