Substrate Arrays as Enzyme Fingerprinting Tools

Jean-Louis Reymond* and Denis Wahler[a]

KEYWORDS:

combinatorial chemistry · enzymes · fingerprints · fluorescent probes · high-throughput screening

1. Introduction

The most fundamental property of any catalyst is selectivity. Selectivity is defined in terms of reaction types, substrate and product range, stereochemical preferences, and operating conditions. The activity pattern of a catalyst, which defines its function, is more important in practical terms than its actual chemical structure. Activity cannot be deduced accurately from structure because of the limited predictive value of chemical theories. Therefore, characterization of a catalyst is an experimental exercise that amounts to recording its activity for a series of substrates and reaction conditions. In contrast to the description of chemical structures, there is no generally accepted formalism for the description of catalytic activity patterns, which can represent extremely diverse and potentially infinitely large data sets. Herein, we discuss the use of substrate arrays as tools for the rapid characterization of enzyme activities and the potential use of such arrays in enzyme discovery and quality control applications.

2. One enzyme, one substrate

2.1. Enzyme assays

The activity of an enzyme is usually measured on the basis of the reaction of a particular substrate and involves monitoring any chemical or physical parameter that changes specifically upon reaction progress. The activity of enzymes is defined with reference to a particular assay and counted in Units (U), which indicate the rate of product formation in micromoles per minute per milligram of enzyme under a given set of conditions (substrate, concentration, solvent, buffer, temperature). The most popular assays are those that produce a spectrophotometric signal and use simple reagents, in particular chromogenic or fluorogenic substrates. A typical example is the use of nitrophenyl esters and glycosides to monitor the activity of esterases and glycosidases, respectively. The advent of recombinant enzyme libraries[1] has made the issue of enzyme assays very pressing in recent years and a number of novel assays have been reported. This subject has been reviewed recently.[2]

2.2. Enzyme activity profiles

The assay of an enzyme with a given substrate provides information about its activity under a given set of conditions. Classical biochemical studies of enzymes start with an exploration of a particular reaction under a series of experimental conditions. Thus, varying substrate concentration usually indicates saturation kinetics, which are interpreted as substrate binding in the Michaelis – Menten model.[3] Measurement of the substrate concentration profile at different pH values then leads to the pH – rate profile, which gives characteristic signatures for the participation of acid-base residues in catalysis, for example the role of aspartate and glutamate residues in acidic xylanases.[4] pH – rate profiles may also be indicative of which reaction step is rate limiting in a particular enzyme, for example, in protein tyrosine phosphatases.[5] Temperature profiles can be interpreted in terms of enzyme denaturation (melting)[6] or temperature-dependent activity,[7] or by use of the Arrhenius equation to determine activation parameters. Arrhenius plots for enzymic rates may be nonlinear and their interpretation can be quite problematic,[8] in particular for membrane-bound enzymes.[9] This problem is caused in part by the temperature dependence of substrate affinity.[10]

2.3. The API ZYM system

The classical biochemical view of the enzyme world as inspired by metabolic pathway analysis is the one enzyme, one substrate correspondence. This view can be extended to the notion of one assay, one enzyme. This idea was developed in the 1960s by Bussière et al. to construct an array of 16 different enzyme substrates for profiling microorganisms.[11] The array, named Auxotab, included chromogenic substrates for lipases and esterases, aminopeptidases, chymotrypsin, trypsin, phosphatases, sulfatases, and β-galactosidases. The assays were formulated as a combined set in a filter-paper format and carried out in parallel on crude microorganism cultures to assess the presence/absence of the corresponding enzymes. The so-called API ZYM system was then developed from this discovery by Daniel Monget at Biomerieux and commercialized,[12] and became a popular tool in microbiology. The API ZYM consists of a series of 19 enzyme assays characteristic for measurement of 19 enzyme activities and one blank as a reference (see Table 1). It has been used broadly to characterize microorganisms.[13] The apparent activities are classified qualitatively from the enzymic tests by

[a] *Prof. Dr. J.-L. Reymond, Dr. D. Wahler*
Departement für Chemie und Biochemie
Universität Bern
Freiestrasse 3, 3012 Bern (Switzerland)
Fax: (+41) 31-631-8057

Table 1. *Enzymes and substrates assayed by the API ZYM system.*[12]

No.	Enzyme	Substrate[a]	pH[b]	color[c]
1	none	None		[d]
2	alkaline phosphatase	2-naphthyl phosphate	8.5	purple
3	esterase C4	2-naphthyl butyrate	7.1	purple
4	lipase C8	2-naphthyl caprylate	7.1	purple
5	lipase C14	2-naphthyl myristate	7.1	purple
6	leucine aminopeptidase	L-leucyl 2-naphthylamide	7.5	orange
7	valine aminopeptidase	L-valyl-2-naphthylamide	7.5	orange
8	cystine aminopeptidase	L-cystyl-2-naphthylamide	7.5	orange
9	trypsin	*N*-benzoyl-DL-arginine 2-naphthylamide	8.5	orange
10	chymotrypsin	*N*-benzoyl-DL-phenylalanine 2-naphthylamide	7.1	purple
11	acid phosphatase	2-naphthyl phosphate	5.4	purple
12	phosphoamidase	naphthol AS Bis-phosphodiamide	5.4	blue
13	α-galactosidase	6-bromo-2-naphthyl-α-D-galactopyranoside	5.4	purple
14	β-galactosidase	2-naphthyl-β-D-galactopyranoside	5.4	purple
15	β-glucuronidase	naphthol AS Bis-β-D-glucuronide	5.4	blue
16	α-glucosidase	2-naphthyl-α-D-glucopyranoside	5.4	purple
17	β-glucosidase	6-bromo-2-naphthyl-β-D-glucopyranoside	5.4	purple
18	N-Acetyl-β-glucosaminidase	1-naphthyl-N-Acetyl-β-D-Glucosaminide	5.4	brown
19	α-mannosidase	6-bromo-2-naphthyl-α-D-mannopyranoside	5.4	purple
20	α-fucosidase	2-Naphthyl-α-L-fucoside	5.4	purple

[a] A fibrous material (filter paper) is impregnated with substrate added as an alcohol solution. [b] A pH stabilizer (tris(hydroxymethyl)aminomethane(Tris) – HCl (> pH 7) or Tris – maleate) is then added and the suface is left in contact with the biological sample to be analyzed for 2 – 4 h at 37 °C. [c] Coloration observed after reaction with a solution of Fast Blue BB (*N*-(4-amino-2,5-diethoxyphenyl)-benzamide) in 25 % aq Tris – HCl containing 10 % (w/w) lauryl sulfate. [d] Control with biological sample only.

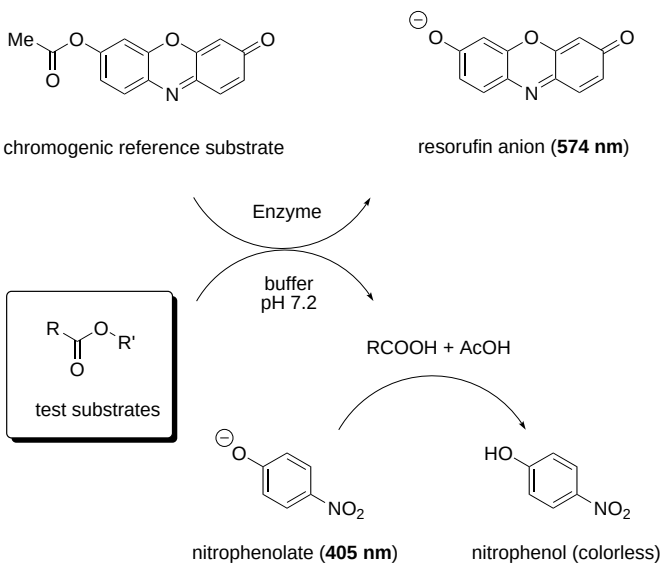
using labels such as α for a strong activity and β for a weak activity, and interpreted in terms of the presence/absence of the corresponding enzymes.[14] Newer, more exhaustive variations on this theme have been developed. The current format includes 32 different assays and comes in miniaturized, preformatted plates.[15]

3. Substrate arrays and fingerprints

3.1. Enzyme substrate profiles

The developement of enzymes as synthetic tools has brought to light that many enzymes accept a range of substrates structurally different from their natural best substrate. In this one enzyme, many substrates view, the reactivity of the enzyme across a range of substrates becomes of interest for the better description of its catalytic properties. The analysis of substrate specificities reveals which types of structures and stereoselectivities are tolerated by a given enzyme.

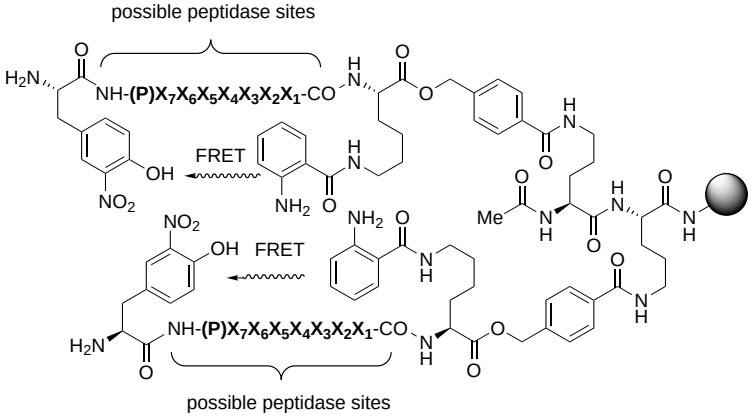
The analysis is particularly facilitated if the enzyme activity with the different substrates can be measured simply in a high-throughput format. Kazlauskas et al. recently reported the substrate profile analysis of four synthetically relevant lipases and esterases across a range of substrates.[16] The rate of hydrolysis of ester substrates was measured by monitoring the pH drop caused by the carboxylic acid product with nitrophenol as a pH indicator. In the so-called Quick E variation,[17] the reaction rates of separate enantiomers are determined by the pH method in the presence of a competing resorufin ester substrate, which is monitored orthogonally at 574 nm, and the relative rates are used to compute the enatiomeric ratio known as the E value (Scheme 1). The data gathered by these assays is then interpreted in terms of a short qualitative description of the enzyme capabilities that can be used to orient a synthetic chemist to one or the other enzyme as a function of the particular target substrate being considered for an enzymic hydrolysis step. The series of substrates used in this study is also described in a recent patent application.[18]



Scheme 1. *Principle of the Quick E method. Microscopic competition between individual test substrates and the reference substrate resorufin acetate is measured as the rate of total acid production (decrease in absorbance at 405 nm) relative to the rate of hydrolysis of the reference substrate (increase in absorbance at 574 nm). These relative rates allow computation of the relative specificity constants k_{cat}/k_M rather than the maximun reaction velocity V_{max} and thus allow prediction of E values when enantiomers are compared. R, R′ = variable groups.*

3.2. Protease profiles by peptide substrate libraries

The substrate profiles of different proteases have been investigated in detail by several groups over the last twelve years by using combinatorially synthesized peptide substrates. The general aim has been to define preferred amino acid sequences for proteases that might also serve for the design of specific inhibitors. For example Meldal et al. reported a combinatorial analysis of the cleavage specificities of subtilisin Carlsberg that used libraries of fluorescence resonance energy transfer (FRET) substrates on resin beads.[19] Libraries of FRET peptides were prepared on an enzyme-compatible resin by using the split-and-mix technique.[20] The entire library was then treated with the protease. Resin beads where the peptide was cleaved by the protease were identified by their fluorescence emission. Approximately 30 beads were fluorescent and the peptides on these beads were sequenced by Edmann degradation. Since the enzymatic cleavage was limited, unreacted peptides were still present on the beads together with the fluorescent cleaved peptides. This situation allowed determination of the sequence of the substrates, the cleavage site by the protease, and the extent of conversion.



Scheme 2. *The FRET peptide libraries used by Meldal for papain cleavage analysis,[21] synthesized on PEGA 4000 resin (PEGA = poly[acryloyl-bis-(aminopropyl)-polyethyleneglycol]). Proline was used at position 8 in one library and omitted in the second, which also omitted aspartate and glutamate as building blocks. 270 000 of approximately 20^7 sequences possible for the resin mass used were effectively prepared by the split and mix strategy. The solid-supported peptides were submitted to proteolytic cleavage directly.*

The results were analyzed with respect to preferred amino acids at P and P′ subsites (amino acid residues in a substrate undergoing cleavage are designated P1, P2, in the N-terminal direction from the cleaved bond and P1′, P2′ in the C-terminal direction). Subtilisin Carlsberg showed sequence preferences that were different for the resin-supported peptides and the soluble peptides that were prepared subsequently. The analysis converged to the consensus sequence Y(NO_2)FQPL-DEK(Abz)GD (Y(NO_2) = L-nitrotyrosine, K(Abz) = ε-(2-aminobenzoyl)-L-lysine). Interestingly, the best substrate at pH 8.5 turned out to have an alanine residue instead of proline at the P2 site, although this residue appears to be proline in almost all peptides analyzed. This shows the difficulty of interpretation of activity patterns in terms of consensus sequences. A more recent and similar study with papain shows the full details of the peptide chemistry (Scheme 2).[21] Thus, a branching lysine is chosen for the enzyme-compatible attachment to the solid phase, with aminobenzamide attached at the ε-amino group of lysine as a fluorescent marker and nitrotyrosine as a fluorescence quencher at the N terminus of the sequence. This study showed that the FRET-labeled sites are strongly coupled, and that a single consensus sequence does not exist.

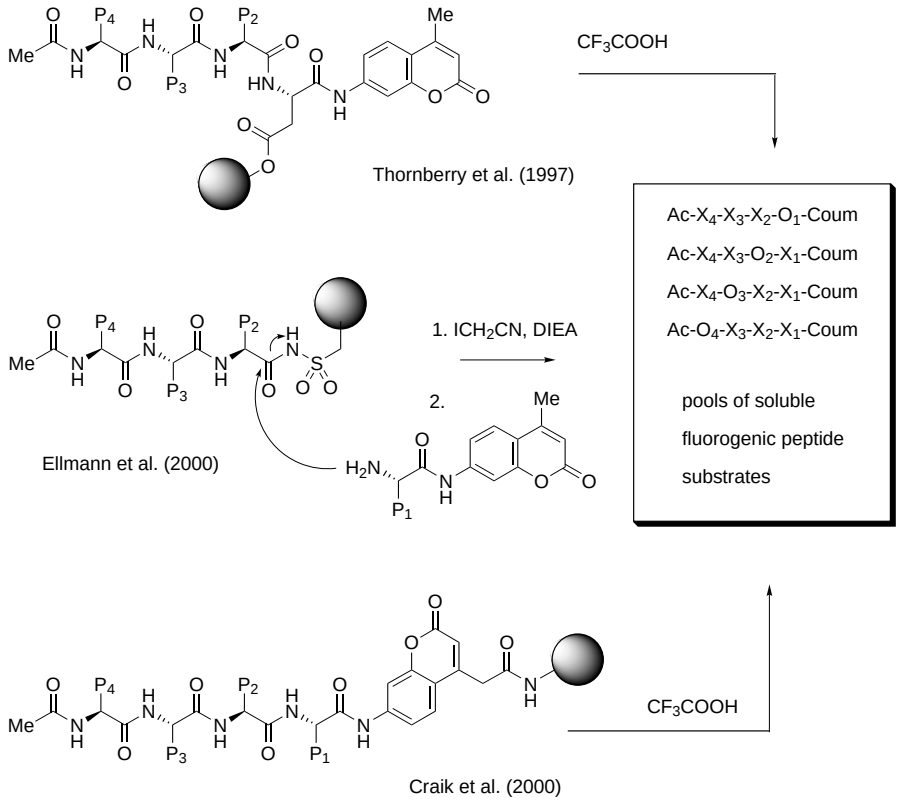
The possible occurence of consensus sequences in peptide substrates for proteases has been analyzed by using combinatorially synthesized positional scanning libraries.[22] A pioneering study was done by Thornberry et al. for scanning the P sites of interleukin-converting enzyme (ICE; also known as caspase-1, a protease involved in apoptosis).[23] *N*-(4-methyl-7-aminocoumarinyl)aspartate (D-Coum) was esterifed onto a solid support and elongated to an *N*-acetylated tetrapeptide. Three sublibraries of 20 pooled peptide mixtures were prepared. These were AcX-XOD-Coum, AcXOXD-Coum, and AcOXXD-Coum, where X represents a mixture of all 20 proteinogenic amino acids (reacted as an isokinetic mixture), O represents a defined amino acid, and Ac is an acetyl group. Considering that aspartyl – coumarinyl amide peptides were known to be excellent substrates for ICE, this study addressed the nature of the preferred residues at sites P2, P3, and P4. The study identified the peptide AcWEHD-Coum, which was found to be a 50-fold better substrate than the previously identified sequence YEAD.[24] The same approach was applied by these authors to analyze the substrate specificities of ICE, nine other caspases, and Granzyme B—all of which are proteases involved in apoptosis—and order them into three different classes according to their activity pattern.[25]

Positional scanning libraries of fluorogenic *N*-coumarinyl amides (Scheme 3) were further developed by Ellman et al. by using a sulfonamide linker strategy that introduces the P1 coumarinylamide as a nucleophile to cleave the synthetic peptide from the resin.[26] The libraries were used to define the specificities of the proteases plasmin and thrombin for positions P2 to P4 by using the reactive lysine at P1.[27] A similar study that used a strategy in which the peptide was linked to the solid support through the coumarin itself was carried out by Craik et al. to analyze an even broader range of proteases for the P1/P4 specificities.[28]

Sheppeck et al. used Ellman's sulfonamide linker chemistry to prepare a 400-member library of tripeptide – coumarinyl amides.[29] By contrast to the pooled peptide libraries described above, the fluorogenic peptides used are single compounds with a well-defined sequence and were all HPLC purified. This approach is much more work-intensive, but has the advantage of delivering readily interpretable results in terms of activities.

3.3. Enzyme fingerprints from substrate arrays

The testing of enzymes with multiple substrates as described above was carried out with the goal of delivering an activity result as clear and simple as possible, ideally perhaps a single, optimal structure of the best substrate. Thus Kazlauskas et al.



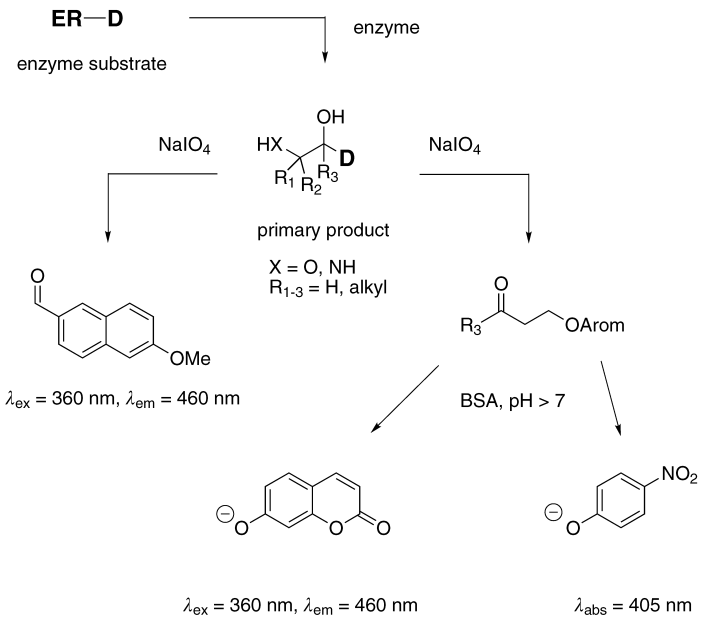
Scheme 3. *Synthesis of positional scanning libraries of fluorogenic peptidyl coumarinylamide for determination of P-site specificities. The peptide mixtures are prepared by solid-phase peptide synthesis and then cleaved from the resin for use as substrates. X_n = isokinetic mixture of proteinogenic amino acids, O_n = defined proteinogenic amino acid (cysteine is omitted and norleucine replaces methionine). Coum denotes either 4-methyl-7-aminocoumarin or 4-carboxamidomethyl-7-aminocoumarin. DIEA* = N,N-*diisopropylethylamine.*

converged on a short description of catalytic capabilities, while the protease profiling efforts converged on consensus sequences. This view adheres to the classical conception of combinatorial chemistry, which uses libraries as intermediates to single, highly active compounds, as well as to the one enzyme, one substrate analysis of the enzyme world as seen by many biochemists.

We recently approached the problem with a slightly different view that consisted of consideration of the entire pattern obtained for the activity of an enzyme with a set of substrates as a possible means to characterize or identify the enzyme.[30] Thus, the activity recorded with each and every substrate used in the test counts in the final result, as is the case in array analyses performed for characterization of expressed mRNA populations of cells by using DNA microarrays. A similar idea of pattern analysis was suggested by Thornberry et al. for their analysis of caspases discussed above[25] although they finally converged on a consensus sequence. A substrate-array analysis of enzyme activity only makes practical sense if it can be reproduced reliably and rapidly. We used our periodate-coupled fluorogenic substrate method to selectively assess the catalytic activity of enzyme samples (Scheme 4).[31] The indirect scheme based on periodate oxidation and optional β-elimination catalyzed by BSA allows use of highly fluorogenic substrates that are particularly resilient to nonspecific reactions.[32] Thus, enzyme activity can be adressed selectively and with high sensitivity.

The general applicability of the periodate oxidation scheme allowed us to prepare a variety of substrates that varied in substitution patterns at the $R_{1\text{-}3}$ groups, type of enzyme reactive groups, and chirality. We prepared substrates with hydrolytically labile functional groups such as amides, esters, carbonates, and epoxides to create arrays suited to analyze the activity of hydrolase enzymes. The different substrates were assembled, together with two reference products, in arrays of either fluorogenic or chromogenic substrates. Enzyme samples were then challenged simultaneously with these arrays which were layed out on microtiter plates. The apparent reaction rates were recorded over two hours with the corresponding microtiter-plate reader instrument and used to compute a complete set of activities for all substrates.

The observed reaction rates are reported graphically as 2D grids of gray-scale squares. The maximum observed rate in the array is set as full black, and the corresponding maximum rate is reported underneath the array. Such gray-scale 2D displays are extremely simple to generate and provide easily



Scheme 4. *Principle of the periodate-coupled assay for hydrolytic enzymes. An enzyme substrate (ER – D) that bears an enzyme reactive group (ER) and a precursor of a fluorogenic or chromogenic group (D) is subjected to the action of an enzyme. Secondary oxidation to activate the chromo/fluorogenic group (D) allows the use of nonactivated enzyme reactive groups (ER) that are silent to nonspecific reactions. BSA = bovine serum albumin, Arom = an aromatic group.*

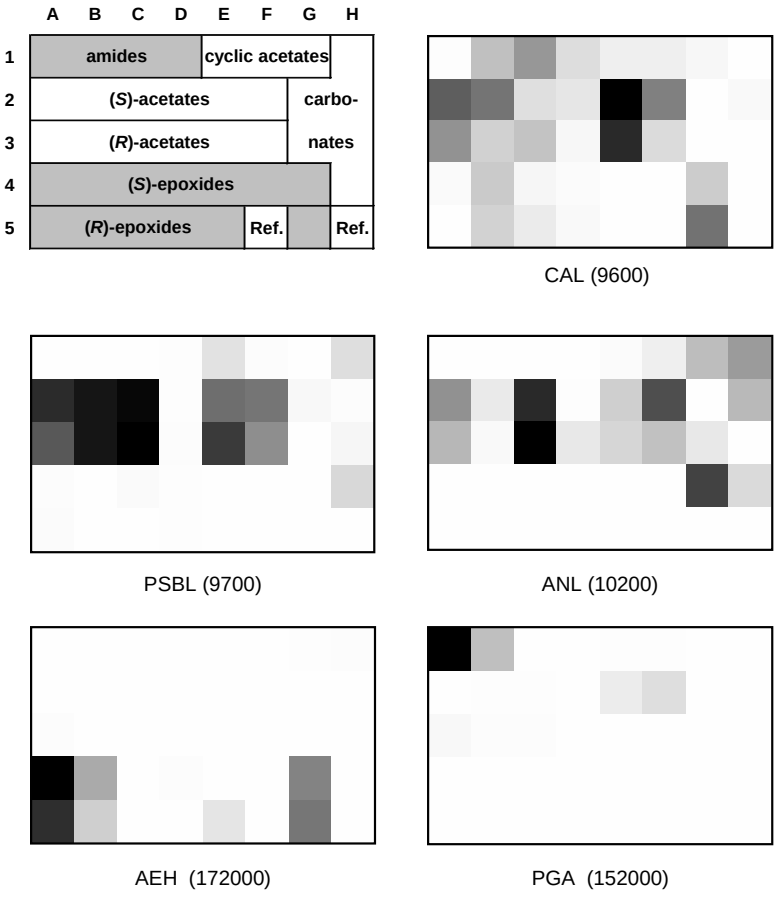


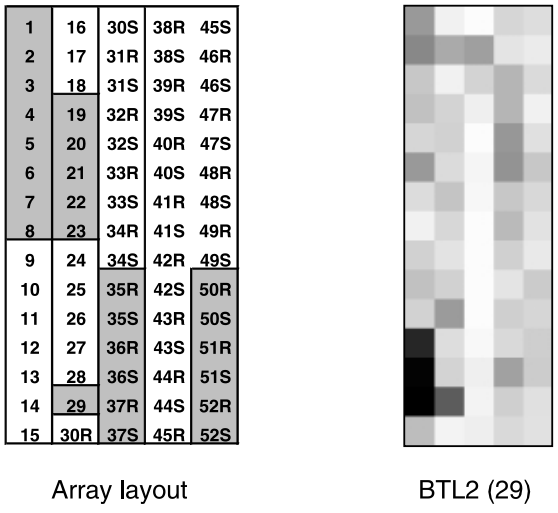
Figure 1. *Enzyme fingerprints from a fluorogenic substrate array. Two reference diol primary products (Ref.) are present in the array. Below each array the enzyme code is shown with the apparent maximum rate ($pM\,s^{-1}$; set as full black in each array) for the formation of 7-hydroxycoumarine (umbelliferone) or 6-methoxy-naphthaldehyde from the fluorogenic substrates. The rates observed in a reference array without enzyme have been substracted from the observed rates with enzymes before generation of the display. The reference diols therefore appear white. Conditions: enzyme (0.1 mg mL^{-1}), substrate (100 μM), aq borate (20 mM; pH 8.8), dimethyl formamide (2.5% (v/v)), BSA (2 mg mL^{-1}), $NaIO_4$ (1 mM), 26 °C. Codes for enzymes: CAL = Candida antarctica lipase F62299; ANL = Aspergillus niger lipase A39,043 – 7; PSBL = Pseudomonas sp. Type B lipoprotein lipase F62336; AEH = Aspergillus niger epoxide hydrolase; PGA = Escherichia coli penicillin G acylase F76427. See Ref.[30] for substrate structures and details.*

identifiable visual patterns. Most importantly, the patterns observed with the enzyme samples, which can be recorded within two hours, are reproducible and may therefore be considered as "fingerprints". Approximately 40 enzyme samples were tested and all gave unique fingerprints with the substrate arrays. Representative examples are shown in Figure 1. The interpretation of enzyme fingerprints is not unequivocal and will depend on the purity of the sample used. Activities not related to the component under consideration might also be apparent in enzyme samples that are not highly purified. These side activities might be useful to identify particular enzyme batches or to trace the presence of contaminants.

3.4. Gray-scale array display

One of the key aspects of array analysis is the graphical display chosen to report the results from the enzyme measurements. The standard approach for display of large sets of data consists of generation of either a series of 2D bar graphs or colored perspective 3D displays. The gray-scale 2D display offers a much more compact alternative that is easy to inspect by eye. It can also be used to represent other data. For example, Kazlauskas's profile of esterases and lipases with a total of 75 substrates can be represented as a 5×15 array. This graphical reanalysis summarizes the observed reactivities (Figure 2).

The peptide library work of Meldal et al.[19b, 21] Ellmann et al.[27] and Craik et al.[28] is readily reported in a gray-



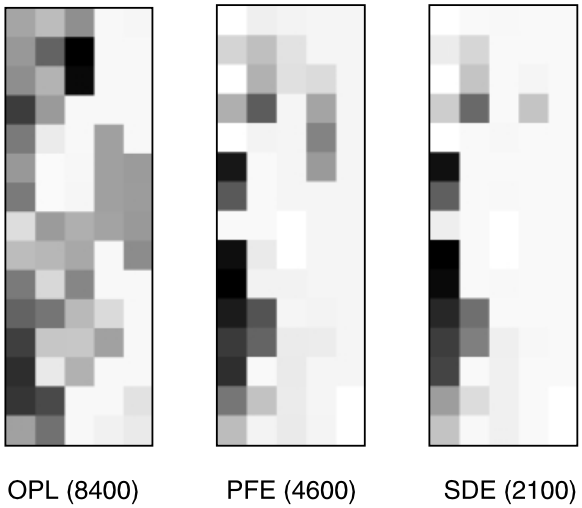


Figure 2. *Array display of esterase profiling results obtained by Kazslauskas et al.[16] The published activities are in U mg^{-1}. The data points were converted to log(1 + activity) to make weaker activities visible and then converted into a gray scale from white, for no activity, to black for maximum activity of the enzyme. The maximum activity observed with each enzyme is given in parentheses next to the enzyme code. Missing data were set to no activity. Enzyme codes: BTL2 = Bacillus thermocatenulatus lipase DSM 730; OPL = Ophiostoma piliferum lipase NRRL 18917; PFE = Pseudomonas fluorescens esterase I SIK-W1; SDE = Streptomyces diastatochromogenes lipase Tü 20. Substrates: **1** – **8**, achiral ethyl esters; **9** – **18**, achiral vinyl esters; **19** – **23**, achiral methyl esters; **24** – **28**, acetates; **29**, tributyrin; **30** – **34**, esters of chiral primary alcohols; **35** – **37**, esters of chiral secondary alcohols; **38** – **49**, methyl and ethyl esters of chiral acids; **50** – **52**, chiral γ- and β-lactones. See the original publication by Kazlauskas for structures.[16]*

scale array for amino acids at each P site (Figure 3). Each square is colored in proportion to the reaction rate or frequency of occurence of peptides with a given amino acid at a given position P (column), relative to the most frequent or fastest reacting amino acid at the particular P site, which is set as full black in that column. These displays may not necessarily represent fingerprints as they cannot be measured rapidly and simultaneously and their direct comparison must also take into account the different settings and methods of each study. Nevertheless, the display again delivers an excellent overview of these complex results. Series of gray-scale displays have been used recently to map the activity of a proteasome against two-position positional scanning libraries of fluorogenic peptide substrates.[33]

4. Active-site-directed enzyme probes

Enzymes, like all proteins, can be separated by their mass by using gel electrophoresis. One can then specifically stain the gel with a specific enzyme substrate and thus obtain confirmation that a given spot corresponds to the enzyme activity.[34] Here enzymic activity is combined with a physicochemical property of the protein for identification. The analysis is much more complex and time consuming than fingerprinting with enzyme substrates since it requires a sequence of sample preparation, electrophoretic separation under well-chosen conditions, and a staining protocol that may involve several washing and revelation operations.

Selective protein labeling methods combined with electrophoretic separation have become relevant with the advent of proteomics, whereby the ensemble of proteins expressed by a cell in a certain state are analyzed in parallel.[35] Typically, a protein sample is first reacted with a covalent modifier that carries a label such as fluorescein or biotin and the sample is then separated by 2D gel electrophoresis. The tagged proteins are visualized by revelation of the tag, for example by treatment of the gel with antifluorescein antibodies or with a labeled streptavidin reagent. Activated disulfides were explored in 1987 as selective labeling agents for cysteine proteases.[36] Recently, Cravatt et al. used a biotinylated fluorophosphonate as an active-site probe for serine proteases (Scheme 5).[37] Bogyo et al. have synthesized pooled peptide libraries similar to those discussed above to generate libraries of epoxide and vinyl sulfone inhibitors and applied them to analyze various cysteine proteases.[38] Sorensen et al. have investigated a library of

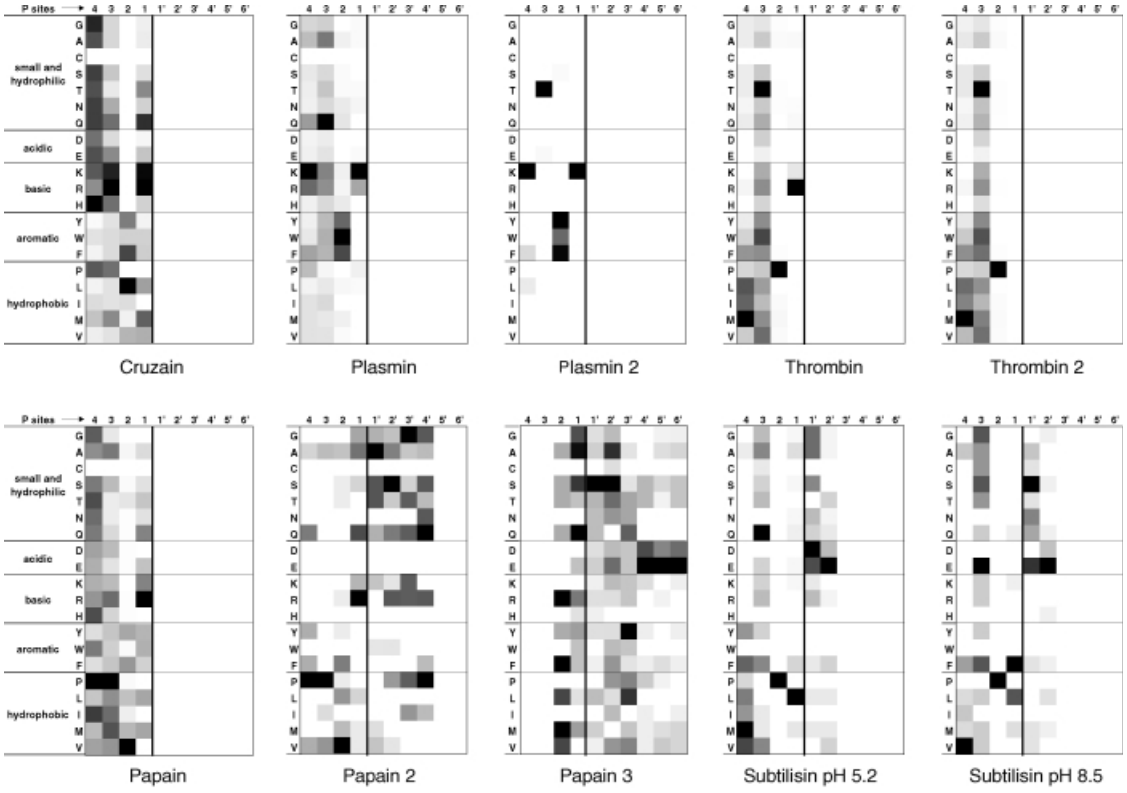
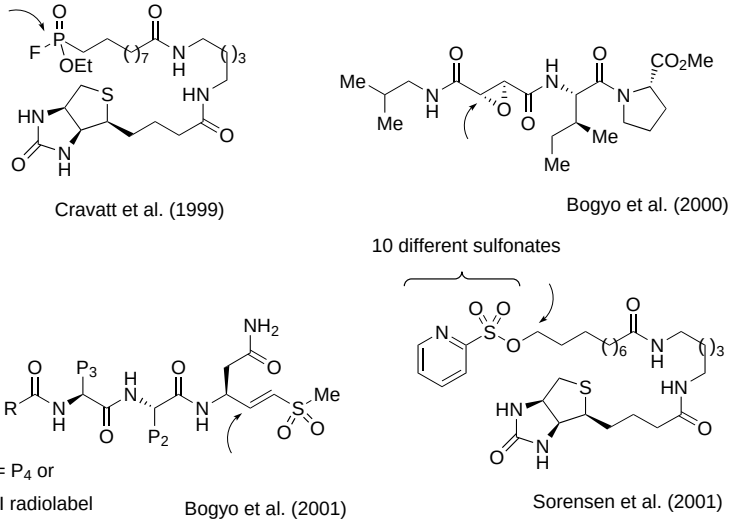


Figure 3. *Gray-scale array display for protease cleavage specificity data. Each column corresponds to one P or P′ site, with the gray scale set as full black for the most frequent (or fastest reacting) residue at this site. Residues and P sites omitted in the studies are set as white. Cruzain, plasmin, and thrombin from Craik et al.[28]: data represent apparent rates from positional scanning combinatorial libraries (PSCL) of* N*-acetylated P1-(4-carboxamidomethyl)-7-coumarinylamide tetrapeptides. P1 was set as the fastest reacting residue for scanning P2, P3, and P4; norleucine data are given in place of methionine. "Plasmin 2" from Ellman et al.[27]: data from* k_{cat}/K_M *of seven optimal* N*-acetylated tetrapeptide (4-methyl)-coumarinylamides; the PSCL data shown in Ref. [27] for plasmin is identical to that of "thrombin 2". Thrombin 2 from Ellman et al.[27]: data represent apparent rates from a PSCL of* N*-acetylated tetrapeptide (4-methyl)-coumarinylamides with P1 set as lysine; norleucine data are given instead of methionine data. Subtilisin at two pH values, and "papain 2" and "papain 3" from Meldal et al.[19b, 21]: data from the relative frequency of occurence of amino acids in solid-phase-bound octa- or nonapeptides with a nitrotyrosine residue at P4, as sequenced on active beads in a FRET solid-phase assay. The nitrotyrosine is strongly preferred at position P4 in papain.*



Scheme 5. *Active-site-directed probes for enzymes and proteins. The probes are reacted with the protein sample in solution and their label visualized by specific staining techniques after electrophoretic separation.*

10 different biotinylated tosylates as nonspecific protein labeling agents and succeeded, with the help of MS analysis and genome mining, in the identification of the enzyme c-ALDH-I (a type I aldehyde dehydrogenase) as a specific labeling target for a biotinylated pyridyl sulfonate in the proteome of rat liver cells.[39] As discussed by these authors,[39] activity-specific staining methods should allow visualization of rare yet important components of the proteome that are masked by other more abundant proteins that co-elute, and are therefore not visible in abundance-dependent identification by trypsin digest/mass spectrometry.

5. Outlook

Enzyme assays are essential tools for enzyme discovery and characterization. They are an important gateway to information such as enzyme mechanisms and enzyme substrate preferences. Data sets of enzymic rates across parameter ranges that include pH, temperature, and substrate profiles can be acquired very simply by using parallel instrumentation if suitable high-throughput screening assays are available. The enzyme assays must be reproducible, reliable, simple and fast, and optimally formulated in an array format. The results can be represented as 2D arrays of grays-cale squares that are particularly easy to inspect visually. These experiments can be faster and less complex that active-site labeling studies coupled to eletrophoretic separations. Data sets generated from enzyme assays are expected to be enzyme-specific and possibly characteristic of the enzyme. They could be used before interpretation as fingerprints for the enzymes, as demonstrated in principle by the substrate arrays discussed above. A practical realization of this idea will require definition of "canonical" structural and physicochemical parameters to be included in a reference array. Fingerprint analyses could be used to analyze the issue of functional convergence in enzymes on a broad scale. On a more practical level, fingerprints might serve as tools for enzyme discovery and for enzyme batch identification and quality control.

The authors thank the University of Bern, the Swiss National Science Foundation, and Protéus SA (Nîmes, France) for financial support.

[1] a) B. Marrs, S. Delagrave, D. Murphy, *Curr. Opin. Microbiol.* **1999**, *2*, 241 – 245; b) W. P. C. Stemmer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747 – 10751; c) W. P. C. Stemmer, *Nature* **1994**, *370*, 389 – 391; d) L. You, F. H. Arnold, *Protein Eng.* **1996**, *9*, 77 – 83; e) M. T. Reetz, K.-E. Jaeger, *Chem. Eur. J.* **2000**, *6*, 407 – 412; f) S. V. Taylor, P. Kast, D. Hilvert, *Angew. Chem.* **2001**, *113*, 3408 – 3436; *Angew. Chem. Int. Ed.* **2001**, *40*, 3310 – 3335.

[2] a) M. T. Reetz, *Angew. Chem.* **2001**, *113*, 292 – 320; *Angew. Chem. Int. Ed.* **2001**, *40*, 284 – 310; b) D. Wahler, J.-L. Reymond, *Curr. Opin. Chem. Biol.* **2001**, *5*, 152 – 158; c) F. Beisson, A. Tiss, C. Rivière, R. Verger, *Eur. J. Lipid Sci. Technol.* **2000**, 133 – 153; d) D. Wahler, J.-L. Reymond, *Curr. Opin. Biotechnol.* **2001**, *12*, 535 – 544.

[3] J. Kraut, *Science* **1988**, *242*, 533 – 540.

[4] M. D. Joshi, G. Sidhu, I. Pot, G. D. Brayer, S. G. Withers, L. P. McIntosh, *J. Mol. Biol.* **2000**, *299*, 255 – 279.

[5] J. M. Denu, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5910 – 5914.

[6] D. J. Hei, D. S. Clark, *Biotechnol. Bioeng.* **1993**, *42*, 1245 – 1251.

[7] Y.-X. Fan, P. McPhie, E. W. Miles, *J. Biol. Chem.* **2000**, *275*, 20301 – 20307.

[8] a) S. Kunugi, H. Hirohara, H. N. Ise, *Eur. J. Biochem.* **1982**, *124*, 157 – 163; b) D. G. Truhlar, A. Kohen, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 848 – 851.

[9] a) G. Lenaz, A. M. Sechi, G. Parenti-Castelli, L. Landi, E. Bertoli, *Biochem. Biophys. Res. Commun.* **1972**, *49*, 536 – 542; b) M. L. Puterman, N. Hrboticky, S. M. Innis, *Anal. Biochem.* **1988**, *170*, 409 – 420.

[10] J. R. Silvius, B. D. Read, R. N. McElhaney, *Science* **1978**, *199*, 902 – 904.

[11] J. Bussière, A. Foucart, L. Colobert, *C. R. Acad. Sci.* **1967**, *264D*, 415 – 417.

[12] D. Monget, P. Nardon, FR 2341865, Paris, **1976**; D. Monget, P. Nardon, US 4277561, **1981**.

[13] M. W. Humble, A. King, I. Philipps, *J. Clin. Pathol.* **1977**, *30*, 275 – 277.

[14] a) E. Gruner, A. Von Graevenitz, M. Altwegg, *J. Microbiol. Methods* **1992**, *16*, 101 – 118; b) P. Garcia-Martos, P. Marin, J. M. Hernandez-Molina, L. Garcia-Agudo, S. Aoufi, J. Mira, *Mycopathologia* **2001**, *150*, 1 – 4.

[15] http://www.biomerieux.com.

[16] A. M. F. Liu, N. A. Somers, R. J. Kazlauskas, T. S. Brush, F. Zocher, M. M. Enzelberger, U. T. Bornscheuer, G. P. Horsman, A. Mezzetti, C. Schmidt-Dannert, R. D. Schmid, *Tetrahedron: Asymmetry* **2001**, *12*, 545 – 556.

[17] a) L. E. Janes, R. J. Kazlauskas, *J. Org. Chem.* **1997**, *62*, 4560 – 4561; b) L. E. Janes, A. C. Löwendahl, R. J. Kazlauskas, *Chem.-Eur. J.* **1998**, *4*, 2324 – 2331.

[18] D. Freije, F. Moris-Varas, I. Brikum, L. Allen, *PCT Int. Appl.* **2001**, WO 0160986.

[19] a) K. Breddam, M. Meldal, *Eur. J. Biochem.* **1992**, *206*, 103 – 107; b) M. Meldal, I. Svendsen, K. Breddam, F. I. Auzanneau, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3314 – 3318.

[20] L. M. Bech, S. B. Sorensen, K. Breddam, *Biochemistry* **1993**, *32*, 2845 – 2852.

[21] P. M. St. Hilaire, M. Willert, M. A. Juliano, L. Juliano, M. Meldal, *J. Comb. Chem.* **1999**, *1*, 509 – 523.

[22] C. Pinilla, J. R. Appel, P. Blanc, R. A. Houghten, *Biotechniques* **1992**, *13*, 901 – 905.

[23] T. A. Rano, T. Timkey, E. P. Peterson, J. Rotonda, D. W. Nicholson, J. W. Becker, K. T. Chapman, N. A. Thornberry, *Chem. Biol.* **1997**, *4*, 149 – 155.

[24] N. A. Thornberry, S. M. Molineaux, *Protein Sci.* **1995**, *4*, 3 – 12.

[25] a) N. A. Thornberry, T. A. Rano, E. P. Peterson, D. M. Rasper, T. Timkey, M. Garcia-Calvo, V. M. Houtzager, P. A. Nordstrom, S. Roy, J. P. Vaillancourt, K. T. Chapman, D. W. Nicholson, *J. Biol. Chem.* **1997**, *272*, 17907 – 17911; b) J. L. Harris, E. P. Peterson, D. Hudig, N. A. Thornberry, C. S. Craik, *J. Biol. Chem.* **1998**, *273*, 27364 – 27373.

[26] J. B. Backes, J. A. Ellman, *J. Org. Chem.* **1999**, *64*, 2322 – 2330.

[27] B. J. Backes, J. L. Harris, F. Leonetti, C. S. Craik, J. A. Ellman, *Nat. Biotechnol.* **2000**, *18*, 187 – 193.

[28] J. H. Harris, B. J. Brackes, F. Leonetti, S. Mahrus, J. A. Ellmann, C. S. Craik, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7754 – 7759.
[29] J. E. Sheppeck II, H. Kar, L. Gosink, J. B. Wheatley, E. Gjerstad, S. M. Loftus, A. R. Zubiria, J. W. Janc, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2639 – 2642.
[30] D. Wahler, F. Badalassi, P. Crotti, J.-L. Reymond, *Angew. Chem.* **2001**, *113*, 4589 – 4592; *Angew. Chem. Int. Ed.* **2001**, *40*, 4457 – 4460.
[31] F. Badalassi, D. Wahler, G. Klein, P. Crotti, J.-L. Reymond, *Angew. Chem.* **2000**, *112*, 4233 – 4236; *Angew. Chem. Int. Ed.* **2000**, *39*, 4067 – 4070.
[32] a) G. Klein, J.-L. Reymond, *Helv. Chim. Acta* **1999**, *82*, 400; b) R. Pérez Carlòn, N. Jourdain, J.-L. Reymond, *Chem. Eur. J.* **2000**, *6*, 4154.
[33] J. L. Harris, P. B. Alper, J. Li, M. Reichsteiner, B. J. Backes, *Chem. Biol.* **2001**, *8*, 1131 – 1141.
[34] a) M. S. Lantz, P. Ciborowski, *Methods Enzymol.* **1994**, *235*, 563 – 594; b) C. U. Hellen, *Methods Enzymol.* **1994**, *241*, 46 – 58; c) C. L. Barbieri, *Exp. Parasitol.* **1988**, *67*, 159 – 166.
[35] a) S. P. Gygi, G. L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9390 – 9395; b) K. H. Lee, *Trends Biotechnol.* **2001**, *19*, 217 – 222.
[36] K. Brocklehurst, D. Kowlessur, M. O'Driscoll, G. Patel, S. Quenby, E. Salih, W. Templeton, E. W. Thomas, F. Willenbrock, *Biochem. J.* **1987**, *244*, 173 – 181.
[37] a) Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694 – 14699; b) D. Kidd, Y. Liu, B. F. Cravatt, *Biochemistry* **2001**, *40*, 4005 – 4015.
[38] a) M. Bogyo, S. Verhelst, V. Bellingard-Dubouchaud, S. Toba, D. Greenbaum, *Chem. Biol.* **2000**, *7*, 27 – 38; b) D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569 – 581; c) T. Nazif, M. Bogyo, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2967 – 2972.
[39] G. C. Adam, B. F. Cravatt, E. J. Sorensen, *Chem. Biol.* **2001**, *8*, 81 – 95.

Received: December 17, 2001
Revised: March 7, 2002 [M334]