JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Fully Functionalized Small-Molecule Probes for Integrated Phenotypic Screening and Target Identification

Justin S. Cisar and Benjamin F. Cravatt*

The Skaggs Institute for Chemical Biology and Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

Supporting Information

ABSTRACT: Phenotypic screening offers a powerful approach to identify small molecules that perturb complex biological processes in cells and organisms. The tendency of small molecules, however, to interact with multiple protein targets, often with moderate to weak affinities, along with the lack of straightforward technologies to characterize these interactions in living systems, has hindered efforts to understand the mechanistic basis for pharmacological activity. Here we address this challenge by creating a fully functionalized small-molecule library whose membership is endowed with: (1) one or more diversity elements to promote interactions with different protein targets in cells, (2) a photoreactive group for UV light-induced covalent cross-linking to interacting proteins, and (3) an alkyne handle for reporter tag conjugation to visualize and identify cross-linked proteins. A library member was found to inhibit cancer cell proliferation selectively under nutrient-limiting (low glucose) conditions. Quantitative chemoproteomics identified MT-ND1, an integral membrane subunit of the ~1 MDa NADH:ubiquinone oxidoreductase (complex 1) involved in oxidative phosphorylation, as a specific target of the active probe. We further demonstrated that the active probe inhibits complex 1 activity in vitro (IC_{50} = 720 nM), an effect that is known to induce cell death in low-glucose conditions. Based on this proof of principle study, we anticipate that the generation and integration of fully functionalized compound libraries into phenotypic screening programs should facilitate the discovery of bioactive probes that are amenable to accelerated target identification and mechanistic characterization using advanced chemoproteomic technologies.

A primary goal of modern chemical biology is to create small-molecule probes that selectively perturb individual biochemical pathways in cells and animals.[1] Such probes can help to elucidate the mechanistic basis for complex physiology and disease as well as serve as starting points for drug development to treat human disorders.[2−5] Phenotypic screening with small-molecule libraries has played a prominent role in the discovery of new chemical probes and drugs.[6] However, determining the biologically relevant targets of small-molecule hits from phenotypic (cell or organismal) screens remains a major technical challenge for multiple reasons.

First, hit compounds often bind to one or more protein targets with low to moderate affinity, and these relatively weak interactions can be difficult to characterize using conventional affinity chromatography methods, where the small molecule is linked to a solid support to enrich binding proteins from cell lysates.[7,8] Recent advances in the synthesis of small-molecule libraries with preinstalled linker elements[9] and improvements in quantitative proteomics[10] have enhanced the scope, sensitivity, and accuracy of such affinity enrichment experiments; but the approaches still involve identifying protein targets from cell lysates, and therefore, small molecule−protein interactions that depend on the integrity of the native cellular environment[11,12] remain difficult to characterize. Genetic approaches for target identification can address this problem,[13,14] but eliminating protein expression may not replicate the action of small molecules, which often produce effects through blockade (or activation) of multiple protein targets and without the physical removal of these proteins from cells. While small-molecule libraries that possess electrophilic groups offer a way to trap interacting proteins in situ through covalent reactivity,[11] we felt that this strategy, which depends on the fortuitous positioning of complementary nucleophilic residues in the binding pockets of target proteins, lacked generality for proteome-wide profiling in phenotypic screens. With these considerations in mind, we sought to develop a chemoproteomic strategy that could be combined with cell-based screening to enable the global and quantitative assessment of reversible small molecule−protein interactions directly in living systems.

Our strategy involves the synthesis of a small-molecule library that is "fully functionalized" to enable direct progression from cell-based screening to target identification without requiring synthetic modification of the biologically active hit compounds. Each library member is equipped with a photoreactive group that is embedded within its structure to permit UV light-induced cross-linking to interacting proteins directly in living cells. Each compound also contains an alkyne, which serves as a click chemistry[15,16] derivatizable handle for visualization, enrichment, and identification of interacting proteins.[17−19] We synthesized an ~30 member library based on two scaffolds: the 5-benzoyl indole (BzIndole) and 7-benzoyl-benzo-1,4-diazepin-2,5-dione (BzBD), where structural diversity was introduced at multiple sites on the core backbone of the molecules (Figure 1A and Figure S1, Supporting Information (SI)). Before proceeding to a phenotypic screen,

Received: May 1, 2012
Published: June 5, 2012

we first evaluated the proteomic interactions of library members in mammalian cells.

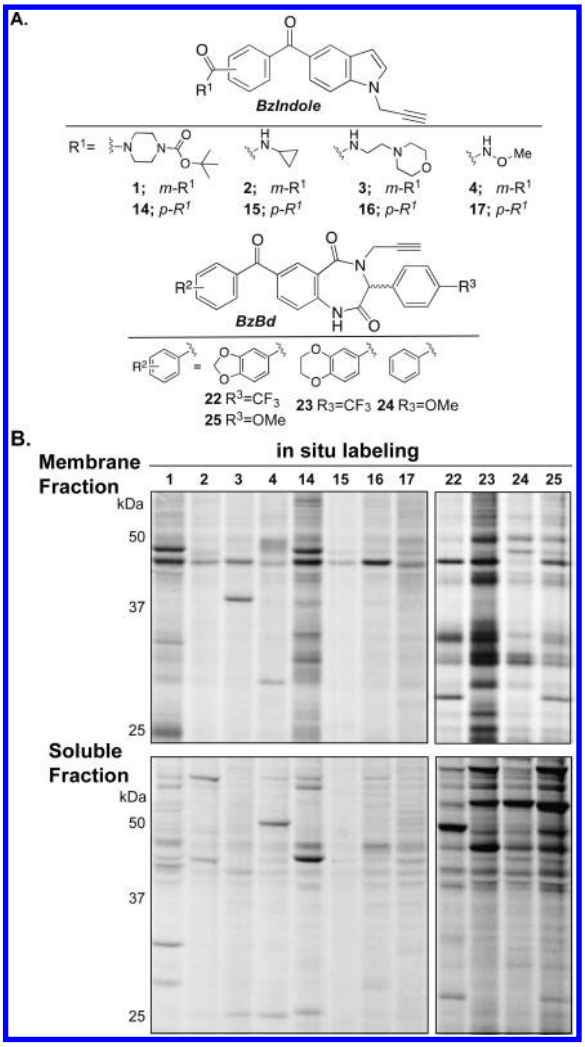

Figure 1. A fully functionalized small-molecule library for integrated cell-based screening and target identification. (A) BzIndole and BzBd scaffolds and structures of representative library members. See Figure S1, SI for structures of all synthesized members of the library. (B) Proteomic profiles for library members (10 μM) following in situ photo-cross-linking in MDA-MB-231 cells. Probe-labeled proteins were visualized by click chemistry conjugation to an azide-rhodamine tag, SDS-PAGE, and in-gel fluorescence scanning (fluorescent gels shown in gray scale).

The human breast cancer cell line MDA-MB-231 was treated with individual probes (10 μM, 40 min), cross-linked in situ with 365 nm light (5 min) and then lysed, and their proteomes were conjugated to rhodamine-azide[16] under click-chemistry conditions and visualized using SDS-PAGE and fluorescent scanning. Individual probes showed markedly distinct protein labeling events (Figure 1B), indicating that the diversity elements introduced into the probe library were sufficient to direct binding to different subsets of proteins in the proteome. Interestingly, these in situ profiles also differed considerably when compared to small molecule−proteome reactions performed in vitro (Figure S2, SI), indicating that the probes interact with different sets of proteins in living cells versus cell lysates.

We next screened the probe library for antiproliferative activity in MDA-MB-231 cells under both normal and low-glucose conditions (8 and 1 mM glucose, respectively). The former condition matches standard cell culture protocols, but the latter may better mimic the nutrient-deprived state encountered in solid tumors, which tend to be poorly vascularized.[20] While none of the compounds affected MDA-MB-231 proliferation in normal glucose, one compound, **1**, profoundly inhibited proliferation in low glucose (Figure 2A and Figure S3, SI). The antiproliferative activity of **1** was also observed in several other human cancer cell lines (Figure 2A) with IC_{50} values ranging from 0.33 to 1.3 μM (Figure 2B and Figure S3, SI).

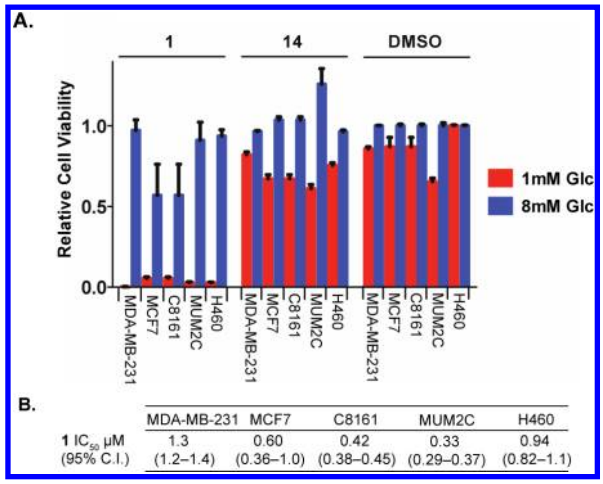

Figure 2. Antiproliferative activity of compound **1** in human cancer cell lines. (A) Inhibition of cancer cell viability in low glucose (1 mM) media after **1** treatment (24 h) as determined by the WST-1 assay. (B) Comparison of cell viability in low (1 mM) and normal (8 mM) glucose for cells treated with 10 μM active (**1**) or control (**14**) probes or vehicle (DMSO). Each cell line's viability was normalized to DMSO-treated cells grown in 8 mM glucose. See Figure S3, SI, for complete screening results.

Compound **1** shares structural features with other inactive library members, perhaps most notably compound **14**, which differs only in the position of the *t*-butoxycarbamoylpiperazine carbonyl group (Figure 1A). We therefore surmised that a comparison of the proteomic targets of **1** versus **14** could reveal proteins relevant for the antiproliferative activity of the former compound. We also synthesized a nonclickable analogue of **1** (compound **37**, Figure 3A) for use in competitive profiling experiments and confirmed that this agent retained antiproliferative activity (IC_{50} ~ 5 μM, Figure S4, SI). Gel-based profiling of proteomes from cells treated with **1**, **14**, or **1** + 5X **37** detected multiple proteins that were preferentially labeled by **1** over **14** and competed by **37** (Figure 3B). To identify these proteins, we performed quantitative proteomics experiments using stable isotope labeling by amino acids in cell culture (SILAC) methods[10] as follows: MDA-MB-231 cells were grown in media containing isotopically heavy- or light-labeled amino acids and then treated with probes **1** and **14**, respectively, or probe **1** ± 5X **37**, respectively. We also performed control experiments where both heavy- and light-labeled cells were treated with probe **1** or with **1** and DMSO, respectively. After in situ photo-cross-linking, cells were lysed and their proteomes conjugated to a biotin-azide tag, and then paired heavy and light proteomic samples were mixed in a 1:1

dx.doi.org/10.1021/ja304213w | J. Am. Chem. Soc. 2012, 134, 10385−10388

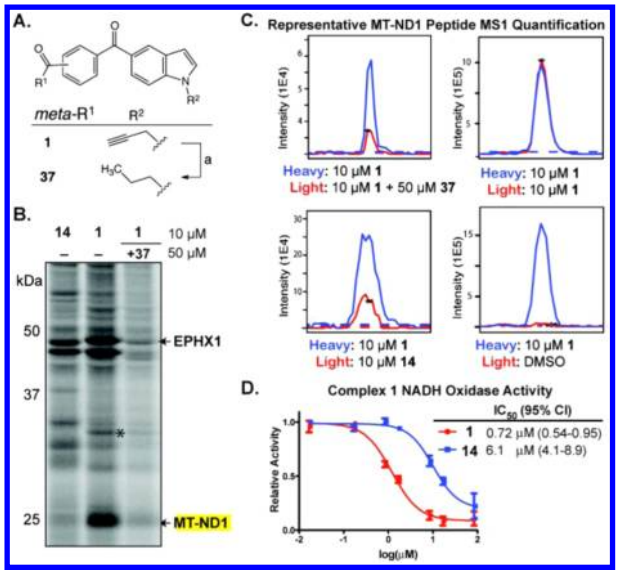

Figure 3. Chemoproteomic identification of proteins that interact with probe **1** in cancer cells. (A) A nonclickable analog of **1** (compound **37**) for competition profiling experiments (reagents: 1 atm H_2, cat. Pd–$BaSO_4$, EtOH). (B) Gel profile comparing proteins labeled by **1** and **14** (10 μM) and **1** competed with 5X **37** (50 μM). Candidate protein bands corresponding to selective targets of **1** identified by LC-MS are denoted on right. (C) Extracted MS1 chromatographs for a representative peptide (ILGYMQLR, 497.28114 m/z) from MT-ND1 shown for **1** ± 5X **37** competition, **1** versus **1**, **1** versus **14**, and **1** versus DMSO (unlabeled) experiments. See Table 1 and Figure S5 and Tables S1 and S2, SI, for more complete presentations of the quantitative proteomic data. (D) Assay for complex 1 NADH oxidase showing selective inhibition by **1** over **14**.

ratio, separated into soluble and membrane fractions, enriched with avidin, digested on a bead with trypsin, and analyzed by LC-MS/MS using an LTQ-Orbitrap instrument. Light and heavy signals were quantified from the parent ion (MS1) peaks using the CIMAGE software[21] to generate SILAC ratios for each detected peptide, and corresponding protein identities were determined from product ion peptide profiles (MS2) using the ProLuCID search algorithm.[22]

Two proteins fit the criteria of being labeled by **1** preferentially over **14** and competed at least 2-fold by 5X **37**: 1) epoxide hydrolase-1 (EPHX1), a 50 kDa microsomal enzyme that metabolizes exogenous and endogenous epoxide metabolites[23] (Table 1 and Figure S5 and Tables S1 and S2, SI); 2) MT-ND1, an integral membrane subunit of the ubiquinone oxidoreductase (complex 1) involved in mitochondrial respiration (Table 1, Figure 3C, and Table S1, SI).[24] MT-ND1 is predicted to be a 35 kDa protein based on amino acid sequence but has been shown in several previous studies to migrate as an ~25 kDa protein by SDS-PAGE.[25−28] Gel-based profiling identified candidate **1**-labeled bands that correspond to both proteins (as deduced by their preferential labeling by **1** over **14** and competition by excess **37**; Figure 3C). Gel-based profiling also detected an additional ~35 kDa target selectively reactive with **1** over **14** (Figure 3C, asterisk), but a corresponding protein target that matches this molecular weight was not detected in our mass spectrometry analyses, perhaps because the comparatively weak probe-labeling signals for this protein precluded sufficient enrichment for its identification. While it is possible that **1** produces its antiproliferative effect through a combination of these protein

Table 1. High-Confidence Protein Targets of Compound 1[a]

		1 vs **14** experiment		**1** vs **1** + 5X **37** experiment	
gene name	MW (kDa)	av ratio	SEM	av ratio	SEM
EPHX1	53	5.19	0.76	2.27	0.33
MT-ND1	35.7	3.00	0.39	2.38	0.42
APMAP	46.5/32.2	1.82	0.07	2.24	0.38
SCCPDH	47.1	1.64	0.17	5.30	0.44
PIGS	61.6	1.38	0.47	2.27	0.76
TRABD	42.3/37.1	1.30	0.20	2.34	0.54
HEBP1	21.1	0.94	0.32	2.30	0.77

[a]Highlighted are two proteins that showed >2-fold higher SILAC ratios in quantitative chemoproteomic studies of **1**- versus **14**-treated and **1**- versus **1** + 5X **37**-treated MDA-MB-231 cells. Additional protein targets that showed >2-fold higher SILAC ratios in versus **1** + 5X **37**-treated MDA-MB-231 cells are also shown.

targets, MT-ND1 stood out, in our mind, as the most likely candidate, considering that other complex 1 inhibitors are known to impair cancer cell viability under low-glucose conditions.[29,30] We therefore assayed **1** and control compound **14** for their ability to inhibit complex 1 activity in purified submitochondrial particles. Complex 1 activity was inhibited more potently by **1** (IC_{50} = 720 nM) than **14** (IC_{50} = 6.1 μM) (Figure 3C), consistent with the relative labeling of MT-ND1 by these compounds in cancer cells (Figure 3B,D). We also confirmed that **1** selectively labels EPHX1 by recombinantly expressing this protein in HEK293T cells (Figure S6, SI). These data, taken together, demonstrate that the antiproliferative agent **1** selectively targets only a handful of proteins in cancer cells, including MT-ND1, the inhibition of which would be predicted to impair mitochondrial respiration and lead to cancer cell death under nutrient-limiting conditions.

The importance of phenotypic screening as a source for new chemical probe and drug discovery is well-recognized.[6] However, determining the mechanism of action for hit compounds from cell-based assays remains a major challenge due, at least in part, to a lack of robust technologies for target identification. Here, we have described a strategy for generating small-molecule libraries that addresses this problem. Each library member is "fully functionalized" to contain photoactivatable and alkyne groups for *in situ* cross-linking to protein targets and chemoproteomic characterization of these small molecule–protein adducts, respectively. As a proof of principle, we identified a small-molecule probe **1** that inhibits hypoglycemic cancer cell proliferation and determined that one of its major cellular targets is the complex 1 component MT-ND1. We further validated that **1** inhibits the NADH oxidase activity of complex 1, similar to other known inhibitors of the complex that elicit glucose-dependent antiproliferative effects in cancer cells. It is noteworthy that previous complex 1 inhibitors have been converted into radiolabeled, photoreactive analogs to ascertain their subunit interactions and binding site information;[25−28] but these studies have often spanned many years in duration, likely due to the technical challenges that accompany incorporation of photo-cross-linkers and radioisotopes into bioactive compounds as well as the difficulties inherent to the biochemical purification and analysis of multisubunit integral membrane protein complexes. By using a fully functionalized library of small molecules in which photoreactive groups, alkyne affinity handles, and diversity elements are embedded into the core structure of the library,

dx.doi.org/10.1021/ja304213w | J. Am. Chem. Soc. 2012, 134, 10385−10388

we have provided a more streamlined platform for uniting the steps of phenotypic screening and target identification.

Projecting forward, we should mention some challenges, perhaps first and foremost the need to create fully functionalized libraries of much greater size and complexity. Here, the development of synthetic methods that allow efficient incorporation of additional photo-cross-linking groups beyond the benzophenone, such as diazirines or aryl azides, would be helpful in expanding the structural diversity of libraries. Another potentially confounding feature is photo-cross-linking yields, which can vary considerably depending on the proximity of the photoreactive group to interacting protein(s). We anticipate that embedding the photoreactive group within the core scaffold of libraries should help to maximize protein interactions and cross-linking yields. Finally, we should emphasize that hit compounds emerging from phenotypic screens with fully functionalized libraries should be viewed as tool compounds as opposed to drug candidates. But, by defining the target profile required to produce a desired biological effect as well as establishing a structure−activity relationship for this activity, we anticipate that the described strategy will set the stage for future medicinal chemistry efforts to convert chemical probes into more drug-like small molecules.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary figures, experimental protocols, synthetic schemes, and compound characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

cravatt@scripps.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Takao and Akemi Yagi who generously provided purified submitochondrial particles. This work was supported by NIH grant GM090294, a Ray Kathren American Cancer Society fellowship (J.S.C.), and a Ruth L. Kirschstein National Research Service Award F32GM095245-01 (J.S.C.).

■ REFERENCES

(1) Paolini, G. V.; Shapland, R. H.; van Hoorn, W. P.; Mason, J. S.; Hopkins, A. L. *Nat. Biotechnol.* **2006**, *24*, 805.
(2) Yagoda, N.; von Rechenberg, M.; Zaganjor, E.; Bauer, A. J.; Yang, W. S.; Fridman, D. J.; Wolpaw, A. J.; Smukste, I.; Peltier, J. M.; Boniface, J. J.; Smith, R.; Lessnick, S. L.; Sahasrabudhe, S.; Stockwell, B. R. *Nature* **2007**, *447*, 864.
(3) Chen, S.; Do, J. T.; Zhang, Q.; Yao, S.; Yan, F.; Peters, E. C.; Scholer, H. R.; Schultz, P. G.; Ding, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17266.
(4) Petrascheck, M.; Ye, X.; Buck, L. *Nature* **2007**, *450*, 553.
(5) Guiguemde, W. A.; Shelat, A. A.; Garcia-Bustos, J. F.; Diagana, T. T.; Gamo, F. J.; Guy, R. K. *Chem. Biol.* **2012**, *19*, 116.
(6) Swinney, D. C.; Anthony, J. *Nat. Rev. Drug Discovery* **2011**, *10*, 507.
(7) Leslie, B. J.; Hergenrother, P. J. *Chem. Soc. Rev.* **2008**, *37*, 1347.
(8) Burdine, L.; Kodadek, T. *Chem. Biol.* **2004**, *11*, 593.
(9) Ahn, Y. H.; Chang, Y. T. *Acc. Chem. Res.* **2007**, *40*, 1025.
(10) Ong, S. E.; Schenone, M.; Margolin, A. A.; Li, X.; Do, K.; Doud, M. K.; Mani, D. R.; Kuai, L.; Wang, X.; Wood, J. L.; Tolliday, N. J.; Koehler, A. N.; Marcaurelle, L. A.; Golub, T. R.; Gould, R. J.; Schreiber, S. L.; Carr, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4617.
(11) Evans, M. J.; Saghatelian, A.; Sorensen, E.; Cravatt, B. F. *Nat. Biotechnol.* **2005**, *23*, 1303.
(12) Salisbury, C. M.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1171.
(13) Wacker, S. A.; Houghtaling, B. R.; Elemento, O.; Kapoor, T. M. *Nat. Chem. Biol.* **2012**, *8*, 235.
(14) Ho, C. H.; Piotrowski, J.; Dixon, S. J.; Baryshnikova, A.; Costanzo, M.; Boone, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 66.
(15) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2596.
(16) Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
(17) Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535.
(18) Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Amer. Chem. Soc.* **2003**, *125*, 4686−4687.
(19) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Amer. Chem. Soc.* **2003**, *125*, 3192−3193.
(20) Vaupel, P.; Kallinowski, F.; Okunieff, P. *Cancer Res.* **1989**, *49*, 6449.
(21) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. *Nature* **2010**, *468*, 790.
(22) Tayo, L. L.; Lu, B.; Cruz, L. J.; Yates, J. R., III *J. Proteome Res.* **2010**, *9*, 2292.
(23) Morisseau, C.; Hammock, B. D. *Annu. Rev. Pharmacol. Toxicaol* **2005**, *45*, 311.
(24) Brandt, U. *Annu. Rev. Biochem.* **2006**, *75*, 69.
(25) Murai, M.; Ishihara, A.; Nishioka, T.; Yagi, T.; Miyoshi, H. *Biochemistry* **2007**, *46*, 6409.
(26) Murai, M.; Sekiguchi, K.; Nishioka, T.; Miyoshi, H. *Biochemistry* **2009**, *48*, 688.
(27) Kakutani, N.; Murai, M.; Sakiyami, N.; Miyoshi, H. *Biochemistry* **2010**, *49*, 4794−4803.
(28) Murai, M.; Mashimo, Y.; Hirst, J.; Miyoshi, H. *Biochemistry* **2011**, *50*, 6901.
(29) Ahmad, I. M.; Aykin-Burns, N.; Sim, J. E.; Walsh, S. A.; Higashikubo, R.; Buettner, G. R.; Venkataraman, S.; Mackey, M. A.; Flanagan, S. W.; Oberley, L. W.; Spitz, D. R. *J. Biol. Chem.* **2005**, *280*, 4254.
(30) Han, M.; Im, D.-S. *Arch. Pharm. Res.* **2008**, *31*, 749.