Development of a Novel Insoluble Tag INSOL for Simplified Protein Preparation and its Application to the Analysis of Autoantibody Response to Cancer Vaccine

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biological Science

Eriko FUKUDA

Table of contents

Abstract

Protein preparation is an important part of proteomic analysis. Tags used for protein preparation are generally soluble tags, such as GST-tag and His-tag. However, when a sufficient amount of the target protein cannot be obtained in the soluble state, a technique for rendering the target protein insoluble by fusing an insoluble tag may be effective. In particular, insoluble tags are used in the following cases: when the overexpression of the target protein is detrimental to the survival and/or proliferation of the host cell in the expression system, or when the soluble/insoluble fraction of the expressed target protein is biased toward the insoluble fraction. The insoluble tags, developed thus far, have two main problems that hinder their use in proteomics analysis. The first is that the main insoluble tag is developed in the *E. coli* expression system, and it takes time to purify the target protein, and it is not suitable for preparation of more than a few hundred proteins. The purification process requires cell disruption and multiple washings of insoluble fractions. The second point is that there are only few target proteins that could be insolubilized (in many cases, several; at the most, several tens of types). Therefore, in this study, I developed a novel insoluble tag, INSOL-tag, that can universally insolubilize any target protein in a cell-free expression system and can be easily purified by centrifugation only. Chapter 1 describes the process of developing the INSOL-tag and the protocol to create a protein array using proteins prepared with INSOL-tag.

In addition, as an example for the effective use of INSOL-tag, I developed a protein array with cancer antigens (INSOL CTA-array) that tends to be insoluble. Cancer antigens are proteins that are specifically expressed in cancer cells and are important because they serve as landmarks when the immune system attacks cancer cells. Cancer vaccine therapy is the latest treatment expected to have the following effects: cancer antigen peptides are administered as vaccines to cancer patients and are taken up by the dendritic cells, followed by an activation of the cancer antigen-specific

cytotoxic T cells that attack the cancer cells. Cancer vaccine therapy activates the immune function, which specifically attacks the cancer cells. Therefore, cancer vaccine therapy is less likely to kill the healthy cells when compared with either chemotherapy or radiation therapy, that is, it is less likely to cause side effects. However, in establishing the use of cancer vaccines as standard therapy, there was a lack of means for evaluating their therapeutic effects in clinical trials. The number of cancer cell-specific cytotoxic T cells is very small, approximately 0.01% of the total lymphocytes, and it is generally difficult to detect them immediately after blood collection. As an alternative, an evaluation method that measures the changes in the type and amount of antibodies produced against cancer antigens, leaked from apoptotic cancer cells due to the action of cytotoxic T cells, i.e., "antigen spreading," has attracted attention. Therefore, I developed a method for measuring "antigen spreading." Chapter 2 describes the results of measuring "antigen spreading" in cancer patients who received cancer vaccine therapy using the protein array with cancer antigens (INSOL CTA-array).

In this study, I developed a new insoluble tag, INSOL-tag, which could insolubilize more than 300 target proteins in multiple categories—that is, the universality of the insolubility of INSOL-tag was shown. I also succeeded in developing a protein array (cancer antigen array, i.e., INSOL CTA-array) specialized for specific categories of antigens that are difficult to prepare.

Since the protein insolubilized with INSOL-tag needs to be dissolved with a surfactant or denaturant, its use is limited to the experimental systems that do not depend on the secondary or higher structural levels of proteins. However, INSOL-tag is highly advantageous, in terms of universality, regardless of the type of target protein, and have a convenient protocol of purification by a single step of centrifugation. The effective applications of INSOL-tag mainly include (1) the preparation of hundreds to tens of thousands of proteins for designing the protein arrays, (2) uniform preparation of antigen proteins in animals for antibody production, and (3) preparation of hundreds to tens of thousands of internal standard proteins for absolute quantification in mass spectrometry.

Up to 101 cancer antigens were spotted on the commercially available protein array, and the variations in the amount of protein bound on the array was as large as about 4×10^4. Whereas the INSOL CTA-array, developed in this study, was spotted with 267 cancer antigens, which is significantly more than that observed in the existing commercial products. Further, the variation in protein binding amount was reduced by about 87 times. This array can be used not only for the evaluation of cancer immunotherapy, including cancer vaccine treatment, but also for searching the markers that are used for the early diagnosis of cancer, selection of treatment candidates, and prediction of recurrence. In conclusion, this study provides useful experimental tools for basic research in molecular biology and tumor immunology.

Abbreviations

4AaCter	C-terminal region of *Bacillus thuringiensis* Cry4Aa toxin
ACPP	acid phosphatase, prostate
CBP	calmodulin binding peptide
CD8	cluster of differentiation 8
CTA	cancer testis antigens
CTL	cytotoxic T lymphocyte
CV	coefficient of variation
DC	dendritic cell
E. coli	*Escherichia coli*
ELP	elastin-like polypeptides
Fh8	*Fasciola hepatica* 8kDa
GST	glutathione-S-transferase
HA	hemagglutinin
HLA	human leukocyte antigen
HuPEX	Human Proteome Expression-Resource
KSI	ketosteroid isomerase
MafG	MAF bZIP transcription factor G
MBP	maltose binding protein
PagP	PhoPQ-activated gene P
PDZ	post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1)
RTX	repeat in toxin
SDS	sodium dodecyl sulfate
SEREX	serological identification of antigens by recombinant expression cloning
SSX2, SSX2B	Synovial sarcoma X breakpoint 2, 2B
SYT	Synovial sarcoma translocated to X chromosome protein
TCR	T cell receptor

General Introduction

Tags for recombinant protein preparation

In proteomics research, preparation of recombinant proteins is very important. When preparing the *in vivo* (*E. coli*, cultured cells) or *in vitro* (cell-free system) recombinant proteins, a purification 'tag' is generally used for purification. As shown in Figure 1, purification tags can be broadly classified into affinity purification tags (soluble tags) and centrifugal purification tags (insoluble tags). Affinity purification tags are peptides or polypeptides that have an affinity toward a particular protein or antibody. Centrifugal purification tags cause a tag-dependent precipitation or aggregation of the target proteins. In addition to purification of the target protein, there are also tags that enable the detection of the target protein or improve its solubility. Although many tags have been developed for decades, there are no tags that can be used on all hosts, and there are no universal tags that allow for the preparation of all recombinant proteins. For this reason, researchers have used multiple tags in combination and developed new tags with improved performances.

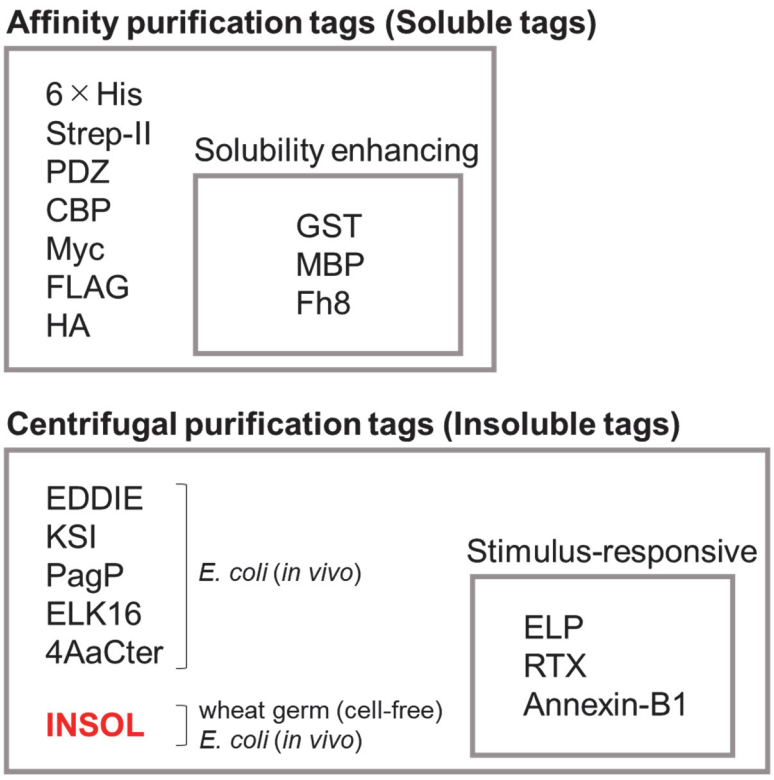


Figure 1 – Tags for recombinant protein purification

Tags that are generally used for protein purification are broadly classified into two groups, and the representative tags belonging to each group are described: PDZ, post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1); CBP, calmodulin binding peptide; HA, hemagglutinin; GST, glutathione-S-transferase; MBP, maltose binding protein; Fh8, *Fasciola hepatica* 8 kDa; KSI, ketosteroid isomerase; PagP, PhoPQ-activated gene P; 4AaCter, C-terminal region of *Bacillus thuringiensis* Cry4Aa toxin; ELP, elastin-like polypeptides; RTX, repeat in toxin.

Affinity purification tags (soluble tags)

Affinity purification tags are fused to the N- or C-terminal of a target protein, and are selectively recovered based on the specific affinity to the binding partner. The most commonly used affinity purification tag worldwide is the polyhistidine tag, namely His-tag (HHHHHH; Kuo, W.H. *et al*. 2011). The advantages of the His-tag are that the resin for purification is inexpensive, smaller in size, has little effect on the structure and function of the target protein, and functions without a specific fold. Conversely, contamination by other proteins containing histidines is the biggest problem in His-tag purification. Strep-tag II (Maertens, B. *et al*. 2015), PDZ domain (Walkup, W.G. 4th. *et al*. 2014), and CBP (Zheng, C.F. *et al*. 1997) are used to purify the target protein by utilizing the affinity with Strep-tactin, peptides like YKQTSV or SIESDV, and calmodulin, respectively. Myc-tag (EQKLISEEDL; Kramer, A. *et al*. 1997), FLAG-tag (DYKDDDDK; Schmidt, P.M. *et al*. 2012) and HA-tag (YPYDVPDYA; Wilson, I.A. *et al*. 1984) are the epitope tags that bind to their corresponding immunoaffinity resins with a high specificity and affinity. These tags have higher purity and yields than the His-tag, but these are suitable only for the small-scale experiments due to the costlier immunoaffinity resins. GST-tag (26 kDa; Smith, D.B. *et al*. 1988), MBP (43 kDa; Kellermann, O.K. *et al*. 1982) is a polypeptide tags that have high specific affinity for small ligands. Fh8 (8 kDa) is derived from a calcium-binding protein from the parasite *Fasciola hepatica* (Costa, S.J. *et al*. 2013a). The mechanism by which Fh8 binds during hydrophobic interaction chromatography is calcium-dependent in a low salt medium (Costa, S.J. *et al*. 2013b). The characteristics of these three tags are that they increase the solubility of the target proteins. Owing to the large size of the tags, the overall yield tends to be lower than that when using shorter tags. Occasionally, it is necessary to cut and remove the tag portion after purification.

Centrifugal purification tags (insoluble tags)

When a recombinant protein is expressed in *E. coli*, the protein may aggregate to form an inclusion body. Although the formation of an inclusion body is often considered a disadvantage, it actually has several advantages, such as enhanced expression level of target protein, high purity, resistance to proteolytic digestion, and relatively easy isolation of inclusion bodies. EDDIE (23 kDa; Cheng, X. *et al*. 2010), the mutant of Npro, is an autoprotease of classical swine fever virus. KSI (ketosteroid isomerase) is a 28 kDa homodimeric protein of *Pseudomonas testosteroni* (Kuliopulos, A. *et al*. 1987). PagP (Hwang, P.M. *et al*. 2012) is a 53 kDa membrane protein of *Pseudomonas aeruginosa*. ELK16 (LELELKLKLELELKLK; Xing, L. *et al*. 2011) is a self-assembling amphipathic peptide. 4AaCter (17 kDa; Hayashi, M. *et al*. 2013) is derived from the C-terminal region of the Cry4Aa toxin of *Bacillus thuringiensis*. These tags improve the expression level of the target protein in the inclusion body of *E. coli* and enable their recovery with high purity. ELP (Banki, M.R. *et al*. 2005) is derived from the native mammalian elastin protein, and it consists of a sequence, VPGXG (X is any amino acid except proline), that repeats 60-110 times. RTX (Shur, O. *et al*. 2013) is a 38 kDa tag consisting of a natural repeat-in-toxin domain (GGXGXDXUX: X is any amino acid, U is hydrophobic amino acid) secreted by the bacterial type I secretion system. Annexin B1 (Ding, F.X. *et al*. 2007) is a 17 kDa calcium-dependent phospholipid protein from *Cysticercus cellulosae*. ELP aggregates above a certain temperature, RTX forms a disordered structure and precipitates in the presence of calcium, while the Annexin B1 binds with the phospholipids in a Ca^{2+} dependent manner to form a precipitate. These three tags are stimulus-responsive insoluble tags. The insoluble tags mainly adopt a strategy of recovering the target protein from the inclusion body of *E. coli*, and require not only the centrifugation of an insolubilized target protein but also cell lysis and washing of the inclusion body. The INSOL-tag, developed in this study, is a 13 kDa insoluble tag derived from the human transcription factor MafG. INSOL-tag, can be used not only in *E. coli* but also in a wheat germ

cell-free expression system. In a cell-free system, the protein of interest can be easily purified by single centrifugation only.

Cancer cells and the immune system

It is estimated that the number of human cells is about 37.2 trillion (Bianconi, E. *et al*. 2013), and from the frequency of cell renewal and the probability of gene copy error, it is thought that thousands of cancer cells are born in healthy people every day that are eliminated by the immune system. In cancer cells, proteins whose expression levels are abnormally increased and proteins that are not expressed in the normal cells, are also expressed. Self-proteins misfolded due to abnormal expression, mutated self-proteins, and self-proteins that are not expressed in normal cells are degraded into peptides by the proteasome complex. The peptides bind to MHC class I in the endoplasmic reticulum and then displayed on the cell surface. When cancer cells encounter $CD8^+$ T cells that recognize the presented peptide, apoptosis of the cancer cells is induced (Figure 2). $CD8^+$ T cells are a type of CTL (cytotoxic T lymphocyte), and the cancer antigens recognized by CTLs are classified in Figure 3. Mutant antigens (light blue and red) expressed as a result of a gene mutation that causes cancer are presented as mutant peptides by HLA class I in cancer cells (Coulie, P.G. *et al*. 1995, Kawakami, Y. *et al*. 2001, Vigneron, N. *et al*. 2002). Antigens (purple) expressed only in cancer and germ cells, represented by cancer testis antigens, are not expressed in other cells, and the corresponding peptide is presented in cancer cells (van der Bruggen, P. *et al*. 1991, De Plaen, E. *et al*. 1994). Products of differentiation-related genes (blue), which are expressed only in certain tissues, may be presented on both, normal and cancer cells (Wölfel, T. *et al*. 1994, Hoek, K.S. *et al*. 2008). Furthermore, a protein (yellow) that is slightly expressed in a healthy cell, may be presented in a larger amount than healthy cells as a result of overexpression in cancer cells (Vigneron, N. *et al*. 2013). The former two are antigens that are specifically expressed in cancer cells, while the latter two are antigens

that are also expressed in other healthy cells. Negative selection in the thymus should kill T cells that bind to self-molecules, but as suggested, this central tolerance is an incomplete mechanism. So far, it has been suggested that the type of proteasome is different between thymic epithelial cells and cancer cells (Nil, A. *et al*. 2004), and that the peptide sequences resulting from the same self-antigen may be different between thymus and cancer.

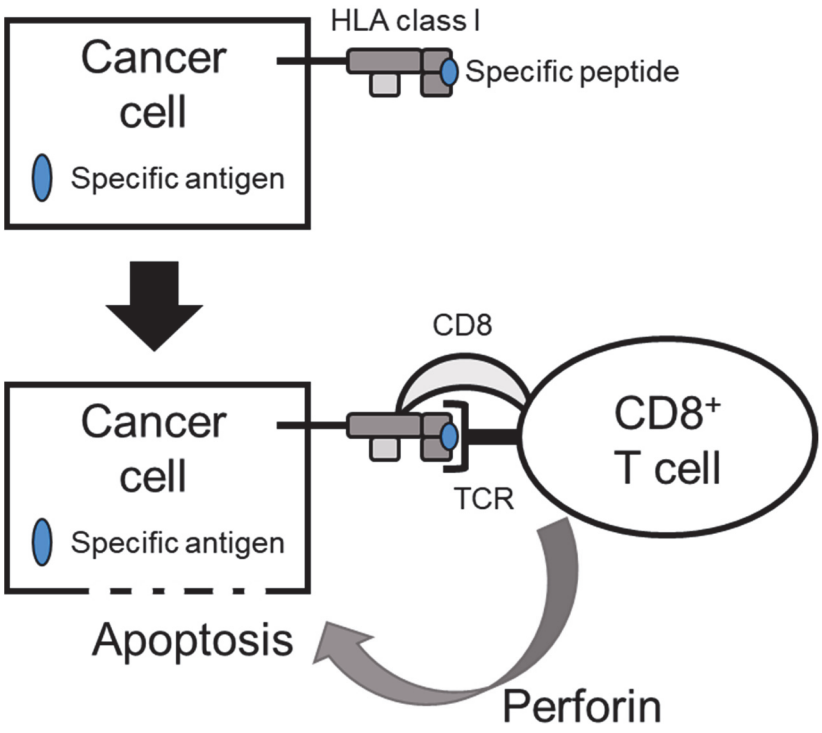


Figure 2 – How CD8⁺ T cell eliminates cancer cells

Specific antigens expressed in cancer cells are degraded by the proteasome, and the specific peptides generated are bound to HLA class I and presented on the cell surface. $CD8^+$ T cells expressing TCRs that recognize a specific peptide secrete molecules such as perforin and induce apoptosis in cancer cells. HLA, human leukocyte antigen; CD8, cluster of differentiation 8; TCR, T cell receptor.

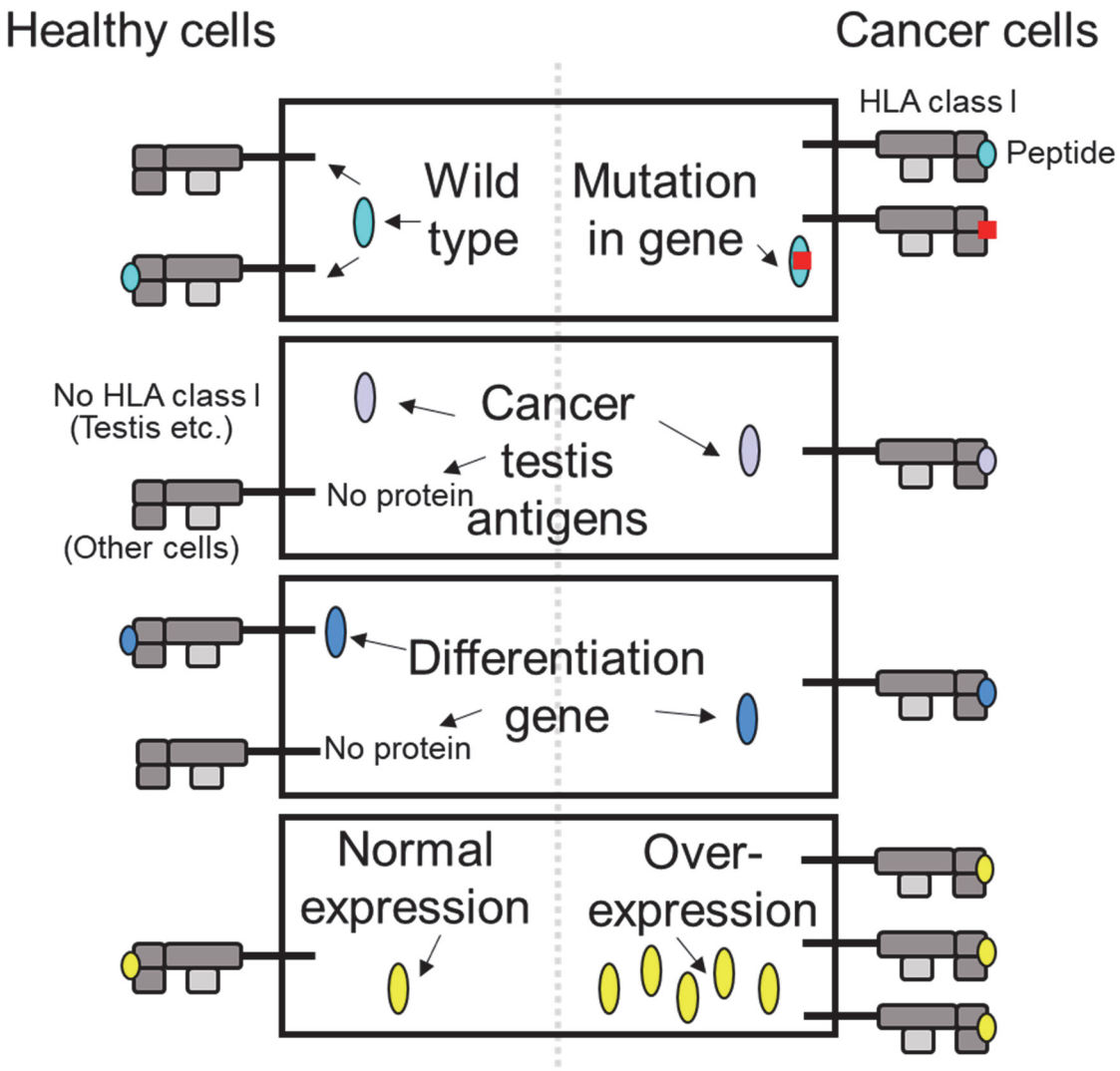


Figure 3 – Classification of cancer antigen-derived peptides recognized by CTL

Schematic drawing showing that the antigens expressed in healthy cells and cancer cells are degraded by a proteasome, and are presented by HLA class I.

Assessment of cancer vaccine treatment and monitoring of anti-cancer antigens

Cancer cells may weaken or abolish the functioning of immune cells to avoid an attack by the immune system (Facciabene, A. *et al*. 2012). Also, if the patient's immunity is weakened, it may not be able enough to attack the cancer cells. A cancer vaccine is a treatment in which an antigen recognized by a cancer-specific CTL is administered to a cancer patient as a vaccine. The aim of the treatment is to activate the patient's inherent ability to attack the cancer cells. As shown in Figure 4, when a vaccine antigen is taken up into a patient's dendritic cells, CTLs ($CD8^+$ T cells) and helper T cells are stimulated by the antigen presentation. It is difficult to monitor the presence of CTLs specific for cancer vaccine antigens and their activation status due to the low frequency of CTLs. Therefore, in recent years, serum anti-vaccine antigen antibodies have been detected as an evidence to the activation of helper T cells. Besides, when a cancer antigen, which is different from a vaccine antigen, leaks as a result of apoptosis of a cancer cell, an autoantibody to the cancer antigen may be produced (this phenomenon is called "antigen spreading"). Measuring antigen spreading is one of the few techniques that provide an indicator for the evaluation of cancer vaccine treatment (Corbière, V. *et al*. 2011). Therefore, in this study, I developed a protein array that can measure the autoantibodies against 276 cancer antigens, and I have further tried to measure the antigen spreading.

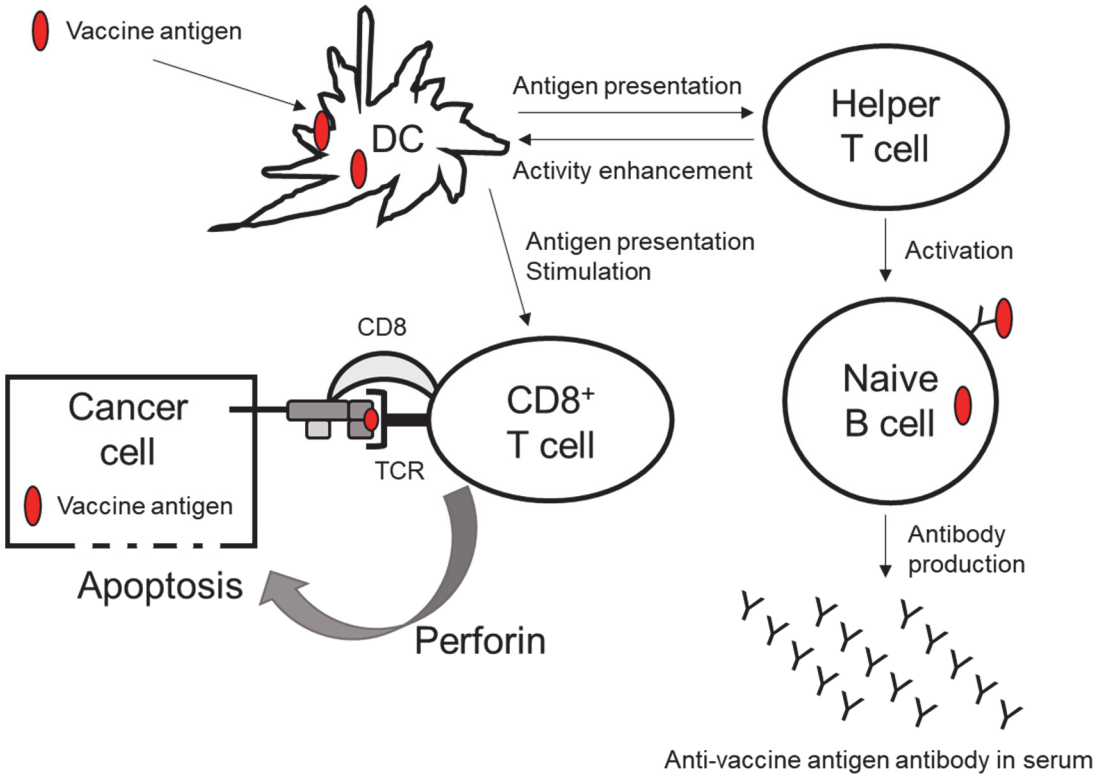


Figure 4 – Functions of T cells and B cells during cancer vaccine administration

When a cancer vaccine antigen (red) is administered to a cancer patient, dendritic cells take up the cancer vaccine antigen, and helper T cells and $CD8^+$ T cells are stimulated downstream by the antigen presentation. Naive B cells, activated by helper T cells, produce anti-vaccine antigen antibodies. DC, Dendritic cell.

Chapter 1

Development of novel INSOL-tag enabling proteome-wide protein handling and application to protein array

Summary

Proteomic analysis requires protein tags that enable high-throughput handling, however, the versatile tags that can be used in *in vitro* expression systems are currently lacking. In the study described in Chapter 1, I developed an insoluble protein tag, INSOL-tag, derived from the human transcription factor MafG. The INSOL-tagged target protein is expressed in a eukaryotic *in vitro* expression system and recovered as a pellet following centrifugation at 19,000 × g for 20 min. Comparisons of the target protein recovery rates of GST-tag and INSOL-tag using 111 cytoplasmic proteins revealed a four-fold increase in the yield of INSOL-tagged proteins. The detection limit of the array was approximately 11.1 pg IgG, and the correlation with ELISA was high ($R^2 > 0.91$). These data suggest that INSOL-tag is a versatile tag that can insolubilize a wide range of target proteins. It is therefore expected to become a powerful tool during comprehensive protein preparations for protein arrays, antibody production, and mass spectrometry.

Introduction

Proteomics analyses, which involve high-throughput expression, purification, and recovery, often result in yield bottlenecks of the target protein due to the low expression level of proteins, low solubilization rate, and poor purification efficiency, when using the standard purification methods. This is because, the optimal conditions for preparation differ between proteins in terms of the expression system and/or host, the type of tag for purification and its fusion position, the expression temperature and time, and the composition of the eluate. More than a dozen different protein or peptide tags have been recently developed; however, no universal tag yet exists, and researchers are therefore required to find the optimal tag for their target protein. A simple protein preparation method that can be used with many types of protein is therefore desired.

There are a number of effective methods that allow purification of a target protein with high purity at low cost, whereby the target protein is insolubilized and easily recovered by centrifugation or filtration. Many tags that insolubilize the target protein have been developed, and these are well documented（Wood, D.W. 2014, Lin, Z. *et al*. 2015, Yadav, D.K. *et al*. 2016). These tags are derived from the enzymes of *Escherichia coli* (Lee, J.H. *et al*. 2000, Derynck, R. *et al*. 1984), other bacteria (Kuliopulos, A. *et al*. 1994, Hwang, P.M. *et al*. 2012), and classical swine fever virus (Achmüller, C. *et al*. 2007), while others are self-assembling peptides (Xing, L. *et al*. 2011) or 5-9 amino acid (a.a.) repeats (Banki, M.R. *et al*. 2005, Shur, O. *et al*. 2013), or derived from the domains required for the hetero-dimerization of human transcription factors (Vidovic, V. *et al*. 2009). ELK16 (16a.a.) and 18A (18a.a.) (Xing, L. *et al*. 2011) are exceptionally short tags, however, most major tags are large, with an average size of about 300 a.a. (range between 125–735a.a.). Each of these tags has characteristics that allows them to be insolubilized while maintaining their enzyme activity, and/or can be insolubilized by stimulation with temperature or calcium after being expressed in a soluble state.

However, most tags have been functionally verified by an *in vivo* expression in *E. coli*, with no examples in the eukaryotic *in vitro* expression systems. Because it takes time to transform and disrupt cells, the *in vivo* expression system of *E. coli* is not suitable for comprehensive expression of 100 or more proteins. Further, the number of examples of insolubilized target proteins remains very limited (from one to a few dozen at most). Proteomic analyses of eukaryotes, including humans, therefore lack insolubilizing tags that can be comprehensively used in the eukaryotic *in vitro* expression systems.

Dr. Naoki Goshima and his colleagues have developed a human full-length cDNA library and expression resource (Human Proteome Expression-resource: HuPEX) to allow a comprehensive expression (Goshima, N. *et al*. 2008) and have used the resource in various proteomic analyses. Among them, Dr. Masatoshi Mori identified an extremely insoluble human protein that cannot be obtained in a soluble state, namely, MafG. In this study, I truncated MafG while maintaining its insolubility, and developed it as an insoluble tag (INSOL-tag) that allows the universal insolubilization of target proteins in a eukaryotic *in vitro* expression system. The INSOL-tagged target proteins can be purified by centrifugation only (Figure 5). Use of the INSOL-tagged protein as an internal standard protein for mass spectrometry was also examined. Arginine residues were inserted into the INSOL-tag, cleaving the tag to a maximum of 6 a.a. via tryptic digestion in order to eliminate the noise from the tag itself in mass spectrometry measurements. As an example of effective use of this INSOL-tag, I developed a method to fabricate the protein array using INSOL-tagged proteins. As a result, INSOL-tagged protein array was found to be highly sensitive and correlated well with ELISA in measuring serum autoantibodies. INSOL-tag is therefore expected to become a powerful tool in comprehensive protein preparation for protein arrays, antibody production, and mass spectrometry.

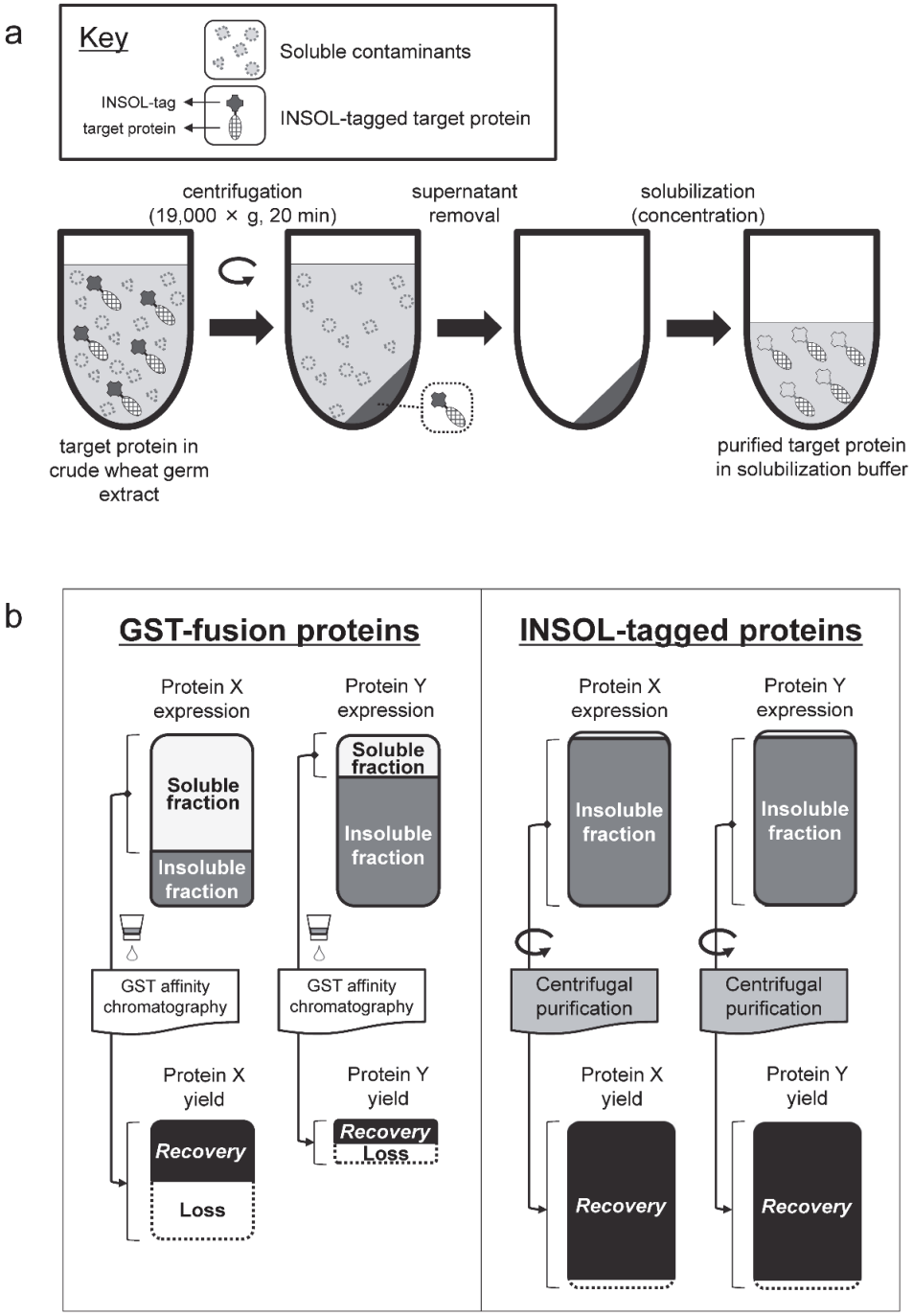


Figure 5 – Schematic illustrations of the proposed new protein purification method using INSOL-tag

(a) Centrifugal purification of recombinant proteins with INSOL-tag. (b) Comparison of GST affinity purification and INSOL-tag centrifugal purification.

Materials and Methods

Strains and plasmids

Escherichia coli DH5α was used as the host strain for the construction of vectors. Human cDNA in the Gateway entry clone was selected from HuPEX (Goshima, N. *et al*. 2008). Destination vectors were used as follows: pGOLD2314 (MafG), pGOLD2320 (A), pGOLD2322 (A + B), pGOLD2324 (B), pGOLD2326 (B + C), pGOLD2328 (C), pGOLD2332 (B + C + R×10), pGOLD2334 (B + C1 + R×6), pGOLD2336 (B + C1 + C2 + R×8), pGOLD2338 (B + C1), pGOLD2340 (B + C1 + C2), and pGOLD2344 (B + C + R×17) for the MafG-derived tags, pEW-5SG and pEW-3GS for the GST-tag (Goshima, N. *et al*. 2008), and pEW-5FG for the FLAG-tag and GST-tag (Goshima, N. *et al*. 2008). To construct pGOLD2314, the gene encoding MafG was obtained by PCR from the MafG entry clone (FLJ81899AAAF). The PCR product was inserted in frame into a self-made destination vector, pMEMORY8190 (GenBank accession number: LC498629), as an *Xho*I-*Kpn*I PCR fragment. To construct pGOLD2320, pGOLD2332, pGOLD2324, pGOLD2326, pGOLD2328, pGOLD2338, and pGOL2340, positions M1 to R55, M1 to S109, R56 to S109, R56 to S162, K110 to S162, R56 to V127, and R56 to K145 in MafG were respectively inserted in frame into pGOLD2314 as *Xho*I-*Kpn*I PCR fragments. To construct vectors with arginines inserted (pGOLD2332, pGOLD2334, pGOLD2336, and pGOL2344), primers designed for site-directed mutagenesis were used.

Expression of recombinant proteins in vitro

Recombinant proteins were produced using the wheat germ cell-free protein expression system as previously described (Goshima, N. *et al*. 2008). Briefly, the LR reaction (Thermo Fisher Scientific, Waltham, MA) was performed using the entry clones and destination vectors, after which DNA fragments containing the SP6 promoter, tags, open reading frame (ORF), and 3′-untranslated region

were amplified by PCR with KOD Dash DNA polymerase (Toyobo, Osaka, Japan). mRNA was generated from the PCR products by SP6 RNA polymerase. *In vitro* translation was carried out by mixing the mRNA with wheat germ extract WEPRO7240G (CellFree Sciences, Yokohama, Japan).

Purification of MafG-fused and INSOL-tagged proteins - centrifugation purification -

An aliquot (~25μl) of each crude protein solution expressed with wheat germ was transferred into a 0.2-ml tube and diluted four-fold with phosphate buffered saline (PBS) containing 1% (v/v) Tween 20 and centrifuged at 19,000 × g for 20 min at 4°C to recover the MafG-fused or INSOL-tagged protein as an insoluble fraction (precipitate). The supernatant was removed and the sediment obtained (insoluble fraction) was designated the purified fraction.

Purification of GST-tagged protein

Purification of 111 recombinant proteins was performed using Glutathione Sepharose 4B (GE Healthcare, Chicago, IL) as previously described, with slight modifications (Takeda, H. *et al*. 2010).

SDS-PAGE

Each sample was treated with NuPAGE™ LDS Sample Buffer (NP0007, Thermo Fisher Scientific) and electrophoresed in NuPAGE Bis-Tris gel (4%–12%) at 200 V for 47 min according to the manufacturer's instructions (Thermo Fisher Scientific). The resolved proteins were stained with Coomassie Brilliant Blue (CBB).

Dilution solution for purification of MafG-fused protein

N-term MafG-fused MAPKAPK3 (FLJ80026AAAF), DAPP1 (FLJ84009AAAF), and Venus were expressed and centrifuged, after which the purity of the target protein in the precipitate was confirmed.

As a negative control, a sample was subjected to translation by adding water instead of mRNA (mock). The crude solution obtained after expression was diluted as shown followed by centrifugation at 19,000 × g for 20 min at 4°C:

No dilution (25 μL crude solution, 0 μL diluent)

1/4 dilution with PBS (*) (25 μL crude solution, 75 μL PBS)

1/4 dilution with PBST+ (**) (25 μL crude solution, 75 μL $PBST^+$)

(*) PBS: 137.0 mM NaCl, 8.1 mM Na_2HPO_4, 2.7 mM KCl, and 1.5mM KH_2PO_4

(**) $PBST^+$: 1% (v/v) Tween 20 in PBS

Precipitates from 25 μL crude solution were separated by SDS-PAGE and stained with CBB. The target protein and other proteins (contaminants) were quantified using a BSA dilution series. LabScan 6.0 on ImageScanner III (GE healthcare) was used to obtain gel images, and ImageQuant TL version 7.0 was used for quantification. The target protein percentage was determined by dividing the target protein by the sum of target protein and contaminants.

Liquid sample holder

The liquid sample holder was formed as previously described (Ogura, T. 2014). It comprised an upper Al holder and a lower acrylic resin portion that maintained the sample solution (wheat germ crude solution or purified INSOL-tag protein suspended in 100 mM phosphate buffer; pH 7.8) at atmospheric pressure between the SiN films (Ogura, T. 2014, Ogura, T. 2015). A 20 nm-thick SiN film supported by a 0.1 mm^2 window in a Si frame (4 mm^2, 0.381 mm thick; Silson, UK) was coated with tungsten (W) using a magnetron sputtering device (Model MSP-30T; Vacuum Device Inc., Japan) as previously described (Ogura, T. 2014, Ogura, T. 2015). The upper W-coated SiN film was attached to the Al holder with double-sided tape, and the W layer on the SiN film was connected to the Al holder with silver conductive ink.

High-resolution SE-ADM and FE-SEM setup

The high-resolution SE-ADM imaging system is based on field-emission SEM (SU5000, Hitachi High-Technologies Co., Japan). The liquid sample holder was mounted onto the SEM stage, and the detector terminal was connected to a pre-amplifier under the holder (Ogura, T. 2015). The electrical signal from the pre-amplifier was passed through a low-pass filter and fed into the external input terminal of the FE-SEM system. The SE-ADM images were captured under 5,000–50,000× magnification with a scanning time of 80 s, working distance of 7 mm, EB acceleration voltage of 2.5 kV, and current of 10 pA.

Image processing and measurements of particle diameter

The SE-ADM images obtained using the SU5000 SEM system were transferred to a personal computer (Intel Core i7, 3.2 GHz, Windows 7). The LPF and scanned signals were processed into high-resolution SE-ADM images using the image-processing toolbox of MATLAB R2014a (Math Works Inc., USA). The original SE-ADM images were filtered through a 2D Gaussian filter (GF) with a kernel size of 11 × 11 pixels and a radius of 1.2σ. The background was removed by subtracting the SE-ADM images from the filtered images using a broad GF (400 × 400 pixels, 200σ). The Feret diameter (the longest distance perpendicular to the two parallel outer tangents to the object) of the particles was measured using ImageJ software (US National Institutes of Health, Bethesda, MD).

Solubilization of INSOL-tagged proteins

INSOL-tagged IgG (IGHG3: FLJ40036AAAF) was expressed and centrifugal purification was performed. After removal of the supernatant, solution Nos. 1 to 5 and 3-1 to 3-4 (Table 1, refer to P. 40) were added, and the precipitates were vortexed for solubilization. The suspensions were

centrifuged at 19,000 × g for 20 min to separate the soluble and the undissolved fractions. Each fraction was separated by SDS-PAGE and quantified as described above. The IgG solubilization rate (%) of each solution was calculated under the assumption that the target protein solubilization amount of solution No. 3 was 100.

Evaluation of the INSOL-tag solubilization solution used to determine the amount of protein binding to the protein array

The solubilized INSOL-tagged IgG solutions described above were spotted onto the active esterified aC plate (Table 1). The amount of protein bound to the aC plate was detected with a primary (anti-MAFG antibody, 1:250, ab154318; Abcam) and secondary antibody (Goat anti-rabbit IgG H&L Alexa Fluor 647, 1:1,000, ab150079; Abcam). The intensity rate (%) of each solution was calculated under the assumption that the target protein binding amount of solution No. 3 was 100.

Protein immobilization on amorphous carbon (aC)

Carboxy group-immobilized aC plates (Nippon Light Metal, Tokyo, Japan) were active-esterized by shaking in 200mM N-Hydroxysuccinimide (#130672; Sigma-Aldrich, St. Louis, MO), 200mM 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (#348-03631; Wako Pure Chemical Industries, Osaka, Japan), and 100 mM phosphate buffer (pH 5.8) for 2 h at 37°C. The aC plates were washed with milliQ water and dried using a compressed-air duster (PC3-5.5TL, CA-30; IAC Co., Ltd., Yokohama, Japan). Protein solution was printed on the aC plates using HORNET-NX (Wako Pure Chemical Industries) equipped with a 1536 pin head. A 0.457-mm diameter pin with a 10-nl slot tip (FP1CS10; V&P Scientific, San Diego, CA) was used. The active-esterized carboxy group and amino group in the proteins were used to cause a dehydration reaction, thereby immobilizing the proteins on the aC plate. The aC plates were stored in the dark at 4°C until use.

INSOL-array assay

Unreacted residues of active ester on the aC plates were blocked with TBST at room temperature (RT) for 15 min with shaking and treated with ImmunoBlock (Sumitomo Dainippon Pharma, Osaka, Japan) at RT for 1 h. To detect immobilized INSOL-tagged IGHG3 or purified human IgG (Thermo Fisher Scientific), the aC plates were incubated for 1 h at RT with Goat anti-Human IgG (H+L) Secondary Antibody Alexa Fluor 647 (1:1,000 dilution; Thermo Fisher Scientific). Anti-MAFG antibody (1:250 dilution, ab154318; Abcam, Cambridge, UK) followed by Goat anti-rabbit IgG H&L Alexa Fluor 647 (1:1,000 dilution, ab150079; Abcam) were used to detect INSOL-tag. To detect autoantibodies in human sera, the aC plates were incubated for 1 h at RT with the sera (1:100 dilution). Goat anti-Human IgG (H+L) Secondary Antibody Alexa Fluor 647 (1:1,000 dilution; Thermo Fisher Scientific) was used as the second antibody. All antibodies and sera were diluted in TBST containing 5% (v/v) ImmunoBlock.

Serum samples

A serum pool from 10 healthy donors (Cat.No. 12181201, Lot No. BJ3529D) was obtained from Cosmo Bio, Ltd. (Tokyo, Japan). Sera from anti-NY-ESO1-positive and anti-MAGEA4-positive patients and gastrointestinal cancer patients were provided by Dr. Hiroshi Shiku of Mie University from the CHP-MAGE-A4 vaccine Phase I + II clinical trial (Clinical trial registration number: UMIN000001999).

ELISA

Ninety-six-well immunoplates (Nunc) coated with recombinant MAGE-A4 protein (20 ng/50 μL/well) and NY-ESO-1 protein (10 ng/50 μL/well) were incubated overnight at 4°C. Plates were

washed with 0.05% Tween 20/PBS and blocked with 1% bovine serum albumin/PBS for 2 h at RT. After washing, 100 μL of serum serially diluted with 1% bovine serum albumin/PBS was added to the plates and incubated at 4°C overnight. After washing, diluted goat antihuman IgG (H + L chain)-HRP (MBL) was added and incubated for 5 h at 4°C. After washing, TMB Substrate Kit (Thermo Fisher Scientific) was added to each well and incubated for 10 min at RT. After adding 0.18 M H_2SO_4, absorbance at 450 nm was measured using a Microplate Reader (model 550; Bio-Rad Laboratories, Hercules, CA).

Array Analysis

The aC plates were scanned at a 10-μm resolution using a GenePix® Professional 4200A microarray scanner (Axon Instruments, Foster City, CA), and scanned images were saved for analysis as 16-bit tiff files using GenePix® Pro ver.6.0 (Axon Instruments). Analysis software Array-Pro Analyzer ver.6.3.1 (Media Cybernetics, Rockville, MD) was used to record the mean by drawing an equally-sized circle around the spots. Net intensity was calculated by subtracting the background (local ring) from each sample intensity. To normalize differences between aC plates, the net intensity of each plate was corrected using the calibration curve of purified human IgG and calculated in units of pg IgG. In addition, Smirnov-Grubbs test ($p < .05$) was performed for each array plate, with antigens that were outliers for both spots regarded as positive. Smirnov-Grubbs test was performed using Ekuseru-Toukei 2012 (Social Survey Research Information, Japan).

Results

Insolubilization of the MafG protein and its ability to insolubilize the target protein

First, MafG was expressed alone to confirm the insolubility of MafG. A wheat germ cell-free system was used for protein expression (Goshima, N. *et al*. 2008). The obtained crude solution was centrifuged to separate it into soluble and insoluble fractions followed by SDS-PAGE analysis (Figure 6 left upper). As a result, MafG itself was easily insolubilized in the wheat germ cell-free expression system. MafG was fused to another target protein to examine the ability to insolubilize the target protein. GST protein was selected as a representative solubility-enhancing tag and expressed by fusing MafG to its N- or C-terminus. As a result, MafG successfully insolubilized the GST protein in all cases (Figure 6 left middle, left lower). Similar experiments were performed using the soluble fluorescent protein Venus; however, N- and C-terminal Venus fusion GST proteins were detected in the soluble fraction (Figure 6 right). These findings suggest that when MafG is fused to another protein as a tag, the protein is obtained in an insolubilized state.

The purification method of the MafG-fused protein was subsequently examined. To do so, MafG-Venus was expressed in a wheat germ cell-free system and centrifuged at 19,000 × g for 20 min. The supernatant was removed and the pellet was solubilized with SDS-PAGE sample buffer and separated by SDS-PAGE, allowing the ratio of MafG-Venus to contaminating proteins to be calculated from the intensity of CBB staining. The purity of MafG-Venus in the pellet was calculated as approximately 43.3% (Table 2). In order to improve the purity, a wheat germ cell-free crude solution was diluted 4-fold with PBS (137.0 mM NaCl, 8.1 mM Na_2HPO_4, 2.7 mM KCl, and 1.5mM KH_2PO_4) or 1% (v/v) Tween 20 in PBS and centrifuged. Using N-term MafG-fused proteins (MAPKAPK3, DAPP1, Venus), the purity of the target protein in the pellet was improved to 55.8%–58.7% (Table 2). This purification method was termed “centrifugal purification.”

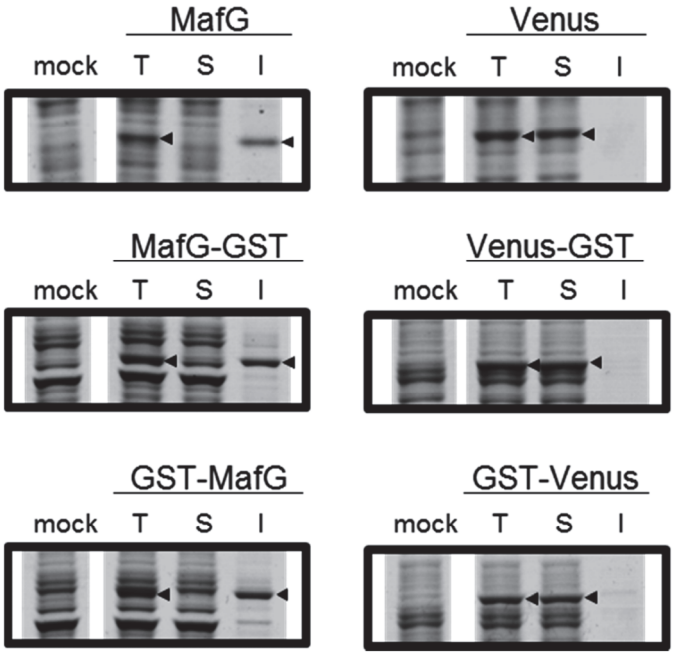


Figure 6 – MafG insolubilization of the target proteins

CBB staining of SDS-PAGE gel for MafG, MafG-GST, GST-MafG, Venus, Venus-GST, and GST-Venus expressed in a wheat germ cell-free system. The soluble and insoluble fractions of the reaction mixture were obtained by centrifugation at 19,000 × g for 20 min at 4°C. The band indicated by an arrowhead represents the target protein. T, total reaction mixture; S, soluble fraction (supernatant); I, insoluble fraction (pellet).

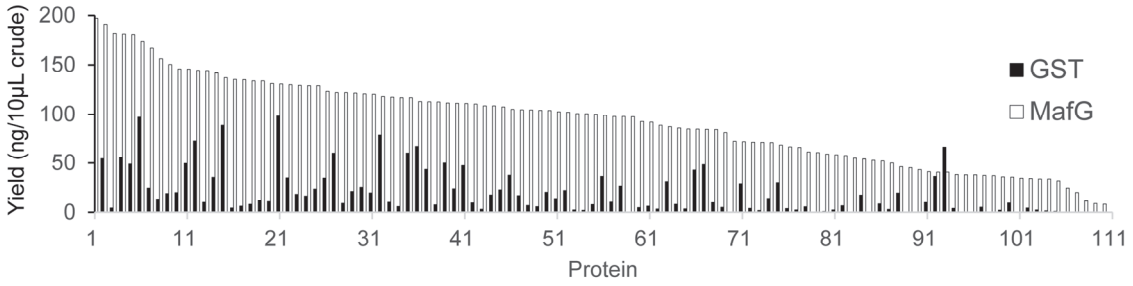
	MAPKAPK3			DAPP1			Venus			mock	
	target protein (ng)	contaminant (ng)	target protein (%)	target protein (ng)	contaminant (ng)	target protein (%)	target protein (ng)	contaminant (ng)	target protein (%)	target protein (ng)	contaminant (ng)
no dilution	287.2	405.1	41.5	209.3	940.8	18.2	347.9	455.8	43.3	-	381.3
1/4 dilution with PBS	291.5	587.6	33.2	463.7	671.4	40.9	399.9	749.0	34.8	-	692.3
1/4 dilution with PBST+	274.9	208.7	56.9	411.3	326.4	55.8	347.8	244.7	58.7	-	147.0

Table 2 – Purification conditions and purity of the MafG-fused protein
PBST+: 1% (v/v) Tween 20 in PBS

Comparison of GST-tag purification and MafG-tag purification (centrifugal purification)

The above results suggest that MafG has a strong ability to insolubilize target proteins; therefore, the insolubilization ability of MafG and the universality of its performance as a purification tag were subsequently determined. To do so, 60 cell-cycle related proteins and 51 signaling-related proteins were selected as target proteins (Table 3, refer to P. 44-47), and expressed as N-terminal GST-tag-fused and N-terminal MafG-fused proteins, respectively. For GST-tag purification, the amount of target protein in the elution fraction during purification with GST affinity chromatography was analyzed, while in MafG-tag purification, the recovery amount of target protein in the pellet following centrifugal purification was analyzed. The yield of target protein for each purification method is shown in Figure 7a, and an example of an SDS-PAGE image of a purified protein is shown in Figure 7b. Although one target protein could not be recovered by MafG-tag purification, the remaining 110 proteins yielded an average of 93.03 ng/10 μL crude extract. A protein that could not be recovered faced problems during amplification of the PCR product as a template for mRNA synthesis. Failed expression of the protein was due to the PCR failure. Meanwhile, GST-tag purification failed to recover the target protein for 17 out of 111 proteins. In addition, the average recovery of target protein for the remaining 94 proteins was only 23.47 ng/10 μL crude extract. These results suggest that it is possible to uniformly recover the target protein by MafG-tag purification. In addition, the yield of target protein with MafG-tag purification was about four times larger on average than with GST-tag purification.

a



b

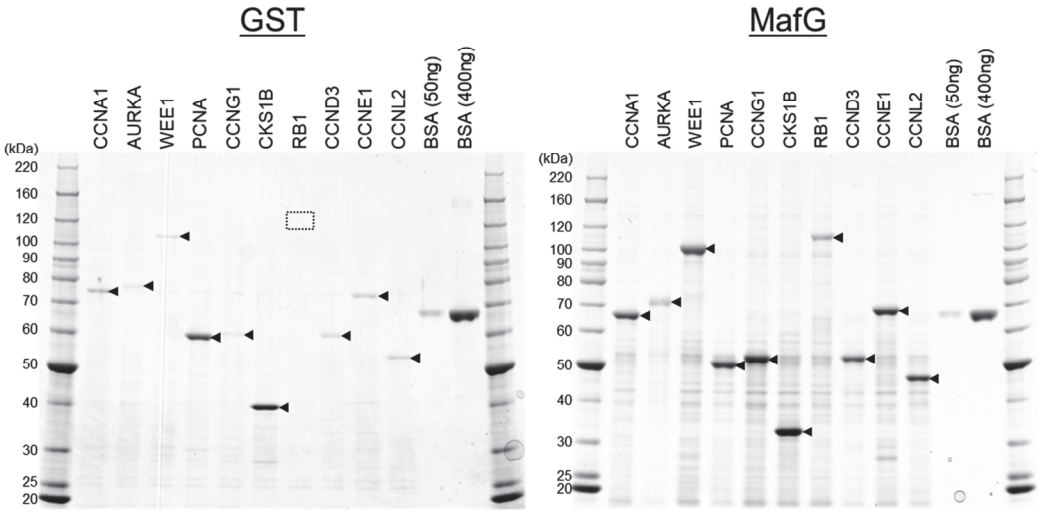


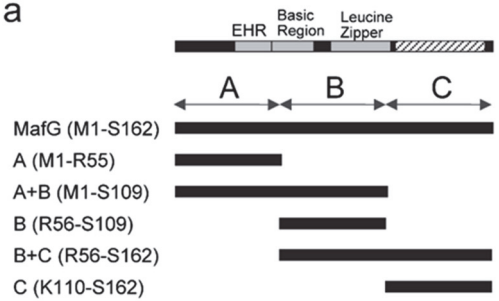
Figure 7 – Comparison of GST-tag purification and MafG-tag purification

(a) Yield comparison of 111 cytoplasmic proteins between GST affinity purification and MafG centrifugal purification. Sixty cell-cycle related proteins and 51 signaling-related proteins involved in the 111 cytoplasmic proteins were expressed in a wheat germ cell-free system. The elution fraction of GST affinity chromatography and pellet fraction of the MafG-fused proteins were analyzed by SDS-PAGE. The yield of each protein (ng/10 μL) was determined using BSA as a standard. (b) Representative example of CBB stained SDS-PAGE patterns of purified proteins obtained with GST affinity purification and MafG-centrifugal purification. Arrowheads represent the target proteins. The dashed box indicates where no band was detected.

Truncation of MafG-tag

MafG is 162 a.a. long, making it a relatively large tag. To identify the area related to insolubilization and improve its versatility as a tag, I attempted truncation. First, MafG was divided in three, A (M1-R55), B (R56-S109), and C (K110-S162), and five fragments of MafG were prepared, A, A + B, B, B + C, and C, as shown in Figure 8a. Each candidate was fused to the N-terminal of Venus or GST, and the amount of target protein recovered in the pellet by centrifugal purification was examined. The relative yield of truncated MafG fused target proteins were compared with full-length MafG-tag. Both Venus (99.5%) and GST (96.2%) had the highest recovery rates with the B + C tag (Figure 8b).

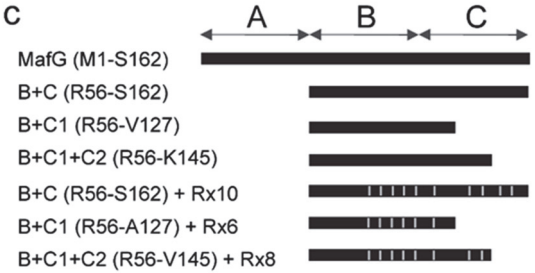
a



b

Venus	Percent yield (%)	GST	Percent yield (%)
MafG-Venus	100.0	MafG-GST	100.0
(A)-Venus	7.8	(A)-GST	12.3
(A+B)-Venus	84.0	(A+B)-GST	81.1
(B)-Venus	74.5	(B)-GST	42.6
(B+C)-Venus	99.5	(B+C)-GST	96.2
(C)-Venus	20.2	(C)-GST	32.6

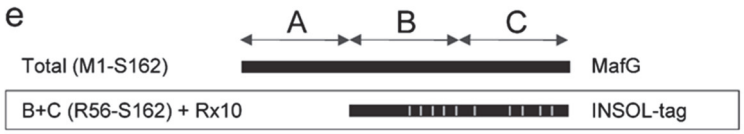
c



d

Venus	Percent yield (%)	GST	Percent yield (%)
MafG-Venus	100.0	MafG-GST	100.0
(B+C)-Venus	99.9	(B+C)-GST	97.6
(B+C1)-Venus	72.9	(B+C1)-GST	56.8
(B+C1+C2)-Venus	77.7	(B+C1+C2)-GST	51.2
(B+C)+Rx10-Venus	90.3	(B+C)+Rx10-GST	102.4
(B+C1)+Rx6-Venus	78.8	(B+C1)+Rx6-GST	64.2
(B+C1+C2)+Rx8-Venus	66.8	(B+C1+C2)+Rx8-GST	63.3

e



INSOL-tag (118 a.a.)

MRRTLKNRGYAASCRVKRVTQKEE***R***LEKQKA***R***ELQQEV***R***EKLASE***R***NASM
KL***R***ELDALRSKYEAL***R***QTFARTVARSPVAPARGPLAAG***R***LGPLVP***R***GKVA
AT***R***SVITIV***R***KSKTDARS

Figure 8 – Systematic truncation analysis of MafG protein for insolubility.

(a) and (c) The MafG truncation mutants analyzed in this study. MafG contains an extended homology region (EHR), a basic region, leucine zipper structure, and a region containing the nuclear-matrix targeting signal (dashed area). First, MafG was divided along its length into regions A, B, and C. Region C was trisected into C1, C2, and C3. White bars represent inserted arginine residues. (b) and (d) Yield comparisons of MafG and MafG truncation mutants as a tag for centrifugal purification. MafG and MafG truncation mutants were fused to the N-terminus of Venus or GST. The proteins were expressed and the pellet was obtained by centrifugation. The purified pellet was analyzed by SDS-PAGE and the yield of each protein was determined. Percent yield (%) indicates the percentage relative to 100% MafG-Venus or MafG-GST. (e) Amino acid sequence of the INSOL-tag. ***R*** represents the inserted arginine-residue.

Further truncation of the MafG-tag and arginine insertion for MS analysis

Further truncation of B + C was subsequently attempted. As shown in Figure 8b, recovery of the target protein with C alone was lower than that with B. The C domain was therefore considered to have low insolubility, and tag candidates were subsequently prepared by removing the latter 1/3 or 2/3 of domain C (B + C1, B + C1 + C2) (Figure 8c). As a result, in both Venus and GST, B + C recovered more than B + C1 and B + C1 + C2 (Figure 8d). At the same time, arginine residues were inserted into the tag so that the tag itself was digested with trypsin when preparing the internal standards for mass spectrometry. In both Venus and GST, even when 10 arginines were inserted into B + C, the recovery amount of (B + C) + R × 10 decreased by only about 10% with B + C and was larger than that of (B + C1) + R × 10 and (B + C1 + C2) + R × 10 (Figure 8d). In the case of (B + C1 + C2) + R × 17 (Figure 9), in which 17 arginines were inserted, the recovery amount of target protein was reduced by an average of 30% to 40%. Arginine is therefore thought to reduce the insolubility of the tag. From the above result, (B + C) + R × 10 was selected as the optimized MafG-tag, termed "INSOL-tag", and used in the subsequent experiments. The a.a. sequence of INSOL-tag is shown in Figure 8e. When an experiment similar to that shown in Figure 7a was performed using INSOL-tag, an equal or higher amount of target protein was recovered compared with MafG-tag (Figure 10).

B+C (R56-S162) + Rx17

MRRTLKNRGYA***R***ASCRVKRVTQKEE***R***LEKQKA***R***ELQ***R***QEV***R***EKLASE***R***N
AS***R***MKL***R***ELD***R***ALRSKYEAL***R***QTFARTVARSPV***R***APARGPLAAG***R***LGP***R***LVP
RGKVAAT***R***SVI***R***TIV***R***KSKTDARS

R: seventeen inserted arginines

Figure 9 – The amino acid sequence of "B+C (R56-S162) + R×17"

Seventeen arginines were inserted into the R56-S162 domain of MafG, such that the number of a.a. residues between two adjacent arginines was six or less. The inserted arginines are indicated by "***R***".

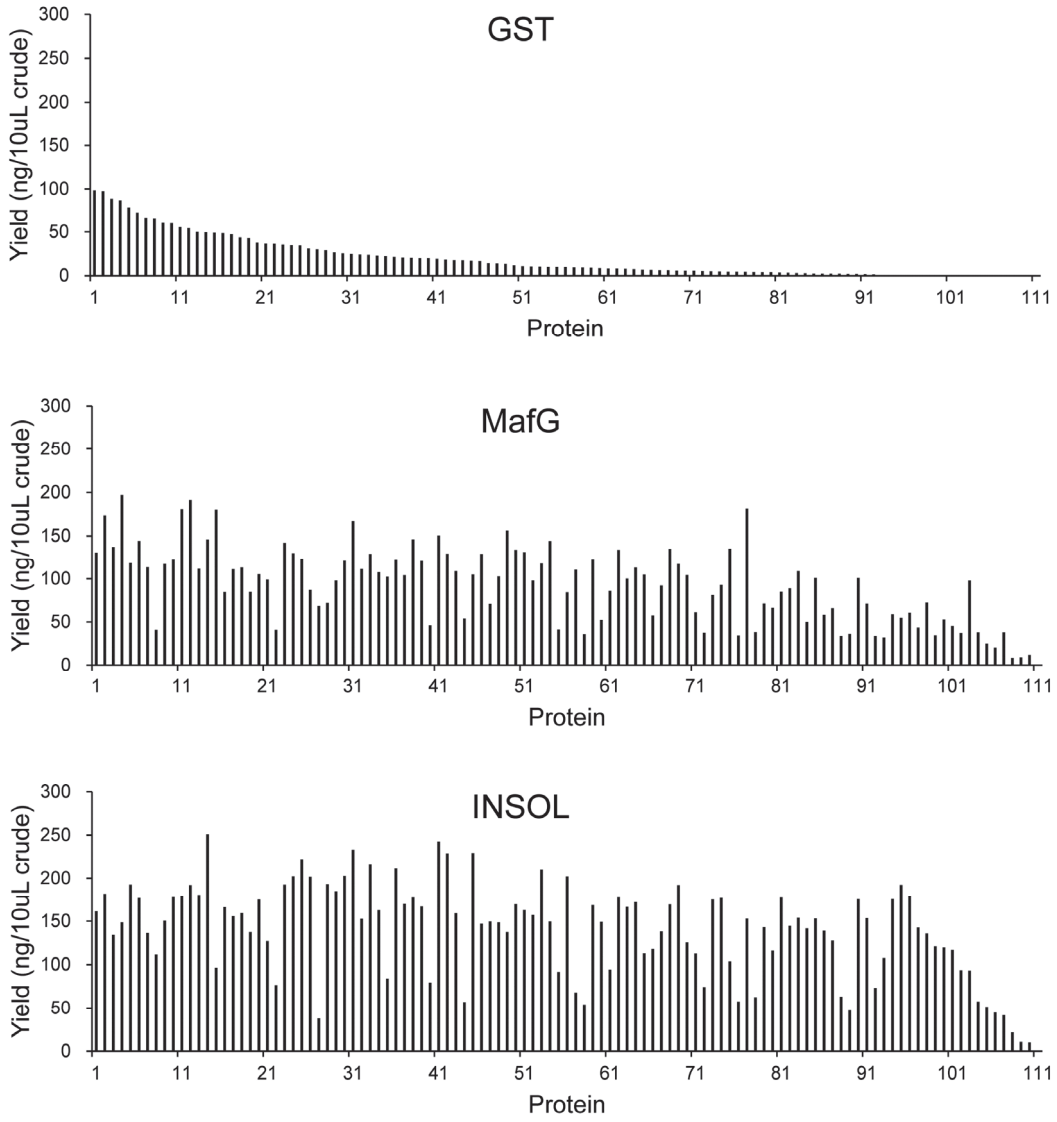


Figure 10 – Recovery amount following purification of 111 proteins with GST-tag, MafG-tag, and INSOL-tag

A total of 111 proteins were fused with the GST-tag, MafG-tag, and INSOL-tag, respectively. Data for GST-tag and MafG-tag are as shown in Figure 7a. The yield per 10 μL crude extract is shown by bar graphs. The proteins are arranged from the left in descending order of yield for GST-tag purification.

Observations of INSOL-tag with scanning electron-assisted dielectric microscopy (SE-ADM)

SE-ADM, unlike conventional electron microscopy techniques, allows observation of undried or stained samples in a liquid holder (Ogura, T. 2015, Senga, Y. *et al*. 2019). When the crude solution of wheat germ cell-free system expressing INSOL-tag was observed with SE-ADM, no particles with a diameter of ≥10 nm were observed (Figure 11a). Next, INSOL-tag recovered in pellet form by centrifugation was vigorously pipetted and suspended in 100 mM phosphate buffer (pH 7.8) and observed again. The INSOL-tag disaggregated into large and small chunks, the smallest unit of which ranged in diameter from 59 to 214 nm (average: 106 nm, SD: 36 nm). Its shape was smooth and spherical rather than chain-like or spinous (Figure 11b).

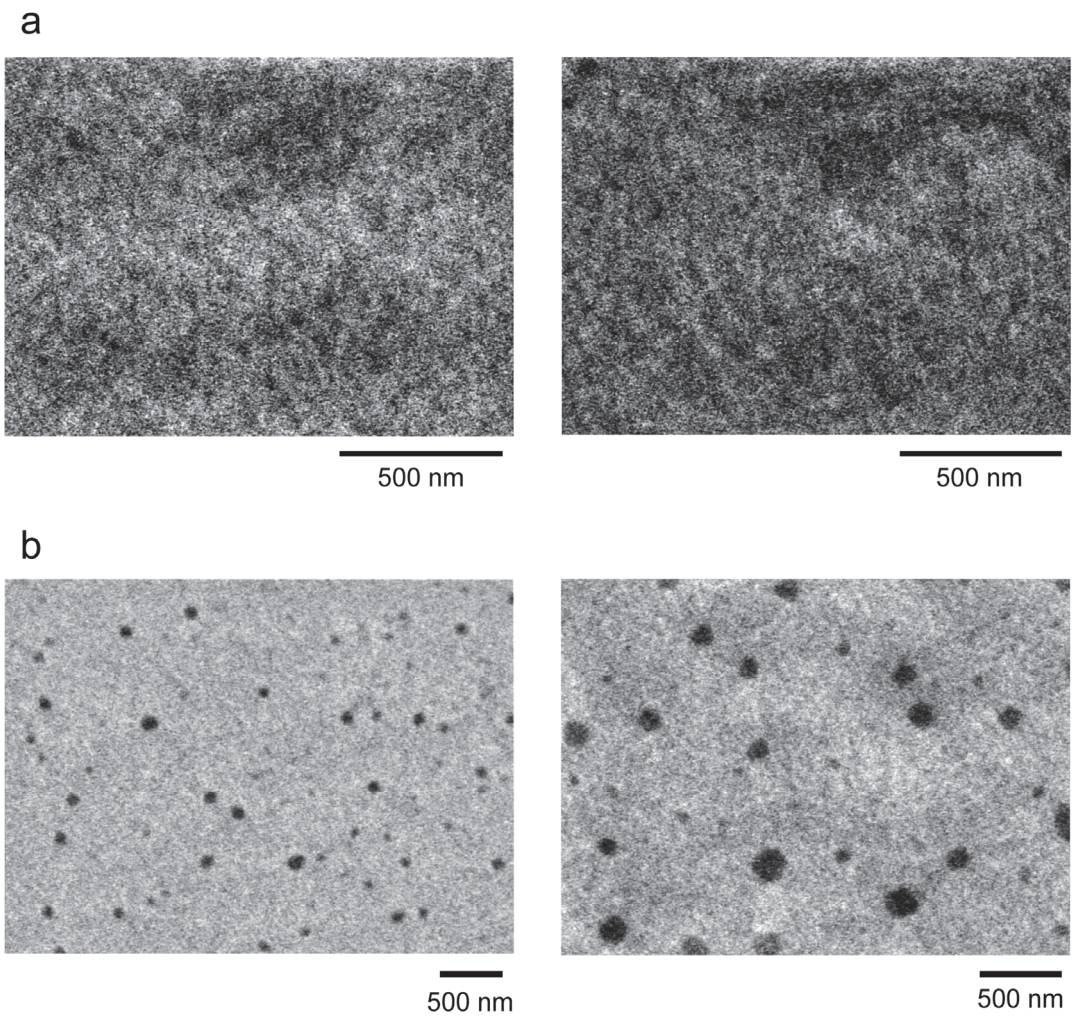


Figure 11 – Scanning electron-assisted dielectric microscopy (SE-ADM) images of INSOL-tag before and after centrifugation

(a) Crude solution of the wheat germ cell-free system expressing INSOL-tag. (b) The INSOL-tag was recovered in pellet form by centrifugation and vigorously pipetted and suspended in 100 mM phosphate buffer (pH 7.8). Scale bars: 500 nm. The left and right panels in (a) and (b) are representative images obtained under the same conditions.

Solubilization of INSOL-tagged proteins for use in the INSOL-tagged protein array

Next, the conditions under which the purified INSOL-tagged target proteins can be solubilized and bound to the aC plate were examined. An aC plate (Nippon Light Metal) with an active ester group on the surface was used for array preparation. The INSOL-tag was fused to the N-terminal side of IgG (IGHG3: FLJ40036AAAF) and expressed in a wheat germ cell-free system, and centrifugal purification was performed. The pellet was solubilized and optimization of the array conditions was performed.

Focusing on the ability to comprehensively solubilize any protein, the SDS was used as a surfactant for solubilization. When target proteins are spotted on the aC plate, the efficiency with which the active ester group on the plate and the amino group of the protein form an amide bond, was highly dependent on the pH. For the solubilization solution I therefore used 100 mM phosphate buffer (pH 6.8, 7.3, 7.8) or 100 mM bicine buffer (pH 8.3) containing 1% SDS. Under all conditions, the amount of solubilization was equivalent to that of sample buffer under normal SDS-PAGE (Table 1, Solution No. 1-5). Each lysate was spotted on an aC plate to examine the amount of protein that bound to the plate. As a result, the amount of target protein was highest at pH 7.8 (Table 1, Solution No. 3).

Although 1% SDS sufficiently solubilized the target protein, the net intensity varied when fluorescence detection was performed with a detection antibody after the protein solution was spotted on the aC plate. When protein solution containing 1% SDS was spotted onto the plate, a large amount of SDS was precipitated to form a mass. This may have been due to the target protein peeling off as well as the mass of SDS during the antigen-antibody reaction. I therefore examined by how much the SDS concentration in the solubilization solution could be reduced. To this end, solubilization amount of INSOL-IGHG3 was compared using 100 mM phosphate buffer (pH 7.8) containing 1, 0.2, 0.04, 0.008, and 0.0016% (w/v) SDS. When the concentration of SDS was 0.04% or more, the solubilized amount was almost equivalent to that of 1% SDS (Table 1, Solution No. 3-1 and 3-2). Subsequently,

each solubilized solution was bound to the aC plate, and the amount of resulting protein was examined. The amount of protein binding was almost equal at 0.04% and 1% (Table 1, Solution No. 3 and 3-2). Furthermore, the variation, which was large with a CV value of 70.8% for 1% SDS, greatly improved to 21.9% for 0.04% SDS. From the above, I decided to use a protein solution solubilized with 0.04% SDS in 100 mM phosphate buffer (pH 7.8) for preparation of the INSOL-tagged protein array.

Solution No.	Composition	IgG solubilization rate(%)	Intensity rate (%)*
1	1%SDS in 100mM phosphate buffer pH6.8	94.1	40.7
2	1%SDS in 100mM phosphate buffer pH7.3	96.6	74.0
3	1%SDS in 100mM phosphate buffer pH7.8	100.0	100.0
4	1%SDS in 100mM bicine pH8.3	99.0	57.6
5	NuPAGE LDS Sample Buffer(1x): 0.5% LDS	96.1	-
3-1	0.2%SDS in 100mM phosphate buffer pH7.8	97.7	60.2
3-2	0.04%SDS in 100mM phosphate buffer pH7.8	94.5	95.1
3-3	0.008%SDS in 100mM phosphate buffer pH7.8	14.8	9.3
3-4	0.0016%SDS in 100mM phosphate buffer pH7.	13.0	8.4

Table 1 – The solubilization conditions for the INSOL-tagged protein

*: As for details, refer to the section entitled "Evaluation of the INSOL-tag solubilization solution used to determine the amount of protein binding to the protein array" in the Materials and Methods.

Quantification and the detection limit of the INSOL-tagged protein array

Purified human IgG (02-7102; Thermo Fisher Scientific) was spotted at 2.0, 3.9, 7.8, 15.6, 31.3, 62.5, 125, 250, 500 pg/spot to quantify and determine the detection limit of the INSOL-tagged protein array prepared with the aC plate. Purified human IgG bound to the aC plate was detected using Alexa Fluor® 647 Goat Anti-Human IgG (H+L) (A-21445, Thermo Fisher Scientific) (Figure 12a).

As a background for signal detection (BG), the signal value of the spot where only a diluted solution of IgG (0.04% SDS in phosphate buffer [pH 7.8]) was spotted was subtracted. A high correlation was observed between the signal value and the amount of IgG (pg) (r = 0.997, R^2 = 0.994, y = 72.627x - 612.720). The SD value of BG was 97.9; however, when the 2SD of BG was set as the detection limit, $y = 97.9 \times 2$ and $x \approx 11.1$. From the above, the detection limit was determined as about 11.1 pg IgG.

Comparison of ELISA and the INSOL-tagged protein array

The antibody detection results of ELISA and the INSOL-tagged protein array were compared using human sera containing various concentrations of anti-MAGEA4 or anti-NY-ESO-1 autoantibodies. In the INSOL-tagged protein array, INSOL-MAGEA4 and INSOL-NY-ESO-1 were respectively bound to the aC plate with an amide bond as shown in Figure 12b. Serum diluted 100 fold was added and antibody binding was detected using Alexa Fluor® 647 Goat Anti-Human IgG (H+L) (1:1,000 dilution, A-21445; Thermo Fisher Scientific). With the anti-MAGEA4 and anti-NY-ESO-1 autoantibodies, the values of ELISA and the INSOL-tagged protein array were highly correlated at R^2 = 0.9127 and 0.9855, respectively (Figure 12c).

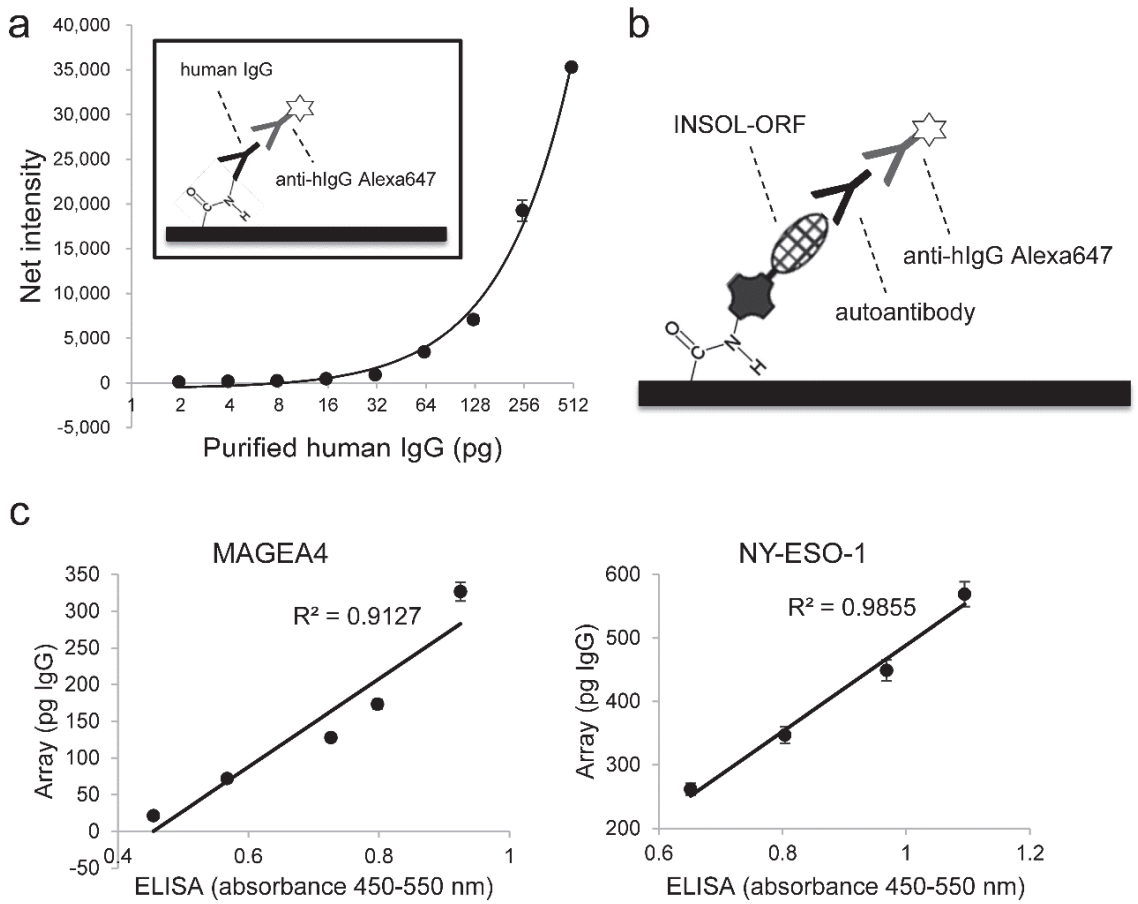


Figure 12 – Limit of detection and dynamic range of the aC protein array

(a) Analyses of the fluorescent intensity and IgG quantity bound to the aC plate (a semi-logarithmic graph which uses a logarithmic scale on the X axis). A series of diluted purified human IgG samples were spotted (n = 16) onto aC plates at 10 nl/spot, followed by the addition of Alexa-647-conjugated anti-human IgG antibody. The regression curve was calculated as: $y = ax + b$, where a = 72.627, b = 612.720, and R^2 = 0.994. anti-hIgG, anti-human IgG antibody. (b) Detection of serum autoantibodies using the INSOL-tagged protein array. The active-esterized surface of the aC plate reacted with amino groups of INSOL-tagged proteins (ORF) to form an amide bond. The INSOL-tagged protein array was incubated with human serum containing autoantibodies, and autoantibody-ORF binding was detected using Alexa-647-conjugated anti-human IgG antibody. (c) Correlation

between INSOL-tagged protein array analyses and ELISA values for serum autoantibody detection. The pg IgG values were calculated from (a).

Table 3 – List of 60 cell-cycle related proteins and 51 signaling-related proteins

Gene symbol	RefSeq ID	Entry clone ID
CCNB2	NP_004692.1	FLJ10542AAAN
CCNL2	NP_112199.2	FLJ14864AAAN
CCND3	NP_001129489.1	FLJ32644AAAN
CCNJ	NP_001127848.1	FLJ35041AAAN
CUL1	NP_003583.2	FLJ38844AAAN
CDC25C	NP_001274512.1	FLJ40391AAAN
RAD21	NP_006256.1	FLJ25655AAAN
WEE1	NP_001137448.1	FLJ16446AAAN
CDK7	NP_001790.1	FLJ22856AAAN
SMC2	NP_001252531.1	FLJ10623AAAN
CCNA1	NP_001104517.1	FLJ50745AAAN
AURKA	NP_940839.1	FLJ80023AAAN
CDK2	NP_001789.2	FLJ80039AAAN
CDK5	NP_004926.1	FLJ80048AAAN
CDKN1A	NP_001207707.1	FLJ80020AAAN
CDH1	NP_004351.1	FLJ92924AAAN
TP53	NP_001119584.1	FLJ92943AAAN
CCND1	NP_444284.1	FLJ93625AAAN
AURKB	NP_004208.2	FLJ80021AAAN
PLK1	NP_005021.2	FLJ80030AAAN
CDK4	NP_000066.1	FLJ80043AAAN
CHEK2	NP_009125.1	FLJ80045AAAN
WEE1	NP_003381.1	FLJ80185AAAN
PCNA	NP_872590.1	FLJ93798AAAN
CDKN1B	NP_004055.1	FLJ92816AAAN
MCM2	NP_004517.2	FLJ01025AAAN
CCNB1	NP_114172.1	FLJ81791AAAN
CCNL1	NP_064703.1	FLJ12912AAAF
CDC25B	NP_001274445.1	FLJ35394AAAF
SMC4	NP_001275682.1	FLJ46264AAAF
CDC2	NP_001777.1	FLJ80084AAAF
CCNI	NP_006826.1	FLJ93016AAAF

CDC7	NP_001127892.1	FLJ93132AAAF
CDC20	NP_001246.2	FLJ93197AAAF
CCNG1	NP_954854.1	FLJ93362AAAF
CKS1B	NP_001817.1	FLJ92030AAAF
CDK6	NP_001138778.1	FLJ94044AAAF
CDC6	NP_001245.1	FLJ94190AAAF
STAG1	NP_005853.2	FLJ51461AAAF
TP53	NP_001263690.1	FLJ50936AAAF
CDC25B	NP_001274445.1	FLJ53333AAAF
SKP2	NP_005974.2	FLJ53740AAAF
CCNC	NP_005181.2	FLJ52012AAAF
CDKN2D	NP_524145.1	FLJ80961AAAF
CHEK1	NP_001107594.1	FLJ80044AAAF
CDT1	NP_112190.2	FLJ80937AAAF
RB1	NP_000312.2	FLJ52431AAAF
CCND3	NP_001751.1	FLJ96475AAAF
MDM2	NP_002383.2	FLJ81856WAAF
CDC25A	NP_001780.2	FLJ81806AAAF
CCNH	NP_001230.1	FLJ81758AAAF
E2F1	NP_005216.1	FLJ82302AAAF
CDKN2A	NP_000068.1	FLJ83039WAAF
CCNF	NP_001752.2	FLJ93902AAAF
CCNA1	NP_003905.1	FLJ95009AAAF
STAG2	NP_001269347.1	FLJ94121AAAF
CCNK	NP_001092872.1	FLJ82599AAAF
CCNJ	NP_001127847.1	FLJ58102AAAF
CCNE1	NP_001229.1	FLJ82186AAAF
CCNL2	NP_001034666.1	FLJ83176AAAF
STAT1	NP_009330.1	FLJ39367AAAN
ATF4	NP_877962.1	FLJ25022AAAN
PTK2	NP_005598.3	FLJ37680AAAN
PTPN7	NP_542155.1	FLJ45281AAAN
GBL	NP_001186103.1	FLJ25896AAAN
RHEB	NP_005605.1	FLJ43457AAAN
TSC1	NP_001155899.1	FLJ50429AAAN

PDK1	NP_002601.1	FLJ80165AAAN
MAPK14	NP_620581.1	FLJ80122AAAN
PRKCB	NP_002729.2	FLJ80152AAAN
MAP2K2	NP_109587.1	FLJ80101AAAN
PRKCG	NP_002730.1	FLJ92607AAAN
RAP1A	NP_001278825.1	FLJ08033AAAN
PRKCA	NP_002728.1	FLJ08071AAAN
SHC1	NP_001123512.1	FLJ08067AAAN
RAF1	NP_002871.1	FLJ92543AAAN
MAP2K4	NP_003001.1	FLJ93529AAAN
ATF2	NP_001243019.1	FLJ46899AAAN
AKT1	NP_001014432.1	FLJ08041AAAN
PRKCH	NP_006246.2	FLJ08174AAAN
PRKCZ	NP_002735.3	FLJ80058AAAN
HRAS	NP_789765.1	FLJ82516AAAN
HRAS	NP_001123914.1	FLJ82516SAAN
AKT2	NP_001617.1	FLJ95460AAAN
ELK1	NP_001107595.1	FLJ93445AAAN
MAPK3	NP_002737.2	FLJ80081AAAN
MAP2K1	NP_002746.1	FLJ76051AAAN
JUN	NP_002219.1	FLJ82448WAAN
MYC	NP_002458.2	FLJ85585SAAN
SOS1	NP_005624.2	FLJ76778AAAN
PRKCD	NP_997704.1	FLJ93717AAAN
PRKCE	NP_005391.1	FLJ94469AAAN
EGFR	NP_005219.2	FLJ76780AAAN
KIAA1303	NP_065812.1	FLJ04039AAAN
PRKAB1	NP_006244.2	FLJ92856AAAF
MAPKSP1	NP_068805.1	FLJ92015AAAF
NCK1	NP_001278928.1	FLJ93089AAAF
RPS6KB1	NP_003152.1	FLJ93319AAAF
SRF	NP_003122.1	FLJ51683AAAF
EIF4EBP1	NP_004086.1	FLJ92286AAAF
PRKCZ	NP_001028754.1	FLJ53316AAAF
STAT5A	NP_001275648.1	FLJ54464AAAF

MKNK2	NP_951009.1	FLJ54773AAAF
GAB1	NP_002030.2	FLJ53999AAAF
GRB2	NP_002077.1	FLJ96637AAAF
CRK	NP_058431.2	FLJ81679AAAF
STAT5B	NP_036580.2	FLJ82377AAAF
MAPK1	NP_620407.1	FLJ58314AAAF
FOS	NP_005243.1	FLJ84847AAAF
STAT1	NP_009330.1	FLJ95929AAAF
PRKCI	NP_002731.4	FLJ08175AAAF

Discussion

INSOL-tag performance

In the purification of 111 cytoplasmic proteins, the mean yields were 19.9, 92.2, and 140.0 ng/10 μL for GST-tag, MafG, and INSOL-tag, respectively (Figure 10). The mean yield increased 1.5-fold when MafG was replaced with INSOL-tag. This was possibly because the a.a. length of the tag was shortened from 162 to 118 a.a., thus, reducing the material consumption for the tag itself during translation. It was shown that the INSOL-tag can insolubilize and recover nearly 120 types of target protein in multiple categories. Thus, when preparing antigens for antibody production using INSOL-tag, it is not necessary to change the purification method for each antigen. Also, when preparing internal standard proteins for mass spectrometry in the field of quantitative proteomics, it will be possible to prepare them in a uniform manner using INSOL-tag. After purification, INSOL-tagged protein must be solubilized with surfactants or denaturing agent such as SDS or guanidine hydrochloride. Therefore, the use of INSOL-tagged proteins are limited to experimental systems that are not sensitive to the structural changes in proteins.

Purity of the purified INSOL-tagged proteins

The purity of the purified INSOL-tagged proteins was less than 60% (Table 2), which seemed to be lower than the GST-tag purification system (Figure 7b). A part of the extra band may be chaperones contained in the wheat germ cell-free system bound to the target protein and collected in a precipitate by centrifugation. The other part may be the insolubilized wheat protein derived from wheat germ cell-free system while the target protein is translated over 18 hours. INSOL-tag can be solubilized with 0.04% (w/v) SDS (Table 1). The critical micelle concentration of SDS is 8×10^{-3} mol/L, i.e. about 0.23 % (w/v) in water at 25°C under atmospheric pressure (Domínguez, A. *et al*.

1997). This suggests that micelle formation of SDS is not required for solubilization of INSOL-tag. To improve the purity of INSOL-tag, it may be effective to add more surfactant during centrifugation and/or wash the pellet more after centrifugation. However, care must be taken not to dissolve INSOL-tag itself.

Insolubilization of MafG in vitro

Understanding why MafG, which functions in a soluble state *in vivo*, is insolubilized *in vitro,* and the aggregation mechanism of INSOL-tag is also important. MafG is expressed *in vivo* at an appropriate amount together with other proteins; however, when expressed in large amounts *in vitro*, it is thought to bind easily to itself via leucine zippers due to its high density. Moreover, when MafG is truncated, in addition to region B, which contains the leucine zipper required for dimerization, region C, which contains the nuclear-matrix targeting signal (Motohashi, H. *et al*. 2011), is also required for insolubilization (Figure 8c, d). MafG and the INSOL-tag are, therefore, thought to form a large mass due to polymerization via leucine zippers and the presence of a nuclear-matrix targeting signal, which can bind to multiple proteins. As shown in the SE-ADM image (Figure 11), the INSOL-tag forms a smooth spherical particle of about 100 nm in diameter. Aggregates of proteins are thought to differ in shape depending on the manner of aggregation. For example, in addition to spherical aggregates, chain- (e.g., amyloid beta) and thorn- (e.g., IgG) like aggregates are observed. The spherical shape of the INSOL-tag particles is thought to be due to polymerization of the INSOL-tag at multiple sites, the aggregates produced in a manner similar to snowball formation.

Conclusions and future prospects

In Chapter 1, a versatile INSOL-tag that can be purified easily via single centrifugation was

developed. The INSOL-tag insolubilizes any target protein in a cell-free eukaryote expression system. It can be used for the comprehensive protein expression, and for preparing proteins that are toxic to the host *in vivo* or difficult to obtain in a soluble state. INSOL-tag therefore has a wide range of applications, including the preparation of proteins for protein arrays, antigens for antibody production, and internal standards for mass spectrometry. It was confirmed that some kinases maintain activity even in their INSOL-tagged state, thus, while tag moiety is aggregated and insolubilized, the structure of the target protein moiety may be maintained. For application of the INSOL-tag technology to functional analysis of proteins, further analysis of how much INSOL-tagged target proteins retain their original enzyme and/or binding activity is required.

In Chapter 1, I also developed a method to fabricate a protein array using INSOL-tagged proteins. The INSOL-tagged protein array succeeded in detecting autoantibodies in human sera, and the detection results were highly correlated with that of ELISA. In the field of tumor immunology, serum autoantibodies against cancer antigens are measured for early diagnosis and evaluation of therapeutic effects of cancer. Among cancer antigens, cancer testis antigens (CTA), which are a large group, are known to be easily insolubilized when expressed as recombinant proteins. Therefore, in the study described in Chapter 2, I developed a protein array spotted with 267 types of INSOL-tagged cancer antigens and measured anti-cancer antigen antibodies in the serum of cancer patients.

Chapter 2

The analysis of autoantibody response to cancer vaccine via the protein array of 267 cancer antigens

Summary

I developed the protein array with INSOL-tagged cancer antigens as an example of the use of INSOL-tag developed in Chapter 1. Among cancer treatment methods, cancer vaccine therapy, in which a particular cancer antigen protein is administered to a patient, is characterized by a few side effects. To date, to evaluate the treatment of cancer vaccine therapy at an early stage of treatment, the method for measuring “antigen spreading” (antibodies made against cancer antigens leaked from dead cancer cells) has been tried. Antigen spreading is associated with tumor shrinkage and metastatic regression in some patients. On the other hand, although the types of antibodies produced after vaccine administration differ from patient to patient, tools for the comprehensive measurement of antibodies against several hundreds of cancer antigens were lacking. In this study, I prepared 267 cancer antigens using INSOL-tag and fabricated the cancer antigen array (INSOL CTA-array). I have succeeded in showing the binding of all 267 antigens to the array plate in an 87 times narrower range than that in the previous method. When autoantibody analysis with INSOL CTA-array was performed in cancer vaccine treated patients, an increase in autoantibodies against the vaccine antigen was observed. Further, eleven autoantibodies against cancer antigens other than vaccine antigens (antigen spreading) were detected. From the above, it was shown that 267 kinds of cancer antigens were successfully purified by INSOL-tag, and INSOL CTA-array was suggested to be useful for cancer vaccine therapy evaluation.

Introduction

Cancer serology has entered a new era in 1981 when Old *et al*. demonstrated the existence of specific autoantibody responses in the serum of cancer patients (Old, L.J. 1981). Initially, a method was used to classify patients by reacting serum on cancer cell lines from cancer patients. So far, hundreds of cancer antigens expressed in cancer cells have been identified, and the autoantibodies against cancer antigens have been investigated for relevance to cancer development and progression. Since the types of cancer antigens vary among cancer types and individuals, a new method called SEREX (serological identification of antigens by recombinant expression cloning) has been developed to comprehensively search for more cancer antigens (Old, L.J. *et al*. 1998). Subsequently, many researchers have also performed a method in which cancer cells are separated by two-dimensional PAGE, followed by Western blotting with serum from cancer patients, and detected spots are measured by mass spectrometry to identify cancer antigens (Brichory, F. *et al*. 2001, Le, Naour, F. *et al*. 2001, Naour, F. *et al*. 2002, Canelle, L. *et al*. 2005, Xia, Q. *et al*. 2005). Several hundreds of cancer antigens were discovered, and then the focus of researchers' interest shifted from the search for new cancer antigens to the analysis of autoantibodies against known cancer antigens. Therefore, a mini-antigen array in which about 10 known cancer antigens are spotted is prepared, and trials for detecting the anti-cancer antigen autoantibodies and applying them to clinical diagnosis have been carried out (Koziol, J.A. *et al*. 2003, Zhang, J.Y. *et al*. 2003, Zhang, J.Y. 2004). However, the positive rate of each anti-cancer antigen autoantibody was found to be as low as 10.8 to 24.6% (Zhang, J.Y. *et al*. 2003), which could be a weak point in diagnosing cancer. This weak point is expected to be overcome by combining the results of autoantibodies against multiple antigens, and further studies have to be carried out. To increase the positive rate of anti-cancer antigen autoantibodies in cancer patients, a protein array that comprehensively contains cancer antigens has been required. However, since cloning and

protein preparation of individual cancer antigens is time-consuming and requires technology, many studies have used arrays of several to tens of cancer antigens. The commercially available Human ProtoArray has a maximum of 101 cancer antigens. The laboratory I belong to has a cDNA library called HuPEX that covers around 80% of the human proteome, and the technology to express proteins with high throughput was established (Goshima, N. *et al*. 2008). In this study, for the analysis of more anti-cancer antigen autoantibodies, I developed a protein array containing 267 cancer antigens (INSOL CTA-array), by using the above-mentioned resources and technologies together with INSOL-tag technology.

Materials and Methods

INSOL-array assay

Unreacted residues of active ester on the aC plates were blocked with TBST at room temperature (RT) for 15 min with shaking and treated with ImmunoBlock (Sumitomo Dainippon Pharma, Osaka, Japan) at RT for 1 h. To detect immobilized INSOL-tagged IGHG3 or purified human IgG (Thermo Fisher Scientific), the aC plates were incubated for 1 h at RT with Goat anti-Human IgG (H+L) Secondary Antibody Alexa Fluor 647 (1:1,000 dilution; Thermo Fisher Scientific). Anti-MAFG antibody (1:250 dilution, ab154318; Abcam, Cambridge, UK) followed by Goat anti-rabbit IgG H&L Alexa Fluor 647 (1:1,000 dilution, ab150079; Abcam) were used to detect INSOL-tag. To detect autoantibodies in human sera, the aC plates were incubated for 1 h at RT with the sera (1:100 dilution). Goat anti-Human IgG (H+L) Secondary Antibody Alexa Fluor 647 (1:1,000 dilution; Thermo Fisher Scientific) was used as the second antibody. All antibodies and sera were diluted in TBST containing 5% (v/v) ImmunoBlock.

Production of the INSOL-CTA array

Recombinant human proteins (267 cancer antigens) were synthesized *in vitro* using INSOL-tag, with centrifugal purification. Each protein was solubilized with 100 mM phosphate buffer containing 0.04% SDS (pH 7.8) and spotted on active-esterified aC plates in duplicate using HORNET-NX (Wako Pure Chemical Industries). All proteins used are listed in Table 4 (refer to P. 65-72).

Array Analysis

The aC plates were scanned at a 10-μm resolution using a GenePix® Professional 4200A microarray scanner (Axon Instruments, Foster City, CA), and scanned images were saved for analysis as 16-bit

tiff files using GenePix® Pro ver.6.0 (Axon Instruments). Analysis software Array-Pro Analyzer ver.6.3.1 (Media Cybernetics, Rockville, MD) was used to record the mean by drawing an equally-sized circle around the spots. Net intensity was calculated by subtracting the background (local ring) from each sample intensity. To normalize differences between aC plates, the net intensity of each plate was corrected using the calibration curve of purified human IgG and calculated in units of pg IgG. In addition, Smirnov-Grubbs test ($p < .05$) was performed for each array plate, with antigens that were outliers for both spots regarded as positive. Smirnov-Grubbs test was performed using Ekuseru-Toukei 2012 (Social Survey Research Information, Japan).

Results

Fabrication of the INSOL-CTA array

To monitor autoantibodies in the sera of cancer patients, INSOL-CTA arrays were fabricated by spotting cancer antigens, mainly cancer testis antigens, on aC plates. Antigens registered in the CT database (Ludwig Institute for Cancer Research, http://www.cta.lncc.br/, Almeida, L.G. *et al*. 2009) and those described in Cheever, M.A. *et al*. 2009 were selected as candidates. Of these antigens, the cDNA resource was available for 267 (Table 4). In addition, dilution series (3.1, 6.3, 12.5, 25.0, 50.0 ng/μL) of purified human IgG and 100 ng/μL BSA as respective positive and negative controls, and INSOL-tag protein solution as a positive control of the tag were also spotted. The 267 antigens were spotted in duplicate, and the controls were spotted in quadruplicate.

Quantitative determination of bound proteins

Because INSOL-tag is derived from MafG, anti-MAFG antibody (ab154318; Abcam) was used to detect proteins bound to the aC plates (Figure 13a, b). The net intensity average of each of the 267 antigens was calculated. The lowest net intensity average was 317 and the highest was 27,626. This low value was significantly higher than the mean value of BSA ($p = 0.013$). When the reproducibility of the protein binding of two spots was examined for all 267 antigens, a high correlation was observed between the first and second spot (R^2 = 0.9464) (Figure 14). In addition, when the reproducibility of protein binding was examined between two different plates, a high correlation was also found between plate 1 and plate 2 (R^2 = 0.9593) (Figure 14).

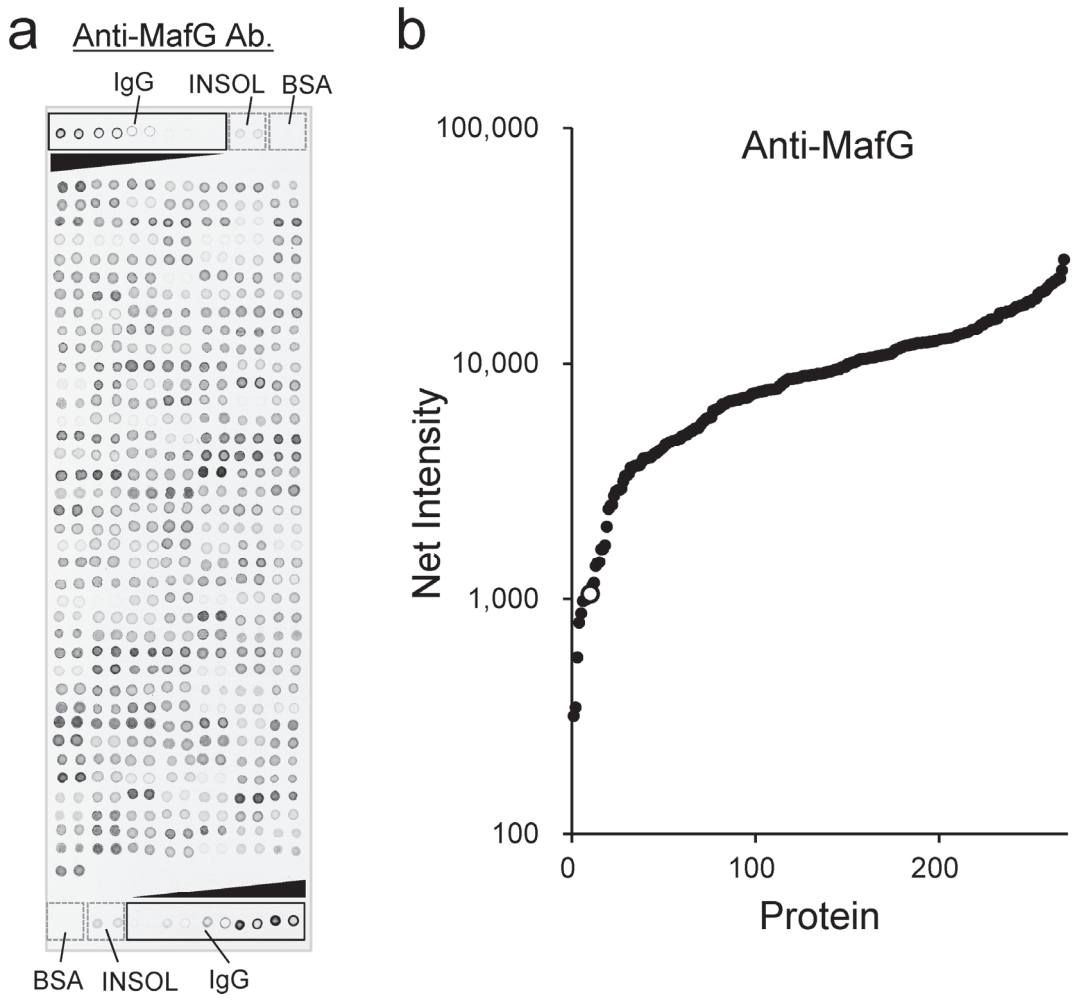


Figure 13 – INSOL-CTA array

(a) INSOL-CTA array of 267 CTA proteins and detection of INSOL-tagged proteins spotted on an aC plate using anti-MafG antibody. BSA, INSOL, and IgG represent BSA (100 ng/μl) as a background, the INSOLtag without an ORF, and purified human IgG (3.1–50 ng/μl) as a positive control, respectively. The 267 antigens were spotted in duplicate and the controls were spotted in quadruplicate on a single aC plate. (b) The amount of INSOL-tagged proteins bound to the aC plate shown as a log plot. An open circle represents the net intensity of the INSOL-tag without an ORF. Black circles represent the net intensities of the INSOLtagged ORFs. The minimum and maximum intensities were 316.8 and 27,625.8, respectively.

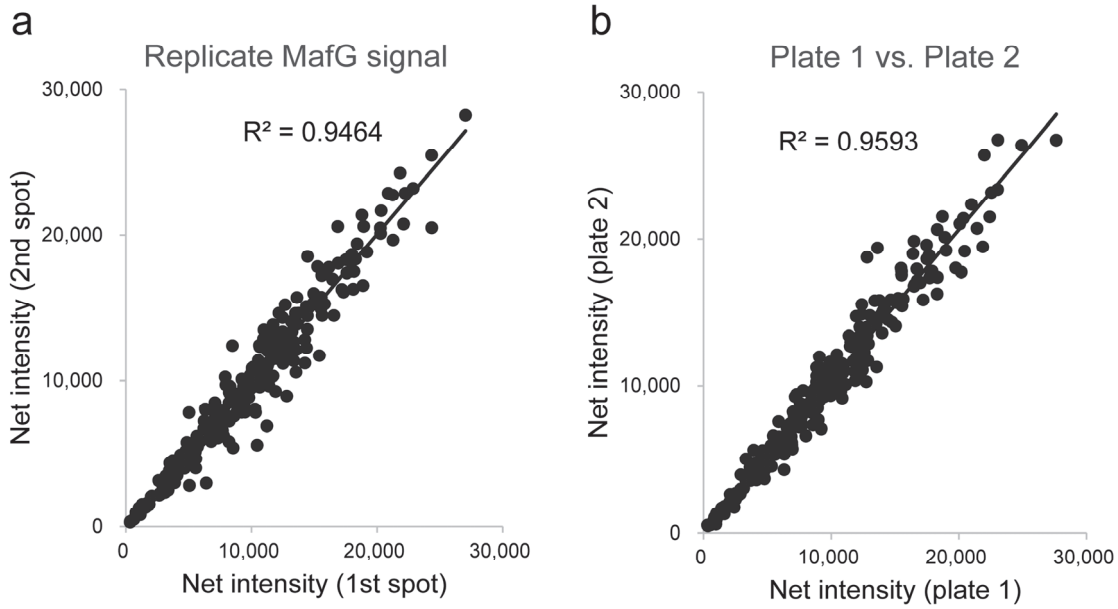


Figure 14 – Evaluation of the reproducibility of protein binding

(a) Proteins were spotted using a single INSOL-CTA array and detected with anti-MafG antibody. The correlation between the first spot (*X* axis) and second spot (*Y* axis) on a single plate is shown for 267 antigens. (b) Proteins were spotted using two different INSOL-CTA arrays and detected with anti-MafG antibody. The average net intensity of duplicate spots on the same aC plate was calculated. The correlation between the average net intensity of the two different INSOL-CTA arrays is shown in the graph. Net intensity was calculated by subtracting the background (local ring) from the raw intensity.

Serum autoantibody monitoring in patients treated with cancer vaccine

As a negative control, pooled sera from 10 healthy humans (12181201, Lot: BJ3529D; Tennessee Blood Services, Memphis, TN) were subjected to the INSOL-CTA array to detect bound autoantibodies. No significant difference in the signal was detected compared with the BSA spot (Figure 15a); however, numerous autoantibodies were detected in the cancer patient sera as shown in Figure 15b. No signals were detected with the secondary antibody alone (Figure 15c).

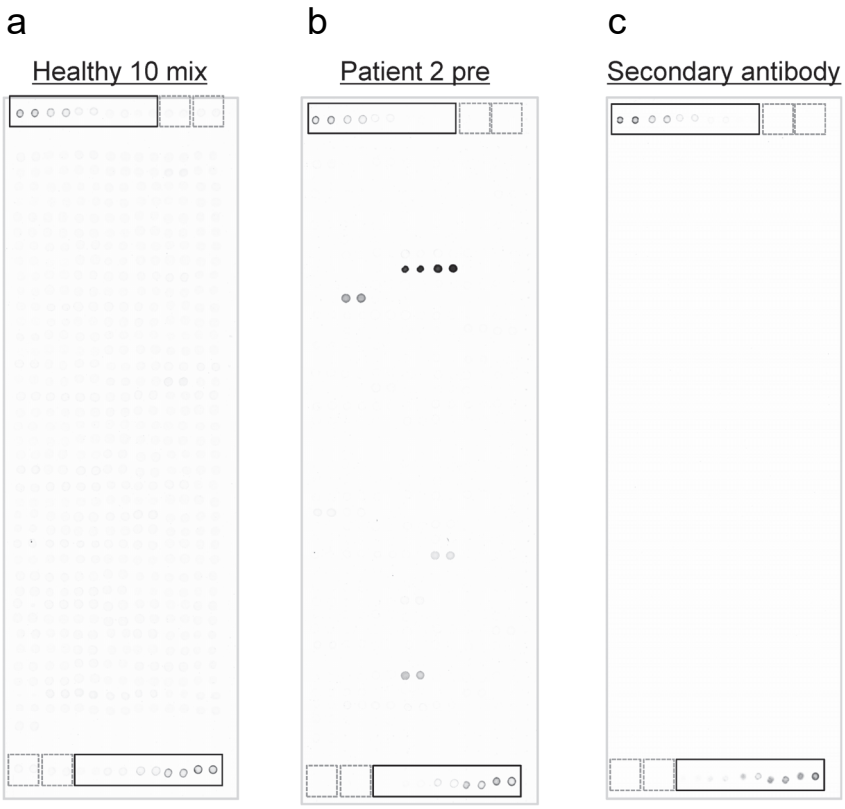


Figure 15 – Representative examples of INSOL CTA-array measurement

(a) Results of the INSOL-CTA array using a serum pool from 10 healthy donors. (b) Representative result of serum autoantibody from a cancer patient. The INSOL-CTA array was carried out using serum from Patient 2, pre-vaccination. Bound IgGs were detected using a secondary antibody (Alexa-647-conjugated anti-human IgG antibody). (c) Results of the INSOL-CTA array using secondary antibody only.

Autoantibody monitoring of serum from the digestive system cancer patients was also performed using the INSOL-CTA array. Patient 1 was a 63-year-old man with colon cancer and Patient 2 was a 69-year-old man with duodenal papillary carcinoma. Sera taken before and after 2, 4, 6, 10 and 13 administrations of MAGEA4 vaccine (10 and 13 only for Patient 2) were used. The signal at each time point was normalized with the standard curve of purified human IgG and calculated in units of pg IgG. For each plate, the Smirnov-Grubbs test ($p < .05$) was performed, and if two spots of an antigen both became outliers, the antigen was regarded as positive. In Patient 1, many positive antigens of the MAGE family were detected, whereas in Patient 2, only a few were detected along with a number of other positive antigens. Five antigens in Patient 1 and 13 in Patient 2 were negative before the vaccination but became positive after administration (Table 5). The results before administration and after 6 and 13 administrations of the vaccine (13 only for Patient 2) are shown in the radar charts in Figure 16. After vaccine administration, antigen spreading was observed; that is, the production of antibodies against antigens whose sequences were not identical to the antibody against the vaccine.

			Intensity (pg IgG)									
			Patient 1				Patient 2					
			No. of vaccination				No. of vaccination					
	Clone ID	Gene Symbol	pre.	2	4	6	pre.	2	4	6	10	13
MAGE family	**FLJ56243AAAF**	**MAGEA4**	n.d.	203	423	568	n.d.	n.d.	n.d.	113	77	n.d.
	FLJ94494AAAF	**MAGEA4**	72	304	737	854	n.d.	n.d.	72	180	123	121
	FLJ95895AAAF	MAGEA1	n.d.	n.d.	95	114	n.d.	n.d.	n.d.	67	n.d.	n.d.
	FLJ76017AAAF	MAGEA2, MAGEA2B	79	96	71	79	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ81624AAAF	MAGEA3	277	273	186	208	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ09168AAAN	MAGEA5	n.d.	94	158	199	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ85809AAAN	MAGEA6	291	268	195	178	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ94585AAAF	MAGEA8	n.d.	n.d.	n.d.	77	n.d.	n.d.	n.d.	n.d.	293	278
	FLJ94880AAAF	MAGEA12	102	148	169	198	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ40242AAAN	MAGEB6	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	168
	FLJ80612AAAF	ACPP	n.d.	n.d.	n.d.	80	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ04202AAAN	ATAD2	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	47	n.d.	n.d.	140
	FLJ04196AAAN	CTAG1A, CTAG1B	n.d.	n.d.	n.d.	n.d.	1224	1385	1419	1073	1286	1555
	FLJ80977AAAF	CTAG2	n.d.	n.d.	n.d.	n.d.	1591	1903	2029	1638	1836	2420
	FLJ81464AAAF	DDX53	n.d.	n.d.	n.d.	n.d.	656	1014	1035	998	1327	2026
	FLJ80078AAAN	LCK	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	118
	FLJ84647AAAN	MYCN	73	109	n.d.	81	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ81394AAAF	PAGE2	n.d.	n.d.	n.d.	n.d.	96	71	n.d.	121	n.d.	243
	FLJ81394SAAN	PAGE2B	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	64	n.d.	120
	100014753_F	PTPN20A	62	98	n.d.	n.d.	194	115	102	88	118	402
	FLJ25396AAAN	RHOXF2	69	100	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	FLJ95609AAAF	RSPH1	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	150
	FLJ82008AAAF	SPAG8	n.d.	n.d.	n.d.	n.d.	n.d.	52	n.d.	n.d.	n.d.	n.d.
	FLJ81708AAAF	SSX2	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	49	67	n.d.	n.d.
	FLJ81708SAAN	SSX2B	n.d.	n.d.	n.d.	n.d.	n.d.	54	n.d.	92	n.d.	n.d.
	100070120_F	TFDP familymember	n.d.	n.d.	n.d.	n.d.	378	590	629	532	667	859
	FLJ80022AAAN	TTK	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	216	218

Table 5 – Serum autoantibody monitoring results obtained using the INSOL-CTA array for two digestive system cancer patients

Sera from pre- and post-vaccinated patients (2, 4, 6, 10, 13 times, respectively) were added to the INSOL-CTA array and bound IgGs were detected. Amounts of bound IgG (pg IgG) were estimated and antigens judged as positive were listed. Color shades represent the gradient of all amounts of bound IgG ranging from white (minimum) to red (maximum). n.d., not detected (not classified as an outlier, Smirnov-Grubbs test, $P < .05$).

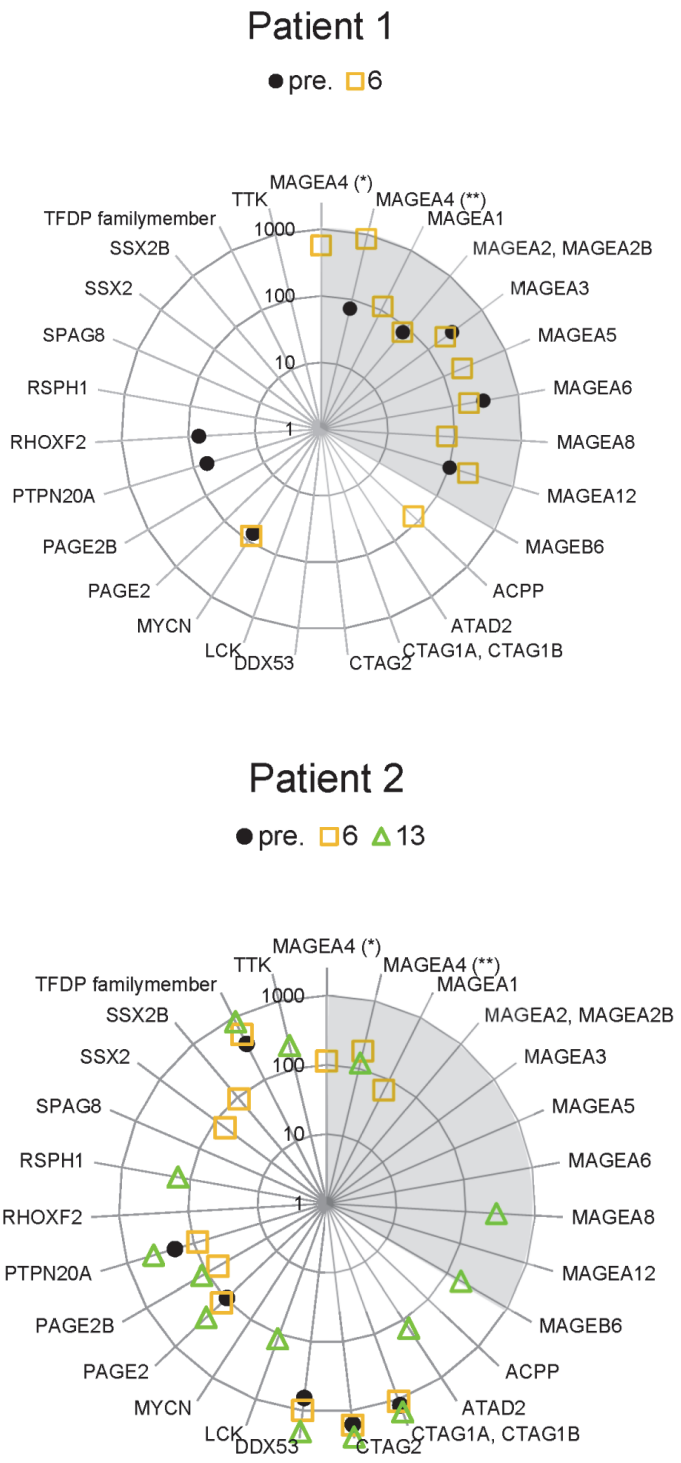


Figure 16 – Radar charts of autoantibody profiling results of vaccinated cancer patients using the INSOL-CTA array.

Detection of autoantibodies in cancer patient sera using the INSOL-CTA array in pg IgG (the amount of bound IgG). Of the 267 antigens included in the INSOL-CTA array, 27 antigens

that were positive in either patient 1 or 2 are described. Solid black circles, before vaccine administration; yellow squares, after six doses of the vaccine; and green triangles, after 13 doses of the vaccine. The unit of the intensity is pg IgG. MAGEA4 (*), FLJ56243AAAF; MAGEA4 (**), FLJ94494AAAF

Table 4 – List of proteins spotted onto the INSOL-CTA array

Gene symbol	RefSeq ID	Entry clone ID
ACPP	NP_001127666.1	FLJ80612AAAF
ACRBP	NP_115878.2	FLJ93199AAAF
ACTL8	NP_110439.2	FLJ32777AAAF
ADAM,ADAM29	NP_001265056.1	FLJ76062AAAF
ADAM2	NP_001265042.1	FLJ80744AAAF
AFP	NP_001125.1	FLJ95694AAAN
ALK	NP_004295.2	FLJ60605AAAN
ANKRD30BP2	XP_011545290.1	FLJ83852WAAF
ANKRD45	NP_940895.1	FLJ45235AAAF
AR	NP_000035.2	FLJ95787AAAN
ARMC3	NP_001269674.1	FLJ52094AAAF
ARNTL	NP_001284651.1	FLJ96151AAAF
ATAD2	NP_054828.2	FLJ04202AAAN
BAGE, BAGE5	NP_872290.1	FLJ85816SAAN
BAGE2, BAGE3	NP_872288.2	FLJ76422AAAF
BAGE4	NP_859055.1	100064189_N
BIRC5	NP_001159.2	FLJ52793AAAF
BIRC5	NP_001159.2	FLJ92184AAAF
BIRC7	NP_647478.1	FLJ81925AAAF
BRDT	NP_001229735.2	FLJ55453AAAF
C15orf60	NP_001035826.1	FLJ44083AAAN
CA9	NP_001207.2	FLJ80657AAAN
CABYR	NP_722452.1	FLJ93748AAAF
CAGE1	NP_001164164.1	FLJ40441AAAF
CALR3	NP_659483.2	FLJ25355AAAN
CASC5	NP_733468.3	100069116_N
CASC5	NP_653091.3	FLJ04197AAAF
CCDC	NP_775827.2	FLJ42119AAAN
CCDC110	NP_689988.1	FLJ52074AAAF
CCDC110	NP_689988.1	FLJ85140AAAF
CCDC33	NP_079331.3	FLJ32855AAAF
CCDC36	NP_001128669.1	FLJ85810WAAN
CCDC62	NP_958843.2	FLJ40344AAAN

CCDC62	NP_958843.2	FLJ85164AAAF
CCNA	NP_003905.1	FLJ95009AAAF
CD276	NP_001019907.1	FLJ78639AAAN
CD276	NP_001019907.1	FLJ85820AAAN
CEA	NP_001278413.1	100003470_N
CEP290	NP_079390.3	100069260_N
CEP290	NP_079390.3	FLJ13615AAAN
CEP55	NP_001120654.1	FLJ96602AAAF
CLOCK	NP_001254772.1	FLJ78672AAAN
COX6B2	NP_653214.2	FLJ85805AAAN
CPXCR1	NP_001171700.1	FLJ94370AAAF
CRISP2	NP_001248751.1	FLJ83529AAAF
CSAG1	NP_001096046.2	FLJ83127AAAF
CSAG2	NP_001123298.1	FLJ82519AAAF
CSPG4	NP_001888.2	FLJ07002AAAN
CT45A1, CT45A2, CT45A4, CT45A6	NP_001017436.1	FLJ25823AAAN
CT45A3	NP_001017435.1	100004214_N
CT45A5	NP_001278459.1	FLJ83136AAAF
CT47A1, CT47A2, CT47A3, CT47A4, CT47A5, CT47A6, CT47A7, CT47A8, CT47A9, CT47A10, CT47A11	NP_775842.2	FLJ81235AAAF
CT62	NP_001096128.1	FLJ85818WAAN
CTAG1A, CTAG1B	NP_001318.1	FLJ04196AAAN
CTAG2	NP_066274.2	FLJ80977AAAF
CTAGE1	NP_758441.2	FLJ82693WAAF
CTAGE5	NP_001234917.1	FLJ33933AAAN
CTCFL	NP_001255970.1	FLJ94642AAAF
CTNNA2	NP_004380.2	FLJ45293AAAN
CXorf48	NP_001026875.1	FLJ84880AAAF
CXorf61	NP_001017978.1	FLJ20611AAAN
CYP1B1	NP_000095.2	FLJ93081AAAN
DCAF12	NP_056212.1	FLJ84132AAAF
DCT/TRP2	NP_001913.2	FLJ75858AAAF
DDX43	NP_061135.2	FLJ94275AAAF
DDX53	NP_874358.2	FLJ81464AAAF

DMRT	NP_068770.2	FLJ94161AAAN
DNAJB8	NP_699161.1	FLJ44080AAAN
DPPA2	NP_620170.3	FLJ81154AAAF
DSCR8	None (Gene ID: 84677)	FLJ04206AAAN
EGFR (EGFRvIII)	NP_005219.2	FLJ76780SAAN
EGFR/HER1	NP_005219.2	FLJ85934SAAF
ELOVL4	NP_073563.1	FLJ30715AAAF
EPCAM	NP_002345.2	FLJ22932AAAN
EPHA2	NP_004422.2	100006424_N
EPHA2	NP_004422.2	FLJ51229AAAN
ERBB2	NP_004439.2	FLJ02812AAAN
ERG	NP_891548.1	FLJ52381AAAF
FAM133A	NP_001164582.1	FLJ37659AAAN
FAM46D	NP_001164045.1	FLJ46269AAAN
FAP	NP_004451.2	FLJ96506AAAF
FATE1	NP_149076.1	FLJ32807AAAN
FBXO	NP_694962.1	FLJ82180AAAF
FMR1NB	NP_689791.1	FLJ25736AAAN
FOLH1	NP_004467.1	FLJ92688AAAF
FOSL1	NP_005429.1	FLJ94314AAAF
FTHL17	NP_114100.1	FLJ09148AAAN
GAGE1	NP_001035753.1	FLJ82197AAAF
GAGE10	NP_001091883.3	100064143_N
GAGE12B	NP_001091888.1	100072367_F
GAGE2A, GAGE2B, GAGE2C, GAGE3, GAGE4, GAGE7, GAGE12I, GAGE8, GAGE2E, GAGE12J, GAGE13, GAGE12C, GAGE12D, GAGE12E, GAGE12F, GAGE12G, GAGE12H	NP_001091877.1	FLJ85817AAAN
GAGE5	NP_001466.1	100003790_N
GAGE6	NP_001467.1	100015388_F
GOLGA6L2	NP_001291317.1	FLJ36144AAAN
GPAT2	NP_997211.2	FLJ16024AAAN
GPATCH2	NP_060510.1	FLJ10252AAAF

HORMAD1	NP_115508.2	FLJ40071AAAN
HORMAD2	NP_689723.1	FLJ25837AAAN
HRAS	NP_001123914.1	FLJ82516SAAN
HSPB9	NP_149971.1	FLJ36169AAAN
IGF2BP3/KOC1	NP_006538.2	FLJ82379AAAF
IGSF11	NP_689751.2	FLJ40348AAAN
IL13RA1	NP_001551.1	FLJ94013AAAF
IL13RA2	NP_000631.1	FLJ80687AAAF
KDM5B	NP_006609.3	FLJ04198AAAN
KDR	NP_002244.1	FLJ09052AAAN
KIAA0100	NP_055495.2	FLJ85813SAAN
KIF2	NP_001284584.1	FLJ54088AAAF
KIF20B	NP_001271188.1	FLJ21975AAAN
KLK3	NP_001639.1	FLJ93384AAAF
KRAS	NP_004976.2	FLJ82576AAAF
LCK	NP_001036236.1	FLJ80078AAAN
LDHC	NP_059144.1	FLJ25452AAAN
LEMD1	NP_001185979.1	FLJ51398AAAF
LEMD1	NP_001001552.3	FLJ22975AAAN
LGMN	NP_001008530.1	FLJ84952WAAF
LIMS1	NP_001180413.1	FLJ94954AAAF
LIPI	NP_945347.2	FLJ82987SAAN
LOC440934	None (Gene ID: 440934)	FLJ04200AAAN
LUZP4	NP_057467.1	FLJ95634AAAF
LY6K	NP_001153826.1	FLJ35226AAAN
LYPD6B	NP_808879.2	FLJ85814SAAN
MAEL	NP_116247.1	FLJ14904AAAN
MAEL	NP_001273306.1	FLJ50290AAAN
MAGEA1	NP_004979.3	FLJ95895AAAF
MAGEA10	NP_001238757.1	FLJ94870AAAF
MAGEA11	NP_001011544.1	FLJ45952AAAN
MAGEA12	NP_001159859.1	FLJ94880AAAF
MAGEA2, MAGEA2B	NP_001269434.1	FLJ76017AAAF
MAGEA3	NP_005353.1	FLJ81624AAAF
MAGEA4	NP_001011550.1	FLJ56243AAAF

MAGEA4	NP_001011550.1	FLJ94494AAAF
MAGEA5	NP_066387.1	FLJ09168AAAN
MAGEA6	NP_787064.1	FLJ85809AAAN
MAGEA8	NP_001159873.1	FLJ94585AAAF
MAGEA9, MAGEA9B	NP_001074259.1	FLJ93838AAAF
MAGEB1	NP_803134.1	FLJ95893AAAF
MAGEB2	NP_002355.2	FLJ93871AAAF
MAGEB3	NP_002356.2	FLJ32879AAAN
MAGEB4	NP_002358.1	FLJ94371AAAF
MAGEB6	NP_775794.2	FLJ40242AAAN
MAGEC2	NP_057333.1	FLJ94157AAAF
MLANA/MART-1	NP_005502.1	FLJ92434AAAF
MORC1	NP_055244.3	FLJ85054AAAF
MSLN	NP_001170826.1	FLJ80624AAAF
MUC1	NP_001191214.1	FLJ60927AAAN
MYCN	NP_001280157.1	FLJ84647AAAN
NLRP4	NP_604393.2	FLJ95931AAAF
NOL4	NP_001185475.1	FLJ34144AAAN
NOL4	NP_001185476.1	FLJ54015AAAF
NR6A1	NP_001480.3	FLJ82687AAAF
NRAS	NP_002515.1	FLJ85658AAAN
NUF2	NP_663735.2	FLJ36029AAAF
NXF2, NXF2B	NP_001093156.1	100003770_N
ODF1	NP_077721.2	FLJ96288AAAF
ODF2	NP_001229282.1	FLJ55346AAAF
ODF3	NP_444510.2	100072437_F
OIP5	NP_009211.1	FLJ81121AAAF
OTOA	NP_733764.1	FLJ35743AAAN
PAGE1	NP_003776.2	FLJ80991AAAF
PAGE2	NP_997222.1	FLJ81394AAAF
PAGE2B	NP_997222.1	FLJ81394SAAN
PAGE3	NP_001290542.1	FLJ03012AAAN
PAGE4	NP_008934.1	FLJ92316AAAF
PAGE5	NP_569734.2	FLJ81047AAAF
PASD1	NP_775764.2	FLJ40233AAAN

PAX3	NP_852123.1	FLJ77133AAAN
PAX5	NP_057953.1	100063578_N
PBK	NP_060962.2	FLJ14385AAAN
PDGFRB	NP_002600.1	FLJ80123AAAF
PIWI	NP_004755.2	FLJ80907AAAN
PIWIL2	NP_001129193.1	FLJ80906AAAN
PLAC1	NP_068568.1	FLJ90605AAAN
PMEL/gp100	NP_008859.1	FLJ35562AAAN
POTEA	NP_001005365.2	100016365_N
POTEB	NP_997238.2	FLJ81454AAAF
POTEC	NP_001131143.1	FLJ97606SAAN
POTED	NP_778146.2	100066427_N
POTEH	NP_001129685.1	100015798_N
PRAME	NP_001278645.1	FLJ26272AAAN
PRM1	NP_002752.1	FLJ92260AAAF
PRM2	NP_002753.2	FLJ92059AAAF
PRSS50	NP_037402.1	FLJ35279AAAN
PRSS54	NP_001292102.1	FLJ25339AAAN
PSCA	NP_005663.2	FLJ80698AAAF
PSMB9	NP_002791.1	FLJ83284WAAF
PTPN20A	NP_056420.3	100014753_F
PTPN20A	NP_056420.3	FLJ02817AAAN
RBM46	NP_659416.1	FLJ84710AAAN
RGS2	NP_056483.3	FLJ85806AAAN
RGS5	NP_003608.1	FLJ96402AAAF
RHOC	NP_001036144.1	FLJ37155AAAN
RHOXF2	NP_115887.1	FLJ25396AAAN
RNF1	NP_112567.2	100066395_N
ROPN1	NP_060048.2	100072423_F
ROPN1	NP_060048.2	FLJ85812WAAN
ROPN1	NP_001012337.1	FLJ55261AAAF
RQCD1	NP_005435.1	FLJ93148AAAF
RSPH1	NP_543136.1	FLJ95609AAAF
SART3	NP_055521.1	FLJ76238AAAN
SEMG1	NP_002998.1	FLJ85808AAAN

SLCO6A1	NP_001275933.1	FLJ16580AAAF
SPA17	NP_059121.1	FLJ92067AAAF
SPACA3	NP_776246.1	FLJ57231SAAN
SPAG	NP_003107.1	40080652_F
SPAG17	NP_996879.1	100015375_F
SPAG6	NP_758442.1	FLJ83094AAAF
SPAG8	NP_758516.1	FLJ82008AAAF
SPAG9	NP_001124000.1	FLJ61493AAAN
SPANXA1, SPANXA2, SPANXC	NP_663695.1	FLJ81395AAAF
SPANXB1, SPANXB2	NP_115850.2	FLJ81300AAAF
SPANXD, SPANXE	NP_115793.1	FLJ81002AAAF
SPANXN1	NP_001009614.1	FLJ03008AAAN
SPANXN2	NP_001009615.1	100014971_N
SPANXN3	NP_001009609.2	FLJ85200AAAF
SPANXN4	NP_001009613.1	FLJ03009AAAN
SPANXN5	NP_001009616.1	FLJ03010AAAN
SPATA19	NP_001278921.1	FLJ25851AAAF
SPEF2	NP_653323.1	FLJ25395AAAN
SPEF2	NP_079143.3	FLJ23164AAAF
SPINLW1	NP_065131.1	FLJ92172AAAF
SPO11	NP_937998.1	FLJ82715AAAF
SSX1	NP_001265620.1	FLJ81661AAAF
SSX2	NP_001265630.1	FLJ81708AAAF
SSX2B	NP_001157889.1	FLJ81708SAAN
SSX3	NP_066294.1	FLJ82512AAAF
SSX4, SSX4B	NP_001030004.1	FLJ81000AAAF
SSX5	NP_066295.3	FLJ81139AAAF
SSX6	NP_001265620.1	FLJ04204AAAN
SSX7	NP_775494.1	100014471_F
SSX8	NP_001265620.1	FLJ56587AAAF
SYCE1	NP_001137235.1	FLJ95894AAAF
SYCP1	NP_001269470.1	FLJ04203AAAN
TAF7L	NP_079161.3	FLJ98209SAAN
TDRD1	NP_942090.1	FLJ16594AAAN
TDRD6	NP_001161831.1	FLJ04199AAAN

TEK	NP_000450.2	FLJ75731AAAF
TEKT	NP_653275.1	FLJ32871AAAN
TERT	NP_937983.2	FLJ01038AAAN
TEX101	NP_113639.4	FLJ80962AAAF
TEX14	NP_112562.3	FLJ04201AAAN
TFDP familymember	NP_057605.3	100070120_F
THEG	NP_057669.1	FLJ81224AAAF
TMEFF1	NP_003683.2	FLJ80534AAAF
TMEFF2	NP_057276.2	FLJ90151AAAN
TMEM108	NP_001129941.1	FLJ80402AAAN
TMPRSS12	NP_872365.1	FLJ81372AAAF
TP53	NP_001119584.1	FLJ92943AAAN
TPPP2	NP_776245.2	FLJ82222AAAF
TPTE	NP_954870.2	FLJ95041AAAF
TSGA10	NP_079520.1	FLJ93879AAAF
TSPY1	NP_001184171.1	FLJ36094SAAN
TSPY2, TSPY3	NP_072095.2	FLJ97373SAAN
TSSK6	NP_114426.1	FLJ80085AAAN
TTK	NP_003309.2	FLJ80022AAAN
TULP2	NP_003314.2	FLJ81202AAAF
TYR	NP_000363.1	FLJ82657AAAF
VENTXP1	None (Gene ID: 139538)	FLJ03007AAAN
WT1	NP_001185481.1	FLJ35849AAAN
WT1	NP_077742.2	FLJ77569SAAN
XAGE1A, XAGE1B, XAGE1C, XAGE1D, XAGE1E	NP_001091073.2	FLJ82544WAAF
XAGE2, XAGE2B	NP_570133.1	FLJ81841AAAF
XAGE3	NP_573440.1	FLJ81420AAAF
XAGE5	NP_570131.1	FLJ03011AAAN
YBX2	NP_057066.2	FLJ95978AAAF
ZNF165	NP_003438.1	FLJ82068AAAF
ZNF6	NP_689790.1	FLJ25735AAAN

Discussion

Quality of INSOL CTA-array

To evaluate the reproducibility of measurement with INSOL CTA-array, the correlation of net intensity was investigated between spots in the same array plate (1st spot and 2nd spot) and between different array plates (plate 1 and plate 2) (Figure 14). These correlations were high and comparable to other high-quality protein arrays (Ramachandran, N. *et al*. 2008). Furthermore, the variation in protein binding amount of INSOL CTA-array was within a narrow range of 87 times (Figure 13b). The commercially available Human ProtoArray (Thermo Fisher Scientific), has a large variation in the amount of protein binding between the minimum 0 and the maximum 42,000 net intensity, whereas the INSOL CTA-array has a more uniform amount of protein binding. Proteins purified by GST-tag are used to fabricate Human ProtoArray (Mattoon, D. *et al*. 2005). It is considered that more uniform amounts of proteins were obtained by centrifugal purification with INSOL-tag than proteins purified with GST-tag by a uniform method. In the *E. coli* expression system used in the SEREX method and *in vivo* proteome separated by two-dimensional electrophoresis, it is difficult to perform equal screening with no bias because the variation in individual protein amounts is extremely large. In protein arrays produced with INSOL-tagged proteins, the bias in protein binding is reduced, which makes it possible to detect the autoantibodies that bind to antigens that are less expressed in cancer cells or *E. coli*.

Characteristics of the INSOL-CTA array in autoantibody profiling of cancer patients

Autoantibodies against antigens specifically expressed in cancer cells arise in cancer patients and are correlated with the onset, progression, and recurrence of cancer (Tan, H.T. *et al*. 2009). To simultaneously identify autoantibodies, SEREX and phage display have been performed for a long

time, resulting in the identification of autoantibodies of various cancers. However, these methods have limitations, such as difficulties in detecting antigens that are expressed at a low level in cancer cells and bias towards antigens that are easily expressed in bacteria or those that are displayed on the phage surface. Accordingly, protein array technology was developed for expanding the scope for the selection of an expression system and allowing individual candidate proteins to be expressed more evenly. Protein arrays also have the advantage of requiring less serum for measurement (*ca.* 25 μL per measurement in the INSOL-CTA array). Recently, Human ProtoArray containing more than 9,000 human proteins were also developed. However, only 101 of 251 genes (276 clones) present in the INSOL-CTA array are included in Human ProtoArray; thus, the INSOL-CTA array covers a wider range of cancer antigens. Although Human ProtoArray has a low coverage of known cancer antigens, it contains many human antigens, so unknown cancer antigens could be found among them. Therefore, Human ProtoArray is important for identifying unknown cancer antigens, while the INSOL-CTA array will be important for monitoring therapeutic effects and recurrence using known cancer antigen profiles.

Autoantibody profiles in sera of digestive system cancer patients

In the serum autoantibody profile of two patients with gastrointestinal cancer revealed by INSOL CTA-array (Table 5), the antibodies detected in both patient 1 and patient 2 were those that bind to MAGEA1, MAGEA4, MAGEA8, and PTPN20A, respectively. In patient 1, anti-MAGEA4 antibody was positive before vaccination, and the antibody titer increased with each vaccination. Anti-MAGEA1, MAGEA5, MAGEA8 and ACPP antibodies became positive after vaccination (antigen spreading in patient 1). On the other hand, patient 2 was negative for the anti-MAGEA4 antibody before vaccination. In addition to the anti-MAGEA4 antibody, a large number of antibodies such as the anti-MAGEA1, MAGEA8, MAGEB6, ATAD2, LCK, PAGE2B, RSPH1, SPAG8, SSX2, SSX2B,

and TTK antibodies became positive after vaccination (antigen spreading in patient 2). Among the above, after the fourth vaccination that the anti-MAGEA4 antibody became positive, many antibodies (9 of 11) turned positive. It is possible that patient 1 who was positive for the anti-MAGEA4 antibody before vaccination had MAGEA4-specific cytotoxic T cells from the beginning, and there were few newly killed cancer cells after vaccination. Conversely, it was suggested that in patient 2, who became positive for the anti-MAGEA4 antibody after vaccination, cytotoxic T cells specific for MAGEA4 were activated for the first time after vaccination. As a result, cancer cells in which MAGEA4 was presented by HLA class I may be newly killed, possibly leading to a large number of antigen spreading. At the time of selecting patients to be vaccinated, in patients 1 and 2, it was confirmed that MAGEA4 antigen was expressed in their cancer cells. If the patient is selected by the criteria of MAGEA4 expressing cancer cells without serum MAGEA4 autoantibody (such as patient 2), the vaccine treatment may be more effective.

In the following, the pathological implications of individual antigens are discussed. ACPP (Acid Phosphatase, Prostate), which became positive after 6 vaccinations in patient 1, is regarded as an important biomarker for evaluating and monitoring prostate cancer (Azumi, N. *et al*. 1991, Wang, Y. *et al*. 2005, Gunia, S. *et al*. 2009). In recent years, ACPP is a potential biomarker for the diagnosis of epithelial ovarian cancer in women (Bae, H. *et al*. 2014). However, it was the first time that the serum anti-ACPP antibody was detected in patients with colon cancer, such as patient 1 in this study. ACPP is known to be expressed in the salivary gland, thymus, lung, kidney, brain, spleen, and thyroid in mice (Quintero, I.B. *et al*. 2007), but its expression in the human large intestine is unknown. Therefore, the relationship between anti-ACPP antibodies and colon cancer needs further investigation. SSX2 (Synovial Sarcoma X Breakpoint 2) and SSX2B, which became positive after 2 or 4 vaccinations in patient 2, are involved in t(X;18)(p11.2;q11.2) translocations detected in almost all synovial sarcoma patients (Guillou, L. *et al*. 2001). As a result of the translocation, SYT (Synovial

Sarcoma Translocated to X Chromosome Protein) on chromosome 18 is fused to one of the SSX genes on the X chromosome. As a result of gene fusion, hybrid proteins are expressed and thought to be involved in cell transformation. So far, a case of SYT/SSX2 translocation has been reported in duodenal synovial sarcoma (Schreiber-Facklam, H. *et al*. 2007). This suggests that there was SSX2 expression in the duodenum in the patient. Based on the above, it is possible that the hybrid protein produced by SYT / SSX2 translocation was expressed in cancer cells in patient 2 with duodenum papillary carcinoma, and anti-SSX2 antibodies were produced as a result of leakage of SSX2 from the broken cancer cells.

Conclusions and prospects

To my best knowledge, the protein array spotted with the world's largest number of cancer antigens was developed in this study. The cDNA clones are owned by the laboratory I belong to, and the simple purification method using INSOL-tag developed in Chapter 1 made this possible. Until now, there were few evaluation methods during the treatment of cancer vaccine therapy. Under such circumstances, it was allowed to profile anti-cancer antigen antibodies comprehensively by using INSOL CTL-array developed in this study. This method is expected to make it possible to diagnose patients who are suitable for cancer vaccine administration before other treatments or to monitor apoptosis of cancer cells due to the effect of the cancer vaccine. In the field of cancer vaccine administration, the effectiveness of comprehensively profiling anti-cancer antigen antibodies needs to be verified by increasing the number of cases in the future.

General Discussion

Positioning of INSOL-tag in protein preparation methods

The technology for preparing recombinant proteins has improved markedly in this decade (Wood, D.W. 2014, Yadav, D.K. *et al*. 2016). Diversification of expression systems, use of various expression control mechanisms, and development of tags with novel functions have contributed to this improvement. Every protein has a unique structure and function. Therefore, there is no universal preparation method that can be used for all proteins, and it is necessary to find an appropriate preparation method for each type of protein. However, the above-mentioned is a case where a protein had to be prepared while maintaining its structure. Unless the experimental system is sensitive to changes in protein structure, even if the target protein is insolubilized, there are further requirements to achieve a higher yield or to prepare hundreds of proteins in high-throughput. INSOL-tag is a novel tag that meets such needs owing to its high versatility and ease of use.

Dr. Goshima and his colleagues have previously expressed 50 representative human proteins in different expression systems and have compared their expression levels: wheat germ (*in vitro*), *E. coli* (*in vitro*), *E. coli* (*in vivo*), and cultured cell (DM2 and CHO) (Goshima, N., *et al*. 2008). Among these expression systems, the wheat germ cell-free expression system was remarkably superior, both in terms of the expression success rate and the expression level. Therefore, I adopted wheat germ (*in vitro*) instead of *E. coli* (*in vitro*) for this study. The combination of INSOL-tag and the wheat germ cell-free expression system may have led to the high recovery efficiency of the target proteins.

Improving the purity of protein purification by INSOL-tag (less than 60%) is one of the future challenges. One solution is to use INSOL-tag in combination with His-tag, which can also be used in the presence of denaturing agents. By solubilizing the precipitate obtained by centrifugal purification with SDS and then performing the second purification with a nickel column, the purity of

INSOL and His-tagged proteins may be improved. Since some INSOL-tagged kinases maintain their enzymatic activity, some of the other proteins may maintain their structure and activity even when INSOL-tag is fused. If INSOL-tag acts like a natural resin, its use as a tool for functional analysis of proteins will expand.

Possibility of INSOL-tag for quantitative proteomics

After the completion of human whole-genome sequencing, gene expression analysis using DNA microarrays has been actively conducted. However, the information obtained from the DNA microarray merely an indirect estimation of the protein expression *in vivo* via the amount of mRNA. As a technique for direct and comprehensive analyses of proteins *in vivo*, exhaustive measurement of cellular proteins by mass spectrometry has become popular (Aslam, B. *et al*. 2017). In the most common proteomics analysis, data sets of proteins contained in samples of different tissues and cells, and data sets before and after performing a specific treatment on the same cells are acquired and compared (Nilsson, T. *et al*. 2010). For comparison of data sets, a quantification technique by mass spectrometry of peptides derived from each protein is important. Quantitative techniques by mass spectrometry include relative quantification and absolute quantification. Absolute quantification using internal standards is considered ideal because of the high reproducibility of analyzes performed at different sites on different days (A known amount of peptide containing a stable isotope is added to the sample to be measured, and this peptide is used as an internal standard to quantify the amount of peptide in the sample.). However, to perform absolute quantification of human cellular proteins, it is necessary to express and purify about 20,000 types of human proteins as internal standards and determine their concentrations by SDS-PAGE.

The INSOL-tag developed in this study is expected to demonstrate its features and be most useful when preparing such an internal standard for mass spectrometry. INSOL-tag, which was

shown to be capable of insolubilizing and recovering about 400 target proteins in this study (proteins in Table2, Table 3 and Table 4), is considered to be highly versatile and applicable to the preparation of human proteome. In a 0.2 ml 8 strips PCR tube, INSOL-tagged protein was purified by centrifugation at 19,000 × g for 20 minutes. However, since purification can be performed by centrifugation at 6,000 × g for 60 minutes, purification using a 96- or 384-well plate is also possible. One disadvantage is that denaturation is required for the insolubilized INSOL-tagged protein, but INSOL-tag purification is sufficient because protein folding is not required for mass spectrometry. From the above, the use of INSOL-tag in a cell-free protein expression system can dramatically improve the throughput of absolute quantitative analysis in mass spectrometry and will allow for the analysis of cell physiology under specific conditions such as cancer or drug-induced changes in cells at the level of the number of molecules.

Significance of autoantibody measurement in cancer patients and prospects

Measurement of autoantibodies to cancer antigens has been attempted since the 1960s (Day, E.D. *et al*. 1965). Subsequently, various methods such as SDS-PAGE, ELISA, SEREX, two-dimensional PAGE, protein array were developed, and major cancer antigens were discovered. Autoantibody signatures in the serum of cancer patients have shown that it is possible to distinguish a variety of cancer patients from healthy individuals and patients with related benign diseases. Thus, the challenge has now begun to monitor the entire process from cancer development to healing with autoantibody profiles using the knowledge and tools developed so far. In this study, I developed a protein array that can be used to monitor the occurrence, outcome of surgical resection, and recurrence of cancer as well as to evaluate cancer vaccine treatment. The analysis suggests the criteria for selecting patients who would be eligible for MAGEA4 vaccine treatment. It is considered that the

high-yield preparation of the antigens using INSOL-tag enabled the measurement of autoantibodies against a large number of cancer antigens more than ever before, leading to this result.

Cancer remains the leading cause of death, and there is a need for early treatment decisions to reduce the rate of death. Measurement of serum autoantibodies can be performed by collecting a small amount of blood so that the physical burden on patients and the economic cost are reduced. In addition, the serum autoantibody profile changes with the course of treatment. Due to its low burden and high disease detection ability, serologic tests will continue to be important. Among immune checkpoint inhibitors that are currently attracting attention in cancer immunotherapy, there is a concern that excessive activation of the autoimmune reaction may cause autoimmune diseases (Yasunaga, M. 2019). For future evaluation of cancer immunotherapy, it will be necessary to include autoimmune disease-related antigens in the protein array in addition to cancer antigens.

Acknowledgements

I express my sincere thanks to Professor Tetsuo Hashimoto at the Graduate School of Life and Environmental Sciences, the University of Tsukuba for his helpful advices and supports to my all activities to prepare the doctoral dissertation. I deeply appreciate Professor Tomoki Chiba, Professor Yuji Inagaki, Professor Kisaburo Nagamune, and Junior Assistant Professor Ikuko Yuyama at the Graduate School of Life and Environmental Sciences, the University of Tsukuba for reviewing the manuscript and providing helpful comments.

I deeply thank my colleagues at National Institute of Advanced Industrial Science and Technology (AIST); Dr. Naoki Goshima for giving me the opportunity and encouragement to carry out this study; Dr. Masatoshi Mori for providing advices for the various experiments; K. Ogawa, T. Ono, M. Shioya for the preparation of recombinant proteins; Dr. Y. Kawamura and M. Takasaka for the creation of the gene list.

I deeply appreciate Dr. Hiroshi Shiku, and Dr. Yoshihiro Miyahara at Mie University for providing sera from the gastrointestinal cancer patients.

I would also like to express my gratitude to my parents and my husband Takeshi Ebina for the ongoing and unwavering support.

I am an employee of AIST. This study was carried out at the Molecular Profiling Research Center for Drug Discovery (molprof), AIST.

References

Achmüller, C., Kaar, W., Ahrer, K., Wechner, P., Hahn, R., Werther, F., Schmidinger, H., Cserjan-Puschmann, M., Clementschitsch, F., Striedner, G., Bayer, K., Jungbauer, A., and Auer, B. (2007) "Npro fusion technology to produce proteins with authentic N termini in E. coli" Nat. Methods. 4(12), 1037-43.

Almeida, L.G., Sakabe, N.J., deOliveira, A.R., Silva, M.C., Mundstein, A.S., Cohen, T., Chen, Y.T., Chua, R., Gurung, S., Gnjatic, S., Jungbluth, A.A., Caballero, O.L., Bairoch, A., Kiesler, E., White, S.L., Simpson, A.J., Old, L.J., Camargo, A.A., and Vasconcelos, A.T. (2009) "CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens" Nucleic. Acids. Res. 37(Database issue), D816-9.

Aslam, B., Basit, M., Nisar, M.A., Khurshid, M., and Rasool, M.H. (2017) "Proteomics: Technologies and Their Applications." J. Chromatogr. Sci. 55(2):182-196.

Azumi, N., Traweek, S.T., and Battifora, H. (1991) "Prostatic acid phosphatase in carcinoid tumors. Immunohistochemical and immunoblot studies." Am. J. Surg. Pathol. 15(8):785-90.

Bae, H., Lim, W., Bae, S.M., Bazer, F.W., Choi, Y., and Song, G. (2014) "Avian prostatic acid phosphatase: estrogen regulation in the oviduct and epithelial cell-derived ovarian carcinomas." Biol. Reprod. 91(1):3.

Banki, M.R., Feng, L., and Wood, D.W. (2005) "Simple bioseparations using self-cleaving elastin-like

polypeptide tags" Nat. Methods. 2(9), 659-61.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F., Perez-Amodio, S., Strippoli, P., and Canaider, S. (2013) "An estimation of the number of cells in the human body." Ann. Hum. Biol. 40(6):463-71.

Brichory, F., Beer, D., Le, Naour, F., Giordano, T., and Hanash, S. (2001) "Proteomics-based identification of protein gene product 9.5 as a tumor antigen that induces a humoral immune response in lung cancer." Cancer. Res. 1;61(21):7908-12.

Canelle, L., Bousquet, J., Pionneau, C., Deneux, L., Imam-Sghiouar, N., Caron, M., and Joubert-Caron, R. (2005) "An efficient proteomics-based approach for the screening of autoantibodies." J. Immunol. Methods. 299(1-2):77-89.

Cheever, M.A., Allison, J.P., Ferris, A.S., Finn, O.J., Hastings, B.M., Hecht, T.T., Mellman, I., Prindiville, S.A., Viner, J.L., Weiner, L.M., and Matrisian, L.M. (2009) "The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research" Clin. Cancer. Res. 15(17), 5323-37.

Cheng, X., Lu, W., Zhang, S., and Cao, P. (2010) "Expression and purification of antimicrobial peptide CM4 by Npro fusion technology in E. coli." Amino. Acids. 39(5):1545-52.

Corbière, V., Chapiro, J., Stroobant, V., Ma, W., Lurquin, C., Lethé, B., van Baren, N., Van, den, Eynde, B.J., Boon, T., and Coulie, P.G. (2011) "Antigen spreading contributes to MAGE vaccination-induced

regression of melanoma metastases." Cancer. Res. 71(4):1253-62.

Costa, S.J., Almeida, A., Castro, A., Domingues, L., and Besir, H. (2013a) "The novel Fh8 and H fusion partners for soluble protein expression in Escherichia coli: a comparison with the traditional gene fusion technology." Appl. Microbiol. Biotechnol. 97(15):6779-91.

Costa, S.J., Coelho, E., Franco, L., Almeida, A., Castro, A., and Domingues, L. (2013b) "The Fh8 tag: a fusion partner for simple and cost-effective protein purification in Escherichia coli." Protein. Expr. Purif. 92(2):163-70.

Coulie, P.G., Lehmann, F., Lethé, B., Herman, J., Lurquin, C., Andrawiss, M., and Boon, T. (1995) "A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma." Proc. Natl. Acad. Sci. U. S. A. 92(17):7976-80.

Day, E.D., Lassiter, S., Woodhall, B., Mahaley, J.L., and Mahaley, M.S. Jr. (1965) "The localization of radioantibodies in human brain tumors. I. Preliminary exploration." Cancer. Res. 25(6):773-8.

De Plaen, E., Arden, K., Traversari, C., Gaforio, J.J., Szikora, J.P., De, Smet, C., Brasseur, F., van der Bruggen, P., Lethé, B., Lurquin, C., et al. (1994) "Structure, chromosomal localization, and expression of 12 genes of the MAGE family." Immunogenetics. 40(5):360-9.

Derynck, R., Roberts, A.B., Winkler, M.E., Chen, E.Y., and Goeddel, D.V. (1984) "Human transforming growth factor-α: Precursor structure and expression in E. coli" Cell 38(1), 287-97.

Ding, F.X., Yan, H.L., Mei, Q., Xue, G., Wang, Y.Z., Gao, Y.J., and Sun, S.H. (2007) "A novel, cheap and effective fusion expression system for the production of recombinant proteins." Appl. Microbiol. Biotechnol. 77(2):483-8.

Domínguez, A., Fernández, A., González, N., Iglesias, E., and Montenegro, L. (1997) "Determination of Critical Micelle Concentration of Some Surfactants by Three Techniques" J. Chem. Educ. 74 (10): 1227-31.

Facciabene, A., Motz, G.T., and Coukos, G. (2012) "T-regulatory cells: key players in tumor immune escape and angiogenesis." Cancer. Res. 72(9):2162-71.

Goshima, N., Kawamura, Y., Fukumoto, A., Miura, A., Honma, R., Satoh, R., Wakamatsu, A., Yamamoto, J., Kimura, K., Nishikawa, T., Andoh, T., Iida, Y., Ishikawa, K., Ito, E., Kagawa, N., Kaminaga, C., Kanehori, K., Kawakami, B., Kenmochi, K., Kimura, R., Kobayashi, M., Kuroita, T., Kuwayama, H., Maruyama, Y., Matsuo, K., Minami, K., Mitsubori, M., Mori, M., Morishita, R., Murase, A., Nishikawa, A., Nishikawa, S., Okamoto, T., Sakagami, N., Sakamoto, Y., Sasaki, Y., Seki, T., Sono, S., Sugiyama, A., Sumiya, T., Takayama, T., Takayama, Y., Takeda, H., Togashi, T., Yahata, K., Yamada, H., Yanagisawa, Y., Endo, Y., Imamoto, F., Kisu, Y., Tanaka, S., Isogai, T., Imai, J., Watanabe, S., and Nomura, N. (2008) "Human protein factory for converting the transcriptome into an in vitro-expressed proteome" Nat. Methods. 5(12), 1011-7.

Guillou, L., Coindre, J., Gallagher, G., Terrier, P., Gebhard, S., de Saint, Aubain, Somerhausen, N., Michels, J., Jundt, G., Vince, D.R., Collin, F., Trassard, M., Le Doussal, V., and Benhattar, J. (2001) "Detection of the synovial sarcoma translocation t(X;18) (SYT;SSX) in paraffin-embedded tissues

using reverse transcriptase-polymerase chain reaction: a reliable and powerful diagnostic tool for pathologists. A molecular analysis of 221 mesenchymal tumors fixed in different fixatives." Hum. Pathol. 32(1):105-12.

Gulley, J.L. (2012) "Therapeutic vaccines: the ultimate personalized therapy?" Hum. Vaccin. Immunother. 9(1), 219-21.

Gunia, S., Koch, S., May, M., Dietel, M., and Erbersdobler, A. (2009) "Expression of prostatic acid phosphatase (PSAP) in transurethral resection specimens of the prostate is predictive of histopathologic tumor stage in subsequent radical prostatectomies." Virchows. Arch. 454(5):573-9.

Hayashi, M., Iwamoto, S., Sato, S., Sudo, S., Takagi, M., Sakai, H., and Hayakawa, T. (2013) "Efficient production of recombinant cystatin C using a peptide-tag, 4AaCter, that facilitates formation of insoluble protein inclusion bodies in Escherichia coli." Protein. Expr. Purif. 88(2):230-4.

Hoek, K.S., Schlegel, N.C., Eichhoff, O.M., Widmer, D.S., Praetorius, C., Einarsson, S.O., Valgeirsdottir, S., Bergsteinsdottir, K., Schepsky, A., Dummer, R., and Steingrimsson, E. (2008) "Novel MITF targets identified using a two-step DNA microarray strategy." Pigment. Cell. Melanoma. Res. 21(6):665-76.

Hwang, P.M., Pan, J.S., and Sykes, B.D. (2012) "A PagP fusion protein system for the expression of intrinsically disordered proteins in Escherichia coli" Protein Expr. Purif. 85(1), 148-51.

Kawakami, Y., Wang, X., Shofuda, T., Sumimoto, H., Tupesis, J., Fitzgerald, E., and Rosenberg, S.

(2001) "Isolation of a new melanoma antigen, MART-2, containing a mutated epitope recognized by autologous tumor-infiltrating T lymphocytes." J. Immunol. 166(4):2871-7.

Kellermann, O.K., and Ferenci, T. (1982) "Maltose-binding protein from Escherichia coli." Methods. Enzymol. 90 Pt E:459-63.

Koziol, J.A., Zhang, J.Y., Casiano, C.A., Peng, X.X., Shi, F.D., Feng, A.C., Chan, E.K., and Tan, E.M. (2003) "Recursive partitioning as an approach to selection of immune markers for tumor diagnosis." Clin. Cancer. Res. 1;9(14):5120-6.

Kramer, A., Keitel, T., Winkler, K., Stöcklein, W., Höhne, W., and Schneider-Mergener, J. (1997) "Molecular basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody." Cell. 91(6):799-809.

Kuliopulos, A., and Walsh, C.T. (1994) "Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides" J. Am. Chem. Soc. 116(11), 4599-607

Kuliopulos, A., Shortle, D., and Talalay, P. (1987) "Isolation and sequencing of the gene encoding delta 5-3-ketosteroid isomerase of Pseudomonas testosteroni: overexpression of the protein." Proc. Natl. Acad. Sci. U. S. A. 84(24):8893-7.

Kuo, W.H., and Chase, H.A. (2011) "A new strategy for the on-column exopeptidase cleavage of poly-histidine tagged proteins." J. Chromatogr. B. Analyt. Technol. Biomed. Life. Sci. 879(28):3028-33.

Le, Naour, F., Brichory, F., Misek, D.E., Bréchot, C., Hanash, S.M., Beretta, L. (2002) "A distinct

repertoire of autoantibodies in hepatocellular carcinoma identified by proteomic analysis." Mol. Cell. Proteomics. 1(3):197-203.

Le, Naour, F., Misek, D.E., Krause, M.C., Deneux, L., Giordano, T.J., Scholl, S., Hanash, S.M.. (2001) "Proteomics-based identification of RS/DJ-1 as a novel circulating tumor antigen in breast cancer." Clin. Cancer. Res. 7(11):3328-35.

Lee, J.H., Kim, J.H., Hwang, S.W., Lee, W.J., Yoon, H.K., Lee, H.S., and Hong, S.S. (2000) "High-level expression of antimicrobial peptide mediated by a fusion partner reinforcing formation of inclusion bodies" Biochem. Biophys. Res. Commun. 277(3), 575-80.

Lin, Z., Zhao, Q., Xing, L., Zhou, B., and Wang, X. (2015) "Aggregating tags for column-free protein purification" Biotechnol. J. 10(12), 1877-86.

Maertens, B., Spriestersbach, A., Kubicek, J., and Schäfer, F. (2015) "Strep-Tagged Protein Purification." Methods. Enzymol. 559:53-69.

Mattoon, D., Michaud, G., Merkel, J., and Schweitzer, B. (2005) "Biomarker discovery using protein microarray technology platforms: antibody-antigen complex profiling." Expert. Rev. Proteomics. 2(6):879-89.

Miyauchi, K., Tsuchikawa, T., Wada, M., Abiko, T., Kyogoku, N., Shichinohe, T., Miyahara, Y., Kageyama, S., Ikeda, H., Shiku, H., and Hirano, S. (2016) "Clinical relevance of antigen spreading pattern induced by CHP-MAGE-A4 cancer vaccination" Immunotherapy 8(5), 527-40.

Motohashi, H., Fujita, R., Takayama, M., Inoue, A., Katsuoka, F., Bresnick, E.H., and Yamamoto, M. (2011) "Molecular determinants for small Maf protein control of platelet production." Mol. Cell. Biol. 31(1), 151-62.

Nil, A., Firat, E., Sobek, V., Eichmann, K., and Niedermann, G. (2004) "Expression of housekeeping and immunoproteasome subunit genes is differentially regulated in positively and negatively selecting thymic stroma subsets." Eur. J. Immunol. 34(10):2681-9.

Nilsson, T., Mann, M., Aebersold, R., Yates, J.R. 3rd., Bairoch, A., and Bergeron, J.J. (2010) "Mass spectrometry in high-throughput proteomics: ready for the big time." Nat. Methods. 7(9):681-5.

Ogura, T. (2014) "Direct observation of unstained biological specimens in water by the frequency transmission electric-field method using SEM" PLoS One 9(3), e92780.

Ogura, T. (2015) "Nanoscale analysis of unstained biological specimens in water without radiation damage using high-resolution frequency transmission electric-field system based on FE-SEM" Biochem. Biophys. Res. Commun. 459(3), 521-8.

Old, L.J. (1981) "Cancer immunology: the search for specificity--G. H. A. Clowes Memorial lecture." Cancer. Res. 41(2):361-75.

Old, L.J., and Chen, Y.T. (1998) "New paths in human cancer serology." J. Exp. Med. 20;187(8):1163-7.

Quintero, I.B., Araujo, C.L., Pulkka, A.E., Wirkkala, R.S., Herrala, A.M., Eskelinen, E.L., Jokitalo, E., Hellström, P.A., Tuominen, H.J., Hirvikoski, P.P., and Vihko, P.T. (2007) "Prostatic acid phosphatase is not a prostate specific target." Cancer. Res. 67(14):6549-54.

Rajagopalan, K., Mooney, S.M., Parekh, N., Getzenberg, R.H., and Kulkarni, P. (2011) "A majority of the cancer/testis antigens are intrinsically disordered proteins" J. Cell. Biochem. 112(11), 3256-67.

Ramachandran, N., Raphael, J.V., Hainsworth, E., Demirkan, G., Fuentes, M.G., Rolfs, A., Hu, Y., and LaBaer, J. (2008) "Next-generation high-density self-assembling functional protein arrays." Nat. Methods. 5(6):535-8.

Schmidt, P.M., Sparrow, L.G., Attwood, R.M., Xiao, X., Adams, T.E., and McKimm-Breschkin, J.L. (2012) "Taking down the FLAG! How insect cell expression challenges an established tag-system." PLoS. One. 7(6):e37779.

Schreiber-Facklam, H., Bode-Lesniewska, B., Frigerio, S., and Flury, R. (2007) "Primary monophasic synovial sarcoma of the duodenum with SYT/SSX2 type of translocation." Hum. Pathol. 38(6):946-9.

Senga, Y., Imamura, H., Ogura, T., and Honda, S. (2019) "In-Solution Microscopic Imaging of Fractal Aggregates of a Stressed Therapeutic Antibody" Anal. Chem. 91(7), 4640-8.

Shur, O., Dooley, K., Blenner, M., Baltimore, M., and Banta, S. (2013) "A designed, phase changing RTX-based peptide for efficient bioseparations" Biotechniques 54(4), 197-206.

Smith, D.B., and Johnson, K.S. (1988) "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase." Gene. 67(1):31-40.

Takeda, H., Kawamura, Y., Miura, A., Mori, M., Wakamatsu, A., Yamamoto, J., Isogai, T., Matsumoto, M., Nakayama, K.I., Natsume, T., Nomura, N., and Goshima, N. (2010) "Comparative analysis of human SRC-family kinase substrate specificity in vitro" J. Proteome Res. 9(11), 5982-93.

Tan, H.T., Low, J., Lim, S.G., and Chung, M.C. (2009) "Serum autoantibodies as biomarkers for early cancer detection." FEBS J. 276(23), 6880-904.

van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A., and Boon, T. (1991) "A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma." Science. 254(5038):1643-7.

Vidovic, V., Prongidi-Fix, L., Bechinger, B., and Werten, S. (2009) "Production and isotope labeling of antimicrobial peptides in Escherichia coli by means of a novel fusion partner that enables high-yield insoluble expression and fast purification" J. Pept. Sci. 15(4), 278-84.

Vigneron, N., Ooms, A., Morel, S., Degiovanni, G., and Van, Den, Eynde, B.J. (2002) "Identification of a new peptide recognized by autologous cytolytic T lymphocytes on a human melanoma." Cancer. Immun. 19;2:9.

Vigneron, N., Stroobant, V., Van, den, Eynde, B.J., and van der Bruggen, P. (2013) "Database of T

cell-defined human tumor antigens: the 2013 update." Cancer. Immun. 15;13:15.

Walkup, W.G. 4th., and Kennedy, M.B. (2014) "PDZ affinity chromatography: a general method for affinity purification of proteins based on PDZ domains and their ligands." Protein. Expr. Purif. 98:46-62.

Wang, Y., Harada, M., Yano, H., Ogasawara, S., Takedatsu, H., Arima, Y., Matsueda, S., Yamada, A., and Itoh, K. (2005) "Prostatic acid phosphatase as a target molecule in specific immunotherapy for patients with nonprostate adenocarcinoma." J. Immunother. 28(6):535-41.

Wilson, I.A., Niman, H.L., Houghten, R.A., Cherenson, A.R., Connolly, M.L., and Lerner, R.A. (1984) "The structure of an antigenic determinant in a protein." Cell. 37(3):767-78.

Wölfel, T., Van, Pel, A., Brichard, V., Schneider, J., Seliger, B., Meyer, zum, Büschenfelde, K.H., and Boon, T. (1994) "Two tyrosinase nonapeptides recognized on HLA-A2 melanomas by autologous cytolytic T lymphocytes." Eur. J. Immunol. 24(3):759-64.

Wood, D.W. (2014) "New trends and affinity tag designs for recombinant protein purification" Curr. Opin. Struct. Biol. 26, 54-61.

Xia, Q., Kong, X.T., Zhang, G.A., Hou, X.J., Qiang, H., and Zhong, R.Q. (2005) "Proteomics-based identification of DEAD-box protein 48 as a novel autoantigen, a prospective serum marker for pancreatic cancer." Biochem. Biophys. Res. Commun. 6;330(2):526-32.

Xing, L., Wu, W., Zhou, B., and Lin, Z. (2011) "Streamlined protein expression and purification using cleavable self-aggregating tags" Microb. Cell Fact. 10, 42-8.

Yadav, D.K., Yadav, N., Yadav, S., Haque, S., and Tuteja, N. (2016) "An insight into fusion technology aiding efficient recombinant protein production for functional proteomics" Arch. Biochem. Biophys. 612, 57-77.

Yasunaga, M. (2019) "Antibody therapeutics and immunoregulation in cancer and autoimmune disease." Semin. Cancer. Biol. pii: S1044-579X(18)30187-1.

Zhang, J.Y. (2004) "Tumor-associated antigen arrays to enhance antibody detection for cancer diagnosis." Cancer. Detect. Prev. 28(2):114-8.

Zhang, J.Y., Casiano, C.A., Peng, X.X., Koziol, J.A., Chan, E.K., Tan, E.M. (2003) "Enhancement of antibody detection in cancer using panel of recombinant tumor-associated antigens." Cancer. Epidemiol. Biomarkers. Prev. 12(2):136-43.

Zheng, C.F., Simcox, T., Xu, L., and Vaillancourt, P. (1997) "A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins." Gene. 186(1):55-60.