ELSEVIER

FEBS Letters 340 (1994) 231–235

FEBS 13738

FEBS LETTERS

Identification of in vivo phosphorylation sites of SET, a nuclear phosphoprotein encoded by the translocation breakpoint in acute undifferentiated leukemia

Yoshifumi Adachi[a], George N. Pavlakis[a], Terry D. Copeland[b,*]

[a]*Human Retrovirus Section and* [b]*Protein Chemistry, ABL-Basic Research Program, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, MD 21702–1201, USA*

Received 25 January 1994

Abstract

SET, the translocation breakpoint-encoded protein in acute undifferentiated leukemia (AUL), is identified as a 39-kDa phosphoprotein found predominantly in the cell nuclei [1994, J. Biol. Chem. 269, 2258–2262]. SET is fused to a putative oncoprotein, CAN, in AUL and is thought to regulate the transformation potential of SET-CAN by its nuclear localization and phosphorylation. We investigated in detail the in vivo phosphorylation of SET. Phosphorylation of SET occurred in all human cell lines examined in vivo, primarily on serine residues. Endoproteinase Glu-C digestion of phosphorylated SET yielded two phosphopeptides. By radiosequencing, we identified the in vivo phosphorylation sites of SET as Ser^{9} and Ser^{24}. The surrounding sequences of Ser^{9} and Ser^{24} contained an apparent consensus site sequence for protein kinase C.

Key words: SET; Nuclear protein; Protein phosphorylation; Translocation; Acute undifferentiated leukemia

1. Introduction

The chromosomal translocation (6;9) (p23;q34) in acute nonlymphocytic leukemia resulted in the formation of a *dek–can* fusion gene [1]. In a case of acute undifferentiated leukemia (AUL), the putative oncogene *can* was fused to a different cellular gene, named *set*, and is assumed to have been activated [2]. SET protein encoded by the *set* gene was recently biochemically identified [3]. Characterization of SET revealed that it is a 39-kDa phosphoprotein found predominantly in the cell nuclei and is expressed ubiquitously in various human cell lines. Analogous to acidic domains in NAP-1, HMG-1, HMG-2, B-23, nucleolin, GAL4, and VP16 [4–10], the acidic motif of SET is thought to serve as a nucleosome/chromatin assembly domain or a transcription activation domain [2,3]. It is known that the variable phosphorylation of nuclear proteins has functional consequences for the regulation of gene expression, replication, chromatin conformation, and nuclear transport by protein–protein and protein–DNA/RNA interactions and constitutes one mechanism by which nuclear events are modulated by external stimuli [11–20]. Similarly, the phosphorylation of SET may be involved in the regulation of SET function(s) in the nuclei and the transformation potential of SET-CAN in AUL. However, little is known about the phosphorylation of SET.

We were therefore interested in investigating in detail the phosphorylation of SET. For this purpose, we isolated ^{32}P-labeled SET from various human cell lines by metabolic labeling and immunoprecipitation, and determined in vivo phosphorylation sites within the SET molecule by radiosequencing. Our findings should be helpful in studying the function of SET and chimeric SET-CAN proteins in the mechanism of AUL leukemogenesis.

*Corresponding author. Fax: (1) (301) 846 1666.

Abbreviations. AUL, acute undifferentiated leukemia; kDa, kilodalton(s); HTLV-I, human T-cell leukemia virus type I; EDTA, ethylenediaminetetraacetic acid; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PVDF, polyvinylidene difluoride; TPCK, L-1-*p*-tosylamino-2-phenylethyl chloromethyl ketone; HPLC, high performance liquid chromatography.

2. Materials and methods

2.1. Cells and metabolic labeling

HUT-102, MT-2, and TL-Su [21] are human T-cell lines infected with human T-cell leukemia virus type I (HTLV-I). HUT-78 [22], H-9 [23], and Jurkat [7] are uninfected human T-cell lines. Raji [21] is a human B-cell line. K-562 [24] and HL-60 [24] are human cell lines derived from erythroleukemia and promyelocytic cells, respectively. HOS [7] and HeLa [7] are human cell lines derived from osteogenic sarcoma and epitheloid carcinoma, respectively. Hematopoietic cell lines were rou-

SSDI 0014-5793(94)00123-D

tinely maintained in RPMI 1640 medium supplemented with 10% fetal calf serum at 37°C in humidified air with 5% CO_2. Cell lines HOS and HeLa were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Metabolic labeling of cells with [^{32}P]orthophosphate (18.5 MBq/ml, >315 TBq/mmol) was conducted by incubating cells (1×10^6 cells/ml) in phosphate-free medium supplemented with 10% fetal calf serum dialyzed against Tris-buffered saline (20 mM Tris-HCl (pH 7.5), 0.15 M NaCl) at 37°C for 2 h as previously described [25,26].

2.2. Peptide synthesis and antibody preparation

Oligopeptides were synthesized with an automated peptide synthesizer (Applied Biosystems model 430A) using *t*-butoxycarbonyl amino acids and *p*-methylbenzhydrylamine resins or using 9-fluorenylmethoxycarbonyl amino acids and PAL supports (Millipore) as described previously [7,26]. The synthetic peptide AQAAKVSKKELNSNC (residues 3–16 of SET) was conjugated to keyhole limpet hemocyanin through the cysteine handle with the bifunctional reagent *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (Pierce). Antiserum against the conjugated peptide was raised in rabbits. An immunoglobulin G (IgG) fraction of antiserum was prepared as described previously [15,24] and used in this study.

2.3. Immunoprecipitation and gel electrophoresis

Metabolically labeled cells were lysed with RIPA buffer (50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 4 mM EDTA, 1% Nonidet P-40, 0.1% sodium deoxycholate, 10 mM $Na_4P_2O_7$, 10 mM NaF, 2 mM Na_3VO_3, 1 mM phenylmethylsulfonyl fluoride) at 4°C for 30 min. After centrifugation at 15,000 rpm (Eppendorf microcentrifuge) for 10 min, the supernatant was incubated with anti-SET serum at 4°C for 2 h. The immunocomplexes were precipitated with protein A-Sepharose CL-4B, then extensively washed with RIPA buffer and LSW buffer (10 mM Tris-HCl (pH 7.6), 10 mM NaCl) as described previously [15,17,25,26]. The immunoprecipitates were subjected to SDS-PAGE on 8–16% gradient polyacrylamide gels according to the procedure of Laemmli [27]. After electrophoresis, gels were stained with Coomassie brilliant blue R-250, dried, and subjected to autoradiography.

2.4. Phosphoamino acid analysis

^{32}P-labeled SET was transferred to an Immobilon PVDF membrane (Millipore) and hydrolyzed in 6 N HCl vapor at 110°C for 2 h. The resulting phosphoamino acid samples were mixed with standards (phosphoserine, phosphothreonine, and phosphotyrosine; 5 μg each) and applied onto cellulose-coated thin-layer glass plates (20 × 20 cm). High-voltage electrophoresis was performed in pH 3.5 buffer (pyridine/acetic acid/water; 1:10:189, by volume) at 4 °C. ^{32}P-labeled phosphoamino acids were identified by autoradiography, and phosphoamino acid standards were identified by spraying the plates with ninhydrin [25,26].

2.5. Endoproteinase digestion and HPLC separation

Endoproteinase Glu-C and TPCK-trypsin were used for digestion of ^{32}P-labeled SET, sequentially. Digestions were carried out at 37°C in 50 mM NH_4HCO_3 (pH 7.8) for 4 h followed by overnight incubation with a second identical dose of proteinase. The samples after each digestion were subjected to reversed-phase HPLC using a C_{18} column (0.46 × 30 cm, Waters) on an LKB HPLC system as previously described [26]. The elution was performed by a linear gradient of 0–40% solvent B for 120 min (solvent A: 0.05% trifluoroacetic acid in water; solvent B: 0.05% trifluoroacetic acid in acetonitrile).

2.6. Radiosequencing

Each phosphopeptide from reversed-phase HPLC was subjected to sequential Edman degradation in a Beckman 890C automated spinning cup sequenator with 10 nmol of horse apomyoglobin (Sigma) carrier as described previously [26]. Fractions collected at each cycle of sequential analysis were transferred to scintillation vials, dried with N_2 and heat, redissolved in 8 ml of Aquasure (DuPont-New England Nuclear), and allowed to equilibrate in the dark for 2 h, and scintillation counted.

3. Results

3.1. In vivo phosphorylation of SET

The SET protein encoded by the *set* gene open reading frame consists of 277 amino acids and has a molecular mass of 39 kDa. A number of various human cell lines express nearly the same level of SET, as recently reported [3]. To examine whether the in vivo phosphorylation of SET occurs ubiquitously, we subjected various human cell lines (HUT-102, HUT-78, MT-2, H-9, TL-Su, Jurkat, Raji, K-562, HL-60, HOS, and HeLa) to metabolic labeling with [^{32}P]orthophosphate followed by immunoprecipitation. As shown in Fig. 1, ^{32}P-labeled SET was immunoprecipitated in all cell lines tested. The incorporation of $^{32}P_i$ into SET showed a marked variation depending upon the cell line. Since the radioactivity associated with SET was highly sensitive to alkaline phosphatase treatment (data not shown) [3], incorporation of the radioactivity was due to protein phosphorylation, but not to ADP-ribosylation.

3.2. Phosphoamino acid analysis of phosphorylated SET

To identify the phosphorylated amino acids present in SET, ^{32}P-labeled SET from various human cell lines were transferred from the gel onto PVDF membrane and subjected to partial acid hydrolysis in 6 N HCl vapor at 110°C for 2 h. Phosphoamino acids were separated by

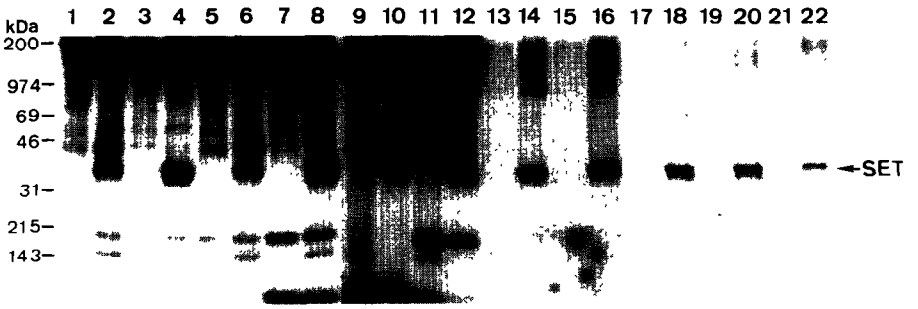

Fig. 1. In vivo phosphorylation of SET. Cells (2×10^6) were metabolically labeled with [^{32}P]orthophosphate for 2 h at 37°C, lysed, and subjected to immunoprecipitation with anti-SET IgG (even-numbered lanes) and preimmune IgG (odd-numbered lanes) as described in Section 2. The cell lines examined were HUT-102 (lanes 1 and 2), HUT-78 (lanes 3 and 4), MT-2 (lanes 5 and 6), H-9 (lanes 7 and 8), TL-Su (lanes 9 and 10), Jurkat (lanes 11 and 12), Raji (lanes 13 and 14), K-562 (lanes 15 and 16), HL-60 (lanes 17 and 18), HOS (lanes 19 and 20), and HeLa (lanes 21 and 22).

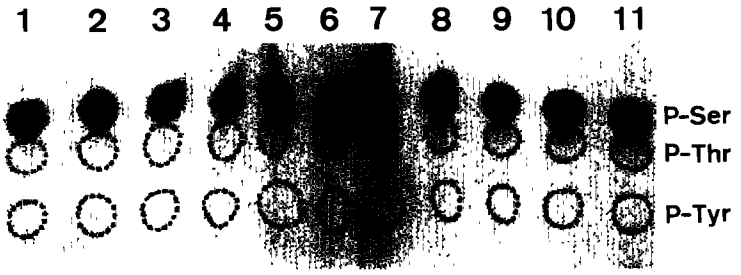

Fig. 2. Phosphoamino acid analysis of SET. ^{32}P-labeled SET from various human cell lines was immunoprecipitated, eluted from gels, and hydrolyzed in 6 N HCl vapor at 110°C for 2 h. Phosphoamino acids were separated by high-voltage electrophoresis and subjected to autoradiography. SET proteins were isolated from lane 1, HUT-102; lane 2, HUT-78; lane 3, MT-2; lane 4, H-9; lane 5, TL-Su; lane 6, Jurkat; lane 7, Raji; lane 8, K-562; lane 9, HL-60; lane 10, HOS; lane 11, HeLa. The positions of unlabeled phosphoserine (P-Ser), phosphothreonine (P-Thr), and phosphotyrosine (P-Tyr) are shown by broken circles.

high-voltage electrophoresis and detected by autoradiography. As shown in Fig. 2, in all cell lines tested, SET was phosphorylated mainly on serine residue(s) in vivo. Phosphothreonine and phosphotyrosine were not detectable in any of the cells examined.

3.3. *Identification of in vivo phosphorylation sites of SET*

To determine the precise location of the phosphorylated serine(s), we prepared ^{32}P-labeled SET by metabolic labeling followed by immunoprecipitation. After extensive digestion with endoproteinase Glu-C, the resulting peptides were separated by C_{18} reversed-phase HPLC (Fig. 3). Three peaks of ^{32}P-radioactivity were isolated. Peak 1 contained $^{32}P_i$. Each ^{32}P-labeled phosphopeptide from peaks 2 and 3 was mixed with myoglobin as carrier protein and subjected to automated Edman degradation in a spinning cup sequenator. Successive degradation and the sequence cycle number were monitored by analysis of phenylthiohydantoin amino acids of carrier protein released at cycles 2 and 11. The amount of radioactivity released during each cycle was monitored by scintillation counting to determine the position(s) of phosphorylated amino acids relative to the N terminus of the peptides (Fig. 4). ^{32}P-radioactivity was released at the second cycle of peak 2 (Fig. 4A). Only one predicted Glu-C peptide (residues 23–25) contained a serine in this position. In contrast, no released radioactivity was observed during 24 cycles in peak 3 (Fig. 4B). This result indicated that the N terminus of this Glu-C peptide was blocked by posttranslational modification such as acetylation and this peptide corresponded to the N-terminal residues 1–12 of SET.

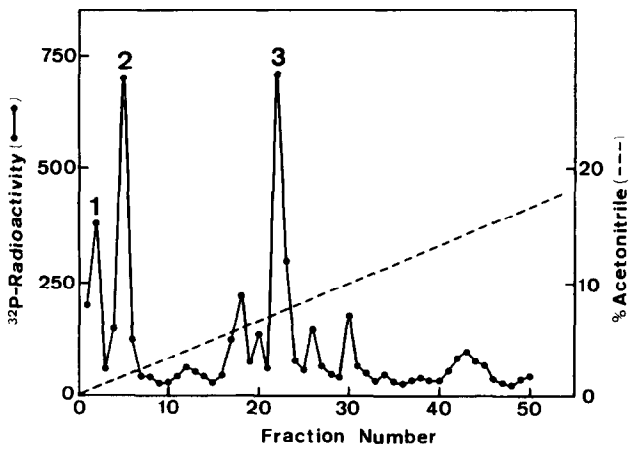

Fig. 3. Isolation of endoproteinase Glu-C-digested phospho-peptides by C_{18} reversed-phase HPLC. ^{32}P-Labeled SET was prepared by immunoprecipitation and SDS-PAGE, then digested with endoproteinase Glu-C. The resulting peptides were applied onto a C_{18} column (0.46 × 30 cm). The phosphopeptides were separated at a flow rate of 1.0 ml/ fraction/min with a linear gradient of 0–40% acetonitrile in 0.05% trifluoroacetic acid as shown by the dashed line. The radioactivity in each fraction was monitored by Cerenkov counting.

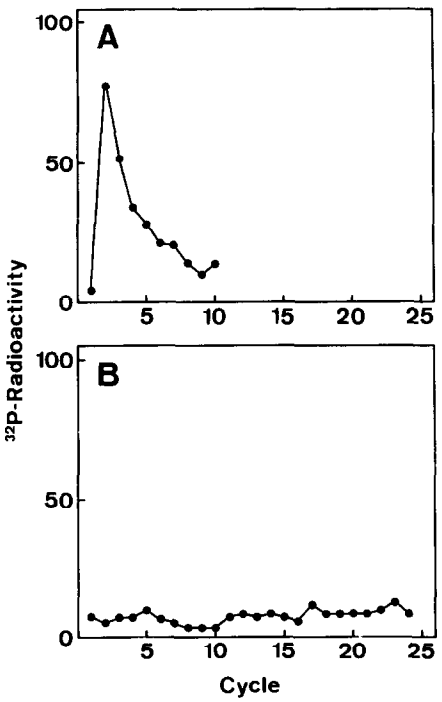

Fig. 4. Radiosequence analysis. Each Glu-C-digested phosphopeptide from reversed-phase HPLC was subjected to automated Edman degradation. The amounts of ^{32}P radioactivity released during each cycle are shown. (A) Peak 2 in Fig. 3; (B) Peak 3 in the same figure.

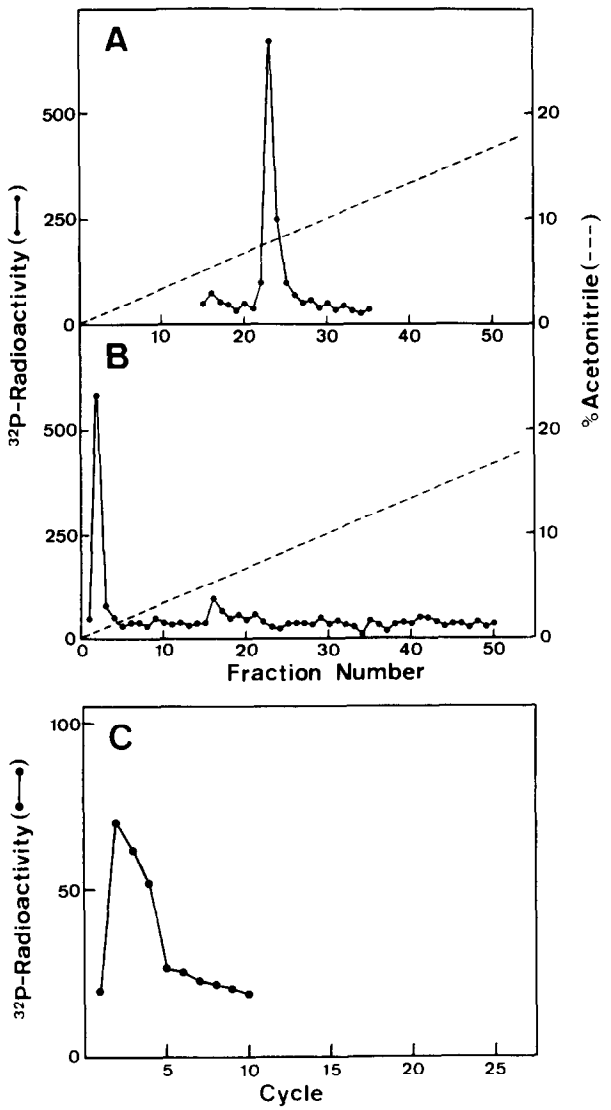

Fig. 5. Analysis of N-terminal Glu-C peptide by tryptic digestion and radiosequencing. (A) Phosphorylated N-terminal Glu-C peptide corresponding to peak 3 in Fig. 3. (B) The phosphopeptide was digested with TPCK-trypsin. The resulting tryptic peptides were separated by C_{18} reversed-phase HPLC as described in Fig. 3. (C) The phosphopeptide isolated in panel B was subjected to radiosequencing.

To identify the phosphorylation site(s) in the N-terminal Glu-C peptide (Figs. 3 and 5A), we subjected this peptide to a secondary digestion with TPCK-trypsin followed by HPLC and radiosequencing. After extensive tryptic digestion, a peak of ^{32}P-radioactivity was detected in the 0.67% acetonitrile fraction of C_{18} reversed-phase HPLC (Fig. 5B). This elution position indicated that the resulting phosphopeptide was highly hydrophilic and corresponded to VS*K (residues 8–10) rather than MS*AQAAK (residues 1–7). This result was also supported by relative retention times on the same reversed-phase HPLC column of the synthetic peptides AcMSAQAAKVSKKE, AcMSAQAAK, and VSK (data not shown). Finally, ^{32}P-labeled peptide in Fig. 5B was subjected to automated Edman degradation followed by scintillation counting. ^{32}P-radioactivity was released at the second cycle (Fig. 5C). This position corresponded to Ser^{9}. Similar analyses with chymotryptic digestion followed by HPLC and radiosequencing also revealed the same results (data not shown). Thus, Ser^{9} and Ser^{24} were identified as in vivo phosphorylation sites of SET (Fig. 6). Moreover, the surrounding sequences of Ser^{9} and Ser^{24} coincide with the consensus sequence of phosphorylation sites of protein kinase C.

4. Discussion

In this report, we examined in detail the in vivo phosphorylation of SET. Phosphorylation of SET was observed in all human cell lines tested (HUT-102, HUT-78, MT-2, H-9, TL-Su, Jurkat, Raji, K-562, HL-60, HOS, and HeLa), although the level of phosphorylation varied depending upon the cell line (Fig. 1). The in vivo phosphorylation of SET therefore occurs in various types of cells ubiquitously. In every case, the major phosphoamino acid of phosphorylated SET was phosphoserine, indicating that SET is a substrate for one or more cellular serine/threonine kinases. Tyrosine kinases are not involved, since phosphotyrosine was not detected.

By analysis of radiosequencing data, we identified in vivo phosphorylation sites of SET as Ser^{9} and Ser^{24} (Figs. 3–6). The surrounding sequences of Ser^{9} and Ser^{24} contain an apparent consensus site sequence for protein kinase C, Ser*-X-Arg/Lys [28]. A wide range of in vivo protein kinase C substrate proteins including nuclear regulatory proteins has been reported [11]. Hence, it is possible that protein kinase C may utilize the suitable substrate sequences on SET and phosphorylate SET in response to physiological stimuli.

Although the function(s) of SET is not known, SET is essential in the mechanism of leukemogenesis in AUL, particularly by activating the putative oncoprotein, CAN, in nuclei and stimulating the transformation po-

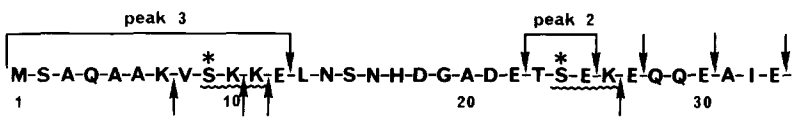

Fig. 6. In vivo phosphorylation sites of SET. Arrows mark the predicted sites of Glu-C (upper) and tryptic (lower) cleavage. Brackets indicate Glu-C-digested phosphopeptides, identified by the numbering system of Fig. 3. Asterisks indicate residues phosphorylated in vivo. The consensus sequences of the sites phosphorylated by PK-C are marked with wavy underlines.

tential of SET-CAN fusion protein [2,3]. It is known that many nuclear proteins, including proteins thought to be involved in gene expression, replication, chromatin conformation, and nuclear transport (e.g. CREB, c-Fos, c-Jun, p53, histone, B-23, and HTLV-I Rex), are also phosphorylated in vivo in response to physiological stimuli and their activities are regulated by phosphorylation [11–20]. By analogy with these nuclear regulatory proteins, the in vivo phosphorylation of SET at Ser^9 and Ser^{24} in nuclei may play a key role for its function by changing structural conformation or affinity for protein–protein or protein–DNA/RNA interaction. The findings described in this report should be helpful in further studying the function of SET and chimeric SET-CAN proteins in the mechanism of leukemogenesis in AUL.

Acknowledgments: We thank Pat Wesdock and Eva Majerova for peptide synthesis, Suzanne Specht for preparation of antibodies, and Anne Arthur for editorial assistance. This work was supported by the National Cancer Institute, Department of Health and Human Services, under Contract NO1-CO-74101 with ABL.

References

[1] von Lindern, M., Fornerod, M., van Baar, S., de Wit, T., Buijs, A. and Grosveld, G. (1992) Mol. Cell. Biol. 12, 1687–1697
[2] von Lindern, M., van Baar, S., Wiegant, J., Raap, A., Hagemeijer, A. and Grosveld, G. (1992) Mol. Cell. Biol. 12, 3346–3355
[3] Adachi, Y., Pavlakis, G.N. and Copeland, T.D. (1994) J. Biol. Chem. 69, 2258–2262
[4] Ishimi, I. and Kikuchi, A. (1991) J. Biol. Chem. 66, 7025–7029
[5] Shirakawa, H. and Yoshida, M. (1992) J. Biol. Chem. 67, 6641–6645
[6] Adachi, Y., Mizuno, S. and Yoshida, M. (1990) J. Chromatogr. 30, 39–46
[7] Adachi, Y., Copeland, T.D., Hatanaka, M. and Oroszlan, S. (1993) J. Biol. Chem. 68, 13930–13934
[8] Lapeyre, B., Bourbon, H. and Amalric, F. (1987) Proc. Natl. Acad. Sci. USA 84, 1472–1476
[9] Brent, R. and Ptashne, M. (1985) Cell 43, 729–736
[10] Hope, I.A. and Struhl, K. (1986) Cell 46, 885–894
[11] Nishizuka, Y. (1986) Science 33, 305–312
[12] Papavassiliou, A.G., Treier, M., Chavrier, C. and Bohmann, D. (1992) Science 58, 1941–1944
[13] Papavassiliou, A.G., Chavrier, C. and Bohmann, D. (1992) Proc. Natl. Acad. Sci. USA 9, 11562–11565
[14] Baudier, J., Delphin, C., Grunwald, D., Khochbin, S., and Lawrence, J.J. (1992) Proc. Natl. Acad. Sci. USA 89, 11627–11631
[15] Adachi, Y., Nosaka, T. and Hatanaka, M. (1990) Biochem. Biophys. Res. Commun. 169, 469–475
[16] Yamamoto, K.K., Gonzalez, G.A., Biggs, W.H. and Montminy, M.R. (1988) Nature 34, 494–498
[17] Nosaka, T., Siomi, H., Adachi, Y., Ishibashi, M., Kubota, S., Maki, M. and Hatanaka, M. (1989) Proc. Natl. Acad. Sci. USA 6, 9798–9802
[18] Chan, P.K., Aldrich, M., Cook, R.G. and Busch, H. (1986) J. Biol. Chem. 61, 1868–1872
[19] Patskan, G.J. and Baxter, C.S. (1985) J. Biol. Chem. 60, 12899–12903
[20] Curran, T., Miller, A.D., Zokas, L. and Verma, I.M. (1984) Cell 6, 259–268
[21] Adachi, Y., Kitahara-Ozawa, A., Sugamura, K., Lee, W.J., Yodoi, J., Maki, M., Murachi, T. and Hatanaka, M. (1992) J. Biol. Chem. 67, 19373–19378
[22] Gazdar, A.F., Corney, D.N., Bunn, P.A., Russell, E.K., Jaffe, E.S., Schechter, G.P. and Guccion, J.G. (1980) Blood 55, 409–417
[23] Popovic, M., Sarngadharan, M.G., Read, E. and Gallo, R.C. (1984) Science 24, 497–500
[24] Adachi, Y., Takano, E., Murachi, T. and Hatanaka, M. (1988) Biol. Chem. Hoppe-Seyler 69, 223–227
[25] Adachi, Y., Ishida-Takahashi, A., Takahashi, C., Takano, E., Murachi, T. and Hatanaka, M. (1991) J. Biol. Chem. 266, 3968–3972
[26] Adachi, Y., Copeland, T.D., Takahashi, C., Nosaka, T., Ahmed, A., Oroszlan, S. and Hatanaka, M. (1992) J. Biol. Chem. 67, 21977–21981
[27] Laemmli, U.K. (1970) Nature 27, 680–685
[28] Kishimoto, A., Nishiyama, K., Nakanishi, H., Uratsuji, Y., Nomura, H., Takeyama, Y. and Nishizuka, Y. (1985) J. Biol. Chem. 60, 12492–12499