A Homologue of the *Mycobacterium tuberculosis* PapA5 Protein, Rif-Orf20, Is an Acetyltransferase Involved in the Biosynthesis of Antitubercular Drug Rifamycin B by *Amycolatopsis mediterranei* S699

Yeping Xiong,[a] Xiumei Wu,[b] and Taifo Mahmud*[a, b]

Since its discovery in the late 1950s, rifamycin has been playing significant roles in combating infectious diseases.[1] Its synthetically modified derivatives remain the principal chemotherapeutical agents used for the treatment of tuberculosis, leprosy, and AIDS-related mycobacterial infections. The potent antibacterial activity of this class of antibiotics is due to their specific inhibition of bacterial DNA-dependent RNA polymerases.[2] However, many strains of *Mycobacterium tuberculosis* have, over time, developed resistance to rifamycin, most of which is due to mutational alterations of the target molecule, the β-subunit of RNA polymerase. This high-level resistance contributes to the rise of tuberculosis cases and increase in the death toll, thus calling for the discovery and development of new antitubercular drugs.

The biosynthesis of rifamycin has been studied extensively by classical isotope-labeled feeding experiments and by contemporary genetic and biochemical approaches.[3–5] The polyketide framework of rifamycin B is assembled from 3-amino-5-hydroxybenzoic acid (AHBA), two molecules of acetate, and eight molecules of propionate. Proansamycin X (Scheme 1) has so far been proposed to be the earliest macrocyclic intermediate in the biosynthesis of rifamycin B. The bioconversion of proansamycin X to rifamycin B involves a series of tailoring reactions, for example, an oxidative rearrangement of the olefinic moiety to a ketal structure; an oxidative removal of C-34a; followed by attachments of side-chain moieties to C-4, C-25, and C-27 as the final decorations of the molecule. Our recent study provides evidence that suggest that acetylation of C-25 takes place prior to methylation of C-27.[6] An *S*-adenosylmethionine-dependent methyltransferase (Rif-Orf14) has been found to be responsible for the methylation reaction, converting 27-*O*-demethylrifamycin SV (DMRSV) to rifamycin SV. However, no obvious candidate gene to encode an acetyltransferase could be located in the 96 kb rifamycin biosynthetic gene cluster.[4] Very recently, Quadri and co-workers[7] reported an acyltransferase activity of *M. tuberculosis* PapA5, a so-called polyketide-associated protein, which had a previously unclear function. The protein was proposed to catalyze diesterification of phthiocerol and phthiodiolone with mycocerosate to give dimycocerosate esters, the components of complex virulence-enhancing lipids produced by *M. tuberculosis*. A homologous gene (*rif-orf20*) was found in the rifamycin gene cluster of *Amycolatopsis mediterranei* S699;[4] hence, it was proposed to be the acetyltransferase of the rifamycin pathway.[7]

To determine whether *rif-orf20* is involved in the acetylation process in rifamycin B biosynthesis by *A. mediterranei* S699, the gene was cloned into pRSET-B (Invitrogen) and recombinantly expressed as a 6XHis-tagged fusion protein in *E. coli* BL21-(DE3)pLysS (Stratagene) under the control of T7 promoter. Recombinant Rif-Orf20 was purified by using a Ni-NTA column and resolved as a 47.2 kDa protein by SDS-PAGE (Figure 1 A). Western blot analysis with anti-Histag antibody (Qiagen) gave distinct signals; this further indicated the over-expression of the his-tagged protein (Figure 1 B). However, due to a relatively low recovery of the soluble his-tagged protein after purification with Ni-NTA columns, enzymatic reactions were mostly carried out with cell-free extracts, with acetyl-CoA (5 mM), and 27-*O*-demethyl-25-*O*-desacetylrifamycin SV (DMDARSV; 2.5 mM) as substrates. DMDARSV was synthesized from DMRSV, an intermediate previously isolated from a rif-*orf14*-inactivated mutant strain of *A. mediterranei* S699.[6] DMRSV was treated with 20 volumes of 5% sodium hydroxide solution in aqueous ethanol (1:1) at room temperature to yield DMDARSV.[8] Analysis of the enzymatic reaction product by ESI-MS and ^{1}H NMR revealed that Rif-Orf20 catalyzes the conversion of DMDARSV to DMRSV (Figure 2 B). Incubation of DMDARSV and acetyl-CoA with cell-free extracts of *E. coli* BL21(DE3)pLysS harboring empty pRSET-B vector as well as incubation of DMDARSV and acetyl-CoA with boiled enzyme did not give DMRSV. It is noteworthy that incubation of DMDARSV with cell-free extract containing Rif-Orf20 in the absent of external acetyl-CoA also gave DMRSV. However, this was found to be due to the presence of acetyl-CoA in the cell-free extracts of *E. coli*. Reactions carried out on dialyzed cell-free extracts in the absent of acetyl-CoA did not give any product.

[a] Y. Xiong, Prof. T. Mahmud
College of Pharmacy, Oregon State University
Corvallis, OR 97331-3507 (USA)
Fax: (+1) 541-737-3999
E-mail: taifo.mahmud@oregonstate.edu

[b] X. Wu, Prof. T. Mahmud
Genetics Program, College of Agricultural Sciences
Oregon State University, Corvallis, OR 97331-2212 (USA)

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/cbic.200400387 ChemBioChem 2005, 6, 834–837

Scheme 1. *Proposed biosynthetic pathway to rifamycin B.*

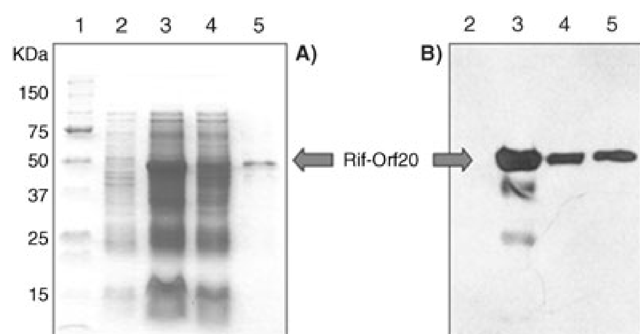

Figure 1. *A) SDS-PAGE and B) Western blot of over-expressed His-tagged Rif-Orf20. 1: protein ladder; 2: cell-free extract of* E. coli/pRSET-B *after IPTG induction; 3: total protein of* E. coli/rif-orf20 *after IPTG induction; 4: cell-free extract of* E. coli/rif-orf20 *after IPTG induction; 5: His-tagged Rif-Orf20 after Ni-NTA column purification.*

In contrast to the PapA5 of *M. tuberculosis*, which recognizes long-chain acyl-CoA thioesters as substrate but does not accept acetyl-CoA or other short-chain acyl-CoA thioesters, Rif-Orf20 efficiently utilizes acetyl-CoA. To gain further insight into Rif-Orf20 substrate selectivity, we explored the enzyme's ability to use different acyl-CoA thioesters. Thus, propionyl-CoA, butyryl-CoA, isobutyryl-CoA (branched-chain), palmitoyl-CoA (long-chain), and benzoyl-CoA (aromatic) were tested. All CoA esters were individually incubated with Rif-Orf20 in the presence of DMDARSV. Analysis of the reaction products by ESI-MS showed that the enzyme utilized propionyl-CoA as substrate, albeit slightly less efficiently (~80%) than acetyl-CoA, to give rise to a propionyl ester derivative of DMDARSV (Figure 2C). No product could be detected in enzyme incubations with the other CoA esters.

The propionyl-CoA incubation product has a molecular mass of 697 Da, 56 amu more than DMDARSV and 14 amu more than DMRSV, and corresponds to a propionylated derivative of DMDARSV. The propionyl group is most likely attached to 25-OH, as this group is the natural acylation site for Rif-Orf20. This was supported by comparisons of its MS/MS fragmentation pattern (*m/z* 696→622→604→394) with those of DMRSV (*m/z* 682→622→604→394) and DMDARSV (*m/z* 640→622→604→394). The ^{1}H NMR spectrum of the propionylated DMDARSV is nearly identical with that of DMRSV, except for the lack of acetyl protons (at 2.08 ppm in that of DMRSV) and the presence of propionyl protons [at 0.63 ppm (t, $J=7.6$ Hz, 3H) and 2.18 ppm (q, $J=7.6$ Hz, 2H)].

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

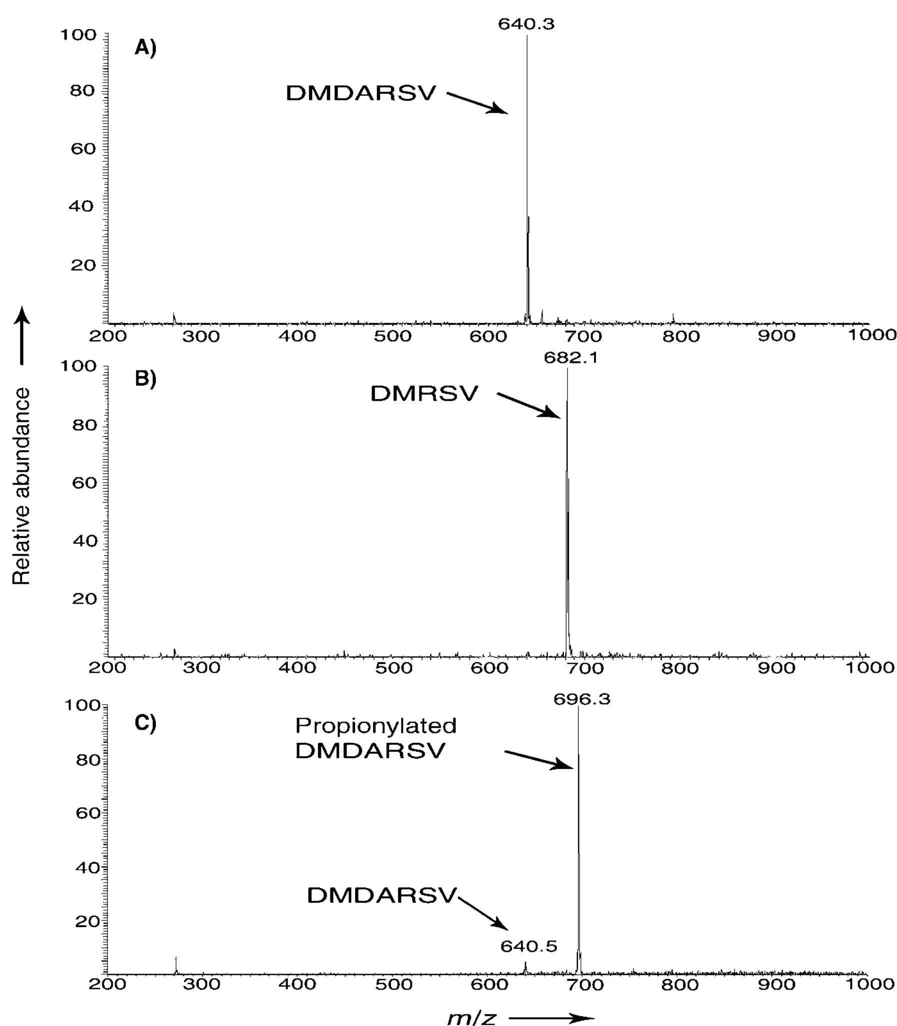

Figure 2. *A) Negative-ion ESI-MS of DMDARSV, B) incubation products of Rif-Orf20 with DMDARSV and acetyl-CoA, and C) incubation products of Rif-Orf20 with DMDARSV and propionyl-CoA.*

This study revealed that Rif-Orf20, a homologue of *M. tuberculosis* PapA5 protein, is an acetyltransferase that is responsible for the conversion of DMDARSV to DMRSV in rifamycin B biosynthesis. This enzyme, together with PapA5, represents a new class of acyltransferases.[9] Rif-Orf20 also utilizes propionyl-CoA as substrate to give rise to a new analogue of rifamycin. However, in contrast to the PapA5 of *M. tuberculosis*, it does not recognize long chain fatty acids as substrate; this suggests that both enzymes have distinct substrate recognition domains.

Experimental Section

Instruments and chemicals: The ^{1}H and ^{13}C NMR spectra were recorded on a Bruker DPX400 FT NMR spectrometer. Low-resolution mass spectra were recorded on a ThermoFinnigan LCQ Advantage (electrospray and atmospheric pressure chemical ionizations) Liquid Chromatograph–Ion Trap Mass Spectrometer. A MaxQ 4000 shaker, Barnstead/Lab-Line, was used for the fermentation. A Bio-Rad SmartSpec 3000 was used for UV measurement of DNA and protein. Chemical reactions were monitored by TLC (silica gel 60 F_{254}, Merck) with detection by UV light or by $Ce(SO_4)_2/H_2SO_4$ solutions. Column chromatography was performed on 230–400 mesh silica gel (Aldrich). For HPLC, a Beckman System Gold Programmable Solvent Module was used with a Beckman System Gold Detector. All chemicals were purchased from EMD Chemicals, Aldrich, Sigma, or Bio-Rad and used without further purification unless otherwise noted. Acetyl-CoA, propionyl-CoA, butyryl-CoA, isobutyryl-CoA, and benzoyl-CoA were purchased from Sigma.

Microorganisms, culture conditions and vectors: The DMRSV-producing strain was generated from *A. mediterranei* S699 by frame-shift inactivation of *rif-orf14*.[6] *E. coli* XL1-Blue (Stratagene) was used as host for subcloning, and *E. coli* BL21-(DE3)pLysS (Stratagene) was used for gene expression. pRSET-B (Invitrogen) was also used for gene expression. *E. coli* strains were grown in LB[10] or LB containing betaine (2.5 mM) and sorbitol (1 M; LBBS)[11] media supplemented with ampicillin (100 μg mL^{-1}) or ampicillin (100 μg mL^{-1})/chloramphenicol (25 μg mL^{-1}) for selection of plasmids.

DNA manipulation: Routine genetic procedures, such as plasmid DNA isolations, restriction endonuclease digestions, alkaline phosphatase treatments, DNA ligations, and other DNA manipulations, were performed according to standard techniques.[10] DNA fragments were excised from agarose gels and residual agarose was removed with the QiaQuick Gel Extraction Kit (Qiagen). PCR was carried out by using Pfx DNA polymerase (Invitrogen) according to the manufacturer's protocol. Sequencing was performed by using the Big Dye RR terminator cycle sequencing kit (PerkinElmer Biosystems), and the gels were run on ABI-3730 sequencers.

Synthesis of DMDARSV: DMRSV (40 mg, 58 mmol) was dissolved in NaOH solution (5 %) in aqueous ethanol (1:1; 800 μL) and stirred for 30 min at room temperature. The reaction mixture was diluted with ice water (1.6 mL), and HCl solution (0.1 N) was added dropwise until the mixture reached pH 3. The product was extracted with ethyl acetate, and the organic solvent was then evaporated. The crude product was subjected to silica gel column chromatography (SiO_2 10 g, EtOAc) to give DMDARSV (32 mg, 47 mmol, 81 %). ^{1}H NMR (D_2O, 400 MHz, 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid, sodium salt (TMSP) was used as internal standard): $\delta = -0.03$ (d, $J = 6.5$ Hz, 3 H, H-34), 0.54 (d, $J = 6.5$ Hz, 3 H, H-33), 0.83 (d, $J = 7$ Hz, 3 H, H-32),[a] 0.90 (d, $J = 7$ Hz, 3 H, H-31),[a] 1.23 (m, 1 H, H-26), 1.45 (m, 1 H, H-24), 1.64 (s, 3 H, H-13), 1.83 (m, 1 H, H-22), 1.96 (s, 3 H, H-

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

14),[b] 2.02 (s, 3H, H-30),[b] 2.25 (m, 1H, H-20), 3.38 (d, $J=11$ Hz, 1H, H-23),[c] 3.69 (d, $J=10.5$ Hz, 1H, H-21),[c] 3.81 (d, $J=9$ Hz, 1H, H-25)[c], 4.34 (m, 1H, H-27), 5.22 (dd, $J=5.5$, 12.5 Hz, 1H, H-28), 6.00 (dd, $J=7$, 15.5 Hz, 1H, H-19), 6.11–6.23 (m, 3H, H-18, H-17, H-29), 7.46 (s, 1H, H-3); ESI-MS *m/z*: 640 $[M-H]^-$. [a], [b] These assignments may be interchanged. [c] Appeared as a doublet because one of the vicinal interproton coupling constants is ~0 Hz.[12]

Gene expression: *Rif-orf20* was amplified from the template pTM1303[6] by using the primers 5′-GAAGATCTTGTGACGCGAGCGGCTGTCCGGCG-3′ (forward) and 5′-GGAATTCTCACAGGTCACGAGTCAGGCGGGCC-3′ (reverse), with a *Bgl*II restriction site introduced into the forward primer and an *Eco*RI site introduced into the reverse primer to facilitate subcloning. PCR reactions were performed in a thermocycler (Eppendorf) under the following conditions: 33 cycles of 30 s at 96 °C, 45 s at 60 °C, and 45 s at 72 °C. The PCR product was purified and digested with *Bgl*II/*Eco*RI, and the resulting 1.12 kb fragment was cloned into pRSET-B and amplified in *E. coli* XL1Blue. The construct was then introduced into *E. coli* BL21-(DE3)pLysS by heat-pulse transformation, and ampicillin/chloramphenicol-resistant transformants were selected. *E. coli* BL21-(DE3)pLysS was grown in LB medium (3 mL) containing ampicillin and chloramphenicol at 37 °C and 300 rpm for 16 h. The culture was then transferred to LBBS medium (20 mL) and incubated at 37 °C and 300 rpm. At an OD_{600} of 0.5, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM, and the incubation was continued at 25 °C and 200 rpm. After 24 h, the cells were harvested by centrifugation at 6500 rpm for 5 min and stored at −80 °C until further use.

Preparation of cell-free extracts containing His_6-tagged recombinant Rif-Orf 20 protein: Tris-HCl buffer pH 7.5 (10 mM, 1 mL) containing dithiothreitol (DTT, 1 mM) and phenylmethylsulfonyl fluoride (PMSF, 1 mM) was added to the thawed cell pellet from 20 mL culture. The mixture was agitated in a Vortex mixer and incubated on ice for 30 min. The suspension was then sonicated at 5 W (3× 30 s), and the cell debris was removed by centrifugation at 10 000 rpm for 10 min. Protein concentration was measured by the Bradford protein microassay with bovine serum albumin as standard. An aliquot of the cell-free extract (0.5 mL) was subjected to dialysis over a 6–8000 molecular weight cut-off membrane (Spectra/Por) in Tris-HCl buffer pH 7.5 (200 mL, 10 mM) containing DTT (1 mM) and PMSF (1 mM) at 4 °C overnight.

SDS-PAGE and Western analysis: For SDS-PAGE, protein solution (10 μL, 80 μg of proteins) was mixed with Laemmli sample buffer (Bio-Rad; 20 μL) containing β-mercaptoethanol (5%) and boiled for 10 min before being loaded onto SDS-PAGE gel. Proteins resolved by SDS-PAGE were electrotransferred to a polyvinylidene difluoride membrane (Millipore, Bedford, MA) [transfer buffer:[13] Tris (14.26 g), glycine (67.82 g), and methanol (15%) in 4 L total volume] overnight at 100 mA. The membrane was blocked with non-fat dry milk (NFDM, 5% *w/v*) in TTBS buffer [Tris-HCl (pH 7.5, 100 mM), NaCl (150 mM), Tween-20 (0.1%)] and incubated with a 1:1000 dilution of primary anti-histag antibody (Qiagen) in NFDM (1%) in TTBS for 2 h at room temperature. After being washing with TTBS buffer, the blot was incubated with a 1:7500 dilution of secondary antibody (peroxidase-conjugated goat anti-mouse IgG (Pierce)) for 45 min at room temperature and then washed 3× with TTBS buffer. The blot was then treated with enhanced chemiluminescence solution (luminol, *p*-coumaric acid and peroxide in Tris buffer pH 8.5 (100 mM)) for 1 min and exposed to X-ray film.

Enzyme assays: The enzymatic reaction mixture for determining the activity of Rif-Orf20 contained DMDARSV (2.5 mM), cell-free extract (50 μL, 0.4 mg of proteins), Tris-HCl buffer pH 7.5 (50 mM), and acetyl-CoA (5 mM) in a total volume of 100 μL. The mixture was incubated at 30 °C for 1 h, acidified with HCl solution (0.1 N, 2.5 μL), and extracted with ethyl acetate (100 μL). The ethyl acetate solution was used directly for mass spectral analysis. Incubations to determine the substrate specificity of Rif-Orf20 contained the same components as above with variables in the use of different acyl-CoA thioesters. A scaled-up enzymatic reaction was carried out with DMDARSV (2.5 M), cell-free extract (1 mL, 8 mg of proteins), Tris-HCl buffer (pH 7.5, 10 mM), and acetyl-CoA (5 M) in a total volume of 2 mL. After 2 h incubation under the same conditions described above, the mixture was worked up to give DMRSV. ^{1}H NMR (D_2O, 400 MHz, 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid, sodium salt (TMSP) was used as internal standard): $\delta=-0.20$ (d, $J=6.5$ Hz, 3H, H-34), 0.70 (d, $J=6.5$ Hz, 3H, H-33), 0.96 (br d, $J\approx6.5$ Hz, 6H, H-31, H-32), 1.21–1.40 (m, 2H, H-24, H-26), 1.78 (s, 3H, H-13), 1.92 (m, 1H, H-22), 2.03 (br s, 6H, H-14, H-30), 2.08 (s, 3H, H-36), 2.38 (m, 1H, H-20), 3.15 (d, $J=10$ Hz, 1H, H-23),[a] 3.83 (d, $J=10$ Hz, 1H, H-21),[a] 3.95 (br d, $J\approx8$ Hz, 1H, H-27), 5.02 (d, $J=10.5$ Hz, 1H, H-25),[a] 5.35 (dd, $J=8$, 12.5 Hz, 1H, H-28), 6.20 (dd, $J=7.5$, 15.5 Hz, 1H, H-19), 6.35 (m, 2H, H-29, H-17), 6.53 (dd, $J=11$, 15.5 Hz, 1H, H-18), 7.20 (s, 1H, H-3). ESI-MS *m/z*: 682 $[M-H]^-$. [a] Appeared as a doublet because one of the vicinal interproton coupling constants is ~0 Hz.[12]

Acknowledgements

The work was supported by a grant from the Medical Research Foundation of Oregon. DNA sequencing was performed by the Center for Gene Research and Biotechnology (CGRB) at Oregon State University. The authors thank Patricia M. Flatt for providing technical assistance and for critical reading of the manuscript.

Keywords: acyltransferase · antibiotics · biosynthesis · *Mycobacterium tuberculosis* · rifamycin

[1] W. Wehrli, *Top Curr. Chem.* **1977**, *72*, 21.
[2] E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S. A. Darst, *Cell* **2001**, *104*, 901.
[3] O. Ghisalba, J. Nuesch, *J. Antibiot.* **1981**, *34*, 64.
[4] P. R. August, L. Tang, Y. J. Yoon, S. Ning, R. Muller, T. W. Yu, M. Taylor, D. Hoffmann, C. G. Kim, X. Zhang, C. R. Hutchinson, H. G. Floss, *Chem. Biol.* **1998**, *5*, 69.
[5] T. Schupp, C. Toupet, N. Engel, S. Goff, *FEMS Microbiol. Lett.* **1998**, *159*, 201.
[6] J. Xu, T. Mahmud, H. G. Floss, *Arch. Biochem. Biophys.* **2003**, *411*, 277.
[7] K. C. Onwueme, J. A. Ferreras, J. Buglino, C. D. Lima, L. E. Quadri, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4608.
[8] N. Maggi, A. Vigevani, R. Pallanza, *Experientia* **1968**, *24*, 209.
[9] J. Buglino, K. C. Onwueme, J. A. Ferreras, L. E. Quadri, C. D. Lima, *J. Biol. Chem.* **2004**, *279*, 30 634.
[10] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York, **1989**.
[11] C. G. Kim, T. W. Yu, C. B. Fryhle, S. Handa, H. G. Floss, *J. Biol. Chem.* **1998**, *273*, 6030.
[12] A. Bacchi, G. Pelizzi, M. Nebuloni, P. Ferrari, *J. Med. Chem.* **1998**, *41*, 2319.
[13] J. Towbin, T. Staehelin, J. Gordon, *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 4350.

Received: November 2, 2004
Published online on March 24, 2005

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim