# THE CHOLINESTERASE-LIKE DOMAIN IS REQUIRED FOR FOLDING AND SECRETION OF THYROGLOBULIN

**by**

**Jaemin Lee**

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in The University of Michigan
2008

Doctoral Committee:

Professor Peter Arvan, Chair
Professor Ronald W. Holz
Associate Professor Kathleen L. Collins
Associate Professor Ursula H. Jakob
Associate Professor Billy Tsai

---

2008

To my mother and father for their love and support

and to my beloved brother’s family

## ACKNOWLEDGEMENTS

First, I would like to thank Dr. Peter Arvan, my thesis mentor for all his support and guidance which allowed me to finish my thesis work. I have had difficulty in progress of the research. I could not overcome it and finish my thesis without his encouragement and patience. I am glad to learn a lot from him to be an independent research scientist. I also enjoyed scientific discussions and other conversations with him. Thank you, Peter.

I also want to express my appreciation to my committee members, Drs. Kathleen Collins, Ronald Holz, Ursula Jakob, and Billy Tsai for their invaluable guidance and suggestions during my thesis period. Thanks to their advices, I could bring my work here.

I thank our collaborators, Drs. Paul Kim and Bruno Di Jeso for sharing scientific ideas and conducting research together. I also want to thank Dr. Zhaohui Xu for helpful discussion about the structure of cholinesterase.

I would like to thank Dr. Amy Chang and people in the Chang lab for getting together in the lab meeting and journal club. I enjoyed and learned a lot during the journal club.

My thanks also go to people in the Arvan lab, both past and present. I want to share my gladness to finish my thesis with them: Xiang Zhao, Jose Ramos-Castaneda (Pepe), Young-nam Park, Saeyoull Cho, Yukihiro Yamaguchi, Regina Kuliawat, Yi Zhang, Xiaofan Wang, Gautam Rajpal, Ming Liu, Roberto Lara-Lemus, Dennis Larkin,

Israel Hodish, Leena Haataja, Aaron Kellogg, and Elizabeth Cole. I will never forget those days with them. I specially want to say thank to Xiang as a friend and former roommate. I also would like to say my sincere thank to Xiaofan for her help and support during last years of my thesis.

Most of all, I must say thanks to my parents and brother's family for their understanding, care, support, and their love to me. Thank you very much, Jaephil, Mom, and Dad. I love you!

## TABLE OF CONTENTS

## LIST OF FIGURES

## ABSTRACT

Thyroid hormone synthesis requires secretion of thyroglobulin, a precursor protein comprising Cys-rich regions I, II, and III (referred to collectively as regions I-II-III) followed by a cholinesterase-like (ChEL) domain. Secretion of mature thyroglobulin requires extensive folding and oligomerization in the ER. Multiple reports have linked mutations in the ChEL domain to thyroid diseases in humans and other animals; these mutations block thyroglobulin from exiting the ER and induce ER stress. The principal hypothesis in this thesis is that thyroglobulin requires the ChEL domain to form its native tertiary and quaternary structure. We report that, in a cell culture system, thyroglobulin with misfolding mutations in the ChEL domain also impairs folding in other domains. A truncated thyroglobulin (regions I-II-III) devoid of the ChEL domain is incompetent for folding and cellular export. However, co-expression of a secretory ChEL domain (which is efficiently secreted) with the truncated thyroglobulin rescued secretion of regions I-II-III by promoting their oxidative maturation and folding. The data indicate that the ChEL domain functions as an intramolecular chaperone. Moreover, a functional ChEL domain engineered to be retained intracellularly co-retains truncated thyroglobulin even as it facilitates the oxidative maturation of regions I-II-III. This indicates that the ChEL domain is also employed as a molecular escort for thyroglobulin. Thyroglobulin becomes a homodimer within the endoplasmic reticulum. The ChEL domain sequence contains predicted helical structures that are homologous to the dimerization helices of acetylcholinesterase. I have found that the ChEL domain can dimerize with itself, just as acetylcholinesterase does, and this is likely to drive the dimerization of wild-type

thyroglobulin. Insertion of an N-linked glycan into one of the putative dimerization helices blocks detectable homodimerization of the isolated ChEL domain. Interestingly, co-expression of upstream regions of thyroglobulin, I-II-III, either *in cis* or *in trans*, overrides the dimerization defect of such a mutant. These data suggest that intermonomer interactions of the ChEL domain of thyroglobulin are enhanced by additional interactions with upstream regions of thyroglobulin. I conclude that the ChEL domain is required to form a native structure of Tg which is essential for export from the endoplasmic reticulum, and for thyroid hormone formation. This role is fulfilled by functioning as an intramolecular chaperone, an escort, and a nidus for thyroglobulin dimerization.

# CHAPTER 1

# INTRODUCTION

## Overview of thyroid hormone biosynthesis

Thyroid hormone regulates a broad range of physiological responses engaged in growth, development, and metabolism [1]. The thyroid hormone is synthesized from a matrix protein, thyroglobulin (Tg). Tg is the most abundant protein in the thyroid gland. Tg is a large homodimeric protein with a molecular weight of 330 kDa per monomer, while the thyroid hormone thyroxine is a small molecule composed of only two di-iodinated tyrosines derived from Tg.

Thyroid hormonogenesis engages a biosynthetic pathway that ultimately encompasses multiple organ systems. Tg gene expression occurs in thyrocytes of the thyroid gland (which resides in the lower part of the neck) and is positively regulated by thyroid stimulating hormone (TSH) released from the pituitary. Transcripton of Tg gene requires thyroid-specific transcription factors TTF-1, TTF-2, and Pax-8, which bind to consensus sequence elements in the Tg promoter. From the 8.3 kb mRNA, Tg is initially synthesized in the endoplasmic reticulum (ER). Tg follows the secretory pathway, via the Golgi complex (Figure 1-1). In the ER, Tg polypeptide folding is initiated, and the protein acquires N-linked sugars and forms a homodimer. After modification of the N-linked sugars in the Golgi complex, Tg is secreted by exocytosis from the apical plasma membrane of thyrocytes. Tg then enters the extracellular lumenal space of the thyroid

gland, surrounded by a tight monolayer of thyrocytes. Upon release at the plasma membrane, tyrosine residues on Tg become iodinated by thyroid peroxidase (TPO). This process requires hydrogen peroxide as an electron acceptor, which is generated by two members of the NADPH oxidase family (Duox1 and Duox2) on the apical membrane of thyrocytes (Figure 1-1). Iodide becomes available to TPO in the thyroid lumen by the action of a series of transporters that reside on the basolateral and apical membranes of thyrocytes. The sodium/iodide symporter is basolateral and assists in iodide uptake from the blood stream. The apical iodide transporter (AIT), pendrin, or other anion channels have been suggested to function as apical iodide transporters [2]. During the process of TPO-catalyzed iodination, there is specific coupling of di-iodinated tyrosyl residues of Tg. At this stage, thyroid hormones, thyroxine (T4) and triiodothyronine (T3) are still embedded within the Tg protein, which is stored at high concentration in the lumen of thyroid follicles. Secreted and stored Tg in the lumen exists primarily as non-covalent homodimers based on studies employing sucrose gradient centrifugation (19S, 660 kDa).

To release T4 or T3 into the blood stream, Tg must be proteolyzed in lysosomes within thyrocytes. This is mediated by endocytosis of Tg from the follicle lumen. Considering the high concentration of iodinated Tg in the follicle, Tg re-uptake might involve fluid-phase endocytosis, which ordinarily does not require any specific receptors. Additionally, some proteins such as N-acetylglucosamine receptor, and megalin, have been described as putative endocytic receptors for Tg, although Tg endocytosis does not seem to depend strongly on such receptors.

After release into blood, thyroxine (T4, which is the major circulatory thyroid hormone), binds to various transport proteins in the circulation. Thyroxine then becomes

available to be further activated by deiodination to triiodothyronine (T3) which is the most active form of the hormone.

**The structure of Tg**

Tg in mammals (e.g. human, rat, mouse, bovine, dog, cat) has been clearly identified and sequenced at the DNA level. Tg in all vertebrates has a similar, conserved protein structure.

Overall, Tg is comprised of four regions (Figure 1-2A) [3]. The first three regions consist of multiple cysteine-rich repeat domains distinguished by their internal homology and exon-intron structure. These regions are called I, II, and III. The C-terminal region is homologous with the entirety of acetylcholinesterase and is called the ChEL domain (CholinEsterase-Like domain).

Region I (Figure 1-2) encompassing the first ~1200 amino residues, has 10 of the 11 "Tg type 1 repeats" that have either six cysteines within the repeat (type 1A) or four cysteines (type 1B). Tg type 1 repeat was named because it was first sequenced in Tg. The type 1 repeat occurs in multiple proteins in evolution, and its internal 3D structure has been characterized. But the structure of type 1 repeats within Tg have not been characterized in the context of the full protein. The Tg type 1 repeat found in other proteins (such as equistatin, nidogen, invariant chain and insulin-like growth factor-binding protein families from sponge to vertebrates) functions in several ways, including as a cysteine protease inhibitor, as well as modulating insulin-like growth factor. The Tg type 1 repeat (especially 1A) has an even number of cysteines with a centered CWCV motif (also CHCA or CYCV in Tg) and a final cysteine (Figure 1-2B). These cysteines

form sequential intrachain disulfide bonds in the pattern as $Cys_1$-$Cys_2$, $Cys_3$-$Cys_4$, $Cys_5$-$Cys_6$. Type 1A repeats are mostly common and and make three disulfide bonds; type 1B repeats have only four cysteines and make only two disulfide bonds (Figure 1-2B).

Region II is a smaller region and contains three type 2 repeats (with two cysteines in each) as well as one final type 1 repeat (Figure 1-2). Region III has several type 3 repeats with various numbers of cysteine (six cysteines for subtype "A" and eight for subtype "B") (Figure 1-2). Type III repeats have no known homology with other proteins from current protein databases. Considering each repeat has an even number of cysteines, it seems highly likely that these cysteines are entirely consumed by intrachain disulfide bonds with "stapling" internally. Such a view is consistent with partial proteolytic digestion patterns using various proteases under native conditions, which tend to cleave between repeats [4].

The ChEL domain of Tg has ~28% identity with acetylcholinesterase [5, 6] (Figure 1-3) and with other cholinesterase-like proteins, called neuroligins (neuroligin 1 – 4). The ChEL domain has six conserved cysteines forming three intrachain disulfide bonds that have been found to be critical to maintain the three dimensional structure of acetylcholinesterase (Figure 1-3). In addition, there is a high degree of conservation of amino acid residues lining the eight β-sheets that comprise the protein backbone (Figure 1-4A). These and additional results strongly suggest that ChEL domain has a similar overall three dimensional structure to that of cholinesterase. However, amino acid residues of the catalytic triad of cholinesterase are not conserved in ChEL, which has been shown to lack cholinesterase activity (Figure 1-4A). What is the structural and functional importance of the cholinesterase-like domain of Tg, or, for that matter, other

cholinesterase-like proteins? Neuroligins, modulators of synapse formation, have recently been identified to have significant overlap of three dimensional structure with that of acetylcholinesterase [7-10] (Figure 1-4B). The cholinesterase-like domain of neuroligins does not have enzymatic activity but encodes a peptide interaction with a physiological partner, neurexin [11]. Indeed, even acetylcholinesterase maintains this peptide interaction function with other proteins such as a snake venom toxin, fasciculin [12] and other neuronal proteins [13].

Another interesting structural similarity shared by acetylcholinsterase and cholinesterase-like proteins is dimerization (Figure 1-4B). It is known that acetylcholinesterase can form a homodimer through the interaction between C-terminal helices of adjacent monomers, producing a four-helix bundle [14]. Recent structural studies reveal that neuroligins also form homodimers using the same four-helix bundle mechanism [7-10]. Conserved dimerizaton of acetylcholinesterase and neuroligins seems functionally important since mutations disrupting dimerization also prevent their function [11, 15]. Importantly, Tg also forms a homodimer and dimerization has been proposed to be a pre-requisite for its secretion [16]. Thus there is reason to think that the ChEL domain of Tg may also engage in both inter- and intra-molecular interactions.

The most proximal Tg type 1 domain may not require the presence of downstream Tg domains in order to complete its folding because a severely truncated N-terminal fragment of Tg can be secreted [17]. It is notable that this N-terminal fragment contains the two tyrosines that have been found necessary to make most thyroid hormone. This may explain why hypothyroidism caused by early termination of translation in Tg, such as is seen Dutch goats, can be relieved by an ample supply of iodine in the diet [18]. But

such studies only underscore the question of the functional role of the ChEL domain in the context of the full length Tg protein.

**Tg folding in the endoplasmic reticulum – N-glycosylation and chaperone interaction**

As a secretory glycoprotein, polysomes direct newly translated Tg to the ER by virtue of its N-terminal signal peptide (19 – 20 amino acids). The nascent polypeptide of Tg is co-translationally translocated into the lumen of the ER through the Sec61 translocon complex. Immediately after entry into the ER, secretory proteins with potential asparagine (N)-linked carbohydrate acceptor sites (N-X-S/T) may acquire core sugar modification by the action of oligosaccharyltransferase (OST), which transfers pre-assembled glucose$_3$-mannose$_9$-N-acetylglucosamine$_2$ ($Glc_3Man_9GlcNAc_2$) core oligosaccharides from the ER membrane dolichol pyrophosphate intermediate to nascent polypeptides. Among the 20 potential glycosylation sites in human Tg, 16 sites are actually glycosylated [19] (13 among 14 in bovine Tg [20]). The number of oligosaccharides may vary (as in porcine Tg [21]), but approximately 10% of total mass of Tg (i.e., ~ 30 kDa of the 330 kDa Tg monomer) comes from covalently bound sugar.

In the ER, the $Glc_3Man_9GlcNAc_2$ core oligosaccharide gets further modified by trimming of three terminal glucoses by the sequential action of ER α-glucosidases I and II. At the stage when nascent polypeptides still contain monoglucosylated oligosaccharides, they start to interact with the lectin-like ER chaperones, calnexin and calreticulin. Calnexin is an integral membrane protein while calreticulin is contained within the ER lumen. Calnexin and calreticulin form chaperone complexes with ERp57,

an ER oxidoreductase that catalyzes formation and isomerization of disulfide bonds. (As a glycoprotein with abundant cysteine-rich domains, nascent Tg interacts with calnexin, calreticulin, and ERp57 during initial folding and such interaction is prevented by the drug castanospermine, which inhibits ER glucosidases and thereby prevents formation of the lectin binding site [22-24]). Tg also interacts with other chaperones such as protein disulfide isomerase (PDI) and immunoglobulin binding protein (BiP, also called GRP78). These interactions begin even before translation is finished [24, 25]. In addition to BiP, other glucose-regulated proteins such as GRP94 and GRP170 also interact with nascent Tg and they (including BiP) seem to prevent aggregation and premature secrection of Tg instead of directly facilitating folding [26, 27].

**Tg folding in the ER – ER quality control and ER-associated degradation**

Further removal of the third and final glucose from monoglucosylated oligosaccharides by glucosidase II prevents glycoproteins from re-binding to calnexin and calreticulin. However, during folding in the ER, glucose can be re-added by the action of UDP-glucose: glycoprotein glucosyltransferase (UGGT), which senses non-native (unfolded) polypeptides. This allows immature forms of glycoproteins to re-enter the ERp57/calnexin/calreticulin pathway, preventing the premature export of incompletely folded Tg. When Tg reaches its mature, native structure, it no longer associates with calnexin/calreticulin or other chaperones so that it is free to move forward from the ER to the Golgi complex. In contrast, misfolded glycoproteins which fail to reach a native conformation are cleared from the ER by a process called “ER-associated degradation” (ERAD). Terminal mannose residues on glycoproteins destined for ERAD

are involved in recognition and degradation of misfolded glycoproteins in part via the action of both ER-resident mannosidase I and anther ERAD enhancing α-mannosidase-like protein (EDEM, Htm1p in yeast) [28, 29]. Inhibitors of mannosidase I such as deoxymannojirimycin and kifunensine prevent misfolded glycoproteins from interacting with EDEM, blocking their degradation. Current evidence suggests that ER-mannosidase I and EDEM are functional components in the ERAD of misfolded Tg [17, 30, 31].

Proteins can be degraded by proteases in the lysosome or by the ubiquitin-proteasome system in the cytosol. Many misfolded secretory proteins degraded by ERAD employ the ubiquitin-proteasome system. As such, these proteins need to move back through the ER membrane so that proteasomes in the cytosol may gain access to them. This "retrotranslocation" across the ER membrane may require integral membrane proteins such as the Sec61 translocon [32]. But the true identity of the "retrotranslocon" is still controversial. Derlin-1 has also been proposed to be a retrotranslocon for ERAD [33]. To achieve retrotranslocation, misfolded proteins are expected to be fully unfolded to pass through the ER membrane, although there may be exceptions [34, 35]. BiP prevents aggregation of proteins in the ER and may help in retrotranslocation by maintaining misfolded proteins in a soluble state. PDI is also implicated as an unfoldase that can facilitate the retrotranslocation of misfolded secretory proteins as well as bacterial toxins [36]. Though PDI is one of the ER oxidoreductases with a thioredoxin motif (C-X-X-C), its actions are not limited to substrates with cysteines or disulfide bonds [37]. Recently, ERdj5 (with a J domain and four thioredoxin domains) has been reported as the first dedicated reductase to help retrotranslocation of ERAD substrates by reducing improper disulfide bonds [38]. Of the molecules mentioned above, degradation

of misfolded Tg requires the ubiquitin-proteasome system as ERAD of Tg is prevented by proteasome inhibitors [17, 30, 31]. Before ERAD, misfolded Tg has prolonged interactions with ER chaperones such as BiP, GRP94, calnexin, PDI, ERp57, and ERp72 [17, 30, 39]. Interestingly, over-expression of PDI facilitates the degradation of misfolded Tg [40]. In this regard, misfolded Tg seems to be a prototypic ERAD substrate.

**Tg folding in the ER – ER stress and the unfolded protein response**

Under conditions of high expression and accumulation of misfolded secretory/membrane proteins in the ER ("ER stress"), cells employ various responses collectively called the "unfolded protein response (UPR)". The function of the UPR is to alleviate ER stress [41]. The UPR is mediated by three different classes of signal transducers in the ER, inositol-requiring protein-1 (IRE1), activating transcription factor-6 (ATF6), and protein kinase RNA (PKR)-like ER kinase (PERK). Each one is an integral membrane protein sensing the folding status in the lumen of the ER to elicit a stress-response pathway. IRE1 oligomerizes in the presence of a high concentration of unfolded proteins in the ER and *trans*-phosphorylates [42]. Oligomerization of IRE1 is induced either directly by unfolded proteins or indirectly via BiP binding to unfolded proteins and thereby releasing BiP inhibition of IRE1. Upon *trans*-phosphorylation, activated IRE1 uses its endonuclease domain to splice the mRNA of X-box binding protein-1 (XBP1, Hac1 in yeast). Translated XBP1, a transcriptional activator, increases the expression of UPR target genes. ATF6 is also retained in the ER by binding of BiP. The ER transmembrane form of ATF6 is also an inactivated form. Under ER stress, ATF6 is released from BiP and follows the secretory pathway to the Golgi, where it

encounters the site 1 protease (S1P) and then the site 2 protease (S2P). Release of the cytoplasmic domain of ATF6 to the cytosol allows this domain to move to the nucleus where it activates a subset of UPR genes. Finally, PERK is a type 1 transmembrane protein with a lumenal domain homologous to that of IRE1 that senses ER stress [43]. Under ER stress, PERK is liberated from BiP, oligomerizes, and *trans*-phosphorylates. PERK then becomes an active kinase that phosphorylates the α-subunit of eukaryotic translation initiation factor-2 (eIF2α) on Serine 51. The phosphorylated eIF2α suppresses further protein translation which helps cells limit proteins that could add to ER stress. Remarkably, the phosphorylated eIF2α actually allows for increased translation of ATF4 which activates the expression of another subset of UPR genes, including the transcriptional factor CHOP (C/EBP-homologous protein). Genes increased in their expression as a consequence of UPR have been identified and many of them (more than 50%) function in the secretory pathway [44].

Cells employ UPR to relieve ER stress but under conditions in which UPR fails to handle the ER stress, cells die by apoptosis [45]. IRE1 can induce apoptosis via interaction of its cytoplasmic domain with tumor necrosis factor receptor-associated factor 2 (TRAF2). This interaction subsequently activates Jun kinase (JNK) which finally leads to apoptosis [46]. Caspase-12 has been described to be activated under ER stress and although it is not yet clear what mechanism induces caspase-12 activation, TRAF2 has been described to interact with caspase-12. CHOP (positively expressed by ATF4) also activates genes which may potentiate apoptosis such as growth-arrest-and-DNA-damage-34 (GADD34) and also suppresses the expression of anti-apoptotic Bcl-2.

Perturbations of folding or increased expression of Tg can cause ER stress and thereby induce UPR [30, 47-49]. Under ER stress, thyroid cells show induced expression of multiple ER chaperones (e.g., calreticulin, BiP, GRP94, ERp29, ERp57, ERp72, and PDI) but not calnexin [30, 47-49]. Increased activities of UPR transducers (XBP1, ATF6, ATF4, and CHOP) also have been described under ER stress [47, 49]. ER stress activates JNK and nuclear factor-κB (NF-κB) in thyroid cell lines [50]. Dominant-negative TRAF2 blocks the activation of NK-κB in thyroid cells under ER stress [50]. Ultimately, increased cell death is observed in thyroid tissue expressing certain misfolded form of Tg [30].

**Tg mutations - Impact on health**

Disorders of thyroid function range from thyroid overactivity (hyperthyroidism, as exemplified by Graves' disease) to hypothyroidism. Hypothyroidism is the most common thyroid disorder worldwide, and environmental iodine deficiency is still the most important underlying factor. But even in regions with sufficient dietary iodine, hypothyroidism, especially congenital hypothyroidism (occurring in 1 to 3,000 to 4,000 live births), is one of the most common neonatal endocrine diseases [51]. Untreated congenital hypothyroidism can cause mental retardation and other growth disorders, which have historically been called "Cretinism" [51].

Congenital hypothyroidism is caused by either thyroid dysgenesis [52] or dyshormonogenesis which means a defect of thyroid hormone synthesis [53]. Considering the complicated pathway of thyroid hormone synthesis, it is not surprising that many genes have been identified as causing hypothyroidism. Thyroid-specific genes

with a defect in hormonogenesis include the TSH receptor, NIS, TPO, Pendred, Duox1, and Tg [53].

Tg mRNA is 8.5 kb; the size of the Tg gene is approximately 300 kb long and consisting of 48 exons [54, 55]. The large size may make the Tg gene unusually vulnerable to mutation. Currently, at least 39 inactivating mutations have been reported in human Tg (Table 1-1) [17, 56, 57]. Among these, 23 are missense mutations, 8 are splicing site mutations, 5 are nonsense mutations, and 3 are caused by single nucleotide deletion or insertion. Around 20% (eight mutations) are found within the ChEL domain of Tg. The disease phenotypes induced by these mutations vary from mild hypothyroidism to severe. Besides hypothyroidism, some mutations (C1058R, C1245R, C1977S, and G2356R) and polymorphisms (P2213L, W2482R, and R2511Q) have been suggested to be associated with thyroid cancer [58, 59]. Tg-related diseases have been reported also in animals such as cattle, goats, sheep, mice, rats, and antelope [56]. Specific Tg mutations causing the disease have been identified in Afrikander cattle (R697stop) [60], Dutch goat (Y296stop) [18], *cog* mouse (L2263P) [61], and *rdw* rat (G2300R) [62, 63].

Chronic autoimmune thyroiditis (e.g., Hashimoto's disease), the most common organ autoimmune disease, is associated with antithyroglobulin antibodies in at least 60% of patients [64]. Thus Tg structure, folding, and antigenicity have definite pathophysiological significance.

**Phylogeny and ontogeny of thyroglobulin**

The functional role of the thyroid hormone is well-described in mammals, especially in humans. For cold-blooded vertebrates, it still remains to be studied. Thyroid tissue is present in all vertebrates. In lamprey, thyroid tissue is not found in the larva period, but can be seen after metamorphosis to the adult. In non-vertebrates such as tunicates (Urochordata) and amphioxus (also known as lancelets, Cephalochordata), no homologous organ has been identified as thyroid tissue. But there are increasing reports showing that thyroid hormone functions in metamorphosis in these pre-vertebrates. Many homologous genes for thyroid hormonogenesis and function have been identified in amphioxus [65]. This suggests that the thyroid hormone system might have evolved earlier than the vertebrates. It is not yet clear whether or not there is a Tg-like gene in amphioxus (the genome sequencing has just been finished but its analysis is still to be done). But Tg-like iodoproteins have been identified in amphioxus and tunicates [66, 67].

Considering that most hormone synthesis occurs at the extreme N-terminus of Tg, it is an interesting question why Tg has evolved as a large molecule. As noted, the size may be disadvantageous because of increased susceptibility to inactivating mutations and autoimmunity. Studies suggest that gene duplication has been engaged in the evolution of Tg structure since 1) cysteine-rich units (type 1, 2, or 3) are present as repeats, 2) repeat, region, and domain boundaries are matched with intron-exon boundaries, and 3) domains with homologous type 1 repeats or homologous ChEL domains are found in many other proteins well before the appearance of vertebrates [68]. It has been suggested that Tg may have evolved for iodine storage (on tyrosine residues), which can be advantageous in an environment on land or fresh water where iodide is a trace element [69]. As noted above, Dutch goats, with an early termination mutation of Tg, still contains two major

hormonogenic sites that can produce thyroid hormone, but survival depends on abundance of dietary iodide [18]. This suggests that the full Tg molecule may be engaged as the central iodide reservoir in the body.

Non-mammalian Tg-like iodoproteins have similar sedimentation coefficients to those of mammalian Tg and have been described even in species where Tg has not yet been cloned [67, 70, 71]. Their amino acid compositions may differ among animals such as lamprey, fishes, reptiles, and mammals but it is notable that cysteine and tyrosine contents are surprisingly constant [70, 72].

Genome sequencing projects recently finished from various taxonomic groups although sequence analysis is incomplete. From zebrafish a nearly full-sized Tg is predicted except for the 9th type I repeat domain (gene bank accession number, XP_694292). Two major hormonogenic tyrosines (Tyr5 and Tyr130) of mammalian Tg are also conserved in predicted zebrafish Tg. Moreover, region II, region III, and the ChEL domain are contained within predicted zebrafish Tg. Though incompletely characterized, regions II-III and ChEL and at least a portion of region I have thus far been DNA-sequenced in the Tg of two pufferfishes (*Tetraodon nigroviridis* and *Takifugu rubripes*). These are consistent with previous reports about the presence of Tg-like iodoproteins from teleosts [70, 72]. Further work is needed to trace the evolutionary origins of the Tg molecule.

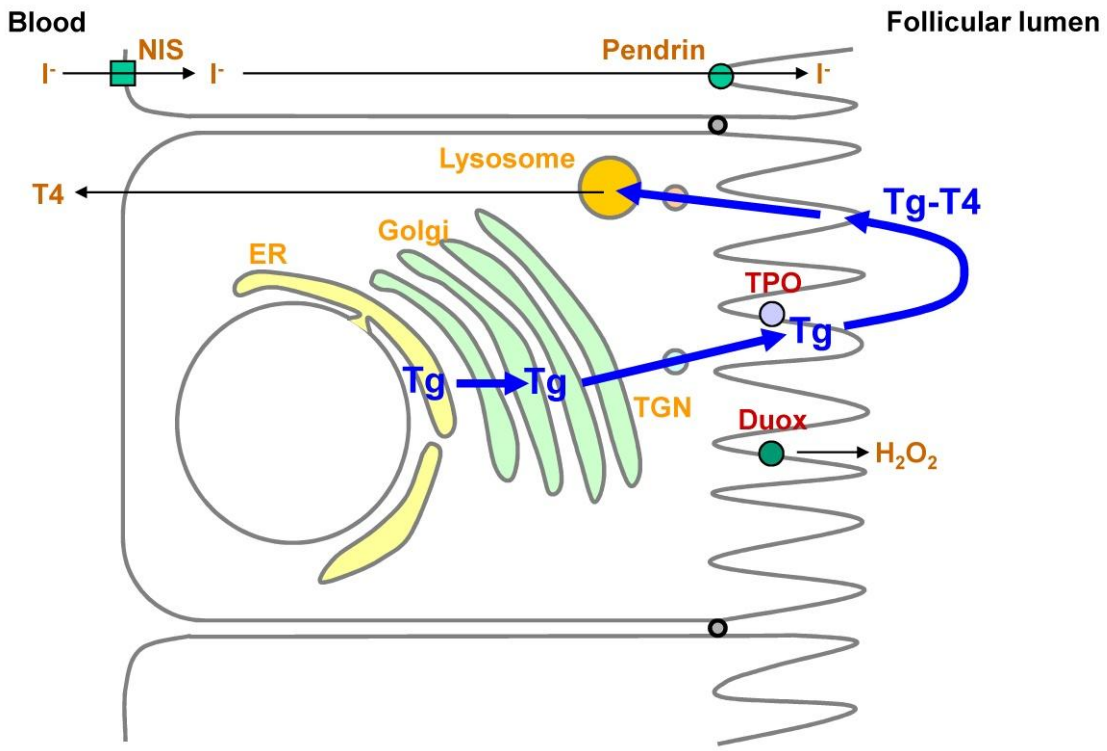


**Figure 1-1. Thyroid hormone synthesis.** Thyrocytes are depicted in simplified view. The expression of thyroglobulin (Tg) is controlled under thyroid specific transcriptional factors such as TTF-1, TTF-2, and Pax-8 which is not shown in this figure. Translation of Tg is through the endoplasmic reticulum (ER) where glycosylation and folding of Tg occur. Also, major quality control of Tg is done in the ER. Note that most of Tg forms a homodimer in the ER. Tg dimer moves through the Golgi and gets further modification of oligosaccharide chains. Tg dimer is sorted in the TGN (*trans*-Golgi network) and secrets through the apical membrane of thyrocytes by exocytosis. Upon apical membrane, thyroid peroxidase (TPO) iodinates tyrosines of Tg in the presence of hydrogen peroxide ($H_2O_2$) generated by Duox1 and Duox2. TPO does following coupling reaction between two di-iodinated tyrosines from Tg to produce thyroxine (T4). Iodide is accumulated in the lumen of thyroid gland by a series of transporters such as sodium/iodide sympoter in the basolateral membrane and pendrin in the apical membrane. Iodide uptake can be happened in the same thyrocytes making Tg and is depicted here as in separated thyrocytes for presentation purpose. Triiodothyronine (T3) is also made in small amount. T4 (and T3) is still embedded within Tg. Free thyroid hormones (T4 and T3) are released from the lysosome where Tg taken by endocytosis is degraded. This pathway shows that traffic and secretion of Tg are important for hormone synthesis.

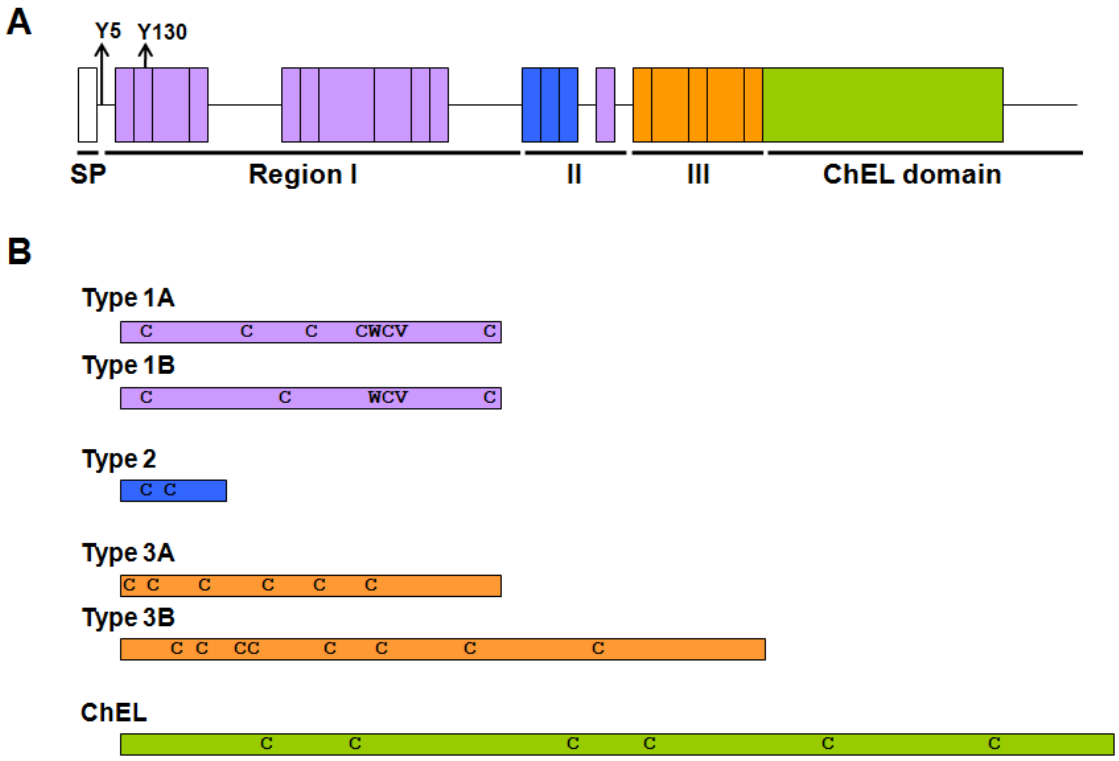


**Figure 1-2. The structure of thyroglobulin. (A)** Protein structure of thyroglobulin (Tg). Region I consists of 10 type 1 repeats and is followed by region II having three consecutive type 2 repeats and 11th type 1 repeat. Region III has five type 3 repeats. The cholinesterase-like domain (ChEL domain) is at C-terminal end. Two major tyrosines (Tyr5 and Tyr130) to form thyroid hormone are shown in N-terminal area of Tg. SP – signal peptide. **(B)** Structure of cysteine-rich repeats (composing region I, II, and III) and ChEL domain. Type 1 repeats can be further categorized into type 1A (6 cysteines) and type 1B (4 cysteines). Type 2 repeats have 2 cysteines in each. Type 3 repeats have two kinds (type 3A with 6 cysteines and type3B with 8 cysteines). Region III has 3 type 3A and 2 type 3B in a way of A-B-A-B-A. The ChEL domain has 6 conserved cysteines with acetylcholinesterase (Figure 1-3). Cysteines in the repeat and ChEL domain are believed to form intrachain disulfide bonds within each repeat unit and domain. Total 122 cysteines are in Tg and believed to form a maximum of 61 intrachain disulfide bonds. No cysteine is involved to form interchain disulfide bond after Tg forms its native structure.

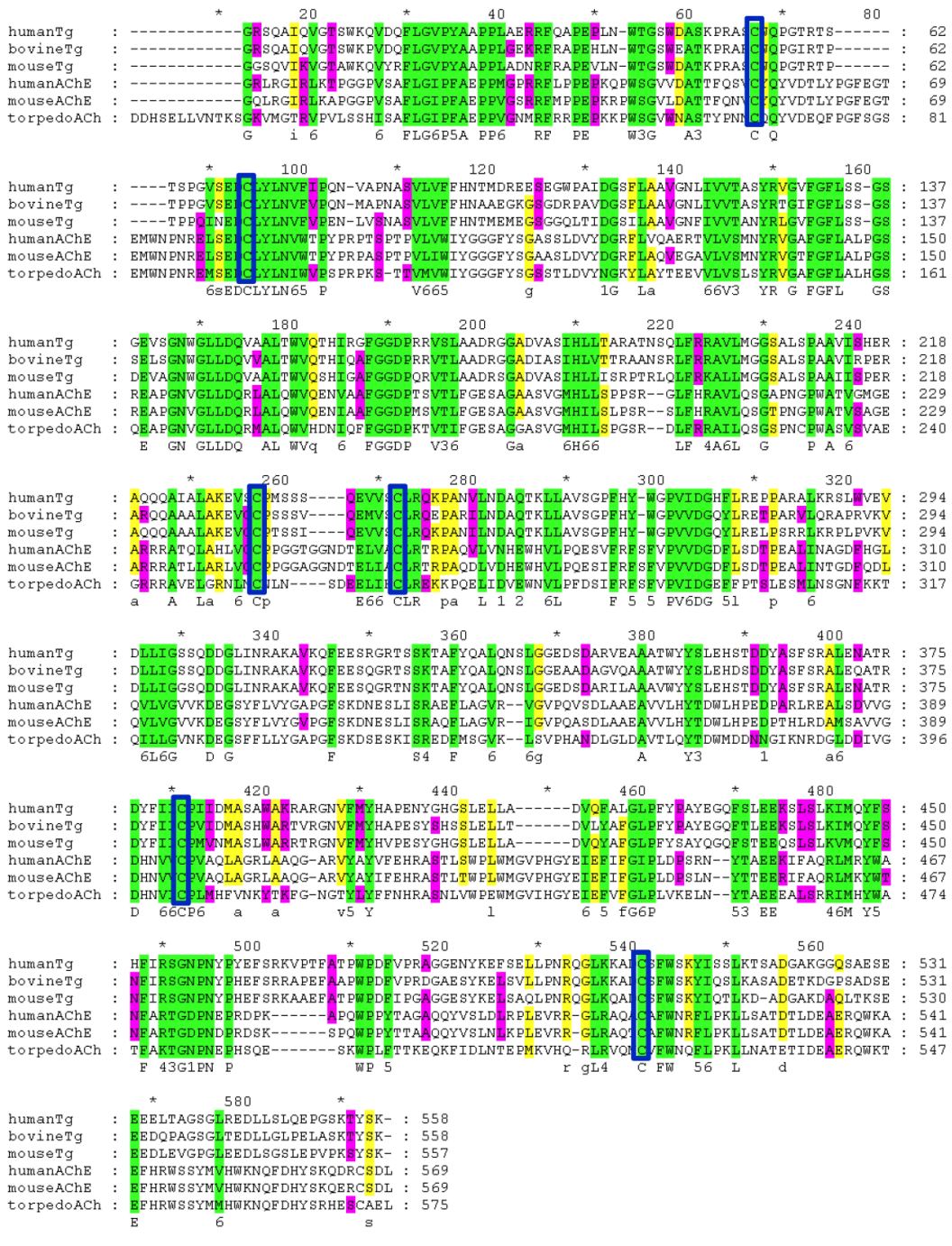


**Figure 1-3. Protein alignment between ChEL domains of Tg and acetylcholinesterases (AChE) from various of species.** Alignment was performed using clustal X software (version 2.0.3) [73]. Colors indicate different homology among proteins, green (represents 100% identity), yellow (over 80%, lower 100%), and purple (over 60%, lower 80%). Boxes indicate conserved six cysteines in AChE and ChEL domain of Tg. Accession numbers used for alignment are P01266 (human Tg), CAA26584 (bovine Tg), NP_033401 (mouse Tg), P22303 (human AChE), NP_033729 (mouse AChE), and CAA27169 (Torpedo AChE).

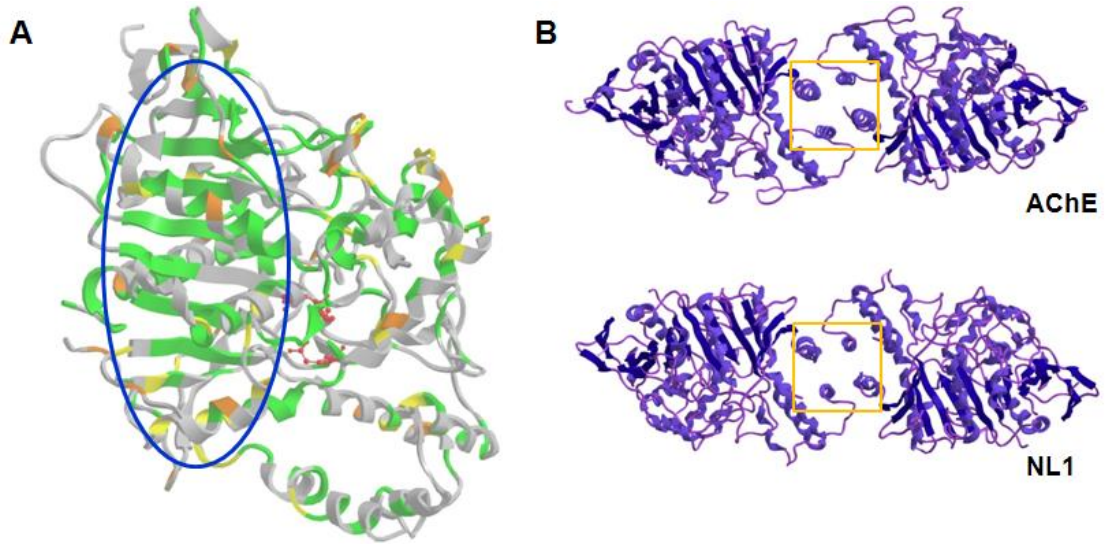


**Figure 1-4. The structure of cholinesterase-like proteins. (A)** Three dimensional structure of monomer of human AChE (PDB ID: 1B41). Coloring was done based on homology in figure 3 using Chemoffice 2006 (Cambridge Soft). Amino acid residues shown as ball and stick with red color are catalytic triad for AChE and none of them are conserved in ChEL domain of Tg. Note that high homology with ChEL of Tg is shown in the backbone structure (circled) in which most are colored with green (represents 100% identity, figure 3). **(B)** The homodimer structure of cholinesterase proteins. The molecule on the top is a dimer structure of mouse AChE (PDB ID: 1J07) and the one on the bottom is from mouse neuroligin 1 (NL1, PDB ID: 3B3Q). Other neuroligins (neuroligin 2 and 4) form a homodimer in similar way with NL1 [7-10]. Note that four-helix bundle (boxed), an area to form a homodimer, is shown in both AChE dimer and NL dimer.

**Table 1-1. List of human thyroglobulin gene mutations.** This list is updated from previously described [56] by adding recent reports [17, 57].

| Exon/intron position | Nucleotide position | Amino acid position |
|---|---|---|
| Exon 2 | c.113G>A | p.R19K |
| Intron 3 | g.IVS3-3C>G | Skipping of exon 4 |
| Exon 5 | c548G>A | p.C164Y |
| Exon 5 | c580T>G | p.C175G |
| intron 5 | g.IVS5+1G>A | Skipping of exon 5 |
| Exon 7 | c.759-760insA | p.L234fsX237 |
| Exon 7 | c.886C>T | p.R277X |
| Exon 8 | c.986A>C | p.Q310P |
| Exon 9 | c.1143delC | p.G362fsX382 |
| Exon 9 | c.2131C>T | p.Q692X |
| Exon 10 | c.2610G>T | p.Q851H |
| Intron 10 | g.IVS10-1G>A | Skipping of exon 11 |
| Exon 11 | c.2969G>A | p.S971I |
| Exon 12 | c.3022C>T | p.R989C |
| Exon 12 | c.3035C>T | p.P993L |
| Exon 14 | c.3229T>C | p.C1058R |
| Exon 17 | c.3790T>C | p.C1245R |
| Exon 20 | c.4310G>A | p.W1418X |
| Exon 21 | c.4397G>A | p.S1447N |
| Exon 22 | c.4537delG | p.D1494fsX1547 |
| Exon 22 | c.4588C>T | p.R1511X or Skipping of exon 22 |
| Exon 24 | c.4820G>T | p.C1588F |
| Intron 24 | g.IVS24+1G>C | Skipping of exon 24 |
| Intron 30 | g.IVS30+1G>T | Skipping of exon 30 |
| Intron 30 | g.IVS30+1G>A | Skipping of exon 30 |
| Exon 31 | c.5690G>A | p.C1878Y |
| Exon 31 | c.5791A>G | p.I1912V |
| Exon 33 | c.5986T>A | p.C1977S |
| Exon 33 | c.6017G>A | p.C1987Y |
| Intron 34 | g.IVS34-1G>C | Skipping of exon 35 |
| Exon 37 | c.6461G>A | p.C2135Y |
| Exon 38 | c.6701C>A | p.A2215D |
| Exon 38 | c.6725G>A | p.R2223H |
| Exon 40 | c.6956G>A | p.G2300D |
| Exon 40 | c.7007G>A | p.R2317Q |
| Exon 41 | c.7121G>T | p.G2355V |
| Exon 41 | c.7123G>A | p.G2356R |
| Intron 45 | g.IVS45+2T>A | Skipping of exon 45 |
| Exon 46 | c.7969C>T | p.Q2638X |

**References for chapter 1**


1. Yen, P.M. (2001). Physiological and molecular basis of thyroid hormone action. Physiol. Rev. *81,* 1097-1142.

2. Carrasco, N. (2005). Thyroid hormone synthesis - thyroid iodine transport. In The Thyroid, 9th Edition, L.E. Braverman and R.D. Utiger, eds. (Philadelphia, Pennsylvania, USA: Lippincott Williams & Wilkins), pp. 37-52.

3. Mercken, L., Simons, M.J., Swillens, S., Massaer, M., and Vassart, G. (1985). Primary structure of bovine thyroglobulin deduced from the sequence of its 8,431-base complementary DNA. Nature *316,* 647-651.

4. Gentile, F., and Salvatore, G. (1993). Preferential sites of proteolytic cleavage of bovine, human and rat thyroglobulin. The use of limited proteolysis to detect solvent-exposed regions of the primary structure. Eur. J. Biochem. *218,* 603-621.

5. Mori, N., Itoh, N., and Salvaterra, P.M. (1987). Evolutionary origin of cholinergic macromolecules and thyroglobulin. Proc. Natl. Acad. Sci. U. S. A. *84,* 2813-2817.

6. Swillens, S., Ludgate, M., Mercken, L., Dumont, J.E., and Vassart, G. (1986). Analysis of sequence and structure homologies between thyroglobulin and acetylcholinesterase: possible functional and clinical significance. Biochem. Biophys. Res. Commun. *137,* 142-148.

7. Arac, D., Boucard, A.A., Ozkan, E., Strop, P., Newell, E., Sudhof, T.C., and Brunger, A.T. (2007). Structures of neuroligin-1 and the neuroligin-1/neurexin-1 beta complex reveal specific protein-protein and protein-Ca2+ interactions. Neuron *56,* 992-1003.

8. Chen, X., Liu, H., Shim, A.H., Focia, P.J., and He, X. (2008). Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions. Nat. Struct. Mol. Biol. *15,* 50-56.

9. Fabrichny, I.P., Leone, P., Sulzenbacher, G., Comoletti, D., Miller, M.T., Taylor, P., Bourne, Y., and Marchot, P. (2007). Structural analysis of the synaptic protein neuroligin and its beta-neurexin complex: determinants for folding and cell adhesion. Neuron *56,* 979-991.

10. Koehnke, J., Jin, X., Budreck, E.C., Posy, S., Scheiffele, P., Honig, B., and Shapiro, L. (2008). Crystal structure of the extracellular cholinesterase-like domain from neuroligin-2. Proc. Natl. Acad. Sci. U. S. A. *105*, 1873-1878.

11. Dean, C., Scholl, F.G., Choih, J., DeMaria, S., Berger, J., Isacoff, E., and Scheiffele, P. (2003). Neurexin mediates the assembly of presynaptic terminals. Nat. Neurosci. *6*, 708-716.

12. Bourne, Y., Taylor, P., and Marchot, P. (1995). Acetylcholinesterase inhibition by fasciculin: crystal structure of the complex. Cell *83*, 503-512.

13. Paraoanu, L.E., and Layer, P.G. (2008). Acetylcholinesterase in cell adhesion, neurite growth and network formation. FEBS J. *275*, 618-624.

14. Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. Science *253*, 872-879.

15. Morel, N., Leroy, J., Ayon, A., Massoulie, J., and Bon, S. (2001). Acetylcholinesterase H and T dimers are associated through the same contact. Mutations at this interface interfere with the C-terminal T peptide, inducing degradation rather than secretion. J. Biol. Chem. *276*, 37379-37389.

16. Park, Y.N., and Arvan, P. (2004). The acetylcholinesterase homology region is essential for normal conformational maturation and secretion of thyroglobulin. J. Biol. Chem. *279*, 17085-17089.

17. Kim, P.S., Lee, J., Jongsamak, P., Menon, S., Li, B., Hossain, S.A., Bae, J.H., Panijpan, B., and Arvan, P. (2008). Defective protein folding and intracellular retention of thyroglobulin-R19K mutant as a cause of human congenital goiter. Mol. Endocrinol. *22*, 477-484.

18. Veenboer, G.J., and de Vijlder, J.J. (1993). Molecular basis of the thyroglobulin synthesis defect in Dutch goats. Endocrinology *132*, 377-381.

19. Yang, S.X., Pollock, H.G., and Rawitch, A.B. (1996). Glycosylation in human thyroglobulin: location of the N-linked oligosaccharide units and comparison with bovine thyroglobulin. Arch. Biochem. Biophys. *327*, 61-70.

20. Rawitch, A.B., Pollock, H.G., and Yang, S.X. (1993). Thyroglobulin glycosylation: location and nature of the N-linked oligosaccharide units in bovine thyroglobulin. Arch. Biochem. Biophys. *300*, 271-279.

21. Franc, J.L., Mallet, B., Lanet, J., and Giraud, A. (1994). The number of oligosaccharides borne by porcine thyroglobulin is variable. Endocrinology *134*, 885-890.

22. Kim, P.S., and Arvan, P. (1995). Calnexin and BiP act as sequential molecular chaperones during thyroglobulin folding in the endoplasmic reticulum. J. Cell Biol. *128*, 29-38.

23. Di Jeso, B., Ulianich, L., Pacifico, F., Leonardi, A., Vito, P., Consiglio, E., Formisano, S., and Arvan, P. (2003). Folding of thyroglobulin in the calnexin/calreticulin pathway and its alteration by loss of Ca2+ from the endoplasmic reticulum. Biochem. J. *370*, 449-458.

24. Di Jeso, B., Park, Y.N., Ulianich, L., Treglia, A.S., Urbanas, M.L., High, S., and Arvan, P. (2005). Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase. Mol. Cell. Biol. *25*, 9793-9805.

25. Kim, P.S., Bole, D., and Arvan, P. (1992). Transient aggregation of nascent thyroglobulin in the endoplasmic reticulum: relationship to the molecular chaperone, BiP. J. Cell Biol. *118*, 541-549.

26. Muresan, Z., and Arvan, P. (1997). Thyroglobulin transport along the secretory pathway. Investigation of the role of molecular chaperone, GRP94, in protein export from the endoplasmic reticulum. J. Biol. Chem. *272*, 26095-26102.

27. Kuznetsov, G., Chen, L.B., and Nigam, S.K. (1997). Multiple molecular chaperones complex with misfolded large oligomeric glycoproteins in the endoplasmic reticulum. J. Biol. Chem. *272*, 3057-3063.

28. Molinari, M., Calanca, V., Galli, C., Lucca, P., and Paganetti, P. (2003). Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. Science *299*, 1397-1400.

29. Oda, Y., Hosokawa, N., Wada, I., and Nagata, K. (2003). EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. Science *299*, 1394-1397.

30. Menon, S., Lee, J., Abplanalp, W.A., Yoo, S.E., Agui, T., Furudate, S., Kim, P.S., and Arvan, P. (2007). Oxidoreductase interactions include a role for ERp72 engagement with mutant thyroglobulin from the rdw/rdw rat dwarf. J. Biol. Chem. *282*, 6183-6191.

31. Tokunaga, F., Brostrom, C., Koide, T., and Arvan, P. (2000). Endoplasmic reticulum (ER)-associated degradation of misfolded N-linked glycoproteins is suppressed upon inhibition of ER mannosidase I. J. Biol. Chem. *275*, 40757-40764.

32. Wiertz, E.J., Tortorella, D., Bogyo, M., Yu, J., Mothes, W., Jones, T.R., Rapoport, T.A., and Ploegh, H.L. (1996). Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. Nature *384*, 432-438.

33. Ye, Y., Shibata, Y., Yun, C., Ron, D., and Rapoport, T.A. (2004). A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. Nature *429*, 841-847.

34. Schelhaas, M., Malmstrom, J., Pelkmans, L., Haugstetter, J., Ellgaard, L., Grunewald, K., and Helenius, A. (2007). Simian Virus 40 depends on ER protein folding and quality control factors for entry into host cells. Cell *131*, 516-529.

35. Okuda-Shimizu, Y., and Hendershot, L.M. (2007). Characterization of an ERAD pathway for nonglycosylated BiP substrates, which require Herp. Mol. Cell *28*, 544-554.

36. Tsai, B., Rodighiero, C., Lencer, W.I., and Rapoport, T.A. (2001). Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. Cell *104*, 937-948.

37. Gillece, P., Luz, J.M., Lennarz, W.J., de La Cruz, F.J., and Romisch, K. (1999). Export of a cysteine-free misfolded secretory protein from the endoplasmic reticulum for degradation requires interaction with protein disulfide isomerase. J. Cell Biol. *147*, 1443-1456.

38. Ushioda, R., Hoseki, J., Araki, K., Jansen, G., Thomas, D.Y., and Nagata, K. (2008). ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. Science *321*, 569-572.

39. Kim, P.S., Kwon, O.Y., and Arvan, P. (1996). An endoplasmic reticulum storage disease causing congenital goiter with hypothyroidism. J. Cell Biol. *133*, 517-527.

40. Forster, M.L., Sivick, K., Park, Y.N., Arvan, P., Lencer, W.I., and Tsai, B. (2006). Protein disulfide isomerase-like proteins play opposing roles during retrotranslocation. J. Cell Biol. *173*, 853-859.

41. Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat. Rev. Mol. Cell Biol. *8*, 519-529.

42. Ron, D., and Hubbard, S.R. (2008). How IRE1 reacts to ER stress. Cell *132*, 24-26.

43. Bertolotti, A., Zhang, Y., Hendershot, L.M., Harding, H.P., and Ron, D. (2000). Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nat. Cell Biol. *2*, 326-332.

44. Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., and Walter, P. (2000). Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell *101*, 249-258.

45. Breckenridge, D.G., Germain, M., Mathai, J.P., Nguyen, M., and Shore, G.C. (2003). Regulation of apoptosis by endoplasmic reticulum pathways. Oncogene *22*, 8608-8618.

46. Urano, F., Wang, X., Bertolotti, A., Zhang, Y., Chung, P., Harding, H.P., and Ron, D. (2000). Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. Science *287*, 664-666.

47. Sargsyan, E., Baryshev, M., and Mkrtchian, S. (2004). The physiological unfolded protein response in the thyroid epithelial cells. Biochem. Biophys. Res. Commun. *322*, 570-576.

48. Kuznetsov, G., Bush, K.T., Zhang, P.L., and Nigam, S.K. (1996). Perturbations in maturation of secretory proteins and their association with endoplasmic reticulum chaperones in a cell culture model for epithelial ischemia. Proc. Natl. Acad. Sci. U. S. A. *93*, 8584-8589.

49. Baryshev, M., Sargsyan, E., Wallin, G., Lejnieks, A., Furudate, S., Hishinuma, A., and Mkrtchian, S. (2004). Unfolded protein response is involved in the pathology of human congenital hypothyroid goiter and rat non-goitrous congenital hypothyroidism. J. Mol. Endocrinol. *32*, 903-920.

50. Leonardi, A., Vito, P., Mauro, C., Pacifico, F., Ulianich, L., Consiglio, E., Formisano, S., and Di Jeso, B. (2002). Endoplasmic reticulum stress causes thyroglobulin retention in this organelle and triggers activation of nuclear factor-kappa B via tumor necrosis factor receptor-associated factor 2. Endocrinology *143*, 2169-2177.

51. Roberts, C.G., and Ladenson, P.W. (2004). Hypothyroidism. Lancet *363*, 793-803.

52. Gillam, M.P., and Kopp, P. (2001). Genetic regulation of thyroid development. Curr. Opin. Pediatr. *13*, 358-363.

53. Gillam, M.P., and Kopp, P. (2001). Genetic defects in thyroid hormone synthesis. Curr. Opin. Pediatr. *13*, 364-372.

54. Baas, F., van Ommen, G.J., Bikker, H., Arnberg, A.C., and de Vijlder, J.J. (1986). The human thyroglobulin gene is over 300 kb long and contains introns of up to 64 kb. Nucleic Acids Res. *14*, 5171-5186.

55. Mendive, F.M., Rivolta, C.M., Moya, C.M., Vassart, G., and Targovnik, H.M. (2001). Genomic organization of the human thyroglobulin gene: the complete intron-exon structure. Eur. J. Endocrinol. *145*, 485-496.

56. Rivolta, C.M., and Targovnik, H.M. (2006). Molecular advances in thyroglobulin disorders. Clin. Chim. Acta *374*, 8-24.

57. Caputo, M., Rivolta, C.M., Esperante, S.A., Gruneiro-Papendieck, L., Chiesa, A., Pellizas, C.G., Gonzalez-Sarmiento, R., and Targovnik, H.M. (2007). Congenital hypothyroidism with goitre caused by new mutations in the thyroglobulin gene. Clin. Endocrinol. (Oxf). *67*, 351-357.

58. Matakidou, A., Hamel, N., Popat, S., Henderson, K., Kantemiroff, T., Harmer, C., Clarke, S.E., Houlston, R.S., and Foulkes, W.D. (2004). Risk of non-medullary thyroid cancer influenced by polymorphic variation in the thyroglobulin gene. Carcinogenesis *25*, 369-373.

59. Hishinuma, A., Fukata, S., Kakudo, K., Murata, Y., and Ieiri, T. (2005). High incidence of thyroid cancer in long-standing goiters with thyroglobulin mutations. Thyroid *15*, 1079-1084.

60. Ricketts, M.H., Simons, M.J., Parma, J., Mercken, L., Dong, Q., and Vassart, G. (1987). A nonsense mutation causes hereditary goitre in the Afrikander cattle and unmasks alternative splicing of thyroglobulin transcripts. Proc. Natl. Acad. Sci. U. S. A. *84*, 3181-3184.

61. Kim, P.S., Hossain, S.A., Park, Y.N., Lee, I., Yoo, S.E., and Arvan, P. (1998). A single amino acid change in the acetylcholinesterase-like domain of thyroglobulin causes congenital goiter with hypothyroidism in the cog/cog mouse: a model of human endoplasmic reticulum storage diseases. Proc. Natl. Acad. Sci. U. S. A. *95*, 9909-9913.

62. Hishinuma, A., Furudate, S., Oh-Ishi, M., Nagakubo, N., Namatame, T., and Ieiri, T. (2000). A novel missense mutation (G2320R) in thyroglobulin causes hypothyroidism in rdw rats. Endocrinology *141*, 4050-4055.

63. Kim, P.S., Ding, M., Menon, S., Jung, C.G., Cheng, J.M., Miyamoto, T., Li, B., Furudate, S., and Agui, T. (2000). A missense mutation G2320R in the thyroglobulin gene causes non-goitrous congenital primary hypothyroidism in the WIC-rdw rat. Mol. Endocrinol. *14*, 1944-1953.

64. Dayan, C.M., and Daniels, G.H. (1996). Chronic Autoimmune Thyroiditis. N. Engl. J. Med. *335*, 99-107.

65. Paris, M., Escriva, H., Schubert, M., Brunet, F., Brtko, J., Ciesielski, F., Roecklin, D., Vivat-Hannah, V., Jamin, E.L., Cravedi, J.P., et al. (2008). Amphioxus postembryonic development reveals the homology of chordate metamorphosis. Curr. Biol. *18*, 825-830.

66. Suzuki, S., and Kondo, Y. (1971). Demonstration of thyroglobulin-like lodinated proteins in the branchial sac of tunicates. Gen. Comp. Endocrinol. *17*, 402-406.

67. Monaco, F., Dominici, R., Andreoli, M., De Pirro, R., and Roche, J. (1981). Thyroid hormone formation in thyroglobulin synthesized in the Amphioxus (Branchiostoma lanceolatum pallas). Comp. Biochem. Physiol. B. Biochem. Mol. Biol. *70*, 341-343.

68. Christophe, D., and Vassart, G. (1990). The thyroglobulin gene evolutionary and regulatory issues. Trends Endocrinol. Metab. *1*, 351-356.

69. Wolff, J. (1964). Transport of Iodide and Other Anions in the Thyroid Gland. Physiol. Rev. *44*, 45-90.

70. Brisson, A., Marchelidon, J., and Lachiver, F. (1974). Comparative studies on the amino acid composition of thyroglobulins from various lower and higher vertebrates: phylogenetic aspect. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. *49*, 51-63.

71. Monaco, F., Andreoli, M., La Posta, A., and Roche, J. (1978). Thyroglobulin biosynthesis in a larval (ammocoete) and adult freshwater lamprey (Lampetra planeri B1.). Comp. Biochem. Physiol. B. Biochem. Mol. Biol. *60*, 87-91.

72. Suzuki, S., Gorbman, A., Rolland, M., Montfort, M.-F., and Lissitzky, S. (1975). Thyroglobulins of cyclostomes and an elasmobranch. Gen. Comp. Endocrinol. *26*, 59-69.

73. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876-4882.

# CHAPTER 2

## THE CHOLINESTERASE-LIKE DOMAIN OF THYROGLOBULIN FUNCTIONS AS AN INTRAMOLECULAR CHAPERONE


### Abstract

Thyroid hormonogenesis requires secretion of thyroglobulin, a protein comprising Cys-rich regions I, II, and III (referred to collectively as region I-II-III) followed by a cholinesterase-like (ChEL) domain. Secretion of mature thyroglobulin requires extensive folding and glycosylation in the ER. Multiple reports have linked mutations in the ChEL domain to congenital hypothyroidism in humans and rodents; these mutations block thyroglobulin from exiting the ER and induce ER stress. We report that, in a cell-based system, mutations in the ChEL domain impaired folding of thyroglobulin region I-II-III. Truncated thyroglobulin devoid of the ChEL domain was incompetent for cellular export; however, a recombinant ChEL protein ("secretory ChEL") was secreted efficiently. Coexpression of secretory ChEL with truncated thyroglobulin increased intracellular folding, promoted oxidative maturation, and facilitated secretion of region I-II-III, indicating that the ChEL domain may function as an intramolecular chaperone. Additionally, we found that the I-II-III peptide was cosecreted and physically associated with secretory ChEL. A functional ChEL domain engineered to be retained intracellularly triggered oxidative maturation of I-II-III but coretained I-II-III, indicating that the ChEL domain may also function as a molecular escort. These insights into the role of the ChEL

domain may represent potential therapeutic targets in the treatment of congenital hypothyroidism.

## Introduction

Thyroid hormone is essential for development and oxidative metabolism in vertebrates. The synthesis of thyroid hormones involves secretion of thyroglobulin (Tg) to the apical extracellular lumen of follicles that serve as the basic functional unit of the thyroid gland. Once secreted, Tg is iodinated via the activity of a thyroid-specific peroxidase to form thyroxine (the major thyroid hormone) involving specific iodinated tyrosyl residues within the Tg polypeptide backbone [1]. No other endogenous proteins of the thyroid are capable of serving as precursors for thyroid hormone synthesis, rendering vertebrate organisms completely dependent on structural information encoded in the Tg molecule — both for directing iodotyrosyl coupling for thyroxine synthesis [2] and for passing ER quality control in the secretory pathway [3]. Cumulatively, steps leading to Tg export from the ER are rate limiting in its overall secretion [4]. Peak secretion of Tg, as measured in cell culture, occurs on the order of 3 hours after synthesis [5].

The Tg cDNA initially encodes 3 regions referred to collectively as region I-II-III, that span 80% of the full-length Tg monomer (2,746 amino acids in mouse Tg), with each region involving multiple Cys-rich repeat domains covalently “stapled” by intradomain disulfide bonds [6]; while the final approximately 500 amino acids of Tg are strongly homologous to acetylcholinesterase [7-9]. The most important site of thyroxine synthesis involves but a small portion of the molecule at the extreme amino terminus of the Tg

protein [10, 11], raising questions about the role of the remaining C-terminal 2,250 amino acids [12]. Indeed, the overall structure of Tg remains unknown; nevertheless, it is clear that along its length, Tg undergoes extensive N-linked glycosylation and monomer folding in the ER [13].

Congenital hypothyroidism with defective Tg has been reported in species ranging from humans to rats and mice. In all 3 species, the cellular phenotype described includes a dilated ER with induction of ER molecular chaperones [14], activation of ER stress signaling pathways [15], and ER-associated degradation (ERAD) of most of the mutant Tg gene product [16]. Of the cases (and animal models) published to date, the cholinesterase-like (ChEL) domain is a very commonly affected site, involving mutations including the newly described A2215D [17, 18]; R2223H [19]; G2300D,R2317Q [20]; G2355V,G2356R; or the skipping of exon 45 (normally encoding 36 residues of the ChEL domain); as well as Q2638Z mutants (where Z represents stop) [21]. In addition are polymorphisms including P2213L and W2482R; and R2511Q that may be associated with an increased incidence of nonmedullary thyroid cancer [22]. Congenital hypothyroidism with defective Tg is caused by the G2300R point mutation localized to the ChEL domain in the *rdw*/*rdw* rat dwarf [23, 24], while we identified L2263P as responsible for the hypothyroidism of the *cog*/*cog* congenital goiter mouse [25]. Indeed, a truncated Tg comprising only region I-II-III and lacking the ChEL domain is blocked within the ER [26]. Together, these findings emphasize that in full-length Tg, a functional ChEL domain is needed to allow Tg to be efficiently transported to the site of thyroid hormone synthesis, in particular, for exit from the ER.

We show in this study, by contrast, that attaching a signal peptide to the ChEL domain is itself sufficient for rapid and efficient intracellular transport of that protein ("secretory ChEL") to the extracellular space. Moreover, by physical association with Tg I-II-III, secretory ChEL rescues *in trans* maximal protein expression, disulfide maturation, and secretion of Tg I-II-III. Indeed, intentional retention of the functional ChEL domain within the ER, while assisting in Tg I-II-III protein expression and disulfide maturation, prevents its intracellular transport. These data point to both a molecular chaperone and a molecular escort function of the ChEL domain in Tg transport for thyroid hormone synthesis.

## Materials and Methods

### Materials

AMS (4-acetamido-4'-maleimidylstilbene-2,2'- disulfonic acid), Lipofectamine 2000, fetal bovine serum, penicillin, and streptomycin were from Invitrogen; Complete protease inhibitor cocktail was from Roche; Brefeldin A, cycloheximide, N-ethylmaleimide, protein G-Agarose, and protein A-Agarose from Sigma Chemical Co.; Endoglycosidase H and Peptide-N-Glycosidase F from New England Biolabs; Trans$^{35}$S-Label from MP Biomedicals; Zysorbin from Zymed Labs; *Trans*IT-LT1 transfection reagent from Mirus. Rabbit polyclonal anti-myc was from Immunology Consultants, Inc. Rabbit polyclonal anti-Tg (containing antibodies against epitopes at both N- and C-terminal regions of the protein) have been previously described [14].

### Site-directed mutagenesis of mouse Tg cDNA

ChEL domain mutations were introduced with the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using the following mutagenic primers: ChEL-KDEL (5'–GCCTGTCCCCAAGAGCTACAGCAAAGACGAGCTATAGGCCGCTTCCCTTTAGTGAGGG –3' and 5'–CCCTCACTAAAGGGAAGCGGCCTATAGCTCGTCTTTGCTGTAGCTCTTGGGGACAGGC–3'); ChEL-myc (5'–CAAGAGCTACAGCAAAGAACAGAAACTGATCTCTGAGGAGGACTTATGATTAATGCTTCG –3' and 5'–CGAAGCATTAATCATAAGTCCTCCTCAGAGATCAGTTTCTGTTCTTTGCTGTAGCTCTTG–3'). Truncated Tg region I-II-III was made by introducing a stop codon at L2169 (5'–GGAAGTCTGGAATCCCTTAGGTCCAATCTGATGTAACATCC–3' and 5'–GGATGTTACATCAGATTGGACCTAAGGGATTCCAGACTTCC–3') within Tg subcloned into pCMS (BD Bioscience – Clontech). To make the secretory ChEL domain, a Sal I restriction site was introduced into 5'-end of ChEL by using a pair of mutagenesis primers (5'–CCTACATCTACCGGAAGTCTGGTCGACCTTTGGTCCAATCTGATGTAACATCC–3' and 5'–GGATGTTACATCAGATTGGACCAAAGGTCGACCAGACTTCCGGTAGATGTAGG–3') with subcloning into the Sal I and Not I sites of pCMS. Each mutation was confirmed by direct DNA sequencing before expression in 293 cells. The prolactin signal peptide in pSPPS1 (kind gift of Dr. V. Lingappa, U. California San Francisco) was fused in frame at the 5'-end of the ChEL cDNA using Nhe I and Sal I sites.

**Cell culture and transfection**

293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum in 6 well plates at 37ºC in a humidified 5% CO2 incubator. Plasmids were transiently transfected using *Trans*IT-LT1 or Lipofectamine 2000 transfection reagent according to the manufacturer's instructions.

**Metabolic labeling and immunoprecipitation**

Transfected 293 cells were starved for 30 min in met/cys-free DMEM, then pulse labeled with 180 μCi/ml $^{35}$S-amino acids for indicated time. The cells were then washed with an excess of cold Met/Cys and chased in complete DMEM. At each time point, cells were lysed in buffer containing 1% NP-40, 0.1% SDS, 0.1 M NaCl, 2 mM EDTA, 25 mM Tris pH 7.4 and protease inhibitor cocktail. Where indicated, 20 mM N-ethylmaleimide was also added to the lysis buffer. For immunoprecipitation, anti-Tg was incubated with cell or media samples overnight at 4ºC, and the immunoprecipitate was recovered with protein A-Agarose. For co-immunoprecipitation studies, samples were incubated overnight at 4ºC with anti-myc antibodies and protein A-Agarose. Immunoprecipitates (or co-precipitates) were washed three times before boiling in SDS sample buffer with or without reducing agent, resolved by SDS-PAGE, and analyzed by fluorography or phosphorimaging.

**Alkylation of Tg cysteine thiols**

Immunoprecipitated Tg was resuspended and incubated in a buffer containing 2% SDS, 50 mM Tris pH7.4 with or without 5 mM AMS for 1 h at 30ºC. The reaction was stopped by boiling in SDS sample buffer plus 0.1 M DTT before analysis by SDS-PAGE.

**Endoglycosidase H or peptide-N-glycosidase digestion**

Immunoprecipitates were boiled for 10 min in denaturing solution with 0.5% SDS containing 1% 2-mercaptoethanol (for Peptide-N-Glycosidase digestion, solution with 0.5% SDS without reducing agent was used) and digested with 250 U of endoglycosidase H in 50 mM sodium citrate, pH 5.5 or 250 U of Peptide-N-Glycosidase in 50 mM sodium phosphate, pH 7.5 with 1% NP-40 for 1 h at 37ºC.

**Results**

**Maturation of recombinant Tg**

During folding, Tg must bury 122 free Cys thiols into intrachain disulfide bonds. Reactivity with 4-acetamido-4'-maleimidylstilbene-2,2'-disulphonic acid (AMS) irreversibly alkylates available Cys thiols, increasing by 0.5 kDa the protein molecular mass per alkylated Cys [27], which can ultimately be detected as a molecular mass shift by SDS-PAGE under reducing conditions. Upon denaturation, Tg secreted from cells (which had already passed ER quality control and received Golgi carbohydrate modifications) showed no detectable reactivity with 5 mM AMS (Figure 2-1, compare lanes 3 and 4). By contrast, newly synthesized intracellular Tg residing primarily in the ER [5] showed substantial AMS reactivity (compare lanes 5 and 6). Cells treated with brefeldin A (BFA; 5 μg/ml) accumulated in the ER newly synthesized Tg that, in terms of

AMS reactivity, seemed equivalent to the sum of secreted and intracellular Tg obtained from untreated cells, revealing the ability to discriminate between distinct folded states. In the case of the *rdw* mutant bearing the G2300R mutation within the ChEL domain of Tg, none of the Tg molecules in the population appeared able to properly mature, as judged by fact that they buried their Cys thiols (Figure 2-1). By 4% SDS-PAGE, a mobility shift of this magnitude (above that of the 330-kDa Tg band) suggests a molecular mass addition of 10–30 kDa, which might potentially represent modification of as many as 20–60 of the first 116 Cys residues of Tg. By contrast, the ChEL domain bearing the mutation has only 6 Cys residues – even modification of all of these residues would amount to a net molecular mass addition of only 3 kDa, which would not be expected to be seen as a significant mobility shift of the Tg band. Thus, these data imply that a mutation within the ChEL domain may impact on the ability to efficiently form disulfide bonds within upstream Tg region I-II-III, which contain the majority of Tg Cys residues.

Recent studies have shown that for a brief period after synthesis, Tg newly synthesized by thyrocytes can be detected in mixed disulfides with a number of ER oxidoreductases: adducts known as forms A, B, and C that are detected by nonreducing SDS-PAGE [28]. Herein, we used nonreducing SDS-PAGE to examine disulfide maturation of recombinant mouse Tg upon transient expression in 293 cells. For wild-type Tg, adducts in the A, B, and C region of the gel (like those reported for endogenous Tg; ref. [28]) could be seen, while an incompletely oxidized Tg monomer called band D was already the major species at the 0 chase time (Figure 2-2A, middle panel). By 1 hour of chase, the predominant Tg form was already a fully oxidized monomer called band E

(Figure 2-2A). Because secretion was ongoing, mature Tg was lost from these cells (data not shown). Under conditions where all forms of Tg were retained in the ER by treatment of cells with BFA, a quantitative profile of oxidative Tg folding intermediates could be observed (BFA; Figure 2-2A, right).

Using PC Cl3 thyrocytes expressing endogenous Tg, we found that the newly identified E isoform is essentially identical to that secreted to the medium. What confounds the nonreducing SDS-PAGE analysis of Tg band mobility is the presence of Golgi-modified N-linked glycans on Tg (Figure 2-3A, right panel). Thus, a comparison of band mobility upon nonreducing SDS-PAGE can be made only after treatment with PNGase F to remove all N-linked oligosaccharides (Figure 2-3A, left panel; a double-headed arrow establishes the identity of the E isoform in cells and media). Similar results were obtained for recombinant Tg expressed in 293 cells when examined without (Figure 2-3B) or with (Figure 2-3C) prior PNGase F digestion.

Tg monomer folding is temperature dependent [29], and Tg secretion is profoundly inhibited at both 20°C and 25°C. At 25°C, there was a slight increase in Tg adducts A, B, and C at the 0 chase time (Figure 2-2B, left). At 20°C there was additional prolongation of Tg adducts and further inhibition of Tg oxidative maturation from folding intermediate D to the mature E form (Figure 2-2B, right). With these results serving as positive controls for Tg folding, we proceeded to examine the oxidative maturation of both *cog* Tg and *rdw* Tg (which bear mutations in the ChEL domain). As shown in Figure 2C, these proteins, even at 2 hours after synthesis, appeared arrested as early folding intermediates, including both Tg adducts and the incompletely oxidized D isoform. While we do not yet know the specific Tg Cys residues involved in either adduct formation or

maturation from D to E forms, these data further imply that function of the ChEL domain is crucial to global folding of the Tg molecule.

**The isolated ChEL domain folds autonomously in the secretory pathway**

The Tg ChEL domain exhibits substantial primary sequence similarity with authentic acetylcholinesterase, a well-studied protein [30] that is fully competent for intracellular transport to the cell surface. We therefore attempted to examine secretory behavior of the isolated ChEL domain driven into the ER via an artificial (prolactin) signal peptide. Indeed, within 4 hours after synthesis, the recombinant secretory ChEL protein was nearly completely released from cells as an endoglycosidase H–resistant species (Figure 2-4, lane 6), indicating that it traversed the intracellular transport pathway as a normal secretory protein.

**The ChEL domain is an intramolecular chaperone for Tg region I-II-III**

We previously reported that Tg region I-II-III is secretion incompetent [26]. To confirm this point, we followed the secretion and sensitivity to endoglycosidase H digestion of newly synthesized I-II-III 4 hours after synthesis. Indeed, while full-length wild-type Tg was efficiently secreted, I-II-III remained intracellular (Figure 2-5A). Persistence of sensitivity to endoglycosidase H digestion (Figure 2-5B) indicated that I-II-III does not have sufficient information for efficient intracellular transport. This was also true even at extended chase times (Figure 2-5C, upper panel). These data support our previous hypothesis that the ChEL domain may be required for normal conformational maturation and intracellular transport of Tg [26].

To better understand the relationship of the I-II-III domains to the ChEL domain, we coexpressed these as separate proteins within the ER. Remarkably, *in trans*, presence of the secretory ChEL domain in the ER rescued the export of I-II-III. This secretion was progressive over time (Figure 2-5C, lower panels).

A likely implication of these findings is that I-II-III directly interacts with the ChEL domain. To examine this, we constructed a secretory ChEL with a carboxyterminal myc-epitope tag. Then either secretory ChEL or secretory ChEL-myc was coexpressed with I-II-III in 293 cells. After metabolic labeling, the chase medium bathing these cells was collected and evenly divided for immunoprecipitation with either anti-Tg or anti-myc (Figure 2-6A). Anti-Tg (lanes 1–3) directly immunoprecipitated both secretory ChEL (or ChEL-myc) as well as the rescued cosecreted I-II-III. Anti-myc (lanes 4–6) could not immunoprecipitate untagged ChEL and also failed to recover any I-II-III. However, anti-myc immunoprecipitated secretory ChEL-myc and quantitatively coprecipitated all of the I-II-III (Figure 2-6A, lane 6) that could be directly immunoprecipitated with polyclonal anti-Tg (lane 3).

These data strongly suggest that nearly 100% of secreted I-II-III associates with secretory ChEL within the ER and remains associated throughout the secretory pathway. Yet newly synthesized I-II-III is secreted with much slower kinetics than the cohort of labeled ChEL molecules synthesized at the same time (Figure 2-5C). I-II-III secretion must be dependent upon subsequently synthesized ChEL molecules, as post-pulse inhibition of further protein synthesis with cycloheximide (Figure 2-6B, lanes 3 and 4) blocked secretion of newly synthesized I-II-III even as it had no effect on the secretion of labeled ChEL. Thus, I-II-III is not ready to interact with ChEL when I-II-III is first made

but evidently needs preliminary folding to acquire competence for ChEL interaction. By contrast, ChEL export can proceed independently of I-II-III (Figure 2-4), decreasing its availability in the ER, which may limit efficiency of I-II-III transport (Figure 2-5C). To test this, we performed cotransfections of cDNAs encoding I-II-III and secretory ChEL at different plasmid ratios (with I-II-III cDNA and total DNA per transfection held constant). As shown in Figure 2-7A, increasing secretory ChEL cDNA increased the expression and secretion of the ChEL protein (bottom panel). Simultaneously, this also increased the quantity of I-II-III that was recovered intracellularly after a 6-hour chase, despite the fact that the transfection quantity of I-II-III cDNA was held constant (upper panel). As the initially synthesized I-II-III (0 chase time) was only marginally affected by coexpression of secretory ChEL (Figure 2-7B), it appears that most of the intracellular increase in I-II-III protein reflects stabilization after synthesis. Indeed, we determined that total I-II-III recovery was increased 100% (i.e., doubled) at 16 hours of chase when secretory ChEL was expressed compared with vector DNA alone (Figure 2-7B). Accompanying the increased recovery in I-II-III was an increase in the amount of I-II-III secretion (Figure 2-7A).

If the ChEL domain were acting as a molecular chaperone for I-II-III, we reasoned that secretory ChEL might be able to enhance oxidative folding of I-II-III. To test this, we designed a secretory ChEL-KDEL construct, in which the KDEL sequence retains the ChEL domain in the early secretory pathway (discussed below). In a pulse-chase format, presence of the ChEL-KDEL protein promoted disulfide maturation of I-II-III from an incompletely oxidized intermediate (reminiscent of intermediate D of full-length Tg; see Figure 2-2) to a more mature (E-like) form (Figure 2-8A, arrows). To

determine whether ChEL domain tertiary structure is required for oxidative rescue of I-II-III, we compared the ChEL-KDEL construct with other secretory ChEL constructs bearing *cog* or *rdw* mutations, which are also retained within the early secretory pathway. As shown in Figure 2-8B, unlike ChEL-KDEL, these mutations within the ChEL domain had little or no beneficial effect on oxidative maturation of I-II-III. By contrast, authentic secretory ChEL (without KDEL) promoted secretion of I-II-III protein that, after deglycosylation with PNGase F, exhibited oxidative maturation comparable to that of the mature E form seen intracellularly (Figure 2-9). Together, the data in Figure 2-8 and Figure 2-9 strongly suggest that a native ChEL domain acts like an intramolecular chaperone, facilitating oxidative maturation of upstream Tg domains.

**The ChEL domain is a molecular escort for Tg region I-II-III**

If the ChEL domain serves only as a true molecular chaperone within Tg, then we would expect that when expressed *in trans*, secretory ChEL (like other molecular chaperones) would show a binding preference for early I-II-III folding intermediates over fully mature molecules. To examine this, we coimmunoprecipitated intracellular I-II-III with secretory ChEL-myc in cells treated with BFA to block intracellular transport in the secretory pathway. In the absence of ChEL, I-II-III exhibited defective oxidative maturation (Figure 2-10, lanes 2–4), while in the presence of secretory ChEL-myc, anti-Tg directly and efficiently immunoprecipitated both immature D and mature E forms of I-II-III (lanes 5–7). The presence of ChEL appeared to increase recovery of I-II-III, consistent with results shown in Figure 7. By contrast, the immunoprecipitation of ChEL with anti-myc seemed inefficient at coprecipitation of the unfolded I-II-III isomer, as

compared with the efficient recovery of the faster-migrating, mature I-II-III (Figure 2-10, lanes 9 and 10). These data along with the persistent ChEL association with I-II-III even after secretion (Figure 2-6A) indicate that ChEL serves as more than an intramolecular chaperone for unfolded Tg molecules. If indeed the ChEL domain helps to physically convey I-II-III export from the ER, then if ChEL transport were to be disrupted, I-II-III transport would be disrupted in parallel. This proved to be the case, as unlike secretory ChEL, secretory ChEL-KDEL permitted secretion of I-II-III only to the extent that the ChEL-KDEL itself escaped from the cells (Figure 2-11, A and B), despite the fact that ChEL-KDEL facilitated the oxidative maturation of I-II-III (Figure 2-8). As predicted by results in Figure 8B, secretory ChEL bearing *cog* or *rdw* mutations also did not support I-II-III secretion (Figure 2-11B), just as these mutations do not permit secretion of intact Tg. Together, the data indicate that the ChEL domain functions as both an intramolecular chaperone and molecular escort for Tg maturation and transport in the secretory pathway.

**Discussion**

Tg has evolved as the unique protein precursor for thyroid hormone synthesis. Specially encoded structural features for thyroid hormone formation and iodide storage [2] are known to come into play once the Tg molecule has been delivered extracellularly, where it comes in contact with the thyroidal iodination machinery [31]. However, structural features encoding Tg transport in the secretory pathway – required for Tg to serve as a prohormone – have been more elusive. In spite of this, it appears that most congenital hypothyroidism with deficient Tg is caused by mutations in which intracellular Tg transport is blocked, rather than by defects in the ability to synthesize

thyroid hormone per se [1, 32]. Evolutionary pressure definitely exists for Tg to be maintained as a secreted protein, as its inability to do so (at least in Merino sheep) compromises viability to a stage where animals cannot reach reproductive age [33, 34]. We now present what we believe to be the first evidence directly implicating the ChEL domain in assisting the intracellular protein transport of Tg.

One way in which the ChEL domain may participate in Tg export is by facilitating Tg folding. We previously suspected that ChEL forms an independently folded domain within Tg, because its replacement by authentic acetylcholinesterase results in a chimera with active cholinesterase function [26]. Yet single point mutations in the ChEL region appear sufficient to prevent formation of disulfide bonds that are made in upstream Tg regions, as evidenced not only by remarkable reactivity with AMS (Figure 2-1), but also by direct assessment of global disulfide maturation (with *cog* and *rdw* mutants being trapped in partially oxidized folding intermediates; Figure 2-2). Using a cleavable signal peptide, we now find that the isolated ChEL domain is a very effective secretory protein in its own right (Figure 2-4), whereas I-II-III alone cannot escape ER retention (Figure 2-5, A and B). The simplest hypothesis (still to be tested) is that exposed chaperone-binding regions on Tg I-II-III confers its ER retention. Remarkably, our current evidence supports that the presence of the ChEL domain in the ER enriches for an oxidatively mature form of I-II-III (Figure 2-8), increases intracellular recovery of I-II-III (Figure 2-7 and Figure 2-10), and allows I-II-III to be secreted (Figure 2-5C) in a form that is both endoglycosidase H resistant (data not shown) and appropriately oxidized (Figure 2-9). Thus, we propose that as long as the ChEL domain is itself properly folded

(Figure 2-11B), it functions as an intramolecular chaperone within the context of full-length Tg, facilitating the efficiency of "on-pathway" protein folding, leading to ER exit.

However, this does not appear to be the whole story, because the ChEL domain is transported with Tg out of the ER. Surprisingly, this is not simply because the ChEL domain is physically contiguous within the Tg protein. Indeed, secretory ChEL, made as a separate protein, is engaged in a direct physical interaction with I-II-III throughout the secretory pathway and into the extracellular space (Figure 2-6A). Nevertheless, most newly synthesized I-II-III does not exit the ER with a secretory ChEL partner that was synthesized during the same pulse-labeling period. Rather, the kinetics of I-II-III secretion are delayed in comparison to the speedy (and efficient) release of coexpressed secretory ChEL (Figure 2-5C). Evidently, intracellular transport of I-II-III must wait for subsequently translated ChEL domain to become available for interaction (Figure 2-6B). Thus, in all likelihood, I-II-III must undergo some preliminary folding prior to its interaction with ChEL. Indeed, the ChEL domain has an even stronger affinity for well-folded I-II-III than it does for immature I-II-III (Figure 2-10). This binding is not accidental but essential to the Tg secretory process. Specifically, retention of ChEL within the ER-Golgi system via an engineered KDEL sequence co-retains the mature I-II-III (Figure 2-11) – even as it permits the ChEL domain to function normally as a molecular chaperone (Figure 2-8). Thus, we are forced to conclude that the ChEL domain also functions as a molecular escort [3] required to convey Tg I-II-III out of the ER.

It is interesting to consider the evolutionary dependence of Tg on these mechanisms of intramolecular rescue, given that the major thyroid hormone–forming site in Tg is at its extreme amino terminus [35]. One might predict that a highly truncated Tg

with fewer disulfide bonds and preservation of crucial amino acids for thyroid hormone formation [10, 11] would allow survival without the need of intramolecular chaperone or escort functions. Indeed, there is speculation that such a short aminoterminal portion of Tg may be all that is required for human survival, which requires ongoing thyroid hormone synthesis [36]. However, evolutionary and environmental circumstances (i.e., iodide unavailability on the earth's crust) have promoted development of a more complex Tg structure for iodide storage in vertebrate organisms [1].

In summary, we propose that intracellular Tg transport for thyroid hormone synthesis is dependent on intramolecular chaperone and escort functions embedded within its ChEL domain.

**Acknowledgements**

This work was supported by National Institutes of Health grant DK40344 (to Dr. P. Arvan). I would like to thank Dr. V. Lingappa (U. of California San Francisco) for the prolactin signal peptide in pSPPS1. We also thank members of the Arvan laboratory, especially Xiang Zhao, for helpful suggestions during the course of this work. This work was published at J. Clin. Invest. (2008) 118: 2950-2958.

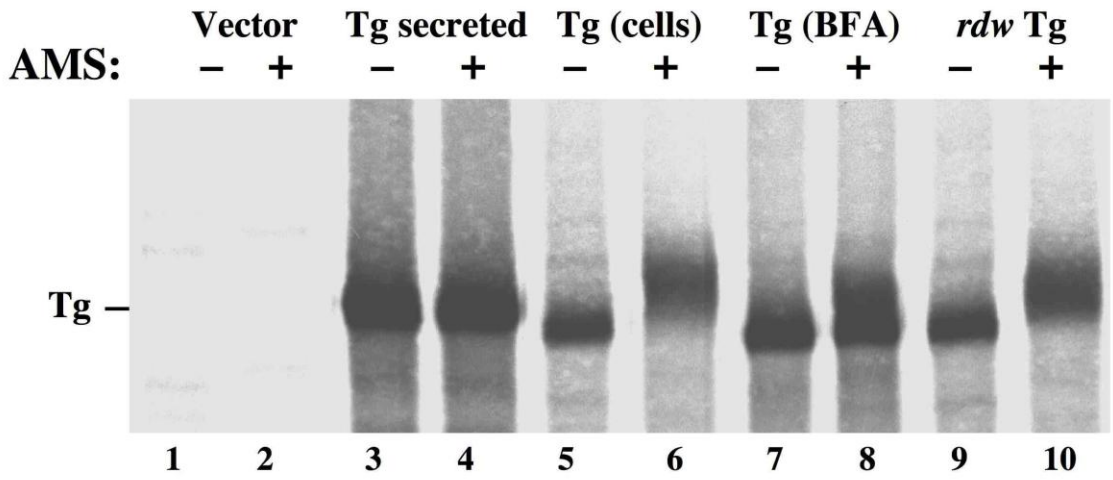


**Figure 2-1. Free Cys thiols in Tg.** 293 cells were transiently transfected with an expression vector encoding wild-type Tg, the *rdw* Tg mutant (G2300R), or empty vector. At 48 hours after transfection, cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for 4 hours in the absence or presence of BFA (5 μg/ml) where indicated. At this time, chase media (lanes labeled "Tg secreted") and cell lysates (all other lanes) were immunoprecipitated with anti-Tg. The immunoprecipitates were denatured in 2× SDS gel sample buffer lacking reducing agents and mock incubated or incubated with AMS (5 mM, 30°C for 1 hour). At the end of the incubation, samples were boiled in the presence of 20 mM DTT and analyzed by 4% SDS-PAGE and fluorography. A slowed mobility (shift up) of the Tg band after AMS is indicative of free reactive thiols in the Tg molecule that are not apparent in secreted, wild-type Tg.

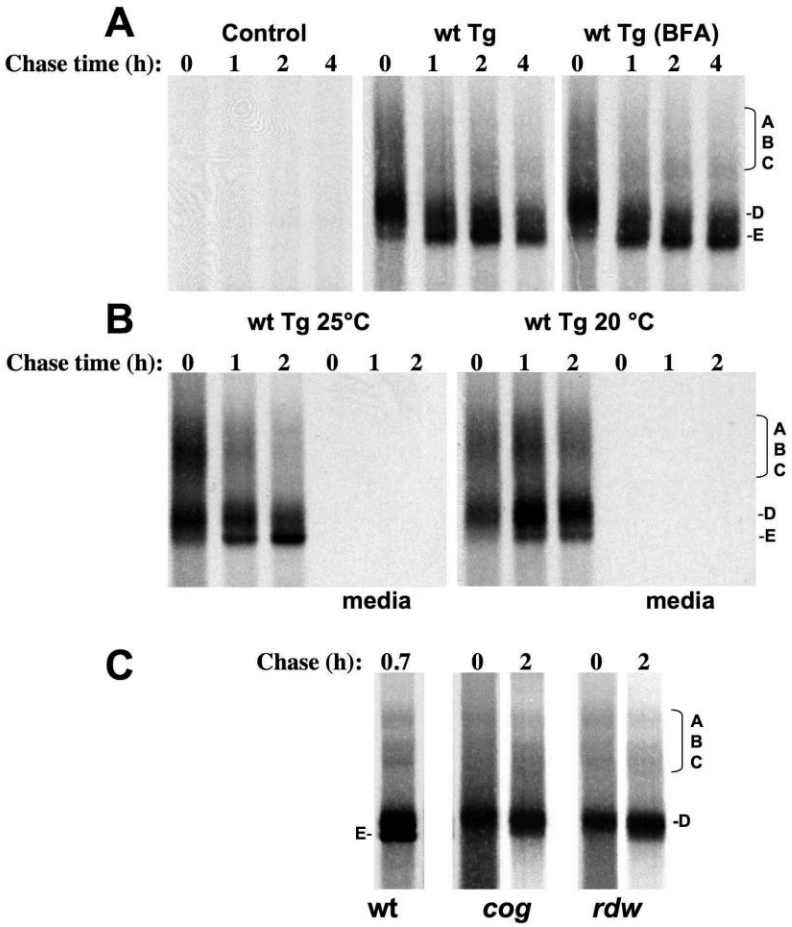


**Figure 2-2. Arrested disulfide maturation of Tg bearing a mutation in the ChEL domain. (A)** 293 cells were transiently transfected with vector DNA encoding wild-type Tg (bearing a C-terminal myc-6xHis epitope tag that does not block Tg secretion; ref. [37]) or empty vector (control [Con]). Transfected cells were then pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased at 37°C for up to 4 hours in the absence or presence of BFA (5 μg/ml) where indicated. At each chase time, cells were lysed, immunoprecipitated with anti-Tg, and analyzed by nonreducing 4% SDS-PAGE and fluorography. A region of the gel enriched in disulfide-linked Tg adducts A, B, and C [28] is indicated. Two additional forms of Tg monomer – folding intermediate D and the E band representing fully oxidized mature Tg – are shown. **(B)** Results of an experiment identical to the one represented in A, except that the cells were chased at either 25°C (left) or 20°C (right). At each time, chase medium was also collected (as indicated), but at these temperatures the media contained no labeled Tg. **(C)** Results of an experiment identical to the one represented in A, except that cells were transfected with plasmid DNA encoding *cog* (L2263P) or *rdw* (G2300R) mutations within the Tg ChEL domain, and the cell lysis buffer included 20 mM *N*-ethylmaleimide. No detectable Tg was secreted, so analysis of the supernatant is not shown.

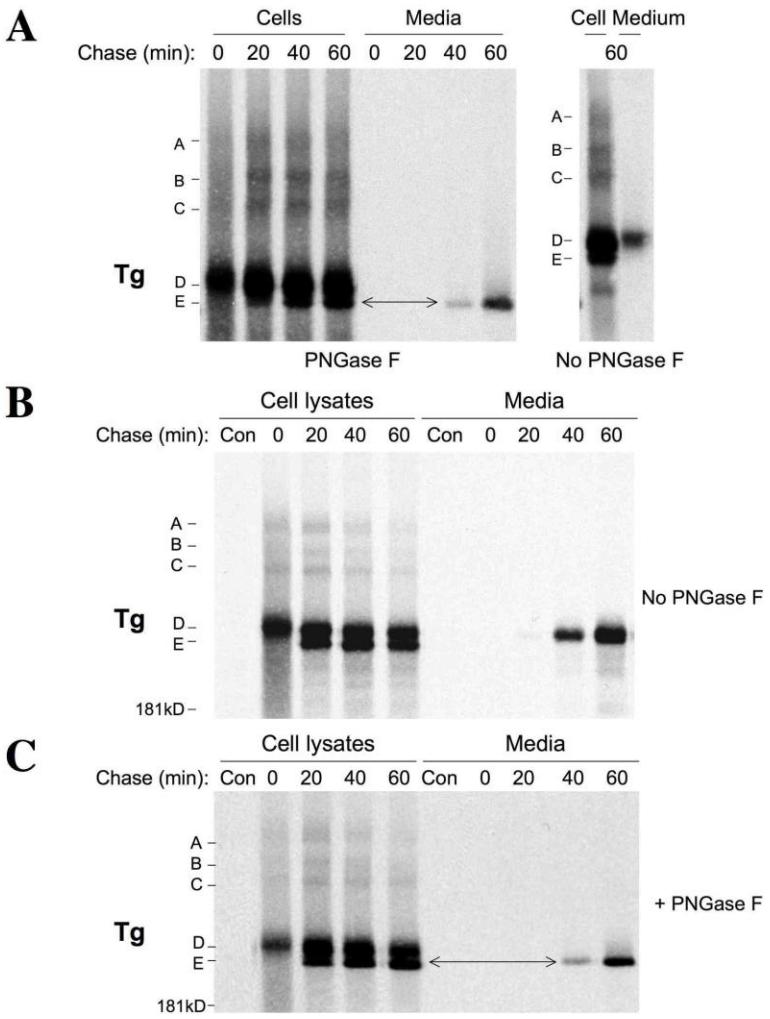


**Figure 2-3. Oxidation state of secreted Tg.** At each chase time, cells were lysed in buffer including 20 mM *N*-ethylmaleimide, and the lysates and chase media immunoprecipitated with anti-Tg and analyzed by nonreducing 4% SDS-PAGE and fluorography. **(A)** PC Cl3 cells were pulse labeled for 10 minutes with $^{35}$S-labeled amino acids and then chased for the times indicated. Immunoprecipitates were either undigested or digested with PNGase F as indicated. Folding intermediates A, B, and C, which have been characterized in previous studies [28], are shown. Also identifed are 2 closely spaced Tg disulfide isomer bands labeled D and E, respectively. Because of glycosylation differences, the mature E isoform does not comigrate with Tg secreted to the medium at 1 hour of chasing. After PNGase F digestion to remove N-glycans, all Tg forms exhibit a faster (shifted-down) mobility. Under these conditions, it is now apparent that secreted Tg comigrates with the intracellular E isomer, identifying the most oxidized band as the most mature folded form of Tg. **(B)** Results of an experiment identical to the one represented in A, except using recombinant Tg expressed in 293 cells, without PNGase F digestion. **(C)** Results of a repeat experiment of that shown in panel B, but including PNGase F digestion. The position of a 181-kDa prestained molecular weight standard is shown at left.

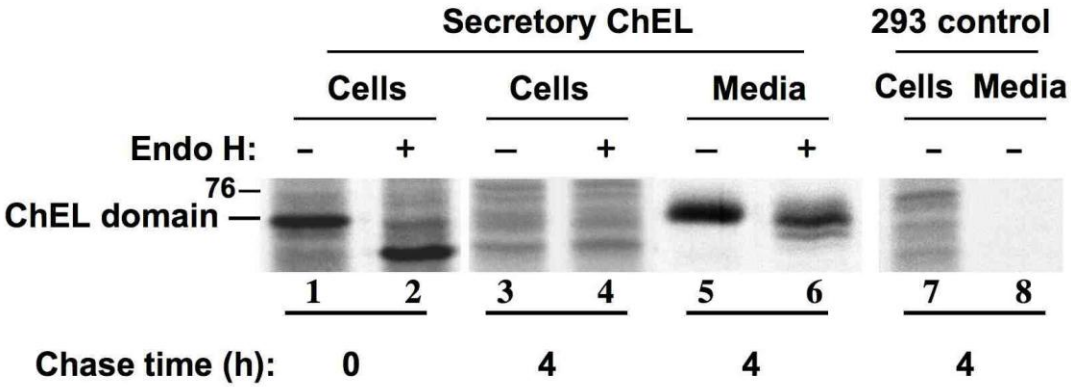


**Figure 2-4. Efficient exit of the isolated Tg ChEL domain from the ER.** 293 cells were transiently transfected with a plasmid encoding the wild-type mouse Tg ChEL domain preceded by the prolactin signal peptide (Secretory ChEL) or were untransfected (293 control). Cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for 0 or 4 hours as indicated, at which time the cells were lysed and both lysates and media immunoprecipitated with a rabbit polyclonal anti-Tg. Immunoprecipitates from transfected cells were divided in 2 equal portions and either mock digested or digested with endoglycosidase H (Endo H). Finally, all samples were analyzed by reducing 5.5% SDS-PAGE and fluorography. The band shift observed after digestion of secretory ChEL from the 0 chase time (shift down, lane 2) is indicative of endoglycosidase H sensitivity and defines ChEL that has not yet reached the Golgi complex; in contrast, none of the secreted ChEL shows the same endoglycosidase H sensitivity, indicating intracellular transport via the Golgi complex. The lanes shown were all run on the same gel, although they are presented noncontiguously. The position of the 76-kDa molecular mass standard is shown at left.

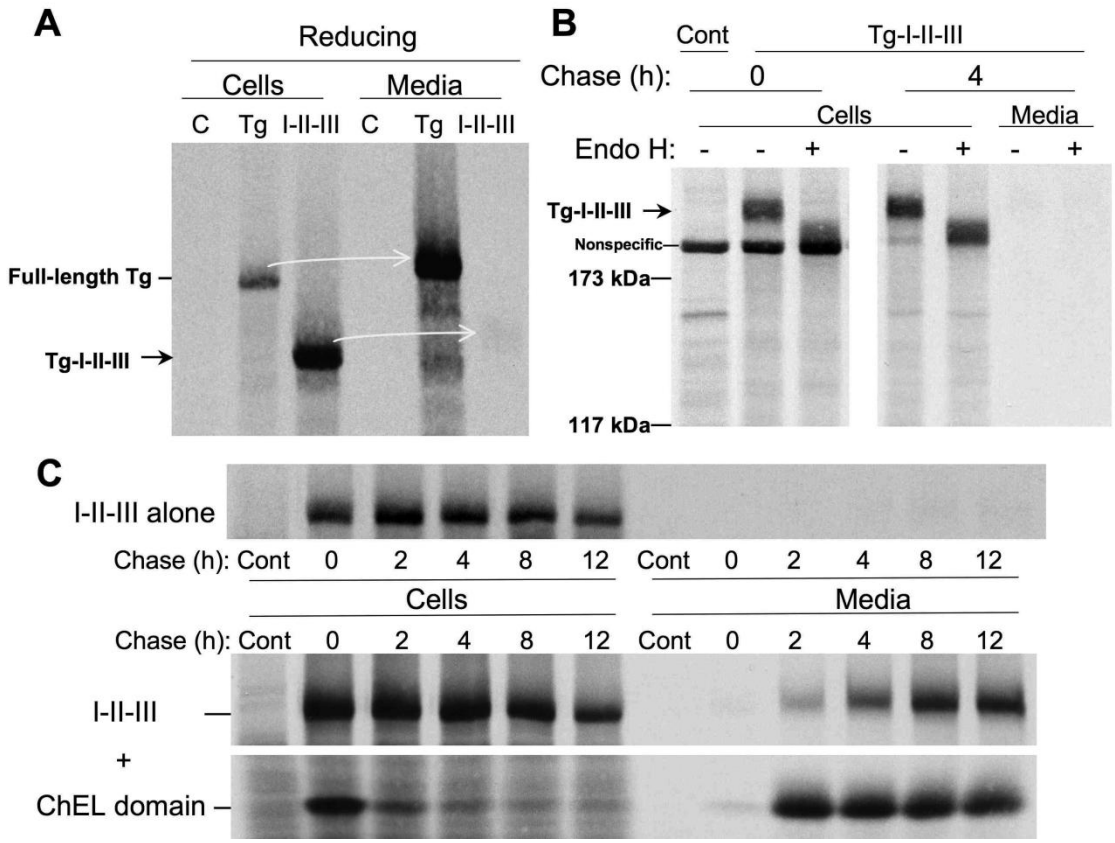


**Figure 2-5. ER exit of Tg region I-II-III in the absence and presence of ChEL domain. (A)** 293 cells were either untransfected (control) or transiently transfected with 2 μg plasmid DNA encoding either full-length wild-type mouse Tg or Tg region I-II-III. Cells were pulse labeled and chased and samples prepared and immunoprecipitated with anti-Tg as in Figure 4, with analysis by reducing 4% SDS-PAGE and fluorography. Arrows highlight secretion (or lack thereof) from cells to media. **(B)** The first lane represents untransfected control cells. The remaining cells transfected to express Tg region I-II-III were pulse labeled as in A and then lysed without or with 4 hours of chasing. Tg immunoprecipitates from cell lysates and media were either mock digested or digested with endoglycosidase H for 1 hour at 37°C, before reducing 4% SDS-PAGE and fluorography. Lanes were run contiguously but have been separated for clarity. The positions of molecular mass standards are shown at left. **(C)** Lower panels: 293 cells were transiently transfected with 0.5 μg plasmid DNA encoding I-II-III cotransfected with 2.5 μg of plasmid DNA encoding the secretory ChEL domain or untransfected controls. Cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for the times indicated. At each chase time, the cells were lysed and both lysates and media immunoprecipitated with polyclonal anti-Tg to recover both I-II-III and secretory ChEL proteins, as revealed by reducing 5.5% SDS-PAGE and fluorography. The upper panel represents a negative control with identical transfection of I-II-III but without cotransfection of the secretory ChEL plasmid.

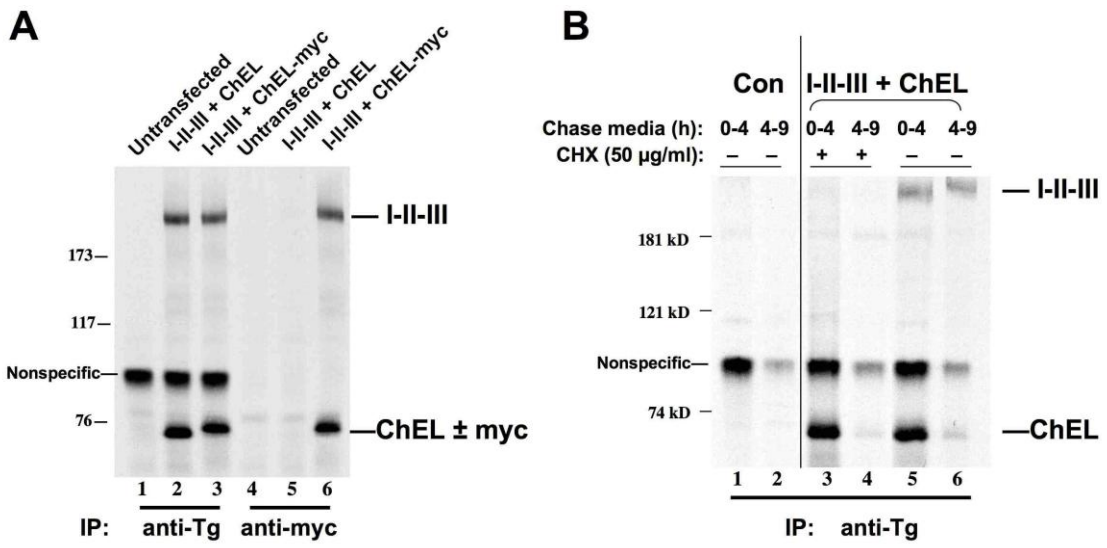


**Figure 2-6. Secreted I-II-III protein is physically associated with secretory ChEL protein. (A)** 293 cells were transiently transfected with 0.5 μg plasmid DNA encoding I-II-III and cotransfected with 2.5 μg of plasmid DNA encoding the secretory ChEL domain either lacking or containing a myc epitope tag, as indicated. The cotransfected cells or untransfected controls were pulse labeled for 30 minutes and chased in complete media, and the secretion after 6 hours was analyzed by immunoprecipitation with anti-Tg or anti-myc. Immunoprecipitates and coprecipitates were analyzed by reducing 5.5% SDS-PAGE and fluorography. Addition of the myc tag slightly retards the SDS-PAGE mobility of the ChEL domain. Note that anti-myc precipitation of ChEL-myc coprecipitates I-II-III. **(B)** Cells untransfected (control) or cotransfected and pulse labeled as in A (I-II-III + ChEL) were chased in complete media for the time intervals shown, in the presence or absence of cycloheximide (CHX). The media were immunoprecipitated with anti-Tg and analyzed by reducing SDS-PAGE and fluorography. These lanes were run contiguously; a black line has been added for clarity to separate the samples. Note that prelabeled ChEL secretion proceeded rapidly in the presence of CHX, but prelabeled I-II-III secretion was blocked. The positions of molecular mass markers are shown at left.

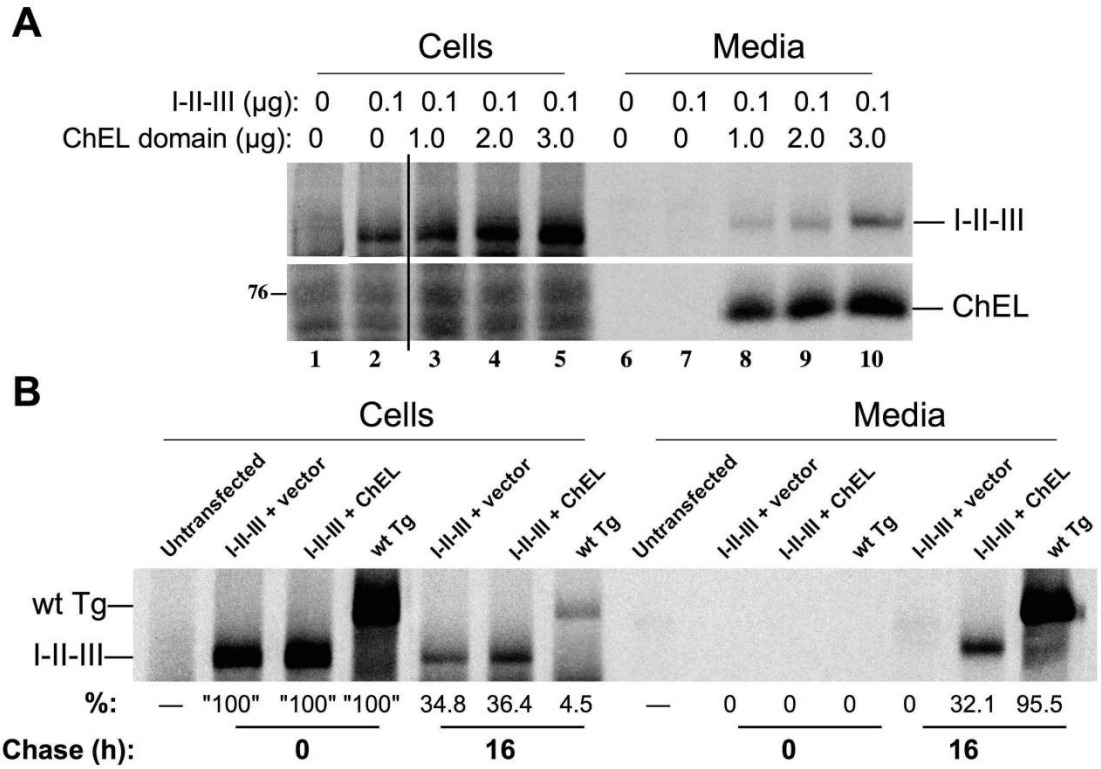


**Figure 2-7. ChEL interaction improves recovery as well as secretion of Tg I-II-III.** **(A)** 293 cells were triply transfected with empty vector plus a plasmid encoding Tg region I-II-III (always 0.1 μg DNA per well) plus a plasmid encoding the secretory ChEL domain (at different DNA levels as shown). DNA in each transfection totaled 3.1 μg per well. Transfected cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for 6 hours, at which time the cells were lysed, and both lysates and media were immunoprecipitated with anti-Tg and analyzed by reducing 5.5% SDS-PAGE and fluorography, as shown. The position of a 76-kDa molecular mass marker is shown at left. The figure has been spliced at the position indicated by a black line (between lanes 2 and 3), but all data were derived from a single exposure of the same gel. **(B)** Cells were either untransfected or transfected with 0.5 μg plasmid DNA encoding I-II-III plus 2.5 μg of vector DNA or that encoding secretory ChEL. As a positive control, 2 μg of plasmid DNA encoding wild-type Tg was transfected in parallel. Cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for either 0 or 16 hours, at which time the cells were lysed and both lysates and media immunoprecipitated with anti-Tg and analyzed by reducing 5.5% SDS-PAGE and phosphorimaging, as shown. The intracellular band density at the 0 chase time was defined as 100%; based on this, the recovery of each band at 16 hours is shown. Total recovery of labeled I-II-III alone at 16 hours was approximately 35%, while total recovery of labeled I-II-III (cells plus media) in the presence of secretory ChEL was approximately 68%.

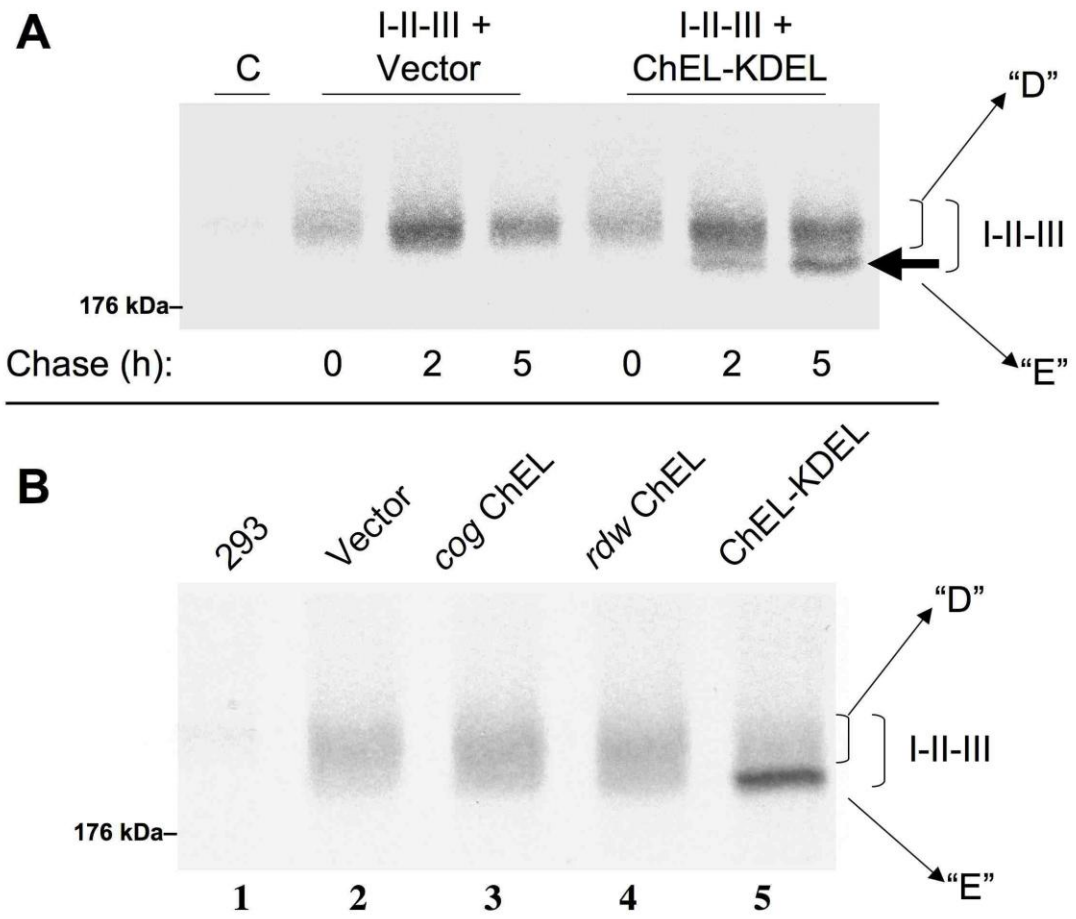


**Figure 2-8. ChEL functions as a molecular chaperone for I-II-III. (A)** Cells were either untransfected (control) or transfected with a plasmid encoding I-II-III plus either empty vector or secretory ChEL-KDEL, as indicated. The transfected cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for the times indicated. At each chase time, cells were immunoprecipitated with anti-Tg and newly synthesized I-II-III analyzed by nonreducing 4% SDS-PAGE and fluorography. Absence of recovery of I-II-III from untransfected control cells is shown at left. The band seen at time 0 and after chasing for 2 hours appears equivalent to that of the D isoform of full-length Tg. In the presence of the ChEL domain in the ER, a faster-migrating band equivalent to that of the mature E isoform of full-length Tg is detected (filled arrow). **(B)** Results of an experiment using the same cotransfection protocol and analysis as in A, except the second plasmid is either vector alone or secretory ChEL domain containing the *cog* mutation, the *rdw* mutation, or the KDEL appendage. Absence of recovery of I-II-III from untransfected 293 cells is shown in lane 1. The position of a 176-kDa molecular mass marker is shown at left.

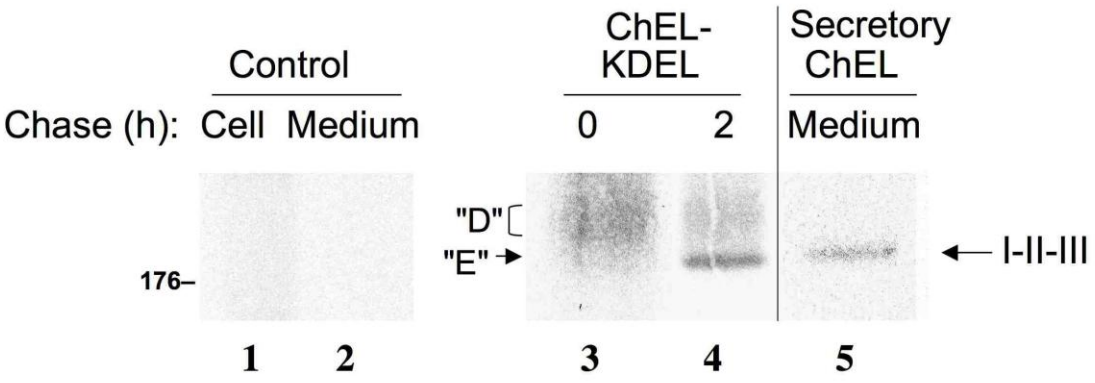


**Figure 2-9. Oxidation state of secreted I-II-III.** 293 cells were transfected either with empty vector (control) or with secretory ChEL bearing or lacking the KDEL appendage. The cells were pulse labeled for 30 min with $^{35}$S-labeled amino acids and then chased for the times indicated (control cells were chased for 2 hours). The cells were lysed, and both lysates and chase medium immunoprecipitated with anti-Tg and subjected to PNGase F digestion to remove N-glycans. The samples were then analyzed by nonreducing 5.5% SDS-PAGE. Under these conditions, it is apparent that secreted I-II-III closely migrates with the intracellular E isomer of I-II-III. A black line has been added to distinguish the medium from the cell lysate. The position of a 176-kDa molecular mass marker is shown at left.

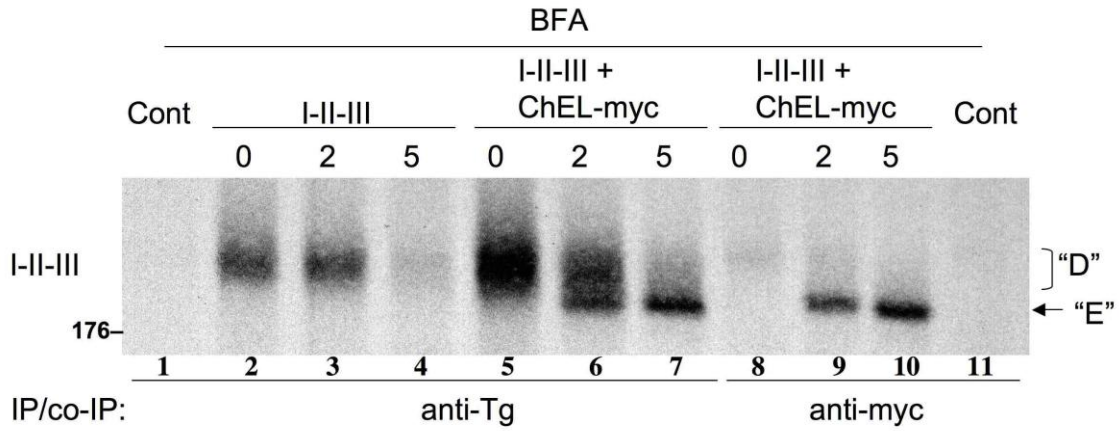


**Figure 2-10. Intracellular association of secretory ChEL with I-II-III.** 293 cells were transfected with empty vector or with plasmid DNAs encoding I-II-III (0.5 μg) plus 1.5 μg of either vector alone (I-II-III) or secretory ChEL with an appended myc epitope tag (ChEL-myc). All cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased for the times shown in the presence of BFA (5 μg/ml). The cells were lysed and immunoprecipitated either anti-Tg (lanes 1–7) or anti-myc (lanes 8–11). Immunoprecipitates were analyzed by nonreducing 5.5% SDSPAGE and fluorography, with the I-II-III region shown. The I-II-III band seen at time 0 appears equivalent to that of the D isoform, while ChELmyc favorably coprecipitates a faster-migrating band equivalent to that of the mature E isoform (arrow). The position of a 176-kDa molecular mass marker is shown at left.

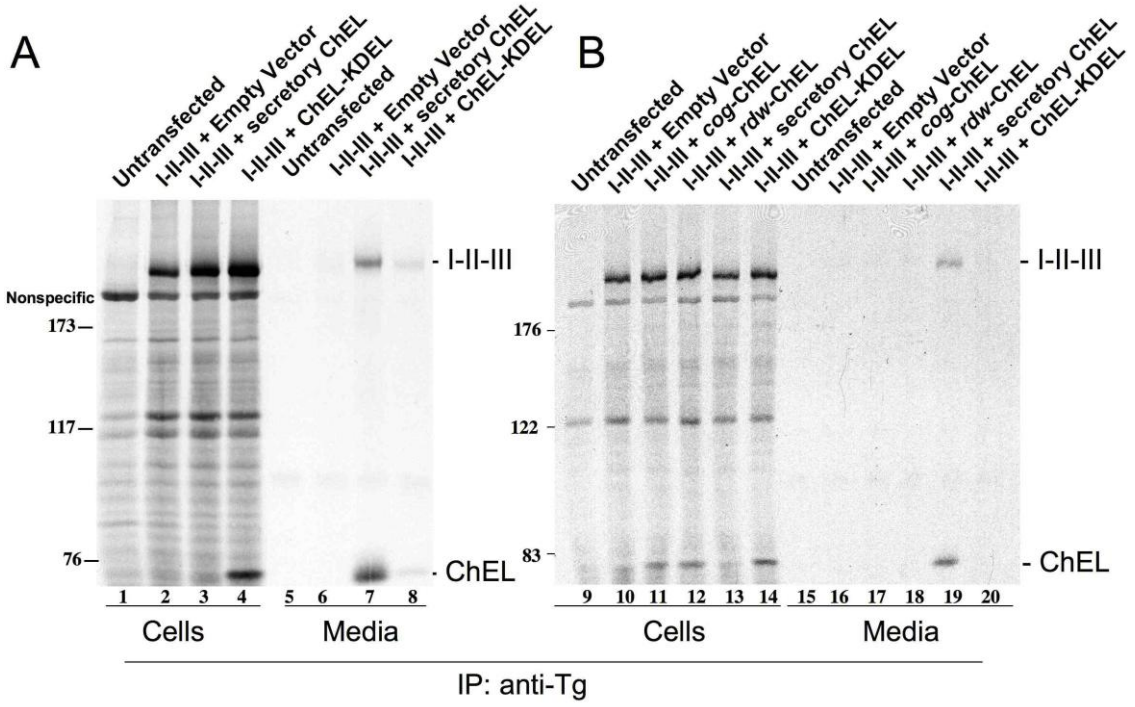


**Figure 2-11. The Tg ChEL domain functions as a molecular escort. (A)** 293 cells were untransfected or transiently cotransfected with 0.5 μg plasmid DNA encoding I-II-III plus 2.5 μg of the constructs indicated. The cells were pulse labeled for 30 minutes with $^{35}$S-labeled amino acids and chased in complete medium for 4 hours. Cell lysates and media were immunoprecipitated with anti-Tg and analyzed by SDS-PAGE and fluorography. Note that anti-Tg recognizes both I-II-III and the ChEL protein; the ChEL-KDEL construct is selectively retained intracellularly, while secretory ChEL is released to the medium. Intracellular retention of ChEL-KDEL causes a parallel retention of I-II-III. **(B)** Cells cotransfected and pulse labeled as in A were chased for 5 hours before analysis as above. Note that the secretory ChEL domain bearing the *cog* or *rdw* mutation cannot support the secretion of I-II-III. The positions of molecular mass markers are shown at left.

**References for chapter 2**


1. Arvan, P., and Di Jeso, B. (2005). Thyroglobulin structure, function, and biosynthesis. In The Thyroid, 9th Edition, L.E. Braverman and R.D. Utiger, eds. (Philadelphia, Pennsylvania, USA: Lippincott Williams & Wilkins), pp. 77-95.

2. den Hartog, M.T., Sijmons, C.C., Bakker, O., Ris-Stalpers, C., and de Vijlder, J.J. (1995). Importance of the content and localization of tyrosine residues for thyroxine formation within the N-terminal part of human thyroglobulin. Eur. J. Endocrinol. *132*, 611-617.

3. Kim, P.S., and Arvan, P. (1998). Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones. Endocr. Rev. *19*, 173-202.

4. Arvan, P., Kim, P.S., Kuliawat, R., Prabakaran, D., Muresan, Z., Yoo, S.E., and Abu Hossain, S. (1997). Intracellular protein transport to the thyrocyte plasma membrane: potential implications for thyroid physiology. Thyroid *7*, 89-105.

5. Muresan, Z., and Arvan, P. (1997). Thyroglobulin transport along the secretory pathway. Investigation of the role of molecular chaperone, GRP94, in protein export from the endoplasmic reticulum. J. Biol. Chem. *272*, 26095-26102.

6. Veneziani, B.M., Giallauria, F., and Gentile, F. (1999). The disulfide bond pattern between fragments obtained by the limited proteolysis of bovine thyroglobulin. Biochimie *81*, 517-525.

7. Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S.S., Friedmann, T., and Taylor, P. (1986). Primary structure of Torpedo californica acetylcholinesterase deduced from its cDNA sequence. Nature *319*, 407-409.

8. Swillens, S., Ludgate, M., Mercken, L., Dumont, J.E., and Vassart, G. (1986). Analysis of sequence and structure homologies between thyroglobulin and acetylcholinesterase: possible functional and clinical significance. Biochem. Biophys. Res. Commun. *137*, 142-148.

9. Mori, N., Itoh, N., and Salvaterra, P.M. (1987). Evolutionary origin of cholinergic macromolecules and thyroglobulin. Proc. Natl. Acad. Sci. U. S. A. *84*, 2813-2817.

10. Marriq, C., Lejeune, P.J., Venot, N., and Vinet, L. (1991). Hormone formation in the isolated fragment 1-171 of human thyroglobulin involves the couple tyrosine 5 and tyrosine 130. Mol. Cell. Endocrinol. *81*, 155-164.

11. Dunn, A.D., Corsi, C.M., Myers, H.E., and Dunn, J.T. (1998). Tyrosine 130 is an important outer ring donor for thyroxine formation in thyroglobulin. J. Biol. Chem. *273*, 25223-25229.

12. van de Graaf, S.A., Ris-Stalpers, C., Pauws, E., Mendive, F.M., Targovnik, H.M., and de Vijlder, J.J. (2001). Up to date with human thyroglobulin. J. Endocrinol. *170*, 307-321.

13. Kim, P.S., Bole, D., and Arvan, P. (1992). Transient aggregation of nascent thyroglobulin in the endoplasmic reticulum: relationship to the molecular chaperone, BiP. J. Cell Biol. *118*, 541-549.

14. Kim, P.S., Kwon, O.Y., and Arvan, P. (1996). An endoplasmic reticulum storage disease causing congenital goiter with hypothyroidism. J. Cell Biol. *133*, 517-527.

15. Baryshev, M., Sargsyan, E., Wallin, G., Lejnieks, A., Furudate, S., Hishinuma, A., and Mkrtchian, S. (2004). Unfolded protein response is involved in the pathology of human congenital hypothyroid goiter and rat non-goitrous congenital hypothyroidism. J. Mol. Endocrinol. *32*, 903-920.

16. Menon, S., Lee, J., Abplanalp, W.A., Yoo, S.E., Agui, T., Furudate, S., Kim, P.S., and Arvan, P. (2007). Oxidoreductase interactions include a role for ERp72 engagement with mutant thyroglobulin from the rdw/rdw rat dwarf. J. Biol. Chem. *282*, 6183-6191.

17. Caputo, M., Rivolta, C.M., Esperante, S.A., Gruneiro-Papendieck, L., Chiesa, A., Pellizas, C.G., Gonzalez-Sarmiento, R., and Targovnik, H.M. (2007). Congenital hypothyroidism with goitre caused by new mutations in the thyroglobulin gene. Clin. Endocrinol. (Oxf). *67*, 351-357.

18. Pardo, V., Rubio, I.G., Knobel, M., Aguiar-Oliveira, M.H., Santos, M.M., Gomes, S.A., Oliveira, C.R., Targovnik, H.M., and Medeiros-Neto, G. (2008). Phenotypic variation among four family members with congenital hypothyroidism caused by two distinct thyroglobulin gene mutations. Thyroid *18*, 783-786.

19. Caron, P., Moya, C.M., Malet, D., Gutnisky, V.J., Chabardes, B., Rivolta, C.M., and Targovnik, H.M. (2003). Compound heterozygous mutations in the thyroglobulin gene (1143delC and 6725G-->A [R2223H]) resulting in fetal goitrous hypothyroidism. J. Clin. Endocrinol. Metab. *88*, 3546-3553.

20. Kitanaka, S., Takeda, A., Sato, U., Miki, Y., Hishinuma, A., Ieiri, T., and Igarashi, T. (2006). A novel compound heterozygous mutation in the thyroglobulin gene resulting in congenital goitrous hypothyroidism with high serum triiodothyronine levels. J. Hum. Genet. *51*, 379-382.

21. Rivolta, C.M., and Targovnik, H.M. (2006). Molecular advances in thyroglobulin disorders. Clin. Chim. Acta *374*, 8-24.

22. Matakidou, A., Hamel, N., Popat, S., Henderson, K., Kantemiroff, T., Harmer, C., Clarke, S.E., Houlston, R.S., and Foulkes, W.D. (2004). Risk of non-medullary thyroid cancer influenced by polymorphic variation in the thyroglobulin gene. Carcinogenesis *25*, 369-373.

23. Kim, P.S., Ding, M., Menon, S., Jung, C.G., Cheng, J.M., Miyamoto, T., Li, B., Furudate, S., and Agui, T. (2000). A missense mutation G2320R in the thyroglobulin gene causes non-goitrous congenital primary hypothyroidism in the WIC-rdw rat. Mol. Endocrinol. *14*, 1944-1953.

24. Hishinuma, A., Furudate, S., Oh-Ishi, M., Nagakubo, N., Namatame, T., and Ieiri, T. (2000). A novel missense mutation (G2320R) in thyroglobulin causes hypothyroidism in rdw rats. Endocrinology *141*, 4050-4055.

25. Kim, P.S., Hossain, S.A., Park, Y.N., Lee, I., Yoo, S.E., and Arvan, P. (1998). A single amino acid change in the acetylcholinesterase-like domain of thyroglobulin causes congenital goiter with hypothyroidism in the cog/cog mouse: a model of human endoplasmic reticulum storage diseases. Proc. Natl. Acad. Sci. U. S. A. *95*, 9909-9913.

26. Park, Y.N., and Arvan, P. (2004). The acetylcholinesterase homology region is essential for normal conformational maturation and secretion of thyroglobulin. J. Biol. Chem. *279*, 17085-17089.

27. Frand, A.R., and Kaiser, C.A. (1999). Ero1p oxidizes protein disulfide isomerase in a pathway for disulfide bond formation in the endoplasmic reticulum. Mol. Cell *4*, 469-477.

28. Di Jeso, B., Park, Y.N., Ulianich, L., Treglia, A.S., Urbanas, M.L., High, S., and Arvan, P. (2005). Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase. Mol. Cell. Biol. *25*, 9793-9805.

29. Kim, P.S., and Arvan, P. (1991). Folding and assembly of newly synthesized thyroglobulin occurs in a pre-Golgi compartment. J. Biol. Chem. *266*, 12412-12418.

30. Morel, N., Leroy, J., Ayon, A., Massoulie, J., and Bon, S. (2001). Acetylcholinesterase H and T dimers are associated through the same contact. Mutations at this interface interfere with the C-terminal T peptide, inducing degradation rather than secretion. J. Biol. Chem. *276*, 37379-37389.

31. Ekholm, R. (1990). Biosynthesis of thyroid hormones. Int. Rev. Cytol. *120*, 243-288.

32. Vono-Toniolo, J., Rivolta, C.M., Targovnik, H.M., Medeiros-Neto, G., and Kopp, P. (2005). Naturally occurring mutations in the thyroglobulin gene. Thyroid *15*, 1021-1033.

33. Falconer, I.R. (1966). Studies of the congenitally goitrous sheep. The iodinated compounds of serum, and circulating thyroid-stimulating hormone. Biochem. J. *100*, 190-196.

34. de Vijlder, J.J. (2003). Primary congenital hypothyroidism: defects in iodine pathways. Eur. J. Endocrinol. *149*, 247-256.

35. Dunn, J.T., and Dunn, A.D. (1999). The importance of thyroglobulin structure for thyroid hormone biosynthesis. Biochimie *81*, 505-509.

36. van de Graaf, S.A., Ris-Stalpers, C., Veenboer, G.J., Cammenga, M., Santos, C., Targovnik, H.M., de Vijlder, J.J., and Medeiros-Neto, G. (1999). A premature stopcodon in thyroglobulin messenger RNA results in familial goiter and moderate hypothyroidism. J. Clin. Endocrinol. Metab. *84*, 2537-2542.

37. Kim, P.S., Lee, J., Jongsamak, P., Menon, S., Li, B., Hossain, S.A., Bae, J.H., Panijpan, B., and Arvan, P. (2008). Defective protein folding and intracellular retention of thyroglobulin-R19K mutant as a cause of human congenital goiter. Mol. Endocrinol. *22*, 477-484.

# CHAPTER 3

## THE CHOLINESTERASE-LIKE DOMAIN OF THYROGLOBULIN, ESSENTIAL IN THYROGLOBULIN TRAFFICKING FOR THYROID HORMONE SYNTHESIS, IS REQUIRED FOR PROTEIN DIMERIZATION


### Abstract

The carboxyl-terminal cholinesterase-like (ChEL) domain of thyroglobulin (Tg) has been identified as critically important in Tg export from the endoplasmic reticulum. In a number of human kindreds suffering from congenital hypothyroidism and in the *cog* congenital goiter mouse and *rdw* rat dwarf models, thyroid hormone synthesis is inhibited because of mutations in the ChEL domain that block protein export from the endoplasmic reticulum. We hypothesize that Tg forms homodimers through noncovalent interactions involving two predicted alpha helices in the ChEL domain that are homologous to the dimerization helices of acetylcholinesterase. This has been explored through selective epitope tagging of dimerization partners, and by inserting an extra, unpaired Cys residue to create an opportunity for intermolecular disulfide pairing. We show that the ChEL domain is necessary and sufficient for Tg dimerization; specifically, the isolated ChEL domain can dimerize with full-length Tg or with itself. Insertion of an N-linked glycan into the putative upstream dimerization helix blocks detectable homodimerization of the isolated ChEL domain. However, interestingly, co-expression of upstream Tg domains, either *in cis* or *in trans*, overrides the dimerization defect of such a mutant. Thus, while the ChEL domain provides a nidus for Tg dimerization, interactions of upstream Tg

regions with the ChEL domain actively stabilizes the Tg dimer complex for intracellular transport.

## Introduction

The synthesis of thyroid hormone in the thyroid gland requires secretion of thyroglobulin (Tg) to the apical luminal cavity of thyroid follicles [1]. Once secreted, Tg is iodinated via the activity of thyroid peroxidase [2], followed by a coupling reaction involving a quinol-ether linkage especially engaging di-iodinated tyrosyl residues 5 and 130 to form thyroxine within the amino-terminal portion of the Tg polypeptide [3, 4]. Preferential iodination of Tg hormonogenic sites is dependent not on the specificity of the peroxidase [5] but on the native structure of Tg [6]. To date, no other thyroidal proteins have been shown to effectively substitute in this role for Tg.

The first 80% of the primary structure of Tg (full-length murine Tg: 2,746 amino acids) involves three regions called I-II-III comprised of disulfide-rich repeat domains held together by intradomain disulfide bonds [7, 8]. The final 580 amino acids of Tg are strongly homologous to acetylcholinesterase [9-11]. Rate-limiting steps in the overall process of Tg secretion involve structural maturation within the endoplasmic reticulum (ER) [12]. Interactions between regions I-II-III and the cholinesterase-like (ChEL) domain have recently been suggested to be important in this process, with ChEL functioning as an intramolecular chaperone and escort for I-II-III export [13]. In addition, Tg conformational maturation culminates in Tg homodimerization [14, 15] with progression to a cylindrical and ultimately, a compact ovoid structure [16-18].

In human congenital hypothyroidism with deficient Tg, the ChEL domain is a commonly affected site of mutation, including the recently-described A2215D [19, 20], R2223H [21], G2300D, R2317Q [22], G2355V, G2356R, and the skipping of exon 45 (which normally encodes 36 amino acids), as well as the Q2638stop mutant [23] (in addition to polymorphisms including P2213L, W2482R, and R2511Q that may be associated with thyroid overgrowth [24]). As best as is currently known, all of the congenital hypothyroidism-inducing Tg mutants are defective for intracellular transport [25]. A homozygous G2300R mutation (equivalent to residue 2298 of mouse Tg) in the ChEL domain is responsible for congenital hypothyroidism in *rdw* rats [26, 27], whereas we identified the Tg-L2263P point mutation as the cause of hypothyroidism in the *cog* mouse [28]. Such mutations perturb intradomain structure [29] and interestingly, block homodimerization [30]. Acquisition of quaternary structure has long been thought to be a prerequisite for efficient export from the ER [31] as exemplified by authentic acetylcholinesterase [32, 33] in which dimerization enhances protein stability and export [34].

Tg comprised only of regions I-II-III (truncated to lack the ChEL domain) is blocked within the ER [29] while a secretory version of the isolated ChEL domain of Tg devoid of I-II-III undergoes rapid and efficient intracellular transport and secretion [13]. Given the striking homology that positions two predicted alpha helices of the ChEL domain to the identical relative positions of the dimerization helices of acetylcholinesterase, these recent findings raise the possibility that ChEL may function as the homodimerization domain of Tg, providing a critical function in maturation for Tg transport to the site of thyroid hormone synthesis [1].

In this chapter, we provide unequivocal evidence for homodimerization of the ChEL domain and "hetero"-dimerization of that domain with full-length Tg, and we provide suggestive evidence that the predicted ChEL dimerization helices provide a nidus for Tg assembly. On the other hand, our data also suggest that upstream Tg regions known to interact with ChEL [13] actively stabilize the Tg dimer complex. Together, I-II-III and ChEL provide unique contributions to the process of intracellular transport of Tg through the secretory pathway.

## Materials and Methods

### Materials

Lipofectamine 2000, Dulbecco's modified Eagle's medium (DMEM), Zysorbin, fetal bovine serum, penicillin, and streptomycin were from Invitrogen; Complete protease inhibitor cocktail was from Roche; Brefeldin A, protein G-Agarose, and protein A-Agarose from Sigma Chemical Co.; Endoglycosidase H from New England Biolabs; Trans$^{35}$S-Label from MP Biomedicals; *Trans*IT-LT1 transfection reagent from Mirus. Rabbit polyclonal anti-myc and anti-GFP were from Immunology Consultants, Inc., monoclonal anti-acetylcholinesterase (MAB303) from Milipore, and monoclonal anti-HA (MMS-101P) from Covance. Rabbit polyclonal anti-Tg (containing antibodies against epitopes at both N- and Cterminal regions of the protein) has been previously described [30].

### Site-directed mutagenesis of mouse Tg cDNA

ChEL domain mutations were introduced with the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using the following mutagenic primers (paired with their complements): D2708C/G2709stop in Tg ("Tg-CD", 5'-CCAGACTTTGAAGGATGCATGTTGAGCCAAGGATGCACAGTTAACC-3'); D2708C in ChEL ("ChEL-CD", 5'–CCAGACTTTGAAGGATGCATGTGGAGCCAAGGATGCACAGTTAACC –3'); ChEL-myc (5'–CAAGAGCTACAGCAAAGAACAGAAACTGATCTCTGAGGAGGACTTATGATTAATGCTTCG –3'); ChEL-HA (5'-CCAAGAGCTACAGCAAATACCCTTACGACGTCCCCGATTACGCGTAGGTTAATGCTTCGC-3'). Truncated Tg regions I-II-III and secretory ChEL were made as described previously [13]. I-II-III-myc and -HA were made using the following primers and their complements (5'-CCGGAAGTCTGAACAGAAGTTGATCTCAGAGGAGGACCTATAGACACCTTCTGTACGC-3' for I-II-III-myc; 5'-CCGGAAGTCTTACCCCTATGACGTCCCAGATTATGCATGATCCACACCTTCTGTACGC-3' for I-II-III-HA). N-glycosylation site (A2538N, V2540T) which is homologous to that produced in acetylcholinesterase [34] was added to Tg-CD and ChEL constructs using the following primer and its complement oligo (5'-GGACTCAGATGCCCGCATCCTTGCTAATGCTACATGGTATTACTCCTTGGAGCACTCC-3'). Each construct was confirmed by direct DNA sequencing before expressed in 293 cells.

**Cell culture and transfection**

293 cells were cultured in DMEM with 10% fetal bovine serum in 6 well plates at 37ºC in a humidified 5% CO2 incubator. Plasmids were transiently transfected using *Trans*IT-LT1 or Lipofectamine 2000 transfection reagent according to the manufacturer's instructions.

**Metabolic labeling and immunoprecipitation**

Transfected 293 cells were starved for 30 min in met/cys-free DMEM, then pulse labeled with 180 μCi/ml $^{35}$S-amino acids. The labeled cells were then washed with an excess of cold Met/Cys and chased in complete DMEM. At each time point, cells were lysed in buffer containing 1% NP-40, 0.1% SDS, 0.1 M NaCl, 2 mM EDTA, 25 mM Tris pH 7.4 and protease inhibitor cocktail cocktail (for cells expressing authentic acetylcholinesterase, lysis buffer contained 1% Triton X-100, 0.1 M NaCl, 2 mM EDTA, 25mM Tris pH 7.4 and protease inhibitors). For immunoprecipitation, anti-Tg or anti-acetylcholinesterase antibodies were incubated with samples overnight at 4ºC, and the immunoprecipitate was recovered with protein A- Agarose (for anti-Tg) or protein G-Agarose (for anti-acetylcholinesterase). For co-immunoprecipitation studies, samples were incubated overnight at 4ºC with anti-myc antibodies and protein A-Agarose. Immunoprecipitates (or co-precipitates) were washed three times before boiling in SDS sample buffer with or without reducing agent, resolved by SDS-PAGE, and analyzed by fluorography or phosphorimaging.

**Endoglycosidase H digestion**

Immunoprecipitates were boiled for 10 min in denaturing solution with 0.5% SDS (or, for Tg D2708C/G2709stop, omitting) containing 1% 2-mercaptoethanol and digested with 250 U of endoglycosidase H in 50 mM sodium citrate, pH 5.5 for 1 h at 37ºC.

## Results

### Use of epitope tagging to follow Tg transport and dimerization

The use of sucrose-velocity gradient centrifugation has been a "gold standard" in examination of the homodimerization of wildtype endogenous Tg [35, 36]. However, for the study of dimerization properties of recombinantly expressed mouse Tg and its domains, we exploited bioengineering methods to tag discrete Tg subunits. Using a (0.5 h) pulse – (5 h) chase protocol, we established that wild-type Tg tagged at the carboxyl-terminus with a green fluorescence protein (GFP) moiety was efficiently secreted, being converted from an endoglycosidase H-sensitive form (large mobility shift upon endo H digestion) to an endo H-resistant form (small mobility shift upon digestion) – precisely as seen for untagged Tg (Figure 3-1A). Adding a triple-myc (3xMyc) tag at the carboxyl-terminus of wild-type Tg produced a similarly well-secreted protein whose release to the media could be blocked by inclusion of the *rdw* (G2298R) Tg mutation (Figure 3-1B). Co-expression of wild-type Tg-GFP with Tg-3xMyc offered one new means to examine Tg dimerization. When Tg-GFP was expressed in 293 cells in the presence or absence of co-expressed Tg-3xMyc, secreted Tg-GFP was positively recovered in the medium, as detected by Western blotting with anti-GFP antibody (Figure 3-1C *right panel*). If the medium was first immunoprecipitated with anti-myc antibody, then Tg-GFP expressed

by itself could no longer be recovered even though it had been secreted (i.e., anti-myc could not immunoprecipitate Tg-GFP, Figure 3-1C, *left panel*). By contrast, when co-expressed, immunoprecipitation of the medium with anti-myc co-precipitated Tg-GFP (Figure 3-1C *left*).

The identical media samples were also analyzed in reverse. Once again, when Tg-3xMyc was expressed in the presence or absence of co-expressed Tg-GFP, secreted Tg-3xMyc was positively recovered as detected by Western blotting with anti-myc (Figure 3-1D *right panel*). If first immunoprecipitated with anti-GFP, then Tg-3xMyc expressed by itself could not be recovered, indicating that anti-GFP did not cross-react with Tg-3xMyc (Figure 3-1D, *left panel*). The anti-GFP antibody was less efficient for Tg-GFP immunoprecipitation (not shown). Nevertheless, to the extent that immunoprecipitation occurred, anti-GFP specifically co-precipitated wild-type Tg-3xMyc when co-expressed with wild-type Tg-GFP (Figure 3-1D *left*). Together, the data in Figure 3-1C and 1D provide strong evidence of dimerization between Tg subunits.

**Putative dimerization sequences in the Tg ChEL domain**

The carboxyl-terminal ChEL domain of Tg is comprised of 581 residues with distinct homology to acetylcholinesterase [10, 11], which undergoes homodimerization via formation of a 4-helix bundle that engages two helices from each monomer [32-34]. Using the PSI-PRED program for prediction of protein secondary structure [37], we examined the sequence of the Tg ChEL domain. We found sequences in the carboxyl-terminal half of the ChEL domain predicted with high confidence to form helical segments (highlighted residues, Figure 3-2). Notably, these sequences are positioned

identically to those for the dimerization helices of acetylcholinesterase [32]. We thus hypothesized that Tg might use this carboxyl-terminal region to encode homodimerization necessary for intracellular protein transport.

**Engagement of the ChEL domain in Tg dimerization**

Evidence indicates that Tg with a dysfunctional ChEL domain cannot homodimerize [30]. To check possible homodimerization of regions I-II-III in the absence of ChEL, we engineered I-II-III-myc and I-II-III-HA tagged proteins. Either individually, or by co-transfection, both constructs were well expressed in 293 cells as detected by immnoblotting with anti-HA or anti-myc (Figure 3-1E *left panel*). Under co-immunoprecipitation conditions, immunoprecipitation with anti-myc recovered the I-II-III-myc protein; however, no co-immunoprecipitation of I-II-III-HA could be detected (Figure 3-1E *right panel*). Co-immunoprecipitation also could not be detected when the anti-HA antibody was employed for immunoprecipitation (not shown). These results (in contrast to those obtained for the isolated ChEL domain, see below) suggest that Tg I-II-III – which is known to be defective in export from the ER [13, 29] – cannot homodimerize. By contrast, in the absence of I-II-III, secretory ChEL with a carboxyl-terminal myc tag is efficiently secreted [13] and current evidence indicates that epitope tagging itself does not block Tg dimerization (Figure 3-1C, D). To examine the dimerization potential of secretory ChEL-myc with full-length wild-type Tg, the two proteins were co-expressed in 293 cells. With or without secretory ChEL-myc, full-length Tg was efficiently secreted as judged by anti-Tg immunoprecipitation (Figure 3-3 *left panel*). In addition, when co-expressed, full-length wild-type Tg could be co-precipitated

from the medium with secretory ChEL-myc (Figure 3-3). While the efficiency of this co-precipitation was low (Figure 3-3, which suggests that for secretory ChEL to dimerize with Tg, it must compete against Tg homodimerization), the data (Figure 3-1E; Figure 3-3) collectively suggest that the isolated ChEL domain is both necessary and sufficient for dimerization of Tg.

Both authentic acetylcholinesterase and the Tg ChEL domain use their 6 Cys residues for intrachain disulfide bonding, while a unique additional Cys residue in authentic acetylcholinesterase (*upward arrow* Figure 3-3A) falls in a peptide extension at the extreme carboxyl-terminus of the monomeric protein just downstream from the ChEL homology region that includes one of dimerization helices [32, 34]. Crystallographic evidence indicates that this unpaired Cys residue can form an intersubunit disulfide bond, which covalently stabilizes acetylcholinesterase homodimers [38]; The presence of such a bond may be exploited as an assay of dimerization potential by nonreducing SDS-PAGE. In 293 cells pulse-labeled with $^{35}$S-amino acids and chased for 3 h, acetylcholinesterase was specifically immunoprecipitated from the chase medium bathing transfected cells, while only background bands were recovered from cells transfected with empty vector (Figure 3-4B). Unrelated to minor heterogeneity in N-linked glycosylation (data not shown), by nonreducing SDS-PAGE, secreted acetylcholinesterase could be recovered at two positions equivalent to monomer and dimer molecular masses (Figure 3-4B). Because dimerization of acetylcholinesterase does not require, nor uniformly employ, the optional intersubunit disulfide bond [39, 40], it is likely that all secreted acetylcholinesterase (Figure 3-4B) is homodimeric although only a portion is stabilized with the intersubunit disulfide bridge.

We wished to exploit this assay to examine dimerization potential of the Tg ChEL domain in the absence of co-expressed full-length wild-type Tg. With this in mind, we mutagenized secretory ChEL to introduce a cysteinyl residue at the carboxyl-terminus of mature murine Tg (Figure 3-4A *downward arrow* equivalent to D2708C,G2709stop mutation) immediately following the helical segment expected to participate in formation of the 4-helix bundle for dimerization. The construct was denoted as "ChEL-CD" to signify the intent to create a construct with "***c***ovalent ***d***imerization" potential. With or without the CD mutation, secretory ChEL proteins exhibited efficient intracellular transport (Figure 3-4C), acquiring endoglycosidase H-resistance within a few hours of synthesis (not shown). However by nonreducing SDS-PAGE secretory ChEL-CD was recovered (in part) as a covalent homodimer (Figure 3-4C *right*). As each of the upstream 6 Cys residues are engaged in evolutionarily conserved intrachain disulfide bonds and thus unavailable for interchain pairing, the acquisition of a new intersubunit covalent bond in ChEL-CD requires engagement of the new unique Cys residue from both monomeric partners, indicating tail-to-tail dimerization of the ChEL domain like that for acetylcholinesterase.

To determine if full-length Tg also engages in tail-to-tail homodimerization, we examined Tg-CD in which the unpaired D2708C (Figure 3-4A) was incorporated into the larger Tg context (Figure 3-5). A D2708S mutant was also prepared as a negative control. First, 293 cells transfected to express Tg-CD were pulse-labeled with $^{35}$S-amino acids and chased for various times in the presence of brefeldin A (BFA) to block intracellular transport in the secretory pathway. Tg immunoprecipitates at each chase time were analyzed by nonreducing SDS-PAGE. At the zero chase time, as series of high molecular

weight intermediates termed "A, B, and C" (Figure 3-5A *left*) have been identified as Tg adducts with resident oxidoreductases of the ER lumen [41]. (These adducts appear as a smear when the cells are lysed in the absence of N-ethylmaleimide or similar aklyating treatment (Figure 3-5A) [13]). The "D" isoform (Figure 3-5A *left*) – which has been shown to be a partially oxidized Tg folding intermediate [13] – was also pronounced at the zero chase time. As reported for wild-type Tg [13], at 1 h of chase there was further maturation of monomers to the fully-oxidized “E” isoform (Figure 3-5A *left*). However, unlike for wild-type Tg, there was the new appearance of Tg-CD homodimers (Figure 3-5A *left panel*). In cells in which secretion was blocked by BFA treatment, intracellular Tg-CD homodimers increased in intensity during the second chase hour (Figure 3-5A *left panel*). From BFA-treated cells, all of these isoforms of Tg migrated as a single band upon reducing SDS-PAGE (not shown).

In the absence of BFA, it was clear that homodimerization of Tg-CD occurred intracellularly even before arrival in the Golgi complex, because a portion of the covalent Tg-CD homodimer had not yet acquired resistance to endoglycosidase H digestion (Figure 3-5A *right*). Such results are consistent with the long-held notion that Tg dimerization occurs in the ER and represents the last structural maturation step required for Tg export to the Golgi complex [14, 15]. Over time, increasing quantities of homodimeric Tg-CD were delivered to the medium (Figure 3-5A *right*). This accumulating amount was exclusively endoglycosidase H-resistant (not shown).

It is known that neither authentic acetylcholinesterase (Figure 3-4B) nor endogenous Tg [41] require an intersubunit disulfide bond for dimer formation. To examine the relationship of the intersubunit disulfide bond of Tg-CD with its

dimerization, we employed sucrose velocity gradient centrifugation. Gradients were loaded from the top and finally collected from the bottom as previously described [41]. In the first three gradients shown in Figure 3-6, the secretion of endogenous Tg from PC Cl3 thyrocytes was compared to that of recombinant wild-type Tg expressed in 293 cells. While 293 cells transfected with empty vector secreted no protein that could be immunoprecipitated with anti-Tg, 293 cells transfected to express wild-type mouse Tg secreted Tg protein that was recovered in the identical dimer fractions as those of the PC Cl3 standard; all running as the 330 kDa band by nonreducing SDS-PAGE. When transfected to express Tg-CD, 293 cells secreted species that ran at both the 660 kDa (covalent dimer) and 330 kDa monomer positions (Figure 3-6 *right*). Importantly, both bands were recovered exclusively in the dimer peak, demonstrating that Tg-CD has a fully preserved ability to homodimerize even without requiring the presence of the covalent intersubunit linkage.

Altogether, the data strongly support that Tg-CD proceeds through the normal Tg folding pathway, achieving homodimerization (and normal intracellular transport), and within the context of the full Tg molecule the extra unpaired cysteine engineered at position 2708 engages in an interchain disulfide bond. We also found that Tg-D2708S,G2709stop was secreted with equally high efficiency to that of wild-type Tg or Tg-CD, excluding any detrimental effects of the G2709stop codon, but this negative control exhibited no covalent dimerization (Figure 3-5B). Thus, given that secretory ChEL-CD (bearing the identical mutation) itself makes covalent homodimers (Figure 3-4C), the accumulated evidence strongly indicates that tail-to-tail dimerization of Tg-CD

monomers creates a juxtaposition that allows the extra unpaired Cys residue of one monomer to partner with the identical Cys residue in the other monomer.

**Perturbation of dimer stability, and suppression of that perturbation**

Previous studies have suggested that introduction of an N-linked glycosylation site into the "α-7/8" helix of authentic acetylcholinesterase can decrease stability of the 4-helix bundle for dimerization [34]. The predicted comparable helix in the ChEL domain of Tg is highlighted in Figure 3-2. Therefore, using a secretory ChEL-myc cDNA template, the **A**-A-**V** sequence shown in Figure 3-2 was mutagenized to **N**-A-**T** within the helical region (this mutant termed "ChEL***G***-myc" to indicate our intent to create a new N-linked ***g***lycosylation site). First, a double-epitope tag strategy was employed similar to that used successfully for full-length Tg (Figure 3-1). Specifically, secretory ChEL tagged with a carboxyl-terminal HA epitope was co-expressed in 293 cells either with secretory ChEL-myc or secretory ChEL***G***-myc. Interestingly, using these constructs, both ChEL-myc and ChEL***G***-myc were well secreted along with ChEL-HA. Anti-myc immunoprecipitation from media was specific and comparably efficient for myc-tagged secretory ChEL or ChEL***G***, but the mobility of the ChEL***G*** protein was slower (Figure 3-7A *upper panel*). [Deglycosylation experiments (not shown) confirmed that the engineered ChEL***G*** glycosylation site was utilized.] Interestingly, in comparison to ChEL-myc, ChEL***G***-myc was deficient for co-precipitation of ChEL-HA (Figure 3-7A *lower panel*).

We next engineered the same N-linked glycosylation site into the covalent dimerizing ChEL-CD construct (Figure 3-4C) to create ChEL***G***-CD. Based on band

mobility, the glycosylation site was indeed utilized in all copies of ChEL***G***-CD as it was for ChEL***G*** (Figure 3-7B *reducing conditions*). However, unlike ChEL-CD (Figure 3-4C), when all ChEL was glycosylated, covalent dimerization could no longer be detected by nonreducing SDS-PAGE (Figure 3-7B *right*). Together, the data in Figure 3-7 support that introduction of an N-glycan into one helix of the presumptive 4-helix bundle, decreases stability of dimerization of the ChEL domain of Tg.

We noted that the fraction of secretory ChEL-CD that undergoes formation of the optional intersubunit covalent bond (Figure 3-4C) is a lesser fraction than that observed for full-length Tg-CD (Figure 3-5A). This suggested the intriguing possibility that additional information provided by upstream Tg regions I-II-III might improve tail-to-tail dimer alignment or stability. Indeed, ChEL exhibits direct physical interactions with I-II-III [13]. To examine the role of I-II-III in the stabilizing homodimers, we expressed the ChEL***G***-CD construct with or without co-expression of Tg regions I-II-III. As before (Figure 3-4C), covalent homodimers of ChEL-CD could be recovered without assistance from other Tg domains (Figure 3-8A, last lane). And as before (Figure 3-7B), the glycosylated ChEL***G***-CD, expressed by itself exhibited diminished dimer stability as measured by a diminished ability to form the intersubunit covalent bond (Figure 3-8A, second lane). However, upon co-expression of I-II-III, which was co-secreted with ChEL***G***-CD, the carboxyl-terminal intersubunit disulfide bond was restored between ChEL***G***-CD homodimers (Figure 3-8A, third lane). The data indicate that *in trans*, Tg regions I-II-III provide added stability to the homodimerization of the Tg ChEL domain.

To explore the significance of these findings, we introduced the same helix glycosylation site within the context of full-length Tg-CD. As expected, Tg***G***-CD utilized

the additional glycosylation site (seen as a slower Tg monomer mobility by SDS-PAGE, Figure 3-8B). While the Tg-CD control was rapidly secreted with approximately two-thirds of the protein containing the covalent intersubunit bond, Tg***G***-CD was nearly as rapidly secreted and the ratio of covalent to noncovalent dimer was only modestly inhibited (Figure 3-7B). Altogether, the data indicate that while the ChEL domain is necessary and sufficient for Tg dimerization, upstream Tg regions I-II-III, either *in cis* or *in trans*, assist in Tg dimer stability.

**Discussion**

Tg transport through the secretory pathway is essential to make Tg available for iodination in the process of thyroid hormone biosynthesis. The structural features in Tg required for its intracellular transport are beginning to be elucidated. For one thing, the Tg ChEL domain must make physical contact with upstream regions I-II-III in a manner not requiring that ChEL be contained within the contiguous Tg polypeptide [13]. For another, Tg homodimerization is required for export from the ER [14, 15]. In this report, we show that secretory ChEL and Tg are both homodimeric proteins (Figure 3-1, Figure **3-4** to Figure **3-8**) that share predicted alpha helical sequences that closely align with the helices critical for homodimerization of acetylcholinesterase (Figure 3-2). Moreover, the isolated ChEL domain can cross-dimerize with intact Tg (Figure 3-3), suggesting that the ChEL domain encodes the minimal information necessary for Tg dimerization. To be effective, these sequences do not need to reside at the extreme C-terminus of Tg because introduction of downstream GFP, myc, or HA tags still allows dimerization via the ChEL domain to take place (Figure 3-1, Figure **3-3**, Figure 3-7A). While the form of GFP we

employed may have oligomerization potential [42], the fact that Tg-GFP cross-dimerized with Tg-3xMyc indicates that dimerization was triggered by Tg sequences rather than by tag sequences.

For acetylcholinesterase, homodimers brought together via a 4-helix bundle [32] may be further stabilized by an intermolecular disulfide bond engaging one unpaired cysteine that follows shortly after the carboxyl-terminal helical domain of each monomer (Figure 3-4A). Introduction of an unpaired Cys residue immediately after the predicted carboxyl-terminal helical sequence in the ChEL domain also allows for covalent bonding of secretory ChEL-CD homodimers (Figure 3-4C) and full-length Tg-CD (Figure 3-5). The simplicity of the one-dimensional nonreducing SDS-PAGE assay of Tg-CD dimerization makes pulse-chase experiments especially convenient. With this approach, Tg dimerization can clearly be shown to occur before acquisition of endoglycosidase H-resistance, and also in cells in which intracellular transport is blocked by BFA treatment (Figure 3-5A). The data are consistent with our longstanding hypothesis that Tg dimerization occurs before export from the ER [14, 15] and that dimerization increases the efficiency of intracellular Tg transport [41]. With or without the intersubunit disulfide bond, Tg-CD has fully preserved dimerization (Figure 3-6) and exportability (Figure 3-8B), but the covalent intersubunit linkage presumably reflects the stability of the dimer, which yields proper tail-to-tail alignment of the monomer polypeptides.

Although structural biology studies of the Tg ChEL domain have not yet begun, based on the foregoing results we propose that the predicted alpha-helical segments within the ChEL domain of Tg are engaged in forming the contact zone essential to the dimer interface. To weaken this putative contact, we introduced a mutation converting the

**A**-A-**V** sequence of the first alpha helix (Figure 3-2) to an **N**-A-**T** glycosylation acceptor site. Presence of the ChEL***G*** mutation in one of the two subunits significantly decreased co-immunoprecipitation efficiency between the dimerization partners (Figure 3-7A), while presence of the same mutation in all copies of ChEL***G***–CD completely eliminated the intersubunit disulfide bond (Figure 3-7B). While these data support a dimerization mechanism involving helical interactions similar to that for acetylcholinesterase, there are reasons to think that Tg homodimer stability rests on more than merely the four-helix bundle. First, the fraction of Tg-CD that makes an intersubunit disulfide bond (Figure 3-5A) is higher than the fraction of secretory ChEL-CD making the same bond (Figure 3-4C). Second, although Tg region I-II-III itself cannot dimerize (Figure 3-1E), when ChEL***G***–CD is co-expressed with I-II-III, not only is I-II-III secretion rescued, but the intersubunit disulfide bond is once again established between ChEL partners (Figure 3-8A). Thus *in trans*, I-II-III contributes to the dimer stability of ChEL***G***. Third, mutation to add the same glycosylation site in full-length Tg***G***-CD does not block efficient Tg secretion (that depends on dimerization [14, 15]) and only slightly decreases the fraction of Tg***G***–CD making the intersubunit disulfide link (Figure 3-8B). Altogether, the evidence points to a reciprocal relationship, i.e., even as the ChEL domain functions as an intramolecular chaperone and escort for Tg regions I-II-III [13], I-II-III also assists in the stability of homodimerization of the ChEL domain.

Multiple small Tg mutations causing congenital hypothyroidism have been identified within the ChEL domain ([19, 20, 27, 28] with others reviewed in [23]). Both defective intramolecular chaperone function [13] and defective Tg dimerization [30] are an expected consequence of such mutations. Additionally, it seems likely that many

additional mutations in Tg regions I-II-III may fail to provide adequate homodimer stability, which could account for why many if not all cases of hypothyroidism with mutant Tg derive from a failure of intracellular transport through the secretory pathway [25]. Conversely, the present studies provide reason to believe that some of these mutants might not be totally dysfunctional, with some Tg ChEL domain mutations perhaps being intramolecularly suppressed in part by upstream Tg domains (Figure 3-8B) while other I-II-III mutations might still be able to heterodimerize with full-length Tg (Figure 3-3). In summary, intracellular Tg transport for thyroid hormone synthesis engages the ChEL domain not only as an intramolecular chaperone and escort [13], but herein we show that ChEL is also a nidus for Tg homodimerization. While the ChEL domain is both necessary and sufficient for Tg transport through the secretory pathway, ChEL interactions with upstream Tg regions stabilize the homodimer. Structural studies will be needed to elaborate the contact sites between monomeric subunits, while additional molecular dissection is needed to identify the site(s) in ChEL that functionally interact with I-II-III and the site(s) in I-II-III that functionally interact with ChEL.

**Acknowledgements**

Dr. X. Wang contributed to Western blotting experiments (Figure 3-1A-D) and Dr. B. Di Jeso contributed to sucrose velocity gradient work (Figure 3-6). I would like to specially thank both of them. This work was supported by National Institutes of Health Grant DK40344 (to Dr. P. Arvan) and in part by the Ministero dell'Università e della Ricerca PRIN 2006069102_004 (to Dr. B. Di Jeso). I thank members of the Arvan laboratory for helpful suggestions during the course of this work.

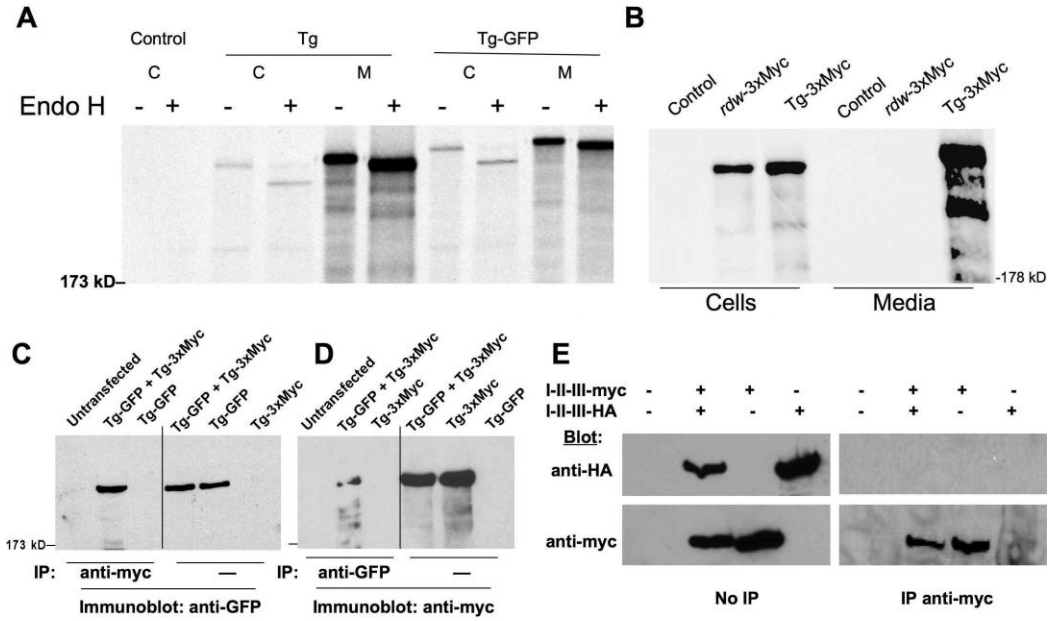


**Figure 3-1. Association of epitope-tagged thyroglobulins.** 293 cells were either untransfected (also called Control), transfected with one epitope-tagged construct, or co-transfected. **(A)** Cells untransfected or transfected with the constructs shown were pulse-labeled for 30 min with $^{35}$S-amino acids and chased for 5 h. Both cells (C) and media (M) were immunoprecipitated with anti-Tg, divided in two equal portions, and either mock-digested ("-") or digested ("+") with endoglycosidase H (Endo H). Tg remaining in the cells was primarily endo H-sensitive. Tg secreted to the medium had a slightly slower mobility, and acquired endo H-resistance (although a portion of the glycans on each Tg molecule remain permanently sensitive to endo H). Tg-GFP migrated more slowly than untagged Tg but otherwise exhibited the same pattern. **(B)** Cells untransfected or transfected with the constructs shown were pulse-labeled for 30 min with $^{35}$S-amino acids and chased for 4 h. Both cells and media were immunoprecipitated with anti-Tg. Tg-3xMyc was efficiently secreted whereas the same construct bearing the *rdw* mutation was blocked intracellularly. **(C)** 293 cells were transfected with the constructs shown and media collected for 24 h. Each sample was divided and analyzed by SDS-PAGE, electrotransfer and immunoblotting with anti-GFP antibody either without (right half of gel) or after immunoprecipitation (IP, left half of gel) with anti-myc antibody. Note that no Tg-GFP is directly immnoprecipitated with anti-myc, whereas Tg-GFP is co-immunoprecipitated when coexpressed with Tg-3xMyc. **(D)** Cells transfected and media collected as in C were analyzed by SDS-PAGE, electrotransfer and immunoblotting with anti-myc either without (right half of gel) or after immunoprecipitation (IP, left half of gel) with anti-GFP. As before, Tg-3xMyc is coprecipitated with Tg-GFP. **(E)** 293 cells were transfected with I-II-III-myc, I-II-III-HA, both, or neither, and the cells were lysed at 48 h in buffer containing 1% NP40 and no SDS. Each sample was divided and analyzed by SDS-PAGE, electrotransfer and immunoblotting with anti-HA (upper panels) and anti-myc (lower panels) either without (left panels) or after immunoprecipitation (IP, right panels) with anti-myc. While I-II-III-myc is readily immunoprecipitated, no co-precipitation of I-II-III-HA can be detected. (A-D were done by Dr. X. Wang)

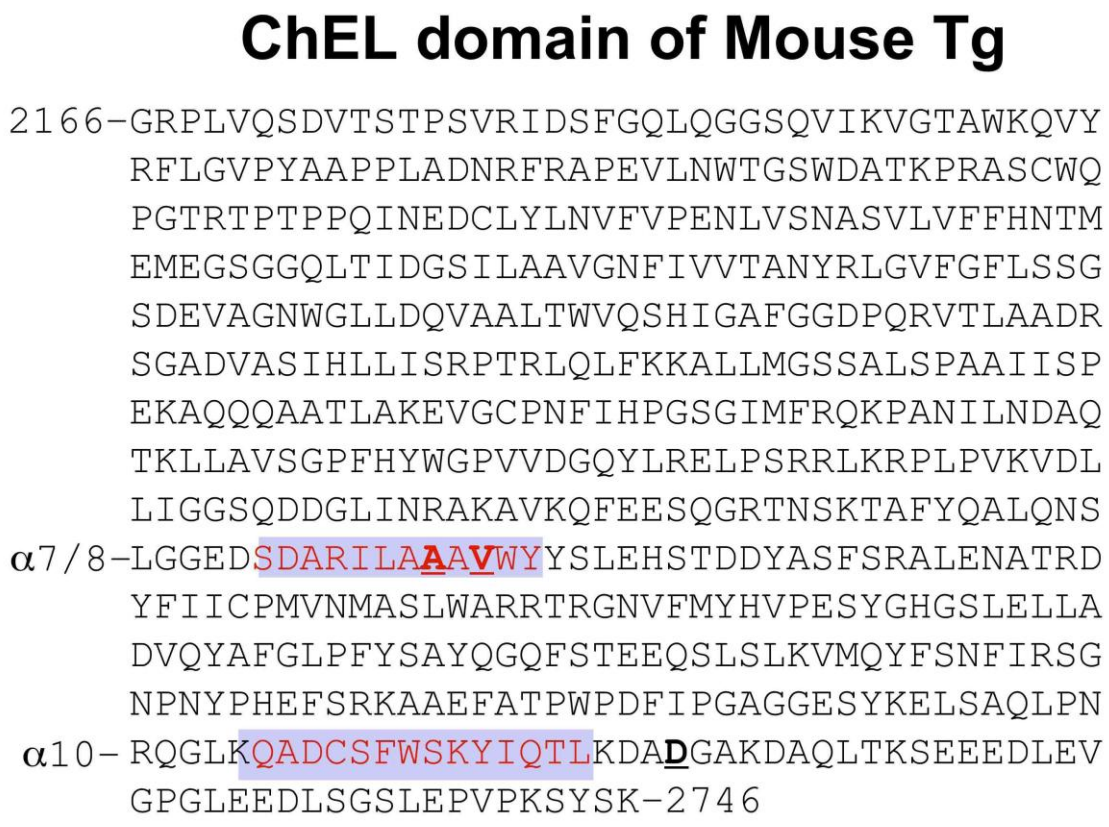


**Figure 3-2. Sequence of the ChEL domain of mouse Tg.** The 581 residue domain extends from residues 2166 to the carboxyl-terminal residue 2746. Based on secondary structural prediction, two alpha-helical segments (highlighted) are predicted to match what is referred to as the α-7/8 helix and α-10 of acetylcholinesterase [32, 34]. Residues that are **bolded** and underlined indicate sites that have been mutagenized. Replacement of the A-A-V sequence with an N-A-T sequence introduced an N-linked glycosylation acceptor site within the α-7/8 helical homology sequence, and the D2708C,G2709stop mutation introduced just downstream of the α-10 helical homology sequence was created to permit formation of an intersubunit disulfide bond.

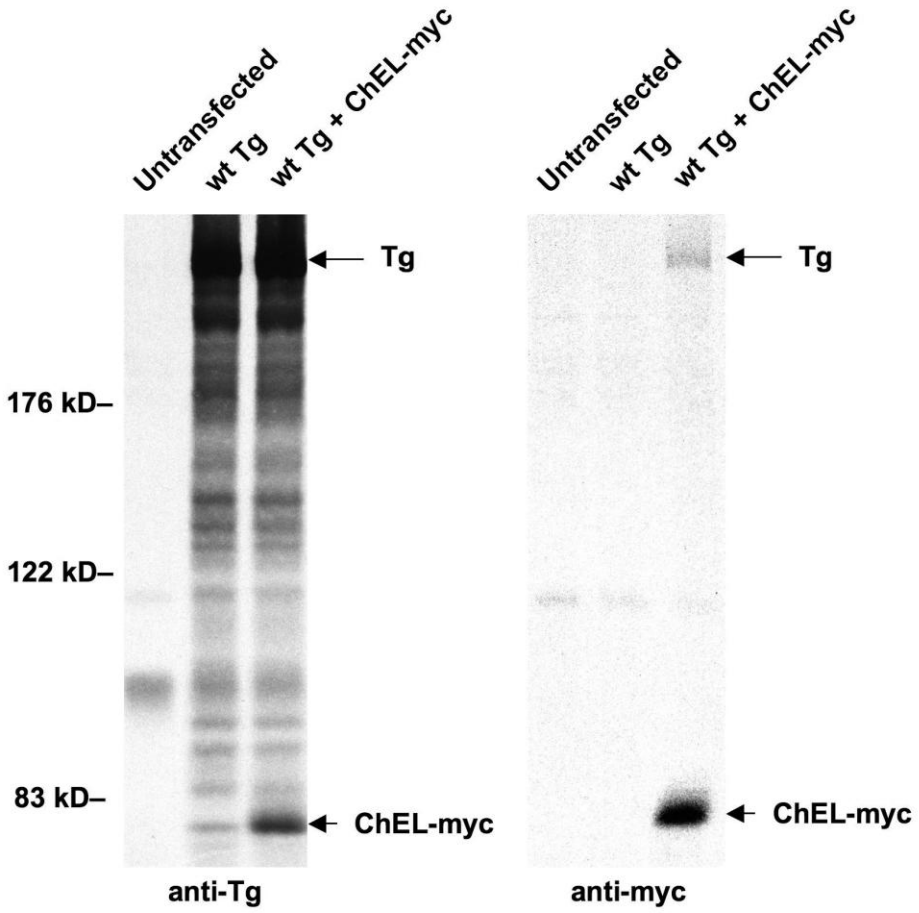


**Figure 3-3. Heterodimerization of the ChEL domain with Tg.** 293 cells were either untransfected or transiently transfected to express wild-type Tg (wt Tg) or that plus secretory ChEL bearing a carboxyl-terminal myc epitope tag (ChEL-myc). Cells were pulse-labeled for 30 min with $^{35}$S-amino acids and chased for 5 h, and media were immunoprecipitated with either anti-Tg (which recovers both proteins, *left panel*) or anti-myc (*right panel*) before SDS-PAGE and fluorography (which, after anti-myc immunoprecipitation, was overexposed ~2-fold). The secretory ChEL-myc protein by itself is known to be rapidly secreted [13]. Immunoprecipitation with anti-myc recovered a fraction of secreted Tg only when co-expressed with secretory ChEL-myc, indicating heterodimerization. The positions of pre-stained molecular weight standards are included (*at left*).

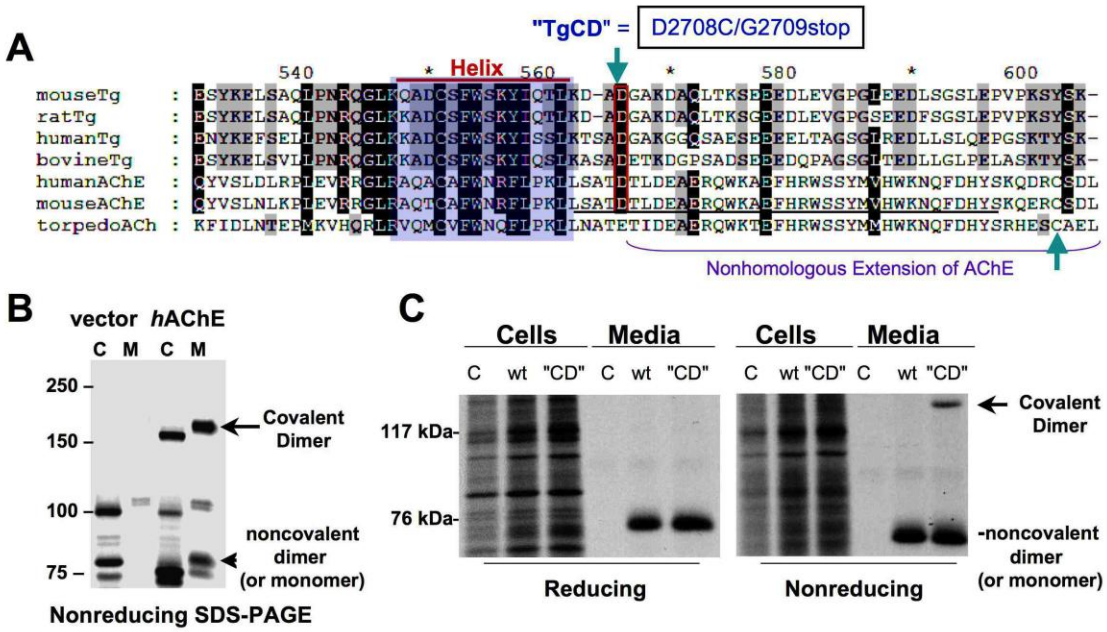


**Figure 3-4. The ChEL domain and authentic acetylcholinesterase (AChE) share capabilities in protein dimerization. (A)** Primary sequence alignment at the carboxyl-terminal end of acetylcholinesterase, indicating a nonhomologous peptide extension that bears an extra, unpaired cysteine (upward arrow) immediately following the α-10 helix that is engaged in homodimerization of acetylcholinesterase. Mouse Tg engineered to terminate with an extra, unpaired cysteine at residue 2,708 (Tg-D2708C,G2709stop = "TgCD") is shown with a downward arrow on the primary sequence. **(B)** 293 cells transiently transfected to express human acetylcholinesterase (*h*AChE) or empty vector were labeled with $^{35}$S-amino acids for 2 h and chased further for 3 h in complete medium before immunoprecipitation of acetylcholinesterase from cell lysate (C) and medium (M) with analysis by nonreducing SDS 8%-PAGE and fluorography. (All bands recovered from cells bearing empty vector represent nonspecific background.) **(C)** 293 cells were either untransfected controls (C) or transiently transfected to express secretory ChEL ("wt", a prolactin signal peptide preceding the ChEL domain of wild-type mouse Tg) or secretory ChEL-CD ("CD"). Cells were pulse-labeled for 30 min with $^{35}$S-amino acids and chased for 4 h, at which time the cell lysates and media immunoprecipitated with a rabbit polyclonal anti-Tg. Immunoprecipitates were analyzed by SDS-PAGE under reducing or nonreducing conditions, as indicated. The positions of covalent ChEL-CD dimer and noncovalent dimer (which runs as monomer by SDS-PAGE) are shown. The positions of pre-stained molecular weight standards are included (*at left*).

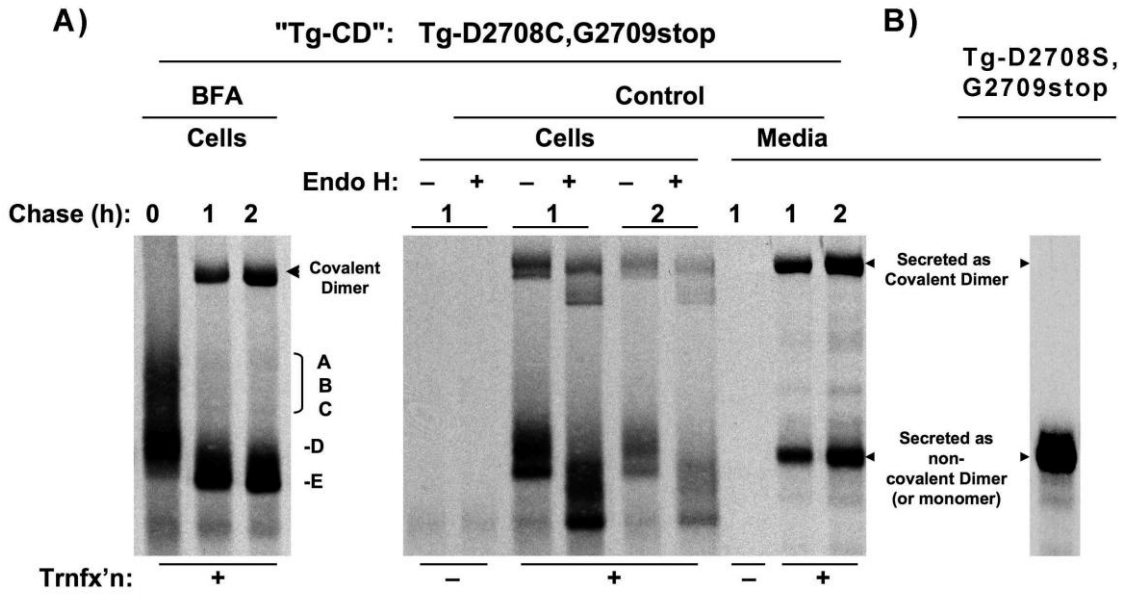


**Figure 3-5. Covalent assembly of Tg-CD indicates tail-to-tail engagement of Tg homodimers. (A)** 293 cells were either mock-transfected (Trnfx'n "-" as indicated at bottom) or transiently transfected with a vector encoding a "Tg-CD" construct containing D2708C,G2709stop. *Left panel*: Cells were pulse-labeled for 30 min with $^{35}$S-amino acids and chased in the presence of brefeldin A (BFA, 5 μg/mL) for the times indicated. Cell lysates were immunoprecipitated with anti-Tg and analyzed by nonreducing SDS 4%-PAGE and fluorography. A smear of high molecular weight bands corresponding to those previously reported to represent Tg adducts with ER oxidoreductases [41] is termed "A, B, and C". A newly-described Tg oxidative folding intermediate "D" [13] is also shown. The positions of covalent Tg-CD dimer and noncovalent dimer (which runs with mature Tg monomer form "E") are shown. *Right panel*: The same experiment but without BFA, in which both cell lysates and chase media were collected at 1 h and 2 h of chase. Immunoprecipitated Tg from each cell lysate was divided into equal portions and either mock-digested or digested with endoglycosidase H before non-reducing SDS-PAGE and fluorography. Intracellular bands near the Tg-CD monomer position are mostly endo H-sensitive (but also include some endo H-resistant Tg-CD that had not made an intersubunit disulfide bond). The intracellular covalent dimers are subdivided into two components: a faster-migrating endo H-sensitive population (indicating dimerization before transport into the Golgi complex) and a slower-migrating population of endo H-resistant Tg-CD. In the media, exclusively the slower-migrating covalent Tg-CD dimer (as well as noncovalent dimer, which runs as monomer) is recovered. **(B)** As a control for the Tg-CD mutation, a D2708S,G2709stop mutant was efficiently secreted but none of the secreted molecules form a covalent dimer.

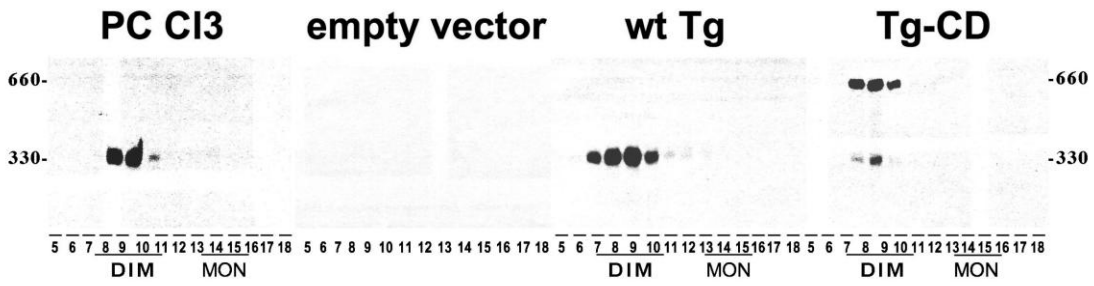


**Figure 3-6. Dimerization of secreted Tg as measured by sucrose velocity gradient centrifugation.** *First gradient*: Endogenous Tg secreted by the metabolically-labeled PC Cl3 thyrocyte cell line. All gradient fractions (collected from the bottom) were immunoprecipitated with anti-Tg antibodies and analyzed by nonreducing SDS-PAGE. The positions of the monomer (MON) and dimer (DIM) peaks are indicated below. *Second gradient*: Media from metabolically-labeled 293 cells transfected with empty vector were immunoprecipitated with anti-Tg and analyzed identically. *Third gradient*: Media from metabolically-labeled 293 cells transfected to express wild-type mouse Tg (wt Tg) were immunoprecipitated with anti-Tg and analyzed identically. Note that all recombinant Tg is secreted in the dimer peak. *Fourth gradient*: Media from metabolically-labeled 293 cells transfected to express Tg-CD were immunoprecipitated with anti-Tg and analyzed identically. Note that all Tg-CD is secreted in the dimer peak, although a portion of these dimers lack the intersubunit disulfide bond. (Sucrose gradient was done by Dr. B. Di Jeso)

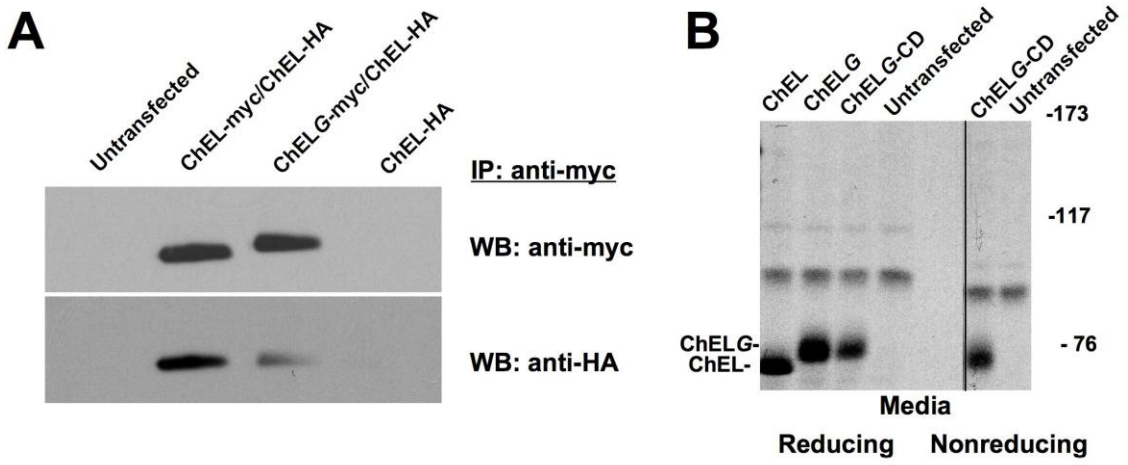


**Figure 3-7. Perturbing ChEL dimer stability by introduction of an N-linked glycosylation site into the α-7/8 helical sequence (Figure 3-2) found to perturb subunit contact in acetylcholinesterase [34].** The **A**-A-**V** to **N**-A-**T** mutagenesis to create an N-linked glycosylation site is highlighted in Figure 3-2. **(A)** 293 cells were either untransfected or transiently transfected to express either secretory ChEL-myc or that bearing the extra glycosylation site (ChEL***G***-myc) in conjunction with an equal amount of plasmid DNA encoding secretory ChEL-HA. Cells expressing ChEL-HA alone were included as negative control. Secretion media were collected for 24 h and immunoprecipitated with anti-myc before SDS-PAGE. Immunoprecipitates were then Western blotted (WB) with either anti-myc (*upper panel*, to demonstrate recovery of ChEL-myc or ChEL***G***-myc) or anti-HA (*lower panel*, to examine the extent of co-precipitation of the dimerization partner). Introduction of an N-glycan slows the mobility of the ChEL-myc band (*upper panel*) and decreases the co-precipitation of ChEL-HA (*lower panel*). **(B)** The ChEL***G*** mutation was introduced into secretory ChEL or secretory ChEL-CD and these constructs were transiently expressed in 293 cells. Secretory ChEL and secretory ChEL***G*** (lacking potential for covalent intersubunit bonding) were included as controls for SDS-PAGE mobility. The cells were metabolically labeled and chased for 4 h; the media were immunoprecipitated with anti-Tg and analyzed by SDS 4%-PAGE under reducing and nonreducing conditions. Note that when ChEL***G***-CD is the sole available partner for dimerization, the protein is still secreted but no covalent dimer can be detected.

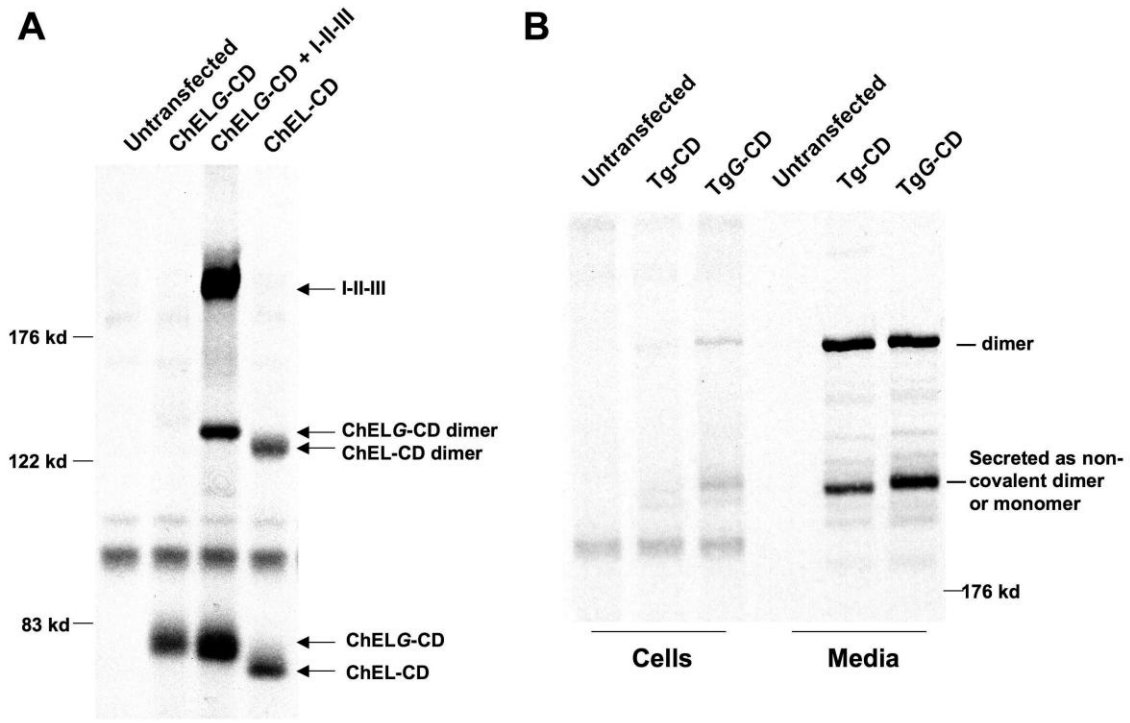


**Figure 3-8. Effect of Tg regions I-II-III on the dimer stability of glycosyated ChEL*G* domain. (A)** 293 cells were transiently transfected either with the secretory ChEL***G***-CD construct alone or cotransfected in the presence of Tg I-II-III. The covalent dimerizing ChEL-CD construct was analyzed in parallel, as a control. Cells were pulse-labeled for 20 min with $^{35}$S-amino acids and chased for 4 h; media was collected and immunoprecipitated with anti-Tg, followed by nonreducing SDS 5.5%-PAGE and fluorography. While secretory ChEL***G***-CD alone produced no covalent homodimer, inclusion of Tg I-II-III increased ChEL***G***-CD secretion with covalent dimerization similar to that of the control, secretory ChEL-CD. **(B)** 293 cells were transiently transfected to express full-length Tg-CD or Tg***G***-CD bearing the extra N-linked glycosylation site in the ChEL domain. Cells were pulse-labeled for 30 min with $^{35}$S-amino acids and chased for 6 h before analysis by Tg immunoprecipitation and nonreducing SDS 4%-PAGE and fluorography. Based on mobility of the Tg***G***-CD band, the extra glycosylation site is clearly used; nevertheless, covalent intersubunit interaction proceeds with only a small decrease in the covalent dimer to monomer ratio.

**References for chapter 3**


1. Arvan, P., and Di Jeso, B. (2005). Thyroglobulin structure, function, and biosynthesis. In The Thyroid, 9th Edition, L.E. Braverman and R.D. Utiger, eds. (Philadelphia, Pennsylvania, USA: Lippincott Williams & Wilkins), pp. 77-95.

2. Taurog, A. (1999). Molecular evolution of thyroid peroxidase. Biochimie *81*, 557-562.

3. Marriq, C., Lejeune, P.J., Venot, N., and Vinet, L. (1991). Hormone formation in the isolated fragment 1-171 of human thyroglobulin involves the couple tyrosine 5 and tyrosine 130. Mol. Cell. Endocrinol. *81*, 155-164.

4. Dunn, A.D., Corsi, C.M., Myers, H.E., and Dunn, J.T. (1998). Tyrosine 130 is an important outer ring donor for thyroxine formation in thyroglobulin. J. Biol. Chem. *273*, 25223-25229.

5. Lamas, L., and Taurog, A. (1977). The importance of thyroglobulin structure in thyroid peroxidase-catalyzed conversion of diiodotyrosine to thyroxine. Endocrinology *100*, 1129-1136.

6. Xiao, S., Dorris, M.L., Rawitch, A.B., and Taurog, A. (1996). Selectivity in tyrosyl iodination sites in human thyroglobulin. Arch. Biochem. Biophys. *334*, 284-294.

7. Veneziani, B.M., Giallauria, F., and Gentile, F. (1999). The disulfide bond pattern between fragments obtained by the limited proteolysis of bovine thyroglobulin. Biochimie *81*, 517-525.

8. van de Graaf, S.A., Ris-Stalpers, C., Pauws, E., Mendive, F.M., Targovnik, H.M., and de Vijlder, J.J. (2001). Up to date with human thyroglobulin. J. Endocrinol. *170*, 307-321.

9. Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S.S., Friedmann, T., and Taylor, P. (1986). Primary structure of Torpedo californica acetylcholinesterase deduced from its cDNA sequence. Nature *319*, 407-409.

10. Swillens, S., Ludgate, M., Mercken, L., Dumont, J.E., and Vassart, G. (1986). Analysis of sequence and structure homologies between thyroglobulin and

acetylcholinesterase: possible functional and clinical significance. Biochem. Biophys. Res. Commun. *137*, 142-148.

11. Mori, N., Itoh, N., and Salvaterra, P.M. (1987). Evolutionary origin of cholinergic macromolecules and thyroglobulin. Proc. Natl. Acad. Sci. U. S. A. *84*, 2813-2817.

12. Arvan, P., Kim, P.S., Kuliawat, R., Prabakaran, D., Muresan, Z., Yoo, S.E., and Abu Hossain, S. (1997). Intracellular protein transport to the thyrocyte plasma membrane: potential implications for thyroid physiology. Thyroid *7*, 89-105.

13. Lee, J., Di Jeso, B., and Arvan, P. (2008). The cholinesterase-like domain of thyroglobulin functions as an intramolecular chaperone. J. Clin. Invest. *118*, 2950-2958.

14. Kim, P.S., and Arvan, P. (1991). Folding and assembly of newly synthesized thyroglobulin occurs in a pre-Golgi compartment. J. Biol. Chem. *266*, 12412-12418.

15. Di Jeso, B., Pereira, R., Consiglio, E., Formisano, S., Satrustegui, J., and Sandoval, I.V. (1998). Demonstration of a Ca2+ requirement for thyroglobulin dimerization and export to the golgi complex. Eur. J. Biochem. *252*, 583-590.

16. Berg, G., Bjorkman, U., and Ekholm, R. (1980). The structure of newly synthesized intracellular thyroglobulin molecules. Mol. Cell. Endocrinol. *20*, 87-98.

17. Berg, G., and Ekholm, R. (1975). Electron microscopy of low iodinated thyroglobulin molecules. Biochim. Biophys. Acta *386*, 422-431.

18. Berg, G., Bjorkman, U., and Ekholm, R. (1980). Conformational change of the thyroglobulin molecule induced by oxidation in vitro. Mol. Cell. Endocrinol. *17*, 139-144.

19. Caputo, M., Rivolta, C.M., Esperante, S.A., Gruneiro-Papendieck, L., Chiesa, A., Pellizas, C.G., Gonzalez-Sarmiento, R., and Targovnik, H.M. (2007). Congenital hypothyroidism with goitre caused by new mutations in the thyroglobulin gene. Clin. Endocrinol. (Oxf). *67*, 351-357.

20. Pardo, V., Rubio, I.G., Knobel, M., Aguiar-Oliveira, M.H., Santos, M.M., Gomes, S.A., Oliveira, C.R., Targovnik, H.M., and Medeiros-Neto, G. (2008). Phenotypic variation among four family members with congenital hypothyroidism caused by two distinct thyroglobulin gene mutations. Thyroid *18*, 783-786.

21. Caron, P., Moya, C.M., Malet, D., Gutnisky, V.J., Chabardes, B., Rivolta, C.M., and Targovnik, H.M. (2003). Compound heterozygous mutations in the thyroglobulin gene (1143delC and 6725G-->A [R2223H]) resulting in fetal goitrous hypothyroidism. J. Clin. Endocrinol. Metab. *88*, 3546-3553.

22. Kitanaka, S., Takeda, A., Sato, U., Miki, Y., Hishinuma, A., Ieiri, T., and Igarashi, T. (2006). A novel compound heterozygous mutation in the thyroglobulin gene resulting in congenital goitrous hypothyroidism with high serum triiodothyronine levels. J. Hum. Genet. *51*, 379-382.

23. Rivolta, C.M., and Targovnik, H.M. (2006). Molecular advances in thyroglobulin disorders. Clin. Chim. Acta *374*, 8-24.

24. Matakidou, A., Hamel, N., Popat, S., Henderson, K., Kantemiroff, T., Harmer, C., Clarke, S.E., Houlston, R.S., and Foulkes, W.D. (2004). Risk of non-medullary thyroid cancer influenced by polymorphic variation in the thyroglobulin gene. Carcinogenesis *25*, 369-373.

25. Vono-Toniolo, J., Rivolta, C.M., Targovnik, H.M., Medeiros-Neto, G., and Kopp, P. (2005). Naturally occurring mutations in the thyroglobulin gene. Thyroid *15*, 1021-1033.

26. Kim, P.S., Ding, M., Menon, S., Jung, C.G., Cheng, J.M., Miyamoto, T., Li, B., Furudate, S., and Agui, T. (2000). A missense mutation G2320R in the thyroglobulin gene causes non-goitrous congenital primary hypothyroidism in the WIC-rdw rat. Mol. Endocrinol. *14*, 1944-1953.

27. Hishinuma, A., Furudate, S., Oh-Ishi, M., Nagakubo, N., Namatame, T., and Ieiri, T. (2000). A novel missense mutation (G2320R) in thyroglobulin causes hypothyroidism in rdw rats. Endocrinology *141*, 4050-4055.

28. Kim, P.S., Hossain, S.A., Park, Y.N., Lee, I., Yoo, S.E., and Arvan, P. (1998). A single amino acid change in the acetylcholinesterase-like domain of thyroglobulin causes congenital goiter with hypothyroidism in the cog/cog mouse: a model of human endoplasmic reticulum storage diseases. Proc. Natl. Acad. Sci. U. S. A. *95*, 9909-9913.

29. Park, Y.N., and Arvan, P. (2004). The acetylcholinesterase homology region is essential for normal conformational maturation and secretion of thyroglobulin. J. Biol. Chem. *279*, 17085-17089.

30. Kim, P.S., Kwon, O.Y., and Arvan, P. (1996). An endoplasmic reticulum storage disease causing congenital goiter with hypothyroidism. J. Cell Biol. *133*, 517-527.

31. Reddy, P.S., and Corley, R.B. (1998). Assembly, sorting, and exit of oligomeric proteins from the endoplasmic reticulum. Bioessays *20*, 546-554.

32. Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. Science *253*, 872-879.

33. Bourne, Y., Taylor, P., and Marchot, P. (1995). Acetylcholinesterase inhibition by fasciculin: crystal structure of the complex. Cell *83*, 503-512.

34. Morel, N., Leroy, J., Ayon, A., Massoulie, J., and Bon, S. (2001). Acetylcholinesterase H and T dimers are associated through the same contact. Mutations at this interface interfere with the C-terminal T peptide, inducing degradation rather than secretion. J. Biol. Chem. *276*, 37379-37389.

35. Seed, R.W., and Goldberg, I.H. (1965). Iodination in relation to thyroglobulin maturation and subunit aggregation. Science *149*, 1380-1382.

36. Seed, R.W., and Goldberg, I.H. (1965). Biosynthesis of thyroglobulin. II. Role of subunits, iodination, and ribonucleic acid synthesis. J. Biol. Chem. *240*, 764-773.

37. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195-202.

38. Velan, B., Grosfeld, H., Kronman, C., Leitner, M., Gozes, Y., Lazar, A., Flashner, Y., Marcus, D., Cohen, S., and Shafferman, A. (1991). The effect of elimination of intersubunit disulfide bonds on the activity, assembly, and secretion of recombinant human acetylcholinesterase. Expression of acetylcholinesterase Cys-580>Ala mutant. J. Biol. Chem. *266*, 23977-23984.

39. Kerem, A., Kronman, C., Bar-Nun, S., Shafferman, A., and Velan, B. (1993). Interrelations between assembly and secretion of recombinant human acetylcholinesterase. J. Biol. Chem. *268*, 180-184.

40. Gough, N.R., and Randall, W.R. (1995). Oligomerization of chicken acetylcholinesterase does not require intersubunit disulfide bonds. J. Neurochem. *65*, 2734-2741.

41. Di Jeso, B., Park, Y.N., Ulianich, L., Treglia, A.S., Urbanas, M.L., High, S., and Arvan, P. (2005). Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase. Mol. Cell. Biol. *25*, 9793-9805.

42. Jain, R.K., Joyce, P.B., Molinete, M., Halban, P.A., and Gorr, S.U. (2001). Oligomerization of green fluorescent protein in the secretory pathway of endocrine cells. Biochem. J. *360*, 645-649.

# CHAPTER 4

# DISCUSSION AND FUTURE DIRECTIONS

The functional and structural importance of thyroglobulin (Tg), which has evolved as a precursor for thyroid hormone synthesis, has begun to emerge after years of biochemical and cell biological studies. It is interesting to consider the complexity of thyroid hormonogenesis, i.e., how many players are involved to make, deliver, and activate thyroid hormone. The thyroid gland, the sole source of thyroid hormone production, requires multiple reactions; 1) iodide transfer by a series of iodide transporters, 2) iodide storage and thyroid hormone synthesis within the Tg molecule, which requires Tg secretion, iodination, and coupling reaction to produce iodothyronines, and 3) endocytosis and proteolysis of Tg to release thyroid hormone to the blood stream.

Tg has some unique features including its large size (330 kDa, monomer) that proceeds to homodimerize (660 kDa). Tg consists of ~2,750 amino acid residues (for monomer) although thyroid hormone is made by conjugation of only two di-iodinated tyrosyl residues. The major hormonogenic site in Tg is limited to the N-terminus of the molecule (Tyr5 – Tyr130). Tg is not an enzyme nor a re-usable molecule, therefore the thyroid gland makes this large protein in order to produce only a couple of thyroid hormones per Tg molecule. Additionally, iodination for the purpose of thyroid hormone formation occurs at the apical surface outside of thyroid cells, with massive storage of

iodinated Tg in the follicular space of thyroid glands. Therefore, production of thyroid hormone is totally dependent upon Tg delivery via the secretory pathway.

I have considered the possibility that the C-terminal region of Tg may serve a primary function to assist in Tg folding and intracellular transport. Previously, mutations in the ChEL domain of Tg (*cog* and *rdw*) have been described to cause hypothyroidism by preventing secretion of Tg [1, 2]. Further, the Tg N-terminal region (I-II-III) expressed without the ChEL domain was reported to fail in secretion [3]. These findings lend support to the thesis that ChEL assists in the intracellular protein transport of Tg by facilitating the formation of a proper native structure of Tg involving reciprocal interactions between ChEL and upstream regions of Tg.

**The ChEL domain functions as an intramolecular chaperone and escort for Tg folding and secretion**

Tg primary structure consists of many repeating units. The exon-intron boundaries and the striking homology of Tg-type 1 repeats, as well as that of the ChEL domain, suggests that Tg structure has evolved by duplication of genes. Domains translated by duplicated genes are generally believed to keep their native structures, and there are reasons to expect that each unit in Tg has its own internal structure (and performs its own folding). The independently folding domains of Tg are expected to start their folding co-translationally with the help of ER chaperones [4, 5]. In GA733-2 (human gastrointestinal carcinoma antigen) and other proteins that contain homologous type 1 Tg repeats, cysteines form consecutive intrachain disulfide bonds within the homologous region [6]. These cysteines are highly conserved and believed to function in

maintaining the compact structure of type 1 repeat-containing proteins. N-terminal fragments of Tg that contain the first three type 1 repeat units can pass ER quality control and be secreted from cells [7]. The ChEL domain also has six conserved cysteines that are important to the native structure of acetylcholinesterase family members (Figure 1-3). Other amino acid residues comprising the backbone structure of acetylcholinesterase are well preserved in the ChEL domain of Tg (Figure 1-4A). All evidence indicates that ChEL forms an independently folded domain within Tg, because its replacement by authentic acetylcholinesterase results in a chimera with active cholinesterase function [3]. Moreover, the ChEL domain (led by a signal peptide) rapidly and efficiently exits the ER and is secreted (Figure 2-4). Regions II-III with its own signal peptide also are well secreted (Figure 4-1B), indicating that these two regions can also fold as a successful unit, suggesting these two regions are unlikely to require assistance from ChEL.

Instead of a linear, string-like structure, Tg is thought to have an oval-like native structure [8]. From this, it is certainly reasonable to hypothesize that some Tg domains may get help from interaction with other domains in order to facilitate folding. Limited proteolysis of Tg also provides clues to solvent-exposed regions versus other regions that are likely to be buried [9, 10]. Understanding the three dimensional structure of Tg could help to explain the epitopes that may help to trigger autoimmune thyroiditis [11].

Final formation of native Tg structure occurs post-translationally. Tg uses most of its 122 cysteines in disulfide pairings. Interchain disulfide bonding has not been described in newly secreted Tg. Therefore, most cysteines form intrachain disulfide bonds (a maximum of 61-intrachain disulfide bonds is possible). The domain structure of Tg readily accounts for most of these. New assays were derived to examine post-

translational maturation of Tg as described in chapters 2 and 3. In order to follow the folding progress of Tg, I used a pulse-chase paradigm and analyzed results by nonreducing SDS-PAGE [4]. For secretory proteins, disulfide maturation has been described as a rate-limiting step in folding process within the ER [12]. Therefore, Tg requires post-translational time to form all of its disulfide bonds, and proper formation of these bonds is tightly correlated with folding progress.

Immediately after its translation, Tg has not yet reached its fully mature form (known as "E") by nonreducing SDS-PAGE (Figure 2-3). Instead, nascent Tg is in complexes with ER oxidoreductases forming bands "A, B, and C" [4] as well as the immature monomer known as "D" (Figure 2-3). In general, only after Tg takes additional conformational maturation steps to the "E" form and also takes steps to form homodimers (chapter 3) can it then be secreted from cells (Figures 2-3 and 3-5).

In chapter 2, one of the hypothesis tested is that global maturation of Tg requires the ChEL domain as an intramolecular chaperone. One possible reason for the ER retention of *cog* Tg and *rdw* Tg is the failure of disulfide maturation from "D" to "E" forms as shown in Figure 2-2C. By contrast, wild-type Tg always progresses to the "E" form prior to its secretion. Misfolding of the ChEL domain caused by *cog* or *rdw* mutation might affect global folding as evidenced by its remarkable reactivity with AMS (Figure 2-1). The influence of the ChEL domain on upstream regions of Tg is further supported by the fact that Tg regions I-II-III alone are incompetent for secretion [3], while that secretion can be rescued by secretory ChEL provided *in trans* (Figure 2-5C). Although Tg regions I-II-III repeat domains are expected to fold independently, they may still need further structural arrangement that depends on the ChEL domain. By

nonreducing SDS-PAGE, I-II-III alone shows unsuccessful progression of its disulfide maturation (Figure 2-8A), persisting as a "D"-like form. This is similar to the arrested progression of full-length Tg bearing mutations in the ChEL domain. Indeed, secretory ChEL *in trans* physically interacts with upstream regions of Tg (Figure 2-6A) and allows I-II-III to progress in disulfide maturation, reaching an "E"-like form (Figure 2-8A). A technical point is that the best nonreducing SDS-PAGE analysis requires removal of sugars by PNGase F (Figure 2-9).

More than 70% of eukaryotic proteins are multidomain proteins [13]. Proteins (domains) are expected to assume their native structure by virtue of their own thermodynamic stabilities [14]. However, cooperative interaction among domains and co-dependency for efficient folding have been described [13, 15]. In this thesis research, I propose that Tg is another example in which interdomain interaction is required for efficient Tg folding and it is mediated by the C-terminal ChEL domain acting as an intramolecular chaperone.

Curiously, the ChEL domain has an even stronger affinity for well-folded I-II-III than it does for immature I-II-III (Figure 2-10). This binding seems essential for the Tg secretory process. Secretory ChEL with an engineered ER-retention motif functions normally as a molecular chaperone (Figure 2-8) yet co-retains the mature I-II-III within the ER-Golgi system (Figure 2-11). Thus, the ChEL domain also functions as a molecular escort [16] required to convey Tg I-II-III out of the ER.

The intramolecular chaperone and escort functions of the ChEL domain are specific and dependent upon structural features of the ChEL domain. Misfolding mutations in secretory ChEL such as *cog* or *rdw* prevent upstream regions of Tg (I-II-III)

to progress through the normal stages of disulfide bond maturation (Figure 2-8B) or to be secreted (Figure 2-11B) even though they can physically interact with regions I-II-III (not shown). Moreover, authentic acetylcholinesterase cannot substitute for the ChEL domain even though it is similar to ChEL in structure. The chimeric protein of Tg in which regions I-II-III connect contiguously with human acetylcholinesterase cannot be secreted even though I can confirm that acetylcholinesterase in the chimeric protein has proper folding based on its enzymatic activity (not shown). Acetylcholinesterase can bind physically with Tg regions I-II-III, but it fails to rescue of secretion of I-II-III (not shown) which contradicts a previous report [3]. Disulfide maturation to the "E" form has not been observed in Tg-acetylcholinesterase chimeras, or in the isolated Tg regions I-II-III when acetylcholinesterase is expressed *in trans* (Wang, X., unpublished data).

In summary of chapter 2, I propose that Tg structural maturations and intracellular transport for thyroid hormone synthesis are dependent on intramolecular chaperone and escort functions embedded within the Tg ChEL domain.

**The ChEL domain and Tg dimerization**

After formation of a compact Tg monomer, Tg proceeds to homodimerization which is required for export from the ER [17, 18]. It is notable that of the established members of the cholinesterase-like family of proteins, acetylcholinesterase and neuroligins are known to form homodimers using the same contact zone involving a four-helix bundle (Figure 1-4B) [19-23]. Evidence suggests that dimerization via this conserved mechanism is essential for the diverse functions of these proteins [24, 25]. In chapter 3, I showed that secretory ChEL and Tg are also both homodimeric proteins

(Figures 3-1, 3-4 to 3-8) sharing predicted alpha helical sequences that closely align with the helices critical for homodimerization of acetylcholinesterase and neuroligins (Figure 3-2).

For acetylcholinesterase, homodimers through the four-helix bundle may be further stabilized by an intermolecular disulfide bridge engaging one unpaired cysteine that follows shortly after the carboxyl-terminal helical domain of each monomer (Figure 3-4A). This allowed me to assay dimerization by formation of a double molecular mass band mobility upon nonreducing SDS-PAGE. Introduction of an unpaired Cys residue immediately after the predicted carboxyl-terminal helical sequence in the ChEL domain allowed for covalent bonding of secretory ChEL-CD homodimers (Figure 3-4C) and full-length Tg-CD (Figure 3-5). The formation of interchain disulfide bonds in secretory ChEL-CD and Tg-CD specifically indicates the formation of a four-helix bundle because none of the cysteines in secretory ChEL (6 Cys residues) forms an interchain disulfide bond as they are each already engaged in one of the three evolutionarily conserved intrachain disulfide bonds. The interchain disulfide made by full-length Tg-CD undoubtedly uses the same Cys residues as the extra cysteine embedded in the ChEL domain. Therefore, ChEL preserves its evolutionarily conserved dimerization, and Tg uses this to stabilize the Tg homodimer. Further, introduction of glycosylation site into one of the predicted helices for dimerization of secretory ChEL (ChEL***G*** or ChEL***G***-CD) decreased the ability of such a protein to interact with the wild-type partner (Figure 3-7A) and completely eliminated the intersubunit disulfide bond in ChEL***G***-CD (Figure 3-7B).

Using a combination of methodologies including pulse-chase analysis in conjugation with nonreducing SDS-PAGE, I observed that Tg dimerization appears

immediately after full maturation of Tg monomers ("E" form) and occurs early, before Golgi arrival (Fig 3-5A). These data are consistent with previous studies of thyroid cells, and support the longstanding hypothesis that Tg dimerization occurs before export from the ER [17, 18].

I have been unable to detect any dimeric interaction between regions I-II-III lacking the ChEL domain (Figure 3-1E), whereas the isolated ChEL domain can cross-dimerize with intact Tg (Figure 3-3). These data suggest that the ChEL domain encodes the minimal information necessary for Tg dimerization. However, Tg homodimer stability appears to depend on more than the four-helix bundle within the ChEL domain, because the presence of I-II-III enhanced the formation of the interchain disulfide bond in ChEL-CD (Figures 3-4C and 3-5A) and overcame the inhibitory effect of the glycosylation mutation on ChEL-CD dimerization, either *in cis* or *in trans* (Figures 3-8). Altogether, the evidence points to an interesting reciprocal structural relationship, i.e., the ChEL domain functions as an intramolecular chaperone for Tg regions I-II-III (chapter 2, [26]) while I-II-III assists in the ChEL domain homodimer stability (chapter 3).

In summary, intracellular Tg transport for thyroid hormone synthesis engages structural maturation of Tg in which the ChEL domain not only functions as an intramolecular chaperone for monomer folding (chapter 2, [26]), but also as a nidus for Tg homodimerization (chapter 3) and an escort that conveys I-II-III out of the ER (chapter 2). The quaternary structure of Tg requires reciprocal interaction between upstream regions of Tg and the ChEL domain facilitating I-II-III folding and stabilizing the ChEL-based homodimer (chapter 3) to ready the protein for intracellular transport.

It is interesting to consider the evolutionary dependence of Tg on the C-terminal ChEL domain, given that the major hormonogenic site of Tg is at its extreme N-terminus [27]. It might be expected that a highly truncated Tg with the N-terminal thyroid hormone-forming site [28, 29] would allow hormone production without the need for intramolecular chaperone or escort functions. Indeed, an N-terminal portion of Tg with hormonogenic tyrosines can be secreted [7] and thyroid hormones (T4 and T3) can theoretically be produced from such a truncated Tg [30]. However, an N-terminal fragment of Tg is not enough to allow animals to survive in an environment with low iodine availability [31]. Therefore, environmental circumstances (such as low iodide availability in freshwater and on land) might have applied evolutionary pressure for development of a more complex Tg molecule capable of iodide storage in vertebrate organisms [32]. It is also consistent with the notion that an early ancestor of land dwelling animals may have been fish designed to live in freshwater in which iodide is lacking. Modern fish have similar sized Tg-like iodoproteins [33, 34] (note the comment on zebrafish Tg in chapter 1).

Bigger proteins such as Tg may face greater problems of misfolding and protein aggregation. Tg may have adopted the ChEL domain in order to minimize misfolding and enable intracellular transport. Among the reported inactivating mutations of Tg (Table 1-1), many are reported to cause a mild disease phenotype [7, 35]. The R19K mutation causes misfolding in the first type 1 repeat and almost completely prevents secretion of an N-terminal fragment of Tg (C175stop). However full Tg with R19K shows only midly impaired secretion [7], consistent with the less severe phenotype of Tg-R19K in patients. Another mutation, Tg-C1977S (C1976S in mouse Tg) causes mild hypothyroidism [35],

and mouse Tg bearing the C1976S mutation still can follow the secretory pathway, albeit with less efficiency (Figure 4-1A, also in [7, 36]). However, secretory II-III bearing the C1976S mutation prevents proper disulfide bonding and completely blocks its secretion (Figure 4-1B). Therefore, the detrimental effect of local domain-specific misfolding in Tg regions I-II-III may be at least partially suppressed in the context of intact Tg (By contrast, local misfolding in the ChEL domain has devastating impact on Tg secretion, as seen in *cog* mice or *rdw* rats [5, 37]). These findings suggest that chaperone function of the ChEL domain helps to protect Tg from potential misfolding caused by upstream mutations. This clearly speaks to the evolutionary importance of maintaining the ChEL domain in full-length Tg.

**Future directions**

**(1) Molecular dissection to identify interacting site(s) in I-II-III and ChEL**

The oxidative maturation and secretion of Tg regions I-II-III are rescued by secretory ChEL, which I believe simulates the intramolecular interaction within intact Tg monomers. In support of this view, the same "D" to "E" maturation occurs in both intact Tg and regions I-II-III (with ChEL *in trans*). It appears that regions I-II-III and secretory ChEL physically interact specifically since the binding stoichiometry between I-II-III and ChEL has been measured by me and is around 1.13 ± 0.27, i.e., the same as when they are contiguously connected in intact Tg. Thus, *in trans*, the interaction between I-II-III and ChEL recapitulates the domain interactions within intact Tg.

What areas in regions I-II-III and ChEL are involved in these interactions? First, co-immunoprecipitation between secretory ChEL and each individually-expressed region

from I-II-III should be performed to identify the interacting regions. Thus far in preliminary experiments, I could not co-immunoprecipitate secretory II-III with ChEL (not shown). So I do not believe that ChEL binds in those regions. In addition, as has already been mentioned, secretory II-III can be efficiently secreted without help of ChEL, and its secretion is not affected by co-expression of secretory ChEL (not shown). These findings suggest that region I, or a hinge region between the repeats, is likely to contain the binding site most important to the interaction with the ChEL domain.

For ChEL itself, alanine scanning could be used to identify potential interacting site(s), since the ChEL domain's globular structure may make it very difficult to use progressive truncation as an experimental approach. Alternatively, ChEL-acetylcholinesterase chimeras could be tried, or neuroligin 1 [25] could be used to explore which conserved (or unconserved) sequences may be involved in the rescue of secretion of I-II-III.

**(2) Disulfide maturation of the Tg monomer**

Considering that the repeat units and the ChEL domain are believed to be independent folding units, the progression of disulfide maturation seen in the "D" to "E" transition by nonreducing SDS-PAGE, which depends on the ChEL domain, is likely to occur as a consequence of posttranslational interactions between distinct Tg regions. Tg has some cysteine-rich areas with no known homology to other proteins in the database. Since fully folded and secreted Tg has no recognizable reaction with alkylating agents such as AMS (Figure 2-1) or PEG-maleimide, the cysteines in the non-homologous hinge regions are believed to also form intrachain disulfide bonds. Conceivably, disulfide bonds

in the non-homologous region could drive the “D” to “E” transition. Sequential mutations of these relatively few non-homologous cysteines could be tried in order to identify the critical disulfide bonds as analyzed by nonreducing SDS-PAGE.

In preliminary experiments, a Tg-C1245R mutation in the non-homologous hinge region, known to cause mild hypothyroidism [36], appears to be one of the potential cysteines to form intrachain disulfide bonds to stabilize the final mature “E” form (Figure 4-2B). Given that mutagenesis of cysteines in other non-homology areas does not prevent the final disulfide maturation (Figure 4-2B), the area after the 10th type 1 repeat, which includes C1245R, may contribute to the shift from “D” to “E” as seen by nonreducing SDS-PAGE. Besides C1245R, at least one other cysteine in this area may contribute to this critical maturation step. Such a view is indeed consistent with the sequence after the 10th type 1 repeat functioning as a true hinge region in the Tg molecule. It may be of interest to mutagenically delete this hinge region, either in full Tg or in I-II-III (in presence of ChEL *in trans*) to determine if this renders Tg refractory to disulfide maturation.

Characterization of I-II-III interactions with ChEL has potential importance for understanding Tg homodimerization. I-II-III alone appears to have no intrinsic propensity to dimerize, but I-II-III enhances ChEL dimerization (also Tg dimerization). Tg dimerization first begins when Tg monomer has matured to the “E” form as seen by nonreducing SDS-PAGE. Thus ChEL may help I-II-III fold, which in turn helps ChEL to make a more stable homodimer. It has been suggested previously that some area in region I (which includes the non-homologous hinge region) might also contribute intermolecular interaction [10].

**(3) Further understanding the role of ER oxidoreductases in Tg folding and secretion**

Secretory proteins require ER chaperones and oxidoreductases for their proper folding and efficient secretion. Disulfide maturation, especially for Tg may be a rate-limiting step in the process [12]. Considering that Tg shows early covalent interactions with ER oxidoreductases [4], it is expected that ER oxidoreductases contribute significantly to the disulfide maturation of Tg. Also, ERp29 has been reported to serve as an escort for Tg [38, 39], just as an ERp29-homologue in fly, Wind, does for Pipe, a glycosaminoglycan-modifying enzyme [40].

Knockdown of these ER oxidoreductases using siRNA could be applied to understand their role in Tg disulfide maturation and secretion. The nonreducing SDS-PAGE approach would allow the observation of changes in Tg maturation as a consequence of oxidoreductase knockdown. In a preliminary experiment, an siRNA against human PDI (designed by G. Rajpal in our laboratory and in Dr. B. Tsai's laboratory) efficiently decreased endogenous PDI in 293 cells ($\geq$ 70% knockdown) and this improved wild-type Tg secretion (not shown). Considering that overexpression of PDI facilitates the degradation of misfolded Tg [41], PDI may work by preventing premature secretion of incompletely folded secretory proteins. This behavior is consistently found in BiP [42] and reported for other ER chaperones.

Molecular dissection of Tg via the expression of discrete regions could give further understanding of the role of ERp29. In preliminary experiments, I did not observe any interaction between secretory ChEL and ERp29 as judged by co-immnopreciptation,

not shown) while I could see interaction between full-length Tg and ERp29 (not shown). Does this mean that ERp29 may bind upstream regions of Tg? I suspect that this might be the case. As seen in Figure 4-3, inhibition of ERp29 function by overexpression of a dominant negative form of ERp29 (NTD ERp29) begins to allow an increased fraction of I-II-III to be secreted (asterisk in Figure 4-3). Currently, it is too early to explain the nature of the ERp29 interaction with I-II-III and the role in which ERp29 is involved in physiological Tg secretion. Co-immunoprecipitation between Tg regions I-II-III and ERp29 and studies of disulfide maturation of I-II-III should be further pursued in order to understand the role of ERp29 in Tg maturation and secretion for thyroid hormone synthesis.

**Acknowledgements**

The preliminary work shown in this chapter was supported by National Institutes of Health Grant DK40344 (to Dr. P. Arvan). I would like to thank Drs. S. Mkrtchian (Karolinska Institute, Sweden) for anti-ERp29 antibody and the plasmid encoding wild-type ERp29. I also would like to thank Dr. B. Tsai, E. Rainey-Barger, and M. Forster for providing the plasmid encoding the N-terminal domain of ERp29 (NTD ERp29) and for helping to design the siRNA for human PDI. I thank G. Rajpal for helping in the siRNA design and conducting the experiment using siRNA. I thank members of the Arvan laboratory for helpful suggestions during the entire course of this thesis project.

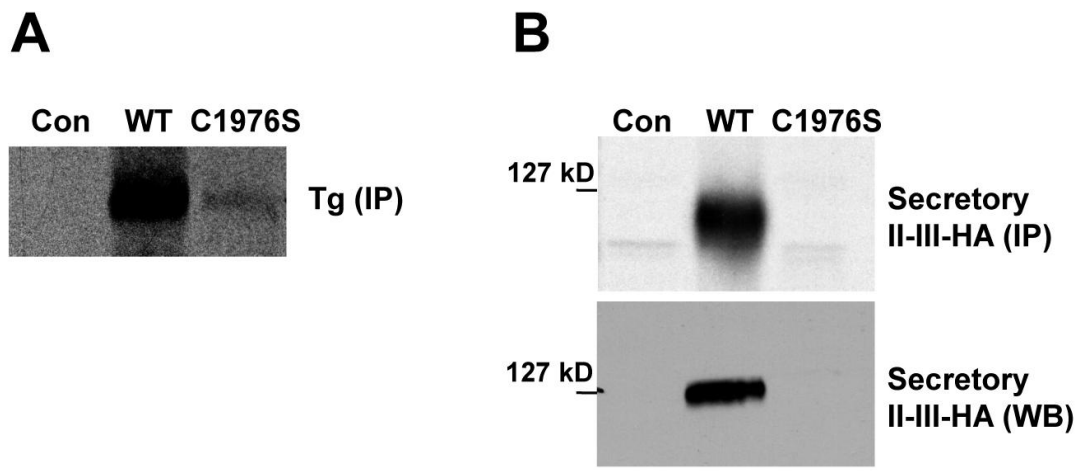


**Figure 4-1. Differences in mutational phenotype of C1976S for secretion of Tg and secretory II-III. (A)** 293 cells were either untransfected (control, [Con]) or transiently transfected with wild-type Tg or Tg C1976S. Cells were pulse-labeled with $^{35}$S-amino acids for 20 min and chased for 4 h in complete media. Media were immunoprecipitated with anti-Tg before SDS-PAGE and fluorography. Tg C1976S has been reported to secret less efficiently than wild-type Tg [7, 36]. **(B)** 293 cells were either untransfected (control, [Con]) or transiently transfected with HA tagged secretory II-III or secretory II-III with C1976S mutation (also tagged with HA). *Top panel*: Cells were pulse-labeled with $^{35}$S-amino acids for 30 min and chased for 5 h in complete media. Media were immunoprecipitated with anti-HA, then analyzed by SDS-PAGE and fluorography. *Bottom panel*: Media were collected for 24 h and analyzed by SDS-PAGE, electrotransfer, and immunoblotting with anti-HA antibody. Note that no detectable secretory II-III with C1976S was found in the media even though it was well expressed in cells (not shown).

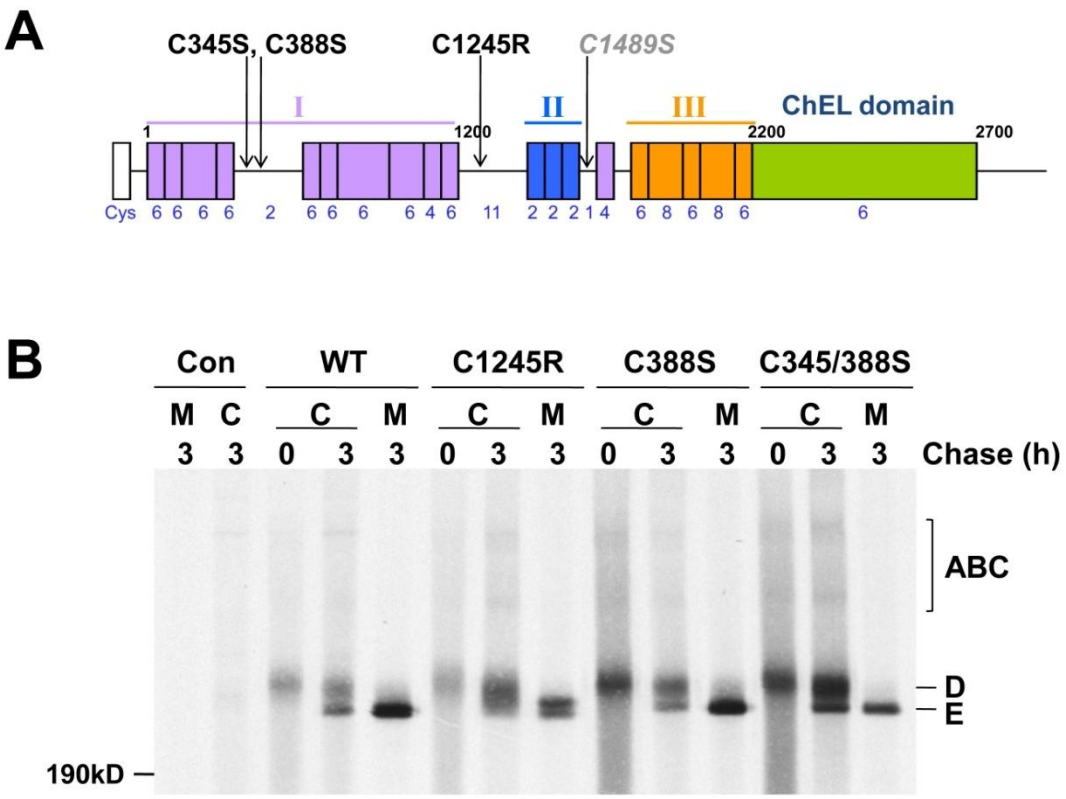


**Figure 4-2. Intrachain disulfide bond mediating final maturation ("D" to "E") of Tg.** **(A)** Tg structure shown with cysteines. Tg has total 122 cysteine and most of them are believed to form intrachain disulfide bonds. Each numbering indicates the number of cysteines in subunits, ChEL domain, and also unknown areas. Some of cysteines in such unknown areas are expected to form intrachain disulfide bonds and stabilize final native structure of Tg. Sequential mutations can be tried to identify such disulfide bonds. The mutations tried preliminarily are shown (down arrow). **(B)** 293 cells were either untransfected (control, [Con]) or transiently transfected with wild-type Tg, Tg C1245R, Tg C388S, or Tg C345/388S. Cells were pulse-labeled with $^{35}$S-amino acids for 20 min and chased in complete media. Cells and media were immunoprecipitated with anti-Tg and treated with PNGase F under nonreducing condition before nonreducing SDS-PAGE and fluorography. Tg C1489S has been tried previously and can reach final disulfide matured form, "E" (not shown). Note that Tg C1245R shows decreased "E" form which was either inhibited to form "E" or could form but not stabilized under SDS-PAGE and can be both.

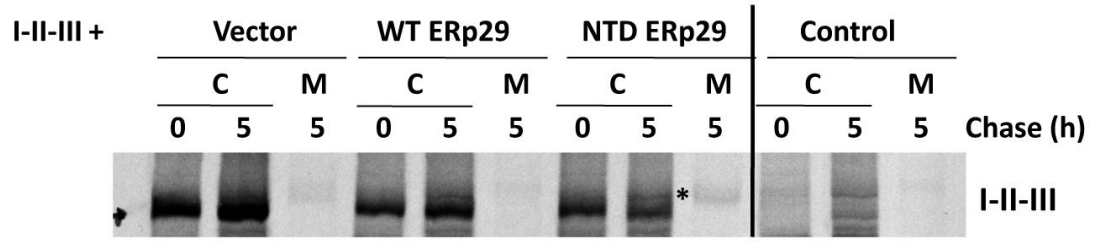


**Figure 4-3. Enhanced secretion of Tg regions I-II-III by suppression of ERp29.** 293 cells were either untransfected (control) or transiently co-transfected with 1µg of plasmid encoding Tg regions I-II-III with 2µg of either blank plasmid or plasmid encoding wild-type ERp29 (WT ERp29) or N-terminal domain of ERp29 with its own ER retention signal (NTD ERp29). Cells were pulse-labeled with $^{35}$S-amino acids for 30 min and chased in complete media. Cells (C) and media (M) were immunoprecipitated with anti-Tg before SDS-PAGE and fluorography. Note that some I-II-III (without ChEL) was rescued in secretion (asterisk) under suppression of ERp29 by dominant negative form of ERp29 (NTD ERp29), but has no change under overexpression of wild-type ERp29. The plasmids encoding wild-type or NTD ERp29 are kindly gift from Dr. Mkrtchian, S. (wild-type ERp29) and Dr. Tsai, B. and Rainey-Barger, E.K. (NTD ERp29).

**References for chapter 4**


1. Medeiros-Neto, G., Kim, P.S., Yoo, S.E., Vono, J., Targovnik, H.M., Camargo, R., Hossain, S.A., and Arvan, P. (1996). Congenital hypothyroid goiter with deficient thyroglobulin. Identification of an endoplasmic reticulum storage disease with induction of molecular chaperones. J. Clin. Invest. *98*, 2838-2844.

2. Hishinuma, A., Furudate, S., Oh-Ishi, M., Nagakubo, N., Namatame, T., and Ieiri, T. (2000). A novel missense mutation (G2320R) in thyroglobulin causes hypothyroidism in rdw rats. Endocrinology *141*, 4050-4055.

3. Park, Y.N., and Arvan, P. (2004). The acetylcholinesterase homology region is essential for normal conformational maturation and secretion of thyroglobulin. J. Biol. Chem. *279*, 17085-17089.

4. Di Jeso, B., Park, Y.N., Ulianich, L., Treglia, A.S., Urbanas, M.L., High, S., and Arvan, P. (2005). Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase. Mol. Cell. Biol. *25*, 9793-9805.

5. Kim, P.S., Kwon, O.Y., and Arvan, P. (1996). An endoplasmic reticulum storage disease causing congenital goiter with hypothyroidism. J. Cell Biol. *133*, 517-527.

6. Chong, J.M., and Speicher, D.W. (2001). Determination of disulfide bond assignments and N-glycosylation sites of the human gastrointestinal carcinoma antigen GA733-2 (CO17-1A, EGP, KS1-4, KSA, and Ep-CAM). J. Biol. Chem. *276*, 5804-5813.

7. Kim, P.S., Lee, J., Jongsamak, P., Menon, S., Li, B., Hossain, S.A., Bae, J.H., Panijpan, B., and Arvan, P. (2008). Defective protein folding and intracellular retention of thyroglobulin-R19K mutant as a cause of human congenital goiter. Mol. Endocrinol. *22*, 477-484.

8. Berg, G., and Ekholm, R. (1975). Electron microscopy of low iodinated thyroglobulin molecules. Biochim. Biophys. Acta *386*, 422-431.

9. Gentile, F., and Salvatore, G. (1993). Preferential sites of proteolytic cleavage of bovine, human and rat thyroglobulin. The use of limited proteolysis to detect solvent-exposed regions of the primary structure. Eur. J. Biochem. *218*, 603-621.

10. Gentile, F., and Palumbo, G. (1993). Trypsin-resistant regions of thyroglobulin: possible relationship with intermonomeric contact site(s). Biochem. Biophys. Res. Commun. *196*, 1120-1126.

11. Gentile, F., Conte, M., and Formisano, S. (2004). Thyroglobulin as an autoantigen: what can we learn about immunopathogenicity from the correlation of antigenic properties with protein structure? Immunology *112*, 13-25.

12. Molinari, M., and Helenius, A. (1999). Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. Nature *402*, 90-93.

13. Han, J.H., Batey, S., Nickson, A.A., Teichmann, S.A., and Clarke, J. (2007). The folding and evolution of multidomain proteins. Nat. Rev. Mol. Cell Biol. *8*, 319-330.

14. Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

15. Shinde, U., and Inouye, M. (2000). Intramolecular chaperones: polypeptide extensions that modulate protein folding. Semin. Cell Dev. Biol. *11*, 35-44.

16. Kim, P.S., and Arvan, P. (1998). Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones. Endocr. Rev. *19*, 173-202.

17. Kim, P.S., and Arvan, P. (1991). Folding and assembly of newly synthesized thyroglobulin occurs in a pre-Golgi compartment. J. Biol. Chem. *266*, 12412-12418.

18. Di Jeso, B., Pereira, R., Consiglio, E., Formisano, S., Satrustegui, J., and Sandoval, I.V. (1998). Demonstration of a Ca2+ requirement for thyroglobulin dimerization and export to the golgi complex. Eur. J. Biochem. *252*, 583-590.

19. Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. Science *253*, 872-879.

20. Arac, D., Boucard, A.A., Ozkan, E., Strop, P., Newell, E., Sudhof, T.C., and Brunger, A.T. (2007). Structures of neuroligin-1 and the neuroligin-1/neurexin-1

beta complex reveal specific protein-protein and protein-Ca2+ interactions. Neuron *56*, 992-1003.

21. Chen, X., Liu, H., Shim, A.H., Focia, P.J., and He, X. (2008). Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions. Nat. Struct. Mol. Biol. *15*, 50-56.

22. Fabrichny, I.P., Leone, P., Sulzenbacher, G., Comoletti, D., Miller, M.T., Taylor, P., Bourne, Y., and Marchot, P. (2007). Structural analysis of the synaptic protein neuroligin and its beta-neurexin complex: determinants for folding and cell adhesion. Neuron *56*, 979-991.

23. Koehnke, J., Jin, X., Budreck, E.C., Posy, S., Scheiffele, P., Honig, B., and Shapiro, L. (2008). Crystal structure of the extracellular cholinesterase-like domain from neuroligin-2. Proc. Natl. Acad. Sci. U. S. A. *105*, 1873-1878.

24. Morel, N., Leroy, J., Ayon, A., Massoulie, J., and Bon, S. (2001). Acetylcholinesterase H and T dimers are associated through the same contact. Mutations at this interface interfere with the C-terminal T peptide, inducing degradation rather than secretion. J. Biol. Chem. *276*, 37379-37389.

25. Dean, C., Scholl, F.G., Choih, J., DeMaria, S., Berger, J., Isacoff, E., and Scheiffele, P. (2003). Neurexin mediates the assembly of presynaptic terminals. Nat. Neurosci. *6*, 708-716.

26. Lee, J., Di Jeso, B., and Arvan, P. (2008). The cholinesterase-like domain of thyroglobulin functions as an intramolecular chaperone. J. Clin. Invest. *118*, 2950-2958.

27. Dunn, J.T., and Dunn, A.D. (1999). The importance of thyroglobulin structure for thyroid hormone biosynthesis. Biochimie *81*, 505-509.

28. Marriq, C., Lejeune, P.J., Venot, N., and Vinet, L. (1991). Hormone formation in the isolated fragment 1-171 of human thyroglobulin involves the couple tyrosine 5 and tyrosine 130. Mol. Cell. Endocrinol. *81*, 155-164.

29. Dunn, A.D., Corsi, C.M., Myers, H.E., and Dunn, J.T. (1998). Tyrosine 130 is an important outer ring donor for thyroxine formation in thyroglobulin. J. Biol. Chem. *273*, 25223-25229.

30. den Hartog, M.T., Sijmons, C.C., Bakker, O., Ris-Stalpers, C., and de Vijlder, J.J. (1995). Importance of the content and localization of tyrosine residues for thyroxine formation within the N-terminal part of human thyroglobulin. Eur. J. Endocrinol. *132*, 611-617.

31. Veenboer, G.J., and de Vijlder, J.J. (1993). Molecular basis of the thyroglobulin synthesis defect in Dutch goats. Endocrinology *132*, 377-381.

32. Taurog, A. (1999). Molecular evolution of thyroid peroxidase. Biochimie *81*, 557-562.

33. Brisson, A., Marchelidon, J., and Lachiver, F. (1974). Comparative studies on the amino acid composition of thyroglobulins from various lower and higher vertebrates: phylogenetic aspect. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. *49*, 51-63.

34. Suzuki, S., Gorbman, A., Rolland, M., Montfort, M.-F., and Lissitzky, S. (1975). Thyroglobulins of cyclostomes and an elasmobranch. Gen. Comp. Endocrinol. *26*, 59-69.

35. Hishinuma, A., Fukata, S., Nishiyama, S., Nishi, Y., Oh-Ishi, M., Murata, Y., Ohyama, Y., Matsuura, N., Kasai, K., Harada, S., et al. (2006). Haplotype Analysis Reveals Founder Effects of Thyroglobulin Gene Mutations C1058R and C1977S in Japan. J. Clin. Endocrinol. Metab. *91*, 3100-3104.

36. Hishinuma, A., Takamatsu, J., Ohyama, Y., Yokozawa, T., Kanno, Y., Kuma, K., Yoshida, S., Matsuura, N., and Ieiri, T. (1999). Two Novel Cysteine Substitutions (C1263R and C1995S) of Thyroglobulin Cause a Defect in Intracellular Transport of Thyroglobulin in Patients with Congenital Goiter and the Variant Type of Adenomatous Goiter. J. Clin. Endocrinol. Metab. *84*, 1438-1444.

37. Kim, P.S., Ding, M., Menon, S., Jung, C.G., Cheng, J.M., Miyamoto, T., Li, B., Furudate, S., and Agui, T. (2000). A missense mutation G2320R in the thyroglobulin gene causes non-goitrous congenital primary hypothyroidism in the WIC-rdw rat. Mol. Endocrinol. *14*, 1944-1953.

38. Magnuson, B., Rainey, E.K., Benjamin, T., Baryshev, M., Mkrtchian, S., and Tsai, B. (2005). ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding. Mol. Cell *20*, 289-300.

39. Baryshev, M., Sargsyan, E., and Mkrtchian, S. (2006). ERp29 is an essential endoplasmic reticulum factor regulating secretion of thyroglobulin. Biochem. Biophys. Res. Commun. *340*, 617-624.

40. Sen, J., Goltz, J.S., Konsolaki, M., Schupbach, T., and Stein, D. (2000). Windbeutel is required for function and correct subcellular localization of the Drosophila patterning protein Pipe. Development *127*, 5541-5550.

41. Forster, M.L., Sivick, K., Park, Y.N., Arvan, P., Lencer, W.I., and Tsai, B. (2006). Protein disulfide isomerase-like proteins play opposing roles during retrotranslocation. J. Cell Biol. *173*, 853-859.

42. Muresan, Z., and Arvan, P. (1998). Enhanced binding to the molecular chaperone BiP slows thyroglobulin export from the endoplasmic reticulum. Mol. Endocrinol. *12*, 458-467.