Identification of the Minimal Lysosomal Enzyme Recognition Domain in Cathepsin D*

Received for publication, June 1, 2005, and in revised form, July 19, 2005. Published, JBC Papers in Press, August 4, 2005, DOI 10.1074/jbc.M505994200

Richard Steet 1, Wang-Sik Lee 1, and Stuart Kornfeld 2

From the Department of Internal Medicine, Washington University School of Medicine, Saint Louis, Missouri 63110

Specific recognition of lysosomal hydrolases by UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase, the initial enzyme in the biosynthesis of mannose 6-phosphate residues, is governed by a common protein determinant. Previously, we generated a lysosomal enzyme recognition domain in the secretory protein glycopepsinogen by substituting in two regions (lysine 203 and amino acids 265-293 of the β loop) from cathepsin D, a highly related lysosomal protease. Here we show that substitution of just two lysines (Lys-203 and Lys-267) stimulates mannose phosphorylation 116-fold. Substitution of additional residues in the β loop, particularly lysines, increased phosphorylation 4-fold further, approaching the level obtained with intact cathepsin D. All the phosphorylation occurred at the carboxyl lobe glycan, indicating that additional elements are required for phosphorylation of the amino lobe glycan. These data support the proposal that as few as two lysines in the correct orientation to each other and to the glycan can serve as the minimal elements of the lysosomal enzyme recognition domain. However, our findings show that the spacing between lysines is flexible and other residues contribute to the recognition marker.

The enzyme UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase (abbreviated "phosphotransferase") 3 has a central role in the targeting of newly synthesized lysosomal hydrolases to lysosomes. This enzyme recognizes a conformation-dependent protein determinant shared by the many different lysosomal hydrolases and then transfers N-acetylglucosamine 1-phosphate from UDP-GlcNAc to selected mannose residues of the high mannose oligosaccharides of the hydrolases. The N-acetylglucosamine residues are subsequently excised by "uncovering enzyme" generating mannose 6-phosphate monoesters that allow high affinity binding to mannose 6-phosphate receptors in the Golgi and translocation to lysosomes (1).

The nature of the common protein recognition determinant has received considerable attention because of its role in this targeting pathway. Our laboratory undertook a molecular dissection of the phosphotransferase recognition marker using a pair of aspartyl proteases, the lysosomal hydrolase cathepsin D and the secretory protein pepsinogen (2, 3). These proteins are 45% identical in amino acid sequence and share similar secondary and tertiary structure but differ in that cathepsin D is highly phosphorylated by phosphotransferase, whereas pepsinogen, the glycosylated form of pepsinogen, is not. By analyzing numerous chimeric proteins, we were able to define a phosphotransferase recognition patch in the carboxyl lobe of cathepsin D formed by two noncontinuous primary sequences, specifically Lys-203 and amino acids 265–293 that form a loop structure directed toward Lys-203 (2). Lysine residues 267 and 293 at the base of the loop made significant contributions to the degree of phosphorylation, implicating lysines as important elements to the recognition patch. It was further observed that the presence of additional regions of cathepsin D enhanced phosphorylation of the chimeric proteases, indicating that the recognition domain is a surface patch that contains multiple interacting sites and perhaps even a second independent site (4–6).

The critical role of lysines in defining the recognition domain has also been supported by studies from the Sahagian laboratory that have demonstrated a significant decrease in the mannose phosphorylation upon mutation of specific pairs of lysine residues in cathepsin D (Lys-203 and Lys-293) and cathepsin L (Lys-54 and Lys-99). Due to the fact that these lysine pairs are spaced 34 Å apart and positioned in a specific orientation relative to the target oligosaccharide in these proteases, they proposed a simple model in which just two critical lysine residues are necessary and sufficient for binding to phosphotransferase (7). However, because residual phosphorylation was observed in the absence of these critical lysine pairs, they pointed out that additional protein elements might be important. Further evidence for the role of lysines as components of the phosphotransferase recognition domain has come from mutagenesis studies of DNase I (8), aspartylglucosaminidase (9), and arylsulfatase (10).

In the current study, we have utilized both a gain-of-function (substituting cathepsin D residues into glycopepsinogen) and a loss-of-function (mutation of cathepsin D residues to alanine) approach in an attempt to determine the minimal phosphotransferase recognition domain. In contrast to our previous work that mostly utilized relatively large segments of sequence, we substituted single residues or small blocks of residues into glycopepsinogen to better analyze the role of individual amino acids in facilitating phosphorylation. Our experiments provide new insight into the nature of the phosphotransferase recognition marker.

MATERIALS AND METHODS

Cell Lines, Plasmids, and Reagents—COS-1 and HeLa cells were obtained from ATCC and maintained in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum and 100 μg/ml penicillin and streptomycin. LoVo cells were also obtained from ATCC and maintained in Ham’s F-12 medium containing 10% fetal bovine serum and 100 μg/ml penicillin and 100 units/ml streptomycin. Human glycopepsinogen cDNA and cathepsin D (bearing a C-terminal myc tag) cDNA were inserted into the pcDNA 3.1(+) expression vector. All mutations in these constructs were introduced using QuikChange site-directed mutagenesis (Stratagene) protocols. In the cathepsin D-myc experiments, lysine residues were always changed to alanine residues. Anti-myc polyclonal antibody (A-14) was from Santa Cruz Biotechnology.
Anti-human pepsinogen antiserum was prepared as described previously (11). [2-3H]Mannose was purchased from PerkinElmer Life Sciences and [35S]methionine/cysteine from MP Biomedicals. Concanavalin A-Sepharose was obtained from Amersham Biosciences, Protein A-agarose from Repligen Corp., and QAE-Sephadex from Sigma. Recombinant endoglycosidase H fused to maltose-binding protein was from New England Biolabs. All other reagents were of the highest quality available and were purchased from Sigma or Fisher.

Transfection and Labeling of Cells—Cells cultured in 60-mm plates for 24 h (80–90% confluent) were transfected with 2–3 μg of DNA using Lipofectamine Plus reagent system (Invitrogen) in the absence of antibiotics. At 24 h post-transfection, cell monolayers were washed and labeled with 200 μCi of [2-3H]mannose in Dulbecco’s modified Eagle’s medium, 10% dialyzed fetal bovine serum, and 1 mM glucose at 37 °C for 2 h. Following addition of glucose and mannose up to 5 mM to stop mannose uptake, the cells were incubated for 4 h before the media was collected and cleared by centrifugation to remove cell debris. Ammonium chloride (10 mM) was routinely present during the labeling and chase periods, because it resulted in higher secretion of most of the mutant constructs.

Immunoprecipitation and Oligosaccharide Analysis—Cell media samples were immunoprecipitated, and oligosaccharides were isolated and analyzed essentially as described in detail previously (6). For the cathepsin D-myc experiments, anti-myc antibody (50 μl) was bound to Protein A-agarose beads prior to immunoprecipitation of labeled cathepsin D from the media. Oligosaccharides released by Endo H digestion of immunoprecipitated protein, which include the phosphorylated molecules, were treated with mild acid to remove any N-acetylglucosamine residues still attached to the phosphate moieties. The percent phosphorylation was calculated as cpm recovered in Endo H-treated oligosaccharides with two mannose 6-phosphate monoesters / cpm recovered in Endo H-resistant oligosaccharides (phosphorylated plus neutral) plus cpm recovered in Endo H-resistant complex oligosaccharides. The values for the complex oligosaccharides were multiplied by 2 to correct for the fact that they contain 3 mannose residues versus an average of 6 mannose residues per high mannose oligosaccharide. In all cases, values for each oligosaccharide pool obtained in routine mock transfection experiments were subtracted prior to percent phosphorylation calculations for the constructs. The mock values for phosphorylated species were 5 cpm or less in the LoVo cell experiments. The analysis of Endo H-released oligosaccharides by QAE-Sephadex chromatography allows for the separation of oligosaccharides bearing one or two mannose 6-phosphate residues.

Acid-activated Autoactivation of Cathepsin D Constructs—HeLa cells transfected with the various cathepsin D-myc constructs were labeled with 0.5 mCi of [35S]methionine/cysteine for 2 h and chased for 4 h by the addition of cold methionine (up to 10 mM). Secreted cathepsin D molecules were immunoprecipitated with anti-myc polyclonal antibody. Washed beads were resuspended in 50 mM Tris-glycine buffer, pH 3.5, in the absence or presence of 500 μM pepstatin A, and incubated for 20 min. at 37 °C. Correctly folded molecules undergo autoactivation with cleavage of their propeptides at this pH. This process is blocked by the presence of the aspartyl protease inhibitor, pepstatin A. Following neutralization with 10–12 μl of 1.5 M Tris-HCl, pH 8.8, labeled molecules were resolved by 8% SDS-polyacrylamide gel electrophoresis and visualized by autoradiography.

RESULTS

Influence of Cell Type on Glycopepsinogen Phosphorylation—We first expressed glycopepsinogen and two constructs in COS, HeLa, and LoVo cells to determine which is the most appropriate to use in the study. LoVo cells lack “uncovering enzyme” activity and are unable to form mannose 6-phosphate monoesters required for binding to mannose 6-phosphate receptors (12). To obtain accurate values of mannose 6-phosphate formation, the transfected cells were labeled with [2-3H]mannose followed by immunoprecipitation of the secreted glycopepsinogen molecules. The [3H]mannose-labeled oligosaccharides were then analyzed for the level of phosphorylation as described previously (6).

The phosphorylation of secreted glycopepsinogen and two constructs (E203K/A267K/Q293K or “3K”, termed construct 5 and “3K + loop segments ABC”; termed construct 8) expressed in COS, HeLa, and LoVo cells is shown in Table One. All three cell types phosphorylated the constructs equally, with LoVo cells producing the highest level of phosphorylation and the highest ratio of oligosaccharides with two mannose 6-phosphate residues versus one mannose 6-phosphate residue. Based on these preliminary results, we utilized LoVo cells for the subsequent studies of modified glycopepsinogen.

Effects of Inserting Cathepsin D Residues on Glycopepsinogen Phosphorylation—in our previous studies, we found that substitution of cathepsin D residues Lys-203 plus amino acids 265–293 into the analogous positions of glycopepsinogen generated a phosphotransferase recognition marker (2, 3). Residues 265–293 form a β loop structure and by substituting the residues of the loop that were calculated to be >20% accessible with alanines (13), we found that Lys-267 was most essential followed by Lys-293 (3). In addition, Lys-203 was absolutely required for this construct to be phosphorylated. The construct used in this particular alanine-scanning mutagenesis experiment contained Lys-203 plus cathepsin D residues 265–319 (3). The impact of just substituting lysines at positions 203, 267, and 293 into glycopepsinogen was examined.

We have now tested a new series of constructs with the goal of identifying the minimal number of residues required to generate an effective phosphotransferase recognition motif (Table Two). The focus was on Lys-203 and the components of the β loop, residues 265–293. Fig. 1 is a schematic representation of cathepsin D showing the location of its
Minimal Lysosomal Enzyme Recognition Elements

TABLE TWO

Effect of cathepsin D residues on glycopepsinogen phosphorylation

All experiments were performed in LoVo cells. Values for the % phosphorylation were determined as described under "Materials and Methods" and represent the fraction of phosphorylated high mannose oligosaccharides within the total pool of high mannose and complex-type glycans. The mock values for phosphorylated species were 5 cpm or less. Constructs were tested 1–5 times each (average values ± S.D. is shown). Average expression denotes the mean total CPM recovered from immunoprecipitation of various constructs.

| Construct | No. of amino acid substitutions | CPM | Oligosaccharide phosphorylation | Ratio 2P/1P | -fold stimulation |
|-----------|----------------------------------|-----|--------------------------------|------------|-----------------|
| Glycopepsinogen (GP) | 0 | 8,200 | 0.022 ± 0.004 | " | 1 |
| #1 E203K | 1 | 16,000 | 0.085 | " | 4 |
| #2: A267K/Q293K | 2 | 5,400 | 0.45 ± 0.01 | 0.04 | 20 |
| #3: E203K/Q293K | 2 | 3,000 | 0.47 | 0.71 | 21 |
| #4: E203K/A267K | 2 | 4,800 | 2.6 ± 0.3 | 0.71 | 116 |
| #5: E203K/A267K/Q293K; “3K” | 3 | 5,700 | 3.4 ± 0.4 | 0.71 | 156 |
| #6: “3K” plus T277K/V281K/P284K | 6 | 7,100 | 5.4 ± 0.5 | 1.4 | 245 |
| #7: “3K” plus loop AB | 12 | 1,100 | 7.7 ± 0.1 | 1.5 | 350 |
| #8: “3K” plus loop ABC | 15 | 3,500 | 9.7 ± 1.0 | 1.8 | 441 |
| Cathepsin D | 19,000 | 42.3 ± 1.7 | 1.5 | 1,923 |

* Radioactivity in phosphorylated species too low to calculate ratio.

FIGURE 1. Structural organization of cathepsin D and alignment of β loop amino acid sequences of human cathepsin D and pepsinogen. The number of residues is designated with respect to cathepsin D. N-Linked glycosylation sites are indicated by arrows. Identical amino acids between cathepsin D (CD) and pepsinogen (GP) are boxed. The position of the five lysine residues present in the β loop structure are noted along with the position of lysine 203 (asterisk). Note that pepsinogen lacks all the lysines.

two N-linked glycans, Lys-203, and the region 265–293. A portion of the loop has been divided into three segments, and the amino acid sequence of cathepsin D and pepsinogen in this region is shown.

As shown in TABLE TWO, glycopepsinogen was barely phosphorylated (0.022% mannose phosphorylation). Addition of a single lysine residue at position 203 of glycopepsinogen (construct 1) resulted in a slight stimulation of phosphorylation (0.085%). Substitution of lysines at positions 267 and 293 (construct 2) resulted in 0.45% oligosaccharide phosphorylation, an increment of 20-fold over that of glycopepsinogen. A similar degree of phosphorylation occurred with construct 3 that contains lysines at positions 203 and 293 (0.47%). Construct 4 with lysines at positions 203 and 267 was phosphorylated 2.6%, a 116-fold increment over the baseline value. This is consistent with the previous finding that Lys-203 and Lys-267 are the most important elements of the recognition marker. However, construct 4 was only phosphorylated one-sixteenth as well as cathepsin D in this LoVo cell expression system (2.6% versus 42.3% phosphorylation). This indicates that other elements are necessary to obtain the optimal phosphotransferase recognition marker.

Construct 5, containing lysines at positions 203, 267, and 293, was phosphorylated 3.4%, a 156-fold stimulation over glycopepsinogen. The cathepsin D β-loop contains three additional lysines at its tip (positions 277, 281, and 284). Construct 6 was prepared to contain all six lysines (203, 267, 277, 281, 284, and 293). Its oligosaccharides were phosphorylated 5.4%, a 245-fold increment over the baseline and 12.8% of that obtained with cathepsin D. Constructs 7 and 8 contain lysines 203, 267, and 293 as well as β-loop segments AB and ABC, respectively. When expressed in LoVo cells, the resultant proteins were phosphorylated 7.7% and 9.7%. It is of note that construct 8 is phosphorylated almost twice as well as construct 6 even though both contain the six lysine residues at the identical positions. This indicates that residues other than lysines contribute to the potency of the recognition determinant.

The two N-linked glycans of glycopepsinogen are located on different lobes of this bilobed molecule: glycan 70 resides on the amino lobe, whereas glycan 199 is on the carboxyl lobe. Previously we reported that Lys-203 plus the β-loop elements preferentially stimulated phosphorylation of the carboxyl lobe oligosaccharide at position 199, although the amino lobe glycan was also phosphorylated (11). To determine if this was the case with the minimal recognition elements, constructs were prepared lacking either the amino- or carboxyl-lobe glycan. Constructs that lacked the amino lobe glycosylation signal at position 70 and contained Lys-203, Lys-267, and Lys-293 (construct 9) or these lysines plus β-loop segments ABC (construct 10) were phosphorylated 9.2% and 21.7%, respectively, a 418-fold and 986-fold increment over glycopepsinogen (Fig. 2). Furthermore, the glycan on construct 10 contained mostly high mannose species with two mannose 6-phosphate residues, indicative of a strong phosphotransferase recognition signal. The carboxyl lobe glycan of cathepsin D is known to acquire two mannose 6-phosphate residues, whereas the amino lobe glycan mostly has species with one mannose 6-phosphate residue (5). Constructs 11 and 12, which
have the same amino acid changes as constructs 5 and 8, respectively, but contain only the amino lobe glycan were undetectably phosphorylated. These findings show that virtually all the phosphorylation of these minimal constructs occurs on the carboxyl lobe glycan.

Effect of Replacing Lysine Residues in Cathepsin D on Phosphorylation—We previously reported that substitution of Lys-203 in cathepsin D with an alanine residue decreased phosphorylation by 15% (4). However, this determination used a Xenopus oocyte expression system and binding to a CI-MPR affinity column as a measure of mannose phosphorylation. This assay could underestimate the effect, because phosphorylation of only one of the two oligosaccharides is required for binding to the affinity column. Cuozzo et al. (7) found that a K203A substitution in cathepsin D resulted in a 42% decrease in phosphorylation using a COS cell expression assay. These authors also reported that a K293A substitution decreased phosphorylation by 13%, whereas K267A had no effect. The double mutant K203A/K293A produced a 69% decrease in phosphorylation.

We analyzed the effect of substituting Lys-203 and the five lysines associated with the β-loop on the phosphorylation of cathepsin D (TABLE THREE). These experiments were performed in HeLa and LoVo cells using a cathepsin D-myc construct that could be immunoprecipitated with anti-myc antibody, thereby avoiding contamination with endogenous cathepsin D. Despite a 10-fold difference in average expression, cathepsin D-myc oligosaccharides were phosphorylated the same in both cell lines (27.9% and 25.5% in HeLa and LoVo cells, respectively). The values for endogenous cathepsin D were higher (35% and 42.3% in HeLa and LoVo cells, respectively) than those obtained for the cathepsin D-myc construct. The lower levels of cathepsin D-myc phosphorylation may be attributed to the presence of the myc tag.

Substitution of Lys-203 with alanine decreased phosphorylation from 27.9% to 10% (64% inhibition) in the HeLa cell system. A similar decrease was observed in LoVo cells (61% inhibition). Replacement of Lys-267 or Lys-293 had lesser effects (20.9 and 19% phosphorylation, respectively). The K203A/K267A and K203A/K293A double mutants were phosphorylated much less well (6.9% and 8.5%, respectively) than the K267A/K293A double mutant (18.2%). The triple mutant (K203A/K267A/K293A or “3KA”) was phosphorylated 5.8% and 3.5% in HeLa and LoVo cells, respectively (79 and 86% inhibition relative to cathepsin D-myc). Constructs containing these three mutations plus mutation of either Lys-277, Lys-281, or Lys-284 to alanine were phosphorylated to the same extent as the K203A/K267A/K293A mutant construct. A similar result was obtained with constructs having lysines 203, 267, 277, 284, and 293 or lysines 203, 267, 281, 284, and 293 mutated to alanines (5.3% and 5.9% phosphorylation, respectively). However, when lysines 277 and 281 were mutated to alanines along with lysines 203, 267, and 293, the resultant protein was only phosphorylated 3.4% (88% inhibition). It is important to note that 3.4% phosphorylation is still 262-fold greater than the 0.013% phosphorylation of glycosipinogen obtained in this cell type, reinforcing the conclusion that other elements within cathepsin D contribute to its interaction with phosphotransferase. The construct with all six lysines mutated to alanines was secreted into the culture medium.

TABLE THREE

| Effect of cathepsin D mutations on oligosaccharide phosphorylation |
|---------------------------------------------------------------|
| All constructs were tested in HeLa cells, and the cathepsin D-myc, K203A, and “3KA” constructs were also tested in LoVo. Constructs were tested two to five times (average values ± S.D. is shown) except for “3KA” plus K281A, “3KA” plus K284A, and “3KA” plus K277A/K284A. Percent inhibition is relative to wild-type values for each cell line. |

| Construct/cell line | Average expression | Oligosaccharide phosphorylation | Ratio 2P/1P | Inhibition |
|---------------------|--------------------|--------------------------------|------------|-----------|
|                     | CPM                | %                              |            |           |
| CathD-myc (wtCD) HeLa | 12,800             | 27.9 ± 3.4                     | 0.50       |           |
| LoVo                | 1,450              | 25.5 ± 1.3                     | 0.93       |           |
| K203A HeLa          | 13,700             | 10 ± 1                         | 0.18       | 64        |
| LoVo                | 1,400              | 10 ± 1                         | 0.33       | 61        |
| K267A               | 6,500              | 20.9 ± 3.1                     | 0.63       | 25        |
| K293A               | 13,300             | 19 ± 1                         | 0.40       | 32        |
| K203A/K267A         | 6,500              | 6.9 ± 0.8                      | 0.13       | 75        |
| K203A/K293A         | 4,000              | 8.5 ± 1.7                      | 0.08       | 70        |
| K267A/K293A         | 1,900              | 18.2 ± 2.3                     | 0.42       | 35        |
| “3KA”: K203A/K267A/K293A HeLa | 12,400 | 5.8 ± 0.2                      | 0.06       | 79        |
| LoVo                | 1,550              | 3.5 ± 0.8                      | 0          | 86        |
| “3KA” + K277A       | 11,800             | 6.6 ± 0.2                      | 0.09       | 76        |
| “3KA” + K281A       | 9,500              | 6.1                             | 0.05       | 78        |
| “3KA” + K284A       | 36,600             | 5.7                             | 0.06       | 80        |
| “3KA” + K277A/K281A | 2,400              | 3.4 ± 0.8                      | 0.13       | 88        |
| “3KA” + K277A/K284A | 2,000              | 5.3                             | <0.05      | 81        |
| “3KA” + K281A/K284A | 1,800              | 5.9 ± 0.1                      | <0.05      | 79        |
Minimal Lysosomal Enzyme Recognition Elements

FIGURE 3. Autoactivation of cathepsin D constructs. Transfected HeLa cells were metabolically labeled with [35S]methionine/cysteine as described under “Materials and Methods.” Secreted cathepsin D molecules were immunoprecipitated, and the beads were resuspended in buffer at pH 3.5 in the absence or presence of pepstatin. Following resolution by SDS-PAGE, bands were visualized by autoradiography. The resulting mature species migrates faster in the gel. The autoactivation is blocked by pepstatin. Following resolution by SDS-PAGE, bands were visualized by autoradiography. This construct follows substitution of lysines at positions 203 and 267. When lysines were placed at positions 203 and 293 or at 267 and 293, the enhancement of phosphorylation was only one-sixth as great, but still 20-fold over that obtained with unaltered glycopepsinogen. The spacing between the α-carbon atoms of Lys-203 and Lys-267 is 40.1 Å, 34.1 Å between Lys-203 and Lys-293, and 12.2 Å between Lys-267 and Lys-293 (TABLE FOUR). Because all three combinations significantly stimulate mannose phosphorylation, it appears that phosphotransferase does not have an absolute requirement for the spacing between stimulatory lysines, at least in the case of glycopepsinogen. Furthermore, although two lysines may represent the minimal components of the recognition marker, it is clear that other elements contribute to the formation of the optimal binding site. Thus, the presence of lysines at positions 203, 267, and 293 is more stimulatory than any combination of two lysines and the addition of three more lysines at positions 203, 267, 277, 281, and 293 likely results from these amino lobe elements.

DISCUSSION

In this study, we sought to clarify the role of lysine residues as elements of the phosphotransferase recognition domain. Our data show that the introduction of just two lysines into glycopepsinogen can stimulate mannose phosphorylation up to 116-fold. However, the location of the lysines relative to each other and to the asparagine-linked high mannose glycan is critical for obtaining the maximal effect. The greatest stimulation occurred with the introduction of lysines at positions 203 and 267. When lysines were placed at positions 203 and 293 or at 267 and 293, the enhancement of phosphorylation was only one-sixth as great, but still 20-fold over that obtained with unaltered glycopepsinogen. The spacing between the α-carbon atoms of Lys-203 and Lys-267 is 40.1 Å, 34.1 Å between Lys-203 and Lys-293, and 12.2 Å between Lys-267 and Lys-293 (TABLE FOUR). Because all three combinations significantly stimulate mannose phosphorylation, it appears that phosphotransferase does not have an absolute requirement for the spacing between stimulatory lysines, at least in the case of glycopepsinogen. Furthermore, although two lysines may represent the minimal components of the recognition marker, it is clear that other elements contribute to the formation of the optimal binding site. Thus, the presence of lysines at positions 203, 267, and 293 is more stimulatory than any combination of two lysines and the addition of three more lysines at positions 203, 267, 277, 281, and 293 likely results from these amino lobe elements.

Sahagian and co-workers (7, 14) have extensively analyzed the role of lysine residues in the phosphorylation of cathepsin D and cathepsin L. In the experiments with cathepsin D, substitution of Lys-203 and Lys-293 with alanines decreased phosphorylation by 42 and 13%, respectively, whereas substitution of Lys-267 with alanine had no effect. When Lys-203 and Lys-293 were both replaced with alanines, phosphorylation decreased to 31% of wild-type cathepsin D (7). Substitution of the other lysines of cathepsin D, either alone or in combination with the Lys-203/Lys-293 modification had either no effect or a small inhibitory effect. They also reported that Lys-54 and Lys-99 of the propiece of cathepsin L were required for efficient phosphorylation of this cysteine protease (14). By examining the crystal structures of cathepsin D and procathepsin L, the Sahagian group observed that the distance between the critical lysine residues in both molecules was 34 ± 0.5 Å (7). They also noted that, in both instances, one of the critical lysines was much closer to the N-linked glycan than the other, although Lys-203 of cathepsin D is only 9.71 Å from Asn-199, whereas Lys-54 of procathepsin L is 25.08 Å from Asn-221. Based on these findings, these investigators proposed a model for the phosphorylation signal consisting of two lysine residues, exposed on the surface of the protein, which are spaced apart by 34 Å and positioned in a specific orientation relative to the target oligosaccharide (7). Because N-linked oligosaccharides are flexible, they would not have to be located at a precise distance from the critical lysines in order for their mannose residues to interact with phosphotransferase (15). They further suggested that the microenvironment of the critical lysines may influence the signals and noted that there appeared to be additional phosphorylation signals as well in view of the fact that removal of the two critical lysines from both procathepsin L and cathepsin D failed to completely abolish mannose phosphorylation.

Our findings support the proposal of Sahagian and colleagues that as few as two lysines in the correct orientation to each other and the car-

| TABLE FOUR | Distance between α-carbons atoms of key lysine and asparagine residues in various lysosomal enzymes |
|------------|-----------------------------------------------------------------------------------------------------|
| Lysosomal enzyme | Residue pair | Distance |
| Cathepsin D | Lys-203, Lys-267 | 40.1 |
| | Lys-203, Lys-293 | 34.1 |
| | Lys-267, Lys-293 | 12.2 |
| | Asn-199, Lys-203 | 9.71 |
| Cathepsin | LLys-54, Lys-99 | 34.1 |
| | Asn-221, Lys-54 | 25.1 |
| DNase I | Lys-50, Lys-74 | 27.1 |
| | Asn-106, Lys-74 | 13.6 |
| | Asn-18, Lys-27 | 14.3 |
| Aspartylglucosaminidase | Lys-183, Lys-214 | 35.0 |

The distances between α-carbons atoms were measured using Swiss-Pdb viewer.
bohydrate chain can serve as a minimal phosphotransferase recognition site. However, our data with the various combinations of Lys-203, Lys-267, and Lys-293 are not consistent with the requirement that the lysines be separated by 34 Å (see TABLE FOUR). The best pair (Lys-203/Lys-267) is separated by 40 Å and Lys-267/Lys-293, which are separated by only 12.2 Å, is as stimulatory as Lys-203/Lys-293, which are spaced 34 Å apart. Furthermore, as discussed above, the presence of additional lysines and other amino acids of the β-loop region enhanced phosphorylation 4-fold over that obtained with Lys-203/Lys-267 alone. This construct (construct 10) phosphorylated the oligosaccharide at Asn-199 half as well as occurs in cathepsin D. These data are in agreement with the previous studies indicating that lysine residues are critical elements of a more extensive recognition marker (4).

The results from studies of three other phosphotransferase substrates address this issue. The phosphorylation of the two Asn-linked oligosaccharides of bovine DNase I requires four lysine residues, Lys-27, Lys-50, Lys-74, and Lys-124. Importantly, Lys-27 selectively facilitates phosphorylation of the Asn-18 glycan, whereas Lys-74 only stimulates phosphorylation of the Asn-106 glycan, indicating that these two lysine residues are components of separate signals (8). Lys-50 also acts on the Asn-106 glycan and it is 27 Å from Lys-74, again illustrating that the spacing between lysines that stimulate phosphorylation of the same glycan is not fixed at 34 Å (16). The efficient phosphorylation of the glycans of aspartylglycosaminidase depends mostly on two lysines (Lys-183 and Lys-214) that are separated by 35 Å, with Lys-177 and Tyr-178 having a lesser role (9). Human arylsulfatase has eight lysine residues, only one of which (Lys-457) has been implicated in the interaction with phosphotransferase (10). The largest distance between Lys-457 and the other lysines is 25Å, establishing that there cannot be a fixed interlysine spacing of 34 Å in this lysosomal enzyme. Because mutation of Lys-457 to alanine only decreased mannose phosphorylation by 70%, there must be other elements in the recognition marker that contribute to efficient phosphorylation.

Taken together, the common feature of all these studies is that selected lysine residues of the target lysosomal enzymes play a critical role in the interaction with phosphotransferase. In all the cases, with the possible exception of arylsulfatase A, two or more lysines serve as critical elements of the binding site. However, the spacing between these lysines is not fixed, nor is the distance of the lysines to the Asn-linked glycan that is phosphorylated. In each instance, mutation of the critical lysines fails to completely abolish mannose phosphorylation, indicating that additional elements play a role in the interaction with phosphotransferase.

Although our study provides important clues into the nature of the recognition domain, it also raises an intriguing question. The finding that just two lysines can govern phosphotransferase binding might predict that some secreted glycoproteins with abundant lysine residues would be good substrates for phosphotransferase. In fact, a survey of three secreted glycoproteins with solved crystal structures (transferrin, ribonuclease B, and heparin cofactor II) reveals multiple surface lysine pairs whose spacing and orientation toward the oligosaccharides are nearly identical to that found in cathepsin D. Why are these secreted glycoproteins not acted upon by phosphotransferase? One can envision several mechanisms to account for this paradox. First, the secreted glycoproteins may contain elements that mask existing phosphotransferase recognition elements (for example, neighboring acidic residues capable of neutralizing key lysines). The existence of inhibitory residues in the vicinity of key lysines has been observed in murine DNase I (16). Second, secreted glycoproteins may have evolved “decoy” lysine pairs that allow binding to phosphotransferase in an orientation that positions their oligosaccharides such that they can’t be efficiently phosphorylated. Third, the secretory glycoproteins may lack non-lysine elements that are required for phosphorylation. In this regard, it should be noted that, because glycopepsinogen and cathepsin D are 45% identical in their amino acid sequence, it is possible that they share non-lysine residues that facilitate phosphorylation. Finally, it may be that many secreted glycoproteins do undergo low levels of mannose phosphorylation that diverts a small percentage of the molecules to lysosomes where they are degraded. Special measures would have to be taken to detect this process. Experiments designed to test some of these possibilities are in progress and should shed further light of the elusive nature of this protein-protein interaction.

Acknowledgments—We thank Walter Gregory for production of the anti-pepsinogen antiserum, Dr. Thomas Baranski for critical reading of the manuscript, and Dr. Jay Ponder for the determination of distances between the α-carbon atoms of various residues.

REFERENCES
1. Kornfeld, S. (1986) J. Clin. Invest. 77, 1–6
2. Baranski, T. J., Faust, P. L., and Kornfeld, S. (1990) Cell 63, 281–291
3. Baranski, T. J., Koelsch, G., Hartsuck, J. A., and Kornfeld, S. (1991) J. Biol. Chem. 266, 23365–23372
4. Baranski, T. J., Cantor, A. B., and Kornfeld, S. (1992) J. Biol. Chem. 267, 23342–23348
5. Cantor, A. B., Baranski, T. J., and Kornfeld, S. (1992) J. Biol. Chem. 267, 23349–23356
6. Dustin, M. L., Baranski, T. J., Sampath, D., and Kornfeld, S. (1995) J. Biol. Chem. 270, 170–179
7. Cuzzo, J. W., Tao, K., Cygler, M., Mort, J. S., and Sahagian, G. G. (1998) J. Biol. Chem. 273, 21067–21076
8. Nishikawa, A., Gregory, W., Frenz, J., Cacia, J., and Kornfeld, S. (1997) J. Biol. Chem. 272, 19408–19412
9. Tikkanen, R., Peltola, M., Oinonen, C., Rouvinen, J., and Peltonen, L. (1997) EMBO J. 16, 6684–6693
10. Yaghoobfar, A., Schestag, F., Dierks, T., and Gieselmann, V. (2003) J. Biol. Chem. 278, 32653–32661
11. Cantor, A. B., and Kornfeld, S. (1992) J. Biol. Chem. 267, 23357–23363
12. Do, H., Lee, W. S., Ghosh, P., Hollowell, T., Canfield, W., and Kornfeld, S. (2002) J. Biol. Chem. 277, 29737–29744
13. Lee, B., and Richards, F. M. (1971) J. Mol. Biol. 55, 379–400
14. Cuzzo, J. W., Tao, K., Wu, Q. L., Young, W., and Sahagian, G. G. (1995) J. Biol. Chem. 270, 15611–15619
15. Warner, J. B., Thalhauser, C., Tao, K., and Sahagian, G. G. (2002) J. Biol. Chem. 277, 41897–41905
16. Nishikawa, A., Nanda, A., Gregory, W., Frenz, J., and Kornfeld, S. (1999) J. Biol. Chem. 274, 19309–19315
Identification of the Minimal Lysosomal Enzyme Recognition Domain in Cathepsin D
Richard Steet, Wang-Sik Lee and Stuart Kornfeld

J. Biol. Chem. 2005, 280:33318-33323.
doi: 10.1074/jbc.M505994200 originally published online August 4, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M505994200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 16 references, 13 of which can be accessed free at http://www.jbc.org/content/280/39/33318.full.html#ref-list-1