DOI 10.1002/pmic.200800021

Proteomics 2008, *8*, 3094–3104

RESEARCH ARTICLE

Chemical proteomic study of isoprenoid chain interactome with a synthetic photoaffinity probe

Ruijun Tian[1, 2, 3]*, *Lingdong Li*[2, 3]*, *Wei Tang*[2, 3], *Hongwei Liu*[1, 2, 3], *Mingliang Ye*[1, 2], *Zongbao K. Zhao*[2]** and *Hanfa Zou*[1, 2]

[1] National Chromatography R&A Center, Dalian Institute of Chemical Physics, CAS, Dalian, China
[2] Division of Biotechnology, Dalian Institute of Chemical Physics, CAS, Dalian, China
[3] Graduate School of the Chinese Academy of Sciences, Beijing, China

A chemical proteomic approach was developed for profiling the noncovalent interactome of isoprenoid chain in the yeast proteome. A chemical probe that harbors a biotin moiety and a photoreactive benzophenone group linked to the terminal of geranyl group was synthesized. Photoaffinity labeling was performed by incubating the *Saccharomyces cerevisiae* proteome and the probe under 365 nm UV light. Thirty proteins were identified by immobilized NeutraAvidin enrichment, on-bead digestion, online 2-D nano-LC/MS/MS identification and semi-quantitative proteomic analysis. As noted by Gene Ontology annotation, the identified proteins demonstrate a wide range of catalytic activity in several biological processes, especially in metabolism and biosynthesis. Further data analysis shows that hydrophobic binding of the synthetic probe is potentially the major interaction force leading to covalent labeling. These results argue that intracellular allosteric interactions conferred by the isoprenoid chain of the corresponding chemical structures may be widespread at an interactomic level.

Received: January 9, 2008
Revised: April 11, 2008
Accepted: April 24, 2008

Keywords:
2-D nano LC / Chemical probe / Interactome / Isoprenoid chain / MS/MS / Yeast proteome

1 Introduction

Isoprenoid chain including dimethylallyl or isopentenyl (5-carbon), geranyl (10-carbon), farnesyl (15-carbon), geranylgeranyl (20-carbon), *etc*., are building blocks for terpenoids, one of the largest family of natural products with enormous structural and stereochemical diversities [1]. Attachment of such a hydrophobic carbon skeleton is a common step for biosynthetic elaboration of natural products such as sterols, carotenoids, and dolichols [2, 3]. Although it can be modified enzymatically to further diversify the structural space, the isoprenoid chains are routinely found intact in a large number of biologically relevant molecules, such as vitamin K2, coenzyme Q, and tocotrienols [4, 5] (Fig. 1). It has been well recognized that these compounds have a wide spectrum of bioactivities with various cellular functions [6]. Large biomolecules are also liable to modification with isoprenoid chain. The attachment of farnesyl and geranylgeranyl moiety to a protein molecule by the corresponding prenyltransferase is a common PTM that occurs on the CXXA box at the C-terminus [7]. Prenylated proteins comprise up to 0.5% of all proteins in mammalian tissues, and have been proved to play a vital role in diverse functions, such as cell growth, differentiations, cytoskeleton structure,

Correspondence: Professor Hanfa Zou, National Chromatography R&A Center, Dalian Institute of Chemical Physics, CAS, Dalian 116023, China
E-mail: hanfazou@dicp.ac.cn
Fax: +86-411-84379620

Abbreviations: CDI, *N,N*-carbonyldiimidazole; **CON,** control sample; **EXP,** experiment sample; **HSP,** heat shock protein; **SCX,** strong cation exchange; **Tween 20,** PEG sorbitan monolaurate

* These authors contributed equally to this work.
** Additional corresponding author: Professor Zongbao K. Zhao, E-mail: zhaozb@dicp.ac.cn

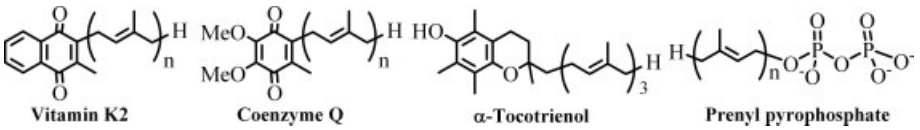

Figure 1. Representative metabolites decorated with isoprenoid chain.

and vesicle trafficking [7]. Consequently, inhibitors of protein prenyltransferase have attracted considerable attention as possible candidates for the design of chemotherapeutic agents.

Although interactions between the core structure of the isoprenoid chain decorated molecule and the corresponding cellular target are important, it is also apparent that the isoprenoid chain should have a contribution to the overall biological function, as it may have interactions with cellular targets *via* different interaction modes. However, it is intrinsically difficult to detect and quantify such interactions, because the all-carbon structure is generally considered as biochemically inert. In other words, the interactions usually result in neither covalent linkage nor ionic interactions, and thus are weaker than other intermolecular interactions. A similar situation is found for studying the interactions between long chain fatty acids and their target proteins. Yet specific hydrophobic interactions between free fatty acids and proteins have been demonstrated in recent years [8].

Recently, chemical proteomics have evolved as a promising approach for profiling small molecule interactome from a complex proteome [9]. Activity-based probes carrying active site-directed reactive groups have been successfully developed for characterizing the interaction of many enzymes with their substrates [10, 11]. A complementary approach to activity-based probes is chemical probes carrying photoaffinity labeling groups. Although the photoaffinity probes have been successfully applied for simple drug–protein interaction studies [12, 13], their potential for profiling the interactome of bioactive small molecules on a global scale has rarely been explored [14, 15].

In this paper, we synthesized a chemical probe that had a benzophenone group attached at the ω-position of a geranyl moiety, to which a biotin moiety was assembled through a linear chain. Our rationale was that those proteins interacting with isoprenoid chain would be covalently trapped when the mixture of the cell lysate and the probe was UV-irradiated at 365 nm. Upon further enrichment *via* avidin–biotin technology, on bead digestion and 2-D nano-LC/MS/MS analysis, the elusive isoprenoid chain interactome would be discovered. Our data indicated that isoprenoid chain attached to the metabolites might interact with a wide range of proteins at an interactomic level.

2 Materials and methods

2.1 Materials

DTT, iodoacetamide (IAA), PEG sorbitan monolaurate (Tween 20), PMSF, and Tris were purchased from Sino-American Biotechnology Corporation (Beijing, China). Urea, ammonium acetate, ammonium bicarbonate, and trypsin were obtained from Sigma (St. Louis, MO, USA). Immobilized NeutrAvidin was obtained from Pierce (Rockford, IL, USA). Protease inhibitor cocktail (EDTA free) was obtained from Roche (Penzberg, Germany). NC membrane was obtained from Invitrogen (Carlsbad, CA, USA). Horseradish peroxidase streptavidin was obtained from Vector (Burlingame, CA, USA). Commercial *S. cerevisiae* proteins, triose phosphate isomerase (T2507), 3-phosphoglycerate kinase (P7643), phosphoglucose isomerase (P5381), and glutathione reductase (G3664) were purchased from Sigma. 3,3′-Diaminobenzidine (DAB) substrate kit was obtained from Zymed (San Francisco, CA, USA). All the other reagents used were of HPLC or analytical reagent grade. Ultrapure water used for preparing solutions was produced by a Milli-Q water system (Millipore, Bedford, MA, USA).

2.2 Synthesis of the photoaffinity probe

As shown in Fig. 2, the synthesis of probe **3** was achieved *via* a coupling step between the precursor **1** [16] and precursor **2** [17] in the presence of *N,N*-carbonyldiimidazole (CDI). Briefly, to the suspension of compound **1** (60 mg, 128 μmol) in 3 mL dry DMF was added CDI (62 mg, 385 μmol) and the mixture was stirred at room temperature for 5 h before 12 μL methanol was added to decompose excess CDI. To this solution the precursor **2** was added directly (68 mg, 160 μmol) and the mixture was stirred for 30 h. Removal the solvent by an oil pump gave a residue which was purified by chromatography on silica gel eluted with iso-prophl alcohol (*i*-PrOH)/$NH_3 \cdot H_2O$ (4:1). The corresponding fractions were collected, concentrated, and lyophilized to give the product **3** as a white powder (36 mg, 32%). ^{1}H NMR (D_2O, 400 MHz): δ = 7.88 (1H, m), 7.75–7.79 (2H, m), 7.52–7.53 (1H, m), 7.31–7.36 (2H, m), 7.13–7.21 (2H, m), 5.21 (1H, m), 5.09 (1H, m), 4.33 (1H, m), 4.30 (2H, s), 4.10 (1H, m), 3.68 (2H, m), 2.88 (3H, m), 2.61 (1H, m), 2.49 (1H, *J* = 12.8 Hz), 2.34 (2H, *J* = 7.8, 21.6 Hz), 1.92 (2H, t, *J* = 7.2 Hz), 1.86 (2H, s), 1.76 (2H, s), 1.00–1.49 (20H, m). ^{13}C NMR (D_2O, 101 MHz): δ = 197.0, 175.4,165.4, 164.9, 137.9, 137.0, 136.1, 134.1, 133.1, 129.7, 128.4, 115.5, 70.9, 66.2, 61.9, 60.2, 55.4, 39.8, 39.2, 38.8, 35.5, 29.8, 29.0, 28.6, 28.2, 27.8, 26.0, 25.4, 24.8, 15.8, 13.5. ^{31}P NMR (D_2O, 162 MHz): δ = 16.3 (d, *J* = 26.1 Hz), −10.4(d, *J* = 26.1 Hz). High resolution MS (HRMS) (ESI) Calcd for $C_{40}H_{53}O_{11}N_3P_2S$ $[M-2H]^{2-}$: 845.2887; Found: 845.2886.

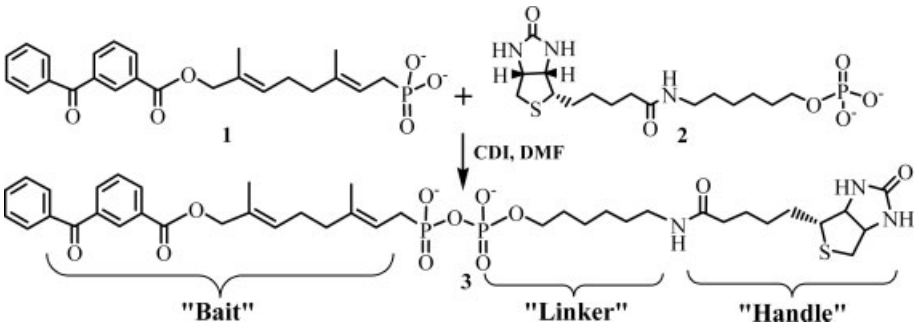

Figure 2. Synthesis of photoaffinity probe **3**.

2.3 Preparation of yeast cell extracts and protein labeling

Yeast *Saccharomyces cerevisiae* AS 2.399 was obtained from the China general microbiological culture collection center, and grown in YPD media (1.0% yeast extract, 1.0% peptone, 2.0% glucose, pH 5.8) at 30°C for 24 h. The cells were harvested by centrifugation, washed with PBS (NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 4.3 mM, KH_2PO_4 1.4 mM, pH 7.4), and stored at −80°C. For cell lysis, approximately 1.5 g of wet cells were ground in a mortar in the presence of liquid nitrogen. The homogenate was suspended in 5 mL of lysis buffer (20 mM phosphate buffer, pH 7.5, 0.2% v/v Tween 20, 1 mM PMSF, 1 mM DTT, 1 × protease inhibitor cocktail) at 4°C and then centrifuged at 20 000 × *g* for 30 min at 4°C. The supernatant was collected and the protein concentration was determined by the Bradford method. Three batches of wet cells were lysated in parallel and final protein concentrations were normalized.

For the photolabeling reaction, 1 mL of yeast proteome samples were incubated with 10 μM compound **3** in the dark at 4°C for 1 h. The samples were then irradiated at 4°C for 2 h using two ZF-I UV lamps (365 nm) at a distance of 5 cm, with rotation. Control samples were similarly handled but with no UV irradiation.

2.4 Streptavidin blot characterization

Total proteins were separated by 12% SDS-PAGE and then transferred to an NC membrane. After being blocked with 5% w/v skimmed milk for 1 h, the membrane was incubated with horseradish peroxidase streptavidin at 25°C for 90 min and washed three times for 15 min with PBS containing 0.1% Tween 20. The membrane was visualized using a DAB kit following manufacturer's instructions.

2.5 NeutrAvidin affinity enrichment and sample preparation for mass spectrometric analysis

After the photolabeling reaction, both the experimental samples (EXPs) and the control samples (CONs) were precipitated by the chloroform–methanol method twice to remove unreacted photoaffinity probe [18]. Precipitated proteins were solubilized in 0.2% w/v SDS, incubated with 100 μL of immobilized NeutrAvidin beads (Pierce) at 4°C for 1.5 h with gentle shaking. Unbound proteins were removed by washing the gel with binding buffer (0.1 M sodium phosphate buffer, pH 7.2, 0.15 M NaCl, 0.1 % w/v SDS) for four times and with pure water twice. Finally, the gels were lyophilized to dryness.

For on-bead digestion, the gels were resuspended in 100 μL of denaturing buffer (8 M urea, 100 mM NH_4HCO_3, pH 8.3). The samples were then mixed with 2 μL of 1 M DTT and incubated at 37°C for 2 h. After that, 10 μL of 1 M IAA was added and incubated at 25°C for an additional 30 min in the dark. The samples were diluted with 400 μL of 100 mM NH_4HCO_3, pH 8.3 and incubated with 20 μg of trypsin at 37°C for 16 h. The digested mixture was desalted with a C18 solid-phase cartridge and lyophilized.

2.6 2-D nano-LC/MS/MS analysis and database searching

The 2-D nano-LC/MS/MS system consisted of a quaternary Surveyor pump and an LTQ linear IT mass spectrometer equipped with a nanospray source (Thermo, San Jose, CA, USA). The temperature of the ion transfer capillary was set at 200°C, and the spray voltage was set at 1.8 kV. All MS and MS/MS spectra were acquired in the data-dependent mode. The mass spectrometer was set that one full MS scan was followed by ten MS/MS scans. The 2-D nano-LC/MS/MS analysis was carried out according to a reported method with minor modifications [19]. Briefly, the tryptic samples were dissolved in 0.1% v/v acetic acid, loaded onto a monolith strong cation exchange (SCX) column (150 μm id × 5 cm) by air pressure within 5 min, and then equilibrated with 0.1% acetic acid/5% ACN. After that, the SCX column with loaded peptide sample was directly connected to an RP column in tandem by a union. The peptides were eluted onto the RP column using a step gradient generated with 1000 mM ammonium acetate in 0.1% acetic acid/5% ACN. The salt concentrations of the buffer for the ten salt steps were: 50, 100, 150, 200, 250, 300, 350, 400, 500, and 1000 mM. After each elution step, a subsequent RPLC/MS/MS was performed.

The acquired MS/MS spectra were searched against the Yeast database downloaded from website [ftp://genome-ftp.stanford.edu/yeast/data_download/sequence/genomic_sequence/orf_protein/orf_trans.fasta.gz] using the Turbo SEQUEST in the BioWorks 3.2 software suite (Thermo). Peptides were searched using fully tryptic cleavage constraints and up to two missed cleavages sites were allowed for tryptic digestion. Cysteine residues were searched as fixed modification of 57.0215 Da, and methionine residues were searched as a variable modification of +15.9949 Da. The mass tolerances are 2 Da for parent masses and 1 Da for fragment masses. Initial searching results were filtered with the following parameters as reported previously by Roth *et al.* [20]: the minimum X_{corr} of 1.8 for a singly charged peptide; 2.5 for a doubly charged peptide; and 3.5 for a triply charged peptides; the minimum ΔC_n cutoff value of 0.08; and more than two peptides measured *per* protein. Furthermore, only the protein identified in at least two out of three samples was saved. For semiquantitative comparison of the proteins identified in EXPs and the CONs, spectral counts for each identified protein from each experiment were extracted, averaged, normalized, and compared as described previously and as noted in the footnotes for Table 1 [20, 21]. The biological processes and molecular functions involved for these identified proteins were assigned using GoMiner, a Web-based application based on Gene Ontology annotation (discovery.nci.nih.gov/gominer) [22, 23].

Table 1. Total spectral counts and proteins identified from the EXPs and the CONs[a)]

Sample name	Total spectral counts	Total identified proteins
EXP 1	1108	134
CON 1	330	40
EXP 2	1008	82
CON 2	1186	124
EXP 3	751	73
CON 3	993	76

a) All data met the following criteria: (i) The minimum Xcorr for singly, doubly and triply charged peptides were 1.8, 2.5, and 3.5, respectively; (ii) the minimum ΔC_n cutoff value was 0.08; (iii) two or more peptides were identified *per* protein.

2.7 *In vitro* labeling experiments

For *in vitro* labeling of *S. cerevisiae* proteins, about 0.2 mg of triose phosphate isomerase (Sigma) was dissolved in the 200 μL of buffer containing 20 mM Tris-HCl and 0.02% w/v BSA, pH 7.5. Protein solution (50 μL) was added to 100 μL of photolabeling reaction solution (20 mM Tris-HCl, 1 mM $MgSO_4$, 10 μM compound **3**, pH 7.5). The mixtures were irradiated at 4°C for 80 and 120 min, respectively, under a UV light at 365 nm. In another set of experiment, commercial *S. cerevisiae* proteins, 3-phosphoglycerate kinase, phosphoglucose isomerase, and glutathione reductase were treated under identical conditions for 2 h. The labeling mixtures were subjected to SDS-PAGE and streptavidin blot treatment as described earlier.

3 Results

3.1 Development and validation of the photoaffinity probe

As shown in Fig. 2, the chemical probe **3** consisted of three parts, the bait, the linker and the handle. The bait was a geranyl moiety extended with a photoactive benzophenone group that also structurally mimic the isoprenoid chain. The biotin moiety was the handle, which facilitates protein enrichment *via* biotin/avidin technology [24, 25]. The linear chain together with the pyrophosphate isostere made the flexible linker. It was important to have a linker so that spatial constraint could not prevent the bait from reaching their respective binding pockets. Synthesis of compound **3** was realized by coupling the fragments **1** [16] and **2** [17] in DMF promoted by CDI. Purification of compound **3** was achieved by silica gel chromatography using *i*-PrOH/$NH_3 \cdot H_2O$ (4:1) as the eluent. The probe **3** in its NH_4^+ salt form was a white powder soluble in water. Good solubility made probe **3** an excellent agent for further exploring hydrophobic interactions, since a hydrophobic probe with poor solubility would require excess detergent and make the explanation of the data more complicated. ^{31}P NMR spectra of the purified probe showed two sets of doublet signals at around 16.3 and −10.4 ppm, respectively. HRMS analysis confirmed the elemental composition of the synthetic structure, and revealed that the pyrophosphate isostere moiety was dissociated at neutral pH.

We next tested the capability of the probe **3** for photoaffinity labeling. Yeast *S. cerevisiae* AS 2.399 cell extracts and the probe were incubated under UV irradiation at 365 nm, the protein mixtures were analyzed by SDS-PAGE and streptavidin blot. As shown in Fig. 3A, soluble proteins were distributed over a wide molecular weight range and showed no discernible differences before and after UV irradiation based on CBB staining. However, the streptavidin blot results indicated that multiple protein bands were labeled, and that the labeling was apparently a time-dependent event (Fig. 3B). As the benzophenone group modifies proteins preferentially on unreactive C–H bonds nearby the excited intermediate with little site specificity [26], irradiation for a longer time would result in more extensive trapping of those targets, as well as extra nonspecific labeling. A recent work also demonstrated that over exposure of the proteome to a benzophenone-based probe slightly increased the labeling of bystander proteins [27]. Therefore, we chose a UV irradiation for 2 h with 10 μM of probe **3** for further investigation.

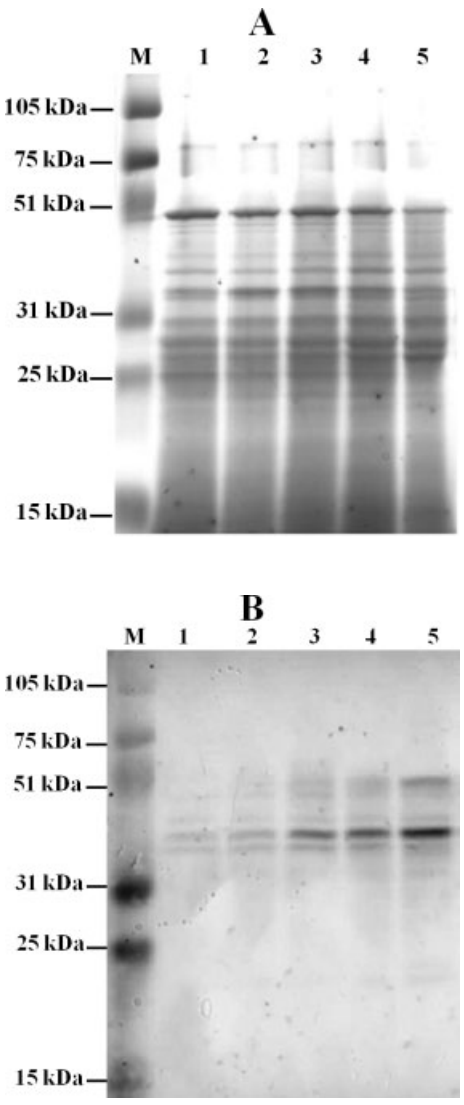

Figure 3. Photoaffinity labeling of yeast proteomes by probe **3**. Soluble yeast proteome (400 μg) was labeled with 10 μM of probe **3** at different time intervals, separated by SDS-PAGE and visualized with CBB staining (A) and streptavidin blot (B). M: protein markers; lane 1: the CON without UV irradiation; lane 2: sample with probe **3** for 1 h; lane 3: sample with probe **3** for 2 h; lane 4: sample with probe **3** for 3 h; lane 5: sample with probe **3** for 4 h.

3.2 Enrichment and identification of photolabeled proteome

To identify proteins interacting with isoprenoid chain, yeast proteomes labeled with probe **3** were subjected to affinity enrichment, on bead digestion and online 2-D nano-LC/MS/MS analysis. *S. cerevisiae* AS 2.399 was cultured in YPD media for 24 h, and cells were divided into three portions for parallel protein extraction. The total proteins were extracted and divided into two portions for the labeling and the control experiment. After photoaffinity labeling reaction, excess probes were removed by chloroform–methanol method. The affinity enrichment of the labeled sample was done using immobilized NeutrAvidin beads due to their minimized nonspecific binding compared with immobilized avidin and streptavidin. A stringent elution in the presence of 0.1–0.2% SDS was used to eliminate nonspecific adsorption of proteins on the beads. Conventionally, the enriched proteins are eluted under harsh denaturing condition, separated by SDS-PAGE, and followed by in gel digestion and MS analysis [28]. However, all of these procedures are time consuming and result in sample loss. In this study, because the powerful 2-D separation coupled with online MS analysis was used, direct on bead digestion was performed to release peptides from these enriched proteins. Finally, 2-D nano-LC/MS/MS analysis was carried out based on the novel monolithic SCX column fractionation, online RP separation, and MS/MS analysis [19]. The results for each sample are summarized in Table 1, suggesting that a large number of peptides and proteins can be identified from each sample.

To improve the reliability of the identification of the proteins, stringent filter criteria were then set: (i) at least two peptides identified *per* protein and (ii) identified in at least two out of three independent experiments. Finally, 71 and 59 proteins were identified from the EXP and the CON, respectively, and 42 proteins were presented in both groups. These results demonstrate the existence of a large number of nonspecific enriched proteins. To this end, a semiquantitative method based on spectral counts as described by Roth *et al.* [20, 21] was applied to compare the relative abundance of each protein in the two groups. We therefore believe that specific labeled proteins could be distinguished from the nonspecific ones.

After meeting the criteria described above, the spectral counts of both EXP and CON were averaged, normalized, and compared. Thirty proteins with EXP/CON spectral count ratio more than 5 were obtained as candidate proteins labeled specifically with the photoaffinity probe (Table 2). Nearly all of the proteins listed in Table 2 were identified uniquely in EXP, except for cytoplasmic chaperone (Hsp90 family) with high EXP/CON ratio of 9. These results confirm the high reliability of the obtained proteomic data.

These proteins were further classified using Gene Ontology annotation. As shown in Fig. 4A, the identified proteins engage in a wide spectrum of biological processes, and many of them are related to cellular metabolism. Further exploration of this category shows that most of these proteins take part in the metabolism or biosynthesis of biomolecules, such as nucleotides, lipids, aromatic compounds, protein, *etc.* Figure 4B shows that catalytic activity is the major molecular function for the identified proteins; 20 of the identified proteins fall into this category. Seven types of enzyme activities were found, especially for the hydrolytic, oxidoreductive, and group transferring activity with more than four candidates for each type. Furthermore, 11 proteins were classified with a wide range of binding activities. The general biological processes and molecular function classification

Table 2. A list of identified proteins meeting the criteria[a)]

Protein name	ORF no.	EXP spectral counts[b)]	CON spectral counts[b)]	EXP/CON ratio[c)]
Oxidoreductase	YGL157W	403.7	0.2	2018.5
Cytoplasmic chaperone of the Hsp90 family	YMR186W	206	0.2	1030
Bifunctional carbamoylphosphate synthetase	YJL130C	201	0.2	1005
Glyceraldehyde-3-phosphate dehydrogenase	YJR009C	158	0.2	790
Elongation factor 2 (EF-2)	YDR385W	138.3	0.2	691.5
3-phosphoglycerate kinase	YCR012W	79.3	0.2	396.5
Alcohol dehydrogenase	YOL086C	63	0.2	315
Basic protein required for normal cell-cycle regulation of histone gene transcription	YBR215W	53.7	0.2	268.5
ATP-dependent DEAD (Asp-Glu-Ala-Asp)-box RNA helicase	YOR204W	33.7	0.2	168.5
Protein involved in structural maintenance of chromosomes	YLR383W	32	0.2	160
Hypothetical protein	YPR011C	29.7	0.2	148.5
Triose phosphate isomerase	YDR050C	27.7	0.2	138.5
Tetrameric phosphoglycerate mutase	YKL152C	26.7	0.2	133.5
Tetradecameric mitochondrial chaperonin	YLR259C	26	0.2	130
Mismatch repair protein	YCR092C	25.7	0.2	128.5
F-box protein	YLR368W	25	0.2	125
Nucleoside diphosphate kinase	YKL067W	25	0.2	125
Minor isoform of pyruvate decarboxylase	YLR134W	24	0.2	120
Subunit of the Anaphase-Promoting Complex/Cyclosome	YOR249C	24	0.2	120
Mitochondrial cruciform cutting endonuclease	YKL011C	22.3	0.2	111.5
Translation elongation factor EF-1 gamma	YKL081W	21	0.2	105
Aspartic protease	YLR121C	19.3	0.2	96.5
Transcription factor	YMR172W	19	0.2	95
Subunit of the condensin complex	YLR086W	18.7	0.2	93.5
Putative protein	YGR117C	17.7	0.2	88.5
Mitochondrial cell death effector	YNR074C	16	0.2	80
Putative protein of unknown function	YHR048W	15.7	0.2	78.5
Protein kinase primarily involved in telomere length regulation	YBL088C	15	0.2	75
Putative RNA helicase related to Ski2p	YGR271W	15	0.2	75
Cytoplasmic chaperone (Hsp90 family)	YPL240C	194.3	21.7	9

a) The identified proteins with EXP/CON ratio of greater than 5 are listed here, and composed by the order of EXP/CON ratio. The complete dataset is available in Table S1 of Supporting Information.

b) The raw spectral counts from three independent experiments were firstly normalized (raw spectral counts for each identified protein were divided by the total spectral count number after meeting the criteria shown in the experimental source, then multiplied by 10 000). Then the normalized values from the three EXPs and CONs were averaged.

c) Ratio of normalized, averaged EXP to CON spectral counts. Normalized, averaged CON spectral counts of zero were changed to 0.2 to avoid division by zero.

show that the affinity probe based on isoprenoid chain has the potential to enrich proteins with diverse transformation capabilities.

3.3 *In vitro* labeling on the identified proteins

Among those identified proteins, yeast triose phosphate isomerase (YDR050C) and 3-phosphoglycerate kinase (YCR012W) were commercially available. We first performed the labeling experiment with triose phosphate isomerase using an identical photolabeling procedure to that for the proteome, and subjected it to an avidin blot. As shown in Fig. 5A, a clear band was observed, indicating that attachment of the biotinylated probe to triose phosphate isomerase was realized upon UV irradiation. More importantly, the modification by compound **3** was also time dependent. Figures 5B and C show additional labeling results with yeast proteins. It was clear that 3-phosphoglycerate kinase gave a positive signal, whilst both phosphoglucose isomerase and glutathione reductase, that are not present in Table 2, indicated no discernible labeling products.

4 Discussion

Metabolites decorated with isoprenoid chain include a large number of intracellular components; yet molecular interactions between those isoprenoid chain and proteins have

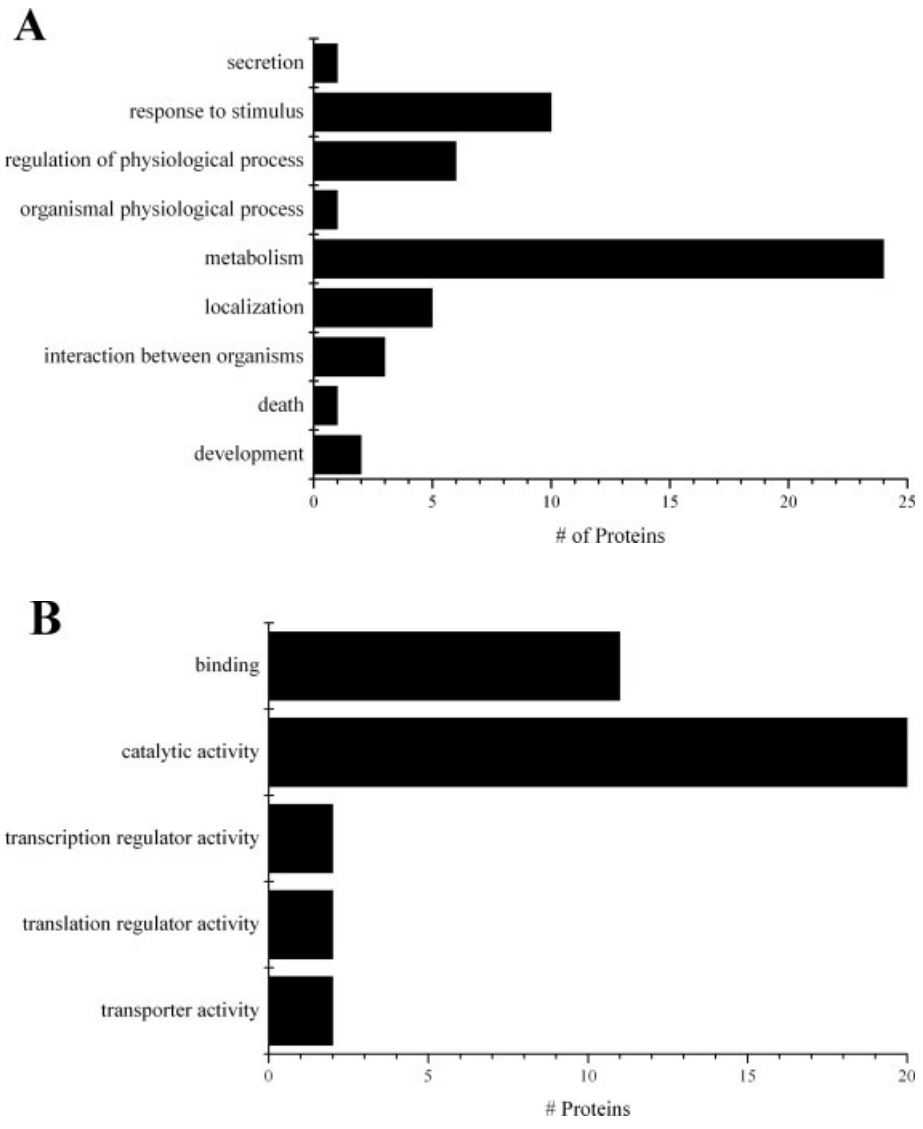

Figure 4. Gene Ontology biological process (A) and molecular function (B) of the identified proteins.

rarely been studied. This is likely because those interactions are hydrophobic interactions, which usually are of a much weaker chemical nature. The determination of such interactions is also technically challenging because limited tools are available to prove hydrophobic interactions between small molecules and proteins.

In this study, photoaffinity probe **3** carrying an isoprenoid chain was synthesized and used for profiling the noncovalent interactome of isoprenoid chains in a yeast proteome. The probe had a photoactivable benzophenone group attached to the terminal of a geranyl chain *via* an ester linkage. The geranyl chain was linked to a biotin-functionalized moiety *via* a pyrophosphate isostere, in which a C–P bond was introduced to block any prenyltransferase activity. Benzophenone was introduced for covalently trapping the proteins because of its chemical stability, hydrophobicity, and photoactivity in ambient UV light [26]. Furthermore, the benzophenone group has significant structure similar to isoprenoid chain. Indeed, a series of photoreactive farnesyl pyrophosphate analogs have been synthesized [16, 29–31], and have been demonstrated as competitive inhibitors of protein farnesyltransferases, suggesting that the benzophenone-modified prenyl chain can bind to the active site where farneyl or geranylgeranyl moiety resides [30, 31]. Structural data based on X-ray crystallographic analysis also showed that the benzophenone-functionalized geranyl moiety mimicked the farnesyl and geranylgeranyl moiety in binding to prenyltransferases [16]. Other photoactivable groups, such as azide and diazirine, might lead to more nonspecific labeling or synthetic difficulties. Also, irradiation at 254 nm for the activation of azide and diazirine group is quite harsh, often leading to photo-destruction target proteins.

Our initial labeling experiments with probe **3** revealed that a large amount of proteins could be identified from the

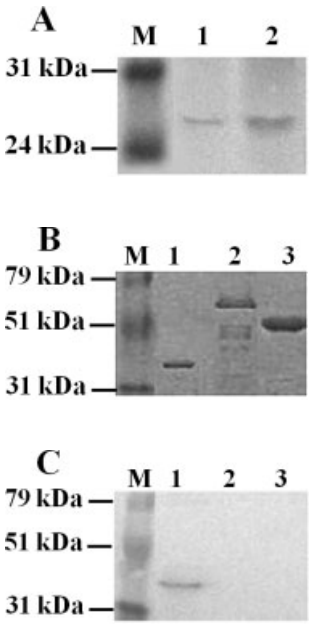

Figure 5. *In vitro* labeling of yeast proteins using 10 μM of compound **3**. (A) Streptavidin blot analysis of triose phosphate isomerase labeled at different incubation times, lane 1: 80 min; lane 2: 120 min. Coomassie staining (B) and streptavidin blot (C) analysis of the labeling results of 3-phosphoglycerate kinase (lane 1), phosphoglucose isomerase (lane 2) and glutathione reductase (lane 3).

control experiment as show in Table 2. Three potential sources may result in nonspecific enrichment. Firstly, the endogenous biotinylated proteins [32], such as acetyl-CoA carboxylase, pyruvate carboxylases and tRNA-binding proteins, are the most high-abundant proteins found in this study. Binding of these proteins on the NeutrAvidin beads will influence the specificity to some extent. Another potential source is nonspecific absorption onto the immobilized avidin beads. This might be largely due to the chemical nature of the graft materials of the beads. In a recent chemical proteomic study on liver and kidney proteomes by Bogyo and coworkers [33], over half of all identified proteins were identified as background proteins after immobilized streptavidin beads enrichment. In another work reported by Roth *et al.* [20], after chemical labeling and immobilized streptavidin enrichment, the vast majority of the identified proteins were identified in both of EXPs and the CONs. Finally, because the benzophenone group modifies proteins preferentially on unreactive C–H bonds with little site specificity [26], the photoaffinity probe may also result in some nonspecific labeling, especially those proteins with similar small molecule-interacting counterparts. A recent work demonstrated that overexposure of the proteome to the benzophenone-based probe slightly increased the labeling of bystander proteins [27]. Hence, as compared to the previous reports [16, 34], both the UV irradiation time and the power of the UV light were reduced in this work to diminish the nonspecific labeling events.

A systematic identification approach generated those candidate proteins with high confidence (Table 2). Five enzymes involved in the glycolysis pathway were identified, including alcohol dehydrogenase, tetrameric phosphoglycerate mutase, 3-phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase, and triose phosphate isomerase. The enrichment of these glycolytic enzymes is likely due to their stronger interactions with the probe. Unfortunately, little biochemical information can be obtained about molecular interactions between metabolites harbored isoprenoid chain and the glycolytic enzymes. Noting that the benzophenone-functionalized geranyl moiety is a hydrophobic and linear-shaped structure, it may be interesting to make the probe resembling the fragment of a long chain fatty acids. Indeed, hydrophobic interactions between protein and lipids or fatty acids have been well documented [7, 35]. It is known that fatty acids had various effects on enzymatic activities of hexokinase [36], glucose 6-phosphate phosphatase [37], fructose-1,6-phosphatase, and phosphofructokinase I [38]. Association of enolase and glyceraldehyde-3-phosphate dehydrogenase of *Lactobacillus crispatus* with the cell wall and lipoteichoic acids have also been reported recently [39]. Therefore, our data suggested that interactions between the glycolytic enzymes and metabolites with isoprenoid chain might have been overlooked.

As noted by Gene Ontology annotation, majority of the identified proteins are classified as having catalytic activity, especially kinase activity. Examples of protein kinases showing interactions with hydrophobic structures are known in the literature. Geranylgeranyl pyrophosphate and its homologs were demonstrated as strong inhibitors of human mevalonate kinase through competitive interactions at the ATP-binding site [40]. In addition, synthetic geranyl pyrophosphate analogs have been reported to inhibit both mevalonate kinase and mevalonate 5-phosphate decarboxylase [41]. Inhibition of mammalian cardiac hexokinase by long chain fatty acids, but activation of the same enzyme by medium chain fatty acids were also reported, although the detailed mechanism remained elusive [36]. Protein kinase from human platelets was found to be activated by free fatty acids [42]. Also, mid- and long-chain fatty acids have been found to activate p38 mitogen-activated protein kinase, which further increased levels of phosphoenolpyruvate carboxykinase and glucose-6-phosphate gene transcripts in mice hepatocytes [43].

Four proteins belonging to the heat shock protein (HSP) family were identified. Members of the HSP family are generally comprised of two functionally coupled domains, the N-terminal ATPase domain and the C-terminal protein binding domain, both of which are central to HSP chaperone activity. It was shown that 3′-sulfogalactolipid specifically bound within the N-terminal domain, leading to a noncompetitive inhibition of HSP70 ATPase activity [35]. In a similar report, Na^+, K^+-ATPase purified from *Pichia pastoris* membranes was shown to have specific interaction with lipids [44]. Therefore, our experimental data indicated that interactions

between HSP proteins and hydrophobilic compounds might be more common than previously thought.

Certain fatty acids have been shown to inhibit hepatitis C virus [45] and brome mosaic virus [46] RNA replication. The reason was tentatively attributed to the requirement for a fatty acid-constructed membrane architecture for RNA replication to occur, but unfortunately concrete biochemical data are not yet available. We believe that the RNA helicases and other related enzymes identified in this study might directly interact with fatty acids or an isoprenoid chain-related structure, leading to regulation of a viral RNA replication.

Given many proteins that were fished out using probe **3**, it was puzzling that some known proteins involved in prenyl transformation, such as protein prenyltransferases, prenylpyrophosphate synthases, *etc*. did not survive our stringent filter criteria. There are some possible explanations for this paradox. One reason is that the abundance of these proteins might be too low, so that they failed to be enriched in the presence of other high abundance counterparts. The other reason is that the probe is not perfect for trapping these proteins. Although similar compounds showed competitive inhibition to protein farnesyltransferase [16, 31], the IC_{50} values were at the lower μM range, suggesting that the binding affinity was rather low. It is also interesting to note that a similar experiment also failed to identify protein prenyltransferases in plant cells [16]. Nevertheless, we have observed some farnesylation related enzymes, such as *p*-hydroxybenzoate:polyprenyl transferase (YNR041C) and farnesyl cysteine-carboxyl methyltransferase (YDR410C), in some of the labeling experiments. The chemical probe can be further optimized to expose the isoprenoid chain so that it can better mimic *in vivo* isoprenoid structure. Successful *in vitro* labeling triose phosphate isomerase and 3-phosphoglycerate kinase have firmly demonstrated the reliability of our method. Further labeling experiment with phosphoglucose isomerase and glutathione reductase afforded no discernible biotin-blot signals, indicating that probe **3** is unlikely random trapping cellular proteins. Therefore, our data suggested that interactions between those identified proteins and probe **3** are likely specific.

It should be emphasized that isoprenoid chain are hydrophobic, linear-shaped chemical constituents. Interactions between isoprenoid chain and the majority of proteins identified in this work are unlikely at the sites where their natural ligands reside. In other words, the isoprenoid chain most likely interacts elsewhere within the protein structural space. Therefore, the biochemical consequences of such interactions are usually noncompetitive inhibitions. In addition, activity disturbance through interaction outside the active site of proteins is generally termed as allosterism [47, 48]. Allosteric regulation using small organic molecules has been recently explored for drug development. For example, the allosteric site of $GABA_A$ receptors has been suggested as a target for CNS disorders drug development [49]. Uncompetitive glutamate racemase inhibitors that bound to a cryptic allosteric site have been identified, providing a basis for design of narrow-spectrum antimicrobial agents [50]. Our data argues that allosteric effects of those metabolites containing isoprenoid chain at an interactomic level might have been underestimated.

In conclusion, we have demonstrated the application of a chemical proteomic approach for studying the isoprenoid chain interactome, and identified 30 proteins with diverse activities and functions from the *S. cerevisiae* proteome. Our data suggest that metabolites decorated with isoprenoid chain are likely to be involved in a wide range of biological processes and that the isoprenoid chain *per se* might impose additional interactions with the target proteins. Based on our observations, two potential directions are identified (i) to address the biological significance of the interactions between isoprenoid chain and the identified proteins and (ii) to explore human and plant samples with similar probes to identify targets with more biological significance. Further work is being undertaken in our laboratory and results will be delivered in due course.

The present work is supported by National Natural Sciences Foundation of China (no. 20735004, 20675081), the China State Key Basic Research Program Grant (2005 CB522701, 2007 CB914102), the China High Technology Research Program Grant (2006AA02A309), the Knowledge Innovation program of CAS (KJCX2.YW.H09), the Knowledge Innovation Program of DICP to H. Z. and CAS "100 Talents" program (Z. Z.).

The authors have declared no conflict of interest.

5 References

[1] Liang, P. H., Ko, T. P., Wang, A. H., Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* 2002, *269*, 3339–3354.

[2] Edwards, P. A., Ericsson, J., Sterols and isoprenoids: Signaling molecules derived from the cholesterol biosynthetic pathway. *Annu. Rev. Biochem.* 1999, *68*, 157–185.

[3] Sagami, H., Kurisaki, A., Ogura, K., Formation of dolichol from dehydrodolichol is catalyzed by NADPH-dependent reductase localized in microsomes of rat liver. *J. Biol. Chem.* 1993, *268*, 10109–10113.

[4] Clarke, C. F., Williams, W., Teruya, J. H., Ubiquinone Biosynthesis in *Saccharomyces cerevisiae*: Isolation and sequence of COQ3, the 3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase gene. *J. Biol. Chem.* 1991, *266*, 16636–16644.

[5] Hofius, D., Sonnewald, U., Vitamin E biosynthesis: Biochemistry meets cell biology. *Trends Plant Sci.* 2003, *8*, 6–8.

[6] Fraser, P. D., Bramley, P. M., The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* 2004, *43*, 228–265.

[7] Kobayashi, M., Mutharasan, R. K., Feng, J., Roberts, M. F., Lomasney, J. W., Identification of hydrophobic interactions between proteins and lipids: Free fatty acids activate phos-

pholipase C delta1 via allosterism. *Biochemistry* 2004, *43*, 7522–7533.

[8] Murthy, S. N., Chung, P. H., Lin, L., Lomasney, J. W., Activation of phospholipase Cepsilon by free fatty acids and cross talk with phospholipase D and phospholipase A2. *Biochemistry* 2006, *45*, 10987–10997.

[9] Han, S. Y., Hwan Kim, S., Introduction to chemical proteomics for drug discovery and development. *Arch. Pharm. (Weinheim)* 2007, *340*, 169–177.

[10] Barglow, K. T., Cravatt, B. F., Activity-based protein profiling for the functional annotation of enzymes. *Nat. Methods* 2007, *4*, 822–827.

[11] Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z. *et al.*, Chemical approaches for functionally probing the proteome. *Mol. Cell. Proteomics* 2002, *1*, 60–68.

[12] Davidson, W., McGibbon, G. A., White, P. W., Yoakim, C. *et al.*, Characterization of the binding site for inhibitors of the HPV11 E1-E2 protein interaction on the E2 transactivation domain by photoaffinity labeling and mass spectrometry. *Anal. Chem.* 2004, *76*, 2095–2102.

[13] Qiu, W. W., Xu, J., Liu, D. Z., Li, J. Y. *et al.*, Design and synthesis of a biotin-tagged photoaffinity probe of paeoniflorin. *Bioorg. Med. Chem. Lett.* 2006, *16*, 3306–3309.

[14] Salisbury, C. M., Cravatt, B. F., Activity-based probes for proteomic profiling of histone deacetylase complexes. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 1171–1176.

[15] Sieber, S. A., Niessen, S., Hoover, H. S., Cravatt, B. F., Proteomic profiling of metalloprotease activities with cocktails of active-site probes. *Nat. Chem. Biol.* 2006, *2*, 274–281.

[16] DeGraw, A. J., Zhao, Z., Strickland, C. L., Taban, A. H. *et al.*, A photoactive isoprenoid diphosphate analogue containing a stable phosphonate linkage: Synthesis and biochemical studies with prenyltransferases. *J. Org. Chem.* 2007, *72*, 4587–4595.

[17] Li, L. D., Tang, W., Liu, W. J., Zhao, Z. B., Synthesis and activity of a functional probe based on photoaffinity-labeled prenyl side-chain probe. *Chin. J. Org. Chem.* 2008, *28*, 489–493.

[18] Wessel, D., Flugge, U. I., A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* 1984, *138*, 141–143.

[19] Wang, F., Dong, J., Jiang, X., Ye, M., Zou, H., Capillary trap column with strong cation-exchange monolith for automated shotgun proteome analysis. *Anal. Chem.* 2007, *79*, 6599–6606.

[20] Roth, A. F., Wan, J., Bailey, A. O., Sun, B. *et al.*, Global analysis of protein palmitoylation in yeast. *Cell* 2006, *125*, 1003–1013.

[21] Wan, J., Roth, A. F., Bailey, A. O., Davis, N. G., Palmitoylated proteins: Purification and identification. *Nat. Protoc.* 2007, *2*, 1573–1584.

[22] Zeeberg, B. R., Feng, W., Wang, G., Wang, M. D. *et al.*, GoMiner: A resource for biological interpretation of genomic and proteomic data. *Genome Biol.* 2003, *4*, R28.

[23] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D. *et al.*, Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 2000, *25*, 25–29.

[24] Bayer, E. A., Wilchek, M., Application of avidin-biotin technology to affinity-based separations. *J. Chromatogr.* 1990, *510*, 3–11.

[25] Bayer, E. A., Wilchek, M., Avidin- and streptavidin-containing probes. *Methods Enzymol.* 1990, *184*, 174–187.

[26] Dorman, G., Prestwich, G. D., Benzophenone photophores in biochemistry. *Biochemistry* 1994, *33*, 5661–5673.

[27] Lamos, S. M., Krusemark, C. J., McGee, C. J., Scalf, M. *et al.*, Mixed isotope photoaffinity reagents for identification of small-molecule targets by mass spectrometry. *Angew. Chem. Int. Ed. Engl.* 2006, *45*, 4329–4333.

[28] Kidd, D., Liu, Y. S., Cravatt, B. F., Profiling serine hydrolase activities in complex proteomes. *Biochemistry* 2001, *40*, 4005–4015.

[29] Turek, T. C., Gaon, I., Distefano, M. D., Strickland, C. L., Synthesis of farnesyl diphosphate analogues containing ether-linked photoactive benzophenones and their application in studies of protein prenyltransferases. *J. Org. Chem.* 2001, *66*, 3253–3264.

[30] Gaon, I., Turek, T. C., Weller, V. A., Edelstein, R. L. *et al.*, Photoactive analogs of farnesyl pyrophosphate containing benzoylbenzoate esters: Synthesis and application to photoaffinity labeling of yeast protein farnesyltransferase. *J. Org. Chem.* 1996, *61*, 7738–7745.

[31] Turek-Etienne, T. C., Strickland, C. L., Distefano, M. D., Biochemical and structural studies with prenyl diphosphate analogues provide insights into isoprenoid recognition by protein farnesyl transferase. *Biochemistry* 2003, *42*, 3716–3724.

[32] Kim, H. S., Hoja, U., Stolz, J., Sauer, G., Schweizer, E., Identification of the tRNA-binding protein Arc1p as a novel target of in vivo biotinylation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 2004, *279*, 42445–42452.

[33] Fonovic, M., Verhelst, S. H. L., Sorum, M. T., Bogyo, M., Proteomics evaluation of chemically cleavable activity-based probes. *Mol. Cell. Proteomics* 2007, *6*, 1761–1770.

[34] Morohashi, Y., Kan, T., Tominari, Y., Fuwa, H. *et al.*, C-terminal fragment of presenilin is the molecular target of a dipeptidic gamma-secretase-specific inhibitor DAPT (N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester). *J. Biol. Chem.* 2006, *281*, 14670–14676.

[35] Whetstone, H., Lingwood, C., 3′-Sulfogalactolipid binding specifically inhibits Hsp70 ATPase activity *in vitro*. *Biochemistry* 2003, *42*, 1611–1617.

[36] Stewart, J. M., Blakely, J. A., Long chain fatty acids inhibit and medium chain fatty acids activate mammalian cardiac hexokinase. *Biochim. Biophys. Acta* 2000, *1484*, 278–286.

[37] Mithieux, G., Bordeto, J. C., Minassian, C., Ajzannay, A. *et al.*, Characteristics and specificity of the inhibition of liver glucose-6-phosphatase by arachidonic acid: Lesser inhibitability of the enzyme of diabetic rats. *Eur. J. Biochem.* 1993, *213*, 461–466.

[38] Ramadoss, C. S., Uyeda, K., Johnston, J. M., Studies on the fatty acid inactivation of phosphofructokinase. *J. Biol. Chem.* 1976, *251*, 98–107.

[39] Antikainen, J., Kuparinen, V., Lahteenmaki, K., Korhonen, T. K., pH-dependent association of enolase and glyceraldehyde-3-phosphate dehydrogenase of *Lactobacillus crispatus* with the cell wall and lipoteichoic acids. *J. Bacteriol.* 2007, *189*, 4539–4543.

[40] Hinson, D. D., Chambliss, K. L., Toth, M. J., Tanaka, R. D., Gibson, K. M., Post-translational regulation of mevalonate kinase by intermediates of the cholesterol and nonsterol

isoprene biosynthetic pathways. *J. Lipid Res.* 1997, *38*, 2216–2223.

[41] Qiu, Y., Li, D., Bifunctional inhibitors of mevalonate kinase and mevalonate 5-diphosphate decarboxylase. *Org. Lett.* 2006, *8*, 1013–1016.

[42] Khan, W. A., Blobe, G. C., Richards, A. L., Hannun, Y. A., Identification, partial purification, and characterization of a novel phospholipid-dependent and fatty acid-activated protein kinase from human platelets. *J. Biol. Chem.* 1994, *269*, 9729–9735.

[43] Collins, Q. F., Xiong, Y., Lupo, E. G., Jr., Liu, H. Y., Cao, W., p38 Mitogen-activated protein kinase mediates free fatty acid-induced gluconeogenesis in hepatocytes. *J. Biol. Chem.* 2006, *281*, 24336–24344.

[44] Haviv, H., Cohen, E., Lifshitz, Y., Tal, D. M. *et al.*, Stabilization of Na+, K+-ATPase purified from *Pichia pastoris* membranes by specific interactions with lipids. *Biochemistry* 2007, *46*, 12855–12867.

[45] Kapadia, S. B., Chisari, F. V., Hepatitis C virus RNA replication is regulated by host geranylgeranylation and fatty acids. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 2561–2566.

[46] Lee, W. M., Ishikawa, M., Ahlquist, P., Mutation of host delta9 fatty acid desaturase inhibits brome mosaic virus RNA replication between template recognition and RNA synthesis. *J. Virol.* 2001, *75*, 2097–2106.

[47] Lindsley, J. E., Rutter, J., Whence cometh the allosterome? *Proc. Natl. Acad. Sci. USA* 2006, *103*, 10533–10535.

[48] Hardy, J. A., Wells, J. A., Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* 2004, *14*, 706–715.

[49] Olsen, R. W., Chang, C. S., Li, G., Hanchar, H. J., Wallner, M., Fishing for allosteric sites on GABA(A) receptors. *Biochem. Pharmacol.* 2004, *68*, 1675–1684.

[50] Lundqvist, T., Fisher, S. L., Kern, G., Folmer, R. H. *et al.*, Exploitation of structural and regulatory diversity in glutamate racemases. *Nature* 2007, *447*, 817–822.

