Simplified proteomics approach to discover protein–ligand interactions

Youngil Chang,¹ Jonathan P. Schleich,² Ross A. VerHeul,² and Chiwook Park^{1,2,3*}

¹Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, Indiana 47907

²Interdisciplinary Life Science Graduate Program, Purdue University, West Lafayette, Indiana 47907

³Bindley Bioscience Center, Purdue University, West Lafayette, Indiana 47907

Received 17 April 2012; Accepted 15 June 2012

DOI: 10.1002/pro.2112

Published online 25 June 2012 proteinscience.org

Abstract: Identifying targets of biologically active small molecules is an essential but still challenging task in drug research and chemical genetics. Energetics-based target identification is an approach that utilizes the change in the conformational stabilities of proteins upon ligand binding in order to identify target proteins. Different from traditional affinity-based capture approaches, energetics-based methods do not require any labeling or immobilization of the test molecule. Here, we report a surprisingly simple version of energetics-based target identification, which only requires ion exchange chromatography, SDS PAGE, and minimal use of mass spectrometry. The complexity of a proteome is reduced through fractionation by ion exchange chromatography. Urea-induced unfolding of proteins in each fraction is then monitored by the significant increase in proteolytic susceptibility upon unfolding in the presence and the absence of a ligand. Proteins showing a different degree of unfolding with the ligand are identified by SDS PAGE followed by mass spectrometry. Using this approach, we identified ATP-binding proteins in the *Escherichia coli* proteome. In addition to known ATP-binding proteins, we also identified a number of proteins that were not previously known to interact with ATP. To validate one such finding, we cloned and purified phosphoglyceromutase, which was not previously known to bind ATP, and confirmed that ATP indeed stabilizes this protein. The combination of fractionation and pulse proteolysis offers an opportunity to investigate protein–drug or protein–metabolite interactions on a proteomic scale with minimal instrumentation and without modification of a molecule of interest.

Keywords: target identification; ligand binding; protein stability; proteolysis; proteomics; ATP

Introduction

Drugs are frequently discovered by their efficacies in cellular or animal models without knowing their molecular targets. When the targets are unknown, it is

not possible to understand how a drug manifests its efficacy. Identification of the molecular targets of a drug provides critical information to decipher the pharmacological mechanism of the drug action.^{1–6} The importance of target identification is greater than ever because the development of chemical genomics and the advances in high-throughput screening have facilitated the discovery of biologically active small molecules at an unprecedented speed.^{7,8}

Target identification is informative not only for drugs with an unknown mode of actions but also for drugs with known targets. It is common that drugs interact with multiple targets, and promiscuous binding may have a critical role in pharmacological

Additional Supporting Information may be found in the online version of this article.

Jonathan P. Schleich's current address is Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232.

Grant sponsor: National Institute of Health; Grant number: GM097528.

*Correspondence to: Chiwook Park, 575 Stadium Mall Drive, Purdue University, West Lafayette, IN 47907.
E-mail: chiwook@purdue.edu

mechanisms of some drugs.⁹⁻¹¹ For example, the anticancer drug imatinib is known to bind targets other than its intended target BCR-ABL kinase. Its ability to bind other targets has made it effective as a therapeutic for other cancers.^{4,12} Moreover, side effects of a drug frequently result from interactions with “off-targets.”¹³ Information on off-targets is valuable in understanding the molecular basis of side effects, which could be useful for the development of new drugs with fewer side effects.¹⁴

Energetics-based target identification is a proteomic approach to identify drug-binding proteins from a mixture of proteins, such as a cell lysate, by exploiting the effect of drug binding on the conformational stability of its binding proteins.¹⁵⁻¹⁷ When a protein binds to a drug, the thermodynamic stability of the protein is increased according to the dissociation free energy of the complex. Also, binding to a drug frequently slows unfolding of the protein. Therefore, one can identify drug-binding proteins by monitoring the change in the thermodynamic stability or the unfolding kinetics of proteins in a cell lysate in the presence of a drug. Energetics-based target identification has many advantages over traditional approaches such as affinity chromatography. The use of a drug without any modification eliminates potential interference with binding to the targets from crosslinking of the drug to a solid support. Also, because the binding occurs in the solution phase, one can control the stringency of the screen easily by changing the concentration of the drug, which is not possible in affinity chromatography. The technical challenge in the application of this principle lies in the development of the methodology to monitor the conformational stability of a multitude of proteins on a proteomic scale.

Recently, we have demonstrated an approach to monitor the change in protein stability upon ligand binding on a proteomic scale by pulse proteolysis.¹⁵ Pulse proteolysis determines the fraction of folded proteins in a mixture of folded and unfolded proteins by selective digestion of unfolded proteins.¹⁸ We have used pulse proteolysis to determine thermodynamic stabilities and unfolding kinetic constants of soluble and membrane proteins under various experimental conditions.¹⁸⁻²¹ Using pulse proteolysis and two-dimensional (2D) gel electrophoresis, we have also monitored urea-induced unfolding of many proteins in a cell lysate simultaneously.¹⁵ By identifying proteins stabilized in the presence of ATP, we have identified not only previously known ATP-binding proteins but also proteins that have not been reported to interact with ATP.¹⁵ This proof-of-principle experiment clearly demonstrates that pulse proteolysis is a powerful tool to discover protein–ligand interactions on a proteomic scale.

Here, we report a surprisingly simple version of energetics-based target identification, which employs

Figure 1. Schematic illustration of the experimental procedure. Proteins in a cell lysate are fractionated by chromatography. Each fraction is incubated in varying concentrations of urea with and without a test ligand. After pulse proteolysis, the remaining proteins in the reaction are analyzed by SDS PAGE. The targets of the test ligand (indicated by arrows) are identified by comparison of the electropherograms of the lanes with (+) and without (–) the ligand.

liquid chromatography and conventional one-dimensional SDS PAGE instead of 2D gel electrophoresis (Fig. 1). This modification significantly reduces the effort for the proteomic analysis after pulse proteolysis. Though having much less resolving power than 2D gel electrophoresis, conventional SDS PAGE is simpler, more quantitative, and more reproducible than 2D gel electrophoresis. The reduction of the complexity of the proteome by fractionation using liquid chromatography allows us to use conventional SDS PAGE to analyze the stabilization of proteins upon ligand binding in an efficient way with coverage comparable to 2D gel electrophoresis. To examine the validity and the reliability of this strategy, we identified ATP-binding proteins in the *E. coli* proteome.

Results

Fractionation of an *E. coli* lysate

To decrease the complexity of the *E. coli* proteome for target identification, the soluble fraction of an *E. coli*

Figure 2. Fractionation of an *E. coli* cell lysate. (A) An *E. coli* K12 lysate was fractionated into 15 fractions by using an anion exchange chromatography. The chromatogram shows the elution monitored by absorbance at 280 nm (A_{280} ; solid line) and the conductivity (dashed line). The location of each fraction is indicated by a vertical tick on the x-axis. (B) The protein contents of fractions were checked by SDS PAGE. The molecular weight markers (M) and the total lysate (L) were loaded as references. Fractions 14 and 15 were analyzed with a separate gel but shown together for comparison.

lysate was fractionated by anion exchange chromatography. The salt gradient was chosen after multiple trials to maximize the distribution of proteins into different fractions. The chromatogram (absorbance at 280 nm) and the salt profile (conductivity of the elution buffer) are shown in Figure 2(A). The eluant from 5 mL to 27 mL was separated into fifteen 1.5 mL fractions. Strong absorbance observed at the end of the elution (~25 mL) is likely due to nucleic acid, as the corresponding fractions were stained strongly by ethidium bromide on an agarose gel after electrophoresis (data not shown). Proteins that were eluted at the washing step were not used in this study.

SDS PAGE of the 15 fractions shows that proteins in the cell lysate are evenly fractionated [Fig. 2(B)]. Most fractions contain 30 to 50 easily identifiable protein bands on the SDS PAGE gels, suggesting that it would be feasible to monitor ATP binding to a significant number of endogenous proteins by pulse proteolysis and SDS PAGE. Fractions 14 and 15 do not contain much protein despite their strong absorbance at 280 nm, which is consistent with the agarose gel electrophoresis result that showed a significant amount of nucleic acid in the fractions (*vide supra*).

Discovering ATP-binding proteins with pulse proteolysis

To identify proteins whose conformational stabilities are affected by ATP, each fraction was incubated overnight in 0 to 4.0 M urea in the presence and absence of 1.0 mM ATP γ S and then digested by pulse proteolysis. We used ATP γ S instead of ATP to minimize the loss of the ligand by hydrolysis. We chose 1.0 mM ATP because the physiological concentration of ATP is on the millimolar scale.^{22,23} As the stabilization of a target protein by a ligand ($\Delta\Delta G^\circ$) is linearly proportional to $\ln([L]/K_d)$,¹⁵ where $[L]$ is the ligand concentration and K_d is the dissociation equilibrium constant, our experimental conditions ensure the identification of target proteins with K_d for ATP significantly smaller than 1.0 mM.

The SDS PAGE of fraction 7 after pulse proteolysis is shown in Figure 3(A) as a representative example. Some bands show clear changes in

Figure 3. Identification of ATP-binding proteins by pulse proteolysis. (A) As a representative example, the pulse proteolysis results of fraction 7 are shown. After incubation with (+) and without (-) 1.0 mM ATP γ S in varying concentrations of urea, the reactions were analyzed by SDS PAGE. Thermolysin is indicated by an arrow. (B) The lanes with and without ATP γ S in 1.0 M urea were scanned, and the absolute value of the binary logarithm of the intensity ratio ($|\log_2(\pm)|$) is plotted against the position on the gel image. Bands showing more than 50% change in intensity are selected as positives (marked with * on the plot and the gel). The 50% change is indicated as a dashed line. The identified proteins from the positive peaks are dnaK, guaB, tufB, and tufB (from left to right).

Table I. Identified Proteins that are Known to Bind ATP

Gene name	Description	Fraction ^a	Role of ATP
pckA	Phosphoenopyruvate carboxykinase	1	Substrate
glnS	Glutaminyl-tRNA synthetase	2	Substrate
pgk	Phosphoglycerate kinase	2, 3	Substrate
pykF	Pyruvate kinase I	3	Substrate
ackA	Acetate kinase A	3, 4	Substrate
aspS	Aspartyl-tRNA synthetase	4	Substrate
gnd	6-Phosphogluconate dehydrogenase	5	Inhibitor
gltA	Citrate synthase	5	Inhibitor
rho	Transcription termination factor rho	5, 6	Substrate
htpG	HSP90	7	Substrate
guaB	IMP dehydrogenase	7	Inhibitor
dnaK	HSP70	7–9	Substrate
lysU	Lysine-tRNA synthetase, inducible	10	Substrate
lysS	Lysine-tRNA synthetase, constitutive	10	Substrate
secA	Protein translocation subunit	11	Substrate
rpoC	RNA polymerase, β' subunit	12	Substrate
rpoB	RNA polymerase, β subunit	12, 13	Substrate
clpB	Protein disaggregation chaperone	12	Substrate
groL	GroEL	12	Substrate
deaD	RNA helicase dead	13	Substrate
hldD	ADP-glyceromanno-heptose 6-epimerase	15	Inhibitor

^a The gel images of samples after pulse proteolysis and the plots of $|\log_2(\pm)|$ for the identified proteins are provided in Supporting Information Fig. 1.

intensities from the incubation with ATP γ S. To analyze the results quantitatively, $|\log_2(\pm)|$, the absolute value of the binary logarithm of the band intensity ratios in the presence of ATP γ S (+) to that in the absence of ATP γ S (–), were calculated from the electropherograms of lanes on SDS PAGE gels. Bands showing more than 50% changes in their intensities were identified from the plot of $|\log_2(\pm)|$. When fraction 7 was incubated in 1.0 M urea, four bands showed more than 50% change in their band intensities from the incubation with 1.0 mM ATP γ S [Fig. 3(B)]. Proteins of the bands that satisfy this selection rule were identified by in-gel digestion and matrix-assisted laser desorption/ionization (MALDI) tandem time-of-flight (TOF) mass spectrometry.

We identified a total of 30 tentative ATP-binding proteins (Tables I and II; Supporting Information Fig. 1). According to EcoCyc, a functional-genomics database of *E. coli*, 21 out of 30 identified proteins are known to interact with ATP (Table I). Seventeen of the 21 proteins in Table I use ATP as a substrate. The remaining four proteins, 6-phosphogluconate dehydrogenase,²⁴ citrate synthase,²⁵ IMP dehydrogenase,²⁶ and ADP-glyceromanno-heptose 6-epimerase,²⁷ have been reported to be inhibited by ATP. The remaining 9 of the 30 identified proteins have not been previously shown to interact with ATP (Table II). Biochemical functions of the nine proteins in Table II are well known but do not require ATP. The fraction of known ATP-binding proteins out of the total identified proteins (21 out of 30, 70%) is quite similar to that of our previous identification of ATP binding proteins using 2D gel electrophoresis (70%).¹⁵

Three identified proteins, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), GroEL, and uridine phosphorylase show a decrease in band intensity in the presence of ATP γ S (Supporting Information Fig. 1). The apparent destabilization of GAPDH upon incubation with ATP has been also observed in our previous identification of ATP-binding proteins using pulse proteolysis and 2D gel electrophoresis.¹⁵ We also have shown that an equilibrium intermediate of GAPDH accumulates in the presence of ATP, which results in the apparent destabilization of the protein.²⁸ It was also reported previously that the C_m

Table II. Identified Proteins that are not Known to Bind ATP

Gene name ^a	Description	Fraction ^a	Ligands ^b
gapA	Glyceraldehyde-3-phosphate dehydrogenase A	1	NAD ⁺
gpmA	Phosphoglyceromutase	1	–
tnaA	Tryptophanase	2, 3	–
udp	Uridine phosphorylase	4	Uridine
acnB	Aconitate hydratase 2	5	–
ahpC	Alkyl hydroperoxide reductase subunit C	5	NAD ⁺ , FAD
tufB	EF-Tu	6–14	GTP
fusA	EF-G	10	GTP
aceE	Pyruvate dehydrogenase	11–14	NAD ⁺

^a The gel images of samples after pulse proteolysis and the plots of $|\log_2(\pm)|$ for the identified proteins are provided in Supporting Information Figure 1.

^b Only ligands structurally related to ATP are listed.

Figure 4. Effect of ATP on unfolding of phosphoglyceromutase. Phosphoglyceromutase was incubated in the absence (●) and presence (○) of 1.0 mM ATP in varying concentrations of urea, and then the unfolded protein was digested by pulse proteolysis. The apparent C_m values of the protein are 1.36M and 1.81M without and with ATP, respectively.

value (the transition midpoint of equilibrium unfolding in a chemical denaturant) of GroEL is decreased upon incubation with ADP.²⁹

Effect of ATP binding on unfolding of phosphoglyceromutase

To validate our identification of novel interactions with ATP, we chose phosphoglyceromutase as a test case. Phosphoglyceromutase catalyzes an isomerization step (3-phosphoglycerate $\rightleftharpoons$ 2-phosphoglycerate) in the glycolytic pathway. ATP binding to this enzyme has not been documented. We overexpressed and purified the enzyme after cloning the gene from *E. coli* genomic DNA. We then performed pulse proteolysis after incubating the purified enzyme in varying concentrations of urea both with and without 1.0 mM ATP. The result clearly showed that ATP increases the thermodynamic stability of the protein. The apparent C_m value is increased from 1.36 M to 1.81 M in the presence of 1.0 mM ATP (Fig. 4). This change in C_m is significant enough to confirm ATP binding by this protein. It is noteworthy that the thermodynamic stability of phosphoglyceromutase is dependent on the protein concentration because this protein is dimeric, and the C_m values we report here only reflect the apparent C_m 's under this experimental condition.³⁰

Discussion

Energetics-based target identification facilitates the discovery of protein–ligand interactions using the change in the conformational stabilities of the target proteins without any modification to the ligand. Previously, we have demonstrated the feasibility of energetics-based target identification by identifying ATP-binding proteins from an *E. coli* lysate with pulse proteolysis and 2D gel electrophoresis.¹⁵ The target

identification method reported here is based on the same principle as the previous approach; the change in conformational stabilities of proteins in the presence of a ligand is probed by pulse proteolysis. However, by reducing the complexity of the proteome by fractionation, we could identify target proteins with conventional SDS PAGE instead of 2D gel electrophoresis.

The use of conventional SDS PAGE instead of 2D gel electrophoresis significantly simplifies the application of energetics-based target identification. First, the use of SDS PAGE is much easier and faster than 2D gel electrophoresis. SDS PAGE takes only a fraction of time and effort that are necessary for 2D gel electrophoresis. Second, SDS PAGE is more reproducible than 2D gel electrophoresis. Even if samples are prepared in an identical way for 2D gel electrophoresis, the resulting 2D gels frequently show differences, which may lead to false positives in our screen. Third, the quantitative comparison of two samples is simpler with SDS PAGE than with 2D gel electrophoresis. The quantitative comparison of two 2D gels requires costly image analysis software or costly reagents for fluorescence tagging.³¹ In SDS PAGE, the comparison of samples is straightforward without the complication of gel-to-gel variation because samples are compared in neighboring lanes [Fig. 3(A)].

The simplicity of SDS PAGE allows us to assay ATP binding at several different urea concentrations simultaneously [Fig. 3(A)]. Energetics-based target identification relies on the change in the fraction of folded proteins upon ligand binding. To be identified by pulse proteolysis, a target needs to be unfolded in the absence of a ligand but folded in the presence of the ligand. Because this condition is satisfied within a certain range of urea concentration specific for each target protein, it is advantageous to perform target identification at several different urea concentrations.¹⁵ In our screen, targets were indeed identified at different urea concentrations (Supporting Information Fig. 1). The application of multiple urea concentrations significantly increases the chance of target identification.

One caveat of the current fractionation approach is that the fractionation by chromatography requires more sample than 2D gel electrophoresis. The approach is suitable for cell lysates that are easy to acquire. To use this approach with smaller volume of cell lysates, the chromatographic procedure may need to be optimized accordingly. As we can perform the proteolysis assays with less than 100 μ L of each fraction, a chromatographic method that can handle small sample volume (1–2 mL) and fraction volumes ($\sim$ 100 μ L) would be useful for a limited amount of a cell lysate.

How many proteins in the *E. coli* proteome are detectable with the combination of fractionation and SDS PAGE? Under the experimental conditions

employed in this study, 30 to 50 proteins were easily detectable by SDS PAGE from each fraction [Fig. 2(B)]. Considering that some highly abundant proteins show up in multiple fractions, we estimate the total number of proteins that we could observe in 15 fractions to be 300 to 500. This coverage is about 10% of the number of genes in *E. coli* genome (4288 genes). Because we have employed anion exchange chromatography, some proteins did not bind to the column and were removed during the washing step before the application of a salt gradient. If the unbound proteins are fractionated again by cation exchange chromatography, the coverage would be increased further and might be comparable to that of typical 2D gel electrophoresis (500–1000 proteins). As a method that is significantly less demanding in instrumentation than 2D gel electrophoresis or mass spectrometry-based proteomics approaches, the proteome coverage the method offers is reasonable.

In this study, we identified a total of 21 proteins that are already known to bind ATP (Table I). According to EcoCyc, ~6% *E. coli* genes are annotated as “ATP binding.” When the number of proteins tested in our assay is estimated at about 500, this number of identified proteins is consistent with the number of ATP-binding proteins we should expect in the fractions. We also identified nine proteins that have not been reported to bind ATP (Table II). It is possible that ATP may have a regulatory role in the function of these proteins or that some of these proteins are multifunctional³² and have an extra biochemical function that requires ATP as a substrate. Interestingly, six out of the nine proteins have biochemical functions that involve binding to a nucleoside (uridine), a nucleotide (GTP), or a cofactor with an adenosine moiety (NAD⁺ and FAD) (Table II). The similarity of these known ligands to ATP suggests that ATP may bind to the binding sites for their cognate ligands, which are not present in our assay. Common metabolites or cofactors similar to the molecule of interest might be added to the screen to their physiological concentrations to test whether or not the interactions between the targets and the ligand are likely to occur in cells also.

We have confirmed that one of these novel ATP-binding proteins, phosphoglyceromutase, is indeed stabilized in the presence of ATP (Fig. 4). The enzyme is not known to interact with ATP, and the catalytic function of the enzyme does not require ATP. However, the stabilization by ATP strongly suggests that the enzyme binds to ATP. We do not know the role of ATP for this enzyme, but it is possible that ATP or a metabolite structurally related to ATP acts as an allosteric regulator of the enzyme. This example showcases the power of our proteomics approach to identify protein–metabolite interactions as well as protein–drug interactions. By uncovering novel interactions between cellular proteins and metabolites, we may be

able to decipher previously unknown biochemical pathways or regulatory networks at the systems level.

Conclusion

We successfully demonstrated the effective combination of fractionation and pulse proteolysis to identify protein–ligand interactions on a proteomic scale. The approach we demonstrate here has many technical advantages compared with traditional approaches for target identification. (1) The test compound does not need to be modified for attachment to a media. (2) The stringency of the assay can be easily controlled by changing the concentration of the test molecule because the degree of target stabilization is dependent on the ligand concentration. (3) The method does not require any expensive isotope tags or cultures with isotope-containing media. (4) The required equipment is minimal. (5) Once fractions are prepared, multiple ligands can be tested in a parallel fashion with the same fractions. Drug or metabolite targets that are identified through this method will provide valuable information for the elucidation of drug mechanisms or protein–metabolite interaction networks in cells.

Materials and Methods

Preparation of K12 cell lysate and fractionation

E. coli K12 MG1655 was grown at 37°C with shaking until the absorbance at 600 nm reached 0.6. The cells were harvested by centrifugation, and the pellets were suspended in 20 mM Tris-HCl buffer (pH 8.0) containing 1.0 mM EDTA and 1.0 mM DTT. The cells were frozen and stored at –20°C overnight and thawed for sonication. The supernatant of the lysate was collected by centrifugation and incubated with 50 U/mL of benzonase and 5 mM of MgCl₂ at 25°C for 1.5 h to hydrolyze nucleic acids. EDTA and NaCl were added to the resulting lysate to the final concentration of 10 mM and 25 mM, respectively. Metabolites and digested nucleic acids in the lysate were removed by desalting with two linearly connected 5 mL HiTrap desalting columns (GE Healthcare; Piscataway, NJ) and 20 mM Tris-HCl buffer (pH 8.0) containing 1.0 mM EDTA, 1.0 mM DTT, and 25 mM NaCl as the elution buffer. The typical protein concentration of the lysate after the desalting step was ~1 mg/mL. The fractionation was performed with Source 15Q anion exchange column (GE Healthcare; Piscataway, NJ, column volume (CV) = 3.9 mL) with NaCl gradient. Fifteen milliliters of the desalted lysate (~15 mg of total protein) was loaded and eluted with three-step linear gradients (25–200 mM NaCl for 1.5 CV, 200–300 mM NaCl for 4 CV, and 300–700 mM for 2 CV) to evenly distribute the proteins into separate fractions. Fifteen 1.5-mL fractions were collected, and used immediately or stored at

–80°C. Chromatography was performed with ÄKTA FPLC system (GE Healthcare; Piscataway, NJ).

Pulse proteolysis

The pulse proteolysis reaction was performed in 20 mM Tris-HCl buffer (pH 8.0) containing 5.0 mM MgCl₂, 1.0 mM DTT, and varying concentrations of urea (0–4.0 M) with or without 1.0 mM ATP γ S. To make the reactions, the fractions were diluted three-fold with an appropriate buffer containing urea. The fractions were not dialyzed individually before pulse proteolysis and contained varying concentrations of NaCl from the NaCl gradient elution. The reactions were incubated overnight at 25°C. Pulse proteolysis was initiated by adding thermolysin to 20 μ g/mL. After 10 min, 50 μ L of the reactions were quenched by adding 7.5 μ L of 600 μ M phosphoamidon, incubating for 1 min, and then adding 7.5 μ L of 500 mM EDTA. The quenched samples were mixed with 15 μ L of 5 $\times$ SDS sample buffer and heated at 95°C for 3 min; 5 $\times$ SDS sample buffer was used to minimize the dilution of the samples. Ten microliters of each sample was analyzed with 13.5% SDS-PAGE gels. The gels were stained with SYPRO Red (Lonza Rockland; Rockland, ME), and scanned with a Typhoon imager (GE Healthcare; Piscataway, NJ). The scanned gels were stained again with colloidal Coomassie Blue G-250³³ for in-gel digestion. The sensitivity of the detection we employed in this study is $\sim$ 10 ng of proteins per band in SDS PAGE. Due to the dilutions during the reaction preparation, quenching, and SDS PAGE sample preparation, the target protein concentration in the fractions should be at least $\sim$ 5 μ g/mL to be detected on the gel.

Electropherogram analysis

Scanned images were analyzed with ImageJ software (available at: <http://rsbweb.nih.gov/ij/>) to produce electropherograms. A small constant value was added to the intensity in the electropherograms to avoid divide-by-zero errors and to suppress noise in the region with low intensities. The data position in one electropherogram of a pair was offset manually to maximize the overlap in band patterns between the pair. The binary logarithm of the ratio of the intensities, $|\log_2(\pm)|$, is defined as the absolute value of $\log_2\left(\frac{\text{intensity of the lane with ATP}\gamma\text{S}}{\text{intensity of the lane without ATP}\gamma\text{S}}\right)$. The 50% change in the intensity was used as the cutoff value to identify positive signals in $|\log_2(\pm)|$ plots. All the positive signals were visually inspected on SDS PAGE gel image to eliminate any false-positive signals from background or unsynchronized peaks.

In-gel digestion and protein identification

In-gel digestion was performed with identified bands as described previously.³⁴ After in-gel digestion, the peptides were desalted with C18 Ultra-Micro Pre-

pTip (Harvard Apparatus; Holliston, MA). Desalted peptide solution (0.5 μ L) was mixed with 0.5 μ L of 5 mg/mL α -cyano-4-hydroxycinnamic acid in 60% acetonitrile and 0.1% trifluoroacetic acid solution and allowed to dry on a MALDI target plate. The peptide sequences were determined with a 4800 Plus MALDI TOF/TOFTM Analyzer (Applied Biosystems; Foster City, CA). Proteins were identified by searching the K12 MG1655 genomic library by using the MASCOT database search engine³⁵ in GPS Explorer (Applied Biosystems; Foster City, CA). All identified proteins had significant MASCOT scores (>100) and GPS Explorer protein confidence indexes (>95%).

Pulse proteolysis of phosphoglyceromutase

The gene coding phosphoglyceromutase was amplified by PCR using K12 MG1655 chromosomal DNA as template. The amplified PCR products were cloned into pAED4 expression vectors. The proteins were overexpressed under the control of the T7 promoter in BL21(DE3)pLysS and purified using two consecutive anion exchange chromatography steps, DEAE Sepharose FF followed by Source 15Q, with an ÄKTA FPLC system (GE Healthcare; Piscataway, NJ).

To confirm the stabilization of phosphoglyceromutase in the presence of ATP, the C_m values of the protein were determined by pulse proteolysis as described previously.¹⁸ Briefly, 0.50 mg/mL phosphoglyceromutase was equilibrated overnight at 25°C in 20 mM Tris-HCl buffer (pH 8.0) containing 10 mM MgCl₂ and varying concentrations of urea with and without 1.0 mM ATP. The equilibrated protein solutions were treated with 0.20 mg/mL thermolysin for 1 min and then quenched by adding EDTA to a final concentration of 13 mM. The remaining proteins in the cell lysates after proteolysis were quantified from the band intensities on SDS PAGE gels stained with SYPRO Red (Lonza Rockland; Rockland, ME) using a Typhoon imager (GE Healthcare; Piscataway, NJ) and ImageJ (available at: <http://rsbweb.nih.gov/ij/>). The C_m values were determined by fitting the band intensities to the following equation:

$$I = I_0 \left\{ \frac{1}{1 + \exp(m(C_m - [\text{urea}]) / RT)} \right\},$$

where I is the observed band intensity, I_0 is the band intensity of the protein digested by pulse proteolysis under native conditions, and m is the dependence of global stability on urea.

Acknowledgment

The authors thank Joseph R. Kasper for helpful comments on this article.

References

1. Ong SE, Schenone M, Margolin AA, Li X, Do K, Doud MK, Mani DR, Kuai L, Wang X, Wood JL, Tolliday NJ,

- Koehler AN, Marcaurelle LA, Golub TR, Gould RJ, Schreiber SL, Carr SA (2009) Identifying the proteins to which small-molecule probes and drugs bind in cells. *Proc Natl Acad Sci USA* 106:4617–4622.
2. Burdine L, Kodadek T (2004) Target identification in chemical genetics: the (often) missing link. *Chem Biol* 11:593–597.
3. Rix U, Superti-Furga G (2009) Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* 5: 616–624.
4. Bantscheff M, Eberhard D, Abraham Y, Bastuck S, Boesche M, Hobson S, Mathieson T, Perrin J, Raida M, Rau C, Reader V, Sweetman G, Bauer A, Bouwmeester T, Hopf C, Kruse U, Neubauer G, Ramsden N, Rick J, Kuster B, Drewes G (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* 25:1035–1044.
5. Godl K, Wissing J, Kurtenbach A, Habenberger P, Blencke S, Gutbrod H, Salassidis K, Stein-Gerlach M, Missio A, Cotten M, Daub H (2003) An efficient proteomics method to identify the cellular targets of protein kinase inhibitors. *Proc Natl Acad Sci USA* 100:15434–15439.
6. Terstappen GC, Schlupen C, Raggiaschi R, Gaviraghi G (2007) Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov* 6:891–903.
7. Mayer TU, Kapoor TM, Haggarty SJ, King RW, Schreiber SL, Mitchison TJ (1999) Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286:971–974.
8. Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, Auld DS (2007) High-throughput screening assays for the identification of chemical probes. *Nat Chem Biol* 3:466–479.
9. Roth BL, Sheffler DJ, Kroeze WK (2004) Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat Rev Drug Discov* 3:353–359.
10. Imming P, Sinning C, Meyer A (2006) Drugs, their targets and the nature and number of drug targets. *Nat Rev Drug Discov* 5:821–834.
11. Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5: 993–996.
12. Heinrich MC, Griffith DJ, Druker BJ, Wait CL, Ott KA, Zigler AJ (2000) Inhibition of c-kit receptor tyrosine kinase activity by STI 571, a selective tyrosine kinase inhibitor. *Blood* 96:925–932.
13. MacDonald ML, Lamerdin J, Owens S, Keon BH, Bilter GK, Shang Z, Huang Z, Yu H, Dias J, Minami T, Michnick SW, Westwick JK (2006) Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nat Chem Biol* 2:329–337.
14. Missner E, Bahr I, Badock V, Lucking U, Siemeister G, Donner P (2009) Off-target decoding of a multitarget kinase inhibitor by chemical proteomics. *ChemBioChem* 10:1163–1174.
15. Liu P-F, Kihara D, Park C (2011) Energetics-based discovery of protein-ligand interactions on a proteomic scale. *J Mol Biol* 408:147–162.
16. West GM, Tucker CL, Xu T, Park SK, Han X, Yates JR, III, Fitzgerald MC (2010) Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc Natl Acad Sci USA* 107:9078–9082.
17. Lomenick B, Hao R, Jonai N, Chin RM, Aghajan M, Warburton S, Wang J, Wu RP, Gomez F, Loo JA, Wohlschlegel JA, Vondriska TM, Pelletier J, Herschman HR, Clardy J, Clarke CF, Huang J (2009) Target identification using drug affinity responsive target stability (DARTS). *Proc Natl Acad Sci USA* 106:21984–21989.
18. Park C, Marqusee S (2005) Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nat Methods* 2:207–212.
19. Na YR, Park C (2009) Investigating protein unfolding kinetics by pulse proteolysis. *Protein Sci* 18:268–276.
20. Kim MS, Song J, Park C (2009) Determining protein stability in cell lysates by pulse proteolysis and Western blotting. *Protein Sci* 18:1051–1059.
21. Schleich JP, Kim MS, Joh NH, Bowie JU, Park C (2011) Probing membrane protein unfolding with pulse proteolysis. *J Mol Biol* 406:545–551.
22. Schneider DA, Gourse RL (2004) Relationship between growth rate and ATP concentration in *Escherichia coli*: a bioassay for available cellular ATP. *J Biol Chem* 279: 8262–8268.
23. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* 5:593–599.
24. de Silva AO, Fraenkel DG (1979) The 6-phosphoglucuronate dehydrogenase reaction in *Escherichia coli*. *J Biol Chem* 254:10237–10242.
25. Danson MJ, Harford S, Weitzman PD (1979) Studies on a mutant form of *Escherichia coli* citrate synthase desensitized to allosteric effectors. *Eur J Biochem* 101: 515–521.
26. Powell G, Rajagopalan KV, Handler P (1969) Purification and properties of inosinic acid dehydrogenase from *Escherichia coli*. *J Biol Chem* 244:4793–4797.
27. Ding L, Seto BL, Ahmed SA, Coleman WG, Jr (1994) Purification and properties of the *Escherichia coli* K-12 NAD-dependent nucleotide diphosphosugar epimerase, ADP-L-glycero-D-mannoheptose 6-epimerase. *J Biol Chem* 269:24384–24390.
28. Liu P-F, Park C (in press) Selective stabilization of a partially unfolded protein by a metabolite. *J Mol Biol*. Available at: <http://dx.doi.org/10.1016/j.jmb.2012.05.044>.
29. Gorovits BM, Horowitz PM (1995) The chaperonin GroEL is destabilized by binding of ADP. *J Biol Chem* 270:28551–28556.
30. Park C, Marqusee S (2004) Analysis of the stability of multimeric proteins by effective ΔG and effective m -values. *Protein Sci* 13:2553–2558.
31. Ünlü M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18: 2071–2077.
32. Jeffery CJ (1999) Moonlighting proteins. *Trends Biochem Sci* 24:8–11.
33. Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9: 255–262.
34. Jimenez CR, Huang L, Qiu Y, Burlingame AL (1998) In-gel digestion of proteins for MALDI-MS fingerprint mapping. *Curr Protoc Protein Sci* 16:14.11–14.15.
35. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567.