Supporting Information

Bobst et al. 10.1073/pnas.1206924109

SI Results

Following is the growth hormone/transferrin fusion protein (GHT) amino acid sequence and tryptic peptides identified from bottom-up analysis of GHT size exclusion chromatography (SEC) fraction 1. The sequence of growth hormone is in italics, the sequence of the linker is in bold, and the rest is the sequence of transferrin. The underlined peptides were identified by peptide mass fingerprint analysis within a mass tolerance of 50 ppm.

MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNA-
MLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTS-
LCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQF-
LRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLED-
GSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRK-
*DMDKVETFLRIVQCRSVEGSCGF***LEAEAAAKEAAAKE**-
AAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKALEVP-
DKTVRWCAVSEHEATKCQSFRDHMKSVIPSDGPSVA-
CVKKASYLDCIRAIAANEADAVTLDAGLVYDAYLAP-
NNLKPVVAEFYGSKEDPQTFYYAVAVVKKDSGFQM-
NQLRGKKSCHTGLGRSAGWNIPIGLLYCDLPEPRKP-
LEKAVANFFSGSCAPCADGTDFPQLCQLCPGCGCSTL-
NQYFGYSGAFKCLKDGAGDVAFVKHSTIFENLANKA-
DRDQYELLCLDNTRKPVDEYKDCHLAQVPSHTVVAR-
SMGGKEDLIWELLNQAQEHFGKDKSKEFQLFSSPHGK-
DLLFKDSAHGFLKVPPRMDAKMYLGYEYVTAIRNLR-
EGTCPEAPTDECKPVKWCALSHHERLKCDEWSVNSV-
GKIECVSAETTEDCIAKIMNGEADAMSLDGGFVYIA-
GKCGLVPVLAENYNKSDNCEDTPEAGYFAVAVVKK-
SASDLTWDNLKGKKSCHTAVGRTAGWNIPMGLLYNK-
INHCRFDEFFSEGCAPGSKKDSSLCKLCMGSGLNLCEP-
NNKEGYYGYTGAFRCLVEKGDVAFVKHQTVPQNTG-
GKNPDPWAKNLNEKDYELLCLDGTRKPVEEYANCH-
LARAPNHAVVTRKDKEACVHKILRQQQHLFGSNVT-
DCSGNFCLFRSETKDLLFRDDTVCLAKLHDRNTYEK-
YLGEEYVKAVGNLRKCSTSSLLEACTFRRP

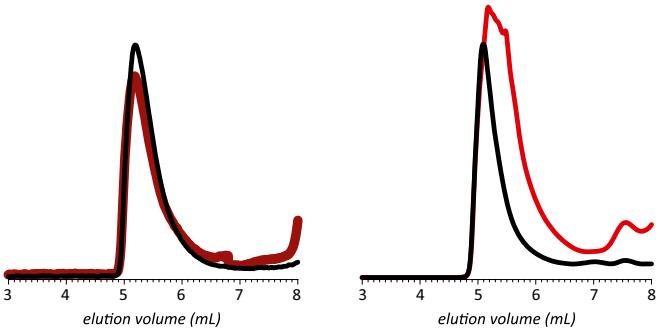

Fig. S1. Shape of the protein peaks eluting at the void volume in SEC chromatograms of the GHTx/BSA (*Left*) and GHTx/transferrin receptor (TfR) (*Right*) mixtures. Black traces show peak shape of GHTx alone. Both mixtures were prepared such that the total mass of GHTx was ~50% of that of either BSA or TfR. There is no statistically significant change in the peak area of GHTx after addition of BSA [relative SD of area under the curve (AUC) is *ca*. 3%], whereas addition of TfR results in noticeable change in the shape and AUC of the void volume peak (*ca*. 85% increase of AUC).

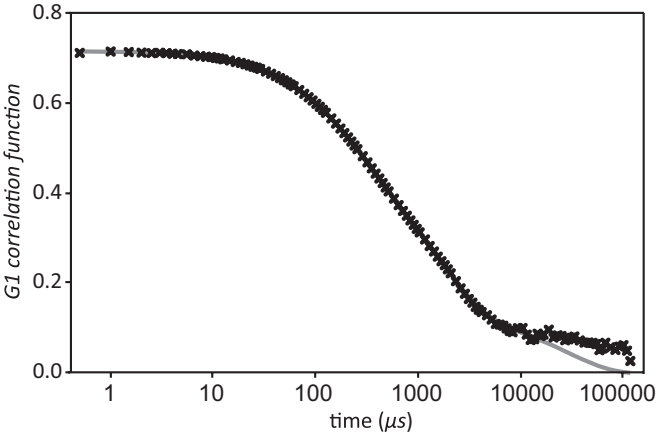

Fig. S2. Distribution fit/Malvern correlation function for dynamic light scattering data shown in Fig. 1*C* for GHTx (size distribution of soluble GHT oligomers).

Table S1. Top six proteins identified in GHTx/TfR peak fraction by Mascot MS/MS ions search engine (p7)

Protein	Protein score	Peptide matches
Cationic trypsin OS = *Bos Taurus*	836	10
Transferrin receptor protein 1 OS = Homo sapiens	94	5
Transferrin OS = Homo sapiens	89	7
Keratin, type I cytoskeletal 9 OS = *Homo sapiens*	85	4
Human growth hormone OS = Homo sapiens	81	2
Keratin, type II cytoskeletal 1 OS = *Homo sapiens*	81	5

The proteins of interest are listed in bold and italics. As noted in the text, GHTx is strongly resistant to proteolysis, which explains the relatively low protein scores. However, all scores from GHTx proteins are well above the significance threshold (scores >46 indicate identity or extensive homology at the 95% confidence level). The proteins not related to GHT or TfR are common contaminants frequently identified by bottom-up analysis of complex protein samples (bovine trypsin, used for proteolytic degradation of the protein sample, and human keratin, frequently introduced to cell culture during sample handling steps).