Phosphoglycosylation catalyzed by UDP-GlcNAC:Ser-protein N-acetylglucosamine-1-phosphotransferase (Ser:GlcNAC phosphotransferase) adds GlcNAco-1-P to peptideyl-Ser of selected Dictyostelium discoideum proteins. Lysosomal cysteine proteinase (CP), proteinase-1(CP7), is the major phosphoglycosylated protein in bacterially grown amoebae. GlcNAc-1-P is added within a Ser-rich domain containing SSS, SGSG, or SGSG repeated motifs that are not found in other papain-like CPs. We studied the substrate specificity of the transferase using peptides containing these motifs and 12 other peptides with one or more Ser residues. Phosphoglycosylation is comparable for all three Dictyostelium CP motifs, but it is not restricted to them. Flanking residues in the other peptides strongly influence phosphoglycosylation efficiency. Dictyostelium microsomal membranes also phosphoglycosylate endogenous acceptors, and some of these acceptors occur as an 18 S complex with the transferase. CP-serine motif peptides inhibit endogenous acceptor phosphoglycosylation weakly (30–40%) at 800 μM, whereas catalytically inactive proteinase-1(CP7) and other non-phosphoglycosylated eukaryotic CPs, lacking the serine domain, inhibit transferase activity at 1–4 μM. SDS denaturation destroys the inhibitory potential of all CPs showing that transferase recognizes a conformation-dependent feature that is shared by all. Proteinase-1(CP7) expressed in Escherichia coli lacks GlcNAc-1-P, but it is a substrate for Ser:GlcNAC phosphotransferase, K_m = 5.6 μM. Thus, Ser:GlcNAC phosphotransferase recognizes both acceptor peptide sequences and a conformational feature of eukaryotic CPs. This may be physiologically important for establishing or maintaining non-overlapping groups of GlcNAc-1-P- and Man-6-P-modified Dictyostelium proteins that reside in functionally distinct endo-lysosomal vesicles.

The lysosomal enzymes of Dictyostelium have two types of carbohydrate modifications. One is a methylated form of Man-6-P on N-linked oligosaccharides (Man-6-POCH_2) (1) and the other, called phosphoglycosylation, is GlcNAc-1-P in a phosphodiester linkage to serine (2). The two modifications appear to be on distinct classes of proteins, indicating that addition of each type is highly selective (3, 4).

A lysosomal cysteine proteinase (CP) from Dictyostelium, proteinase-1(CP7), is the major carrier of GlcNAc-1-P in bacterially grown vegetative cells (2, 3). Dictyostelium CPs have highly conserved domains that are found in other eukaryotic CPs of the papain family, but those expressed during vegetative growth also have a novel serine-rich domain containing repeated motifs of SGSG, SSSS, and SGSG that are not found in other members of the CP family (5, 6). One or more of these motifs is probably the site GlcNAc-1-P addition since deletion of this entire Ser-rich domain in proteinase-1(CP7) abolishes the recognition by a GlcNAc-1-P-specific monoclonal antibody (6).

Protein glycosylation is initiated when a glycosyltransferase recognizes an acceptor amino acid in the context of neighboring residues, e.g., the NXS(T) sequence for N-linked glycosylation (7). O-Linked glycosylation initiated by αGalNAc transferases involves recognition of surrounding amino acids and aspects of their secondary structure that generate a continuum of efficient to inefficient acceptor glycosylation sites (8, 9). Glycosyltransferases can also be specific for certain classes of proteins. Examples include the selective addition of Man-6-P residues to high mannose-type N-linked chains of mammalian lysosomal enzymes (10, 11) and the modification of complex-type N-linked chains of pituitary hormones with GalNAc-4-SO_4 (12–14). In both cases, the transferase binds generic oligosaccharides with low affinities but binding improves 100–1000-fold when these chains are presented in the context of their natural protein acceptors (12–14). These glycosyltransferases have a catalytic site and a second site for recognition of conformation-specific features of the acceptor proteins (13, 14).

UDP-GlcNAC:Ser-protein N-acetylglucosamine-1-phosphotransferase (Ser:GlcNAC phosphotransferase) is responsible for phosphoglycosylation of Dictyostelium proteins. Although this enzyme has been partially purified and characterized from Dictyostelium earlier (15), very little is known about its substrate specificity. Since our previous results showed that CPs are modified by GlcNAc-1-P (3, 4), we wanted to determine if Ser:GlcNAC phosphotransferase specifically recognizes features of the CPs and/or one of the Ser motifs found in Dictyostelium. We found that the transferase has a preference for certain Ser-containing peptide sequences, but most importantly, it shows a high affinity for binding eukaryotic CPs of the papain family. The latter may be the first step in selecting proteins for phosphoglycosylation and may help to distinguish them from proteins modified with Man-6-P.

* This work was supported by NIGMS RO1 32485 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: The Burnham Institute, La Jolla Cancer Research Center, 10901 North Torrey Pines Rd., La Jolla, CA 92037. Tel.: 619-646-3142; Fax: 619-646-3193; E-mail: hudson@ljcrf.edu.

A lysosomal cysteine proteinase (CP)_1 from Dictyostelium, proteinase-1(CP7), is the major carrier of GlcNAc-1-P in bacterially grown vegetative cells (2, 3). Dictyostelium CPs have highly conserved domains that are found in other eukaryotic CPs of the papain family, but those expressed during vegetative growth also have a novel serine-rich domain containing repeated motifs of SGSG, SSSS, and SGSG that are not found in other members of the CP family (5, 6). One or more of these motifs is probably the site GlcNAc-1-P addition since deletion of this entire Ser-rich domain in proteinase-1(CP7) abolishes the recognition by a GlcNAc-1-P-specific monoclonal antibody (6).

Protein glycosylation is initiated when a glycosyltransferase recognizes an acceptor amino acid in the context of neighboring residues, e.g., the NXS(T) sequence for N-linked glycosylation (7). O-Linked glycosylation initiated by αGalNAc transferases involves recognition of surrounding amino acids and aspects of their secondary structure that generate a continuum of efficient to inefficient acceptor glycosylation sites (8, 9). Glycosyltransferases can also be specific for certain classes of proteins. Examples include the selective addition of Man-6-P residues to high mannose-type N-linked chains of mammalian lysosomal enzymes (10, 11) and the modification of complex-type N-linked chains of pituitary hormones with GalNAc-4-SO_4 (12–14). In both cases, the transferase binds generic oligosaccharides with low affinities but binding improves 100–1000-fold when these chains are presented in the context of their natural protein acceptors (12–14). These glycosyltransferases have a catalytic site and a second site for recognition of conformation-specific features of the acceptor proteins (13, 14).

UDP-GlcNAC:Ser-protein N-acetylglucosamine-1-phosphotransferase (Ser:GlcNAC phosphotransferase) is responsible for phosphoglycosylation of Dictyostelium proteins. Although this enzyme has been partially purified and characterized from Dictyostelium earlier (15), very little is known about its substrate specificity. Since our previous results showed that CPs are modified by GlcNAc-1-P (3, 4), we wanted to determine if Ser:GlcNAC phosphotransferase specifically recognizes features of the CPs and/or one of the Ser motifs found in Dictyostelium. We found that the transferase has a preference for certain Ser-containing peptide sequences, but most importantly, it shows a high affinity for binding eukaryotic CPs of the papain family. The latter may be the first step in selecting proteins for phosphoglycosylation and may help to distinguish them from proteins modified with Man-6-P.

The abbreviations used are: CP, cysteine proteinase; E-64, trans-epoxysuccinyl-l-leucylamido(-4-guanidino) butane; MALDI-MS, matrix-assisted laser desorption ionization-mass spectrometry; DAN, 1,5-diaminonaphthalene; HABA, [2-(4-hydroxyphenylazo)-benzoic acid]; C8, N-octanoyl; HPLC, high pressure liquid chromatography; GST, glutathione S-transferase; Ser:GlcNAC phosphotransferase, UDP-GlcNAC: Ser-protein N-acetylglucosamine-1-phosphotransferase.

This paper is available on line at http://www.jbc.org
EXPERIMENTAL PROCEDURES

MATERIALS—UDP-[6-3H]GlcNAc 60 Ci/ml was purchased from American Radiolabeled Chemicals, St. Louis, MO. Peptide 1 (N-octanoyl (C8)-GS-GSSGSGSGS), peptide 2 (N-octanoyl (C8)-GS-GSSGSGS), and peptide 3 (N-octanoyl (C8)-GS-GSSGSGS) were from Tama Laboratories, Houston, TX. Other peptides used were gifts from Dr. Robert Hill (Duke University Medical Center, Durham, NC) and Dr. Ake Elhammar (Upjohn Co., Kalamazoo, MI). Bonded C18 reversed phase silica gel 15 μm Plasmid pGEX 5X1 was purchased from Pharmacia Biotech Inc. and FLAG-monomonal antibody M2 was from Kodak. All other chemicals were from Sigma.

Cells—Dictyostelium discoideum strain AX-2 grown in synthetic HL-5 medium and harvested in mid log phase was used for all work (16).

Ser-GlcNAc Phosphotransferase Assay—Membranes were prepared from axenically growing cells at 4 × 109 cells/ml by sonication and high speed centrifugation as detailed earlier (17). Assay of Ser-GlcNAc phosphotransferase in solubilized membranes with synthetic peptides was performed essentially as described earlier (18), except that after stopping the reaction, the products were analyzed on C18 columns using the vacuum manifold system from Qiagen. The peptides were used at 4–5-fold molar excess with respect to the respective K$_m$ within the linear range for protein concentration and time. Assay of transferase with endogenous acceptors was performed as detailed (15). For competition with purified CPs in the endogenous assay, the proteinases were preincubated with either cystatin from egg white or E-64 for 1 h at room temperature and then added to the assay mixture as a complex with the inhibitor. CPs were denatured by adding a 2-fold excess of SDS over protein and heating at 100 °C for 5 min. Increasing concentrations of purified CP7 overexpressed in Dictyostelium (21) and Ser-GlcNAc phosphotransferase activity (fractions 7–12 from sucrose density gradient which had very low activity on endogenous acceptors) to determine the K$_m$ of the transferase for the fusion protein. The fusion protein was denatured by 2-fold excess of SDS and heated at 100 °C for 5 min, and a 4–5-fold excess of bovine serum albumin was added to quench the excess SDS. This mixture was then incubated in the presence of the transferase to determine the K$_m$ of denatured fusion protein. A similar incubation of the transferase with SDS-denatured bovine serum albumin showed that the transferase activity is still retained under these conditions.

Construction of the Plasmid for Mature-FLAG-CP7-GST—The plasmid pGEX 5X1 was used for generating the fusion protein, FLAG-CP7-GST. The complete sequence of the CP7 gene, cprG, in Dictyostelium has been published and the FLAG epitope was introduced before the cloning site. The polymerase chain reaction product, mature-FLAG-CP7, was cut with restriction endonucleases and ligated into the vector pGEX 5X1 which was prepared using the same enzymes. The construct was transformed into E. coli strain XL1-Blue using the rapid transformation procedure according to the manufacturer’s directions.

Expression and Purification of the Mature FLAG-CP7 Fusion Protein—A single colony of E. coli XL1-blue containing the plasmid encoding mature-FLAG-CP7-GST was incubated in LB medium. The cells were grown at 37°C until the A$_{600}$ reached 1.0 and then isopropyl-1-thio-β-D-galactopyranoside was added at a final concentration of 1 mM, and the culture was further maintained at 37°C for 3 h. The cells were harvested and subjected to extensive sonication in TTBS buffer. The incubation was added to the cells containing the fusion protein and collected by centrifugation at 10,000 × g for 10 min. The pellet containing the fusion protein was resuspended in 0.1 M Tris, pH 8.0, containing 8 mM urea, 0.1 M NaH$_2$PO$_4$, 1% Triton X-100 and dialyzed extensively against TTBS buffer overnight at 4°C. The dialyzed material was diluted 30-fold and applied to a DEAE column equilibrated with 5 mM NaCl, 20 mM Tris-HCl, pH 7.5. The column was washed extensively with the same buffer, and the fusion protein was eluted with 20 mM NaCl in 20 mM Tris-HCl, pH 7.5. The fractions containing the fusion protein were pooled, concentrated by ultrafiltration, and stored at −20 °C. The purified protein was separated on a 12.5% SDS gel and stained with Coomassie Blue and also analyzed by Western blot using either anti-FLAG monoclonal antibody (M2) or GlcNAc-1-P-specific mAb, AD7.5 (3). This was followed by incubation with anti-mouse alkaline phosphatase conjugate and visualized with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium.

Chromatographic Analyses—TLC was performed on cellulose plates and developed in ethyl acetate:acetic acid:water in 6:3:4 ratio (v/v). Sugar standards were detected by silver nitrate spray (19), and the phosphorylated sugar standards were detected by molybdate-pchloric acid spray (19). QAE-Sepharose and C18 reverse phase chromatographies were carried out as described (20). HPLC analyses of anionic compounds were done using a 4.6 mm × 25 cm Microsorb-MV amino-bonded silica column with a linear gradient of 0.01–0.25 m ammonium formate, pH 6.0, containing 40% acetonitrile for 60 min. The flow rate was 0.5 ml/min.

MALDI-MS—Peptide substrates and Ser-GlcNAc phosphotransferase products were analyzed by MALDI-MS using a Kratos MALDI I system. Peptides were dissolved in 50% acetonitrile at 0.1 mg/ml; transferase peptide products were dissolved in a minimal volume of 50% acetonitrile (about 5–20 μM). Matrices used in these studies were as follows: 1) DAN, a saturated solution of 1.5-diaminonaphthaleine in 50% acetonitrile; 2) o-2-nitrophenylhydrazine, a saturated solution of o-nitrobenzaldehyde in 50% acetonitrile; 3) 2,5-dihydroxy-benzoic acid, 50 mM 2,5-dihydroxybenzoic acid in 50% acetonitrile, and 4) HABA, a saturated solution of [2-(4-hydroxyphenylazo)-benzoic acid] in 50% acetonitrile.

Mild Acid/Base Hydrolysis—Mild acid treatment was done on the peptide and endogenous products with 0.01 N HCl at 100 °C for 15 min. Mild base hydrolysis was performed at room temperature with 0.1 N NaOH for 4 h. The acid/base hydrolysates were neutralized, desalted, and analyzed by C18 reverse phase chromatography or QAE-Sephadex anion exchange chromatography as detailed above.

Sucrose Density Gradient Centrifugation—Membranes were prepared from log phase cells grown to 4 × 109 cells/ml, and the post-nuclear supernatants were solubilized with 1% octyl glucoside and centrifuged at 80,000 rpm for 15 min in a TLA-100.2 rotor. The detergent-solubilized extract (10 mg/ml protein) contained the majority of the Ser-GlcNAc phosphotransferase activity on both peptide and endogenous acceptors. 1.0 ml of solubilized extract was layered on a 10–40% sucrose gradient made in 50 mM Tris/HCl, pH 7.5, with 0.1% octyl glucoside and centrifuged at 32,000 rpm either for 3 or 16 h using an SW 40 rotor. 0.5-ml fractions were collected by piercing the bottom of the tube with a 25-gauge needle. The fractions were assayed for protein (21) and Ser-GlcNAc phosphotransferase activity using both peptide 1 (CSGSSSSSGSS) and endogenous acceptors. Total mouse RNA (Ambion, Inc., Austin, TX) was also separated on a similar sucrose gradient to determine the position of 28 S and 18 S ribosomal RNA used as markers.

RESULTS

Phosphotransferase Activity on Peptides That Resemble the Serine-containing Domains of Dictyostelium CPs—Phosphorylated CPs from vegetative Dictyostelium cells have a 25–110-amino acid serine-rich domain that is not found in other eukaryotic CPs (5, 6). The domain has three repeated motifs of N-$\epsilon$-octanoyl (C8)-GS-GSSGSGSGS, N-$\epsilon$-octanoyl (C8)-GS-GSSGSGS, and N-$\epsilon$-octanoyl (C8)-GS-GSSGSGS. All were expressed in E. coli XL1-blue using the rapid transformation protocol according to the manufacturer’s directions.

Expression and Purification of the Mature FLAG-CP7 Fusion Protein—A single colony of E. coli XL1-blue containing the plasmid encoding mature-FLAG-CP7-GST was incubated in LB medium. The cells were grown at 37°C until the A$_{600}$ reached 1.0 and then isopropyl-1-thio-β-D-galactopyranoside was added at a final concentration of 1 mM, and the culture was further maintained at 37°C for 3 h. The cells were harvested and subjected to extensive sonication in TTBS buffer. The incubation was added to the cells containing the fusion protein and collected by centrifugation at 10,000 × g for 10 min. The pellet containing the fusion protein was resuspended in 0.1 M Tris, pH 8.0, containing 8 mM urea, 0.1 M NaH$_2$PO$_4$, 1% Triton X-100 and dialyzed extensively against TTBS buffer overnight at 4°C. The dialyzed material was diluted 30-fold and applied to a DEAE column equilibrated with 5 mM NaCl, 20 mM Tris-HCl, pH 7.5. The column was washed extensively with the same buffer, and the fusion protein was eluted with 20 mM NaCl in 20 mM Tris-HCl, pH 7.5. The fractions containing the fusion protein were pooled, concentrated by ultrafiltration, and stored at −20°C. The purified protein was separated on a 12.5% SDS gel and stained with Coomassie Blue and also analyzed by Western blot using either anti-FLAG monoclonal antibody (M2) or GlcNAc-1-P-specific mAb, AD7.5 (3). This was followed by incubation with anti-mouse alkaline phosphatase conjugate and visualized with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium.

Chromatographic Analyses—TLC was performed on cellulose plates and developed in ethyl acetate:acetic acid:water in 6:3:4 ratio (v/v). Sugar standards were detected by silver nitrate spray (19), and the phosphorylated sugar standards were detected by molybdate-pchloric acid spray (19). QAE-Sepharose and C18 reverse phase chromatographies were carried out as described (20). HPLC analyses of anionic compounds were done using a 4.6 mm × 25 cm Microsorb-MV amino-bonded silica column with a linear gradient of 0.01–0.25 m ammonium formate, pH 6.0, containing 40% acetonitrile for 60 min. The flow rate was 0.5 ml/min.
GlcNAc-1-P, they were treated with either mild acid or base to release label which was analyzed by QAE-Sephadex ion exchange or C18 reversed phase chromatography. Mild acid hy-

**TABLE I**

| Peptides   | $K_m$ (μM) | $V_{max}$ (pmol/min/mg) | $V_{max}/K_m$ |
|------------|------------|------------------------|---------------|
| 1. C8GSSSSSSSG | 194        | 56                     | 0.29          |
| 2. C8GSGSGSGS | 126        | 27                     | 0.21          |
| 3. C8SGSQGSGQ | 129        | 30                     | 0.23          |


drolysis sufficient to cleave labile phosphodiester bonds released 88–94% of the radiolabel which co-migrated exclusively with authentic GlcNAc on cellulose TLC plates (not shown). Base hydrolysis released 86–97% of $^3$H label, and all of it co-migrated with GlcNAc-1-P by TLC (not shown). The base-released material bound to QAE-Sephadex, and 85% was neutralized by digestion with alkaline phosphatase. These results are consistent with the presence of [$^3$H]GlcNAc-1-P in a phosphodiester linkage to the peptidylserine and is similar to the products characterized in previous studies (2, 15, 18).

**HPLC and MALDI-MS Analysis of Peptide Products**—

[$^3$H]GlcNAc-labeled peptides were analyzed by anion exchange HPLC. Peptide 1 gave two major incompletely resolved peaks, and peptides 2 and 3 gave single peaks (Fig. 1A). These products were purified by preparative anion exchange HPLC and

**FIG. 1.** HPLC and MALDI-MS analysis of peptide products. A, the products of peptides 1–3 were purified by C18 cartridges and subjected to anion exchange HPLC on amino columns. The $K_m$ and $V_{max}$ were determined by double-reciprocal plots of reaction rates at different peptide concentrations.
The activity of Ser:GlcNAc phosphotransferase was measured with peptides having single or multiple Ser residues flanked by various amino acids. The peptides were used at 800 μM, and the assay was performed at room temperature and at pH 7.5. \( K_m \) and \( V_{max} \) were determined as described in Table I.

| Peptide     | Specific activity \( \text{pmol/min/mg} \) | \( K_m \) (μM) | \( V_{max} \) (μmol/min/mg) | \( V_{max}/K_m \) |
|-------------|---------------------------------|-------------|------------------|-----------------|
| VLGSIV     | 12                              |             |                  |                 |
| VLGSIV     | 21                              |             |                  |                 |
| VKSEASSFI  | 5                               |             |                  |                 |
| PPDAASAAPLR| 10                             | 75          | 161              | 0.215           |
| PPDDGSGGPLR| 27                             |             |                  |                 |
| PPDDVSVVPLR| 7                              |             |                  |                 |
| PPDAASSAPLR| 203                            | 156         | 155              | 1.19            |
| APARSSPP   | 11                              |             |                  |                 |
| APPPSSLPSP | 9                               |             |                  |                 |
| PPAGSSAPG  | 226                            | 476         | 230              | 0.483           |
| PPASSAPG   | 62                             |             |                  |                 |
| RTPPP      | 13                              |             |                  |                 |
| RSPPP      | 7                               |             |                  |                 |

* \( K_m \) or \( V_{max} \) not determined as it was linear up to 800 μM of peptide concentration.

TABLE II
Ser:GlcNAc phosphotransferase has a preference for certain peptide sequences

The activity of Ser:GlcNAc phosphotransferase was measured with peptides having single or multiple Ser residues flanked by various amino acids. The peptides were used at 800 μM, and the assay was performed at room temperature and at pH 7.5. \( K_m \) and \( V_{max} \) were determined as described in Table I.

The activity of Ser:GlcNAc phosphotransferase was measured with peptides having single or multiple Ser residues flanked by various amino acids. The peptides were used at 800 μM, and the assay was performed at room temperature and at pH 7.5. \( K_m \) and \( V_{max} \) were determined as described in Table I.

analyzed by MALDI-MS to determine the number of GlcNAc-1-P residues. Peptide 1 product did not yield any desorbed ions in the expected mass range (700–2000) when analyzed using four different matrices in either the positive ion or negative ion mode. Products of peptides 2 and 3 gave strong signals in the negative ion mode using the DAN and HABA matrices, respectively. Fig. 1B shows the negative ion mass spectra for peptides 2 and 3 (panels i and iii) and the HPLC-purified transferase products (panels ii and iv). The calculated mass for peptide 2 is 720.8 and that for peptide 3 is 862.9. Panel i shows two major signals at 719.2 and 815.3, corresponding to peptide 2 (expected [M – H] \(^+\) = 719.8) and its trifluoroacetyl ester (+ 96 mass units). Panel ii shows the HPLC-purified transferase products of peptide 2 with negative ion masses of 1002.6 and 1099.5 (Δm/z = 96.1; trifluoroacetyl ester). These ions are 283 mass units larger than the corresponding peptide ions shown in panel i, as expected for GlcNAc-1-phosphoryl derivatives (mass = 2832). Panel iii shows a major negative ion at 861.9 for peptide 3 (expected [M – H] \(^+\) = 861.9). Panel iv shows the major negative ion at 1145.6 for the HPLC-purified peptide 3 transferase product, 283.7 mass units larger than the peptide, as expected for the GlcNAc-1-phosphoryl derivative. A second product ion at 1438.2 is 296.6 mass units larger than the major product ion and, therefore, is probably not due to an additional GlcNAc-1-phosphoryl group. Thus, MALDI-MS shows only one GlcNAc-1-P per peptide.

Phosphotransferase Activity on Other Serine-containing Synthetic Peptides—None of the three peptide motifs is highly preferred as an acceptor, but the neighboring amino acids are expected to influence recognition and efficiency. We tested a limited series of 12 peptides having single or multiple Ser residues and different flanking amino acids. None of these is found in the Ser-rich domain of the known Dictyostelium CPs. As shown in Table II, glycoamidase values vary by at least 50-fold. In a homologous series of peptides containing a single Ser flanked by various amino acids, it was clear that flanking by Ala (PAAASAPALR) was much better than either Gly (PPDGSGGPLR) or Val (PPDVSVVPLR). A peptide with multiple contiguous Ser residues flanked by single Ala residues (PPDAASSAPLR) was the best acceptor having a 5-fold better efficiency than the Ser motifs seen in the Dictyostelium CPs but nearly the same \( K_m \). Serines flanked by basic amino acids (APARSSPP) or by Pro (APPSSLPSP) residues were poor acceptors. This limited survey clearly shows that flanking amino acids influence transferase activity. However, the specificity of phosphoglycosylation seen in Dictyostelium suggested that other features might also contribute to substrate recognition.

Ser:GlcNAc Phosphotransferase Activity on Endogenous Acceptors—Detergent-solubilized Dictyostelium microsomal membranes incorporate label from UDP-[\(^3\)H]GlcNAc into endogenous acceptors. Incorporation is linear for up to 2 h and from 1.6–6.5 mg/ml protein as reported in earlier studies (15, 18).

In Dictyostelium, UDP-[\(^3\)H]GlcNAc can be used by two different GlcNAc-1-P transferases as follows: one is Ser:GlcNAc phosphotransferase, and the other is UDP-N-acetylglucosamine:lysosomal enzyme N-acetylgalactosamine-1-phosphotransferase. The latter transfers GlcNAc-1-P to selected mannose residues on N-linked oligosaccharides of lysosomal enzymes to generate the Man-6-P marker (22). A GlcNAc transferase adds an “intersecting” GlcNAc residue to high mannose N-linked chains (23). The N-linked glycopeptides bind to concanavalin A-Sepharose, whereas GlcNAc-1-P-Ser peptides do not (15). Under our assay conditions, >90% of the label is incorporated into glycopeptides that do not bind to the lectin column and are therefore likely to be Ser:GlcNAc phosphotransferase products.

The \(^3\)H-labeled products were degraded by base and mild acid treatments as described above to characterize their linkage to the acceptors. Base released 86% of the label of which 80% bound to QAE-Sephadex before but not after digestion with alkaline phosphatase. TLC analysis of the base and acid-hydrolyzed material gave exclusively [\(^3\)H]GlcNAc-1-P and GlcNAc, respectively (not shown). These results confirmed that most of the product is [\(^3\)H]GlcNAc-1-P in a diester linkage to the endogenous acceptors, presumably through serine residues.

Ser:GlcNAc Phosphotransferase Specifically Recognizes a Native Dictyostelium Lysosomal CP—Solubilized membranes and UDP-[\(^3\)H]GlcNAc were mixed with increasing amounts of proteinase-1(CP7) as a potential acceptor for GlcNAc-1-P. Rather than stimulating overall incorporation, the added protein inhibited incorporation into endogenous acceptors (Fig. 2A). To eliminate the possibility that inhibition resulted from residual proteolytic activity (24), proteinase-1(CP7) was inactivated with several CP-specific inhibitors prior to incubation. However, this inactivation made proteinase-1(CP7) an even better inhibitor of transferase activity (Fig. 2A and Table III). Thus, proteolysis does not cause the loss of transferase activity.

Proteinase-1(CP7) gave slightly different transferase inhibition curves depending on the CP inhibitor used. E-64, a covalent inhibitor (25), gave a 50% inhibition at 1.6 μM as compared with leupeptin, a reversible peptide inhibitor (26) which gave 50% inhibition at 2.8 μM. The best inhibitor was cystatin, a small protein inhibitor with high affinity for CPs (27), which forms a complex that inhibited transferases at only 0.4 μM (Fig. 2A, Table III). None of the inhibitors by themselves inhibit Ser:GlcNAc phosphotransferase activity when assayed using either endogenous or peptide acceptors. The CP-inhibitor complexes also do not accelerate the rate of UDP-GlcNAc breakdown (not shown). Significantly, when proteinase-1(CP7) is boiled in SDS and the residual SDS complexed with bovine serum albumin, the denatured CP no longer inhibits endogenous acceptor phosphoglycosylation. Since control incubations with SDS alone do not cause inhibition, this finding shows that a conformational feature of proteinase-1(CP7) is required to inhibit transferase activity.

Ser:GlcNAc Phosphotransferase Also Recognizes Native Mammalian and Plant Cathepsins—Except for their Ser-rich domains, Dictyostelium CPs are highly homologous to other CPs (5, 6). If Ser:GlcNAc phosphotransferase recognizes a por-
Ser:GlcNAc phosphotransferase recognizes a conformation feature of eukaryotic cysteine proteinases. Ser:GlcNAc phosphotransferase assay was performed using endogenous acceptors in the presence of indicated amounts of added purified CPs (A and B) or acceptor peptides 1 and 2 (C). Where indicated, the CP was first inactivated with cystatin, E-64, or leupeptin (Leu). The residual Ser:GlcNAc phosphotransferase activity was calculated from a control incubation without any exogenous CP. P1, proteinase-1(CP7), Cath 02, cathepsin 02; Cath S, cathepsin S.

**TABLE III**

| Proteins | 50% inhibition |
|----------|---------------|
| P1       | 10.0          |
| P1 + cystatin | 0.4       |
| P1 + leupeptin | 2.8      |
| P1 + E-64 | 1.6          |
| Papain   | 12.0          |
| Papain + cystatin | 6.5     |
| Cathepsin S + cystatin | 3.0  |
| Cathepsin 02 + cystatin | 2.1  |
| Cathepsin L + cystatin | 1.5  |
| Procathepsin B + cystatin | 2.0  |
| Bromelain + E-64 | 10.0 |
| Bromelain | None at 26    |
| Pepsin   | None at 90    |
| Thryroglobulin | None at 10 |
| Bovine serum albumin | None at 100 |
| Ovalbumin | None at 140   |

Endogenous Acceptor Complex—Peptide acceptors are efficiently glycosylated by well-washed, sonicated membrane preparations whether detergents are added to the assay or not, showing that the peptides have full access to the transferase. This is also true of endogenous acceptors; they are phosphoglycosylated nearly as well (75%) when detergents are omitted from the assay (Table IV). Since any freely soluble endogenous acceptors would be removed by extensive washing, this finding may mean that transferase and endogenous acceptors are both membrane-bound and mobile in the plane of the membrane or that they form a complex. To explore this possibility, washed membranes were detergent-solubilized and centrifuged to remove particles and aggregates >100S. The clear solution containing solubilized proteins was applied to a 10–40% sucrose gradient, and fractions were assayed for Ser:GlcNAc phosphotransferase activity using both peptide and endogenous acceptors. Fig. 3A shows a gradient centrifuged for 3 h. Under these conditions the activities for endogenous and peptide acceptors are clearly separated from each other with very little overlap. The peptide activity has a calculated apparent S value of >60, whereas endogenous acceptors ran at <60 S. Since most of the protein solubilized by the detergent co-fractionated with the activity for endogenous acceptors, the transferase may be using solubilized acceptors rather than those in the putative complex. Therefore, the sample was centrifuged for 16 h, and under these conditions, most of the soluble protein is well separated from the major peak of endogenous acceptor activity. About 30% of the input activity on endogenous acceptors is recovered as a peak having an apparent size ~18 S based on the position of 18 S and 28 S ribosomal RNA markers (Fig. 3B). These results show that the Ser:GlcNAc phosphotransferase activity on peptide and endogenous acceptors is physically separable and that a portion of the transferase probably exists as a complex with endogenous acceptors. The nature of this association is unknown.

Endogenous acceptor activity sediments further in the gradient with increasing centrifugation time, but this does not occur for peptide acceptor activity. It remains at nearly the same position in the gradient regardless of the centrifugation time. At 3 h, the calculated density of the peak is 1.075 g/ml and at 16 h it is 1.095 g/ml (Fig. 3B). Similar results are seen if fractions 7–12 are pooled from the 3-h gradient (Fig. 3A) and centrifuged for 16 h (Fig. 3C); the transferase activity still bands at 1.090 g/ml. This probably means that the transferase is at or near its isopycnic density. Since this is much lower than the density of pure protein (1.33–1.44 g/ml), the transferase is probably associated with cellular lipids or octyl glucoside present in the gradient.

CPs Inhibit the Ser:GlcNAc Phosphotransferase Activity for Peptide Acceptors—Cystatin-inactivated CPs also inhibit phosphoglycosylation of peptide 1 (C6GSSSSSSG) using endoge-
Table IV
Ser-GlcNAc phosphotransferase is closely associated with endogenous CP acceptors

| Sample | Detergent | Peptide 1 | Endogenous |
|--------|-----------|-----------|-------------|
| 1      | 1 ×       | 115       | 6.4         |
| 2      | 1 ×       | 119       | 8.0         |
| 3      | 0         | 102       | 5.7         |
| 4      | 0.1 ×     | 110       | 6.0         |
| 5      | 0.2 ×     | 93        | 7.1         |
| 6      | 0.5 ×     | 100       | 7.0         |
| 7      | 1 ×       | 117       | 7.2         |

Crude microsomal membranes were prepared as detailed under “Experimental Procedures” (sample 1) or the membrane fraction was washed three times by centrifugation at 100,000 × g for 15 min with excess 50 mM Tris/HCl, pH 7.5, buffer (samples 2–7). Ser:GlcNAc phosphotransferase activity from these microsomal membranes was assayed using peptide 1 and endogenous acceptors, either in the presence or absence of various detergent concentrations. The results are the average of two separate determinations. 1 × detergent concentration is 1 mM sodium deoxycholate and 0.17% Lubrol.

Expression and Purification of Mature-FLAG-CP7 Fusion Protein—Cystatin-complexed proteinase-1 (CP7) is the best inhibitor giving 50% inhibition at 0.6 μM, and the cathepsins are effective at 2–4 μM (Fig. 4), showing that CPs are recognized equally well by trans-ferase whether it is associated with endogenous acceptors or not. Additional kinetic experiments with CPs were unable to confidently resolve the type of inhibition occurring on peptide acceptors.

Numerous attempts at expressing full-length CP7 in S. cerevisiae and P. pastoris were unsuccessful. These included GST fusion products, with or without a variety of epitope tags (Myc, FLAG, and His) introduced at several locations in the molecule. Only trace amounts of the protein were sometimes produced. Similar results were seen in trying to express these constructs in E. coli, where expression was usually lethal. However, we were able to successfully overexpress a mature form of proteinase-1 (CP7) as a FLAG-tagged GST fusion protein shown in Fig. 5A.

Mature-FLAG-CP7-GST (fusion protein) was expressed only in isopropyl-1-thio-β-D-galactopyranoside-induced cells (Fig. 5B, lanes 1 and 2), where it accounted for ~20% of the cell protein and was mostly found in inclusion bodies. Although the protein contained a GST tag, the soluble form did not bind to glutathione beads and required purification by DEAE-Sephadex chromatography (Fig. 5B, lane 3). The final yield was 48 mg/liter culture. The purified fusion protein reacts with anti-FLAG mAb, M2 (lane 4), but not with GlcNAc-1-P mAb, AD7.5 (lane 5) indicating that it is not phosphoglycosylated in E. coli. The apparent molecular mass, 65 kDa, is near its theoretical mass of 63 kDa.

FLAG-CP7-GST Fusion Protein Is an Acceptor for Ser-GlcNAc Phosphotransferase—Solubilized Dictyostelium membranes incubated with purified fusion protein and UDP-GlcNAc produced a new band located just below the 82-kDa marker that reacted with the GlcNAc-1-P antibody AD7.5 Fig. 5B (lane 7), whereas membranes incubated in the absence of fusion protein lacked this band (lane 6). FLAG antibody M2 confirmed that the new band is the fusion protein (lane 9) which is absent from the membrane preparation incubated without fusion protein (lane 8). The product of [3H]UDP-GlcNAc-labeled fusion protein was also confirmed to be [3H] GlcNAc-1-P by the same acid and base lability criteria used for peptide and endogenous acceptors. The fusion protein had a K_m of 5.6 μM (Fig. 5C and inset) when incubated with endogenous acceptor-free trans-ferase. This is comparable to the CP concentration needed for 50% inhibition on endogenous acceptors (Table III) and is 27-fold better than peptides with Ser-rich domains. The fusion protein appears to have at least some of the CP conformation recognized by the trans-ferase since SDS-denatured fusion protein is not an acceptor when used up to 20 μM (Fig. 5C).
The specificity of Ser:GlcNAc phosphotransferase appears to result from recognition of preferred serine peptides and a conformational feature shared by papain-like cysteine proteinases. Since GlcNAc-1-P-Ser and Man-6-P are mutually exclusive formational feature shared by papain-like cysteine proteinases. Since GlcNAc-1-P-Ser and Man-6-P are mutually exclusive, it seemed likely that one or more of the three serine motifs is modified. The in vitro results presented here show that a single GlcNAc-1-P can be added to the peptides (Fig. 1) with about equal efficiency (Table I), but no single motif is highly preferred. Other peptides with single or multiple serine residues in contexts different than those found in the CP sequences are also acceptors. They show a 50-fold preference range and provide evidence that the flanking amino acids strongly influence glycosylation. A serine flanked by alanines is a 13-fold better acceptor than serine flanked by valines (Table II). No “phosphoglycosylation consensus” sequence is apparent from this limited survey, but peptide phosphoglycosylation is clearly not restricted to only those motifs found in the Dictyostelium CPs. Threonine-containing peptides were not analyzed in detail since earlier studies showed that GlcNAc-1-P is added to Ser residues (2, 15).

We considered the possibility that Ser:GlcNAc phosphotransferase might specifically recognize structural features of the Dictyostelium CPs thus providing an added degree of specificity. Inhibition of endogenous acceptor phosphoglycosylation by exogenous proteinase-1(CP7), but not by SDS-denatured protein, suggested that the transferase recognizes a specific conformation of proteinase-1(CP7) (Fig. 2). Inhibition of transferase improved when the protein was complexed with various CP inhibitors which may induce or stabilize a preferred conformation at neutral pH (7.5) where many CPs are unstable. However, the inhibitory potency of cystatin-inactivated CP was similar at pH 6.0 and pH 7.5 (not shown), arguing against stabilization of a pH-dependent inactivation.

**DISCUSSION**

**Substrate Specificity of Ser:GlcNAc Phosphotransferase**

The residual Ser:GlcNAc phosphotransferase activity is plotted as % control that contained no CP.

![Graph showing CPs inhibit the Ser:GlcNAc phosphotransferase activity on peptides](image)

**FIG. 4.** CPs inhibit the Ser:GlcNAc phosphotransferase activity on peptides. Ser:GlcNAc phosphotransferase activity containing very low activity on endogenous acceptors (fractions 7–12 from Fig. 3A) was assayed with peptide 1 (C8GSSSSSSG) in the presence of increasing amounts of cystatin-inactivated proteinase-1(CP7) or mammalian cathepsins. The residual Ser:GlcNAc phosphotransferase activity is plotted as % control that contained no CP.

**FIG. 5.** Mature FLAG-CP7 fusion protein is a much better acceptor for Ser:GlcNAc phosphotransferase than peptides. A, the GST fusion protein construct was made by an in-frame fusion of GST to the N terminus of the mature CP7 (amino acids (aa) 109). The FLAG epitope was inserted at amino acids 282 immediately before the start of the Ser-rich domain (6). The locations of the active site histidine (H) and cysteine (C) residues are indicated. B, lane 1, uninduced E. coli containing fusion protein; lane 2, same as lane 1 after induction with isopropyl-1-thio-β-D-galactopyranoside; lane 3, 0.2 μg of purified fusion protein separated on 12.5% SDS-polyacrylamide gel electrophoresis and stained with Coomassie Blue; lane 4, 1 ng of purified fusion protein probed with anti-FLAG antibody (M2) after blotting on nitrocellulose; lane 5, 1 ng of purified fusion protein probed with GlcNAc-1-P-specific antibody AD7.5, after blotting; Dictyostelium membrane preparations incubated with UDP-GlcNAc either in the absence (lanes 6 and 8) or presence (lanes 7 and 9) of fusion protein and analyzed by immunoblotting with GlcNAc-1-P antibody AD7.5 (lanes 6 and 7) or anti-FLAG M2. C, increasing concentrations of native or SDS-denatured fusion protein were incubated with a Ser:GlcNAc phosphotransferase fraction that contained very low activity on endogenous acceptors (fractions 7–12, Fig. 3A). The activity, V₀ (pmol/ml/min) was plotted against substrate concentration [S] (μM). The inset shows a Lineweaver plot to determine Kₘ of 5.6 μM.
Initial attempts to use proteinase-1(CP7) as an acceptor for GlcNAc-1-P were unsuccessful, probably because the glycosylation sites were already occupied, but expression of a mature form of a FLAG-tagged GST fusion protein was successful. This protein does not have proteolytic activity and cannot bind to cystatin or glutathione beads, indicating that it does not have a native conformation. Nevertheless, the \( K_m \) of this acceptor (5.6 \( \mu \)M) is comparable to those seen for other CPs. The additional finding that SDS denaturation abolishes its acceptor activity argues that the fusion protein retains sufficient conformation to be recognized reasonably well by the transferase (Fig. 5). This may prove useful for affinity purification of the transferase.

It is clear that cytosol-free microsomal membranes contain both transferase and endogenous acceptors since they show nearly 75% of their full activity in the absence of any detergents (Table IV). Sucrose gradient fractionation of solubilized membranes separates transferase activities for the peptide and endogenous acceptors. We do not know the basis of the separation, but the transferase without endogenous acceptors appears to have a low buoyant density suggesting an association with cellular lipids. The low density is unlikely to result from residual large membranous aggregates, since the initial centrifugation conditions were designed to pellet particles >100 S. The transferase and endogenous acceptors appear to cofractionate as a complex with an apparent size of ~18 S based on rRNA standards. More work will be needed to define the nature of this interaction and to determine if this is a true molecular complex.

Cystatin-inactivated proteinase-1(CP7) and the other CPs inhibit peptide glycosylation using endogenous acceptor-free Ser:GlcNAc phosphotransferase at similar concentrations to those for endogenous acceptors (Fig. 4). However, >800 \( \mu \)M peptide is necessary to inhibit 30–40% of the activity on endogenous acceptors, suggesting that the peptides cannot compete or displace endogenous acceptors. The transferase has a considerable preference for recognizing endogenous acceptors over the individual peptides that it modifies. Although the transferase prefers CPs, the wide range of activities on peptide acceptors probably means that GlcNAc-1-P addition is selective but not exclusive for CPs. Together, these results suggest that the transferases may have separate but interacting or overlapping recognition sites, one for the peptide and the other for the preferred conformation of the CPs. More refined analyses will be needed to resolve this complex recognition of these multiple substrates.

Only three protein specific glycosyltransferases have been identified. These include the UDP-N-acetylgalcosamine:lyosomal enzyme N-acetylgalcosaminyltransferase (10, 11), UDP-GalNAc:glycoprotein hormone N-acetylgalactosaminyltransferase (12–14), and UDP-Glc:glycoprotein glucosyltransferase (29). The structures produced by these glycosyltransferase have important biological functions such as lysosomal targeting (10, 11), regulation of circulatory half-life (30), and quality control of newly synthesized glycoproteins (31). In Dictyostelium, Ser:GlcNAc phosphotransferase and the UDP-N-acetylgalcosamine:lyosomal enzyme N-acetylgalcosamine-1-phosphotransferase donate GlcNAc-1-P to mutually exclusive populations of proteins. These proteins are located in different vesicles of the endo-lysosomal system and function sequentially in the digestion of phagocytosed bacteria (4). This intriguing correlation suggests that the two sugar modifications may help to establish or maintain the segregation of these proteins. The preferential recognition of eukaryotic CPs provides a physical basis for this potentially important glycosylation of selected proteins in Dictyostelium.

Acknowledgments—We thank the following people for the generous gift of purified cathepsins: Dr. D. Bromme for cathepsins S, L, and O2; Dr. J. S. Mort for human procathepsin B; and Dr. R. W. Mason for cathepsin L. We also thank Dr. Åke Elhammer and Dr. Robert Hill for the various synthetic peptides and Dr. Guy Salvesen for critical review of the manuscript.

REFERENCES

1. Freeze, H. H. (1997) in Comprehensive Biochemistry (Neuberger, A., and Van Deenen, L. L. M., eds) Vol. 29B, pp 91–105, Elsevier Science, The Netherlands
2. Gustafson, G. L., and Gander, J. E. (1984) Methods Enzymol. 107, 172–183
3. Mehta, D. P., Ichikawa, M., Salimath, P. V., Etchison, J. R., Haak, R., Manzi, A., and Freeze, H. H. (1996) J. Biol. Chem. 271, 10987–10993
4. Souza, G. M., Mehta, D. P., Lammertz, M., Rodriguez-Parras, J., Wu, R., Cardelli, J. A., and Freeze, H. H. (1997) J. Cell Sci. 110, 2239–2248
5. Souza, G. M., Hirai, J., Mehta, D. P., and Freeze, H. H. (1995) J. Biol. Chem. 270, 28838–28845
6. Ord, T., Adessi, C., Wang, L., and Freeze, H. H. (1997) Arch. Biochem. Biophys. 339, 64–72
7. Kornfeld, R., and Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631–664
8. Elhammer, A. P., Poorman, R. A., Brown, E., Maggiora, L. G., Hoogerheide, J. G., and Kezdy, F. J. (1993) J. Biol. Chem. 268, 10029–10038
9. Hanssen, J. E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, O. J., Hansen, J.-E. S., and Brunak, S. (1995) Biochem. J. 308, 801–813
10. Lang, L., Reitman, M., Tang, J., Robert, R. M., and Kornfeld, S. (1984) J. Biol. Chem. 259, 14663–14771
11. Kornfeld, R., and Mellman, I. (1989) Annu. Rev. Cell Biol. 5, 483–525
12. Smith, P. L., and Baenziger, J. U. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 329–322
13. Baenziger, J. U. (1994) FASEB J. 8, 1019–1025
14. Mazzola, S. M., Hooper, L. V., and Baenziger, J. U. (1996) J. Biol. Chem. 271, 12117–12120
15. Merello, S., Parodi, J. A., and Cousse, R. (1995) J. Biol. Chem. 270, 7281–7287
16. Sussman, M. R. (1987) Methods Cell Biol. 26, 9–29
17. Freeze, H. H., Hindsuval, O., and Ichikawa, M. (1992) J. Biol. Chem. 267, 4431–4439
18. Freeze, H. H., and Ichikawa, M. (1995) Biochem. Biophys. Res. Commun. 208, 384–389
19. Dawson, R. M. C., Elliott, D. C., Elliott, W. H., Jones, K. M. (1986) J. Biol. Chem. 261, 6475–6480
20. Real, S., Cardelli, J. A., and Freeze, H. H. (1997) J. Biol. Chem. 272, 189–198
21. Capasso, J. M., Capasso, A., and Kaplan, A. (1986) Biochim. Biophys. Acta 860, 151–152
22. Barrett, J. R., and Freeze, H. H. (1996) Glycobiology 6, 177–189
23. Miller, G. L. (1959) Anal. Chem. 31, 964–967
24. Lang, L., Cousse, R., and Kornfeld, S. (1986) J. Biol. Chem. 261, 6520–6525
25. Capasso, J. M., Capasso, A., and Kaplan, A. (1996) Biochim. Biophys. Acta 1281, 15–22
26. Barrett, J. R., and Freeze, H. H. (1996) Biochemistry 35, 402–404