Chapter 4

Applications of the *In Vitro* Virus (IVV) Method for Various Protein Functional Analyses

Noriko Tabata, Kenichi Horisawa and
Hiroshi Yanagawa

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/55483

1. Introduction

The complete decoding of the human genome in 2003 raised expectations for targeted drug design and personalized medical care based on genetic information and also for gene therapy of disease. To reach these goals, it is necessary to develop new technologies capable of identifying drug-target genes from enormous volumes of genome information, validating the biological functions, and comprehensively analyzing protein functions and interactions that have hitherto been studied individually. We first developed an mRNA display, termed the *in vitro* virus (IVV) method [1-3], and a C-terminal protein labeling method [4,5]that provided the required technology (Fig. 1).

The IVV method, originally developed for evolutionary protein engineering based on *in vitro* translation systems, was subsequently applied for the analysis of various protein interactions. In the IVV method, the genotype molecule (mRNA) is linked to the phenotype molecule (protein) through puromycin in a cell-free translation system [3]. We have developed this method into a stable, efficient, and high-throughput technique that allows simple selection without any requirement for post-translational work [3], unlike previous systems. An additional technological advance is the elimination of the need to express and purify bait proteins for downstream protein-protein interaction studies. Our totally *in vitro* cell-free co translation system provides a simpler solution that is suitable for high-throughput, genome-wide analysis, as baits are synthesized within each reaction. Moreover, as cotranslation of bait and prey proteins is favorable for the formation of multi-protein complexes, this approach offers a better chance to obtain a more comprehensive data set, including both direct and indirect interactions, in a single experiment. Here, we provide an overview of this system and discuss its advantages for analyses of various protein interactions including protein-protein, DNA(RNA)-protein, peptide-protein, drug-protein, and antigen-antibody interactions.

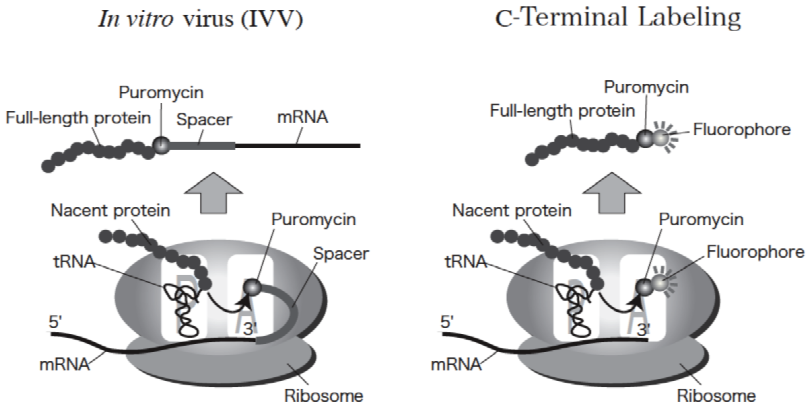

Figure 1. *In vitro* virus(IVV) formation and C-terminal labeling on the ribosome. Puromycin at the 3′-terminal end of a spacer ligated to mRNA can enter the ribosomal A-site to bind covalently to the C-terminal end of the encoded full-length protein in the ribosomal P-site. The same property of puromycin to develop a highly specific labeling system for proteins was also utilized, in which puromycin derivatives bearing a fluorescein moiety are used to label the C-terminal end of a full-length protein.

The antibiotic puromycin [6], which is an analogue of the 3′ end of TyrtRNATyr [7], acts in both prokaryotes and eukaryotes [8] as an inhibitor of peptidyl transferase [9]. It has two modes of inhibitory action. The first is by acting as an acceptor substrate that attacks peptidyl-tRNA (donor substrate) in the P site to form a nascent peptide [10]. The second is by competing with aminoacyl-tRNA for binding to the A′ site, which is the binding site of the 3′ end of aminoacyl-tRNA within the peptidyl transferase active site [11]. It has been reported that the polypeptides released by puromycin are not full-length proteins [9]. Similarly, it has been shown that growing peptide chains on ribosomes are transferred to the α-amino group of puromycin, which interrupts the normal reaction of peptide bond formation [10]. Therefore, these conventional studies suggested that puromycin is a non-specific inhibitor of protein synthesis as a result of competition with aminoacyl-tRNA. However, since most of the studies on puromycin were performed at relatively high concentrations, the behavior of puromycin at lower concentrations was still an open question at that time. It had been reported that full-length protein that fails to be released from ribosomes at the final stage of protein synthesis, requires treatment with puromycin or RFs to be released [12]. These results led us to hypothesize that puromycin at very low concentrations, which would not effectively compete with aminoacyl-tRNA [2], might act as a noninhibitor and bond specifically to full-length protein at the stop codon. Indeed, we confirmed that puromycin and its derivatives bond only to full-length protein at very low concentrations (such as 0.04 μM), where they are "non-inhibitors" of protein synthesis, by using ^{32}P-labeled rCpPuro (2′-ribocytidylyl-(3'→5')- puromycin) in an *E. coli* S30 extract cell-free translation system [2]. Our results provided the first evidence that specific bonding of puromycin to the full-length protein occurs at the stop codon during the process of termination of protein synthesis at very low concentration. In other words, puromycin at sufficiently low concentrations only has the opportunity to be bonded to proteins at

http://dx.doi.org/10.5772/55483

a stop codon, where it does not need to compete with aminoacyl-tRNA. Since termination is a relatively slow step involving a translational pause in eukaryotes [13] and *E. coli* [14], it is possible that puromycin even at very low concentrations can bind to the A' site, and compete with RFs to release the full-length protein from ribosomes. Accordingly, under these conditions, puromycin can be incorporated specifically at the C-terminus of the full-length protein. This concept is the basis of so-called puromycin technology [2,3]. The combination of this puromycin technology with *in vitro* translation has yielded novel methods, such as IVV and C-terminal labeling techniques, for protein selection and screening [15,16], fluorescence labeling [2,5], and affinity purification (pull-down assay), as well as protein chips for proteomics [5,17], and superior methods for evolutionary protein engineering.

2. Application of the IVV method for analyzing various protein interactions

The IVV method is applicable for analyzing protein–protein [15,16,18,19] DNA(RNA)–protein [20-22], peptide–protein [23-27], drug–protein [28, 29] and antigen–antibody [30,31] interactions. As a result, we are able to explore protein complexes, transcription factors, RNA-binding proteins, bioactive peptides, drug-target proteins, and antibodies (antibodies will be discussed later) (Fig. 2). library and conjugation with a polyethylene glycol (PEG) spacer with puromycin, cell-free translation to form IVV and interaction of bait attached to beads and prey IVV library to form complexes, IVV selection with bait, and RT-PCR to amplify mRNA tags. Selection rounds are repeated until sufficient enrichment is obtained, followed by cloning and sequencing, and protein sequences are decoded.

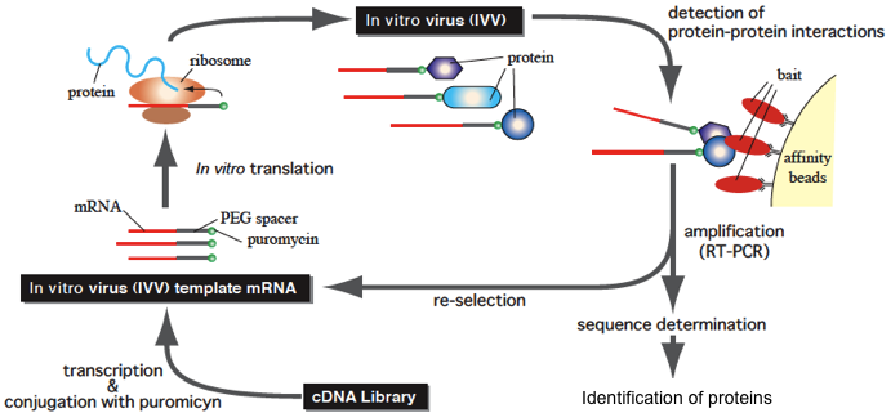

Figure 2. IVV selection based on cell-free co translation. The IVV selection procedure is composed of transcription from a cDNA

2.1. DNA-protein interaction

The specific interactions between cis-regulatory DNA elements and transcription factors are critical components of transcriptional regulatory networks [32,33]. The whole genome and complete cDNA sequences contain large numbers of transcription factors and their binding DNA sequences, and thus comprehensive analysis of DNA-transcription factor interactions is expected to provide a deep understanding of the mechanisms of cell proliferation, developmental processes in tissue morphogenesis and disease. Currently, combined use of chromatin immunoprecipitation (ChIP) assay with DNA microarrays (ChIP-chip)[34] has become the most widely used high-throughput method for discovering cis-regulatory DNA elements for a transcription factor. In contrast, development of high-throughput methods for discovering transcription factors for a cis-regulatory DNA element remains at an early stage. Although the yeast one-hybrid method [35] and phage display [36] are attractive candidates, these methods are not easily scalable because of the use of living cells. In addition, as over-expression of transcription factors often affects cellular metabolism, such transcription factors are difficult to screen. In order to circumvent these difficulties, we focused on a totally *in vitro* mRNA display technology such as IVV method [1-3,15,16] for the discovery of DNA–protein interactions.

Comprehensive analysis of DNA–protein interactions is important for mapping transcriptional regulatory networks at a genome-wide level. We employed the IVV method for *in vitro* selection of DNA-binding protein heterodimeric complexes [20]. Under improved selection conditions using a TPA-responsive element (TRE) as a bait DNA, known interactors c-fos and c-jun were simultaneously enriched about 100-fold from a model library (a 1:1:20 000 mixture of c-fos, c-jun and gst genes) after one round of selection. Furthermore, almost all of the AP-1 family genes, including c-jun, c-fos, junD, junB, atf2 and b-atf, were successfully selected from an IVV library constructed from a mouse brain poly A^+ RNA after six rounds of selection. These results indicate that the IVV selection system can identify a variety of DNA-binding protein complexes in a single experiment. Since almost all transcription factors form hetero-oligomeric complexes to bind with their target DNA, this method should be most useful to search for DNA-binding transcription factor complexes.

2.2. Peptide–protein interaction

Peptides are powerful tools for disrupting protein-protein interactions because the large interacting surfaces and the high specificity of these peptides lead to fewer adverse side effects when they are used as pharmaceutical agents [37]. As previously reported, several peptides that inhibit the MDM2-p53 interaction have been identified from randomized peptide libraries using phage display [38]. Hu et al. identified a 12-amino-acid (aa) peptide (LTFEHYWAQLTS), DI, that could inhibit not only the MDM2-p53 interaction, but also the MDMX-p53 interaction more effectively than Nutlin-3, a small molecular inhibitor of the MDM2-p53 interaction [39]. An MDM2 homologue, MDMX is highly expressed in tumors, and it binds to and negatively regulates p53 [40]. Furthermore, DI expressed with recombinant adenovirus as a thioredoxin-fused protein could activate the p53 pathway both *in vitro* and *in vivo*. However, DI was not

sufficiently optimized because it was selected by phage display from a 12-mer random library (4.161,015 possible members) with a size of ~10^8 that did not cover all of the possible sequences.

To overcome this problem, we performed *in vitro* selection of MDM2-binding peptides from random peptide libraries using the IVV method [1-3,15]. This system based on cell-free translation is a potent method for screening large peptide libraries (10^{13} unique members) and is able to cover all of the possible sequences in a 10-mer random library. We applied the IVV method to identify a highly optimized peptide that could disrupt the MDM2-p53 complex from a random library containing all of the possible sequences by dividing the selection process into two stages. We also verified that a selected peptide could inhibit the MDM2-p53 interaction in living cells and block tumor cell growth.

We identified an optimal peptide named MIP that inhibited the MDM2-p53 and MDMX-p53 interactions 29- and 13-fold more effectively than DI, respectively (Fig. 3) [23]. Adenovirus-mediated expression of MIP fused to the thioredoxin scaffold protein in living cells caused stabilization of p53 through its interaction with MDM2, resulting in activation of the p53 pathway. Furthermore, expression of MIP also inhibited tumor cell proliferation in a p53-dependent manner more potently than did DI. These results show that two-stage, the peptide selection by IVV method [1-3,15] is useful for the rapid identification of potent peptides that target oncoproteins.

Bcl-X_L, an antiapoptotic member of the Bcl-2 family, is a mitochondrial protein that inhibits activation of Bax and Bak, which commit the cell to apoptosis, and it therefore represents a potential target for drug discovery. Peptides have potential as therapeutic molecules because they can be designed to engage a larger portion of the target protein with higher specificity. We selected 16-mer peptides that interact with Bcl-X_L from random and degenerate peptide libraries using the IVV method [24]. The selected peptides have sequence similarity with the Bcl-2 family BH3 domains, and one of them has higher affinity (IC_{50} = 0.9 μM) than Bak BH3 (IC_{50} = 11.8 μM) for Bcl-X_L *in vitro*. We also found that GFP fusions of the selected peptides specifically interact with Bcl-X_L, localize in mitochondria, and induce cell death. Further, a chimeric molecule, in which the BH3 domain of Bak protein was replaced with a selected peptide, retained the ability to bind specifically to Bcl-X_L. These results demonstrate that this selected peptide specifically antagonizes the function of Bcl-X_L and overcomes the effects of Bcl-X_L in intact cells. Thus, the IVV method is a powerful technique to identify peptide inhibitors with high affinity and specificity for disease-related proteins.

The importin α/β pathway mediates nuclear import of proteins containing the classical nuclear localization signals (NLSs). Although the consensus sequences of the classical NLSs have been defined, there are still many NLSs that do not match the consensus rule and many nonfunctional sequences that match the consensus. We identified six different NLS classes that specifically bind to distinct binding pockets of importin α. By screening of random peptide libraries using the IVV method, we selected peptides bound by importin α and identified six classes of NLSs, including three novel classes (Table 1). Two noncanonical classes (class 3 and class 4) specifically bound to the minor binding pocket of importinα, whereas the classical monopartite NLSs (class 1 and class 2) bound to the major binding pocket. Using a newly developed universal green fluorescent protein expression system, we

found that these NLS classes, including plant-specific class 5 NLSs and bipartite NLSs, fundamentally require regions outside the core basic residues for their activity and have specific residues or patterns that confer distinct activities between yeast, plants, and mammals. Furthermore, amino acid replacement analyses revealed that the consensus basic patterns of the classical NLSs are not essential for activity, and more unconventional patterns, including redox-sensitive NLSs, were generated. These results explain the causes of NLS diversity. The defined consensus patterns and properties of importin α-dependent NLSs provide useful information for identifying NLSs [25-27].

Clone No.	Peptide sequence	Frequency
X12-1 (MIP)	PRFWEYWLRLME	10
X12-2	KSFQQYWQELML	9
X12-3	KTFEEYWLMLMS	4
X12-4	PSFWEHWVELML	4
X12-5	KRFQDYWSELML	3
$p53_{17-28}$	ETFSDLWKLLPE	-
	19 23 26	

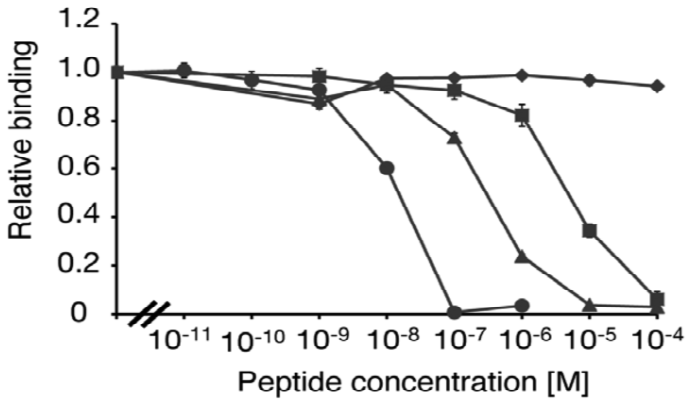

Figure 3. A multiple-sequence alignment of the selected peptides(upper) and inhibition of MDM2-p53 interactions by synthetic peptides(lower). Peptide sequences selected from randomized mRNA displayed peptide libraries were aligned using the ClustalW program. 12-mer peptides after five rounds of selection. Fixed amino acid residues are shown in bold. Values below the displayed $p53_{17-28}$ sequence indicate the position of amino acids in the p53 protein. MDM2 was generated by an *in vitro* transcription reaction, bound to His6-p53 immobilized on copper-coated plates in the presence of various concentrations of synthetic MIP(circle), DI (triangle), 3A (diamond) or p53 (square) peptides and quantified by ELISA.

http://dx.doi.org/10.5772/55483

NLS class	Consensus sequence[a]
Class 1	KR(K/R)R, K(K/R)RK
Class 2	(P/R)*XX*KR(^DE)(K/R)
Class 3	KR*X*(W/F/Y)*XX*AF
Class 4	(R/P)*XX*KR(K/R)(^DE)
Class 5	LGKR(K/R)(W/F/Y)
Bipartite	KRX_{10-12}K(KR)(KR)[b] KRX_{10-12}K(KR)*X*(K/R)[b]

[a]Sequence representation is as follows: (^DE), any amino acid exept Asp or Glu; X_{10-12}, any 10-12 amino acids.

[b]For more optimal patterns, acidic residues should be rich in the central linker region and rare in the terminal linker region, whereas basic and hydrophobic residues should be rare in the central region and proline-rich in the terminal region.

Table 1. Consensus sequence of six classes of importin α-dependent NLS

2.3. Drug-protein interaction

Despite the introduction of newly developed drugs, such as lenalidomide and bortezomib, multiple myeloma is still difficult to treat and patients have a poor prognosis. In order to find novel drugs that are effective for multiple myeloma, we tested the antitumor activity of 29 phthalimide derivatives against several multiple myeloma cell lines. Among these derivatives, 2-(2,6-diisopropylphenyl)-5-amino-1H-isoindole-1,3-dione (TC11) was found to be a potent inhibitor of tumor cell proliferation and an inducer of apoptosis *via* activation of caspase-3, 8 and 9. This compound also showed *in vivo* activity against multiple myeloma cell line KMS34 tumor xenografts in ICR/SCID mice. To identify TC11-binding proteins, we used the IVV method [1-3,15]. We first prepared a cDNA library derived from KMS34 cells, because our data suggested that KMS34 cells were the most sensitive to TC11. As a bait, biotinylated TC11 was immobilized on a microfluidic chip and TC11-binding proteins were selected. Although the 4-amino group of TC11, which was experimentally inferred to be critical for the activity, was biotinylated *via* a linker, the biotinylation hardly affect the antitumor activity. Among 11 candidate TC11-binding proteins identified by the IVV method after 4 rounds of selection, we focused on the nucleolar phosphoprotein nucleophosmin (NPM). Sequencing revealed that three selected NPM clones, designated 1–183 NPM, encoded the 183 NH_2-terminal amino acids of NPM, which include the oligomerization domain and a part of the histone-binding domain. The enrichment efficiency of the NPM clones was confirmed to be 10^4-fold after 4 rounds of selection by RT-PCR. NPM is a multifunctional protein involved in both tumorigenesis and tumor suppression [41]; for example, it regulates cell proliferation and centrosome dupulication [42] and stabilizes oncoprotein Myc [43] and tumor-suppressor protein p53 [44]. Therefore, we hypothesized that NPM is involved in TC11-induced apoptosis of tumor cells. Immunofluorescence and NPM-knockdown studies in HeLa cells suggested that TC11 inhibits centrosomal clustering by inhibiting the centrosomal-regulatory function of NPM, thereby inducing multipolar mitotic cells, which undergo apoptosis. NPM may become a novel target for development of antitumor drugs active against multiple myeloma [28].

We screened 46 novel anilinoquinazoline derivatives for activity to inhibit proliferation of a panel of human cancer cell lines. Among them, Q15 showed potent *in vitro* growth-inhibito-

ry activity towards cancer cell lines derived from colorectal cancer, lung cancer and multiple myeloma. It also showed antitumor activity towards multiple myeloma KMS34 tumor xenografts in lcr/scid mice *in vivo*. Unlike the known anilinoquinazoline derivative gefitinib, Q15 did not inhibit cytokine-mediated intracellular tyrosine phosphorylation. To elucidate the mechanism through which Q15 inhibits proliferation of tumor cells, we set out to identify Q15-binding proteins by means of the IVV method [1-3,15]. We prepared a cDNA library derived from total RNA of human colon carcinoma SW480 cells, because, like other tumor cells, SW480 cells were sensitive to Q15. Proteins that bind to biotinylated Q15 immobilized on beads were selected using the IVV method. From the library obtained after 5 rounds of selection, we analyzed the DNA sequences of 100 clones. Among them, we obtained six clones of a fragment of the Luzp5/NCAPG2 gene encoding hCAP-G2$_{262\text{-}476}$ containing the HEAT (Huntingtin, elongation factor 3, a subunit of protein phosphatase 2A, TOR lipid kinase) repeat domain. Although three other clones were obtained redundantly, they were confirmed to be false-positive clones by means of binding assay (data not shown). hCAP-G2 is a subunit of condensin II complex [45,46], which is regarded as a key player in mitotic chromosome condensation [47]. Immunofluorescence study indicated that Q15 compromises normal segregation of chromosomes, and therefore might induce apoptosis. Thus, our results indicate that hCAP-G2 is a novel therapeutic target for development of drugs active against currently intractable neoplasms [29].

3. Application of the IVV method for comprehensive interactome network analyses

Interactome networks are essential for complete systems-level descriptions of cells. Large-scale protein-protein interactions (PPIs) are integral in the analysis of topological and dynamic features of interactome networks. We introduce large-scale interactome network analyses using a combination of the IVV method and a biorobot for 50 human transcription factors (TFs).

Comprehensive analysis of PPIs is an important task in the field of proteomics, functional genomics and systems biology. PPIs are usually analyzed by means of biochemical methods such as pull-down assay and co-immunoprecipitation, yeast two-hybrid (Y2H) assay [35] and phage display [36]. Recently, the combined use of mass spectrometry (MS) with an affinity tag [48] has made biochemical methods more comprehensive and reliable. However, the testable interaction conditions are restricted by the properties of the biological sources. The Y2H assay is one of the major tools used in the discovery and characterization of PPIs [49]. However, the results of Y2H analyses often include many false positives due to auto-activating bait or prey fusion proteins [50] and interactions of proteins that are toxic to yeast cells cannot be examined. Phage display, the most widely used display technology [51], is an effective alternative, because the interactions between libraries and target proteins occur *in vitro*, allowing optimal conditions to be used for many different target proteins. However, the detectability of very low copy number proteins by phage display is still limited, because phage libraries are produced in living bacteria [52]. Totally *in vitro* display technologies such as ribosome display [53], the IVV method [1-3,15] and DNA display [54] can circumvent the above difficulties,

http://dx.doi.org/10.5772/55483

because they do not need living cells. As a model bait protein, we chose the basic leucine zipper (bZIP) domain of Jun protein, an important transcription factor, to screen Jun interactors from a mouse brain cDNA library. By performing iterative affinity selection and sequence analyses, we selected 16 novel Jun-associated protein candidates in addition to four known interactors. By means of real-time PCR and pull-down assay, 10 of the 16 newly discovered candidates were confirmed to be direct interactors with Jun *in vitro*. Furthermore, interaction of 6 of the 10 proteins with Jun was observed in cultured cells by means of co-immunoprecipitation and observation of subcellular localization. These results demonstrate that this *in vitro* display technology is effective for the discovery of novel protein–protein interactions and can contribute to the comprehensive mapping of protein–protein interactions.

Furthermore, in a single experiment using bait Fos, more than 10 interactors, including not only direct, but also indirect interactions, were enriched. Further, previously unidentified proteins containing novel leucine zipper (L-ZIP) motifs with minimal binding sites identified by sequence alignment as functional elements were detected as a result of using a randomly primed cDNA library. Thus, we consider that this simple IVV selection system based on cell-free cotranslation could be applicable to high-throughput and comprehensive analysis of PPI and complexes in large-scale settings involving parallel bait proteins.

Interactome networks are essential for complete systems-level descriptions of cells. Large-scale PPIs are integral in the analysis of topological and dynamic features of interactome networks [55]. Several attempts to collect large-scale PPI data have been initiated using various model organisms [56-61] and were subsequently conducted in humans [62-64]. Traditionally, protein interaction data are collected using high-throughput *in vivo* expression tools based on the yeast two hybrid (Y2H; [49]) and tandem affinity purification-mass spectrometry (TAP-MS; [65]) methods. Experiments of this nature have provided large-scale PPI data, but they have only generated information on interacting partners, without considering binding domains in detail. In the field of systems biology, a further understanding of cellular networks will require more complete data sets describing the underlying physical interactions between cellular components. Thus, it is important to identify not only the binding partners, but also the interacting domains at the amino acid level [66]. In fact, the idea of mapping the interacting regions (IRs) involved in a PPI has been previously suggested in connection with several large-scale screens [67]. Our IVV method of analyzing PPIs [15,16] is well suited for large-scale, high-throughput mRNA display of the domain-based interactome using a randomly primed cDNA library, and we were able to achieve the first large-scale mapping of human IR data at the domain level for TF-related protein complexes. Functional domains were easily extracted based on the identified sequences using a randomly primed prey library as a non-biased representation [15]. Bait mRNA templates were prepared *in vitro* [15], and large-scale IVV method was performed using a biorobot that can simultaneously execute up to 96 selections. Fifty human TF-related proteins were used as bait, and a human brain cDNA library was used as prey. A modified high-throughput version of IVV selection was employed [15]. Integration of large-scale PPI data with other data sets, such as 3D structural information [68] and expression data [69], is necessary to identify the possible functions of interaction networks [68]. Large-scale IR data sets are expected to reflect functional domains and to indicate the biological roles of the

network without the need to integrate additional data. We confirmed the reliability and accuracy of our data by performing pull-down assays [15] and by examining the overlap between our results and known PPI domains using Pfam search [70]. The core data set (966 IRs; 943 PPIs) displayed a verification rate of 70%. Analysis of the IR data set revealed the existence of IRs that interact with multiple partners (Fig. 4). Furthermore, these IRs were preferentially associated with intrinsic disorder. This finding supports the hypothesis that intrinsically disordered regions play a major role in the dynamics and diversity of TF networks through their ability to structurally adapt to and bind with multiple partners. Accordingly, this domain-based interaction resource represents an important step in refining protein interactions and networks at the domain level and in associating network analysis with biological structure and function.

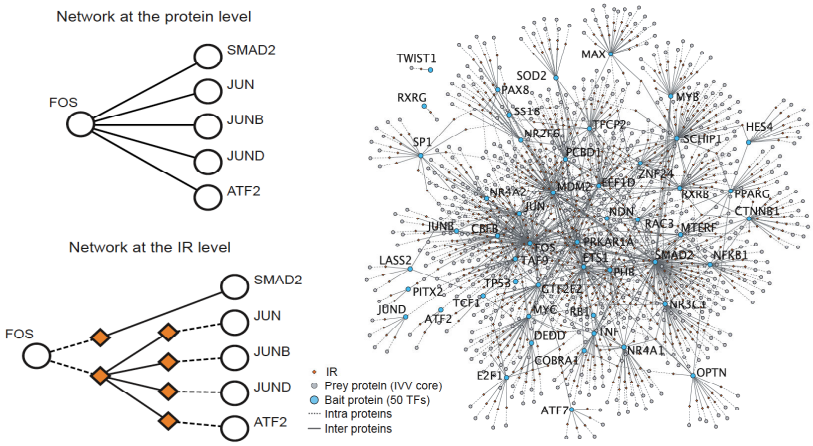

Figure 4. A transcription factor network at the interaction region(IR) level developed using IVV data. Graphic expression of the PPI network at the IR level(right). Interacting interfaces of the proteins, determined as IRs by IVV experiments, are drawn on the graph as diamond-shape nodes (IR nodes). Broken and solid lines indicate 'intra-' and 'inter-' protein edges, respectively. The graph contains 1,572 nodes (842 IR nodes and 730 protein nodes) and 842 intra-protein edges. Note that overlapping IRs are merged into a single node in the constructed network. An example of an underlying network graph at the IR level. Graphical expression of the FOS network at the protein level (upper left). PPIs are simply expressed by nodes indicating proteins and edges that connect them. Graphical expression of the FOS network at the IR level (lower left). A leucine zipper region of the FOS protein exclusively interacts with leucine zipper regions of other proteins (JUN, JUNB, JUND and ATF2). In addition, a region distinct from the leucine zipper in the FOS protein interacts with SMAD2.

4. Ultrahigh enrichment of antibodies by the IVV method on a microfluidic chip

Rapid preparation of monoclonal antibodies with high affinity and specificity is required in diverse fields from fundamental molecular and cellular biology to drug discovery and

http://dx.doi.org/10.5772/55483

diagnosis [71]. In addition to classical hybridoma technology, *in vitro* antibody-display technologies [30,53,72-75] are powerful approaches for isolating single chain Fv (scFv) antibodies from recombinant antibody libraries. However, these display techniques require several rounds of affinity selection (typically, the library size is 10^7~10^{12}, while the enrichment efficiency is 10~10^3-fold per round). Recently, microfluidic systems have been developed for high-throughput protein analysis [76], since they offer the advantages of very low sample volumes, rapid analysis, and automated recovery of captured analytes for further characterization. However, there have been few attempts to combine microfluidic systems with *in vitro* antibody-display technologies so far. We showed that a microfluidic system can be combined with the IVV method [1-3,15] and employed for scFv selection from naïve and randomized scFv libraries with ultrahigh efficiency of 10^6- to 10^8-fold per round [31].

4.1. *In vitro* selection of antibodies from a naïve scFv library

The IVV selection of scFv was performed on a Biacore microfluidic chip. Since the diversity of the mouse scFv library prepared from mouse spleen poly A^+ RNAs is estimated to be 10^6~10^8, while IVV method [15] allows screening of ~10^{12} molecules, we also introduced random point mutations into the scFv library. We chose p53 (human tumor suppressor protein) and MDM2 (human murine double minute) proteins as model antigens that were immobilized on the Biacore sensor chip. The selection experiment was performed on the microfluidic chip, and selected scFv genes were amplified by reverse transcription (RT)-PCR and identified by cloning and sequencing. Unexpectedly, the recovered anti-p53 and anti-MDM2 scFv sequences converged on a single sequence and two sequences, respectively, after only two rounds of selection. These clones showed high affinity, but also low antigen-specificity, in pull-down assays, and so we examined the clones obtained after a single round of selection in each case. When the binding activities of 29 (anti-p53) and 20 (anti-MDM2) clones with distinct sequences were examined by means of pull-down assays, P1-93 and M1-19 showed high specificity against the respective antigens among p53, MDM2 and BSA (Fig. 5A). The amino-acid sequences of P1-93 and M1-19 are shown in Fig. 5B. In competitive ELISA, both clones dose-dependently inhibited the ELISA signal (Fig. 5C), and Scatchard plots revealed that the K_Ds of P1-93 and M1-19 were 22 nM and 5.9 nM, respectively. The K_Ds of P1-93 and M1-19 were also determined by surface plasmon resonance (SPR) as 12 nM and 4.3 nM, respectively (Fig. 5D). The values obtained by the two different methods are similar.

4.2. *In vitro* evolution of scFv

Further, we performed *in vitro* evolution of scFv with higher affinity against MDM2 from a randomly mutated M1-19 scFv library. We applied on-rate or off-rate selection as a selection pressure for *in vitro* affinity maturation with the Biacore instrument: the on-rate selection was performed by controlling flow rate, and the off-rate selection was carried out by using a prolonged washing process on the sensor chip. After one round of selection, the recovered scFv genes were cloned and sequenced, and the K_Ds were evaluated by competitive ELISA (Figs. 6A and 6D). We obtained four mutants with higher affinity for MDM2 (K_D = 0.7-3.8 nM) than the progenitor M1-19 from 22 distinct clones. The strongest binder, M1-19a, was con-

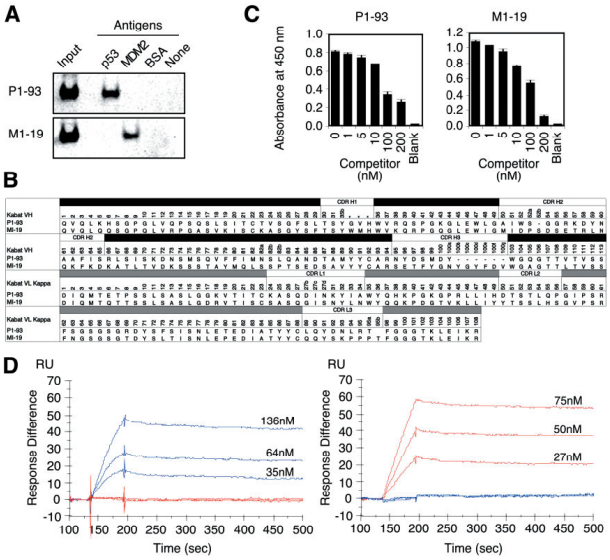

Figure 5. The selected scFvs anti-p53 P1-93 and anti-MDM2 M1-19. (A) Pull-down assays of the anti-p53 scFv P1-93 (top) and anti-MDM2 scFv M1-19 (bottom) using p53-, MDM2- or BSA-immobilized beads. Recovered scFv with FLAG-tag was detected with the anti-FLAG antibody. (B) Predicted amino-acid sequences of the V_H (black bar) and V_L (gray bar) regions of anti-p53 scFv P1-93 and anti-MDM2 scFv M1-19. (C) Competitive ELISA. P1-93 or M1-19 was pre-incubated with a competitor (0~200 nM free antigen) and allowed to bind to antigen-immobilized plates. After washing, remaining scFvs were detected with the anti-T7 tag antibody. (D) Biacore sensorgrams of the purified P1-93 (left: 31 kDa) and M1-19 (right: 32 kDa) using a p53- (blue lines: 55 kDa) or MDM2-immobilized (red lines: 66 kDa) sensor chip. The measurements were performed under conditions of 450 RU of the ligand and at a flow rate of 60 μl/min. To determine dissociation constants, three different concentrations (35, 64, and 136 nM for P1-93 and 27, 50, and 75 nM for M1-19) of the monomeric scFvs were injected.

firmed to have a higher on-rate and lower off-rate than M1-19 by SPR (Fig. 6B) and was also confirmed to recognize only the antigen protein MDM2 in crude cell lysates by Western blotting (Fig. 6C). These results indicated that the selected scFv had high enough affinity and specificity for practical use. Although the mutations of the selected scFvs were distributed among the whole sequences and no "consensus" mutations were identified, the mutation Y100bH within V_H CDR3 may contribute to the improved affinity and specificity, because this region is usually important for binding with antigens.

4.3. Ultrahigh efficiency of protein selection

Surprisingly, our results indicated that positive clone(s) were efficiently enriched through only one or two rounds of selection from a large library containing ~10^{12} molecules, implying ultrahigh efficiency of the method. To estimate the enrichment efficiency, we performed model experiments using a mixture of two kinds of scFv genes. The P1-93 (anti-p53) or M1-19 (anti-MDM2) gene was mixed with an anti-fluorescein scFv gene [30]('Flu' as a negative control) at a ratio of 1:10^2, 1:10^4, 1:10^6 or 1:10^8, and subjected to one round of the IVV selection on the sensor

http://dx.doi.org/10.5772/55483

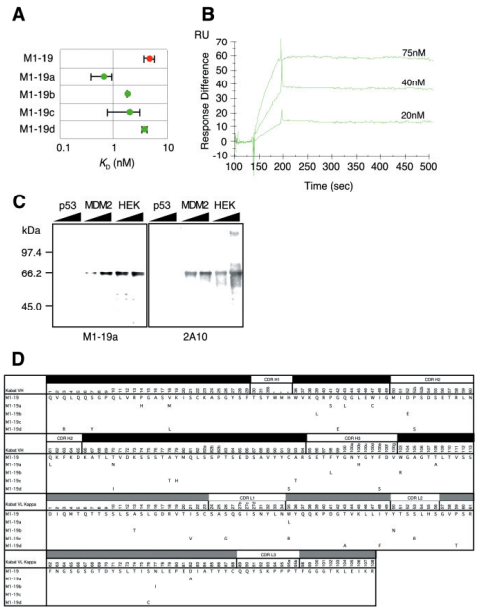

Figure 6. *In vitro* evolution of anti-MDM2 scFv M1-19. (A) Dissociation constants (K_Ds) of M1-19 (red circle) and the four mutant scFvs (green circles) were obtained by means of Scatchard plots of the competitive ELISA data. (B) Biacore sensorgrams of the mutant M1-19a (the amount of immobilized antigen was 450 RU). M1-19a had a higher on-rate and lower off-rate (k_a = 2.5 x 10^5/Ms, k_d = 8.6 x 10^5/s, K_A = 3.0 x 10^9/M, K_D = 0.34 nM) than the progenitor M1-19 (k_a = 5.5 x 10^4/Ms, k_d = 2.3 x 10^4/s, K_A = 2.4 x 10^8/M, K_D = 4.3 nM; see Fig. 5D: red lines on right side). (C) Recombinant p53 and MDM2 proteins (10 and 25 ng), and HEK-293 cell lysates (1 and 2.5 μg) were analyzed by Western blotting using M1-19a (left) or commercially available anti-MDM2 2A10 (right) as a control, respectively. (D) Predicted amino-acid sequences of the V_H (black bar) and V_L (gray bar) regions of the mutant scFvs M1-19a-d.

chip. The selection of the 1:10^6 mixture of P1-93:Flu genes and the 1:10^8 mixture of M1-19:Flu genes each resulted in a roughly 1:1 final gene ratio (Fig. 7A), indicating enrichment efficiencies of 10^6- and 10^8-fold per round, respectively. Furthermore, we confirmed that not only protein-protein (antigen-antibody) interactions, but also protein-DNA and protein-drug interactions were selected by our method with high enrichment efficiencies of >10^6-fold (Figs. 7B and 7C). Since the enrichment efficiencies of these model experiments with a usual agarose resin were only 10~10^3-fold per round, the enrichment efficiency was improved 10^3~10^5-fold over previous methods. Furthermore, we confirmed that the IVV system using a photocleavable linker between mRNA and protein is useful for *in vitro* selection of epitope peptides, recombinant antibodies, and drug-receptor interactions [77].

Although Biacore instruments have so far been utilized mainly to analyze biomolecular interactions by SPR, a few researchers have used this approach to fish for affinity targets from a randomized DNA library [78], phage-displayed protein libraries [79,80], or a ribosome-displayed antibody library [81]. However, the enrichment efficiency in these applications was not high. Why, then, was ultrahigh efficiency achieved in the present protein selection by IVV method? The IVV is a relatively small object pendant from its encoding RNA moiety, which

is about ten times larger. Thus, nonspecific adsorption of RNA on solid surfaces is potentially significant. The matrix of the Biacore sensor chip consists of carboxymethylated dextran covalently attached to a gold surface and poorly binds nucleic acid molecules, since both materials are negatively charged. In contrast, phage display and ribosome display involve large protein moieties (coat proteins or ribosome), so the use of the sensor chip may not improve the enrichment efficiency in these cases.

It should be noted that the ultrahigh enrichment efficiency made it difficult to set the number of selection rounds at a level that is appropriate to remove all non-binders as well as to pick all binders with various affinities from a library. If the number of selection rounds is too small, many negative sequences will be cloned; on the other hand, excess rounds of selection will yield only a single sequence with the highest affinity. In this study, we obtained 20-30 different sequences, including P1-93 and M1-19, with high antigen-specificity after a single round of selection, while we obtained only one or two negative sequences with high affinity but low antigen-specificity from the 10^6-10^8 library after two (probably excess) rounds of selection (>10^{12}-fold).

In summary, we achieved ultrahigh efficiencies (10^6~10^8-fold per round) of protein selection by IVV method with the microfluidic system. We obtained scFvs with high affinity and specificity from a naïve library by IVV selection for the first time. It took only three days to perform each selection experiment, including activity evaluation by ELISA. Although preparation of target materials of high quality is required, we anticipate this simple method to be a starting point for a versatile system to facilitate high-throughput preparation of monoclonal antibodies for analysis of proteome expression and detection of biomarkers, high-throughput

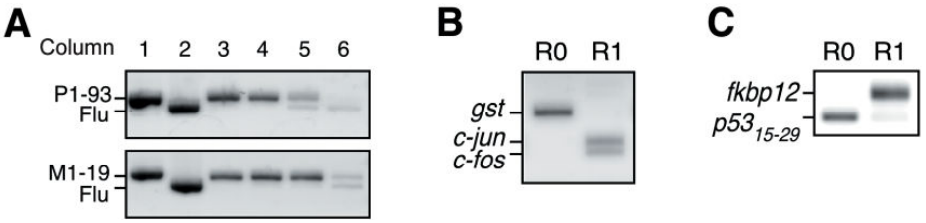

Figure 7. IVV selection of protein interactions on the Biacore sensor chip. (A) The P1-93 (912 bp: column 1 top) or M1-19 (936 bp: column 1 bottom) gene was mixed with an anti-fluorescein scFv gene [30](Flu: 888 bp: column 2) at a ratio of 1:10^2 (column 3), 1:10^4 (column 4), 1:10^6 (column 5) or 1:10^8 (column 6). The mixtures were subjected to IVV selection on the sensor chip conjugated with antigens p53 (top) or MDM2 (bottom). The RT-PCR products amplified from fractions after one round of selection were analyzed by agarose gel electrophoresis. (B) *In vitro* selection of protein-DNA interactions was performed using a mixture of three genes with N-terminal T7 tag and C-terminal FLAG-tag coding sequences: c-fos (349 bp), c-jun (394 bp), gst (597 bp). The template RNAs of c-fos, c-jun and gst (negative control) were mixed at a ratio of 1:1:10^6. The mixtures were translated and the resulting IVV libraries were selected on the sensor chip conjugated with bait DNA (AP-1)[20]. The RT-PCR products amplified from fractions before (R0) or after one round (R1) of selection were analyzed by agarose gel electrophoresis. (C) *In vitro* selection of protein-drug interactions was performed using a mixture of two genes with N-terminal T7 tag and C-terminal FLAG-tag coding sequences: fkbp12 (448 bp) and a p53 (15-29 aa) fragment (175 bp). The template RNAs of fkbp12 and the p53 fragment (negative control) were mixed at a ratio of 1:10^6. The mixtures were translated and the resulting IVV libraries were selected on the sensor chip conjugated with bait drug (FK506)[71]. The RT-PCR products amplified from fractions before (R0) or after one round (R1) of selection were analyzed by agarose gel electrophoresis.

http://dx.doi.org/10.5772/55483

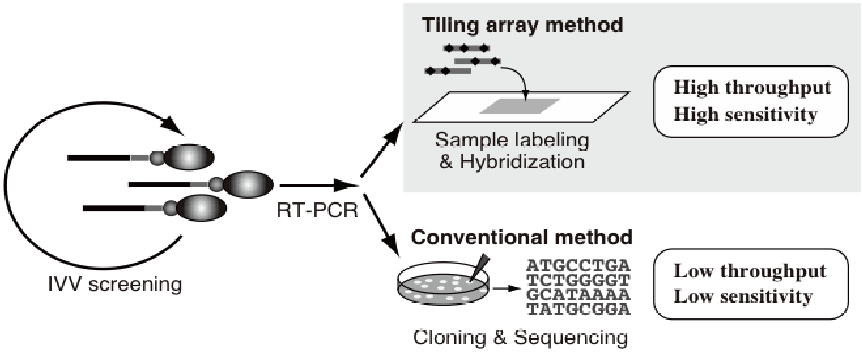

Figure 8. Schematic representation of the tiling array method and a conventional method for sequence determination of cDNAs from IVV screening.

analysis of protein-protein, protein-DNA and protein-drug interactions in proteomic and therapeutic fields, and rapid evolution of novel artificial proteins from large randomized libraries that often require ten or more rounds of selection.

5. Highly sensitive, high-throughput cDNA tiling arrays for detecting protein interactions selected by the IVV method

The most serious bottleneck in the IVV method has been in the final decoding step to identify the selected protein sequences. This step is usually achieved by cloning in bacteria and DNA sequencing using Sanger sequencers, but the following difficulties arise: 1) Only a limited number of clones can be analyzed, and thus positive candidates whose contents in the selected library are less than a threshold determined by the number of analyzed clones are lost as false negatives. 2) Positive sequences with low contents in a library can be enriched by iterative rounds of affinity selection, but lower-affinity binders compete with higher-affinity binders and therefore drop out of the screening. 3) DNA fragments which are injurious to cloning hosts, e.g., cytotoxic sequences, may be lost. 4) Cloning and sequencing of a huge number of copies of selected sequences is redundant, cost-ineffective, and time-consuming.

A DNA microarray is an efficient substitute for the cloning and sequencing processes to overcome the above limitations (Fig. 8). The combined use of a tiling array [82] representing ORF sequences with the IVV method would provide a completely *in vitro* platform for highly sensitive and parallel analysis of protein interactions. It should be possible to detect enrichment of cDNA fragments of selected candidates even with low contents or low affinity. However, for the analyses, the tiling arrays should be custom-designed specifically for the IVV screening, because the DNA fragments from the IVV method have unique characteristics, e.g., short length and functional region-concentrated distribution [16]. In this chapter, we introduce a

highly sensitive, high-throughput protein interaction analysis procedure combining the IVV method with tiling arrays.

First, we designed a custom oligo DNA microarray as follows: 1) Oligonucleotide probes of 50-mer in length were used. This is the preferred length for microarray probes, because shorter probes result in low sensitivity and longer probes produce non-specific signals [83]. 2) There should be no gaps between the probes. A contiguous linear series of data is required to recognize a signal peak in the algorithm for tiling array analysis, as described below, so the probes must be densely arranged. 3) mRNA sequences were employed for the tiling array. Only coding regions are required for the purpose of protein-interaction analysis, so other genomic sequences, e.g., introns, control regions and non-coding RNAs, were not employed.

Second, we also improved the method for labeling of cDNA samples. Usually, double-stranded DNA samples for a tiling array analysis are labeled by using random primers [84]. However, cDNA fragments selected from a randomly fragmented cDNA library [16] seem to be too short for efficient labeling by random priming. Indeed, in a test analysis with a tiling array using the random priming labeling method, we failed to detect any of the previously detected positive controls. Therefore we employed another labeling procedure [85], in which sense-strand-labeled

RNAs were produced by one-step *in vitro* transcription using a SP6 promoter attached to cDNA fragments from IVV screening.

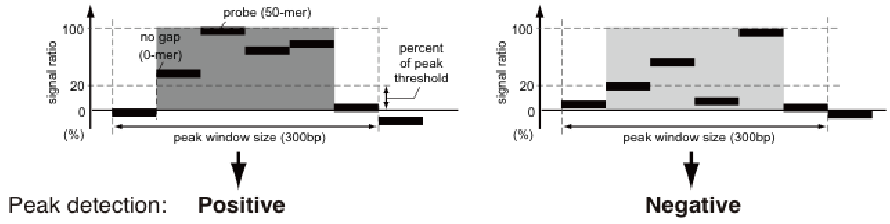

Figure 9. Windowed threshold detection algorithm for signal peak detection.

Third, we developed a detection algorithm for specific signal peaks from raw data. After iterative rounds of IVV screening, the resulting cDNA libraries in the presence and absence of bait protein, called bait (+) and bait (-) library, respectively, are labeled with fluorescence dyes using the above method, and hybridized separately. The ratios of the signal intensities from the experiments in the presence and absence of bait were calculated. Next, we searched for signal peaks in the data using the "windowed threshold detection" algorithm (Fig. 9). This algorithm looks for at least four data points that are above a threshold value within a window. These points were grouped together and presented as a peak. We used the following parameters in the algorithm: peak window size, 300 bp; percent of peak threshold, 20% of maximum data in each mRNA sequence. The value of

http://dx.doi.org/10.5772/55483

each peak was the maximum value of the data points in that peak. Only reproducible peaks in the duplicated data were collected as candidates.

As an actual model study, we performed protein-protein interaction screening for mouse Jun protein [86], a transcription factor containing a bZIP domain, using the combined IVV and tilling array method. For this study, we constructed a novel custom microarray containing ~1,600 ORF sequences of known and predicted mouse transcription-regulatory factors (334,372 oligonucleotides) [16,87,88] to analyze cDNA fragments from IVV screening for Jun-interactors, and named it the Transcription-Factor Tiling (TFT) array. From the 5th-round DNA library of the IVV screening in the presence and absence of a bait Jun protein, we obtained labeled RNAs and hybridized them onto the TFT array [89].

Positive signal peaks were collected using the windowed threshold detection algorithm; the total number of peaks was 647 on 545 mRNA sequences (some of the mRNA sequences included multiple peaks) [89]. An example is shown in Fig. 10. To distinguish between true positives and false positives, specific enrichment of the selected candidate was validated by real-time PCR. Among the top 10 percent of the peaks (64 regions), specific enrichment of 35 peaks was confirmed in the screening (white bars in Fig. 11A). The data indicate that the appropriate threshold for distinguishing between true positives and noise in the microarray signal is a signal ratio of 3~4. The 35 candidates identified in the present study include all of the 20 Jun-interactors identified in our previous studies using conventional cloning and sequencing [16,88]. Furthermore, the 35 candidates include eight well-known Jun-associated proteins, which is double the number in the previous study, in which four known Jun-interactors were obtained (white bars of Fig. 11B) [16,88,89]. In other words, 15 proteins including four known Jun-interactors were newly detected using the TFT arrays.

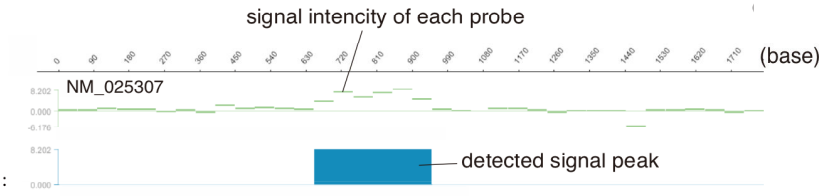

Figure 10. An example of probe signals and a detected signal peak on a gene.

Finally, we used *in vitro* pulldown assay and the surface plasmon resonance method to confirm the physical association of the 11 newly discovered candidates with Jun. As a result, ten of the 11 tested candidates exhibited specific interaction with the bZIP domain of Jun. Although most of the above tested interactions seem to be very weak, we considered that the interactions are true positives, because all of the candidates except for one contain leucine-heptad repeats in the selected regions, and such repeats are an important motif for heterodimerization with Jun [89].

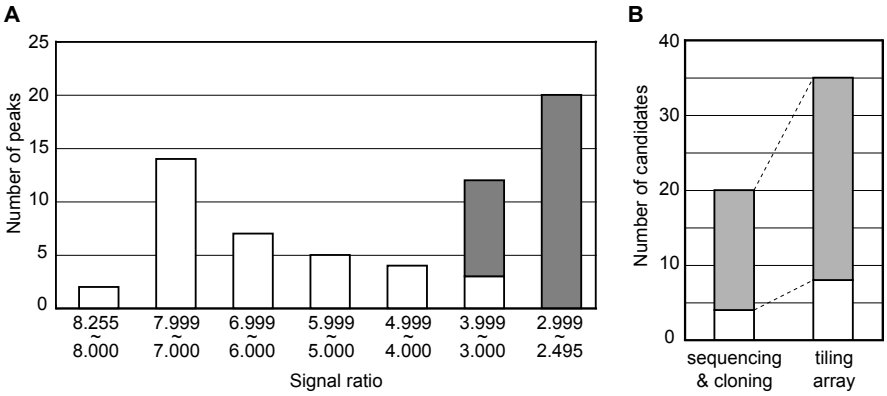

Figure 11. Data from the TFT array. (A) The top 10% of candidates were confirmed by real-time PCR. White and gray indicate numbers of enriched and non-enriched candidates, respectively. (B) Numbers of known (white) and newly selected (gray) proteins from conventional sequencing and the TFT arrays.

Previous studies and our survey revealed that the cDNA library used in this screening contained 29 known Jun-interactors [89,90]. Of these proteins, four (14%) and eight (28%) were detected by conventional sequencing and by the TFT array method, respectively. Thus, the TFT array method provides a remarkable increase in the number of identified interactors and this confirms the value of our new methodology as a screening tool for protein interactions. While the coverage was increased considerably, the accuracy did not decrease. Specifically, the number of false positives did not increase: the rates of confirmation of proteins by *in vitro* pull-down assays in the previous and present studies were 75% and 74%, respectively [90]. Undetected remaining interactors were considered to be false negatives. Mismatching of the selection conditions, e.g., salts, detergents, and pH, or the bait construct, e.g., length, region, and tags, might inhibit these interactions.

For quantitative analysis, the abundance ratios of 35 specifically selected candidates in the initial and screened cDNA libraries were determined by real-time PCR, and the enrichment rates (abundance ratio in the 5th round library per that in the initial library) were also calculated [88,89]. The abundance of the 15 newly found candidates (excluding four cases) was less than the theoretical threshold determined from the results of our previous study (an analysis of 451 clones). In order to detect the least abundant candidate (1.3×10^{-4}% of the screened cDNA library) by cloning and sequencing, it would have been necessary to analyze at least 1.0×10^{6} clones. These results indicate that our new method is more sensitive, higher-throughput and more cost-effective than the previous method.

From the standpoint of the detection sensitivity, the combined use of the IVV method with tiling arrays provides an extremely sensitive method for protein-interaction analysis, because even a very weakly expressed target could be detected in this study. In the cDNA library before IVV screening, the content of fragments of the selected region of the least abundant known Jun-binder was 1.2×10^{-7}%. If one mRNA molecule existed per cell, the content of a fragment of the gene would be about 1.2×10^{-5} to 5.9×10^{-5}% (we employed reported parameters for this calculation [91]). Thus, the content of the least abundant mRNA in the initial library corre-

http://dx.doi.org/10.5772/55483

sponds to about one molecule per 20 to 100 cells. This suggests that this gene is expressed at a very low level in a cell type that is a minor component of the tested tissue. It is noteworthy that targets expressed at such low levels can be detected without the need for a cell purification procedure, e.g., collection of somatic stem cells by flow cytometry. The high sensitivity of our method may allow access to targets which would be hard to analyze with other existing tools, such as the TAP method [48].

In summary, we have applied tiling array technology, which has previously been used for ChIP-chip assays and transcriptome analyses, to protein-interaction analysis with the IVV method. Compared with previous results obtained with cloning and sequencing, the use of the tiling array greatly increased sensitivity. This method can detect targets expressed at extremely low levels. This highly sensitive and reliable method has the potential to be used widely, because the tiling array approach can easily be extended to a genome-wide scale, even though the search space is limited in tiled sequences.

6. Conclusion

We have developed an mRNA display technology, named the *in vitro* virus (IVV) method, as a stable and efficient tool for analyzing various protein functions. The IVV method is applicable for exploring protein complexes, transcription factors, RNA-binding proteins, bioactive peptides, drug-target proteins and antibodies, as well as *in vitro* protein evolution from random-sequence and block-shuffling libraries. We further developed a large-scale and high-throughput IVV screening system utilizing a biorobot, microfluidic tip, and tiling array. Here we reviewed applications of the IVV method for protein functional analyses.

Acknowledgements

We thank Nobuhide Doi, Etsuko Miyamoto-Sato, Hideaki Takashima, Shunichi Kosugi, Masamichi Ishizaka, Toru Tsuji, Nobutaka Matsumura, Seiji Tateyama, Shigeo Fujimori, Hironobu Kimura, Isao Fukuda, Hirokazu Shiheido, Ichigo Hayakawa, Hiroaki Genma, Mayuko Tokunaga, Yuko Yonemura-Sakuma, Yoko Ogawa, Kazuyo Masuoka, Naoya Hirai and Masako Hasebe who are collaborators in the studies introduced in this review. We also thank Profs. Hideyuki Saya, Yutaka Hattori, Masaru Tomita and Taketo Yamada, and Drs Tatsuya Hirano and Takao Ono for valuable advice and discussion.

Author details

Noriko Tabata, Kenichi Horisawa and Hiroshi Yanagawa

Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, Yokohama, Japan

References

[1] Nemoto N, Miyamoto-Sato E, Husimi Y, Yanagawa H. *In vitro* virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome *in vitro*. FEBS Lett 1997;414: 405-408.

[2] Miyamoto-Sato E, Nemoto N, Kobayashi K, Yanagawa H. Specific bonding of puromycin to full-length protein at the C-terminus. Nucleic Acids Res 2000;28: 1176-1182.

[3] Miyamoto-Sato E, Takashima H, Fuse S, Sue K, Ishizaka M, et al. Highly stable and efficient mRNA templates for mRNA-protein fusions and C-terminally labeled proteins. Nucleic Acids Res 2003;31: e78.

[4] Nemoto N, Miyamoto-Sato E, Yanagawa H. Fluorescence labeling of the C-terminus of proteins with a puromycin analogue in cell-free translation systems. FEBS Lett 1999; 462: 43-46.

[5] Doi N, Takashima H, Kinjo M, Sakata K, Kawahashi Y et al. Novel fluorescence labeling and high-throughput assay technologies for *in vitro* analysis of protein interactions. Genome Res 2002;12: 487-492.

[6] Yarmolinsky MB, De La Hara,G.L. Inhibition by puromycin of amino acid incorporation into protein. Proc Natl Acad Sci USA 1959;45:1721-1729.

[7] Takeda Y, Hayashi S, Nakagawa H, Suzuki F. The effect of puromycin on ribonucleic acid and protein synthesis. J Biochem 1960;48:169-177.

[8] Nemeth AM, de la Haba GL. The effect of puromycin on the developmental and adaptive formation of tryptophan pyrrolase. J Biol Chem 1962;237:1190-1193.

[9] Nathans D. Puromycin inhibition of protein synthesis: Incorporation of puromycin into peptide chains. Proc Natl Acad Sci USA 1964;51:585-592.

[10] Traut RR, Monro RE. The puromycin reaction and itsw reaction to protein synthesis. J Mol Biol 1964;10:63-72.

[11] Steiner G., Kuechler E, Barta A. Photo-affinity labelling at the peptidyl transferase centre reveals two different positions for the A- and P-sites in domain V of 23S rRNA. EMBO J 1988;7:3949-3955.

[12] Kudlicki W, Odom OW, Kramer G, Hardesty B. Activation and release of enzymatically inactive, full-length rhodanese that is bound to ribosomes as peptidyl-tRNA. J Biol Chem 1994;269:16549-16553.

[13] Wolin SL, Walter P. Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J 1988;7:3559-3569.

[14] Bjornsson A, Isaksson LA. Accumulation of a mRNA decay intermediate by ribosomal pausing at a stop codon. Nucleic Acids Res 1996;24:1753-1757.

[15] Miyamoto-Sato E, Ishizaka M, Horisawa K, Tateyama S, Takashima H, et al. Cell-free cotranslation and selection using *in vitro* virus for high-throughput analysis of protein-protein interactions and complexes. Genome Res 2005;15:710-717.

[16] Horisawa K, Tateyama S, Ishizaka M, Matsumura N, Takashima H et al.: *In vitro* selection of Jun-associated proteins using mRNA display. Nucleic Acids Res 2004;32: e169.

[17] Kawahashi Y, Doi N, Takashima H, Tsuda C, Oishi Y et al. *In vitro* protein microarrays for detecting protein-protein interactions: Application of a new method for fluorescence labeling of proteins. Proteomics 2003;3:1236-1243.

[18] Miyamoto-Sato E, Ishizaka M, Fujimori S, Hirai N, Masuoka K et al. A comprehensive resource of interacting protein regions for refining human transcription factor networks: Domain-based interactome. PLoS ONE 2010;5: e9289.

[19] Suzuki H, Alistair RRF, Nimwegen E, Daub CO, Balwierz PJ. et al. The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet 2009;41:553-562.

[20] Tateyama, S, Horisawa, K, Takashima, H, Miyamoto-Sato, E, Doi, N et al. Affinity selection of DNA-binding protein complexes using mRNA display. Nucleic Acids Res 2006;34: e27.

[21] Horisawa K, Imai T, Okano H, Yanagawa H. 3′-Untranslated region of doublecortin mRNA is a binding target of the Musashi1 RNA-binding protein. FEBS Lett 2009;583:2429-2434.

[22] Horisawa K, Imai T, Okano H, Yanagawa H. The Musashi family RNA-binding proteins in stem cells. Biomol Concepts 2010;1:59-66.

[23] Shiheido H, Takashima H, Doi N, Yanagawa H. mRNA display selection of an optimized MDM2-binding peptide that potently inhibits MDM2-p53 interaction. PLoS ONE 2011;6:e17898.

[24] Matsumura N, Tsuji T, Sumida T, Kokubo M, Onimaru M et al. mRNA display selection of a high-affinity Bcl-X_L-specific binding peptide. FASEB J 2010;24:2201-2210.

[25] Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E et al. Six classes of nuclear localization signals specific to different binding grooves of importinα. J Biol Chem 2009;284:478-485.

[26] Kosugi S, Hasebe M, Tomita M,Yanagawa H. Systematic identification of cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. Pro Natl Acad Sci USA 2009;106: 10171-10176.

[27] Kosugi S, Hasebe M, Entani T, Takayama S, Tomita M et al. Design of peptide inhibitors for importin α/β nuclear import pathway by activity-based profiling. Chem & Biol 2008;15:940-949.

[28] Shiheido H, Terada F, Tabata N, Hayakawa I, Matsumura N, et al. A phthalimide derivative that inhibits centrosomal clustering is effective on multiple myeloma. PLoS ONE 2012;7,e38878.

[29] Shiheido H, Naito Y, Kimura H, Genma H, Takashima H. An anilinoquinazoline derivative inhibits tumor growth through interaction with hCAP-G2, a subunit of condensing II. PLOS ONE 2012;7:e44889.

[30] Fukuda I, Kojoh K, Tabata N, Doi N, Takashima H et al. *In vitro* evolution of single-chain antibodies using mRNA display. Nucleic Acids Res 2006;34: e127.

[31] Tabata N, Sakuma Y, Honda Y, Doi N, Takashima H. et al. Rapid antibody selection by mRNA display on a microfluidic chip. Nucleic Acids Res 2009;37: e64.

[32] Pollack JR, Iyer VR. Characterizing the physical genome. Nature Genet 2002;32:515–521.

[33] Yu H, Luscombe NM, Quian J, Gerstein M. Genomic analysis of gene expression relationships in transcriptional regulatory networks. Trends Genet 2003;19: 422–427.

[34] Barski A, Frenkel B. ChIP display: novel method for identification of genomic targets of transcription factors.Nucleic Acids Res 2004;32, e104.

[35] Luo Y, Vijaychander S, Stile J, Zhu,L. Cloning and analysis of DNA-binding proteins by yeast one-hybrid and one-two-hybrid systems. Biotechniques 1996;20:564–568.

[36] Hagiwara H, Kunihiro S, Nakajima K, Sano M, Masaki H. et al. Affinity selection of DNA-binding proteins from yeast genomic DNA libraries by improved phage display vector. J Biochem 2002;132:975–982.

[37] Borghouts C, Kunz C, Groner B. Current strategies for the development of peptide-based anti-cancer therapeutics. J Peptide Sci 2005;11: 713–726.

[38] Bottger V, Bottger A, Howard SF, Picksley SM, Chene P. et al. Identification of novel mdm2 binding peptides by phage display. Oncogene 1996;13: 2141–2147.

[39] Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F. et al. *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. Science 2004;303: 844–848.

[40] Hu B, Gilkes DM, Chen J. Efficient p53 activation and apoptosis by simultaneous disruption of binding to MDM2 and MDMX. Cancer Res 2007;67:8810–8817.

[41] Grisendi S, Mecucci C, Falini B, Pandolfi PP. Nucleophosmin and cancer. Nat Rev Cancer 2006;6: 493–505.

[42] Okuda M. The role of nucleophosmin in centrosome duplication. Oncogene 2002;21: 6170–6174.

[43] Li Z, Boone D, Hann SR. Nucleophosmin interacts directly with c-Myc and controls c-Myc-induced hyperproliferation and transformation. Proc Natl Acad Sci USA 2008;105:18794–18799.

[44] Colombo E, Marine JC, Danovi D, Falini B, Giuseppe P. Nucleophosmin regulates the stability and transcriptional activity of p53. Nat Cell Biol 2002;4: 529–533.

[45] Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, et al. Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell 2003;115: 109–121.

[46] Hirano T. Condensins: Organizing and segregating the genome. Curr Biol 2005;15:R265–275.

[47] Hudson DF, Marshall KM, Earnshaw WC. Condensin: Architect of mitotic chromosomes. Chromosome Res 2009;17:131–144.

[48] Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E. et al. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods 2001;24:218-229.

[49] Fields S, Song O. A novel genetic system to detect protein-protein interactions. Nature 1989;340:245-246.

[50] Vidalain PO, Boxem M, Ge H, Li S, Vidal, M. Increasing specificity in high-throughput yeast two-hybrid experiments. Methods 2004;32:363-370.

[51] Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 1985;228:1315-1317.

[52] Bradbury AR, Marks JD. Antibodies from phage antibody libraries. J Immunol Methods 2004;290: 29-49.

[53] Hanes J, Pluckthun A. *In vitro* selection and evolution of functional proteins by using ribosome display. Proc Natl Acad Sci USA 1997; 94:4937-4942.

[54] Yonezawa M, Doi N, Kawahashi Y, Higashinakagawa T, Yanagawa, H. DNA display for *in vitro* selection of diverse peptide libraries. Nucleic Acids Res 2003;31: e118.

[55] Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. Nature 2001;411: 41-42.

[56] Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS. et al. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature 2000;403:623-627.

[57] Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M. et al. A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc Natl Acad Sci USA 2001;98:4569-4574.

[58] Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 2002;415:141-147.

[59] Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B. et al. A protein interaction map of *Drosophila melanogaster*. Science 2003;302:1727-1736.

[60] Li S, Armstrong CM, Bertin N, Ge H, Milstein S. et al. A map of the interactome network of the metazoan *C. elegans*. Science 2004;303:540-543.

[61] Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, et al. Interaction network containing conserved and essential protein complexes in *Escherichia coli*. Nature 2005;433:531-537.

[62] Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A. et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature 2005;437:1173-1178.

[63] Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH. et al. A human protein-protein interaction network: a resource for annotating the proteome. Cell 2005;122:957-968.

[64] Ewing RM, Chu P, Elisma F, Li H, Taylor P, et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 2007;3: 89.

[65] Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M. et al. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol 1999;17:1030-1032.

[66] Hakes L, Pinney JW, Robertson DL, Lovell SC. Protein-protein interaction networks and biology--what's the connection? Nat Biotechnol 2008;26:69-72.

[67] Fromont-Racine M, Rain JC, Legrain P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nat Genet 1997;16:277-282.

[68] Kim PM, Lu LJ, Xia Y, Gerstein MB. Relating three-dimensional structures to protein networks provides evolutionary insights. Science 2006;314: 1938-1941.

[69] Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF. et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 2004;430:88-93.

[70] Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V. et al. Pfam: clans, web tools and services. Nucleic Acids Res 2006;34:D247-251.

[71] Hoogenboom, H.R. Selecting and screening recombinant antibody libraries. Nat Biotechnol 2005;23:1105-1116.

[72] Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. J. Mol. Biol 1991;222:581-597.

[73] He M, Taussig, M.J. Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for *in vitro* display and evolution of antibody combining sites. Nucleic Acids Res 1997;25:5132-5134.

[74] Reiersen H, Lobersli I, Loset GA, Hvattum E, Simonsen B. et al. Covalent antibody display--an *in vitro* antibody-DNA library selection system. Nucleic Acids Res 2005;33:e10.

[75] Boder ET, Midelfort KS,Wittrup KD. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. Proc Natl Acad Sci USA 2000;97:10701-10705.

[76] Lion N, Rohner TC, Dayon L, Arnaud IL, Damoc E. et al. Microfluidic systems in proteomics. Electrophoresis 2003;24:3533-3562.

[77] Doi N, Takashima H, Wada A, Oishi Y, Nagano T. et al. Photocleavable linkage between genotype and phenotype for rapid and efficient recovery of nucleic acids encoding affinity-selected proteins. J Biotechnol 2007;131:231-239.

[78] Hao D, Ohme-Takagi M, Yamasaki K. A modified sensor chip for surface plasmon resonance enables a rapid determination of sequence specificity of DNA-binding proteins. FEBS Lett 2003;536:151-156.

[79] Yamamoto Y, Tsutsumi Y, Yoshioka Y, Nishibata T, Kobayashi K. et al. (2003) Site-specific PEGylation of a lysine-deficient TNF-alpha with full bioactivity. Nat Biotechnol 2003;21:546-552.

[80] Malmborg AC, Dueñas M, Ohlin M, Söderlind E, Borrebaeck CA. Selection of binders from phage displayed antibody libraries using the BIAcore™ biosensor. J. Immunol. Methods 1996;198:51-57.

[81] Yuan Q, Wang Z, Nian S, Yin Y, Chen G. et al. Screening of high-affinity scFvs from a ribosome displayed library using BIAcore biosensor. Appl. Biochem. Biotechnol 2009;152:224-234.

[82] Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, et al. Large-scale transcriptional activity in chromosomes 21 and 22. Science 2002;296(5569): 916-919.

[83] Nuwaysir EF, Huang W, Albert TJ, Singh J, Nuwaysir K. et al. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. Genome Res 2002 ;12:1749-1755.

[84] Kim TH, Barrera LO, Zheng M, Qu C, Singer MA. et al. A high-resolution map of active promoters in the human genome. Nature 2005;436:876-880.

[85] 't Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT. Fluorescent labelling of cRNA for microarray applications. Nucleic Acids Res 2003;31: e20.

[86] Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK. et al. Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. Science 1997;238:1386-1392.

[87] Gunji W, Kai T, Sameshima E, Iizuka N, Katagi H. et al. Global analysis of the expression patterns of transcriptional regulatory factors in formation of embryoid bodies

using sensitive oligonucleotide microarray systems. Biochem Biophys Res Commun 2004;325:265-275.

[88] Horisawa K, Doi N, Takashima H, Yanagawa H. Application of quantitative real-time PCR for monitoring the process of enrichment of clones on *in vitro* protein selection. J Biochem 2005;137:121-124.

[89] Horisawa K, Doi N, Yanagawa H. Use of cDNA tiling arrays for identifying protein interactions selected by *in vitro* display technologies. PLoS ONE. 2008;3:e1646.

[90] Chinenov Y, Kerppola TK Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. Oncogene 2001;20:2438-2452.

[91] Campbell NA. Biology, 4th Ed. New York: The Benjamin/Cummings Publishing; 1996