Dill, K., & Allerhand, A. (1979) *J. Biol. Chem. 254*, 4524–4531.
Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J.-C., Fournet, B., Spik, G., & Montreuil, J. (1978) *Eur. J. Biochem. 87*, 323–329.
Eschenfelder, V., Brossmer, R., & Friebolin, H. (1975) *Tetrahedron Lett.*, 3069–3072.
Friebolin, H., Baumann, W., Keilich, G., Ziegler, D., Brossmer, R., & Von Nicolai, H. (1981) *Hoppe-Seyler's Z. Physiol. Chem. 362*, 1455–1463.
Goux, W. J., Perry, C., & James, T. L. (1982) *J. Biol. Chem. 257*, 1829–1835.
Harris, P. L., & Thornton, E. R. (1978) *J. Am. Chem. Soc. 100*, 6738–6745.
Jaques, L. W., Glant, S., & Weltner, W., Jr. (1980) *Carbohydr. Res. 80*, 207–211.
Jennings, H. J., Lugowski, C., & Kasper, D. L. (1981) *Biochemistry 20*, 4511–4518.
Keilich, G., Ziegler, D., Brossmer, R., Friebolin, H., Baumann, W., Kurz, D., Wajda, R., Weibhaar, G., & Von Nicolai, H. (1982) *Fresenius' Z. Anal. Chem. 311*, 393–394.
Kobata, A. (1979) *Anal. Biochem. 100*, 1–14.
McCleary, B. V. Taravel, F. R., & Cheetham, N. W. H. (1982) *Carbohydr. Res. 104*. 285–297.
Messer, M., Trifonoff, E., Collins, J. G., & Bradbury, J. H. (1982) *Carbohydr. Res. 102*, 316–320.
Montreuil, J., & Vliegenthart, J. F. G. (1979) *Proceedings of the International Symposium on Glycoconjugates, 4th*, Woodshole, MA, pp 35–78, Academic Press, New York.
Nunez, H. A., Matsuura, F., & Sweeley, C. C. (1981) *Arch. Biochem. Biophys. 212*, 638–643.
Prohaska, R., Koerner, T. A. W., Jr., Armitage, I. M., & Furthmayr, H. (1981) *J. Biol. Chem. 256*, 5781–5791.
Rana, S. S., Barlow, J. J., & Matta, K. L. (1982) *Carbohydr. Res. 101*, 245–253.
Roy, R., & Jennings, H. J. (1983) *Carbohydr. Res. 112*, 63–72.
Schauer, R. (1982) *Adv. Carbohydr. Chem. Biochem. 40*, 131–234.
Shashkov, A. S., Arbatsky, N. P., Derevitskaya, V. A., & Kochetkov, N. K. (1979) *Carbohydr. Res. 72*, 218–221.
Sillerud, L. O., & Yu, R. K. (1983) *Carbohydr. Res. 113*, 173–188.
Sillerud, L. O., Yu, R. K., & Schafer, D. E. (1982) *Biochemistry 21*, 1260–1271.
Srivastava, V. K., Sondheimer, S. J., & Schuerch, C. (1980) *Carbohydr. Res. 86*, 203–214.
Veh, R. W., Michalski, J.-C., Corfield, A. P., Sander-Wewer, M., Gies, D., & Schauer, R. (1981) *J. Chromatogr. 212*, 313–322.
Veluraja, K., & Rao, V. S. R. (1980) *Biochim. Biophys. Acta 630*, 442–446.
Vliegenthart, J. F. G. (1979) *Adv. Exp. Med. Biol. 125*, 77–91.
Vliegenthart, J. F. G., Van Halbeek, H., & Dorland, L. (1981) *Pure Appl. Chem. 53*, 45–77.

Neoglycoproteins: In Vitro Introduction of Glycosyl Units at Glutamines in β-Casein Using Transglutaminase†

Sau-chi Betty Yan and Finn Wold*

ABSTRACT: Exploring different methods for preparing neoglycoproteins with a specific number of oligosaccharides in specific positions, we have used guinea pig liver transglutaminase to incorporate glycosyl units into glutamine residues in β-casein. In order to prevent ε-(γ-glutamyl)lysine cross-link formation, the lysine residues of β-casein were first blocked either by amidination with ethyl acetimidate or by acylation with succinic anhydride. The glycosyl donor substrates prepared for this work were maltotriose reductively aminated with cadaverine, *N*-(Glc-Glc-glucitol-1)-cadaverine, and an asparaginyl nonasaccharide from ovalbumin modified with a 6-aminohexanoyl group at the α-amino group. The transglutaminase-catalyzed incorporation of these two donors into the β-casein derivatives was monitored in comparison to the incorporation of the commonly used transglutaminase substrate dansylcadaverine under conditions of optimal incorporation (multiple additions of enzyme, large excess of donor, and long incubation time). For both dansylcadaverine and Glc-Glc-Glc(OH)-cadaverine, 5 and 8 mol of donor were incorporated per mol of amidinated and succinylated β-casein, respectively. Competition experiments showed that the two donor substrates are incorporated into the same glutamine sites. Partial sequencing of the glycosylated β-casein permitted the identification of glutamine residues 56, 79, 167, 175, and 194 as the primary sites of incorporation in amidinated casein with residues 54 and 182 as possible sites for partial glycosylation. The results are consistent with a specific glycosylation of only selected glutamines in this transglutaminase-catalyzed process. The bulkier nonasaccharide derivative was also found to be a glycosyl donor in the transglutaminase reaction, but in this case the incorporation was lower (a maximum of 4 mol/mol) than for the other donor substrates, and multiple distinct bands were observed upon sodium dodecyl sulfate gel electrophoresis of the glycosylated product.

Lectin–sugar interactions are important in various biological recognition and communication processes such as receptor-mediated endocytosis and cell–cell interactions (Barondes, 1981). The specificity determinants of these interactions are not well understood, but recent studies, in particular those on the binding of glycoconjugates by hepatic lectins and the subsequent internalization of the glycoproteins, strongly suggest that the total process in terms of both binding and the subsequent biological consequences of binding is complex (Ashwell & Harford, 1982). These studies suggest that the efficacy of binding and clearance of glycoproteins depend not only on the

† From the Department of Biochemistry and Molecular Biology, The University of Texas Medical School, Houston, Texas 77225. *Received December 7, 1983.* This work was supported by U.S. Public Health Service Grant GM 31305 and by a grant from the Robert A. Welch Foundation.

0006-2960/84/0423-3759$01.50/0

sequence of short sugar chains but also on the spacing of multiple copies of these chains (Stowell et al., 1980; Van Lenten & Ashwell, 1972; Wong et al., 1978; Baenziger & Maynard, 1980), and it appears that the spacing of the carbohydrate moieties can be expressed both through the branching of the oligosaccharide and through the relative positions of the carbohydrate moieties on the protein backbone. One approach to investigate the role of these different aspects of glycoprotein structure in lectin–sugar interactions is to prepare neoglycoproteins with mono- or multiantennary oligosaccharides displayed either monomerically or multimerically in a specific, predetermined array. The current strategies of neoglycoprotein synthesis (Lee & Lee, 1982) are chemical methods and are mostly limited to the incorporation of mono- and dissaccharides into proteins or to the introduction of a single oligosaccharide (Mencke & Wold, 1982). If more than one oligosaccharide were to be incorporated, it would be very difficult to control the number and site of the sugar attachment by the available methods. To try to overcome this difficulty, we have developed an enzymatic method for the preparation of multivalent neoglycoproteins by exploiting the very nature of enzymic reactions, namely, their high degree of specificity.

Transglutaminase (Folk, 1983) catalyzes the reaction

$$\text{R-CONH}_2 + \text{R'-NH}_2 \overset{Ca^{2+}}{\rightleftharpoons} \text{R-CONHR'} + \text{NH}_3$$

in which R-$CONH_2$ represents the acceptor, a protein-bound glutamine, and R'-NH_2, the donor, an alkylamine; in vivo the common amine donor is probably a protein-bound lysine side chain. It appears that transglutaminase has a very stringent sequence specificity requirement for the acceptor site but can recognize a wide variety of alkylamines as donors, showing high affinity for straight chain aliphatic amines with a chain length of six carbons (Lorand et al., 1979). Because of this unique specificity characteristic, we have explored the use of guinea pig liver transglutaminase to prepare neoglycoproteins in which a specific number of oligosaccharides are incorporated into specific sites in the protein. For the studies reported here β-casein was used as the acceptor protein, and for most of the work, the artificial substrate *N*-(Glc-Glc-glucitol-1)-cadaverine was used as the donor. Much of the previous work with transglutaminase has been done with dansylcadaverine[1] as donor (Lorand et al., 1968), and all the reactions reported here were carried out in parallel with this well-characterized donor substrate for comparison.

Materials and Methods

The following chemicals were obtained from Sigma Chemical Co.: bovine casein (purified powder), cadaverine, monodansylcadaverine, and maltotriose. Fresh guinea pig livers, shipped over wet ice, were from Pel-Freeze Biologicals. Sodium cyanoborohydride was from Aldrich Chemical Co. Trypsin, TPCK treated, and *S. aureus* V8 were from Millipore Chemical Co. 9-Fluorenylmethyl choloroformate was from Pierce Chemical Co.

Preparation of Bovine β-Casein as Substrate. Bovine β-casein was purified from casein powder according to the procedure as outlined by McKenzie (1967). The β-casein was determined to be more than 95% pure by SDS–polyacrylamide gel electrophoresis and the amino-terminus determination by Edman degradation. To prevent cross-linking of β-casein by transglutaminase, the ε-amino group of all lysines of β-casein was modified either with succinic anhydride (Lorand et al., 1971; Chu et al., 1969) or ethyl acetimidate (Wofsy & Singer, 1963).

Purification of Guinea Pig Liver Transglutaminase. Transglutaminase was purified from fresh guinea pig liver according to Connellan et al. (1971) and was more than 95% pure according to SDS–polyacrylamide gel electrophoresis. The enzyme was stored frozen in small aliquots at −20 °C with no loss of activity after 1 year.

Quantitative Analysis of the Incorporation of Dansylcadaverine into β-Casein. The incorporation of dansylcadaverine into succinylated or amidinated β-casein by purified guinea pig liver transglutaminase was quantitified according to Lorand et al. (1968) with minor modifications: 60 μM β-casein was incubated with a 30-fold excess of dansylcadaverine in pH 7.5 Tris buffer containing 10 mM $CaCl_2$ at room temperature for 10 h with four additions of 25 μg of transglutaminase. The reaction was stopped by addition of trichloroacetic acid to a final concentration of 10%. The precipitated, dansylated β-casein was washed 4 times with ethanol/ether (1:1) to remove unbound dansylcadaverine. The dansylcadaverine-containing β-casein was hydrolyzed at 110 °C with 6 N HCl in vacuo. Various concentrations of dansylcadaverine were treated with the same hydrolysis conditions to serve as standards. The dansylcadaverine released by hydrolysis from the protein was quantified by fluoresence as measured in a Hitachi Perkin-Elmer MPF-2A fluoresence spectrophotometer. The absorption coefficients for β-casein and dansylcadaverine are 4.6 (1% protein) and 4.64×10^6 M^{-1} cm^{-1} at 280 and 326 nm, respectively.

Synthesis of a Simple Sugar–Alkylamine Substrate for Transglutaminase. *N*-(Glc-Glc-glucitol-1)-cadaverine [$(Glc)_2NH_2$] was synthesized by reductive amination (Gray, 1974) of maltotriose with cadaverine. Maltotriose (0.4 mmol), cadaverine (2 mmol), and sodium cyanoborohydride (10 mmol) were dissolved in 5 mL of 0.2 M phosphate buffer, pH 8.0. The reaction was incubated at 37 °C for 5 days. Excess cadaverine was used to favor the formation of the desired monosubstituted cadaverine product. Dowex 50W-X2 (25 mL) equilibrated in water was added to stop the reaction. After the H_2 evolution had subsided, the resin was washed with 500 mL of water. The resin was packed into a column, and $(Glc)_2NH_2$ was eluted by a gradient of 0–2 N NaCl. $(Glc)_2NH_2$ was detected by the phenol–sulfuric acid test (Dubois et al., 1956). The product was further purified and desalted by gel filtration chromatography on a Sephadex G-10-120 column (1.5 × 115 cm) and a Sephadex G-25-80 column (1.5 × 115 cm). The pure $(Glc)_2NH_2$ ($C_{23}H_{46}O_{15}N_2$, M_r 590) was obtained in 40% yield (1.6 mmol); its identity was confirmed by fast atom bombardment mass spectrometry in which the protonated molecular ion at 591 was the only significant peak detected.

Preparation of a Complex Oligosaccharide Derivative as Substrates for Transglutaminase. Asparaginyl glycopeptide (fraction AC-C, a mixture of $Man_7GlcNAc_2$-, $Man_5GlcNAc_4$-, and $Man_4GlcNAc_5$-) from ovalbumin was isolated by cation-exchange chromatography according to Huang et al. (1970), with modifications suggested by other workers (Tai et al., 1975; Conchie & Strachan, 1978). The amino terminus of the asparaginyl oligosaccharide was modified with 6-aminohexanoic acid by the following procedure. The amino group of 6-aminohexanoic acid was protected by 9-fluorenylmethyl chloroformate (Fmoc-Cl) according to Chang et al. (1980). Then the carboxyl group of the 6-[[[(fluorenylmethyl)oxy]carbonyl]amino]hexanoic acid was activated by *N*-hydroxysuccinimide according to the procedure of An-

[1] Abbreviations: dansyl, 5-(dimethylamino)naphthalene-1-sulfonyl; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane.

derson et al. (1964). In the final steps, the amino terminus of the asparaginyl oligosaccharide (4 μmol in 200 μL of 30 mM $NaHCO_3$) was acylated with 11.5 μmol of the activated 6-[[[(fluorenylmethyl)oxy]carbonyl]amino]hexanoic acid (dissolved in 200 μL of dioxane). After 20 h at room temperature the [6-[[[fluorenylmethyl)oxy]carbonyl]amino]hexanoyl]asparaginyl oligosaccharide was extracted with water, and the 9-fluorenylmethyl formate group was removed with piperdine/dimethylformamide (1:1) at room temperature (Anderson et al., 1964). The yield of the product, (6-aminohexanoyl)asparaginyl oligosaccharide was 2.9 μmol. Since glycosidic bonds are sensitive to acidic conditions, acid-labile amino-protecting groups such as carbobenzoxy chloride and di-*tert*-butyl dicarbonate were avoided for the above reaction scheme. Fmoc-Cl was chosen because the protected amino group can be easily deblocked by mild alkali. 6-[[[(Fluorenylmethyl)oxy]carbonyl]amino]hexanoic acid and its *N*-hydroxysuccinimide ester migrated on silica gel $G60F_{254}$TLC plates with R_f of 0.42 and 0.29, respectively. Ether was the ascending solvent. The structure of the ester was also confirmed by proton NMR (data not shown).

Sequence of Peptides by Edman Degradation. Sequencing was carried out with a Beckman Model 890D automatic sequencer using either the 0.1 M Quadrol (Beckman No. 030176) or the DMAA (Beckman No. 110377) program. Polybrene (3 mg) was used for all peptide samples. The 2-anilino-5-thiazolinone derivatives of the amino acids were converted to the corresponding 3-phenyl-2-thiohydantoins (PTH) by heating for 10 min at 80 °C with 0.2 mL of 1 N HCl containing 0.1% ethanethiol. Every effort was made to carry out the conversion as promptly as possible, but samples collected in the evening could be left for as much as 15 h prior to the conversion. The PTH-amino acids were identified by HPLC with a μBondapak C_{18} column (Waters). The Waters Model 440 liquid chromatography system was equipped with a Model 660 solvent programmer. Eluting solvents were the following: solvent A, 10 mM sodium acetate, pH 5.05, and 1% acetonitrile; solvent B, 100% acetonitrile. Curve 5 of the solvent programmer was chosen to bring the mobile phase solvent A/solvent B mixture from start (v/v 75:15) to finish (v/v 15:75) in 20 min at a rate of 1 mL/min. To identify α-PTH-ε-amidinated lysine, the final concentration of the mobile phase was increased to solvent A/solvent B of 15:100 (v/v). PTH-amino acids were quantified with the Waters WISP 710B date module system. Since the sequence of bovine β-casein was already published (Dumas et al., 1972), no attempt was made to quantify PTH-Ser or PTH-Thr by HI hydrolysis. Also because of the known primary structure of β-casein, it was possible to subject mixtures of peptides to simultaneous sequencing without extensive purification. The recovery of PTH-Gln from the spinning cup sequenator by manual conversion with 1 N HCl at 80 °C for 10 min is about 45% PTH-Gln and 45% PTH-Glu. The 2-anilino-5-thiazolinone derivatives of glycosylated Gln was not extracted from the spinning cup by butyl chloride. Thus, any glutamine residue in the β-casein that is an acceptor substrate for transglutaminase will give a zero recovery of PTH-Gln and PTH-Glu from the peptide by automated Edman degradation.

Amino Acid Analysis. Peptides or proteins were hydrolyzed in vacuo for 20 h at 110 °C with 6 N HCl containing 0.1% phenol. The amino acids were identified and quantified by a Beckman Model 121C amino acid analyzer.

Others. Slab SDS–polyacrylamide gel electrophoresis was performed according to O'Farrell (1975). Unless otherwise specified, gradient gels of 9–14% with a stacking gel of 4.75% were used. For dansylcadaverine-labeled protein, the gels were visualized with UV light without fixing or staining (Lorand et al., 1972).

Results and Discussion

The basic premise for this work is that transglutaminase is specific for only a limited number of glutamine residues in the acceptor protein. Although little is known about the specificity determinants for the transglutaminase reaction, the data in the literature certainly are consistent with the premise that the enzyme indeed is highly specific (Folk, 1983). Caution must be exercised when interpreting these data. Most transglutaminase assays with native acceptor proteins have not been conducted for the purpose of achieving maximum incorporation of donor. In some studies, the data could be misleading in that the (γ-glutamyl)lysine cross-links have not been determined. As a general rule, however, it appears that for native acceptor proteins, only one or a few (indeed, sometimes none) glutamines are modified (Brenner & Wold, 1978). When the same acceptor proteins are denatured, incorporation of donor increases but remains much lower than predicted if all glutamines are reactive. The tentative model then is that a certain characteristic sequence information is the minimum requirement for the modification of a given glutamine by transglutaminase and that this proper sequence may be cryptic unless it is properly "exposed" through conformational manipulations (denaturation).

To examine the use of transglutaminase in the preparation of neoglycoproteins, several aspects of the reaction needed to be established as outlined by the following questions: Will simple sugar–alkylamine and complex oligosaccharide–alkylamine derivatives act as amine donors for the enzyme Are they as efficient in terms of incorporation as is the well-characterized dansylcadaverine and are they incorporated into the same acceptor sites? How many of the 21 glutamine residues in β-casein are active acceptor substrates in the "native" and in the denatured protein? If any specificity is observed, does it reflect the desired all-or-none effect (certain residues are always completely modified, others not modified at all) or a possible spectrum of reactivities in which residues are factionally modified in a mostly random selection? The results are presented to provide answers to these questions.

Bovine β-casein was selected as acceptor protein for this work because the primary sequence of this protein is known (Dumas et al., 1972) and because it has been a widely used substrate for transglutaminase. To prevent cross-link formation, all the incorporation experiments were carried out with β-casein in which all the lysine residues had been either amidinated or succinylated. Amidination of proteins is a mild process, and the charges of the protein remain unchanged. Thus, amidinated β-casein should retain most, if not all, of the native conformation and represents our native substrate. Succinylation on the other hand reverses the charge of lysine side chain; at least for some proteins this drastic change in charge has been shown to result in unfolding (Habeeb, 1967), and we consider this our denatured substrate.

A typical experiment involving the glycosylation of β-casein with $(Glc)_2NH_2$ is shown in Figure 1, and the comparative extent of incorporation of $(Glc)_2NH_2$ and dansylcadaverine is summarized in Table I. The results in Table I show that succinylated β-casein definitely incorporated more monodansylcadaverine and $(Glc)_2NH_2$ than did amidinated β-casein and that the same amount of $(Glc)_2NH_2$ was incorporated into succinylated and amidinated β-casein as dansylcadaverine. The fact that $(Glc)_2NH_2$ was shown to compete effectively with monodansylcadaverine for acceptor sites on bovine β-

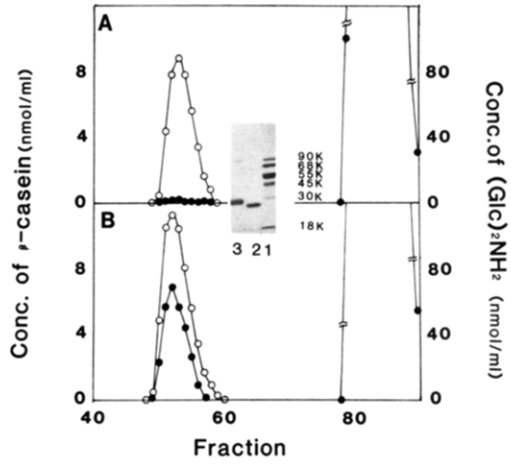

FIGURE 1: Glycosylation of β-casein with $(Glc)_2NH_2$ using the transglutaminase reaction. Amidinated β-casein (60 μM) was incubated with a 300-fold excess of $(Glc)_2NH_2$ in pH 7.5 Tris buffer containing 10 mM $CaCl_2$ at room temperature for 10 h with (panel B) or without (panel A) four additions of 25 μg of transglutaminase. The protein (○) was separated from unreacted $(Glc)_2NH_2$ (●) by Sephadex G-25-80 gel filtration chromatography. $(Glc)_2NH_2$ concentration was measured by the phenol–sulfuric acid test and β-casein concentration by its absorbance at 280 nm or by amino acid analysis. Analysis of $(Glc)_2NH_2$ content of fractions making up the protein peak in (B) gave internally consistent values averaging 5.2 mol of $(Glc)_2NH_2$/mol of β-casein. Peak protein fraction of panel A (lane 2) and panel B (lane 3) were analyzed by SDS–polyacrylamide gel electrophoresis. The glycosylated β-casein migrated with an apparent molecular weight of 30 000 in comparison with M_r 24 000 for unmodified β-casein. [The position of (dansylcadaverine)$_5$-β-casein was identical with that of the unmodified control.] Lane 1 contains molecular weight standards: transferrin, M_r 90 000; bovine serum albumin, M_r 68 000; amylase, M_r 55 000 and 52 000; ovalbumin, M_r 45 000; DNase, M_r 30 000; myoglobin, M_r 17 900. Glycosylation of succinylated β-casein was performed in exactly the same fashion; the results are summarized in Table I.

Table I: Incorporation of Monodansylcadaverine or $(Glc)_2NH_2$ into β-Casein by Transglutaminase[a]

	mol incorporated/mol of β-casein	
	dansylcadaverine	$(Glc)_2NH_2$
amidinated β-casein	5.1	5.2
succinylated β-casein	7.9	7.8

[a] The ε-amino group of all the lysines of β-casein were modified by either succinic anhydride or ethyl acetimidate. The procedures for quantifying dansylcadaverine and $(Glc)_2NH_2$ incorporation are described under Materials and Methods and in the legend to Figure 3, respectively. There was no further increase in incorporation of dansylcadaverine after an additional 6 h of reaction or with further addition of transglutaminase.

casein (Figure 2) provides strong evidence that the two different donors are incorporated at the same glutamine acceptor sites. From these results we conclude that simple sugar derivatives can be effectively incorporated into proteins by transglutaminase, that there are glutamine residues in the native β-casein that are inaccessible to transglutaminase but are exposed and can be modified in the "denatured" β-casein, and that in both cases a homogeneous product containing five and eight $(Glc)_2NH_2$, respectively, appears to be the end product of the reaction.

For this method to have general application in neoglycoprotein preparation, it should also be possible to use complex oligosaccharides as donors. When the asparaginyl oligosaccharide (fraction AC-C) from ovalbumin (Tai et al., 1975) derivatized by 6-aminohexanoic acid was used as donor, it was found to be incorporated less efficiently than was $(Glc)_2NH_2$. Amidinated β-casein, after prolonged incubation with this derivative and transglutaminase, was analyzed by SDS–PAGE

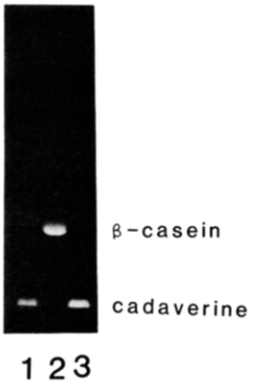

FIGURE 2: Test for competition between $(Glc)_2NH_2$ and dansylcadaverine for acceptor sites in β-casein. Amidinated β-casein (60 nmol) was incubated with dansylcadaverine (500 nmoles) and transglutaminase in the presence (lane 1) or absence (lane 2) of $(Glc)_2NH_2$ (3000 nmol) under the reaction conditions described under Materials and Methods. An aliquot of the reaction was fractionated by gradient sodium dodecyl sulfate–polyacrylamide gel electrophoresis. The gel was visualized under UV light without fixing or staining. Lane 3 represents a control reaction without $CaCl_2$.

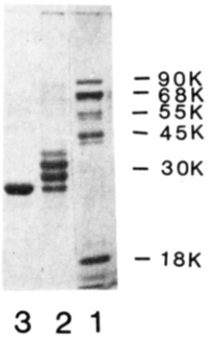

FIGURE 3: Incorporation of complex oligosaccharide into amidinated β-casein by transglutaminase reaction. Amidinated β-casein (16 nmol) was incubated with 1.25 μmol of (6-aminohexanoyl)asparaginyl oligosaccharide at room temperature in 50 mM Tris buffer and 10 mM $CaCl_2$, pH 7.5, for 20 h with 5 μg of transglutaminase (added in 1-μg increments every 4 h). An aliquot of the reaction was analyzed by SDS–PAGE (lane 2). A control reaction without transglutaminase was carried out under the same conditions (lane 3). The molecular weight standards in lane 1 are the same as those given in Figure 1.

(Figure 3). In addition to the native β-casein band, there were four higher molecular weight protein bands, suggesting that as many as 4 mol of oligosaccharide were incorporated per mol of amidinated β-casein. All the four higher molecular weight species bound tightly to a concanavalin A–Sepharose column and could be eluted with 100 mM methyl α-mannoside (data not shown). In contrast to the reaction with $(Glc)_2NH_2$, however, we were unable to push the glycosylation reaction of amidinated β-casein with the oligosaccharide to completion (i.e., 5 mol of oligosaccharide/mol of casein). The reason for this finding is not immediately clear to us; the most likely explanation is that the total donor group in this case is so large that the incorporation at one site precludes incorporation at a neighboring site. This explanation may also be consistent with the observation of multiple products; if each step of incorporation at the five available sites is random, each step would determine which and how many of the remaining sites would be accessible for subsequent steps.

The most critical issue regarding the utility of this glycosylation reaction is the specificity question. Are the donor groups linked to distinct glutamine residues or are they fractionally distributed on all the 21 glutamines of β-casein in a random fashion? In order to answer this question, the modified β-casein and its fragmentation products were sub-

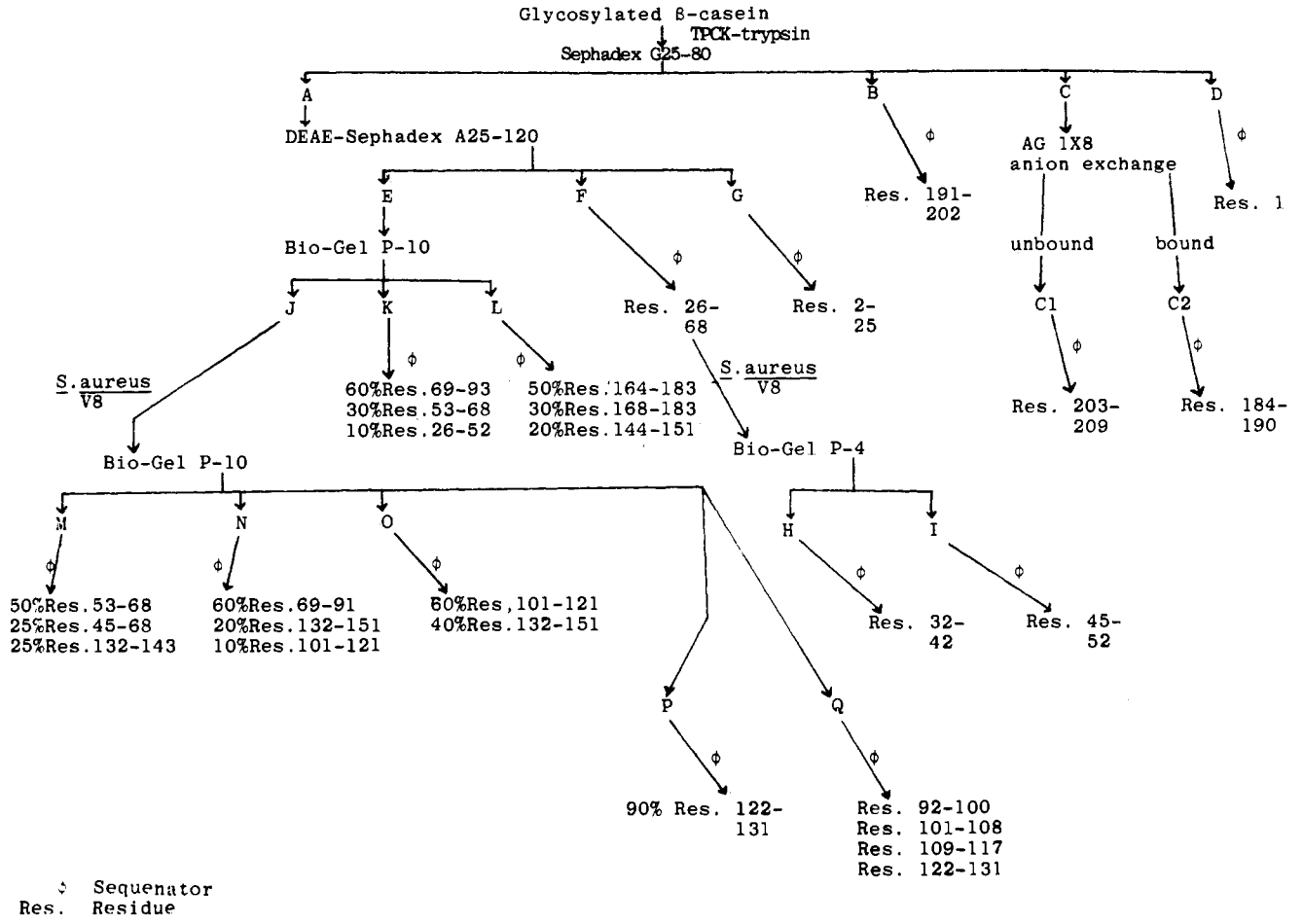

FIGURE 4: Fragmentation of glycosylated β-casein by trypsin and *S. aureus* V8 and the fractionation of the peptides. A total of 900 nmol of amidinated β-casein was glycosylated with $(Glc)_2NH_2$ in the presence of guinea pig liver transglutaminase. The glycosylated β-casein was separated from excess $(Glc)_2NH_2$ and transglutaminase by Sephadex G-75-120 gel filtration. The glycosylated β-casein was digested with TPCK–trypsin (1% w/w) for 4 h at 37 °C in 50 mM Tris, pH 7.5. Tryptic peptides were subjected to purification by different column chromatography or further digestion by *S. aureus* V8. Digestions with *S. aureus* V8 were carried out at 37 °C for 6 h in 50 mM NH_4HCO_3, pH 7.8.

jected to sequencing. Amidinated β-casein modified with dansylcadaverine was only sparingly soluble in aqueous solution, and tryptic peptides of the dansylated protein were also very insoluble, making it very difficult to locate the acceptor sties in this derivative. We considered it a special bonus of the glycosylation work when it was found that the derivative modifed with $(Glc)_2NH_2$ exhibited no such solubility problems, and all the sequence data presented here for locating the acceptor sites in amidinated β-casein were consequently obtained with $(Glc)_2NH_2$-modified casein.

PTH-glutamine can be totally converted to PTH-glutamate in 1 N HCl at 80 °C in 1 h, and to prevent deamidation of glutamine residues in the glycosylated β-casein, acidic conditions such as cyanogen bromide fragmentation were avoided. Preparation of peptides for sequencing was consequently limited to proteolytic digestion around neutral pH and is summarized in Figure 4. Since the lysine residues had been amidinated, TPCK–trypsin should cleave only at the arginine residues. However, we found that there were many unexpected and often incomplete cleavages in addition to the cleavages at the C-terminus of arginines, making the peptide purification a formidable task. These nonspecific cleavages of β-casein by TPCK–trypsin has been observed by other workers (Richardson & Mercier, 1979). As a result of this phenomenon, some of the data were obtained with peptide mixtures, and as a consequence we have less confidence in the data for residues 54, 56, 72, 79, 89, 141, 146, 149, 167, 175, and 182 than for the other ones. Furthermore, we were unable to recover peptide 158–163 from fractionation of tryptic peptides and consequently have no information regarding Gln-160.

In spite of these considerations, we feel that the data summarized in Table II provide convincing evidence that the incorporation of $(Glc)_2NH_2$ is quite specific. Five glutamine residues (56, 79, 167, 175, and 194) are more than 90% modified; residues 54 and 182 appear to be partially modified while the rest of the glutamines gave close to a 100% recovery of PTH-Gln + PTH-Glu. The high recovery of PTH-Glu + PTH-Gln for residues 38, 39, and 40 is due to the carry-over effect of the preceding glutamine or glutamate residues in sequencing the peptide by Edman degradation (see Figure 5). It seems reasonable to propose that the two partially glycosylated sites (54 and 182) in the amidinated protein correspond to two of the three extra acceptor sites observed for the succinylated derivative. If this is so, it suggests that these sites have the proper covalent structure as substrates for transglutaminase and that they because of conformational differences in amidinated and succinylated β-casein are only partially exposed in the former, while fully exposed in the latter. There are some discrepancies in sequence assignments between our data and the original published work (Dumas et al., 1972) as shown in Figure 5. These disagreements have been suspected previously by others (Richardson & Mercier, 1979) and are now substantiated by our data.

By examining the sequences around the seven acceptor sites in β-casein (Figure 5), we noticed that there is an amino acid with an electron-rich side chain right next to each of the

Table II: Summary of Sequence Data on the Recovery of Glutamine Residues from Glycosylated β-Casein

glutamine residue	% recovery of glutamine as PTH-E + PTH-Q	acceptor[a] site	no.[b]	peptide fragment[c]
34	60 + 34 = 94	–	1	peptide F
38	111 + 50 = 165	–	1	peptide F
39	74 + 53 = 127	–	1	peptide F
40	50 + 66 = 116	–	1	peptide F
46	54 + 56 = 110	–	1	peptide I
54	24 + 34 = 68	±	3	peptide pool M
56	6 + 9 = 15	+	2	peptide pool M
72	26 + 55 = 81	–	1	peptide pool K
79	3 + 2 = 5	+	1	peptide pool K
89	28 + 79 = 107	–	2	peptide pool K and N
123	38 + 55 = 93	–	2	peptide P
141	25 + 57 = 82	–	1	peptide pool O
146	32 + 95 = 127	–	2	peptide pool O
149	34 + 44 = 78	–	1	peptide pool O
160	no data			
167	3 + 8 = 11	+	2	peptide pool L[f]
175	6 + 4 = 10	+	2	peptide pool L[d]
182	13 + 22 = 35	±	2	peptide pool L
188	30 + 70 = 100	–	1	peptide C2
194	0 + 0 = 0	+	2	peptide B
117	100 + 0		1	peptide pool O[e]
195	100 + 0		3	peptide B[e]

[a] (–) not acceptor; (+) acceptor; (±) partial acceptor. [b] Number of trials. [c] Peptide fragment from which sequence data were obtained. Preparation and designation of the peptide fragments are described in detail in Figure 4. [d] This residue was assigned as glutamic acid in the original sequence work published by Dumas et al. (1972). Our data indicated it was glutamine and was glycosylated. [e] These residues were assigned as glutamine in the original sequence work by Dumas et al. (1972). Our data showed that these residues were glutamic acids instead. [f] This glutamine in native β-casein was identified as an acceptor for factor XIIIA, and this same glutamine in a synthetic pentadecapeptide corresponding to residue 161–175 of β-casein was an acceptor for guinea pig liver transglutaminase (Folk, 1983).

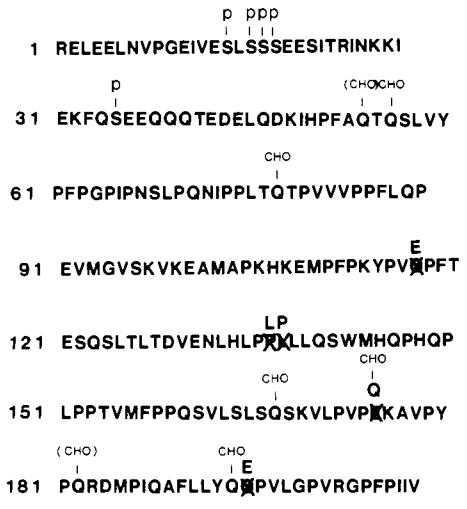

FIGURE 5: Location of glycosylated glutamine residues in the sequence of β-casein. The sequence of bovine β-casein as reported by Dumas et al. (1972). β-Casein is a phosphoprotein, and the sites of phosphorylation are designated with "P". Glutamines that were glycosylated with $(Glc)_2NH_2$ are indicated in the sequence with "CHO". Glutamines-54 and -182 were partially glycosylated. "X" marks the position where there is a discrepancy in sequence assignment between our data and the original published work by Dumas et al. (1972). The residue found in this work is inserted above the X.

acceptor glutamine residues. Five of these seven adjacent amino acids are hydroxy amino acids, namely, serine, threonine, or tyrosine. These seven acceptor sites are located in either random coil or β-bend of the predicted secondary structures of β-casein (Creamer et al., 1981). Space filling models of -Gln-Ser-, -Ser-Gln-, -Gln-Tyr-, and -Tyr-Gln- indicate that the hydroxy group of serine or tyrosine can be brought close enough to interact with the amide group of the glutamine side chain as long as the dipeptide is not engaged in α-helix or β-sheet structure. Further work will determine whether these observations are relevant as specificity determinants in the acceptor substrate of guinea pig liver transglutaminase.

Acknowledgments

We are grateful to Dr. Richard M. Caprioli of The University of Texas Health Science Center Analytical Chemistry Center for providing the FAB-mass spectrometer analyses, to Christopher C. Q. Chin of the UTHSC Protein Sequencing Laboraory for advise and assistance in carrying out the sequencing work, and to Anne DeForest for preparing the manuscript.

References

Anderson, G. W., Zimmerman, J. E., & Callahan, F. M. (1964) *J. Am. Chem. Soc. 86*, 1839–1842.

Ashwell, G., & Harford, J. (1982) *Annu. Rev. Biochem. 51*, 531–554.

Baenziger, J. V., & Maynard, Y. (1980) *J. Biol. Chem. 255*, 10001–10012.

Barondes, S. H. (1981) *Annu. Rev. Biochem. 50*, 207–231.

Brenner, S. C., & Wold, F. (1978) *Biochim. Biophys. Acta 522*, 74–83.

Chang, C. D., Waki, M., Ahmad, M., Meienhofer, J., Lundell, E. O., & Haug, J. D. (1980) *Int. J. Pept. Protein Res. 15*, 59–66.

Chu, F. S., Crary, E., & Bergdoll, M. S. (1969) *Biochemistry 8*, 2890–2896.

Conchie, J., & Strachan, I. (1978) *Carbohydr. Res. 63*, 193–213.

Connellan, J. M., Chung, S. I., Whetzel, N. K., Bradley, L. M., & Folk, J. E. (1971) *J. Biol. Chem. 246*, 1093–1098.

Creamer, L. K., Richardson, T., & Parry, D. A. D. (1981) *Arch. Biochem. Biophys. 211*, 689–696.

Dubois, M., Gilles, K. A., Hamilton, J. B., Rebers, P. H., & Smith, F. (1956) *Anal. Biochem. 28*, 350–356.

Dumas, B. R., Brignon, G. M., Grosclaude, F., & Mercier, J. C. (1972) *Eur. J. Biochem. 25*, 505–514.

Folk, J. E. (1983) *Adv. Enzymol. Relat. Areas Mol. Biol. 54*, 1–56.

Gray, G. R. (1974) *Arch. Biochem. Biophys. 163*, 426–428.

Habeeb, A. F. S. A. (1967) *Arch. Biochem. Biophys. 121*, 652.

Harford, J., & Ashwell, G. (1982) in *The Glycoconjugates* (Horowitz, M. I., Ed.) Vol. IV, pp 27–55, Academic Press, New York.

Huang, C. C., Mayer, H. E., & Montgomery, R. (1970) *Carbohydr. Res. 13*, 127–137.

Lee, Y. C., & Lee, R. T. (1982) in *The Glycoconjugates* (Horowitz, M. I., Ed.) Vol. IV, pp 57–83, Academic Press, New York.

Lorand, L., Rule, N. G., Ong, H. H., Furlanetto, A., Downey, J., Oner, N., & Bruner-Lorand, J. (1968) *Biochemistry 7*, 1214–1223.

Lorand, L., Lockridge, O. M., Campbell, L. K., Myhrman, R., & Bruner-Lorand, J. (1971) *Anal. Biochem. 44*, 221–231.

Lorand, L., Chenoweth, D., & Gray, A. (1972) *Ann. N.Y. Acad. Sci. 202*, 155–171.

Lorand, L., Parameswaran, K. N., Stenberg, P., Tong, Y. S., Velasco, P. T., Jonsson, N. A., Mikiver, L., & Moses, P. (1979) *Biochemistry 18*, 1756–1765.

McKenzie, H. A. (1967) *Adv. Protein Chem. 22*, 55–234.
Mencke, A. J., & Wold, F. (1982) *J. Biol. Chem. 257*, 14799–14805.
O'Farrell, P. T. (1975) *J. Biol. Chem. 250*, 4006–4021.
Richardson, B. R., & Mercier, J. C. (1979) *Eur. J. Biochem. 99*, 285–297.
Stowell, C. P., Lee, R. T., & Lee, Y. C. (1980) *Biochemistry 19*, 4904–4908.
Tai, T., Yamashita, K., Ogata-Arakawa, M., Koide, N., Muramatsu, T., Iwashita, S., Inoue, Y., & Kobata, A. (1975) *J. Biol. Chem. 250*, 8569–8675.
Van Lenten, L., & Ashwell, G. (1972) *J. Biol. Chem. 247*, 4633–4640.
Wofsy, L., & Singer, S. J. (1963) *Biochemistry 2*, 104–116.
Wong, K. L., Debanne, M. T., Hatton, M. W. C., & Regoeczi, E. (1978) *Int. J. Pept. Protein Res. 12*, 27–37.

Photoaffinity Labeling of the α_1-Adrenergic Receptor Using an ^{125}I-Labeled Aryl Azide Analogue of Prazosin†

Christine E. Seidman, Hans-Jürgen Hess, Charles J. Homcy,* and Robert M. Graham

ABSTRACT: α_1-Adrenergic receptor probes, which can be radioiodinated to yield high specific activity radioligands, have been synthesized and characterized. 2-[4-(4-Aminobenzoyl)piperazin-1-yl]-4-amino-6,7-dimethoxyquinazoline (CP63,155), an arylamine analogue of the selective α_1-adrenergic antagonist prazosin, and its iodinated derivative, 2-[4-(4-amino-3-[^{125}I]iodobenzoyl)piperazin-1-yl]-4-amino-6,7-dimethoxyquinazoline ([^{125}I]CP63,789), bind reversibly and with high affinity (K_D = 1 nM and 0.6 nM, respectively) to rat hepatic membrane α_1-adrenergic receptors. Conversion of [^{125}I]CP63,789 to the aryl azide yields a photolabile derivative, 2-[4-(4-azido-3-[^{125}I]iodobenzoyl)piperazin-1-yl]-4-amino-6,7-dimethoxyquinazoline ([^{125}I]CP65,526), which prior to photolysis binds competitively and with high affinity (K_D = 0.3 nM). Binding of [^{125}I]CP63,789 and [^{125}I]CP65,526 (prior to photolysis) is rapid and saturable. Both ligands identify similar α_1-adrenergic receptor binding site concentrations as the parent probe, [^{3}H]prazosin. Specific binding by these iodinated ligands is stereoselective and inhibited by a variety of adrenergic agents with a specificity typical of the α_1-adrenergic receptor. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and autoradiography of [^{125}I]CP65,526-labeled rat hepatic membranes reveal major protein species with molecular weights of 77K, 68K and 59K. Each protein binds adrenergic ligands with stereoselectivity and with a specificity typical of the α_1-adrenergic receptor. Inclusion of multiple protease inhibitors during membrane preparation prior to SDS–PAGE does not alter the labeling of these peptides. Smaller peptides with molecular weights of 42K and 31K display prazosin-inhibitable [^{125}I]CP65,526 binding. Labeling of these protein species with [^{125}I]CP65,526 is not inhibitable by other adrenergic agonists or antagonists. They are thus unlikely to represent subunits of the receptor. These findings confirm and extend our observations on the subunit composition of the receptor determined with the purified protein and indicate the utility of these novel high-affinity radioiodinated probes as tools for more detailed elucidation and comparison of the molecular properties of the receptor in a variety of tissues.

Considerable progress has been made in the isolation, purification, and molecular characterization of a number of hormone and drug receptors (Homcy et al., 1983; Momoi & Lennon, 1982; Schneider et al., 1982) including the α_1-adrenergic receptor (Graham et al., 1982a,b). To further aid in the molecular characterization of this subtype of the α-adrenergic receptor, we synthesized and characterized a photoaffinity label, which upon photolysis covalently and specifically incorporates into the receptor binding site (Hess et al., 1983).

We here report on the development of a high-affinity α_1-selective probe, which can be radioiodinated and purified by high-performance liquid chromatography (HPLC) to yield a compound of high specific activity (2175 Ci/mmol, assumed). Conversion of this radioiodinated derivative to the aryl azide yields a radiolabeled photoaffinity probe, selective for the α_1-adrenergic receptor. This probe can be used to readily identify and characterize the minute quantities of receptor present in most tissues and cells.

Materials and Methods

Carrier free Na^{125}I was purchased from Amersham, Arlington Heights, IL. Premixed sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) standards were purchased from Pharmacia and iodinated according to the method of Hunter & Greenwood (1962). Phenylmethanesulfonyl fluoride (PMSF), sodium azide (NaN_3), bacitracin, and soybean trypsin inhibitor (STI) were from Sigma. Sodium nitrite ($NaNO_2$) was obtained from Fisher, acetonitrile (CH_3CN) was from Baker, and 24-mm filters (no. 32 glass) were from Schleicher & Schuell. X-ray film (XAR-5) was from Kodak and was developed in an X-omat M20 processor (Kodak). Image-intensifying screens were from Du Pont. The

† From the Cellular and Molecular Research Laboratory, Cardiac Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114 (C.E.S., C.J.H., and R.M.G.), and Pfizer Central Research, Groton, Connecticut 06340 (H.-J.H.). *Received November 1, 1983.* C.E.S. is supported by NIH Training Grant HL-070208. C.J.H. and R.M.G. are Established Investigators, American Heart Association (Grants AHA 80-148 and 82-240, respectively). These studies were supported in part with NIH Grants HL-19259 and NS-19583, a grant-in-aid from the American Heart Association (83-1242, in part with funds from the Massachusetts Affiliate), and a grant from the R.J. Reynolds Co.

* Address correspondence to this author at the Cellular and Molecular Research Laboratory, Cardiac Unit, Massachusetts General Hospital, Boston, MA 02114.

0006-2960/84/0423-3765$01.50/0