Crystal Structure and Autoactivation Pathway of the Precursor Form of Human Tripeptidyl-peptidase 1, the Enzyme Deficient in Late Infantile Cereoid Lipofuscinosis

Received for publication, September 8, 2008, and in revised form, November 7, 2008 Published, JBC Papers in Press, November 26, 2008, DOI 10.1074/jbc.M806943200

Jayita Guhaniyogi1,2, Istvan Sohar1, Kalyan Das1, U.N. Michael Stock1,2,3, and Peter Lobel1,2

From the 1Center for Advanced Biotechnology and Medicine, 2Department of Biochemistry, and 3Department of Pharmacology, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, and the 4Department of Chemical Biology, Rutgers University, and the 5Howard Hughes Medical Institute, Piscataway, New Jersey 08854

Late infantile neuronal ceroid lipofuscinosis (LINCL)3 is a neurodegenerative lysosomal storage disease of childhood that presents typically between the ages of 2 and 4 years with the onset of seizures. Disease progression is reflected by blindness, dementia, mental retardation, and an increase in the severity of seizures. LINCL is always fatal, and the life span of patients is typically 6–15 years. LINCL is caused by mutations in TTP1 (previously named CLN2, for ceroid lipofuscinosis neuronal type 2 gene) (1), which normally encodes a lysosomal protease, tripeptidyl-peptidase 1 (TPP1, EC 3.4.14.9) (2, 3).

There is currently no treatment of demonstrated efficacy for LINCL, but promising progress is being made in some directions. Proof-of-principle for virus-mediated gene therapy has been established in a mouse model of LINCL, with a significant improvement in disease phenotype achieved with the use of adenovirus vectors expressing TPP1 (4–7). Afflicted children have also been treated with adenovirus vectors, although it is too soon to determine whether significant clinical benefits have been achieved in these early trials (8). Enzyme replacement therapy, an approach that has proven successful in a number of other lysosomal storage diseases, has also been investigated in LINCL. Purified recombinant human TPP1 that contains the mannose 6-phosphate lysosomal targeting modification can be taken up by LINCL fibroblasts where it degrades storage material (9), and the protein has been introduced into the cerebrospinal fluid of the LINCL mouse model via intraventricular injection, resulting in significant uptake into the brain and some correction of neuropathology (10).

For therapeutic approaches that rely upon replacing a mutant gene product with a functional protein via recombinant methods, e.g. gene and enzyme replacement therapy, a thorough understanding of the biological and biophysical properties of the protein in question are essential for success. Thus, for LINCL, considerable effort has been directed toward the investigation of TPP1, and as a result, this is a well characterized

Late infantile neuronal ceroid lipofuscinosis neuronal type 2 gene) (1), which normally encodes a lysosomal protease, tripeptidyl-peptidase 1 (TPP1, EC 3.4.14.9) (2, 3).

The on-line version of this article (available at http://www.jbc.org/) contains Figs. S1–S8, Tables S1–S4, and additional references.

1 Investigator of the Howard Hughes Medical Institute. To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-4844; Fax: 732-235-5289; E-mail: stock@cabm.rutgers.edu.

2 To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-5032; Fax: 732-235-5289; E-mail: lobel@cabm.rutgers.edu.

3 The abbreviations used are: LINCL, late infantile ceroid lipofuscinois; TPP1, tripeptidyl-peptidase 1; CH0, Chinese hamster ovary; MOPS, 3-(N-morpholino)propanesulfonic acid; PDB, Protein Data Bank; TLS, translation/libra-

Homologous protein search; GlcNAc, N-acetyl-β-glucosamine; r.m.s.d., root mean square deviation; PB, Poisson-Boltzmann; APBS, adaptive PB solver; MS, mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; BisTris, 2-(bis[2-hydroxyethyl]amino)-2-(hydroxymethyl)-propane-1,3-diol; Endo H, endoglycosidase H; SAD, single wavelength anomalous dispersion.

The atomic coordinates and structure factors (code 3EDY) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The on-line version of this article (available at http://www.jbc.org/) contains Figs. S1–S8, Tables S1–S4, and additional references.

1 Investigator of the Howard Hughes Medical Institute. To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-4844; Fax: 732-235-5289; E-mail: stock@cabm.rutgers.edu.

2 To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-5032; Fax: 732-235-5289; E-mail: lobel@cabm.rutgers.edu.

3 The abbreviations used are: LINCL, late infantile ceroid lipofuscinosis; TPP1, tripeptidyl-peptidase 1; CH0, Chinese hamster ovary; MOPS, 3-(N-morpholino)propanesulfonic acid; PDB, Protein Data Bank; TLS, translation/libra-

Homologous protein search; GlcNAc, N-acetyl-β-glucosamine; r.m.s.d., root mean square deviation; PB, Poisson-Boltzmann; APBS, adaptive PB solver; MS, mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; BisTris, 2-(bis[2-hydroxyethyl]amino)-2-(hydroxymethyl)-propane-1,3-diol; Endo H, endoglycosidase H; SAD, single wavelength anomalous dispersion.

The on-line version of this article (available at http://www.jbc.org/) contains Figs. S1–S8, Tables S1–S4, and additional references.

1 Investigator of the Howard Hughes Medical Institute. To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-4844; Fax: 732-235-5289; E-mail: stock@cabm.rutgers.edu.

2 To whom correspondence may be addressed: 679 Hoes Lane, Piscataway, NJ 08854. Tel.: 732-235-5032; Fax: 732-235-5289; E-mail: lobel@cabm.rutgers.edu.

3 The abbreviations used are: LINCL, late infantile ceroid lipofuscinosis; TPP1, tripeptidyl-peptidase 1; CH0, Chinese hamster ovary; MOPS, 3-(N-morpholino)propanesulfonic acid; PDB, Protein Data Bank; TLS, translation/libra-

Homologous protein search; GlcNAc, N-acetyl-β-glucosamine; r.m.s.d., root mean square deviation; PB, Poisson-Boltzmann; APBS, adaptive PB solver; MS, mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; BisTris, 2-(bis[2-hydroxyethyl]amino)-2-(hydroxymethyl)-propane-1,3-diol; Endo H, endoglycosidase H; SAD, single wavelength anomalous dispersion.
enzyme at the functional and molecular levels (reviewed in Refs. 11, 12). **TPP1** encodes a 563-residue preproprotein with a cleavable N-terminal 19-residue signal sequence. The proenzyme (residues 20–563) is a soluble monomer that undergoes proteolytic cleavage in the lysosome, converting the zymogen to an active, mature protease (residues 196–563) (1). Studies on purified pro-TPP1 demonstrate that maturation is autocatalytic in vitro (13, 14) but may involve other proteases in vivo (15). TPP1 is glycosylated, and its N-linked oligosaccharides have been implicated in maturation, activity, targeting, and stability of the processed enzyme (16, 17).

TPP1 is a serine protease (14) that possesses two catalytic functions as follows: a primary tripeptidyl exopeptidase activity with a pH optimum of ~5.0 that catalyzes the sequential release of tripeptides from the unsubstituted N termini of substrates (15). TPP1 is glycosylated, and its N-linked oligosaccharides have been determined for three bacterial members of this family (22). High resolution crystal structures of both free and inhibitor-bound complexes have been determined for three bacterial members of this family (sedolisin (23–26), kumamolisin (27, 28), and kumamolisin-As (29, 30)), and for one (kumamolisin), the structure of a mutant, inactive precursor form has also been obtained (28). These proteins share a common subtilisin-like fold, an octahedrally coordinated calcium-binding site, and an active site that contains an unusual Ser-Glu-Asp (SED) catalytic triad, rather than the Ser-His-Asp (SHD) triad of subtilisin (31, 32). Chemical modification studies of TPP1 have revealed that Ser479 is the active site nucleophile (14). Modeling studies suggest that Glu272 and Asp276 complete the catalytic triad and that Asp360 is homologous to the conserved Asn in the subtilisin family in its role in stabilization of the oxyanion of the tetrahedral intermediate during catalysis (33). Site-directed mutagenesis studies are consistent with these conclusions (14, 34).

A detailed understanding of the tertiary structure of TPP1 may have implications for developing or improving therapeutic strategies. First, a high resolution model would provide the basis for targeted protein engineering efforts to design TPP1 derivatives with increased half-life prior to and/or upon delivery to the lysosome. Successful creation of a longer lived TPP1 molecule could significantly enhance gene or enzyme replacement approaches to LINCL. Second, a structural model for TPP1 could be valuable in designing derivatives tagged with protein transduction domains to facilitate crossing of the blood-brain barrier for delivery to the central nervous system from the bloodstream. In this study, we present the crystal structure of the proform of human TPP1 at 1.85 Å resolution. This model provides novel insights into the structural basis for the pH-induced auto-activation of the proform of TPP1. A structure of glycosylated pro-TPP1 has been independently determined, displaying features similar to those of deglycosylated TPP1.

**EXPERIMENTAL PROCEDURES**

**Protein Production**—Production and purification of recombinant human TPP1 from a Chinese hamster ovary (CHO) expression system were conducted using modifications of methods described previously (9). Cells were cultured in 1050-cm² roller bottles for 9 days with serum-free Dulbecco’s modified Eagle’s/F-12 medium containing 1 μg/ml kifunensin. Conditioned media were harvested and adjusted to contain 2.5 M NaCl as well as 20 mM Tris, pH 8.0, and 20 μM CaCl₂ (TC Buffer). All subsequent steps were conducted at 4 °C unless noted otherwise. The solution was clarified by centrifugation and filtration and then applied to a butyl-Sepharose 4 fast-flow column (GE Healthcare). After washing with 2.5 M NaCl in TC Buffer, the column was eluted with a 7.5-column volume linear gradient of 2.5 to 0 M NaCl in TC Buffer. Fractions containing most of the pro-TPP1 were concentrated and equilibrated with 20 mM NaCl/TC Buffer by diafiltration. The solution was filtered, loaded onto a Mono Q12 column (Bio-Rad), and washed with 4 column volumes of 20 mM NaCl in TC Buffer. The column was eluted with 10 column volumes of a linear gradient of 20 to 515 mM NaCl in TC Buffer. Peak fractions containing pro-TPP1 were adjusted to contain 100 mM sodium citrate, pH 5.5. Protein was deglycosylated by incubation with Endo Hf (New England Biolabs, 26,500 units per mg of pro-TPP1) at 25 °C for 15 h followed by 1 h at 37 °C. The solution was buffer exchanged and repurified by Mono Q12 chromatography as described above. Fractions containing pro-TPP1 were concentrated and applied to a Superdex 75 column equilibrated in 150 mM NaCl in TC Buffer to separate pro-TPP1 monomer from a minor amount of aggregated material.

A similar procedure was used to isolate glycosylated pro-TPP1, omitting kifunensin from the cell culture media and the Endo H digestion and second Mono Q12 chromatography steps. Hexahistidine-tagged human pro-TPP1 was produced using an insect cell system and purified as described previously (14) except that all column buffers contained 20 μM CaCl₂, and a final Superdex 75 gel filtration step was performed as described above.

**TPP1 Activation Kinetics**—TPP1 proenzyme was diluted to the desired concentration using 0.1% Triton X-100, 150 mM NaCl, 20 mM Tris, pH 8.0. For experiments using hexahistidine TPP1, the protein was preincubated with 1 mM E-64 to inhibit an insect cell thiol protease contaminant that was present in trace amounts in some preparations. Proenzyme (20 μl) was mixed with 180 μl of 0.1% Triton X-100, 150 mM NaCl, and 100 mM sodium formate adjusted with HCl to the indicated pH. The reaction was incubated at 25 °C for the indicated time, and the activation process was terminated by diluting 1 part sample into 9 parts ice-cold 0.1% Triton X-100, 150 mM NaCl, 500 mM sodium acetate, pH 5.0, frozen on dry ice, and stored at −80 °C. The amount of enzymatically active TPP1 (total final TPP1 concentration typically 1.5 nM) was measured using the previously described kinetic fluorescence assay (35) using 200 μM Ala-Ala-Phe-AMC substrate in 0.1% Triton X-100, 150 mM NaCl, 500 mM sodium acetate, pH 5.0. Data for activation conducted at pH 4.0 were fit to a one-phase exponential association model using built in functions in Prism version 5.01 (GraphPad
software). The activation reactions at pH 2.5 to 3.5 were fit to an irreversible consecutive reaction model where the inactive TPP1 zymogen (Z) is converted to an intermediate (I) with partial activity, which is then converted into a fully active mature form (M) as shown in Reaction 1,

\[ Z \rightarrow I \rightarrow M \]

**K\textsubscript{1} \quad K\textsubscript{2}**

**REACTION 1**

Using this model, one can calculate the activity at any given time from Equation 1,

\[ \nu_t = v_{\text{max}} (\alpha_1 [Z]_t + \alpha_2 [I]_t + \alpha_3 [M]_t) \]

(Eq. 1)

Here \( \nu_t \) denotes the observed TPP1 enzymatic activity at any given time; \( v_{\text{max}} \) denotes the enzymatic activity of the preparation when existing as the fully activated form; \( \alpha \) denotes the relative enzymatic activity of the different forms, and for any given time \([Z]_t + [I]_t + [M]_t\) is a constant that corresponds to the total added TPP1, with all existing as Z at time \( t = 0 \). Data were fit using the integrated rate equations given in Ref. 36 using Prism version 5.01, with the following constraints: no activity for the zymogen (\( \alpha_1 = 0 \)), partial activity for the intermediate form (0 \( \leq \alpha_2 \leq 1 \)), and maximal activity for the mature form (\( \alpha_3 = 1 \)).

**Protein Chemistry**—Pro-TPP1 was activated as described above. For matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), reactions were terminated by mixing with an equal volume of ice-cold solution containing 0.5% trifluoroacetic acid, 50% acetonitrile, mixed with sinapinic acid matrix, and analyzed using an Applied Biosystems/MDS-Sciex 4800 mass spectrometer operated in the linear mode. For SDS-PAGE analysis, reactions were terminated using reducing sample buffer and fractionated on 10% NuPAGE BisTris gels using MOPS running buffer (Invitrogen). Gels were either stained using Sypro Ruby (Molecular Probes) or transferred to membranes, stained with Coomassie Blue, and relevant regions of the membrane excised and sequenced as described previously (9, 14).

**Crystallization of Pro-TPP1**—Purified Endo H-deglycosylated pro-TPP1 was dialyzed into TC buffer containing 100 mM NaCl, concentrated to 7 mg/ml using a 30,000 centrifugal device (Millipore), and filtered through 0.22-μm pore size cellulose acetate filters (Corning Costar). Initial crystallization conditions were identified by high throughput screening at the Hauptman Woodward Medical Research Institute, Buffalo, NY. In the optimized condition, crystals of pro-TPP1 ranging in size between 50 × 100 × 100 μm and 50 × 100 × 300 μm appeared in ~3 days when set up and grown at 4 °C using hanging drop diffusion by mixing 1 μl of pro-TPP1 solution at 5 mg/ml and 1 μl of reservoir solution containing 0.1 M citrate, 4−7% polyethylene glycol 6000, pH 5.0.

**Data Collection, Structure Determination, and Refinement**—

Data for two native, one mercury-derivatized, and one lead-derivatized crystals were collected to 1.85 Å at X4C at the National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY. The lead and mercury derivatives were obtained by soaking the crystals overnight at 4 °C in reservoir solutions containing heavy atom salts at concentrations of 2.5 mM Pb(NO\textsubscript{3})\textsubscript{2} and 0.5 mM C\textsubscript{2}H\textsubscript{5}HgCl. Cryoprotection was achieved with passage of the crystals held in cryoloops through 1-μl drops of 50% ethylene glycol for 1 s and flash-cooling in liquid N\textsubscript{2} stream. For all crystals, 180° of data with 1° oscillation angle were collected. All four data sets were processed and scaled in the primitive orthorhombic P2\textsubscript{2}121 space group. Furthermore, attempts for solving the structure by multiple isomorphous replacement using multiple heavy atom derivatives were unsuccessful because of lack of isomorphism between the datasets (supplemental Table S1). Subsequently, the crystal structure was solved by single wavelength anomalous dispersion (SAD) technique using the mercury dataset. Four mercury sites were obtained from the anomalous differences using SHELX programs (38) with the HKL2MAP software (39). The sites were used for SAD phasing at 2.5 Å resolution.

**TABLE 1**

| PDB ID | Data collection |
|--------|----------------|
| JEDY   |                 |

**Data collection and refinement statistics**

Values corresponding to highest resolution shells are shown in parentheses.

| Resolution (Å)   | 50.00–1.83 (1.90–1.83) |
|------------------|-------------------------|
| \( R_{\text{sym}} \) | 0.079 (0.457)           |
| \( R_{\text{free}} \) | 0.088 (0.515)           |
| Completeness (%) | 99.1 (93.3)             |
| Redundancy       | 7.6 (8.8)               |

**Phasing**

Figure of merit after SAD phasing at 2.50 Å (figure of merit after phase extension to 1.85 Å)

Pearson correlation coefficient, \( R_{\text{p.i.m.}} \), and identity, \( I_{\text{ot}} \), for the best model, and the overall B-value for the final model, 21.88 Å\(^2\)
Structural Analyses—Superposition of structural models was accomplished using LSQKAB (54) from the CCP4 program suite version 6.0 (44). Structural comparisons were made using the Secondary Structure Matching server (55), and overall root mean square deviation (r.m.s.d.) values were obtained using the r.m.s.d. calculation algorithm within CNS version 1.2 (53). The protein–protein interaction server (56, 57), the iMOLTALK server (58), and the WHATIF web interface (59) were used for analyses of interdomain and lattice contacts. Buried surface areas were calculated using CNS version 1.2 (53), and all structural figures were generated using PyMOL version 1.0 (60).

Electrostatic Potential Calculations—Analysis of pH dependence on interdomain binding energy in pro-TPP1 was accomplished using the Adaptive Poisson–Boltzmann Solver (APBS) version 0.5.1 (61). The PDB2PQR server (62, 63) implementing the PARSE forcefield and the PROPKA program (64) to assign specific protonation states at different pH conditions (pH 2.5–8.0) was used to convert the pro-TPP1 protein data bank (PDB) coordinates to three sets of PQR files, one for the intactzymogen (residues 20–563) ($Z_{\text{pH}}$), one for the propiece (residues 20–195) ($P_{\text{pH}}$), and the third for the mature domain (residues 196–563) ($M_{\text{pH}}$) for a given pH condition. These pH-specific PQR files were used as input for electrostatic binding energy calculations at that pH using a linearized Poisson–Boltzmann equation (PB) using cubic B-spline discretization of the charge distributions. PB calculations were performed at 298 K with a dielectric constant of 78.0 for water and 2.0 for the protein interior. The ion concentrations were set to 0.15 M with ionic radius of 2.0 Å. Care was taken to remove self-interaction energies by using exactly the same grid spacing, length and center, determined by the psize.py script provided in the APBS distribution. The electrostatic binding energy at a given pH ($BE_{\text{pH}}$) is calculated by APBS as shown in Equation 2,

$$BE_{\text{pH}} = Z_{\text{pH}} - (P_{\text{pH}} + M_{\text{pH}})$$

(Eq. 2)

RESULTS AND DISCUSSION

Production of Deglycosylated TPP1—Initial attempts to obtain suitable crystals from previously characterized glycosylated recombinant human pro-TPP1 produced in insect cells (14) and in CHO cells (9) were unsuccessful (data not shown). Given that heterogeneity in glycosylation may have been a complicating factor, we tried different methods to deglycosylate the protein produced from CHO cells. In our hands, peptide-N-glycosidase F was not completely effective in removing oligosaccharides unless the protein was denatured, suggesting that steric hindrance prevented cleavage of the GlcNAc–Asn bond. Endo H, which cleaves between the first two Asn-linked GlcNAc residues, was partially effective in deglycosylating the native molecule, but some N-linked glycans were retained, reflecting the presence of oligosaccharides that had been processed to Endo H–resistant forms (15, 65). We therefore cultured the CHO cells in the presence of the endoplasmic reticulum mannosidase I inhibitor kifunensin, which prevents further processing of N-linked glycans to Endo H–resistant forms (66). When treated with Endo H, pro-TPP1 produced in this manner yielded a single species by SDS-PAGE analysis, corresponding to a deglycosylated form in which single single Asn-linked GlcNAc monosaccharides remain, as confirmed by the crystal structure (supplemental Fig. S1 and see below).

Properties of Deglycosylated TPP1—Glycosylation has been reported to be critical for the folding, processing, and/or stability of TPPI in cultured cells (17, 67). In this study, it was therefore important to compare the functional properties of deglycosylated TPP1 prepared for structural studies and the previously well characterized glycosylated native protein. Purified pro-TPP1 is a zymogen that undergoes low pH–triggered autoactivation and processing (9, 13, 14, 16, 68). The kinetics for autoactivation of deglycosylated pro-TPP1 were essentially identical to those for glycosylated protein produced from untreated CHO cells or a hexahistidine-tagged protein produced from insect cells (Fig. 1). Rates were greatly accelerated at lower pH values and showed biphasic kinetics at pH 3.5 and below (Fig. 1).

We also measured the activity of preactivated glycosylated and deglycosylated TPP1 using two different fluorogenic substrates, Ala–Ala–Phe–AMC and Arg–Nle–Nle–AMC. Initial rates were determined as a function of substrate concentration and pH. The $k_{\text{cat}}$ values ($V_{\text{max}}$/protein concentration) for the deglycosylated and glycosylated proteins were similar (Fig. 2, A and B). Although it appeared that the deglycosylated protein had

---

5 J. Clardy, personal communication.
slightly higher activity, the difference fell within the variation
that we typically observe with independent preparations of gly-
cosylated TPP1. The Michaelis constant ($K_m$) (Fig. 2, A and B)
and substrate specificity constant ($k_{cat}/K_m$) (Fig. 2C) of the
glycosylated and deglycosylated protein were essentially indistin-
guishable. Our results do not rule out a role for glycosylation in
folding or trafficking of pro-TPP1, nor do they address the role
of the terminal Asn-linked GlcNAc that remains after Endo
H-deglycosylation. However, they do demonstrate that Endo
H-deglycosylated TPP1 is equivalent in terms of autocalytic
and enzymatic function to fully glycosylated TPP1.

Crystal Structure of Pro-TPP1 Reveals a Conserved Subtilisin-
like Fold—We have determined the crystal structure of the
inactive precursor of Endo H-deglycosylated form of TPP1 at
1.85 Å resolution (supplemental Fig. S1 and Fig. 3). The data
processed in the P212121 space group with one molecule in the
asymmetric unit with a Matthews coefficient and solvent con-
tent of 2.37 Å³/Da and 48.0%, respectively. The structure is
markedly similar to the glycosylated form of pro-TPP1 deter-
mined at 2.35 Å resolution in the P21212 space group with two
molecules in the asymmetric unit (PDB ID 3EE6). Superposition
of C$\beta$ coordinates of deglycosylated pro-TPP1 with those of the
two chains of glycosylated pro-TPP1 yield r.m.s.d values of 0.61
and 0.66 Å, respectively, with differences in positioning of a few
solvent exposed loops and conformations of a couple of active
site residues (as discussed below). Pro-TPP1 exhibits a charac-
teristic bi-domain precursor structure with a C-shaped prodo-
main (residues Ser20–Phe169) tightly packed against a more
globular catalytic domain (residues Leu196–Pro563) connected
by a 24-residue linker (Gly$^{172}$–Gly$^{195}$). The propiece has an
antiparallel-$\alpha$/antiparallel-$\beta$ fold and the catalytic domain has a
typical subtilisin-like fold.

FIGURE 1. Autoactivation of pro-TPP1. Highly purified human recombinant
TPP1 preparations (filled circles, hexahistidine-tagged pro-TPP1 from an
insect cell expression system; open circles, fully glycosylated pro-TPP1 from
CHO cells; and open squares, Endo H-deglycosylated pro-TPP1 from kifunen-
sin-treated CHO cells) were incubated at the indicated pH for the specified
time before stopping the reaction and measuring TPP1 activity as described
under “Experimental Procedures.” Data for each experiment were normalized
to plateau values, all of which were comparable. The curves shown are fit to
the reaction curves for the hexahistidine-tagged pro-TPP1 preparation as
described under “Experimental Procedures.”

FIGURE 2. Enzymatic properties of glycosylated and Endo H-deglycosyl-
ated TPP1. CHO cell-produced enzyme preparations described in Fig. 1 were
incubated at pH 3.5 for 2 h at 37 °C. Enzyme activity measurements were
conducted using eight different concentrations of either Ala-Ala-Phe-AMC
(800 to 6.25 μM) or Arg-Nle-Nle-AMC (400 to 3.12 μM) substrate. Final reac-
tions contained 150 mM NaCl, 0.1% Triton X-100, and either 100 mM sodium
formate, pH 2.5, or sodium acetate (all other pH values) adjusted to the indi-
cated pH using HCl. Kinetic constants were calculated using the built-in
Michaelis-Menten curve fitting functions in Prism version 5.01. Open symbols
denote deglycosylated protein; filled symbols denote glycosylated protein. A
and B, circles, $k_{cat}$; squares, $K_m$. C, circles, Ala-Ala-Phe-AMC substrate; squares,
Arg-Nle-Nle-AMC substrate.
Structure of Pro-TPP1

The N-terminal portion (residues 172–179) of the propiece linker is well ordered and threads through the substrate-binding groove making strong short distance contacts (one hydrogen bond, one salt bridge, and three hydrophobic interactions) with the catalytic domain. In contrast, the remaining portion of the linker (residues 180–195) loops out of the interface making only minimal long range van der Waal contacts. The conformation of this portion of the linker is stabilized by a few strong lattice contacts involving residues 180–190 as discussed below. Interestingly, the well structured nature of the N-terminal portion of the linker is supported by identical conformations and its involvement in interdomain contacts in both glycosylated and deglycosylated structures. The flexible nature of the C-terminal portion of the linker is emphasized by its involvement in lattice contacts in the deglycosylated structure and by its absence in the glycosylated structure, possibly because of disorder. Furthermore, a TLS motion analysis (48–50) for evidence of thermal motion detects significant flexibility in the linker region as compared with the rest of the protein (supplemental Fig. S2). Such flexibility may be important in positioning the linker for autocatalytic cleavage during zymogen activation (see below).

The structure has marked similarities with those of other S53 family members (24, 27, 28, 30) in terms of the overall fold, barring minor differences in the presence or absence of a few secondary structural elements in both the propiece and mature domain and orientation of the linker (supplemental Figs. S3 and S4). Superpositions of Cα coordinates of corresponding regions of pro-TPP1 with propiece of prokumamolisin (PDB ID 1T1E) (28), and mature enzymes of sedolisin (PDB ID 1GA4) (24), kumamolisin (PDB ID 1GTG) (27), and kumamolisin-As (PDB ID 1SN7) (30) yield r.m.s.d. values of 2.17, 2.07, 1.81, and 1.92 Å, respectively. Compared with prokumamolisin, the propiece and mature domain in pro-TPP1 are more compact with 550 Å² more buried surface area than the former, whereas the respective linkers are oriented at an angle of ~30° (supplemental Fig. S4). The total buried surface area between the propiece and mature domain in pro-TPP1 is 2200 Å² with 22 direct and 10 water-mediated hydrogen bonds, 3 salt bridges, and 6 hydrophobic interactions (supplemental Fig. S5 and supplemental Table S2).

Key features of the S53 family members, including Cα positions of catalytic residues forming the charge transfer complex (viz. Ser475, Glu272, Asp276, and Asp160), two cis-peptide bonds, and the calcium-binding site are conserved. Three cis-Pro residues are found in pro-TPP1 (Pro79, Pro450, and Pro561), of which the former two are completely conserved in the bacterial homologs (supplemental Fig. S3). The structural role and binding partners of the octahedrally coordinated calcium ion, including the side chain carboxyl groups of Asp517 and Asp451, backbone carbonyl groups of Val15, Gly339, and Gly541, and a bound water molecule, are conserved within the S53 family. In addition, three disulfide linkages unique to pro-TPP1 (viz. Cys111–Cys122, Cys365–Cys526, and Cys522–Cys537) stabilize the fold. The lone Cys527 is completely buried within the core of the mature domain and is within hydrogen bonding distance of the side chain carboxyl of Asp276 (see below).

Endo H treatment leaves one GlcNAc moiety at each N-linked glycosylation site. Mutagenesis data indicate that all five potential glycosylation sites of TPP1 (viz. Asn210, Asn222, Asn286, Asn313, and Asn443) are utilized (67). We find clear evidence for single GlcNAc moieties at all sites except for Asn222, a surface residue, which is clearly not modified in our structure. The structure of the glycosylated form of pro-TPP1 also shows that Asn222 is not utilized (PDB ID 3EE6). It is formally possible that this site is heterogeneously glycosylated and that crystallization conditions have selected out the four-sugar glycosylated form from a population of four- and five-sugar forms. In our structure, the GlcNAc moiety linked to Asn286 participates in lattice contacts with a symmetry-related molecule (supplemental Fig. S6 and supplemental Table S3).

Active Site of Pro-TPP1 and Maintenance of the Zymogen in the Inactive State—The active site in pro-TPP1 contains an SED (Ser475–Glu272–Asp276) catalytic triad, similar to all other members of the S53 family of serine peptidases and contrasting with the SHD catalytic triad of canonical serine peptidases. The replacement of the His residue with the acidic residue Glu in the S53 family allows the active site serine to function in nucleophilic attack of substrate at acidic pH (14), with the Glu abstracting the hydroxyl proton from the serine (14, 23–26, 34). In addition, there is a second conserved Asp residue (Asp160), designated here as D’, in place of a conserved Asn in the subtilisin family. The role of the latter is to bind and stabilize the
The crystal structure of pro-TPP1 reveals a number of features that help maintain thezymogen in an inactive state, with the active site assuming a geometry that is suboptimal for catalysis. Low pH-mediated pro-TPP1zymogen maturation is likely to be associated with different positioning of key elements at the active site as well as different protonation states of active site carboxylates that might favor conformations poised for catalysis.

Although the C’ positions of key catalytic residues, including SEDD’ are superimposable with those of corresponding residues in all previously determined structures of bacterial homologs (24, 27, 28, 30), the side chain conformations and hydrogen bond network at the active site in pro-TPP1 are different (Fig. 4). As revealed by several high resolution crystal structures of free and inhibitor-bound mature enzymes of sedolisin, kumamolisin, and kumamolisin-As, the geometry of an active site poised for catalysis is characterized by a charge transfer complex involving the key SED residues with hydrogen bonds between one of the side chain carboxyl oxygens of the Glu with the hydroxyl of the Ser and the other with one of the carboxyl oxygens of the Asp. In contrast, in pro-TPP1, side chains of both Glu272 and Asp276 are directed away from each other and from the active site Ser475. Such alternate rotameric conformations of these catalytic residues allow hydrogen bonds distinct from those characteristic of a charge transfer complex (Fig. 4). PROPKA predictions (64) yield $pK_a$ values of 3.6 and 6.4 for Glu272 and Asp276 in the mature enzyme, respectively (supplemental Table S4). From this it follows that such rotameric conformations for the Glu and the Asp, as observed in pro-TPP1, might arise from electrostatic charge repulsions and could likely be favored at neutral pH, at which both of these residues are predicted to be deprotonated and negatively charged. However, at the crystallization pH condition, pH 5.0, only Glu272 and not Asp276 is predicted to be deprotonated. The active site Glu is also found in dual conformations in some inhibitor complexes in structures of homologs crystallized at higher pH (25). In addition, the conformations of Glu272 and Asp276 in the glycosylated form of pro-TPP1 are similar to those found in mature enzymes of bacterial
Structure of Pro-TPP1

homologs. These results suggest that the acidic residues at the active site most likely exist in equilibrium between alternate rotameric states and that different conditions favoring different pH-driven protonation states increase the probability of one state over another.

Furthermore, in both structures of the glycosylated and Endo H-deglycosylated forms of pro-TPP1, the side chain hydroxyl group of the active site Ser$^{375}$ is also involved in a strong hydrogen bond with the side chain hydroxyl of Ser$^{324}$ at 2.6 Å. Interestingly, this Ser (designated here as $S'$) is completely conserved within the S53 family (supplemental Fig. S3). It is noteworthy that in the crystal structures of mature enzymes of the bacterial homologs, the region of the protein containing the conserved $S'$ residue exhibits a different conformation and consequently a distinct hydrogen bond geometry with the active site Ser (Fig. 4).

For instance, in mature kumamolisin, the $S'$ residue Ser$^{316}$ is in a different rotameric conformation, and its C$\alpha$ position is displaced by 1.5 Å away from that of the pro-TPP1 $S'$ Ser$^{324}$. This geometry supports a hydrogen bond between the backbone carbonyl of the active site Ser$^{466}$ and the side chain hydroxyl of Ser$^{316}$ at 2.8 Å (backbone–side chain bond) and another between the side chain hydroxyl of the active site Ser$^{466}$ and the backbone carbonyl of Ser$^{316}$ at 3.5 Å (side chain–backbone bond), in place of a strong side chain–side chain hydrogen bond in pro-TPP1. With lack of mutational data on this residue, its contribution to the catalytic machinery is not certain but is worth consideration because of its complete conservation. In pro-TPP1, Ser$^{324}$ is located at the tip of the mb4 strand. The differences in positioning of structural elements of this region compared with those in mature enzymes are likely propagated as a result of interdomain contacts involving residues C-terminal to Ser$^{324}$, viz. within the mb4–mb5 loop and the mb5 strand (region III in supplemental Fig. S5).

The side chain conformation of the D' residue Asp$^{360}$ constituting the oxyanion hole is indistinguishable from the corresponding residue in homologous structures except in prokumamolisin where the corresponding Asp$^{352}$ is in a different rotameric conformation. In addition, a bound water molecule (water 3) occupies a similar position as the backbone carbonyl of the scissile peptide bond of a substrate as observed upon superposition of the pro-TPP1 catalytic domain with an inhibitor bound homologous structure (27). This water makes strong hydrogen bonds with the side chain oxygens of active site Ser$^{475}$ and Asp$^{360}$ (Fig. 4).

Zymogens are commonly kept in their inactive states by occlusion of their active sites by propeptides that function as competitive inhibitors, occupying the substrate-binding site in a nonhydrolyzable configuration (reviewed in Ref. 75). To test whether the pro-TPP1 linker can function in this role, we docked a model hexapeptide substrate to the catalytic domain in which the side chain conformations of Glu$^{372}$ and Asp$^{376}$ were reoriented to assumed active conformations, based on the structures of mature enzymes of bacterial homologs. This hexapeptide was readily accommodated in the absence of the propiece linker. However, when the linker was superimposed on the docked model, the P3 and P3' residues overlap with linker residues Pro$^{378}$ and Arg$^{383}$, respectively, suggesting steric occlusion. The docking study further suggests that the linker residue Arg$^{183}$ would hinder approach of the substrate to the S2' and S3' subsites (Fig. 5).

In addition, a putative salt bridge between the positively charged Arg$^{183}$ and negatively charged Asp$^{360}$ is also likely to render a nonfunctional oxyanion hole. This is consistent with the $pK_a$ of 4.7 predicted for Asp$^{360}$ in the zymogen form (supplemental Table S4). A corresponding salt bridge between Arg$^{165}$ and the oxyanion hole Asp$^{352}$ observed in prokumamolisin has been implicated in substrate-mediated inactivation (28). However, in pro-TPP1 this salt bridge is absent (>5.0Å), presumably because of crystal lattice contacts, positively charged Arg$^{183}$ of the linker is likely to participate in electrostatic interactions with negatively charged Asp$^{360}$ of the oxyanion hole, potentially restricting its role in general acid and base catalysis.

**Structural Basis for Tripeptidyl Peptidase Activity**—The substrate-binding subsite and the active site pocket are likely to undergo subtle changes in the mature enzyme with different rotameric conformations of catalytic triad and neighboring residues (such as Ser$^{324}$, see above). Nevertheless, a docked model of a hexapeptide substrate AAFFAA revealed that the Asp$^{327}$ side chain is a likely candidate for binding to the positively charged unsubstituted N termini of P3 residues of substrates, thus facilitating release of tripeptides after catalysis (Fig. 5). Asp$^{327}$ is completely conserved in all mammalian TPP1 proteins but not among the bacterial homologs (supplemental Fig. S3).

Interestingly, in the latter, a neighboring Trp side chain blocks this region. In pro-TPP1, the Asp$^{327}$ carboxyl is 5.3 Å away from the N terminus of the modeled hexapeptide,
although this is likely to change in the mature domain. In addition, pK_a predictions using the PROPKA program (64) indicate that the pK_a of Asp327 is 6.2 when it is buried as part of pro-TPP1, whereas it is 3.9 when exposed in the mature enzyme. From this it follows that at pH < 3.9 it is likely to be protonated when its binding to the positively charged N terminus of a substrate is suboptimal. This correlates with the K_a profile of the enzyme, which reveals a substantial increase in K_a below pH 4.0 (Fig. 2). Such a role for Asp327 was also suggested from the previously predicted homology model for TPP1 (33), supported by site-directed mutagenesis (34).

Model for Autoproteolytic Endopeptidase Cleavage of Pro-TPP1—MALDI-TOF-mass spectrometric (MS) analysis revealed differences in the size of the liberated propeptide at different pH values (Fig. 6A). This indicates that thezymogen can utilize three different endopeptidase cleavage sites (Fig. 6B) as follows: 1) Ser181–Leu182 site within pH range 4.0–4.5; 2) Gln189–Val190 site within pH range 3.0–4.0; and 3) Asp168–Phe169 site within pH range 2.5–4.0. Cleavages at sites 1 and 2 have been reported previously (13).

Calculations using APBS electrostatic analysis software (61) with pro-TPP1 domains revealed that the binding energy between the pro-piece and mature domain is more positive (i.e. interactions are less stable) with lower pH (Fig. 6C). A similar profile is also observed for calculations done using prokumamolisin coordinates (data not shown). This phenomenon of pH-induced effects on binding is intuitive. Key salt bridge interactions as well as hydrogen bonds contributed by ionizable residues may be disrupted at low pH where acidic residues are likely to become protonated. This conclusion is supported by a recent report that the extent of inhibition of TPP1 activity by the prodomain decreases with lower pH (68). In addition, complex formation between kumamolisin-As E78H/
Structure of Pro-TPP1

D164N mutant mature domain and wild-type propiece was observed to be less at lower pH (29). These results are consistent with a pH-induced change in binding between the propiece and the mature domain. In pro-TPP1, the binding interface has many electrostatic bonds and ionizable residues with $pK_a$ values within the physiological pH range (Fig. 6B and supplemental Fig. S6 and supplemental Tables S2 and S4). Predictions of $pK_a$ values of ionizable residues at the interdomain interface reveal that these residues cluster into groups that map to specific regions of the binding interface (supplemental Fig. S6 and supplemental Table S4). These regions are associated with distinct ranges of $pK_a$ values, suggesting pH-specific disruption of domain contacts.

We propose a model of activation that takes into the account the pH-induced change in binding together with the pH-specific cleavage site preference (Fig. 6D). In the previously determined prokumamolisin structure, the cleavage site His$^{171}$–Phe$^{172}$ is positioned close to the active site Ser (Ser$^{460}$) (supplemental Fig. S4). This corresponds to cleavage site 1 (Ser$^{181}$–Leu$^{182}$) in pro-TPP1, which is relatively close to the active site Ser$^{473}$, with the carbonyl carbon of the scissile peptide being 12.0 Å away. In the absence of the lattice interactions that alter the position of the linker in our crystals grown at pH 5.0, this site is likely to approach Ser$^{473}$. However, under the conditions that we use in our in vitro activation reactions, autocatalytic cleavage at pH 4.5 and above is extremely slow, with the intact pro-TPP1 zymogen persisting for days (data not shown). It is likely that tight binding between the prodomain and catalytic domain at this pH holds the zymogen in a conformation that favors the nonoptimal active site geometry, and that the small amount of cleavage observed might arise from local breathing motions that allow acquisition of a conformationally active catalytic site. However, given that the prodomain retains high affinity for the catalytic domain and remains associated even after cleavage (68), the equilibrium would favor the inactive conformation. At lower pH, the prodomain-catalytic domain interface could become destabilized, allowing assumption of a more conformationally optimal active site in the catalytic domain, as well as reorientation of the prodomain and linker. After endoproteolytic cleavage, the liberated propeptide as well as portions of the linker region still attached to the catalytic domain could then be further digested by the free catalytic domain (see below).

Consistent with this model, rates of pro-TPP1 activation are greatly accelerated in the presence of denaturants such as urea (supplemental Fig. S7) and guanidine hydrochloride (data not shown). In addition, the environment of late endosomes and lysosomes may be conducive to destabilization of the prodomain-catalytic domain complex, allowing rapid activation at the physiological pH of these organelles (pH 6 to 4.5) (76, 77). For instance, glycosaminoglycans have been shown to accelerate autoactivation of pro-TPP1 at mildly acidic pH (16), as do liposomes containing bis(monoacylglycerol)phosphate (also called lysobisphosphatidic acid), an anionic lipid enriched in late endosomes (78). Nonetheless, decreasing the pH to optimize the protonation states of residues in the catalytic center, combined with reorientation of the catalytic serine and glutamic acid that would be favored by destabilization of the prodomain-catalytic domain interaction, are likely to be key elements in acquisition of activity.

Pathway for Pro-TPP1 Processing in Vitro—Analysis of hexahistidine-tagged glycosylated recombinant TPP1 using SDS-PAGE (Fig. 7A), chemical sequencing (Fig. 7C), and MALDI-TOF MS (data not shown) corroborates the MS analysis of the Endo H-deglycosylated untagged TPP1 (Fig. 6A), and it also reveals that autocatalytic processing involves a series of pH-dependent endoproteolytic cleavages and tripeptidyl aminopeptidase cleavages. At pH 4.0, the predominant initial cleavage products correspond to an intermediate starting at Val$^{190}$ (Fig. 7C) and residues 20–189 of the propeptide (Fig. 7A, band marked F2 and MALDI-TOF MS; data not shown). Over time, the intermediate undergoes two tripeptidyl-peptidase cleavages, yielding an intermediate starting at Thr$^{193}$ and the mature protein starting at Leu$^{196}$ (Fig. 7C), whereas the liberated propeptide is degraded (Fig. 7A). At pH 2.5, the predominant initial cleavage products correspond to an intermediate starting at Phe$^{169}$ (Fig. 7A, band marked *, and C) and residues 20–168 of the propeptide (Fig. 7A, band marked F3 and MALDI-TOF MS; data not shown). With increasing time, the intermediate undergoes a second endoproteolytic cleavage to liberate the intermediate form starting at Val$^{190}$ with subsequent TPP1 cleavage as above (Fig. 7C). Initial endoproteolytic cleavages of peptide bonds Asp$^{168}$–Phe$^{169}$ and Gln$^{189}$–Val$^{190}$ occur both at pH 3.0 and 2.5 (Fig. 7A and C). In addition to the tripeptidyl-peptidase cleavages of the form with an N-terminal Val$^{190}$ observed at pH 4.0, there is also processing of the form starting at Phe$^{169}$, yielding intermediates with N-terminal Gly$^{172}$ and Arg$^{173}$.

The observed cleavage pattern can be summarized by the pathway shown in Fig. 7D. This is entirely consistent with the known exopeptidase specificity of TPP1, where peptides with prolines at either P1 or P1’ are extremely poor substrates (20), thus preventing further exopeptidyl digestion past Arg$^{173}$. The Leu$^{182}$ N terminus produced by cleavage at pH 4.5 (Fig. 6A) (13) could not be processed by progressive tripeptidyl peptidase cleavages to yield mature TPP1 because of both frame and specificity considerations, and would be predicted to undergo a second endoproteolytic cleavage at Gln$^{189}$–Val$^{190}$ to yield the form with an N-terminal Leu$^{196}$. The liberated propeptide has an N terminus SYSP that is unlikely to be efficiently digested by TPP1 exopeptidase activity. Thus, in the absence of any additional proteases, its hydrolysis probably involves TPP1 endoproteolytic cleavage. Activation at pH 4.0 can be accelerated by other proteases (data not shown), and given the protease-rich environment of the lysosome, it is likely that downstream hydrolysis of the initial propeptide cleavage product is not rate-limiting in vivo.

It has been recently reported that the propeptide can inhibit the activity of mature TPP1 at neutral and mildly acidic pH, with an IC$_{50}$ in the micromolar range at pH 4.0 (68). Our results on comparing zymogen cleavage and acquisition of TPP1 activity are consistent with this. There was a marked lag between the two events at pH 4.0 (Fig. 7B, $t_{1/2}$ cleavage ~20 min; $t_{1/2}$ activation...
In contrast, the appearance and disappearance of the free propeptide could be fit with a two-step consecutive irreversible reaction model (36), with the second phase having a $t_{1/2}$ of $7\text{ h}$, paralleling that of acquisition of TPP1 activity. Thus, the lag is likely to reflect the time required for proteolysis of the inhibitory prodomain. The lag at lower pH values was much less pronounced, with the difference in the half-times between the first phase of the activation reaction and zymogen cleavage differing by at most a factor of 2 and both processes occurring rapidly (Fig. 7B; pH 2.5, 3.0, and 3.5; $t_{1/2}$ cleavage 3, 6, and 20 s; and $t_{1/2}$ activation 5, 12, and 20 s, respectively). This is consistent with rapid dissociation of the cleaved propeptide at more acidic pH values.

**Biomedical Relevance**—The structure of pro-TPP1 provides some insights into disease-causing mutations in LINCL (discussed in the accompanying paper by Pal et al. (81) (supplemental Fig. S8). There are currently 57 mutations listed in the NCL mutation data base. Roughly half (29 mutations) are missense

---

**FIGURE 7.** Proteolytic processing and autoactivation of pro-TPP1 zymogen. A, autoproteolytic processing. Activation reactions were conducted using 1–1.5 $\mu M$ pro-TPP1, and aliquots (typically 0.5–1 $\mu g$) were analyzed by SDS-PAGE as described under “Experimental Procedures.” Z, position of $66$-kDa glycosylated pro-TPP1 zymogen; asterisk, $49$-kDa intermediate; M, $46$-kDa glycosylated mature TPP1; F₁, F₂, and F₃, fragments cleaved at sites 1, 2, and 3, respectively, with assignment of F₁ being based on MALDI-TOF analysis of deglycosylated TPP1 (Fig. 6A). B, comparison of proteolytic processing and acquisition of TPP1 activity. Two to four gels for each pH were used for the analysis and background-corrected intensity data normalized for each gel. The green upward pointing triangles indicate disappearance of zymogen at each time point, normalized to the intensity at $t_0$ (e.g. $1/Z_{t_0}$). The red downward pointing triangles indicate levels of liberated propeptide ($F_1 + F_2 + F_3$), normalized to intensity of the highest value obtained for the reaction time course. Open circles indicate TPP1 activity measurements on activation reactions initiated using 1.5 $\mu M$ zymogen. For comparison, data from Fig. 1 conducted using 15 nM protein are presented using shaded circles. Although the time course was similar at both concentrations for pH 4 and 3.5, there were noticeable differences at pH 2.5 and 3.0. For the latter reactions, the initial phase for acquisition of activity was similar, but the second phase was noticeably slower for the high concentration reactions. The reason for this is not entirely obvious, but it may reflect experimental considerations such as some reassociation of the prodomain after the activation reactions were arrested by dilution into pH 5.0 buffer. Similarly, the faster time course and coincidence of the acquisition of enzymatic activity and endoproteolysis at pH 4.0 that we observed previously (14) may reflect differences in experimental design, where the reactions were terminated by dilution into pH 4.0 buffer and then assayed immediately for TPP1 activity. Under these conditions, the propeptide would be expected to be largely dissociated from the mature protein. C, N-terminal sequence analysis of autocatalytically processed TPP1. Activation reactions were conducted using 1.5 $\mu M$ pro-TPP1. Protein (17 pmol/time point) was resolved by SDS-PAGE and transferred to membranes. Bands of molecular mass ranging between 46 and 49 kDa were analyzed by chemical sequencing. The table lists the amount of each form in picomoles as estimated from the sequencing data (see “Experimental Procedures” for details). D, model for pro-TPP1 processing. Symbols are as in C and Fig. 6 (see text for details).
mutations, some of which could respond to small molecule chaperone therapy directed at stabilizing the aberrant protein and thus rescuing function. For these, the structure of pro-TPP1 may help in identifying such small molecule therapeutics. However, most individuals with LINCL would not benefit from this approach as two mutations that encode totally nonfunctional proteins represent the majority of disease alleles (79). One is an invariant 3′ splice acceptor mutation in intron five that introduces a frameshift upstream of the Phe<sup>169</sup> codon; the other is a nonsense mutation at codon Arg<sup>208</sup> that does not appear to efficiently respond to potential therapeutics that promote translation readthrough (80).

Based on these and other considerations, any widely applicable therapy will need to target sufficient amounts of TPP1 to the brain to digest both newly arriving substrates and pre-existing lysosomal storage material. For enzyme replacement therapy using peripherally administered protein, it will be necessary to modify pro-TPP1 so that it can cross the blood-brain barrier (e.g. creation of a chimeric protein containing elements that allow transcytosis). In addition, for both enzyme and gene therapy approaches, it would be extremely useful to engineer functional TPP1 derivatives that have a longer half-life in the lysosome, thereby increasing the effectiveness of properly targeted protein. Knowledge of the structure of TPP1 should be invaluable for such efforts.

Acknowledgments—We thank the Steinfeld and Sheldrick laboratories for sharing their manuscript and coordinates prior to submission; Bill Canfield for suggesting the use of and providing kifunensin to produce Endo H-sensitive pro-TPP1; John Taylor for advice and syntheses for sharing their manuscript and coordinates prior to submission; hong Wang, Caifeng Zhao, and Haiyan Zheng for invaluable assistance; and David Sleat and Vikas Nanda for helpful discussions.

REFERENCES

1. Sleat, D. E., Donnelly, R. J., Lackland, H., Liu, C. G., Sohar, I., Pullarakt, R. K., and Lobel, P. (1997) Science 277, 1802–1805
2. Rawlings, N. D., and Barrett, A. J. (1999) Biochim. Biophys. Acta 1429, 496–500
3. Vines, D. J., and Warburton, M. J. (1999) FEBS Lett. 443, 131–135
4. Cabrera-Salazar, M. A., Roskelley, E. M., Bu, J., Hodges, B. L., Yew, N., Dodge, J. C., Shihabuddin, L. S., Sohar, I., Sleat, D. E., Scheule, R. K., Davidson, B. L., Cheng, S. H., Lobel, P., and Passini, M. A. (2007) Mol. Ther. 15, D160–D164
5. Sohar, I., Sleat, D. E., and Lobel, P. (2004) in Handbook of Proteolytic Enzymes, W. H. Freeman & Co., New York, p. 37, pp. 1893–1896 Elsevier/Academic Press, San Diego
6. Chang, M., Cooper, J. D., Sleat, D. E., Cheng, S. H., Dodge, J. C., Passini, M. A., Lobel, P., and Davidson, B. L. (2008) Mol. Ther. 16, 649–656
