

Kinetics of Proteolytic Reactions in Nanoporous Materials

Hongyan Bi,[†] Liang Qiao,[‡] Jean-Marc Busnel,[†] Baohong Liu,^{‡,*} and Hubert H. Girault^{†,*}

Laboratoire d'Electrochimie Physique et Analytique, Ecole Polytechnique Fédérale de Lausanne, Station 6, CH-1015 Lausanne, Switzerland, and Department of Chemistry, Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, P.R. China

Received May 5, 2009

Proteolysis with proteases preloaded within nanopores of porous material is a very fast process, where proteins can be digested in minutes compared to the conventional bulk enzyme reactions taking place over hours. To model this surprising phenomenon, a modified sequential proteolytic mechanism has been developed to simulate the kinetics of the reaction. Digestion of myoglobin was used as an example to show the high efficiency of the in-nanopore enzymatic reaction, while angiotensin 1 and ACTH (1–14) were selected as model peptides to validate the theoretical considerations. The proteolytic peptides were quantified by capillary electrophoresis and sequenced by mass spectrometry using bottom-up strategy. The simulation clearly shows that the major factor for the very fast digestion kinetics observed stems from a peptide confinement effect, where the generated peptides are trapped within a confined space for further proteolysis to the final products. On the other hand, the ingress and diffusion of the proteins into the porous cavity can accelerate or limit the first proteolytic step requiring the encounter between the substrates and enzymes. The present model can be widely applied to different enzyme catalyzed reactions for high-throughput protein profiling, and can promote the study of enzyme reactions occurring inside the cell.

Keywords: proteolysis • nanoporous material • nanoconfinement effect • kinetics simulation

Introduction

Over the past few years, proteomics has emerged as a rapid developing and open-ended endeavor, benefiting from the tremendous recent progress in biotechnology and success in genomics.¹ To understand the protein primary sequence,² post-translational modifications² as well as protein–protein interaction,³ a series of technical tools have been developed. One powerful strategy for the analysis of complex protein samples is mass spectrometry (MS) based bottom-up proteomics,¹ combining the techniques of sequence-specific enzymatic cleavage of proteins, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) or high performance liquid chromatography (HPLC) separation, MS or MS/MS analysis, and database search.

Nowadays, mass spectrometry including electrospray ionization mass spectrometry (ESI-MS) and matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is generally carried out with a rather high sensitivity. The detection limit of these techniques can easily reach low femtomole amounts of peptides.⁴ One critical step that still limits the overall sensitivity of the MS-based proteomics approach is proteolysis.⁴ In-solution digestion is an inherent slow process and proteins at submicromolar concentrations are very difficult

to be digested, simply due to the mass transport considerations associated to a two-body encounter probability.^{4,5} In extracted biological samples, many proteins are naturally expressed at a low abundance. Therefore, to realize global analysis of proteins, the reaction time may be prolonged or the enzyme amount may be increased. However, the presence of a high enzyme-to-protein ratio inevitably induces a strong enzyme autolysis which can result in the suppression of MS signals of the analyte.^{4–7} To address these issues, many enzyme immobilization based methods have been developed to realize highly efficient proteolysis by immobilizing the proteases on matrices, such as capillaries and microchips, and then by flowing the protein solutions over them.^{8–14} Additionally, the immobilized enzyme systems can also be used as a model to study membrane-bound enzymes reactions occurring in the cell.

Among various matrices, nanoporous silica is environmentally acceptable, structurally stable, and resistant to microbial attacks, and has therefore drawn much interest for its biological applications.^{15–23} Indeed, such porous materials can provide supports for biochemical reactions by adsorbing reactants into the inner-pores and confining the following reactions in nanoscale domains.^{24,25} The specific surface area of these materials is large and the pore size can be tuned from few to hundreds nanometers, which can then easily host trypsin, a small globular protein (23–24 kDa) with a diameter of about 4 nm, and also larger enzyme molecules.¹⁵ Recently, we have designed a digestion system confined in the nanochannels of a porous silica where the reaction efficiency was greatly increased.^{26,27} The gist of this method is to entrap the enzyme

* To whom correspondence should be addressed. H.H.G.: tel, +41 21 6933145; fax, +41 21 6933667; e-mail, hubert.girault@epfl.ch. B.L.: tel, +86-21-65642405; fax, +86-21-65641740; e-mail, bhliu@fudan.edu.cn.

[†] Ecole Polytechnique Fédérale de Lausanne.

[‡] Fudan University.

in a nanopore reactor where its local concentration is *de facto* very high. As a result, the in-nanopore reaction kinetics always operates in an enzyme excess mode, that is, at its maximum rate and is pseudo-first-order with respect to the proteins or peptides.

Herein, the mechanistic aspects of proteolysis reactions with enzymes entrapped within nanopores are proposed to simulate the kinetics of the highly efficient process. To realize this purpose, a novel enzymatic nanoreactor was fabricated by loading trypsin in the nanopores of macroporous ordered siliceous foams (MOSF), which has an average pore size of ~ 100 nm that is large enough to entrap any enzymes and allow an easy transport of substrates into active reaction sites.²⁸ By incubating proteins together with a suspension of the enzymatic nanoreactors, the sequential time-consuming protein/peptide–enzyme encounter processes that may be even controlled by diffusion are replaced by a process of protein ingress in the MOSF—in pore diffusion—successive proteolytic reactions—peptide egress. Indeed, protease digestion involves a series of consecutive reactions cleaving the peptide chain at specific well-targeted points, such as arginine and lysine for trypsin. To illustrate this point, the classical sequential scheme,^{29–32} which considers that the enzyme cleaves one protein or peptide at a time and that the generated peptide fragments diffuse away from the substrate, was employed and modified to simulate the proteolysis process. Digestion of myoglobin was employed here as an example to illustrate the high efficiency of the nanoconfined proteolysis. Two small peptides, angiotensin I, with one active digestion site, and ACTH (1–14), with two active sites, were employed to validate the simulation. A capillary electrophoresis (CE) system was used to quantify the products and a coupled MS analysis was used for the identification.

Experimental Section

Maple Calculation. The kinetic model is based on a set of differential equations that can be solved analytically with commercial software (Maple 12, Waterloo Maple, Inc.) on a personal computer (1.67 GHz Mobile DualCore Intel Core 2 Duo T5500, 1024 MB DDR2–667 DDR2 SDRAM, Latitude D630 Dell).

Materials. Myoglobin (from horse heart, 95%) and α -cyano-4-hydroxycinnamic acid (CHCA, 99%) were obtained from Sigma-Aldrich (Steinheim, Germany). Trypsin (M_w , 23 800 Da, from bovine pancreas) was from AppliChem GmbH (Darmstadt, Germany). ACTH(1–14) (SYSMEHFRWGKPVG, M_w , 1680.91) and angiotensin I (DRVYIHPFHL, M_w , 1296.5) were purchased from Bachem (Bubendorf, Switzerland). Ammonium bicarbonate (NH_4HCO_3 , >98%) was obtained from Fluka-Garantie (Buchs, Switzerland). Acetonitrile (ACN, $\geq 99.5\%$) was from Riedel-de Haën GmbH (Seelze, Germany). Acetic acid (HAc, $\geq 99.5\%$) was purchased from Fluka (Buchs, Switzerland) and trifluoroacetic acid (TFA, 99.8%) was from Acrös Organics (Geel, Belgium). Hydroxypropylcellulose (HPC) was from Acrös (Chemie Brunschwig AG, Basel, Switzerland). Macroporous ordered siliceous foams (MOSF)²⁸ were obtained from the Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials (Fudan University, Shanghai, China). All reagents were used as received without further purification. Deionized water (18.2 M Ω cm) was produced by an alpha Q-Millipore System (Zug, Switzerland).

Nanoconfined and In-Solution Proteolysis. The scanning electron microscope Philips XL 30 FEG-SEM was employed for

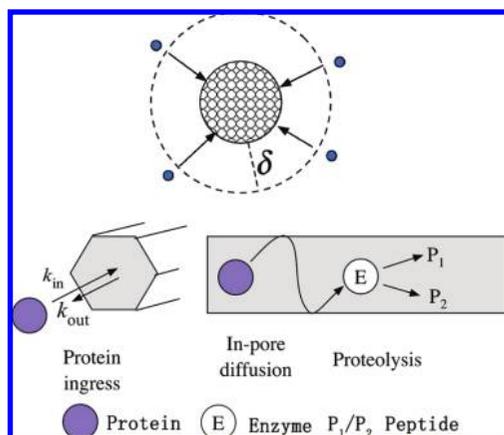
the SEM images of MOSF. The enzymes were previously entrapped in the nanopores of MOSF. In detail, 4 mg/mL MOSF aqueous solution was first stirred at room temperature for 1 h, and then this resultant solution was used to dissolve trypsin at 20 ng/ μL with persistent stirring at room temperature for 1 h. Proteins or peptides (as substrate) were dissolved in 25 mM NH_4HCO_3 (pH 7.8) at 20 ng/ μL . Then, the solutions of substrate and enzyme were mixed by different volume to adjust enzyme-to-substrate ratio at 1:30 (g/g). In case of in-solution digestion, the procedure was kept, including the enzyme-to-substrate ratio, but using free enzymes. Incubation was performed on a Thermo Shaker (TS-100, Biolabo scientific instruments) at 37 °C and 1200 rpm shaking. After different incubation times, 30 μL of sample solution was collected and added with 3 μL of 10% HAc to stop the digestion and then stored at -20 °C.

Capillary Electrophoresis Analysis of the Digested Peptides. To study the kinetics, digestion product separation was performed using a HP 3D Capillary Electrophoresis System equipped with diode array detection and computer control via HP ChemStation software. Fused-silica capillary (50 μm i.d., 375 μm o.d.) was obtained from BGB Analytik AG (Böckten, Switzerland) and coated with HPC following a previously reported procedure³³ to suppress the electro-osmotic flow and limit sample adsorption onto capillary walls. A recently introduced robotized iontophoretic fraction collection system was used to couple CE and MS.³⁴

Mass Spectrometry Methods and Data Analysis. All the digested products were analyzed using a Bruker Microflex MALDI-TOF mass spectrometer. The products were dropped on a commercial MALDI target plate and dried at room temperature. Then, CHCA matrix was dropped as an overlayer and dried under the same conditions. The matrix was 4 mg/mL CHCA in ACN/ H_2O /TFA (50/49.9/ 0.1% (v/v)). Before identification, the MS instrument was calibrated by external calibration using a standard peptide mixture. The instrumental parameters were optimized to obtain good mass spectra. Mass spectrometric data analysis was performed with flex Analysis (Bruker) and Mascot searching engine. The masses and intensities of special peaks were read out using the flex Analysis software, and then sent to Mascot for database searching, where the experimentally measured peptide masses were compared with the theoretical ones for a specific Swiss-Prot entry. The m/z tolerance was set as 0.5, and the minimum signal/noise for reading peaks was set as 6.

Results and Discussion

Enzymatic Proteolysis in Nanoporous Matrices. Different from conventional in-solution digestion, which can be represented as a succession of protein/peptide–enzyme encounter reactions that are diffusion limited, proteolysis confined within a nanopore is a multistep process combining protein ingress, in-pore diffusion, cleavage, peptide in-pore diffusion, further cleavage, and finally peptide egress, as shown in Scheme 1. Enzymes were first entrapped into the porous matrices before being added to the substrate solution. Because the in-pore concentration of enzyme is really high, as soon as one substrate molecule diffuses into the nanopore, it will have a high encounter probability with the entrapped enzymes. Compared with the free enzymatic reaction where the generated peptides can diffuse away from the protease back to the bulk solution, in the case of enzyme-loaded porous matrices, the generated peptides are trapped in the nanopores and their further

Scheme 1. Schematic Representation of the Proteolysis Using MOSF Entrapped Enzyme

encounter probability with the enzyme becomes much higher, and consequently, the longer the peptide, the larger the entrapment effect. Small peptides, however, have a higher escape probability, and may diffuse back to the bulk solution, where they can be collected and analyzed by mass spectrometry.

Here, myoglobin, a protein predicted with 21 cleavage sites for tryptic digestion, was used to illustrate the high efficiency of the present digestion strategy. Macroporous ordered siliceous foams (MOSF) were employed for entrapping enzymes due to its excellent three-dimensional structure of hexagonally²⁸ arrayed columns where the in-pore diffusion is a pseudolinear process. The SEM images of MOSF are shown in Supporting Information SI-1. The peptide mass fingerprinting (PMF) results of in-solution and in-nanoreactor digestion are compared in Figure 1. After incubation at 37 °C for 30 min under stirring, 24 peptides were successfully identified with very high signal-to-noise (S/N) using MALDI-MS from the nanopore-based proteolysis of 20 μg/mL myoglobin (Figure 1a). In contrast, only 7 peptides were identified with a strong noise using the in-solution digestion strategy under the same condition (Figure 1b) and 15 peptides were identified with good S/N using a typical overnight (21 h) in-solution digestion (Figure 1c). The trypsin-to-protein ratios for both in-solution and nanoreactor based proteolysis are all fixed at 1:30 (g/g) after experimental optimization. The detailed information, including sequence and the number of miss-cleavage sites, about these identified peptides was listed in Supporting Information (SI-3). Considering the PMF results and the incubation time, we can estimate that the required reaction time is decreased at least by ~40 times using the entrapped enzyme media.

Sequential Mechanism in Solution. To analyze the fast reaction rates observed when using the enzyme-loaded MOSF, a modified sequential mechanism^{29–32} was employed to model the proteolysis kinetics. In this mechanism, the enzyme combines with a substrate, such as a protein or a large peptide, at its active amino acid, and then cleaves it into two smaller peptides at a time. Different from models, such as the zipper model, which assume that a terminal low molecular weight peptide containing no active group is cleaved away in each reaction,²⁹ here we consider that the reaction can happen randomly at any given reactive target points of the substrate, where both big and small peptides may be generated at a time. This process can be schematically represented as

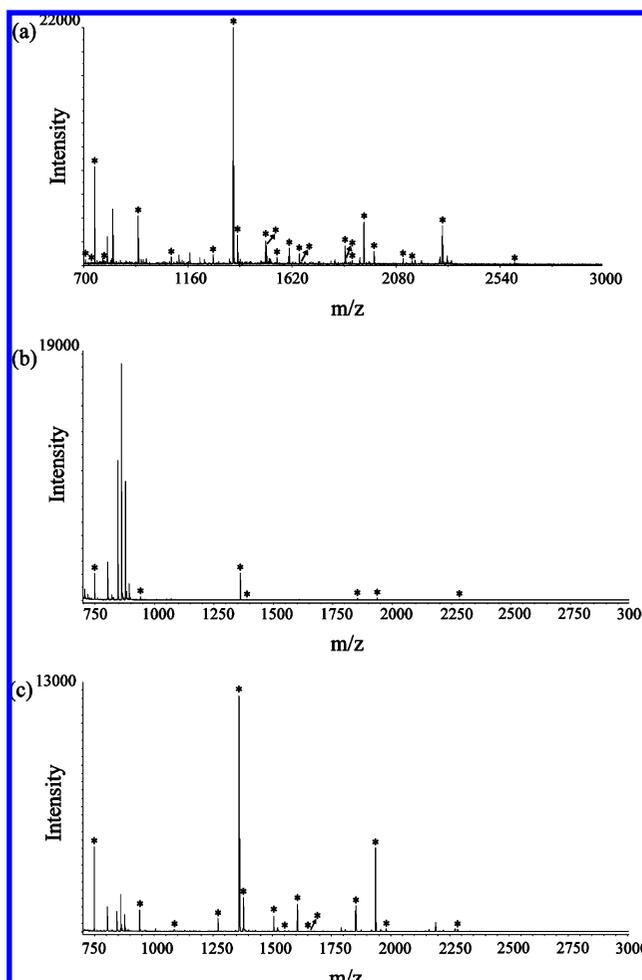
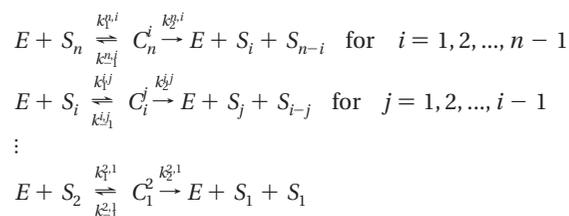


Figure 1. PMF spectra of peptides obtained from in-pore tryptic digestion of myoglobin after 30 min incubation (a), and from a typical in-solution tryptic digestion of myoglobin after 30 min (b) and 21 h (c) incubation. The peaks marked with asterisk (*) represent identified peptides of myoglobin. Each digestion was 20 ng/μL myoglobin (in 25 mM NH₄HCO₃) incubated with 1/30 (g/g) trypsin with or without MOSF (133 ng/μL), at 37 °C and under 1200 rpm shaking.



(1)

The primary substrate protein or peptide including $n-1$ active amino acids, called S_n , combines with an enzyme E to form a complex C_n^i . The complex then generates two substrate fragments S_i and S_{n-i} . S_i can be expressed by a probability distribution, $P_n^i (=C_n^i/\sum C_n^i)$. Subsequently, S_i or S_{n-i} would further react with the enzyme E again to form another complex C_i^j or C_{n-i}^j that in turn is converted into smaller fragments S_j . This process continues until the final fragment S_1 is produced, which cannot be digested any more.

Because of the complex protein structure and various reaction constants, an accurate rate equation of the present sequential mechanism is mathematically complicated. Indeed,

during the proteolysis, many different peptides, all with i active digestion sites can be generated from diverse mother substrates because of the different amino acid sequences. For a simplified analytical calculation, several approximations are made. Here, we suppose that the protein is completely denatured; we neglect the differences between amino acid residues and assume that all trypsin molecules bind with the active sites nearly in the same manner; therefore, the reaction constants k_1 , k_{-1} , and k_2 are very similar for all the reactions. These assumptions are not likely in practice; however, it is necessary to simplify the problem and would not take trouble for the understanding of the basic processes of enzymatic reactions. With these assumptions, it is considered that the enzyme–substrate reaction would randomly happen at any active sites, meaning that the probability distribution factor P^j can be approximately assumed as $2/i - 1$, and all the peptides generated with i active amino acids can be classified into one group named S_i . Therefore, the rate equations for the modified sequential mechanism are

$$\begin{aligned}
 P^j &= \frac{2}{i-1} \\
 \frac{dS_n}{dt} &= k_{-1}C_n - k_1ES_n \\
 \frac{dS_i}{dt} &= k_{-1}C_i - k_1ES_i + \sum_{j=i+1}^n k_2C_j^i = k_{-1}C_i - k_1ES_i + \\
 &\quad \sum_{j=i+1}^n k_2 \frac{2}{j-1} C_j \quad \text{for } i = 2, 3, \dots, n-1 \quad (2) \\
 \frac{dS_1}{dt} &= \sum_{i=2}^n k_2C_i^1 = \sum_{i=2}^n k_2 \frac{2}{i-1} C_i \\
 \frac{dC_i}{dt} &= k_1ES_i - k_{-1}C_i - k_2C_i \quad \text{for } i = 2, 3, \dots, n \\
 \frac{dE}{dt} &= - \sum_{i=2}^n \frac{dC_i}{dt}
 \end{aligned}$$

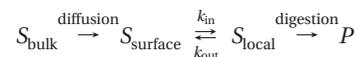
The complexes C_i are intermediates. Thus, the rate eqs 2 can be simplified using the classical steady-state approximation as

$$\begin{aligned}
 \frac{dC_i}{dt} &= k_1ES_i - k_{-1}C_i - k_2C_i = 0 \quad \text{for } i = 2, 3, \dots, n \\
 \frac{dS_n}{dt} &= - \frac{k_1k_2}{k_{-1} + k_2} ES_n = - \frac{k_2}{K_M} ES_n \\
 \frac{dS_i}{dt} &= - \frac{k_1k_2}{k_{-1} + k_2} ES_i + \sum_{j=i+1}^{i+1} \frac{2}{j-1} \frac{k_1k_2}{k_{-1} + k_2} ES_j = \\
 &\quad - \frac{k_2}{K_M} ES_i + \sum_{j=i+1}^{i+1} \frac{2}{j-1} \frac{k_2}{K_M} ES_j \quad \text{for } i = 2, 3, \dots, n-1 \\
 \frac{dS_1}{dt} &= \sum_{i=n}^2 \frac{2}{i-1} \frac{k_1k_2}{k_{-1} + k_2} ES_i = \sum_{i=n}^2 \frac{2}{i-1} \frac{k_2}{K_M} ES_i
 \end{aligned} \quad (3)$$

K_M is the Michaelis–Menten constant. With the knowledge of the concentration of enzyme E and the initial concentration of substrate S_0 , the differential eqs 3 can be solved conveniently, using, for example, the Maple software. In the case of in-solution digestion, the concentrations of protein/peptides (S_j) and enzyme (E) employed in the eqs 3 are simply the bulk concentrations.

Sequential Mechanism in Nanopores. In the case of nanoconfined proteolysis, the situation can be simulated where the

initial substrate S_n would diffuse through the linear pore and react with the inside entrapped enzyme, while the generated digest products S_i ($i = 1, 2, \dots, n - 1$) would be confined to this nanopore for subsequent digestion, as shown in Scheme 1. On the surface of MOSF, there should be an equilibrium for substrate S_n between diffusion, ingress/egress and digestion:



The pseudospherical diffusion around the MOSF allows us to write the following steady-state equation for (S_{surface}):

$$\frac{D}{\delta}(S_{\text{bulk}} - S_{\text{surface}}) = k_{\text{in}}S_{\text{surface}} - k_{\text{out}}S_{\text{local}}$$

where k_{in} and k_{out} are the ingress and egress rate constants ($\text{m}\cdot\text{s}^{-1}$), δ is the thickness of the diffusion layer around the MOSF material that can be considered in a first approximation as the radius of the MOSF particle, and D is the protein diffusion coefficient in water ($\text{m}^2\cdot\text{s}^{-1}$) as shown in Scheme 1. For S_{local} , the first proteolysis step can be written as

$$k_{\text{in}}S_{\text{surface}} - k_{\text{out}}S_{\text{local}} = \frac{k_2}{K_M} E_{\text{local}}S_{\text{local}}$$

By elimination, we can express the local concentration of substrate as a function of the bulk value as

$$S_{\text{local}} = \frac{K_{\text{ads}}}{1 + \frac{k_2(1 + D_a)E_{\text{local}}}{k_{\text{out}}K_M}} S_{\text{bulk}} = \eta S_{\text{bulk}}$$

where here the Michaelis–Menten constant accounts also for the in-pore diffusion, where K_{ads} is the adsorption constant ($= k_{\text{in}}/k_{\text{out}}$), and where D_a is the Damköhler number defined by

$$D_a = \frac{k_{\text{in}}\delta}{D}$$

If $D_a \gg 1$, the diffusion rate is limiting. For $D_a \ll 1$, the ingress rate is limiting, while for $D_a \approx 1$, both rates are comparable. η is the equivalent of the effectiveness parameter defined for porous media.³⁵ When there is no specificity between matrices and substrates, $K_{\text{ads}} \leq 1$, and η is a factor always less than 1. When there is strong affinity between matrices and substrates, such as electrostatic interactions, hydrophilic/hydrophobic interactions or chemical interactions, $K_{\text{ads}} \gg 1$; η can be a factor bigger than 1. The proteolysis kinetics equations for entrapped enzyme in porous matrices are then

$$\begin{aligned} \frac{d(S_n)_{\text{local}}}{dt} &= -\eta \frac{k_2}{K_M} E_{\text{local}}(S_n)_{\text{bulk}} \\ \frac{d(S_j)_{\text{local}}}{dt} &= -\frac{k_2}{K_M} E_{\text{local}}(S_j)_{\text{local}} + \sum_{j=n-1}^{i+1} \frac{2}{j-1} \frac{k_2}{K_M} E_{\text{local}}(S_j)_{\text{local}} + \\ &\quad \eta \frac{2}{n-1} \frac{k_2}{K_M} E_{\text{local}}(S_n)_{\text{bulk}} \quad i = 2, 3, \dots, n-1 \\ \frac{d(S_1)_{\text{local}}}{dt} &= \sum_{i=n-1}^2 \frac{2}{i-1} \frac{k_2}{K_M} E_{\text{local}}(S_i)_{\text{local}} + \eta \frac{2}{n-1} \frac{k_2}{K_M} E_{\text{local}}(S_n)_{\text{bulk}} \end{aligned} \quad (4)$$

For comparison with the in-solution digestion condition, the local concentration change can be converted into a change in bulk scale. Considering that all the enzymes are immobilized in the nanospace and all the generated S_i ($i = 1, \dots, n-1$) are trapped in the nanopores, it is reasonable that

$$\begin{aligned} \frac{d(S_i)_{\text{local}}}{dt} &= \frac{V_{\text{bulk}}}{V_{\text{local}}} \frac{d(S_i)_{\text{bulk}}}{dt} \quad \text{for } i = 1, 2, \dots, n \\ (S_i)_{\text{local}} &= \frac{V_{\text{bulk}}}{V_{\text{local}}} (S_i)_{\text{bulk}} \quad \text{for } i = 1, 2, \dots, n-1 \\ E_{\text{local}} &= \frac{V_{\text{bulk}}}{V_{\text{local}}} E_{\text{bulk}} \end{aligned}$$

where V_{bulk} is the bulk volume and V_{local} is the total pore volume of the porous material, which is V_{MOSF} in the presented example, thereby eqs 4 can be simplified as

$$\begin{aligned} \frac{d(S_n)_{\text{bulk}}}{dt} &= -\eta \frac{k_2}{K_M} E_{\text{bulk}}(S_n)_{\text{bulk}} \\ \frac{d(S_j)_{\text{bulk}}}{dt} &= -\frac{k_2}{K_M} \frac{V_{\text{bulk}}}{V_{\text{MOSF}}} E_{\text{bulk}}(S_j)_{\text{bulk}} + \\ &\quad \sum_{j=n-1}^{i+1} \frac{V_{\text{bulk}}}{V_{\text{MOSF}}} \frac{2}{j-1} \frac{k_2}{K_M} E_{\text{bulk}}(S_j)_{\text{bulk}} + \\ &\quad \eta \frac{2}{n-1} \frac{k_2}{K_M} E_{\text{bulk}}(S_n)_{\text{bulk}} \quad i = 2, \dots, n-1 \\ \frac{d(S_1)_{\text{bulk}}}{dt} &= \sum_{i=n-1}^2 \frac{V_{\text{bulk}}}{V_{\text{MOSF}}} \frac{2}{i-1} \frac{k_2}{K_M} E_{\text{bulk}}(S_i)_{\text{bulk}} + \\ &\quad \eta \frac{2}{n-1} \frac{k_2}{K_M} E_{\text{bulk}}(S_n)_{\text{bulk}} \end{aligned} \quad (5)$$

Simulation of In-Solution and In-Nanopore Digestion.

From the kinetics equations, we see that two factors can greatly influence the final proteolysis efficiency of the in-nanopore digestion comparing with the in-solution one: the effectiveness factor η and the $V_{\text{bulk}}/V_{\text{MOSF}}$ ratio. On the surface of porous matrices, there is an equilibrium between substrate diffusion and in-pore reaction. When the diffusion is very fast and $k_{\text{in}} \gg k_{\text{out}}$, the overall kinetics is limited by the proteolysis rate, thereby the effectiveness factor η is ≥ 1 . When there is no specific affinity between matrices and substrates, $k_{\text{in}} \approx k_{\text{out}}$, the final process is limited by diffusion and the enzyme reaction cannot run at its maximum rate, and therefore, the η is lower than unity.³⁵ This factor indicates that, for the first step of protein digestion, the reaction rate with the entrapped enzyme is indeed depended on the diffusion and ingress/egress process, being either accelerated or limited. However, the nanoconfined digestion can benefit further peptide digestion. The generated products from first cleavage, which should still be bigger

peptides, can be restricted in the nanodomains with enzyme of high local concentration, which is indeed $V_{\text{bulk}}/V_{\text{MOSF}}$ times higher than the bulk concentration, and thereby the final products can be instantaneously generated after continuous digestion. This phenomenon is named peptide nanoconfinement effect. In contrast, for in-solution digestion, the primary products from the first cleavage diffuse back to the solution and the following steps of the digestion suffer again from a low encounter probability.

The peptide nanoconfinement effect was first evaluated by simulation. Because peptides with no or few miscleavage sites are the most important in the proteins structure identification by using the MS based bottom-up proteomics strategy, the generation of S_1 is used to evaluate the proteolysis efficiency. Considering that the MOSF has a pore volume of $1.3 \text{ cm}^3/\text{g}^{28}$ and that the final concentration of MOSF is 0.13 mg/mL , the local concentration of enzymes or restricted peptides should be increased by ~ 6000 ($V_{\text{bulk}}/V_{\text{MOSF}}$) times. As shown in Figure 2, proteins with 1, 2, and 6 reactive cleavage sites are considered. For simulated calculation, the quantities in this instance are specified as: $E = 1 \text{ mg/mL}$, $S_{n0} = 1 \text{ mg/mL}$ and $k_2/K_M = 1$ ($\text{mg/mL})^{-1}(\text{h})^{-1}$, where the same bulk concentration of enzyme and substrate are considered for both in-solution and in-nanopore digestion. From Figure 2a, it is found that nanospace confined digestion efficiency is only depended on the value of the effectiveness parameter η when the substrate has only one active cleavage site. If there is no specific affinity between matrices and substrates ($\eta \leq 1$), the in-solution digestion would be even more efficient, illustrating the lack of peptides nanoconfinement effect. For proteins with 2 or several active sites, this kind of nanoconfinement effect can be observed and is more significant for the substrate with more cleavage sites, Figure 2b,c. Meanwhile, the effectiveness factor would still strongly influence the digestion efficiency. The smaller value of η , the slower kinetics is observed. For different substrates, varied values of η can be obtained depending on the Michaelis–Menten constant K_M and the Damköhler number D_a . Therefore, it can be concluded that the peptides nanoconfinement effect is widely effective for proteins with several active cleavage sites. However, for some substrates with really small effectiveness factor, this kind of effect for accelerating proteolysis is not visible.

CE-MS Validation of the Peptides Nanoconfinement Effect. To validate the simulation, a capillary electrophoresis (CE) system was employed to realize product quantification and separation. Two simple peptides with one and two tryptic cleavage sites, respectively, angiotensin 1 (DRVYIHPFHL, theoretical M_w , 1296.5) and ACTH(1–14) (SYSMEHFRWKGKPVG, theoretical M_w , 1680.9), were selected as models. After performing digestion with free trypsin or MOSF-entrapped trypsin for various durations and further using glacial HAc to stop the digestion, the generated products were analyzed by CE and CE-MS. We note that both the in-solution and in-nanopore digestion were performed at the same substrate and enzyme bulk concentrations for comparison.

For angiotensin 1, a peptide with one tryptic cleavage site “R”, two products, DR and VYIHPFHL, without any cleavage sites are expected to be generated by tryptic proteolysis. As shown in Figure 3a, without digestion, one single peak of substrate was observed on the electropherogram, while several other peaks corresponding to products from peptide proteolysis were newly observed after a typical 23 h in-solution digestion. Subsequently using fraction collection after CE, the product

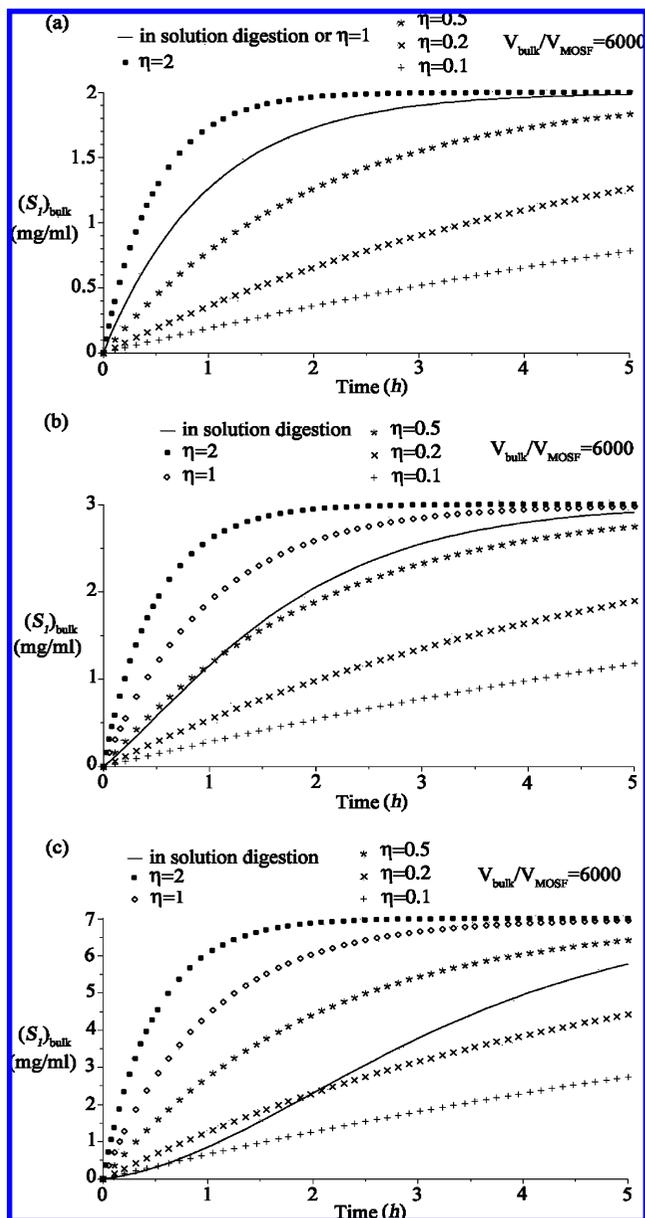


Figure 2. Predicted peptides generation from sequential mechanism: the bulk concentration of generated peptides (S_1) plotted as a function of time. The quantities in this instance are specified as $E = 1 \text{ mg/mL}$, $S_{n0} = 1 \text{ mg/mL}$, $k_2/K_M = 1 \text{ (mg/mL)}^{-1}(\text{h})^{-1}$ and $n = 2$ (a), 3 (b), and 7 (c).

VYIHPFHL (M_w , 1025) was identified by MALDI-MS (Figure 3b). It corresponds to the peak migrating after the substrate on the electropherogram as indicated in the figure by the “#” mark. The other product, DR, is not suitable for MALDI-MS sequence analysis due to its low molecular weight. Herein, neutrally coated capillaries with lengths of 35.2/26.7 cm or 51.6/43.1 cm (total/effective) \times 50 μm i.d. were used while a voltage of 25 kV was applied for CE separation. For CE-MS collection, a 43 cm long capillary and a voltage of 20.83 kV were used to maintain the electric field and the migration times relatively constant as compared to CE-UV experiments. Still, a longer sample injection time was used for more sample loading for MS detection, but not affecting the migration time (data not shown). Once the peak corresponding to the product has been identified, the corrected area (ratio of peak area to migration time) of the peaks of interest in CE-UV experiments were

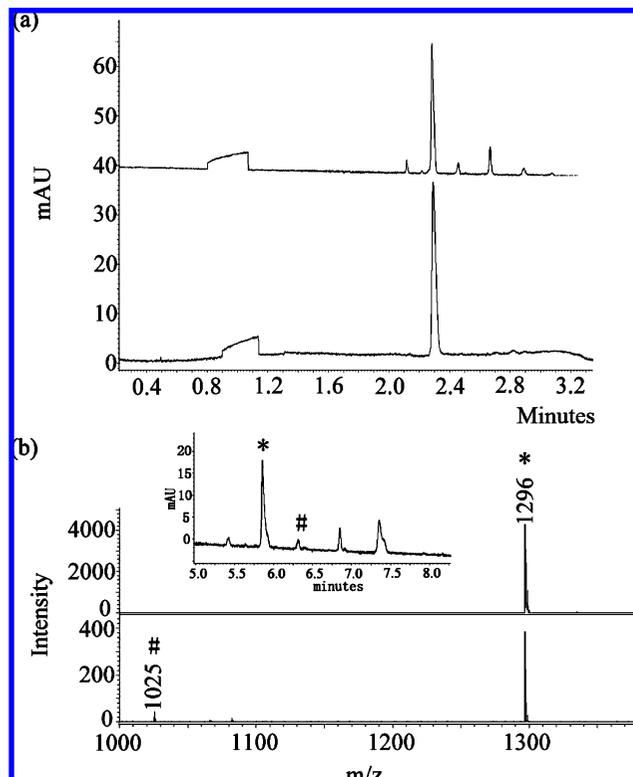


Figure 3. (a) Electropherogram of angiotensin 1 (bottom) and CE separation of its proteolysis products after 23 h in-solution digestion (upper). (b) Mass spectra of the 2 sequential fractions migrated from 5.5 to 6.17 min (upper) and from 6.17 to 6.5 min (bottom) corresponding to the inset electropherogram, using CHCA as matrix in positive ion mode. The peaks marked with an asterisk (*) represent the parent peptide, while the peaks marked with a pound symbol (#) represent the product of VYIHPFHL. Digestion condition: 20 ng/ μL substrate (in 25 mM NH_4HCO_3) was digested with 1/30 (g/g) trypsin with or without MOSF (133 ng/ μL). Capillary length is 35.2/26.7 cm (total/effective) for analysis shown in panel a and 51.6/43.1 cm (total/effective) for analysis shown in panel b inset. CE separation condition (panel a): 25 kV, 200 nm, 25 $^\circ\text{C}$, 25 s injection under 50 mbar, and a 10% HAc aqueous solution as a background electrolyte. Same CE conditions except 20.83 kV of applied voltage, 43.0 cm of capillary length and 125 s of injection period for CE-MS collection (panel b inset).

employed to represent the relative quantity of VYIHPFHL in each digest solution and further plotted as a function of the incubation time to show the reaction kinetics. As shown in Figure 5a, the in-solution digestion shows quicker reaction kinetics than that from the proteolysis using entrapped enzyme. This situation is in accordance with the theoretical simulation, where the effectiveness factor η limits the proteolysis efficiency in entrapped enzyme system for the first step of cleavage.

In case of ACTH(1–14), three final products without any reactive digestion sites of SYSMEHFR (M_w , 1056.16), WGK (M_w , 389.45), and PVG (M_w , 271.32), and two intermediate products with one reactive digestion site of SYSMEHFRWGK (M_w , 1427.60) and WGKPVG (M_w , 642.76), can be theoretically generated during proteolysis. The CE spectra of ACTH(1–14) with and without digestion are presented in Figure 4a. Similar to angiotensin 1, only one peak corresponding to the parent peptide with a migration time of \sim 6.3 min was observed without digestion, while several peaks were obtained after digestion. Because of the limitation of MALDI-MS detection

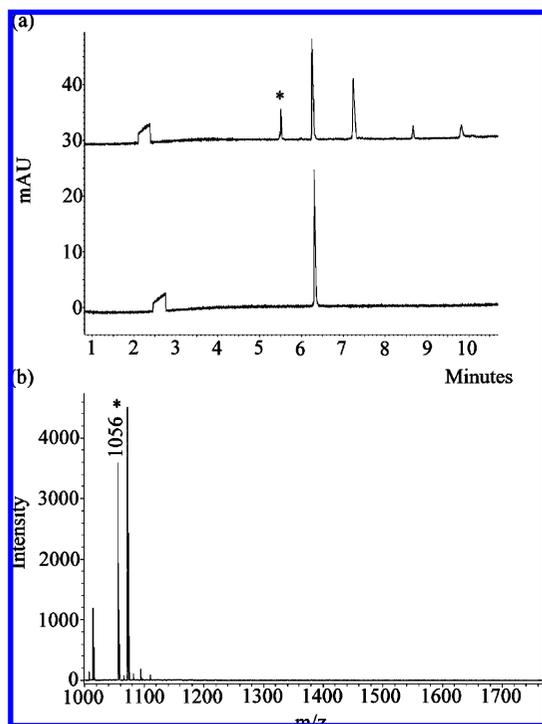


Figure 4. (a) Electropherogram of ACTH(1–14) and CE separation of its digestion products after 23.5 h in-solution digestion (from bottom to upper). (b) Mass spectrum of the fraction migrated between 5.5 and 6 min corresponding to upper panel a, using CHCA as matrix in positive ion mode. The peaks marked with asterisk (*) represent the product of SYSMEHFR. Capillary length is 51.6/43.1 cm (total/effective) for CE separation. Digestion concentration and other separation and collection conditions are the same as Figure 3.

in the low molecular mass range, only the SYSMEHFR final product was identified by CE-MS and then considered for the kinetics study. On the trace obtained by CE-UV, this product corresponds to the peak detected before the parent substrate, as marked on the CE spectrum by the asterisk (*) (Figure 4). Figure 5b shows the production kinetics of SYSMEHFR from in-solution and entrapped-enzyme based proteolysis. Illustrating the performance of the nanoconfinement effect, it is proven that proteolysis for a two-cleavage site containing peptide is much quicker when the enzyme is entrapped in the nanoporous materials compared to classical conditions using free enzymes. The experimental findings are in good accordance with the theoretical model. As investigated here, the digestion kinetics of two simple peptides provide clear evidence for the proposed kinetics model.

As compared to in-solution digestion, two factors are essential when considering nanopore-based enzymatic reactions, the nanoconfinement effect factor $V_{\text{bulk}}/V_{\text{MOSF}}$, which only depends on the concentration and total pore volume of the porous materials, and the effectiveness factor η that may sometimes limit the proteolysis efficiency in such enzyme-entrapped systems. As shown in Supporting Information SI-2, similar reaction kinetics were observed for nanoconfinement factors ranging from 100 to 10 000, indicating that the final reaction rate is restricted by the relatively slow first-step of cleavage. Thus, a wide range of matrix concentrations can be employed for the nanodomain confined proteolysis. The limitations induced by η can potentially be eliminated by using functionalized matrices, which would allow introducing specific

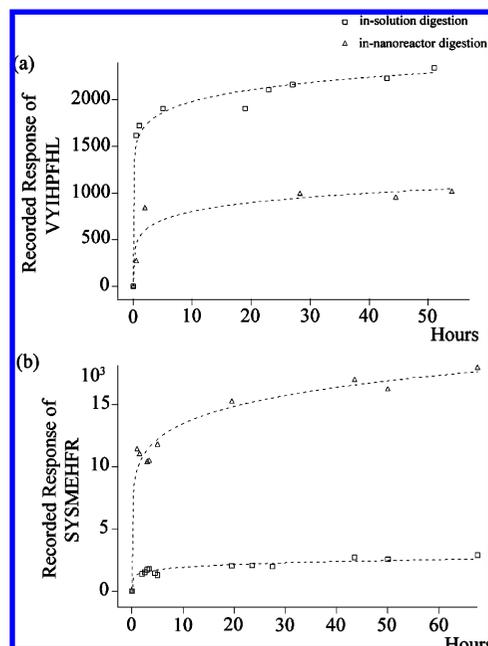


Figure 5. The ratios of peak area-to-retention time of VYIHPFHL from angiotensin 1 (a) and of SYSMEHFR from ACTH(1–14) (b) plotted as a function of digestion time for in-solution digestion (square spots) and proteolysis by using immobilized enzyme (triangle spots). Digestion concentration and other separation and collection conditions can be referred to Figure 3.

affinities, such as electrostatic, hydrophilic/hydrophobic, or chemical interactions. By accelerating the diffusion of the substrates within the nanopores, a value of η higher than 1 would be achieved. With this understanding, highly efficient enzymatic reactions would be widely achieved by introducing varied matrices for different substrates.

Conclusion

We have developed a modified kinetics model on the base of sequential mechanism, to explain and simulate the highly efficient proteolysis performed by nanoporous matrix entrapped enzymes. Validated by theoretical calculations, CE-UV and CE-MS characterization and identification, the peptide nanoconfinement effect is shown to be a key factor for the high-efficiency of enzyme-confined proteolysis, while the diffusion limitation may sometimes be a negative factor and cannot be negligible. Also, the present investigation may certainly help to understand many enzyme catalyzed reactions occurring in various environments such as a cell and shows great potential for improving high-throughput protein profiling in proteomics research.

Acknowledgment. The authors are grateful to Guochu Deng (Laboratoire de céramique, EPFL) for assisting SEM detection and also thank Prof. Chengzhong Yu's group (Fudan University) for providing MOSF. H.H.G. is grateful to Fudan University for a visiting Professor Fellowship.

Supporting Information Available: The SEM images of MOSF, the influence of nanoconfinement factor on the digestion kinetics, and the detailed information of peptides identified from the in-nanodomain and in-solution digestion of myoglobin. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.
- (2) Ptacek, J.; Devgan, G.; Michaud, G.; Zhu, H.; Zhu, X. W.; Fasolo, J.; Guo, H.; Jona, G.; Breitskreutz, A.; Sopko, R.; McCartney, R. R.; Schmidt, M. C.; Rachidi, N.; Lee, S. J.; Mah, A. S.; Meng, L.; Stark, M. J. R.; Stern, D. F.; De Virgilio, C.; Tyers, M.; Andrews, B.; Gerstein, M.; Schweitzer, B.; Predki, P. F.; Snyder, M. *Nature* **2005**, *438*, 679–684.
- (3) Rual, J. F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G. F.; Gibbons, F. D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D. S.; Zhang, L. V.; Wong, S. L.; Franklin, G.; Li, S. M.; Albala, J. S.; Lim, J. H.; Fraughton, C.; Llamas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R. S.; Vandenhaute, J.; Zoghbi, H. Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M. E.; Hill, D. E.; Roth, F. P.; Vidal, M. *Nature* **2005**, *437*, 1173–1178.
- (4) Doucette, A.; Craft, D.; Li, L. *Anal. Chem.* **2000**, *72*, 3355–3362.
- (5) Craft, D.; Doucette, A.; Li, L. *J. Proteome Res.* **2002**, *1*, 537–547.
- (6) Peterson, D. S.; Rohr, T.; Svec, F.; Frechet, J. M. J. *J. Proteome Res.* **2002**, *1*, 563–568.
- (7) Peterson, D. S.; Rohr, T.; Svec, F.; Frechet, J. M. J. *Anal. Chem.* **2002**, *74*, 4081–4088.
- (8) Kato, M.; Sakai-Kato, K.; Jin, H. M.; Kubota, K.; Miyano, H.; Toyo'oka, T.; Dulay, M. T.; Zare, R. N. *Anal. Chem.* **2004**, *76*, 1896–1902.
- (9) Li, Y.; Yan, B.; Deng, C. H.; Yu, W. J.; Xu, X. Q.; Yang, P. Y.; Zhang, X. M. *Proteomics* **2007**, *7*, 2330–2339.
- (10) Liu, Y.; Qu, H. Y.; Xue, Y.; Wu, Z. L.; Yang, P. Y.; Liu, B. H. *Proteomics* **2007**, *7*, 1373–1378.
- (11) Liu, Y.; Xue, Y.; Ji, J.; Chen, X.; Kong, J.; Yang, P. Y.; Girault, H. H.; Liu, B. H. *Mol. Cell. Proteomics* **2007**, *6*, 1428–1436.
- (12) Zhang, Y. H.; Liu, Y.; Kong, J. L.; Yang, P. Y.; Tang, Y.; Liu, B. H. *Small* **2006**, *2*, 1170–1173.
- (13) Liu, Y.; Zhong, W.; Meng, S.; Kong, J. L.; Lu, H. J.; Yang, P. Y.; Girault, H. H.; Liu, B. H. *Chem.–Eur. J.* **2006**, *12*, 6585–6591.
- (14) Liu, Y.; Lu, H. J.; Zhong, W.; Song, P.; Kong, J.; Yang, P.; Girault, H. H.; Liu, B. H. *Anal. Chem.* **2006**, *78*, 801–808.
- (15) Jarzebski, A. B.; Szymanska, K.; Bryjak, J.; Mrowiec-Bialon, J. *Catal. Today* **2007**, *124*, 2–10.
- (16) Fan, J.; Lei, J.; Wang, L. M.; Yu, C. Z.; Tu, B.; Zhao, D. Y. *Chem. Commun.* **2003**, 2140–2141.
- (17) Fan, J.; Yu, C. Z.; Gao, T.; Lei, J.; Tian, B. Z.; Wang, L. M.; Luo, Q.; Tu, B.; Zhou, W. Z.; Zhao, D. Y. *Angew. Chem., Int. Ed.* **2003**, *42*, 3146–3150.
- (18) Lei, J.; Fan, J.; Yu, C. Z.; Zhang, L. Y.; Jiang, S. Y.; Tu, B.; Zhao, D. Y. *Microporous Mesoporous Mater.* **2004**, *73*, 121–128.
- (19) Qiao, S. Z.; Yu, C. Z.; Xing, W.; Hu, Q. H.; Djojoputro, H.; Lu, G. Q. *Chem. Mater.* **2005**, *17*, 6172–6176.
- (20) Goradia, D.; Cooney, J.; Hodnett, B. K.; Magner, E. *J. Mol. Catal. B: Enzym.* **2005**, *32*, 231–239.
- (21) Hudson, S.; Magner, E.; Cooney, J.; Hodnett, B. K. *J. Phys. Chem. B* **2005**, *109*, 19496–19506.
- (22) Deere, J.; Magner, E.; Wall, J. G.; Hodnett, B. K. *Chem. Commun.* **2001**, 5, 465–466.
- (23) Hudson, S.; Cooney, J.; Hodnett, B. K.; Magner, E. *Chem. Mater.* **2007**, *19*, 2049–2055.
- (24) Shui, W. Q.; Fan, J.; Yang, P. Y.; Liu, C. L.; Zhai, J. J.; Lei, J.; Yan, Y.; Zhao, D. Y.; Chen, X. *Anal. Chem.* **2006**, *78*, 4811–4819.
- (25) Fan, J.; Shui, W. Q.; Yang, P. Y.; Wang, X. Y.; Xu, Y. M.; Wang, H. H.; Chen, X.; Zhao, D. Y. *Chem.–Eur. J.* **2005**, *11*, 5391–5396.
- (26) Qiao, L.; Liu, Y.; Hudson, S. P.; Yang, P. Y.; Magner, E.; Liu, B. H. *Chem.–Eur. J.* **2008**, *14*, 151–157.
- (27) Qian, K.; Wan, J. J.; Qiao, L.; Huang, X. D.; Tang, J. W.; Wang, Y. H.; Kong, J. L.; Yang, P. Y.; Yu, C. Z.; Liu, B. H. *Anal. Chem.* **2009**, *81*, 5749–5756.
- (28) Wang, H. N.; Zhou, X. F.; Yu, M. H.; Wang, Y. H.; Han, L.; Zhang, J.; Yuan, P.; Auchterlonie, G.; Zou, J.; Yu, C. Z. *J. Am. Chem. Soc.* **2006**, *128*, 15992–15993.
- (29) Srividhya, J.; Schnell, S. *Comput. Biol. Chem.* **2006**, *30*, 209–214.
- (30) Hollands, T. R.; Fruton, J. S. *Biochemistry* **1968**, *7*, 2045–2053.
- (31) Bull, H. B.; Currie, B. T. *J. Am. Chem. Soc.* **1949**, *71*, 2758–2760.
- (32) Sachdev, G. P.; Fruton, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3424–3427.
- (33) Shen, Y. F.; Smith, R. D. *J. Microcolumn Sep.* **2000**, *12*, 135–141.
- (34) Busnel, J.; Josserand, J.; Lion, N.; Girault, H. H. *Anal. Chem.* **2009**, *81*, 3867–3872.
- (35) Shuler, M. L.; Kargi, F. *Bioprocess Engineering: Basic Concepts*; Prentice Hall: Upper Saddle River, NJ, 2002.

PR9003954