Comparison of a Novel Cell-based Reporter Assay and a Competitive Binding ELISA for the Detection of Thyrotropin-Receptor (TSHR) Autoantibodies (TRAb) in Graves' Disease Patients

A thesis presented to

the faculty of

the Russ College of Engineering and Technology of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Science

Misako Hata

March 2010

This thesis titled

Comparison of a Novel Cell-based Reporter Assay and a Competitive Binding ELISA for the Detection of Thyrotropin-Receptor (TSHR) Autoantibodies (TRAb) in Graves' Disease Patients

by

MISAKO HATA

has been approved for

the Department of Chemical and Biomolecular Engineering

and the Russ College of Engineering and Technology by

_____

Douglas J. Goetz

Professor of Chemical and Biomolecular Engineering

_____

Dennis Irwin

Dean, Russ College of Engineering and Technology

ABSTRACT

MISAKO HATA, M.S., March 2010, Biomedical Engineering

Comparison of a Novel Cell-based Reporter Assay and a Competitive Binding ELISA for the Detection of Thyrotropin-Receptor (TSHR) Autoantibodies (TRAb) in Graves' Disease Patients (66 pp.)

Director of Thesis: Douglas J. Goetz

The pathogenesis factor of Graves' Disease (GD) has been widely accepted as autoantibodies against thyrotropin receptor (TSHR) over-stimulating follicular cells to produce excess thyroid hormones. For the last few decades, the competitive binding assay for TSHR antibody (TRAb) has been the most commonly used assay for the differential diagnosis of GD. The competitive binding assay measures the heterogeneous mixture of TRAbs in the patients' sera that prevent labeled thyroid stimulating hormone (TSH) or monoclonal stimulating TRAb from binding to the fixed human recombinant or porcine TSHRs.

In this study, a new cell based reporter assay with chimeric human TSHR (Mc4) was evaluated against the third generation competitive binding Enzyme-Linked Immuno Sorbent Assay (ELISA). Mc4 utilizes its mechanism to detect only the simulating TRAbs in the patent sera that directly correlate with GD hyperthyroidism. Furthermore, a Mc4 predicate, a cell based reporter assay with human wild type (hWT) TSHR (CHO-Luc), was evaluated. This study conducted comparisons of these three assays on the same group of GD patients (n = 200) and healthy blood donors (HBD) (n = 40).

Overall sensitivities given the sample provider's diagnosis as the reference standard were similar with all three assays (84.0 – 73.4%). Mc4 had the second highest sensitivity (79.5%) without misdiagnosing normal controls (specificity = 100%). Sensitivity comparison was ambiguous since some of the 200 GD specimens had high TSH and might have been receiving antithyroid drug treatments which interfered with the assay results. When GD positive groups were divided with TSH levels, agreements of all the assay results were the highest within the very low TSH (TSH < 0.01 μIU/mL) group. Interpretation of TRAb ELISA gave different performance measures (sensitivity and specificity) within the same sample set.

Approved: _____

Douglas J. Goetz

Professor of Chemical and Biomolecular Engineering

ACKNOWLEDGMENTS

The progression of my research work, career, and school could not be done with the support from my advisor, committee members, corporate members from Diagnostic Hybrids, and friends and family.

Firstly, I would like to express my gratitude to my thesis advisor Dr. Douglas J Goetz. He has arranged a wonderful educational and career opportunity for me. Working in the real-world laboratory as an intern and joining the biomedical engineering program was an exceptional opportunity. He has also supported me in the completion of this thesis work. Thank you Dr. Goetz.

I would also like to thank Dave Scholl Ph.D, the president of Diagnostic Hybrids, for welcoming me and allowing me to use company resources and precious clinical samples for this research. Also, I would like to thank my supervisors Ruth Kelly Ph.D. and Jeff Houtz and my other colleagues at Diagnostic Hybrids. This academic research has shown the product was genuine. I wish the best for Thyretain™.

I would like to thank all the committee members Drs. Frank Schwartz, Valerie Young and Monica Burdick for all the support, suggestions and answering my questions. And thank you to Tom Riggs, the biomedical engineering program assistant, for all the information you have provided me. I would also like to extend a special thank you to Dr. Leonard Kohn, a pioneer of this assay and distinguished research professor, for patiently answering my questions.

Lastly I would like to thank my husband, Greg Emery. It has been a painful few years going through school, my career change, and raising children. Thank you for

sharing house work, being a great support, and editing incomprehensible papers. And thank you to my daughters Grace and Violet for being good girls.

None of this would have been possible without you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Autoimmune Disease

Autoimmune diseases are caused by the body's immune system attacking a part of itself instead of foreign substances such as viruses or bacteria [1]. Autoimmune diseases are classified into two general groups: multi-system autoimmune diseases and organ-specific autoimmune diseases [2]. Multi-system autoimmune diseases are, for example, systemic lupus erythematosus (SLE), neonatal lupus and Sjögren's syndrome. These autoimmune diseases share the same autoantibody specificity against ribonucleoprotein (RNP) particles but have different symptoms in each patient [2]. The relationship between causality and symptoms of the multi-system autoimmune disease is not well known and symptoms tend to affect the whole body system. On the other hand, organ specific autoimmune diseases have a specific target by the immune system and it is relatively easier to find biomarkers (autoantibodies) associated with underlying clinical symptoms [2]. In this thesis, the focus is on the thyroid gland, specifically an autoimmune disorder targeted to thyrotropin receptors (TSHR) on the surface of the thyroid follicular cells.

Autoimmune Thyroid Disease (AITD)

The John's Hopkins Autoimmune Research center website (based on [3]) lists the thyroid specific autoimmune diseases. Graves' disease (first) and Hashimoto's thyroiditis (third) are in the top three prevalent autoimmune diseases in the United States. Autoimmune Thyroid Disease (AITD) is one of the well-studied organ-specific autoimmune diseases involved in lymphocytic infiltration of the thyroid gland, producing different kinds of autoantibodies of the thyroid components [4]. In general, the presence of certain combinations of thyroid autoantibodies are loosely related to clinical features

of the AITD, except Graves' disease (GD) which is directly caused by the presence of one functional type (stimulating) of thyrotropin receptor antibodies (TRAb) [5-7].

The thyroid is an endocrine organ which absorbs iodine from food, attaches the iodine to two tyrosine molecules and converts it principally into thyroxine which is the circulating thyroid hormone which helps regulate the body's metabolism [8]. There can be two metabolic consequences due to AITD: one is hyperthyroidism (over-production of thyroxine) resulting in a hyper-metabolic state and the other is hypothyroidism, resulting in inadequate production of thyroxine and a hypo-metabolic state. Hashimoto's Thyroiditis, results from a lymphocytic attack of the thyroid, which in general leads chronic inflammation of the thyroid, can lead to the destruction of the thyroid, is the most common cause of hypothyroidism [9]. GD, the focus of this thesis, results of autoantibodies directed against the TSH receptor and autonomously stimulates thyroid hormone synthesis and release resulting in hyperthyroidism.

Graves' Disease

What is Graves' Disease?

GD is an autoimmune thyroid disorder leading to an overactive thyroid state called hyperthyroidism [10]. The other causes of hyperthyroidism include; toxic thyroid adenoma, toxic multinodular goiter, subacute thyroiditis, and iodine containing dyes or medications [10, 11]. The over production of thyroxine results in increased metabolic rate and sensitivity to beta adrenergic receptors. Symptoms of hyperthyroidism are weight loss, proximal muscle weakness, rapid heart rate, fatigue, and tremors. Typical cosmetic signs are goiter (enlargement of the thyroid), exophthalmos (prominence of the eyes), and in rare cases pretibial myxedema (abnormal inflammation of the skin on the shins) [7, 10, 11]. A subgroup of GD patients also develops Graves' ophthalmopathy (GO) which is

due to abnormal inflammation of the retro-orbital tissues behind the eye causing the anterior displacement of the eyes out of the orbit. Most patients with GO are hyperthyroid however some patients exhibit normal thyroid hormone levels which is called Euthyroid GO [8]. Overall the hallmark of GD is the presence of autoantibodies against TSHR [5, 6].

Mechanism of Graves' Disease

In healthy individuals, the pituitary gland, the controller of the thyroid gland, produces the thyroid stimulating hormone (TSH) which binds to the TSHR to stimulate the thyroid gland (follicular cells) to produce the thyroid hormones. These thyroid hormones are called triiodothyronine (T3) and thyroxine (T4) [8]. Figure 1 shows TSH and T3/T4 and the location of the organs. T3 and T4 are released in the blood stream (endocrine system) to circulate and control the metabolism of all the cells in the body [8]. A healthy individual maintains a certain ratio of T3 and T4 and T4 is the principle circulating thyroid hormone. T4 is converted into T3 (one less iodine) by monodeiodination in target cells and is much more potent than T4 [8]. Only the "free" circulating forms T3 and T4 are biologically active. "Free" means it is unbound to a carrier protein called thyroid binding globulin (TBG), ready to be used, and reflects a better measure of hormone levels for diagnosis [8].

A positive/negative feedback mechanism (Figure 2) controls production of T3 and T4 detected by the pituitary gland and hypothalamus. The hypothalamus controls the pituitary gland's TSH production by releasing the thyroid releasing hormone (TRH). When T3 and T4 levels are low, this is sensed by the hypothalamus and pituitary which then release TRH and TSH subsequent stimulation of the thyroid to release T3 and T4 to

restore normal hormone levels. When T3 and T4 levels are high, both TRH and TSH production is decreased, the thyroid is no longer stimulated and release of T3 and T4 is reduced to maintain the euthyroid state [8].

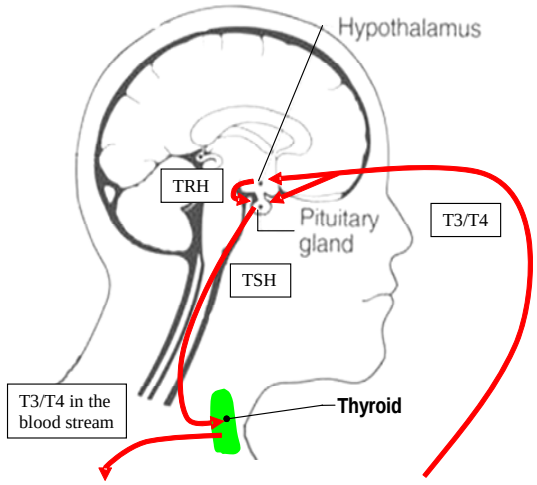

Figure 1. Feedback mechanism of the TRH, TSH, and T3/T4 of a healthy individual and location of the organs. [Adapted from the copy right free clip art (0512-0712-1719-0225) on www.clipartguide.com.]

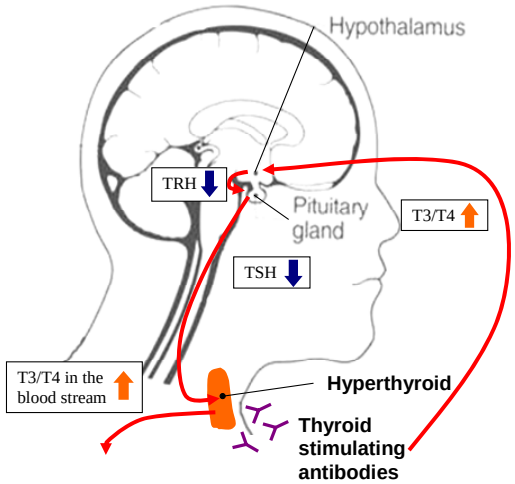

Figure 2. GD hyperthyroid state. Stimulating autoantibodies disturb the feedback mechanism and constantly stimulate the thyroid. [Source same as Figure 1.]

In the case of GD, the patient'simmune system is triggered to produce antibodies towards their own thyroid. Thyroid stimulating autoantibody (TSAb) is one of the functional autoantibodies directed against TSHR (schematically shown in Figure 2) [5-8, 10, 11]. TSAbs, acting like TSH, bind to TSHR stimulating the follicular cells. As a result, TSAbs stimulate the thyroid to produce too much T3 and T4 [8] which is autonomous to the hypothalamic/pituitary regulation of the thyroid previously described. Figure 2 shows a hyperthyroid state with an increased level of T3/T4. The hypothalamus and pituitary gland detect the high concentration of T3 and T4 in the circulating blood and reduce TSH production in an attempt to suppress thyroid activity. However, the presence of TSAb, which bind and stimulate the thyroid with longer duration compared to native TSH, constantly over-stimulates the GD patient's thyroid leading to the disease's symptoms. Therefore the typical diagnostic pattern of the GD patient is a low TSH levels in the serum, high levels of both biologically-active free T3 and free T4[10, 12], and the presence of the TRAb's which are directly causing the GD.

Diagnosis of GD

ATA Guidelines and TSHR Autoantibody Measurement

The American Thyroid Association (ATA) Guidelines are considered the general standard practice for the screening of thyroid dysfunction. According to the ATA [12], a low (less than 0.1 μU/mL) TSH in patients' serum usually indicates cases of hyperthyroidism. For accurate hyperthyroidism detection, free T4 and optional free T3 assays are recommended in addition to the TSH measurement [12]. Interestingly, ATA's recommendations do not mention TRAb measurements although TRAb assays are

commonly used to differentiate Graves' disease from other forms of hyperthyroidism. TRAb measurement is sometimes recommended for special cases such as pregnancy [13, 14] where the antibody can be transferred via the placenta to the baby or to diagnose Euthryoid GO [5, 14]. There seem to be inconsistent opinions about use of TRAb measurement for differential diagnosis of GD. Part of the reason is the assay technologies and the other reason is TRAb heterogeneity discussed in later sections.

Concept of Molecular Diagnostics

Historically, a long-acting thyroid stimulator (LATS) was discovered in GD patients' sera in 1956 and was identified as a cause of the thyroid stimulating activities [15]. Later, the LATS were found to be a form of immunoglobulin G (IgG) capable of binding to the TSHR and mimicking endogenous TSH [16]. Currently it is widely accepted that GD is thereby an autoimmune disease and caused by the action of TSAbs in the patient sera leading to a hyperthyroid state[5-8, 10, 11, 17, 18]. Even so, detection of TSAbs is debated by clinicians [5, 10, 12, 19]. One of the reasons is that measurement of the other hormone levels (TSH, freeT4 and T3) in addition to the clinical feature, is said to be adequate to diagnose GD [10, 14]. Also, there has been a debate about the accuracy of the TSHR antibody detection methods [20] as well as molecular biological interaction between heterogeneous TRAbs in the patients' sera and TSHRs, which underlies the assay principles [21]. Therefore, there are two hindrance factors for optimizing TSAb measurements for differential diagnosis of GD: (i) Complexity of the molecular mechanism of TSHR and its autoantibodies, (ii) Accuracy of the assay due to technologies and underlying assay principles.

TSHR Autoantibodies

GD patients have been shown to different clones of TSHR autoantibodies directed at different locations of the TSHR[22]. The actions of the different TRAbs are thought to vary due to the different binding sites. As a consequence there are different types of TRAbs: stimulating, blocking and neutral [21]. Stimulating, also called TSAbs, are one of the TRAbs which stimulates the thyrocytes for T3/T4 production mainly through the cyclic adenosine monophosphate (cAMP) pathway [22] and cause GD. Diagnosis of GD is simple if patients only develop TSAbs which activate only cAMP cascades. However, a minor group of AITD patients also develop blocking, or thyroid blocking antibodies (TBAbs) which interfere with TSH and sometimes TSAbs. In these cases, the TSHRs on thyrocytes are “blocked” from T3/T4 production, leading to hypothyroidism [23]. TBAb are often found in primary myxedema patients (hypothyroidism), a subgroup of Hashimoto’s thyroiditis, and a small subgroup of GD patients who develop both TBAb and TSAb [21, 24, 25]. Neutral autoantibodies, which neither stimulate nor block thyroid stimulating activities, have also been found by a few studies [22, 26]. However, neutral autoantibodies have not yet been clearly identified in studies with a large number of patients being screened [27].

In addition to the different types of TRAbs, there may be variations in signaling cascades after TRAbs bind to TSHRs. A recent paper by Syed [28] has shown a unique signaling cascade (“c-Raf-ERK-p90RSK”) in addition to the traditional “G protein - cAMP protein kinase A/ERK” using FRTL-5 cells with monoclonal TSAbs and TBAbs.

However, these signaling cascades interact and are not mutually exclusive. TSAbs mainly stimulate the G protein – cAMP protein kinase A/ERK pathway.

TSHR

TSHRs, which are the principle autoantigens of GD are expressed on the surface of thyroid epithelial cells (also known as follicular cells, principal cells or thyrocytes) and respond to TSH [8]. TSHR is a unique receptor compared to the family of "G-protein coupled receptors with seven transmembrance spanning domains of the cAMP generators" [22]. There is evidence that TSHRs are expressed in a variety of other human or animal tissues such as adipocytes, fibroblasts and bone cells even on embryonic stem cells [29, 30].

The TSHR is a gene composed of a total of 764 amino acids translated from a single mRNA (Figure 3) [27]. The primary and secondary structures of TSHR have been known since 1989; however, the tertiary and quaternary structures which could delineate the confomation of an autoantibody and TSHR subunits have long been debated among researchers [31]. There are four domains in the TSHR: i) N-terminal domain (1-22), ii) leucine-rich domain (LCD) (53-277), iii) hinge region (277-418), and iv) transmembrane domain (419 – 764). Each domain structure is well studied; however, the total TSHR structure on the cell surface as a multimer is not known and there are controversies with the crystal structure of the related receptor such as follicle stimulating hormone receptor, partial crystal structure of TSHR and functional studies by mutagenesis [31].

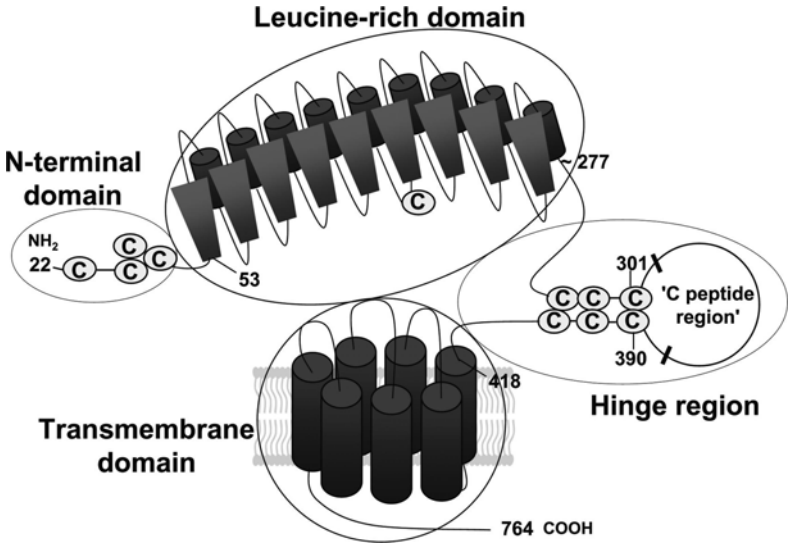

Figure 3. A schematic model of a TSHR showing four components: N-terminal domain, Leucine-rich domain, hinge region and transmembrane domain. The hinge region includes a ‘c peptide region’ which is cleaved by an unknown enzyme and leads to A sub unit (LRD + N-terminal domain) separation from the transmembrane domain. (Figure adapted with kind permission from the author and Mary Ann Liebert, Inc., publishers [31]).

Mutagenesis (point-, deletion or chimeric) studies of the TSHR receptor were popular in the 1990’s in an attempt to determine where the TRAb’s bind to the receptor to trigger the cell activities. Chimeric receptors in which the target region of the TSHR was substituted with a homologous receptor of a different species were preferred over the mutagenesis study in order to keep highly conformational TSHRs close to the original shape [23]. Based on these studies, TSH, TSAb, and TBAb binding regions on the receptor were determined. This is illustrated in Figure 4. This model contradicts a binding study using the labeled and affinity purified TRAbs [32] and crystallized TSHR-TRAb or TSHR-TSH pictures [27] in which TSH and TRAbs were found to bind to the N-terminal end of TSHR at the LRD. Taking these differences in TSHR-TRAb epitope studies in

consideration, commercial assays have been developed. In the next section, details of two major commercially used assay principles are discussed.

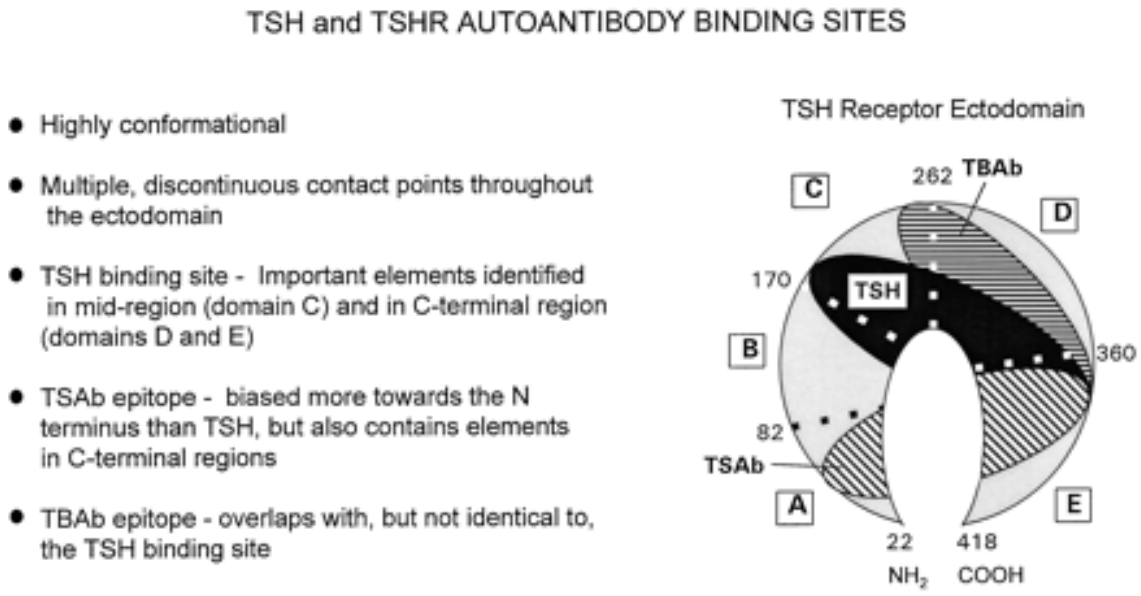

Figure 4. The right figure shows a schematic view of general binding sites of TSAb, TBAb and TSH based on the TSHR mutation studies. Important points about the TRAb and TSHR interactions are outlined. (Figure adopted with kind permission from the endocrine society publishers. [23])

Concept of Molecular Recognition Assays

TSHR Autoantibody Detection Methods

A variety of *in vitro* tests for detecting minute amounts of TSHR autoantibodies in patients' sera are available. Immunoassays utilize antibody-antigen binding reaction to detect the target autoantibodies biochemically. Conventional immunoassay methods have two types: immunochemical methods and immunometric methods [33]. Immunochemical methods, for example Immunoprecipitation Assays (IPA), qualitatively detect

autoantibodies resulting in a binary result, either positive (present) or negative (not present). Immunometric methods such as radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA) and chemiluminescent immunoassay (CLIA) indirectly quantify the autoantibody concentrations [33]. Then, a cut-off value determines if the TRAb presence is positive or negative. The purpose of an immunoassay is to detect the existence and quantify the target substance in the biological specimen. There is another category of tests called bioassays which are utilized for TRAb detection. Unlike the immunoassay, bioassay utilizes live cells to measure the reaction, in this case stimulus) of the TRAb existing in the GD patients' sera.

Table 1 summarizes a list of major TRAb detection methods for differential diagnosis of GD used in clinical labs and research labs [34]. Bioassays, competitive binding assays (also called indirect, TSH inhibition assays in the Table 1) and direct binding assays are the three major categories for detecting TRAbs. To date, controversy still remains regarding which assay most accurately reflects the biological activities and clinical relevance of GD [34]. Although listed in Table 1, direct binding assays are considered unreliable due to the very low concentration of the autoantibody existing in patients' sera [34]. In addition to the list in Table 1, the latest autoantibody detection method is the flow cytometry analysis, which investigates individual cell morphology and expression with labeled antibodies. However, this method is out of scope of this thesis and routine use for clinical set up since special equipment and skills are necessary for diagnosis without standardized cut-off.

Table 1

List of Major TRAb Detection Methods and Their Principles. Table adopted with kind permission of the author and Springer Science and Business Media. [34] of *Table 2 (p. 51)*

Assay & Antibody	Source of TSHR	Detection system(s)
Bioassays		
LATS; LATS-P	In vivo mouse assay	Release of ^{131}I from thyroid gland
TSAb (TSI)	Human thyroid cells Rat FRTL5 cells CHO cells (human TSHR)	cAMP generation Iodide transport Chemiluminescence
TBAb (TSBAb)	Human thyroid cells Rat FRTL5 cells CHO cells (human TSHR)	Inhibition of cAMP release (or iodide uptake) by standard TSH concentration
Indirect, TSH inhibition assays		
TBI	Thyroid membranes CHO cells (human TSHR) Soluble TSHR (porcine; human)	Inhibition of ^{125}I-bovine TSH binding
Direct binding assays		
DTAb	^{35}S-TSHR cell free translates	Immunoprecipitation
	^{125}I-labeled hTSHR	Immunoprecipitation
	TSHR-ectodomain-His tail	ELISA
	TSHR-289-6 His tail	Protein A capture of antibodies; detection by anti-His antibody
	TSHR labeled with Fab to C-terminus	Protein A precipitation of antibodies

Therefore there are two major types of widely used assays in the clinical setting detecting TRAbs in the patients' sera for differential diagnosis of GD [5]. One is a competitive binding assay and the other is a bioassay. The comparison of these two types of assays with different principles is the focus of this thesis. The principle of each assay with the underlying latest technology is explained in the following sections.

Principle of Competitive Binding Assay

The competitive binding assay utilizes labeled TSH or monoclonal TSAb competing against TRAbs existing in patient sera for the fixed amount of TSHRs. As shown in Figure 5, antibodies existing in patient sera and labeled binding element provided in the assay compete for the detergent solubilized receptors attached to the bottom of the tube or well. By measuring the labeled TSH or monoclonal TSAb, M-22 for newer assays, occupying the receptor after washing steps, the concentration of the autoantibody existing in the patient sera is estimated indirectly by a standard curve constructed with known concentrations of TRAbs. Several versions of this type of assay are available. These versions include better labeled TSHR binding elements (from TSH to monoclonal TSAb), origin of the receptor (from porcine to recombinant human), labeling method (RIA, ELISA, CLIA and electrochemiluminescence immunoassay (ECLIA)), and detection method (from radio receptor immunoprecipitation assay to the latest beads based liquid phase immunoassay). These modifications have increased the sensitivity and specificity of this type of assay.

However, the competitive binding assay only measures the aggregate of TRAbs existing in the patient sera competing with the labeled substance for the fixed TSHR position [5]. If a patient happened to have a mixture of TSAbs and TBAbs, the assay cannot differentiate the types of autoantibodies.

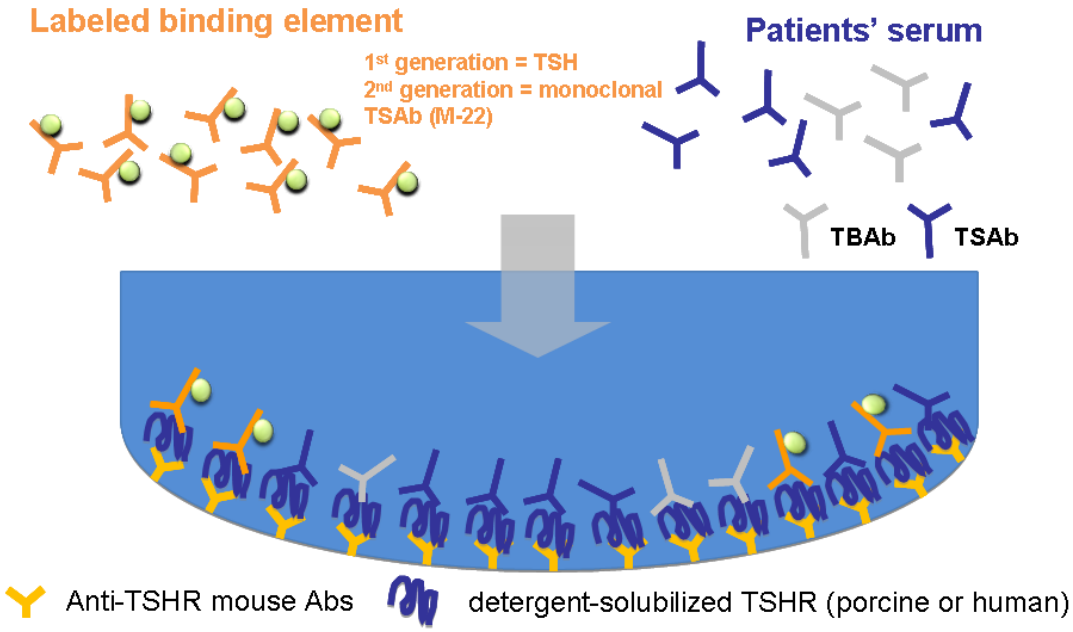

Figure 5. Simplified view of the competitive binding assay. The detergent-solubilized TSHRs are attached on the bottom of a tube or well. In the tube or well, patient's sera and labeled antibodies are mixed to compete and measure the labeled antibody occupancy to estimate the autoantibody titer in the patient's sera.

Antibody Estimation

Competitive binding assay measures the labeled TSH or monoclonal TSAb (M-22), occupying the fixed amount of TSHRs. The reading format may be absorbance, optical density, luminescence, fluorescence, radioactivity depending on the assay. There are two calculation methods to estimate the amount of TSHR Abs in patients' sera: percent inhibition and calibration method. It is up to the users to determine which methods to use, although some manufacturers suggest one method over the other. The percent inhibition method measures how much the patient serum inhibits labeled TSH or monoclonal TSAb with respect to the negative control value. The percent inhibition method is calculated by plugging the negative control (provided by manufacturer)

reading and a sample value reading in Equation [1] to yield percent inhibition by the patient sera [35].

$$\mathbf{100} \times \left(\mathbf{1} - \frac{\text{test sample absorbance 450 nm}}{\text{negative control absorbance 450 nm}} \right) \quad [1]$$

The calibration method requires calibrators, a set of known concentrations of the TRAb from human sera (NIBSC 90/672) standardized by National Institute for Biological Standards and Control (NIBSC). NIBSC standardized samples are also considered a World Health Organization (WHO) standard. A set of calibrators can be used to calculate the standard curve to interpolate the concentration of the antibodies in patent sera. The standard curve [36] can be linear or a non- linear curve. If a non- linear curve is used, there are several non- linear regression curve fitting models such as 4PL, 5PL, or logit [37]. Again, each assay manufacturer may suggest which (linear or non-linear) regression curves to use.

Evolution of Competitive Binding Assays

Competitive binding assays have three major variable components: i) the receptor, ii) the competitive binding element, and iii) the label on the competitive binding element. First, receptors are derived from two species: human (recombinant) or porcine for most of the commercial assays. Secondly, competitive binding elements are the labeled material mixed with patients' sera which compete for the TSH or monoclonal TSAb binding epitopes on the receptor. The earlier generation assay had bovine or porcine TSH; however, the creation of the monoclonal TSAb (M-22) from a lymphocyte of a severe GD patient by Sanders et al [38] replaced the TSH. Use of M-22 has increased the sensitivity of the assay due to common shared epitopes with patient's TRAbs on TSHR

[39]. Third, labels on competitive binding elements are important for usability of the assay. The earlier assays had radioactive material (I^{125}) attached to the TSH which required extra care and steps to conduct the assay. The latest assay format is an ECLIA in a machine platform which can detect TRAbs in 30 min [40]. Table 2 lists the commercially available competitive binding assays. These are broken down into 3 generations of assays. The first generation are the liquid phase competitive immunoprecipitation assays. The 2^{nd} and 3^{rd} generations are solid support competitive inhibition assays using TSH (2^{nd} generation) or TSAb (stimulating TRAb) (3^{rd} generation) as the competitive binding element. Each manufacturer has a slightly different format and there has been debate in the literature which assay accurately detects the patients' TRAb existence with respect to endocrinologists' diagnosis based on the other hormone levels (e.g. free T3/T4).

Table 2

A List of the 3 generations of Competitive Binding Assays and Each Major Manufacturer's Assay Format.

	#	Product Name	i) Receptor	ii) Competitive Binding element	iii) Detection method	Manufacturer	Calibrator	cut-off value	Reference
1st generation	1	TRAK-Assay	porcine thyroid membrane extracts	Unknown	RRA immunoprecipitation	Brahms Diagnostica	Unknown	>15% (set by a reference lab), > 9.9% for (Morgenthaler 2002)	Muroi 2007, Morgenthaler 2002, Zimmermann-Belsing 2002, Costagliola 1999
	2	TSH-REZAK	porcine	porcine TSH	RRA immunoprecipitation	MEDIPAN Diagnostics	human serum	>= 15 IU/L, 10-15 IU/L (gray zone), < 10 IU/L (negative)	Schott 2000
	3	KRONUS TRAb	porcine	Unknown	RRA immunoprecipitation	KRONUS	not specified	not specified	Preissner 2003
2nd generation	4	KRONUS TRAb CT RIA	porcine	bTSH	RIA solid phase	KRONUS	90/672	not specified	Preissner 2003
	5	KRONUS-ELISA	porcine	bTSH	ELISA solid phase	KRONUS	90/672	> 1.5 IU/L (positive), 1.5 - 1.0 IU/L (gray zone), < 1.0 IU/L (negative)	Preissner 2003
	6	TRAK-ELISA	porcine	bTSH	ELISA solid phase	EUROIMMUN	90/672	> 2.0 IU/L (positive), 2.0 - 1.8 IU/L (gray zone), < 1.8 IU/L (negative)	Hermsen 2008
	7	DYNOtest TRAK human	human	bTSH	RIA solid phase	Brahms Diagnostica	90/672	> 1.5 IU/L (positive), 1.5 - 1.0 IU/L (gray zone), < 1.0 IU/L (negative)	Muroi 2007, Villalta 2004, Preissner 2003, Morgenthaler 2002, Schott 2000,
	8	LUMItest TRAK human	human	bTSH	CLIA solid phase	Brahms Diagnostica	5 points calibrator 90/672	> 1.5 IU/L (positive), 1.5 - 1.0 IU/L (gray zone), < 1.0 IU/L (negative)	Hermsen 2008, Liu 2008, Kamijo 2005, Villalta 2004, Zimmermann-Belsing 2002
3rd generation	9	Medizym®TRAb clone (TRAb ELISA)	porcine	human monoclonal TRAb M-22	ELISA solid phase - biotinylated M22 with streptavidin-polyperoxidase as the secondary tag	MEDIPAN Diagnostics	5 points calibrator 90/672	> 0.4 IU/L (positive), 0.3- 0.4 IU/L (gray zone), < 0.3 IU/L (negative)	
	10	TSH Receptor Autoantibody 3rd Generation ELISA Kit	porcine	human monoclonal TRAb M-22	ELISA solid phase - biotinylated M22 with streptavidin-polyperoxidase as the secondary tag	RSR Limited	5 points calibrator 90/672	> 0.4 IU/L (positive), < 0.3 IU/L (negative)	Kamijo 2005
	11	TRAb-Fast-ELISA	porcine	human monoclonal TRAb M-22	ELISA solid phase direct-labeled M-22 with peroxidase	EUROIMMUN	5 points calibrator 90/672	> 2.0 IU/L (positive), 1.8 - 2.0 IU/L (gray zone), < 1.8 IU/L (negative)	Liu 2008, Zophel 2009
	12	KRONUS TRAb ELISA	porcine	human monoclonal TRAb M-22	ELISA solid phase direct-labeled M-22 with peroxidase	KRONUS	4 points calibrator 90/672	> 1.0 IU/L (positive), < 1.0 IU/L (negative)	
	13	Roche ECLusys	porcine	human monoclonal TRAb M-22	ECLIA liquid phase ruthenium labelled M-22 Automated on the ECLusys platform	Roche	2 points calibrator 90/672	> 1.75 IU/L (positive)	Hermsen 2009

chemiluminescent immunoassay (CLIA), Luminescence immunoassay (LIA), Electrochemiluminescence immunoassay (ECLIA)

First vs. Second Generation Assay Comparison

Many comparison studies with large numbers of clinical samples have been published as new generation assays emerge on the market. Several studies compared commercially available 1st and 2nd generation assays. The 1st generation assay is a radio receptor immunoprecipitation method (or radio receptor assay, RRA) in which TRAbs in patient sera and ^{125}I labeled TSH compete for detergent solubilized porcine TSHR receptors in suspension (liquid phase). The bound receptors with labeled TSH are centrifuged to count the label in the pallet. The 2nd generation assay has porcine or human recombinant TSHRs fixed onto the bottom of the tube or plate (solid-phase) so that unbound materials can be washed to reduce the non-specific binding.

Muroi [41], Morgenthaler [42], Schott [43], and Costagliola [44] compared 1st and 2nd generation of TRAb assays with ^{125}I read out. Costagliola [44] and Zimmermann-Belsing [45] compared 1st generation RRA against 2nd generation competitive binding assay with chemiluminecent read out. The main differences between these two assays are the origin of the receptors (porcine used in 1st generation and recombinant human used in 2nd generation) and detection methods (immunoprecipitation RRA used in 1st generation and solid-phase competitive binding immunoassay used in 2nd generation). All the studies agreed that 2nd generation TRAb assays have higher sensitivity and specificity.

However in a GD remission study by Zimmermann-Belsing [45] and Muroi [41] both 1st and 2nd generation assays did not show usability for predicting GD remission after antithyroid drug treatment (ATD). In addition, Bolton et al. [46] compared the 1st and 2nd generation assays of the different formats (RRA immunoprecipitation and

ELISA) with 56 GD sera sample sets. This study did not use the commercially available assay. The result did not show any difference and it was concluded that the newer ELISA format is as good as the RRA method. However, in commercial assay comparisons, 2nd generation human receptor assays yielded better diagnostic sensitivity compared to the porcine receptor based on the immunoprecipitation RRA [42-44].

Comparisons amongst Second Generation Assays

With the majority of the studies[42-44] suggesting that the 2nd generation TRAb assay have superior diagnostic sensitivity, the comparison studies between 2nd generation assays with different detecting systems have been compared. Preissner et al at Mayo Clinic [47] conducted the comparison of one 1st generation (KRONUS TRAb, row 3 in Table 2) and three 2nd generation commercially available assays (KRONUS coated tube (CT) RIA in row 4, KRONUS ELISA in row 5 and BRAHMS TRAK RIA in row 7) with 32 untreated GD patients. In Preissner's [47] comparison, BRAHMS TRAK (in row 7) and KRONUS ELISA (in row 5) resulted in about the same sensitivity (93.8% / 93.3%) out of non-treated GD pool determined by the percent inhibition method. BRAHMS TRAK (in row 7) sensitivity was even higher against KRONUS ELISA (in row 5) (96.9% / 93.3%) if the gray zone was scored as positive. It is noteworthy that if the determination index was percent inhibition KRONUS ELISA (in row 5) has high sensitivity. However if the calibrator included in the kit was used to estimate the antibody titer and the gray zone was treated as negative, then sensitivity became far worse (from 93.3% to 66.7%). KRONUS TRAb (in row 3) and CT RIA (in row 4) had lower sensitivities and specificities compared to BRAHMS TRAK (in row 7) and KRONUS

ELISA(in row 5). Similarly, KRONUS TRAb (in row 3), KRONUS CT RIA's (in row 4) diagnostic power worsened when the calibration method was used compared to when the percent inhibition method was used. This is probably due to the calibrator range included in the kit.

Another study by Villalta [48] compared 2nd generation BRAHMS' DYNO test TRAK human (in row 7 in Table 2) and LUMI test TRAK human (in row 8), both of which have human recombinant TSHR receptors. The difference between the two tests was the label detection mechanism. The DYNO test is a RIA assay and the LUMI test is a CLIA. The study concluded that both tests had the same sensitivity, specificity and coefficient of variance (CV). Hermsen's study [49] compared the 2nd generation Lumitest TRAK human by BRAHMS (in row 8) and the 2nd generation TRAK-ELISA by EUROIMMUN (in row 6). His study compared the receptor difference (human vs. porcine), and reading formats (CLIA and ELISA). Based on the manufacturers' cut-offs, CLIA by BRAHMS with human receptors (in row 8) yielded slightly better sensitivity, 95.2% against 92.9% by the porcine based EUROIMMUN ELISA (in row 6).

Comparison Studies with Latest Third Generation Assays

A major technical breakthrough came when Sanders et al. developed a human monoclonal TSAb with high stimulation activity and TSH inhibition capability [38]. All the 3rd generation commercial assays utilize a monoclonal TSAb called M-22 instead of TSH so that polyclonal TSAbs in the patient serum compete with M-22 for similar binding sites.

A study [17] with the 3rd generation ELISA (row 10 in Table 2) with M-22 based on Smith's assay [39] and 2nd generation assay DYNOtest TRAK human by BRAHMS (row 7) has shown very high sensitivity and specificity for both. The 3rd generation assay detected 243, and the 2nd generation assay detected 242 out of 244 untreated GD patients. A very high sensitivity (99.6% by 3rd generation assay in row 10 and 99.2% by 2nd generation assay in row 7) with a non treated GD group was obtained at the optimal cut-off determined by the receiver operating curve (ROC) [50]. The reading format was the percent inhibition method (optimal cut-off was around 15% by the percent inhibition method). The study found that both methods detect TRAb; however, the 3rd generation assay (in row 10) had better specificity (96.7% by 3rd generation against 93.9% by 2nd generation (in row 7)) if the painless thyroiditis patient group (119) was added as "TSAb negative" patients. On the other hand, Liu [51] compared the 2nd generation Lumitest TRAK human (in row 8) by BRAHM's and the 3rd generation EUROIMMUN TRAb-fast-ELISA (in row 11). The 2nd generation CLIA, TRAb human (in row 8) had higher sensitivity (95.2%) than the 3rd generation TRAb-Fast-ELISA (in row 11; 92.9%). And also CLIA, TRAb human (row 8) had higher specificity (100%) than TRAb-Fast-ELISA (in row 11; 97.3%). However CLIA TRAb human (in row 8) had a gray zone (1.0 – 1.5 IU/L) considered as "negative" for calculating sensitivity and specificity. These comparison studies of 3rd and 2nd generation assays have yielded contradictory results. Zophel [52] pointed out that there are differences in the ELISA format between Smith's ELISA (in row 10)[39] used in Kamijo's study [17] and EUROIMMUN TRAb-fast-ELISA (in row 11). The original Smith assay requires longer incubation time and the

secondary labeling (biotinylated M22 with streptavidin-polyperoxidase label) whereas EUROIMMUN TRAb-fast-ELISA (in row 11) has a direct peroxidase-label on M-22. As a result, EUROIMMUN TRAb-fast-ELISA (in row 11) had faster assay completion time; however, the sensitivity of the assay may be compromised due to the direct labeling. A confirmation study is necessary to compare the original (in row 10) [39] and faster version (TRAb-fast-ELISA in row 11 or KRONUS ELISA in row 12) of the assay.

Although the sensitivity and specificity of the competitive assays appear to increase as newer versions emerge [17, 39, 41-43, 46, 51, 53], some studies set their own optimal cut-off. Sensitivity and specificity varies depending upon whether the user established cut-off or a manufacturer recommended cut-off is used. Also a problem with the competitive binding assay is that there is more than one calculation method (a percent inhibition and calibration methods using a standard curve with different curve fitting models) to determine TRAb concentration. Again, these assays only detect the total amount of antibodies present in the patient serum without differentiating TSAb and TBAb and it has been difficult to utilize the TRAb values for GD remission study after ATD treatment.

KRONUS ELISA Assay Principle

In this thesis ELISA (TRAb ELISA in row 11 in Table 2) manufactured by KRONUS (Boise ID) is used in the assay comparison. The KRONUS ELISA is based on the 3rd generation competitive binding between direct peroxidase-labeled TSAb monoclonals (M-22) and TRAbs existing in patients' sera. The M-22 was produced from a heterohybridoma [38] fused with the peripheral blood lymphocytes of a GD patient and

mouse/human hybrid cell line (K6H6/B5). The assay begins by adding a start buffer to the detergent solubilized and lyophilized porcine TSHR coated well. Next, patient serum is added to the wells along with a set of calibrators which makes a standard curve to estimate the TRAb concentration and negative and positive controls provided by the manufacturer. One-hour incubation time allows any existing TRAbs the opportunity to bind unoccupied TSHRs. After the incubation and a wash step, M-22, which competitively binds with patients' TRAbs, is added. The peroxidase labeled M-22 binds to unoccupied TSHR during 25 minutes of incubation (refer to Figure 5). Adding chromogenic substrates (3,3',5,5'-Tetramethylbenzidine (TMB)) to each well changes the color. Results from these tests are measured by Optical Density (OD) values and expressed as percent inhibition from Equation 1 or titers interpolated from a standard curve based on NIBSC 90/672 calibrators with concentrations at 1, 2, 8 and 40 IU/L. The KRONUS ELISA product insert recommended the cut-off at 1 IU/L [35].

Principle of Bioassay

A bioassay, specifically a bioreporter assay, is another category of assay used for differential diagnosis of GD. These assays utilize live animal cells stably transfected with human TSHR. The recombinant TSHR sequence has a reporter gene encoded that responds to binding of the TSHR [23, 54]. Unlike competitive binding assays which measure the amount of TRAbs in the patients' sera, bioassays measure stimulus of the TSAbs within the heterogeneous polyclonal mixture of TRAbs [21]. In theory, the measurement of bioassay correlates with the patients' TSAb activity on their thyrocytes.

TSAbs in a patient's serum binding to the ectodomain of the transfected human TSHRs on live animal cells results in the activation of G protein leading to cAMP pathway and inositol phosphate (IP) pathway activation. cAMP stimulates cAMP dependent protein kinase A (PKA) in the cytosol. Stimulated PKA phosphorylates the nuclear transcription factor CREB, activating cAMP response genes [55] Bioreporter assays contain a promoter and a corresponding reporter gene to show a stimulation signal. The cAMP pathway activation results in the production of a chemiluminescent signal, which is measured by a luminometer. Although TSAbs are known to stimulate both cAMP and IP pathways the amount of TSAb necessary to stimulate IP pathway is much more than the amount to stimulate cAMP pathway [20]. Moreover, clinical effects of the IP pathway have not been defined [20]. Therefore bioreporter assays measuring cAMP stimulation by the TSAbs are widely accepted.

The assay measures only TSAbs activating cAMP in the patient serum, which reflects pathogenesis of GD. Although a bioassay seems an attractive alternative, its availability is limited and somewhat specialized [5]. The drawbacks of the bioassays are i) Clinical labs need cell culture facility, ii) the assay takes time due to the cell culture, and iii) Up until 2009, there were no widely available and reliable manufactured reporter gene bioassays [21]. Again in the United States, it is often considered that in addition to the TRAb titer in the patient sera measured by competitive binding assay and clinical symptoms with other hormone measurement is "good enough" [5, 10, 12, 19].

Despite these drawbacks, bioassay for GD has many merits. TSAb activity measurement with administration of ATDs is one application for the bioassay. In past

research, TRAb measurement, even with refined competitive binding assays, did not give positive recommendations for ATD studies [21, 41, 43, 45]. One study with bioassay with a chimeric receptor has shown promising results for ATD prediction [56]. Also, reliable TSAb measurements are expected for monitoring pregnant GD patients and newborns. TSAb measurement is useful since pregnant GD patients tend to have counteractive TBAbs in TRAb mixtures and the fine-tuning of ATD is necessary to avoid the TRAb affecting the fetus, leading to conditions such as neonatal hyperthyroidism [13, 57]. Besides, if GD is reflective of TSAb activity and clinical communities' concerns are convenience and reliability of the bioassay, a fast reliable bioassay can replace the TRAb assay for differential diagnosis of GD.

Evolution of Bioassay

Historically, LATS (former nomenclature for TRAbs) were identified by injecting a patient's serum into guinea pigs and the effect of the TSAbs was measured by radioiodine released from a guinea pig's thyroid *in vivo* [23]. Since then, in order to understand TSH and the TSAb stimulating mechanism in the thyroid, cells were tested on various bioassays using animals *in vivo*, animal thyroid cells or slices [23]. After the invention of cloning technology, animal cells expressing the recombinant TSHR were widely used.

FRTL-5 Rat thyroid cells and Chinese Hamster Ovary Firefly luciferase bioreporter (CHO-Luc) cells stably transfected hTSHR are utilized for research purposes to measure the effect of the TRAbs [47]. However, the sensitivity and specificity of FRTL-5 was not as good as the 2nd or 3rd generation of the TRAb assays [47]. Also,

correlations of TRAb assays and reporter gene bioassays were low. FRTL-5 and CHO-Luc bioassays were also used for identifying the binding sites of TSH, TSAb and TBAb[18, 23-25, 58-61].

CHO-Luc Assay Principle

The CHO-Luc assay uses living cells to detect the stimulus of TSHR by TSAbs in the patient serum (CHO-Luc; Diagnostic Hybrids, Athens OH). CHO-Luc is genetically engineered CHO recombinant cell line that expresses WT human TSHR [25]. CHO-Luc cells exposed to a GD patient's serum (containing TSAbs) results in the activation of the cAMP pathway and production of chemiluminescent signal, which is measured on a luminometer. Results from the CHO-Luc assay are expressed as a percent above a reference control. The CHO-Luc assay is categorized, by the Food and Drug Administration (FDA), an analyte specific reagent (ASR). An ASR is defined as:

> "Antibodies, both polyclonal and monoclonal, specific receptor proteins, ligands, nucleic acid sequences, and similar reagents which, through specific binding or chemical reaction with substances in a specimen, are intended to use in a diagnostic application for identification and quantification of an individual chemical substance or ligand in biological specimens."

in Code of Federal Regulations (CFR) 864.4020 of the FDA. In other words, CHO-Luc assay is "available," however only used for unofficial research purposes.

Besides CHO-Luc being an ASR, CHO-Luc assay detects heterogeneous mixture of the TRAb in the patient sera. A minor population of GD patients is known to have both TSAbs and TBAbs. If the mixture of TSAbs and TBAbs is added to the CHO-Luc, TBAbs interfere with TSAbs to inhibit the TSAb stimulation measured in the assay. As a result, patients who developed TRAbs could be diagnosed as negative due to unknown amount of inhibiting signals. Also, competitive binding assays detect TBAbs in the myxedema patients and a subgroup of Hashimoto's thyroiditis. In order to overcome this problem and provide fast and reliable reporter gene bioassays Diagnositc Hybrids (DHI), in collaboration with Interthyr, has developed a second generation of bioassay termed Mc4 (Thyretain^{TM}) explained in the following section.

Mc4 Assay Principle

In order to improve the specificity of the cell-based assay for the detection of TSAbs only, a new recombinant cell line, CHO-Mc4, was engineered by DHI in collaboration with Interthyr [25]. The principle of the CHO-Mc4 cell line is similar to that of CHO-Luc; however, the CHO-Mc4 cell line contains a chimeric TSHR receptor (with mouse Luteinizing hormone (LH)) and was engineered to contain a deletion in the carboxyl-terminus of the TSHR (261-370) (Figure 4) in a region reported to be bound by TBAbs and TSH based on the chimeric receptor study [25]. The loci of most of the TSAb epitopes are known to be intact [34]. Theoretically, based on the chimeric receptor studies [18, 25, 58-61] with clinical samples, the correlation between positive results in this assay and the existence of TSAbs in a patient's serum should be high since the chimeric receptor eliminates most of the TBAb and TSH binding sites.

CHAPTER 2: OBJECTIVE

The study outlined in this thesis focused on the evaluation of the Mc4 assay and a comparison of performance with the KRONUS ELISA assay (in row 5 in Table 2). Comparison studies with a 3rd generation and earlier generations of competitive binding assays have been done in the past [17, 39, 40, 47]. However, a comparison of the 3rd generation competitive binding ELISA and the cell-based assay, CHO-Luc and Mc4, to our knowledge, has not been done. The goal of this thesis project was to conduct a side by side comparison of these three assays, and to evaluate the overall performance of all three assays when used to test sera from the same set of GD patients. Overall, the results from this study may add more information about differences amongst the assays with different principles.

CHAPTER 3: SUBJECTS AND METHOD

Subjects

Two hundred forty serum samples (mean age, 42.7; range, 2.9 - 80.7 years; 183 females) were obtained from Professor George J. Kahaly at Gutenberg University Medical Center Department of Medicine I in Mainz, Germany. The informed consent processes were conducted based on the protocols by the local ethical committee, and sera were transported, aliquoted and stored at -80°C in the DHI laboratory. The patient information was blinded at the time of experiments.

The first two hundred samples (#1 - #200) were divided into two groups based on their TSH levels retrospectively. One hundred sixty two samples (mean age, 46.6 years; range, 12.7 – 80.7 years; 132 females) came from the low TSH GD patients based on the diagnosis by physicians at Gutenberg University. One hundred sixty two patients' TSH levels were low at the time of blood withdrawal (<0.4 μIU/mL). Some of the 38 patients (mean age, 44.1 years; range, 14.7 – 64.1years; 30 females) had high TSH levels for indeterminate reasons (refer to Figure 6). Forty samples (mean age, 25.4 years; range, 2.9 – 68.4 years; 21 females) came from HBD. Figure 6 shows the scatter plot of the TSH levels of the 240 samples which came in order from the sample provider. Group 1 is the first 162 low TSH GD positive; group 2 is the 38 ambiguously defined patients; and group 3 is the last 40 healthy samples.

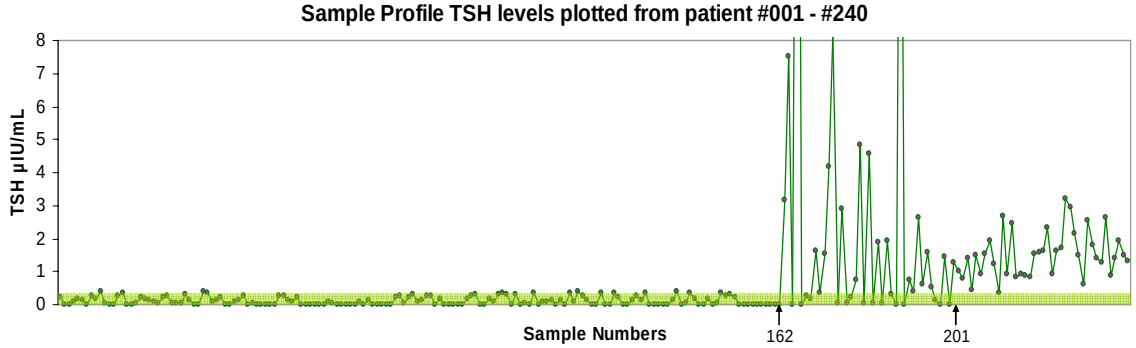

Figure 6. A scatter plot of TSH level from sample #1 to #240. First 162 patients had less than 0.4 µU/mL TSH levels. Samples from # 162 to #200 (shown in the arrow) had TSH levels higher then typical non-treated GD patients. Some of the serum had very high TSH levels (TSH up to 29.63 µU/mL). Samples from #201 to #240 are obtained from HBD.

The samples came with auxiliary serum data which include TSH, T4, T3, anti-Tg, and anti-TPO. None of the donors were pregnant. It is important to note that some of the patients were treated with anti-thyroid drugs at the time of blood withdrawal. However, the individual treatment data and duration information was not provided. It is unfortunate that clinical data was not provided to explain the high TSH levels and treatment (if one was received). However, this study focuses on the agreements among the different assays within the same sample group.

Laboratory Tests

CHO-Luc

The CHO-Luc assay comes with all the reagents necessary. Following the manufacturer's specification, the assay has a 96 well plate format and each plate can measure 13 specimens, positive, reference and negative controls (each with triplicate measurements). It takes a total of 2 and half days to obtain the reading. At day 1, frozen

CHO-Luc cells are thawed in a warm water bath. Thawed cells in suspension are immediately mixed with 5 ml per plate growth media (GM). 100 µl of cell suspension in GM per well are planted into the middle 48 wells and incubated for 24 hours in an incubator with 5% CO_2, 37 C°, and 90% humidity. At day 2, after 24 hours of incubation, confluency of the cells is checked by microscope. GM is discarded and 100 µl of starvation media (SM) per well is added after one wash with SM. The plates are incubated for the next 24 hours. At day 3, frozen serum with positive, negative and reference controls are thawed in a warm water bath. Controls and samples (40 µl each) are prepared and mixed with 400 µl warmed reaction buffer (RB) at room temperature (1:11 dilution). After vortexing the samples and RB mix, the 100 µl of the mixture per well is added to each well in triplicate following the DHI recommended format. The plate with all the control and samples are incubated again for 4 hours. After the incubation, 75 µl of thawed luciferase detection cell lysis buffer at room temperature is added to each well. After 5 minutes of lysis time, the plate is read by a luminometer. Each reading is divided by the reference control reading of the plate to obtain a signal to reference ratio (SRR%).

Due to ASR's flexibility, the cutoff with this sample set was set at 150% (SRR%) by the ROC method (in Appendix A). Two to four plates of cells per day were planted in a batch which could measure 26 – 52 samples.

Mc4

The Mc4 assay comes with all the reagents necessary and is conducted according to manufacturer specification. The Mc4 assay takes over night (less than 24 hours) from

start to finish. Mc4 cells need 16 – 18 hours for incubation omitting the starvation step required for the CHO-Luc assay. At day 1, frozen Mc4 cells are thawed in a warm water bath. Thawed cells in suspension are immediately mixed with 5 ml per plate of GM. 100 µl of cells in GM per well are planted into the middle 48 wells incubated for 16 – 18 hours in an incubator with 5% CO_2, 37 C˚, and 90% humidity. At day 2, after checking confluency of the cells, frozen serum with positive, negative and reference controls are thawed in a warm water bath. Controls and samples (40 µl each) are prepared, and mixed with 400 µl warmed RB at room temperature. After vortexing the samples and RB mix, the 100 µl of the mixture per well is added to each well in triplicate following the DHI recommended format. The plate with all the control and samples is incubated again for 3 hours. After the incubation, 75 µl of thawed luciferase detection cell lysis buffer at room temperature is added to each well. After 10 minutes of lysis time, the plate is read by a luminometer. Each reading is divided by the reference reading of the plate to obtain a SRR%. The sera is considered to be positive when SRR% >= 140%.

Two plates of cells per day were planted in a batch which could measure 26 samples.

KRONUS ELISA

The 3rd generation of the FDA approved competitive binding assay KRONUS TRAb ELISA was conducted according to manufacturer specification. All reagents come with the kit. The assay takes two and half hours with additional time for sample preparation. The product insert (PI) recommends centrifuging the sample sera at 10,000 – 15,000 g for 5 minutes if the serum is cloudy or particulate matters are observed. All

samples were centrifuged prior to the ELISA assay to eliminate the potential of non-specific binding after thawing the sera in a warm water bath. The assay requires 150 µl serum per test; therefore, 200 – 180 µl of serum is centrifuged just in case for the lipemic samples. When a lipid layer is observed, the samples are carefully pipetted out to avoid the lipid layer.

The strips of the 96 well format TSH receptor coated wells are reconstituted with 75 µl start buffer at room temperature. The NIBSC 90/672 TRAb calibrators (concentrations at 1, 2, 8 and 40 IU/L), positive and negative controls came with the kit, and patients' serum, all for 75 µl each are added in duplicate wells. The covered plate is incubated for an hour at room temperature on a plate shaker set at 500 rpm. After the incubation, the plate is shaken out to remove the serum and buffer and then washed with the wash buffer once. After washing, 100 µl reconstituted M-22-peroxidase are added to each well. The plate is incubated at room temp for 25 minutes. After the second incubation, M-22-peroxidace is shaken out following 2 washes with the wash buffer and 1 wash with deionized water. The 100 µl peroxidase substrate is added to the each well by a repeating pipette. Immediately after adding substrate, the plate is incubated 25 minutes in the dark. Incubation in the dark is required for the substrate to develop the blue color. After the incubation, 50 µl of stop solution is added to each well. Within 5 minutes after adding the stop solution, the plates are shaken briefly on the plate shaker (5 seconds), then 450 nm absorbance is read by the ELISA plate reader blanked against a well with 100 µl substrate and 100 µl stop solution.

The PI defers the cut off determination to the users. The options are percent inhibition, linier and 4PL estimations. Since KRONUS PI states that the recommended cut-off is 1 IU/L, a linear standard line estimation was chosen based on the calibrators included in the kit.

Analysis and Rationale

All three assays' end results are binary classified (either positive or negative) derived from a continuous measurement either below or above the cut-off. The same sample set was used for this comparison and test results were scored using pre-determined cut-off values recommended by the manufacturer except the CHO-Luc assay. For the CHO-Luc assay, due to ASR status, a cut-off value was determined by the ROC method (refer to Appendix A). All three tests results (in SRR% or IU/L) were plotted out by GRAPH PAD prism software.

Sensitivity and specificity were calculated based on the 200 "GD positive" samples diagnosed and provided by Gutenberg University against 40 samples from healthy blood donors (HBD) also provided by the Gutenberg University. Sensitivity is the probability of positives correctly diagnosed by referring to the third party diagnosis as a reference standard. Specificity is the probability of negatives correctly diagnosed by referring to the third party diagnosis as a reference standard. The calculation is shown in Table 3. GD is caused by TRAbs; therefore, assays seek its existence or stimulation. However, ATD treatment is known to suppress autoantibody production. Therefore, it is already known that sensitivity and specificity would not be as high compared to other

published studies due to the fact that unknown portions of patients might have been already treated with ATD [34].

For the main focus of this study, KRONUS ELISA and Mc4 were compared in 2 by 2 metrics called a contingency table in order to observe discordances (false positives and false negatives in the middle table in Table 3). The FDA guidelines for clinical trials advise 2 by 2 metrics that show positive percent agreement (PPA) and negative percent agreement (NPA) when one device is compared with the other predicate device. This approach was adopted for providing data in a manner that easily allows comparison of the assay results. The right table in Table 3 shows the example of the PPA and NPA calculation. For example, when 20 samples were tested, 10 of them came out positive by predicate assay and 9 of them came out positive by the evaluating assay. The agreement between them is 90% (9 / 10 =0.9). One disagreement is considered a “false negative” with respect to the predicate assay. NPA is calculated the same way. If PPA and NPA are high, the two assays are considered similar. The discordance between the two tests may arise from the difference between the assays or difference in TRAb population.

In addition, if one device is compared to the reference standard which confirms true disease status (usually a biopsy or physicians diagnosis) the same calculation of the positive portion is called sensitivity and the negative portion is called specificity.

Table 3

Two by Two Contingency Table Format.

Comparing one assay to the reference standard (Se and Sp)

		True Disease Status / Drs' diagnosis	
		Positive	Negative
Assay	Positive	True Positive	False Positive
	Negative	False Negative	True Negative
		Sensitivity TP/(TP+FN)	Specificity TN/(FP+TN)

Comparing one assay to the predicate assay (PPA and NPA)

		Predicate Assay	
		Positive	Negative
Assay	Positive	True Positive	False Positive
	Negative	False Negative	True Negative
		PPA TP/(TP+FN)	NPA TN/(FP+TN)

Example

		Predicate Assay	
		Positive	Negative
Assay	Positive	9	2
	Negative	1	8
		PPA= 90% 9/(9+1)	NPA= 80% 8/(8+2)

Safety Issues

All the assays used in this study required human sera provided by DHI. DHI protocols for human sera handling, biohazards disposal, and general lab practice were followed. DHI operates under a set of self-documented guidelines to meet or exceed the Occupational Safety and Health Administration (OSHA) requirements.

CHAPTER 4: RESULTS

Sensitivity and Specificity

Two hundred “GD positive” patients’ sera diagnosed and provided by Gutenberg University were considered as a reference standard. The patient group is not typical “non treated GD”, especially since there were some patients who had high TSH (Figure 6). The 200 patients were not categorized into subgroups of GD since clinical features of the patients were not provided. Forty negative samples from healthy blood donors (HBD) were also tested to see specificity. Table 4 lists the sensitivity and specificity of all three assays. KRONUS ELISA has two different ways to interpret the plate reading and sensitivity and specificity for each method was listed in Table 4. Therefore sensitivity and specificity calculated with/without including the sample that fell in the indeterminate range as positive was tabulated.

Table 4

Sensitivity and Specificity Based on 200 GD Defined by the Third Party and 40 HBD Samples Provided by the Same Organization.

	Sensitivity	**Specificity**
Mc4	79.5% (159/200)	100% (0/40)
CHO-Luc	84.0% (168/200)	98% (1/40)
KRONUS ELISA (IU/L)	73.4% (146/199)	100% (0/40)
KRONUS (% inhibition)	80.9% (161/199)	95% (2/40)

The results show that a considerable portion (16 – 27%) of the “GD” patients showed negative results (no TRAb existence). Sensitivities were low compared to other

published literature [17, 39, 47, 51] although some of the literature derived optimal cut-off based on the ROC curves and only had a non-treated GD patients group. This low sensitivity is likely due to the differences in GD populations used in this study compared to previous studies. In this study, it is unknown whether or not the GD patients had been treated whereas in previous studies the population was untreated GD patients. This observation highlights the importance of using the same sample population when comparing different assays. That said, the fact that Mc4/CHO-Luc had similar, if not greater sensitivity to KRONUS ELISA, and KRONUS ELISA has been observed to have ~93.3% [47] sensitivity on defined non-treated GD patients, it is reasonable to speculate that the bioassay would have a similar if not grater sensitivity in that patient population.

In Table 4, CHO-Luc gave the highest sensitivity following KRONUS ELISA by the percent inhibition method; however, KRONUS ELISA based on percent inhibition method misdiagnosed two HBD as positive, and CHO-Luc misdiagnosed one HBD as positive out of 40 HBD samples. Mc4 had the highest sensitivity without sacrificing the specificity.

Two by Two Comparisons

Table 5 shows the agreements and disagreements of the results in the Mc4 and KRONUS ELISA. The table is the comparison of all 239 samples (KRONUS ELISA had one sample quantity not sufficient (QNS)). The PPA is very high (95.9%), which indicates that both assays have high sensitivities of detecting the TSAbs presents in the GD patients' sera. There is a relatively high percentage (19 / 93 = 20.4%) of "false positive" results which means Mc4 scored positive but KRONUS ELISA scored

negative. If GD indicates the existence of TSAb causing hyperthyroid state, it does not make sense that samples diagnosed as GD does not have TRAbs measured by KRONUS ELISA. Mc4 seems to pick up more stimulators in the patient sera (out of 199 GD) than KRONUS ELISA using calibrators. The 19 "false positives" may arise from borderline patients underestimated (judged as negative) by the calibrator method discussed in the next section.

Table 5

The 2 by 2 Table Comparing Mc4 and KRONUS ELISA Based on 199 GD (one QNS) and 40 HBD Samples.

2x2 (Quantitative) all 240 Samples

		KRONUS ELISA cut off at 1IU/L	
		Positive	Negative
Mc4 cut off at 140%	Positive	140	19
	Negative	6	74
Total:		146	93
		PPA	**NPA**
		95.9%	**79.6%**
Overall Agreement:			**89.5%**

The 6 "false negatives" may arise from heterogeneity of the TRAbs. When stimulator, blocker and neutral TRAbs co-exist in a patient serum, competitive binding assays will detect the mixture of TRAbs, whereas Mc4 may result as negative if a

stimulator is not present. The portion of the “false negatives” were small (6) compared to the “false positives” (19).

Moreover, two of the 6 “false negative” samples included lipemic samples (samples 91 and 93) and may have been interfered with labeled M-22 binding and generated as positive by ELISA. However, these two samples were positive by all other assays except Mc4. One sample (sample 107) was negative by Mc4 but positive by CHO-Luc and KRONUS ELISA. Again, without clinical picture of the patient, it cannot be concluded whether discordance is due to the assay mechanism or the different biological components.

Table 6 shows the agreements and disagreements of the results in the Mc4 and CHO-Luc. The larger number of “false negative” (14) was observed compared to Mc4 vs. KRONUS ELISA companion. The 14 “false negative” may arise from higher variation and SSR% obtained by CHO-Luc when compared with Mc4. The 14 “false negative” discordances may or may not be due to the receptor (WT vs. chimeric) differences. Seven samples out of 14 “false negatives” were positive only by CHO-Luc but all negative by the other assays.

Table 6

The 2 by 2 Tables Comparing the Agreements of Mc4 and CHO-Luc Based on 200 GD and 40 HBD samples.

2x2 (Quantitative) all 240 Samples

		CHO-Luc cut off 150%	
		Positive	Negative
Mc4 cut off at 140%	Positive	155	4
	Negative	14	67
Total:		169	71
		PPA	**NPA**
		91.7%	**94.4%**
Overall Agreement:		**92.5%**	

With Respect to TSH Value

One of the problems with this 200 GD samples were that some of the patients had relatively high TSH with respect to typical non treated GD patients. Therefore, 200 GD patients were divided by different TSH levels: TSH < 0.01 µU/mL, TSH <= 0.4 µU/mL, TSH > 0.4 µU/mL, and HBD groups. TSH < 0.01 µU/mL were arbitrarily chosen since some of the assay comparison literature [17] with high sensitivity and specificity had only non treated GD patients with TSH less than 0.01 µU/mL. The numbers of samples scored as positives (negatives) with all three assays are listed in Table 7.

Table 7

Numbers of Positives and Negatives Scored by all Three Assays are Listed within the Defined Groups of GD Patients with Respect to TSH levels. In this table KRONUS ELISA was scored based on the calibrator method (positive as more than 1 IU/L).

	TSH < 0.01	TSH <= 0.4	TSH > 0.4	HBD
All Positive	54	124	13	0
All Negative	2	25	1	38
Disagreements	8	31	6	1
QNS or not measured	1	1	0	0
Total Samples	65	181 (one HBD)	20	39
% of agreements	87.7%	82.9%	70.0%	97.4%

Table 8

Numbers of Positives and Negatives Scored by All Three Assays are Listed within the Defined Groups of GD Patients with Respect to TSH Levels. In this table KRONUS ELISA was scored based on the percent inhibition method (positive as 15%).

	TSH < 0.01	TSH <= 0.4	TSH > 0.4	HBD
All Positive	58	132	14	0
All Negative	2	23	1	36
Disagreements	4	25	5	3
QNS or not measured	1	1	0	0
Total Samples	65	181 (one HBD)	20	39
% of agreements	93.8%	86.2%	75.0%	92.3%

When percent inhibition method was used the agreements of the three assays increased, from 87.7% to 93.8% within the TSH < 0.01 µU/mL group. The results were separated by TSH levels. For comparison the number of samples diagnosed as positive for each assay are listed in Table 9.

Table 9

Within each Group of Samples with Different TSH Level. This table lists how many of the samples are scored as positive by each assay.

Total # of samples within the TSH range	TSH < 0.01	TSH <= 0.4	TSH > 0.4	HBD
	65	181 (one HBD)	20	40
Numbers of samples scored as positive				
Mc4	62	141	18	0
CHO-Luc	60	151	17	1
KRONUS (%inhibition) one QNS	56	145	15	2
KRONUS (IU/L) one QNS	60	132	14	0

The numbers of samples scored as positive within the low TSH (65) samples (presumably untreated GD patients), Mc4 was the highest (62) compared with other assays.

Inter plate Coefficient of Variation

From the assays of the 240 samples, CVs of positive and negative controls provided by the manufacturer were compared to determine how much plate to plate signal differences in the final readings are present. Mc4 and KRONUS ELISA had almost the same CV values for their positive and negative controls. Negative control CV%

((standard deviation / mean)*100%) were higher due to the smaller mean values with respect to the variation.

Table 10

Between Plate CV% of Positive Controls and Negative Controls Provided by the Manufacturer.

	Number of plates	Positive control CV%	Negative control CV%
CHO-Luc	19	26.8%	13.4%
Mc4	20	12.0%	22.5%
KRONUS ELISA calibrated	8	11.5%	23.1%
KRONUS ELISA % ihhibition	8	6.3%	N/A

CHAPTER 5: DISCUSSION

Three assays were compared side by side experiments. The 2 by 2 table comparison had relatively large “false positives” by Mc4 vs. KRONUS ELISA and relatively large “false negatives” by Mc4 vs. CHO-Luc based on the KRONUS ELISA and CHO-Luc as predicate device.

Difference in the Percent Inhibition and Calibration Method

The KRONUS product insert indicates several different methods to interpret the results as discussed in Chapter 1. Antibody Estimation. After conducting the assay, the linear standard line estimation and percent inhibition method were compared. The results are in shown in

Table 11. Seventeen discordances were observed with 239 samples (one Quantity Not Sufficient (QNS)) samples. Although the data came out of the same plate, different calculation methods generated the different diagnosis. The tendency was that linear standard line estimation with calibrators tends to generate negative results (less than 1 IU/L), whereas the percent inhibition method results in positive (more than 15% recommended by the KRONUS customer service). In addition although the percent inhibition method had high sensitivity the method sacrificed specificity risking a healthy person diagnosed as positive (2 samples came out positive from 40 HBD group). KRONUS ELISA calibration method had a tendency to show negative results when judging gray zone positive evaluated by the other assays (data not shown).

Table 11

The 2 by 2 Tables Showing the Agreements of Linear Calibration and % Inhibition Methods Based on 199 GD and 40 HBD Samples.

2x2 (Quantitative)

		KRONUS ELISA cut off at 1 IU/L		
		Positive	Negative	
KRONUS ELISA cut off at 15%	Positive	146	17	
	Negative	0	76	
Total:		146	93	239
		PPA	**NPA**	
		100.0%	**81.7%**	

Lipemic Samples

Several lipemic samples generated very inconsistent results evaluated by the 4 different assays. It was difficult to differentiate if lipemic samples interfered with the assay or there were other factors (TBAbs included in the TRAbs). Since clinical information was not available, no assumption could not be made. Although lipid screening and removal was attempted, it was hard to remove the entire lipid in the serum due to the limited amount of sera. Especially for the ELISA assay, lipid may have interfered with the assay affecting results. This is a considerable factor for clinical laboratories which handles large volumes of sera.

Overall Mc4 Performance

Considering the inconsistency observed by the CHO-Luc and KRONUS ELISA calculation methods, Mc 4 had comparable CV% and high sensitivity compared with other competitive binding assays. Mc4 had shown the comparability and utility of the bioassay.

CHAPTER 6: CONCLUSION

In this study, two bioassays and the 3rd generation competitive binding assay (KRONUS ELISA) were compared side by side. All the results were dichotomous (positive or negative) converted based on the manufacturer's suggestions. The results of three assays were compared. Overall sensitivity was similar, except two assays (the KRONUS ELISA percent inhibition method and CHO-Luc) had lower specificity which may generate the problems for large volume screening. Mc4, the latest bioassay has comparable results showing promising utility for the clinical setting.

A significant discordant result seemed to have been created by inconsistent results due to the KRONUS ELISA having several ways to interpret the signal. If calibrator used for KRONUS ELISA, the result disagreed with the other three assays, however if the percent inhibition method was used more positive results were concordant but specificity became lower risking diagnosing the HBD as positive. As explained in Chapter 1 Evolution of Competitive Binding Assays, 5 points calibrator assay instead of 4 points might generate higher sensitivity (In Table 2 rows 9, 10 and 11) instead of 4 points used in KRONUS ELISA assay (row 12). Or, the original version of the 3rd generation ELISA assay (In Table 2 rows 9 and 10) may be superior to the direct labeling (rows 11 and 12). More comparison studies on the same sample set are necessary to determine if slight differences in the assay format will make a large difference in performance.

In about one third of the GD samples, discordances of binary results (positive or negative) were observed in at least one out of all three assays. Several discordant results were retested and analyzed by dilution studies (data not included in thesis) showing

possible TBAb existence in a small percentage of the patients. Combination of the assay with different principles seemed to show TBAb containing sera which may explain the clinical symptoms of the euthyroid GD patients. Further studies with large sets of samples with defined clinical information are necessary to define the heterogeneity of the TRAb and further typing the GD patients as well as monitoring the ATD treatment, pregnant GD patients. Currently there is no data correlating bioassay stimulation signal and patients symptom with respect to TRAb concentration. In the future, a large bank of assay results with patient info may lead to delineate the characteristics of sub population of GD.

REFERENCES

1. Eisenbarth, G.S., *Immunobiology of Autoimmunity*, in *Molecular mechanisms of endocrine and organ specific autoimmunity*, G.S. Eisenbarth, Editor. 1999, R.G. Landes: Austin TX. p. 1-2.
2. Pollard, K.M., *Autoimmunity*, in *Autoantibodies and autoimmunity: molecular mechanisms in health and disease*, K.M. Pollard, Editor. 2006, Wiley-VCH: Weinheim, Germany. p. 8.
3. Jacobson, D.L., et al., *Epidemiology and Estimated Population Burden of Selected Autoimmune Diseases in the United States.* Clinical Immunology and Immunopathology, 1997. **84**(3): p. 223.
4. Weetman, A.P., *Autoimmune thyroid disease.* Autoimmunity, 2004. **37**(4): p. 337.
5. Ajjan, R. and A. Weetman, *Techniques to quantify TSH receptor antibodies.* Nature Clinical Practice Endocrinology & Metabolism, 2008. **4**: p. 461-468.
6. Davies, T.F., et al., *Thyrotropin receptor-associated diseases: from adenomata to Graves disease.* The Journal Of Clinical Investigation, 2005. **115**(8): p. 1972.
7. Weetman, A.P., *Graves' disease.* The New England Journal of Medicine, 2000. **343**(17): p. 1236.
8. Shomon, M.J., *Living well with Graves' disease and hyperthyroidism: what your doctor doesn't tell you-- that you need to know*. 2005, NewYork NY: HarperCollins. 427.
9. Siegenthaler, W., *Differential diagnosis in internal medicine*. Disorders of the Head and Neck, ed. G.A. Spinas, P. Ott, and S.J. Stoeckli. 2007, NY: Thieme. 1104.
10. Brent, G.A., *Graves' Disease.* The New England Journal of Medicine, 2008. **358**(24): p. 2594-2605.
11. Gross, M.D., et al., *Radioiodine therapy of Grave's disease*, in *Graves' Disease Pathogenesis and Treatment*, B. Rapoport and S.M. McLachlan, Editors. 2000, Kluwer Academic Publishers: Norwell MA. p. 154 -155.
12. Ladenson, P.W., et al., *American Thyroid Association Guidelines for Detection of Thyroid Dysfunction.* Archives of Internal Medicine, 2000. **160**(11): p. 1573-1575.
13. Kamijo, K., *TSH-receptor Antibodies Determined by the First, Second and Third Generation Assays and Thyroid-stimulating Antibody in Pregnant Patients with Graves' Disease.* Endocrine Journal, 2007. **54**(4): p. 619.
14. Volpe, acute, and Robert, *B-Lymphocytes in Autoimmune Thyroid Diseases (AITD).* Reviews in Endocrine and Metabolic Disorders, 2000. **1**(1/2): p. 79.
15. Adams DD, P.H., *Abnormal responses in the assay of thyrotropin.* Proceedings of the University of Otago Medical School, 1956. **34**: p. 11-12.
16. Smith, B.R., S.M. McLachlan, and J. Furmaniak, *Autoantibodies to the Thyrotropin Receptor*. Endocrine Reviews, 1988. **9**(1): p. 106-121.
17. Kamijo, K., K. Ishikawa, and M. Tanaka, *Clinical evaluation of 3rd generation assay for thyrotropin receptor antibodies: The M-22-biotin-based ELISA initiated by Smith.* Endocrine Journal, 2005. **52**: p. 525-529.

18. Kohn, L.D. and N. Harii, *Thyrotropin Receptor Autoantibodies (TSHRAbs): Epitopes, Origins and Clinical Significance.* Autoimmunity, 2003. **36**(6/7): p. 331.
19. Davies, T.F., *Thyroid-Stimulating Antibodies Predict Hyperthyroidisma.* Journal of Clinical Endocrinology & Metabolism, 1998. **83**(11): p. 3777-3781.
20. Gupta, M.K., *Thyrotropin-receptor antibodies in thyroid diseases: advances in detection techniques and clinical applications.* International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine, 2000. **293**(1-2): p. 1-29.
21. Sinclair, D., *Analytical aspects of thyroid antibodies estimation.* Autoimmunity, 2008. **41**(1): p. 46.
22. Michaleka, K., et al., *TSH receptor autoantibodies.* Autoimmunity Reviews, 2009. **Article in Press, Corrected Proof**.
23. Rapoport, B., et al., *The Thyrotropin (TSH)-Releasing Hormone Receptor: Interaction with TSH and Autoantibodies.* Endocrine Reviews, 1998. **19**(6): p. 673-716.
24. Kim, W.B., et al., *The Prevalence and Clinical Significance of Blocking Thyrotropin Receptor Antibodies in Untreated Hyperthyroid Graves' Disease.* Thyroid, 2000. **10**(7): p. 579-586.
25. Tahara, K., et al., *Immunoglobulins from Graves' disease patients interact with different sites on TSH receptor/LH-CG receptor chimeras than either TSH or immunoglobulins from Idiopathic myxedema patients.* Biochemical and biophysical research communications, 1991. **179**: p. 70-76.
26. Tonacchera, M., et al., *Patient with monoclonal gammopathy, thyrotoxicosis, pretibial myxedema and thyroid-associated ophthalmopathy; demonstration of direct binding of autoantibodies to the thyrotropin receptor.* European Journal of Endocrinology, 1996. **134**(1): p. 97-103.
27. Smith, B.R., J. Sanders, and J. Furmaniak, *TSH Receptor Antibodies.* Thyroid, 2007. **17**(10): p. 923-938.
28. Morshed, S.A., R. Latif, and T.F. Davies, *Characterization of Thyrotropin Receptor Antibody-Induced Signaling Cascades.* Endocrinology, 2009. **150**(1): p. 519-529.
29. Ma, R., R. Latif, and T.F. Davies, *Thyrotropin-Independent Induction of Thyroid Endoderm from Embryonic Stem Cells by Activin A.* Endocrinology, 2009. **150**(4): p. 1970-1975.
30. Davies, T., R. Marians, and R. Latif, *The TSH receptor reveals itself.* The Journal of Clinical Investigation, 2002. **110**(2): p. 161.
31. Rapoport, B. and S.M. McLachlan, *The Thyrotropin Receptor in Graves' Disease.* Thyroid, 2007. **17**(10): p. 911.
32. Morgenthaler, N.G., S.C. Ho, and W.B. Minich, *Stimulating and Blocking Thyroid-Stimulating Hormone (TSH) Receptor Autoantibodies from Patients with Graves' Disease and Autoimmune Hypothyroidism Have Very Similar Concentration, TSH Receptor Affinity, and Binding Sites.* Journal of Clinical Endocrinology & Metabolism, 2007. **92**(3): p. 1058-1065.

33. Tozzoli, R., *Recent advances in diagnostic technologies and their impact in autoimmune diseases.* Autoimmunity Reviews, 2007. **6**(6): p. 334.
34. Rapoport, B. and S.M. McLachlan, *Thyroid Autoantibodies in Graves' disease*, in *Graves' Disease Pathogenesis and Treatment*, B. Rapoport and S.M. McLachlan, Editors. 2000, Kluwer Academic Publishers: Norwell MA. p. 43-66.
35. *TSH Receptor Autoantibody (TRAb) ELISA Kit [product insert]*. 2008, KRONUS: Boise, ID.
36. Wang, D. and S.J. Soong, *Comparisons of titer estimation methods for multiplexed pneumococcal opsonophagocytic killing assay.* Computational Statistics & Data Analysis, 2008. **52**(11): p. 5022-5032.
37. Law, B., *Immunoassay: a practical guide*. 1996, Boca Raton, FL: CRC Press. 222.
38. Sanders, J., et al., *Human monoclonal thyroid stimulating autoantibody.* The Lancet, 2003. **362**(9378): p. 126.
39. Smith, R., Bernard, et al., *A New Assay for Thyrotropin Receptor Autoantibodies.* Thyroid, 2004. **14**(10): p. 830-835.
40. Hermsen, D., et al., *Technical evaluation of the first fully automated assay for the detection of TSH receptor autoantibodies.* Clinica Chimica Acta, 2009. **401**(1-2): p. 84.
41. Muroi, R., et al., *Thyrotropin receptor antibodies measured by second-generation assay are highly specific for Graves' disease and correlate with serum thyrotropin levels in patients with Graves' disease in remission and during maintenance therapy.* Jichi Medical University journal, 2007. **30**: p. 19-28.
42. Morgenthaler, N.G., et al., *Detection of low titre TBII in patients with Gravesâ€™ disease using recombinant human TSH receptor.* Clinical Endocrinology, 2002. **57**(2): p. 193.
43. Schott, M., et al., *Detecting TSH-Receptor Antibodies with the Recombinant TBII Assay: Technical and Clinical Evaluation.* Hormone and Metabolic Research, 2000. **32**(10): p. 429.
44. Costagliola, S., et al., *Second Generation Assay for Thyrotropin Receptor Antibodies Has Superior Diagnostic Sensitivity for Graves' Disease.* Journal of Clinical Endocrinology & Metabolism, 1999. **84**(1): p. 90-97.
45. Zimmermann-Belsing, T., et al., *Use of the 2nd generation TRAK human assay did not improve prediction of relapse after antithyroid medical therapy of Graves' disease.* European Journal of Endocrinology, 2002. **146**(2): p. 173-177.
46. Bolton, J., et al., *Measurement of Thyroid-stimulating Hormone Receptor Autoantibodies by ELISA.* Clin Chem, 1999. **45**(12): p. 2285-2287.
47. Preissner, C.M., et al., *Comparison of Thyrotropin-Receptor Antibodies Measured by Four Commercially Available Methods with a Bioassay That Uses Fisher Rat Thyroid Cells.* Clinical Chemistry, 2003. **49**(8): p. 1402-1404.
48. Villalta, D., et al., *Analytical and diagnostic accuracy of "second generation" assays for thyrotrophin receptor antibodies with radioactive and chemiluminescent tracers.* Journal of Clinical Pathology, 2004. **57**(4): p. 378-382.

49. Hermsen, et al., *Comparison of a Solid Phase Human- versus Porcine-Thyrotropin Receptor-based Immunoassay for the Measurement of Thyrotropin Receptor Antibodies in Patients with Thyroid Diseases.* Experimental and Clinical Endocrinology and Diabetes, 2008. **116**: p. S59.
50. Hanley, J.A. and B.J. McNeil, *The meaning and use of the area under a receiver operating characteristic (ROC) curve.* Radiology, 1982. **143**: p. 29-36.
51. Liu, C., et al., *Comparison of M22-based ELISA and Human-TSH-Receptor-based Luminescence Assay for the Measurement of Thyrotropin Receptor Antibodies in Patients with Thyroid Diseases.* Hormone and Metabolic Research, 2008. **40**(07): p. 479.
52. ouml, et al., *M22 based (manual) ELISA for TSH-receptor antibody (TRAb) measurement is more sensitive than 2nd generation TRAb assays.* Clinica Chimica Acta, 2009. **403**(1-2): p. 266.
53. Kamijo, K., *TSH-Receptor Antibody Measurement in Patients with Various Thyrotoxicosis and Hashimoto's Thyroiditis: a Comparison of Two Two-Step Assays, Coated Plate ELISA Using Porcine TSH-Receptor and Coated Tube Radioassay Using Human Recombinant TSH-Receptor.* Endocrine Journal, 2003. **50**(1): p. 113.
54. Chazenbalk, G.D., et al., *Evidence for negative cooperativity among human thyrotropin receptors overexpressed in mammalian cells.* Endocrinology, 1996. **137**(11): p. 4586-4591.
55. Kondo, T., S. Ezzat, and S.L. Asa, *Pathogenetic mechanisms in thyroid follicular-cell neoplasia.* Nature Reviews Cancer, 2006. **6**(4): p. 292.
56. Kim, T.Y., et al., *Epitope Heterogeneity of Thyroid-Stimulating Antibodies Predicts Long-Term Outcome in Graves' Patients Treated with Antithyroid Drugs.* Journal of Clinical Endocrinology & Metabolism, 2003. **88**(1): p. 117-124.
57. Kung, A.W.C., K.S. Lau, and L.D. Kohn, *Epitope Mapping of TSH Receptor-Blocking Antibodies in Graves' Disease That Appear during Pregnancy.* Journal of Clinical Endocrinology & Metabolism, 2001. **86**(8): p. 3647-3653.
58. Grasso, Y.Z., et al., *Epitope Heterogeneity of Thyrotropin Receptor-Blocking Antibodies in Graves' Patients as Detected with Wild-Type versus Chimeric Thyrotropin Receptors.* Thyroid, 1999. **9**(6): p. 531-537.
59. Kim, W.B., et al., *Changes in Epitopes for Thyroid-Stimulating Antibodies in Graves' Disease Sera During Treatment of Hyperthyroidism: Therapeutic Implications.* Journal of Clinical Endocrinology & Metabolism, 1997. **82**(6): p. 1953-1959.
60. Watanabe, Y., et al., *Subtypes of Anti-TSH Receptor Antibodies Classified by Various Assays Using CHO Cells Expressing Wild-Type or Chimeric Human TSH Receptor.* Thyroid, 1997. **7**(1): p. 13-19.
61. Kim, W.B., et al., *Epitopes for thyroid-stimulating antibodies in Graves' sera: a possible link of heterogeneity to differences in response to antithyroid drug treatment.* Journal of Clinical Endocrinology & Metabolism, 1996. **81**(5): p. 1758-1767.

APPENDIX A: CHO-LUC ROC ANALYSIS

ROC curve is constructed to determine the optimal cut-off point where the test has given the highest true positive rate and the lowest false positive rate based on the diagnosis given by the sample provider as a reference [50]. Table 12 shows the data which generated the ROC curve on the right side. The table shows the list of sensitivities and specificities with respect to the alternative cut-off points. The optimal point is SRR% at 150% where sensitivity is 82.7% and specificity is 97.5%. The data did not include the 38 patients' data some of which had high TSH in Figure 6. For the cut-off analysis, abnormal GD patients were avoided not to skew the data.

Table 12

The ROC Curve and Data. The optimal point is set at 150% where sensitivity and specificity are optimal.

Cut off	*Neg*	*Pos*	*Se*	*Sp*	*1-Sp*
			1	*0*	*1*
90%	6	162	100.0%	0.15	0.85
100%	11	159	98.1%	0.275	0.725
110%	19	155	95.7%	0.475	0.525
120%	26	147	90.7%	0.65	0.35
130%	35	144	88.9%	0.875	0.125
140%	38	136	84.0%	0.95	0.05
150%	39	134	82.7%	0.975	0.025
160%	40	129	79.6%	1	0
170%	40	126	77.8%	1	0
180%	40	124	76.5%	1	0
190%	40	123	75.9%	1	0
200%	40	123	75.9%	1	0
210%	40	121	74.7%	1	0
220%	40	116	71.6%	1	0
230%	40	114	70.4%	1	0
240%	40	111	68.5%	1	0
250%	40	110	67.9%	1	0
260%	40	110	67.9%	1	0
270%	40	108	66.7%	1	0
280%	40	107	66.0%	1	0
290%	40	107	66.0%	1	0
300%	40	106	65.4%	1	0
310%	40	0	0.0%	1	0