

國立清華大學生物資訊與結構生物研究所

碩士論文

Institute of Bioinformatics and structural Biology

National Tsing Hua University

Master Thesis

分析以昆蟲病毒系統表達之凝血酶敏感蛋白區域包含 7A 的功能性片段

Analysis of THSD7A functional domain expressed by baculoviral system

研究生：吳曉鈞 Hsiao-Chun Wu (g9680541)

指導教授：莊永仁 Dr. Yung-Jen Chuang

中華民國九十九年六月

摘要

內皮細胞的方向性遷移對血管及其形態的發育扮演重要的角色。它受到許多引導因子嚴密的調控。先前，我們發現一個新穎蛋白：凝血酶敏感蛋白區域包含 7A (Thrombospondin Type-1 Domain Containing 7A; THSD7A) 可能參與上述過程，但是其結構與功能的關聯機制尚未被深入了解。利用生物資訊分析的結果顯示，THSD7A 被預測為一個包含十一個凝血酶敏感蛋白重複區域 (TSR) 以及一個 RGD 序列的穿膜糖蛋白。這些特徵暗示 THSD7A 可能參與細胞的移動或細胞與胞外基質的交互作用。在我們先前的研究中指出，當斑馬魚體內 THSD7A 的表現減少時，會導致體節間血管 (ISV) 不正常的分歧與生長延遲。另一方面，我們觀察到在血管前端的內皮尖端細胞 (endothelial tip cell) 上生成過多的絲狀偽足 (filopodia)，其形成被視為調控遷移方向的關鍵因素。同時，我們也發現 THSD7A 的轉錄產物會表現在斑馬魚的神經系統中。基於上述的發現，我們假設 THSD7A 可能是一個從神經組織分泌出的血管引導蛋白，並且會在血管發育過程中扮演外源性負向調控內皮尖端細胞方向性遷移的因子。本篇論文研究的主要目的是利用活體外的分析方式找出 THSD7A 的功能性片段。首先，為了釐清人類 THSD7A 的轉譯後修飾與其同源異構體 (isoform)，我們藉由哺乳類細胞株表現其蛋白。我們亦利用桿狀病毒-昆蟲細胞表現系統成功建構 THSD7A 截切片段 TF2 之載體並表現蛋白。最後，我們利用穿透式細胞遷移試驗 (transwell migration assay) 與管柱生成試驗 (tube-like formation assay) 兩種與血管發育相關的活體外實驗，測試 TF2 的功能。結果顯示，THSD7A 是一個會被分泌到細胞外的糖蛋白。TF2-His 融合蛋白可以成功地藉由培養在 Express 5 培養基的 High5 昆蟲細胞表現並以適當的條件純化之。然而，我們發現 TF2 不會調控人類臍帶靜脈內皮細胞的遷移與管柱生成能力。這個結果暗示我們 THSD7A 的功能性片段可能不是座落在 TF2 上。

Abstract

Directed migration of endothelial cell (EC) plays a critical role in vascular growth and patterning processes. It is tightly regulated by various guidance cues and molecules. Previously, we identified a novel protein, Thrombospondin Type-1 Domain Containing 7A (THSD7A), which may involved in this process, but not much is known regarding its structural-functional mechanism. By bioinformatic analysis, THSD7A was predicted as a membrane glycoprotein which contains eleven thrombospondin-type-1 repeats (TSR) and one RGD motif. These characters imply THSD7A may be involved in cell migration and cell-to-ECM interactions. In our previous study, we found out that the down-regulation of zTHSD7A (THSD7A homolog in zebrafish) gene expression resulted in abnormal branching and stalled of intersegmental vessel (ISV). On the other hand, the endothelial tip cell at the leading front of growing vessel showed excessive filopodia formation which was thought to guide its direction. We also discovered that *Thsd7a* transcript was expressed in zebrafish nervous system. Based on these findings, we hypothesize that THSD7A may be a vessel guidance protein secreted by neural tissue and act as an exogenous negative regulator of the directed migration of endothelial tip cell during vascular growth. The specific aim of this thesis study is focusing on analyzing the functional domain of THSD7A *in vitro*. First, recombinant full-length human THSD7A was expressed by mammalian cells to reveal the isoforms of THSD7A and modification. Baculoviral vectors containing truncated fragment 2 (TF2) of THSD7A was then constructed and used to express the target protein region by insect cells system. The resulting TF2 were then tested with two *in vitro* angiogenic assays, which are the transwell migration assay and tube-like formation assay. These results demonstrated that there was a secreted form of THSD7A which was glycosylated. His-tagged TF2 (TF2-His) can indeed be expressed by High5 insect cells in Express 5 medium successfully, and an optimal purification procedure was established. However, we found TF2 has no activity on regulating the migration and tube-formation of HUVEC. This suggests that the functional domain of THSD7A may not locate within the TF2 region.

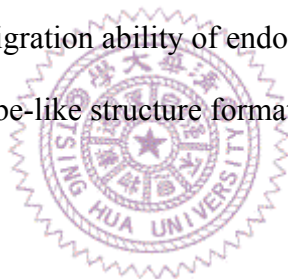
List of contents

摘要.....	I
Abstract.....	II
List of contents	II
List of figures.....	V
List of supplemental figures	VI
1. Introduction.....	1
1.1 Angiogenesis and its regulatory factors	1
1.2 The process of angiogenesis and endothelial cell directed migration.....	1
1.3 Thrombospondin structural homology repeat (TSR) and TSR-containing proteins	2
1.4 Thrombospondin Type 1 Domain Containing 7A (THSD7A) and its potential role in modulation of endothelial cells directed migration.....	2
1.5 N-glycosylation and its function.....	4
1.6 Baculovirus-insect cell expression system.....	4
1.7 The objective of this study	6
2. Materials and Methods.....	7
2.1 Construction of protein expression vectors.....	7
2.2 Cell cultures of HEK293T, Sf9, High5 and HUVEC	7
2.3 Transfection of mammalian cells	8
2.4 Mammalian cell lysis and secretory protein collection.....	8
2.5 Cotransfection of insect cells with virus DNA and transfer plasmid DNA	9
2.6 Plaque assay	9
2.7 PCR analysis of recombinant virus DNA	11
2.8 Virus production and scale up	12
2.9 End-point dilution	12
2.10 Time course of recombinant protein expression	12

2.11 Western Blotting.....	13
2.12 Purification of TF2 recombinant protein	14
2.13 Silver staining	15
2.14 HUVEC isolation	15
2.15 Functional assay of TF2 recombinant protein.....	16
3. Result s	17
3.1 Analysis of the putative human THSD7A.....	17
3.2 Human THSD7A is a secreted protein.....	17
3.3 Human THSD7A is N-glycosylated in mammalian cells	18
3.4 Generation of TF2 recombinant virus	19
3.5 Expression and identification of TF2 recombinant protein in High5 cells	21
3.6 TF2 recombinant protein has an N-glycosylated form in High5 cells.....	23
3.7 Purification of TF2 recombinant protein by nickel affinity chromatography	24
3.8 Validation of the purified TF2 recombinant protein	25
3.9 TF2 recombinant protein does not affect HUVEC migration and tube formation	26
4. Discussion.....	27
5. References	31

List of figures

Figure 1: The predicted Thrombospondin Type1 Domain Containing 7A	36
Figure 2: Human THSD7A is an N-glycoprotein, and it can be a secreted protein.	37
Figure 3: The construction of truncated human THSD7A expression vector	38
Figure 4: The morphology of 4-7 days post-transfected Sf9 insect cells.....	39
Figure 5: Verify the recombinant virus DNA by polymerase chain reaction.....	40
Figure 6: The optimal expression time point and MOI of TF2-His can be recognized	41
Figure 7: TF2-His expressed by insect cells have an N-glycosylated form.....	42
Figure 8: The establishment of an optimized purification protocol with stepwise elution of TF2-His recombinant protein.....	43
Figure 9: Identification of the TF2 recombinant protein	44
Figure 10: TF2 does not affect migration ability of endothelial cells.....	45
Figure 11: TF2 does not affect tube-like structure formation of endothelial cells.....	46



List of supplemental figures

Supplement 1: THSD7A may act as a repulsive regulator during ISV angiogenesis <i>in vivo</i>	47
Supplement 2: Overview the process of recombinant protein production.....	51
Supplement 3: Blue plaque formation reveals the potential of Sf9 cells infected by TF2 recombinant virus.....	52
Supplement 4: Excessive MOI resulted in similar expression level of two forms of TF2 and have poor secretion ability	53
Supplement 5: The impurities of TF2 elution is contaminated by the damaged cells.	54
Supplement 6: The structure of TF2 may be destructed after truncation.....	55
Supplement 7: The predicted binding domain contain a TSR and a CD36-binding mitif.....	56



1. Introduction

1.1 Angiogenesis and its regulatory factors

Angiogenesis, the recruitment of new blood vessels originate from pre-existing capillary plexus ¹, is a fundamental process in embryonic developmental stage, some postnatal physiological events and in several pathological events like tumor angiogenesis. In normal condition, it is tightly regulated under the balance of positive and negative regulatory factors ²⁻⁴, many of them have been proved to exist in extracellular matrix or circulations, acting as a chemotaxis signal ⁵. Imbalance of these factors results in various pathological conditions such as neovascularization, hemorrhage and age-related macular degeneration ⁶⁻⁸. The most important angiogenic factor is vascular endothelial growth factor (VEGF). VEGF and its receptors are well-characterized positive regulators of angiogenesis in developmental ⁸⁻¹² and pathological processes ^{12,13}. On the other hand, negative regulators like thrombospondin-1 (TSP-1) has been known as the first identified endogenous inhibitor of angiogenesis ¹⁴, the general functions of TSP-1 were to inhibit the migration, proliferation, apoptosis and tubulogenesis of endothelial cells (EC) *in vitro* ¹⁵⁻¹⁸ and inhibit new blood vessel formation *in vivo* ¹⁸⁻²¹. The comprehensive understanding of endogenous regulators would contribute to the understanding of the angiogenic homeostasis *in vivo*, and most importantly, it would accelerate the development of therapeutic strategies for abnormal angiogenesis.

1.2 The process of angiogenesis and endothelial cell directed migration

Angiogenesis is generally activated by a gradient of angiogenic factors binding to EC receptors to activate endothelial tip cell, and following by EC proliferation, directional decision and ECM remodeling, then finally form the blood vessel. Endothelial cell directed migration is a key step in angiogenesis. The migration process of activated EC consists of: (1) filopodia formation to sensing the direction of chemoattractants; (2) Lamellipodia formation at the leading edge to extend cell body; (3) Focal adhesion assembly of the newly extension site to attach to the extracellular matrix (ECM); (4) Stress fiber formation and further connect to focal adhesion sites to induce cell

body contraction and movement, disassembly of the connections to release rear edge of the cell body; and finally, (5) recycling of these regulatory factors to continue the process. The first activated EC was known as the tip cell, it has shown that the chemoattractants guides the migration of the cell via extension and traction of filopodia of the tip cell ²²⁻²⁴.

1.3 Thrombospondin structural homology repeat (TSR) and TSR-containing proteins

Numerous fragments of large proteins have been proved as endogenous inhibitors of angiogenesis ²⁵. One well-known domain was called thrombospondin structural homology repeats (TSR) which was derived from thrombospondin-1 (TSP-1). TSR is an evolutionarily-conserved domain which was defined to contain a hydrophobic cluster with three conserved tryptophans (WxxWxxW), a basic cluster with two conserved arginines (RxR) and six conserved cysteines ²⁶. The function of TSR was thought to inhibit EC migration ²⁷ and angiogenesis ^{28,29}. Besides thrombospondin-1, there are numerous TSR-containing proteins which are commonly involved in regulation of cell adhesion, migration and tissue remodeling ^{20,30-35}.

1.4 Thrombospondin Type 1 Domain Containing 7A (THSD7A) and its potential role in modulation of endothelial cells directed migration

THSD7A was predicted as a multiple TSR-containing protein whose protein sequence is conserved among the vertebrates such as *H. sapiens*, *P. troglodytes*, *M. musculus* and *D. rerio* (zebrafish). Previous data from our group showed that knockdown of *Thsd7a* impaired intersegmental vessels (ISV) angiogenesis of zebrafish, which implied that THSD7A regulated vessel patterning *in vivo* (unpublished data from Chieh-Huei Wang). When we proceeded to examine the endothelial tip cells in detail, we observed excessive filopodia extensions on tip cells (unpublished data from Chian-Huei Wang). It has been proved that endothelial tip cell guides its way through the modulation of filopodia ²². Additionally, the lost of suppression signal resulted in increased tip cell filopodia ²⁴. These evidences supported that the tip cell lost its sense of direction

when we knockdown THSD7A, this implied that THSD7A guides leading front of growing vessel through modulation the direction of activated tip cell migration. Next, we investigated whether THSD7A caused this effect via modulation of EC mobility. The up-regulation and down-regulation of endogenous THSD7A in human umbilical vein endothelial cells (HUVEC) showed that THSD7A inhibited EC migration and tube-like structure formation *in vitro* ³⁶. Accordingly, these data suggests that THSD7A may involve in angiogenesis through the coordination of endothelial cell directed migration.

Furthermore, there are evidences indirectly implied the possible relationship between THSD7A and cell movement. First, we found that C-terminus of THSD7A colocalized with integrin α V β 3 and paxillin at focal adhesion sites in HUVEC ³⁶. When inspecting the C-terminal peptide sequence of THSD7A, we found it contains a RGD motif, which is known to interact with integrin. Additionally, a large numbers of research also indicated that the binding of RGD can active integrin at focal adhesion sites. These sites are connected to actin cytoskeleton to drive cell migration ^{37,38}. Moreover, paxillin have been shown to modulate actin filament remodeling at focal adhesion sites, too ^{39,40}. These evidences give us the hint that THSD7A may be participated in α V β 3 and paxillin mediated EC migration. Second, when actin polymerization was inhibited, actin formed aggregates and THSD7A was stuck into actin aggregates. This result indicated that THSD7A may involve in actin polymerization. Third, the knockdown of THSD7A in HUVEC caused multiple broad Lamellipodia-like formation (unpublished data from Pei-Tsu Su, Zih-Yin Lai). These evidences supported the connection between THSD7A and EC migration.

To study THSD7A *in vivo*, we surveyed the *Thsd7a* transcript expression pattern in zebrafish. The data showed that it can be detected in the central nervous system (unpublished data from Chieh-Huei Wang). This result is consistent with the expression of *Thsd7a* ortholog in mouse embryo (data from Kuo Lab, Stanford University). Although THSD7A was found to be a neuron-derived protein in zebrafish, it regulated the formation of vessels. Previous studies demonstrated that the patterning of vascular system showed high similarity to the nervous system ^{41,42}. Furthermore, it has been proved that neural-derived protein can regulate vessel formation and

EC migration, which was known as neural-driven angiogenesis^{43,44}. In conclusion, it seems that THSD7A may act as an exogenous factor which derived from neuronal cells, and participate in neural-driven angiogenesis through regulation of directed EC migration (supplement 1). However, it should be further validate whether the hypothetical secreted form of THSD7A may exist, and which fragment or TSR of THSD7A may participate in this process.

1.5 N-glycosylation and its function

A growing body of evidences suggests that the precise addition of N-glycans to proteins play two pivotal roles. First, it helps with the folding and assembly of proteins in the ER, and promotes protein stability and secretion⁴⁵⁻⁵¹. Each of these function which mentioned above were closely linked to each other. For example, incorrect folding has shown to disturb the secretion of proteins^{52,53}. Second, the carbohydrates on proteins outside of the cell provide particular recognition “tags” for interactions with various external effectors, and affect both its signaling and function consequently^{48-51,54,55}. More and more studies revealed that protein glycosylation can regulate cell migration⁵⁶⁻⁵⁹. On the contrary, if the glycosylation of protein is incorrect, it may lead to low stability, incorrect folding, low secretion, and furthermore, affect the function of the protein⁶⁰⁻⁶⁴. Besides, there is a statistic data demonstrated that the relationship between glycosylation and folding: incorrect protein glycosylation will lead to 80% improper folding⁴⁸.

1.6 Baculovirus-insect cell expression system

The baculovirus expression system have been studied and used extensively since 1983 after an eukaryotic expressing vector was developed from *Autographa California* multiple nucleopolyhedro virus (AcMNPV) and used to produce biologically active β -interferon⁶⁵. To date, numerous bioactive proteins have been produced by baculovirus expression system successfully⁶⁶⁻⁶⁸, it has been used to produce protein pharmaceuticals and vaccines for medical usage for decades. The most special character of baculovirus is the formation of polyhedrin after infected host

cells, and the sequentially formation of occlusion body (OB) inside the nucleus, this is the reason why it belongs to multiple nucleopolyhedro virus (MNPV). The reason for why baculovirus can be developed to be a popular extracellular protein expression system is due to their strong promoters. For example, the polyhedrin and p10 genes both can be drive by their strong promoter, but they belong to non-essential structure genes ⁶⁹. Without these genes, the propagation of baculovirus in host cells will not be affected. On the other words, these genes can be displaced by other foreign gene. This gene can be drove and expressed in a large quantity. Sf9, Sf21 and High5 insect cells are commonly used for baculovirus expression. Except for its strong promoter which can drive exogenous gene in high efficiency and lead to high protein expression levels, this expression system has other potential advantages: (1) Baculovirus has the capacity to insert 10kbp exogenous gene due to their greatest genome. (2) Genetically, the hosts of baculovirus were limited in invertebrate, so the operation of baculovirus shows high safety. (3) It contains higher eukaryotic post-translational modification system which is capable of produce mammalian-like glycoproteins. (4) It can excise the signal peptide correctly and it has higher proper folding rate. (5) It is benefit for producing exogenous protein with their serum free and high-density suspension culture. Although there are many other kinds of expression systems, such as: bacteria, yeast and mammalian cells ⁷⁰, compared insect cells with bacteria and yeast expression systems, it has more complete post-translational modification. Furthermore, it is capable of producing secreted protein but not forming inclusion bodies, and it provides proper environment for protein folding and S-S bond formation than bacteria, thus it has better chance to get biological activity secreted proteins. Compared with mammalian cells, it contains higher expression level and it allows great quantity of suspension culture, and it can be cultured in serum free circumstances, it seems that it has the higher opportunity to purify high yield of purified protein by easier purification steps. In conclusion, baculovirus-insect cell system seems to be an optimal choice for functional exogenous protein expression because it can make up the deficiency of other systems, and to produce high purity of functional protein in a cost-effective manner. Bac-N-BlueTM Baculovirus expression system (invitrogen) was used in this study, the principle of recombinant protein expression were demonstrated in supplemental figure

(supplement 2).

1.7 The objective of this study

In summary, we hypothesize that THSD7A may be a secreted protein which may act as a repulsive regulator during ISV angiogenesis *in vivo* (supplement 1). To simplify the circumstances, I try to study whether THSD7A inhibit EC migration and angiogenesis *in vitro*? Here, endothelial cell (EC) was used as the cell model. I first investigate whether there is a secreted form of THSD7A. If it does exist, what functional domain of THSD7A is responsible for directing EC migration *in vitro*? As a whole, the understanding on THSD7A functional domain in EC directed migration could be valuable in developing new therapeutic agents for inhibiting pathological angiogenesis.



2. Materials and Methods

2.1 Construction of protein expression vectors

Full length human THSD7A containing endogenous signal peptide was constructed in pCMV-Tag4C expression vector. Truncated fragment TF2 with His-tag or GST-tag and GST only were constructed in pMelBacA (Invitrogen), a baculovirus transfer vector for secretion of recombinant proteins.

2.2 Cell cultures of HEK293T, Sf9, High5 and HUVEC

Human embryonic kidney 293T cells were cultured in Dulbecco's Modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum (Gibco), 100 units/ml penicillin (Gibco), 100 µg/ml streptomycin (Gibco), 0.25 µg/ml fungizone (BioSource) and 3.7 g/L sodium bicarbonate in a humidified 5% CO₂ incubator at 37°C. Sf9 (*Spodoptera frugiperda*) insect cells were purchased from invitrogen and maintained in complete TNM-FH insect medium (Sigma-Aldrich) supplemented with 5% heat-inactivated FBS, 100 units/ml penicillin, 100 µg/ml streptomycin and 0.25 µg/ml fungizone, then cultured in a non-humidified environment incubator at 28°C. For large scale suspension cell culture, 0.1% Pluronic® F-68 was added into the culture medium as cushioning reagent. High5 insect cells were cultured in Express Five® SFM (Gibco) serum-free medium and supplemented with 100 units/ml penicillin, 100 µg/ml streptomycin, 0.25 µg/ml fungizone and 10 µg/ml gentamycin, then cultured in a non-humidified environment incubator at 28°C. Another medium used was EX-CELL™ 405 (safc biosciences) supplemented with 100 units/ml penicillin, 100 µg/ml streptomycin, 0.25 µg/ml fungizone and 0.35 g/L of sodium bicarbonate. Human umbilical vein endothelial cells (HUVEC) were maintained in M200 supplemented with low serum growth supplement (LSGS; Cascade Biologic) 100 units/ml penicillin, 100 µg/ml streptomycin and 2.5 µg/ml fungizone at 37°C in ambient air with 5% CO₂. HUVEC were used between passages 2 and 5 in all experiments.

2.3 Transfection of mammalian cells

HEK293T cells were transfected with relevant vectors using jetPEITM transfection reagent (Polyplus-transfection) according to the manufacturer's instructions. 3×10^6 HEK293T cells were seeded in 10 cm dishes overnight before transfection, approximately 50% cell confluence would form at the day of transfection. Endotoxin-free plasmid DNA of pCMV-Tag4C and pCMV-Tga4C-THSD7A were prepared by using the endo-free plasmid DNA midi kit (Invitrogen) to prevent endotoxin-induced cell death. Briefly, 16 μ l jetPEITM and 8 μ g plasmid DNA were respectively added in 500 μ l of 150 mM NaCl_(aq), added the jetPEITM solution to the DNA containing solution once. The 1 ml mixture was incubated for 20 minutes at room temperature and then added in each dish, incubated the dishes at 37°C in a CO₂ incubator. After 48 hours, cell lysate could be collected for the following experiments.

2.4 Mammalian cell lysis and secretory protein collection

1. Secretory protein collection

Culture medium was collected and kept on ice. Amicon[®] Ultra-4 Centrifugal Filter Devices (Millipore) was used to concentrate culture medium by centrifuging at $4,000 \times g$ in a swinging-bucket rotor at 4°C, the final retentate volume of the medium was 100 μ l. The concentrated medium was collected from the tube and transferred to a new tube for later use.

2. Adherent cell protein collection

Cell lysate was extracted by the RIPA buffer (25 mM Tris·HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 1 mM PMSF and 1 \times protease inhibitor cocktails (Roche)). After the culture medium was removed from adherent cells, the cells were washed twice with cold PBS. Cold RIPA buffer was added to the cells and kept it on ice. Then, swirled the dish to distribute the buffer evenly and incubated on 4°C shaker for 30 minutes. Cells lysate was collected by using a cell scraper and passed through a 25-gauge needle for several times. The samples were centrifuged at $16,000 \times g$ for 15 minutes at 4°C to pellet the cell debris. The supernatant were

collected and transferred to a new tube for further analysis.

2.5 Cotransfection of insect cells with virus DNA and transfer plasmid DNA

1. Seed culture dishes with insect cells

2×10^6 log-phase Sf9 cells in 2 ml TNM-FH medium were seeded into a 60 mm dish. The dishes were rocked back and forth gently, then side to side to evenly distribute cells. The cells were allowed to attach to the bottom of the dish for at least 30 minutes.

2. Recombination of virus DNA and transfer plasmid DNA

1.5 ml tube containing 10 μ l (5.5 μ g) of the Bac-N-BlueTM DNA was centrifuged and then added the following reagents: 1 ml of Grace's insect medium (without supplements or FBS), 4 μ l of the transfer plasmid (1 μ g/ μ l) and 20 μ l of the Cellfectin[®] Reagent. This transfection mixture was mixed gently and incubated at room temperature for 15 minutes. In the same time, removed the medium from the cell monolayer carefully, then washed the monolayer by adding 2 ml of fresh Grace's medium without supplement or FBS and then removed the medium from the cell again, after that, the transfection mixture was added into the dishes, the dishes were then incubated at room temperature for 4 hours on a platform rocker providing a smooth side to side rocking motion. Following the 4 hours incubation period, 1 ml of complete TNM-FH medium was added to the dishes, sealed the dishes with parafilm and incubated at 28°C for at least 72 hours.

3. Harvest transfection virus stock

After 72 hours, harvested 2 ml of medium from transfected dish to a sterile 15 ml tube, this transfection virus stock was filtered and stored at 4°C. To further confirm that the transfection was successful, 3 ml of fresh complement TNM-FH medium was added to the transfected dishes and incubated for additional 96 hours, checked it again 4-7 days posttransfection.

2.6 Plaque assay

1. Prepare agarose solution

Low melting agarose powder was dissolved into sterile water to final concentration 2% and then autoclaved. Agarose solution was equilibrated and put into 47°C water bath.

2. Seed culture dishes with insect cells

3×10^6 log-phase Sf9 cells in 2 ml TNM-FH medium were seeded into a 60 mm dish. The dishes were rocked back and forth gently, then side to side to evenly distribute cells. The cells were allowed to adhere for 45 minutes.

3. Prepare TNM-FH/X-gal mixture

X-gal (20 mg/ml) and 2 × TNM-FH medium were mixed. The final concentration of X-gal was 400 µg/ml. The mixture was then filtered by 0.22 µm filter.

4. Prepare virus serial dilutions

For each infection condition, 1.5 ml sterile tubes were labeled from 10^{-1} to 10^{-3} . 450 µl TNM-FH medium was added to each tube and then 50 µl cotransfect medium was added to the first tube then mixed thoroughly using a vortex mixer. Serial dilutions were performed by transferring 50 µl to the subsequent tube and mixed. A fresh pipet was used at each time.

5. Infect monolayer cells

Several plates were checked under a light microscope to confirm that the cell confluence of each plate is 70% and the cells were evenly distributed. Aspirated off 1.2 ml medium from the plates using a sterile pipette tip carefully and do not disrupt the cell monolayer. 0.4 ml of the virus dilution was added to each plate, rocked plate back and forth gently, then side to side to evenly distribute virus. Plates were sealed by parafilm and rocked the plates gently by low speed shaker for one hour at room temperature.

6. Overlay infected cells with agarose

About one hour later, 2 ml of 2 × TNM-FH/X-gal mixture and 2 ml of 2% agarose solution were mixed, the temperature of the mixture was checked at the same time and should be warm enough to stay in the liquid state, but cool enough for the cells. Next, medium was aspirated off carefully from the edge of the plate using sterile pipette tip, then 1% agarose solution was added slowly from the edge of the plate. The plate was kept untroubled for a while and the agarose was

allowed to solidify. Another overlay was prepared during this time. After the agarose of all plates were solidified, sealed plates with parafilm and incubated at 28°C for 7-10 days.

2.7 PCR analysis of recombinant virus DNA

1. Amplify virus

2.5×10^5 log-phase Sf9 cells in 0.8 ml TNM-FH medium were seeded into a 24-well plate. Bleu plaques were picked and put into each well by using sterile pipette tips, white plaque should be picked as negative control. To ensure proper isolation, plaques should be picked from plates containing fewer than 50 plaques. 24-well plate was incubated at 27°C for 3 days. After 3 days, 0.4 ml of medium containing virus were harvested from each well and transferred to a sterile 1.5 ml tube, centrifuged at $1,000 \times g$ for 5 minutes at 4°C to remove cell debris, then stored at 4°C, it is used to isolate virus DNA and perform PCR analysis of the putative recombinant virus. The 24-well plate should be kept at 27°C until all of the cells lysed (5-7 days), then the rest portion of the medium was collected and centrifuged at $1,000 \times g$ for 5 minutes at 4°C to remove cell debris. Stored at 4°C and prevented from light, the passage one (P-1 virus) recombinant virus stock will be used to generate a large-scale, high-titer virus stock.

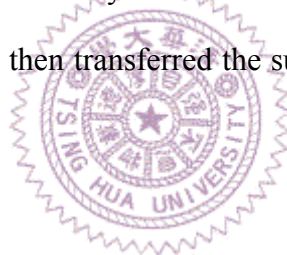
2. Isolate virus DNA and PCR analysis

Cell debris was centrifuged at 5,000 rpm for 10 minutes at room temperature and the supernatant was transferred to a new sterile 1.5 ml tube. Added 0.38 ml cold (4°C) 20% polyethylene glycol (PEG) in 1 M NaCl, inverted twice and allowed to stand at room temperature for 30 minutes, then centrifuged at maximum speed for 20 minutes at room temperature. Discarded the supernatant and 50 µl of sterile water was added, the sides of the tubes was washed carefully to ensure that all of the virus particles were resuspended. Added 5 µl of Proteinase K (5-10 mg/ml) and incubated at 50°C for 1 hour. Virus DNA was extracted by 55 µl of phenol-chloroform reagent (1:1). Then, centrifuged at maximum speed for 10 minutes at room temperature and transferred the upper aqueous phase to a new sterile tube. DNA was precipitated by the addition of 1/10 volume of 3 M sodium acetate, 2.5 µl of glycogen and 2 volumes of 100% ethanol, then incubated at -20°C for

overnight. The next day, centrifuged at maximum speed for 25 minutes at 4°C and washed the pellet with 70% ethanol, then centrifuged again and removed all traces of ethanol. Pellet obtained was resuspended in 8 µl of sterile water. These DNA was used as the PCR template.

2.8 Virus production and scale up

2×10^6 log-phase Sf9 cells in 5 ml TNM-FH medium were seeded into two 25 cm² flasks. Added 20 µl of the P-1 virus stock to each flask and incubated at 27°C for about 5-10 days until more than half of the cells lysed. All the lysed cells and medium were collected by sterile pipet and stored at 4°C. This is P-2 virus stock. Then, added 2.5 ml of the P-2 virus stock to a 250 ml suspension culture of log-phase Sf9 cells (2×10^6 cells/ml) in TNM-FH containing Pluronic® F-68, then incubated the suspension culture at 27°C with constant stirring (120 rpm) for 7-10 days, harvested when more than half of the cells lysed. Medium was centrifuged at $1,000 \times g$ for 20 minutes at 4°C to remove cell debris, then transferred the supernatant to a sterile tube, filtered and stored at 4°C, this is P-3 virus stock.



2.9 End-point dilution

Tenfold serial dilutions of the P-3 virus stock were prepared from 10^{-2} to 10^{-9} , then trypsinized Sf9 cells and diluted the cells to a concentration of 1×10^5 cells/ml with complete TNM-FH. 110 µl of the each virus dilution and 1100 µl of the Sf9 cells were mixed, 100 µl of the mixture was seeded into a 96 well plate for ten repeats, seeded another well of cells without any virus as uninfected controls. The plate was incubated at 27°C for about 7 days. After 7 days, added 1 µl of X-gal (20 mg/ml) to each well, about 30 minutes later, the color of the medium would change, blue wells and white wells were counted for virus titer calculation.

2.10 Time course of recombinant protein expression

Three spinner flasks with High5 insect cells at a density of 2×10^6 cells/ml were prepared,

the amount of viruses were calculated and added in each spinner flasks, for example, MOI (multiplicity of infection) =1, MOI=5 or MOI=10 were used. The spinner flasks were incubated at 27°C with a constant spin rate of 120 rpm. Removed 1 ml of the cells and transferred to a 1.5 ml tube every 12 hours. The cells were pelleted at 800 × g for 10 minutes at 4°C. The samples were kept on ice to prevent proteolysis. Transferred the supernatant to a new tube and labeled. The cell pellet and the supernatant were stored at -70°C to reduce proteolysis.

2.11 Western Blotting

10% polyacrylamide gels were prepared and used to analyze protein samples. Proteins were mixed with 3× sample dye containing 2% (v/v) β-mercaptoethanol (Sigma-Aldrich) and incubated at 95°C for 10 minutes, the samples were then spun down and loaded into each performed wells sequentially. The Novex® Sharp™ Pre-stained protein standard was used as the marker of protein molecular weight. Gel was stacked at 70 volts and run at 140 volts. After gel electrophoresis, western blot was performed according to the Trans-Blot® Electrophoretic Transfer Cell Instruction Manual (Bio-Rad). Proteins in the gel were transferred to a 0.22 μm polyvinylidene difluoride membrane (Millipore) in ice-cooled transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol). This step was run at 100 volts for about 1.5 hours. Then, 5% NFDM in TBST (200 mM Tris-base, 1.5 M NaCl, pH7.6, 0.002% Tween-20) or 3% BSA (Sigma-Aldrich) were used to block the PVDF for 1 hour at room temperature. The membranes were probed with each primary antibody and incubated at 4°C for overnight with gentle agitation. The next day, primary antibodies were removed and the membranes were washed 4 times with TBST for 10 minutes on an orbital shaker. After washing, TBST was then removed and replaced with HRP-conjugated secondary antibodies in 5% NFDM/TBST and incubated for 1 hour at room temperature with gentle agitation. Next, secondary antibodies were removed and the membranes were washed 4 times with TBST for 10 minutes again. ECL solution A and solution B (Perkin Elmer) were mixed before use and added onto the membrane for 1 ~ 2 minutes, it should be protected from light. This analysis was then completed by draining the membrane, mounted on film in developing cassette and exposed for

varied duration of times.

2.12 Purification of TF2 recombinant protein

1. Column packing

TF2-His recombinant protein was purified by using Ni-Sepharose-6-Fast-Flow (GE Healthcare) resin, the resin were preswollen in 20% ethanol. Packed 3 ml of the resin into a packing reservoir and washed the resin with 10 ml sterile water three times, then rinsed the resin with 10 ml binding buffer (20 mM Tris·Cl, 500 mM NaCl, 0 mM imidazole, pH7.9) three times.

2. Protein sample preparation

After High5 cells were infected at MOI=5 for 72 hours, the suspension culture was centrifuged at $1,500 \times g$ for 30 minutes at 4°C. The supernatant was collected and kept on ice. Protease inhibitor PMSF was added into the supernatant, the final concentration of PMSF was 1 mM. Then, protein sample was dialyzed overnight against dialysis buffer (20 mM Tris, 500 mM NaCl) using SnakeSkin™ Pleated Dialysis Tubing (Thermo). Next day, the protein sample was filtered by 0.45 µm filter and then added imidazole to give a final concentration of 15 mM. Incubated protein sample with binding buffer rinsed Ni-Sepharose-6-Fast-Flow resin by using the rotor at 10 rpm for 1 hour at 4°C.

3. Recombinant protein purification

After 1 hour incubation, packed the resin into a packing reservoir and then washed the resin with binding buffer containing different concentrations of imidazole sequentially. The resin were washed with 12 ml binding buffer (20 mM Tris·Cl, 500 mM NaCl, pH7.9) containing 0 mM imidazole, 12 ml binding buffer containing 20 mM imidazole and 12 ml binding buffer containing 40 mM imidazole. Then, rinsed the resin with 0.9 ml binding buffer containing 500 mM imidazole to elute TF2 recombinant protein for five times. The purified TF2 recombinant protein was dialyzed against dialysis buffer (20 mM Tris·Cl, 150 mM NaCl, pH7.4) by Dialysis tubing cellulose membrane (Sigma-Aldrich) overnight at 4°C. The next day, TF2 protein was transferred to a 1.5 ml

tube and stored at 4°C for future functional analysis.

2.13 Silver staining

The solutions for silver staining were prepared before use. SDS-PAGE gel was soaked in 125 ml fixing solution (40% ethanol, 10% acetic acid) for 30 minutes with gentle agitation and then decanted the solution. Soaked the gel in 125 ml sensitizing solution (30% ethanol, 1.25% w/v glutardialdehyde, 2% w/v sodium thiosulphate, 68% w/v sodium acetate) for 30 minutes with gentle agitation and then decanted the solution. The gel was washed 3 times with distilled water for 5 minutes. Next, the gel was soaked in 125 ml reacting solution (2.5% w/v silver nitrate, 0.148% w/v formaldehyde) for 20 minutes with gentle agitation and then decanted the solution. The gel was washed 2 times with distilled water for 1 minute. Soaked the gel in 125 ml of fresh developing solution (25% w/v sodium carbonate, 0.148% w/v formaldehyde) for 2-5 minutes, agitated slowly until band appears, do not over developed because the reaction would be continued. Decanted the solution, then soaked the gel in 125 ml of the stopping solution (1.46% w/v EDTA- $\text{Na}_2 \cdot 2\text{H}_2\text{O}$) for 10 minutes, agitated slowly. The gel was washed 3 times with distilled water for 5 minutes.

2.14 HUVEC isolation

HUVEC were isolated from fresh umbilical cord without any clamp marks and needle holes. The umbilical vein was washed with 10 ml PBS twice slowly to remove the residual blood and clots, then filled with collagenase and fixed by hemostats, then incubated in 37°C PBS for 7 minutes, vein endothelial cell sheets would released from the inner lining of the blood vessel during this time. Removed the hemostat and collected the effluent into a sterile 50 ml tube, the enzyme activity was terminated by DMEM and then centrifuged at $200 \times g$ for 5 minutes. Cell pellet was resuspended into completed EC growth medium and seeded in a 0.2% gelatin-coated flask, the cells were cultured the next day.

2.15 Functional assay of TF2 recombinant protein

1. Cell migration assay

The migration ability of HUVEC was assessed by BD Falcon™ Cell Culture Inserts (BD Biosciences) with an 8 µm mean pore size membrane. The membrane was coated with 0.2% gelatin at 37°C for 60 minutes. HUVEC were trypsinized and suspended in supplemented M200 medium to final concentration of HUVEC was 5×10^5 cells/ml, 200 µl of cells were loaded into each upper well, TF2-His was added in lower chamber from 0.33 nM to 33 nM. The chambers were incubated for 5 hours in a humidified 5% CO₂ incubator at 37°C. Fixed the cells with methanol for 15 minutes and then removed the cells from the upper well by using cotton swabs, the membrane was washed with PBS twice and stained with Hoechst 33342 (invitrogen) for 15 minutes at room temperature. The stained cells were imaged by CCD and five random fields were counted at 100 × magnification, each condition was performed in triplicate and repeated at least three times.

2. Tube-like formation assay

Basement membrane matrix (BD Biosciences) was thawed at 4°C, 48-well plate was pre-cooled at -20°C and placed on ice, coated 100 µl of basement membrane matrix into each well and allowed to solidify at 37°C for 30 minutes. HUVEC were trypsinized and suspended in supplemented M200 medium, then seeded on matrigel-coated well to final concentrations from 0.33 nM to 33 nM for 4 hours in a humidified 5% CO₂ incubator at 37°C. Images were captured and the tube length was quantified at 100 × magnification, each concentration was assayed in triplicate and repeated at least three times.

3. Results

3.1 Analysis of the putative human THSD7A

The peptide sequence of putative human THSD7A precursor encodes a polypeptide chain of 1657 amino acids with a calculated molecular weight of 180 kDa approximately. Domain analysis of THSD7A was performed on PROSITE web servers which showed that THSD7A contains eleven thrombospondin type-1 (TSP-1) repeats, one CD36 binding site (CSKTCG) and one RGD motif, in addition to this, THSD7A was predicted as a transmembrane protein which comprised a signal peptide (Figure 1). This result revealed that the function of THSD7A may be an extracellular signal transmitter which involved in cell-to-cell or ECM-to-cell interactions. The post-translational modification of THSD7A was predicted from UniProtKB / Swiss-Prot which displayed that there are fourteen potential N-glycosylation sites (Figure 1). Since almost all extracellular proteins are N-glycosylated, this prediction showed highly consistent with the domain analysis result, and it further implicated the fundamental interactions of THSD7A may be based on carbohydrate-carbohydrate, carbohydrate-protein or glycoprotein-glycoprotein interactions.

3.2 Human THSD7A is a secreted protein

From previous studies, we demonstrated that *Thsd7a* transcript was detected in the central neuron system of developing zebrafish (unpublished data from Chieh-Huei Wang). However, if we down regulated THSD7A expression in zebrafish embryo, it impaired the angiogenesis of ISVs (unpublished data from Chieh-Huei Wang). Since angiogenesis was modulated by many kinds of angiogenic factors, it appeared that THSD7A may likely be a secretory protein which was derived from the neuron cells and acted as an angiogenic protein to regulate the morphogenesis of ISVs. The predicted human *Thsd7a* was found to encode a protein with a putative signal peptide at the N-terminus. The subcellular localization of THSD7A was thus predicted using this signal peptide sequence on several web servers. The consensus outcome of the predictions indicated that THSD7A is most likely to be an extracellular protein (Presidi, PENCE PA-SUB and BaCelLo), with lower

possibility to be a mitochondrial protein (iPSORT and TargetP 1.1), or on the periplasmic (CELLO) or plasma membrane (MultiLoc). For these reasons, it was hypothesized that THSD7A may be a secretory protein. To investigate that, human THSD7A protein was constructed (Figure 2a) and expressed by means of transient expression in human embryonic kidney 293T cells. Two days after transfection, the culture medium and cell lysate were collected and analyzed by western blotting. Different antibodies that can recognize different epitopes through the THSD7A peptide sequence were used in this experiment. Figure 2 shows the localization pattern of THSD7A in different fractions of HEK cell lysates. The result of α -tubulin was used to distinguish medium from pellet (Figure 2b, right panel; left panel). It was found that when THSD7A protein expression was up-regulated, a 260 kDa band can be recognized in lysates of cell pellets by IDS2 (Inter-TSP1-Domain-2), IDS9 and CTE antibodies (Figure 2b, left panel, lane 2). On the contrary, no band can be detected in the lysates of cell pellets of vector-only control (Figure 2b, left panel, lane 1). On the other hand, THSD7A can be identified in culture medium by IDS2 and IDS9 antibodies but not by CTE when it was over-expressed in HEK cells (Figure 2b, right panel, lane 2), but in the vector-only control of culture medium, THSD7A was non-detectable (Figure 2b, right panel, lane 1). It seems that there is another secreted form of THSD7A which contains IDS2 and IDS9 recognition sites. Furthermore, based on the previous bioinformatic analysis (Figure 1), we have derived that the hypothetical THSD7A is a membrane protein, whose transmembrane domain is located in the C-terminus of THSD7A, and both IDS2 and IDS9 recognition sites are located in the predicted extracellular domain. In other words, the extracytosolic domain of THSD7A may be cleaved and secreted (Figure 2c).

3.3 Human THSD7A is N-glycosylated in mammalian cells

The predicted core polypeptide of THSD7A precursor is 180 kDa, but the recognized THSD7A in cell lysate is approximately 260 kDa, which means that the post-translational modification of THSD7A would need to add around 80 kDa onto its molecular mass. Since the predicted THSD7A has many potential N-glycosylation sites (Figure 1), I used tunicamycin, which

is an effective N-glycosylation blocker, to examine whether the higher protein molecular weight was caused by N-glycosylation. To investigate that, human THSD7A protein was expressed by means of transient expression in human embryonic kidney 293T cells. Two days after transfection, the culture medium and cell lysate were collected and analyzed by western blotting. IDS2, IDS9 and CTE antibodies were used to detect THSD7A in this experiment. The data showed that the treatment of tunicamycin declined approximately 50 kDa from 260 kDa THSD7A in cell lysate when recognized by all three antibodies (Figure 2b, left panel, lane 3). It demonstrated that THSD7A is indeed a N-glycosylated protein when expressed by mammalian cells, but the molecular weight of de-glycosylated THSD7A cannot be predicted precisely based on this Western blot result. On the other hand, no band can be detected in the culture medium fraction (Figure 2b, right panel, lane 3).

3.4 Generation of TF2 recombinant virus

From previous studies, we demonstrated that THSD7A may modulate leading front of growing vessel via controlling EC migration. Additionally, it was found that the extracellular domain of THSD7A can be secreted (Figure 2b, left panel, lane 3). Thus, THSD7A was expressed to identify its functional domain. Previous study indicated that the expression level of E. coli, insect and mammalian cell expressed full-length THSD7A was low. So, the extracellular domain was truncated into three fragments: TF1 (truncated fragment 1), TF2 and TF3 without destroy the complete TSRs (Figure 3a). In this thesis study, I chose TF2 to identify the functional domain of THSD7A because of its abundance of TSRs. TF2 is located on the putative full-length THSD7A nucleotide sequence from 1254 bp to 2877 bp, namely the amino acid sequence from 418 to 959. The hypothetical molecular weight of TF2 is around 60 kDa. Since THSD7A was proved to be a N-glycosylated protein, baculovirus expression system was selected to over-express TF2 in the insect cell system. This expression system not only provides mammalian-like post-translational modification of protein N-glycosylation but also enables a more economical approach to produce recombinant proteins. Composite sequence of GST-TF2, TF2-His and GST only control were sub-cloned into pMelBac

expression vector which containing a N-terminal insect honeybee melittin secretion signal peptide (HBM) and P_{PH} (the promoter of polyhedrin) successfully (Figure 3b). Three plasmids were then recombined with *Autographa californica* Multiple Nucleopolyhedrosis Virus (AcMNPV) DNA before co-transfection into the Sf9 insect cells subsequently. The result from a time lapse experiment showed cell growth and cell morphology change of Sf9 after co-transfection. Briefly, the cessation of cell growth can be observed at the initial 5 days (Figure 4b), some cells (wild-type virus infected cells) will contain viral occlusions which appear as refractive crystals in Sf9 were formed about 5-7 days (Figure 4c-4d, black arrow), virus infected cell will lysis and detached from culture dish, some cell will release the occluded viruses after 7-10 days of virus infection afterward (Figure 4d, right panel, arrowhead). As a result, the characteristics described above represented a successful transfection. Recombinant virus clones were then separated from WT virus clones by plaque assay, in which 10⁻¹ to 10⁻³ dilutions of virus mixture and no virus control were tested in my experiment. Round-shaped Sf9 (supplement 3a) became flat-shaped (supplement 3b) after 3-4 days of culture in a regular procedures of all conditions. Then, the blue plaques appeared during 5-10 days post-infection, which indicated a successful infection (supplement 3c, upper panel). Additionally, when I compared the Sf9 morphologies between control and infected cells in enlarged view, control cells showed no significant changed from 3-4 days post-infection (supplement 3c, middle panel), but the infected cells marked with the blue plaques had larger cell bodies with vacuole-like structures (supplement 3c, bottom panel). Plates with 10⁻² virus dilution showed non-crowded distribution but sufficient quantity of blue plaques in all three types of recombinant virus was found in my experiments. These plates were then used to harvest recombinant virus clones (supplement 3c, upper panel). The resulting virus DNA was verified by polymerase chain reaction to distinguished recombinant virus from the unwanted WT virus. In this PCR diagnosis analysis, three types of DNA patterns which represented three kinds of vectors can be seen in the result of GST-TF2 virus preparation, which includes (1) recombinant virus only (Figure 5b, lane 3), (2) WT virus only (Figure 5b, lane 2), and (3) WT virus contaminated of recombinant virus (Figure 5b, lane 1). On the other hand, the diagnosis analysis of GST and TF2-His viral constructs showed

all recombinant virus patterns (Figure 5c-5d). It appeared that because the distribution of GST-TF2 plaques were too crowded in the plate, the clone picking during the harvest procedure was not able to prevent WT virus contamination.

3.5 Expression and identification of TF2 recombinant protein in High5 cells

To demonstrate TF2 recombinant protein expression, insect cell lysates and mediums were collected every 12 hours during 1 to 4 post-infection days under MOI=1, MOI=3, MOI=5 or MOI=10. The IDS4 antibody was then used in Western blot to analyze these samples. Different parameters were tested to optimized TF2 recombinant protein production. The finding are described as follow: GST-TF2 expressed by suspended High5 insect cells, which cultured in Express 5 medium, showed high expression level in cell lysates. However, it was non-secreted into culture medium even though it has an engineered secreted signal peptide (Figure 6a). This result is the same as then GST-TF2 expressed by the suspended Sf9 insect cells, which was cultured in TNM-FH medium (Figure 6b).

I then compared the GST-TF2 expression level in the High5 and Sf9 cell lines. Even though the MOI for the Sf9 cells was increased to MOI=3 and MOI=5, the expression level of High5 under MOI=1 and MOI=3 performed better than the Sf9 cells. Subsequently, I tried another culture medium, EX-CELL 405, for High5 insect cells in the attempt to solve the protein non-secreted problem. Even though numerous researchers have expressed a wide range of recombinant proteins by EX-CELL 405 medium successfully, I could not detect target protein expression in either cell lysate or culture medium after the cells was infected at MOI=3. A positive control were used to validate the experimental procedure was also shown in the Western blotting (Figure 6c). This result demonstrated that the property of culture medium have a great influence in protein expression even though the same type of insect cell line was used. In my experiment, Express 5 medium performed better than the EX-CELL 405 medium for GST-TF2 expression in the High5 insect cell line. Finally, the protein expression of His tagged TF2 was examined. Conditions with MOI=1, MOI=5 and MOI=10 were used, and the resulting production samples were harvested at 0, 24, 36, 48, 60 and 72

hours post-infection (hpi). Secreted form of TF2-His can be identified in culture medium (Figure 6d, right panels). This result indicated that the key factor in regulating the cell's protein secretion ability depends on the property of the fusion tag on the TF2.

In generally, the causes of non-secreted GST recombinant protein could be: (1) MOI used for viral infection was too high, (2) insect cell line selection problems, (3) excessive expression of large protein induced formation of insoluble inclusion bodies, and finally, (4) the characteristic of recombinant protein itself. While various combinations of MOIs, cell lines and culture conditions were tested in this study, we have yet to produce solid improvement in GST-TF2 secretion. On the other hand, GST is a relative larger protein (26 kDa, 218 a.a.) compared with the 6 × His tag. Therefore, the non-secretion may be caused by the high molecular weight and fusion protein folding of the recombinant GST-TF2 presented in this study.

Next, the expression level of TF2-His at different MOIs: MOI=1, MOI=5 and MOI=10 throughout different culturing time points in High5 insect cells were compared. The expression of TF2-His was increased both in cell lysate and supernatant along the time until 72 hpi in all MOIs (Figure 6d). MOI=5 showed highest protein expression level both in cell lysate and supernatant (Figure 6d, middle panels) as compared to MOI=1 (Figure 6d, upper panels) and MOI=10 (Figure 6d, bottom panels). This data revealed the fact that the higher virus titer did not lead to higher protein expression level either in cell lysate or culture medium. There is one possible explanation for this observation: the infection titer may be too high so the High5 cells were lysed before they could generate enough recombinant protein, and the released protease may degrade the secreted TF2 protein.

Taken together, these results demonstrated the optimized secreted protein expression condition was to express TF2 recombinant protein in High5 insect cells and cultured with Express 5 growth medium by using TF2-His recombinant virus. And the insect cells shall be infected at MOI=5 for 72 hours (Figure 6d, right middle panel, lane 6).

Interestingly, I noticed that there are two forms of GST-TF2 that can be recognized in cell pellets which weighted approximately 90 kDa and 110 kDa (Figure 6a-6b), it was similar to the

pattern of TF2-His which weighted approximately 62 kDa and 70 kDa (Figure 7a, lane 1). The 90 kDa form of GST-TF2 and the 62 kDa form of TF2-His were the predicted precursor forms, but the nature of the other two forms of TH2-His were still unknown. I assumed these products could be different cleavage forms or different post-translational modification forms of TF2 recombinant protein. Furthermore, only the 70 kDa form of TF2-His can be secreted into culture medium (Figure 7a, lane 2). To check whether the secretion was caused by protein glycosylation, I further examined the different forms of TF2 in the next section.

3.6 TF2 recombinant protein has an N-glycosylated form in High5 cells

In this section, I again used tunicamycin, which is known to block N-glycosylation, to see the effect on the molecular mass of the 70 kDa isoform of TF2-His. Cell pellets were collected and analyzed by Western blot. The result showed that the expression level of 70 kDa form of TF2-His was declined drastically (Figure 7b, lane 2). As expected, it demonstrated that TF2-His can be modified by N-glycosylation in the insect cells, and the molecular weight was raised from 62 kDa (Figure 7b, lane 2) to 70 kDa (Figure 7a, lane 2). For the same reason, GST-TF2 may be modified by N-glycosylation to raise the molecular weight from 90 kDa to 110 kDa. In conclusion, there are indeed glycosylated form and unglycosylated precursor form of TF2 recombinant protein expressed in the insect cells. This is consistent with the post-translational modification findings in the mammalian cell system (Figure 2).

In conclusion, the predicted full-length THSD7A may contain fourteen N-glycosylation sites. The molecular weight shift caused by N-glycosylation is approximately 50 kDa. On the other hand, TF2 was predicted to contain four N-glycosylation sites, the molecular shift caused by N-glycosylation is approximately 10 kDa.

Additionally, although I have observed that there are two forms of TF2-His existed in the High5 cell pellets (62 kDa precursor form and 70 kDa glycosylated form; Figure 7), the expression ratio of these two forms from different batch of expression was different. Approximately 59% of TF2-His expression showed that the expression level of glycosylated form was more prominent

than the precursor form, but the other 41% showed that the expression level of both two forms were the same. Two trials of expression preparation were compared. In this experiment, I used the same batch of TF2-His virus for protein expression, but the results were different (Figure 6d, supplement 4). In the first expression trial (MOI=1, MOI=5 and MOI=10), the expression level of glycosylated form was higher than the precursor form (Figure 6d). But in the second expression trial (MOI=16.6, MOI=83 and MOI=166), the expression level of two forms of TF2-His were the same, and they have poor secretion ability and lower cell survival rate than the first trial (supplement 4). In my opinion, the possible explanation for this inconsistency may be the High5 cells were not “healthy” enough at the time of infection in the second trial. For example, the passage numbers of High5 cells is too high. Second, the cells are not in the logarithmic phase thus the cell viability is not high enough. Third, the insect cell suspension culture was not handled quickly enough. All these factors may disturb protein expression, and result in lower ratio of secreted form in culture medium.

3.7 Purification of TF2 recombinant protein by nickel affinity chromatography

Since TF2 recombinant protein was His-tagged, TF2-His was purified by nickel affinity chromatography. In a standard procedure, crude protein sample was incubated with nickel conjugated beads first. The beads were then washed by various low concentrations of imidazole to eliminate non-specific proteins but the target protein may lose at the same time. After several wash, the beads were treated with high concentration of imidazole to elute TF2-His. Higher concentration of imidazole led to high yield of TF2-His but low purity. For this reasons, various concentrations and volumes of wash and elution buffer were tested because different parameters will lead to different levels of protein purity. In this section, an optimized protocol for TF2-His recombinant protein purification was established after various conditions were tried. At first, 12 ml of wash buffer with different salt concentration were used to wash the beads: 0 mM, 20 mM, 40 mM, 60 mM. Then, I eluted the proteins with 250 mM elution buffer. These procedures were analyzed by silver stain and Western blotting using same volumes of protein sample. My results showed that there were impurities in 250 mM elution except for the 70 kDa TF2-His (data not shown). Then, I

proceeded to increase the wash concentration from 60 mM to 75 mM. But the result indicated that TF2-His was lost in this concentration. Thus, the change of washing buffer cannot eliminate the impurities significantly (data not shown). Next, I tried to decrease the concentration of elution buffer from 250 mM to 125 mM, and eliminated the 75 mM wash step. My result revealed that the total yield of TF2-His was decreased, but the level of impurities was still the same (data not shown). Since the impurities was difficult to eliminate, I switched to focus on producing higher quantities of TF2-His. 500 mM imidazole was tried in the elution step and the wash volumes of 40 mM washing buffer was increased from 12 ml to 20 ml. Finally, although TF2-His still contains the impurities, the purity of TF2-His became higher because 500 mM imidazole remarkably increased the concentration of TF2-His but not the impurities (Figure 8, panel a). This result was checked by Western blotting at the same time. Concentrations of TF2-His were very low at initial preparation steps (Figure 8b, lane 1-4) and wash steps (Figure 8b, lane 5-10). However, I was able to concentrate the TF2-His later (Figure 8b, lane 12-13). As a result, an optimized protocol was established to purify the 70 kDa isoform of TF2-His expressed by insect cells.

3.8 Validation of the purified TF2 recombinant protein

After purification, the purified TF2-His recombinant protein was further validated by anti-IDS4 antibody, anti-6xHis antibody and ConA. Meanwhile, the purity of the TF2-His was verified by using silver stain analysis. Two preparations, crude (protein sample from culture medium before purification) and elute (protein sample after purification) with the same volume, were analyzed and compared. The results showed that proteins recognized by anti-IDS4 and anti-6xHis antibodies displayed identical patterns. There is only a 70 kDa band that can be detected in the crude sample (Figure 9b-c, lane 1). However, three bands can be recognized in purified TF2-His, which includes the 70 kDa major band, and additional 62 kDa and 42 kDa minor bands (Figure 9b-c, lane 2). On the other hand, when ConA (which can recognize mannose of N-glycosylation) was used to detected TF2-His, only the 70 kDa band can be detected in the elute preparation (Figure 9d, lane 2). Moreover, multiple bands can be recognized in crude sample (Figure 9d, lane 1) by ConA.

Finally, silver stain analysis of elute preparation showed multiple bands, including the dominant 70 kDa band and the weaker 62 kDa and 42 kDa (Figure 9a, lane 2) bands. These bands could also be detected in the crude sample as expected (Figure 9a, lane 1). Taken together, these results further confirmed the identity of the purified protein sample and showed that the purified protein was predominant the 70 kDa glycosylated form of TF2-His.

3.9 TF2 recombinant protein does not affect HUVEC migration and tube formation

After the truncated form of THSD7A, TF2-His, was purified, it was used in functional assays to explore whether it has an effect on endothelial directed migration and tube formation. TF2-His recombinant protein at various concentrations: 25 ng/ml (0.33 nM), 250 ng/ml (3.33 nM), 1250 ng/ml (16.67 nM), 2500 ng/ml (33.33 nM) were tested in the angiogenic functional assays. First, transwell assay was used to analyze HUVEC migration rate in the presence of TF2-His. VEGF was used as positive control in this assay (data not shown). After 5.5 hours of incubation, migrated endothelial cells were stained and photographed. Quantification of migrated cells was presented as relative values to the control cells treated with PBS only. The result showed that there was no significant difference in migrated cell numbers at all concentrations that were tested (Figure 10b), even though 2500 ng/ml TF2-His was used (Figure 10a). Second, the endothelial tube-formation assay was performed. Quantification of total tube length was presented as relative values to the control cells which treated with PBS only. After 4 hours of incubation period, the tube-like structures on the plate were photographed. The results showed no significant differences at various concentrations of TF2-His (Figure 11a-b). In summary, the addition of recombinant TF2-His protein did not affect HUVEC migration or tube formation.

4. Discussion

In retrospect, the initial goal of this thesis study is to investigate the functional domain of THSD7A involved in angiogenesis of endothelial cells. Thus, I expressed and purified truncated fragment 2 of THSD7A to perform the angiogenic functional analysis. First, a protocol to express and purify TF2 recombinant protein was established. Through trial and error, the conditions to generate the target protein with most favorable purity and yield were identified. Nevertheless, based on the silver stain and Western blotting analysis, we can still identify unexpected proteins other than TF2-His (Figure 9a-c, lane 2) in the preparation. Although the target protein purity could be further improved by changing the imidazole concentration during purification, the impurities could not be completely removed.

How to improve the protein purity? The solution may arise after knowing the source and identify of the impurities. I suspect that the impurities, 62 kDa and 42 kDa bands, may be derived from the non-secreted portion of insect cells (supplement 5). This is supported by the finding that the 62 kDa band may be the unglycosylated form of TF2-His. On the other hand, the 42 kDa band may arise from degraded protein attacked by active proteases released from the dead/damaged cells.

Since the purity of TF2 recombinant protein varied from one preparation to the other, there could still be many different factors affecting the outcome on purity and yield. One possible cause could be the viral titer. Some researchers have shown that although lower infection titer may cost more time to harvest, it could generate high yield of the target protein and reduced cell death during expression process⁷¹⁻⁷³. So, decreasing infection titer shall be tried in the future.

After we established the baculovirus expression system for THSD7A-TF2 recombinant protein, we performed HUVEC-based angiogenic assays to study its function. Various concentrations of TF2 from 0.33 nM to 33.33 nM were used in this study. Unfortunately, our results demonstrated that TF2 can not affect endothelial cell migration or tube-like formation ability.

There are some possible explanations for such negative result:

1. The insect cell lacks the ability to generate hybrid and complex types of N-glycans:

Although many researchers have successfully purified recombinant protein with biological activities using the baculovirus/insect cell expression system⁶⁵⁻⁶⁸, there are still some restrictions of this system due to the N-glycosylation differences as compared to the mammalian system. The three main types of N-glycosylation in mammalian cells are: oligomannose, complex and hybrid. In previous studies, researchers have indicated the N-glycans have significant effect on EC migration during signaling transduction steps in angiogenesis. It was shown that the inhibition of hybrid and complex types of N-linked oligosaccharides could also disrupt the formation of capillary tubes *in vitro*^{64,74}. The negative results could hence arise from the deficient capacity of the insect cells to process N-linked glycans into hybrid or complex type structures.

Furthermore, a growing body of evidences suggests that the precise addition of carbohydrates to proteins will promote protein stability, folding, secretion and affect both its functional activity⁴⁵⁻⁵¹. If the glycosylation of our target protein is incorrect, it may lead to loss of function during the protein expression and purification process. In this study, we have suggested that the 42 kDa impurities may be caused by the low stability of TF2-His protein. Additionally, when we over-expressed the TF2 recombinant protein in the insect system, we always observed a large portion of glycosylated proteins cannot be secreted (Figure 7a). This is supported by the tunicamycin treatment experiment, in which I found the deglycosylated THSD7A cannot be secreted into the culture medium (Figure 2a, right panel, lane 3). Thus, the level of the precise protein glycosylation can influence the ratio of the secreted portion to the non-secreted portion of the target protein⁶³. Furthermore, incorrect glycosylation will lead to 80% improper folding⁴⁸ and disturb the secretion of protein, too^{52,53}. In summary, the baculovirus expressed system may not be a suitable system for TF2-His if glycosylation is critical for its activity.

2. HUVEC-based assay system is deficient in VEGF (vascular endothelial growth factor):

THSD7A was found to contain many TSRs. Other TSR-containing protein is found to inhibit VEGF-induced angiogenesis through the inhibition of PI3K signaling, which leads to the reduced EC survival and EC migration in HUVEC^{75,76}. In this study, although we performed functional assays using the standard protocol used by many angiogenesis studies, the true effect of THSD7A

may only be seen when other activation factors were presented in the assay.

3. Structure of the functional domain may be destructed:

THSD7A is a novel protein without proven 3D structure. It is also a very large protein, which is difficult to predict its 3D structure by past bioinformatic tools at the start of this thesis study. In our original experimental design, THSD7A was divided into three parts without the exact knowledge on its protein structure. This approach posed a potential flaw since we cannot know whether the functional region of TF2 is folded the same as the native form or not. Fortunately, we can now predict the 3D structure of THSD7A on a newly developed tool called I-TASSER. I-TASSER (<http://zhang.bioinformatics.ku.edu/I-TASSER/>) is a web tool which enables structure prediction of high molecular weight protein. The predicted native THSD7A showed like a globular structure (supplement 6a), three truncated fragments was labeled on it: TF1 (green), TF2 (red) and TF3 (white). Six TSRs on TF2 was labeled on native THSD7A, too (supplement 6b). The structure of TF2 only was also predicted and labeled six TSRs on it (supplement 6c). Compared with native and truncated form of TF2, the TSRs were abnormality in truncated form. It showed that the structure of TF2 was destructed after truncation, thus may destroy the possible functional domain on it.

4. The functional domain of TSR is not located in TF2:

Since TF2 showed no activity in our angiogenic assay, a simple explanation may be that the functional domain of THSD7A is not located within TF2. It may hide in other region of the protein. The previous structure predicted result also showed that THSD7A contain a predicted protein binding domain on TF3 (supplement 7a), it comprise a complete TSR (supplement 7b). This TSR was predicted to contain a CD36 binding motif which may bind to cell surface receptor CD36. This information can be very useful in our future approach to design the expression vector used to re-express THSD7A.

The problem on recombinant THSD7A expression and purification will be solved eventually. This thesis study has nonetheless established a protocol for the expression and purification of mammalian protein for our research group. In the future, the laboratory could always make further

improvement based on the protocol. Future researchers should try to figure out whether insect cells can generate hybrid and complex glycosylation or not. They could also try to use advanced tools to predict the functional domain of THSD7A and to express the recombinant protein in mammalian cell system simultaneously.



5. References

1. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
2. Folkman, J. Endogenous angiogenesis inhibitors. *APMIS* **112**, 496-507 (2004).
3. Nyberg, P., Xie, L. & Kalluri, R. Endogenous inhibitors of angiogenesis. *Cancer Res* **65**, 3967-3979 (2005).
4. Carmeliet, P. Angiogenesis in life, disease and medicine. *Nature* **438**, 932-936 (2005).
5. Folkman, J. Endogenous inhibitors of angiogenesis. *Harvey Lect* **92**, 65-82 (1996).
6. Wong, C.G., Rich, K.A., Liaw, L.H., Hsu, H.T. & Berns, M.W. Intravitreal VEGF and bFGF produce florid retinal neovascularization and hemorrhage in the rabbit. *Curr Eye Res* **22**, 140-147 (2001).
7. Ferrara, N., Mass, R.D., Campa, C. & Kim, R. Targeting VEGF-A to treat cancer and age-related macular degeneration. *Annu Rev Med* **58**, 491-504 (2007).
8. Ferrara, N., Gerber, H.P. & LeCouter, J. The biology of VEGF and its receptors. *Nat Med* **9**, 669-676 (2003).
9. Yancopoulos, G.D., *et al.* Vascular-specific growth factors and blood vessel formation. *Nature* **407**, 242-248 (2000).
10. Ferrara, N. VEGF and the quest for tumour angiogenesis factors. *Nat Rev Cancer* **2**, 795-803 (2002).
11. Folkman, J. & Kalluri, R. Cancer without disease. *Nature* **427**, 787 (2004).
12. Matsumoto, T. & Claesson-Welsh, L. VEGF receptor signal transduction. *Sci STKE* **2001**, re21 (2001).
13. Shweiki, D., Itin, A., Soffer, D. & Keshet, E. Vascular endothelial growth factor induced by hypoxia may mediate hypoxia-initiated angiogenesis. *Nature* **359**, 843-845 (1992).
14. Dameron, K.M., Volpert, O.V., Tainsky, M.A. & Bouck, N. Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1. *Science* **265**, 1582-1584 (1994).
15. Guo, N., Kruttsch, H.C., Inman, J.K. & Roberts, D.D. Thrombospondin 1 and type I repeat peptides of thrombospondin 1 specifically induce apoptosis of endothelial cells. *Cancer Res* **57**, 1735-1742 (1997).
16. Hsu, S.C., *et al.* Inhibition of angiogenesis in human glioblastomas by chromosome 10 induction of thrombospondin-1. *Cancer Res* **56**, 5684-5691 (1996).
17. Nor, J.E., *et al.* Thrombospondin-1 induces endothelial cell apoptosis and inhibits angiogenesis by activating the caspase death pathway. *J Vasc Res* **37**, 209-218 (2000).
18. Jimenez, B., *et al.* Signals leading to apoptosis-dependent inhibition of neovascularization by thrombospondin-1. *Nat Med* **6**, 41-48 (2000).
19. Wang, Y., Wang, S. & Sheibani, N. Enhanced proangiogenic signaling in thrombospondin-1-deficient retinal endothelial cells. *Microvasc Res* **71**, 143-151 (2006).
20. Good, D.J., *et al.* A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin. *Proc Natl Acad Sci U S A* **87**, 6624-6628 (1990).

21. Tolsma, S.S., *et al.* Peptides derived from two separate domains of the matrix protein thrombospondin-1 have anti-angiogenic activity. *J Cell Biol* **122**, 497-511 (1993).
22. Gerhardt, H., *et al.* VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia. *J Cell Biol* **161**, 1163-1177 (2003).
23. Gerhardt, H., *et al.* Neuropilin-1 is required for endothelial tip cell guidance in the developing central nervous system. *Dev Dyn* **231**, 503-509 (2004).
24. Suchting, S., *et al.* The Notch ligand Delta-like 4 negatively regulates endothelial tip cell formation and vessel branching. *Proc Natl Acad Sci U S A* **104**, 3225-3230 (2007).
25. Karagiannis, E.D. & Popel, A.S. Anti-angiogenic peptides identified in thrombospondin type I domains. *Biochem Biophys Res Commun* **359**, 63-69 (2007).
26. Huwiler, K.G., Vestling, M.M., Annis, D.S. & Mosher, D.F. Biophysical characterization, including disulfide bond assignments, of the anti-angiogenic type 1 domains of human thrombospondin-1. *Biochemistry* **41**, 14329-14339 (2002).
27. Sharghi-Namini, S., *et al.* The first but not the second thrombospondin type 1 repeat of ADAMTS5 functions as an angiogenesis inhibitor. *Biochem Biophys Res Commun* **371**, 215-219 (2008).
28. Silverstein, R.L. & Febbraio, M. CD36-TSP-HRGP interactions in the regulation of angiogenesis. *Curr Pharm Des* **13**, 3559-3567 (2007).
29. Simantov, R. & Silverstein, R.L. CD36: a critical anti-angiogenic receptor. *Front Biosci* **8**, s874-882 (2003).
30. Adams, J.C. & Tucker, R.P. The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. *Dev Dyn* **218**, 280-299 (2000).
31. Bork, P. The modular architecture of a new family of growth regulators related to connective tissue growth factor. *FEBS Lett* **327**, 125-130 (1993).
32. Karagiannis, E.D. & Popel, A.S. Peptides derived from type I thrombospondin repeat-containing proteins of the CCN family inhibit proliferation and migration of endothelial cells. *Int J Biochem Cell Biol* **39**, 2314-2323 (2007).
33. Silverstein, R.L. The face of TSR revealed: an extracellular signaling domain is exposed. *J Cell Biol* **159**, 203-206 (2002).
34. Ji, W.R., *et al.* Characterization of kringle domains of angiostatin as antagonists of endothelial cell migration, an important process in angiogenesis. *FASEB J* **12**, 1731-1738 (1998).
35. Yamaguchi, N., *et al.* Endostatin inhibits VEGF-induced endothelial cell migration and tumor growth independently of zinc binding. *EMBO J* **18**, 4414-4423 (1999).
36. Wang, C.H., *et al.* Thrombospondin type I domain containing 7A (THSD7A) mediates endothelial cell migration and tube formation. *J Cell Physiol* **222**, 685-694 (2010).
37. Giancotti, F.G. Complexity and specificity of integrin signalling. *Nat Cell Biol* **2**, E13-14 (2000).
38. Klemke, R.L., *et al.* Regulation of cell motility by mitogen-activated protein kinase. *J Cell Biol* **137**, 481-492 (1997).

39. Turner, C.E. Paxillin and focal adhesion signalling. *Nat Cell Biol* **2**, E231-236 (2000).
40. Hu, Y.L. & Chien, S. Dynamic motion of paxillin on actin filaments in living endothelial cells. *Biochem Biophys Res Commun* **357**, 871-876 (2007).
41. Carmeliet, P. & Tessier-Lavigne, M. Common mechanisms of nerve and blood vessel wiring. *Nature* **436**, 193-200 (2005).
42. Mukouyama, Y.S., Shin, D., Britsch, S., Taniguchi, M. & Anderson, D.J. Sensory nerves determine the pattern of arterial differentiation and blood vessel branching in the skin. *Cell* **109**, 693-705 (2002).
43. Zukowska, Z., Grant, D.S. & Lee, E.W. Neuropeptide Y: a novel mechanism for ischemic angiogenesis. *Trends Cardiovasc Med* **13**, 86-92 (2003).
44. Movafagh, S., Hobson, J.P., Spiegel, S., Kleinman, H.K. & Zukowska, Z. Neuropeptide Y induces migration, proliferation, and tube formation of endothelial cells bimodally via Y1, Y2, and Y5 receptors. *FASEB J* **20**, 1924-1926 (2006).
45. Pratap, J., Rajamohan, G. & Dikshit, K.L. Characteristics of glycosylated streptokinase secreted from *Pichia pastoris*: enhanced resistance of SK to proteolysis by glycosylation. *Appl Microbiol Biotechnol* **53**, 469-475 (2000).
46. Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369 (2001).
47. Parodi, A.J. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochem J* **348 Pt 1**, 1-13 (2000).
48. Mitra, N., Sinha, S., Ramya, T.N. & Surolia, A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem Sci* **31**, 156-163 (2006).
49. Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019-1049 (2004).
50. Lee, J., Park, J.S., Moon, J.Y., Kim, K.Y. & Moon, H.M. The influence of glycosylation on secretion, stability, and immunogenicity of recombinant HBV pre-S antigen synthesized in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* **303**, 427-432 (2003).
51. Zhao, Y.Y., *et al.* Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer Sci* **99**, 1304-1310 (2008).
52. Mosmann, T.R. & Williamson, A.R. Structural mutations in a mouse immunoglobulin light chain resulting in failure to be secreted. *Cell* **20**, 283-292 (1980).
53. Dobson, C.M. Protein folding and misfolding. *Nature* **426**, 884-890 (2003).
54. Schachner, M. & Martini, R. Glycans and the modulation of neural-recognition molecule function. *Trends Neurosci* **18**, 183-191 (1995).
55. Tsuda, T., Ikeda, Y. & Taniguchi, N. The Asn-420-linked sugar chain in human epidermal growth factor receptor suppresses ligand-independent spontaneous oligomerization. Possible role of a specific sugar chain in controllable receptor activation. *J Biol Chem* **275**, 21988-21994 (2000).
56. Guo, H.B., Lee, I., Kamar, M., Akiyama, S.K. & Pierce, M. Aberrant N-glycosylation of beta1 integrin causes reduced alpha5beta1 integrin clustering and stimulates cell migration.

Cancer Res **62**, 6837-6845 (2002).

57. Grewal, P.K., Holzfeind, P.J., Bittner, R.E. & Hewitt, J.E. Mutant glycosyltransferase and altered glycosylation of alpha-dystroglycan in the myodystrophy mouse. *Nat Genet* **28**, 151-154 (2001).
58. Guo, H.B., Lee, I., Bryan, B.T. & Pierce, M. Deletion of mouse embryo fibroblast N-acetylglucosaminyltransferase V stimulates alpha5beta1 integrin expression mediated by the protein kinase C signaling pathway. *J Biol Chem* **280**, 8332-8342 (2005).
59. Isaji, T., *et al.* Introduction of bisecting GlcNAc into integrin alpha5beta1 reduces ligand binding and down-regulates cell adhesion and cell migration. *J Biol Chem* **279**, 19747-19754 (2004).
60. Soderquist, A.M. & Carpenter, G. Glycosylation of the epidermal growth factor receptor in A-431 cells. The contribution of carbohydrate to receptor function. *J Biol Chem* **259**, 12586-12594 (1984).
61. Takahashi, M., Tsuda, T., Ikeda, Y., Honke, K. & Taniguchi, N. Role of N-glycans in growth factor signaling. *Glycoconj J* **20**, 207-212 (2004).
62. Yokoe, S., *et al.* The Asn418-linked N-glycan of ErbB3 plays a crucial role in preventing spontaneous heterodimerization and tumor promotion. *Cancer Res* **67**, 1935-1942 (2007).
63. Zhou, W. & Tsai, H.M. N-Glycans of ADAMTS13 modulate its secretion and von Willebrand factor cleaving activity. *Blood* **113**, 929-935 (2009).
64. Pili, R., *et al.* The alpha-glucosidase I inhibitor castanospermine alters endothelial cell glycosylation, prevents angiogenesis, and inhibits tumor growth. *Cancer Res* **55**, 2920-2926 (1995).
65. Smith, G.E., Summers, M.D. & Fraser, M.J. Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Mol Cell Biol* **3**, 2156-2165 (1983).
66. Giese, N., May-Siroff, M., LaRochelle, W.J., van Wyke Coelingh, K. & Aaronson, S.A. Expression and purification of biologically active v-sis/platelet-derived growth factor B protein by using a baculovirus vector system. *J Virol* **63**, 3080-3086 (1989).
67. Gillespie, L.S., Hillesland, K.K. & Knauer, D.J. Expression of biologically active human antithrombin III by recombinant baculovirus in *Spodoptera frugiperda* cells. *J Biol Chem* **266**, 3995-4001 (1991).
68. Smith, G.E., *et al.* Modification and secretion of human interleukin 2 produced in insect cells by a baculovirus expression vector. *Proc Natl Acad Sci U S A* **82**, 8404-8408 (1985).
69. Knebel, D., Lubbert, H. & Doerfler, W. The promoter of the late p10 gene in the insect nuclear polyhedrosis virus *Autographa californica*: activation by viral gene products and sensitivity to DNA methylation. *EMBO J* **4**, 1301-1306 (1985).
70. Hodgson, J. Expression systems: a user's guide. Emphasis has shifted from the vector construct to the host organism. *Biotechnology (N Y)* **11**, 887-893 (1993).
71. Gotoh, T., Miyazaki, Y., Chiba, K. & Kikuchi, K. Significant increase in recombinant protein production of a virus-infected Sf-9 insect cell culture of low MOI under low dissolved oxygen conditions. *J Biosci Bioeng* **94**, 426-433 (2002).

72. Hu, Y.C., Tsai, C.T., Chang, Y.J. & Huang, J.H. Enhancement and prolongation of baculovirus-mediated expression in mammalian cells: focuses on strategic infection and feeding. *Biotechnol Prog* **19**, 373-379 (2003).
73. Kim, J.S., *et al.* Production of recombinant polyhedra containing Cry1Ac fusion protein in insect cell lines. *J Microbiol Biotechnol* **17**, 739-744 (2007).
74. Fuster, M.M. & Esko, J.D. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* **5**, 526-542 (2005).
75. Calzada, M.J., *et al.* Alpha4beta1 integrin mediates selective endothelial cell responses to thrombospondins 1 and 2 in vitro and modulates angiogenesis in vivo. *Circ Res* **94**, 462-470 (2004).
76. Short, S.M., *et al.* Inhibition of endothelial cell migration by thrombospondin-1 type-1 repeats is mediated by beta1 integrins. *J Cell Biol* **168**, 643-653 (2005).



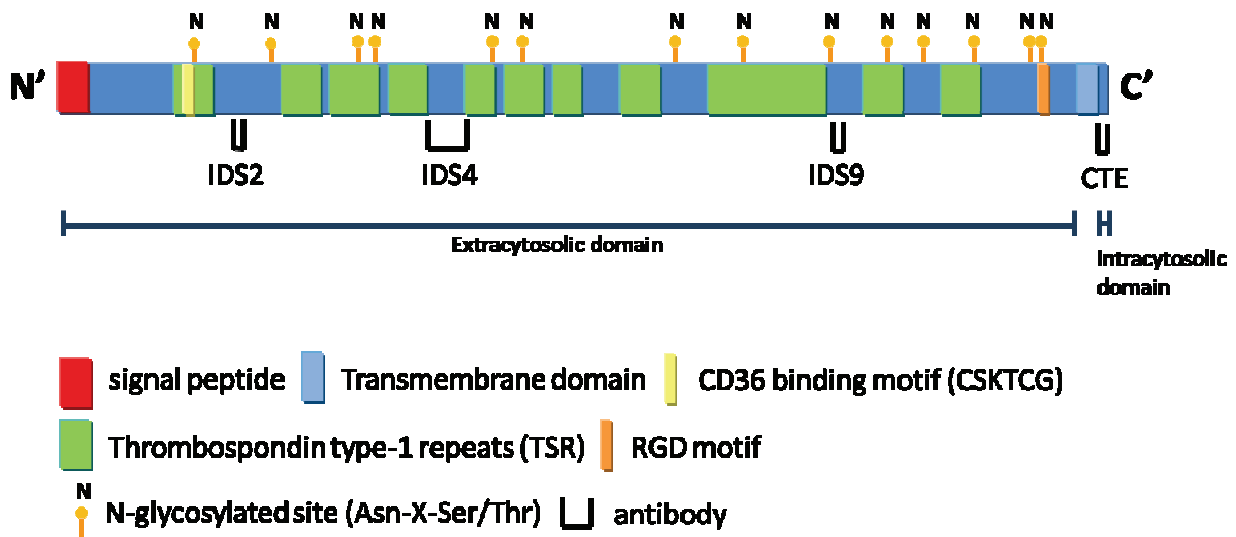


Figure 1: The predicted Thrombospondin Type1 Domain Containing 7A

Domain structures of human THSD7A were predicted from a deduced 1657-amino acid sequence (180kDa). It is a transmembrane protein which contains N-terminal signal peptide, eleven Thrombospondin type-1 repeats (TSR), CD36 binding motif (CSKTCG) and one RGD motif. Potential glycosylation sites are also displayed.

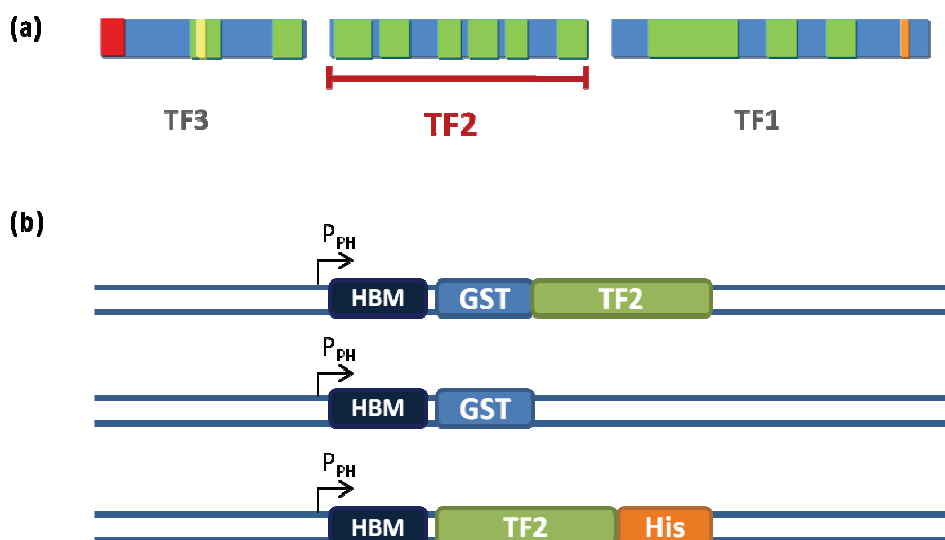


Figure 3: The construction of truncated human THSD7A expression vector

Schematic represents the extracellular domain of THSD7A was truncated into three fragments: TF1, TF2 and TF3 (a). Constructs of TF2 were cloned into pMelBac vector which includes a fused N-terminal insect honeybee melittin secretion signal peptide (HBM) and P_{PH} (the promoter of polyhedrin). Three constructs were performed including GST-tagged TF2, GST only as the control vector and His-tagged TF2 (b). The predicted molecular weight of core polypeptide of TF2 is 62kDa.

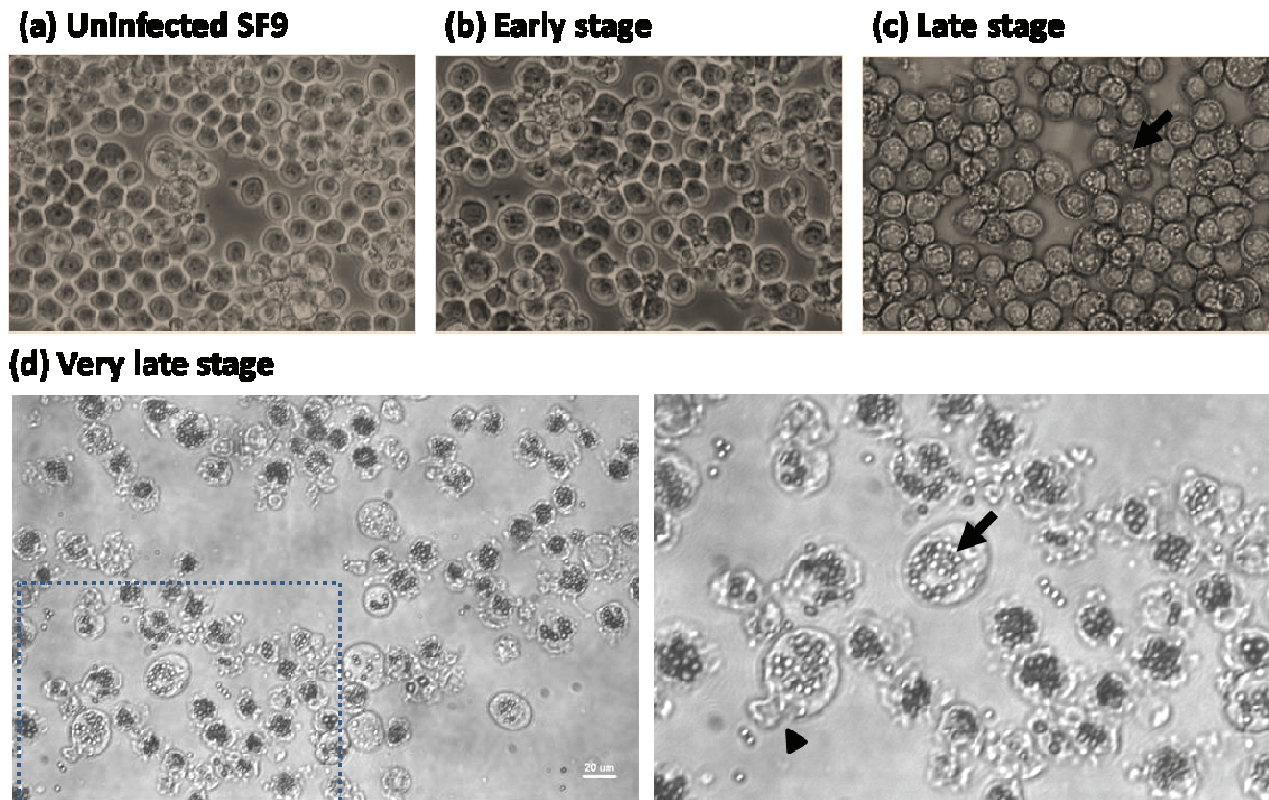


Figure 4: The morphology of 4-7 days post-transfected Sf9 insect cells

Time lapse experiments were carried out on Sf9 insect cells infected by baculovirus which generated after transfection. Panel (a) shows the cell morphology of Sf9 before virus infection, panel (b)-(d) shows different stages after virus infection on Sf9. Compared with panel (a), the growth of Sf9 after infection at the initial few days (b) is not obviously, cessation of cell growth can be observed. Few days later, viral occlusions formed in few cells (wild-type virus infected cells) which appear as refractive crystals in Sf9 are indicated by arrows (c)-(d). At the very late stage of the virus infection, cell lyzed and detached from culture dish, among these, some cell released the occluded virus which indicated by arrowhead (d). These characteristics described above represented a successful transfection.

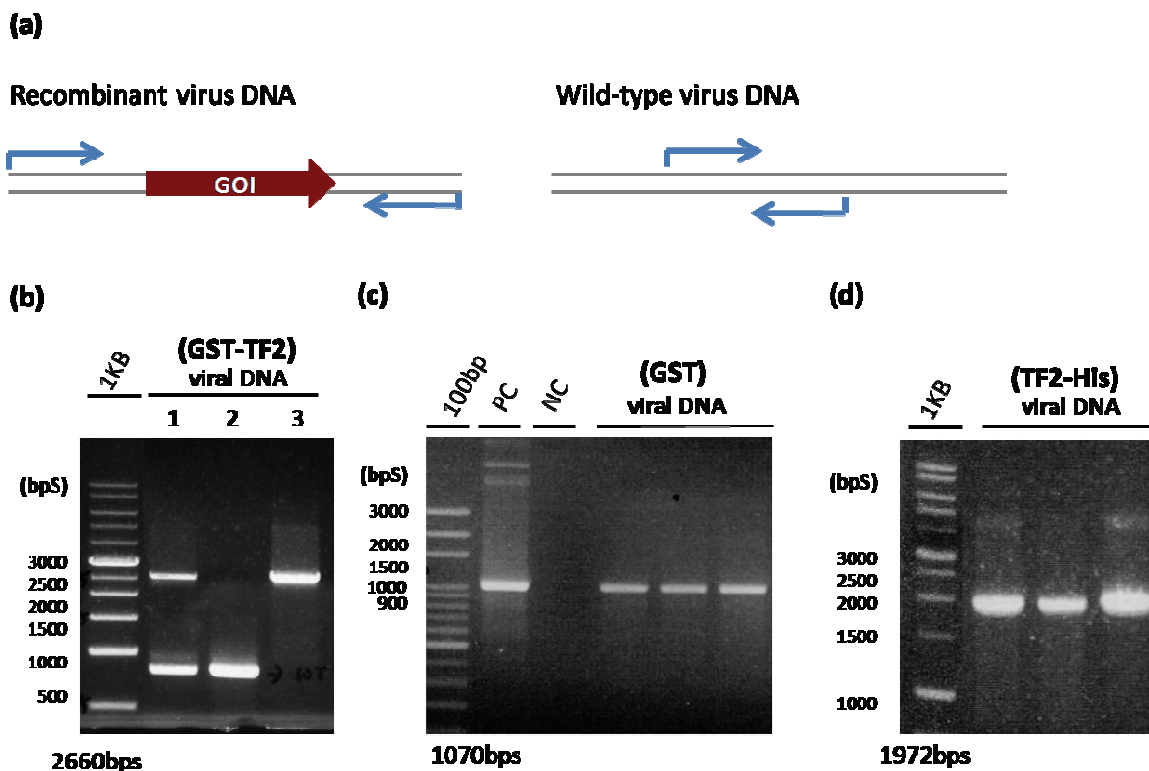


Figure 5: Verify the recombinant virus DNA by polymerase chain reaction

TF2 recombinant virus was checked by polymerase chain reaction (PCR). To verify all recombinant viruses which contained specific inserts, we performed PCR analysis by using primers that flank the polyhedrin region at both sides of insertion site after isolation of viral DNA **(a)**. The following are the results of GST-TF2 **(b)**, GST **(c)** and TF2-His **(d)**.

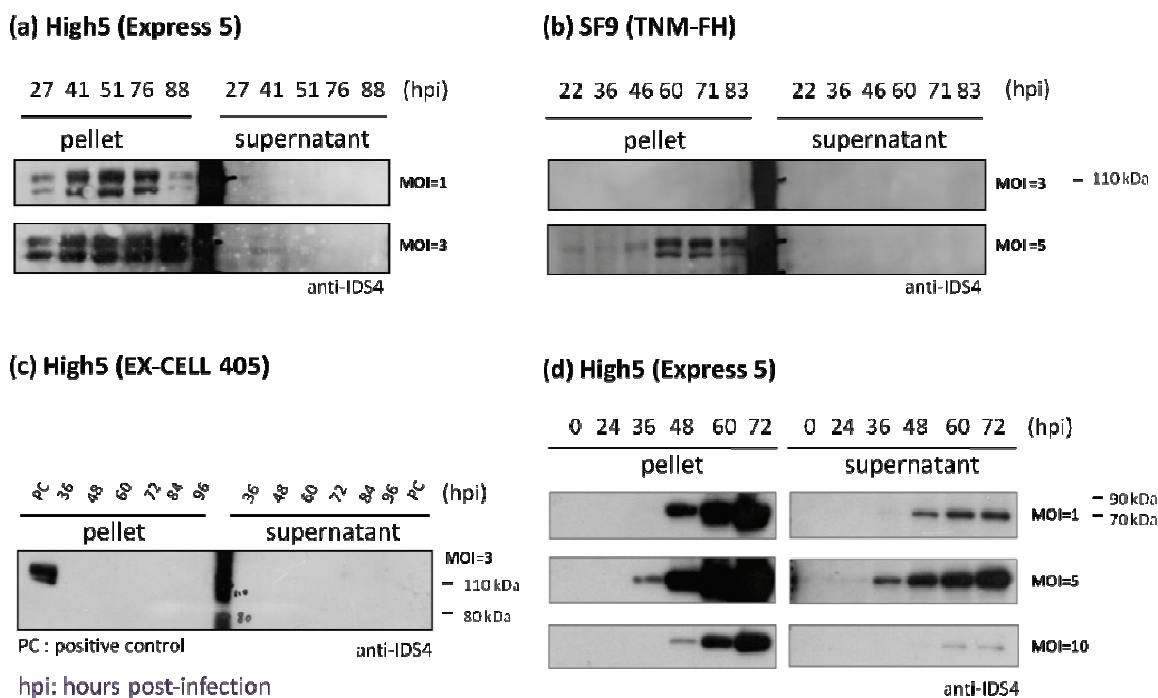


Figure 6: The optimal expression time point and MOI of TF2-His can be recognized

Time course analysis of the recombinant protein expression in High5 or Sf9 insect cells, anti-IDS4 antibody was used to detect TF2-His recombinant protein. Panel (a)-(c) displays the pattern of GST-TF2 recombinant protein expression and (d) displays the pattern of TF2-His recombinant protein expression. Panel (a)-(b) shows GST-TF2 protein express in Express5 medium by different insect cell types ((a): High5, (b): Sf9), but the protein expression pattern observed is the same, that two forms of GST-TF2 can be found in cell pellet, but secreted form of GST-TF2 was low or absent in culture media. Panel (c) represents GST-TF2 protein express in Ex-cell 405 medium by High5 cells, neither pellet nor medium have the recombinant protein expression. GST-TF2 expressed in pellet by using Express 5 medium and High5 cells was used as the positive control. On the other hand, TF2-His was expressed both in pellet and culture medium when protein expressed in Express 5 medium by High5 cells, and the optimal expression time point and MOI can be recognized as MOI=5, hpi=72 (d). MOI = multiplicity of infection, hpi = hours post-infection.

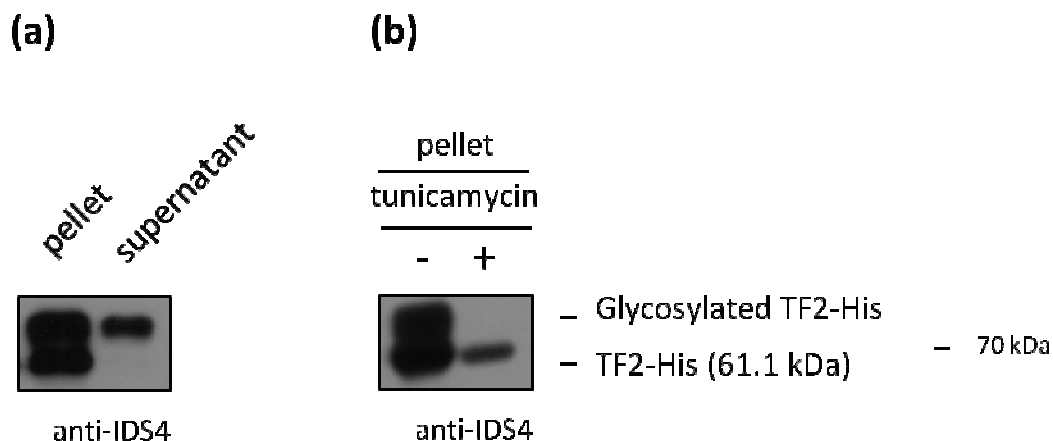


Figure 7: TF2-His expressed by insect cells have an N-glycosylated form

Two forms of TF2-His can be recognized by protein expressed by insect cells in pellet, but only the high molecular mass form of TF2-His can be secreted into medium, the molecular weight difference between this two forms is approximately 10 kDa **(a)**. Tunicamycin is an N-glycosylation blocker, which inhibits N-glycosylation of TF2-His, and the synthesis of glycoform TF2-His (upper band) was entirely blocked. This result shows that TF2-His is a highly N-glycosylated protein **(b)**.

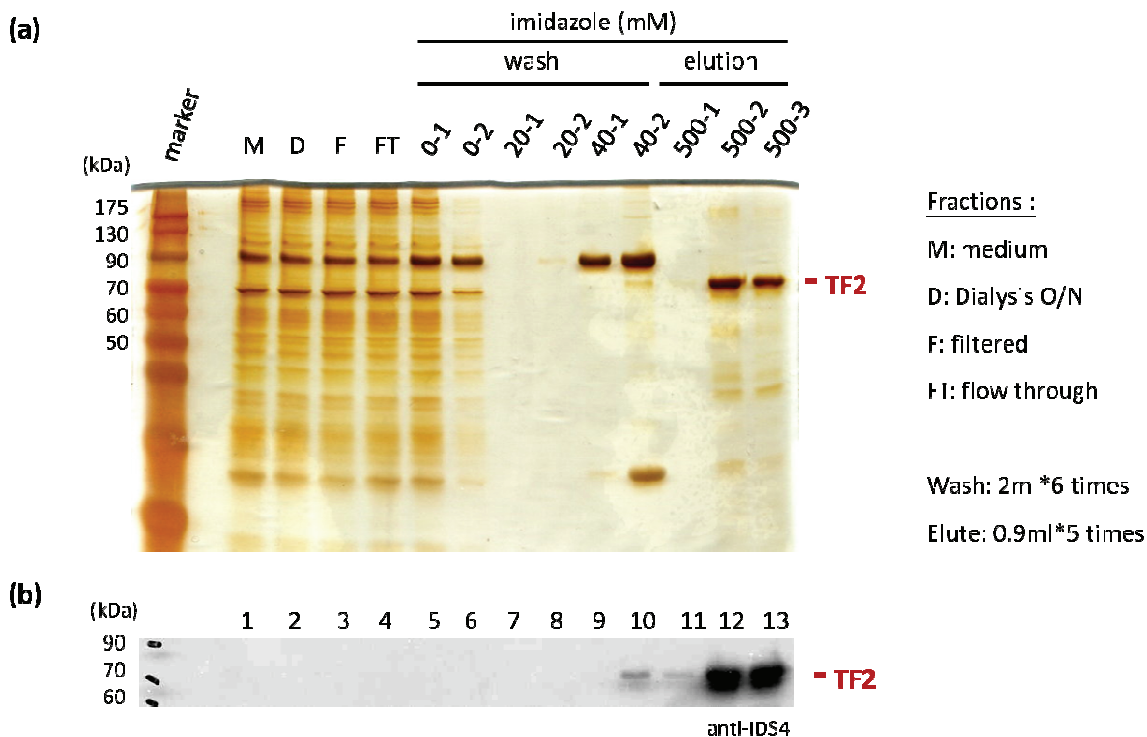


Figure 8: The establishment of an optimized purification protocol with stepwise elution of TF2-His recombinant protein.

The purification procedures of TF2-His were analyzed. TF2-His recombinant protein mainly eluted in 2-3 fractions with an imidazole concentration of 500 mM. Silver stain analysis **(a)** and Western blot analysis **(b)** reveals that the major product of elution is TF2-His recombinant protein.

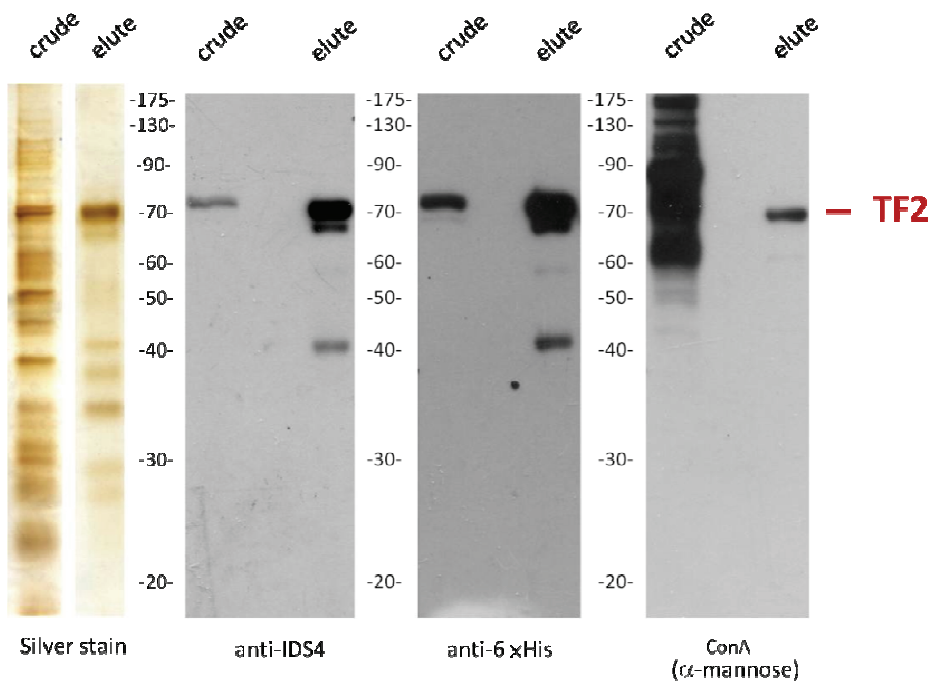
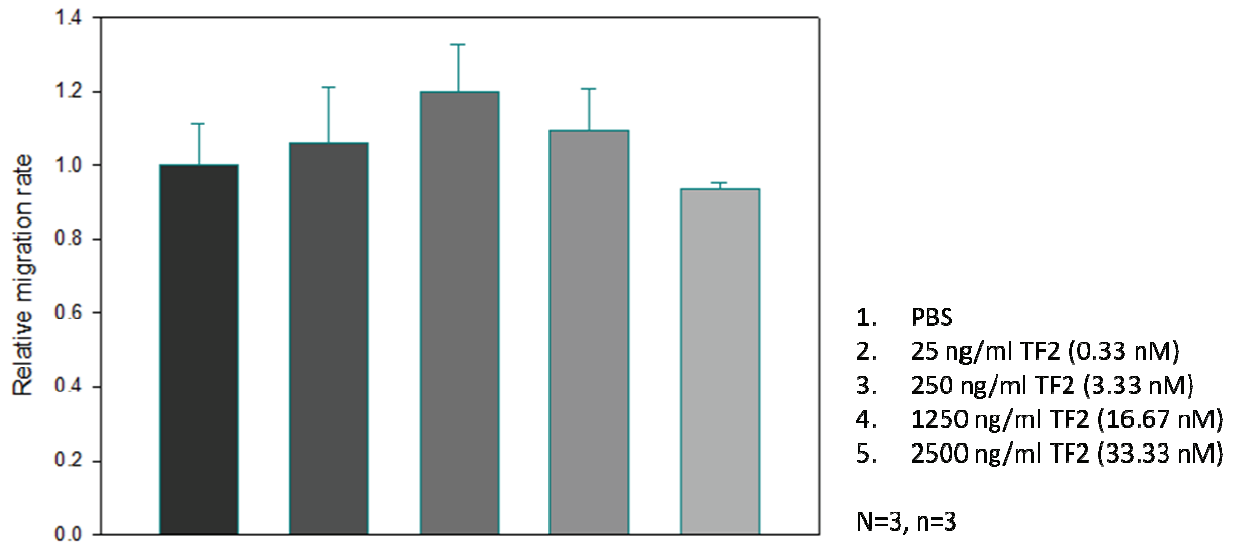


Figure 9: Identification of the TF2 recombinant protein

Identify TF2-His recombinant protein with different antibodies before (crude) and after (elute) purification. Panel (a) shows the result of silver stain. The same samples are also distinguished by Western blot by using anti-IDS4 (b), anti-6×His antibodies (c) and ConA (d).

(a)



(b)

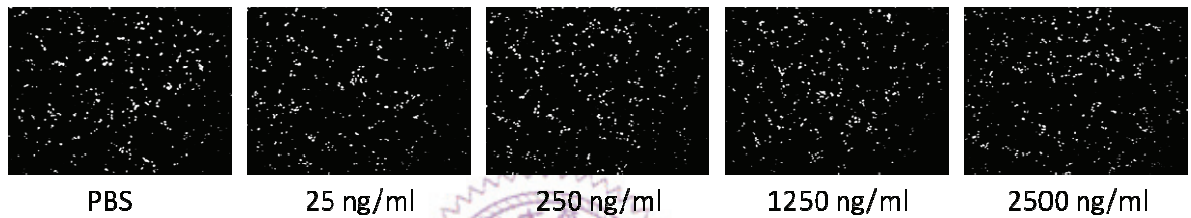
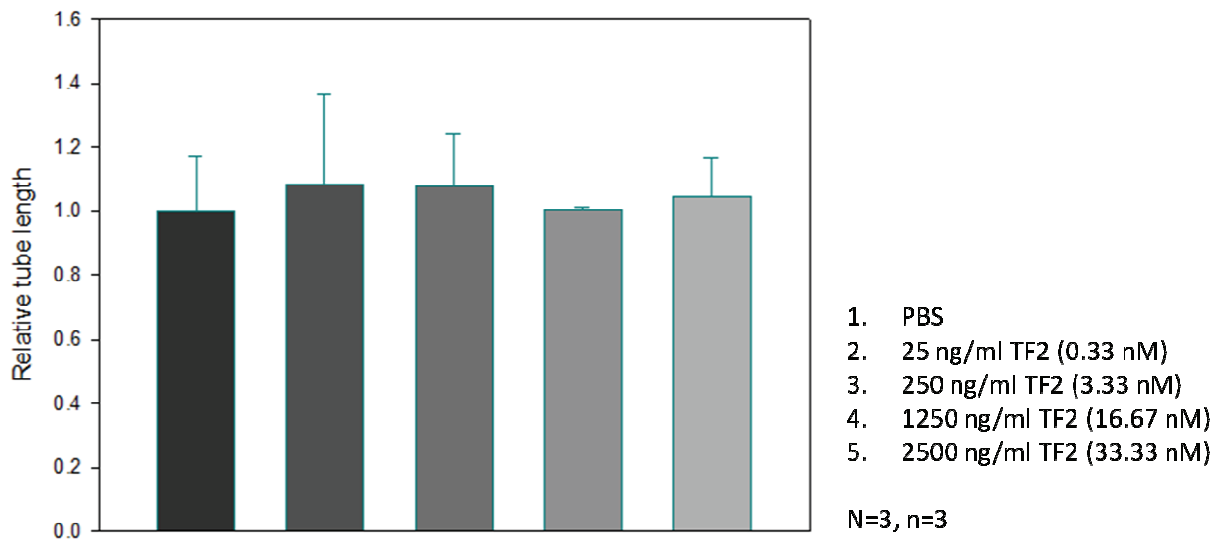


Figure 10: TF2 does not affect migration ability of endothelial cells

The figure shows that the effects of TF2 recombinant protein on directed migration of human umbilical vein endothelial cell (HUVEC). Migration ability of HUVEC was measured by transwell assay, changing the concentrations of TF2 recombinant protein in lower chamber (25, 250, 1250 and 2500 ng/ml), PBS was used as the negative control and 50 ng/ml VEGF was used as positive control. After 5.5 hours incubation period, migrated cells were stained with Hoechst 33342 and then photographed (b). Quantification of migrated cells was represented as relative values to the control cells treated with PBS in panel (a) (VEGF not shown). Five fields were counted in triplicate from each sample and three independent experiments were performed. Columns: average of three times of experiments, bars: standard deviation.

(a)



(b)

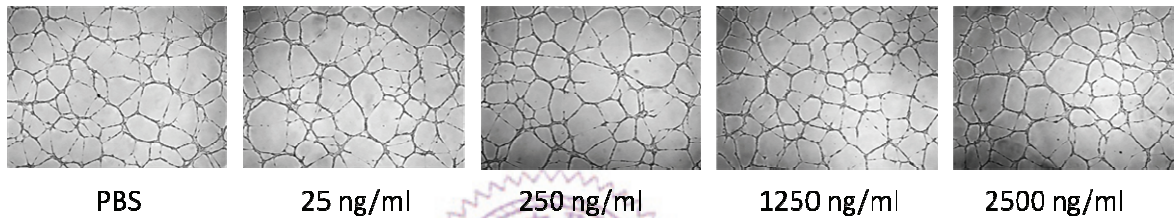
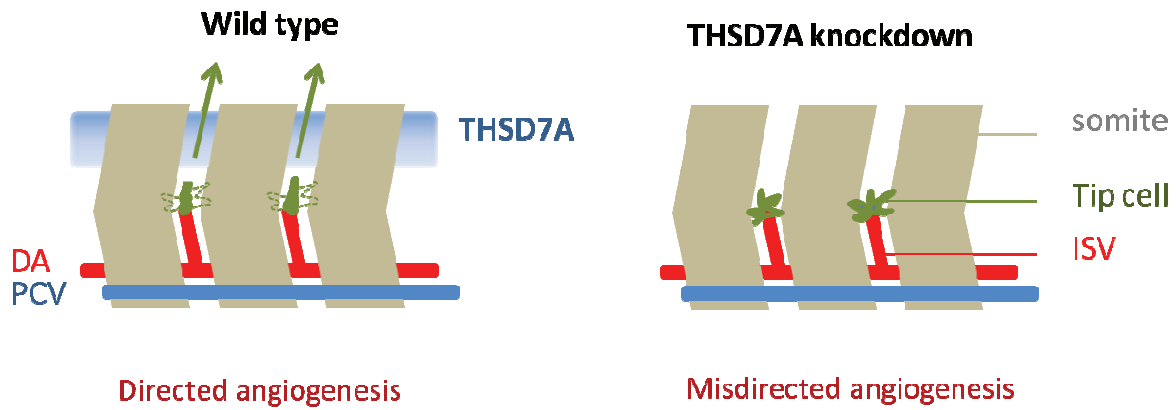


Figure 11: TF2 does not affect tube-like structure formation of endothelial cells

This figure shows that the effects of TF2 recombinant protein on migration of human umbilical vein endothelial cell (HUVEC). Cells were treated with PBS or different concentrations of TF2 recombinant protein (25, 250, 1250 and 2500 ng/ml) during the 4 hours incubation period on matrigel, then photographed the results **(b)**. Quantification of tube formation was carried out by measuring the sum tube length from three independent experiments **(a)**, it is presented as relative values to the control cells treated with PBS. Columns: average of three times of experiments, bars: standard deviation.



Supplement 1: THSD7A may act as a repulsive regulator during ISV angiogenesis *in vivo*.

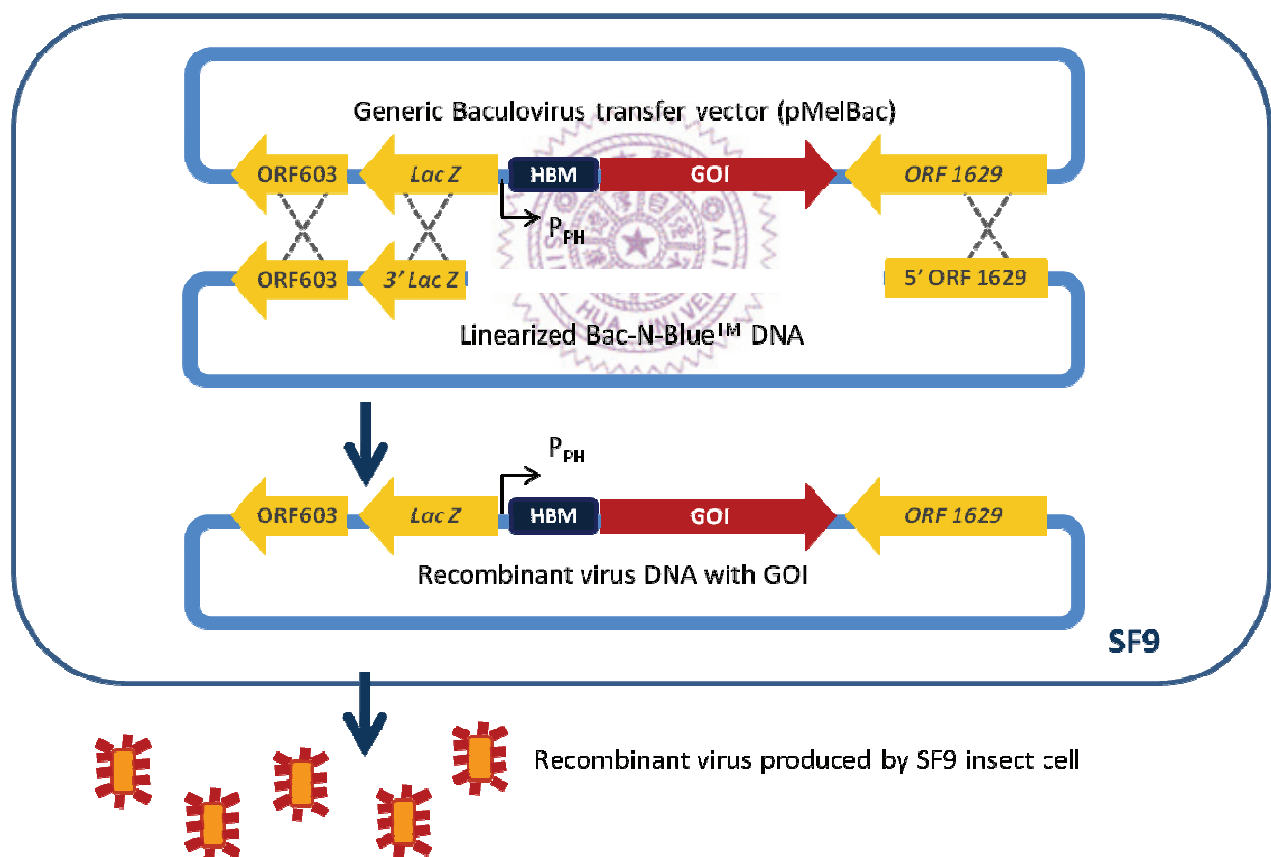
In normal condition, THSD7A inhibited excessive filopodia formation and promoted directed migration of growing vessel. But when the absence of THSD7A, the growing vessel lost its direction and lead to misdirected angiogenesis. So, we hypothesized that THSD7A may act as a repulsive regulator during ISV angiogenesis.



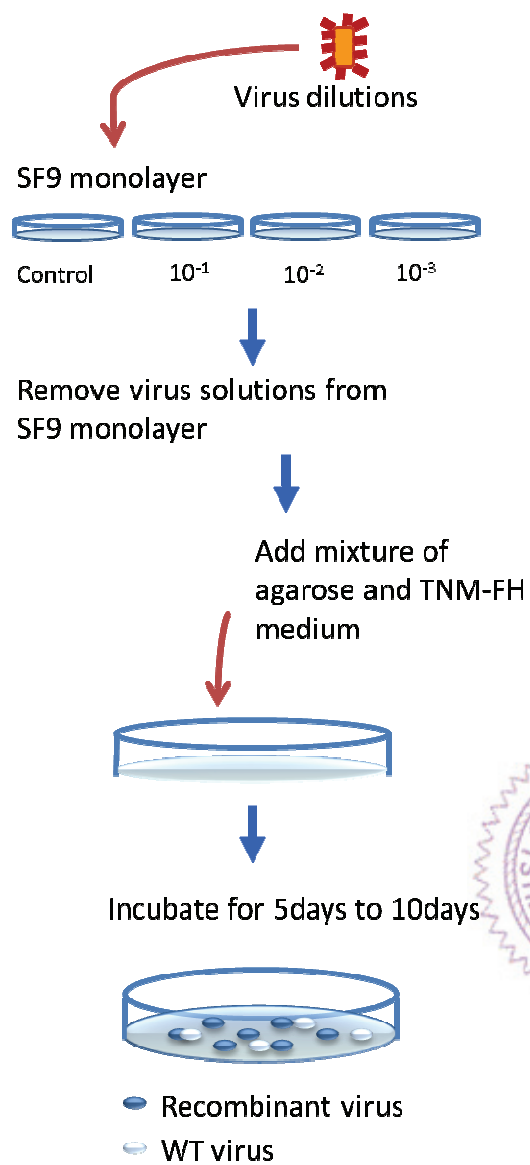
(a) Overview the processes of recombinant protein production

step	Action	Strategy	Insect cell	Figure
1	Construction of expression vector	Gene cloning into pMelBac		Figure 3 (b)
2	Recombinant virus generation	Co-transfection (pMelBac, linearized Bac-N-Blue™ DNA)	SF9	Supplement 1 (b) Figure 4
3	Virus clone isolation	Plaque assay	SF9	Supplement 1 (c)
4	P1 virus generation	Blue plaque incubation	SF9	Supplement 1 (d)
5	Virus clone confirmation	PCR analysis of virus DNA		Supplement 1 (e) Figure 5
6	Recombinant virus enlargement	Virus infection (P-1 → P-2 → P-3)	SF9	Supplement 1 (f)
7	Virus titer confirmation	End-point dilution	SF9	Supplement 1 (g)
8	Recombinant protein production	Virus infection	High5, Sf9	Figure 6

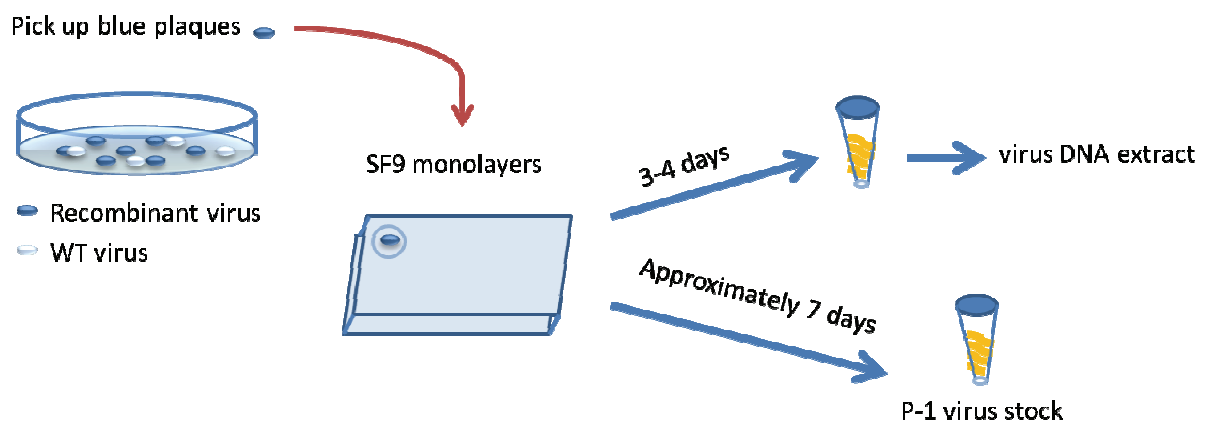
(b) Recombination



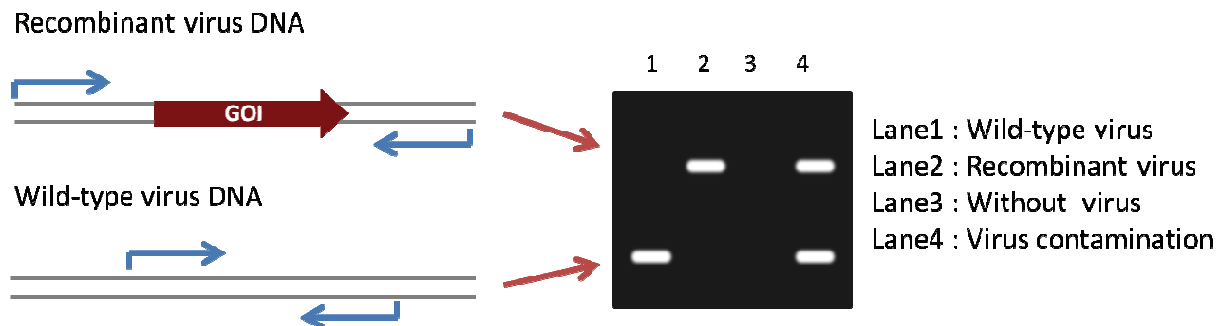
(c) Plaque assay



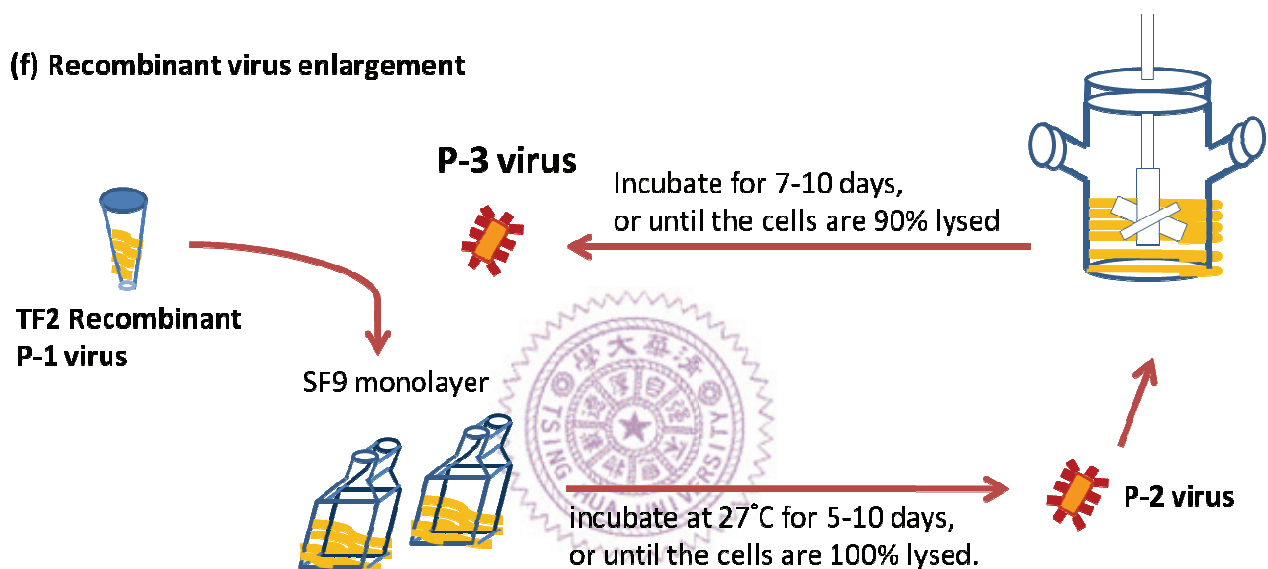
(d) P-1 virus generation



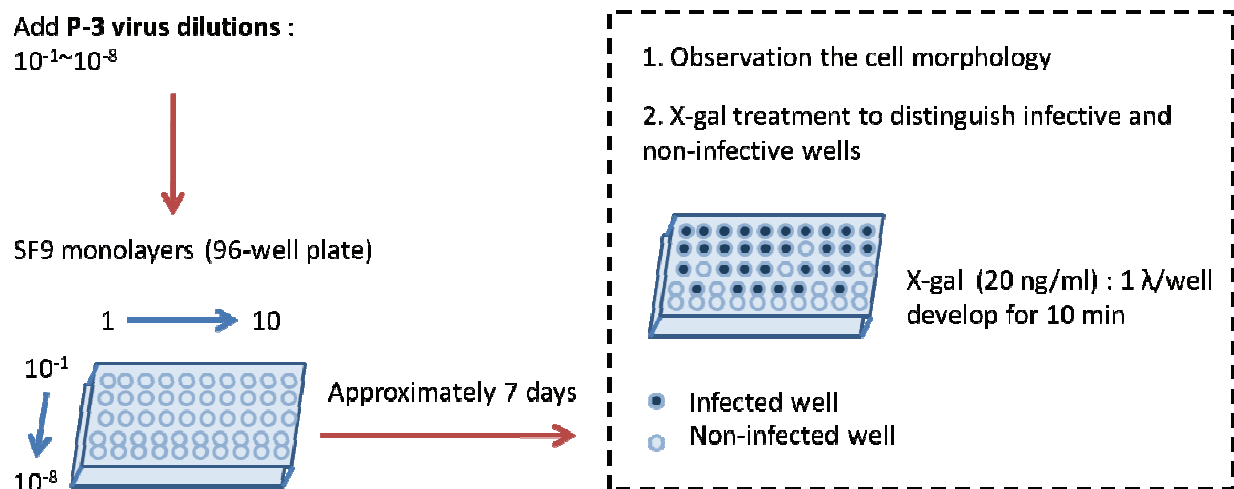
(e) PCR analysis of virus DNA



(f) Recombinant virus enlargement



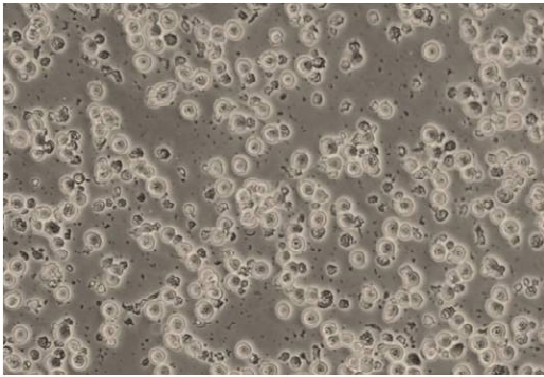
(g) Titer the recombinant virus (End-point dilution)



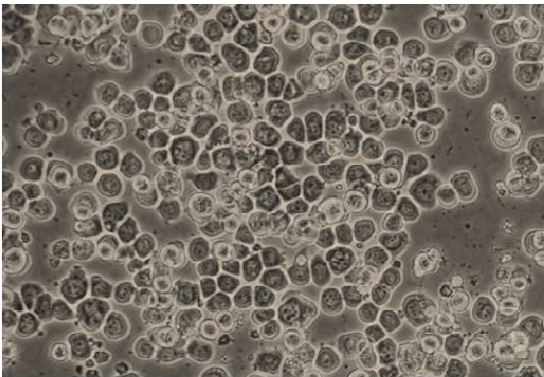
Supplement 2: Overview the process of recombinant protein production

Bac-N-BlueTM Baculovirus expression system was used to produce TF2 recombinant protein in this study. The reason for using baculovirus system in this study was described (section 1.6). The experimental processes were as follows: **(a)** Overview the processes of recombinant protein production. Step 1-7 shows the procedures for recombinant virus production. Step 8 shows the procedure for recombinant protein production. First, TF2 containing GST or His tag was constructed into pMelBac plasmids (Figure 3). **(b)** Then, co-transfected pMelBac and linearized Bac-N-BlueTM DNA (polyhedrin gene was replaced) into Sf9 insect cells, these two plasmids have the chance to be recombined and generated a recombinant virus DNA with TF2 gene which can be driven by a strong promoter called P_{PH} without forming polyhedrin. On the other hand, wild-type virus DNA may exist, too. It refers to circular form Bac-N-BlueTM DNA (uncut Bac-N-BlueTM DNA) that will produce polyhedrin. The produced virus mixture was collected in this process. **(c)** Next, plaque assay was used to separate recombinant virus from wild-type virus. **(d)** Blue plaques were then picked and incubated with Sf9 to generate P-1 virus. **(e)** Then, extracted virus DNA of different virus clones. PCR analysis was performed by using primers that flank the polyhedrin region at both sides of the insertion site, the recombinant virus and wild-type will perform different DNA patterning. Thus, the recombinant virus can be distinguished from wild-type virus. Four kinds of patterning represents four conditions may exist: wild-type virus only (e, lane 1), recombinant virus only (e, lane 2), without any virus (e, lane 3) and virus contamination (e, lane 4). **(f)** After the TF2 recombinant virus clone was picked, P-1 virus was enlarged to P-3 virus. **(g)** End-point dilution was then used to titer the virus (virus counting was based on David R. O'Reilly *et al.*, Baculovirus Expression Vectors/ A Laboratory Manual, 1992, 133-134). Finally, the virus was used to infect High5 or Sf9 insect cells to produce recombinant protein (Figure 6).

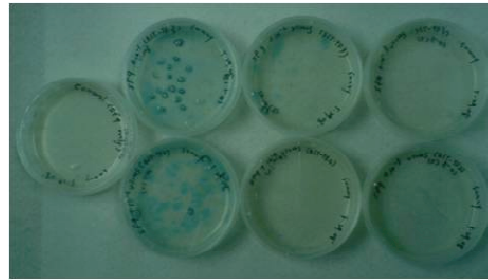
(a) 1 day post-Infection



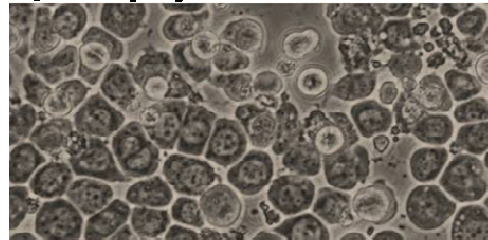
(b) 4 days post-Infection



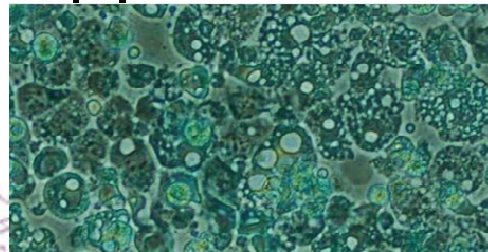
(c) 5-10 days post-Infection



w/o Blue plaque

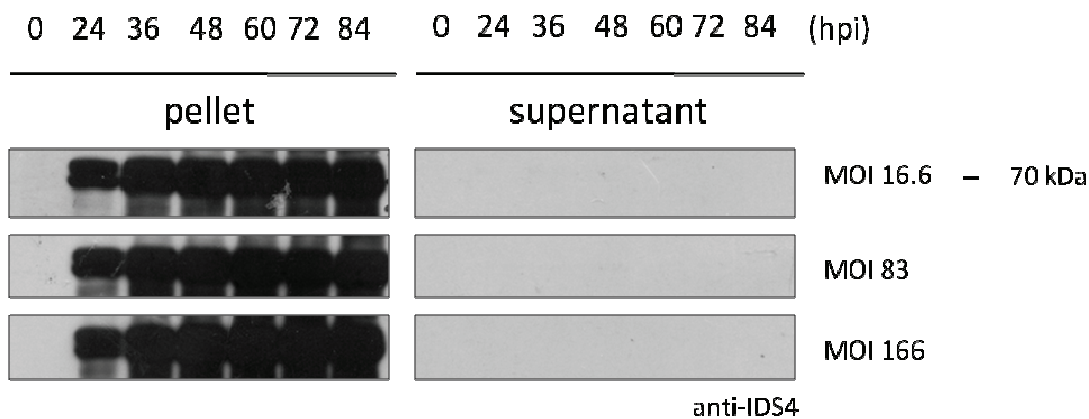


Blue plaque



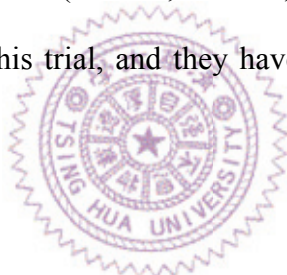
Supplement 3: Blue plaque formation reveals the potential of Sf9 cells infected by TF2 recombinant virus

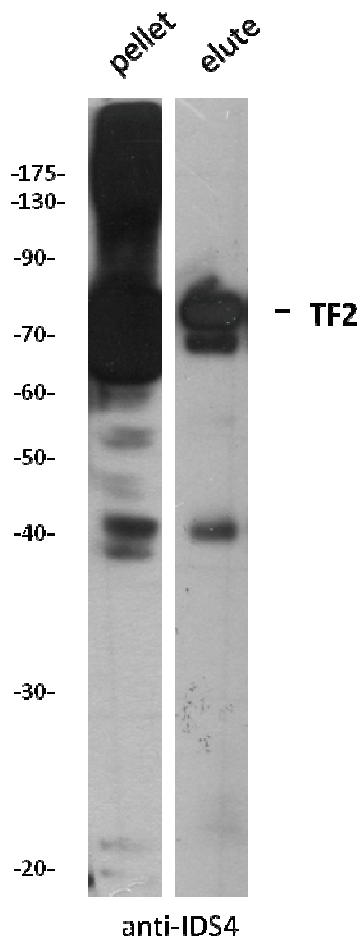
Recombinant virus clones were separated from WT virus clones by plaque assay, in which 10^{-1} to 10^{-3} dilutions of virus mixture and no virus control were tested in this experiment. Round-shaped Sf9 (supplement 3a) became flat-shaped (supplement 3b) after 3-4 days of culture in a regular procedures of all conditions. The blue plaques appeared during 5-10 days post-infection, which indicated a successful infection (supplement 3c, upper panel). Compared the Sf9 morphologies between control and infected cells in enlarged view, control cells showed no significant changed from 3-4 days post-infection (supplement 3c, middle panel), but the infected cells marked with the blue plaques had larger cell bodies with vacuole-like structures (supplement 3c, bottom panel). Plates with 10^{-2} virus dilution showed non-crowded distribution but sufficient quantity of blue plaques in all three types of recombinant virus.



Supplement 4: Excessive MOI resulted in similar expression level of two forms of TF2 and have poor secretion ability

Same batch of TF2-His virus was used for protein expression at higher MOI (MOI=16.6, MOI=83 and MOI=166) compared with Figure 6d (MOI=1, MOI=3, MOI=5). The expression level of two forms of TF2-His were the same in this trial, and they have poor secretion ability and lower cell survival rate than Figure 6d.

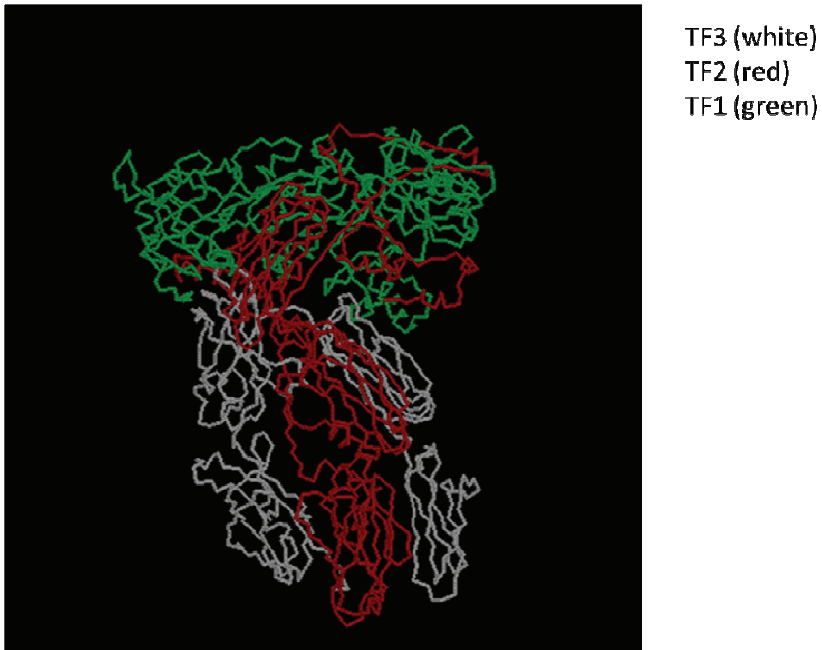




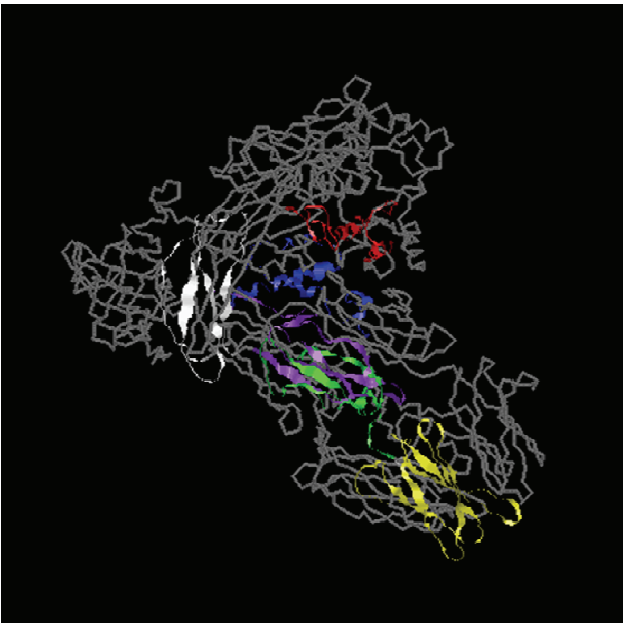
Supplement 5: The impurities of TF2 elution is contaminated by the damaged cells.

Compared pellet with eluted protein, the impurities, 62 kDa and 42 kDa bands (supplement 5, lane 2), seemed to be derived from the non-secreted portion of insect cells (supplement 5, lane 1). This is supported by the finding that the 62 kDa band may be the unglycosylated form of TF2-His.

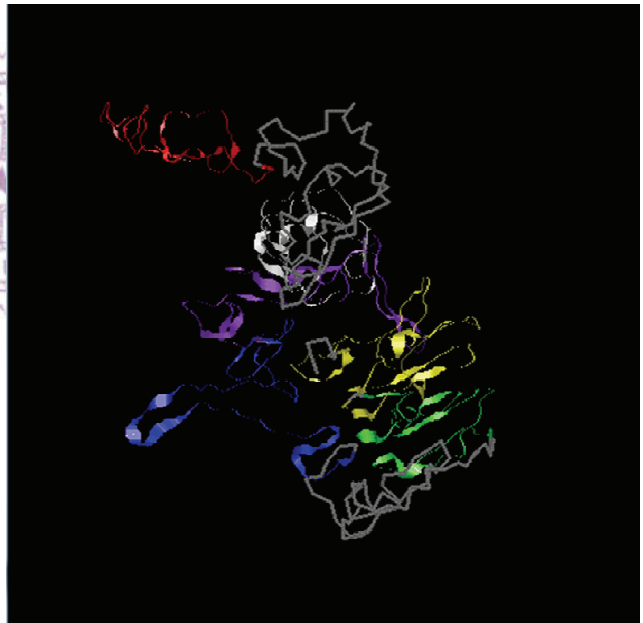
(a) THSD7A predicted by I-TASSER (TF3, TF2, TF1)



(b) Predicted THSD7A (TF2 containing 6*TSR)



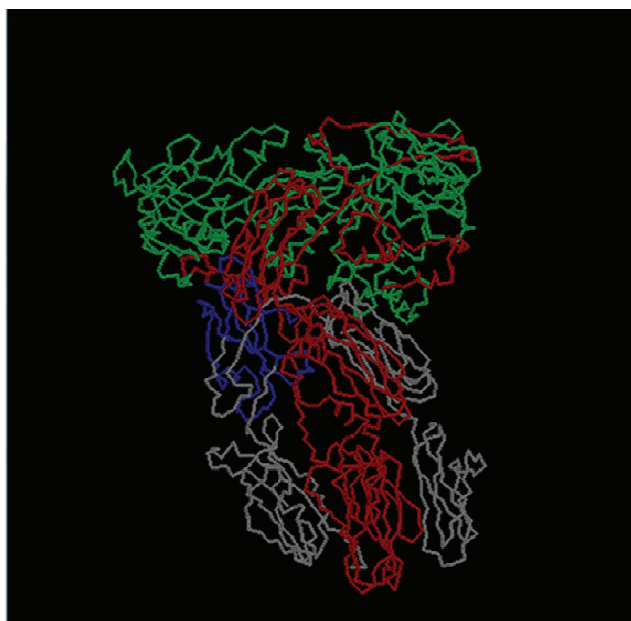
(c) Predicted TF2 (6*TSR)



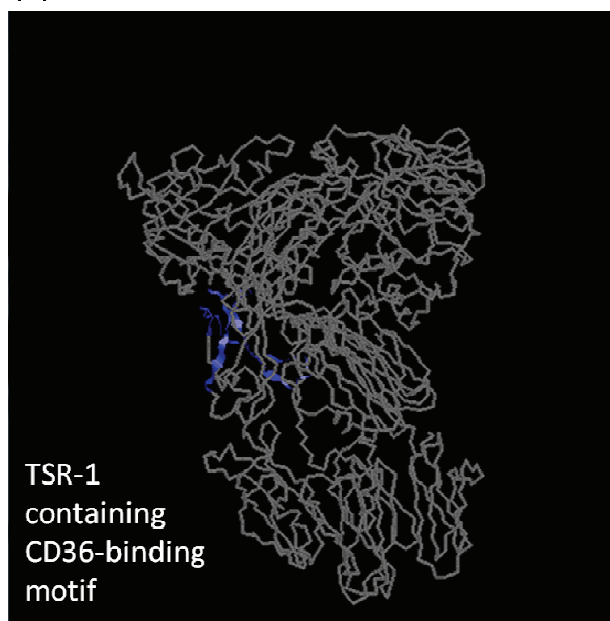
Supplement 6: The structure of TF2 may be destroyed after truncation

The predicted native THSD7A showed like a globular structure **(a)**, three truncated fragments was labeled on it: TF1 (green), TF2 (red) and TF3 (white). Six TSRs on TF2 was labeled on native THSD7A, too **(b)**. The structure of TF2 only was also predicted and labeled six TSRs on it **(c)**. Compared with native and truncated form of TF2, the TSRs were abnormality in truncated form. It showed that the structure of TF2 was destroyed after truncation.

(a) Predicted binding site (label blue)



(b) TSR-1 on THSD7A



Supplement 7: The predicted binding domain contain a TSR and a CD36-binding motif

The structure predicted result showed that THSD7A contain a predicted protein binding domain on TF3 (a), it comprise a complete TSR (b). This TSR was predicted to contain a CD36 binding motif which may bind to cell surface receptor CD36.