Proteomics 2003, *3,* 2135–2141 DOI 10.1002/pmic.200300597

Xiaolian Gao[1]
Xiaochuan Zhou[2]
Erdogan Gulari[3]

[1]Chemistry Department, University of Houston, Houston, TX, USA
[2]Atactic Technologies & Xeotron, Houston, TX, USA
[3]Chemical Engineering Department, University of Michigan, Ann Arbor, MI, USA

Light directed massively parallel on-chip synthesis of peptide arrays with t-Boc chemistry

Peptide and peptidomimetic molecule arrays are emerging powerful tools for parallel screening of binding in proteomics and pharmaceutical discovery research. Up to now the common method of preparing peptide arrays was based on spotting on glass using a library of presynthesized peptides. However, due to the large number of monomers (amino acids) it is not possible to have combinatorial libraries which include all combinations of natural and synthetic amino acids. We describe a very flexible on-chip oligopeptide synthesis method which uses the well developed t-Boc based solid state synthesis chemistry. A very high degree of flexibility is achieved by using light photo generated acids and maskless projection lithography for spatially directed deprotection. Use of microfluidic chips enables moderately high densities, short reaction times and off-the-shelf chemicals. Examples are given from synthesis of metal ion binding peptides and epitope binding assys.

Keywords: Microfluidics / On-chip synthesis / Peptide / Photo generated acid / t-Boc PRO 0597

1 Introduction

Microarrays of peptides and peptidomimetics, which make it possible to determine thousands of binding events in a massively parallel fashion, are upcoming and most powerful tools of proteomics and pharmaceutical research. Initially, microarrays were developed for high-throughput gene expression analysis, but their value in proteomics is now being realized [1–5]. Large scale synthesis of peptides was made possible by the development of solid phase synthesis [6]. Peptide or peptidomimetic microarrays are prepared by immobilizing libraries of peptides on a solid support by either high-throughput spotting or light-directed *in situ* synthesis. Given the fact that there are 20 natural and several synthetic amino acids, it is obvious that complete combinatorial libraries for even peptides of five or six amino acids are not possible. However, one advantage of peptide libraries is that peptides stored in the library can be purified after synthesis; *in situ* synthesis eliminates the need to maintain a separate peptide library.

Correspondence: Dr. Erdogan Gulari, Chemical Engineering Department, University of Michigan, 2300 Hayward, 3074 Dow, Ann Arbor, MI 48109, USA
E-mail: gulari@umich.edu
Fax: +1-734-763-0459

Abbreviations: PGA, photo generated acid; **PGAP**, photo generated acid precursor

Until recently, light directed *in situ* synthesis of peptide libraries was based on specially synthesized amino acids with photolabile protection groups [7, 8] using photolithography. Although it is promising for high-density microarrays, this process has practical limitations. Photomasks are expensive (as many as 20 photomasks can be required for the addition of one amino acid), and their fabrication requires a semiconductor grade clean room. Besides, the preparation of the amino acid monomers containing photolabile protection groups is difficult and has not been demonstrated for all the 20 natural amino acids. As a result, the overall process is not cost effective. Every time a change needs to be made, a new set of masks has to be prepared. Furthermore, the chemistry used is not as efficient and clean as the conventional solid phase synthesis chemistry. Recently, we have made several innovations that make photolithographic peptide synthesis a standard laboratory process [9–12]. The remainder of this article will describe the process in detail and give examples of peptide arrays we have synthesized.

2 Materials and methods

Our method for parallel synthesis of peptides relies on four innovations: (i) use of photo generated acid precursors (PGAPs) for removal of acid labile protection groups; (ii) conventional t-Boc protected amino acid monomers; (iii) digital projection lithography; and (iv) microfluidic

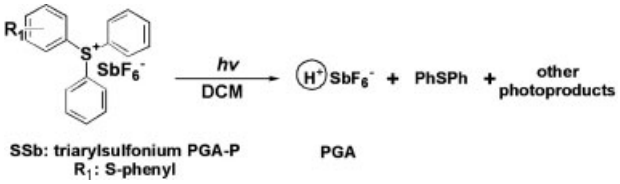

Figure 1. Photo acid generation from SSb (triaryl sulfonium sulfide salt of antimony hexafluoride) PGAP, in solution after light irradiation. This is the gating step in the parallel synthesis process.

chips. In order to synthesize on-chip in a massively parallel fashion, a mechanism to control the location of the reaction on the chip is necessary. In our method, this control is accomplished by introducing photo generated acid [13] (Fig. 1) to replace the conventionally used TFA or TCA for the deprotection reaction, which activates the amino functional group. Since the absorption maxima of most of the photoacid precursors are in the UV region, it is also necessary to use an activator (or energy transfer agent) such as perylene [14] if the wavelength of the light source is longer than 365 nm.

The only deviation from conventional t-Boc solid phase synthesis is that our scheme uses photo generated acid (PGA) in the deprotection step instead of TFA [6].

Being a key component of our parallel synthesis, digital photolithography allows programmable light activation of chemical reactions [15, 16]. A text or Excel file containing the desired amino acid sequences as a function of position on the microarray is loaded into the projection control computer, which then generates the necessary light patterns for synthesis. The generated light patterns are projected onto the chip using a regular digital micro-mirror projector (TI-Instruments, Dallas, TX, USA) coupled with a high-pressure mercury or mercury-xenon light source. The lower wavelength limit of micro-mirror light projectors is 400 nm. Thus, if a regular micro-mirror projector is used, it is necessary to use PGAP compounds that are sensitive above 400 nm. Alternatively, one can use digital projectors specifically made to work in the UV region. During the deprotection step of each synthesis cycle, the light is projected onto the chip in a predetermined pattern to cause the formation of acid in the illuminated spots. A schematic diagram of the digital photolithography and the complete synthesis system is shown in Fig. 2. For photolithography, we have also used a violet diode laser emitting at 405 nm successfully with a commercially available laser light writing system [11].

Since t-Boc based synthesis chemistry is carried out with liquid reagents, it presents a serious challenge for miniaturization and parallel synthesis; even if one can replace

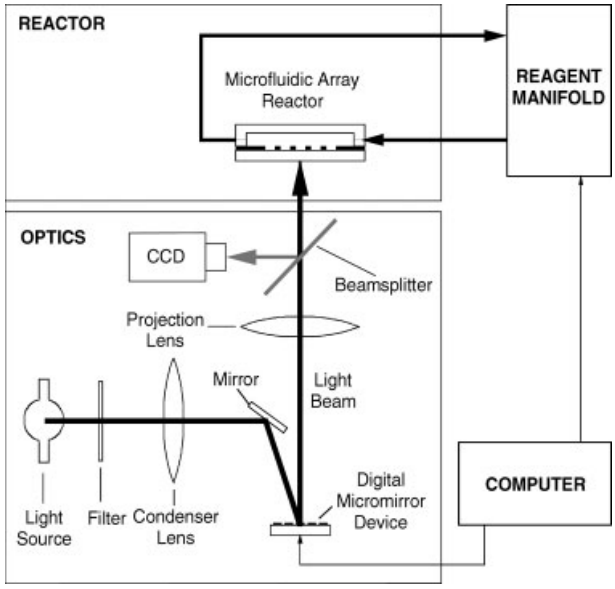

Figure 2. A schematic diagram of the programmable synthesis reactor.

the TFA with PGA. The two most important problems are: isolation of one spot from the other to prevent diffusion of the PGA to unwanted areas, and delivery of equal amounts of reagents over the whole microarray. Typical diffusion coefficients of the chemicals and PGA used are $\sim 10^{-6}$ cm^2; thus, over the time scales involved for the deprotection reaction ($\sim$1 min) they can diffuse by more than 100 microns. For this purpose, we developed microfluidic chip reactors. In its simplest form, the microfluidic chip uses hydrophobic isolation of synthesis spots from each other [10, 11]. With hydrophobic isolation, the surface of the whole chip is flooded with one reagent at a time and the excess reagent is drained by gravity, leaving uniform droplets of reagents behind as shown in Fig. 3. While this approach is simple and elegant, it has some limitations. Despite the fact that the vapor phase is saturated, the isolated liquid droplets have a relatively short life time of $\sim$1–2 min and hydrophobic patterning loses its hydrophobicity after five to six amino acid addition cycles. A better solution is to use a sealed microfluidic reactor with physical isolation of reaction chambers from each other. Figure 4 shows a schematic diagram of the microfluidic reactor.

As mentioned above, our light directed synthesis of peptides on-chip follows the standard t-Boc chemistry; which consists of: (i) derivatizing the surface with a protected linker such as aminopropyltriethoxysilane; (ii) filling the reactor with a solution of the photo generated acid precursor in dichloromethane; (iii) light-directed acid generation at the desired reaction chambers to deprotect the linker; (iv) flushing and neutralizing the generated acid;

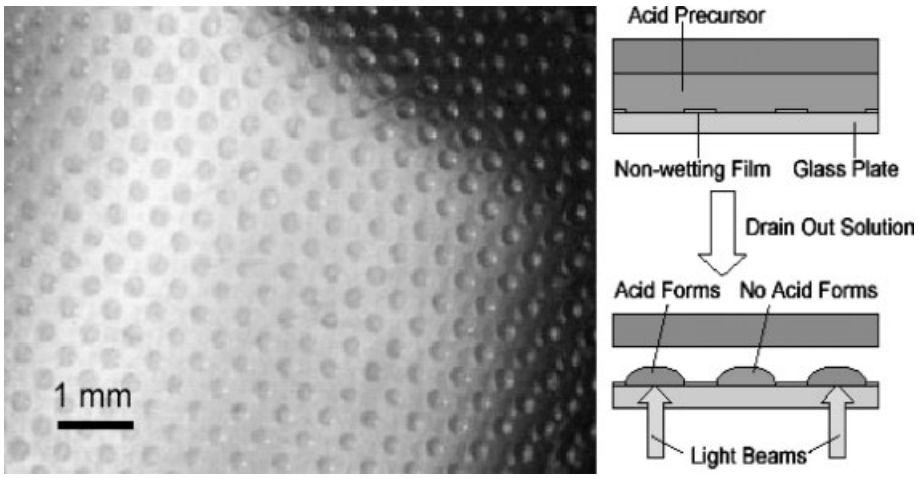

Figure 3. Micro droplet array, formed by the use of hydrophobic patterning, and explanation of the liquid droplet reaction chamber isolation during acid generation.

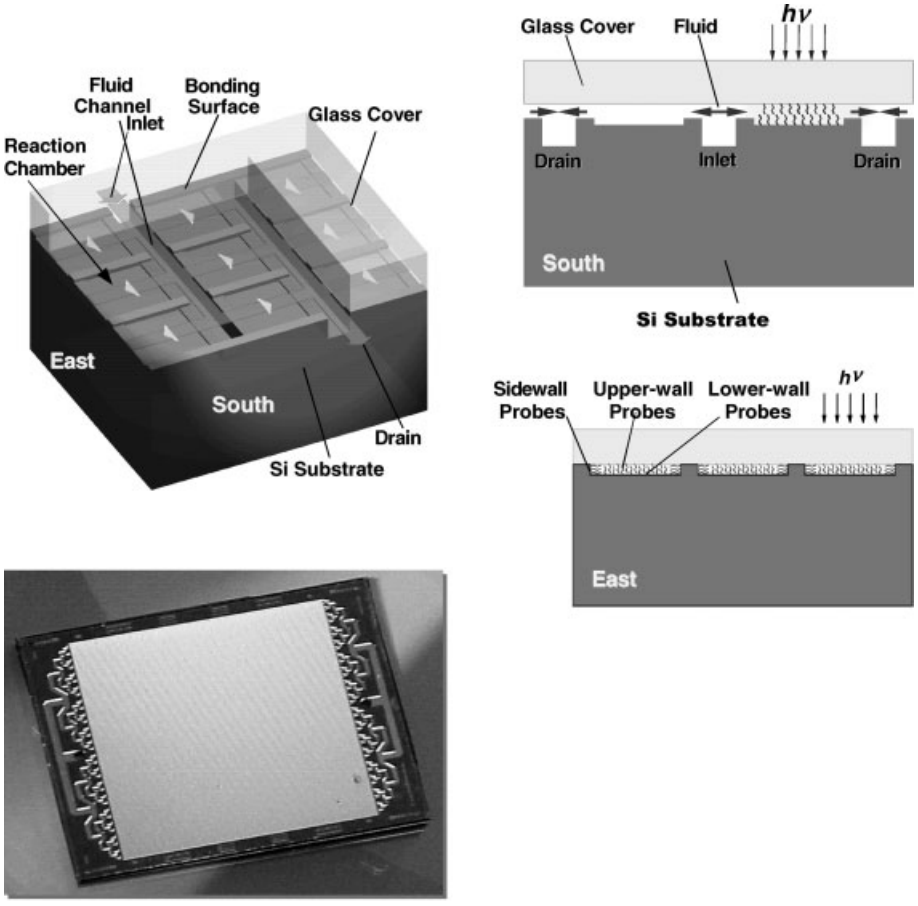

Figure 4. Schematic diagram of how the microfluidic chip works (top panels) and picture of an actual microfluidic chip with ~30 000 reaction chambers (bottom panel). The structures seen at both ends of the reactor are reagent delivery and drain manifolds. Approximate size 1.4 × 1.7 cm.

and (v) coupling of the first amino acid. Capping of unreacted linkers and side chain deprotection are then carried out according to the standard peptide synthesis procedure [17]. This cycle is repeated for each step by the number of amino acids to be incorporated in that step. Figure 5 schematically shows how different peptides are grown on the surface, one amino acid at a time. The cycle times associated with the on-chip synthesis are actually shorter than the cycle times observed with resin-based synthesis. Typically, deprotection reactions

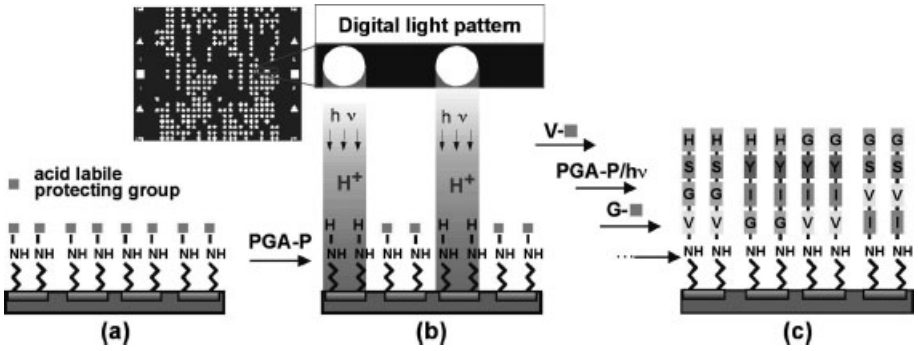

Figure 5. Sequential synthesis of peptides on the chip surface by the addition of amino acids.

can be carried out in ~2–5 min and the coupling reaction in ~10 min, which makes combinatorial synthesis of peptide microarrays of 10–20 amino acids feasible.

3 Results

3.1 Metal binding peptide assays [11]

To test the feasibility of a peptide based toxic metal sensor in drinking water, four oligopeptides dns-Glu-Cys-Glu-Glu, H_2N-Glu-Glu-Glu-Glu, dns-Cys-Cys-Cys-Cys and dns-Gly-Gly-Gly-Gly (dns-dansyl chloride, were incorporated into the sequence as a chromophore for detection purposes) were simultaneously synthesized using the hydrophobic isolation method and a 405 nm laser with a laser writer. The first tetra peptide is known to have specific Pb(II) ion binding capability as detected by an increase in the fluorescence intensity of dns after binding [18]. The location of the four peptides in the microarray are shown in Fig. 6A and the fluorescence image of the chip is shown in Fig. 6B after exposure to 1 mM $PbCl_2$ solution in buffer. As seen clearly in Fig. 6B, the synthesized sequence on the surface binds Pb(II). We have since synthesized other peptide arrays to detect a number of heavy metals and aromatic compounds in solution.

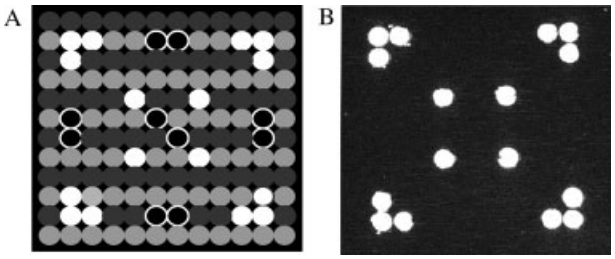

Figure 6. Microarray of tetrapeptides for Pb(II) binding: (a) pattern of localized parallel synthesis of dansyl-labeled tetrapeptides, dns-Glu-Cys-Glu-Glu (○, white dot), dns-Glu-Glu-Glu-Glu (●, black dot), dns-Cys-Cys-Cys-Cys (●, light gray) and dns-Gly-Gly-Gly-Gly (●, dark gray); (b) fluorescence image of dansyl-labeled tetrapeptides simultaneously synthesized on microarray after exposure to 1 mM $PbCl_2$.

3.2 Epitope binding assays [10]

One major application of a peptide library is to use it in immunoassays [19]. Figure 7 displays results of FITC-PAb240 binding to a peptide/peptidomimetic microchip synthesized by incorporation of 28 different monomers in 45 synthesis cycles and containing 2304 sequences, in which there are 281 unique sequences, and each sequence is at least a triplicate. These sequences were designed for mutational analysis of the RHSVV epitope sequence, and positional scanning of p53 in the region covering RHSVV. These mutation analyses confirm the amino acid analogs 2-aminobutyric acid, norvaline and norleucine (Abu, Nva and Nle, compounds f–h, Fig. 7A, B) can be substituted for naturally occurring residues. The analysis of these results provides new insight into the sequence for PAb240 epitope binding: (a) serine and valine can only be L-stereoisomers (Fig. 7A, rows 1–3, column labeled with a/b); (b) arginine cannot be replaced by glycine and glutamine; (c) serine is highly selective and cannot be replaced by any of the amino acids used (Fig. 7A, row 1); (d) the valine binding site is highly hydrophobic as shown by its affinity for several residues containing C2 to C4 alkyl side chains (Fig. 7A, rows 2–6, X = I, L, f (Abu), g (Nva), and h (Nle)), while sarcosine, β-alanine, alanine and glycine show no binding (Fig. 7A, rows 2–6); (e) the valine binding site exhibits steric specificity. The Abu residue is a valine analog but missing one of the γ-methyls (side chain = ethyl), and it retains ~70% binding affinity compared to valine in either positions. Nva and Nle, which contain linear propyl and butyl side chains, exhibit reduced binding affinity. The longer the side chain, the greater the reduction in binding. A cyclic alkyl side chain (cyclohexyl, Cha) almost completely suppresses the binding. A comparison of RHS<u>X</u>V and RHSV<u>X</u> (X = Abu, Nva or Nle) suggests that the alkyl side chain effect is more profound when X is next to the serine, away from the end of the epitope sequence. Another finding of this study was the strength of binding of various positional sequences. The same study [10] also looked at the importance of the position of the binding motif.

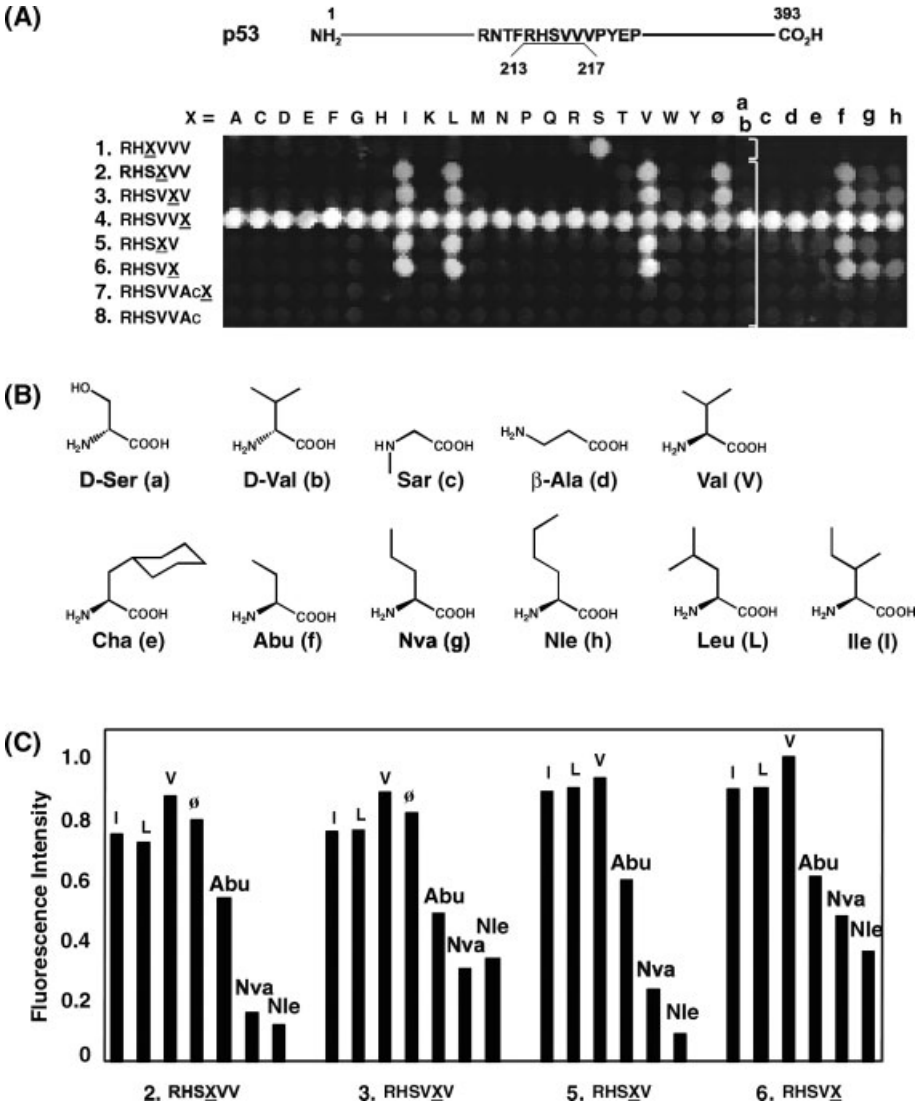

Figure 7. PAb240 binding to the peptide and peptidomimetic microchip for synthesis validation and antibody epitope analog discovery. (A) Position of PAb240 epitope within the human p53 sequence and the fluorescent image of the sequences containing 20 natural and eight unnatural amino acids, which are indicated by the labels at the top of the figure, following the order of 20 amino acids (letters A–Y), ø (skipping the reaction at this position), and a–h (the eight synthetic monomers). The specific positions of these incorporated residues are indicated by a symbol X (rows 1–8 on left). The sequences in the column labeled with a and b contain D-serine (a) and D-valine (b) with D-serine substituting for L-serine in row 1 and D-valine in the X position in the rest of the sequences. (B) Chemical structures of the eight unnatural amino acids as well as valine, isoleucine, and leucine, drawn for comparison purposes. (C) Fluorescent intensity plot of rows 2, 3, 5 and 6 from the image shown in (A).

These results are unambiguously confirmed by comparison to those of phage display [20, 21] identifying the known RHSVV epitope sequence and leading to the discovery of new peptidomimetic antibody recognition elements.

3.3 Stepwise yield

One of the fundamental issues in any synthesis platform is stepwise yield. Yield is an especially critical issue when one is trying to grow oligomers of reasonable length (>10), a stepwise yield of 90% gives only 35% true length decamers. To obtain a quantitative measure of the stepwise yield, the terminal amine groups in the growing amino acid oligomers were either labeled with 4(5)-carboxy fluorescin diisobutyrate in selected positions during the stepwise addition of amino acids or capped with acetic anhydride. The chip was illuminated with the same light source used to generate PGAs and the resulting intensity was measured with a cooled CCD camera (Apogee, Auburn, CA, USA). The results are shown in Fig. 8. Comparison of the relative fluorescence intensity of the growing oligomers gives a stepwise yield of 98+%.

4 Discussion

The widespread acceleration in the use of DNA microarrays and gene chips for genomic analysis [22] demonstrates the usefulness of miniaturized spatially addressa-

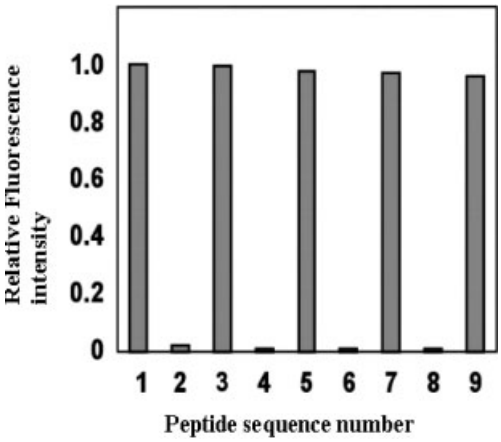

Figure 8. Plot of relative fluorescence intensities for 20 sets of redundant data points for sequences 1–9 given below. FR, fluorescin; Ac, Acetyl. (1)Ahx-βAla-βAla-NH-FR; (2) Ahx-βAla-βAla-Ac; (3) Ahx-βAla-βAla-Val-NH-FR; (4) Ahx-βAla-βAla-Val-Ac; (5) Ahx-βAla-βAla-Val-Val-NH-FR; (6) Ahx-βAla- βAla-Val-Val-Ac; (7) Ahx-βAla-βAla-Val-Val-Ser(Bzl)-NH-FR; (8) Ahx-βAla-βAla-Val-Val-Ser(Bzl)-Ac; (9) Ahx-βAla-βAla-Val-Val-Ser(Bzl)-His (Bom)-NH-FR.

ble microarrays for high-throughput genomic analysis. Despite the fact that peptides and peptide analogs are among the most tested drug candidates and biological intermediates, the choices and prices available in DNA microarrays are not yet available for peptides. Therefore, it will be highly desirable to have access to these potentially powerful tools, especially those which are flexible in sequence design to allow the incorporation of a broad range of building blocks. Peptide microarrays possessing these features can complement random peptide phage libraries [23–26] which can generate 10^6–10^{15} peptides for functional assays, but are routinely limited to natural amino acids and are commonly biased in sequence expression space.

In this article, we have described a simple and highly efficient method for parallel synthesis of peptides and peptide analogs on silicon-glass microchips of 1–2 cm^2 size. The approach demonstrated herein enables massively parallel measurements of molecular complex formation/dissociation under identical conditions, and thus is particularly well suited for high-throughput screening assays that rely on relative quantitation measurements. The amino acids and other chemicals used are all available off-the-shelf without needing any special synthesis. While the method described here is based on t-Boc chemistry for peptides, the general mechanism of gating the deprotection reaction through photo generated chemicals is broadly applicable to synthesis of DNA, RNA, peptide and carbohydrate or other small molecule microarrays. This versatility is difficult to achieve with monomers protected by photolabile groups [27]. Another advantage of using photo generated acids is the non-linear nature of the deprotection reaction. The reaction rate increases very rapidly above a certain threshold of acid concentration and becomes acid concentration independent quickly [10]. This feature makes it possible to use short light exposure times and eliminates cross-talk due to stray light initiated reactions in neighboring reaction sites [28] and gives higher fidelity and yield. In comparison, deprotection by cleaving off the photolabile groups is linear in light intensity and exposure.

The use of microfluidic chips has several benefits; the chips are sealed and can not be easily contaminated like the spotted microarrays; the small dimensions of the flow channels as well as reaction sites increase the mass transfer rates which lead to faster reaction times. Both the deprotection time as well as the coupling time for peptide synthesis on the microchip is at least a factor of two to three times shorter than the resin based synthesis. In terms of the array densities, microfluidic chips used in our investigations are intermediate between spotted arrays and high-density arrays of Affymetrix. We have made arrays with densities of 144–30 000 chambers/cm^2. The density can probably be pushed up by another factor of two; but at that density, extreme cleanliness is required with respect to dust because the dimensions of the flow channels are of the order of 5 microns, which can get plugged easily by suspended dust leading to problems of uniformity.

5 Concluding remarks

At the present, the cost of producing peptide arrays by the method described in this paper is relatively high (compared to DNA chips) due to the longer times required by the peptide synthesis reactions; however, we believe that it is feasible to decrease the synthesis time of a 20 mer peptide to less than eight hours. This, coupled with high-throughput synthesis, has the potential to bring the cost of the peptide chips to the level of DNA chips and make them economically viable for everyday use in research.

This research was supported by grants from NIH/NCI, NSF, Texas ATP, Michigan Life Sciences Corridor Fund, the National Foundation for Cancer Research, and the R. A. Welch Foundation.

Received May 19, 2003

6 References

[1] MacBeath, G., Schreiber, S. L., *Science* 2000, *289*, 1760–1763.

[2] Mezzasoma, L., Bacarese-Hamilton, T., Di Cristina, M., Rossi, R. *et al.*, *Clin. Chem. 2000, 48, 121–130.*

[3] Zhu, H., Klemic, J. F., Chang, S., Bertone, P. *et al.*, *Nat. Genet.* 2000, *26*, 282–289.

[4] Zhu, H., Bilgen, M., Bangham, R., Hall, D. *et al.*, *Science* 2001, *293*, 2101–2105.

[5] De Wildt, R., Mundy, C., Gorick, B., Tomlinson, I., *Nat. Biotechnol.* 2000, *18*, 989–994.

[6] Merrifield, R. B., *Science* 1986, *232*, 341–347.

[7] Jung, G. (Ed.), *Combinatorial Peptide and Nonpeptide Libraries. A Handbook*, VCH, New York 1996.

[8] Fodor, S. P. A., Leighton, J., Pirrung, M. C., Stryer, L. *et al.*, *Science* 1991, *251*, 767–773.

[9] Gao, X., Zhou, X., Gulari, E., Method and apparatus for chemical and biochemical reactions using photo-generated reagents, US Patent 6.426, 184.

[10] Pellois, J. P., Zhou, X., Srivannavit, O., Zhou, T. *et al.*, *Nat. Biotechnol.* 2002, *20*, 922–926.

[11] Komolpis, K., Srivannavit, O., Gulari, E., *Biotechnol. Prog.* 2002, *18*, 641–646.

[12] Pellois, J. P., Wang, W., Gao, X., *J. Comb. Chem.* 2000, *2*, 355–360.

[13] MacDonald, S. A., Wilson, C. G., Frechet, J. M., *Acc. Chem. Res.* 1994, *27*, 151–157.

[14] Crivello, J. V., Lam, J. H. W., *J. Polymer Sci.* 1979, *17*, 1059–1065.

[15] Gao, X., Le Proust, E., Zhang, H., Srivannavit, O. *et al.*, *Nucleic Acids Res.* 2001, *29*, 4744–4750.

[16] Singh Gasson, S., Green, R. D., Yue, Y., Nelson, C. *et al.*, *Nat. Biotechnol.* 1999, *17*, 974–978.

[17] Stewart, J. M., Young, J. D. (Eds.), *Solid Phase Peptide Synthesis*, Pierce Chemical Co., Rockford, IL 1984.

[18] Deo, S., Godwin, H. A., *J. Am. Chem. Soc.* 2000, *122*, 174–175.

[19] Fodor, S. P. A., Stryer, L., Winker, J. L., Holmes, C. P. *et al.*, Affymetrix, United States Patent 6, 124, 102, 2000.

[20] Stephen, C. W., Lane, D. P., *J. Mol. Biol.* 1992, *225*, 577–583.

[21] Stephen, C. W., Helminen, P., Lane, D. P., *J. Mol. Biol.* 1995, *248*, 58–78.

[22] The chip challenge, *Nat. Genet.* 1999, *21*, 61–62.

[23] Devlin, J. J., Panganiban, L. C., Devlin, P. E., *Science* 1999, *249*, 404–406.

[24] Scott, J. K., Smith, G. P., *Science* 1999, *249*, 386–390.

[25] Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R., Cesareni, G., *J. Mol. Biol.* 1999, *222*, 301–310.

[26] Cwirla, S. E., Peters, E. A., Barrett, R. W., Dower, W. J., *Proc. Natl. Acad. Sci. USA* 1990, *87*, 6378–6382.

[27] Bochet, C., *Tetrahedron Lett.* 2000, *41*, 6341–6346.

[28] Garland, P. B., Serafenowski, P. J., *Nucleic Acids Res.* 2002, *30*, e99.