

Mass spectrometry study of a transferrin-based protein drug reveals the key role of protein aggregation for successful oral delivery

Cedric E. Bobst^a, Shunhai Wang^a, Wei-Chiang Shen^b, and Igor A. Kaltashov^{a,1}

^aDepartment of Chemistry, University of Massachusetts, Amherst, MA 01003; and ^bDepartment of Pharmacology and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA 90033

Edited* by George H. Lorimer, University of Maryland, College Park, MD, and approved July 10, 2012 (received for review April 24, 2012)

A recently designed human growth hormone/transferrin fusion protein (GHT) remains one of the very few examples of a protein capable of eliciting measurable therapeutic response after oral administration. To better understand the underlying factors that resulted in this rare success of nonparenteral protein drug delivery, we analyzed proteolytic stability and receptor binding properties of this protein, the key factors in overcoming the primary barriers to successful oral delivery. Analysis of GHT by a combination of size exclusion chromatography and mass spectrometry revealed that a significant protein population exists in an oligomeric (GHTx) state in addition to the anticipated monomer (GHT1). These states of GHT were evaluated for their survivability in stomach-like conditions, as well as their ability to bind transferrin receptor (TfR). Our results reveal an exceptional stability of GHTx, as well as the preserved ability to bind TfR, a critical first step in crossing the epithelial-intestinal barrier through receptor-mediated transcytosis.

protein therapeutic | electrospray ionization | native electrospray ionization MS | protein/receptor interaction | recombinant fusion protein

The spectacular success of protein-based therapeutics in the past decade stems from the high selectivity and specificity with which these medicines interact with their therapeutic targets. However, despite significant investments and extensive research and development efforts, the progress has been disappointingly slow in developing biopharmaceuticals for oral administration. Two major challenges encountered in this field are inability of the vast majority of proteins to cross physiological barriers, such as the epithelial-intestinal barrier, and the extreme lability of most proteins in the gastrointestinal (GI) tract (1). As a result, the route of administration of the ever-growing number of protein therapeutics is still limited to i.v. or s.c. injection. A mechanism of oral delivery—an alternative administration strategy that is convenient and non-invasive—is often seen as the “holy grail” of protein therapy (2, 3). A rare success story in this field is a recently designed growth hormone/transferrin fusion protein (GHT) (4), which has been shown to be able to overcome the obstacles faced by oral biopharmaceuticals, and demonstrated biological activity after oral administration (4). Tethering a drug payload to human serum transferrin (Tf) had been recognized for some time as an attractive method to enable its precise delivery to the target cells by taking advantage of receptor-mediated endocytosis (5). A similar but somewhat less frequently acknowledged feature of Tf trafficking is its ability to undergo receptor-mediated transcytosis (RMT) in certain polarized cells (6, 7). Although the ability of this protein to interact with transferrin receptor (TfR) expressed on the surface of such cells may be sufficient to transport Tf (of Tf-based drugs) across the blood-brain barrier (8), the apparent success of GHT in crossing the intestinal-epithelial barrier is less straightforward and warrants detailed investigation, because it may provide important clues toward successful design of the next generation of oral biopharmaceuticals. Indeed, although the presence of TfR throughout the GI tract has been firmly established (9), it remains unclear whether the growth hormone segment of GHT interferes

with its ability to bind TfR. More importantly, it remains a mystery as to what allows GHT to escape the common fate of nearly all proteins in the extreme environment of the stomach, acid hydrolysis and proteolysis by pepsin. In this work we approached these questions using mass spectrometry (MS)-based methods to probe various aspects of GHT structure and behavior that are critical for its apparent success as an orally administered therapeutic.

Results

On the basis of its amino acid sequence (*SI Results*), GHT has a calculated mass of 104.8 kDa, which is well within the range of MS. However, initial attempts to analyze GHT [expressed in human embryonic kidney cells (4)] directly by electrospray ionization (ESI) MS were not successful, and the sample was subsequently analyzed by size exclusion chromatography (SEC), which resolves GHT into two distinct fractions (Fig. 1*A*). Peak A elutes very close to the column's void volume, whereas the elution volume of peak B is consistent with the calculated mass of monomeric GHT. ESI MS analysis of fraction B carried out at neutral pH and physiological ionic strength (150 mM) provides a mass of 103.8 ± 1.4 kDa that corresponds to the monomeric form of GHT (Fig. 1*B*); this fraction was therefore designated as GHT1. The narrow charge state distribution exhibited by GHT1 ions clearly suggests that the protein maintains compact conformation in solution (10), consistent with the notion of both Tf and growth hormone segments being natively folded. The shoulder peak in GHT chromatogram (coeluting with control Tf) is likely to correspond to heat shock protein 70 (Hsp70), a ubiquitous mammalian chaperone actively expressed under stress conditions (11). Because its elution time is close to that of GHT1, traces of this protein can also be seen in the GHT1 fraction (ionic signal in the *m/z* region 3500–4500 in Fig. 1*B*).

Considering the high level of GHT abundance established by SDS/PAGE, it seemed probable that this protein, at least to some degree, was also a constituent of fraction A. Dynamic light scattering measurements indicate that the components of this fraction are large species, the majority of which are submicron particles (Fig. 1*C* and *SI Text*). Although the molecular species of such size cannot be analyzed directly by conventional ESI MS, their composition can be established using the so-called “bottom-up” MS analysis (12), whereby the proteins of interest are digested using endoproteases, followed by proteolytic peptide analysis by LC/MS/MS and search of a database to identify the precursor protein(s). The bottom-up method developed and

Author contributions: C.E.B. and I.A.K. designed research; C.E.B. and S.W. performed research; W.-C.S. contributed new reagents/analytic tools; C.E.B., S.W., W.-C.S., and I.A.K. analyzed data; and C.E.B. and I.A.K. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: kaltashov@chem.umass.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1206924109/-DCSupplemental.

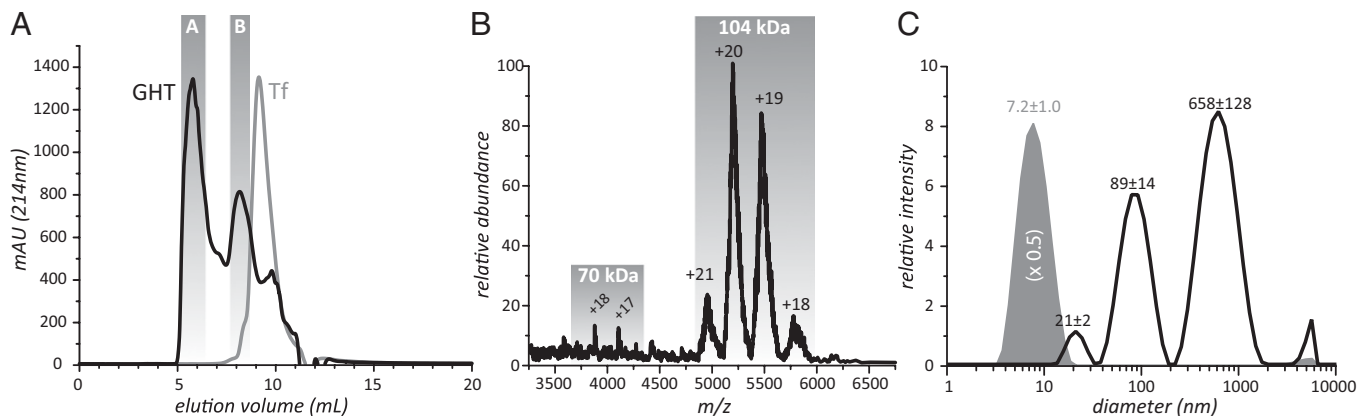


Fig. 1. (A) SEC chromatograms of GHT (black trace) and Tf standard (gray trace). (B) ESI MS of SEC fraction B. (C) Dynamic light scattering analysis of SEC fraction A (black trace) and a reference Tf sample (gray).

optimized with Tf in this work includes reduction and alkylation of cysteine residues under denaturing conditions before digestion with trypsin. Despite reducing and denaturing steps, GHT fraction A proved significantly more resistant to proteolysis compared with Tf, and a considerably higher quantity of trypsin was used to produce a satisfactory yield of proteolytic fragments. Analysis of the data using the amino acid sequence of GHT peptide identified fragments derived from both Tf and the growth hormone segments of GHT; additionally, several peptides were identified from the helical linker segment of GHT (more details in *SI Text*). This mass fingerprint analysis unequivocally confirms the presence of GHT in submicron particles making up SEC fraction A, suggesting that they are likely to be large soluble GHT aggregates; these species were collectively designated GHTx to reflect their oligomeric nature.

One of the most intriguing aspects of GHT is its apparent ability to survive the harsh stomach environment, a property that was conjectured to be a result of the Tf moiety somehow providing protection against proteolysis (13). However, Tf alone does not

display any anomalous antiproteolytic stability against pepsin, even at suboptimal pH and/or temperature (14, 15). To better understand the surprising stability of GHT in the stomach, susceptibility of GHT1, GHTx, and a Tf control to hydrolysis were examined under strongly acidic conditions in the presence of pepsin at 37 °C. The extent of proteolysis was assessed by monitoring absorbance (214 nm) at the peak elution volume of the intact protein by SEC (Fig. 2 A and B). The stability of these three samples is readily compared in the kinetic plot of Fig. 2C. The control Tf sample is efficiently degraded, with <10% of intact protein remaining after 30 min and no detectable protein after 1.5 h, consistent with the previous reports of its susceptibility to efficient digestion with pepsin (14, 15). GHT1 displayed only slightly higher stability than Tf, with only 15% of the protein surviving after 1.5 h of exposure to acid and pepsin. In stark contrast, GHTx was remarkably stable, displaying only a negligible decrease in abundance over the first 5 h of digestion, and more than 50% of the protein surviving after 21 h.

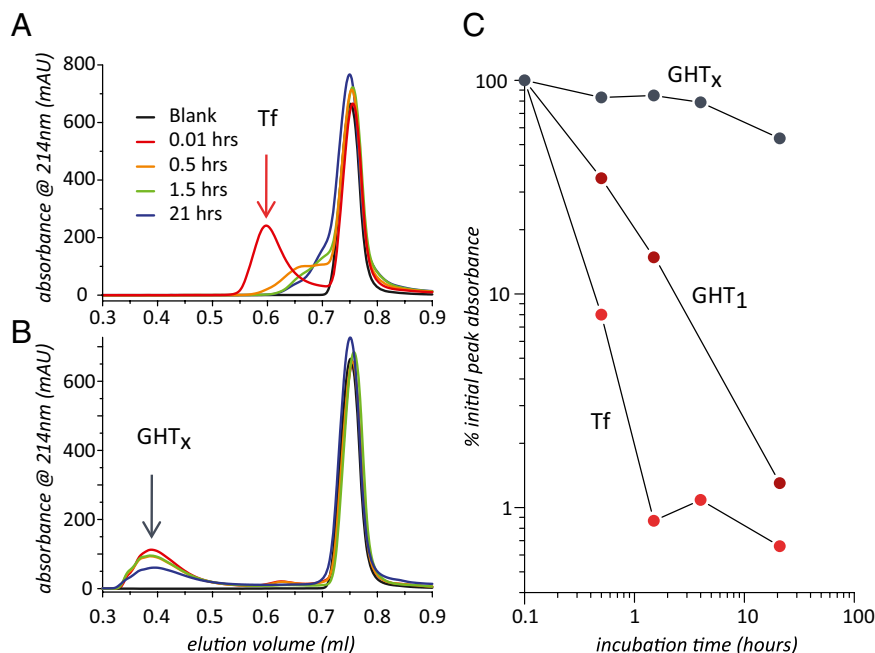


Fig. 2. Progress of TF (A) and GHTx (B) hydrolysis with pepsin at pH 3.5 monitored by SEC and kinetics of TF, GHT1, and GHTx hydrolysis (C).

signal from the mass spectrum shown in Fig. 3*B* and vice versa (with all protein concentrations in low- and submicromolar ranges).

Unfortunately, these relatively straightforward approaches cannot be readily applied to determine whether TfR also recognizes GHTx. Even unbound GHTx, let alone its putative complex with TfR, cannot be detected by ESI MS owing to its large size and heterogeneity (*vide supra*). SEC alone is not helpful either, because free GHTx elutes with the column void, and a shift to an earlier retention is simply not possible even if it associates with TfR. Nevertheless, we note that the addition of GHTx to an excess of TfR does result in an increase in the peak area of the early-eluting species (Fig. 4*A*), consistent with the notion of a complex formation between TfR and GHTx (a control experiment in which an excess amount of albumin was added to GHTx did not result in detectable change of the area of the early eluting SEC peak; more details in *SI Text*). Isolation of this early-eluting peak followed by the bottom-up analysis of its composition reveals the presence of TfR. LC/MS/MS analysis of the tryptic digest of this fraction (Fig. 4*B*) positively identified several proteolytic fragments derived from TfR (in addition to the previously matched GHT fragments; more details in *SI Text*) using the Mascot search engine. This allowed the presence of TfR to be unequivocally established in the early-eluting SEC fraction using the probability-based protein identification criteria (17). Identification of one such TfR fragment (T71, LTTDFGNAEK) is illustrated in Fig. 4 *C* and *D*, where the MS/MS pattern of one of the digestion products of the early-eluting SEC fraction of GHTx/TfR mixture closely matches that of a tryptic fragment T71 derived from the reference TfR sample (Fig. 4*C*). The elution time of this peptide on a C₁₈ column (25 min) is also identical to that of the T71 tryptic fragment derived from TfR in the absence of GHTx. Because free TfR elutes at a much later time in SEC, cross-contamination of the early-eluting SEC fraction with unbound receptor can be ruled out, and association with GHTx is the only possible explanation for its presence in the early-eluting fraction (highlighted in red in Fig. 4*A*).

Discussion

This work provides very important insight into molecular processes responsible for the apparent success of GHT as an orally administered biopharmaceutical and the role of soluble aggregates. Facile formation of large soluble aggregates endows this fusion protein with remarkable stability toward the very aggressive stomach environment without obliterating its ability to recognize TfR, a critical first step in crossing the intestinal–epithelial barrier by taking advantage of RMT. Although the monomeric form of the protein is also recognized by TfR, it shows only marginal antihydrolytic stability in the stomach-like environment, and it is not yet clear whether GHT1 also contributes to GHT activity *in vivo*. In addition to the dramatic enhancement of the antihydrolytic properties, oligomerization of GHT may provide another important advantage *vis-à-vis* crossing the epithelial–intestinal barrier. Indeed, chemically cross-linked oligomeric Tf had been shown to alter the endocytotic pathway to favor a slow recycling compartment and transcytosis (18, 19). Although the current paradigm attaches an unquestionably negative connotation to a wide range of protein aggregation phenomena, particularly in the biopharmaceutical arena, examples begin to emerge whereby aggregation of protein drugs can be used to enhance their therapeutic properties (20). The work presented in this article questions the negative stigma attached to the aggregation processes by clearly demonstrating that

they may be exploited in the design of efficient orally administered protein therapeutics. Obviously a large amount of work needs to be done to address safety and pharmacokinetic properties of soluble protein aggregates, but the ultimate goal of creating a new class of oral protein drugs no longer seems so elusive.

Materials and Methods

SEC and Dynamic Light Scattering. The mobile phase used for all SEC runs consisted of 140 mM ammonium acetate and 10 mM ammonium bicarbonate, pH adjusted to 7.2 with formic acid. SEC for GHT separation and GHTx binding to TfR was performed using a TSKgel G3000swxl; GHT1 binding to TfR was performed on a TSKgel G2000swxl. The dimensions of both columns were 30 × 7.8 cm, and they were run at 1 mL/min. Standards used for calibration were purchased from Sigma (St. Louis, MO): salmon sperm DNA (void volume), thyroglobulin, Tf, BSA, carbonic anhydrase, and cytochrome c. Dynamic light scattering data for GHTx SEC fractions and a control sample of Tf (5 mg/mL) in SEC buffer were collected on a Zetasizer Nano ZS and analyzed with Zetasizer software version 6.20 (Malvern).

Offline ESI MS. Protein samples (1–5 μM) in 140 mM ammonium acetate and 10 mM ammonium bicarbonate, pH adjusted to 7.2 with formic acid, were loaded into borosilicate metal coated emitters (Thermo Fisher Scientific, Pittsburgh, PA) and directly infused into a QStar-XL (AB SCIEX, Toronto, Canada) hybrid quadrupole/time-of-flight mass spectrometer equipped with a nanospray source.

Incubation with Pepsin. Protein stability under stomach-like conditions was assessed by incubating the target protein (0.2–0.25 mg/mL) with a protein to pepsin ratio of 300:1 (by weight) in a solution of 85 mM ammonium acetate and 6 mM ammonium bicarbonate, pH adjusted to 3.5 with HCl, and placed in a 37 °C water bath. At the designated time points a 10-μL aliquot was removed and injected on a TSKgel Super SW3000 30 × 1.2 cm (Tosoh, King of Prussia, PA) using the mobile phase described above and run at 65 μL/min.

Bottom-Up Analysis. Samples were denatured with 6M Guanidine HCl, reduced with TCEP (tris(2-carboxyethyl)phosphine) and DTT (2 mM and 4 mM, respectively), at 50 °C for 20 min, followed by alkylation of cysteine residues with iodoacetamide (12 mM at 50 °C for 20 min in the dark) and digested with trypsin using a published procedure (21) [buffer exchanged into 50 μL of 100 μM ammonium bicarbonate, followed by addition of trypsin to a 20:1 substrate-to-enzyme ratio (except for GHTx samples, which had an estimated ratio of 5:1), and incubated at 37 °C for 24 h]. A Vivaspin 500 (Sartorius Stedim Biotech, Goettingen, Germany) concentrator with a molecular weight cutoff of 10 kDa was used as the reaction container for all steps. The resulting peptides were analyzed by LC/MS/MS using an LC Packings Ultimate (Dionex/Thermo Fisher Scientific) nano-LC system coupled to the above-mentioned mass spectrometer. The samples were resolved on a C18 column (Acclaim PepMap 100 C18, 3 μm, 75 μm i.d. × 15 cm) at a flow rate of 0.2 μL/min with a gradient as follows: 0–10% solvent B in 6 min, 10–50% solvent B in 34 min, 50–100% (vol/vol) solvent B in 3 min, 100% solvent B in 5 min, followed by 0% solvent B in 12 min. Mobile phase A was 0.1% formic acid, and mobile phase B was acetonitrile with 0.1% formic acid. The MS was operated in a data-dependent mode, whereby each MS1 scan was followed by two MS/MS scans, in which the two most abundant ions detected in the MS1 scan were fragmented by collision-activated dissociation (CAD). The precursor ion selection window was set at unit resolution in all MS/MS experiments. Data were analyzed using the public Mascot (Matrix Science) server for database searches or the Biotools 1.1.1 package within Analyst QS 1.4.1 (AB SCIEX) for sequence targeted analysis.

ACKNOWLEDGMENTS. We thank Dr. Yisheng Xu and Prof. Paul L. Dubin for providing access to dynamic light scattering equipment and help with data acquisition, and Arseniy Belov for help with acquiring the negative control size exclusion chromatography data. This work was supported by National Institutes of Health Grants R01 GM061666 (to I.A.K.) and R01 GM063647 (W.-C.S.).

- Morishita M, Peppas NA (2006) Is the oral route possible for peptide and protein drug delivery? *Drug Discov Today* 11:905–910.
- Shen W-C (2003) Oral peptide and protein delivery: Unfulfilled promises? *Drug Discov Today* 8:607–608.
- Moeller EH, Jorgensen L (2008) Alternative routes of administration for systemic delivery of protein pharmaceuticals. *Drug Discov Today* 5:e89–e94.
- Amet N, Wang W, Shen WC (2010) Human growth hormone–transferrin fusion protein for oral delivery in hypophysectomized rats. *J Control Release* 141:177–182.
- Daniels TR, Delgado T, Helguera G, Penichet ML (2006) The transferrin receptor part II: Targeted delivery of therapeutic agents into cancer cells. *Clin Immunol* 121:159–176.
- Zhang Y, Pardridge WM (2001) Rapid transferrin efflux from brain to blood across the blood–brain barrier. *J Neurochem* 76:1597–1600.
- Widera A, Norouziyan F, Shen WC (2003) Mechanisms of TfR-mediated transcytosis and sorting in epithelial cells and applications toward drug delivery. *Adv Drug Deliv Rev* 55:1439–1466.

