



NPR

Identifying the cellular targets of natural products using T7 phage display

Journal:	<i>Natural Product Reports</i>
Manuscript ID	NP-HIG-10-2015-000128.R1
Article Type:	Highlight
Date Submitted by the Author:	24-Jan-2016
Complete List of Authors:	Piggott, Andrew; Macquarie University, Department of Chemistry and Biomolecular Sciences Karuso, P; Macquarie University, Department of Chemistry and Biomolecular Sciences

SCHOLARONE™
Manuscripts



Natural Product Reports

ARTICLE

Identifying the cellular targets of natural products using T7 phage display

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Andrew M. Piggott^a and Peter Karuso^{a*}

While Nature continues to deliver a myriad of potent and structurally diverse biologically active small molecules, the cellular targets and modes of action of these natural products are rarely identified, significantly hindering their development as new chemotherapeutic agents. This article provides an introductory tutorial on the use of T7 phage display as a tool to rapidly identify the cellular targets of natural products and is aimed specifically at natural products chemists who may have only limited experience in molecular biology. A brief overview of T7 phage display is provided, including its strengths, weaknesses, and the type of problems that can and cannot be tackled with this technology. Affinity probe construction is reviewed, including linker design and natural product derivatisation strategies. A detailed description of the T7 phage biopanning procedure is provided, with valuable tips for optimising each step in the process, as well as advice for identifying and avoiding the most commonly encountered challenges and pitfalls along the way. Finally, a brief discussion is provided on techniques for validating the cellular targets identified using T7 phage display.

Introduction

Natural Products

Natural products have undisputedly been an invaluable source of therapeutic drugs in the past, and despite a substantial decline in natural products research and development by pharmaceutical companies in recent years, 26% of the new drugs approved from 1981–2010 were either natural products or natural product derivatives.¹ A further 24% of these new drugs were synthetic, but designed as ligands to proteins that were previously identified through their binding to natural products; i.e. as mimics of natural product pharmacophores. The remarkable success of natural products as drugs and drug leads can be explained by the hypothesis that essentially all natural products are privileged structures² and therefore every natural product must have a biological target; our challenge is to identify the biological target of each natural product. If this could be achieved in a routine way, many new drug targets would be identified and more drugs would be developed based on natural product structures. Importantly, once the protein target of a natural product is discovered, that protein is identified as 'druggable' and the full repertoire of medicinal and combinatorial chemistry can be applied to the structure-based design, development and optimisation of a new drug.

The classic approach to finding a natural product's binding partner involves a series of linear steps that must start from some sort of known biological activity and then proceed through a process of slowly narrowing down the possible modes of action until only one possibility remains. For example, Jung and co-workers isolated new and known sesterterpenes from the sponge *Psammocinia* sp. through a bioassay-guided process.³ The brine shrimp lethality assay employed suggested to the authors that a mixture of strobilinin and felixinin act as cytotoxic agents. A panel of cancer cell lines was thus tested and selective activity for the mixture was found against SK-MEL-2 melanoma cells at low micromolar concentrations. A common mode of anticancer activity involves the inhibition of DNA replication, so the authors tested their compounds in a ³¹P-based DNA replication assay and found that the compounds inhibited SV40 DNA replication *in vivo*. Inhibition of topoisomerase or DNA polymerase can affect DNA replication so the authors tested their compounds against two enzymes *in vitro* (topoisomerase I and DNA polymerase R) by previously established assays and found modest activity against both enzymes at 5 and 10 μM respectively. The results have since appeared in several reviews, which state that strobilinin and felixinin are ligands for topoisomerase and DNA polymerase.^{4–10}

The series of linear steps employed by the authors highlights some of the problems of a classical approach to the determination of natural product modes of action. At each step, a series of assumptions must be made about the underlying mechanism that depends intrinsically upon the current state of knowledge: Only activity for which a test exists

^a Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW 2109, Australia.

* Corresponding Author. Ph: +61-2-9850-8290; Email: peter.karuso@mq.edu.au

ARTICLE

Natural Product Reports

can be explored, initiating a somewhat circular argument. For all other activities, as Wittgenstein put it “*Wovon man nicht sprechen kann, darüber muß man schweigen*”.¹¹ In the traditional approach, one can only find activity in places where it can be found; that is to say, in systems that we already know about and using assays that are already established. As noted above, 24% of new drugs are synthetic compounds targeted to biomolecules that were identified as binding to a natural product. To increase this number, what is needed are unbiased genome-wide methods that are capable of agnostically identifying natural product-binding proteins.¹²⁻¹⁴

Genome-wide approaches can be classified as either *forward chemical proteomics*, in which a cell's proteome acts as a library from which the natural product's binding partner is isolated, or *reverse chemical proteomics*, in which the cell's genome (transcriptome) is used as a starting point.^{12,13,15} In both cases, selection is based on an interaction between the natural product and a protein library (proteome) but in the latter, the protein is attached in some way to its encoding gene.

Forward Chemical Proteomics

The most common method of linking natural products to their cellular targets is forward chemical proteomics, whereby the natural product is either fluorescently/radioactively labelled and used to identify binding partners from protein electrophoresis gels, or is immobilised onto a solid support and used to “pull-down” binding partners from entire proteomes. This approach was first used in 1965 to identify a previously undescribed 57 kDa protein binding partner for the cytotoxic natural product colchicine.¹⁶⁻¹⁸ This protein was later shown to be the subunit of microtubules and was named “tubulin”.¹⁹ Subsequently, forward chemical proteomics has been used to identify cellular targets for numerous other natural products, including the immunosuppressants cyclosporin²⁰⁻²³ and FK506,²⁴⁻²⁷ the cytotoxins didemnin B²⁸⁻³⁰ and halichondrin B,³¹ the antibiotics fumagillin^{32,33} and myriocin,^{34,35} and many others,^{13,15,36} including an off-target for acivicin.³⁷

Affinity purification (forward chemical proteomics) remains one of the most important methods for isolating natural product binding proteins, as it can directly reveal the physical interactions between small molecules and their biomolecular targets. However, due to the size of the proteome (2×10^6 for humans), the dynamic range of protein concentrations ($>10^{12}$ for plasma) and the range of affinities small molecule–protein interactions (pM–mM), affinity purification faces a specificity problem: how to identify the true target(s) of a natural product in the face of a large number of weak or non-specific interactions.³⁸

Reverse Chemical Proteomics

Due to the specificity problem, it is common to isolate the most abundant rather than the most avid binding partners when using forward chemical proteomics to identify the

cellular targets of natural products.^{28,29} This can be problematic for proteins that are expressed in very low concentrations at any given time in a cell. Reverse chemical proteomics addresses this problem by employing a physical link between each protein and its encoding gene. Using this so-called “genotype–phenotype link”,³⁹ low copy-number proteins that are recovered using a natural product affinity support can be replicated using the attached gene and subjected to subsequent pull-down experiments. This iterative approach results in amplification of weak signals (low copy-number proteins) and therefore allows the most avid binding partner to be identified.

Over the past few years, a number of strategies have been developed to screen random peptide libraries for sequences targeting specific proteins.⁴⁰⁻⁴⁵ Peptides with binding properties can be serially selected and amplified based on a physical link with the encoding cDNA. These technologies include phage, mRNA/DNA, viral, ribosomal, and cell display technologies, with phage display being the best understood and most commonly used. In contrast, the display of human proteins on the surface of cells, organelles, and viruses is far less common but has very recently provided biologists with important tools for the directed evolution of proteins (e.g., enzymes and antibodies)⁴⁶⁻⁵² and analysis of protein–protein interactions.⁵³⁻⁵⁶ However, the use of display technologies for the unbiased isolation of proteins that bind natural products is an underutilised technique that holds considerable potential in chemical biology.

Phage Display

The display of peptides on the surface of bacteriophages (phages) was first developed with the *Escherichia coli*-specific filamentous phage, M13.⁵⁷ M13 phages replicate without killing their host by assembling in the periplasm of the bacteria and secreting through the outer membrane. The most widely used format for ‘phage display’ is through fusion of the pIII or pVIII coat proteins of M13 phage to the peptide of choice. These coat proteins (pIII and pVIII) tolerate only *N*-terminal display as the C-terminus is inside the phage particle. This means that the display of cDNA libraries of eukaryote proteins is not possible because the bacterial ribosome first translates the foreign cDNA and terminates when it reaches the 3'-stop codon; thus never translating the actual coat protein. The only *C*-terminal display system, where the coat protein is translated first is with coat protein pVI.⁵⁸ There are strict limitations on the size of the chimera using pVI and consequently this system has not proved popular.

In an alternative strategy, Cramer and Suter designed a pIII-based cDNA display system that utilised the leucine-zipper interaction of c-Fos and c-Jun. Here, the cDNA library is fused to c-Fos and incorporated into the M13 genome. The pIII protein is constitutively fused to c-Jun and, in the cytoplasm, the two proteins associate and form a disulfide bond, covalently linking the target protein to pIII via c-Fos and c-

Jun.⁵⁹ As the pIII protein is also required for bacteriophage assembly and infectivity, there are some limits to the size of the protein that can be successfully expressed and displayed without causing an unacceptable bias toward phage with no gene insert. A related limitation is that M13 phage particles must be secreted through the bacterial membrane, which can act as a barrier to some clones and potentially bias selections. These problems have led to the search for alternate display systems that are lytic and facilitate C-terminal display to overcome the stop codon problem.

T7 Phage Display

The T7 phage display system invented by Rosenberg *et al.* and commercialised by Novagen (Merck-Millipore) avoids the stop codon problem by utilising C-terminal display on the gp10B protein.⁶⁰ T7 phage is also lytic, so display and reproduction are not dependent on secretion through the bacterial membrane. The size of the displayed protein can be quite large, with functional enzymes of at least 1200 amino acids having been successfully displayed. The number of copies per phage particle is also adjustable from 0.1–15 out of 415 copies of gp10B in the phage capsid, allowing multivalent interactions. These phage also have a rapid life cycle and are resistant to severe conditions such as 5 M NaCl, 4 M urea, 2 M guanidine-HCl, 100 mM DTT and a pH range between 4 and 10, making them perfect for application in a chemistry laboratory. However, they can be effectively destroyed using UV light or bleach.

The first example, and clear indication of the potential of phage display to isolate protein binding partners for small molecules, was demonstrated by Austin, who used the immunosuppressant drug FK506 to isolate its known protein target (FKBP) from a T7 phage display library.⁶¹ Initially, FK506 was biotinylated at the allyl group by dihydroxylation of the double bond (OsO_4), cleavage of the resulting diol (NaIO_4), reduction of the aldehyde ($\text{LiAlH}(\text{OEt})_3$) and activation of the alcohol (*N,N'*-disuccinimidyl carbonate). This site was chosen because it was already well known that the allyl group is exposed to solvent in the FK506-FKBP crystal structure, so derivatisation should not interfere with binding to FKBP.⁶² The biotinylated analogue (**1**) was immobilised onto monomeric avidin-agarose beads to produce an affinity column. Lysate from T7 phages displaying a human brain cDNA library was affinity captured onto an FK506-derivatised column, washed and eluted with free biotin. The eluate was transfected into fresh *E. coli* and the new sub-library precipitated with PEG to remove biotin from the previous elution step. Finally, the phages were resuspended in buffer and absorbed onto a fresh FK506 affinity column for the next round of selection. Following each round of selection (biopanning), random phage plaques were picked and their DNA inserts amplified by PCR. After the second round of selection, all clones picked gave rise to a DNA band of 450 bp, suggesting the library had converged. DNA sequencing of these bands revealed identical copies of a full-length, in-frame gene coding for human FKBP.

Therefore, it appeared that after only two rounds of selection with the FK506 affinity resin, phage particles displaying FKBP on their surface had been amplified selectively to become the dominant members of the library. However, it was later discovered that the FKBP clone isolated was actually a synthetic positive control, rather than from the brain cDNA library.⁶³ Cross-contamination is a common problem associated with all phage display systems. T7 phages are particularly hardy and can survive on surfaces for many years. However, using only disposable plastic ware, minimising the formation of aerosols, and disinfecting exposed surfaces with bleach or UV light can reduce the risk of contamination. Austin repeated the experiment, performing seven rounds of selection and taking particular care to avoid contamination. Sixteen random clones were picked after the sixth round of selection and their DNA inserts amplified by PCR. Five of these clones were found to be identical and contained the entire coding sequence of FKBP1a.⁶³ This number increased to eleven out of sixteen clones in round seven. The results of this experiment showed that it is possible to isolate the binding partner for a small epitope such as a natural product using a cDNA library displayed on T7 bacteriophage, thereby providing an important proof of concept.

Subsequently, T7 phage display has been employed to identify cellular targets for a number of other natural products (Table 1). Once the cellular target for a natural product is known, T7 phage display can also be used to identify novel isoforms of that target. For example, we have used T7 phage display to identify new FKBP isoforms from a plant (*Arabidopsis thaliana*) cDNA T7 library, and a bacterial (*Pseudomonas stutzeri*) gDNA T7 library.⁶⁴

<Table 1; full page width>

Procedures

In this tutorial, we will take the reader through the process of reverse chemical proteomics using T7 phage display, from preparation of the affinity supports through biopanning to validation of hits. We have also provided a comprehensive step-by-step guide to this process as Supporting Information, including a list of equipment and reagents required, full details of reagent preparation, volumes, incubation times and temperatures, as well as helpful hints to avoid the most commonly encountered pitfalls along the way.

<Figure 1; full page width>

Construction of Natural Product Affinity Probe

The first step in natural product target identification using T7 phage display is construction of an affinity probe. Affinity probes generally contain three domains – a tag, a linker and a natural product (Figure 1).¹⁴ The tag can consist of either a radioactive/fluorescent label, to allow visualisation of bound proteins on an electrophoresis gel, or a solid-phase bead/surface, to allow affinity purification of bound proteins. Frequently, biotin is used as the tag as it allows both affinity

purification, using a streptavidin resin, and in-gel visualisation, using streptavidin coupled to a reporter enzyme such as horseradish peroxidase. The natural product and tag are usually separated from each other by a linker, which can be simply an alkyl chain, a peptide or a polyethylene glycol (PEG).¹⁴ Alkyl chain-based linkers are chemically stable, but fold up upon themselves in aqueous environments due to hydrophobic effects. Peptide-based linkers are hydrophilic, but susceptible to proteolysis in biological media. PEG-based linkers are the most commonly used in affinity probes as they are both chemically stable and hydrophilic, with the alternating C–O dipoles ensuring the linker adopts an extended conformation in aqueous environments. The chosen linker should also be as long as possible to provide sufficient room for large proteins to bind to the immobilised natural product, as the binding site for a particular small molecule may be deeply buried in the protein target.

The most difficult step in constructing a natural product affinity probe is attaching the tag and linker to the natural product at a site that does not interfere with its biological activity. In most cases, natural products are highly functionalised molecules that are available in only very limited supply, making synthetic modifications particularly challenging. Frequently, the most reactive functional group on a natural product (and hence the easiest place to attach a linker) is essential for its biological activity and cannot be modified. Ideally, it is best to functionalise a C–H bond and leave all other functional group untouched. There have been recent initiatives that will make this a reality in the near future.⁶⁵ While this aspect of identifying cellular targets for natural products is beyond the scope of this article, we refer readers to an excellent review in this journal by Daniel Romo.⁶⁶ While probe construction critically depends on the structure of the natural product under investigation, there are a number of developing areas of research that may make this easier in the future. For example, Marty Burke's 'synthesis machine' promises to make automated natural product and derivatised natural products a reality.⁶⁷ In parallel, the chemoselective/site selective derivatisation of natural products is an active area of research^{66,68–70} as is mild C–H bond activation to introduce new linkage sites onto complex natural products.^{71–73}

If the biological activity of a natural product is known or suspected, it is possible to assay the completed affinity probe or a simple derivative that masks the site of linker attachment, to ensure this activity has not been lost. However, if we wish to take a completely unbiased and agnostic approach to natural product target discovery, it is clearly not possible to determine whether synthetic modifications to a natural product will have any impact on binding to its (as yet unknown) cellular target. The best way to increase the probability of obtaining a functional affinity probe is to attach the tag and linker to several different locations on the natural product and pool the resulting affinity probes. This approach is also useful if the natural product must be presented to the cellular target in a particular orientation.

Immobilisation of Natural Product Affinity Probe

Before affinity selection from a T7 phage display library can begin, the natural product affinity probe must be immobilised onto a solid support. For biotinylated probes, the most commonly used solid supports are NeutrAvidin/streptavidin-coated microtitre plates, agarose resins or magnetic beads. For this tutorial, we will assume a NeutrAvidin-coated microtitre plate is being used. In such a case, a 100× stock solution of the affinity probe is prepared in DMSO and then diluted with a suitable buffer to give a working concentration containing <1% DMSO. The NeutrAvidin-coated plates are then incubated with this working solution to immobilise the affinity probe onto the solid support. Finally, the plate is washed with buffer to remove unbound probe and refrigerated until required. This procedure is then repeated for an appropriate negative control compound, which usually consists of the biotinylated linker terminated with a small fragment of the natural product of interest.

Selection of T7 host strain

The two most commonly used *E. coli* host strains for T7 phage display are BLT5403 and BLT5615, both of which carry an ampicillin resistance plasmid that also supplies extra T7 capsid protein (gp10A). In BLT5403, gp10A is driven by a T7 promoter, yielding large amounts of gp10A upon T7 infection. In BLT5615, gp10A is driven by a lacUV5 promoter, which requires the addition of IPTG to induce capsid production. Four variants strains of BLT5616 are also available for specific applications. BLT5615*rna* has the nonessential RNase I gene deleted, allowing RNA to be used as a probe. Origami B 5615 contains mutations in the thioredoxin reductase and glutathione reductase genes, creating a cytoplasmic environment compatible with the formation of disulfide bonds, thereby facilitates the expression of active, soluble proteins in *E. coli*. Rosetta 5615 enhances the expression of eukaryotic proteins by supplying tRNAs for codons rarely used in *E. coli* (AUA, AGG, AGA, CGG, CUA, CCC and GGA). Rosetta-gami 5615 combines the features of Origami B 5615 and Rosetta 5615.

Growth of T7 Lysates

To obtain a T7 sublibrary suitable for biopanning, an *E. coli* host strain (e.g. BLT5615) is first inoculated with a small aliquot of the primary T7 cDNA/gDNA library and the culture is incubated until cell lysis occurs (typically 1.5 h), as indicated by the culture transitioning from cloudy to almost completely clear over a 3–5 min period. Once lysis is complete, the culture is immediately cooled on slushy ice to minimise degradation of displayed proteins. The addition of a protease inhibitor cocktail at this stage is worth considering, although this may cause problems if the cellular target of the natural product turns out to be a protease. The chilled phage lysate is then centrifuged to precipitate cellular debris and the supernatant is decanted into a clean tube ready for biopanning. A mild detergent (e.g. 0.05% Tween 20) can be added to the clarified

lysate to reduce non-specific binding in subsequent biopanning experiments. For long-term storage, the phage lysate can be kept in 10% glycerol at $-80\text{ }^{\circ}\text{C}$.

Biopanning

To reduce the number of non-specific binders present in the initial T7 sublibrary, an aliquot of the clarified phage lysate is first added to a microtitre plate containing the immobilised control compound and left to incubate for 1 h. The depleted lysate is then transferred to a plate containing the immobilised natural product probe and left to incubate for 2 h. The lysate is then discarded and the plate is washed quickly with an appropriate wash buffer. Finally, the retained phages are eluted either specifically, by addition of free natural product or cleavage of the probe linker, or non-specifically with 1% SDS. The eluted phages (Round 1 sublibrary) are then transfected into fresh *E. coli* host cells and the entire biopanning process is repeated. Generally, 5 or 6 rounds of biopanning are sufficient for the library to converge to a manageable number of clones. In the initial T7 library, there will only be a relatively small number of clones displaying proteins capable of binding to the probe, so washing too aggressively at the start may result in all target phages being lost. Therefore, the stringency of the washing step (number of washes and total washing time) should be progressively increased with each round of biopanning.

<Figure 2>

Titration of Eluted T7 Phages

Once biopanning is complete, it is useful to determine the number of phage particles recovered from each round of selection. To calculate phage titres, the eluted phage solution from each round of biopanning is serially diluted from 10^{-1} to 10^{-8} and a small aliquot of each dilution is dropped onto an agar plate inoculated with a lawn of *E. coli* host cells. The plate is then incubated until small clear areas (plaques) are clearly visible against the background lawn of bacteria (Figure 2). The dilutions that contain a countable number of plaques (usually 5–50) are identified and the number of discrete plaques visible in these dilutions are counted. The phage titres are then back-calculated to give plaque-forming units per millilitre (pfu/mL), taking into account the volume of phage solution used and the dilution factor. Finally, the calculated titres are plotted against the selection round (Figure 2). Generally, a sigmoidal curve can be fitted to the data, with a large increase in phage titre indicative of a successful selection. A titre that gradually decreases as a function of selection round is generally indicative of a failed experiment.

Gel Electrophoresis of T7 Sublibraries

An indication of the success or failure of a biopanning campaign can also be obtained by PCR amplification of the phage DNA inserts present after each round of affinity selection (Figure 3). The initial T7 library contains a large number of unique phage clones and hence the amplified DNA inserts appear as a long smear on the DNA gel. The appearance of one or more intense bands in the DNA gel coinciding with a

marked increase in phage titre suggests selective enrichment of one or more specific phage clones by the natural product affinity probe.

<Figure 3>

Picking T7 Plaques

The final step in the T7 phage display process is to identify the cellular targets that were enriched by the natural product affinity probe. This can be achieved by isolating a number of individual phage clones and sequencing each of their DNA inserts to reveal the identity of the proteins being displayed on the phage surface. For maximum diversity of cellular targets, the sublibrary immediately following the dramatic increase in phage titre should be used for picking plaques. Alternatively, the final sublibrary can be used to identify the most avid binding target. In either case, an aliquot of phage lysate containing ~ 250 plaques is spread over a lawn of *E. coli* host cells and the plate is incubated until clear plaques are clearly visible. Micropipette tips are then used to stab the centre of each plaque to obtain sufficient phage lysate for PCR amplification and DNA sequencing. T7 phages displaying larger proteins tend to grow more slowly than those displaying smaller proteins, or no proteins at all (überphage). When picking plaques, it is best to aim for the smaller ones as these are more likely to display intact proteins.

<Figure 4>

Digesting an aliquot of the PCR-amplified DNA inserts from each individual phage plaque with a frequent cutting restriction endonuclease and separating the fragments by gel electrophoresis allows a unique DNA fingerprint to be generated for each clone (Figure 4). This provides a rapid and cost-effective method of determining whether two clones contain the same DNA insert, closely related inserts or completely different inserts, thereby minimising the number of DNA sequencing reactions required. The restriction endonuclease *Hinf*I, which cuts DNA at the sequence G|ANTC, is robust, relatively inexpensive and produces a suitable number of fragments from cDNA inserts in the 100–1000 bp size range.

<Figure 5>

After sequencing clones containing unique DNA inserts, a BLAST nucleotide search is performed to identify the genes encoded and hence the proteins being displayed (Figure 5). The T7 phage vector restriction site is located near the beginning of the DNA sequence (**GAATTC** for *Eco*R1). The first codon in the same reading frame as the gp10B coat protein is **AAT**. The foreign DNA insert must be in the same reading frame as the coat protein, which can be determined by counting in groups of three to the start codon (**ATG**). The presence of both start and stop (**TAA/TGA/TAG**) codons in the correct reading frame indicates that a full-length protein was displayed, while the absence of one or both of these codons indicates that a smaller protein fragment was displayed. If the BLAST search shows “Plus/Minus” in the strand column, this indicates that the sequence has been inserted backwards into

the T7 phage vector and was most likely not displayed correctly.

Validation

Having identified one or more cellular targets for a natural product using T7 phage display, the protein–small molecule interaction must be fully characterised and validated. The simplest target validation method is the on-phage binding assay,⁷⁴ in which equal numbers of the rescued phage clone (displaying the putative target protein) and a negative control phage (e.g. überphage) are added to plates derivatised with either the biotinylated natural product or a control compound and are incubated with varying concentrations of free natural product. After washing and elution, the phage titres from each experiment are used to calculate the protein–small molecule binding affinity (K_D) from a four-parameter logistic nonlinear regression model.⁷⁵

For most other validation methods, the putative target protein must be first cloned and overexpressed in a suitable host. In some cases, the entire gene encoding the cellular target will be contained within the rescued T7 clones and can be directly transferred to a suitable expression vector (e.g. *E. coli* BL21). However, if only a fragment of the gene is available, the full-length gene will first need to be recovered from the source organism using suitable primers.

Surface plasmon resonance (SPR)⁷⁶ is a mature technique that can provide an estimate of the protein–small molecule association and dissociation kinetics using the biotinylated natural product immobilised on a streptavidin-coated SPR chip. However, the process of derivatising and immobilising the natural product may alter the interaction kinetics and the values obtained may not be indicative of the true *in vivo* situation. MicroScale Thermophoresis (MST) is a newer technique that relies on the directed movement of particles in a microscopic temperature gradient. Any change in the hydration shell of biomolecules due to changes in their structure/conformation results in a relative change in movement along the temperature gradient, which can be used to determine binding affinities. Importantly, MST allows measurement of interactions directly in solution without the need of immobilisation to a surface. Analytical ultracentrifugation (AU),⁷⁷ which relies on concentration gradients produced by the differential sedimentation of species with different molecular weights, can also be used to quantify protein–small molecule interactions without the need to derivatise the small molecule. A variety of other label-free biosensors are available for quantifying binding interactions, although these vary greatly in terms of cost, maturity and accessibility.⁷⁸

The enthalpy and stoichiometry of the protein–small molecule interaction can be measured using isothermal titration calorimetry (ITC),⁷⁹ without the need to modify the protein or natural product. X-ray diffraction of the natural product co-

crystallised with the protein target can also provide valuable information regarding the site and mechanism of binding, which can facilitate subsequent rational design of therapeutic agents based on the natural product scaffold. However, this can also be a time-consuming and costly process and is best tackled in collaboration with an experienced crystallographer. Finally, if the target is an enzyme, a K_i value can be determined against a known substrate.

The most challenging (and important) step in the validation process is determining the overall significance of the protein–small molecule interaction in the context of the whole organism – only then can the true mode of action of the natural product be inferred. Biochemical pathways are highly interconnected and perturbing a single protein in a pathway can result in far more complex effects than anticipated. On the other hand, functional redundancy in cellular systems may render the protein–small molecule interaction irrelevant, with closely related proteins taking over the function of the inhibited cellular target. Quantitative transcriptomics⁸⁰ (e.g. microarray, RNA-seq) and proteomics⁸¹ (e.g. SILAC, ICAT, iTRAQ) can be useful in comparing overall differences in levels of transcription and translation caused by exposure to the natural product, although natural variability in expression levels can make the data obtained difficult to interpret. RNA interference (RNAi)⁸² can also be used to observe phenotypic consequences of silencing the gene(s) encoding the cellular target. In most cases, the approach taken to determine mode of action will need to be specifically tailored to each natural product. Clearly, effective multidisciplinary collaborations are critical to the success of such endeavours.

Limitations

While T7 phage display is a powerful tool that can rapidly link a natural product to its cellular target, the technique does have some significant limitations that must be evaluated before embarking on a biopanning campaign. Post-translational modifications of T7 phage-displayed eukaryotic proteins, including *N*- and *C*-terminal tailoring, phosphorylation, glycosylation and prenylation, are not possible in a bacterial host. The claimed upper size limit for proteins displayed on the T7Select vector is 1200 amino acids (~132 kDa; 3600 bp DNA insert),⁸³ although in practise the value is likely to be somewhat smaller. Very large proteins, multi-domain proteins and membrane-associated proteins are unlikely to be displayed correctly on the surface of T7 phage, although fragments, single protein domains and extracellular binding loops may retain the ability to bind to the natural product/small molecule. This has been shown by Austin for the α -domain of domain of F1-ATP synthase⁸⁴ and by us for a single protein from the human ribosome.⁷⁴ Indeed, identifying a protein fragment or domain that retains the ability to bind to the natural product can provide valuable information regarding the site and mechanism of binding.

Statistically, two-thirds of the foreign genes in most T7 phage libraries are inserted out of frame with the phage coat protein gene, leading to the display of random peptide fragments that are not encoded by the attached gene. If non-directional cloning methods are used, half of the clones will also have the foreign genes inserted backwards. Consequently, only one in six clones in a typical T7 library has the potential to display a functional protein. This number is further decreased by proteolysis of displayed peptides, poor expression in the bacterial host due to differences in codon usage, incorrect processing of the phage capsule, failure of the inserted gene product to fold correctly without the mammalian endoplasmic reticulum or chaperone proteins, and very much less likely, toxicity of the expressed protein to the host. Some of these issues are at least theoretically addressed with the Rosetta-gami cell lines, which provide thioredoxin reductase and glutathione reductase to create a cytoplasmic environment compatible with the formation of disulfide bonds, and which also supply tRNAs for codons rarely used in *E. coli*. The addition of a protease inhibitor cocktail can reduce proteolysis of displayed peptides, although this may cause problems if the target of the natural product is itself a protease.

Many of the theoretical limitations of T7 phage display described above are actually inconsequential in practice due to the very large number of phage particles that can be obtained in a small volume. For example, a typical T7 phage lysate contains 10^{12} clones/mL, so even if only 0.1% of the clones are displaying useful proteins, this is still more than enough capacity to display a typical eukaryotic genome with excellent coverage. Another significant advantage of T7 phage display is its speed, with the entire biopanning procedure outlined above taking only 1 week to complete. Consequently, even if T7 phage display is unable to identify the cellular target of a particular natural product, this conclusion can be reached very quickly and the researcher can move on to the next technique or next compound with very little time lost.

Conclusions

Chemical proteomics approaches, especially affinity chromatography, dominate the methods used to identify natural product binding proteins. However, it is often difficult to minimise contamination due to non-specific binding of proteins to an affinity matrix resulting in a large amount of background and the inevitable isolation of the most common binding partners instead of the most avid binding partner. Reverse chemical proteomics approaches, such as phage display, are relatively mature and powerful techniques to correlate natural products to their binding proteins. While not widely used, T7-phage display is relatively easy to adopt into a chemistry lab and has a number of distinct advantages over forward chemical proteomics and other reverse chemical proteomics approaches.¹² These include a robust expression system, rapid life cycle, ease of use and that all analyses are performed at the gene level through DNA sequencing. In this tutorial we have aimed to provide clear instructions specifically

for chemists who are interested in identifying the protein target(s) of their natural products. We would also be happy to collaborate with groups around the world to achieve the same result.

Acknowledgements

This work was supported by the Australian Research Council (DP130103281, FT130100142) and Macquarie University.

Notes and references

- G. M. Cragg and D. J. Newman, *Biochim. Biophys. Acta*, 2013, **1830**, 3670-3695.
- D. J. Newman and G. M. Cragg, in *Natural Products: Discourse, Diversity, and Design*, eds. A. Osbourn, R. J. Goss and G. T. Carter, John Wiley & Sons, Inc, Hoboken, NJ, USA, 2014, ch. 21, pp. 397-411.
- K. Choi, J. Hong, C.-O. Lee, D.-K. Kim, C. J. Sim, K. S. Im and J. H. Jung, *J. Nat. Prod.*, 2004, **67**, 1186-1189.
- X.-L. Hou, Z. Yang, K.-S. Yeung and H. N. C. Wong, *Prog. Heterocycl. Chem.*, 2005, **17**, 142-171.
- J. W. Blunt, B. R. Copp, M. H. G. Munro, P. T. Northcote and M. R. Prinsep, *Nat. Prod. Rep.*, 2006, **23**, 26-78.
- Y. Liu, S. Zhang and P. J. M. Abreu, *Nat. Prod. Rep.*, 2006, **23**, 630-651.
- N. Dixon, L. S. Wong, T. H. Geerlings and J. Micklefield, *Nat. Prod. Rep.*, 2007, **24**, 1288-1310.
- Y. Liu, L. Wang, J. H. Jung and S. Zhang, *Nat. Prod. Rep.*, 2007, **24**, 1401-1429.
- S. Hussain, M. Slevin, S. Matou, N. Ahmed, M. I. Choudhary, R. Ranjit, D. West and J. Gaffney, *Angiogenesis*, 2008, **11**, 245-256.
- C. Zhang and Y. Liu, *J. Nat. Med.*, 2015, **69**, 255-266.
- L. Wittgenstein, *Ann. Nat. Phil.*, 1921, **14**, 185-262.
- P. Karuso, in *Comprehensive Natural Products II*, eds. H.-W. Liu and L. Mander, Elsevier, Oxford, 2010, vol. 9, ch. 14, pp. 513-567.
- A. M. Piggott and P. Karuso, *Comb. Chem. High Throughput Screen.*, 2004, **7**, 607-630.
- A. M. Piggott and P. Karuso, *Mar. Drugs*, 2005, **3**, 36-63.
- H. J. Jung and H. J. Kwon, *Arch. Pharmacol. Res.*, 2015, **38**, 1627-1641.
- E. W. Taylor, *J. Cell Biol.*, 1965, **25**, 145-160.
- G. G. Borisy and E. W. Taylor, *J. Cell Biol.*, 1967, **34**, 525-533.
- G. G. Borisy and E. W. Taylor, *J. Cell Biol.*, 1967, **34**, 535-548.
- H. Mohri, *Nature*, 1968, **217**, 1053-1054.
- M. M. Merker and R. E. Handschumacher, *J. Immunol.*, 1984, **132**, 3064-3070.
- R. E. Handschumacher, M. W. Harding, J. Rice, R. J. Drugge and D. W. Speicher, *Science*, 1984, **226**, 544-547.
- N. Takahashi, T. Hayano and M. Suzuki, *Nature*, 1989, **337**, 473-475.
- M. Tropschug, I. B. Barthelmess and W. Neupert, *Nature*, 1989, **342**, 953-955.
- J. J. Siekierka, S. H. Y. Hung, M. Poe, C. S. Lin and N. H. Sigal, *Nature*, 1989, **341**, 755-757.
- J. J. Siekierka, M. J. Staruch, S. H. Y. Hung and N. H. Sigal, *J. Immunol.*, 1989, **143**, 1580-1583.
- M. W. Harding, A. Galat, D. E. Uehling and S. L. Schreiber, *Nature*, 1989, **341**, 758-760.
- J. Liu, J. D. Farmer, Jr., W. S. Lane, J. Friedman, I. Weissman and S. L. Schreiber, *Cell*, 1991, **66**, 807-815.
- C. M. Crews, J. L. Collins, W. S. Lane, M. L. Snapper and S. L. Schreiber, *J. Biol. Chem.*, 1994, **269**, 15411-15414.

- 29 C. M. Crews, W. S. Lane and S. L. Schreiber, *Proc. Nat. Acad. Sci. USA*, 1996, **93**, 4316-4319.
- 30 M. B. Potts, E. A. McMillan, T. I. Rosales, H. S. Kim, Y.-H. Ou, J. E. Toombs, R. A. Brekken, M. D. Minden, J. B. MacMillan and M. A. White, *Nat. Chem. Biol.*, 2015, **11**, 401-408.
- 31 M. J. Towle, K. A. Salvato, J. Budrow, B. F. Wels, G. Kuznetsov, K. A. Aalfs, S. Welsh, W. Zheng, B. M. Seletsky, M. H. Palme, G. J. Habgood, L. A. Singer, L. V. DiPietro, Y. Wang, J. J. Chen, D. A. Quincy, A. Davis, K. Yoshimatsu, Y. Kishi, M. J. Yu and B. A. Littlefield, *Cancer Res.*, 2001, **61**, 1013-1021.
- 32 N. Sin, L. Meng, M. Q. W. Wang, J. J. Wen, W. G. Bornmann and C. M. Crews, *Proc. Nat. Acad. Sci. USA*, 1997, **94**, 6099-6103.
- 33 W. T. Lowther, D. A. McMillan, A. M. Orville and B. W. Matthews, *Proc. Nat. Acad. Sci. USA*, 1998, **95**, 12153-12157.
- 34 Y. Miyake, Y. Kozutsumi, S. Nakamura, T. Fujita and T. Kawasaki, *Biochem. Biophys. Res. Commun.*, 1995, **211**, 396-403.
- 35 J. K. Chen, W. S. Lane and S. L. Schreiber, *Chem. Biol.*, 1999, **6**, 221-235.
- 36 T. Böttcher, M. Pitscheider and S. A. Sieber, *Angew. Chem. Int. Ed.*, 2010, **49**, 2680-2698.
- 37 J. Kreuzer, N. C. Bach, D. Forler and S. A. Sieber, *Chem. Sci.*, 2015, **6**, 237-245.
- 38 W. Zheng, G. Li and X. Li, *Arch. Pharm. Res.*, 2015, **38**, 1661-1685.
- 39 K. M. Specht and K. M. Shokat, *Curr. Opin. Cell Biol.*, 1999, **14**, 155-159.
- 40 K. Nakanishi, T. Sakiyama, Y. Kumada, K. Imamura and H. Imanaka, *Curr. Proteomics*, 2008, **5**, 161-175.
- 41 R. Rangel, L. Guzman-Rojas, L. G. Le Roux, F. I. Staquicini, H. Hosoya, E. M. Barbu, M. G. Ozawa, J. Nie, K. Dunner, Jr., R. R. Langley, E. H. Sage, E. Koivunen, J. G. Gelovani, R. R. Lobb, R. L. Sidman, R. Pasqualini and W. Arap, *Nat. Commun.*, 2012, **3**, 788.
- 42 J. K. Scott and L. Craig, *Curr. Opin. Biotechnol.*, 1994, **5**, 40-48.
- 43 M. Szardenings, *J. Recept. Sig. Transd.*, 2003, **23**, 307-349.
- 44 B. Westerlund-Wikstrom, *Int. J. Med. Microbiol.*, 2000, **290**, 223-230.
- 45 P. Monaci, F. Bartoli, G. Di Zenzo, M. Nuzzo and L. Urbanelli, *Tumor Targeting*, 1999, **4**, 129-142.
- 46 M. S. Packer and D. R. Liu, *Nat. Rev. Genet.*, 2015, **16**, 379-394.
- 47 A. H. Badran and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2015, **24**, 1-10.
- 48 M. D. Lane and B. Seelig, *Curr. Opin. Chem. Biol.*, 2014, **22**, 129-136.
- 49 A. Plückthun, *Methods Mol. Biol.*, 2012, **805**, 3-28.
- 50 L. Yuan, I. Kurek, J. English and R. Keenan, *Microbiol. Mol. Biol. Rev.*, 2005, **69**, 373-392.
- 51 C. J. Buchholz, L. J. Duerner, S. Funke and I. C. Schneider, *Comb. Chem. High Throughput Screen.*, 2008, **11**, 99-110.
- 52 E. T. Boder and K. D. Wittrup, *Methods Enzymol.*, 2000, **328**, 430-444.
- 53 S. A. Gai and K. D. Wittrup, *Curr. Opin. Struct. Biol.*, 2007, **17**, 467-473.
- 54 E. M. Phizicky and S. Fields, *Microbiol. Rev.*, 1995, **59**, 94-123.
- 55 A. Rothe, R. J. Hosse and B. E. Power, *FASEB J.*, 2006, **20**, 1599-1610.
- 56 S. S. Sidhu, W. J. Fairbrother and K. Deshayes, *ChemBioChem*, 2003, **4**, 14-25.
- 57 G. P. Smith, *Science*, 1985, **228**, 1315-1317.
- 58 L. S. Jespers, J. H. Messens, A. De Keyser, D. Eeckhout, I. Van Den Brande, Y. G. Gansemans, M. J. Lauwereys, G. P. Vlasuk and P. E. Stanssens, *Bio/Technology*, 1995, **13**, 378-382.
- 59 R. Cramer and M. Suter, *Gene*, 1993, **137**, 69-75.
- 60 A. Rosenberg, K. Griffin, F. W. Studier, M. McCormick, J. Berg, R. Novy and R. Mierendorf, *inNovations*, 1996, **6**, 1-6.
- 61 P. P. Sche, K. M. McKenzie, J. D. White and D. J. Austin, *Chem. Biol.*, 1999, **6**, 707-716.
- 62 K. P. Wilson, M. M. Yamashita, M. D. Sintchak, S. H. Rotstein, M. A. Murcko, J. Boger, J. A. Thomson, M. J. Fitzgibbon, J. R. Black and M. A. Navia, *Acta Crystallogr. D Biol. Crystallogr.*, 1995, **51**, 511-521.
- 63 P. P. Sche, K. M. McKenzie, J. D. White and D. J. Austin, *Chem. Biol.*, 2001, **8**, 399-400.
- 64 A. M. Piggott, A. M. Kriegel, R. D. Willows and P. Karuso, *Bioorg. Med. Chem.*, 2009, **17**, 6841-6850.
- 65 H. M. L. Davies and D. Morton, *Chem. Eng. News*, 2015, **93**, 32-33.
- 66 O. Robles and D. Romo, *Nat. Prod. Rep.*, 2014, **31**, 318-334.
- 67 J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. Palazzolo, J. W. Lehmann, G. F. Morehouse and M. D. Burke, *Science*, 2015, **347**, 1221-1226.
- 68 C. A. Lewis and S. J. Miller, *Angew. Chem. Int. Ed.*, 2006, **118**, 5744-5747.
- 69 I. Sattler, S. Grabley and R. Thiericke, in *Drug Discovery from Nature*, eds. S. Grabley and R. Thiericke, Springer-Verlag, Berlin, 1999,, ch. 11, pp. 191-214.
- 70 C.-Y. Zhou, J. Li, S. Peddibhotla and D. Romo, *Org. Lett.*, 2010, **12**, 2104-2107.
- 71 J. Yamaguchi, A. D. Yamaguchi and K. Itami, *Angew. Chem. Int. Ed.*, 2012, **51**, 8960-9009.
- 72 S. M. Paradine, J. R. Griffin, J. Zhao, A. L. Petronico, S. M. Miller and M. C. White, *Nat. Chem.*, 2015, **7**, 987-994.
- 73 J. M. Howell, K. Feng, J. R. Clark, L. J. Trzepkowski and M. C. White, *J. Am. Chem. Soc.*, 2015, **137**, 14590-14593.
- 74 A. M. Piggott and P. Karuso, *ChemBioChem*, 2008, **9**, 524-530.
- 75 N. Bindslev, *Drug Acceptor Interactions*, CoAction Publishing, Stockholm, Sweden, 2008.
- 76 J. Homola, *Chem. Rev.*, 2008, **108**, 462-493.
- 77 M. Arkin and J. D. Lear, *Anal. Biochem.*, 2001, **299**, 98-107.
- 78 M. A. Cooper, ed., *Label-Free Biosensors: Techniques and Applications*, Cambridge University Press, Cambridge, 2009.
- 79 J. E. Ladbury, G. Klebe and E. Freire, *Nat. Rev. Drug Discov.*, 2010, **9**, 23-27.
- 80 F. Ozsolak and P. M. Milos, *Nat. Rev. Genet.*, 2011, **12**, 87-98.
- 81 B. Domon and R. Aebersold, *Nat. Biotechnol.*, 2010, **28**, 710-721.
- 82 M. Ghildiyal and P. D. Zamore, *Nat. Rev. Genet.*, 2009, **10**, 94-108.
- 83 Novagen T7Select System Manual TB178 Rev.D 0311JN, 2011, Merck Biosciences, 1-24.
- 84 S. N. Savinov and D. J. Austin, *Comb. Chem. High Throughput Screen.*, 2001, **4**, 593-597.
- 85 J. Izaguirre-Carbonell, H. Kawakubo, H. Murata, A. Tanabe, T. Takeuchi, T. Kusayanagi, S. Tsukuda, T. Hirakawa, K. Iwabata, Y. Kanai, K. Ohta, M. Miura, K. Sakaguchi, S. Matsunaga, H. Sahara, S. Kamisuki and F. Sugawara, *Sci. Rep.*, 2015, **5**, 15136.
- 86 N. H. Kim, N. B. Pham, R. J. Quinn, J. S. Shim, H. Cho, S. M. Cho, S. W. Park, J. H. Kim, S. H. Seok, J.-W. Oh and H. J. Kwon, *Int. J. Biol. Sci.*, 2015, **11**, 1088-1099.
- 87 T. Shishido, M. Hachisuka, K. Ryuzaki, Y. Miura, A. Tanabe, Y. Tamura, T. Kusayanagi, T. Takeuchi, S. Kamisuki, F. Sugawara and H. Sahara, *Eur. J. Immunol.*, 2014, **44**, 3220-3231.
- 88 S. Tsukuda, T. Kusayanagi, E. Umeda, C. Watanabe, Y.-t. Tosaki, S. Kamisuki, T. Takeuchi, Y. Takakusagi, I. Shiina and F. Sugawara, *Bioorg. Med. Chem.*, 2013, **21**, 311-320.
- 89 X. Yu, P. Zhao, L. Zhang and Y. Zhang, *Biomed. Chromatogr.*, 2013, **27**, 1574-1579.
- 90 J. Woolard, W. Vousden, S. J. Moss, A. Krishnakumar, M. V. R. Gammons, D. G. Nowak, N. Dixon, J. Micklefield, A.

- Spannhoff, M. T. Bedford, M. A. Gregory, C. J. Martin, P. F. Leadlay, M. Q. Zhang, S. J. Harper, D. O. Bates and B. Wilkinson, *Chem. Sci.*, 2011, **2**, 273-278.
- 91 D. Manita, Y. Toba, Y. Takakusagi, Y. Matsumoto, T. Kusayanagi, K. Takakusagi, S. Tsukuda, K. Takada, Y. Kanai, S. Kamisuki, K. Sakaguchi and F. Sugawara, *Bioorg. Med. Chem.*, 2011, **19**, 7690-7697.
- 92 H. J. Jung, J. S. Shim, J. Lee, Y. M. Song, K. C. Park, S. H. Choi, N. D. Kim, J. H. Yoon, P. T. Mungai, P. T. Schumacker and H. J. Kwon, *J. Biol. Chem.*, 2010, **285**, 11584-11595.
- 93 Q.-L. He, H. Jiang, F. Zhang, H.-B. Chen and G.-L. Tang, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 3995-3998.
- 94 H. Kim, L. Deng, X. Xiong, W. D. Hunter, M. C. Long and M. C. Pirrung, *J. Med. Chem.*, 2007, **50**, 3423-3426.
- 95 W. Boehmerle, U. Splittgerber, M. B. Lazarus, K. M. McKenzie, D. G. Johnston, D. J. Austin and B. E. Ehrlich, *Proc. Nat. Acad. Sci. USA*, 2006, **103**, 18356-18361.
- 96 J. S. Shim, J. Lee, H.-J. Park, S.-J. Park and H. J. Kwon, *Chem. Biol.*, 2004, **11**, 1455-1463.
- 97 Y. Jin, J. Yu and Y. G. Yu, *Chem. Biol.*, 2002, **9**, 157-162.

ARTICLE

Natural Product Reports

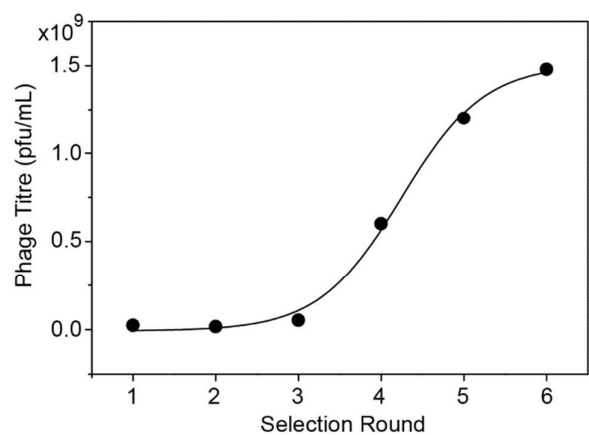
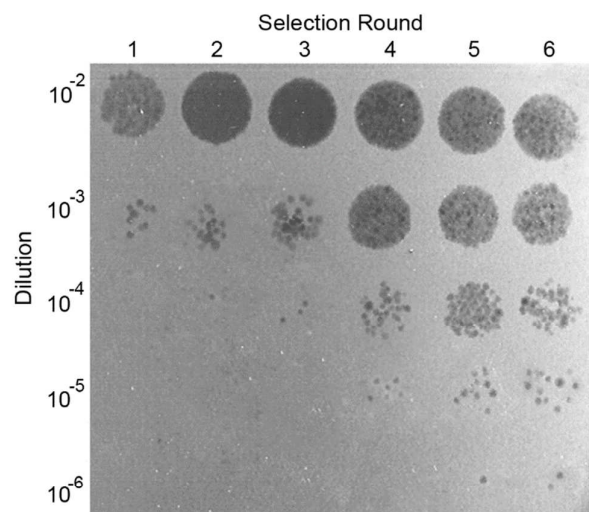


Figure 2. Titrating T7 phage following 6 rounds of affinity selection

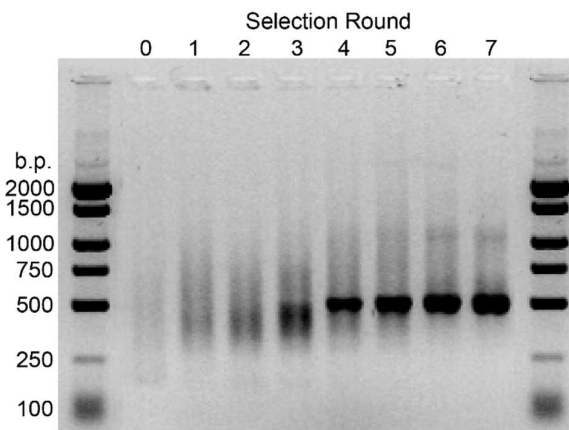


Figure 3. Gel electrophoresis of PCR products amplified from T7 phage sublibraries following 7 rounds of affinity selection.

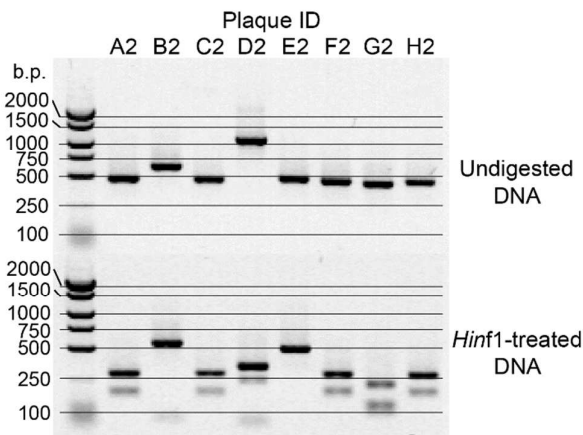


Figure 4. Gel electrophoresis of undigested and *Hinf*1-digested PCR products amplified from individual T7 phage plaques following 6 rounds of affinity selection.

DNA Sequence:

GGGGACC **GAATTC**AAGCGAGAGACAGAGAGAAGATATAGAG**ATGGGTG**
TGGAGAAGCAAGTCATCAGACCCGGCAACGGTCCCAAACCCGCTCCGG
GCCAAACCGTCACCGTCCATTGCACCGGATTCGGGAAAGATGGTGATC
TATCCCAAGAAGTCTGGAGTACAAAGGACGAGGGACAAAAACCTTCT
CGTTCCAAATCGGTAAGGTGCTGTGATCAAAGGATGGGATGAAGGCG
TTATCGGAATGCAAAATGGAGAGGTTGCTCGCTTCGGGTGCTCATCGG
ATTACGCATATGGCGTGGTGGATTTCCGGCATGGGGATCCAGCCAA
ACTCTGTTCTCGACTTTGAAATCGAAGTACTGAGCGTGCAGTAATGAT
 CTCCCCAGCTCCTAAACTGTGGGTAGATTA

BLAST Search:

Arabidopsis thaliana peptidyl-prolyl isomerase FKBP12 mRNA, complete cds
 Sequence ID: [ref|NM_125831.2|](#) Length: 632 Number of Matches: 1

Range 1: 58 to 454		GenBank	Graphics	▼ Next Match	▲ Previous
Score	Expect	Identities	Gaps	Strand	
734 bits(397)	0.0	397/397(100%)	0/397(0%)	Plus/Plus	
Query 18	GAGAGACAGAGAGAAGATATAGAGATGGGTGGAGAAAGTCATCAGACCCGGCAAC	77			
Sbjct 58	GAGAGACAGAGAGAAGATATAGAGATGGGTGGAGAAAGTCATCAGACCCGGCAAC	117			
Query 78	GATCCCAAAACCCGCTCCGGGCAAAACCGTCACCGTCCATTGCACCGGATTCGGGAAAGAT	137			
Sbjct 118	GATCCCAAAACCCGCTCCGGGCAAAACCGTCACCGTCCATTGCACCGGATTCGGGAAAGAT	177			
Query 138	GATGATCTATCCCAAGTCTGGAGTACAAAGGACGAGGGACAAAAACCTTTCTCGTTC	197			
Sbjct 178	GATGATCTATCCCAAGTCTGGAGTACAAAGGACGAGGGACAAAAACCTTTCTCGTTC	237			
Query 198	CAATCGGTAAGGTGCTGTGATCAAGGATGGGATGAAGCGTTATCGGAATGCAAAAT	257			
Sbjct 238	CAATCGGTAAGGTGCTGTGATCAAGGATGGGATGAAGCGTTATCGGAATGCAAAAT	297			
Query 258	GGAGAGTTGCTCGCTTCGGGTGCTCATCGGATTACCGATATGGCGCTGGTGGATTCGG	317			
Sbjct 298	GGAGAGTTGCTCGCTTCGGGTGCTCATCGGATTACCGATATGGCGCTGGTGGATTCGG	357			
Query 318	GCATGGGGATCCAGCCAAACTCTGTTCTCGACTTTGAAATCGAAGTACTGAGCGTGCAG	377			
Sbjct 358	GCATGGGGATCCAGCCAAACTCTGTTCTCGACTTTGAAATCGAAGTACTGAGCGTGCAG	417			
Query 378	TAATGATCTCCCCAGCTCCTAAACTGTGGGTAGATTA	414			
Sbjct 418	TAATGATCTCCCCAGCTCCTAAACTGTGGGTAGATTA	454			

Translated Proteins:

Frame 1: **NSSERQREDIEMGVKQVIRPGNGPKPAPGQTVTVHCTG**
FGKDGDLKQKFWSTKDEGQKPFQIGKGAVIKGWDEGV
IGMQIGEARLRCSDDYAYGAGGFPAWGIQPNVLDLFEI
EVLSVQ

Frame 2: IQARDREKI

Frame 3: FKRETTERRYRDGCGEASHQTRQRSQTRSGPNRHRPLHRI
 RERW

Figure 5. DNA sequencing of T7 phage plaques. In this example, a clone displaying FKBP12 has been rescued from an *Arabidopsis thaliana* T7 library using a biotinylated FK506 probe. The rescued clone is full-length, in the correct reading frame and contains a leader sequence of 11 amino acids at the N-terminus of the protein. The BLAST search shows a perfect match with the sequence in GenBank.

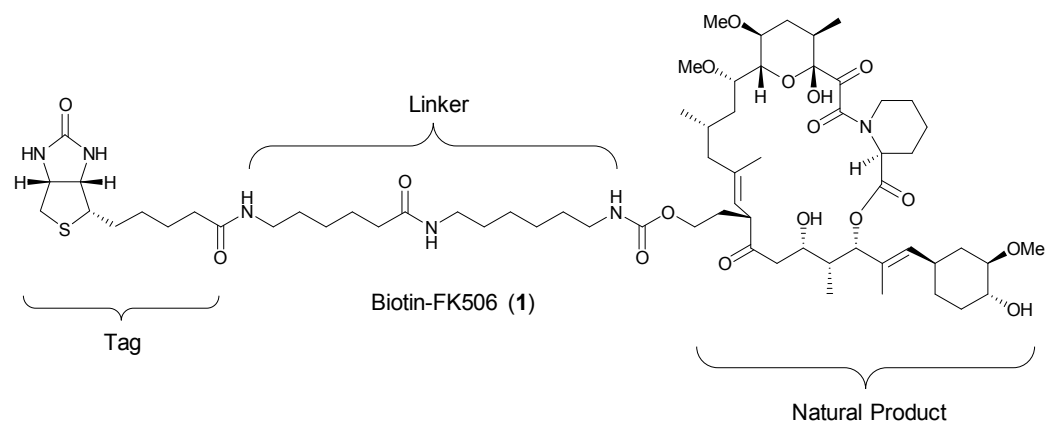
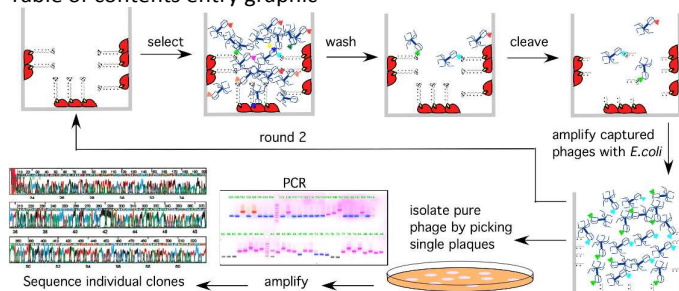


Figure 1. Components of a natural product affinity probe, as exemplified by biotin-FK506 (1).⁶¹

Table 1. Cellular targets for natural products identified using T7 phage display

Year	Natural product	Target identified	Ref.
2015	sulfoquinovosyl diacylglyceride	focal adhesion kinase (FAK)	85
2015	(-)- β -O-methylsynephrine	nucleoporin 153 kDa (NUP153)	86
2014	pyrenocine B	EpsinR	87
2013	ridaifen B	Grb10 interacting GYF protein 2 (GIGYF2)	88
2013	doxorubicin	methyl-CpG binding protein 2 (MeCP2)	89
2011	borrelidin	spliceosome-associated protein formin binding protein 21 (FBP21)	90
2011	camptothecin	human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1)	91
2010	terpestacin	ubiquinol-cytochrome c reductase binding protein (UQCRB)	92
2009	FK506	FKBP2, FKBP3, <i>Arabidopsis thaliana</i> FKBP, <i>Pseudomonas stutzeri</i> FKBP	64
2008	cyclosporine A	cyclophilin A (CypA), cyclophilin B (CypB)	93
2008	kahalalide F	ribosomal protein S25 (RPS25)	74
2007	demethylasterriquinone B1	glyceraldehyde 3-phosphate dehydrogenase (GADPH)	94
2006	taxol	neuronal Ca ²⁺ sensor 1 (NCS-1)	95
2004	HBC	calmodulin	96
2002	doxorubicin	human nucleolar phosphoprotein 140 (hNopp140)	97
1999	FK506	FKBP1a	61

Table of contents entry graphic



A description of the T7 phage biopanning procedure is provided with tips and advice suitable for setup in a chemistry laboratory.