Substrate specificity of kinase

DOI: 10.1002/anie.200700195

High-Throughput Identification of Substrate Specificity for Protein Kinase by Using an Improved One-Bead-One-Compound Library Approach**

*Yun-Gon Kim, Dong-Sik Shin, Eun-Mi Kim, Hyung-Yeon Park, Chang-Soo Lee, June-Hyung Kim, Bon-Su Lee, Yoon-Sik Lee, and Byung-Gee Kim**

Understanding the correct biological functions of numerous kinases through the substrate-specific identification of each kinase has become a key point in the characterization of the kinase–substrate relationship, identification of the signal-transduction pathways in the human kinome map, and elucidation of their related disease-metastasis mechanisms in detail.[1,2] However, the identification of the optimal substrate specificity of an individual kinase has been performed by using a limited or fixed peptide-library pool and needed laborious and time-consuming substrate-peptide decoding procedures. For instance, several attempts have been made to identify the substrate specificity of a protein tyrosine kinase (PTK) by a microarray-based deletion, an alanine-scanning peptide library,[3] and a positional-scanning peptide library with radio-labeled adenosine 5′-triphosphate (ATP).[4] Unfortunately, the fixed-position methods[3,4] have difficulties in determining the complete peptide sequences for the kinase substrate because they are based on single-site optimization for the amino acid sequences adjacent to the phosphorylated tyrosine. Moreover, the chemical ladder sequencing based on Edman degradation is unsuitable as an application in a high-throughput format owing to its repetitious procedures.[5]

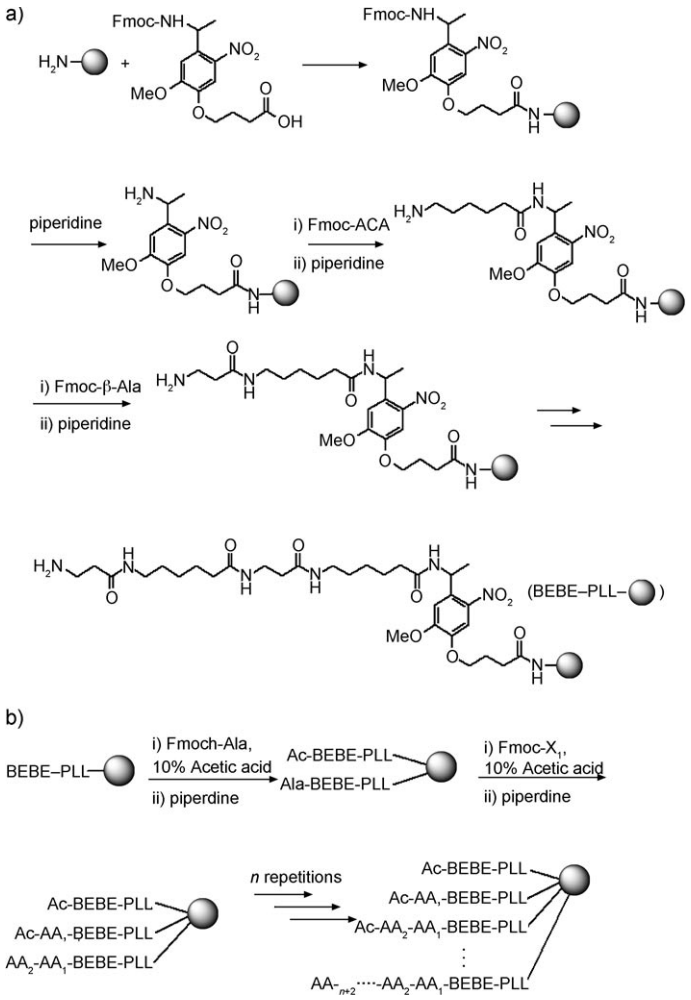
Herein, a new high-throughput system was developed to identify the substrate specificity of PTKs, $p60^{c\text{-}src}$ and ZAP-70, by using a fully randomized "one-bead one-compound" (OBOC) combinatorial peptide library[5,6] constructed by ladder synthesis and MALDI-TOF MS. This platform technology enables profile peptide sequences to be phosphorylated by various kinases, which, in turn, can provide the information to identify putative proteins that actually participate in PTK signaling pathways. Previously, several on-bead phosphorylation assays have been reported for the detection of kinase activity and screening of kinase inhibitors.[7] Compared with the conventional approaches, the advantages of this method are 1) the use of a fully randomized peptide library (over 100000), which was bound to 30 mg of the beads designed for kinase substrate-specific identification, made possible by a core-shell-type resin with a higher peptide concentration on one bead as compared with that found on a commercial resin (TentaGel); 2) a remarkably faster decoding procedure by using MALDI-TOF MS and ladder peptide synthesis; 3) an antibody-based color assay without the use of radioactive ^{32}P or ^{33}P; and 4) direct identification of actual substrate proteins of the kinase in vivo in cellular regulatory networks as well as substrate specificity.

The OBOC ladder peptide library was constructed by the "split and pool" method on a HiCore resin (Scheme 1).[8] Excluding a fixed position of the tyrosine residue and the N- and C-terminal alanine residues, four residues of amino acids were randomized with 18 natural amino acids except for cysteine and tyrosine. Both the alanine residues and a β-alanine-ε-aminocaproic acid–β-alanine-ε-aminocaproic acid (BEBE) spacer were inserted to give a nonrestricted physical environmental condition for the kinase reaction and to avoid overlap with the matrix (2,5-dihydroxybenzoic acid) peaks for the MALDI-TOF MS analysis. Thus, the resulting peptide library sequence was Ala-X-X-Tyr-X-X-Ala-BEBE-NH_2 (X = any naturally occurring amino acid except Cys and Tyr). The ladder-synthesized peptide library (30 mg), which contained over 100000 randomized sequences and was sufficient to cover all the variations of amino acids at the four positions, was used to identify the substrate specificity of the protein tyrosine kinase. After the kinase reaction of the random peptide libraries was left for 4 h at 30 °C with gentle shaking, an alkaline phosphatase (AP)-conjugated anti-phosphotyrosine antibody was used to identify the phosphorylated

[*] Y.-G. Kim, Dr. D.-S. Shin, E.-M. Kim, Prof. Y.-S. Lee, Prof. B.-G. Kim
Interdisciplinary Program for Biochemical Engineering and Biotechnology,
School of Chemical and Biological Engineering
Seoul National University
Seoul 151-742 (Korea)
Fax: (+82) 2-874-1206
E-mail: byungkim@snu.ac.kr
Dr. H.-Y. Park, Prof. B.-S. Lee
Department of Chemistry
Inha University
Incheon 402-751 (Korea)
Prof. C.-S. Lee
Department of Chemical and Biological Engineering
Chungnam National University
Daejeon, 305-764 (Korea)
Prof. J.-H. Kim
Department of Chemical Engineering
Dong-A University
Busan 604-714 (Korea)

[**] We would like to deeply acknowledge the critical reading of our manuscript and insightful discussion by Prof. Kit. S. Lam. This work was supported by the ERC program of MOST/KOSEF (R11-2000-075-03001-0) and the Korea Ministry of Science & Technology, which carries out development of protein chip technology Projects (M10415010006-04N1501-00610).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.



Scheme 1. a) Introduction of a photolabile linker and a spacer group: Fmoc-protected 4-[4-(1-aminoethyl)-2-methoxy-5-nitrophenoxy] butyric acid (Fmoc-PLL; PLL = photolabile linker) was coupled to the resin, and the BEBE spacer was synthesized on the resin by using benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP)-mediated Fmoc chemistry. b) Preparation of a ladder peptide on HiCore resin: Fmoc–amino group (Fmoc-X) and 10% (mol/mol) acetic acid mixture were used as building blocks so that a certain portion of amino groups were acetylated by using acetic acid to generate ladder members on the resin. Ac = acetyl, ACA = aminocaproic acid.

peptide beads by virtue of their color changes after incubation with the AP substrate solution. The dark-red, phosphorylated beads were simply detected by the use of an inverted microscope. Control experiments were performed to determine whether anti-phosphotyrosine antibodies produced false-positive results. The peptide sequences were exactly determined by MALDI-TOF MS, but the isobaric residues (e.g., Ile and Leu or Gln and Lys) could not, as expected, be differentiated by MS. Three-dimensional structure modeling of the optimal kinase substrate sequence was carried out as described in the Supporting Information.

To use the fully randomized peptide library in an experiment, the quantity of peptides obtained from one bead (< 100 μm) must be large enough to be analyzed by MALDI-TOF MS. In the previous work, the core-shell-type resin, in which all amino groups are located at its surface, showed a superior photolytic cleavage yield as compared with TentaGel resins.[8] The core-shell-type resin enabled us to obtain reliable MALDI-TOF MS spectra containing ions that corresponded to the full-length peptide as well as to ladder sequences of the peptide at the 50 pM level. Mass spectra showing the exact ladder sequences could be accurately calculated and interpreted as the kinase-reactive peptide sequences (Figure 1). These results showed that the quantity of the peptides obtained from one bead is sufficient to be sequenced by MALDI-TOF MS. In addition, the PLL was quite efficient to accelerate the rate of the cleavage reaction and collect peptides for MALDI-TOF MS analysis. Rapid photocleavage (10 min) by 365-nm UV irradiation was more suitable for the large-scale bioassay system than that of the conventional cyanogens bromide (CNBr)-promoted cleavage method (over 12 h).[9] The solid-phase screening method using PLL has been previously carried out for the identification of inhibitors of a recombinant cysteine protease, but in this case, the polyethylene glycol polyacrylamide (PEGA) resin (300–800 μm) was too large to cover the complexity ($18^4 = 1.05 \times 10^5$) of the peptide library bound to 30 mg of resin.[10]

To validate the feasibility of this method, $p60^{c\text{-}src}$ was chosen as its substrate specificity is well known and has been previously tested in this system.[11] A total of 100 red-colored beads were selected from the OBOC peptide library after phosphorylation by $p60^{c\text{-}src}$, and 81 peptides (81%) were unambiguously analyzed by high quality MALDI-TOF MS spectra. The phosphorylated peptide sequences found are summarized in the Supporting Information, and the critical residues for the phosphorylation were determined statistically

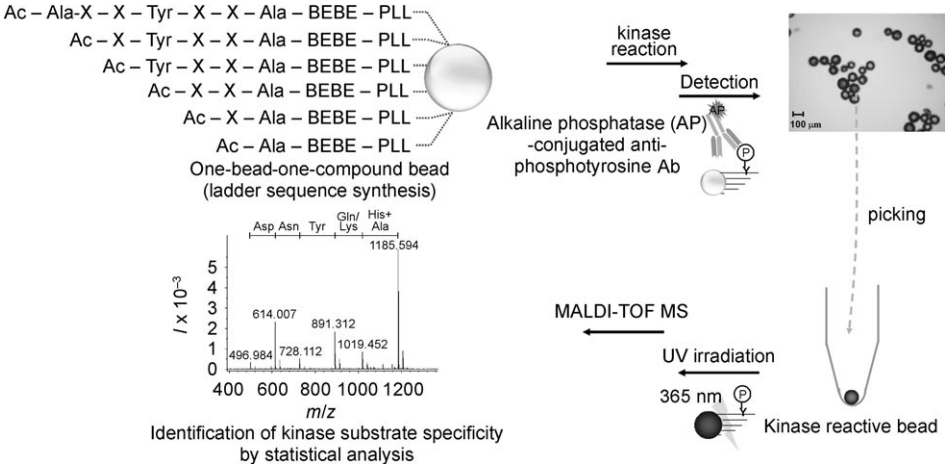


Figure 1. Experimental workflow illustration for determining substrate specificities of $p60^{c\text{-}src}$ and ZAP-70 tyrosine kinase. Ab = antibody.

based on the peptide sequencing results (Figure 2). As expected from previous studies, the substrate specificity of p60$^{c\text{-}src}$ showed that acidic amino acids (Glu and Asp) at the +1 and +2 positions and Ile at the −1 position are necessary for efficient phosphorylation of the tyrosine residue.

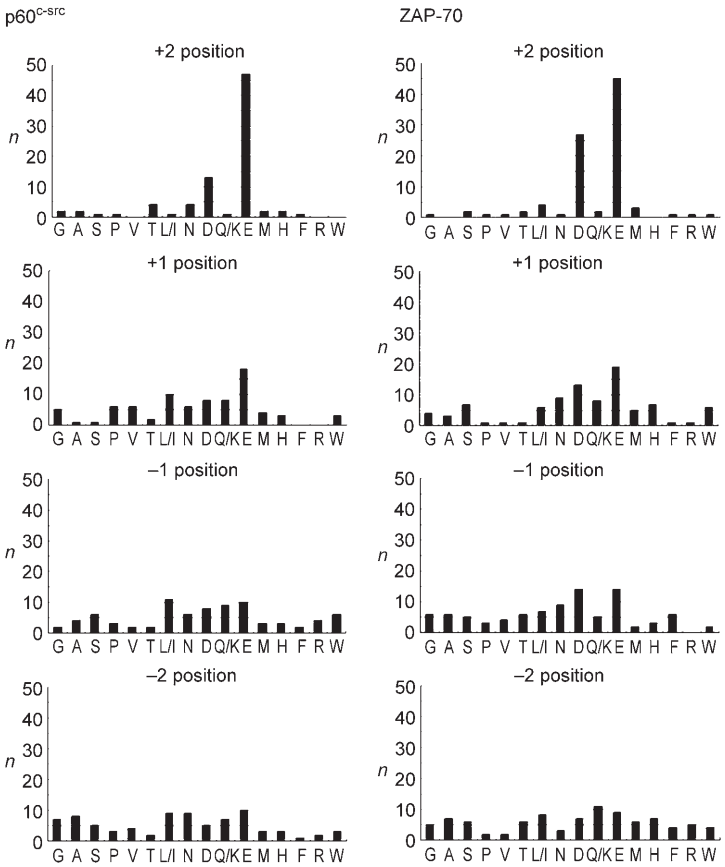


Figure 2. Substrate specificities of the p60$^{c\text{-}src}$ and ZAP-70 tyrosine kinases. The *y* axis represents the number of appearances of a particular amino acid at a certain position, and the *x* axis indicates the one-letter amino acid code. The isobaric amino acids, leucine/isoleucine (L/I) and glutamine/lysine (Q/K) are indicated as a pair. (+: carboxy-group terminal to phosphotyrosine of the optimal phosphopeptides; −: amino-group terminal to the phosphotyrosine in the optimal phosphopeptides). *n* = number of appearances.

To expand the utility of this system, the identification of substrate specificity for the previously uncharacterized ZAP-70 tyrosine kinase was carried out. To the best of our knowledge, this is the first report on the substrate specificity of ZAP-70 kinase by using a combinatorial peptide library. ZAP-70 tyrosine kinase is well known as a T-cell-specific PTK and plays a critical role in the activation of the T cell via its antigen receptor (TCR) and is therefore a rational target for immunomodulatory therapies.[12] The substrate specificity of ZAP-70 was identified in the same manner as for the p60$^{c\text{-}src}$ substrate identification (Figure 1 and the Supporting Information). In this case, the peptide sequencing efficiency by MADLI-TOF MS analysis was 92%. Interestingly, glutamic acid (Glu) at the +2 and +1 position appeared most frequently (49% and 21%) according to the sequence analysis. Aspartic acid (Asp) at the +2 position was the second most frequently appearing residue (29%). Moreover, Glu and Asp were shown to be the most likely amino acids present at the −1 position, but there was no selectivity at the −2 position. The statistical data suggest that ZAP-70 tyrosine kinase recognizes the acidic amino acids (Glu and Asp) at both sides of the tyrosine residue.

To confirm the screened results of real peptide phosphorylation, a modified image subtraction approach was used as described in the Supporting Information.[13] After the substrate screening for ZAP-70 kinase, three (AEIYEDA, ADTYDEA, and AHTYEDA) of the screened peptides were individually synthesized on the beads. Their phosphorylation reactions were re-examined to verify not only that the selection results of the beads of phosphorylated peptides were not false-positive results, but also that their selection was not a result of nonspecific binding to the anti-phosphotyrosine antibody (Figure 3a and b show only the case of AEIYEDA). As a negative control, beads incubated without kinase were used and the subsequent procedure was the same as for the kinase assay (Figure 3a). This result clearly demonstrated that there was no nonspecific interaction between the acidic residue and the anti-phosphotyrosine antibody and that the kinase specifically recognized the substrate. Furthermore, the phosphorylation on the tyrosine residue of the selected peptide (AEIYEDA) was re-examined by using MALDI-

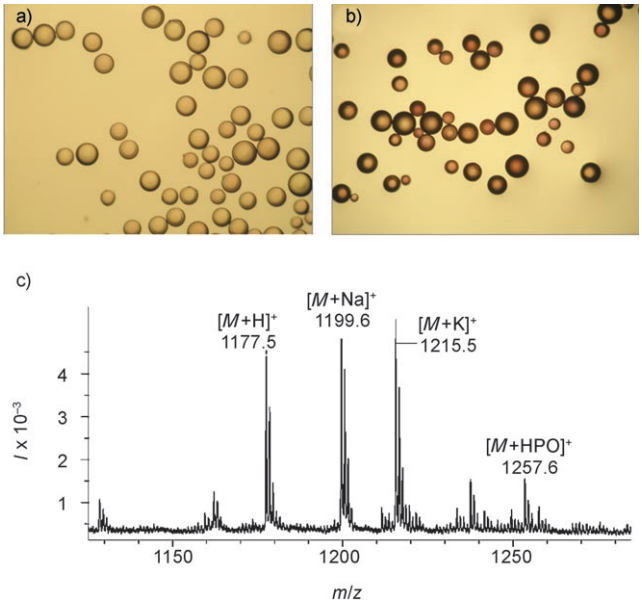


Figure 3. a) Microscopic image of negative-control beads (AEIYEDA) without nonspecific binding by an anti-phosphotyrosine antibody. b) Microscopic image of positive coloring beads (AEIpYEDA) by the anti-phosphotyrosine antibody. c) The phosphorylated ZAP-70 tyrosine kinase substrate, AEIpYEDA, was analyzed by MALDI-TOF MS in positive-ion mode.

TOF MS after performing a kinase reaction on the bead (Figure 3c). The phosphorylated peptide peak was successfully identified at m/z 1257.6 and is attributed to the addition of a phosphate group to the tyrosine residue arising from the reaction of the ZAP kinase. Thus, this result shows that the substrate peptides obtained from positively colored beads were indeed phosphorylated and that they were specifically recognized by the anti-phosphotyrosine antibody.

In addition, the annotation of potential target proteins for the $p60^{c\text{-}src}$ and ZAP-70 kinase signal-transduction pathway was undertaken by using the SwissProt database (Table 1).

Table 1: The top-scoring matched proteins to the substrate sequences of PTK $p60^{c\text{-}src}$ and ZAP-70 among human proteins in the SwissProt database.

Protein tyrosine kinase	Identified sequence	Protein name(s)
$p60^{c\text{-}src}$	AKYDE	epidermal growth factor receptor substrate
	INYEE	G-protein coupled receptor
	AIYEE	Src-activating and signalling molecule protein
	NPYIE	C-Rel proto-oncogene protein
	DTYDE	T-cell receptor zeta chain Proto-oncogene CBL-B/c-CBL
	EIYED	proto-oncogene Vav
ZAP-70	LIYDE	proto-oncogene tyrosine protein kinase receptor ret (C-ret)
	HTYED	ephrin type-A receptor 5 (Tyrosine-protein kinase receptor EHK-1)

The c-src kinase is a multifunctional protein that regulates oncogenesis through direct interaction with various cellular factors and non-receptor proteins.[14a] In this study, the identified $p60^{c\text{-}src}$ substrate sequences were matched with the kinase target proteins in human cancer and various signaling pathways.[14b,c] In addition, the Blast search with the screened ZAP-70 kinase active peptide sequences was in perfect agreement with the motifs predicted for the T-cell-receptor zeta chain,[15a] proto-oncogene Cbl,[15b] and Vav proteins[15c] that have been proved experimentally to interact with ZAP-70. Thus, these results proved that this method can be used for the direct identification of target proteins for the phosphorylation process of various kinases in vivo as well as tabulation of their substrate specificities.

In summary, a high-throughput approach was described for the identification of the substrate specificity of PTKs by use of the improved OBOC peptide library synthesized with ladder sequences and analyzed by MALDI-TOF MS. The small size of the HiCore resin (< 100 μm) used in constructing the combinatorial peptide library enabled us to obtain the fully randomized peptides with reasonably manageable amounts of resin (30 mg). The quantity of peptides released from one bead is sufficient to be sequenced by MALDI-TOF MS, and the photoelution of the ladder peptides on HiCore resin facilitates more-rapid and convenient sequencing of the substrate peptides. In addition, in spite of the use of solid-phase phosphorylation, the identified substrate sequences matched well with sequences surrounding the actual phosphorylation sites in vivo when compared with a Blast analysis by using the SwissProt database.[16] This new strategy has a great potential to be applied not only to analyze the substrate specificity of other general tyrosine kinases as well as serine/threonine kinases, but also to analyze substrate specificities of other enzyme systems, such as glycosyltransferases, proteases, transglutaminases, etc.

Received: January 16, 2007
Revised: April 27, 2007
Published online: June 6, 2007

Keywords: enzymes · high-throughput screening · mass spectrometry · peptides · substrate specificity

[1] F. A. al-Obeidi, J. J. Wu, K. S. Lam, *Biopolymers* **1998**, *47*, 197–223.
[2] S. A. Johnson, T. Hunter, *Nat. Methods* **2005**, *2*, 17–25.
[3] M. Uttamchandani, E. W. Chan, G. Y. Chen, S. Q. Yao, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2997–3000.
[4] J. E. Hutti, E. T. Jarrell, J. D. Chang, D. W. Abbott, P. Storz, A. Toker, L. C. Cantley, B. E. Turk, *Nat. Methods* **2004**, *1*, 27–29.
[5] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.
[6] K. S. Lam, M. Lebl, V. Krchnak, *Chem. Rev.* **1997**, *97*, 411–448.
[7] a) S. E. Martin, B. R. Peterson, *J. Pept. Sci.* **2002**, *8*, 227–233; b) S. Akita, N. Umezawa, T. Higuchi, *Org. Lett.* **2005**, *7*, 5565–5568; c) O. von Ahsen, U. Bomer, *Chembiochem* **2005**, *6*, 481–490; d) J. J. Slon-Usakiewicz, J. R. Dai, W. Ng, J. E. Foster, E. Deretey, L. Toledo-Sherman, P. R. Redden, A. Pasternak, N. Reid, *Anal. Chem.* **2005**, *77*, 1268.
[8] H. Kim, J. K. Cho, W. J. Chung, Y. S. Lee, *Org. Lett.* **2004**, *6*, 3273–3276.
[9] A. Song, J. Zhang, C. B. Lebrilla, K. S. Lam, *J. Am. Chem. Soc.* **2003**, *125*, 6180–6188.
[10] P. M. St Hilaire, L. C. Alves, F. Herrera, M. Renil, S. J. Sanderson, J. C. Mottram, G. H. Coombs, M. A. Juliano, L. Juliano, J. Arevalo, M. Meldal, *J. Med. Chem.* **2002**, *45*, 1971–1982.
[11] Z. Songyang, K. L. Carraway, 3rd, M. J. Eck, S. C. Harrison, R. A. Feldman, M. Mohammadi, J. Schlessinger, S. R. Hubbard, D. P. Smith, C. Eng, et al., *Nature* **1995**, *373*, 536–539.
[12] D. Qian, A. Weiss, *Curr. Opin. Cell Biol.* **1997**, *9*, 205–212.
[13] A. Lehman, S. Gholami, M. Hahn, K. S. Lam, *J. Comb. Chem.* **2006**, *8*, 562–570.
[14] a) R. Ishizawar, S. J. Parsons, *Cancer Cell* **2004**, *6*, 209–214; b) J. S. Biscardi, R. C. Ishizawar, C. M. Silva, S. J. Parsons, *Breast Cancer Res.* **2000**, *2*, 203–210; c) Y. C. Ma, J. Huang, S. Ali, W. Lowry, X. Y. Huang, *Cell* **2000**, *102*, 635–643.
[15] a) N. Isakov, R. L. Wange, W. H. Burgess, J. D. Watts, R. Aebersold, L. E. Samelson, *J. Exp. Med.* **1995**, *181*, 375–380; b) M. L. Lupher, Jr., K. A. Reedquist, S. Miyake, W. Y. Langdon, H. Band, *J. Biol. Chem.* **1996**, *271*, 24063–24068; c) S. Katzav, M. Sutherland, G. Packham, T. Yi, A. Weiss, *J. Biol. Chem.* **1994**, *269*, 32579–32585.
[16] M. B. Yaffe, *Nat. Methods* **2004**, *1*, 13–14.

