A New Chemical Probe for Proteomics of Carbohydrate-Binding Proteins

Lluis Ballell,[a] Kirstin J. Alink,[a] Monique Slijper,[b] Cees Versluis,[b] Rob M. J. Liskamp,[a] and Roland J. Pieters*[a]

One of the main consequences of the completion of the human and other genome-sequencing projects has been the fact that the challenge in biological research has shifted towards the assignment of functions to the numerous gene products. The field of proteomics contributes to this process by developing and applying novel methods for the global analysis of protein expression in biological samples of high complexity.[1] Proteomics research typically starts by reducing the sample complexity through multidimensional separation methods based on the unique characteristics of the proteins or their peptides, such as isoelectric point, molecular mass or hydrophobicity, followed by identification and quantification by mass spectrometry. This methodology, however, has some disadvantages as it may overlook, for example, low-abundance or extremely lipophilic proteins that are physiologically important. Moreover, these methods are based on the detection of protein abundance rather than protein activity, whereas numerous forms of post-translational regulation[2] take place that turn the proteins into their active forms.

Only recently new synthetic molecular probes have been developed that are capable of selectively capturing enzymes by means of their catalytic activity.[3] This "activity-based" approach holds promise for the study of enzyme families for which irreversible inhibitors can be designed, but cannot be applied to proteins devoid of catalytic activity. Amongst those, the lectins are a particularly intriguing group. These proteins recognise carbohydrate motifs and are involved in multiple important cellular-recognition processes.[4] Considering their importance, we envisaged the need for developing new synthetic molecular probes for the functional study of lectins in biological systems. We here report on the development of a new chemical probe that allows the covalent capture and subsequent analysis of members of an important lectin family: the galectins.[5] Galectins are a subfamily of lectins known to bind monovalent β-galactoside-containing motifs. Around 14 mem-

[a] *Dr. L. Ballell, K. J. Alink, Prof. Dr. R. M. J. Liskamp, Dr. R. J. Pieters*
Department of Medicinal Chemistry
Utrecht Institute for Pharmaceutical Sciences, Utrecht University
P.O. Box 80082, 3508 TB, Utrecht (The Netherlands)
Fax: (+31) 30-253-6655
E-mail: r.j.pieters@pharm.uu.nl.

[b] *Dr. M. Slijper, C. Versluis*
Department of Biomolecular Mass Spectrometry
Utrecht Institute for Pharmaceutical Sciences
and Bijvoet Center for Biomolecular Research, Utrecht University
Sorbonnelaan 16, 3584 CA, Utrecht (The Netherlands)

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

 DOI: 10.1002/cbic.200400209

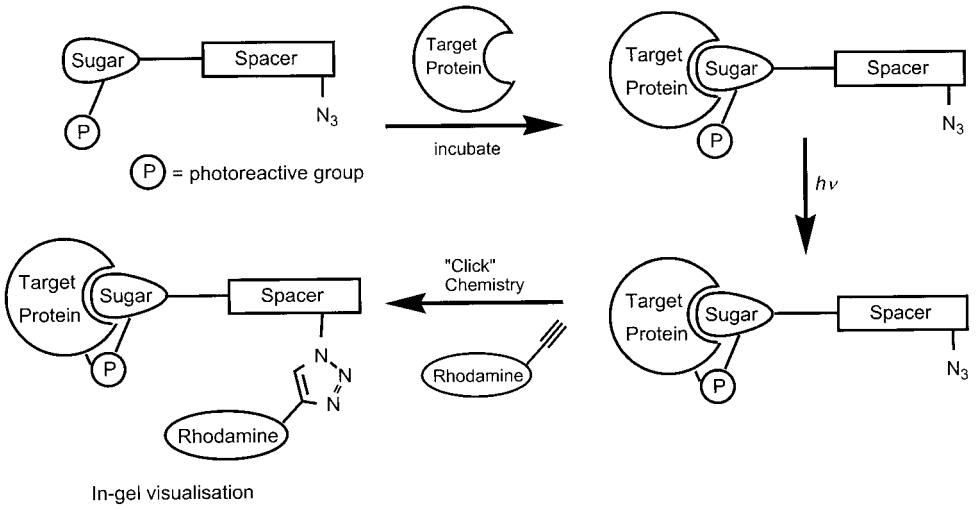
bers of the family have been characterised in mammals and they are known to be implicated in many different processes such as development, differentiation, cell–cell adhesion, cell–matrix interaction, growth regulation, apoptosis, RNA splicing and tumour metastasis.[6,7] Gaining insight into the levels of galectins in biological samples is highly relevant as their levels are linked to, for example, malignancy.[8] In a neuroblastoma model, galectin-1 induced tumour cell apoptosis, while galectin-3 blocked the effect.[9]

The probes designed in our lab are based on the combined application of three elements: 1) a ligand containing a photoaffinity label,[10] 2) a chemoselective ligation motif[11] and 3) a chemical tag for visualisation (Scheme 1). Several examples exist in the literature in which some of these elements have been successfully employed.[12] A chemoselective ligation method proved useful in in vivo enzyme profiling without the disturbance of a bulky fluorescent tag,[12e] and a photoaffinity label was shown to capture kinases.[12b] To our knowledge this is the first report in which all three of the mentioned elements are employed together to generate a new chemical probe.

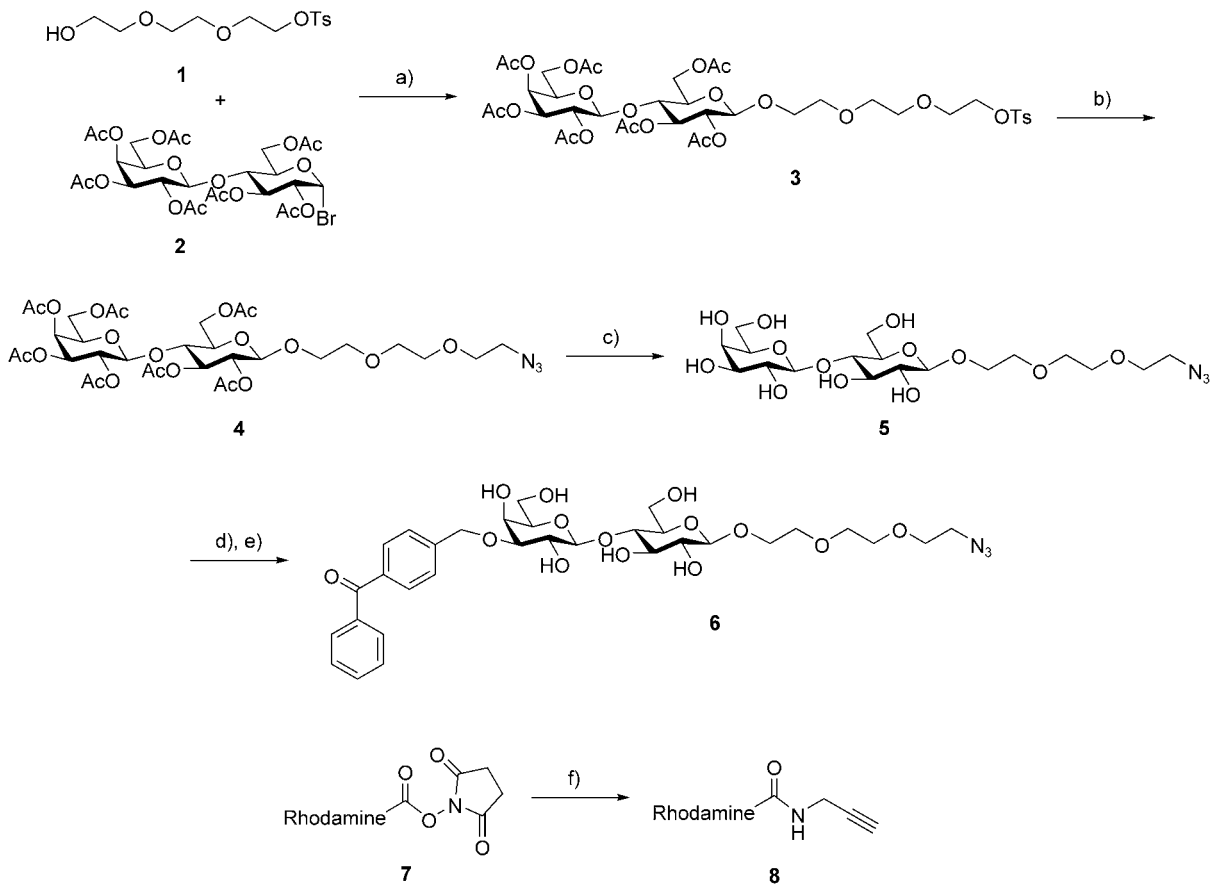
The design of the probe was based on available crystal structures. Close inspection of the Lac and LacNAc complexes of galectin-1[13] and galectin-3[14] revealed the possibility of elaborating the 3-OH position of the galactose unit without interference with the binding of the sugar. While the 4′-OH is engaged in hydrogen bonding, substituents at the 3′-OH could run parallel to the surface of the protein. The orientation of the sugar is conserved in both crystal structures, and elaboration of the 3′-OH has already been exploited to synthesise galectin-3 inhibitors with enhanced activity.[15] We thus decided to position the photoreactive group at the 3′-OH of the lactoside. Benzophenone was chosen, since it can be activated in a reversible manner, it does not react with protic solvents and it only modifies C–H bonds within 3 Å of the carbonyl oxygen.[16] Due to the highly lipophilic nature of the benzophenone substituent and in order to minimise the overall hydrophobic character of the probe, a water-solubilising linker containing ethylene glycol units was incorporated. Finally, an azido functionality was included with the purpose of taking advantage of the highly efficient 1,3-dipolar chemoselective copper-catalysed cycloaddition between azides and alkynes, denoted as "click" chemistry.[11] All of these features are present in compound **6**.

The synthesis of **6** (Scheme 2) started with the mono-tosylation of commercially available tri(ethylene glycol) to yield the mono-substituted spacer **1**.[17] Glycosylation of **1** with hepta-*O*-acetyl-lactosyl bromide **2**[18] under IBr-promoted conditions gave only the β-*O*-glycosylated product **3.** After displacement of the tosyl group with sodium azide to give **4**,[19] the acetates were fully removed to yield the deprotected lactoside **5**.[19] Microwave-assisted, tin-mediated regioselective 3-*O*-alkylation[20] with 4-(bromomethyl)benzophenone gave the desired 3′-*O*-substituted photoprobe **6** in high yield. The alkyne–rhodamine conjugate **8**, for in-gel visualisation of trapped proteins, was prepared from commercially available succinimidyl ester **7.**

In order to asses the effect of photoactivation on the protein, the probe was irradiated in the presence of human galectin-1 (360 nm, 30 min), and the ability of the isolated protein (after ultrafiltration) to bind a carbohydrate chip surface was measured. A 40% decrease in the equilibrium SPR signal was observed when compared to unmodified protein. To unequivocally establish the covalent attachment of the probe to the target protein, galectin-1 was irradiated in the presence of a 50-fold excess of probe **6** (2 mM), and the resulting mixture was ultrafiltrated (cut-off 5 kDa) to eliminate excess labelling reagent. The filtrate was diluted tenfold in 0.1% (*v*/*v*) acetic acid and 50% (*v*/*v*) acetonitrile in H_2O, then analysed by ESI-MS (Figure 1). MS data showed that around 70–80% of the protein was labelled with **6**. The 14.8 kDa galectin-1 peaks were always accompanied by peaks corresponding to the addition of one probe molecule **6** (Figure 1). Under these conditions, the possibility of formation of noncovalent adducts can be excluded. Repeating the experiment with a 20-fold excess of competing lactose (with respect to **6**) resulted in complete prevention of covalent attachment of the probe to the protein;



Scheme 1. *General outline for the recognition-based selective tagging of proteins. The initial binding event is followed by covalent capturing through photoactivation and subsequent chemoselective ligation for in-gel visualisation.*



Scheme 2. *Synthesis of chemical probe* ***6*** *and rhodamine derivative* ***8****. Reagents and conditions: a) IBr, DCM/CH_3CN, 61%; b) NaN_3, DMF, reflux, 95%; c) NaOMe, MeOH, 93%; d) Bu_2SnO, Benzene/CH_3CN (5:1), μW, 5 min., 150°C, followed by e) 4-(Bromomethyl)benzophenone, Bu_4NI, Benzene/CH_3CN (5:1), μW 10 min., 150°C, 71%; f) propargyl bromide·HCl, iPr_2NEt, CH_2Cl_2, 94%.*

this clearly showed that binding-site recognition was necessary for the covalent attachment of the probe to the protein (data not shown).

After validating the photoinduced covalent attachment of **6** to galectin-1, we focused on the "bio-orthogonal" (inert to chemical functionalities present in common biological samples) chemoselective attachment of the labelled protein to a reporter group that would allow visualisation by attachment of a rhodamine dye to the targeted protein. Our choice was the copper-catalysed 1,3-dipolar cycloaddition "click" reaction between azides and alkynes.[11] The azido functionalised probe **6** was treated with the alkyne-containing rhodamine derivative **8** in the presence of $CuSO_4$, a tris(triazolyl)amine ligand,[21] and copper wire for 1 h in phosphate buffer (pH 8); this resulted in the total consumption of alkyne **6** to generate the triazolyl-containing adduct. Once optimal conditions for the "click" reaction were determined, we proceeded to analyse its potential application in the presence of the azide-containing labelled protein. Thus galectin-1 was incubated with probe **6** under 360 nm light, and the resulting mixture was ultrafiltrated to eliminate excess labelling reagent. The protein was treated with the alkyne-derivatised rhodamine **8** (2 mM, twofold excess with respect to probe), and, after ultrafiltration, the solution was run on an 18% SDS-PAGE gel. Visualisation by a laser-induced fluorescence scanner produced the image depicted in Figure 2 (lane 2). After visualisation, the same gel was immediately stained with Coomassie blue to reveal both labelled and unlabelled protein. It could be clearly observed that the rhodamine tag had been incorporated into galectin-1. In order to rule out unspecific tagging of the protein, the same sample preparation was repeated with denatured protein. As shown in Figure 2 (lane 1), no labelling was observed in that case; this confirmed that specific binding is required for galectin-1 labelling.

In order to further support the carbohydrate binding-site dependent nature of the labelling, galectin-1 and galectin-3 together with cytochrome C and glyceraldehyde-3-phosphate dehydrogenase were subjected to the same photolabelling and tag-attachment routine as described before to generate the image observed in Figure 3. Selective labelling of the galactose-recognising proteins was observed.

In conclusion, we have described the synthesis and validation of a new chemical probe for the study of galectins based on the combined application of photoaffinity labelling, chemoselective ligation and affinity-tag visualisation. Placement of the photoaffinity label was based on the crystal structures of galectins and proved to be a crucial factor, since attachment of the label to the reducing end of lactose (see Supporting Infor-

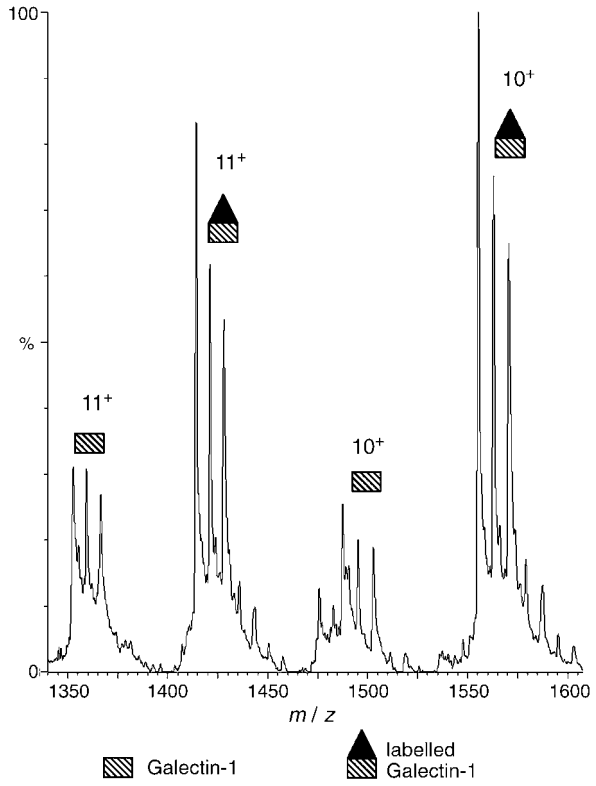


Figure 1. *Section of an ESI-MS spectrum for labelling experiments. Galectin-1 was incubated in phosphate buffer with* ***6*** *under 360 nm light for 45 min. After ultrafiltration, two major sets of triple peaks are observed (even pure galectin-1 was always observed as triple peaks since the protein was not 100% homogeneous). The mass difference is consistent with the expected result from covalent addition (675.7 Da corresponding to the mass of the probe (693.7 Da) minus the loss of one molecule of H_2O). Relative integration of the two signals gave an estimated 70–80% yield for the photoactivation step.*

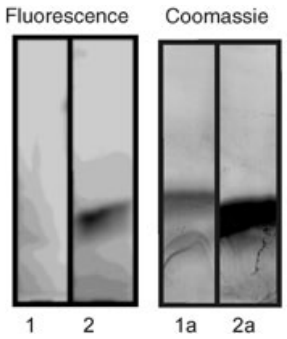


Figure 2. *Specific labelling of galectin-1 after the two-step (photoactivation and chemoselective ligation) procedure; excitation 532 nm, emission 560 nm. Lanes 1 and 1a: pure galectin-1 where the protein has been denatured prior to incubation with* ***6****; lanes 2 and 2a: result of galectin-1 (active) photoincubation with* ***6****.*

mation) yielded no detectable protein labelling. Our method, which includes design aspects, complements the application of the Sulfo-Bed cross-linking agents that were conjugated to (neo)glycoproteins and shown to capture bacterial adhesin proteins.[22] Whether the design aspect of the placement of the photoreactive group is required might depend on the targeted

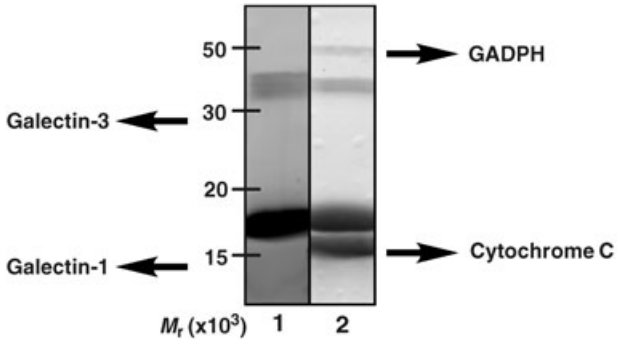


Figure 3. *SDS-PAGE of the products obtained after photoaffinity labelling and rhodamine tag attachment of a mixture of galectin-1, galectin-3, cytochrome C and glyceraldehyde-3-phosphate dehydrogenase. Lane 1: fluorescence image; lane 2: Coomassie staining.*

lectin or lectin class. This novel probe will now be applied in a proteomic study in complex biological samples to help unravel different aspects of galectin biology. The approach is amenable to many protein families, since only the existence of a photoreactive version of a ligand/inhibitor is needed. The use of large molecular constructs is also minimised, as the result of the use of chemoselective ligation methods to incorporate the reporter group. To those attractive features must be added its simplicity and sensitivity, the small amounts of probe required and the ease of synthesis, thus providing a quick and inexpensive way for many laboratories to analyse particular subproteomes of interest.

Acknowledgements

We thank Dr. M. J. E. Fischer for help with SDS-PAGE work. This work was supported by the European Union in the form of a Marie Curie Individual Fellowship to L.B. (contract number MCFI-2002-00249) and the Netherlands Proteomics Centre.

Keywords: chemoselectivity · galectins · glycoconjugtes · photoaffinity labeling · proteomics

[1] R. Aebersold, M. Mann, *Nature* **2003**, *422*, 198–207.
[2] K. Voseller, L. Wells, G. W. Hart, *Biochimie* **2001**, *83*, 575–581.
[3] N. Jessani, B. J. Cravatt, *Curr. Opin. Chem. Biol.* **2004**, *8*, 1–6; M. Ichikawa, Y. Ichikawa, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1769–1773.
[4] *Essentials of Glycobiology*, A. Varki, R. Cummings, J. Esko, H. Freeze, G. W. Hart, J. Marth, CSHL Press, New York, **1999**.
[5] a) H.-J. Gabius, *Eur. J. Biochem.* **1997**, *243*, 543–576; b) S. H. Barondes, D. N. W. Cooper, M. A. Gitt, H. Leffler, *J. Biol. Chem.* **1994**, *269*, 20807–20810.
[6] A. Danguy, I. Camby, R. Kiss, *Biochim. Biophys. Acta* **2002**, *1572*, 285–293.
[7] A. Hittelet, H. Legendre, N. Nagy, Y. Bronckart, J. C. Pector, I. Salmon, P. Yeaton, H. J. Gabius, R. Kiss, I. Camby, *Int. J. Cancer* **2003**, ***103***, 370–379.
[8] H. Lahm, S. André, A. Hoeflich, H. Kaltner, H.-C. Siebert, B. Sordat, C.-W. von der Lieth, E. Wolf, H.-J. Gabius, *Glycoconjugate J.* **2003**, *20*, 227–238.
[9] J. Koptiz, C. Von Reitzenstein, S. André, H. Kaltner, J. Uhl, V. Ehemann, M. Cantz, H.-J. Gabius, *J. Biol. Chem.* **2001**, *276*, 35917–35923.
[10] S. A. Fleming, *Tetrahedron* **1995**, *51*, 12479–12520.

[11] a) W. G. Lewis, L. G. Green, F. Grynszpan, Z. Rádic, P. R. Carlier, P. Taylor, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 1095 – 199; *Angew. Chem. Int. Ed.* **2002**, *41*, 1053 – 1057; b) C, W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057 – 3064.
[12] a) G. C. Adam, C. D. Vanderwal, E. J. Sorensen, B. F. Cravatt, *Angew. Chem.* **2003**, *42*, 5480 – 5484, *Angew. Chem. Int. Ed.* **2003**, *42*, 5480 – 5484; b) M. C. Hagenstein, J. H. Mussgnug, K. Lotte, R. Plessow, A. Brockhinke, O. Kruse, N. Sewald, *Angew. Chem.* **2003**, *115*, 5793 – 5796; *Angew. Chem. Int. Ed.* **2003**, *42*, 5635 – 5638; c) H. Ovaa, P. F. van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebiger, A. M. C. H. van den Nieuwendijk, P. J. Galardy, G. A. van der Marel, H. L. Ploegh, H. S. Overkleeft, *Angew. Chem.* **2003**, *115*, 3754 – 3757; *Angew. Chem. Int. Ed.* **2003**, *42*, 3626 – 3629; d) A. E. Speers, B. F. Cravatt, *ChemBioChem* **2004**, *5*, 41 – 47; e) A. E. Speers, G. C. Adam, B. F. Cravatt, *J. Am. Chem. Soc.* **2003**, *125*, 4686 – 4687.
[13] "Crystal Structure Analysis of Human Galectin-I" (Protein Data Bank code 1GZW) M. I. F. Lopez-Lucendo, D. Solis, H. Kaltner, H.-J. Gabius, A. Romero, unpublished results.
[14] J. Seetharaman, A. Kanigsberg, R. Slaaby, H. Leffler, S. H. Barondes, J. M. Rini, *J. Biol. Chem.* **1998**, *273*, 13 047 – 13 052.
[15] P. Sorme, Y. Qian, P. G. Nyholm, H. Leffler, U. J. Nilsson, *ChemBioChem* **2002**, *3*, 183 – 189.
[16] G. Dormán, G. D. Prestwich, *Trends Biotechnol.* **2000**, *18*, 64 – 77.
[17] Y. Nakatsuji, N. Kameda, M. Okahara, *Synthesis* **1987**, *3*, 280 – 281.
[18] K. P. R. Kartha, H. J. Jennings, *J. Carbohydr. Chem.* **1990**, *9*, 777 – 781.
[19] J. Li, S. Zacharek, X. Chen, J. Wang, W. Zhang, A. Janczuk, P. G. Wang, *Bioorg. Med. Chem.* **1999**, *7*, 1549 – 1558.
[20] These regio-selective alkylation reactions were found to proceed more rapidly (from 5–10 h to 5 min) and with higher yields (>80%) by the use of microwave irradiation. L. Ballell, J. A. F. Joosten, F.-A. el Maate, R. M. J. Liskamp, R. J. Pieters, *Tetrahedron. Lett.* **2004**, 45, 6685 – 6687.
[21] Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192 – 3193.
[22] D. Ilver, A. Arnqvist, J. Ögren, I.-M. Frick, D. Kersulyte, E. T. Incecik, D. E. Berg, A. Covacci, L. Engstrand, T Borén, *Science* **1998**, *279*, 373 – 377.

Received: June 20, 2004
Published online: December 2, 2004

