

Development of multidimensional liquid chromatographic
methods hyphenated to mass spectrometry, preparation
and analysis of complex biological samples

Dissertation

zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie und Werkstoffwissenschaften
der Universität des Saarlandes

von

Nathanaël DELMOTTE

Saarbrücken

2007

Tag des Kolloquiums : 12. Juli 2007

Dekan : Prof. Dr. Uli Müller

Mitglieder des

Prüfungsausschusses : Prof. Dr. Christian Huber
Dr. Alain Van Dorsselaer
Prof. Dr. Rolf Müller
Prof. Dr. Laurence Sabatier
Prof. Dr. Marie-Claire Hennion
Dr. Andreas Tholey

Développement de méthodes chromatographiques liquides multidimensionnelles couplées à la spectrométrie de masse, préparation et analyse d'échantillons biologiques complexes

Development of multidimensional liquid chromatographic methods hyphenated to mass spectrometry, preparation and analysis of complex biological samples

Thèse

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR

DE STRASBOURG

par

Nathanaël DELMOTTE

Soutenue le 12 juillet 2007 devant la commission d'examen :

Dr. Alain VAN DORSSELAER

Directeur de thèse

Prof. Christian HUBER

Co-directeur de thèse

Prof. Marie-Claire HENNION

Rapporteur externe

Prof. Rolf MÜLLER

Rapporteur externe

Prof. Laurence SABATIER

Rapporteur interne

Dr. Andreas THOLEY

Examineur

This doctoral thesis was performed under the joint supervision of Prof. Dr. Christian Huber and Dr. Alain Van Dorsselaer.

Ce travail de thèse a été réalisé dans le cadre d'une co-tutelle sous la direction du Prof. Dr. Christian Huber et du Dr. Alain Van Dorsselaer.

Die vorliegende Dissertation wurde unter der Leitung von Prof. Dr. Christian Huber und Dr. Alain Van Dorsselaer durchgeführt.

Prof. Dr. Christian HUBER

Instrumentelle Analytik und Bioanalytik

Universität des Saarlandes

Postfach 15 11 50

D-66041 Saarbruecken

Tel: +49 (0) 681 302 3433 Fax: 49 (0) 681 302 2963 E-mail: christian.huber@mx.uni-saarland.de

Dr. Alain VAN DORSELAER

Institut Pluridisciplinaire Hubert Curien

Sciences Analytiques et Interactions Ioniques et Biomoléculaires

Laboratoire de Spectrométrie de Masse Bio-Organique

Université Louis Pasteur

ECPM, 25 Rue Becquerel, F- 67087 Strasbourg Cedex 2

Tel: +33 (0) 3 90 24 27 83 Fax: +33 (0) 3 90 24 27 81 E-mail: vandors@chimie.u-strasbg.fr

Acknowledgements

First of all, I would like to thank my both supervisors, **Prof. Dr. Christian Huber** and **Dr. Alain Van Dorsseleer** for giving me the opportunity to graduate in their lab. I am grateful for the guidance and support. Herzlichen Dank!

Prof. Dr. Laurence Sabatier, **Prof Dr. Marie-Claire Hennion**, **Prof Dr. Rolf Müller**, and **PD Dr. Andreas Tholey** are acknowledged for accepting to spend some of their precious time to correct and evaluate this work.

I also want to thank **Prof. Dr. Dr. h.c. Heinz Engelhardt**, **PD Dr. Frank Steiner**, and **Dr. Markus Martin** for their extremely valuable answers to any of my questions.

I am grateful to **Dr. Uwe Kobold**, **Dr. Thomas Meier**, **Dr. Andreas Gallusser**, **Dr. Albert Geiger**, **Thomas Dülffer**, **Sabine Kerkenbusch**, and **Thomas Weidner** from **Roche Diagnostics** for the financial support as well as for their assistance, input, and advice.

My sincere thanks go to **Reiner Wintringer**, “Windy”. Without him, the instruments would not be running so good, and the everyday life in the lab would not be so peaceful!

I kindly thank **Christa Göllen**, **Gabriele Krug**, and **Véronique Trimbour** for their help in the intricacies of the German and French red tape!

I warmly thank my first colleagues **Dr. Bettina Mayr**, **Dr. Hansjörg Toll**, and **Dr. Christian Schley** for their good advice, shrewdness, and tricks.

My very special thanks go to **Maria Lasaosa** for the very pleasant (and successful!) team-working, and **Prof. Dr. Elmar Heinzle** for permitting this cooperation.

I also want to thank my “Hiwi-Team”: **Silke Ruzek**, **Volker Neu**, and **Thomas Jakoby**, as well as **Benny Kneissl** for the useful bioinformatics software.

Of course (!) I also thank **Verena Fraaß**, **Andreas Leinenbach**, **Katja Melchior**, and **Jens Mohr**. I cannot forget **Dr. Anis Mahsunah**, **Eva Luxenburger**, **Iris Gostomski**, **Manuela Hügel**, **Rainer Geiss**, **Devid Hero**, **Bilgin Vatansever**, **Dr. Leena Suntornsuk**, **Sascha Quinten**, **Patrick Riefer**, and **Nathalie Selevsek**. I am grateful for all their encouragements, futile discussions, and jokes; I really enjoyed the friendly work atmosphere!

Many thanks go to **Dr. Christine Schaeffer** and **Dr. Jean-Marc Strub** for their help and assistance during my stays in Strasburg.

Thanks to **Hans-Peter Skohoutil**, **Norbert Ochs**, **Jens Wiegert**, and **Robin Adolph** for the manufacturing and supply of many un-purchasable hardware!

My thanks go to **Dionex**, **LCPackings** and **Bruker Daltonics** for the instruments and to **Dr. Aleš Štrancar** from **BIA Separations** for the CIM disks.

Last but not least, my hearty thanks go to the surrounding of my **family** and **friends** for their joyful and constant support!

Table of contents

Acknowledgements	7
Table of contents	9
Table of abbreviations, acronyms, and symbols	15
Preliminaries	17
English version.....	18
Version française	19
Résumé substantiel	20
Deutsche Version.....	22
Chapter I: Goal of the thesis	23
Chapter II: Theoretical part	27
1 Biological samples	28
1.1 Biological fluids.....	28
1.2 Cells and protein cell extracts	30
2 Structure and properties of proteins	32
2.1 Amino acids, peptides and proteins	32
2.2 Roles of proteins.....	36
2.3 Proteome	37
3 High-performance liquid chromatography for the separation of biomolecules	40
3.1 Reversed-phase and ion-pair reversed-phase HPLC	42
3.2 Ion-exchange HPLC	44
3.3 Affinity chromatography.....	46
3.4 Size-exclusion chromatography	47
3.5 Restricted access materials	48
4 Methods of detection	50
4.1 Immunodetection methods	50
4.2 Mass spectrometry	51

4.2.1	Principle of electrospray ionization.....	51
4.2.2	Quadrupole mass analyzer	54
4.2.3	Quadrupole ion trap mass analyzer	55
4.3	Identification of peptides and proteins with mass spectrometry and algorithmic computation	58
Chapter III: Development of monolithic immuno-adsorbers for the isolation of biomarkers from human serum		61
1	Introduction	62
1.1	Aim of the work.....	62
1.2	Investigated biomarkers	64
1.2.1	Myoglobin as biomarker of heart infarct.....	64
1.2.2	NT-proBNP as biomarker of heart insufficiency	65
2	Materials and methods	66
2.1	Chemicals and samples	66
2.2	Analytical setups.....	67
2.2.1	Isolation of biomarkers with CIM disks.....	67
2.2.2	High-performance liquid chromatography-mass spectrometry	67
2.2.3	Detection and quantitation of NT-proBNP with immunoassays	69
2.3	Preparation of affinity CIM disks	70
2.3.1	Preparation of anti-myoglobin- and anti-NT-proBNP CIM disks via direct immobilization	70
2.3.2	Preparation of affinity CIM disks via streptavidin-biotin anchorage	71
3	Isolation of myoglobin from human serum by affinity chromatography.....	74
3.1	Isolation of myoglobin at high concentration	74
3.2	Isolation of myoglobin at low concentration	76
3.3	Conclusions	78
4	Isolation of NT-proBNP from human serum by affinity chromatography.....	79
4.1	Evaluation of the loadability of anti-NT-proBNP-CIM disk	79
4.2	Isolation of NT-proBNP from human serum at 125 fmol/ μ L.....	80
4.3	Isolation of NT-proBNP from human serum at 7.8 fmol/ μ L.....	82

4.4	Stability and bath-to-batch reproducibility of anti-NT-proBNP-CIM disks	84
4.5	Calibration curve with anti-NT-proBNP-CIM disk	86
4.6	Hyphenation of anti-NT-proBNP-CIM disk with mass spectrometry	88
4.6.1	Hyphenation with PS-DVB monolithic trap column	89
4.6.2	Hyphenation with Pepmap™ C18 trap column	90
4.6.3	Discussion	92
5	Conclusions	94

Chapter IV: Development of an on-line SPE-HPLC-ESI-MS method for the analysis of drugs in whole blood hemolysates..... **95**

1	Introduction	96
2	Materials and methods	98
2.1	Chemicals and instruments	98
2.2	Preparation of blood hemolysates	99
2.3	Columns	99
2.3.1	Modification of LiChrospher ADS material	99
2.3.2	Packing of columns	100
2.4	Analytical setups	100
2.4.1	One-dimensional HPLC setup	100
2.4.2	Two-dimensional HPLC setup	100
3	Choice of the stationary phase	102
3.1	Hemoglobin exclusion with physical barrier and uniform surface topochemistry: ChromSpher Biomatrix	104
3.2	Hemoglobin exclusion with physical barrier and dual surface topochemistry: LiChrospher ADS	106
3.3	Hemoglobin exclusion with chemical barrier and uniform surface topochemistry: Capcell Pak	108
3.4	Hemoglobin exclusion with chemical barrier and dual surface topochemistry: Bioptic AV-2, SPS, and Biotrap 500 MS	110
3.5	Conclusions and next steps	114
4	Modification of LiChrospher ADS with amino-dextran	116
5	Modification of LiChrospher ADS with poly-D-lysine and polyethyleneimine	119

6	Evaluation of LiChrospher ADS at different pH	122
7	Evaluation of Biotrap 500 MS at pH 10.7.....	127
7.1	Retention of hemoglobin on Biotrap 500 MS at pH 10.7	127
7.2	Extraction of analytes with Biotrap 500 MS at pH 10.7.....	128
8	Quantitation of tetracycline hydrochloride in human whole blood hemolysates.....	132
8.1	Limit of detection	132
8.2	Carry-over.....	134
8.3	Calibration curve.....	135
9	Conclusions.....	138
Chapter V: Development and evaluation of multidimensional HPLC-MSsystems for proteome analysis		139
1	Materials and methods	141
1.1	Chemicals.....	141
1.2	Preparation of tryptic digests.....	141
1.3	Analytical setups for the first separation steps.....	142
1.4	Second separation step and data acquisition	142
1.5	Data processing and evaluation.....	143
2	Experimental.....	145
2.1	A classical 2D-HPLC-MS setup: SCX x IP-RP-HPLC.....	145
2.1.1	Proteome analysis of <i>C. glutamicum</i> with SCX x IP-RP-HPLC	146
2.1.2	Fractionation repeatability after peptide separation with SCX.....	147
2.2	A new 2D-HPLC-MS setup: RP-HPLC x IP-RP-HPLC	147
2.2.1	Proteome analysis of <i>C. glutamicum</i> with RP x IP-RP.....	149
2.2.2	Fractionation repeatability after separation with RP-HPLC at pH 10.0..	149
3	Results and discussion	150
3.1	Peptide separation.....	150
3.1.1	Peptide separation with SCX.....	150
3.1.2	Peptide separation with RP-HPLC at pH 10.0	151
3.1.3	Peptide separation with IP-RP-HPLC at pH 2.1.....	152

3.2	Peptide distribution	153
3.2.1	Peptide distribution with SCX.....	153
3.2.2	Peptide distribution with RP-HPLC at pH 10.0.....	157
3.3	Proteome coverage	159
3.4	Identification confidence.....	160
3.5	Repeatability of the dimensions of separation	162
3.5.1	Repeatability of peptide separation and identification with IP-RP-HPLC-ESI-MS/MS at pH 2.1.....	162
3.5.2	Separation and fractionation repeatability with SCX.....	164
3.6	Dimension orthogonality.....	170
3.7	Complementarity of SCX x IP-RP- and RP x IP-RP-HPLC methods.....	173
3.8	Ability of 2D-HPLC setups to analyze proteomes	176
Chapter VI: References		179
Appendix		189

Table of abbreviations, acronyms, and symbols

\cap	set-theoretic intersection, “and”
U	set-theoretic union, “or”
\	set-theoretic difference, “without”
2D	two-dimensional
Å	Ångström, 0.1 nm
AC	affinity chromatography
ACN	acetonitrile
ADS	alkyl-diol-silica
ASCII	American Standard Code for Information Interchange
C18	octadecyl
CE	capillary electrophoresis
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
CID	collision-induced dissociation
CIM	convective interaction medium
Da	Dalton
DC	direct current
DMSO	dimethyl sulfoxide
EC	enzyme commission number
ECLIA	electrochemiluminescence immunoassay
EIC	extracted ion chromatogram
ELISA	enzyme-linked immunosorbent assay
ESI	electrospray ionization
e.g.	<i>exempli gratia</i> , “for example”
Fig.	figure
FT-ICR	Fourier transform ion cyclotron resonance mass spectrometer
g	g-force, 9.80665 m/s ²
GC	gas chromatography
GE	gel electrophoresis
GRAVY	grand average of hydropathy
HFBA	heptafluorobutyric acid
HPAC	high-performance affinity chromatography
HPLC	high-performance liquid chromatography
IEC	ion-exchange chromatography
i.d.	internal diameter
IP-RP-HPLC	ion-pair reversed-phase chromatography
LC-MS	liquid chromatography-mass spectrometry
MALDI	matrix-assisted laser desorption/ionization
MOWSE	molecular weight search
MRM	multiple reaction monitoring
MS	mass spectrometry
MS/MS	tandem mass spectrometry
M _w	molecular weight
n.a.	not available
NT-proBNP	N-terminal prohormone brain natriuretic peptide
PFF	Peptide Fragment Fingerprinting
pH	potential of hydrogen

pl	isoelectric point
PMF	Peptide Mass Fingerprinting
ppm	parts per million
PS-DVB	poly(styrene-co-divinylbenzene)
RAM	restricted access material
RF	radio frequency
RP-HPLC	reversed-phase chromatography
rpm	revolutions per minute
RTICC	reconstructed total ion current chromatogram
SCX	strong cation exchange
SEC	size-exclusion chromatography
SIM	selected ion monitoring
SPE	solid phase extraction
Tab.	table
TEA	triethylamine
TEAA	triethylammonium acetate
TEAB	triethylammonium bicarbonate
TFA	trifluoroacetic acid
TOF-MS	time-of-flight mass spectrometer
Tris	trishydroxymethylaminomethane
UV	ultraviolet
VIS	visible
vs.	<i>versus</i>

Preliminaries

Abstracts

Key-words

French and German translations

English version

Development of multidimensional liquid chromatographic methods hyphenated to mass spectrometry, preparation and analysis of complex biological samples.

Immunoabsorbers based on monolithic epoxy-activated CIM disks have been developed in order to target biomarkers of heart diseases. The developed immunoabsorbers permitted to selectively isolate myoglobin and NT-proBNP from human serum. Anti-NT-proBNP-CIM disks permitted a quantitative isolation of NT-proBNP at concentrations down to 750 amol/ μ L in serum ($R^2 = 0.998$).

Six different restricted access materials have been evaluated with respect to their ability to remove hemoglobin from hemolysates. Experiments at different pH revealed that the retention of hemoglobin can be drastically diminished at pH 10.7. Because of better chemical stability at high pH, the polymeric Biotrap 500 MS RAM column was optimized for the analysis of hemolysates. The setup permits to quantitatively extract antibiotics from whole blood hemolysates at biologically relevant concentrations (200 pg/ μ L), and without carry-over of hemoglobin.

A new 2D-HPLC-ESI-MS/MS setup for proteome analysis was developed. It consisted of a peptide separation by RP-HPLC at pH 10.0, followed by IP-RP-HPLC at pH 2.1. This new setup was compared with a classical SCX x IP-RP-HPLC setup. Separation repeatability is similar with both setups. The orthogonality between methods of separation is higher in the SCX x IP-RP-HPLC approach than in the RP x IP-RP-HPLC scheme. However, the better peptide distribution and separation efficiency achieved with the RP x IP-RP-HPLC setup permitted to identify significantly more peptides than with the classical SCX x IP-RP-HPLC setup. Both approaches are complementary and a combination of both setups permits to identify more peptides than replicate injections performed with a single setup.

Key-words:

Affinity chromatography, biomarker, CIM, complementarity, *Corynebacterium glutamicum*, hemoglobin, hemolysate, hyphenation, mass spectrometry, micro-HPLC, monolithic column, multidimensional liquid chromatography, NT-proBNP, orthogonality, proteome analysis, RAM, repeatability, serum, trap-column

Version française

Développement de méthodes chromatographiques liquides multidimensionnelles couplées à la spectrométrie de masse, préparation et analyse d'échantillons biologiques complexes.

Des immunoadsorbants ont été développés à partir de disques CIM monolithiques pour l'analyse de biomarqueurs impliqués dans des maladies cardio-vasculaires. Les colonnes développées ont permis d'isoler sélectivement la myoglobine et le NT-proBNP du sérum humain. Les colonnes anti-NT-proBNP ont permis l'isolation quantitative du NT-proBNP ($R^2 = 0,998$) à des concentrations jusqu'à 750 amol/ μ L de sérum.

Six matériaux à accès restreints ont été évalués en fonction de leur aptitude à exclure l'hémoglobine d'hémolysats sanguins. Des injections à différents pH ont montré que la rétention de l'hémoglobine est drastiquement restreinte à pH 10,7. En raison d'une bonne stabilité à pH basique, la colonne polymérique Biotrap 500 MS RAM a été retenue pour l'extraction d'antibiotiques d'hémolysats sanguins. Des extractions quantitatives d'analytes à faibles concentrations (200 pg/ μ L) ont été réalisées sans effet mémoire d'hémoglobine sur la colonne.

Un nouveau système 2D-HPLC-ESI-MS/MS pour l'analyse protéomique a été développé. Le système est composé d'une séparation par RP-HPLC à pH 10,0, suivie d'une séparation par IP-RP-HPLC à pH 2,1. Ce nouveau système a été comparé à un système conventionnel SCX x IP-RP-HPLC. L'orthogonalité des méthodes de séparation est plus élevée dans l'approche SCX x IP-RP-HPLC que dans le schéma RP x IP-RP-HPLC. Cependant, en raison d'une meilleure distribution des peptides et d'une meilleure efficacité de séparation, le système RP x IP-RP-HPLC permet d'identifier significativement plus de peptides. Les deux approches sont complémentaires et une combinaison des deux systèmes permet d'identifier plus de peptides que des analyses répétées par un système unique.

Mots-clefs:

Analyse protéomique, biomarqueur, chromatographie d'affinité, chromatographie liquide multidimensionnelle, CIM, colonne monolithique, colonne d'enrichissement, complémentarité, *Corynebacterium glutamicum*, couplage, hémoglobine, hémolysat, micro-HPLC, NT-proBNP, orthogonalité, RAM, répétabilité, sérum, spectrométrie de masse

Résumé substantiel

Des immunoadsorbenters ont été développés à partir de colonnes disques CIM monolithiques pour l'analyse de biomarqueurs impliqués dans des maladies cardio-vasculaires (myoglobine, NT-proBNP). Pour chacun des antigènes étudiés, des anticorps ont été greffés avec succès sur le disque polymérique. Les colonnes développées ont permis d'isoler sélectivement la myoglobine et le NT-proBNP du sérum humain. La myoglobine a été sélectivement isolée puis détectée dans des échantillons de sérum jusqu'à 250 fmol/ μ L. Cependant, la capacité des anticorps greffés n'est pas suffisante pour l'analyse d'échantillons cliniques. L'analyse frontale d'une colonne anti-NT-proBNP révèle l'aptitude de l'immunoadsorbent à lier jusqu'à 250 pmol de NT-proBNP, ce qui est suffisant pour l'analyse d'échantillons cliniques. De plus, les colonnes anti-NT-proBNP ont une très bonne stabilité temporelle (> 18 mois) et leur préparation présente une très grande reproductibilité inter lots. Les colonnes anti-NT-proBNP ont permis l'isolation quantitative du NT-proBNP ($R^2 = 0,998$) à des concentrations jusqu'à 750 amol/ μ L de sérum, ce qui correspond à des concentrations en NT-proBNP dans le sérum de patients gravement malades. Le couplage des immunoadsorbenters à un spectromètre de masse a été réalisé pour des concentrations jusqu'à 7,8 fmol/ μ L en implémentant une colonne d'enrichissement de type PepmapTM C18.

Six matériaux à accès restreints (RAM) ont été évalués en fonction de leur aptitude à exclure l'hémoglobine d'hémolysats sanguins. Globalement toutes les colonnes présentent les mêmes propriétés : à pH 2,1 une adsorption significative ainsi qu'un effet mémoire conséquent de l'hémoglobine dans des injections consécutives. La dérivatisation du matériau LiChrospher ADS avec des polymères neutres ou positivement chargés tels que l' amino-dextrane, la polyéthylèneimine et la polylysine ne permet pas de diminuer l'adsorption des protéines sur la colonne. Des injections à différents pH (2,1-10,7) ont montré que la rétention de l'hémoglobine est drastiquement restreinte à pH basique (10 mmol/L éthanolamine, pH 10,7). En raison de sa bonne stabilité à pH basique, la colonne polymérique Biotrap 500 MS RAM a été retenue pour l'extraction d'antibiotiques d'hémolysats sanguins. Des échantillons réels dopés à la tétracycline ont été analysés avec le système SPE-HPLC-ESI-MS développé. La détection a été assurée par un spectromètre de masse quadropôle linéaire fonctionnant en mode SIM. Des extractions quantitatives d'analytes à faibles concentrations (jusqu'à 200 pg/ μ L) ont été réalisées sans effet mémoire d'hémoglobine sur la colonne Biotrap 500 MS.

Un système chromatographique de type SCX x IP-RP-HPLC a été développé pour analyser des mélanges complexes de peptides. Des fractions ont été collectées après l'échangeur cationique et les peptides fractionnés ont été ensuite séparés par IP-RP-HPLC à pH 2,1 dans la seconde dimension chromatographique. Le nombre de charges positives portées par les peptides apparaît être un paramètre crucial pour la rétention des peptides sur la colonne échangeuse d'ions. L'addition d'acétonitrile dans les éluants permet de supprimer les interactions secondaires entre les peptides et la phase stationnaire (interactions hydrophobes). Une bonne répétabilité en terme d'identification des peptides est obtenue : 54,0 % des peptides sont identifiés au minimum trois fois lors de quintuples injections. L'analyse des échantillons en triplicata apparaît comme un bon compromis entre la quantité de peptides identifiés et le temps d'analyse. Un nouveau système 2D-HPLC-ESI-MS/MS pour l'analyse protéomique a été développé. Le système est composé d'une séparation par RP-HPLC à pH 10,0, suivie d'une séparation par IP-RP-HPLC à pH 2,1. Bien que de la triéthylamine a été utilisée pour obtenir le pH basique, aucun appariement d'ions n'est observé et la rétention des peptides est gouvernée par les interactions solvophobes. Comme la seconde dimension chromatographique est identique dans les deux systèmes, il a été possible de comparer entre-elle la première dimension de chacun des systèmes (SCX vs. RP-HPLC à pH basique). Des répétabilités similaires ont été observées. L'orthogonalité des méthodes de séparation est plus élevée dans l'approche SCX x IP-RP-HPLC que dans le schéma RP x IP-RP-HPLC. Cependant, en raison d'une meilleure distribution des peptides et d'une meilleure efficacité de séparation, le système RP x IP-RP-HPLC permet d'identifier significativement plus de peptides. Les deux approches sont complémentaires et une combinaison des deux systèmes permet d'identifier plus de peptides que des analyses répétées par un unique système. Aucun des deux systèmes n'est discriminatoire à l'égard du pI ou de la masse des protéines. Le nombre d'enzymes identifiées révèle la capacité des systèmes 2D-HPLC-ESI-MS/MS développés à analyser des protéomes complets tels que celui de *C. glutamicum*.

Deutsche Version

Entwicklung von mehrdimensionalen, flüssigkeitschromatographischen Methoden gekoppelt mit Massenspektrometrie, Vorbereitung und Analyse von komplexen biologischen Proben.

Immunoadsorber, die auf monolithischen Epoxy-aktivierten CIM-Disks basieren, wurden entwickelt, um Biomarker für Herzkrankheiten nachweisen zu können. Die entwickelten Immunoadsorber erlaubten eine selektive Isolierung von Myoglobin und NT-proBNP aus menschlichen Serum. Anti-NT-proBNP-CIM-Disks ermöglichten die quantitative Isolierung von NT-proBNP mit Konzentrationen bis zu 750 amol/ μ L im Serum ($R^2 = 0,998$).

Sechs verschiedene „Restricted Access Materials“ wurden im Hinblick auf ihre Fähigkeit, bzgl. der Entfernung von Hämoglobin aus Hämolysaten untersucht. Experimente bei unterschiedlichen pH-Werten ergaben, dass die Retention von Hämoglobin bei einem pH-Wert von 10,7 deutlich verkleinert werden kann. Aufgrund ihrer höheren chemischen Stabilität bei höheren pH-Werten wurde die polymere „Biotrap 500 MS RAM“ für die Analyse von Hämolysaten optimiert. Die Methode ermöglicht die quantitative Extraktion von Antibiotika aus Gesamtblut Hämolysaten mit biologisch relevanten Konzentrationen (200 pg/ μ L), ohne die Verschleppung des Hämoglobins.

Eine neue 2D-HPLC-ESI-MS/MS-Methode wurde für die Proteomanalyse entwickelt. Sie bestand aus einer Peptid-Trennung mittels RP-HPLC bei einem pH-Wert von 10,0 und anschließender IP-RP-HPLC bei einem pH-Wert von 2,1. Anschließend wurde diese neue Methode mit einer klassischen SCX x IP-RP-HPLC-Methode verglichen. Die Orthogonalität zwischen den beiden Trennmethode bei der SCX x IP-RP-HPLC ist hierbei höher als bei der entsprechenden RP x IP-RP-HPLC-Methode. Allerdings erlaubt die bessere Verteilung der Peptide und die bessere Trenneffizienz der SCX x IP-RP-HPLC-Methode die Identifizierung einer höheren Anzahl an Peptiden. Beide Methoden sind komplementär, und eine Kombination beider Methoden erlaubt die Identifizierung einer größeren Anzahl an Peptiden als wiederholte Injektionen bei einer eindimensionalen Methode.

Schlagwörter:

Affinitätschromatographie, Anreicherungssäule, Biomarker, CIM, *Corynebacterium glutamicum*, Hämoglobin, Hämolysat, Hochleistungs-Flüssigchromatographie, Komplementarität, Kopplung, Massenspektrometrie, mehrdimensionale Mikro-HPLC, monolithische Säule, NT-proBNP, Orthogonalität, Proteomanalyse, RAM, Serum, Wiederholbarkeit

Chapter I

Goal of the thesis

I. Goal of the thesis

The outstanding development of new analytical techniques in the last thirty years has revolutionized our way to handle and analyze biological molecules and macromolecules. Thus, the introduction of fused silica capillaries ^[1] in the late seventies permitted the development and the miniaturization of numerous separation techniques such as capillary gas chromatography (GC), high-performance liquid chromatography (HPLC) and capillary electrophoresis (CE). In the late eighties, advances in mass spectrometry (MS) and more particularly soft ionization methods permitted to analyze high-molecular ions such as proteins and peptides. Electrospray ionization (ESI) developed by J. Fenn ^[2] and matrix-assisted laser desorption/ionization (MALDI) by F. Hillenkamp, M. Karas and K. Tanaka ^[3,4] are now the most common ionization modes for bio-molecular compounds. The miniaturization of HPLC columns (75 – 200 μm i.d.) and the decrease of mobile phase flow rates (0.2 – 3 $\mu\text{L}/\text{min}$) facilitated the hyphenation of liquid chromatography with mass spectrometry. Finally, peptide sequencing based on gas-phase fragmentation was achieved with tandem mass spectrometry (MS/MS). Fragmentation rules established by K. Biemann ^[5] and new algorithms or search engines such as Mascot ^[6,7] and Sequest ^[8] finalized a fully automatic interpretation of huge MS/MS data sets.

Biological samples are extremely complex and no single dimensional analytical technique is able to handle them. The numerousness of the analytes is mostly responsible for such a high complexity. Thus, in case of biological fluids such as urine, serum or blood hemolysates, several thousands of components have already been reported. The high dynamic range (in terms of concentration) of the analytes also increases the difficulties of analysis. Dynamic ranges of $1\text{-}10^8$ or of $1\text{-}10^{10}$ are common ^[9]. As a matter of fact, a sample simplification is required to get reliable and reproducible identifications. This can be achieved by combining several separation steps ^[10-13] (e.g. 2D-HPLC) with a specific and sensitive detection technique ^[14-16] (e.g. MS/MS or ELISA).

The analysis of a biological sample can be required for various reasons. First, one aim of studies can lie in the detection and/or the quantitation of only a single or a handful of analytes. For instance, the analytes may be biomarkers in a diagnostics study. In case of heart infarct, only proteins specific to heart diseases may be

researched. In a therapeutic study, the analytes of interest may be antibiotics in blood or drug degradation products in urine. This first approach is characterized by fishing out few analytes of interest. On the other hand, some studies require the detection of as many as possible components present in the sample. Thus, the differential analysis of two cell states is performed by checking in two cell extracts the presence (or the absence) of all presumably expressed proteins. The development of this second approach also called holistic strategy (holistic, greek *holos*, “whole”) is only possible because of the latest technical advancements. Obviously, these two different analytical goals are antagonist and require different setups.

In this context, the aim of the thesis was to develop HPLC-MS methods allowing meaningful analyses of biological samples. The methods should be established with model compounds but evaluated on real biological samples (serum, whole blood, protein cell extract). They should also be able to respond to common bioanalytical challenges (biomarker detection, drug targeting, proteome analysis) in classical and holistic strategies and should be easily transposable to similar samples.

Chapter II

Theoretical part

II. Theoretical part

Medical diagnostics and therapeutical studies are mostly performed on biological materials such as blood and its components, tissue and tissue fluids, excreta and secreta. To avoid dissolution steps and to assure a high compatibility of the investigated samples with the chromatographic setups, only native fluid materials (such as human whole blood and human serum) or protein cell extracts were investigated. These biological materials are described in the following chapter. Structures and properties of the most relevant components of these biological fluids, namely peptides and proteins, are also presented. Finally, a theoretical description of the separation and detection methods employed in this work is given.

1 Biological samples

1.1 Biological fluids

Blood is the fluid consisting of plasma, blood cells, and platelets that is circulated by the heart through the vascular system, carrying oxygen and nutriments to and waste materials away from all body tissues. A healthy human adult has around 5 liters blood in his body, which represents around 7 % of its whole weight. Optically, oxygenated blood is bright red. This color is explained by the presence of oxygenated iron in the hemoglobin of red blood cells. Human hemoglobin is built up with four subunits, each subunit being a globular protein embedded with a heme group. The name hemoglobin, as a concatenation of heme and globin, reflects this particular structure. The heme itself consists of a protoporphyrin group and an iron atom, being responsible for the binding of oxygen. The structure of a heme is depicted in Fig. 1. The molecular weight of undissociated hemoglobin is about 68,000 Da (four 17,000 Da subunits).

The lysis of red blood cells leads to the release of hemoglobin into the surrounding fluid. The name hemolysate refers to a whole blood sample in which the red blood cells have been lysed. Thus, hemolysates differ from whole blood only by the free circulation of hemoglobin (and other species originally contained in red blood cells) in the liquid phase of the medium.

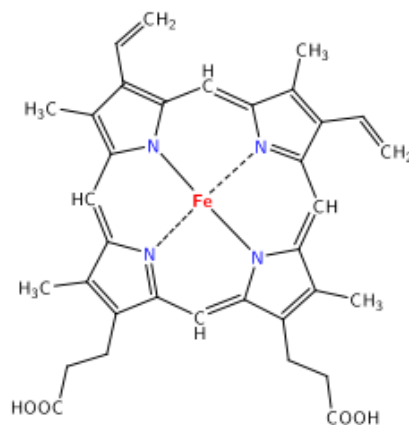


Fig. 1. Structure of a heme, consisting of a protoporphyrin group and an iron atom.

Plasma can be defined as the liquid phase of blood and consists of the residual fluid obtained after blood clotting (approximately 55 % of blood volume). Plasma consists of water (91 %) and proteins (7 %) but also contains salts, glucids, lipids, vitamins, and hormones. The total protein concentration is usually evaluated to 60 to 85 g/L in healthy adult plasma ^[17]. Ten plasma proteins represent about 90 % of this value as depicted in Fig. 2. Albumin (up to 60 %) and immunoglobulins (about 35 %) are the major plasma proteins whereas fibrinogen (about 5 %) is responsible of blood coagulation. The removal of fibrinogen permits to obtain serum. Thus, serum differs from plasma only by the absence of fibrinogen. The clear yellowish color of serum is explained by the formation of bilirubin during the degradation process of hemoglobin. The physical and chemical properties of blood and its derivatives are highly influenced by various salts dissolved at different concentrations. Sodium chloride represents by itself 75 % of the total salt amount. Buffer properties of bicarbonate are responsible for the slightly basic pH of blood (7.4). Potassium and calcium are not directly responsible for blood stability but permit to maintain vital functions such as nerve excitability and muscle contractions.

Blood and blood derivatives are very sensitive to coagulation and protein precipitation. To avoid such phenomena some sample handling procedures must be followed. For instance, freeze-thaw cycles should be avoided and samples should only be diluted in physiological buffers or salt solutions.

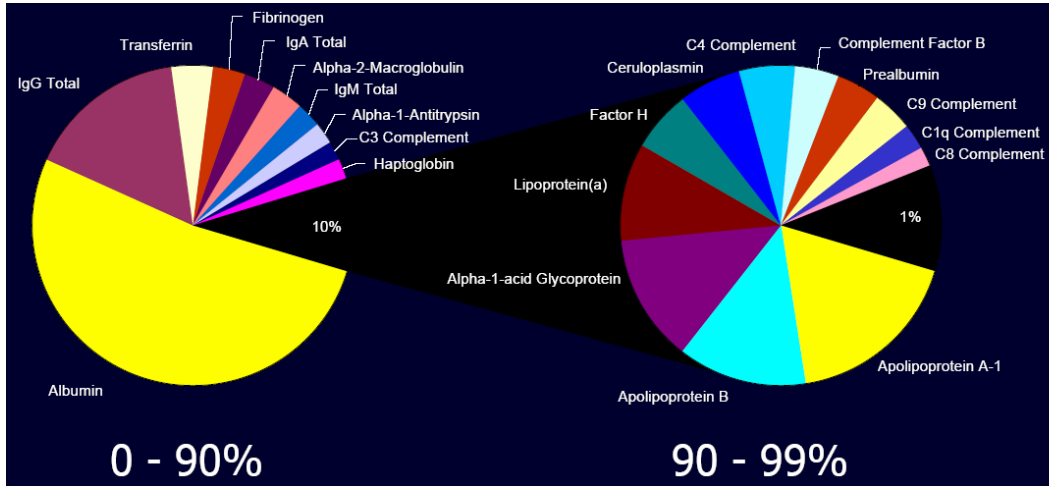


Fig. 2. The major proteins in human plasma, classified as a function of their concentration.

1.2 Cells and protein cell extracts

The cell (latin *cellula*, “small room”) is the structural and functional unit of all living organisms. A cell is an own entity able to take in nutrients, to convert these nutrients into energy, and finally to carry out specialized functions. Cells contain their own hereditary information and are consequently able to reproduce themselves. Some organisms, such as bacteria, consist of a single cell, whereas other organisms are made of multiple cells. For instance, the human body consists of 220 different types of cells and biological tissues. In the case of multicellular organisms, the different cells are differentiated and specialized for one particular task. Thus bone marrow cells produce blood cells, while muscle cells are responsible for the production of muscle tissue.

Two kinds of cells can be distinguished: prokaryotic cells and eukaryotic cells. The major difference resides in the lack of a cell nucleus in prokaryotic cells, whereas eukaryotic cells contain a membrane-delineated compartment that houses the genetic information. Prokaryotic cells are mostly singletons, whereas eukaryotic cells are usually found in multi-cellular organisms.

For analytic purposes, cells are often lysed. Cell lysis refers to the burst of a cell and to the resulting release of its content. Depending on the kind of cells, lyses are achieved in different manners. However, all mechanisms are leading to the disruption of the cellular membrane. Few of them are listed bellow and quickly described ^[18]:

- cells are placed in a hypotonic environment (e.g. deionized water). Water osmotically diffuses into the cells and the cells burst. This procedure can be enzymatically improved (e.g. with lysozyme)
- by repetitive freeze-thaw cycles, cellular membranes are distorted until disruption
- small cell layers are dessicated at 20-30 °C and finally ground in a mortar
- by immersing cells in a cold and water miscible organic solvent, lipids of the cellular membranes are removed and the cells burst
- with a vibration cell/mixer mill, cells are placed in a steel container with glass beads and violently agitated
- under rapid variations of pressure produced by ultrasonic waves, cells are rapidly destroyed
- cells are pressed under pressure until membrane disruption (French press)

Protein cell extracts correspond to the proteins that are released during cell lysis. Depending on the protocol used to perform protein collection, protein cell extracts may contain only a part or the whole proteins originally present in the cells. For instance, membrane proteins, which are particularly difficult to be extracted, may not be present in protein cell extracts.

2 Structure and properties of proteins

2.1 Amino acids, peptides and proteins

Proteins are macromolecules constituted of an enchainment of building blocks called amino-acids. All amino acids are made up of one central carbon atom (α) with a tetrahedral configuration. The four groups bound to this α -carbon atom consist of an amine group, a carboxyl group, a hydrogen atom, and a so-called side chain. Amino acids differentiate themselves in the structure of the side chains. Because of differences in terms of chemical composition, length, and functional groups, the side chains are responsible for the various chemical properties of the amino acids. One distinguishes amino acids with aliphatic, aromatic, and heterocyclic side chains but also with hydroxyl-, sulfhydryl-, carboxyl-, and amido groups. Twenty amino acids are occurring in natural proteins. One-letter- and three-letter symbols have been introduced to refer to these proteogenic amino acids ^[19]. They are listed in Tab. 1. Except for glycine, the α -carbon atom is an asymmetric center. One can thus distinguish D- and L- isomers. Only L- isomers are present in the human body.

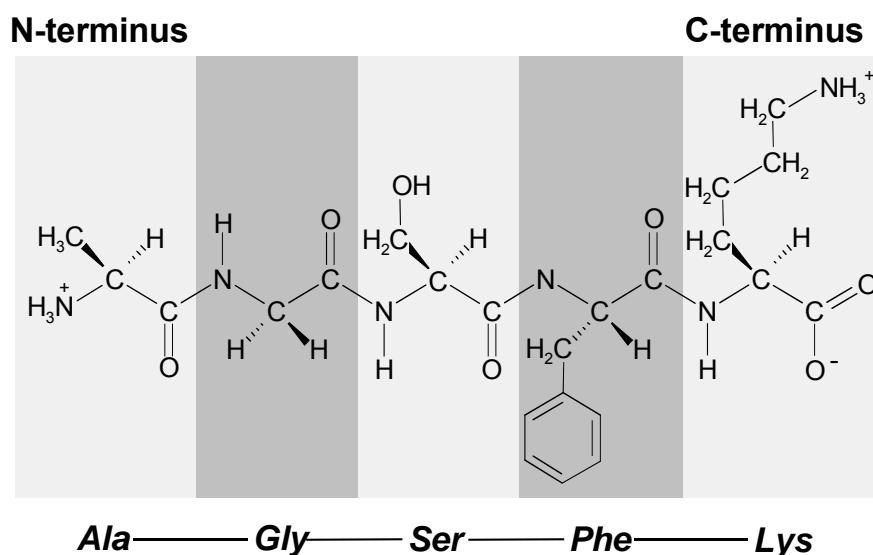
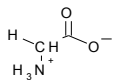
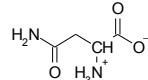
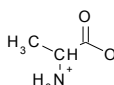
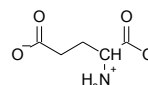
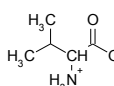
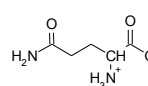
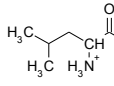
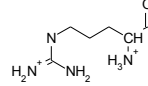
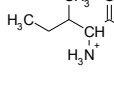
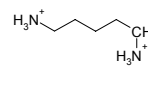
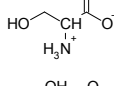
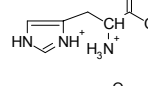
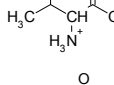
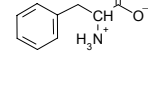
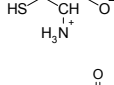
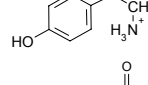
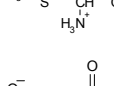
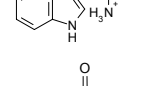
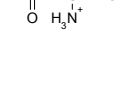
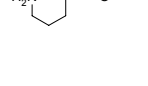


Fig. 3. Amino-acid sequence of the pentapeptide AGSFK ^[20].

Two consecutive amino acids are bound together by coupling the α -carboxyl group of the first amino acid with the α -amino group of the second amino acid. The resulting amide bond is often referred as peptide bond. During the formation of the bond, water is removed, and what remains of each amino acid is called an amino-acid residue. A five amino-acid sequence is depicted in Fig. 3. The term oligopeptides is

employed to refer to amino-acid enchainments up to 20 residues; those with more are called polypeptides. Over 50 residues, the term protein is generally used.

Tab. 1. Name, 3-letter symbol, 1-letter symbol and structure of the 20 proteinogenic amino acids.

glycine	Gly (G)		asparagine	Asn (N)	
alanine	Ala (A)		glutamic acid	Glu (E)	
valine	Val (V)		glutamine	Gln (Q)	
leucine	Leu (L)		arginine	Arg (R)	
isoleucine	Ile (I)		lysine	Lys (K)	
serine	Ser (S)		histidine	His (H)	
threonine	Thr (T)		phenylalanine	Phe (F)	
cysteine	Cys (C)		tyrosine	Tyr (Y)	
methionine	Met (M)		tryptophan	Trp (W)	
aspartic acid	Asp (D)		proline	Pro (P)	

The peptide bond is characterized by high kinetic stability (resistance to hydrolysis) and also by double bond properties. This double bond character results in a planar arrangement and six atoms are present in the same plane. They consist of the α -carbon atom and the α -carboxyl group of the first amino acid, and also of the α -amino group and the α -carbon atoms of the second amino acid. On the opposite, the bonds between the α -carbon atoms and the amino- and carboxyl groups are single bonds, resulting in free rotation around their axes. These free rotations permit the folding of peptides and proteins in different manners. Generally four structure levels are differentiated: the primary-, secondary-, tertiary-, and quaternary structures.

The primary structure of a protein, also called sequence, corresponds to the order in which the different amino acids are assembled. The amino-acid sequence is by convention written from the amino terminus (N-terminus) to the carboxyl terminus (C-terminus) ^[19].

The secondary structure corresponds to the arrangement in space of the amino-acid chain. The first elements of secondary structure, the alpha helix and the beta sheet were suggested in 1951 by Linus Pauling ^[21-23]. The alpha helix structures are stabilized by hydrogen bonds between amino- and carboxyl groups of the protein backbone. The side chains are directed to the outside and can interact with each other or with the medium. Angles between peptide bonds are 80°. The helix is made of around 3.7 amino-acid residues per spiral turn. By stretching an alpha helix a beta sheet can be obtained. In the latter, two anti-parallel chains are observed. All the atoms of the peptide bond are present in the same plane, but the α -carbon atoms are simultaneously located in two different planes. The periodicity is about 7 Å. Schematic representations of alpha-helix and beta-sheet structures are represented in Fig. 4. Beta-elbow structures are also found. They consist of four consecutive amino acids for which an hydrogen bond between the carboxyl group of the first amino acid and the amino group of the fourth amino acid is formed. Random coils are also observed. These structures are non-periodic structures but the valence angles are still observed.

The tertiary structure corresponds to the spatial arrangement of the different secondary structures. This three-dimensional structure is specific to each protein and is maintained with chemically different bonds. Disulfide bonds, established between two cysteine residues are particularly strong because of their covalent nature. The tertiary structure of a protein depends on the surrounding medium. Parameters such as solvent, ionic strength, viscosity and concentration have influence on the conformational structure of the protein. For instance, in aqueous medium non-polar amino acids are tending to avoid water, whereas polar amino acids are presented on the outside. The presence of prosthetic groups is also influencing the three-dimensional structure of proteins. A prosthetic group is a molecule bond to a protein, but which is not constituted of amino acids (e.g. the heme group of hemoglobin). The complexity of proteins is also increased by post-translational modifications such as acetylation, phosphorylation, and glycosylation.

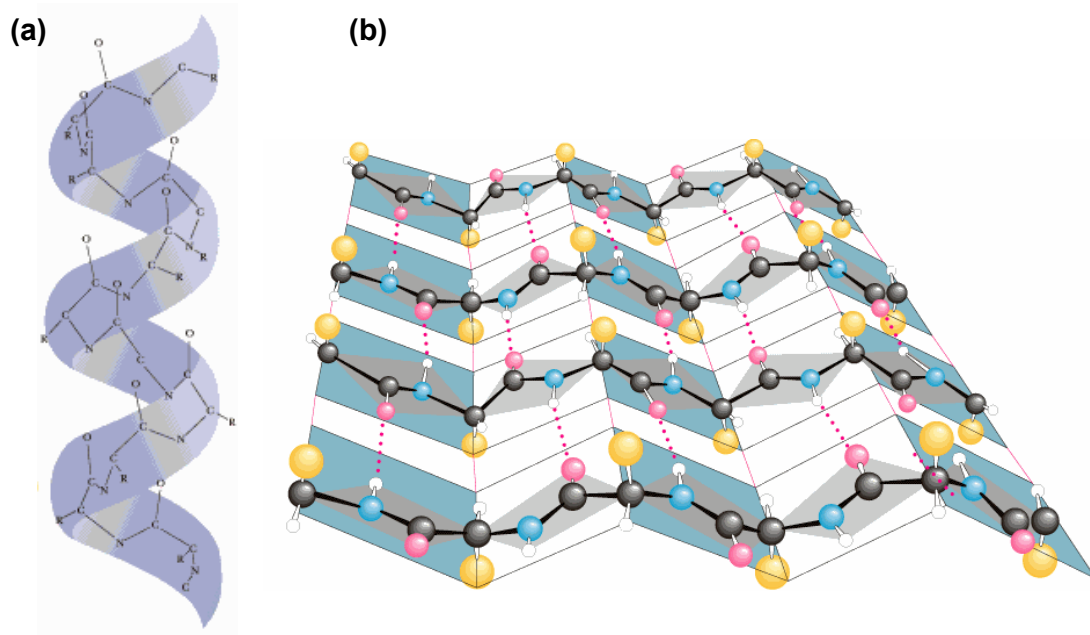


Fig. 4. Schematic representation of (a) alpha-helix- and (b) beta-sheet structures.

Several protein subunits with defined tertiary structures are able to interact together to form a multi-protein complex, which is usually called quaternary structure. Generally, non covalent bonds are maintaining such structures.

Fibrous proteins and globular proteins are often distinguished. Fibrous proteins are often structural proteins (e. g. collagen and fibrin) with folded or spiral chains, whereas globular proteins hold their name from their spherical or elliptic structure. Fibrous proteins are generally non soluble in water. On the contrary, globular proteins (e.g. albumin) are in many cases water soluble.

Because of their amino-acid constitution, peptides and proteins are amphoteric molecules. They possess both acidic and basic groups. Thus the charge of a protein depends not only on the number of acidic or basic groups present in the sequence, but also on the pH of the solution in which the protein is present. The pH at which a protein carries no net electrical charge is named isoelectric point (pI).

The hydrophilicity/hydrophobicity of peptides is generally evaluated by computing the so-called grand average of hydropathy (GRAVY). The GRAVY value for a peptide or a protein is calculated as the sum of the hydropathy values of all the amino acids, divided by the number of residues in the sequence. The term hydropathy (strong feeling about water), and the hydropathy scale were first introduced in 1982 by J. Kyte and R. Doolittle^[24]. The hydropathy scale is reproduced in Tab. 2.

Tab. 2. Hydropathy scale according to J. Kyte and R. Doolittle ^[24].

side-chain	hydropathy index	side-chain	hydropathy index
isoleucine	4.5	serine	-0.8
valine	4.2	tyrosine	-1.3
leucine	3.8	proline	-1.6
phenylalanine	2.8	histidine	-3.2
cysteine	2.5	glutamic acid	-3.5
methionine	1.9	glutamine	-3.5
alanine	1.8	aspartic acid	-3.5
glycine	-0.4	asparagine	-3.5
threonine	-0.7	lysine	-3.9
tryptophan	-0.9	arginine	-4.5

2.2 Roles of proteins

Proteins are playing important roles in life processes. For instance some proteins work as biocatalyzers (enzymes) to synthesize and construct cell elements. Proteins and particularly structural proteins are responsible together with nucleic acids for more than 2/3 of the anhydrous mass of a cell.

They are present in membranes, tendons, muscles, and connective tissues. As regulation substances, proteins are essential during reaction processes and cell differentiation. Numerous peptides and proteins are involved in the control of genes and during neurobiological processes. In vertebrate species, immunoglobulins are assuring the body defense by identifying and neutralizing foreign objects such as bacteria or viruses. The antibodies, with their Y-shaped structure, are essential in therapeutic methods and diagnostics. Each antibody possesses a paratope specific for one particular epitope of an antigen. These lock-key matching properties permit antigen and antibody to specifically bind together. Other proteins are responsible for the transport and the storage of non soluble substances, metal atoms or oxygen. Some proteins may be disease agents and toxic for other organisms (toxins).

Plasma proteins have numerous functions. All together they contribute to the colloid osmotic pressure of the blood. The most important proteins in human plasma are depicted in Fig. 2. Proteins present in human serum can be sorted in the following eight design/function groups ^[9]:

1. proteins secreted by solid tissues and which act in plasma. They are also called „classical plasma proteins“. They are bigger than the kidney filtration cut-off (~ 45 kDa) and more than 50,000 different molecular forms are present in plasma
2. immunoglobulins. Because of their complexity, immunoglobulins form a unique class of proteins. Around 10 millions of different sequences are present in the blood of a human adult
3. long-distance receptor ligands. In this group classical peptide- and protein hormones are classified (e.g. insulin and erythropoietin)
4. local receptor ligands. They are generally small proteins responsible for interactions between neighbor cells. Concentrations of local receptor ligands in plasma are very low (some tens of pg/mL)
5. temporary passengers. They correspond to non-hormone proteins which are transported in plasma from their secretion site to their primary function site
6. tissue leakage products. The proteins are present in blood because of cell damage or death. Theoretically all proteins present in an organism can leak into plasma. In the case of human plasma up to 50,000 proteins may be present. In this group, the most important diagnostics markers are classified, e.g. troponin, Brain Natriuretic Peptide (BNP), and myoglobin used to diagnose cardiac disease
7. aberrant secretions. These proteins are released from tumors or diseased tissues
8. foreign proteins. They do not originate from the organism itself but from infectious organisms or parasites

Human plasma possesses the largest and the most complex set of proteins (proteome) present in the human body.

2.3 Proteome

The term “proteome” was first coined by Mark Wilkins in 1995 ^[25]. A proteome refers to the total set of proteins expressed in a given cell (or tissue) at a given time (under well defined conditions). The proteome is larger than the genome, because one gene is responsible for the expression of more than one protein. Thus for human beings 20,000 to 25,000 genes are leading to the possible expression of about 500,000

proteins ^[26]. Because of their various functions, proteins are present at different concentrations in cells or body fluids. Their dynamic range is extremely wide. Dynamic ranges of $1-10^8$ or of $1-10^{10}$ are usual ^[9].

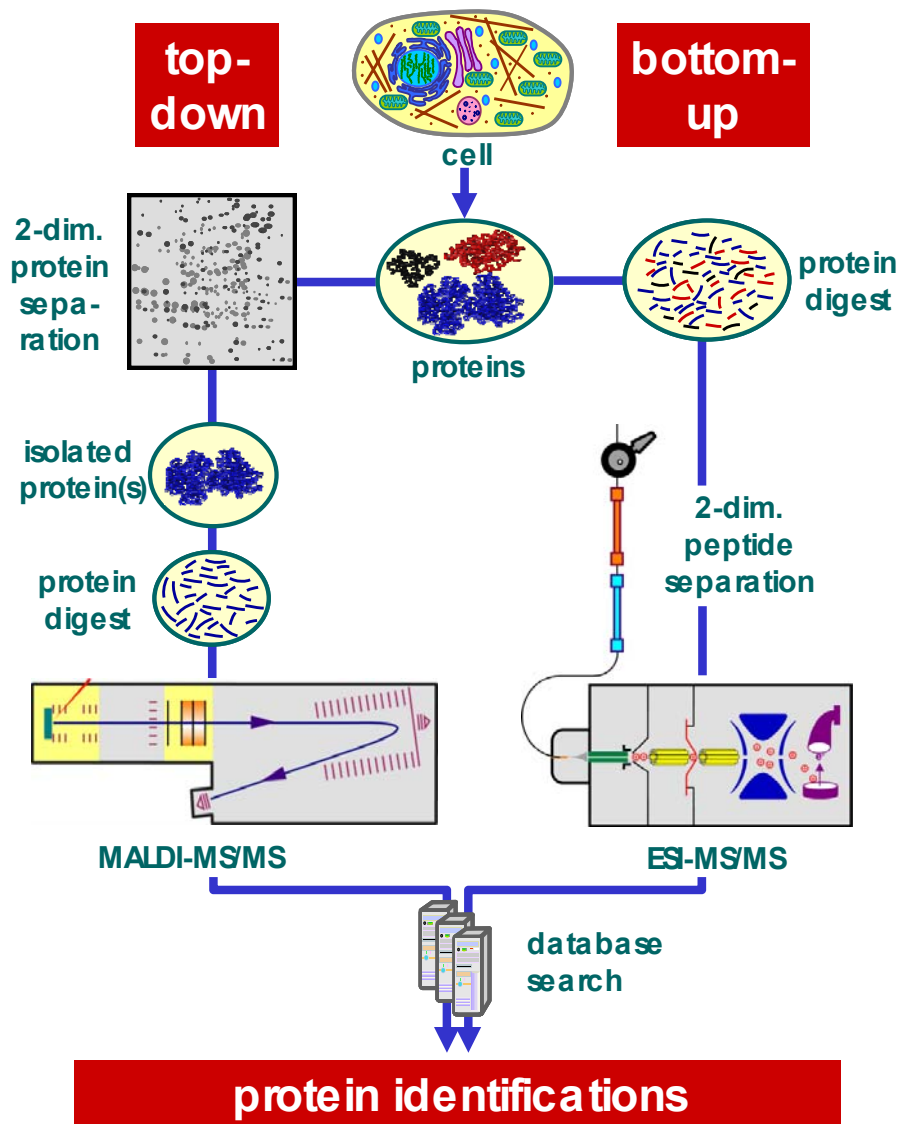


Fig. 5. Strategies for proteome analysis: top-down proteomics (left) and bottom-up proteomics (right). Reproduced from Huber ^[27].

The study of a proteome is termed "proteomics". Because of the high complexity of proteomes, proteomics require the development of multi-dimensional analytical methods. Two strategies for proteome analysis have been explored. A scheme of these two strategies is depicted in Fig. 5. In the top-down approach, proteomes are mostly analyzed by two-dimensional gel electrophoresis (2D-GE). Proteins are first separated as a function of their isoelectric point followed by a size-dependent

separation step. The protein spots are usually extruded and submitted to tryptic digestion before analysis by mass spectrometry. In the bottom-up approach (or shotgun proteomics), proteins are digested immediately after their isolation. The resulting set of peptides is generally analyzed by multi-dimensional chromatography hyphenated to mass spectrometry.

3 High-performance liquid chromatography for the separation of biomolecules

Combination of several chromatographic separations requires a good comprehension of the mechanisms occurring in each separation step. The large size and the ampholytic nature of peptides and proteins are among others ^[28] important parameters to be considered in terms of stationary phase or elution conditions.

The major difference between the chromatographic separation of small organic and biomolecules is that peptides and proteins are able to simultaneously adsorb on different places of the stationary phase ^[29]. As a consequence, the complete desorption of a biomolecule only occurs when the elution strength of the eluents is high enough to take off all the adsorption points of the biomolecule on the stationary phase. These adsorption/desorption mechanisms are often described as an “on-off” model ^[30]. Multisite adsorption of biomolecules results in highly negative slopes in plots depicting the logarithm of the retention factor k as a function of the eluent strength (e.g. percentage of acetonitrile for reversed-phase chromatography). A comparison for nitrobenzene (123.1 Da) and lysozyme (16,951.5 Da) is depicted in Fig. 6.

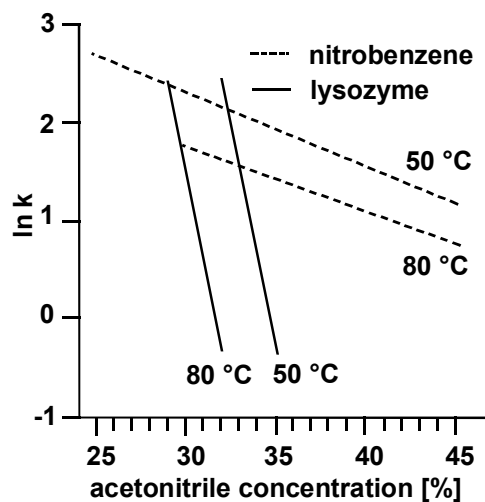


Fig. 6. Plots of the logarithmic retention factor vs. the concentration of acetonitrile in the reversed-phase chromatography of lysozyme (solid lines) and nitrobenzene (dashed lines) on a polystyrene stationary phase. Reproduced from Chen ^[31].

Multisite adsorption also results in the focusing of biomolecules at the column head. Consequently, it is possible to load large volumes of biomolecular samples without

band broadening. Because of their large size, biomolecules diffuse very slowly in porous stationary phases (Fig. 7a). They need more time than small organic molecules to enter and to exit the pores and this results in slow mass transfer. To avoid this phenomenon, it is possible to reduce the size of the particles of the stationary phase (Fig. 7b). However, this results in a lower permeability of the column.

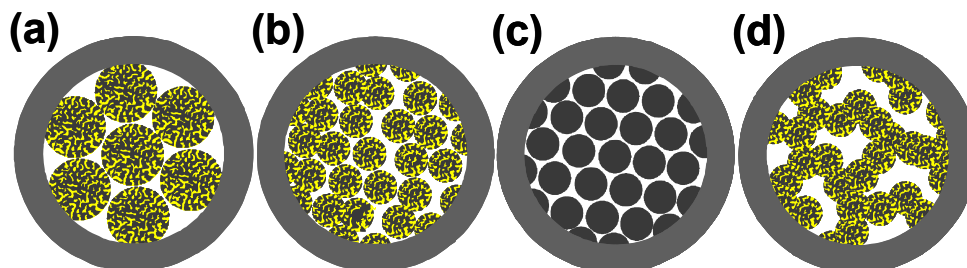


Fig. 7. Comparison of four different column packings ^[20]. (a) Spheric, porous, and large particles (high permeability, large interparticle voids, low efficiency); (b) spheric, porous, and small particles (low permeability, small interparticle voids, high efficiency); (c) spheric, non-porous, and small particles (low permeability, small interparticle voids, high efficiency, small loadability), (d) monolithic phase (high permeability, high efficiency).

One can also use non-porous material (Fig. 7c) but at the cost of a decrease in terms of column loadability. Stationary phases particularly well designed for the separation of biomolecules are the so-called monoliths (Fig. 7d). Because of their structure, monolithic columns combine high permeability and high efficiency. A monolith (or a continuous bed volume) is a unitary porous structure formed by *in situ* polymerization ^[32-35]. Because of the monolithic structure, the whole mobile phase is pumped through the pore channels ^[36]. Thus the mass transfer of the analytes is supported by convective flow in the macropore channels ^[37] (> 50 nm) and the diffusion dependent mass transfer occurs only in the micro- and in the mesopores (< 2 nm, and 2 - 50 nm, respectively). Monolithic columns also present a high permeability and consequently permit rapid separations of biomolecules at high flow rates ^[38-42].

To better understand the requirements of each separation step, the principles of the chromatographic methods ^[43-46] employed in our work are summarized in Tab. 3 and described in the next sections.

Tab. 3. Chromatographic separation methods combined in this work.

method of separation	acronym	principle of separation
<i>non-interactive methods</i>		
size-exclusion chromatography	SEC	differences in molecule size
<i>interactive methods</i>		
ion-exchange chromatography	IEC	electrostatic interactions
reversed-phase chromatography	RPC	solvophobic interactions
ion-pair reversed-phase chromatography	IP-RPC	electrostatic and solvophobic interactions
affinity chromatography	AC	biospecific interactions

3.1 Reversed-phase and ion-pair reversed-phase HPLC

In reversed-phase high-performance liquid-chromatography (RP-HPLC) the analytes are separated in a polar medium (e.g. water). Retention occurs due to solvophobic interactions between the analytes and a hydrophobic stationary phase. Solid phases used in RP-HPLC are mostly alkylated silica materials (e.g. octadecyl silica) or hydrophobic organic polymers (e.g. PS-DVB). The elution of the retained analytes takes place through the addition of an organic solvent (e.g. acetonitrile or methanol) in the elution mixture. This thermodynamically facilitates the desorption of the analytes from the stationary phase. Polar acids such as phosphoric acid, hydrochloric acid, and formic acid are usually added to the mobile phase in order to denature the proteins. Peak symmetries are thus improved by avoiding secondary interactions between biomolecules and residual silanol groups of silica-based stationary phases. Under such conditions, the retention of the analytes is mostly due to solvophobic interactions^[47]. A schematic representation of such interactions is depicted in Fig. 8.

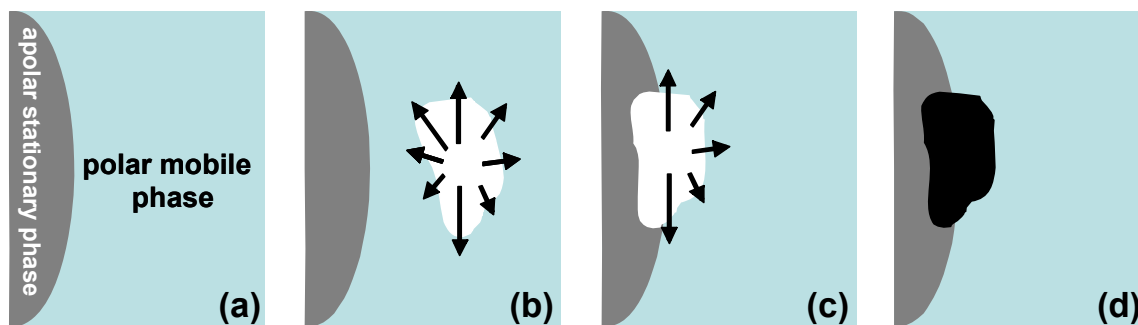


Fig. 8. Schematic representation of a solvophobic effect.

The introduction of an apolar analyte in a polar mobile phase is energetically unfavorable. The generation of a cavity implicates the destruction of numerous hydrogen bonds, which is energetically not counterbalanced by the solvation of the analyte in the solvent (a) and (b). The adsorption of an apolar analyte on the apolar stationary phase requires the destruction of less hydrogen bonds and is consequently energetically more favorable (c). Adsorption consequently occurs on the stationary phase (d).

In ion-pair reversed-phase high-performance liquid chromatography, relatively hydrophobic acids (e.g. trifluoroacetic acid, heptafluorobutyric acid) or bases (triethylammonium acetate) are added to the eluents. Consequently, biomolecules are not only separated as a function of their hydrophoby but also as of their charge. Different models have been proposed (ion-pair model, dynamic ion-exchange model) but nowadays the non-stoichiometric model is generally accepted. The hydrophobic groups of the deprotonated acids or bases (amphiphiles) are adsorbed on the non polar stationary phase and an electrical surface potential is generated ^[48]. The concentration of the ion-pair reagent is generally not high enough to completely cover the stationary phase. Biomolecules are then retained on the stationary phase because of electrostatic and solvophobic interactions ^[49-52]. The principle of ion-pair reversed-phase chromatography is schematically depicted in Fig. 9a.

Reversed-phase and ion-pair reversed-phase chromatography are very well designed to separate peptides and proteins with high efficiency and high peak capacity. This is mostly due to the absence of conformation isomers in the media during the separation process. The chromatographic conditions (e.g. 0.05 % TFA) are indeed harsh enough to denaturate most of the proteins. This separation process can consequently not be used for the fractionation of protein mixtures performed in order to collect biologically active proteins. The separation of nine peptides with ion-

pair reversed-phase high-performance liquid chromatography is depicted in Fig. 9b.

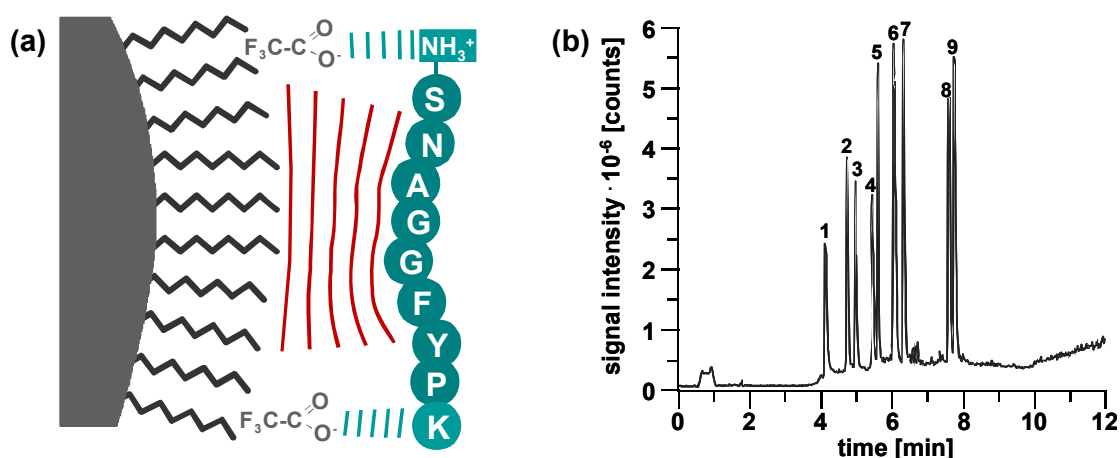


Fig. 9. (a) Schematic representation of electrostatic and solvophobic interactions in ion-pair reversed-phase chromatography. (b) IP-RP-HPLC-MS analysis of a 9-peptide mixture ^[20].

Column, PS-DVB, 60 x 0.20 mm i.d.; mobile phase, (A) H_2O + 0.05 % TFA, (B) ACN + 0.05 % TFA; gradient, 0-50 % B in 15 min; flow rate, 2.5 $\mu\text{L}/\text{min}$; 50 $^\circ\text{C}$; detection, ESI-MS; sample, solution containing 1=bradykinin fragment 1-5, 2=[Arg8]-vasopressin, 3=methionine enkephalin, 4=leucine enkephalin, 5=oxytocin, 6=bradykinin, 7=LHRH, 8=bombesin, 9=substance P; injection, 1 ng of each peptide.

The eluents (e.g. water, methanol, and acetonitrile) and the modifiers (e.g. TFA, and TEAA) used in RP-HPLC and IP-RP-HPLC are volatile. This volatility permits to hyphenate the chromatographic separation step with mass spectrometric detection by using interfaces such as electrospray ionization ^[53;54] or matrix-assisted laser desorption ionization ^[55;56]. The advantages of a mass spectrometric detection are detailed in section 4.2 of this chapter.

3.2 Ion-exchange HPLC

In ion-exchange high-performance liquid-chromatography (IEX-HPLC) the analytes submitted to separation are positively or negatively charged. One speaks of cation-exchange HPLC or anion-exchange HPLC, respectively. The positive (resp. negative) analytes electrostatically interact with negative (resp. positive) functional groups grafted on the stationary phase. For peptides and proteins, the interactions with the stationary phase may occur with amino-acid residues inside the peptide backbone but also with the carboxyl group at the C-terminus and the amine group at the N-terminus of the sequence. Amino-acid residues electrostatically interacting in IEX-HPLC are the acidic amino acids (aspartic acid, and glutamic acid) and the basic

amino acids (arginine, lysine, and histidine). A schematic representation of the electrostatic interactions taking place in ion-exchange chromatography is depicted in Fig. 10.

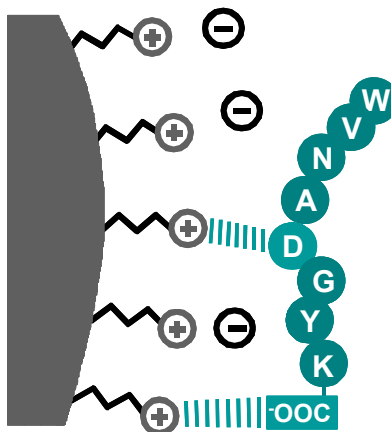


Fig. 10. Electrostatic interactions between an anion-exchange stationary phase and the octapeptide WVNADGYK ^[20].

Because of their amphoteric properties, proteins can be separated by anion-exchange and cation-exchange HPLC. At $\text{pH} > \text{pI}$, proteins have a negative net charge and can be separated by anion-exchange HPLC. On the contrary, at $\text{pH} < \text{pI}$, proteins have a positive net charge and they can be separated by cation-exchange HPLC. The choice between cation-exchange and anion-exchange HPLC is determined by the pI value of the proteins and the pH value at which the separation should be performed ^[57]. To get reproducible retention of proteins on the stationary phases, the difference $|\text{pH} - \text{pI}|$ should be higher than one pH unit. Protein separations using (a) anion-exchange and (b) cation-exchange HPLC are depicted in Fig. 11.

The elution of the analytes is performed by increasing the ionic strength of the mobile phase. This is mostly achieved by performing a salt gradient (e.g. sodium chloride) during the course of elution. Consequently, the eluents and as a matter of fact IEX-HPLC are generally non compatible with hyphenation to mass spectrometry. Organic solvents (e.g. acetonitrile) are sometimes added to the aqueous eluents to lower secondary hydrophobic interactions between analytes and stationary phase ^[58].

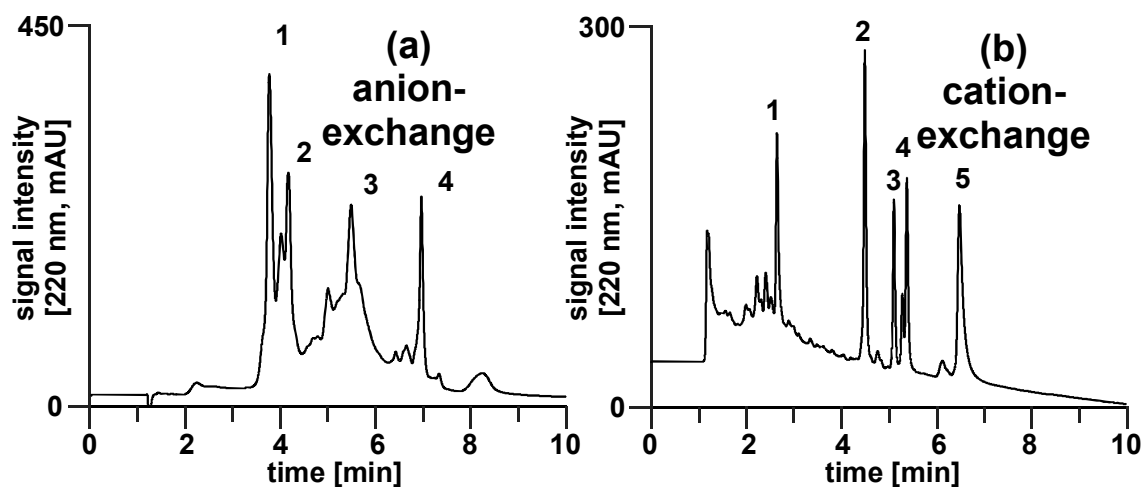


Fig. 11. Separation of proteins with anion-exchange and cation-exchange HPLC ^[20]. (a) Column, 250 x 4.0 mm i.d. ProPac SAX-10, 10 μ m; mobile phase, 10-min gradient of 0-0.50 mol/L NaCl in 20 mmol/L Tris-HCl, pH 8.0; flow rate, 1.0 mL/min; temperature, ambient; detection, UV at 220 nm; sample, 10 μ L of a solution containing 1=conalbumin (130 μ g/mL), 2=transferrin (200 μ g/mL), 3=ovalbumin (500 μ g/mL), 4=trypsin inhibitor (170 μ g/mL); (b) column, 250 x 4.0 mm i.d. ProPac SCX-10, 10 μ m; mobile phase, 10-min gradient of 0.50 mol/L NaCl in 50 mmol/L Na₂HPO₄, pH 6.0; flow rate, 1.0 mL/min; temperature, ambient; detection, UV at 220 nm; sample, 10 μ L of a solution containing 1=trypsinogen (400 μ g/mL), 2= α -chymotrypsinogen A (70 μ g/mL), 3=ribonuclease A (300 μ g/mL), 4=cytochrome C (70 μ g/mL), 5=lysozyme (100 μ g/mL).

During a separation performed with ion-exchange HPLC, the three-dimensional structures of biomolecules are usually preserved. Proteins are not denatured as long as the pH of the eluent does not attain harsh values. Under these conditions, biomolecules collected after IEX-HPLC are still biologically active. On the other hand, the three-dimensional structures of the proteins and the resulting different conformations are mostly responsible for broad chromatographic peaks. For this reason, the efficiency and the peak capacity in ion-exchange HPLC are significantly smaller than in reversed-phase- and ion-pair reversed-phase HPLC.

3.3 Affinity chromatography

Affinity chromatography is based on highly-specific biological interactions between two species such as antigen and antibody, glycoprotein and lectin, or enzyme and co-enzyme. One of the partners is covalently bound to the stationary phase, whereas the other (the analyte) is present in the sample. The biospecific interaction between the two partners is used to selectively bind the analyte to the column. The elution is

performed either with competitive adsorption or with conformational modification. The latter can be achieved by variation of the pH or of the ionic strength. A schematic representation of affinity chromatography is depicted in Fig. 12.

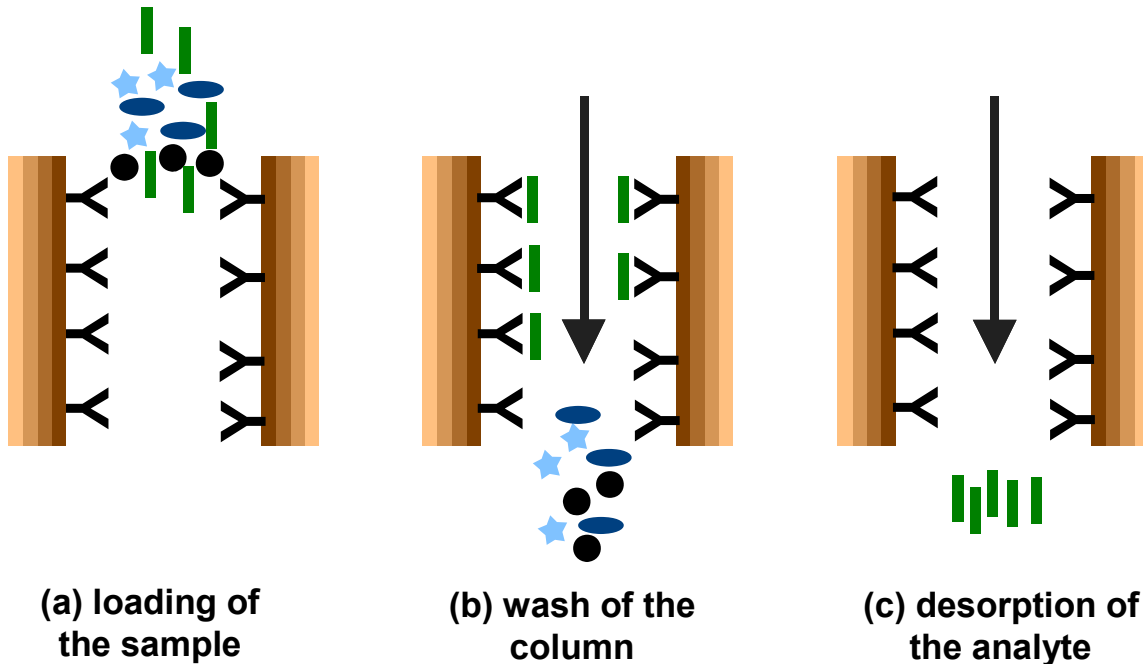


Fig. 12. Principle of affinity chromatography ^[59].

Affinity chromatography is the separation method with the highest selectivity, enabling the isolation of a single biomolecule (e.g. protein) from very complex samples (e.g. serum) ^[60;61]. This high selectivity is due to a very specific affinity interaction between both interacting molecules. Electrostatic interactions but also hydrogen bonding, hydrophobic interactions and tight steric fit of the interacting molecules are involved.

3.4 Size-exclusion chromatography

In size-exclusion chromatography, analytes are separated as a function of their size (hydrodynamic volume). The stationary phase consists of a porous material with defined pore sizes. Molecules over a critical size are too voluminous to enter into the pores and are eluting in the solvent front. Molecules under the exclusion limit not only move between the particles of the stationary phase but also penetrate into the pores. Thus big analytes elute at the beginning, whereas small analytes elute at the end of the chromatographic process, as illustrated in Fig. 13. The mobile phase acts as a

solvent and has no influence on the separation. Any adsorptive interactions between the surface of the chromatographic support and the analytes have to be avoided in order not to bias the size dependence of elution.

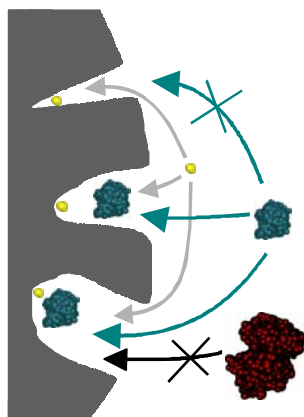


Fig. 13. Principle of size-exclusion chromatography ^[27].

3.5 Restricted access materials

Restricted access materials (RAM, term introduced in 1991 ^[62]) have been developed in order to perform on-line sample preparation and to get rid of the handling of untreated biological samples. Restricted access materials permit the direct injection of complex biological fluids (e.g. serum) into a HPLC system. On a RAM column, small proteins or organic molecules are extracted whereas the majority of proteins are flowing through the column. This result is achieved by the combination of different chromatographic modes. Various stationary phases are nowadays available and different mechanisms have been developed ^[63-65]. Generally, a combination of size-exclusion chromatography and reversed-phase HPLC is employed. The outside surface of the stationary phase is constituted of unreactive, hydrophilic groups, whereas the inner surface of the stationary phase is grafted with hydrophobic material. Large biomolecules can not access the inner surface of the chromatographic support and are consequently not retained. On the contrary, small biomolecules or small organic molecules penetrate into the pores, where they are retained ^[66-68]. A schematic representation of a restricted access material is depicted in Fig. 14.

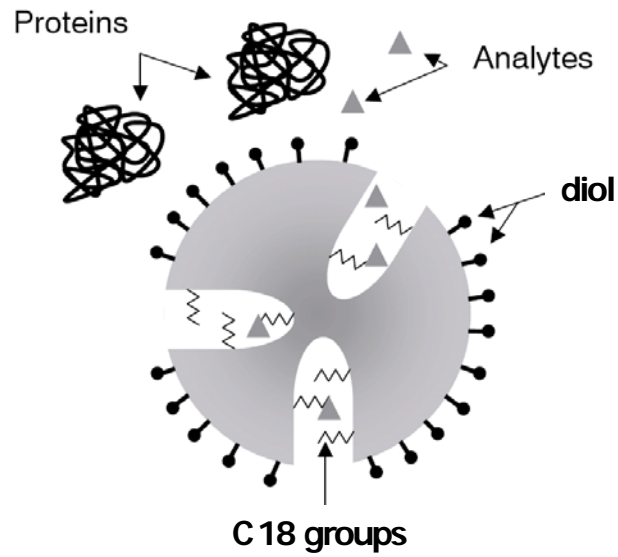


Fig. 14. Schematic representation of the restricted access medium LiChrospher ADS. Large proteins can not access the inner surface of the chromatographic support. Analytes can penetrate into the pores, where they are retained ^[65].

4 Methods of detection

In the next section, a brief description of the different methods of detection utilized in this work is given. First, immunodetection methods are described. They are generally highly specific and have very low limits of detection. However, they require a good knowledge of the analyte and the production of new antibodies for each new investigated analyte. On the other side, mass spectrometric detection is more generic. The hyphenation of mass spectrometry with chromatography is also relatively easy. Mass spectrometry is nowadays the detection method of choice in analytics and bioanalytics.

4.1 Immunodetection methods

Enzyme-linked immunosorbent assays (ELISA) were first described in the seventies [15;16]. These biological assays permit to specifically detect and quantify biomolecules present at low concentrations in complex samples. Generally, antibodies specific for an antigen are bound to the surface of test tubes or in wells of a microtiter plate (Fig. 15a). The complex sample is then introduced and the antibodies specifically catch and bind the antigen of interest (Fig. 15b).

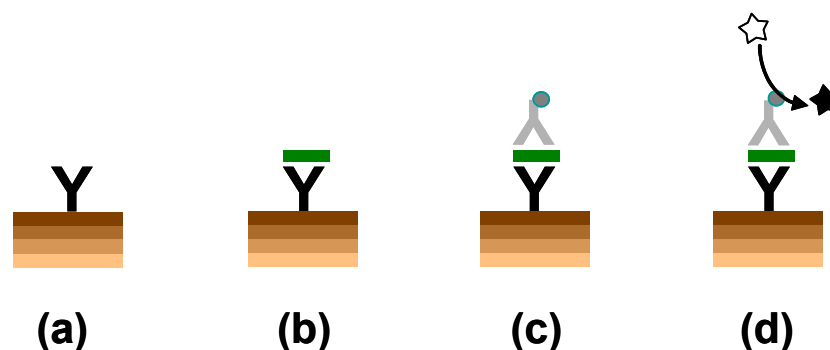


Fig. 15. Schematic representation of a sandwich ELISA.

(a) Plate coating with catching antibodies; (b) sample loading and capture of antigen; (c) addition of detecting antibodies and binding to antigen; (d) substrate addition and conversion in a detectable form.

Since the other constituents of the sample are not retained by the antibodies, it is possible to wash them out with an appropriate buffer. This mechanism is similar to the one utilized in affinity chromatography. Detection and quantitation are then performed by means of antibodies specific to another epitope than the one already occupied by the catching antibodies. The detecting antibodies are themselves

covalently bound to an enzyme. A sandwich structure (catching antibody – antigen – detecting antibody - enzyme) is obtained (Fig. 15c). Finally, a substrate is added in the medium and is converted to a detectable (chromogenic or fluorescent) form by the enzyme (Fig. 15d).

The main practical advantages of this technique are the ease of handling, the possibility to work with fully automated systems, and the rapidity of the detection (micro-ELISA reader) ^[69]. Moreover, very low limits of detection are achieved ^[70]. However, the development of two types of antibodies specific to different epitopes of the antigen are time consuming and expensive.

Numerous variations of ELISA have been developed. Among them, ECLIA (electrochemiluminescence immunoassay) is also based on a sandwich principle. However, the catching antibodies are immobilized on magnetic beads through the formation of a biotin-streptavidin complex. The detecting antibodies are covalently bound to a tris(2,2'-bipyridyl)ruthenium(II)-complex, which luminesced by electric energy ^[71].

4.2 Mass spectrometry

Mass spectrometry presents numerous advantages. First, this is a very sensitive method of detection. Signals have already been detected for analytes down to the zeptomol level ^[72]. Mass spectrometry also gives under special conditions (e.g. tandem mass spectrometry) structural information on the analyte. Finally, with the introduction in the late eighties of two revolutionary ionization techniques, namely electrospray ionization ^[2] and matrix-assisted laser desorption/ionization ^[3;4], it has become possible to hyphenate mass spectrometry with high-performance liquid chromatography for the analysis of large biomolecules.

4.2.1 Principle of electrospray ionization

The term electrospray refers to the dispersion of a liquid in numerous small charged droplets with the help of an electric field. Thus, electrospray-ionization of an analyte is achieved by two processes:

- the formation of electrically charged droplets from the in-coming sample solution
- the transfer of the ions from the droplets to the gas phase under the influence of a strong electric field.

In practice, the sample solution is continuously pumped through a capillary tube into the ESI source at flow rates typically ranging from a few nL/min to several $\mu\text{L}/\text{min}$. The application of a high-voltage potential (2-6 kV) between the end of the capillary tube and the counter-electrode (end plate) results in the electrolysis of the liquid and induces a charge accumulation at the liquid surface located at the tip of the capillary (Fig. 16a). The charges repulse themselves and a decrease of the surface tension of the sample solution occurs. A so-called Taylor cone is formed and a thin liquid filament protrudes. The liquid filament finally breaks when the Rayleigh limit (the charge repulsion equals the surface tension) is exceeded.

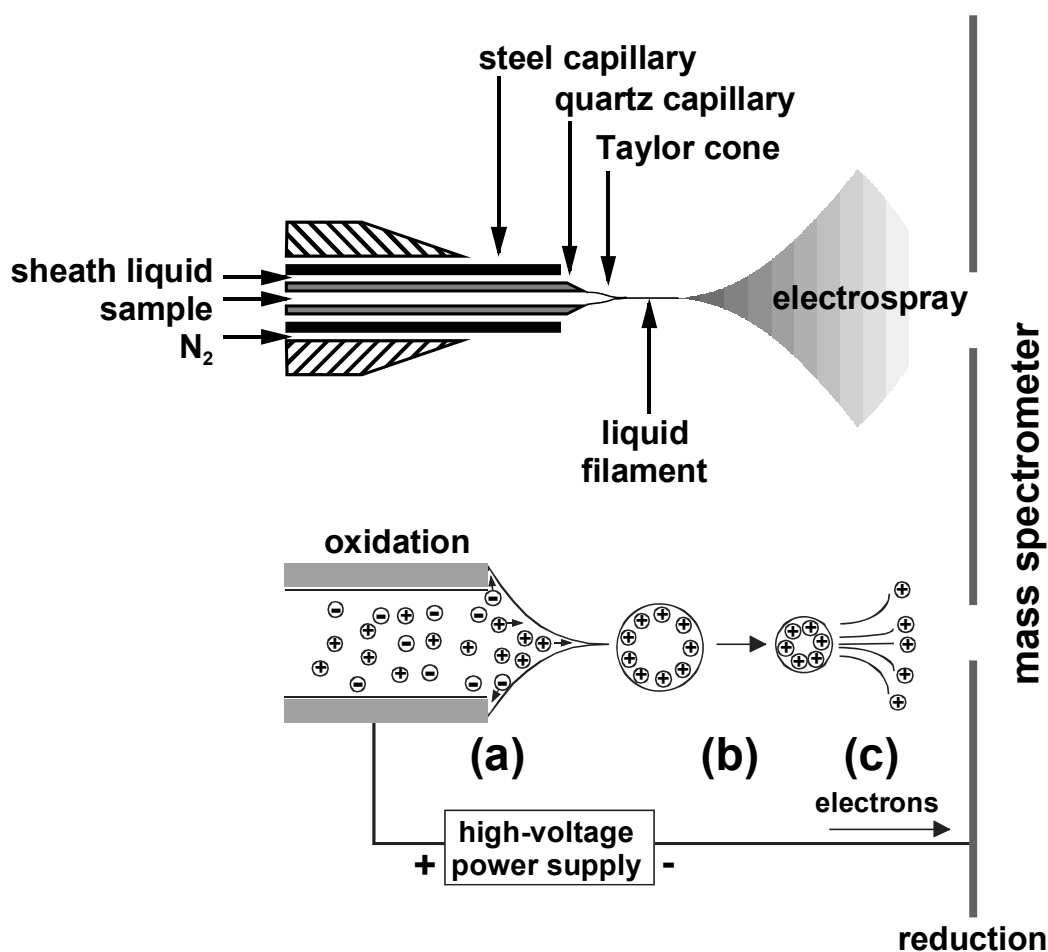


Fig. 16. Schematic representation of the ESI process at macroscopic (up) and microscopic (down) scale ^[59].

The droplets shrink with the evaporation of the solvents (Fig. 16b) and finally ions find themselves in the gas phase (Fig. 16c). The real mechanism is not yet completely understood. According to the so-called charge residue model ^[73] the

droplets shrink until only one single ion is present in the droplet. Whereas according to the ion evaporation model, the ions are released from the droplets with Coulomb explosions ^[74].

The hyphenation of ESI-MS to HPLC requires a compatibility of the eluent used during the separation step with the ionization process. For separations performed with reversed-phase HPLC or ion-pair reversed-phase HPLC the concentration of organic solvent determines the retention of the analytes on the column and can consequently not be arbitrarily changed. For this reason an organic solvent is sometimes added to the mobile phase after the column and before the electro spray source. This liquid is either introduced parallel to the capillary tube with a concentric arrangement ^[75;76] or introduced with a T-piece ^[77].

In pneumatically assisted ESI, the liquid entering in the ESI interface is surrounded by a third concentric tube permitting the introduction of a so-called spraying gas in order to support and stabilize the spray.

Electrospray ionization is especially well adapted to ionize large molecules (e.g. proteins, peptides, and nucleic acids) ^[78-80]. ESI permits to bring the analytes in the gas phase without inducing fragmentation. Another advantage of ESI is that during the ionization process, multi-charged ions are produced. Thus also big molecules, having a mass far above the detection range of the mass analyzer, can be detected. The value measured in the analyzer is indeed not the mass of the analytes but the m/z value with m the mass of the ion and z the charge of the ion. Consequently, an analyte with a mass of around 100,000 Da and possessing 100 charges is detected at around m/z 1,000. For this reason, molecules up to several millions Da can be analyzed with mass spectrometry despite limitations of quadrupole- and ion-trap analyzers.

4.2.2 Quadrupole mass analyzer

Quadrupole mass analyzers consist of four parallel metal rods arranged symmetrically. Ideally, the four rods should have the shape of a hyperbola in cross section but cylindrical rods practically approximate a hyperbolic-field. Diagonally opposite rods are electrically connected together. Both a direct current (DC) and an oscillating radio frequency (RF) signal are applied across the rods, adjacent rods having opposite charge (Fig. 17a).

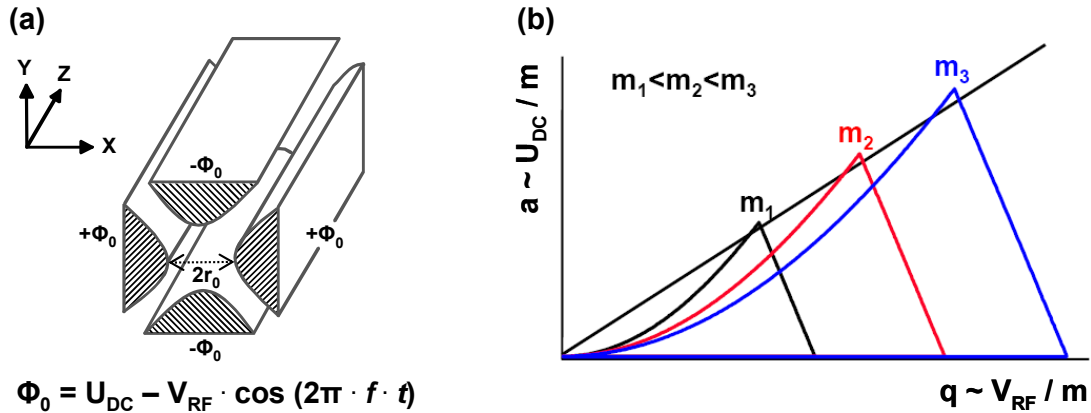


Fig. 17. (a) Quadrupole with hyperbolic rods and applied potential; (b) stability areas as a function of U_{DC} and V_{RF} for ions with different masses ^[81;82].

The potential Φ_0 between two adjacent rods can be described by $\Phi_0 = U_{DC} - V_{RF} \cdot \cos(2\pi f t)$ with f being the frequency of the RF potential. By applying this potential Φ_0 to each pair of rods, a quadrupolar electric field results.

Each point (x, y, z) in the electrical field is exposed to the potential

$$\Phi_{(x,y)} = \Phi_0 \cdot \frac{x^2 - y^2}{r_0^2}$$

with r_0 being the half of the distance between two opposite rods.

Thus, $\Phi_{(x,y)}$ does not depend on z and the ions entering the analyzer are only submitted to accelerations following the x - and y -directions. The trajectory of an ion is then described by the so-called Mathieu equations ^[83;84], and it is possible to define in a (U_{DC}, V_{RF}) plot, stability regions for which the coordinates (x, y) of an ion remain smaller than r_0 (Fig. 17b). Ions present in such a region have stable trajectories in the quadrupole and will successfully traverse the quadrupole to reach the detector.

All other ions have unstable trajectories and will collide with the rods at some point. By maintaining a constant U_{DC}/V_{RF} ratio, a straight operating line of the analyzer is obtained as depicted in Fig. 17b. By scanning along this operating line, a successive detection of different masses occurs. The higher the slope, the better the resolution. By applying $U_{DC} = 0$, all the ions with a m/z higher than the one defined by V_{RF} are stable (x and $y < r_0$). Under these conditions the ions are systematically brought back to the center of the rods, even if they were deflected by a collision. This focusing effect is important to increase the transmission of ions after collisions. Quadrupole mass analyzers present a high ion transmission from the interface to the detector. They are also easy to use and the calibration is long-term stable.

4.2.3 Quadrupole ion trap mass analyzer

The principle of quadrupole ion trap mass analyzers was already described by W. Paul and H.S. Steinwedel in 1960 ^[85]. However, the first commercial analyzer was released only 20 years later because of technical reasons ^[86]. In an ion trap analyzer, the ions are captured (trapped) in an electric field. In this electric field, the ions move on stable trajectories. By varying the electric field the ions can be selectively ejected from the ion trap to the detector as a function of their m/z . Contrary to a quadrupole, where only ions with a specific m/z are able to pass through the analyzer, the whole ions are captured in the ion-trap analyzer and are only afterwards selectively ejected to the detector as a function of their m/z . A schematic representation of an electrospray-ion trap mass spectrometer is depicted in Fig. 18.

The quadrupole ion trap consists of a ring electrode and two hyperbolic end-cap electrodes. The ions can enter and exit the ion trap through two small holes in the middle of each end-cap electrode. The ions produced in the ion source are focused and transported to the ion trap by means of skimmers and octapole lenses. In order to slow down the incoming ions, a helium pressure of $\sim 3 \times 10^{-3}$ bar is maintained in the ion trap. Thus the kinetic energy of the ions is reduced by collisions of the ions with helium atoms, and the ions are trapped. The application of a variable voltage on the ring electrode and on the end-cap electrodes results in a quadrupole electric field in the trap. The ions are then stabilized and move on trajectories defined by the Mathieu equations ^[87].

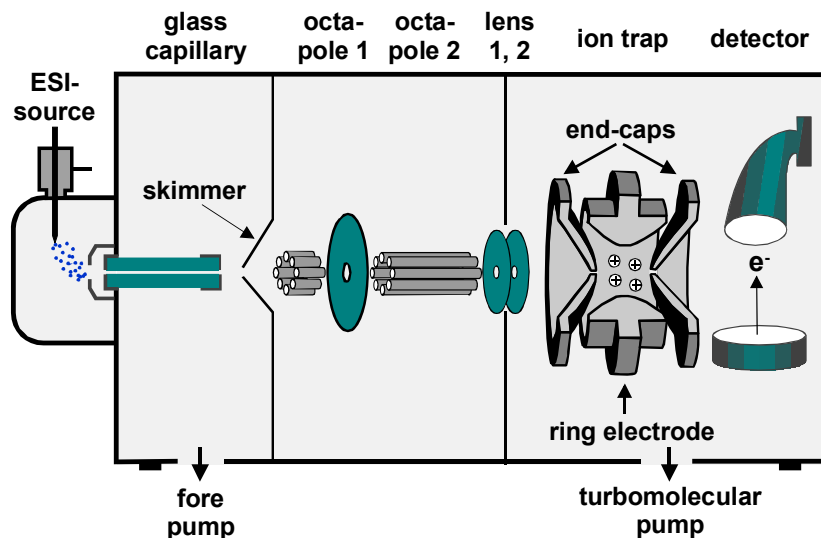


Fig. 18. Schematic representation of an electrospray-ion trap mass spectrometer.

During a measurement cycle, the ions are accumulated in the ion trap for a short period of time, mostly between 0.1 and 100 ms. Subsequently, the entrance of the ion trap is closed for in-coming ions by changing electric potentials. Thus once the ion trap is filled, no other ion can penetrate. This permits first to avoid a too high charge concentration in the ion trap and a resulting decrease of the mass accuracy. This also prevents any interference between two consecutive measurements.

The sequential ejection of the ions happens by the application of a mass selective resonance frequency to the end-cap electrodes. The oscillation amplitudes of the ions increase, and finally a destabilization of the ion trajectories occurs and the ions are ejected. The out-going ions are finally detected with a multiplier, which converts and amplifies the signal into an electric current.

One of the advantages of the ion trap is the possibility to perform tandem mass spectrometry experiments (MS/MS). This permits the production of fragment ions by means of collision-induced dissociations. In this modus, ions with an m/z of interest (called precursor ions) are isolated in the ion trap, whereas the other ions are ejected by the application of resonance frequencies. The kinetic energy of the precursor ions is then increased up to a value at which the ions collide with the helium atoms also present in the trap. During these collisions, enough potential energy is transferred to the ions to induce fragmentation. The fragments (product ions) are then sequentially ejected from the trap and detected. An alternative is to isolate in the trap a product ion and to fragment it by performing new collision-induced dissociations. This

process is theoretically n times repeatable (MS^n) but is practically limited because of a signal decrease by each MS cycle. A schematic representation of the different steps to perform an MS/MS analysis is depicted in Fig. 19.

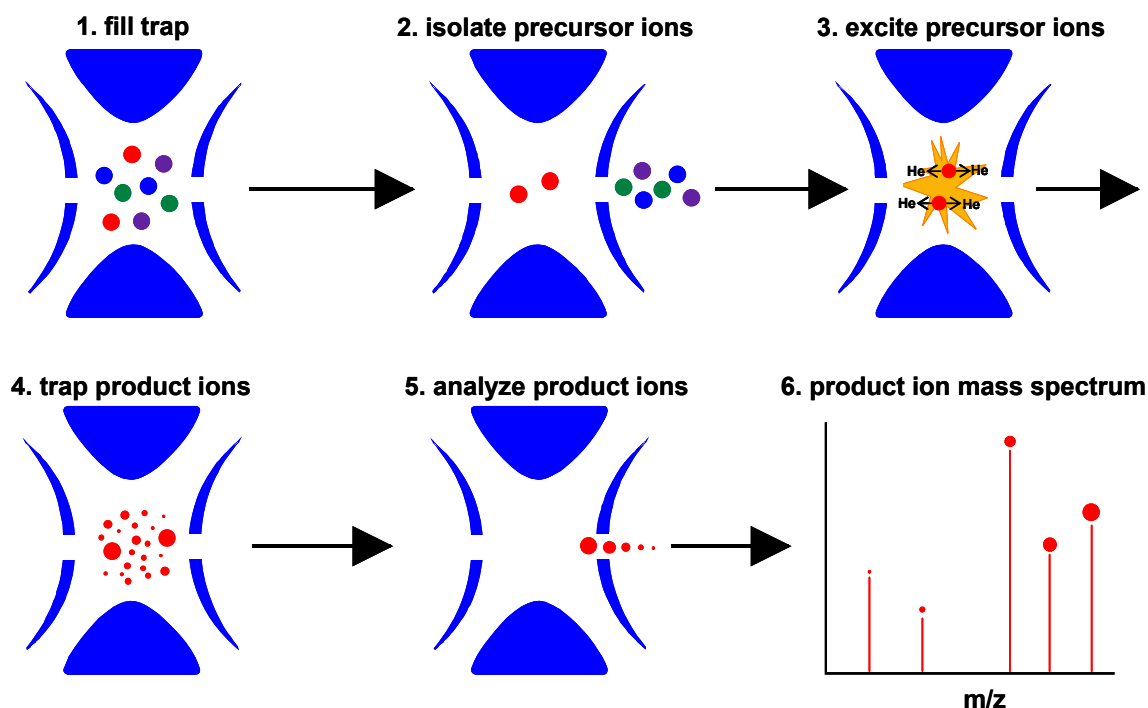


Fig. 19. Schematic representation of an MS/MS cycle with a quadrupole ion trap mass spectrometer.

The advantages of an ion trap are the high scan speed (10 to 20 times quicker than a quadrupole mass spectrometer) and the possibility to perform up to 10 successive MS/MS cycles. As a result, ion traps are the analyzers of choice for hyphenation with chromatographic systems, which require high data acquisition rates but also product ion mass spectra for structure elucidation and identification of the analytes.

4.3 Identification of peptides and proteins with mass spectrometry and algorithmic computation ^[88]

Two different approaches are nowadays used to identify peptides and proteins with mass spectrometry detection. Both the Peptide Mass Fingerprinting (PMF) and the Peptide Fragment Fingerprinting (PFF) approaches are based on the comparison of measured mass spectra with theoretical peak lists saved in a data bank. Generally, the complex protein mixture one wants to analyze is previously enzymatically digested with a specific enzyme (e.g. trypsin) and the resulting complex peptide mixture is separated by two-dimensional high-performance liquid chromatography before introduction in the mass spectrometer.

In the Peptide Mass Fingerprinting (PMF) approach, a peptide is identified by measuring its m/z and by comparing it to a list of m/z calculated from a list of peptides generated from *in silico* digestion of the proteins in a database. Protein identification occurs when one or more peptides matching a part of the protein sequence are identified. In this approach, the identification of a peptide is only based on a single m/z measurement. To get high-confidence results, it is of importance to get very accurate m/z values (3-5 ppm mass deviation) and to use high-resolution mass analyzers such as time-of-flight (TOF) or Fourier Transform Ion Cyclotron Resonance (FT-ICR) being able to differentiate lysine (m/z 128.095) from glutamine (m/z 128.059). Anyways, to get the best mass accuracy it is recommended to perform internal mass recalibration. This procedure is, however, time consuming for the operator and seldom automated.

In the Peptide Fragment Fingerprinting (PFF) approach, the complex protein mixture to be analyzed is digested and separated as in the previously described PMF approach. The difference appears in the mass spectrometric detection. After MALDI or ESI ionization, peptides are fragmented in the gas phase. The fragments obtained in the gas phase are finally separated, detected, and compared to sets of theoretical fragments. Theoretical fragments are obtained from *in silico* fragmentation of peptides, according to fragmentation rules and to the type of instrument used for the fragmentation ^[89-92]. In this approach, a peptide is identified by the mass of the intact precursor peptide ion, but also by the m/z of each fragment observed in the spectrum. For this reason it is possible to get unambiguous peptide and protein identification without high accuracy mass instruments (e.g. quadrupole ion trap).

At low collision energy (< 200 eV), fragmentations generally occur at three different locations of the peptide backbone: just before, just after, or exactly at the location of the peptide bonds. If the charge remains on the N-terminus of the sequence, so-called a_n , b_n , and c_n ions are obtained, whereas x_n , y_n , and z_n ions are formed when the charge is carried on the C-terminus of the sequence. The common nomenclature for sequence ions in peptide mass spectra is depicted in Fig. 20. Some other fragmentations are sometimes observed. They usually correspond to elimination of water, ammonia, and carbon monoxide along the peptide backbone but eliminations from the side chains can also happen.

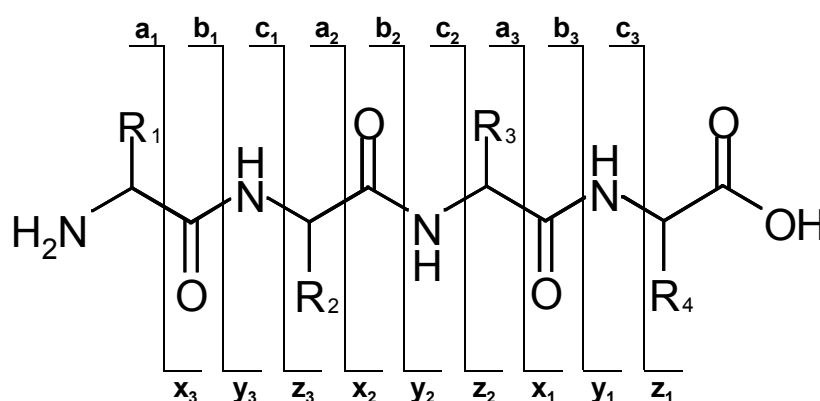


Fig. 20. Collision induced fragmentations along the peptide backbone ^[93].

Numerous algorithms have been developed for fully automated interpretation of huge amounts of MS and MS/MS data. Some of them are listed in Tab. 4. For PFF, two search engines are leading on the market: Sequest from Jimmy Eng and John Yates ^[8], and Mascot from Matrix Science ^[6;7]. Generally the operator only needs to enter search parameters (e.g. protease used for the digestion, chemical modification, investigated organism, mass tolerance) and to load the mass spectra as peak lists (e.g. ASCII format) into the search engine. After computation the operator gets a list of identified peptides/proteins. In the case of Mascot, the confidence of the identifications is evaluated by the so-called MOWSE (molecular weight search) scoring algorithm. A MOWSE score is computed for each peptide/protein. The MOWSE score is defined as $-10 \cdot \log(P)$ with P the probability to assign a random hit as a positive hit. For instance, a MOWSE score 200 signifies that the probability to have a random hit is 10^{-20} .

Tab. 4. List of on-line available search engines for protein identification with PMF ^[94].

search engine	web site
Aldente	http://www.expasy.org/tools/aldente/
Mascot	http://www.matrixscience.com
Protein prospector	http://prospector.ucsf.edu/
Profound	http://prowl.rockefeller.edu/prowl-cgi/profound.exe
PeptideSearch	http://www.narrador.embl-heidelberg.de
DARWIN	http://www.cbrg.ethz.ch/darwin

Because the MOWSE score is only an approximation, the level of false positives is usually independently tested by performing a search in a reverse database ^[95]. In a reverse database, each polypeptide sequence of the forward database is written in reverse from last residue to first. Consequently, the reverse database has the same size, the same number of entries, and the same amino-acid distribution than the forward database. The hits obtained by performing a search in the reverse database are truly random hits. Assuming to have the same number of random hits by performing a search in the forward version of the database, one can evaluate the confidence of your peptide/protein identifications by computing the ratio:

$$\frac{\textit{number of hits obtained in the reverse database}}{\textit{number of hits obtained in the forward database}}$$

The obtained value should approximately correspond to the confidence threshold chosen for the Mascot search (generally 95 %).

Chapter III

Development of monolithic
immunoabsorbers for the
isolation of biomarkers
from human serum

III. Development of monolithic immunoadsorbents for the isolation of biomarkers from human serum

1 Introduction

1.1 Aim of the work

High-performance affinity chromatography (HPAC) is one of the most sensitive and selective methods for the isolation of single compounds from highly complex biological samples [61, 96]. Already described in 1953 by Leonard Lerman [97], affinity chromatography permits nowadays the specific extraction of proteins from samples as complex as serum.

However, to cover a large range of analytes the high specificity of HPAC implicates the use of a very high number of affinity columns having different selectivities. Consequently, affinity chromatography is not very useful for large-scale proteomic and peptidomic analysis. Nevertheless, affinity chromatography finds two prominent applications in proteomics.

The first application consists in the depletion of the high-abundant proteins from human serum. Antibodies against albumin, transferrin, and immunoglobulins are immobilized on a stationary phase. Then, the sample (e.g. serum) is injected into the affinity column. The flowthrough containing the compounds of interest is collected for further analyses, whereas major serum proteins are retained on the column. After elution of high-abundant proteins and regeneration of the column, a new sample can be injected. Commercially available columns allow the elimination of > 99 % of the targeted proteins for up to 200 injections without showing any decrease in performance [98,99].

Affinity chromatography finds a second major application for the quantitative analysis of biomarkers. In the past decade, the analysis of whole proteomes has triggered the search for biomarkers of disease. Because simultaneous quantitation of all the proteins present in samples as complex as human serum is not possible, quantitation should be only based on few proteins specific to one biological or diagnosis question. Biomarkers are typically very low concentrated: ng/mL concentration range for

common tumor markers, and pg/mL concentration range for heart disease markers [100]. With the actual status of technique, there is no simple way that mass-spectrometry based approaches can routinely quantify such biomarkers without reducing the complexity of the sample [101]. Because only few biomarkers are introduced yearly into the market [100], a lot of time and money can be invested for the investigation and the analysis of each biomarker. In such conditions, affinity-based approaches can be developed for each biomarker and affinity chromatography plays a major role. Antibodies specific to the biomarker of interest are immobilized on a stationary phase. The sample (e.g. serum) is injected into the affinity column and all proteins are flowing through, except the biomarker specifically retained on the column. After elution, the biomarkers are collected for further analyses.

Monolithic supports are materials comprising a continuous bed volume. Because of their structure, monolithic columns combine high permeability and high efficiency. Almost no diffusion resistance is observed during mass transfer (see chapter II, section 3). Consequently, antigens present in a sample/mobile phase can quickly interact with antibodies immobilized on such supports [102, 103]. Monolithic supports are thus the materials of choice to immobilize antibodies for affinity chromatography. Convective Interaction Media (CIM) monoliths are now available as flat disks [104]. Several monolithic disks bound with different antibodies/ligands can be stacked into a single cartridge, permitting the easy development of immunoabsorbers for conjoint liquid chromatography [105].

In this context, the aim of our work in cooperation with Roche Diagnostics (Penzberg, Germany) was to develop monolithic CIM immunoabsorbers, permitting to target biomarkers of heart disease at relevant concentrations. Our work was focused on two biomarkers: myoglobin and NT-proBNP.

1.2 Investigated biomarkers

1.2.1 Myoglobin as biomarker of heart infarct

Myoglobin ^[106] is a cytoplasmic protein which involved in the transport of oxygen. Myoglobin also permits oxygen storage in primary aerobic working muscles such as heart. Human myoglobin (M_w 17,041.9 / 17,052.5 Da) is made of a single polypeptide chain of 153 amino acids:

```
GLSDGEWQLV LNVWGKVEAD IPGHGQEVLI RLFKGHPE TL EKFDKFKHLK  
SEDEMKASED LKKHGATVLT ALGGILKKG HHEAEIKPLA QSHATKHKIP  
VKYLEFISEC IIQVLQSKHP GDFGADAQGA MNKALELFRK DMASNYKELG  
FQG
```

The tertiary structure of the protein is mostly α -helical; eight α -helical portions are separated by unarranged structures (see Fig. 21). The dimensions of the protein measured in solution are: 4.5 x 3.5 x 2.5 nm. The property of myoglobin to bind oxygen molecules is due to the presence of a heme group. This non-covalent group made of a porphyrin ring is able to complex iron and is responsible for the red color of myoglobin. Two histidine residues inside the natural protein play a decisive role for the binding of oxygen to the heme group.

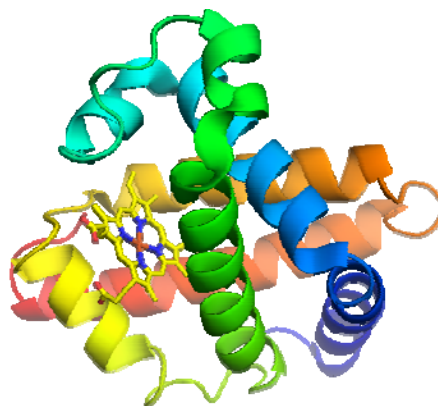


Fig. 21. Three-dimensional representation of human myoglobin.

When a muscle (e.g. heart muscle) is injured, myoglobin is released into serum ^[107]. Such a case occurs during myocardial infarction. Therefore, myoglobin is an important factor in the diagnosis of acute myocardial infarction ^[108], early reinfarction ^[109], and successful reperfusion following lysis therapy ^[110]. Two hours after infarction, the plasma myoglobin concentration rises, and after 6-9 hours a

concentration peak is reached (200 – 1,000 ng/mL plasma). After 12-24 hours the myoglobin concentration recovers a value of 3 -65 ng/mL [111, 112]. A method for the absolute quantitation of myoglobin in human serum has already been established [59;88;113;114] but myoglobin was not isolated with immunoabsorber. In the present work, the isolation of human hemoglobin from serum with immunoaffinity has been investigated.

1.2.2 NT-proBNP as biomarker of heart insufficiency

Brain Natriuretic Peptide (BNP) is a vasoreactive peptide hormone implicated in coronary heart disease, arterial hypertension, valvular disease, and primary myocardial disease. N-terminal prohormone Brain Natriuretic Peptide (NT-proBNP) and BNP are obtained after cleavage into two parts of the prohormone Brain Natriuretic Peptide (proBNP) by the protease furin (see Fig. 22). BNP is biologically active but very unstable in plasma. NT-proBNP is much more stable (a few days in blood). To circumvent this problem of stability, the concentration of BNP is generally evaluated by determining the concentration of NT-proBNP. In the present work, NT-proBNP has been retained as biomarker model [115]. The concentration of NT-proBNP in plasma is about 30 pg/mL for healthy people. In case of left ventricular diastolic and/or systolic dysfunctions, concentrations higher than 1,000 pg/mL are achieved [116, 117]. Thus, plasma NT-proBNP is recognized as a sensitive and specific marker for diagnosis of left ventricular dysfunction [118].

Two types of NT-proBNP were provided by Roche Diagnostics: recombinant histidine-tagged NT-proBNP expressed in *Escherichia coli* and synthetic NT-proBNP. Their sequences and properties are the following:

Histidine-tagged NT-proBNP (1-66): M_w 8521.2 / 8526.3 Da, pI 7.24

MRGSHHHHHH GSHPLGSPGS ASDLETSGLQ EQRNHLQGLK SELQVEQTSL
EPLQESPRPT GVWKSREVAT EGIRGHR

Synthetic NT-proBNP (1-76): M_w 8434.4 / 8439.4 Da, pI > 9.0

HPLGSPGSAS DLETSGLQEQ RNHLQGLKSE LQVEQTSLEP LQESPRPTGV
WKSREVATEG IRGHRKXVLY TLRAPR, with X = norleucine

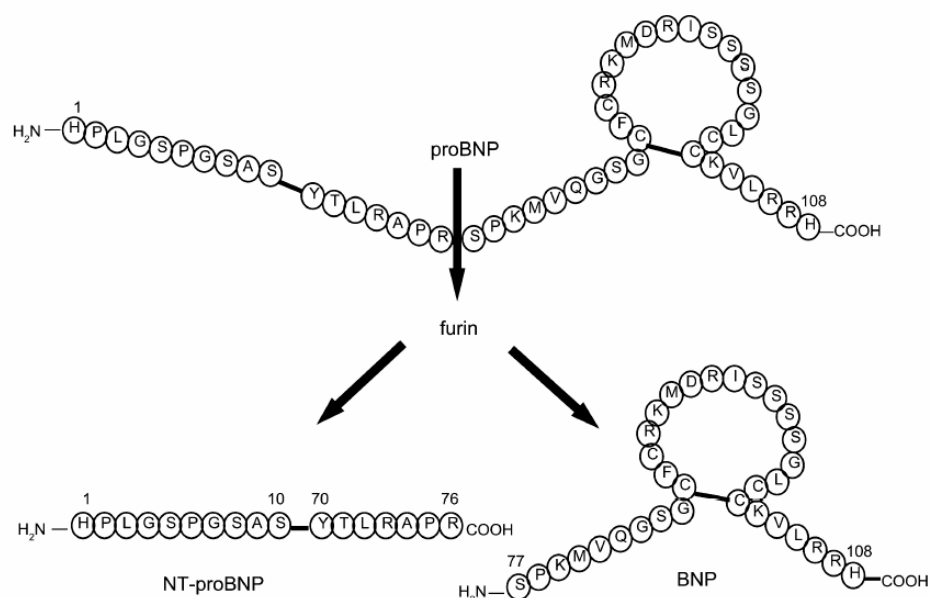


Fig. 22. Schematic drawing of proBNP showing enzymatic cleavage into biologically active BNP and NT-proBNP. Reproduced from Hall ^[119].

2 Materials and methods

2.1 Chemicals and samples

Deionized water (18.2 MΩ cm) was prepared with a Purelab Ultra Genetic system (Elga, Griesheim, Germany). Acetonitrile (E Chromasolv) was purchased from Sigma-Aldrich (Steinheim, Germany). Analytical reagent grade sodiumdihydrogenphosphate-1-hydrate, suprapur acetic acid (100 %), and sodium hydroxyde (> 99 %) were obtained from Merck KGaA (Darmstadt, Germany). Sodium chloride was supplied by Grüssing GmbH (Filsum, Germany). Guanidine hydrochloride (≥ 99.0 %), ortho-phosphoric acid (> 85 %), heptafluorobutyric acid (≥ 99.0 %), and trifluoroacetic acid (≥ 99.5 %) were purchased from Fluka (Buchs, Switzerland). Analytical reagent grade hydrochloric acid was purchased from Fisher Scientific (Schwerte, Germany). Ubiquitin from bovine erythrocytes was supplied by Sigma (Schnelldorf, Germany). Human myoglobin and myoglobin depleted serum were obtained from the Institute for Reference Materials and Measurements of the European Union (Geel, Belgium). Recombinant streptavidin, His-tagged NT-proBNP, synthetic NT-proBNP, anti-myoglobin antibodies, and anti-NT-proBNP polyclonal antibodies were supplied by

Roche Diagnostics (Penzberg, Germany).

Anti-NT-proBNP antibodies were polyclonal antibodies specific against amino acids [1-21] from NT-proBNP. The antibodies were extracted from the serum of sheep immunized with NT-proBNP. Procedures for the preparation and the biotinylation of the antibodies have already been published elsewhere ^[117]. A 1.67 mg/mL streptavidin solution was prepared by dissolving 5 mg of a streptavidin lyophilizate in 3.0 mL of a 0.5 mol/L, pH 8.0 sodium phosphate solution.

Epoxy-activated CIM disks with 6.0 mm diameter, 2.0 mm thickness and column volume of 56 μ L were kindly obtained from Aleš Štrancar (BIA Separations d.o.o., Ljubljana, Slovenia). The basis material of the CIM disks was poly(glycidyl methacrylate-co-ethylene glycol dimethacrylate). The polymer material has a volume of 33.6 μ L and a mass of 23 mg per disk. The polymer material is chemically stable at pH 1-14 and buffer concentrations 0-8 mol/L. Consequently, very harsh conditions (e.g. 100 mmol/L hydrochloric acid medium) can be employed to release the biomarkers from the antibodies.

Fractions were evaporated with a vacuum concentrator (model 5301) supplied by Eppendorf AG (Hamburg, Germany).

2.2 Analytical setups

2.2.1 Isolation of biomarkers with CIM disks

The following setup, particularly well designed for flow rates of 30-150 μ L/min, was used to run monolithic CIM disks. It consisted of a low-pressure Rheos 2000 HPLC system (Flux Instruments, Basel, Switzerland), a degasser from Knauer GmbH (Berlin, Germany), and a Rheodyne injection system (Rohnert Park, CA, USA) with a 1,000 μ L external loop. UV detection was monitored at 214 nm with a 45 nL Z-shaped capillary detection cell (Ultimate, LC Packings, Amsterdam, the Netherlands). Fractions collected after the separation performed on monolithic CIM disks were then analyzed either with HPLC-ESI-MS or with immunoassays.

2.2.2 High-performance liquid chromatography-mass spectrometry

The separation setup used to analyze fractions collected or biomarkers trapped after CIM disks consisted of a capillary/nano system from LCPackings (Amsterdam, The Netherlands), equipped with an Ultimate low-pressure gradient micro-pump (model

Ultimate), a Switchos micro-column 10-port switching unit with loading pump (model Switchos), and a micro-injector (model Famos). Trap columns (PS-DVB monolith 5 x 0.2 mm i.d., and Pepmap™ C18 5 x 0.3 mm i.d., 5 μm) were obtained from LCPackings (Amsterdam, the Netherlands). Analytical column (60 x 0.1 mm i.d.) was a poly-styrene/divinylbenzene (PS-DVB) monolith prepared according to the published procedure^[120].

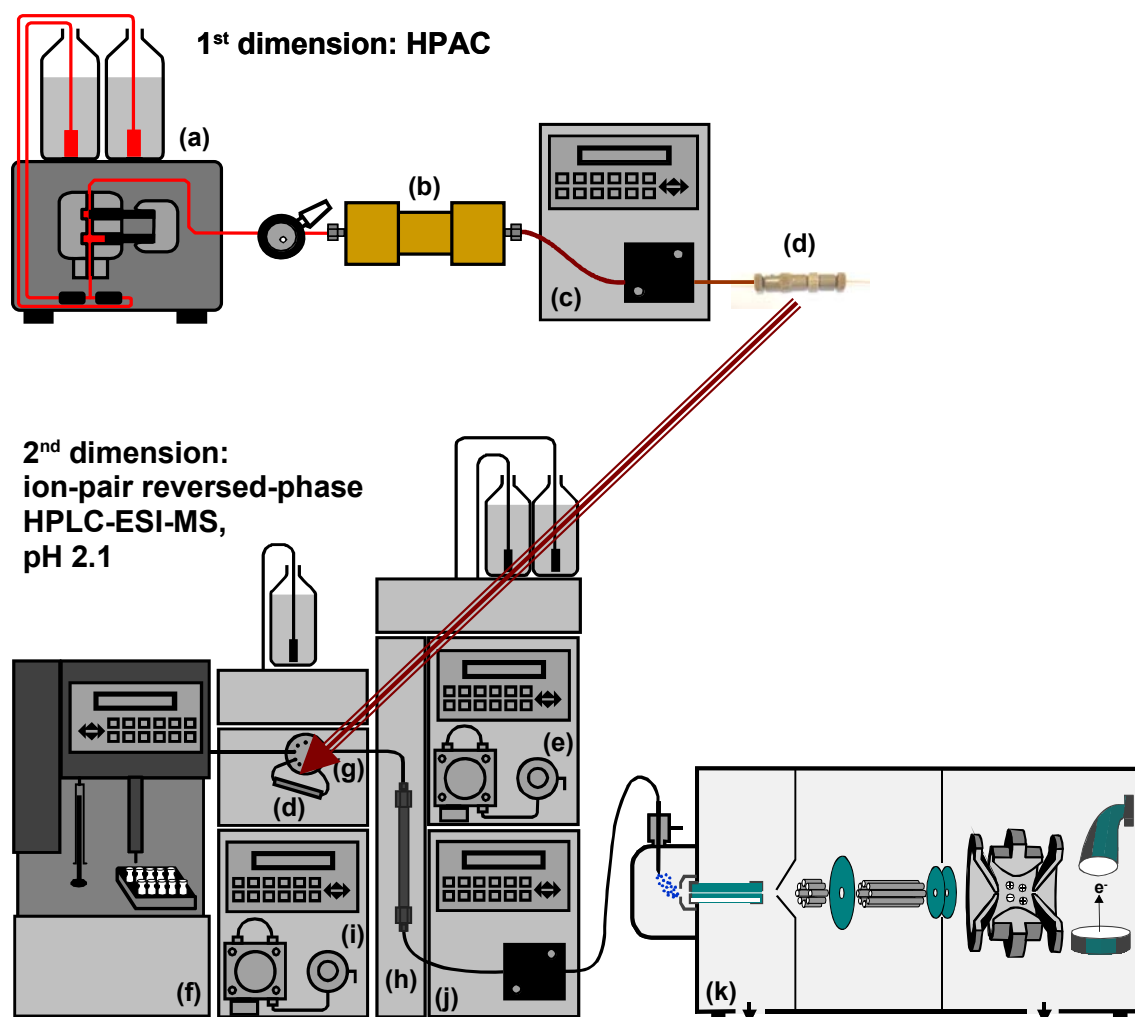


Fig. 23. Instrumental setup for off-line, two-dimensional biomarker detection by HPAC x IP-RP-HPLC-ESI-MS.

(a) Pumping system for affinity enrichment; (b) affinity CIM disk; (c) UV detector for monitoring the first dimension; (d) 10 x 0.20 mm i.d. monolithic trap column trap column; (e) pumping system for IP-RP separation; (f) autosampler; (g) 10-port switching valve; (h) 60 x 0.10 mm i.d. monolithic separation column; (i) pump for loading and washing; (j) UV detector for monitoring the second dimension; (k) electro-spray-ion trap mass spectrometer.

Fractions were collected after the UV detector (c) and injected into the second dimension with the autosampler (f), or NT-proBNP was trapped on a column (d) finally mounted on a 10-port switching valve (g).

Eluents were degassed with helium. The trap column permitted to inject high sample volumes (e.g. 10 μL) without peak broadening. It also permitted the analysis of samples containing traces of sodium chloride without failing the source of the mass spectrometer. An ion-trap mass spectrometer (model esquire HCT) from Bruker Daltonics (Bremen, Germany) with a modified ESI-ion source (spray capillary: fused silica capillary, 0.090 mm o.d., 0.020 mm i.d.) was utilized as detector.

The instrument was operated in MS mode. MS spectra were recorded in positive ion mode with an electrospray voltage of 3,500 V. The heated capillary temperature was set to 300°C. The following mass spectrometric parameters were applied: ultra scan 50 – 3,000 m/z ; scan speed, 26,000 m/z per s; full scan, 500 – 1,500 m/z ; ion polarity, positive; trap drive, 93.2; octapole RF amplitude, 88.5 Vpp; lens 2, -36.1 V; capillary exit, 253.8 V; nebulizer gas, 20 psi; dry gas, 4 L/min; end-plate high-voltage offset, -500 V; ICC target, 70,000; maximum accumulation time, 200 ms.

A schematic representation of the setup is depicted in Fig. 23.

2.2.3 Detection and quantitation of NT-proBNP with immunoassays

Fractions collected after affinity-CIM disks were analyzed with immunoassays performed with an Elecsys 2010 instrument (Roche Diagnostics, Penzberg, Germany). Detection and quantitation are based on an electrochemiluminescence immunoassay (ECLIA) ^[71]. Biotinylated polyclonal antibodies against epitope (1-21) are bound through a streptavidin-biotin complex to magnetic beads. These antibodies are used to catch NT-proBNP. Polyclonal antibodies against epitope (39-51) are covalently bound to a tris(2,2'-bipyridyl)ruthenium(II)-complex and are employed for detection. The reaction mixture is aspirated into a measuring cell, where the microparticles are magnetically captured onto the surface of an electrode. The application of a voltage to the electrode induces a chemiluminescent emission, which is measured by a photomultiplier. The results are determined via a calibration curve. NT-proBNP concentrations can be measured between 0.6 amol/ μL (5 fg/ μL) and 4,130 amol/ μL (35,000 fg/ μL). If the concentration of NT-proBNP is out of range, the sample is diluted with human NT-proBNP depleted serum ^[117, 121].

2.3 Preparation of affinity CIM disks

Two different schemes were followed for the preparation of affinity CIM disks. Antibodies specific to biomarkers were either directly bound to the monolithic support or immobilized via the formation of a streptavidin-biotin complex^[105;122].

2.3.1 Preparation of anti-myoglobin- and anti-NT-proBNP CIM disks via direct immobilization

An epoxy-CIM disk was mounted into the column housing and washed for 15 min at a flow rate of 70 $\mu\text{L}/\text{min}$ with bidistilled water to eliminate the storage liquid, ethanol-water (20/80). Then, the epoxy-CIM disk was equilibrated with 0.50 mol/L sodium phosphate, pH 8.0, for 25 min at a flow rate of 70 $\mu\text{L}/\text{min}$. Antibody solutions (400 μL of 5.0 mg/mL anti-myoglobin antibody in 150 mmol/L sodium chloride solution or 3,000 μL 3.34 mg/mL in 166 mmol/L potassium phosphate, pH 8.5, for anti-NT-proBNP) were pumped through the epoxy-CIM disk at a flow rate of 15 $\mu\text{L}/\text{min}$. The effluent was collected in a glass tube positioned at the column outlet. Then, the affinity-CIM disk was removed from the housing, immersed in 3 mL of the collected antibody solution and stored at room temperature overnight. A schematic representation of the immobilization of the antibodies on the stationary phase is depicted in Fig. 24. The opening of the epoxy-ring is achieved by a nucleophilic attack of an amine group of the antibodies.

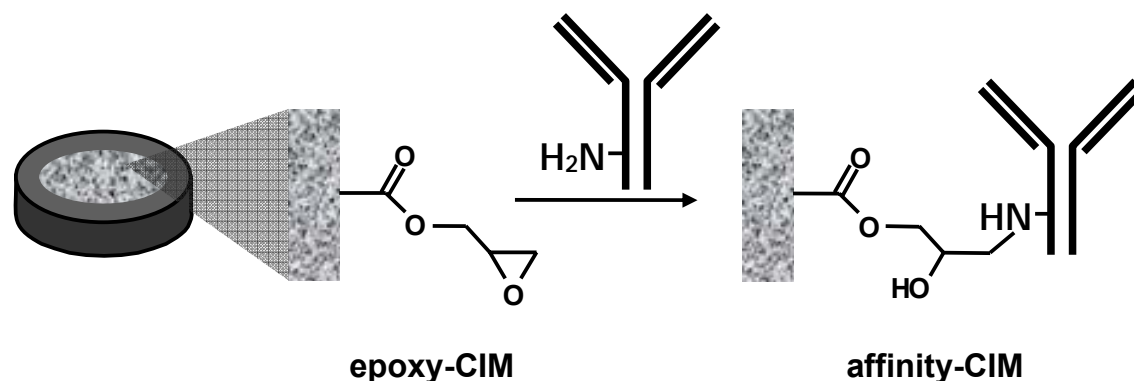


Fig. 24. Preparation of monolithic immunoadsorbents: direct immobilization of antibodies on an epoxy-activated CIM disk.

The anti myoglobin- or anti-NT-proBNP-CIM disk was washed from non-bound antibody molecules with 1,500 μL of 0.5 mol/L sodium phosphate, pH 8.0 (70

$\mu\text{L}/\text{min}$). Housing and connecting tubes were also washed with sodium phosphate, pH 8.0 and purged with air. After antibody immobilization, the anti-NT-proBNP- and anti-myoglobin-CIM disks were washed with 1,500 μL of 0.50 mol/L sodium phosphate, pH 8.0 at a flow rate of 70 $\mu\text{L}/\text{min}$. To block residual epoxy groups, 5.0 mL of 1.0 mol/L ethanolamine were pumped through the column at a flow rate of 70 $\mu\text{L}/\text{min}$. The CIM disk was immersed in 1.0 mol/L ethanolamine solution overnight at room temperature. Subsequently, the disk was washed with 2.5 mL of 0.50 mol/L sodium phosphate, pH 8.0, 1.0 mol/L NaCl at 70 $\mu\text{L}/\text{min}$, followed by 4.0 mL of 0.5 mol/L sodium phosphate, pH 8.0, at 70 $\mu\text{L}/\text{min}$, and finally with 2.5 mL 30 mmol/L sodium chloride at 70 $\mu\text{L}/\text{min}$.

2.3.2 Preparation of affinity CIM disks via streptavidin-biotin anchorage

The epoxy-CIM disk was mounted into the housing and washed for 15 min at a flow rate of 70 $\mu\text{L}/\text{min}$ with bidistilled water to eliminate traces the storage liquid, ethanol-water (20/80). The epoxy-CIM disk was equilibrated for 25 min at a flow rate of 70 $\mu\text{L}/\text{min}$ with 0.50 mol/L sodium phosphate, pH 8.0. Then, 3 mL of a 1.67 mg/mL streptavidin solution were pumped through the epoxy-CIM disk at a flow rate of 15 $\mu\text{L}/\text{min}$. The effluent was collected in a glass tube positioned at the column outlet. Then the CIM disk was removed from the housing, immersed in 3 mL of the collected streptavidin solution, and stored at room temperature overnight.

The streptavidin-CIM disk was washed from non-bound streptavidin molecules with 1,500 μL of 0.50 mol/L sodium phosphate, pH 8.0 (70 $\mu\text{L}/\text{min}$). Housing and connecting tubes were also washed with sodium phosphate, pH 8.0 and purged with air. A biotinyl-anti-NT-proBNP antibody solution consisting of 0.50 mg antibodies present in 500 μL 50 mmol/L potassium phosphate, 150 mmol/L sodium chloride, pH 7.0 was prepared. The absorbance of the antibody solution was measured in quintuplicate at 280 nm using a UV-Vis spectrophotometer (Model DU 7400, Beckman, Fullerton, CA). The solution was then infused over the streptavidin-CIM disk at 10 $\mu\text{L}/\text{min}$. The solution eluting at the outlet of the column was collected and its absorbance was measured in quintuplicate. This procedure was repeated three times. The total absorbance was reduced by about 80 % (see Fig. 26), meaning that 80 % of the antibodies may have reacted with streptavidin. In total, approximately

0.40 mg of antibody were fixed on the disk. It corresponds to 17.4 mg/g polymer material. These values are in accordance with values published elsewhere ^[103,123]. A schematic representation of the immobilization of the antibodies on the stationary phase is depicted in Fig. 25. The epoxy-ring opening is achieved by a nucleophilic attack of an amine group of streptavidin.

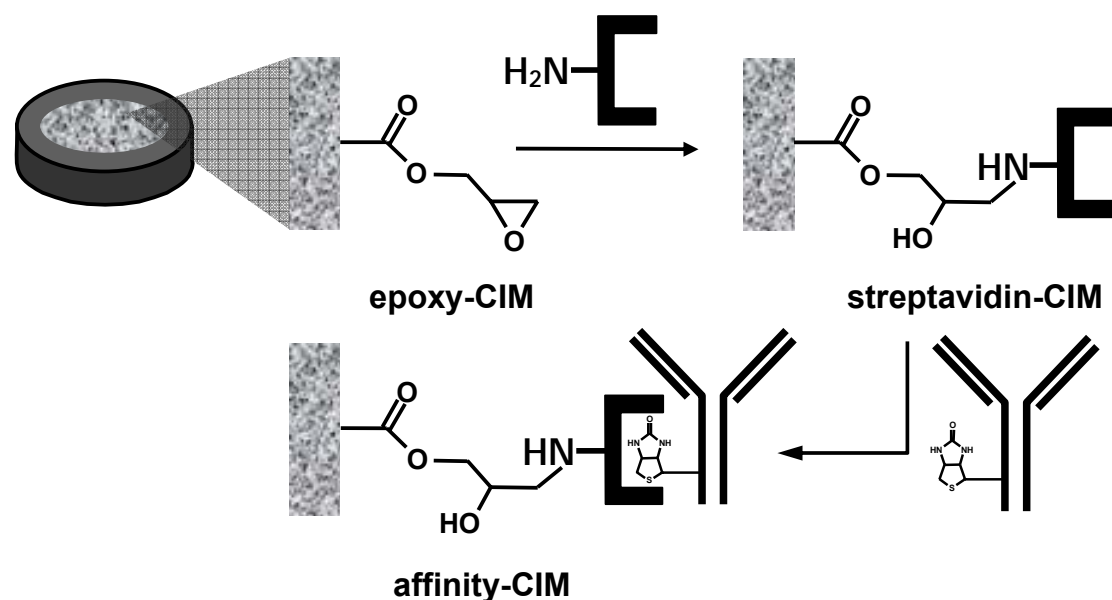


Fig. 25. Preparation of monolithic immunoadsorbents: immobilization of antibodies on an epoxy-activated CIM disk via streptavidin-biotin complex.

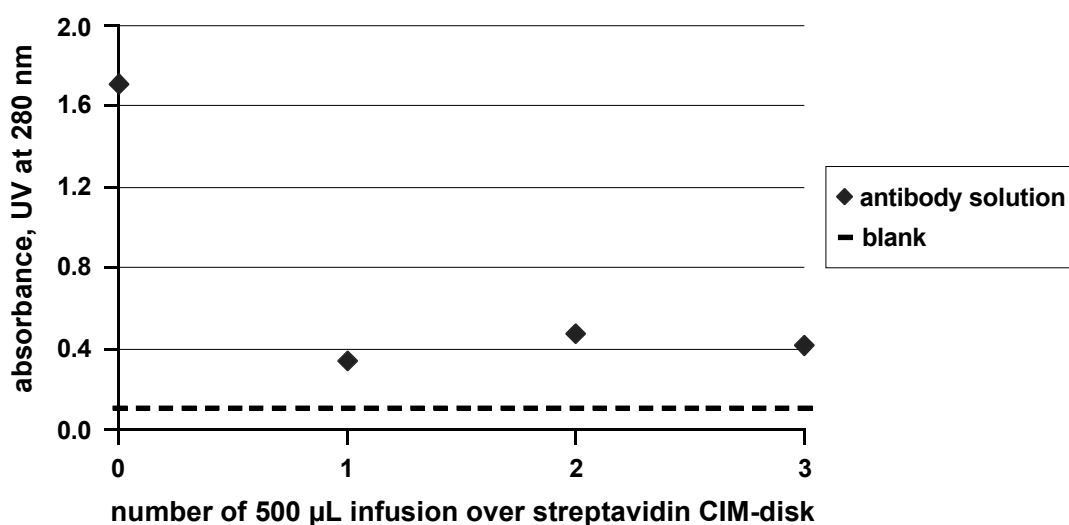


Fig. 26. Absorbance decrease of the biotinyl-anti-NT-proBNP antibody solution after infusion over the streptavidin CIM-disk. Detection with UV at 280 nm.

After antibody immobilization, the biotinyl-anti-NT-proBNP-streptavidin-CIM disk was washed with 1,500 μL of 0.50 mol/L sodium phosphate, pH 8.0 at a flow rate of 70 $\mu\text{L}/\text{min}$. To block residual epoxy groups, 5.0 mL of 1.0 mol/L ethanolamine were pumped through the column at a flow rate of 70 $\mu\text{L}/\text{min}$. The CIM disk was immersed in the 1.0 mol/L ethanolamine solution overnight at room temperature. Subsequently, the disk was washed with 2.5 mL of 0.50 mol/L sodium phosphate, pH 8.0, 1.0 mol/L NaCl at 70 $\mu\text{L}/\text{min}$, followed by 4.0 mL of 0.50 mol/L sodium phosphate, pH 8.0, at 70 $\mu\text{L}/\text{min}$, and finally with 2.5 mL 30 mmol/L sodium chloride at 70 $\mu\text{L}/\text{min}$.

3 Isolation of myoglobin from human serum by affinity chromatography

3.1 Isolation of myoglobin at high concentration

In 30 μL of myoglobin depleted serum, 75 pmol human myoglobin were spiked. The concentration of human myoglobin in serum was consequently 2.5 pmol/ μL . This value was far above biologically relevant values (0.5 to 60 fmol/ μL), but allowed trouble-free detections of myoglobin with MS even in large fraction volumes.

The sample was diluted up to 150 μL with 150 mmol/L NaCl and injected over the anti-myoglobin-CIM disk at 30 $\mu\text{L}/\text{min}$. After 85 min, the pump was switched to deliver 5 % CH_3COOH at 120 $\mu\text{L}/\text{min}$ and 120- μL fractions were collected every minute. Finally the column was re-equilibrated for further injections by pumping 150 mmol/L NaCl at 30 $\mu\text{L}/\text{min}$. The chromatogram is depicted with dashed lines in Fig. 27.

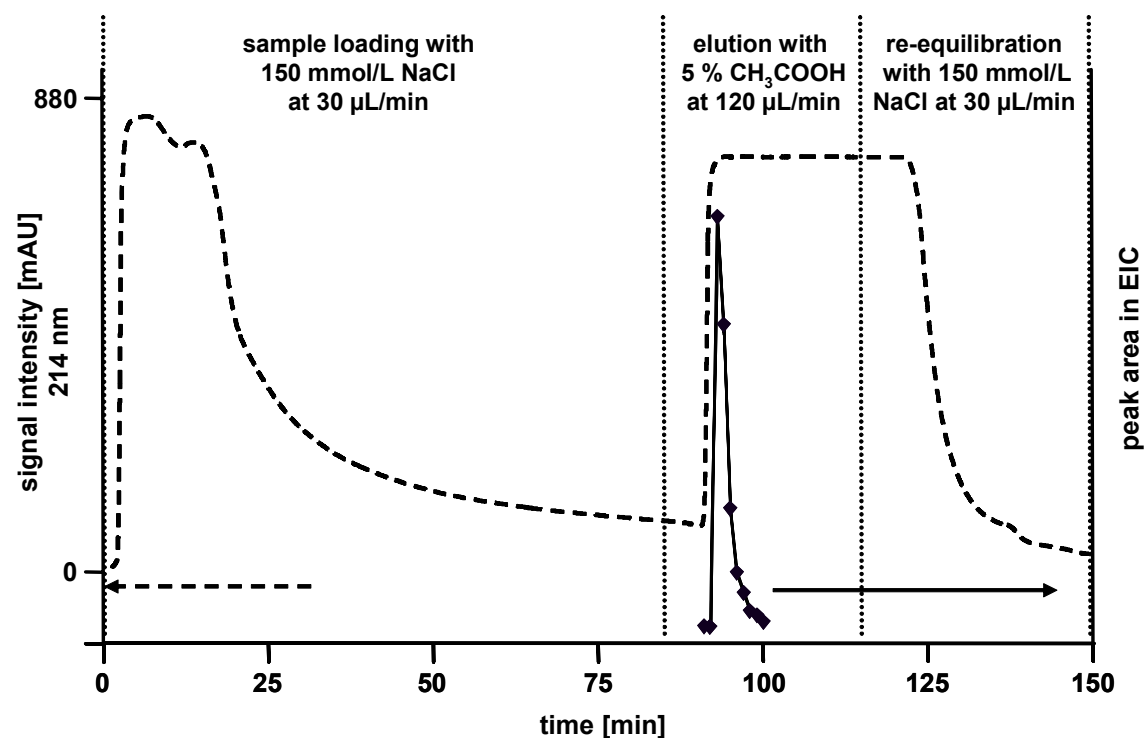


Fig. 27. Elution profile of a serum sample spiked with 2.5 pmol/ μL myoglobin. 85 min loading on mini CIM disk IV: mobile phase, 150 mmol/L NaCl; flow rate, 30 $\mu\text{L}/\text{min}$; room temperature; sample, 75 pmol human myoglobin spiked in 30 μL human serum and diluted to 150 μL with 150 mmol/L sodium chloride. 30 min elution: mobile phase, 5 % CH_3COOH ; flow rate, 120 $\mu\text{L}/\text{min}$; room temperature

Proteins of serum are flowing through the column at the beginning of the separation.

Elution is performed with acetic acid which explains the increase in intensity observed during the whole elution window.

Ten μL of the collected fractions were then injected without further pre-treatment into the RP-HPLC-ESI-MS system described in section 2.2.2 and equipped with a 10 x 0.2 mm PS-DVB trap- and a 60 x 0.2 mm PS-DVB analytical column. A representative reconstructed total ion current chromatogram and a mass spectrum corresponding to myoglobin are depicted in Fig. 28 and in Fig. 29.

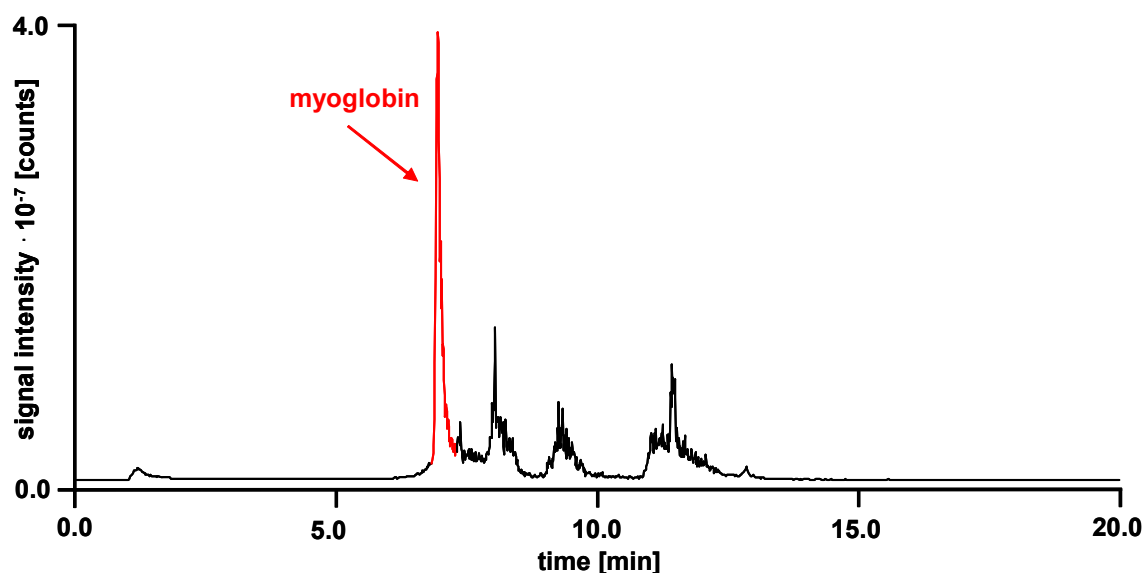


Fig. 28. Reconstructed total ion current chromatogram of a fraction collected after mini CIM disk IV corresponding to the elution of human myoglobin with $\text{H}_2\text{O}/\text{CH}_3\text{COOH}$ 95:5 (v/v). 3-min trapping: column, PS-DVB, 10 x 0.2 mm i.d.; mobile phase, 0.10 % HFBA in H_2O , isocratic elution; flow rate, 10.0 $\mu\text{L}/\text{min}$; room temperature; sample, 10 μL of fraction 92-93' collected after the anti-myoglobin-CIM disk. RP-HPLC-ESI-MS analysis: column, PS-DVB, 60 x 0.2 mm i.d.; mobile phase, (A) 0.050 % TFA in H_2O , (B) 0.050 % TFA in ACN; linear gradient, 0-60 % B in 9.0 min; flow rate, 2.5 $\mu\text{L}/\text{min}$; temperature, 55°C.

Peak areas corresponding to intact human myoglobin were calculated from extracted ion chromatograms of m/z 1004.0, 948.3, 898.4, 853.6, 813.0, and 776.1 with m/z widths of ± 0.5 . These m/z correspond to intact myoglobin 17 to 22 times protonated, respectively. Before integration, peaks were smoothed using a Gauss algorithm and a smoothing width of 1.5 s. The resulting elution curve is depicted with solid lines in Fig. 27. A very sharp elution is observed, proving that myoglobin molecules were retained on the anti-myoglobin-CIM disk. In order to evaluate the recovery of human myoglobin after the monolithic immunoadsorber, a reference solution of myoglobin

(312 fmol/ μL in 5 % CH_3COOH) was injected in the RP-HPLC-ESI-MS setup. By comparing the peak areas computed for the fractions collected after the anti-myoglobin-CIM disk and the peak area computed for the reference, it was possible to calculate the myoglobin recovery after the monolithic immunoadsorber. A myoglobin yield of 32 % was computed. Such a recovery is acceptable but may not be sufficient for the analysis of low-concentrated samples. Two reasons may explain this value. First, the capacity of the immunoadsorber to bind myoglobin may be not sufficient enough for the high amount of myoglobin injected. The other reason is that the antibodies may be not well suited for affinity chromatography (e.g. inappropriate binding constants).

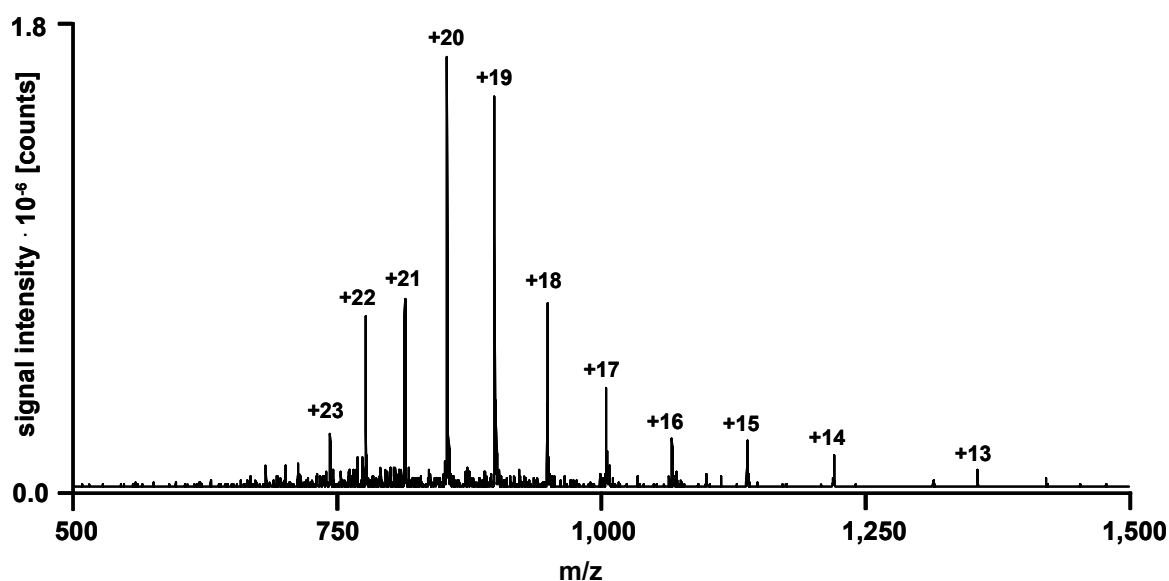


Fig. 29. Mass spectrum of human myoglobin observed at 7.0 min. Conditions as in Fig. 28. The different charge states of myoglobin are indicated.

3.2 Isolation of myoglobin at low concentration

In order to check if the anti-myoglobin-CIM disk can be utilized for real samples, a similar experiment as before was performed with a myoglobin concentration of 250 fmol/ μL in serum. Thus, 75 pmol myoglobin were spiked into 300 μL myoglobin-depleted serum. The sample was diluted up to 600 μL with 150 mmol/L NaCl and injected over the anti-myoglobin-CIM disk at 50 $\mu\text{L}/\text{min}$. After 70 min, the pump was switched to deliver 5 % CH_3COOH at 120 $\mu\text{L}/\text{min}$ and eight 120- μL fractions corresponding to the elution of myoglobin were collected.

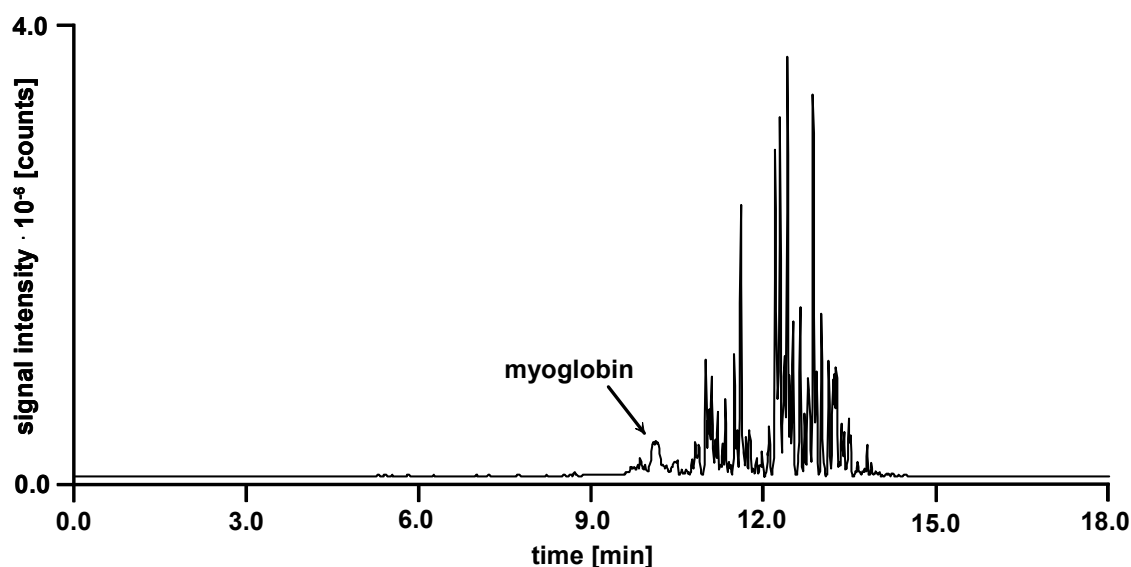


Fig. 30. Extracted ion chromatogram of a fraction collected after mini CIM disk IV corresponding to the elution of human myoglobin with $\text{H}_2\text{O}/\text{CH}_3\text{COOH}$ 95:5 (v/v). 4-min trapping: column, PS-DVB, 10 x 0.2 mm i.d.; mobile phase, 0.10 % HFBA in H_2O , isocratic elution; flow rate, 10.0 $\mu\text{L}/\text{min}$; room temperature; sample, 10 μL of fraction 74-75' collected after the anti-myoglobin-CIM disk. RP-HPLC-ESI-MS analysis: column, PS-DVB, 60 x 0.1 mm i.d.; mobile phase, (A) 0.050 % TFA in H_2O , (B) 0.050 % TFA in ACN; linear gradient, 0-60 % B in 9.0 min; flow rate, 0.8 $\mu\text{L}/\text{min}$; room temperature.

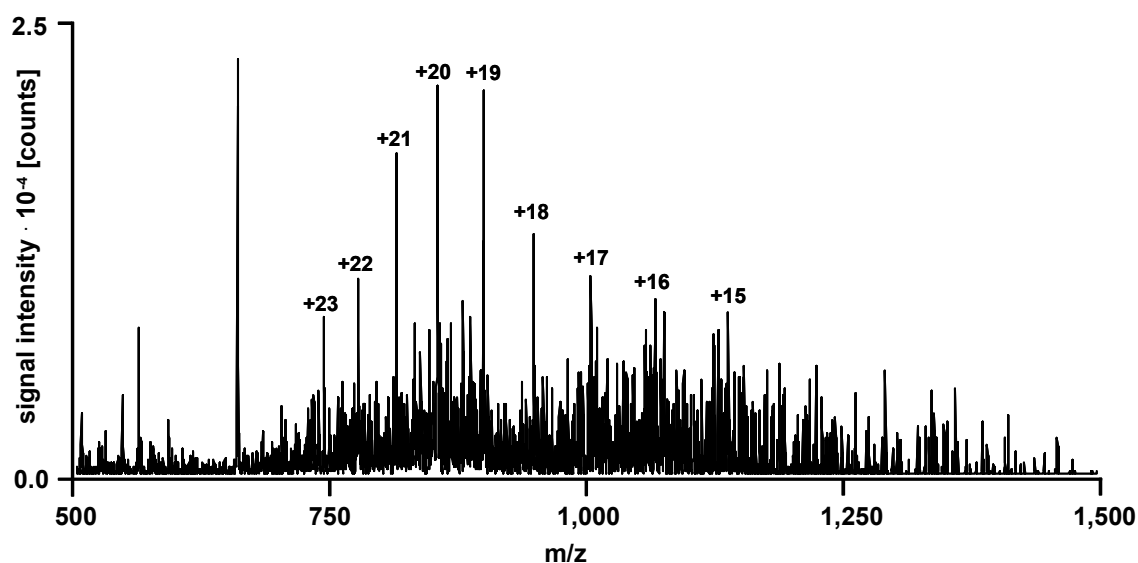


Fig. 31. Mass spectrum of human myoglobin observed at 10.0 min. Conditions as in Fig. 30. The different charge states of myoglobin are indicated.

Ten μL of the collected fractions were then injected without further pre-treatment into the RP-HPLC-ESI-MS system. Myoglobin was detected in the fractions. However, as it can be observed in the extracted ion chromatogram (m/z 1219.0, 1137.8, 1066.8, 1004.1, 948.3, 898.4, 853.6, 813.0, 776.1, 742.4, and 711.5 with m/z widths of ± 0.3 and Gaussian smoothing width of 1.5 s) depicted in Fig. 30, the signal attributed to myoglobin is near the limit of detection of the mass spectrometer. The corresponding mass spectrum is depicted in Fig. 31.

3.3 Conclusions

Anti-myoglobin antibodies were successfully bound to an epoxy-activated monolithic CIM disk. The antibodies still presented affinity to myoglobin once bound to the stationary phase. Myoglobin was selectively isolated from human serum with the developed anti-myoglobin-CIM disk. At high concentration (2.5 pmol/ μL), a recovery of about 32 % was achieved. Myoglobin was successfully isolated and detected from serum samples at concentrations down to 250 fmol/ μL . However, the binding strength of the immunoadsorber is not sufficient for the analysis of real samples. Because the antibodies themselves (and not the amount of antibodies bound on the stationary phase) appeared to be responsible for this low binding capacity, it was decided to focus our work on another biomarker: NT-proBNP.

4 Isolation of NT-proBNP from human serum by affinity chromatography

4.1 Evaluation of the loadability of anti-NT-proBNP-CIM disk

The loadability of anti-NT-proBNP CIM disk was evaluated by frontal analysis ^[124]. The amount of NT-proBNP bondable on the disk was evaluated from breakthrough curves as follows. Anti-NT-proBNP-CIM disk was washed with 30 mmol/L sodium chloride at 10 μ L/min using a syringe pump. The column effluent was monitored using a UV detector with a 45 nL Z-shaped capillary detection cell. For the determination of the column hold-up volume, 5.0 % acetic acid was infused and the strong increase in UV absorption at 214 nm was taken as the column hold-up time (5.25 min). After washing of the column with 500 μ L 30 mmol/L sodium chloride at 70 μ L/min, direct infusion of 50 ng/ μ L (5.84 pmol/ μ L) NT-proBNP solution was performed at 10 μ L/min and the UV signal was observed at 214 nm. The obtained elution profile is depicted in Fig. 32. The inflection point was determined at 9.50 min. The amount of his-tagged NT-proBNP bound on the column was then evaluated to $(9.50 - 5.25) \times 10 \times 5.84 \approx 250$ pmol.

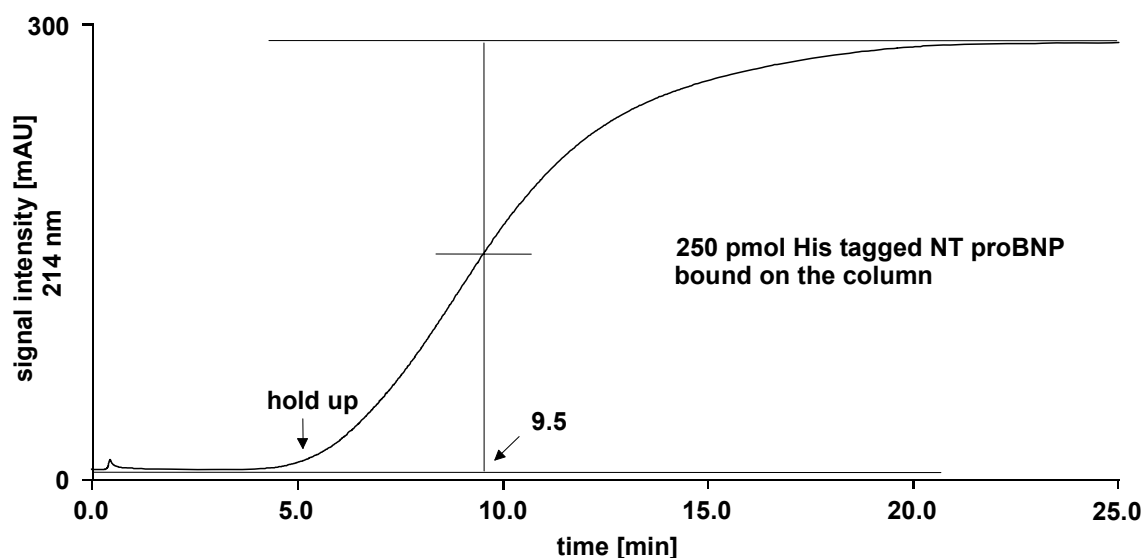


Fig. 32. Breakthrough curve of his-tagged NT-proBNP infused on anti-NT-proBNP-CIM disk. Sample, 50 ng/ μ L (5.84 pmol/ μ L) his-tagged NT-proBNP in 30 mmol/L NaCl; direct infusion; flow rate, 10 μ L/min; room temperature; UV detection, 214 nm. In order to check that NT-proBNP was not retained on the column because of

unspecific interactions with the stationary phase, a 5.84 pmol/ μ L solution of ubiquitin was infused under the same conditions. The inflection point was determined at 5.75 min (see Fig. 33). This value is very similar to the hold-up time of the system (5.25 min). The small difference (only $\sim 10\%$ of the time difference observed by infusing NT-proBNP) is attributed to the imprecision of the method.

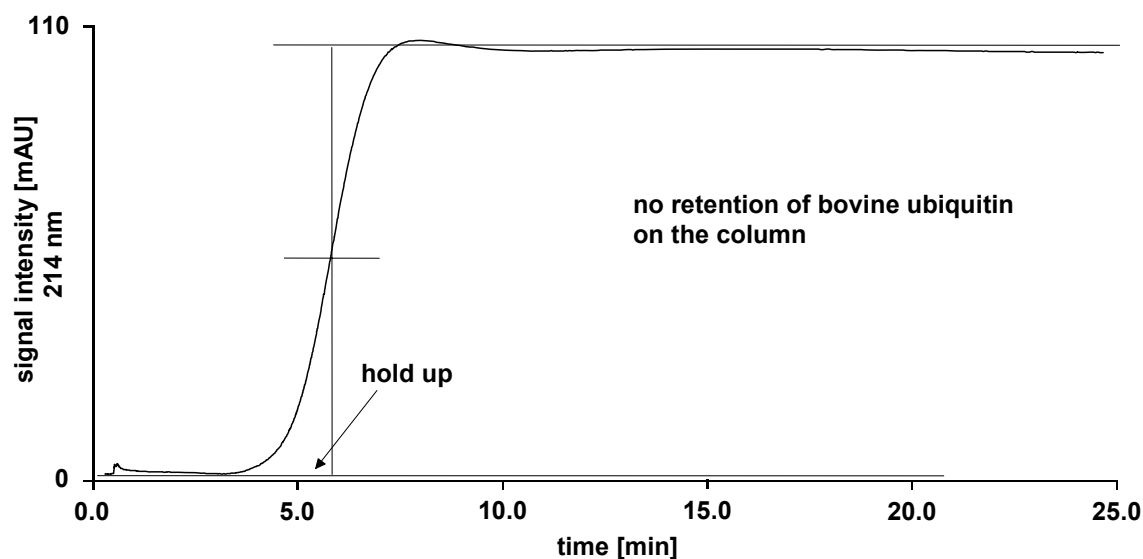


Fig. 33. Breakthrough curve of ubiquitin infused on anti-NT-proBNP-CIM disk. Sample, 5.84 pmol/ μ L ubiquitin in 30 mmol/L NaCl; direct infusion; flow rate, 10 μ L/min; room temperature; UV detection, 214 nm.

4.2 Isolation with CIM disk of NT-proBNP from human serum at 125 fmol/ μ L

In 300 μ L serum, 37.5 pmol synthetic NT-proBNP were spiked. The concentration of synthetic NT-proBNP in serum was consequently 125 fmol/ μ L. The sample was diluted to a final volume of 600 μ L with 150 mmol/L NaCl and was injected over the anti-NT-proBNP-CIM disk. The loading and washing step was performed with 150 mmol/L NaCl at 50 μ L/min for 102 min. Elution was performed with pumping of 100 mmol/L HCl at 120 μ L/min. During the whole chromatographic process, 300- μ L fractions (corresponding to 6.0 min or 2.5 min) were collected. The chromatogram is depicted in Fig. 34.

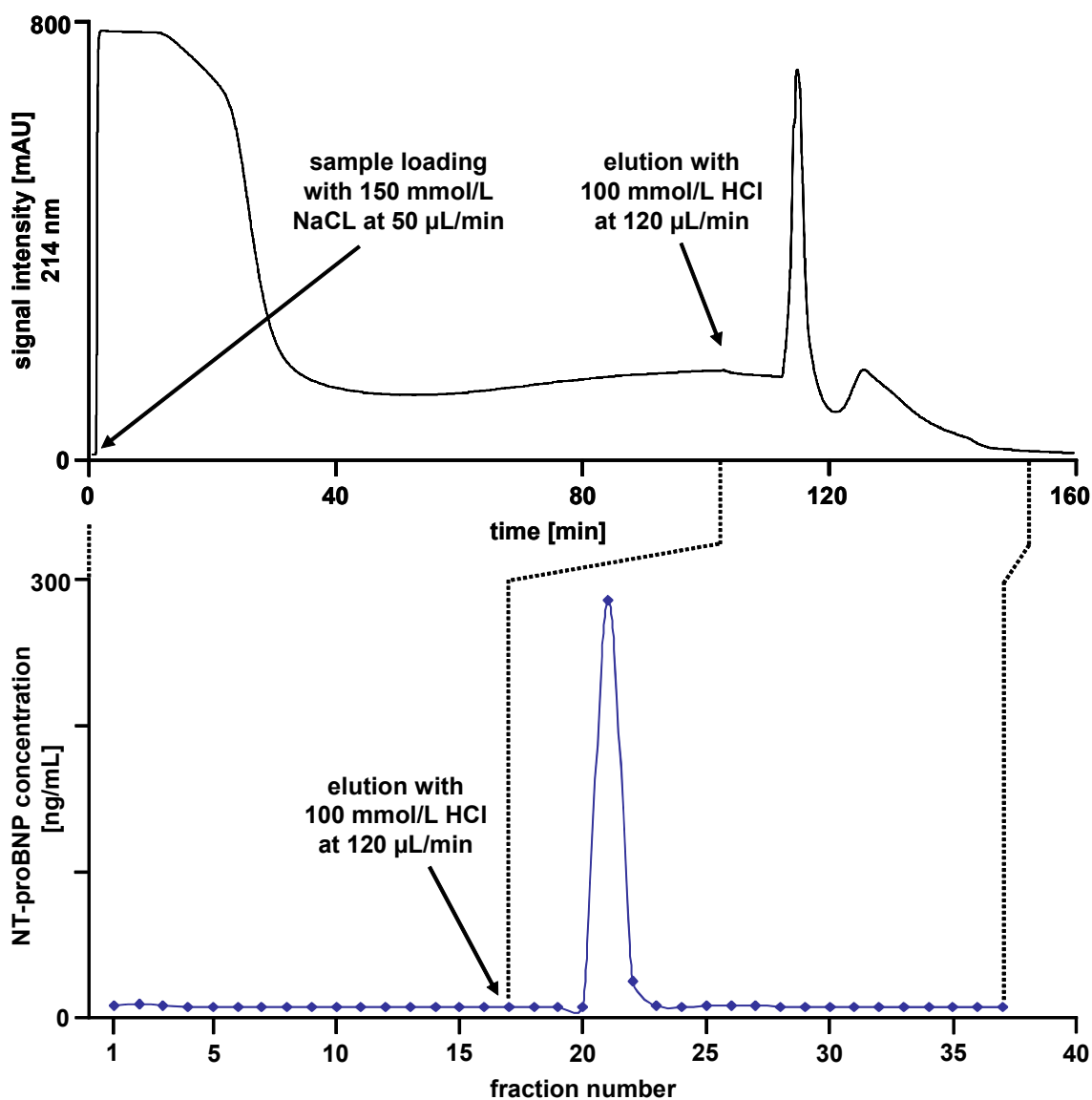


Fig. 34. (up) UV chromatogram of synthetic NT-proBNP spiked in serum and injected over anti-NT-proBNP-CIM disk and (down) elution profile. 102-min loading step: mobile phase, 150 mmol/L NaCl; flow rate, 50 $\mu\text{L}/\text{min}$; room temperature; sample, 37.5 pmol synthetic NT-proBNP spiked in 300 μL serum and diluted to 600 μL in 150 mmol/L NaCl; fractions, 17 \times 6.0 min. Elution: mobile phase, 100 mmol/L HCl; flow rate, 120 $\mu\text{L}/\text{min}$; room temperature; fractions, 20 \times 2.5 min. Concentration of NT-proBNP in the collected fractions was determined with ECLIA.

The analysis of the fractions with ECLIA revealed a very clear elution profile. Most of the synthetic NT-proBNP (92 % of the 37.5 pmol injected) were detected in the 7 first fractions of the elution step. Very low concentrations of NT-proBNP were detected in the flow-through and washing steps. The results are summarized in Tab. 5 and depicted in Fig. 34. These results show the ability of anti-NT-proBNP-CIM disk to

isolate NT-proBNP from serum samples.

Tab. 5. Concentration of synthetic NT-proBNP in 300- μ L fractions collected after anti-NT-proBNP-CIM disk. The concentrations were determined with ECLIA. The loading and washing step is corresponding to fractions # 1-16, and the elution starts with fraction # 17.

Fraction number	NT-proBNP [pg/mL]	Fraction number	NT-proBNP [pg/mL]
1	1,198	20	32
2	2,537	21	31
3	1,075	22	301,960
4	327	23	18,652
5	187	24	598
6	63	25	37
7	41	26	1,050
8	38	27	1,100
9	34	28	717
10	33	29	105
11	33	30	28
12	34	31	30
13	27	32	31
14	32	33	29
15	31	34	29
16	32	35	34
17	32	36	33
18	32	37	29
19	30	38	30

4.3 Isolation with CIM disk of NT-proBNP from human serum at 7.8 fmol/ μ L

In order to check the capacity of the anti-NT-proBNP-CIM disk to isolate NT-proBNP at lower concentrations, 2.35 pmol of synthetic NT-proBNP were spiked in 300 μ L serum (synthetic NT-proBNP concentration of 7.8 fmol/ μ L in serum). Then the sample was diluted to a final volume of 600 μ L with 150 mmol/L NaCl. The solution was injected over the anti-NT-proBNP-CIM disk. The loading and washing step was performed with 150 mmol/L NaCl at 50 μ L/min for 102 min. Elution was performed by pumping 100 mmol/L HCl at 120 μ L/min. The corresponding chromatogram is depicted in Fig. 35.

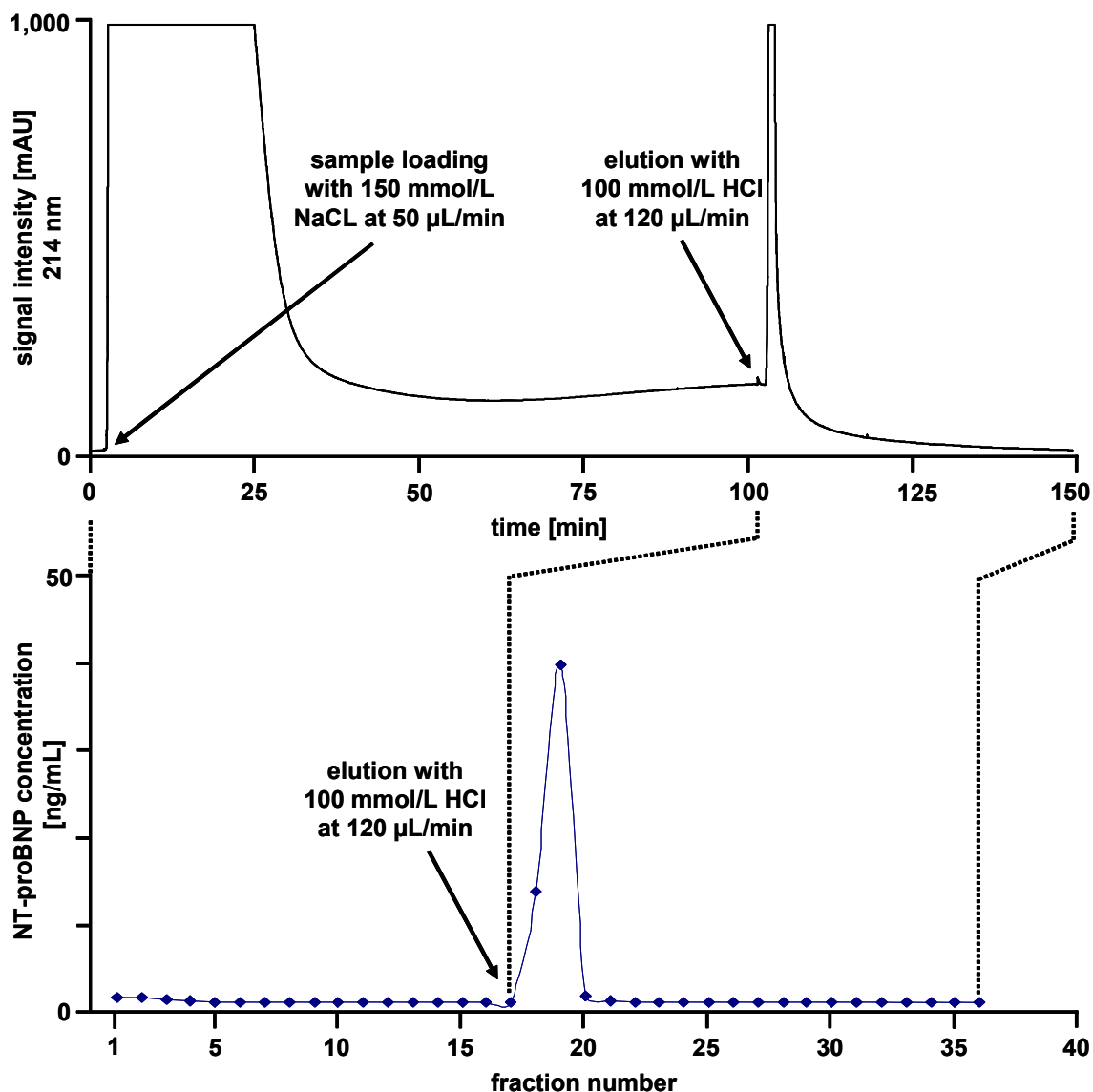


Fig. 35. (up) UV chromatogram of synthetic NT-proBNP spiked in serum and injected over anti-NT-proBNP-CIM disk and (down) elution profile. 102-min loading step: mobile phase, 150 mmol/L NaCl; flow rate, 50 μ L/min; room temperature; sample, 2.35 pmol synthetic NT-proBNP spiked in 300 μ L serum and diluted to 600 μ L in 150 mmol/L NaCl; fractions, 17 x 6.0 min. Elution: mobile phase, 100 mmol/L HCl; flow rate, 120 μ L/min; room temperature; fractions, 19 x 2.5 min. Concentration of NT-proBNP in the collected fractions was determined with ECLIA.

During the separation, fractions were collected and analyzed with ECLIA as in section 4.2. The analysis of the fractions revealed a very clear elution profile. Most of the synthetic NT-proBNP (85 % of the 2.3 pmol injected) were detected in the first 5 fractions of the elution step. Very low concentrations of NT-proBNP were detected in the flow-through and washing steps (see Tab. 6 and Fig. 35).

Tab. 6. Concentration of synthetic NT-proBNP in 300- μ L fractions collected after anti-NT-proBNP-CIM disk. The concentrations were determined with ECLIA. The loading and washing step is corresponding to fractions # 1-16, and the elution starts with fraction # 17.

Fraction number	NT-proBNP [pg/mL]	Fraction number	NT-proBNP [pg/mL]
1	488	20	688
2	561	21	140
3	300	22	61
4	119	23	46
5	82	24	26
6	36	25	14
7	15	26	13
8	12	27	11
9	12	28	11
10	10	29	9
11	13	30	13
12	11	31	10
13	11	32	11
14	10	33	13
15	11	34	11
16	9	35	14
17	9	36	12
18	13,419	37	10
19	40,540		

4.4 Stability and bath-to-batch reproducibility of anti-NT-proBNP-CIM disks

Anti-NT-proBNP-CIM disk showed very good properties in terms of elution profile and sample recovery. However, the implementation of the CIM disk in a routine analysis setup requires a stability of the immunoadsorber and also a good batch-to-batch reproducibility. This was checked by injecting 2.35 pmol synthetic NT-proBNP spiked in 300 μ L serum (7.8 fmol/ μ L in serum) on anti-NT-proBNP-CIM disks from different batches. The samples were diluted to a final volume of 600 μ L with 150 mmol/L NaCl and loaded with 150 mmol/L NaCl at 50 μ L/min for 102 min. The elution was performed by pumping 100 mmol/L HCl at 120 μ L/min. Fractions were collected and analyzed with ECLIA as previously described.

Three anti-NT-proBNP-CIM disks from three different batches were evaluated. Two disks were tested after preparation and another after 18 months. The results are

depicted in Fig. 36. For all three affinity disks very similar elution profiles were obtained. A slight band broadening worsening is observed for the immunoadsorber after 18 months (more than 35 loading/elution cycles) but the elution profile is still well defined. The reproducibility from batch to batch appears excellent. Therefore anti-NT-proBNP-CIM disks are well-suited for high through-put analyses of NT-proBNP in human serum.

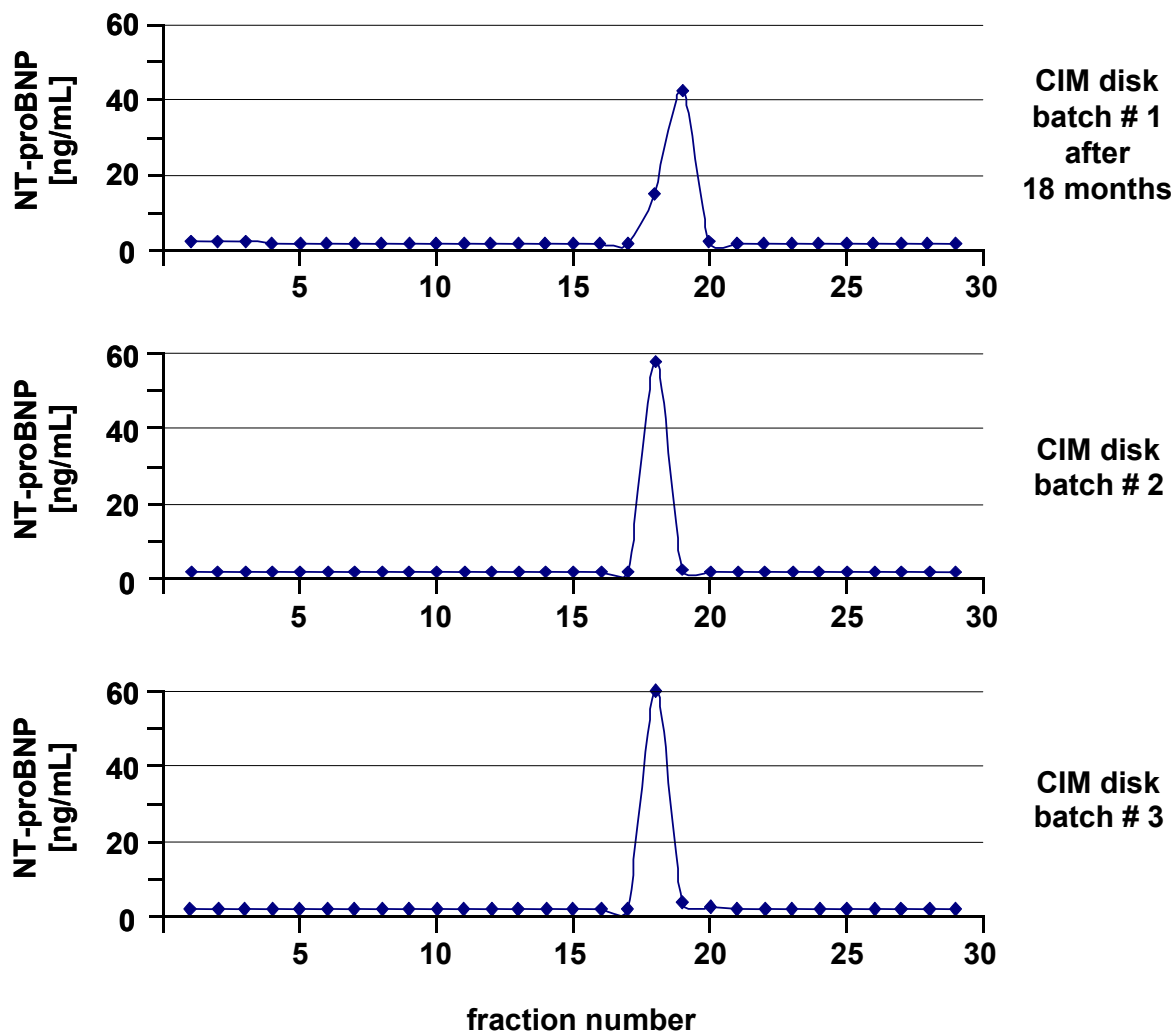


Fig. 36. Elution profile of synthetic NT-proBNP with three anti-NT-proBNP-CIM disks from three different batches. Conditions as in Fig. 35.

4.5 Calibration curve with anti-NT-proBNP-CIM disk

Serum concentrations of NT-proBNP in patients suffering from systolic and diastolic dysfunctions are in the high amol/ μL order of concentration. To test the ability of CIM disk immunoadsorbents to quantitatively extract NT-proBNP from real samples, serum aliquots were spiked with NT-proBNP at concentrations down to 750 amol/ μL , and were injected over an anti-NT-proBNP-CIM disk. NT-proBNP recoveries were evaluated by ECLIA analyses of collected fractions.

After equilibration of the affinity CIM disk with 150 mmol/L NaCl, a blank run consisting of an injection of 300 μL of human serum diluted to 600 μL with 150 mmol/L NaCl was performed. The loading and washing steps were performed with 150 mmol/L NaCl at 50 $\mu\text{L}/\text{min}$ for 102 min. The “elution” was performed by pumping 100 mmol/L HCl at 120 $\mu\text{L}/\text{min}$. 300- μL fractions were collected as in section 4.2 and analyzed with ECLIA.

The results of the ECLIA measurements for the blank injection are depicted in Fig. 37. One observes that the immunoadsorber is not completely free from synthetic NT-proBNP from previous injections. NT-proBNP naturally present in serum may also contribute to this signal. However, the amount of NT-proBNP detected in the elution peak was of about 10 fmol to be compared with 225 fmol ($\sim 4\%$) injected in the system for the first calibration point. It was concluded that the column can be utilized without further washing cycles for the injection of the standards.

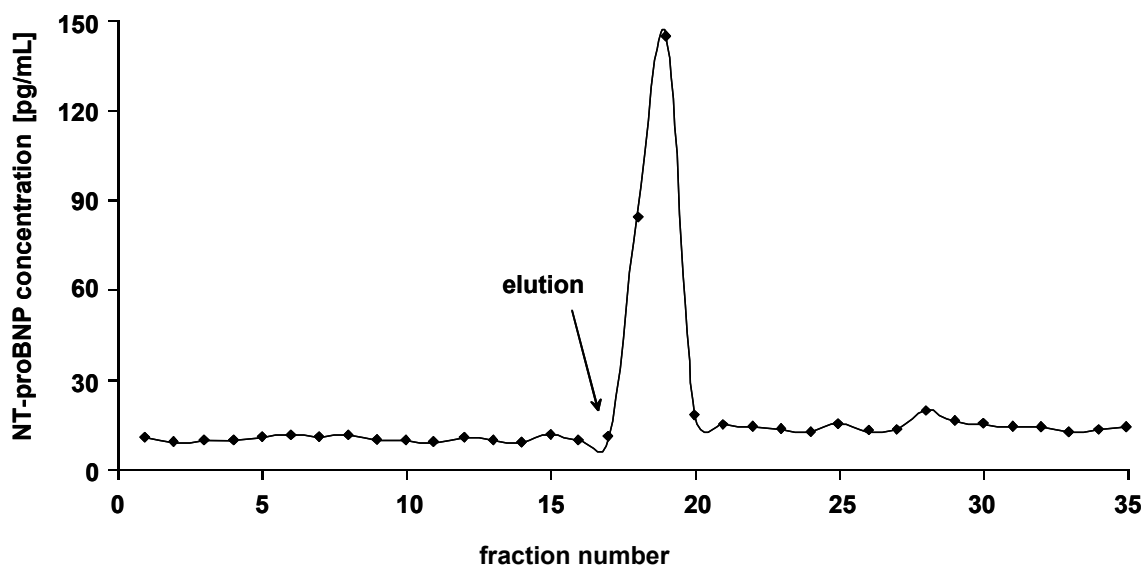


Fig. 37. Blank elution profile of serum with an anti-NT-proBNP-CIM disk. Sample, 300 μL serum diluted to 600 μL in 150 mmol/L NaCl. Other conditions as in Fig. 35.

Under the same conditions, 300 μL from synthetic NT-proBNP serum solutions (synthetic NT-proBNP concentrations: 750 $\text{amol}/\mu\text{L}$, 1,000 $\text{amol}/\mu\text{L}$, 2,000 $\text{amol}/\mu\text{L}$, and 4,000 $\text{amol}/\mu\text{L}$) were injected over the immunoadsorber. The collected fractions were also analyzed with ECLIA. The elution profiles were similar for the four tested concentrations. They are depicted in Fig. 38. Synthetic NT-proBNP is mostly retained on the column during the loading and washing step. In all cases, a sharp elution occurs immediately after the switch to 100 mmol/L HCl. At 4,000 $\text{amol}/\mu\text{L}$, a slight flowing-through is observed in the five first fractions. In all cases, the amount of synthetic NT-proBNP eluting in fractions # 17 to # 22 corresponds to more than 70 % of the total amount of NT-proBNP detected during the chromatographic run. It appears that even at low concentrations, anti-NT-proBNP-CIM disks are very well designed for the enrichment of NT-proBNP from serum.

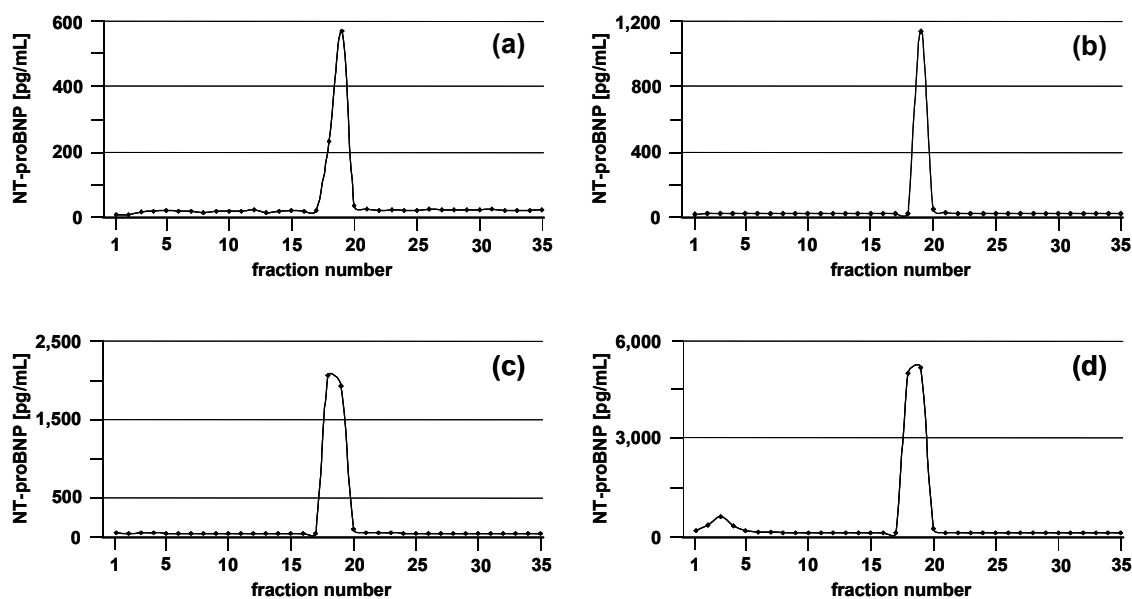


Fig. 38. Elution profiles of synthetic NT-proBNP spiked in serum and injected over anti-NT-proBNP-CIM disk. Samples, 300 μL human serum spiked with synthetic NT-proBNP at (a) 750 $\text{amol}/\mu\text{L}$, (b) 1,000 $\text{amol}/\mu\text{L}$, (c) 2,000 $\text{amol}/\mu\text{L}$, and (d) 4,000 $\text{amol}/\mu\text{L}$. Conditions as in Fig. 35.

A calibration curve, corresponding to the amount of synthetic NT-proBNP detected in the elution peak (fractions # 17 to # 22) as a function of the concentration of synthetic NT-proBNP in human serum was computed and is plotted in Fig. 39.

The observed correlation ($R^2 > 0.99$) proves the ability of anti-NT-proBNP-CIM disks to quantitatively extract synthetic NT-proBNP from human serum at relevant

concentrations for diagnostics purposes. More generally, these results prove the ability of immunoadsorbents based on monolithic supports to quantitatively extract proteins of interest from biological complex matrixes.

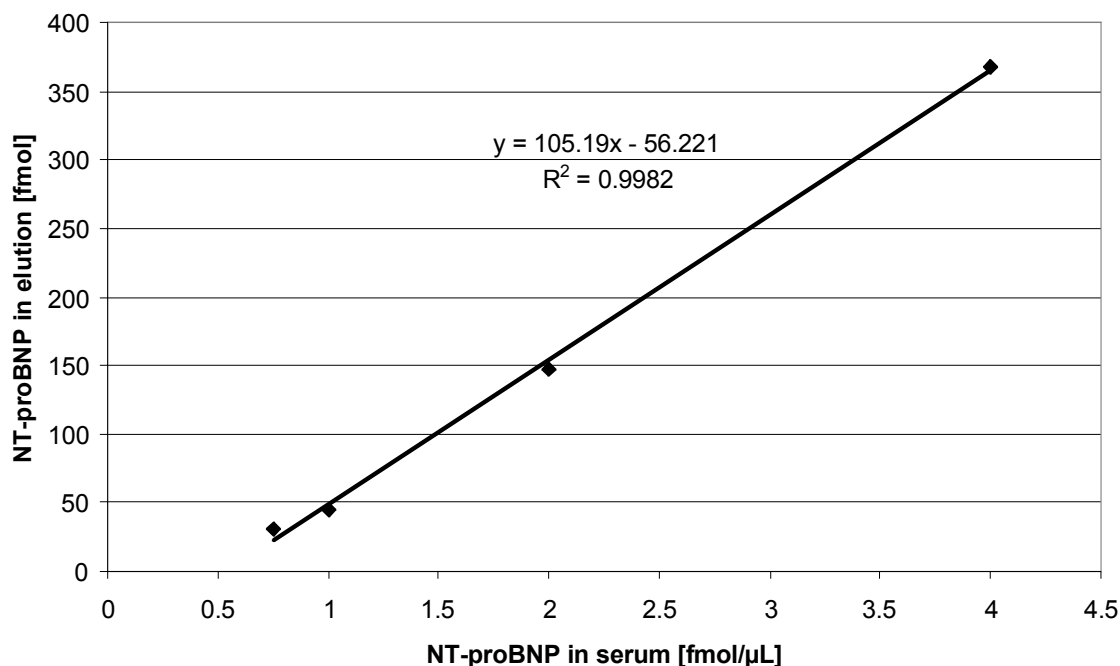


Fig. 39. Synthetic NT-proBNP recovered in elution peak (fractions # 17 to # 22) after anti-NT-proBNP-CIM disk as a function of synthetic NT-proBNP spiked in human serum. A very good linear correlation is observed.

4.6 Hyphenation of anti-NT-proBNP-CIM disk with mass spectrometry

In the previous sections (4.1-4.5), the ability of anti-NT-proBNP-CIM disks to reproducibly and quantitatively extract NT-proBNP from human serum down to 750 amol/μL has been pointed out. However, the detection was performed with ECLIA, a method of detection which requires two new sets of antibodies for each sample (catching and detecting antibodies). In the following section the hyphenation of anti-NT-proBNP-CIM disk with a more generic method of detection, namely mass spectrometry, is investigated. Hyphenation of anti-NT-proBNP-CIM disk with MS was performed by trapping the effluent from the CIM disk onto a small trap column, which was finally mounted in a HPLC-ESI-MS setup. Trapping was investigated with two trap columns: a PS-DVB 5 x 0.2 mm i.d. monolithic-, and a PepmapTM C18 5 x 0.3 mm i.d., 5 μm column.

4.6.1 Hyphenation with PS-DVB monolithic trap column

Synthetic NT-proBNP was spiked into serum to a concentration of 125 fmol/ μ L. 300 μ L spiked serum were diluted up to 600 μ L with 150 mmol/L NaCl and injected over anti-NT-proBNP-CIM disk. The sample was loaded for 60 min with 150 mmol/L NaCl at a flow rate of 50 μ L/min. The washing step was performed with 150 mmol/L NaCl at 500 μ L/min. Then, the elution was performed by hyphenating a monolithic PS-DVB 5 x 0.2 mm column to anti-NT-proBNP-CIM disk and by pumping 100 mmol/L HCl at 15 μ L/min for 90 min. It was not possible to set a higher flow rate, because of the resulting high back pressure (130 bar) incompatible with CIM disk certifications (50 bar). The monolithic column was then decoupled from anti-NT-proBNP-CIM disk and mounted as a trap column in the RP-HPLC-ESI-MS setup (see 2.2.2). Synthetic NT-proBNP was eluted in back-flush from the trap column and finally separated from residual serum proteins with a monolithic PS-DVB 60 x 0.1 mm column with an acetonitrile gradient. Synthetic NT-proBNP was observed in the extracted ion chromatogram (Fig. 40) and in mass spectrum (Fig. 41).

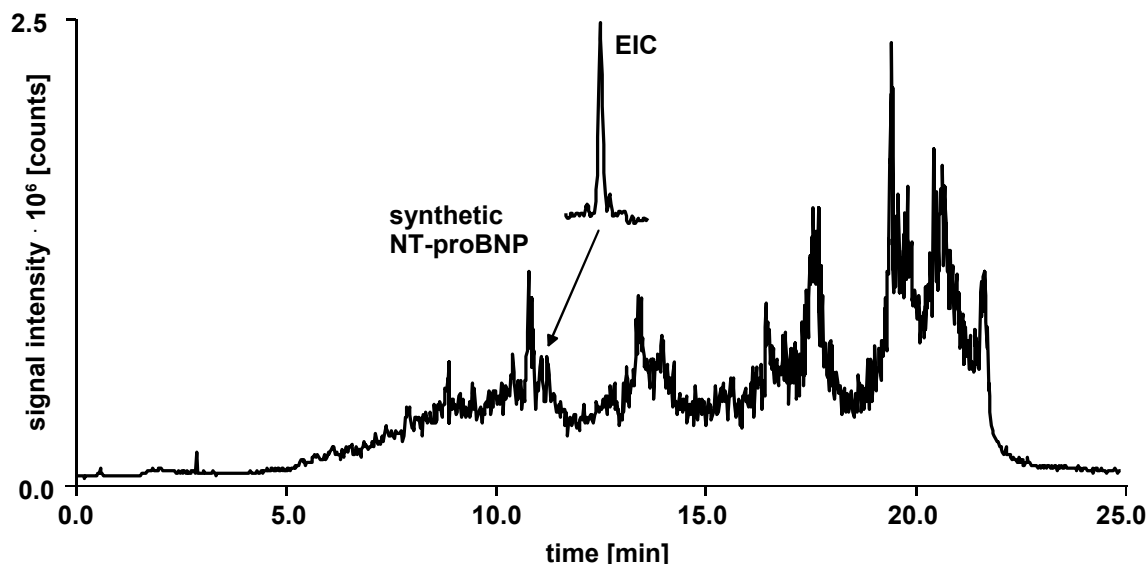


Fig. 40. Reconstructed total ion current chromatogram of synthetic NT-proBNP. Sample, 37.5 pmol synthetic NT-proBNP spiked in 300 μ L serum and diluted to 600 μ L with 150 mmol/L NaCl, injected over anti-NT-proBNP-CIM disk, eluted with 100 mmol/L HCl PS-DVB, 5 x 0.2 mm i.d. column. The column was decoupled from anti-NT-proBNP-CIM disk and implemented as a trap column for RP-HPLC-ESI-MS. Mobile phase, 0.10 % HFBA in H₂O, isocratic elution; flow rate, 10.0 μ L/min; room temperature. RP-HPLC-ESI-MS analysis: column, PS-DVB, 60 x 0.1 mm i.d.; mobile phase, (A) 0.050 % TFA in H₂O, (B) 0.050 % TFA in ACN; linear gradient, 0-80 % B in 20.0 min; flow rate, 0.8 μ L/min; 55°C.

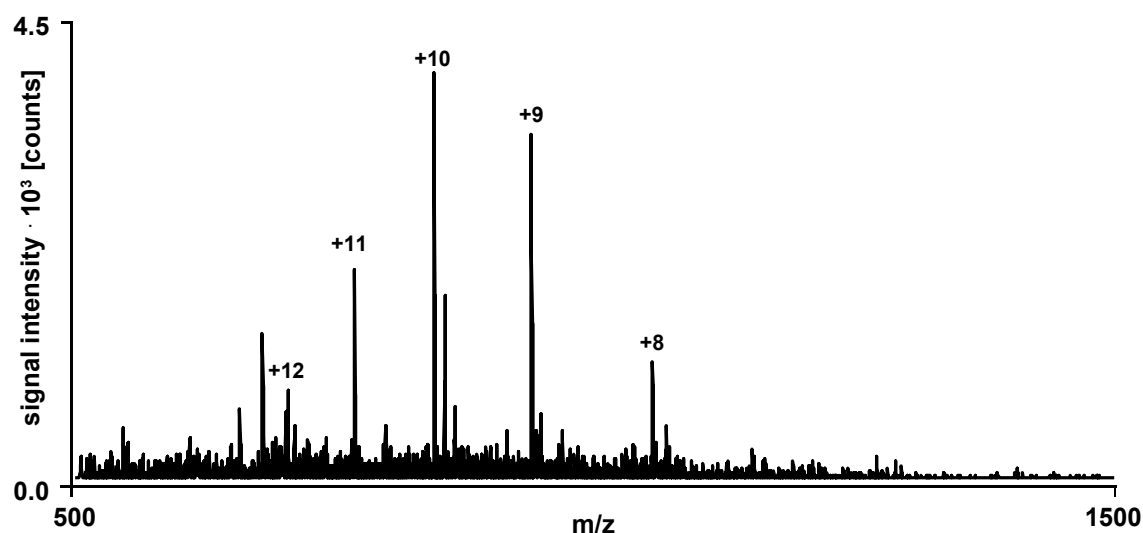


Fig. 41. Mass spectrum of synthetic NT-proBNP. Sample, 125 fmol/ μ L synthetic NT-proBNP in serum; trap column, PS-DVB, 5 x 0.2 mm i.d.; eluents, gradient, flow rate, and temperature as in Fig. 40.

The same experiment was performed with serum samples spiked with synthetic NT-proBNP at lower concentrations. However, NT-proBNP was no longer detected.

4.6.2 Hyphenation with PepmapTM C18 trap column

Hyphenation of anti-NT-proBNP-CIM disk was also investigated with a 5 x 0.3 mm PepMapTM C18 trap column. The higher permeability of the PepmapTM column in comparison with the PS-DVB column, permitted to perform elution from anti-NT-proBNP-CIM disk at flow rates up to 50 μ L/min without leakage from CIM disk housing (110 bar). Synthetic NT-proBNP serum solutions were prepared and injected into the HPLC system as depicted in Fig. 23. The samples were loaded and washed with 150 mmol/L NaCl at 50 μ L/min. After 105 min, a PepmapTM C18 trap column was hyphenated to anti-NT-proBNP-CIM disk and 100 mmol/L HCl was pumped at 50 μ L/min for 30 min. The trap column was then dismantled from CIM disk and implemented as a trap column in the RP-HPLC-ESI-MS system previously described. Synthetic NT-proBNP was eluted in back-flush with an acetonitrile gradient and finally detected with ESI-MS.

Synthetic NT-proBNP concentrations down to 7.8 fmol/ μ L in serum were detected (see Fig. 42 and Fig. 43).

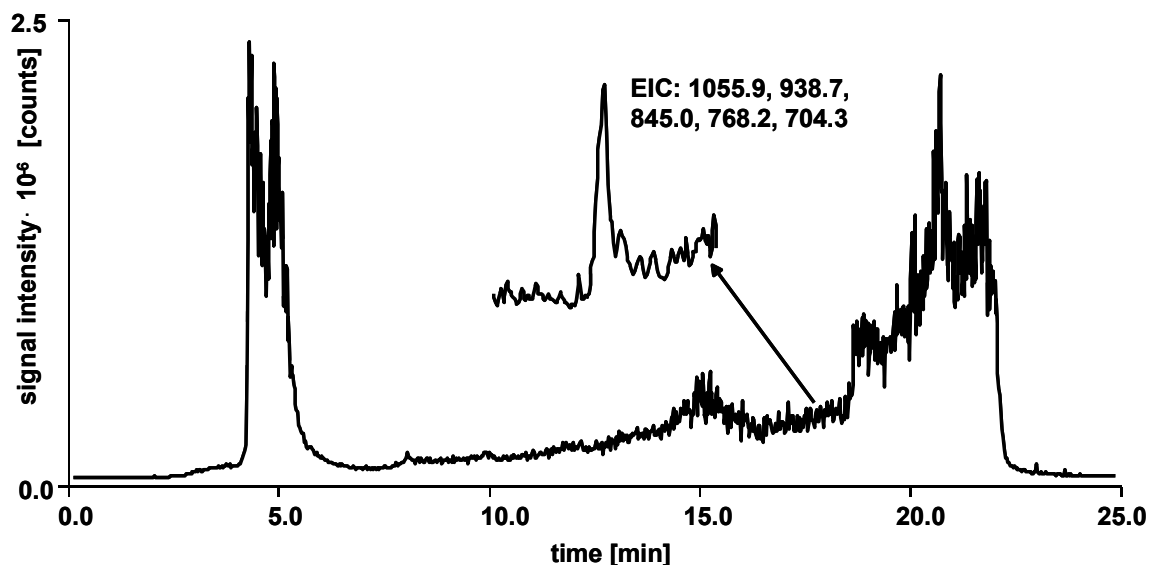


Fig. 42. Reconstructed total ion current chromatogram of synthetic NT-proBNP. Sample, 2.35 pmol synthetic NT-proBNP spiked in 300 μ L serum and diluted to 600 μ L with 150 mmol/L NaCl, injected over anti-NT-proBNP-CIM disk, eluted with 100 mmol/L HCl PepmapTM C18, 5 x 0.3 mm i.d. column. The column was decoupled from CIM disk and implemented as a trap column for RP-HPLC-ESI-MS. Mobile phase, 0.10 % HFBA in H₂O, isocratic elution; flow rate, 10.0 μ L/min; room temperature. RP-HPLC-ESI-MS analysis: mobile phase, (A) 0.050 % TFA in H₂O, (B) 0.050 % TFA in ACN; linear gradient, 0-50 % B in 15.0 min; flow rate, 0.8 μ L/min; room temperature.

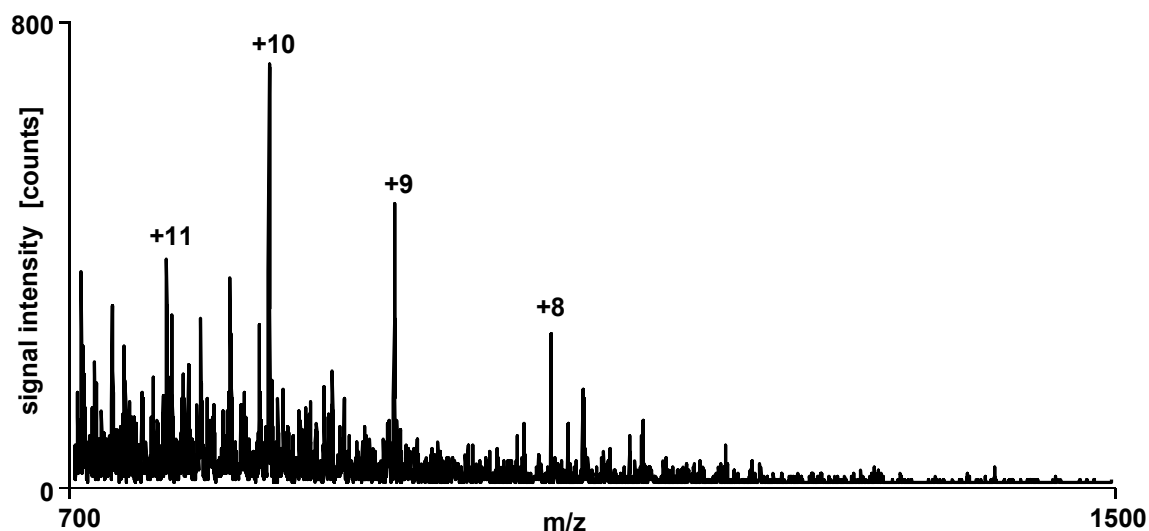


Fig. 43. Mass spectrum of synthetic NT-proBNP. Sample, 7.8 fmol/ μ L synthetic NT-proBNP in serum injected over anti-NT-proBNP-CIM disk and trapped PepmapTM C18, 5 x 0.3 mm i.d. column. eluents, gradient, flow rate, and temperature as in Fig. 42.

4.6.3 Discussion

As shown in Fig. 42 and in Fig. 43, it was possible to detect synthetic NT-proBNP at concentrations down to 7.8 fmol/ μ L in human serum with trapping on PepmapTM, and ESI-MS detection. This corresponds to a gain sensitivity of 16 in comparison with the sensitivity observed with the monolithic PS-DVB 5 x 0.2 mm trap column (see 4.6.1).

The reason of this gain is probably due to the better analyte transfer from anti-NT-proBNP-CIM disk to the PepmapTM C18 trap column (50 μ L/min, no leakage) than from anti-NT-proBNP-CIM disk to the monolithic PS-DVB trap (10-15 μ L/min, critical pressure). The hydrophobicity of the stationary phase may play a minor role. The PepmapTM C18 trap column seems to be the trap column of choice for analyte transfer from CIM disks because of the low produced back pressure.

However, strong synthetic NT-proBNP carry-over was observed with the PepmapTM C18 trap column (data not shown), which is not wished for sample quantitation. As no carry-over was observed with PS-DVB trap columns, a PS-DVB based stationary phase with low back pressure should be the best trap column.

Sample recovery with trapping and RP-HPLC-ESI-MS appears to be smaller than the recovery observed with ECLIA detection (85 %, see 4.3) at similar concentrations. Injection of 300 μ L of synthetic NT-proBNP at 7.8 fmol/ μ L in serum with 85 % recovery correspond to approximately 2.0 pmol synthetic NT-proBNP eluting from CIM disk during the elution process. This value is far above the detection limit of the mass spectrometer (some fmol), and a very intensive signal should be obtained with ESI-MS detection. Different phenomena can be involved in the low recovery of synthetic NT-proBNP from serum samples with trapping and analysis with RP-HPLC-ESI-MS. The following factors can be mentioned:

- insufficient analyte transfer from anti-NT-proBNP-CIM disk to PepmapTM C18 trap column
- insufficient loadability of PepmapTM; especially with serum samples (bleeding of the antibodies, incomplete depletion of high-abundant proteins; serum albumin, ferroxidase, and thrombin have already been detected)
- incomplete elution of NT-proBNP from the PepmapTM trap column; strong carry-overs were already observed
- degradation of NT-proBNP; degradation products of NT-proBNP were already

detected in neutral and acidic solutions; degradations only occur at N- and C-termini and have consequently no impact on ECLIA detection (epitopes for catching and detecting antibodies in the middle of the protein sequence)

- ion-suppression in the ESI source (co-elution of proteins, impurities)

5 Conclusions

Immunoabsorbers based on monolithic epoxy-activated CIM disks have been developed in order to target biomarkers implicated in heart diseases (myoglobin, and NT-proBNP). In both cases, antibodies were successfully bound to the polymeric disk material. The developed immunoabsorbers permitted to selectively isolate myoglobin and NT-proBNP from human serum. Myoglobin was successfully isolated and detected from serum samples at concentrations down to 250 fmol/ μ L. However, the binding capacity of the antibodies was not sufficient for the analysis of clinical samples. Frontal analysis of anti-NT-proBNP-CIM disk revealed the ability of the immunoabsorber to bind up to 250 pmol NT-proBNP. This capacity is highly sufficient for the analysis of clinical samples (some pmol). Anti-NT-proBNP-CIM disks show a very good stability over 18 months, and an excellent batch-to-batch reproducibility has been observed. Anti-NT-proBNP-CIM disks permitted a quantitative isolation of NT-proBNP at concentrations down to 750 amol/ μ L in serum ($R^2 = 0.998$). This concentration corresponds to NT-proBNP concentrations in serum of highly ill patients. Hyphenation of CIM immunoabsorbers with mass spectrometry has been achieved for concentrations down to 7.8 fmol/ μ L.

Chapter IV

Development of an on-line SPE-
HPLC-ESI-MS method for the
analysis of drugs in whole blood
hemolysates

IV. Development of an on-line SPE-HPLC-ESI-MS method for the analysis of drugs in whole blood hemolysates

1 Introduction

For medical diagnostic and therapeutic purposes, determination of drug concentrations in biological fluids is essential. Drugs are generally monitored in blood and its components but also in biological fluids such as saliva, urine, synovial fluid, and cerebrospinal fluid. These fluids are difficult to be analyzed due to their complexity. The complexity is not only due to the numerousness of the analytes but also to their high dynamic range (in terms of concentration) ^[9]. For the analysis of whole blood time consuming sample preparation is required. Drugs of interests are indeed usually bound to erythrocytes and lysis of the erythrocytes is required. Generally, protein precipitation with ZnSO₄/MeOH is performed and followed by centrifugation. Only after these steps of preparation the sample can be injected into a HPLC chain. Such sample preparations are labor-intensive and can not be automated at low costs. They are generally considered as the time-limiting step in the analytical process and are undesirable for routine procedures (especially for analyses with more than 100 samples a day). In order to circumvent these problems, Roche Diagnostics (Penzberg, Germany) has developed a hemolysis reagent permitting to obtain homogeneous hemolysates. After addition of the reagent to whole blood, a clear hemolysate is obtained in few minutes and the sample can be injected without further treatment into a HPLC-MS setup.

In order to address small molecules of interest (e.g. immunosuppressive compounds such as rapamycin), high-abundant proteins such as hemoglobin (120-170 mg/mL) and albumin (60 mg/mL) have to be eliminated from the hemolysate. This can be achieved by performing on-line solid phase extraction (SPE) ^[125-127] before separation and detection with HPLC-ESI-MS. Some experiments have already permitted to detect 40 µg/mL rapamycin in hemolysate by Roche Diagnostics. However, high carry-over, especially of hemoglobin, has been observed in the following injections. Such a carry-over is not compatible with high-throughput quantitative analysis.

In this context, the aim of our work in cooperation with Roche Diagnostics was to develop an on-line SPE-HPLC-ESI-MS method, permitting to target drugs at relevant concentrations. The specifications of the setup were the following:

1. injection of whole blood hemolysate over an “appropriately” designed material
2. elimination of hemoglobin and proteins
3. quantitative analysis of targets (small drugs)
4. no carry-over of hemoglobin in a consecutive blank injection

2 Materials and methods

2.1 Chemicals and instruments

Deionized water (18.2 M Ω cm) was prepared with a Purelab Ultra Genetic system (Elga, Griesheim, Germany). Acetonitrile (E Chromasolv), human hemoglobin (H7379-5G), and N,N'-disuccinimidylcarbonate were purchased from Sigma-Aldrich (Steinheim, Germany). Trifluoroacetic acid ($\geq 99.5\%$), DMSO, 4-(dimethylamino)-pyridin, imidazole ($\geq 99.5\%$), ammonium acetate ($> 99\%$), ethanolamine ($\geq 99.0\%$), and potassium thiocyanate ($\geq 98\%$) were purchased from Fluka (Buchs, Switzerland). Analytical reagent grade sodiumdihydrogen-phosphate-1-hydrate, 1-methyl-octyl-pyrrolidinium chloride, and Tris-HCl were supplied by Merck KGaA (Darmstadt, Germany). Aminodextran (M_w 40,000 Da), 5-methyltetrahydrofolic acid (CAS # 68792-52-9; Sigma M0132), and whole blood samples were obtained from Roche Diagnostics (Penzberg, Germany). Poly-D-lysine hydrobromide (M_w 30,000 – 70,000 Da), carbonic anhydrase from bovine erythrocytes, β -lactoglobulin A and B, human insulin, ubiquitin from bovine erythrocytes, tetracycline hydrochloride (CAS # 64-75-5), human hemoglobin (H7379-5G, batch # 095K7540), and human serum (H1388-20mL, batch # 026K0467) were supplied by Sigma (Schnelldorf, Germany). Ovalbumin, and L-ascorbic acid were purchased from Serva Electrophoresis (Heidelberg, Germany). 1,4-Dioxane was purchased from Riedel de Haen (Seelze, Germany).

Polyethyleneimine 25 kDa and polyethyleneimine 60 kDa were obtained from the workgroup of Prof. Wenz (Department of Organic Macromolecular Chemistry, Saarland University, Saarbrücken, Germany). Polyethyleneimine 25 kDa is available by BASF (Ludwigshafen am Rhein, Germany) under the name Lupasol G 500 as a 40 weight percent aqueous solution (N_{prim.}:N_{sec.}:N_{tert.} = 1:1.1:0.7). Polyethyleneimine 60 kDa is available by Acros (Geel, Belgium) as a 60 weight percent aqueous solution (N_{prim.}:N_{sec.}:N_{tert.} = 1:2:1).

Solution pH was measured with a pH SenTix 61 electrode and a pH-meter pH 537 supplied by WTW (Weilheim, Germany).

Vivaclear mini 0.5 clarifying filter membranes (0.8 μ m PES) were obtained from Vivascience AG (Goettingen, Germany).

2.2 Preparation of blood hemolysates

Four % lysis reagent was prepared by mixing 0.7 g 1-methyl-octyl-pyrrolidinium to 0.3 g potassium thiocyanate in 25 mL H₂O. Blood hemolysates were obtained at neutral conditions by mixing under gentle agitation whole blood and lysis reagent (300 µL whole blood, 2.7 mL 150 mmol/L NaCl, and 3 mL 4 % lysis reagent). After a few minutes the blood sample became clear and ready for injection to the HPLC-MS system.

2.3 Columns

ChromSpher Biomatrix was purchased from Varian, Inc. (Middelburg, The Netherlands). LiChrospher ADS was obtained from Merck KGaA (Darmstadt, Germany). Capcell Pak was supplied by Shiseido Co (Tokyo, Japan) and Bioptic AV-2 was purchased from GL Sciences (Torrance, CA, USA). Biotrap 500 MS was obtained from ChromTech Ltd (Cogleton, United Kingdom) and SPS C18 was purchased from Regis Technologies, Inc (Morton Grove, IL, USA).

2.3.1 Modification of LiChrospher ADS material

LiChrospher ADS bulk material was activated with N,N'-disuccinimidyl-carbonate. 1 g bulk material was weighted in a 50 mL reaction tube and 10 mL water free dioxane was added. 2 mL N,N'-disuccinimidyl-carbonate diluted in DMSO (0.6 mol/L) and 2 mL 4-dimethylaminopyridin in DMSO (0.6 mol/L) were added. The mixture was left for reaction at room temperature under agitation for 6 hours. Finally the mixture was centrifugated (10 min at 5,000 g). The supernatant was removed, and the particles were resuspended in 10 mL of a DMSO-dioxane solution (1:2.5) and agitated for 10 min. The procedure was repeated twice. After a last centrifugation, the succinimidyl-carbonated particles were desiccated overnight (16 hours) at 4°C (vacuum desiccator with sicapent).

The activated particles were then coupled with aminodextran T40. 10 mL of a cold aminodextran T40 solution (10 mg/mL in 0.1 mol/L sodium phosphate buffer, pH 7.5) was added to the previously activated and desiccated particles. The mixture was agitated overnight (16 hours) at 4°C. After centrifugation, the particles were washed in phosphate buffer for 10 min under agitation. The procedure was repeated twice. The modified particles were finally stored in 10 mL 20 mmol/L Tris-HCl buffer, pH 7.0.

2.3.2 Packing of columns

Modified or unmodified particles were packed in stainless steel 30 x 2.0 mm column housings supplied by Bischoff Chromatography (Leonberg, Germany). The particles were suspended and pushed at 3 mL/min with Tris buffer. After 20 mL, the packing solution was switched to acetonitrile/water 80:20 (v/v). The columns were closed at each extremity with two 2 µm stainless steel filters.

2.4 Analytical setups

2.4.1 One-dimensional HPLC setup

The different RAM columns were implemented in a HPLC system from Dionex (Idstein, Germany). Solvent delivery was performed by a P680 HPLC pump. Sample injection was performed by an ASI-100 Automated Sample Injector. Signals were monitored by a UVD340U diode array detector. The analytical setup is depicted in Fig. 44.

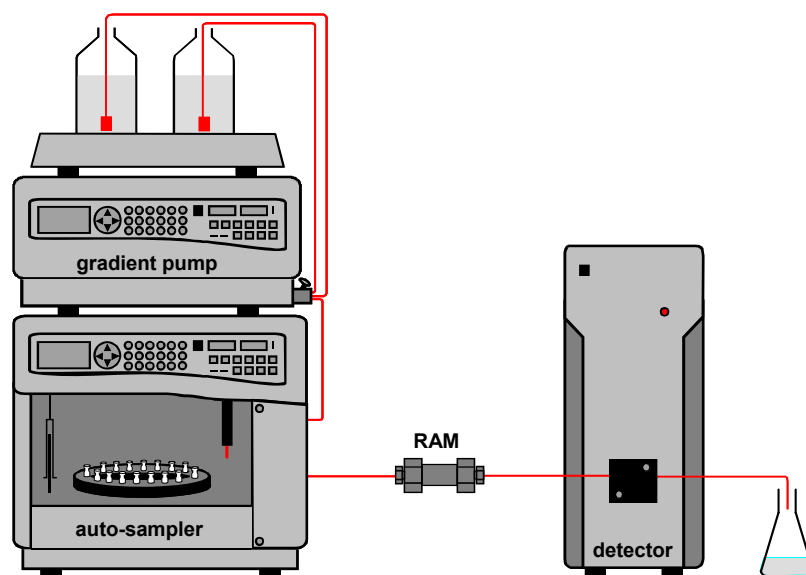


Fig. 44. One-dimensional analytical HPLC setup used to test RAM columns.

2.4.2 Two-dimensional HPLC setup

The 2D-HPLC-MS setup was assembled as follows. The loading pump was a P680 HPLC pump from Dionex (Idstein, Germany) and the injection system consisted of a Rheodyne 6-port-valve (7725) mounted with a 1 mL external loop (Rohnert Park, CA, USA). Hemoglobin monitoring was performed with a diode array detector UVD340U (Dionex) implemented after the RAM column. The switching unit consisted of a

Rheodyne 6-port-valve (7000). The elution was performed over a Prontosil 300-5-C18-H 5 μm (125 x 2.0 mm) analytical column from Bischoff Chromatography (Leonberg, Germany). Eluent delivery was assured with a Rheos 2000 pump from Flux Instruments (Basel, Switzerland). MS detection of the analytes was performed with a Thermo Finnigan Surveyor MSQ supplied by Dionex (Germering, Germany). A scheme of the analytical setup is depicted in Fig. 45.

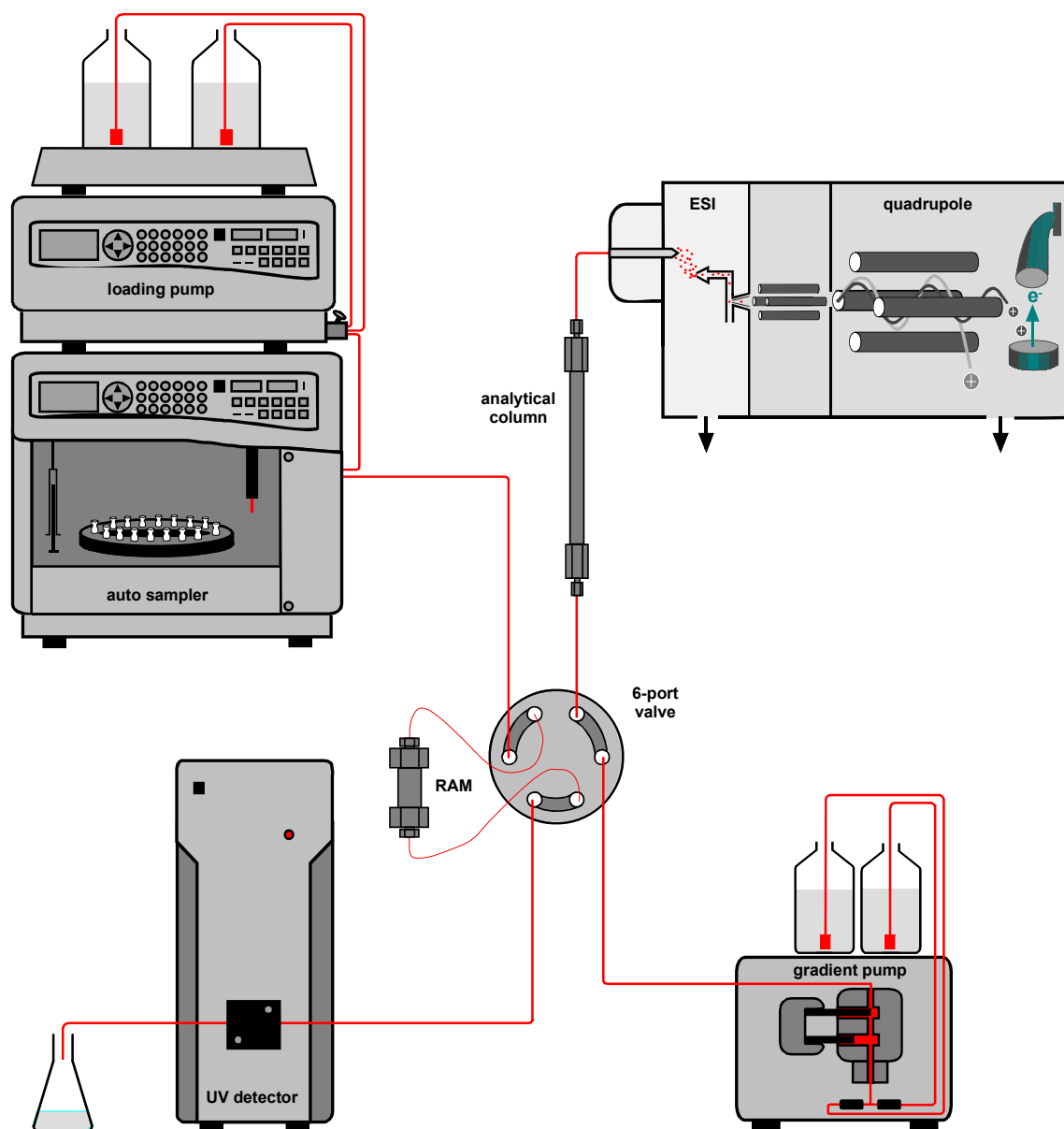


Fig. 45. Two-dimensional experimental setup used to extract analytes from hemolysates. Hemoglobin is monitored after RAM with UV-Vis detection. Analytes are detected after the analytical column with mass spectrometry.

3 Choice of the stationary phase

The so-called Restricted Access Materials are stationary phases of choice to analyze small analytes in complex biological matrixes with on-line SPE-HPLC. Macromolecules are excluded and can only interact with the outer surface of the material, whereas small analytes of interest access the inner surface of the pores and are retained. Different exclusion mechanisms have been developed and numerous media are now available on the market. Classifications of RAM have already been proposed ^[63;65]. They are generally based on the type of mechanism involved to exclude the macromolecules. A schematic classification is depicted in Fig. 46.

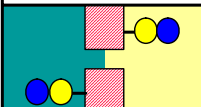
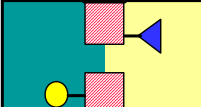
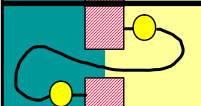

inside outside	exclusion barrier	surface topochemistry	commercial products
	physical	uniform	ChromSpher 5 Biomatrix (Chrompack)
	physical	dual	ISRP GFFII (Regis Technologies) LiChrospher ADS (Merck KGaA)
	chemical	uniform	Hisep (Supelco) Capcell Pak MF (Shiseido)
	chemical	dual	Bioptic AV-2 (GL Sciences) BioTrap 500 MS (ChromTech) SPS (Regis Technologies)

Fig. 46. Classification of restricted access materials (RAM) ^[65].

The exclusion is achieved either by a physical or by a chemical exclusion barrier. Physical exclusion barriers consist of small pore diameters (60 – 120 Å), whereas chemical exclusions are achieved by anchoring a polymer-network at the surface of the material. In both cases, macromolecules do not have the possibility to penetrate into the pores. Uniform or dual surface topochemistries are employed. In order to check if some stationary phases are better designed to exclude hemoglobin from whole blood hemolysate, six commercially available columns were chosen and tested under the same conditions. The name and characteristics of the columns are summarized in Tab. 7. Each type of exclusion mechanism (chemical vs. physical exclusion barrier and uniform vs. dual surface topochemistry) was represented.

Tab. 7. Characteristics of the different tested RAM columns.

	ChromSpher Biomatrix Varian	LiChrospher ADS Merck	Capcell Pak Shiseido	Bioptic AV-2 GL Sciences	Biotrap 500 MS Chromtech	SPS C18 Regis
stationary phase	porous silica	porous silica	porous silica	porous silica	hydrophobic polymer	porous silica
particle size [μm]	5	25	5	5	n.a.	5
pore size [Å]	120	60	80	n.a.	n.a.	100
length [mm]	50	25	50	50	13	50
diameter [mm]	4.6	4.0	4.6	4.6	4.0	2.1
pH stability	2.0-8.0	2.0-7.5	2.0-7.5	2.0-7.5	2.0-11.0	2.5-7.5

Hemoglobin samples were prepared by dissolving human hemoglobin to a concentration of 150 mg/mL in 0.05 % aqueous trifluoroacetic acid (TFA). 10 μL of the hemoglobin solution were injected and eluted for 15 min isocratically with 0.05 % TFA, followed by a gradient of 0 - 100 % acetonitrile/0.05 % TFA in 30 min. The column effluent was monitored at 396 nm (the maximum of absorption of hemoglobin). Finally the columns were washed and regenerated by 100 % acetonitrile/0.05 % TFA for 2 min and 0.05 % aqueous TFA for 13 min. Blank injections consisting of 10 μL 0.05 % aqueous TFA were performed after each hemoglobin injection in order to check the carry-over of hemoglobin. For each HPLC run, the peaks corresponding to hemoglobin flowing-through and hemoglobin retained on the column were integrated. Integration values are only approximated values because of the saturation of the detector over 1,000 mAU. Ratios between hemoglobin retained and hemoglobin flowing through were also computed. Carry-overs were also determined for each column by performing two consecutive blank injections.

3.1 Hemoglobin exclusion with physical barrier and uniform surface topochemistry: ChromSpher Biomatrix

The first column to be evaluated was ChromSpher Biomatrix from Varian. The physical exclusion of macromolecules is achieved by pore sizes of 120 Å. In-pore surface and external surface chemistries are similar. They consist of a combination of diol- and hydrophobic groups. Because of the arrangements of the diol- and the hydrophobic groups, only small analytes can access to hydrophobic surfaces, whereas macromolecules can only interact with diol groups. A schematic representation of ChromSpher Biomatrix material is depicted in Fig. 47.

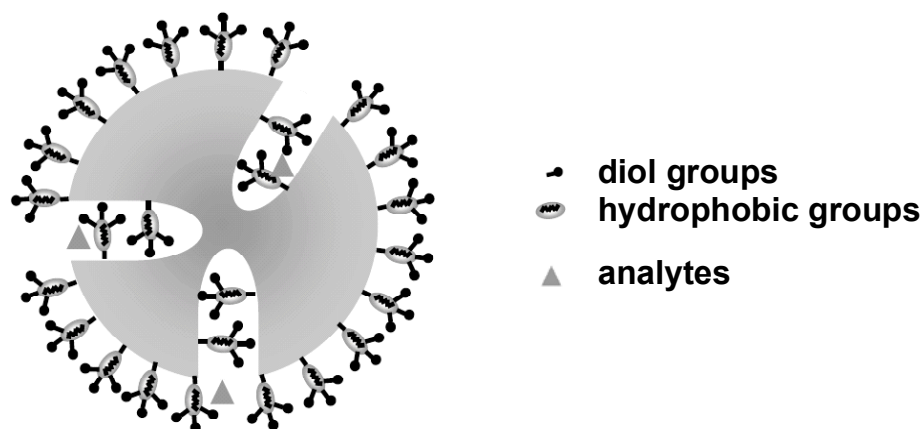


Fig. 47. Schematic representation of the topochemistry of ChromSpher Biomatrix ^[65].

Chromatograms corresponding to an injection of hemoglobin and to the first consecutive blank injection are depicted in Fig. 48 and in Fig. 49. One observes that hemoglobin is not only flowing through the column, but is also retained on the stationary phase. This signifies that hemoglobin is not completely excluded. The elution of hemoglobin occurred in a broad peak at about 20 % acetonitrile in the eluent. Carry-over is observed in a consecutive blank injection at a similar acetonitrile proportion.

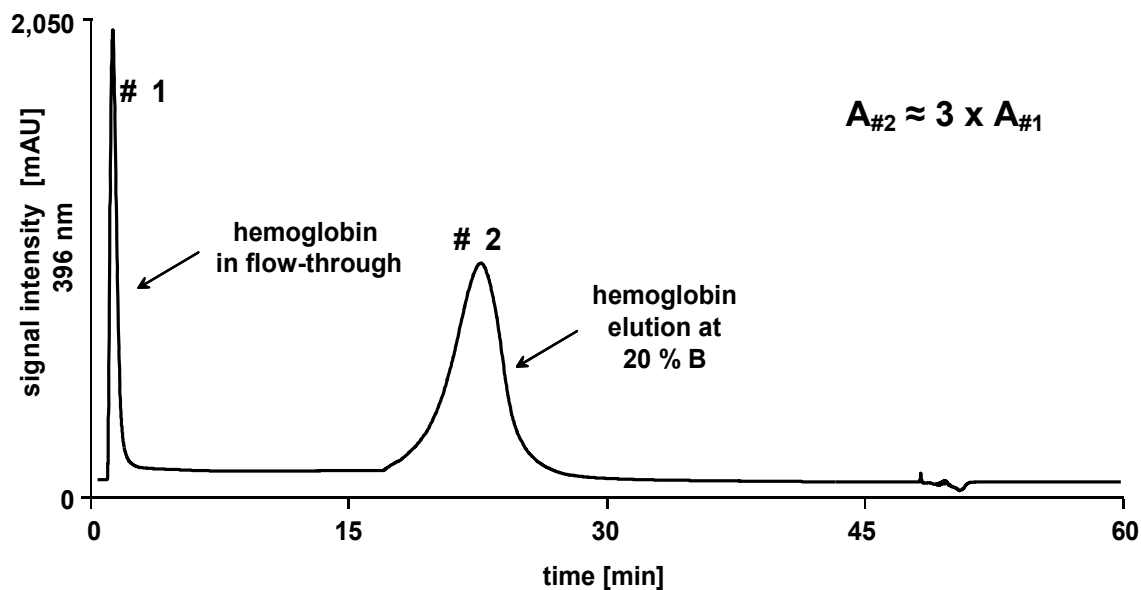


Fig. 48. Retention of hemoglobin on ChromSpher Biomatrix Varian. Sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

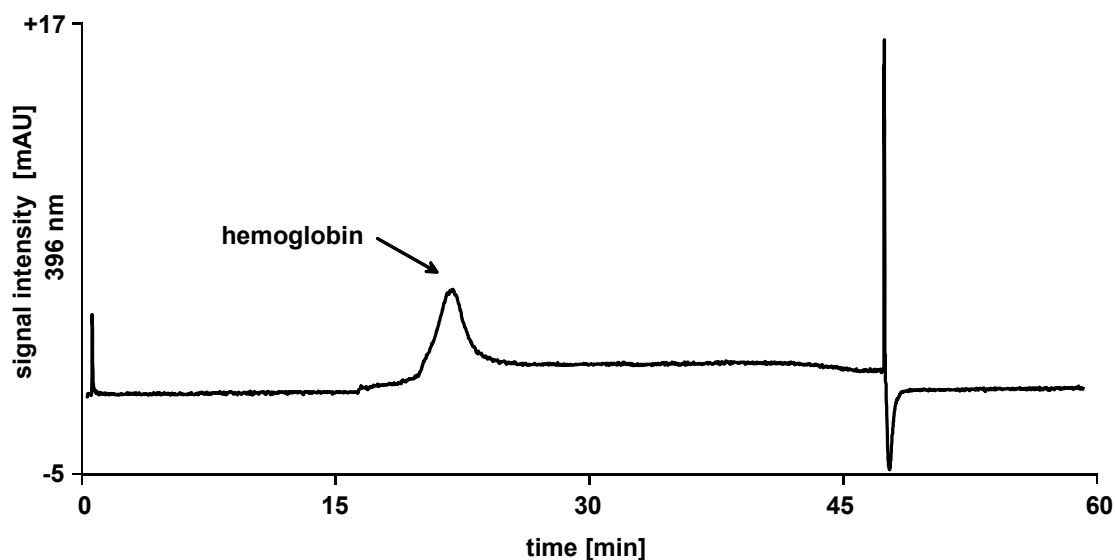


Fig. 49. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, ChromSpher Biomatrix Varian; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

3.2 Hemoglobin exclusion with physical barrier and dual surface topochemistry: LiChrospher ADS

The physical exclusion of macromolecules is achieved by pore sizes of 60 Å in the LiChrospher ADS material. The acronym ADS signifies alkyl-diol-silica and summarizes the two different groups present at the surface of the material. The inner surface of the pores is coated with C18 chains, whereas diol groups are present on the external surface of the particles (see Fig. 14 and Fig. 50). Contrary to small analytes, macromolecules can not penetrate into the pores and can not be retained by interactions with C18 chains.

Chromatograms corresponding to an injection of hemoglobin and to the first consecutive blank injection are depicted in Fig. 51 and in Fig. 52. As also observed with the ChromSpher Biomatrix column, hemoglobin is not only flowing through the column but is also retained on the stationary phase. Hemoglobin elutes in a broad peak at about 40 % acetonitrile in the eluent. Carry-over of hemoglobin is observed in a consecutive blank injection at a similar acetonitrile proportion.

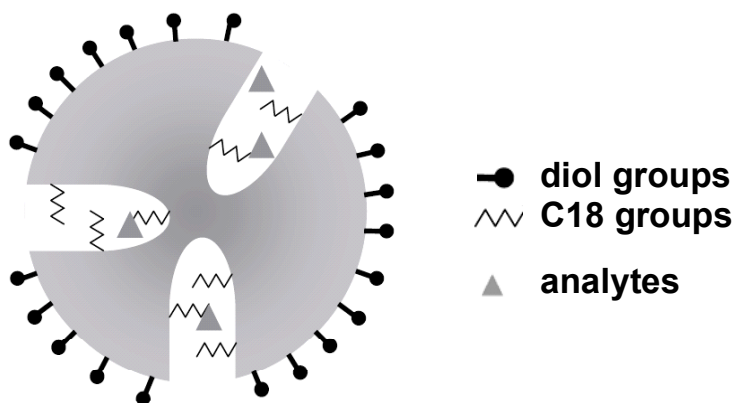


Fig. 50. Schematic representation of the topochemistry of LiChrospher ADS ^[65].

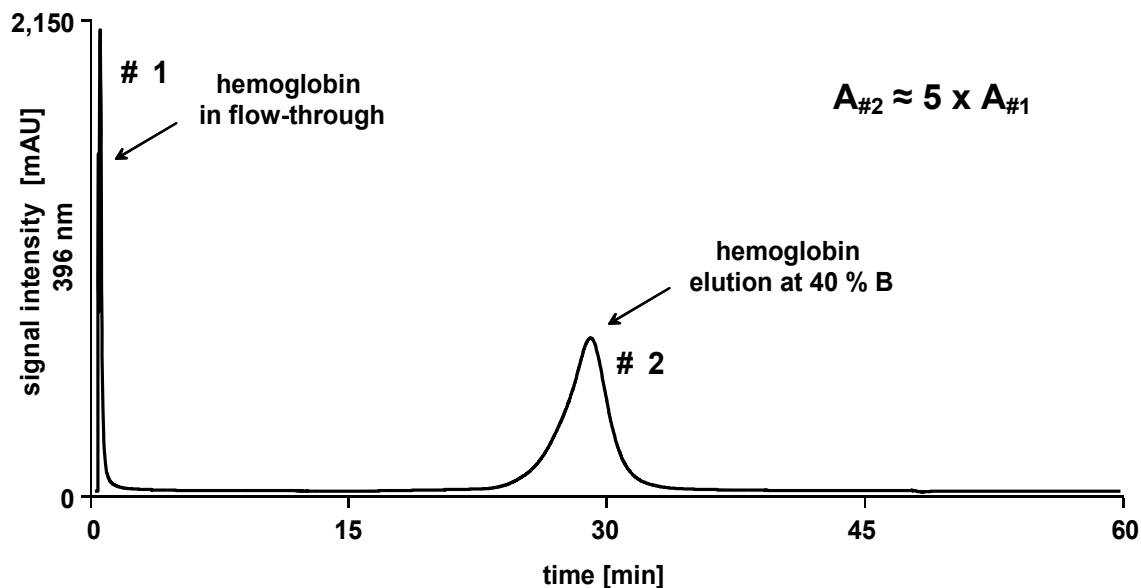


Fig. 51. Retention of hemoglobin on LiChrospher ADS Merck; sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

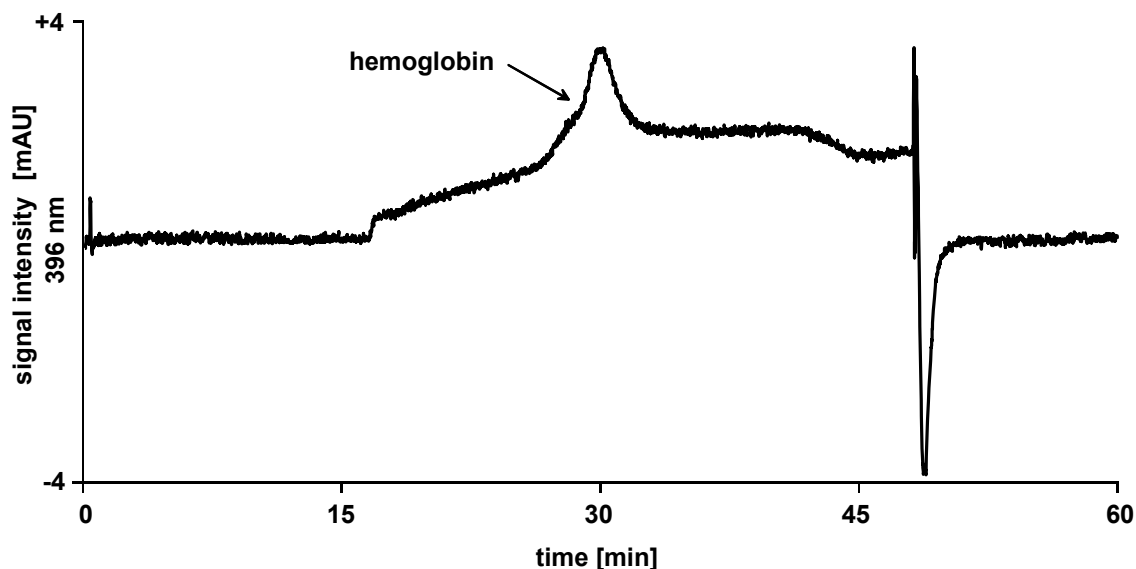


Fig. 52. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, LiChrospher ADS Merck; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

3.3 Hemoglobin exclusion with chemical barrier and uniform surface topochemistry: Capcell Pak

In Capcell Pak materials, the silica support is covered with a uniform thin film of silicone polymer (capsule-type packing material). Chemical exclusion of macromolecules is achieved by polyoxyethylene chains grafted at the surface of the silicone layer. These long chains shield hydrophobic groups, which are only accessible to small analytes. Macromolecules can not really interact with the hydrophobic groups and flow through the column. A schematic representation of Capcell Pak is depicted in Fig. 53.

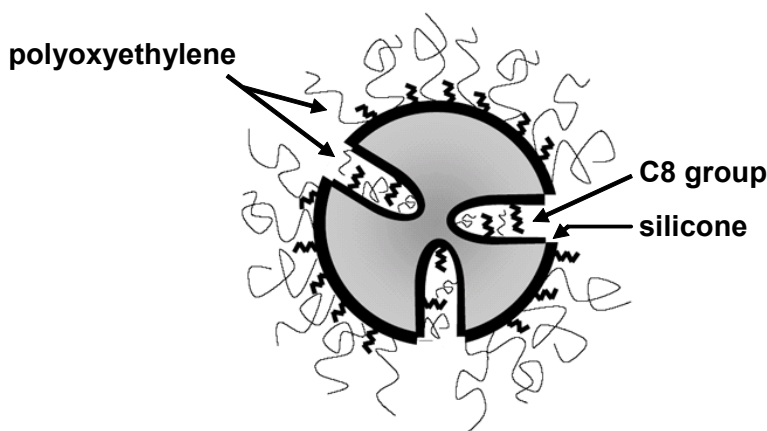


Fig. 53. Schematic representation of the topochemistry of Capcell Pak ^[65].

Chromatograms corresponding to an injection of hemoglobin and to the first consecutive blank injection are depicted in Fig. 54 and in Fig. 55. As also observed with the columns with physical exclusion barrier, hemoglobin is flowing through the column but is also retained on the stationary phase. However, the amount of hemoglobin retained on the column appears to be much higher in the case of the Capcell Pak column ($A_{\#2} \approx 30 \times A_{\#1}$) as compared to the previously tested columns ($A_{\#2} \approx 3-5 \times A_{\#1}$). Hemoglobin is eluted from the stationary phase in a broad peak at about 40 % acetonitrile in the eluent. Hemoglobin carry-over is also observed in a consecutive blank injection.

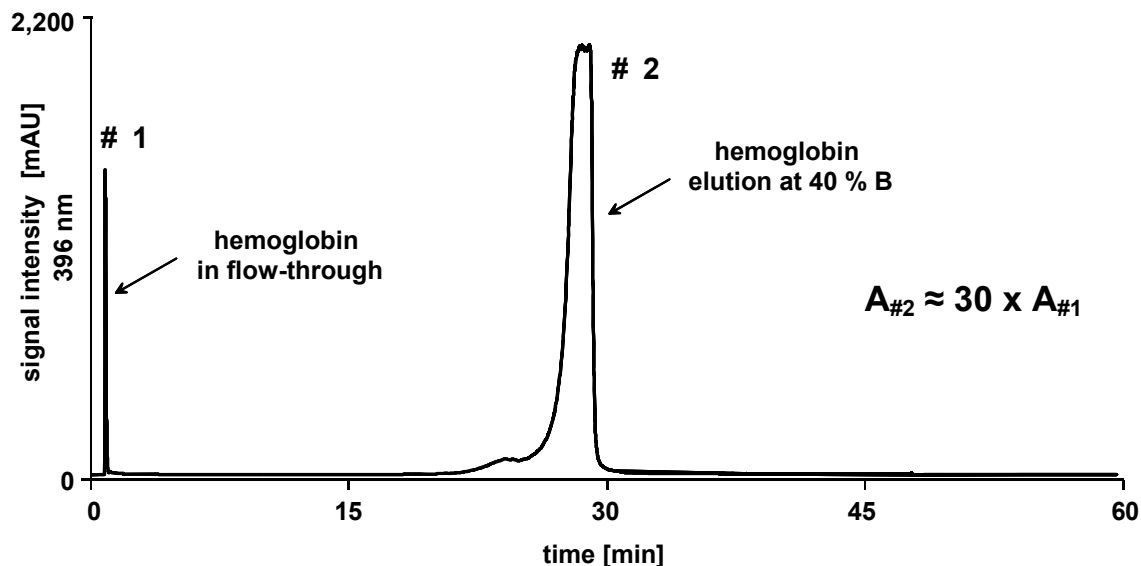


Fig. 54. Retention of hemoglobin on Capcell Pak Shiseido; sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

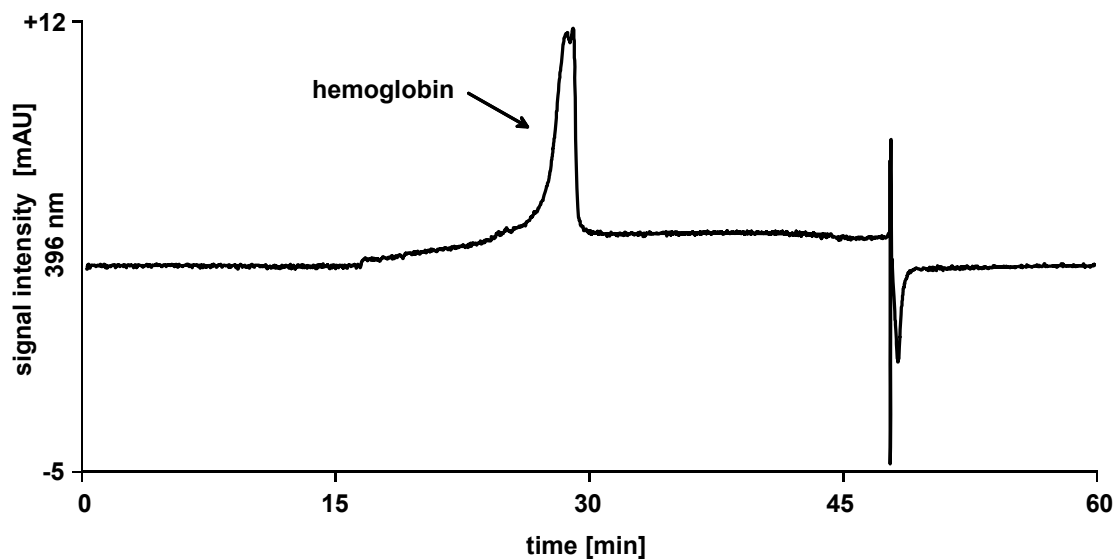


Fig. 55. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, Capcell Pak Shiseido; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

3.4 Hemoglobin exclusion with chemical barrier and dual surface topochemistry: Bioptic AV-2, SPS, and Biotrap 500 MS

The three last materials evaluated in terms of human hemoglobin exclusion and carry-over are excluding macromolecules by combining a chemical barrier and different chemistries inside and outside the pores. For Bioptic AV-2, denatured avidin molecules are grafted at the outer surface of the stationary phase. For SPS, polyoxoethylene chains are utilized (as for the Capcell Pak column). In case of Biotrap 500 MS, α_1 -acid glycoprotein molecules are bound at the outer surface of the particles. For all three materials, hydrophobic groups such as C18 chains (SPS) are present at the inner surface of the pores. A generic representation of the three columns is depicted in Fig. 56.

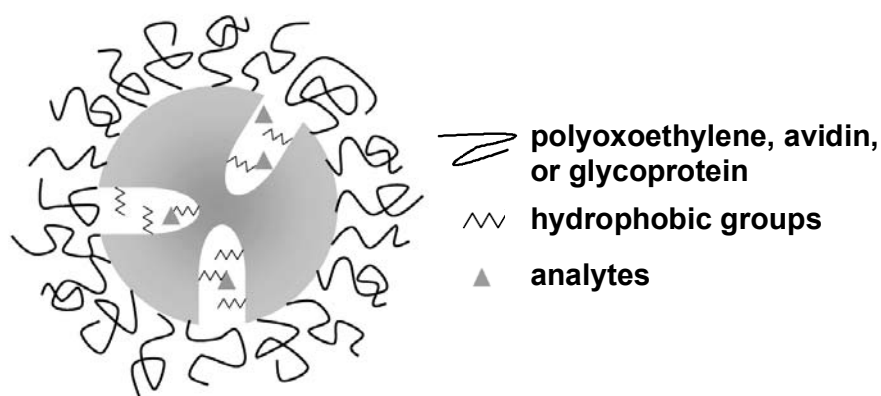


Fig. 56. Schematic representation of Bioptic AV-2, SPS, and Biotrap 500 MS ^[65].

In all three cases, hemoglobin was observed in the flow-through but hemoglobin was also retained on the column. Elution of hemoglobin occurred at 20 %, 45 % and 30 % acetonitrile in the eluents for Bioptic AV-2, SPS, and Biotrap 500 MS materials, respectively. With all three columns, carry-over of hemoglobin was observed in consecutive blank injections. Chromatograms of hemoglobin and blank injection are depicted in Fig. 57 and in Fig. 58 for Bioptic AV-2, in Fig. 59 and in Fig. 60 for SPS, and in Fig. 61 and in Fig. 62 for Biotrap 500 MS, respectively.

No significant difference was observed between the three materials with chemical exclusion and dual topochemistry. However, stronger retention of hemoglobin was observed with the SPS C18 material. The elution peak was also very broad (more than 10 min). The carry-over appeared more pronounced with the SPS C18 column.

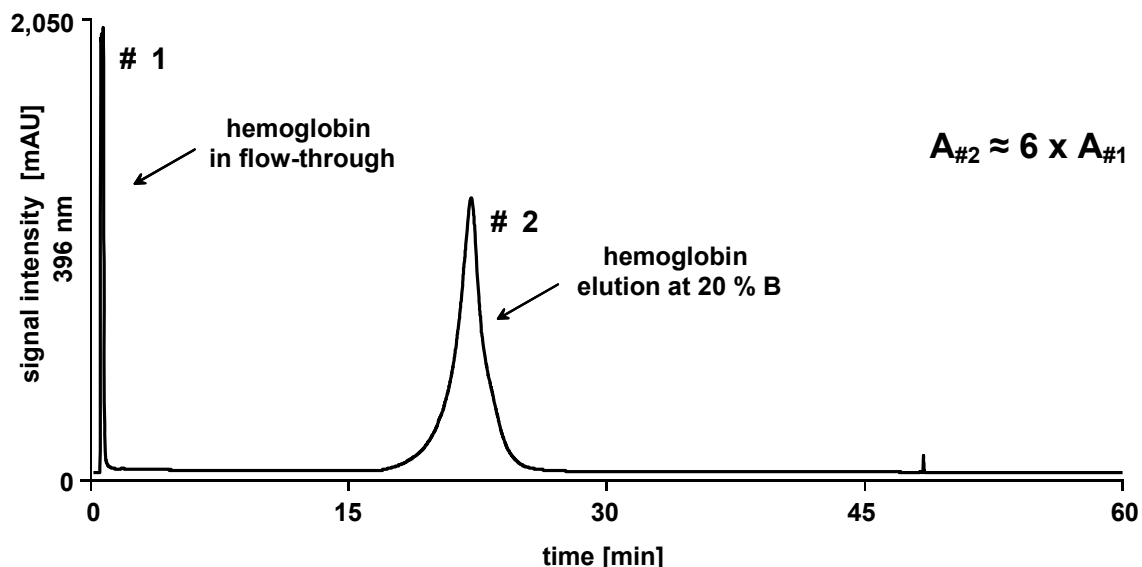


Fig. 57. Retention of hemoglobin on Bioptronic AV-2 GL Sciences Inc.; sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

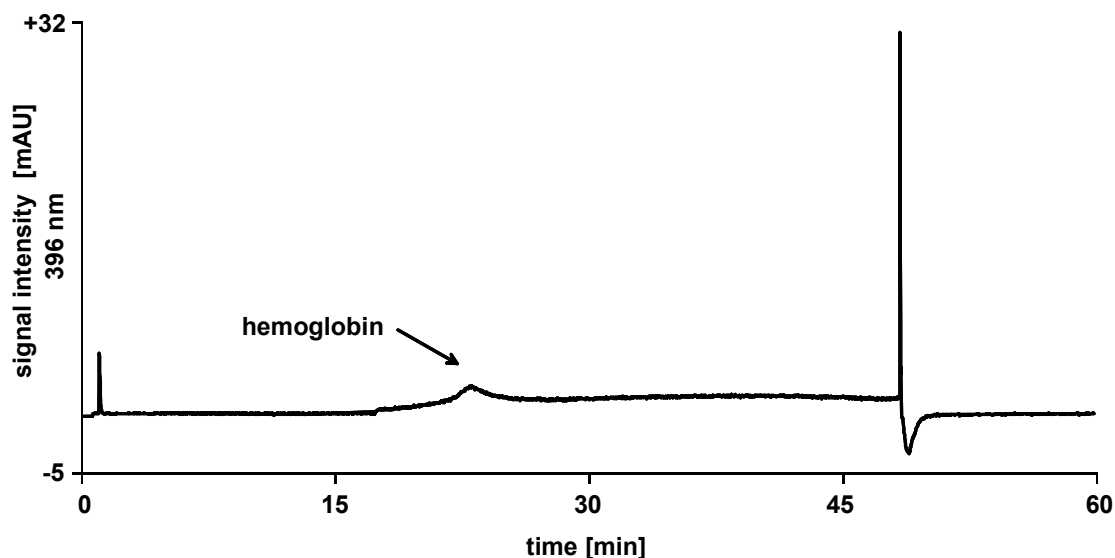


Fig. 58. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, Bioptronic AV-2 GL Sciences Inc.; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.0 mL/min; room temperature; detection at 396 nm.

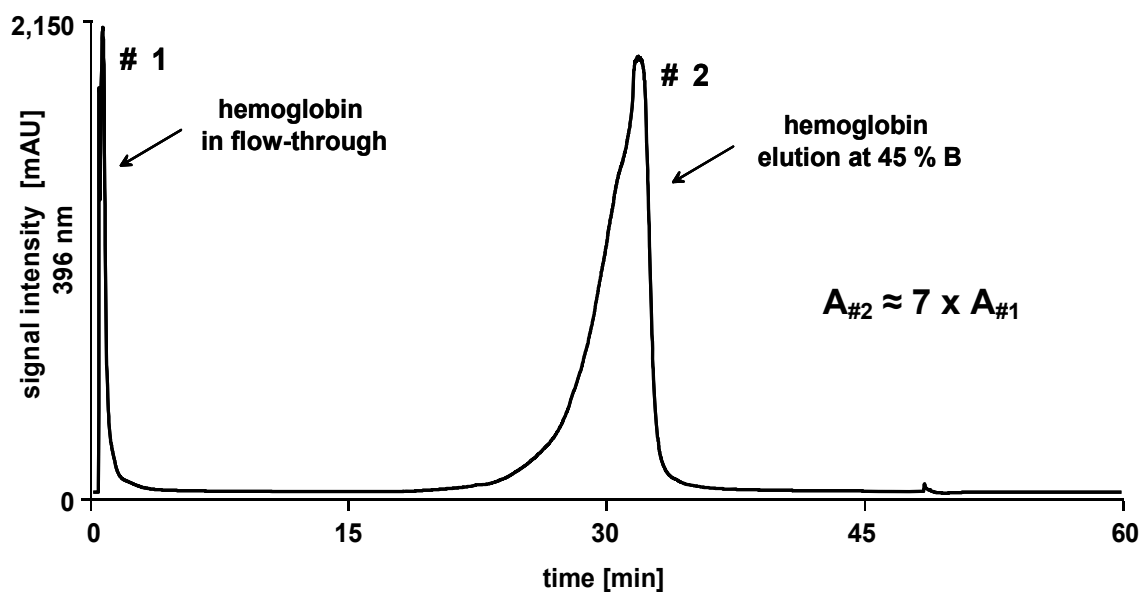


Fig. 59. Retention of hemoglobin on SPS C18 Regis; sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 0.5 mL/min; room temperature; detection at 396 nm.

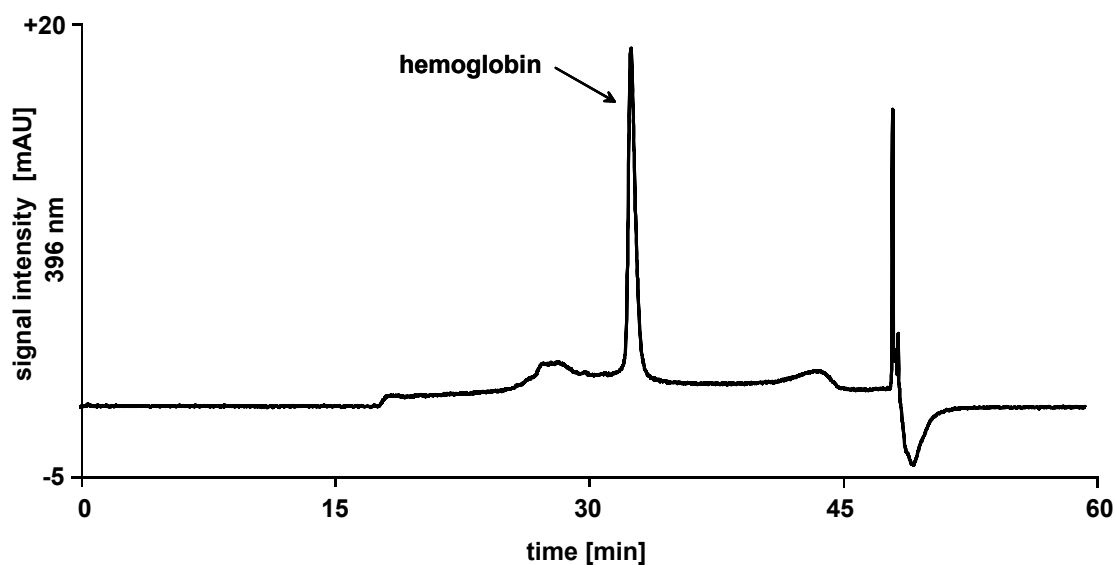


Fig. 60. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, SPS C18 Regis.; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 0.5 mL/min; room temperature; detection at 396 nm.

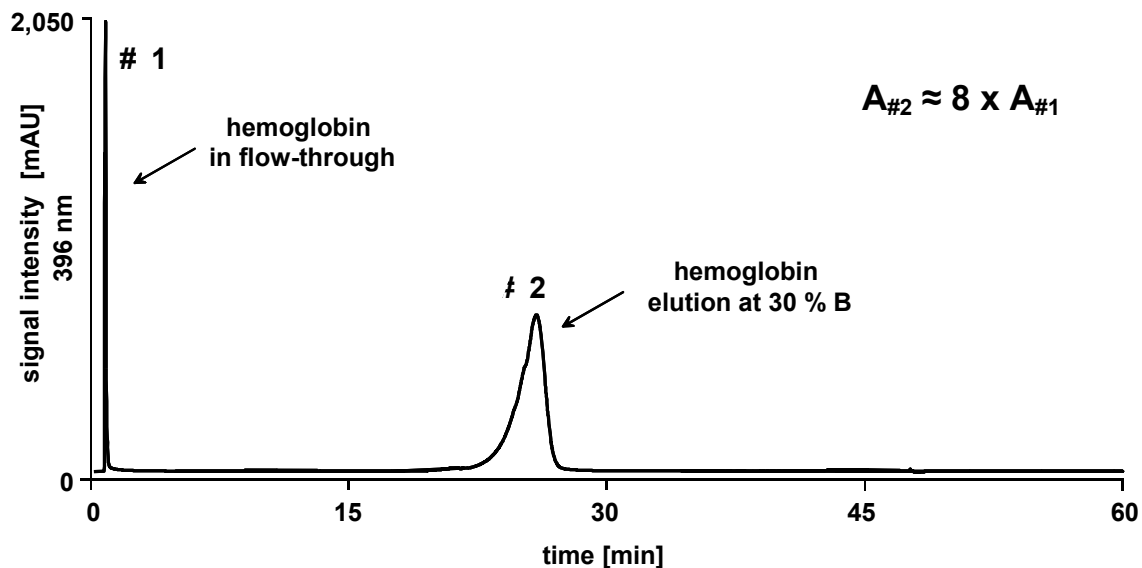


Fig. 61. Retention of hemoglobin on Biotrap 500 MS Chromtech; sample, 150 mg/mL hemoglobin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

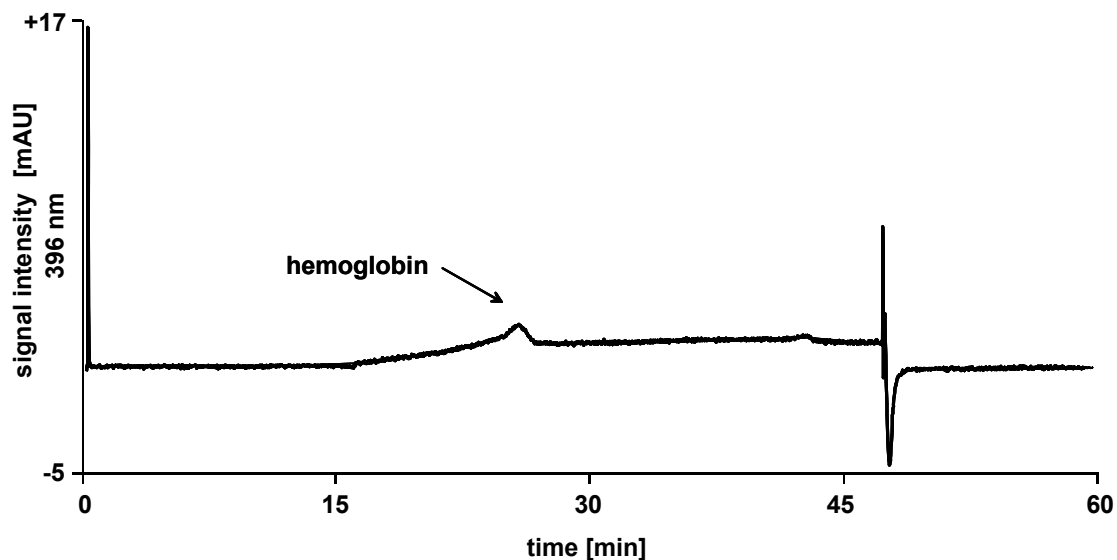


Fig. 62. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, Biotrap 500 MS.; sample, 10 μ L 0.05 % aqueous TFA; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

3.5 Conclusions and next steps

Whatever the principle of macromolecule exclusion, the elution profiles for all six columns looked quite similar. Part of the hemoglobin eluted in the flow-through, while the rest reproducibly eluted as a relatively broad peak at 20-40 % acetonitrile concentrations. In all cases, most of the hemoglobin was retained on the column (peak # 2) as compared to elution in the flow-through (peak # 1). The following injections of pure water/0.05 % TFA showed relatively small peaks for hemoglobin in the flow-through, but significant peaks of retained hemoglobin. Carry-overs are in the same order of magnitude for the six columns. However, the exclusion of hemoglobin appears to be most efficient with ChromSpher Biomatrix and with LiChrospher ADS material (area peak # 2 / area peak # 1), and very ineffective for Capcell Pak. The comparison of the six columns in terms of hemoglobin carry-over and hemoglobin retention is summarized in Tab. 8.

Tab. 8. Hemoglobin carry-over [%] at pH 2.1 in peak # 2 for six different RAM columns. Injections # 0 correspond to injections of 10 μ L of 150 mg/mL hemoglobin and are considered as a reference for each column. Injections # 1 and # 2 are blank injections directly following injections # 0.

	peak area of retained hemoglobin [%]					
	ChromSpher Biomatrix Varian	LiChrospher ADS Merck	Capcell Pak Shiseido	Bioptic AV-2 GL Sciences	SPS C18 Regis	Biotrap 500 MS Chromtech
injection # 0	100	100	100	100	100	100
injection # 1	0.52	0.20	0.35	0.10	0.20	0.11
injection # 2	0.32	0.05	0.03	0.03	0.12	0.05
area peak # 2 /area peak # 1 injection # 0	3	5	30	6	7	8

The areas of the peaks corresponding to retained hemoglobin (peak # 2) were also plotted in a histogram. As observed in Fig. 63, the higher retention was observed with SPS C18. It is not possible to directly compare the results of Capcell Pak and

ChromSpher Biomatrix with the results of the other restricted access materials because freshly prepared hemoglobin solutions were employed.

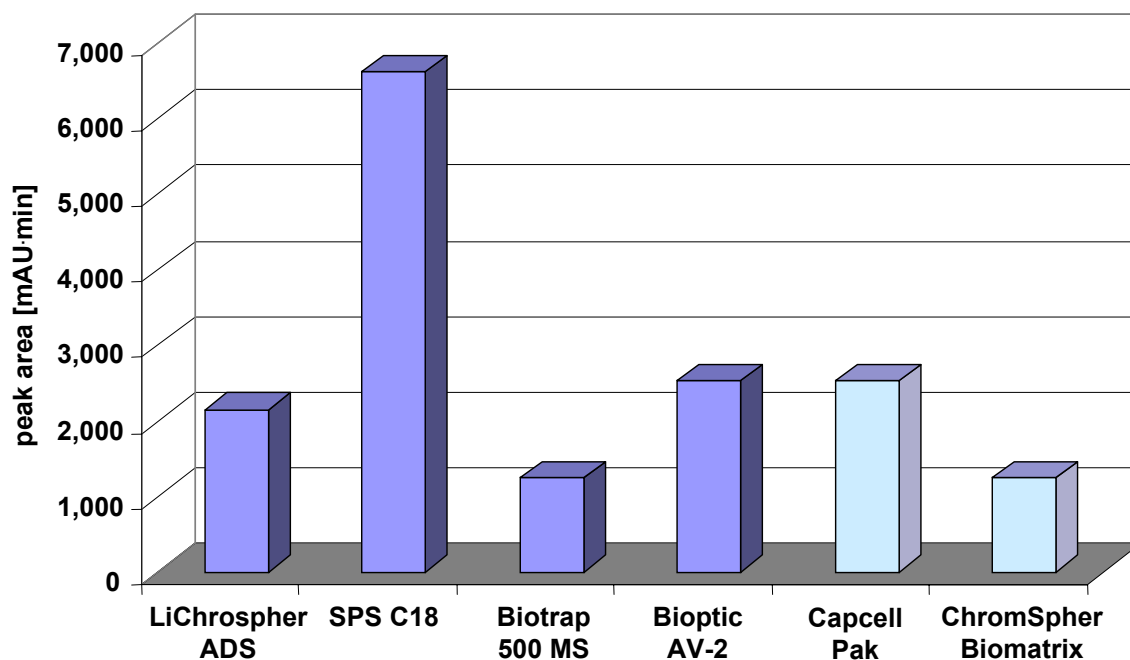


Fig. 63. Hemoglobin retention on six different restricted access media. A different 150 mg/mL hemoglobin solution was used to test Capcell Pak and ChromSpher Biomatrix.

To explain the retention of hemoglobin on the stationary phase, two phenomena are taken into consideration. The first reason is that the pores are too large, permitting hemoglobin to enter and adsorb on the C18 groups or even interact with silanol groups. The second reason is that some unspecific interaction such as silanol/ion-exchange interaction may take place on the outer surface of the material.

An approach to diminish the amount of macromolecules binding on the inner surface of the pores is to sterically hinder the entrance into the pores. The most adapted material to our challenge seemed to be LiChrospher ADS from Merck. The reason is that this restricted access medium has the smallest pores of all materials available on the market with pore dimensions of $60 \pm 10 \text{ \AA}$ ($6 \pm 1 \text{ nm}$). Moreover, diol groups on the outer surface are suitable for further functionalization. A modification of the LiChrospher ADS material and the influence on the retention of hemoglobin is investigated in the next section.

4 Modification of LiChrospher ADS with aminodextran

Retention of hemoglobin may take place in the pores of the LiChrospher ADS material, where C18 chains are grafted to the silica based stationary phase. It signifies that the entrance of the pores (6 – 8 nm) may be too huge and permit at least partial penetration of hemoglobin molecules (M_w 16,000 – 18,000 Da). The grafting of large and voluminous compounds (e. g. aminodextran) on the outer surface of the stationary phase should reduce the entrance space of the pores and diminish the amount of macromolecules accessing the hydrophobic chains inside the pores. A reaction scheme depicting the derivatization of LiChrospher ADS particles with aminodextran is represented in Fig. 64.

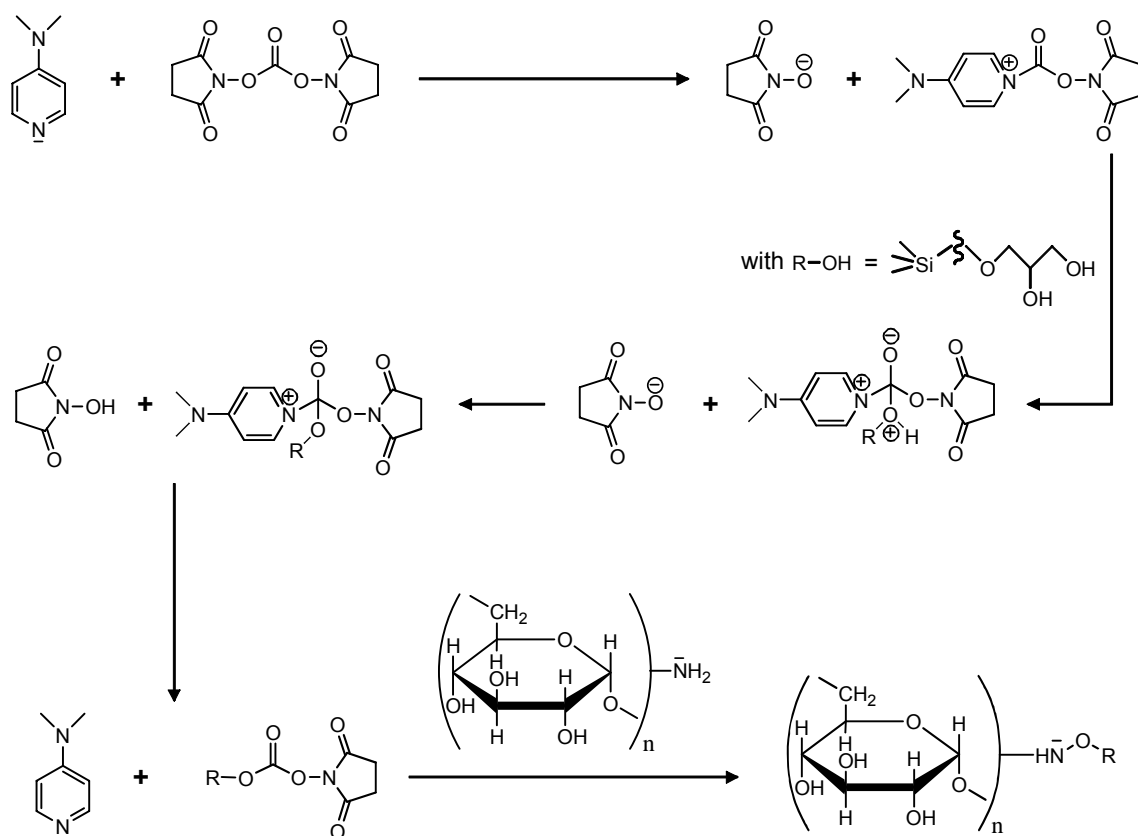


Fig. 64. Derivatization of LiChrospher ADS with aminodextran.

An in-house 30 x 2.0 mm column was packed with unmodified LiChrospher ADS material and mounted in the one-dimensional HPLC system (see 2.4.1). 10 μL of five

0.5 mg/mL protein solutions were successively injected. The proteins (insulin, ubiquitin, β -lactoglobulin A+B, ovalbumin, hemoglobin) were covering a large range of masses: from 5,800 to 45,000 Da. The proteins were eluted for 5 min isocratically with 0.05 % aqueous TFA, followed by a gradient of 0-100 % acetonitrile/0.05 % TFA in 10 min. An in-house 30 x 2.0 mm column packed with modified LiChrospher ADS material with aminodextran was also subjected to the same chromatographic runs. For each chromatographic run, peak areas corresponding to proteins retained on the column were computed. Values are only approximated because of difficulties of integration due to very large fronting and tailing. The different results are summarized in Tab. 9.

Tab. 9. Retention of proteins on a 30 x 2.0 mm in-house packed LiChrospher ADS column and on an in-house modified LiChrospher ADS material packed in a 30 x 2.0 mm column ^a.

protein	M _w (Da)	protein amount [mAU·min] retained on LiChrospher ADS	protein amount [mAU·min] retained on modified LiChrospher ADS
insulin	~ 5,800	10.2	1.2
ubiquitin	~ 8,600	4.7	not detected
β -lactoglobulin A+B	~ 18,400	1.7	not detected
ovalbumin	~ 43,000	not detected	not detected
hemoglobin	~ 64,500 subunits: 16,000	14.1	16.3

^a Samples, 0.5 mg/mL protein diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 5.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 10.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 214 nm, 396 nm for hemoglobin.

For small proteins, which can penetrate at least part into a fraction of the pores to be retained on the C18 surface, RAM material modification shows a significant reduction of protein retention. Insulin adsorption is thus 10 fold smaller on the modified as compared to the unmodified material. Ubiquitin and lactoglobulins are no longer retained.

Ovalbumin, not retained on the unmodified material, is also not detected on the modified phase. However, hemoglobin adsorption is equivalent on both materials. Thus, it appears that the modification of Lichrospher ADS material with aminodextran is effective to suppress pore penetration of small proteins, but the modification does not seem to influence the retention of large proteins. It suggests, that adsorption of small proteins (e.g. insulin) occurs in the pores and is reduced by steric obstruction, whereas retention of large proteins is due to nonspecific adsorption at the outer surface of the restricted access material.

5 Modification of LiChrospher ADS with poly-D-lysine and polyethyleneimine

Nonspecific interactions between macromolecules and the outer surface of the restricted access material seem to explain the retention of large proteins. In order to check this hypothesis, LiChrospher ADS was derivatized with chemical structures (amines) presenting positively charged groups under acidic conditions (e.g. pH 2.1). Because hemoglobin also carries positive charges at low pH (pI 6.8), electrostatic interactions should permit to reject hemoglobin from the stationary phase and avoid non-specific retention.

LiChrospher ADS bulk material was derivatized with polyethyleneimine 25 (M_w 25,000 Da), polyethyleneimine 60 (M_w 60,000 Da), and poly-D-lysine. Then, columns were packed and successively mounted in the one-dimensional HPLC setup (see 2.4.1). 10 μ L of the 0.5 mg/mL protein solutions previously mentioned were successively injected over each column. The proteins were eluted for 5 min isocratically with 0.05 % TFA, followed by a gradient of 0-100 % acetonitrile/0.05 % TFA in 5 min. Unmodified LiChrospher ADS material was also tested to get reference values.

Tab. 10. Retention of proteins on a 30 x 2.0 mm in-house packed LiChrospher ADS column and on in-house modified LiChrospher ADS materials packed in 30 x 2.0 mm columns ^a.

			peak area of retained protein peak [mAU' min]				
	MW (Da)	pI	unmodified	amino dextran	poly-ethylene-imine 25	poly-ethylene-imine 60	polylysine 30-70
insulin	~ 5,800	5.5	2.3	2.3	38.1	14.2	12.6
ubiquitin	~ 8,600	6.6	0.1	0.1	6.8	1.6	4.4
hemoglobin	~ 16,000	6.8	2.5	1.4	2.3	1.5	4.8
ovalbumin	~ 45,000	5.2	0.6	0.8	2.5	1.0	17.5

^a Samples, 0.5 mg/mL protein diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 5.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 5.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 214 nm, 396 nm for hemoglobin.

For each chromatographic run, peak areas corresponding to proteins retained on the column were computed. The different results are summarized in Tab. 10. No significant differences were observed between the unmodified LiChrospher ADS material and the material derivatized with aminodextran. For insulin and ubiquitin these results are in contradiction with the values presented in the previous section (see section 4). However, in the present case, a very steep gradient was utilized for elution (0 – 100 % acetonitrile in 5 min) and the integration of the peaks was much easier because of much more discernable peaks. To sum up, modification of LiChrospher ADS material with aminodextran has no real impact on the amount of retained proteins.

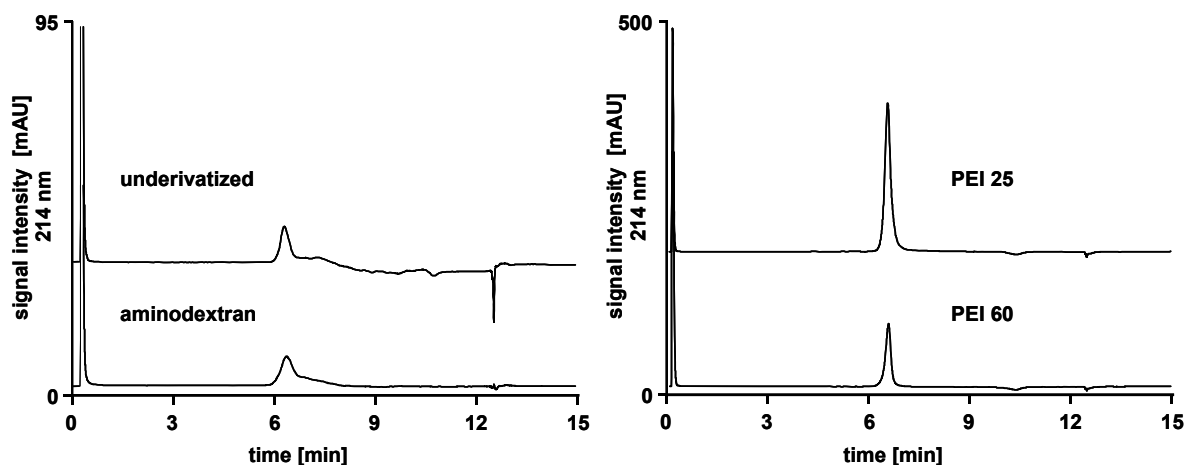


Fig. 65. Influence of the derivatization of LiChrospher ADS material on the retention of insulin. Columns, 30 x 2.0 mm in-house packed columns with underivatized LiChrospher material, material derivatized with aminodextran, polyethyleneimine 25, and polyethyleneimine 60. Sample, 0.5 mg/mL insulin diluted in 0.05 % aqueous TFA; 10 μ L injection; mobile phase, (A) 0.05 % TFA in H₂O, (B) 0.05 % TFA in ACN; 5.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 5.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 214 nm.

Derivatization of LiChrospher ADS with polyethyleneimine 25 and 60 results in increased adsorptions of small proteins (e.g. insulin and ubiquitin). No significant influence is observed for larger proteins. No correlation was observed between pI of proteins and increases in adsorption. Higher amounts of adsorbed proteins were observed with modifications performed with polyethyleneimine 25 than with polyethyleneimine 60. Derivatization of LiChrospher ADS with poly-D-lysine results in increased adsorption for all four proteins. However, no correlation was observed between size of proteins and increase in adsorption. Chromatograms obtained for

insulin with unmodified LiChrospher ADS material, and material derivatized with aminodextran, and polyethyleneimines are depicted in Fig. 65. The chemical composition of modified and unmodified LiChrospher ADS materials was checked with elemental analyses (Model Vario EL, Elementar Analysensysteme GmbH, Hanau, Germany). The results are summarized in Tab. 11. Significant differences are observed for materials derivatized with polyethyleneimines. An increase of the N concentration of about 0.3 % is observed, proving the successful derivatization of the LiChrospher ADS material with polyethyleneimine 25 and polyethyleneimine 60.

Tab. 11. Elemental analyses of LiChrospher ADS materials.

	N [%]	C [%]	H [%]
unmodified material	0.02	20.65	3.16
activated material	0.02	22.46	4.75
material modified with aminodextran	0.04	22.50	3.67
material modified with polyethyleneimine 25	0.32	21.72	3.67
material modified with polyethyleneimine 60	0.29	21.72	3.68
material modified with polylysine	0.06	21.55	3.60
aminodextran	0.23	40.23	7.45
polyethyleneimine 25	14.89	24.10	8.10
polyethyleneimine 60	16.10	27.88	8.94

As a conclusion, modifications of LiChrospher ADS with positively charged polymers do not diminish but increase the adsorption of positively charged proteins (chromatographic runs performed at pH 2.1). Thus, electrostatic interactions do not seem to play a major role in the exclusion/retention of proteins on the LiChrospher ADS material. Hydrophobic interactions between proteins and hydrophobic regions of the polymers grafted on the stationary phase may explain these results. Derivatization of LiChrospher ADS either with aminodextran (neutral at pH 2.1), or with polyethyleneimines and poly-D-lysine (positively charged at pH 2.1) did not permit to significantly reduce the adsorption of hemoglobin on the stationary phase.

6 Evaluation of LiChrospher ADS at different pH

To better understand the adsorption of hemoglobin on LiChrospher ADS particles, solutions of hemoglobin were injected at various pH. For each pH value, amounts of hemoglobin flowing through (peak # 1) and retained on the column (peak # 2) were computed. Carry-over was also evaluated by performing blank injections consecutive to injections of hemoglobin.

Ten μL of a 150 mg/mL hemoglobin solution were injected over an unmodified LiChrospher ADS column mounted in the one-dimensional HPLC setup. Elution was performed with a linear gradient (0-100 % B in 30 min) and followed by a wash step (2 min at 100 % B). The separation was performed at pH 6.6 ($\text{H}_2\text{O}/\text{ACN} + 5 \text{ mmol/L}$ ammonium acetate), pH 8.3 ($\text{H}_2\text{O}/\text{ACN} + 5 \text{ mmol/L}$ imidazole), and pH 10.7 ($\text{H}_2\text{O}/\text{ACN} + 10 \text{ mmol/L}$ ethanolamine). Detection was performed at 396 nm.

At pH 6.6, hemoglobin is present in the flow-through (100 %), elutes at around 15 % B (24 %), and is also present in the wash peak (4 %). However, this separation is not always reproducible and the quantitation of the retained hemoglobin is rather difficult because of the large width and the flatness of the peaks. Accurate quantitation of hemoglobin in the flow-through is also difficult because of the high-intensity of the signal. Hemoglobin in the wash peak presents a carry-over and was previously not observed at pH 2.1 with TFA (see Fig. 51). A chromatogram of hemoglobin separated at pH 6.6 is depicted in Fig. 66.

At pH 8.3, hemoglobin is present in the flow-through (100 %) and also elutes at around 30 % B (35 %) in a broad peak but is not present in the wash step. A chromatogram is depicted in Fig. 67. Carry-over in the flow-through and at 35 % B is observed in a blank injection following the injection of hemoglobin at pH 8.3 (see Fig. 68) but no longer in a second blank injection.

Finally, LiChrospher ADS was tested at pH 10.7, although pH 10.7 is far above the stability of silica-based stationary phases. Hemoglobin was present in the flow-through (100 %) and also at 20 % B (20 %). Practically no memory effect was observed in a following blank injection. The corresponding chromatograms are depicted in Fig. 69 and in Fig. 70.

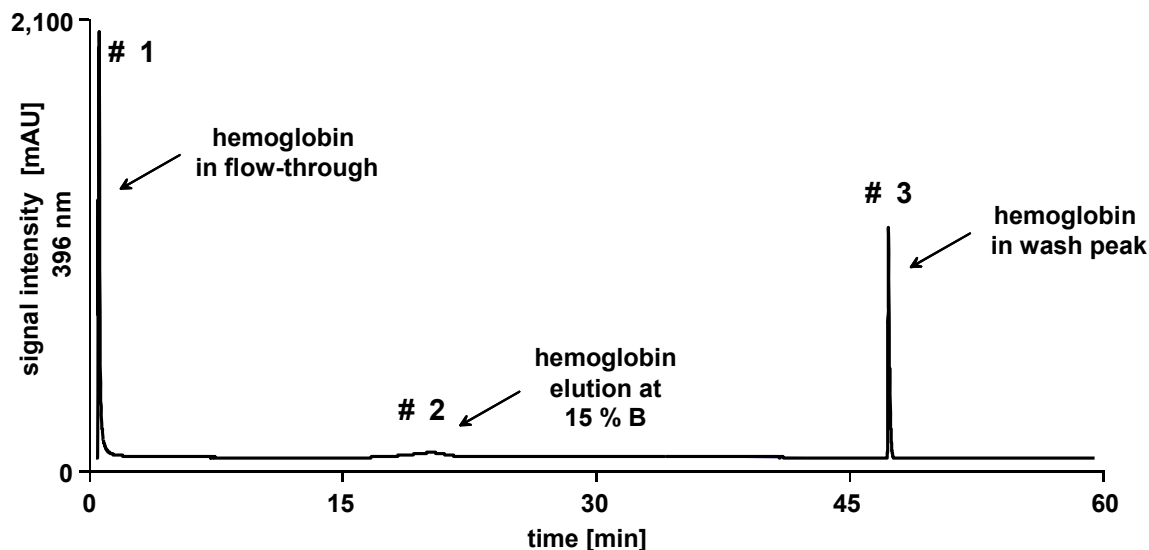


Fig. 66. Chromatogram of hemoglobin at pH 6.6. Column, LiChrospher ADS Merck; sample, 150 mg/mL hemoglobin diluted in 5 mmol/L ammonium acetate; 10 μ L injection; mobile phase, (A) 5 mmol/L ammonium acetate in H₂O, (B) 5 mmol/L ammonium acetate in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

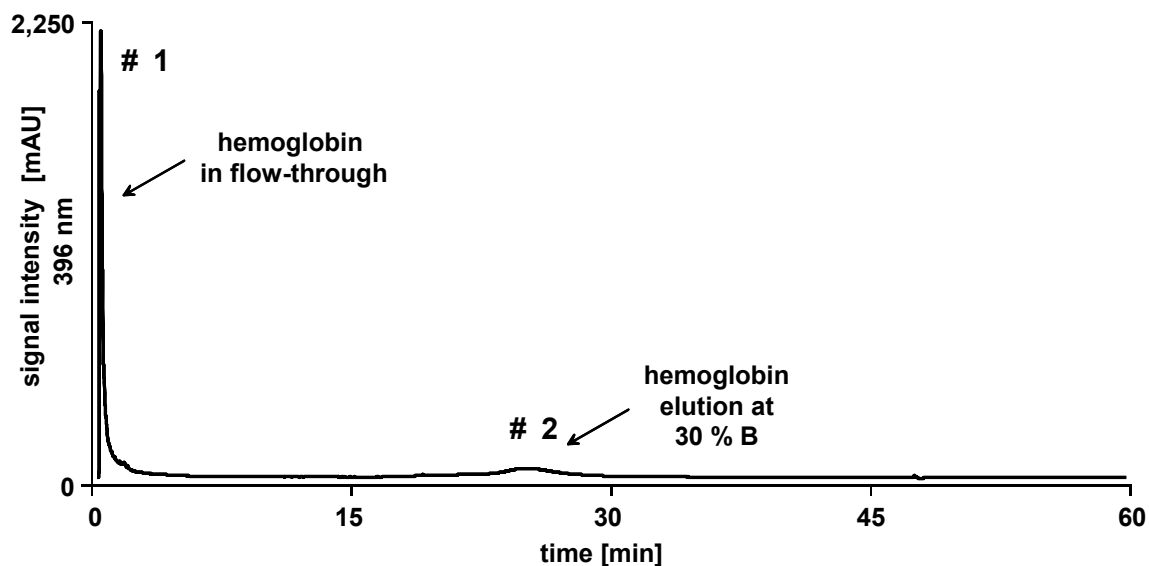


Fig. 67. Chromatogram of hemoglobin at pH 8.3. Column, LiChrospher ADS Merck; sample, 150 mg/mL hemoglobin diluted in 5 mmol/L imidazole; 10 μ L injection; mobile phase, (A) 5 mmol/L imidazole in H₂O, (B) 5 mmol/L imidazole in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

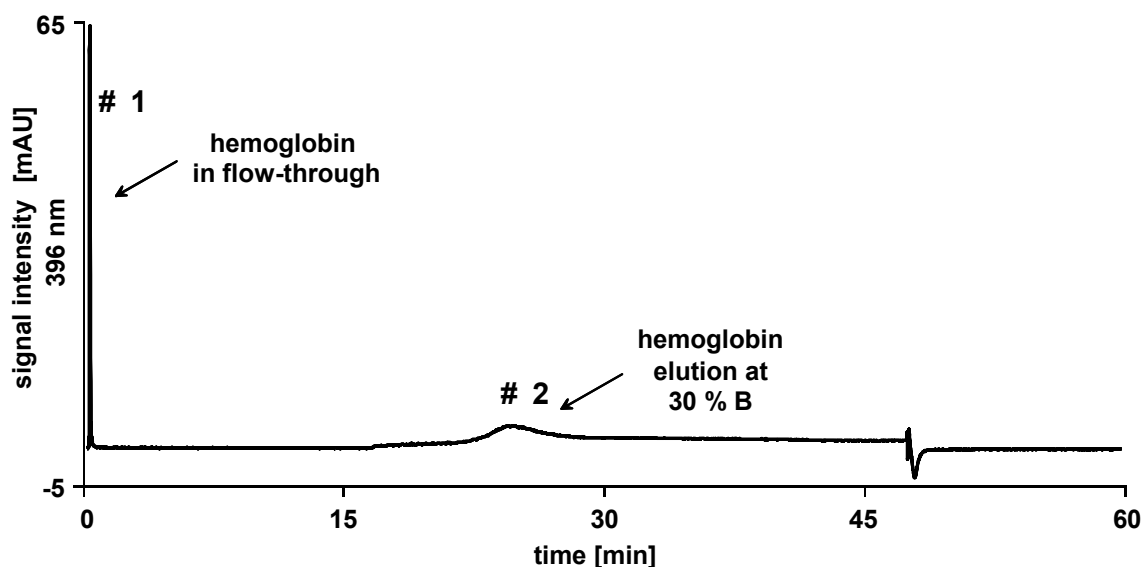


Fig. 68. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, LiChrospher ADS Merck; sample, 5 mmol/L imidazole; 10 μ L injection; mobile phase, (A) 5 mmol/L imidazole in H₂O, (B) 5 mmol/L imidazole in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

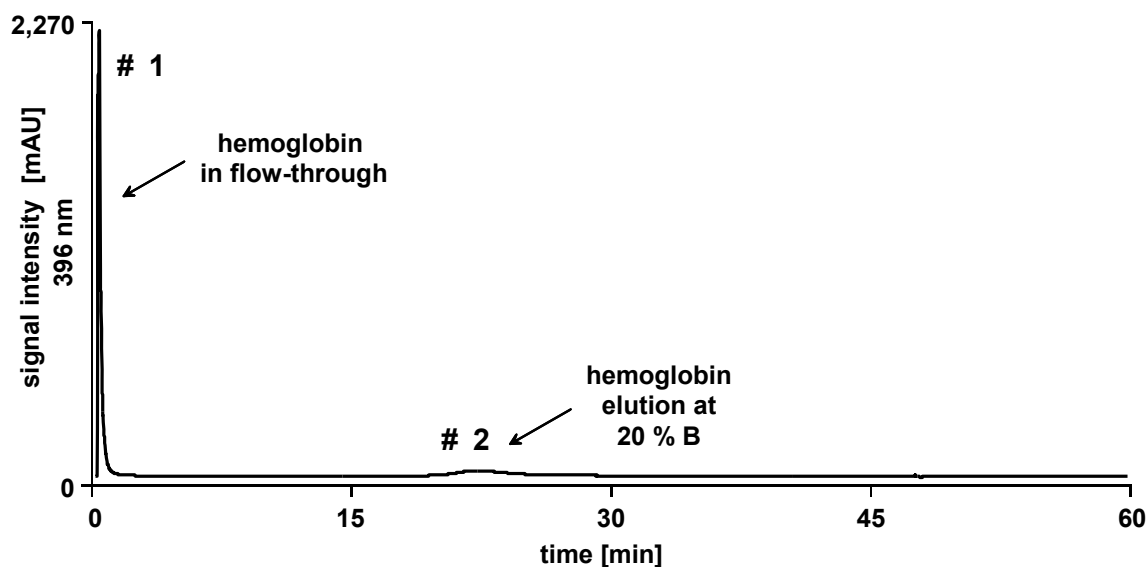


Fig. 69. Chromatogram of hemoglobin at pH 10.7. Column, LiChrospher ADS Merck; sample, 150 mg/mL hemoglobin diluted in 10 mmol/L ethanolamine; 10 μ L injection; mobile phase, (A) 10 mmol/L ethanolamine in H₂O, (B) 10 mmol/L ethanolamine in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

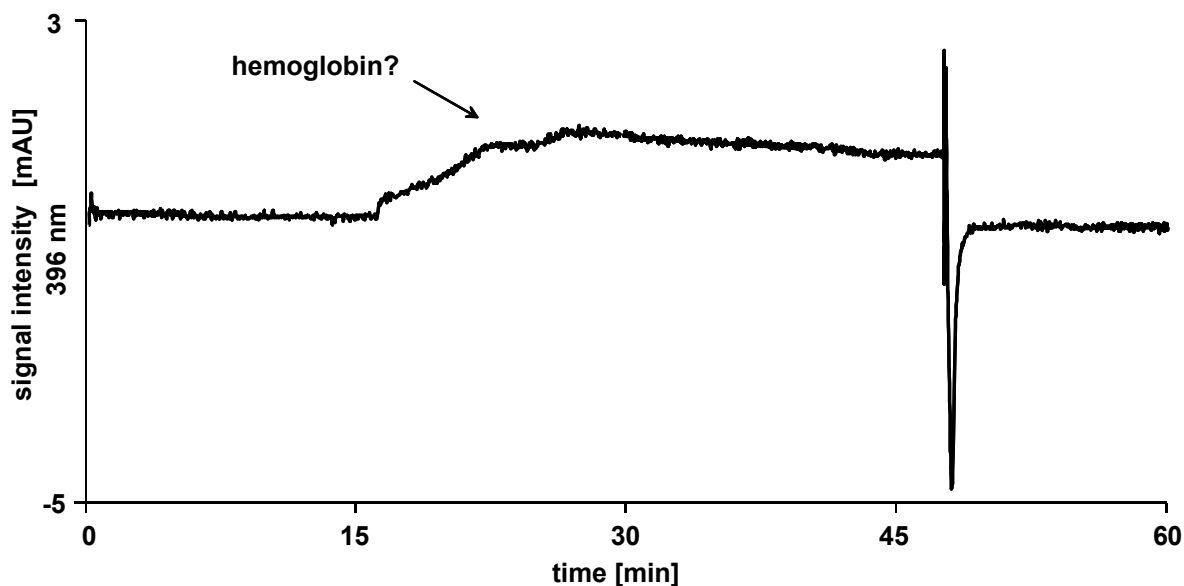


Fig. 70. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, LiChrospher ADS Merck; sample, 10 mmol/L ethanolamine; 10 μ L injection; mobile phase, (A) 10 mmol/L ethanolamine in H₂O, (B) 10 mmol/L ethanolamine in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 30.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

Despite difficulties of quantitation because of very large and flat peaks, the tests performed at four different pH show that the higher the pH, the lower the retention and the carry-over of hemoglobin. The results are summarized in Tab. 12.

Tab. 12. Hemoglobin exclusion and carry-over on a LiChrospher ADS column used at different pH.

pH	ratio A _{#2} / A _{#1}	hemoglobin in flow-through	hemoglobin retained on the column	hemoglobin in wash peak	carry-over
2.1	5.0	X	X	-	X
6.6	0.2	X	X	X	X
8.3	0.3	X	X	-	disappear after 1 injection
10.7	0.2	X	X	-	not detected

Practically no carry-over was observed at pH 10.7. This pH value appears to be optimal to avoid adsorption of hemoglobin and memory effects. However, silica-based columns are not stable at this pH and routine analyses can not be envisaged at such a high pH. For this reason, LiChrospher ADS can not be routinely operated under such basic conditions. However, among the set of the six restricted access media previously tested, one of them (Biotrap 500 MS from Chromtech) is a polymer-based material (certified stable up to pH 11) and will be tested at high pH in further experiments.

7 Evaluation of Biotrap 500 MS at pH 10.7

7.1 Retention of hemoglobin on Biotrap 500 MS at pH 10.7

Hemoglobin retention and carry-over were evaluated with Biotrap 500 MS at pH 10.7 (H₂O/ACN + 10 mmol/L ethanolamine). The same analytical setup as before was used (see 2.4.1). The gradient was 0-100 % acetonitrile in 15 min followed by a wash step consisting of 2 min at 100 % acetonitrile.

By injection of 10 μ L of a 150 mg/mL hemoglobin solution, most of the hemoglobin is flowing through the column (~ 80 %). The retained molecules (~ 20 %) are eluted with about 20 % acetonitrile. The corresponding chromatogram is depicted in Fig. 71. In a blank injection following the injection of hemoglobin practically no carry-over is observed as depicted in Fig. 72.

High pH values seem to be really a key point for the break-through of hemoglobin in the flow-through volume of RAM columns as observed with LiChrospher ADS from Merck or with Biotrap 500 MS from Chromtech.

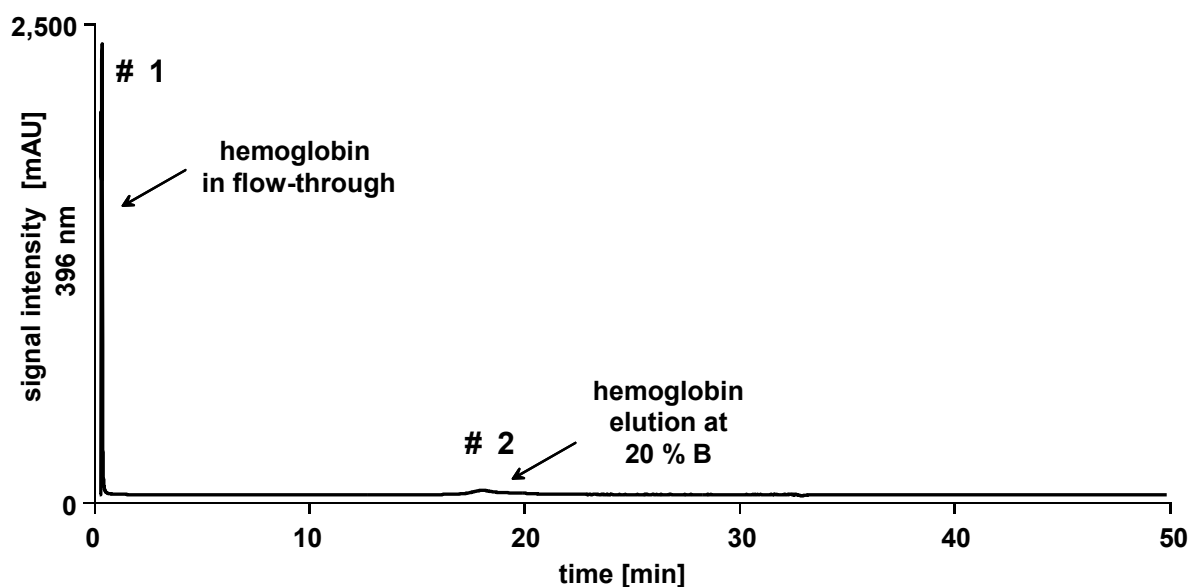


Fig. 71. Retention of hemoglobin on Biotrap 500 MS Chromtech; sample, 150 mg/mL hemoglobin diluted in 10 mmol/L ethanolamine; 10 μ L injection; mobile phase, (A) 10 mmol/L ethanolamine in H₂O, (B) 10 mmol/L ethanolamine in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 15.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

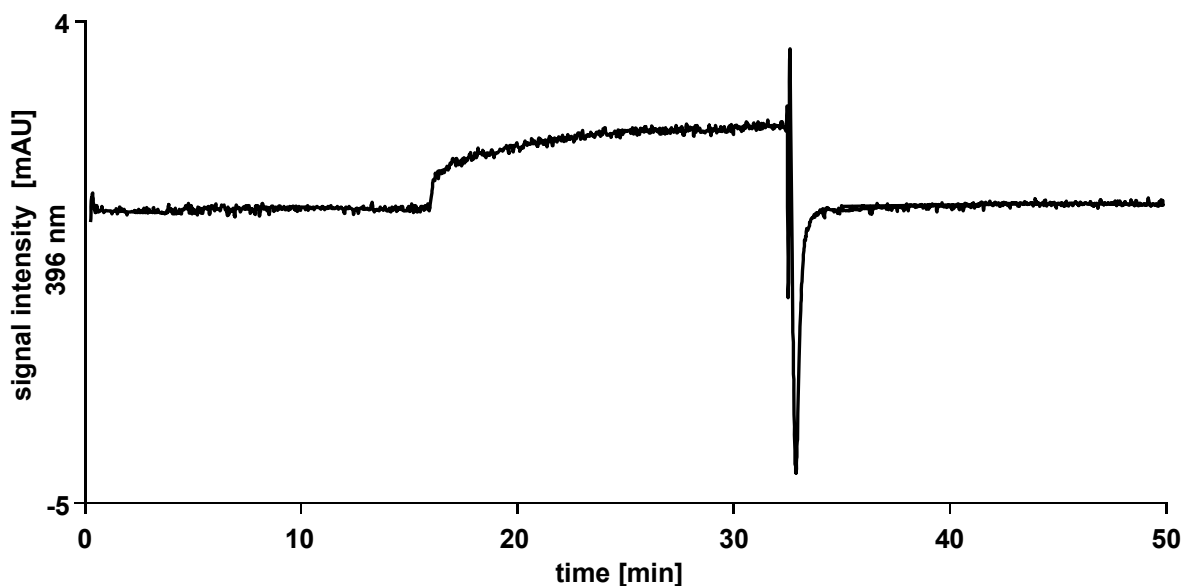


Fig. 72. Hemoglobin carry-over: first blank injection after hemoglobin sample. Column, Biotrap 500 MS Chromtech; sample, 10 mmol/L ethanolamine; 10 μ L injection; mobile phase, (A) 10 mmol/L ethanolamine in H₂O, (B) 10 mmol/L ethanolamine in ACN; 15.0 min isocratic at 100 % A, then linear gradient, 0-100 % B in 15.0 min, and 2.0 min at 100 % B; flow rate, 1.5 mL/min; room temperature; detection at 396 nm.

7.2 Extraction of analytes with Biotrap 500 MS at pH 10.7

Two analytes (one neutral and one charged) were chosen as model compounds to characterize the extraction capability of Biotrap 500 MS at pH 10.7 (10 mmol/L ethanolamine). The first analyte was tetracycline hydrochloride. This antibiotic, the structure of which is depicted in Fig. 73, is water soluble and neutral at high pH. Its molecular formula is C₂₂H₂₄N₂O₈.HCl and its molecular weight is 480.90 g/mol. Mass spectrometric detection can be performed by monitoring *m/z* 445. The quantitation of tetracycline in biological fluids is generally achieved by HPLC-MS/MS, after protein precipitation^[128]. Tetracycline and derivatives are widely used in swine production^[129]. Concentrations are generally in the ng/ μ L range^[130-132].

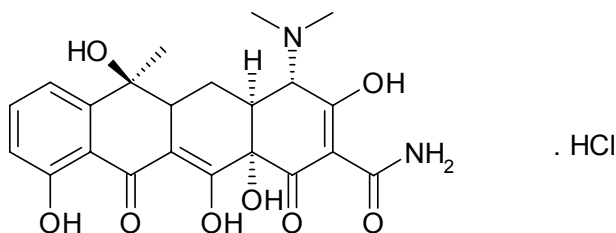


Fig. 73. Structure of tetracycline hydrochloride.

The second analyte was 5-methyltetrahydrofolic acid. Its structure is depicted in Fig. 74. Its molecular formula is $C_{20}H_{23}N_7Na_2O_6 \cdot HCl$ and its molecular weight is 503.42 g/mol. Mass spectrometric detection can be performed by monitoring m/z 460. At high pH (and particularly during the extraction step with 10 mmol/L ethanolamine), the molecule is present as a doubly negatively charged compound. 5-methyltetrahydrofolic acid is very sensitive to oxidation but can be stabilized by spiking ascorbic acid [133]. The concentration of 5-methyltetrahydrofolate is ranging from 2 to 20 ng/mL in serum and from 100 to 200 ng/mL in erythrocytes. Quantitation is generally performed by solid phase extraction followed by LC-MS/MS or microbiologic assay [134;135].

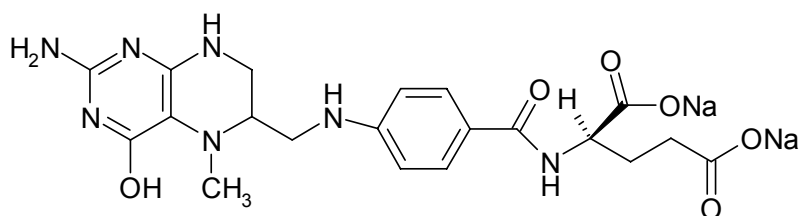


Fig. 74. Structure of 5-methyltetrahydrofolic acid disodium salt.

A synthetic whole blood solution was prepared by spiking human hemoglobin at 175 mg/mL in human serum. Before use and to avoid HPLC system plugging, the solution was filtrated over a Vivaclear mini 0.5 clarifying filter membrane.

Sixty ng tetracycline were spiked in 300 μ L of the synthetic blood solution. Thus the concentration of tetracycline was 200 pg/ μ L. The sample was diluted to 900 μ L with addition of 600 μ L of 10 mmol/L ethanolamine. The sample was injected over Biotrap 500 MS and washed for 5 min with 10 mmol/L ethanolamine at 3.2 mL/min. Then the switching valve was turned and elution occurred at 300 μ L/min over the C18 column. Eluent A was 0.05 % aqueous TFA and eluent B was 0.05 % TFA in ACN. The linear gradients were 0 to 50 % B in 7.5 min, followed by 50 to 100 % B in 2.5 min and isocratic conditions at 100 % B for 2 min. Meanwhile the connections of the injection system were washed by performing a linear gradient of acetonitrile (0 to 100 % in 4 min). Detection was performed in Single Ion Monitoring mode on m/z 445.2 ± 1.0 . The corresponding chromatogram is depicted in Fig. 75.

This experiment shows that it is possible to extract neutral analytes such as 4-tetracycline from “synthetic blood hemolysates” with Biotrap 500 MS under high pH

conditions, compatible with low hemoglobin carry-over. Tetracycline at concentrations down to 200 pg/ μ L in “synthetic blood hemolysate” was detected in these preliminary experiments with a linear quadrupole mass analyzer. This value can be undoubtedly reduced by performing MRM detections on triple quadrupole analyzers.

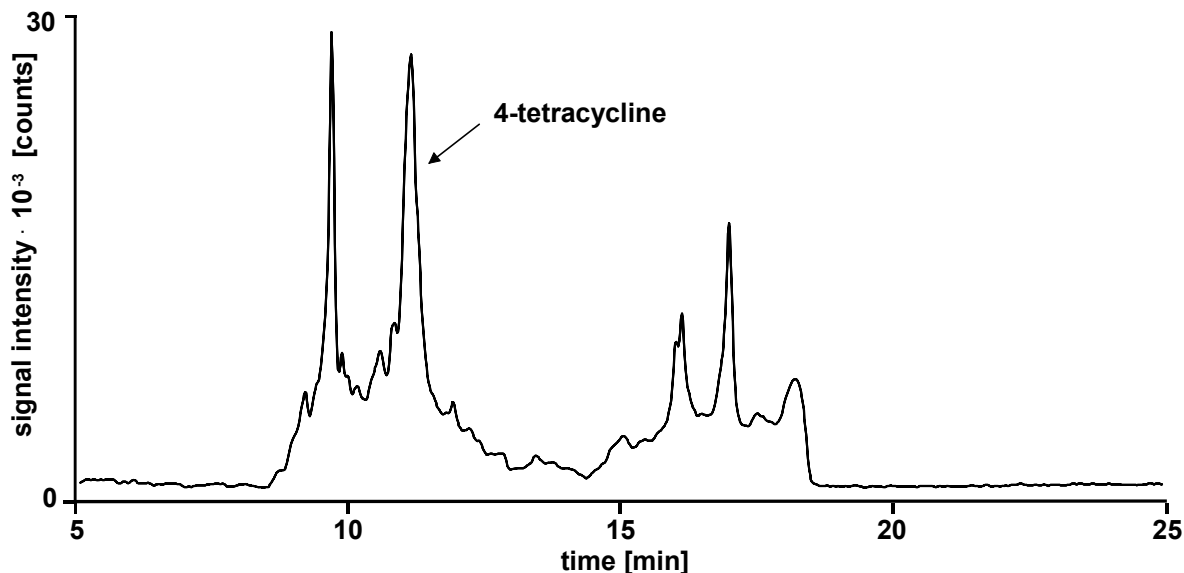


Fig. 75. Extraction of 200 pg/ μ L tetracycline hydrochloride from “synthetic blood hemolysate” sample (175 mg/mL hemoglobin spiked in serum).

5-min extraction: column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; sample, 300 μ L of 200 pg/ μ L tetracycline hydrochloride spiked in “synthetic blood hemolysate” and diluted 1/3 with 10 mmol/L ethanolamine prior to injection.

Back flush elution: column, Prontosil 300-5-C18-H 5 μ m, Bischoff, 125 x 2.0 mm i.d.; mobile phase, (A) 0.05 % aqueous TFA, (B) 0.05 % TFA in ACN; linear gradient, 0-50 % B in 7.5 min; 50-100 % B in 2.5 min, isocratic conditions at 100 % B for 2 min, flow rate, 0.3 mL/min; room temperature. Detection with Surveyor MSQ in selected ion monitoring mode targeted on m/z 445.2 \pm 1.0.

We intended to perform the same experiment with 5-methyltetrahydrofolate as an analyte. However, the analyte did not show any retention on the Biotrap 500 MS column because of the predominance of negative species at pH 10.7 (see Fig. 76). To circumvent the inability of Biotrap 500 MS to retain charged compounds (mostly negatively charged because of high pH values), the use of ion pair reagents such as butyldimethylammonium bicarbonate could be considered.

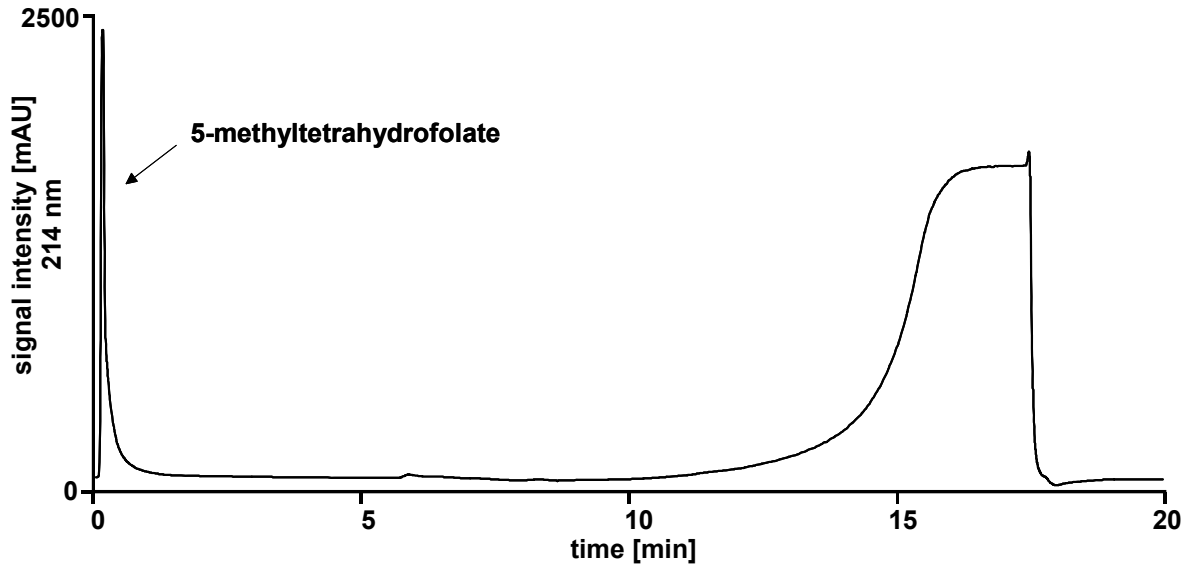


Fig. 76. Injection at pH 10.7 of 5-methyltetrahydrofolate over Biotrap 500 MS. Column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phases, (A) 10 mmol/L ethanolamine, (B) 10 mmol/L ethanolamine in ACN; gradient: isocratic conditions at 100 % A for 5 min, then 0-100 % B in 10 min, followed by isocratic conditions at 100 % B for 2 min; flow rate, 1.5 mL/min; room temperature; sample, 10 μ g 5-methyltetrahydrofolate and 10 μ g ascorbic acid.

8 Quantitation of tetracycline hydrochloride in human whole blood hemolysates

The extraction of tetracycline hydrochloride from “synthetic blood hemolysates” (serum spiked with hemoglobin) has already been achieved with Biotrap 500 MS at pH 10.7. However, the applicability of the developed method to real samples was not proven. This has been investigated on real hemolysates as follows.

Whole blood hemolysate samples were injected over Biotrap 500 MS mounted in the two-dimensional HPLC-MS setup. After flowing through of hemoglobin, the analyte of interest was transferred to mass spectrometric detection by switching the 6-port valve.

8.1 Limit of detection

Thirty ng tetracycline hydrochloride were spiked in 3 mL whole blood hemolysate (150 μ L blood + 1,350 μ L 150 mmol/L NaCl + 1,500 μ L 4 % lysis reagent). The resulting concentration of tetracycline hydrochloride corresponds to 200 pg/ μ L in whole blood. Then, 2.5 mL of the sample were injected over Biotrap 500 MS and washed with 10 mmol/L ethanolamine at 3.2 mL/min. After 18 min the 6-port valve was switched and elution occurred over the analytical column at 300 μ L/min. (A: H₂O + 0.05 % TFA, B: ACN + 0.05 %, linear gradient 10-30 % B in 7.5 min followed by 30-100 % B in 2.5 min and isocratic conditions for 2 min). Mass spectrometric detection was performed by monitoring m/z 445.2 \pm 1.0. Single ion monitoring chromatograms corresponding to a blank run and to an injection of 200 pg/ μ L tetracycline hydrochloride are depicted in Fig. 77 and in Fig. 78. As observed, tetracycline was still well detected at 200 pg/ μ L in whole blood hemolysate. However, the neighboring elution of a compound with the same m/z does not permit to detect much less concentrated samples.

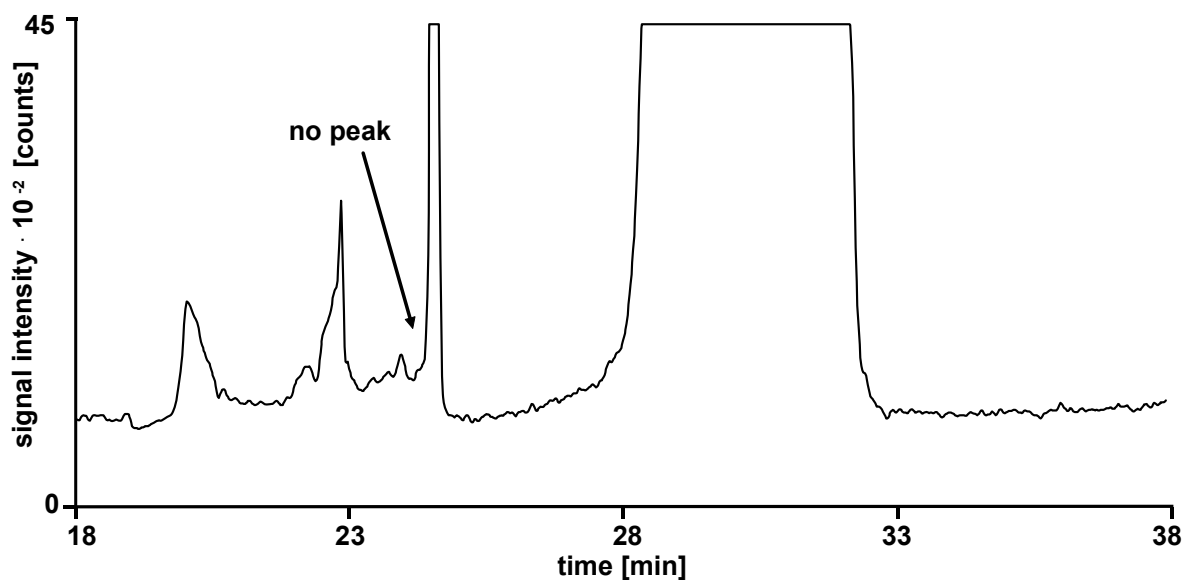


Fig. 77. Blank injection of whole blood hemolysate.

18-min extraction: column, Biotrap 500 MS, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; sample, 2.5 mL of a blood hemolysate obtained by mixing 150 μ L whole blood, 1,350 μ L 150 mmol/L NaCl and 1,500 μ L 4 % lysis reagent.

Back flush elution: column, Prontosil 300-5-C18-H 5 μ m, 125 x 2.0 mm i.d.; mobile phase, (A) 0.05 % aqueous TFA, (B) 0.05 % TFA in ACN; linear gradient, 10-30 % B in 7.5 min; 30-100 % B in 2.5 min, isocratic conditions at 100 % B for 2 min, flow rate, 0.3 mL/min; room temperature. Detection in SIM mode targeted on m/z 445.2 \pm 1.0.

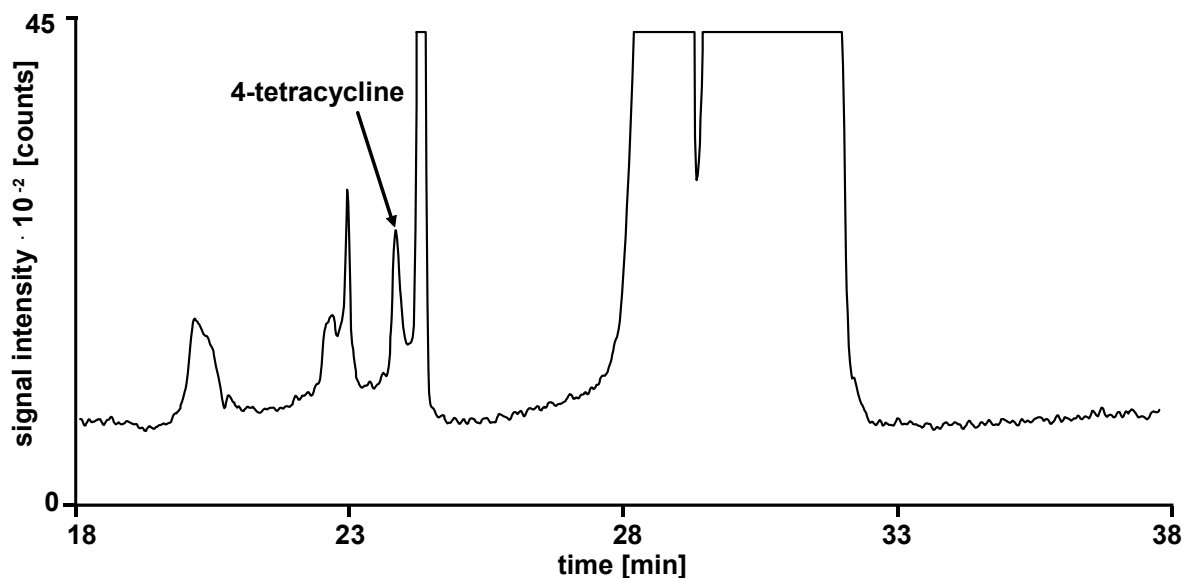


Fig. 78. Extraction of tetracycline from whole blood hemolysate.

Conditions as in Fig. 77; sample, 2.5 mL of a blood hemolysate obtained by mixing 150 μ L whole blood, 1,350 μ L 150 mmol/L NaCl and 1,500 μ L 4 % lysis reagent and spiked with 30 ng tetracycline. Tetracycline concentration in blood was 200 μ g/ μ L.

8.2 Carry-over

To check the carry-over of hemoglobin on biotrap 500 MS, 2.5 mL of a whole blood hemolysate were injected into the HPLC-MS system. The loading step on Biotrap 500 MS was performed with 10 mmol/L ethanolamine at 3.2 mL/min for 18 min. The 6-port valve was then switched and back flush elution was performed with a gradient of ACN + 0.05 % TFA (10-30 % B in 7.5 min, 30-100 % B in 2.5 min, 100 % B for 2.0 min, and 0 % B for 7.9 min). During the back flush elution, the sample loop was washed by performing a gradient of ACN + 10 mmol/L ethanolamine (0 % B for 2 min, 0-100 % B in 4 min, 100 % B for 6 min, and 0 % B for 7.9 min).

The carry-over of hemoglobin on Biotrap 500 MS was checked by performing a blank chromatographic run (see Fig. 80) after such an injection of hemolysate (see Fig. 79). Detection was performed at 396 nm.

No carry-over was observed. This result is similar to a blank chromatographic run performed after an injection of “synthetic blood hemolysate” (hemoglobin spiked in human serum) as described in section 7. The lysis of erythrocytes with the lysis reagent do not seem to influence the carry-over of hemoglobin on Biotrap 500 MS at pH 10.7.

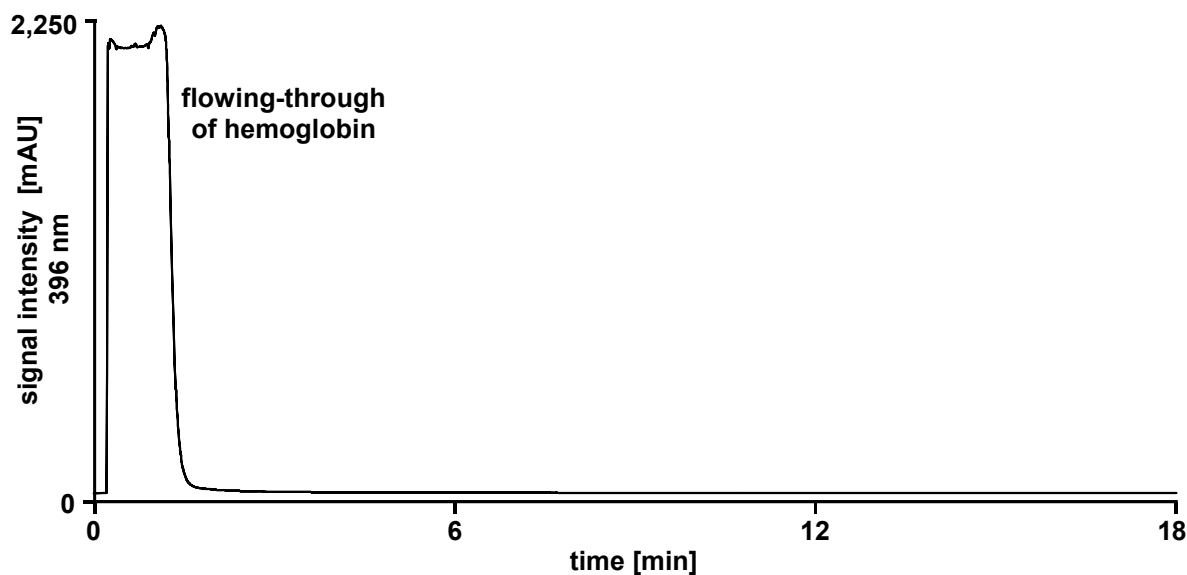


Fig. 79. Flowing-through of hemoglobin on Biotrap 500 MS.

18-min extraction: column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; sample, 2.5 mL of a blood hemolysate obtained by mixing 150 μ L whole blood, 1,350 μ L 150 mmol/L NaCl and 1,500 μ L 4 % lysis reagent; detection at 396 nm.

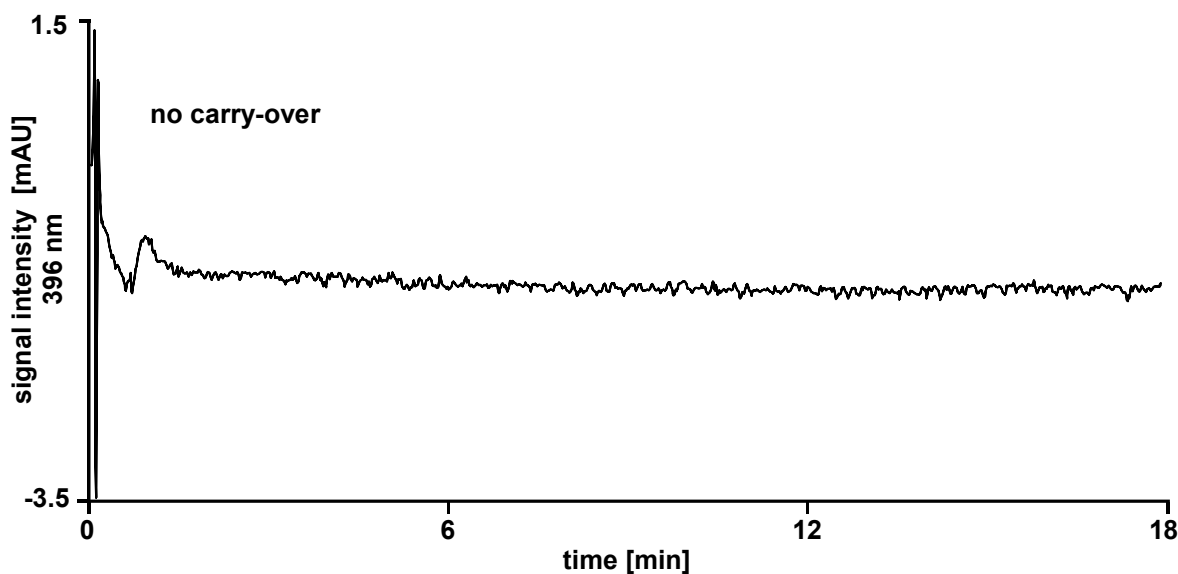


Fig. 80. Blank chromatographic run on Biotrap 500 MS after an injection of hemolysate.

18-min extraction: column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; previous sample, 2.5 mL of a blood hemolysate obtained by mixing 150 μ L whole blood, 1,350 μ L 150 mmol/L NaCl and 1,500 μ L 4 % lysis reagent; detection at 396 nm.

8.3 Calibration curve

Five samples consisting of tetracycline hydrochloride spiked in blood hemolysates were prepared. The concentrations of tetracycline hydrochloride in whole blood were 200 μ g/ μ L, 600 μ g/ μ L, 1,000 μ g/ μ L, 1,500 μ g/ μ L, and 2,000 μ g/ μ L. The corresponding concentrations in the hemolysate solutions to be injected were: 9.5 μ g/ μ L, 28.5 μ g/ μ L, 47.5 μ g/ μ L, 70.9 μ g/ μ L, and 94.3 μ g/ μ L. For each sample, 2.5 mL were injected and analyzed after an 8-min extraction. Peak areas corresponding to tetracycline were integrated. Values are only approximated because of difficulties of integration due to co-elution of tetracycline with another compound, and peak broadening at high concentrations. Example chromatograms are depicted in Fig. 81.

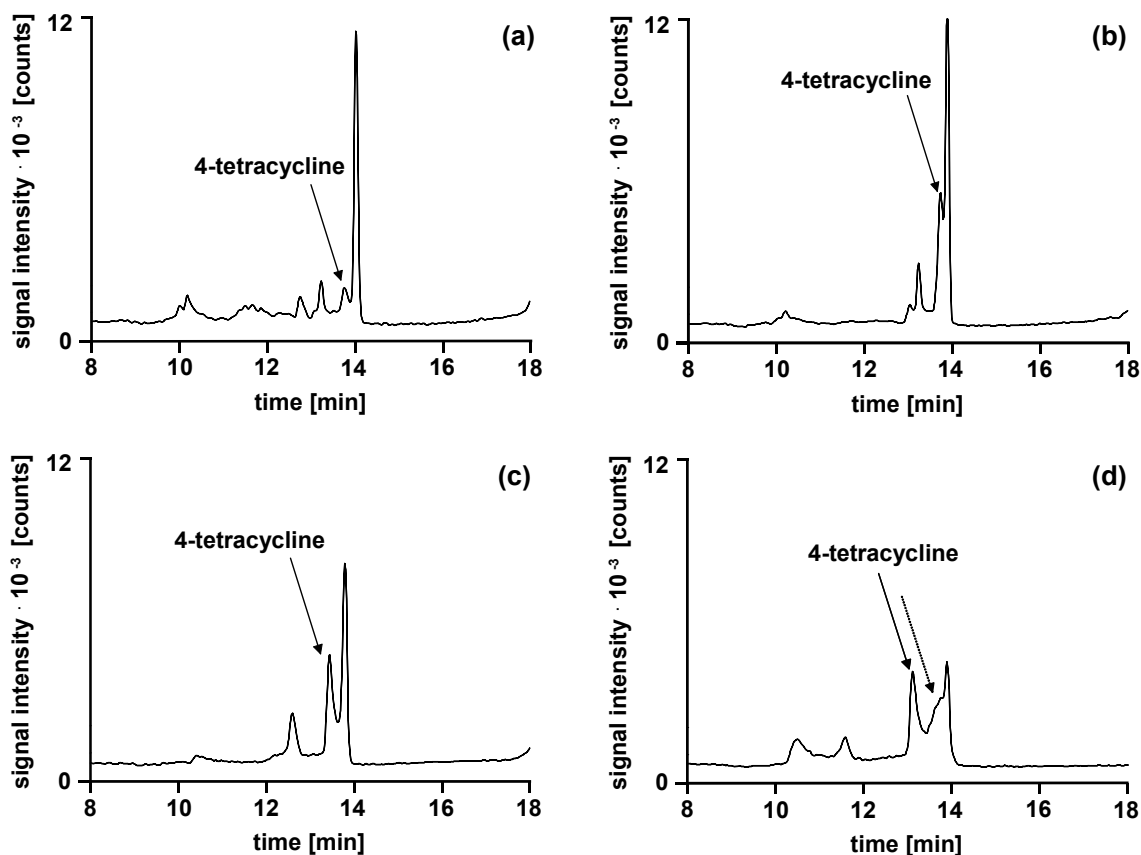


Fig. 81. Extraction of tetracycline from human whole blood hemolysates. 8-min extraction: column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; samples, 2.5 mL of blood hemolysates containing tetracycline at (a) 9.5 pg/ μ L, (b) 28.5 pg/ μ L, (c) 47.5 pg/ μ L, and (d) 70.9 pg/ μ L. Back flush elution: column, Prontosil 300-5-C18-H 5 μ m, Bischoff, 125 x 2.0 mm i.d.; mobile phase, (A) 0.05 % aqueous TFA, (B) 0.05 % TFA in ACN; linear gradient, 10-30 % B in 7.5 min; 30-100 % B in 2.5 min, isocratic conditions at 100 % B for 2 min, flow rate, 0.3 mL/min; room temperature. Detection with Surveyor MSQ in selected ion monitoring mode targeted on m/z 445.2 \pm 1.0.

A calibration curve was computed and is plotted in Fig. 82. A linear correlation is observed ($R^2 = 0.945$), proving the ability of Biotrap 500 MS to quantitatively extract antibiotics from whole blood hemolysates at physiologically meaningful concentrations (200 pg/ μ L in whole blood). Multiple reaction monitoring (MRM) detections on triple quadrupole analyzers should permit to avoid integration errors due to coelution of the analyte with other compounds. MRM detections should also permit to get lower limits of detection.

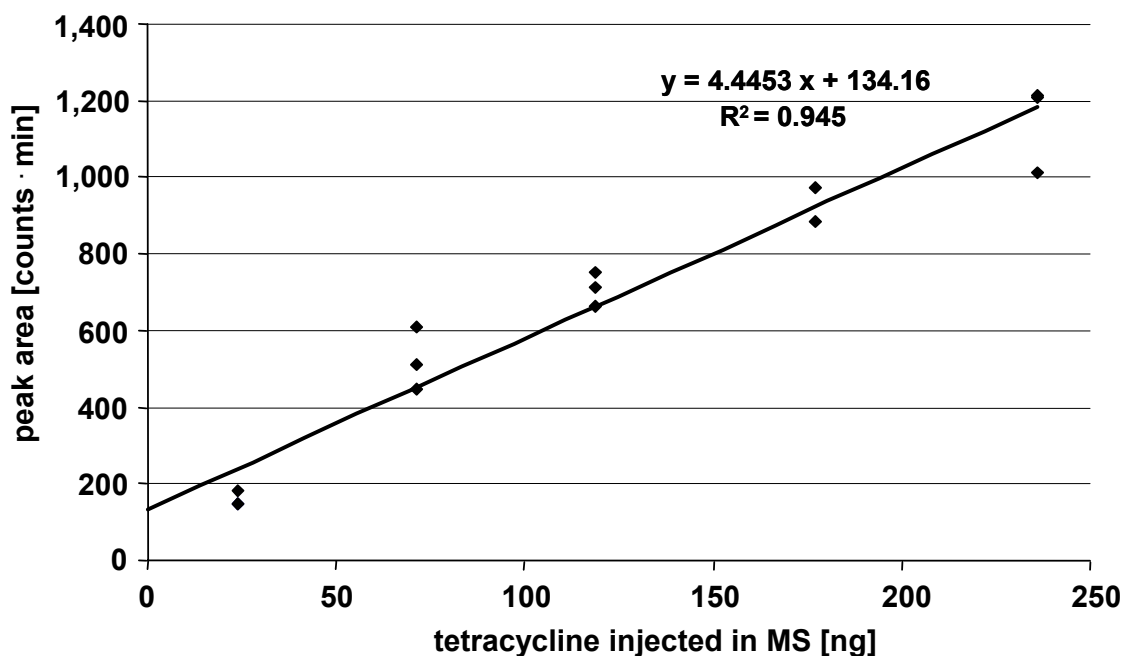


Fig. 82. Tetracycline recovery in human whole blood hemolysates.

8-min extraction: column, Biotrap 500 MS, Chromtech, 13 x 4 mm i.d.; mobile phase, (A) 10 mmol/L ethanolamine; flow rate, 3.2 mL/min; room temperature; samples, 2.5 mL of blood hemolysates containing tetracycline at 9.5 pg/ μ L, 28.5 pg/ μ L, 47.5 pg/ μ L, 70.9 pg/ μ L, and 94.3 pg/ μ L.

Back flush elution: column, Prontosil 300-5-C18-H 5 μ m, Biscoff, 125 x 2.0 mm i.d.; mobile phase, (A) 0.05 % aqueous TFA, (B) 0.05 % TFA in ACN; linear gradient, 10-30 % B in 7.5 min; 30-100 % B in 2.5 min, isocratic conditions at 100 % B for 2 min, flow rate, 0.3 mL/min; room temperature. Detection with Surveyor MSQ in selected ion monitoring mode targeted on m/z 445.2 \pm 1.0.

A linear correlation is observed but peak integration was rather difficult because of co-elution of tetracycline with another compound.

9 Conclusions

Six different restricted access materials have been evaluated with respect to their ability to remove hemoglobin from hemolysates. In general, all six columns showed similar behavior: at pH 2.1, RAM materials showed significant adsorption and memory effects for hemoglobin.

Derivatization of LiChrospher ADS material with aminodextran did not decrease the adsorption of proteins. Derivatizations with polyethyleneimines and polylysine generally increased the adsorption of proteins on the LiChrospher ADS material. However, no correlation between properties of proteins (e.g. molecular weight, pI) and adsorption was observed.

Experiments at different pH (2.1-10.7) revealed that the retention of hemoglobin can be suppressed at alkaline pH (10 mmol/L ethanolamine, pH 10.7). A small amount of hemoglobin was still retained; however, carry-over between different injections was completely eliminated. Because of better chemical stability at elevated pH, the polymeric Biotrap 500 MS RAM column was optimized for the analysis of hemolysates.

Tetracycline spiked into real hemolysates could be extracted with Biotrap 500 MS at alkaline pH, and detected down to 200 pg/ μ L concentrations using a linear quadrupole mass spectrometer operated in selected ion monitoring mode. The ability of the developed setup to quantitatively extract antibiotics from whole blood hemolysates at biologically relevant concentrations ^[128,129], and without carry-over of hemoglobin was proven. Better detection limits and less interference with matrix compounds should be achievable using selected reaction monitoring tandem mass spectrometry.

Chapter V

Development and evaluation of
multidimensional HPLC-MS
systems for proteome analysis

V. Development and evaluation of multidimensional HPLC-MS systems for proteome analysis

In the field of proteomics, the detection of all the components present in a sample may be required. For instance, differential analysis of two cell states is performed by checking in two cell extracts the presence (or the absence) of all the presumably expressed proteins. Such a holistic strategy (holistic, greek *holos*, “whole”) requires analytical methods able to separate and identify a huge set of proteins and peptides. In this framework, two multidimensional setups for the separation of complex mixtures of peptides were developed and evaluated with tryptic digests of protein cell extracts of *Corynebacterium glutamicum*.

Corynebacterium glutamicum is a gram positive soil bacterium used for the industrial production ^[136] of L-glutamate and L-lysine (1,500,000 to/a, and 700,000 to/a, respectively). Aside this evident industrial interest, *C. glutamicum* was chosen as a model organism for various reasons. Although related to pathogenic bacteria such as *Corynebacterium diphtheriae* or *Mycobacterium tuberculosis*, *C. glutamicum* is itself not pathogenic and did not require any special security equipment in the lab. Metabolic pathways of *C. glutamicum* are known ^[137] and the genome was available in database ^[138-140]. The annotation of around 3,000 proteins permitted us to perform peptide and protein identification with tandem mass spectrometry and database search.

The first investigated scheme of separation followed a classical off-line strong cation exchange (SCX) separation, followed with ion-pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC) at pH 2.1. The second separation scheme was a very new approach based on a peptide separation with RP-HPLC at pH 10.0, a fraction collection, and a separation with IP-RP-HPLC at pH 2.1. Peptide identification was achieved in both setups with electrospray ionization tandem mass spectrometry (ESI-MS/MS).

In the following chapter, both approaches are described and evaluated. The new RP x IP-RP separation scheme is compared to the classical SCX x IP-RP approach in terms of peptide separation, proteome coverage, dimension repeatability, dimension orthogonality, and method complementarity.

1 Materials and methods

1.1 Chemicals

Deionized water (18.2 MΩ cm) was prepared with a Purelab Ultra Genetic system (Elga, Griesheim, Germany). Acetonitrile (E Chromasolv) and 2-mercaptoethanol (> 98 %) were purchased from Sigma-Aldrich (Steinheim, Germany). Analytical reagent grade sodiumdihydrogen-phosphate-1-hydrate, acetic acid (100 %), and triethylamine (> 99 %) were obtained from Merck KGaA (Darmstadt, Germany). Sodium chloride was supplied by Grüssing GmbH (Filsum, Germany). Urea (≥ 99.5 %), ortho-phosphoric acid (85 %), ammonium hydrogencarbonate (≥ 99.5 %), ammonium formate (> 97 %), iodacetic acid (≥ 99.5 %), formic acid (88-91 %), heptafluorobutyric acid (≥ 99.0 %), and trifluoroacetic acid (≥ 99.5 %) were purchased from Fluka (Buchs, Switzerland). Sequencing grade modified trypsin was supplied by Promega (Madison, WI, USA) and Slide-A-Lyzer dialysis cassettes by Perbio Science (Bonn, Germany).

1.2 Preparation of tryptic digests

Protein cell extracts of *Corynebacterium glutamicum* (wild type) were kindly obtained from the workgroup of Prof. E. Heinzle (Biochemical Engineering, Saarland University, Saarbrücken, Germany). Protein concentrations were evaluated by Bradford tests ^[141]. Proteins were denaturated by addition of 120 µL 8 mol/L urea, 500 mmol/L ammonium hydrogencarbonate to 200 µL of cell lysate. The sample was incubated for one hour at 37°C under gentle agitation (600 rpm). The disulfide bridges were then reduced by addition of 12 µL 300 mmol/L dithiothreitol, sample degazing under argon, and incubation for 2 hours at 37°C at 600 rpm. After sample cooling, 8 µL 2 mol/L iodacetic acid were added for carboxymethylation and the sample was incubated for 30 minutes at room temperature under light protection. Excess of iodacetic acid was eliminated with 16 µL 1 mol/L 2-mercaptoethanol and incubation for 20 minutes at room temperature. Finally the sample was dialyzed for 16 hours against 1 L distilled water in a dialysis cassette (3,500 MW cut-off membranes). Trypsin (1 µg for 50 µg protein) was activated for 30 minutes at 30°C in 50 mmol/L acetic acid and incubated in the protein solution for a 24-hour digestion. Finally, the digestion was quenched by addition of trifluoroacetic acid (1.0 % (v/v) end

concentration). The protein digest was centrifuged for 5 minutes at 13,000 rpm. The supernatant was collected and split into aliquots.

1.3 Analytical setups for the first separation steps

A first chromatographic setup designed for flow rates of 1 mL/min consisted of an Agilent 1050 series HPLC system (Waldbronn, Germany), and a Rheodyne injection system (Rohnert Park, CA, USA) with a 400 μ L external loop. UV detection was monitored at 214 nm with a Spectromonitor 3100 from Milton Roy (Ivyland, PA, USA). The flow cell had an optical pathway of 10 mm (volume: 14 μ L). Eluents were degassed with helium. This setup was used to run a 250 x 4.0 mm ProPacTM SCX-10 column preceded by a 100 x 4.0 mm guard column from Dionex (Idstein, Germany).

A second chromatographic setup adapted for flow rates of 200-250 μ L/min consisted of a low-pressure Rheos 2000 HPLC system (Flux Instruments, Basel, Switzerland), a degasser from Knauer GmbH (Berlin, Germany), and an injection system from Rheodyne (Model 7725, Rohnert Park, CA, USA) with a 400 μ L external loop. UV detection was monitored at 214 nm with a capillary detector 433 purchased from Kontron AG (Zuerich, Switzerland). The flow cell had an optical pathway of 5 mm (volume: 1 μ L). This setup was used to run a 250 x 2.0 mm ProPacTM SCX-10 column preceded by a 50 x 2.0 mm guard column from Dionex (Idstein, Germany), a 200 x 2.1 mm PolySULFOETHYL column from PolyLC (Southboro, MA, USA) preceded by a 10 x 2.1 mm guard column, and a 150 x 2.0 mm 3 μ m C18 Gemini column from Phenomenex (Aschaffenburg, Germany).

After collection, the fractions were evaporated with a vacuum concentrator (model 5301) supplied by Eppendorf AG (Hamburg, Germany).

1.4 Second separation step and data acquisition

The second separation setup consisted of a 2D capillary/nano system from LCPackings (Amsterdam, The Netherlands), equipped with an Ultimate low-pressure gradient micro-pump (model Ultimate), a Switchos micro-column 10-port switching unit with loading pump (model Switchos), and a micro-injector (model Famos). Trap- (10 x 0.2 mm) and analytical columns (60 x 0.1 mm) were poly-

styrene/divinylbenzene (PS-DVB) monoliths prepared according to the already published procedure ^[120]. Eluents were degassed with helium.

An ion-trap mass spectrometer (model esquire HCT) from Bruker Daltonics (Bremen, Germany) with a modified ESI-ion source (spray capillary: fused silica capillary, 0.090 mm o.d., 0.020 mm i.d.) was utilized as detector. The instrument was operated in data-dependent mode. MS/MS spectra were recorded in positive ion mode with an electrospray voltage of 3,500 V and fragmentation amplitude ramped from 0.5 to 3.0 V. The heated capillary temperature was set to 300°C. The following mass spectrometric parameters were applied for automated peptide identifications by data-dependent tandem mass spectrometry: mass range mode, ultra scan 50 – 3,000 *m/z*; scan speed, 26,000 *m/z* per s; full scan, 450 – 1,500 *m/z*; ion polarity, positive; trap drive, 93.2; octapole RF amplitude, 88.5 Vpp; lens 2, -36.1 V; capillary exit, 253.8 V; nebulizer gas, 20 psi; dry gas, 4 L/min; end-plate high-voltage offset, -500 V; ICC target, 70,000; maximum accumulation time, 200 ms; precursor ions auto MS⁽ⁿ⁾, 3; MS averages, 5 spectra; MS/MS scan range, 200 – 2,000 *m/z*; active exclusion, after 2 spectra for 0.50 min; MS/MS fragmentation amplitude, 1.5 V; smart fragmentation, on (30 – 200 %); absolute threshold MS/MS, 4,500.

1.5 Data processing and evaluation

Mass spectra processing was performed with Data Analysis 3.3 from Bruker Daltonics (Bremen, Germany). Database searches were performed against an in-house database containing the sequenced proteins of *Corynebacterium glutamicum* ATCC 13032 Kitasato. The database (2,993 entries) was downloaded from the Institute for Genomic Research (TIGR) under <http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi>. The engine software was Mascot 2.1 based on the MOWSE algorithm (Matrix Science, London, UK) ^[6;7]. The following search parameters were applied: taxonomy, all entries; fixed modification, cysteine carboxymethylation; variable modification, methionine oxidation; enzyme, trypsin; peptide tolerance, ± 1.3 Da; MS/MS tolerance, ± 0.3 Da; maximum number of missed cleavages, 1.

A protein was positively identified with a significance threshold of 0.05, meaning that random hits (so called false positives) occurred with a frequency lower than 5 %. The 95 % significance level corresponded to a MOWSE score of 23. The ion score cut-off

was set at 23. This value was already over the identification threshold for some peptides, leading to loss of peptides by the identification process. However, this strong cut-off value was utilized to ensure protein identifications based on very reliable peptide hits.

2 Experimental

2.1 A classical 2D-HPLC-MS setup: SCX x IP-RP-HPLC

The first separation scheme utilized for the analysis of *C. glutamicum* consisted of a strong cation exchange separation hyphenated to an ion-pair reversed-phase separation at pH 2.1. A schematic representation of the setup is depicted in Fig. 83.

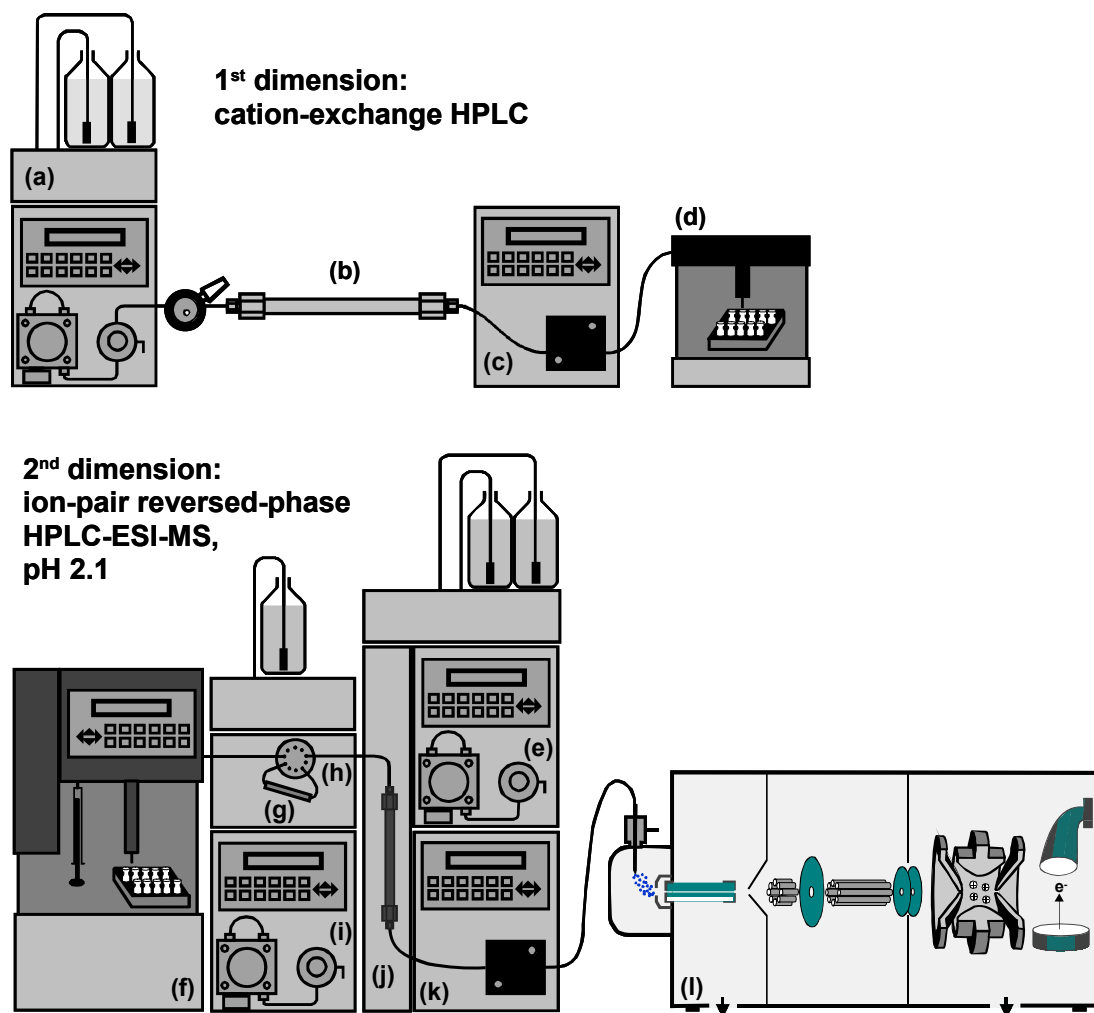


Fig. 83. Instrumental setup for off-line, two-dimensional peptide separations by SCX-HPLC x IP-RP-HPLC-ESI-MS ^[20].

(a) Pumping system for SCX separation; (b) strong cation exchange column; (c) UV detector for monitoring the first dimension; (d) fraction collector; (e) pumping system for IP-RP separation; (f) autosampler; (g) 10 x 0.20 mm i.d. monolithic trap column; (h) 10-port switching valve; (i) pump for loading and washing; (j) 60 x 0.10 mm i.d. monolithic separation column; (k) UV detector for monitoring the second dimension; (l) electrospray-ion trap mass spectrometer.

The detection was performed by electrospray ionization tandem mass spectrometry. The hyphenation of both methods of separation was achieved following an off-line approach: fractions were collected after the first separation step and injected after partial evaporation in the second chromatographic system. The use of a small trap column permitted the injection of large volumes of fractions without leading to peak broadening in the second dimension of separation. The trap column was also used to desalt the fractions before injection into the mass spectrometer. This consequently avoided salt contamination in the ion source of the mass spectrometer. Different cation exchangers were implemented in the setup and several peptide separations were performed in order to avoid data misinterpretation.

2.1.1 Proteome analysis of *C. glutamicum* with SCX x IP-RP-HPLC

Approximately 280 µg of a tryptic digest of *C. glutamicum* protein cell extract were injected over a 250 x 4.0 mm ProPac™ SCX-10 column. Sample loading was performed at 1 mL/min with (A) 5 mmol/L NaH₂PO₄, pH 3.0, 20 % ACN and elution with (B) 5 mmol/L NaH₂PO₄, pH 3.0, 20 % ACN, 500 mmol/L NaCl. The gradient was 0 - 3 % B in 9 min, 3 - 10 % B in 8 min, and 10 - 100 % B in 4 min. The very shallow salt gradient was combined with increasing fraction volumes to get higher homogeneity, in terms of peptide amount, between the collected fractions. Fractions were collected as follows: 250-µL fractions till 7.25 min, 500-µL fractions between 7.25 and 13.25 min, and 1,000-µL fractions until the end of the chromatographic separation. The fractions were evaporated to a final volume of 100 µL (2.5-, 5- and 10-fold concentration for the 250-, 500- and 1,000-µL fractions, respectively).

The second separation step was performed by loading and washing for 4 min 10 µL of the evaporated fractions on the 10 x 0.2 mm PS-DVB trap column with H₂O + 0.10 % heptafluorobutyric acid at 10 µL/min. The switching valve was then commuted and a back flush elution over the 60 x 0.1 mm PS-DVB analytical column was performed at 0.75 µL/min. Mobile phases consisted of (A) 0.05 % trifluoroacetic acid in water (pH 2.1) and (B) 0.05 % trifluoroacetic acid in acetonitrile. Peptides were separated at room temperature by using a gradient from 0 to 20 % B in 60 min, followed by isocratic conditions at 100 % B for 3 min. A total of 44 fractions collected between 1.25 min and 21.25 min were analyzed in triplicate, leading to 132 HPLC-MS/MS runs.

2.1.2 Fractionation repeatability after peptide separation with SCX

Approximately 90 µg of a tryptic digest of *C. glutamicum* protein cell extract were injected in triplicate over a 50 + 250 x 2.0 mm ProPac™ SCX-10 column. The flow rate was set to 250 µL/min. Mobile phases were (A) 5 mmol/L NaH₂PO₄, pH 3.0, 20 % ACN and (B) 5 mmol/L NaH₂PO₄, pH 3.0, 20 % ACN, 500 mmol/L NaCl. The gradient was 0 - 100 % B in 30 min, followed by isocratic conditions at 100 % B for 5 min. 250-µL fractions were collected per hand every minute and evaporated to 100 µL (2.5-fold concentration). Five consecutive fractions were selected and analyzed by injecting 10 µL of each in the IP-RP-HPLC-ESI-MS/MS system. Each fraction was analyzed in quintuplicate to avoid data misinterpretation due to potential bad repeatability of the second separation and/or detection step (e.g. ion suppression in the electrospray source).

Fractionation repeatability was also evaluated with a 10 + 200 x 2.1 mm PolySULFOETHYL column. The flow rate was set to 250 µL/min and 90 µg of *C. glutamicum* peptides were injected in triplicate. Sample loading was performed for 10 min with (A) 10 mmol/L ammonium formate, pH 3.0, 25 % acetonitrile. Elution occurred under a gradient of (B) 500 mmol/L ammonium formate, pH 6.8, 25 % acetonitrile: 0 – 50 % B in 40 min, 50 – 100 % B in 10 min, followed by isocratic conditions at 100 % B for 10 min. 250-µL fractions were collected per hand every minute, evaporated to 100 µL (2.5-fold concentration), and 5 consecutive fractions were selected and analyzed such as the fractions collected after the ProPac™ SCX-10 column.

2.2 A new 2D-HPLC-MS setup: RP-HPLC x IP-RP-HPLC

The second separation scheme utilized for the analysis of *C. glutamicum* consisted of a reversed-phase separation at pH 10.0 hyphenated to an ion-pair reversed-phase separation at pH 2.1. As in the classical SCX x IP-RP-HPLC approach, the hyphenation of both methods of separation was achieved off-line, and the detection was performed by electrospray ionization tandem mass spectrometry. Fractions were also collected after the first separation step and injected after partial evaporation in the second chromatographic system. Although sample desalting was not required because of the absence of salt in the eluents of the first dimension of separation (elution with acetonitrile), a small trap column was used to quickly inject large volumes of fractions in the second dimension of separation.

The feasibility of a two-dimensional separation of peptides with different pH in first and second separation dimensions was already proven on small set of peptides [142-144] but no literature was found for large scale separations of peptides. In both dimensions, hydrophobic interactions between stationary phase and peptides occur. However, the separation of peptides at high and low pH is achievable because of the amphiphile structure of peptides. Since peptides are charged molecules comprised of ionizable basic and acidic functional groups, the change of mobile phase pH has a pronounced effect on their retention behavior. At high pH, most of the basic sites are neutral and the acidic groups are negatively charged. On the contrary, under low pH conditions, the basic sites are positively charged and the acidic groups are neutral. These differences in terms of charges can be combined with different retention mechanisms (e.g. RP and IP-RP) and stationary phases (e.g. C18 and PS-DVB) to obtain separation systems with high peak capacity. A schematic representation of such a chromatographic system is depicted in Fig. 84. Silica-based stationary phases are usually only stable up to pH 8 in aqueous solutions. Consequently, polymer-based stationary phases are often used for the separation of peptides at high pH. However, silica-based stationary phases stable up to pH 12 have been developed in the last years [145] and are now commercially available (e.g. Waters, Phenomenex). Such a column was used at pH 10.0 to separate the peptides in the first chromatographic step.

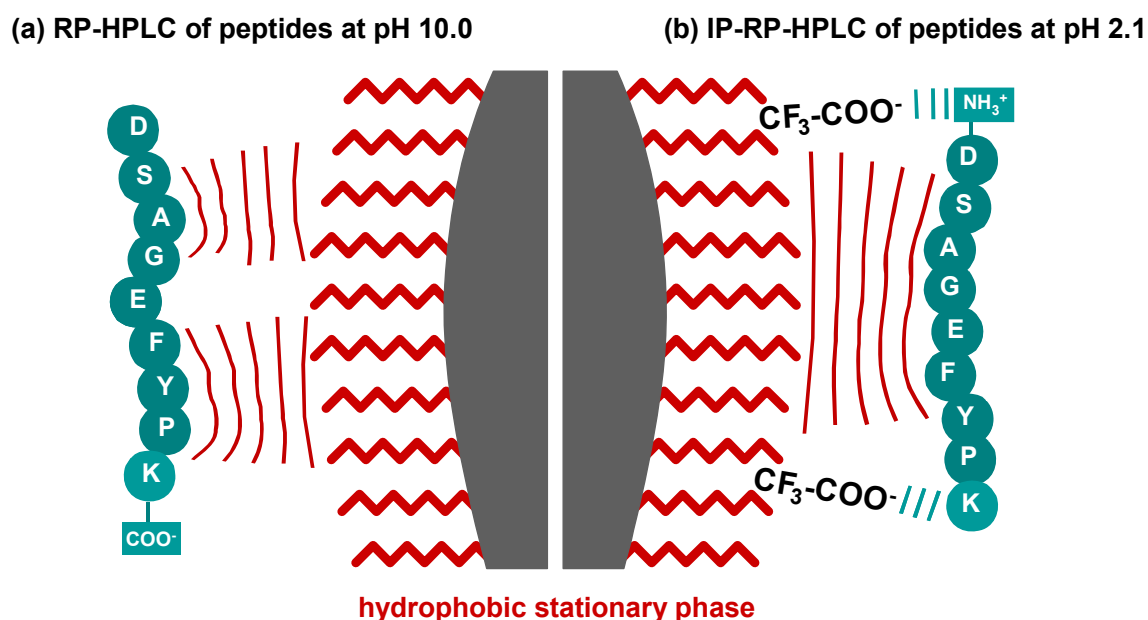


Fig. 84. Separation of peptides by (a) reversed-phase- under alkaline conditions and (b) ion-pair reversed-phase chromatography under acidic conditions.

2.2.1 Proteome analysis of *C. glutamicum* with RP x IP-RP

Approximately 280 µg peptides from *C. glutamicum* were injected over a 150 x 2.0 mm Gemini C18 column. Sample loading was performed at 200 µL/min with (A) 72 mmol/L triethylamine titrated to pH 10.0 with acetic acid. Elution was performed with (B) 72 mmol/L triethylamine, 52 mmol/L acetic acid in acetonitrile. Triethylamine is known to act as ion-pair reagent^[146;147]; however under these conditions reversed-phase mechanism predominates over ion-pair reversed-phase mechanism (see 3.2.2). The gradient was 0 - 55 % B in 55 min, followed by isocratic conditions at 100 % B for 2 min. 200-µL fractions were collected every minute. Acetonitrile was eliminated by evaporating the fractions to a final volume of 20 µL (10-fold concentration). Fractions were finally taken up with 105 µL of 0.10 % aqueous heptafluorobutyric acid.

Thirty-one fractions collected between 14 min and 45 min were analyzed in triplicate (total of 93 HPLC-MS runs). Columns, mobile phase solvents, flow rates and gradients were identical to the one used for the SCX fraction analysis (see 2.1.1). Because the fractions did not contain salts, the switching valve was commuted directly after sample transfer (2.5 min).

2.2.2 Fractionation repeatability after peptide separation with RP-HPLC at pH 10.0

Approximately 90 µg of *C. glutamicum* peptides were injected in triplicate over the 150 x 2.0 mm Gemini C18 column. Sample loading was performed at 200 µL/min with (A) 72 mmol/L triethylamine titrated to pH 10.0 with acetic acid. Elution was performed with (B) 72 mmol/L triethylamine, 52 mmol/L acetic acid in acetonitrile. The gradient was 0 - 60 % B in 35 min, 60 – 100 % B in 10 min, followed by isocratic conditions at 100 % B for 2 min. 400-µL fractions were collected per hand every two minutes and evaporated to 20 µL to eliminate acetonitrile (20-fold concentration). Five consecutive fractions were taken up with 95 µL 0.10 % aqueous heptafluorobutyric acid and analyzed in quintuplicate with the IP-RP-HPLC-ESI-MS/MS system previously described (see 2.1.1).

3 Results and discussion

3.1 Peptide separation

In the following section, peptide separations in the first chromatographic dimensions (SCX and RP-HPLC at pH 10.0) and in the second chromatographic dimension (IP-RP-HPLC at pH 2.1) are investigated.

3.1.1 Peptide separation with SCX

The chromatogram of the peptide separation performed with SCX in order to analyze the proteome of *C. glutamicum* is depicted in Fig. 85.

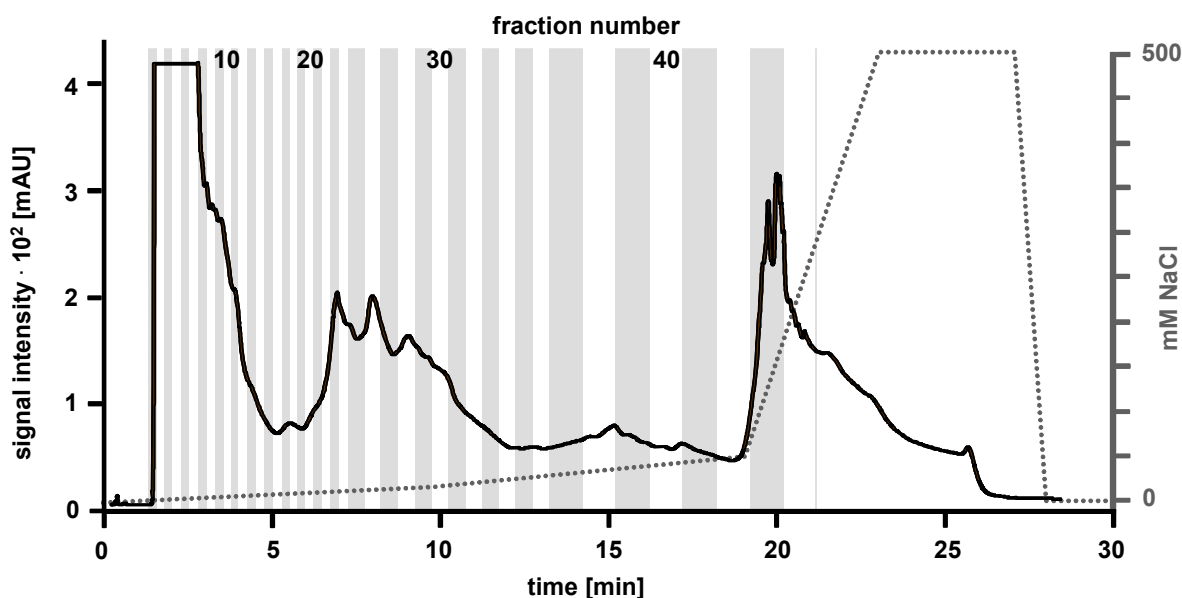


Fig. 85. Strong cation exchange fractionation of a digested *C. glutamicum* lysate (first chromatographic dimension).

Column, 100 + 250 x 4 mm i.d. ProPac™ SCX-10; mobile phase, (A) 5 mM NaH₂PO₄, pH 3.0, 20 % acetonitrile, (B) 0.50 mol/L NaCl in eluent (A); gradient, 0-3 % B in 9.0 min, 3-10 % B in 8 min, 10-100 % B in 4 min; flow rate, 1 mL/min; temperature, 25°C; detection, UV at 214 nm; sample, 280 µg tryptic digest of a *C. glutamicum* protein cell extract; fractions, 24 x 0.25 min, 12 x 0.50 min, 8 x 1.00 min.

As observed, the strong cation exchanger does not show a high peak capacity and most of the peptides are eluting within a narrow gradient strength window (0 - 50 mmol/L NaCl). This is partially explained by the fact that separations in SCX are generally based on the charge of the analytes. Practically, peptides are not individually separated but sets of peptides with similar charges (1+, 2+, 3+, and 4+) are separated. Although the salt gradient was very shallow, most of the peptides

were eluting in two sets (fractions 1-12, and fractions 19-34). To get a higher homogeneity in terms of peptide amount between the collected fractions, fractions at the very beginning of the chromatographic process were collected every 0.25 min, whereas fractions with less peptides were collected only every 1.00 min.

3.1.2 Peptide separation with RP-HPLC at pH 10.0

The same amount of peptides was injected into the RP-HPLC system at pH 10.0. The obtained chromatogram is depicted in Fig. 86. Peaks are observed over a broad separation window (over 45 min), highlighting a good peptide separation with a linear gradient of acetonitrile. Previous experiments showed an absence of peptides during the first 15 min of the separation and consequently fractions were not collected at the beginning of the separation. Peaks observed in this time window may be partly attributed to chemicals used during the extraction and/or the digestion of the proteins. Because peptides appeared to be continuously eluted from the column, fractions were collected with the same periodicity (1.00 min fractions).

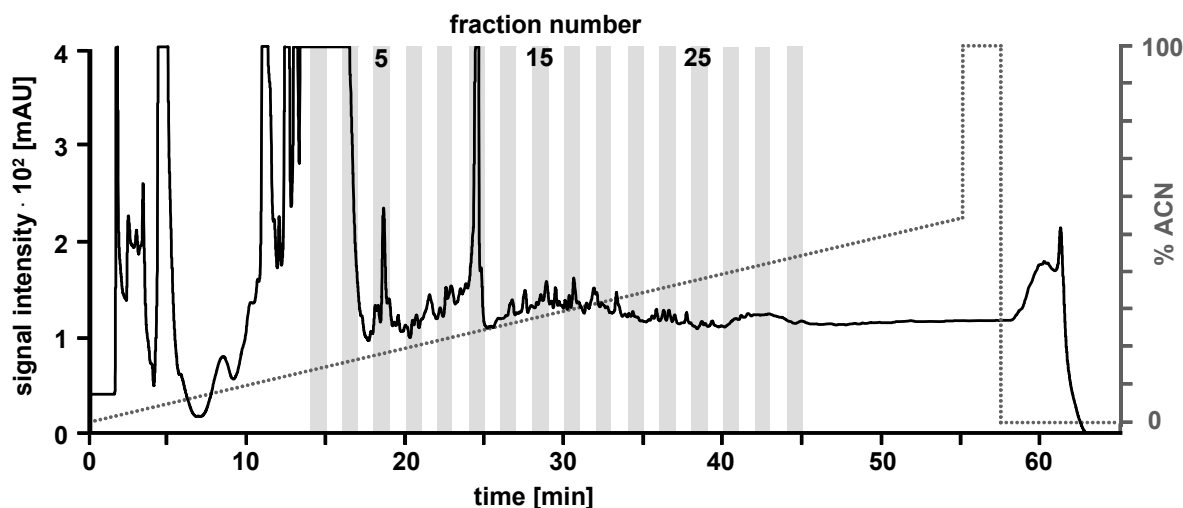


Fig. 86. Reversed-phase fractionation at pH 10.0 of a digested *C. glutamicum* lysate (first chromatographic dimension).

Column, 150 x 2.0 mm i.d. 3 μ m Gemini C18; mobile phase, (A) 72 mmol/L triethylamine titrated to pH 10.0 with acetic acid, (B) 72 mmol/L triethylamine and 52 mmol/L acetic acid in acetonitrile; gradient, 0-55 % B in 55.0 min; flow rate, 200 μ L/min; temperature, 25°C; detection, UV at 280 nm; sample, 280 μ g tryptic digest of a *C. glutamicum* protein cell extract; fractions, 31 x 1.00 min.

3.1.3 Peptide separation with IP-RP-HPLC at pH 2.1

All fractions collected after separation with strong cation-exchange- and reversed-phase HPLC at pH 10.0 were analyzed with IP-RP-HPLC at pH 2.1 and detected with ESI tandem mass spectrometry. Typical reconstructed total ion current chromatograms are depicted for three consecutive fractions in Fig. 87. A mass spectrum (see Fig. 88a) corresponds to each point of such chromatograms. For each mass spectrum, most intensive ions are selected and submitted to fragmentation (collision-induced dissociations). Fragments are obtained (see Fig. 88b) and finally compared to a database for peptide identification. The principle of peptide identification with Peptide Fragment Fingerprinting (PFF) is outlined in chapter II, section 3.13.

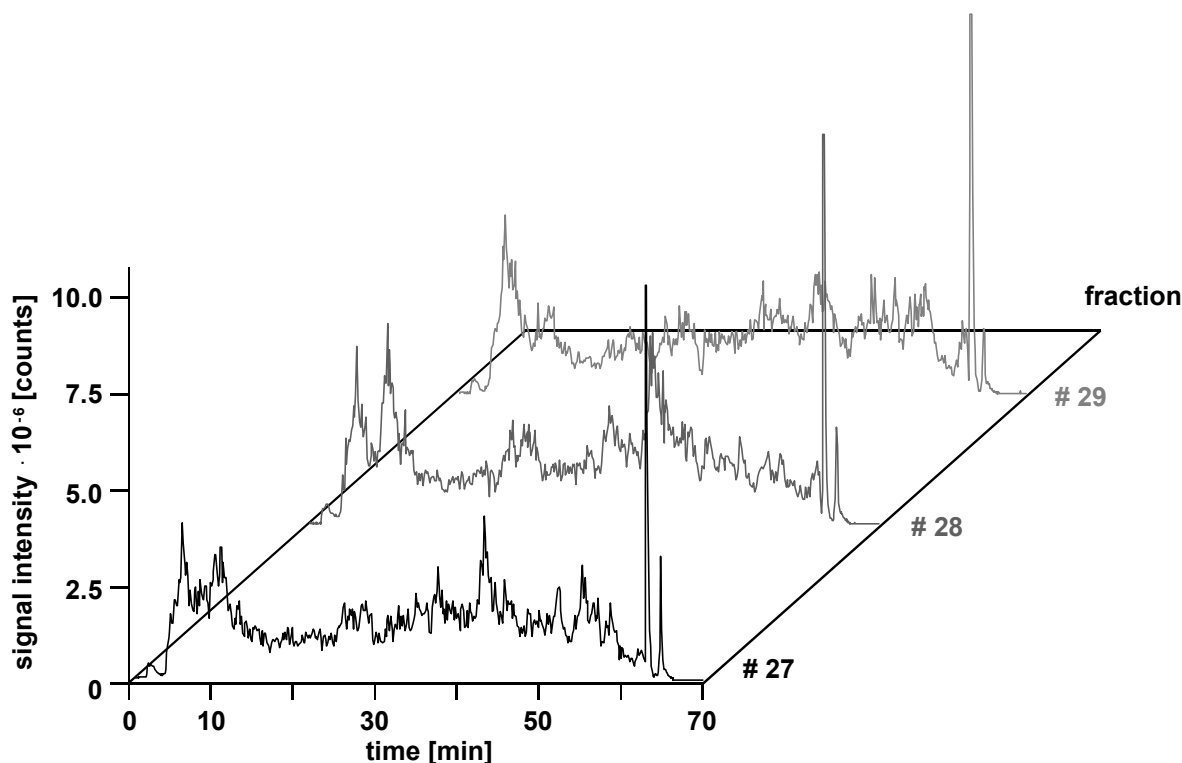


Fig. 87. Ion-pair reversed-phase separation at pH 2.1 of three consecutive cation-exchange fractions of *C. glutamicum*.

Columns, 10 mm \times 0.20 mm i.d. monolithic PS-DVB preconcentration column, and 60 mm \times 0.20 mm i.d. monolithic PS-DVB separation column; loading solvent, 0.10 % aqueous heptafluorobutyric acid; trapping time, 4.0 min; loading flow rate, 10 μ L/min; mobile phase: (A) 0.050 % aqueous trifluoroacetic acid, (B) 0.050 % trifluoroacetic acid in acetonitrile; gradient, 0–20 % B in 60 min; flow rate, 0.750 μ L/min; temperature of preconcentration- and separation column, 25 $^{\circ}$ C; detection, ESI-MS/MS; sample, 10 μ L tryptic peptides of *C. glutamicum*, fraction 27, 28, and 29 from strong cation-exchange chromatography, 10 μ L injected.

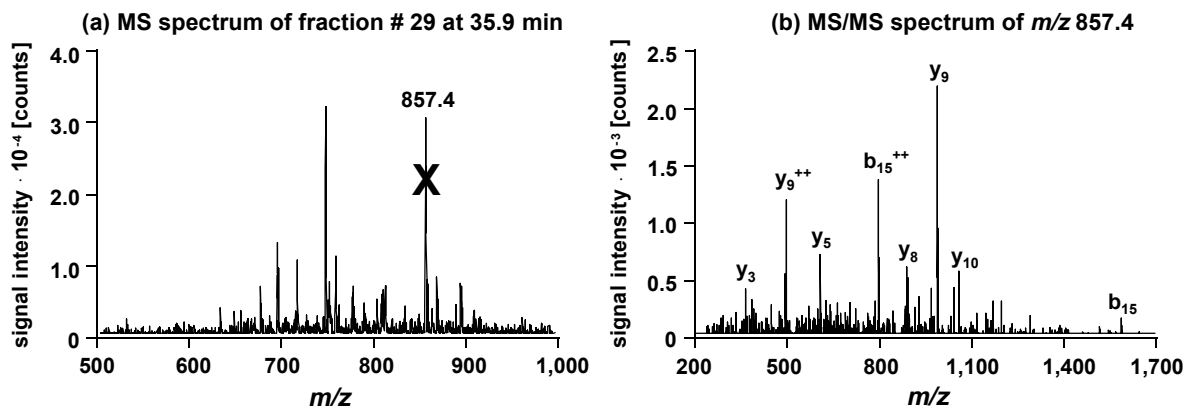


Fig. 88. (a) MS spectrum of cation exchange fraction # 29 at 35.9 min and (b) MS/MS spectrum of m/z 857.4. Chromatographic conditions as in Fig. 87 and MS parameters as described in section 1421.4.

3.2 Peptide distribution

In the following section, the peptide distributions obtained after separation with SCX and RP-HPLC at pH 10.0 are discussed.

3.2.1 Peptide distribution with SCX

The 44 fractions, collected after the strong cation exchanger, were injected in triplicate into the second separation dimension. For each HPLC-MS/MS run, a database search was performed, leading to 132 different identification reports. For the three replicates of each fraction, the three identification reports were compared and a nonredundant list of identified peptides was established. Then the number of unique identified peptides as a function of the fraction number was plotted. The obtained distribution is depicted in Fig. 89a. Peptides are identified all over the 44 fractions. As expected, the peptide hits are not continuously distributed and most of the peptides are identified in two distinguished time windows (bimodal distribution). The bimodal distribution shows maxima in fractions 1 and 25, and minima in fractions 16 and 44. The same data interpretation was performed at the protein level and a similar profile was obtained (see Fig. 89b). Maxima were observed for fractions 7 and 27 and minima for fractions 18 and 44.

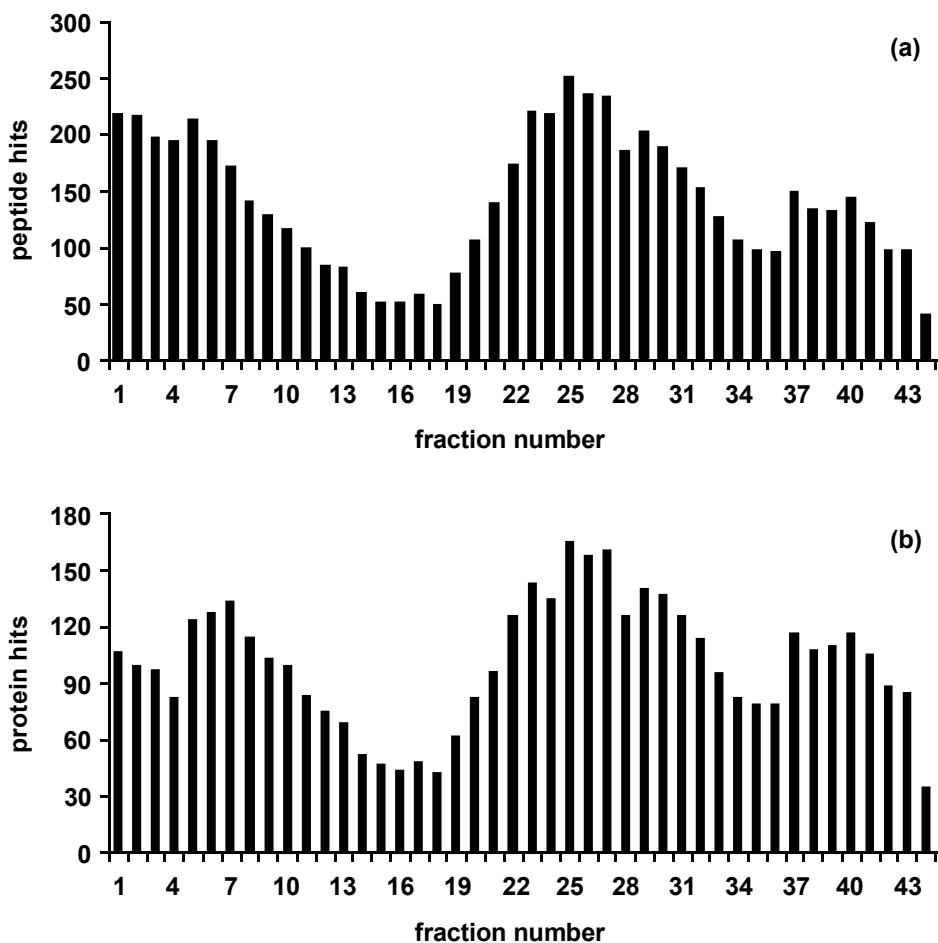


Fig. 89. (a) Peptide hit and (b) protein hit distribution over SCX fractions. Data correspond to three replicate injections in the second chromatographic dimension. Redundant hits within replicates have been eliminated but redundant hits between fractions are present.

As already discussed, the bimodal distribution is attributed to the charge states of the tryptic peptides ^[20;148]. Doubly charged peptides are eluting at the very beginning of the separation, whereas peptides carrying three or more charges are retained on the column and are eluted at higher salt concentrations. In order to verify this hypothesis, for each identified peptide of each fraction (Σ 6,254), the number of positively charged residues was computed. Because the separation was performed at pH 3.0, positive residues were the N-terminus of the peptide (pK_a 9-11) and the three basic amino acids: arginine (pK_a 12.5), lysine (pK_a 10.5), and histidine (pK_a 6.0). At pH 3.0, no residue is negatively charged. Consequently, retention of peptides on the cation exchanger is mostly determined by the number of positive charges in the amino-acid sequence and no correlation should be obtained in a plot representing pI as a function of the fraction number. As observed in Fig. 90, a slight tendency of

increasing pI values with increasing fraction numbers appears. However, no real correlation is observed.

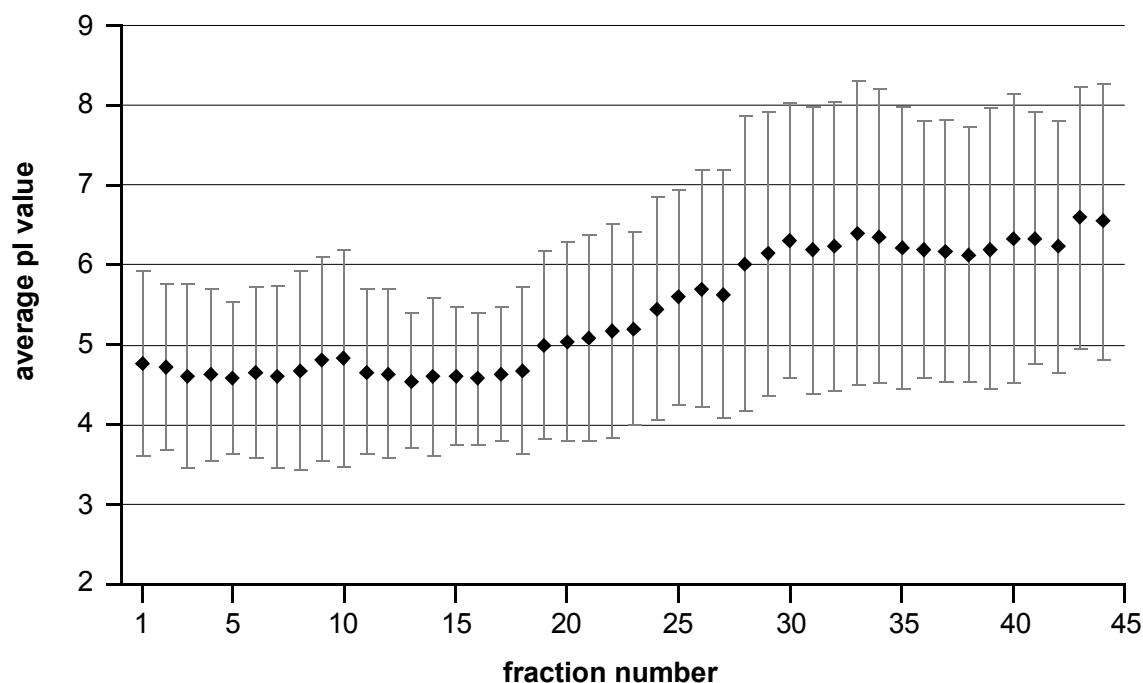


Fig. 90. Plot representing the average pI value of the identified peptide as a function of the fraction number. The pI values were computed with the compute pI/Mw tool on-line available under www.expasy.org, and arithmetically averaged.

On the other hand, if we assume that the positive charges carried by the peptides are mainly responsible for the peptide distribution, low charged peptides should elute at the beginning of the separation, whereas peptides carrying more charges should elute latter. In Fig. 91, elution profiles for peptide charged 1+ to 6+ are depicted. The elution of doubly charged peptides at the very beginning of the separation process is confirmed. The elution of the triply charged peptides occurs later (from fraction # 19). The elution of some three-time charged peptides in the first two fractions may be explained by a column overloading. Very few peptides with only one single positive charge were detected. They correspond to C-terminus fragments of proteins which do not contain any basic residues.

Acetonitrile was utilized in the eluents during the SCX separation in order to diminish hydrophobic interactions (secondary interactions) between peptides and the polymeric stationary phase of the cation exchanger. In order to check the minor role of such interactions, the hydropathy index (GRAVY index) of each identified peptide was computed using parameters published elsewhere ^[24]. Then, arithmetic averages

of GRAVY values were computed for each fraction and plotted in Fig. 92. Despite a slight tendency of decrease, no correlation was observed. This reveals the negligible role of hydrophobic interactions occurring in the separation process.

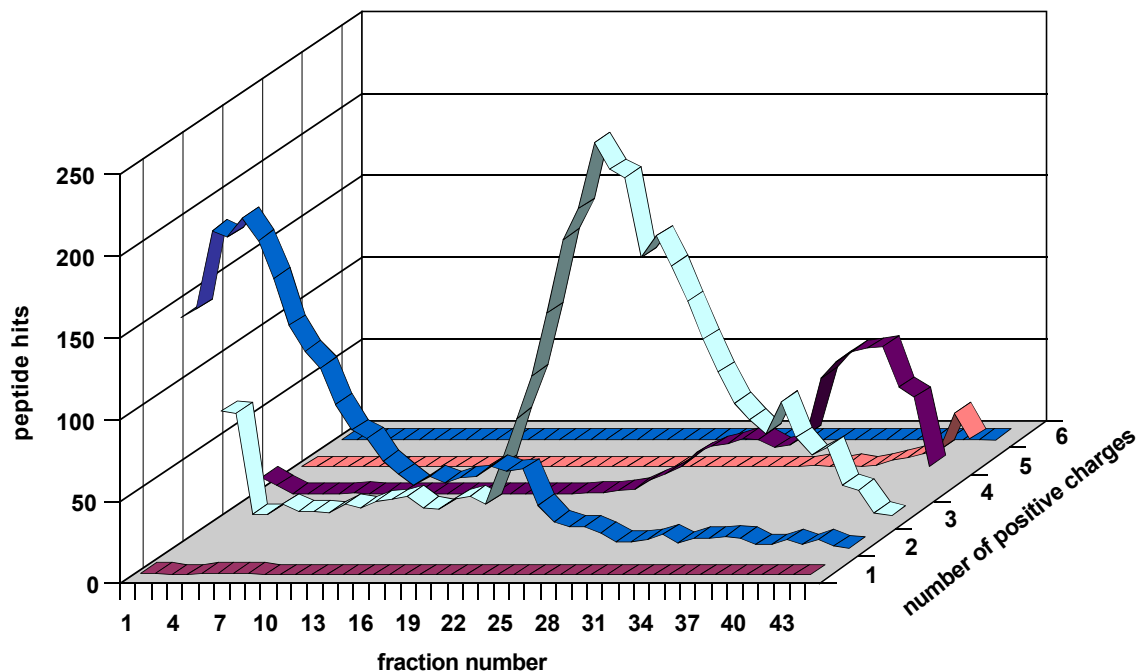


Fig. 91. Dependence between the number of positive charges carried by a peptide and the retention time of the peptide in a strong cation-exchange separation.

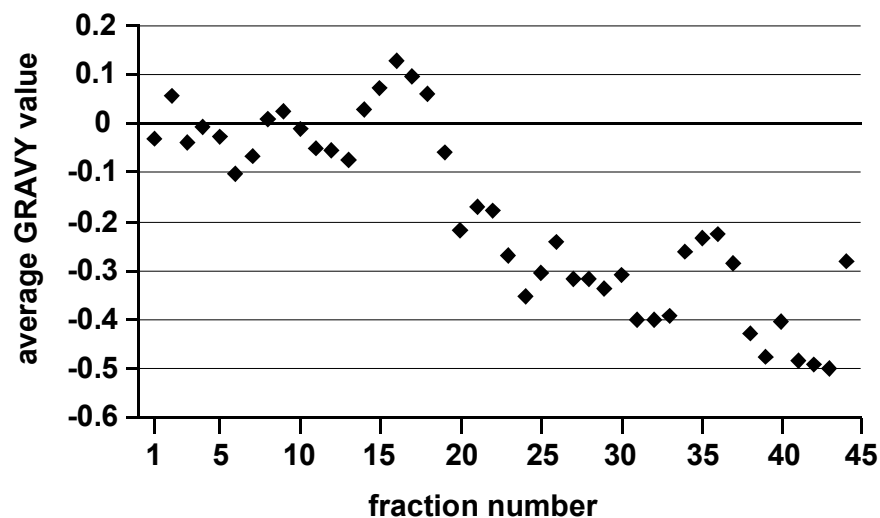


Fig. 92. Plot representing the average GRAVY value of the identified peptides as function of the fraction number.

It results from these different computations that the most important parameter, influencing the retention of the peptides in the cation exchanger, is the number of positive charges carried by the peptides. Negative charges and hydrophobic properties play a minor role.

3.2.2 Peptide distribution with RP-HPLC at pH 10.0

The 31 fractions collected after the reversed-phase separation at pH 10.0 were injected in triplicate into the second separation dimension. As for the SCX fractions, for each HPLC-MS/MS run a database search was performed. 93 identification reports were obtained. Redundant hits within replicates were filtered out and for each fraction a list of non redundant identified peptides was established. The peptide hit distribution is plotted in Fig. 93a.

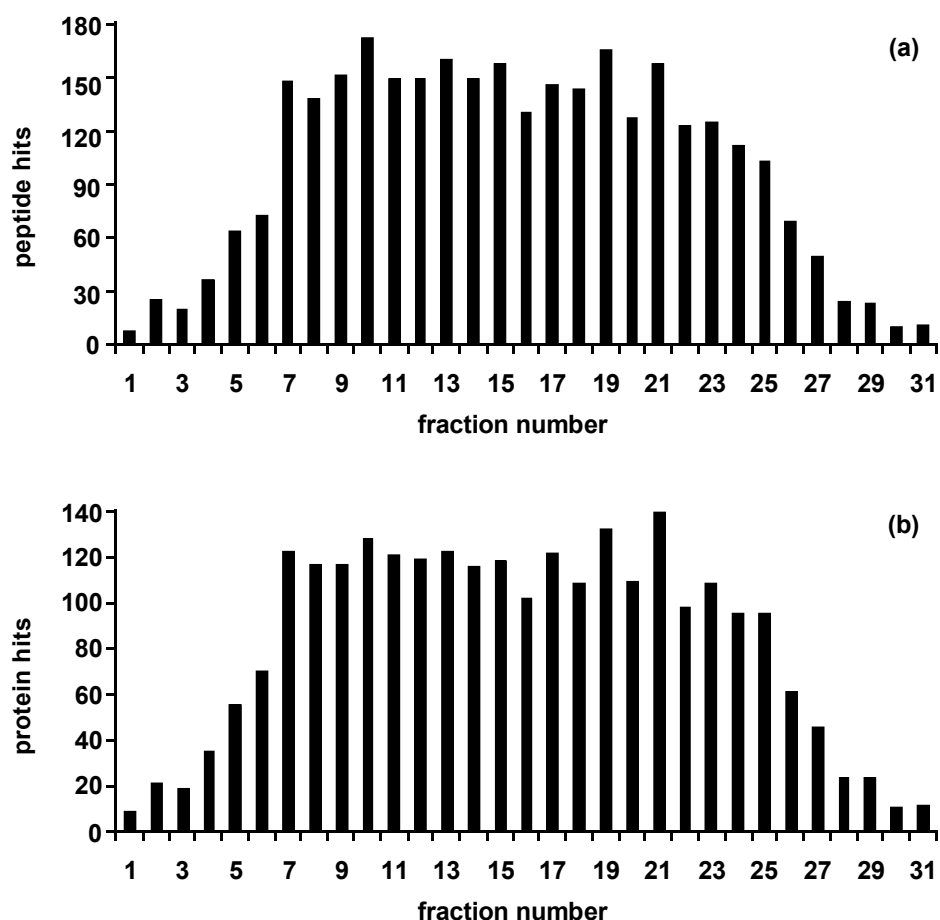


Fig. 93. (a) Peptide hit and (b) protein hit distribution over RP fractions. Data correspond to three replicate injections in the second chromatographic dimension. Redundant hits within replicates have been eliminated but redundant hits between fractions are occurring.

Peptides are identified all over the 31 fractions. The peptide hit distribution is quite similar to a Gauss curve with a truncated maximum. After an increase over 6 fractions, the number of peptide hits is stable over 18 fractions, and finally decreases

over 7 fractions. The same data interpretation was performed at the protein level and a very similar profile was obtained (see Fig. 93b).

In order to check the influence of the charge state of the peptides on retention, the number of negative charges was computed for each peptide. Because the first dimension of separation was performed at pH 10.0, peptides were carrying negative charges at C-terminus (pK_a 1.8-2.4), and at aspartic acid (pK_a 3.9) and glutamic acid (pK_a 4.3) residues. For each charge state (1- to 7-), the average retention time of the peptides (expressed as an average fraction of elution) in the first dimension was computed. It was then possible to plot a diagram representing the mean elution time as a function of the charge state of the peptides. As depicted in Fig. 94, a very good correlation was observed.

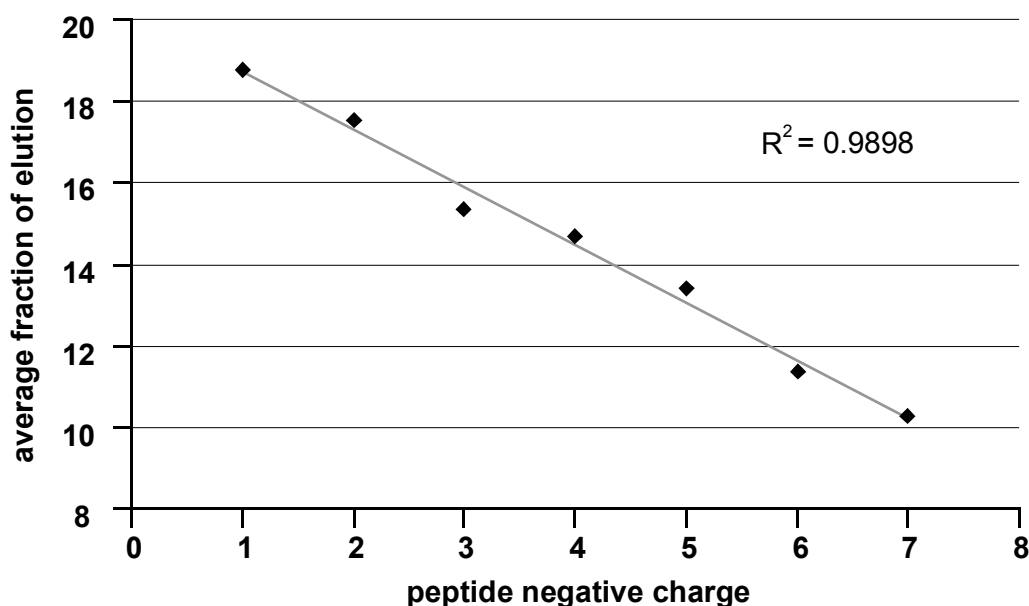


Fig. 94. Plot representing the average fraction of elution for negative charged peptides.

The more charged the peptides, the shorter the elution. This leads to the assumption that the reversed-phase mechanism predominated over the ion-pair reversed-phase mechanism during the separation. Under such conditions, peptides carrying many charges are presenting much more affinity to the mobile phase than low-charged peptides. Consequently, peptides carrying many charges are eluting first and peptides with few charges are eluting later.

In order to check this hypothesis, the GRAVY index of each identified peptide was computed. Then, for each fraction, the arithmetic average of the GRAVY values of all the detected peptides was calculated. Finally, these average GRAVY values were plotted as a function of the fraction number (see Fig. 95).

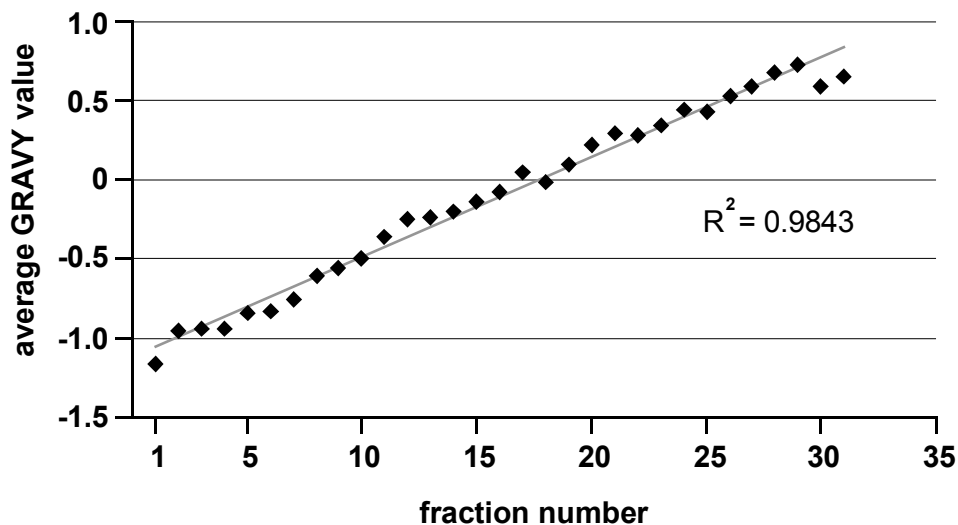


Fig. 95. Plot representing the average GRAVY value of the identified peptides as a function of the fraction of elution.

A linear correlation was observed ($R^2 = 0.9843$). The plot reveals that less hydrophobic peptides were eluted at low acetonitrile content, whereas more hydrophobic peptides were eluted at higher concentration of acetonitrile. This signifies that the retention of the peptides was strongly influenced by solvophobic interactions and it confirms the predominant role of the reversed-phase mechanism. Triethylamine was not concentrated enough to deal as an ion-pair reagent.

3.3 Proteome coverage

The proteome analysis of *C. glutamicum* with the SCX x IP-RP-HPLC setup (132 HPLC-ESI-MS runs) led to 6,254 peptide hits. After removal of redundant peptide hits between the fractions, a total of 2,398 unique peptides were identified. With the RP x IP-RP-HPLC setup (93 HPLC-ESI-MS runs), 3,124 peptide hits were obtained and after removal of redundancies, 2,708 unique peptides were identified. Thus, in comparison with the SCX approach, ca. 13 % more unique peptides were identified (2,708 vs. 2,398) with ca. 50 % less peptide hits (3,124 vs. 6,254) in the RP x IP-RP-HPLC approach. The significant decrease in redundancies with the RP x IP-RP-HPLC approach (1.15 hits per unique peptide) in comparison with the SCX x IP-RP-

HPLC approach (2.61 hits per unique peptide) is partially explained by the analysis of less (31 vs. 44) but bigger (1.00-min vs. 0.25-, 0.50-, and 1.00-min) fractions in the RP x IP-RP approach. However, the weak retention of low charged peptides (1+ and 2+) on the cation exchanger requires such a fractionation to avoid the collection of too complex fractions, which can be hardly analyzed because of strong ion suppression in ESI. The flowing-through of some peptides during the SCX separation is illustrated by the fact that more than 45 % of the unique identified peptides are already detected in the first 5 min of the separation. To sum up, the new RP x IP-RP-HPLC approach permits to significantly increase the number of identified peptides in a significantly shorter time of analysis (124 vs. 176 hours).

At the protein level, the SCX x IP-RP-HPLC setup led to 4,544 protein hits. After removal of redundant protein hits between the fractions, a total of 695 unique proteins were identified. The RP x IP-RP-HPLC approach led to the identification of 745 unique proteins, which corresponds to 7 % more identifications than with the classical SCX x IP-RP method.

The proteome coverage can be defined as the ratio between the number of detected proteins and the number of potentially expressed proteins. The number of potentially expressed proteins corresponds to the number of proteins annotated for a given genome, but not necessarily present in the protein cell lysate. In the case of the present study on *C. glutamicum*, the number of potentially expressed proteins corresponds to the number of entries in the database (2,993). Both analytical methods gave relevant proteome coverage: 23.2 % with the classical SCX x IP-RP-HPLC setup and 24.9 % with the RP X IP-RP-HPLC setup. Considering that many proteins were not expressed in the cells and that proteins were lost during sample preparation (extraction, dialysis), these values are very promising for further biologically relevant applications.

3.4 Identification confidence

No set of identified peptides in a Peptide Fragment Fingerprinting approach is free of false-positive hits. The presence of false-positive identifications is intrinsic to the method of identification and mostly results from the algorithms utilized to identify the peptides. Algorithms compare experimental MS/MS spectra with theoretical fragmentation patterns and return the best peptide match for each MS/MS spectrum within the database. Because best matches are not necessarily perfect matches, a

portion of the identifications are incorrect peptide sequence assignments due to coincidental similarities in MS/MS fragmentation patterns. In the present study, the identification confidence threshold was set to 95 % by applying a MOWSE score cutoff of 23. However, the algorithm may always return random hits and the validity of the identifications has to be *a posteriori* checked. To address this issue, all the experimental mass lists were tested against nonsense protein sequences obtained by reversing the protein sequences of *C. glutamicum* (reverse sequence database strategy ^[95]).

The peak lists generated in the SCX x IP-RP-HPLC approach returned 2,398 peptide hits and 695 protein hits after a search in the forward database, whereas 38 peptide hits and 38 protein hits were randomly obtained from a search in the reverse database. At the peptide level, a false-positive identification rate of 1.6 % was computed, while a false-positive identification rate of 5.5 % was obtained at the protein level. With the RP x IP-RP-HPLC strategy, very similar results were obtained: 40 peptides and 40 proteins were randomly identified in the reverse database. The corresponding false-positive identification rates were 1.5 % for peptides and 5.4 % for proteins. These values are in agreement with the 95 % identification confidence threshold set during the MASCOT searches and reveals that no systematic error occurred during the identification process. The imposing difference between false-positive identification rates at peptide and protein levels are explained by the fact that more peptides may identify a single protein. Such peptides are very usual for true identifications in a forward database, but false protein identification in a reverse database is rarely achieved with more than one peptide.

Nowadays a generally accepted rule only validates protein identifications which are based on the detection of more than one peptide of the protein sequence ^[149;150]. In the SCX x IP-RP-HPLC approach, 414 proteins fulfilled this criterion. With the RP x IP-RP-HPLC setup, 468 proteins were identified with more than one peptide, meaning that the new separation approach permitted to identify 13 % more proteins than the classical approach under such a criterion of data validation.

Another criterion which influences the identification of peptides and proteins is the peptide tolerance used to perform the database search. For all the identifications performed in this work, a peptide tolerance of ± 1.3 Da was applied. This value is rather high but the aim of the present work was not to identify proteins at the highest confidence level but to develop and evaluate multidimensional HPLC-ESI-MS/MS

methods. Moreover, if some protein misidentifications occur, they should statistically happen at comparable levels in both HPLC-ESI-MS/MS approaches. This was checked by reducing the peptide tolerance from ± 1.3 Da down to ± 0.3 Da and by maintaining the identification confidence threshold at 95 % with a MOWSE score cutoff of 17. In this case, 679 proteins were identified with the SCX x IP-RP-HPLC setup and 718 proteins with the RP x IP-RP-HPLC system. The protein identification decreases at similar levels: 2.3 % and 3.6 % respectively.

3.5 Repeatability of the dimensions of separation

3.5.1 Repeatability of peptide separation and identification with IP-RP-HPLC-ESI-MS/MS at pH 2.1

In order to check the repeatability of the second dimension of both HPLC-ESI-MS/MS setups, samples consisting of tryptic peptides of *C. glutamicum* were injected in quintuplicate into the IP-RP-HPLC-ESI-MS/MS system. For each replicate analysis an identification search was performed with Mascot. For each sample, the five identification reports were compared with each other at the peptide level. Some peptides were identified five times, whereas others were only once detected. Finally, frequencies of peptide identification were computed for each sample. In order to get statistically meaningful data, such analyses were pursued for 15 different samples and arithmetic averages were calculated. The data is schematically represented in a pie-chart in Fig. 96. A quite good repeatability of the HPLC-ESI-MS/MS analyses is observed: 30.6 % of the peptides were detected in all five replicate analyses. However, 28.7 % of the peptides were detected only once. Still too complex samples leading to peptide co-elution and ion suppression in the ESI source may partially explain these results. The peptide distribution in two major groups (54.0 % of the peptides are identified three or more times, whereas 28.7 % of the peptides are detected only once) is also due to the dynamic range in concentration of the peptides. Proteins, and as a matter of fact peptides, are not concentrated at the same level in the sample. Consequently some peptides are highly abundant and present at concentrations far above the limit of detection of the mass spectrometer, whereas the concentrations of low abundant peptides are ranging around the limit of detection of the mass spectrometer. As a result, high abundant peptides are quasi

always detected, whereas peptides at concentrations ranging around the limit of detection of the mass spectrometer are very rarely identified.

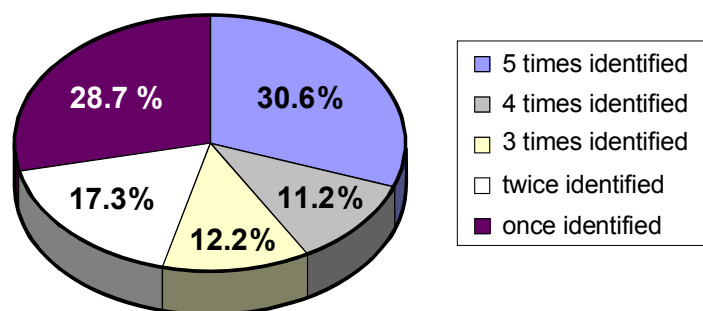


Fig. 96. Pie-chart representing the repeatability of peptide identification with IP-RP-HPLC-ESI-MS/MS of five replicate analyses of *C. glutamicum* tryptic digests. The values are arithmetic averages computed for 15 different samples.

The computed data also prove that repetitive analyses help to significantly increase the number of identified peptides. In order to estimate a meaningful number of replicates one should perform, the cumulative amount of identified peptides as a function of the replicate number was computed for quintuplicate analyses. The results are depicted in Fig. 97.

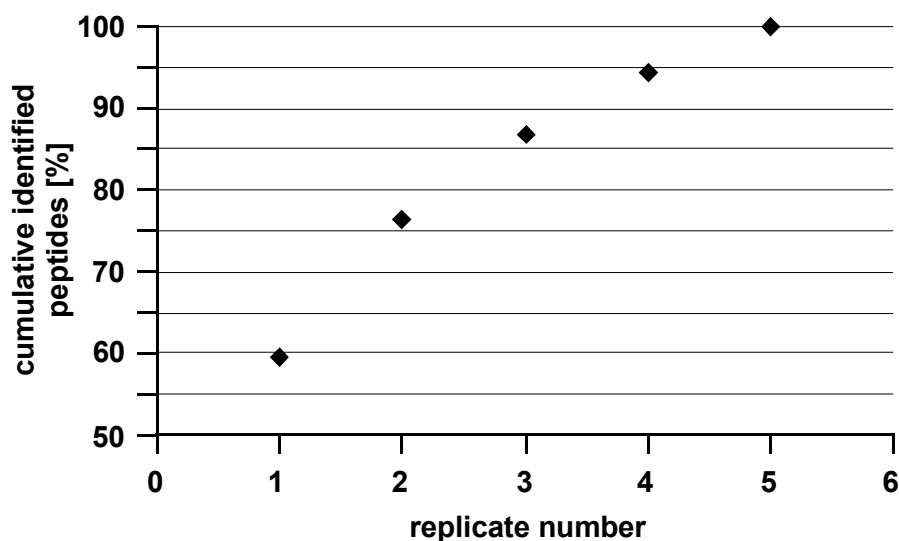


Fig. 97. Percent of identified peptides as function of the number of replicates for quintuplicate IP-RP-HPLC-ESI-MS/MS analyses at pH 2.1.

One observes that the first injection permits to identify more than 55 % of the peptides. With three replicates, more than 85 % of the peptides are identified. The fourth and the fifth injections still yield an increase in terms of peptide identification but

the gain is inferior to 10 % per injection. No absolute rules should be set and the operator should determine for each sample and analytical aim how much time can be invested to perform replicate injections. However, the analysis of a sample in triplicate appears to be a good compromise in terms of peptide identification and time consumption. This value was the one chosen for the identification study (see 3.3 Proteome coverage), whereas five replicate measurements were performed in order to determine the repeatability of the first dimension of both two-dimensional HPLC setups (see next sections).

3.5.2 Separation and fractionation repeatability with SCX

The repeatability of the first dimension of the classical SCX x IP-RP-HPLC setup was investigated. Three injections of a same tryptic digest of *C. glutamicum* proteins were performed on a cation exchanger and fractions were collected in triplicate. Each collected fraction was injected in quintuplicate into the second separation dimension. The MS/MS data of the five replicates were merged together and for each fraction collected in the first separation step, one single peptide identification analysis was performed. These multiple injections in the IP-RP-HPLC-ESI-MS/MS system permitted to diminish data misinterpretation due to insufficient repeatability of the second dimension.

Consecutive fractions containing between 90 and 200 peptides were collected after the cation exchanger and analyzed with IP-RP-HPLC-ESI-MS/MS. Because the SCX separation was performed in triplicate, it was possible to compute for each fraction the number of peptides identified in three, in two, or in one of the replicates. An arithmetic average of the identification redundancies was computed for five fractions. In order to check that the results were not dependent on the cation exchanger, the same experiment was performed on two different cation exchangers (ProPacTM SCX, and PolySULFOETHYL). The results are depicted in Fig. 98 and in Fig. 99. With both columns very similar results were obtained. Most of the peptides (41 – 45 %) were identified in all three replicates, showing a quite good repeatability of the separation and fractionation setup. The amount of peptides only once detected was rather high (37 – 39 %) and peptides twice detected were the less numerous (22 – 16 %).

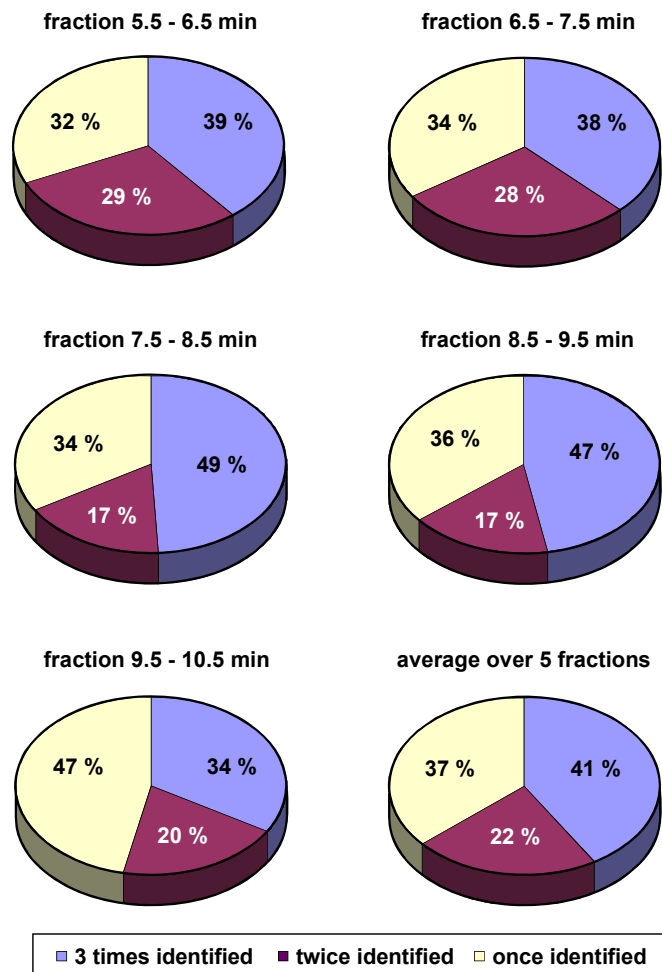


Fig. 98. Repeatability of the separation and fractionation step with the ProPac™ SCX-10 column. Each pie-chart corresponds to one fraction collected in triplicate after the cation exchanger and analyzed in quintuplicate with IP-RP-HPLC-ESI-MS/MS. The percentages correspond to the amount of peptides once, twice or three times identified.

Because quintuplicate analyses should have permitted to identify most of the peptides present in a fraction, the identification distribution is mostly attributed to the first dimension of the setup. Factors leading to a limited repeatability are various. First, peptides may not be eluted from the cation exchanger in a repeatable manner. Such differences are explained by numerous factors: e.g. limited repeatability of gradient formation, small variations in column temperature, or equilibration of the column. Errors may also occur during the collection of the fractions. This is of particular importance for small fractions and for fractionations performed per hand. Finally, fractions are partially evaporated in polypropylene vials before injection in the second separation dimension and adsorption phenomena can take place ^[151].

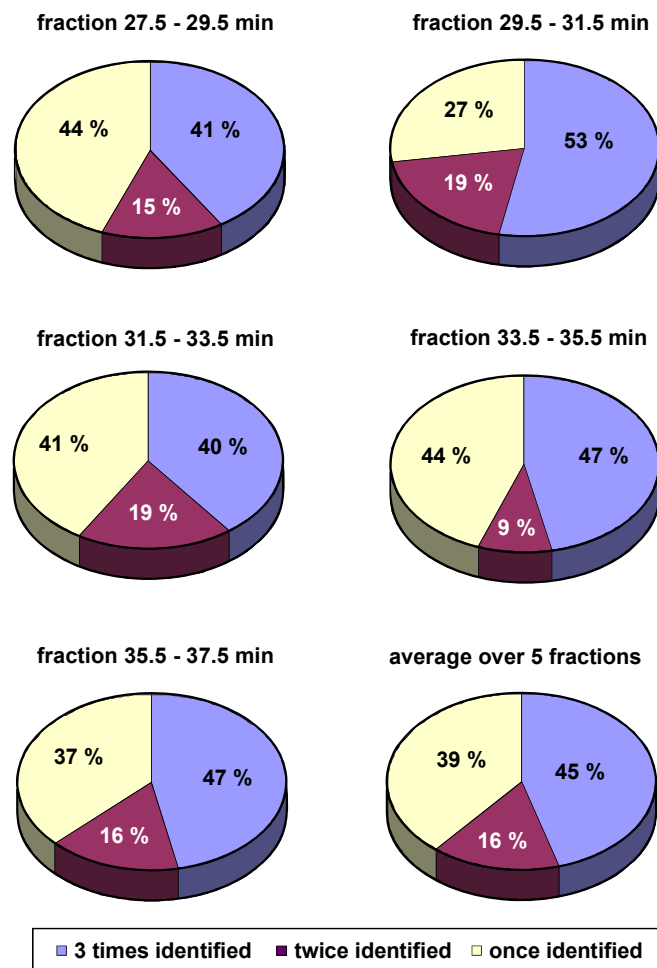


Fig. 99. Repeatability of the separation and fractionation step with the PolySULFOETHYL column. Each pie-chart corresponds to one fraction collected in triplicate after the cation exchanger and analyzed in quintuplicate with IP-RP-HPLC-ESI-MS/MS. The percentages correspond to the amount of peptides once, twice or three times identified.

It is not possible to accurately determine the importance of each source of imprecision but the influence of small differences in terms of retention time and fraction collection was evaluated. Delay in the elution results in the collection of peptides in later fractions. The five fractions collected after the cation-exchanger were consecutive fractions. Consequently, by merging the identification results of the five fractions, one gets a set of identified peptides independent of variations in retention/collection time. For each replicate injection in the first dimension of separation, such a data examination was performed. Finally, the three peptide sets were compared and the amounts of peptides once, twice, or three times identified were computed. A schematic representation of the computing process permitting to

determine the influence of the separation and fractionation step is depicted in Fig. 100. The results are depicted for the ProPac™ SCX-10 column in Fig. 101a, and for the PolySULFOETHYL column in Fig. 101b.

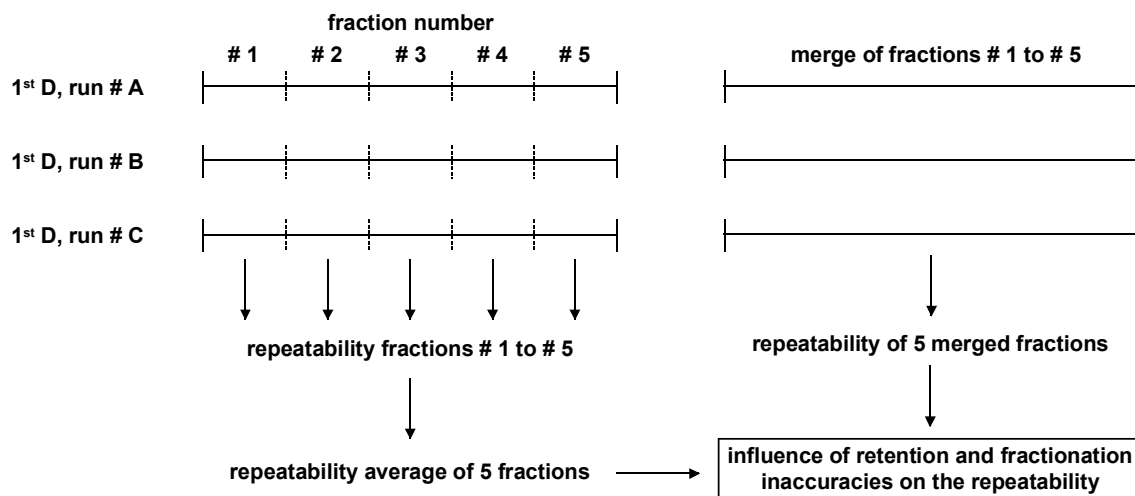


Fig. 100. Determination of the influence of retention and fractionation inaccuracies on the repeatability of peptide identification.

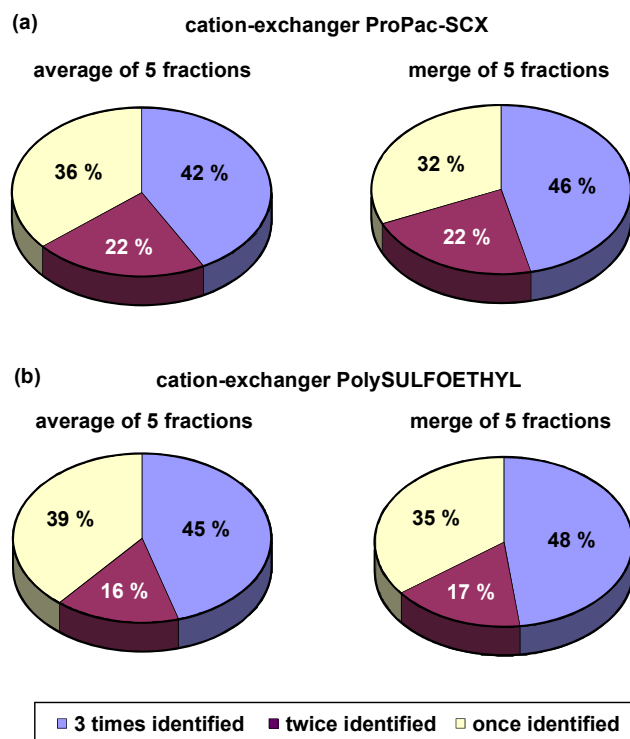


Fig. 101. Repeatability of the separation and fractionation step with (a) the ProPac™ SCX-10 and (b) the PolySULFOETHYL column. Pie-charts on the left side represent the average repeatability observed for five fractions. Pie-charts on the right side correspond to the repeatability obtained after a merge of the same five fractions.

For both columns, higher repeatability with merged fractions was observed as compared with averages of isolated fractions. In case of the ProPac™ SCX-10 column, three times identified peptides represent 42 % of the identified peptides in isolated fractions whereas they are 46 % in merged fractions. With the PolySULFOETHYL column, the amount of twice or three times identified peptides is increased from 61 % to 65 % by merging the fractions. These results show that small changes in terms of peptide retention or fractionation have a real incidence on the repeatability of the first separation step of the classical SCX x IP-RP-HPLC-MS/MS setup. Similar repeatabilities were computed for two different columns (operated under different conditions), allowing the generalization of the observed results to setups based on other cation exchangers.

3.5.3 Separation and fractionation repeatability with IP-RP-HPLC at pH 10.0

As performed in the previous section with cation-exchangers, the repeatability of the first dimension of separation and of the fractionation step was determined with the column used to separate the peptides with RP-HPLC at pH 10.0. Thus, three replicate injections of the same tryptic peptide mixture of *C. glutamicum* were performed. Moreover, five fractions were collected in the same retention time window for each replicate. The fractions were analyzed in quintuplicate with the IP-RP-HPLC-ESI-MS/MS system. The results are depicted in Fig. 102. Most of the peptides were detected three times (42 %). Peptides once identified represent 34 %, and peptides twice identified represent 24 % of all identified peptides. These percentages are very similar to those obtained with cation exchangers (41 – 45 %, 16 - 22 %, 37 - 39 % for once, twice, and three times identified peptides, respectively). Thus, the newly established setup presents the same performances than classical SCX setups in terms of repeatability of peptide identification.

In order to check the influence of the fractionation itself, identification reports were merged for the five consecutive fractions as described in the previous section (see 3.5.2). Finally, the percentages of peptides identified once, twice, and three times were computed. The results are depicted in pie-charts in Fig. 103.

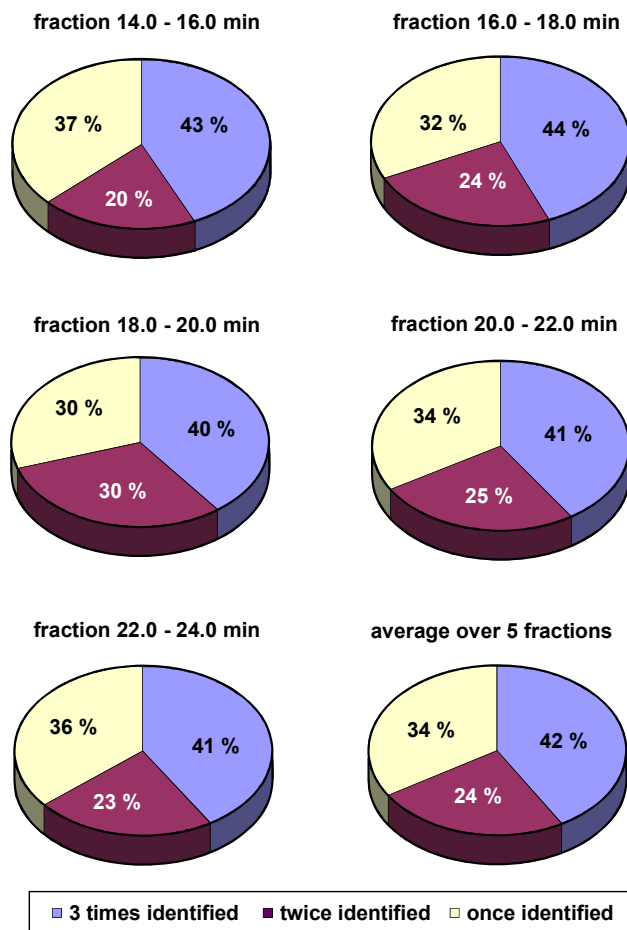


Fig. 102. Repeatability of separation and fractionation with the Gemini column. Each pie-chart corresponds to one fraction collected in triplicate after the C18 column, and analyzed in quintuplicate with IP-RP-HPLC-ESI-MS/MS. The percentages correspond to the amount of peptides identified once, twice, or three times.

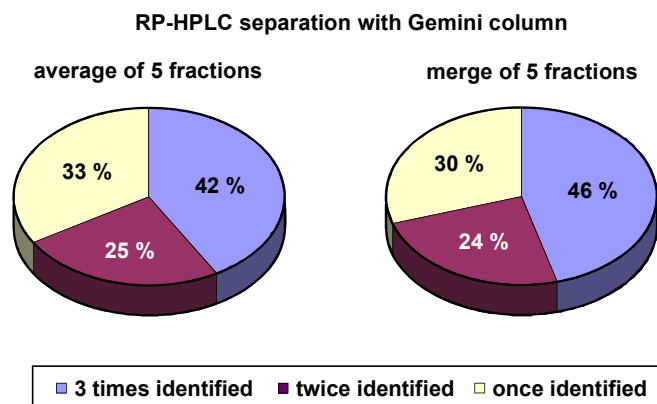


Fig. 103. Repeatability of the separation and fractionation step with the C18 Gemini column. Pie-chart on the left side represents the average repeatability observed for five fractions. Pie-chart on the right side corresponds to the repeatability obtained after a merge of the same five fractions.

By merging the identification reports, 4 % more peptides (46 % instead of 42 %) were identified in all three replicates. This result is very similar to those obtained with cation exchangers (46 % instead of 42 % with ProPac™ SCX-10, and 48 % instead of 45 % with PolySULFOETHYL). It permits to conclude the inaccuracies, that occur during the separation and fractionation step with the new RP x IP-RP-HPLC setup, are equivalent to those made with classical SCX x IP-RP-HPLC setups.

3.6 Dimension orthogonality

For both setups, the orthogonality between the two dimensions of separation was evaluated. This can be relatively easily represented in a plot by depicting each peptide hit as a function of its retention time in both chromatographic dimensions [142;143]. In such a plot, the x-axis represents the retention time of the first dimension of separation, and the y-axis represents the retention time of the second dimension of separation. Because fractions were collected at relatively small intervals (0.25 – 1.00 min), the retention time of a peptide in the first dimension of separation approximately corresponds to the middle point of the collection time window of the fraction in which the peptide was identified. Such plots were computed for the classical SCX x IP-RP-HPLC and the new RP x IP-RP-HPLC approaches. They are depicted in Fig. 104 and in Fig. 105.

The data used to plot the maps are presented in sections 3.2 and 3.3. The sample was a tryptic digest of a protein cell extract of *C. glutamicum*. Redundancies within a fraction are not depicted, but peptides identified in different fractions are taken into account. In case of redundant identifications for peptides within a fraction, the retention time retained to plot the map corresponds to the retention time of the peptide with the best identification (highest MOWSE score). To identify peptides, peaks were integrated in the chromatogram of the second dimension of separation (IP-RP-HPLC at pH 2.1, y-axis) until the peak corresponding to the wash peak and regeneration of the column. It explains the clear-cut peptide distribution at 65 min and 72 min in Fig. 104 and in Fig. 105, respectively.

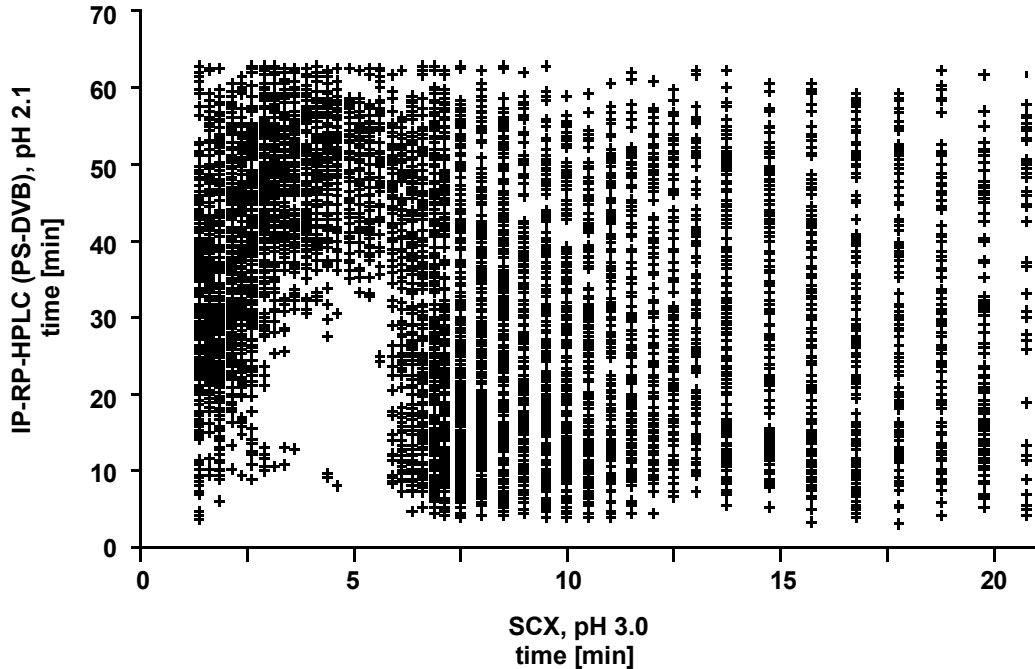


Fig. 104. Orthogonality map obtained for the SCX x IP-RP-HPLC approach. Each peptide hit is represented as a cross. The 6,254 peptide hits used to plot the map are presented in sections 3.2 and 3.3.

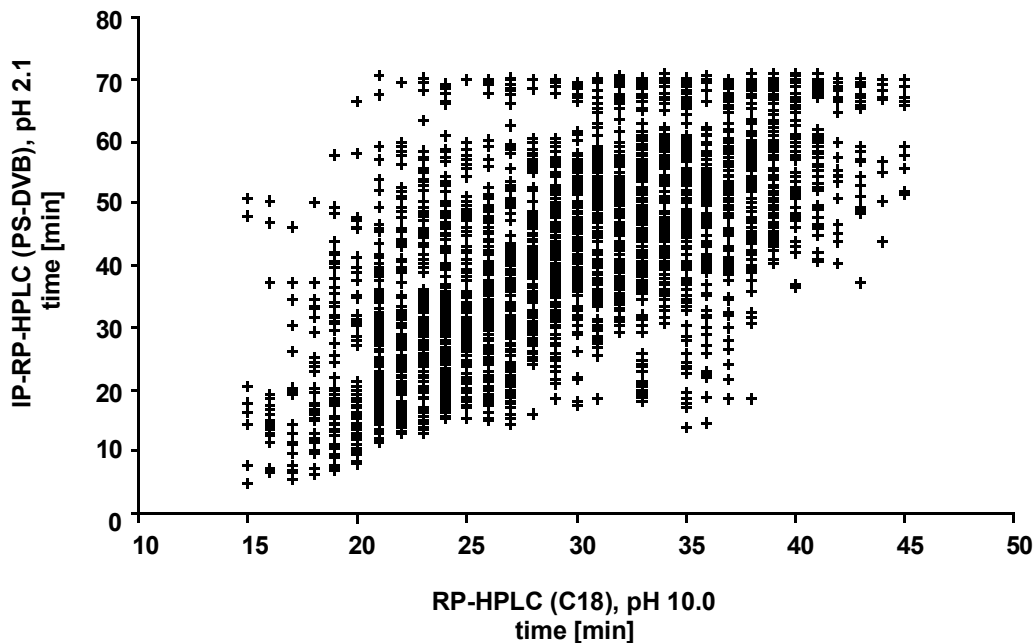


Fig. 105. Orthogonality map obtained for the alkaline RP-HPLC x acidic IP-RP-HPLC approach. Each peptide hit is represented as a cross. The 3,124 peptide hits used to plot the map are presented in sections 3.2 and 3.3.

In the case of the SCX x IP-RP-HPLC approach, a very good orthogonality is observed. Peptide hits are dispersed over the whole diagram map. The higher

density of the peptide distribution in the first 7 min of the SCX separation in comparison with the following 6 min and the last 8 min is explained by the collection of fractions at increasing periods (24 x 0.25 min, 12 x 0.50 min, and 8 x 1.00 min). A small zone for fractions collected between minutes 3 and 6 does not contain any hits. During this period of time, a steep decrease in the number of identified peptides is observed. It corresponds to a salt concentration at which 2+ peptides have already been eluted and 3+ peptides are still retained on the cation exchanger (see Fig. 91).

In the RP-HPLC separation scheme under basic and acid conditions, peptide hits are also well dispersed over the diagram map. However, a clear trend is observed. The later the elution in the first dimension, the later the elution in the second dimension. It reveals that despite the use of different pH (10.0 vs. 2.1) and different stationary phases (silica-based C18 vs. PS-DVB), both dimensions of separation are not truly orthogonal. This result has been expected: the elution was performed for both separations by performing a linear gradient of acetonitrile. The orthogonality of the two RP dimensions should be increased by combining basic and acidic conditions with different eluents (acetonitrile vs. methanol).

Despite a lower orthogonality, the RP x IP-RP-HPLC setup permitted to identify significantly more peptides than the classical SCX x IP-RP-HPLC approach (see 3.3 Proteome coverage). The fact that both RP dimensions are not truly orthogonal does not seem to be a critical factor. The increase in terms of fraction homogeneity and sample decomplexification appears to be much more relevant than dimension orthogonality to better identify peptides.

3.7 Complementarity of SCX x IP-RP-HPLC and RP x IP-RP-HPLC methods

Peptide identification reports obtained with both 2D-HPLC-MS/MS setups were compared (see section 3.3 Proteome coverage). A nonredundant list of identified peptides was extracted from these two identification reports and Σ 3,659 unique peptides were identified. Peptides were finally classified into three categories: peptides only identified with the SCX x IP-RP-HPLC setup (Σ 951 peptides), peptides only identified with the RP x IP-RP-HPLC system (Σ 1,262 peptides), and peptides identified with both setups (Σ 1,447 peptides). The results are depicted in Fig. 106.

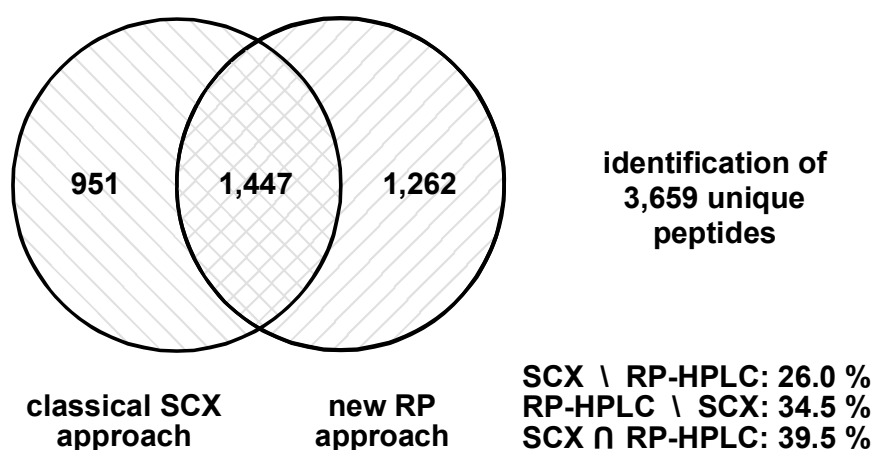


Fig. 106. Intersection diagram of peptides of *C. glutamicum* identified with a classical SCX x IP-RP-HPLC setup and a new RP x IP-RP-HPLC setup.

Peptides detected with only one single setup are representing 60.5 % of the identified peptides (SCX \ RP: 26.0 %, and RP \ SCX: 34.5 %). Only 39.5 % of the peptides were detected with both setups and at first sight both methods are complementary. Thus performing an analysis of *C. glutamicum* with RP x IP-RP-HPLC in addition to a classical SCX x IP-RP-HPLC analysis permits an increase of 52.6 % in terms of unique identified peptides (3,659 vs. 2,398). However, the identification of Σ 3,659 peptides was performed with six analyses in the second separation step, and no longer with three replicates. The supplementary identifications are due to additional injections in the IP-RP-HPLC-ESI-MS/MS system but also to the complementarity of both methods of separation. In order to check the percentage of additional identifications just obtained by increasing the number of analyses, an extrapolation of Fig. 97 was computed. By three from a total of five replicates, 87 % of the peptides are identified with the IP-RP-HPLC-ESI-MS/MS setup. By three from a total of six

replicates, a rigorous estimation is that 80 % of the peptides are identified. With this approximation, the 2,398 peptides identified by means of three replicates represent approximately 80 % of the peptides possibly identified by performing six replicate injections in the classical SCX x IP-RP-HPLC setup. It is then possible to compute the theoretical number of peptides identified by performing six replicate injections in the classical SCX x IP-RP-HPLC setup (~ 2,998). The percent of peptides additionally identified for reason of the complementarity of both setups is obtained by computing:

$$\frac{\text{identifications with 3 injections in each setup} - \text{identifications with 6 replicates in 1 setup}}{\text{identifications with 6 replicates in 1 setup}} \cdot 100$$

The gain in terms of peptide identification due to the complementarity of both setups was evaluated to 22 %. This signifies that the analysis in triplicate of one sample with two different setups is more advantageous than the analysis in sextuplicate with one setup.

Identification reports at the protein level were also compared and the numbers of proteins identified with one setup or with both setups were computed. The results are depicted in Fig. 107. By combining both setups 871 unique proteins were identified (the list is available as appendix). This value corresponds to 29.1 % of the possibly expressed proteins in *C. glutamicum* (see 3.3 Proteome coverage). The combination of the new RP x IP-RP-HPLC setup with the classical SCX x IP-RP-HPLC setup permitted to increase the amount of unique identified proteins by 25 % (695 vs. 871). 126 proteins were just identified with SCX x IP-RP-HPLC and 176 proteins were just identified with RP x IP-RP-HPLC. The amount of proteins identified with both setups is rather high (65.3 %) in comparison with the amount of peptides identified with both setups (39.5 %). This is explained by the fact that different peptides may lead to the identification of one single protein. Consequently, the number of proteins identified with more than one peptide increases by combining both setups. A commonly accepted criterion for the identification of a protein is the necessity to detect more than one single peptide from the protein sequence ^[149;150]. 414 proteins fulfill this criterion in the SCX x IP-RP-HPLC approach and 468 proteins with the RP x IP-RP-HPLC setup (see 3.4 Identification confidence). By combining both methods, 585 proteins are identified by at least two peptides, which corresponds to a gain of 41 % as compared with a single analysis based with the SCX x IP-RPHPLC setup. To sum

up, the combination of both methods of analysis does not only increase the amount of identified proteins but also leads to significantly more confident identifications.

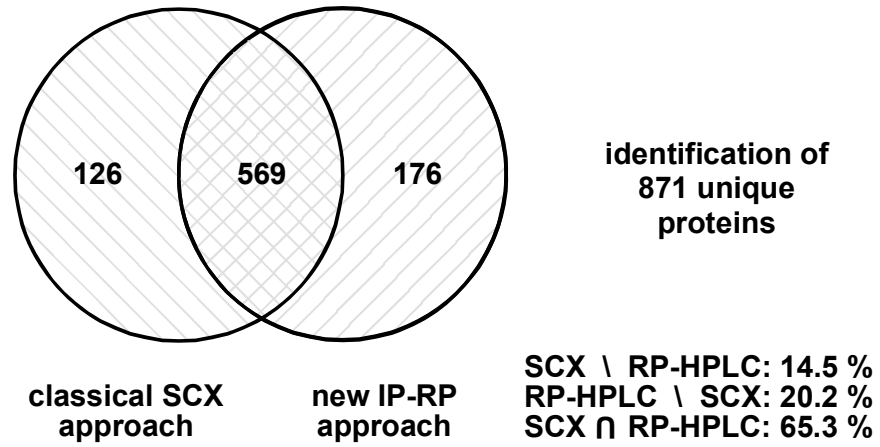


Fig. 107. Intersection diagram of proteins of *C. glutamicum* identified with a classical SCX x IP-RP-HPLC setup and a new RP x IP-RP-HPLC setup.

3.8 Ability of 2D-HPLC setups to analyze proteomes

The aim of this work was to develop and evaluate multi-dimensional HPLC-MS systems and not to compare them with other systems of separation such as two-dimensional gel electrophoresis. However, in order to illustrate that the 2D-HPLC systems developed in this work are not discriminating against proteins of particular pI or molecular mass, the proteins identified with both 2D-HPLC setups were plotted in a $MW = f(pI)$ diagram (see Fig. 108a). A similar diagram was also plotted for the whole proteins potentially expressed in *C. glutamicum* (see Fig. 108b).

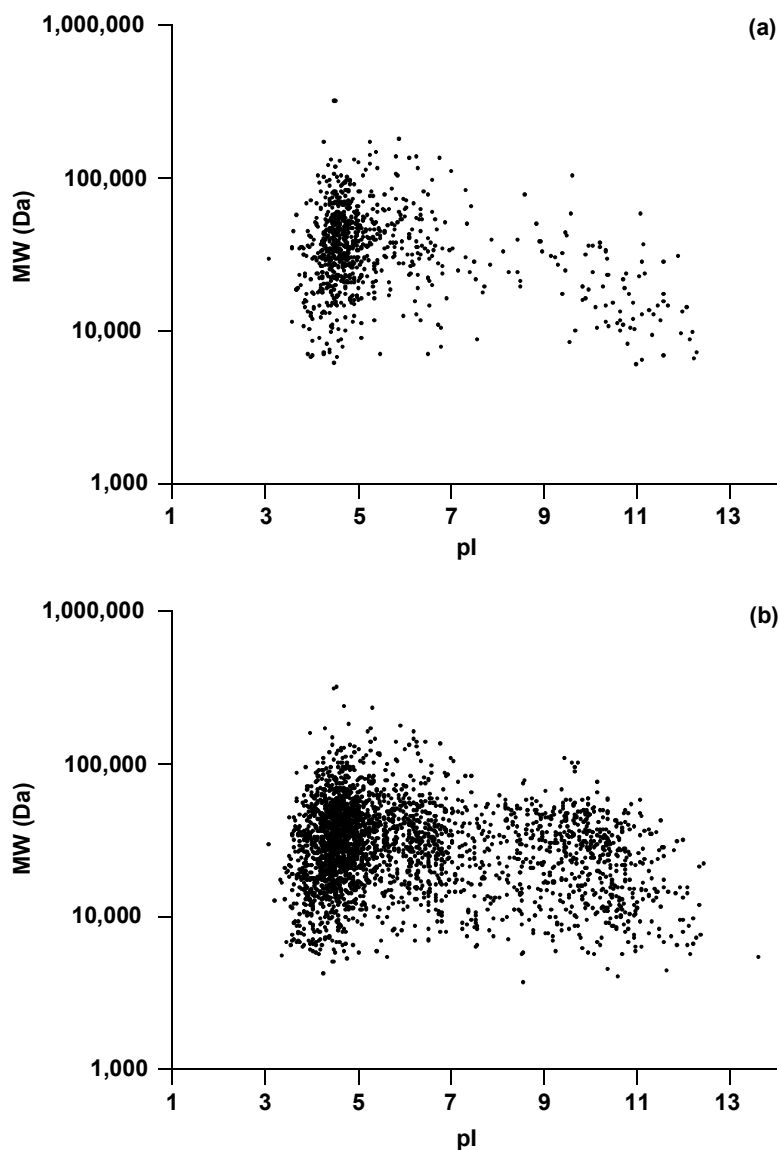


Fig. 108. $MW=f(pI)$ plots representing the (a) identified proteins with both 2D-HPLC setups and (b) the potentially expressed proteins of *C. glutamicum*.

A similar pattern was obtained for both cases, proving that the 2D-HPLC-ESI-MS/MS methods can be used to analyze complete proteomes and not only sets of proteins of particular properties. Finally the amount of enzymes detected by combining both two-dimensional HPLC setups was computed. Enzymes are proteins of particular interest. Their expression or absence in a protein cell extract gives indeed useful data to understand the metabolic or biosynthetic pathways taking place in the production of analytes of interest (e.g. L-glutamate and L-lysine for *C. glutamicum*) [152;153]. 334 enzymes were identified. It corresponds to 56.7 % of the whole possibly expressed enzymes in *C. glutamicum*, proving the ability of the 2D-HPLC-ESI-MS/MS setups to analyze whole proteomes in the frame of biologically relevant challenges. The detection of enzymes studied in an already published work about the production of L-lysine [153] with *C. glutamicum* ATCC 13287 (mutant) was checked. Enzymes involved in central carbon metabolism, energy metabolism, and respiratory chain were detected (see Tab. 13).

Tab. 13. Biologically relevant enzymes of *C. glutamicum* detected by 2D-HPLC-ESI-MS-MS.

locus	enzyme description
phosphotransferase uptake	
NCgl2553	phosphotransferase system IIC component [2.7.1.69]
glycolysis	
NCgl0817	glucose-6-phosphate isomerase [5.3.1.9]
NCgl1202	6-phosphofructokinase [2.7.1.11]
NCgl2673	fructose-bisphosphate aldolase [4.1.2.13]
NCgl1524	triosephosphate isomerase [5.3.1.1]
NCgl1526	glyceraldehyde-3-phosphate dehydrogenase [1.2.1.12]
Krebs cycle	
NCgl0630	citrate synthase [4.1.3.7]
NCgl1482	aconitase A [4.2.1.3]
NCgl0634	monomeric isocitrate dehydrogenase (NADP+) [1.1.1.42]
NCgl1084	2-oxoglutarate dehydrogenases, E1 component [1.2.4.2]
NCgl2477	succinyl-CoA synthetase beta subunit [6.2.1.5]
NCgl0967	fumarase [4.2.1.2]
NCgl2297	malate/lactate dehydrogenase [1.1.1.37]
pentose phosphate pathway	
NCgl1514	glucose-6-phosphate 1-dehydrogenase [1.1.1.49]
NCgl1516	6-phosphogluconolactonase [3.1.1.31]
NCgl1396	6-phosphogluconate dehydrogenase, family 1 [1.1.1.44]
NCgl1513	transaldolase [2.2.1.2]
NCgl1512	transketolase [2.2.1.1]
anaplerosis	
NCgl0659	pyruvate carboxylase [6.4.1.1]
NCgl1523	phosphoenolpyruvate carboxylase [4.1.1.31]
NCgl2765	phosphoenolpyruvate carboxykinase (GTP) [4.1.132]
NCgl2904	malic enzyme [1.1.1.40]
NCgl2247	malate synthase [4.1.3.2]
ATP formation	
NCgl1162	F0F1-type ATP synthase delta subunit [3.6.1.34]
respiratory chain	
NCgl2437	heme/copper-type cytochrome/quinol oxidase, subunit 1 [1.9.3.1]

4 Conclusions

A classical SCX x IP-RP-HPLC setup was established in order to analyze very complex mixtures of peptides. Fractions were collected after the cation exchanger and peptides were separated with IP-RP-HPLC at pH 2.1 in the second dimension. The number of positive charges carried by the peptides appears to be the crucial parameter for the retention of peptides on the cation exchanger. The addition of acetonitrile in the eluents permits to suppress secondary interactions (hydrophobic interactions) between the peptides and the cation exchanger. A quite good repeatability in terms of peptide identification was obtained (by quintuplicate injections 54.0 % of the peptides are identified three or more times). Triplicate analyses appear to be a good compromise in terms of peptide identification and time consumption. A new two-dimensional HPLC-ESI-MS/MS setup was developed for proteome analysis. It consisted of a peptide separation by RP-HPLC under basic conditions (pH 10.0) followed by IP-RP-HPLC under acidic conditions (pH 2.1). Although triethylamine was used to set the alkaline pH, ion-pair mechanism is not taking place during the separation and solvophobic interactions are determining the peptide retention (insufficient triethylamine concentration for ion-pair mechanism). Because the second separation step was identical in both setups, it was possible to compare the first dimension of both methods (SCX vs. RP-HPLC at pH 10.0) with each other. The repeatability of the separations is similar with both setups. The orthogonality between methods of separation is higher in the SCX x IP-RP-HPLC approach than in the RP x IP-RP-HPLC scheme. However, the better peptide distribution and the better separation efficiency achieved with the RP x IP-RP-HPLC setup permitted to identify significantly more peptides than with the classical SCX x IP-RP-HPLC setup. Both approaches are complementary and consequently a combination of both setups permits to identify more peptides than replicate injections performed with a single setup. Both setups are not discriminating against pI and molecular weight of the proteins. Finally, the computation of the number of identified enzymes reveals the ability of the 2D-HPLC-ESI-MS/MS setups to analyze whole proteomes in the frame of biologically relevant challenges.

Chapter VI

References

VI. References

1. Dandeneau, R. D.; Zerenner, E. H. *HRC & CC* **1979**, 2, 351-6.
2. Fenn J. B. *J. Am. Soc. Mass. Spectrom.* **1993**, 4, 524.
3. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, 60, 2299-301.
4. Tanaka, K.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, 2, 151.
5. Biemann, K. *Annu. Rev. Biochem.* **1992**, 61, 977-1010.
6. Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, 3, 327-32.
7. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, 20, 3551-67.
8. Eng, J. K.; McCormack, A. L.; Yates, J. R. III *J. Am. Soc. Mass Spectrom.* **1994**, 5, 976-89.
9. Anderson, N. L.; Anderson, N. G. *Mol. Cell. Proteomics* **2002**, 1, 845-67.
10. Opiteck, G. J.; Lewis, K. C.; Jorgenson, J. W.; Anderegg, R. J. *Anal. Chem.* **1997**, 69, 1518-24.
11. Liu, H.; Berger, S. J.; Chakraborty, A. B.; Plumb, R. S.; Cohen, S. A. *J. Chromatogr. B* **2002**, 782, 267-89.
12. Kachman, M. T.; Wang, H.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Anal. Chem.* **2002**, 74, 1779-91.
13. Wall, D. B.; Kachman, M. T.; Gong, S.; Hinderer, R.; Parus, S.; Misek, D. E.; Hanash, S. M.; Lubman, D. M. *Anal. Chem.* **2000**, 72, 1099-111.
14. Mann, M.; Wilm, M. *Anal. Chem.* **1994**, 66, 4390-9.
15. Engvall, E.; Perlman, P. *Immunochemistry* **1971**, 8, 871-4.
16. Engvall, E.; Jonsson, K.; Perlmann, P. *Biochim. Biophys. Acta* **1971**, 251, 427-34.
17. Dimeski, G.; Barnett, R. J. *Crit. Care Resusc.* **2005**, 7, 12-5.
18. Lottspeich, F.; Engels J. W. *Bioanalytik*, Spektrum: Heidelberg, 2006.

19. *Eur. J. Biochem.* **1984**, *138*, 9-37.
20. Schley, C., Universität des Saarlandes, Saarbrücken, Germany, 2007.
21. Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 729-40.
22. Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 251-6.
23. Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205-11.
24. Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105-32.
25. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. *Electrophoresis* **1995**, *16*, 1090-4.
26. Altairac, S.; Bairoch, A.
<http://www.expasy.org/prolune/dossiers/prolune018.shtml> .
27. Huber, C. G.; Schley, C.; Delmotte, N. *Proteomics and Peptidomics: New Technology Platforms Elucidating Biology*, Marko-Varga, G., Ed.; Elsevier: Amsterdam, 2005.
28. Creighton, T. E. *Proteins: Structures and molecular properties*, 2nd ed.; W. H. Freeman and Company: New York, 1993.
29. Geng, X.; Regnier, F. E. *J. Chromatogr.* **1984**, *296*, 15-30.
30. Heftman, E. *Chromatography*, 5th ed.; Elsevier: Amsterdam, 1992.
31. Chen, H.; Horváth, C. *Anal. Meth. Instr.* **1993**, *1*, 213-22.
32. Hansen, L. C.; Sievers, R. E. *J. Chromatogr.* **1974**, *99*, 123-33.
33. Hjerten, S.; Liao, J.-L.; Zhang, R. *J. Chromatogr.* **1989**, *473*, 273-75.
34. Svec, F.; Fréchet, J. M. J. *Anal. Chem.* **1992**, *64*, 820-22.
35. Minakuchi, H.; Nakanishi, K.; Soga, N.; Ishizuka, N.; Tanaka, N. *Anal. Chem.* **1996**, *68*, 3498-501.
36. Viklund, C.; Svec, F.; Fréchet, J. M. J. *Chem. Mater.* **1996**, *8*, 744-50.
37. Petro, M.; Svec, F.; Fréchet, J. M. J. *J. Chromatogr. A* **1996**, *752*, 59-66.

38. Huber, C. G.; Kleindienst, G.; Bonn, G. K. *Chromatographia* **1997**, *44*, 438-48.
39. Oberacher, H.; Huber, C. G. *Trends Anal. Chem.* **2002**, *21*, 166-73.
40. Premstaller, A.; Oberacher, H.; Walcher, W.; Timperio, A.-M.; Zolla, L.; Chervet, J.-P.; Cavusoglu, N.; Van Dorsselaer, A.; Huber, C. G. *Anal. Chem.* **2001**, *73*, 2390-96.
41. Walcher, W.; Oberacher, H.; Troiani, S.; Hölzl, G.; Oefner, P.; Zolla, L.; Huber, C. G. *J. Chromatogr. B* **2002**, *782*, 111-25.
42. Huber, C. G.; Oefner, P. J.; Bonn, G. K. *Chromatographia* **1993**, *37*, 653-58.
43. Deyl, Z.; Miksik, I.; Tagliaro, F.; Tesarova, E. *Advanced Chromatographic and Electromigration Methods in BioSciences*, Elsevier: Amsterdam, 1998.
44. Kastner, M. *Protein liquid chromatography*, Elsevier: Amsterdam, 2000.
45. Hearn, M. T. W. *HPLC of Proteins, Peptides and Polynucleotides*, VCH Publishers: New York, 1991.
46. Mant, C. T.; Hodges, R. S. *High-Performance Liquid Chromatography of Peptides and Proteins*, CRC Press: Boca Ranton, 1991.
47. Horvath, C.; Melander, W.; Molnar, I. *J. Chromatogr.* **1976**, *125*, 129-56.
48. Patthy, M. *J. Chromatogr. A* **1994**, *660*, 17-23.
49. Melander, W.; El Rassi, Z.; Horváth, C. *J. Chromatogr.* **1989**, *469*, 3.
50. Bartha, A.; Stahlberg, J. *J. Chromatogr. A* **1994**, *668*, 255-84.
51. Bartha, A.; Stahlberg, J. *J. Chromatogr.* **1990**, *535*, 181-87.
52. Chen J.-G.; Weber, S. G.; Glavina, L. L.; Cantwell, F. F. *J. Chromatogr. A* **1993**, *656*, 549-76.
53. Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. *Anal. Chem.* **1985**, *57*, 675-79.
54. Banks, J. F. *J. Chromatogr. A* **1996**, *743*, 99-104.
55. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53-68.

-
56. Karas, M.; Bahr, U. *Trends Anal. Chem.* **1990**, *9*, 321.
 57. Kopaciewicz, W.; Rounds, M. A.; Fausnaugh, J.; Regnier, F. E. *J. Chromatogr.* **1983**, *266*, 3-21.
 58. Mant, C. T.; Litowski, J. R.; Hodges, R. S. *J. Chromatogr. A* **1998**, *816*, 65-78.
 59. Mayr, B. M., Universität des Saarlandes, Saarbrücken, Germany, 2005.
 60. Cuatrecasas, P.; Wilchek, M.; Anfinsen, C. B. *Proc. Natl. Acad. Sci. USA* **1968**, *61*, 636-43.
 61. Weller, M. G. *Fresenius J. Anal. Chem.* **2000**, *366*, 635-45.
 62. Desilets, C. P.; Rounds, M. A.; Regnier, F. E. *J. Chromatogr.* **1991**, *544*, 25-39.
 63. Souverain, S.; Rudaz, S.; Veuthey, J.-L. *J. Chromatogr. B* **2004**, *801*, 141-56.
 64. Hagemeyer, E.; Boos, K.-S.; Schlimme, E.; Lechtenböcker, K.; Kettrup, A. *J. Chromatogr.* **1983**, *268*, 291-95.
 65. Boos, K.-S.; Grimm, C.-H. *Trends Anal. Chem.* **1999**, *18*, 175-80.
 66. Walles, M.; Mullett, W. M.; Pawliszyn, J. *J. Chromatogr. A* **2004**, *1025*, 85-92.
 67. Papp, R.; Mullett, W. M.; Kwong, E. *J. Pharm. Biomed. Anal.* **2004**, *36*, 457-64.
 68. Machtejevas, E.; John, H.; Wagner, K.; Ständker, L.; Marko-Varga, G.; Forssmann, W.-G.; Bischoff, R.; Unger, K. K. *J. Chromatogr. B* **2004**, *803*, 121-30.
 69. Forest, J. C.; Masse, J.; Lane, A. *Clin. Biochem.* **1998**, *31*, 81-8.
 70. Tahara, T.; Usuki, K.; Sato, H.; Ohashi, H.; Morita, H.; Tsumura, H.; Matsumoto, A.; Miyazaki, H.; Urabe, A.; Kato, T. *Br. J. Haematol.* **1996**, *93*, 783-8.
 71. Blackburn, G. F.; Shah, H. P.; Kenten, J. H.; Leland, J.; Kamin, R. A.; Link, J.; Peterman, J.; Powell, M. J.; Shah, A.; Talley, D. B.; Tyagi, S. K.; Wilkins, E.; Wu, T.-G.; Massey, R. J. *Clin. Chem.* **1991**, *37*, 1534-9.

72. Andren, P. E.; Emmett, M. R.; Caprioli, R. M. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 867-69.
73. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240.
74. Iribarne, J. V.; Thomson, A. *J. Chem. Phys.* **1976**, *64*(6), 2287-94.
75. Smith, R. D.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1988**, *60*, 1948-52.
76. Warriner, R. N.; Craze, A. S.; Games, D. E.; Lane, S. J. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1143-49.
77. Lee, E. D.; Mück, W. M.; Henion, J. D.; Covey, T. R. *J. Chromatogr.* **1988**, *474*, 21-37.
78. Snyder, P. A. *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry*, American Chemical Society: 1996.
79. Cole, R. B. *Electrospray Mass Spectrometry: Fundamentals, Instrumentation & Applications*, John Wiley & Sons, Inc.: New York, 1997.
80. Mora, J. F.; Van Berkel, G. J.; Enke, C. G.; Cole, R. B.; Martinez-Sanchez, M.; Fenn, J. B. *J. Mass. Spectrom.* **2000**, *35*, 939-52.
81. De Hoffmann, E., Charette, J., and Stroobant, V. John Wiley & Sons: 2001.
82. Watson J.T. Raven Press: 85.
83. Dawson, P. H. and Bingqi, Y. 84; Vol. 56, pp. 41-50.
84. Dawson, P. H.; Bingqi, Y. *Int. J. Mass Spectrom. Ion Process.* **1984**, *56*, 25-39.
85. Paul W. and Steinwedel, H. S. 60.
86. Stafford Jr., G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 589-96.
87. McLuckey, S. A.; Berkel, G. J.; Goeringer, D. E.; Glish, G. L. *Anal. Chem.* **1994**, *66*, 737-43.
88. Delmotte, N.; Mayr, B.; Leinenbach, A.; Reinert, K.; Kohlbacher, O.; Klein, C.; Huber, C. G. *Dagstuhl Seminar Proceedings, Computational Proteomics* **2005**, 05471.

-
89. Biemann, K. *Anal. Chem.* **1986**, *58*, 1288A-300A.
 90. Biemann, K.; Papayannopoulos, I. A. *Acc. Chem. Res.* **1994**, *27*, 370-78.
 91. Marto, J. A.; Brame, C. J.; Ficarro, S. B.; White, F. M.; Shabanowitz, J.; Hunt, D. F. *Encyclopedia of mass spectrometry* **2004**, *2*, 165-76.
 92. Wang, P.; Kish, M. M.; Wesdemiotis, C. *Encyclopedia of mass spectrometry* **2004**, *2*, 139-51.
 93. Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
 94. Dainese, P.; James, P. *Proteome Research: Mass Spectrometry*, James, P., Ed.; Springer: Berlin, 2001.
 95. Huttlin, E. L.; Hegeman, A. D.; Harms, A. C.; Sussman, M. R. *J. Proteome Res.* **2007**, *6*, 392-8.
 96. Larsson, P. O.; Glad, M.; Hansson, L.; Mansson, M. O.; Ohlson, S.; Mosbach, K. *Adv. Chromatogr.* **1983**, *21*, 41-84.
 97. Lerman, L. S. *Nature* **1953**, *172*, 635-6.
 98. Guerreiro, N.; Charmont, S. *Agilent Technologies Application Note* **2004**.
 99. Brand, J.; Haslberger, T.; Zolg, W.; Pestlin, G.; Palme, S. *Proteomics* **2006**, *6*, 3236-42.
 100. Zolg, J. W.; Langen, H. *Mol. Cell. Proteomics* **2004**, *3*, 345-54.
 101. Zolg, W. *Mol. Cell. Proteomics* **2006**, *5*, 1720-6.
 102. Rucevic, M.; Clifton, J. G.; Huang, F.; Li, X.; Callanan, H.; Hixson, D. C.; Josic, D. *J. Chromatogr. A.* **2006**, *1123*, 199-204.
 103. Lim, Y. P.; Josic, D.; Callanan, H.; Brown, J.; Hixson, D. C. *J. Chromatogr. A* **2005**, *1065*, 39-43.
 104. Strancar, A.; Barut, M.; Podgornik, A.; Koselj, P.; Josic, D.; Buchacher, A. *LC GC* **1998**, 660-70.
 105. Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85-8.

106. Stryer, L. *Biochemistry*, 4th ed.; W. H. Freeman and Company: New York, 1995.
107. Kagen, L.; Scheidt, S.; Roberts, L.; Porter, A.; Paul, H. *Am. J. Med.* **1975**, *58*, 177-82.
108. Bakker, A. J.; Koelemay, M. J.; Gorgels, J. P.; van Vlies, B.; Smits, R.; Tijssen, J. G.; Haagen, F. D. *Eur. Heart J.* **1994**, *15*, 45-53.
109. Honda, Y.; Katayama, T. *Int. J. Cardiol.* **1984**, *6*, 325-39.
110. Lavin, F.; Kane, M.; Forde, A.; Gannon, F.; Daly, K. *Br. Heart J.* **1995**, *73*, 422-7.
111. Cloonan, M. J.; Bishop, G. A.; Wilton-Smith, P. D.; Carter, I. W.; Allan, R. M.; Wilcken, D. E. *Pathology* **1979**, *11*, 689-99.
112. Bhayana, V.; Cohoe, S.; Pellar, T. G.; Jablonsky, G.; Henderson, A. R. *Clin. Biochem.* **1994**, *27*, 395-406.
113. Gröpl C.; Lange, E.; Reinert, K.; Kohlbacher, O.; Sturm, M.; Huber, C. G.; Mayr, B. M.; Klein, C. *Lecture Notes in Bioinformatics, Proceedings of the First Symposium on Computational Life Sciences (CLS 2005)*, Berthold, M.; Glen, R.; Diederichs, K. Kohlbacher, O. F. I., Eds.; Springer, Heidelberg: 2005.
114. Mayr, B. M.; Kohlbacher, O.; Reinert, K.; Sturm, M.; Gropl, C.; Lange, E.; Klein, C.; Huber, C. G. *J. Proteome Res.* **2006**, *5*, 414-21.
115. Hunt, P. J.; Richards, A. M.; Nicholls, M. G.; Yandle, T. G.; Doughty, R. N.; Espiner, E. A. *Clin. Endocrinol.* **1997**, *47*, 287-96.
116. De Lemos, J. A.; McGuire, D. K.; Drazner, M. H. *Lancet* **2003**, *362*, 316-22.
117. Karl, J., Lill, H., Stahl, P., Krueger, K., Borgya, A., and Gallusser, A. European Patent Office, 2006.
118. Richards, A. M.; Nicholls, M. G.; Yandle, T. G.; Frampton, C.; Espiner, E. A.; Turner, J. G.; Buttimore, R. C.; Lainchbury, J. G.; Elliott, J. M.; Ikram, H.; Crozier, I. G.; Smyth, D. W. *Circulation* **1998**, *97*, 1921-9.
119. Hall, Ch. *The European Journal of Heart Failure* **2004**, *6*, 257-60.
120. Premstaller, A.; Oberacher, H.; Huber, C. G. *Anal. Chem.* **2000**, *72*, 4386-93.

-
121. Roche Diagnostics *proBNP Elecsys kit, 03121640/122* **2006**, 1-5.
 122. Hendrickson, W. A.; Pahler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2190-4.
 123. Gupalova, T. V.; Lojkina, O. V.; Palagnuk, V. G.; Totolian, A. A.; Tennikov, T. B. *J. Chromatogr. A* **2002**, *949*, 185-93.
 124. Hage, D. S. *J. Chromatogr. B* **2002**, *768*, 3-30.
 125. Hennion, M.-C. *J. Chromatogr. A* **1999**, *856*, 3-54.
 126. Gilar, M.; Bouvier, E. S.; Compton, B. J. *J. Chromatogr. A* **2001**, *909*, 111-35.
 127. Krishnan, T. R.; Ibrahim, I. *J. Pharm. Biomed. Anal.* **1994**, *12*, 287-94.
 128. Sokol, J.; Matisova, E. *J. Chromatogr. A* **1994**, *669*, 75-80.
 129. Beaudry, F.; del Castillo, J. R. *Biomed. Chromatogr.* **2005**, *19*, 523-8.
 130. Gordon, J. M.; Walker, C. B.; Murphy, J. C.; Goodson, J. M.; Socransky, S. S. *J. Clin. Periodontol.* **1981**, *8*, 117-21.
 131. Gordon, J. M.; Walker, C. B.; Murphy, J. C.; Goodson, J. M.; Socransky, S. S. *J. Periodontol.* **1981**, *52*, 609-12.
 132. Ylitalo, P.; Hinkka, H.; Neuvonen, P. J. *Eur. J. Clin. Pharmacol.* **1977**, *12*, 367-73.
 133. Pfeiffer, C. M.; Fazili, Z.; McCoy, L.; Zhang, M.; Gunter, E. W. *Clin. Chem.* **2004**, *50*, 423-32.
 134. Fazili, Z.; Pfeiffer, C. M. *Clin. Chem.* **2004**, *50*, 2378-81.
 135. Fazili, Z.; Pfeiffer, C. M.; Zhang, M.; Jain, R. *Clin. Chem.* **2005**, *51*, 2318-25.
 136. Hermann, T. *J. Biotechnol.* **2003**, *104*, 155-72.
 137. Sahm, H.; Eggeling, L.; de Graaf, A. A. *Biol. Chem.* **2000**, *381*, 899-910.
 138. Kalinowski, J.; Bathe, B.; Bartels, D.; Bischoff, N.; Bott, M.; Burkovski, A.; Dusch, N.; Eggeling, L.; Eikmanns, B. J.; Gaigalat, L.; Goesmann, A.; Hartmann, M.; Huthmacher, K.; Kramer, R.; Linke, B.; McHardy, A. C.; Meyer, F.; Mockel, B.; Pfefferle, W.; Puhler, A.; Rey, D. A.; Ruckert, C.; Rupp, O.;

- Sahm, H.; Wendisch, V. F.; Wiegrabe, I.; Tauch, A. *J. Biotechnol.* **2003**, *104*, 5-25.
139. Krubasik, P.; Kobayashi, M.; Sandmann, G. *Eur. J. Biochem.* **2001**, *268*, 3702.
140. Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M., and Ozaki, A. European Patent Application, 2001.
141. Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248-54.
142. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Anal. Chem.* **2005**, *77*, 6426-34.
143. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *J. Sep. Sci.* **2005**, *28*, 1694-703.
144. Toll, H.; Oberacher, H.; Swart, R.; Huber, C. G. *J. Chromatogr. A* **2005**.
145. Wyndham, K. D.; O'Gara, J. E.; Walter, T. H.; Glose, K. H.; Lawrence, N. L.; Alden, B. A.; Izzo, G. S.; Hudalla, C. J.; Iraneta, P. C. *Anal. Chem.* **2003**, *75*, 6781-8.
146. Wu, A. B.; Huang, M. C.; Ho, H. O.; Yeh, G. C.; Sheu, M. T. *Biomed. Chromatogr.* **2004**, *18*, 443-9.
147. Uesugi, T.; Sano, K.; Uesawa, Y.; Ikegami, Y.; Mohri, K. *J. Chromatogr. B Biomed. Sci. Appl.* **1997**, *703*, 63-74.
148. Schley, C.; Altmeyer, M.; Müller, R.; Swart, R.; Huber, C. G. *J. Proteome Res.* **2006**, *5*, 2760-68.
149. Bradshaw, R. A. *Mol. Cell. Proteomics* **2005**, *4*, 1223-5.
150. Bradshaw, R. A.; Burlingame, A. L.; Carr, S.; Aebersold, R. *Mol. Cell. Proteomics* **2006**, *5*, 787-8.
151. Inouye, S. *Appl. Microbiol.* **1973**, *25*, 279-83.
152. Kiefer, P.; Heinzle, E.; Zelder, O.; Wittmann, C. *Appl. Environ. Microbiol.* **2004**, *70*, 229-39.
153. Kromer, J. O.; Sorgenfrei, O.; Klopprogge, K.; Heinzle, E.; Wittmann, C. *J. Bacteriol.* **2004**, *186*, 1769-84.

Appendix

Nonredundant list of proteins of
C. glutamicum ATCC 13032
identified with
2D-HPLC-ESI-MS/MS

Appendix

Nr.	locus	protein description	mass (Da)	score
1	NCgl0245	isopropylmalate/homocitrate/citramalate synthase [4.1.3.12]	68,435	54,752
2	NCgl1094	methionine synthase II [2.1.1.14]	81,611	44,876
3	NCgl0480	elongation factor Tu	43,998	30,315
4	NCgl1235	phosphoglycerate dehydrogenase [1.1.1.95]	55,384	20,050
5	NCgl1304	ribosomal protein S1	53,998	16,733
6	NCgl0935	enolase [4.2.1.11]	44,979	15,515
7	NCgl1255	glucan phosphorylase [2.4.1.1]	90,728	14,740
8	NCgl0802	fatty-acid synthase [2.3.1.85]	315,398	13,985
9	NCgl0478	elongation factor G	78,075	12,933
10	NCgl2167	pyruvate dehydrogenase, decarboxylase component [1.2.4.1]	102,820	10,644
11	NCgl1950	ribosomal protein S2	30,158	10,256
12	NCgl0460	ribosomal protein L1	25,031	10,120
13	NCgl0472	DNA-directed RNA polymerase beta subunit/160 kD subunit [2.7.7.6]	147,899	9,584
14	NCgl2673	fructose-bisphosphate aldolase [4.1.2.13]	37,365	9,527
15	NCgl0490	ribosomal protein L2	31,164	9,116
16	NCgl1165	F0F1-type ATP synthase beta subunit [3.6.1.34]	52,621	8,439
17	NCgl0488	ribosomal protein L4	23,594	7,854
18	NCgl2375	ABC-type transporter, periplasmic component	49,475	7,609
19	NCgl0520	ribosomal protein L15	15,417	7,387
20	NCgl1084	2-oxoglutarate dehydrogenases, E1 component [1.2.4.2]	138,844	7,151
21	NCgl0954	glycine hydroxymethyltransferase [2.1.2.1]	46,570	7,110
22	NCgl2702	70 kDa heat shock chaperonin protein	66,247	7,062
23	NCgl0541	ribosomal protein L17	17,507	6,920
24	NCgl0659	pyruvate carboxylase [6.4.1.1]	123,433	6,787
25	NCgl2826	superoxide dismutase [1.15.1.1]	22,088	6,430
26	NCgl0670	acyl-CoA carboxylase [6.3.4.14]	63,554	6,372
27	NCgl2621	chaperonin GroEL	57,319	6,036
28	NCgl0795	citrate synthase [4.1.3.7]	48,956	6,021
29	NCgl2894	myo-inositol-1-phosphate synthase	39,373	5,938
30	NCgl1163	F0F1-type ATP synthase alpha subunit [3.6.1.34]	60,149	5,726
31	NCgl1409	NADH dehydrogenase, FAD-containing subunit [1.6.99.3]	51,042	5,708
32	NCgl1999	glutamate dehydrogenase/leucine dehydrogenase [1.4.1.4]	49,363	5,701
33	NCgl0725	ribosome-associated protein Y	24,558	5,481
34	NCgl0754	pyridoxine biosynthesis enzyme	33,635	5,405
35	NCgl1512	transketolase [2.2.1.1]	75,222	5,231
36	NCgl0539	ribosomal protein S4	23,314	5,157
37	NCgl1482	aconitase A [4.2.1.3]	102,510	5,126
38	NCgl1346	argininosuccinate synthase [6.3.4.5]	44,223	5,060
39	NCgl0491	ribosomal protein S19	10,454	4,944
40	NCgl0634	monomeric isocitrate dehydrogenase (NADP+) [1.1.1.42]	80,148	4,770
41	NCgl2409	3-oxoacyl-(acyl-carrier-protein) synthase [2.3.1.85]	316,847	4,739
42	NCgl1524	triosephosphate isomerase [5.3.1.1]	27,360	4,598
43	NCgl0540	DNA-directed RNA polymerase alpha subunit/40 kD subunit [2.7.7.6]	36,706	4,482
44	NCgl1976	ribosomal protein S16	17,826	4,448
45	NCgl0556	ribosomal protein L13	16,328	4,430
46	NCgl0572	co-chaperonin GroES	10,763	4,290
47	NCgl0487	ribosomal protein L3	23,203	4,247
48	NCgl2133	glutamine synthase [6.3.1.2]	53,383	4,075
49	NCgl0286	cAMP-binding domain containing protein	25,030	4,014
50	NCgl0355	dihydrolipoamide dehydrogenase/glutathione oxidoreductase-like protein [1.8.1.4]	50,852	3,951
51	NCgl0827	phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase [2.1.2.3]	56,015	3,942
52	NCgl2008	pyruvate kinase [2.7.1.40]	51,671	3,825
53	NCgl2480	acetyl-CoA hydrolase [2.8.3.-]	54,541	3,686
54	NCgl2377	ABC-type transporter, ATPase component	40,348	3,624

Nr.	locus	protein description	mass (Da)	score
55	NCgl1525	3-phosphoglycerate kinase [2.7.2.3]	42,787	3,623
56	NCgl0471	DNA-directed RNA polymerase beta subunit/140 kD subunit [2.7.7.6]	128,199	3,509
57	NCgl0537	ribosomal protein S13	13,884	3,372
58	NCgl2473	cysteine synthase [4.2.99.8]	33,449	3,361
59	NCgl1949	translation elongation factor Ts	29,320	3,271
60	NCgl0837	ribosomal protein L31	10,003	3,173
61	NCgl0469	ribosomal protein L7/L12	13,293	3,138
62	NCgl1858	phosphoenolpyruvate-protein kinase [2.7.3.9]	59,977	3,082
63	NCgl2198	class II glycyl-tRNA synthetase [6.1.1.14]	53,243	3,069
64	NCgl2217	4-alpha-glucanotransferase [2.4.1.25]	78,881	3,019
65	NCgl1513	transaldolase [2.2.1.2]	38,363	3,017
66	NCgl0493	ribosomal protein S3	28,150	3,007
67	NCgl0967	fumarase [4.2.1.2]	49,849	2,943
68	NCgl0494	ribosomal protein L16/L10E	15,777	2,921
69	NCgl0496	ribosomal protein S17	10,574	2,876
70	NCgl2126	dihydrolipoamide acyltransferase [2.3.1.61]	70,862	2,870
71	NCgl0390	phosphoglycerate mutase 1 [5.4.2.1]	27,287	2,844
72	NCgl1224	ketol-acid reductoisomerase [1.1.1.86]	36,252	2,838
73	NCgl0033	peptidyl-prolyl cis-trans isomerase (rotamase) [5.2.1.8]	20,111	2,837
74	NCgl0492	ribosomal protein L22	12,962	2,831
75	NCgl1136	homoserine dehydrogenase [1.1.1.3]	46,526	2,825
76	NCgl2048	methionine synthase II [2.1.1.14]	44,042	2,768
77	NCgl1385	FHA-domain-containing protein	15,393	2,735
78	NCgl0303	cold shock protein	7,286	2,690
79	NCgl2715	sulfate adenylate transferase subunit 1	46,901	2,659
80	NCgl0794	phosphoserine aminotransferase [2.6.1.52]	40,137	2,617
81	NCgl2340	aminopeptidase N [3.4.11.2]	96,326	2,603
82	NCgl1523	phosphoenolpyruvate carboxylase [4.1.1.31]	103,308	2,491
83	NCgl0579	putative inosine-5-monophosphate dehydrogenase	39,862	2,491
84	NCgl0582	GMP synthase [6.3.5.2]	56,242	2,386
85	NCgl0677	detergent sensitivity rescuer dtsR2	58,308	2,385
86	NCgl2293	valyl-tRNA synthetase [6.1.1.9]	101,754	2,380
87	NCgl0468	ribosomal protein L10	17,945	2,343
88	NCgl0516	ribosomal protein L6	19,324	2,336
89	NCgl2360	cystathionine gamma-synthase [4.2.99.9]	41,806	2,314
90	NCgl0247	aspartokinase [2.7.2.4]	44,959	2,285
91	NCgl1547	carbamoylphosphate synthase large subunit [6.3.5.5]	120,600	2,262
92	NCgl1926	predicted dehydrogenase [1.1.99.16]	54,972	2,246
93	NCgl0517	ribosomal protein L18	14,736	2,243
94	NCgl2191	glucosamine 6-phosphate synthetase [2.6.1.16]	67,464	2,190
95	NCgl2719	putative ferredoxin/ferredoxin-NADP reductase	50,157	2,161
96	NCgl2329	FKBP-type peptidyl-prolyl cis-trans isomerase	49,726	2,151
97	NCgl1222	thiamine pyrophosphate-requiring enzyme [4.1.3.18]	67,036	2,129
98	NCgl2585	ATPase with chaperone activity, ATP-binding subunit	101,451	2,115
99	NCgl0578	inosine monophosphate dehydrogenase [1.1.1.205]	53,559	2,113
100	NCgl2494	phosphoribosylaminoimidazol (AIR) synthetase [6.3.3.1]	38,233	2,036
101	NCgl0313	class III Zn-dependent alcohol dehydrogenase [1.2.1.-]	39,365	2,014
102	NCgl0807	hypothetical protein	10,143	2,002
103	NCgl0499	ribosomal protein L14	13,388	1,997
104	NCgl2879	ribosomal protein L9	15,930	1,966
105	NCgl2582	L-2,3-butanediol dehydrogenase	27,047	1,932
106	NCgl0345	UDP-N-acetylglucosamine enolpyruvyl transferase [2.5.1.7]	47,255	1,912
107	NCgl1041	peroxiredoxin [1.11.1.-]	17,970	1,912
108	NCgl0248	aspartate-semialdehyde dehydrogenase [1.2.1.11]	36,306	1,911
109	NCgl0237	aspartate transaminase [2.6.1.1]	46,605	1,901

Appendix

Nr.	locus	protein description	mass (Da)	score
110	NCgl2499	phosphoribosylformylglycinamide (FGAM) synthase, synthetase domain [6.3.5.3]	81,430	1,883
111	NCgl2915	leucyl-tRNA synthetase	106,886	1,882
112	NCgl0477	ribosomal protein S7	17,528	1,866
113	NCgl0501	ribosomal protein L5	21,775	1,814
114	NCgl2165	hypothetical protein	15,476	1,799
115	NCgl1557	translation elongation factor P	20,628	1,767
116	NCgl2280	ribosomal protein L21	10,852	1,766
117	NCgl0976	putative fructose-1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase	35,521	1,731
118	NCgl1262	3-isopropylmalate dehydratase large subunit [4.2.1.33]	52,020	1,729
119	NCgl0251	catalase [1.11.1.6]	58,731	1,716
120	NCgl0719	S-adenosylhomocysteine hydrolase [3.3.1.1]	52,624	1,704
121	NCgl2509	adenylosuccinate lyase [4.3.2.2]	52,319	1,693
122	NCgl1871	hypothetical protein	50,498	1,690
123	NCgl0625	O-acetylhomoserine sulfhydrylase [4.2.99.10]	46,724	1,684
124	NCgl2765	phosphoenolpyruvate carboxykinase (GTP) [4.1.1.32]	67,238	1,660
125	NCgl1573	aspartyl-tRNA synthetase [6.1.1.12]	67,201	1,627
126	NCgl2287	nucleoside diphosphate kinase [2.7.4.6]	14,791	1,591
127	NCgl1900	polyribonucleotide nucleotidyltransferase [2.7.7.8]	81,395	1,565
128	NCgl2443	ribonucleotide reductase alpha subunit	81,433	1,557
129	NCgl1396	6-phosphogluconate dehydrogenase, family 1 [1.1.1.44]	51,818	1,506
130	NCgl0610	ABC-type transporter, periplasmic component	31,780	1,505
131	NCgl0515	ribosomal protein S8	14,282	1,505
132	NCgl2793	seryl-tRNA synthetase [6.1.1.11]	46,616	1,501
133	NCgl1501	ABC-type transporter, ATPase component	27,623	1,475
134	NCgl1910	translation initiation factor 2	103,523	1,455
135	NCgl2669	adenylosuccinate synthase [6.3.4.4]	46,914	1,430
136	NCgl0671	thiosulfate sulfurtransferase [2.8.1.1]	33,531	1,419
137	NCgl1244	glutamyl- and glutamyl-tRNA synthetases [6.1.1.17]	55,414	1,399
138	NCgl2328	ATP-dependent Clp protease proteolytic subunit 1 [3.4.21.92]	21,253	1,394
139	NCgl2500	phosphoribosylformylglycinamide (FGAM) synthase, glutamine amidotransferase domain [6.3.5.3]	23,641	1,394
140	NCgl2579	carbonic anhydrase [4.2.1.1]	22,331	1,387
141	NCgl1886	phage shock protein A (IM30)	30,411	1,380
142	NCgl2747	PLP-dependent aminotransferase [2.6.1.1]	48,899	1,380
143	NCgl0533	adenylate kinase [2.7.4.3]	19,415	1,376
144	NCgl1219	dihydroxyacid dehydratase/phosphogluconate dehydratase	65,213	1,373
145	NCgl1607	threonyl-tRNA synthetase [6.1.1.3]	76,933	1,363
146	NCgl2521	thiamine pyrophosphate-requiring enzyme [1.2.2.2]	62,375	1,362
147	NCgl1915	ABC-type transporter, periplasmic component	57,711	1,349
148	NCgl0658	dihydrolipoamide dehydrogenase [1.8.1.4]	49,800	1,326
149	NCgl1545	hypothetical protein	11,860	1,321
150	NCgl1456	hypothetical protein	14,245	1,298
151	NCgl0880	hypothetical protein	37,649	1,279
152	NCgl0500	ribosomal protein L24	11,182	1,267
153	NCgl0817	glucose-6-phosphate isomerase [5.3.1.9]	59,130	1,264
154	NCgl1024	quinolinate synthase [1.4.3.-]	47,000	1,261
155	NCgl1152	transcription termination factor	83,976	1,259
156	NCgl2337	ribose 5-phosphate isomerase RpiB [5.3.1.26]	17,151	1,258
157	NCgl1177	1,4-alpha-glucan branching enzyme [2.4.1.18]	82,715	1,254
158	NCgl2530	predicted hydrolase of the HAD superfamily	29,633	1,252
159	NCgl1053	predicted membrane GTPase involved in stress response	68,775	1,246
160	NCgl2449	putative Zn-NADPH,quinone dehydrogenase [1.1.1.1]	35,877	1,245
161	NCgl0558	phosphomannomutase	46,441	1,238
162	NCgl2368	ABC-type transporter, duplicated ATPase component	62,336	1,232
163	NCgl2297	malate/lactate dehydrogenase [1.1.1.37]	34,887	1,219

Nr.	locus	protein description	mass (Da)	score
164	NCgl0573	chaperonin GroEL	56,711	1,200
165	NCgl0486	ribosomal protein S10	11,463	1,189
166	NCgl2261	ribosomal protein S20	9,547	1,189
167	NCgl0905	phosphoribosylpyrophosphate synthetase [2.7.6.1]	35,818	1,185
168	NCgl1164	F0F1-type ATP synthase gamma subunit [3.6.1.34]	35,718	1,185
169	NCgl2586	inositol-monophosphate dehydrogenase	50,903	1,162
170	NCgl2475	predicted acetyltransferase	10,681	1,158
171	NCgl2068	isoleucyl-tRNA synthetase [6.1.1.5]	117,621	1,148
172	NCgl0171	cold shock protein	7,283	1,147
173	NCgl2810	L-lactate dehydrogenase [1.1.1.27]	34,500	1,142
174	NCgl2098	3-Deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase	51,587	1,142
175	NCgl1570	alanyl-tRNA synthetase [6.1.1.7]	96,456	1,129
176	NCgl1223	acetolactate synthase, small subunit [4.1.3.18]	18,737	1,123
177	NCgl0776	ABC-type cobalamin/Fe ³⁺ -siderophore transport system, periplasmic component	35,726	1,118
178	NCgl1202	6-phosphofructokinase [2.7.1.11]	37,614	1,117
179	NCgl1947	ribosome recycling factor	20,736	1,102
180	NCgl1336	phenylalanyl-tRNA synthetase beta subunit [6.1.1.20]	89,729	1,101
181	NCgl0518	ribosomal protein S5	22,714	1,094
182	NCgl1503	predicted iron-regulated ABC-type transporter SufB	53,633	1,091
183	NCgl1561	chorismate synthase [4.6.1.4]	43,764	1,088
184	NCgl2116	asparagine synthase	72,517	1,079
185	NCgl0317	nucleoside-diphosphate-sugar epimerase	33,574	1,063
186	NCgl1676	hypothetical protein	13,220	1,059
187	NCgl0489	ribosomal protein L23	11,081	1,056
188	NCgl0620	5,10-methylene-tetrahydrofolate dehydrogenase [1.5.1.5]	30,363	1,056
189	NCgl2015	phosphoribosylformimino-5-aminoimidazole carboxamide ribonucleotide (ProFAR) isomerase [5.3.1.16]	26,651	1,046
190	NCgl0134	hypothetical protein	62,509	1,041
191	NCgl0388	acyl-CoA synthetase [6.2.1.3]	62,706	1,024
192	NCgl0726	preprotein translocase subunit SecA	95,461	1,014
193	NCgl1353	universal stress protein UspA	16,160	1,008
194	NCgl2508	phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase [6.3.2.6]	33,260	988
195	NCgl2698	NAD-dependent aldehyde dehydrogenase [1.2.1.3]	55,246	987
196	NCgl1960	ribosomal protein L19	12,867	985
197	NCgl2682	ATPase with chaperone activity, ATP-binding subunit	93,174	973
198	NCgl0226	hypothetical protein	7,072	969
199	NCgl0005	DNA gyrase (topoisomerase II) B subunit [5.99.1.3]	76,082	964
200	NCgl2718	putative nitrite reductase	63,302	956
201	NCgl1320	hypothetical protein	33,325	955
202	NCgl1472	methylmalonyl-CoA mutase, N-terminal domain/subunit [5.4.99.2]	65,566	949
203	NCgl1352	tyrosyl-tRNA synthetase [6.1.1.1]	46,532	944
204	NCgl0834	ribosomal protein L28	8,884	936
205	NCgl1541	S-adenosylmethionine synthetase [2.5.1.6]	44,265	912
206	NCgl2123	branched-chain amino acid aminotransferase [2.6.1.42]	40,454	912
207	NCgl2773	putative polyketide synthase	172,491	908
208	NCgl1343	PLP-dependent aminotransferase [2.6.1.11]	41,470	905
209	NCgl2511	phosphoribosylamine-glycine ligase [6.3.4.13]	43,750	899
210	NCgl2555	glucosamine-6-phosphate isomerase [5.3.1.10]	27,435	889
211	NCgl0351	predicted UDP-glucose 6-dehydrogenase [1.1.1.22]	43,030	879
212	NCgl0187	FAD/FMN-containing dehydrogenase	53,087	879
213	NCgl1845	Mn-dependent transcriptional regulator	25,527	871
214	NCgl2026	pullulanase	94,565	870
215	NCgl0002	DNA polymerase III beta subunit [2.7.7.7]	42,557	867
216	NCgl0012	DNA gyrase (topoisomerase II) A subunit [5.99.1.3]	94,918	859
217	NCgl1211	Asp-tRNAAsn/Glu-tRNA ^{Gln} amidotransferase B subunit	54,548	851

Appendix

Nr.	locus	protein description	mass (Da)	score
218	NCgl2365	predicted thioesterase	17,584	839
219	NCgl0495	ribosomal protein L29	8,760	829
220	NCgl2886	transcriptional regulator	17,958	829
221	NCgl1263	3-isopropylmalate dehydratase small subunit [4.2.1.33]	22,185	829
222	NCgl2077	UDP-N-acetylmuramate-alanine ligase [6.3.2.8]	51,135	824
223	NCgl1335	phenylalanyl-tRNA synthetase alpha subunit [6.1.1.20]	38,508	801
224	NCgl1996	hypothetical protein	29,569	795
225	NCgl2774	putative acyl-CoA synthetase	67,873	785
226	NCgl2628	hypothetical protein	131,909	777
227	NCgl2620	hypothetical protein	36,079	775
228	NCgl1305	phosphotransferase system IIC component, glucose/maltose/N-acetylglucosamine-specific	72,699	772
229	NCgl0385	hypothetical protein	17,819	758
230	NCgl1982	nitrogen regulatory protein PII	12,240	757
231	NCgl2269	hypothetical protein	17,959	757
232	NCgl1423	glycosyltransferase	30,115	754
233	NCgl1132	arginyl-tRNA synthetase [6.1.1.19]	59,802	753
234	NCgl1502	predicted iron-regulated ABC-type transporter SufB	42,370	752
235	NCgl1585	histidyl-tRNA synthetase [6.1.1.21]	47,271	748
236	NCgl1448	phosphoribosyl-ATP pyrophosphohydrolase [3.6.1.31]	9,816	740
237	NCgl1073	ADP-glucose pyrophosphorylase [2.7.7.27]	43,949	728
238	NCgl2737	putative membrane protease subunit	34,823	726
239	NCgl1182	electron transfer flavoprotein beta-subunit	27,204	716
240	NCgl2716	sulfate adenyltransferase subunit 2	34,384	708
241	NCgl1938	gcpE-like protein	41,815	704
242	NCgl2984	thioredoxin reductase [1.6.4.5]	34,408	703
243	NCgl0519	ribosomal protein L30/L7E	6,859	695
244	NCgl2904	malic enzyme [1.1.1.40]	41,024	691
245	NCgl0186	putative dehydrogenase	26,706	688
246	NCgl1835	transcriptional regulator [2.7.1.63]	26,671	679
247	NCgl2259	membrane GTPase LepA	68,911	676
248	NCgl2806	hypothetical protein	31,744	676
249	NCgl0857	methionyl-tRNA synthetase [6.1.1.10]	67,867	672
250	NCgl2973	hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	28,217	670
251	NCgl0328	nitroreductase [1.6.99.3]	21,209	669
252	NCgl1890	hypothetical protein	10,601	663
253	NCgl2277	aldo/keto reductase [1.1.1.-]	31,038	656
254	NCgl1153	protein chain release factor A	39,711	654
255	NCgl1836	DNA-directed RNA polymerase sigma subunit SigA	54,738	653
256	NCgl1901	ribosomal protein S15P/S13E	10,304	647
257	NCgl0811	inositol monophosphatase family protein	27,555	646
258	NCgl2528	D-2-hydroxyisocaproate dehydrogenase	35,351	630
259	NCgl1347	argininosuccinate lyase [4.3.2.1]	51,126	627
260	NCgl1162	F0F1-type ATP synthase delta subunit [3.6.1.34]	28,865	624
261	NCgl2279	ribosomal protein L27	9,435	621
262	NCgl2453	phosphoglucomutase [5.4.2.2]	59,229	610
263	NCgl1321	hypothetical protein	21,953	609
264	NCgl2327	ATP-dependent Clp protease proteolytic subunit 2 [3.4.21.92]	23,081	609
265	NCgl1980	signal recognition particle GTPase	58,356	606
266	NCgl2045	1,4-alpha-glucan branching enzyme	65,812	600
267	NCgl1057	ferredoxin 3	11,946	597
268	NCgl0032	hypothetical protein	29,405	584
269	NCgl1123	hypothetical protein	19,283	583
270	NCgl1371	16S rRNA uridine-516 pseudouridylate synthase [4.2.1.70]	35,783	572
271	NCgl2216	long-chain acyl-CoA synthetase (AMP-forming) [6.2.1.3]	67,134	572
272	NCgl2289	predicted acetyltransferase	12,646	566

Nr.	locus	protein description	mass (Da)	score
273	NCgl2881	ribosomal protein S6	11,010	565
274	NCgl0946	transcription elongation factor	18,803	560
275	NCgl2656	acetate kinase [2.7.2.1]	43,123	556
276	NCgl0562	hypothetical protein	30,143	551
277	NCgl1895	predicted hydrolase of the metallo-beta-lactamase superfamily	77,786	548
278	NCgl0988	predicted GTPase	39,047	545
279	NCgl0902	ribosomal protein L25	21,749	544
280	NCgl0371	formyltetrahydrofolate hydrolase [3.5.1.10]	34,466	542
281	NCgl2272	gamma-glutamyl phosphate reductase [1.2.1.41]	45,870	541
282	NCgl1087	shikimate 5-dehydrogenase [1.1.1.25]	28,969	538
283	NCgl1183	electron transfer flavoprotein alpha-subunit	32,124	538
284	NCgl2658	putative ferredoxin/ferredoxin-NADP reductase [1.18.1.2]	50,098	533
285	NCgl0641	exonuclease III	34,123	528
286	NCgl2594	lysyl-tRNA synthetase class II [6.1.1.6]	59,064	522
287	NCgl2497	acyl-CoA hydrolase	37,788	515
288	NCgl0878	hypothetical protein	12,720	514
289	NCgl1170	lactoylglutathione lyase	16,905	504
290	NCgl1548	carbamoylphosphate synthase small subunit [6.3.5.5]	42,224	500
291	NCgl2930	indole-3-glycerol phosphate synthase [5.3.1.24]	50,736	496
292	NCgl2075	cell division GTPase	47,244	492
293	NCgl2021	histidinol dehydrogenase [1.1.1.23]	46,801	489
294	NCgl2578	NAD-dependent aldehyde dehydrogenase [1.2.1.-]	51,111	487
295	NCgl1572	hypothetical protein	44,453	482
296	NCgl2072	hypothetical protein	16,969	480
297	NCgl1199	putative Asp-tRNA ^{Asn} /Glu-tRNA ^{Gln} amidotransferase A subunit [6.3.5.-]	52,735	479
298	NCgl2534	hypothetical protein	12,045	479
299	NCgl0799	Na ⁺ /proline, Na ⁺ /panthothenate symporter or related permease	57,348	467
300	NCgl1966	predicted RNA binding protein	81,674	462
301	NCgl0654	uracil phosphoribosyltransferase [2.4.2.9]	22,567	461
302	NCgl2070	cell division initiation protein	38,662	461
303	NCgl1447	ATP phosphoribosyltransferase [2.4.2.17]	30,269	457
304	NCgl2422	metal-dependent hydrolase [3.1.6.1]	27,700	457
305	NCgl2804	hypothetical protein	11,939	456
306	NCgl0285	Zn-dependent hydrolase	31,088	451
307	NCgl1919	prolyl-tRNA synthetase [6.1.1.15]	64,408	451
308	NCgl0396	exopolyphosphatase	30,287	450
309	NCgl2607	inorganic pyrophosphatase [3.6.1.1]	17,894	450
310	NCgl2474	serine acetyltransferase [2.3.1.30]	19,538	450
311	NCgl1516	6-phosphogluconolactonase [3.1.1.31]	24,463	449
312	NCgl0145	lactoylglutathione lyase-like protein	31,212	447
313	NCgl0714	phosphomannomutase [5.4.2.8]	49,016	434
314	NCgl1880	RecA/RadA recombinase	40,261	433
315	NCgl0933	hypothetical protein	13,327	433
316	NCgl0738	hypothetical protein	7,952	432
317	NCgl0324	Zn-dependent alcohol dehydrogenase [1.1.1.2]	37,783	430
318	NCgl1471	methylmalonyl-CoA mutase, N-terminal domain/subunit [5.4.99.2]	80,061	429
319	NCgl2139	threonine synthase [4.2.99.2]	53,014	424
320	NCgl2431	nicotinic acid phosphoribosyltransferase	48,053	417
321	NCgl0538	ribosomal protein S11	14,377	416
322	NCgl2436	phosphoserine phosphatase [3.1.3.3]	46,475	415
323	NCgl1344	ornithine carbamoyltransferase [2.1.3.3]	34,423	414
324	NCgl1340	acetylglutamate semialdehyde dehydrogenase [1.2.1.38]	37,243	412
325	NCgl2117	hypothetical protein	12,261	412
326	NCgl1833	hypothetical protein	11,032	411
327	NCgl2281	ribonuclease E [3.1.4.-]	113,294	410

Appendix

Nr.	locus	protein description	mass (Da)	score
328	NCgl0160	hypothetical protein	35,681	410
329	NCgl2842	hypothetical protein	32,318	405
330	NCgl0314	Zn-dependent hydrolase or glyoxylase	22,912	405
331	NCgl2970	ABC-type transport systems, periplasmic component	36,677	402
332	NCgl2663	phosphoribosylglycinamide formyltransferase 2	44,138	402
333	NCgl2817	L-lactate dehydrogenase	45,686	401
334	NCgl0557	ribosomal protein S9	19,711	399
335	NCgl0459	ribosomal protein L11	15,378	399
336	NCgl2602	GTP cyclohydrolase I [3.5.4.16]	22,132	396
337	NCgl1141	nitrate reductase beta chain [1.7.99.4]	60,252	382
338	NCgl1055	hypothetical protein	31,018	377
339	NCgl0853	glycosidase [3.2.1.54]	43,352	376
340	NCgl2336	hypothetical protein	30,491	376
341	NCgl0697	ABC-type transporter, periplasmic component	45,223	375
342	NCgl1109	putative helicase	80,542	374
343	NCgl1373	predicted GTPase	60,531	373
344	NCgl2246	hypothetical protein	30,557	371
345	NCgl1825	hypothetical protein	22,167	368
346	NCgl1072	predicted glycosyltransferase	45,194	368
347	NCgl0748	hypothetical protein	51,603	367
348	NCgl1913	hypothetical protein	19,444	365
349	NCgl2874	thioredoxin	13,563	365
350	NCgl0906	N-acetylglucosamine-1-phosphate uridyltransferase	50,485	364
351	NCgl1430	Xaa-Pro aminopeptidase [3.4.13.9]	40,324	362
352	NCgl0948	uncharacterized protein	32,842	362
353	NCgl2415	RNase PH [2.7.7.56]	26,348	362
354	NCgl2170	hypothetical protein	34,347	358
355	NCgl2929	anthranilate phosphoribosyltransferase [2.4.2.18]	36,611	357
356	NCgl1237	isocitrate/isopropylmalate dehydrogenase [1.1.1.85]	36,123	356
357	NCgl1090	hypothetical protein	29,766	351
358	NCgl2659	predicted acyltransferase	20,218	347
359	NCgl2490	hypothetical protein	24,575	346
360	NCgl2834	two-component system, response regulator	22,362	346
361	NCgl1023	nicotinate-nucleotide pyrophosphorylase [2.4.2.19]	29,416	344
362	NCgl2501	phosphoribosylformylglycinamide (FGAM) synthase, PurS component	8,739	344
363	NCgl1558	Xaa-Pro aminopeptidase	38,969	339
364	NCgl0534	methionine aminopeptidase [3.4.11.18]	28,059	337
365	NCgl0273	hypothetical protein	16,625	336
366	NCgl2156	Zn-ribbon protein	26,347	335
367	NCgl1526	glyceraldehyde-3-phosphate dehydrogenase [1.2.1.12]	36,197	334
368	NCgl1559	3-dehydroquinate synthetase [4.6.1.3]	38,950	332
369	NCgl0678	detergent sensitivity rescuer dtsR1	58,787	332
370	NCgl0151	predicted metalloendopeptidase	71,356	326
371	NCgl2657	phosphotransacetylase [2.3.1.8]	49,285	322
372	NCgl0341	predicted pyridoxal phosphate-dependent enzyme	41,885	320
373	NCgl2128	lipoate synthase	39,693	319
374	NCgl2554	beta-fructosidase [3.2.1.26]	48,101	318
375	NCgl2470	UDP-N-acetylglucosamine enolpyruvyl transferase [2.5.1.7]	44,518	318
376	NCgl2537	trehalose-6-phosphatase [3.1.3.12]	28,088	316
377	NCgl1529	predicted P-loop-containing kinase	34,863	315
378	NCgl2790	glycerol kinase [2.7.1.30]	56,284	310
379	NCgl2205	GTPase	33,399	310
380	NCgl0937	hypothetical protein	20,236	309
381	NCgl2676	orotate phosphoribosyltransferase [2.4.2.10]	19,433	309
382	NCgl0626	carbon starvation protein, predicted membrane protein	81,212	299

Nr.	locus	protein description	mass (Da)	score
383	NCgl0832	ribosomal protein S14	11,725	299
384	NCgl0337	putative ATPase involved in chromosome partitioning	50,080	298
385	NCgl1731	hypothetical protein	11,257	296
386	NCgl1844	DNA-directed RNA polymerase sigma subunit SigB	37,533	293
387	NCgl2135	hypothetical protein	14,515	292
388	NCgl1142	anaerobic dehydrogenase [1.7.99.4]	139,511	289
389	NCgl2083	UDP-N-acetylmuramyl tripeptide synthase [6.3.2.13]	54,274	288
390	NCgl1446	aspartate ammonia-lyase [4.3.1.1]	57,881	288
391	NCgl2813	predicted flavoprotein	20,219	285
392	NCgl0865	FAD/FMN-containing dehydrogenase [1.1.1.28]	63,799	285
393	NCgl2310	predicted hydrolase/acyltransferase [4.1.1.44]	26,748	284
394	NCgl0833	ribosomal protein L33	6,459	280
395	NCgl2120	NaMN,DMB phosphoribosyltransferase [2.4.2.21]	37,594	278
396	NCgl2492	predicted aminomethyltransferase related to GcvT	40,104	276
397	NCgl0327	dTDP-D-glucose 4,6-dehydratase [4.2.1.46]	38,251	275
398	NCgl1241	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase [5.3.3.-]	29,157	274
399	NCgl1549	dihydroorotase [3.5.2.3]	48,145	274
400	NCgl2349	hypothetical protein	45,125	269
401	NCgl1299	DNA polymerase I [2.7.7.7]	96,769	269
402	NCgl1909	ribosome-binding factor A	16,459	265
403	NCgl1485	predicted nucleoside-diphosphate-sugar epimerase	23,806	265
404	NCgl2943	hypothetical protein	22,790	264
405	NCgl2709	Zn-dependent alcohol dehydrogenase [1.1.1.1]	37,195	263
406	NCgl1362	CTP synthase [6.3.4.2]	61,000	262
407	NCgl2210	molecular chaperone	41,101	261
408	NCgl2150	hypothetical protein	67,983	258
409	NCgl1450	Methionine synthase I, cobalamin-binding domain [2.1.1.13]	134,077	256
410	NCgl1860	putative fructose-1-phosphate kinase [2.7.1.11]	34,043	256
411	NCgl2053	dehydrogenase [1.-.-.]	31,207	255
412	NCgl2219	Zn-dependent oligopeptidase [3.4.15.5]	73,835	255
413	NCgl1941	hypothetical membrane protein	16,357	255
414	NCgl1898	dihydrodipicolinate reductase [1.3.1.26]	26,066	254
415	NCgl2446	NAD synthase [6.3.5.1]	30,408	253
416	NCgl0842	molybdopterin biosynthesis enzyme	20,146	253
417	NCgl0049	NAD-dependent aldehyde dehydrogenase [1.2.1.16]	51,974	252
418	NCgl1829	hypothetical protein	30,674	252
419	NCgl0458	transcription antiterminator	34,743	250
420	NCgl2145	hypothetical protein	17,471	249
421	NCgl0710	nucleoside-diphosphate-sugar pyrophosphorylase [2.7.7.13]	38,785	248
422	NCgl0916	gamma-glutamyltranspeptidase [2.3.2.2]	68,396	245
423	NCgl1823	hypothetical protein	15,793	245
424	NCgl2985	thioredoxin	11,934	242
425	NCgl2831	hypothetical protein	23,847	242
426	NCgl0709	predicted glycosyltransferase	31,539	241
427	NCgl1932	methionine aminopeptidase [3.4.11.18]	33,082	241
428	NCgl2274	glutamate 5-kinase [2.7.2.11]	38,602	241
429	NCgl0366	hypothetical protein	14,889	238
430	NCgl2091	5,10-methylenetetrahydrofolate reductase [1.7.99.5]	36,205	237
431	NCgl2227	PLP-dependent aminotransferase [2.6.1.-]	41,053	237
432	NCgl1553	hypothetical protein	17,817	236
433	NCgl2403	bacterioferritin comigratory protein	17,352	231
434	NCgl1166	F0F1-type ATP synthase epsilon subunit [3.6.1.34]	13,076	228
435	NCgl1824	hypothetical protein	27,008	227
436	NCgl0304	topoisomerase IA [5.99.1.2]	109,583	227
437	NCgl2767	predicted S-adenosylmethionine-dependent methyltransferase	28,846	225

Appendix

Nr.	locus	protein description	mass (Da)	score
438	NCgl1074	hypothetical protein	22,769	225
439	NCgl2772	acetyl-CoA carboxylase beta subunit [6.4.1.3]	56,010	224
440	NCgl2847	hypothetical protein	52,326	224
441	NCgl2932	tryptophan synthase alpha chain [4.2.1.20]	29,115	224
442	NCgl1948	uridylate kinase [2.7.4.-]	26,398	222
443	NCgl2518	two-component system, response regulator	26,334	222
444	NCgl1855	SOS-response transcriptional repressor [3.4.21.88]	27,283	219
445	NCgl0647	tryptophanyl-tRNA synthetase [6.1.1.2]	37,861	217
446	NCgl1316	universal stress protein UspA	15,528	217
447	NCgl1876	glutamate ABC-type transporter, periplasmic component	31,822	217
448	NCgl1912	transcription terminator	35,884	216
449	NCgl1442	aspartyl aminopeptidase	44,989	216
450	NCgl1058	PLP-dependent aminotransferase [2.6.1.1]	39,740	216
451	NCgl1355	predicted sugar phosphatase of the HAD superfamily	33,846	214
452	NCgl2159	protein-tyrosine-phosphatase [3.1.3.48]	17,285	213
453	NCgl0793	hypothetical protein	34,696	212
454	NCgl2507	protease II	78,987	212
455	NCgl1554	hypothetical protein	16,376	208
456	NCgl0416	delta-aminolevulinic acid dehydratase-like protein [4.2.1.24]	36,944	206
457	NCgl0242	predicted glutamine amidotransferase	26,531	206
458	NCgl2178	beta-lactamase class C	29,897	205
459	NCgl2414	xanthosine triphosphate pyrophosphatase	22,142	205
460	NCgl1369	rhodanese-related sulfurtransferase	30,654	204
461	NCgl0353	cell wall biogenesis glycosyltransferase	31,533	201
462	NCgl2175	predicted sugar phosphatase of the HAD superfamily	29,620	200
463	NCgl1318	predicted nucleoside-diphosphate-sugar epimerase	29,473	198
464	NCgl1514	glucose-6-phosphate 1-dehydrogenase [1.1.1.49]	57,645	198
465	NCgl1267	D-alanine-D-alanine ligase [6.3.2.4]	38,709	196
466	NCgl0957	hypothetical protein	14,716	195
467	NCgl2333	hypothetical protein	9,045	194
468	NCgl2531	hypothetical protein	16,594	193
469	NCgl0679	biotin-(acetyl-CoA carboxylase) ligase [6.3.4.15]	29,047	193
470	NCgl2788	UDP-galactopyranose mutase [5.4.99.9]	46,152	190
471	NCgl0698	ABC-type transporter, ATPase component	36,167	189
472	NCgl1479	protoheme ferro-lyase [4.99.1.1]	40,844	189
473	NCgl0276	hypothetical protein	6,163	189
474	NCgl0403	porphobilinogen deaminase [4.3.1.8]	32,163	188
475	NCgl1422	hypothetical protein	15,283	188
476	NCgl0420	uroporphyrinogen-III decarboxylase [4.1.1.37]	40,563	188
477	NCgl1354	TPR-repeat-containing protein	49,704	185
478	NCgl2046	threonine dehydratase [4.2.1.16]	46,775	185
479	NCgl2616	rhodanese-related sulfurtransferase	10,400	185
480	NCgl0950	3-Deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase [4.1.2.15]	39,281	183
481	NCgl1995	hypothetical protein	19,304	183
482	NCgl0092	hypothetical protein	103,422	183
483	NCgl0630	citrate synthase [4.1.3.7]	42,626	178
484	NCgl2439	ferritin-like protein	18,055	176
485	NCgl0583	hypothetical protein	15,785	176
486	NCgl2086	predicted S-adenosylmethionine-dependent methyltransferase	36,786	173
487	NCgl2247	malate synthase [4.1.3.2]	82,485	169
488	NCgl2840	transcriptional regulator	23,845	169
489	NCgl2425	transcriptional regulator	18,634	169
490	NCgl0536	translation initiation factor IF-1	8,369	169
491	NCgl2553	phosphotransferase system IIC component [2.7.1.69]	69,336	169
492	NCgl2975	putative copper chaperone	6,976	168

Nr.	locus	protein description	mass (Da)	score
493	NCgl2124	leucyl aminopeptidase [3.4.11.1]	52,755	167
494	NCgl2362	hemoglobin-like protein	15,147	167
495	NCgl2880	single-stranded DNA-binding protein	23,287	166
496	NCgl2402	hypothetical protein	10,149	162
497	NCgl2020	histidinol-phosphate aminotransferase/tyrosine aminotransferase [2.6.1.9]	39,894	161
498	NCgl1196	NAD-dependent DNA ligase [6.5.1.2]	74,892	161
499	NCgl0188	hypothetical protein	8,763	161
500	NCgl0900	glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase [1.2.1.12]	53,229	159
501	NCgl1493	ABC-type transporter, duplicated ATPase component	59,062	158
502	NCgl2016	glutamine amidotransferase [2.4.2.-]	23,390	157
503	NCgl1155	putative translation factor (SUA5)	22,786	156
504	NCgl2001	hypothetical protein	14,778	156
505	NCgl2104	1-acyl-sn-glycerol-3-phosphate acyltransferase [2.3.1.51]	27,184	154
506	NCgl2604	hypoxanthine-guanine phosphoribosyltransferase	22,766	153
507	NCgl1322	excinuclease ATPase subunit	105,037	153
508	NCgl0503	aldo/keto reductase [1.1.1.-]	30,028	152
509	NCgl1500	selenocysteine lyase	45,946	152
510	NCgl0422	glutamate-1-semialdehyde aminotransferase [5.4.3.8]	45,793	150
511	NCgl2110	Rieske Fe-S protein	45,446	150
512	NCgl1213	predicted oxidoreductases	39,694	150
513	NCgl1515	hypothetical protein	34,833	150
514	NCgl2437	heme/copper-type cytochrome/quinol oxidase, subunit 1 [1.9.3.1]	65,233	149
515	NCgl1584	glycerol-3-phosphate dehydrogenase	62,940	149
516	NCgl0765	archaeal fructose-1,6-bisphosphatase	27,934	149
517	NCgl1466	phospholipid-binding protein	18,979	148
518	NCgl1326	ribosomal protein L20	14,774	148
519	NCgl0392	two-component system, response regulators	26,042	146
520	NCgl0730	5-enolpyruvylshikimate-3-phosphate synthase	45,673	146
521	NCgl2471	hypothetical protein	21,123	146
522	NCgl2551	rRNA methylase [2.1.1.-]	33,723	146
523	NCgl0624	homoserine O-acetyltransferase [2.3.1.11]	41,385	145
524	NCgl1357	predicted rRNA methylase	29,700	145
525	NCgl1827	deoxyxylulose-5-phosphate synthase	67,954	145
526	NCgl1342	N-acetylglutamate kinase [2.7.2.8]	33,645	144
527	NCgl0830	hypothetical protein	30,729	144
528	NCgl1944	predicted Fe-S-cluster redox enzyme	40,355	144
529	NCgl0846	UDP-glucose pyrophosphorylase [2.7.7.9]	33,763	143
530	NCgl1544	guanylate kinase [2.7.4.8]	22,619	143
531	NCgl1551	pyrimidine operon attenuation protein [2.4.2.9]	20,939	142
532	NCgl0361	succinate dehydrogenase/fumarate reductase Fe-S protein	27,269	142
533	NCgl1070	SAM-dependent methyltransferase	30,807	142
534	NCgl2146	hypothetical protein	24,226	142
535	NCgl2295	molecular chaperone	52,101	141
536	NCgl1850	transcriptional regulator	35,269	140
537	NCgl2148	glutamine synthase [6.3.1.2]	50,524	140
538	NCgl2502	glutathione peroxidase [1.11.1.9]	17,752	140
539	NCgl1510	NADPH,quinone reductase [1.6.5.5]	35,156	140
540	NCgl0099	predicted oxidoreductase [1.1.1.91]	33,981	140
541	NCgl2440	transcriptional regulator	27,264	139
542	NCgl1133	diaminopimelate decarboxylase [4.1.1.20]	47,614	138
543	NCgl1543	DNA-directed RNA polymerase subunit K/omega	10,328	135
544	NCgl1591	adenine/guanine phosphoribosyltransferase [2.4.2.7]	19,609	135
545	NCgl2082	UDP-N-acetylmuramyl pentapeptide synthase [6.3.2.15]	53,407	134
546	NCgl0046	hypothetical FHA-domain-containing protein	31,356	134
547	NCgl2355	PLP-dependent aminotransferase [2.6.1.11]	49,458	133

Appendix

Nr.	locus	protein description	mass (Da)	score
548	NCgl0651	hypothetical protein	45,325	133
549	NCgl0808	hypothetical protein	8,827	132
550	NCgl0333	serine protease [3.4.21.26]	73,824	132
551	NCgl0120	transcriptional regulator	40,867	132
552	NCgl0421	protoporphyrinogen oxidase [1.3.3.4]	45,730	131
553	NCgl1443	RecB family exonuclease	32,508	131
554	NCgl1896	dihydrodipicolinate synthase/N-acetylneuraminase lyase [4.2.1.52]	31,300	129
555	NCgl0338	protein-tyrosine-phosphatase [3.1.3.48]	22,104	128
556	NCgl2631	hypothetical protein	48,669	128
557	NCgl0105	deoR family transcriptional regulator of sugar metabolism	28,076	127
558	NCgl2931	tryptophan synthase beta chain [4.2.1.20]	45,001	126
559	NCgl1351	hypothetical protein	6,790	125
560	NCgl1587	peptidyl-prolyl cis-trans isomerase (rotamase) [5.2.1.8]	29,791	125
561	NCgl1536	pentose-5-phosphate-3-epimerase [5.1.3.1]	23,913	124
562	NCgl0831	ribosomal protein S18	9,749	122
563	NCgl2979	putative polynucleotide polymerase	55,721	121
564	NCgl0183	hypothetical protein	53,204	120
565	NCgl1178	glycosidase [3.2.1.1]	75,629	120
566	NCgl1868	diaminopimelate epimerase [5.1.1.7]	29,590	118
567	NCgl2645	exonuclease III [3.1.11.2]	29,759	118
568	NCgl1071	beta-fructosidase [3.2.1.26]	54,737	118
569	NCgl2105	glucose kinase [2.7.1.2]	34,455	118
570	NCgl1364	integrase	33,172	117
571	NCgl0223	prephenate dehydrogenase	36,384	116
572	NCgl0603	predicted nucleoside-diphosphate-sugar epimerase	60,606	115
573	NCgl1439	hypothetical protein	56,869	115
574	NCgl0707	SNF2 family helicase	106,877	115
575	NCgl2263	ankyrin repeat containing protein	14,148	114
576	NCgl2401	amidase	20,342	113
577	NCgl2552	cysteinyI-tRNA synthetase [6.1.1.16]	51,189	112
578	NCgl0448	peptidase E	22,975	109
579	NCgl1233	hypothetical protein	67,439	109
580	NCgl2037	maltooligosyl trehalose synthase	90,572	109
581	NCgl1861	phosphotransferase system, fructose-specific IIC component [2.7.1.69]	70,631	109
582	NCgl1528	hypothetical protein	36,164	109
583	NCgl0450	2-succinyl-6-hydroxy-2, 4-cyclohexadiene-1-carboxylate synthase [4.1.1.71]	57,484	107
584	NCgl0930	hypothetical protein	7,503	107
585	NCgl1588	hypothetical protein	19,082	106
586	NCgl2292	folylpolyglutamate synthase	49,285	105
587	NCgl2890	hypothetical protein	35,747	105
588	NCgl0215	histidinol-phosphate aminotransferase/tyrosine aminotransferase [2.6.1.9]	36,659	105
589	NCgl1532	riboflavin synthase beta-chain [2.5.1.9]	16,639	104
590	NCgl2927	anthranilate synthase component I [4.1.3.27]	56,581	103
591	NCgl2080	UDP-N-acetylmuramoylalanine-D-glutamate ligase [6.3.2.9]	49,330	103
592	NCgl2013	imidazoleglycerol-phosphate synthase and cyclase hisF	27,344	103
593	NCgl0702	hypothetical protein	23,505	102
594	NCgl0293	hypothetical membrane protein	19,420	101
595	NCgl1874	2-methylthioadenine synthetase	58,013	101
596	NCgl2099	hypothetical protein	19,256	101
597	NCgl0727	hypothetical protein	14,507	100
598	NCgl2701	molecular chaperone GrpE	23,666	100
599	NCgl2988	putative cell division protein ParB	40,611	99
600	NCgl1200	siderophore-interacting protein	31,269	99
601	NCgl1049	arsenate reductase	12,996	98
602	NCgl2584	hypothetical protein	11,037	98

Nr.	locus	protein description	mass (Da)	score
603	NCgl1078	ATPase	39,147	97
604	NCgl1022	cysteine sulfinatase desulfurase	37,996	97
605	NCgl1533	GTP cyclohydrolase II [3.5.4.25]	46,144	95
606	NCgl1703	site-specific DNA methylase or	40,070	95
607	NCgl0843	large-conductance mechanosensitive channel	14,526	95
608	NCgl2493	hypothetical protein	8,203	95
609	NCgl1274	N6-adenine-specific methylase	21,126	94
610	NCgl0779	ABC-type cobalamin/Fe3+-siderophore transport system, ATPase component	27,988	93
611	NCgl0706	type II restriction enzyme, methylase subunits	171,795	93
612	NCgl0112	panthothenate synthetase [6.3.2.1]	29,928	93
613	NCgl0875	ABC-type transporter	67,218	92
614	NCgl0797	acetyl-CoA carboxylase beta subunit [6.4.1.2]	51,375	92
615	NCgl2495	glutamine phosphoribosylpyrophosphate amidotransferase	55,411	92
616	NCgl0982	penicillin tolerance protein	35,812	91
617	NCgl0118	predicted hydrolase of the HAD superfamily	22,701	91
618	NCgl2595	panthothenate synthetase [6.3.2.1]	28,047	90
619	NCgl2777	putative esterase	70,746	90
620	NCgl1117	putative helicase	115,849	90
621	NCgl0229	queuine/archaeosine tRNA-ribosyltransferase [2.4.2.29]	46,602	89
622	NCgl0873	dimethyladenosine transferase [2.1.1.-]	31,754	89
623	NCgl1441	predicted SAM-dependent methyltransferase involved in tRNA-Met maturation [2.1.1.77]	31,316	89
624	NCgl2358	oxidoreductase [1.1.1.36]	25,330	88
625	NCgl0622	flotillin-like protein	49,718	88
626	NCgl2186	hypothetical protein	69,595	88
627	NCgl2989	putative cell division protein ParA	33,340	87
628	NCgl2472	regulatory-like protein	30,937	86
629	NCgl0513	hypothetical protein	40,843	86
630	NCgl1940	1-deoxy-D-xylulose 5-phosphate reductoisomerase [1.1.1.-]	40,990	85
631	NCgl1597	Holliday junction resolvase DNA-binding subunit	21,579	84
632	NCgl0980	exonuclease VII small subunit [3.1.11.6]	9,014	84
633	NCgl1956	hypothetical protein	12,197	83
634	NCgl2359	transcriptional regulator	27,655	83
635	NCgl2796	hypothetical protein	13,157	83
636	NCgl1303	SAM-dependent methyltransferase	29,263	83
637	NCgl0085	urea amidohydrolase (urease) alpha subunit [3.5.1.5]	61,471	82
638	NCgl1239	predicted signal-transduction protein	69,294	82
639	NCgl2344	hypothetical protein	6,022	81
640	NCgl0672	hypothetical protein	39,705	81
641	NCgl0767	protein chain release factor B	41,169	80
642	NCgl2118	hypothetical membrane protein	29,860	80
643	NCgl2934	hypothetical protein	29,015	80
644	NCgl2897	starvation-inducible DNA-binding protein	18,344	80
645	NCgl0618	ABC-type Fe3+-siderophores transport system, periplasmic component	35,113	80
646	NCgl1846	UDP-glucose 4-epimerase [5.1.3.2]	35,300	79
647	NCgl1586	putative Zn-dependent hydrolase	23,352	78
648	NCgl0616	SIR2 family NAD-dependent protein deacetylase	28,842	78
649	NCgl0336	hypothetical esterase	39,597	78
650	NCgl0657	metal-dependent amidase/aminoacylase/carboxypeptidase	42,346	77
651	NCgl1650	transcriptional regulator	26,925	76
652	NCgl0721	two-component system response regulator	24,985	75
653	NCgl0716	phosphomannose isomerase [5.3.1.8]	42,924	75
654	NCgl0766	archaeal fructose-1,6-bisphosphatase	31,109	75
655	NCgl0098	proline dehydrogenase [1.5.1.12]	126,476	75
656	NCgl1106	putative mutT-like protein	14,747	74

Appendix

Nr.	locus	protein description	mass (Da)	score
657	NCgl2354	hypothetical protein	70,099	74
658	NCgl0737	putative helicase	46,128	74
659	NCgl2571	transcriptional regulator	22,245	74
660	NCgl1015	putative PEP phosphonmutase [2.7.8.23]	26,570	73
661	NCgl1599	hypothetical protein	27,105	73
662	NCgl1325	ribosomal protein L35	7,250	73
663	NCgl1461	dihydroorotate dehydrogenase [1.3.3.1]	39,537	72
664	NCgl0315	dehydrogenase [1.1.1.100]	22,549	72
665	NCgl2782	membrane-associated phospholipid phosphatase	17,610	72
666	NCgl2787	predicted flavoprotein involved in K ⁺ transport	67,005	71
667	NCgl0703	hypothetical protein	94,038	70
668	NCgl1568	predicted periplasmic solute-binding protein	41,934	70
669	NCgl0326	dTDP-4-dehydrorhamnose 3,5-epimerase or related enzyme	49,123	70
670	NCgl2195	chromosome segregation ATPase	72,306	69
671	NCgl0239	DNA polymerase III, gamma/tau subunits	82,519	68
672	NCgl1484	GMP synthase [6.3.5.2]	27,965	68
673	NCgl1475	membrane protease subunit	47,287	67
674	NCgl0576	hypothetical protein	35,206	67
675	NCgl2273	putative phosphoglycerate dehydrogenase	33,032	67
676	NCgl2730	putative peptidase	47,859	66
677	NCgl0952	hypothetical protein	19,108	66
678	NCgl1994	dsRNA-specific ribonuclease [3.1.26.3]	27,464	65
679	NCgl0414	uroporphyrinogen-III synthase/methylase [2.1.1.107]	64,458	65
680	NCgl0751	hypothetical protein	18,570	64
681	NCgl2755	hypothetical protein	16,026	64
682	NCgl0743	predicted NAD-binding component of Kef-type K ⁺ transport system	38,197	64
683	NCgl1606	diadenosine tetraphosphate (Ap4A) hydrolase	19,504	61
684	NCgl2980	hypothetical protein	36,633	61
685	NCgl0064	ATPase related to phosphate starvation-inducible protein PhoH	50,226	61
686	NCgl2275	predicted GTPase	53,646	61
687	NCgl2889	hypothetical protein	11,388	60
688	NCgl0180	PAS/PAC domain containing protein [2.7.3.-]	16,294	60
689	NCgl2376	hypothetical protein	46,596	60
690	NCgl0507	putative formate dehydrogenase [1.2.1.2]	79,130	60
691	NCgl0575	DNA-directed RNA polymerase specialized sigma subunit	20,712	58
692	NCgl0431	hypothetical protein	11,668	58
693	NCgl1025	ADP-ribose pyrophosphatase	23,999	58
694	NCgl2875	copper chaperone	7,099	57
695	NCgl0157	NAD-dependent aldehyde dehydrogenase [1.2.1.27]	53,726	56
696	NCgl2304	ATP-dependent protease Clp, ATPase subunit	47,202	56
697	NCgl0148	hypothetical protein	32,574	56
698	NCgl2624	hypothetical protein	35,217	56
699	NCgl0570	hypothetical protein	59,991	56
700	NCgl2319	protocatechuate 3,4-dioxygenase beta subunit [1.13.11.1]	32,465	56
701	NCgl2794	transcriptional regulator	29,524	56
702	NCgl2809	pyruvate kinase-like protein	67,820	55
703	NCgl2687	alkanal monooxygenase	37,859	55
704	NCgl0628	uncharacterized protein	31,225	55
705	NCgl0372	deoxyribose-phosphate aldolase [4.1.2.4]	22,810	55
706	NCgl0070	hypothetical protein	14,552	54
707	NCgl1902	inosine-uridine nucleoside N-ribohydrolase	33,301	54
708	NCgl0094	nucleoside phosphorylase [3.2.2.4]	53,016	54
709	NCgl1012	Mg-chelatase subunit ChII	50,050	54
710	NCgl0604	deoxyribodipyrimidine photolyase	55,259	54
711	NCgl1738	hypothetical protein	20,271	54

Nr.	locus	protein description	mass (Da)	score
712	NCgl1159	F0F1-type ATP synthase a subunit [3.6.1.34]	30,361	54
713	NCgl0399	hypothetical protein	7,117	54
714	NCgl2798	putative phosphoglycerate mutase	24,823	54
715	NCgl0415	hypothetical protein	16,200	53
716	NCgl0347	cell wall biogenesis glycosyltransferase	38,980	53
717	NCgl1387	hypothetical protein	23,482	52
718	NCgl1324	translation initiation factor IF3	21,763	52
719	NCgl0784	hypothetical protein	7,011	52
720	NCgl2896	hypothetical protein	55,380	50
721	NCgl1539	N-formylmethionyl-tRNA deformylase [3.5.1.31]	18,647	50
722	NCgl2990	glucose-inhibited division protein B	24,204	49
723	NCgl1894	hypothetical protein	23,266	49
724	NCgl1315	helicase subunit of the DNA excision repair complex	81,231	49
725	NCgl2035	DNA polymerase III epsilon subunit [2.7.7.7]	41,713	48
726	NCgl1384	preprotein translocase subunit SecA	83,566	47
727	NCgl1974	RimM protein	18,606	47
728	NCgl2477	succinyl-CoA synthetase beta subunit [6.2.1.5]	41,855	47
729	NCgl1906	hypothetical protein	31,118	46
730	NCgl0938	exopolyphosphatase [3.6.1.11]	34,974	46
731	NCgl0673	hypothetical protein	15,169	46
732	NCgl1277	ABC-type transporter, permease component	34,564	46
733	NCgl2385	short-chain dehydrogenase	28,629	46
734	NCgl2251	choline-glycine betaine transporter	68,335	45
735	NCgl1609	hypothetical protein	18,696	45
736	NCgl1748	periplasmic serine protease	35,266	44
737	NCgl0523	NAD-dependent aldehyde dehydrogenase	52,581	44
738	NCgl1391	hypothetical protein	34,478	44
739	NCgl2242	2-5 RNA ligase	20,685	43
740	NCgl0882	enoyl-CoA hydratase/carnithine racemase [4.2.1.17]	36,565	43
741	NCgl1137	homoserine kinase [2.7.1.39]	32,658	43
742	NCgl0720	thymidylate kinase [2.7.4.9]	22,364	42
743	NCgl2278	hypothetical protein	23,007	41
744	NCgl1161	F0F1-type ATP synthase b subunit [3.6.1.34]	21,082	41
745	NCgl1552	predicted SulA family nucleoside-diphosphate sugar epimerase	42,380	41
746	NCgl0476	ribosomal protein S12	13,479	41
747	NCgl1184	cysteine sulfinatase desulfinate	39,122	40
748	NCgl0731	hypothetical protein	23,535	40
749	NCgl1732	hypothetical protein	8,316	40
750	NCgl0243	UDP-N-acetylmuramyl tripeptide synthase	45,223	39
751	NCgl1527	hypothetical protein	35,712	39
752	NCgl1341	Ornithine acetyltransferase [2.3.1.35]	40,037	38
753	NCgl1670	hypothetical protein	140,349	38
754	NCgl2438	ribonucleotide reductase beta subunit	38,032	38
755	NCgl0684	phosphoribosylaminoimidazole carboxylase [4.1.1.21]	16,981	38
756	NCgl2640	hypothetical protein	42,702	38
757	NCgl2268	phosphoglycerate mutase	26,304	38
758	NCgl2783	hypothetical protein	73,798	37
759	NCgl0820	hypothetical helicase [3.6.1.-]	85,353	37
760	NCgl2919	2-hydroxyhepta-2,4-diene-1,7-dioatesomerase	30,597	37
761	NCgl2324	DNA-binding HTH domain containing protein	97,714	37
762	NCgl1368	acetyltransferase	21,497	37
763	NCgl0446	dihydroxynaphthoic acid synthase [4.1.3.36]	35,634	36
764	NCgl1834	archaeal fructose-1,6-bisphosphatase	29,964	36
765	NCgl0565	putative membrane protein	55,688	36
766	NCgl2465	ABC-type transporter, ATPase component	30,811	36

Appendix

Nr.	locus	protein description	mass (Da)	score
767	NCgl0358	predicted transcriptional regulator	53,574	35
768	NCgl1197	hypothetical protein	23,647	35
769	NCgl1542	phosphopantothenoylcysteine synthetase/decarboxylase	44,061	35
770	NCgl1830	dUTPase [3.6.1.23]	15,911	35
771	NCgl0892	peptide chain release factor 3	60,502	34
772	NCgl0632	transcriptional regulator	26,830	34
773	NCgl0755	predicted glutamine amidotransferase	21,317	34
774	NCgl1275	phosphopantetheine adenyltransferase	17,680	34
775	NCgl1734	hypothetical protein	19,833	34
776	NCgl2208	phosphate starvation-inducible protein PhoH	38,800	34
777	NCgl2487	histone acetyltransferase HPA2-like protein	32,115	34
778	NCgl1011	hypothetical protein	72,022	33
779	NCgl1618	hypothetical protein	67,824	33
780	NCgl1240	DNA polymerase III epsilon subunit	24,328	33
781	NCgl2152	galactokinase [2.7.1.6]	46,537	33
782	NCgl2872	hypothetical protein	8,062	32
783	NCgl2151	hypothetical protein	6,802	32
784	NCgl2200	Fe ²⁺ /Zn ²⁺ uptake regulation protein	16,029	32
785	NCgl2751	deoxycytidine deaminase [3.5.4.13]	20,676	32
786	NCgl0194	hypothetical protein	23,322	32
787	NCgl2986	N-acetylmuramoyl-L-alanine amidase [3.5.1.28]	45,935	31
788	NCgl2147	glutamine synthetase adenyltransferase [2.7.7.42]	116,554	31
789	NCgl0343	predicted glycosyltransferase	32,326	31
790	NCgl0456	geranylgeranyl pyrophosphate synthase [2.5.1.30]	37,799	31
791	NCgl0700	helicase family member	180,670	31
792	NCgl0823	predicted transcriptional regulator	19,120	31
793	NCgl1839	hypothetical protein	63,046	31
794	NCgl1474	hypothetical protein	21,441	31
795	NCgl1270	uracil DNA glycosylase [3.2.2.-]	26,928	31
796	NCgl1924	hypothetical protein	51,737	30
797	NCgl1488	cation transport ATPase [3.6.1.-]	95,418	30
798	NCgl0074	permease	50,244	30
799	NCgl1051	hypothetical protein	20,177	30
800	NCgl0031	ABC-type transporter, ATPase component	27,477	30
801	NCgl0340	predicted nucleoside-diphosphate sugar epimerase	64,410	30
802	NCgl1150	molybdenum cofactor biosynthesis enzyme	41,660	30
803	NCgl1181	hypothetical protein	42,076	30
804	NCgl2766	hypothetical membrane protein	36,271	30
805	NCgl2801	hypothetical membrane protein	24,063	30
806	NCgl2510	PLP-dependent aminotransferase [2.6.1.1]	42,638	30
807	NCgl0924	transcription-repair coupling factor - superfamily II helicase	133,782	29
808	NCgl1460	hypothetical protein	35,146	29
809	NCgl0838	hypothetical protein	6,616	29
810	NCgl1061	tetrahydrodipicolinate N-succinyltransferase [2.3.1.117]	31,383	29
811	NCgl0770	tmRNA-binding protein	18,993	29
812	NCgl2040	hypothetical protein	39,482	29
813	NCgl2606	D-alanyl-D-alanine carboxypeptidase [3.4.16.4]	44,717	29
814	NCgl1189	spermidine synthase	34,634	29
815	NCgl2129	hypothetical membrane protein	28,482	28
816	NCgl0398	pyrroline-5-carboxylate reductase [1.5.1.2]	28,629	28
817	NCgl2350	ABC-type transporter, duplicated ATPase component	59,194	28
818	NCgl1209	ABC-type transport system, periplasmic component	35,908	28
819	NCgl1309	inosine-uridine nucleoside N-ribohydrolase [3.2.2.1]	34,075	28
820	NCgl0455	oxidoreductase	46,016	28
821	NCgl0354	acetyltransferase	22,482	28

Nr.	locus	protein description	mass (Da)	score
822	NCgl1085	ABC-type transporter, duplicated ATPase component	135,573	28
823	NCgl1550	aspartate carbamoyltransferase, catalytic chain [2.1.3.2]	33,931	27
824	NCgl0810	thymidylate synthase [2.1.1.45]	30,385	27
825	NCgl0272	predicted phosphohydrolase [3.1.-.-]	33,262	27
826	NCgl2909	D-amino acid dehydrogenase subunit	44,736	27
827	NCgl2482	phosphate uptake regulator	28,286	27
828	NCgl2928	anthranilate synthase component II [4.1.3.27]	22,267	27
829	NCgl0364	hypothetical protein	35,537	27
830	NCgl0128	ankyrin repeat protein	40,042	27
831	NCgl0740	hypothetical protein	27,740	27
832	NCgl0155	sugar kinases, ribokinase family [2.7.1.4]	34,957	27
833	NCgl1865	GTPase	52,861	27
834	NCgl2812	predicted hydrolase of the HAD superfamily	23,505	27
835	NCgl0394	ABC-type transport system permease component	89,883	27
836	NCgl1904	pseudouridine synthase [4.2.1.70]	31,825	27
837	NCgl2692	hypothetical protein	42,276	27
838	NCgl1715	hypothetical protein	91,924	26
839	NCgl2649	predicted epimerase	29,873	26
840	NCgl1706	hypothetical protein	57,266	26
841	NCgl0305	hypothetical protein	24,388	26
842	NCgl1412	predicted phosphoribosyltransferase	17,821	26
843	NCgl2243	sugar kinase [2.7.1.15]	30,204	26
844	NCgl0199	selenocysteine lyase	43,289	26
845	NCgl2448	hypothetical protein	15,761	26
846	NCgl2221	putative trehalose synthase	69,639	25
847	NCgl0158	hypothetical protein	32,543	25
848	NCgl1272	RecG-like helicase [3.6.1.-]	77,789	25
849	NCgl2176	hypothetical protein	38,229	25
850	NCgl0159	thiamine pyrophosphate-requiring enzyme	68,665	25
851	NCgl2153	exoribonucleases	52,712	25
852	NCgl2188	bacterial type DNA primase [2.7.7.-]	71,344	25
853	NCgl0061	4-oxalocrotonate tautomerase-like protein	17,275	25
854	NCgl1875	glutamate ABC-type transporter, ATPase component	27,512	24
855	NCgl1334	rRNA methylase [2.1.1.34]	20,847	24
856	NCgl0041	serine/threonine protein kinase [2.7.1.-]	50,401	24
857	NCgl2309	acetyl-CoA acetyltransferase	42,738	24
858	NCgl0234	hypothetical dioxygenase protein	34,751	24
859	NCgl0257	predicted transcriptional regulator	12,794	24
860	NCgl1066	putative dihydropteroate synthase [2.5.1.15]	29,812	24
861	NCgl2717	phosphoadenosine phosphosulfate reductase	25,636	24
862	NCgl2655	serine/threonine protein kinase	91,248	24
863	NCgl0839	two-component system, response regulator	26,087	24
864	NCgl2893	efflux system protein	63,883	24
865	NCgl2158	predicted phosphatase [3.1.3.18]	24,867	24
866	NCgl2079	bacterial cell division membrane protein	59,225	24
867	NCgl2308	transcriptional regulator	27,697	24
868	NCgl1806	integrase	44,101	23
869	NCgl1506	ABC-type transporter, ATPase component	36,912	23
870	NCgl2276	xanthine/uracil permease	70,314	23
871	NCgl2057	23S RNA-specific pseudouridylate synthase [4.2.1.70]	33,427	23