

Mapping a Disordered Portion of the Brz2001-Binding Site on a Plant Monooxygenase, DWARF4, Using a Quartz-Crystal Microbalance Biosensor-Based T7 Phage Display

Yoichi Takakusagi,* Daisuke Manita, Tomoe Kusayanagi, Jesus Izaguirre-Carbonell, Kaori Takakusagi, Kouji Kuramochi, Kazuki Iwabata, Yoshihiro Kanai, Kengo Sakaguchi, Fumio Sugawara

Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, Chiba, Japan.
*Present address: National Cancer Institute, National Institutes of Health, Bethesda, Maryland.

ABSTRACT

In small-molecule/protein interaction studies, technical difficulties such as low solubility of small molecules or low abundance of protein samples often restrict the progress of research. Here, we describe a quartz-crystal microbalance (QCM) biosensor-based T7 phage display in combination use with a receptor–ligand contacts (RELIC) bioinformatics server for application in a plant Brz2001/DWARF4 system. Brz2001 is a brassinosteroid biosynthesis inhibitor in the less-soluble triazole series of compounds that targets DWARF4, a cytochrome P450 (Cyp₄₅₀) monooxygenase containing heme and iron. Using a Brz2001 derivative that has higher solubility in 70% EtOH and forms a self-assembled monolayer on gold electrode, we selected 34 Brz2001-recognizing peptides from a 15-mer T7 phage-displayed random peptide library using a total of four sets of one-cycle biopanning. The RELIC/MOTIF program revealed continuous and discontinuous short motifs conserved within the 34 Brz2001-selected 15-mer peptide sequences, indicating the increase of information content for Brz2001 recognition. Furthermore, an analysis of similarity between the 34 peptides and the amino-acid sequence of DWARF4 using the RELIC/MATCH program generated a similarity plot and a cluster diagram of the amino-acid sequence. Both of these data highlighted an internally located disordered portion of a catalytic site on DWARF4, indicating that this portion is essential for Brz2001 recognition. A similar trend was also noted by an analysis using another 26 Brz2001-selected peptides, and not observed using the 27 gold electrode-recognizing control peptides, demonstrating the reproducibility and specificity of this method. Thus, this affinity-based strategy enables high-throughput detection of the small-molecule-

recognizing portion on the target protein, which overcomes technical difficulties such as sample solubility or preparation that occur when conventional methods are used.

INTRODUCTION

Identification of small-molecule binding sites on protein is an essential task in medicinal chemistry, chemical biology, and pharmacology. This information serves as a molecular basis for understanding not only the mode of action for the small molecule of interest, but also the biological phenomena associated with the small molecule. While recent proteomics approaches have enabled comprehensive identification of small-molecule/protein interactions in theory, physicochemical properties of experimental samples may restrict their interactions. These are divided mainly into two aspects of problems imposed by those of (1) small molecules and (2) proteins. The former can include low solubility in solvent, low levels of detection on analytical methods due to the small molecular mass of compounds, and inaccessibility of the immobilized small molecular size of compounds to internally located binding sites on larger target proteins. The latter can include the requirement for large amounts of the protein, determination of co-crystallization conditions, molecular size restrictions of the analytical method, and low solubility in solvents that is mainly due to protein misfolding when genetically engineered. Thus, research on undruggable proteome that is a result of these technical limitations should be largely continued.¹

A biosensor-based screening of T7 phage-displayed random peptide is a useful strategy for detecting small-molecule/protein interactions, which could circumvent those technical limitations.^{2–6} This approach employs a high-sensitivity quartz-crystal microbalance (QCM) biosensor that can monitor the decrease of alternating voltage-induced intrinsic crystal vibrations in real time, which indicates a mass increase in the gold electrode deposited on the crystal of the QCM sensor chip (Fig. 1A). The discovery of QCM phenomena in the air phase, which is described as the Sauerbrey's equation, goes back to the original report in 1959.⁷ Owing to the reiterated physical and engineering refinements over the past half-century, high-resolution QCM biosensors have been developed. They make it possible to weigh the pico- to nanogram scales of bound

ABBREVIATIONS: IR, infrared spectra; QCM, quartz-crystal microbalance; RELIC, receptor–ligand contacts; SAM, self-assembled monolayer; SDS, sodium dodecyl sulfate.

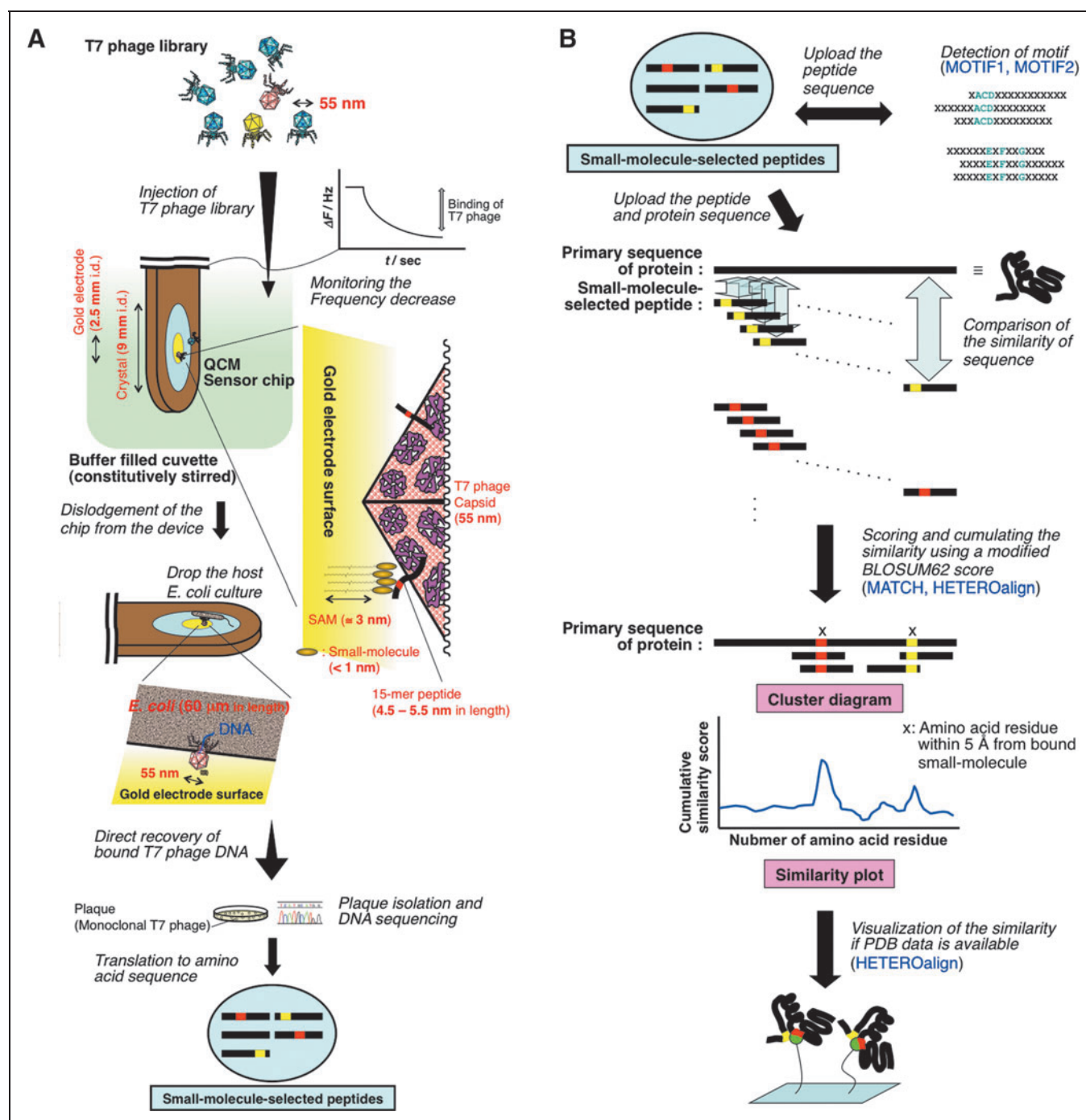


Fig. 1. A schematic representation of a quartz-crystal microbalance (QCM) biosensor-based T7 phage display and an analysis of the small-molecule-recognized peptide sequence using bioinformatics program in a receptor–ligand contacts (RELIC) suite. **(A)** One-cycle biopanning of a small-molecule-recognizing peptide sequence on a T7-phage capsid. The T7-phage library is injected into the buffer filled cuvette, and the frequency decrease on binding of the T7 phage to the small molecule on gold is monitored in real time. The 15-mer random peptide (4.5–5.5 nm in length) on the T7-phage capsid (55 nm) can sterically contact with any small molecules (<1 nm) immobilized on the gold electrode. The bound T7-phage DNA is directly recovered by host *Escherichia coli* infection. After plaque isolation, a part of the peptide-encoding region is sequenced according to the general procedure. **(B)** Illustration of the algorithmic analysis using the MOTIF, MATCH, or HETEROalign program in RELIC suite. The continuous (MOTIF1) or discontinuous (MOTIF2) conserved motif in the small-molecule-selected peptides is detected to evaluate the success of the selection. The subset of short peptides is then used to MATCH or HETEROalign (if PDB data is available) to calculate the weak similarity between the short peptide sequences and a primary protein sequence. The score is calculated by a modified BLOSUM62 amino-acid matrix within each window of five amino acids in length across the entire length of the protein sequence and then cumulatively plotted as a similarity plot. The mapping of the amino-acid sequence is generated at the portion of maximal similarity score.

molecules even in aqueous solutions, and have been widely applied to studies in life and material sciences.^{8–10} The use of this platform in a T7 phage display enabled a rapid capture of small-molecule recognizing T7 phages on the gold by sub-minutes of monitoring the frequency decrease, with no need of repeated rounds of biopanning.^{3,4} This also corresponds to the fact that the QCM-based selection eliminated hitherto inevitable problems associated with phage distribution bias during the amplification process on classical methods. Furthermore, subsequent use of a host *Escherichia coli* culture realized a rapid and direct recovery of DNA from the trace amount of bound T7 phage, without elute condition exploration. By sequencing the peptide-encoding region of phage DNA, the corresponding amino-acid sequence can be clearly determined. In other words, the T7 phage particle plays a role as a transducer for the highly sensitive biosensor to sequence the trace amount of peptide bound to the small molecule on the gold electrode. Significantly, unlike proteins, the use of random shorter peptides can allow sterical contact with any compounds of small molecular size fixed on the gold electrode (Fig. 1A).^{3,4}

In 2004, Rodi and Makowski and their coworkers launched the receptor–ligand contacts (RELIC) bioinformatics server, and they have since extensively studied the annotation of a small-molecule binding site on protein by a similarity search using affinity-selected random peptides.^{11–16} According to their earlier studies, phage-displayed short peptides, most of which are basically disordered, can mimic the disordered loops in naturally occurring proteins that are considered an essential portion in many small-molecule/protein interactions.^{17–19} Indeed, there are several examples of ligands that occupy sites on proteins which are not the classic “substrate binding site” as well as cases where analogous inhibitors do not bind analogously.^{20–22} Based on these observations, they have established a suite of bioinformatics programs designed for random peptide phage display. Once the amino-acid sequence of affinity-selected peptides is determined, the use of an appropriate program in RELIC with the sequence allows not only the qualitative and quantitative assessment of the affinity-selected peptide population, but also the annotation of small-molecule binding sites on protein that is de-

pendent on the weak similarity with the limited short peptide sequence data (Fig. 1B).^{11,13,23,24}

In this work, the QCM-based one-cycle biopanning in combination use with RELIC bioinformatics analysis has been extended into plant molecular biology.^{3,4} Brz2001 (Fig. 2A) is a specific DWARF4 inhibitor in plant hormone brassinosteroid biosynthesis.^{25,26} Brz2001 belongs to the triazole series of compound that shows low solubility in solvents, as represented by antifungal agents such as fluconazole. In addition, recent studies have revealed that the target protein DWARF4 is a C22 hydrolase which is classified into plant cytochrome P450 (Cyp₄₅₀) monooxygenase super family and contains heme and iron in an interior location of the enzyme where it catalyzes the hydroxylation reaction.²⁷ This fact implies that the DWARF4 protein requires these cofactors for folding into a functionally correct three-dimensional structure when genetically engineered. Furthermore, it may be also challenging to obtain the amount of purified plant enzyme for interaction analysis by NMR experiment or X-ray crystallography. Using this plant Brz2001/DWARF4 system, which may be time consuming and labor-intensive to refine the direct interaction properties on classical methods, we show that this rapid approach reveals disordered loops within catalytic sites on DWARF4 as an essential portion for Brz2001 recognition. This affinity-based strategy could be applicable for any small-molecule/protein systems under the identical protocol, overcoming the technical difficulties of conventional methods.

MATERIALS AND METHODS

QCM Apparatus and Reagents

The 27-MHz QCM apparatus (AffinixQ) and ceramic sensor chip (SiO₂, 0.06 mm thick, 9 mm in diameter, 64 mm²; Au, 0.1 mm thick, 2.5 mm in diameter, 4.9 mm²) were purchased from Initium Inc. T7 select 10-3 OrientExpressTM cDNA Cloning System was from Novagen. The oligonucleotide and primer were purchased from InvitrogenTM Custom Oligonucleotides. Klenow DNA polymerase I and ExoSAP-IT were purchased from USB Corporation. Ex Taq DNA polymerase was obtained from TaKaRa. ABI PRISM[®] BigDyeTM Terminator Cycle Sequencing Kit was from Applied Biosystems.

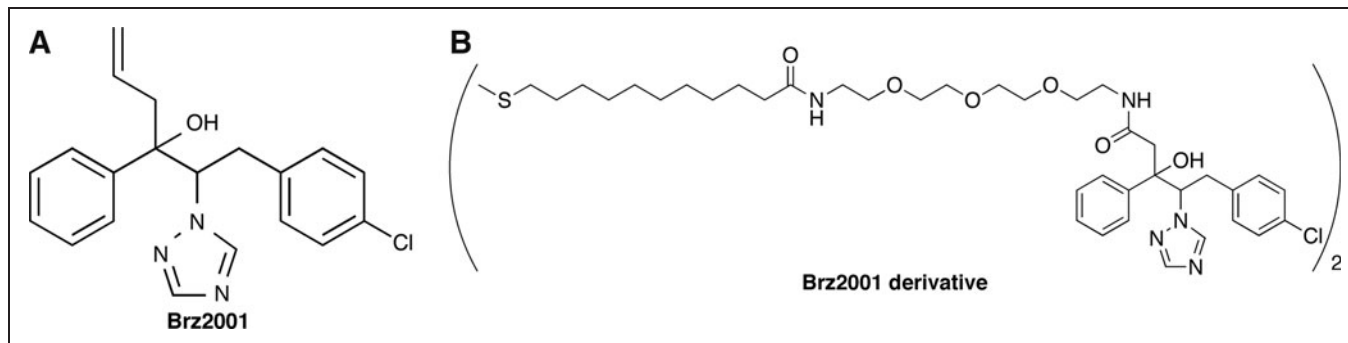


Fig. 2. Structure of (A) Brz2001 and (B) a Brz2001 derivative that forms a self-assembled monolayer (SAM) on a gold surface. The molecule comprises four units: (i) sulfide bond for chemisorption on the gold electrode surface of the sensor chip via thiol–Au interactions; (ii) a C11 alkyl chain for accumulation of the Brz2001 on the gold; (iii) diethylene glycol (DEG) for reducing the nonspecific binding (mainly via hydrophobic interaction) with the gold surface or linkers; and (iv) Brz2001 to act as bait during biopanning.

Construction of the T7 Phage-Displayed Random Peptide Library

For the preparation of a duplex DNA library, oligonucleotide GGG GAT CCG AAT TCT (NNK)₁₅ TGA AAG CTT CTC GAG GG (0.056 pM) and CCC TCG AGA AGC TTT CA (0.56 pM) were mixed with Klenow buffer, heated to 95°C for 5 min, and annealed by slowly cooling the mixture to 37°C. The single-stranded regions were converted to duplex DNA by continuing the incubation at 37°C for 2 h in the presence of dNTPs (2.5 mM) and Klenow enzyme (0.5 mU/μL). After the reaction, double-stranded DNA was recovered by EtOH precipitation. The obtained DNA was then digested separately with *EcoRI* and *HindIII*

restriction enzyme and inserted into the T7 select 10-3b vector according to the manufacturer's instructions. The primary titer of this T7 phage library was 1.6×10^7 pfu/mL. For the screening procedure, the phage library was amplified up to 1.7×10^{10} pfu/mL using *E. coli* (BLT5615) as the host strain.

The QCM Biosensor-Based T7 Phage Display

The protocol is summarized in Table 1. A ceramic sensor chip was attached on the oscillator of a 27-MHz QCM apparatus (AffinixQ), and the intrinsic frequency in the air phase was recorded before compound immobilization. After detaching the chip, a 20 μL aliquot

of a Brz2001 derivative solution (1 mM in 70% EtOH) was dropped onto the gold electrode of the ceramic sensor chip and left for 1 h under a humid and shaded atmosphere at room temperature. The surface of the electrode was washed for 10 min in buffer (10 mM Tris-HCl, pH 8.0, and 200 mM NaCl), which was stirred at 1,000 rpm at 25°C (AffinixQ, Initium Inc.). The sensor chip was set up for the QCM apparatus and recorded the decreased frequency in the air phase to measure the immobilized Brz2001 amount. The Brz2001 derivative was immobilized at ~11,000 Hz (330 ng) on average. The QCM sensor was immersed into the cuvette containing 8 mL of buffer and then allowed to fully stabilize. An aliquot of 8 μL of a T7 phage library (1.7×10^{10} pfu/mL) was injected into the cuvette. Frequency changes, caused by binding to the Brz2001 immobilized on the gold electrode surface, were then monitored for 10 min. For the recovery of bound phages, 20 μL of log-phase host *E. coli* (BLT5615) solution, which was cultured for 30 min at 37°C in the presence of 50 μg/mL of carbenicillin and 1 mM of IPTG beforehand (OD₆₀₀=1.0), was dropped onto the gold electrode and then incubated at 37°C for 30 min. Another 200 μL of LB medium was then added to the resulting solution, which was then subjected to plaque isolation. After that, the gold electrode surface was swabbed with 1% sodium dodecyl sulfate, treated with 5 μL of piranha solution (concentrated H₂SO₄: 30% H₂O₂=3: 1) for 5 min, washed with dH₂O, and then dried.

Plaque Isolation

The recovered phage solution, which contains 5×10^5 – 5×10^6 pfu/mL of T7 phages on average, was diluted from 10¹ to

Table 1. Protocol for the QCM Biosensor-Based T7-Phage Display and Analysis of RELIC Bioinformatics Server

Step	Parameter	Value	Description
1	Sensor chip preparation	20 μL	1 mM Brz2001 derivative forming SAM
2	Setup of QCM apparatus	8 mL	10 mM Tris-HCl, pH 8.0, 200 mM NaCl
3	T7-phage library injection	8 μL	Injection into the buffer-filled cuvette
4	Monitoring the QCM sensor	10 min	Monitoring the frequency decrease for 10 min
5	Recovery of bound phage	20 μL	Dropping the host <i>E. coli</i> culture on the gold electrode
6	Phage plaque isolation	3 h	Incubation at 37°C
7	PCR and DNA sequencing	50 μL	20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 6 mM MgSO ₄
8	Analysis of RELIC server		Uploading the peptide sequence (.txt) according to the instruction and executing the program
9	Sensor chip regeneration	5 μL	Swabbing the gold electrode surface by 1% SDS and then dropping the piranha solution

Step Notes

1. A 20 μL aliquot of a Brz2001 derivative solution (70% EtOH) that forms SAM was dropped onto the gold electrode of the ceramic sensor chip and left for 1 h under a humid and shaded atmosphere at room temperature. The immobilized amount on gold was measured as the decreased frequency in the air phase after immobilization.
2. The sensor chip was attached on oscillator and immersed; the buffer-filled cuvette was constitutively stirred at 1000 rpm at 25°C (AffinixQ, Initium Inc.).
3. After stabilizing the QCM sensor, the 8 μL of T7 phage library (final > 10⁸ pfu) was injected into the cuvette.
4. Frequency decrease by mass increase on binding of T7 phage to Brz2001 on the gold electrode was monitored for 10 min.
5. The sensor chip was detached from the apparatus, and 20 μL of log-phase and T7 phage-infective host *Escherichia coli* (BLT5615) solution was dropped onto the gold electrode to incubate for 30 min at 37°C with shaking. The resulting solution was added to 200 μL of LB medium.
6. The 100 μL of each diluted solution (10¹ to 10⁶ times) was plated onto LB/carbenicillin plate with 3 mL of prewarmed top agarose and 100 μL of host *E. coli* solution, and incubated for 3 h at 37°C.
7. Each plaque was isolated and suspended into 50 μL phage extraction buffer. Peptide-encoding region on T7 phage DNA was amplified by PCR using specific primers and then sequenced.
8. Resulting amino-acid sequence of peptide was listed in .txt file and uploaded to execute an appropriate program in RELIC bioinformatics server according to the instructions.
9. After swabbing the gold electrode surface by 1% SDS, 5 μL of piranha solution (concentrated H₂SO₄: 30% H₂O₂=3: 1) was dropped and left for 5 min, and then washed by dH₂O.

QCM, quartz-crystal microbalance; SAM, self-assembled monolayer; RELIC, receptor ligand contacts; SDS, sodium dodecyl sulfate; PCR, polymerase chain reaction.

10^6 times. A 100 μ L aliquot of each solution was mixed with 200 μ L of log-phase host *E. coli* solution and 3 mL of prewarmed top agarose [10 g/L Bacto tryptone, 5 g/L yeast extract, 5 g/L NaCl, 6 g/L agarose], and then seeded onto an LB/carbenicillin plate. The plate was incubated at 37°C for 3 h to form individual phage plaque.

PCR and DNA Sequencing

Plaques were randomly picked from LB plates, and each was dissolved in 50 μ L of phage extraction buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 6 mM $MgSO_4$). The phages were disrupted by heating the extract to 65°C for 10 min. To amplify the peptide-encoding region, PCR was performed according to the following PCR mixture: 0.3 μ L phage solution, 0.05 μ L of the 100 μ M forward primer 5'-TGC TAA CTT CCA AGC GGA CC-3' and the 100 μ M reverse primer 5'-AAA AAC CCC TCA AGA CCC GTT TA-3', 0.25 mM dNTP, 1 μ L of 10 $\times$ Ex Taq DNA polymerase buffer, and 1.25 U of Ex Taq DNA polymerase. The mixture was diluted up to 10 μ L using double-distilled water. PCR condition was 25 cycles each of 94°C for 60 s, 50°C for 30 s, and 72°C for 30 s. The 5 μ L of products were treated with 2 μ L of ExoSAP-IT (digested at 37°C for 15 min and then inactivated at 80°C for 15 min) and 70% EtOH precipitated. The products were then sequenced using ABI Prism BigDye Terminator Cycle Sequencing Kit and on an ABI Prism3100 Genetic Analyzer (ABI) according to the manufacturer's protocols.

Analysis Using RELIC Server

The stand-alone type RELIC programs were kindly provided by Dr. Lee Makowski (Northeastern University, Boston, MA). AAFREQ, MOTIF1 and 2, and MATCH programs in RELIC were used according to their manuals for analysis of Brz2001-selected peptides and detection of the Brz2001-binding site on DWARF4.¹³

RESULTS AND DISCUSSION

To immobilize Brz2001 (MW: 353.85) on the gold electrode of the QCM sensor chip, we used a derivative that forms a self-assembled monolayer on gold (Fig. 2B, Supplementary Fig. S1; Supplementary Data are available online at www.liebertpub.com/adt).^{28–30} Apart from its improved high solubility (>10 mM) in 70% EtOH, the larger molecular mass of derivative (MW: 1,490.78) enabled a tight attachment of the compound on the gold and thereby increased the sensitivity of QCM at least thrice greater than that using avidin-biotin immobilization by piggyback or mass enhancer effect.^{3,4,31–33} Figure 3A shows the representative sensorgram after injecting a T7 phage-displayed random peptide library into a buffer-filled cuvette where the sensor chip was immersed (Fig. 1A). The decreased frequency 10 min after the injection of T7 phage library was 108 ± 15 Hz; whereas after the injection of control T7 phage, it was 37 ± 15 Hz (Fig. 3B). The sensor chip was dislodged from the device, and the bound T7 phage DNA was directly recovered by host *E. coli* BLT5615 infection. After the plaque isolation, T7 phage DNA of the peptide-encoding region was sequenced using a specific primer and translated into an amino-acid sequence (Fig. 1A). In total, four sets of one-cycle biopanning identified 34 of the 15-mer peptides that potentially recognize Brz2001 (Table 2).

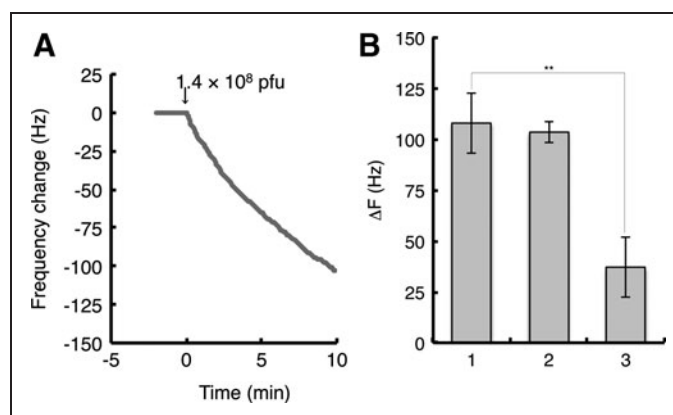


Fig. 3. The QCM biosensor-based one-cycle biopanning. **(A)** A representative QCM sensorgram obtained by an affinity selection of peptide on the QCM apparatus (AffinixQ) using the Brz2001 derivative and a random peptide T7 phage library (8 μ L; 1.4×10^8 pfu). The Brz2001-immobilized ceramic sensor chip generated by SAM was attached to the QCM apparatus. After injecting the T7-phage library at the indicated concentration, the frequency decrease was monitored for 10 min. **(B)** Frequency decrease 10 min after the injection. 1, T7-phage library versus Brz2001; 2, T7-phage library versus gold electrode; 3, Monoclonal T7 phage (displaying no 15-mer peptide) versus Brz2001. Data are means $\pm$ SD ($n=3$). ** $P < 0.01$.

In general, small molecules appear to show weak interactions with peptides via surface-surface contacts of continuous or discontinuous motifs of several amino-acid residues. Testing the interaction between individual peptides with small molecules should be experimentally limited to some extent, because there may be a large number of potential small-molecule-recognizing peptides within a library. Furthermore, affinity between immobilized small molecules and short peptides is too weak (K_D : 10^{-4} – 10^{-5} M range) to characterize each affinity status. Nonetheless, algorithmic and heuristic approaches using the RELIC program enabled rapid extraction of sequential information that is potentially associated with the small-molecule recognition from a subset of affinity-selected peptides, without checking the affinity status of individual peptides.^{11,13,23,24} Table 3 shows continuous motifs that were detected from 34 Brz2001-selected peptides by the use of the RELIC/MOTIF1 program. In addition to the exact match of short amino-acid stretches, conservative motifs are efficiently extracted from the limited peptide population. Indeed, a pair of six-continuous motifs, 6 kinds of five-continuous and 25 of four-continuous motifs, were detected in the subset of peptides (Table 3). Furthermore, the RELIC/MOTIF2 program, which searches for patterns of three amino acids and does not allow conservative amino-acid substitutions, allows identical gap lengths, highlighted VXXFXF (among five individual peptides), CXXVXL, FXVXXG, SXXXVXXXXS, VXLXV (four peptides; Table 4), and others (two or three peptides; data not shown). These motifs were not significantly detected within the 103 peptide sequences (Supplementary Table S1) arbitrarily selected from the unscreened parent library. In addition, frequency of amino acids C F G S V, which are the components of the continuous and discontinuous motifs, also

Table 2. Brz2001-Selected 15-mer Peptide Sequences After Four Sets of Individual One-Cycle Biopanning

No.	Sequence	No.	Sequence	No.	Sequence	No.	Sequence
1	VLLAVWSVFHERSWL	10	CLIVCLFIVRCGGCF	19	GCNVVMCPDFCLPAR	28	LGWVGFFPDFPSVLL
2	SLYLPCVLPDLRFRM	11	CGFAFYVWLMVMW	20	QFFFSDDGSPPEVG	29	CSQGLLHSAVFALCF
3	MLVCIFFFGMLIMF	12	FCVCFGCTIMVLL	21	SSCNLVAAHQFCGD	30 ^a	LSLGSFGVSYGVSG
4	ASPLIVAAACSSYFL	13	AFVSGADNVNCLHCA	22	LFISIFVLAFCFALV	31	SCLAWLNCSSGGLWS
5	FCVNGTTCVIGWV	14	TDVDFLPNIHVSG	23 ^a	SVFCVSGSSSHWLE	32 ^a	FCVNNGTCSVWRCD
6	RVFFVIFLSDLIWPF	15	SNAPCIVDMSSPLVH	24	YSCDAYVGVGVLCI	33 ^a	GCVFILWGISSSHDA
7	IFDALSLVCLSVQPS	16 ^a	SASCFTGGMVRHVHR	25 ^a	LHLIGLAASSGLSDL	34	CVPSFGFVLPNRLAA
8	LCCSVNLELSPSMGG	17	VRPHPLARSLPVAR	26	ICNSVSLFVFGPYGA		
9 ^a	CDFLCVRSVAGTMVA	18	CVYPAFTPFQYTLL	27	FLPGIGYASVLCSSG		

T7-phage particles were arbitrarily extracted from the resulting solution and then analyzed by PCR amplification of the DNA encoding the fusion peptide. The PCR products were then sequenced.

^aThe peptide directly highlights a disordered portion on DWARF4 using MATCH program (Fig. 5A).

increased 1.1- to 1.23-fold through the selection, although it decreased or did not change in the other amino acids (*Supplementary Tables S2 and S3*). Collectively, these results indicate that the information content for Brz2001 recognition significantly increased through the affinity selection for Brz2001.

Although the mode of docking of Brz2001 to DWARF4 has still been unsolved by NMR experiment or X-ray crystallography, the subsequent use of the RELIC/MATCH program with the Brz2001-selected peptide sequences enabled the prediction of polypeptide portions involved in Brz2001 binding. This program calculates a weak similarity between small-molecule-selected short peptides and a longer target polypeptide, even in the absence of a clearly identifiable consensus sequence motif or a three-dimensional structure of the target (Fig. 1B).^{3,4,11,13} Figure 4A shows a similarity plot between 34 Brz2001-selected random peptides and the DWARF4 amino-acid sequence (M1-L513). This plot was calculated by a modified BLOSAM62 amino-acid matrix within each window of five amino acids in length across the entire length of the protein sequence.^{12,14–16} The similarity was cumulatively scored and plotted at each amino-acid residue. Furthermore, scores calculated using 103 unscreened parent random peptides (*Supplementary Table S1*) were subtracted as a background. As a result, a portion containing S438 in DWARF4 was detected as a maximal similarity score by employing 34 Brz2001-selected peptide sequences. As seen in Figure 5A, this analysis also generated a cluster diagram, indicating that this portion is essential for Brz2001 recognition. A similar trend was noted in the analysis using another 26 Brz2001-selected peptides (Figs. 4B and 5B; *Supplementary Table S4*). In this case, other portions containing N379, P407, and P454, respectively, were also pinpointed, in addition to S438. By contrast, such maximal scores were not observed by analysis using 27 gold electrode-

recognizing control peptides (Fig. 4C; *Supplementary Table S5*). These results indicate the reproducibility and specificity of this method. Furthermore, a part of the highlighted amino-acid residues in 15-mer peptides overlapped with the potential Brz2001-recognizing motifs detected in analysis using RELIC/MOTIF-1 and -2 programs (Tables 3 and 4). Thus, this strategy could be applicable to annotate a small-molecule-recognizing portion on proteins in the case when the three-dimensional structure is unavailable. Significantly, the small-molecule binding site that is internally located on protein can be targeted by this approach. Furthermore, this method is free from both naturally and experimentally occurring protein mutations, deletions, and other modifications that may influence the protein folding and fail to target identification in some cases.

According to the information about human cytochrome 2C9, one of a Cyp₄₅₀ monooxygenase homologous to DWARF4 (E value: 6e-12 on standard blastp) and being available in the three dimensional structure [PDB ID: 1R90], the primary sequence of the highlighted portion is potentially disordered flexible loops that comprise a substrate-binding site within the vicinity of the catalytic site that is mediated by heme, iron, and a cysteine of the heme axial ligand (Fig. 5A, B; *Supplementary Fig. S2*).^{34,35} Indeed, DWARF4 inhibitors having a triazole ring interfere with the redox activity via this portion, which result in the inhibition of brassinosteroid biosynthesis and growth suppression of plants ranging in concentration from 10⁻⁴ to 10⁻⁶ M (K_D : 10⁻⁵–10⁻⁶ M).^{25–27,36,37} Dynamic mobilization of the flexible portions that have been first detected in this experiment seems to be essential during docking of Brz2001 to the catalytic site. In addition, several other amino-acid residues in DWARF4 were detected by this analysis (Fig. 4A, B). These residues could contribute to the binding of Brz2001 to DWARF4, or might just be false positives,

Table 3. Four to Six Continuous Motifs Extracted from Brz2001-Affinity-Selected Peptides Using the MOTIF1 Program in the RELIC Suite

Six continuous motifs			Five continuous motifs											
No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif
3	1	MLVCIF	3	10	MLLIM	4	4	LIVAA	5	11	CIGW	22	3	ISIFV
10	2	LIVCIF	11	10	LMVVM	21	5	LVLAA	31	2	CLAW	26	5	VSLFV
			12	10	IMVLL	5	5	NGTTC	10	5	CLFIV			
			12	11	MVLLL	32 ^a	5	NGTTC	33 ^a	2	CVFIL			
Four continuous motifs														
No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif	No.	Pos.	Motif
1	1	VLLA	3	2	LVCI	5	12	IGWV	8	7	LELS	16 ^a	6	TGGM
4	4	LIVA	7	7	LVCL	28	1	LGWV	15	7	VDMS	31	10	SGGL
21	5	LVLA	10	3	IVCL	31	3	LAWV	9 ^a	11	GTMV	17	5	PGLA
1	2	LLAV	24	12	VLCI	6	6	IFLS	27	8	ASVL	27	3	PGIG
25 ^a	3	LIGL	3	8	FGML	22	1	LFIS	10	6	LFIV	22	7	VLAF
1	4	AVWS	31	4	AVWL	6	8	LSDL	22	5	IFVL	28	3	VWGF
31	12	GLWS	3	9	GMLL	25 ^a	12	LSDL	33 ^a	3	VFIL	23 ^a	8	SSSH
2	1	SLYL	31	4	AVWL	7	6	SLVC	12	9	TIMV	33 ^a	10	SSSH
22	4	SIFV	4	2	SPLI	27	9	SVLC	28	12	SVLL	30 ^a	6	GFV
26	6	SLFV	15	11	SPLV	8	1	LCCS	14	12	VSFG	34	6	GFVL
2	5	PCVL	5	1	FCVS	27	11	LCCS	30 ^a	9	VSYG			
15	4	PCIV	23 ^a	3	FCVS									

^aThe peptide directly highlights a disordered portion on DWARF4 using MATCH program.

Table 4. Three Discontinuous Motifs Extracted from Brz2001-Affinity-Selected Peptides Using the MOTIF2 Program in the RELIC Suite

No.	Motif	No.	Motif	No.	Motif	No.	Motif	No.	Motif
3	MLVCIFFGMILLIMF	8	LCCSVNLELSPSMGG	5	FCVSNGTTCVCIGW	4	ASPLIVAAACSSYFL	1	VLLAVWSVFHERSWL
22	LFISIFVLAFCFALV	10	CLIVCLFIVRCGGCF	10	CLIVCLFIVRCGGCF	23	SVFCVSGSSSHWLIE	7	IFDALSLVCLSVQPS
26	ICNSVSLFVFGPYGA	26	ICNSVSLFVFGPYGA	30	LSLGSGFWSYGVSG	30 ^a	LSLGSGFWSYGVSG	11	CGFAFYVWLMVVMW
28	LGWVGFFPDPSVLL	31	SCLAWLNCSGGLWS	32	FCVNNGTTCVWRCD	31	SCLAWLNCSGGLWS	26	ICNSVSLFVFGPYGA
34	CVPSFGFVLPNRLAA								

Motifs conserving more than 4 peptides are presented in bold.

^aThe peptide directly highlights a disordered portion on DWARF4 using MATCH program.

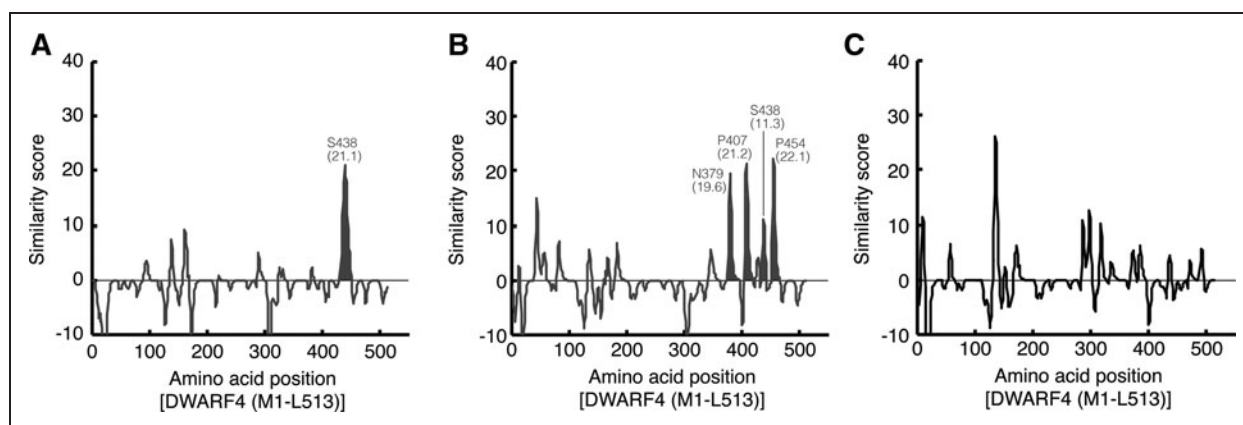


Fig. 4. Analysis of the similarity of amino-acid sequence between Brz2001-selected peptides and DWARF4 using the MATCH program in the RELIC suite: similarity plot. Scores for the sequence of DWARF4 (M1-L513) against the sequences of 34 (A) or 26 (B) peptides selected for affinity to Brz2001, or 27 for gold electrode (control) (C) was calculated using RELIC/MATCH program. The similarity scores of random peptides chosen without affinity selection (*Supplementary Table S1*) have been subtracted from these scores to remove library bias. The portion of maximal similarity score is shown with solid shading in the area under the curve.

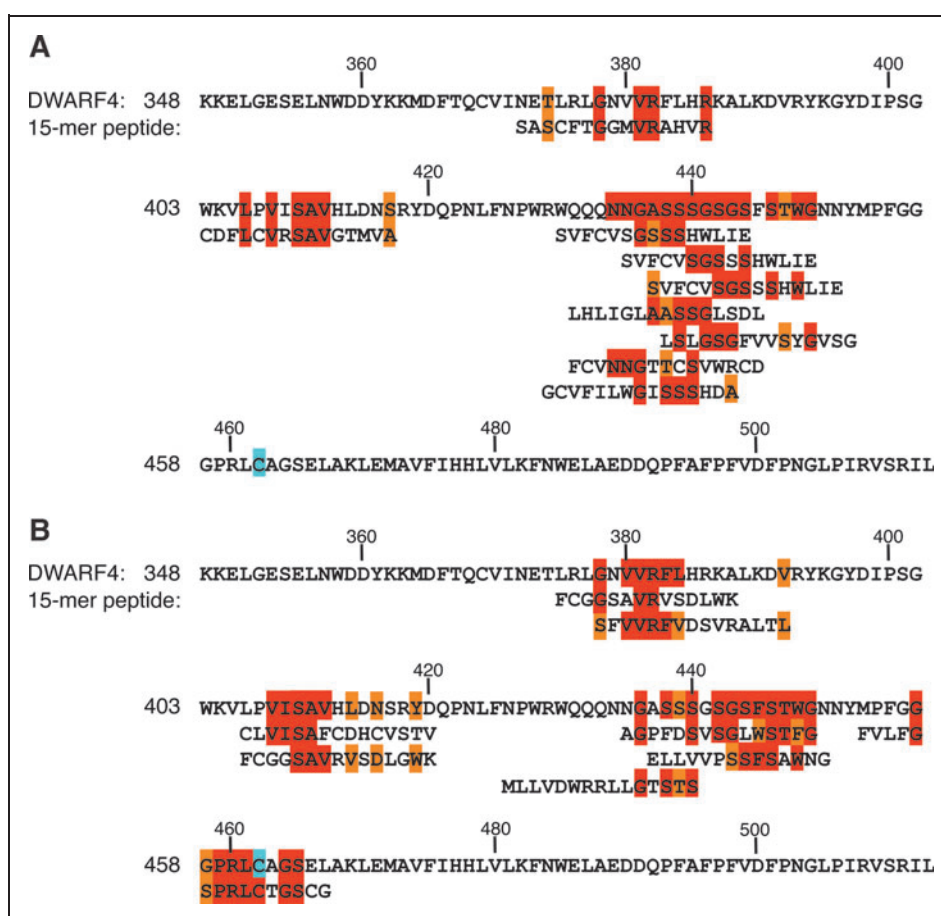


Fig. 5. Analysis of the similarity of amino-acid sequence between Brz2001-selected peptides and DWARF4 using the MATCH program in the RELIC suite: cluster diagram. A highlighted mapping of amino-acid sequence between Brz2001-affinity-selected 34 (A) or 26 (B) of peptides and DWARF4 at the portion of maximal similarity. Residues exhibiting identity or similarity to the protein sequence are highlighted in red or orange. The axial ligand of C462 coordinated with iron ion is shown in cyan. $1 \text{ Hz} = 0.62 \text{ ng/cm}^2$.

as can be seen in the control experiment (Fig. 4C). Conversely, undetected portions or residues involved in the binding of Brz2001 might be present. The introduction of the position of linker on Brz2001, or genetically modified diversity or length of library peptides, might improve these properties. Along with precise docking studies, further experiments make it possible to expand the utility of this approach.

Since brassinosteroid has been discovered as a plant growth-promoting hormone, extensive studies have been conducted about its biosynthesis as well as the production of inhibitors as a tool for biosynthesis regulation.^{25–27,36–41} Brz2001 is one of the synthetic compounds that specifically inhibits DWARF4 activity. These findings in our work will greatly contribute to better understanding the precise mode of docking as well as further designing the DWARF4 inhibitor for studies in brassinosteroid biosynthesis. Moreover, experimental methodology in this work, which could be beyond the technical difficulties of the conventional approach, will be a valuable tool for detecting a small-molecule binding site on a target protein as well as expanding the druggable proteome in a wide range of small-molecule-related studies, including plant molecular biology.

CONCLUSION

In this work, the QCM biosensor-based T7 phage display in combination use with bioinformatics analysis has extended into plant molecular biology. A screening of a library of T7 phage-displayed 15-mer random peptides was undertaken to identify 34 Brz2001-recognizing peptides. Subsequent qualitative and quantitative assessment of affinity-selected peptides, which is based on an algorithmic and heuristic approach using RELIC programs, predicted the increase in the information content of Brz2001 recognition through the selection, highlighting continuous and discontinuous peptide motifs. Furthermore, a similarity search using the RELIC/MATCH program with Brz2001-selected peptides has identified a potential disordered portion of DWARF4 as a scaffold during Brz2001 docking, even though the three-dimensional structure of this enzyme has been unavailable. A similar trend was also noted by an analysis using another 26 Brz-selected peptides and not observed using 27 control peptides, demonstrating the reproducibility and specificity of this method. Beyond the technical limitations associated with physicochemical properties of small-molecule or protein samples, this affinity-based methodology could be applicable to any small-molecule/protein systems under the identical protocol as a high-throughput method for the scanning of small-molecule-recognizing portion. Experimental methodology in this work will greatly contribute to the wide range of small-molecule-related studies, including plant molecular biology.

ACKNOWLEDGMENTS

The authors thank Dr. Tadao Asami (University of Tokyo, Tokyo, Japan) for providing Brz2001, and Dr. Lee Makowski (Northeastern University, Boston, MA) for providing the stand-alone type RELIC programs. This work was partially supported by a Grant-in-Aid for Scientific Research (The Ministry of Education, Culture, Sports, Science, and Technology of Japan, Japan Society for the Promotion of Science).

DISCLOSURE STATEMENT

The authors declare no conflict of interests.

REFERENCES

1. Crews CM: Targeting the undruggable proteome: the small molecules of my dreams. *Chem Biol* 2010;17:551–555.
2. Takakusagi Y, Takakusagi K, Sugawara F, Sakaguchi K: Use of phage display technology for the determination of the targets for small-molecule therapeutics. *Expert Opin Drug Discov* 2010;5:361–389.
3. Takakusagi Y, Takakusagi K, Sugawara F, Sakaguchi K: [Validation of small-molecule/protein interactions by the T7 phage display strategy using a quartz-crystal microbalance device]. *Tanpakushitsu Kakusan Koso* 2009;54:1203–1209.
4. Takakusagi Y, Kuramochi K, Takagi M, et al.: Efficient one-cycle affinity selection of binding proteins or peptides specific for a small-molecule using a T7 phage display pool. *Bioorg Med Chem* 2008;16:9837–9846.
5. Takakusagi Y, Kuroiwa Y, Sugawara F, Sakaguchi K: Identification of a methotrexate-binding peptide from a T7 phage display screen using a QCM device. *Bioorg Med Chem* 2008;16:7410–7414.
6. Takakusagi Y, Takakusagi K, Kuramochi K, Kobayashi S, Sugawara F, Sakaguchi K: Identification of C10 biotinylated camptothecin (CPT-10-B) binding peptides using T7 phage display screen on a QCM device. *Bioorg Med Chem* 2007;15:7590–7598.
7. Sauerbrey G: The use of quartz oscillators for weighing thin layers and for microweighing. *Z Phys* 1959;155:206–222.
8. Speight RE, Cooper MA: A survey of the 2010 quartz crystal microbalance literature. *J Mol Recognit* 2012;25:451–473.
9. Becker B, Cooper MA: A survey of the 2006–2009 quartz crystal microbalance biosensor literature. *J Mol Recognit* 2011;24:754–787.
10. Cooper MA, Singleton VT: A survey of the 2001 to 2005 quartz crystal microbalance biosensor literature: applications of acoustic physics to the analysis of biomolecular interactions. *J Mol Recognit* 2007;20:154–184.
11. Makowski L: Quantitative analysis of peptide libraries. In: *Phage Nanobiotechnology*. Petrenko VA, Smith GP, (eds.), pp. 33–54. RSC Publishing, Cambridge, United Kingdom, 2011.
12. Carter DM, Gagnon JN, Damlaj M, et al.: Phage display reveals multiple contact sites between FhuA, an outer membrane receptor of *Escherichia coli*, and TonB. *J Mol Biol* 2006;357:236–251.
13. Mandava S, Makowski L, Devarapalli S, Uzubell J, Rodi DJ: RELIC—a bioinformatics server for combinatorial peptide analysis and identification of protein-ligand interaction sites. *Proteomics* 2004;4:1439–1460.
14. Rodi DJ, Agoston GE, Manon R, Lapcevic R, Green SJ, Makowski L: Identification of small molecule binding sites within proteins using phage display technology. *Comb Chem High Throughput Screen* 2001;4:553–572.
15. Rodi DJ, Makowski L: Similarity between the sequences of taxol-selected peptides and the disordered loop of the anti-apoptotic protein, Bcl-2. *Pac Symp Biocomput* 1999:532–541.
16. Rodi DJ, Janes RW, Sanganeer HJ, Holton RA, Wallace BA, Makowski L: Screening of a library of phage-displayed peptides identifies human bcl-2 as a taxol-binding protein. *J Mol Biol* 1999;285:197–203.
17. Rezaei-Ghaleh N, Blackledge M, Zweckstetter M: Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *ChemBioChem* 2012;13:930–950.
18. Metallo SJ: Intrinsically disordered proteins are potential drug targets. *Curr Opin Chem Biol* 2010;14:481–488.
19. Uversky VN, Dunker AK: Understanding protein non-folding. *Biochim Biophys Acta* 2010;1804:1231–1264.
20. Hadfield AT, Diana GD, Rossmann MG: Analysis of three structurally related antiviral compounds in complex with human rhinovirus 16. *Proc Natl Acad Sci USA* 1999;96:14730–14735.
21. Curry S, Mandelkow H, Brick P, Franks N: Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat Struct Biol* 1998;5:827–835.

22. Mattos C, Rasmussen B, Ding X, Petsko GA, Ringe D: Analogous inhibitors of elastase do not always bind analogously. *Nat Struct Biol* 1994;1:55–58.
23. Makowski L, Soares A: Estimating the diversity of peptide populations from limited sequence data. *Bioinformatics* 2003;19:483–489.
24. Rodi DJ, Soares AS, Makowski L: Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries. *J Mol Biol* 2002;322:1039–1052.
25. Sekimata K, Kimura T, Kaneko I, et al.: A specific brassinosteroid biosynthesis inhibitor, Brz2001: evaluation of its effects on Arabidopsis, cress, tobacco, and rice. *Planta* 2001;213:716–721.
26. Min YK, Asami T, Fujioka S, Murofushi N, Yamaguchi I, Yoshida S: New lead compounds for brassinosteroid biosynthesis inhibitors. *Bioorg Med Chem Lett* 1999;9:425–430.
27. Asami T, Mizutani M, Fujioka S, et al.: Selective interaction of triazole derivatives with DWF4, a cytochrome P450 monooxygenase of the brassinosteroid biosynthetic pathway, correlates with brassinosteroid deficiency in planta. *J Biol Chem* 2001;276:25687–25691.
28. Cheng CI, Chang YP, Chu YH: Biomolecular interactions and tools for their recognition: focus on the quartz crystal microbalance and its diverse surface chemistries and applications. *Chem Soc Rev* 2012;41:1947–1971.
29. Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM: Self-assembled monolayers of thiolates on metals as a form of nanotechnology. *Chem Rev* 2005;105:1103–1169.
30. Wink T, van Zuilen SJ, Bult A, van Benkom WP: Self-assembled monolayers for biosensors. *Analyst* 1997;122:43R–50R.
31. Takakusagi Y, Takakusagi K, Ida N, et al.: Binding region and interaction properties of sulfoquinovosylacylglycerol (SQAG) with human vascular endothelial growth factor 165 revealed by biosensor-based assays. *MedChemComm* 2011;2:1188–1193.
32. Mao X, Yang L, Su XL, Li Y: A nanoparticle amplification based quartz crystal microbalance DNA sensor for detection of *Escherichia coli* O157:H7. *Biosens Bioelectron* 2006;21:1178–1185.
33. Henne WA, Doorneweerd DD, Lee J, Low PS, Savran C: Detection of folate binding protein with enhanced sensitivity using a functionalized quartz crystal microbalance sensor. *Anal Chem* 2006;78:4880–4884.
34. Wester MR, Yano JK, Schoch GA, et al.: The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J Biol Chem* 2004;279:35630–35637.
35. Williams PA, Cosme J, Ward A, Angove HC, Matak Vinkovic D, Jhoti H: Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 2003;424:464–468.
36. Bajguz A, Asami T: Suppression of *Wolffia arrhiza* growth by brassinazole, an inhibitor of brassinosteroid biosynthesis and its restoration by endogenous 24-epibrassinolide. *Phytochemistry* 2005;66:1787–1796.
37. Bajguz A, Asami T: Effects of brassinazole, an inhibitor of brassinosteroid biosynthesis, on light- and dark-grown *Chlorella vulgaris*. *Planta* 2004;218:869–877.
38. Yoshimitsu Y, Tanaka K, Fukuda W, et al.: Transcription of DWARF4 plays a crucial role in auxin-regulated root elongation in addition to brassinosteroid homeostasis in *Arabidopsis thaliana*. *PLoS One* 2011;6:e23851.
39. Sekimata K, Ohnishi T, Mizutani M, et al.: Brz220 interacts with DWF4, a cytochrome P450 monooxygenase in brassinosteroid biosynthesis, and exerts biological activity. *Biosci Biotechnol Biochem* 2008;72:7–12.
40. Nagata N, Asami T, Yoshida S: Brassinazole, an inhibitor of brassinosteroid biosynthesis, inhibits development of secondary xylem in cress plants (*Lepidium sativum*). *Plant Cell Physiol* 2001;42:1006–1011.
41. Nagata N, Min YK, Nakano T, Asami T, Yoshida S: Treatment of dark-grown *Arabidopsis thaliana* with a brassinosteroid-biosynthesis inhibitor, brassinazole, induces some characteristics of light-grown plants. *Planta* 2000;211:781–790.

Address correspondence to:
Kengo Sakaguchi, PhD, or Fumio Sugawara, PhD
Department of Applied Biological Science
Faculty of Science and Technology
Tokyo University of Science
2641 Yamazaki
Noda, Chiba 278-8510
Japan

E-mail: kengo@rs.noda.tus.ac.jp or sugawara@rs.noda.tus.ac.jp