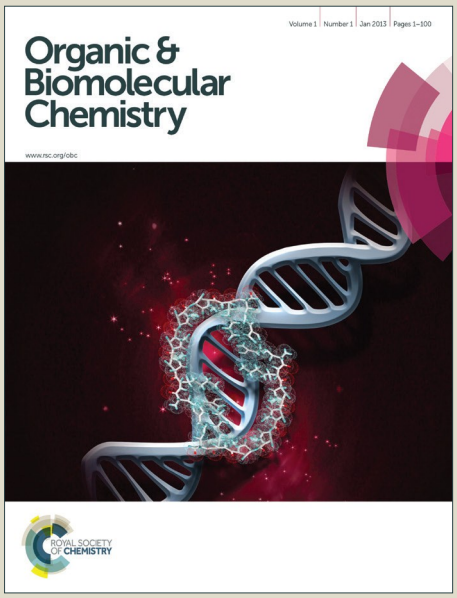
Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: B. Söveges, T. Imre, T. Szende, Á. L. Póti, G. B. Cserép, T. Hegedus, P. Kele and K. Németh, *Org. Biomol. Chem.*, 2016, DOI: 10.1039/C6OB00810K.



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

View Article Online
DOI: 10.1039/C6OB00810K

Journal Name

ARTICLE

Systematic study of protein labeling by fluorogenic probes using cysteine targeting vinyl sulfone-cyclooctyne tags

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

B. Söveges,[a] T. Imre,[b] T. Szende,[a] Á. L. Póti,[c] G. B. Cserép,[a] T. Hegedűs,[d] P. Kele[a] and K. Németh[a]*

Fluorescent tagging of proteins via accessible cysteine residues is of paramount importance. In this study, model proteins of interest (mitogen-activated protein kinases) were labeled successfully in native state on their free thiols by direct fluorescent derivatization, or in a sequential manner where conjugation of the site specific linker and the fluorophore is carried out in two steps. To this end we designed and prepared two novel chemical reporters carrying vinyl sulfone as Cys targeting function and cyclooctyne motifs, suitable for subsequent conjugation with fluorogenic azides via copper free strain-promoted azide-alkyne click chemistry. Direct and sequential labeling reaction steps were analyzed by native PAGE, capillary zone electrophoresis and tandem mass spectrometry. The efficiency of tagging was correlated with solvent accessibility of the Cys residues. Our results indicated that conjugation of native proteins by vinyl sulfone linkers was fast and thiol-selective. Subsequent click reaction with fluorogenic dyes generates intensive fluorescence signal and fulfills all requirements of bioorthogonality.

Introduction

Fluorescent labeling is undoubtedly one of the most popular techniques for monitoring biological structures and processes.[1, 2] Side chains of some naturally occurring proteinogenic amino acids or either termini of proteins offer suitable platforms for targeting.[3, 4] More specific, site selective fluorescent tagging can be carried out on rare amino acids like cysteine and tyrosine.[5-10] Such amino acid side-chains are efficiently modulated via their thiol or phenolic hydroxyl group using specific reagents.[3, 11] Generally, maleimides or iodoacetamides are applied for chemical ligation in Cys labeling schemes. In spite of the fact that vinyl sulfones show superior performance compared to maleimides or iodoacetamides vinyl sulfone based reagents are rarely applied.[12, 13] They show excellent reactivities towards thiols with nearly quantitative yields without the formation of by-products. Furthermore both the reagents and the products are stable and have good water-solubility. In the last decade bioorthogonal modulation schemes emerged as excellent means for fluorescent labeling of proteins.[14-17] These bioorthogonal reagents should suit the following main requirements: (1) the kinetic, thermodynamic, and metabolic stability of the reactants and products, (2) non-toxicity for living organisms, (3) high selectivity, specificity and efficiency under physiological conditions (ambient temperature and pressure, neutral pH, aqueous conditions) in other words either of the partners should exclusively react with each other without any interference with other surrounding functions present in live cells.

Lately, we demonstrated the development of Cys specific vinyl sulfone linkers bearing terminal alkyne as well as the efficiency of a sequentially driven two-step modification method of peptides using a fluorogenic dye in comparison with one step direct labeling schemes[10] (Fig. 1). These proof-of-concept experiments underlined the advantages of

[a] Research Centre for Natural Sciences of Hungarian Academy of Sciences, Institute of Organic Chemistry, Chemical Biology Research Group; H-1117 Budapest, Magyar tudósok krt. 2. Hungary. Tel.: +36 1 382 6659; E-mail address: nemeth.krisztina@ttk.mta.hu
[b] Research Centre for Natural Sciences of Hungarian Academy of Sciences, Institute of Organic Chemistry, MS Metabolomics Research Group; H-1117 Budapest, Magyar tudósok krt. 2.
[c] Research Centre for Natural Sciences of Hungarian Academy of Sciences, Institute of Enzymology, Protein Research Group; H-1117 Budapest, Magyar tudósok krt. 2. Hungary.
[d] MTA-SE Molecular Biophysics Research Group, Department of Biophysics and Radiation Biology, Semmelweis University; Tuzolto u. 37-47, H-1094 Budapest, Hungary.

Electronic Supplementary Information (ESI) available: Direct and sequential fluorescent labeling of other proteins of interest are demonstrated. Detailed syntheses routes and characterization data of molecules are available.
See DOI: 10.1039/x0xx00000x

sequentially driven modification schemes as these furnished only one single fluorescent species i.e. the labeled peptide providing the possibility for using no-wash conditions.

As a continuation of this work we aimed at further investigating the potential of such vinyl sulfone based linkers and extending the applicability of these chemical reporters to protein labeling schemes. However, proteins represent more challenging scaffolds as their rigid conformation compared to peptides, and their cysteine residues are often hardly accessible. In this work we intended to optimize the conditions of the derivatization scheme in terms of time-scale analysis of the reaction and the number of tagged cysteines. Furthermore, we made efforts to administer some novel linker molecules in order to eliminate the requirement for the toxic Cu(I) catalyst in the click reaction (copper-catalyzed azide-alkyne cycloaddition (CuAAC)) between the linker and dye. The use of copper as a catalyst is often compromised in CuAAC schemes due to the presence of chelating moieties in proteins (e.g. His-tagged)[18]. To this end, we synthesized novel vinyl sulfone linkers bearing cyclooctyne moieties, which enables the azide-alkyne reaction to proceed via strain-promoted azide-alkyne cycloaddition (SPAAC).[11] We have tested two cyclooctyne motifs, a slower reacting, easily accessible CyO,[19] (methyl 4-(cyclooct-2-yn-1-ylmethyl)benzoate) (2.4×10^{-3} $M^{-1}s^{-1}$)[17] and a synthetically more challenging, but fast reacting COMBO (2-methoxycarbonyl-5,6,9,10-tetrahydro-7,8-dehydro-benzocyclooctene) (0.8 $M^{-1}s^{-1}$) that was developed in our laboratory.[20, 21] Labeling schemes were compared using fluorescent vinyl sulfone reagents enabling direct labeling and sequentially applicable non-fluorescent linkers in combination with fluorogenic dyes.

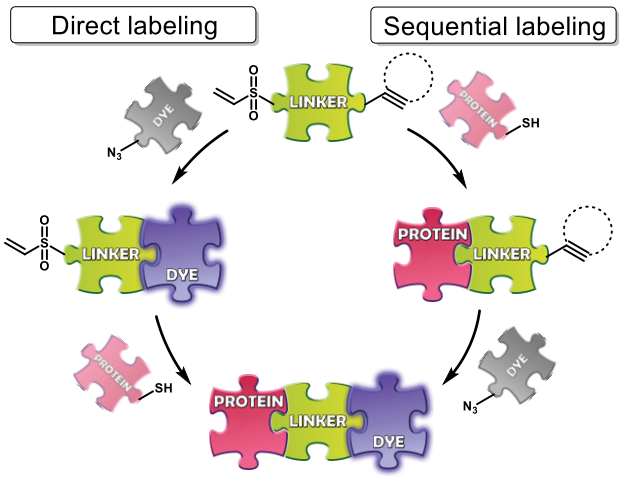


Fig. 1 Comparison of direct and sequential fluorescent modification schemes.

Results and discussion

View Article Online
DOI: 10.1039/C6OB00810K

Direct labeling of p38 with fluorescent linker-dye module

The recombinant intracellular protein, p38α/MAPK14 (called p38 in the followings) (44 kDa; pI 5.6), a member of mitogen-activated protein kinase family was selected as model protein, which is a key regulator element in signal transduction pathways.[22-24] More importantly, it contains four reduced thiol groups, two of which are surface-accessible (Fig. 2A).

When p38 was treated with a cysteine-specific fluorescent linker-dye conjugate[10] ($L_{TA}D_1$: Fig. 2C) an intensive fluorescent signal evolution was observed by capillary zone electrophoresis with laser induced fluorescent detection (CZE-LIF) that reached maximum value within 90 minutes (Fig. 2B and C). Labeling was found to be specific and covalent as $L_{TA}D_1$ with pre-alkylated p38 gave only a very small peak in the electropherogram compared to untreated p38 (Fig. 2B). The disadvantage of this labeling scheme is that the unreacted fluorescent reagent contributes to the overall signal depending on the efficiency of the separation from the excess of $L_{TA}D_1$.

A systematic tandem mass-spectrometric survey was carried out involving trypsin digested plain p38 and $L_{TA}D_1$ treated p38 with or without iodoacetamide (IAM) pretreatment. IDA (Information Dependent Analysis) LC-MS experiments proved that the labeling was specific to thiol residues. Moreover it was also revealed that the Cys 119 residue was the primary subject of labeling and Cys 162 being the secondary target of $L_{TA}D_1$ as the amount of these free thiol containing peptide fragments decreased by 97 ± 2 % and 60 ± 15 %, respectively (n = 4) in LC-MS/MRM (Multiple Reaction Monitoring) measurements (Fig. 3). Accordingly, the average number of bound $L_{TA}D_1$ per p38 protein was 1.57 based on the MS results being in good agreement with labeling efficiency (1.46 ± 0.18 (n = 5)) estimated from the ratio of absorbancies at 420 nm and 280 nm, monitored by UV-Vis spectrophotometry (ESI).

Apart from these surface accessible thiol groups, derivatization of the remaining Cys residues together with other amino acids were insignificant. These findings are in good agreement with the X-ray structure of p38 (PDB ID: 1A9U),[22] which suggest similar order of the thiols (accessible surface area (ASA): 91 $Å^2$, 23 $Å^2$, 2 $Å^2$ and 1 $Å^2$ for Cys 119, 162, 211 and 39, respectively).

Furthermore, in case of a mutant p38 ($p38^{C162S}$) where one of the cysteines was replaced by serine[25] the intensity of labeling signal was comparable (ASA of Cys 119: 103 $Å^2$) to that of the wild type protein as established by CE-LIF and native polyacrylamide gel electrophoresis (CN-PAGE) (Fig. 2B; Fig. 5).

In another instance, similarly intensive and specific signal could be generated on another key signaling protein, i.e. ERK2 (extracellular signal-regulated kinase 2), when treated with $L_{TA}D_1$ (Fig. 2, Fig. 4). ERK2 is very similar to p38 regarding its size and *pI* (44 kDa, *pI* 6.3) but the former has seven free thiol groups, four in the inner region, two on the surface and one in the active center of the protein (see ASA values in Table S1).[26, 27]

View Article Online
DOI: 10.1039/C6OB00810K

Journal Name

ARTICLE

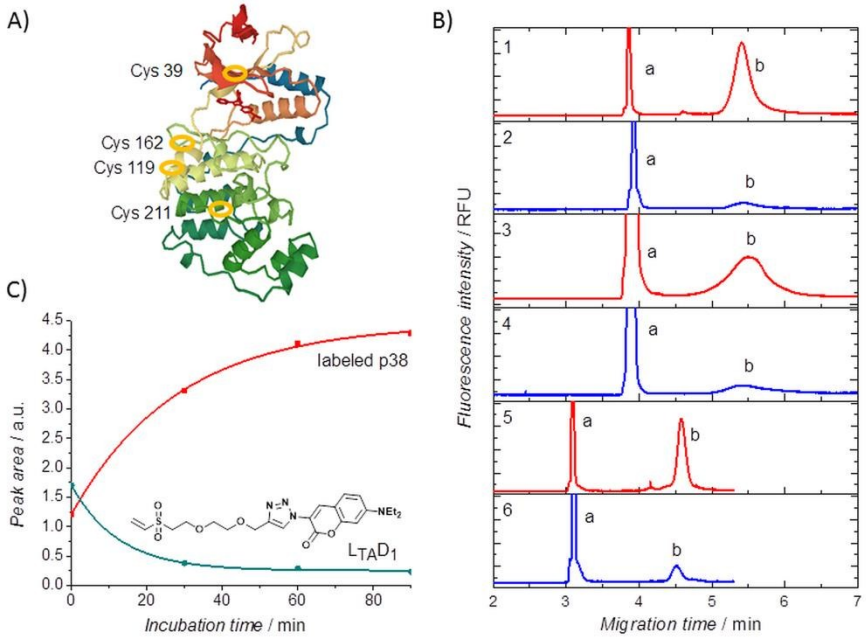


Fig. 2 A: Structure of p38 with free thiol groups (Protein Data Bank ID: 1A9U).[22] B: Capillary zone electrophoresis with laser induced fluorescent detection of fluorescent tagging of p38 (1, 2), p38^{C162S} (3, 4) and ERK2 (5, 6) with $L_{TA}D_1$; [proteins with (2, 4, 6) or without (1, 3, 5) alkylation pretreatment]. Peak 'a' corresponds to signal of $L_{TA}D_1$ and peak 'b' to the labeled protein. For details see Experimental section. C: Time-scale analysis of the derivatization process of p38 followed by the increasing area of peak 'b' and decreasing area of peak 'a' in CZE-LIF. Insert: Structure of the fluorescent tag $L_{TA}D_1$.

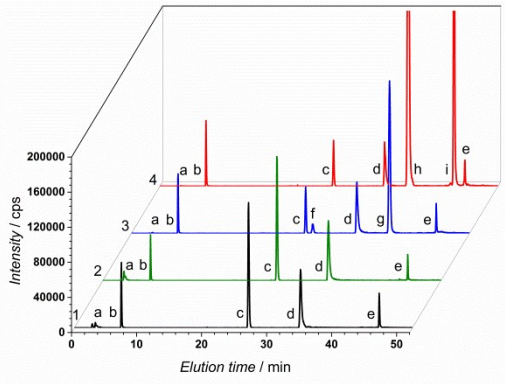


Fig. 3 Representative MRM chromatograms of tryptic digests of p38 (1), alkylated p38 labeled with L_{COMBO} (2) p38 labeled with L_{COMBO} (3) and p38 labeled with $L_{TA}D_1$ (4). Peaks were assigned by their MRM transitions. Peaks are numbered by their elution order: a: peptide fragment [119-121]; b: internal standard – peptide fragment [46-49]; c: peptide fragment [150-165]; d: peptide fragment [24-45]; e: peptide fragment [190-220]; f: peptide fragment [119-121] conjugated with L_{COMBO}; g: peptide fragment [150-165] conjugated with L_{COMBO}; h: peptide fragment [119-121] conjugated with $L_{TA}D_1$; i: peptide fragment [150-165] conjugated with $L_{TA}D_1$.

Though p38 was labeled efficiently and selectively, it should be mentioned that labeling schemes of other proteins used as models in the preliminary phase of this study brought our attention to a series of limitations (see ESI). More particularly, it was observed that during Cys-specific labeling, some proteins (e.g. albumin and α1-acid glycoprotein) showed strong ligand binding features where the undesirable non-covalent labeling contributed to the overall fluorescent signal. Furthermore, the accessibility of the Cys residues can further limit the efficiency of the fluorescent derivatization of native proteins.

Sequential labeling of p38 with new linkers and fluorogenic azide dyes

There are two options to circumvent the obstacle of the false positive labeling due to the non-covalent interactions in case of proteins, which have ligand binding capacity. One of the resolutions would be the separation of the ligand-bound linker-dye construct – which is a very time-consuming process whilst the protein can degrade. There is a more elegant

View Article Online
DOI: 10.1039/C6OB00810K

solution, however, i.e. to apply fluorogenic dye(s) in sequential labeling route. The main advantage of driving the labeling reaction in a sequential manner using fluorogenic dyes is that the dyes applied are not fluorescent until unreacted even if adsorbed non-specifically at biological surfaces[28-32] therefore provides no-wash procedures. To this end we developed bifunctional linkers (L_{COMBO}, L_{CyO}) with vinyl sulfone warhead and cyclooctyne moieties as biorthogonal functional groups (Fig 4.; see details of the synthesis in the Experimental Section Scheme 1). These strained rings allow copper free SPAAC reactions with fluorogenic azide dyes.

Testing of fluorogenic azide dyes. Our objective was to elaborate the performance of our cyclooctynylated vinyl sulfone linkers in combination with fluorogenic dyes in Cys-specific labeling schemes. Firstly, we have reacted our new linkers and six fluorogenic azide dyes (D_1-D_6)[10, 28, 29, 32] (Table 1). As expected, the substantial difference between the reactivity of the two cyclooctyne platforms[19, 20] resulted in less intense fluorescence increase in case of L_{CyO} than for L_{COMBO} (Table 1) therefore we decided to use L_{COMBO} for further studies.

It is crucial that the fluorogenic dyes are stable under the labeling conditions or their decomposition product is non-fluorescent. To get a deeper insight into the stability of our fluorogenic azides[10, 28, 29, 32] in reducing media, we have conducted a systematic screening study in order to assess the most suitable fluorogenic dyes. The fluorogenic performance of azide-quenched systems can be compromised in many ways. It is known, that azides are prone to reduction in the presence of strong reducing agents (ME, TCEP). However, there are also instances in the literature that cysteine itself[33] – even in peptides and proteins – can reduce azides (Table 1). Furthermore, it should be kept in mind that cytoplasm represents a highly reductive medium (cf. glutathione (GSH)) where the application of azide functions (or other bioorthogonal chemical reporters e.g. unnatural amino acids) needs careful attention as loss of azide integrity can lead to the appearance of non-specific signal (Table 1). In order to check the reducing effects of proteins containing reduced Cys, our azides D_1-D_6 were mixed with native albumin (BSA) – having a reduced cysteine in position 34. Evolution of intense fluorescent signals was observed in case of D_1, D_2, D_3 and D_6 in the presence of native BSA while no fluorescence could be detected using prealkylated form of the protein. These results indicated that fluorescent decomposition products generated upon the reaction with the reduced Cys and those compounds associated non-covalently to the protein giving a false positive signal of selective labeling. On the contrary, in the case of D_4 and D_5 only slight or insignificant emission was detected in the presence of reductive BSA in comparison with control experiments using non-reducing prealkylated BSA (Table 1, Table S2) demonstrating that these dyes remained intact.

The most appropriate dyes were selected from the series of D_1-D_6 based on their stability and reactivity. We ranked the dyes based on the changes in fluorescence intensity values i.e. we preferred the smallest increase in the presence of reducing agents – including BSA – referring to the best stability as well

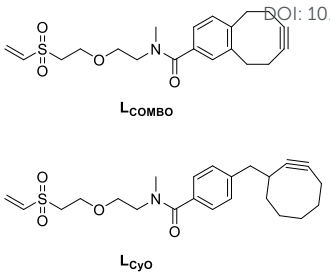
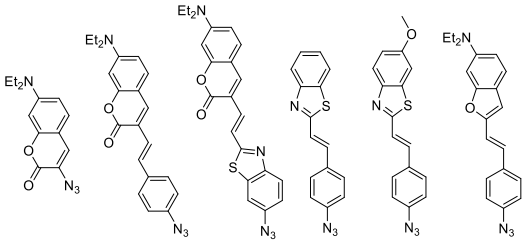


Fig. 4 Structures of the bifunctional linkers

Table 1 The evolution of fluorescent signal of dyes (D_1-D_6) incubated with L_{CyO} or L_{COMBO} and changes of intensity of fluorescent signal in the presence of reducing agents after two hours is shown.

		Fluorogenic azide dyes					
		D_1[a]	D_2[a]	D_3[a]	D_4[b]	D_5[c]	D_6[a]
							
		Changes in fluorescence intensity *(I/I_0)*					
Linkers	L_{CyO}	0.9	1.1	0.8	1.4	2.4	0.9
	L_{COMBO}	1.5	1.4	1.1	8.0	10.3	3.2
Reducing agents	TCEP	7.3	1.4	1.5	1.2	1.5	1.6
	ME	1.6	2.0	1.6	1.0	1.4	2.0
	Cys	2.3	2.5	2.6	0.9	1.2	2.6
	GSH	2.2	3.1	1.5	0.9	1.2	16.2
	BSA	2.4	11.2	18.9	4.0	7.0	52.2

as the greatest increase that corresponds to the highest reactivity with L_{COMBO} and L_{CyO} and the highest fluorogenicity.

Accordingly, dyes D_5 and D_4 and D_1 showed the highest stability and the largest reactivity. The fluorescence generated upon reaction with L_{COMBO} was slighter for D_4 than D_5, which is in good agreement with quantum yield data published previously.[29] Consequently, we have selected D_5 for further applications. In addition, in order to compare results obtained with $L_{TA}D_1$ we used dye D_1 as well.

Conjugation of linkers to the protein. We have applied the same reaction protocol for treatment of p38 with cyclooctynylated linker L_{COMBO} as with $L_{TA}D_1$. The specificity of L_{COMBO} binding to the protein was confirmed by MS (Fig. 3). As expected, the two surface accessible Cys (119 and 162) were targeted at similar modification ratios as observed with $L_{TA}D_1$ (95 ± 3 % and 68 ± 5 %, respectively, n = 3). Furthermore, we checked by subsequent addition of $L_{TA}D_1$ to cyclooctyne-treated p38 for any residual accessible thiol groups on p38. To

View Article Online
DOI: 10.1039/C6OB00810K

our delight, no substantial fluorescent signal could be detected for these p38 samples on SDS-PAGE suggesting that derivatization with the linker was complete within two hours (Fig. S2).

Labeling of linker conjugated p38 with fluorogenic azide dyes. We have successfully conducted sequential labeling schemes using L_{COMBO} and D_1 or D_5 monitored by CZE-LIF and CN-PAGE (Fig. 5, Fig. S3). $p38^{C162S}$ and ERK2 gave similar results upon tagging with L_{COMBO} and D_1 or D_5 (Fig. 5). It is worth mentioning that the emission wavelength of conjugated D_1 was found to be strongly dependent on the structure of the linker used but indifferent of the protein applied.

Specificity of the tagging was tested on pre-alkylated p38. Covalent modifications were identified exclusively at accessible cysteine residues as demonstrated by MS experiments. Only slight non-covalent binding was experienced in case of native p38 by CZE-LIF and CN-PAGE (Figure S3) indicating the existence of ligand binding effects, similarly to albumin (Fig. S4), although to a much less extent. Time scale analysis showed (Fig. 6) that the click reaction is fast and completed within 20 min in case of mixing L_{COMBO} with D_5, and slightly slower for the protein-L_{COMBO} conjugate with D_5.

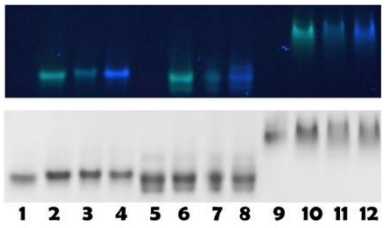


Fig. 5 Fluorescent tagging of p38, $p38^{C162S}$ and ERK2 with $L_{TA}D_1$, and with L_{COMBO} and dyes D_1 and D_5 monitored by CN-PAGE. Upper gel photo refers to fluorescent emission (λ_{ex} = 365 nm) and lower to the Coomassie staining. Samples 1: native, untreated p38 2: p38 labeled with $L_{TA}D_1$; 3: p38 conjugated with L_{COMBO} and labeled with D_1; 4: p38 conjugated with L_{COMBO} and labeled with D_5; 5: native, untreated $p38^{C162S}$ 6: $p38^{C162S}$ labeled with $L_{TA}D_1$; 7: $p38^{C162S}$ conjugated with L_{COMBO} and labeled with D_1; 8: $p38^{C162S}$ conjugated with L_{COMBO} and labeled with D_5; 9: native, untreated ERK2; 10: ERK2 labeled with $L_{TA}D_1$; 11: ERK2 conjugated with L_{COMBO} and labeled with D_1; 12: ERK2 conjugated with L_{COMBO} and labeled with D_5.

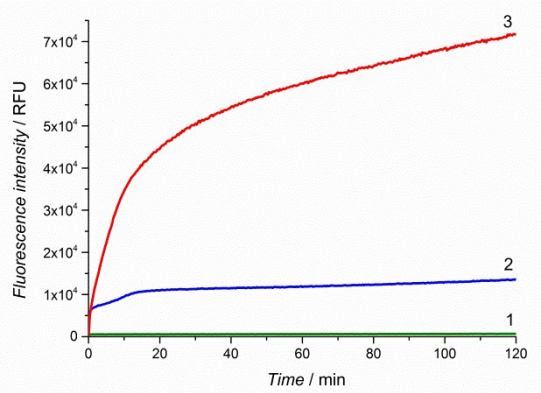


Fig. 6 Time-scale analysis of the click reaction of L_{COMBO} (2) and p38-L_{COMBO} (3) conjugate with D_5 compared to D_5 alone (1) monitored by fluorescent imaging (λ_{ex}: 355 nm/λ_{em}: 450 nm; n = 3).

The conjugation process followed a normal saturation type pattern with a rapid increase in fluorescence intensity at high reactant concentrations that reached a plateau after 30 minutes.

Noteworthy and a pleasant finding is the fact, that fluorescence intensity of L_{COMBO}-D_5 adduct was substantially less intensive compared to p38-L_{COMBO}-D_5 conjugate since the fluorescent intensity of D_5 increases 20 and 100 times upon tagging with L_{COMBO} and with p38-L_{COMBO} conjugate, respectively.

Conclusions

Herein we have demonstrated that cycloocytynylated thiol-specific vinyl sulfone warhead bearing chemical reporter(s) are suitable to derivatize appropriate protein platforms. The bioconjugation reaction of the native proteins is fast and selective. Subsequent click reaction with carefully selected fluorogenic dyes ultimately provides the generation of an intensive fluorescence signal and fulfills all requirements of bioorthogonality. Therefore, we propose the application of a new vinyl sulfone and COMBO functionalized thiol specific linker for the installation of a fast reacting bioorthogonal cyclooctyne motif. The COMBO moiety can be efficiently targeted with carefully selected fluorogenic dyes to achieve intensive and specific fluorescence.

Experimental

Organic syntheses

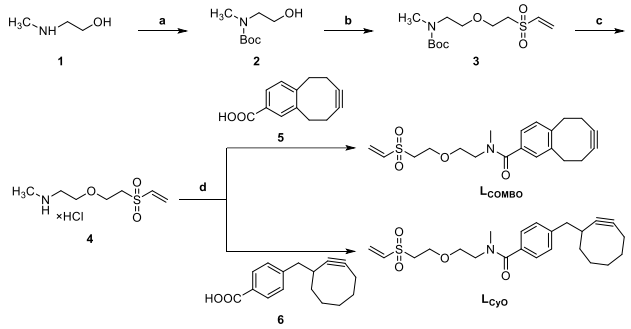
Unless otherwise indicated, all starting materials were obtained from commercial suppliers (Sigma-Aldrich, Fluka, Merck, Alfa Aesar, Reanal, Molar Chemicals, Fluorochem) and used without further purification. CyO and COMBO were synthetized according to literature procedures.[19, 20] $L_{TA}D_1$ was prepared as described previously.[10]

Analytical thin-layer chromatography (TLC) was performed on silica gel 60 F254 precoated aluminium TLC plates from Merck. Visualization of TLC samples was performed with a 254/365 nm UV lamp and/or $KMnO_4$ stain. Column chromatography was carried out with silica gel (0.06-0.2 mm) from Zeochem.

NMR spectra were recorded on a Varian Inova 500 MHz and Varian 600 MHz NMR spectrometers. Chemical shifts (δ) are given in parts per million (ppm) using solvent signals as the reference. Coupling constants (J) are reported in Hertz (Hz). Splitting patterns are designated as s (singlet), d (doublet), t (triplet), qr (quadruplet), qv (quintuplet), m (multiplet), dd (doublet of a doublet), td (triplet doublet), dt (doublet triplet), brs (broad singlet). The exact masses were determined with an Agilent 6230 time-of-flight mass spectrometer. IR spectra were obtained on a Bruker Alpha-P spectrometer on a single-reflection diamond ATR unit.

View Article Online
DOI: 10.1039/C6OB00840K

Scheme 1 Synthesis of L_{COMBO} and L_{CyO}: a) di-*tert*-butyl dicarbonate, DCM, 25°C, 14 h, 67 %; b) cat. ^{t}BuOK, DVS, THF, N_2, 25°C, 45 min, 52 %; c) CH_3COCl, MeOH, 0-25°C, 4 h, 98 %; d) HBTU, HOBt×H_2O, EDIPA, DMF, 25°C, 3 h, 20 % (L_{COMBO}) and 47 % (L_{CyO}).



Syntheses of cyclooctynylated linkers (L_{COMBO}, L_{CyO}). Boc-protected 2-(methylamino)-ethanol (**2**) was subjected to a Michael-type of addition with divinyl sulfone (DVS) in the presence of catalytic ^{t}BuOK to afford **3** in medium yield. Removal of the Boc protecting group was effected by treatment with trifluoroacetic (TFA) acid to furnish **4**. However, it was noted that the product was always slightly contaminated upon TFA mediated deprotection.

(L_{COMBO}, L_{CyO}) were prepared using COMBO-COOH (**5**)[20] or CyO-COOH (**6**)[19] under standard peptide coupling conditions (EDIPA, HOBt, HBTU – N,N-diisopropylethylamine, Thus, we applied in situ generated methanolic HCl solution, which furnished analytically pure hydrochloride salt of the free amine (**4**) in quantitative yield. Bifunctional chemical reporters hydroxybenzotriazole, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, respectively) (Scheme 1; for more details see ESI).

Fluorogenic dyes. Dyes D_1-D_6 (Table 1) have been prepared as described previously.[10, 28, 29, 32] Stock solutions were dissolved in 10 mM in DMSO.

Model Proteins. Human serum albumin, bovine serum albumin, carbonic anhydrase I (from human erythrocytes), superoxide dismutase I (from bovine erythrocytes) and α_1-acid glycoprotein were purchased from Sigma (St. Louis, MO, USA). All kinase protein constructs were recombinantly expressed in Escherichia coli Rosetta (DE3) pLysS (Novagen) cells with standard techniques.[25] p38α, p38α^{C162S} mutant and ERK2 were expressed with *N*-terminal cleavable hexahistidine tag. Recombinant proteins were purified with standard chromatographic techniques. The procedure involved a Ni-affinity chromatography step that was followed by an ion exchange step on a Resource Q column (GE Healthcare). The elution buffer contained 20 mM Tris (pH 8.0), 500 mM NaCl, 10 % glycerol, and 2 mM TCEP. Concentrations of stock solutions were 20 mg/mL (500 µM), 10 mg/mL (250 µM) and 13 mg/mL (325 µM) for p38, p38^{C162S} as well as for ERK2, respectively.

Calculation of solvent accessibility of Cys residues in p38. Atoms other than the target proteins' were removed from the structures. The cleaned structure files were submitted to accessible surface area (ASA) and relative ASA (RSA) calculations using DSSP[34] and the dssp Python package [http://jbloom.github.io/mapmuts/dssp_module.html] using the 'Tien2013' scale of maximum accessible surface area of amino acids.[35]

Conjugation processes

Conjugation of model proteins and $L_{TA}D_1$. 50 µM protein and 300 µM $L_{TA}D_1$ were mixed and incubated for 2 hours at room temperature in dark and in 50 mM sodium HEPES buffer (pH 8.0).

Conjugation of model proteins and linkers. 100 µM protein and 1 mM linker (L_{CyO} or L_{COMBO}) were mixed and incubated for 2 hours at room temperature in dark and in 50 mM sodium HEPES buffer (pH 8.0). The linker excess was removed by SpinPrep column (Sigma, St. Louis, MO, USA) filled with Sephadex G-25 "Fine" desalting gel (Pharmacia Fine Chemicals, Sweden).

Labeling of model proteins conjugated with linker with fluorogenic azide dyes. 50 µM protein and 300 µM fluorogenic dyes were mixed and incubated for 2 hours at room temperature in dark and in 50 mM sodium HEPES buffer (pH 8.0).

Alkylation of model proteins. The model proteins were alkylated with 700 times excess of iodoacetamide (IAM) for 1 hour at room temperature in dark and in 50 mM sodium HEPES buffer (pH 8.0). The unreacted reagent was removed by SpinPrep column (Sigma, St. Louis, MO, USA) filled with Sephadex G-25 "Fine" desalting gel (Pharmacia Fine Chemicals, Sweden).

Analytical methods

Capillary electrophoresis. Background electrolyte (BGE) components Tris (Tris(hydroxymethyl)-aminomethane and Tris-HCl, HEPES (N-[2-hydroxyethyl]piperazine-N'-[2-ethansulfonic acid]), were purchased from Sigma (St. Louis, MO, USA) sodium hydroxide and phosphoric acid were purchased from Merck GmbH (Darmstadt, Germany).

Capillary electrophoresis was performed with an Agilent Capillary Electrophoresis 3D^{CE} system (Agilent Technologies, Waldbronn, Germany) applying bare fused silica capillary having a 64.5 cm total and 56 cm effective length with 50 µm I.D. (Agilent Technologies, Santa Clara, CA, USA). On-line absorption at 200 nm was monitored by DAD UV-Vis detector. Laser (λ_{ex}: 355 nm) (Teem Photonics, Meylan, France) induced fluorescent (LIF) signal was observed by ZetaLIF Discovery detector (Picometrics, Labege, France). The capillary was thermostated at 25°C. Between measurements, the capillary

View Article Online
DOI: 10.1039/C6OB00810K

was rinsed subsequently with 0.1 M HCl, 1.0 M NaOH, 0.1 M NaOH and distilled water for 3 minutes each and with BGE for 5 minutes. 500 mM Tris-HCl (pH 9.0) buffer was used for p38 samples and 200 mM sodium phosphate (pH 9.0) for the other proteins. Samples were injected by 5×10^3 Pa pressure for 6 sec. Runs were performed in the positive-polarity mode with 30 kV.

Predicted acidic dissociation constant (*pKa*) of $L_{TA}D_1$ dye and theoretical isoelectric point (*pI*) of the proteins were calculated by ChemAxon's Marvin Calculator Plugin software (version 5.9.2; URL: http://www.chemaxon.com/marvin) and by ExPASy Bioinformatics Resorce Portal (URL:http://web.expasy.org/compute_pi/), respectively.

Clear native polyacrylamide gelelectrophoresis. Samples were diluted with sample buffer (62.5 mM Tris-HCl (pH 6.8) + 25 % glycerol + 0.01 % bromophenolblue) in 1:1 ratio. The size of the native polyaclyamide gels were 8.5 cm×7.5 cm×0.1 cm. They were the combination of 4 % concentration and 10 % separation PAGE gels (acrylamide:bisacrylamide ratio was 37.5:1; and gel buffer was 375 mM Tris-HCl (pH 8.8). 5.0 µg of protein was loaded per lane. Separations were carried out in Mini-Protean Tetra Cell (Bio-Rad, Hercules, CA, USA) using 25 mM Tris/192 mM Glycine; (pH 8.3) as running buffer and 150 V voltage for 70 min. Gels were documented by Fluorchem FC2 Imager (Alpha Innotech, San Leandro, CA, USA) at λ_{ex}: 365 nm. Furthermore, gels were stained for proteins with Coomassie-Brillant-Blue.

UV-Vis spectroscopy. Calibrations for p38 and dye-$L_{TA}D_1$ were obtained at 280 nm and 420 nm in quartz cuvette with 1 mm light pathlength by Jasco 7800 UV-Vis spectrophotometer. Dye/protein ratio was estimated from A_{420}/A_{280} with corrections according to the ε values obtained (for more details see ESI).

Screening of reduction of fluorogenic azide dyes by fluorescent detection. The effect of reducing agents TCEP, ME, Cys (300 µM), glutathione (5 mM) and BSA (50 µM) – with or without prealkylation – on fluorogenic azide dyes (300 µM; D_1-D_6) were studied based on the increase of their fluorescent signals (λ_{ex}: 355 nm/ λ_{em}: 410 nm; λ_{ex}: 355 nm/ λ_{em}: 450 nm; and λ_{ex}: 410 nm/ λ_{em}: 500 nm) by Biotek Synergy 2 Cytation 3 imaging plate reader with Gen5 software version 2.07 (Biotek Winooski, VT, USA) in round bottom, 384-well plate for fluorescence (Corning, Corning, NY, USA). In comparison, the evolution of fluorescent signal of the dyes (D_1–D_6) upon conjugation with L_{CyO} and L_{COMBO} after two hours of incubation was determined.

Time-trace analysis with monitoring the fluorescent signal. Derivatization reaction (20 µL) between the fluorogenic azide dye (300 µM; D_5) and the L_{COMBO} (50 µM) or p38 (50 µM) conjugated with L_{COMBO} were monitored based on the generation of the fluorescent signal (λ_{ex}: 355 nm; λ_{em}: 450 nm) by Biotek Synergy 2 Cytation 3 imaging plate reader with Gen5 software version 2.07 (Biotek Winooski, VT, USA) in round bottom, 384-well plate for fluorescence (Corning, Corning, NY, USA).

Mass Spectrometry. Tryptic digestion method was adapted from our former publication.[36] Briefly, 50 µL of p38 (2 mg/mL, 50 µM, 2.5 nmol) with or without pretreatment and 10 µL 0.2 % (w/v) RapiGest SF (Waters, Milford, USA) solution buffered with 50 mM ammonium bicarbonate (NH_4HCO_3) were mixed. 1.5 µL of 45 mM DTT in 100 mM NH_4HCO_3 were added and kept at 60°C for 30 min. After cooling the sample to room temperature, 2 µL of 100 mM iodoacetamide in 100 mM NH_4HCO_3 were added and placed in the dark for 30 min. The reduced and alkylated protein was then digested by 5 µL (2 mg/mL) trypsin (the enzyme-to protein ratio was 1 : 10) (Sigma, St. Louis, MO, USA). The sample was incubated at 37 °C for 60 min. To degrade the surfactant, 7 µL of formic acid (500 mM) solution was added to the digested p38 sample to have the final concentration between 30 and 50 mM (pH ≈ 2) and was incubated at 37°C for 45 min. For LC-MS analysis, the acid treated sample was centrifuged for 10 min at 13 000 rpm.

QTRAP 6500 triple quadruple – linear iontrap mass spectrometer, equipped with a Turbo V Source in electrospray mode (AB Sciex, CA, USA) and a Perkin Elmer Series 200 micro LC system (Massachusetts, USA) was used for LC-MS/MS analysis. Data acquisition and processing were performed using Analyst software version 1.6.2 (AB Sciex Instruments, CA, USA). Chromatographic separation was achieved by Vydac 218 TP52 Protein & Peptide C18 column (250 mm × 2.1 mm, 5µm). Sample was eluted with a gradient of solvent A (0.1 % formic acid in water) and solvent B (0.1 % formic acid in ACN). Flow rate was set to 0.2 mL/min. The initial conditions for separation were 5 % B for 4 min, followed by a linear gradient to 30 % B by 40 min, from 41 to 50 min 90 % B, from 51 to 55 min 90 % B is retained; from 56 to 60 min back to initial conditions with 5 % eluent B. The injection volume was 10 µL (300 pmol on column). The MS/MS system was operated under multiple reaction monitoring (MRM) mode, with a dwell time of 100 ms for each transition, using nitrogen as the collision gas, set at "high".

Acknowledgements

The authors thank Attila Reményi and Anita Alexa for the conductive discussions. The authors would like to thank András Herner for preparation of some of the applied fluorogenic dyes. Ilona Kawka, Marianna Rakács and Holly Sofka are greatly acknowledged for the technical support. The research was financially supported by the Hungarian Academy of Sciences and the Hungarian Scientific Research Fund (Lendület Grant - LP 2013-55/2013 and OTKA-NN-110214, OTKA-NN-116265).

DOI: 10.1039/C6OB00810K

References

1 R. Wombacher and V. W. Cornish, *J Biophotonics*, 2011, **4**, 391.

2 G. B. Cserép, A. Herner and P. Kele, *Methods and Applications in Fluorescence*, 2015, **3**, 042001.

3 C. S. McKay and M. G. Finn, *Chem. Biol.*, 2014, **21**, 1075.

4 G. Zhang, S. Zheng, H. Liu and P. R. Chen, *Chem. Soc. Rev.*, 2015, **44**, 3405.

5 G. B. Cserép, A. Herner, O. S. Wolfbeis and P. Kele, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 5776.

6 A. Izquierdo-Alvarez and A. Martinez-Ruiz, *J Proteom*, 2011, **75**, 329.

7 O. Koniev, G. Leriche, M. Nothisen, J. S. Remy, J. M. Strub, C. Schaeffer-Reiss, A. Van Dorsselaer, R. Baati and A. Wagner, *Bioconjug. Chem.*, 2014, **25**, 202.

8 A. J. Meyer and T. P. Dick, *Antioxid. Redox Signal.*, 2010, **13**, 621.

9 C. Yin, F. Huo, J. Zhang, R. Martinez-Manez, Y. Yang, H. Lv and S. Li, *Chem. Soc. Rev.*, 2013, **42**, 6032.

10 G. B. Cserép, Z. Baranyai, D. Komáromy, K. Horváti, S. Bősze and P. Kele, *Tetrahedron*, 2014, **70**, 5961.

11 K. Lang and J. W. Chin, *ACS Chem. Biol.*, 2014, **9**, 16.

12 J. Morales-Sanfrutos, F. J. Lopez-Jaramillo, F. Hernandez-Mateo and F. Santoyo-Gonzalez, *J. Org. Chem.*, 2010, **75**, 4039.

13 J. Morales-Sanfrutos, J. Lopez-Jaramillo, M. Ortega-Munoz, A. Megia-Fernandez, F. Perez-Balderas, F. Hernandez-Mateo and F. Santoyo-Gonzalez, *Org. Biomol. Chem.*, 2010, **8**, 667.

14 R. K. Lim and Q. Lin, *Chem. Commun. (Camb.)*, 2010, **46**, 1589.

15 M. Grammel and H. C. Hang, *Nat. Chem. Biol.*, 2013, **9**, 475.

16 D. M. Patterson, L. A. Nazarova and J. A. Prescher, *ACS Chem. Biol.*, 2014, **9**, 592.

17 X. Chen and Y.-W. Wu, *Org. Biomol. Chem.*, 2016.

18 V. Hong, S. I. Presolski, C. Ma and M. G. Finn, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 9879.

19 N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi, *ACS Chem. Biol.*, 2006, **1**, 644.

20 B. R. Varga, M. Kállay, K. Hegyi, S. Béni and P. Kele, *Chem. Eur. J.*, 2012, **18**, 822.

21 M. C. Huber, A. Schreiber, P. von Olshausen, B. R. Varga, O. Kretz, B. Joch, S. Barnert, R. Schubert, S. Eimer, P. Kele and S. M. Schiller, *Nat Mater*, 2015, **14**, 125.

22 Z. Wang, B. J. Canagarajah, J. C. Boehm, S. Kassisa, M. H. Cobb, P. R. Young, S. Abdel-Meguid, J. L. Adams and E. J. Goldsmith, *Structure*, 1998, **6**, 1117.

23 A. Cuadrado and A. R. Nebreda, *Biochem. J.*, 2010, **429**, 403.

24 E. K. Kim and E. J. Choi, *Arch. Toxicol.*, 2015, **89**, 867.

25 A. Garai, A. Zeke, G. Gogl, I. Törő, F. Fördős, H. Blankenburg, T. Barkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi, *Sci. Signal*, 2012, **5**, ra74.

26 W. Peti and R. Page, *Protein Sci.*, 2013, **22**, 1698.

27 T. Kinoshita, M. Warizaya, M. Ohori, K. Sato, M. Neya and T. Fujii, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 55.

28 A. Herner, G. Estrada Girona, I. Nikic, M. Kállay, E. A. Lemke and P. Kele, *Bioconjug. Chem.*, 2014, **25**, 1370.

29 A. Herner, I. Nikic, M. Kállay, E. A. Lemke and P. Kele, *Org. Biomol. Chem.*, 2013, **11**, 3297.

30 P. Shieh, M. J. Hangauer and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2012, **134**, 17428.

31 A. Nadler and C. Schultz, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 2408.

32 K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill and Q. Wang, *Org. Lett.*, 2004, **6**, 4603.

33 J. V. Staros, *Trends Biochem. Sci.*, 1980, **5**, 320.

34 W. G. Touw, C. Baakman, J. Black, T. A. H. te Beek, E. Krieger, R. P. Joosten and G. Vriend, *Nucleic Acids Res.*, 2014, **43**, D364.

35 A. Porollo, M. Z. Tien, A. G. Meyer, D. K. Sydykova, S. J. Spielman and C. O. Wilke, *PLoS One*, 2013, **8**, e80635.

36 T. Imre, G. Schlosser, G. Pocsfalvi, R. Siciliano, E. Molnár-Szöllősi, T. Kremmer, A. Malorni and K. Vékey, *J. Mass Spectrom.*, 2005, **40**, 1472.