bioanalytical chemistry

Andreas Manz Nicole Pamme Dimitri Iossifidis

Imperial College Press

bioanalytical chemistry

This page intentionally left blank

bioanalytical chemistry

Andreas Manz Nicole Pamme Dimitri Iossifidis

Imperial College London



Imperial College Press

Published by

Imperial College Press 57 Shelton Street Covent Garden London WC2H 9HE

Distributed by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

BIOANALYTICAL CHEMISTRY

Copyright © 2004 by Imperial College Press

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 1-86094-370-5 ISBN 1-86094-371-3 (pbk)

Typeset by Stallion Press Email: enquiries@stallionpress.com

Printed in Singapore.

CONTENTS

Prefe List	ace of Abbreviations	ix vi
Cha	nter 1 Biomolecules	1
1 1		1
1.1	Amino Acids, Pepudes and Proteins	1
	1.1.1 Allillo Acids 1.1.2 Pentides and Proteins	27
12	Nucleic Acids	14
1.2	1.2.1 The Structure of Nucleic Acids	15
	1.2.2 Synthesis of Proteins	20
1.3	Biomolecules in Analytical Chemistry	22
	1.3.1 Classical Analytical Chemistry	22
	1.3.2 Limitations of Classical Analytical Chemistry	22
	1.3.3 Bioanalytical Chemistry	23
Cha	pter 2 Chromatography	29
2.1	The Principle of Chromatography	29
2.2	Basic Chromatographic Theory	31
2.3	Application of Liquid Chromatography for Bioanalysis	34
	2.3.1 Reversed Phase Liquid Chromatography (RP-LC)	34
	2.3.2 Ion Exchange Chromatography (IEC)	37
	2.3.3 Affinity Chromatography	40
	2.3.4 Size Exclusion Chromatography (SEC)	42
Cha	pter 3 Electrophoresis	47
3.1	Principle and Theory of Electrophoresis	48
	3.1.1 Electrophoretic Mobility	49
	3.1.2 Joule Heating	50
	3.1.3 Electroosmotic Flow (EOF)	50
	3.1.4 Separation Efficiency and Resolution	54
3.2	Gel Electrophoresis (GE)	56
	3.2.1 Instrumentation for Gel Electrophoresis	57
	3.2.2 Modes of Gel Electrophoresis	63
	3.2.3 Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS–PAGE)	63
	Enculophotosis (SES TROE)	05

	3.2.4	Isoelectric Focussing (IEF)	64
	3.2.5	Two-Dimensional Gel Electrophoresis (2D-GE)	67
3.3	Capill	ary Electrophoresis (CE)	69
	3.3.1	Capillary Electrophoresis Instrumentation	70
	3.3.2	Capillary Zone Electrophoresis (CZE)	75
	3.3.3	Capillary Isoelectric Focussing (CIEF)	76
	3.3.4	Micellar Electrokinetic Chromatography (MEKC)	77
	3.3.5	Capillary Gel Electrophoresis (CGE)	82
Cha	pter 4	Mass Spectrometry	85
4.1	The P	rinciple of Mass Spectrometry	85
	4.1.1	Ionisation	86
	4.1.2	Mass Analyser	86
	4.1.3	Detector	87
4.2	Matrix	x Assisted Laser Desorption Ionisation – Time of Flight	
	Mass	Spectrometry (MALDI-TOF/MS)	87
	4.2.1	Ionisation Principle	87
	4.2.2	Mass Analysis in Time-of-Flight Analyser	90
	4.2.3	Detection of Ions	92
	4.2.4	Resolution	92
	4.2.5	Sample Pretreatment	93
	4.2.6	Applications of MALDI	94
4.3	Electr	ospray Ionisation Mass Spectrometry (ESI-MS)	97
	4.3.1	Ionisation Principle	98
	4.3.2	ESI – Source and Interface	99
	4.3.3	Quadrupole Analyser	100
	4.3.4	Applications of ESI-MS	101
Cha	pter 5	Molecular Recognition:	
		Bioassays, Biosensors, DNA-Arrays and	
		Pyrosequencing	109
5.1	Bioas	says	110
	5.1.1	Antibodies	111
	5.1.2	Antigens	113
	5.1.3	Antibody-Antigen Complex Formation	114
	5.1.4	Assay Formats	115
	5.1.5	Home Pregnancy Test	120
	5.1.6	Enzyme Immunoassays (EI and ELISA)	121
5.2	Bioser	nsors	125
	5.2.1	Bioreceptors	126

	5.2.2 Transducers	127
	5.2.3 The Blood Glucose Sensor	128
5.3	DNA Binding Arrays	131
	5.3.1 The Principle of DNA Arrays	131
	5.3.2 Fabrication of DNA Arrays	132
	5.3.3 Development and Analysis of a DNA Array	134
	5.3.4 DNA Sequencing with Arrays	134
	5.3.5 Other Applications of DNA Arrays	136
5.4	DNA Identification by Pyrosequencing	136
	5.4.1 The Principle of Pyrosequencing	137
	5.4.2 Sample Preparation and Instrumentation	140
	5.4.3 Applications of Pyrosequencing	140
Cha	pter 6 Nucleic Acids:	
	Amplification and Sequencing	143
6.1	Extraction and Isolation of Nucleic Acids	143
	6.1.1 CsCl Density Gradient Centrifugation	144
	6.1.2 Total Cellular DNA Isolation	145
	6.1.3 RNA Isolation – The Proteinase K method	145
6.2	Nucleic Acid Amplification – The Polymerase Chain	
	Reaction (PCR)	146
	6.2.1 The Principle of PCR	146
	6.2.2 The Rate of Amplification During a PCR	149
	6.2.3 Reagents for PCR	151
	6.2.4 Real-Time PCR	153
	6.2.5 Reverse Transcription – PCR (RT-PCR)	155
6.3	Nucleic Acid Sequencing	156
	6.3.1 The Use of Restriction Enzymes in Sequencing	156
	6.3.2 The Chemical Cleavage method	1.50
	(The Maxam-Gilbert method)	158
	6.3.3 The Chain Terminator method (The Sanger or	1(0
6.4	Dideoxy method)	162
6.4	KINA Sequencing	100
Cha	pter 7 Protein Sequencing	169
7.1	Protein Sequencing Strategy	170
7.2	End-group Analysis	170
	7.2.1 <i>N</i> -terminal Analysis (Edman Degradation)	171
	7.2.2 <i>C</i> -terminal Analysis	172
7.3	Disulfide Bond Cleavage	175

7.4 Separation and Molecular Weight Determination of the Protein		
	Subunits	177
7.5	Amino Acid Composition	
7.6	Cleavage of Specific Peptide Bonds	
	7.6.1 Enzymatic Fragmentation	180
	7.6.2 Chemical Fragmentation Methods	183
7.7	Sequence Determination 18	
7.8	Ordering of Peptide Fragments 186	
7.9	Determination of Disulfide Bond Positions 18	
7.10	Protein Sequencing by Mass Spectrometry	187

Index

189

Preface

In a time when sequencing the human genome has just recently been completed, when Nobel prizes are awarded to inventors of bioanalytical instrumentation and when the reading of journals such as *Science* or *Nature* has become ever more difficult to the chemist due to the flood of molecular biology terminology appearing in these groundbreaking publications ... At exactly this time, it seems imperative to provide a small introductory textbook covering the most frequently used instrumental methods of analytical chemistry in molecular biology. The increasingly interdisciplinary nature of modern research makes it essential for researchers of different backgrounds to have at least a minimal understanding of neighbouring sciences if they are to communicate effectively.

For many years, Professor Manz has presented a "bioanalytical chemistry" course at Imperial College, whilst being acutely aware of the lack of a suitable textbook for this subject. Of course, each individual subunit could be found in yet another biochemistry, mass spectrometry, separations or analytical chemistry textbook. However, considering the importance of biomolecules in recent academic and industrial research, it is somewhat surprising that this is not yet reflected in current analytical chemistry textbooks. In the light of these facts, it seems appropriate for us to write a new book concerning the various aspects of biomolecular analysis.

This book is aimed primarily at chemistry students, but is also intended to be a useful reference for students, lecturers and industrial researchers in biological and medicinal sciences who are interested in bioanalysis techniques. It is assumed that the basic principles and instrumental techniques of analytical chemistry are already common knowledge. An important objective of this book is to give an appreciation of how analytical methods are influenced by the properties that are peculiar to biomolecules. The priorities that govern the choice of instrumental techniques for the analysis of molecules such as DNA and proteins are radically different to those applicable to classical analytical chemistry (see Summary of Chapter 1). Whereas samples containing small molecules can be characterised by gas or liquid chromatography, when it comes to DNA sequencing or proteomic analysis, there is a sudden need for sheer separation power. Hence, students must have as clear an understanding of isoelectric focussing or 2D slab gel separation as they would of conventional chromatography. Other methods described in this book may be completely new to the chemist. For example, the polymerase chain reaction used for DNA amplification or the Sanger reaction for DNA sequencing, where low yield chemical reactions are performed to generate hundreds of products.

In the first chapter of this book, a general introduction to biomolecules is given. This is followed by several chapters describing various instrumental techniques and bioanalytical methods. These include: electrophoresis, isoelectric focussing, MALDI-TOF, ESI-MS, immunoassays, biosensors, DNA arrays, PCR, DNA and protein sequencing. Instead of being a comprehensive reference or textbook, it is intended that this book should provide introductory reading, perhaps alongside a taught course. A list of references is given at the end of each chapter, should further information be required on any particular subject.

Hopefully, this book will be well received by both teachers and students, particularly in a time when techniques of bioanalysis should be familiar to every chemistry graduate.

The authors would like to thank Dr. Alexander Iles for his comments on the manuscript.

Andreas Manz, Nicole Pamme, Dimitri Iossifidis London, March 12, 2003

List of Abbreviations

2D-GE	two-dimensional gel electrophoresis
А	Adenine
α	selectivity factor
Ab	antibody
ABTS	2,2'-azino-bis (ethyl-benzothiazoline-6-sulfonate)
ac	alternating current
α-CHCA	α -cyano-4-hydroxy-cinnamic acid
AChE	acetylcholine esterase
ADT	adenosine diphosphate
Ag	antigen
AIDS	acquired immunodeficiency syndrome
Ala	Alanine
AMP	adenosine monophosphate
AN	aggregation number
AP	alkaline phosphatase
APS	adenosine phosphosulphate
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
c	concentration
С	Cytosine
C%	degree of cross-linking
CCD	charged coupled device
cDNA	complementary DNA
CE	capillary electrophoresis
CGE	capillary gel electrophoresis
CHAPS	3-[(cholamido propyl) dimethyl
	ammonio]-1-propane sulphonate
CI	chemical ionisation
CID	collision-induced dissociation
CIEF	capillary isoelectric focussing
СМ	carboxy methyl
CMC	critical micelle concentration

CNBr	cyanogen bromide
CTAB	cetyltrimethylammonium bromide
CTAC	cetyltrimethylammonium chloride
Cys	Cysteine
CZE	capillary zone electrophoresis
D	diffusion coefficient
Da	Dalton
DAD	diode array detector
dATP	deoxyadenine triphosphate
dATP-αS	deoxyadenine α -thio-triphosphate
dc	direct current
dCTP	deoxycytosine triphosphate
ddNTP	2',3'-dideoxynucleotide triphosphate
DEAE	diethyl aminoethyl
dGTP	deoxyguanine triphosphate
DHBA	2,5-dihydroxy benzoic acid
DMS	dimethyl sulphate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNDP	deoxynucleotide diphosphate
dNMP	deoxynucleotide monophosphate
dNTP	deoxynucleotide triphosphate
DoTAB	dodecyl trimethyl ammonium bromide
ΔpI	resolution (in isoelectric focusing)
dsDNA	double stranded DNA
DTT	dithiothreitol
dTTP	deoxythymine triphosphate
ϵ	dielectric constant
E	electric field strength
e	electron charge
EI	electron impact ionisation
EI	enzyme imunoassay
$E_{\rm kin}$	kinetic energy
ELISA	enzyme-linked immunosorbent assay
EOF	electroosmotic flow
ESI	electrospray ionisation
Fab	antigen binding fragment of Ig
FAB	fast atom bombardment
Fc	crystallisable fragment of Ig
Fef	electric force
F_{fr}	frictional force
FRET	fluorescence resonance energy transfer
FWHM	full width at half maximum

G	Guanine
GC	gas chromatography
GE	gel electrophoresis
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GOx	glucose oxidase
GPC	gel permeation chromatography
Н	height equivalent of a theoretical plate
η	viscosity
hCG	human chorionic gonadotropin
His	Histidine
HIV	human immunodeficiency virus
HPCE	high performance capillary electrophoresis
HPG	human genome project
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
i.d.	inner diameter
IEC	ion exchange chromatography
IEF	isoelectric focussing
Ig	Immunoglobulin
Ile	Isoleucine
IPG	immobilised pH gradient
IR	infrared
k′	capacity factor
k_3	turnover of an enzyme
K _{eq}	equilibrium constant of antibody-antigen complex formation
Km	Michaelis-Menten constant
L	length (of capillary, colum or gel)
λ	wavelength
LC	liquid chromatography
Leu	Leucine
LIF	laser induced fluorescence
Lys	Lysine
m	mass
М	molar, mol L^{-1}
m/z	mass-to-charge ratio
MALDI	matrix assisted laser desorption ionisation
μ_{app}	apparent mobility
MECC	micellar electrokinetic capillary chromatography
MEKC	micellar electrokinetic chromatography
μ_{EOF}	electroosmotic mobility
μ_{ep}	electrophoretic mobility

$\mu_{ep,AVE}$	average electrophoretic mobility of two analytes
Met	Methionine
mM	millimolar
mRNA	messenger RNA
MS	mass spetrometry
MS/MS	tandem mass spectrometry
μ_{tot}	total mobility
MW	molecular weight
Ν	plate number
N_0	initial number of DNA molecules in PCR
N _m	number of DNA molecules in PCR
NMR	nuclear magnetic resonance
ODS	octadecyl silane
OPA	ortho-phthalaldehyde
ox.	oxidised
PA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
pН	potentium hydrogenis
Phe	Phenylalanine
pI	isoelectric point
PICT	phenylisothiocyanate
pK	dissociation constant
ppb	parts per billion
PPi	pyrophosphate
ppm	parts per million
RP	reversed phase
Pro	Proline
PSD	post source decay
PTH	phenylthiohydantoin
q	charge of ion
QT-PCR	quantitative PCR
r	ionic/molecular radius
red.	reduced
RNA	ribonucleic acid
R_S	resolution
RT	reverse transcription
RT-PCR	reverse transcription polymerase chain reaction
S	signal intensity
s^2	peak dispersion
SA	sinapinic acid

SC	sodium cholate
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
Ser	Serine
SLD	soft laser desorption
SNP	single nucleotide polymorphism
ssDNA	single stranded DNA
STC	sodium taurocholate
STS	sodium tetradecyl sulphate
t	migration time
Т	Thymine
Τ%	total gel concentration
<i>t</i> ₀	zero retention time
Taq	Thermus aquaticus
TFA	trifluoroacetic acid
Thr	Threonine
<i>t</i> _{mc}	retention time of micelles
TOF	time of flight
t_R	retention time
TRIS	tris (hydroxylmethyl)-aminomethane
tRNA	transfer RNA
Trp	Tryptophan
Tyr	Tyrosine
u	flow rate
U	Uracil
UV	ultraviolet
V	applied voltage
v	migration velocity
V_0	inter particle volume
Val	Valine
$v_{\rm EOF}$	velocity of electroosmotic flow
$v_{ m ep}$	electrophoretic velocity
V_{g}	volume of gel particles
$V_{\rm i}$	intrinsic volume
vis	visible
$v_{\rm MC}$	velocity of micelles
$V_{\rm R}$	retention volume
$V_{\rm t}$	total volume
W	peak width
Z	ion charge
ζ	zeta potential

Chapter 1

BIOMOLECULES

In this chapter, you will learn about ...

- ♦ ... the biomolecules that are most commonly analysed in bioanalytical chemistry: amino acids, proteins and nucleic acids.
- ♦ ... the structure of these biomolecules and their physical and chemical characteristics.
- ♦ ... some of the functions of these biomolecules and how they interact with each other in the cell.

Chemists are likely to be familiar with certain biomolecules such as carbohydrates and lipids from their organic chemistry lectures. However, many do not have a clear understanding of the composition and function of other biomolecules such as proteins and DNA. This chapter introduces the biomolecules, which are the target of the analytical methods described in the following chapters.

1.1 Amino Acids, Peptides and Proteins

Amino acids are the building blocks for peptides and proteins and play an important part in metabolism. 20 different amino acids are found in living organisms. They can connect to each other via peptide bonds to form long chains. Proteins may consist of thousands of amino acids and can have molecular weights of up to several million Dalton (Da). Shorter chains of up to a few hundred amino acids are referred to as peptides. The sequence of the amino acids within the molecule is essential for the structure and function of proteins and peptides in biological processes.

1.1.1 Amino Acids

The general structure of an amino acid is shown in Fig. 1.1. It consists of a tetrahedral carbon atom (C-alpha) connected to four groups: a basic amino group $(-NH_2)$, an acidic carboxyl group (-COOH), a hydrogen atom (-H) and a substituent group (-R), which varies from one amino acid to another. The amino group is in the alpha position relative to the carboxyl group, hence the name α -amino acids. Amino acids are chiral with the exception of glycine, where the R substituent is a hydrogen atom. All natural amino acids have the same absolute configuration: the L-form in the Fischer convention or the S-form according to the Cahn-Ingold-Prelog rules, with the exception of cysteine, which has the R-configuration.

Amino acids can be classified according to their substituent R groups (Fig. 1.2 to Fig. 1.8): in *basic amino acids*, R contains a further amino group, whereas in *acidic amino acids*, R contains a further carboxyl group. In addition, there are *aliphatic*, *aromatic*, *hydroxyl containing* and *sulfur containing amino acids* according to the nature of the substituent, as well as a *secondary* amino acid.

For convenience, the names for amino acids are often abbreviated to either a *three symbol* or a *one symbol short form*. For example, Arginine can be referred



Fig. 1.1. General structure of an α -L-amino acid.



Fig. 1.2. Basic amino acids.



Fig. 1.3. Acidic amino acids.



Fig. 1.4. Aliphatic amino acids.



Fig. 1.5. Aromatic amino acids.



Fig. 1.6. Sulfur containing amino acids.



Fig. 1.7. Amino acids with an alcoholic hydroxyl group.



Fig. 1.8. Secondary amino acid.

to as Arg or R and Glycine can be shortened to Gly or G. The abbreviations for the 20 natural amino acids are listed in Table 1.1. These naturally occurring amino acids are the building blocks of peptides and proteins. Any particular amino acid is not likely to exceed 10 % of the total composition of a protein (see Table 1.1).

Amino acids can also be classified according to their polarity and charge at pH 6 to 7, which corresponds to the pH range found in most biological systems. This is often referred to as the *physiological pH*. *Non-polar amino acids* with no