CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Protein Glycosylation via Sulfur Fluoride Exchange (SuFEx) Chemistry: The Key Role of a Fluorosulfate Thiolactoside

Authors: Cristina Nativi, Alberto Marra, Jiajia Dong, Tiancheng Ma, Stefano Giuntini, Elisa Crescenzo, Linda Cerofolini, Marco Martinucci, Claudio Luchinat, Marco Fragai, and Alessandro Dondoni

To be cited as: *Chem. Eur. J.* 10.1002/chem.201803912

Link to VoR: http://dx.doi.org/10.1002/chem.201803912

Supported by
ACES

WILEY-VCH

10.1002/chem.201803912

WILEY-VCH

FULL PAPER

Protein Glycosylation via Sulfur Fluoride Exchange (SuFEx) Chemistry: The Key Role of a Fluorosulfate Thiolactoside

A. Marra, [b] J. Dong, [c] T. Ma, [c] S. Giuntini, [a],[d] E. Crescenzo, [a],[d] L. Cerofolini, [a],[d] M. Martinucci, [a] C. Luchinat, [a],[d] M. Fragai, *[a],[d] C. Nativi*[a] and A. Dondoni*[e]

Dedicated to Professor K. Barry Sharpless on the occasion of his 77th birthday

Abstract: Protein glycosylation is the most complex post-translational modification process. More than 50% of human cells proteins are glycosylated, while bacteria such as *E. coli* do not have this modification machinery. Indeed, the carbohydrate residues in natural proteins affect their folding, immunogenicity, and stability toward proteases, besides controlling biological properties and activities. It is therefore important to introduce such structural modification in bioengineered proteins lacking the presence of carbohydrate residues. This is not a trivial as it requires reagents and conditions compatible with protein's stability and reactivity. We report herein on the introduction of lactose moieties in two natural proteins, namely ubiquitin (Ub) and L-asparaginase II (ANSII). The synthetic route employed is based on the Sulfur(VI) Fluoride Exchange (SuFEx) coupling of a lactose tethered arylfluorosulfate (Lact-Ar-OSO_2F) with the ε-NH_2 group of lysine residues of the proteins. This metal free click SuFEx reaction relies on the properties of the fluorosulfate employed which is easily prepared in multigram scale from available precursors and reacts chemoselectively with the ε-NH_2 group of lysine residues under mild conditions. Thus, iterative couplings of Lact-Ar-OSO_2F to Ub and ANSII, afforded multiple glycosylations of these proteins so that up to three and four Lact-Ar-OSO_2 groups were introduced in Ub and ANSII, respectively, via the formation of a sulfamoyl (OSO_2-NH) linkage.

[a] Dr. S. Giuntini, E. Crescenzo, Dr. L. Cerofolini, M. Martinucci, Prof. C.Luchinat, Prof. M. Fragai, Prof. C. Nativi
Department of Chemistry
University of Florence
via della Lastruccia, 3-13 Sesto F.no (FI) 50019 Italy
E-mail: fragai@cerm.unifi.it; cristina.nativi@unifi.it
[b] Prof. A. Marra
Institut des Biomolécules Max Mousseron (IBMM), UMR 5247, Université de Montpellier, CNRS, Ecole Nationale Supérieure de Chimie de Montpellier
8 Rue de l'Ecole Normale, 34296 Montpellier- cedex 5, France
[c] Prof. J. Dong, T. Ma
Key Laboratory of Organofluorine Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry
University of Chinese Academy of Sciences, Chinese Academy of Sciences
345 Lingling Road, Shanghai 200032, P. R. China
[d] Prof. C. Luchinat, Prof. M. Fragai, Dr. S. Giuntini, E. Crescenzo, Dr. L. Cerofolini
CERM and CIRMMP
via Luigi Sacconi, 6, 50019 Sesto F.no (FI), Italy
[e] Prof. A. Dondoni
Interdisciplinary Center for the Study of Inflammation
University of Ferrara,
Italy

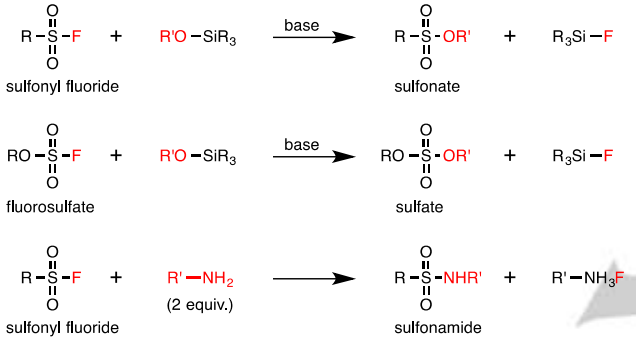
Supporting information for this article is given via a link at the end of the document.

Introduction

Protein glycosylation is an important post-translational modification in eukaryotic organisms and is biosynthetically mediated by numerous glycosyltransferases.[1] The extensive occurrence of such modification is supported by the fact that 50% of proteins in human cells are glycosylated[2] and 70% of the total of therapeutic proteins, that are currently in clinical trials, are glycoproteins.[3] Indeed the carbohydrate residues are not just decorative elements as they profoundly affect protein folding, immunogenicity, and stability toward proteases besides controlling biological properties and activities.[4] The latter functions operate mainly in cellular processes, such as cellular recognition and adhesion, cell growth and differentiation.[5] In this way glycoproteins are involved in vital biological events, either detrimental (inflammation, cancer metastasis, viral and bacterial infection),[6] or beneficial (immune response, fertilization). Studies on these issues are made problematic because the isolation of glycoproteins with a well-defined carbohydrate structure from natural sources is difficult and even with current techniques virtually impossible. This is due to two main factors, that is: *i)* the oligosaccharide micro heterogeneity produced by the presence of various glycoforms, and *ii)* the very labile *N*- and *O*-acetal bonds, which are the most diffuse connections between the terminal carbohydrate of the oligosaccharide fragment and asparagine or serine/threonine residue, respectively. Therefore, it is difficult to identify the specific oligosaccharide structures that are responsible of an individual biological process or protein function. Hence there is an urgent need for the preparation of non-natural glycoproteins exhibiting robust anomeric linkages and a well-defined structure and composition. Another reason for such request comes from impact in medicine. In fact, it is not surprising that a substantial fraction of the currently approved protein pharmaceuticals need to be properly glycosylated to exhibit optimal therapeutic efficacy.[7] In some cases the glycosylation of natural proteins lacking carbohydrate fragments can lead to new properties as precursors to valuable drugs or it is used to produce glycovaccines.[8] Moreover, several studies demonstrated that glycosylation induces various effects on the stability of protein pharmaceuticals.[9] Therefore substantial efforts are being made

10.1002/chem.201803912

for the efficient glycosylation of proteins,[10] and for developing biophysical methodologies to characterize these protein derivatives.[11] Actually glycosylation of proteins is not a trivial task as it requires a chemoselective chemical ligation that introduces the carbohydrate moieties under conditions which are compatible with the protein environments such as room temperature and aqueous medium, as well as the absence of a metal catalyst. The wise choice in the arsenal of click reactions may offer suitable solutions to those requests. Indeed, in addition to the prototypical click reaction,[12] represented by the copper catalysed alkyne azide cycloaddition,[13] many other efficient click processes are available,[14] For instance, a few years ago another good process for click chemistry was revived from an unrecognized state[15] by Sharpless and co-workers.[16] This methodology is based on the reactivity of sulfonyl fluorides and fluorosulfates with O- and N-nucleophiles and therefore was named as sulfur fluoride exchange (SuFEx) (Scheme 1).



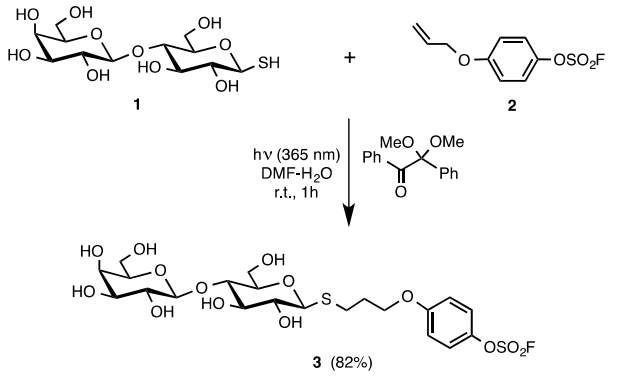
Scheme 1. SuFEx reactions of fluoro derivatives with silyl ether and amine.

Applications of SuFEx chemistry dealing with small molecules and biomacromolecules have been reported in recent papers.[17] Based on these premises and given our interest in peptide and protein glycosylation by click chemistry,[18] we report below on the coupling of a glycosylated fluorosulfate to lysine residues of ubiquitin (Ub) and *E. Coli* L-asparaginase II (ANSII). Ub is a 76-amino acid protein, that incorporates seven lysine residues, and which is well known for its role in post-translational modification (PTM) of many proteins. The use of glycosylated Ub would expand the scope of the ubiquitination as the modified protein would be also affected by the presence of the carbohydrate moiety. Also ANSII is a lysine rich enzyme, approved for the clinical use against acute lymphoblastic leukaemia (ALL). This tetrameric protein featuring a molecular mass of about 140 kDa, is well known for its ability in catalysing the hydrolysis of asparagine, an essential amino acid for protein biosynthesis in leukemic cells, into ammonia and aspartate. Two main disadvantages in the therapeutic use of this biological drug are the need for frequent intramuscular injections and the very high rate of allergic reactions. Fortunately, the modification of the drug by pegylation (PEG) has been shown to reduce these inconveniences while retaining the anti-leukemic effectiveness[.[19] and the original structural features of the protein.[20] Another modification of ANSII that turned out to be beneficial on its properties as a drug is the glycation by lactose mediated by sodium cyanoborohydride, i.e. a reductive coupling, which was reported forty years ago.[21] In fact, the glycated enzyme showed increased thermal stability and resistance to proteolytic cleavage. These promising results support our program of performing an approach for the glycosylation of ANSII using new chemistry based on the click SuFEx strategy.

Results and Discussion

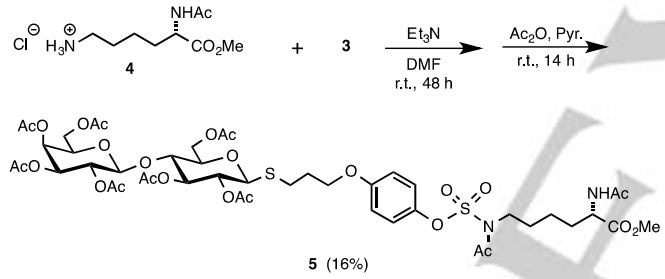
The initial part of our work consisted in the preparation of a glycosylated reagent containing the SO_2F functionality suitable for the SuFEx reaction with the ε-amino group of lysine residues natively present in Ub and ANSII. As recently the SuFEx reaction of a polyethylene glycol fluorosulfate (PEG-OSO_2F) with bovine serum albumin (BSA) has been reported,[22] we decided to employ a glycosylated fluorosulfate for our SuFEx approach .[23] No examples of protein functionalization by fluorosulfate derivatives of biomolecules (sugar, amino acids, etc.) are known to date. Despite its strong potential, the SuFEx approach has been scarcely exploited for the functionalization of proteins. In 2016, the coupling of a single polyethylene glycol fluorosulfate (PEG-OSO_2F) with the lysine residues of commercially available bovine serum albumin (BSA) was reported by Averick and co-workers.[22] In the same year, the Sharpless group demonstrated that PEG-fluorosulfate derivatives can react[23a] chemoselectively with the phenolic hydroxyl of tyrosine residues within the binding site of intracellular lipid binding proteins and, more recently, that various aryl fluorosulfates show a similar reactivity toward other proteins.[23b] Therefore, no examples of protein functionalization by fluorosulfate derivatives of biomolecules (e.g. sugar and amino acids) are known to date. Since fluorosulfate bearing carbohydrates have never been described, we envisaged the synthesis of a reagent featuring a sugar unit anomerically linked to the fluorosulfate functionality through a suitable aryl tether. Hence the lactose fluorosulfate **3** was prepared in 82% yield by photoinduced thiol-ene coupling (TEC)[24] of the 4-allyloxyphenyl fluorosulfate[25] **2** with the readily accessible 1-thio-β-D-lactose[26] **1** in the presence of 2,2-dimethoxy-2-phenylacetophenone (DPAP) as photoinitiator (Scheme 2). Quite remarkably the reagent **3** featured the sugar residue linked to the arylfluorosulfate moiety through anomeric carbon-sulfur bond. This linkage would provide the glycosylated products from the SuFEx coupling with high stability toward hydrolases. Moreover, it turned out that the lactose fluorosulfate **3** showed sufficient water solubility and stability[27] to be safely used in reactions with the target proteins in aqueous medium under mild basic conditions and at room temperature.

10.1002/chem.201803912



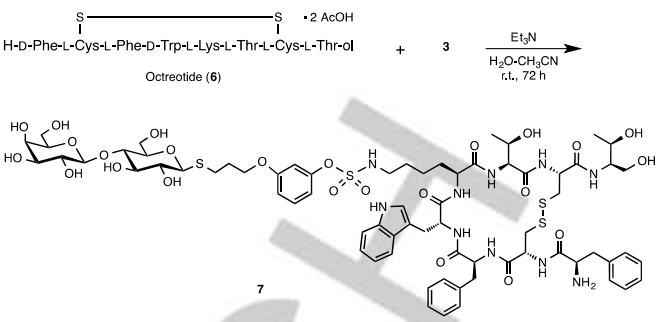
Scheme 2. Synthesis of the lactose fluorosulfate **3**.

In order to prove the SuFEx reaction of **3** with the ε-amino group of lysine residues, two model reactions were carried out. The first experiment involved the coupling of **3** with protected lysine. Indeed, the reaction of **3** (2 equiv.) with *N*-acetyl-L-lysine methyl ester hydrochloride **4** in the presence of Et_3N (3 equiv.) in DMF as the solvent (r.t., 48 h), followed by acetylation, afforded the sulfamate derivative **5** in 16% isolated yield. This result demonstrated the occurrence of an effective SuFEx coupling of the OSO_2F group of **3** with the ε-NH_2 of lysine.



Scheme 3. Synthesis of the glycosylated lysine **5**.

Another evidence for the above chemoselectivity came from the coupling of **3** (10 equiv.) with the octapeptide Octreotide[28] bis-acetate **6** featuring a lysine residue in its body. Indeed, this reaction carried out at room temperature in the presence of Et_3N afforded the product **7** (Scheme 4). The structure of **7** derived from the SuFEx reaction of **3** with the ε-NH_2 of lysine residue is in agreement with that of the product formed in the aza-Michael coupling of Octreotide with N-phenylvinylsulfonamide derivatives.[29] Notably the introduction of a lactose residue in Octreotide bears its own value as this modification may affect the pharmacokinetic properties of the peptide, that is in fact employed for the treatment of acromegaly, a pathology that is characterized by hormonal disorder in adult age patients.



Scheme 4. Synthesis of the glycosylated peptide **7**.

Conjugation of the lactose fluorosulfate with the proteins. Having demonstrated that the fluorosulfate **3** is enough reactive with the free ε-NH_2 of lysine residues thus leading to a sulfamoyl linkage, the SuFEx reaction of **3** was then carried out on Ub and ANSII, as both these proteins bear lysine residues. In both cases, the reactions were carried out in phosphate buffer (pH 8) and at room temperature with an excess of fluorosulfate **3**. The glycosylation was monitored by gel electrophoresis and the excess of **3** removed by dialysis. The product of reaction between **3** and Ub was investigated through MALDI mass spectrometry and NMR for the determination of the number of derivative molecules attached to each protein molecule, and the conjugation pattern of each lysine residue.

The mass spectrum of ubiquitin after the reaction with **3** displays three main peaks (Figure S1): *i)* one peak corresponds to the mass of the free protein, *ii)* a second peak to the mass of ubiquitin functionalized with a single molecule of disaccharide derivative, and *iii)* a third to the mass of ubiquitin with two molecules of disaccharide derivative. The most intense peak among the three is that corresponding to the free protein. This result suggests the presence in solution of a heterogeneous mixture of functionalized protein. In particular, from zero up to two molecules of disaccharide can be covalently linked to each molecule of Ub, involving the different solvent exposed NH_2 groups. The 2D ^1H-^{15}N HSQC NMR spectrum collected on the ubiquitin sample after the derivatization reaction confirm the presence of the heterogeneous population of isomers (Figure 1). In particular, the ^1H-^{15}N cross-peaks corresponding to the unreacted protein are still present and intense; concomitantly new H-N cross-peaks (one or more for each original cross-peak) at different values of chemical shift appeared. These new signals can be attributed to the H-N backbone atoms of the new species generated after the reaction with **3**. For most of the cross-peaks, the intensity of the new signals is lower than that of the unreacted species, in agreement with the results of the mass spectrometry analysis. Unfortunately, the new cross-peaks corresponding to the sulfamate protons of the reacted lysine sidechains could not be detected because of the pH of the buffered solution (pH = 8) that ensures the chemical stability of the conjugate, and the low pKa value of the sulfamate moiety.

10.1002/chem.201803912

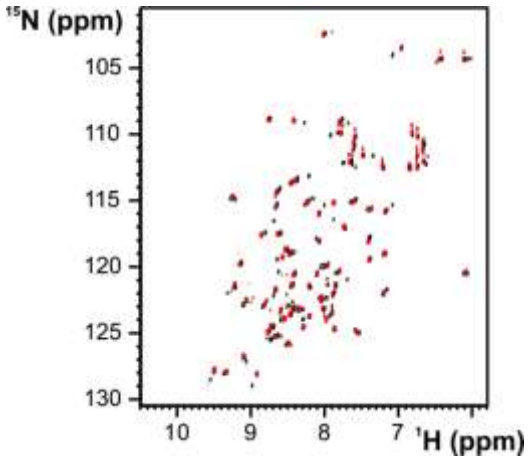


Figure 1. Superimposition of the 2D ^{1}H-^{15}N HSQC spectra of free ubiquitin (red) and ubiquitin after one SuFEx reaction with the disaccharide derivative (black). In the latter spectrum signals corresponding to the unreacted ubiquitin are still visible. The spectra were recorded at 298 K and 950 MHz.

Considering the presence of unreacted lysine residues still available for functionalization on the protein surface, the SuFEx reaction was repeated two more times on the same protein sample. The mass spectrum of the protein, subjected to three successive SuFEx reactions with **3**, shows four peaks (Figure S2): the same three as after a single SuFEx reaction, and an additional one corresponding to ubiquitin with three molecules of disaccharide derivative. Interestingly, the relative intensity of these peaks changes after each SuFEx reaction, and finally, the most intense peak in the mass spectrum corresponds to Ub with one molecule of disaccharide derivative conjugated to the protein. To shed light on the conjugation degree of each lysine residue, an NMR analysis has been carried out on samples of Ub before and after the SuFEx reactions. In the 1D ^{1}H NMR spectrum of the protein after three successive SuFEx reactions with **3**, the signals are broader than the signals in the spectrum of the free protein (Figure S3) and of the protein after a single reaction. This can be explained by an increase of both the hydrodynamic volume and the heterogeneity of the conjugated protein population. A more detailed analysis of the aliphatic region of the spectrum around 0 ppm indicates the presence of an additional signal at the resonance frequency of -0.05 ppm (Figure 2). This new peak, nearby the signal assigned to the methyl group (Qδ1) of Leu-50, corresponds to the same methyl group experiencing a different chemical environment because of the conjugation of the neighboring Lys-48 with disaccharide chains. The presence in the spectra of two signals with the same intensity and corresponding to the free and reacted forms of the protein indicates for Lys-48 a ratio of about 1:1 between the reacted and non-reacted species. Of note, this peak after only a single SuFEx reaction is barely visible.

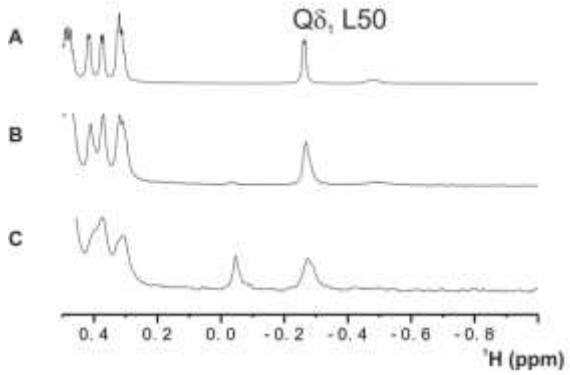


Figure 2. 1D ^{1}H NMR spectra of free ubiquitin (A), ubiquitin after one SuFEx reaction with **3** (B), and ubiquitin after three successive SuFEx reactions with the disaccharide derivative (C). The assignment of the methyl group (Qδ1) of Leu-50 of free ubiquitin is reported in the spectrum. The spectra were recorded at 298 K and 950 MHz.

The 2D ^{1}H-^{15}N HSQC spectrum acquired on the protein sample after three successive SuFEx reactions with the disaccharide derivative shows a similar pattern of cross-peaks of that after a single SuFEx reaction, but with different relative signal intensity (Figure S4). In particular, the new cross-peaks are increased in intensity. This phenomenon is sizable for some lysine residues (Lys-11, Lys-33 and Lys-48). For Lys-11 and Lys-33 the cross-peak of the unreacted species is almost completely disappeared after three successive SuFEx reactions, while the signal intensity of the new cross-peaks corresponding to the reacted lysine residue are sizable increased. Conversely, the cross-peak corresponding to the reacted form of Lys-48 is barely distinguishable in the spectrum after a single SuFEx reaction, but it shows the same intensity of the peak of the unreacted species after three SuFEx reactions (Figure S5). This is in agreement with what observed also for the methyl group of Leu-50 in the monodimensional spectra (see above).

A more detailed analysis on the conjugation sites of Ub has been provided by the analysis of the chemical shift perturbation. As expected, the protein residues experiencing the largest effects are the lysines and residues in their close proximity (Figure S6-S7). For a semi-quantitative evaluation of the conjugation degree of each lysine residue, the relative intensity of the cross-peaks corresponding to the reacted and unreacted species within the same spectrum has been analyzed in detail and graphically reported in Figure 3.[11]

10.1002/chem.201803912

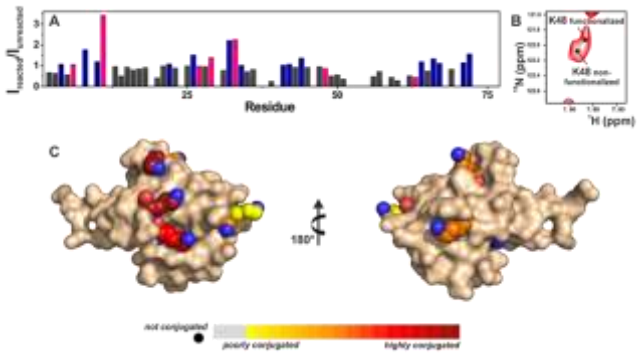


Figure 3. (A) Graphical representation of the intensity ratio per residue of the reacted species with lactose fluorosulfate **3** and the unreacted one present in the same spectrum after three successive SuFEx reactions with **3**. The lysine residues are in magenta and the residues experiencing the largest variation are in blue. (B) Enlargement of the 2D ^{1}H-^{15}N HSQC spectrum of ubiquitin after three successive SuFEx reactions with **3**. On the spectrum is indicated an example of the cross-peaks corresponding to the functionalized and non-functionalized Lys-48 residue used to evaluate the signal intensity ratio. (C) Surface representation of ubiquitin (pdb code: 1UBQ) with the lysine residues and the N-terminus color coded according their conjugation degree.

A similar integrated strategy has been used to analyze the pharmacologically relevant ANSII conjugated with the same thio-lactose derivative. The comparison of the 1D ^{1}H spectra collected before and after each SuFEx reaction of **3** with the protein, shows a group of signals between 3 and 4.5 ppm, which increases progressively in intensity after each reaction (Figure 4).

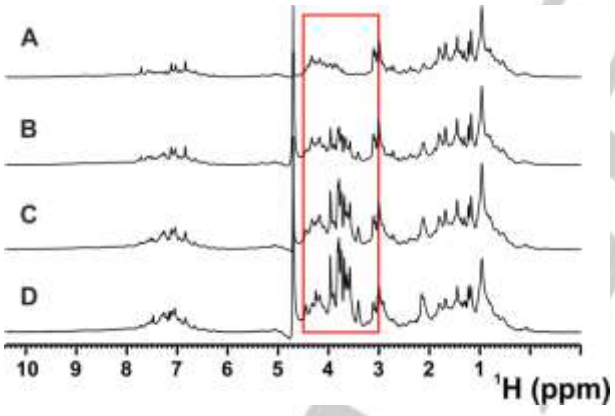


Figure 4. 1D ^{1}H NMR spectra of free ANSII (A), ANSII after one (B), two (C) and three (D) successive SuFEx reactions with **3**. The red square highlight the region between **3** and 4.5 ppm, where particularly emerge some of the protons of the lactose moiety. The spectra were recorded at 310 K and 900 MHz.

Unfortunately, the very large molecular weight of ANSII makes it unfeasible the analysis of the products of the reactions by 2D ^{1}H-^{15}N HSQC. However, as previously observed for Ub, the analysis carried out by mass spectrometry shows that the SuFEx reaction on ANSII results in a heterogeneous mixture of positional glycosides (Figure S8). Interestingly, the signal of the free ANSII is completely missing in MALDI spectrum and from the broad signal, the peaks of the protein glycosylated with one, two, three and four disaccharide units per monomer of ANSII can be distinguished.

The introduction of lactose residues is expected to improve the stability of ANSII against proteolytic enzymes. To assess the improved stability of the glycoconjugated protein, samples of ANSII functionalized with **3** have been incubated at 37 °C with trypsin or with the catalytic domain of human matrix metalloproteinase 1 (MMP-1) and monitored up to 3 days (see Supporting Information). The effect of the proteolytic enzymes has been monitored by SDS-PAGE gel. From the visual inspection of the gels, it was immediately evident that the conjugation with **3** prevents the proteolytic degradation of ASNII by trypsin that conversely takes place when the free protein is incubated with trypsin under the same experimental conditions (Figure S10 and S11). Moreover, the conjugation with Lact-Ar-OSO2 groups is also capable to slow down significantly the degradation of ASNII treated with MMP-1, a very efficient proteolytic enzyme involved in the proteolytic degradation of extracellular proteins and glycoproteins (Figure S12).

Conclusions

In summary, we have developed an efficient protocol for the introduction of lactose groups on pharmaceutical relevant biomolecules. This was made possible by using a typical SuFEx process involving the coupling of a lactose bearing fluorosulfate with the free ε-amino groups of lysine moieties incorporated into the proteins. In this way, multiple glycosylation of Ub and ANSII were carried out being the lactose residues linked to the lysine fragments through a sulfamoyl tether. The glycosylation of ANSII is expected to slow down the clearance of the enzyme in vivo, thus increasing its half-life with a beneficial reduction of the drug administration frequency. Due to a remarkable versatility, the SuFEx might be suitable for the decoration of challenging proteins with a panel of different saccharides or biologically relevant residues.

Experimental Section

Reactions were monitored by TLC on silica gel 60 F_{254} with detection by charring with sulfuric acid. Flash column chromatography was performed on silica gel 60 (40-63 μm). Optical rotations were measured at 20 ± 2 °C in the stated solvent; $[\alpha]_D$ values are given in deg mL g^{-1} dm^{-1}. ^{1}H (400 MHz), ^{13}C (100 MHz), and ^{19}F (376 MHz) NMR spectra were recorded in the stated solvent at room temperature unless otherwise specified. MALDI-TOF (Bruker Ultraflex III) and high resolution mass spectrometry (Waters Micromass Q-TOF) analyses of compounds **3** and **5** were performed at the *Laboratoire de Mesures Physiques*, IBMM – University of Montpellier. The photoinduced thiol-ene reaction was carried out in a glass vial located 2.5 cm away from the household UV-A lamp apparatus equipped with four 15 W tubes (1.5 x 27 cm each). The commercially available photoinitiator 2,2-dimethoxy-2-phenylacetophenone (DPAP) was used without further purification.

Lactose fluorosulfate 3. A solution of thiol **1** (1.433 g, 4.00 mmol), fluorosulfate **2** (1.207 g, 5.20 mmol), and DPAP (102 mg, 0.40 mmol) in DMF (5 mL), partially concentrated under vacuum (ca. 0.1 mbar) to remove the traces of Me_2NH, and H_2O (1 mL) was irradiated (λ_{max} 365 nm) under vigorous stirring at room temperature for 1 h and then concentrated. The residue was triturated with CH_2Cl_2 (2 x 10 mL) and the precipitate, recovered by decantation, afforded, after drying under vacuum, **3** (1.937 g, 82%) as an amorphous solid; $[\alpha]_D$ = -15.3 (*c* 1.1, H_2O). ^1H NMR (D_2O): δ 7.38 and 7.07 (2 d, 4H, *J* = 9.0 Hz, Ar), 4.47 (d, 1H, $J_{1,2}$ = 10.1 Hz, H-1), 4.37 (d, 1H, $J_{1',2'}$ = 7.6 Hz, H-1'), 4.36-4.14 (m, 2H, $CH_2CH_2CH_2O$), 3.88-3.42 (m, 10H), 3.34-3.29 (m, 1H), 2.96-2.79 (m, 2H, CH_2S), 2.11-2.04 (m, 2H, $CH_2CH_2CH_2$). ^{13}C NMR (D_2O): δ 158.0 (C), 143.4 (C), 121.8 (CH), 115.6 (CH), 103.0 (CH), 85.5 (CH), 78.5 (CH), 78.2 (CH), 75.9 (CH), 75.3 (CH), 72.6 (CH), 72.2 (CH), 70.9 (CH), 68.6 (CH), 66.8 (CH_2), 61.0 (CH_2), 60.3 (CH_2), 29.3 (CH_2), 26.7 (CH_2). ^{19}F NMR (D_2O): δ 36.31. HRMS (ESI/Q-TOF) *m/z* calcd for $C_{21}H_{32}FO_{14}S_2$ $(M+H)^+$ 591.1218, found 591.1220.

Lactosyl-lysine 5. To a solution of *N*-acetyl-L-lysine methyl ester hydrochloride **4** (24 mg, 0.10 mmol) and lactose fluorosulfate **3** (118 mg, 0.20 mmol) in DMF (250 μL), partially concentrated under vacuum (ca. 0.1 mbar) to remove the traces of Me_2NH, was added Et_3N (42 μL, 0.30 mmol). The solution was kept at room temperature for 48 h and then concentrated. A solution of the residue in pyridine (1 mL) and acetic anhydride (1 mL) was kept at room temperature for 14 h and then concentrated. The residue was eluted from a column of silica gel with 1:1 cyclohexane-AcOEt, then AcOEt, to give **5** (17.5 mg, 16%) as a colorless syrup; $[\alpha]_D$ = +2.7 (*c* 0.6, $CHCl_3$). ^1H NMR ($CDCl_3$): δ 7.16-7.12 and 6.94-6.90 (2 m, 4H, Ar), 6.05 (d, 1H, *J* = 7.8 Hz, NH), 5.38 (dd, 1H, $J_{3',4'}$ = 3.4 Hz, $J_{4',5'}$ = 0.9 Hz, H-4'), 5.24 (dd, 1H, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz, H-3), 5.13 (dd, 1H, $J_{1',2'}$ = 7.8 Hz, $J_{2',3'}$ = 10.5 Hz, H-2'), 4.98 (dd, 1H, H-3'), 4.97 (dd, 1H, $J_{1,2}$ = 10.0 Hz, H-2), 4.59 (ddd, 1H, *J* = 5.2, 7.5, 7.8 Hz, C*H*NH), 4.52 (d, 1H, H-1), 4.51 (d, 1H, H-1'), 4.19-4.07 (m, 4H, 2 H-6, 2 H-6'), 4.06 (t, 2H, *J* = 6.0 Hz, $CH_2CH_2CH_2O$), 3.90 (ddd, 1H, $J_{5',6'a}$ = 6.2 Hz, $J_{5',6'b}$ = 7.6 Hz, H-5'), 3.80 (dd, 1H, $J_{4,5}$ = 9.8 Hz, H-4), 3.77 (s, 3H, OCH_3), 3.66-3.62 (m, 3H), 2.94-2.79 (m, 2H, CH_2S), 2.49, 2.18, 2.13, 2.09, 2.076, 2.071, 2.070, 2.05, and 1.99 (9 s, 27H, 9 Ac), 2.12-2.06 (m, 2H, $CH_2CH_2CH_2$), 1.87-1.53 and 1.36-1.23 (2 m, 6H). ^{13}C NMR ($CDCl_3$): δ 172.9 (C), 170.4 (C), 170.3 (C), 170.2 (C), 170.1 (C), 169.9 (C), 169.8 (C), 169.7 (C), 169.1 (C), 158.1 (C), 142.6 (C), 122.9 (CH), 115.6 (CH), 101.1 (CH), 83.8 (CH), 77.2 (CH), 76.8 (CH), 76.2 (CH), 73.7 (CH), 71.0 (CH), 70.7 (CH), 70.2 (CH), 69.1 (CH), 66.6 (CH), 66.5 (CH_2), 62.1 (CH_2), 60.8 (CH_2), 52.5 (CH_3), 52.0 (CH), 47.8 (CH_2), 31.9 (CH_2), 29.4 (CH_2), 28.2 (CH_2), 27.0 (CH_2), 24.8 (CH_3), 23.2 (CH_3), 22.2 (CH_2), 20.86 (CH_3), 20.82 (CH_3), 20.76 (CH_3), 20.67 (CH_3), 20.65 (CH_3), 20.53 (CH_3). HRMS (ESI/Q-TOF) *m/z* calcd for $C_{46}H_{65}N_2O_{25}S_2$ $(M+H)^+$ 1109.3318, found 1109.3315.

Lactosyl-Octreotide 7. To a solution of Octreotide bis-acetate **6** (8.0 mg, 7.0 μmol) and lactose fluorosulfate **3** (41 mg, 70.0 μmol) in H_2O (100 μL) and CH_3CN (50 μL) was added Et_3N (4 μL, 28.0 μmol). The solution was kept at room temperature for 72 h and then concentrated. The residue was purified by HPLC (Zorbax SB-C18 column, eluent H_2O_2-CH_3CN, gradient from 80:20 to 20:80). The pure derivative **7** (5 mg) was isolated as glassy solid in 45 % yield.

Expression and purification of ^{15}N-labeled Ubiquitin. Escherichia coli BL21(DE3) Gold cells were transformed with pET-21a(+) plasmid encoding ubiquitin gene. The cells were cultured in a 15N-labeled minimal medium supplemented with 0.1 mg mL−1 of ampicillin, grown at 310 K, until A600 reached 0.6, then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside. Cells were further grown at 310 K overnight and then harvested by centrifugation at 7500 rpm for 15 min at 277 K. The pellet was suspended in 20 mM Tris-HCl, pH 7.0 buffer (20 mL per liter of culture) supplemented with 12.5 μg mL−1 DNAse, 500 μg mL−1 lysozyme, 20 mM MgSO4 and protease inhibitors, and incubated at 277 K for 15 min upon magnetic stirring. The suspension was sonicated alternating 30 sec of sonication and 3 min of resting for 10 cycles, then centrifuged at 40000 rpm for 30 min at 277 K. The supernatant was treated with small aliquots of an HClO4 solution to adjust the pH down to 4.5. The solution was again centrifuged at 40000 rpm for 30 min at 277 K to remove the precipitate. The protein in solution was then purified by cationic-exchange chromatography using a HiPrep SP FF 16/10 column (GE Healthcare Life Science) preliminarly equilibrated with 20 mM AcONa, pH 4.5 buffer. The protein was eluted in the same buffer with a linear 0-500 mM NaCl gradient. Fractions containing pure ubiquitin were identified by Coomassie staining SDS-PAGE gels. A further purification was performed by size-exclusion chromatography using a HiLoad 16/60 Superdex 75 pep grade column (GE Healthcare Life Science). The elution was achieved in 50 mM phosphate, pH 8.0 buffer.

Expression and purification of ANSII Escherichia coli C41(DE3) cells were transformed with pET-21a(+) plasmid encoding ANSII gene. The cells were cultured in LB medium supplemented with 0.1 mg mL^{-1} of ampicillin, grown at 310 K, until A600 reached 0.6, then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside. They were further grown at 310 K overnight and then harvested by centrifugation at 7500 rpm for 15 min at 277 K. The pellet was suspended in 10 mM Tris-HCl, pH 8.0, 15 mM EDTA, 20% sucrose buffer (60 mL per liter of culture) and incubated at 277 K for 20 min upon magnetic stirring. The suspension was centrifuged at 10000 rpm for 30 min and the supernatant discarded. The recovered pellet was re-suspended in H_2O milli-Q (60 mL per liter of culture) and newly incubated at 277 K for 20 min under magnetic stirring. Again the suspension was centrifuged at 10000 rpm for 30 min. The pellet was discarded, and the supernatant treated with ammonium sulfate. Still under magnetic stirring solid ammonium sulfate was added in aliquots up to 50% saturation. The precipitate was removed by centrifugation, then further ammonium sulfate was added up to 90% saturation to trigger the precipitation of ANSII, which was recovered again by centrifugation. The precipitated ANSII was re-dissolved in a minimal amount of 20 mM Tris-HCl, pH 8.6 buffer and dialyzed extensively against the same buffer. ANSII was purified by anionic-exchange chromatography using a HiPrep Q FF 16/10 column (GE Healthcare Life Science). The protein was eluted in 20 mM Tris-HCl, pH 8.6 buffer with a linear 0-1 M NaCl gradient. Fractions containing pure ANSII were identified by Coomassie staining SDS-PAGE gels, then pooled and further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 pep grade column (GE Healthcare Life Science). The elution was performed in 50 mM phosphate, pH 8.0 buffer.

Functionalization of Ub and ANSII with lactose fluorosulfate 3. The SuFEx reaction of functionalization with **3** was achieved by mixing 800 μL of a 500 mM phosphate, pH 8.0, solution containing 6.2 mg mL^{-1} of the protein (725 μM Ub or 180 μM ANSII) with 200 μL of a 0.50 M solution of **3** in water. Each reaction was performed at 298 K for 48 hours, then the mixture was washed with 50 mM phosphate buffer, pH 8.0.

NMR measurements Solution NMR spectra of Ub and ANSII were recorded on Bruker AVANCEIII-HD NMR spectrometers, operating at 900 and 950 MHz 1H Larmor frequency, equipped with triple resonance cryo-probes, at 298 K and 310 K, respectively. Protein samples were in water buffered solution [50 mM phosphate buffer, pH 8] with a protein concentration of ~ 0.5 and 0.25 mM for Ub and ANSII, respectively. However, after consecutive reactions with lactose fluorosulfate 3, the protein concentration was slightly reduced. 1D ^1H and 2D ^1H-^{15}N HSQC NMR spectra were recorded on Ub samples after one and three SuFEx reactions; whereas only 1D ^1H NMR spectra were performed on ANSII samples after a single, two and three SuFEx reactions. All the spectra were processed with the Bruker TopSpin software package, and analyzed with the program CARA.[30] The assignment of the spectra of Ub was based on

the data reported in the Biological Magnetic Resonance Data Bank under the accession code 6457.[31]

Acknowledgements

The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy center, as well as the EC Contracts iNext No. 653706. The authors acknowledge the support of the University of Florence CERM-TT and the Recombinant Proteins JOYNLAB. The FP7 WeNMR (project No. 261572) and H2020 West-Life (project No. 675858) are also acknowledged.

Keywords: SuFEx reaction • glycosylation • Ubiquitin • Asparaginase • flurosulfate thiolactoside

[1] a) J. Marth, B. Lowe, A. Varki in *Essentials of Glycobiology*, (Eds.: A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, G. W. Hart, J. Marth), Cold Spring Harbor Laboratory Press, New York, **1999,** Chapters 8 and 17; b) P. Sears, C. H. Wong, *Cell. Mol. Life Sci*. **1998**, *54*, 223–252; c) H.-J. Gabius, S. André, H. Kaltner, H.-C. Siebert, *Biochim. Biophys. Acta* **2002**, *1572*, 165–177.

[2] a) R. Apweiler, H. Hermjakob, N. Sharon, *Biochim. Biophys. Acta* **1999**, *1473*, 4–8; b) C.-H Wong *J. Org. Chem.* **2005**, *70*, 4219-4225.

[3] N. Sethuraman, T. A. Stadheim, *Curr. Opin. Biotechnol.* **2006**, *17*, 341–346.

[4] O. Seitz, *Chembiochem* **2000**, *1*, 214–246.

[5] a) A. Varki, *Glycobiology* **1993**, *3*, 97–130; b) R. A. Dwek, *Chem. Rev.* 1996, *96*, 683-720; c) C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357–2364; d) P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, R. A. Dwek, *Science* **2001**, *291*, 2370–2376.

[6] S. Kornfeld, *J. Clin. Invest.* **1998**, *101*, 1293–1295.

[7] a) A. M. Sinclair, S. Elliott, *J Pharm Sci* **2005**, *94*, 1626–1635; b) D. T.-Y Liu, *Trends Biotechnol.*, **1992**, *10*, 114-119.

[8] P. Costantino, R. Rappuoli, F. Berti, *Expert Opin Drug Discov* **2011**, *6*, 1045–1066.

[9] R. J. Solá, K. Griebenow, *J Pharm Sci* **2009**, *98*, 1223–1245.

[10] a) B. G. Davis, *Chem. Rev.* **2002**, *102*, 579–602; b) S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105–1109; c) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131–163.

[11] S. Giuntini, E. Balducci, L. Cerofolini, E. Ravera, M. Fragai, F. Berti, C. Luchinat, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 14997–15001.

[12] a) H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2004–2021; b) H. C. Kolb, K. B. Sharpless, *Drug Discov. Today* **2003**, *8*, 1128–1137.

[13] a) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596–2599; b) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.

[14] R. K. Iha, K. L. Wooley, A. M. Nyström, D. J. Burke, M. J. Kade, C. J. Hawker, *Chem. Rev.* **2009**, *109*, 5620–5686.

[15] V. Gembus, F. Marsais, V. Levacher, *Synlett* **2008**, *2008*, 1463–1466.

[16] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 9430–9448.

[17] For a collection of references dealing with the application of SuFEx (a) or related group (b) on chemical biology, see: a) T. Guo, G. Meng, X. Zhan, Q. Yang, T. Ma, L. Xu, K. B. Sharpless, J. Dong, *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 2605–2610; b) C. Baggio, L. Gambini, P. Udompholkul, A. F. Salem, A. Aronson, A. Dona, E. Troadec, F. Pichiorri, M. Pellecchia, *J. Med. Chem.* **2018**, DOI 10.1021/acs.jmedchem.8b00810.

[18] a) A. Dondoni, A. Massi, P. Nanni, A. Roda, *Chemistry* **2009**, *15*, 11444–11449; b) M. Lo Conte, S. Staderini, A. Marra, M. Sanchez-Navarro, B. G. Davis, A. Dondoni, *Chem. Commun. (Camb.)* **2011**, *47*, 11086–11088; c) A. Dondoni, A. Marra, *Chem Soc Rev* **2012**, *41*, 573–586.

[19] M. L. Graham, *Adv. Drug Deliv. Rev.* **2003**, *55*, 1293–1302.

[20] a) E. Ravera, S. Ciambellotti, L. Cerofolini, T. Martelli, T. Kozyreva, C. Bernacchioni, S. Giuntini, M. Fragai, P. Turano, C. Luchinat, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 2446–2449; b) S. Giuntini, L. Cerofolini, E. Ravera, M. Fragai, C. Luchinat, *Sci Rep* **2017**, *7*, 17934.

[21] J. W. Marsh, J. Denis, J. C. Wriston, *J. Biol. Chem.* **1977**, *252*, 7678–7684.

[22] S. Li, D. Cohen-Karni, L. T. Beringer, C. Wu, E. Kallick, H. Edington, M. J. Passineau, S. Averick, *Polymer* **2016**, *99*, 7–12.

[23] a) Chen, W.; Dong, J.; Plate, L.; Mortenson, D. E.; Brighty, G. J.; Li, S.; Liu, Y.; Galmozzi, A.; Lee, P. S.; Hulce, J. J.; Cravatt, B. F.; Saez, E.; Powers, E. T.; Wilson, I. A.; Sharpless, K. B.; Kelly, J. W. *J. Am. Chem. Soc.* **2016**, *138*, 7353-7364. b) Mortenson, D. E.; Brighty, G. J.; Plate, L.; Bare, G.; Chen, W.; Li, S.; Wang, H.; Cravatt, B. F.; Forli, S.; Powers, E. T.; Sharpless, K. B.; Ian A. Wilson, I. A.; Kelly, J. W. *J. Am. Chem. Soc.* **2018**, *140*, 200-210.

[24] A. Dondoni, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 8995–8997.

[25] Compound **2** was prepared by reaction of 4-(allyloxy)phenol with sulfuryl fluoride (SO_2F_2) gas (see Supplementary Information) according to the method described in ref. 16. For the preparation of fluorosulfates, see also:E. Zhang, J. Tang, S. Li, P. Wu, J. E. Moses, K. B. Sharpless, *Chem. Eur. J.* **2016**, *22*, 5692–5697.

[26] N. Floyd, B. Vijayakrishnan, J. R. Koeppe, B. G. Davis, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7798–7802.

[27] A solution of **3** in D_2O containing $NaHCO_3$ remained unaltered after 96 hours at room temperature.

[28] Octreotide **6** is produced by Bioindustria, Laboratorio Italiano Medicinali, LIM (e-mail: DT@bioindustria.it) and commercialized as a drug against acromegaly.

[29] H. Chen, R. Huang, Z. Li, W. Zhu, J. Chen, Y. Zhan, B. Jiang, *Org. Biomol. Chem.* **2017**, *15*, 7339–7345.

[30] R. Keller, *The Computer Aided Resonance Assignment Tutorial (CARA)*, CANTINA Verlag, **2004**.

[31] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 6836–6837.

10.1002/chem.201803912

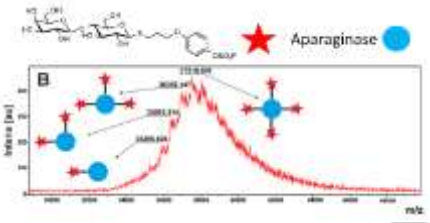
WILEY-VCH

Entry for the Table of Contents

Layout 1:

FULL PAPER

The metal free click SuFEx reaction is here employed for the coupling of a lactosyl fluorosulfate with the free ε-amino groups of lysine moieties incorporated into Ubiquitin (Ub) and into the validated drug Asparaginase II (ANSII). The glycosylation of ANSII slowing down the clearance of the enzyme in vivo, increases its half-life with a beneficial reduction of its administration frequency.



A. Marra, [b] *J. Dong,* [c] *T. Ma,* [c] *S. Giuntini,* [a],[d] *E. Crescenzo,* [a],[d] *L. Cerofolini,* [a],[d] *M. Martinucci,* [a] *C. Luchinat,* [a],[d] *M. Fragai,* *[a],[d] *C. Nativi*[a] and A. Dondoni*[e]

Page No. – Page No.

Protein Glycosylation via Sulfur Fluoride Exchange (SuFEx) Chemistry: The Key Role of a Fluorosulfate Thiolactoside