

Direct screening of solution phase combinatorial libraries encoded with externally sensitized photolabile tags

Rudresha Kottani, Roman A. Valiulin, and Andrei G. Kutateladze*

Department of Chemistry and Biochemistry, University of Denver, Denver, CO 80208-2436

Communicated by Howard E. Zimmerman, University of Wisconsin, Madison, WI, July 28, 2006 (received for review June 1, 2006)

Solution phase combinatorial chemistry holds an enormous promise for modern drug discovery. Much needed are direct methods to assay such libraries for binding of biological targets. An approach to encoding and screening of solution phase libraries has been developed based on the conditional photorelease of externally sensitized photolabile tags. The encoding tags are released into solution only when a sought-for binding event occurs between the ligand and the receptor, outfitted with an electron-transfer sensitizer. The released tags are analyzed in solution revealing the identity of the lead ligand or narrowing the range of potential leads.

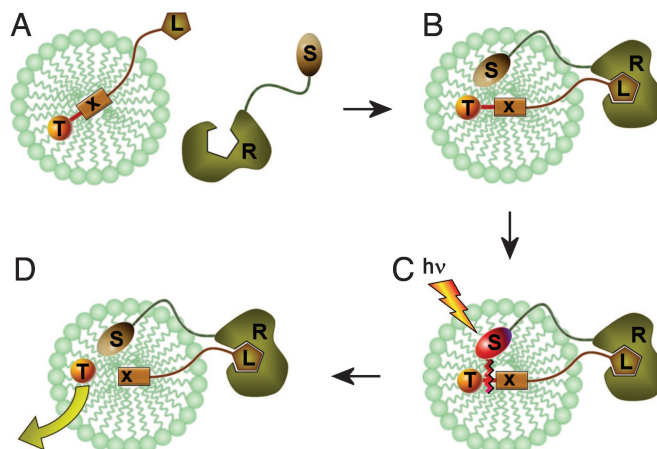
dithiane tag | photoinduced externally sensitized fragmentation | solution library encoding and screening

High-throughput combinatorial chemistry is critical for modern drug discovery, with the most prolific adaptation being the split-and-mix synthesis of libraries immobilized on polymeric beads (1, 2). These one-bead-one-compound libraries are assayed for binding of biological targets through fluorescence-guided mechanical segregation of the winning beads. The required mechanical manipulations impose a lower limit on the size of the particle used as a solid support. It is further recognized that solution phase combinatorial chemistry holds even greater promise, as it is compatible with both divergent and convergent multistep synthetic schemes, and not constrained to linear synthesis (3). However, its immense potential has not yet been fully realized, partly due to the complexity of assaying solution mixtures for binding. Although various iterative deconvolution methods (4) have been developed to synthesize and screen soluble sublibraries, all of them require redundant synthetic steps and are time consuming and expensive.

We have developed a methodology for direct screening of solution phase libraries encoded with photolabile tags that does not depend on mechanical manipulation or iterative deconvolution.

Results and Discussion

Our tagging methodology is based on dithiane adducts of carbonyl compounds, which are capable of efficient photoinduced fragmentation, but only in the presence of an external electron transfer (ET) sensitizer. At the core of this approach is the concept we term photolabile scaffolds for molecular recognition: binary molecular systems in which the sensitizer and the dithiane-based photocleavable fragments are each tethered to the respective components of a host-guest molecular recognition pair (5–7). In these pairs, photoinduced fragmentation is contingent on a molecular recognition event, which is necessary to arm the system, making it photolabile. This general concept has now been developed into a methodology for screening of encoded libraries, both unsupported or immobilized on nanosized carriers too small for mechanical handling. Scheme 1 gives a general outline for screening the micelle-solubilized library of ligands L, encoded with the tethered tags T. The receptor R is outfitted with an ET-sensitizer S, and incubated with the library (A). The host-guest binding brings the sensitizer into the vicinity



Scheme 1.

of the adduct (B), which, upon irradiation (C), triggers the release of dithiane tags (D).

Our strategy is to encode individual molecules one tag at a time. The k th library member L_k can be encoded with a set of tags $\{T\}_k$, for example, T_{k1} , T_{k2} , and T_{k3} , in the following fashion: one fraction of L_k molecules is encoded with T_{k1} , another fraction is encoded with T_{k2} , and yet another is encoded with T_{k3} , such that L_k is present in the solution as three subpopulations: L_k -tether- T_{k1} , L_k -tether- T_{k2} , and L_k -tether- T_{k3} . In this case, irradiation yields the desired result due to the fact that all of the tags encoding individual molecules of the bound ligand are collectively released into the solution, revealing the identity of the lead compound.

The following considerations were taken into account in the designing of tags. The tags should (i) be amendable for detection at very low concentrations by ubiquitous analytical techniques; (ii) not possess any functional groups implicated in biological interactions; (iii) not interfere with synthetic steps; (iv) be easy to separate from the aqueous screening environment, for example, due to significant hydrophobicity; and (v) be accessible via a simple and straightforward synthetic approach. We will demonstrate that 2-alkyl substituted dithianes satisfy these requirements and can be used for encoding as binary digits, i.e., 1 indicates that free dithiane is present in the solution and 0 indicates that no dithiane is found.

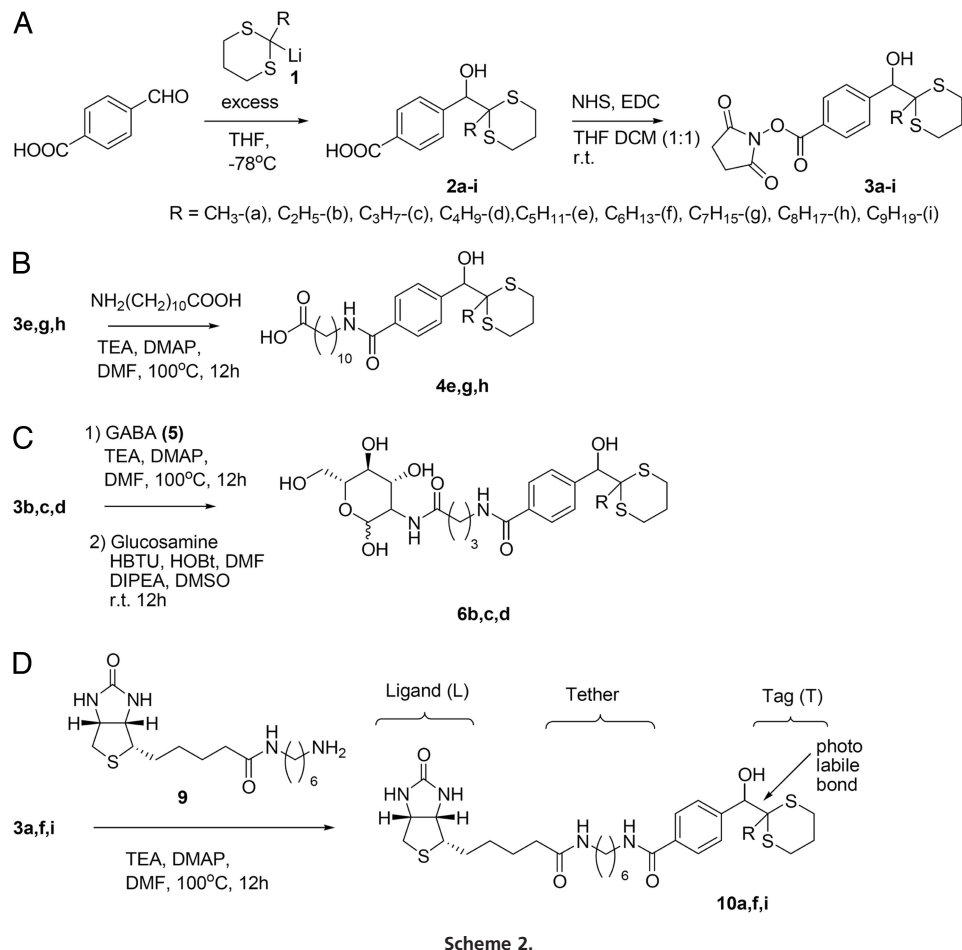
Author contributions: R.K. and R.A.V. performed research; A.G.K. designed research; A.G.K. analyzed data; and A.G.K. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: ET, electron transfer; GCMS, gas chromatography mass spectrometry; EI, electron impact ionization.

*To whom correspondence should be addressed. E-mail: akutatel@du.edu.

© 2006 by The National Academy of Sciences of the USA



The ability to detect tags at very low levels is critical for any successful library encoding scheme. We have found that dithianes are particularly suitable for subpicomolar detection by gas chromatography with mass-selected detector (GCMS) with electron impact ionization. A better than 10:1 signal-to-noise ratio is achieved at a level of 500 femtomoles of alkyl dithianes per injection. In theory, for a library of one million tagged compounds, averaging 1–2 kDa, only a few milligrams are needed for the binding assay. The adducts of 2-alkyl dithianes and 4-formylbenzoic acid were chosen as tagging modules because alkyl dithianes add readily to nonenolizable aromatic aldehydes and the carboxylate serves as a practical handle to tether the tag to a ligand. A series of *N*-hydroxysuccinimide esters **3a–i** were synthesized as shown in Scheme 2A (for experimental details, refer to *Supporting Materials and Methods*, which is published as supporting information on the PNAS web site).

To prove the concept we chose the archetype host-guest system, biotin-avidin, and synthesized a model minilibrary comprising three members (Scheme 2B–D), each encoded with a set of three dithiane tags: a carboxylate **4e,g,h** encoded with 2-pentyl, 2-heptyl and 2-octyl dithianes (decimal 208; binary 11010000), a sugar **6b,c,d** (ethyl-, propyl-, and butyl; 14; 1110), and biotin **10a,f,i** (methyl-, hexyl-, and nonyl; 289; 100100001). In this model library, the encoding of individual members was intended not to overlap, to accurately assess the validity of the method and ensure unambiguous interpretation of the screening results.

The receptor, ImmunoPure avidin (Pierce, Rockford, IL), was outfitted with xanthone as an ET-sensitizer. The *N*-hydroxysuccinimide ester of 11-(xanthone-2-carboxamido)undecanoic acid

was coupled to avidin in a 20 mM sodium phosphate buffer according to a described procedure (8), with subsequent purification on a Sephadex G-25 column. The degree of immobilization was quantified by UV spectroscopy to be 0.77, indicating that, on average, each tetramer of avidin was carrying approximately three tethered xanthone carboxylates. Tethering the xanthone-based sensitizer to avidin is, in a way, no different from outfitting a receptor with a fluorophore for the conventional on-the-bead fluorescence-guided assays.

To ensure fidelity of the method we compartmentalized the screening volume with micellar detergent, dodecyl phosphocholine (DPC), preventing indiscriminate collisional quenching of avidin-tethered xanthone by unbound molecules and thus limiting the ET sensitization exclusively to the bound host–guest complex. Besides the fact that amphiphiles displaying phosphocholine are common in biological settings, utilization of DPC micelles offers a number of additional benefits: (i) it ensures that the screening is always compatible with aqueous media, regardless of aqueous solubility of the tested libraries; (ii) it allows us to center the design of the tagging system around readily available hydrophobic alkyl dithianes, which can be selectively extracted after irradiation with hexane or other nonpolar solvents for unobstructed GCMS analysis; (iii) it spatially segregates the photofragmentation chemistry from molecular recognition, eliminating potential interference between them; (iv) it restricts the photochemistry to the micelle interior improving quantum efficiency of fragmentation, as it is known to increase in the nonpolar environment; and (v) the micelle-assisted design offers an option of solubilizing certain target proteins, which are not water soluble. Although not an issue with avidin, this

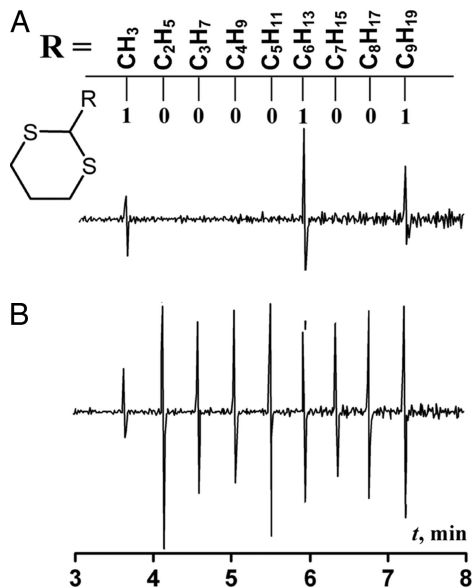


Fig. 1. The first derivative GCMS single ion monitoring (SIM) traces. (A) The trace encoding biotin in binary 100100001, obtained after the photolytic assay. (B) The trace for all nine alkyl dithianes at 1 pmol per injection.

functionality may prove critical in assaying insoluble membrane proteins.

A typical screening procedure involved solubilization of the minilibrary, ≈ 0.5 mg per tagged compound, in a 0.6-ml aqueous solution containing 60 mg of DPC. To this clear micellar solution, 0.5 ml of avidin–xanthone conjugate was added, so the final concentrations were 0.7 mM of each library member carrying one tag (6.3 mM total), 23 μ M protein, and 155 mM DPC. The micelle-embedded molecules had apparent translational diffusion coefficients of 7×10^{-7} cm²·s⁻¹, as measured with spin-echo pulse field gradient (PFG) ¹H NMR. This corresponds to the hydrodynamic radius of ≈ 3 nm, indicating that the occupied DPC micelles did not deviate much from their original 5- to 5.5-nm size. The resulting micellar solution was incubated in an orbital shaker for 1 h, purged with argon for 45 min, and irradiated for 4 h by using a 335-nm long pass filter. Then the mixture was extracted with 0.5 ml of hexane, concentrated to 100 μ l and analyzed by GCMS. Fig. 1 clearly shows that only the biotin-encoding tags, namely methyl-, hexyl-, and nonyl-dithianes (binary 100100001, read from left to right), were detected in the chromatogram. The other six dithiane tags encoding glucosamine and aminoundecanoic acid were not discernible at all, attesting to the high fidelity of the assay.

The integrated intensity of dithiane peaks in most experiments was comparatively uniform within 30–50%. The quantum yield of alkyl dithiane photo-release from ketone adducts increases in small increments of 2–3% in a homologous series, leveling off for the higher alkyls (9). The GCMS sensitivity of dithianes detection also varies insignificantly. If needed, these small variations can be offset by adjusting the quantities of individual tags used for encoding of each step.

To demonstrate that the compartmentalization requirement is not necessarily strict, one micelle in the described screening experiment contained on average two tagged library molecules (assuming that the aggregation number of DPC is 50–60; ref. 10). In theory, if the protein-tethered sensitizer indiscriminately releases both tags from the two occupants of the bound micelle, the integrated intensity of the false peaks in the chromatogram should constitute more than one third of the correct peak's intensity. Experimentally we did not detect any false tags, with

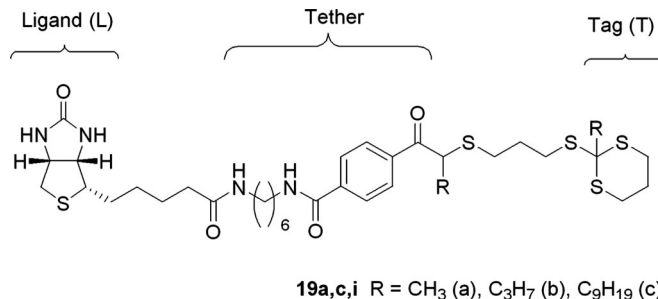


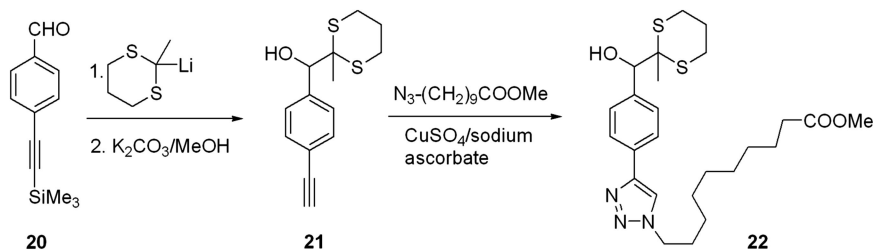
Fig. 2. Thio ortho esters-encoded biotin.

the signal to noise ratio of the single-ion monitoring current exceeding 20:1. This shows that either the sensitizer discriminates between the bound and nonbound occupants of the micelle, preferentially triggering the release of the bound tag, or the micelles containing two biotin molecules bind much better than the micelles containing only one biotin, in which case false release is not at all possible. It is also conceivable that both factors operate concurrently, improving the fidelity of the method. If needed, a one-molecule–one-micelle compartmentalization can be readily achieved in practice by increasing the detergent concentration.

The generality of this approach is not limited to the tagging assemblies based on dithiane–aldehyde adducts. Potentially, any externally sensitized fragmentation reaction can be used in such binding assays. We achieved similar results with different tags comprised of thio ortho esters, 2-alkylthio-dithianes (11). The model minilibrary was tagged with nine thio ortho esters (19), three tags per library member, with biotin encoded by a different decimal 261, binary 10000101 (Fig. 2). The DPC micellar solution of the library was incubated with the avidin–xanthone conjugate, irradiated and extracted with hexane. The GCMS trace again showed only the dithianes encoding biotin. None of the other six dithianes were detected in the hexane extract after irradiation.

In order for this screening methodology to afford maximal information content, the encoding strategy is to tag each reagent used at each step of combinatorial synthesis, so that every library member is encoded with the same total number of tags, equal to the number of synthetic steps. As an example, a library of pentapeptides with 7 possible amino acid residues, containing $7^5 = 16,807$ members, would need $5 \times 7 = 35$ tags for encoding. The same 35 tags, of course, encode a library of $5^7 = 78,125$ heptapeptides with five possible residues per position. The one-tag-per-reagent encoding is much more informative and less ambiguous than the conventional binary scheme suitable for the libraries immobilized on polymeric beads (12, 13). For example, tags T₁ through T₇ each encode the first-step reagents A₁ through A₇; tags T₈ through T₁₄ encode the second position, and so on. In the best case scenario, when the assayed library contains a lead ligand with much higher binding affinity than the next closest competitor, our photoassay identifies such a ligand in just one screening. In cases when a library contains several high-affinity ligands with very similar K_D values, analysis of multiple GCMS peaks within each individual range, i.e., tags 1–7, tags 8–14, etc., should dramatically reduce the number of possibilities, providing statistically significant information about the most preferred components/reagents at each position/synthetic step. If needed, a much smaller sublibrary can be synthesized and the lead compound determined.

Of more importance is the actual binding, not the mathematics of deconvolution. In a typical case of fluorescence-guided on-the-bead assays, the multiple hits are not due to the fact that the respective K_D values are nearly identical, but rather because the



Scheme 3.

receptor is overloaded and free to saturate several binding equilibria. Subsequent dilution of the receptor to very low concentrations should help limit the number of the identified winners, with the constraint being the dithiane detection limit. As we demonstrated, we can confidently detect subpicomolar amounts of dithianes per injection by using conventional analytical instrumentation. Extracted dithianes can be further pre-concentrated before GC injection, allowing for handling the full physiological range of K_D values, down to subnanomolar.

In an early classical example by Lam *et al.* (14), a large pentapeptide library of 2,476,099 members was synthesized and screened against a monoclonal antibody, known to bind the β -endorphin epitope (YGGFL). The authors identified four peptide ligands with K_D values under 37 nM: YGGFQ (15.0 nM), YGGFL–native ligand (17.5 nM), YGGFA (32.9 nM), and YGGFT (36.9 nM). The K_D for the next best ligand, YGGLS, was 20 times greater (726 nM). In a case like this, our photo-assay methodology, with a sensitizer-tethered antibody, would produce one peak in each range of tags encoding positions 1–4, YGGF, and four peaks in the range of tags encoding position 5. Thus, in a single solution-phase screening, all four winners can be unambiguously identified.

Dynamic encoding of libraries, prepared by the split-pool approach, is normally achieved with a coupling scheme orthogonal to the used synthetic steps. The tagging synthons, esters 3, can be coupled with the library molecules via tethers containing primary amino groups, as in conventional polypeptide encoding (15, 16), which involves the appropriate cycling of the BOC or Fmoc protection. Alternatively, the esters 3 can be used for dynamic tagging in conjunction with other efficient coupling reactions, such as Staudinger ligation (17–19) or Sharpless' click chemistry (20).

To this end we coupled 3a with propargyl amine, providing the acetylenic tagging component S20 for the azide-acetylene click pair. Another readily available tag series is prepared from commercially available acetylenic benzaldehyde as shown in Scheme 3. Whereas the Staudinger ligation produces the benzamide linkage, found in compounds 4, 6, and 10, for which we have proven the concept, acetylene–azide click chemistry yields triazoles, which needed further testing to ascertain their compatibility with photoinduced fragmentation and the release of dithiane tags. Methyl 10-azidodecanoate, emulating an azide-tethered library member, was “clicked” onto adduct 21, forming triazole 22, which upon benzophenone sensitization released methyl dithiane with a quantum efficiency very similar to the parent (unsubstituted) benzaldehyde adduct. Triazole 22 was unchanged after prolonged irradiation in the absence of the ET-sensitizer, showing no self-cleavage at wavelengths >330 nm. This was a critically important finding because premature self-cleavage is detrimental to screening, as it produces false positives.

In conclusion, we have described here a methodology for encoding and screening of combinatorial libraries based on externally sensitized photolabile tags. The approach does not depend on mechanical handling of the beads, allowing for the

screening of unsupported solution phase libraries and, potentially, libraries immobilized on nano-carriers, including dendrimers. This methodology was tested on a model minilibrary, for which the high-fidelity of screening was demonstrated. We also have prepared acetylene-bearing tags for encoding via click chemistry and have outlined a general approach to dynamic tagging of libraries. Our findings have the potential to fundamentally transform combinatorial encoding and screening of solution phase libraries, as currently there are no direct methods to assay solution mixtures for binding.

Materials and Methods

Common reagents were purchased from Sigma-Aldrich (St. Louis, MO) and used without further purification. THF was refluxed over and distilled from potassium benzophenone ketyl before use. 1H and ^{13}C NMR spectra were recorded at 25°C on a Varian Mercury 400 MHz instrument, in $CDCl_3$, $DMSO-d_6$, or CD_3OD using tetramethylsilane as an internal standard. Pulsed Field Gradient NMR studies were carried out with the Varian Performa I PFG module and a four-nuclei autoswitchable PFG probe. Column chromatography was performed on silica gel, 70–230 mesh. The UV-visible spectra were recorded on a Beckman DU-640 spectrophotometer. Irradiations were carried out in a carousel Rayonet photo reactor (Southern New England Ultra Violet, Branford, CT), equipped with RPR-3500 UV lamps (300- to 400-nm spectral distribution of irradiance density with a maximum at 350 nm) and a 335-nm long-pass solution filter. Gas chromatography was performed on a Varian Saturn 2100 T Ion-Trap instrument with electron ionization. Selective ion monitoring (m/z 119, 74) was used to detect dithiane tags after fragmentation.

Dithiane–Aldehyde Adducts 2(a–i). Dithiane–aldehyde adducts 2(a–i) were synthesized according to the general procedure: n-BuLi (14.6 ml, 23.3 mmol, 1.6 M solution in THF) was added at $-25^\circ C$ to a solution of 2-alkyl-1,3-dithiane (23.3 mmol) in dry THF (40 ml) under nitrogen atmosphere. The solution was stirred for 2 h, the temperature was reduced to $-78^\circ C$, and 4-formylbenzoic acid (0.5 g, 3.33 mmol) in 20 ml of THF was added. The solution was stirred at $-78^\circ C$ for an additional 2 h and then allowed to warm to room temperature. The reaction was quenched with saturated aqueous ammonium chloride (20 ml) and extracted twice with 20 ml of ethyl acetate. The aqueous layer was acidified with 5% HCl and extracted with ethyl acetate (100 ml), and the extract was dried over Na_2SO_4 . The solvent was removed in vacuum. The crude product was recrystallized from toluene to furnish adducts 2a–i in 90–96% yields.

Tagging of biotin, glucosamine and aminoundecanoic acid with the encoding pendants 2a–i was carried out according to modified literature procedures by using EDC or DIPEA coupling reagents. All members of the two model minilibraries (10a–i, 16f,g,h, 18b,d,e, and 19a,c,i) were individually purified and characterized by 1H and ^{13}C NMR. For detailed synthetic procedures and spectra, see *Supporting Materials and Methods*.

Xanthone Conjugate of Avidin. 100 μ l of **13** (5 mg/ml in DMSO) was added to a solution of avidin (Pierce, Rockford, IL) (10 mg, 0.147 μ mol) in 1 ml of 20 mM sodium phosphate buffer. The mixture was gently shaken for 2 h at ambient temperature, followed by purification on a Sephadex G-25 column. The degree of immobilization was quantified spectroscopically to be 0.77 as described in *Supporting Materials and Methods*.

Photochemical Screening. In a typical screening experiment, 0.5 mg each of **10a**, **6b**, **6c**, **6d**, **4e**, **10f**, **4g**, **4h**, and **10i** were added to a solution of DPC (Anatrace, Maumee, OH) (60 mg) in 0.6

ml of D₂O. The mixture was stirred for 24 h at ambient temperature until a clear micellar solution was obtained, and 0.5 ml of the avidin–xanthone conjugate was added. This solution was incubated with shaking for 1 h, degassed with argon for 45 min, and irradiated for 4 h by using Rayonet reactor (RPR-3500 lamps) with a 335-nm long-pass solution filter. The micellar solution was extracted with 0.5 ml of hexane, concentrated to 0.1 ml, and analyzed by GCMS with single ion monitoring.

This work was supported by National Science Foundation Grant CHE314344 and National Institutes of Health Grant GM067655.

1. Lam KS, Lebl M, Krchňák V (1997) *Chem Rev* 97:411–448.
2. Lam KS, Liu R, Miyamoto S, Lehman AL, Tuscano JM (2003) *Acc Chem Res* 36:370–377.
3. Baldino CM (2000) *J Combinatorial Chem* 2:89–103.
4. Konings DAM, Wyatt JR, Ecker DJ, Freier SM (1997) *J Med Chem* 40:4386–4395.
5. Mitkin OD, Kurchan AN, Wan Y, Schiwal BF, Kutateladze AG (2001) *Org Lett* 3:1841–1844.
6. Mitkin O, Wan Y, Kurchan A, Kutateladze A (2001) *Synthesis* 7:1133–1142.
7. Wan Y, Angleson JK, Kutateladze A. G (2002) *J Am Chem Soc* 124:5610–5611.
8. Wilchek M, Bayer, EA, eds (1990) *Methods Enzymol* 184:176.
9. Gustafson TP, Kurchan AN, Kutateladze AG (2006) *Tetrahedron* 62:6574–6580.
10. Brown LR, Bösch C, Wüthrich K (1981) *Biochim Biophys Acta* 642:296–312.
11. Valiulin RA, Kottani R, Kutateladze AG (2006) *J Org Chem* 71:5047–5049.
12. Ohlmeyer MHJ, Swanson RN, Dillard LW, Reader JC, Asouline G, Kobayashi R, Wigler M, Still WC (1993) *Proc Natl Acad Sci USA* 90:10922–10926.
13. Borchardt A, Still WC (1994) *J Am Chem Soc* 116:373–374.
14. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ (1991) *Nature* 354:82–84.
15. Franz AH, Liu R, Song A, Lam KS, Lebrilla CB (2003) *J Comb Chem* 5:125–137.
16. Liu R, Marik J, Lam KS (2002) *J Am Chem Soc* 124:7678–7680.
17. Saxon E, Bertozzi CR (2000) *Science* 287:2007–2010.
18. Saxon E, Armstrong JI, Bertozzi CR (2000) *Org Lett* 2:2141–2143.
19. Nilsson BL, Kiessling LL, Raines RT (2000) *Org Lett* 2:1939–1941.
20. Kolb HC, Finn MG, Sharpless KB (2001) *Angew Chem Int Ed* 40:2004–2021.