

# Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity

Brent R Martin<sup>1,2</sup>, Ben N G Giepmans<sup>1,3,4</sup>, Stephen R Adams<sup>1</sup> & Roger Y Tsien<sup>1,5,6</sup>

**Membrane-permeant biarsenical dyes such as FIAsh and ReAsH fluoresce upon binding to genetically encoded tetracysteine motifs expressed in living cells<sup>1,2</sup>, yet spontaneous nonspecific background staining can prevent detection of weakly expressed or dilute proteins<sup>2,3</sup>. If the affinity of the tetracysteine peptide could be increased, more stringent dithiol washes should increase the contrast between specific and nonspecific staining. Residues surrounding the tetracysteine motif were randomized and fused to GFP, retrovirally transduced into mammalian cells and iteratively sorted by fluorescence-activated cell sorting for high FRET from GFP to ReAsH in the presence of increasing concentrations of dithiol competitors. The selected sequences show higher fluorescence quantum yields and markedly improved dithiol resistance, culminating in a >20-fold increase in contrast. The selected tetracysteine sequences, HRWCCPGCCKTF and FLNCCPGCCMEP, maintain their enhanced properties as fusions to either terminus of GFP or directly to  $\beta$ -actin. These improved biarsenical-tetracysteine motifs should enable detection of a much broader spectrum of cellular proteins.**

Biarsenical-tetracysteine labels are analogous to fluorescent protein fusions<sup>1</sup>, yet offer several unique capabilities such as correlative fluorescence and electron microscopy (EM)<sup>4</sup>, determination of protein age by multicolor fluorescence pulse-chase<sup>4,5</sup>, chromophore-assisted light inactivation (CALI) for spatiotemporal inactivation of proteins<sup>6,7</sup> and numerous other *in vitro* applications<sup>2</sup>. Additionally, the tetracysteine sequence consists of only a few amino acids, and is thus far smaller and potentially less perturbative than incorporation of an autofluorescent protein<sup>8–10</sup>. Furthermore, biarsenical-tetracysteines are detectable immediately following tetracysteine translation<sup>11,12</sup>, allowing visualization of early events in protein synthesis, in contrast to the intrinsic delays required for fluorescent protein maturation<sup>13</sup>. Several other protein fusion partners can also trap distinctive tags in or on living cells<sup>14</sup>; however, these proteins either are an order of magnitude larger than tetracysteine motifs, require secondary processing enzymes, lack a general ability to label intracellular targets, or have no demonstrated expanded functionality.

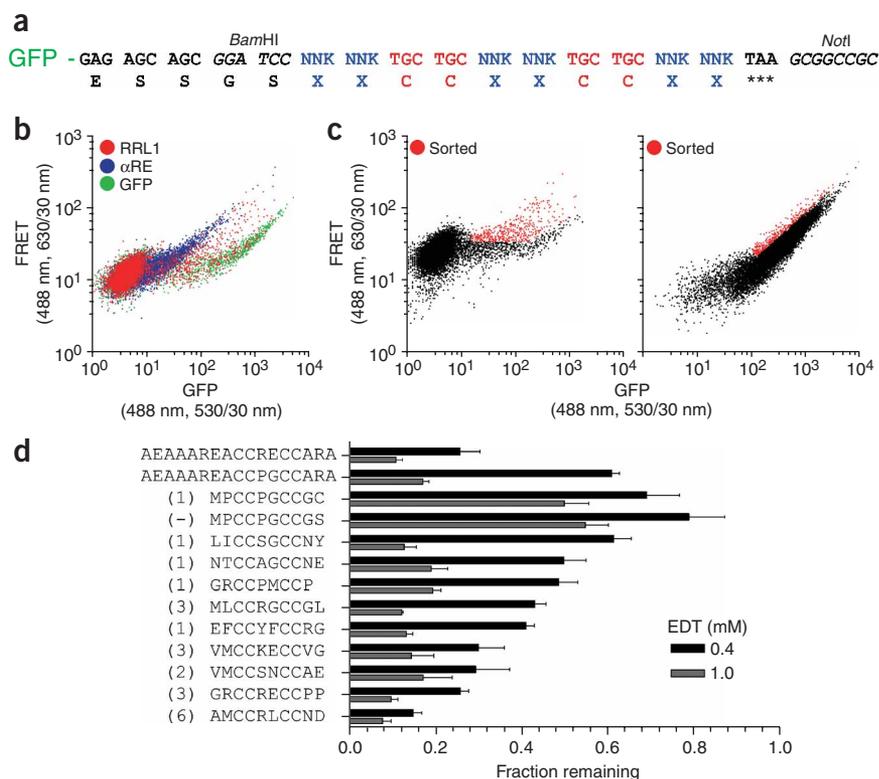
The earliest designs of tetracysteine sequences were intended to encourage  $\alpha$ -helicity under the assumption that the biarsenical dye would ideally fit into the *i*, *i*+1, *i*+4, and *i*+5 positions of an  $\alpha$ -helix<sup>1</sup>. With these sequences, nonspecific biarsenical background staining was estimated to equal the fluorescence of labeled protein of several micromolar<sup>2,3</sup>. Partial reduction of the background fluorescence was achieved by increasing the concentration of the dithiols 1,2-ethanedithiol (EDT) or 2,3-dimercaptopropanol (BAL) in washes to remove thiol-dependent background or by including nonfluorescent dyes to block hydrophobic binding sites<sup>15</sup>. When the helix-breaking amino acids proline and glycine were inserted between the dicysteine (CC) pairs to create a hairpin, the resulting tetracysteine substantially enhanced the affinity and contrast of FIAsh-labeled tetracysteine fusion proteins in cells, increasing the tolerable concentration of dithiol competitors without detrimental loss of specific fluorescence<sup>2</sup>. However, only a few pairs of amino acids were tested between the cysteines, and the surrounding residues were left unaltered, maintaining the  $\alpha$ -helical bias.

To optimize the tetracysteine sequence for improved ReAsH affinity and fluorescence, we have developed a retrovirally transduced mammalian cell-based library approach for fluorescence selection of optimal residues surrounding the tetracysteine motif by fluorescence-activated cell sorting (FACS). Other complementary approaches, such as surface display on phage or bacteria<sup>16</sup>, screen libraries for high-affinity binders *in vitro*, thus disregarding maintenance of desirable fluorescence properties *in vivo*. By carrying out these selections in the reducing environment of mammalian cytosol, we intended to avoid disulfide formation and evolve peptides with improved specificity, activity, toxicity and expression in the environment most relevant to their final application.

We created our first library, ReAsH Retroviral Library 1 (RRL1), by ligating a semi-randomized oligonucleotide cassette to the C terminus of green fluorescent protein (GFP) in a retroviral cloning vector (Fig. 1a). NIH3T3 cells were infected with the recombinant viral library at a low multiplicity of infection (MOI), stained with ReAsH and analyzed by flow cytometry. Measurement of the GFP quench and GFP-sensitized FRET (fluorescence resonance energy transfer) to ReAsH emission allows for determination of the kinetics and extent

<sup>1</sup>Department of Pharmacology, <sup>2</sup>Biomedical Sciences Graduate Program, <sup>3</sup>Department of Neurosciences, <sup>4</sup>National Center for Microscopy and Imaging Research, <sup>5</sup>Department of Chemistry and Biochemistry, and <sup>6</sup>Howard Hughes Medical Institute, University of California, San Diego, 9500 Gilman Dr., La Jolla, California 92093-0647, USA. Correspondence should be addressed to R.Y.T. (rtsien@ucsd.edu).

Received 20 April; accepted 18 August; published online 11 September 2005; doi:10.1038/nbt1136



**Figure 1** RRL1 selection for improved tetracysteine sequences. **(a)** Schematic of RRL1 cloning strategy. Unique restriction sites (italic), randomized codons (blue) and cysteine codons (red) are indicated. **(b)** FACS analysis of ReAsH binding by FRET. NIH3T3 cells virally transduced with either GFP-RRL1 (red), GFP- $\alpha$ RE (blue) or GFP alone (green) after ReAsH staining and a 30-min 0.1 mM BAL wash. ReAsH binding is characterized by an increase in FRET (630/22 nm) and a decrease in GFP (530/30 nm) fluorescence, which appear on a log scale as a shifts upwards and leftwards. **(c)** RRL1 FACS selections. Cells collected (red) in sort 1 (left) and sort 4 (right). **(d)** Sequence results and analysis for dithiol resistance. Unique sequences isolated in the RRL1 selection are listed, with the number of identical clones isolated in parenthesis. (–) indicates an additional peptide deliberately generated rather than isolated from the selection. The dithiol resistance of ReAsH fluorescence is shown for each sequence determined from live cell imaging experiments. Measurements represent the average of five or more cells after acute treatment with 0.4 mM and 1.0 mM EDT to ReAsH-labeled cells. Background-subtracted total ReAsH fluorescence before washing is normalized to 1, representing the cysteine in the final randomized position of MPCCPGCCGC to a serine had no effect on dithiol resistance, proving the last cysteine was a fortuitous nonparticipant in arsenical binding.

of ReAsH labeling in a single cell<sup>2</sup>. Specific ReAsH binding was detectable in cells expressing GFP fused to AEEAAARECCRCRARA<sup>1</sup> ( $\alpha$ RE), our first-generation tetracysteine sequence, and RRL1 cells, but not in cells expressing GFP alone (Fig. 1b). Notably, the RRL1 cells showed varying levels of FRET after washing with dithiols, suggesting that different amino acid combinations near the tetracysteine can modulate the dithiol resistance and/or fluorescence properties of the complex. FRET-positive RRL1 cells were collected and expanded (Fig. 1c, left). To discriminate higher-affinity peptides, we performed three additional rounds of sorting, each time increasing the selection pressure by escalating the dithiol concentration during washing. Finally, we sorted single cells (Fig. 1c, right) on a 96-well plate.

Sequence analysis of the isolated clones established ten new tetracysteine sequences (Fig. 1d). MPCCPGCCGC was highly resistant to EDT, maintaining 50% of its total ReAsH fluorescence in the face of 1.0 mM EDT, whereas  $\alpha$ RE and AEEAAARECCPGCCARA<sup>2</sup> ( $\alpha$ PG), our second-generation hairpin tetracysteine, both retained less than 20% (Fig. 1d). The next best four peptides all contained either internal prolines or internal glycines, corroborating the superiority of hairpin turns over helical conformations.

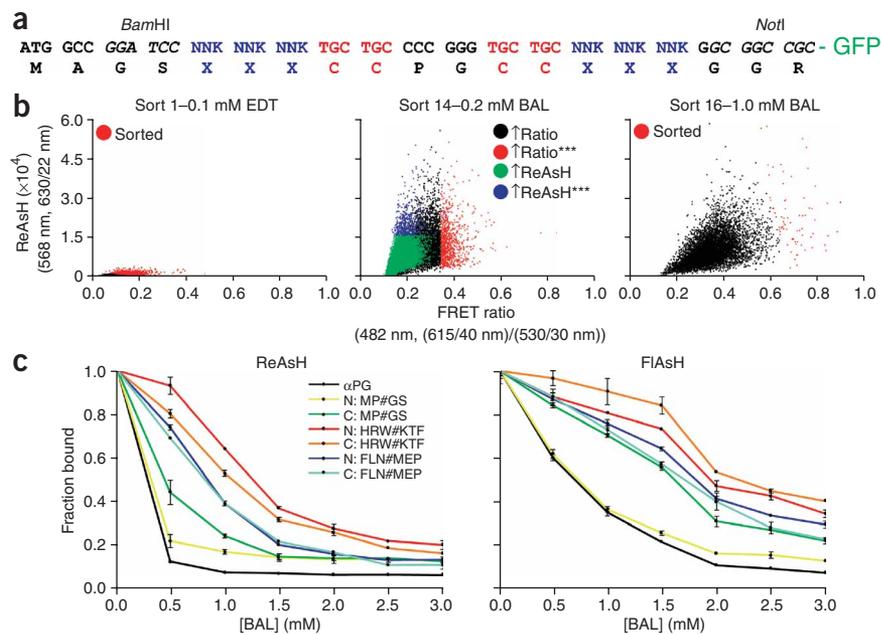
The results from the RRL1 selection confirmed the consensus sequence CCPGCC and verified the usefulness of the mammalian cell-based library approach for optimization of the tetracysteine motif. In an effort to further optimize the ReAsH binding tetracysteine peptide, we devised a new library, RRL2, fixing PG as the internal residues while randomizing the three external residues on either side of the tetracysteine, XXXCCPGCCXXX; we define this hereafter by the abbreviation XXX#XXX (# = CCPGCC) (Fig. 2a).

Three hundred million HEK293 cells were infected with RRL2 virus and presorted for GFP expression, isolating 30 million cells. A 568-nm laser was added to directly excite ReAsH, allowing cells to be sorted

based on two criteria: FRET ratio (GFP-sensitized ReAsH emission divided by GFP emission) and directly excited ReAsH emission. The GFP-positive cells were stained with ReAsH and sorted for multiple rounds, each round selecting the best 10–15% of the total population (Fig. 2b). Such moderate-stringency sorts eliminate unfavorable cells over the course of several selections, each time amplifying the pool of selected cells in culture for better sampling of each improved genotype. After ten rounds, the population fell into two categories: one showing high ReAsH fluorescence and a low FRET ratio, and the other with a high FRET ratio but lower ReAsH fluorescence. Each phenotype was simultaneously separated by sorting with two streams, one pool for each phenotype. After four more rounds of sorting, cells were washed with a low concentration of dithiol and sorted into 96-well plates to determine the composition of each population (Fig. 2b, middle).

After sequencing and analysis, the clones with high ReAsH intensity and low FRET ratio showed massive overexpression of the tetracysteine-GFP fusion, but weak dithiol resistance and poor labeling efficiency (data not shown). By excluding the GFP excitation data during the selections, protein expression levels were left uncorrected, leading to selection for overexpression rather than high affinity. In contrast, the population with the high FRET ratio and lower ReAsH fluorescence was dominated by sequences with dithiol resistance equal to or better than that of MP#GS (see Supplementary Fig. 1 online) while still expressing high levels of the fusion protein.

After two more rounds of selection with higher stringency, single cells were sorted onto 96-well plates for clonal expansion (Fig. 2b, right). All 22 clones isolated converged on three sequences, HRW#KTF, FLN#MEP and YRE#MWR. Each peptide was tested as both N- and C-terminal fusions to GFP (Emerald) and CFP (Cerulean<sup>17</sup>), and exhibited substantially improved dithiol resistance



**Figure 2** RRL2 selection and analysis of the optimized flanking residues. **(a)** Schematic of RRL2 cloning strategy. Notation as in **Figure 1a**. **(b)** RRL2 sort history. The FRET ratio is plotted versus ReAsH intensity in individual cells after a given BAL wash. In sort 1 (left), cells with high FRET ratios and ReAsH intensities were collected (red). Sort 14 (middle) shows the separation of cells with high FRET ratio ( $\uparrow$ Ratio, red/black) from those with high ReAsH intensity ( $\uparrow$ ReAsH, blue/green) and the corresponding sorted fraction (\*\*\*). In sort 16 (right), the final clones were selected from the top few percent in the high-FRET-ratio population. **(c)** Dithiol resistance of final optimized tetracysteines. BAL titration of ReAsH- (left) and FIAsh-labeled (right) N- and C-terminal optimized tetracysteine fusions to GFP and Cerulean, respectively. Tetracysteine color notation is the same in both ReAsH and FIAsh titrations. Dithiol resistance is shown as the average fraction of the FRET ratio remaining after incremental washes with high concentrations of BAL, shown with corresponding standard deviations derived from three or more duplicate wells on a 96-well plate.

for both ReAsH and FIAsh (**Fig. 2c**), while showing little preference for either biarsenical and confirming the weaker dithiol resistance of ReAsH as compared to FIAsh<sup>2</sup>. Upon ReAsH labeling, cells expressing YRE#MWR fused to the N-terminus of GFP, but neither of the other two sequences, quickly formed tiny, subcellular, highly red-fluorescent aggregates capable of evading even the highest-dithiol washes. This sequence was therefore excluded from further analysis, though the ability to precipitate a protein in living cells merely through addition of a permeant fluorogenic small molecule may be useful in other contexts.

In certain cell types, a small fraction of the tetracysteine-GFP protein becomes palmitoylated, blocking biarsenical binding and targeting the fusion protein to the plasma membrane. Fusion of a single Flag, hemagglutinin (HA) or Myc epitope upstream of the tetracysteine or addition of the palmitoylation inhibitor 2-bromopalmitate<sup>18</sup> blocked artifactual membrane localization without affecting dithiol resistance (see **Supplementary Fig. 2** online). In several cases, including those of  $\beta$ -actin and  $\alpha$ -tubulin, genetic fusion to a cellular protein prevents membrane localization and ReAsH binding to the tetracysteine is unobstructed.

Increasing the FRET ratio of acceptor to donor emissions can be accomplished by altering the orientation of the GFP and ReAsH chromophores as well as by increasing the fluorescence quantum yield of ReAsH directly. To address this, the mammalian-expressed tetracysteine-GFP proteins were affinity purified to homogeneity from clonal lysates using FIAsh-agarose beads<sup>2</sup>, then labeled with ReAsH *in vitro*. The selected peptides increased the quantum yield of ReAsH from 0.28 to 0.47 (see **Supplementary Table 1** online). Both optimized sequences retain much of the improved fluorescence properties when transferred to the C terminus of GFP. Replacement of the GFP with CFP permitted similar analysis of FIAsh, which likewise showed marked improvement in fluorescence quantum yield associated with the optimized tetracysteines.

Next, we tested whether the improved dithiol resistance and higher quantum yields could improve ReAsH contrast in cells. Populations of ReAsH-labeled HEK293T cells expressing distinct tetracysteine-GFP fusions were analyzed by flow cytometry using dual laser excitation to monitor both GFP and ReAsH fluorescence. FLN#MEP cells showed

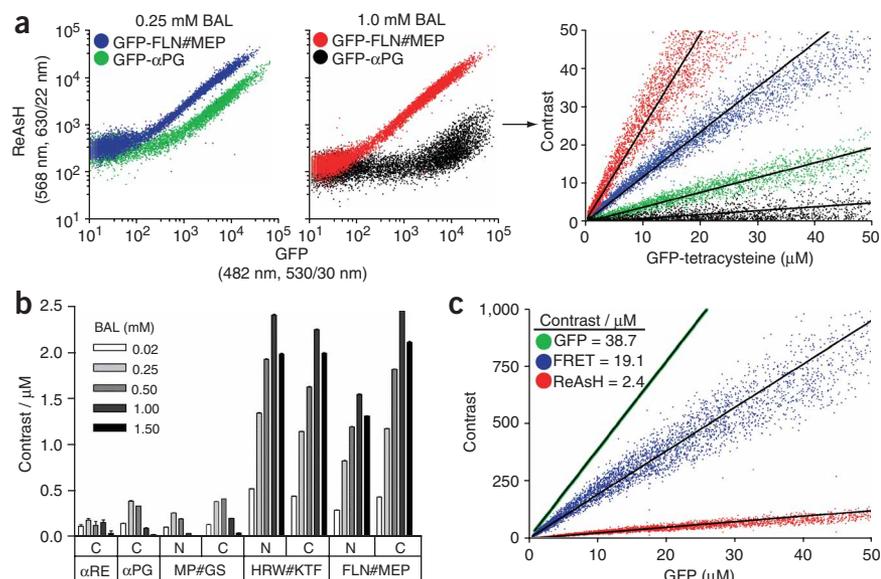
improved dithiol resistance and enhanced ReAsH fluorescence when compared to  $\alpha$ PG cells (**Fig. 3a**, left). As expected, the contrast is linearly proportional to the amount of tetracysteine-GFP in the cell. The slope of the linear regression of contrast versus the tetracysteine-GFP concentration normalizes for expression, allowing direct comparison of the effectiveness of different staining conditions, dithiol washes or tetracysteine sequences (**Fig. 3a**, right).

The improved dithiol resistance and quantum yields of HRW#KTF and FLN#MEP considerably increase the contrast of the tetracysteine-biarsenical complex in living cells (**Fig. 3b**). These effects are most obvious after washing with high concentrations of dithiol, which minimize the nonspecific background staining without affecting the desired specific fluorescence. The optimized tetracysteine peptides increase ReAsH contrast  $\sim$ 20-fold and  $\sim$ 6-fold over the  $\alpha$ RE and  $\alpha$ PG sequences, respectively (**Fig. 3b**).

Despite the improvements described above, the absolute contrast of ReAsH under optimal conditions is still approximately 16-fold lower than that of GFP (**Fig. 3c**). Fusion of a tetracysteine-tagged CFP (labeled with FIAsh or ReAsH) or GFP (labeled with ReAsH) to a cellular protein allows biarsenical fluorescence to be monitored by FRET from the donor fluorescent protein to the nearby, genetically fused biarsenical-tetracysteine acceptor. By excluding nonspecific biarsenicals from being excited, FRET-mediated detection increases ReAsH contrast almost eightfold, making it only twofold less than that for GFP (**Fig. 3c**). Fusion of a tetracysteine barely increases the size of the fluorescent protein, but potentially provides the extended functionalities of the biarsenical-tetracysteine system, such as fluorescence pulse-chases<sup>4,5</sup>, CALI<sup>6,7</sup> and EM photoconversion<sup>4</sup>.

To test whether the optimized tetracysteines retain greater contrast when fused to biologically relevant proteins, not just GFP, we transduced primary human foreskin fibroblasts for stable expression of MP#GS, FLN#MEP or HRW#KTF as N-terminal tetracysteine-GFP fusions to  $\beta$ -actin. After ReAsH labeling and washing with a high concentration of dithiol, actin stress fibers were easily identified by GFP fluorescence, but ReAsH fluorescence was only visible with FLN#MEP and HRW#KTF, not with MP#GS (**Fig. 4a**). Next, the optimized sequences were directly fused to  $\beta$ -actin without

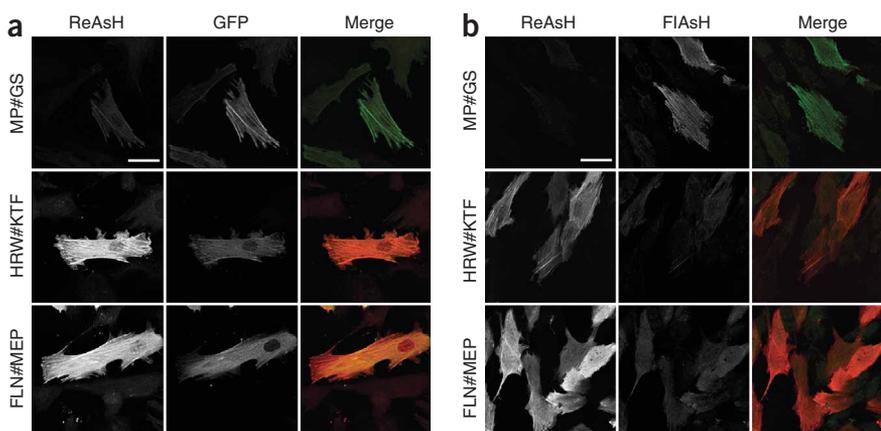
**Figure 3** Contrast improvement quantified by flow cytometry. **(a)** Comparison of ReAsH and GFP fluorescence in HEK293T cells expressing  $\alpha$ PG (green/black) and FLN#MEP (blue/red) after dithiol washes. Overlay of populations comparing ReAsH and GFP values after 0.25 mM BAL (left) or 1.0 mM BAL washes (middle). ReAsH contrast in a single tetracycline-GFP expressing cell is defined as specific ReAsH fluorescence divided by the mean nonspecific ReAsH fluorescence determined from stained, nontransduced cells. Background ReAsH fluorescence drops after the high-BAL wash. The calculated contrast, comparing the GFP concentration to ReAsH contrast in single cells, is shown with corresponding color notation (right). Linear regression lines are shown for each population, demonstrating the steeper slope of FLN#MEP (red/blue) versus  $\alpha$ PG (green/black). **(b)** Concentration-independent analysis of tetracycline contrast. Bars represent the slope of the linear regression and the corresponding standard error. N and C terminal fusions of the optimized tetracycline sequences were compared to the  $\alpha$ -helical sequences at five different concentrations of BAL. Moving the tetracycline to the C terminus of GFP had little effect on the overall contrast of HRW#KTF and modestly improved the contrast of FLN#MEP. **(c)** Contrast of GFP (green), FRET from GFP to ReAsH (blue) and ReAsH (red). The slope of each linear regression is noted in the legend, corresponding to the contrast achieved per micromolar of recombinant protein. GFP contrast is calculated from GFP-transduced HEK293T cells relative to nontransduced cells. FRET and ReAsH contrast is calculated from GFP-FLN#MEP-transduced cells labeled with ReAsH and washed with 1.0 mM BAL.



an intervening GFP. Tetracycline expressing cells were first labeled with ReAsH, then washed with a high concentration of dithiol. Subsequently, cells were labeled with FIAsh to fill all tetracycline vacancies caused by the high-dithiol wash and analyzed for both FIAsh and ReAsH fluorescence. FLN#MEP and HRW#KTF were capable of resisting high dithiol washes and retaining ReAsH fluorescence, whereas ReAsH labeling of MP#GS was completely replaced by FIAsh (**Fig. 4b**). Clearly, addition of FLN#MEP or HRW#KTF as the flanking residues of CCPGCC radically increases the dithiol resistance of the tetracycline-biarsenical complex, providing better contrast in combination with high concentration dithiol washes, independent of fusion to GFP.

We currently do not have a complete molecular explanation of the higher affinity of the final improved tetracyclines, HRW#KTF and FLN#MEP. Alanine scanning suggests that large aromatic residues are useful (see **Supplementary Fig. 3** online), possibly shielding the biarsenical-tetracycline complex from competing dithiols. Charge provides no simple explanation, because HRW#KTF has a charge of +2 to +3, whereas FLN#MEP has a -1 charge, and the key residues contributing to the high dithiol resistance are neutral (see **Supplementary Fig. 3**). Further understanding of the affinity and photophysical properties of the complexes will likely require high-resolution structures. In the meantime, we prefer FLN#MEP over HRW#KTF, mainly because the former gives somewhat higher quantum yields.

Most attempts to use FACS to screen libraries have been carried out in bacteria<sup>19–21</sup>, yet screening in mammalian cells is the best guarantee that the resulting optimizations will function properly in mammalian cells. Additionally, many genetically encoded reporters detect mammalian-specific biochemistry, making bacterial optimization infeasible. The ability to screen large mammalian cell libraries will be a crucial tool in the further improvement of genetically encoded fluorescent reporters.



**Figure 4** Fusion of optimized tetracyclines to  $\beta$ -actin. **(a)** Tetracycline-GFP- $\beta$ -actin fusions. Confocal images comparing ReAsH labeled tetracycline-GFP-actin fusions in human primary fibroblasts after a wash with 0.75 mM BAL. Substantial GFP fluorescence is expected even in the absence of dithiol washing, because the FRET efficiency is less than 100%. **(b)** Tetracycline- $\beta$ -actin fusions. Cells were labeled with ReAsH, washed with 0.75 mM BAL, and relabeled with FIAsh to bind free sites, then weakly washed with 0.1 mM BAL to reduce FIAsh background. We concluded that the length of the experiment was too short to detect substantial new protein expression, as determined by ReAsH/FIAsh pulse-chase time course experiments or addition of cycloheximide (data not shown). Scale bar, 40  $\mu$ m.

## METHODS

**RRL1 production.** Oligonucleotide primers sequences are listed in **Supplementary Table 2** online. Phosphorylated, degenerate oligonucleotides (Primer1, Primer2) were annealed at a final concentration of 50  $\mu\text{M}$  in T4 DNA ligase buffer (NEB). The retrovirus vector pCLNCX (Imgenex) was modified to destroy the *Bam*HI site upstream of the second CMV promoter. EGFP was amplified by PCR (Primer3, Primer4) to introduce cloning sites and a short linker, then ligated into the modified pCLNCX. Vector DNA was digested with *Bam*HI and *Not*I (NEB) twice, then ligated with an optimized concentration of the annealed library overnight. The ligation reaction was purified and concentrated using a QIA-Quick column (Qiagen), then electroporated into TG1 cells (Stratagene). Based on serial dilutions of the electroporated cells, the library contained  $2.6 \times 10^7$  members, or 2.4% of the total nucleotide diversity. After overnight growth in 250 ml LB with ampicillin, the plasmid library was purified and cotransfected, using calcium phosphate, with pCL-Eco (Imgenex) into 80% confluent HEK293 cells in a 10-cm dish. The medium was replaced after 24 h, and 48 h later 8 ml of virus was harvested, filtered and frozen in liquid nitrogen. NIH3T3 cells ( $6.5 \times 10^6$ ) were infected with 1 ml of  $1.2 \times 10^6$  green fluorescence units (GFU) per ml virus in the presence of 8  $\mu\text{g/ml}$  polybrene (Sigma).

**RRL2 production.** Degenerate oligonucleotides (Primer5, Primer6), each containing 18 bp of complementarity over the region encoding for the CCPGCC motif, were annealed at a final concentration of 50  $\mu\text{M}$ . Overhangs were filled in with Klenow fragment ( $3' \rightarrow 5'$  exonuclease-negative) (NEB). After heat inactivation, DNA was extracted with phenol/chloroform and precipitated with ethanol. The product was digested with *Bam*HI and *Not*I, extracted with phenol/chloroform and concentrated in a Microcon YM-10 column (Millipore), then separated by nondenaturing PAGE. The digested fragment was excised and electroeluted (Bio-Rad Model 422 Electro-Eluter, 12–15-kDa cut-off), extracted with phenol/chloroform and concentrated. In RRL2, NIH3T3 cells were replaced with the smaller, more robust and more highly expressing HEK293 cells and the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE)<sup>22</sup> was added to the  $3'$  untranslated region of the enhanced folding mutant Emerald GFP<sup>23</sup>. These improvements increased the expression of the tetracycline-GFP library, amplifying the mean signal from a single cell by >10-fold. The WPRE element was amplified by PCR from BluescriptII SK+ WPRE (Stratagene; Primer7, Primer8) and cloned between the *Xho*I and *Cl*aI sites of pCLNCX. Emerald GFP was amplified by PCR (Primer9, Primer10) to introduce restriction sites into the N terminus of GFP, digested with *Hind*III and *Xho*I, and ligated into the pCLNCX-WPRE vector. Next, the GFP vector was digested with *Bam*HI and *Not*I, extracted with phenol/chloroform and precipitated with ethanol. An optimized concentration of cassette was ligated with 15  $\mu\text{g}$  of vector, extracted with phenol/chloroform and precipitated with ethanol in the presence of 20  $\mu\text{g}$  yeast tRNA (Ambion). The ligation product was electroporated into Electro-Ten Blue cells (Stratagene) and grown overnight in 1 liter LB with ampicillin. The RRL2 plasmid library was calculated to contain  $8.6 \times 10^8$  members, or 80% of the total nucleotide diversity. Purified plasmid was cotransfected with an equal concentration of pCL-Ampho (Imgenex) into four 10-cm plates of 80% confluent HEK293 cells with Fugene6 (Roche Diagnostics). After virus production and storage at  $-80^\circ\text{C}$ , thawed virus titered by flow cytometry contained  $2.5 \times 10^6$  GFU/ml in HEK293 cells. Three hundred million HEK293 cells were infected at an MOI of 0.14.

**ReAsH staining and flow cytometry.** Cells were cultured in DMEM supplemented with 10% FBS, 100 units/ml penicillin G, 100  $\mu\text{g/ml}$  streptomycin and 2.5  $\mu\text{g/ml}$  amphotericin B. Confluent cells were stained with 0.25  $\mu\text{M}$  or 0.5  $\mu\text{M}$  ReAsH and 10  $\mu\text{M}$  EDT in HBSS (Hanks balanced saline solution supplemented with 2 g/l glucose and 20 mM HEPES) for 1 h, rinsed once with HBSS, then incubated with specific concentrations of dithiol (EDT or BAL) in HBSS for 30 min at room temperature. Cells were then trypsinized, pelleted and resuspended in HBSS and then sorted into 15-ml tubes containing 5 ml medium with 30% FBS. Sorted cells were pelleted and cultured in medium containing 0.5 mg/ml G418. Single cells were collected in 96-well plates with 0.3 ml medium per well supplemented with 10% Optimem and an additional 10% FBS. RRL1 was sorted on a MoFlow flow cytometer (Cytomation) using a single 488-nm laser and 530/30-nm and 630/22-nm emissions. RRL2 was performed on a Becton Dickinson FACSVantage DiVa with two lasers, 482 nm

(530/30-nm and 615/40-nm emission) and 568 nm (630/22-nm emission). Laser alignment changes on successive rounds of selection were normalized using readings of alignment beads.

**Sequence retrieval and mutagenesis.** Total RNA was isolated using Trizol Reagent (Invitrogen) and reverse transcribed with ImPromII reverse transcriptase (Promega) using a gene-specific primer in the  $3'$  UTR of pCLNCX (Primer11), then amplified by PCR (Primer11, Primer12). The PCR product was gel purified and sequenced with either a  $5'$  or a  $3'$  pCLNCX primer. All unique PCR products from RRL1 were digested with *Hind*III and *Not*I and subcloned into pCDNA3 (Invitrogen). MP#GC was mutated to MP#GS by PCR (Primer7, Primer13). The RT-PCR products of several unique clones from RRL2 were digested with *Hind*III and *Xho*I for ligation into the pCLNCX-WPRE vector. N-terminal YRE#MWR-GFP was subcloned onto the C terminus of the GFP from RRL1 using *Bam*HI and *Xho*I sites. Next, the tetracycline-GFP fusion was amplified by PCR to introduce a stop codon before the *Not*I site (Primer12, Primer14), digested and ligated into pCLNCX-WPRE. Other tetracycline clones were then substituted by subcloning. C-terminal MP#GS was amplified by PCR to remove the stop codon (Primer7, Primer15), then ligated to the N terminus of RRL2 GFP with *Hind*III/*Not*I sites, and finally subcloned to the RRL2 vector using *Bam*HI/*Xho*I sites. Alanine substitutions were generated by a modified QuikChange (Stratagene) protocol using HRW#KTF and FLN#MEP as N-terminal fusions to Emerald GFP, then subcloned into pCLNCX-WPRE for virus production.

**Dithiol titrations.** RRL1 unique clones were analyzed for dithiol resistance in transiently transfected HeLa cells by fluorescence microscopy, using the following filters: GFP (excitation 480/30 nm, emission 535/25 nm), FRET (excitation 480/30 nm, emission 635/55 nm), ReAsH (excitation 540/25 nm, emission 635/55 nm), dichroic 505LP. Maximal ReAsH fluorescence was reached in approximately 30–45 min, as determined by a lack of further GFP fluorescence quenching. EDT or BAL were dissolved in DMSO and premixed with HBSS, then added in increasing concentrations to cells at intervals of 15–20 min. The average ReAsH fluorescence in nontransfected cells was subtracted from tetracycline-GFP-expressing cells. RRL2 clonal cell lines were plated in multiple wells on 96-well plates with black walls and clear bottoms and grown to confluency. Fluorescence measurements were taken on a Tecan Sapphire fluorescence plate reader at each fluorophore's specific wavelengths. Cells were measured before staining to obtain an initial baseline fluorescence, then stained with ReAsH for 1 h. The staining solution was replaced with HBSS containing 10  $\mu\text{M}$  EDT to measure the maximal fluorescence. Next, various concentrations of dithiol in HBSS were added to predetermined wells and incubated for 30 min before the final measurement. Background fluorescence was determined using nontransduced HEK293 cells and subtracted from each individual wavelength.

**Protein purification and fluorescence measurements.** A confluent monolayer of clonal tetracycline tagged GFP or CFP HEK293 cells was lysed in RIPA buffer (1% Nonidet P-40, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 2 mM EDTA, 0.1 M sodium phosphate, pH 7.2) supplemented with 10 mM sodium 2-mercaptoethanesulfonate (MES, Aldrich), 1 mM tris-(2-carboxyethyl)phosphine HCl (TCEP, Molecular Probes) and Complete protease inhibitor (Roche Diagnostics). All subsequent solutions contained 10 mM MES and 1 mM TCEP. Cell debris was pelleted and the soluble cell lysate was filtered through a 0.45- $\mu\text{m}$  syringe filter, then diluted 1:2 in PBS. Prewashed FLaSH beads<sup>2</sup> were added to the lysate and incubated at  $4^\circ\text{C}$  overnight, then pelleted and washed three times with PBS, then once in PBS supplemented with 0.1 mM EDT or BAL. Protein was eluted with 0.25 M DTT (Invitrogen), then buffer-exchanged four times in PBS by centrifugation in a 30 kDa-cutoff Microcon filter. Next, protein was labeled with 3–5-fold excess biarsenical overnight at  $4^\circ\text{C}$ , then buffer exchanged four times with 0.1 mM BAL in PBS to remove free biarsenical. No other bands were detectable by SDS-PAGE, as determined by either direct ReAsH fluorescence or Coomassie blue staining. Mass spectrometry analysis of purified, trypsin-digested protein revealed a loss of expected mass observed in N-terminal-most tetracycline-containing peptide, corresponding to N-terminal methionine cleavage and N-terminal acetylation. FLaSH and ReAsH quantum yields were determined using

dichlorofluorescein in 0.1 N NaOH ( $\Phi = 0.92$ ) or rhodamine 101 in ethanol ( $\Phi = 1.0$ , then corrected for refractive index), respectively, as standards.

**Determining ReAsH contrast using flow cytometry.** HEK293T cells (ATCC) were transduced with retrovirus encoding a tetracycline-GFP fusion, stained 2 d later with 0.5  $\mu\text{M}$  ReAsH and 10  $\mu\text{M}$  EDT for 1 h, then washed once with Dulbecco's PBS containing 1 mM EDTA and trypsinized. Suspended cells were divided into aliquots in various concentrations of BAL, incubated for 30 min, then pelleted and resuspended in HBSS supplemented with 20  $\mu\text{M}$  BAL to prevent further ReAsH binding, and analyzed on a Becton Dickinson FACS-Vantage DiVa flow cytometer. One aliquot was washed with 10 mM BAL to remove all ReAsH so as to measure the maximal GFP fluorescence of the population, and hence the average GFP quench at the tested BAL concentrations. Lower-staining concentrations of ReAsH slightly enhanced contrast, yet resulted in a lower fraction of labeled tetracycline, and subsequently less signal. Inclusion of Disperse Blue 3 (ref. 15) or long overnight labeling procedures did not enhance ReAsH contrast. GFP contrast was measured in the absence of ReAsH labeling to avoid background contributions from nonspecific background ReAsH labeling. Quantum FITC MESF High Level beads (Bangs Laboratories) were analyzed and peak averages were plotted against the fluorophore number to produce a standard curve. HEK293T cell volume was determined by measuring the average cell diameter of suspended cells by differential interference contrast (DIC) microscopy. Data were calculated using the following formulae (where  $\Phi$  = fluorescence quantum yield,  $\epsilon$  = extinction coefficient,  $A$  = absorbance, laser wavelengths = 482 nm (for GFP and FRET) and 568 nm (for ReAsH), peak  $A_{\text{Emerald GFP}} = 484$  nm, peak  $A_{\text{Fluorescein}} = 490$  nm,  $V$  = cell volume =  $0.9 \pm 0.2$  pl, and  $\text{GFP}_{\text{event}} = \text{GFP fluorescence signal from a single event}$ ,  $T$  = transduced = GFP positive cell, and  $\text{NT}$  = nontransduced HEK293T cell):

$$B = \text{Brightness} = (\Phi_{\text{GFP}} * \epsilon_{\text{GFP}} * A_{\text{GFP:482nm}} / A_{\text{GFP:484nm}}) / (\Phi_{\text{Fluorescein}} * \epsilon_{\text{Fluorescein}} * A_{\text{Fluorescein:482nm}} / A_{\text{Fluorescein:490nm}})$$

$$Q = \text{Quench} = (\text{Mean GFP fluorescence at given [BAL]}) / (\text{Mean GFP fluorescence at 10 mM BAL})$$

From standard curve,  $K$  = moles fluorescein/relative fluorescent units

$$[\text{GFP}] = \text{GFP}_{\text{event}} * K * B / (Q * V)$$

$$\text{ReAsH Contrast} = (\text{ReAsH}_T - \text{Mean ReAsH}_{\text{NT}}) / (\text{Mean ReAsH}_{\text{NT}})$$

$$\text{GFP Contrast} = (\text{GFP}_T - \text{Mean GFP}_{\text{NT}}) / (\text{Mean GFP}_{\text{NT}})$$

Contrast was plotted against GFP concentration and slopes were calculated from linear regression analysis using  $1/X^2$  weighting to minimize percentage errors.

**Production and analysis of  $\beta$ -actin fusions.** HEK293 total RNA was isolated and reverse transcribed using an oligo d(T) primer to generate cDNA. The human  $\beta$ -actin cDNA was amplified by PCR (Primer16, Primer17) using *Taq* polymerase (Roche), digested with *HindIII/XhoI* and ligated in the N-terminal FLN#MEP pCLNCX-WPRE vector. The resulting plasmid was used as a PCR template to generate an in-frame fusion of actin (Primer18, Primer17) and FLN#MEP GFP pCLNCX-WPRE (Primer12, Primer19). The vector and the two PCR products were digested (pCLNCX; *HindIII/XhoI*, 4cys-GFP; *HindIII/BglII*, actin; *BglIII/XhoI*) and ligated. The final actin or GFP-actin fusions were subcloned into N-terminal HRW#KTF and MP#GS pCLNCX-WPRE vectors. Human primary fibroblasts were transduced with tetracycline- $\beta$ -actin and tetracycline-GFP- $\beta$ -actin fusions and maintained without drug selection to preserve heterogeneous expression levels. Cells were grown and imaged on glass-bottom poly-D-lysine-coated dishes (MatTek). All labeling and imaging was performed at 37 °C. Cells were labeled with 0.5  $\mu\text{M}$  ReAsH and 10  $\mu\text{M}$  EDT in DMEM for 1 h, then rinsed and incubated with 0.75 mM BAL for 30 min. Tetracycline-actin fusions were further labeled with 0.5  $\mu\text{M}$  FIAsh and 10  $\mu\text{M}$  EDT for 1 h, then washed with 0.1 mM BAL. Next, the washing medium was rinsed away several times and replaced with medium lacking phenol red.

Imaging was performed on a Bio-Rad MRC 1024 confocal system and maintained at 37 °C using a Biotech lens heater and a 3-cm water circulation-based dish heater. Cells were imaged in a single large focal plane positioned at the cell base through a 63 $\times$  oil objective at 1,024  $\times$  1,024 resolution with the following settings: ReAsH: excitation 568 nm, emission 585 nm LP; FIAsh: excitation 488 nm, emission 540/30 nm; GFP: excitation 488 nm, emission 522/35 nm. After the high-concentration dithiol washes required for optimal biarsenical-tetracycline contrast, proliferation of primary human foreskin fibroblasts slows down for awhile, with the cells dividing only once in 72 h as compared to two divisions for untreated cells. Overall, the cells retain high viability and actin dynamics are unaffected and remain observable for several days after labeling.

ReAsH and FIAsh are commercially available through Invitrogen, marketed as Lumio-Red and Lumio-Green, respectively.

*Note: Supplementary information is available on the Nature Biotechnology website.*

#### ACKNOWLEDGMENTS

We thank the VA Research Flow Cytometry Core Facility at UCSD and W. Coyt Jackson for flow cytometry assistance, Larry Gross for mass spectrometry support, Paul Steinbach and Qing Xiong for laboratory assistance, Thomas Hope (University of Illinois at Chicago) for the Bluescript SK<sup>+</sup> WPRE plasmid, and members of the Tsien laboratory and FIAshers group for helpful discussions. Some of the work included here was conducted at the National Center for Microscopy and Imaging Research, which is supported by US National Institutes of Health grant RR04050 (to Mark H. Ellisman, University of California, San Diego). This work was supported by NIH NS27177, GM 72033, Department of Energy DE-FG03-01ER63276 and the Howard Hughes Medical Institute.

#### AUTHOR CONTRIBUTIONS

B.R.M. performed all experiments with the exception of those described in **Figure 4**, which were done in collaboration with B.N.G.G., and the characterization of synthetic peptide complexes, performed by S.R.A. Experiments were designed and analyzed with contributions from all authors. The manuscript and figures were prepared by B.R.M., B.N.G.G. and R.Y.T.

#### COMPETING INTERESTS STATEMENT

The authors declare competing financial interests; see the *Nature Biotechnology* website for details.

Published online at <http://www.nature.com/naturebiotechnology/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Griffin, B.A., Adams, S.R. & Tsien, R.Y. Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269–272 (1998).
- Adams, S.R. *et al.* New biarsenical ligands and tetracycline motifs for protein labeling *in vitro* and *in vivo*: synthesis and biological applications. *J. Am. Chem. Soc.* **124**, 6063–6076 (2002).
- Stroffekova, K., Proenza, C. & Beam, K.G. The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflugers Archiv. Eur. J. Physiol.* **442**, 859–866 (2001).
- Gaietta, G. *et al.* Multicolor and electron microscopic imaging of connexin trafficking. *Science* **296**, 503–507 (2002).
- Ju, W. *et al.* Activity-dependent regulation of dendritic synthesis and trafficking of AMPA receptors. *Nat. Neurosci.* **7**, 244–253 (2004).
- Marek, K.W. & Davis, G.W. Transgenically encoded protein photoinactivation (FIAsh-FALI): Acute inactivation of synaptotagmin I. *Neuron* **36**, 805–813 (2002).
- Tour, O., Meijer, R.M., Zacharias, D.A., Adams, S.R. & Tsien, R.Y. Genetically targeted chromophore-assisted light inactivation. *Nat. Biotechnol.* **21**, 1505–1508 (2003).
- Andresen, M., Schmitz-Salue, R. & Jakobs, S. Short tetracycline tags to  $\beta$ -tubulin demonstrate the significance of small labels for live cell imaging. *Mol. Biol. Cell* **15**, 5616–5622 (2004).
- Hoffmann, C. *et al.* A FIAsh-based FRET approach to determine G-protein coupled receptor activation in living cells. *Nat. Methods* **2**, 171–176 (2005).
- Panchal, R.G. *et al.* *In vivo* oligomerization and raft localization of Ebola virus protein VP40 during vesicular budding. *Proc. Natl. Acad. Sci. USA* **100**, 15936–15941 (2003).
- Rice, M.C., Bruner, M., Czymmek, K. & Kmiec, E.B. *In vitro* and *in vivo* nucleotide exchange directed by chimeric RNA/DNA oligonucleotides in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **40**, 857–868 (2001).
- Rice, M.C., Czymmek, K. & Kmiec, E.B. The potential of nucleic acid repair in functional genomics. *Nat. Biotechnol.* **19**, 321–326 (2001).
- Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).

14. Chen, I. & Ting, A.Y. Site-specific labeling of proteins with small molecules in live cells. *Curr. Opin. Biotechnol.* **16**, 35–40 (2005).
15. Griffin, B.A., Adams, S.R., Jones, J. & Tsien, R.Y. Fluorescent labeling of recombinant proteins in living cells with FIAsH. *Methods Enzymol.* **327**, 565–578 (2000).
16. Benhar, I. Biotechnological applications of phage and cell display. *Biotechnol. Adv.* **19**, 1–33 (2001).
17. Rizzo, M.A., Springer, G.H., Granada, B. & Piston, D.W. An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.* **22**, 445–449 (2004).
18. Webb, Y., Hermida-Matsumoto, L. & Resh, M.D. Inhibition of protein palmitoylation, raft localization, and T cell signaling by 2-bromopalmitate and polyunsaturated fatty acids. *J. Biol. Chem.* **275**, 261–270 (2000).
19. Cormack, B.P., Valdivia, R.H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38 (1996).
20. Daugherty, P.S., Iverson, B.L. & Georgiou, G. Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* **243**, 211–227 (2000).
21. Nguyen, A.W. & Daugherty, P.S. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360 (2005).
22. Zufferey, R., Donello, J.E., Trono, D. & Hope, T.J. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.* **73**, 2886–2892 (1999).
23. Cubitt, A.B., Woollenweber, L.A. & Heim, R. Understanding structure-function relationships in the *Aequorea victoria* green fluorescent protein. *Methods Cell Biol.* **58**, 19–30 (1999).