Stability and Function of Interdomain Linker Variants of Glucoamylase 1 from *Aspergillus niger*†

Jørgen Sauer,‡ Trine Christensen,§,|| Torben P. Frandsen,‡,⊥ Ekaterina Mirgorodskaya,#,▽ Kirsten A. McGuire,‡,▼ Hugues Driguez,♦ Peter Roepstorff,# Bent W. Sigurskjold,§ and Birte Svensson*,‡

Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark, Department of Biochemistry, August Krogh Institute, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen Ø, Denmark, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense University, ampusvej 55, DK-5230 Odense M, Denmark, and Centre de Recherches sur les Macromolécules Végétales, CNRS, F-38041 Grenoble, Cedex 9, France

Received March 13, 2001; Revised Manuscript Received May 31, 2001

ABSTRACT: Several variants of glucoamylase 1 (GA1) from *Aspergillus niger* were created in which the highly *O*-glycosylated peptide (aa 468−508) connecting the $(\alpha/\alpha)_6$-barrel catalytic domain and the starch binding domain was substituted at the gene level by equivalent segments of glucoamylases from *Hormoconis resinae*, *Humicola grisea*, and *Rhizopus oryzae* encoding 5, 19, and 36 amino acid residues. Variants were constructed in which the *H. resinae* linker was elongated by proline-rich sequences as this linker itself apparently was too short to allow formation of the corresponding protein variant. Size and isoelectric point of GA1 variants reflected differences in linker length, posttranslational modification, and net charge. While calculated polypeptide chain molecular masses for wild-type GA1, a nonnatural proline-rich linker variant, *H. grisea*, and *R. oryzae* linker variants were 65 784, 63 777, 63 912, and 65 614 Da, respectively, MALDI-TOF-MS gave values of 82 042, 73 800, 73 413, and 90 793 Da, respectively, where the latter value could partly be explained by an *N*-glycosylation site introduced near the linker C-terminus. The k_{cat} and K_m for hydrolysis of maltooligodextrins and soluble starch, and the rate of hydrolysis of barley starch granules were essentially the same for the variants as for wild-type GA1. β-Cyclodextrin, acarbose, and two heterobidentate inhibitors were found by isothermal titration calorimetry to bind to the catalytic and starch binding domains of the linker variants, indicating that the function of the active site and the starch binding site was maintained. The stability of GA1 linker variants toward GdnHCl and heat, however, was reduced compared to wild-type.

Glucoamylase (1,4-α-D-glucan glucohydrolase, EC 3.2.1.3, GA)[1] hydrolyzes α-1,4 and with approximately 700-fold lower activity α-1,6 glucosidic linkages to release β-D-glucose from the nonreducing ends of starch and related polysaccharides and oligosaccharides (*1*−*4*). GA1 from *Aspergillus niger* consists of three parts: (i) a catalytic domain (aa 1−440); and (ii) a highly glycosylated linker (aa 441−508) that at the C-terminus connects to (iii) an antiparallel β-sheet starch binding domain (aa 509−616) (*5*−*7*). While the structure of the entire *A. niger* GA1 is currently not available, the high-resolution crystal structure of the catalytic domain of the 96% identical *Aspergillus awamori* var. *X100* GA1 showed an $(\alpha/\alpha)_6$-barrel fold (*8*). GA is produced by *A. niger* in two forms: GA1 containing all three of the above-mentioned regions and GA2 that lacks the starch binding domain (*5*). GA1 and GA2 possess very similar properties in their action toward soluble substrates (*9*), while GA2, in contrast to GA1, hydrolyzes granular starch very slowly (*5*). The function and structure of the catalytic and starch binding domains have been thoroughly investigated (*10*), but the dynamics of the cooperation between the domain and the three-dimensional structure of intact GA1 remain to be determined.

The *O*-glycosylated linker of GA1 has been attributed roles in stability, adsorption onto raw starch granules, and secretion (*11*−*14*). Recently, addition of SBD to a suspension of insoluble corn starch containing GA2 was found to increase the rate of starch hydrolysis, suggesting that the SBD

† This work was supported by the Danish Research Councils' Committee on Biotechnology (No. 9502014) and by the European Union 4th Framework Programme on Biotechnology (BIO4-CT98-0022).

* To whom correspondence should be addressed. Phone: +45 3327 5345. Facsimile: +45 3327 4708. E-mail: bis@crc.dk.

‡ Carlsberg Laboratory.

§ August Krogh Institute, University of Copenhagen.

|| Current address: Department of Chemistry, Duke University, Durham, NC 27708.

⊥ Current address: Novozymes, Novo Allé, DK-2880 Bagsværd, Denmark.

University of Southern Denmark, Odense University.

▽ Current address: Boston University School of Medicine, 715 Albany St., R-806 Boston, MA 02118-2526.

▼ Current address: NatImmune, Symbion Science Park, Fruebjergvej 3, DK-2100 København Ø, Denmark.

♦ Centre de Recherches sur les Macromolécules Végétales, CNRS.

[1] Abbreviations: GA1, glucoamylase 1 (aa 1−616) from *Aspergillus niger*; GA2, glucoamylase 2 (aa 1−512/514), a truncated form of glucoamylase without the starch binding domain; LV, linker variant; DSC, differential scanning calorimetry; ITC, isothermal titration calorimetry; SBD, starch binding domain; β-CD, β-cyclodextrin (cyclomaltoheptaose); L0 and L14, heterobidentate ligands containing acarbose linked directly to β-cyclodextrin and via a spacer of 14 Å, respectively.

10.1021/bi010515i CCC: $20.00
Published on Web 07/12/2001

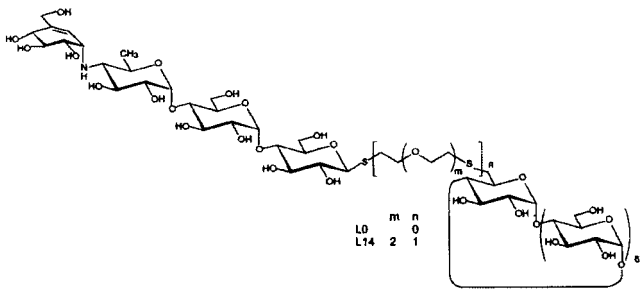

FIGURE 1: Structures of the heterobidentate ligands containing acarbose and β-cyclodextrin joined by an oligo(ethylene glycol) spacer 14 Å in length (L14) or containing no spacer (L0).

disrupted α-glucan chain interactions and thus increased the accessibility to the enzyme of the α-glucan chains in hydrolysis of granular starch (*15*). Earlier measurement by isothermal titration calorimetry of GA1 binding of synthetic heterobidentate inhibitors consisting of acarbose connected to β-cyclodextrin through oligo(ethylene glycol) spacers of different lengths (Figure 1; *16*) demonstrated simultaneous association to the active site of the catalytic domain and one of the two starch binding sites on the SBD. This indicates that the two domains are close to each other in the solution conformation of GA1 (*17*). The binding of the heterobidentate inhibitors furthermore shortened the hydrodynamic long axes of GA1 as shown by dynamic light scattering measurements (*16*), whereas scanning tunneling microscopy images earlier had suggested that the two domains were far apart and connected by the linker region in an extended conformation (*18*).

Genes encoding GA1 variants containing new linkers were generated in the present work and were expressed in the methylotrophic yeast *Pichia pastoris*, a system established recently and shown to efficiently produce recombinant GA1 having essentially the same catalytic properties as GA1 from *A. niger* (*19*). Small differences, however, were found in the extent of glycosylation, which apparently did not influence enzymatic activity (*19*). In the present study, the series of *A. niger* GA1 protein variants produced had the most highly *O*-glycosylated part of the linker region replaced with naturally occurring linker sequences of glucoamylases from *Humicola grisea*, *Rhizopus oryzae*, and *Hormoconis resinae* as well as one unnatural proline-rich linker. The parent enzymes thus included *Rhizopus oryzae* GA where SBD is situated in front of the catalytic domain (*20*), the thermostable *Humicola grisea* GA (*21*), and *Hormoconis resinae* GA having unusually high activity toward α-1,6 glucosidic linkages (*22*). While these glucoamylases possess highly homologous catalytic and starch binding domains, they show large variations in both length and sequence of linker regions (*23*). By exchanging the *A. niger* wild-type GA1 linker with corresponding regions from other glucoamylases and a proline-rich artificial linker, we aimed at examining the role of the linker region with regard to thermostability, domain interaction, and catalytic specificity. Also we pursue the impact of variations in the number of glycosylation sites in the linker region on the above-mentioned parameters. Indeed, such linkers have potential in future combination of new functionalities, and therefore insight into the structure/function relationships of linkers is relevant not only for glucoamylases but also for application as a biotechnological tool.

MATERIALS AND METHODS

Chemicals, Proteins, and Kits. Plasmid DNA for construction and transformation was prepared using the GFX kit, and purification of PCR fragments and restriction fragments was done using the GFX PCR or gel band purification kit, all from Amersham Pharmacia Biotechnology (Uppsala, Sweden). Restriction endonucleases, T4-DNA ligase, calf intestinal phosphatase, and PFU polymerase were obtained from Promega (Madison, WI). Oligodideoxynucleotides were purchased from DNA Technology A/S (Århus, Denmark), and sequencing was performed using the BigDye sequencing kit and an ABI PRISM 310 capillary sequencer (Perkin-Elmer, Germany). The maltooligosaccharides maltose through maltoheptaose, isomaltose, soluble starch, the glucose oxidase kit, and guanidine hydrochloride were obtained from Sigma (St. Louis, MO). Acarbose was a generous gift of Drs. D. Schmidt and E. Truscheit (Bayer AG, Wuppertal, Germany), and acarbose−Sepharose was prepared essentially as previously described (*24*). The heterobidentate inhibitors L0 and L14 (Figure 1) were synthesized previously (*16*). GA1 and GA2 from *A. niger* were purified from the commercial enzyme AMG 200L (Novo Nordisk, Bagsværd, Denmark), and wild-type GA1 overexpressed in *A. niger* was prepared as described earlier (*5*, *9*, *25*). Polyclonal antibody against *A. niger* GA1 was a laboratory stock.

Bacterial and Yeast Strains, Growth Conditions, and Plasmids. Escherichia coli strain DH5α (*26*) grown in LB (Luria broth), 100 mg/L ampicillin (for solid media 15 g/L agarose was added) was used as host during construction of plasmids. *Pichia pastoris* GS115 (Invitrogen, San Diego, CA) was grown and used for expression of the GA1 linker variants essentially as previously described (*19*). The pBS$^-$ derivative pXWT containing the complete GA1 encoding sequence including propeptide and signal peptide (*19*) was used as template for creating the linker replacement variants. The *P. pastoris* expression vector pHIL-D2 (Invitrogen) harboring the constructed linker replacement variants was used to produce secreted GA1 variants.

Construction of Linker Variants and Transformation of P. pastoris. The region encoding amino acids 468−508 of *A. niger* GA1 was replaced with the corresponding regions of homologous glucoamylase genes from *R. oryzae*, *H. grisea*, and *H. resinae*. In the latter construct, the linker coding region was moreover elongated to encode the proline-rich sequence extensions PSPSPY, PSPSPYPSPSPY, and PYPYPY, respectively (Figure 2). The gene encoding the linker replacement variant containing the *R. oryzae* GA linker was created in two steps by PCR (Figure 3). In the first step, the oligodeoxynucleotide primer 1 was used together with the pBS$^-$-specific reverse primer with pXWT as template, generating a PCR fragment encoding the *N*-terminal part of GA1 from *A. niger* and 27 amino acids of the 37 amino acid long linker from *R. oryzae* GA. The complete *R. oryzae* linker variant gene was obtained by ligating a 856 bp *Pst*I/*Bst*XI-cleaved PCR product generated using the oligodeoxynucleotide's primer 2 together with the reverse primer and the PCR product from the first reaction as template into pXWT, which was previously cleaved with the same enzymes and treated with calf intestinal phosphatase. The variant GA1 genes containing the *H. grisea*, *H. resinae*, and three elongated *H. resinae* linkers (Figure 3) were created

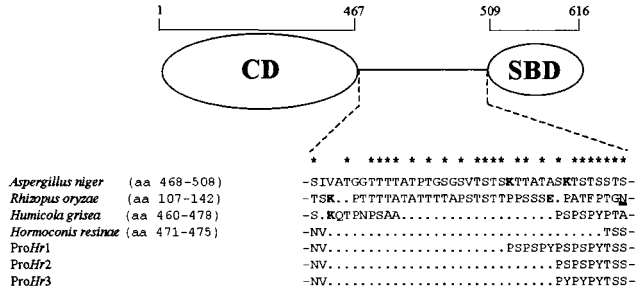

FIGURE 2: Schematic representation of the constructed *A. niger* GA1 linker variants showing the replacing encoded sequences. The catalytic domain (CD; aa 1−467) and starch binding domain (SBD; aa 509−616) are connected through a highly *O*-glycosylated linker segment (aa 468−508). The alignment shows linker regions from fungal glucoamylases used to replace the *A. niger* GA1 linker. Linker variant genes encoding the Pro*Hr*1, Pro*Hr*2, and Pro*Hr*3 GA1s were created by inserting proline-rich motifs into the *H. resinae* linker variant GA1. Numbers in parentheses represent the corresponding wild-type sequences. Charged amino acids in the linker region are marked in boldface type, and asterisks above the alignment indicate glycosylated serine and threonine in *A. niger* GA1 (*6*). The putative *N*-glycosylation site in the *R. oryzae* linker variant is underlined.

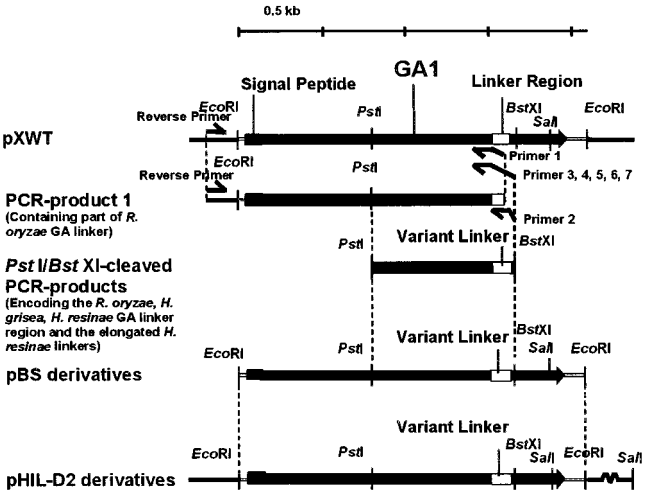

FIGURE 3: Construction of linker variant genes. Genes and sequences of interest are indicated as black or white boxes. The primers used are displayed on the figure as arrows. Important restriction enzyme sites used during construction are indicated (see Materials and Methods for detailed explanation).

in one step using primers 3, 4, 5, 6, and 7, respectively, together with the pBS$^-$-specific reverse primer and pXWT as template, generating, following *Pst*I/*Bst*XI cleavage of the resulting PCR fragment, 802, 760, 778, 796, and 778 bp fragments. These fragments were subsequently ligated into pXWT, previously cleaved with the same enzymes and treated with calf intestinal phosphatase (Figure 3).

The genes encoding GA1 linker variants were cloned into the *P. pastoris* expression vector pHIL-D2 by ligating an *Eco*RI fragment from the pXWT derivatives into the vector previously linearized with the same enzyme and treated with calf intestinal alkaline phosphatase. Subsequently, the correct orientation of the insert relative to the AOX promoter was checked with the restriction enzyme *Sal*I for which a site is situated in the 3′ end of the linker variant genes, close to another *Sal*I site situated in the vector.

For transformation of *P. pastoris*, 5 μg of each of the pHIL-D2 derivatives was prepared and digested 2 h with *Not*I at 37 °C, ethanol-precipitated, resuspended in 5 μL of sterile water, and used for transformation of *P. pastoris* by electroporation as previously described (*19*). Positive transformants were selected on plates as colonies able to synthesize histidine (*His*$^+$), but unable to grow on methanol as sole carbon source (*Mut*$^-$).

Expression of Linker Variant GA. For small-scale expression tests, transformant colonies were grown in 14 mL Falcon tubes (Becton Dickenson, New York) in 2 mL of BMGY (6.7 g/L yeast nitrogen base without amino acids, 20 g/L peptone, 10 g/L yeast extract, 1% glycerol, 0.4 mg/L D-biotin, 0.1 M potassium phosphate, pH 6) for 48 h at 30 °C followed by centrifugation (1500*g*, 5 min) and resuspension in BMMY (6.7 g/L yeast nitrogen base without amino acids, 20 g/L peptone, 10 g/L yeast extract, 0.4 mg/L D-biotin, 0.75% methanol). After 48 h of induction at 30 °C, the culture supernatants were isolated by centrifugation, and GA activity was measured using maltose as substrate. Large-scale production of linker variant GAs was initiated by growing a *P. pastoris* transformant starter culture of 50 mL of BMGY in 1 L flasks under vigorous shaking (New Brunswick Lab Shaker, 200 rpm) for 24 h at 30 °C. The starter culture was used (5 × 10 mL) to inoculate 5 × 1 L of BMGY in 5 L flasks and grown until an OD_{600} of approximately 25. The cells were harvested by centrifugation at 1500*g* for 5 min and resuspended in 2 × 0.5 L of BMMY in 2 × 5 L flasks. After 48 h, cells were harvested as before and resuspended in 1 L of fresh BMMY in a 5 L flask for a second induction of 48 h. The combined supernatants from the first induction culture batches were stored at 4 °C for subsequent combination with the second induction supernatant to a total volume of 2 L.

Purification of Linker Variants GAs. Culture supernatant (2 L) was concentrated to 250 mL (Pellicon cell; Amicon membrane, cutoff 10 kDa), and the medium was simultaneously exchanged by 0.1 M sodium acetate, pH 4.5, 0.5 M NaCl. The concentrate was divided into two portions, each of which was applied onto an acarbose−Sepharose column (diameter 2.6 cm, containing 5 mL of resin) followed by elution with 1.7 M Tris-HCl and dialysis against 20 mM piperazine, pH 5.4, 0.15 M NaCl, 0.002% NaN_3 (w/v). The final purification on HiLoad Q-Sepharose was done as previously described (*9*, *24*). The purity of GA1 variants was assessed by SDS−PAGE (gradient gel 10−15%), IEF (pH 3.5−5), and Western Blot using GA1 antiserum with the Phast-System (Amersham-Pharmacia, Uppsala, Sweden) according to the manufacturer's recommendation. The protein concentration was determined either by aid of amino acid analysis or spectrophotometrically using $\epsilon_{280} = 1.37 \times 10^5$ cm^{-1} M^{-1} for GA1 and 1.09×10^5 cm^{-1} M^{-1} for GA2 (*24*). The GA1 value was also used for the linker variants as the content of Trp was unchanged from wild-type. Amino acid analysis was performed on acid hydrolysates of 10−50 μg of GA (6 M HCl; 110 °C; 24 h) using an LKB model Alpha Plus amino acid analyzer. The N-terminal sequence was determined using an Applied Biosystems 470A protein sequencer equipped with an on-line PTH-analyzer.

Enzyme Assays. The activity of GA1 linker variants in culture supernatants and during purification was measured toward 28 mM maltose as substrate in 50 mM sodium acetate, pH 4.5 at 45 °C, and quantifying the formed glucose by the glucose oxidase method reading the absorbance in

microtiter plates using a CERES UV900 (BIOTEK) ELISA reader (*25*).

Kinetic parameters k_{cat} and K_m were detemined for the maltooligosaccharides maltose through maltoheptaose, isomaltose, and soluble starch. Initial rates of hydrolysis by 2.1−100 nM enzyme at 12 different substrate concentrations in the range (1/10)K_m and 8K_m were fitted to the Michaelis−Menten equation using the program ENZFITTER (*27*).

Hydrolysis of Starch Granules. Barley starch granules (4 g/L; Primalco, Finland) were incubated with 20 nM enzyme in 50 mM sodium acetate, pH 4.5 (2 mL), at 45 °C in 10 mL round-bottom tubes in a vertical shaker. Aliquots of 100 μL were removed and frozen immediately. Following thawing and discarding residual starch granules after centrifugation (20000*g*, 20 min), released glucose was quantified using the glucose oxidase method (*25*).

Thermostability. GA1 variants and wild-type (0.6 μM) were incubated for 5 min at 5 °C intervals in the temperature range 30−90 °C in 50 mM sodium acetate, pH 4.5, and snap-frozen in liquid nitrogen. Following thawing, the residual activity was determined toward 28 mM maltose as above. T_m was defined as the temperature corresponding to 50% loss of the initial activity.

Guanidine Hydrochloride Denaturation. Enzyme (0.25 μM) was incubated at 25 °C for 1 h at varying concentrations of GdnHCl (0−7.8 M) in 50 mM sodium acetate, pH 4.5. Protein unfolding was monitored fluorometrically by exciting at 280 nm and monitoring the emission at 360 nm using a Perkin-Elmer LS50 luminescence spectrometer. The fraction of unfolded protein was calculated from the equation: $f_u = (F_N - F_{obs})/(F_N - F_U)$, where F_N, F_U, and F_{obs} are the fluorescence intensities for completely folded, completely unfolded, and partially unfolded protein at a given GdnHCl concentration, respectively.

Mass Spectrometry. MALDI-TOF was performed on a Voyager-Elite PerSeptive Biosystem Inc. (Framingham, MA) instrument in the linear mode using delayed extraction. Positive ions were accelerated to 20 kV in a dual-step ion source from a matrix of 9:1 2,5-dihydroxybenzoic acid (25 g/L) and 2-hydroxy-5-methoxybenzoic acid (25 g/L) (Aldrich, Milwaukee, WI) dissolved in a 2:1 mixture of water and acetonitrile containing 0.1% trifluoroacetic acid (*36*, *37*) after a delay of 160 μs. The samples were prepared by mixing (1 μL) with matrix (1 μL) directly on the target followed by air-drying. The spectra were calibrated externally using singly and doubly charged ions of bovine serum albumin (Sigma).

Isothermal Titration Calorimetry (ITC). ITC experiments were performed using an MCS isothermal titration calorimeter from MicroCal, Inc. (Northampton, MA) (*28*). Titrations were carried out at 27 °C in 50 mM sodium acetate, pH 4.5. Briefly, enzyme (9−35 μM) was titrated with 21 portions of 13 μL of ligand dissolved in the above buffer, covering the ranges 65−350 μM for acarbose and the heterobidentate inhibitors, and 0.3−1 mM for β-cyclodextrin. Calibration of the instrument was accomplished using electrical heat pulses. Acquisition of titration thermograms and integration of the heat signals were controlled by the Observer and Origin ver. 5.0 software, respectively, from MicroCal, Inc. Blank titrations of ligand into buffer were used to correct for heat of dilution, and nonlinear regression analysis of binding isotherms was performed as previously described (*17*, *29*).

Differential Scanning Calorimetry (DSC). DSC experiments were performed on an MSC differential scanning calorimeter from MicroCal, Inc., with a cell volume of 1.1934 mL. Calibration of temperature was done using two hydrocarbons with known melting temperature supplied by MicroCal, Inc., in sealed steel capillary tubes. Heat calibration was done using electrical pulses. The scan rate was 60 °C h^{-1}, and an external pressure of 3 bar was applied over the cells. Protein solutions between 15 and 24 μM were prepared by dialyzing against 50 mM sodium phosphate, pH 7.5, and then degassing by stirring under vacuum prior to scanning. Buffer scans were subtracted from protein scans, and the molar heat capacity was obtained (C_p) by normalizing the known protein concentration and cell volume of the calorimeter.

RESULTS

Design of Linker Variants. The segment in the GA1 gene from *A. niger* encoding amino acid residues 468−508 was replaced by DNA encoding the counterparts of glucoamylases from *R. oryzae*, *H. grisea*, and *H. resinae* as well as artificial sequences rich in proline (Pro*Hr*1, -2, and -3) by PCR-insertion of the sequences into the cDNA of *A. niger* GA1 (Figures 2 and 3). In GA1 of *A. niger*, this highly *O*-glycosylated C-terminal part of the domain linker region contains 39 amino acids whereas the constructed variants encode linkers ranging from 5 to 36 amino acid residues (Figure 2). The corresponding GA1 variant proteins were produced in a *P. pastoris* system by using the pHIL-D2 expression vector. The variants were secreted into the culture medium with the correctly processed N-terminal sequence most probably because the constructs comprise the wild-type *A. niger* GA1 signal sequence (see below) (*19*).

Production, Chemical Characterization, and Mass Spectrometry of GA1 Linker Variants. Testing for production of linker variant GA1s in *P. pastoris* minicultures revealed that those containing a 5 amino acid long replacing linker (from *H. resinae*), or 2 different 11 amino acid proline-rich linkers (variants Pro*Hr*2 and Pro*Hr*3) comprising the short *H. resinae* linker (Figure 2), were not detected in the culture medium. Since it has previously been indicated that the linker region from aa 470 to aa 507 was important for secretion of GA1 (*12*), an attempt was made to detect these linker variant proteins intracellularly by Western blotting on cell extracts. However, no immunoreactive protein was detected (results not shown). The GA1 variants with the *Rhizopus*, the *Humicola* sequences, and the longer of the artificial proline-rich linkers (Pro*Hr*1) were produced in large-scale cultures in amounts of 200−300 mg/L compared to 400 mg/L of wild-type GA1 (*19*). No attempt was made to correlate the OD_{600} of the culture or gene-dose with the amount of produced protein. The variants were readily purified to apparent homogeneity from culture supernatants by acarbose−Sepharose affinity chromatography followed by Q-Sepharose anion exchange chromatography (*9*, *24*). The N-terminal sequence of all enzyme variants was Ala-Thr-Leu-Asp-Ser-Trp-Leu-Ser-Asn-Glu, identical to wild-type GA1 (*6*), indicating correct N-terminal processing. The amino acid composition of the variants agreed with that calculated from the corresponding amino acid sequences (results not shown).

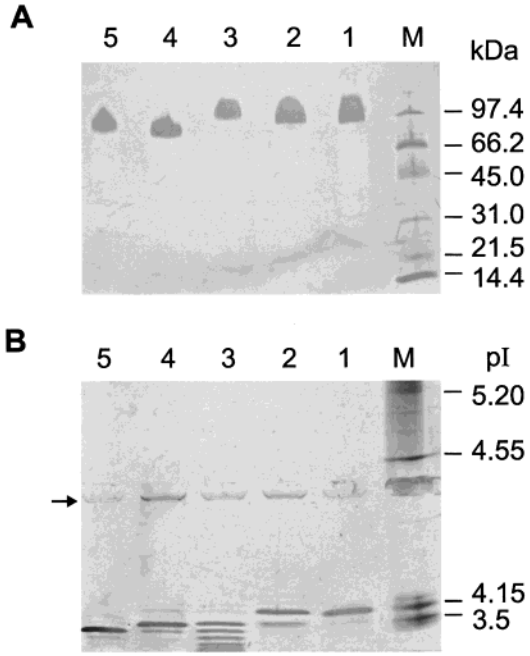

FIGURE 4: Silver-stained SDS−PAGE (A) and IEF (B) gels of the GA1 linker variants. 1 μg and 4 μg of protein were applied on SDS−PAGE or IEF gels, respectively. Lane 1, *A. niger* GA1 purified from a commercial enzyme preparation; lane 2, recombinant GA1 from *P. pastoris*; lane 3, *R. oryzae* linker variant GA; lane 4, *H. grisea* linker variant GA1; and lane 5, Pro*Hr*1 linker variant GA1. The following molecular mass markers were used in SDS−PAGE: phosphorylase *b* (97.4 kDa), serum albumin (66.2 kDa), ovalbumin (45.0 kDa), carbonic anhydrase (31.0 kDa), soybean trypsin inhibitor (21.5 kDa), lysozyme (14.5 kDa). The p*I* markers used for IEF are α-glucoamylase (p*I* 3.5), glucose oxidase (p*I* 4.15), soybean trypsin inhibitor (p*I* 4.55), and β-lactoglobulin A (p*I* 5.20). The point of application on the IEF gel is marked with an arrow.

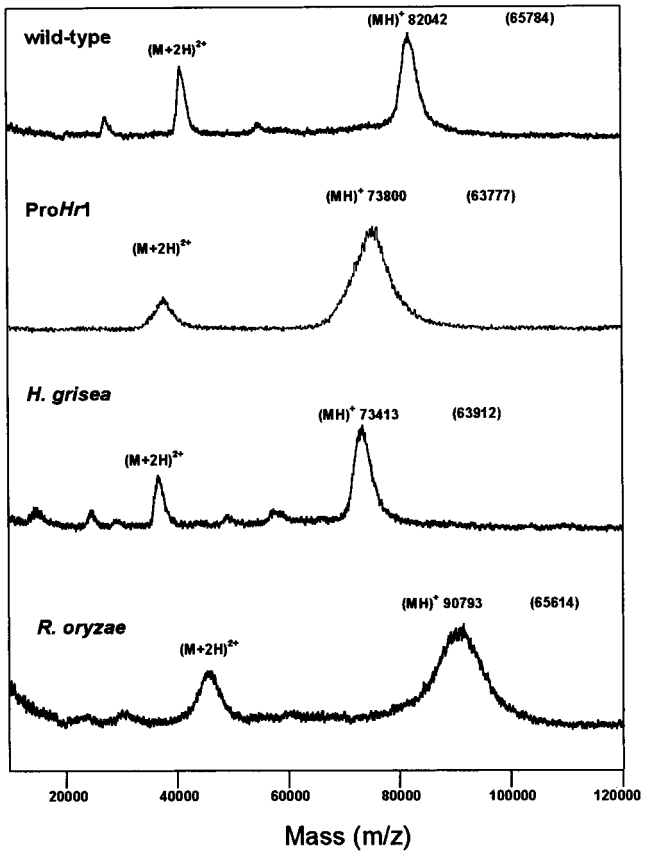

FIGURE 5: MALDI-TOF-MS spectra of recombinant wild-type and linker variant GA1s. The determined masses are indicated on the spectra together with calculated masses (in parentheses) from the amino acid sequence for the unglycosylated polypeptide chain of the different GA1 forms.

The relative size of Pro*Hr*1, *H. grisea*, and *R. oryzae* linker variant GA1s as evaluated by SDS−PAGE (Figure 4A) together with wild-type GA1 produced in *P. pastoris* and wild-type GA1 purified from a commercial glucoamylase showed that the apparent sizes of these variants differed in a way which was not in accordance with the known amino acid sequence differences. The calculated masses of the polypeptide chains were 63 777, 63 912, and 65 614 Da, respectively, compared to 65 784 Da for wild-type GA1. Thus, the order of migration on SDS−PAGE, based on the size of the polypeptide chain, should be ascending from wild-type to the variant containing the 17 amino acid long proline-rich linker (Pro*Hr*1 GA1). Both the *R. oryzae* and this Pro*Hr*1 variant, however, migrated more slowly in SDS−PAGE than expected (Figure 4A, lanes 3 and 5). These relative differences were confirmed by MALDI-TOF-MS. Moreover, the shape of the recorded mass spectra indicates considerably higher heterogeneity for the Pro*Hr1* and *R. oryzae* variants compared to the *H. grisea* variant and wild-type GA1.

The average masses based on singly and doubly charged ions were found for wild-type produced from *P. pastoris*, the Pro*Hr*1, *H. grisea*, and *R. oryzae* linker variant GA1s to be 82 042, 73 800, 73 413, and 90 793 Da, respectively (Figure 5). This suggests glycosylation contributed 16 258, 10 023, 9501, and 25 179 Da to the mass of the proteins. The number of potential *O*-glycosylated serines and threonines in the replaced part of the linker region of wild-type GA1 is 25 compared to 7, 6, and 21 for Pro*Hr*1, *H. grisea*, and *R. oryzae* linker variant GA1s, respectively. In wild-type GA1 from *A. niger*, the vast majority of potential *O*-glycosylation sites from Ser468 to Ser508, both inclusive, are indeed glycosylated (*6*). Thus, since the *R. oryzae* linker variant only contains 21 Ser and Thr, which is 4 less than wild-type, the high mass difference of 25 179 Da cannot likely be attributed only to *O*-glycosylation of the linker region, but most probably originates from other modifications. In this variant, a putative *N*-glycosylation site was introduced at the C-terminal asparagine of the substituted linker sequence, followed by Cys and Thr of SBD, and thus this site was probably utilized. A mass increase on a single *N*-glycosylation site has been identified in heterologous expression of the gene encoding rDER f 1 house dust mite allergen, giving rise to a similar large size increase (*30*). Moreover, additional modifications in the form of phosphorylation as suggested by the occurrence of unusually acidic forms (see below) may account for a minor mass increase. The calculated polypeptide chain masses of Pro*Hr*1 and *H. grisea* linker variant GA1s were 63 777 and 63 912 Da, respectively. Thus, the polypeptide of the *H. grisea* variant, containing a 19 amino acid linker with 6 potential *O*-glycosylation sites, is 134 Da larger than the Pro*Hr*1 linker variant of 17 amino acids containing 7 potential *O*-glycosylation sites. Both SDS−PAGE (Figure 4A) and mass spectrometry, however, indicated a higher molecular mass for the Pro*Hr*1 GA1, MALDI-TOF-MS (Figure 5) giving 387 Da higher mass for this compared to *H. grisea* linker variant, suggesting a higher degree of glycosylation.

Table 1: Kinetic Parameters of GA1 and GA1 Linker Variants for Hydrolysis of Maltooligosaccharides, Isomaltose, and Soluble Starch[a]

	Rhizopus oryzae LV			*Humicola grisea* LV		
substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
maltose (α-1,4)	10.2 ± 0.2	1.25 ± 0.06	8.2 ± 0.4	8.9 ± 0.2	1.32 ± 0.06	6.7 ± 0.3
maltotriose	32.5 ± 0.6	0.37 ± 0.02	88 ± 5	31.5 ± 0.3	0.36 ± 0.009	88 ± 2
maltotetraose	44.9 ± 0.4	0.11 ± 0.003	409 ± 11	46.5 ± 0.5	0.10 ± 0.004	465 ± 19
maltopentaose	45.2 ± 0.2	0.11 ± 0.002	411 ± 8	46.0 ± 0.4	0.10 ± 0.003	460 ± 14
maltohexaose	47.6 ± 0.4	0.11 ± 0.003	433 ± 12	50.6 ± 0.6	0.11 ± 0.004	460 ± 18
maltoheptaose	56.7 ± 0.7	0.10 ± 0.004	567 ± 24	55.7 ± 0.8	0.12 ± 0.005	464 ± 21
isomaltose (α-1,6)	0.35 ± 0.01	33.9 ± 2.3	0.010 ± 0.001	0.33 ± 0.02	34.0 ± 3.2	0.009 ± 0.001
soluble starch	60.2 ± 0.09	0.15 ± 0.005[b]	401 ± 13[c]	56.7 ± 0.7	0.17 ± 0.008[b]	334 ± 16[c]
	Pro*Hr*1 LV			wild-type		
substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
maltose (α-1,4)	9.2 ± 0.3	1.05 ± 0.1	8.8 ± 0.6	8.2 ± 0.3	1.12 ± 0.12	7.3 ± 0.8
maltotriose	28.9 ± 0.5	0.30 ± 0.01	96 ± 4	27.8 ± 0.5	0.29 ± 0.06	96 ± 4.6
maltotetraose	48.7 ± 0.5	0.10 ± 0.003	487 ± 15	47.0 ± 0.8	0.10 ± 0.005	470 ± 25
maltopentaose	46.2 ± 0.4	0.12 ± 0.003	385 ± 12	43.2 ± 0.8	0.12 ± 0.006	360 ± 19
maltohexaose	48.4 ± 0.4	0.11 ± 0.003	440 ± 13	49.1 ± 0.7	0.10 ± 0.005	491 ± 26
maltoheptaose	55.5 ± 0.7	0.12 ± 0.005	463 ± 20	56.0 ± 0.6	0.12 ± 0.004	467 ± 16
isomaltose (α-1,6)	0.30 ± 0.02	30.0 ± 4.7	0.010 ± 0.002	0.29 ± 0.013	33.3 ± 5.3	0.009 ± 0.001
soluble starch	55.8 ± 0.5	0.19 ± 0.007[b]	294 ± 11[c]	55.5 ± 1.5	0.21 ± 0.0007[b]	264 ± 7.2[c]

[a] Assays were performed in 50 mM sodium acetate, pH 4.5 at 45 °C, using 12 different substrate concentrations between (1/10)K_m and 8K_m. [b] K_m expressed in % soluble starch. [c] k_{cat}/K_m expressed in s^{-1} $\%^{-1}$.

Recombinant wild-type GA1 produced in *P. pastoris* and the GA1 purified from a commercial preparation produced by *A. niger* both have isoelectric points of 4.2 with minor more acidic forms being present in the recombinant GA1 (Figure 4B). The p*I* values of the different GA1 linker variants, however, deviate from that of the wild-type GA1 (Figure 4B), reflecting differences in the content of charged residues. The replaced region in wild-type GA1 linker contains two lysines, whereas linker sequences introduced from *R. oryzae* and *H. grisea* GA have one lysine and one glutamic acid, and one lysine, respectively, and the Pro*Hr*1 linker variant has no charged side chains (Figure 2). Remarkably, the *R. oryzae* variant preparation contains several more acidic forms (Figure 4B) presumably due to phosphorylation of mannosyl residues, a phenomenon previously seen with recombinant protein produced in *P. pastoris* (*31*).

Enzymatic Properties of GA1 Linker Variants. The Michaelis−Menten kinetic parameters of the glucoamylase variants for maltooligosaccharides are listed in Table 1. *A. niger* glucoamylase forms show the same activity toward maltose whether the SBD is present (*9*). The linker variant and wild-type GA1s also have very similar kinetic parameters for hydrolysis of the α-1,4-linked maltose through maltoheptaose, and the α-1,6-linked isomaltose. Thus, k_{cat} for maltose was in the range of 8.2−10.2 s^{-1}, increasing in the order of wild-type, *H. grisea*, Pro*Hr*1, and *R. oryzae* linker variant GA1s. K_m values were in the range 1.05−1.32 mM, increasing for Pro*Hr*1, wild-type, *R. oryzae*, and *H. grisea* linker variant GAs. The longer oligosaccharides were superior substrates, k_{cat} for maltotriose varying from 27.8 to 32.5 s^{-1}, wild-type having the lower and again the *R. oryzae* linker variant GA1 the higher value; K_m values were in the range 0.29−0.37 mM; again, wild-type and the Pro*Hr*1 linker variant GA1 had the lowest values. The kinetic parameters for maltotetraose, -pentaose, and -hexaose were all very similar, k_{cat} being in the range 43.2−50.6 s^{-1} with no enzyme form being consistently superior or inferior toward all three maltooligosaccharides. k_{cat} for maltoheptaose was for all enzyme forms 55.5−56.7 s^{-1} and K_m 0.10−0.12 mM. Even though the three-dimensional structure of the catalytic domain of glucoamylase is known (*8*), it is not possible to ascribe such small variations in kinetic parameters to the impact of the modest differences in glycosylation of the different forms, or to global conformational differences caused by the different linker sequences. No significant differences were found in activity toward isomaltose (Table 1). For different forms, the relative specificity toward the α-1,4-linked maltose over the α-1,6-linked isomaltose ($[k_{cat}/K_m]_{maltose}/([k_{cat}/K_m]_{isomaltose}$) varied in the range from 690 to 880, which is similar to wild-type GA1 that has a value of 838.

Table 2: Hydrolysis of Barley Starch Granules

enzyme	activity[a] (%)
wild-type GA1 produced in *P. pastoris*	100
wild-type GA2 from *A. niger*	<0.1
wild-type GA1 from *A. niger*	100
R. oryzae linker variant	90
H. grisea linker variant	94
Pro*Hr*1 linker variant	98

[a] The experiment was performed at 45 °C in 50 mM sodium acetate, pH 4.5, containing 4 mg/mL barley starch granules.

The finding that the linker variant engineering caused very little effect on the enzymatic activity was also observed for the polymeric substrates soluble starch and barley starch granules (Tables2 and 3). K_m and k_{cat} for soluble starch increased and decreased, respectively, both in the order of *R. oryzae*, *H. grisea*, Pro*Hr*1 linker variants, and wild-type GA1 (Table 1). The *R. oryzae* linker variant GA1 has the higher activity for soluble starch (Table 1). The shorter the variant linker sequence, the higher was the rate toward barley starch granules although all rates were within a 98−90% range of GA1 wild-type (Table 2). The correlation between

Table 3: Activation Energy, Unity Temperature, and Unfolding Enthalpy of the Irreversible Transition of the Catalytic Domain of Different Glucoamylase Forms

cell content[a]	E_a (kJ mol^{-1})	T_u (°C)	ΔH (kJ mol^{-1})
GA1 (*P.p*)	314 ± 6	65.9 ± 0.1	2021 ± 28
GA1 (*A.n*)	209 ± 2	64.9 ± 0.1	1614 ± 19
GA1 (r*A.n*)	274 ± 4	66.8 ± 0.1	2115 ± 26
Rhizopus oryzae LV	341 ± 6	62.7 ± 0.1	1683 ± 27
Humicola grisea LV	331 ± 3	63.4 ± 0.1	1798 ± 14
Pro*Hr*1 LV	360 ± 7	62.0 ± 0.1	1004 ± 17

[a] GA1 (*P.p*), wild-type GA1 produced in the methylotrophic yeast *P. pastoris*; GA1 (*A.n*), GA1 from *A. niger*; GA1 (r*A.n*), recombinant GA1 produced in *A. niger*; and the three linker variants (LV).

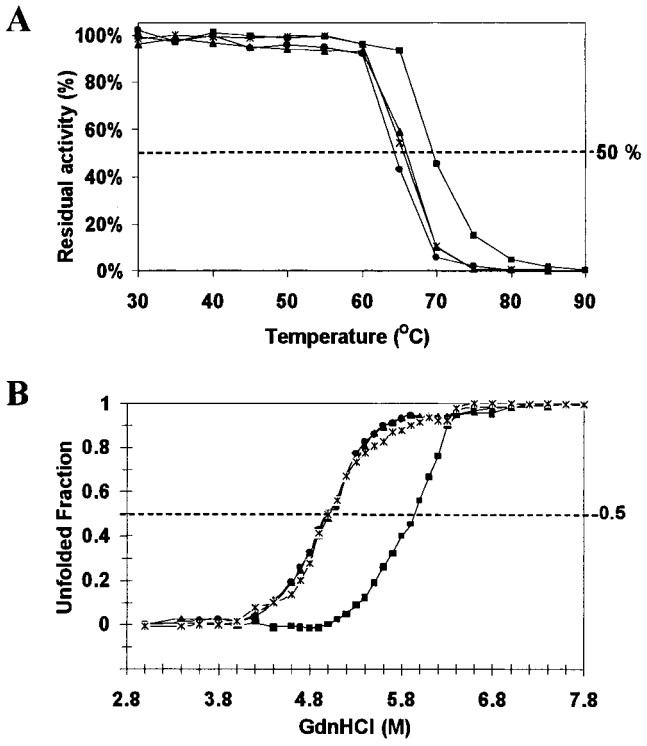

FIGURE 6: Thermostability (A) and denaturation by GdnHCl (B) for wild-type GA1 (■), GA1 Pro*Hr*1 (*), *H. grisea* (●), and *R. oryzae* (▲) linker variants produced in *P. pastoris*. The experiments were performed as described under Materials and Methods.

activity and linker length has most likely no real significance.

Stability of GA1 Linker Variants. The conformational stability of the GA linker variants was investigated by three different methods: (i) activity measurement following incubation at elevated temperatures, (ii) GdnHCl denaturation, and (iii) differential scanning calorimetry (DSC). The temperature stability of the linker variants was clearly reduced relative to that of wild-type GA1. Thus, measurements of residual activity after incubation in the range 30−90 °C showed T_m values of 66, 64, and 65 °C for the *R. oryzae*, *H. grisea*, and ProHr1 linker variants as compared to 70 °C for wild-type GA1 produced in *P. pastoris* (Figure 6A). Similarly, a decrease in the conformational stability of all variants was clearly apparent from GdnHCl unfolding studies (Figure 6B). The denaturation of the linker variants started at approximately 4.2 M GdnHCl compared to 5.0 M for wild-type GA1. Only minor differences, however, were found among the linker variants.

DSC previously showed that the catalytic domain of *A. niger* GA1 undergoes irreversible heat denaturation according

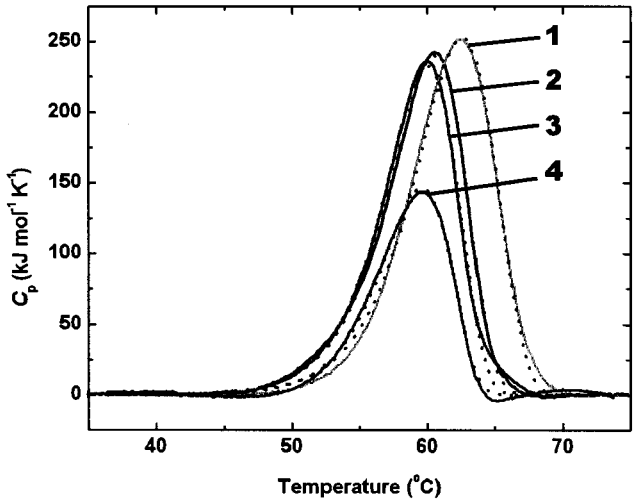

FIGURE 7: Unfolding recorded by DSC for wild-type GA1 (1), *H. grisea* (2), *R. oryzae* (3), and Pro*Hr*1 (4) linker variants produced in *P. pastoris*. The solid curves are experimentally observed DSC curves, and the dotted lines represent curve fits to a one-step irreversible unfolding mechanism. The thermodynamic parameters of the unfolding process are listed in Table 4.

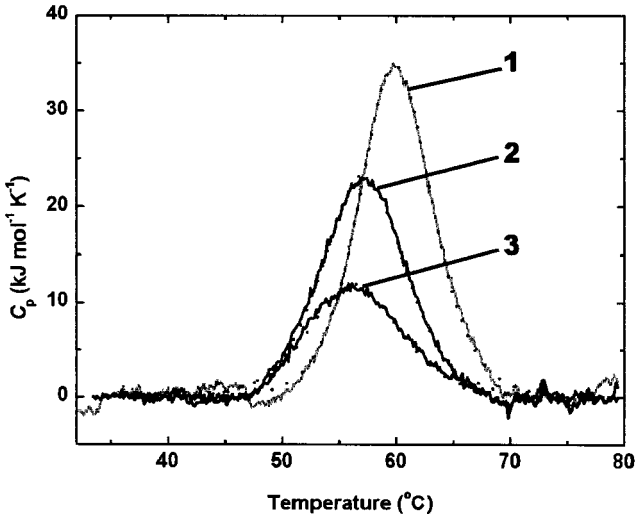

FIGURE 8: Analysis of the reversible unfolding of SBD. Solid lines are the experimentally obtained DSC curves for wild-type GA1 (1), *H. grisea* (2), and *R. oryzae* (3) linker variants produced in *P. pastoris*. The dotted lines represent fits to a non-two-state model. The calculated thermodynamic parameters are listed in Table 5.

to a one-step mechanism (*14*, *32*), while SBD unfolds reversibly. DSC profiles of GA1 wild-type and linker variants (Figure 8) are described by the activation energy, E_a, the unity temperature, T_u, which is the temperature where the rate constant of unfolding is 1 min^{-1}, and the enthalpy of unfolding, ΔH (Table 3). This also contains the enthalpy of the unfolding of SBD. It is clear that ΔH and T_u are smaller for the linker variants than for the wild-type GA1, confirming that the catalytic domains of the variants are less stable. In rescans, only one peak from the reversible unfolding of SBD apperas. This unfolding is characterized by the denaturation temperature, T_d, the calorimetric enthalpy, ΔH, and the van't Hoff enthalpy, ΔH_{vH}, which is determined from the peak shapes assuming a two-state unfolding mechanism (*33*). If the transition follows a two-state mechanism on the experimental time scale, then ΔH and ΔH_{vH} will be identical. If the transition follows a multistate mechanism, then $\Delta H_{vH} >$

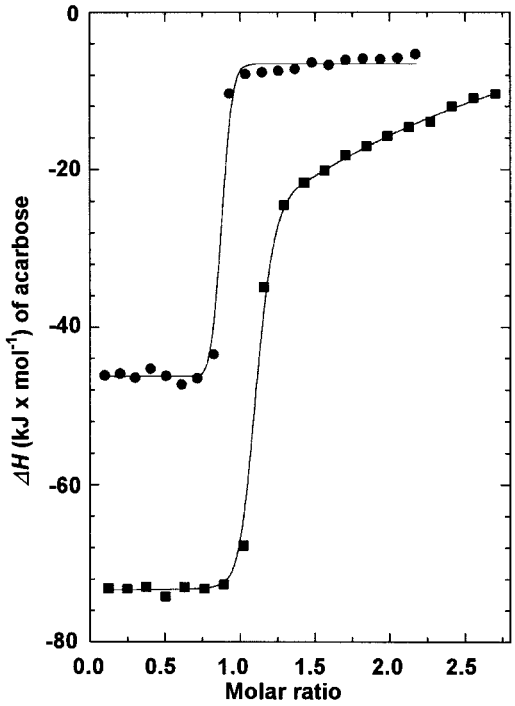

FIGURE 9: Binding isotherms of binding of acarbose (●) and L0 (■) to GA1 glucoamylase produced in *P. pastoris*.

Table 4: Denaturation Temperature and Calorimetric and van't Hoff Enthalpies of Unfolding of the Reversible Transition of SBD from Different Glucoamylases and Linker Variants

cell content[a]	T_d (°C)	ΔH (kJ mol^{-1})	ΔH_{vH} (kJ mol^{-1})
GA1 (*P.p*)	59.9 ± 0.1	281 ± 2	448 ± 3
GA1 (*A.n*)	57.0 ± 0.1	328 ± 2	477 ± 3
GA1 (r*A.n*)	58.9 ± 0.1	271 ± 1	444 ± 3
Rhizopus oryzae LV	57.2 ± 0.1	234 ± 1	372 ± 2
Humicola grisea LV	55.8 ± 0.1	171 ± 1	313 ± 3

[a] GA1 (*P.p*), wild-type GA1 expressed in the methylotrophic yeast *P. pastoris*; GA1 (*A.n*), GA1 from *A. niger*; GA1 (r*A.n*), GA1 recombinantly expressed in *A. niger*; and the two linker variants (LV).

ΔH, and if interdomain or intersubunit cooperativity takes place, then $\Delta H_{vH} < \Delta H$ (*33*). The DSC scans are shown in Figure 9 and the thermodynamic parameters are summarized in Table 4. As for the catalytic domains, the SBDs in the variants are destabilized as compared to SBD in wild-type GA1, as indicated by the lower values of T_d and ΔH. Furthermore, all SBD forms unfold according to a mechanism that is not a simple two-state mechanism since $\Delta H_{vH} > \Delta H$ (*33*).

ITC Analysis of Ligand Binding to GA1 Linker Variants. A series of ligand binding experiments with acarbose, β-CD, and the heterobidentate inhibitors L0 and L14 (Figure 1) to the variants of GA1 were carried out by using ITC. In L0, the acarbose and β-CD moieties are joined directly by a β-thioglucosidic bond between the anomeric carbon of acarbose and one of the C-6 methylene groups in β-CD, and in L14 they are connected by an oligo(ethylene glycol) spacer—$SCH_2OCH_2CH_2OCH_2CH_2S$—which in its extended conformation is 14 Å long. In the case of GA1 from *A. niger*, these double-headed ligands bind simultaneously to the active site on the catalytic domain and one of the starch binding sites on SBD to form a 1:1 complex (*17*). L0 and L14 therefore can probe if the domain−domain interaction and a functional geometry of the catalytic and starch binding sites are established in the GA1 linker variants.

Acarbose, L0, and L14 all bind to the recombinant wild-type GA1 and the *H. grisea*, *R. oryzae*, and Pro*Hr*1 GA1 linker variants with affinity given by K_a, exceeding 10^8 M^{-1} (Table 5), which is too tight by at least 1 order of magnitude for accurate measurement of the association constant (K_a). As the transition of the isotherm of acarbose binding to GA1, however, is much sharper than that of L0 binding (Figure 9), the latter interaction is substantially weaker. The binding isotherm of L0 also shows a more pronounced biphasic profile than the binding isotherm of acarbose. This is due to ligand binding of the β-CD moiety in one of the sites on SBD after L0 binding bidentately to both the catalytic site and one of the two β-CD binding sites on SBD. As a consequence of this behavior, all of the L0 and L14 binding isotherms have been fitted using a model with two independent binding sites in which the strongest site has been taken as the bidentate binding involving the active site on the catalytic domain and one of the binding sites on SBD. The association constant of acarbose is approximately 10^{12} M^{-1} (*34*, *35*). This was measured by displacement ITC using 1-deoxynojirimycin as inhibitor. In direct concentrations with the concentrations used here, the upper limit for the determination of association constants is approximately 10^8 M^{-1}. Binding of L0 and L14, however, was reported previously to be with a K_a of 10^7-10^8 M^{-1} (*17*), which is about 1 order of magnitude less than what is observed in the present work. As the previous ITC experiments (*17*) were carried out a using narrower concentration range of ligands, the secondary binding to the SBD site was not resolved. This may have led to an underestimation of the affinities of about an order of magnitude, corresponding to a free energy change of approximately 6 kJ mol^{-1}.

For the three variants and GA1 wild-type, the enthalpy of binding of L0 and L14 is approximately the sum of the enthalpies of binding acarbose and β-CD, respectively, as was also previously established for the wild-type (*17*). This means that all wild-type and variant forms examined bind L0 and L14 bidentately, i.e., in both the catalytic site and a starch binding site. The stoichiometries are in concordance with this mode of binding all close to 1:1.

In all cases, a titration with acarbose following that with L0 or L14 was also performed. These thermograms showed small peaks only corresponding to heats of dilution, thus showing that the catalytic sites were indeed filled by L0 or L14. Subsequent titrations with β-CD resulted in a 1:1 stoichiometry showing that only one of the binding sites of SBD was vacant (Table 5). Thus, both double-headed ligands bind with the acarbose moiety in the catalytic site and the β-CD moiety in one of the starch binding sites of SBD, yielding an overall 1:1 stoichiometry (*17*). This is also supported since the enthalpy contributions from binding of the acabose and β-CD separately are roughly additive for L0 and L14. L0 with no spacer arm is able to bind with both of its moieties, even to the Pro*Hr*1 linker variant having the shortest linker (Figure 2). No further attempts were made to prove that only binary complexes are formed between the heterobidentate ligands and the enzymes.

Table 5: Association Constant, Binding Energy, and Apparent Stoichiometry of the Complexation between Acarbose, β-Cyclodextrin, and the Heterobidentate Ligands L0 and L14 and Recombinant GA1 Linker Variants (LV)

cell content	ligand	K (M^{-1})	ΔG (kJ mol^{-1})	ΔH_{bind} (kJ mol^{-1})	$T\Delta S°$ (kJ mol^{-1})	apparent stoichiometry
GA1 (*P.p*)[a]	acarbose	$\geq 10^8$	≤ -46	-40.0 ± 0.4	≤ -6	0.85 ± 0.02
GA1 (*P.p*)	β-CD	$(8.6 \pm 1.3) \times 10^4$	-28.4 ± 0.4	-43.6 ± 2.5	$+15.2 \pm 2.5$	1.33 ± 0.05
GA1 (*P.p*)	L0	$\geq 10^8$	≤ -46	-72.0 ± 0.4	$\leq +26$	1.05 ± 0.01
GA1 (*P.p*) + L0	β-CD	$(3.2 \pm 0.7) \times 10^4$	-25.9 ± 0.5	-33.6 ± 2.9	$+7.7 \pm 2.9$	0.95 ± 0.10
GA1 (*P.p*)	L14	$\geq 10^8$	≤ -46	-83.8 ± 0.9	$\leq +38$	0.96 ± 0.01
GA1 (*P.p*) + L14	β-CD	$(5.0 \pm 2.0) \times 10^4$	-27.0 ± 1.0	-31.8 ± 5.7	$+4.8 \pm 5.8$	0.99 ± 0.16
H. grisea LV	acarbose	$\geq 10^8$	≤ -46	-32.0 ± 8.2	≤ -15	0.95 ± 0.25
H. grisea LV	β-CD	$(8.4 \pm 1.3) \times 10^4$	-28.3 ± 0.4	-48.1 ± 2.5	$+19.8 \pm 2.5$	1.27 ± 0.05
H. grisea LV	L0	$\geq 10^8$	≤ -46	-63.3 ± 0.7	$\leq +17$	0.92 ± 0.01
H. grisea LV + L0	β-CD	$(9.7 \pm 2.0) \times 10^3$	-22.9 ± 0.5	-49.8 ± 6.1	$+26.9 \pm 6.2$	1.20 ± 0.26
H. grisea LV	L14	$\geq 10^8$	≤ -46	-79.5 ± 1.5	$\leq +34$	0.97 ± 0.02
H. grisea LV + L14	β-CD	$(4.1 \pm 0.9) \times 10^4$	-26.5 ± 0.5	-32.2 ± 19.1	$+5.6 \pm 19.1$	0.71 ± 0.15
R. oryzae LV	acarbose	$\geq 10^8$	≤ -46	-35.9 ± 1.4	≤ -10	1.08 ± 0.02
R. oryzae LV	β-CD	$(5.3 \pm 0.5) \times 10^4$	-27.2 ± 0.2	-63.3 ± 2.6	$+36.1 \pm 2.6$	1.37 ± 0.04
R. oryzae LV	L0	$\geq 10^8$	≤ -46	-72.9 ± 0.6	$\leq +27$	1.14 ± 0.01
R. oryzae LV + L0	β-CD	$(1.8 \pm 0.2) \times 10^4$	-24.5 ± 0.3	-43.2 ± 9.7	$+18.7 \pm 9.7$	1.04 ± 0.13
R. oryzae LV	L14	$\geq 10^8$	≤ -46	-73.6 ± 1.3	$\leq +28$	1.25 ± 0.01
R. oryzae LV + L14	β-CD	$(6.7 \pm 1.2) \times 10^4$	-27.8 ± 0.5	-30.3 ± 3.1	$+2.5 \pm 3.1$	1.06 ± 0.06
Pro*Hr*1 LV	acarbose	$\geq 10^8$	≤ -46	-36.9 ± 0.5	≤ -9	0.93 ± 0.02
Pro*Hr*1 LV	β-CD	$(8.1 \pm 2.2) \times 10^4$	-28.2 ± 0.7	-41.6 ± 5.3	$+13.4 \pm 5.4$	1.50 ± 0.13
Pro*Hr*1 LV	L0	$\geq 10^8$	≤ -46	-71.7 ± 3.4	$\leq +26$	0.92 ± 0.03
Pro*Hr*1 LV + L0	β-CD	$(7.8 \pm 2.0) \times 10^3$	-22.4 ± 0.6	-79.3 ± 4.9	$+56.9 \pm 5.0$	1.00 ± 0.23
Pro*Hr*1 LV	L14	$\geq 10^8$	≤ -46	-67.8 ± 0.7	$\leq +22$	1.14 ± 0.01
Pro*Hr*1 LV + L14	β-CD	$(3.6 \pm 0.4) \times 10^4$	-26.2 ± 0.3	-35.7 ± 2.8	$+9.5 \pm 2.8$	1.04 ± 0.06

[a] GA1 (*P. p*), wild-type GA1 produced *Pichia pastoris*.

DISCUSSION

The present study describes in detail the role in function and stability of the highly glycosylated linker region connecting the catalytic and starch binding domains in glucoamylase from *A. niger*. The versatility of the bidomain architecture was investigated by replacement of Ser468−Ser508 of this region with various natural glucoamylase homologue sequences and artificial linker sequences using PCR mutagenesis. The replacing segments originated from different sources: the single case known with SBD in front of the catalytic domain *Rhizopus* glucoamylase (*20*), a thermostable glucoamylase from *Hormoconis grisea* (*21*), and a glucoamylase with unusually high activity for the α-1,6-linked substrates from *Humicola resinae* (*21*), the latter two both having SBD positioned C-terminal to the catalytic domain. The high proline content in the linker region of the thermostable glucoamylase motivated the introduction of nonnaturally occurring proline-rich sequences extending the five amino acid residue long linker of the *H. resinae* enzyme. The glucoamylase variants were characterized with respect to successful expression and secretion from the host *P. pastoris*, degree of posttranslational modifications, enzymatic properties on oligosaccharide, soluble and insoluble polysaccharide substrates, thermostability of the activity and of the conformation (DSC), conformational stability against guanidine hydrochloride, and capacity to interact with double-headed synthetic inhibitors targeted to both the catalytic and the starch binding domain.

It was shown previously for linker-truncated forms of GA1 from *A. niger* produced in *Saccharomyces cerevisiae* that such forms were not secreted but resided intracellularly when aa 466−512 were deleted. When aa 466−483 were deleted, the variant was secreted at a level of 20% of wild-type GA1 and had reduced thermostability. When aa 485−512 were deleted, the level of secretion was 60% of wild-type GA1, and the thermostability was reduced (*11*). In a similar study (*12*), a deletion of aa 470−507 prevented secretion. In the present study, indeed the shorter *H. resinae*, Pro*Hr*2, and Pro*Hr*3 linker variant GA1s representing the replacing sequences NVTSS, NVPSPSPYTSS, and NVPYPYPY, respectively, were not detected in the culture supernatant. Furthermore, slightly lower amounts of the other linker variants were produced compared to wild-type GA1. The recombinant wild-type GA1 was thus produced at approximately 400 mg/L, while linker variants yielded 200−300 mg/L. It has previously been demonstrated that the number of integrated copies of genes into *P. pastoris* affects the amounts of other recombinant proteins produced (*36*), but no attempt to make a correlation was done here. Another explanation may be that altered stability of the protein accounted for part of the decrease in the amount of protein found in the culture supernatant.

The very similar Michaelis−Menten kinetic parameters determined for the variant GA1s for the substrates maltose through maltoheptaose, soluble potato starch, and isomaltose, and their mutually similar rate of hydrolysis of barley starch granules, indicated that the catalytic site was not significantly affected by the variations in the linker sequences. Thus, the access of smaller and larger substrates to the active site appeared essentially unchanged. This is in agreement with the fact that GA1 and GA2, the latter is the natural form that lacks the SBD but has the linker region intact, hydrolyze soluble substrates with the same efficiency, indicating that SBD has no influence on catalysis using these substrates. As the ability to hydrolyze raw barley starch granules was also unaffected by the linker replacements, indeed the catalytic domain and SBD cooperated well despite the rather large structural variation of the introduced linkers.

The apparent only small impact of the connecting sequence was further substantiated by the binding thermodynamics of

the heterobidentate inhibitors L0 and L14. It was concluded that both domains of wild-type and linker variant GA1s bound simultaneously to the acarbose and the β-cyclodextrin parts of the double-headed inhibitor molecules with essentially the sum of the enthalpies of binding for the individual ligands, acarbose and β-cyclodextrin. However, in the case of the L0 inhibitor complex with Pro*Hr*1 GA1, the variant with the shortest of the replacing interdomain sequences, binding of a free β-CD molecule to the second site of SBD was more enthalpically and less entropically favorable than formation of the corresponding β-CD complexes with L0−GA1 wild-type and L14−Pro*Hr*1 GA1 linker variant, respectively. This may reflect that access was hindered to the second SBD site in the L0 complex in case of the Pro*Hr*1 variant, while binding of the spacer-extended inhibitor L14 confers a global structure to this GA1 variant bidentate inhibitor complex that allows interaction at the second SBD site as unhindered as in the case of the L0−GA1 wild-type complex. The Gibbs free energy change is not significantly affected, since the effects on the enthalpy and entropy virtually compensate each other.

It is considered that structural flexibility is allowed for the linker region without adversely affecting the function of the bidomain enzyme. This is in keeping with the hypothesis that SBD of GA1 both is confined to binding onto starches and has the capacity to disentangle the α-glucan chains of the ordered structure of solid starch to facilitate the function of the catalytic domain (*15*). The linker connecting the two domains keeps the local concentration of substrate high for the efficient catalytic degradation. It may be, however, that the catalytic domain and SBD cooperate in the dual function, as the level of GA2 activity toward granular starch in the presence of isolated SBD is lower than of GA1 (*15*).

Most remarkably, the linker variants clearly had reduced stability as compared to GA1 wild-type. Thus, the residual activity following incubation at elevated temperatures gave a 5−6 °C reduction in T_m, and GdnHCl denaturation starts 0.8 M lower than the 5 M GdnHCl for the GA1 wild-type. These two methods, however, showed no significant differences for individual linkers.

In DSC unfolding experiments, the GA1 wild-type enzyme displayed higher unity temperature T_u and enthalpy of unfolding ΔH than the linker variants. Moreover, the reversible unfolding of SBD took place with significantly lower T_d and smaller ΔH, indicating destabilization imposed by the homologue and artificial linker sequences, respectively. Remarkably, the GA1 variant Pro*Hr*1, i.e., the shorter of the variants obtained, gave no heat capacity peak in the rescan corresponding to changes due to reversible unfolding of SBD. This suggests that the short linker prevented the refolding of SBD by cooling after being unfolded. Some other variation occurred among the linker variant GA1s; thus, the unfolding of the *R. oryzae* variant with a relatively longer linker displayed slightly lower stability than both the wild-type and *H. grisea* GA1 linker variants. As the wild-type has the longest linker, this implies that linker length is not the only critical aspect of the destabilization. Glycosylation of the putative *N*-glycosylation site at the C-terminus of the *R. oryzae* linker most probably occurred as judged from the mass increase of the variant as compared to the GA1 wild-type. This perhaps contributes to the destabilization as the bulky *N*-glycosidically linked unit may interrupt the stabilizing intramolecular interactions. This is in agreement with the stability of different forms of isolated SBD being measured by DSC being dependent on the length of the retained portion of the linker sequence (*14*).

In conclusion, exchanging the most intensively *O*-glycosylated part (aa 468−508) of the linker region in GA1 from *A. niger* with homologue linkers from related bidomain fungal glucoamylases or an artificial proline-rich linker was essentially without effect on the catalytic properties of GA1. The stability of both catalytic and starch binding domains, however, was clearly decreased. Seen in that context, the linker region functions as a connecting peptide between two domains having individual capacity to work on substrates. The catalytic domain is exposed to an elevated concentration of substrate due to the fixation of SBD, which possibly also serves to facilitate access to substrate through its disentangling of glucan chains of solid starch substrates. Since the ability to use starch as a substrate is virtually unchanged, this implies that there is no strict linker−structure-dependent cooperation between the two domains in that natural function. This stabilizing effect of the linker seen on the *A. niger* GA1 has probably evolved slowly and by interaction between specific groups of the linker and the domains. The linker may also hold the domains in correct position relative to each other, allowing specific interdomain stabilizing contacts.

ACKNOWLEDGMENT

We are grateful to Sidsel Ehlers for excellent technical assistance and to Bodil Corneliussen and Lone Sørensen for performing amino acid and N-terminal sequence analyses.

SUPPORTING INFORMATION AVAILABLE

Table 1 containing the sequence of the primers used for PCR construction of the linker variants (1 page). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hiromi, K., Hamauzu, Z. I., Takahashi, K., and Ono, S. (1966) *J. Biochem. (Tokyo) 59*, 411−418.
2. Hiromi, K., Takahashi, K., Hamauzu, Z. I., and Ono, S. (1966) *J. Biochem. (Tokyo) 59*, 469−475.
3. Sierks, M. R., and Svensson, B. (1994) *Protein Eng. 7*, 1479−1484.
4. Fierobe, H.-P., Stoffer, B. B., Frandsen, T. P., and Svensson, B. (1996) *Biochemistry 35*, 8696−8704.
5. Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T., and Ottesen, M. (1982) *Carlsberg Res. Commun. 47*, 55−69.
6. Svensson, B., Larsen, K., Svendsen, I., and Boel, E. (1983) *Carlsberg Res. Commun. 48*, 529−544.
7. Sorimachi, K., Jacks, A. J., Le Gal-Coëffet, M.-F., Williamson, G., Archer, D. B., and Williamson, M. P. (1996) *J. Mol. Biol. 259*, 970−987.
8. Aleshin, A. E., Golubev, A., Firsov, L. M., and Honzatko, R. B. (1992) *J. Biol. Chem. 267*, 19291−19298.
9. Stoffer, B., Frandsen, T. P., Busk, P. K., Schneider, P., Svendsen, I., and Svensson, B. (1993) *Biochem. J. 292*, 197−202.
10. Sauer, J., Sigurskjold, B. W., Christensen, U., Frandsen, T. P., Mirgorodskaya, E., Harrison, M., Roepstorff, P., and Svensson, B. *Biochim. Biophys. Acta* (in press).
11. Libby, C. B., Cornett, C. A., Reilly, P. J., and Ford, C. (1994) *Protein Eng. 7*, 1109−1114.
12. Semimaru, T., Goto, M., Furukawa, K., and Hayashida, S. (1995) *Appl. Environ. Microbiol. 61*, 2885−2890.

13. Goto, M., Tsukamoto, M., Kwon, I., Ekino, K., and Furukawa, K. (1999) *Eur. J. Biochem. 260*, 596−602.
14. Christensen, T., Svensson, B., and Sigurskjold, B. W. (1999) *Biochemistry 38*, 6300−6310.
15. Southall, S. M., Simpson, P. J., Gilbert, H. J., Williamson, G., and Williamson, M. P. (1999) *FEBS Lett. 447*, 58−60.
16. Payre, N., Cottaz, S., Boisset, C., Borsali, R., Svensson, B., Henrissat, B., and Driguez, H. (1999) *Angew. Chem., Int. Ed. Engl. 38*, 974−977.
17. Sigurskjold, B. W., Christensen, T., Payre, N., Cottaz, S., Driguez, H., and Svensson, B. (1998) *Biochemistry 37*, 10446−10452.
18. Kramer, G. F. H., Gunning, A. P., Morris, V. J., Belshaw, N. J., and Williamson, G. (1993) *J. Chem. Soc., Faraday Trans. 89*, 2595−2602.
19. Fierobe, H.-P., Mirgorodskaya, E., Frandsen, T. P., Roepstorff, P., and Svensson, B. (1997) *Protein Expression Purif. 9*, 159−170.
20. Ashikari, T., Nakamura, Y., Tanaka, N., Kiuchi, Y., Shibano, Y., Tanaka, T., Amachi, T., and Yoshizumi, H. (1986) *Agric. Biol. Chem. 50*, 957−964.
21. Tosi, L. R. O., Francisco, H. F., and Jorge, J. A. (1993) *Can. J. Microbiol. 39*, 846−852.
22. Fagerström, R. (1991) *J. Gen. Microbiol. 137*, 1001−1008.
23. Coutinho, P. M., and Reilly, P. J. (1997) *Proteins: Struct., Funct., Genet. 29*, 334−347.
24. Clarke, A. J., and Svensson, B. (1984) *Carlsberg Res. Commun. 49*, 559−566.
25. Frandsen, T. P., Dupont, C., Lehmbeck, J., Stoffer, B., Sierks, M. R., Honzatko, R. B., and Svensson, B. (1994) *Biochemistry 33*, 13808−13816.
26. Hanahan, D. (1983) *J. Mol. Biol. 166*, 557−580.
27. Leatherbarrow, R. J. (1987) *Enzfitter, A Non-Linear Regression Data Analysis Program for IBM PC*, Elsevier Science, Amsterdam, The Netherlands
28. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L.-N. (1989) *Anal. Biochem. 179*, 131−137.
29. Bundle, D. R., and Sigurskjold, B. W. (1994) *Methods Enzymol. 247*, 288−305.
30. Best, E. A., Stedman, K. E., Bozic, C. M., Hunter, S. W., Vailes, L., Chapman, M. D., McCall, C. A., and McDermott, M. J. (2000) *Protein Expression Purif. 20*, 462−471.
31. Montesino, R., Nimtz, M., Quintero, O., Garcia, R., Falcon, V., and Cremata, J. A. (1999) *Glycobiology 9*, 1037−1043.
32. Khan, S. M. A., Reilly, P. J., and Ford, C. (2000) *Starch/Staerke 52*, 385−397.
33. Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167−241.
34. Svensson, B., and Sierks, M. R. (1992) *Carbohydr. Res. 227*, 29−44.
35. Sigurskjold, B. W., Berland, C. R., and Svensson, B. (1994) *Biochemistry 33*, 10191−10199.
36. Karras, M., Ehring, H., Nordhoff, E., Stahl, B., Strupat K., Hillenkamp, F., Grehl, M., and Krebs, B. (1993) *Org. Mass. Spectrom. 28*, 1476−1481.
37. Karras, M. and Hillenkamp, F. (1988) *Anal. Chem. 60*, 2299−2301.

BI010515I