研究生物化學特性和結構特性來探討黑腹果蠅蛋白質酪氨酸亞硫酸化酵素

學生: 林資翔　　指導教授: 楊裕雄 教授

國立交通大學分子醫學與生物工程研究所碩士

摘要

蛋白質酪氨酸亞硫酸化酵素位於細胞內反式高爾基氏網，催化蛋白質中酪氨酸亞硫酸化反應，為一調控細胞外蛋白質交互影響力之重要因子，且調控許多生理上重要功能，例如：發炎作用，人類免疫系統缺乏病毒的入侵及甲狀腺機能減退侏儒症。但因缺乏同質性蛋白質酪氨酸亞硫酸化酵素來瞭解其生化上特性，使其在分子層級上之資訊所知甚少。在我的論文研究中，利用一凝血蛋白脢去除融合蛋白- 轉錄延長因子，首次能夠得到同質性黑腹果蠅蛋白質酪氨酸亞硫酸化酵素。藉由此瞭解黑腹果蠅蛋白質酪氨酸亞硫酸化酵素之酵素動力學，蛋白質四級結構，酵素穩定度和受質調控的特性。經分子篩層析法指出於溶液下其具有兩種結構, 且在鹽與甘油的存在下得以穩定，目前已可以將其分離用於日後的研究。在人類與黑腹果蠅蛋白質酪氨酸亞硫酸化酵素上之點突變 H269Q，H267Q 並不會影響其比活性，但在大腸桿菌內表現量大幅降低導致其總活性隨之大幅減少。而相同的點突變被報導在家鼠上會造成侏儒症，推測可能的原因是此點突變會影響酵素的穩定度或表現量。

Biochemical and structural characterization of *Drosophia melanogaster* tyrosylprotein sulfotransferase

Student: Tsu-Hsiang Lin Advisor: Prof. Yuh-Shyong Yang

Department of Biological Science and Technology and Institute of Molecular Medicine and Bioengineering, National Chiao Tung University, Hsinchu, Taiwan, ROC

ABSTRACT

Protein tyrosine sulfation, mediated by tyrosylprotein sulfotransferase (TPST) that resides in *trans-Golgi* network, is a key modulator of extracellular protein-protein interactions and consequently regulates various physiological functions including inflammation, HIV infection, and hypothyroidism related dwarfism. Limited information at molecular level is available due to the lack of homogenous TPST for detailed biochemical characterization. In this study, a truncated *Drosophila melanogaster* TPST (*Dm*TPST) was first prepared following thrombin proteolysis to remove NusA fusion protein. The kinetics, structure, stability, and substrate regulation of *Dm*TPST were characterized. The result of gel filtration indicated that there were two configurations of *Dm*TPST were simultaneously presented in the solution and could be isolated for future studies. *Dm*TPST can be stabilized with salt and glycerol. A hypothyroidism-related mutation in *Dm*TPST and *h*TPST2 did not cause any loss of

specific activity. However, the total TPST activity was significantly decreased following its expression in *E. coli*. Similar mutation has been reported to cause dwarfism in mouse. It is proposed that such mutation may affect the stability or expression of TPST.

ACKNOWLEDGEMENT

“對酒當歌，人生幾何？譬如朝露，去日苦多。”這首曹操的短歌行前兩句正是我目前的心情寫照。相對於一般的碩士生，我在碩士的求學生涯多花了一年時間。當然禍兮福所倚，這也讓我更努力向實驗室的同仁們學習更多知識與做人態度。因此我需要感謝的人太多了，首先我的指導教授楊裕雄老師，很感謝你當初願意收我這個沒有生化背景的學生，也容許我一再的失敗。並且提供我們一個充滿自由研究思考的環境，讓我們可以自動自發的思考與研究，相信這對我以後工作是一個不能或缺的能力。

我也很感謝王雲銘教授，汪宏達教授，陳俊榮教授能擔當我的口試委員，並對我的論文提出相當多的意見，讓我知道我仍有相當進步的空間。尤其汪宏達教授對我的論文提出相當多的建議，讓我思考到許多我所慮不足之處，相信你的建議不只對目前的碩士生涯，也對我以後的人生多有助益。同時我也感謝很多幫過我的學長們，國家同步輻射的江庭蔚學長，很遺憾沒有結出TPST蛋白質晶體，沒有幫到你真是讓我覺得有點對不起你。已畢業的交大光電所博士研究生張育維學長已很感謝你之前教我不少關於 CMOS phtodiode 的相關知識。有你們的幫助也讓我的學習的更加多元豐富。

LEPE 實驗室的夥伴們感謝你們陪我走過這段有著歡樂與辛酸的路。最先要感謝的學長，當然是在我加入 TPST 這組後，一路帶我做實驗和傳授我許多知識

的呂陸宜(LE)大學長，同時也跟我分享很多人生的觀念跟態度，不時的指正我的缺失，對你的感激之情難以言喻。人超好的普普學長，在我當初沒有題目時，帶我做 2-D electrophoresis，至今我都記得當時的情景。豁達的的蕭博士程允學長，你不藏私的給予許多人生觀念與態度，讓我對你也十分的佩服，另外就是你也與我有類似的經歷，呵。愛打籃球的小志學長，你對秀華學姊的求婚是實在是我碩士生活中難忘的經驗。爬山涉水的淵仁學長，你讓我很羨慕，因爲你讓我見識到原來博士生也可以這麼健康地去很多地方做戶外活動。我們實驗室招牌-小胖, 感謝提供我實驗生活很多有趣的畫面，你的心地很好，是個好人(你又多一張卡了，恭喜)。大美女咏馨小姐，多謝你常幫我們安排研究生活的娛樂活動，不過有一陣子實在喝太大了，我喝到有點胃痛。我的房東的兒子-少東，有你在的 party 決不冷場，不過你喝茫後的樣子我有點不敢恭維。以上三位以後 TPST 組的未來就靠你們發揚光大拉，希望之後我回來吃尾牙時你們都有豐碩的成果。有緣的康寧, 希望你早點找到有緣人拉，趕快脫離單身。糟糕的 A 哥燦，恭喜你也敖出頭畢業拉，希望你之後找到適合你的工作。小蘋果欣怡，很可惜沒有跟你出去玩過，希望你之後開心點有朝氣，可以跟下面那位看齊，呵。立志當貴婦的小萍老師，你的無厘頭讓實在很有趣，不過你的努力我們也有看到。可愛的綺綺，你也是很認真，最近也越來越有學姊的架式拉。格鬥天王奇叡，趕快補點身體吧，感覺你快被榨乾了，不過你稱住這一人小組也挺利害的。

最後不能避俗，一定要感謝我的爸媽們從小養育我跟我弟，有你們的栽培才

有今天的我們，我們會努力不讓你們失望的。還有老弟，你不但是我弟也可以陪我聊很多心事，並可談古論今，很高興我們都有此見識。在此我 cite 陸宜的名言跟大家分享:

磨練才能熟練，投入才能深入，付出才能傑出！

CONTENTS **PAGE**

CONTENTS OF FIGURES AND TABLES

FIGURES

TABLES

ABBREVIATIONS

Abbreviation and Symbol	Full name
ε	Absorption (exitinction) coefficient
A_{280}	Absorption at 280 nm
A_{600}	Absorption at 600 nm
CCR5	Chemokine (C-C motif) receptor 5
D. melanogaster	*Drosophila melanogaster*
*Dm*TPST	*Drosophila melanogaster* tyrosylprotein sulfotransferase
EDC	Ethyl-3-(3-dimethylaminopropyl) Carbodiimide HCl
gp120	Glycoprotein 120
BL21 (DE3)	*E. coli* BL21(DE3)
HIV	Human immunodeficiency virus
*h*TPST	*Human* tyrosylprotein sulfotransferase
k_{cat}	Turnover number
kDa	Kilodaton
K_m	Michaelis constant
MALDI-TOF	Matrix-assisted laser desorption ionization-time of flight mass
MES	2-[N-morpholino] ethanesulfonic acid
PAGE	Polyacrylamide gel electrophoresis
PAP	Adenosine 3',5'-diphosphate
PAPS	3'-phosphoadenosine 5'-phosphosulfate
PSGL-1	P-selectin glycoprotein ligand-1
PTM	Post-translational modification
SDS	Sodium dodecyl sulfate
SULTs	Sulfotransferases
TSH	Thyroid stimulating hormone
TSHR	Thyroid stimulating hormone receptor
TPST	Tyrosylprotein sulfotransferase
V_{max}	Maximum velocity
FPLC	Fast protein liquid chromatography

INTRODUCTION

Sulfation is a widespread biological reaction responsible for many important physiological functions, such as hormone regulation, signal transduction, viral entry, and molecular recognition (Kansas et al., 1996; Wang et al., 2005; Ueoka et al., 2000). Sulfotransferases, whose chemical reaction is somewhat similar to kinases, catalyze the transfer of a sulfuryl group (SO_3^-) from a donor molecule, usually 3'-phosphoadenosine 5'-phosphosulfate (PAPS), to a variety of amine and hydroxyl substrates as nucleophiles (Fig. 1). In vertebrates, there are two classes of 3sulfotransferases: cytosolic and membrane-associated sulfotransferases. Cytosolic sulfotransferases catalyze small endogenous and exogenous compounds, such as drugs, steroid hormones, chemical carcinogens, bile acids, and neurotransmitters (Chapman et al., 2004). Membrane-associated sulfotransferases catalyze the sulfation of macromolecules, such as carbohydrates, peptides and proteins, and are mainly membrane-bound forms localized in Golgi apparatus. Although there are enormous amount of sulfated proteins in biological system, very little information about their biological functions either on metabolic pathways or physiological significances is available. While sulfation is vital for various physiological regulations, hydrolysis of sulfate esters catalyzed by arylsulfatase (ARS) also linked to many important cellular functions including bioactivation of endogenous compounds, cellular degradation, and

modulation of signaling pathways (Hanson et al., 2004). In particular, how the interplay between various members of sulfotransferases and ARS enzyme families regulates the availability and biological activity of xenobiotics and endogenous molecules remains poorly understood.

Pal peptides, such as gastrin, phyllokinin, cholecystokinin, and caerulein (Gregory et al., 1964; Anastasi et al., 1966; Mutt et al., 1968; Anastasi et al., 1968). Post-translational tyrosine *O*-sulfation of proteins was mediated by the enzyme, namely tyrosylprotein sulfotransferase (TPST, EC 2.8.2.20), which localizrotein tyrosine sulfation was first observed by Bettelheim in bovine fibrinopeptide B in 1954 (Bettelheim et al., 1954). The enzyme catalyzes the transfer of sulfate group to the hydroxyl group of a tyrosine residue to form a tyrosine sulfate ester and a 3'-phosphoadenosine-5'-phosphosphate (PAP) from the universal sulfate donor adenosine 3-phosphate 5- phosphosulfate (PAPS) (**Fig. 1**) (Lee et al., 1983). In 1960s, the protein sulfation was detected as tyrosine *O*-sulfate in severed in *trans*-Golgi network (Lee et al., 1983). Protein tyrosine sulfation has been known to take place in a variety of organisms including prokaryotes and multicellular species (Lee et al., 1983). The target proteins belong to the classes of lysosomal proteins, secretory, and plasma membrane, which reflects their intracellular localizations. As compared to phosphorylation, there is much less information in sulfation either on its biochemical

characterizations or biological functions. So far there are two distinct human TPSTs, namely *h*TPST1 and *h*TPST2, have been identified. They are in similar size (370–377 residues) (Ouyan et al., 1998) and share 66% identity in primary sequence (Supplementary Fig. S2). Each TPST cDNA encodes a sequence with type II transmembrane domain, continues with a short N-terminal cytoplasmic domain and a luminal catalytic domain (Moore et al., 2003). Moreover, each has six conserved luminal cysteine residues and two *N*-glycosylation sites (Mishiro et al., 2006). In addition, it is proposed that only one TPST gene in *Drosophila melanogaster* (*Dm*TPST) via genomic analysis (Moore, 2003).

Tyrosine-sulfated proteins play important roles in many physiological and pathological processes including hormonal regulation hemostasis, inflammation and infectious diseases (Moore, 2003; Kehoe et al., 2000). For the majority of these proteins, the specific function of protein tyrosine sulfation is not well understood (Moore, 2003). Tyrosine sulfation has been implicated in intracellular trafficking (Friederich et al., 1988) and proteolytic processing (Bundgaard et al., 1995) of certain secreted proteins. Many studies indicated that tyrosine sulfation is a key modulator of extracellular protein-protein interactions (Moore, 2003; Kehoe et el., 2000). For example, tyrosine sulfation on the leukocyte adhesion molecule, P-selectin glycoprotein ligand-1 (PSGL-1), reinforces the binding affinity with P-selectin on

activated vascular endothelium (Wilkins et al., 1995). Also, it has been demonstrated that several sulfated tyrosine residues in the N-terminal domain of the chemokine receptor CCR5 is required for optimal binding of the chemokines RANTES, MIP-1a and MIP-1b (Farzan et al., 2000). Furthermore, tyrosine sulfation of CCR5, a major HIV-1 co-receptor, is also critical for its ability to interact with the HIV-1 envelope glycoprotein gp120 and mediates viral entry into host cells (Bannert et al., 2001). A missense mutation of a highly conserved region of the tyrosylprotein sulfotransferase 2 (*TPST-2*) gene in growth-retarded (*grt*) mouse resulted in an autosomal recessive, fetal-onset, severe thyroid hypoplasia-related TSH hyporesponsiveness. It has been found that TPST-2 has a high degree of substrate preference for TSH receptor (TSHR); however, no TPST-2 activity is detected in *grt*-mutated mice. Consequently it will lead to a loss-of-function on TSH-TSHR signal transduction pathway (Sasaki et al., 2007).

TPSTs have been purified from several mammalian tissues such as bovine adrenal medulla, rat liver, and human liver (Niehrs et al., 1990; Ramaprasad et al., 1998; Young et al., 1990). The recombinant TPSTs from mammalian cells, Chinese hamster ovary cell line have also been reported (Danan et al., 2010). We developed the prokaryotic expression system utilizing *E. coli* as host to purify TPST with high throughput, homogeneity, and confidence. The recombinant TPST has been reported

to couple with the PAPS generating system to produce the desired tyrosine sulfated proteins (Lu et al., unpublished). However, fusion protein-free *h*TPST was not acquired because of *h*TPST was digested to fragment by thrombin but not that for *Dm*TPST. In this study, a method using bovine thrombin protease to remove fusion protein was developed. Various biochemical and structural characterization of *Dm*TPST were investigated. Moreover, the hypothyroidism-related mutation (H266Q) in *m*TPST2 was studied using *Dm*TPST and *h*TPST2 as model.

MATERIALS AND METHODS

Materials

T4 DNA ligase, BamHI, XhoI restriction endonucleases were purchased from New England Biolabs (Ipswich, MA, USA). Oligonucleotide primers and peptides were individually synthesized by Mission Biotech Co., Ltd. (Taiwan) and Genemed Synthesis Inc.(San Antonio, USA). MES, Trizma base, sodium phosphate, NaCl, imidazole, glycerol, bovine thrombin, and pyrophosphatase were products of Sigma (St. Louis, MO, USA). Blue Dextran 2000, Albumin, Ovabumin, Chymotrypsinofen A, Ribonuclease A, Aldolase, HisTrap fastflow sepharose, Hitrap Q sepharose fastflow, and Sephacryl S-100HR were purchased from Pharmacia Biotech GE Healthcare (Uppsala, Sweden). Sodium [^{35}S]sulfate was purchased from PerkinElmer (Boston, MA, USA). Cellulose thin-layer chromatographic plates were obtained from Merck & Co., Inc. (Whitehouse Station, NJ, USA). All other reagents were the highest grade and commercially available.

Methods

Sequence alignment and transmembrane domain analyses - The sequence alignment was performed by ClustalW and sorted shading by BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html). The residue colored in red was

the predicted transmembrane domain calculated by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html).

Vector construction – All primers for site directed mutagenesis: sense (5' - ggtgctgcaccatgaggagttca - 3') and antisense (5'- TGAACTCCTCCTGGTGCAGC ACC - 3'). The *Dm*TPST and *Dm*TPST (H269Q) were subcloned into pET43a expression vectors. The cDNA of potential cytosolic domain (29-377) of *Dm*TPST predicted above was amplified by PCR through specific primers designed to contain XhoI restriction site (5'-tgaagaattcgacgccgccaacgagctctcctc -3') in the sense and the antisense one consisted of EcoRI restriction site (5'-tgccctcgagctctcccacagcattcgattggc -3'). cDNA fragment was inserted into the EcoRI/XhoI doubly-restriction sites and then confirmed using ABI Prism 377 DNA sequencer (Applied Biosystems, Foster City, CA) following the standard protocol.

Expression and purification – A single colony of BL21 (DE3) consisted of pET-43a plasmid with *Dm*TPST cDNA was cultured in LB medium containing ampicillin at 37 ^{o}C. The 1 mM ITPG was added to induce *Dm*TPST expression while the bacterium reached mid-long growth (A_{600} 0.8-1.0) for 16 hr at 20 ^{o}C in a shaking incubator. The cells were harvested by centrifugation at 13400 g for 30 min at 4°C and the pellet was disrupted by sonication in IMAC5 buffer (50mM Tris-HCl at pH 8.0,

500mM NaCl, 5mM imidazole, and 10% glycerol). The Ni-NTA sepharose was used to purify NusA-*Dm*TPST and further digested by bovine thrombin for 3 hours at 4 ^{o}C to get rid of NusA. NusA-free *Dm*TPST was purified by Hitrap Q sepharose, and the protein purity was determined by SDS-polyacrylamide gel electrophoresis.

In-gel digestion and identification by MALDI-TOF - The spots of interest were excised and digested in gel with trypsin according to standard procedure (Shevchenko et al., 1996). The digested samples were analyzed by MALDI-TOF, and the results were analyzed by Mascot software using NCBI and Swissprot as databases.

Stability assay and optimal preservation – The *Dm*TPST was treated with NaCl gradient (0mM, 50mM, 100mM, 150mM, 200mM) for 2hrs at 4°C , using SDS-PAGE electrophoresis pattern to indicate the result. To reveal the relationship between quaternary structure and salt stability, the *Dm*TPST was treated with 0mMNaCl and 500mM NaCl for 2hr, was determined by Sephacryl S-100 HR. The thermal stability was determined by analyzed SDS-PAGE electrophoresis pattern, *Dm*TPST was incubated at 20°C and 4°C for 3 days in optimal condition (50mM Tris, 150mM NaCl, 10%glycerol).

Gel filtration – The quaternary structure of *Dm*TPST, peak1 of *Dm*TPST, peak2 of *Dm*TPST were analyzed by monitoring the Sephacryl S-100 HR elution pattern and

apparent elution volume (V_e), which was used to confirm the estimated molecular weight of *Dm*TPST. The molecular weight marker was used Blue Dextran 2000 as V0 and the calibration was consisted of albumin (67 kDa), ovabumin (43 kDa), chymotrypsinofen A (25 kDa), and ribonuclease A (13.7 kDa)]. Total amount of 1 mg *Dm*TPST in the buffer of 50 mM Tris-HCl at pH 8.0, 150 mM NaCl, and 10% glycerol was injected into Sephacryl S-100 HR at a rate of 1 ml per min. The Sephacryl S-100 HR separated *Dm*TPST peak1 and *Dm*TPST peak2 was re-injected into Sephacryl S-100 HR to determine the relationship between two peaks. Furthermore, the *Dm*TPST was completely separated by Sephacryl S-100 HR-Sephacryl S-200 HR joining.

EDC crosslinking – 6μg *Dm*TPST was treated with 50mM EDC on total volume 20μl. at various temperatures for 2, 5 and 6 hours. The SDS-PAGE electrophoresis was used to analyze the result.

Activity assay - The recombinant TPST activity was determined using the radiation of [^{35}S]PAPS as donor and transferred the sulfate group to substrate, PSGL-1.The coupled-enzyme (*h*PAPSS-1 and TPST) radioactive assay was newly established for the measurement of TPST activity in our lab (Liu et al., unpublished). The standard assay was composed of 50 mM MES at pH 6.5, 5 mM

beta-mercaptoethanol, 4 mM inorganic [^{35}S]SO_4^{2-}, 1 mM $MgCl_2$, 1 mM ATP, 120μM PSGL-1 peptide (ATEYEYLDYDFL), 1 μg recombinant *h*PAPSS-1, 1 unit (unit = mole product/min) pyrophosphatease and then incubated for 15 minutes at 37°C to generate saturated [^{35}S]PAPS. After this pre-incubation, purified TPST was added to initiate the reaction of protein tyrosine sulfation for 45 minutes at 37°C in a final volume of 20 μl. The reactions were terminated by heating at 95°C for 2 minutes. The supernatant was collected and analyzed by spotting 2 μl aliquot of the reaction mixture onto a cellulose thin-layer chromatographic (TLC) plate and developed with n-butanol/pyridine/formic acid/water (5:4:1:3; by volume) as the solvent system. The dried plate was exposed with Kodak BioMax MR film which provided the optimal resolution for ^{35}S autoradiography. For the enzyme kinetic assay, the concentration of PSGL-1 varied from 0.16 to 120 μM. Results of kinetic experiments were analyzed using nonlinear regression to fit the appropriate equation to the data. Kinetic data obtained from non-inhibitory experiments were individually fit to Michaelis-Menten Equation 1 (Cornish-Bowden, 1995). The rate constants (K_m and V_{max}) were obtained using SigmaPlot 2001, V7.0 and Enzyme Kinetics Module, V1.1 (SPSS Inc., Chicago, IL). Data used represent mean values derived from three determinations.

$$v = V[S]/(K_m + [S]) \quad (1)$$

RESULT

High sequence alignment and homology of TPST– Through the use of the bioinformatic tool, we determined that *h*TPST2 and *Dm*TPST shared ~56 % identity of sequences with a similar length (**Fig. 2**). The identity of sequences between *h*TPST2 and *m*TPST-2 was too high to reach 96 %, thus *h*TPST2 replaced *m*TPST-2 in the hypothyroidism-related mutation experiment.

Cloning, expression and purification of TPST – *h*TPST2 and *Dm*TPST cDNA were subcloned to *pET-43a (+)* expression vectors to render the maximal soluble protein (Fig. 3a and Fig. 4 and **Fig. 5**). Although the fusion protein, NusA, facilitated TPST protein folding, excessive residual buildup (60 kDa) might have influenced the *Dm*TPST catalysis (**Fig. 3b**). The NusA-TPST was divided into NusA and TPST owing to the thrombin digestive site between the two proteins (lane 3 in **Fig. 5**). Only *Dm*TPST was suitable for acquiring *fusion free* enzymes, because *h*TPST2 had been digested by thrombin protease, but it was unsuitable for *h*TPST-2 (**Fig. 4** and **Fig. 5**). NusA and *Dm*TPST were separated by Hitrap Q sepharose according to differences in anion ion exchange (lane 4 in **Fig. 4**). As the purification table of *Dm*TPST (**Table 1**) shows, there was approximately 10 % recovery and 0.6 mg homogeneous *Dm*TPST in a single batch (2.5 liter LB broth cultivation). The purified *Dm*TPST was identified by

trypsin in-gel digestion following MS analysis (**Fig. 6**).

Various stability assays – *Dm*TPST was found to be stable in the optimal buffer containing 50 mM Tris-base at pH 8, 200 mM NaCl, 10 % glycerol (**Fig. 7a**). Low salt concentration and absence of glycerol, contributed to the instability of *Dm*TPST (**Fig. 6a**). In addition, *Dm*TPST was even more unstable at 20 °C than 4 °C, with the incubation (**Fig. 6b**). The protease inhibitor cocktail prevented the degradation of the protein from uncertain proteolytic digestion (**Fig. 6b**).

Analysis of structure – Figure 9a shows two forms of *Dm*TPST in sephacryl S-100 HR elution pattern, with both identified as *Dm*TPST according to SDS-PAGE electrophoresis (**Fig. 9b**). Under non-reducing, two forms of *Dm*TPST had equal migration of electrophoresis. The corresponding molecular weight of standard proteins related to elution volume is shown in **Fig. 10** and **TABLE 3**, and the calculated values of *Dm*TPST were similar to the theoretical molecular weight of the dimer and monomer of *Dm*TPST (**TABLE 4**). However, the EDC treated TPST revealed that only the monomer form of *Dm*TPST existed, because the EDC treated TPST showed only one band apparent molecular weight of TPST-36kDa (**Fig. 11**). Furthermore, the sephacryl S-100 HR- sephacryl S-200 HR were used to separate two structures of *Dm*TPST (**Fig. 12** and **Fig. 13**), and two structures of *Dm*TPST were not in equilibrium with each other, as shown **in Fig. 14** shown. Although sephacryl S-100

HR- sephacryl S-200 HR had better separation than sephacryl S-100 HR, the length of the two forms of TPST were wider than used. Neither forms of TPST had an effect under high concentrations of sodium chloride; but degradation was caused by low sodium chloride (**Fig. 15)**.

TPST activity assay – Our lab developed a coupled enzyme assay to generate freshly saturated [^{35}S] PAPS, which we applied to TPST catalysis (**Fig. S3**). The catalytic efficiency of *h*PAPSS-1 in generating PAPS was a great deal higher than the TPST exhausted PAPS (Liu et al., unpublished). The sulfated peptide was separated by thin-layer chromatography (TLC) and probed via liquid scintillation analyzer. In a standard assay, 0.12 M P-selectin glycoprotein ligand-1 (PSGL1) peptide was used as the substrate for 1.5 μg TPST, shown in **Fig. 16**. For the kinetics assay the K_m and k_{cat} values for NusA-free *Dm*TPST were 42.1 μM and 0.32 min^{-1}, respectively (**TABLE 5**). In addition, the catalytic efficiency (k_{cat}/K_m) of NusA-free *Dm*TPST was similar to that with NusA-TPST (Wang et al., unpublished).

Hypothyroidism-related point mutation of TPST – Previous studies showed that the H267Q mutation of mouse TPST-2 resulted in dysfunctional enzymatic activity (Sasaki et al., 2007). The detailed mechanism, however, remains unclear. Multiple sequence alignment revealed that the H267Q of mouse TPST-2 was highly

conserved among various species (Sasaki et al., 2007). The H269Q of *Dm*TPST and H267Q in human TPST-2 was examined (**Fig. 17).** The expression profile revealed that the protein expression of *Dm*TPST-H269Q and *h*TPST2-H267Q provided very little soluble protein, (**TABLE 6)**. However, the specific activity of purified *h*TPST2-H267Q and *Dm*TPST-H269Q was slightly less than wild type TPST, with few soluble enzymes (**TABLE 6**). This indicated that the hypothyroidism-related point mutation had led to an expression of less TPST, but the activity of TPST remained.

DISCUSSION

It was found that only one TPST gene in *Drosophila melanogaster* (*Dm*TPST) via genomic analysis (Moore, 2003) was found. *Dm*TPST sequence identity compared to *h*TPST1 and *h*TPST2 was 57 % and 61 % respectively (Fig. 2); moreover, approximately 75 % of known human genes associated with disease have a recognizable match to the genetic code of fruit flies, and 50 % of fly protein sequences have mammalian analogues. *Drosophila melanogaster* is one of the most studied organisms in biological research, particularly in genetics and developmental biology.

In previous studies, TPST was purified from mammalian tissue, or purified recombinant TPST from mammalian cells. However, the amount of TPST purified from eukaryote was too limited for the molecular based characterization to be studied. This was the cause of the difficulty in performing further research. Our lab developed an expression system utilizing *E. coli* as a host in the purification of hTPST-2, optimized to provide high throughput, homogeneity, and confidence (Lu et. al., unpublished). However, fusion protein free *h*TPST could not be acquired from fusion protein – *h*TPST, because the TPST was fragmented due to thrombin digestion. In contrast, *Dm*TPST was much more tolerant of thrombin digestion than *h*TPST. Thus, according to the results of Hitrap Q sepharose separation, fusion protein free *Dm*TPST

was homogenously purified through thrombin digestion and Hitrap Q sepharose separation. In addition, the net charge of *Dm*TPST was weakly negative.

The fact that NusA-free *Dm*TPST was found to be increasingly unstable under low salt concentration with an absence of glycerol might be attributable to the disequilibrium of the salt bridges among *Dm*TPST, which further destructs the architecture of peptide formation. Previous research indicated that zebrafish TPST activity decreased at temperatures higher than 37 ^{o}C (Emi et al., 2004). Our research also indicated that a protease caused TPST to degrade into fragments. The protease digestion level of TPST at 20 ^{o}C was more serious than it was at 4°C.

Gel filtration analysis revealed that the truncated transmembrane *Dm*TPST had two structures; perhaps TPST with two sharps or two quaternary structures. The non-reducing SDS-phage indicated that the covalent bond was not the reason for the structures, and this led to the formation of two structures of *Dm*TPST. In addition, no dimer quaternary structure of TPST was found with EDC formed amide covalent bond between the two enzymes. Dynamic equilibrium was not observed between the two structures, and sodium chloride had no effect on either of them. As a result, we propose that *Dm*TPST may have only monomer quaternary structures with two sharps, and of course, dynamic equilibrium between the two structures was not observed either salt effect. The function and characterization of the two sharps of the *Dm*TPST

were the next issues we had to deal with.

Hypothyroidism-related point mutation nearly leads to a different expression profile on *Dm*TPST-H269Q that much less level expression on either supernatant or pellet of *E. coli* cultivation. The mutation codon, CAG, is used extensively in *E. coli* systems; therefore, the factor of codon usage in the expression system could be excluded. The protein expression level of *h*TPST2-H267Q showed a similar situation to that with a very less expression of protein. The enzyme activity of H267Q of *h*TPST2 however, showed no apparent difference from wild type, as shown in Table 3. Previous research indicated that the enzyme activity of *m*TPST2-H266Q had been eliminated (Sasaki et al., 2007). The identity of the protein sequence between human and mouse TPST-2 was 96 %, which might contribute to the sharp conflict. Although the less protein expression of the H266Q of *m*TPST2, might be attributed to detection of *m*TPST2 by western blot in this study (Sasaki, et al. 2007). Also, the expression system of *m*TPST2-H266Q was cell culture and that differed from ours in this research. The activity of *m*TPST2-H267Q was examined amidst contamination from cell lysate; therefore, the results might be questionable and insensitive for the detection of TPST activity.

So far, the structure of TPST2 is still unavailable, and difficult to computationally

model. We use the MODELLER server to model the TPST2 with the 1xv1, the human sulfotransferase SULT1B1, as the template. We were able to roughly understand the relative structure and regulatory residue from this modeled structure. According to this modeled structure, the H266Q was localized at the surface of TPST and excluded from the active packed site as shown in the modeling of mTPST2 (**Fig. S2**). Up to this point, the possible reason that a mutation could lead to lower expression levels, might affect the structural stability. As a result, we proposed a pathway of hypothyroidism-related point mutation. A mutation on TPST may lead to structural instability of TPST, thus TPST would be fragmented or unable to folded TPST. This could lead to a lack of TPST expression decreasing sulfation on TPST substrate-TSHR. Sulfation less TSHR has weaker interaction with TSH, thus blocking the downstream signal transduction (**Fig. 18**).

In this study, we were the first to purify and identify *Dm*TPST with enzymatic activity. The compound stability of *Dm*TSPT was both examined. The structure of *Dm*TPST *in vitro* may also be determined by the existence of two sharp structures. A hypothyroidism-related mutation, H269Q in TPST was not competent to be translationally expressed, which led to a loss of activity. Further study to uncover and characterize the two sharp structures and the role of His-269 on mechanism will be investigated in detail, in future studies.

REFERENCES

1. Anastasi, A., Bertaccini, G., and Erspamer, V. (1966) Pharmacological data on phyllokinin (bradykinyl-isoleucyl-tyrosine *O*-sulphate) and bradykinyl-isoleucyl-tyrosine. *Br. J. Pharmacol. Chemother.* **27:**479-485

2. Anastasi, A., Erspamer, V., and Endean, R. (1968) Isolation and amino acid sequence of caerulein, the active decapeptide of the skin of hyla caerulea. *Arch. Biochem. Biophys.* **125:**57-68

3. Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., Choe, H., Sodroski, J. (2001) CC and CX3C Chemokines Differentially Interact with the N Terminus of the Human Cytomegalovirus-encoded US28 Receptor. *J. Exp. Med.* **194:** 1661–1673

4. Bettelheim, F. R. (1954) Tyrosine-*O*-sulfate in a peptide from fibrinogen. *J. Am. Chem. Soc.* **76:**2838-2839.

5. Beisswanger, R., Corbeil, D., Vannier, C., Thiele, C., Dohrmann, U., Kellner, R., Ashman, K., Niehrs, C., and Huttner, W. B. (1998) Existence of distinct tyrosylprotein sulfotransferase genes: molecular characterization of tyrosylprotein sulfotransferase-2. *Proc. Natl. Acad. Sci. U. S. A.* **95:**11134-11139.

6. Bundgaard, J. R., Vuust, J., Rehfeld, J. F. (1995) Post-translational Modifications

of Proteins: Tools for Functional Proteomics. *EMBO. J.* **14:** 3073–3079

7. Chapman, E., Best, M. D., Hanson, S. R., and Wong, C. H. (2004) Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed. Engl.* **43:**3526-3548

8. Cornish, B. A. (1995) Analysis of Enzyme Kinetic Data pp 118–122, *Oxford University Press, Oxford.*

9. Cormier, E. G., Persuh, M., Thompson, D. A., Lin, S. W., Sakmar, T. P., Olson, W. C., Dragic, T. (2000) An Aptamer That Neutralizes R5 Strains of Human Immunodeficiency Virus Type 1 Blocks gp120-CCR5 Interaction. *Proc. Natl. Acad. Sci. USA* **97:** 5762–5767

10. Danan, L.M., Yu, Z., Ludden, P.J., Jia, W., Moore, K.L., Leary, J.A. (2010) Catalytic Mechanism of Golgi-Resident Human Tyrosylprotein Sulfotransferase-2: A Mass Spectrometry Approach. *J Am Soc Mass Spectrom.* **9:**1633-1642

11. Farzan, M., Vasilieva, N., Schnitzler, C. E., Chung, S., Robinson, J., Gerard, N. P., Gerard, C., Choe, H., Sodroski, J. (2000) A Tyrosine-sulfated Peptide Derived from the Heavy-chain CDR3 Region of an HIV-1-neutralizing Antibody Binds gp120 and Inhibits HIV-1 Infection. *J. Biol. Chem.* **275:** 33516–33521

12. Friederich, E., Fritz, H. J., Huttner, W. B. (1988) Tyrosylprotein sulfotransferase

in rat submandibular salivary glands. *J. Cell Biol*. **107:** 1655–1667

13. Gerard, N. P., Gerard, C., Sodroski, J., Choe, H. (1999) A Tyrosine-sulfated Peptide Derived from the Heavy-chain CDR3 Region of an HIV-1-neutralizing Antibody Binds gp120 and Inhibits HIV-1 Infection. *Cell* **96:** 667–676

14. Gerard, N. P., Gerard, C., Sodroski, J., Choe, H. (1999) A Highly Conserved Arginine in gp120 Governs HIV-1 Binding to Both Syndecans and CCR5 via Sulfated Motifs. *Cell* **96:** 667–676

15. Gregory, H., Hardy, P. M., Jones, D. S., Kenner, G. W., and Sheppard, R. C. (1964) The antral hormone gastrin. Structure of gastrin. *Nature* **204:**931-933

16. Kansas, G. S. (1996) Selectins and their ligands: current concepts and controversies. *Blood* **88:**3259-3287

17. Kehoe, J. W., Bertozzi, C. R. (2000) Got Sulfate? Luring Axons This Way and That. *Chem Biol* **7:** R57–R61

18. Komori, R., Amano, Y., Ogawa-Ohnishi, M., and Matsubayashi, Y. (2009) Identification of tyrosylprotein sulfotransferase in Arabidopsis. *Proc. Natl. Acad. Sci*. **106:**15067-15072

19. Lee, R. W., and Huttner, W. B. (1983) Tyrosine-O-sulfated proteins of PC12 pheochromocytoma cells and their sulfation by a tyrosylprotein sulfotransferase. *J. Biol. Chem.* **258:**11326-11334

20. Liu, C.C., and Schultz, P. G. (2006) Recombinant expression of selectively sulfated proteins in Escherichia coli. *Nat Biotechnol.* **24:**1436-1440

21. Liu, C.C., Cellitti, S.E., Geierstanger, B.H., and Schultz, P.G. (2009) Efficient expression of tyrosine-sulfated proteins in E. coli using an expanded genetic code. *Nat. Protoc.* **4:**1784-1789.

22. Lu, L.Y., Hsieh, Y.C., Liu, M.Y., Lin, Y.H., Chen, C.J., Yang, Y.S. (2002) Identification and Characterization of Two Amino Acids Critical for the Substrate Inhibition of Human Dehydroepiandrosterone Sulfotransferase (SULT2A1). *J. Biol. Chem.* **41:**12959-12966

23. Mishiro, E., Liu, M.Y., Sakakibara, Y., Suiko, M., Liu M.C. (2004) Zebrafish tyrosylprotein sulfotransferase: molecular cloning, expression, and functional characterization. *Biochem Cell Bio.* **82(2):** 295-303.

24. Mishiro, E., Sakakibara, Y., Liu, M. C., and Suiko, M. (2006) Differential enzymatic characteristics and tissue-specific expression of human TPST-1 and TPST-2. *J. Biochem. (Tokyo)* **140:**731-737

25. Moore, K. L. (2003) The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **278:**24243-24246

26. Mutt, V., and Jorpes, J. E. (1968) Structure of porcine cholecystokinin-pancreozymin. Cleavage with thrombin and with trypsin. *Eur. J. Biochem.* **6:**156-162

27. Niehrs, C., and Huttner, W. B. (1990) Purification and characterization of tyrosylprotein sulfotransferase. *EMBO J.* **9:**35-42

28. Ouyang, Y. B., and Moore, K. L. (1998) Molecular cloning and expression of human and mouse tyrosylprotein sulfotransferase-2 and a tyrosylprotein sulfotransferase homologue in Caenorhabditis elegans. *J. Biol. Chem.* **273:**24770-24774

29. Ramaprasad, P., and Kasinathan, C. (1998) Isolation of tyrosylprotein sulfotransferase from rat liver. *Gen. Pharmac.* **30:**555-559

30. Sasaki, N., Hosoda, Y., Nagata, A., Ding, M., Cheng, J.M., Miyamoto, T., Okano, S., Asano, A., Miyoshi, I., Agui, T. (2007) A mutation in Tpst2 encoding tyrosylprotein sulfotransferase causes dwarfism associated with hypothyroidism. Mol Endocrinol. **21(7):**1713-21.

31. Ueoka, C., Kaneda, N., Okazaki, I., Nadanaka, S., Muramatsu, T., and Sugahara,

K. (2000) Neuronal cell adhesion, mediated by the heparin-binding neuroregulatory factor midkine, is specifically inhibited by chondroitin sulfate E. Structural ans functional implications of the over-sulfated chondroitin sulfate. *J. Biol. Chem.* **275:**37407-37413

32. Wang, L., Fuster, M., Sriramarao, P., and Esko, J. D. (2005) Endothelial heparan sulfate deficiency impairs L-selectin- and chemokine-mediated neutrophil trafficking during inflammatory responses. *Nat. Immunol.* **6:**902-910

33. Wilkins, P. P., Moore, K. L., McEver, R. P., Cummings, R. D. (1995) Detection and Purification of Tyrosine-sulfated Proteins Using a Novel Anti-sulfotyrosine Monoclonal Antibody. J. Biol. Chem. **270:** 22677–22680

34. Young, W. F. Jr. (1990) Human liver tyrosylsulfotransferase. *Gastroenterology.* **99:**1072-1078

FIGURES

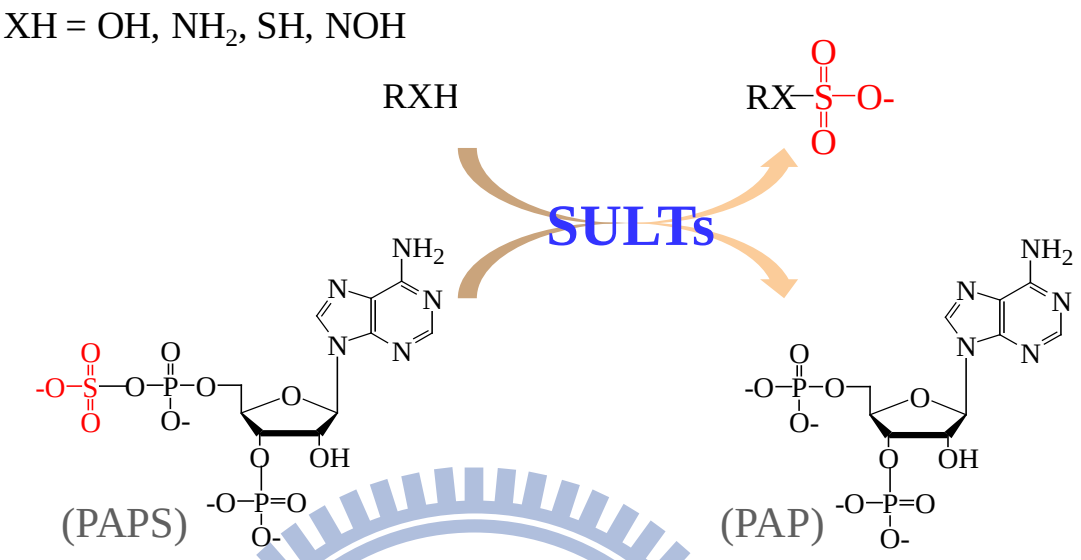


Figure 1. Sulfotransferases catalyzation mechanisam. Sulfotransferases catalyze the transfer of a sulfuryl group (SO_3^-) from a donor molecule, usually 3'-phosphoadenosine 5'-phosphosulfate (PAPS), to a variety of amine and hydroxyl substrates as nucleophiles, resulting in the formation of a substrates O_4-sulfate ester and 3', 5'-ADP (PAP).

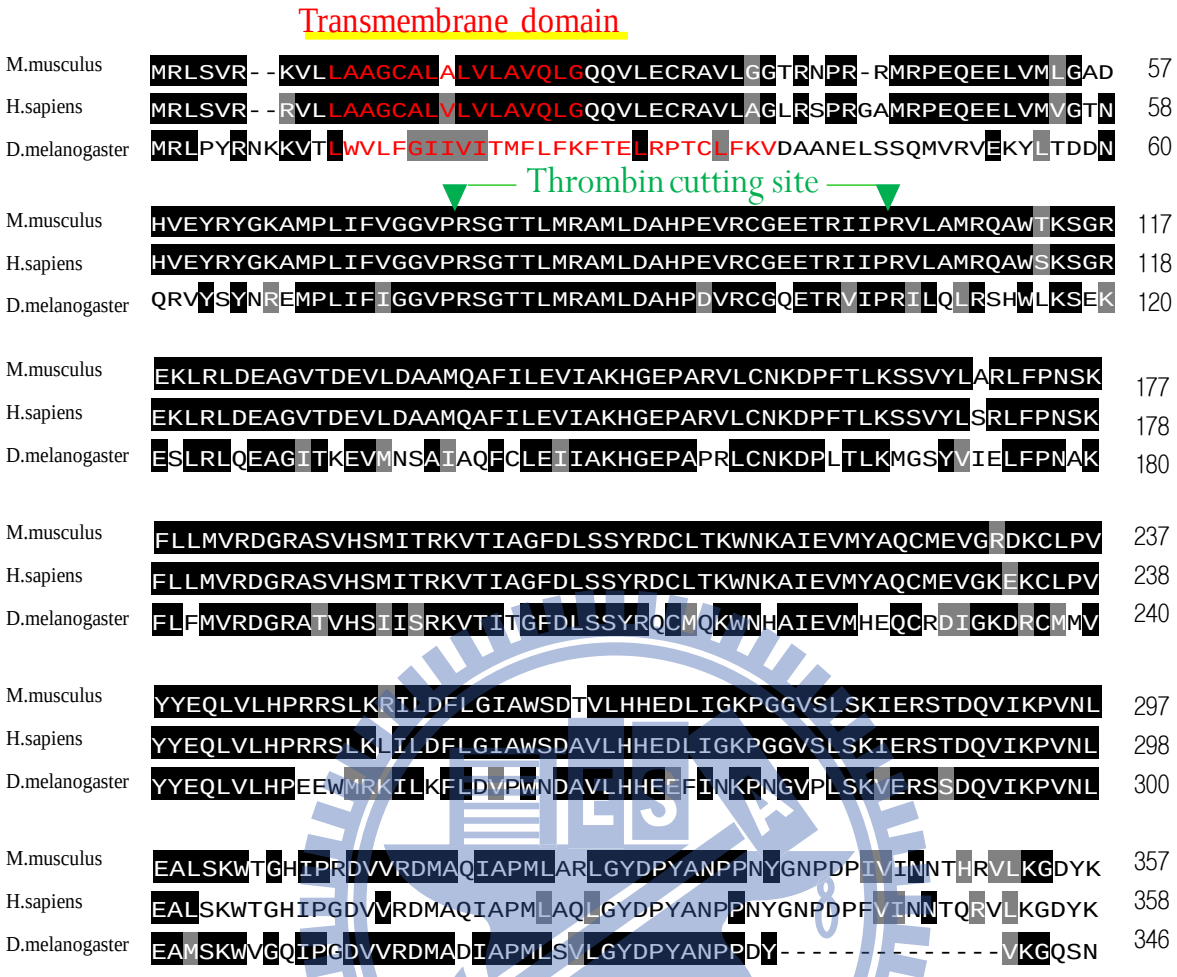
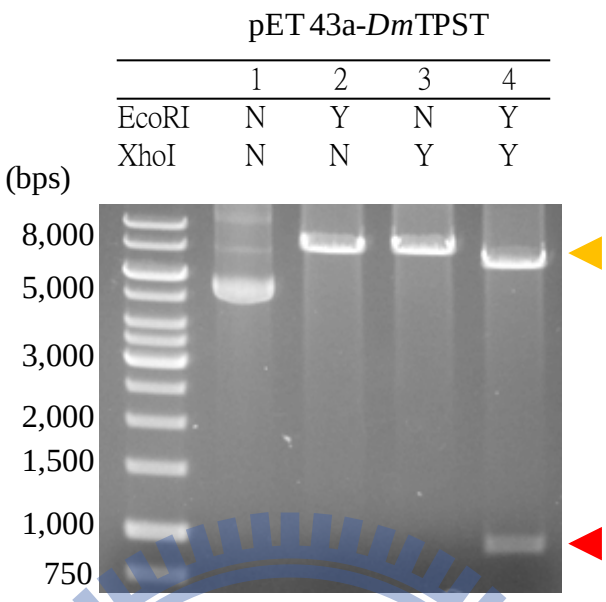


Figure 2. Multiple alignment of amino acid sequences of TPSTs. Sequence alignment and transmembrane domain analysis of *human* TPST-2 (*h*TPST2), *Mouse musculus* (*m*TPST2), and *Drosophila melanogaster* TPST (*Dm*TPST). The sequence alignment was performed by ClustalW and sorted shading by BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html), the sequence identity of *m*TPST2 to *h*TPST2 and *Dm*TPST are 94% and 54%, individually. The black background indicated identity to each other and the gray one meant conserved

substitutions. The residue colored in red was the predicted transmembrane domain calculated by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html). The thrombin autolytic cleavage site was identified by FindPept and indicated as the green triangle (http://au.expasy.org/tools/findpept.html).

(a)



(b)

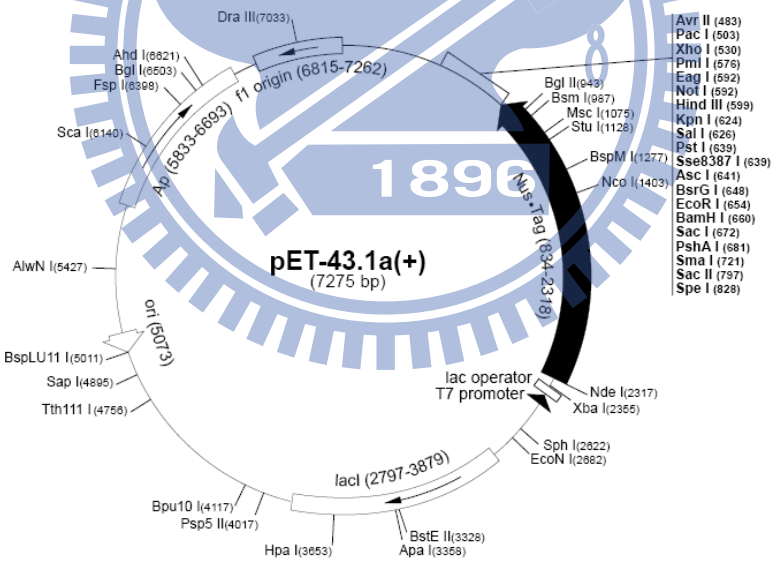


Figure 3. *DmTPST* Clone to plasmid-*pET 43a* (a) pET43a-*DmTPST* along, with EcoRI, XhoI, EcoRI+XhoI, respectively were by agarose gel electrophoresis (1%). (b) Vector *pET43a*, *Dm*TPST constructed between EcoRI and xhoI restriction site. This figure was acquired from Merck company.

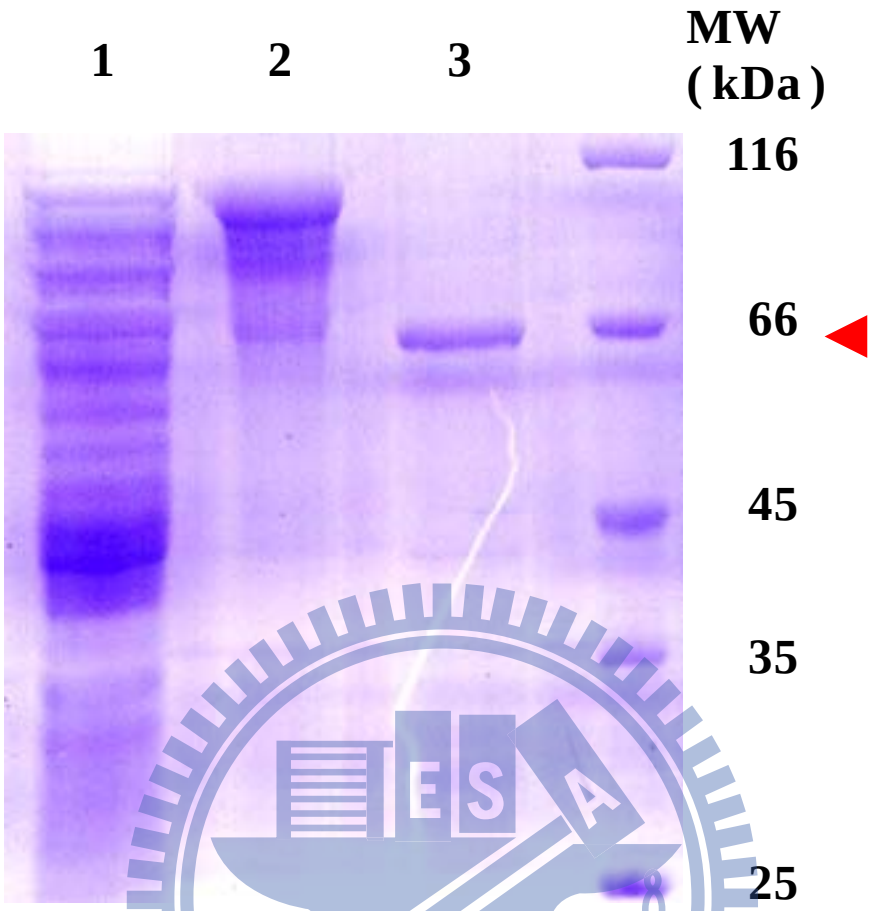


Figure 4. ***h*TPST2 purification.** The purification process of *h*TPST2 was as follows, 1.crude extract of TPST-cultivated host bacteria; 2.Ni-NTA column purification from crude extract; 3.NusA-*h*TPST2 digested with thrombin. The red triangle indicated fusion protein NusA, which the molecular weight is approximate 60kDa.

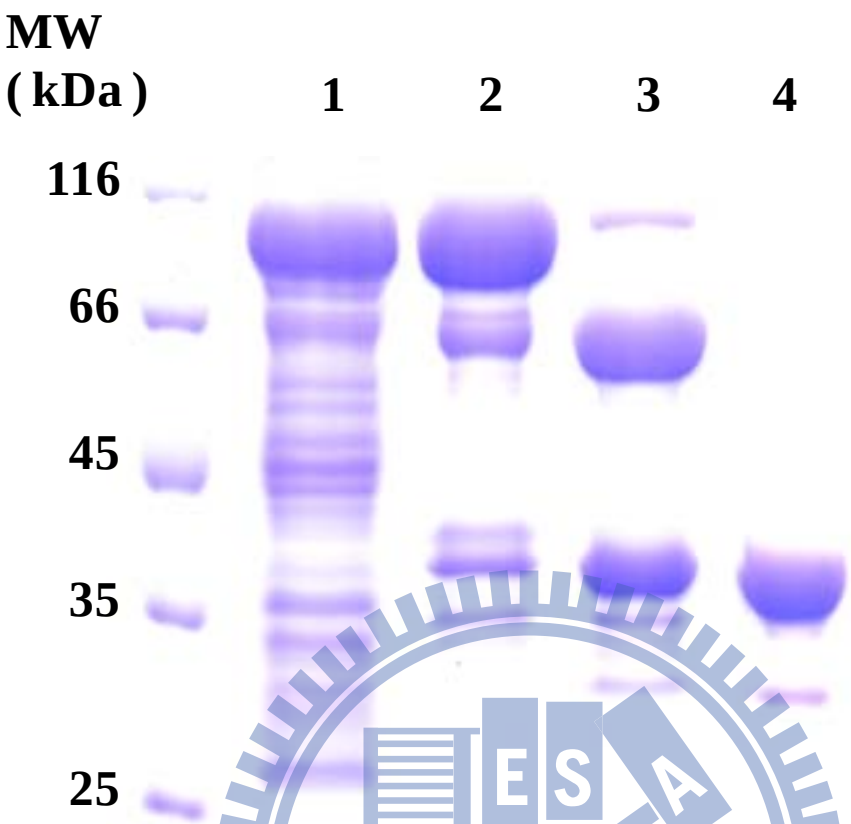
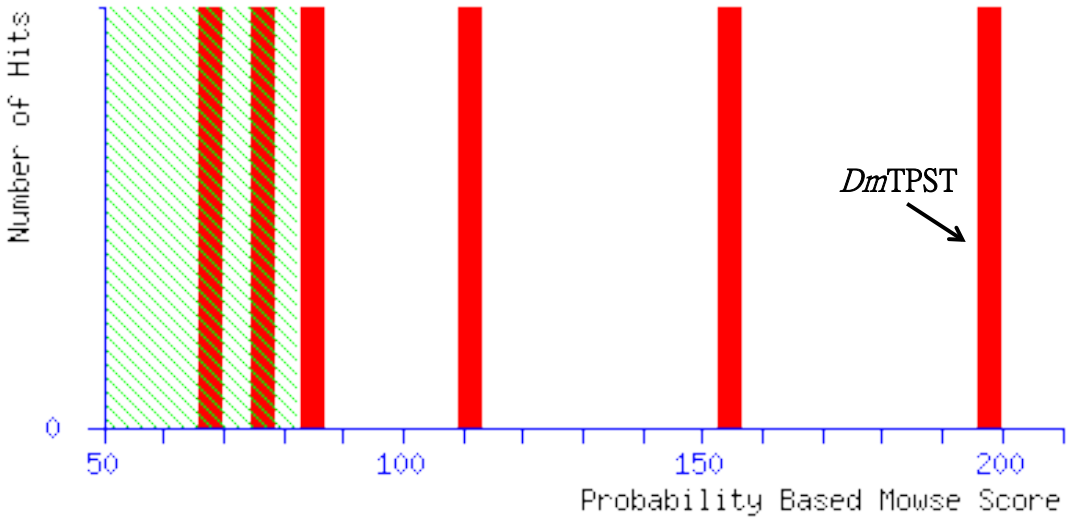


Figure 5. *Dm*TPST purification. Overall purification process of NusA-free *Dm*TPST was as follows, 1. crude extract of TPST-cultivated host bacteria; 2. Ni-NTA column purification from crude extract; 3. NusA-*Dm*TPST digested with thrombin; 4. purified NusA-free *Dm*TPST by HiTrap Q column.



1 MRLPYRNKKV TLWVLFGIIV ITMFLFKFTE LRPTCLFKVD AANELSSQMV

51 RVEK**YLTDDN QR**VYSYNR**EM PLIFIGGVPR** SGTTLMR**AML DAHPDVR**CGQ

101 ETRVIPRILQ LRSHWLKSEK **ESLRLQEAGI TKEVMNSAIA QFCLEIIAK**H

151 GEPAPR**LCNK DPLTLKMGSY VIELFPNAK**F LFMVRDGR**AT VHSIISRKVT**

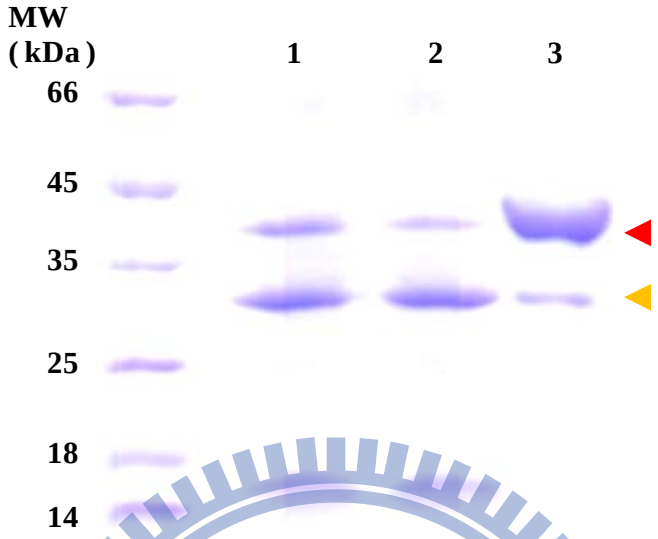
201 **ITGFDLSSYR** QCMQK**WNHAI EVMHEQCR**DI GKDRCMMVYY EQLVLHPEEW

251 MRKILKFLDV PWNDAVLHHE EFINKPNGVP LSKVERSSDQ VIKPVNLEAM

301 SK**WVGQIPGD VVR**DMADIAP MLSVLGYDPY ANPPDYVKGQ SNAVGE

Figure 6. The identification of *Dm*TPST. The score of *Dm*TPST was 198, and was shown as the arrow indicated. The red and bold typefaces were the peptides fingerprinted in the assay. Only indicated the identity or extensive homology ($p<0.05$), the vertical bars outside the shaded green region in the histogram.

(a)



(b)

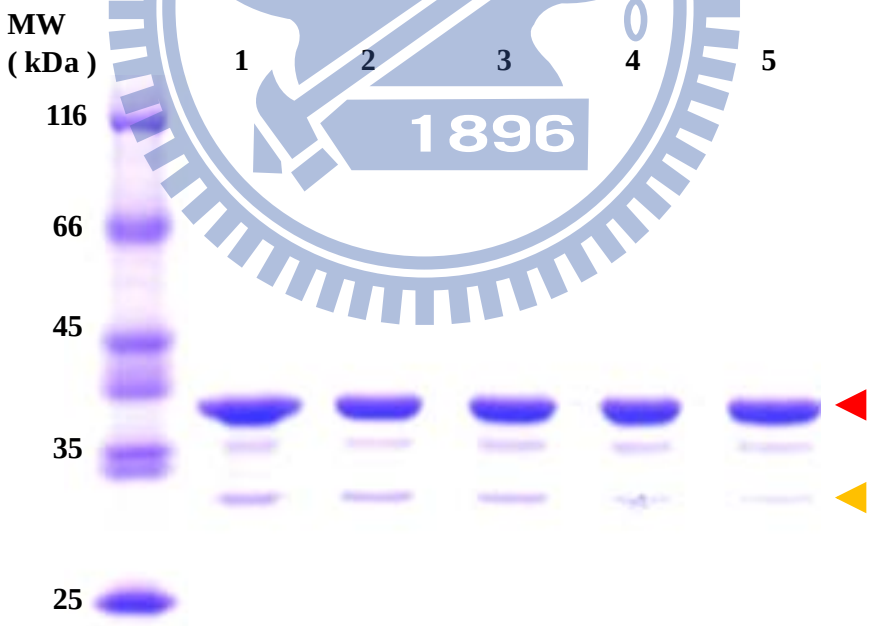


Figure 7. *Dm*TPST stability and optimal preserved buffer. (a) *Dm*TPST (3.5ug) was incubation with different buffer (1. 50mM tris, 150mM NaCl, 2. 50mM tris, 10%glycerol, 3. 50mM tris, 150mM NaCl, 10%glycerol) for 1hr at 4 ^{o}C. (b) *Dm*TPST (2μg) was treated with NaCl gradient (1. 0mM, 2. 50mM, 3. 100mM, 4. 150mM, 5. 200mM) for 2hrs at 4°C , using SDS-PAGE electrophoresis pattern to indicate all results. The red and orange triangle indicated *Dm*TPST and *Dm*TPST fragment, individually.

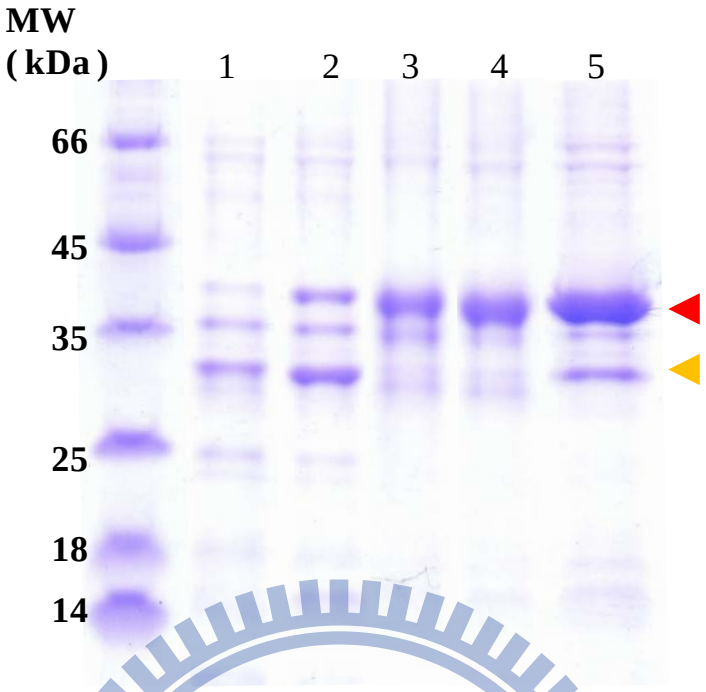
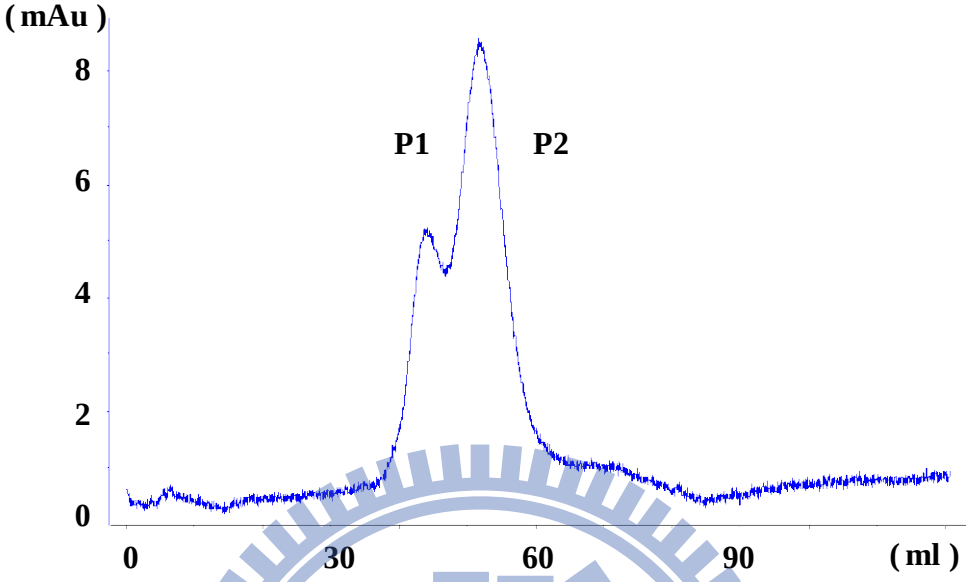
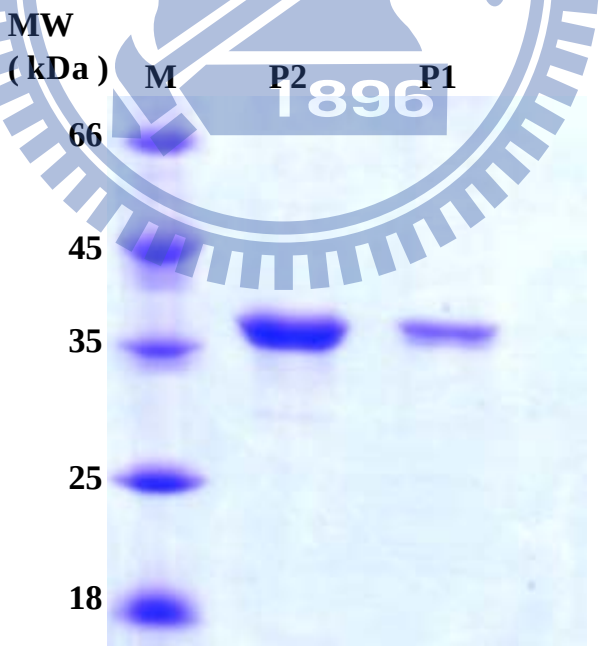


Figure 8. Thermal effect study of *Dm*TPST. *Dm*TPST (3.5μg) was incubated at two temperatures (1 and 3 at 20°C; 2 and 4 at 4°C) for 3 days in the optimal buffer, also number 3 and 4 was treated with protease inhibitor cocktail. Number 5 was control that had no any treatment for the enzyme. The red and orange triangle indicated *Dm*TPST and *Dm*TPST fragment, individually.

(a)



(b)



(c)

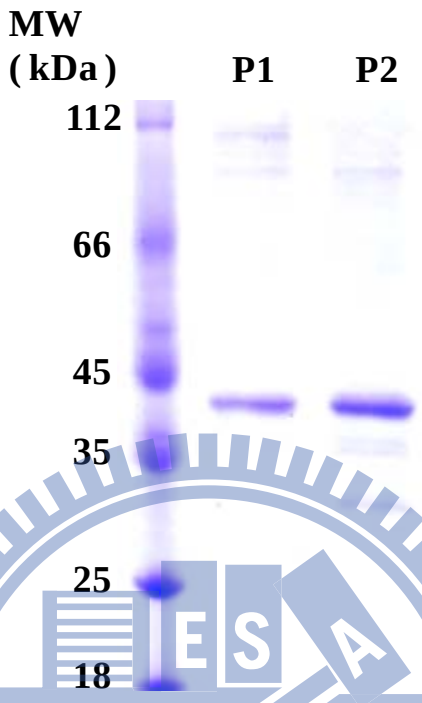


Figure 9. The quaternary structure of *Dm*TPST. The quaternary structure of *Dm*TPST (1mg) was determined by sephacryl S-100 HR in an optimal buffer, as the FPLC elution pattern showed that (a) two peaks (P1 and P2) were eluted at total volume 120ml. (b) And peak1 and peak2 were identical component- *Dm*TPST, was indicated by reducing SDS-PAGE. (c) Also, peak1 and peak2 were analyzed by non-reducing SDS-PAGE.

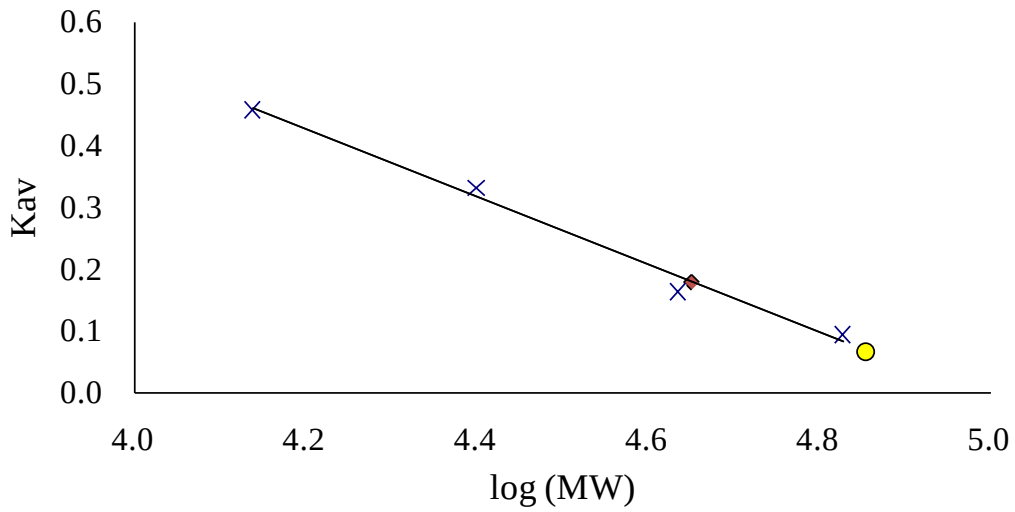


Figure 10. The standard proteins for gel filtration analysis. A Calibration curve which defined the relationship between the logarithm of their respective molecular weights and the elution volumes of a set standard, was determined by sephacryl S-100 HR. Using blue dextran as Vo, Ribonuclease A(15.6 KDa), Chymotrypsinofen A(19.4 KDa), Ovabumin(47.6 KDa), Albumin(62.9 KDa),were as calibration standard. Two peaks of *Dm*TPST were circle and diamond sharps, respectively.

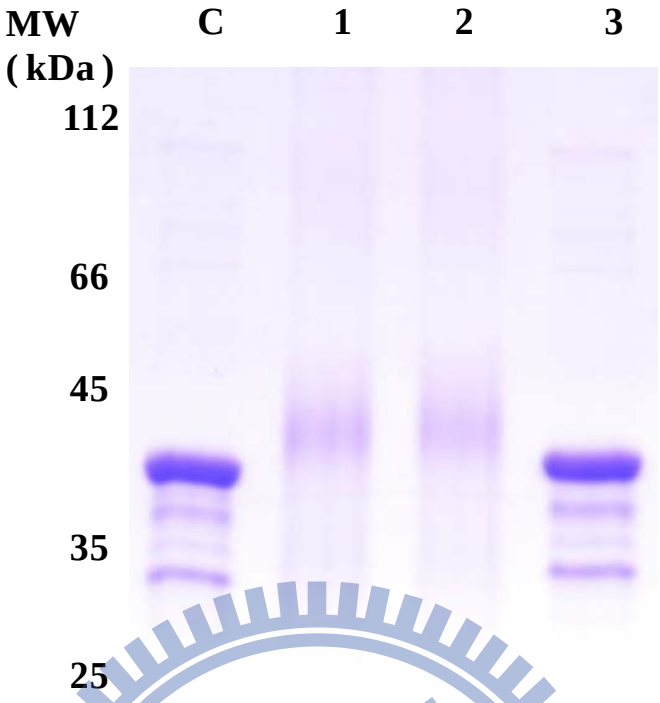


Figure 11. Identification of two peaks of TPST. EDC is used to yield stable amide bonds between two proteins to demonstrate the interaction of two proteins. 6μg *Dm*TPST was treated with 50mM EDC on total volume 20μl. (C) was the control which had no treatment. TPST was treated with EDC at 25°C for 2hr (1), 25°C for 5hr (2) and 4°C for 6hr (3).

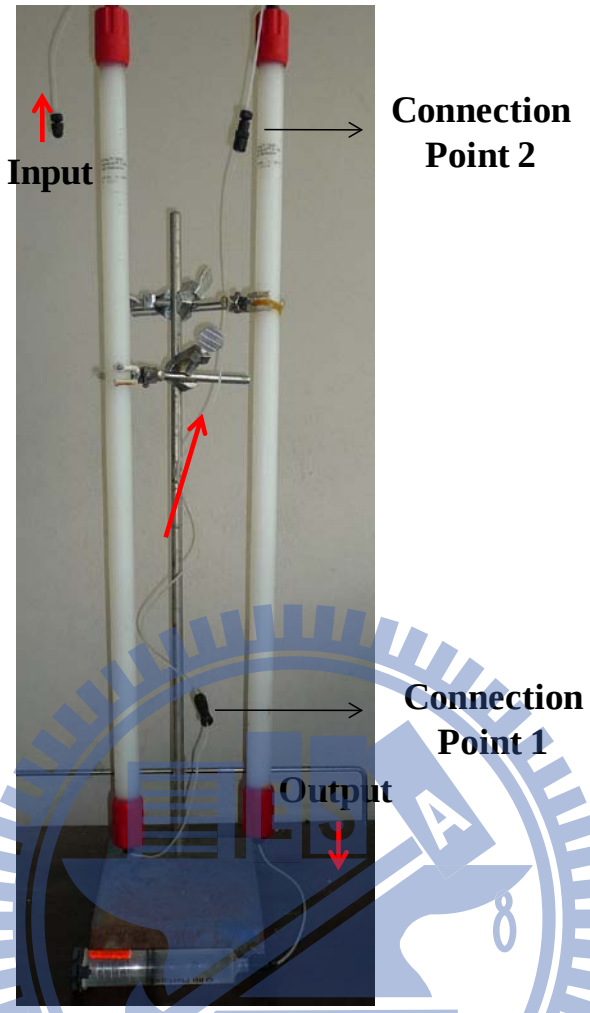


Figure 12. Picture of sephacryl S-100 HR- sephacryl S-200 HR connection. For joining sephacryl S-100 HR and sephacryl S-200 HR, a pipe connected sephacryl S-100 HR and sephacryl S-200 HR with connection point 1 and connection point 2, individually. The sample was inputted at sephacryl S-200 HR side, and outputted at sephacryl S-100 HR side. The red arrow indicates the sample flow way.

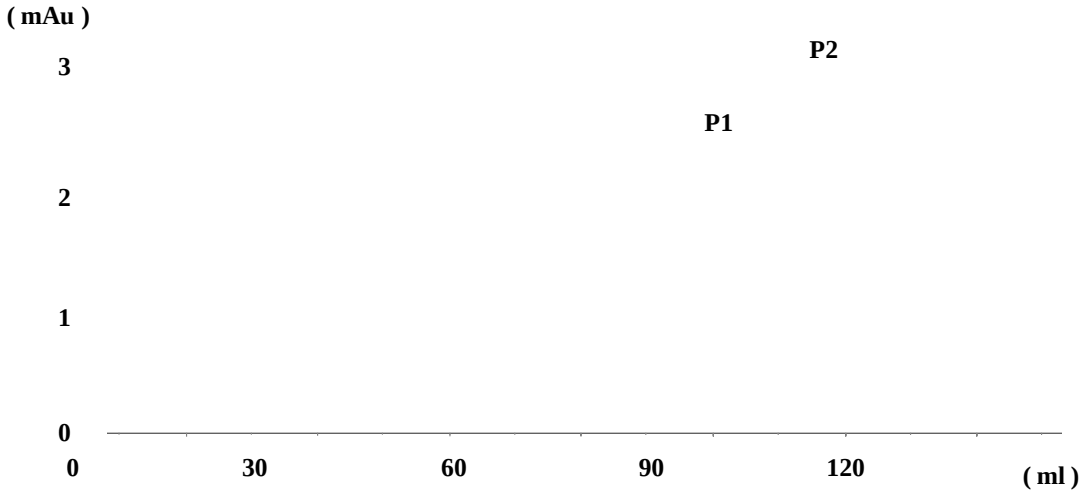
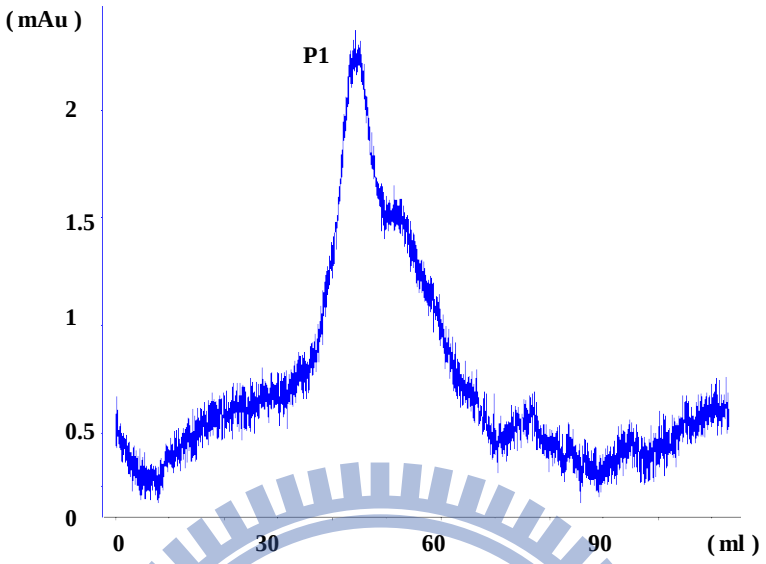


Figure 13. Separation of two forms structure of *Dm*TPST. *Dm*TPST(1 mg) was separation to peak1 and peak2 by sephacryl S-100 HR- sephacryl S-200 HR at optimal buffer 120ml.The peak1and peak 2 was eluted 101ml and 117ml, individually.

(a)



(b)

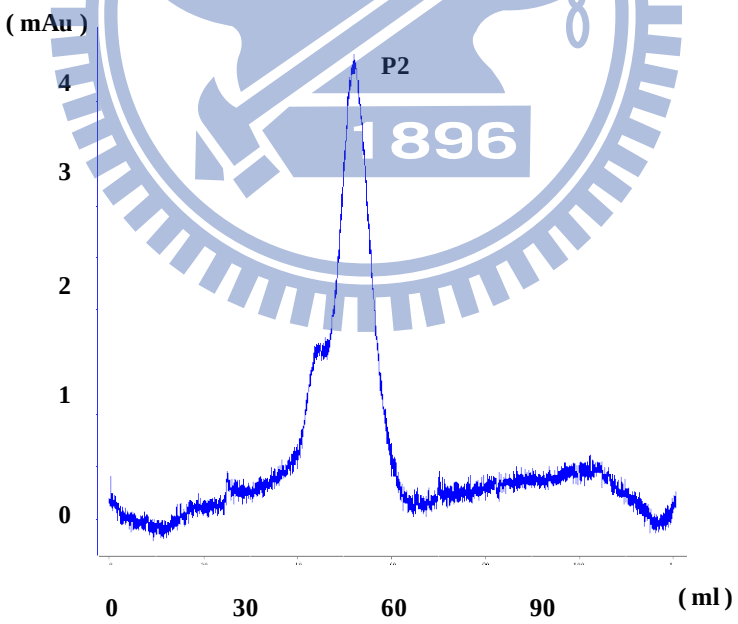
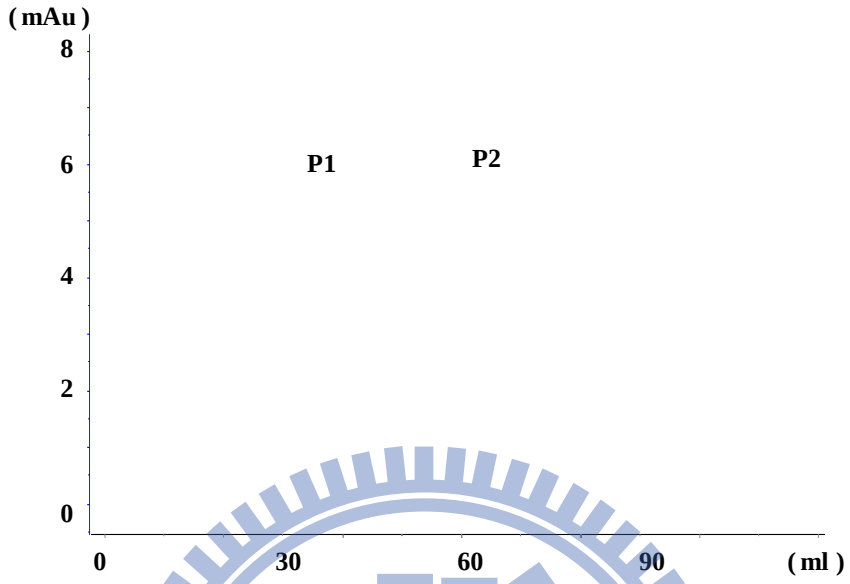


Figure 14. The two forms structure of *Dm*TPST. The *Dm*TPST (a) peak 1(0.23 mg) and (b) peak 2(0.38 mg) was separated by sephacryl S-100 HR- sephacryl S-200 HR, and re-injected into sephacryl S-100 HR at optimal buffer 120ml, individually.

(a)



(b)

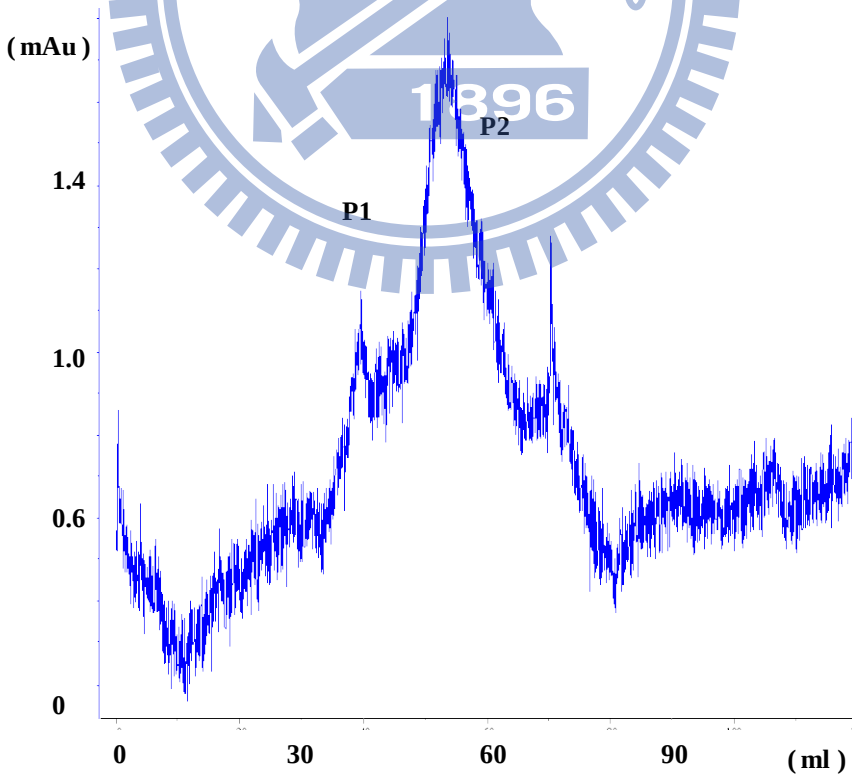
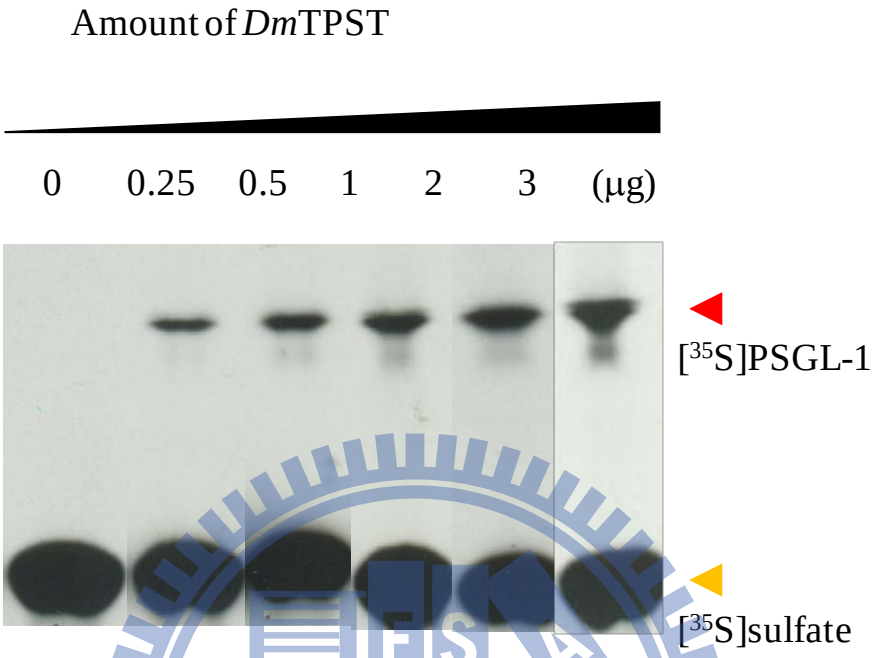


Figure 15. The influence of Sodium chloride on quaternary structure of *Dm*TPST.

The *Dm*TPST treated with (a) 500mM NaCl, and (b) 0mM NaCl to determine the influence on quaternary structure of *Dm*TPST, were analyzed by sephacryl S-100 HR at total volume 120 ml.

(a)


Amount of *Dm*TPST
0
0.25
0.5
1
2
3
(μg)
[35S]PSGL-1
[35S]sulfate

(b)

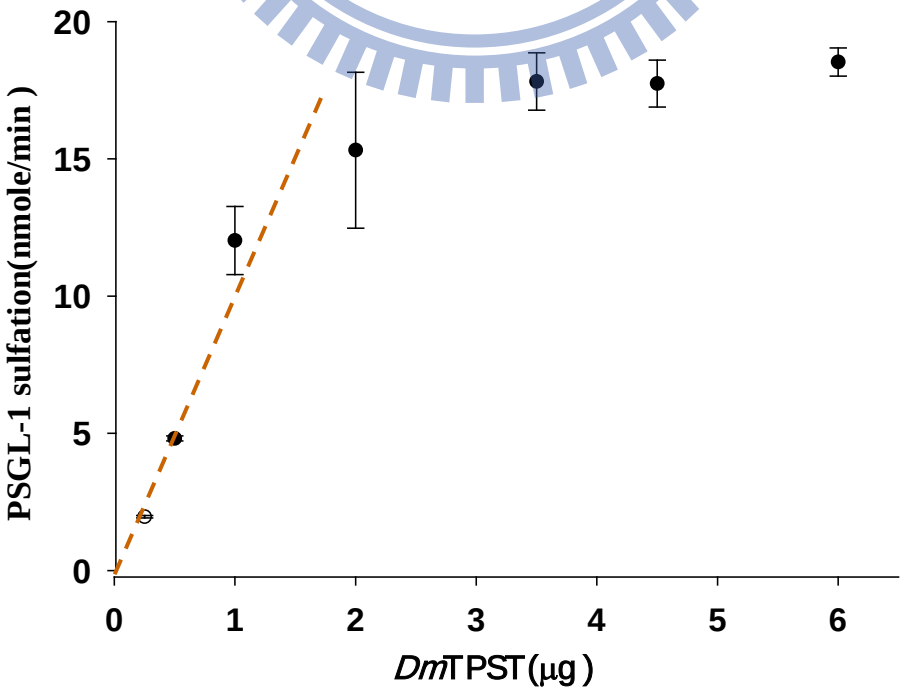

PSGL-1 sulfation(nmole/min)
20
15
10
5
0
0
1
2
3
4
5
6
*Dm*TPST (μg)

Figure 16. Reaction optimization for *Dm*TPST. (a) TPTS activity were measured by monitoring the sulfated substrate-[^{35}S]PSGL-1, which separated by thin-layer chromatography (TLC). . (b) With the enzyme gradient [^{35}S]PSGL-1 was increased. The dash line means the optimal reaction for TPST kinetics assay, and the red and orange triangle indicate the [^{35}S]PSGL-1 and [^{35}S]sulfate, individually

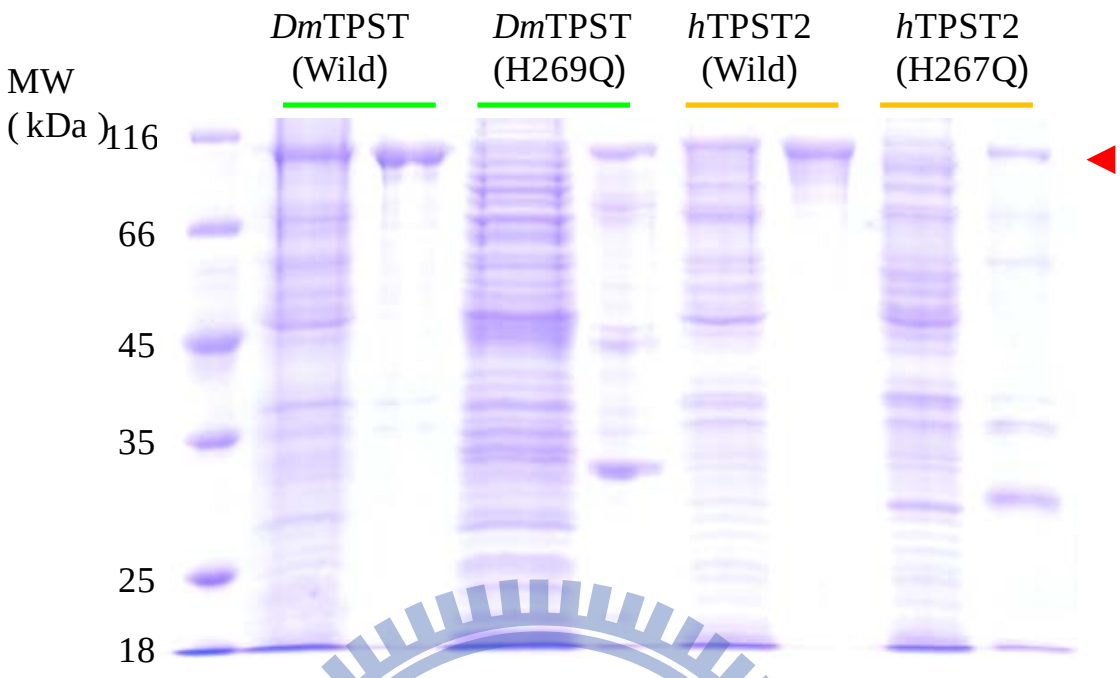


Figure 17. The expression levels of TPST mutation and wild type. The crude extract (left) and Ni-NTA column purified (right) of *Dm*TPST (wild), *Dm*TPST-H269Q, *h*TPST2 (wild), *h*TPST2-H2697Q, following as the order. The crude extract was 20 μg and purified TPST is 2μg. The red triangle indicated that the NusA-TPST was approximated 102 kDa.

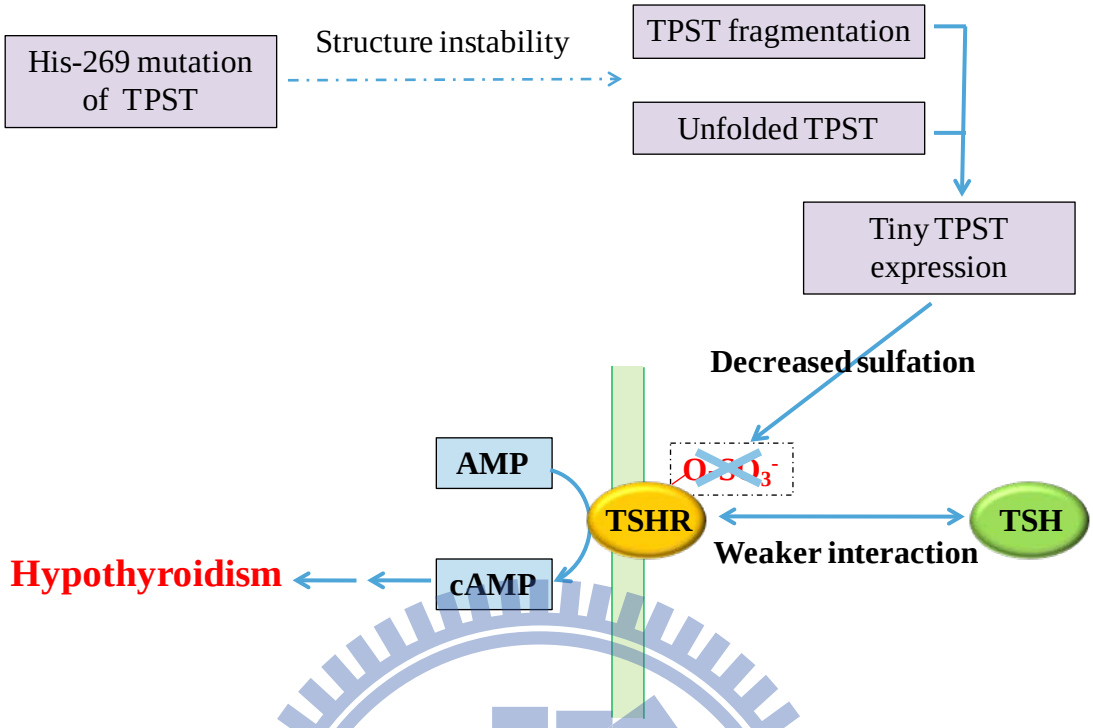


Figure 18. Possible Pathway of TPST-related Hypothyroidism. We proposed a pathway of hypothyroidism-related point mutation. A mutation on TPST may lead to structural instability of TPST, thus TPST would be fragmented or unable to folded TPST. This could lead to a lack of TPST expression decreasing sulfation on TPST substrate-TSHR. Sulfation less TSHR has weaker interaction with TSH, thus blocking the downstream signal transduction

TABLES

Table 1. Purification table of *h*TPST2.

	Total protein (mg)	Total activity (nmole.min^{-1})	Specific activity (nmol.min^{-1}.mg^{-1})	Yield (%)
Crude	**218.5**	**29.0**	**0.3**	**100**
Ni-NTA column	**2.7** [b]	**6.6**	**2.4** [a]	**23**
Thrombin digestion	**2.7** [b]	**N.D.** [c]	**N.D.** [c]	**N.D.** [c]

TPST activity was measured as indication under" Experimental Procedure."

[a] The specific activity of TPST through Ni-NTA column was calculated without NusA fusion tag.

[b] The total protein of TPST through Ni-NTA column showed without NusA fusion tag.

[c] N.D. mean TPST was too fragmented to detect enzyme activity

Table 2. Purification table of fusion tag-free *Dm*TPST.

Step	Total protein (mg)	Total activity ($nmole.min^{-1}$)	Specific activity ($nmol.min^{-1}.mg^{-1}$)	Yield[c] (%)	Purification[d] fold
Crude extract	436.2	45	0.10	100	1
Ni-NTA column	2.6 [b]	19	6.9 [a]	42	69
Hitrap Q column	0.6	5	7.71	10	75

[a] The specific activity of TPST through Ni-NTA column was calculated without NusA fusion tag.

[b] The total protein of TPST through Ni-NTA column showed without NusA fusion tag.

[c] Yield= total activity of product / total activity of crude extrude) * 100%

[d] Purification fold=specific activity of product / specific activity of crude extract

TPST activity was measured as indication under" Experimental Procedure."

Table 3. The standard protein for gel filtration analysis

S-100

Vo[a] (Blue Dextran 2000)：40.83			Vt[b](ml)	: 120
Name	**MW**	**log**	**Ve[c]**	**Kav[d]**
(standard)	**(kDa)**	**(MW)**	**(ml)**	
Albumin	**67**	**4.83**	**48.34**	**0.09**
Ovabumin	**43**	**4.63**	**53.81**	**0.16**
Chymotrypsinofen A	**25**	**4.40**	**67.07**	**0.33**
Ribonuclease A	**13.7**	**4.14**	**77.05**	**0.46**

[a] The Blue Dextran 2000 (2000 kDa) was as Vo, which retention volume was 40.83ml.

[b] Vt was the volume of column volume as 120 ml.

[c] Ve was the eluted volume of standard protein.

[d] Kav was calculated by equation: Kav = (Ve – Vo)/(Vt – Vo).

Table 4. Estimated MW of two quaternary structure of *Dm*TPST

Name	Theoretical	log	Kav[a]	Apparent[b]
	(kDa)	(MW)		(kDa)
***Dm*TPST-P1**	**76.6**	**4.58**	**0.07**	**70.7**
***Dm*TPST-P2**	**38.3**	**4.58**	**0.18**	**44.5**

[a] Kav was calculated by equation : Kav = (Ve – Vo)/(Vt – Vo) .

[b] The *Dm*TPST apparent molecular weight was acquired from the log(MW) and Kav relationship of a set standards.

Table 5. Comparison with kinetic constants of *Dm*TPST and *h*TPST2[a].

Enzyme	Kinetics		
	k_{cat}	K_m	k_{cat}/K_m
	(min^{-1})	(μM)	(M^{-1}sec^{-1})
***Dm*TPST**	**0.32 ± 0.013**	**42.1±5.3**	**126.7**
NusA-*Dm*TPST	**0.16 ± 0.007**	**12.0±2.5**	**222.2**
NusA-*h*TPST-2	**0.11 ± 0.006[b]**	**19.5±3.1[b]**	**73.9[b]**

[a] TPST activity was measured as indication under" Experimental Procedure."

[b] The kinetic constants of *h*TSPT-2 wild type was from Lu et al. unpublish.

Table 6. Purification efficient of *Dm*TPST- H269Q and *h*TPST2 -H267Q.

	Type	Step	Total protein (mg)	Total activity (nmole.min^{-1})	Specific activity (nmol.min^{-1}.mg^{-1})	Yield (%)	Purification fold
*Dm*TPST	Wild	Crude	436.2	45	0.1	100	1
		Ni-NTA	2.6[b]	19	6.9[a]	42	69
	H269Q	Crude	373.5	0.1	0.0002[a]	100	1
		Ni-NTA	0.5	0.1	0.2[a]	134	1027
*h*TPST-2	Wild	Crude	218.5	29	0.3	100	1
		Ni-NTA	2.7	7	2.4	23	7
	H267Q	Crude	512.0	0.1	0.0002	100	1
		Ni-NTA	1.2	0.4	0.3	330	1446

[a] The specific activity of TPST through Ni-NTA column was calculated without NusA fusion tag.

[b] The total protein of TPST through Ni-NTA column showed without NusA fusion tag.

TPST activity was measured as indication under" Experimental Procedure."

SUPPLEMENTARY DATA

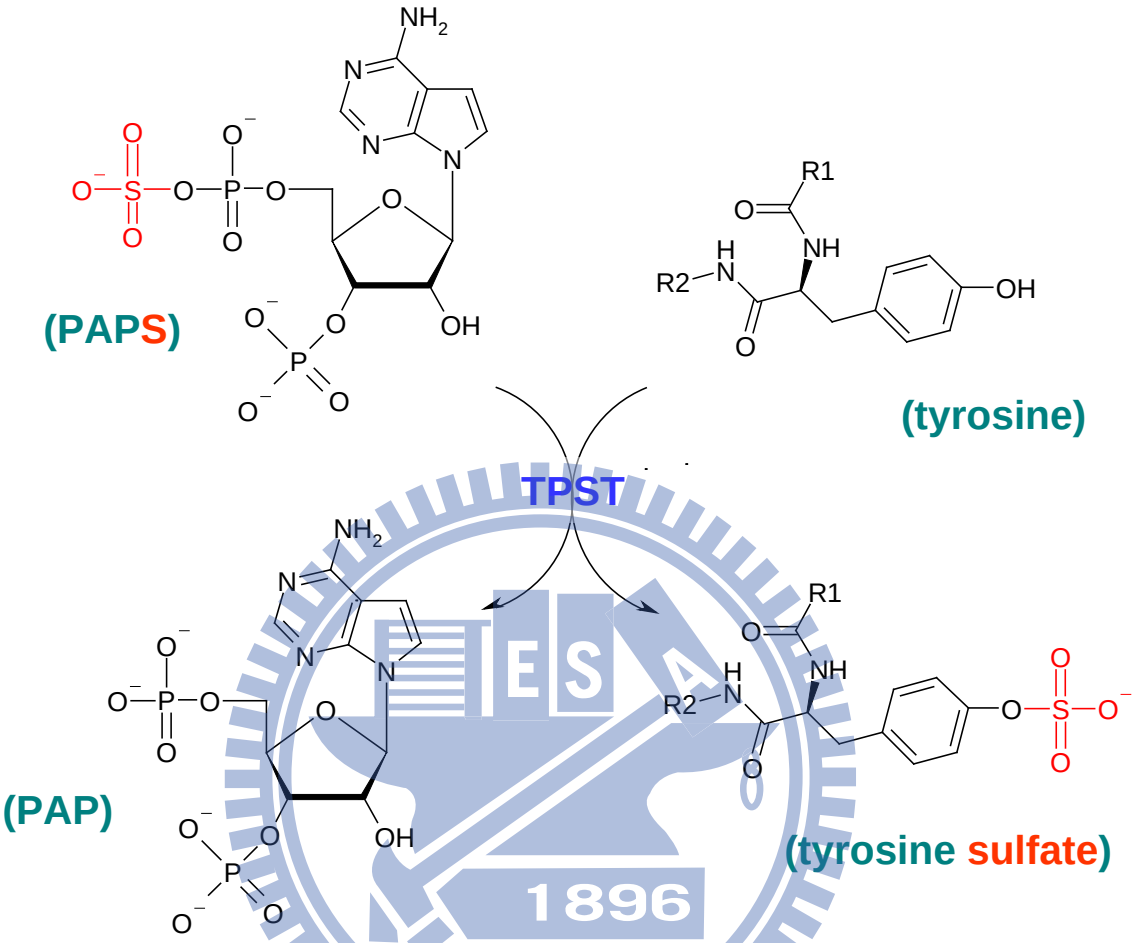


Figure S1. TPST catalytic reaction. TPST utilizes the unique sulfate donor, PAPS, and transfers sulfate group to tyrosine residue amongst the specific domain in proteins or peptides. The TPST and PAPS denote the tyrosylprotein sulfotransferase and 3'-phosphoadenosine 5'-phosphosulfate, respectively. This figure is supplied by ABD Lu- Yi Lu.

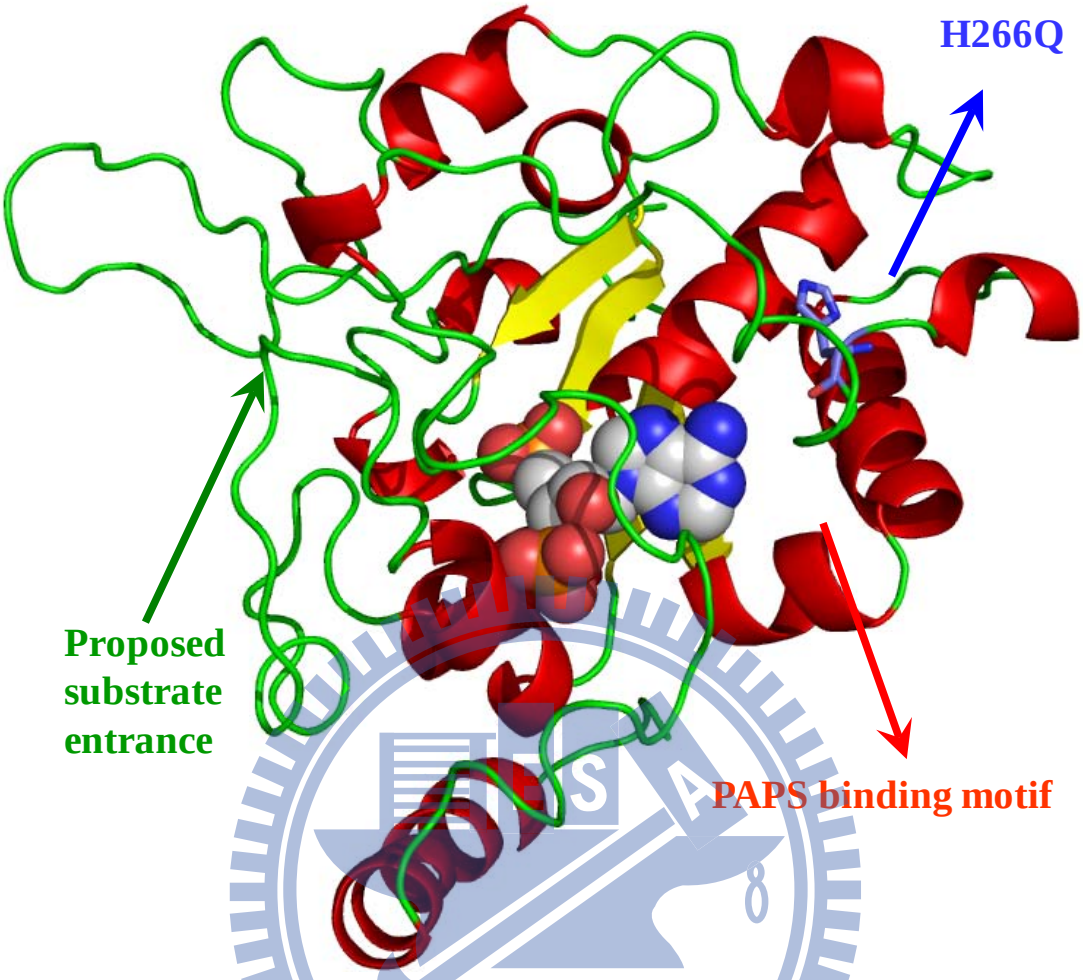


Figure S2. Simulation modeling of mTPST-2. We use the MODELLER server to model the TPST2 and use the 1xv1, the human sulfotransferase SULT1B1, to be the template. Overall the modeling score of this template is 0.94 and overwhelmingly higher than other template structures. Mostly the secondary structures are well-modeled and the PAP binding site is converged. This figure is supplied by ABD Lu- Yi Lu.

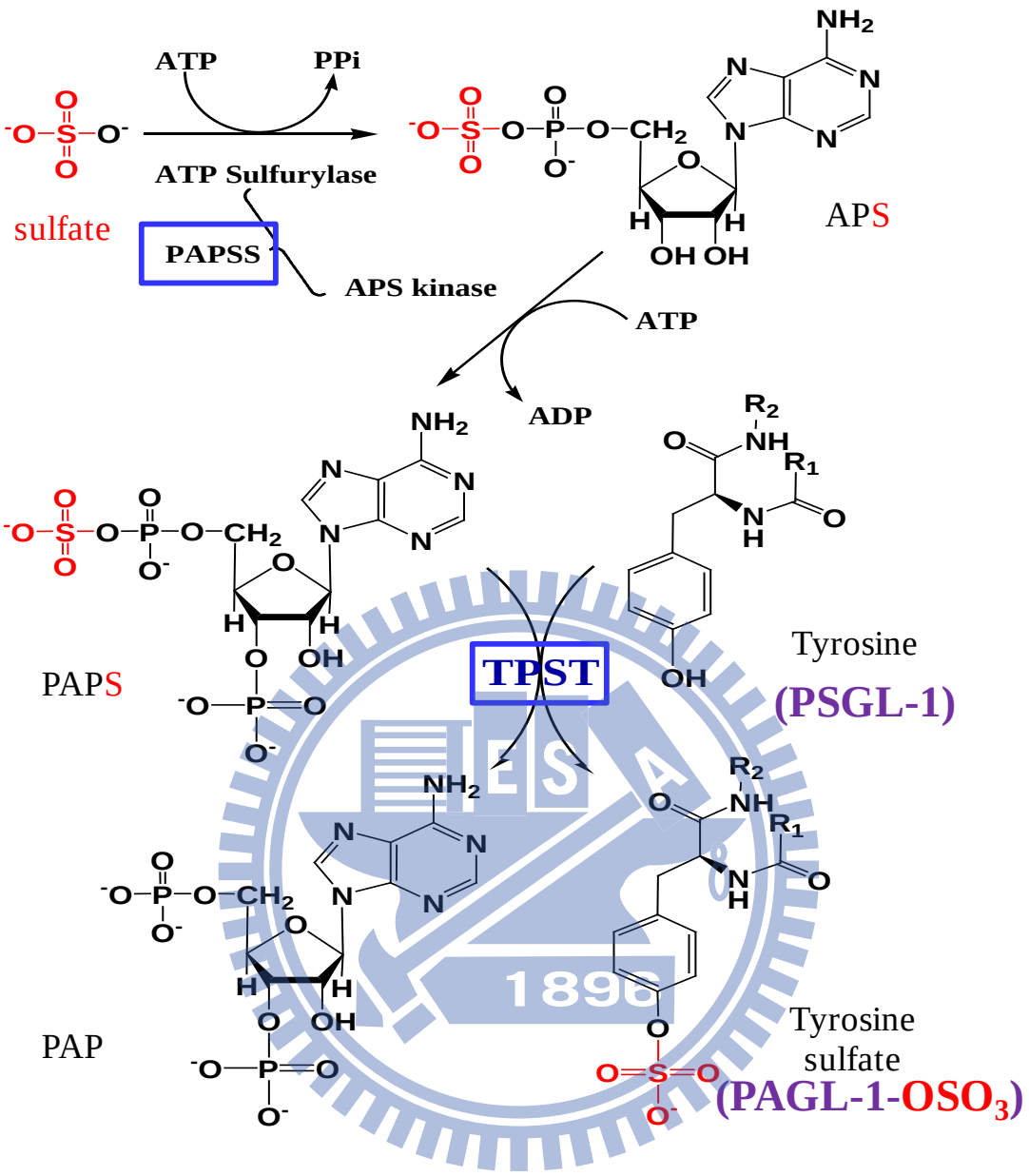


Figure S3. Scheme for the determination of TPST activity. The biosynthesis of PAPS from ATP and SO_4^{2-} was catalyzed by PAPSS, a bifunctional enzyme contains ATP sulfurylase and APS kinase activities (Step A). TPST transferred a moiety of sulfuryl group of the saturated PAPS generated from Step A to protein acceptors (Step B).