

THIOL-TARGETED MICROSPRAY MASS SPECTROMETRY OF PEPTIDES AND PROTEINS THROUGH ON-LINE EC-TAGGING

THÈSE N° 3525 (2006)

PRÉSENTÉE LE 12 MAI 2006

À LA FACULTÉ SCIENCES DE BASE

Laboratoire d'électrochimie physique et analytique

SECTION DE CHIMIE ET GÉNIE CHIMIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Loïc DAYON

Ingénieur chimiste diplômé de l'E.N.S. de Chimie de Mulhouse,
DEA de Chimie, Université de Haute-Alsace, Mulhouse, France
et de nationalité française

acceptée sur proposition du jury:

Prof. H. Vogel, président du jury
Prof. H. Girault, directeur de thèse
Prof. R. Aebersold, rapporteur
Prof. C. Amatore, rapporteur
Prof. H. Lashuel, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Lausanne, EPFL

2006

*La vérité appartient à ceux qui la cherchent
et non point à ceux qui prétendent la
détenir.*

Marie Jean Antoine Nicolas de Caritat
marquis de Condorcet (1743-1794)

REMERCIEMENTS

Je tiens à remercier chaleureusement le professeur Hubert Girault pour m'avoir donné l'opportunité et les moyens d'effectuer ce travail de thèse au sein de son laboratoire. La confiance et la sympathie qu'il m'a accordées, les conseils et les enseignements qu'il m'a prodigués m'ont été d'une aide extrêmement précieuse. Il m'a transmis, entre autres choses, la rigueur et la persévérance nécessaires au travail du chercheur. C'est ainsi que j'ai pu confirmer, pendant ces trois ans et demi de thèse, mon goût pour la recherche, initié par le docteur Christoph Boss et le professeur Jacques Eustache lors de mon diplôme d'études approfondies de chimie.

Je voudrais ensuite remercier le docteur Christophe Roussel pour ses nombreux enseignements, notamment dans le domaine de l'électrochimie, et l'intérêt général qu'il a porté à mon travail. Je suis sincèrement reconnaissant au docteur Jacques Josserand pour son soutien et ses conseils ainsi qu'au professeur Henrik Jensen pour son concours lors de ma première année de thèse. Je remercie également le docteur Tatiana Rohner pour ses recommandations et pour m'avoir transmis ses connaissances acquises dans le cadre de sa thèse, et le docteur Niels Lion pour sa disponibilité, ses conseils et ses nombreux enseignements. Je tiens à exprimer ma gratitude au professeur Anna Bratjer-Toth pour son appui, son enthousiasme ainsi que pour ses messages d'encouragement et son accueil chaleureux en Floride.

Par ailleurs, je remercie sincèrement le docteur Andrea Lionello et Michel Prudent pour leur amitié, leur sympathie ainsi que pour leur soutien. Nos discussions, d'ordre scientifique ou non, ont été des moments très agréables et j'espère qu'il y en aura beaucoup d'autres.

Je suis également reconnaissant aux membres de mon jury de thèse, les professeurs Ruedi Aebersold, Christian Amatore et Hilal Lashuel pour leurs remarques, leurs enseignements, leurs conseils et l'intérêt qu'ils ont bien voulu porter à ma démarche. Je remercie tout aussi vivement le professeur Horst Vogel, président du jury.

Je voudrais également remercier Hoang-Trang Lam pour la relecture d'un de mes chapitres de thèse, Valérie Devaud et le docteur Jean-Pierre Abid pour leur aide, et le docteur Nicolas Eugster pour ses conseils à la préparation des séances d'exercices de cinétique chimique. Je remercie Maria Szuman, Sandra Jeanneret et Claudine Bovey pour leur assistance.

Enfin, je remercie inévitablement et de tout mon cœur mes parents, Annick et Gérard, ma sœur, Claire, ma famille ainsi que Mélanie et mes amis pour leur soutien sans faille et tous les bons moments que nous avons passés à Lausanne, Paris, Montbéliard, Mulhouse, Lyon et Montréal... Ils sont pour beaucoup dans ma persévérance et l'aboutissement de ce travail. Je leur dédicace ce manuscrit.

RESUME

Les stratégies de modification spécifique d'acides aminés de protéines sont courantes en analyse protéomique. Étant données ses propriétés nucléophiles et son occurrence dans le protéome, la cystéine est une cible très intéressante.

Une source de nébulisation électrospray en spectrométrie de masse a été utilisée pour générer électrochimiquement des espèces réactives pour la cystéine. Cette technique a permis la modification en ligne de la fonction thiol de la L-cystéine par la 1,4-benzoquinone produite par électro-oxydation de la 1,4-hydroquinone présente en solution. Dans ce travail, plusieurs hydroquinones substituées ont été étudiées comme agents inductifs du marquage. Les constantes de vitesse de la réaction d'addition entre la L-cystéine et les benzoquinones substituées ont été déterminées par méthodes électrochimiques. Les valeurs déterminées sont apparues en adéquation avec la nature du substituant. Les avancements apparents de la réaction de marquage qui ont ensuite été évalués par analyse par spectrométrie de masse se sont montrés très dépendants de l'aptitude à la nébulisation des agents de marquage.

Le marquage électrochimique en ligne des cystéines avec détection par spectrométrie de masse a été étudié au niveau peptidique avec les différentes hydroquinones substituées. La méthoxycarbonyl-1,4-hydroquinone s'est révélée comme le meilleur agent de marquage puisqu'elle a permis de marquer l'ensemble des cystéines de peptides contenant jusqu'à trois cystéines. La détermination du nombre de cystéines présentes dans les peptides est possible grâce au caractère non quantitatif de la méthode. En effet, la présence simultanée du signal du peptide initial et des produits de marquage permet le comptage des cystéines *via* la détection du déplacement de masse relatif au marqueur. L'identification de l'albumine de sérum bovin et de l' α -lactalbumine humaine par empreinte massique de leurs peptides

protéolytiques s'est trouvée améliorée par comptage de leurs cystéines. Cette procédure permettant de restreindre la recherche dans les bases de données de séquence de protéines a également été appliquée à l'identification d'un mélange de quatre protéines.

Parallèlement, la nature électrolytique de la source microspray a été soulignée. L'équation de Levich a notamment été vérifiée comme première approximation des courants limites de convection-diffusion correspondants à l'oxydation des hydroquinones à l'électrode de la source de nébulisation électrospray.

Une simulation par éléments finis du marquage successif des cystéines dans les peptides a été développée. L'obtention des profils de concentration des agents de marquage, des espèces initiales et marquées le long du canal microfluidique de la source électrospray a permis de dégager les paramètres influençant le processus. Dans la mesure où le contrôle de l'avancement de la réaction est essentiel au comptage des cystéines par spectrométrie de masse, ce travail théorique a établi une gamme de conditions optimales de fonctionnement.

Le marquage électrochimique des cystéines dans les protéines a ensuite été réalisé afin de sonder leur environnement. Un modèle analytique a été développé pour le calcul rapide des avancements de marquages attendus en sortie de canal, avant nébulisation électrospray. La comparaison des résultats de marquage d'une protéine et de sa forme réduite a souligné l'influence de la structure 3D de la protéine sur l'accessibilité de l'agent marquant.

La modification chimique directe des cystéines libres par des benzoquinones a également été évaluée. Ces agents d'alkylation ont été employés pour l'isolation des peptides avec cystéines par chromatographie diagonale, l'hydrophobicité de l'agent de marquage induisant un déplacement du temps de rétention des peptides cibles.

Cette étude a montré que l'électrochimie inhérente au processus d'électrospray pouvait servir à la modification de molécules contenant des cystéines, notamment après séparation par chromatographie liquide. Des stratégies analytiques ont été développées pour tirer avantage de cette modification électrochimiquement induite.

Mots-clés : Mass spectrometry, electrospray ionization, tagging, electrochemistry, cysteine, proteomics.

ABSTRACT

Modification strategies targeting specific amino acids in proteins are widespread in proteomic analysis. Cysteine residues have received deep consideration in view of their nucleophilic properties and their occurrence in the proteome.

A recently developed micro-electrospray emitter for mass spectrometry was used to electrogenerate species reactive towards specific residues in biomolecules. When spraying L-cysteine in the presence of hydroquinone, the thiol cysteine moiety reacts *via* a 1,4-Michael addition with the benzoquinone electrochemically generated at the electrode. A series of electrogenerated selective electrophiles based on substituted benzoquinones was characterized as tags for L-cysteine. The rate constants pertaining to the addition of L-cysteine onto the benzoquinones were determined through electrochemical techniques. It was shown that the rate constants are primarily dependent on the electronic nature of the substituents. The apparent tagging extents observed for L-cysteine in microspray mass spectrometry experiments were shown to be highly dependent of the ionization efficiencies of the tag.

The on-line mass spectrometric electrochemical tagging (EC-tagging) of cysteine residues was studied for peptides. The EC-tagging was tested with the different hydroquinones on an undecapeptide containing one cysteine residue. Methoxycarbonyl-1,4-hydroquinone was shown to be the most efficient probe and revealed to be suitable to count cysteine units in peptides containing up to three cysteines. The number of cysteines corresponds to the number of characteristic mass shifts observed from the unmodified peptide. The identification of bovine serum albumin and human α -lactalbumin digest samples in a peptide mapping strategy were greatly improved by the application of the EC-tagging technique as post-column treatment. Indeed, the determination of cysteine content in the

tryptic peptides provides powerful supplementary information to the masses. The tagging method was applied to the determination of four proteins in a model mixture.

In parallel, the microspray emitter was characterized as an electrolysis flow cell for the EC-tagging of peptides. The Levich equation was validated as a first approximation for the calculation of the convection-diffusion limiting current in the device.

A finite element simulation of the multi-tagging process of peptides was developed to yield the relative distribution and concentration of tags, untagged and tagged species in the microchannel. The main chemical parameters determining the kinetics of the labelling were assessed and discussed considering the microfluidic aspects of the process. The control of the tagging extent allows the simultaneous mass spectrometric analysis of both the unmodified and of the modified peptide(s). This theoretical work has established the range of optimum conditions for the determination of the number of cysteines in peptides containing up to five cysteine groups.

The mass spectrometric EC-tagging of cysteine residues in proteins was studied to probe the cysteine environment. An analytical model was developed to calculate rapidly the tagging extent before the spray event. Experiments with unmodified proteins and their chemically reduced forms have highlighted the strong effect of the cysteine site reactivity on the tagging efficiencies. This study has shown relevant parameters for such on-line electrochemical derivatization / mass spectrometric detection strategies.

The chemical derivatization of cysteines by benzoquinone reagents was also investigated. These alkylating reagents revealed efficient for diagonal liquid chromatography to isolate cysteinyl peptides by the retention time shifts due to the hydrophobicity of the tags.

The work has demonstrated that the inherent electrochemistry of the electrospray can be employed as post column treatment to derivatize cysteinyl biomolecules. Analytical strategies have been developed to take advantage of this electrochemically-controlled modification.

Keywords: Mass spectrometry, electrospray ionization, tagging, electrochemistry, cysteine, proteomics.

LIST OF ABBREVIATIONS

AC	Alternative current
ALICE	Acid-labile isotope-coded extractants
APTA	3-acrylamidopropyltrimethylammonium
BAP	Biotin pentylamine
β -hCG	Human chorionic gonadotropin- β
BLA	β -lactoglobulin A from bovine milk
BQ	Benzoquinone compounds
BSA	Bovine serum albumin
CE	Capillary electrophoresis
cICAT	"Cleavable" isotope-coded affinity tag
CID	Collision-induced dissociation
CK	Creatine phosphokinase from rabbit muscle
COFRADIC	Combined fractional diagonal chromatography
CRM	Charged residue model
CV	Cyclic voltammetry
DC	Direct current
DEP	Diethylpyrocarbonate
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DTNB	5,5'-Dithiobis(2-nitrobenzoic acid) known as Ellman's reagent
DTT	1,4-Dithio-DL-threitol
ECE	Electrochemical-chemical-electrochemical

EC-tagging	Electrochemical tagging
EDT	1,2-Ethanedithiol
ESI	Electrospray ionization
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
GE	Gel electrophoresis
GIST	Global internal standard technique
GlaD	Guanido-labelling derivatization
HFBA	Heptafluorobutyric acid
HPLC	High performance liquid chromatography
HQ	Hydroquinone compounds
ICAT	Isotope-coded affinity tag
IEM	Ion evaporation model
IGOT	Isotope-coded glycosylation-specific tagging
IMAC	Immobilized metal affinity chromatography
iTRAQ	Isotope tags for relative and absolute protein quantification
LC	Liquid chromatography
LSV	Linear sweep voltammetry
MALDI	Matrix-assisted laser desorption ionization
MCAT	Mass-coded abundance tagging
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NBS	Nitrobenzenesulfonyl
NHS	<i>N</i> -Hydroxysuccinimide
NMR	Nuclear magnetic resonance
OMIU	<i>O</i> -methylisourea
P	L-cysteine, cysteinyl peptide or cysteinyl protein
PE	Polyethylene
PET	Polyethylene terephthalate
PhIAT	Phosphoprotein isotope-coded affinity tag
PIC	Phenylisocyanate

PQ _i	L-cysteine, cysteinyl peptide or cysteinyl protein tagged <i>i</i> times
PSD	Post source decay
PST	Protein sequence tag
PTM	Post-translational modification
QAT	Quaternary amine tag
QTOF	Quadrupole time-of-flight
QUEST	Quantification using enhanced sequence tags
RF	Radio frequency
RP	Reverse-phase
SALSA	Scoring algorithm for spectral analysis
SCE	Saturated calomel electrode
SCX	Strong cation-exchange
SELECT	Selective extraction of labelled entities by charge derivatization and tandem mass spectrometry
SEM	Scanning electron microscopy
SIM	Selected ion monitoring
SMCC	Succinimidyl 4-(<i>N</i> -maleimidomethyl)cyclohexanecarboxylate
SRM	Selected reaction monitoring
SWV	Square wave voltammetry
TBP	Tri- <i>n</i> -butylphosphine
TFA	Trifluoroacetic acid
TMS	Tetramethylsilane
TNBS	2,4,6-Trinitrobenzenesulfonic acid
TOF	Time-of-flight
TQ	Triple-quadrupole
UV	Ultra-violet
VICAT	Visible isotope-coded affinity tag
XPS	X-ray photoelectron spectroscopy

LIST OF SYMBOLS

$2h$	Channel height	m
A	Electrode surface	m^2
$\alpha\text{-TE}\%$	Apparent tagging extent	0.01
C	Concentration	$\text{mol}\cdot\text{m}^{-3}$ and $\text{M} = \text{mol}\cdot\text{L}^{-1}$
C_i	Concentration of the species i	$\text{mol}\cdot\text{m}^{-3}$ and $\text{M} = \text{mol}\cdot\text{L}^{-1}$
d	Channel width	m
D	Diffusion coefficient	$\text{m}^2\cdot\text{s}^{-1}$
D_i	Diffusion coefficient of the species i	$\text{m}^2\cdot\text{s}^{-1}$
δ	Chemical shift	ppm
δ	Mesh size	m
ΔE_p	$ E_{pa} - E_{pc} $ in CV	V
ΔE_s	Potential step	V
ΔE_{SW}	Step amplitude in SWV	V
E_{pa}^0	Oxidation peak potential	V
E_{pc}^0	Reduction peak potential	V
$\alpha\text{-TE}\%$	Effective tagging extent	0.01
F	Faraday constant	$96485 \text{ C}\cdot\text{mol}^{-1}$
F_V	Flow rate	$\text{m}^3\cdot\text{s}^{-1}$
I or i	Current	A
I_{lim}	Convection-diffusion current	A
I_p	Oxidation peak current for an ECE mechanism	A

I_p^0	Oxidation peak current for an E mechanism	A
I_{spray}	Electrospray current	A
k	Rate constant for a reaction with an overall order n	$(\text{M}^{-1})^{n-1} \cdot \text{s}^{-1}$
k_{cys}	Rate constant for a reaction with an overall order n implying L-cysteine	$(\text{M}^{-1})^{n-1} \cdot \text{s}^{-1}$
k_i	Rate constant for a i consecutive reaction with an overall order n	$(\text{M}^{-1})^{n-1} \cdot \text{s}^{-1}$
k_{ox}	Oxidation rate constant	$\text{m} \cdot \text{s}^{-1}$
k_p	Rate constant for a reaction with an overall order n implying a given cysteine in a protein	$(\text{M}^{-1})^{n-1} \cdot \text{s}^{-1}$
k_{red}	Reduction rate constant	$\text{m} \cdot \text{s}^{-1}$
k_{true}	Real rate constant for a reaction with an overall order n	$(\text{M}^{-1})^{n-1} \cdot \text{s}^{-1}$
L	Electrode length	m
l	Electrode width	m
λ_{max}	Maximum absorption wavelength	m
L_{ch}	Channel length	m
m	Mass	kg and Da
M	Molar mass	$\text{g} \cdot \text{mol}^{-1}$
m/z	Mass-to-charge ratio	Th
n	Number of consecutive reactions	1
N	Number of electrodes	1
n_{BQ}	Chemical amount of electrogenerated benzoquinone BQ	mol
N_{BQ}	Flux rate of electrogenerated benzoquinone BQ	$\text{mol} \cdot \text{s}^{-1}$
N_p	Flux rate of incoming peptide or protein P	$\text{mol} \cdot \text{s}^{-1}$
pH	pH	1
pI	Isoelectric point	1
$\text{p}K_a$	$\text{p}K_a$	1
$\text{p}K_{a_{\text{cys}}}$	$\text{p}K_a$ of the thiol function of L-cysteine	1
$\text{p}K_{a_p}$	$\text{p}K_a$ of the thiol function of a given cysteine in a protein	1
P_s	Mesh Péclet number	1

R_e	Reynolds number	1
R_i	Rate of generation or consumption of i	$\text{M}\cdot\text{s}^{-1}$
σ_p	Hammet constant (para-substituent)	1
t	Time	s
t_e	Residence time of the species i on the top of the electrode	s
TE	Tagging extent	1
$TE\%$	Tagging extent	0.01
t_R	Retention time	s
\mathbf{v}	Fluid velocity vector	$\text{m}\cdot\text{s}^{-1}$
v	Fluid velocity	$\text{m}\cdot\text{s}^{-1}$
\bar{v}	Mean flow velocity	$\text{m}\cdot\text{s}^{-1}$
v	Rate of reaction	$\text{M}\cdot\text{s}^{-1}$
V	Volume	m^{-3}
V_e	Volume element over the electrode	m^{-3}
v_{\max}	Maximum fluid velocity	$\text{m}\cdot\text{s}^{-1}$
x	Concentration evolution	M
x, X	Cartesian space coordinate	m
y, Y	Cartesian space coordinate	m
$Yield_{\text{MS}}$	Mass spectrometric tagging yield	0.01
$Yield_{\text{true}}$	Bulk tagging yield	0.01
z	Stoichiometric number of electrons involved in an electrode reaction	1
z, Z	Cartesian space coordinate	m

TABLE OF CONTENTS

CHAPTER I. Introduction	1
1. Tagging methods are everywhere!	1
2. A propitious field to tagging methods	2
2.1. Mass spectrometry in proteomics	3
2.2. Proteomic analytical strategies	5
2.3. Tagging methods for MS-based proteomics	5
3. Taking advantage of the electrospray process	6
3.1. Electrospray ionization	6
3.2. Electrospray sources for mass spectrometry	8
3.3. On-line electrochemical tagging: state of the art	9
4. Objective of the work	10
Bibliography	12
CHAPTER II. Chemical modifications for targeted proteomics: mass spectrometry-based strategies	15
1. Introduction	15
2. Specific amino acid targets	16
2.1. Cysteine modification	16
2.2. Lysine modification	24
2.3. Arginine modification	27
2.4. Tryptophan modification	28
2.5. Methionine modification	29
2.6. Histidine modification	30

3. N- and C-terminal targets.....	31
3.1.N-terminal labelling.....	31
3.2.C-terminal labelling.....	38
4. Derivatization of post-translational modification sites.....	39
4.1.Serine and threonine phosphorylation	39
4.2.Glycolysation	43
4.3.Tyrosine nitration	45
5. Conclusion.....	45
Bibliography	47
CHAPTER III. Hydroquinone probes for on-line electrochemical tagging of cysteine residues: mechanism and kinetics.....	53
1. Introduction	53
2. Experimental	55
2.1. Reagents	55
2.2. Chemical and electrochemical set-up	56
2.3. Synthesis	57
2.4. Digital simulations.....	58
3. Results and discussions.....	59
3.1. Electrochemical induced tagging of L-cysteine	59
3.2. ¹ H NMR study of the addition on benzoquinone	65
3.3. On-line electrochemical tagging of L-cysteine during mass spectrometry .	67
3.4. On-line electrochemical tagging of β -lactoglobulin A during ESI-MS	69
4. Conclusion.....	73
Bibliography	74
CHAPTER IV. On-line counting of cysteine residues in peptides during electrospray ionization by electrogenerated tags and their application to protein identification.....	77
1. Introduction	77

2. Experimental.....	80
2.1. Chemicals.....	80
2.2. Microchip device.....	80
2.3. MS-set-up.....	81
2.4. Digestion and peptide analysis.....	81
2.5. Protein identification	82
3. Results and discussions.....	82
3.1. Tagging of chorionic gonadotropin (109-119) amide	82
3.2. Tagging of peptides containing several cysteines	84
3.3. Tandem mass spectrometry of peptide-quinone adducts.....	89
3.4. Mass fingerprinting of bovine serum albumin and human α -lactalbumin using on-line electrochemical tagging.....	93
3.5. Mass fingerprinting of a mixture of 4 proteins using on-line electrochemical tagging	97
4. Conclusion.....	98
Bibliography	100
Annexe: peptide masses found for BSA, peptide masses found for human α - lactalbumin and peptide masses found for a mixture of BSA, human α -lactalbumin, ovalbumin and β -lactoglobulin from bovine milk.....	102
CHAPTER V. Characterization of polyethylene terephthalate microspray emitter for mass spectrometric EC-tagging.....	105
1. Introduction	105
2. Experimental	106
2.1. Chemicals.....	106
2.2. Microchip device.....	107
2.3. Electrochemical set-up.....	107
2.4. Mass spectrometry.....	108

3. Results and discussions.....	109
3.1. Characterization.....	109
3.2. Microelectrode-array	115
4. Conclusion.....	120
Bibliography	122
CHAPTER VI. Electrochemical tagging of peptides in a microchannel during microspray mass spectrometry: numerical simulation of consecutive reactions with multiple cysteines	125
1. Introduction	125
2. Computational methods	127
2.1. Numerical technique.....	127
2.2. Scaling	128
3. Results and discussions.....	130
3.1. Microspray characteristic.....	130
3.2. Numerical model	132
3.2.1. Model.....	132
3.2.2. Assumptions.....	133
3.3. Kinetics of multi-tagging	134
3.4. Optimization of multi tagging.....	139
4. Conclusions	142
Bibliography	144
Annexe: finite-element formulation.....	146
CHAPTER VII. Probing cysteine in proteins by mass spectrometric EC-tagging.....	149
1. Introduction	149
2. Experimental	151
2.1. Maple calculation	151
2.2. Materials	151
2.3. Reduction of proteins.....	151

2.4. EC-tagging	151
2.5. Chemical labelling and ionization experiments	152
2.6. Micro-fabricated ESI emitter	152
2.7. MS set-up	152
2.8. Diffusion coefficient measurements	153
3. Results and discussions	153
3.1. Mechanism and kinetics	153
3.1.1. Analytical kinetic model	154
3.1.2. Numerical validation of the model	157
3.1.3. Experimental validation of the model	158
3.2. EC-tagging of proteins	159
3.2.1. Ionization effects	161
3.2.2. Variations of pK_a and steric hindrance	162
3.3. Multi-EC-tagging of reduced β -lactoglobulin A	163
4. Conclusion	167
Bibliography	169
Annexe: Maple programme for calculating five-cysteine-containing protein	172
CHAPTER VIII. Chemical modification with benzoquinone tags	173
1. Introduction	173
2. Experimental	175
2.1. Chemicals	175
2.2. Derivatization protocol	176
2.3. Mass spectrometry	176
2.4. High performance liquid chromatography	177
3. Results and discussions	177
3.1. Benzoquinone reagents: reactivity and selectivity	177
3.2. Modification of peptides: application to diagonal chromatography	181

3.3. Modification of proteins: counting of cysteine residues.....	185
4. Conclusion.....	187
Bibliography	196
CHAPTER IX. Conclusions and perspectives.....	191
APPENDIX I.....	197
APPENDIX II	203
APPENDIX III.....	205
APPENDIX IV	207
CURRICULUM VITAE.....	209

CHAPTER I. *Introduction*

1. Tagging methods are everywhere!

Nowadays, tagging methods are everywhere. The term “tag” appeared in the 15th century to design a “small hanging piece” but the meaning “label” was first recorded in 1835. When searching for the terms “tagging”, “tag”, “tags” or “tagged”, you find many fields associated. Medicine, biology, biochemistry and chemistry use tagging techniques as well as linguistics. Recently, tagging methods have appeared to index and assign material on the Web, also known sometimes as “folksonomies” (a portmanteau of “folk” and “taxonomy”). The electronic tagging of a person or a vehicle by attachment of an electronic device has been introduced in 2003 in the United Kingdom to monitor the whereabouts of offenders.

In linguistics, part of speech tagging is the process of marking the words in a text with their corresponding parts of the speech. The common application is met in the early years of school when pupils identify nouns, verbs, adjectives and so on. However, it refers generally to computer algorithms to do much the same, *i.e.* the computer analysis of languages. Indeed, many sub-classes of parts of the speech exist and the word-forms can be ambiguous.¹ In biology, tagging of animals with more or less sophisticated tags including satellite tags permits to follow the displacement of species. As an example, the tagging of Pacific pelagics is a research project that studies the journey across the open ocean of cetaceans, turtles, sharks, tunas and salmon.^{2,3} In medicine, nuclear radiology uses very small amounts of radioactive materials or radiopharmaceuticals that serves as tags to examine organ function and structure. Such a technology helps to diagnose and treat abnormalities of a disease such as thyroid cancer. For instance, positron emission tomography helps to visualize the metabolism of the heart muscle, *i.e.* the biochemical changes taking place in the body.^{4,5}

In biochemistry, and more particularly in the growing field of proteomics, tags can help for purification, detection and quantification of proteins and peptides. Many chemical tagging strategies have been carried out as shown by the large number of publications related. The work described in this dissertation assesses a new tagging technique based on the electrochemically-induced generation of tags that are specific for cysteine moieties in proteins and peptides during mass spectrometry (MS).

2. A propitious field to tagging methods

J.J. Berzelius first mentioned the word protein (from the Greek *prôtos*, first) in 1838 in a letter to G.J. Mulder. In 1839, G.J. Mulder coined the term in an article of the *Journal für praktische Chemie* to design “ the organic substance which is present in all constituents of the animal body, also in the plant kingdom”.⁶ However, this “substance” was known for thousands years since in ancient Egypt (3rd Dynasty, *c.a.* 2600 B.C.), glue (*i.e.* gelatine from bones and skins) was used as adhesive for furniture. Proteins are one of the classes of biological macromolecules, as nucleic acids, lipids and polysaccharides that are the primary constituents of living entity. They are composed of amino acids joined by peptide bonds. The term proteome, which is a dynamic entity, was coined in 1994 and was defined as the protein complement of the genome. Studying the proteome became known as proteomics. It now evokes not only all the proteins in a given cell, but also the set of all the protein isoforms and modifications, the interaction between them, the structural description of proteins and their higher-order complexes⁷, hence the three following domains⁸:

Profiling proteomics: Profiling proteomics is related to the identification of proteins present in a biological sample or of proteins that have differential expression state between different samples.

Functional proteomics: Functional proteomics focuses on the determination of protein functions as regards the presence of specific groups and the study of protein-ligand interaction.

Structural proteomics: Structural proteomics deals with the knowledge of 3D protein structure and the elucidation of interacting amino acids in protein complexes that are primordial for the understanding of protein function in the cell. X-ray crystallography and computational modelling are ones of its main methods.

Proteomics must deal with issues such as limited and variable sample material, sample degradation, vast dynamic range and post-translational modification. However, proteomics is set to have a profound impact on the clinical diagnosis and drug discovery through the understanding of disease processes, the discovery and development of new biomarkers for diagnosis and early detection of disease, and the accelerated drug development by effective strategies to evaluate therapeutic effect and toxicity.⁹

2.1. Mass spectrometry in proteomics

MS has become the method of choice for the detection of proteins from complex samples.¹⁰ The introduction of effective laser desorption ionization time-of-flight (MALDI-TOF) MS and electrospray ionization (ESI) tandem mass spectrometry (MS/MS) have enabled this evolution. In addition, the concept of correlating MS data with genomic sequence databases has revealed to be perfectly suited to identify proteins.

To provide good quality proteomic analysis, mass analysers have to present several features, explicitly mass accuracy, sensitivity, resolution and capability to generate tandem mass spectra. TOF analysers are commonly coupled to MALDI, which allows the transfer of analytes from a crystalline matrix to the gas phase. TOF analysers exploit the distinct velocities that charged ions of different masses have when accelerated under an electric field. Accelerated ions flying through a field-free tube under *vacuum* hit the detector at different time intervals according to their velocities, *i.e.* their m/z . The resolution of the TOF analyser was improved for instance by the development of the reflectron that now enable mass accuracy down to 10 to 50 ppm and 5000 to 15000 resolution. In MALDI-TOF, the fragmentation of molecules is achieved by post source decay (PSD) that consists in providing excess of laser energy to the sample. However such fragmentations are often poor quality. A hybrid instrument, the quadrupole time-of-flight (QTOF) apparatus, has been developed to

overcome this drawback by combination of quadrupoles to perform collision-induced dissociation (CID) before the TOF analysis.^{11,12}

Several analysers are commonly combined to ESI, namely the triple-quadrupole (TQ), the ion trap and the Fourier transform ion cyclotron resonance (FTICR). The TQ is composed of three consecutive sets of quadrupole. A quadrupole consists of four rods, ideally with a parabolic section, that are used to generate an electric field. By combination of a direct and an alternative current (DC and AC), the quadrupole filters specific m/z . When applying only an AC, it acts as an ion guide. In the TQ, the first and third quadrupoles serve either as mass filter or ion guide. The second quadrupole is set for ion guidance and CID. It is also possible to generate simple mass spectra as well as tandem mass spectra to identify peptides and proteins. The TQ presents low resolution and mass discrimination. Although it was the first analyser to have been used in proteomics, more sensitive analysers such as the ion on trap have supplanted it.¹³

Ion trap analysers utilize a combination of electrodes to accumulate ions in a space defined by the electric field. The ion trap can be seen like a cylindrical quadrupole. The captured ions are accumulated in the trap where they describe 3D orbits, cutting figures of eight. Once enough ions are accumulated, they are ejected from the trap by changing the radio frequency (RF) amplitude, which is related to m/z . MS or MS/MS can be performed. Ion trap mass spectrometers are robust, sensitive and relative low cost. They allow mass accuracy from 50 to 300 ppm and 2000 resolution.¹⁴

FTICR-MS can achieve ultra-high resolution (> 1000000) and mass accuracy (< 5 ppm). Ions are injected from the source into a container called the “cell” which is located in a strong magnetic field. The axis of the field is perpendicular to the velocities of the ions, resulting in a cyclotron motion of the ions submitted to the Lorentz force. The circular orbits of ions are then increased by application of a RF potential that in addition focuses them in packets. Once the ions are trapped on a stable orbit, their detection occurs when passing by two detector plates, inducing small current called the image current. After amplification, the induced current that is in fact composed of all the currents induced by the ions (at respective frequencies related to their m/z) is handled by FT analysis to obtain the mass spectrum.¹⁵

2.2. Proteomic analytical strategies

The study of proteins through MS can rely on diverse analytical strategies. In top-down proteome characterization strategies, individual proteins are selected for direct analysis. At the contrary, the bottom-up analysis consists of cleaving the protein into peptides (from the Greek *pepsis*, digestion) fragments by chemical or enzymatic treatment. For protein identification, two bottom-up methods are generally used. The first one is based on the mass measurements of all the peptides fragments whereas the second relies on the MS fragmentation of one or several of the proteolytic peptides. Peptide-mass mapping or mass fingerprinting qualified the first of these approaches since the identification is made on a set of peptide masses that are characteristics and constitutes a “fingerprint” of the original cleaved protein. Peptide sequencing is used to name the second approach that relies on the information given by tandem mass spectra provided by ESI-MS/MS experiments. Intact protein analyses are generally less effective for protein identification than measurements at the peptide level. However, whereas bottom-up approach suffers from incomplete sequence coverage, the top-down approach brings a relevant solution to the examination of site-specific mutations and post-translational modifications of proteins.¹⁶

Historically, 2D gel electrophoresis is a powerful technique and has been widely used and combined with MS detection. Nevertheless, it suffers from the lacks of capability for studying low abundance proteins and hydrophobic proteins that easily precipitate at their isoelectric point (pI). The gel-free analysis of complex mixture is a combination of one- or two-dimensional liquid chromatography (LC) with MS and MS/MS detection. Such a technique has the advantage of reducing sample lost and contamination. Lower abundance proteins are therefore easily studied. The on-line coupling of liquid chromatography with ESI-MS renders the approach convenient for automation.¹³

2.3. Tagging methods for MS-based proteomics

Mass spectrometers are poor quantitative devices because of the different ionization efficiencies of the analytes of a more or less complex sample. The relationship between the amount of analyte and the measured signal intensity in MS is consequently complex and incompletely understood. The differential analysis of complex samples in a gel-

free bottom-up approach can be achieved by isotope tagging of peptides that allows the quantification of changes in peptide amounts between two extracts by MS. The isotope-coded affinity tagging¹⁷ and the global N-terminal isotope coding^{18,19} are two examples of the use of such tags.

In addition, tagging methods in proteomics have many other applications. They can serve for separation, isolation, enrichment and detection. A short overview of these methods is presented in the chapter II of the dissertation.

3. Taking advantage of the electrospray ionization process...

3.1. Electrospray ionization

The phenomenon of ESI has been known about for 250 years since the works of a French clergyman physicist named J.-A. Nollet. For instance, he observed that a person electrified by connection to a high voltage generator would not bleed normally if he were cut himself and that the blood would spray from the wound. In 1968, M. Dole *et al.* described the idea of electrospraying an analyte solution in MS.^{20,21} J.B. Fenn and co-workers demonstrated the potential of this idea and developed ESI as a true interface for MS²² to study large and fragile biomolecules.^{23,24}

In positive ionization mode ESI-MS, a solution of the analytes is passed through a capillary, which is held at a high potential. The effect of the high electric field is to engender a spray of highly charged droplets, which pass down a potential gradient and pressure gradient towards the mass spectrometer. During the transport, the droplets reduce in size by evaporation of the solvent or by Coulomb explosion, resulting in the production of fully desolvated ions.

It is convenient to divide the electrospray process into three stages: droplet formation, droplet shrinkage and gaseous ion formation. Assuming a positive potential, positive ions in solution accumulate at the surface to establish a Taylor cone (Figure 1). At a high enough imposed field, the cone is drawn to a thread which produced positively charged

droplets when the surface tension is exceeded by the applied electrostatic force. The evaporation and fission of the droplets depends on the initial size and charge of the initial droplets produced by ESI.²⁵

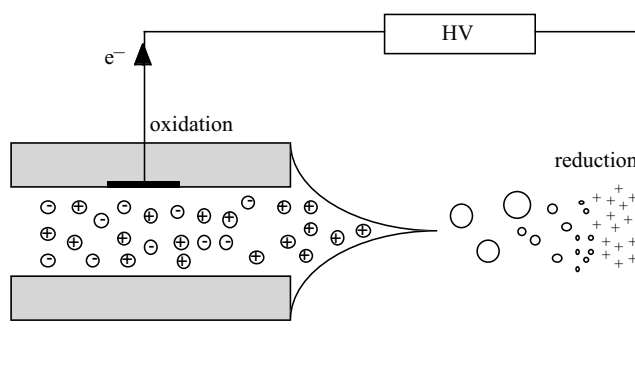


Figure 1. Electrospray ionization process in positive ionization mode.

On the second stage, a generation of offspring droplets occurs. To explain this process, two theories have been proposed. M. Dole considered that there is a succession of solvent evaporations and Coulomb explosions.²⁰ Fission (Coulomb explosion) occurs at the point at which the magnitude of the charge is sufficient to overcome the surface tension holding the droplets together. The condition is expressed by the Raleigh's stability limit:

$$q_R = 8\pi(\epsilon_0\gamma R_R^3)^{1/2} \quad (\text{Equation 1})$$

where R_R and q_R are respectively the radius and the charge of the droplet, ϵ_0 is the permittivity of the *vacuum* and γ is the surface tension of the solvent. The so-called charged residue model (CRM) holds that the sequence of Raleigh instabilities followed by periods of evaporation produces ultimate droplets, each of which contains only one solute. In the last stage, that molecule becomes a free gas phase ion by retaining some of its droplet's charge as the last of its solvent evaporates.

A second mechanism for gas-phase ion production has been described by J.V. Iribarne and B.A. Thomson.²⁶ The model is called the ion evaporation model (IEM). The formation of single-ion droplets occurs at the surface of the emitted droplets. An offspring

droplet made up of one-solvated ions is pulled out of the initial droplet by the repulsions between the charged ion and the other charges of the droplet. Then, the evaporation occurs, resulting in gas phase single-ions.

In positive ionization mode, positive droplet emission continuously carries off positive ions. The requirement for charge balance and the fact that only electrons can flow through the metal wire supplying the electric potential lead to the conclusion that the ESI process involves electrochemistry.²⁷ In other words, the ESI device can be viewed as an electrolysis cell. An electrochemical oxidation occurs at the positive electrode (*i.e.* at the liquid-metal interface of the capillary). This reaction supplies positive ions to the solution by oxidation of water in acidic media ($2 \text{H}_2\text{O} \rightarrow \text{O}_2 (\text{g}) + 4 \text{H}^+ + 4 \text{e}^-$) for instance. Conversion of atoms from the metal electrode to positive metal ions can also take place.

3.2. Electrospray sources for mass spectrometry

It is clear that optimal sensitivity of detection using electrospray is achieved by delivering reduced flow rates and several laboratories have devoted efforts in this direction. An electrospray capillary with an orifice diameter of 1-2 μm is estimated to yield droplets with diameters of $< 200 \text{ nm}$.^{28,29} The efficiency of conversion of condensed-phase analyte is superior than conventional-scale sources.³⁰ The low flow rate has the major advantage to increase the sampling time considerably. Thus, instrumental parameters, such as the collision energy for MS/MS experiments of complex peptide mixtures, can be optimized.²⁹

Microspray (100 to 500 $\text{nL}\cdot\text{min}^{-1}$) and nanospray (1 to 100 $\text{nL}\cdot\text{min}^{-1}$) emitters have been fabricated with glass and silica capillaries, which were interfaced with microfluidic chips.³¹⁻³³ In parallel, microtechnologies have offered many options to fabricate emitters based on various materials like polyimide³⁴, polydimethylsiloxane³⁵, polycarbonate³⁶, silicon³⁷, silicon dioxide³⁸ and parylene.³⁹ Diverse sizes and shapes that affect the spray generation were patterned, as the recent nib like design.⁴⁰

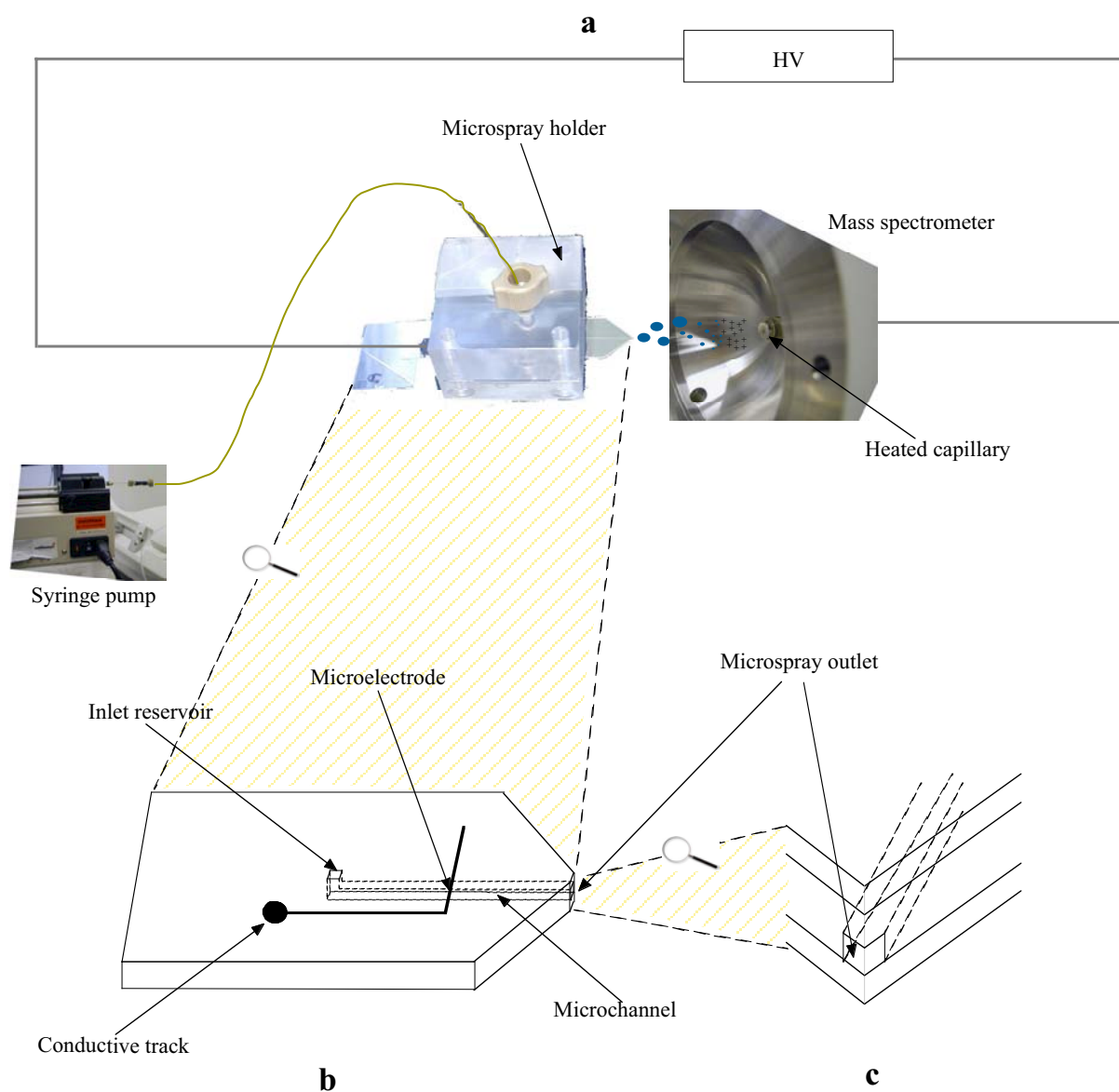


Figure 2. MS set-up comprising the microspray emitter placed in front of the entrance of the mass spectrometer (a). Scheme of the microspray emitter (b). A zoom caption of the microspray outlet is depicted (c).

3.3. On-line electrochemical tagging: state of the art

H.H. Girault's group has developed a polyethylene terephthalate (PET) microspray emitter with an integrated microband electrode (Figure 2; see manuscript appendix I for detailed description of the fabrication).⁴¹ The on-line electrochemical tagging of cysteines has been developed using this microspray emitter. Taking advantage of the

electrolytic nature of ESI, the microspray device was used to electrogenerate tags for thiol groups. In a positive ionization mode, compounds such as 1,4-benzoquinone can be generated at the electrode from 1,4-hydroquinone to react *via* a 1,4-Michael addition with free cysteines in proteins.⁴² It was shown that the geometrical and operational parameters of the interface play on the tagging degree.⁴³ The electrochemical mechanism and the rate constant of the addition of L-cysteine onto 1,4-benzoquinone were determined.⁴⁴

4. Objective of the work

This work deals with the development of the on-line electrochemical tagging (EC-tagging). Several improvements of the technique are presented and application to protein studies is carried out. Capabilities and limitations of the EC-tagging are discussed.

Chapter II is a short overview of the tagging methods currently used in proteomics. It highlights the need of such methods for quantification and separation purposes.

The improvement of the EC-tagging efficiency is assessed in chapter III using substituted 1,4-hydroquinones. The rate constant of the reactions with L-cysteine is determined through electrochemical methods. The EC-tagging efficiency on L-cysteine is evaluated in MS.

In chapter IV, the mass spectrometric EC-tagging of peptides is carried out. The EC-tagging of peptides was applied to the identification of model proteins through the counting of cysteines units in their tryptic peptides by MS.

Chapter V is dedicated to the characterization of the PET emitter for EC-tagging purposes. The influence of geometric and physic parameters of the emitters is discussed.

A finite-element model of the EC-tagging is presented in chapter VI. The study describes the consecutive tagging reactions that happen in the microchannel and assesses the theoretical capabilities and limitations of the EC-tagging method.

Chapter VII is devoted to the EC-tagging of free cysteines in intact protein level. The method is shown to provide information on the cysteine environment by measurement of the EC-tagging extent through MS measurements.

In chapter VIII, the use of benzoquinone as modification reagents of cysteines is evaluated. Preliminary results on the use of such tagging reagents in diagonal liquid chromatography are presented.

For clarity, the chapters of the thesis are written in an independent manner. Experimental parts are for instance sometimes redundant from a chapter to another to avoid fastidious research through the manuscript.

Bibliography

- 1 L. Tanabe and W.J. Wilbur, *Bioinformatics*, 2002, 18, 1124-1132.
- 2 B.A. Block, S.L.H. Teo, A. Walli, A. Boustany, M.J.W. Stokesbury, C.J. Farwell, K.C. Weng, H. Dewar and T.D. Williams, *Nature*, 2005, 434, 1121-1127.
- 3 K.C. Weng, P.C. Castilho, J.M. Morrisette, A.M. Landeira-Fernandez, D.B. Holts, R.J. Schallert, K.J. Goldman and B.A. Block, *Science*, 2005, 310, 104-106.
- 4 S.S. Gambhir, *Nat. Rev. Cancer*, 2002, 2, 683-693.
- 5 C.Y.J. Shu, S.L. Guo, Y.J. Kim, S.M. Shelly, A. Nijagal, P. Ray, S.S. Gambhir, C.G. Radu and O.N. Witte, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, 102, 17412-17417.
- 6 G.J. Mulder, *J. Prakt. Chem.*, 1839, 16, 129-152.
- 7 M. Tyers and M. Mann, *Nature*, 2003, 422, 193-197.
- 8 D. Figeys, *Anal. Chem.*, 2003, 75, 2891-2905.
- 9 S. Hanash, *Nature*, 2003, 422, 226-232.
- 10 R. Aebersold and M. Mann, *Nature*, 2003, 422, 198-207.
- 11 A.V. Loboda, A.N. Krutchinsky, M. Bromirski, W. Ens and K.G. Standing, *Rapid Commun. Mass Spectrom.*, 2000, 14, 1047-1057.
- 12 A. Shevchenko, A. Loboda, A. Shevchenko, W. Ens and K.G. Standing, *Anal. Chem.*, 2000, 72, 2132-2141.
- 13 D. Figeys, *Industrial Proteomics: Application for Biotechnology and Pharmaceuticals*, 1st ed., John Wiley & Sons, New York, 2005, pp. 1-62.
- 14 E. de Hoffmann, J. Charette and V. Stroobant, *Spectrométrie de Masse: Cours et Exercices*, Dunod, Paris, 1999.
- 15 M.P. Barrow, W.I. Burkitt and P.J. Derrick, *Analyst*, 2005, 130, 18-28.
- 16 B. Bogdanov and R.D. Smith, *Mass Spectrom. Rev.*, 2005, 24, 168-200.
- 17 S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, 17, 994-999.
- 18 M. Munchbach, M. Quadroni, G. Miotto and P. James, *Anal. Chem.*, 2000, 72, 4047-4057.
- 19 A. Chakraborty and F.E. Regnier, *J. Chromatogr. A*, 2002, 949, 173-184.

- 20 M. Dole, L.L. Mack and R.L. Hines, *J. Chem. Phys.*, 1968, 49, 2240-2249.
- 21 L.L. Mack, P. Kralik, A. Rheude and M. Dole, *J. Chem. Phys.*, 1970, 52, 4977-4986.
- 22 M. Yamashita and J.B. Fenn, *J. Phys. Chem.*, 1984, 88, 4451-4459.
- 23 J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong and C.M. Whitehouse, *Science*, 1989, 246, 64-71.
- 24 J.B. Fenn, *Angew. Chem.-Int. Edit.*, 2003, 42, 3871-3894.
- 25 J.F. de la Mora and I.G. Loscertales, *J. Fluid Mech.*, 1994, 260, 155-184.
- 26 J.V. Iribarne and B.A. Thomson, *J. Chem. Phys.*, 1976, 64, 2287-2294.
- 27 G. Diehl and U. Karst, *Anal. Bioanal. Chem.*, 2002, 373, 390-398.
- 28 M.S. Wilm and M. Mann, *Int. J. Mass Spectrom. Ion Process.*, 1994, 136, 167-180.
- 29 M. Wilm and M. Mann, *Anal. Chem.*, 1996, 68, 1-8.
- 30 R. Juraschek, T. Dulcks and M. Karas, *J. Am. Soc. Mass Spectrom.*, 1999, 10, 300-308.
- 31 D. Figeys, Y.B. Ning and R. Aebersold, *Anal. Chem.*, 1997, 69, 3153-3160.
- 32 I.M. Lazar, R.S. Ramsey, S. Sundberg and J.M. Ramsey, *Anal. Chem.*, 1999, 71, 3627-3631.
- 33 J.J. Li, T.L. Tremblay, C. Wang, S. Attiya, D.J. Harrison and P. Thibault, *Proteomics*, 2001, 1, 975-986.
- 34 V. Gobry, J. van Oostrum, M. Martinelli, T.C. Rohner, F. Reymond, J.S. Rossier and H.H. Girault, *Proteomics*, 2002, 2, 405-412.
- 35 J.-S. Kim and D.R. Knapp, *Electrophoresis*, 2001, 22, 3993-3999.
- 36 K.Q. Tang, Y.H. Lin, D.W. Matson, T. Kim and R.D. Smith, *Anal. Chem.*, 2001, 73, 1658-1663.
- 37 G.A. Schultz, T.N. Corso, S.J. Prosser and S. Zhang, *Anal. Chem.*, 2000, 72, 4058-4063.
- 38 J. Sjødahl, J. Melin, P. Griss, A. Emmer, G. Stemme and J. Roeraade, *Rapid Commun. Mass Spectrom.*, 2003, 17, 337-341.
- 39 L. Licklider, X.Q. Wang, A. Desai, Y.C. Tai and T.D. Lee, *Anal. Chem.*, 2000, 72, 367-375.
- 40 S. Le Gac, C. Cren-Olive, C. Rolando and S. Arscott, *J. Am. Soc. Mass Spectrom.*, 2004, 15, 409-412.

- 41 T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.
- 42 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 43 T.C. Rohner, J. Josserand, H. Jensen and H.H. Girault, *Anal. Chem.*, 2003, 75, 2065-2074.
- 44 C. Roussel, T.C. Rohner, H. Jensen and H.H. Girault, *ChemPhysChem*, 2003, 4, 200-206.

CHAPTER II. *Chemical modifications for targeted proteomics: mass spectrometry-based strategies*

1. Introduction

Although many efforts have been made to develop genomic technologies, the study of proteins remains an essential pathway to comprehend biological processes. Indeed, the gene expression does not correlate univocally with the protein expression. Indeed, post-translational modification (PTM) and truncation of proteins are few examples of the complexity encountered at the protein level. In a decade, the proteomic field has developed widely thanks to the concomitant development of several technical areas. In particular, mass spectrometry (MS) and the development of soft ionization techniques have enabled the setting up of platforms to provide large-scale, rapid and sensitive detection of both peptides and proteins. However, due to the tremendous complexity of a proteome, sample preparation before the final analysis is recommended not to exceed the capacity of the analytical systems.

The targeting of specific structural features is a practical approach to reduce this complexity. Chemical tagging methods have been carried out to target specific amino acids or PTM in proteins and peptides to facilitate the proteome analysis while preserving the information requisite for the exactness of the analysis. On the one hand, the chemical tagging strategies enable specific selection by affinity clean up for instance; they bring, on the other hand, implements for the mass spectrometric detection, which is one of the central detection techniques in proteomics. In that sense, these tags are so-called “mass tags” since they serve for the relative quantification levels by comparative stable-isotope labelling or for the

modification of the ionization efficiency and subsequent fragmentation behaviour by MS and tandem MS (MS/MS).

The twenty proteinogenic amino acids present limited choice of functional groups to label. The nucleophilic character of the thiol group of cysteine and the free amino group of lysine are consequently the target of many derivatization protocols. Some recent reviews have been published dealing complementarily with the chemical derivatization used in proteomics.^{1,2} In this chapter, the chemical modifications that are commonly employed for mass spectrometry-based tagging are presented. The cysteine, lysine, arginine, tryptophan, methionine and histidine targeting is reviewed. Peptide terminal labelling *via* the modification of primary amine or carboxylic group is in keeping with the general need of global analysis of the samples to insure sufficient sequence and proteome coverage. The tagging of PTM sites is as well essential due to the significance of such biological events and the sub-stoichiometric concentrations of associated proteins. Though phosphorylated-site targeting has received much more attention, glycosylation and tyrosine nitration targeting protocols have appeared. Herein, a particular consideration is given to the chemical reactions involved. Some of the MS-based tagging methods are comprehensively exposed and the advantages and the drawbacks of these strategies are addressed.

2. Specific amino acid targets

2.1. Cysteine modification

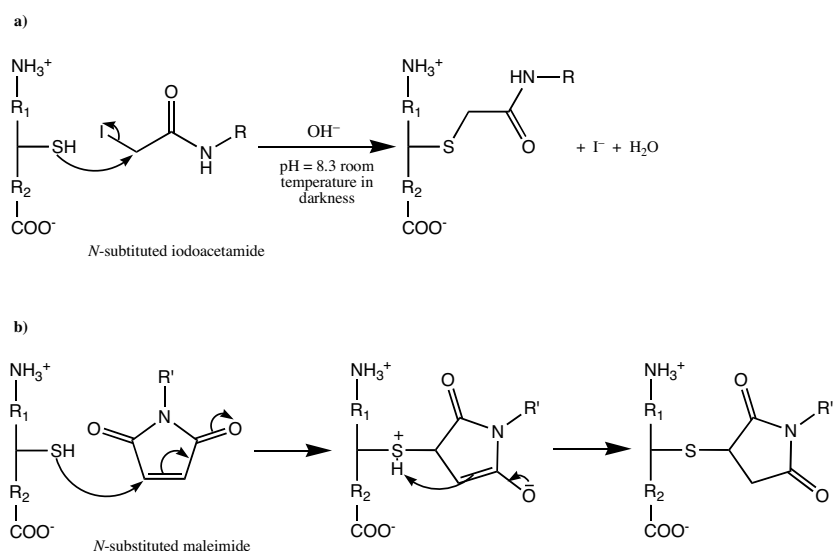
Cysteines are important for metal coordination, catalysis and protein structure by forming disulfide bonds but have been shown to be involved in modulation of protein activity and signalling events *via* reactions such as redox events, chelation of transition metals (mainly Zn^{2+} , Mn^{2+} and Cu^{2+}) or *S*-nitrosylation (*i.e.* catalysed transfer of nitric oxide NO to the thiol).

Cysteine is a frequent target of modifications since the thiol function can react specifically by nucleophilic substitution or nucleophilic addition. Modifying the cysteine amino acid was first introduced for 2D gel purpose. The alkylation of cysteine moieties

prevents indeed the random regeneration of disulfide bridges that induces many more spots in the final 2D map³ and avoids the cysteine-containing proteins to be retained by polyacrylamide gels.⁴

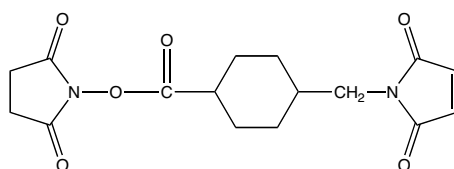
Structural studies of cysteinyl proteins:

Cysteine modification has been achieved for the rapid determination of cysteine number in proteins or peptides by mass spectrometric (MS) detection. For instance, the reaction with iodoacetamide has served for cysteine counting of complex peptide extracts (Scheme 1a).⁵ The mass spectrometric analysis is performed before and after the chemical derivatization. Cysteinyl species can be identified and the number of cysteines can then be determined by the resulting mass shift. In recent works, the derivatization with *N*-ethylmaleimide has been used to identify the number of cysteines in monoamine oxidases (Scheme 1b).⁶ Maleimide reactions are sulfhydryl specific as at pH = 7 the reaction proceeds 1000 times faster with thiols than with amines.^{7,8} At higher pH, the reaction with amine may also take place.⁹ The counting of cysteine units can reveal itself as valuable information in the identification of proteins through confrontation to protein databases.⁴



Scheme 1. Reactions of cysteine with *N*-substituted iodoacetamide (nucleophilic substitution SN2) (a) and with *N*-substituted maleimide (1,4-Michael addition) (b).

Chemical cross-linking of proteins with the MS analysis of the products of the reaction serves to elucidate 3D protein structures and interacting sequences in protein complexes. Cross-linking reagents, which present at least a reactive site targeting thiol functional groups on one of the spacer extremities, are common (Scheme 2). Maleimide but also disulfide compounds able to participate in a thiol-disulfide exchange are used in this case.¹⁰



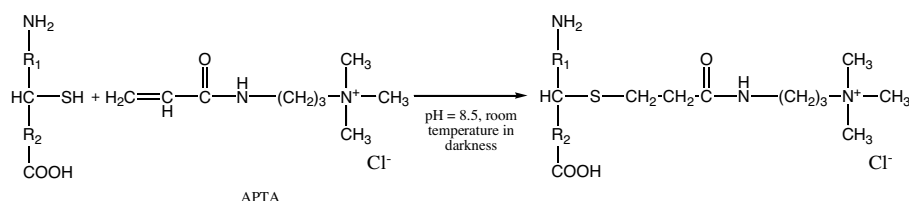
Scheme 2. Succinimidyl 4-(*N*-maleimidomethyl)cyclohexanecarboxylate (SMCC) for the cross-linking thiols and amines.

Chemical modifications of cysteines in proteins have extensively been used for probing protein topology and dissecting the functional importance of specific amino acid residues. According to the denaturation state of proteins for instance, *N*-ethylmaleimide and iodoacetamide have served to investigate the thiol accessibility by MS.^{11,12} The MS analysis is carried out directly on the protein. Followed by the study of the proteolytic peptides to localise the cysteinyl active sites, the reactivity of specific cysteine is accessed.

Selection of cysteinyl peptides:

An important feature of cysteine is that it is a low abundance amino acid. According to *in silico* studies of *E. coli*, yeast and the human proteome, cysteine is expected to occur individually in only 10 to 20% of the tryptic peptides from a proteome but in 90% of the proteins in a proteome.¹³ Then, the selection of cysteine-containing peptides reduces sample complexity, providing at least one peptide from 90% of the proteome. The use of tags to select cysteinyl biomolecules is common as shown by many proteomics application in the

following. The derivatization of cysteine by affinity tags has been performed with biotinylated reagents as a mean of selection. Tagged peptides are retained by avidin affinity chromatography and are finally eluted in acidic media to be analysed by LC-MS.



Scheme 3. Reaction of 3-acrylamidopropyltrimethylammonium chloride (APTA) with a cysteine residue.

More recently, the thiol 1,4-Michael addition on the double bond of APTA tag (3-acrylamidopropyltrimethylammonium chloride; Scheme 3), a quaternary amine tag (QAT), has allowed the selection *via* strong cation-exchange (SCX) of cysteine-containing peptides.¹⁴

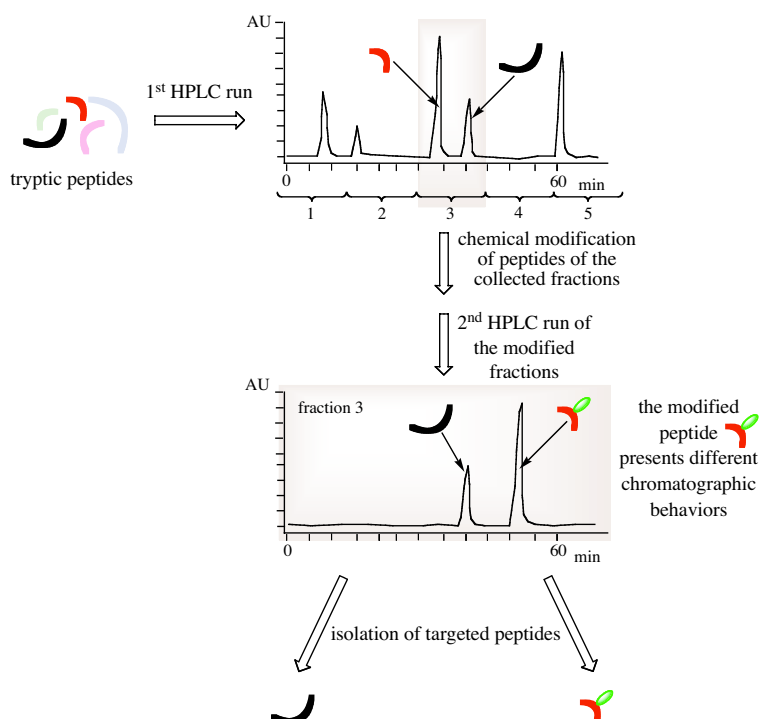
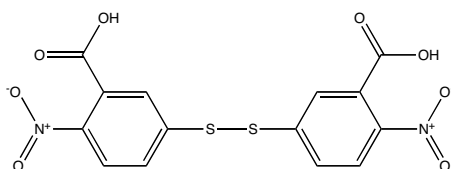


Figure 1. General strategy for the isolation of specific peptides by combined fractional diagonal chromatography (COFRADIC).

Diagonal chromatography is an interesting mean for the isolation of cysteine-containing peptides. K. Gevaert *et al.* have developed a method for the isolation of sets of representative peptides out of complex mixtures, called combined fractional diagonal chromatography (COFRADIC; Figure 1).¹⁵ In this COFRADIC approach, a reversible labelling method *via* disulfide exchange is carried out. 5,5'-Dithiobis(2-nitrobenzoic acid) is used to modify cysteine residues prior to proteolytic digestion (Scheme 4).



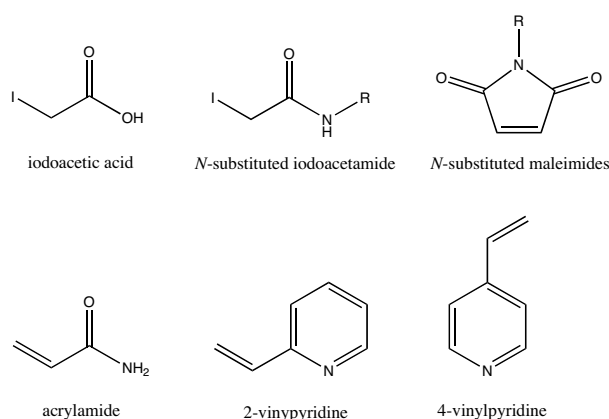
Scheme 4. Structure of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB, *i.e.* Ellman's reagent) used for disulfide exchange.

The obtained digested peptides are then of two kinds: cysteine-containing peptides are altered whereas cysteine-free peptides are unchanged. The complex mixture is separated during a primary high-pressure liquid chromatography (HPLC) run. The collected fractions are then submitted to chemical reduction by 1,4-dithio-DL-threitol (DTT). The covalently linked label being so removed, the cysteine-containing peptides actually become more hydrophilic. In a second HPLC run, the cysteinyl targeted peptides shift out of the primary collection interval and can be isolated for further analysis.

Relative quantification of cysteine-containing peptides:

The differential quantification in MS of cysteine-containing peptide by mass tags introduction is a widespread technique. The relative quantification between two samples is achieved by labelling one sample with a light and the other sample with a heavy form of the tag. Non-deuterated and deuterated acrylamide^{16,17}, 2- and 4-vinylpyridine¹⁸, *N*-methyl and *N*-ethyl-maleimide¹⁹, *N*-ethyliodoacetamide²⁰, *N*-tert-butyl-2-iodoacetamide and 2-iodo-*N*-

phenylacetamide²¹, for instance, have been used in this way (Scheme 5). After tagging, the two samples are mixed and further analysed by mass spectrometry. As the two labels do not induce difference in ionization efficiency, the relative quantification of cysteine-containing peptides is obtained. The limitation of the technique appears when a chromatographic separation is performed after the labelling since the introduction of the two different labels can induce two different chromatographic behaviours. It seems to be very dependent on the number and the location of heavy isotopes.



Scheme 5. Usual cysteine alkylating reagents for proteomics reported in the literature. The use of 3-bromopropylamine has been described for protein sequence analysis but has not been employed for MS-based proteomics.^{22,23} As well, redox indicators like quinone compounds initiating 1,4-Michael additions have been currently used for the detection of thiols.²⁴

The combination of affinity selection associated with the relative quantification of cysteine-containing peptides through the derivatization with specific tags has been carried out. R. Aebersold and co-workers have introduced the isotope-coded affinity tag (ICAT).^{25,26} The reagent consists of three components: a thiol reactive group that reacts specifically with cysteine, a polyether linker that can be synthesized as isotopically normal or heavy, and a biotin affinity tag. The strategy adopted is described in Figure 2. The technique enables an enrichment of cysteine-containing peptides of a mixture of two extracts, the first one having been primarily tagged by the light and the second by the heavy form of the reagent. After the release of the targeted peptides, LC-MS analysis is performed which provides the relative

quantification between the proteins of each extracts. The ICAT reagents are commercially supplied by Applied Biosystems (Foster City, CA, USA) and major bioinformatics softwares (like ProFound²⁷ and Mascot²⁸ employed in chapter IV) take it into account, making the tag being used regularly.

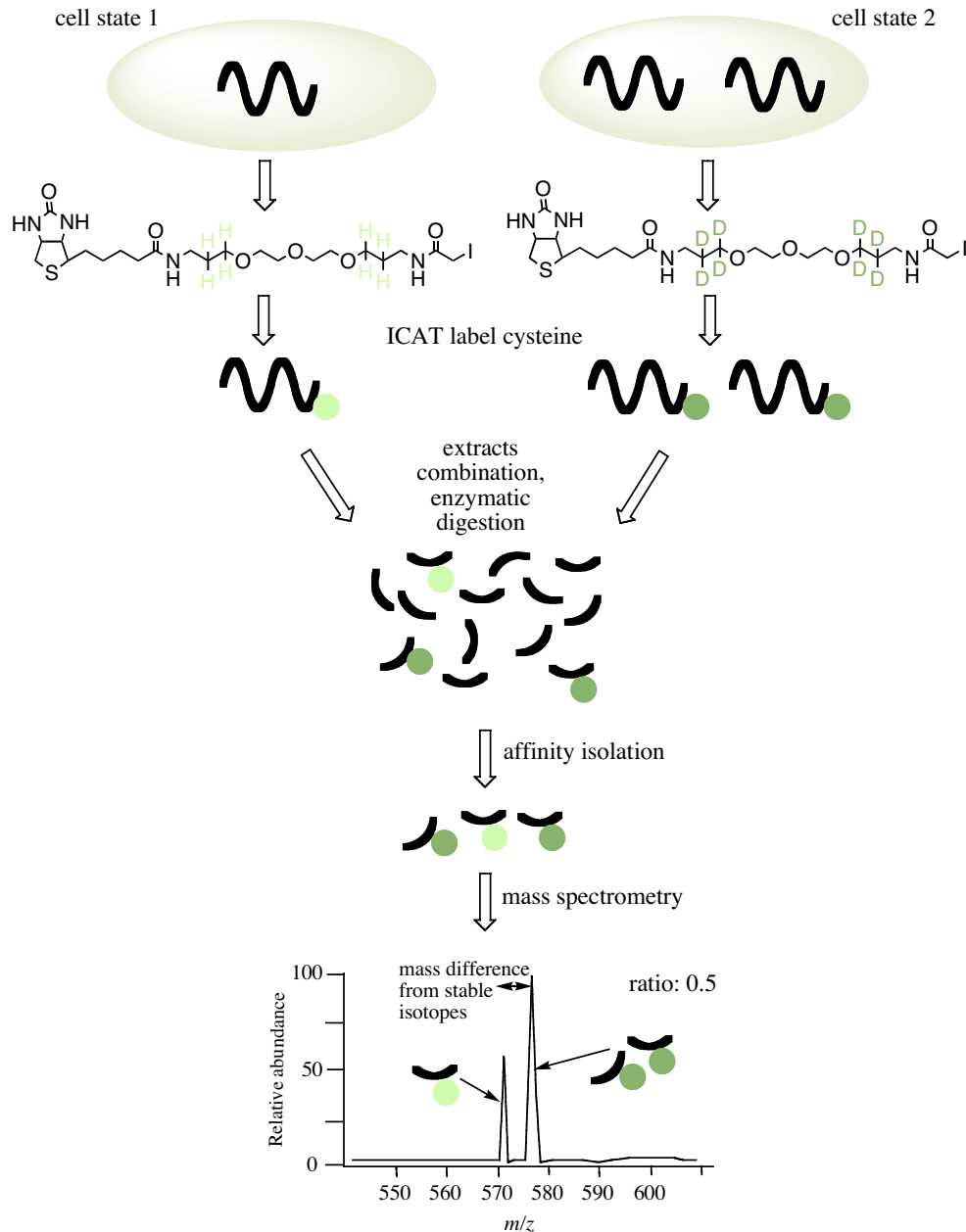
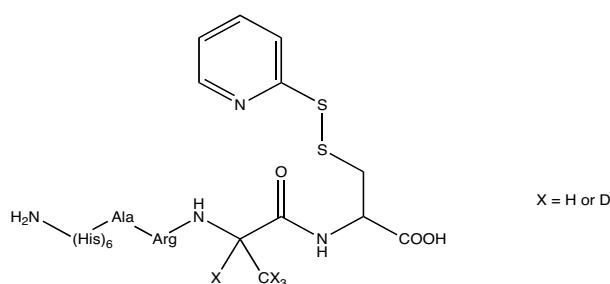


Figure 2. The isotope- coded affinity tag (ICAT) strategy is described for the enrichment and the relative quantification of cysteinyl peptides. The figure is adapted from S.P Gygi *et al.*²⁹

Nevertheless, F.E. Regnier and co-workers have pointed out that LC separation can occur between peptides labelled by the d_0 - and d_8 -forms, distorting the subsequent quantification. They proposed the use of isotope ^{12}C and ^{13}C labels that do not exhibit such chromatographic separation.^{30,31} Additionally, the presence of the hydrophobic biotin moiety narrows the targeted peptide elution zone during LC. The important mass of the reagent (442 Da for the light form) can be limiting when using ion trap mass spectrometers. Some people notice that the tag can complicate MS/MS patterning in addition because of loss of fragments from the tag.

Modified versions of the ICAT namely the “solid-phase” ICAT, the “cleavable” ICAT (cICAT) and the visible isotope-coded affinity tag (VICAT) have been developed to deal with the possible limitations of use. The “solid-phase” ICAT³² relies on photocleavable labels immobilized on glass beads. Once isolated on the beads, the tagged peptides are liberated under UV irradiation. Simplification of complex mixture is realized without biotinylation. The cICAT^{33,34} has been commercialised by Applied Biosystems and presents the same characteristics of original ICAT. However, it uses ^{12}C and ^{13}C instead of ^1H and ^2H isotopes and the linker can be broken in acidic media, thus avoiding chromatographic and MS/MS patterning issues described above. Lastly, the VICAT³⁵ combines ^{14}C isotope coding, photocleavable linker and allows absolute quantification of peptides of interest through ^{14}C scintillation counting and mass spectrometry. Other cysteine-specific affinity tags for quantitative proteomics have been recently introduced based on solid-phase capture approaches like the acid-labile isotope-coded extractants (ALICE).³⁶

The HysTag method combines an isotopically labelling of cysteine and complexity-reducing agent by immobilized metal affinity chromatography (IMAC).³⁷ The tag (*i.e.* $\text{H}_2\text{N}-(\text{His})_6\text{-Ala-Arg-Ala-Cys}(2\text{-thiopyridyldisulfide})\text{-COOH}$) is composed of an affinity ligand (His_6 -tag), a tryptic cleavage site (Arg-Ala) that release residues 1-8 of the reagent and Ala-9 residue that contains four (d_4) or no (d_0) deuterium atoms (Scheme 6).



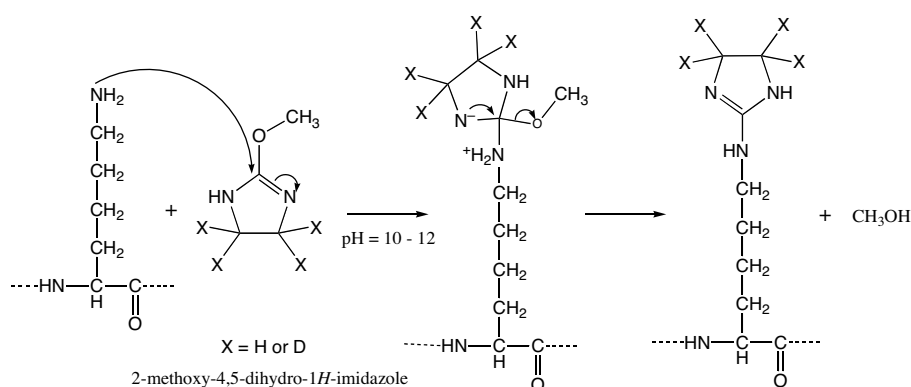
Scheme 6. Structure of the HysTag.

The tag reacts with proteolytic cysteinyl peptides *via* disulfide exchange. After selection on an immobilized Ni^{2+} column, the sample is once again submitted to trypsin digestion that release the part of the tag that is no more necessary. Analysis by LC-MS of the sample allows quantification of the extract respectively labelled by isotopically d_4 and d_0 tags. The authors reported that no separation occur between the isotopically labelled compounds and that the tag is not fragmented during MS/MS, thus not complicated the sequencing.

2.2. Lysine modification

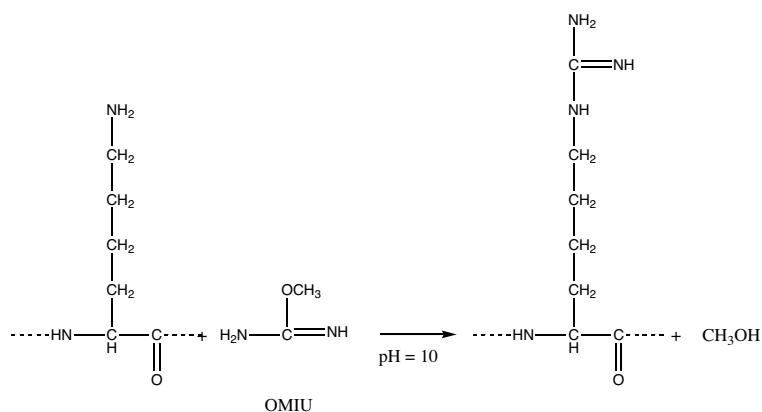
Because trypsin is the mostly used enzyme in bottom-up proteomics and cleaves protein at the C-terminus of lysine and arginine, many peptides have potentially one lysine in such strategies. Moreover, in matrix-assisted laser desorption ionization (MALDI), the lysine-containing peptides has been shown to produce ions detected in lesser abundance than the ones containing an arginine residue that is more basic.³⁸ Therefore, the insertion of basic residue at the lysine is valuable to insure better sequence coverage and several electrophilic tags have been employed. Because trypsin usually does not cleave C-terminal to a modified lysine, the derivatization must be performed after protein digestions.

Lysine-containing peptides can be specifically tagged using 2-methoxy-4,5-dihydro-1*H*-imidazole and relatively quantified thanks to the use of its deuterated form (Scheme 7).³⁹



Scheme 7. Specific reaction of lysine with 2-methoxy-4,5-dihydro-1*H*-imidazole.

The tag does not react with the N-terminal since the pK_a of the lysine side chain is higher ($pK_a = 10.53$ for the amino acid). The MS intensity signal of the tagged peptides is enhanced compared to its initial form because the basicity of the peptides increases thanks to the tag. The MS/MS patterning is facilitated when the lysine is at the peptide C-terminal (see section 3.2. of this chapter), which is relevant when trypsin is used in the digestion. The tag is sold by Agilent Technologies (Palo Alto, CA, USA).

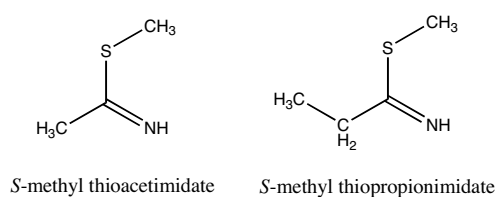


Scheme 8. *O*-methylisourea (OMIU) modifies lysine into homoarginine (guanidination).

The specific guanidination of lysine by *O*-methylisourea (OMIU) (Scheme 8) has been used for lysine counting as a means of further constraining protein search in database^{40,41}, lysine blocking prior to the sulfonation of N-termini⁴² (see also 3.1. of this

chapter), sequencing of peptides and their relative quantification by LC-ESI-MS termed mass-coded abundance tagging (MCAT).⁴³ In the latter method, the authors noticed that there is only a slight temporal difference in the elution of the modified and unmodified peptide pairs in reverse-phase liquid chromatography (RP-LC) (the modified peptide eluted after the unmodified one). Moreover, the quantification suffers from the intrinsic ionization efficiency difference between the lysine and homoarginine-containing peptides in electrospray ionization MS (ESI-MS).⁴⁴ To cope with these issues, F.L. Brancia *et al.* have recently introduced *O*-methylisourea incorporating C and N isotopes that enables relative quantification.⁴⁵ Their guanidino-labelling derivatization (GLaD) strategy was successfully tested for model protein digests.

The relative quantification of lysine-containing peptides has been also realized by differential amidination of lysine by *S*-methyl thioacetimidate and *S*-methyl thiopropionimidate (amidination reaction) (Scheme 9). The approach is called quantification using enhanced sequence tags (QUEST)⁴⁶ but it is restricted to MALDI-MS since separation between the two tagged peptide analogues should occur during LC. The technique enhanced ion-yields for lysine-containing peptide and that way greatly increased the efficacy of peptide mass mapping.



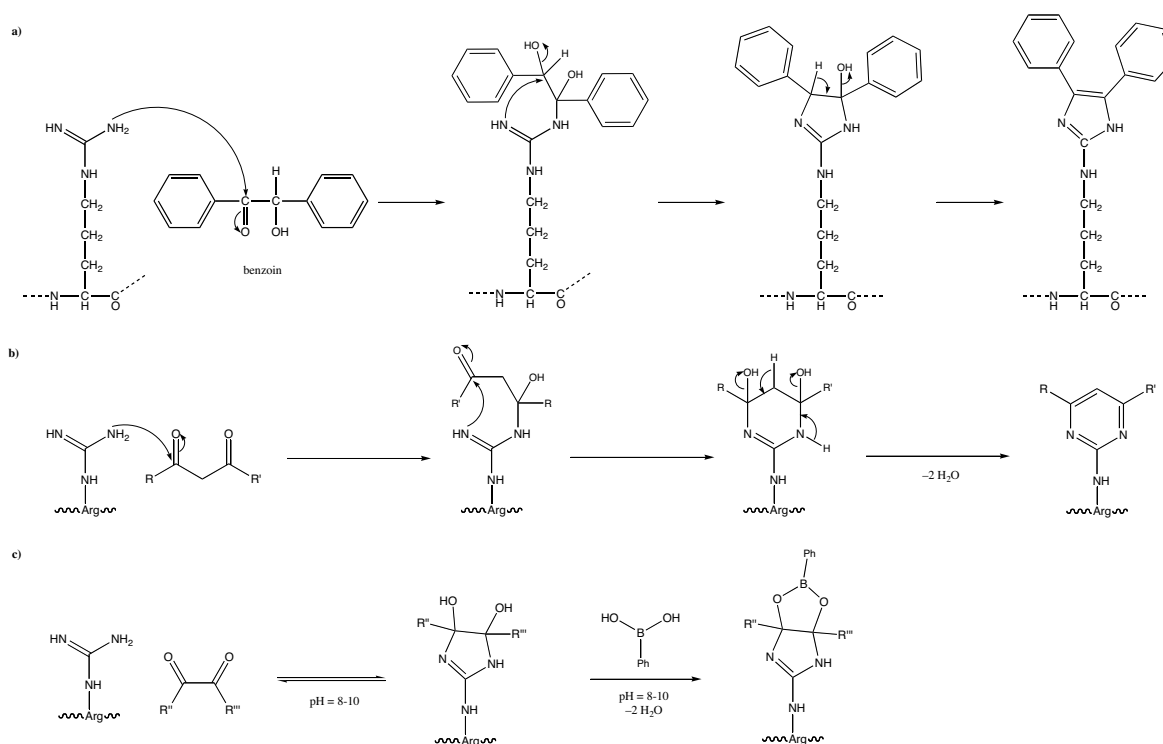
Scheme 9. Reagents used in QUEST for the amidination of lysine.

The number of reagents for lysine increases when working directly with protein. Indeed, at that level they are the only primary amine in the chain in addition of a single N-terminus. Chemical modification of lysine in combination with mass spectrometric detection of entire proteins followed by their tryptic peptides analysis was used to characterize the

protein surface topology and the reactive fundamental lysine residues. Acetylation and succinylation of lysine- ϵ -amino (see section 3.1. of this chapter) have been used in that case to tag entire proteins.^{47,48} Recently, *N*-hydroxysuccinimidobiotin (NHS-biotin) has been used to evaluate conformational feature induced by peptide binding in a given class of protein.⁴⁹ Cross-linking reagents with functionalities such functionalities are available to target amino groups in proteins.

2.3. Arginine modification

Because of the strong basicity of arginine ($pK_a = 12.48$ for the amino acid), the presence of this amino acid greatly influences the ionization efficiency and the fragmentation of peptides under collision-induced dissociation (CID). As specified before (see section 2.2.), ionization efficiency implies prevalence of tryptic peptides with terminal arginine over peptides with terminal lysine.



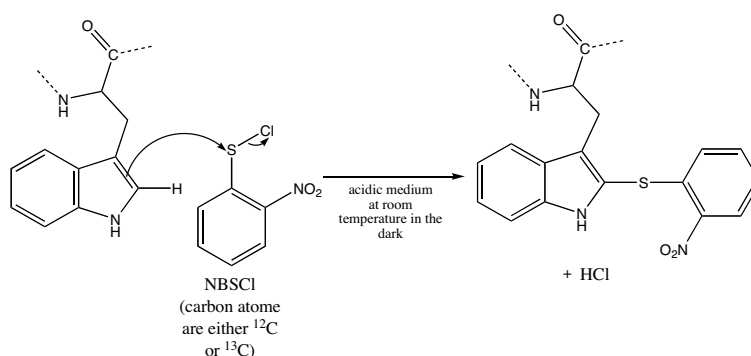
Scheme 10. Reactions towards arginine of benzoin (a), β -dicarbonyl compounds (b) and α -dicarbonyl compounds followed by reaction with phenylboronic acid (c).

The basicity of arginine residues may be reduced by reaction with benzoin that is selective for guanidine and monosubstituted guanidino compounds (including arginine containing peptides) in alkaline solution (Scheme 10).⁵⁰ This derivatization has been used for the HPLC determination of arginine-containing peptides because of the strong fluorescence of the 2-substituted amino-4,5-diphenylimidazole adduct.

Reducing the basicity of arginine is considered detrimental to the overall sensitivity of MS-based analysis. However, in order to simplify MS/MS fragmentation, it has been shown that derivatization with acetylacetone⁵¹ or 1,1,1-trifluoro-2,4-pentanedione⁵² can be helpful. Deactivation of the arginine group leads to a complete suppression of C-terminal ions (*i.e.* x, y, z ions) and to an enhancement of signal-to-noise ratios of some N-terminal ions (*i.e.* a, b, c ions). Recently, the reaction of guanidino group with α -dicarbonyl compounds (*i.e.* 2,3-butanedione) followed by the reaction with an aryl boronic acid (Scheme 10) has been carried out to see the influence of such tags on the MS/MS patterning of arginine-containing peptides.⁵³⁻⁵⁵ The surface topology-probing has been carried out using the reactivity toward 1,2-cyclohexanedione (Scheme 10).⁴⁷

2.4. Tryptophan modification

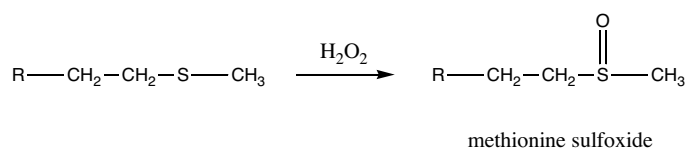
Tryptophan is a less abundant amino acid (even less abundant than cysteine residues) and is one of the most hydrophobic residues among the 20 amino acids. 2-Nitrobenzenesulfonyl chloride (NBS-Cl) has been used to label tryptophan residues (Scheme 11).⁵⁶ This type of reagent has indeed been found to react with tryptophan and cysteine residues in proteins and peptides. Of the two residues, the labelling of tryptophan prevails over cysteine in acidic media. After derivatization, tryptophan is rendered even more hydrophobic by the NBS tag. Then, tryptophan-containing peptides can be selected with a Sephadex or RP column. The method allows quantitative proteome analysis using the heavy (contains six ¹³C) and light (contains six ¹²C) forms of the reagent.



Scheme 11. The reaction of tryptophan with 2-nitrobenzenesulfonyl chloride (NBSCl) is depicted.

2.5. Methionine modification

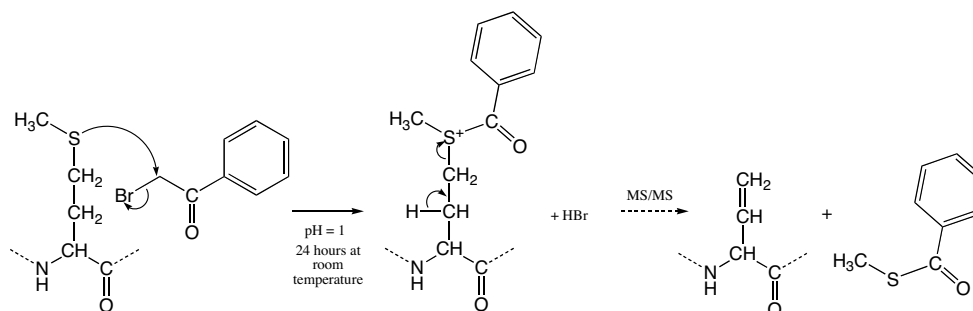
Methionine can be modified by oxidation. Methionine-containing peptides were oxidised by hydrogen peroxide for instance (Scheme 12). The retention time of the modified peptides generally changes a few minutes in RP-LC. Diagonal chromatography (see Figure 1 for the general procedure description) has been used for the isolation of these peptides. The digest proteome is separated by RP-LC and fractions are collected and treated by hydrogen peroxide. Each fraction is then rechromatographed. The retention time of methionine-containing peptides is generally reduced from 3 to 6 min, enabling their isolation.⁵⁷



Scheme 12. Methionine residues are oxidised by hydrogen peroxide.

The approach called selective extraction of labelled entities by charge derivatization and tandem mass spectrometry (SELECT) has been realised by specific alkylation of methionine in very acidic media (Scheme 13).⁵⁸ In fact, this is possible because the rate of alkylation of the thioether sulfur of methionine is independent of the pH whereas the reactivity of other nucleophiles (*i.e.* cysteine, lysine and histidine) decreases with the pH. The strategy is based on the formation of a “fixed charge” sulfonium ion that results in

characteristic neutral loss fragmentation under CID that can be rapidly and sensitively identified out of complex mixture of peptides (Scheme 13).

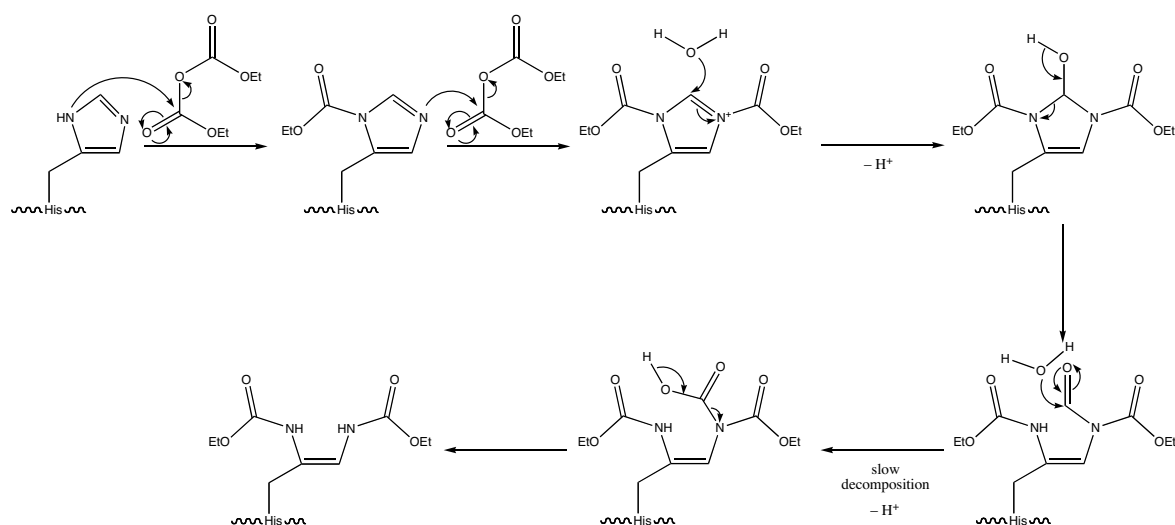


Scheme 13. Modification of methionine by 2-bromoacetophenone. The fragmentation products induced by CID MS of the derivatized compound are shown.

Reaction with heavy and light form of 2-bromoacetophenone has been shown to provide the quantitative analysis of differential protein expression, *via* measurement of the relative abundances of the ions generated by the neutral loss induced by the labels observed in a triple quadrupole mass spectrometer.

2.6. Histidine modification

Histidine residues are involved in various functions of proteins as proton donors and acceptors and especially as metal chelating agent. In myoglobin and hemoglobin for instance, histidine binds the central ions of the heme groups. Diethyl pyrocarbonate (DEP) has been used for the modification of histidine. This derivatization has been successfully applied in order to study the different reactivity of histidine residues and to distinguish them by MS. The drawbacks of this reaction are that side reactions with other nucleophilic amino acids and bismodifications can occur (Scheme 14).⁵⁹ Moreover, the constitutional isomers obtained from the reaction on different nitrogen do not allow the strategy to be of general used for proteomics applications.



Scheme 14. Reaction of histidine with an excess of diethyl pyrocarbonate, showing several possibly obtained products. The reaction is known as the Bamberger's imidazole cleavage. The constitutional intermediate isomers, that cannot be distinguished by simple MS but isolated by RP-HPLC for instance, are not shown.

The selection of histidine-rich peptide *via* Cu(II)-IMAC has been carried out and such peptides quantified.⁶⁰ However, the tagging step of this method consists in the acylation of the N-terminus of tryptic peptides to quantify the targeted peptides (see section 3.1.2. of this chapter).

3. *N*- and *C*-terminal targets

3.1. *N*-terminal labelling

N-terminal tagging strategies

The two usual reactions for modification of amines are acylation or alkylating reaction, which mostly occur rapidly and in high yield to give amide or secondary amine. In this section, a method that aims at determining the N-terminal peptide of a protein is described to reduce the sample complexity by transforming every protein in a unique peptide. Several strategies that target proteolytic peptides presenting specific N-terminal pattern are also presented.

The combined fractional diagonal chromatography (COFRADIC, see Figure 1) was described for the isolation of N-terminal peptides (Figure 3).⁶¹ The selection of N-terminal peptide reduces sample complexity. A good coverage of the proteome is ensured since each protein provides one N-terminal peptide.

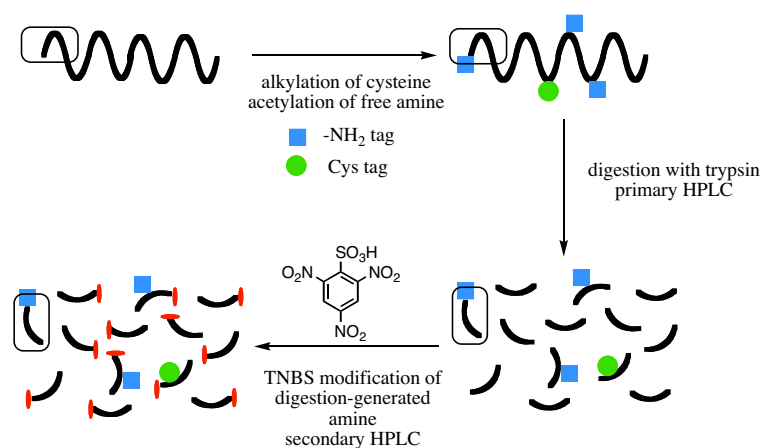


Figure 3. COFRADIC strategy for the isolation of the unique N-terminal peptides of proteins. In the last step, the internal peptides are rendered more hydrophobic and can be discarded. The N-terminal peptide that appears in a box is collected and analysed.

In the procedure, the cysteines are first alkylated with iodoacetamide and the free amine functions are acetylated (see section 3.1.2. of this chapter). The proteins are digested with trypsin, which now only cleaves at arginine residues. The internal peptides with free N-terminus and the N-terminal peptides with blocked N-terminus resulting are then separated by RP-HPLC. Each collected fraction is treated with 2,4,6-trinitrobenzenesulfonic acid (TNBS) that modified the free N-terminal of internal peptide. In a second RP-HPLC run, peptides altered with TNBS have become more hydrophobic and exhibit shifted retention times. The species with unchanged retention times are identified as the N-terminal peptides that can be isolated and further analysed by LC-MS/MS. The approach for identifying proteins through the mass spectrometric analysis of their N-terminal peptide should be feasible using highly accurate mass spectrometers.

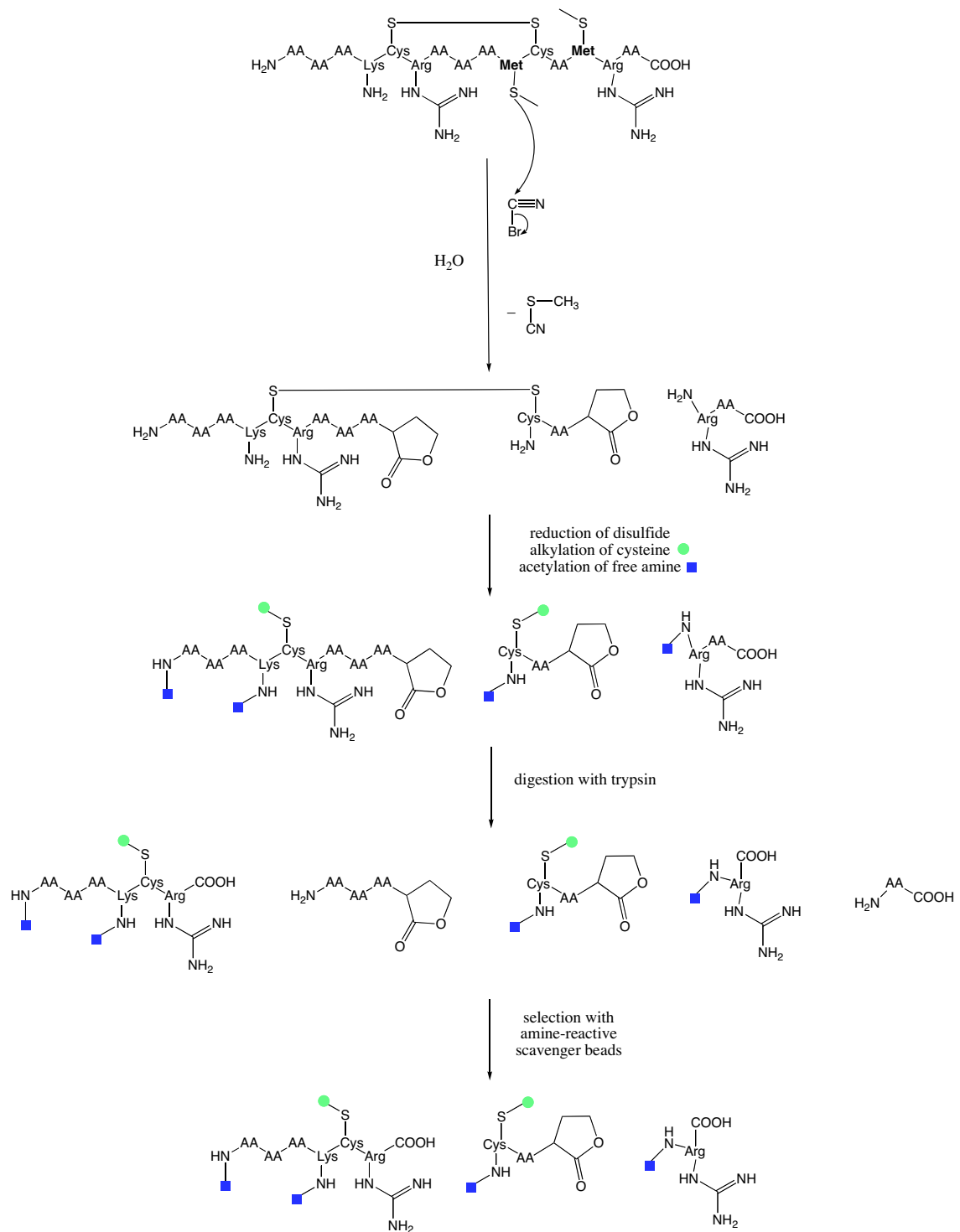


Figure 4. PST procedure, that enables the selection of N-terminal fragments from CNBr cleavage peptides (see in the text for detailed explanation).

C. Hamon and co-workers have developed the protein sequence tag (PST) methodology that enables the isolation of N-terminal peptides from cyanogen bromide (CNBr) cleaved proteins (Figure 4).^{62,63} CNBr cleaves proteins at the methionines. Disulfide bonds are reduced and the thiol residues are alkylated with iodoacetamide. The primary amines are blocked with a basic tag, as *N,N*-dimethylglycine *N*-hydroxysuccinimide ester. This tag is easily protonated, which is an advantage for MS ionization techniques or SCX chromatography. As lysine has become inaccessible to trypsin, tryptic digestion generates peptides with arginine at the C-terminus. All the peptides with primary amines that result from this digestion can be retained on an amine-reactive scavenger resin. The end result of the process is a pool of peptides that represents the N-terminal fragments from each CNBr cleavage peptide. This procedure gives a mixture of peptides of reduced complexity that can be directly analysed by LC-MS in a gel-free approach. The PST technology has been shown to be robust for the analysis of hydrophobic proteins. It increases the number of such proteins identified.

N-terminal serine and threonine tagging have been realized in order to capture peptides presenting this characteristic since the distribution of such tryptic peptides in the proteome is quite similar to that of cysteinyl peptides.⁶⁴ The mild oxidation of the N-terminal 1,2 amino alcohols of serine and threonine gives an aldehyde or a ketone respectively (by this way methionine are oxidised too). The resulting peptide is then derivatized with biocytin hydrazide^ξ for the selection on streptavidin-coated magnetic beads.

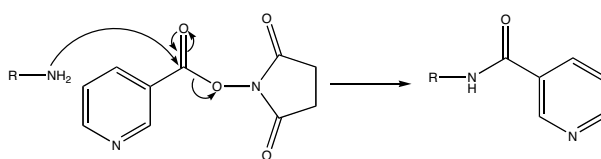
Global N-terminal isotope coding

The previously presented labelling techniques for proteomics target specific amino acids or peptides. If on one hand they reduce sample complexity, on the other hand

^ξ *i.e.* biotinyl- L-lysine hydrazide (-CO-NH-NH₂).

these methods inevitably lead to reduced sequence and proteome coverage. Another approach consists in the systematic labelling of all peptides at their N-terminus.

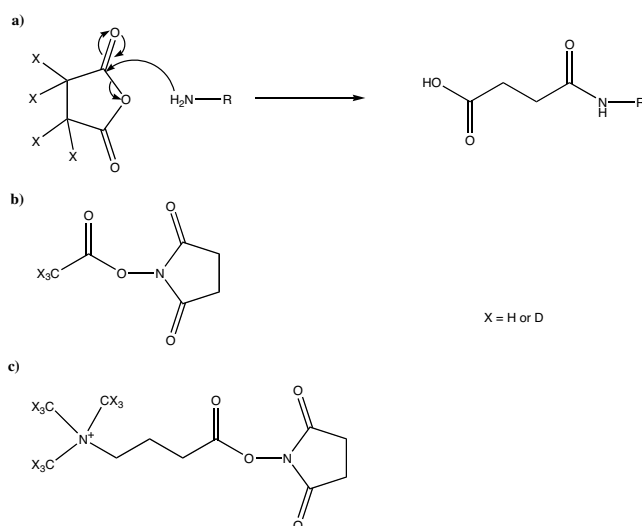
M. Munchbach *et al.* have introduced a generic approach for the labelling of peptides at their N-terminus. Nicotinyl-*N*-hydroxysuccinimide ester and its deuterated form, which provides 4 mass unit difference, have been used for global comparative quantification and facilitated *de novo* sequencing (Scheme 15).⁶⁵ Briefly, both protein samples are digested. The lysines are blocked by succinylation. The first sample is derivatized with the light form of the tag whereas the other is modified with the heavy tag. The two samples are then combined and analysed by MALDI time of flight MS (MALDI-TOF-MS). In addition of the relative quantification, the tag moiety directs the fragmentation since the highly basic nature of the nicotinyl group greatly increases the relative yield of the b-ion series compared to the native peptide. The b-ions are easily identified as doublets linked to the isotopic pattern whereas y-ions appears as singlets.



Scheme 15. Reaction of amine with nicotinyl-*N*-hydroxysuccinimide ester.

The global internal standard technique (GIST) is a stable-isotope labelling of primary amines with *N*-acetoxy succinimide^{66,67} or succinic anhydride¹³ and their deuterated analogues introduced by F.E. Regnier and co-workers (Scheme 16). It has been shown that the GIST strategy uniformly labels all peptides in a tryptic digest, accurately predicts the degree of change in protein expression in a biological system and can be used with either MALDI-MS or ESI-MS. The technique has been associated with inverse labelling⁶⁸ for accurate quantification of up and down regulation.¹⁴ Briefly, the inverse labelling for the quantification of two extracts consists in dividing each of the two extracts into two parts that are separately labelled with the heavy and the light form of the tag. The heavy labelled part of the first extract is mixed with the light labelled part of the second extract and *vice versa*. The

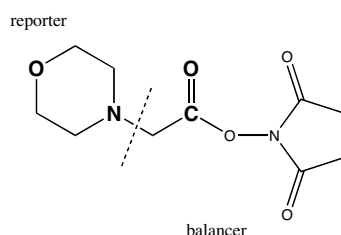
two mixtures are then analysed by LC-MS. If the first mixture reveals an increase in isotope ratio, then the second mixture must reveal a decrease, meaning that the two extracts have really different expression states. The advantage of inverse labelling is the ability to differentiate between proteolysis artefacts and very large change in concentration. More recently, 4-trimethylammonium butyrate has been used as labelling reagent.⁶⁶ Because acetylation of primary amines with *N*-acetoxy succinimide increases the peptide acidity by removal of a positive charge, lysine-containing peptides become harder to detect. By the introduction of the quaternary amine terminating tag, the problem is addressed.



Scheme 16. Acylation reagents used for the GIST technology. Succinic anhydride (a), *N*-acetoxy succinimide (b) and 4-trimethylammonium butyrate (c) are used as hydrogenated or deuterated labels.

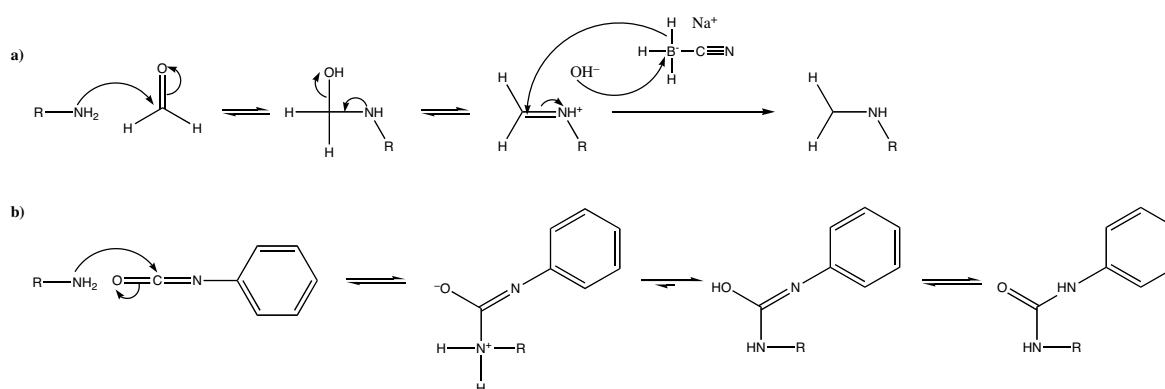
The new isotope tags for relative and absolute quantification (iTRAQ) commercialised by Applied Biosystems are amine modifying labeling reagents for multiplexed relative and absolute protein quantification. These isobaric tags contain both reporter and balancer group (Scheme 17). The reporter group serves for quantification. During CID it is cleaved to yield an isotope series that is representative of the amount of a single peptide of known mass from each of up to four samples. The balancer group is depleted of the same stable isotope, which provides each isotopic tag at the same mass. The different extracts (four at the maximum) are labeled with one of the iTRAQ reagents. They are gathered and

subjected to LC-MS or LC-LC-MS. Each mass peptide is then submitted to MS/MS for its quantification in each extract. As with GIST, the method targets all the peptides but reveals itself time-consuming because of the need of MS/MS for each peptide. Several applications have already been reported.⁶⁹⁻⁷²



Scheme 17. iTRAQ reagents. The isotopes are in bold type.

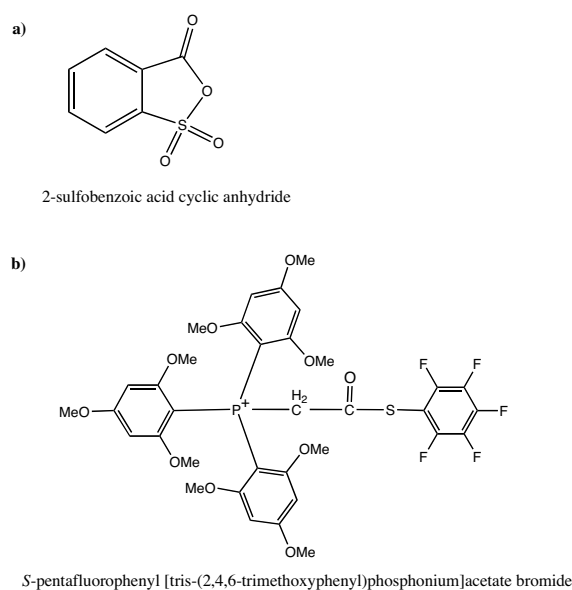
Reductive amination has also been described for the labelling of N-termini and lysine by formaldehyde in presence of sodium cyanoborohydride (Scheme 18a).⁷³ Phenyl isocyanate (PIC) has been shown to be specific for the labelling of N-termini at pH = 8 since thiols are prior alkylated (Scheme 18b).⁷⁴ Carboxyl and hydroxyl groups can react only at low pH and the reactivity of the ϵ -amino group of lysine is lower than that of the N-terminal amine.



Scheme 18. Reductive amination (a) and reaction of amine with phenyl isocyanate (b).

Tagging for tandem mass spectrometry

Improving mass spectrometric sequencing *via* the labelling of N-terminal amino groups of peptides is common as previously seen. The two basic techniques when only this purpose is in view are the sulfonation and the attachment of phosphonium groups (Scheme 19).



Scheme 19. Examples of reagents for the attachment of a negative (a) or a positive (b) charge at the N-terminus of peptides.

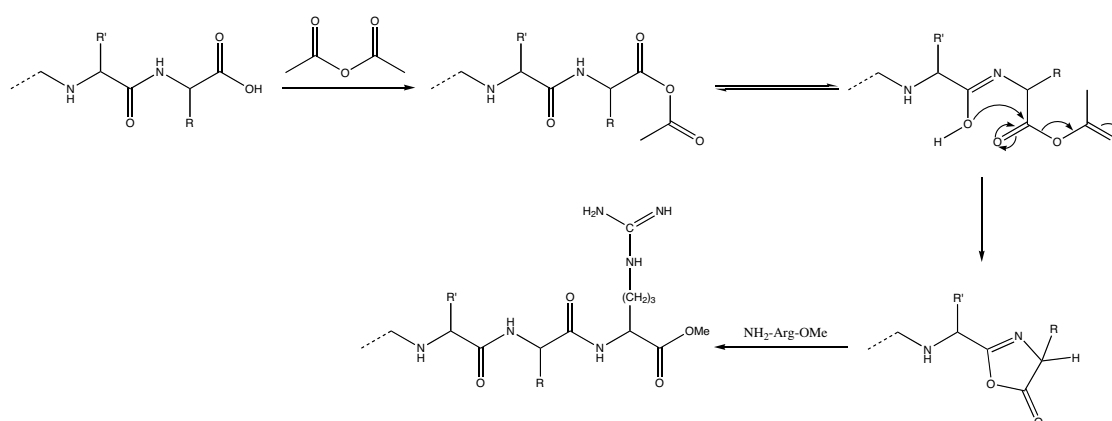
The attachment of the negatively charged labels by sulfonation suppresses the N-terminal fragments ions and y-ions are only observed.⁷⁵ The opposite procedure, *i.e.* attaching a positive charge on the N-terminus, can be carried out, thereby causing the predominance of a- and b-ions.^{76,77} In this way, phosphines or phosphonium salts are used.

3.2. C-Terminal labelling

The esterification of carboxylic acids of tryptic peptides has been performed with methanol and its deuterated analogues.⁷⁸ As with the systematic labelling of amines, the C-terminal isotope coding technique enables the tagging of almost all the peptides in the sample.

Using this approach, the authors have developed an algorithm that derives peptide sequence *de novo* by comparison of tandem mass spectra of d_0 and d_3 -peptide methyl esters.

The conversion of peptides C-terminal carbonyl group to oxazolone has been developed for enhancing the response of a peptide in MALDI-MS.⁷⁹ The reaction consists of dehydration with acetic anhydride to give an oxazolone, followed by reaction with amino acid derivatives such as arginine methyl ester (Scheme 20). The authors are eager to exploit the technique to discern the C-terminal tryptic peptide of a protein that is almost insensitive to MALDI-MS in general. Moreover, the incorporated tag directs the fragmentation, accentuating z- and b-type ions under post-source decay that should facilitate the data interpretation.



Scheme 20. The oxazolone-based C-terminal derivatization of a peptide is depicted.

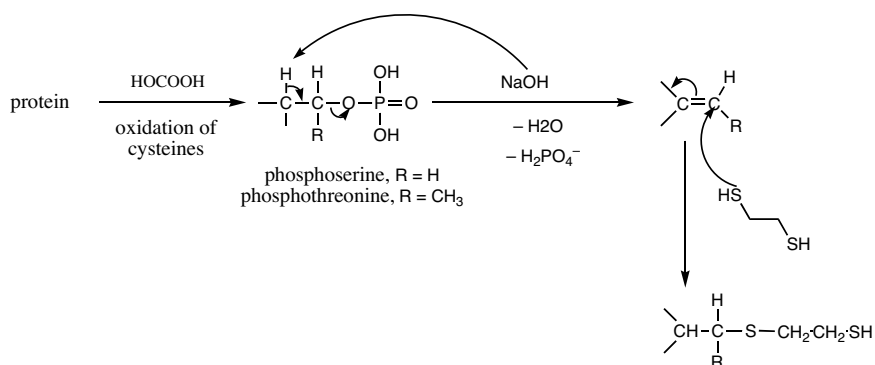
4. Derivatization of post-translational modification sites

4.1. Serine and threonine phosphorylation

The phosphorylation of amino acids that happens mainly on serine, threonine and tyrosine residues regulates a vast number of biological processes like signal transduction, cell differentiation, proliferation, metabolism, proteins production and apoptosis. Because of the

low abundance of phosphopeptides after proteolytic digestion, the enrichment of such peptides is required, all the more so as the 100% coverage of the protein sequence is difficult to obtain. The identification of phosphorylated species in peptides mixture is further complicated by the mass spectrometric properties of the highly acidic phosphate moieties that imply lower ionization efficiency or suppression phenomena.

Several methods targeting the serine and threonine phosphorylation sites are based on the selective chemical modification of their phosphate group. Once the elimination of the phosphate is realized, an addition of thiol or amine is performed on the double bond (Scheme 21).



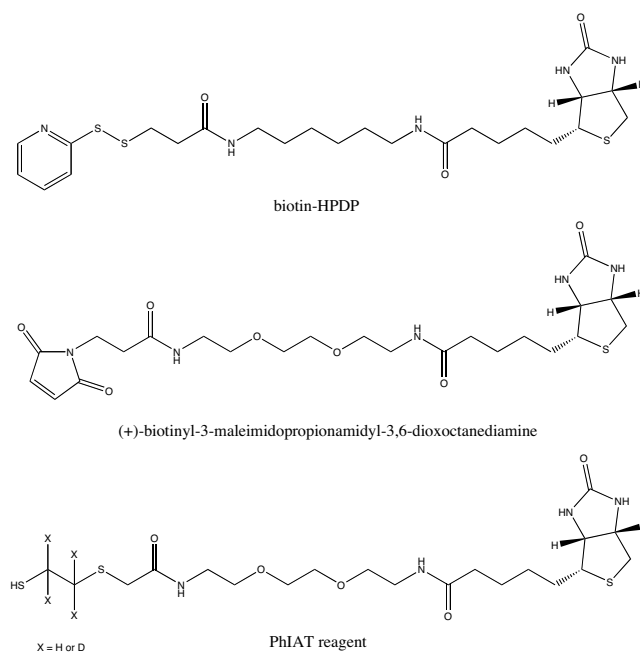
Scheme 21. Chemical elimination/addition modification of phosphorylated serine and threonine. Isomeric considerations are not taken into account.

Whereas the elimination in strong basic media either in the presence of sodium hydroxide or barium hydroxide (β -elimination is catalysed by Ba²⁺ ions⁸⁰), many nucleophilic addition reagents have been used. C. Klemm *et al.* present several of them in a recent study dedicated to the enhancement of ionization efficiency of phosphorylated peptides in MALDI.⁸¹ The mostly employed addition reagent remains 1,2-ethanedithiol (EDT).^{82,83} The EDT reagent presents two thiol functions. One is dedicated to the addition reaction and the other one can serve as a mean of selection.

Besides, in the case of phosphorylated serine, the nucleophilic 1,4-Michael addition happens on the carbon initially carrying the phosphate group. The racemization of

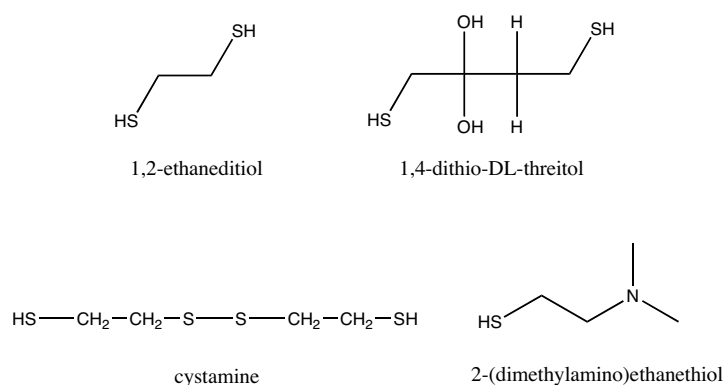
the α carbon happens during its re-protonation. Using EDT, the diastereoisomeric modified peptides obtained (in the case of the single serine amino acid derivatization, enantiomers are obtained) can be eventually differentiated during a chromatographic separation. Some works based on RP-HPLC have shown several products coming from a phosphorylated serine peptide.^{80,84,85} It subsequently implies that the MS signal intensity of the modified peptides is divided when such separation is over. In the case of threonine, the attacks on the *Si* and *Re* side of the double bond imply the formation of four diastereoisomeric peptides. It can explain the difficulty to recover peptide phosphorylated on threonine.

W. Weckwerth *et al.* have used the deuterated form of EDT in order to be able to comparatively quantify by LC-MS and consequently identify phosphoproteins by the mass shift.⁸⁶ Another approach for identification consists in the derivatization with a mixture of two different alkanethiols.⁸⁷ Phosphopeptide-specific derivatives are distinguished due to their characteristic ion-pair signature but the MALDI-MS intensity signal is divided twice (more if the peptide contains more than one phosphorylation site). Even if in both cases the derivatized peptides present a better ionization efficiency than their initial phosphorylated form, the problem of sample complexity still remains.



Scheme 22. Biotin chemistry used for the isolation of phosphorylated peptides.

The enrichment analysis of phosphorylated peptides has been proposed by converting the phosphate into a biotin moiety. The free thiol group of the peptide modified by EDT is put to react with (+)-biotinyl-3-maleimidopropionamidyl-3,6-dioxoctanediamine⁸⁸ or biotin-HPDP (Pierce, Rockford, IL, USA) (Scheme 22).⁸⁵ The targeted peptides are enriched by avidin chromatography. M.B. Goshe *et al.* adopted this strategy and combined it with the use of deuterated EDT for relative quantification of phosphorylated state. The method is called the phosphoprotein isotope-coded affinity tags (PhIAT) (Scheme 22).^{89,90} A.J. Thompson *et al.* decided to combine IMAC selection on iron(III) beads of phosphopeptides with the β -elimination/Michael addition of 2-aminoethanedithiol (Scheme 23) in order to select and facilitate peptide identification and sequencing by MS/MS.⁹¹ The relative quantification of phosphoprotein via Ga(III) selection and GIST has also been proposed.⁶⁶



Scheme 23. Commonly used reagents for the addition reaction following the elimination of the phosphate group.

In parallel, different groups have used directly the properties of the free thiol group remaining to isolate the peptide either with iodoacetyl group immobilized on glass bead (in this case, the addition reagent was cystamine (Scheme 23), which was further reduced by DTT liberating the free thiol).⁹² DTT (Scheme 23), which is more soluble than EDT in water, is widely employed too. It presents two asymmetric carbons (*N.B.* the commercial product is a racemic mixture). Its use in RPLC-ESI-MS based methods over the derivatization is

tendentious since four diastereoisomers can have been formed in serine case (eight in threonine case), possibly distorting the endeavour of differential quantification of samples. Such separation steps have to be avoided. DTT was successfully used for isolation on thiol-Sepharose resin⁹³ and quantification by MALDI-MS.⁹⁴

H. Steen and M. Mann employed 2-(dimethylamino)ethanethiol as addition reagent that introduces an additional basic functional group at the former phosphorylation site.⁹⁵ The thioether oxidation provides a sulfoxide by controlled addition of H₂O₂. The generated 2-(dimethylamino)ethanesulfoxide derivatives improve the effective ionization. They give rise to a characteristic fragment ion of $m/z = 122.06$ Th upon low energy CID that enables detection by precursor ion scanning^ξ in positive mode.

4.2. Glycosylation

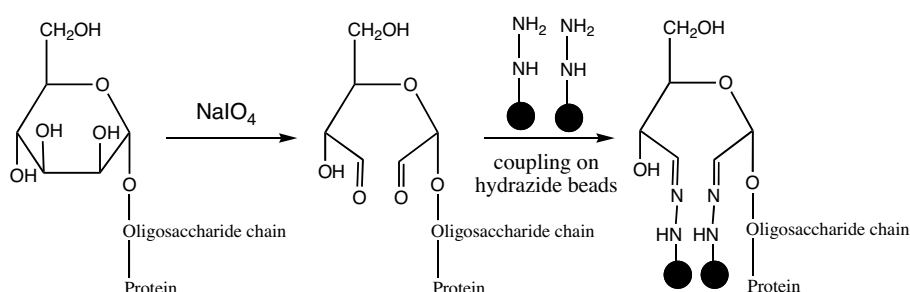
Glycosylation is the most frequently observed PTM. This PTM is associated with cellular regulation and function⁹⁷ and implies with many diseases like Alzheimer's disease, cancer and diabetes. Two types of glycosylation exist. The *O*-glycosylation is the attachment of mono- or oligosaccharides to the hydroxyl group of serine, threonine and sometimes to hydroxyproline. The N-linked glycosylation happens at the amide group of asparagines when a serine or a threonine residue is located after the following amino acid. It occurs while the protein is still on the ribosome (*i.e.* during protein synthesis). The N-linked glycosylation is also not really a PTM. However, deletion or addition of sugar units can occur as PTM, in both *O*- and N-linked glycosylation. Glycosylation is complex because of the numerous variation of the attached glycans. At a first stage of glycoproteomics, only the identification of such PTM is required.

^ξ precursor ion scan is also called selected ion monitoring (SIM) or selected reaction monitoring (SRM) depending on whether performed in single or tandem quadrupole mode. This method takes advantage of certain characteristic losses by ions of interest for their detection. As example, in SIM mode, the quadrupole analyzer is set to pass only a fragment with m/z characteristic of the ion. Therefore, only when the target ion is present in the source will a signal be detected. Since no scanning of the quadrupole is necessary, this is a very sensitive technique.⁹⁶

GIST technology has been coupled to lectin affinity chromatography have been combined for differential proteomics.^{98,99} The tryptic peptides are acetylated using the light or heavy *N*-acetoxy succinimide before separation on a lectin column. Lectins are plant proteins, which binds to glycan contained in peptides and proteins. Many lectins exist which are specific to certain glycosylation pattern. The glycopeptides are then deglycosylated by PNGase F (a glycosidase specifically cleaving N-linked glycans) to get rid of the heterogeneous oligosaccharides and identified by MS.

Another method consists firstly in the selection of tryptic glycopeptides by lectin affinity. The selected peptides are then deglycosylated by PNGase F in H₂¹⁶O and/or H₂¹⁸O to incorporate of a stable isotope tag for facilitate identification¹⁰⁰ or subsequent relative quantification of samples or a mixture.¹⁰¹ This approach was called isotope-coded glycosylation-specific tagging (IGOT)¹⁰¹ and has been employed for the identification of glycosylation sites in the human bile.¹⁰²

R. Aebersold and co-workers oxidised cis-diol group with periodate and the glycoproteins are coupled to hydrazide beads (Scheme 24). The glycoproteins are then digested on the beads. After washing, only the glycosylated peptides remain on the beads. The relative quantification of two samples is then realised by using succinic anhydride and its deuterated analogue. The released of the peptides is done with PNGase F, followed by LC-ESI-MS/MS or LC-MALDI-MS.¹⁰³



Scheme 24. Oxidation of a carbohydrate to an aldehyde followed by covalent coupling to hydrazide resin.

Like phosphate groups, the O-linked glycosides can be subjected to β -elimination under basic conditions. The 1,4-Michael addition of DTT or biotin pentylamine (BAP) allows the isolation of glycopeptides on thiol-Sepharose column or by biotin-avidin affinity followed respectively by MALDI-MS or LC-ESI-MS/MS.¹⁰⁴

4.3. Tyrosine nitration

Nitrotyrosine is often found in proteins and plays a direct role in cellular signalling. The variation in nitrotyrosine concentration has been linked to oxidative stress diseases for instance.¹⁰⁵ The approach described by G. Nikov *et al.* consists in the reduction of the nitrotyrosine to aminotyrosine using $\text{Na}_2\text{S}_2\text{O}_4$.¹⁰⁶ The amine then created, the $\text{p}K_a$ of which is lower than that of lysine, is specifically linked to a cleavable biotin tag. After selection by avidin affinity chromatography, the tag is cleaved and the peptides further analysed by MS.

5. Conclusion

This short overview of the currently chemical modifications for MS-based proteomics demonstrates the ample use and the possibilities that offer such derivatization steps for the reduction of sample complexity, the identification of protein candidates, the quantification of protein amounts between two extracts. The protein topology mapping, *de novo* sequencing or interpretation of tandem mass spectra is also possible. Nevertheless, all these tagging techniques have a defined dynamic range whereas the expression profile of the proteins covers a few orders of magnitude. For lower abundant protein, the tagging reactions can fail to provide a linear response. The relative quantification may so be inaccurate all the more so as side reactions can result from the tagging reaction.

In this chapter, the relevance and the basis of chemical modifications have been proposed. The following of the dissertation is focused on the development of a new

derivatization method using the inherent electrochemistry of the electrospray. The on-line electrochemical tagging (EC-tagging) will be shown to allow model proteomic application.

Bibliography

- 1 H. Mirzaei and F. Regnier, *J. Chromatogr. B*, 2005, 817, 23-34.
- 2 A. Leitner and W. Lindner, *J. Chromatogr. B*, 2004, 813, 1-26.
- 3 B. Herbert, M. Galvani, M. Hamdan, E. Olivieri, J. MacCarthy, S. Pedersen and P.G. Righetti, *Electrophoresis*, 2001, 22, 2046-2057.
- 4 S. Sechi and B.T. Chait, *Anal. Chem.*, 1998, 70, 5150-5158.
- 5 S. Neitz, M. Jurgens, M. Kellmann, P. Schulz-Knappe and M. Schrader, *Rapid Commun. Mass Spectrom.*, 2001, 15, 1586-1592.
- 6 F. Hubalek, J. Pohl and D.E. Edmondson, *J. Biol. Chem.*, 2003, 278, 28612-28618.
- 7 G. Gorin, P.A. Matic and G. Doughty, *Arch. Biochem. Biophys.*, 1966, 115, 593-597.
- 8 D.G. Smyth, W. Konigsberg and O.O. Blumenfeld, *Biochem. J.*, 1964, 91, 589-595.
- 9 C.F. Brewer and J.P. Riehm, *Anal. Biochem.*, 1967, 18, 248-255.
- 10 A. Sinz, *J. Mass Spectrom.*, 2003, 38, 1225-1237.
- 11 E.J. Bures, J.O. Hui, Y. Young, D.T. Chow, V. Katta, M.F. Rohde, L. Zeni, R.D. Rosenfeld, K.L. Stark and M. Haniu, *Biochemistry*, 1998, 37, 12172-12177.
- 12 G.B. Hegy, C.H.L. Shackleton, M. Carlquist, T. Bonn, O. Engstrom, P. Sjöholm and H.E. Witkowska, *Steroids*, 1996, 61, 367-373.
- 13 S.H. Wang, X. Zhang and F.E. Regnier, *J. Chromatogr. A*, 2002, 949, 153-162.
- 14 D.Y. Ren, S. Julka, H.D. Inerowicz and F.E. Regnier, *Anal. Chem.*, 2004, 76, 4522-4530.
- 15 K. Gevaert, B. Ghesquiere, A. Staes, L. Martens, J. Van Damme, G.R. Thomas and J. Vandekerckhove, *Proteomics*, 2004, 4, 897-908.
- 16 S. Sechi, *Rapid Commun. Mass Spectrom.*, 2002, 16, 1416-1424.
- 17 S. Gehanne, D. Cecconi, L. Carboni, P.G. Righetti, E. Domenici and M. Hamdan, *Rapid Commun. Mass Spectrom.*, 2002, 16, 1692-1698.
- 18 R. Sebastiano, A. Citterio, M. Lapadula and P.G. Righetti, *Rapid Commun. Mass Spectrom.*, 2003, 17, 2380-2386.
- 19 S. Niwayama, S. Kurono and H. Matsumoto, *Bioorg. Med. Chem. Lett.*, 2001, 11, 2257-2261.

- 20 M. Shen, L. Guo, A. Wallace, J. Fitzner, J. Eisenman, E. Jacobson and R.S. Johnson, *Mol. Cell. Proteomics*, 2003, 2, 315-324.
- 21 C. Pasquarello, J.C. Sanchez, D.F. Hochstrasser and G.L. Corthals, *Rapid Commun. Mass Spectrom.*, 2004, 18, 117-127.
- 22 R.A. Jue and J.E. Hale, *Anal. Biochem.*, 1994, 221, 374-378.
- 23 R.A. Jue and J.E. Hale, *Anal. Biochem.*, 1993, 210, 39-44.
- 24 P.C. White, N.S. Lawrence, J. Davis and R.G. Compton, *Anal. Chim. Acta*, 2001, 447, 1-10.
- 25 D.K. Han, J. Eng, H.L. Zhou and R. Aebersold, *Nat. Biotechnol.*, 2001, 19, 946-951.
- 26 S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, 17, 994-999.
- 27 http://prowl.rockefeller.edu/profound_bin/WebProFound.exe.
- 28 http://www.matrixscience.com/search_form_select.html.
- 29 S.P. Gygi, B. Rist and R. Aebersold, *Curr. Opin. Biotechnol.*, 2000, 11, 396-401.
- 30 R.J. Zhang, C.S. Sioma, R.A. Thompson, L. Xiong and F.E. Regnier, *Anal. Chem.*, 2002, 74, 3662-3669.
- 31 R.J. Zhang and F.E. Regnier, *J. Proteome Res.*, 2002, 1, 139-147.
- 32 H.L. Zhou, J.A. Ranish, J.D. Watts and R. Aebersold, *Nat. Biotechnol.*, 2002, 20, 512-515.
- 33 Y. Oda, T. Owa, T. Sato, B. Boucher, S. Daniels, H. Yamanaka, Y. Shinohara, A. Yokoi, J. Kuromitsu and T. Nagasu, *Anal. Chem.*, 2003, 75, 2159-2165.
- 34 K.C. Hansen, G. Schmitt-Ulms, R.J. Chalkley, J. Hirsch, M.A. Baldwin and A.L. Burlingame, *Mol. Cell. Proteomics*, 2003, 2, 299-314.
- 35 Y. Lu, P. Bottari, F. Turecek, R. Aebersold and M.H. Gelb, *Anal. Chem.*, 2004, 76, 4104-4111.
- 36 Y.C. Qiu, E.A. Sousa, R.M. Hewick and J.H. Wang, *Anal. Chem.*, 2002, 74, 4969-4979.
- 37 J.V. Olsen, J.R. Andersen, P.A. Nielsen, M.L. Nielsen, D. Figeys, M. Mann and J.R. Wisniewski, *Mol. Cell. Proteomics*, 2004, 3, 82-92.
- 38 E. Krause, H. Wenschuh and P.R. Jungblut, *Anal. Chem.*, 1999, 71, 4160-4165.

- 39 E.C. Peters, D.M. Horn, D.C. Tully and A. Brock, *Rapid Commun. Mass Spectrom.*, 2001, 15, 2387-2392.
- 40 F.L. Brancia, A. Butt, R.J. Beynon, S.J. Hubbard, S.J. Gaskell and S.G. Oliver, *Electrophoresis*, 2001, 22, 552-559.
- 41 J.A. Karty, M.M.E. Ireland, Y.V. Brun and J.P. Reilly, *J. Proteome Res.*, 2002, 1, 325-335.
- 42 T. Keough, M.P. Lacey and R.S. Youngquist, *Rapid Commun. Mass Spectrom.*, 2000, 14, 2348-2356.
- 43 G. Cagney and A. Emili, *Nat. Biotechnol.*, 2002, 20, 163-170.
- 44 F.L. Brancia, M.E. Openshaw and S. Kumashiro, *Rapid Commun. Mass Spectrom.*, 2002, 16, 2255-2259.
- 45 F.L. Brancia, H. Montgomery, K. Tanaka and S. Kumashiro, *Anal. Chem.*, 2004, 76, 2748-2755.
- 46 R.L. Beardsley and J.P. Reilly, *J. Proteome Res.*, 2003, 2, 15-21.
- 47 D. Suckau, M. Mak and M. Przybylski, *Proc. Natl. Acad. Sci. U.S.A.*, 1992, 89, 5630-5634.
- 48 M.O. Glocker, C. Borchers, W. Fiedler, D. Suckau and M. Przybylski, *Bioconjugate Chem.*, 1994, 5, 583-590.
- 49 G.J. Carven and L.J. Stern, *Biochemistry*, 2005, 44, 13625-13637.
- 50 H. Cui, J. Leon, E. Reussaet and A. Bult, *J. Chromatogr. A*, 1995, 704, 27-36.
- 51 S. Dikler, J.W. Kelly and D.H. Russell, *J. Mass Spectrom.*, 1997, 32, 1337-1349.
- 52 B. Spengler, F. Luetzenkirchen, S. Metzger, P. Chaurand, R. Kaufmann, W. Jeffery, M. Bartlett-Jones and D.J.C. Pappin, *Int. J. Mass Spectrom.*, 1997, 169, 127-140.
- 53 A. Leitner and W. Lindner, *Anal. Chim. Acta*, 2005, 528, 165-173.
- 54 A. Leitner and W. Lindner, *Anal. Chem.*, 2005, 77, 4481-4488.
- 55 A. Leitner and W. Lindner, *J. Mass Spectrom.*, 2003, 38, 891-899.
- 56 H. Kuyama, M. Watanabe, C. Toda, E. Ando, K. Tanaka and O. Nishimura, *Rapid Commun. Mass Spectrom.*, 2003, 17, 1642-1650.
- 57 K. Gevaert, J. Van Damme, M. Goethals, G.R. Thomas, B. Hoorelbeke, H. Demol, L. Martens, M. Puype, A. Staes and J. Vandekerckhove, *Mol. Cell. Proteomics*, 2002, 1, 896-903.

- 58 G.E. Reid, K.D. Roberts, R.J. Simpson and R.A.J. O'Hair, *J. Am. Soc. Mass Spectrom.*, 2005, 16, 1131-1150.
- 59 M. Kalkum, M. Przybylski and M.O. Glocker, *Bioconjugate Chem.*, 1998, 9, 226-235.
- 60 D.Y. Ren, N.A. Penner, B.E. Slentz and F.E. Regnier, *J. Proteome Res.*, 2004, 3, 37-45.
- 61 K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas and J. Vandekerckhove, *Nat. Biotechnol.*, 2003, 21, 566-569.
- 62 K. Kuhn, A. Thompson, T. Prinz, J. Muller, C. Baumann, G. Schmidt, T. Neumann and C. Hamon, *J. Proteome Res.*, 2003, 2, 598-609.
- 63 T. Prinz, J. Muller, K. Kuhn, J. Schafer, A. Thompson, J. Schwarz and C. Hamon, *J. Proteome Res.*, 2004, 3, 1073-1081.
- 64 D. Chelius and T.A. Shaler, *Bioconjugate Chem.*, 2003, 14, 205-211.
- 65 M. Munchbach, M. Quadroni, G. Miotto and P. James, *Anal. Chem.*, 2000, 72, 4047-4057.
- 66 L. Riggs, E.H. Seeley and F.E. Regnier, *J. Chromatogr. B*, 2005, 817, 89-96.
- 67 A. Chakraborty and F.E. Regnier, *J. Chromatogr. A*, 2002, 949, 173-184.
- 68 Y.K. Wang, Z.X. Ma, D.F. Quinn and E.W. Fu, *Anal. Chem.*, 2001, 73, 3742-3750.
- 69 M. Hardt, H.E. Witkowska, S. Webb, L.R. Thomas, S.E. Dixon, S.C. Hall and S.J. Fisher, *Anal. Chem.*, 2005, 77, 4947-4954.
- 70 L. DeSouza, G. Diehl, M.J. Rodrigues, J.Z. Guo, A.D. Romaschin, T.J. Colgan and K.W.M. Siu, *J. Proteome Res.*, 2005, 4, 377-386.
- 71 L.H. Choe, K. Aggarwal, Z. Franck and K.H. Lee, *Electrophoresis*, 2005, 26, 2437-2449.
- 72 K. Aggarwal, L.H. Choe and K.H. Lee, *Proteomics*, 2005, 5, 2297-2308.
- 73 J.L. Hsu, S.Y. Huang, N.H. Chow and S.H. Chen, *Anal. Chem.*, 2003, 75, 6843-6852.
- 74 D.E. Mason and D.C. Liebler, *J. Proteome Res.*, 2003, 2, 265-272.
- 75 T. Keough, R.S. Youngquist and M.P. Lacey, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, 96, 7131-7136.
- 76 J.R. Strahler, Y. Smelyanskiy, G. Lavine and J. Allison, *Int. J. Mass Spectrom.*, 1997, 169, 111-126.

- 77 Z.H. Huang, J. Wu, K.D.W. Roth, Y. Yang, D.A. Gage and J.T. Watson, *Anal. Chem.*, 1997, 69, 137-144.
- 78 D.R. Goodlett, A. Keller, J.D. Watts, R. Newitt, E.C. Yi, S. Purvine, J.K. Eng, P. von Haller, R. Aebersold and E. Kolker, *Rapid Commun. Mass Spectrom.*, 2001, 15, 1214-1221.
- 79 T. Nakazawa, M. Yamaguchi, K. Nishida, H. Kuyama, T. Obama, E. Ando, T. Okamura, N. Ueyama, K. Tanaka and S. Norioka, *Rapid Commun. Mass Spectrom.*, 2004, 18, 799-807.
- 80 M.F. Byford, *Biochem. J.*, 1991, 280, 261-265.
- 81 C. Klemm, S. Schroder, M. Gluckmann, M. Beyermann and E. Krause, *Rapid Commun. Mass Spectrom.*, 2004, 18, 2697-2705.
- 82 H.E. Meyer, E. Hoffmannposorske, H. Korte and L.M.G. Heilmeyer, *FEBS Lett.*, 1986, 204, 61-66.
- 83 H. Jaffe, Veeranna and H.C. Pant, *Biochemistry*, 1998, 37, 16211-16224.
- 84 C.F.B. Holmes, *FEBS Lett.*, 1987, 215, 21-24.
- 85 M. Adamczyk, J.C. Gebler and J. Wu, *Rapid Commun. Mass Spectrom.*, 2001, 15, 1481-1488.
- 86 W. Weckwerth, L. Willmitzer and O. Fiehn, *Rapid Commun. Mass Spectrom.*, 2000, 14, 1677-1681.
- 87 M.P. Molloy and P.C. Andrews, *Anal. Chem.*, 2001, 73, 5387-5394.
- 88 Y. Oda, T. Nagasu and B.T. Chait, *Nat. Biotechnol.*, 2001, 19, 379-382.
- 89 M.B. Goshe, T.D. Veenstra, E.A. Panisko, T.P. Conrads, N.H. Angell and R.D. Smith, *Anal. Chem.*, 2002, 74, 607-616.
- 90 M.B. Goshe, T.P. Conrads, E.A. Panisko, N.H. Angell, T.D. Veenstra and R.D. Smith, *Anal. Chem.*, 2001, 73, 2578-2586.
- 91 A.J. Thompson, S.R. Hart, C. Franz, K. Barnouin, A. Ridley and R. Cramer, *Anal. Chem.*, 2003, 75, 3232-3243.
- 92 H.L. Zhou, J.D. Watts and R. Aebersold, *Nat. Biotechnol.*, 2001, 19, 375-378.
- 93 D.T. McLachlin and B.T. Chait, *Anal. Chem.*, 2003, 75, 6826-6836.
- 94 A. Amoresano, G. Marino, C. Cirulli and E. Quemeneur, *Eur. J. Mass Spectrom.*, 2004, 10, 401-412.

- 95 H. Steen and M. Mann, *J. Am. Soc. Mass Spectrom.*, 2002, 13, 996-1003.
- 96 http://www.cco.caltech.edu/~ppmal/sample_prep/term_def.html.
- 97 T.W. Rademacher, R.B. Parekh and R.A. Dwek, *Annu. Rev. Biochem.*, 1988, 57, 785-838.
- 98 L. Xiong, D. Andrews and F. Regnier, *J. Proteome Res.*, 2003, 2, 618-625.
- 99 R.Q. Qiu and F.E. Regnier, *Anal. Chem.*, 2005, 77, 2802-2809.
- 100 L. Xiong and F.E. Regnier, *J. Chromatogr. B*, 2002, 782, 405-418.
- 101 H. Kaji, H. Saito, Y. Yamauchi, T. Shinkawa, M. Taoka, J. Hirabayashi, K. Kasai, N. Takahashi and T. Isobe, *Nat. Biotechnol.*, 2003, 21, 667-672.
- 102 T.Z. Kristiansen, J. Bunkenborg, M. Gronborg, H. Molina, P.J. Thuluvath, P. Argani, M.G. Goggins, A. Maitra and A. Pandey, *Mol. Cell. Proteomics*, 2004, 3, 715-728.
- 103 H. Zhang, X.J. Li, D.B. Martin and R. Aebersold, *Nat. Biotechnol.*, 2003, 21, 660-666.
- 104 L. Wells, K. Vosseller, R.N. Cole, J.M. Cronshaw, M.J. Matunis and G.W. Hart, *Mol. Cell. Proteomics*, 2002, 1, 791-804.
- 105 S.A.B. Greenacre and H. Ischiropoulos, *Free Radic. Res.*, 2001, 34, 541-581.
- 106 G. Nikov, V. Bhat, J.S. Wishnok and S.R. Tannenbaum, *Anal. Biochem.*, 2003, 320, 214-222.

CHAPTER III. *Hydroquinone probes for on-line electrochemical tagging of cysteine residues: mechanism and kinetics*[§]

1. Introduction

The term proteome appeared in the nineties to design the ensemble of proteins related to a genome. The study of the proteome (*i.e.* proteomics) has seen a tremendous growth in popularity over the last decade as illustrated by the number of publication related to this field. Protein structure, function and interactions is the focus of many research studies in the field of analytical chemistry.^{1,2} The discovery and the development of soft ionization techniques in mass spectrometry (MS) has led to new tools for studying peptides and proteins.³ MS-based study of proteins is a discipline made possible by the availability of gene and genome sequence databases and technical and conceptual advances in many areas.⁴

The amino acid sequence of a given peptide or protein can be elucidated by MS or tandem MS (MS/MS) techniques either on the native biomolecule (“top down”) or on the tryptic digest of the biomolecule (“bottom up”). When dealing with complex samples, it is often advantageous to couple a separation step to the MS analysis as illustrated by the success of hyphenated techniques incorporating for instance gel electrophoresis (GE, including off-gel electrophoresis)^{5,6}, capillary electrophoresis (CE)⁷ or micro liquid chromatography (μ LC).⁸

[§] based on L. Dayon, C. Roussel and H. H. Girault, *Chimia*, 2004, 58, 204-207 as an invited article and C. Roussel, L. Dayon, H. Jensen and H. H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.

To reduce the complexity of the protein mixtures, sample preparation steps have been developed to target specific amino acids and post-translational modifications by so-called tagging strategies. These strategies have also opened the way to relative quantification using stable-isotope-labelled tags. In this approaches, nucleophilic amino acids such as cysteine have received a special attention due to their reactivity against various electrophilic compounds. Thus, various alkylation compounds such as iodoacetamide, acrylamide or different vinylpyridines have been developed to react specifically with cysteine.^{4,9-12} Nevertheless, in many cases, the alkylation process is a time consuming step¹¹ and can lead to side reactions with the medium when the optimum reaction time is exceeded.

The chemical tagging of proteins is usually carried out as a simple addition of a molecular tag to a solution containing the protein(s) of interest followed by product analysis. These methods may therefore be termed sequential in the sense that tagging and analysis are performed as two independent steps.

Recently, a new dynamic method based on an electrochemically induced coupling of an electrogenerated electrophile (benzoquinone) with free cysteine units in proteins has been introduced.¹³⁻¹⁵ In this method, the electrochemical tagging (EC-tagging) is performed within a microspray interface for MS, thus preventing an additional analysis step. More importantly, it has been shown that the geometrical and operational parameters of the specifically designed microspray chip containing a carbon microelectrode can effectively be used to control the tagging efficiency.¹³ The electrochemical mechanism occurring in the microspray has been elucidated, and the rate constant pertaining to the reaction of L-cysteine with benzoquinone determined.¹⁵

The objective of the present work has been to further develop this method by characterising new tags based on substituted hydroquinones and to improve the EC-tagging reaction. A series of mono-substituted hydroquinones (i.e. methyl-, methoxy-, methoxycarbonyl- and nitro-1,4-hydroquinone) have been studied by electrochemical techniques in order to describe the mechanistic details of the EC-tagging. Furthermore, the site-specific rate constants corresponding to the homogeneous reaction between the electrogenerated substituted benzoquinones and L-cysteine have been determined and

analysed. Finally, the microspray MS tagging efficiencies have been studied using L-cysteine and β -lactoglobulin A as substrates. For the latter, it is concluded that all the electrogenerated tags react selectively with the free cysteine residue of β -lactoglobulin A under the experimental conditions used in the microspray MS experiments. Investigations on the on-line generation of electrophiles reacting with free cysteine residues are reported.

2. Experimental

2.1. Reagents

L-cysteine ($\geq 99.5\%$), myoglobin from Horse heart ($> 90\%$) acetic acid ($> 99.5\%$), lithium trifluoromethanesulfonate (purum), silver nitrate ($> 99.5\%$), potassium hydroxide ($\geq 86\%$) and diethylether ($> 99.5\%$) were purchased from Fluka (Buchs, Switzerland). Methanol ($\geq 99.8\%$) and chloroform (99.0-99.4%) were purchased from Riedel-de Haën (Seelze, Germany). Methanol ($\geq 99.8\%$) and sodium sulfate ($\geq 99\%$) were purchased from Merck (Darmstadt, Germany). Methyl-1,4-hydroquinone (99%) and methoxy-1,4-hydroquinone (99%) were purchased from Acros (Pittsburgh, PA, USA). Methoxycarbonyl-1,4-hydroquinone (methyl-2,5-dihydroxybenzoate 99%) was purchased from Aldrich (Milwaukee, WI, USA). Nitro-1,4-hydroquinone was purchased from Frinton Laboratories (Vineland, NJ, USA). Bradykinin (99%), Met-enkephalin-Arg-Phe (98%) and β -lactoglobulin A (from Bovine milk) were purchased from Sigma (St. Louis, MO, USA). Leu-enkephalin ($> 99\%$) was purchased from Bachem (Bubendorf, Switzerland). We have used deionised water (18.2 M Ω -cm, Milli-Q system, Millipore (Bedford, MA, USA)).

Concerning the used peptides and proteins, the amino acid composition is given below: Leu-enkephalin (G (2), F, L, Y), Met-enkephalin-Arg-Phe (R, G (2), M, F (2), Y), bradykinin (R (2), G, F (2), P (3), S), β -lactoglobulin A (A (14), C (1 free and 4 involved in disulfide bonds), D (11), E (16), F (4), G (3), H (2), I (10), K (15) L (22), M (4), N (5), P (8), Q (9), R (3), S (7), T (8), V (10), W (2), Y (4)) and myoglobin (A (15), D (8), E (13), F (7), G

(15), H (11), I (9), K (19), L(17), M (2), N (2), P (4), Q (6), R (2), S (5), T (7), V(7), W (2), Y (2)).

2.2 Chemical and electrochemical set-up

UV-Visible spectroscopy: UV-Visible spectra were recorded on an Ocean Optics (Dunedin, FL, USA) UV-Visible spectrometer (Chem 2000-UV-VIS). The samples were diluted at 50 μM in CHCl_3 or at 250 μM in $\text{MeOH} / \text{H}_2\text{O} / \text{AcOH}$ 50% / 49% / 1%.

^1H NMR spectroscopy: ^1H NMR spectra were recorded on a Bruker-Spectrospin AC 200 spectrometer (Fälladen, Switzerland) operating at 200 MHz for ^1H NMR and 50 MHz for ^{13}C NMR. Chemical shifts were measured relative to TMS in CDCl_3 or ($\text{CD}_3\text{OD} / \text{D}_2\text{O} / \text{CD}_3\text{COOD}$ 50% / 49% / 1%) as solvent. The chemical shift, coupling constants and integrals (5% relative accuracy) obtained for the adducts (chemically generated, see 2.3.) were used to determine the product composition and relative yields as previously described for the addition of thiols on benzoquinones.¹⁶

Mass Spectrometry: Mass spectra were recorded on a Finnigan LCQ duo Ion Trap apparatus (San José, CA, USA) using a specially designed microchip. The electrode was $70 \times 25 \mu\text{m}^2$, the channel was $35 \times 30 \mu\text{m}^2$, the distance from the electrode to the tip was 1.8-2 cm. The flow rate was $250 \text{ nL}\cdot\text{min}^{-1}$. The applied voltage was 3.5-4 kV. The concentration of L-cysteine and hydroquinones were 0.2 mM and 20 mM respectively prepared in degassed solutions and each experiment was repeated at least three times. The concentrations of β -lactoglobulin A and the hydroquinones were 5 μM / 5 mM and 5 μM / 2.5 mM. The concentrations of peptide and hydroquinones were 50 μM and 20 mM. The chip was fabricated as previously described¹⁷ (see also appendix I for a comprehensive description). In the MS experiments, no supporting electrolyte was used and in all cases the solutions were purged with nitrogen to remove oxygen. It is rather important to work with deoxygenated solutions as small amounts of oxygen may oxidise the hydroquinone into the benzoquinone form, and thereby cause unwanted tagging reactions to occur. The mass spectra were deconvoluted using the biomass deconvolution software from Thermo Finnigan.

Electrochemistry: Analytical experiments were performed in an undivided cell fitted with a saturated calomel reference electrode (SCE), glassy carbon working electrode (diameter 3 mm) and a platinum wire counter electrode, using as electrochemical solvent a mixture of (MeOH / H₂O / AcOH 50% / 49% / 1%) with 0.1 M of lithium trifluoromethanesulfonate. The solvent was degassed under nitrogen. The cyclic voltammograms (CV's), linear sweep voltammograms (LSV's) and square wave voltammograms (SWV's) were recorded on a Metrohm Autolab PGSTAT 30 (Herisau, Switzerland) and the different conditions are reported on the different figures. The carbon electrode was carefully polished before each potential scan.

2.3. Synthesis

Benzoquinones 2b-d: The synthesis of the benzoquinones was performed according to literature procedures.^{18,19} The first step consists in the synthesis of the oxidative agent Ag₂O from AgNO₃.

In a 500 mL flask, 2 g (12 mmol) of AgNO₃ is dissolved in 200 mL of distilled water. The flask is then put in an ice bath for 5 minutes under stirring. 673 mg (12 mmol) of solid KOH is added in small portions. Stirring is maintained for 15 minutes. The formed solid of Ag₂O (1.23 g, 91%) is filtrated on a glass-fritted disk and washed three times with distilled water and put in the oven at 80°C for 2 hours.

To a solution of 1 mmol of hydroquinone **1** in 3.3 mL of Et₂O containing 413 mg (2.91 mmol, 2.91 eq) of Na₂SO₄, was added 777 mg (3.35 mmol, 3.35 eq) of Ag₂O. The mixture was subsequently stirred for 3 hours. After filtration, the collected solvent fractions were evaporated in *vacuum* at room temperature. The obtained precipitate was dissolved in *n*-hexane under ultrasonic irradiation. Subsequently, the solution was cooled using an ice bath, leading to a pure crystallized solid.

Methoxy-1,4-benzoquinone **2b**: 72 mg, (52%) of a yellow solid: ¹H NMR (CDCl₃) δ 3.85 ppm (s, 3H, O-CH₃), 5.95 ppm (m, 1H, 3-H), 6.75 ppm (m, 2H, 5-H and 6-H); UV-Visible (CHCl₃) λ_{max} = 253 nm.

Methyl-1,4-benzoquinone **2c**: 46 mg (37%) of a yellow solid: ^1H NMR (CDCl_3) δ 2.10 ppm (s, 3H, CH_3), 6.65 ppm (m, 1H, 3-H), 6.75 ppm (m, 2H, 5-H and 6-H); UV-Visible (CHCl_3) $\lambda_{\text{max}} = 248$ nm.

Methoxycarbonyl-1,4-benzoquinone **2d**: 106 mg (64%) of a orange solid: ^1H NMR (CDCl_3) δ 3.95 ppm (s, 3H, O- CH_3), 6.85 ppm (m, 2H, 5-H and 6-H), 7.15 ppm (m, 1H, 3-H); UV-Visible (CHCl_3) $\lambda_{\text{max}} = 249$ nm.

Adducts **4b-d**: The pure benzoquinone form **2** (5 mg) and cysteine **3** (5 mg) were dissolved in 600 μL of deuterated solvent (CD_3OD / D_2O / CD_3COOD 50% / 49% / 1%) and give the following results for the aromatic protons:

Adducts coming from methoxy-1,4-benzoquinone **2b**: A mixture of the 5-position (47%, $\delta = 6.65$ ppm (s, 1H, H^6), $\delta = 7.05$ ppm (s, 1H, H^3)) and 6-position (53%, $\delta = 6.55$ ppm (d, 1H, $J = 2.5$ Hz, H^5), $\delta = 6.60$ ppm (d, 1H, $J = 2.5$ Hz, H^3)) was obtained.

Adducts coming from methyl-1,4-benzoquinone **2c**: A mixture of the 6-position (74%, $\delta = 6.70$ ppm (d, 1H, $J = 2.7$ Hz, H^3), $\delta = 6.90$ ppm (d, 1H, $J = 2.7$ Hz, H^5)) and 5-position (26%, $\delta = 6.80$ ppm (s, 1H, H^3), $\delta = 7.00$ ppm (s, 1H, H^6)) was obtained.

Adducts coming from methoxycarbonyl-1,4-benzoquinone **2d**: The addition occurs on position 3 (85%, $\delta = 6.95$ ppm (d, 1H, $J = 8.9$ Hz, H^6), $\delta = 7.00$ ppm (d, 1H, $J = 8.9$ Hz, H^5)), 6 (13%, $\delta = 7.35$ ppm (d, 1H, $J = 2.9$ Hz, H^5), $\delta = 7.35$ ppm (d, 1H, $J = 2.9$ Hz, H^3)) and 5 (2%, $\delta = 6.90$ ppm (s, 1H, H^6), $\delta = 7.30$ ppm (s, 1H, H^3)). In all cases, the regiochemistry of addition of L-cysteine was determined using the values of the coupling constant of the two protons present in **4b-d** (i.e. J_{ortho} : 6-10 Hz, J_{meta} : 1-3 Hz, J_{para} : 0-1 Hz).

2.4. Digital simulations

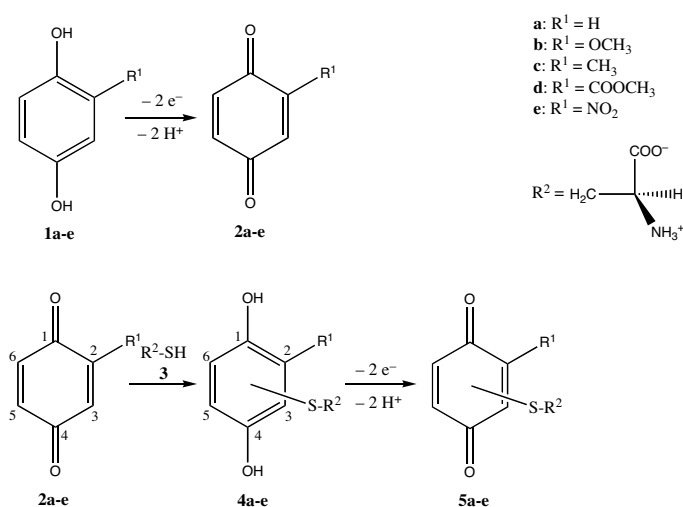
Electrochemical simulations were carried out using the commercially available software Digisim 3.0 from Bioanalytical Systems (West Lafayette, IN, USA). The simulation is based on an expanding grip (expanding grid factor 0.5) and the Rudolph algorithm²⁰ using

voltage step of 1.8 mV. No IR drop was assumed (the experiments were performed using IR compensation). A simple mechanism was assumed in which benzoquinone is generated in a two electron irreversible oxidation (the standard heterogeneous rate constant was chosen so that the CV was irreversible at all scan rates used) followed by the addition of L-cysteine. The adduct was assumed to be oxidised at a potential similar to that of the hydroquinone. The standard curves were obtained by plotting I_p / I_p^0 (I_p is the current in the presence of L-cysteine and I_p^0 is the current in the absence of L-cysteine) as a function of the scan rate. A given rate constant was assumed for each excess factor (i.e. concentration of L-cysteine divided by the hydroquinone concentration).

3. Results and discussion

3.1. Electrochemical tagging of L-cysteine

The EC-tagging of L-cysteine has previously been studied using non-substituted hydroquinone¹⁵ as well as other thiols and experimental conditions.^{15,21-27} The overall mechanism can be described as in Scheme 1 for the experimental conditions used in the present work.



Scheme 1. Mechanism of the EC-tagging of substituted hydroquinones by L-cysteine.

Depending on the acidity of the solution, the electrogenerated benzoquinone electrophile may be present in a protonated form. The compounds under study are 1,4-hydroquinone **1a**, methoxy-1,4-hydroquinone **1b**, methyl-1,4-hydroquinone **1c**, methoxycarbonyl-1,4-hydroquinone **1d** and nitro-1,4-hydroquinone **1e** respectively. It is expected that an electron pair attracting group will increase the electrophilicity of the electrogenerated substituted benzoquinone (relative to the non-substituted one) whereas an electron pair donating group will decrease it.²⁸

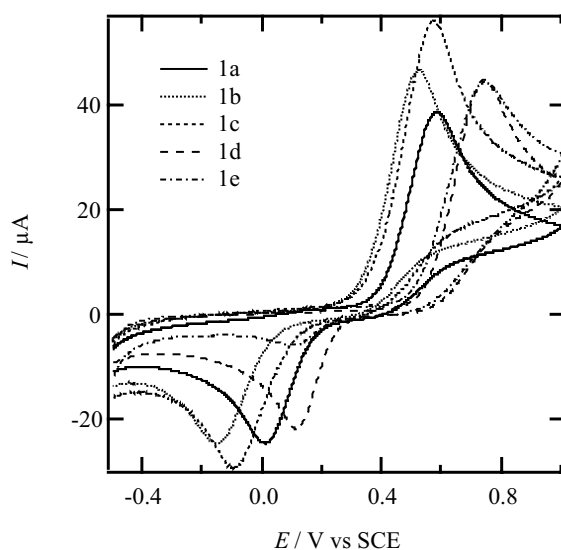


Figure 1. Cyclic voltammograms recorded at $75 \text{ mV}\cdot\text{s}^{-1}$ using a glassy carbon working electrode of 2 mM of the hydroquinones **1a-1e** employed in the present study.

The electronic effects of the substituents can be studied by cyclic voltammetry of the substrates. In Figure 1 are shown the cyclic voltammograms (CV's) of the hydroquinones **1a-e** recorded in the microspray medium (i.e. MeOH / H₂O / AcOH 50% / 49% / 1%) at 75 mV/s using a glassy carbon working electrode in a classical electrochemical cell. The CV's appear electrochemically irreversible (slow electrode reaction kinetics) in the present medium with oxidation waves from 0.50 to 0.75 V vs. SCE and corresponding reduction waves of the benzoquinone forms from -0.15 to 0.11 V vs. SCE. The position of the oxidation and

reduction peaks depends on the electronic structure of the substituents. Indeed, as shown in Figure 2 the reduction and oxidation peak potentials (E_{pa}^0 and E_{pc}^0) correlate well with the Hammett coefficient, σ_p .

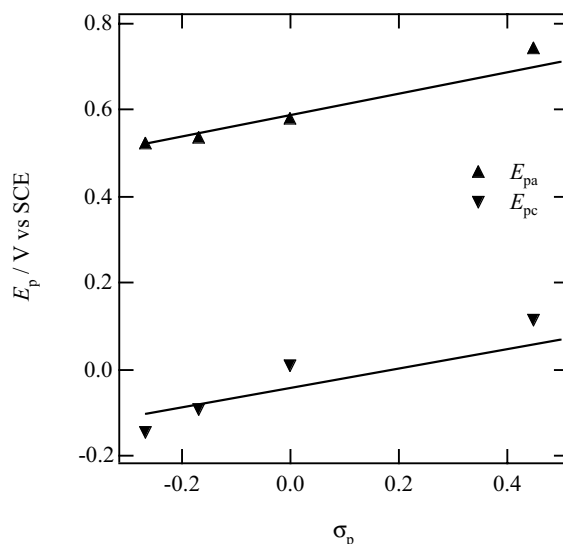


Figure 2. Correlation of the peak potentials recorded at a scan rate of $75 \text{ mV}\cdot\text{s}^{-1}$ corresponding to the oxidation of the hydroquinones **1a-d** (E_{pa}^0) and reduction of the benzoquinones **2a-d** (E_{pc}^0) with the Hammett para-substituent constant, σ_p , as tabulated in the literature.²⁹

The CV's displayed in Figure 1 indicate that the hydroquinones are all stable on the time scale of a single experiment. However, the peak corresponding to the reduction of the benzoquinone **2e** appears much smaller than the others suggesting a degradation of **2e** in the medium. This phenomenon was further confirmed by multiscan analysis performed over longer times and it was therefore decided to exclude **1e** from further electrochemical investigations. Multiscan experiments of the other substrates did not reveal degradation in the medium on the timescale of a CV (data not shown). According to the literature, the degradation process of **2e** may be ascribed to a polymerization reaction.³⁰ The long-term stability (during at least 3 hours) of the remaining hydroquinones was checked using CV and UV-spectroscopy. The corresponding chemically synthesized (see experimental) benzoquinones **2b-d** are stable in the microspray solution on the time scale of a CV, but after a few minutes, linear sweep voltammetry (LSV) and UV measurements indicated the

formation of degradation products of **2d**. NMR data obtained in the deuterated microspray medium ($\text{CD}_3\text{OD} / \text{D}_2\text{O} / \text{CD}_3\text{COOD}$ 50% / 49% / 1%) confirmed the observations.

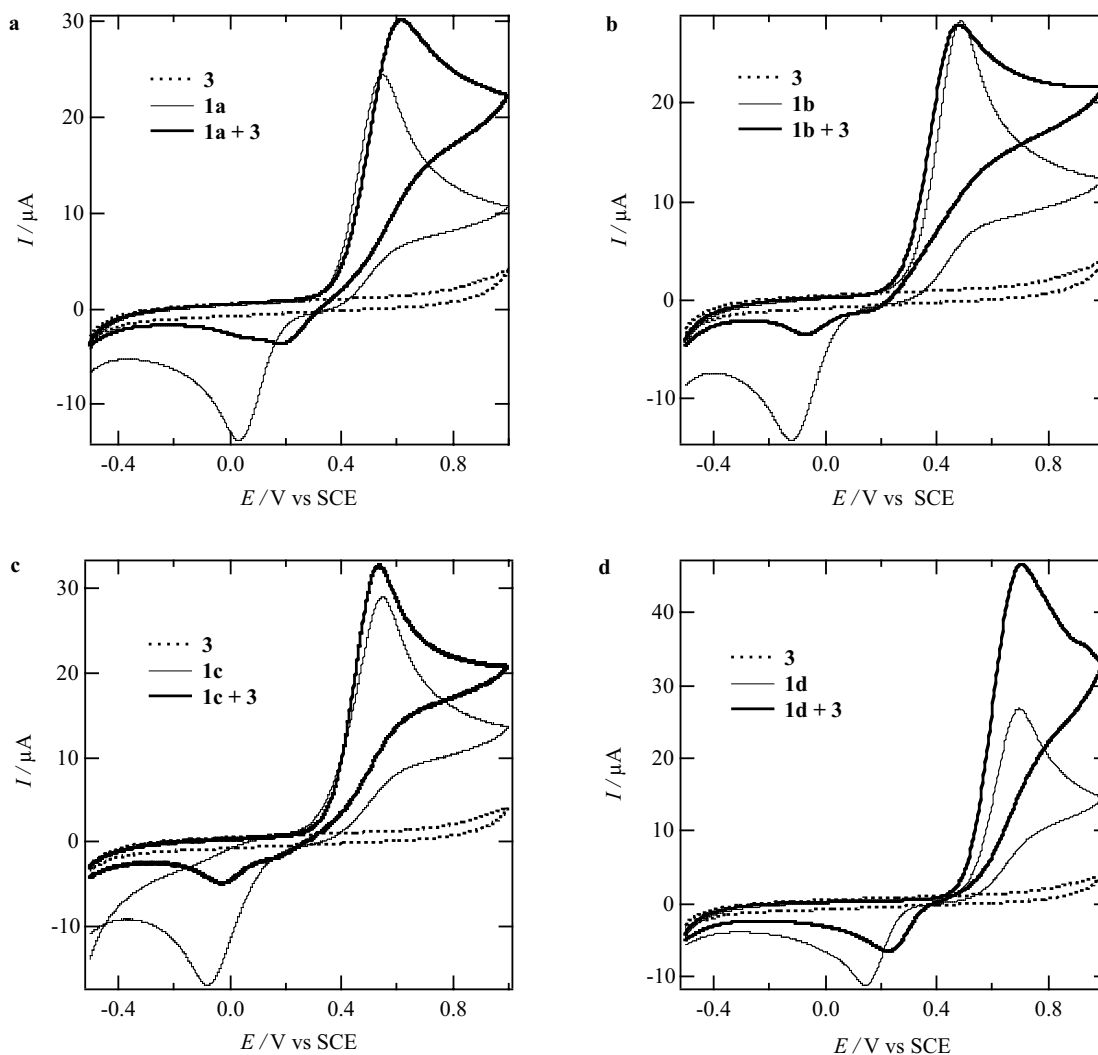


Figure 3. Cyclic voltammograms recorded at $25 \text{ mV}\cdot\text{s}^{-1}$ using a glassy carbon working electrode of 2 mM L-cysteine (dotted line), 2 mM of the hydroquinone (full line) and the hydroquinone (2 mM) in the presence of 2 mM L-cysteine (bold line). (a) 1,4-Hydroquinone **1a**. (b) Methoxy-1,4-hydroquinone **1b**. (c) Methyl-1,4-hydroquinone **1c**. (d) Methoxycarbonyl-1,4-hydroquinone **1d**.

When L-cysteine was added to the electrochemical cell, the electrochemical behaviour changed depending on the scan rate and amount of L-cysteine. As shown in Figure 3, the reduction wave that corresponds to the benzoquinone decreases but at the same time the

oxidation peak current corresponding to the hydroquinone increases. These characteristics are in accordance with an addition of L-cysteine on the electrogenerated benzoquinone form as previously described for the non-substituted hydroquinone.¹⁵ The electrochemical adduct behaviours were studied by recording the CV's of the products obtained from mixing an equimolar amount of the benzoquinone form and L-cysteine.

The adducts are oxidised at a potential very similar to the oxidation potential of the hydroquinone in agreement with a Hammett substituent constant close to zero for thioethers.²⁹ The reduction potential of the benzoquinone form of the adducts is also close to that of the non-substituted benzoquinone form. In order to investigate if a further addition of L-cysteine may take place, the CV's of the primary adduct were recorded in the presence of L-cysteine, but only small changes in the CV's were detected even at low scan rates. Furthermore, no signs of double addition products could be detected in microspray MS. The double addition will be neglected in the following.

From Figure 3, it can be observed that the oxidation peak current corresponding to the hydroquinones in the presence of L-cysteine is higher than in the absence of L-cysteine. This phenomenon is easily understood as the addition product is in fact oxidised at a potential similar to that of the substituted hydroquinones. The current increase in the presence of L-cysteine (*i.e.* I_p / I_p^0 , I_p^0 is the peak current in the absence of L-cysteine) can be used to determine the homogeneous rate constant corresponding to the addition of L-cysteine on the electrogenerated benzoquinones.¹⁵ The procedure for extracting the rate constant involves digital simulations of the CV's corresponding to the experimental conditions (*i.e.* scan rate and concentrations of hydroquinone and L-cysteine). In principle, the rate constants can then be extracted from a direct comparison of the CV's. However, as the reduction wave corresponding to the reduction of the benzoquinone adduct seems to be hampered by adsorption processes and/or degradation of some of the compounds, it was decided to focus on the analysis of the ratio I_p / I_p^0 as a function of the scan rate. Figure 4a and 4b show the experimental values (markers) for **1b** and **1d** as well as the corresponding simulated curves (lines). The details of the simulations are described in the experimental section.

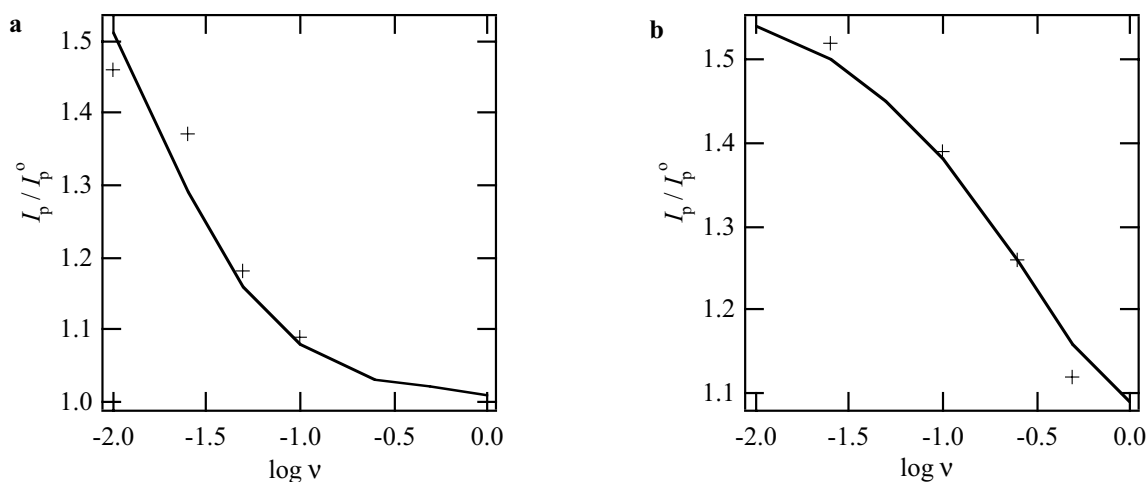


Figure 4. Simulated standard curves corresponding to the mechanism depicted in Scheme 1 (solid line). The markers correspond to experimental points. (a) Methoxy-1,4-hydroquinone (2 mM) and L-cysteine (8 mM). (b) Methoxycarbonyl-1,4-hydroquinone (2 mM) and L-cysteine (1 mM).

Compound	1a		1b		1c		1d	
k^1 ($M^{-1}\cdot s^{-1}$)	210 ¹		50 ¹		50 ¹		5000 ¹	
Position ²	Relative yield ³	Rate constant ⁴ ($M^{-1}\cdot s^{-1}$)	Relative yield ³	Rate constant ⁴ ($M^{-1}\cdot s^{-1}$)	Relative yield ³	Rate constant ⁴ ($M^{-1}\cdot s^{-1}$)	Relative yield ³	Rate constant ⁴ ($M^{-1}\cdot s^{-1}$)
3	25%	52.5	< 2%	n.d. ⁵	< 2%	n.d. ⁵	85%	4250
5	25%	52.5	47%	23.5	26%	13	2%	100
6	25%	52.5	53%	26.5	74%	37	13%	650

¹ The overall rate constant obtained by fitting to simulated standard curves as described in the text. ² The position on the benzoquinone ring as described in scheme 1. ³ Relative yields calculated from ¹H NMR data (relative accuracy $\pm 5\%$). ⁴ Local rate constants calculated from the relative yields and the overall rate constant. ⁵ Not determined due to the low abundance of this compound.

Table 1. The overall and local rate constants corresponding to the reaction between electrogenerated benzoquinones and L-cysteine.

In the case of **1c**, the fit to the simulated standard curves was of a very poor quality especially at high scan rates and consequently no reliable rate constants could be extracted using this method. However, it may be noted from Figure 3 that the shape of the CV's in the case of **1b** and **1c** appears very similar at low scan rates. In particular, the reduction peaks of both the adduct and the non-reacted benzoquinone appear in the cyclic voltammograms with similar intensities. Based on these observations, it can be concluded that the rate constants for addition of L-cysteine to methyl- and methoxy-1,4-hydroquinone are of the same order of magnitude. The different observed kinetics constants are assembled in Table 1.

3.2. ¹H NMR study of the addition on benzoquinone

According to the structure of each substituted benzoquinone, the cysteine addition can take place at three different positions (i.e. 3, 5 and 6 as defined in Scheme 1), whereas in the case of the non-substituted benzoquinone, there are four equivalent positions. In principle, the observed overall rate constants are therefore the sum of up to three rate constants corresponding to distinctly different reactions. The substitution patterns and individual rate constants can be determined by product studies of a quantitative conversion of the benzoquinone form with L-cysteine in the medium. ¹H NMR analysis of the product mixture has been used to provide information on product structure and relative amounts. The pure benzoquinone form **2** and cysteine **3** were mixed in equal quantity (w/w) in the previously mentioned microspray deuterated solvent (CD₃OD / D₂O / CD₃COOD 50% / 49% / 1%). The different spectra were recorded after 10 minutes (see experimental) and compared with the spectra of pure benzoquinone (obtained in the same deuterated solvent) in order to rule out degradation products arising from reactions with the solvent. As it happens, the regiochemistry is highly dependent on the substituent. The reaction between **2b** and L-cysteine **3**, leads to two different products corresponding to the addition on positions 5 (47%) and 6 (53%). Using **2c** a mixture of positions 5 (26%) and 6 (74%) was obtained. In the case of the **2d**, the addition occurs on positions 3 (85%), 5 (2%) and 6 (13%).

Based on the ¹H NMR spectra of pure benzoquinones (see experimental part), it seems that for **2d**, the proton H³ is the most deshielded ($\delta = 7.15$ ppm), pointing out a low electronic density at position 3 in accordance with 85% addition. For benzoquinone **2b** and

2c, the same proton H^3 is situated at 5.95 ppm and 6.65 ppm respectively leading to a higher electronic density on site 3. It may be also noted that the chemical shift values for H^5 and H^6 are located at 6.75-6.85 ppm, quite close to the chemical shift of the simple benzoquinone ($\delta = 6.80$ ppm). Therefore, the substituent has a major effect on the proton H^3 in accordance with the dependence of the rate constant on the substituent. Nevertheless, it must be pointed out that the selectivity of the addition onto the benzoquinone ring is strongly dependent on the nature of the thiol.¹⁶

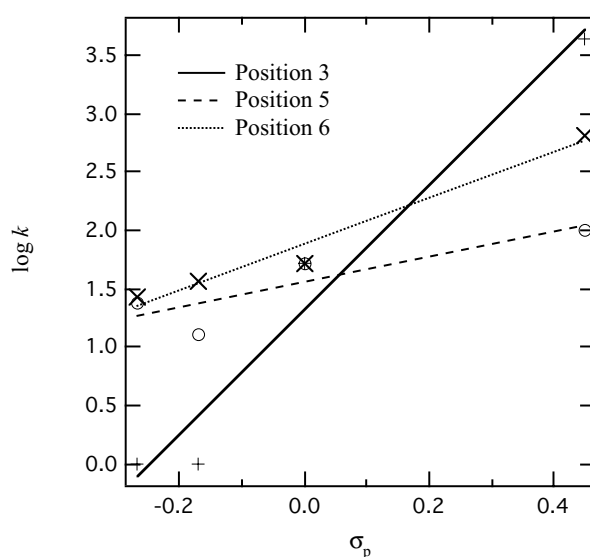


Figure 5. Correlations between the site-specific rate constants and the Hammett substituent constant, σ_p . For the non-substituted hydroquinone, the rate constants are similar for all positions (i.e. for $\sigma_p = 0$, the rate of addition on position 3, 5, 6 are the same).

The relative yields and the overall rate constant can be used to determine the reactivity of each position on the benzoquinone ring as summarized in Table 1. As expected, electron-withdrawing substituents increase the reactivity whereas electron-donating substituents tend to decrease the reactivity. The results can be presented in a graphical form as shown in Figure 5. Admittedly, it is a limited data set, but it is apparent that the rate constants for positions 5 and 6 exhibit a correlation with the Hammett substitution parameter, σ_p , which is often used to describe the effect of a substituent on the reactivity.²⁸ For position 3 only two

points are available, but the fact that no addition was found to occur at this position for **2b** and **2c** confirms the deactivating nature of the electron-donating substituents. It is also apparent that, as expected, the influence of the substituent decreases with the distance from the reactive site on the benzoquinone ring (i.e. position 5 is less influenced than position 6 which is again less influenced than position 3). The correlations may be used for a rough prediction of the reactivity and substitution patterns of benzoquinones as tags for L-cysteine units in biomolecules.

3.3. On-line electrochemical tagging of L-cysteine during mass spectrometry

In electrospray ionization, a high voltage is applied between the spray source and the MS. Recently, a new polymer microspray interface incorporating a carbon microelectrode was developed using laser photoablation.¹⁷ When a miniaturised spray interface (i.e. with a high surface to volume ratio) is employed, the high voltage may be used to electrogenerate compounds that can subsequently be detected by the MS as the interface effectively operates as a channel flow electrolysis cell.^{14,31-33} A similar procedure was used in the current investigation and as seen in Figure 6, all the substrates were able to electrotag L-cysteine as all the adducts can be observed in the mass spectra. However, when the relative abundances of L-cysteine ($[MH^+]$: $m/z = 122$ Th) and the adducts are compared, it is evident that the apparent tagging extents are not similar. In Table 2 are given the average values of the apparent tagging extent averaged for 70 minutes indicating that it increases in the order: **1c** (R = CH₃) < **1a** (R = H) < **1b** (R = OCH₃) < **1d** (R = CO₂CH₃). The compound **1e** (R = NO₂) caused the electrospray to be unstable, presumably due to an electro-polymerisation as also suggested by the electrochemical investigation.³⁰ According to the electrochemical studies, the apparent tagging extents are surprising as the rate constants corresponding to the addition of L-cysteine onto the benzoquinones follows the electronic effect of the substituent (i.e. **1d** > **1a** > **1c** ≈ **1b**). However, an important factor to consider is the ionization efficiency in microspray MS of the products.

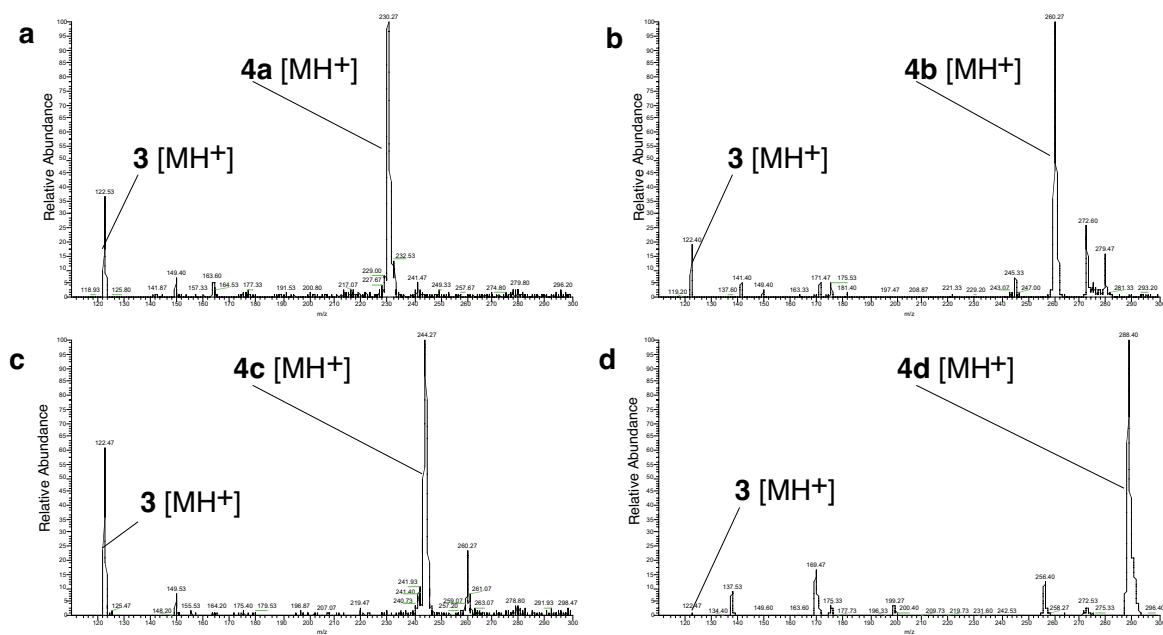


Figure 6. Microspray mass spectra recorded by infusing a mixture of 0.2 mM of L-cysteine and 20mM of the hydroquinone. (a) 1,4-Hydroquinone **1a**. (b) Methoxy-1,4-hydroquinone **1b**. (c) Methyl-1,4-hydroquinone **1c**. (d) Methoxycarbonyl-1,4-hydroquinone **1d**.

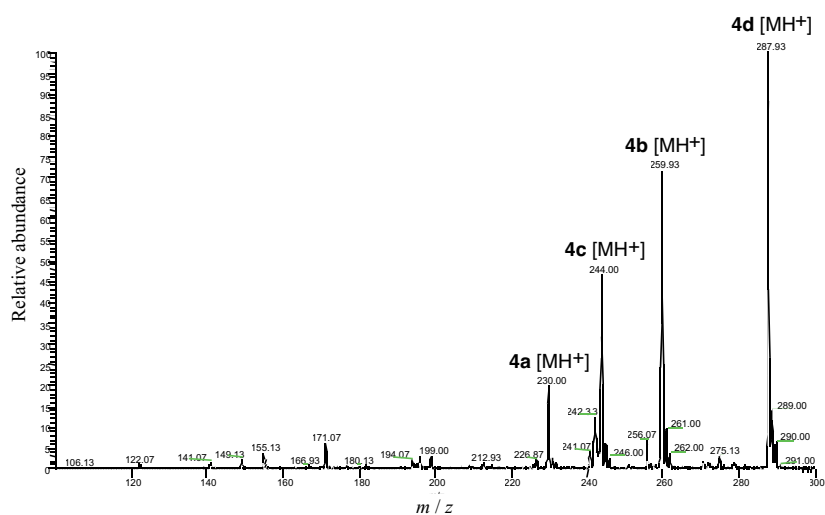


Figure 7. Microspray mass spectrum of a solution containing 0.5 mM of each of the adducts (i.e. **4a-d**).

Figure 7 shows an mass spectrum, obtained by infusing an equimolar mixture of the chemically generated adducts **4a-d**, that demonstrates that the ionization efficiency increases as: **4a** (R = H, [MH⁺]: m/z = 230 Th) < **4c** (R = CH₃, [MH⁺]: m/z = 244 Th) < **4b** (R = OCH₃, [MH⁺]: m/z = 260 Th) < **4d** (R = CO₂CH₃, [MH⁺]: m/z = 288 Th). It may thus be concluded that without accurate knowledge of the actual ionization efficiency, the real tagging extent cannot be determined on the basis of the mass spectra alone.

Due to the fact that single nucleophile amino acids sometimes show a weak MS signal, it was decided to test the selectivity of hydroquinones **1b-d** toward L-cysteine using several small peptides without cysteine residues but containing other nucleophilic amino acids. Leu-enkephalin, Met-enkephalin-Arg-Phe and bradykinin were used. In every case, no tag was observed confirming the good selectivity of hydroquinones **1b-d** toward L-cysteine under the present ESI conditions.

3.4. On-line electrochemical tagging of β -lactoglobulin A during ESI-MS

The on-line tagging method can be used for studies on biomolecules incorporating reactive thiol moieties. A proof of principle was already provided by investigations of the on-line EC-tagging of β -lactoglobulin A, a protein containing one free cysteine residue, by hydroquinone **1a**.^{14,15} The reaction was performed in a non-degassed medium using the same microspray microchip interface.¹⁴ The ratio protein/hydroquinone was taken equal to 1/1000 (5 μ M / 5 mM). The β -lactoglobulin A/hydroquinone mixture was stored in an eppendorf tube and sampled at different incubation time. The apparent EC-tagging extent, which was found to be 35% at the beginning, reached 100% after more than 3 hours. The extent increase was explained by the possible oxidation of hydroquinone **1a** into benzoquinone **2a** by the oxygen present in solution.

In a new set of experiments, EC-tagging reactions were reinvestigated using degassed solutions of hydroquinones **1a-d** and β -lactoglobulin A to compare and quantify the tagging efficiency. As previously observed for hydroquinone **1a**, the EC-tagging involving hydroquinones **1b-d** was found to be selective toward cysteine. Using myoglobin, a protein without any cysteine residue, the EC-tagging reaction was not observed. In degassed conditions using β -lactoglobulin A, the apparent EC-tagging extents are lower than the one

obtained with hydroquinone **1a** in non-degassed mode (Table 2). Methoxy-1,4-hydroquinone **1b** and methoxycarbonyl-1,4-hydroquinone **1d** reveal the best tagging efficiency.

Compound	1a	1b	1c	1d
EC-tagging extent of Cysteine ¹	76% ⁴	87% ⁴	63% ⁴	99% ⁴
EC-tagging extent of β -lactoglobulin A ²	23% ⁴	27 to 52% ^{4,5}	26% ⁴	31% ⁴
EC-tagging extent of β -lactoglobulin A ³	26% ⁴	-	-	32% ⁴

¹ Ratio hydroquinones / cysteine: 20 mM / 0.2 mM. ² Ratio hydroquinones/ β -lactoglobulin A: 5 mM / 5 μ M. ³ Ratio hydroquinones / β -lactoglobulin A: 2.5 mM / 5 μ M. ⁴ The apparent tagging extent during a microspray MS experiment was measured by the ratio of the abundance of the tagged product over the sum of the abundances of L-cysteine **3** or β -lactoglobulin A and the tagged product. The EC-tagging reactions were performed using degassed solutions for 10 minutes. ⁵ The apparent tagging extent evolves from 27 to 52% in 80 minutes of infusion.

Table 2. The overall apparent MS EC-tagging extents obtained using a polymeric microspray interface incorporating a carbon electrode.

In order to clarify the influence of oxygen on the tagging efficiency, some electrochemical investigations using square wave voltammetry (SWV)³⁴ were performed. According to the different oxidation potential of each hydroquinone (Figure 1), the evolution of hydroquinones **1a**, **b** and **d** was studied in a non-degassed ESI medium using 10 μ M of starting material (the oxidation of hydroquinones is observable in the μ M range). The signal of hydroquinones **1a**, **b**, and **d** decreases as a function of time (Figure 8). It can be attributed to their oxidation into the benzoquinone form. As a proof, the appearance of benzoquinone was also monitored and compared with the benzoquinone signals obtained from pure synthesized ones (see experimental part). Whatever the case, the curves reach a plateau, leading to a maximum benzoquinone production of 10% (*ca.* 1 μ M) for **1a**, 15% (*ca.* 1.5 μ M) for **1b** and 5% (*ca.* 0.5 μ M) for **1d**.

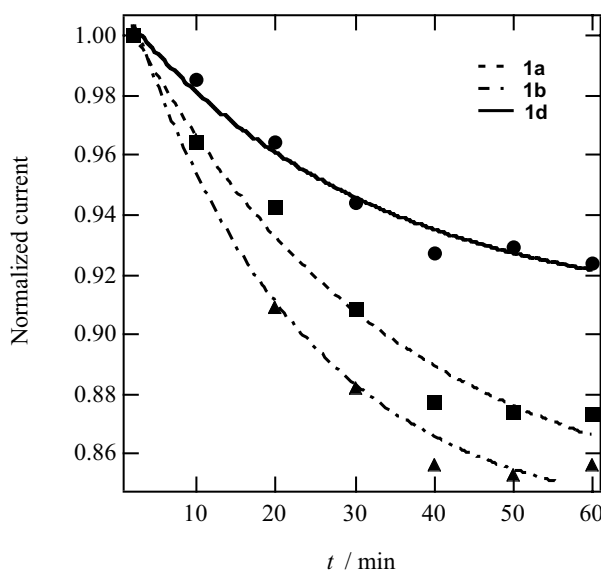


Figure 8. Time evolution of the normalized current for hydroquinones **1a**, **1b** and **1d**. The normalized current was taken as a relative intensity $I_p(t) / I_p(t = 2 \text{ min})$ and was determined by square wave voltammetry using $10 \mu\text{M}$ of each hydroquinones and as main parameters: potential step $\Delta E_s = 5 \text{ mV}$, step amplitude $\Delta E_{\text{SW}} = 25 \text{ mV}$ and frequency $f = 10 \text{ Hz}$.

Based on the SWV studies, it is clear (concentration of β -lactoglobulin A of $5 \mu\text{M}$) that the natural oxidation of hydroquinones plays a significant role on the final tagging extent. For example, using hydroquinone **1a**, the apparent tagging extent was found to be 23% in absence of oxygen instead of 35%¹⁴ in the presence of oxygen. The 12% apparent extent increase is in good agreements with what was predicted by SWV studies (10%). According to these investigations, the possibility to use **1b** was abandoned because this particular hydroquinone is the most oxidable one, leading to an evolution of the EC-tagging efficiency during time (Table 2) even if the medium is initially well degassed (*N.B.* the experiments are not performed under controlled atmosphere). We therefore focused on hydroquinone **1d**, which is on one hand, the most reactive toward L-cysteine and on the other hand the less reactive toward oxygen.

In order to improve the signal obtained from EC-tagging of β -lactoglobulin A by **1d**, several ratios protein / methoxycarbonyl-1,4-hydroquinone **1d** were tested (Table 2). The

optimisation leads to a 1/500 ratio, which produces the best apparent EC-tagging extent and reduces the noise present in the spectra. Figure 9 shows the obtained mass spectra for the electrochemical tagging of β -lactoglobulin A by **1d**. This last study confirms that methoxycarbonyl-1,4-hydroquinone **1d** is at present the most convenient probe for EC-tagging of cysteine residues in proteins.

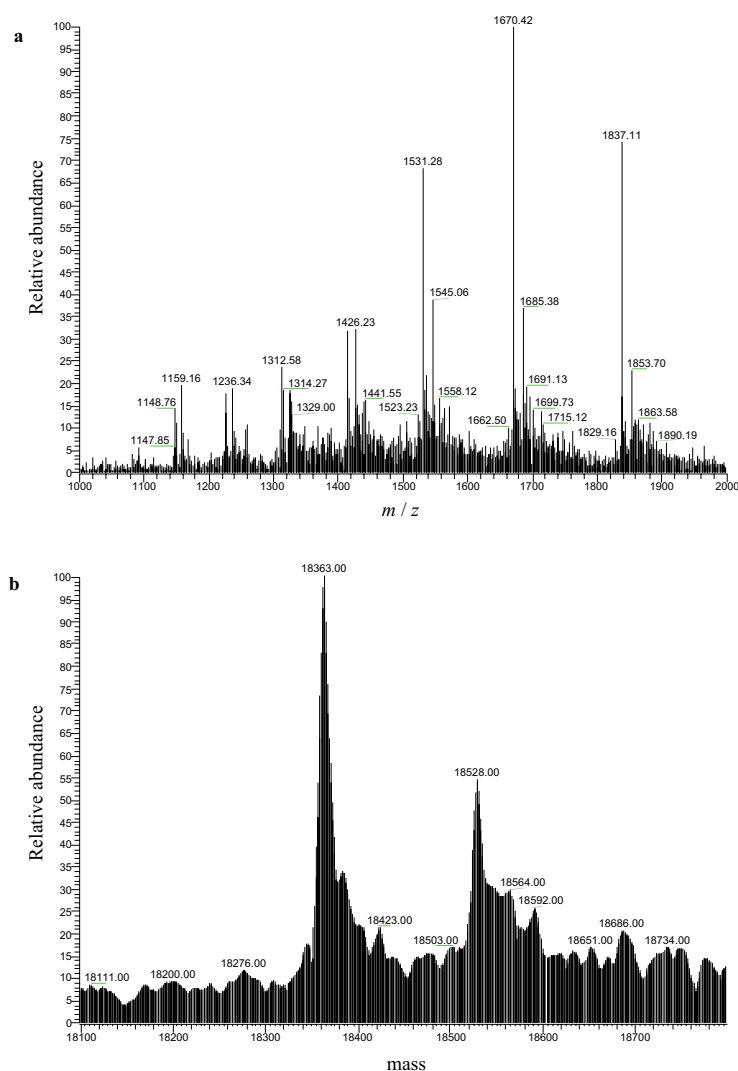


Figure 9. Microspray mass spectra obtained by infusing a mixture of 5 μ M β -lactoglobulin A in the presence of 2.5 mM carboxymethyl-1,4-hydroquinone **1d** (a). The corresponding deconvoluted spectrum is also represented (b).

4. Conclusion

A selective cysteine tagging method has been described. The procedure relies on the electrogeneration of selective electrophiles, which react with free cysteines. In order to have an accurate and well-characterised analytical tool, it is of paramount importance that the kinetic and mechanistic details of the tagging reaction are well understood. In this work, we have therefore focussed on a thorough electrochemical characterisation and determination of the relevant rate constants. In particular, the site-specific reactivity of a given position on the benzoquinone ring was found to be highly dependent on the substituents.

The apparent extents for L-cysteine tagging observed in microspray MS using a polymeric microspray interface have been measured showing that the ionization efficiencies of the benzoquinone-L-cysteine adduct are highly dependent on the substituent on the benzoquinone ring. In the case of methoxycarbonyl-1,4-hydroquinone the on-line tagging reaction is close to being quantitative, i.e. 100% tagging efficiency. The methodology was further validated using the protein β -lactoglobulin A. Additional microspray MS experiments using peptides and myoglobin containing other nucleophilic amino acid residues than L-cysteine have underlined the selectivity of the studied tags.

The proof of the concepts and methods presented in this work could provide a new approach for the analysis of complex biological mixtures. By on-line selective EC-tagging of free L-cysteine units, the data interpretation is easier. A particular point concerns the fact that both the native and shifted distributions are observed in MS. Thus, both the native and tagged sample can be studied in only one experiment, whereas other techniques require the preparation and analysis of two samples.^{35,36} Finally, the dynamic nature of the current technique can lead to a direct measure of the relative extent of the tagging reaction. This is a particularly important point as the extent is directly related to the reactivity and therefore to the 2- and 3-dimensional structure of the protein.

Bibliography

- 1 D. Figeys, *Anal. Chem.*, 2003, 75, 2891-2905.
- 2 S.D. Patterson and R.H. Aebersold, *Nature Genet.*, 2003, 33, 311-323.
- 3 R. Aebersold and D.R. Goodlett, *Chem. Rev.*, 2001, 101, 269-295.
- 4 M. Hamdan and P.G. Righetti, *Mass Spectrom. Rev.*, 2002, 21, 287-302.
- 5 P.E. Michel, F. Reymond, I.L. Arnaud, J. Josserand, H.H. Girault and J.S. Rossier, *Electrophoresis*, 2003, 24, 3-11.
- 6 A. Pandey and M. Mann, *Nature*, 2000, 405, 837-846.
- 7 C.Y. Quang, A. Malek and M.G. Khaledi, *Electrophoresis*, 2003, 24, 824-828.
- 8 D.A. Wolters, M.P. Washburn and J.R. Yates, *Anal. Chem.*, 2001, 73, 5683-5690.
- 9 R.A. Jue and J.E. Hale, *Anal. Biochem.*, 1993, 210, 39-44.
- 10 R.A. Jue and J.E. Hale, *Anal. Biochem.*, 1994, 221, 374-378.
- 11 N. Lundell and T. Schreitmuller, *Anal. Biochem.*, 1999, 266, 31-47.
- 12 S. Sechi and B.T. Chait, *Anal. Chem.*, 1998, 70, 5150-5158.
- 13 T.C. Rohner, J. Josserand, H. Jensen and H.H. Girault, *Anal. Chem.*, 2003, 75, 2065-2074.
- 14 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 15 C. Roussel, T.C. Rohner, H. Jensen and H.H. Girault, *ChemPhysChem*, 2003, 4, 200-206.
- 16 H.S. Wilgus, J.W. Gates, E. Frauenglass, E.T. Jones and R.F. Porter, *J. Org. Chem.*, 1964, 29, 594-600.
- 17 T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.
- 18 A. Klemenc, G. Ofner and H. Wirth, *z. Anorg. Allg. Chem.*, 1951, 265, 220-228.
- 19 M. Nakazaki and K. Naemura, *J. Org. Chem.*, 1981, 46, 106-111.
- 20 M. Rudolph, D.P. Reddy and S.W. Feldberg, *Anal. Chem.*, 1994, 66, A589-A600.
- 21 B.A. Brookes, P.C. White, N.S. Lawrence and R.G. Compton, *J. Phys. Chem. B*, 2001, 105, 6361-6366.
- 22 A. Digga, S. Gracheva, C. Livingstone and J. Davis, *Electrochem. Commun.*, 2003, 5, 732-736.

- 23 G. Hignett, S. Threlfell, A.J. Wain, N.S. Lawrence, S.J. Wilkins, J. Davis, R.G. Compton and M.F. Cardosi, *Analyst*, 2001, 126, 353-357.
- 24 N.S. Lawrence, J. Davis and R.G. Compton, *Talanta*, 2000, 52, 771-784.
- 25 N.S. Lawrence, J. Davis and R.G. Compton, *Talanta*, 2001, 53, 1089-1094.
- 26 N.S. Lawrence, J. Davis, L. Jiang, T.G.J. Jones, S.N. Davis and R.G. Compton, *Analyst*, 2000, 125, 661-663.
- 27 N.S. Lawrence, J. Davis, F. Marken, L. Jiang, T.G.J. Jones, S.N. Davies and R.G. Compton, *Sens. Actuators B*, 2000, 69, 189-192.
- 28 M.B. Smith and J. March, *March's Advanced Organic Chemistry*, 5th ed., John Wiley and Sons, New York, 2001.
- 29 C. Hansch, A. Leo and R.W. Taft, *Chem. Rev.*, 1991, 91, 165-195.
- 30 K. Yamamoto, T. Asada, H. Nishide and E. Tsuchida, *Bull. Chem. Soc. Jpn.*, 1990, 63, 1211-1216.
- 31 A.T. Blades, M.G. Ikononou and P. Kebarle, *Anal. Chem.*, 1991, 63, 2109-2114.
- 32 G.J. Van Berkel and F.M. Zhou, *Anal. Chem.*, 1995, 67, 2916-2923.
- 33 F.M. Zhou and G.J. Van Berkel, *Anal. Chem.*, 1995, 67, 3643-3649.
- 34 J.G. Osteryoung and R.A. Osteryoung, *Anal. Chem.*, 1985, 57, A101-A110.
- 35 G.C. Adam, E.J. Sorensen and B.F. Cravatt, *Mol. Cell. Proteomics*, 2002, 1, 781-790.
- 36 G. Cagney and A. Emili, *Nature Biotechnol.*, 2002, 20, 163-170.

CHAPTER IV. *On-line counting of cysteine residues in peptides during electrospray ionization by electrogenerated tags and their application to protein identification*[§]

1. Introduction

Protein identification is one of the main tasks in proteomics, which refers to the study of the protein content of a biological sample. With this aim in view, some identification procedures based on mass spectrometry (MS)^{1,2} have been developed, such as the study of the peptides coming from the proteolytic digestion of proteins^{3,4}. In a typical gel-free bottom-up approach², the proteolysis can lead from 30 to more than 100 peptides per digested proteins.⁵ A protocol based on reverse phase high performance liquid chromatography mass spectrometry (RP-HPLC-MS) is commonly used to separate and analyse the generated peptides.^{6,7} The parent protein is found thanks to the mass fingerprint of the tryptic peptides and comparison with DNA and protein databases.⁸

Besides, it is advantageous to use supplementary information (in addition to the peptides masses) to constrain the database search. Sechi and Chait⁹ have demonstrated that protein identification can be greatly improved when knowing the content of one specific amino acid residue within tryptic peptides since it rules out many potential matches. It can be

[§] based on L. Dayon, C. Roussel, M. Prudent, N. Lion and H. H. Girault, *Electrophoresis*, 2005, 26, 238-247.

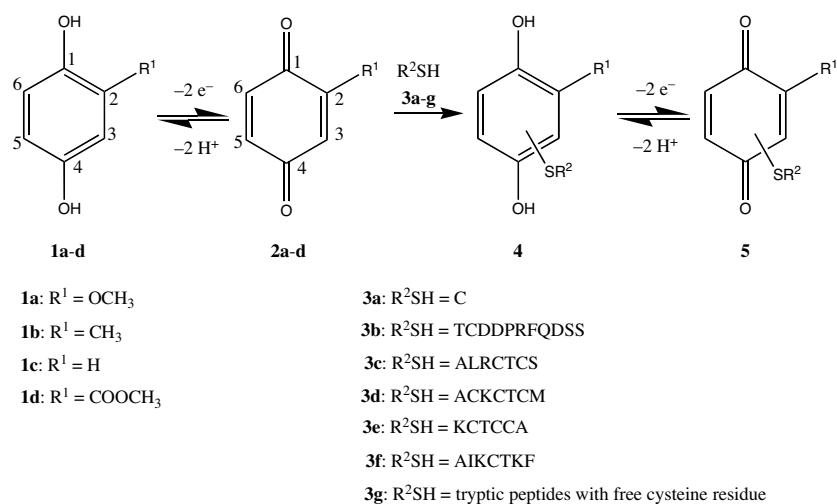
used for a large range of proteins when quantifying a well-chosen amino acid in proteomes.¹⁰ Cysteine, contained in 89.3%, or histidine, contained in 83% of all proteins in humans, are good candidates.⁵ The knowledge that a peptide contains at least one cysteine residue, together with a mass accuracy analysis of 0.5 Da, reduces by a factor of 2.5-10, the number of possible protein candidates from which such a peptide is derived.¹¹

The chemical tagging of free cysteines by iodoacetamide or by other alkylating reagents¹²⁻¹⁴ was first introduced in protein 2D-gel electrophoresis procedures.¹⁵ Indeed, cysteine alkylation avoids in particular the formation of acrylamide adducts during electrophoresis. The tagging of free cysteines can be also used for the counting of cysteines in peptides. MS methods to quantify cysteines are based on the mass shift between the modified and the untagged biomolecule or can rely upon the isotopic distribution pattern specific to the alkylated peptides.^{9,16-18} They often require a time consuming sample preparation¹⁹, and when used with a quantitative label, a division of the sample, one half unlabelled, the other labelled.

The recently developed on-line electrochemical tagging (EC-tagging) of free cysteine residues using a microspray chip²⁰ as electrospray ionisation (ESI) source is a practical alternative to the previous sequential methods. Such a process yields instantaneously both tagged and untagged biomolecules directly from a single sample. A microspray device with an integrated carbon electrode is used as an electrolysis cell^{21,22} to electrogenerate species able to react specifically with cysteine moieties in acidic spray condition.²³ The tagging of free cysteine by hydroquinones coupled to MS detection involves an electrochemical-chemical-electrochemical (ECE) mechanism as it was previously described (Scheme 1).²³⁻²⁵ Under flow conditions the second electrochemical oxidation has no time to occur.

The method was originally applied to the target protein β -lactoglobulin A with only one free cysteine residue using hydroquinone.²³ The obtained spectrum presented two peak distributions: the unmodified one, corresponding to the starting biomolecule and a shifted one, corresponding to the adduct resulting from the labelling of the free cysteine residue contained in the protein. Several studies were conducted to optimize the tagging reaction and several substituted hydroquinones, presenting different reactivities toward L-cysteine, were employed. The kinetics of the addition of L-cysteine was determined by

electrochemical studies and the final on-line tagging efficiency was investigated by MS measurements (chapter III).²⁶



Scheme 1. Reaction mechanism of the EC-tagging of cysteines.

In this chapter, the study of substituted hydroquinones as tags for cysteine-containing peptides was carried out. The reactivity/tagging efficiency was investigated on a target peptide, the human chorionic gonadotropin- β (109-119) amide (β -hCG) containing one cysteine residue, using several hydroquinones: methoxy-1,4-hydroquinone, methyl-1,4-hydroquinone, 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone. Detailed investigations were then pursued on the EC-tagging using 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone for the counting of cysteine units in peptides. Peptides containing up to three cysteines were tested and the tagging of a peptide containing two consecutive cysteine residues was studied. The EC-tagging with methoxycarbonyl-1,4-hydroquinone was thus applied to bovine serum albumin (BSA) and human α -lactalbumin digestions in a gel-free peptide mapping strategy in order to highlight the improvement of the protein identification by the knowledge of the cysteine content of some tryptic peptides in comparison with the classical procedure. The post-column EC-tagging was then tested for the identification of four proteins in a model mixture.

2. Experimental

2.1. Chemicals

1,4-Hydroquinone (> 98%), ammonium bicarbonate (> 98%), 1,4-dithio-DL-threitol (DTT) (≥ 99.5) and myoglobin from horse heart (> 90%) were purchased from Fluka (Büchs, Switzerland), methyl-1,4-hydroquinone (99%), methoxy-1,4-hydroquinone (99%) from Acros (Fair Law, NJ, USA), methoxycarbonyl-1,4-hydroquinone (methyl-2,5-dihydroxybenzoate, 99%) from Aldrich (Milwaukee, WI, USA) and iodoacetamide, bradykinin (99%), Met-enkephalin-Arg-Phe (98%), bovine serum albumin, human α -lactalbumin (> 90%), albumin from chicken eggs (ovalbumin, > 98%) and β -lactoglobulin A from bovine milk from Sigma (St. Louis, MO, USA). Chorionic gonadotropin- β (109-119) amide (> 94%), Leu-enkephalin (> 99%) and KCTCCA (70%) were bought from Bachem (Bubendorf, Switzerland). Porcine trypsin was from Promega (Madison, WI, USA). Methanol (> 99.8%, Riedel-de Haën, Seelze, Germany), acetic acid (> 99.5%, Fluka), trifluoroacetic acid (TFA) (99%, Merck, Darmstadt, Germany) and acetonitrile (> 99.5%, Fluka) were used without any further purification. Deionised water (18.5 M Ω -cm) was prepared using a Milli-Q system from Millipore (Bedford, MA, USA). Solvents used for high performance liquid chromatography were water HPLC grade from Sds (Peypin, France) and acetonitrile ultra gradient HPLC grade from J.T. Baker (Deventer, Netherlands). Synthetic peptides AIKCTKF, ALRCTCS and ACKCTCM (>70%) were prepared by Catherine Servis at the Institut de Biochimie (Faculté de Médecine, Epalinges, Switzerland).

2.2. Microchip device

The fabrication of the microspray interface has been previously described (see also appendix I for a comprehensive description).²⁰ Briefly, a polyethylene terephthalate (PET) substrate (100 μm thick Melinex[®] sheet from Dupont (Wilmington, DE, USA)) was photoablated with a UV excimer laser (ArF 193 nm from Lambda Physik (Göttingen, Germany)). The carbon electrode was 70 \times 25 μm^2 integrated in the channel of 35 \times 30 μm^2 section. The distance from the electrode to the outlet of the microchip was 1.8-2 cm.

2.3. MS set-up

A LCQ DUO ion trap mass spectrometer (Finnigan, San José, CA, USA) was used. The heated capillary was kept at 200 °C. In each experiment, the ion transmission parameters are optimised automatically in order to improve the detection of the analyte of interest (the unmodified peptide). The ESI interface was removed and the microchip holder was mounted on the probe slide adapter of the mass spectrometer. The device was coupled to a syringe pump (kdScientific, Holliston, MA, USA) to introduce the solution. The flow rate was set to 250 nL·min⁻¹ and the voltage applied was 3.5-4 kV. The distance from the outlet of the microchip to the entrance of the spectrometer varied between 1 and 2 cm in order to optimise the signal and the trap injection time.

In a typical experiment, a given peptide was mixed with a given hydroquinone in order to obtain respectively concentration of 50 µM and 20 mM. The spray medium used was MeOH / H₂O / CH₃COOH 50% / 49% / 1%. The solutions were not degassed.

2.4. Digestion and peptide analysis

2 mg of protein were dissolved in 2 mL of ammonium bicarbonate solution (5 mM, pH = 8). 1.23 mg of DTT (4 mM) was added, and then 20 µg of trypsin (protein ratio of 1:100 (w/w)). For the mixture of four proteins, 1 mg of each protein was dissolved in 4 mL of buffer solution (40 µg of trypsin was used). The digestion was run at 37 °C for 4 hours. The solution was lyophilised and the obtained solid was then diluted in 400 µL of H₂O and 100 µL were injected for separation. Before separation, the BSA digest sample was divided in two equal parts, and one of them was subjected to iodoacetamide alkylation (+1.41 mg of iodoacetamide (38 mM), reaction performed at 45 °C for 30 minutes in the dark). The HPLC separation (Alliance 2690 Waters (Mildford, MA, USA) system using Millennium software) was performed using a C₁₈ reverse phase column (Nucleosil 100-5 C18 from Macherey-Nagel, Düren, Germany). The separation was run for 90 minutes using a gradient of solvent H₂O / TFA 99.88% / 0.12% (solvent A) and acetonitrile (solvent B). The gradient was run as follows: from 0 to 5 minutes 100% of A, then to 62% of A and 38% of B at 65 minutes, 43% of A and 57% of B at 75 minutes, 20% of A and 80% of B at 90 minutes at a flow rate of 1

mL·min⁻¹. Fractions were collected every minute. To 90 µL of a given fraction, 10 µL of 2-methoxycarbonylhydroquinone 200 mM in CH₃CN / H₂O / TFA 50% / 49% / 1% were added to be analysed by microspray MS. To 90 µL of a fraction of the alkylated sample was added 10 µL of CH₃CN / H₂O / TFA 50% / 49% / 1%.

2.5. Protein identification

Single protein samples were identified using Mascot search software.^{27,28} The model mixture was recovered using ProFound.^{29,30} Mascot search was performed with the following parameters: charge states of +1, +2 and +3, 1 maximum miscleavage, possible oxidation of methionine as variable modification, no fixed modification (except for the iodoacetamide alkylated sample for which carbamidomethyl was chosen), peptide mass tolerance of ± 0.4 Da. The main chosen parameters for the mixture are 1 miscleavage, tolerance of 0.4 Da, taxonomy Chordata and the search was focused with one charge state peptides. The search was performed in SwissProt database.³¹ In all tagging experiments, the cysteine content information was only entered for peptides detected as containing some cysteines. The information on the supposed absence of cysteine was not specified.

3. Results and discussions

3.1. Tagging of chorionic gonadotropin-β (109-119) amide

The cysteine selective EC-tagging method²⁶ was successfully assessed with a target peptide containing one cysteine residue, human chorionic gonadotropin-β (109-119) amide (**3b**) ($M = 1269.30 \text{ g}\cdot\text{mol}^{-1}$) using methoxy-1,4-hydroquinone (**1a**), methyl-1,4-hydroquinone (**1b**), 1,4-hydroquinone (**1c**) and methoxycarbonyl-1,4-hydroquinone (**1d**) (Scheme 1).

The EC-tagging reaction was monitored (Figure 1) using the microspray interface described above and a classical MS medium (MeOH / H₂O / AcOH 50% / 49% / 1%). The tagging occurs with all hydroquinones **1a-d**, and the presence of both the untagged peptide as

well as the tagged product was observed (Figure 1). The presence of sodium and potassium adducts on the peptides accounts for extra peaks in the spectra.

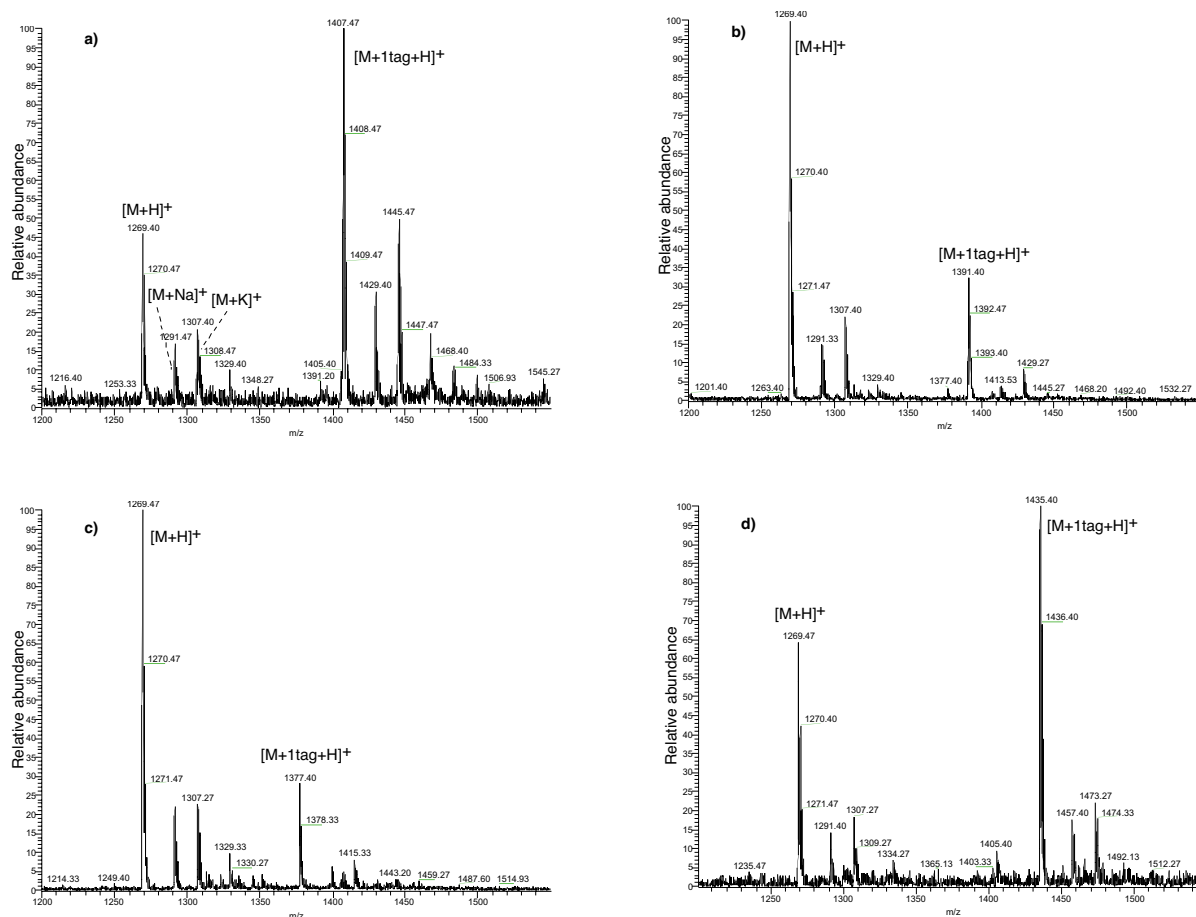


Figure 1. Mass spectra of chorionic gonadotropin β (109-119) amide infused at 250 nL·min⁻¹ in the presence of **1a** (a), **1b** (b), **1c** (c) and **1d** (d).

The tagging efficiency for each probe was compared, by calculating the average apparent tagging extent on 40 minutes of infusion (0.55, 0.24, 0.22 and 0.56 for **1a-d** respectively; the tagging extent was calculated as the ratio of the abundance of the tagged peptide over the sum of the untagged and tagged peptide on mass spectra). The tagging extents follow the electronic effect of each substituent excepted for **1a**, which should normally exhibit the lowest extent.

The particular reactivity of methyl-1,4-hydroquinone (**1a**) had been previously observed with L-cysteine (**3a**) and attributed to a better ionization efficiency of the formed adduct **3a+2a** compared with adducts coming from the reaction with **2b** and **2c**.²⁶ However, no ionization differences could be seen between the peptide and the peptide+benzoquinone adducts. As previously observed, the higher reactivity of **1a** could be explained by its natural oxidation by oxygen in the spray solution, leading to a chemical tagging in the infusion line prior to the EC-tagging.²⁶ As a proof, the tagging extent involving **1a** was observed to evolve through time, from 0.38 to 0.64 after 40 minutes of infusion (MS experiments were not run under inert atmosphere).

3.2. Tagging of peptides containing several cysteines

Since lots of proteins contain many cysteine residues, the probability to find a peptide coming from a protein digestion having more than one cysteine is relevant. Thus, in order to check if the method could be applied to the quantification of cysteine units in tryptic peptides, the on-line EC-tagging of peptides containing more than one cysteine residue as well as a peptide presenting two consecutive cysteines was investigated. In the following, we focussed only on the reaction of **1c** and **1d** with several peptides. This choice was justified by the fact that, on one hand **1d** always provided the best tagging efficiency because of the kinetics of the addition reaction and, on the other hand, **1d** is the less naturally oxidable hydroquinone compound.²⁶ Hydroquinone (**1c**) is also taken as reference since it is the simplest, and the most studied species for on-line EC-tagging.^{23,25,26} Moreover, it is not quickly oxidised in the presence of oxygen.²³ First, the EC-tagging of two synthetic peptides was studied: one contains two cysteine residues (ALRCTCS, $M = 752.90 \text{ g}\cdot\text{mol}^{-1}$) and the other presents three cysteine moieties (ACKCTCM, $M = 758.99 \text{ g}\cdot\text{mol}^{-1}$).

Hydroquinone compounds **1c** and **1d** were investigated for the EC-tagging of ALRCTCS (**3c**). As can be observed in the spectra of Figure 2, **1d** provides again the best tagging efficiency, evidenced by the fact that the second tagging peak presents a good intensity compared to the one obtained with **1c**.

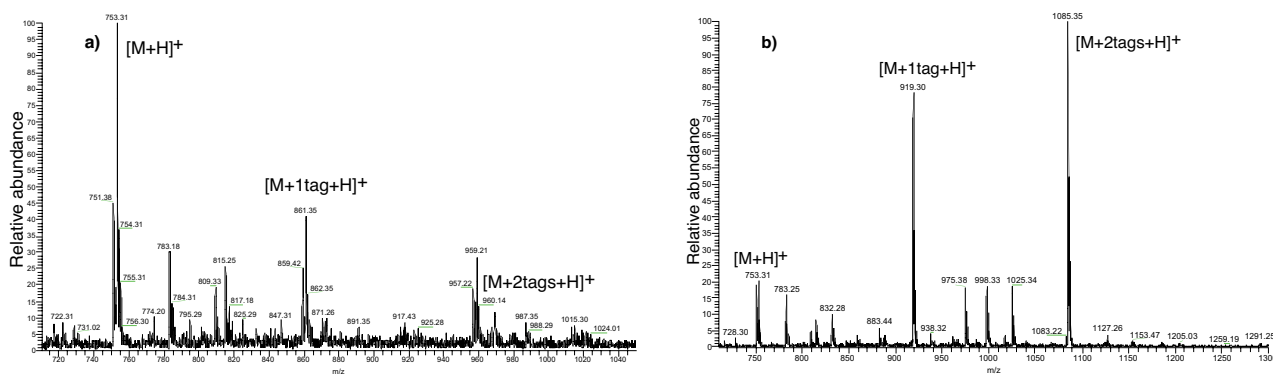


Figure 2. Mass spectra of peptide ALRCTCS infused in the presence of **1c** (a) and **1d** (b).

Then, ACKCTCM (**3d**) was infused in the presence of hydroquinone compounds **1c** and **1d**. The tagging results were followed by MS. Using **1c**, it was not possible to obtain the tagging of all cysteine residues (Figure 3a). In contrast, the use of **1d** showed that it is possible to tag the three residues at the same time, providing the quantification of free cysteines present in this peptide. Indeed, the signals coming from the singly, doubly and triply tagged peptide are well observable on the mass spectrum; the presence of the untagged peptide signal is also visible (Figure 3b).

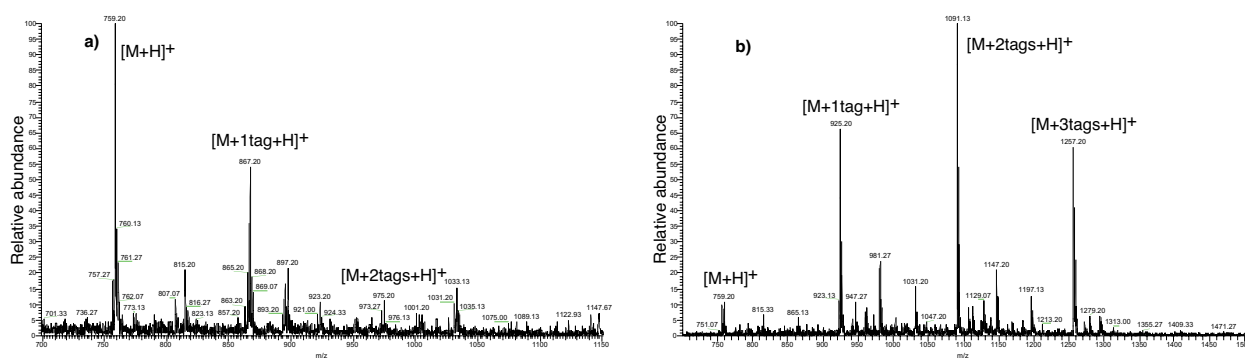


Figure 3. Mass spectra of peptide ACKCTCM infused in the presence of **1c** (a) and **1d** (b).

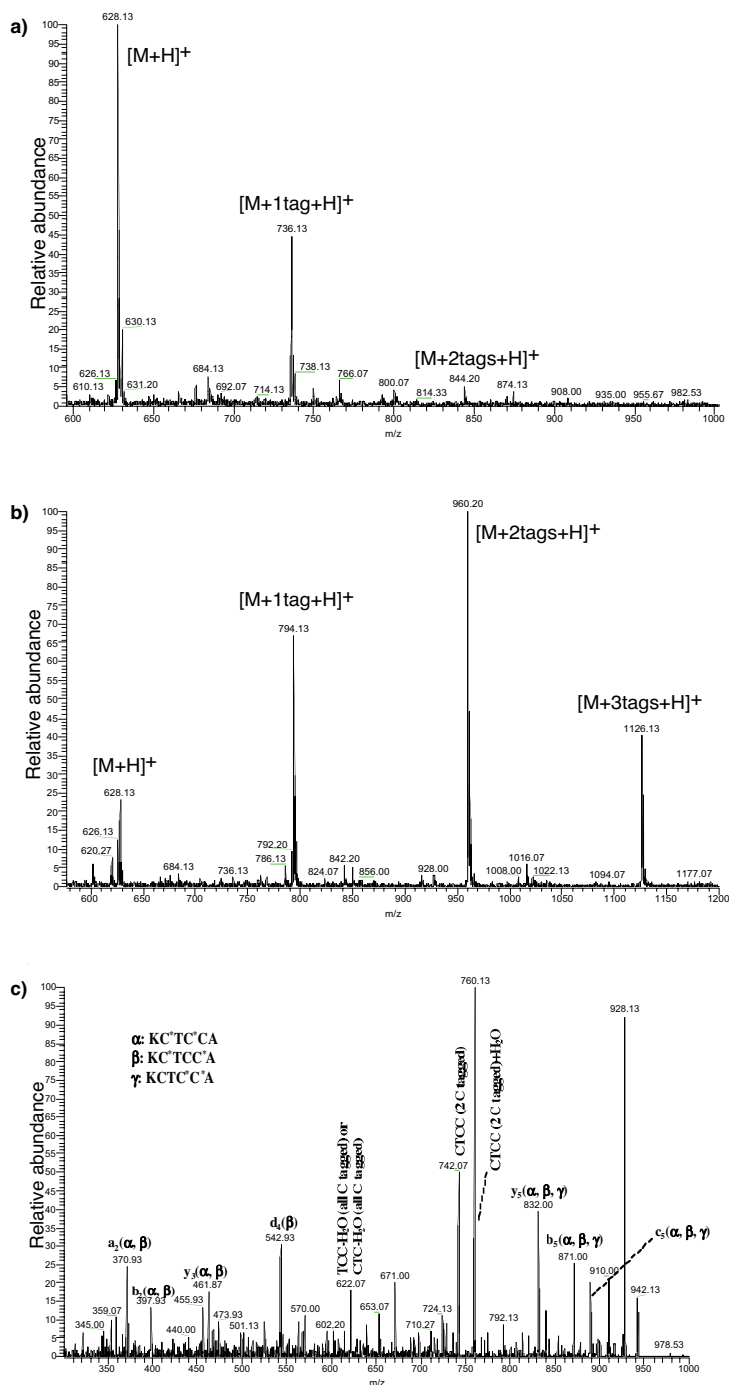


Figure 4. Mass spectra of peptide KCTCCA infused in the presence of **1c** (a) and **1d** (b). Tandem mass spectrum at $m/z = 960.2$ Th ($[M+2tags+H]^+$) corresponding to the two times tagged peptide (c). 40% collision energy was employed. α , β and γ peptides represent the different two tagged peptide possibilities (* means that the cysteine is labelled). The peak at $m/z = 928.1$ Th corresponds to the loss methanol from the tag or the loss of sulphur.³²

The EC-tagging scheme was assessed in the challenging case of the synthetic peptide KCTCCA (**3e**) ($M = 627.81 \text{ g}\cdot\text{mol}^{-1}$), containing three cysteines, two of which being consecutive. Despite possible steric hindrance, Figure 4a-b shows clearly that the tagging happens in an efficient manner with **1d** whereas with **1c** only the singly tag peak is well observed. Yet, on a set of experiments, one could notice that the triply tagged peptide peak was sometimes poorly intense (*i.e.* the triply tagged compound is not greatly produced). As a matter of facts, the tagging of two consecutive cysteines may be difficult.

According to Figure 4c, the tandem MS (MS/MS) of the peak at $m/z = 960$ Th (corresponding to the doubly tagged peptide $[\text{M}+2\text{tags}+\text{H}]^+$) let think that the two consecutive cysteine residues are not tagged at the same time (*i.e.* it is easier to tag distant cysteine residues). For the doubly tagged peptide, three possible adducts can be generated. Indeed, α peptide (KC*TC*CA), β peptide (KC*TCC*A) and γ peptide (KCTC*C*A) can be formed, in which the asterisk represents the cysteine residue labelled by methoxycarbonyl-1,4-hydroquinone (**1d**). The predicted peptide fragments obtained in the three cases were calculated and gathered together in Table 1.³³ The mass fragments obtained by MS/MS do not match with those relative to the peptide tagged on the two consecutive cysteine moieties (Figure 4c, Table 1). They all correspond to fragments of α and β peptides, which are tagged on distant residues. Thus, MS/MS experiments of on-line tagged products provide significant information for the localisation of the cysteines.³⁴ These observations can be explained by steric hindrance reasons. When the first cysteine residue is tagged, the surrounding others sites become less accessible to the probe and especially the consecutive ones. Nevertheless, as the triply tagged product is all the same well observed (Figure 4b), the good reactivity of methoxycarbonyl-1,4-hydroquinone (**1d**) is able to counterbalance the previously mentioned steric hindrance factor.

This set of results confirms that the method is quite convenient due to its non-quantitative efficiency in such conditions, since it does not require the preparation of two samples to access the number of cysteines. Moreover, the EC-tagging is quasi-instantaneous compared to the chemical labelling which is often a time consuming step.¹¹ The tool reveals itself as an efficient way to quantify cysteine in peptides.

Fragment type	α Peptide ^a (KC*TCC*A) fragment mass (Da)	β Peptide ^a (KC*TCC*A) fragment mass (Da)	γ Peptide ^a (KCTC*C*A) fragment mass (Da)
a_n	a ₅ : 843.7 a ₄ : 740.7 a ₃ : 471.4 a ₂ : 370.3	a ₅ : 843.7 a ₄ : 574.6 a ₃ : 471.4 a ₂ : 370.3	a ₅ : 843.7 a ₄ : 574.6 a ₃ : 305.4 a ₂ : 204.3
b_n	b ₅ : 871.7 b ₄ : 768.6 b ₃ : 499.4 b ₂ : 398.3	b ₅ : 871.7 b ₄ : 602.6 b ₃ : 499.4 b ₂ : 398.3	b ₅ : 871.7 b ₄ : 602.6 b ₃ : 333.4 b ₂ : 232.3
c_n	c ₅ : 888.8 c ₄ : 785.6 c ₃ : 516.5 c ₂ : 415.4	c ₅ : 888.8 c ₄ : 619.6 c ₃ : 516.5 c ₂ : 415.4	c ₅ : 888.8 c ₄ : 619.6 c ₃ : 350.5 c ₂ : 249.4
x_n	x ₅ : 858.6 x ₄ : 589.5 x ₃ : 488.4 x ₂ : 219.2	x ₅ : 858.6 x ₄ : 589.5 x ₃ : 488.4 x ₂ : 385.2	x ₅ : 858.6 x ₄ : 755.5 x ₃ : 654.4 x ₂ : 385.2
y_n	y ₅ : 832.6 y ₄ : 563.5 y ₃ : 462.4 y ₂ : 193.3	y ₅ : 832.6 y ₄ : 563.5 y ₃ : 462.4 y ₂ : 359.3	y ₅ : 832.6 y ₄ : 729.5 y ₃ : 628.4 y ₂ : 359.3
z_n	z ₅ : 816.6 z ₄ : 547.1 z ₃ : 446.4 z ₂ : 177.2	z ₅ : 816.6 z ₄ : 547.1 z ₃ : 446.4 z ₂ : 343.2	z ₅ : 816.6 z ₄ : 713.1 z ₃ : 612.4 z ₂ : 343.2

^a the peptide can be tagged twice in three different manners (α , β and γ peptide). * represents a tagging site.

Table 1. Predicted fragment masses ($[M+H]^+$) from MS/MS experiments for KCTCCA peptide tagged on two cysteine residues.[§]

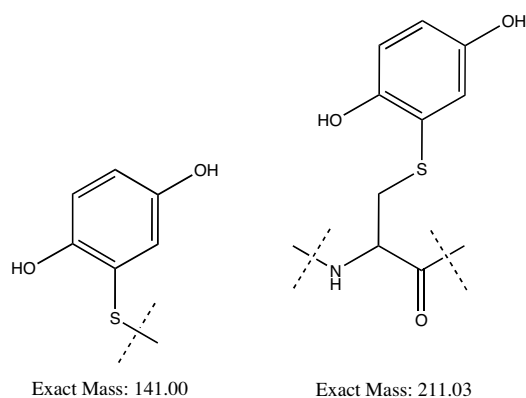
[§] nomenclature for sequence ions is given in the manuscript appendix III.

To confirm that the on-line tagging can be used in other classical spray media and coupled to HPLC separation, synthetic peptide ACKCTCM was analysed in CH₃CN / H₂O / TFA 50% / 49.9% / 0.1% with **1d**. The EC-tagging occurs as in the classical medium used for on-line EC-tagging. Since the method can be quite useful for the study of protein digests, the tagging of ACKCTCM was also performed in CH₃CN / H₂O / TFA 20% / 79.9% / 0.1% and CH₃CN / H₂O / TFA 80% / 19.9% / 0.1% media. The tagging remains efficient: the three cysteines are labelled at the same time (data not shown).

The selectivity for cysteine residues was tested in these different conditions on several peptides (bradykinin, Leu-enkephalin and Met-enkephalin-Arg-Phe) and myoglobin from horse heart, which are cysteine free. As previously obtained in MeOH / H₂O / CH₃COOH 50% / 49% / 1%²⁶, the tagging reaction was found to be selective in the experimental conditions.

3.3. Tandem mass spectrometry of peptide-quinone adducts

The fragmentation of the singly adducted peptides was then studied to identify specific characteristics in the mass spectra that could be used to corroborate the peptide targets of the tags in complex mixture. D.E. Mason and D.C. Liebler have previously characterized 1,4-benzoquinone-adducted model peptides by MS/MS with collision-induced dissociation (CID).³⁴ They showed that sequence information obtained from modified peptides is comparable to that of their unmodified analogues and allows unambiguous localisation of the modified cysteine residue. In addition, characteristic fragmentation of benzoquinone-adducted tryptic peptides establishes the presence of thiol-benzoquinone adducts. A unique ion pair separated by 141 or 142 Da corresponds to β -elimination of benzoquinol-S and benzoquinol-SH. An alternate ion pair of 211 Da shows a fragmentation at the peptide bond on either side of the adducted cysteine (Scheme 2).



Scheme 2. Characteristic mass losses of 1,4-benzoquinone-adducted cysteine residue.

In our study, CID was performed during the EC-tagging of chorionic gonadotropin- β (109-119) amide (TCDDPRFQDSS amide) and AIKCTKF (**3f**) ($M = 810.02 \text{ g}\cdot\text{mol}^{-1}$) on the peptides and the corresponding adducts with **1c** and **1d**. For 1,4-benzoquinone-adducted and methoxycarbonyl-1,4-benzoquinone-adducted peptides, the m/z values of the modified b_n and y_n ions are shifted by 108 and 166 Th respectively (Figure 5 and 6).

The localization of cysteine residue is facilitated and allows the distinction of fragments. For instance, for chorionic gonadotropin- β (109-119) amide, the peak at $m/z = 835.2 \text{ Th}$, which could be attributed to b_7 or y_7 , is unambiguously attributed to y_7 since the fragment appears without mass shifts in tandem mass spectra of 1,4-benzoquinone-adducted and methoxycarbonyl-1,4-benzoquinone-adducted peptides (Figure 5).

The characteristic mass loss of 142 Da for 1,4-benzoquinone-SH becomes 200 Da for methoxycarbonyl-1,4-benzoquinone-SH. For AIKCTKF, an ion pair separated by 200 Da is noted from $[M-H_2O]^+$ ($m/z = 958.1 \text{ Th}$) to $m/z = 758.2 \text{ Th}$ (Figure 6c). In addition, a mass loss of 32 Da (from $m/z = 976.1$ to 944.1 Th) corresponds to the loss of methanol from the tag. The adduct-specific markers thus should confirm the identification of the adducted peptides detected by the EC-tagging.

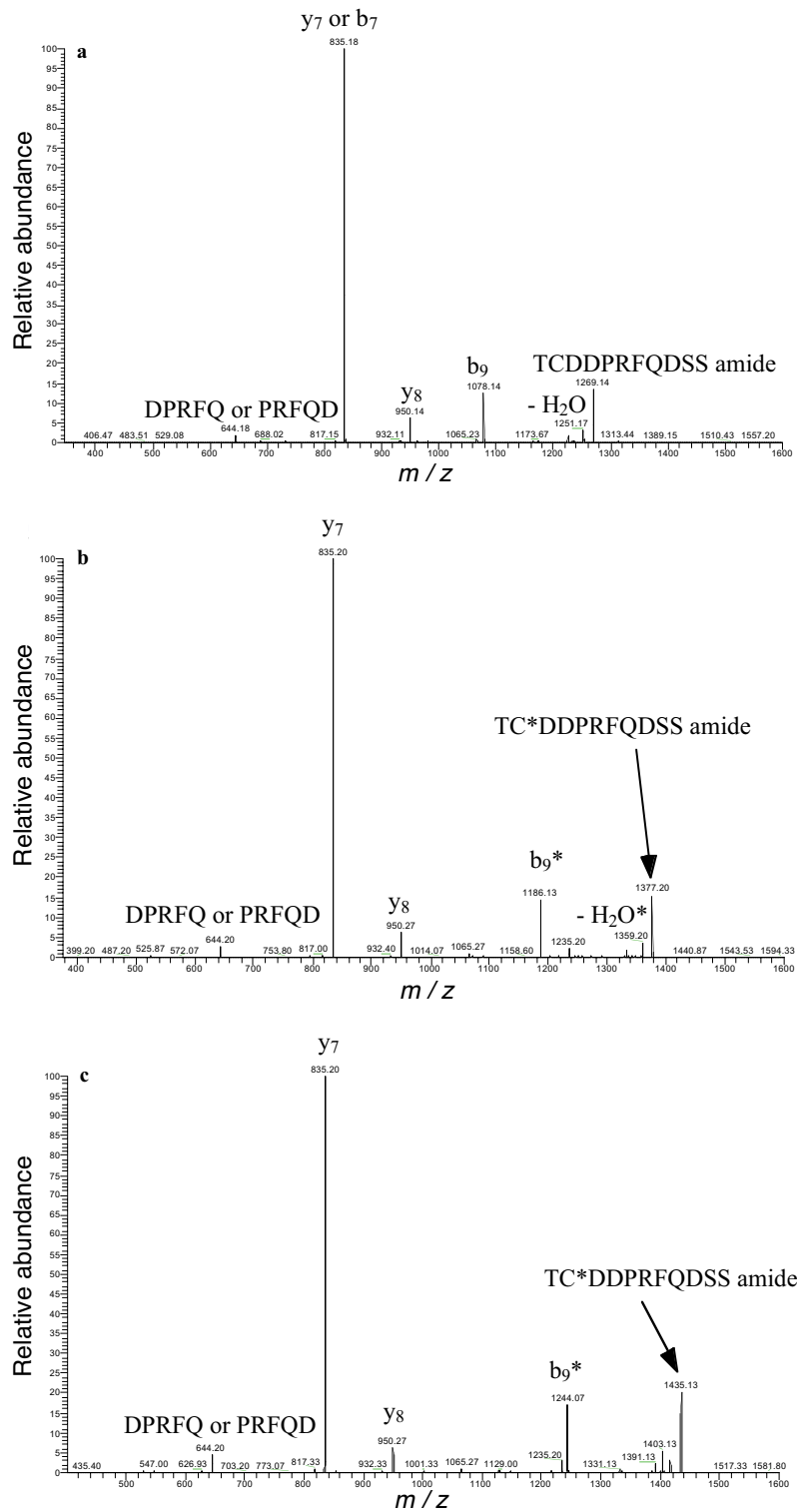


Figure 5. Tandem mass spectra of chorionic gonadotropin- β (109-119) amide (a), 1,4-benzoquinone-adducted (b) and methoxycarbonyl-1,4-benzoquinone-adducted peptides (c). Asterisks correspond to fragments containing the tag.

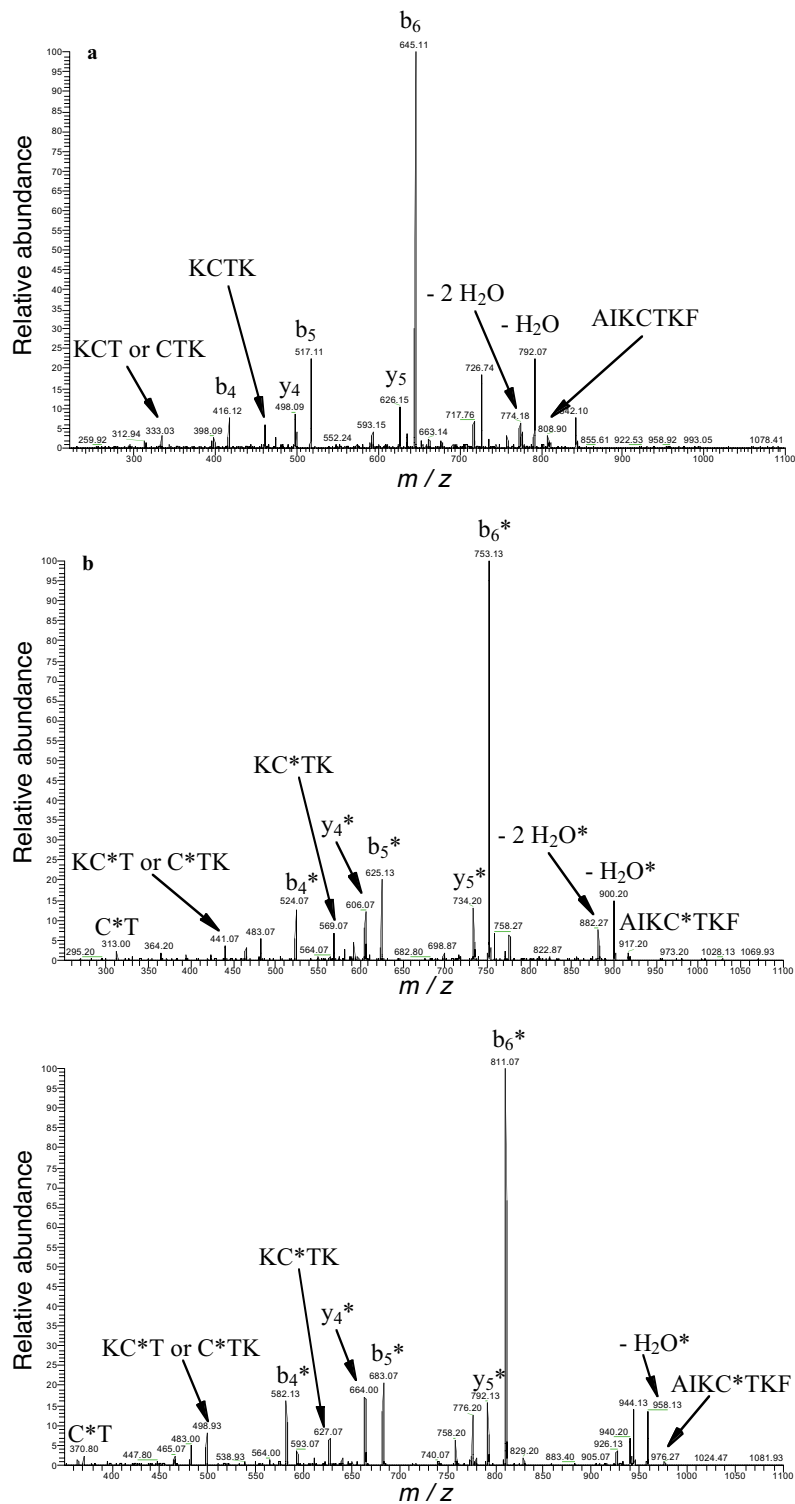


Figure 6. Tandem mass spectra of peptide AIKCTKF (a), 1,4-benzoquinone-adducted (b) and methoxycarbonyl-1,4-benzoquinone-adducted peptides (c). Asterisks correspond to fragments containing the tag.

3.4. Mass fingerprinting of bovine serum albumin and human α -lactalbumin using on-line electrochemical tagging

As a proof of principle, the cysteine quantification tagging method was applied to a typical mass fingerprinting experiment of BSA, which contains 35 cysteines.^{3,4} BSA was reduced, digested and the tryptic peptides (**3g**) separated by RP-HPLC. The fractions corresponding to retention times between 10 and 75 minutes on the chromatogram (Figure 7a) were analysed by chip-ESI-MS with the EC-tagging using methoxycarbonyl-1,4-hydroquinone (**1d**).

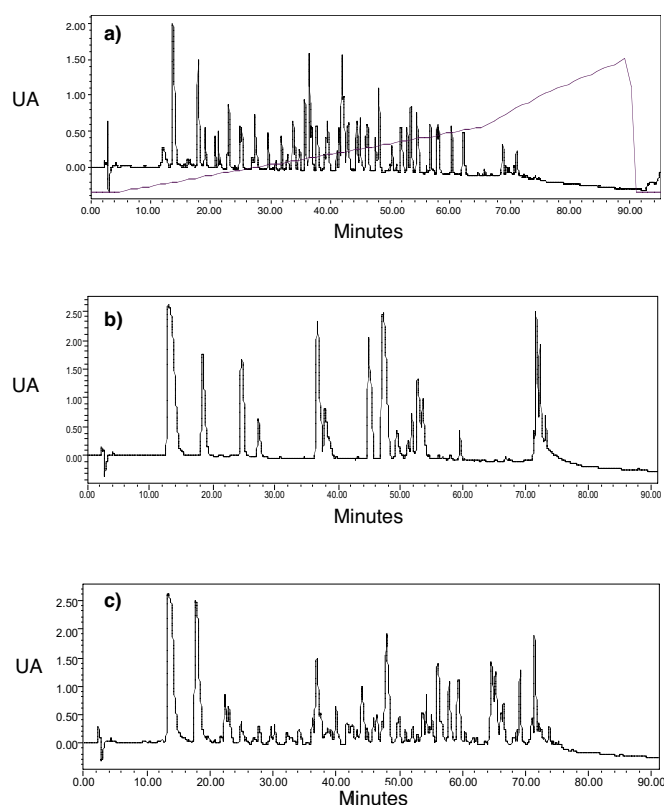


Figure 7. Chromatogram at 214 nm of BSA digest and elution gradient used (acetonitrile evolution gradient is given) (a), α -lactalbumin (b) and a mixture of BSA, α -lactalbumin, ovalbumin and β -lactoglobulin (c).

As a result, 22 peptides were tagged (Figure 8 and see Annexe 1), two of which corresponding to a doubly charge state of another tagged peptide. 7 of them were unidentified and 13 belonged to BSA. The tagged peptides peaks were well defined. 3 peptides belonging to BSA were tagged twice, showing the relevance of the study of the tagging of peptides with more than one cysteine.

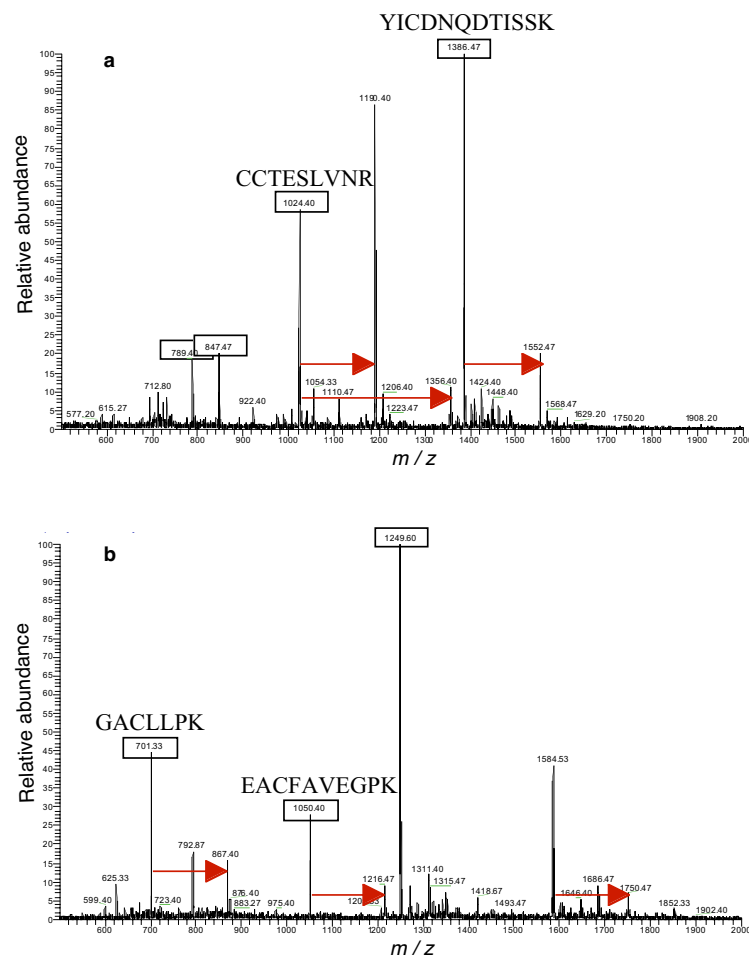


Figure 8. Fractions 34 (a) and 40 (b) of digested BSA infused in the presence of **1d**. Arrows show mass shift corresponding to the tag. Box m/z correspond to BSA matched peptides.

The entire list of peptides (Annexe 1) was submitted to Mascot sequence query²⁸ for a search in SwissProt³¹ database without giving any information on the taxonomy of the studied protein (see experimental part for the search parameters).

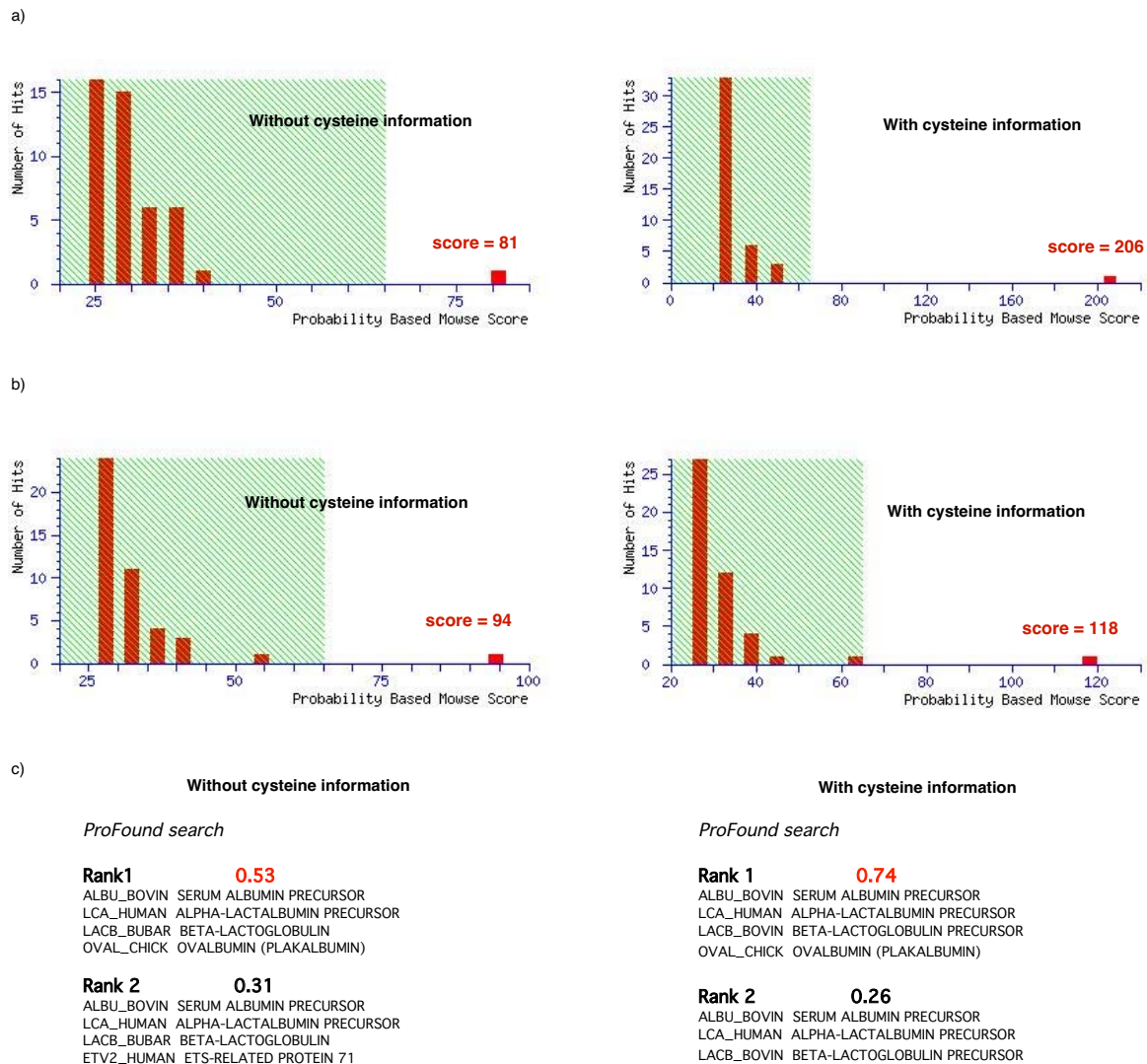


Figure 9. Identification scores without and with the cysteine content information content obtained thanks to on-line EC-tagging for BSA (a) and α -lactalbumin (b) using Mascot and for a mixture of BSA, α -lactalbumin, ovalbumin and β -lactoglobulin (c) using ProFound.

Without giving information on cysteine content, BSA was identified with a score of 81 and protein coverage of 74% (63 peptides matched). When the information on the cysteine content of the tagged peptides was added (the number of cysteines is specified for

cysteinyl containing peptides by “comp(number[C])” entered after the peptide mass), the score raised to 206 (Figure 9a).

In comparison, the identification procedure was carried out on an iodoacetamide alkylated BSA sample (see experimental part). In this case, BSA was identified with a score of 83 (coverage 61%, 56 peptides attributed to BSA). As this last experiment did not provide the same coverage from the precedent one, we decided to simulate the alkylation by iodoacetamide (+57 Da) of all cysteine-containing peptides based on the peptide list of the untagged BSA search. The score gave 89 (coverage 74%). Thus, the alkylation with iodoacetamide does not bring the same information to constrain the search as the indication of the number of cysteines. The determination and indication of peptide cysteine content^{9,18} with EC-tagging method was found to greatly improve the protein identification compared to first, the untagged sample and second, the sample alkylated with iodoacetamide.

Moreover, when only the 22 tagged peptides were entered with their cysteine content, the score was 160 with 22% coverage (13 peptides found) proving that the cysteine content information is a powerful data and let think that the selection of cysteine containing peptides coupled to the on-line EC-tagging method could be considered.

Similarly, human α -lactalbumin, which contains only 8 cysteines, was subjected to the mass fingerprinting EC-tagging experiment (Figure 7b). 23 peptides were found by MS analysis (Annexe 2). Without specifying the cysteine content, the score was 94 with 44% coverage (12 peptides corresponded to the considered protein). When cysteine information was added, the score reached 118 (Figure 9b). This time, only 11 peptides were matched, due to the fact that peptide at $m/z = 1693.5$ Th ($[M+H]^+$) exhibited only one tag instead of two ($m/z = 2025.5$ Th ($[M+2tags+H]^+$)) because of the limitation of the mass spectrometer to $m/z < 2000$ Th. As this peptide contained in fact two cysteine residues, Mascot did not identify it as belonging to the protein of interest. Taking into account two cysteines in this peptide for the Mascot search, the score would have reached 147.

When the only 6 tagged peptides were entered, the score was 98 (16% coverage). Only 4 peptides were recognized to belong to the protein. One peptide, at $m/z = 1693.5$ Th

($[M+H]^+$) was tagged once instead of twice (mass range limitation of the apparatus). This peptide was not taken into account by the software. The other one was generated by 2 miscleavages and was excluded by Mascot (1 maximum miscleavage was specified for the search). As a conclusion, it illustrates well the fact that cysteine content knowledge is powerful additional information in database query.

3.5. Mass fingerprinting of a mixture of 4 proteins using on-line electrochemical tagging

BSA, human α -lactalbumin, ovalbumin and β -lactoglobulin from bovine milk were submitted to reduction, proteolytic digestion followed by RP-HPLC separation (Figure 7c and see Annexe 3). The identification search was performed using ProFound³⁰, which allows searching for a protein mixture. The taxonomy Chordata was chosen for the search. Without tagging information, the mixture of four proteins was found with a probability of 53%. When cysteine information was specified in extra settings function, the probability reached 74% and the discrimination compared to other possibilities was enhanced (Figure 9c). Moreover, the time search was improved (2.52 s and 1.56 s without and with tagging information respectively). For ovalbumin, 6 peptides were matched without being tagged. For β -lactoglobulin, 1 cysteine residue was detected whereas in BSA, 8 cysteines were found and 7 in α -lactalbumin. The peptides, which were expected to be tagged because of their cysteine content, were labelled in all cases.

Another interesting characteristic is that the method potentially allows differentiation of isobaric peptides. For instance, the peptide at $m/z = 933.3$ Th ($[M+H]^+$) could be attributed to both β -lactoglobulin and α -lactalbumin based on mass measurement only at the resolution of the ion trap. With the additional information on cysteine content, it could be unambiguously attributed to α -lactalbumin. As another example, the MS signal situated at $m/z = 701.3$ Th ($[M+H]^+$) could be attributed to two peptides by mass alone, one coming from β -lactoglobulin and the other one from BSA which contains one cysteine. In that case, the information on cysteine content was found as well to allow a correct attribution.

4. Conclusion

The successful on-line EC-tagging of several cysteine residues in peptides is possible with methoxycarbonyl-1,4-hydroquinone whereas hydroquinone and substituted hydroquinones bearing an electron-donating group are not efficient enough. Indeed, peptides containing up to three cysteine residues are successfully tagged with methoxycarbonyl-1,4-hydroquinone, even if two cysteine moieties are consecutive. The counting of cysteine residues is directly obtained from one experiment through the direct observation of untagged and tagged products in a single mass spectrum.

The EC-tagging of peptides was applied for the identification of proteins through their proteolytic digest thanks to peptide mass fingerprinting. The procedure is simple and efficient and does not cause loss of sample. The EC-tagging is a quasi-instantaneous reaction contrary to any off-line alkylation step. Several proteins and a model mixture of 4 proteins were successfully identified thanks to the determination of peptide cysteine content by MS EC-tagging. Even if the original peptide signal peak is divided according to the number of tags (by three at the maximum for a peptide containing two cysteine units for instance), the convenient information on cysteine content is greatly valuable. We showed that the identification score is improved compared to the case of untagged or iodoacetamide alkylated digest samples. The discrimination with other candidates is also enhanced and the search time is reduced. The procedure is capable of enhancing mass spectrometric peptide mapping strategies.

The model experiments, despite the large amount of starting material, are quite realistic since the analysed HPLC fractions, off-line collected, present peptide concentrations finally delivered to the MS comparable to those which would have been obtained from typical proteomics nanoLC procedures. However, the experiments described above were restricted to protein identification through peptide mass fingerprinting, which limits the approach to very simple protein mixtures, because of instrumental limitations: no on-line capillary or nanoLC system nor query tools (such as SEQUEST) were available to identify proteins from more

complex mixtures in a true shotgun approach, coupling large-scale sequencing projects with the amino acid sequence information that can be gleaned from MS/MS.^{6,35}

To push the technology a step further, the tag influence on the MS/MS sequencing should be investigated more systematically for gel-free bottom-up application. In any case, as the unmodified peptide is still present, MS/MS experiments could be performed on it as well as on the non-cysteinylyl containing peptides. The method has to be validated in a real shotgun approach with adequate analysis system in order to study complex protein mixtures. In particular, the power of the approach has to be explored in conjunction with automated data dependent MS/MS analysis, and should prove particularly useful for peptides that are not sequenced by MS/MS (two or three most abundant peptides in each spectrum).

The automation of the technique will be envisaged. It should be realized by the constant infusion of 2-methoxycarbonylhydroquinone (which was shown to be stable in solution²⁶) thanks to a T junction between the LC separation step and the microESI emitter. The tagging efficiency control can be adapted through the chip parameters or by the spray current regulation. Bioinformatics tools well adapted to the on-line EC-tagging method could be also developed.

Bibliography

- 1 R. Aebersold and M. Mann, *Nature*, 2003, 422, 198-207.
- 2 S.D. Patterson and R.H. Aebersold, *Nature Genet.*, 2003, 33, 311-323.
- 3 A.E. Ashcroft, *Nat. Prod. Rep.*, 2003, 20, 202-215.
- 4 K. Gevaert and J. Vandekerckhove, *Electrophoresis*, 2000, 21, 1145-1154.
- 5 F.E. Regnier, L. Riggs, R.J. Zhang, L. Xiong, P.R. Liu, A. Chakraborty, E. Seeley, C. Sioma and R.A. Thompson, *J. Mass Spectrom.*, 2002, 37, 133-145.
- 6 D.A. Wolters, M.P. Washburn and J.R. Yates, *Anal. Chem.*, 2001, 73, 5683-5690.
- 7 M.P. Washburn, D. Wolters and J.R. Yates, *Nature Biotechnol.*, 2001, 19, 242-247.
- 8 D. Fenyo, *Curr. Opin. Biotechnol.*, 2000, 11, 391-395.
- 9 S. Sechi and B.T. Chait, *Anal. Chem.*, 1998, 70, 5150-5158.
- 10 H. Zhang, W. Yan and R. Aebersold, *Curr. Opin. Chem. Biol.*, 2004, 8, 66-75.
- 11 M. Hamdan and P.G. Righetti, *Mass Spectrom. Rev.*, 2002, 21, 287-302.
- 12 D.C. Brune, *Anal. Biochem.*, 1992, 207, 285-290.
- 13 R.A. Jue and J.E. Hale, *Anal. Biochem.*, 1993, 210, 39-44.
- 14 N. Lundell and T. Schreitmuller, *Anal. Biochem.*, 1999, 266, 31-47.
- 15 B. Herbert, M. Galvani, M. Hamdan, E. Olivieri, J. MacCarthy, S. Pedersen and P.G. Righetti, *Electrophoresis*, 2001, 22, 2046-2057.
- 16 M. Adamczyk, J.C. Gebler and J. Wu, *Rapid Commun. Mass Spectrom.*, 1999, 13, 1813-1817.
- 17 S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold, *Nature Biotechnol.*, 1999, 17, 994-999.
- 18 D.R. Goodlett, J.E. Bruce, G.A. Anderson, B. Rist, L. Pasa-Tolic, O. Fiehn, R.D. Smith and R. Aebersold, *Anal. Chem.*, 2000, 72, 1112-1118.
- 19 M. Galvani, M. Hamdan, B. Herbert and P.G. Righetti, *Electrophoresis*, 2001, 22, 2058-2065.
- 20 T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.
- 21 G.J. Van Berkel and F.M. Zhou, *Anal. Chem.*, 1995, 67, 2916-2923.
- 22 F.M. Zhou and G.J. Van Berkel, *Anal. Chem.*, 1995, 67, 3643-3649.

- 23 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 24 G. Hignett, S. Threlfell, A.J. Wain, N.S. Lawrence, S.J. Wilkins, J. Davis, R.G. Compton and M.F. Cardosi, *Analyst*, 2001, 126, 353-357.
- 25 C. Roussel, T.C. Rohner, H. Jensen and H.H. Girault, *ChemPhysChem*, 2003, 4, 200-206.
- 26 C. Roussel, L. Dayon, H. Jensen and H.H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.
- 27 D.N. Perkins, D.J.C. Pappin, D.M. Creasy and J.S. Cottrell, *Electrophoresis*, 1999, 20, 3551-3567.
- 28 http://www.matrixscience.com/search_form_select.html.
- 29 W.Z. Zhang and B.T. Chait, *Anal. Chem.*, 2000, 72, 2482-2489.
- 30 http://prowl.rockefeller.edu/profound_bin/WebProFound.exe.
- 31 <http://www.expasy.org/sprot/>.
- 32 R.M. Silverstein, G.C. Basler and T.C. Morill, *Spectrometric Identification of Organic Compounds*, 5th ed., John Wiley & sons, New York, 1991.
- 33 K. Biemann, *Method Enzymol.*, 1990, 193, 886-887.
- 34 D.E. Mason and D.C. Liebler, *Chem. Res. Toxicol.*, 2000, 13, 976-982.
- 35 W.H. McDonald and J.R. Yates, *Dis. Markers*, 2002, 18, 99-105.

Annexe 1: Peptide masses found for BSA

517.20, 520.80, 545.33, 571.80, 597.33, 604.40, 609.27, 649.33, 649.33 comp(1[C]), 653.40, 660.33, 665.33, 689.33, 693.33 comp(2[C]), 701.33 comp(1[C]), 712.33, 720.47, 733.40, 740.40, 752.33, 760.40 comp(1[C]), 789.40, 792.87, 808.93, 818.33, 819.33, 820.47, 832.40, 834.40, 834.87 comp(1[C]), 838.87, 841.40 comp(1[C]), 847.47, 886.33, 900.87, 912.87 comp(1[C]), 922.40, 926.40, 927.47, 945.4, 982.00, 990.47, 1002.40, 1011.33 comp(1[C]), 1014.53 comp(1[C]), 1015.40 comp(1[C]), 1023.47, 1024.40 comp(2[C]), 1039.47 comp(2[C]), 1045.27, 1050.40 comp(1[C]), 1128.47, 1133.40, 1142.60, 1159.00 comp(1[C]), 1163.53, 1194.33, 1201.53, 1218.53, 1249.60, 1283.60, 1304.67, 1305.60, 1341.53, 1349.40 comp(2[C]), 1362.60, 1364.4 comp(2[C]), 1367.53, 1386.53 comp(1[C]), 1399.53, 1479.73, 1490.67, 1491.53, 1497.60 comp(1[C]), 1519.60 comp(1[C]), 1520.60, 1554.47 comp(1[C]), 1567.67, 1568.60, 1584.53 comp(1[C]), 1637.53, 1639.87, 1648.67, 1667.67 comp(1[C]), 1699.67, 1799.47, 1823.73, 1851.67, 1888.67.

Annexe 2: Peptide masses found for human α -lactalbumin

513.07, 523.13, 549.20, 595.40, 601.27, 607.20, 651.27, 664.27 comp(1[C]), 729.40, 754.93, 778.80, 888.40, 933.33 comp(1[C]), 1010.33, 1048.20 comp(1[C]), 1102.33, 1111.33, 1161.33 comp(1[C]), 1164.27, 1614.30, 1693.50 comp(1[C]), 1806.47 comp(1[C]), 1900.13.

Annexe 3: Peptide masses found for a mixture of BSA, human α -lactalbumin, ovalbumin and β -lactoglobulin from bovine milk

508.20, 545.27, 549.20 comp(1[C]), 573.27, 580.27, 601.27, 647.40, 649.27, 651.27, 664.13 comp(1[C]), 673.27, 674.27, 701.33, 701.40, 729.33, 729.40, 733.40, 789.33, 820.33, 837.40, 841.33 comp(1[C]), 847.40, 853.33, 888.40, 903.47, 916.27, 922.33, 927.27, 933.33

comp(1[C]), 949.33, 959.33, 982.40, 990.33, 1003.27, 1003.33, 1014.40, 1024.27
comp(2[C]), 1048.33 comp(1[C]), 1102.33, 1102.40, 1142.93, 1161.47 comp(1[C]), 1163.40,
1193.53, 1249.33, 1283.47, 1305.47, 1345.60, 1349.33 comp(2[C]), 1362.47 comp(1[C]),
1399.40, 1479.53, 1519.40 comp(1[C]), 1567.47, 1571.47, 1572.60, 1614.33, 1635.33,
1639.73, 1658.47 comp(1[C]), 1680.53, 1684.54, 1687.53, 1692.60, 1693.33 comp(1[C]),
1694.53, 1806.47 comp(1[C]), 1858.600.

CHAPTER V. *Characterization of polyethylene terephthalate microspray emitter for mass spectrometric EC-tagging*

1. Introduction

Microfluidic systems are increasingly used in biology, chemistry and biotechnology.¹⁻³ Indeed, the small diffusion distance and the large surface-to-volume ratio provide shorter time scale and more efficient processes. For instance, microchannel reactor devices have been developed for analytical and synthetic purposes⁴, and combined with micromixing tools.⁵ In proteomics particularly, microfluidic reactors have been used for measuring enzyme kinetics⁶ and performing enzymatic synthesis.⁷ On-line peptide mapping has been achieved by the hyphenation of tryptic digestion in a microchannel device and electrospray ionization mass spectrometric (ESI-MS) detection.⁸ Other MS hyphenated techniques like on-line on-chip post-column derivatization reactions have been performed for enhancing the degree of ionization of amines, aldehydes and ketones.⁹

Miniaturized electrospray interfaces fabrication has been widely investigated for ESI-MS. These devices have been shown to provide better sensitivity than standard ESI sources by the improvement of desolvation and ionization efficiency.^{10, 11} Most work has employed glass or silicon substrates, which are hydrophilic and lead to the wetting of the surface, preventing the formation of a well focussed electric field essential for the generation of a stable electrospray.^{12, 13} Polymer-made microchips addressed this issue and are attractive

for various reasons including the increased reproducibility of the devices, the device quality and the low cost batch production.¹⁴

The manufacturing of microchip interface for on-chip combinations of electrochemistry and sheathless ESI-MS has recently been described with an array of microcoil electrodes integrated in the ESI emitter.¹⁵ A thin channel, planar electrode emitter device has also been used for the study and the control of electrochemical oxidation of several compounds in positive ESI mode. The electrode material, the electrode area, the channel height above the electrode and the solution flow rate have been evaluated as regards the control of the mass transport to the electrode surface.¹⁶

In our laboratory, a microspray emitter comprising a microband carbon electrode has been developed based on the micromachining of polyethylene terephthalate (PET) substrates by ultra violet (UV) laser ablation.¹⁷ From this technology, the inherent electrochemistry of the ESI¹⁸ has been exploited to perform electrochemically-assisted reaction within the interface. On-line electrochemical tagging (EC-tagging) has been developed for the modification of proteins¹⁹ and peptides.²⁰ In this chapter, the characterization of the microspray interface is carried out in terms of EC-tagging. The EC-reactor capabilities of the device were studied and the evaluation of the efficiency of the tagging of cysteines in peptides by electrogenerated benzoquinone tags (BQ) is proposed. A new microspray chip design with an on-chip microband-electrode-array was investigated as regards the tagging efficiency of the cysteine in chorionic gonadotropin- β (109-119) amide.

2. Experimental

2.1. Chemicals

1,4-Hydroquinone (> 98%) and lithium trifluoromethanesulfonate (purum) were purchased from Fluka (Büchs, Switzerland). Methoxycarbonyl-1,4-hydroquinone (methyl-2,5-dihydroxybenzoate, 99%) and ferrocenemethanol (97%) were from Aldrich (Milwaukee, WI, USA). Chorionic gonadotropin- β (109-119) amide (> 94%) was bought from Bachem

(Bubendorf, Switzerland). Methanol (> 99.8%, Riedel-de Haën, Seelze, Germany) and acetic acid (> 99.5%, Fluka) were used without any further purification. Deionised water (18.2 M Ω ·cm) was prepared using a Milli-Q system from Millipore (Bedford, MA, USA). Synthetic peptide AIKCTKF (>70%) was prepared by Catherine Servis at the Institut de Biochimie (Faculté de Médecine, Epalinges, Switzerland).

2.2. Microchip device

The fabrication of the microspray interface has been previously described (see also appendix I for a comprehensive description).¹⁷ Briefly, a PET substrate (100 μ m thick Melinex[®] sheet from Dupont, Wilmington, DE, USA) was photoablated with a UV excimer laser (ArF 193 nm from Lambda Physik, Göttingen, Germany). The dimensions of the electrodes, that were made of carbon paste Electrador ED 5000 (Electra polymer & Chemicals Ltd. Tonbridge, UK), were 70 \times 25 μ m² integrated in the channel of 35 \times 30 μ m² section. The distance from the electrode (*N.B.* the electrode which is the furthest away from the outlet in the case of the microelectrode-arrays) to the outlet of the microchip was 2 cm. For microelectrode-arrays, the distance between each electrode was 1 mm. The dimensions were controlled using a measuring microscope MM40 from Nikon (Egg, Switzerland). The surface topology was investigated by scanning electron microscopy (SEM) (Philips XL30 FEG from Philips Electron Optics, Eindhoven, the Netherlands).

2.3. Electrochemical set-up

The electrochemical measurements were recorded on an Autolab PGSTAT 30 potentiostat from Metrohm (Hesirau, Switzerland). Cyclic voltammetry (CV) was carried out using an undivided cell with a protected saturated calomel electrode as reference electrode, a glassy carbon or a carbon screen-printed (with Electrador ED 5000) electrode as working electrode and a platinum wire as counter electrode. The diffusion coefficient of ferrocenemethanol was measured by chronoamperometry, using the same set-up with a glassy carbon electrode (3 mm diameter) as working electrode. The working electrode was carefully polished with a suspension of 0.3 μ m alumina (Buehler, Lake Bluff, IL, USA) before every experiment. The potential was set for 120 s at 0.5 V and the oxidation current was recorded

every 1 s. The diffusion coefficient was extracted from the Cottrell equation $I(t) = f(t^{-1/2}) = zFD^{1/2}AC\pi^{-1/2}t^{-1/2}$ where C and D are respectively the bulk concentration and the diffusion coefficient of the electroactive species and A is the surface of the electrode. The diffusion coefficient of ferrocenemethanol was found to be $5.9 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$ in MeOH / H₂O / AcOH 50% / 49% / 1%.

The microspray emitters were characterized by chronoamperometry at 1.5 V (the working potential was determined by linear sweep voltammetry (LSV) at $20 \text{ mV} \cdot \text{s}^{-1}$) for 240 s with a solution of ferrocenemethanol 1 mM and LiCF₃SO₃ 0.1 M in MeOH / H₂O / AcOH 50% / 49% / 1%. The oxidation current was recorded every 1 s. The microchip was mounted on a holder with a stainless steel tube used for the fluidic connection and as counter electrode (two-electrode set-up). The working electrode was the inlaid carbon electrode of the microchip. The device was coupled to a syringe pump (kdScientific, Holliston, MA, USA) to introduce the solution and the flow rate (F_V) varied from 100 to 1000 nL·min⁻¹. Before each experiment the microchips were cleaned with five alternative successive steps of 2 V for 60 s and -2 V for 60 s under a water flow of 250 nL·min⁻¹.

2.4. Mass spectrometry

A LCQ DUO ion trap mass spectrometer (Finnigan, San José, USA) was used. The heated capillary was kept at 200 °C. In each experiment, the ion transmission parameters were optimised automatically in order to improve the detection of the analyte of interest (the unmodified peptide). The ESI interface was removed and the microchip holder was mounted on the probe slide adapter of the mass spectrometer. The device was coupled to a syringe pump to introduce the solution. The flow rates varied from 125 to 500 nL·min⁻¹ and the applied voltage was 3.5-4 kV. The distance from the outlet of the microchip to the entrance of the spectrometer varied between 1 and 2 cm in order to optimise the signal and the trap injection time. The spray current (I_{spray}) was set at 120 nA.

The spray medium was MeOH / H₂O / AcOH 50% / 49% / 1%. The concentration of the peptide was 50 μM and the concentration of hydroquinone compounds (HQ) varied from 250 μM to 20 mM. The solutions were not degassed.

3. Results and discussions

In EC-tagging, cysteine-containing peptides or proteins (P) infused in the presence of hydroquinone compounds (HQ) are tagged by the benzoquinones (BQ) generated at the microband electrode of the electrospray source (see appendix I for comprehensive description of the microspray emitter). The electrospray source used in MS is a two-electrode, controlled-current electrolysis flow cell. For a better understanding and control of the EC-tagging, the characterisation of the microspray emitter is essential.

3.1. Characterization

Electrochemical behaviour:

Carbon electrodes have many desirable features such as electrochemical inertness over a wide range of potential, high hydrogen and oxygen evolution overvoltages, low background currents and high electrical conductivity. The carbon ink used for the microspray emitter fabrication is composed of three principal components: organic polymer binder, solvent and graphite. During the curing process, the solvent is evaporated, leaving a printed binder and carbon mix. The properties of this carbon ink are different from a glassy carbon electrode as confirmed by cyclic voltammetry (CV) experiments (Figure 1). For instance, the difference between the oxidation and reduction peak potentials of ferrocenemethanol (ΔE_p) is 83 mV and 137 mV using respectively the glassy carbon electrode and the screen-printed electrode. This different behaviour may be attributed to electron transfer hindrance and changes in the diffusion properties due to the microelectrode character of the carbon ink electrode. The exposure to the UV light from the laser excimer that occurs during the microspray emitter fabrication may activate the electrode surface by removal of organic substances.²¹ However, the electron transfer rates should remain low compared to the glassy carbon electrode all the more so as the time of light exposure on the carbon is especially short during the fabrication of the microspray emitter (see appendix I).¹⁷

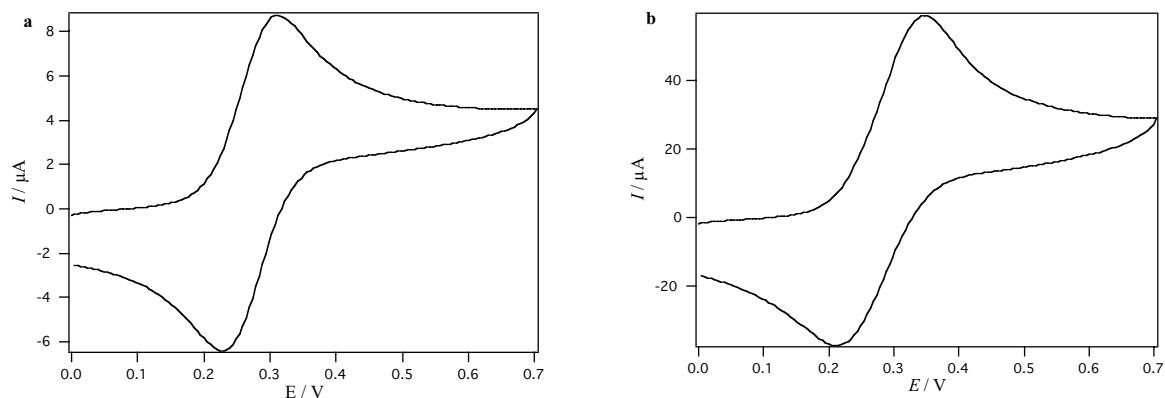


Figure 1. CV at $50 \text{ mV}\cdot\text{s}^{-1}$ of ferrocenemethanol 1 mM in MeOH / H₂O / AcOH 50% / 49% / 1% with LiCF₃SO₃ 0.1 M using a glassy carbon electrode (a) or a carbon screen printed with Electrador ED 5000 electrode (b) as working electrode.

After evaluation of the electrode material, the electrochemical behaviour of the microspray emitter was tested by linear sweep voltammetry (LSV) and chronoamperometry.²² The LSV of ferrocenemethanol shows that the plateau of the oxidation is not attained ideally because of the large IR drop intrinsic to the two-electrode set-up (Figure 2).

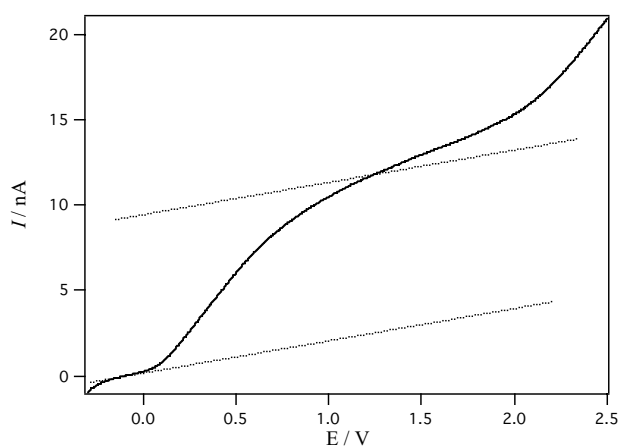


Figure 2. LSV at $20 \text{ mV}\cdot\text{s}^{-1}$ of ferrocenemethanol 1 mM in MeOH / H₂O / AcOH 50% / 49% / 1% with LiCF₃SO₃ 0.1 M in the microspray emitter at a flow rate (F_V) of $250 \text{ nL}\cdot\text{min}^{-1}$.

The choice of ferrocenemethanol for the electrochemical studies instead of hydroquinone compounds is justified by the electrochemical reversibility and the low oxidation potential of the compound. Because of the slow electronic exchange and their higher oxidation potential, 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone used for the EC-tagging of cysteines cannot be considered directly in the hydro-organic spray medium for the characterization of the microspray emitter, especially due to the IR drop that shifts the oxidation potential to more positive value when using a two-electrode set-up. The steady state feature of the flow electrolysis was confirmed, as the current intensity of the pseudo plateau does not depend on the scan rate of the LSV.

Theoretically, the convection-diffusion limiting current (I_{lim}) in a 2D laminar Poiseuille flow for a band microelectrode can be described by the following equation known as Levich equation²³⁻²⁵

$$I_{\text{lim}} = 0.925zFCl(LD)^{2/3} \left(\frac{F_V}{h^2d} \right)^{1/3} \quad (\text{Equation 1})$$

where z is the number of electrons per oxidized molecule, F is the Faraday constant, C and D are respectively the bulk concentration and the diffusion coefficient of the electroactive species, l and L the width and the length of the electrode, F_V the pressure-driven flow rate, $2h$ and d the height and the width of the channel.

To check if Equation 1 applies in the microspray emitter, chronoamperometry experiments were performed. The method is notably less influenced by IR drop when reaching the diffusion-controlled current. The curves $I = f(t)$ (*i.e.* the current (I) was recorded as a function of time (t)) were obtained with ferrocenemethanol at a difference in voltage (DV) of 1.5 V under a flow rate of 250 nL·min⁻¹ (*i.e.* the typical flow rate used for EC-tagging) (Figure 3a). The curves reveal that adsorption of the electroactive species occurs on the electrode material as the current continuously decreases with time. Passivation of the electrode occurs, as the steady state value is never reached. The experimental I_{lim} will also be approximated at relatively short times because of this electrode deactivation.

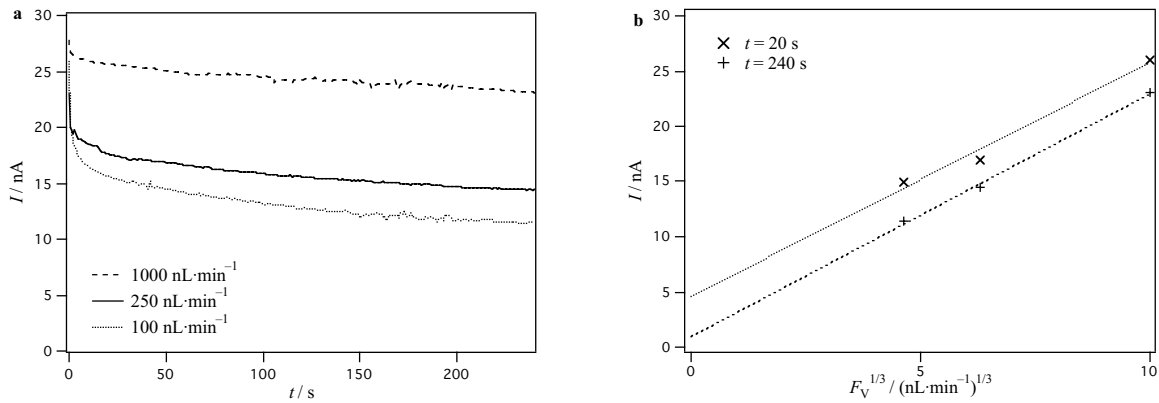


Figure 3. Chronoamperometric response of ferrocenemethanol 1 mM in MeOH / H₂O / AcOH 50% / 49% / 1% with LiCF₃SO₃ 0.1 M at 1.5 V in the microspray emitter at flow rates (F_V) of 100, 250 and 1000 nL·min⁻¹ (a). Measured convection-diffusion limiting (I_{lim}) current value at $t = 20$ and 240 s as a function of $F_V^{1/3}$ (b).

At a flow rate of 250 nL·min⁻¹, a current value of 17 nA was measured at $t = 20$ s in the emitter whereas the calculation gives an I_{lim} of 20 nA (15% of difference). Actually, Equation 1 is defined for an infinite plane system ($l \gg L$). The edge effects of the lateral walls of the microchannels are not negligible, resulting in a smaller experimental 3D value under pressure-driven flow. Experiments were then realized at different flow rates (Figure 1). Measurements gave $I_{lim} = 14$ and 26 nA for flow rates of 100 and 1000 nL·min⁻¹ respectively instead of 15 and 33 nA according to Equation 1 (respectively 7% and 21% of difference). The difference between the measurement and the model decreases according to the flow rate due to the decrease of the sharpness of the parabolic flow profile (*i.e.* lower heterogeneity of velocities). Besides, with a slightly different microchip design, P. Morier *et al.* have reported superior measured currents than calculated ones at low flow rates (*i.e.* 50-500 nL·min⁻¹).²⁶ In Figure 3b, the long-term current is showed to follow the Levich equation. The slope of the lines in Figure 3b provides the geometric factor of the microspray emitter. In view of the whole results, the Levich equation can be used as a first approximation of the I_{lim} in the following of the dissertation. Furthermore, the equation has been previously validated for similar microsystems^{26,27}, even in a similar case of protruding microband electrodes.²⁸

MS characterization:

The microspray emitter has been previously characterized as an efficient ionization source, providing a stable MS signal with time. The device can be used several times without failure in simple infusion of peptides and proteins and insures a good detection limit.¹⁷ Herein, the characterization of the microspray as a tool for the mass spectrometric EC-tagging is led. The tagging efficiency can be measured by MS as the apparent extent of the reaction (*a-TE%*), calculated as 100 fold the intensity of the peak of the tagged compounds over the sum of the intensity of the peaks of the untagged and tagged compounds. Experiments were performed with chorionic gonadotropin β (109-119) amide and AIKCTKF. Methoxycarbonyl-1,4-hydroquinone and 1,4-hydroquinone were used as EC-tagging reagents.

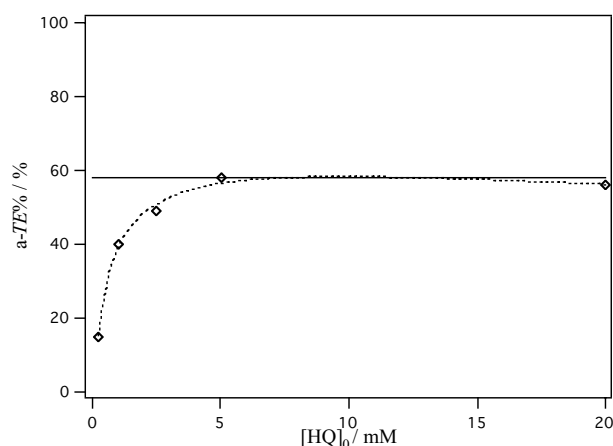


Figure 4. Apparent EC-tagging extent (*a-TE%*) for chorionic gonadotropin β (109-119) amide (50 μ M) as a function of the initial concentration of methoxycarbonyl-1,4-hydroquinone ($[HQ]_0$).

In Figure 4, the influence of the initial concentration of HQ ($[HQ]_0$) on the EC-tagging efficiency of chorionic gonadotropin β (109-119) amide is reported at a flow rate of 250 $\text{nL}\cdot\text{min}^{-1}$. The *a-TE%* increases with $[HQ]_0$ and reaches a plateau. Actually, when the spray current ($I_{\text{spray}} = 120 \text{ nA}$) is lower than the I_{lim} , the concentration of the produced BQ is fixed by the I_{spray} . In other words, the condition $I_{\text{lim}} > I_{\text{spray}}$ implies a 100% faradaic yield of

HQ, which is the most oxidable species present in solution.²⁹ The maximum a-TE% is reached for $[HQ]_0 = 5$ mM, which correlates with the calculated I_{lim} values of Table 1 (*i.e.* I_{lim} attains I_{spray}). According to these results, no parasitic reaction seems happen during the mass spectrometric EC-tagging. In addition, no adsorption occurs in view of the stability of the a-TE% with time.

$[HQ]_0$ / mM	0.25	1	2.5	5	20
I_{lim} / nA ^a	7	29	72	145	578

^a according to the Equation 1 where $C = [HQ]_0$, $D = 3.5 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$ determined in the experimental part of chapter VII for methoxycarbonyl-1,4-hydroquinone, $l = 25 \cdot 10^{-6} \text{ m}$, $L = 70 \cdot 10^{-6} \text{ m}$, $F_V = 250 \text{ nL} \cdot \text{min}^{-1} = 4.166 \cdot 10^{-12} \text{ m}^3 \cdot \text{s}^{-1}$, $2h = 35 \cdot 10^{-6} \text{ m}$ and $d = 30 \cdot 10^{-6} \text{ m}$.²³⁻²⁵ The Faraday constant (F) is 96500 C and two electrons are involved in the electrochemical process ($z = 2$).

Table 1. Theoretical convection-diffusion limiting current (I_{lim}) calculated according to the initial concentration of methoxycarbonyl-1,4-hydroquinone ($[HQ]_0$).

In Figure 5, the influence of the channel length (L_{ch}) from the electrode to the tip is studied for the EC-tagging of the synthetic peptide AIKCTKF.

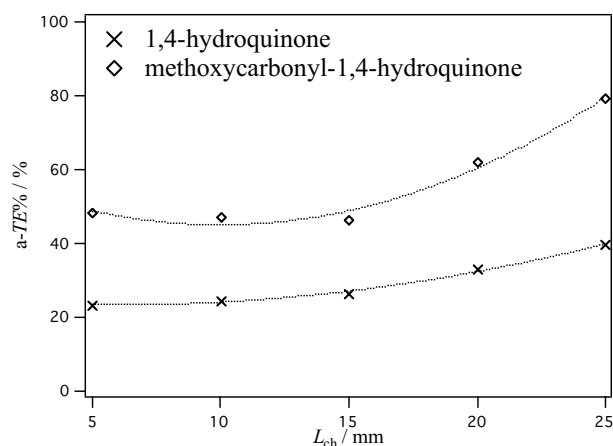


Figure 5. Apparent EC-tagging extent (a-TE%) of AIKCTKF (50 μM) using methoxycarbonyl-1,4-hydroquinone or 1,4-hydroquinone (2.5 mM) as a function of the microchannel length (L_{ch}) from the electrode to the tip.

The curves show that for short L_{ch} , the α -TE% is almost constant. The Taylor cone and the ESI plume may contribute to the tagging either by the efficient mixing of reactive species or the generation of BQ by oxygen contribution since the ESI process happens at the air. For 1.5-2.5 cm-long channels, the increase of the tagging is clearly observed with L_{ch} (*i.e.* the time of reaction). The chemical tagging reaction in the channel is also effective. With methoxycarbonyl-1,4-hydroquinone, the increase of the α -TE% is larger due to the greater reactivity of the generated BQ toward the 1,4-Michael addition of cysteine, which is kinetically controlled. The maximum tagging is not reached at 2.5 cm yet. In fact, the α -TE% depends highly on the reactivity of the cysteine residue in the molecular environment as well as on the ionization efficiency of the peptides and the tagged analogue in ESI-MS (see chapter VII for a comprehensive consideration of these factors).

3.2. Microelectrode-array

The influence of the electrode area on the EC-tagging efficiency was studied. Microelectrode-arrays were fabricated with three, five and eight electrodes of equal dimension. Figure 6 confirms the reproducibility of the electrode size and surface. For an eight-electrode-array, scanning electron microscopy (SEM) pictures of each electrode show similar surfaces.

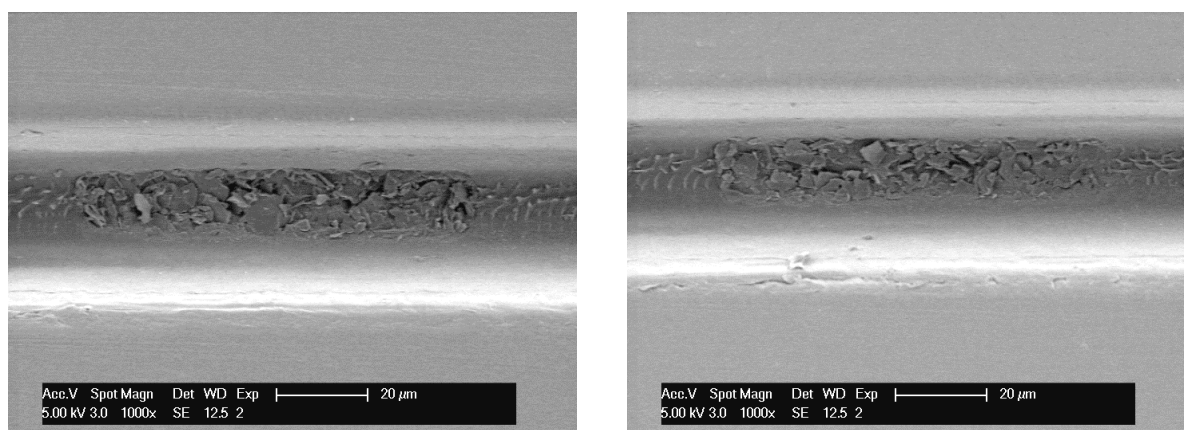


Figure 6. SEM picture of two different electrodes of an eight-microelectrode-array microspray emitter.

The microelectrode-array microspray emitters were evaluated by chronoamperometry at 1.5 V with ferrocenemethanol as previously described. As shown in Figure 7a, I_{lim} clearly increases with the number of microelectrodes (N). According to Equation 1, $I_{\text{lim}} = 20, 43$ and 82 nA should be obtained if the equivalent electrode length (L') is assumed in accordance to N (*i.e.* $L' = NL$). Comparison with the experimental current values clearly shows that this simple operation does not fit (Figure 7a) and that the microelectrode-array system behaves more complexly. As an explanation, the depletion in the concentration of the redox substrate along the channel can imply a decrease of the I_{lim} at the successive electrodes. Figure 7b illustrates that Equation 1 must be effectively corrected. Although the plots of $I = f(N^{2/3})$ fit as lines, the lines do not go through the origin of the graph. Simulation of the process with multiple microband electrodes may be carried out to investigate the process and draw the relevant associated parameters.

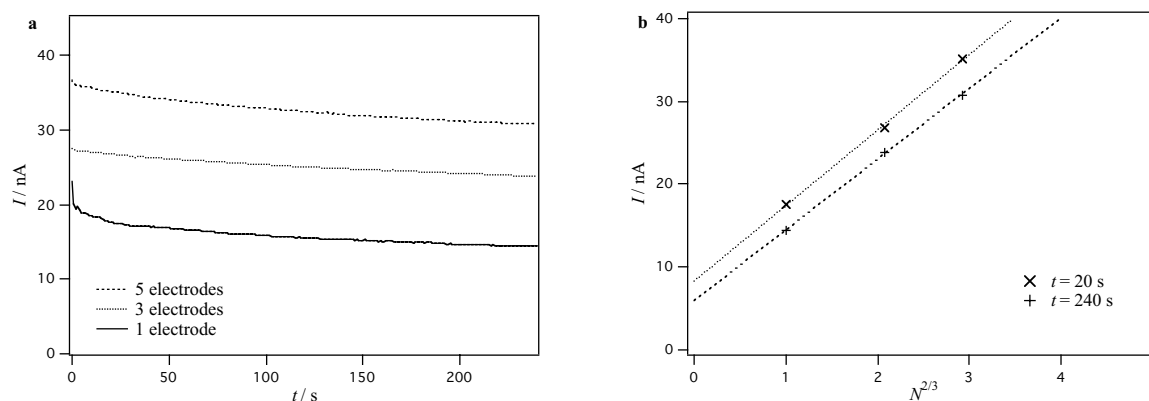


Figure 7. Chronoamperometric response of ferrocenemethanol 1 mM in MeOH / H₂O / AcOH 50% / 49% / 1% with LiCF₃SO₃ 0.1 M at 1.5 V at a flow rate (F_V) of 250 nL·min⁻¹ for microspray emitters presenting one, three and five inlaid microelectrodes (a). Measured convection-diffusion limiting (I_{lim}) current value at $t = 20$ and 240 s as a function of $N^{2/3}$ (b).

The microelectrode-array chips were tested for mass spectrometric EC-tagging. As before, I_{spray} was set at 120 nA and the flow rate was 250 nL·min⁻¹. At $[\text{HQ}]_0 = 20$ mM, the curve reveals a slightly decrease of the a-TE% as a function of N . In fact, in the electrode-

array configuration, the total I_{spray} is divided into currents i_n between each electrode as illustrated in Figure 8. It was assumed that I_{lim} is equal at each electrode and all i_n contribute to faradaic events.

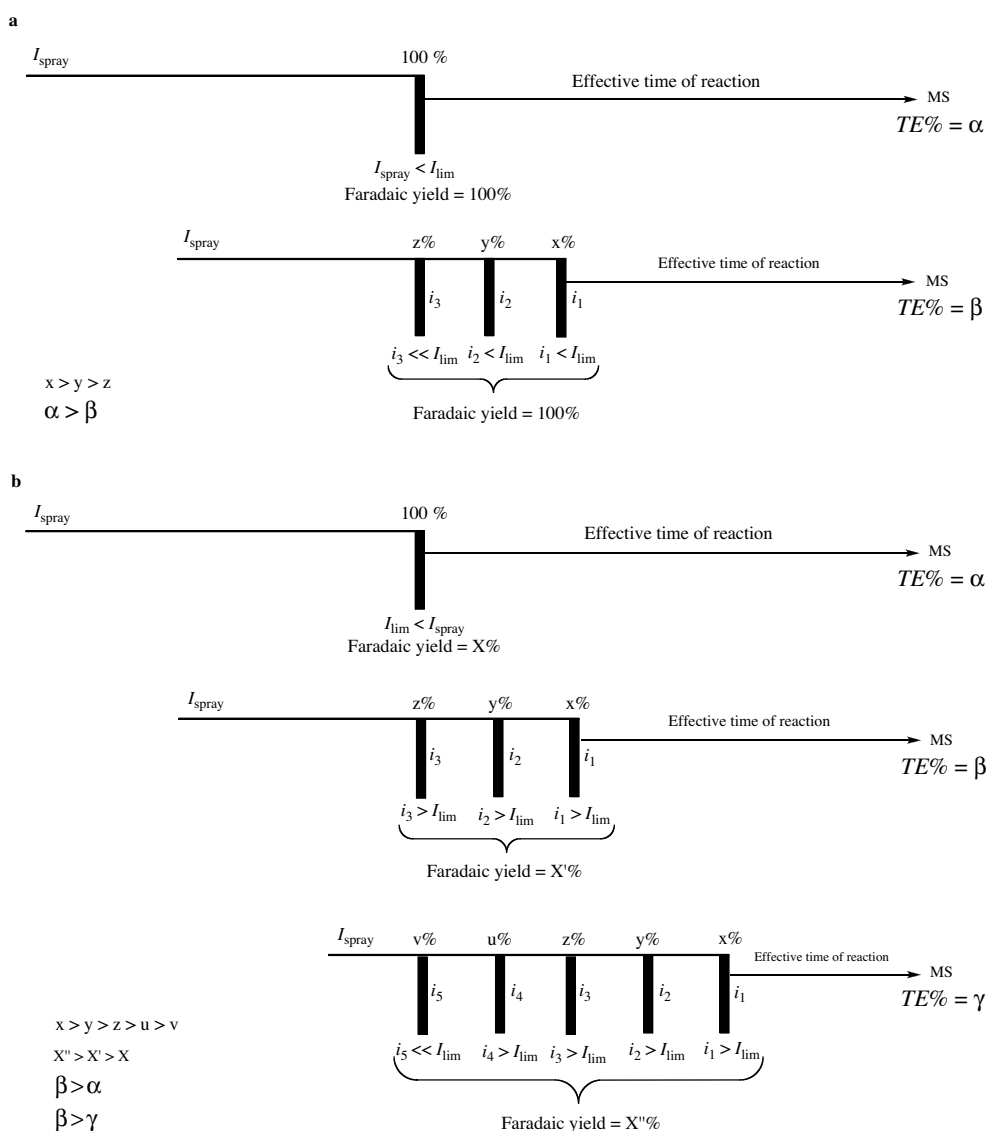


Figure 8. Scheme assumed to describe the influence of the number of electrodes on the EC-tagging extent ($TE\%$). When $I_{\text{spray}} < I_{\text{lim}}$ for the simple one-electrode microspray emitter, the increase of the number of electrodes (N) leads to a decrease of the effective tagging reaction time and maybe bypass currents. The $TE\%$ decreases (a). When $I_{\text{spray}} > I_{\text{lim}}$ for the simple one-electrode microspray emitter, the increase of N allows the increase of the faradaic yield and of the $TE\%$ (b).

These i_n are not equal according to the position of the electrode. It can be assumed that the larger currents are provided to the electrodes positioned closer to the MS ground. The decrease of the EC-tagging extent ($TE\%$) with N can be simply explained by the decrease of the time of chemical tagging reaction (Figure 8). While the electrode that provides the bigger current (*i.e.* where the bigger amount of BQ is produced) gets closer from the end of the channel according to N , the effective time of tagging reaction decreases. The final a- $TE\%$ is therefore decreased (see chapter VI for a comprehensive simulation study of the kinetics of the EC-tagging). Moreover, bypass currents may exist. In fact, the series circuit model of the ESI³⁰ may be much more complicated in this microelectrode-array configuration.

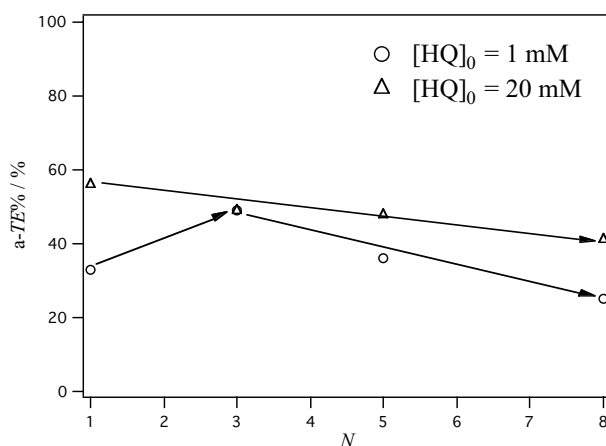


Figure 9. Apparent EC-tagging extent (a- $TE\%$) for chorionic gonadotropin β (109-119) amide ($50 \mu\text{M}$) using methoxycarbonyl-1,4-hydroquinone at 1 mM and 20 mM as a function of the number of microelectrodes (N) inlaid in the microchannel.

Working with high enough $[\text{HQ}]_0$ is sufficient to obtain a paramount a- $TE\%$ with the classic microspray emitter (*i.e.* one electrode). Nevertheless, working with high $[\text{HQ}]_0$ has been shown to be possibly detrimental to the MS analysis of proteins due to the decrease of signal to noise (see chapter III).²⁹ The EC-tagging of chorionic gonadotropin β (109-119) amide was carried out with methoxycarbonyl-1,4-hydroquinone at $[\text{HQ}]_0 = 1 \text{ mM}$ in order to see the influence of N when smaller $[\text{HQ}]_0$ are involved. At this concentration, the theoretical I_{lim} is 29 nA for the simple one-microelectrode microspray emitter. As the increase of the

electrode area (*i.e.* increase of N) augments I_{lim} theoretical value (Figure 8), the faradaic yield will followed, up to reaching 100% ($I_{\text{lim}} = I_{\text{spray}} = 120$ nA). Therefore, the $a\text{-TE}\%$ is expected to increase. The $a\text{-TE}\%$ is reported in Figure 9 according to N . For the five- and eight-microelectrode-array chips, a decrease of the tagging efficiency is obtained as previously observed at $[\text{HQ}]_0 = 20$ mM. However, this time, the best EC-tagging is achieved with the three-microelectrode-array microspray emitter. It finally confirms that the use of microelectrode-arrays is pertinent to improve and optimize the EC-tagging.

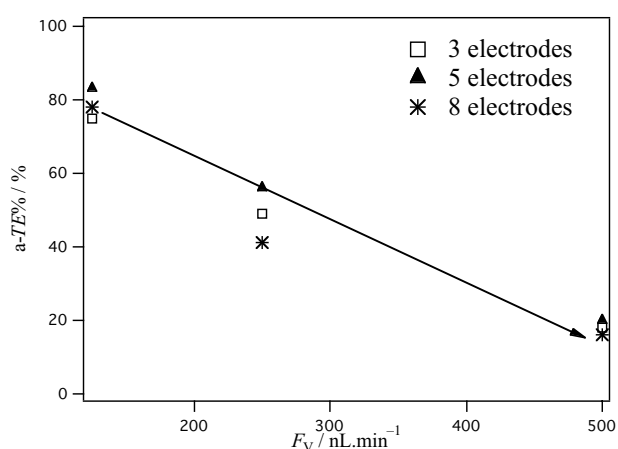


Figure 10. Apparent EC-tagging extent ($a\text{-TE}\%$) of chorionic gonadotropin β (109-119) amide (50 μM) in the presence of methoxycarbonyl-1,4-hydroquinone (20 mM) as function of the flow rate (F_v).

At last, experiments at different fluid flow rates were performed to see the influence on the $a\text{-TE}\%$. Experiments were performed from 125 to 500 $\text{nL}\cdot\text{min}^{-1}$. In Figure 10 are reported the data for the different microelectrode-arrays. The $a\text{-TE}\%$ increases while the flow rate decreases. Nevertheless, the lowest applicable flow rates are governed by the geometry of the microspray emitter. Using our polymer device, the spray was unstable when the flow rate was less than 100 $\text{nL}\cdot\text{min}^{-1}$. At 125 $\text{nL}\cdot\text{min}^{-1}$, the $a\text{-TE}\%$ reached 80% . When decreasing the flow rate, the generation of BQ is increased per volume unit (*i.e.* the residence time on the top of the electrode is increased). In addition, the residence time in the microchannel is longer (*i.e.* the time of the chemical reaction), providing superior $a\text{-TE}\%$.

Whatever the N constituting in the array, the results are quite similar, confirming the previous data when working with high initial concentration of HQ such as $[HQ]_0 = 20 \text{ mM}$.

4. Conclusion

The fabrication of the microspray device was shown to be well controlled and highly reproducible as regards the geometry and the surface properties. The EC-tagging experiments were operated between 125 to 500 $\text{nL}\cdot\text{min}^{-1}$. The electro spray and the EC-tagging process were demonstrated to be stable during hours. As regards the long-term stability of the devices, the microspray emitters were found to suffer from a decrease in efficiencies of EC-tagging after intensive utilization while simple spraying remained almost effective. Inadequate cleaning, fouling process and therefore parasitic reactions might be the source of it. The microfabricated microspray was shown to be an effective electrolysis flow cell. Because of the direct coupling to ESI-MS, the technique can be applied to the study of multiple electrochemical reaction as oxidative processes of biologically active compounds. The electrochemical behaviour of the interface was investigated with a model electrochemical system and the Levich equation was validated as a first approximation for the calculation of convection-diffusion limiting current in the microspray device.

In addition, a microelectrode-array design was tested. It was shown that microelectrode-arrays allow decreasing the initial concentration of the electrochemical tag. For instance, the best tagging efficiency was obtained using a three-microelectrode-array chip with methoxycarbonyl-1,4-hydroquinone at an initial concentration of 1 mM. A compromise between the time of the microchip fabrication and the need to reduce the concentration of electroactive probe has to be considered according to the experiment. At last, the results clearly showed that the EC-tagging is largely dependent on the flow rate of the fluid. Low flow rates increase the global electrochemical production of benzoquinone and the chemical tagging. The optimization of the residence time seems to be the proper way to control the EC-tagging degree. Finally, the microspray device can reveal useful for the study of the physical

chemistry of the electrospray ionization. The design of smaller microspray interface can be effectively carried out to range the nanospray scale.

Bibliography

- 1 P.A. Auroux, D. Iossifidis, D.R. Reyes and A. Manz, *Anal. Chem.*, 2002, 74, 2637-2652.
- 2 N. Lion, T.C. Rohner, L. Dayon, I.L. Arnaud, E. Damoc, N. Youhnovski, Z.Y. Wu, C. Roussel, J. Josserand, H. Jensen, J.S. Rossier, M. Przybylski and H.H. Girault, *Electrophoresis*, 2003, 24, 3533-3562.
- 3 D.R. Reyes, D. Iossifidis, P.A. Auroux and A. Manz, *Anal. Chem.*, 2002, 74, 2623-2636.
- 4 M. Miyazaki, K. Yamashita, Y. Yamaguchi, T. Honda, H. Nakamura, M. Fujii and H. Maeda, *New J. Chem.*, 2004, 28, 1622-1626.
- 5 A.D. Stroock, S.K.W. Dertinger, A. Ajdari, I. Mezic, H.A. Stone and G.M. Whitesides, *Science*, 2002, 295, 647-651.
- 6 G.H. Seong, J. Heo and R.M. Crooks, *Anal. Chem.*, 2003, 75, 3161-3167.
- 7 K. Kanno, H. Maeda, S. Izumo, M. Ikuno, K. Takeshita, A. Tashiro and M. Fujii, *Lab Chip*, 2002, 2, 15-18.
- 8 H.L. Wu, Y.P. Tian, B.H. Liu, H.J. Lu, X.Y. Wang, J.J. Zhai, H. Jin, P.Y. Yang, Y.M. Xu and H.H. Wang, *J. Proteome Res.*, 2004, 3, 1201-1209.
- 9 V. Spikmans, S.J. Lane, B. Leavens, A. Manz and N.W. Smith, *Rapid Commun. Mass Spectrom.*, 2002, 16, 1377-1388.
- 10 M. Wilm and M. Mann, *Anal. Chem.*, 1996, 68, 1-8.
- 11 G.A. Valaskovic, N.L. Kelleher, D.P. Little, D.J. Aaserud and F.W. McLafferty, *Anal. Chem.*, 1995, 67, 3802-3805.
- 12 R.S. Ramsey and J.M. Ramsey, *Anal. Chem.*, 1997, 69, 1174-1178.
- 13 Q.F. Xue, F. Foret, Y.M. Dunayevskiy, P.M. Zavracky, N.E. McGruer and B.L. Karger, *Anal. Chem.*, 1997, 69, 426-430.
- 14 J.S. Rossier, N. Youhnovski, N. Lion, E. Damoc, S. Becker, F. Reymond, H.H. Girault and M. Przybylski, *Angew. Chem.-Int. Edit.*, 2003, 42, 54-58.
- 15 G. Liljegren, A. Dahlin, C. Zettersten, J. Bergquist and L. Nyholm, *Lab Chip*, 2005, 5, 1008-1016.

- 16 G.J. Van Berkel, K.G. Asano and V. Kertesz, *Anal. Chem.*, 2002, 74, 5047-5056.
- 17 T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.
- 18 T.C. Rohner, N. Lion and H.H. Girault, *Phys. Chem. Chem. Phys.*, 2004, 6, 3056-3068.
- 19 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 20 L. Dayon, C. Roussel, M. Prudent, N. Lion and H.H. Girault, *Electrophoresis*, 2005, 26, 238-247.
- 21 M.D. Osborne, B.J. Seddon, R.A.W. Dryfe, G. Lagger, U. Loyall, H. Schafer and H.H. Girault, *J. Electroanal. Chem.*, 1996, 417, 5-15.
- 22 C. Amatore, M. Belotti, Y. Chen, E. Roy, C. Sella and L. Thouin, *J. Electroanal. Chem.*, 2004, 573, 333-343.
- 23 V.G. Levich, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, 1962, pp. 112-116.
- 24 P.R. Unwin and R.G. Compton, *Comprehensive Chemical Kinetics*, Vol. 29, Eds.: R.G. Compton and A. Hammet, Elsevier, Amsterdam, 1989, pp. 173-193.
- 25 H.H. Girault, *Analytical and Physical Electrochemistry*, 1st ed., EPFL Press, Lausanne, 2004, pp. 285-286.
- 26 P. Morier, C. Vollet, P.E. Michel, F. Reymond and J.S. Rossier, *Electrophoresis*, 2004, 25, 3761-3768.
- 27 I.E. Henley, K. Yunus and A.C. Fisher, *J. Phys. Chem. B*, 2003, 107, 3878-3884.
- 28 R. Ferrigno, P.F. Brevet and H.H. Girault, *J. Electroanal. Chem.*, 1997, 430, 235-242.
- 29 C. Roussel, L. Dayon, H. Jensen and H.H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.
- 30 G.S. Jackson and C.G. Enke, *Anal. Chem.*, 1999, 71, 3777-3784.

CHAPTER VI. *EC-tagging of peptides in a microchannel during microspray mass spectrometry: numerical simulation of consecutive reactions with multiple cysteines*^ξ

1. Introduction

In proteomics, the identities of most proteins represented in sequence databases can be determined by correlating mass spectrometric data with databases.¹ To narrow down possible matching candidates, specific searching constraints are needed. The mass mapping or mass fingerprinting of peptides derived from proteolytic digestion of a protein provides the basic constraint.²⁻⁷ However, a sufficient coverage of the protein (*i.e.* the determination of sufficient number of proteolytic peptides coming from a protein) is required to unambiguously identify a protein.

Tandem mass spectrometric (MS/MS) analysis of peptides in mixtures is the most common and most restrictive constraint use in addition to mass mapping.^{8,9} In most cases a collision-induced dissociation (CID) spectrum from a single peptide through electrospray (ESI) MS/MS is then sufficient to conclusively identify a protein. In this technique, peptides ions are sequentially selected for MS/MS out of a mixture.^{10,11} However, in the case of a complex mixture, the generation of CID spectra for all the components fails because of time

^ξ based on L. Dayon, J. Josserand and H. H. Girault, *Phys. Chem. Chem. Phys.*, 2005, 7, 4054-4060, promoted as an “hot article” in *Physical Chemistry and Chemical Physics*.

limitation. Automated analysis is thus routinely programmed to give the priority to peptides having the highest ion current. Therefore, when the mixture is complex, many detected ions but presenting low intensity are not fragmented, reducing the dynamic range of the method.

Procedures capable of enhancing the identification procedure are of great value. The accuracy in the masses determination is a mean of constraining the database searching.^{12,13} The molecular weight of the protein, the isoelectric point of the protein or of the tryptic peptides¹⁴, the presence of rare amino acid such as cysteine, methionine, or tryptophan in the peptide sequence bring powerful information to enhance the matching. The counting of cysteine in peptide using the isotopic signature of a specific tag for thiol^{12,15} or the differential analysis of unlabelled and labeled cysteinyl samples¹⁶ can be used to significantly improved protein identification by database searching.

The on-line electrochemical tagging (EC-tagging) of peptides by probes electrogenerated at the microband electrode of a microspray emitter¹⁷ have enabled the on-line counting of cysteine units in peptides¹⁸ to improve the processes of identification of model proteins.¹⁹ The tagging reaction happens in the microspray emitter that behaves as a flow electrolysis cell and is quasi-instantaneous. As post-column treatment, the technique would provide powerful information on cysteine content, notably for low intensity peptides that would not be selected for MS/MS. Here, the cysteine content determination relies on the control of the extent of the reaction to insure a minimum amount of the unmodified peptide together with a minimum amount of the fully tagged peptide. The number of cysteine units is determined by mass spectrometry as the number of mass shift(s) from the unmodified peptide. Because the starting peptide should not be totally converted to guaranty the success of the analysis, understanding the kinetics of the flow reaction is relevant to control the process and to generalize the technique.

To numerically simulate the EC-tagging of peptides, multi-stage chemical reactions should be considered in a fluid flow.^{20,21} Validated by comparison to previous works^{22,23}, a finite-element convection-diffusion-reaction model for an electrochemical

mechanism²⁴ has been also further developed for the consecutive tagging by markers electrogenerated at the bottom of a microchannel (2 cm length). Some numerical simulations are carried out for the addition from one up to five tags (multi-step consecutive reaction) to determine the major chemical and kinetic parameters involved in the microfluidic process. The present theoretical work establishes the range of optimal conditions to achieve the counting of cysteine units in peptides, in order to apply the counting/identification technique to complex protein mixtures.

2. Computational methods

2.1. Numerical technique

The finite-element formulation was generated on the numerical software Flux-Expert[®] (Astek Rhône-Alpes, Grenoble, France), in a 2D Cartesian form. It was operated on a Dell PC, 2 Go RAM (Red Hat Linux).

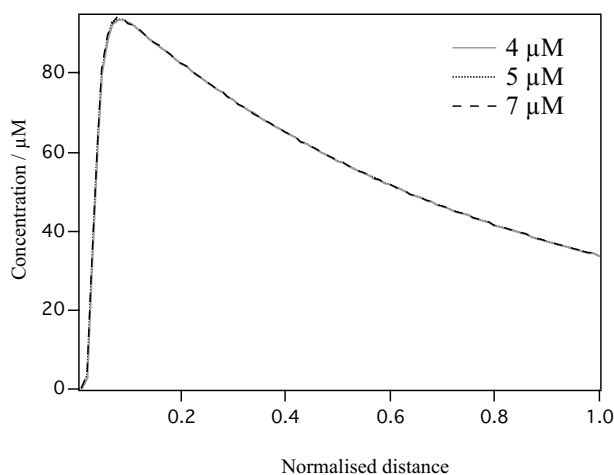


Figure 1. Meshing validation by numerical comparison of BQ concentration along the channel with main mesh size of 4 µm, 5 µm and 7 µm.

The mesh size adopted for the channel was $\delta = 5 \mu\text{m}$ and it was reduced (*i.e.* $0.4 \mu\text{m}$) at the electrode extremities in order to take into account the edge effects. The resulting mesh Péclet number $P_S = \bar{v} \delta / D$ was found to be 50 and 14 for the protein or peptide (P) and the benzoquinone (BQ) respectively. The error induced by a $7 \mu\text{m}$ channel mesh (compared to a $4 \mu\text{m}$ one) was checked to be 0.01 and 0.05% for BQ and P respectively at the end of the channel (Figure 1). It confirms the mesh Péclet number limit of 100 determined in the literature.²⁵ The total number of mesh is 30000, leading to a number of freedom degrees of 240000 for the eight unknowns in the five-tagging case. The conservation of species flux (BQ and P) was verified.

2.2. Scaling

Because of the 2 cm channel experimental length, a scaling was necessary to have an acceptable mesh Péclet number while keeping a meshing grid and a matrix size that can be numerically treated. Then, a 0.5 cm long geometry was taken but the channel height ($2h$) and the electrode length were kept the same (the electrode length remains quasi-negligible compared to channel length, *i.e.* the ratio is 1/50 instead of 1/200 for the experimental conditions). The flow rate was divided four times to insure the same residence time in the channel and the same transversal diffusion time. The incoming HQ concentration was adapted to have the same flux ratio between the incoming P and the BQ generated at the electrode. With the Levich equation^{26,27}, one can express a proportionality relationship between the flux rate of produced BQ at the electrode N_{BQ} and the flow rate $F_v = \bar{v} 2hd$ (where \bar{v} is the mean flow velocity, $2h$ and d the height and the width of the channel):

$$N_{\text{BQ}} \propto [\text{HQ}]_0 F_v^{1/3} \quad (\text{Relation 1})$$

From Relation 1, if the flow rate is divided 4 times, the flux of BQ is divided by $4^{1/3} = 1.6$. It has been checked by simulation that the ratio of flux of BQ at the electrode for the normal case ($\bar{v} = 4 \cdot 10^{-3} \text{ m} \cdot \text{s}^{-1}$) over the scaling by 4 ($\bar{v} = 1 \cdot 10^{-3} \text{ m} \cdot \text{s}^{-1}$) is 1.589 (≈ 1.6). Moreover, the flux of the protein N_P incoming into the channel is given by:

$$N_P = [\text{P}]_0 F_v \quad (\text{Relation 2})$$

Whatever the flow rate scaling, the ratio of these flux must be conserved:

$$\frac{N_{\text{BQ}}}{N_{\text{P}}} \propto \frac{[\text{HQ}]_0 F_{\text{V}}^{-2/3}}{[\text{P}]_0} \quad (\text{Relation 3})$$

To conserve the ratio $N_{\text{BQ}} / N_{\text{P}}$, if the flow rate is divided by 4, the initial concentration $[\text{HQ}]_0$ must be divided by $4^{2/3}$. The mean concentration $[\text{BQ}]$ remains unchanged by this operation as:

$$[\text{BQ}] = \frac{N_{\text{BQ}}}{F_{\text{V}}} \propto [\text{HQ}]_0 F_{\text{V}}^{-2/3} \quad (\text{Relation 4})$$

The scaling was fully validated by the numerical comparisons of concentration of BQ and PQ_3 along the channel for a scaling by 2 (channel length of 10 mm) and by 4 (channel length of 5 mm), which appear in Figures 2 and 3. Numerical simulation studies were pursued with the scaling of 4.

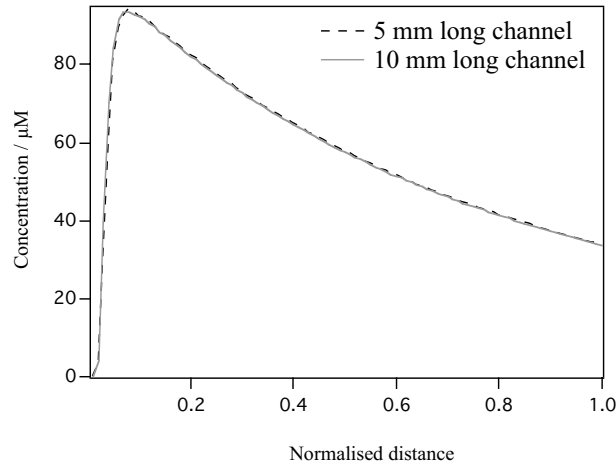


Figure 2. Scaling validation by numerical comparison of BQ concentrations along the channel for a scaling by 2 and by 4.

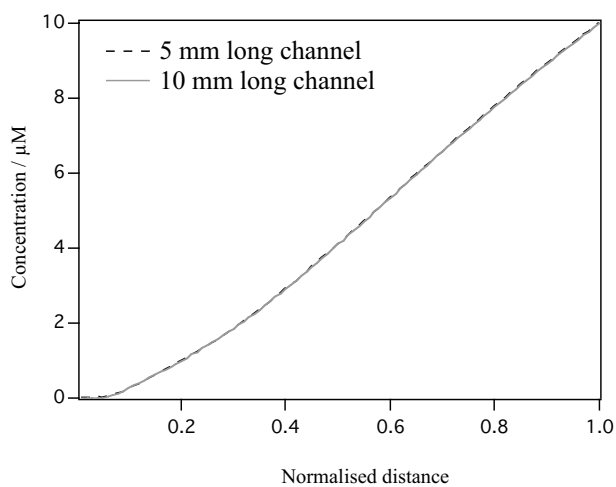


Figure 3. Scaling validation by numerical comparison of PQ_3 concentrations (three-cysteine target simulation) along the channel for a scaling by 2 and by 4.

3. Results and discussions

3.1. Microspray characteristic

The MS experimental set-up comprised the emitter shown in Figure 4, which behaves as an electrolytic-flow-cell reactor.

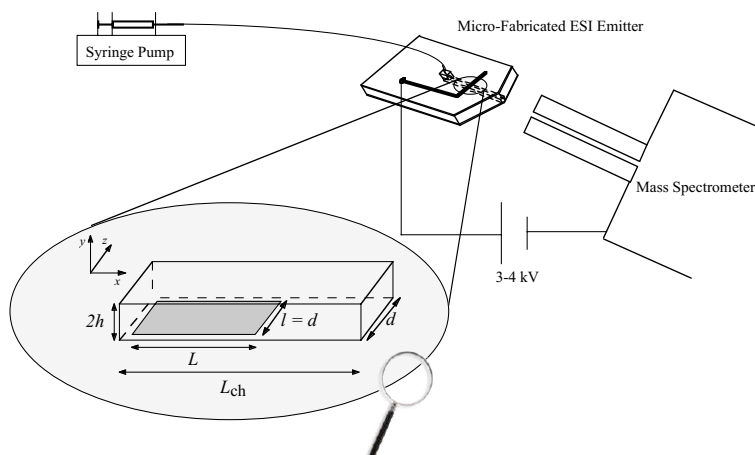


Figure 4. MS set-up using the microspray emitter. A zoom caption on the electrode is depicted.

The analyte is mixed with the electro-active probe prior to pressure-driven infusion through a capillary tube to the microspray emitter. The probes studied are 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone (HQ in general).

The overall tagging process can proceed *via* an Electrochemical-Chemical-Electrochemical (ECE) mechanism (Figure 5) where:

- the first Electrochemical reaction is the oxidation of HQ in 1,4-benzoquinone or methoxycarbonyl-1,4-benzoquinone (BQ in general)
- the Chemical reaction is the 1,4-Michael addition of the thiol functional group of the cysteine of a protein or a peptide (P) on the BQ ring that yields the products (PQ_i), where i represents the number of cysteines being tagged.
- the final Electrochemical reaction involves the oxidation of the adduct, but under the present flow conditions this second electrochemical oxidation has no time to occur.²⁴ It implies that only one thiol addition on the benzoquinone core can happen. Therefore, a simple EC mechanism will be considered in the present study.

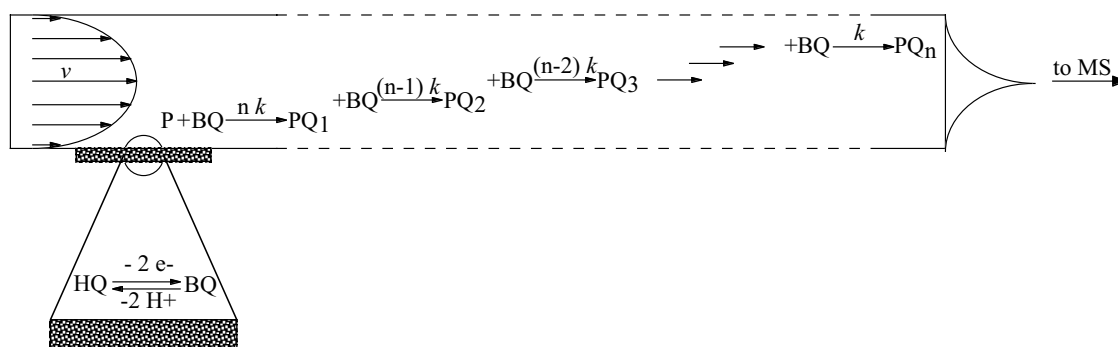


Figure 5. Scheme of the model including the consecutive reactions of tagging.

3.2. Numerical model

3.2.1. Model

The present model is developed for a five-cysteine-containing species at maximum. It addresses the convection-diffusion-reaction of the eight species i here considered (respectively HQ, BQ, P, PQ₁₋₅ for the hydroquinone, the benzoquinone, the protein or peptide, and the five possible successive degrees of the tagged protein or peptide). This transient model (Equation 1) is applied in a steady state regime to a 2D cross-section of the geometry (see assumptions and Figure 4):

$$\frac{\partial c_i}{\partial t} + \nabla \cdot (-D_i \nabla c_i + \mathbf{v} c_i) = R_i \quad (\text{Equation 1})$$

where c_i is the concentration of the species i , D_i its diffusion coefficient, \mathbf{v} is the fluid velocity vector and R_i is the rate of generation or consumption of i (Figure 5). The nabla symbol is used to simplify the notation. In the annexe, the global forms of these local equations are described for each species, by using the Galerkin formulation (finite-element method).

The tagging conditions and numerical parameters are given in Table 1 and in the computational section (Flux-Expert[®] software).

$2h$	35 μm	D_P	$1 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$
L	70 μm	D_Q	$3.5 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$
L_{ch}	0.5 cm ^a	$\bar{\mathbf{v}}$	$1 \cdot 10^{-3} \text{ m} \cdot \text{s}^{-1}$ ^a
$[\text{HQ}]_0$	0.993 mM ^a	v_{max}	$1.5 \cdot 10^{-3} \text{ m} \cdot \text{s}^{-1}$ ^a
$[\text{P}]_0$	50 μM		

^a scaled value

Table 1. Numerical reference parameters. A scaling has been performed for numerical simulations. The real values before scaling are: $L_{\text{ch}} = 2\text{cm}$, $[\text{HQ}]_0 = 2.5 \text{ mM}$ and $\bar{\mathbf{v}} = 4 \cdot 10^{-3} \text{ m} \cdot \text{s}^{-1}$. According to previous works, the width of the channel is 30 μm and the flow rate is 250 $\text{nL} \cdot \text{min}^{-1}$.^{18,19,24,28}

3.2.2. Assumptions

- The electrochemical reaction is assumed to be limited by the diffusion (fast electrode reaction) and HQ is assumed to be the only species oxidised at the electrode.
- The solution is assumed to be sufficiently diluted and isothermal so that the viscosity and the density of the fluid can be considered as not modified by concentration or temperature variations. The diffusion coefficients of the species are also treated as uniform on the entire study domain.
- The channel walls are considered to be smooth and the eventual migration effects due to the applied potential are neglected.
- The width d of the channel is assumed to be much larger than its height $2h$ so that the velocity gradient in the third dimension can be neglected (2D Cartesian assumption to overcome numerical limitations).
- The fluid is assumed to be Newtonian and its velocity is described according to a Poiseuille profile (laminar flow conditions, $Re = 0.035$).
- The reactivity of cysteines is taken to be equal at every site of the biomolecule (equal to that of cysteine amino acid) and any other parasite reactions are neglected.
- The numerical simulations of the tagging reaction are considered only in the channel, thereby neglecting the reactions in the Taylor cone or in the ESI plume. Previous experimental MS mono-tagging of synthetic peptide AIKCTK carried out with microchannel emitters of variable channel length have clearly showed that the channel contribution is predominant for a channel length of 2 cm with the present flow rate (see chapter V).
- The flow rate of the fluid is a fixed parameter. The simulations correspond to a flow rate of $250 \text{ nL}\cdot\text{min}^{-1}$ (*i.e.* a mean flow velocity $\bar{v} = 4\cdot 10^{-3} \text{ m}\cdot\text{s}^{-1}$ before scaling).

3.3. Kinetics of multi-tagging

The parameters playing a key role in the tagging final efficiency are investigated using the finite-element model (see computational methods). The multi-tagging process in a laminar flow is evaluated in terms of tagging extent TE ($= (\sum_{i=1\dots n} [PQ_i]) / [P]_0 = ([P]_0 - [P]) / [P]_0$, *i.e.* the consumption of the protein or peptide P).

The addition rate constants are assumed to be equal to those corresponding to the addition of L-cysteine on 1,4-benzoquinone and methoxycarbonyl-1,4-benzoquinone (respectively 210 and 5000 $M^{-1}\cdot s^{-1}$ in methanol / water / acetic acid 50% / 49% / 1%).^{28,29} For simplification, the diffusion coefficient of all the quinone probes are taken to be equal to $3.5\cdot 10^{-10} m^2\cdot s^{-1}$, that in fact corresponds to the diffusion coefficient of methoxycarbonyl-1,4-hydroquinone in methanol / water / acetic acid 50% / 49% / 1%.³⁰ The mean diffusion coefficient of the target biomolecule P was chosen as $1\cdot 10^{-10} m^2\cdot s^{-1}$.³¹ As the target biomolecule concentration is uniform at the inlet, its diffusion coefficient does not affect the final adduct amount (the diffusion coefficient of BQ plays an important role as BQ generated at the electrode diffuses from the bottom of the channel along the flow).²⁴ The initial concentrations are taken in accordance to previous experimental works that showed valuable analysis of proteins and peptides (Table 1).^{19,28}

Mono-tagging. Simulations were first run for a species containing a single thiol function. In Figure 6, a comparison of species distributions along the channel in its central portion ($y = h$) is proposed for the two probes considered for a single-cysteine target. The formation of adducts PQ_1 becomes efficient with methoxycarbonyl-1,4-hydroquinone ($k = 5000 M^{-1}\cdot s^{-1}$), inducing consequent consumption of both the biomolecule and the electrogenerated BQ.

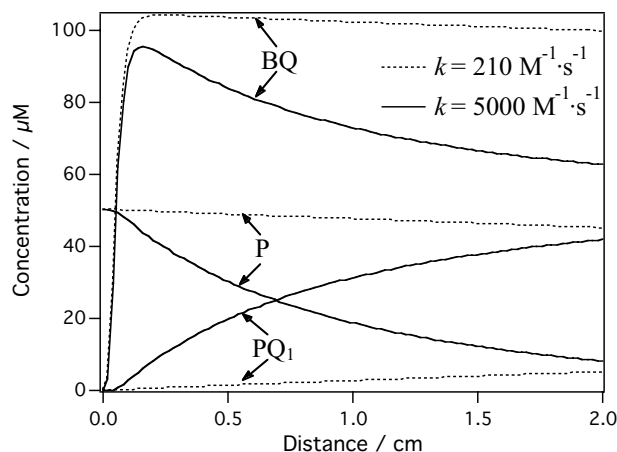
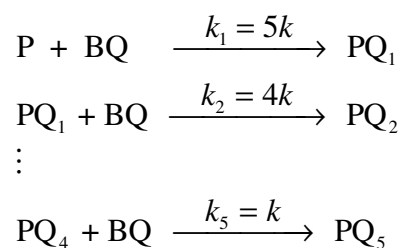


Figure 6. Distribution along the channel ($y = h$) of the species for a single-cysteine biomolecule with $k = 210 \text{ M}^{-1}\cdot\text{s}^{-1}$ and $k = 5000 \text{ M}^{-1}\cdot\text{s}^{-1}$. Initial concentrations of HQ and P are 2.5 mM and 50 μM respectively.

Multi-tagging. The same kinetic comparison was made for a biomolecule with five cysteine units (each cysteine site is considered to have the same labelling rate constant). For $k = 210 \text{ M}^{-1}\cdot\text{s}^{-1}$ (Figure 7a), the first adduct PQ_1 is the only species produced in a reasonable amount. The production is higher than in the case of a singly-cysteinyll target (15.3 μM instead of 5.0 μM at the outlet of the channel) since here, it possesses five cysteine groups. In fact, the probability for the probe to react with a cysteine in a target possessing five cysteines is five-fold higher than the probability to react with cysteine in a singly-cysteinyll target. The first step of the consecutive tagging reactions presents an apparent rate constant $k_1 = 5k$ since the rate law is here formulated as a function of the biomolecule concentration $[\text{P}]$ (Figure 5). Therefore, the reaction rate is multiplied as shown below.



Part of the first adduct PQ_1 reacts with BQ to give the successive adducts, the production of which is limited to PQ_2 for the present k value.

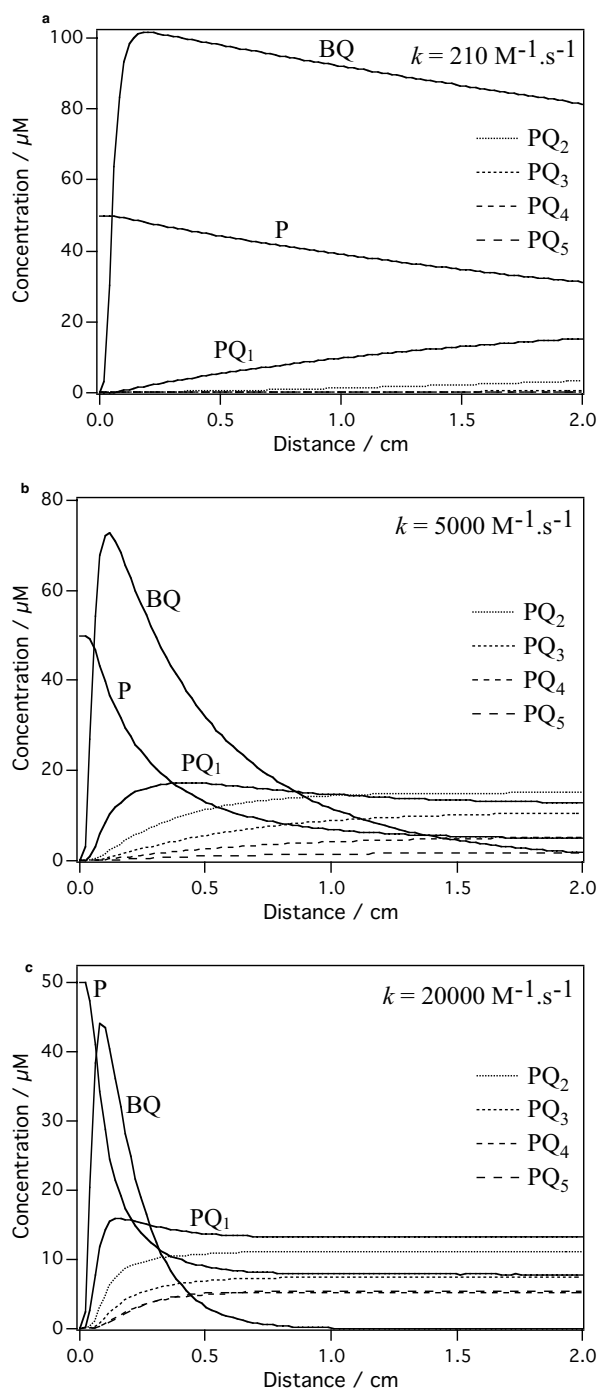


Figure 7. Distribution along the channel ($y = h$) of the species for a biomolecule with five cysteine units with $k = 210 \text{ M}^{-1}\cdot\text{s}^{-1}$ (a), $k = 5000 \text{ M}^{-1}\cdot\text{s}^{-1}$ (b) and $k = 20000 \text{ M}^{-1}\cdot\text{s}^{-1}$ (c). Initial concentrations of HQ and P are 2.5 mM and 50 μM respectively.

For $k = 5000 \text{ M}^{-1}\cdot\text{s}^{-1}$ (Figure 7b), almost all the BQ generated at the electrode is consumed at the end of the channel. The production of all adducts is enhanced, apart from the first one whose concentration decreased beyond a distance of 5 mm from the electrode (*i.e.* 1.25 s of reaction) to feed the following additions. The second adduct PQ_2 is the most favoured species at the end of the channel (15 μM). The fifth adduct PQ_5 is also observable in small amounts. Simulations with a rate constant $k = 20000 \text{ M}^{-1}\cdot\text{s}^{-1}$ were performed as shown in Figure 7c. The reaction is quite fast and the thermodynamic equilibrium is reached in about 2.5 s (*i.e.* 1 cm from the electrode). The electrogenerated benzoquinone, which is correlated to the BQ initial concentration, is in deficit with respect to cysteine units. It implies that the species P, PQ_1 , PQ_2 , PQ_3 , PQ_4 and PQ_5 remain unchanged for $x > 1 \text{ cm}$ (no more BQ to react). This result is not only the effect of the addition rate constant but also a consequence of the multi-tagging that amplifies the tendency.

Figure 8 reports the influence of the HQ initial concentration on the $TE (= (\sum_n [\text{PQ}_n]) / [\text{P}]_0 = ([\text{P}]_0 - [\text{P}]) / [\text{P}]_0)$ of a species containing five cysteine units.

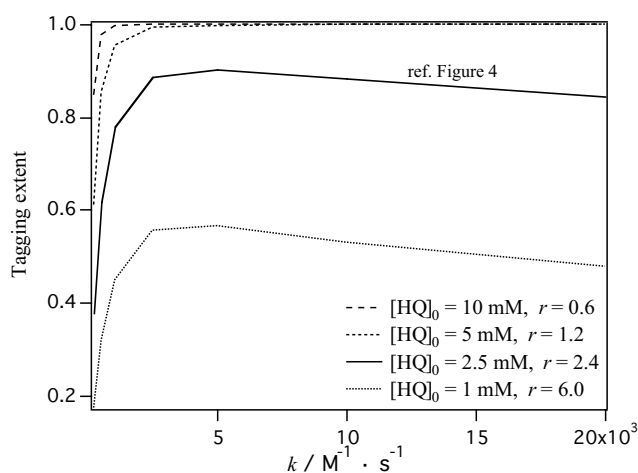


Figure 8. Tagging extent TE (*i.e.* consumption of P) obtained for several rate constants and initial HQ concentrations in the case of a five-cysteine biomolecule tagging (the ratio r between the flux of cysteine units and that of electrogenerated BQ is also indicated). *N.B.* Assuming a flat profile of velocity, the tagging extent is 0.984 with $[\text{HQ}]_0 = 2.5 \text{ mM}$ and $k = 5000 \text{ M}^{-1}\cdot\text{s}^{-1}$ for a five-cysteinyll target (instead of 0.902).

When the BQ is in excess ($[HQ]_0 \geq 6$ mM), a total conversion of P is obtained for k values above $2500 \text{ M}^{-1}\cdot\text{s}^{-1}$. On the other hand, for initial HQ concentrations of 1 and 2.5 mM, the electrochemically-produced BQ is in deficit with respect to the cysteine moieties and, for high k values, BQ is found to be totally consumed before the end of the channel when the protein is still present (see Figure 7c). The high kinetics limits the consumption of P because the following steps consume BQ rapidly.

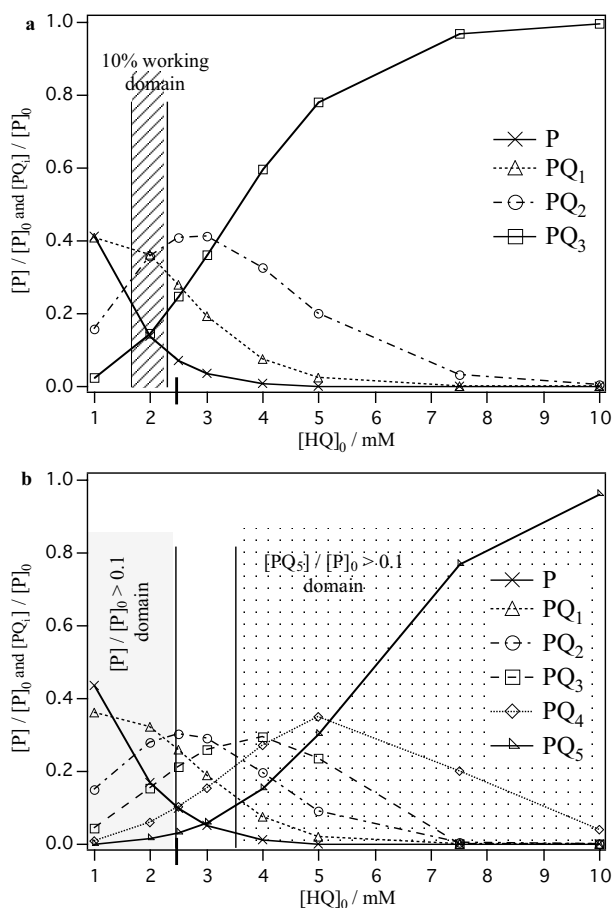


Figure 9. Evolution of species ratio at the end of the channel for several HQ concentrations in case of a three- and a five-cysteine target (a and b respectively). The domain where both $[P] / [P]_0 > 0.1$ and $[PQ_3] / [P]_0 > 0.1$ (10% working domain) is indicated by hatching. A bold line marks the reference concentration.

3.4. Optimization of multi-tagging of peptides

The initial concentration of HQ plays a key role in the final tagging degree since it controls the production of the BQ markers (in the present assumption of diffusion control of the current). In Figure 9a and 9b, the evolution of the species at the end of the channel is given according to $[HQ]_0$ for biomolecules containing three and five cysteines respectively. As expected, the higher the $[HQ]_0$, the higher the consumption of P and the production of the completely tagged species (*i.e.* PQ_3 and PQ_5 for peptides containing three and five cysteine groups respectively). In the previous section, a target with five cysteine units was taken for emphasizing kinetics and better understanding the multi-stage process. A peptide with three cysteine groups serves as reference since such peptides are more probable in proteomic analysis. To enable the MS counting of cysteines in peptides in future experiments, we have decided to impose the following criteria for the simulation: both the proportion of P ($= 100 \cdot [P] / [P]_0$) and of the completely tagged specie PQ_n ($= 100 \cdot [PQ_n] / [P]_0$) should be above 10% in order to be detectable by MS. In Figure 9, the corresponding working domains mark these conditions. For a three-cysteine peptide (Figure 9a) the two conditions are compatible ($1.755 \text{ mM} < [HQ]_0 < 2.275 \text{ mM}$), but the two working domains do not intersect for a five-cysteine peptide (Figure 9b).

The initial concentration of HQ drives the tagging rate and thereby the proportion of species at the end of the channel. When $[HQ]_0$ is too high, there is not enough P left and when it is too low, there is not enough PQ_n produced. As shown in Figure 10, tagging reaction times t in the channel could thus be chosen to make the two conditions expressed above to be compatible. The 10% working conditions are represented in Figure 10a for the three-cysteine peptide. The domain shows that high concentrations of HQ imply working over small times to insure that P is not totally consumed. Nevertheless, a wide range of concentration is then compatible with these short times: 0.6 s of reaction allows working between 10 and 20 mM of initial HQ. When working with a longer reaction time, the concentration possibilities are narrower but the analysis can be done at many times (for instance, the analysis is possible from 2 to more than 5 s for an initial concentration of 2 mM). In other words, when the apparent kinetics become slower, the fixed conditions are satisfied for wider time ranges (the scale is widespread due to the second order kinetic law).

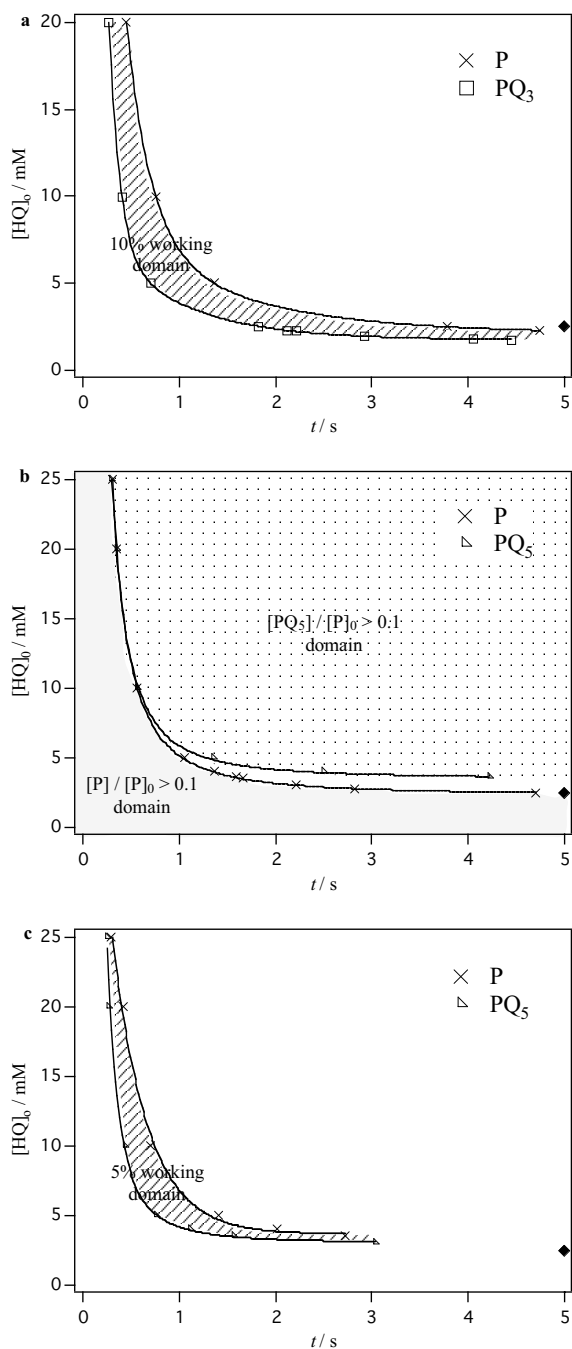


Figure 10. Reaction time within the channel to obtain $100 \cdot [P] / [P]_0 = 10\%$ (\times) and $100 \cdot [PQ_n] / [P]_0 = 10\%$ (\square or \triangle for $n = 3$ and 5 respectively) for a three- and a five-cysteine target (a and b respectively). Time needed to obtain $100 \cdot [P] / [P]_0 = 5\%$ (\times) and $100 \cdot [PQ_5] / [P]_0 = 5\%$ (\triangle) for a five-cysteine-containing target (c). Hatching indicates the domains where both the conditions are satisfied. The reference case is indicated by \blacklozenge .

For the five-stage tagging (Figure 10b), the conditions are almost never compatible except at high concentrations and for very restrictive time ranges (for $[HQ]_0 = 25$ mM, the reaction time must be fixed at 0.3 s). When the criteria level is reduced to 5%, a much wider working domain appears for the five-stage reaction (Figure 10c). The technique appears applicable to peptides containing five or less cysteines since the co-existence of P and PQ_n is no more likely when more cysteines are present as BQ is stocked in the intermediate species.

For successful cysteine counting, the initial concentration of P is also a key feature to consider, all the more so as $[P]_0$ is not really controlled by the manipulator. Indeed, in proteomic analysis, the amounts of tryptic peptides derive from the protein concentrations, which are quite variable from one to another protein of a complex mixture.

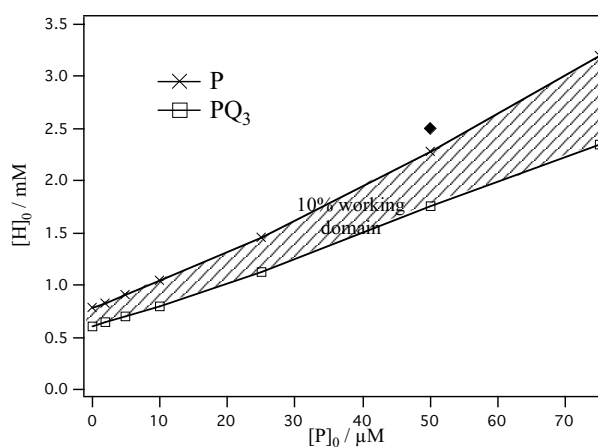


Figure 11. The initial hydroquinone concentration to have $100 \cdot [P] / [P]_0 = 10\%$ (\times) and $100 \cdot [PQ_3] / [P]_0 = 10\%$ (\square) at the end of the channel is represented depending on the initial concentration $[P]_0$. The hatching zone between the curves indicates the 10% working domain where both the conditions are satisfied. The reference case is indicated by \blacklozenge .

In the simulated data of Figure 11, for a given reaction time $t = 5$ s (*i.e.* $L_{ch} = 2$ cm), concentration $[P]_0$ varies from 0.1 to 75 μ M. The initial concentrations $[HQ]_0$ that provide 10% of P and PQ_3 according to each initial concentration $[P]_0$ are reported. Between

the two curves, the hatched domain indicates that both conditions are satisfied. The reference point (◆) shows that $[HQ]_0 = 2.5 \text{ mM}$, which was taken in the previous part, provides good analysis of three-cysteine peptides from 55 to 100 μM . Defining the ratio $\Delta[HQ]_0 / [HQ]_0$ for the 10% working domain, 0.51 and 0.31 are obtained for $[P]_0 = 0.1$ and 75 μM respectively showing that the working interval slightly decreases with the concentration. The nature of the curve proves that the ratio $[HQ]_0 / [P]_0$ can not be used as term to predict the tagging extent in general. This is further confirmed by several numerical studies at different concentrations while keeping the ratio $[HQ]_0 / [P]_0$ constant (Figure 12).

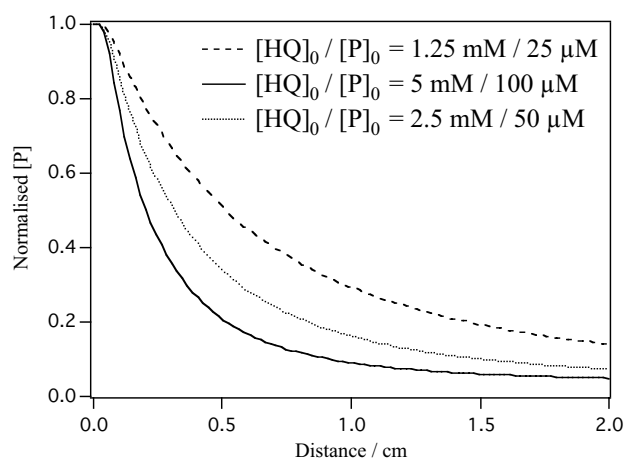


Figure 12. Kinetics of the tagging (evolution of $[P] / [P]_0$) for several initial concentrations of HQ and P in the case of a three-cysteine biomolecule. The ratio $[HQ]_0 / [P]_0$ is kept constant.

4. Conclusions

The influence of the tagging rate constant as well as the impact of the probe and target concentrations has been simulated for a single and a five-step tagging reaction in a microchannel. The finite-element model has shown how strongly the tagging efficiency derives from the number of consecutive tagging reactions (*i.e.* the number of cysteine moieties in the protein or peptide determining the apparent kinetics of each consecutive reaction).

The non-suppression of the unmodified molecule signal is essential to guaranty the success of the on-line MS counting of cysteine moieties. On the other hand, the residence time and the probe concentration should be sufficient to achieve all the tagging degrees. Numerical simulations have been used to determine the optimal conditions for this. It has been shown that the cysteine counting is possible for up to five units. When the molecule possesses more cysteine residues, the co-existence of both the untagged and of the fully tagged molecule is no more likely whatever the residence time. To overcome this, the introduction of a recognition signal could be added to the tag. Some mass tags presenting characteristic isotopic patterns could thus be employed.

Bibliography

- 1 A.E. Ashcroft, *Nat. Prod. Rep.*, 2003, 20, 202-215.
- 2 W.J. Henzel, C. Watanabe and J.T. Stults, *J. Am. Soc. Mass Spectrom.*, 2003, 14, 931-942.
- 3 J.R. Yates, S. Speicher, P.R. Griffin and T. Hunkapiller, *Anal. Biochem.*, 1993, 214, 397-408.
- 4 P. James, M. Quadroni, E. Carafoli and G. Gonnet, *Biochem. Biophys. Res. Commun.*, 1993, 195, 58-64.
- 5 D.J.C. Pappin, P. Hojrup and A.J. Bleasby, *Curr. Biol.*, 1993, 3, 327-332.
- 6 W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley and C. Watanabe, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, 90, 5011-5015.
- 7 M. Mann, P. Hojrup and P. Roepstorff, *Biol. Mass Spectrom.*, 1993, 22, 338-345.
- 8 G.E. Reid and S.A. McLuckey, *J. Mass Spectrom.*, 2002, 37, 663-675.
- 9 D.A. Wolters, M.P. Washburn and J.R. Yates, *Anal. Chem.*, 2001, 73, 5683-5690.
- 10 S.D. Patterson and R. Aebersold, *Electrophoresis*, 1995, 16, 1791-1814.
- 11 M. Mann and M. Wilm, *Anal. Chem.*, 1994, 66, 4390-4399.
- 12 D.R. Goodlett, J.E. Bruce, G.A. Anderson, B. Rist, L. Pasa-Tolic, O. Fiehn, R.D. Smith and R. Aebersold, *Anal. Chem.*, 2000, 72, 1112-1118.
- 13 K.R. Clauser, P. Baker and A.L. Burlingame, *Anal. Chem.*, 1999, 71, 2871-2882.
- 14 B.J. Cargile and J.L. Stephenson, *Anal. Chem.*, 2004, 76, 267-275.
- 15 S. Sechi and B.T. Chait, *Anal. Chem.*, 1998, 70, 5150-5158.
- 16 S. Neitz, M. Jurgens, M. Kellmann, P. Schulz-Knappe and M. Schrader, *Rapid Commun. Mass Spectrom.*, 2001, 15, 1586-1592.
- 17 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 18 L. Dayon, C. Roussel and H.H. Girault, *Chimia*, 2004, 58, 204-207.
- 19 L. Dayon, C. Roussel, M. Prudent, N. Lion and H.H. Girault, *Electrophoresis*, 2005, 26, 238-247.
- 20 M.J. Clifford, S.M. Cox and E.P.L. Roberts, *Chem. Eng. J.*, 1998, 71, 49-56.
- 21 S.M. Cox, M.J. Clifford and E.P.L. Roberts, *Physica A*, 1998, 256, 65-86.

- 22 K. Harriman, D.J. Gavaghan, P. Houston, D. Kay and E. Suli, *Electrochem. Commun.*, 2000, 2, 576-585.
- 23 J.A. Alden and R.G. Compton, *J. Phys. Chem. B*, 1997, 101, 9741-9750.
- 24 T.C. Rohner, J. Josserand, H. Jensen and H.H. Girault, *Anal. Chem.*, 2003, 75, 2065-2074.
- 25 V. Mengeaud, J. Josserand and H.H. Girault, *Anal. Chem.*, 2002, 74, 4279-4286.
- 26 V.G. Levich, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, 1962, pp. 112-116.
- 27 P.R. Unwin and R.G. Compton, *Comprehensive Chemical Kinetics*, Vol. 29, Eds.: R. G. Compton and A. Hammet, Elsevier, Amsterdam, 1989, pp. 173-193.
- 28 C. Roussel, L. Dayon, H. Jensen and H.H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.
- 29 C. Roussel, T.C. Rohner, H. Jensen and H.H. Girault, *ChemPhysChem*, 2003, 4, 200-206.
- 30 L. Dayon, C. Roussel and H.H. Girault, *J. Proteome Res.*, 2006, 5, 793-800.
- 31 S. Metsamuuronen, S.P. Reinikainen and M. Nystrom, *Desalination*, 2002, 149, 453-458.

Annexe: finite-element formulation

The general Equation 1 is treated for a maximum five successive tagging ($n = 5$). It takes into account the diffusion-convection-reaction equations for the eight considered species. The order of writing of the $c_i c_j$ is chosen to maximize the diagonal part of the matrix of unknowns shown at the end of the paragraph ($c_i c_j$ being in the column j and $c_j c_i$ being in the column i).

$$\frac{\partial c_{\text{HQ}}}{\partial t} + \nabla \cdot (-D_{\text{Q}} \nabla c_{\text{HQ}} + \mathbf{v} c_{\text{HQ}}) = -k_{\text{ox}} c_{\text{HQ}} + k_{\text{red}} c_{\text{BQ}} \quad (\text{Equation 1A})$$

$$\begin{aligned} \frac{\partial c_{\text{BQ}}}{\partial t} + \nabla \cdot (-D_{\text{Q}} \nabla c_{\text{BQ}} + \mathbf{v} c_{\text{BQ}}) = & +k_{\text{ox}} c_{\text{HQ}} - k_{\text{red}} c_{\text{BQ}} - k_1 c_{\text{P}} c_{\text{BQ}} - k_2 c_{\text{PQ}_1} c_{\text{BQ}} \\ & - k_3 c_{\text{PQ}_2} c_{\text{BQ}} - k_4 c_{\text{PQ}_3} c_{\text{BQ}} - k_5 c_{\text{PQ}_4} c_{\text{BQ}} \end{aligned} \quad (\text{Equation 2A})$$

$$\frac{\partial c_{\text{P}}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{P}} + \mathbf{v} c_{\text{P}}) = -k_1 c_{\text{BQ}} c_{\text{P}} \quad (\text{Equation 3A})$$

$$\frac{\partial c_{\text{PQ}_1}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{PQ}_1} + \mathbf{v} c_{\text{PQ}_1}) = +k_1 c_{\text{BQ}} c_{\text{P}} - k_2 c_{\text{BQ}} c_{\text{PQ}_1} \quad (\text{Equation 4A})$$

$$\frac{\partial c_{\text{PQ}_2}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{PQ}_2} + \mathbf{v} c_{\text{PQ}_2}) = +k_2 c_{\text{BQ}} c_{\text{PQ}_1} - k_3 c_{\text{BQ}} c_{\text{PQ}_2} \quad (\text{Equation 5A})$$

$$\frac{\partial c_{\text{PQ}_3}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{PQ}_3} + \mathbf{v} c_{\text{PQ}_3}) = +k_3 c_{\text{BQ}} c_{\text{PQ}_2} - k_4 c_{\text{BQ}} c_{\text{PQ}_3} \quad (\text{Equation 6A})$$

$$\frac{\partial c_{\text{PQ}_4}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{PQ}_4} + \mathbf{v} c_{\text{PQ}_4}) = +k_4 c_{\text{BQ}} c_{\text{PQ}_3} - k_5 c_{\text{BQ}} c_{\text{PQ}_4} \quad (\text{Equation 7A})$$

$$\frac{\partial c_{\text{PQ}_5}}{\partial t} + \nabla \cdot (-D_{\text{P}} \nabla c_{\text{PQ}_5} + \mathbf{v} c_{\text{PQ}_5}) = +k_5 c_{\text{BQ}} c_{\text{PQ}_4} \quad (\text{Equation 8A})$$

where $k_{\text{ox, red}}$ are defined on the electrode surface only, and k_{1-5} is defined over all the domain. To take into account the number of free cysteines without introducing the cysteine concentration $C_{m=1\dots n, n \leq 5} = (n-m) [\text{PQ}_m]$ (where m is the adduct rating and n is the number of possible adducts), the rate constant is written as $k_{m=1\dots n, n \leq 5} = (n+1-m) k$. These expressions are derived in the global general form (Equations 9A and 11A), using the Galerkin's

formulation (multiplication by a projective function α and integration on the domain of study Ω).

$$\iint_{\Omega} \alpha \left[\frac{\partial c_i}{\partial t} + \nabla \cdot (-D_i \nabla c_i + \mathbf{v} c_i) - R_i \right] d\Omega = 0 \quad (\text{Equation 9A})$$

The convection term is derived taking into account the continuity equation $\nabla \cdot \mathbf{v} = 0$. Decomposing the product between α and the divergence, the second order derivative of (Equation 9A) (divergence of the gradient...) becomes:

$$\alpha \nabla \cdot (-D_i \nabla c_i) = \nabla \cdot (-\alpha D_i \nabla c_i) + D_i \nabla \alpha \cdot \nabla c_i \quad (\text{Equation 10A})$$

Applying Equation 10A in Equation 9A and using the Ostrogradsky theorem, the divergence term is rejected at the boundary (Equation 11A), where it expresses the diffusion flux conditions of the specie i . In the present case of study, this boundary condition equals to zero (no diffusion flux at the boundaries of the domain).

$$\iint_{\Omega} \left[\alpha \frac{\partial c_i}{\partial t} + D_i \nabla \alpha \cdot \nabla c_i + \alpha \mathbf{v} \cdot \nabla c_i - \alpha R_i \right] d\Omega = \int_{\partial\Omega} \alpha D_i \frac{\partial c_i}{\partial n} dl = 0 \quad (\text{Equation 11A})$$

Equation 11A is applied to the present problem (Equations 1-8A) and is expressed in the local matrix form (matrix of unknowns, where c_{1-8} are respectively the concentrations of HQ, BQ, P, PQ_{1,2,3,4,5}):

$$\left[L_i \right] \left[\frac{\partial c_i}{\partial t} \right] + \left[M_{ij} \right] \left[c_i \right] = 0 \quad (\text{Equation 12A})$$

where L_i is the transient term (not used in the present calculations). The M_{ij} steady state matrix is described below. All the gradients are written in a nabla form ∇ . The function β is the interpolation function of the unknown. The position in the matrix of the reaction terms is chosen to maximize the weight of the diagonal.

HQ	BQ	P	PQ ₁	PQ ₂	PQ ₃	PQ ₄	PQ ₅
$D_Q \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_{ox} \beta$	$-\alpha k_{red} \beta$						
$-\alpha k_{ox} \beta$	$D_Q \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_{red} \beta$ $+\alpha k_1 c_P \beta$ $+\alpha k_2 c_{PQ_1} \beta$ $+\alpha k_3 c_{PQ_2} \beta$ $+\alpha k_4 c_{PQ_3} \beta$ $+\alpha k_5 c_{PQ_4} \beta$						
		$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_1 c_{BQ} \beta$					
		$-\alpha k_1 c_{BQ} \beta$	$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_2 c_{BQ} \beta$				
			$-\alpha k_2 c_{BQ} \beta$	$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_3 c_{BQ} \beta$			
				$-\alpha k_3 c_{BQ} \beta$	$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_4 c_{BQ} \beta$		
					$-\alpha k_4 c_{BQ} \beta$	$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$ $+\alpha k_5 c_{BQ} \beta$	
						$-\alpha k_5 c_{BQ} \beta$	$D_P \nabla \alpha \nabla \beta$ $+\alpha v \nabla \beta$

CHAPTER VII. *Probing cysteine reactivity in proteins by mass spectrometric EC-tagging*^ξ

1. Introduction

In proteins, cysteine residues are important for metal coordination, catalysis and protein structure by forming disulfide bonds. Moreover, crucial cysteine residues are involved in modulation of protein activity and signalling events via redox reactions, chelation of transition metals and *S*-nitrosation. Cysteine is also the binding site in human albumin for biological and clinical small molecules such as platinum(II) anticancer drugs.^{1,2} The reactivity of cysteine in proteins is complex for reasons such as steric hindrance, charge distribution and solvation. It varies from one protein to another and specific competition with glutathione makes the system even more complex in biological fluids. The antioxidant character of cysteinyl proteins depends highly on the protein structure that for instance prevent the formation of disulfide bonds in albumin or reduce the antioxidant capacity of hemoglobin in comparison to glutathione.^{3,4}

As nucleophiles, thiols have a reaction rate that depends on the protonation state of the sulfhydryl group. The primary structure of the biomolecule influences the thiol reactivity since the pK_a of the thiol is strongly dependent on the charged residues in the vicinity of the cysteine. It has been shown that Michael-type addition of sulfhydryl-containing peptides onto unsaturated groups have higher or lower rates when positive charges and negative charges are respectively in the vicinity, resulting to a decrease or an increase of pK_a .⁵

^ξ based on L. Dayon, C. Roussel and H.H. Girault, J. Proteome Res., 2006, 5, 793-800.

In electrospray ionization (ESI) mass spectrometry (MS), the application of an electric field to generate the spray has led the consideration of emitters as on-line electrochemical flow-cells. Then, the inherent electrochemical aspect of electrospray⁶⁻⁸ has opened the way to the study of electrochemically induced reactions⁹⁻¹² like rearrangements of biological molecules.¹³⁻¹⁵ Moreover, electro-active probes have been developed to chemically derivatize weakly ionisable compounds¹⁶⁻²¹, and to label specific amino acids.²² Recently, we have developed a polymer micro-ESI emitter comprised of a microband electrode.²³ This micro-flow-cell was shown to be an efficient controlled-current electrochemical flow-cell compared to many commercially available ESI sources.²⁴ Additionally, the upstream position of the microband electrode within the micromachined flow channel is a major advantage in the electrogeneration of tags to bind to molecules flowing above the electrode toward the Taylor cone. The oxidation of hydroquinone derivatives on the microband anode was studied to tag cysteine residues in peptides *via* a selective 1,4-Michael addition. When controlling both the electrode mass transport and the kinetics of the addition reaction, the application of on-line counting of cysteines in peptides to the identification of proteins by peptide mass mapping was achieved.^{25,26}

In this chapter, a method to probe cysteinyl sites in proteins is described. An analytical kinetic model is developed to predict tagging extents at the end of the microchannel prior to the Taylor cone. We show that the MS measurements of the extent of the EC-tagging reaction depend mainly on the reactivity of cysteine residues in the proteins and not on the ionization properties of adducts. The EC-tagging of β -lactoglobulin A (one free cysteine unit) was studied and compared with the multi-tagging of its reduced form (i.e. five free cysteine units after reduction of the two disulfide bonds). Creatine phosphokinase and reduced insulin were probed by EC-tagging to complete the study.

2. Experimental

2.1. Maple[®] calculation

The kinetic model is based on a set of differential equations that can be solved analytically with commercial software (Maple[®] 9.5, Waterloo Maple Inc.) on an iMac G5 (2GHz PowerPC, 512MB DDR400 SDRAM from Apple, Cupertino, CA, USA) (see annexe).

2.2. Materials

Insulin from bovine pancreas, creatine phosphokinase (CK) from rabbit muscle and β -lactoglobulin A from bovine milk (> 90%) were purchased from Sigma (St Louis, MO, USA). Tri-*n*-butylphosphine (TBP, 97%) and methoxycarbonyl-1,4-hydroquinone (methyl 2,5-dihydroxybenzoate) (99%) were from Aldrich (Milwaukee, WI, USA). 1,4-Hydroquinone (> 98%), 1,4-benzoquinone (\geq 98%), L-cysteine (> 99.5%), lithium trifluoromethanesulfonate (purum), *N,N*-dimethylformamide (DMF, \sim 99%), acetic acid (AcOH, 99.5%) were from Fluka (Büchs, Switzerland). Methanol (MeOH, > 99.8%) was bought from Riedel-de Haën (Seelze, Germany). Acetonitrile (HPLC grade) was from Sds (Peypin, France). Deionised water (18.2 M Ω -cm) was prepared using a Milli-Q system from Millipore (Bedford, MA, USA). Methoxycarbonyl-1,4-benzoquinone was synthesized as already described.²⁷

2.3. Reduction of proteins

To reduce the disulfide bridges, 1 mg of purified protein was dissolved in 900 μ L of H₂O. 100 μ L of TBP at 10% in DMF ($4 \cdot 10^{-5}$ mol) was added to the protein solution according to the literature.²⁸ The mixture was stirred between 60 and 90 min. Then, the resulting mixture was lyophilized overnight to get rid of the excess reagent (TBP)²⁹ that can react with benzoquinone compounds.³⁰ Without lyophilization, the EC-tagging of the protein was indeed not obtained. The proteins were redissolved in degassed MeOH / H₂O / AcOH 50% / 49% / 1%.

2.4. EC-tagging

The proteins at 50 μ M were sprayed with methoxycarbonyl-1,4-hydroquinone or 1,4-hydroquinone 2.5 mM in degassed MeOH / H₂O / AcOH 50% / 49% / 1%. The total

acquisition time of the spectra was 1 min and the EC-tagging was checked to be stable during 25 min. A few minutes were needed to reach the steady state of the process. The apparent tagging extents (α -TE%) and the MS tagging yield ($Yield_{MS}$) were calculated as 100 fold the sum of the peak intensity of the product(s) over the sum of the peak intensity of the product(s) and the starting protein.

2.5. Chemical labelling and ionization experiments

The chemical formation of adducts was performed by adding benzoquinone BQ as a solid or in acetonitrile solution (from 250 to 2500 μ M). The ratio BQ / protein or cysteine was varied from 0.05 to 0.7. The mixtures were analysed by MS after 30-45 min of reaction. The reaction is rapid and is assumed to be total^{27,31} to calculate the yields $Yield_{MS}$ as 100 fold the sum of the peak intensity of the product(s) over the sum of the peak intensity of the product(s) and the starting protein.

2.6. Micro-fabricated ESI emitter

The polymer microspray fabrication has been already described in detail (see also appendix I for comprehensive description).²³ The surface of the electrode integrated at the bottom of the channel was $70 \times 25 \mu\text{m}^2$ (the length L was 70 μm , and the width l was 25 μm). The channel length from the electrode to the tip was $L_{\text{ch}} = 2$ cm and the cross-section was $30 \times 35 \mu\text{m}^2$ (the width d was 30 μm and the height $2h$ was 35 μm).

2.7. MS set-up

A LCQ DUO ion trap mass spectrometer (Finnigan, San José, CA, USA) was used. The heated capillary was kept at 200 °C. In each experiment, the ion transmission parameters were optimized automatically in order to improve the detection of the analyte of interest (the unmodified biomolecule). The ESI interface was removed and the microchip holder was mounted on the probe slide adapter of the mass spectrometer. The device was coupled to a syringe pump (kdScientific, Holliston, MA, USA) to introduce the solution. The flow rate F_v was set to 250 $\text{nL}\cdot\text{min}^{-1}$ and the voltage applied was 3-4 kV. The distance from the outlet of the microchip to the entrance of the spectrometer varied between 1 and 2 cm in order to set the electrospray current and to optimize the signal and the trap injection time. The applied spray current I_{spray} was set at 120 nA.

2.8. Diffusion coefficient measurements

The electrochemical measurements were recorded on an Autolab PGSTAT 12 potentiostat from Metrohm (Herisau, Switzerland) using an undivided cell filled with a protected saturated calomel electrode as reference electrode, a glassy carbon electrode (3 mm diameter) as working electrode and a platinum wire as counter electrode. The electrochemical cell was filled with the spray solution (MeOH / H₂O / AcOH 50% / 49% / 1%) at 0.1 M lithium trifluoromethanesulfonate and 2 mM hydroquinone. The working electrode was carefully polished with a suspension of 0.3 μm alumina (from Buehler, Lake Bluff, IL, USA) before every experiment. Diffusion coefficients of 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone were measured by chronoamperometry. The potential was set for 60 s at 0.8 V and at 1 V for 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone respectively and the oxidation current was recorded (every 0.5 s). Diffusion coefficients were extracted from the slope of the straight line $I(t) = f(t^{0.5})$ and were found to be $4 \cdot 10^{-10}$ and $3.5 \cdot 10^{-10}$ m²·s⁻¹ for 1,4-hydroquinone and methoxycarbonyl-1,4-hydroquinone. This difference can be attributed to the strong hydration of 1,4-hydroquinone in water compared to that of methoxycarbonyl-1,4-hydroquinone.

3. Results and Discussions

3.1. Mechanism and kinetics

Figure 1 shows a schematic representation of a microchip electrospray emitter comprising basically a flow channel and a band electrode located at the bottom of the microchip. During electrospray from the microchip at a flow rate F_V of 250 nL·min⁻¹, benzoquinone tags (BQ) are generated at a microband electrode by oxidation of hydroquinone compounds (HQ) and react specifically with the thiol groups of the proteins (P) in the flow channel. The products of the reaction at the end of the channel are analyzed continuously by ion trap mass spectrometry. The addition of BQ tags on thiols has a very large equilibrium constant.^{27,31}

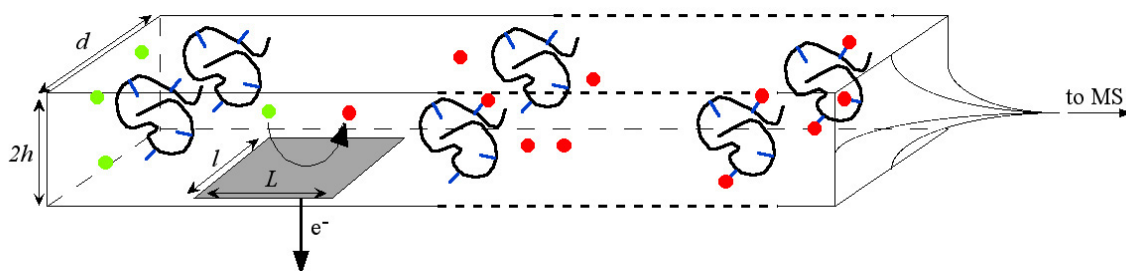
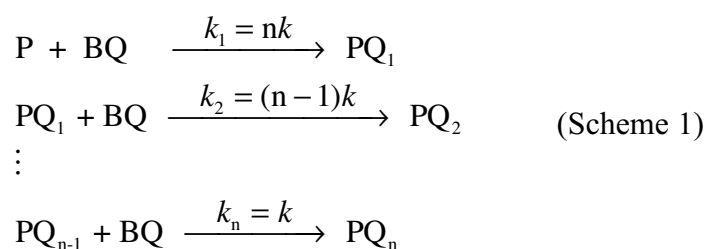


Figure 1. Schematic representation of the EC-tagging of cysteinyl proteins (P) by benzoquinone (BQ, ●) tags electrogenerated at the microspray electrode from hydroquinone (HQ, ●).

According to the number of cysteines n in P, adducts (PQ_i) are successively formed ($i = 1 \dots n$). The mechanism for multi-cysteine-containing protein can be described by



where PQ_i are representative of the tagged state but not of a defined molecule (*i.e.* for multi-cysteine-containing protein, PQ_1 represents several regio-isomers). It is assumed that every thiol group presents equal reactivity.

3.1.1. Analytical kinetic model

- *Electrochemical step.* In the positive ionization mode, the reaction taking place at the high voltage microband electrode³² as illustrated in Figure 1 is the oxidation of HQ considering its redox potential.²⁷ In the flow-cell microspray emitter, the convection-diffusion limiting current I_{lim} in a 2D laminar Poiseuille flow can be classically calculated by³³⁻³⁵

$$I_{lim} = 0.925zFCI(LD)^{2/3} \left(\frac{F_V}{h^2d} \right)^{1/3} \quad (\text{Equation 1})$$

where z is the number of electrons per oxidized molecule, F is the Faraday constant, C and D are respectively the bulk concentration and the diffusion coefficient of the electroactive species, l and L the width and the length of the electrode, F_V the pressure-driven flow rate, $2h$ and d the height and the width of the channel.

The electro spray current I_{spray} is set by controlling the flow rate, the distance between the chip tip and the entrance of the MS, and the high voltage applied by the MS⁸ (see Experimental). When the imposed electro spray current I_{spray} value is greater than the calculated HQ oxidation convection-diffusion limiting current I_{lim} , we can use Faraday's law to calculate the quantity of BQ tags produced at the electrode according to

$$n_{\text{BQ}} = \frac{I_{\text{lim}} t_e}{zF} \quad (\text{Equation 2})$$

where $t_e = L/\bar{v}$ is the residence time of species on the top of the electrode and \bar{v} is the mean flow velocity.

General parameters	1,4-Hydroquinone	Methoxycarbonyl-1,4-hydroquinone
$F_V = 250 \text{ nL}\cdot\text{min}^{-1}$	$D = 4 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$ ^b	$D = 3.5 \cdot 10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$ ^b
$\bar{v} = 4.3 \text{ mm}\cdot\text{s}^{-1}$	$I_{\text{lim}} = 81 \text{ nA}$ ^c	$I_{\text{lim}} = 74 \text{ nA}$ ^c
$[\text{HQ}]_0 = 2.5 \text{ mM}$	$[\text{BQ}]_0 = 101 \text{ }\mu\text{M}$ ^d	$[\text{BQ}]_0 = 93 \text{ }\mu\text{M}$ ^d

^a The real cross-section geometry of the channel is in fact trapezoidal.³² To bring this correction to the model, the width of the channel d was taken as $d_m = (d + l) / 2$. The reaction time t in the full channel is then 4.7 s.

^b see Experimental.

^c according to Equation 1.

^d according to Equation 2.

Table 1. Parameters^a for calculation of species concentrations in the channel (*N.B.* $I_{\text{spray}} = 120 \text{ nA}$).

Conversely, when the imposed electro spray current I_{spray} value is lower than the calculated HQ oxidation current I_{lim} , the quantity of BQ tags produced at the electrode is

$$n_{\text{BQ}} = \frac{I_{\text{spray}} t_e}{zF} \quad (\text{Equation 3})$$

Considering a fast transversal diffusion above the electrode, the initial concentration of BQ tags ($[\text{BQ}]_0$) can be assimilated to its mean concentration in the volume element over the electrode ($V_e = L \times d \times 2h$) after a time t_e (Table 1 and Figure 1). Post-electrode oxidation of HQ by the other species electrogenerated at the electrode such as oxygen is neglected.

- *Chemical step.* For a simplified analytical calculation of the extent of tagging in the microchip set-up, complete mixing is assumed to be ideal in the finite volume V_e . The reaction occurs in this volume translating with the flow along the microchannel. In the case of the simplest addition between BQ tags and a protein P containing one cysteine the rate law follows a first order kinetics for each reactant:^{36,37}

$$v = -\frac{d[\text{BQ}]}{dt} = -\frac{d[\text{P}]}{dt} = \frac{d[\text{PQ}_1]}{dt} = k[\text{BQ}][\text{P}] \quad (\text{Equation 4})$$

v is the rate of the reaction, k is the rate constant and $[\text{BQ}]$, $[\text{P}]$ and $[\text{PQ}_1]$ represent respectively the concentration of the electrogenerated BQ tags, the concentration of a protein P containing one cysteine residue and the concentration of the single-cysteine-containing product PQ_1 at the time t . The integrated rate law gives

$$\frac{1}{[\text{BQ}]_0 - [\text{P}]_0} \ln \left[\frac{[\text{P}]_0([\text{BQ}]_0 - x)}{[\text{BQ}]_0([\text{P}]_0 - x)} \right] = kt \quad (\text{Equation 5})$$

where $[\text{BQ}]_0$ and $[\text{P}]_0$ are the initial concentrations of the electrogenerated BQ tags (*i.e.* at the electrode) and of the single-cysteine-containing protein, and $x = [\text{PQ}_1] = [\text{BQ}]_0 - [\text{BQ}] = [\text{P}]_0 - [\text{P}]$.

This calculation can be applied to the consecutive stages when the protein possesses several cysteine units (general mechanism written above). The use of the Maple[®] software is then convenient to solve the differential equation systems, as the analytical solutions are too cumbersome to derive manually (see the program for a five-step chemical reaction, *i.e.* five-cysteine-containing protein, in the annexe).

3.1.2. Numerical validation of the model

We have developed in the previous chapter a finite-element model treating the emitter as a flow chemical reactor to study the consecutive reactions in the microchannel and to optimize the conditions for cysteine counting in peptides.³⁸ The comparison of the concentrations [BQ], [P] and [PQ_n] at the end of the channel from the analytical kinetic model shows a very good correlation with the simulated data, validating the present analytical kinetic model as illustrated in Figure 2 for methoxycarbonyl-1,4-hydroquinone.

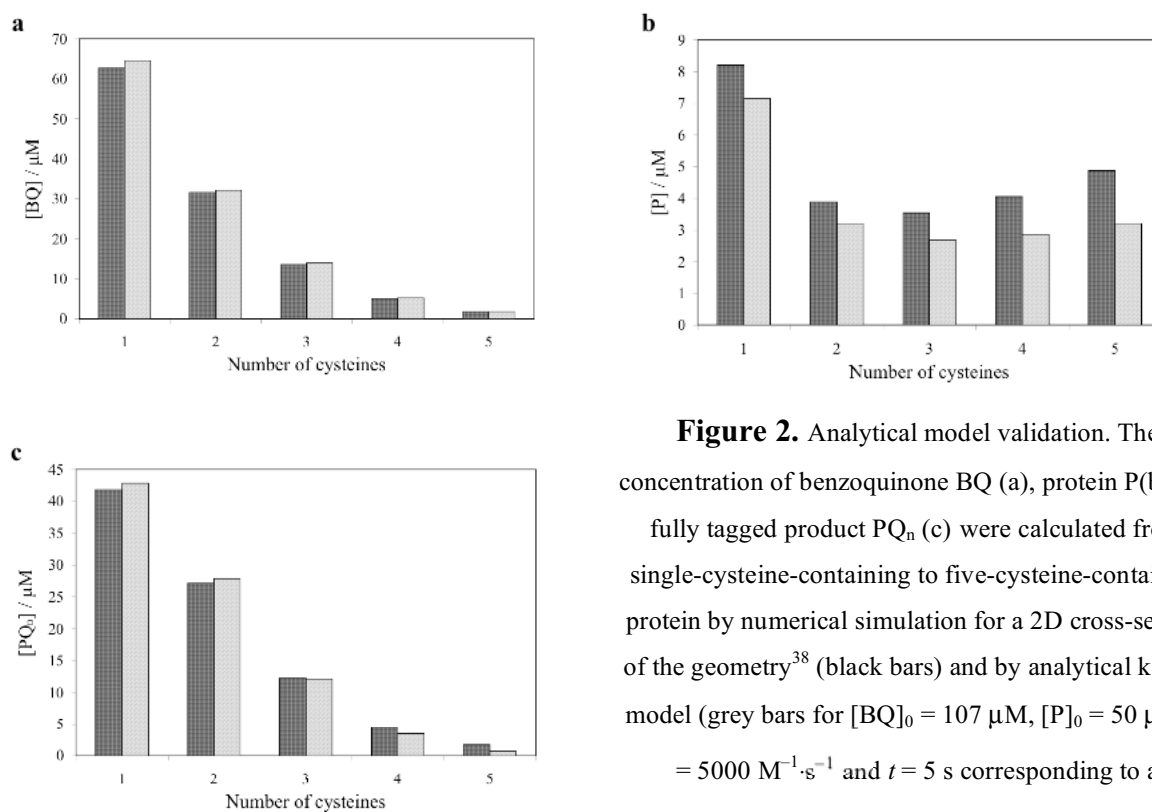


Figure 2. Analytical model validation. The concentration of benzoquinone BQ (a), protein P(b) and fully tagged product PQ_n (c) were calculated from single-cysteine-containing to five-cysteine-containing protein by numerical simulation for a 2D cross-section of the geometry³⁸ (black bars) and by analytical kinetic model (grey bars for [BQ]₀ = 107 μM , [P]₀ = 50 μM , k = 5000 $\text{M}^{-1}\cdot\text{s}^{-1}$ and t = 5 s corresponding to a rectangular geometry with $l = d = 30 \mu\text{m}$).

3.1.3. Experimental validation of the model

Tagging of L-cysteine ($[P]_0 = 200 \mu\text{M}$) by methoxycarbonyl-1,4-hydroquinone ($[\text{HQ}]_0 = 20 \text{ mM}$) has been previously investigated and the Michael addition rate constant has been determined to be very large in the spray medium MeOH / H₂O / AcOH 50% / 49% / 1%.^{27,37} The apparent EC-tagging extent, $\alpha\text{-TE}\%$ was 99%. To take into account MS ionization phenomena induced by the tag^{10,39}, a calibration curve was determined giving the MS response as a function of the true mixture composition (Figure 3).

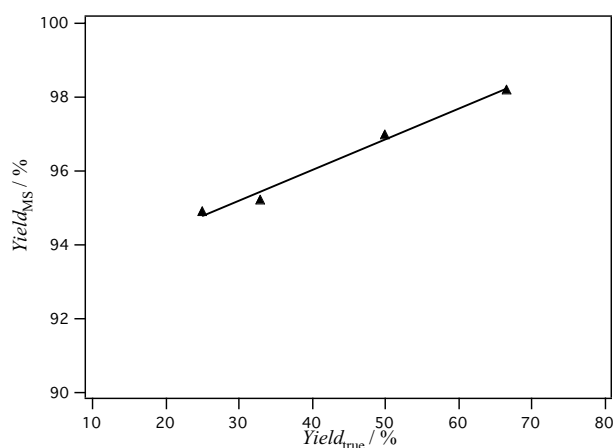


Figure 3. Calibration curves of the MS tagging yield ($Yield_{\text{MS}}$) as a function of the bulk yield ($Yield_{\text{true}}$) for the reaction of L-cysteine with methoxycarbonyl-1,4-benzoquinone. The mixtures were prepared by adding BQ equivalent (from 0.25 to 0.7) to 200 μM of L-cysteine. The fitting equation is $Y = 92.68 + 0.083514 X$.

The effective $\alpha\text{-TE}\%$ (*i.e.* in bulk) extracted from the curve is 76%. The analytical kinetic model predicts a 67% $\text{TE}\%$ value for tagging with methoxycarbonyl-1,4-hydroquinone, which correlates rather well with experimental $\alpha\text{-TE}\%$ of 76%. These data validate the use of the analytical kinetic model to predict the tagging efficiency.

Hydroquinone compound	k_{cys} ($\text{M}^{-1} \cdot \text{s}^{-1}$) ^a	Tagging extent (%)		
		Analytical $TE\%$ ^b	Apparent $a\text{-}TE\%$ ^c	Effective $e\text{-}TE\%$ ^c
Methoxycarbonyl-1,4-hydroquinone	5000	79	31	26
1,4-Hydroquinone	210	9	25	5

^a rate constant of Michael addition of L-cysteine in MeOH / H₂O / AcOH 50% / 49% / 1%.^{27,37}

^b extents were calculated thanks to Equation 5 with initial concentration of Table 1, a time residence $t = 4.7$ s and rate constants relative to L-cysteine in MeOH / H₂O / AcOH 50% / 49% / 1%.

^c experimental data were obtained with β -lactoglobulin A, which possesses one free cysteine (Figure 6).

Table 2. Tagging extent obtained by calculation and MS experiments for a single-cysteine-containing biomolecule. The addition reaction time in the microchannel is 4.7 s.

3.2. EC-tagging of proteins

The analytical kinetic model for EC-tagging has been applied to the study of β -lactoglobulin A tagged by methoxycarbonyl-1,4-benzoquinone. This tag has been shown to be efficient because of its reactivity and the stability of its associated hydroquinone form, even in the presence of oxygen.²⁷ It is therefore an ideal probe to tag cysteine in proteins and the physical parameters relative to the tag are given in Table 1. The tagging extent $TE\%$ (Table 2) predicted by the kinetic model for a singly cysteinyl biomolecule is calculated using the rate constant measured for the tagging of L-cysteine ($TE\% = 100 \cdot [PQ_1] / ([P] + [PQ_1]) = 100 \cdot [PQ_1] / [P]_0$).

From an experimental point of view, short residence times in the microchannel are required to avoid the reaction running to completion, so that simultaneous MS detection of both the untagged and tagged proteins can be obtained with a mass shift corresponding to the mass of the benzoquinone tags.

The apparent $a\text{-}TE\%$ (*i.e.* the consumption of P) for the EC-tagging of β -lactoglobulin A (one free cysteine) by methoxycarbonyl-1,4-hydroquinone at a concentration ratio of 50 μM / 2.5 mM was found to be 31%. The mass spectrum obtained is given in Figure 4.

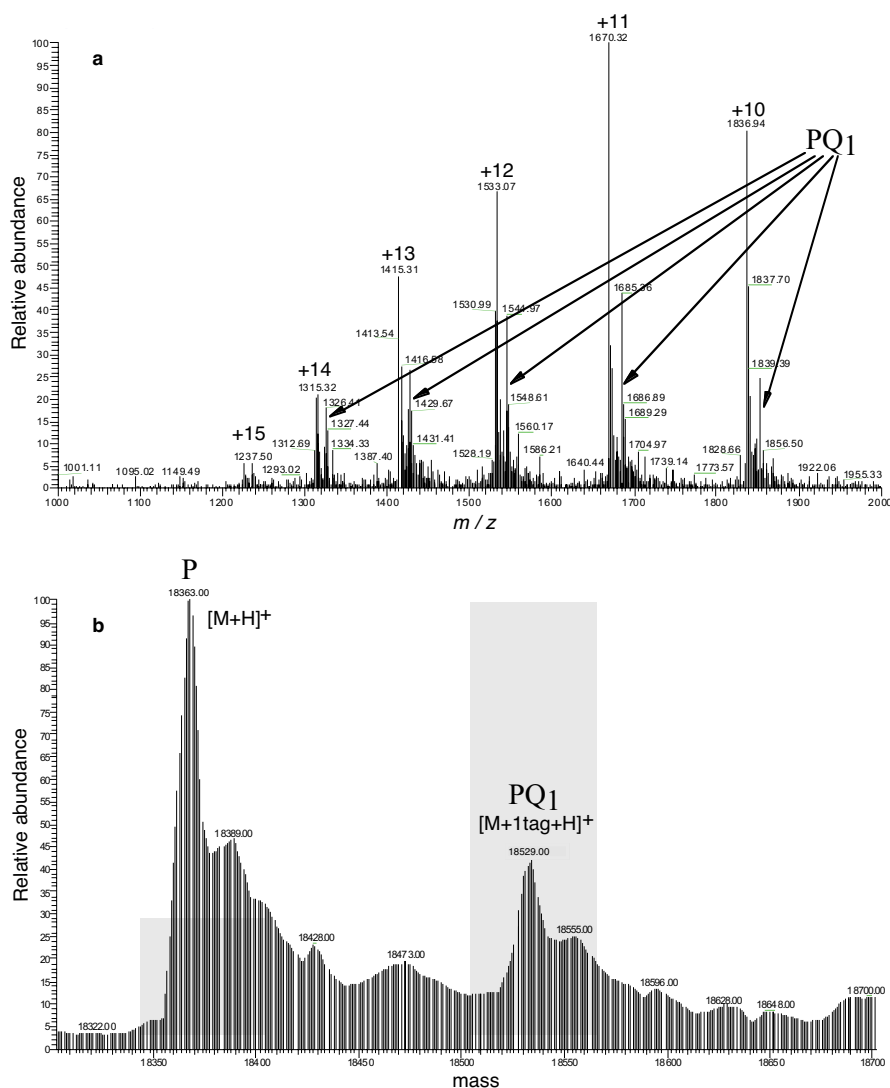


Figure 4. Microspray mass spectrum of β -lactoglobulin A (50 μ M) infused with methoxycarbonyl-1,4-hydroquinone (2.5 mM) (a). Deconvolution mass spectrum (b). Grey bars give the predicted distribution of the species according to the analytical model. P and PQ₁ indicate respectively the untagged and tagged proteins.

In comparison with the analytical model (Table 2), methoxycarbonyl-1,4-hydroquinone provides a much lower conversion than expected ($31 \pm 10\%$ a-TE% instead of 79% by calculation). This decrease in the reactivity of the addition reaction compared to that of the simple amino acid can be explained by several factors including difference in ionizations, variation of the local pK_a in proteins and steric hindrance.⁴⁰

3.2.1. Ionization effects

The tag can induce a difference in MS ionization efficiency. However, the mass spectrum of an equimolar mixture of β -lactoglobulin A and its tagged analogue shows a signal enhancement for the tagged compound over that of the untagged protein (see Figure 5). This effect is similar to that observed in Figure 3 for L-cysteine.

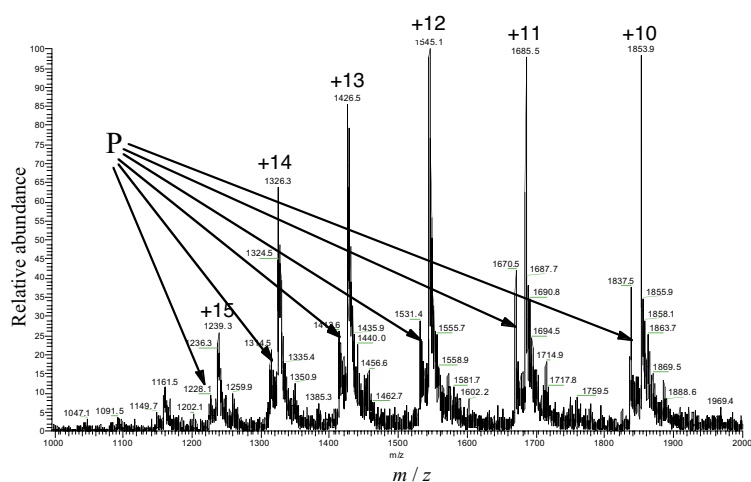


Figure 5. Microspray mass spectrum of a solution containing 50 μM β -lactoglobulin A and 25 μM of methoxycarbonyl-1,4-benzoquinone after 30 minutes of reaction in the ESI medium. The main observed distribution (with indicated charge states) corresponds to the free-thiol tagged β -lactoglobulin A (PQ_1). The non-shifted distribution ($m/z = 1413.6, 1531.4, 1670.5, 1837.5\dots$) from the untagged β -lactoglobulin A (P) represents around 30% of that of the tagged β -lactoglobulin A.

Calibration curves of the MS response as a function of a known mixture of β -lactoglobulin A and its associate adduct were carried out (Figure 6) and an effective $e\text{-TE}\%$ of 26% was determined. Therefore, the observed decrease of the tagging efficiency for β -lactoglobulin A cannot be attributed to a difference in ionization.

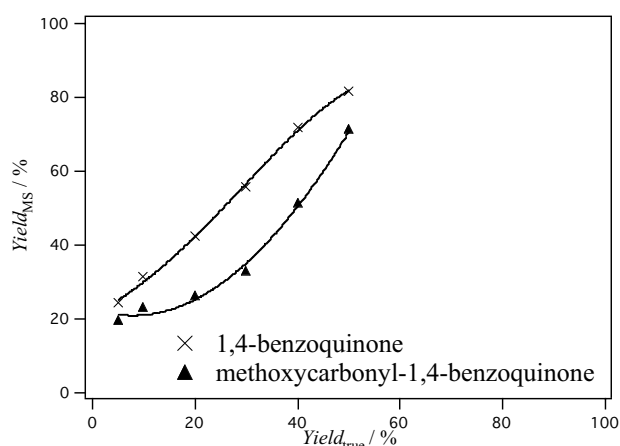


Figure 6. Calibration curves for the MS tagging yield ($Yield_{MS}$) as a function of the bulk yield ($Yield_{true}$) for the reaction of β -lactoglobulin A with methoxycarbonyl-1,4-benzoquinone and 1,4-benzoquinone. The mixtures were prepared by adding BQ equivalent (from 0.05 to 0.5) to 50 μ M of β -lactoglobulin A. The fitting of the curves gives $Y = 613.57 + 592.61 \exp(-((X - 7.8623) / (142.26))^2)$ and $Y = 8.917 + 77.954 \exp(-((X - 61.596) / (45.194))^2)$ for methoxycarbonyl-1,4-benzoquinone and 1,4-benzoquinone respectively.

3.2.2. Variation of pK_a and steric hindrance

The thiol pK_a is influenced by the peptidic chain primary structure.⁴¹ At a given pH, a decrease or an increase of the pK_a lead respectively to an increase or a decrease in the observed addition rate according to the Bronsted equation:

$$k = (10^{pK_a - pH} + 1)^{-1} k_{true} \quad (\text{Equation 6})$$

where k_{true} is the rate constant when the thiol is fully deprotonated.^{5,42}

According to Equation 5, the e-TE% of 26% yields an observed rate constant $k = k_p = 760 \text{ M}^{-1} \cdot \text{s}^{-1}$. Having evaluated the addition rate constant, Equation 6 can then be used to determine the local pK_{ap} compared to that of L-cysteine. Indeed, we can write assuming that k_{true} is the same for both the amino acid and the protein:

$$\frac{k_p}{k_{cys}} = \frac{10^{pK_{acys} - pH} + 1}{10^{pK_{ap} - pH} + 1} \quad (\text{Equation 7})$$

where k_{cys} is the addition rate constant determined with L-cysteine (Table 2) and $\text{p}K_{a_{\text{cys}}}$ is the thiol $\text{p}K_a$ of L-cysteine. This equation yields a local $\text{p}K_{a_{\text{p}}}$ of 8.8 when taking $\text{p}K_{a_{\text{cys}}} = 8$ for L-cysteine and $\text{pH} = 3.3$ (spray medium).

For comparison, we also tagged β -lactoglobulin A with 1,4-hydroquinone and in this case the e-TE% measured (Table 2) yields a rate constant value of $k_{\text{p}} = 105 \text{ M}^{-1}\cdot\text{s}^{-1}$. Considering Equation 7 and the fixed thiol $\text{p}K_{a_{\text{p}}}$ value of β -lactoglobulin A in the medium, Equation 8 must be satisfied:

$$\left(\frac{k_{\text{p}}}{k_{\text{cys}}} \right)_{\text{methoxycarbonyl-1,4-hydroquinone}} = \left(\frac{k_{\text{p}}}{k_{\text{cys}}} \right)_{1,4\text{-hydroquinone}} \quad (\text{Equation 8})$$

This ratio is however found three times larger for 1,4-hydroquinone than for methoxycarbonyl-1,4-hydroquinone showing the role of steric hindrance after consideration of $\text{p}K_a$ effects. These results clearly show that the decrease of the tagging rate of β -lactoglobulin A by methoxycarbonyl-1,4-hydroquinone from the predicted value of 79% to the experimental value of 26% does not only result from the thiol increased $\text{p}K_{a_{\text{p}}}$ in the protein but also from other factors. The steric hindrance acting on the EC-tag can be considered to have a relevant influence on the decrease of the kinetics. Indeed, free access of the probe to the cysteine is notably hindered by the protein chain even in denaturing solvents as the 3D structure is conserved by the presence of disulfide bridges even at low pH.^{43,44}

In conclusion, the low tagging extent e-TE% of β -lactoglobulin A is basically associated to kinetic limiting factors linked with the cysteine environment (*i.e.* the $\text{p}K_a$ value and the steric hindrance). The EC-tagging is shown to probe the general thiol reactivity by the mass spectrometric measurement of the extent of the tagging reaction.

3.3. Multi-EC-tagging of reduced β -lactoglobulin A

To explore the effect of steric hindrance, we studied the tagging of β -lactoglobulin A, but after chemically reducing it. Indeed, the cleavage of the two disulfide bonds of the protein will alter completely the tertiary structure, due to almost complete unfolding of the protein chain.

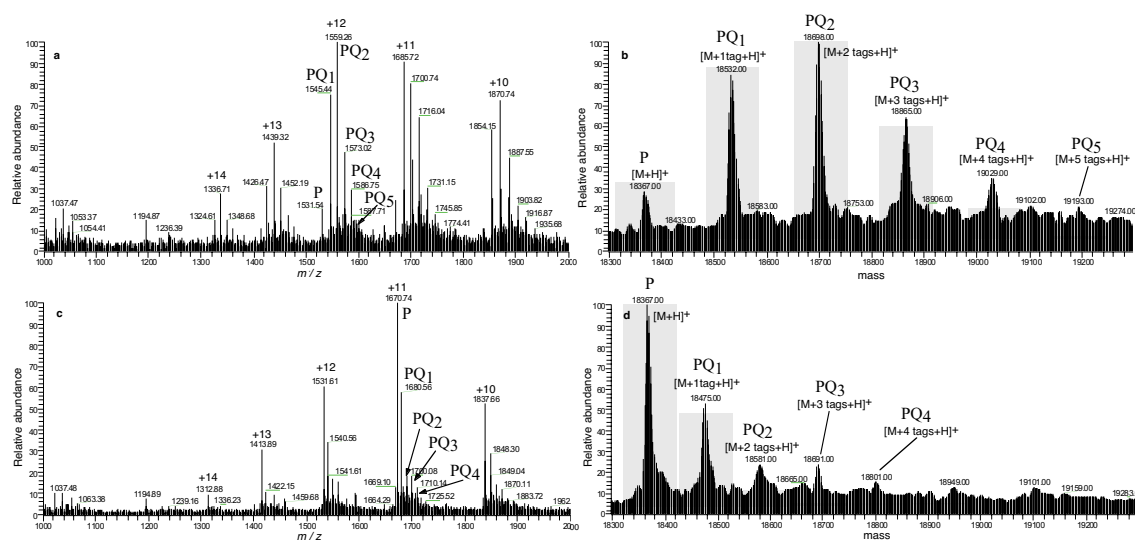


Figure 7. Microspray mass spectra of reduced β -lactoglobulin A (50 μ M) infused with methoxycarbonyl-1,4-hydroquinone (2.5 mM) (a) and 1,4-hydroquinone (2.5 mM) (c). Deconvolution mass spectra of reduced β -lactoglobulin A (50 μ M) infused with methoxycarbonyl-1,4-hydroquinone (2.5 mM) (b) and 1,4-hydroquinone (2.5 mM) (d). Grey bars give the predicted distributions of the species according to the analytical model. P and PQ_{1...5} indicate respectively the untagged and successively tagged proteins.

Reduced β -lactoglobulin A was EC-tagged using methoxycarbonyl-1,4-hydroquinone and 1,4-hydroquinone as shown in Figure 7a-b and c-d respectively. As predicted by the analytical kinetic model, methoxycarbonyl-1,4-hydroquinone provides clearly a very efficient cysteine labelling when compared to 1,4-hydroquinone. Indeed, the tagging of the five cysteines of reduced β -lactoglobulin A can be observed as shown in Figure 7b. By comparing mass spectra of reduced and non-reduced forms of the protein, the method gives access to the number of disulfide bonds. Besides, comparison with the kinetic model that fits perfectly (Figure 7b and 7d) shows clearly the role of the steric hindrance for the non-reduced protein.

The multi-EC-tagging of proteins by methoxycarbonyl-1,4-hydroquinone was further tested with creatine phosphokinase (CK) and reduced insulin. By reduction, insulin

splits into two polypeptidic chains. The B-chain contains mainly basic amino acids of the protein, conferring to the polypeptide a higher ionization in the spray medium (MeOH / H₂O / AcOH 50% / 49% / 1%) compared to that of the A-chain which is not detected.⁴⁵ The B-chain of insulin is efficiently EC-tagged by methoxycarbonyl-1,4-hydroquinone since the untagged species is no longer observed. The fully tagged product appears greatly enhanced relative to the 1,4-hydroquinone reaction (see Figure 8).

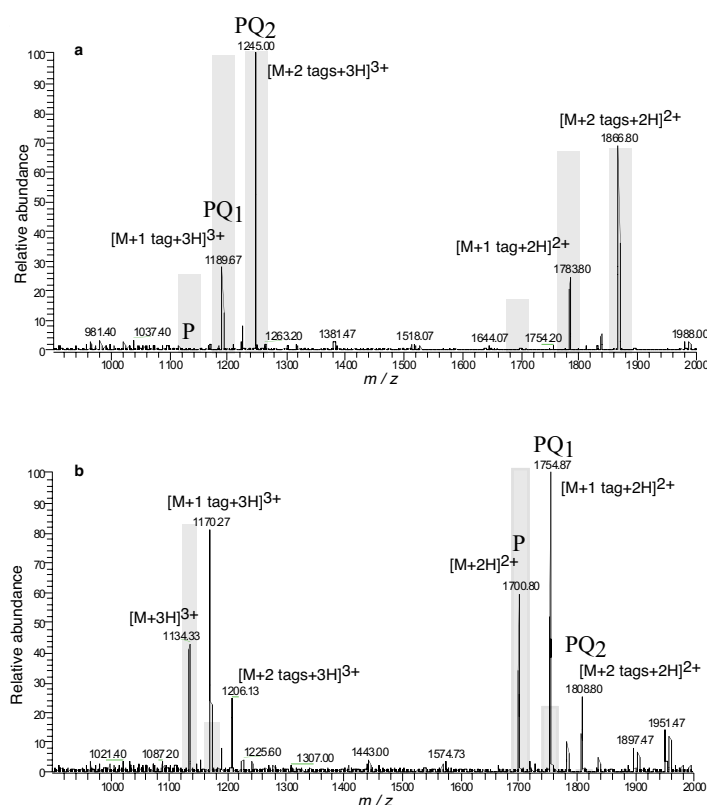


Figure 8. Microspray mass spectra of reduced insulin (50 μ M) infused with methoxycarbonyl-1,4-hydroquinone (2.5 mM) (a) and 1,4-hydroquinone (2.5 mM) (b). Only the B-chain was detected. Grey bars give the predicted distributions according to the analytical model. P, PQ₁ and PQ₂ indicate respectively the untagged, singly- and doubly-tagged proteins.

CK, which contains four free cysteine residues⁴⁶, is EC-tagged on every residues. As for the previous case, MS gives peaks ranging from the untagged protein to the quadruply tagged protein, thus providing the on-line counting of the cysteine units (Figure 9).

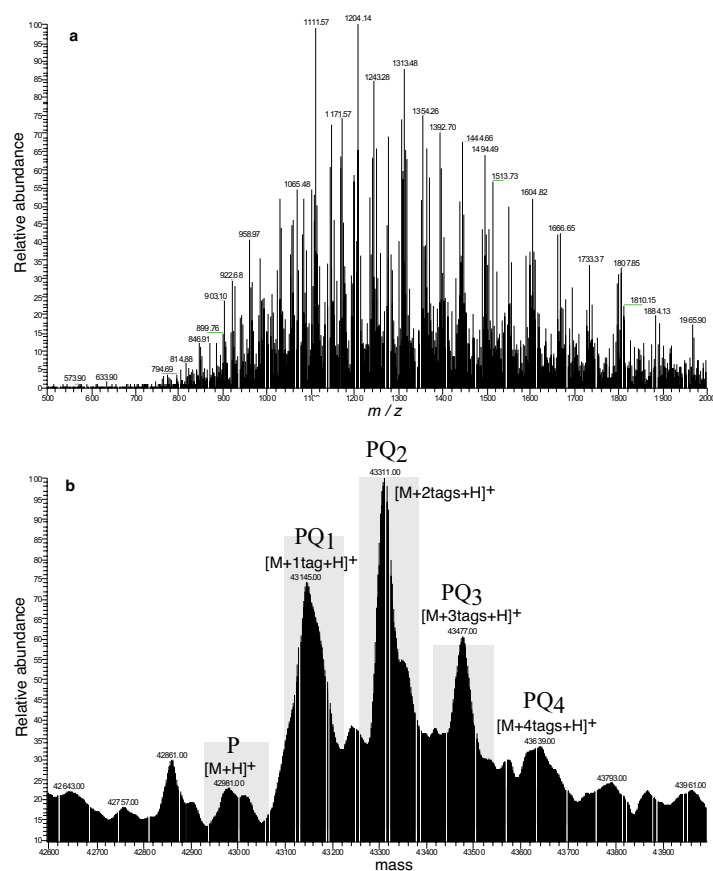


Figure 9. Microspray mass spectrum of creatine phosphokinase (50 μ M) infused with methoxycarbonyl-1,4-hydroquinone (2.5 mM) (a). Deconvolution mass spectrum (b). Grey bars give predicted the distribution of the species according to the analytical model. P and PQ_{1...4} indicate respectively the untagged and successively tagged proteins.

The experiments also provide information on the CK structure in the spray medium as under native conditions, only one cysteine unit is accessible for labelling.⁴⁷ These results therefore show that CK unfolds efficiently in the spray medium (MeOH / H₂O / AcOH 50% / 49% / 1%) to afford the quadruple tagging reaction.

This last study demonstrates that methoxycarbonyl-1,4-hydroquinone is suitable for the multi-EC-tagging at the protein level. For multi-cysteine-containing unfolded protein,

the analytical model fits perfectly well with the apparent tagging (Figures 6 and 7). Whatever the individual cysteine reactivity, the increase of the rates (*i.e.* nk) especially and compensation phenomena between higher and lower reactive thiols reduce the divergence.

4. Conclusion

The EC-tagging technique, by non-quantitative labelling, was shown to allow quick detection of cysteine-containing proteins by detection of the specific mass shift(s) corresponding to the marker. An analytical kinetic model taking into account the electrochemical and the chemical steps of the labelling reaction was developed to predict the tagging extent before MS analysis. Validated by comparison with numerically simulated data, the model was shown to fit perfectly well to the EC-tagging experiments on unfolded multi-cysteinyll-containing protein.

In addition, experiments with unmodified proteins and their chemically reduced forms show the strong effect of the cysteine site reactivity on the EC-tagging efficiencies. Chemically reducing the proteins prior to tagging leads both to a successive multi-tagging of the cysteine units liberated by the reduction of disulfide bonds and to an increase in the reaction kinetics. Methoxycarbonyl-1,4-hydroquinone was found to be efficient to tag up to five cysteines in biomacromolecules.

When decoupled from ionization factors eventually induced by the tag, the EC-tagging by methoxycarbonyl-1,4-hydroquinone reveals itself as an useful tool to probe the thiol reactivity of a single-thiol-containing protein. A similar study may be carried out for singly cysteinyl peptides that are by far the most common cysteinyl peptides resulting from proteolytic digestion of the proteome.⁴⁸ For peptides or small molecules presenting negligible steric constraint, the EC-tagging based on mass spectrometric detection could thus directly provide the thiol pK_a values that are commonly determined by potentiometric and electrophoretic means.

EC-microreactors directly coupled to ESI-MS are attractive tools that can be extensively used for probing the surface accessibilities of specific amino acids. The

development of specifically targeted tags will help providing information about the protein microenvironment.

Bibliography

- 1 A.J. Stewart, C.A. Blindauer, S. Berezenko, D. Sleep, D. Tooth and P.J. Sadler, *FEBS J.*, 2005, 272, 353-362.
- 2 B.P. Esposito and R. Najjar, *Coord. Chem. Rev.*, 2002, 232, 137-149.
- 3 P. Di Simplicio, F. Franconi, S. Frosali and D. Di Giuseppe, *Amino Acids*, 2003, 25, 323-339.
- 4 Z. Shaked, R.P. Szajewski and G.M. Whitesides, *Biochemistry*, 1980, 19, 4156-4166.
- 5 M.P. Lutolf, N. Tirelli, S. Cerritelli, L. Cavalli and J.A. Hubbell, *Bioconjugate Chem.*, 2001, 12, 1051-1056.
- 6 J.F. de la Mora, G.J. Van Berkel, C.G. Enke, R.B. Cole, M. Martinez-Sanchez and J.B. Fenn, *J. Mass Spectrom.*, 2000, 35, 939-952.
- 7 G. Diehl and U. Karst, *Anal. Bioanal. Chem.*, 2002, 373, 390-398.
- 8 T.C. Rohner, N. Lion and H.H. Girault, *Phys. Chem. Chem. Phys.*, 2004, 6, 3056-3068.
- 9 V. Kertesz, H.T. Deng, K.G. Asano, R.L. Hettich and G.J. Van Berkel, *Electroanalysis*, 2002, 14, 1027-1030.
- 10 T.Y. Zhang, S.P. Pali, J.R. Eyler and A. Brajter-Toth, *Anal. Chem.*, 2002, 74, 1097-1103.
- 11 T.Y. Zhang and A. Brajter-Toth, *Anal. Chem.*, 2000, 72, 2533-2540.
- 12 H.T. Deng and G.J. Van Berkel, *Anal. Chem.*, 1999, 71, 4284-4293.
- 13 M.C.S. Regino and A. Brajter-Toth, *Anal. Chem.*, 1997, 69, 5067-5072.
- 14 H.T. Deng and G.J. Van Berkel, *Electroanalysis*, 1999, 11, 857-865.
- 15 P.H. Gamache, D.F. Meyer, M.C. Granger and I.N. Acworth, *J. Am. Soc. Mass Spectrom.*, 2004, 15, 1717-1726.
- 16 J.M.E. Quirke, C.L. Adams and G.J. Van Berkel, *Anal. Chem.*, 1994, 66, 1302-1315.
- 17 J.M.E. Quirke, Y.L. Hsu and G.J. Van Berkel, *J. Nat. Prod.*, 2000, 63, 230-237.
- 18 B. Seiwert, H. Henneken and U. Karst, *J. Am. Soc. Mass Spectrom.*, 2004, 15, 1727-1736.
- 19 G.J. Van Berkel and K.G. Asano, *Anal. Chem.*, 1994, 66, 2096-2102.

- 20 G.J. Van Berkel, J.M.E. Quirke and C.L. Adams, *Rapid Commun. Mass Spectrom.*, 2000, 14, 849-858.
- 21 G.J. Van Berkel, J.M.E. Quirke, R.A. Tigani, A.S. Dilley and T.R. Covey, *Anal. Chem.*, 1998, 70, 1544-1554.
- 22 C. Roussel, L. Dayon, N. Lion, T.C. Rohner, J. Josserand, J.S. Rossier, H. Jensen and H.H. Girault, *J. Am. Soc. Mass Spectrom.*, 2004, 15, 1767-1779.
- 23 T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.
- 24 T.C. Rohner, J.S. Rossier and H.H. Girault, *Electrochem. Commun.*, 2002, 4, 695-700.
- 25 L. Dayon, C. Roussel, M. Prudent, N. Lion and H.H. Girault, *Electrophoresis*, 2005, 26, 238-247.
- 26 L. Dayon, C. Roussel and H.H. Girault, *Chimia*, 2004, 58, 204-207.
- 27 C. Roussel, L. Dayon, H. Jensen and H.H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.
- 28 C. Carru, L. Deiana, S. Sotgia, G.M. Pes and A. Zinellu, *Electrophoresis*, 2004, 25, 882-889.
- 29 J.H. Bourell, K.P. Clauser, R. Kelley, P. Carter and J.T. Stult, *Anal. Chem.*, 1994, 66, 2088-2095.
- 30 W.M. Abdou, Y.O. ElKhoshnieh and A.A. Kamel, *J. Chem. Res.-S*, 1996, 326-327.
- 31 O. Crescenzi, G. Prota, T. Schultz and L.J. Wolfram, *Tetrahedron*, 1988, 44, 6447-6450.
- 32 J.S. Rossier, R. Ferrigno and H.H. Girault, *J. Electroanal. Chem.*, 2000, 492, 15-22.
- 33 H.H. Girault, *Analytical and Physical Electrochemistry*, 1st ed., EPFL Press, Lausanne, 2004, pp. 285-286.
- 34 P.R. Unwin and R.G. Compton, *Comprehensive Chemical Kinetics*, Vol. 29, Eds.: R. G. Compton and A. Hammet, Elsevier, Amsterdam, 1989, pp. 173-193.
- 35 V.G. Levich, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, 1962, pp. 112-116.
- 36 A.R. Amaro, G.G. Oakley, U. Bauer, H.P. Spielmann and L.W. Robertson, *Chem. Res. Toxicol.*, 1996, 9, 623-629.
- 37 C. Roussel, T.C. Rohner, H. Jensen and H.H. Girault, *ChemPhysChem*, 2003, 4, 200-206.

- 38 L. Dayon, J. Josserand and H.H. Girault, *Phys. Chem. Chem. Phys.*, 2005, 7, 4054-4060.
- 39 D.Y. Ren, S. Julka, H.D. Inerowicz and F.E. Regnier, *Anal. Chem.*, 2004, 76, 4522-4530.
- 40 W.P. Jencks and J. Carriuolo, *J. Am. Chem. Soc.*, 1960, 82, 1778-1786.
- 41 O.A. Bizzozero, H.A. Bixler and A. Pastuszyn, *Biochim. Biophys. Acta-Protein Struct. Molec. Enzym.*, 2001, 1545, 278-288.
- 42 B.A. Grzybowski, J.R. Anderson, I. Colton, S.T. Brittain, E.I. Shakhnovich and G.M. Whitesides, *Biophys. J.*, 2000, 78, 652-661.
- 43 S. Uhrinova, M.H. Smith, G.B. Jameson, D. Uhrin, L. Sawyer and P.N. Barlow, *Biochemistry*, 2000, 39, 3565-3574.
- 44 K.R. Babu, A. Moradian and D.J. Douglas, *J. Am. Soc. Mass Spectrom.*, 2001, 12, 317-328.
- 45 J.L. Stephenson, B.J. Cargile and S.A. McLuckey, *Rapid Commun. Mass Spectrom.*, 1999, 13, 2040-2048.
- 46 H.R. Wang, J.H. Bai, S.Y. Zheng, Z.X. Wang and H.M. Zhou, *Biochem. Biophys. Res. Commun.*, 1996, 221, 174-180.
- 47 N. Tanaka, T. Tonai and S. Kunugi, *Biochim. Biophys. Acta-Protein Struct. Molec. Enzym.*, 1997, 1339, 226-232.
- 48 N. Lion, PhD thesis, Ecole Polytechnique Fédérale de Lausanne, 2006.

Annexe: Maple program to solve the kinetic equations for the tagging of a five-cysteine-containing protein

```

> restart:deq1:=diff(PQ1(t),t)=5*k*PP(t)*BQ(t)-4*k*PQ1(t)*BQ(t);
> deq2:=diff(PQ2(t),t)=4*k*PQ1(t)*BQ(t)-3*k*PQ2(t)*BQ(t);
> deq3:=diff(PQ3(t),t)=3*k*PQ2(t)*BQ(t)-2*k*PQ3(t)*BQ(t);
> deq4:=diff(PQ4(t),t)=2*k*PQ3(t)*BQ(t)-k*PQ4(t)*BQ(t);
> deq5:=diff(PQ5(t),t)=k*PQ4(t)*BQ(t);
> deq6:=diff(PP(t),t)=-5*k*PP(t)*BQ(t);
> deq7:=diff(BQ(t),t)=-5*k*PP(t)*BQ(t)-4*k*PQ1(t)*BQ(t)-3*k*PQ2(t)*BQ(t)-
2*k*PQ3(t)*BQ(t)-k*PQ4(t)*BQ(t);
> sys:={deq1,PQ1(0)=0,deq2,PQ2(0)=0,deq3,PQ3(0)=0,deq4,PQ4(0)=0,deq5,PQ5(0)=0
,deq6,PP(0)=P0,deq7,BQ(0)=BQ0};
> fcns:={PQ1(t),PQ2(t),PQ3(t),PQ4(t),PQ5(t),PP(t),BQ(t)};
> rep:=dsolve(sys,fcns);
> assign(rep);
> P0:=50e-6;
> BQ0:=93e-6;
> k:=5000;
> s1:=eval(BQ(t), t=4.7);
> evalf(s1);
> s2:=eval(PP(t), t=4.7);
> evalf(s2);
> s3:=eval(PQ5(t), t=4.7);
> evalf(s3);

```

CHAPTER VIII. *Chemical modification with benzoquinone tags*

1. Introduction

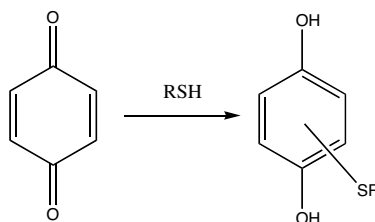
Living systems are dynamic and their development, survival, proliferation and necrosis are characteristic of each system. The identification and quantification of the proteins linked to stimuli is the basis of proteomics. Disease and regulatory events are in this way recognised as up- and down-regulation of certain proteins. The modification of amino acids has been tremendously used for relative quantification of parallel cellular extracts. Isotopic tags, like the isotope coded affinity tag (ICAT)¹, have enabled the relative quantification of proteins through mass spectrometry (MS). Nevertheless, specific modifications of amino acids are often performed for other reasons, namely the isolation of compounds of interest and the detection of characteristic feature. These modification strategies aim at reducing sample complexity and simplifying the analysis of data. They can be applied at the protein level or at the proteolytic peptide level.

Diagonal chromatography is based on a very simple observation. When a sample is chromatographed in the x direction of a paper and dried, and then rechromatographed in the perpendicular y direction using the same eluting condition, the analytes appear on the chromatogram in a diagonal line. Compounds that are chemically modified between the two runs in a way that their partition coefficient is affected, will consequently move from the diagonal and be specifically recognised. In 1966, J.R. Brown and B.S. Hartley described such a technique for the isolation of cysteine peptides by diagonal paper electrophoresis, by

exposing the paper to vapours of performic acid, converting cysteine to cysteic acid before the reseparation.² Similar procedures were employed later for the isolation of peptides carrying free amines³, and methionine-⁴, histidine-⁵ and tyrosine-containing peptides.⁶ The diagonal chromatography technique has been recently implemented for high performance liquid chromatography (HPLC) and has been employed for targeted proteomic isolation of analytes (a review has been recently published⁷). Indeed, fractions collected from a first HPLC run are chemically treated and rechromatographed. The modified analytes present then a different retention time (t_R) during the second run. Methionine-containing peptides and cysteine-containing peptides, respectively oxidised with oxygen peroxide and oxidised by Ellman's reagent (before the first run) and further reduced by tris(2-carboxyethyl)phosphine (TCEP) (before the second run), have been specifically isolated.^{8,9} In the same way, K. Gevaert *et al.* have demonstrated that the method so-called combined fractional diagonal chromatography (COFRADIC) is able to reduce the complexity of proteomic samples while keeping essential information for the identification when targeting the N-terminal tryptic peptide of proteins.¹⁰ F.E. Regnier and co-workers have investigated the potential of the methods for the study of post-translational modifications by enzymatically removing the phosphorylation modification before the second HPLC. They proposed optimized conditions for the isolation of peptides phosphorylated on serine and threonine using heptafluorobutyric acid (HFBA) as pairing agent. Whereas tyrosine phosphorylation was identified with diagonal chromatography, the deglycosylation of glycopeptides was shown to provide too small changes in t_R .¹¹ K. Gevaert *et al.* have used this approach to analyse the phosphoproteome of human HepG2 hepatocytes.¹²

D.E. Mason and D.C. Liebler proposed a protocol for the derivatization of cysteinyl peptides with 1,4-benzoquinone in distilled water (Scheme 1).¹³ In addition, they studied the fragmentation of benzoquinone-peptide adducts during electrospray ionization (ESI) MS and co-developed a pattern recognition algorithm called SALSA (scoring algorithm for spectral analysis) for the large screening of tandem MS (MS/MS) spectra for fragmentation characteristics indicative of specific peptide modification.¹⁴ The reaction of

benzoquinone (BQ) compounds with cysteine has been shown to be rapid and selective for the thiol group when performed in acidic medium.¹⁵



Scheme 1. The 1,4-addition of Michael for thiols on 1,4-benzoquinone is depicted.

In this chapter, an evaluation of the use of BQ compounds for the isolation of cysteine peptides by diagonal chromatography is proposed. The derivatization occurs between the first run and the second HPLC run, taking advantage of the possibility to perform the derivatization reaction in acidic conditions. Several substituted 1,4-benzoquinones, with increasing hydrophobicity, were evaluated as regards their selectivity, their reactivity as well as the t_R shift they induced on cysteinyl peptides. In a second part, the possibility to use BQ tags to modify cysteine directly on proteins is envisaged. The counting of cysteine in proteins thanks to BQ tagging was performed by sequential analysis of the proteins followed by that of the specifically modified analogues as described in the literature with other alkylating reagents.^{16, 17} The portion of the thiols involved in disulfide bridges can also be determined when operating the modification before and after the reduction of the disulfide bonds.

2. Experimental

2.1. Chemicals

1,4-Benzoquinone ($\geq 98\%$), methyl-1,4-benzoquinone ($\geq 98\%$), 2,5-dimethyl-1,4-benzoquinone ($\geq 98\%$) and L-cysteine ($> 99.5\%$) were from Fluka (Büchs, Switzerland) and 1,4-naphtoquinone hydrate (97 %) and tri-*n*-butylphosphine (TBP, 97%) were purchased from

Aldrich (Milwaukee, WI, USA). Methoxycarbonyl-1,4-benzoquinone was synthesized as previously described.¹⁵ Chorionic gonadotropin- β (109-119) amide (> 94%), Leu-enkephalin (> 99%), KCTCCA (70%) and angiotensin I triacetate salt (> 99%) were bought from Bachem (Bubendorf, Switzerland). Synthetic peptides AIKCTKF (> 70%), ALRCTCS (> 70%), ACKCTCM (> 70%) were prepared by Catherine Servis at the Institut de Biochimie (Faculté de Médecine, Epalinges, Switzerland). Bradykinin (99%), Met-enkephalin-Arg-Phe (98%), creatine phosphokinase (CK) from rabbit muscle, β -lactoglobulin A (BLA) from bovine milk (>90%) and myoglobin from horse heart (> 90%) were purchased from Sigma (St Louis, MO, USA). Methanol (> 99.8%, Riedel-de Haën, Seelze, Germany) trifluoroacetic acid (TFA, 99%, Merck, Darmstadt, Germany), *N,N*-dimethylformamide (DMF, ~ 99%, Fluka) and acetic acid (> 99.5%, Fluka) were used without any further purification. Deionised water (18.2 M Ω) was prepared using the Milli-Q system from Millipore (Bedford, MA, USA). Water (UV-HPLC) and acetonitrile (HPLC-gradient grade) were bought from Panreac Quimica S.A. (Barcelona, Spain).

2.2. Derivatization protocol

Proteins at 1 mg·mL⁻¹ or peptides at 50 μ M were derivatized by mixing with 5% (*V/V*) of BQ reagents at 100 mM in CH₃CN. The reaction was performed at room temperature until 60 min but followed every 15 minutes.

As regards the derivatization protocol used for diagonal chromatography, the collected fractions of peptides (at a concentration of 1-2 mM) were modified by addition of 1% (*V/V*) of BQ reagents at 100 mM in CH₃CN. The reaction was performed at room temperature for 60 min. For the further derivatization with the phosphine, 1% (*V/V*) of TBP 400 mM in DMF 10% was added. The mixture was stirred for 40 min.

2.3. Mass spectrometry

Mass spectra were recorded on a Finnigan LCQ duo ion trap apparatus (San José, CA, USA). The collected HPLC fractions were analysed off-line without further sample preparation using the classical ESI source.

2.4. High performance liquid chromatography

The HPLC separations were performed with an Alliance 2690 system from Waters (Milford, MA, USA.) comprising a Photodiode Array Detector 996. To test their stability, the BQ were analysed on an Acclaim 120 C18, 5 μ m column from Dionex (Sunnyvale, CA, USA) with an isocratic elution of CH₃CN at a flow rate of 1 mL·min⁻¹.

To evaluate the capabilities of cysteinyl peptide derivatization by BQ tags in diagonal chromatography, 1 μ mol of peptide was injected in the HPLC system. Fractions of 1-2 mL were collected and divided into two samples. 100 μ L of the first sample was directly reinjected into the HPLC. The second sample was modified with BQ. 100 μ L of the modified mixture was rechromatographed. Whereas the analyte of the first rechromatographed sample gave the same t_R than the initially chromatographed peptide, the modified sample revealed a shifted t_R . The peptides and modified peptides were analysed using a Nucleosil 100-5 C18 column from Macherey-Nagel (Düren, Germany). The separation was run for 60 min using a gradient of H₂O/TFA 99.88% / 0.12% (solvent A) and CH₃CN (solvent B). The gradient was run as follows: 0-3 min 100% A, then to 62% A and 38% B at 43 min, 43% A and 57% B at 50 min and 20% A, 80% B at 60 min at a flow rate of 1 mL·min⁻¹.

3. Results and discussions

3.1. Benzoquinone reagents: reactivity and selectivity

Stability of benzoquinone compounds:

The stability of 1,4-benzoquinone, methyl-1,4-benzoquinone, methoxycarbonyl-1,4-benzoquinone, 2,5-dimethyl-1,4-benzoquinone and 1,4-naphtoquinone was studied in CH₃CN and in H₂O on the basis of few hours. It was followed by HPLC using CH₃CN as solvent. In CH₃CN, all the compounds were found to be stable. Nevertheless apart from 1,4-benzoquinone, all the other BQ were found to degrade slightly in H₂O after 60 min. Methoxycarbonyl-1,4-benzoquinone was strongly degraded in this medium.

Selectivity of benzoquinone compounds:

The selectivity of the BQ for cysteine residues was studied using myoglobin, a protein containing every amino acids apart from cysteine. The reaction of the protein with an excess of BQ was controlled during 60 min, checking the products by MS. In acidic solvents (see experimental part), the BQ were found to be selective for the thiol of cysteine. They were not selective when the reaction was performed in deionised H₂O (Figure 1). Besides, the electrophilic 1,4-benzoquinone was shown to react with nucleophilic lysine and histidine of cytochrome *c* in the literature.^{18, 19}

Methoxycarbonyl-1,4-benzoquinone was revealed to react incompletely with other groups than the thiol even, in acidic medium. Mass shifts corresponding to the addition of methoxycarbonyl-1,4-benzoquinone were detected in the deconvolution mass spectrum of myoglobin. The high reactivity of methoxycarbonyl-1,4-benzoquinone, which is suitable for electrochemical tagging (EC-tagging) because of the short time of reaction in the microchannel^{15, 20}, is too big to use this BQ as a selective chemical reagent. Methoxycarbonyl-1,4-benzoquinone was thus abandoned for the following application.

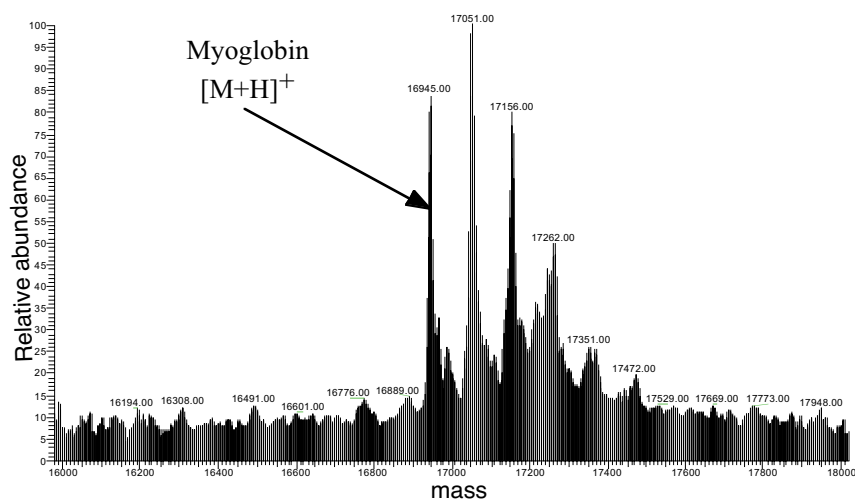


Figure 1. Deconvolution mass spectrum of myoglobin 1mg·mL⁻¹ derivatized with 1,4-benzoquinone (5 mM) in H₂O for 15 min.

The reaction of methoxycarbonyl-1,4-benzoquinone with other nucleophilic amino acids was studied with several peptides by MS control of the products (Table 1). In MeOH / H₂O / AcOH 50% / 49% / 1%, it was evaluated that, in addition to the thiol of cysteine, methoxycarbonyl-1,4-benzoquinone reacts preferentially with the alcohol function of serine and threonine. Indeed, extra peaks of variable intensity were seen in the presence of these amino acids in the peptidic sequence. When increasing the acidity of the medium using a strong acid like TFA, the selectivity for cysteine could be almost recovered (data not shown obtained with myoglobin).

<i>Peptide Sequence</i>	<i>Number of additions</i>
YGGFL	0
YGGFMRF	0
DRVYIHPFHL	+ 1 (2%)
RPPGFSPFR	+ 1 (46%)
ACKCTCM	3 + 1 (19%)
ALRCTCS	3 + 1 (31%) + 1 (1%)
AIKCTKF	1 + 1 (8%)
KCTCCA	3 + 1 (25%)

Table 1. The number of additions of methoxycarbonyl-1,4-benzoquinone on several peptides was followed by MS during 30 min of reaction. The reaction with cysteine is quantitative. Apart from the product of the reaction with cysteine, extra peaks (indicated as + 1) corresponding to supplementary adducts can be seen in the mass spectra.

Reactivity of benzoquinone compounds:

The general reactivity of all the BQ for cysteine was studied with chorionic gonadotropin- β (109-119) amide in MeOH / H₂O / AcOH 50% / 49% / 1%. The reaction was monitored using MS and the rapid and total disappearance of the starting peptide was noticed.

The mass spectra revealed that the oxidation of the formed adducts can occur with an excess of the reagents (see the following sections).

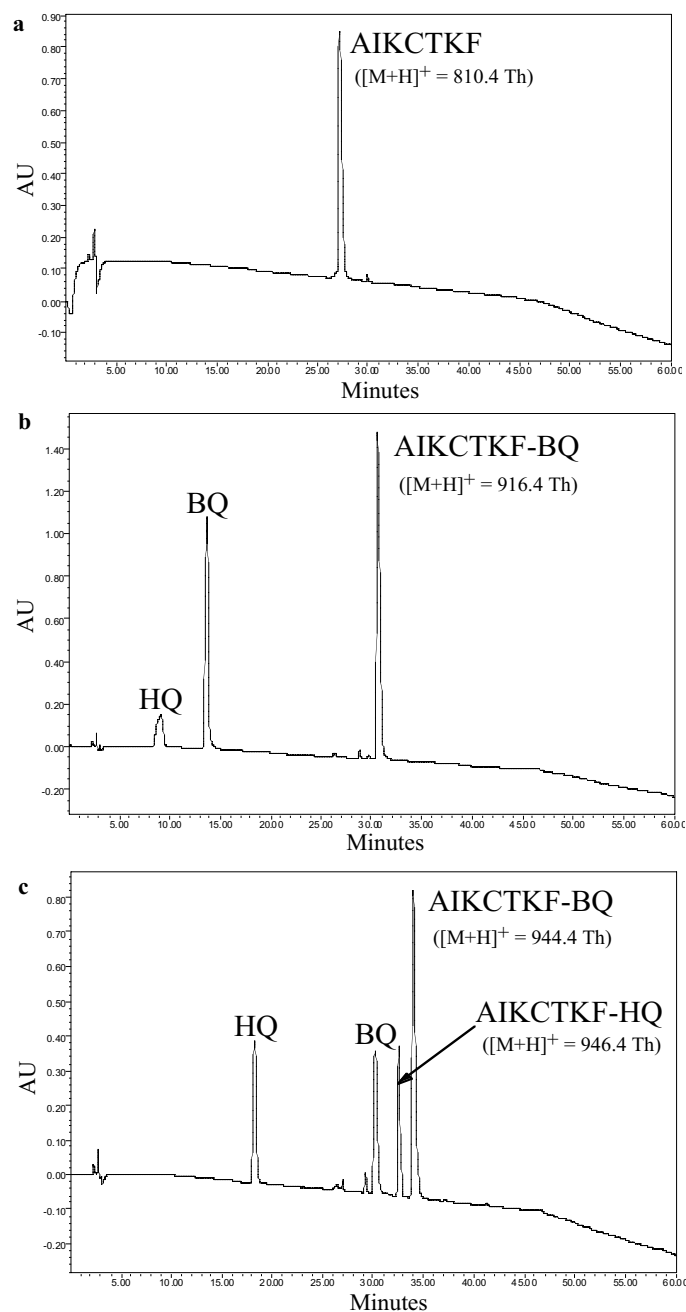


Figure 2. Reversed-phase chromatograms at 214 nm of AIKCTKF (a), AIKCTKF modified with 1,4-benzoquinone (b) and 2,5-dimethyl-1,4-benzoquinone (c). The benzoquinone and hydroquinone compounds are symbolised respectively by BQ and HQ in general.

3.2. Modification of peptides: application to diagonal chromatography

The tagging of peptides with BQ was carried out for the isolation of cysteines-containing peptides *via* HPLC. 1,4-Benzoquinone, methyl-1,4-benzoquinone and 2,5-dimethyl-1,4-benzoquinone with increasing hydrophobicity were employed to observe the influence of the chemical modification of the peptides on the retention time (t_R). Synthetic peptide AIKCTKF was chromatographed by HPLC using a gradient of 60 min optimized for the separation of tryptic peptides. The fraction relative to the peptide was collected, treated by a BQ reagent and rechromatographed. The chromatograms obtained with and without derivatization are depicted in Figure 2.

The addition of the thiol onto methyl-1,4-benzoquinone has been previously shown to occur on different sites of the BQ core.¹⁵ Isobaric adducts were thus separated during LC and identified by MS. Methyl-1,4-benzoquinone is consequently not convenient for clean derivatization. When using 1,4-benzoquinone as modifier, one product was quantitatively formed. Nevertheless, the adduct was found to be oxidised, exhibiting 2 Da of difference with the expected product (Figure 3).

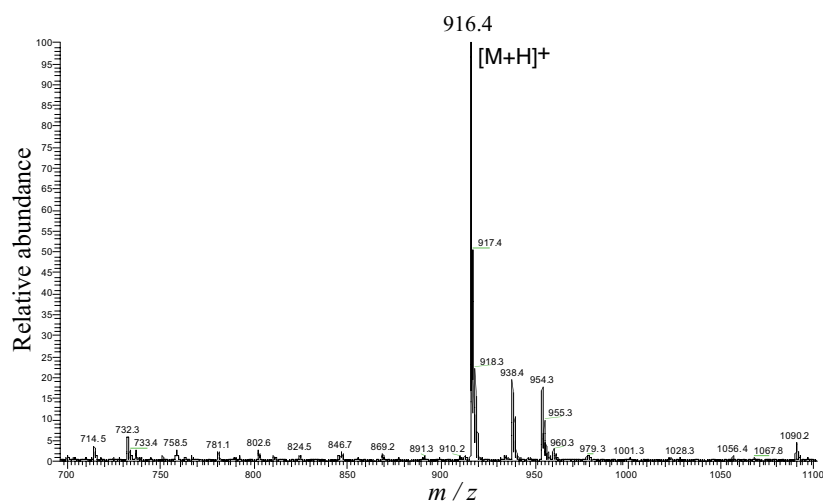
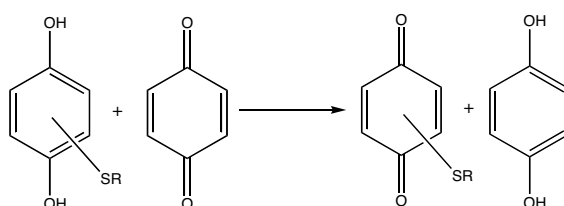


Figure 3. Mass spectrum of the product ($t_R = 30.6$ min) coming from the reaction of the AIKCTKF ($[M+H]^+ = 810.4$ Th) containing fraction with 1,4-benzoquinone ($M = 108.09$ g·mol⁻¹).

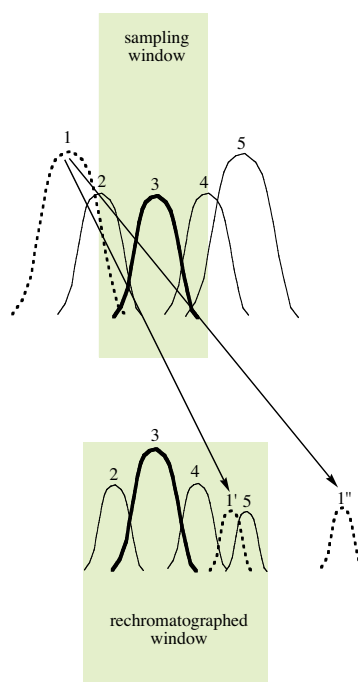
A control experiment by reaction of the obtained adduct with L-cysteine showed that an addition of the L-cysteine occurs. It confirmed that the oxidised form of the adduct is obtained. In fact, the excess BQ is able to oxidise the hydroquinone (HQ) added to the peptide to provide the BQ added peptide (Scheme 2). Besides, HQ formed by the reaction appears in the chromatogram (Figure 2b).



Scheme 2. The excess of BQ is able to re-oxidised the HQ adduct.²¹

When 2,5-dimethyl-1,4-benzoquinone served as tagging reagent, the adducts obtained were the oxidised ones in majority and the non oxidised ones ($t_R = 32.6$ min) as illustrated in Figure 2c. The oxidation by the excess reagent is less efficient than before. As regards the addition, the reaction is almost total. Only slight amount of the unmodified peptide can be seen.

The initial product was found at $t_R = 27.2$ min. After modification, the adduct came out at $t_R = 30.6$ min and 34.0 min respectively when modified with 1,4-benzoquinone and 2,5-dimethyl-1,4-benzoquinone. Due to the higher hydrophobicity of the latter, the shift in the t_R is larger. Respective t_R shifts of 3.4 and 6.8 min show that the BQ tags can be potentially used for the recognition of cysteinyl peptides. Indeed, the minimum time shift of 3.4 min obtained with 1,4-benzoquinone is theoretically sufficient to be used in diagonal chromatography when sampling collection of 1 min for the first dimension.¹¹ The fraction collection from the first dimension of diagonal chromatography will be much wider and will show multiple analytes in the second HPLC run. For successful isolation, the modified peptides must have a change in the t_R sufficient to go out from the overlap region in order to be recognised from false positive going out from the initial time interval (Scheme 3).



Scheme 3. The constraints imposed by diagonal chromatography are illustrated according to F.E. Regnier and co-workers.¹¹ For complex mixture samples, several components are detected in one fraction. When this fraction is rechromatographed, its time window becomes wider than the initial interval. The modification of the peptide must induce a sufficient shift in the t_R (modified peptide 1'') in such a way that the modified peptide does not fall in the rechromatographed window (modified peptide 1').

Peptides that present two and three cysteines were studied in the same way with 1,4-benzoquinone. ALRCTCS ($t_R = 21.7$ min, $[M+H]^+ = 753.3$ Th) was rechromatographed after modification. The double adduct was obtained at $t_R = 31.7$ min under its oxidised form ($m/z = 965.3$ Th). ACKCTCM ($[M+H]^+ = 759.2$ Th) was also tested. The triply tagged product was recovered ($m/z = 1077.1$ Th). The t_R shift was 12.4 min (from 23.6 to 36.0 min). Few side products were found before the desired product, that surely correspond to incomplete reaction (Figure 4). Similar experiments with 2,5-dimethyl-1,4-benzoquinone showed that the reaction is less efficient. In fact, several products with different oxidation states were obtained.

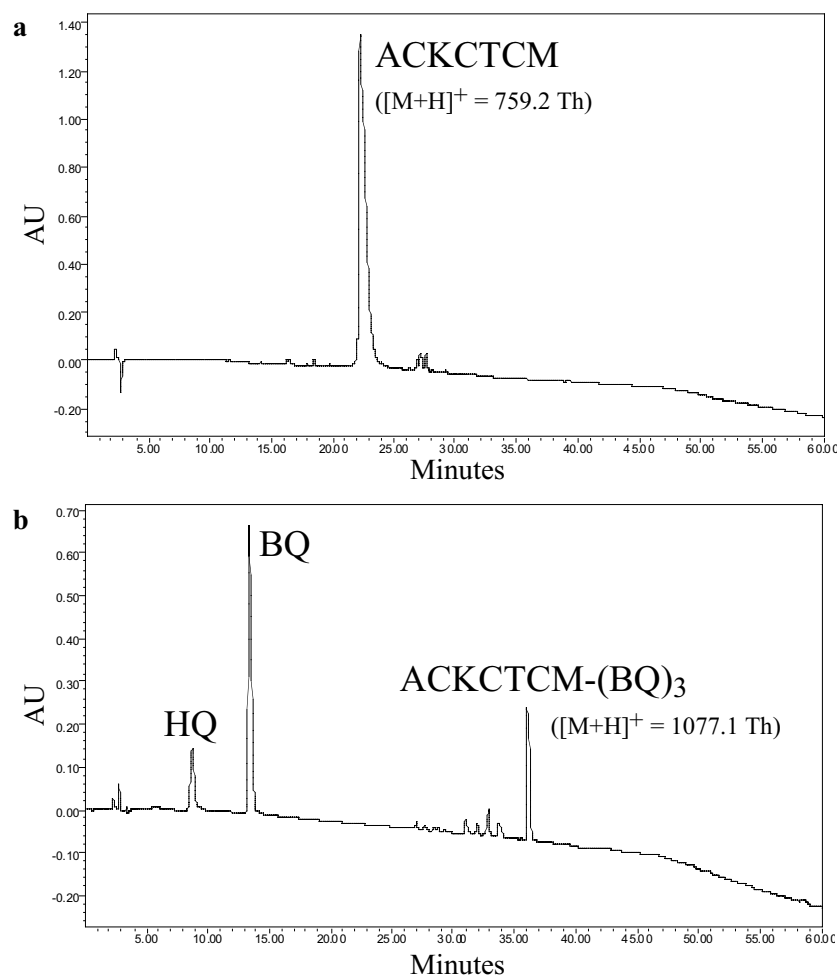


Figure 4. Reversed-phase chromatograms at 214 nm of ACKCTCM (a) and ACKCTCM modified with 1,4-benzoquinone (b). 1,4-Benzoquinone and 1,4-hydroquinone compounds are symbolised respectively by BQ and HQ.

The capability to use 1,4-benzoquinone as modification agent for diagonal chromatography has been demonstrated. The t_R shift induced by the modifier is sufficient to unambiguously identify cysteinyl peptides. The re-oxidation of the added HQ by the excess of the BQ reagent could be advantageously exploited. By further reaction with tri-*n*-butylphosphine²² (TBP) for instance after the BQ derivatization, the introduction of a permanent charge was achieved (Figure 5). This additional modification may provide an enhancement in the MS ionization efficiency of weakly ionisable peptides. The TBP addition increases even more the hydrophobicity. Moreover, the re-oxidation of the added HQ into BQ

can be exploited to access the number of cysteine in the peptide by comparison of the isolated compound before and after reaction with TBP (thiols like cysteine can also be employed).

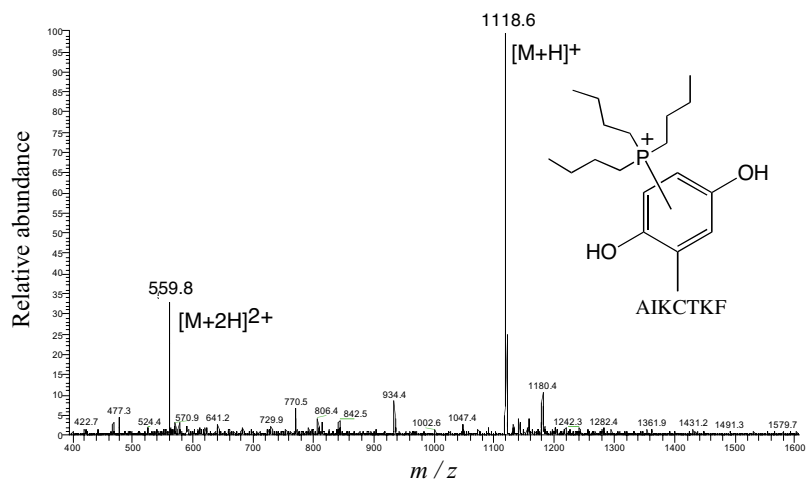


Figure 5. Mass spectrum of AIKCTKF derivatized with 1,4-benzoquinone, followed by further modification of the product with TBP.

Another interesting feature of the BQ derivatization is that the product is active under UV and presents a characteristic maximum of absorption ($\lambda_{\max} = 237$ nm). When scanning the wavelength, it is then possible to have a higher certainty on the isolated peptides.

3.3. Modification of proteins: counting of cysteines residues

BQ were evaluated as new alkylating reagents of cysteines in proteins. Creatine phosphokinase (CK) from rabbit muscle and β -lactoglobulin A (BLA) in different acidic media (MeOH / H₂O / AcOH 50% / 49% / 1%, CH₃CN / H₂O / TFA 20-50-80% / 79.9-49.9-19.9% / 0.1%) were put to react with 1,4-benzoquinone and 1,4-napthoquinone (Figure 6).

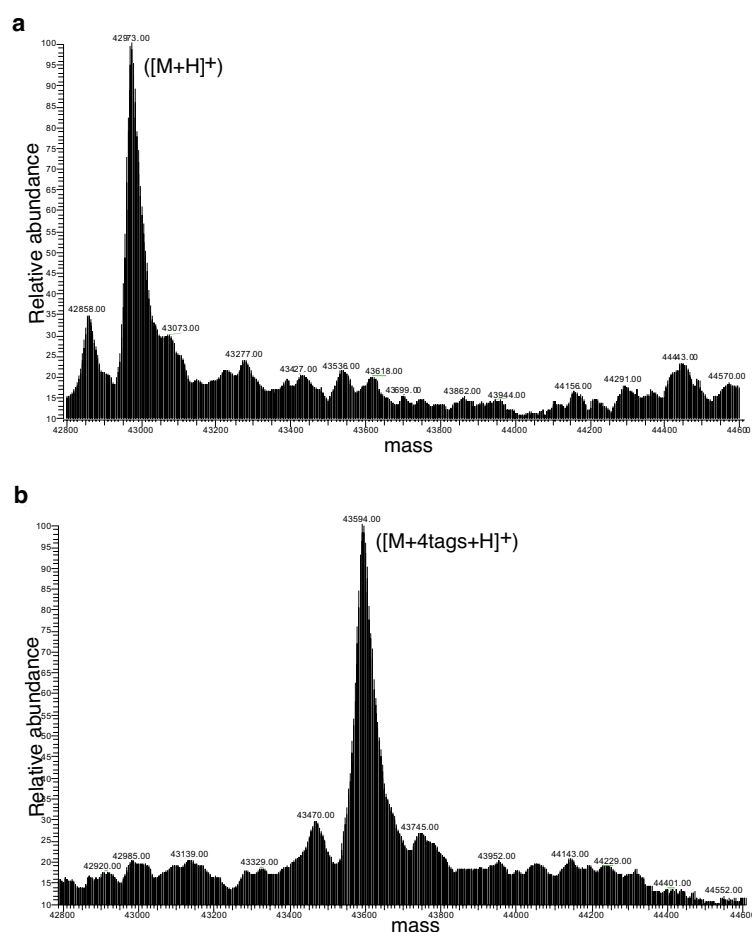


Figure 6. Deconvolution mass spectra of creatine phosphokinase (a) and its analogues modified with 1,4-naphtoquinone in MeOH / H₂O / AcOH 50% / 49% / 1% for 15 min (b).

1,4-Benzoquinone was found to provide a rapid modification of the two proteins that was complete after 15 min of reaction. Conversely, with 1,4-naphtoquinone, the reaction was less efficient and was incomplete even after 60 min when the reaction medium was essentially composed of CH₃CN. Although 1,4-naphtoquinone provides a larger mass addition per cysteine residue (158 Da), the simplest 1,4-benzoquinone revealed to be a better reagent. In CK, it gives access to the four free cysteines. In reduced BLA²⁰, the five cysteines were all modified after 15 min of reaction. By comparison of the mass spectra of the unmodified, tagged form and tagged reduced form of the protein, the method allows the rapid counting of cysteine units directly at the protein level. The method is actually pushed a step further by studying the modification of the reduced recombinant human serum albumin (r-HAS)

containing 35 cysteines. To study this larger protein (~ 66 kDa) an ESI time of flight (TOF) is required. First investigation has indicated that several products are formed. They may result from incomplete modification or may be relative to several oxidation states.

4. Conclusions

1,4-Benzoquinone has revealed to be a good alkylating reagent of cysteine groups in peptides. In acidic medium, it guarantees quantitative reactivity and 100% specificity. This interesting characteristic renders the BQ tags attractive when a reversed-phase LC separation is over, since in proteomics, acidic gradients are used for the analysis of peptides coming from proteolytic digestion. Furthermore, the reaction does not require any washing step. The derivatization has been shown to enable the isolation of cysteinyl peptides when combined to diagonal liquid chromatography ensuring a sufficient t_R shift for unambiguous analysis. Indeed, the introduction of the tag renders the peptides more hydrophobic. It was demonstrated that the use of 2-methyl and 2,5-dimethyl-1,4-benzoquinone provides greater t_R shifts but leads to unclean or incomplete reactions. Complex mixtures of tryptic peptides may be studied to fully validate the use of the BQ for the isolation of cysteine-containing peptides.

In proteins, the reaction with BQ has been shown to provide the cysteine content by sequential analysis of untagged and tagged samples. Nevertheless, tagging of reduced recombinant human serum albumin, which contains 35 cysteine units, by 1,4-benzoquinone, resulted in disperse ESI-MS signals. Control experiments seemed to indicate that the dispersion might be caused by gradual re-oxidation of the HQ tag by the free BQ that was added in excess over the substrate. Another explanation is that the tagging of cysteines happens incompletely. Experiments are ongoing including variations of the molar ratios between the substrate and BQ (driving reaction to completeness and producing completely oxidized attached tag). Reversing the oxidation by addition of reducing agent and changing the solvent conditions may also be foreseen.

Bibliography

- 1 S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, 17, 994-999.
- 2 J.R. Brown and B.S. Hartley, *Biochem. J.*, 1966, 101, 214-228.
- 3 P.J.G. Butler, J.I. Harris, B.S. Hartley and R. Leberman, *Biochem. J.*, 1967, 103, P78-P79.
- 4 J. Tang and B.S. Hartley, *Biochem. J.*, 1967, 102, 593-599.
- 5 W.H. Cruickshank, T.M. Radhakrishnan and H. Kaplan, *Can. J. Biochem.*, 1971, 49, 1225-1232.
- 6 W.H. Cruickshank, B.L. Malchy and H. Kaplan, *Can. J. Biochem.*, 1974, 52, 1013-1017.
- 7 K. Gevaert, P. Van Damme, L. Martens and J. Vandekerckhove, *Anal. Biochem.*, 2005, 345, 18-29.
- 8 K. Gevaert, B. Ghesquiere, A. Staes, L. Martens, J. Van Damme, G.R. Thomas and J. Vandekerckhove, *Proteomics*, 2004, 4, 897-908.
- 9 K. Gevaert, J. Van Damme, M. Goethals, G.R. Thomas, B. Hoorelbeke, H. Demol, L. Martens, M. Puype, A. Staes and J. Vandekerckhove, *Mol. Cell. Proteomics*, 2002, 1, 896-903.
- 10 K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas and J. Vandekerckhove, *Nat. Biotechnol.*, 2003, 21, 566-569.
- 11 P.R. Liu, C.L. Feasley and F.E. Regnier, *J. Chromatogr. A*, 2004, 1047, 221-227.
- 12 K. Gevaert, A. Staes, J. Van Damme, S. De Groot, K. Hugelier, H. Demol, L. Martens, M. Goethals and J. Vandekerckhove, *Proteomics*, 2005, 5, 3589-3599.
- 13 D.E. Mason and D.C. Liebler, *Chem. Res. Toxicol.*, 2000, 13, 976-982.
- 14 B.T. Hansen, J.A. Jones, D.E. Mason and D.C. Liebler, *Anal. Chem.*, 2001, 73, 1676-1683.
- 15 C. Roussel, L. Dayon, H. Jensen and H.H. Girault, *J. Electroanal. Chem.*, 2004, 570, 187-199.
- 16 F. Hubalek, J. Pohl and D.E. Edmondson, *J. Biol. Chem.*, 2003, 278, 28612-28618.

- 17 S. Neitz, M. Jurgens, M. Kellmann, P. Schulz-Knappe and M. Schrader, *Rapid Commun. Mass Spectrom.*, 2001, 15, 1586-1592.
- 18 M.D. Person, D.E. Mason, D.C. Liebler, T.J. Monks and S.S. Lau, *Chem. Res. Toxicol.*, 2005, 18, 41-50.
- 19 M.D. Person, T.J. Monks and S.S. Lau, *Chem. Res. Toxicol.*, 2003, 16, 598-608.
- 20 L. Dayon, C. Roussel and H.H. Girault, *J. Proteome Res.*, 2006, 5, 793-800.
- 21 A. Digga, S. Gracheva, C. Livingstone and J. Davis, *Electrochem. Commun.*, 2003, 5, 732-736.
- 22 W.M. Abdou, Y.O. ElKhoshnieh and A.A. Kamel, *J. Chem. Res.-S*, 1996, 326-327.

CHAPTER IX. *Conclusions and perspectives*

The electrochemical tagging (EC-tagging) of free cysteines in peptides and proteins has been developed and optimized. The technique relies on the electrochemistry inherent to the electrospray ionization (ESI) process. The high voltage applied between a microspray emitter and a mass spectrometer is used to generate tags at the electrode of the microspray that rapidly react with cysteines before the spray event. The products of the labelling reaction are continuously analysed by mass spectrometry (MS).

The EC-tagging consists in the modification of free cysteines by electrogenerated benzoquinone tags. In a first time, the selection of an efficient tagging reagent was carried out. The kinetics of the addition reaction with L-cysteine was determined for several substituted benzoquinones. Cyclic voltammetry (CV) and digital simulation were used to determine the rate constants of the reaction. Basically, the electrochemical-chemical-electrochemical (ECE) mechanism of the overall tagging reaction was exploited to extract the rate of production of the benzoquinone-adducted L-cysteine. Methoxycarbonyl-1,4-benzoquinone was found to provide an addition rate constant of $5000 \text{ M}^{-1}\cdot\text{s}^{-1}$ (*i.e.* one order of magnitude higher than the simple 1,4-benzoquinone) as the electron attracting group renders the reagent more reactive toward the nucleophilic addition of the thiol on the benzoquinone core. During this investigation, it was underlined that the ionization capability of the tags can be exploited to increase the final apparent tagging extent measured by MS. Consequently, derivatization coupled with ESI-MS enables the enhancement of the detection levels of molecules. In this way, it may be valuable to enlarge the panel of tagging reagents. The study of labelling reagents such as p-aminophenol (Figure 1) or dopamine may be carried out.

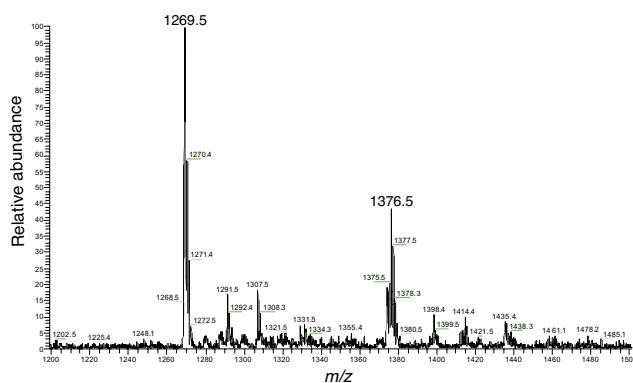


Figure 1. Mass spectrum of the EC-tagging of chorionic gonadotropin- β (109-119) amide ($M = 1269.30$ g·mol $^{-1}$) 50 μ M by p-aminophenol ($M = 109.13$ g·mol $^{-1}$) 2.5 mM in MeOH / H $_2$ O / AcOH 50% / 49% / 1%.

The on-line EC-tagging was employed to label cysteinyl peptides before MS analysis. By control of the tagging reaction, the technique was shown to give access to the cysteine content of peptides by simultaneous analysis of peptides and products of the tagging. Mass shift(s) from a peptide to its tagged analogues indicates how many tags are attached (*i.e.* how many cysteines are present in the sequence). The on-line cysteine counting was used to restrain database search in the identification process of model proteins. The potential of the method was proved. The next development will be the automation of the technique. Complex protein mixtures may then be studied to evaluate the capabilities of the EC-tagging as a post-column stage for wide applications in proteomics. In addition, the EC-tagging modification technique might be coupled to the exact mass measurement of peptides by Fourier transform ion cyclotron MS. The combination of both the mass accuracy and cysteine content will allow the reduction of the number of peptides needed for accurate identification of proteins.

The microspray emitter behaves as an electrolysis flow cell. Characterization of the emitter was carried out from an electrochemical point of view. The Levich equation for a laminar flow was validated as a first approximation for the calculation of the limiting convection-diffusion currents of oxidation in the microchannel of the microspray emitter. The influence of parameters such as the flow rate of the fluid and the concentration of the tags were studied. Low flow rates provide better apparent tagging extents during MS analysis, but the lower limit of applicable flow rate is basically governed by the microspray geometry. New

microspray emitter could thus be fabricated in order to decrease the feeding flow rate. Electrospray emitters allowing $\text{nL}\cdot\text{min}^{-1}$ scale flow rates would be suitable.

Numerical simulations were carried out to obtain the distributions of the successive tagged species along the microchannel. This study showed that the counting of cysteines in peptides is theoretically possible up to five cysteines since the coexistence of both the untagged and fully tagged peptides is achieved. Moreover, a range of concentration of both the tags and the peptides was determined to ensure efficient working conditions. This work is fundamental for automation of the technique in a classical gel-free bottom up approach. In addition, numerical simulation of consecutive reactions in a fluid flow is of interest for diverse microfluidic application, especially in the field of microchemical reactors.

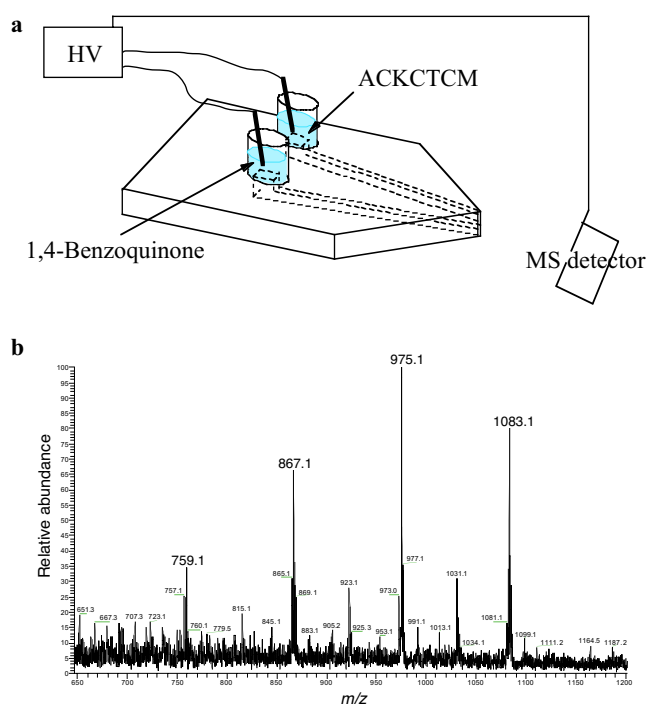


Figure 2. Scheme of the dual channel microspray emitter set-up (a). Mass spectrum of ACKCTCM ($M = 758.99 \text{ g}\cdot\text{mol}^{-1}$) $50 \mu\text{M}$ in H_2O tagged by 1,4-benzoquinone ($M = 108.09 \text{ g}\cdot\text{mol}^{-1}$) 5 mM in $\text{MeOH} / \text{AcOH} 99\% / 1\%$ using the dual channel set up (b).

An analytical model to predict rapidly the EC-tagging extent at the end of the microchannel was developed. By comparison of the model with MS measurement, the reactivity of cysteine residues in proteins can be evaluated. Indeed, the experimental tagging extent depends on both the reactivity and the accessibility of the cysteine. This measurement is therefore representative of every cysteinyl molecule as the reactivity depends on the amino acid sequence.

Many cysteine-labelling reagents are routinely used. Nevertheless, benzoquinone compounds have the characteristic to be appropriate for the modification of cysteine residues in acidic media. Diagonal chromatography was utilized to isolate cysteine-containing peptides. Basically, modification of peptides by 1,4-benzoquinone was performed between two identical liquid chromatography separations. The modified peptides exhibit different retention time (t_R) during the second run with shifts from at least 3 minutes. Actually, the t_R shift depends on both the number of cysteines and the initial hydrophobic properties of the peptides. In parallel, chemical derivatization experiments using a dual channel microspray emitter have also been carried out (Figure 2). In this special configuration, cysteinyl peptides and 1,4-benzoquinone are mixed at the nozzle of the chip. The tagging extent is controlled by supply of the benzoquinone reagent through the flow rate regulation.

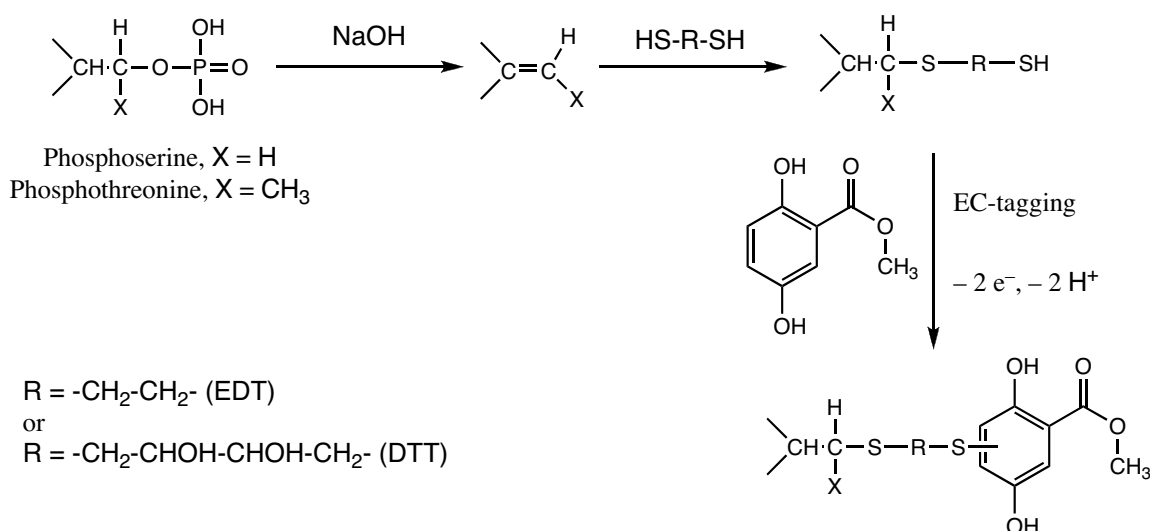


Figure 3. Strategy for the study of phosphorylated peptides integrating of the EC-tagging method.

Post-translational modifications (PTMs) change protein properties by proteolytic cleavage or by addition of a modifying group to certain amino acids. PTMs of a protein can determine its activity state, interaction with other proteins and localization. Many strategies of derivatization of phosphorylated sites are based on the elimination of the phosphate followed by addition of a dithiol like 1,2-ethanedithiol (EDT) (Figure 3). In this way, the hyphenation of the EC-tagging may reveal interesting to facilitate the detection of phosphorylated peptides and will allow the counting of the phosphorylated sites. Preliminary experiments have been carried out. The EC-tagging was performed on a phosphorylated peptide first subjected to elimination and addition of 1,4-dithio-DL-threitol (DTT) (Figure 3). The experiment has shown that the oxidation of the gem-diols (pinacols) in the microspray emitter is predominant over the oxidation of methoxycarbonyl-1,4-benzoquinone.

The coupling of electrochemistry and MS is directly feasible through the nature of the electrospray process. It might reveal as a complement to traditional electrochemical techniques to study electrochemical organic and inorganic reactions. Actually, reaction products can be directly analysed by MS. The on-chip oxidation may reveal itself especially suitable to replicate oxidations occurring in biological processes. The electrochemistry inherent to the electrospray may be integrated to various analytical strategies based on MS detection.

APPENDIX I. *Polyethylene terephthalate microspray emitter fabrication*

The fabrication of the microspray emitter (Figure 1) for mass spectrometry (MS) has been precisely described by T.C. Rohner *et al.*¹ Since the assembly of such devices is a key part of the EC-tagging work, it is relevant to recall the description of the emitter.

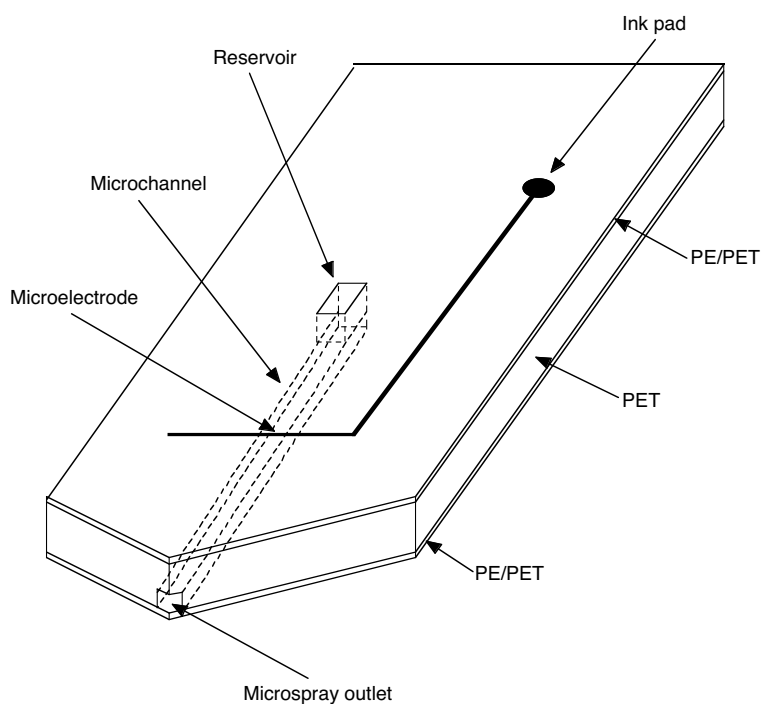


Figure 1. Microspray emitter.

¹ T.C. Rohner, J.S. Rossier and H.H. Girault, *Anal. Chem.*, 2001, 73, 5353-5357.

The process relies on the laser ablation of polymers. In 1982, the first reports of laser ablation of polymers were published almost simultaneously by Y. Kawamura *et al.*² and R. Srinivasan *et al.*³ R. Srinivasan went on to become a leader in the field, introducing for instance the terms of laser ablation and ablative photodecomposition. The removal of polymer material occurs at a well-defined fluence. In the process, the energy of a laser pulse is used to break bonds in a polymer substrate and to remove the decomposed fragments from the ablation area. The ablation is commonly carried out with an excimer laser that delivers pulses in the UV. The name excimer was first used as a contraction for excited dimer like xenon (Xe₂), krypton (Kr₂) or argon (Ar₂) but it now describes any molecular laser with a dissociative lower laser level. Typical lasers of that kind are argon fluoride, (ArF, 193 nm), krypton fluoride (KrF, 248 nm), xenon fluoride (XeF, 351 nm) and xenon chloride (XeCl, 308 nm). The mechanism of the laser ablation is still a rather controversial issue. It is generally agreed that the energy of laser pulses is initially transformed into electronic excitation. However, the subsequent energy transfer can take variety of forms. The photochemical models assume that the electronic excitations result directly in bond breaking whereas the photothermal models presume that the electronic excitations thermalize, yielding thermally broken bonds. In the photophysical models both thermal and nonthermal processes are taken into account.⁴ The surface chemistry of polymer can change tremendously when modified by laser ablation. Grazing angle X-ray photoelectron spectroscopy (XPS) studies of polyethylene terephthalate (PET) have shown that the ratio oxygen/carbon diminished dramatically after the ablation⁵ and that the formation of alcohol and aldehyde groups occurs.⁶ The redeposit of charged ejected material, results in an increase of the hydrophilicity.⁷

² Y. Kawamura, K. Toyoda and S. Namba, *Appl. Phys. Lett.*, 1982, 40, 374-375.

³ R. Srinivasan and V. Maynebantou, *Appl. Phys. Lett.*, 1982, 41, 576-578.

⁴ T. Lippert and J.T. Dickinson, *Chem. Rev.*, 2003, 103, 453-485.

⁵ S. Lazare, P.D. Hoh, J.M. Baker and R. Srinivasan, *J. Am. Chem. Soc.*, 1984, 106, 4288-4290.

⁶ H. Watanabe and M. Yamamoto, *J. Appl. Polym. Sci.*, 1999, 71, 2027-2031.

⁷ C.M. Chan, T.M. Ko and H. Hiraoka, *Surf. Sci. Rep.*, 1996, 24, 3-54.

In the typical set-up for the fabrication of microfluidic devices, the polymer substrate is mounted on a xy -table, perpendicularly to the laser beam (Figure 2). A mask defines the ablation area. In the dynamic mode, the complete pattern is drilled by translation of the xy -table, moving the sample during irradiation. In the static mode, no translation of the substrate is performed.⁸

The fabrication of the microspray emitter is performed as followed. The ablation stages are realised in a clean room. On one side of the 100 μm thick PET sheet^a, two perpendicular channels are drilled in the dynamic mode to pattern an “L” design (the width is 0.7 cm and the length is 4.5 cm) as shown in Figure 3a.



Figure 2. Laser excimer set-up (photograph reprinted with permission from V. Devaud and W. Baer).

The fluence and the pulse frequency of the ArF excimer laser^b are fixed (200 mJ, 50 Hz). The depth of the channel is thus adjusted by setting the speed of the xy -table. The channel shape is inspected and the dimensions are controlled using a microscope enabling the measurement of the length, the width and the depth of the structure. After ablation, the chip is

⁸ M.A. Roberts, J.S. Rossier, P. Bercier and H. Girault, *Anal. Chem.*, 1997, 69, 2035-2042.

^a Melinex[®] sheet from Dupont, Wilmington, DE, USA, $E_{\text{threshold}} = 40 \text{ mJ}\cdot\text{cm}^{-2}$.

^b from Lambda Physik, Göttingen, Germany.

cleaned with methanol or isopropanol and dried under nitrogen. The “L” designed channel is filled with carbon ink^c which is dried at 80 °C for 45 min (Figure 3b). For insuring the stability and the isolation of the conductive track, the substrate is then laminated (Figure 3c). In the lamination process, a thin PET foil coated with a melting adhesive layer of polyethylene (PE)^d is rolled onto the PET sheet with a heated roller (130 °C and 3 bars for 3 s). The chip is put at 80 °C for 45 min to insure the coating.

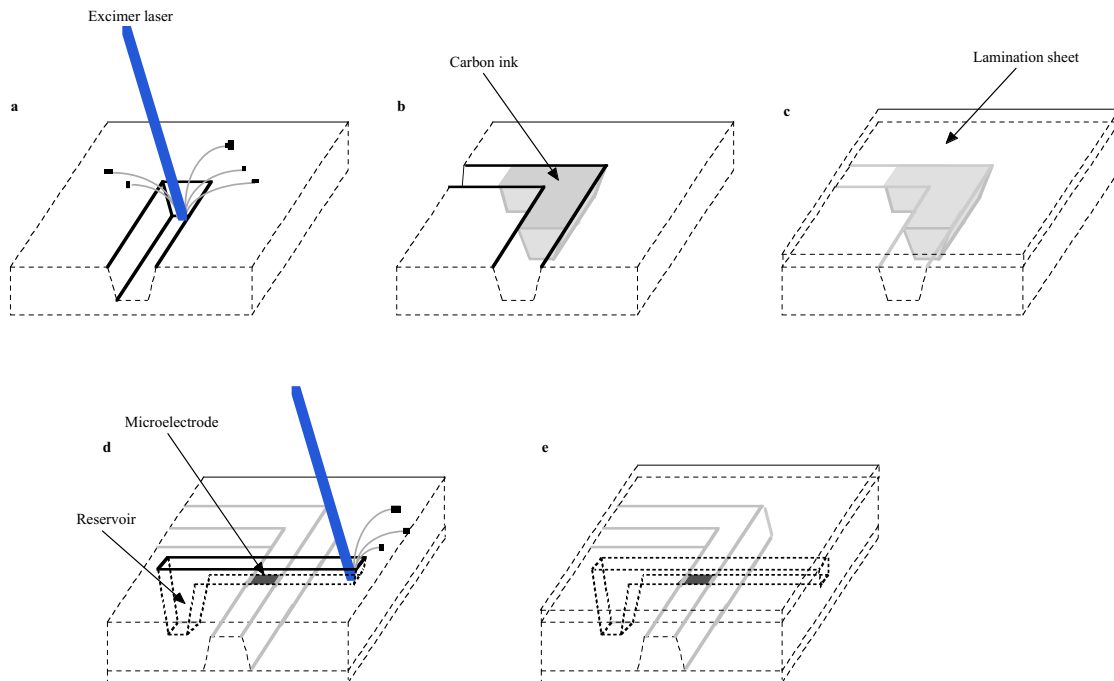


Figure 3. Main steps of the fabrication of the microspray emitter. The PET substrate is photoablated (a). The channel is fill with carbon paste (b). The track is sealed with a lamination sheet (c). On the other side of the substrate, the infusion microchannel is drilled, revealing the microelectrode (d). The infusion microchannel is closed by lamination (e).

The second ablation step of the fabrication is carried out on the opposite side of the PET sheet. Perpendicularly to the width of the “L” channel, a channel is drilled (Figure 3d). In fact, the depth is adjusted in order to reach the bottom of the carbon conductive

^c carbon Electrador ED 5000 paste from Electra polymers & chemicals Ltd., Tonbridge, UK.

^d PET, 25 μm thick with a 10 μm thick adhesive PE from Morane, Oxon, UK.

channel. Thus a microelectrode with controlled dimension is generated at the bottom of the channel that will serve for the infusion of the analyte as illustrated by the scanning electron microscopy (SEM)^c picture of Figure 4.

At the ends of that channel, a reservoir is made by static laser shootings that drill the PET sheet and the lamination right through. After cleaning as previously described, the PET sheet is laminated to close the channel and the chip is tempered in the oven (Figure 3e).

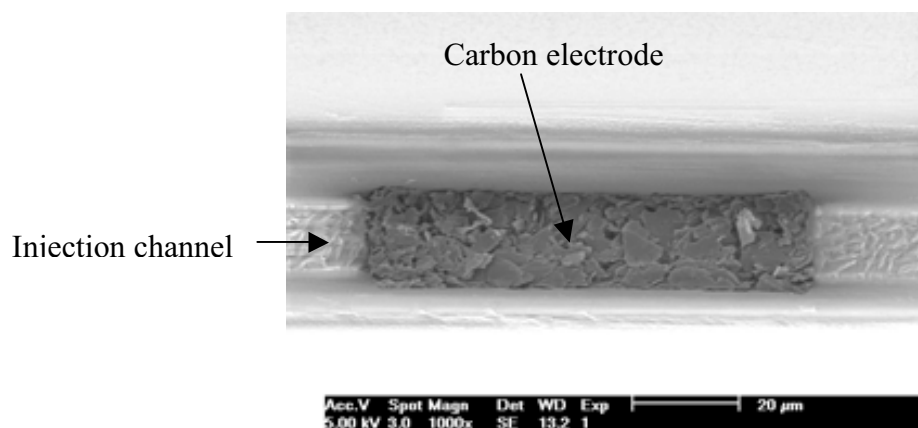


Figure 4. SEM picture at the inlaid electrode of the microchannel.

The contact between the power supply and the microelectrode is obtained through the conductive track and pad (Figure 1). This latter is obtained by removal of a part of the lamination and deposit of carbon ink which is let dried for 24 hours at room temperature. The final V-shaped nozzle is obtained by cutting the sheet with scissors.

A top view of the microelectrode is presented in Figure 4. The electrode appears slightly protruding and its roughness may reveal an increase of the surface area after the laser treatment. Numerical simulations have shown that the system can be controlled either by hemicylindrical or by linear diffusion in the direction of the channel, depending on the

^c Philips XL30 FEG from Philips Electron Optics, Eindhoven, the Netherlands.

geometric factors.⁹ The surface topography, crystallinity and wettability of the ablated PET polymer have been previously studied.¹⁰ For the dynamic ablation, an inhomogeneous surface is obtained and the surface has been characterized to be hydrophilic and highly wettable because of the redeposit fragments. In our case, because of the cleaning step, we should be closer to the results obtained in static mode revealing hydrophobic behaviour with a poor wettability. The cross section of the PET microchannel has a trapezoidal shape due to the nature of the photoablation process.

For MS coupling, the microspray is placed in a Plexiglas holder and positioned in front of the mass spectrometer (Figure 5).

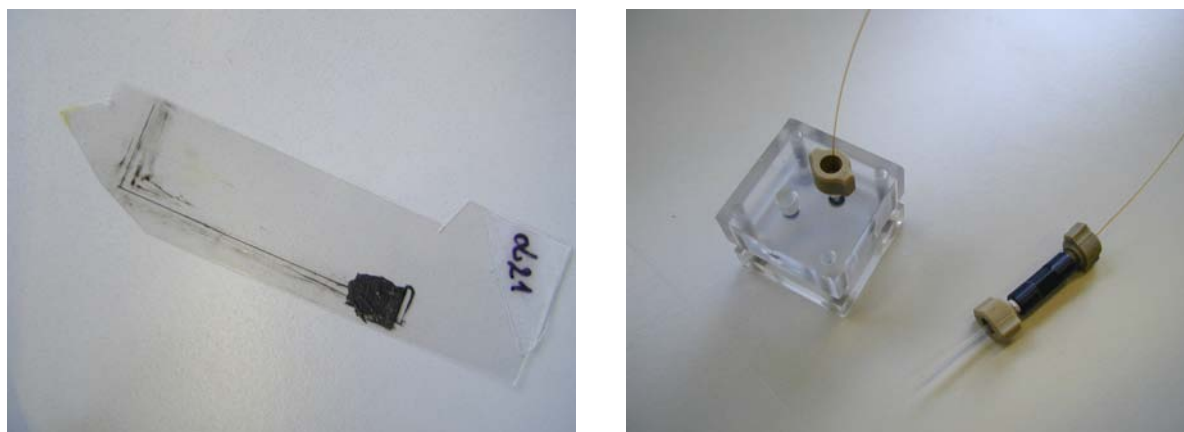
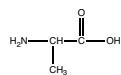
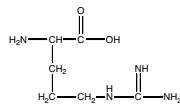
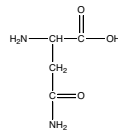
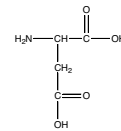
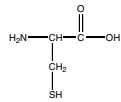
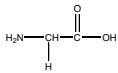
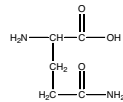
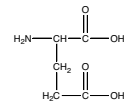
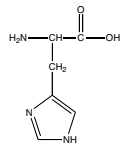
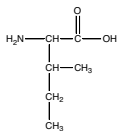
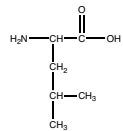
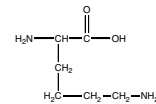
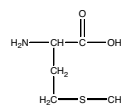
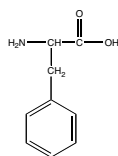
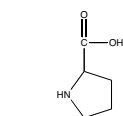
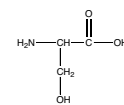
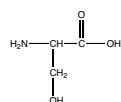
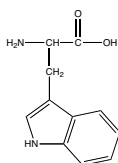
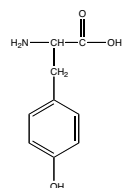
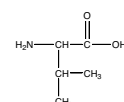


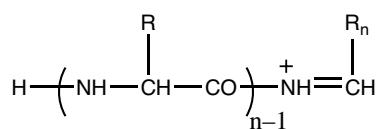
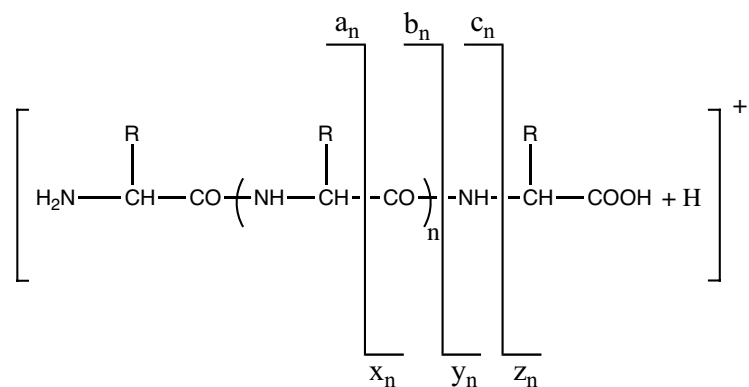
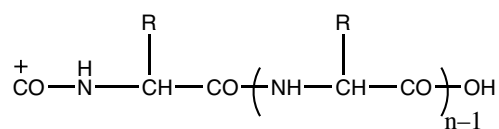
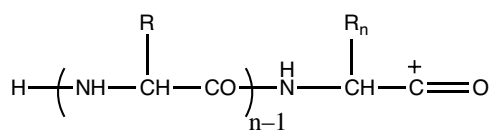
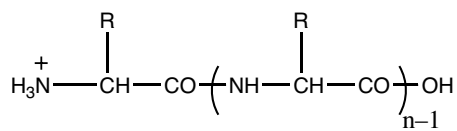
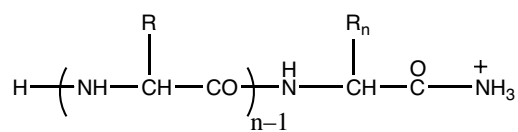
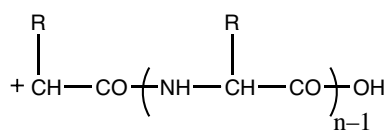
Figure 5. Photographs of the microspray emitter (a) and the holder used for pressure-driven flow infusion (b).

⁹ J.S. Rossier, M.A. Roberts, R. Ferrigno and H.H. Girault, *Anal. Chem.*, 1999, 71, 4294-4299.

¹⁰ J.S. Rossier, P. Bercier, A. Schwarz, S. Loidant and H.H. Girault, *Langmuir*, 1999, 15, 5173-5178.

APPENDIX II. *Amino acids*Alanine (Ala or **A**)Arginine (Arg or **R**)Asparagine (Asn or **N**)Aspartic acid (Asp or **D**)Cysteine (Cys or **C**)Glycine (Gly or **G**)Glutamine (Gln or **Q**)Glutanic acid (Glu or **E**)Histidine (His or **H**)Isoleucine (Ile or **I**)Leucine (Leu or **L**)Lysine (Lys or **K**)Methionine (Met or **M**)Phenylalanine (Phe or **F**)Proline (Pro or **P**)Serine (Ser or **S**)Threonine (Thr or **T**)Tryptophan (Trp or **W**)Tyrosine (Tyr or **Y**)Valine (Val or **V**)

**APPENDIX III. Nomenclature for sequence ions in
tandem mass spectra of peptides**

 a_n  X_n  b_n  Y_n  c_n  Z_n

APPENDIX IV. *List of peptides used*

Peptide sequence		Name
ACKCTCM	H-Ala-Cys-Lys-Cys-Thr-Cys-Met-OH	
AIKCTKF	H-Ala-Ile-Lys-Cys-Thr-Lys-Phe-OH	
ALRCTCS	H-Ala-Leu-Arg-Cys-Thr-Cys-Ser-OH	
DRVYIHPFHL	H-Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-OH	Angiotensin I
KCTCCA	H-Lys-Cys-Thr-Cys-Cys-Ala-OH	
RPPGFSPFR	H-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg-OH	Bradykinin
TCDDPRFQDSS	H-Thr-Cys-Asp-Asp-Pro-Arg-Phe-Gln-Asp-Ser-Ser-OH	Chorionic gonadotropin- β (109-119)
YGGFL	H-Tyr-Gly-Gly-Phe-Leu-OH	Leu-enkephalin
YGGFMRF	H-Tyr-Gly-Gly-Phe-Met-Arg-Phe-OH	Met-enkephalin-Arg-Phe

CURRICULUM VITAE

Personal

- Name: **Loïc DAYON**
- Citizenship: French
- Date of birth: 16th April 1978

Academic

2003-2006 **Ecole Polytechnique Fédérale de Lausanne** (Switzerland)
Preparation of the PhD ès Sciences

2002 **Ecole Nationale Supérieure de Chimie de Mulhouse** (France)
'Engineer' Diploma in Chemistry

2002 **Université de Haute-Alsace** (Mulhouse, France)
Master Degree in Chemistry

1996 **Lycée Condorcet** (Paris 9^{ème}, France)
'A' Level in Math, Physics, Chemistry and Language

Professional

2002 (6 months) **Actelion Pharmaceuticals** Allschwil, Switzerland
Research placement: synthesis of aspartic protease inhibitors

2001 (5 months) **Roche Vitamines France** Village-Neuf, France
Engineer placement: optimization of cantaxanthine and riboflavine-5'-phosphate synthesis

2000 (6 months) **F. Hoffmann-La Roche** Basel, Switzerland
Technician placement: quality control and insurance of pharmaceutical products

Languages

French (mother tongue)

English (very good level; TOIEC with 810/1000)

German

Personal interest

Swimming, ski, surf, baseball, basketball, running, ballroom dancing, music, cinema

Publications

L. Dayon, C. Roussel, H.H. Girault, **J. Proteome Res.**, 2006, 5, 793-800.

L. Dayon, J. Josserand, H.H. Girault, **Phys. Chem. Chem. Phys.**, 2005, 7, 4054-4060.

L. Dayon, C. Roussel, M. Prudent, N. Lion, H. H. Girault, **Electrophoresis**, 2005, 26, 238-247.

C. Roussel, L. Dayon, N. Lion, T. C. Rohner, J. Josserand, J. S. Rossier, H. Jensen, H. H. Girault, **J. Am. Soc. Mass Spectrom.**, 2004, 15, 1767-1779.

C. Roussel, L. Dayon, H. Jensen, H. H. Girault, **J. Electroanal. Chem.**, 2004, 570, 187-199.

L. Dayon, C. Roussel, H. H. Girault, **Chimia**, 2004, 58, 204-207.

N. Lion, T. C. Rohner, L. Dayon, I. L. Arnaud, E. Damoc, N. Youhnovski, Z. Y. Wu, C. Roussel, J. Josserand, H. Jensen, J. S. Rossier, M. Przybylski, H. H. Girault, **Electrophoresis**, 2003, 24, 3533-3562.

Oral presentation

Modification en ligne des cystéines libres par utilisation d'une source nanospray en spectrométrie de masse, *7^{èmes} Journées Francophones des Jeunes Physico-Chimistes*, Monastir, Tunisia, March 2004.

Poster presentations

Pittcon 2006, Orlando, USA, March 2006.

Congrès 2005 de la Société Suisse de Chimie, Lausanne, Switzerland, October 2005.

1^{er} Symposium de Chimie et Biologie Analytiques, Montpellier, France, September 2005.

2003 Congress of the Swiss Proteomics Society, Basel, Switzerland, December 2003.

Congrès 2003 de la Société Suisse de Chimie, Lausanne, Switzerland, October 2003.

20^{ième} journées Française de Spectrométrie de Masse, Toulouse, France, September 2003.

Awards

Best poster presentation Analytical Chemistry section at the **Swiss Chemical Society (Société Suisse de Chimie)** in 2003.

