筑波大学
University of Tsukuba

博士 (人間生物学) 学位論文
PhD dissertation in Human Biology

Targeting novel oncogene THG-1 with macrocyclic peptides for squamous cell carcinoma therapeutic

(扁平上皮癌におけるタンパク質間相互作用を標的としたマクロサイクリックペプチドの解析)

2019

筑波大学グローバル教育院

School of Integrative and Global Major in
University of Tsukuba
PhD program in Human Biology

Sophie T. PhuongDung Tran

Targeting novel oncogene THG-1 with macrocyclic peptides for squamous cell carcinoma therapeutic

Sophie T. PhuongDung Tran

Table of Contents

ABSTRACT

Purpose:

Ineffective current treatments for squamous cell carcinoma (SCC) is one of the main reasons for its low 5-year survival. Transforming growth factor beta-1 stimulated clone 22 - isoform 4 (Tsc22D4/ THG-1) is overexpressed in more than 80% of SCC specimens including esophagus SCC, lung SCC and cervical SCC while strictly resided in mitotically active basal layer of normal squamous epithelial tissue. Knockout of THG-1 caused reduction in cancer cells growth, invasion and tumorigenesis. Experimentally, protein-protein interactions (PPIs) were suggested as THG-1 mechanism to promote cancer progression. Well-known functional interactions of cellular regulators, including: Keap1, PHD2, TBLR1 and NRBP1, are disrupted in the present of THG-1. In my research, THG-1 specific macrocyclic peptide ligands were identified and their application for research involving THG-1 as well as their potential use for disrupting THG-1 PPIs were evaluated.

Materials and methods:

The modified Random non-standard Peptide Integrated Discovery (RaPID) system was employed for screening of macrocyclic peptides against purified recombinant THG-1 protein. In this system, D-stereochemistry, unusual side chains and N- methylation containing macrocyclic peptides were generated and screened against THG-1. Identified macrocyclic peptides were used for fluorescein staining, pull-down and ELISA-like peptide sorbent assay. Moreover, PPI inhibition activity of identified macrocyclic peptide was evaluated both in vitro and in cells.

Results

By using the modified RaPID system, I successfully identified macrocyclic peptides that specifically bind to both recombinant THG-1 and endogenous THG-1. R4-1 macrocyclic peptide successfully pulled-down both recombinant THG-1 as well as endogenous THG-1 protein in esophagus squamous cell carcinomas cell lines. In addition, fluorescein-tagged R4-1 specifically stained THG-1 in the THG-1 overexpressed HaCaT cells and displayed low non-specific staining in the negative control HaCaT-mock cells. Especially, R4-1 showed antagonist activity to THG-1 and Keap1 interaction *in vitro* by NanoBiT and co-IP assay. In addition, R4-1 treatment downregulated Nrf2's target genes expression and caused reduced cells proliferation.

Discussion

Application of identified macrocyclic peptide ligands against THG-1 in staining, ELISA-like peptide sorbent assay and pull-down in replacement for antibody would reduce the cost, and consumed-time for those procedure in laboratory. Most important, identified macrocyclic peptide R4-1 is THG-1 PPI antagonist. While THG-1 and Keap1 interaction implies important role in cancer progression, R4-1 inhibitory activity may serve as therapeutic for treatment of squamous cell carcinomas.

ABBREVIATIONS

BLI	Bio-layer interferometry
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue
cDNA	complementary DNA
CPP	cell penetrating peptide
DIPEA	N, N-Diisopropylethylamine
DMF	Dimethyl fumarate
DMSO	Dimethyl sulfoxide
E.*coli*	*Escherichia coli*
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked Immunosorbent assay
ESCC	Esophagus squamous cell carcinomas
FACS	Fluorescence activated cell sorting
FCS	Fetal calf serum
FIT	Flexible in vitro translation
HBTU	(2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEK293T	Human embryonic kidney 293T
HNSCC	Head and neck squamous cell carcinomas
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography

IB	Immunoblotting
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Keap1	Kelch-like ECH associated protein 1
LSCC	Lung squamous cell carcinomas
MALDI-TOF	Matrix-assisted laser desorption/ionization- time of flight
MMT	Monomethoxytrityl
mRNA	messenger RNA
NHS	N-Hydroxysuccinimide
NMP	N-Methyl-2-pyrrolidone
NQO1	NAD(P)H dehydrogenase quinone 1
Nrf2	Nuclear- factor (NF) related factor 2
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGD	6-phosphogluconate dehydrogenase, decarboxylating
PHD2	prolyl hydroxylase domain-containing protein 2
PPI	Protein- Protein Interaction
qPCR	Quantitative real-time PCR
RaPID	Random non-standard Peptide Integrated Discovery
RT-PCR	Reverse transcription-PCR
SCC	Squamous cell carcinomas

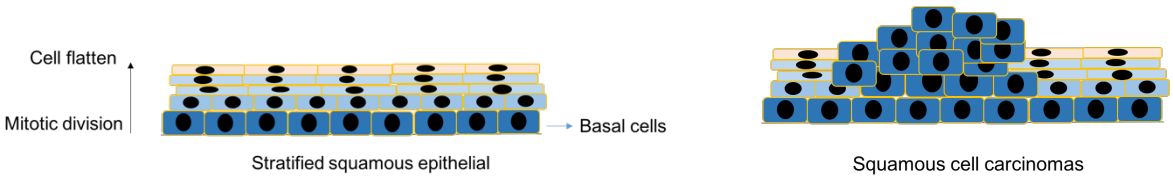
SDS	Sodium lauryl sulfate
TBLR1	transducin β-like protein 1 (TBL1)- related protein 1
TFA	Trifluoroacetic acid
THG-1	TSC Homologous Gene 1
TIS	Triisopropylsilane
TKT	Transketolase
TP53	Tumor protein p53
TSC22	Transforming growth factor-β (TGF-β) stimulated clone 22

CHAPTER I: GENERAL INTRODUCTION

1. *Squamous cell carcinoma*

Squamous epithelial cells exist in the tissues of the external surfaces of the body, the upper aero digestive tract and the lower female genital tract. Squamous tissues are usually exposed to oxygen and needed a cytoprotective mechanism for oxidative stress. Squamous tissue has a stratified structure with a stem-cells- and progenitor-cells containing basal layer. Accumulated mutations and alteration of genomic DNA in stem or progenitor cells in basal layer lead to cancer development, especially mutations in genes that regulate oxidative stress (Rock J. R. et al, 2010). Marker genes that associate with basal cells characteristic could be useful for diagnostic and therapeutic against SCC.

Figure 1.1: Structure of stratified squamous epithelia and squamous cell carcinomas tissue



Squamous cell carcinoma (SCC) are most frequently found cancer type of skin, oral, head and neck, esophagus, cervical and lung cancer. Esophagus SCC accounts for 90% of esophagus cancer, and caused 400,000 deaths in 2012 (ACS, 2011). Lung SCC is major types of non-small cell lung cancer and kills approximate 50,000 people annually in United State (ACS, 2014). Eighty percent of cancer of the uterus cervix are squamous cell carcinomas (Parkin DM., et al., 2012) with 27,600 new cases reported and 265,700 deaths worldwide (Torre LA., et al., 2012). Five

hundred fifty thousand new cases were reported annually with 300,000 deaths caused by head and neck cancer, while among those 90% are squamous cell carcinomas (Jemal A., et al, 2011).

a. Esophagus squamous cell carcinomas (ESCC)

(i) Epidemiology

Esophagus cancer is the 8th most common cancer worldwide and the 6th leading cause of cancer-related deaths in the world. Among its two subtype, adenocarcinoma and SCC, SCC account for 80% of cases of esophagus cancer worldwide (Kamangar F., et al, 2006). ESCC is predominant in Asia, Africa, South America, Southern Europe and among African American in North America. Eighty percent of cases occur in developing countries (Rusti AK. and El-Serag HB., 2014).

ESCC is rare at young age, and usually occurs in the seventy and eighty years of life. Incidence of ESCC is three times higher in blacks compared to whites. The difference in distribution of ESCC indicate that ethic and genetic factors as well as lifestyle all play roles in the development of ESCC.

(ii) Risk factor

Environmental risk factor:

Alcohol and tobacco are two among the highest risk factors for ESCC. The risk of ESCC is twice in current smoking group than a group containing non-smoker (Rustgi AK. And El-Serag HB., 2014). Tabacco contains carcinogenic substances such as nicotine-derived nitrosamine and N'-nitrosonornicotine. Exposure of esophagus mucosa to carcinogens has been proposed as the cause of ESCC.

Incidence of ESCC also three to five times higher in people who consume alcohol. Alcohol is absorbed by the upper gastrointestinal tract and can be metabolized into acetaldehyde. Acetaldehyde can react with deoxyguanosine (dG) in DNA, which might result in single- and double-stranded breaks, point mutations, chromosome aberrations (Brooks PJ., et al, 2014).

Genetic risk factor:

A whole-exome sequencing study in 192 ESCC tumors and adjacent normal tissue of patients in north-central China reveals that TP53 mutations occurred in 90% of patients (Zhang L., et al, 2015). Epidermal growth factor receptor (EGFR) is found to be overexpressed in 60-76% of ESCCs (Zhang W., et al, 2014). EGFR overexpression would increase its signaling transduction cascades such as Akt, MAPK, and JNK, which lead to cell proliferation, migration, and tumorigenesis. Mutations also were reported in genes that regulate the cell cycle with low frequencies (2%-10%) such as CDKN2A (encodes p16), PIK3CA, NFE2L2 (Nrf2), NOTCH1 and NOTCH3 (Song Y., et al, 2014).

(iii) Treatment

Over the last several decades, the five-year survival rate remains to range from 15% to 25% (Pennathur A. et al, 2013). Currently, early stage ESCC with negligible metastasis can be cured with endoscopic resection and ablation. However, without prognostic factors, early stage ESCC is asymptomatic and difficult to detect (Campbell F., et al. 2007). ESCC is often diagnosed in late stages and is accompanied by with extensive metastasis to lymph nodes in the cervical, thoracic, and abdominal regions. In Japan, cisplatin plus 5-fluorouracil (5-FU) is used as the standard chemotherapy for locally

advanced ESCC. Chemotherapy followed by surgery obtained 55% of 5-year rate survival (Ando N., et al. 2012).

Event though, EGFR is the most studied molecular target in SCC, there is little evidence supporting the advantages of EGFR-targeting therapy for ESCC. In phase 3 clinical trial, additional of anti-EGFR monoclonal antibody to chemoradiotherapy did not provide significant survival benefit (Crosby T., et al, 2013). PD-1(programmed cell death 1) -PD-L1 (programmed cell death ligand 1) positive patients had significantly poorer outcomes than ESCC patients without this activation pathway. Recently, strategies for targeting this pathway have been developed and a clinical trial is underway (Ohigashi Y., et al, 2005).

b. Lung SCC (LSCC)

(i) Epidemic

Lung cancer is the leading cause of cancer death with 1.8 million new cases were estimated in 2012 (Ferley J., et al, 2013). The 5-year survival rate in the United State for lung cancer is 15.6% (Dela-Cruz CS. et al., 2011). Lung cancer can be divided into small cell lung cancer (15% of cases) and non-small cell lung cancer (85% of cases). Non-small cell lung cancer is further divided into adenocarcinomas (38.5% of cases), squamous cell carcinomas (20% of cases) and large cell carcinomas (2.9% of cases).

(ii) Risk factor

Environmental factor

LSCC incidence is positively correlated with smoking, 91% of LSCC was attributed to cigarette smoking (Molina JR., et al, 2008). SCC was the most predominant type of non-

small cell lung cancer in the past century. However, from the 1960s and 1970s, with the changes in cigarette design and composition, smoke inhalation pattern, and improved diagnostic approaches, there was a shift from SCC into adenocarcinomas (Drilon A. et al, 2012).

Genetic risk factor

Genes involved in oxidative stress and squamous differentiation were found to be frequently altered in LSCC. Nineteen percent of patients of LSCC possess Nrf2 mutations at its Keap1 interaction motif. Keap1 loss of function mutation was reported in 12% of patients (CGA research network, 2012). The alteration in stress defense system results in activating of many antioxidation enzymes which bring benefit to cancer cells survival. Similar to ESCC, 81% of patients had TP53 mutations.

SOX2 and TP63 expression have also been reported at a rate of 21% and 16%, respectively. SOX2 expression was significantly higher in LSCC compared with adenocarcinomas in immunohistochemistry staining (Wilbertz T. et al., 2011). N-terminal-truncated (ΔN) P63 is the most common expressed isoform in squamous epithelial and lung carcinomas (Hibi K., et al, 2000). ΔNP63 antagonizes TP53 and serves as a stem-cell factor in basal cells of squamous epithelia.

(iii) Treatment

Currently, gemcitabine/cisplatin is used as chemotherapy for LSCC with relative benefit compared to pemetrexed/cisplatin (Derman BA., et al, 2015). In addition, nab-paclitaxel is also used in chemotherapy for non-small cell lung carcinomas (NSCLC).

Even EGFR mutations do not occur in LSCC (Rekhtman et al., 2012), the addition of Nactitumumab, an EGFR IgG1 monoclonal antibody against EGFR, showed improved overall survival, progression-free survival, and disease control rate (Thatcher N., et al., 2014).

Similar to ESCC, immune evasion may play an important role for LSCC. The anti-PD-L1 antibody and the anti PD-1 monoclonal antibody are under clinical trial. The anti PD-1, pembrolizumab response rate was 18-25% in the squamous histology NSCLC in phase I trial.

c. Head and neck SCC (HNSCC)

(i) Epidemiology

There were 500,000 new cases of HNSCC reported annually (Torre LA., et al, 2015). HNSCC can arise within the oral cavity, oropharynx, hypopharynx, nasopharynx, and larynx.

HNSCC incident is higher in men compared to women. There are 20 out of 10,000 men affected by HNSCC in France, Hong Kong, the Indian subcontinent, Central and Eastern Europe (Stenson KM., 2018). The Indian subcontinent has a higher incident of oral SCC, while Hong Kong has a higher rate of nasopharyngeal SCC, pharyngeal/ or laryngeal SCC is more common in the other populations.

From 1980 to 2011, there was a steady increase of oropharyngeal SCC and declining of hypopharynx and larynx SCC. The mortality of oropharyngeal SCC and laryngeal SCC is higher in African American compared to white.

(ii) Risk factor

Environmental factor

Exposure to tobacco, tobacco-like products, and alcohol increases the risk of developing HNSCC. Heavy smokers have 5 to 25 times higher risk of HNSCC than non-smokers. Smokeless tobacco still associates with a higher incidence of HNSCC (Wyss B., et al, 2016). Passive smoke exposure also relates to a higher risk of HNSCC. These patients exposed to environmental tobacco smoke in both workplace and home, the incidence is higher in women with tongue cancer.

Infection of papillomavirus (HPV) is also a risk factor of HNSCC, especially oropharyngeal SCC. More than 50% of oropharyngeal SCC patients have HPV-positive (Marur S., et al, 2016). HPV positive HNSCC patients are typically middle-aged, non-smoking men of higher social status. Fortunately, HPV-positive HNSCC has a better prognosis compared to patients with HPV-negative tobacco-related HNSCC.

Genetic factor

HNSCC genomes have a high degree of instability. Both HPV (+) and HPV (-) HNSCC have focal amplification for 3q26/28, which includes transcriptional factors TP63, SOX2, and PIK3CA. In addition, TP53 mutation was found in 72% of patients, PIK3CA, NOTCH1, and KMT2D mutations were reported at 21%, 19% and 18% respectively (CGA network, 2015).

HPV (+) tumor has the signature deletions and truncations mutation in TNF receptor-associated 3 (TRAF3). This aberrant promotes NF-κB signaling. HPV (-) tumor harbors co-amplifications of 11q13 (CCND1, FADD and CTNN) and 11q22 (BIRC2 and YAP). Those

cause the dysregulation in the cell cycle, cell death and NF-κB signaling (CGA network, 2015)

HPV (-) tumor has predominant mutation in CUL3 (6%), Keap1 (5%) and Nrf2 (16%) compared to HPV (+). These findings suggest the Nrf2 oxidative-stress pathway perturbation is tobacco-related signature (CGA network, 2015).

(iii) Treatment

Surgery, radiation, and chemotherapy are used alone or in combination for the treatment of HNSCC depended on disease stages. For early-stage HNSCC, surgery or radiation alone are applied. For patients with advanced stages (stages III and IV), surgery together with platinum-based chemoradiation, with or without chemotherapy, are applied. However, removing a tumor in HNSCC would leave patients with difficulties in speaking and eating.

Cisplatin is used as the chemo-agents for HNSCC. The administration of cisplatin with radiation showed a definite benefit for patients (Blanchard P., et al, 2011). Combination of EGFR monoclonal antibody cetuximab and panitumumab with radiotherapy also indicated the improved overall survival (29.3 months in single treatment compared to 49 months) (Bonner JA., et al, 2010).

In head and neck SCC, common perturbations were also observed such as EGFR gene amplification, upregulation of HER2, HER3, and MET (Cancer Genome Atlas Network, 2015). Among pathways, targeted agents against NOTCH1, MET-PI3KCA-mTOR (Cohen EE., et al, 2005), EGFR-Ras-Raf-MEK (Galloway TJ., et al, 2015), and Wnt/β-catenin are under investigation (Cancer Genome Atlas Network, 2015). However, the response rate

only ranging from 10% to 15% and have no meaningful clinical benefit (Marur S. and Forastiere AA., 2016).

In general, chemotherapy showed low success rate in treatment for squamous cell carcinomas patients while current investigated targeted therapies only limited to a small portion of patients. Currently, the trial of targeted therapies only focuses on well-known molecular pathways such as EGFR, FGFR, AKT… which only account less than 15% of molecular perturbation in squamous cell carcinomas. There is an urgent need in investigating the molecular signature of squamous cell carcinomas and developing of targeted agents for more effective treatments. In addition, similar to most cancers, squamous cell carcinomas are incurable in the late stage of diagnostic, highly associated genetic perturbation could be used as prognostic value and bring great benefits to patients.

2. *Cellular oxidative and electrophilic stress defense system and its perturbation in SCC*

Many of SCC harbor perturbation in oxidative stress regulation system (Keap1-Nrf2 system). Tobacco and alcohol exposure are important environmental factors correlated with SCC incidents. Thereby, the regulation of stress defense system is important in maintaining homeostatic and disturbing in this system would be a critical step in cancer progression.

a. Cellular oxidation and electrophilic stress defense system

Reactive oxygen species (ROS, oxygen-derived radical and non-radical substances) are continuously produced by an organism's normal use of oxygen. Free radicals are very unstable and tend to form stable compounds with other molecules (lipid, DNA, protein…). As a result, damages in the biological system can occur, such as the breakdown of lipids, the inactivation of

enzymes, the introduction of changes (mutations) in DNA, the destruction of cellular membrane…

The generated ROS are rapidly removed in a normal cell. However, if a high level of ROS is generated, cells will enter oxidative stress state. Cells with oxidative stress have a high influx of Ca^{2+} and iron. Excessive free Ca^{2+} leads to DNA fragmentation subsequently leads to apoptosis. In addition, ROS can directly induce several changes in expression of TP53, Caspase cascade, NF-κB and cytochrome *c*… whose function regulates programmed cell death (Redza-Dutordoir M., and Averill-Bates D.A., 2016)

In order to counteract oxidative stress, cells produce antioxidants, which can be either enzymatic or non-enzymatic. By donating their own electrons to ROS, the antioxidants neutralize the adverse effects. Non-enzymatic antioxidants include vitamins such as vitamin E, vitamin C. Superoxide dismutase (SOD) and glutathione peroxidase (GPx) belong to enzymatic antioxidants that can eliminate ROS. The followings are catalytic equations of SOD, GPx.

$$2\,O_2^{\cdot-} + 2\,H^+ \xrightarrow{SOD} H_2O_2 + O_2$$

$$2\,H_2O_2 \xrightarrow{CAT} 2\,H_2O + O_2$$

$$H_2O_2 + 2\,GSH \xrightarrow{GPx} 2\,H_2O + GSSG$$

The glutathione peroxidase system includes glutathione peroxidase, glutathione reductase, cofactor glutathione (GSH) and nicotinamide adenosine dinucleotide phosphate (NADPH). Other antioxidant enzymes, such as glutathione reductase (Gsr), NAD(P)H: quinone oxidoreductase 1(Nqo1), UDP-glucuronyl transferase (Ugt), and thioredoxin (Txn) reductase (Txrnd), sulfiredoxin

(Srx) and glutathione-S-transferases (Gsts), responsible for recycling thiols or facilitating the excretion of oxidized and reactive secondary metabolites. γ-Glutamyl cysteine synthase (γGCS), which is composed of glutamate cysteine ligase (Gcl) catalytic (Gclc) and modifier (Gclm) subunits is essential for the biosynthesis of GSH. In addition, several stress response proteins such as heme-oxygenase-1(HO1), metallothioneins and heat-shock protein also play important roles.

In cells, the production of antioxidant enzymes is regulated at a transcriptional level. Gene promoter of several antioxidants and cytoprotective detoxifying enzymes and proteins contain the antioxidant response elements (ARE) with the core sequences 5′ -TGAG/CnnnGC - 3′. One important transcription factor, that controls the ARE-mediated transcriptional response, is NF-E2 related factor 2 (Nrf2), which has been extensively studied in the past few years. Nrf2 binds to the ARE and upregulates several antioxidant genes such as NQO1, HO1, Gpx2, Trxdn…

b. NF-E2 related factor 2 (Nrf2)

Nrf2 is a cap 'n' collar basic leucine zipper(CNC-bzip) transcription factor. Nrf2 has six Neh (Nrf2-ECH homology) domains. Neh1 contains cnc-bZIP region, which is responsible for dimerization and DNA binding. Neh4 and Neh5 bind to cAMP response element (CREB)-binding protein, therefore activating transcription. Neh2 is highly conserved among vertebrate and responsible for its binding to Keap1, a Nrf2 negative regulator.

Figure 1.2: Schematic of Nrf2 protein domains

Nrf2 was first cloned due to its binding to NEF2-binding motif and potential in

erythropoiesis and platelet development (Moi P., et al,1994). However, it was found to be non-essential for blood cell differentiation, but to induce the drug metabolism enzymes including glutathione S-transferase (GST), NAD(P)H: quinone oxidoreductase 1 (NQO1) (Itoh K., et al, 1997). Promoter of Nrf2 induced drug metabolism enzymes contain the antioxidant response element (ARE) (Nguyen T. et al, 2003). The binding of Nrf2 to DNA requires the heterodimerization to activation transcription factor (ATF) and/ or AP- family of leucine zipper protein, including Jun (c-Jun, Jun-D, and Jun-B) and small Maf (MafG, MafK, and MafF).

The antioxidant defense system is regulated by Nrf2 through various mechanisms: induction of catabolism superoxide and peroxides, generation of oxidized cofactor and proteins, expression of antioxidant proteins, increasing in redox transport, metal-chelation (Ma Q, 2013). The function and structure of Nrf2 are evolutionary conserved. Compared to human Nrf2 protein, mouse Nrf2 has 82.4% homology score, rat 83%, cow 89%, dog 88.9%, chicken 67% and zebrafish 49% (Maher J. and Yamamoto M., 2010). Increased in expressing of drug metabolism enzymes by Nrf2 leads to detoxification and elimination of numerous exogenous and some endogenous chemicals.

Removing of Nrf2 in mice resulted in increased susceptibility to chemicals and pathogen with oxidative stress pathology (Motohashi H. and Yamamoto M., 2004) (Chan K. et al, 2001) (Walter DM. et al, 2008). Increasing in Nrf2 activity correlates with cellular protecting against oxidative damage (Klaassen CD and Reis M. 2010) (Talalay P. et al, 2003). High Nrf2 protein level negatively correlates with the prognostic of cancer patients (Sporn MB., Liby KT., 2012)

Transient Nrf2 activation in normal cells responding to stress brings health benefits. However, persistent Nrf2 activation in cancer cells has detrimental effects by enhancing cancer

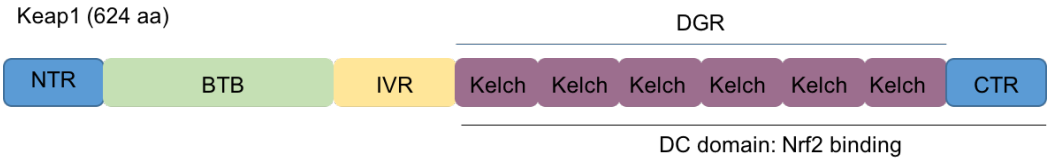
cell drug resistance, metabolic activity and aggressive tumorigenic activity. Thereby, tightly regulating is important for body homeostatic. Disturbing in this system can lead to cancer development and progression.

c. Regulation of antioxidant defense system

Important roles of Nrf2 suggest a tight regulation of its expression. However, Nrf2 mRNA expression is broad and independent from inducers. So, post-translational regulation of Nrf2 expression is required. At basal condition, Nrf2 is constantly degraded with half-life approximately 20 minutes (Itoh K. et al, 2003) (Kobayashi A. et al, 2004)

Under the normal cellular condition, Nrf2 is kept transcriptionally inactive by binding to its negative regulator Kelch-like ECH associated protein 1 (KEAP1). KEAP1 acts as an adaptor to recruit cullin3 (CUL3) to the complex. The ubiquitin-conjugating enzyme (E2) is recruited to Cul3-C terminal by RING-box protein 1. E2 catalyzes polyubiquitinations of Nrf2 protein on its lysine in Neh2 domain. Thereby, Nrf2 ubiquitination occurs, resulting in a low level of Nrf2 in many cell lines.

Figure 1.3: Schematic of Keap1 protein domains. (NTR: N-terminal region, BTB: Broad complex, Tramtrack, and Bric-a-Brac, IVR: Intervening region, DGR: Double Glycine repeat, CTR: C-terminal region)

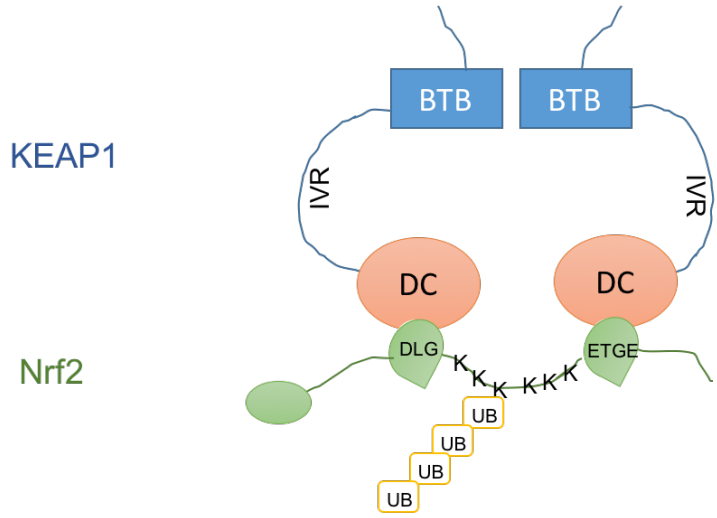


Nrf2 interacts with Keap1 at 1:2 ratio at ETGE and DLG motif (Tong KI., et al, 2006). Co-

crystallization of Keap1 DC domain (Fig. 1.3) and ETGE peptide or DLG peptide revealed the identical binding pocket (Padmanadhan et al, 2008), suggested the homodimerization of Keap1. The binding affinity of Keap1 with ETGE motif has two order magnitude higher than that with Nrf2 DLG motif (Tong KI. Et al, 2007). According to the hinge and latch model, ETGE motif of Nrf2 binds to one Keap1 in homodimer complex first, thereby promotes binding of DLG motif to other Keap1. As the results, the six lysine residues of Nrf2 face out and are exposed for polyubiquitination.

Keap1 has 4 domains: the BTB (Broad-complex, Tramtrack, and Brick à Brac), IVR (intervening region), DGR (double glycine region) and CTR (C-terminal region). The combination of DGR and CTR forms DC domain with β-propeller structure, which interacts with Nrf2. Keap1 dimerization crystallized structure was reported in 2010 by Ogura and colleagues. Keap1 dimerization roughly resembled a pair of cherry with two spherical bodies connected by a thin linker. The NTR and BTB form the stem and linker structure which occupy 13.5% of the whole volume. The globular body occupies 86.5% and consists of IVR and DC domains (Ogura et al., 2010).

Figure 1.4: Schematic describes interaction of Keap1 and Nrf2



Keap1 is a cysteine-rich protein, with 27 residues in human case. Keap1 cysteines are easily modified by oxidants and electrophiles in vitro. Among those, C151, C272, and C288 have been shown to be important for Keap1 functions by alternating its conformation (Taguchi K., et al, 2011). Two critical IVR cysteines are C272 and C288. Due to its proximity to DC domains in a globular structure, covalent modification of cysteines in IVR may affect the structural integrity of DC domain, which eventually alters its interaction with Nrf2. Cysteine C151 in BTB domain may important for the interaction with E3 ligase.

In response to electrophiles and ROS, which have the ability to interact with a sulfhydryl group, Keap1 cysteines are covalently modified and its mediated repression of Nrf2 is diminished. Thereby, Nrf2 accumulates and translocates into the nucleus, where it activates ARE- genes. The Nrf2-Keap1 system plays an important role in response to electrophiles and oxidative stress due to Keap1 high sensitivities to a low amount of electrophile. Keap1-deficient mice have constitutively active Nrf2 and hyperkeratosis in the stomach and the esophagus, which leads to postnatal lethality (Wakabayashi N., *et al*., 2003). The effect of Keap1 knockout was nullified by the knockout of Nrf2. This indicates the Nrf2 constitutive activation is the main cause in this lethal phenotype.

d. Perturbation in SCC

Evidence was reported and indicated the contribution of Keap1/ Nrf2 pathway dysfunction in the development of squamous cell carcinomas of the esophagus (cancer genome atlas, 2017), lung (Cancer Genome Atlas, 2012), head and neck (Cancer Genome Atlas, 2015). Keap1, Nrf2 and its target genes are positively associated with prognostic with higher expression

compared to adjacent normal tissue in squamous cell carcinomas of the larynx (Li et al, 2016), oral (Huang CF., et al., 2013), head and neck (Solis LM., et al.,2010). Alteration of Keap1 and Nrf2 expression may cause by mutations in either Keap1 or Nrf2, thereby affects this interaction and leads to the accumulation of Nrf2 as well as its target genes.

Keap1 mutations were found in the N-terminus or BTB domain which affect Keap1's ability to bind with Cullin 3 and ubiquitination complex. Others mutations occur at Kelch domain that disrupt Keap1 and Nrf2 interaction, were also reported (Padmanabhan B., et al, 2006). In LSCC, Yoo et al. reported that among 49 specimens, there were 6% mutation located at N-terminal and Kelch domain (Yoo et al, 2012). Consistently, Jeong et al. reported 6 out of 42 specimens bared Keap1 mutation (16%) (Jeong et al., 2017). In ESCC, 2.9% of Keap1 mutations among 490 tumors was reported by Du et al (Du P. et al, 2017).

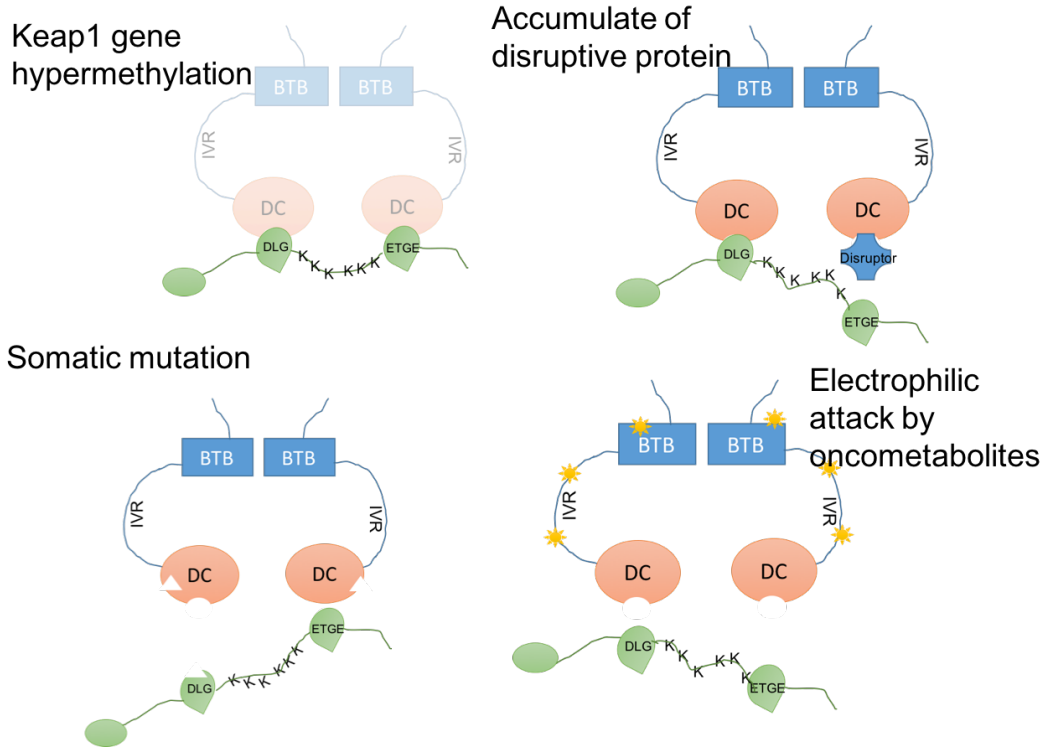
Two hundred seventy-four coding mutations were reported in Keap1 and 389 were reported in Nrf2 (COSMIC, http:// cancer.sanger.ac.uk/cosmic). Nrf2 mutations were reported exclusively in ETGE and DLG domains. Nrf2 mutations were reported in 15% mutation in LSCC (Cancer Genome Atlas, 2012). Analyzing 1145 cancer tissues from carcinomas of the oesophagus, skin, uterine cervix, lung, larynx, breast, colon, stomach, prostate, bladder, ovary…, Kim et al reported 95.5% of found Nrf2 mutations occurred in squamous cell carcinomas (Kim YR., et al, 2010). This data indicated that Nrf2 perturbation could be considered as SCC feature.

By immunohistochemistry, Kim and colleagues reported the Nrf2 expressing in the nucleus of 50% of the ESCC and 47% of the skin SCC. However, among 35 ESCC analyzed specimens, only eight (22.8%) has the Nrf2 mutations. The other 27 (77.2%) have neither Keap1 nor Nrf2 mutations, this indicated the unidentified mechanism of Nrf2 upregulations (Kim YR. et

al, 2010). Similarly, 50% of gastric cancer have neither Keap1 nor Nrf2 mutations but displayed high Nrf2 expression in the nucleus (Yoo et al, 2012).

In fact, there are several mechanisms underlying the constitutively active of Nrf2 including epigenetic silencing of Keap1 expression, transcriptional induction of Nrf2 by K-Ras, post-translational modification of Keap1 and accumulation of disruptor protein such as p62.

Figure 1.5: Mechanism for constitutive accumulation of Nrf2. (1) Hyper-methylation of Keap1 promoter leads to reduction of Keap1 mRNA. (2) Accumulation of disruptor protein which decrease Keap1 and Nrf2 interaction and results in Nrf2 accumulation. (3) Somatic mutations of Keap1 or Nrf2 protein which disrupt their interaction. (4) Electrophilic attack by oncometabolites caused structural change in Keap1.



p62 is a reported disruptor protein of Nrf2 and Keap1 interaction. P62 is a stress-inducible protein with multiple functions in the regulation of cell growth, survival, and proliferation via its

ability to interact with several key proteins of various signaling. In non-canonical Keap1-Nrf2 pathway, p62 competitively binds to Keap1, which leads to Nrf2 accumulation (Katsuragi Y et al, 2016). Several reports discussed on P62 accumulation in liver cancer causing hyper-proliferation of cancer cells and aggressive carcinomas phenotype (Xiang X., et al, 2017; Kessler Sm., et al, 2015). P62 induces chromosomal instability in hepatocellular carcinomas through reactive oxygen species (ROS) production (Kessler Sm., et al, 2015), which in turn leads to more undifferentiated tumors.

In cancer cells, p62 is accumulated in the cytoplasm, together with enriched Nrf2 protein level. Notably, p62 is Nrf2 target gene. The question arises regarding the cause of p62 accumulation, whether p62 is only a consequence of Nrf2 positive feedback loop. The research in our laboratory focuses on a novel disruptor protein TSC22D4 (THG-1) with reported ability to interacts with Keap1 at ETGE motif, suggested the competitive binding to Keap1 against Nrf2 (Hast BE. Et al, 2013 and Tamir TY., et al, 2016). Unlike p62, THG-1 is not a target of Nrf2.

3. *TSC- homologous gene 1 (THG-1/TSC22D4)*

TSC-22 family is an evolutionarily conserved cytoplasm protein including TSC22D1, TSC22D2, TSC22D3 and TSC22D4, which is characterized by conserved TSC-box and a leucine zipper motif in the middle. TSC-22 family is reported interacting with other proteins involving in many cellular processes (Liang F. et al, 2016). TSC-22 (TSC22D1.2) interacts with TβR1, thereby inhibits Smad7/Smurfs induced ubiquitination and degradation of the receptor (Yan X. et al, 2011). Interaction between TSC22D1.2 and p53 was reported to prevent p53 ubiquitination (Yoon C.H. 2012). TSC22D2 was reported to interact with PKM2 and inhibits colorectal cancer cell growth (Liang F., et al,2016). TSC22D3 interacts with NF-κB and inhibits T-cell activation

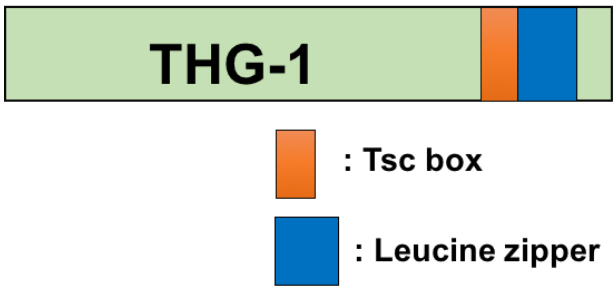
(Ayroldi E., et al. 2001). TSC22D3 also reported to interact with Ras (Ayroldi E., et al, 2007) and p53 (Ayroldi E., et al, 2015). Even though the TSC-22 family has the ability to interact with several proteins as well as macromolecule complexes, the full understanding of the interactome remains to be elucidated (Liang F., et al. 2016). There is only one paper reported the interactome of TSC22D2 (Liang F., et al. 2016), none for TSC22D1, TSC22D3, TSC22D4. Even in TSC22D2 case, more experiments for characterizing and validating the function of interactions need to be carried out.

TSC22D4 (THG-1) is a member of the TSC22 family. THG-1 was first identified as TSC22 interacting protein and proposed to have transcriptional functions (Krester H., et al, 1999). However, THG-1 does not contain DNA binding regions. THG-1 was suggested to regulate gene expression by homo- and/or hetero-dimerization with others leucine zipper-containing factors. THG-1 was found to upregulate in cancer cachexia, hyperosmotic stress in murine inner medullary collecting duct cells, and various SCC specimens (Figure1.7 A). THG-1 knock-down in SCC cell lines reduced their proliferation, invasion, and tumorigenesis (Figure 1.7 B, C). Except for heterodimerization with TSC22, no further details clarify THG-1 molecular mechanism.

a. Structural information

THG-1 is encoded by a gene on chromosome 7 with 395 amino acids. Similar to other TSC22 members, THG-1 structure includes a leucine-zipper domain and TSC box.

Figure 1.6: Schematic on THG-1 structure.



Leucine zipper domain is a common three-dimension structural motif with alpha- helix conformation with leucine residue at every seventh position. Leucine zipper domain is the heterodimerization domain of the bZIP (basic-region leucine zipper) class of transcription factors. However, leucine zipper was also reported in proteins other than transcription factor, suggested its involvement in protein-protein interaction modulation (Hakoshima T., 2014). Dimerization of protein through leucine zipper domain forms by the interaction of the hydrophobic core involving leucine side chain.

The leucine zipper domain in THG-1 may be important for its homo/hetero-dimerization with other proteins containing leucine zipper domain. Although the leucine zipper suggests THG-1's ability to heterodimerize, the specific partners are decided by other residues, which interfere with the electrostatic interaction between two helices (Hakoshima T, 2014).

Tsc box is the domain adjacent to leucine zipper domain. It was reported to be important for the nuclear translocation and TSC22 function in inhibiting Ras-transformation (Nakamura M., te la, 2012) and in suppressing of cell division (Hashiguchi A., et al, 2007).

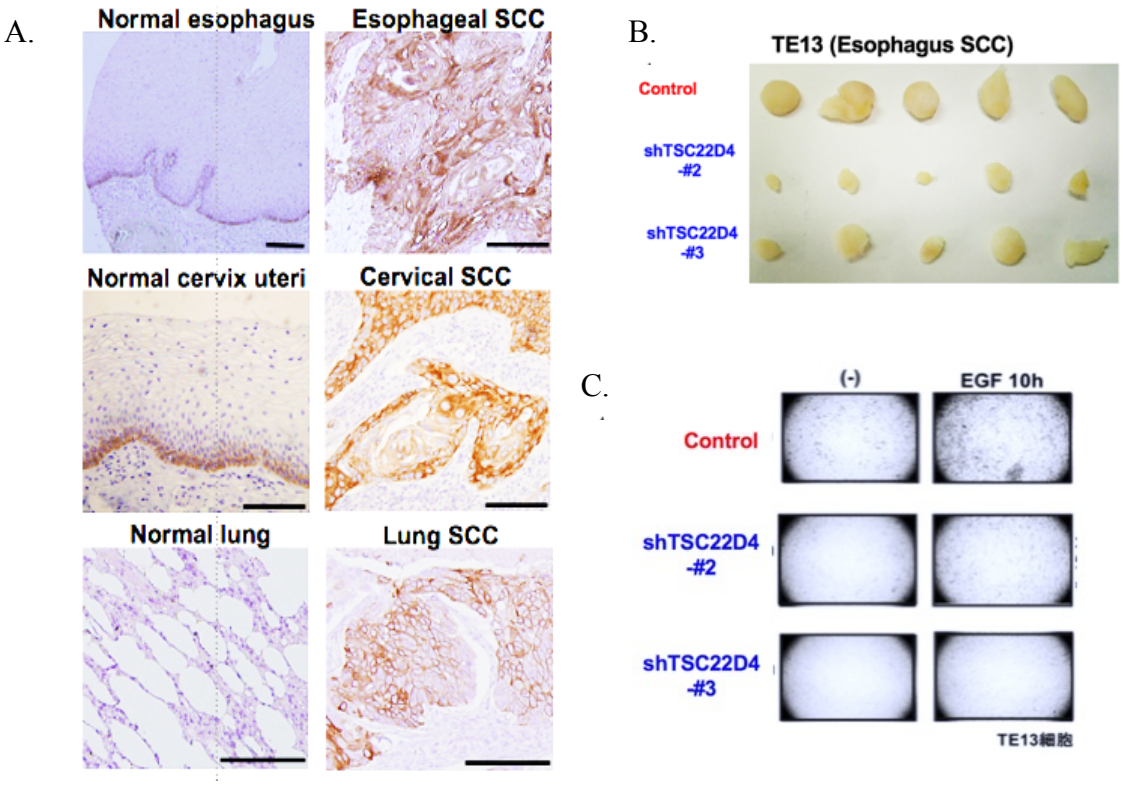
The ETGE motif is most frequently occurs in Keap1 binding proteins (Hast B. E. et al, 2013). THG-1 is reported to contain the ETGE motif and found to interact with Keap-1, thereby strongly activate Nrf2-mediated transcription (Hast B. E. et al, 2013). Consistent with that, our data indicated that THG-1 is highly expressed in SCC and sequestered Keap1 from Nrf2. By this

mean, Nrf2 activation is prolonged in SCC and leads to advantages of cancer cells in surviving and development (Figure 11).

b. Pathology distribution and implementation in cancer

THG-1 was found to be upregulated in cancer cachexia (Jones A., et al, 2013), hyperosmotic stress in murine inner medullary collecting duct cells (Fiol DF. Te al, 2007), and various SCC specimens (Suzuki H., et al, 2018). THG-1 expression is strictly controlled in the basal layer of the squamous tissue (Figure 1.7A, left panel). The basal layer of the squamous epithelium is an important reservoir for stem and proliferating cells, THG-1 special spatial pattern implies its importance in the regulation of cell proliferation and maintenance of homeostasis. Interestingly, THG-1 expression is broadly expressed in squamous cell carcinoma (Figure 1.7A, right panel). THG-1 expression is consistent with the pathology structure of ESCC where basal cells lose their ability to differentiate properly and fill the entire epithelial compartment [Gabber HE et al, 2000]. THG-1 knockdown in TE13 (ESCC cell line) caused a reduction in cell growth and ability to form xenograft tumor (Figure 1.7B). Additionally, reduced invasive ability had been observed in THG-1 knockdown cell line (Figure 1.7C)

Figure 1.7: THG-1 has an important role in cancer cell proliferation, tumorigenesis, and invasiveness. (A) Immunohistochemistry staining against THG-1 (B) Tumors formation of TE13 and THG-1 KD TE13 cells in xenograft mice. (C) Invasion of TE13 and TSC22D4 KD TE13 cells.



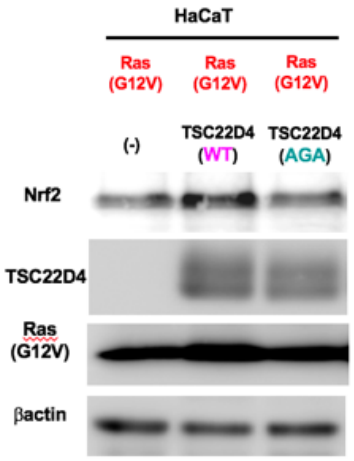
Suzuki H., et al. unpublished data

c. Mechanism of action and potential function

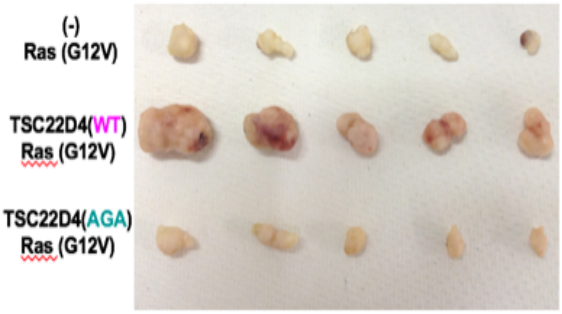
THG-1 regulates oxidative stress signaling through direct binding with Keap1, reducing its availability to Nrf2. Both our data and Hast B.E et al data indicated Nrf2 medicated transcription activation in the presence of THG-1 (Hast B.E. et al, 2013). THG-1 mutation at Keap1 binding motif ETGE impairs not only cancer cells' growth, but also xenograft tumor development significantly (Figure 1.8).

Figure 1.8: Role of THG-1 in Nrf2 accumulation and tumorigenesis (A) Nrf2, THG-1 expressing in HaCaT cells expressing Ras(G12V), THG-1 plus Ras(G12V), and THG-1(AGA) plus Ras(G12V). (B) Tumor formation of Ras(G12V), THG-1 plus Ras(G12V), and THG-1 (AGA) plus Ras(G12V) cells in nude mice.

A.



B.



Suzuki H., et al. unpublished data

4. *THG-1 as a therapeutic target in squamous cell carcinomas*

a. Reduction of Nrf2 levels in Nrf2 addicted cancer

Inhibition of Nrf2 signaling is a promising therapeutic approach for the treatment of Nrf2-addicted cancer type. Inhibiting of Nrf2 protein production, Nrf2 medicated transcription or restoring Keap1 function are main approaches.

Nrf2 inhibition:

Isolated from *Brucea javanica* plant, Brusatol is a type of degraded diterpenoid used in the treatment of cancer, amoebic dysentery, malaria. Brusatol acts by inhibiting DNA, RNA and protein synthesis. In a recent report by Ren et al, brusatol enhanced Nrf2 ubiquitination without changing Keap1 protein level, likely due to the modulation of ubiquitination-mediated protein degradation system (Ren D. et al, 2010)

Halofuginone is a less toxic derivative of febrifugine, which is isolated from Chinese herb Dichroa febrifuge. Upon the treatment of halofuginone, Nrf2 mRNA did not change but its

protein level was significantly decreased. By inhibiting prolyl-tRNA-synthetase, halofuginone inhibits Nrf2 protein synthesis. Even though halofuginone is not specific to Nrf2, a strong effect was shown on Nrf2 due to its short half-life and fast turn-over (Tsuchida K., et al, 2016)

ML385 is a small molecule that selectively binds to the Neh1 domain of Nrf2, thereby inhibiting its interaction with DNA and sMAF protein (Singh A. et al, 2016). ML385 is used together with chemotherapy like Doxorubicin or Taxol and aimed for Nrf2- addicted non-small cell lung cancer with Keap1 mutations.

Target Keap1- Nrf2 disruptor proteins: p62, THG-1

Nrf2 is an attractive target for cancer therapy. However, it can be bifunctional: Nrf2 acts as a tumor suppressor protects normal cells from carcinogens and Nrf2 as oncogene promotes the drugs resistance of cancer cells. As, Nrf2 is not specific to cancer cells, either inhibiting or inducing its functions can lead to severe side effects. In this regards, it should be concerned that the long-term application of Nrf2 inhibitor may increase body susceptibility to oxidative stress and cancer development.

Recover of Keap1 expression, or disruption of disruptor proteins (such as p62) could be considered a better strategy for specific modulating Nrf2 pathway. In 2017, Wang L. and his colleagues reported the use of Verteporfin for inhibition of p62 function, therefore inhibition of Nrf2 constitutive activation (Wang L., et al, 2017). Verteporfin, a benzoporphyrin derivative, was approved by the FDA for photodynamic therapy in eliminating abnormal blood vessel in the eyes. Verteporfin induces the formation of p62 crosslinked, which losses the ability to bind to ubiquitinated protein (Donohue E., et al, 2014). In their paper, Wang L. et al reported the

inhibition of Nrf2 pathway through the aggregation of p62 both in vitro and in xenograft tumor tissues of prostate cancer (Wang L., et al, 2017).

However, the clinical role of p62 in SCC and its correlation with Nrf2 are unclear. In 2012, Inoue D. and colleagues reported the independent prognostic role of p62 and Nrf2, where Nrf2 status strongly was associated with clinical outcome of both adenocarcinomas and squamous cell carcinomas of non-small-cell lung cancer but p62 was only found to be significantly correlated with adenocarcinomas prognostic. Additionally, Nrf2 and p62 were reported to accumulate in 34% and 37% of NSCLC respectively (109 specimens), the accumulation of them is not well correlated (Inoue D., et al, 2012). They suggested that there was an unknown mechanism causes Nrf2 stabilization, as Keap1/ Nrf2 mutations were not detected, Keap1 expression was not repressed and only Nrf2 was stained but not p62 (Inoue D., et al, 2012).

Unlike p62, THG-1 was found to be accumulated in more than 80% of ESCC, LSCC, and cervical SCC but did not show in adenocarcinomas counterpart (Suzuki H., et al, unpublished). The disruption of THG-1 and Keap1 interaction can serve as therapy for squamous cell carcinomas patients with Keap1 expression but neither Keap1/Nrf2 mutations or p62 accumulation.

b. Targeting THG-1 protein-protein interaction and current screening system

i. *Targeting protein-protein interaction*

Estimated, there are 150,000-500,000 interactions in human cells (Venkatesan K et al, 2009). Understanding of interaction networks allows us to learn about the function of proteins and their evolution in the involving networks. To elucidate the molecular pathways associated with PPI networks, thousands of protein-binding and inhibiting reagents are required. However,

PPIs are considered as challenging targets. Binding interfaces of PPIs are often large, flat and featureless. The contact surfaces of PPIs are around 1500 to 3000 $Å^2$ (Jones S., et al, 1996) which are difficult to be cover by small molecules with an average coverage of 300-1000 $Å^2$ (Cheng AC., et al, 2007). The interfaces of PPIs often dominate by hydrophobic patches and lack of substantial grooves or deep pockets, those characteristics are not favorable for small molecules binding (Miller S., 1989).

So far, PPIs have been mostly targeted by therapeutic proteins. For instance, Durvalumab, currently in phase 3 clinical trial, inhibits the interaction between its target PD-L1 with PD-1 and CD 80. Thereby, Durvalumab may help to overcome the immunosuppressive effect of PD-L1 (Reichert J. M., 2017). Olaratumab is approved by the FDA in Oct 2016 in conjunction use with doxurumicine for soft-tissue sarcoma. By blocking the interactions of subunit alpha of platelet-derived growth factor receptors, Olaratumab inhibits tumor cells growth (Teyssonneau D. and Italiano A., 2017).

However, the large size of antibodies usually causes their limited tissue penetration and prevents them from reaching intracellular targets. In addition, proteins therapeutic are usually administrated parentally due to their poor stability. There is an obvious need for novel-drug classes that can target intracellular PPIs with comparable pharmacological properties to antibodies but with enhanced permeability and stability. With the limitations of above-mentioned drugs modalities, attention was turned to new chemical space from 700 to 1900 Da, in which peptide fall into the category (Josephson K., et al, 2014).

In general, peptides are selective and efficacious ligands that bind specifically to their targets. This property improves their safety, tolerability and efficacy profiles in humans.

Compared to antibodies, peptides are much smaller, thereby exhibit lower immunogenicity. On the other hand, the cost and complexity for the production of peptide therapeutic are also lower. In addition, peptides are amiable to a chemical modification to further increase their pharmacological properties such as permeability, stability, specificity…. In their review in 2013, Kaspar and Reichert stated that there were more than 60 FDA-approved peptide medicine, approximately 140 peptide drugs in clinical trials and more than 500 in preclinical development. However, linear peptides face significant challenges including short half-life and fast elimination, orally unavailable, low membrane permeability (Fosgerau K. and Hoffmann T., 2015)

ii. *Macrocyclic peptides*

Natural peptide products, that possess biological functions, usually have the macrocycle structure. Approximately, 3% of more than 100,000 natural-product secondary metabolites are macrocycles (Wessjohann, L.A. et al, 2005). Analyzing of natural-derived drugs, that were approved from 1976- 2006, showed a high proportion of macrocyclic products. This indicated their importance for drugs discovery and development (Ganesan A., 2008). From 2006- 2015, there were nine approved cyclic peptide drugs which account 3% of marketed drugs during the time (Zorzi A. et al, 2017).

Firstly, cyclization decreases the conformational freedom of the peptide, thereby improves their specificity and affinity (Drigger E.M., et al, 2008). There have been successful examples of macrocyclic peptide drugs to address challenging target like protein-protein interactions. For instance, cyclosporine A exhibits an ability to block cyclophilin A and calcineurin interaction, thereby inhibits calcineurin activity and decreases the production of cytokines (Matsuda S, Koyasu S, 2000).

Second, macrocyclic peptides are a less susceptible substrate for proteases, thereby have higher stability. In 2000, Fairlie D. P. et al cyclized the substrates of HIV-1 protease and concluded that cyclization significantly decreases catalytic activities of protease on those substrates. Macrocyclic peptides do not fit into endopeptidase active sites of proteases, which usually refer to a β-sheet conformation (Tyndall, J. D. A., Nall T., and Fairlie D. P., 2005).

Thirdly, cyclization increases the internal hydrogen bonds formation, thereby enhances the permeability of macrocyclic peptides. In general, internal hydrogen bonds reduce the energy cost for desolvating the peptides during insertion into the membrane (Rezai T., et al, 2006). Internal hydrogen bonds could be used to mask the hydrogen bond donor, shield polarity from the surrounding environment, therefore increase macrocyclic peptides permeability (Over B., et al, 2014).

Despite those advantages, macrocyclic peptides have been underexplored for the discovery of novel drugs. Majority of current macrocyclic peptide drugs are derived or based on natural sources. However, there is a reluctance in an investigation of natural products. The structural complexity of naturally derived substances causes difficulties in analog synthesis. For the last decade, many researchers have been synthesized and screened large libraries of a natural product-like macrocyclic peptide to cover this class of therapeutic drugs and to meet the medical needs (Table 1). Thus far, only one in nine approved cyclic peptides, peginesatide, was the first to be develop *de novo* (Zorzi A. et al, 2017).

Macrocycle-based library approaches	
Approach	Company/organization
Cyclopeptides: MATCH cyclopeptides	Tranzyme Pharma (now Ocera Therapeutics)
Protein epitope mimetics (PEM)	Polypor
Stapled peptide	Aileron Therapeutics
DNA-encoded cyclopeptides	Ensemble Therapeutic
RNA-encoded cyclopeptides	PeptiDream
	Ra pharmaceutical
Bicyclic peptides	Bicycle Therapeutic
	Pepscan
Nacellins	Encycle therapeutics
Cell-permeable cyclopeptides	Circle Pharma
SICLOPPS	University of Southampton
Cyclic cell-penetrating peptides	The Ohio State University
Phylomer peptides	Phylogica
Non peptide-based macrocycles: DOS macrocycles	Broad Institute
MacroFinder	Polyphor
Nanocyclix	Oncodesign
CMRT	Cyclenium Pharma

(Zorzi A. et al, 2017)

c. Random non-standard Peptide Integrated Discovery (RaPID) system

i. *General introduction*

Recent year, many researchers have developed several platform technologies to screen a large combinatorial library of cyclic peptides against proteins of interest such as mRNA display, phage display, one bead-two compound, diversity-oriented synthesis (Quian Z., et al, 2017). Those high through-put methods decrease cost and time to discover lead compounds for drug development. Among those, mRNA-display is an *in-vitro* selection technology based on linkage between genotype (mRNA) and translated phenotype (peptide). Selected peptides can be easily identified by sequencing encoded nucleic acid tag. mRNA-display has been introduced in 1997 (Roberts, R.W. and Szostak, J. W.) and successfully applies until now due to their advantages in library size (~10^{10}), the ease of library generating and reduced time-consuming.

However, the usage of mRNA-display was limited by the usage of canonical amino acids as well as the technical difficulty of cyclization. Hence, various technologies were developed to incorporate non-standard amino acids on to peptides and to cyclize peptides for selection. Disulfide bond formation, Cl-acetyl thioalylation, Dehydroalanine Michael addition, Dibromoxylene thiolakylation, "Click" cyclioaddition and NHS-disuccinimidyl glutarate crosslinking are six among cyclization strategies that compatible with mRNA display (Josephson K., et al, 2014). Notably, Suga H. and colleagues have developed and successfully implemented in numerous projects, the method for spontaneous generating of stable thioether- cyclization macrocyclic peptides (Goto Y., et al, 2008). This system, Random non-standard Peptide Integrated Discover (RaPID), is based on the ability of RNA "flexizyme" to reprogram the genetic

code for the incorporation of non-canonical amino acids in combination with mRNA-display (Ohuchi M., et al, 2007, Goto Y., et al, 2011, Morimoto J., et al, 2011).

Flexizymes are de novo tRNA acylation ribozymes that capable of charging any amino acid onto desired tRNAs regardless of the anti-codon and thus result in desired tRNAs charged with non-proteinogenic amino acids. Prepared tRNAs are combined with *in-vitro* translation system, which composes of the purified ribosome, essential recombinant translation factors, tRNAs, ribonucleotide triphosphates, amino acids, and other small organic and inorganic molecules required for translation. Cognate aminoacyl-tRNA synthetase and certain amino acid are excluded to produce vacant codons. The vacant codons then can be reassigned with prepared corresponding tRNA charged with a non-proteinogenic amino acid. Using flexizymes, genetic code reprogramming is achieved and the translation process is called "flexible *in vitro* translation (FIT)".

The RaPID system is the integration of the FIT system and mRNA-displayed for the screening of peptide ligands with macrocyclic scaffold and desired modified amino acids. The RaPID system can be used and was used to identify bioactive natural product-like peptidomimetics.

ii. *Advantages*

The RaPID system allows direct incorporation of non-proteinogenic amino acid during the selection. Incorporation of non-proteinogenic amino acid into the sequence of peptides further enhances their resistance against proteases. In addition, the modification of peptides backbone, achieved by non-proteinogenic amino acid such as N-methylation, can result in reducing the

number of hydrogen-bond donors. As mentioned earlier, reducing in hydrogen-bond donor often aid benefit in cellular permeability and oral availability.

With the power of the FIT system for genetic code reprogramming, libraries that contain an enormous number of unique non-proteinogenic containing peptides (billions to trillions) can be created. The high diversity of the library increases the chance to identify bioactive compounds with desired pharmaco-properties.

The thioether macrocyclization occurs spontaneously after translation without the requirement of other treatment. Therefore, the preparation of macrocyclic peptide can be achieved easily with almost 100% efficiency. The RaPID system provides a robust and fast platform to identify the macrocyclic peptide ligands for the protein of interest.

iii. Reported studies

With the use of the RaPID system, researchers can directly select for macrocyclic peptide against targets of interests. Thereby, ring-size, confirmation, and sequences are optimized to acquire high affinity and specificity ligands. This system speeds up the drug discovery process and has been reported for successful identification of macrocyclic peptide ligands for cocrystallization (Tanaka Y., et a, 2013, Hipolito C.J., et al, 2013), and high affinity bioactive compounds against VEGFR2 (Kawakami, Ishizawa et al. 2013), SIRT2 deacetylase (Yamagata, Goto et al. 2014), and E6AP (Yamagishi, Shoji et al. 2011), Akt (Hayashi Y. et al, 2012) , and histone demethylase (Kawamura A. et al, 2017).

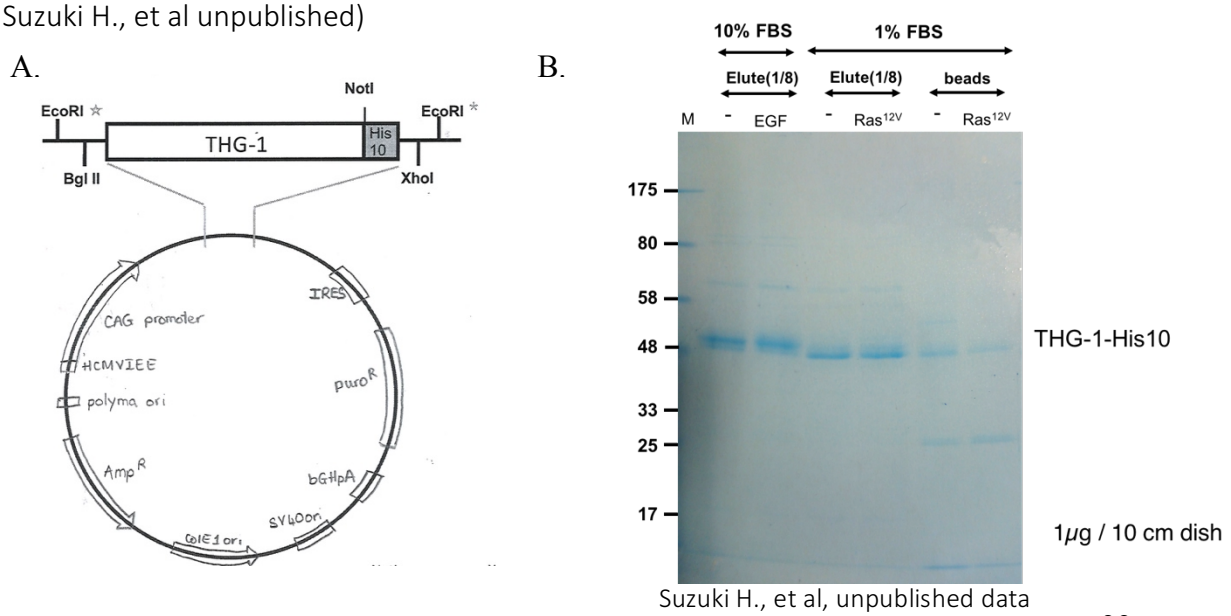
CHAPTER II: PRODUCTION OF THG-1 PROTEIN

1. *General introduction*

Production of high-quality THG-1 protein is critical for the success of the macrocyclic peptide selection against this protein. Each round of selection requires 100 pmol of protein molecules, with around 7 rounds for each selection attempt and several single clone assays. Around 500 μg of highly purified THG-1 protein is required to complete a selection attempt.

THG-1 cDNA was previously extracted from TE13 cell line and cloned into pCAGIP-FL-THG-1-His10 vector by Dr. Suzuki H (Figure 2.1 A). The plasmid was expressed in 293T cells and protein was purify by Ni^{2+} purification with the estimated yield of 1 μg/ 10 cm culture dish (Figure 2.1 B). In the consideration of this low yield and the expense of mammalian cell culture, we decided to switch into an *E.coli* expression system.

Figure 2.1: **THG-1 cloning and expressing in the mammalian cell line.** A. Plasmid structure of THG-1 expressing vector pCAGIP-FL-THG1-His10. B. Coomassie blue stained SDS-PHAGE gel of purified THG-1 from 293T cell transfected with pCAGIP-FL-THG-1-His10. (data adapted from Suzuki H., et al unpublished)



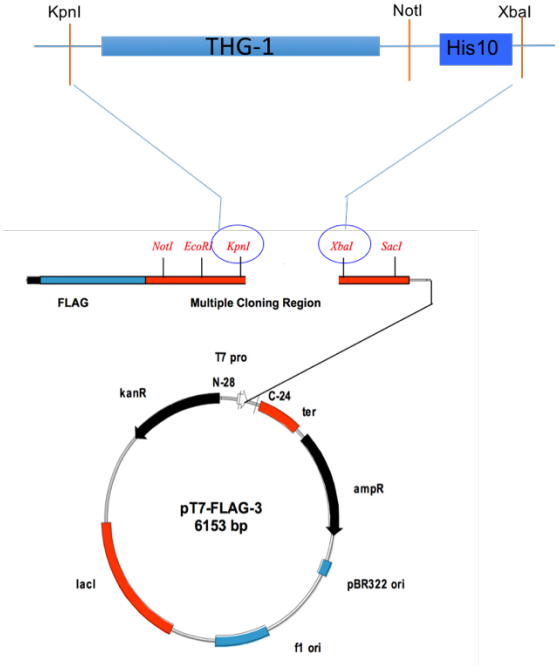
Suzuki H., et al, unpublished data

E. coli can be easy to culture in large amount with inexpensive and simple media. Additionally, *E. coli* can grow, multiply very rapidly and produce heterologous protein. Recombinant protein production includes five steps: isolation of gene of interest, insertion of the gene into expressing vector, the transformation of the cells and selection of genetically modified organism, protein expression, and protein extraction.

2. *Plasmid construction*

THG-1 encoded gene was PCR amplified from the pCAGIP-FL-THG-1-His10 plasmid with the forward primer KpnI-THG1 and the reverse primer XbaI-THG1. KpnI-THG1 was designed complemented with 5′ end of THG-1 sequence with *KpnI* DNA sequence added at 5′ end of the primer. XbaI-THG1 was designed complemented with His10 sequence with *XbaI* DNA sequence added at 5′ end of the primer. The resulting PCR product was digested with *KpnI* and *XbaI,* subsequently cloned into pT7-FLAG-3 plasmid cut with same restriction enzymes. Figure 2.2 depicted the process of cloning THG-1 encoded gene into pT7-FLAG-3 plasmid.

Figure 2.2: structure of THG-1 expressing plasmid pT7-FLAG-3-THG-1 using *E. coli* system.

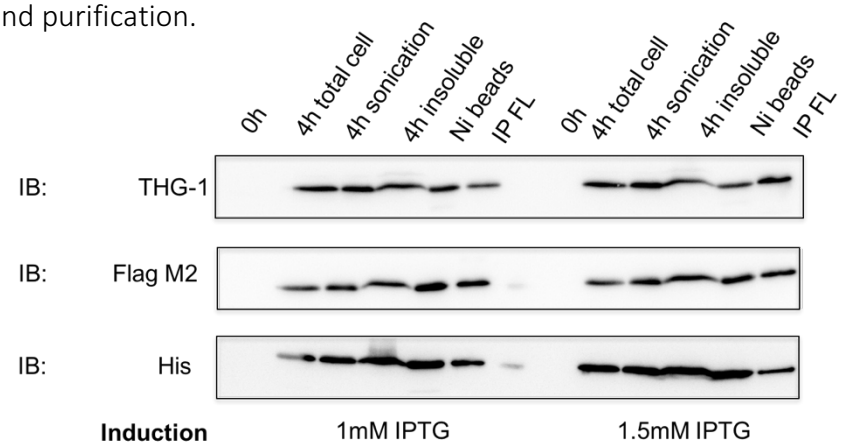


Because there is only 2-3 base at the end of PCR products, this may reduce binding of restriction enzyme. Therefore, in order to enhance efficiency in cutting the PCR gene fragment with *KpnI* and *XbaI* restriction enzymes, PCR product was cloned into T-easy vector and amplified in DH5α *E. coli.* Even though pT7-FLAG-3 plasmid was cut with two different restriction enzymes, plasmid self-ligation occurred with a high frequency which resulted in no insertion of the gene of interest. To overcome this problem, the cut plasmid was treated with Alkaline phosphatase to remove phosphate monoester prior to the ligation.

3. Transformation and expression

pT7-FLAG-3-THG-1 plasmid requires *E. coli* cells containing the DE3 lysogen as a source of T7 polymerase. BL21 (DE3) was selected as an expression host for THG-1 protein. Western-blot analysis of SDS-PAGE gel of cells lysate of BL21(DE3) transformed with pT7-FLAG3-THG-1 plasmid confirmed the expression of THG-1 protein in these host cells. The expression was not different between induction with 1 mM or 1.5 mM of IPTG and THG-1 protein can be pulled down with either Ni^{2+} beads or Flag antibody (Figure 2.3)

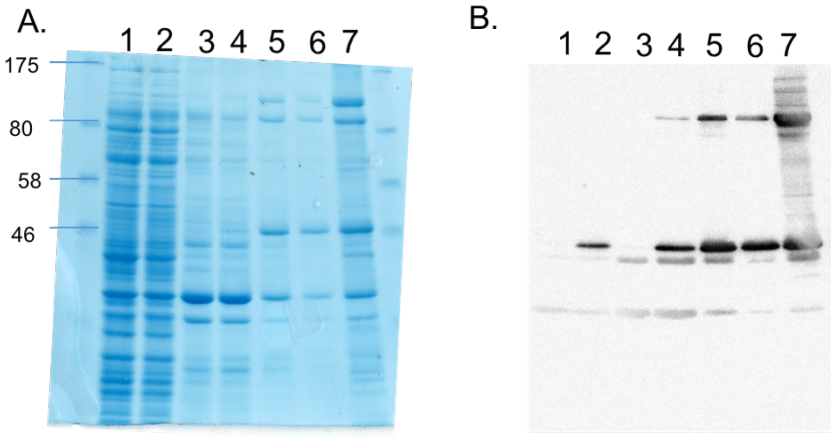
Figure 2.3: pT7-FLAG3-THG-1 expression in BL21 (DE3) *E.coli.* BL21 (DE3) was transformed with pT7-FLAG3-THG-1 plasmid until OD_{600} reach 0.5. 1 mM and 1.5 mM IPTG was used for protein expression induction. Samples were collected at different stages of the protein extraction and purification.



4. *Purification of the expressed protein*

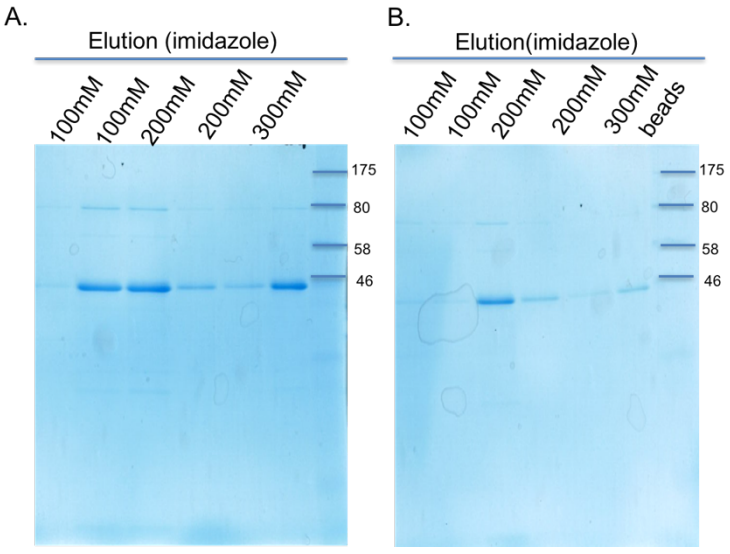
After protein induction for 4 hours, cells were collected and disrupted by sonication. The cell lysate was incubated with Ni^{2+} beads and low amount of imidazole (50 mM) for removing non-specifically bound protein. The purification was successful, however, collected THG-1 protein was fragmented and the productivity was low with only 30 µg/500 ml culture (Figure 2.4).

Figure 2.4: Purification of THG-1 produced in BL21 (DE3). A. Coomassie brilliant blue (CBB) staining of the SDS-PAGE gel of the samples collected after different steps of purification process. B. Western-blot analysis of SDS-PAGE of the same samples in CBB stained gel. Samples (1) cell lysate before induction, (2) cell lysate after 4 hours of induction, (3) elution from Ni^{2+} beads with 100 mM imidazole, (4) 1^{st} elution sample from Ni^{2+} beads with 200 mM imidazole, (5) 2^{nd} elution sample from Ni^{2+} beads with 200 mM imidazole, (6) elution from Ni^{2+} beads with 300 mM imidazole, (7) remaining protein on beads after elution.



Analyzing of THG-1 protein sequences revealed that 8% of its coding sequences were rare codons in *E. coli*. BL21-Codon Plus (DE3)-RIL and BL21-Codon Plus (DE3)-RP were tested for expression. With same culture and induction condition as BL21 (DE3), BL21-Codon Plus (DE3)-RP showed higher production yield (300 µg/500 mL culture) with higher purity (Figure 2.5).

Figure 2.5: Purification of THG-1 produced in BL21-Codon Plus (DE3)-RP and BL21-Codon Plus (DE3) –RIL. Coomassie blue (CBB) staining of the SDS-PAGE gel of samples collected after different steps of the purification process. A. BL21-Codon Plus (DE3)-RP (extra copy of tRNA gene argU (AGA, AGG), proL (CCC)). B. BL21-Codon Plus (DE3)-RIL (extra copy of tRNA gene argU (AGA, AGG), ileY (AUA), leuW (CUA)). Elution buffer containing different amount of imidazole (100 mM x 2 times, 200 mM x 2 times, 300 mM).

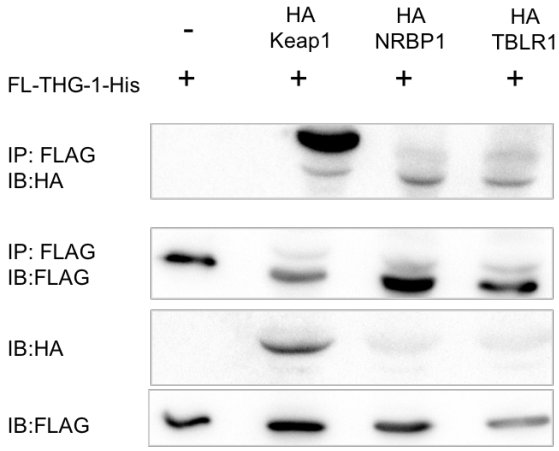


5. *Protein-interactions of purified recombinant THG-1 protein*

THG-1 protein produced in *E. coli* may lack of post-translational modifications which may be important for its protein-protein interactions. Purified THG-1 protein was mixed with 293T cell

lysate overexpressed with Keap-1, PHD2, NRBP1 or TBLR1. Results indicated that purified recombinant THG-1 protein can interact with those proteins. (Figure 2.6).

Figure 2.6: Co-immunoprecipitation of purified recombinant THG-1 protein with some of its identified interacting proteins including Keap1, NRBP1, PHD2, and TBLR1.



6. *Conclusion*

A critical requirement for a successful selection of active ligands against a particular protein of interest is obtaining high quality and quantity of the targeted protein. By experimenting with both mammalian and *E. coli* system, we demonstrated the advantage of *E. coli* system in producing recombinant THG-1 protein with a yield of 300 μg/ 500 mL of culture. Purified THG-1 protein showed interactions with its identified interacting proteins: Keap1, NRBP1, PHD2, and TBLR1. Different cell types, condition for cells culture, protein induction and procedure of purification were evaluated and the best combination was reported.

7. *Materials*

Primers:

KpnI-THG1: 5’ CCG GTA CCT T AGC GGG GGC AAG AAG AAG 3’

XbaI-THG1: 5’ GGT CTA GAT TCA GTG ATG GTG ATG ATG 3’

8. *Experimental conditions*

a) Construction of Flag3-THG-1-His10 protein

The THG-1 gene was PCR amplified from previously cloned pCAGIP-FL-THG-1-His10 plasmid provided by Dr. Suzuki Hiroyuki using KpnI-THG1 and XbaI-THG1 primers. The resulting dsDNA was cloned in pGEM-T Easy vector (Promega). Resulted vector was transformed to E. coli strain DH5α. pGEM-T Easy- THG-1 plasmid and pT7-FLAG 3 plasmid was digested with KpnI and SacI. The 3’ end of cut pT7-FLAG 3 plasmid was dephosphorylated by shrimp alkaline phosphatase (Takara). The products were ligated and transformed into *E.coli* strain DH5α. The resultant pT7-FLAG 3-THG1-His10 expressing vector was transformed to *E.coli* strain BL21-Codonplus(DE3)-RP. The transformed bacteria clone was cultured in 2ml of LB-Amp medium for overnight at 37°C. The bacterial culture was mixed with 2ml of 40% glycerin and aliquot into 200μl stock for storage at –80°C until use.

b) Plasmid purification

50 mL of LB medium containing 100 μg/mL Ampicillin were inoculated with 1 colony of *E.coli* DH5α containing plasmid pT7-FLAG-3-THG-1-His10 and incubated overnight at 37°C while shaking at 220 rpm. The cells were spun down at 5 000 rpm for 10 minutes and the supernatant was discarded. Cells were lysed with 1% SDS in 0.2M NaOH. The plasmid purification was performed using a Wizard midiprep DNA purification (Promega) following its protocols. The plasmid was eluted with 300 μL of TE buffer and store at -20°C. The DNA concentration was measured using a spectrophotometer.

c) Transformation

1 μL plasmid DNA was added to 100 μL of competent BL21 (DE3), BL21-Condon Plus (DE3)-RP, BL21-Condon Plus (DE3)-RIL . Cells were kept on ice for 30 minutes, then heated at 42°C for 40s and back on ice for 2 min. 100 μL of transformation mix was transferred to LB agar plate containing ampicillin. The LB agar plate was kept at 37°C overnight.

d) Protein expression

One colony was picked from the agar plate and grown in 150 mL of LB medium containing 100 μg/mL ampicillin with shaking at 220 rpm overnight. In the following day, 10 mL of precultured medium was added to 1 L of LB medium containing 100 μg/mL ampicillin and incubated at 20°C with shaking at 220 rpm. At an OD600 of 0.5, protein expression was induced with 1 mM IPTG and the cells was kept at 20°C with shaking at 220 rpm for 4 hours. Cell culture was harvested by centrifuge at 5000 rpm for 10 min at 4°C. The supernatant was discarded and pellet was stored at -80°C until use.

e) Purification

i. Cell lysis

The cell pellet was resuspended in 300 mL 1x PBS with 1% Triton and 0.5 mM PMSF as a protease inhibitor. Cells were lysed by sonication with 10 short bursts of 10 sec followed by intervals of 30 sec for cooling. Follow by centrifugation at 8 000 rpm at 4°C for 15 minutes, the supernatant was kept and the pellet was discarded.

ii. Pull-down

Nickel beads were washed once with PBS. 50 mM of imidazole was added to the cell lysis. The cell lysis was incubated with nickel beads at 4°C with rotation at 150 rpm for 2 hours. Beads were collected and washed 3 times with PBS.

iii. Elution

Protein-bound beads were incubated with elution buffer (300 mM imidazole in 20 mM Tris-HCl pH=7.5, 150 mM NaCl) for 1 hour. The supernatant was collected and elution was repeated 3 more times.

iv. Buffer exchange and concentration

Collected supernatant was dialyzed in 1 L of storage buffer (20 mM Tris-HCl pH=7.5, 150 mM NaCl) overnight at 4°C. The protein sample was concentrated using an Amicon Ultra-15 centrifuge filter unit (Merck Millipore). After every 5 minutes of centrifugation, the sample was mixed by pipetting to prevent protein aggregation and absorption onto the filter membrane.

f) Coomassie Brilliant blue (CBB) staining and Western-blot

Samples were collected and subjected to SDS-PAGE. For CBB staining, the gel was incubated with staining solution (0.6 mM Coomassie brilliant blue in 50% methanol, 10% acetic acid) for 1 hour. The gel was then de-stained with a solution containing 25% methanol, 75% acetic acid). For western-blotting, proteins were electrotransferred onto PVDF membrane (Millipore, Burling ton, MA, USA). Anti-His (WAKO), FLAG (M2; Sigma), HA (3F10, Roche), THG-1 (Sigma) were used as primary antibodies for immunoblotting. Reacted antibodies were detected using an enhanced chemiluminescence detection system (WAKO).

g) Co-immunoprecipitation

The plasmids encoding HA-Keap1, HA-PHD2, HA-TBLR1, HA-NRBP1 were transfected into 293T cells ($2x10^5$ cells/ well 12 wells dish) using PEI reagent. Twenty-four hours after transfection, the cells were washed once with PBS and lysed by NP40 lysis buffer (1%NP-40, 10 mM Tris-HCl (pH 7.4), 150 mM NaCl). Cell debris was removed and the lysate was collected and incubated

with purified FL-THG-1-His10 protein. Anti-FLAG -antibody was added into the solution for 1 hour at 4°C and the immune complexes were precipitated by incubation with protein G-Sepharose beads for 30 minutes at 4°C. Collected beads then were washed with NP40 lysis buffer for 3 times. The immunoprecipitated proteins and aliquots of the total cell lysates were subjected to western blotting.

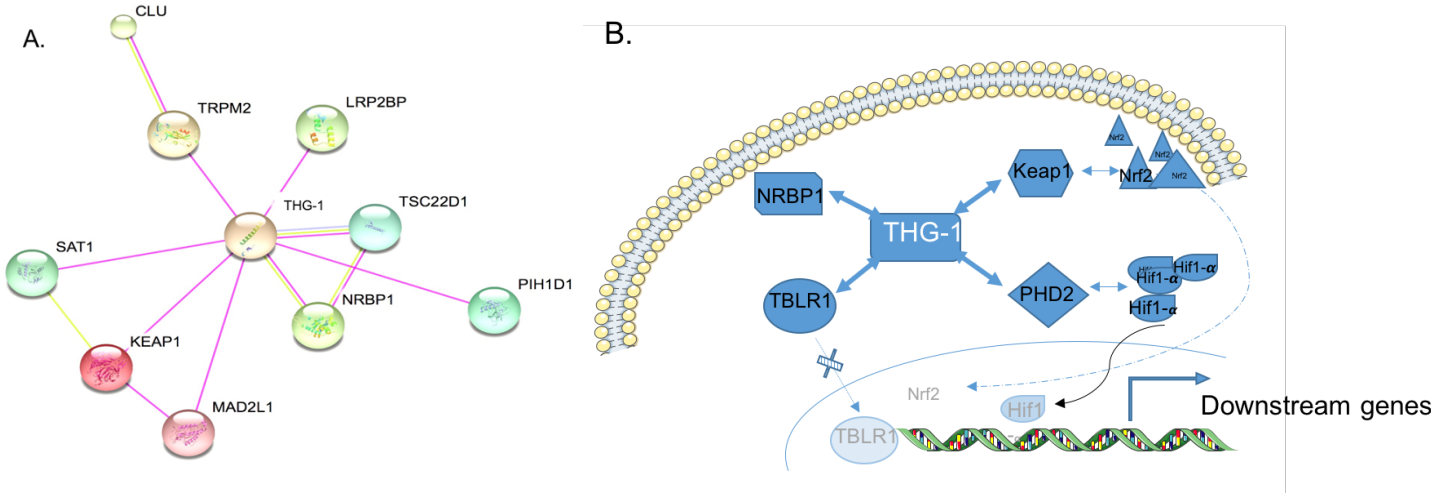
CHAPTER III: IDENTIFICATION OF MACROCYCLIC PEPTIDE LIGANDS

1. *Introduction*

a. Inhibition of THG-1 protein-protein interaction

THG-1 was found to be upregulated in various SCC tumors and found to be important for cancer cell proliferation, tumorigenesis, and invasion. THG-1 does not possess a DNA-binding domain and is mainly found in the cytoplasm. The mechanism by which THG-1 promotes its tumorigenesis activity is unclear. In the STRING database for protein-protein interaction reported several interacting proteins of THG-1 including Keap1, NRBP1, TSC22D1 [1] (Figure 3.1 A). By analyzing THG-1 co-precipitation proteins using mass spectrophotometry, several interacting proteins of THG-1 including Keap1, NRBP1, PHD2, TBLR1 were determined and investigated (Figure 3.1 B)

Figure 3.1: THG-1 interacting partners. A. THG-1 interacting network from STRING database. B. THG-1 interactions that are research focused in the lab.



As discussed previously in general Introduction (Chap I. 3c), THG-1 interaction with Keap1 is critical for the accumulation of Nrf2 in cancer cells, which leads to survival advantages of cancer cells. To further demonstrate the role of this interaction in cancer cells development and validate the possibility of target THG-1 for squamous cell carcinomas therapy, my research aims to develop THG-1 specific ligands that can serve as THG-1-Keap1 interaction antagonist.

b. Non-standard peptide screening against THG-1 by RaPID system

With the aforementioned advantages of macrocyclic peptides as drug entities for targeting protein-protein interaction, I here utilized the RaPID system to construct and screen thioether macrocyclic peptides against THG-1 protein.

2. Setting up of selection

a. Design of thioether macrocyclic peptides

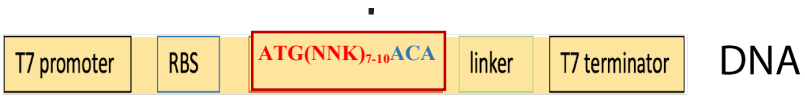
In my research, $NNK_{7\text{-}10}$ libraries (N: A/T/C/G; K: G/T) (Fig. 3.2A) was employed for the selection process. Those libraries encoded random peptides with 7- 10 amino acids in between the initiation amino acid and cysteine at the end. The initiation methionine codon (CAU) was reprogramed using ClAc-$_D$-Tryptophan. ClAc- $_D$-Tryptophan was pre-charged onto $tRNA^{fMet}_{CAU}$ by flexizyme eFx. During translation, methionine is removed from the reconstituted cell-free translation system and the CAU codon will become vacant. Charge ClAc-$_D$-W-$tRNA^{fMet}_{CAU}$ was added into the translation mixture, ClAc-$_D$-W was reassigned into CAU initiation codon (Figure 3.2 B).

The formation thioether bond between ClAc- of ClAc-$_D$-W-$tRNA^{fMet}_{CAU}$ group and SH-group of cysteine occurs spontaneously upon the completion of peptide translation needless the addition

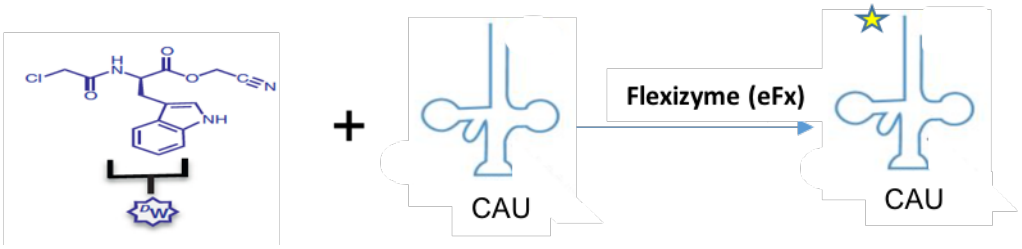
of any additional reagents (Figure 3.2 C). Therefore, macrocyclic peptide libraries with high diversity are prepared easily by preparing mRNA libraries from synthesized DNA.

Figure 3.2: **The preparation of thioether macrocyclic peptide library**. A. The schematic on the design of DNA library encoded macrocyclic peptides. B. Amino acylation of $tRNA^{fMet}_{CAU}$ using an activated ClAc-$_D$-Tryptophan and flexizyme eFx. C. General scheme depicting the cyclization process.

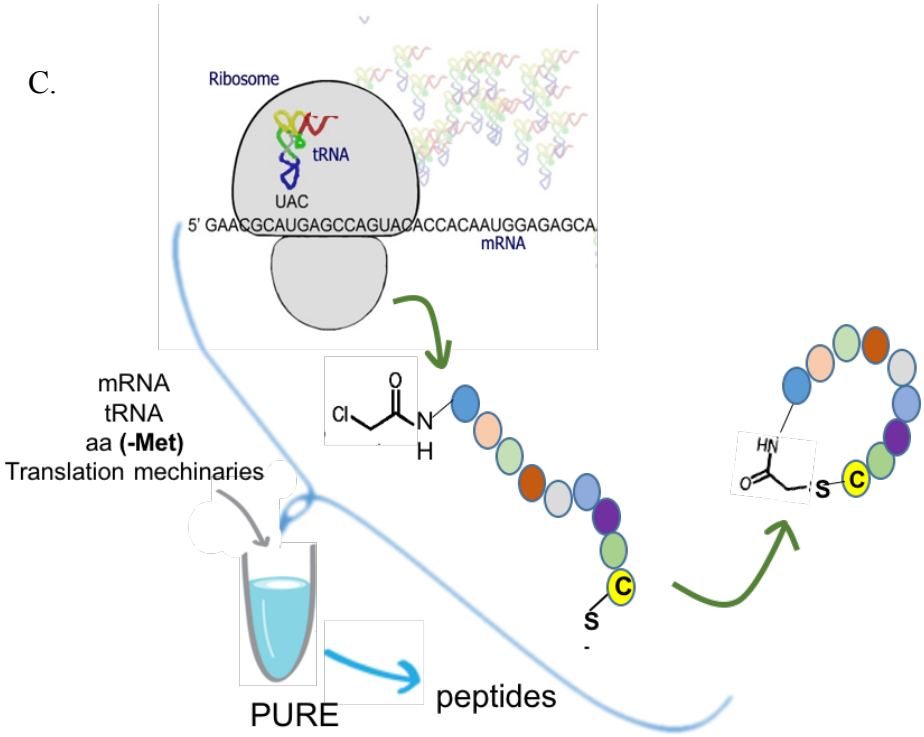
A.



B.



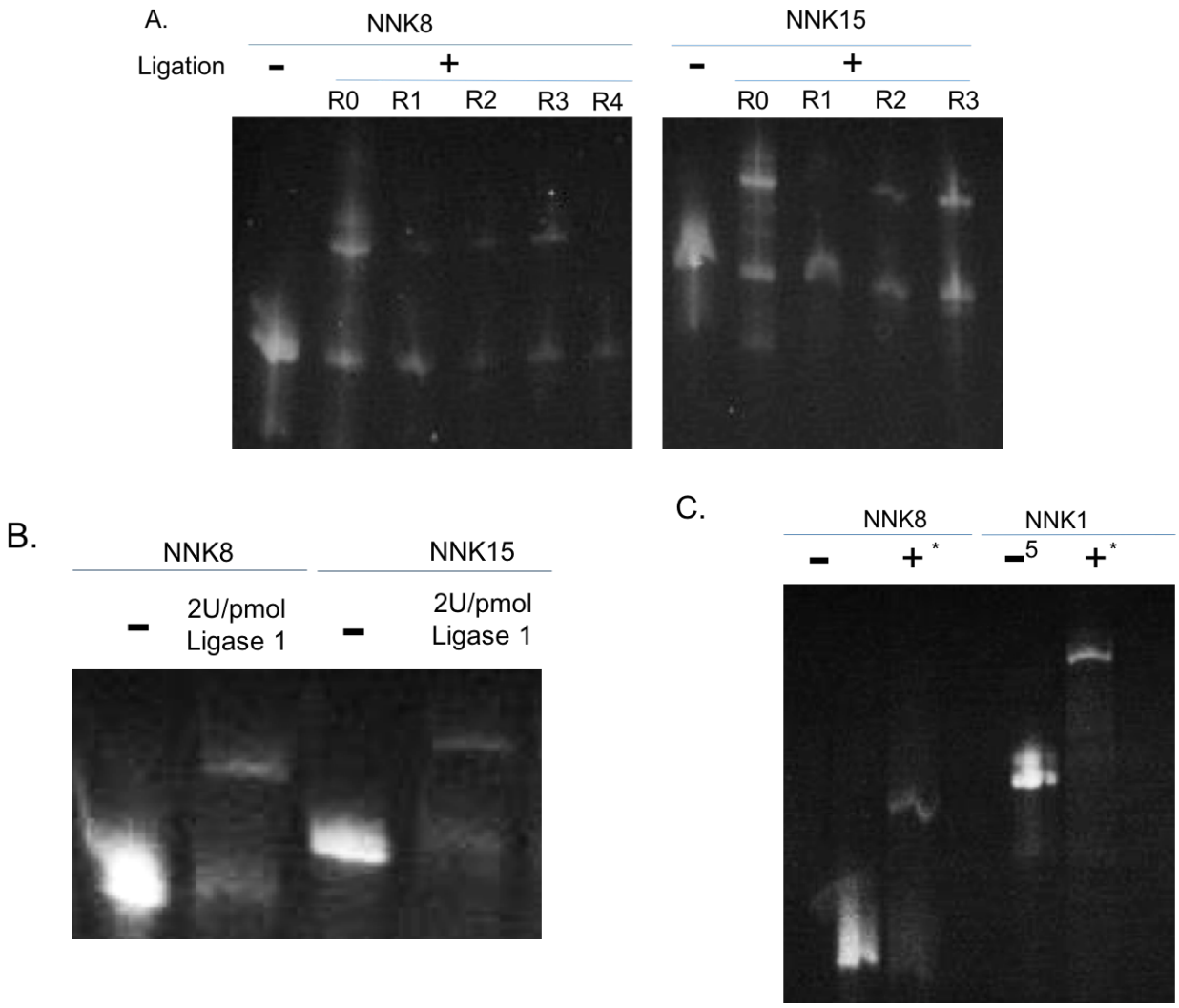
C.



b. Ligation efficiency improvement

In the RaPID system, one critical step for successful screening is the ligation of mRNA libraries with Puromycin linker (Pu-linker), which is responsible for conjugating mRNA (genotype) with its translated peptide (phenotype). Our system employed commercial T4 RNA ligase (NEB) with the stock concentration of 10 units/μL. According to established protocol, 0.6 units of T4 RNA ligase was used for ligation of 1 pmol of mRNA with 0.2 pmol of Pu-linker. With the same conditions for ligation, I recognized the decreasing in ligation efficiency over the selection process with efficiency around 0-30% (Figure 3.3 A). The reason is not well understood; it may result from the carryover phenol during mRNA extraction after transcription reaction. It may also depend on sequences of mRNA in libraries. I tested different conditions for ligation including temperature, hybridization facilitating, molecular crowding (PEG), splint DNA ligation (data not shown). I also tested the higher concentration of T4 RNA ligase 2 U/ pmol of mRNA and high T4 RNA ligase together with polynucleotide kinase (Figure 3.3 B &C).

Figure 3.3: Efficiency of the ligation between mRNA library and Pu-linker. A. Ligation efficiency over different rounds of screening process. R0: initial library ligation, R1-4: different rounds during screening. B. Ligation efficiency of NNK8 and NNK15 using 1.6 U/pmol of T4 RNA ligase. C. Ligation efficiency of NNK8 and NNK 15 using 2 U/ pmol of T4 RNA ligase and 1U/pmol polynucleotide kinase.

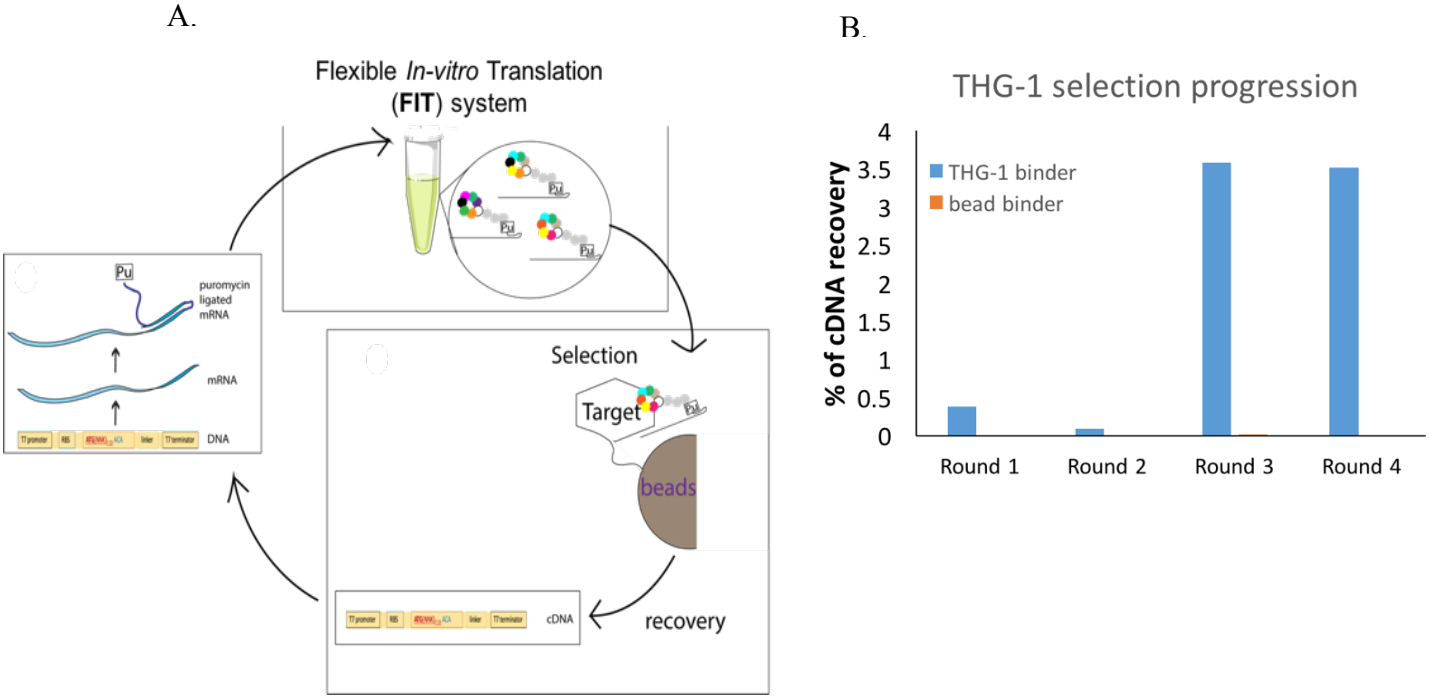


3. *Results of in-vitro selection of macrocyclic peptides*

In the first round of screening, mRNA pools of $NNK_{7\text{-}10}$ were ligated with Pu-linker. Ligated libraries were translated by the FIT system supplemented with pre-charged ClAc-$_{D}$-W-$tRNA^{fMet}_{CAU}$. Peptides-mRNA conjugated libraries were reverse transcribed and mixed with immobilized THG-1 on His-tag magnetic Dynabeads. cDNA of bound peptides were recovered and used as templates for the next screening cycle (Figure 3.4 A). After the 3rd round of screening, a

significant increase was observed, and after 4th round, the cDNA was sequenced using MiSeq (Illumina) (Figure 3.4 B).

Figure 3.4: **Selection of macrocyclic peptides against a human THG-1 recombinant protein.** A. Scheme of mRNA display selection of macrocyclic peptide ligands produced by the FIT system against THG-1 protein. B. Progress during the in vitro selection of macrocyclic peptides from NNK7-10 random library (~10^{12} in library size, where **N** is A, T, G or C, and **K** is G or T) against THG-1 recombinant protein. Recovery of cDNA from each round of selection was determined by quantitative PCR.



Revolution trajectory was established based on the sequences of peptides and their frequencies of occurrence over the selection process to show the evolution of peptides over the screening process (Figure 3.5 A). Revolution trajectory clearly showed the accumulation of some

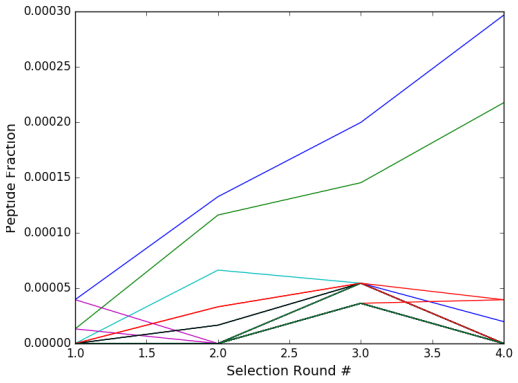
peptides together with the recession of the majority of sequences. Top sequences is the most accumulated peptide over 4 rounds of selection (Table 3.1) showed no common motif among them (Figure 3.5 B).

Peptide name	Amino acid sequences	Frequency	Percentage
R4-1	wLPTKLIRHCGSGSGS	24	0.05%
R4-2	wDKLSSPNPTTCGSGSGS	20	0.04%
R4-4	wKVAHVTPKCGSGSGS	15	0.02%
R4-5	wTRYHHVMICGSGSGS	12	0.02%
R4-7	wRYNQVDQCGSGSGS	11	0.02%
R4-8	wEKWSSPNSYCGSGSGS	11	0.02%
R4-9	wSRYHQDVCGSGSGS	9	0.02%
R4-10	wWRKKGNPTTLCGSGSGS	7	0.01%

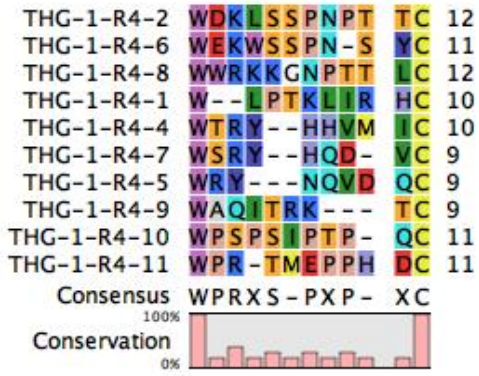
Table 3.1: THG-1 binding macrocyclic peptides in an enriched ClAc-$_D$-W library. Sequences of the top selected macrocyclic peptide after 4 rounds of selection against THG-1, which are determined by next the generation sequencing.

Figure 3.5: A. Selection evolution trajectory. B. Alignment of enriched peptides sequences. C. Target specificity of the cloned macrocyclic peptide. Binding efficiency between cloned macrocyclic peptide and proteins was analyzed by quantification of the cDNA recovered after incubation with immobilized protein using quantitative PCR.

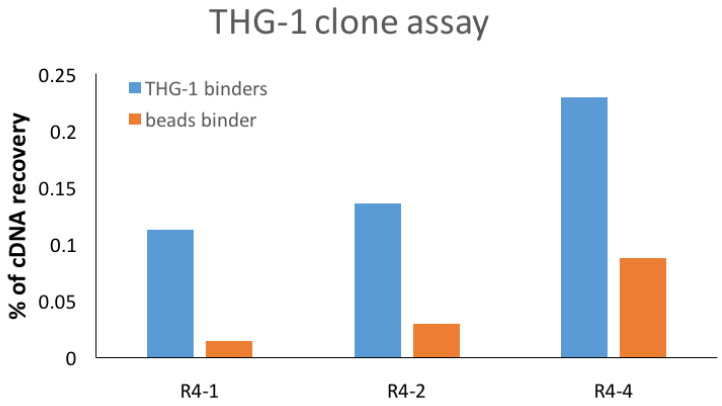
A.



B.



C.



The top 3 peptides were selected for clone assay against THG-1. Clone assay was carried out similar to the process used for a single round of selection with the initial library, but only contained a specific mRNA encoded a particular peptide. Peptide-mRNA solution of each peptide was separately subjected to THG-1 immobilized beads or free-beads. The clone assays indicated the binding of R4-1, R4-1 and R4-4 peptides against THG-1 as the recovered cDNA in THG-1 specific binding was higher than non-specific binding (beads binders) (Figure 3.4 C.)

4. Conclusion

The RaPID system was successfully employed to select for macrocyclic peptides that bind against THG-1. Even though, the RaPID system was established since 2007 with several published articles about ligands and active compounds against the targeted protein, all of them were selected in the Suga H. lab or with a homemade reconstitute translation system, T4 RNA ligase. Macrocyclic peptide screening established in this research was carried out using commercialized in vitro translation system (PURExpress) and T4 RNA ligase (NEB).

Screening showed significant enrichment at 3rd round. Sequences of several enriched peptides were identified using the next generation sequencing. Among them, R4-1, R4-2, and R4-4 were confirmed to bind to THG-1 in a clone assay. Unfortunately, sequences alignment of top enriched peptides did not reveal any common motif or characteristic for rational design for peptide improvement (if needed).

5. *Material*

Name	Sequences 5'-3'
T7-CH-F46	TAATACGACTCACTATAGGGTTGAACTTTAAGTAGGAGATATATCC
Lib-XNNK7-15C	GACCCAGACCCAGACCCACA(KNN)$_{7-15}$CATGGATATATCTCCTACTTAAAG
CGS3-CH-R39	TTTCCGCCCCCCGTCCTA**AGACCCAGACCCAGACCCACA**
CGS3-CH-R24	
Fx-F	GGCG*TAATACGACTCACTATA*GGATCGAAAGATTTCCGC
eFxR45	ACCTAACGCTAATCCCCTTTCGGGGCCGCGGAAATCTTTCGATCC
eFxR18	ACCTAACGCTAATCCCCT
fMetE-F	*GTAATACGACTCACTATAG*GCGGGGTGGAGCAGCCTGGTAGCTCGTCG
fMetE-R1	GAACCGACGATCTTCGGGTTATGAGCCCGACGAGCTACCAGGCTGCTC
fMetE-R2	TGGTTGCGGGGGCCGGATTTGAACCGACGATCTTCGGGT

6. *Experimental conditions*

- Preparation of eFx

PCR was carried out using Taq DNA polymerase (NEB) with 100 μM Fx-F, 100 μM eFx-R18, and 1 μM efx-R45. The resulting products were purified by phenol-chloroform extraction and ethanol precipitation. The DNA product was transcribed by HiScribe T7 Quick High Yield RNA synthesis kit (NEB) according to the manufactures instructions. Resulted mRNA was purified by denaturing PAGE and diluted in RNase-free water to a 250μM concentration.

- Preparation of $tRNA^{fMet}_{CAU}$

PCR was carried out using Taq DNA polymerase (NEB) with 100 μM fMetE-F, 100μM fMetE-R2 and 1 μM fMetE-R1. The resulting products were purified by phenol-chloroform extraction and ethanol precipitation. The DNA product was transcribed by HiScribe T7 Quick High Yield RNA synthesis kit (NEB). Resulted mRNA was purified by denaturing PAGE and diluted in RNase-free water to a 250 μM concentration.

- Preparation of $ClAc^{D}W\text{-}tRNA^{fMet}_{CAU}$

The following reaction was performed to make 125 pmol of $ClAc^{D}W\text{-}tRNA^{fMet}_{CAU}$. According to the required aminoacyl-tRNA, this reaction was scaled up. 2.1 μl of MilliQ water was mixed with 0.7 μl each of 500 mM HEPES-KOH buffer pH 7.5, 250 μM eFx and 250μM $tRNA^{fMet}_{CAU}$. The solution was heated at 98°C for 2 min and cooled down at RT for 5 minutes. 1.4 μl of 3 M $MgCl_2$ was added and the mixture was further incubated at RT for 5 minutes prior to transferring to on ice. Then 1.4 μl of 25mM $ClAc^{D}W$ in DMSO was added and the reaction was incubated on ice for two hours. The reaction was stopped and $ClAc^{D}W\text{-}tRNA^{fMet}_{CAU}$ was precipitated by adding 3.3 μl 0.3 M NaOAc, 8.3 μl 99.8% EtOH, followed by centrifugation at 15000rpm for 15 minutes. The

pellet was rinsed twice with 50μl 0.1M NaOAc in 70%EtOH and once with 50 μl 70%EtOH. Pellet ClAc^{D}W-tRNA$^{fMet}_{CAU}$ was stored at -80°C until use.

- Preparation of mRNA sequence library for expression of ^{D}W-NNK-library

The following procedure was used to prepare dsDNA NNK7-15 library. 100 pmol T7-CH-F46 was annealed with 100 pmol XNNK$_{(7\text{-}10)}$C-R(66-75) and extended by Taq DNA polymerase. The product was diluted 100 times with PCR reaction mixture and amplified using 500 pmol each of T7-CH-F46 and CGS3-CH-R39. The PCR product was purified using phenol-chloroform extraction and followed by ethanol precipitation. The dsDNA was transcribed using HiScibe T7 Quick High Yield RNA synthesis Kit (NEB) according to manufacture instruction. The resulting mRNA was gel purified by denaturing PAGE and dissolved in RNase free water at the concentration of 10 μM.

- Ligation of Puromycin linker to mRNA library

60 pmol of 10 μM mRNA library was incubated with 90 pmol of 7.5 μM puromycin linker, 1X ligase buffer, 10% DMSO, 1mM ATP, 2 U/pmol T4RNA ligase1 (NEB) and 1U/pmol T4 Polynucleotide Kinase (WAKO). The reaction was incubated for two hours at RT. 5μl of ligation product was subjected for visualization using denaturing PAGE and EtBr staining to ensure ligation efficiency. Ligated mRNA library was purified using phenol-chloroform-isopropanol and precipitated with ethanol. Products were dissolved in RNase free water.

- Immobilization of recombinant THG-1-His 10

Varying amounts of recombinant THG-1 protein (100 pmol, 200 pmol, 500 pmol) was incubated with 1 μl of anti-His magnetic Dynabeads (Invitrogene) for 15 minutes at 4°C. Then, supernatant and beads were separated. An equal amount of 2XSDS sample buffer was added, samples were heated at 98°C for 2 minutes and subjected to SDS-PAGE, followed by Coomassie

Blue Staining. The concentration that fully saturated the magnetic beads with a substantial amount of protein left in the supernatant (200 pmol) were selected for the selection process.

- Translational synthesis of ^{D}W-NNK-library cyclic peptide

PURExpress In Vitro protein Synthesis Kit (NEB) was used for translation. For every 7.5 pmol of puromycin-ligated mRNA, the following amount of kit components were used: 1 μl of SolA (-tRNA,a.a), 1.5 μl of SolB (-RF1,2,3), 0.5 μl of tRNA, 0.5 μl of 5 mM amino acid (-Met), pellet of 125 pmol of ClAc^{D}W-tRNA$^{fMet}_{CAU}$. The reaction mixture was incubated at 37°C for 30 minutes, followed by 12 minutes incubation at RT for enhancing conjugation between peptide and mRNA via puromycin. Then 1 μl of 100 mM EDTA was added to the translation mixture and the solution was further incubated at 37°C for 30 minutes to facilitate peptide cyclization.

- Selection of the random peptide libraries

The puromycin-ligated-mRNA was translated in 5 μl of translation reaction as described above. The mRNA was then reverse transcribed by incubation the mixture at 37°C for 60 minutes in addition with 0.4 μl of 250 mM $MgCl_2$, 0.2 μl of 2.5 mMdNTPs, 2 μl of 5X reaction buffer and 0.4 μl of Reverse Transcriptase, RNase H Minus, Point Mutant (M-MLV RT (H–) (Promega). Then, 1X PBS was added to the resulting mixture to make up the volume of 30 μl. After removing of EDTA and $MgCl_2$ by gel permeation chromatography (Sephadex® G-25 DNA GR), the solution was mix with an equal volume of 2X blocking buffer (0.2%BSA in 1x PBS). In order to remove the His-tag protein in the translation system and bead binding peptides, the random peptides were repeatedly mixed with 1 μl of anti-His Dynabeads (Invitrogen) and incubated for 10 minutes at 4°C for six times. Then, the THG-1 binding peptides were selected by mixing the remaining library with immobilized THG-1 for 30 minutes at 4°C. The supernatant was then removed and

the beads were washed for five times with 1X PBS. Then, 50 μl of stripping buffer (PCR reaction mixture containing CGS3-CH-R39 and T7-CH-F46 primers without Taq DNA polymerase) was added to the beads. The solution was heated at 98°C for 2 minutes and the supernatant was transferred into new tubes. 0.5 μl of the recovered solution was used for real-time PCR to quantify the amount of DNA recovered after the binding selection. The remaining DNA was amplified by PCR with 0.25 μl (1.25 U) of Taq DNA polymerase, that was added into the recovered solution. The amplified DNA was transcribed to produce mRNA library enriched with THG-1 binding sequences, which was used for the next selection round. After the fourth round of selection, recovery rate stayed plateau compared to third round. Selection was stoped and deep sequencing was performed.

- MALDI-TOF mass analysis

The analysis samples were desalted (if required) using C-tip (Nikkyo Technos) and analyzed using linear positive modes in an Axima TOF/TOF (Shimadzu)

- Next generation sequencing analysis of selection recovered library

First, PCR was performed with Rd1T7.F71 and an13Rd2.R49 to add primer hybridization sequence. Then, products were used for second PCR with P5S5**Rd1.F57 and Rd2N7**P7.R52 to add index and immobilization sequences. In 2^{nd} PCR, different indexes were used for different samples for their identity. PCR products were purified using NucleoSpin Gel and PCR Clean-up (Takara) according to manufacture instruction. Samples DNA were quantified using Qbit dsDNA BR Assay kit (Thermo Fisher Scientific) and diluted to approximate 10nM. All samples were mixed together using with equal volume, the mixture then was quantified and diluted into 4 nM using a buffer of 0.1% Tween in10 mM Tris-HCl (pH 8.5). 5 μl of mix DNA samples were denatured with

an equal amount of 0.2 M NaOH. Sequencing was performed using Miseq Reagent Kit v3 (Illumina) with PhiX control v3 (Illumina).

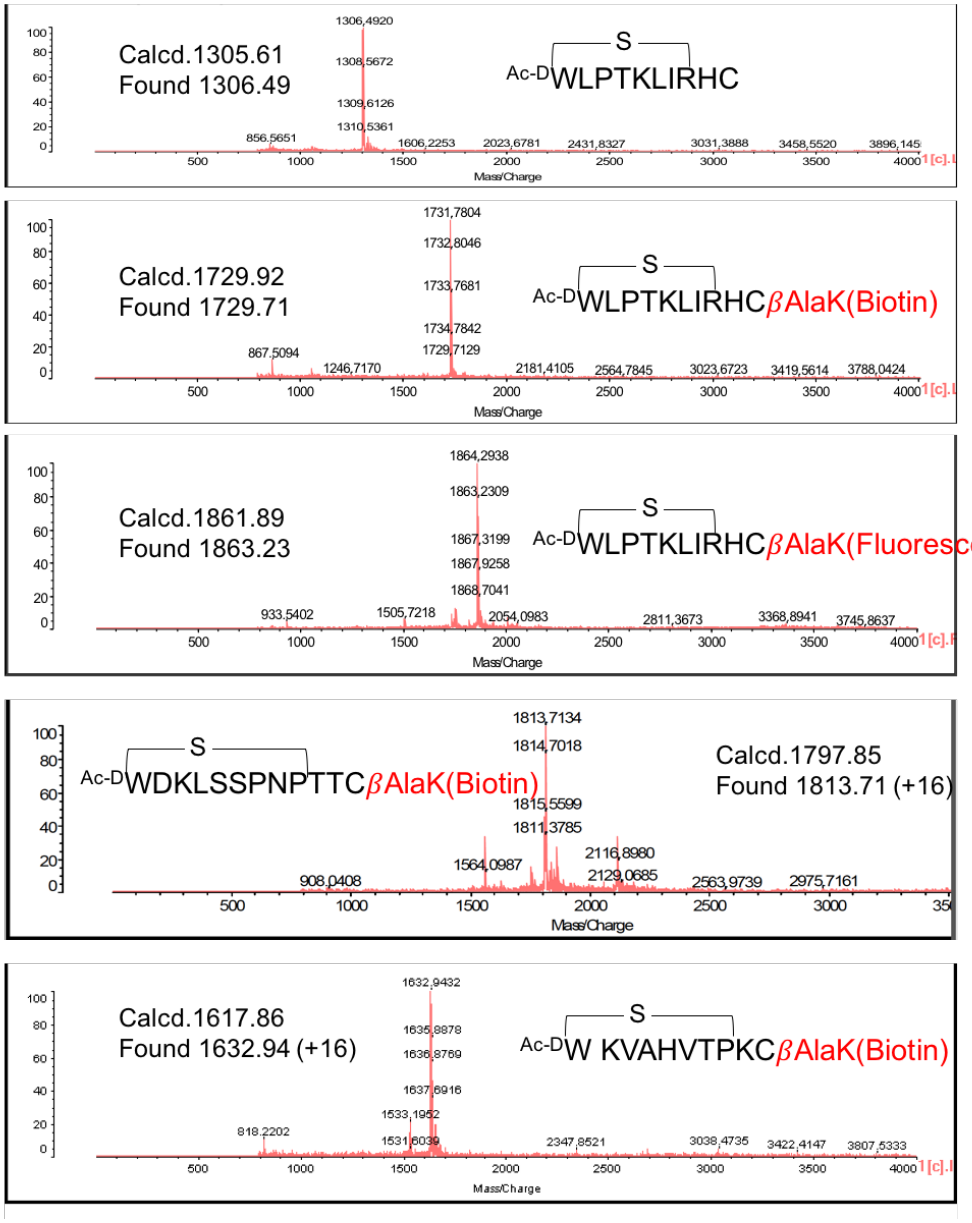
CHAPTER IV: CHARACTERIZATION OF THE MACROCYCLIC PEPTIDES

1. *The binding kinetic of the macrocyclic peptides*

In order to evaluate the binding affinity and biological activities of identified macrocyclic peptides to THG-1 protein, I synthesized those peptides and their derivatives (including biotin-tagged and fluorescein tagged) by Fmoc solid-phase peptide synthesis (Figure 4.1)

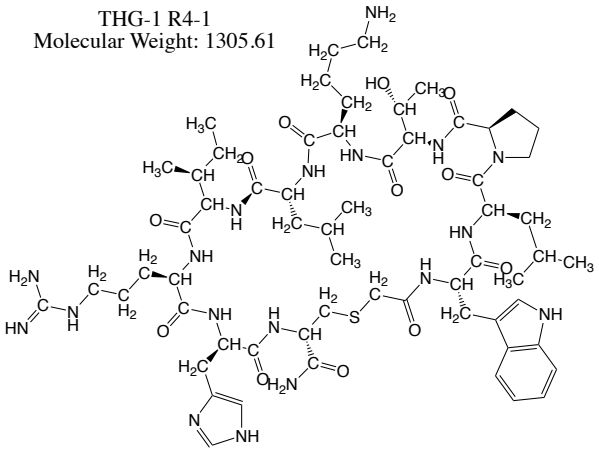
Figure 4.1: Chemical synthesis of identified macrocyclic peptides and their derivatives. A. Amino acid sequences and MALDI-TOF mass analysis of chemically synthesized peptides. B. Chemical structure of R4-1

A.



B.

THG-1 R4-1
Molecular Weight: 1305.61

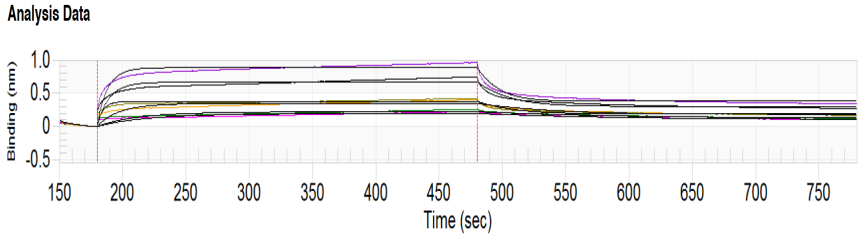


The synthesized peptides were analyzed for binding kinetic with THG-1 by Bio-Layer Interferometry (BLItz, fortéBIO). His-THG-1 protein was immobilized on biosensor tip using His-tag, biosensor then dipped in macrocyclic peptides solution. Binding of ligands to a protein causes a change in the interference pattern of light reflected between two surfaces: biosensor tip and the internal reference layer. Binding affinity analysis revealed that R4-1 peptide bound to THG-1 with a fast association ($1.99e^{6}$) and slow dissociation ($1.3e^{-1}$ s^{-1}), resulting in affinity of 65.6 nM (Figure 4.2).

Figure 4.2: **Binding kinetic analysis of R4-1 peptide**. A. Sensorgram of R4-1 with His-THG-1. B. The amino acid sequence and their calculated K_d values measured from BLItz.

A.

B.



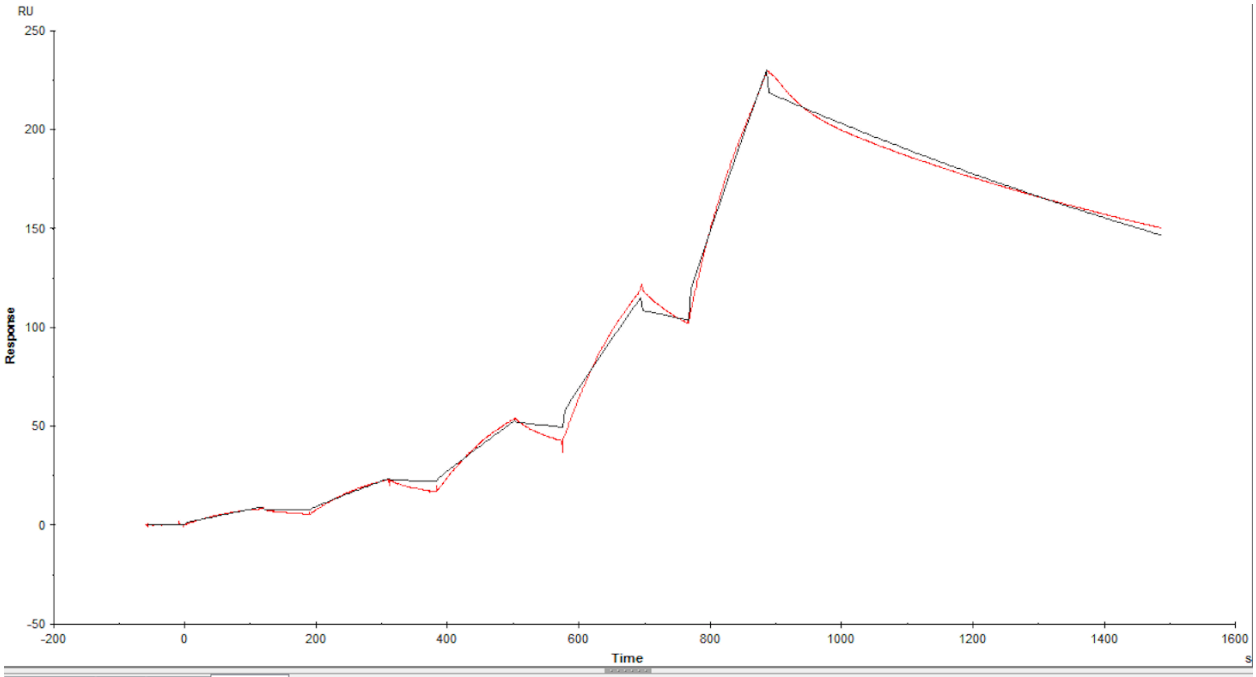
Name	Peptide sequences	Scaffold	K_d (nM)	K_a ($M^{-1}s^{-1}$)	K_d (s^{-1})
R4-1	wLPTKLIRHC	cyclic	65.6	$1.99e^{6}$	$1.3e^{-1}$

However, we observed high non-specific binding in the sensor control even though in different concentrations of BSA and Tween surfactant were experimented to reduce non-specific binding. High non-specific binding to the sensor chip causes difficulty in fitting measurement

data to a model equation for correct calculation of binding kinetic. Therefore, another method is needed for a more accurate measurement of binding kinetics.

Surface plasma resonance (SPR) was employed to measure the binding kinetics. In this experiment, His-tagged-THG-1 protein was immobilized on the sensor chip and different concentrations of R4-1B flowed through the sensor under 0.1% DMSO in PBS. Even though, the data showed that THG-1 binds to R4-1B, the binding kinetic cannot be determined correctly. The response unit (Rmax) was unexpectedly higher compare to it should be for a small size molecule like macrocyclic peptide (Figure 4.3).

Figure 4.3: Binding kinetic analysis of R4-1 peptide measured by SPR.



Quality Control | Report | Residuals | Parameters

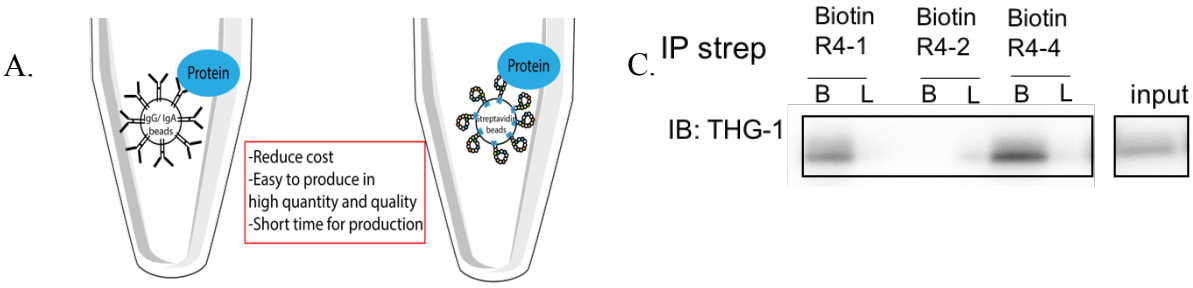
Curve	ka (1/Ms)	T(ka)	kd (1/s)	T(kd)	Rmax (RU)	T(Rmax)	Conc (M)	tc	T(tc)	f (ul/min)	RI (RU)	T(RI)
	260.9	17	6.692E-4	3.0E+2	4266.0	17		1.505E+8	0.84			
Cycle: 7							6.25E-08			30	1.2	13
							1.25E-07				0.8	8.3
							2.5E-07				1.3	12
							5E-07				7.4	62
							1E-06				13.1	1.1E+2

2. *Usage of peptide in pull-down, ELISA-like experiment, fluorescein staining and immunohistochemistry staining*

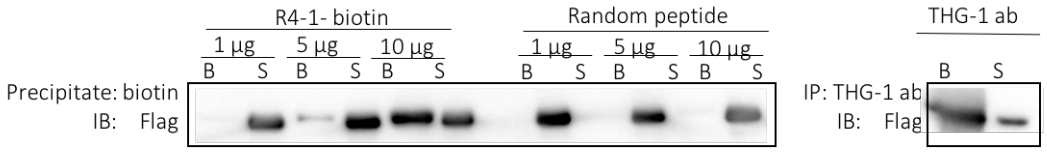
a. *Peptide dependent affinity pull-down*

THG-1 protein can be pull-downed in an R4-1-dependent manner when biotin-tagged R4-1 is immobilized on Streptavidine beads. The negative control showed no non-specific binding of THG-1 to Streptavidine beads coated with a control peptide (Figure 4.3). R4-1 peptide was dissolved in DMSO. In order to prevent the effect of DMSO on the conformation of THG-1 protein, biotin-tagged R4-1 was pre-incubated with Streptavidine beads in advanced. Resultant beads were then incubated with a buffered solution containing THG-1 protein. A limitation of this experimental system is the unidentified quantities of R4-1 macrocyclic peptide bound on defined Streptavidine beads.

Figure 4.3: Peptido-precipitation of THG-1 recombinant protein using R4-1 biotinylated peptide. A. Schematic of peptido-precipitation. Increasing amounts of R4-1 tagged with biotin was immobilized on Streptavidin beads. Then, R4-1 immobilized beads were incubated with THG-1 overexpressing HEK293T lysate. Beads were collected and washed with PBS. Both beads and supernatant were subjected to SDS-PAGE and WB analysis. B. Western-blot analysis of R4-1 peptide-dependent precipitation of purified *E.coli* recombinant His-THG-1. C. Western-blot analysis of R4-1 peptide-dependent precipitation of purified 293 overexpressed THG-1.



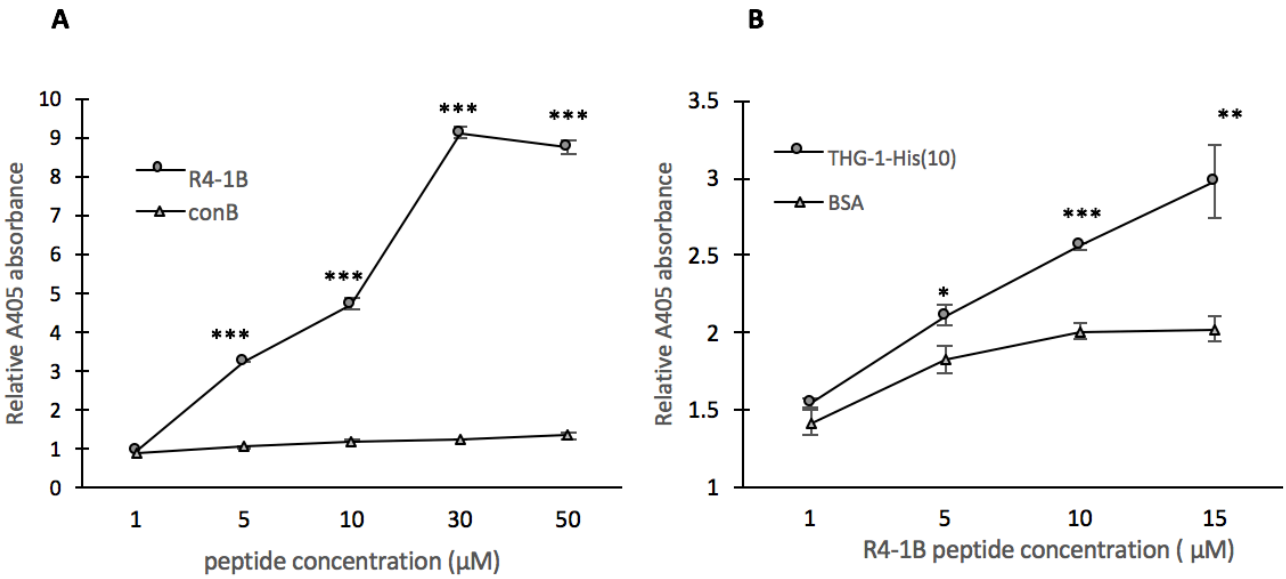
B.



b. ELISA-like protein quantification

R4-1B also was used in an experiment similar to enzyme-linked immunosorbent assay (ELISA). Varying concentration of R4-1B or Biotin-tagged control peptide solution were incubated for 2 hours in THG-1 pre-coated Nunc-immuno-plate (Thermo-Scientific). Avidin was used for linking biotin-horseradish peroxidase (HRP) and R4-1B. As the concentration of R4-1B increased, the signal was increased. This phenomenon was not observed in random peptide case, as well as in plate coated with BSA only (Figure 4.4)

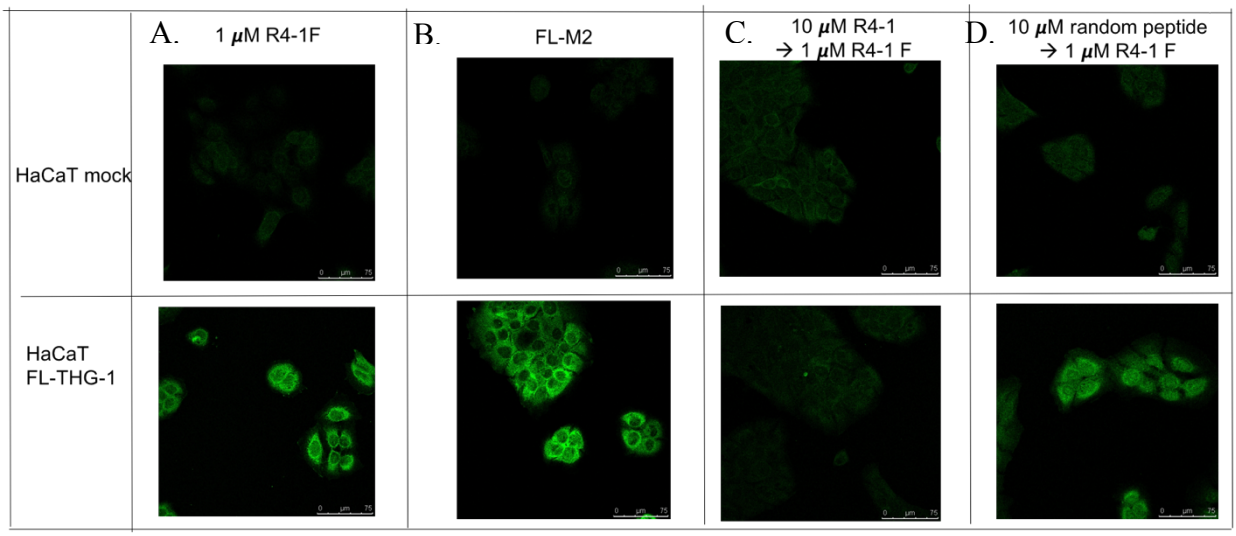
Figure 4.4: Usage of macrocyclic peptides for Enzyme-linked peptide-sorbent assay (ELPA). A. Relative signal read-out from 50 ng/ well His-THG-1 pre-coated Nunc-immuno-plate incubated with R4-1B peptide or control peptide. B. Relative signal read-out from 50 ng/ well His-THG-1 or control protein pre-coated Nunc-immuno-plate incubated with R4-1B peptide. C. Relative signal read-out from different amount of THG-1 protein pre-coated Nunc-immuno-plate incubated with 15 µM R4-1B peptides



c. *Fluorescein imaging*

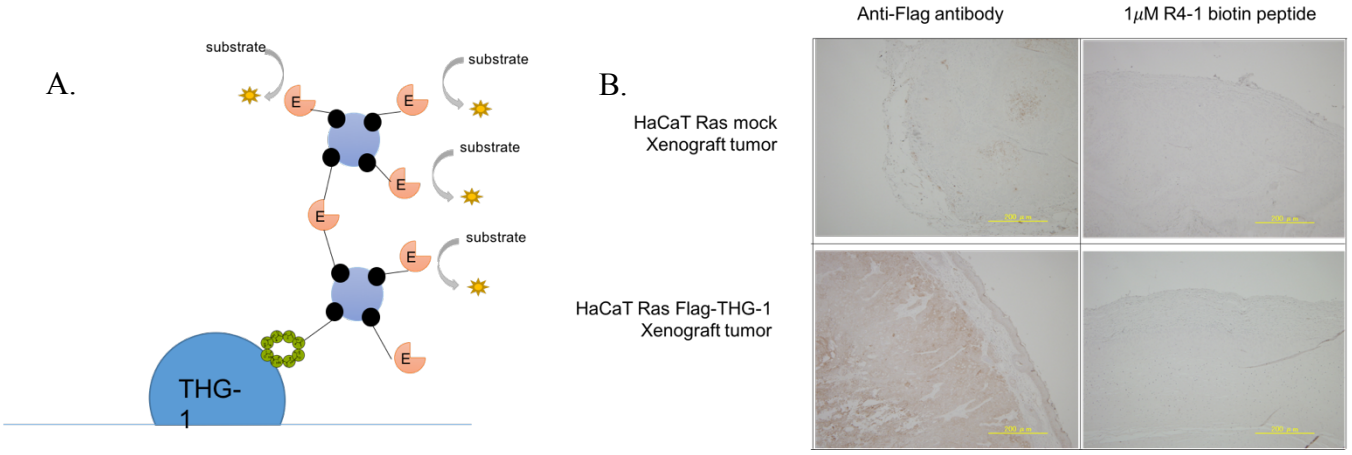
I prepared fluorescein tagged R4-1, refer as R4-1F, and used it to visualized the localization of THG-1 protein in HaCaT Ras- mock/ HaCaT Ras-THG-1 cells. R4-1F showed specifically staining to THG-1 as shown by the localization that was consistent with the staining pattern produced by an anti-Flag antibody (N-terminus tag of overexpressed THG-1). No staining was found on HaCaT Ras-mock cells, which does not express THG-1 (Figure 4.5 A, B). To further confirm the specificity of binding, a 10-folds concentration of non-tagged R4-1 or random macrocyclic peptide was incubated with cells before stained with R4-1F. The incubation with non-tagged R4-1 diminished the staining of R4-1F while this phenomenon was not observed in condition using control macrocyclic peptide (Figure 4.5 C, D).

Figure 4.5: **Laser confocal microscope image of HaCaT Ras-mock and HaCaT Ras-FL-THG1 cells stained with fluorescein-tagged R4-1 (R4-1F) and Flag-M2 antibody.** A. Cells stained with 1 μM R4-1F. B. Cells stained with FL-M2. C. Cells incubated with 10 μM R4-1, then stained with 1 μM R4-1F to access non-specific binding.



Iwasaki K et al. presented the staining of high-density MCF7 using macrocyclic peptide selected against EpCAM (Iwasaki K. et al, 2015). The report suggested an advantage in using macrocyclic peptide above to stain high-density cells, tumor tissue (Iwasaki K. et al, 2015). This could be explained when associating macrocyclic peptide size (~1-2 kDa) and antibody size (>40kDa) with their penetration ability. Without access to frozen tissue, I attempted to see if there is any difference between R4-1 stained and antibody stained on paraffin-embedded tumors. Biotin-tagged R4-1 (R4-1B) was used to stained embedded xenograft tumor from HaCaT-Ras mock and HaCaT-Ras THG1 in a manner similar to immunohistochemistry, where color development was depended on the reaction between Horse peroxidase (HPR) and its substrate. Avidin was used to linked R4-1B and HRP-biotin. Results are shown in Fig. 4.6. However, staining was not observed in using 1 μM of R4-1B. The usage of R4-1B for histochemistry staining may need different optimization process including peptide concentration, antigen retrieval method, incubation time, *etc*.

Figure 4.6: **Peptide-histochemistry staining of xenograft tumor from HaCaT-Ras-mock and HaCaT Ras-THG1. A.** Schematic of biotin-tagged- R4-1 dependent histochemistry staining. The flag-M2 antibody was used as a positive control. HaCaT-Ras-mock is negative control for non-specific binding of the peptides.



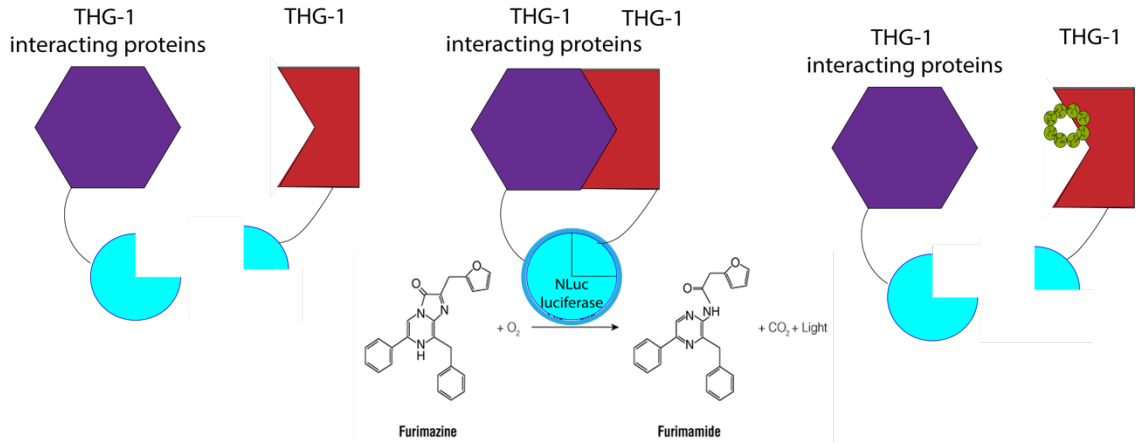
3. *Evaluation of selected macrocyclic peptides as THG-1 protein-protein interaction inhibitors*

As mentioned earlier, THG-1 is suggested to modulate its functions through protein-protein interactions (PPIs). Demonstrated in our lab, THG-1 interacts with Keap1 and possesses an important role in tumorigenesis (Suzuki H., et al, unpublished data). A THG-1-Keap1 interaction antagonist is required for further validation on the role of this interaction. To assess whether R4-1 has the ability to disrupt THG-1 PPIs, particularly THG-1 and Keap1, I employed the Nano-BiT system (Dixon AS., et al, 2015). In the NanoBiT system, complementary fragments of NanoLuc are genetically fused onto either Keap1 and THG-1. The intrinsic affinity of NanoLuc fragments is 190 μM, which is outside the ranges of PPI. The activity of NanoLuc depends on the complex formation caused by fused protein pairs: Keap1 and THG-1.

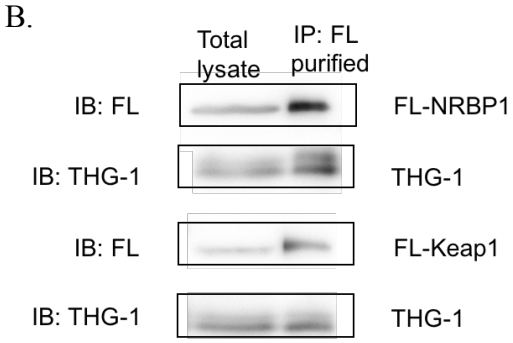
Figure 4.7: **NanoBiT system to evaluate the antagonist activity of the identified macrocyclic peptide**. A. Schematic of NanoBiT assay. B. Western-blot analysis of purified NanoBiT

complex. Anti-Flag antibody was used to affinity-precipitate FL-NRBP1/ FL-Keap1. THG-1 was co-precipitated

A.



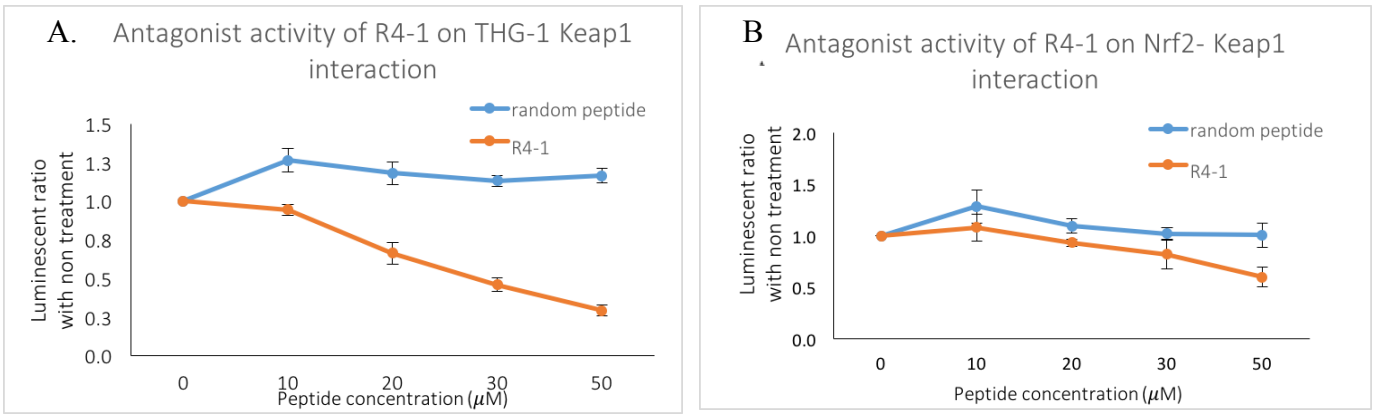
B.

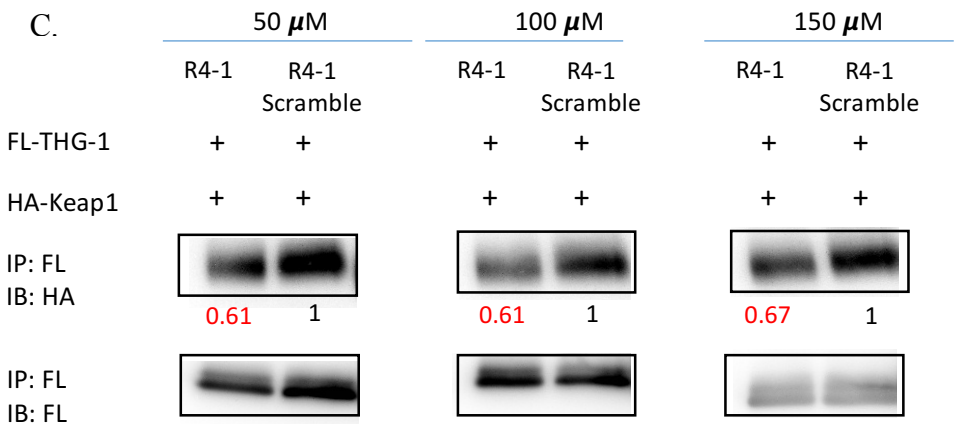


In current research and development, there are examples of compounds that disrupt Keap1 and Nrf2 interaction. However, this leads to the accumulation of Nrf2 and would only be applicable for anti-inflammatory diseases but not in squamous cell carcinomas. Results demonstrated that R4-1 can disrupt THG-1 specific interaction with Keap1 (Figure 4.8A) but does not affect Keap1 and Nrf2 interactions. Specific targeting of THG-1 would serve as a potential therapeutic with less side effect. According to our data, embryonic knock out of THG-1 is not

lethal and mice development was mostly normal with the thinner skin layer. Protein atlas indicated

Figure 4.8: **R4-1 disrupts THG-1 and Keap1 interaction in vitro.** A. R4-1 inhibits Keap1 and THG-1 interaction B. but not Keap1 and Nrf2 interaction. C. Western blot analysis of THG-1 and pulled down Keap1 in presence of R4-1.





4. *Methods*

a. Solid phase chemical synthesis

Fmoc-amino acid, peptide coupling reagents, Rink-Amide resin (Novabiochem) and solvents were reagent grade and used without further purification unless otherwise noted. Peptides were synthesized using Syro I instrument (Biotage).

Rink-Amide resin (54 mg, 25 mmol) in the synthesized reactor was pre-swollen in 1 mL of DMF (dimethylformamide) for 30 min. The process of deprotection and amino acid coupling was carried out automatically and repeated until N-terminal D-Tryptophan was coupled and their Fmoc was removed. Briefly, Fmoc deprotection was carried out by incubated resin with 40% piperidine in N, N-dimethylformamide (DMF) with agitation for 3 min, then repeated once for 12 minutes. After washing with DMF (1mL, 6 times), coupling was achieved when resin was incubated with appropriate amino acid and 0.19 M 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and 1-hydroxybenzentriazole (HOBt), and 0.38 M N-N-diisopropylethylamine (DIPEA) in DMF with vortexing for 40 min.

After washing the resin with DMF, peptides were cleaved from the resin and side-chain de-protected by a solution of trifluoroacetic acid (TFA): triisopropyl silane (TIS): water (925:25:50) with rotation at room temperature for 3 hours. Cleaved peptides were precipitated by adding diethyl ether, centrifuged, and washed with diethyl ether (3 mL, five times). Crude peptides were dissolved in 50% water: acetonitrile 0.1% TFA. Triethylamine was used to raise pH to approximately 10, and the solution was incubated at room temperature for 1 hour with shaking to facilitate the cyclization via thioether bond formation between N-terminal chloroacetamide group and cysteine sulfhydryl group. Then peptides solution was acidified by TFA to approximately pH=3 and purified by reverse-phase HPLC (Shimadzu).

b. <u>Fluorescein/ biotin conjugation to peptide</u>

Peptide chemical synthesis was carried out similarly as previously described until N-terminal Fmoc was removed. Monomethoxyltrityl (MMT) protecting group on lysine was removed by incubating the peptide-bound resin in 1% TFA, 1% TIS in dichloromethane (10 min, 9 times). The resin then was washed with dichloromethane, N-Methyl-2-pyrrolidone (NMP), and 10% DIPEA in NMP respectively. 0.1 M NHS-Fluorescein/ NHS-biotin in 10% DIPEA in NMP for 1 hour. After washing the resin with dichloromethane and DMF, peptide cleavage was carried out normally as described previously (ChapIV-5a).

c. MALDI-TOF analysis of synthesized peptides

MALDI-TOF was performed using reflectron positive modes in an Axima TOF/TOF (Shimadzu)

d. BLItz kinetic measurement

Biosensors were pre-hydrated in binding buffer (PBS 0.05% Tween 0.1% BSA) for 30 min before the experiment, followed by initial baseline washing for 30-sec in binding buffer. The sensor tip was then transferred to the His-THG-1 containing drop holder at a concentration of 0.15 μg/μL for a 120-sec loading step. After 30-sec baseline dip in binding buffer, the binding kinetics was measured by dipping THG-1 coated sensor tip in a peptide solution for 120-sec association step. The subsequent dissociation step was performed by dipping the sensor tip in fresh binding buffer for 120-sec.

e. ELISA like peptide sorbent experiment

THG-1 or control protein was coated on Nunc-immuno-plate (Thermo-scientific) by incubating 1 μg/ mL protein solutions at 4°C overnight. After washing the plates, they were blocked by incubating with 2% BSA in PBS for 1 hour at 37°C. Peptide solutions at various concentrations (0.3, 0.5 and 1 μM) were incubated in the plate for 2 hours. After washing by PBS

(200 μL, 3 times), plate was incubated with ABC complex (Vectastain, VECTOR) according to manufacture instruction. Color development was achieved by adding ABTS substrate (Sera Care), the reaction was stopped by adding an equal volume of 1% SDS. Absorbents were recorded at OD_{405} using Microplate reader MTP-300 (Corona Electric).

f. Inhibition of THG-1 and Keap1 co-precipitation by R4-1

FL-THG-1 and HA-Keap1 were separately transfected into 293T cells. After 24 hours, cells were collected and lysed with NP40 lysis buffer. FL-THG-1 was immunoprecipitated by anti-Flag antibody (M2, SIGMA) and IgG beads from cell lysate. After washing beads (PBS, 3 times), different concentration R4-1 or scramble peptide in PBS were added to beads and incubated for 30 min at 4°C. Cell lysate from Keap1 overexpressed 293T was added into the mixture and incubated for an additional 30 min. Beads were then collected and washed 3 times with PBS, then subjected to SDS-PAGE for Western-blot analysis

g. Peptido-precipitation

Chemically synthesized R4-1 biotin was immobilized on M-280 streptavidin beads (Dynabeads, Life Technologies). The R4-1 biotin-beads were incubated with purified Flag-THG-1 from *E.coli* or from 293T cells at 4°C for 0.5 h. Beads and supernatant were separated and beads were washed with cold 1x PBS three times. Samples were analyzed by 10% SDS-PAGE. THG-1 was visualized using anti-Flag (M2, SIGMA) mouse monoclonal antibody and anti-mouse HRP-conjugated secondary antibody followed by chemiluminescent detection.

h. Peptido- fluorescein staining and confocal microscopy

HaCaT or HaCaT-Flag-THG-1 cells were cultured on 25x35 mm cover glass on 6 well dish at 37 °C under 5% CO_2 overnight. Culture medium was removed and cells were washed once with

1x PBS. Cells were fixed and permeabilized by 200 μL/ well 100% cold acetone at -20 °C for 5 min. After washing 3 times with 1x PBS, 1 μM of R4-1 fluorescein-labeled macrocyclic peptide or mouse anti-Flag antibody (M2, SIGMA) in 100 μL of 1x PBS was added and incubated at room temperature for 1 hour. After washing 3 times with PBS, Alexa Fluor 488- conjugated mouse anti-mouse IgG antibody (1:200) in 1x PBS was added to the anti-Flag antibody incubated wells and incubated for an additional 1 hour. R4-1 fluorescein (R4-1F) incubated wells were kept in 1xPBS. All cells were washed 3 times with PBS before mounting and imaging using a spectral confocal microscope (Leica). For competitive assay with non-labeled peptides, cells were treated with 10 μM non-labeled peptides, washed 3 times with 1xPBS before the incubation with 1 μM R4-1 fluorescein-labeled peptides.

i. Immunoprecipitation of NanoBiT complex

LgBiT-FLAG-Keap1 and SBiT-THG-1, LgBiT-FLAG-Keap1 and SBiT-Nrf2 were stably expressed in HEK 293 T cells (Nano KT and Nano KN cells). NanoKT and Nano KN cells were cultured on three 10 cm dishes until reach 90% confluence. Cells were washed with 1xPBS and lysed with NP40 lysis buffer. The supernatant was collected by centrifuging cells at 130,000 rpm for 15 min at 4°C. 2 μL of mouse anti-Flag antibody (M2, SIGMA) was added into the supernatant and incubated for 1 hour at 4°C with agitation. 25 μL of mouse IgG agarose beads were added and incubated for an additional 30 minutes. The LgBiT-FLAG-Keap1 and SBiT-THG-1 complex was eluted by 100 μg/ml of FLAG peptide (WAKO.)

j. NanoBiT assay

Purified NanoBiT complexes were incubated with R4-1 peptide at various concentrations at 4°C for 1 hour. Protein-protein interaction was measured by NanoGlo Luciferases assay system

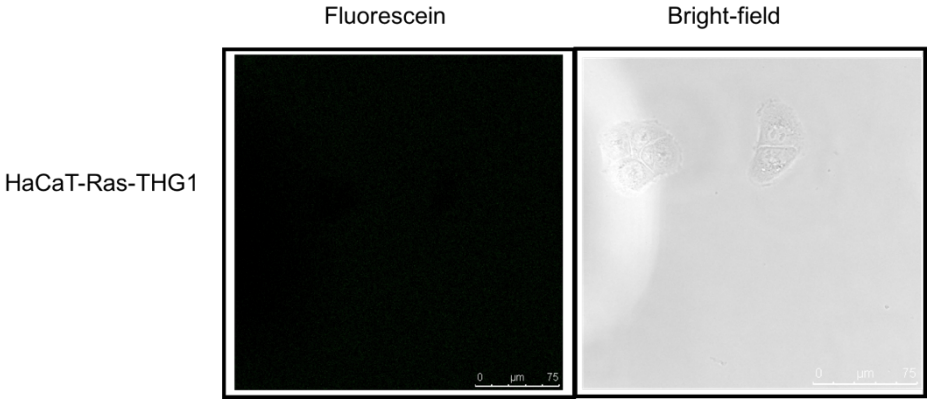
(Promega) according to instructed protocol. Briefly, NanoGlo substrate was diluted 100 times in 2x-diluted buffer. An equal volume of prepared substrate combined with incubated NanoBiT complexes was added to a white 96-well plate in triplicate. Luminescent emission was scanned using a MicroLumat Plus plate reader (Berthod).

CHAPTER V: IN CELL ACTIVITY OF MACROCYCLIC PEPTIDE

1. *Cellular permeability analysis of R4-1 peptide in HaCaT-Ras-THG-1 cells*

Peptide usually has a size bigger than 500 Da and highly hydrophilic, therefore it is generally poor in cell-permeability. Thus, R4-1 peptide may not exhibit its inhibitory activity to THG-1 and Keap1 interaction due to its poor cell-penetration. To confirm this is true or not, R4-1F was used to evaluate cell-penetration in HaCaT-Ras-THG1 cells. Cells were incubated with 1 μM R4-1F for 3 hours, then microscopic observation was performed. No fluorescein signal was detected indicating R4-1F is not cell-permeable (Figure 5.1).

Figure 5.1 : Microscopic observation of HaCaT-Ras-THG1 cells incubated with 1μM R4-1F for 3 hours

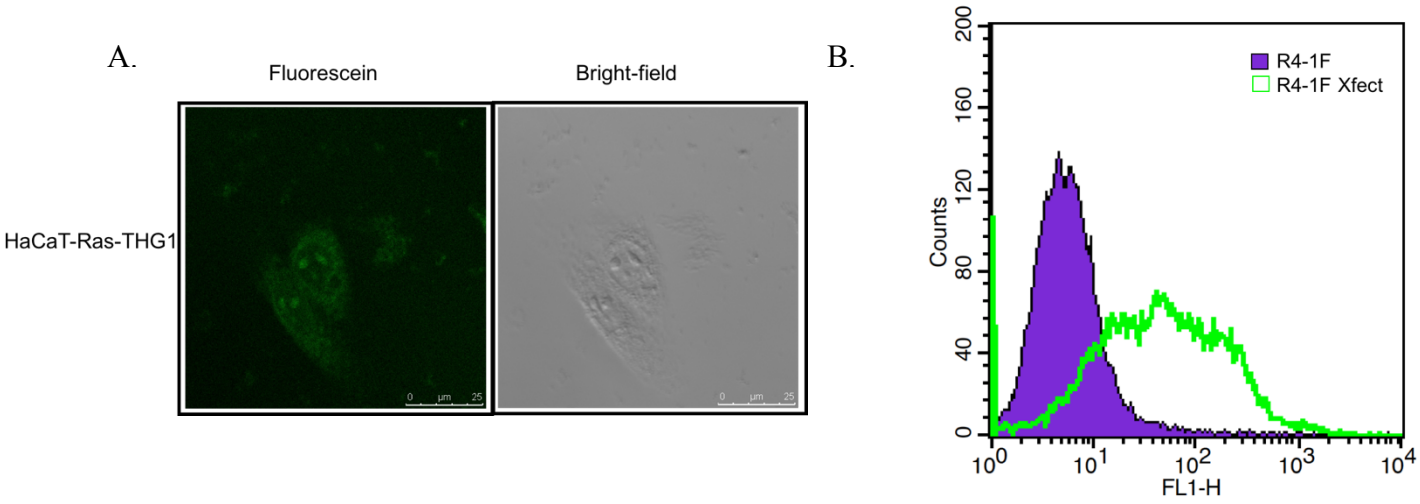


2. *Transfection of the macrocyclic peptide*

In order to deliver R4-1 into cells, I employed the Xfect protein transfection reagent (Takara). Xfect protein transfection reagent includes an amphipathic peptide that contains both hydrophobic and hydrophilic domains. Amphipathic peptide interacts with protein/ peptide of interest and cell membrane, thereby promoting the internalization of protein/peptide of interest into cells by caveolae-mediated endocytosis. Results indicated successful transfection of R4-1

into HaCaT-Ras-THG-1 cells. As shown in figure 5.2 A, fluorescein signal of R4-1F was detected inside cells with a dispersed pattern. According to Fluorescence-activated cell sorting (FACS) data, the transfection efficiency of approximately 90% can be achieved (Figure 5.2B).

Figure 5.2: **Transfection of R4-1F into cells using Xfect protein transfection reagent.** A. Microscopic observation of live HaCaT-Ras-THG1 cells incubated with R4-1F in Xfect transfection reagent. B. FACS data of transfected HaCaT-Ras-THG1 cells compared with cells incubated with R4-1F only.

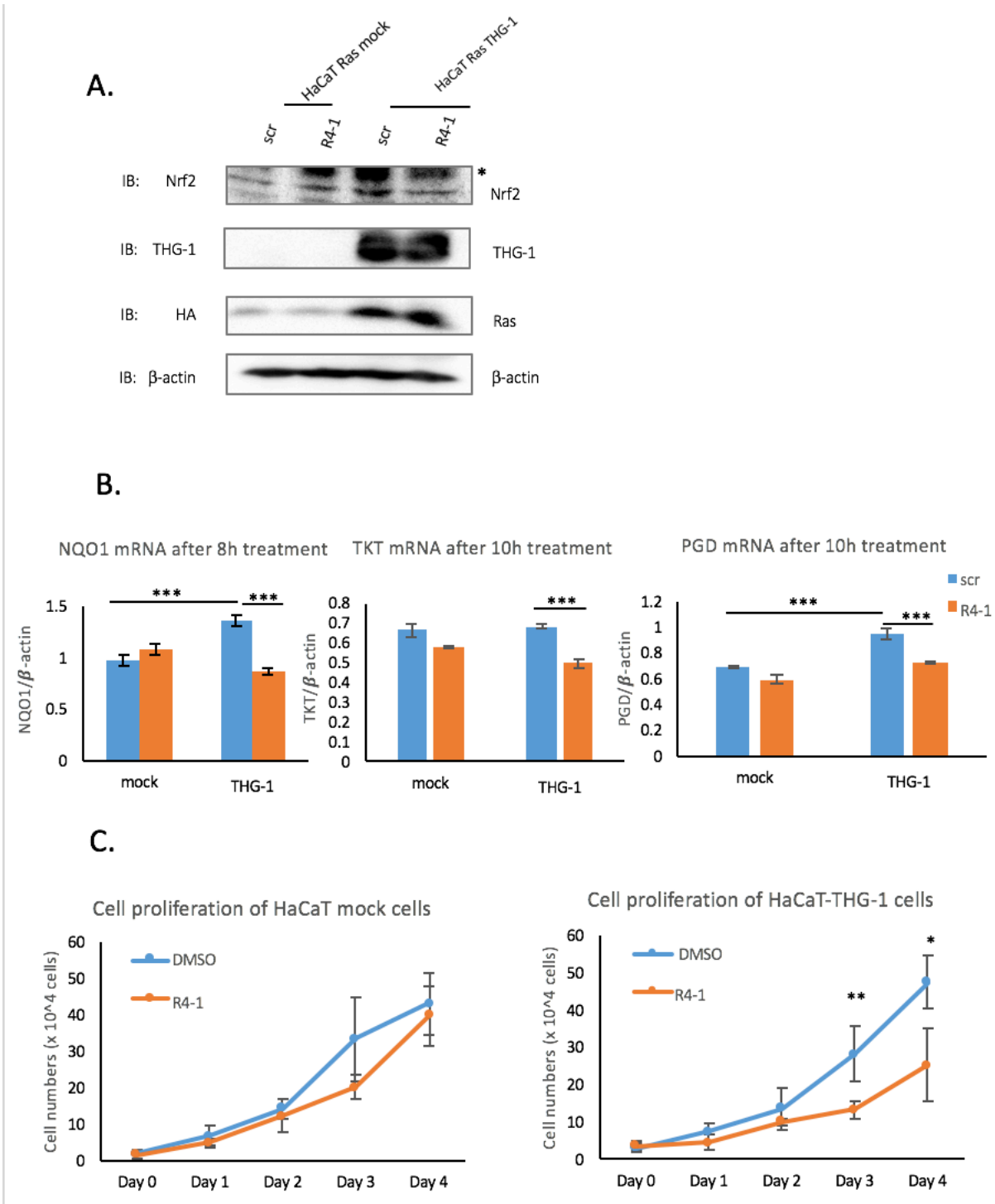


Results showed that Xfect can carry R4-1F inside cells and importantly, fluorescein signal showed a diffused pattern in the cytoplasm, not in the endocytoplasmic vesicles. This is important for R4-1 to be functional in cells, the localization is together with THG-1. With high transfection efficiency, R4-1 is expected to show clear effects on the expression level of Nrf2, a Keap1 target protein. However, it is important to note that the amount of actual transfected peptide inside cells compared to the total used amount cannot be determined. Therefore, the effective dosage for effect could be different from the usage doses. Transfection of R4-1 into cells can only be used to assess R4-1 effect on THG-1 and Keap1 interaction qualitatively.

3. Effect of R4-1 on Keap1 and THG-1 interaction when transfected into HaCaT-Ras THG1 cells

THG-1 sequesters Keap1 from its interaction with Nrf2, thereby increases Nrf2 protein level as well as Nrf2 target genes expression. To test the effect of R4-1 disrupting this interaction of THG-1 and Keap1 in cells, R4-1 was transfected inside cells and Nrf2 protein level, as well as its target genes mRNA, were evaluated.

Figure 5.3: Disruption of THG-1 and Keap1 interaction by R4-1 resulted in reducing of Nrf2 activity. A. Western-blot analysis of Nrf2 protein level in R4-1 transfected HaCaT-Ras-THG-1 and HaCaT-Ras- mock cells (*: non-specific band). B. mRNA expression level of NQO1, TKT, PGD, target genes of Nrf2. C. The proliferation of HaCaT-Ras-mock (left) and HaCaT-Ras-THG-1 (right) when treated with DMSO or R4-1.



Two micrograms of macrocyclic peptide R4-1 was transfected inside both HaCaT-Ras-mock and HaCaT-Ras-THG-1 cells using Xfect protein transfection. After 4 hours of incubation with

peptide transfection reagents, the cell culture medium was changed and cells were incubated for an additional 4 hours or 6 hours. Proteins were collected by direct cell lysis with 2x SDS and subjected to Western blot. As showed in Fig. 5.3 A, Nrf2 protein level was not changed in HaCaT-Ras-mock cells when treated with either R4-1 or scramble peptide. On the other hand, the Nrf2 protein level was reduced in R4-1 treated HaCaT-Ras-THG-1 cells comparing with scramble peptide treatment. Furthermore, Nrf2 target genes expression levels were evaluated. Figure 5.3 B showed the reducing of NQO1, TKT and PGD mRNA level in R4-1 treated HaCaT-Ras-THG-1 cells comparing with scramble treated cells. Consistent with Western blot data, this impact was not observed when comparing R4-1 and scramble peptide treatment in HaCaT-Ras-mock cells. Those results indicate R4-1 peptide's ability to restore Nrf2 protein and its target genes expression in THG-1 dependent manner. Effect of R4-1 macrocyclic peptide on cell proliferation was also evaluated and results are displayed in figure 5.3 C. In HaCaT-Ras- mock cells, no differences between the proliferation of cells treated with DMSO and R4-1 peptides. However, in HaCaT-Ras-THG-1 cells, cells treated with R4-1 peptide had lower proliferation rate compared with DMSO treated cells. From those data, R4-1 can reduce Nrf2 protein and its targets gene, therefore, affect proliferation of THG-1 expressing cells. On the other hand, no effect was observed in cells without THG-1 expression. In conclusion, with its ability to disrupt THG-1 and Keap1 interaction, R4-1 treatment cancels Nrf2 protein and target genes accumulation in cells in THG-1 dependent manner.

4. Cell-penetrating peptide: R4-1CPP-F

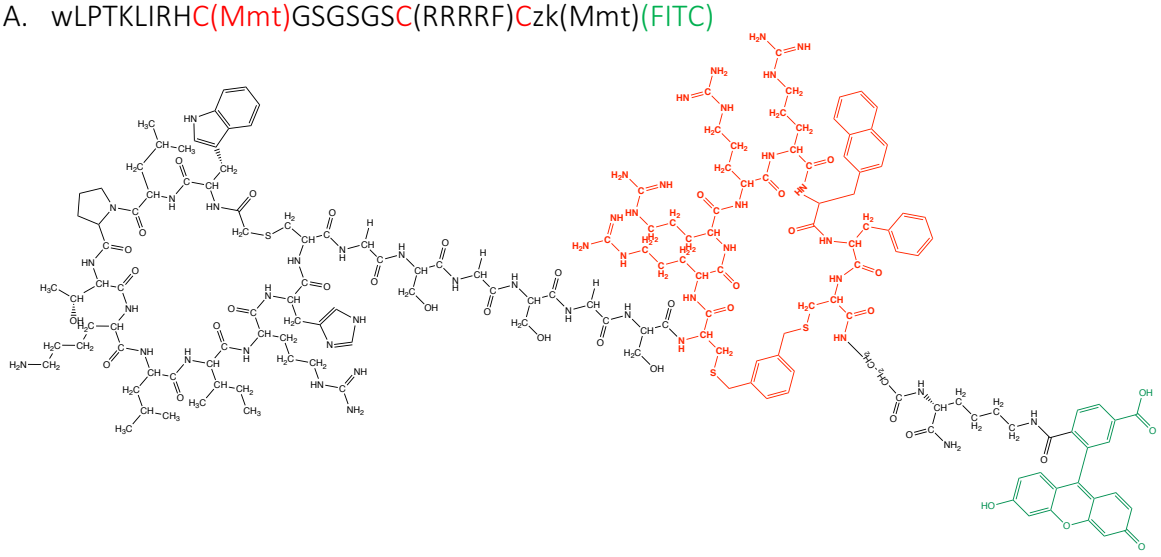
In order to simplify the delivery of macrocyclic peptide into cells, R4-1 was conjugated with cyclic heptapeptide cyclo FΦRRRRRRQ (cFΦR4, where Φ is L-2-napthylalanine). cFΦR4 were

reported to have the cell-penetrating ability through endocytosis. Strikingly, unlike most cell-penetrating peptide (CPP), cFΦR4 can escape early endosome into the cytoplasm (Quian Z., et al, 2014). With 4-12 times higher delivery efficiency, superior serum stability, cFΦR4 was suggested to be a useful delivery cargo.

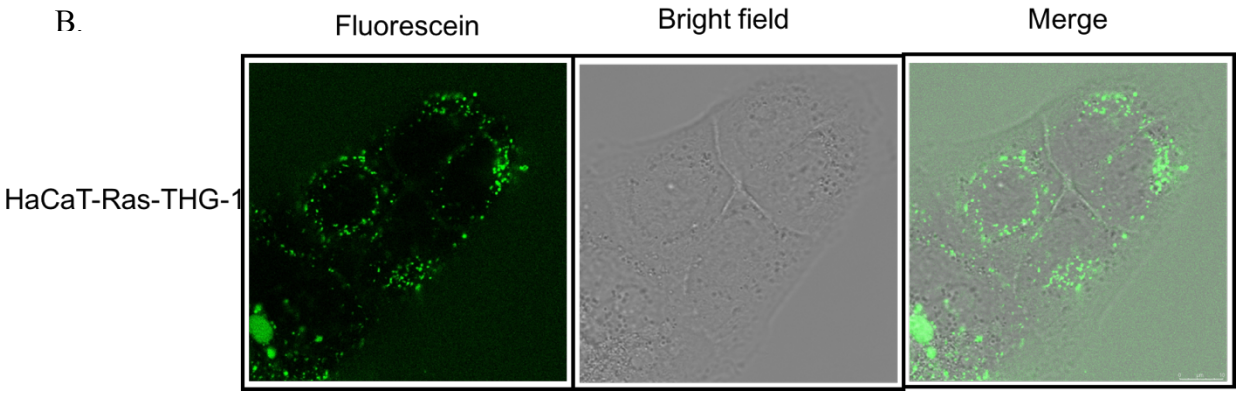
Synthesized fluorescein tagged-R4-1- cFΦR4 (R4-1CPP-F) conjugated (Figure 5.4A) was incubated with HaCaT-Ras-THG-1 cells for 4 hours. Microscopic observation and FACs were performed. As expected, without any transportation cargo, fluorescein signal accumulation was significantly increased inside cells (Figure 5.4 B). FACs data indicated that positive cells account for more than 90% of the total population when using 1 μM R4-1CPP-F (Figure 5.4C). However, the fluorescein pattern was not dispersed throughout cytoplasm at the time of detection (Figure 5.4 B). The punctate pattern indicated endocytoplasmic vesicles trap. Therefore, a time course experiment was performed to confirm whether R4-1CPP-F could escape endosome and deliver its activity. In figure 5.4 D, after 6 and 10 hours of incubation of cells with R4-1CPP-F, dot pattern was replaced by a more dispersed pattern. However, the fluorescein intensity was also reduced. From this results, a conclusion cannot be made on the escaping efficiency of R4-1CPP-F in the cytoplasm.

Figure 5.4: Cell-penetration of R4-1CPP. A. Chemical structure of R4-1CPP-F. B. Microscopic observation of live HaCaT-Ras-THG1 cells incubated with 1μM R4-1CPP-F for 4 hours. C. Fluorescein-activated cell sorting (FACs) data of HaCaT-Ras-THG1 cells incubated with 1μM R4-1CPP-F compared with cells incubated with non-treated cells. D. Microscopic observation of live HaCaT-Ras-THG1 cells incubated with 1μM R4-1CPP-F for 4, 6 and 10 hours

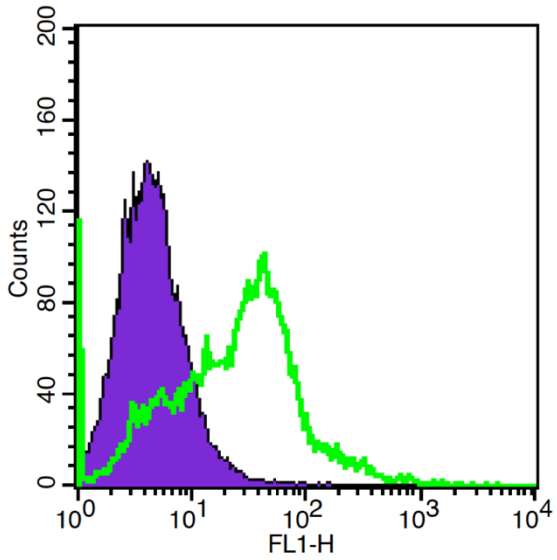
A. wLPTKLIRHC(Mmt)GSGSGSC(RRRRF)Czk(Mmt)(FITC)



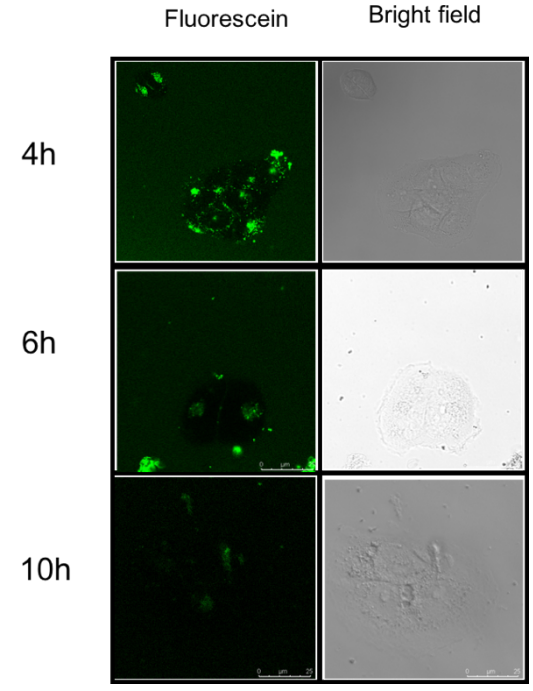
B.



C.



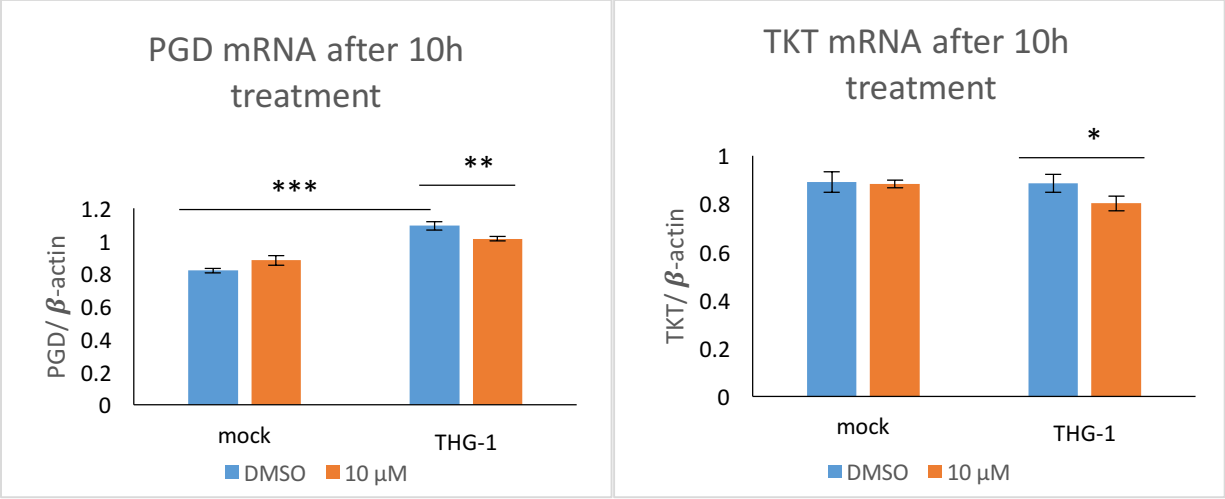
D.



5. *Effect of R4-1CPP-F on Keap1 and THG-1 interaction in HaCaT-Ras-THG-1 cells.*

Ten micromolar (10 μM) of R4-1CPP-F were incubated in the culture medium of HaCaT-Ras-mock or HaCaT-Ras-THG-1 for 10 hours. Total mRNA was extracted and cDNA was generated. As showed in Fig. 5.5, consistent with R4-1 transfection data, treatment of R4-1CPP-F reduced expression of Nrf2 target genes (TKT, PGD) in HaCaT-Ras-THG-1 but not HaCaT-Ras-mock when compared with DMSO treatment. Noticeable, different treatment duration and different cellular stress level lead to a different response in mRNA expression and may require a different amount of R4-1CPP-F (data not shown). This phenomenon may due to the difference in the mRNA expression level of genes or due to escaping efficiency of R4-1CPP-F. In addition, precipitation of R4-1CPP-F in culture medium was observed in 30 μM treatment, increasing of peptide treatment concentration may not be considered as a solution.

Figure 5.5: **Disruption of THG-1 and Keap1 interaction by R4-1CPP-F resulted in reducing of Nrf2 activity.** mRNA expression of PGD and TKT, target genes of Nrf2 when HaCaT-Ras-mock and HaCaT-Ras-THG-1 cells are treated with DMSO or 10 μM of R4-1CPP-F.



In this preliminary data of using R4-1CPP-F for treating HaCaT-Ras-THG-1 cells, optimization of treatment conditions need to carry out. In addition, the endosome escaping efficiency of R4-1CPP-F into cell cytoplasm is unclear. Living cells NanoBiT assay will be useful to confirm whether R4-1CPP-F can successfully escape endosome, an optimum time for escaping, and the effect of R4-1CPP-F on THG-1 and Keap1 interaction. Further experiments need to be performed.

6. Materials

Primers	Sequence 5'-3'
NQO1-F	AGAAGAGAGGATGGGAGGTA
NQO1-R	TGGTGATAGAAAGCAAGGTC
TKT-F	GCTGAACCTGAGGAAGATCA
TKT-R	TGTCGAAGTATTTGCCGGTG
PGD-F	ATATAGGGACACCACAAGACGG
PGD-R	GCATGAGCGATGGGCCATA
β-actin-F	GCTCATAGCTCTTCTCCAGGG
β-actin-R	TGAACCCTAAGGCCAACCGTG

7. Methods

a. Peptide transfection

Peptide transfection was carried out using Xfect Protein Transfection (Takara) follow the manufacture protocols. Briefly, 1x 10^6 cells were cultured in 6 wells plate overnight. In one Eppendorf tube, 1x Xfect transfection reagent is mixed with deionized water. In another Eppendorf tube, 2 μg of R4-1 was mixed with Xfect protein buffer. Then, R4-1 solution was

added into diluted 1x Xfect transfection reagent and incubated for 30 minutes at room temperature. During the incubation time, the culture medium was changed into 400 μL serum-free medium/ well. Add transfection reagent/ peptide mixture to cells in serum-free medium. After 4 hours of incubations, the medium was aspirated and changed into the original culture medium. Cells were incubated for addition 4 hours or 6 hours.

b. Protein extraction and Western-blot analysis

Culture medium was aspirated and cells were washed with PBS. Cells were solubilized by 2x SDS buffer. Collected protein solutions were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were electrotransferred to a mixed ester nitrocellulose membrane (Hybond-C Extra, GE Healthcare) and subjected to immunoblot analysis. Anti-Nrf2 (H-300, Santa Cruz Biotechnology), anti-β-actin (Sigma), anti-hemagglutinin (HA) (3F10, Roche Applied Science), anti-FLAG (M2, Sigma) were used as primary antibodies for the immunoblot analysis. Reacted primary antibodies were detected with horseradish peroxidase-conjugated anti-mouse IgG or anti-rabbit IgG antibodies (GE Healthcare) and SuperSignal West Dura Extended Duration Substrate (Thermo Scientific)

c. mRNA extraction and cDNA generation

Total RNA was prepared using ISOGENII (Nippon Gene). Reverse transcription was performed using High-capacity RNA-to-cDNA Master mix (Applied Biosystems).

d. qPCR

Quantitative PCR was performed using qPCR MasterMix for SYBER Green I (Applied Biosystems) with T7-CH-F46 and CGS3-CH-R39 primers and the ABI7500 Fast Sequence Detection system.

e. FACs

293T cells were cultured with cells density 2 x 10^5 cells/ well in 12 wells plate. Cells were transfected with R4-1F as described previously or incubated with 10 μM of R4-1CPP-F for 4 hours. Cells were treated with trypsin (Sigma) to obtain single cells. After washing with PBS, cells were spun down and suspended in 1% FCS-PBS. The samples were analyzed using BD FACSAria (BD Bioscience) and BD FACSDiva software (100 μm Sorp AriaII 5B 3R 3V 5YG). Negative control for R4-1F transfection was R4-1F incubation without the transfection reagent. Negative control for R4-1CPP-F was DMSO only.

CHAPTER VI: GENERAL CONCLUSION

1. *Establishment of selection system*

The technology involves flexizyme for reprogramming of translation and incorporation of non-canonical amino acid was introduce since 2007 (Ohuchi M., et al, 2007). The RaPID system based on the ability of RNA "flexizyme" to reprogram the genetic code for the incorporation of non-canonical amino acids in combination with mRNA-display. This system has been employ to screen for specific ligands with desired modifications against various proteins of interest. However, previously reported studies utilizing the RaPID system employed many lab –prepared components. In this research, I performed the RaPID selection using commercially available reagents except for flexizyme and chloroacetyl-D-tryptophan. Establishment of a protocol using commercially available enzymes increases the accessibility of the RaPID system to many laboratories that do not have enough resources, particularly personnel, to produce the various kinds of enzymes and other reagents required for full reconstitution of a translation system.

2. *Successful identification of THG-1 binding macrocyclic peptide*

Antibodies have been used for a long time as a detection probe as well as therapeutics. However, the selection and production of monoclonal antibodies are technically challenging. On the other hand, screening of macrocyclic peptide ligands using RaPID system usually involves 6 iterative rounds and require approximately two weeks for the whole process. In this study, after the establishment of the screening system, I successfully identified different macrocyclic peptides that bind to THG-1.

3. *Application of macrocyclic peptides*

Identified MCPs that bind to THG-1 were used for fluorescein staining, ELISA-like protein sorbent assay for quantification, and protein pull-down—tasks which have conventionally used antibodies. Direct staining by fluorescein-tagged MCPs had comparable or better specificity than that of indirect immunofluorescent staining. It will simplify and reduce the time consumed for fluorescein imaging.

R4-1B, R4-2B, and R4-4B all displayed ability to pull-down recombinant THG-1, to varying degrees. Their abilities to pull-down THG-1 overexpressed in HEK 293 and endogenous THG-1 in TE13 required larger quantities of MCPs. This difference may be caused by the difference in valency. Peptides are much easier to synthesize the multimers or derivatives. Multi-valent MCPs may become a better tool even for protein precipitation than antibodies.

4. *Bio-activity of identified macrocyclic peptide R4-1*

R4-1 macrocyclic peptide does not only display binding to THG-1 but also exhibits the ability to disrupt THG-1 and Keap1 interaction both in vitro and in cell assay. When delivered inside cells that overexpressed with THG-1, R4-1 can successfully restore the expression of Nrf2 and its target genes (NQO1, TKT, PGD). Cell proliferation was not affected in R4-1 treatment in HaCaT-Ras-mock cells, while displayed the proliferation inhibition effect in HaCaT-Ras-THG-1 cells. Those data suggested that targeting THG-1 and Keap1 interaction is a potential strategy for reducing cancer cells growth. In addition, R4-1 is a potential lead product for targeting this interaction.

5. *Chemical modification of macrocyclic peptide*

Cell penetration sequence was attached to the sequence of R4-1 without affecting its ability to disrupt THG-1 and Keap1 interaction. Incubation of 10 μM of R4-1CPP-F in culture medium for

10 hours showed a reduction in TKT and PGD mRNA level in HaCaT-Ras-THG-1 compared with DMSO control. Moreover, such an effect is not observed in HaCaT-Ras-mock. However, more characterization of R4-1CPP needs to be performed: cytosolic localization of R4-1CPP, bioavailability, serum stability, etc.

CHAPTER VII: FUTURE PERSPECTIVE

This research identified THG-1 macrocyclic peptide ligand, R4-1 with the ability to disrupt THG-1 and Keap1 interaction. Results from this research validate the ability to target THG-1 as a therapeutic strategy for squamous cell carcinomas. There will be several questions needed to be answered and multiple research strategies needed to be carried out to bring R4-1 finding into clinical application. Most important, binding characteristic between R4-1 and THG-1 need to be confirmed, including binding affinity and binding location. Cocrystallization of R4-1 and THG-1 can be used to understand binding between THG-1 and R4-1. From the results of cocrystallization, modification of R4-1 to improve its serum stability, permeability, specificity, and affinity will be easier to carry out and with more confidence.

REFERENCES

1. Ando N., Kato H., Igaki H., et al. A randomized trial comparing postoperative adjuvant chemotherapy with cisplatin and 5-fluorouracil versus preperative chemotherapy for advanced squamous carcinomas of the thoracic esophagus. Ann Surg Oncol. 19:68-74. (2012)
2. Ayroldi E., Migliorati G., Bruscoli S., et al. Modulation of T-cell activation by glucocorticoid-induced leucine zipper factor via inhibition of nuclear factor kappaB. Blood. 98: 743-753. (2001)
3. Blanchard P., Baujat B., Holostenco V., et al. MACH-NC collaborative group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): a comprehensive analysis by tumor site. Radiother Oncol. 354: 567-578. (2011).
4. Bonner JA., Harari PM., Giralt J., et al. Radiotherapy plus cetuximad for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. Lancet Oncol. 11: 21-28. (2010)
5. Borcoman E. and Tourneau CL. Pembrolizumab in cervical cancer: the lasted evidence and medical usefulness. Ther Adv Med Oncol. 9: 431-439. (2017)
6. Borel JF., Kis ZL., Beveridge T. The history of the discovery and development of Cyclosporin (Sandimmune®)". In Merluzzi VJ, Adams J. The search for anti-inflammatory drugs case histories from concept to clinic. Boston: Birkhäuser. 27–63(1995).
7. Brooks PJ., Zakhari S. Acetaldehyde and the genome beyond nuclear DNA adducts and carcinogenesis. Environ Mol Mut. 55: 77-91. (2014)

8. Cancer Genome Atlas network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 517: 576-582. (2015).
9. Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 489: 519–525. (2012).
10. Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 517: 576–582. (2015).
11. Cancer.org. Atlanta: American Cancer Society. Cancer Facts and Figures (2014). Available from http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/
12. Chan K., Han XD., An YW. An important function of Nrf2 in combating oxidative stress: detoxification of acetotaminophen. Proc. Natl. Acad. Sci. 98: 4611-16. (2001).
13. Cheng AC., Coleman RG., Smyth KT., et al. Structural- based maximal affinity model predicts small-molecule druggability. Nature Biotech. 25: 71-75. (2007).
14. Cheng TY. D., Cramb SM., Baade PD., et al. The international epidemiology of lung cancer: lastest trends, disparities, and timor characteristics. J Thorac Oncol. 11: 1653-1671. (2016).
15. Crosby T., Hurt CN., Falk S., et al. Chemoradiotherapy with or without ceuximad in patients with oesophageal cancer (SCOPE1): a multicenter, phase 2/3 randomized trial. Lancet Oncol. 14: 627-637. (2013).
16. Dela-Cruz CS., Tanoue LT., and Matthay RA. Lung cancer: epidemiology, etiology and prevention. Clin Chest Med. 32: 605-644. (2011).
17. Dixon AS., Schwinn MK., Hall MP., et al. NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. ACS Chem Biol. 11: 400-408. (2016)

18. Driggers EM., Hale SP., Lee J., and Terrett NK. The exploration of macrocycles for drug discovery-an underexploited structural class. Dugs Discov. 7: 608- 624. (2008).
19. Drilon A., Rekhtman N., Ladanyi M., Paik P. Squamous cell carcinomas of the lung: emerging biology, controversies, and the promise of targeted therapy. Lancet Oncol. 14: e148-26. (2012).
20. Du P., Huang P., Huang X., et al. Comprehensive genomic analysis of esophagus squamous cell carcinomas reveals clinical relevance. Scientific report. 7: 15324-333. (2017).
21. Ferlay, J., Soerjomataram I., Dikshit R., et al. GLOBOCAN 2012v1.0, Cancer Incidence and Mortality Worldwide. IARC Cancer Base No. 11. Lyon, France: International Agency for Research on Cancer. (2013).
22. Fiol DE., and Kultz D. Osmotic stress sensing and signaling in fishes. FEBS J. 274: 109-124. (2013).
23. Fosgerau K. and Hoffmann T. Peptide therapeutics: current status and future directions. Drug Discov. Today. 20: 122-128 (2015).
24. Ganesan A. The impact of natural products upon modern drug discovery. Curr Opin Chem Biol. 12: 306-317. (2008)
25. Goehler H., Lalowski M., Stelzl U., et al. A protein interaction network links GTI1, an enhancer of huntingtin aggregation, to huntington's disease. Mol Cell. 15: 853-865. (2004).
26. Hakoshima T. Leucine zipper. Encyclopedia of life science. (2014).
27. Hashiguchi A., Hitachi K., Inui M., et al. TSC-box is essential for the nuclear localization and antiproliferative effect of XTSC-22. Develop. Growth Differ. 49: 197-204. (2007).
28. Hast BE., Goaldfarb D., Mulvaney KM., et al. Proteomic analysis of ubiquitin ligase KEAP1

reveals associate proteins that inhibit NRF2 ubiquitination. Cancer Res. 73: 2199-2210. (2013).

29. Hayashi Y., Morimoto J., Suga H. In vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. ACS Chem. Biol. 7: 607–613. (2012).
30. Hayes J.D. and McMahonM. Nrf2 and Keap1 mutations: permanent activation of an adaptive response in cancer. Trends biochem Sci. 34: 176-188. (2009).
31. Hibi L., Trinh B., Patturajan M., et al. AIS is an oncogene amplify in squamous cell carcinomas. Proc Natl Acad Sci. 97: 5462-67. (2000).
32. Hipolito CJ., Bashirudin NK., Suga H. Protein cocrystallization molecule originating from in vitro selected macrocyclic peptide. Curr Opin Struct Biol. 26: 24-31. (2014).
33. Hipolito CJ., Tanaka Y., Katoh T., et al. A macrocyclic peptide that serves as a cocrystallization ligand and inhibits the function of a MATE family transporter. Molecules. 18: 10514-10530. (2013).
34. TSC22D4 protein-STRING interaction network. Available at: https://string-db.org/network/9606.ENSP00000300181
35. Huang CF., Zhang L., Ma SR., et al. Clinical significance of KEAP1 and NRF2 in oral squamous cell carcinoma. PLoS One. 8: e83479. (2013).
36. Inoue D., Suzuki T., Mitsushi Y., et al. Accumulation of p62/SQSTM1 is associated with poor prognosis in patients with lung adenocarcinomas. Cancer Sci. 2013:760-766. (2012).
37. Itoh K., Wakabayashi N., Katoh Y., Ishii, et al. Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles. Genes Cells. 8**:** 379-391. (2003).

38. Jemal A., Bray F., Center MM., et al. Global cancer statistics. CA Cancer J Clin. 61:69-90. (2011).

39. Jeong Y., Hoang NT., Lovejoy A., et al. Role of Keap1/ Nrf2 and Tp53 mutations in lung squamous cell carcinoma development and radiation resistant. AACR Cancer Discovery. 7: 86-101. (2016).

40. Jones A., Friedrich K., Rohm M., et al. TSC22D4 is a molecular output of hepatic wasting metabolism. EMBO Mol Med. 5: 294-308. (2013).

41. Jones S., Thorton JM. Principles of protein-protein interactions. Proceedings of the National Academy of Sciences of the United State of America. 93: 13-20. (1996).

42. Josephson K., Ricardo A., Szostak JW. mRNA display: from basic principles to macrocycle drug discovery. Drug discovery today. 19: 388-399. (2014).

43. Kamangar F., Dorces GM., Anderson WF. Pattern of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol. 24: 2137-2150. (2006).

44. Kaspar AA. and Reichert JM. Future directions for peptide therapeutic development. Drug Discov. Today. 18: 807-817. (2013).

45. Katsuragi Y., Ichimura Y. and Komatsu M. Regulation of the Keap1-Nrf2 pathway by p62/SQSTM1. Curr Oppi Toxicol. 1: 54-61. (2016).

46. Kawakami T., Ishizawa T., Fujino T., et al. In vitro selection of multiple libraries created by genetic code reprogramming to discover macrocyclic peptides that antagonize VEGFR2 activity in living cells. ACS Chem Biol. 8: 1205-1214. (2013).

47. Kawamura A., Munzel M., Kojima T., et al. Highly selective inhibition of histone demethylases

by de novo macrocyclic peptides. Nature Comm. 8: 14773. (2017).

48. Kessler SM., Laggai S., Schultheiss CS., et al. IMP2/ p62 induces genomic instability and an aggressive hepatocellular carcinoma. Cell Death Dis. 6: 31894. (2015).
49. Kester, H. A., Belanchetot C., den Hertog J., et al. Transforming Growth Factor-stimulated Clone-22 Is a Member of a Family of Leucine Zipper Proteins That Can Homo- and Heterodimer and Has Transcriptional Repressor Activity. Journal of Biological Chemistry 274: 27439-27447. (1999).
50. Kim YR., Oh JE., Kim MS., et al. Oncogenic Nrf2 mutations in squamous cell carcinomas of the oesophagus and skin. J Pathol. 220: 446-451. (2010).
51. Klaassen CD., Reisman SA. Nrf2 rescue: effects of the antioxidative/ electrophilic response on the liver. Toxicol. Appl. Pharmacol. 244: 57-65. (2010).
52. Kobayashi A., Kang MI., Okawa H., et al. Oxidative stress sensor Keap1 function as an adaptor for Cul3-based E3 ligase to regulate the proteasomal degradation of Nrf2. Mol. Cell. Biol. 24: 7130-39. (2004).
53. Landschulz WH., Johnson PF., McKnight SL. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science. 240: 1759–1764. (1988).
54. Liang F., Li Q., Li Z., et al. TSC22D2 interacts with PKM2 and inhibits cell growth in colorectal cancer. International j. of Oncology. 49: 1046-1056. (2016).
55. Luther A., Moehle K., Chevalier E., et al. Protein epitope mimetic macrocycles as biopharmaceuticals. Curr Opin Chem Biol. 38: 45-51. (2017).
56. Ma Q. Role of Nrf2 in oxidative stress and toxicity. Annu Rev Pharmacol Toxicol. 53: 401-426. (2013).

57. Maher J., Yamamoto M. The rise of antioxidation signaling- the evolution and hormetic actions of Nrf2. Toxicol. Appl. Pharmacol. 244: 4-15. (2010).
58. Marur S., Forastiere AA. Head and neck squamous cell carcinomas: update on epidemiology, diagnosis and treatment. Mayo Clin Proc. 91: 386-396. (2016).
59. Matsuda S, Koyasu S. Mechanisms of action of cyclosporine. Immunopharmacology. 47: 119–25. (2000).
60. Miller S. The structure of interfaces between subunits of dimeric and tetrameric proteins. Protein Eng. 3: 77-83. (1989).
61. Moi P., Chan K., Asusnis I., Cao A., Kan YW. Isolation of NFE2-related factor 2 (Nrf2), an NF-E2 like basic leucine zipper transcriptional activator that binds to tandem NF-E2/ AP1 repeat of the β-globin locus control region. Proc. Natl. Acad. Sci. 91: 9926-30. (1994).
62. Molina JR., Yang P., Cassivi SD., et al. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 83: 584-594. (2008)
63. Motohashi H. and Yamamoto M. Nrf2-Keap1 defines a physiological important stress response mechanism. Trends Mol. Med. 10: 549-57. (2004).
64. Nakamura M., Kitaura J., Enomoto Y., et al. Transforming growth factor-β-stimulated clone-22 is a negative feedback regulator of Ras/Raf signaling: implications for tumorigenesis. Cancer Sci. 103: 26-33. (2011).
65. Nantel A; Quatrano RS. Characterization of three rice basic/leucine zipper factors, including two inhibitors of EmBP-1 DNA binding activity. The Journal of Biological Chemistry. 271: 31296–31305. (1996).

66. Nguyen T., Sherrat PJ., Picket CB. Regulatory mechanism controlling gene expression mediated by the antioxidant response element. Annu. Rev. Pharmacol. Toxicol. 43: 233-60. (2003).

67. Ohh M., Park CW., Ivan M., et al. Ubiquitination of hypoxia inducible factor requires direct binding to the beta domain of von Hippel- Lindau protein. Nat Cell Biol. 2: 423-427. (2000).

68. Ohigashi Y., Sho M., Yamada Y., et al. Clinical significant of programed death-1 ligand -1 and programed death -1 ligand -2 expression in human esophageal cancer. Clin Cancer Res. 11: 2947-2953. (2005).

69. Ohta T., Ijima K., Miyamoto M., et al. Loss of Keap1 function activates Nrf2 and provides advanatages for lung cancer cell growth. Cancer Res. 68: 1303-1309. (2008)

70. Ohuchi M., Murakami H., Suga H. The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. Curr Opin Chem Biol. 11: 135-144. (2007).

71. Okawa H., Motohashi H., Kobayashi A., et al. Hepatocyte-specific deletion of the keap1 gene activates Nrf2 and confers potent resistance against acute drug toxicity. Biochem Biophys Res Commun. 339: 79-88. (2006).

72. Over B., McCarren P., Artursson P., et al. Impact of stereospecific intramolecular hydrogen bonding on cell permeability and physicochemical properties. J. Med. Chem. 57: 2746-2754. (2014).

73. Padmanabhan B., Tong KI., Ohta T., et al. Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. Mol Cell. 21: 689-700. (2006).

74. Penathur A., Gibson MK., Jobe BA. and Luketick AD. Oesophagus carcinomas. Lancet 381: 400-412. (2013).

75. Quian Z, LaRochelle JR., Jiang B., et al. Early endosomal escape of a cyclic cell-penetrating peptide allows effective cytosolic cargo delivery. Biochemistry. 53: 4034-4046. (2014).
76. Reichert J. M. Antibodies to watch in 2017. MABS. 9: 167-181. (2017).
77. Reisman SA., Yeager RL., Yamamoto M. and Klaassen CD. Increase Nrf2 activation in liver from Keap1 knowdown mice increase expressing of cytoprotective genes that detoxifies electrophiles more than those that detoxify reactive oxygen species. Toxicol Science. 108: 35-47. (2009).
78. Rekhman N., Paik PK., Arcila ME., et al. Clarifying the spectrum of driver oncogene mutations in biomarker-verified squamous carcinomas of lung: lack of EGFR/ KRAS and presence of PIK2CA/AKT1 mutations. Clin Cancer Res. 18: 1167-1176. (2012).
79. Ren D., Villeneuve NF., Jiang T., et al. Brusatol enhances the efficacy of chemotherapy by inhibiting the Nrf2-mediated defense system. PNAS. 108: 1433-1438. (2010).
80. Rezai T., Bock JE., Zhou MV., et al. Conformation flexibility, internal hydrogen bonding, and passive membrane permeability: successful in silico prediction of the relative permeabilities of cyclic peptides. JACS. 128: 14073-14080. (2006).
81. Rock JR., Randell SH., and Hogan BLM. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. Disease model and mechanism. 3: 545-556. (2010).
82. Sanchez P., Carcer GD., Sandovol IV., et al. Localization f atypical protein kinase C isoform into lysosome-targeted endosomes through interaction with p62. Mol Cell. Biol. 18: 3069-3080. (1998).

83. Shimizu H., Sato K., Berberich T., et al. LIP19, a basic region leucine zipper protein, is a Fos-like molecular switch in the cold signaling of rice plant. Plant and Cell Physiology. 46: 1623–34. (2005).

84. Singh A., Venkannagari S., Oh KH., et al. Small molecule inhibitor of Nrf2 selectively intervenes therapeutic resistence in KEAP1-defecient NSCLC tumors. ACS Chem Biol. 11: 3214-3225. (2016).

85. Solis LM., Behrens C., Dong W., et al: Nrf2 and Keap1 abnormalities in non-small cell lung carcinoma and association with clinicopathologic features. Clin Cancer Res. 16: 3743-3753. (2010).

86. Song Y., Li L., Ou Y., et al. Identification of genomic alteration oesophageal squamous cell cancer. Nature. 509: 91-95. (2014).

87. Sporn MB., Liby KT. Nrf2 and cancer: the good the bad and the important of context. Nature Review Cancer. 12: 564-571. (2012).

88. Sykiotis GP., and Bohmann D. Stress-activated cap'n'collar transcription factors in aging and human disease. Science signaling. 3: 1-22. (2010).

89. Taguchi K., Maher JM., Suzuki T., et al. Genetic analysis of cytoprotection function supported by graded expression of Keap1. Mol. Cell Biol. 30: 3016-3026. (2010).

90. Taguchi K., Motohashi H., Yamamoto M. Molecular mechanism of Keap1 and Nrf2 pathway in stress response and cancer evolution. Gene cells. 18: 23-140. (2011).

91. Talalay P., Dinkova-Kostova AT., Holzclaw WD. Importance of phase II gene regulation in protection against electrophile. Adv Enzyme Regul. 43: 121-134. (2003).

92. Tamir TY., Mulvaney KM., and Major MB. Dissecting the Keap1/Nrf2 pathway through

proteomics. Curr Opin Toxicol. 1: 118-124. (2016)

93. Tanaka, Y., Hipolito CJ., Maturana A., et al. Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. Nature. 496: 247–251. (2013).
94. Teyssonneau D. and Italiano A. Olaratumab for soft-sarcoma. Expert Opinion on Biological Therapy. 17: 1019-1025. (2017).
95. Thatcher N., Hirch FR., Szczesna A., et al. A randomized, multicenter, open-label, phase III study or gemcitabine-cisplatin (GC) chemotherapy plus necitumumab (IMC-11F8/LY3012211) versus GC alone in the first-line treatment of patients with stage IV squamous non-small cell lung cancer . J Clin Oncol. 32: 8008. (2014).
96. The Cancer Genomes Atlas research network. Comprehensive genomic characterization of squamous cell lung cancer. Nature. 489: 519-525. (2012).
97. The Huntington's disease collaborative research group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 72: 971-983. (1993).
98. Tong KI., Katoh Y., Kusunori H., et al. Keap1 recruits Neh2 through binding to ETGE and DLG motif: characterization of the two site molecular recognition model. Mol. Cell Biol. 26: 2887-900. (2006).
99. Tong KI., Padmanabhan B., Kobayashi A., et al. Different electrostatic potentials define ETGE and DLG motifs as hinge and latch in oxidative stress respose. Mol. Cell Biol. 27: 7511-7521. (2007).
100. Torre LA., Bray F., Siegel RL., et al. Global cancer statistics, 2012. CA Cancer J Clin. 65: 87–108. (2015).

101. Tsuchida K., Tsujita T., Hayashi M., et al. Halofuginone enhances the chemo-sensitivity of cancer cells by suppressing Nrf2 accumulation. Free Radic Biol and Med. 103: 236-247. (2017).
102. Venkatesan K., Rual JF., Vazquez A., et al. An empirical framework for binary interactome mapping. Nat Methods 6: 83-90. (2009).
103. Wakabayashi N., Itoh K., Wakabayashi J., et al. Keap1- null mutation leads to postnatal lethality due to constitutive active of Nrf2. Nat Genet. 35: 238-245. (2003).
104. Walter DM., Cho HY., Kleeberger SR. Oxidative stress and antioxidant in the pathogenic of pulmonary fibrosis: the potential role of Nrf2. Antioxid. Redox. Signal. 10: 321-32. (2008).
105. Wang L., Kim D., Wise JTF., et al. P62 as a therapeutic target for inhibition of autophagy in prostate cancer. The Prostate. 78:390-400. (2017).
106. Wessjohann, LA., Ruijter E., Riverna DG., Brandt W. What can a chemist learn from nature's macrocycles? A brief, conceptual view. Mol. Divers. 9: 171-186. (2005).
107. Wilbertz T., Wagner P., Petersen K., et al. SOZ2 gene amplification and protein overexpression are associated with better outcome in squamous cell lung cancer. Mod Phathol. 24: 944-53. (2011).
108. Wyss AB., Hashibe M., Lee YA., et al. Smokeless tobacco use and the risk of head and neck cancer: Pooled analysis of US studies in the INHANCE Consortium. Am J Epidemiol. 184: 703-716. (2016).
109. Xiang X., Qin HG., You XM., et al. Expression of P62 in hepatocellular carcinoma involving hepatitis B virus infection and aflatoxin B1 exposure. Caner Med. 6: 2357-2369. (2017).
110. Yamagata K., Goto Y., Nishimasu H., et al. Structural basis for potent inhibition of SIRT2

deacetylase by a macrocyclic peptide inducing dynamic structural change. Structure 22: 345-352. (2014).

111. Yamagishi Y., Shoji I., Miyagawa S., et al. Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. Chem Biol. 18: 1562-1570. (2011).
112. Yoon CH., Rho SB., Kim ST., et al. Crucial role of TSC-22 in preventing the proteosomal degradation of p53 in cervical cancer. PLoS One. 7: e42006. (2012).
113. Zhang L., Zhou Y., Cheng C., et al. Genetic analyses reveal mutational signatures and frequently altered genes in esophagus squamous cell carcinomas. Am J Hum Gent. 96: 597-611. (2015).
114. Zhang W., Zhu H., Liu X., et al. Epidermal growth factor is a prognosis predictor in patients with esophagus squamous cell carcinomas. Ann Thorac Surg. 98: 513-519. (2014).
115. Zorzi A., Deyle K., Heinis C. Cyclic peptide therapeutics: past, present and future. Chem. Biol. 38: 24-29. (2017).

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Prof. Mitsuyasu Kato for his devoted guidance, valuable suggestion, and encouragement throughout my Ph.D. research. Your dedication, patience, experiences, and knowledge have led me through obstacles and challenges during this work. Your enthusiasm and vision in research keep me on track, have a strong belief in the future of science and my contribution.

I would like to express my thank to Associate Prof. Hyroyuki Suzuki for accepting me as a member of your group. Thank you for your suggestions and details experiments explanation. Thank you for always keep track of my works.

I would like to show my appreciation to Assistant Prof. Christopher J Hipolito for kindly explain to me all of my questions and wondering in chemistry as well as in researcher life. I couldn't accomplish those projects without your supporting. Through discussion with you, I have learned a lot.

I also express my deep thanks to Prof. Peter ten Dijke for your insight and generous help. I also express my gratitude to Associate Prof. Yuki Goto, Assistant Prof. Naohiro Terasaka for valuable assisting and discussion during my work at Suga's laboratory.

I am also indebted to other members of Prof. Kato's research group for their stimulating discussion and encouragement.

Finally, a huge thanks to my family who are strongly reliable and deeply supporting me. I would also thank my husband Mr. Nguyen Thanh Luan for endless cherishing and encouragement.