Introduction of an *N*-Glycosylation Cassette into Proteins at Random Sites: Expression of Neoglycosylated FGF

BioTechniques 27:576-590 (September 1999)

Atsuko Yoneda, Masahiro Asada, Masashi Suzuki and Toru Imamura

National Institute of Bioscience and Human Technology, Ibaraki, Japan

ABSTRACT

We developed a method for introducing an N-glycosylation cassette into proteins at random sites by constructing cDNAs and expressing it in mammalian cells. The protocol entails four steps: (i) generation of cDNAs that contain single, randomly-located blunt end cuts; (ii) ligation of N-glycosylation cassettes into the blunt end cuts in three-frame formats; (iii) selection of the cDNA clones encoding N-glycosylated proteins; and (iv) subcloning into an expression vector for transfection and expression in mammalian cells. This method was evaluated using secreted fibroblast growth factor (FGF) as a model protein. Several secreted FGF cDNA clones, each containing an AsnLeuSer-coding sequence at a random site, were obtained. When these clones were expressed in mammalian cells, some of the secreted FGFs were found to be N-glycosylated. The method described here should also be applicable for random introduction of functional oligopeptide/polypeptide cassettes into virtually any protein of interest.

INTRODUCTION

Earlier studies have shown that the oligosaccharide moieties of glycoproteins play crucial roles in a variety of cellular processes. Protein glycosylation modifies many of the characteristics of the protein core, including its secretion, stability, clearance, interactions with receptors or ligands and oligomerization (13).

Introduction of *N*-glycosylation sequence(s) into selected proteins, by intentional substitution or insertion of amino acids, has provided substantial information about the functions of oligosaccharides in glycoproteins and about the structures of the core proteins. Such studies have included: (i) investigation of the two-dimensional topological arrangement of intrinsic membrane proteins in the lipid bilayer (12); (ii) restoration of lectin activity by introduction of a novel *N*-glycosylation sequence into a nonactive, nonglycosylated ricin B chain mutant (14); (iii) mapping domains for ligand binding by a low-affinity nerve growth factor receptor (p75) (1); and (iv) assessment of lysozyme stabilization (11). The results from these studies all indicate that introduction of *N*-glycan might be a promising method for obtaining protein mutants with moderately altered characteristics.

To date, however, remodeling proteins by *N*-glycosylation has required detailed information about the molecular structure of the subject protein and time-consuming construction of a large number of mutants. Our goal was to develop a method for introducing *N*-gly-

cosylation sequences into proteins without having to pinpoint the insertion site. With the protocol described here, we successfully used a secreted form of fibroblast growth factor (FGF) as the target protein and a potential *N*-glycosylation sequence as the inserted functional sequence/domain. Further, this protocol should also be applicable to other proteins and other functional sequence/domain of interest.

MATERIALS AND METHODS

Deoxyribonuclease I (DNase I), restriction enzymes, SUPREC™-01 and DNA Ligation Kit version 2 were purchased from Takara Shuzo (Tokyo, Japan). SeaKem® GTG® agarose was from FMC BioProducts (Rockland, ME, USA). Cloned *Pfu* DNA Polymerase was from Stratagene (La Jolla, CA, USA). GENECLAN® II was from BIO 101 (Vista, CA, USA). DNA Sequencing Kit Dye Terminator Cycle Sequencing Ready Reaction with AmpliTaq® DNA Polymerase FS was from PE Biosystems (Foster City, CA, USA). *E. coli* DH5α was from TOYOBO (Osaka, Japan). DIG DNA Labeling Kit, DIG Luminescent Detection Kit, digoxigenin-POD Fab fragment and recombinant peptide *N*-glycosidase F (PNGase) were from Roche Molecular Biochemicals (Basel, Switzerland). Heparin sepharose-CL6B and Hybond™-N+ was from Amersham Pharmacia Biotech (Uppsala, Sweden). Manganese chloride, chloramphenicol sulfate and ampicillin sulfate were from Wako Pure

Research Report

Chemicals (Osaka, Japan). Geneticin disulfate salt and RPMI 1640 powder were from Sigma (St Louis, MO, USA). Fetal bovine serum (FBS) was from Filtron Pty. Ltd. (Brooklyn, Australia).

Cell Culture

Mouse pro-B Ba/F3 cells were purchased from the Riken gene bank (Tsukuba, Ibaraki, Japan). The cells were grown in RPMI 1640 supplemented with 10% FBS and 10% WEHI-3 cell-conditioned medium (Riken gene bank) used as a source of interleukin-3.

Construction of a Model Protein to Introduce *N*-Glycosylation Sequence

An oligonucleotide encoding human FGF-1 α was amplified by polymerase chain reaction (PCR) using primers No. 967 (5'-gcgtc gacag cgcta attac aagaa gccc aactc-3') and No. 630 (5'-ccgaa ttgga attct ttaat cagaa gagac tgg-3') and human FGF-1 cDNA as a template. The PCR product was digested with *Sal*I and *Eco*RI and subcloned into the *Sal*I/*Eco*RI site of pBluescript® II KS+ (Stratagene). This construct was designated FGF-1 α /pBS. A double-stranded (ds)DNA encoding a Kozak sequence and the signal peptide of mouse FGF-6 was amplified by PCR using primers No. 1048 (5'-gcgtc gaccc accat gtccc gggga gcagg acgtg ttacg ggcac gctgc aggtc ctctg ctcc-3') and No. 968 (5'-gcgat atcca gtacg gtgcc gttgg cgcg-3') and mouse FGF-6 cDNA as a template. The PCR product was digested with *Sal*I and *Eco*RV and subcloned into the *Aor*51HI and *Sal*I sites of FGF-1 α /pBS. This construct, encoding a secreted FGF [secFGF(N₃₅)] that contains an *N*-glycosylation sequence derived from the original FGF-6 signal peptide, was designated secFGF(N₃₅)/pBS. As our purpose was to detect *N*-glycosylation on the cassette newly introduced into the protein, the endogenous Asn₃₅ residue was substituted with Gln as described below. A dsDNA encoding a Kozak sequence and the signal peptide of mouse FGF-6, in which Asn₃₅ was substituted to Gln₃₅, was amplified by PCR using primers No. 105 (5'-gcgtc gaccc accat gtc-3') and No. 124 (5'-gcgat atcca gtacg gtgcc ttggg cgcg-3') and mouse FGF-6 cDNA as a template. The

PCR product was digested with *Sal*I and *Eco*RV and subcloned into the *Aor*51HI and *Sal*I sites of FGF-1 α /pBS. This construct was designated secFGF/pBS. This cDNA construct was used in all the following study as the core construct to which the *N*-glycosylation cassette was introduced. Although the chimeric FGF protein encoded by this cDNA still contains one potential *N*-glycosylation sequence from the original FGF-1 primary structure (Asn₁₄₉GlySer), *N*-glycosylation of this sequence was found to be of very low level (M. Asada et al., unpublished data).

DNase I Digestion

Ten micrograms of secFGF/pBS in 100 μ L of 50 mM Tris-HCl (pH 7.5) with 100 μ g/mL bovine serum albumin (BSA) and 10 mM MnCl were digested with DNase I (1.4 mU) at 37°C. Aliquots of the reaction mixture were sampled at various times and mixed with 1 μ L of 0.5 M EDTA (pH 8.0) to stop the reaction. The digests were resolved by electrophoresis on 1.4% agarose gels containing ethidium bromide in 1 $\times$ 40 mM Tris-acetate with 1 mM EDTA (pH 8.0). Appropriately sized DNA was excised from the gel, purified using SUPREC-01 and dissolved in 10 mM Tris-HCl with 1 mM EDTA (pH 8.0).

Preparation of Probes and Southern Blotting

Probes used for Southern blot analysis were prepared from the secFGF cDNA digest using restriction enzymes. secFGF cDNA was amplified by PCR using secFGF/pBS as a template with primers No. 115 (5'-aaca aagct gggt aacggg-3') and No. 116 (5'-gccgc ttag aacta gtggat-3'). Aliquots of the product were digested with *Nae*I/*Sal*I or *Hin*fI/*Eco*RI: the former yielded a 103-bp fragment that served as the 5'-terminal section of the secFGF cDNA, and the latter yielded a 69-bp fragment that served as the 3'-terminal section. The fragments were purified and labeled with digoxigenin using a DIG DNA labeling kit according to the manufacturer's instructions (Roche Molecular Biochemicals). DNA transfer onto nylon Hybond-N+ membranes, hybridization with probes and detection using anti-

digoxigenin-POD Fab fragment were performed as described previously (9).

Preparation of cDNA Cassettes that Encoded the Potential *N*-Glycosylation Sequence

Three inserts with differing reading frames, each encoding a potential *N*-glycosylation sequence (AsnLeuSer) and a chloramphenicol resistance gene (Cm^r) removable by digestion with *A*f/II, were generated by PCR. To accomplish this, a Cm^r cDNA template was amplified with *P*fu DNA polymerase using the following primers. For the 3n fragment, the primers were forward (5'-gcgcg cgta acta agcat cggca cgtaa gaggt tccaa ctt-3') and reverse (5'-atata agcgc ttaag ttacg ccccg ccctg cact ca-3'); for the 3n+1 fragment, forward (5'-gcgcg cggtt aaact taagc atcgg cacgt aagag gttcc aactt-3') and reverse (5'-atata tatcc cgggc ttaag ttacg cccgc cctgc cact-3'); and for the 3n+2 fragment, forward (5'-atata tcggc aactt aagca tcggc acga agagg ttcca ac-3') and reverse (5'-atata tatgc cggtt taagt tacgc cccgc cctgc cactca-3'). Each PCR product was digested with Klenow enzyme (Takara Shuzo) and ligated into the *Eco*RV site of pBluescript SK+ treated with calf intestine alkaline phosphatase (Takara Shuzo), and each insert fragment was cleaved from its respective construct by digestion with the following sets of enzymes. The 3n fragment was obtained by digestion with *Hinc*II and *Aor*51HI; 3n+1 by digestion with *Pme*I and *Sma*I; and 3n+2 by digestion with *Nru*I and *Nae*I. Each fragment was then purified using GENECLEAN II and resuspended in TE buffer.

Ligation of a cDNA Cassette Encoding an *N*-Glycosylation Sequence into a DNase I-Treated Plasmid

The inserts were then individually ligated with DNase-treated plasmids. *E. coli* DH5 α were transformed by each product and selected for chloramphenicol (Cm) and ampicillin (Amp) resistance. Plasmids were prepared by the alkaline-sodium dodecyl sulfate (SDS) method and digested with *A*f/II to remove Cm^r. Plasmids free of the Cm^r gene were purified and recircularized by ligation under dilute conditions. The recircularized plasmids were digested

Research Report

with *Sall/EcoRI* to isolate the inserts containing the cDNA encoding the secFGF proteins with *N*-glycosylation sequences. The inserts were then further purified using SUPREC-01 enabling subsequent subcloning into an expression vector.

Construction of secFGF cDNA Harboring an *N*-Glycosylation Sequence at Predetermined Sites

For purposes of comparison, two additional forms of secFGF with an *N*-glycosylation cassette were constructed. The "A" form harbored a potential *N*-glycosylation sequence at the N terminus of the FGF-1 α portion. Its cDNA was constructed by the overlapping extension method using PCR and cloned *Pfu* DNA polymerase. The 5' half fragment was generated by PCR using primers No. 115 and No. Ntr (5'-attag cactt aagtt atcca gtacg gtgcc ttg-3') and secFGF/pBS as a template, while the 3' half was generated using the same template and primers No. 116 and No. Ntf (5'-ctgga taact taagt gctaa ttaca agaag ccc-3'). Equal aliquots of the reactants were mixed, and PCR was performed using primers No. 115 and No. 116. The final product was digested with *Sall/EcoRI*, purified and subcloned between the *Sall/EcoRI* sites of pBluescript KS+ vector. The "F" form harbored a potential *N*-glycosylation sequence at the C terminus of secFGF. The cDNA for F-form secFGF was generated by PCR using cloned *Pfu* DNA polymerase, secFGF/pBS as a template and primers No. 115 and C term reverse (5'-gcgcg cgaat tccta gctta agtta tcaga agaga ctggc ag-3'). The reaction product was digested with *Sall/EcoRI*, purified and subcloned between the *Sall/EcoRI* sites of pBluescript KS+ vector.

Sequence Analysis and Construction of Expression Vectors

Nucleotide sequences were analyzed using the dideoxynucleotide chain termination method with a Model 373S DNA sequencer (PE Biosystems). Selected cDNA clones were isolated by digestion with *Sall/EcoRI* and subcloned between the *Sall/EcoRI* sites of pMEX-neo, an eukaryotic expression vector.

Figure 1. Introduction of an *N*-glycosylation cassette into the polypeptide core of FGF at random sites. (A) Schematic diagram of the procedure for random insertion of an AsnLeuSer coding sequence. The secFGF is a chimera composed of the mutated signal peptide from mouse FGF-6 and the N-terminal truncated form of human FGF-1. The cDNA was cloned in pBlueScript (pBS). Digestion with DNase I in the presence of Mn²⁺ generates secFGF cDNAs containing a single blunt end cut at various sites. *N*-glycosylation cassettes in three different frames containing Cm^r were ligated into the blunt end cuts. *E. coli* were transformed by the ligation products, and colonies were selected on Cm/Amp LB plates. Plasmids prepared from selected colonies were digested with *AflIII* to remove Cm^r, recircularized and cloned.

Research Report

Transfection

Ba/F3 cells [1.5×10^6 in 150 μ L of phosphate-buffered saline (PBS)] were transfected with 8 μ g of expression plasmid by electroporation and suspended in RPMI 1640 medium supplemented with 10% FBS and 10% WEHI-3-conditioned medium. After two days in culture, the transfected cells were provided with fresh growth medium containing geneticin disulfate salt (1 mg/mL) and cultured for an additional two weeks to select stable transfectants.

Analysis of the Expressed FGF Proteins

To detect FGF proteins secreted into the culture medium, 40 μ L of heparin sepharose CL6B gel equilibrated with PBS were added to 2 mL of the conditioned medium from each transfectant culture and incubated at 4°C overnight. For the transfectants B, C, D and E, 9

mL of the conditioned medium was used. After washing with PBS twice, the beads were then boiled in 40 μ L of SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer for 10 min, and the dissociated proteins were analyzed by SDS-PAGE.

To detect intracellular FGF proteins, the cells (ca. 5×10^5) were lysed on ice for 5 min in 100 μ L of lysis buffer, consisting of 10 mM Tris-buffered saline (TBS) (pH 7.5) containing 1% Nonidet® P-40 (Iwai Kagaku, Ltd., Tokyo, Japan). The lysates were centrifuged, and the supernatants were recovered, mixed with SDS-PAGE sample buffer and boiled for 10 min. The samples were then centrifuged again, and the same quantity of FGF proteins in cleared supernatants were resolved by SDS-PAGE on 13% polyacrylamide gels attached with 3% stacking gels (6). After electrophoresis, the resolved proteins were electroblotted onto nitrocellulose membranes (Optitran® BA-S 85

reinforced NC, 0.45 μ m; Schleicher & Schuell, Dassel, Germany). The membranes were blocked with 3% skim milk (dehydrated; Becton Dickinson Microbiology Systems, Sparks, MD, USA) in TBS for 1 h at room temperature and then incubated with 1 μ g/mL anti-FGF-1 monoclonal antibody (mAb-1) (3) in TBS containing 3% skim milk overnight at 4°C. The membranes were then washed three times with 0.05% Tween® 20 in TBS and incubated at room temperature for 1 h with affinity-purified goat anti-mouse IgG conjugated with horseradish peroxidase (Chemicon International, Temecula, CA, USA). After washing, the signals were detected using an ECL™ Detection Kit (Amersham Pharmacia Biotech).

PNGase Digestion

The expressed FGF proteins contained in the lysates of the transfectants were adjusted to equal amounts and

Figure 1. (B) *N*-glycosylation cassettes in three different frames. The cDNA cassettes for introducing the potential *N*-glycosylation sequence (AsnLeuSer) were designed in three different frames. To minimize addition of excessive amino acids, the AsnLeuSer-coding sequence (aacttaagc) was placed on both ends of the Cm^r gene so that when digested with Afl/II (which recognizes cctaag), the resultant reconstituted sequence would encode AsnLeuSer after ligation. The figure illustrates examples for various results of DNase I digestion; note that besides the AsnLeuSer sequence, introduction of the cassettes adds only a single amino acid at most following removal of Cm^r.

Research Report

digested with 0.2 U of peptide *N*-glycosidase F (PNGase) overnight in a final volume of 50 μ L. The reaction products were resolved by SDS-PAGE and visualized by immunoblot as described above.

RESULTS

DNase I Digestion of secFGF cDNA Generating Single Cuts at Random Sites

Human FGF-1 was selected as a model protein into which a potential *N*-glycosylation sequence would be introduced. To enable the expressed proteins to be efficiently secreted, a signal peptide was introduced at the N terminus along with a single alanine used as a linker. An oligonucleotide encoding mouse FGF-6 signal peptide (the N terminal 40 amino acids) was fused to the 5' terminus of cDNA of FGF-1 α , an N-terminal truncated form of FGF-1 (22–155) with full mitogenic activity, as shown in Figure 1A.

The critical step in the procedure is to randomly introduce a potential *N*-glycosylation sequence into the polypeptide core of secFGF (Figure 1). Therefore, we first established a protocol for creating single blunt cuts at random sites in the secFGF cDNA using DNase I, which in the presence of Mn^{2+} cleaves the two DNA strands at approximately the same site yielding fragments that are blunt-ended or that have protruding termini of only one or two nucleotides (8). Controlled digestion at 37°C for selected times generated a single-cut form (Figure 2A, lanes 2–7, arrow b) and various quantities of relaxed form (Figure 2A, lanes 2–7, arrow a). For example, digestion for 1 min left a small amount of the uncut form remaining (Figure 2A, lane 2, smear band indicated by the white bracket). On the other hand, digestion for 6 min resulted in a large number of multiple cut forms (Figure 2A, lane 7, indicated by the

black bracket). The optimum protocol for DNase I digestion of secFGF cDNA was thus determined to be the following. Digest 10 μ g of DNA with 1.4 mU of DNase I at 37°C for 3 min in the presence of 10 mM Mn^{2+} .

To test whether the DNase I digestion had cut the plasmid randomly, the DNA bands corresponding to the single-cut form (Figure 2A, lanes 3–6, arrow b) were excised from the gel, and the DNA was purified, resolved on an agarose gel, transferred to a nylon membrane and hybridized with a digoxigenin-labeled probe targeting the 3' terminus of secFGF cDNA (Figure 2B, lane 1). No smear bands appeared below the size of the linear plasmid, which confirmed the integrity of the single-cut form (Figure 2B, lane 1, arrow). The same results were obtained using the 5' terminus as the probe (data not shown). When the purified, single-cut DNA was further digested with

Figure 2. Generation of a single blunt cut at random sites by digestion with DNase I. (A) Time course of digestion with DNase I. The secFGF/pBS plasmid was digested with DNase I at 37°C in the presence of Mn^{2+} . Aliquots of the digests were sampled at selected times and resolved by agarose electrophoresis. Lane 1, lambda *Hind*III digest size marker; lanes 2, 3, 4, 5, 6 and 7, reaction product after digestion for 1, 2, 3, 4, 5 and 6 min, respectively. Arrows "a" and "b" denote the relaxed and single-cut forms of the plasmid, respectively. The white and black brackets indicate smear bands containing various forms of uncut and multiple-cut plasmid, respectively. (B) Confirmation of the random distribution of the sites digested with DNase I. Plasmids with single-cuts, prepared as in the panel A, were resolved by agarose electrophoresis and purified (lane 1). The same aliquots were then digested with *Eco*RI (lane 2) or with *Sa*I and *Eco*RI (lane 3). The digests were resolved by agarose electrophoresis, transferred to a nylon membrane and hybridized with the 3'-terminal fragment of secFGF cDNA labeled with digoxigenin. The signal was detected with anti-digoxigenin antibody, labeled with peroxidase and visualized by ECL. The arrow indicates the position of the single-cut form, and the arrowhead indicates the position of secFGF cDNA. The long and short brackets indicate smear bands containing cDNA cut at random sites.

*Eco*RI, which was unique in the plasmid, it was fractionated into smear bands (Figure 2B, lane 2, long bracket), indicating that the sites of DNase I digestion were distributed randomly over the plasmid. To further confirm that the secFGF cDNA was cut randomly, the purified single-cut DNA was digested with *Eco*RI and *Sal*I, an enzyme combination that digests both ends of the secFGF cDNA. This reaction generated a population of variably-sized DNA fragments (Figure 2B, lane 3, short bracket), all of which were smaller than secFGF (Figure 2B, lane 3, black ar-

rowhead) and indicative of the random distribution of DNase I cleavage sites.

Preparation of cDNA Cassettes Encoding a Potential *N*-Glycosylation Sequence

To introduce the potential *N*-glycosylation sequence (AsnLeuSer) into the blunt-end cuts created by DNase I digestion, we designed cDNA cassettes in three possible frames that encode AsnLeuSer when matched (Figure 1B). The cassettes also contained the *Cm^r* gene used for selection. To block excessive

Figure 3. Identification of cassette insertion sites in the cDNA. Cassette insertion sites, as revealed by direct nucleotide sequencing of multiple clones, are indicated by arrows and arrowheads. Each clone contained a single cassette; arrows indicate insertion sites where AsnLeuSer was encoded in frame. The clones with gray-colored arrows were constructed using the overlapping extension method.

Research Report

addition of amino acids, the AsnLeuSer-coding sequence was divided in two by the *Cm^r* gene so that after removal of this selection gene, the entire cassette encoded only one amino acid at most, in addition to AsnLeuSer. With three sets of oligonucleotide primers that differed in frame, three different cassettes were generated by PCR using the *Cm^r* gene as a template (Figure 1B). The PCR-generated cassettes were first cloned into the *EcoRV* site of pBluescript plasmids, and a plasmid was selected for each frame only after sequence verification. The respective *N*-glycosylation cassettes in the three-frame format were then cleaved from each plasmid using sets of restriction enzymes as described in Materials and Methods.

Ligation of *N*-Glycosylation Cassettes and the Selection of Clones Encoding secFGF with an *N*-Glycosylation Sequence

The *N*-glycosylation cassettes were ligated with the single-cut plasmids originally encoding secFGF. The ligation products were then used to transform *Escherichia coli* cells, and approximately 400 independent transformant colonies were obtained by selection on plates containing both ampicillin and chloramphenicol. Plasmids were prepared from the individual colonies, and the sizes of the inserts were examined by the digestion with *XhoI/XbaI*.

Fifty clones were found to contain secFGF mutants of approximately 1250 bp, suggesting that the *N*-glycosylation cassettes were successfully inserted into the open reading frame of secFGF. In contrast, some clones were found to contain inserts that were equal in size to secFGF or at least smaller than 1250 bp. These smaller inserts likely derived from double-cut secFGF cDNA. The mutant secFGF cDNA clones were digested with *Afl*III to remove *Cm^r* and were recircularized by self-ligation under dilute conditions. Unsuccessful clones were not pursued further. Nucleotide sequences of the recircularized clones were determined by restriction enzyme digestion and direct sequencing, which identified 21 mutated secFGF clones that had successfully incorporated AsnLeuSer-coding nucleotides without losing the original secFGF nucleo-

tide sequence. Indeed, 4 of the 21 clones encoded secFGF, in which the potential *N*-glycosylation sequence (AsnLeuSer) was encoded in frame at random sites (Figure 3, arrows B, C, D and E). Table 1 summarizes the local amino acid sequences of the four clones deduced from their nucleotide sequences. For clones A and F, the *N*-glycosylation consensus sequence was respectively inserted at the N or C terminus in the same format as the other clones.

Expression of the *N*-Glycosylation Sequence-Introduced secFGF Proteins in Mammalian Cells

The inserts encoding secFGF with the attached potential *N*-glycosylation sequences (AsnLeuSer) were individually subcloned from the cloning vector to a pMEXneo expression vector (4) and then transfected into Ba/F3 cells. The culture medium and cell bodies of the transfectants were collected and analyzed for expression of secFGF proteins. Analysis of the cell lysates revealed that all of the cDNA clones were successfully translated into proteins exhibiting secFGF antigenicity (Figure

Table 1. Local Sequence of Six secFGF Clones into Which an *N*-Glycosylation Cassette Was Introduced

Clone	Intact secFGF Sequence	Created Sequence
A	LLDAN	LLD NLS AN
B	RDRSD	RDT NLS RSD
C	DQHIQ	DQ QNL SPIQ
D	GEVYI	GE ANL SLYI
E	LEENH	LE ANL SQNH
F	VSSD*	VSSD NLS *

The inserted amino acids are indicated by the bold font, and the *N*-Glycosylation sequence are underlined. D* and S* denote Asp and Ser at the C-termini, respectively.

4A). Further, most of these proteins were secreted by the transfectants; one exception, the translate of clone B was not detected in the conditioned culture medium (Figure 4B).

Analysis of the secreted proteins on SDS-PAGE followed by immunoblotting revealed that portions of the C, D and E forms of the secFGF mutants and

Figure 4. Analysis of the potentially *N*-glycosylated proteins expressed by Ba/F3 cells. (A) Cell lysates. The cell lysate of each transfectant was resolved by SDS-PAGE, and the FGF proteins were detected using anti-FGF-1 monoclonal antibody. The closed arrowheads indicate the position of *N*-glycosylated proteins of FGF proteins, and the open arrowhead indicates the position for the simple secFGF protein. (B) Conditioned medium. The secFGF forms containing the *N*-glycosylation sequence were expressed in Ba/F3 cells. The same quantity of the secFGF proteins in the conditioned medium of each transfectant were resolved by SDS-PAGE, and the FGF proteins were detected using anti-FGF-1 monoclonal antibody. (C) *N*-glycosylation. The lysates of the transfectants were either not treated (-) or digested with PNGase (+). The reaction products were resolved by SDS-PAGE and detected with anti-FGF-1 antibody. The same quantity of total FGF protein was analyzed for each construct.

Research Report

secFGF(N₃₅) all migrated as though they were of greater molecular mass than the simple secFGF protein, suggesting the presence of modification groups on these proteins. In contrast, modification of the A and F forms were not apparent (Figure 4B). Interestingly, the B-form mutant was detected in the lysate as a heavily modified protein (Figure 4A), even though it was not detected in the conditioned medium (Figure 4B).

To investigate the nature of the modifications found on the various secFGF mutants, they were digested with PNGase and subjected to SDS-PAGE (Figure 4C). After digestion, all of the mutants, including the B form, migrated on SDS-PAGE in a manner indicating the same molecular mass as the original secFGF. This confirmed that the modifications found on all of the mutants were *N*-glycans. Because the potential endogenous *N*-glycosylation

sequence at Asn₁₄₉ was modified at very low level in our earlier studies (M. Asada et al., unpublished data), most likely *N*-modifications of the C, D and E forms took place only on the newly introduced *N*-glycosylation cassettes. With respect to the B-form mutants, the excessively high glycosylation levels and the appearance of higher molecular mass bands (Figure 4A) suggests that the RDTNLSRSD sequence created in this construct (Table 1) is a very efficient acceptor for *N*-glycosylation.

DISCUSSION

Since the discovery of the *N*-glycosylation consensus sequence (7), its addition to or deletion from a large number of proteins has been attempted only by targeted mutagenesis. However, detailed structural information has been required to successfully carry out the

insertion. Random introduction of an *N*-glycosylation sequence into a protein may thus be a powerful new approach to remodeling proteins for which structural data are not available or structural changes are not predictable. As there is no practical limitation on the degree to which the scale of our protocol can be expanded, it could, in principal, provide a cDNA library for all neoglycosylated proteins. Because the method does not depend on the presence of specific restriction enzyme sites, multiple clones expressing proteins with novel glycosylation sites can be generated by a single approach. This procedure is most useful for characterizing secreted proteins or transmembrane proteins with only one transmembrane domain. These proteins impose little constraint on the position at which the glycosylation site might be inserted. However, transmembrane proteins with several transmembrane domains allow the

CONFERENCE ANNOUNCEMENT: OCTOBER 24–26 1999 **Molecular Library Technologies At The Millennium**

DATE: Oct. 24–26, 1999 **PLACE:** Holiday Inn, Chinatown, San Francisco, CA, USA
Tel: 415-433-6600; 800-243-1135
Organizers: Dennis Burton, PhD, James W. Larrick, MD, PhD, and James D. Marks, MD, PhD
Sponsor: Palo Alto Institute of Molecular Medicine
Financial Support: BioTechniques/Eaton Publishing, Prolifaron, Active Motif,
Genentech, Phylos, Dyax/Target Quest
Details: Palo Alto Institute of Molecular Medicine, 2462 Wyandotte Street, Mountain View, CA 94043 USA
Website: www.pano.com/paimm
Tel: 650-694-1420 **Fax:** 650-694-7717 **E-mail:** paimm@netgate.net

❖ Keynote Speakers ❖

Sidney Brenner, *Molecular Sciences Institute, Berkeley*; Stuart Kauffman, *BiosGroup*
James Wells, *Sunesis*; Greg Winter, *MRC*

❖ Other Speakers ❖

Andreas Plückthun, *Zurich*; John McCafferty, *CAT*; James D. Marks, *UCSF*;
Andrew Bradbury, *Los Alamos*; Ian Tomlinson, *MRC*; William Dower, *Affymax*,
Carlos Barbas, *Scripps*; Tim Clackson, *Ariad*; Hennie Hoogenboom, *Dyax/Target Quest*
Robert Charles Ladner, *Dyax*; Arne Skerra, *Munich*; Kim D. Janda, *Scripps*; Jon A. Ellman, *Berkeley*
Nebosja Janjic, *NexStar*; K. Dane Wittrup, *Illinois*; Michael Taussig, *Babraham*
Garry Nolan, *Stanford*; Roger Brent, *Molecular Sciences Institute*
Stephen Michnick, *Montreal*; Andrew Griffiths, *MRC*; Pim Stemmer, *Maxygen*
Richard Roberts, *CalTech*; Frances Arnold, *CalTech*; Richard Houghton, *Torrey Pines*

This meeting will honor Cara Marks, PhD, structural biologist, Assistant Professor at UCSF and wife of Jim Marks.
Cara died unexpectedly in May of this year. Proceeds from the meeting will endow the Cara Marks Memorial
Scholarship Fund at the Crowden School to support musical and academic development of children ages 9 to 14.

Research Report

insertion of a glycosylation site only outside the hydrophobic membrane-embedded sequences. PCR-based methods, which offer the possibility to insert additional amino acids at specific sites, will work better in these cases.

Recently, a method for a random insertion of a fluorescent protein into the regulatory subunit of cAMP-dependent protein kinase was developed to obtain information about the enzyme's functional domains (2). However, it is not applicable for introducing a short peptide like AsnLeuSer, nor does it provide a method to adjust translation frames. Methods for random insertion of a short peptide has not been reported. Thus, the system described here offers many advantages over the earlier methods (10).

In our experiments, the *N*-glycosylated proteins comprised only a fraction of the expression products for most of the constructs. Kasturi et al. (5) reported that the efficiency of the *N*-glycosylation of a rabies virus glycoprotein

increased when AsnXaaSer was substituted with AsnXaaThr. Therefore, using an *N*-glycosylation cassette with higher efficiency than our present AsnLeuSer cassette may improve the yield of glycosylated forms. However, efficiency and specificity of glycosylation is also dependent on many cellular factors including the expression levels of glycosyl transferases and transporters, and the activity of the synthetic pathway for the sugar precursors. This makes establishing a glycoprotein production system capable of complete addition of a uniform oligosaccharide moiety a subject for continued future study.

Finally, because no FGF-specific or *N*-glycosylation sequence-specific structural characteristics were necessary by the protocol, with minor modification, the method described here should also be applicable to any other combination of protein and oligopeptide/polypeptide cassette of interest.

ACKNOWLEDGMENTS

This research was supported in part by the Complexed Carbohydrate project of the Agency of Industrial Science and Technology of Ministry of International Trade and Industry to T.I., M.S. and M.A. A.Y. is supported by a NEDO fellowship. Text was edited by MST editing company.

REFERENCES

1. **Baldwin, A.N. and E.M. Shooter.** 1995. Zone mapping of the binding domain of the rat low affinity nerve growth factor receptor by the introduction of novel *N*-glycosylation sites. *J. Biol. Chem.* 270:4594-4602.
2. **Biondi, R.M., P.J. Baehler, C.D. Reymond and M. Veron.** 1998. Random insertion of GFP into the cAMP-dependent protein kinase regulatory subunit from *Dictyostelium discoideum*. *Nucleic Acids Res.* 26:4946-4952.
3. **Imamura, T., S. Oka, T. Tanahashi and Y. Okita.** 1994. Cell cycle-dependent nuclear localization of exogenously added fibroblast growth factor-1 in BALB/c 3T3 and human vascular endothelial cells. *Exp. Cell Res.* 215:363-372.
4. **Imamura, T., Y. Tokita and Y. Mitsui.** 1992. Identification of a heparin-binding growth factor-1 nuclear translocation sequence by deletion mutation analysis. *J. Biol. Chem.* 267:5676-5679.
5. **Kasturi, L., J.R. Eshleman, W.H. Wunner and S.H. Shakin-Eshleman.** 1995. The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence *N*-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. *J. Biol. Chem.* 270:14756-14761.
6. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
7. **Marshall, R.D.** 1974. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochem. Soc. Symp.* 40:17-26.
8. **Melgar, E. and D.A. Goldthwait.** 1968. Deoxyribonucleic acid nucleases. II. The effects of metals on the mechanism of action of deoxyribonuclease I. *J. Biol. Chem.* 243:4409-4416.
9. **Ozawa, K., M. Seo and T. Imamura.** 1997. A quantitative method for evaluation of FGF family and FGF receptor family gene expression by RT-PCR. *Brain Res. Brain Res. Protoc.* 1:211-216.
10. **Ross-Macdonald, P., A. Sheehan, G.S. Roeder and M. Snyder.** 1997. A multipurpose transposon system for analyzing protein production, localization, and function in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 94:190-195.
11. **Ueda, T., H. Iwashita, Y. Hashimoto and T. Imoto.** 1996. Stabilization of lysozyme by introducing *N*-glycosylation signal sequence. *J. Biochem.* 119:157-161.
12. **Vannier, B., X. Zhu, D. Brown and L. Birnbaumer.** 1998. The membrane topology of human transient receptor potential 3 as inferred from glycosylation-scanning mutagenesis and epitope immunocytochemistry. *J. Biol. Chem.* 273:8675-8679.
13. **Varki, A.** 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3:97-130.
14. **Zhan, J., M.d. Sousa, J.A. Chaddock, L.M. Roberts and J.M. Lord.** 1997. Restoration of lectin activity to a non-glycosylated ricin B chain mutant by the introduction of a novel *N*-glycosylation site. *FEBS Lett.* 407:271-274.

Received 10 February 1999; accepted 3 May 1999.

Address correspondence to:

Dr. Toru Imamura
Biosignaling Department
National Institute of Bioscience and Human
Technology
Higashi 1-1, Tsukuba
Ibaraki 305-8566, Japan
Internet: imamura@nibh.go.jp